



دانشگاه صنعتی شریف
دانشکده‌ی مهندسی مکانیک

کلاس تدریس‌یار دینامیک مولکولی
تبدیل انرژی

عنوان:

تمرین سوم

نگارش:

محمد عرفان حمدی - ۹۹۲۰۹۰۴۱

مدرس:

حسین شایگانی

آبان ۱۴۰۰

فهرست مطالب

۱	پیش‌پردازش داده‌ها	۱
۱	۱-۱ اصلاح ساختار پروتئین	۱
۲	۱-۲ قرار دادن پروتئین در Water Box	۲
۵	۲ پردازش پروتئین	۵
۵	۲-۱ ریلکس کردن پروتئین با استفاده از NAMD	۵
۷	۲-۲ اعمال نیروی ثابت به پروتئین	۷
۹	۲-۳ ایجاد تغییرات در فایل conf	۹

فهرست شکل‌ها

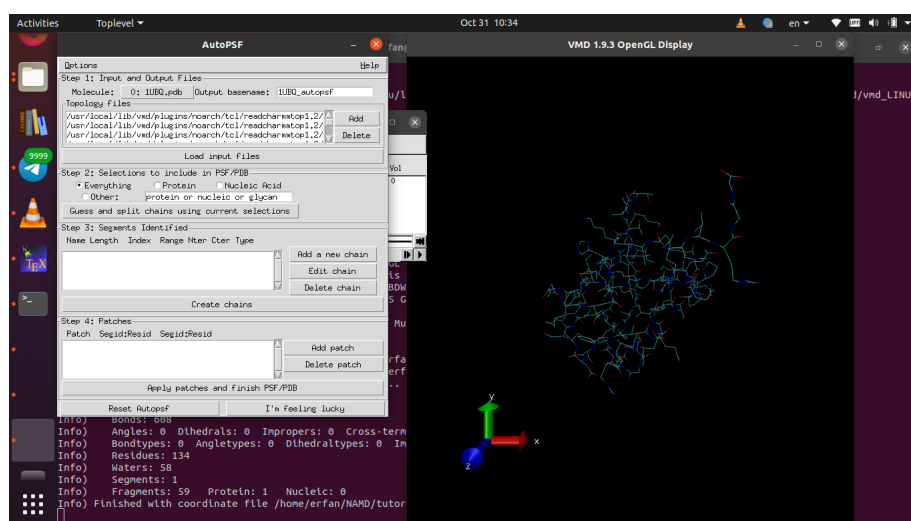
۱-۱	ماژول نرم افزار VMD	۱
۲-۱	استفاده از ماژول نرم افزار VMD	۳
۳-۱	پروتئین احاطه شده در یک مکعب پر از آب	۴
۱-۲	قرار دادن فایل‌های مورد نیاز در پوشه مربوطه	۵
۲-۲	فایل تنظیمات	۶
۳-۲	نتیجه خروجی ریلکسیشن	۶
۴-۲	افزودن نقاط جدید به ساختار پروتئین	۹
۵-۲	ساختار فایل‌ها در پوشه common	۹
۶-۲	ساختار فایل‌ها در پوشه pulldf	۱۰
۷-۲	تنظیمات اجرای شبیه‌سازی NAMD	۱۰
۸-۲	پارامترهای شبیه‌سازی	۱۱
۹-۲	تعیین پارامترهای مربوط به شبیه‌سازی	۱۱
۱۰-۲	لاگ نمایش داده شده در صفحه ترمینال	۱۲
۴۱۱۱-۲	فریم از حرکت پروتئین	۱۲

فصل ۱

پیش پردازش داده‌ها

۱-۱ اصلاح ساختار پروتئین

با توجه به نقص‌هایی که ممکن است در روش اندازه گیری به کار گرفته شده وجود داشته باشد، نیاز است که یک سری اصلاحات اولیه‌ای روی ساختار پروتئین انجام شود و فایل PSF مورد نیاز تهیه شود. بدین منظور میتوان از ماژول‌های نرم افزار VMD نیز استفاده کرد که برای این کار در نوار ابزار به بخش Extensions رفته و در بخش Modeling گزینه Automatic PSF Builder را انتخاب کرد.



شکل ۱-۱: ماژول نرم افزار VMD

در این جا روش دوم که با استفاده از کد نویسی در ترمینال است شرح داده می‌شود. در ابتدا اتم‌های پروتئین انتخاب شده و در متغیر ubq ریخته میشوند. سپس فایل PDB مدنظر ساخته می‌شود.

```
set ubq [atomselect top protein]
```

```
$ubq writepdb ubqp.pdb
```

سپس پکیج‌های مورد نیاز بارگزاری میشوند و فایل توپولوژی و تغییر نام‌های مورد نیاز برای نرم افزار VMD که بتواند قابل استفاده باشد انجام میگیرد.

```
package require psfgen
```

```
topology top_all27_prot_lipid.inp
```

```
pdbalias residue HIS HSE
```

```
pdbalias atom ILE CD1 CD
```

```
segment U {pdb ubqp.pdb}
```

در قدم بعدی مختصات اتم‌های هیدروژن حدس زده شده و سپس فایل‌های PDB و PSF ساخته میشوند.

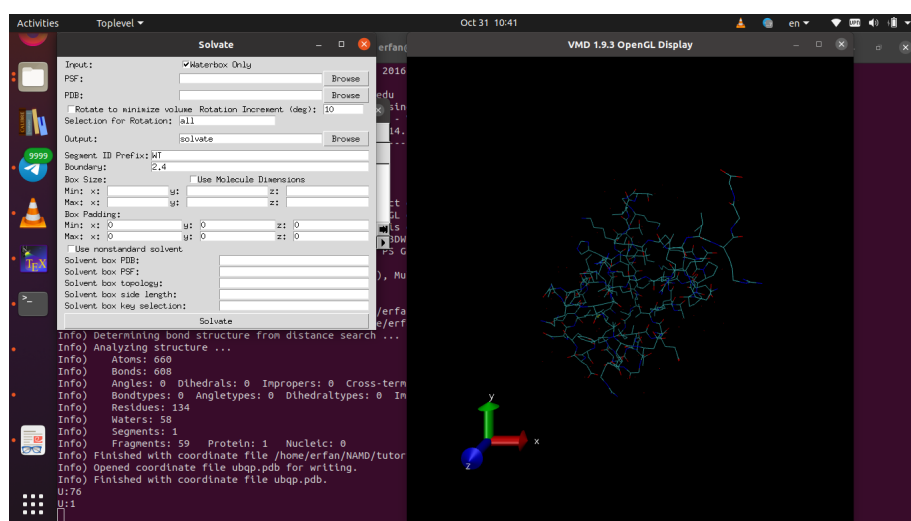
```
guesscoord
```

```
writepdb ubq.pdb
```

```
writesf ubq.psf
```

۲-۱ قرار دادن پروتئین در Water Box

حال لازم است که پروتئین را در یک جعبه آب قرار بدهیم تا در ادامه به ریلکس کردن آن پردازیم. برای این منظور میتوان از ماژول‌های نرم افزار VMD و یا کد نویسی استفاده کرد. به منظور استفاده از ماژول‌ها کافیست که در نوار ابزار به بخش Extensions رفته و در بخش Modeling گزینه Add Solvation Box را انتخاب کرد. و در صفحه باز شده فایل PDB و PSF مد نظر را انتخاب کرد.



شکل ۱-۲: استفاده از ماژول نرم افزار VMD

در این تمرین از روش دوم یعنی نوشتن کدها استفاده می‌شود. برای اینکار از بخش Extensions گزینه TK console انتخاب می‌شود که ترمینال برنامه‌نویسی در نرم افزار VMD است. سپس دستورات زیر وارد می‌شود.

```
package require solvate
```

ابتدا پکیج مورد نیاز فراخوانی می‌شود.

```
solvate ubq.psf ubq.pdb -t 5 -o ubq_wb
```

در این دستور، گزینه `-t` فاصله‌ای که باید پروتئین در بزرگترین بعد خود از دیواره باکس خود بگیرد را تعیین میکند. و گزینه `-o` نشانگر این است که فایل‌های خروجی باید پیشوند مشخص شده را داشته باشند.

```
set everyone [atomselect top all]
```

```
measure minmax $everyone
```

```
min : {10.442000389099121 7.947000026702881 -5.327000141143799}
```

```
max : {51.11899948120117 50.23099899291992 40.895999908447266}
```

در این دو دستور ابتدا همه اتم‌ها انتخاب و سپس دورترین و نزدیکترین نقاط نسبت به مبدا محاسبه میشوند. که میتوان با استفاده از این داده‌ها مختصات وسط پروتئین را نیز محاسبه کرد.

$$[x_c, y_c, z_c] = \frac{\min + \max}{2} \quad (1-1)$$

$$[x_c, y_c, z_c] = [30/775, 29/085, 17/785] \quad (2-1)$$

میتوان برای اینکار از دستور زیر هم استفاده کرد.

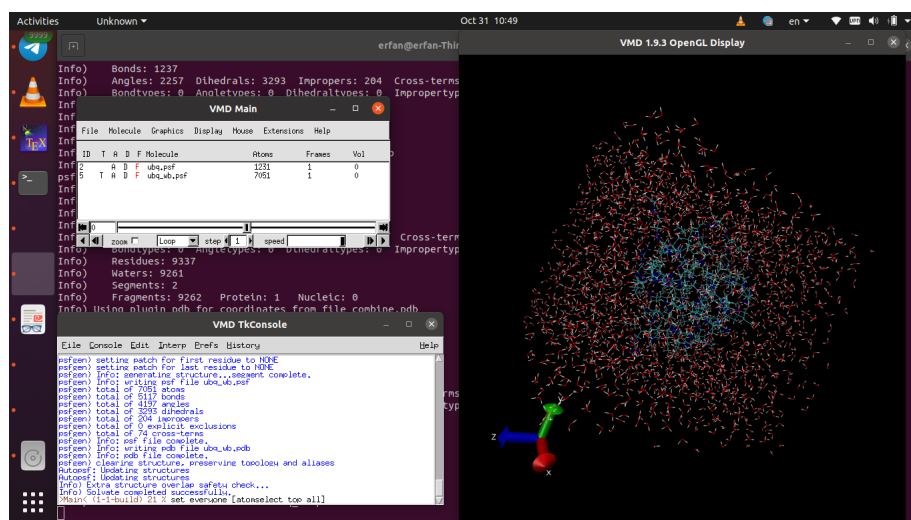
```
measure center $everyone
```

سپس اینبار اتم‌های پروتئین انتخاب شده و در یک فایل PDB ذخیره میشوند.

```
set selprotein [atomselect top protein]
```

```
$selprotein writepdb common/ubq_ww_eq.pdb
```

حال میتوان این فایل را در نرم افزار VMD باز کرد.



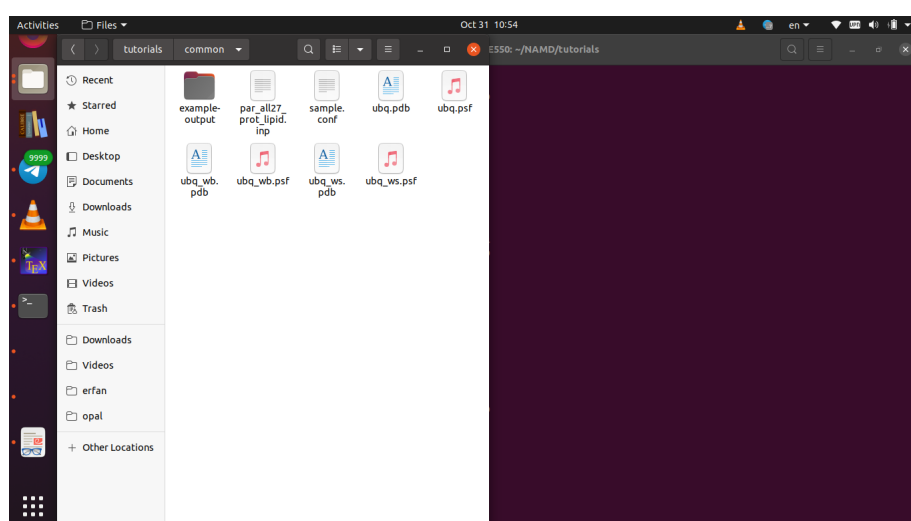
شکل ۱-۳: پروتئین احاطه شده در یک مکعب پر از آب

فصل ۲

پردازش پروتئین

۱-۲ ریلکس کردن پروتئین با استفاده از NAMD

حال که پروتئین در ظرف آب قرارگرفت لازم است که با استفاده از نرم افزار NAMD آن را ریلکس کرد. بدین منظور در پوشه‌ای که دارای فایل‌های کانفیگوریشن مربوطه به NAMD است رفته و بررسی میکنیم که آدرس فایل‌ها درست تنظیم شده باشند. که در اینجا بایستی پوشه common باشد. پس لازم است که فایل‌های مربوط به جعبه آب و فایل‌های PDB و PSF در پوشه مربوطه کپی شود.



شکل ۱-۲: قرار دادن فایل‌های مورد نیاز در پوشه مربوطه

حال به پوشه دارای فایل تنظیمات رفته و ترمینال را در آن اجرا میکنیم،

```

erfan@erfan-ThinkPad-E550: ~/NAMD/tutorials/t1-3-box
#####
## JOB DESCRIPTION
#####
# Minimization and Equilibration of
# Ubiquitin in a Water Box
#####
## ADJUSTABLE PARAMETERS
#####
structure      ../common/ubq_wb.psf
coordinates     ../common/ubq_wb.pdb
set temperature 310
set outputname  ubq_wb_eq
firsttimestep   0
#####
## SIMULATION PARAMETERS
#####
# Input
paratypeCharmm on
parameters     ../common/par_all27_prot_lipid.tnp
temperature     $temperature
#####
# Force-Field Parameters
exclude        scaled1-4
1-4scaling     1.0
cutoff         12.0
switching      on
switchdist     10.0
1,1
Top

```

شکل ۲-۲: فایل تنظیمات

سپس با اجرای دستور زیر عملیات ریلکسیشن روی پروتئین انجام می‌شود.

```
namd2 ubq_wb_eq.conf > log.txt
```

این دستور، نرم افزار NAMD را اجرا کرده و سپس نتیجه خروجی را در علاوه بر اینکه در ترمینال نشان می‌دهد، در فایل لاگ نیز ذخیره میکند.

```

erfan@erfan-ThinkPad-E550: ~/NAMD/tutorials/t1-3-box
ENERGY: 2400 238.2178 698.3340 307.2110 38.1383 -23956.5436 1619.8226 0.0000 0.0000
4529.8618 -16524.9581 310.0558 -21854.8199 -16508.2020 310.5537 945.3530 902.9239 69292.100
0 -16.1636 -15.2767
PRESSURE: 2500 -583.336 -146.794 109.555 -146.794 -340.122 349.607 109.555 349.607 -725.013
GPRESSURE: 2500 -406.389 -228.051 127.485 -372.26 -430.781 217.057 97.022 160.081 -654.395
PRESSAVG: 2500 34.981 -160.189 -41.028 -160.189 7.4722 208.916 -41.028 208.916 -133.058
GPRESSAVG: 2500 31.8524 -160.245 -50.3903 -167.538 10.7879 216.402 -23.7748 210.163 -133.928
ENERGY: 2500 260.4386 668.8358 306.7591 46.1166 -23915.6172 1585.5688 0.0000 0.0000
4584.8985 -16543.0156 313.0229 -21127.9142 -16522.4254 312.4912 -522.4994 -497.1615 69632.397
7 -30.2015 -30.4293
WRITING EXTENDED SYSTEM TO RESTART FILE AT STEP 2500
WRITING COORDINATES TO DCD FILE ubq_wb_eq.dcd AT STEP 2500
WRITING COORDINATES TO RESTART FILE AT STEP 2500
FINISHED WRITING RESTART COORDINATES
The last position output (seq=2500) takes 0.001 seconds, 86.215 MB of memory in use
WRITING VELOCITIES TO RESTART FILE AT STEP 2500
FINISHED WRITING RESTART VELOCITIES
The last velocity output (seq=2500) takes 0.000 seconds, 86.215 MB of memory in use
PRESSURE: 2600 -46.2413 111.639 194.93 111.639 443.392 135.447 194.93 135.447 -209.037
GPRESSURE: 2600 -61.4103 432.929 254.614 112.332 226.189 151.114 174.384 54.3261 55.764
PRESSAVG: 2600 252.944 -53.8185 109.9 -53.8185 -225.956 62.9101 109.9 62.9101 -105.36
GPRESSAVG: 2600 250.336 -63.6089 115.467 -42.9867 -231.412 61.3275 117.272 60.9889 -102.652
ENERGY: 2600 215.4449 673.2014 306.3243 42.9128 -23997.3963 1542.0512 0.0000 0.0000
4600.8040 -16610.6577 314.9116 -21217.4617 -16596.8815 312.7311 62.7080 106.8472 69058.893
3 -26.1243 -25.9096
WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 2600
CLOSING EXTENDED SYSTEM TRAJECTORY FILE
WRITING COORDINATES TO OUTPUT FILE AT STEP 2600
CLOSING COORDINATE DCD FILE ubq_wb_eq.dcd
WRITING VELOCITIES TO OUTPUT FILE AT STEP 2600
=====
Wallclock: 168.500381 CPUTime: 167.923096 Memory: 86.417969 MB
[Partition 0][Node 0] End of program
erfan@erfan-ThinkPad-E550: ~/NAMD/tutorials/t1-3-box$

```

شکل ۲-۳: نتیجه خروجی ریلکسیشن

نتایج در پوشه box3-1- ذخیره میشوند.

۲-۲ اعمال نیروی ثابت به پروتئین

حال که پروتئین ریلکس شده است میتوان محاسبات مد نظر را روی آن انجام داد. در اینجا ابتدا اتم کربن از residue اول را ثابت کرده و سپس درجهت برداری که اتم کربن residue اول را به اتم کربن residue آخر متصل میکند، نیروی ثابت به پروتئین اعمال می شود. بدین منظور لازم است که بردار اعمال نیرو محاسبه شود. این بردار از تفریق مختصات اتم آخر از اتم اول بدست می آید.

ابتدا اما لازم است که پارامترهای مربوط به ساکن بودن یا متحرک بودن residue را با پارامتر beta مشخص کرد. اگر این پارامتر دارای مقدار ۱ باشد به معنی ثابت بودن است و مقدار ۰ به معنی امکان تحرک داشتن.

```
set allatoms [atomselect top all]

$allatoms set beta 0

set fixedatom [atomselect top "resid 1 and name CA"]

$fixedatom set beta 1

$allatoms set occupancy 0

set smdatom [atomselect top "resid 76 and name CA"]

$smdatom set occupancy 11.54
```

نیروهای بین اتمی نیز با پارامتر occupancy تعیین می شود که در اینجا تمامی اتم ها به غیر از اتم کربن آخرین residue نیروی بین اتمی شان مقدار ۰ قرار داده شده است.

حال برای محاسبه جهت بردار نیرو ابتدا مختصات اتم آخر را از اتم اول کم کرده و سپس آن را تقسیم بر اندازه اش کرده تا بردار نرمال بدست آید. این کار با استفاده از این دستور ممکن است و میتوان آن را نیز به صورت دستی محاسبه کرد. ابتدا به صورت دستی محاسبه میکنیم و سپس نتیجه را با استفاده از دستور مقایسه میکنیم. مختصات اتم کربن residue اول و آخر.

```
resid 1 CA : {26.240999 24.927999 3.279000} resid 76 CA : {36.798000 39.930000 32.4
```

حال این دو مختصات شروع و پایان را از همدیگر کم میکنیم و بر اندازه این بردار تقسیم میکنیم.

$$v = [-۱۰/۵۵۵۷۰۰۱ - ۱۵/۰۰۲۰۰ - ۲۹/۱۷۴۹۹۸] \quad (۱-۲)$$

$$|v| = ۲۷/۱۵۷۷۴۸ \quad (۲-۲)$$

$$\frac{v}{|v|} = ۰/۳۸۸, ۰/۵۵۴۲, ۱/۰۷۴ \quad (۳-۲)$$

برای محاسبه از دستورهایی نرم افزار VMD نیز میتوان استفاده کرد.

```
set smdpos [lindex [$smdatom get {x y z}] 0]
```

```
set fixedpos [lindex [$fixedatom get {x y z}] 0]
```

```
vecnorm [vecsub $smdpos $fixedpos]
```

در این دستور ابتدا پارامترهای مربوط به مختصات اتم ثابت و متحرک در یک متغیر ذخیره شده و سپس در دستور آخر ابتدا این دو پارامتر را از هم دیگر کم کرده و سپس نرمال آن را حساب میکند.

```
{0.3063295295636324 0.43530883445678653 0.8465627194491839}
```

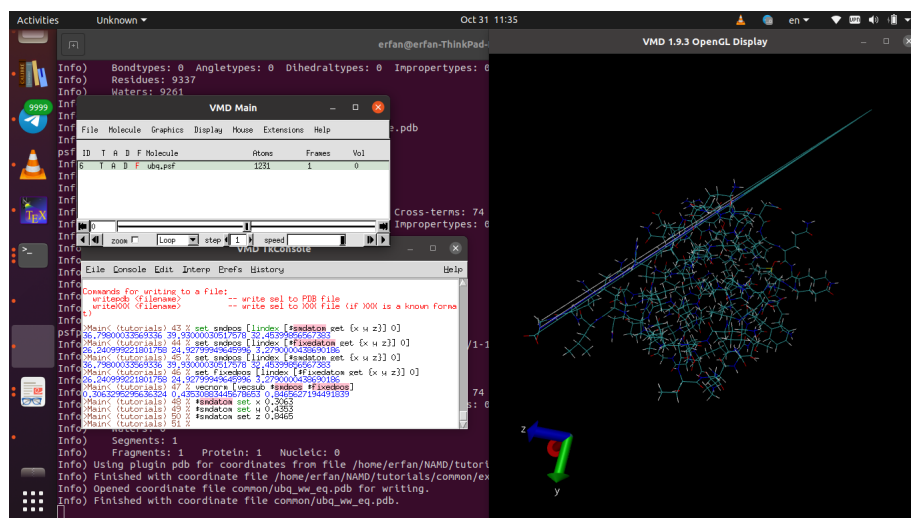
حال این پارامترهای بدست آمده در متغیرهای جدا گانه x, y, z ذخیره میشوند.

```
set x 0.30632
```

```
set y 0.43530
```

```
set z 0.84656
```

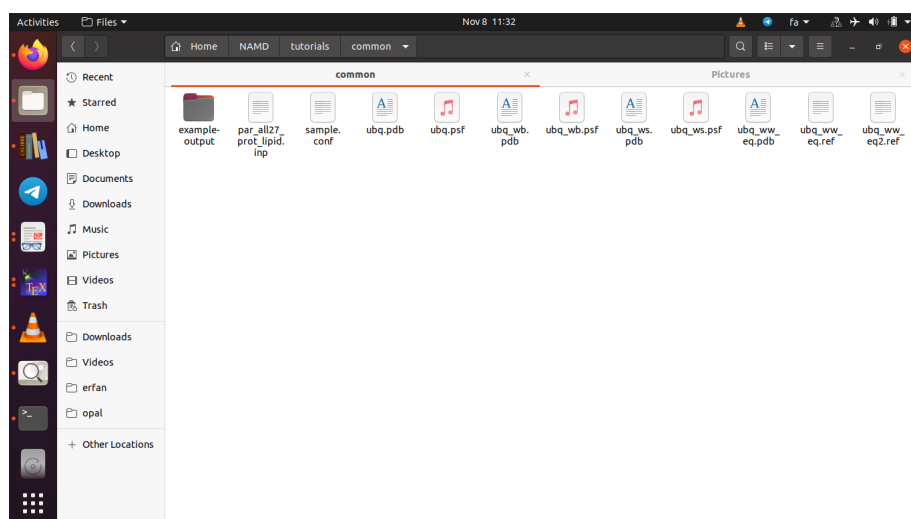
در نمایش پروتئین به این صورت تغییری ایجاد می شود.



شکل ۲-۴: افزودن نقاط جدید به ساختار پروتئین

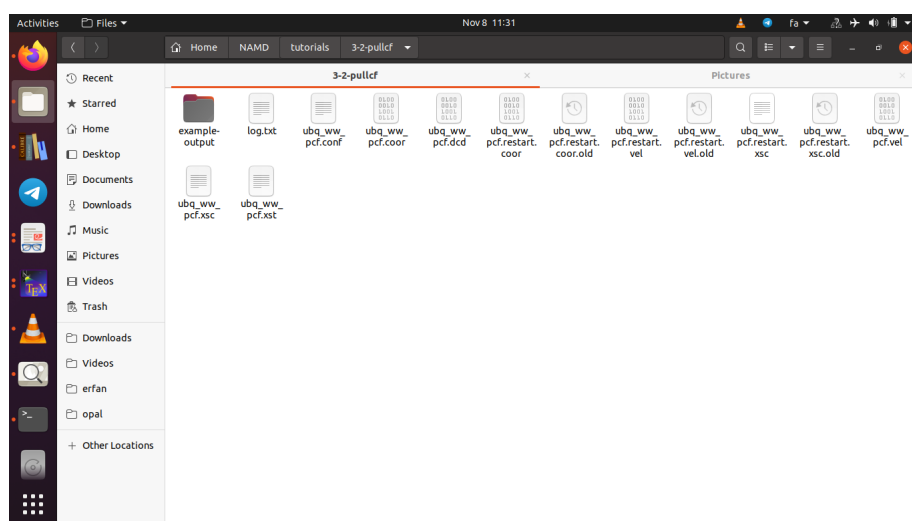
۲-۳ ایجاد تغییرات در فایل conf

بدین منظور ساختار پوشه‌ها را بدین صورت تغییر می‌دهیم که فایل‌های ساختاری و مختصات در پوشه common و فایل‌های conf را در پوشه pull_cf قرار می‌دهیم.



شکل ۲-۵: ساختار فایل‌ها در پوشه common

در همین صفحه نتایج حاصل از اجرای نرم افزار NAMD هم مشاهده می‌شود.



شکل ۲-۶: ساختار فایل‌ها در پوشه pulcf

به منظور انجام پردازش‌های لازم توسط نرم افزار NAMD لازم است که دستورات و شرایط اولیه لازم در فایل conf. داده شود.

در بخش اول Parameters Adjustable مسیر مربوط به فایل‌های ساختاری و مختصات و پیشوند فایل‌های ذخیره را مشخص میکنیم. سپس دمایی که در آن واکنش قرار است انجام بگیرد را مشخص میکنیم. سپس در پارامتر firsttimestep مشخص میکنیم که از کدام استپ زمانی شروع به بررسی کند.

```
#####
## ADJUSTABLE PARAMETERS ##
#####

structure      ../common/ubq.psf
coordinates    ../common/ubq_ww_eq.pdb
outputName     ubq_ww_pcf

set temperature 310

# Continuing a job from the restart files
if {0} {
  set inputname      myInput
  binCoordinates     $inputname.restart.coor
  binVelocities      $inputname.restart.vel ;# remove the "temperature" entry if you use this!
  extendedSystem     $inputname.xsc
}

firsttimestep    0
```

شکل ۲-۷: تنظیمات اجرای شبیه‌سازی NAMD

سپس در مرحله بعد سایر تنظیمات شامل مشخص کردن فایل توپولوژی و غیره را با ۰ یا ۱ کردن پارامتر داخل if تنظیم میکنیم.

```
#####
## SIMULATION PARAMETERS ##
#####

# Input
paratypecharmm      on
parameters           ../common/par_all27_prot_lipid.inp

# NOTE: Do not set the initial velocity temperature if you
# have also specified a .vel restart file!
temperature          $temperature

# Periodic Boundary conditions
# NOTE: Do not set the periodic cell basis if you have also
# specified an .xsc restart file!
if {0} {
  cellBasisVector1    20.0  0  0
  cellBasisVector2    0  20.0  0
  cellBasisVector3    0  0  50.0
  cellOrigin          0  0  0
}
wrapWater            on
wrapAll              on
```

شکل ۲-۸: پارامترهای شبیه‌سازی

سپس در نهایت با مشخص کردن فایل نهایی که دارای بردار نیرو است و اینکه ما می‌خواهیم به چه مقداری این شبیه‌سازی انجام شود، فایل را ذخیره می‌کنیم.

```
#####
## EXTRA PARAMETERS ##
#####

# Put here any custom parameters that are specific to
# this job (e.g., SMD, TclForces, etc...)

constantforce yes
constforcefile ../common/example-output/ubq_wv_eq2.ref

#####
## EXECUTION SCRIPT ##
#####

# Minimization
if {0} {
  minimize           100
  reinitvels         $temperature
}

run 20000 ;# 100ps
```

شکل ۲-۹: تعیین پارامترهای مربوط به شبیه‌سازی

سپس در نهایت به منظور اجرای شبیه‌سازی در لینوکس در پوشه‌ای که حاوی فایل `conf` بود دستور زیر را وارد می‌کنیم که به معنای اجرای شبیه‌سازی با استفاده از پارامترهای موجود در فایل تنظیمات و ذخیره سازی لاگ آن است.

```
namd2 sample.conf > log.txt
```

نتایج شبیه‌سازی را میتوان از طریق ترمینال در حین اجرای نرم افزار نیز مشاهده کرد.

```

erfan@erfan-ThinkPad-E550: ~/NAMD/tutorials/t1-3-box
ENERGY: 2400 238.2178 698.3340 387.2110 38.1383 -23956.5436 1619.8226 0.0000 0.0000
0 4529.8618 -16524.9581 310.8558 -21054.8199 -16508.2020 310.5537 945.3538 982.9239 69292.180
-16.1636 -15.2767

PRESSURE: 2500 -582.336 -146.794 109.555 -146.794 -340.122 349.667 109.555 349.667 -725.013
GPRESSURE: 2500 -406.388 -228.051 127.485 -372.26 -430.701 217.857 97.0922 160.801 -654.395
PRESSAVG: 2500 34.981 -160.189 -41.028 -160.189 7.4722 208.916 -41.028 208.916 -133.058
GPRESSAVG: 2500 31.8524 -160.245 -50.3903 -167.538 10.7879 216.402 -23.7748 210.163 -133.928
ENERGY: 2500 260.4360 668.8358 306.7591 46.1166 -23915.6172 1505.5688 0.0000 0.0000
7 4584.8985 -16543.0156 313.8229 -21127.9142 -16522.4254 312.4912 945.3538 982.9239 69292.180
-30.2015 -30.4293

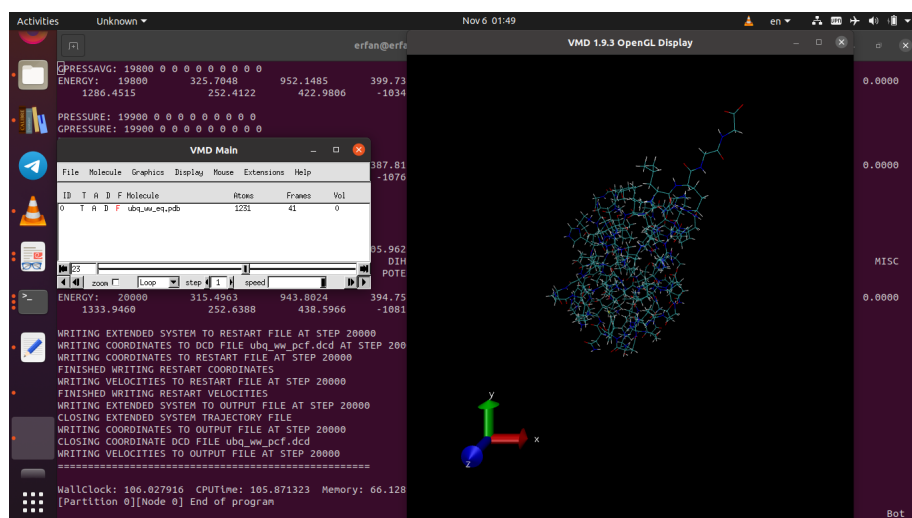
WRITING EXTENDED SYSTEM TO RESTART FILE AT STEP 2500
WRITING COORDINATES TO DCD FILE ubq_wb_eq.dcd AT STEP 2500
WRITING COORDINATES TO RESTART FILE AT STEP 2500
FINISHED WRITING RESTART COORDINATES
The last position output (seq=2500) takes 0.001 seconds, 86.215 MB of memory in use
WRITING VELOCITIES TO RESTART FILE AT STEP 2500
FINISHED WRITING RESTART VELOCITIES
The last velocity output (seq=2500) takes 0.000 seconds, 86.215 MB of memory in use
PRESSURE: 2600 -46.2413 111.639 194.93 111.639 443.392 135.447 194.93 135.447 -209.027
GPRESSURE: 2600 -61.4103 432.929 254.014 112.332 326.188 151.114 174.384 54.5261 55.764
PRESSAVG: 2600 252.944 -53.8185 109.9 -53.8185 -225.956 62.9101 109.9 62.9101 -105.36
GPRESSAVG: 2600 256.336 -63.6809 115.467 -42.9867 -231.412 61.3275 117.272 60.9889 -102.052
ENERGY: 2600 215.4449 673.2014 306.3243 42.9128 -23997.3963 1542.0512 0.0000 0.0000
3 4600.8040 -16610.6577 314.9116 -21217.4617 -16590.8815 312.7311 945.3538 982.9239 69292.180
-26.1243 -25.9096

WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 2600
CLOSING EXTENDED SYSTEM TRAJECTORY FILE
WRITING COORDINATES TO OUTPUT FILE AT STEP 2600
CLOSING COORDINATE DCD FILE ubq_wb_eq.dcd
WRITING VELOCITIES TO OUTPUT FILE AT STEP 2600
=====
WallClock: 168.500381 CPUtime: 167.923096 Memory: 86.417969 MB
erfan@erfan-ThinkPad-E550: ~/NAMD/tutorials/t1-3-box$

```

شکل ۲-۱۰: لاگ نمایش داده شده در صفحه ترمینال

در نهایت میتوان نتایج حرکت پروتئین را با بارگذاری فایل dcd. بر روی یک فایل ساختاری بدست آورد.



شکل ۲-۱۱: فریم از حرکت پروتئین

مراجع

- [1] B. Alberts. *Molecular Biology of the Cell (page 173, 473)*. Garland Science, 2014.
- [2] J. Berg. *Biochemistry 8th ed.* Freeman and Co, 2015.