

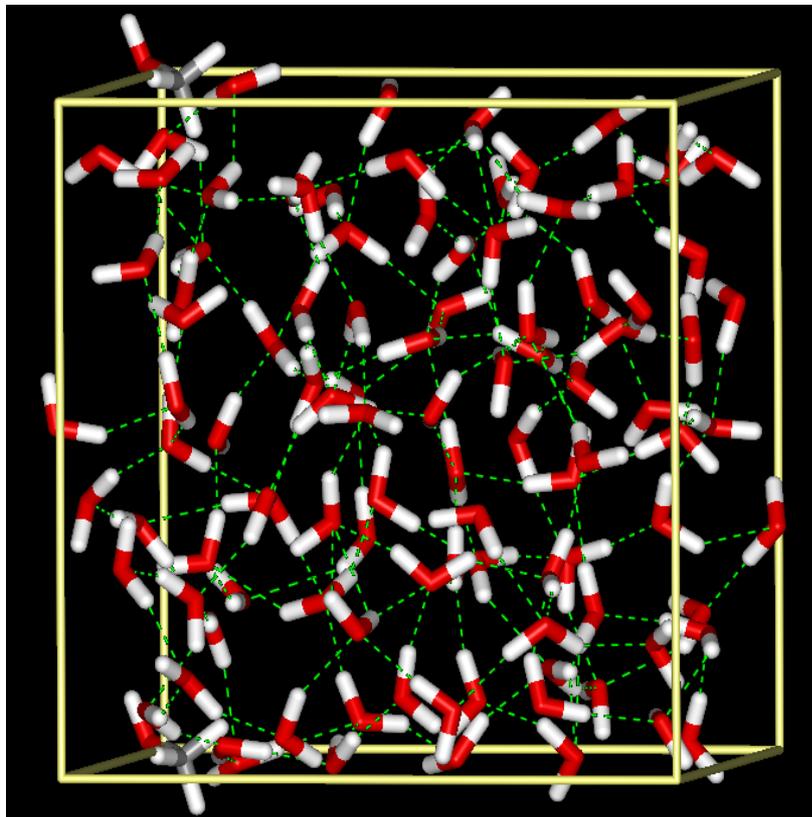
Simulação Computacional de Líquidos

Profa. Kaline Coutinho

kaline@if.usp.br

Instituto de Física da USP

Aula 20: Iniciando simulações com dinâmica molecular de soluto em vácuo e solução aquosa com o programa GROMACS



ROTEIRO PARA PRÁTICA DE SIMULAÇÃO COMPUTACIONAL

Assunto: Simulação de molécula flexível com o Método Dinâmica Molecular com o programa GROMACS: DCI

Kaline Coutinho (kaline@if.usp.br)
Instituto de Física da USP, Brasil

I. Introdução

As simulações de moléculas flexíveis com o programa GROMACS utilizam o método de dinâmica molecular (MD). Vamos usar o ensemble NVT para simulação da molécula isolada e NPT para simulação em água. Para utilizar um tempo de integração maior ($\Delta t = 0.002 \text{ ps} = 2 \text{ fs}$) vamos usar vínculos nas ligações químicas que envolvam átomos de hidrogênio, mas outra opção é usar um tempo de integração menor ($\Delta t = 0.0001 \text{ ps} = 0.1 \text{ fs}$) sem vínculos.

II. Cálculos quânticos para otimização de geometria e cargas atômicas

Antes de iniciar as simulações é recomendado realizar cálculos com mecânica quântica (QM) para obter a geometria otimizada da molécula. Neste tutorial vamos utilizar a molécula mostrada na Figura 1, que chamaremos de DCI, com carga 0 e estado de spin singleto, ou seja, multiplicidade 1. Utilizando o programa GAUSSIAN com a palavra-chave OPT a geometria foi otimizada utilizando o método B3LYP e conjunto de funções de base cc-pVDZ, e depois já com a geometria otimizada, outro cálculo foi realizado para obter as cargas atômica usando o ajuste do potencial eletrostático com o método Merz-Singh-Kollman (palavra-chave Pop=MK) usando o mesmo método/base da otimização. Estas cargas atômicas da molécula devem ser calculadas em vácuo e em solução aquosa com modelo contínuo (palavra-chave SCRF=(PCM, solvente=water)) para analisar o quanto a molécula pode se polarizar ao mudar de um ambiente apolar (vácuo), para um ambiente polar (água) e assim, identificar se este efeito é importante e deve ser considerado. Um aumento do momento de dipolo, ou cargas atômicas maiores de 30% já começa a ser relevante.

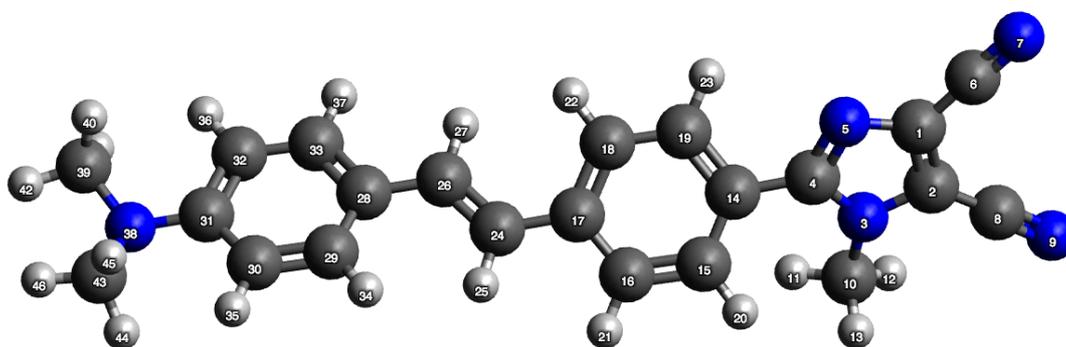


Figura 1: Visualização da geometria otimizada e numeração dos átomos que será usada.

Atenção com a numeração: os grupos químicos devem ser numerados em sequência de forma que cada átomo seguinte esteja ligado com um átomo que apareceu anteriormente para não dar problema na forma de gerar o campo de força OPLS no site do LigParGen. Veja que neste caso, o grupo imidazol, N_2C_3 , é numerado primeiro, depois os grupos $(\text{CN})_2$ e CH_3 que estão ligados. Em

seguida, aparece o anél aromático ligado ao imidazol, depois a cadeia carbônica (CH)₂, depois o próximo anél aromático e por final o grupo amino N(CH₃)₂.

III. Preparando os arquivos de entrada para simulação

Após os cálculos quânticos é necessário gerar um arquivo em pdb com a geometria otimizada. Sugerimos gerar um arquivo no formato xyz que é mais simples e depois transformar para o pdb. O formato xyz é da seguinte forma: 1^a. linha: número de átomos; 2^a. linha: comentário; 3^a. linha em diante: símbolo atômico e coordenadas cartesianas (x, y e z) em Å. Como mostrado o arquivo abaixo dci-opt.xyz:

```
46
dci OPT(MP2/CC-PVDZ) q(Density Pop=MK)
C -6.110097 1.000568 -0.284350
C -6.296048 -0.297657 0.226960
N -5.032299 -0.841768 0.310437
C -4.149280 0.115995 -0.150528
N -4.782484 1.243730 -0.508876
C -7.140258 1.959110 -0.547595
N -8.012313 2.740161 -0.756966
C -7.467294 -1.001865 0.620789
N -8.434715 -1.606200 0.961451
C -4.730247 -2.150975 0.887887
H -3.719578 -2.120236 1.316122
H -5.464648 -2.362249 1.678672
H -4.786457 -2.938486 0.120598
C -2.692531 -0.056802 -0.211635
C -2.084661 -1.259750 -0.634174
C -0.688063 -1.356785 -0.698395
C 0.145681 -0.267260 -0.352356
C -0.480754 0.948283 0.023118
C -1.872424 1.052105 0.101300
H -2.698179 -2.109137 -0.951911
H -0.228234 -2.295269 -1.029731
H 0.132292 1.829021 0.240720
H -2.348524 1.996309 0.383282
C 1.602932 -0.427992 -0.413022
H 1.964182 -1.297872 -0.977466
C 2.509883 0.377008 0.211500
H 2.137619 1.191103 0.849171
C 3.967285 0.238223 0.154977
C 4.638971 -0.555523 -0.804078
C 6.031721 -0.674609 -0.810120
C 6.838043 -0.005227 0.151638
C 6.166029 0.820852 1.091298
C 4.771192 0.930786 1.088534
H 4.068040 -1.070839 -1.584256
H 6.492535 -1.290827 -1.584998
H 6.729140 1.385695 1.837068
H 4.289837 1.574736 1.834611
N 8.224741 -0.165195 0.177873
C 8.996888 0.826873 0.912001
H 8.850648 1.860198 0.533032
H 8.732306 0.810085 1.981008
H 10.063304 0.569857 0.833701
C 8.861600 -0.685429 -1.023616
H 8.508799 -1.706332 -1.238755
H 8.679216 -0.057053 -1.920596
H 9.945773 -0.743395 -0.848506
```

O arquivo dci-opt.xyz pode gerar um arquivo no formato pdb usando o programa Open Babel que é gratuito e pode ser instalado facilmente no Linux e Windows. O comando para fazer esta mudança de formato é o seguinte:

```
$ babel -ixyz dci-opt.xyz -opdb dci-opt.pdb
1 molecule converted
17 audit log messages
```

O arquivo dci-opt.pdb será usado no servidor LigParGen (link: <http://zarbi.chem.yale.edu/ligpargen/>) para gerar o arquivo de topologia do campo de força OPLS. Na Figura 2 mostramos a janela do servidor LigParGen e as opções sugeridas.

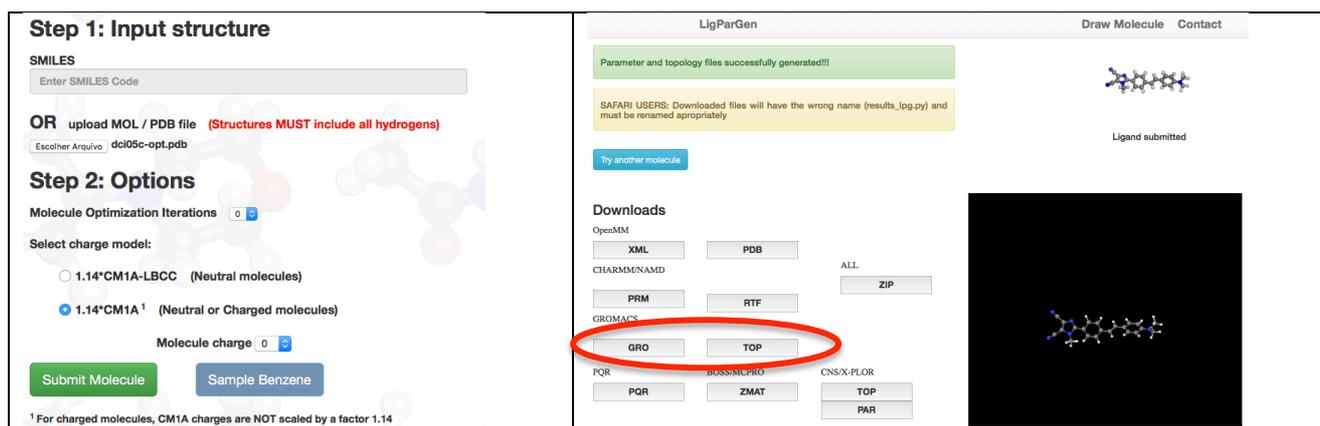


Figura 2: Visualização do site LigParGen com as opções de submissão (esquerda). Clicar no botão *Submit Molecule*; e visualização do site após submissão (direita). Clicar nos botões *GRO* e *TOP* do *GROMACS* para fazer download destes dois arquivos gerados.

Após utilizar o servidor LigParGen, são necessários baixar os arquivos GRO e TOP. Ambos arquivos são gerados para utilização no programa GROMACS. Assim, sugerimos renomear estes arquivos baixados do site LigParGen para: *dci.gro* (arquivo com a geometria inicial para uma simulação com o programa GROMACS) e *dci.itp* (arquivo com a topologia para uma simulação com o programa GROMACS).

Deve-se editar o arquivo *dci.itp* para colocar o nome da molécula com 3 letras maiúsculas (nos blocos [moleculetype] e [atoms]) e as cargas calculadas com QM (no bloco [atoms]). Note que quatro casas decimais são suficientes para cargas atômicas e é possível fazer ajustes nas cargas atômicas para colocar valores médios em átomos semelhantes, exemplo os Hs do grupo CH₃, mas no final dos arredondamentos e aproximações das cargas atômicas, não esqueça de somar total estas cargas para conferir se a carga total está correta. Veja um exemplo dos blocos modificados no arquivo *dci-ok.itp*. Neste caso não fizemos ajustes nas cargas atômicas, apenas truncamento na quarta casa decimal e para zerar a carga total colocamos as diferenças nos Hs dos 2 grupos CH₃.

IV. Realizando a simulação da molécula flexível em vácuo

Agora com os arquivos de entrada para o programa GROMACS, vamos realizar uma simulação em vácuo (no ensemble NVT com densidade de 0.001 g/cm³, ou seja com uma caixa cúbica de lado 8.4 nm). Nesta simulação manteremos todas as ligações químicas livres para vibrar. Então a simulação usará um passo de simulação de 0.1 fs e tempo total de 1 ns. São necessários os seguintes arquivos para realizar a simulação:

- *dci.top* (com definição do campo de força, nomes dos arquivos de topologia e quantidade de cada tipo de molécula),
- *dci-ok.itp* (com a topologia e parâmetros do potencial ligado: bond, angle, dihedral e improper, e potencial não-ligado: Lennard-Jones e Coulomb),
- *min.mdp* (com as palavras chaves para minimização da configuração inicial) e
- *nvt.mdp* (com as palavras chaves para as simulações).

Exemplo de arquivo *dci.top*:

```
; The force-field files to be included
#include "oplsaa.ff/forcefield.itp"

; The solute topology file to be included
#include "dci-ok.itp"

[ system ]
DCI in vacuum
```

```
[ molecules ]
DCI 1
```

Exemplo dos blocos modificados (nome da molécula “DCI” e cargas obtidas com QM) do arquivo de topologia gerado com o LigParGen: dci-ok.itp

```
[ moleculetype ]
; Name          nrexcl
DCI              3
[ atoms ]
; nr      type  resnr  residue  atom  cgnr  charge  mass
1  opl_800  1    DCI    C00    1    0.2345  12.0110
2  opl_801  1    DCI    C01    1   -0.3427  12.0110
3  opl_802  1    DCI    N02    1    0.1520  14.0070
4  opl_803  1    DCI    C03    1    0.3333  12.0110
5  opl_804  1    DCI    N04    1   -0.4954  14.0070
6  opl_805  1    DCI    C05    1    0.3430  12.0110
7  opl_806  1    DCI    N06    1   -0.4186  14.0070
8  opl_807  1    DCI    C07    1    0.4857  12.0110
9  opl_808  1    DCI    N08    1   -0.4425  14.0070
10 opl_809  1    DCI    C09    1   -0.3112  12.0110
11 opl_810  1    DCI    HOA    1    0.1181  1.0080
12 opl_811  1    DCI    HOB    1    0.1575  1.0080
13 opl_812  1    DCI    HOC    1    0.1433  1.0080
14 opl_813  1    DCI    COD    1    0.0230  12.0110
15 opl_814  1    DCI    COE    1   -0.1294  12.0110
16 opl_815  1    DCI    COF    1   -0.2160  12.0110
17 opl_816  1    DCI    COG    1    0.1913  12.0110
18 opl_817  1    DCI    COH    1   -0.1186  12.0110
19 opl_818  1    DCI    COI    1   -0.2197  12.0110
20 opl_819  1    DCI    HOJ    1    0.1101  1.0080
21 opl_820  1    DCI    HOK    1    0.1447  1.0080
22 opl_821  1    DCI    HOM    1    0.1457  1.0080
23 opl_822  1    DCI    HON    1    0.1558  1.0080
24 opl_823  1    DCI    COO    1   -0.2242  12.0110
25 opl_824  1    DCI    HOP    1    0.1464  1.0080
26 opl_825  1    DCI    COQ    1   -0.2723  12.0110
27 opl_826  1    DCI    HOR    1    0.1379  1.0080
28 opl_827  1    DCI    COS    1    0.2594  12.0110
29 opl_828  1    DCI    COT    1   -0.2136  12.0110
30 opl_829  1    DCI    COU    1   -0.3266  12.0110
31 opl_830  1    DCI    COV    1    0.4279  12.0110
32 opl_831  1    DCI    COW    1   -0.3450  12.0110
33 opl_832  1    DCI    COX    2   -0.2198  12.0110
34 opl_833  1    DCI    HOY    2    0.1469  1.0080
35 opl_834  1    DCI    H0Z    2    0.1759  1.0080
36 opl_835  1    DCI    H10    2    0.1881  1.0080
37 opl_836  1    DCI    H11    2    0.1458  1.0080
38 opl_837  1    DCI    N12    2   -0.2669  14.0070
39 opl_838  1    DCI    C13    2   -0.1364  12.0110
40 opl_839  1    DCI    H14    2    0.0655  1.0080
41 opl_840  1    DCI    H15    2    0.0617  1.0080
42 opl_841  1    DCI    H16    2    0.1037  1.0080
43 opl_842  1    DCI    C17    2   -0.1106  12.0110
44 opl_843  1    DCI    H18    2    0.0539  1.0080
45 opl_844  1    DCI    H19    2    0.0622  1.0080
46 opl_845  1    DCI    H1A    2    0.0962  1.0080
```

Exemplo de arquivo min.mdp:

```
integrator      = steep          ; Steepest descent algorithm for energy minimization
dt              = 0.0001
nsteps         = 5000
emtol          = 1              ; [kJ/mol.nm] convergency criterium

; Output control frequency
nstxout        = 10
nstvout        = 10
nstlog         = 10
nstenergy      = 10
nstxtcout      = 10
xtc_precision  = 10

; Neighborsearching and short-range nonbonded interactions
pbc            = xyz
cutoff-scheme  = verlet
```

Exemplo de arquivo nvt.mdp:

```
integrator      = md              ; leap-frog integrator
dt              = 0.0001          ; 0.1 fs timestep
nsteps         = 10000000        ; total time 1 ns
constraints     = none           ;
; OR
;constraints    = h-bonds        ;
;constraint-algorithm = LINCS    ;
;continuation   = no            ; create constraints to the start configuration or yes continue with the previous
constraints

; Generate initial velocities
gen-vel        = yes             ; Maxwell-Boltzmann distribution
gen-temp       = 300            ; initial temperature
gen-seed       = 4537566        ; Random seed

; Energy control
comm-mode      = Angular        ; remove translation and rotation of CM (used in gas phase)
; OR
```

```

;comm-mode           = Linear           ; remove translation of CM (used in liquid phase)
nstcomm             = 100              ; frequency for CM remove
comm-grps          = system            ; CM of the whole system
energygrps         = system

; Output control frequency
nstxout            = 1000
nstvout            = 1000
nstfout           = 1000
nstlog             = 1000
nstenergy          = 1000
nstxtcout         = 1000
xtc_precision     = 1000

; Neighborsearching and short-range nonbonded interactions
pbc                = xyz
cutoff-scheme      = verlet
nstlist            = 20
ns_type            = grid
rlist              = 2.0

; Long range correction for Coulomb
;coulombtype        = PME
; OR
coulombtype        = Generalized-Reaction-Field
epsilon_rf         = 1.5              ; dielectric of vacuum
rcoulomb           = 1.7             ; Coulomb cut-off

; Long range correction for Lennard-Jones
vdw_type           = Cut-off
rvdw               = 1.7             ; Lennard-Jones cut-off
rvdw-switch        = 1.6             ; Lennard-Jones starting the switch-off
vdw-modifier        = potential-switch
DispCorr           = EnerPres        ; Long range dispersion corrections for Energy and Pressure

; Thermostat
tcoupl             = v-rescale        ; Stochastic modified Berendsen thermostat
tc-grps            = system          ; Coupled the whole system together
tau_t              = 0.5             ; Time constant for temperature coupling in ps
ref_t              = 300             ; Target temperature
ld-seed            = 7650987         ; Random seed

```

Para executar a simulação com MD, primeiro é necessário copiar o diretório de topologia para o seu diretório de execução, usando o seguinte comando:

```
$ cp -r /usr/local/gromacs/share/gromacs/top/oplsaa.ff .
```

Em seguida, edite o arquivo `dci.gro` gerado pelo LigParGen para centralizar o soluto no meio da caixa cúbica de simulação de tamanho 8.4 nm, usando o seguinte comando:

```
$ gmx editconf -f dci.gro -o dci-ok.gro -box 8.4
```

Agora com os 3 arquivos de entrada (`min.mdp`, `dci-ok.gro` e `dci.top`) é possível gerar o arquivo binário `dci-min.tpr` que será utilizado para realizar a minimização da configuração inicial necessária antes da simulação propriamente dita, usando os seguintes comandos:

```
$ gmx grompp -f min.mdp -c dci-ok.gro -p dci.top -o dci-min.tpr
$ gmx mdrun -s dci-min.tpr -deffnm dci-min-run -v >& dci-min-run.out
```

Agora com os 3 arquivos de entrada (`nvt.mdp`, `dci-min-run.gro` e `dci.top`) é possível gerar o arquivo binário `dci-nvt.tpr` que será utilizado para realizar a simulação com MD, usando os seguintes comandos:

```
$ gmx grompp -f nvt.mdp -c dci-min-run.gro -p dci.top -o dci-nvt.tpr -maxwarn 1
$ gmx mdrun -s dci-nvt.tpr -deffnm dci-nvt-run -v >& dci-nvt-run.out
```

Atenção: a opção `-maxwarn 1` deve ser usada devido a opção `comm-mode = Angular` que retira a translação e rotação do centro de massa do soluto que só deve ser usado para simulações de uma única molécula ou um único aglomerado de moléculas. Sugerimos que leia com atenção os NOTES, WARNING e ERRORS que são escrito na tela após o comando `gmx grompp`.

Esta execução levou cerca de 10 min em um micro com processador i3 quadricore executando em 1 processador.

V. Realizando análises da simulação da molécula flexível em vácuo

Para verificar como a molécula DCI se deformou ao longo da simulação é preciso transformar o arquivo `dci-nvt-run.trr` no formato `gro`, que pode ser visualizado no programa VMD, usando o seguinte comando:

```
$ gmx trjconv -f dci-nvt-run.trr -s dci-nvt.tpr -o dci-nvt-traj.gro -pbc mol
$ vmd dci-nvt-traj.gro
```

A análise de várias propriedades como: energia potencial (opção 10), energia cinética (opção 11), energia total (opção 12), temperatura (opção 14), pressão (opção 15) e muitas outras, pode ser feita usando o seguinte comando e o arquivo de saída `energy.xvg` pode ser visualizado no programa `xmgrace`:

```
$ gmx energy -f dci-nvt-run.edr -s dci-nvt.tpr
$ xmgrace energy.xvg
```

Ver mais análises no tutorial do Julio Ortiz.

VI. Realizando a simulação da molécula flexível em solução

Com os arquivos de entrada para o programa GROMACS, vamos realizar uma simulação em solução aquosa (no ensemble NPT com densidade inicial próxima de 1.0 g/cm^3 , ou seja com uma caixa cúbica de lado 3.2 nm com cerca de 1000 moléculas de água). Nesta simulação manteremos as ligações químicas com átomos de hidrogênio rígidas para acelerarmos a simulação usando um passo de simulação de 1 fs e tempo total de 5 ns. São necessários os seguintes arquivos para realizar a simulação:

- `dci-w.top` (com definição do campo de força, nomes dos arquivos de topologia e quantidade de cada tipo de molécula),
- `dci-ok.itp` (com a topologia e parâmetros do potencial ligado: bond, angle, dihedral e improper, e potencial não-ligado: Lennard-Jones e Coulomb),
- `min.mdp` (com as palavras chaves para minimização da configuração inicial) e
- `npt-w.mdp` (com as palavras chaves para as simulações).

Exemplo de arquivo `dci-w.top`:

```
; The force-field files to be included
#include "oplsaa.ff/forcefield.itp"

; The solute and solvent topology files to be included
#include "dci-ok.itp"
#include "oplsaa.ff/tip3p.itp"

[ system ]
DCI in water
[ molecules ]
DCI      1
SOL     1000
```

Exemplo de arquivo `npt-w.mdp`:

```
integrator      = md           ; leap-frog integrator
dt              = 0.001        ; 1 fs timestep
nsteps         = 5000000       ; total time 5 ns
;constraints    = none
; OR
constraints     = h-bonds      ;
constraint-algorithm = LINCS    ;
lincs-iter      = 3           ; correct rotation on the constrained bond
```

```

continuation      = no          ; create constraints to the start configuration or yes continue with the previous constraints
; Generate initial velocities
gen-vel          = yes         ; Maxwell-Boltzmann distribution
gen-temp         = 300        ; initial temperature
gen-seed         = 4537566    ; Random seed
; Energy control
;comm-mode       = Angular     ; remove translation and rotation of CM (used in gas phase)
; OR
comm-mode        = Linear      ; remove translation of CM (used in liquid phase)
nstcomm          = 100        ; frequency for CM remove
comm-grps        = system      ; CM of the whole system
energygrps       = DCI SOL
; Output control frequency
nstxout          = 1000
nstvout          = 1000
nstfout          = 1000
nstlog           = 1000
nstenergy        = 1000
nstxtcout        = 1000
xtc_precision    = 1000
; Neighborsearching and short-range nonbonded interactions
pbc              = xyz
cutoff-scheme    = verlet
nstlist          = 20
ns_type          = grid
rlist            = 1.7
; Long range correction for Coulomb
coulombtype      = PME
; OR
;coulombtype     = Generalized-Reaction-Field
;epsilon_rf      = 78.5        ; dielectric of water
rcoulomb         = 1.4        ; Coulomb cut-off
; Long range correction for Lennard-Jones
vdw_type         = Cut-off
rvdw             = 1.4        ; Lennard-Jones cut-off
rvdw-switch      = 1.3        ; Lennard-Jones starting the switch-off
vdw-modifier     = potential-switch
DispCorr         = EnerPres   ; Long range dispersion corrections for Energy and Pressure
; Thermostat
tcoupl           = v-rescale   ; Stochastic modified Berendsen thermostat
tc-grps          = system      ; Coupled the whole system together
tau_t            = 0.5        ; Time constant for temperature coupling in ps
ref_t            = 300        ; Target temperature (K)
ld-seed          = 7650987    ; Random seed
; Barostat
Pcoupl           = berendsen
Pcoupltype       = isotropic   ; Pressure in xyz equally
tau_p            = 0.5        ; Time constant for pressure coupling in ps
ref_p            = 1.0        ; Target pressure (bar)
compressibility   = 4.51e-5    ; Compressibility for water

```

Agora executar os seguintes comandos para gerar a caixa inicial do soluto solvatado em água, depois minimizar e por fim simular:

```

$ gmx editconf -f dci-min-run.gro -o dci-box.gro -box 3.2
$ gmx solvate -cp dci-box.gro -p dci-w.top -o dci-w.gro

```

```

Output configuration contains 3145 atoms in 1034 residues
Volume          :      32.768 (nm^3)
Density         :      961.234 (g/l)
Number of SOL molecules:      1033

```

```

$ gmx grompp -f min.mdp -c dci-w.gro -p dci-w.top -o dci-min-w.tpr
$ gmx mdrun -s dci-min-w.tpr -deffnm dci-min-w-run -v >& dci-min-w-run.out
$ gmx grompp -f npt-w.mdp -c dci-min-w-run.gro -p dci-w.top -o dci-npt-w.tpr
$ gmx mdrun -s dci-npt-w.tpr -deffnm dci-npt-w-run -v >& dci-npt-w-run.out

```

Para visualizar a trajetória, usar o seguinte comando:

```

$ gmx trjconv -f dci-npt-w-run.trr -s dci-npt-w.tpr -o dci-npt-traj.gro -pbc mol
$ vmd dci-npt-traj.gro

```

Esta execução levou cerca de 40 h em um micro com processador i3 quadricore executando em 1 processador.