大学院等高度副プログラム:基礎理学計測学「先端的研究法：低温電子顕微鏡」
2020/08/31-2020/09/04

CiCLE
Cyclic Innovation for Clinical Empowerment

Tutorial of Relion 3.1 : Innexin-6 2D-3D classification

Innexin-6を題材とした実習

Takeshi Kawabata (Institute for Protein Research, Osaka Univ.)
川端　猛　（大阪大学　蛋白質研究所）

2020/09/02

大阪大学

INSTITUTE for PROTEIN RESEARCH

MPI process, threads and GPU

Number of MPI process: プロセスの数

Number of threadsの数: スレッドの数

・Unixでは、いくつかのプロセスが同時に動いています。(topで確認)

・各プロセスはそれぞれいくつかのスレッドを使います。

・同じプロセスのスレッドは同じメモリを共有します。

・[全プロセス数]×[1プロセスが使うスレッド数] <= [CPUのスレッド総数]

・各プロセスがGPUを使います。[プロセス数] >= [GPUの枚数]
複数のプロセスが一つのGPUを使うことも可能。

・Refine3Dでは、必ず偶数個のGPU, [プロセス数] = [GPUの枚数] + 1

Reference-free 2D class averaging

[2GPUs, 35-50 min]

File Jobs Schedules | I/O | CTF | Optimisation | Sampling | Helix | Compute | Running

Import
Motion correction
CTF estimation
Manual picking
Auto-picking
2D classification
3D initial model
3D classification
3D auto-refine
3D multi-body
CTF refinement
Bayesian polishing
Mask creation
Join star files
Particle subtraction
Post-processing

2D Classification

Input images STAR file: tract/job010/particles.star ? Browse
Continue from here: ? Browse

Extract/refauto_manu_100mic_64pix/particles.star

※1GPU, 3threadsで
VIRT 47.7GBだが、
RES 600MBほど。4枚
同時でも負荷はそれ
ほど高くない。
44分ほど計算時間
がかかる。

I/O | CTF | Optimisation | Sampling | Helix | Compute | Running

Number of classes: 20 (20)
Regularisation parameter T: 2
Number of iterations: 25
Use fast subsets (for large data sets)? No
Mask diameter (A): 160 (160)
Mask individual particles with zeros? Yes
Limit resolution E-step to (A): -1

Alias :"refauto_manu_100mic_64pix"

I/O view | Job actions | Current: refauto_manu_100mic_64pix | Display:

I/O | CTF | Optimisation | Sampling | Helix | Compute | Running

Use parallel disc I/O? Yes
Number of pre-read particles: 3
Pre-read images into RAM? No
Scratch directory:
Cache images through disc? No
Use GPU acceleration? Yes
Which GPUs to use: 0:1

Input your GPU numbers.
Let's use 2 GPUs !

Input your GPU numbers.
Let's use 2 GPUs !

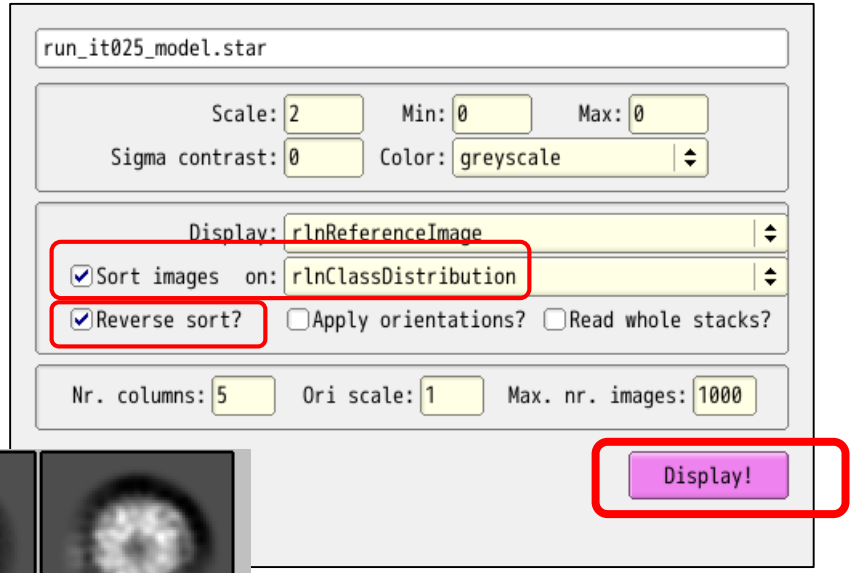
I/O | CTF | Optimisation | Sampling | Helix | Compute | Running

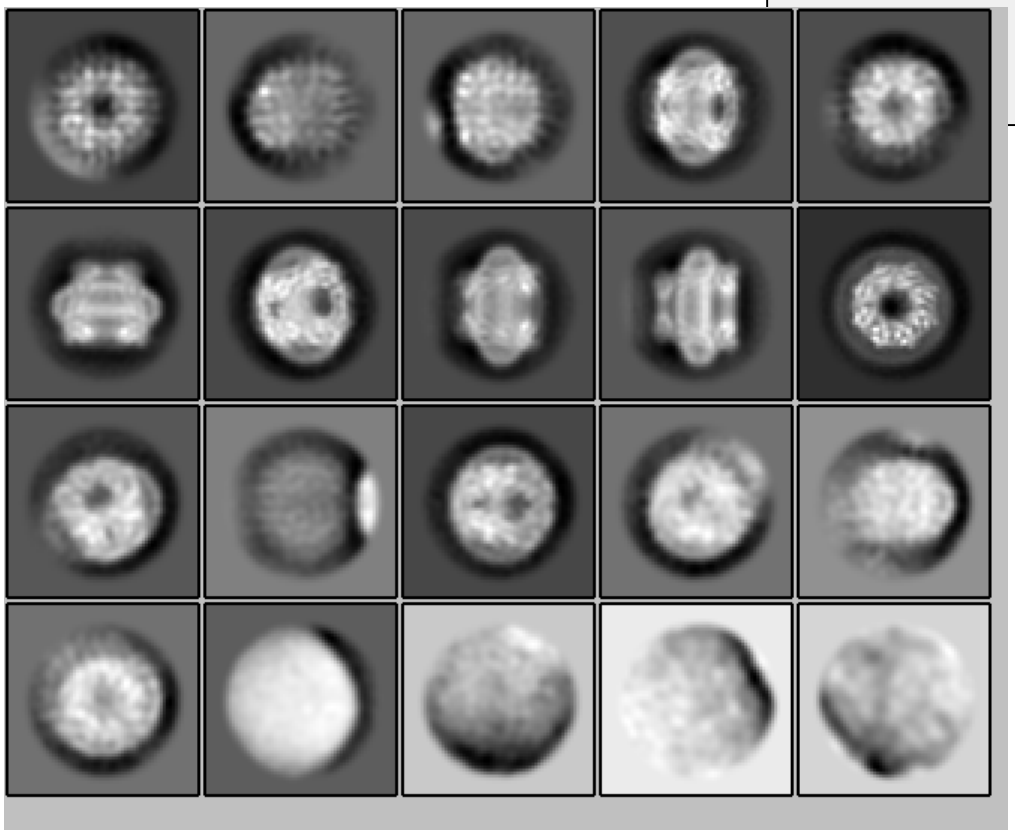
Number of MPI procs: 3 (3)
Number of threads: 2 (2)
Submit to queue? No
Queue name: openmpi
Queue submit command: qsub
Standard submission script: relion-3.1/scripts/qsub.csh ? Browse
Minimum dedicated cores per node: 24
Additional arguments:

Run!

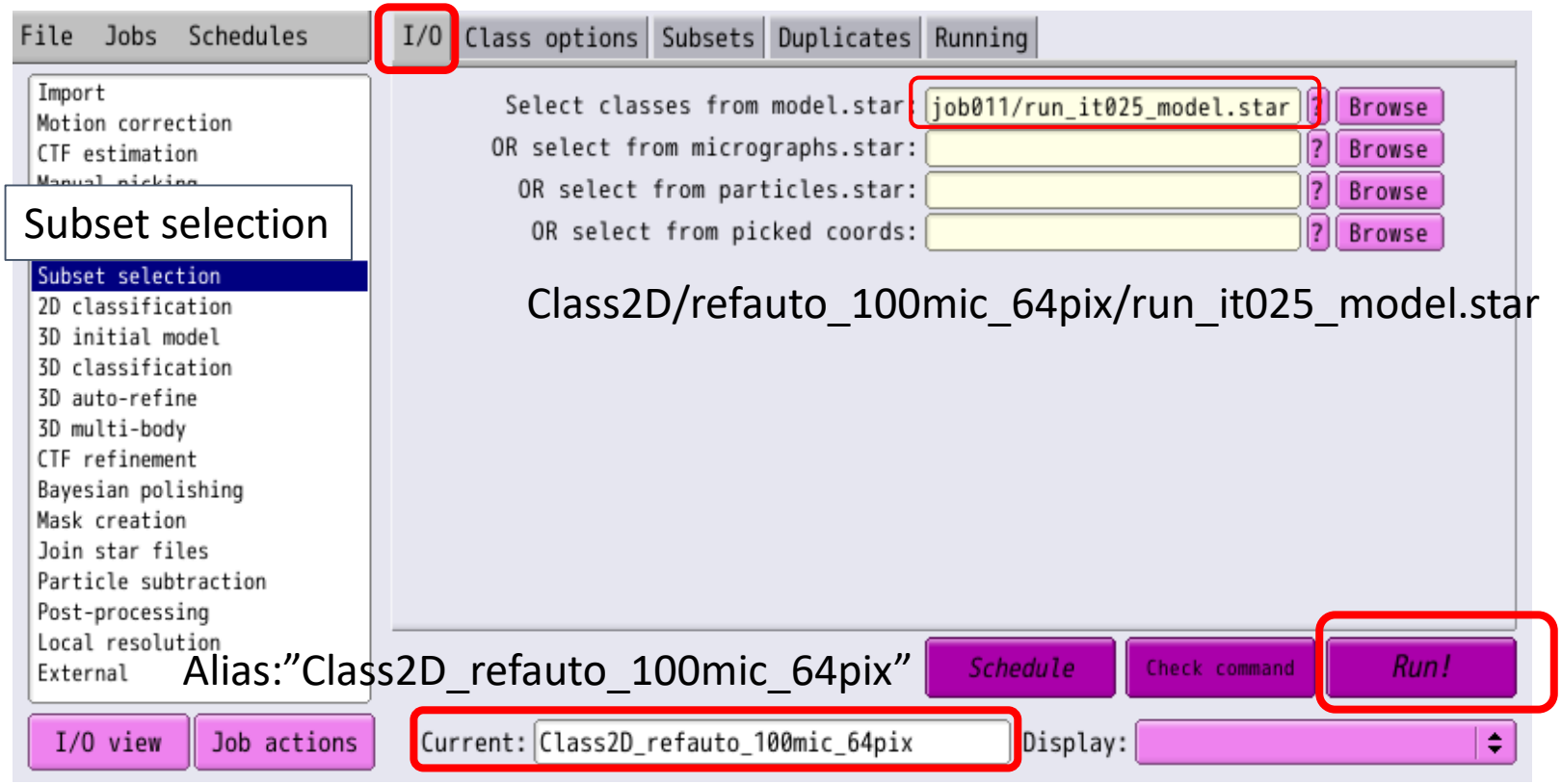
[#MPI] = [#GPU] +1

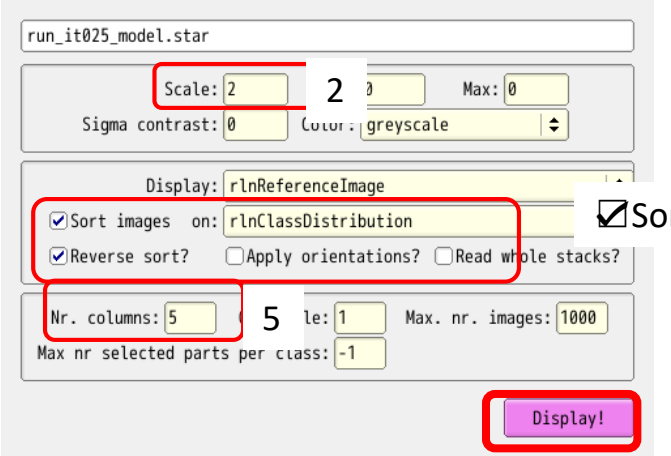
Reference-free 2D class averaging





2. Reference-free 2D class averaging: Subset selection



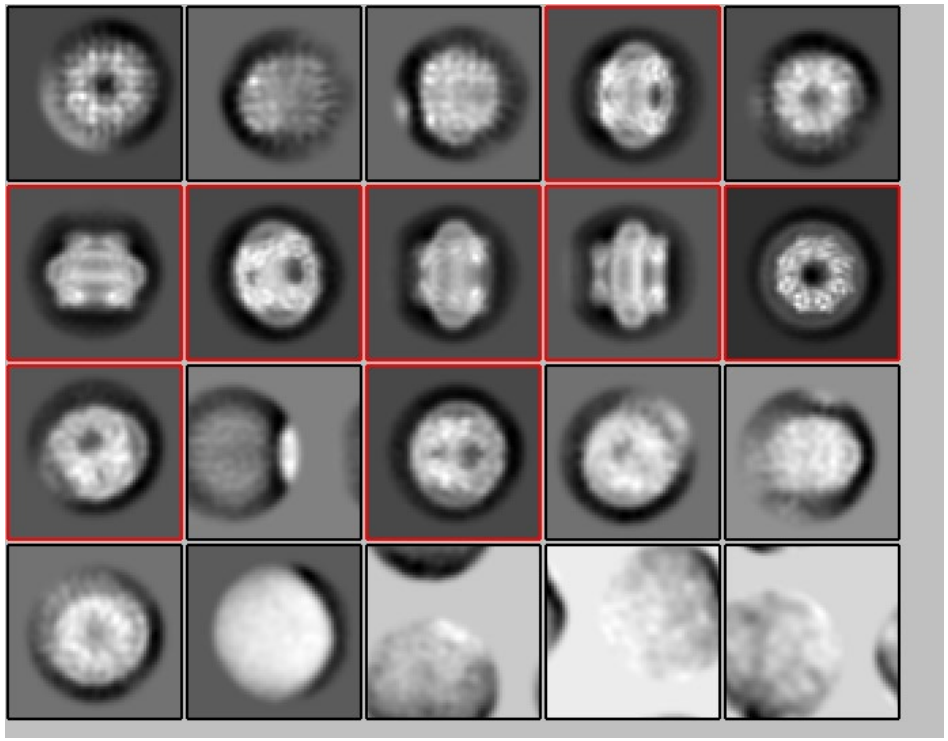


☑Sort images on: rlnClassDistribution

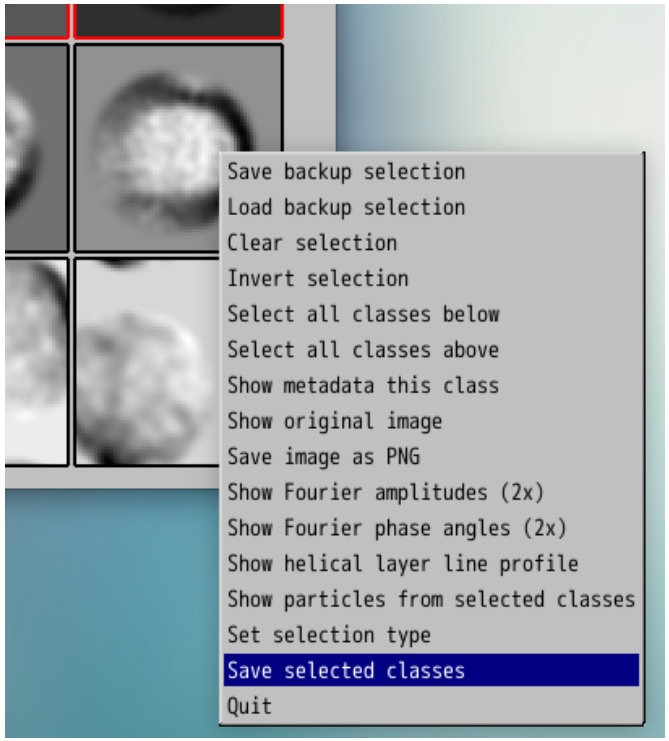
クラスに属する粒子数の数でソート

2. Reference-free 2D class averaging

Among the 20 classes, select a few nice-looking classes, by the left button of the mouse. Selected classes are indicated by red boxes.



Click the right mouse button, and Select [Save selected classes]



Total number of particles from 100 micrographs is **60292**
i.e. on average there were 603 particles per micrograph.

Saved Select/job012/class_averages.star with 8 selected images.
Saved Select/job012/particles.star with **25617** selected particles.

3. De novo 3D model generation

[1GPU, 13 min]

File Jobs Schedules | I/O | CTF | Optimisation | SGD | Compute | Running

Import
Motion correction
CTF estimation
Manual picking
Auto-picking
Particle extraction
Subset selection
2D classification
3D initial m
3D classifi
3D auto-refine
3D multi-body
CTF refinement
Bayesian polishing
Mask creation
Join star files
Particle subtraction
Post-processing
Local re
External

3D Initial model

Input images STAR file: elect/job012/particles.star ? Browse

Select/Class2D_refauto_100mics_64pix/particles.star

*relion_refine*が実行される。
VIRT 47.4GB
RES 482 MB
計算時間は6分ほど。

Alias : C8_Class3D_refauto_100mics_64pix

I/O view | Job actions | Current: C8_Class2D_refauto_100mics_64pix | Display:

I/O | CTF | Optimisation | SGD | Compute | Running

Number of classes: 1 — 1
Mask diameter (A): 160 — 160
latten and enforce non-negative solvent? Yes
Symmetry: C8 — C8
Initial angular sampling: 15 degrees
Offset search range (pix): 6
Offset search step (pix): 2

CTF | Optimisation | SGD | Compute | Running

Number of initial iterations: 50
Number of in-between iterations: 200
Number of final iterations: 50
Write-out frequency (iter): 10
Initial resolution (A): 35
Final resolution (A): 15
Initial mini-batch size: 100
Final mini-batch size: 500
Increased noise variance half-life: -1

Input your GPU number

Optimisation | SGD | Compute | Running

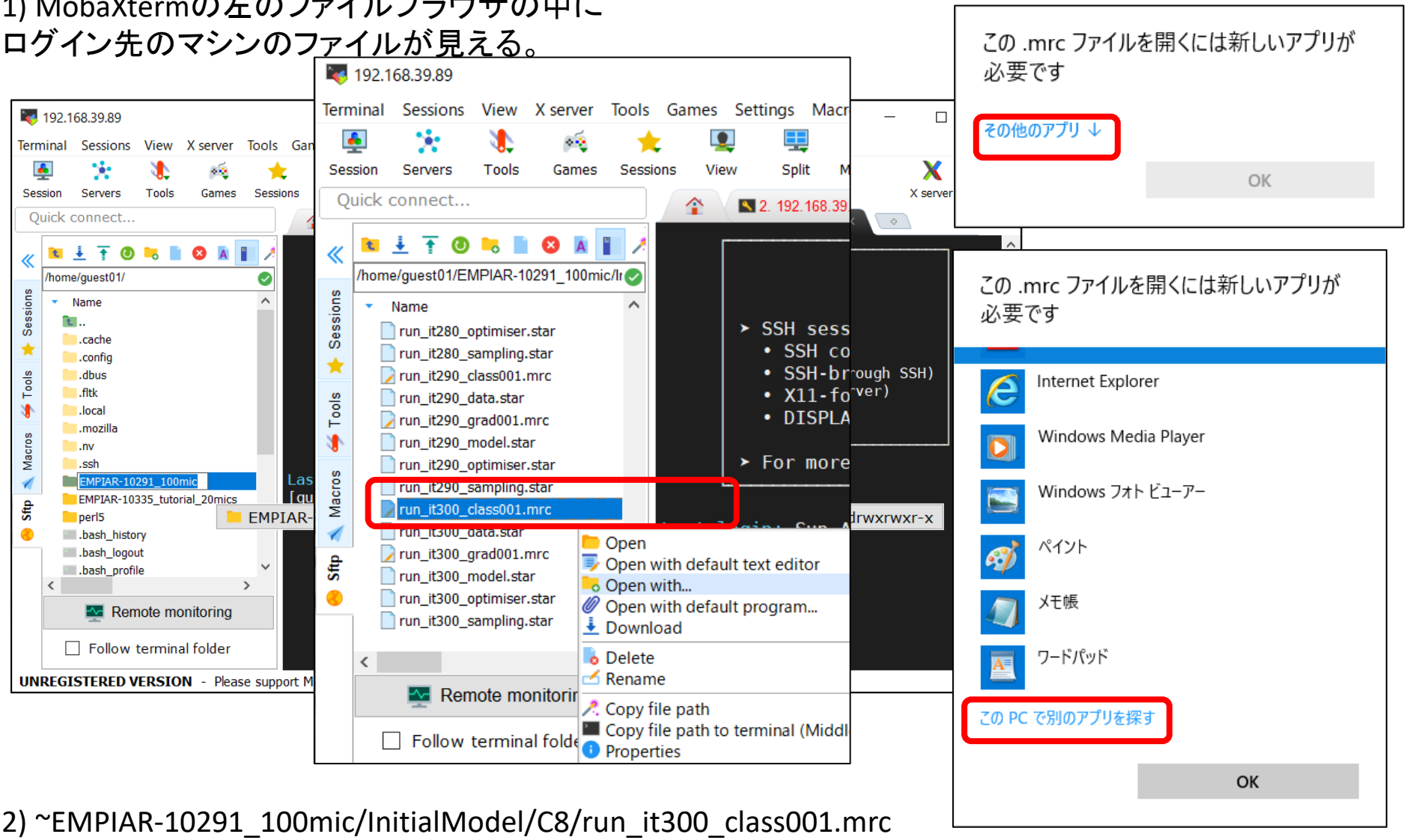
Use parallel disc I/O? Yes
Number of pooled particles: 3
Skip padding? No
Skip gridding? No
re-read all particles into RAM? No
particles to scratch directory:
terations through disc? No
Use GPU acceleration? Yes — Yes
Which GPUs to use: 0

tion | SGD | Compute | Running

er of MPI procs: 1 — 1
mber of threads: 6 — 6
Submit to queue? No
Queue name: openmpi
Queue submit command: qsub
rd submission script: relion-3.1
cated cores per node: 24
Addi
Run!

3. 3D密度マップ(*.mrc)の表示方法

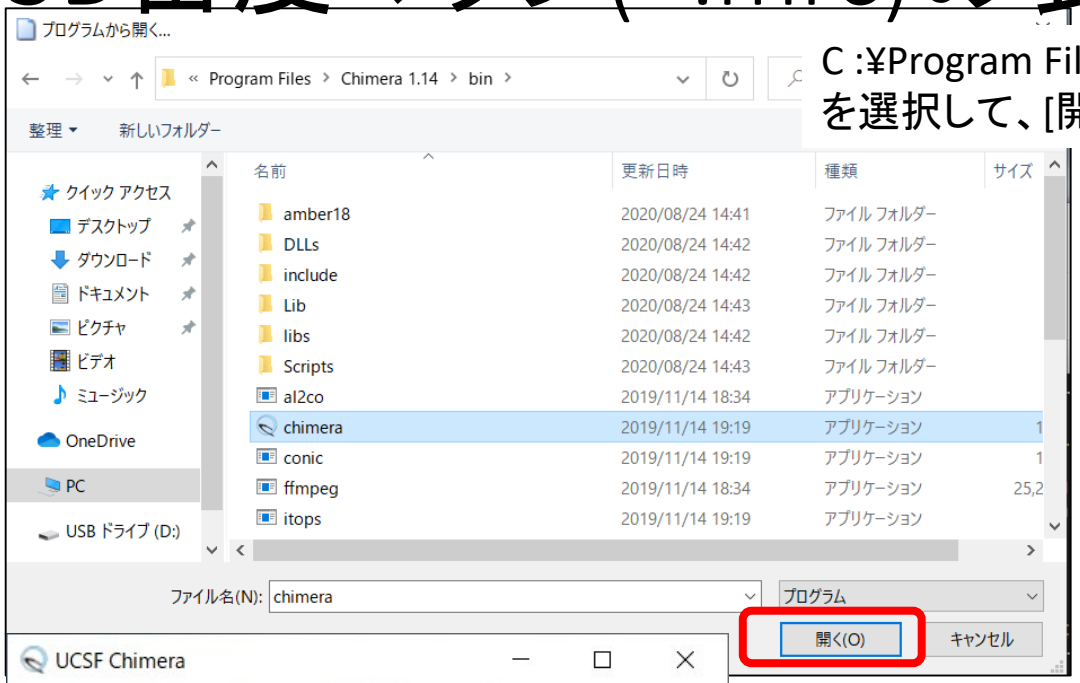
1) MobaXtermの左のファイルブラウザの中に
ログイン先のマシンのファイルが見える。



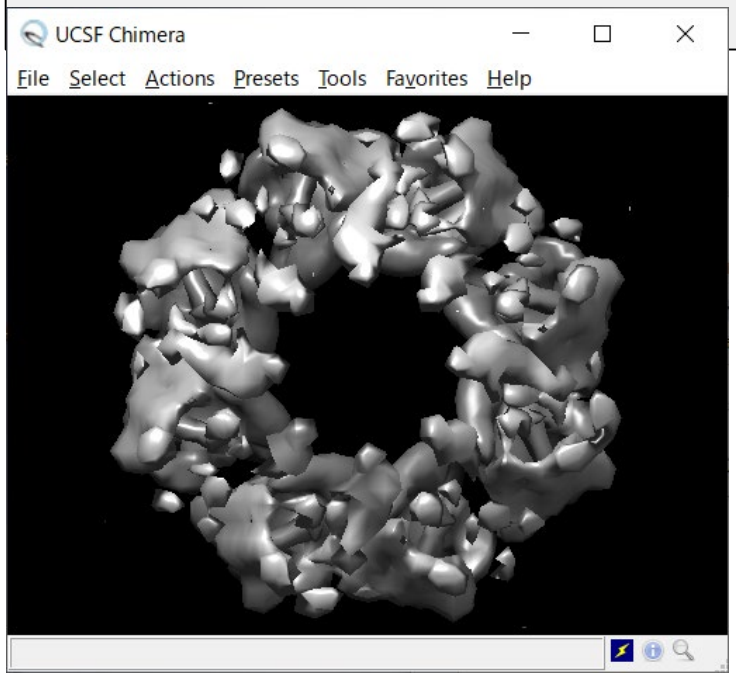
2) ~EMPIAR-10291_100mic/InitialModel/C8/run_it300_class001.mrc
を選択し、マウス右クリックのポップアップで、[Open with...]を選択

3) .mrcファイルを開くアプリの設定を行う

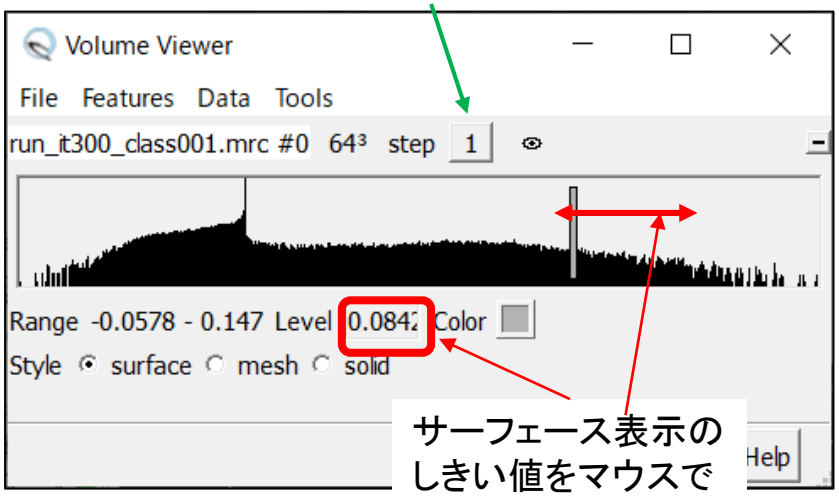
3. 3D密度マップ(*.mrc)の表示方法



C :¥Program Files ¥ Chimera1.14¥bin¥chimera
を選択して、[開く]



格子点のステップを調節可能。
2の場合2x2x2の格子点を一つにまとめて表示



サーフェース表示の
しきい値をマウスで
調節可能

4.1 Unsupervised 3D classification

[2GPUs, 45 min]

Select/Class2D_refauto_100mics_64pix/particles.star

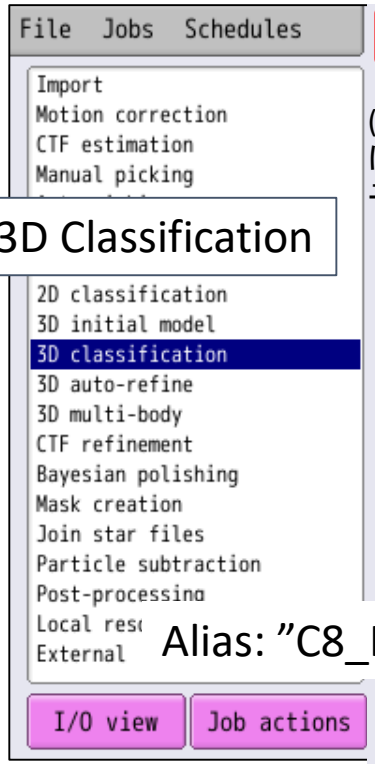
(選択した11個のクラスに属する1156個の粒子)

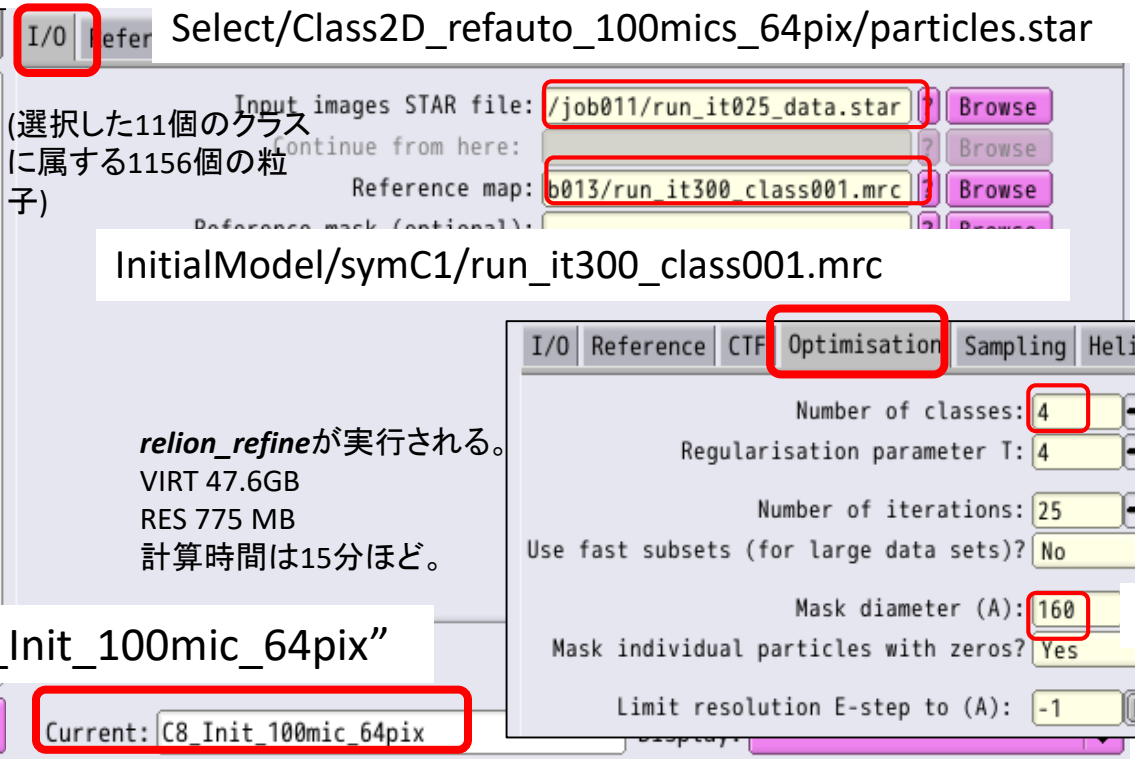
3D Classification

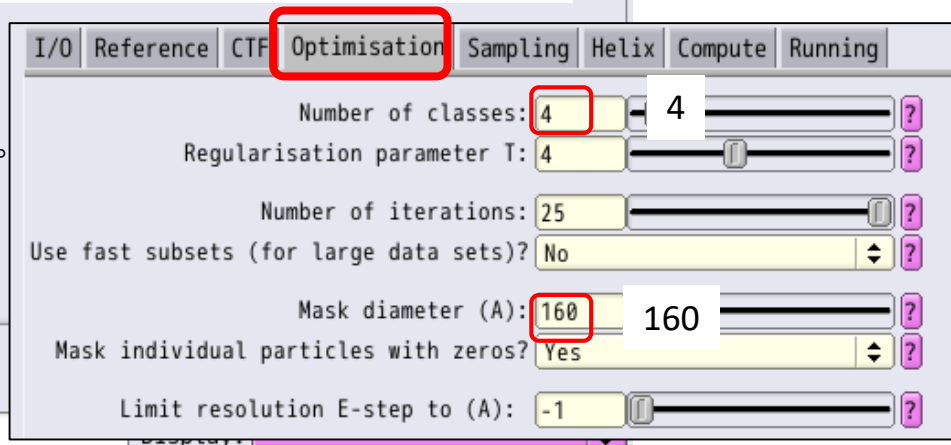
InitialModel/symC1/run_it300_class001.mrc

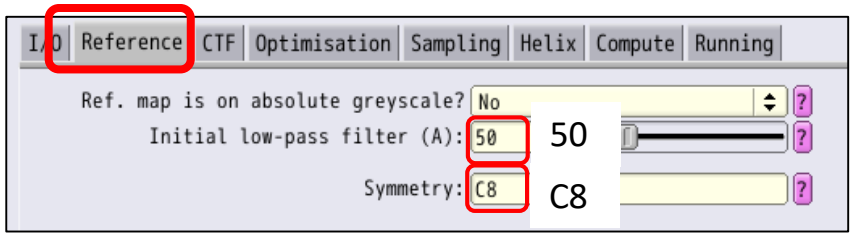
*relion_refine*が実行される。
VIRT 47.6GB
RES 775 MB
計算時間は15分ほど。

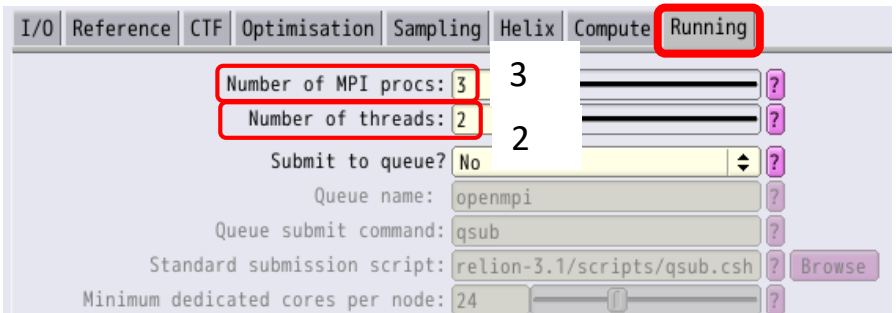
Alias: ”C8_Init_100mic_64pix”

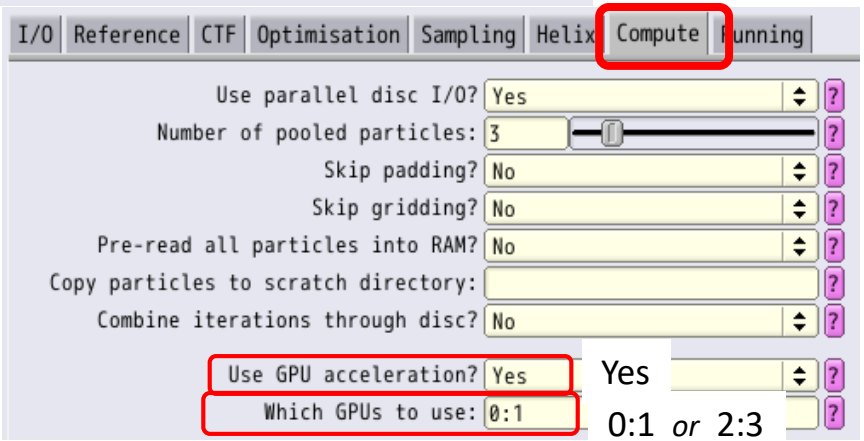












Input your GPU numbers.
Let’s use 2 GPUs !

4.1 Unsupervised 3D classification

RELION-3.1-beta-commit-9575df: /home/guest01/EMPIAR-10291_100mic (embox2)

File Jobs Schedules

I/O | Reference | CTF | Optimisation | Sampling | Helix | Compute | Running

Import
Motion correction
CTF estimation
Manual picking
Auto-picking
Particle extraction
Subset selection
2D classification
3D initial model
3D classification
3D auto-refine
3D multi-body
CTF refinement
Bayesian polishing
Mask creation
Join star files
Particle subtraction
Post-processing
Local resolution
External

Number of MPI procs: 3
Number of threads: 2
Submit to queue? No
Queue name: openmpi
Queue submit command: qsub
Standard submission script: relion-3.1/scripts/qsub.csh Browse
Minimum dedicated cores per node: 24
Additional arguments:

Schedule | Check command | Continue!

I/O view | Job actions | Current: 014: Class3D/C8_Init_100mic_64pix/ | Display:

Finished jobs
014: Class3D/C8_Init_100mic_64pix/
013: InitialModel/C8_Class2D_refauto_1
012: Select/Class2D_refauto_100mic_64p
011: Class2D/refauto_manu_100mic_64pix
010: Extract/refauto_manu_100mic_64pix
009: AutoPick/manu_100mic_64pix/
008: AutoPick/manu_5mic_64pix/
007: Select/manupick_5mic_64pix/
006: Class2D/manupick_5mic_64pix/

Running jobs

Scheduled jobs

Input
011: Cl
013: ni

Output

in: run_it025_data.star
in: run_it300_class001.mrc
out: run_it025_data.star
out: run_it025_model.star
out: run_it025_class002.mrc
out: run_it025_class003.mrc
out: run_it025_class004.mrc

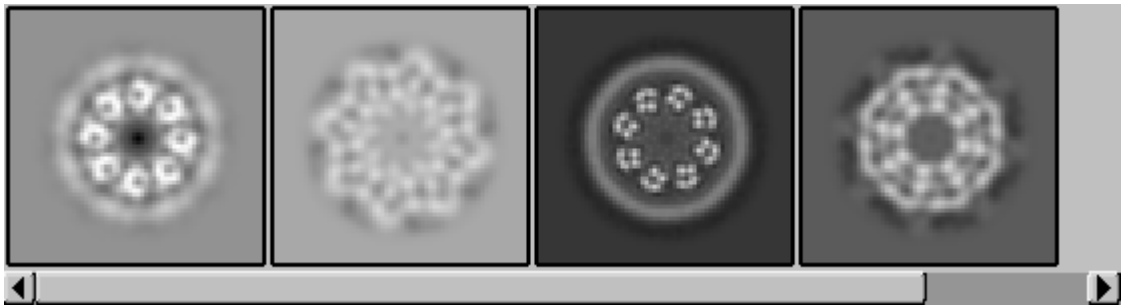
run_it025_model.star

Scale: 2 Min: 0 Max: 0
Sigma contrast: 0 Color: greyscale

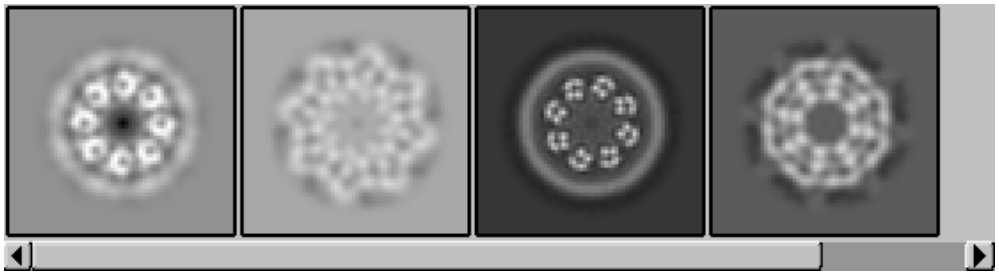
Display: rlnReferenceImage
Sort images on: rlnClassDistribution
Reverse sort? Apply orientations? Read whole stacks?

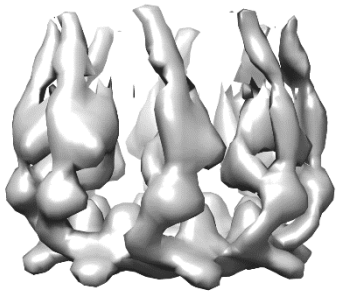
Nr. columns: 5 Ori scale: 1 Max. nr. images: 1000

Display!



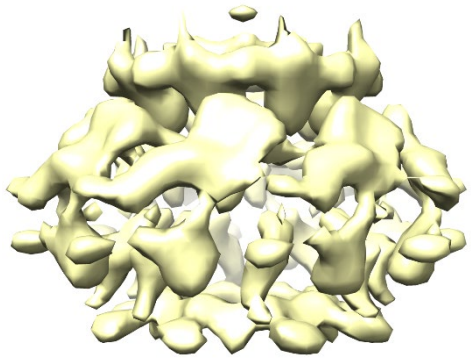
4.1 Unsupervised 3D classification





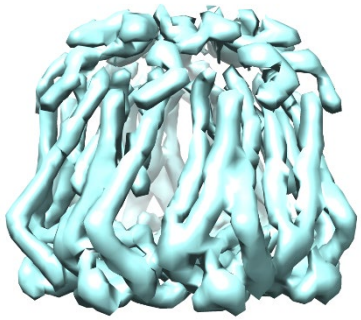
run_it025_class001.mrc

ClassDistribution
0.240319



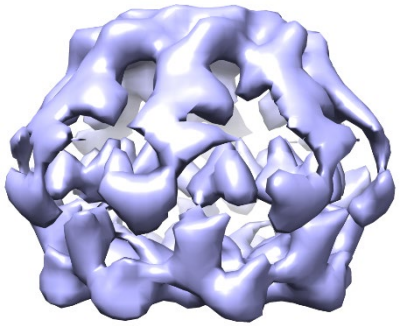
run_it025_class002.mrc

ClassDistribution
0.198251



run_it025_class003.mrc

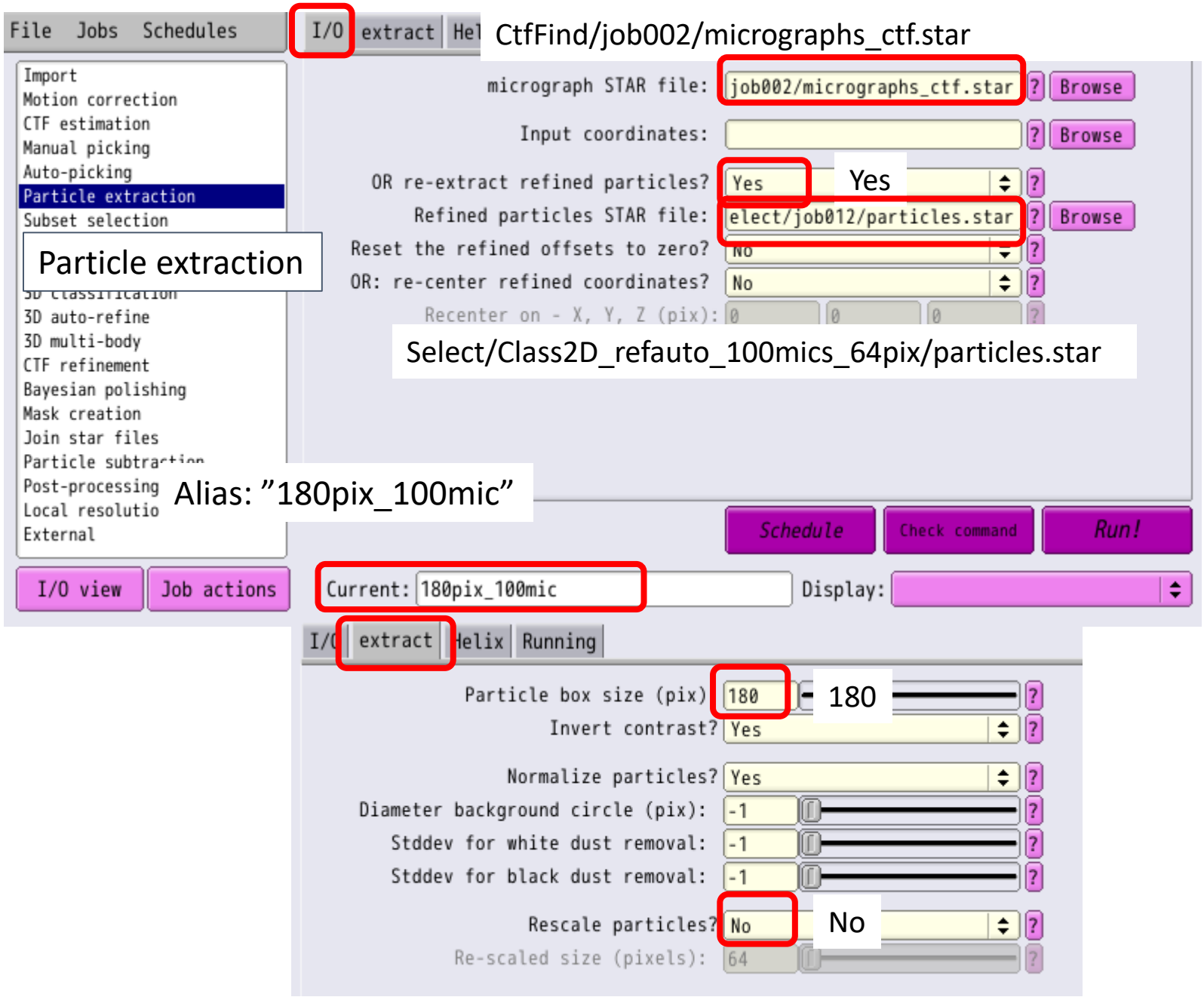
ClassDistribution
0.349774



run_it025_class004.mrc

ClassDistribution
0.211656

Extract particles with 180 x 180 pixels



Convert 64^3map into 128^3 map

Type following UNIX commands:

```
[guest01@embox~]$ cd
[guest01@embox~]$ cd EMPIAR-10291_100mic
[guest01@embox EMPIAR-10290_100mic]$  cd Class3D/C8_Init_100mic_64pix
[guest01@embox C8_Init_100mic_64pix]$  relion_image_handler --i
run_it025_class003.mrc --o run_it025_class003_box180.mrc --new_box 180
--angpix 3.465 --rescale_angpix 1.232
```

Type the class number of the your best looking class !!
It may be **class001** or **class002** or **class003** or **class004**.

5. High resolution 3D refinement

[2GPUs, 26 min]

Extract/180pix_100mic/particles.star

3D auto-refine

Class3D/C8_Init_100mic_64pix/run_it025_class**003**_box180.mrc

Type the class number of the your best looking class !!
It may be **001** or **002** or **003** or **004**.

Tutorial says low-pass filter is 50Å, but 30 Å is better for this case.

※3つの***relion_refine_mpi***のジョブが実行される。そのうち2つのジョブがVIRT 34.5GB、RES 1.5GBずつメモリを使う。計算時間は35分ほど。

Alias: 180pix_100mic

※GPU1枚でも実行は開始されるが、20分後に、GPUのout of memoryのエラーで、不正終了する。

Input your GPU numbers.
Let's use 2 GPUs !

[Number of MPI proc] should be odds number more than 2 (**3** or **5** or **7** ....)

5. High resolution 3D refinement

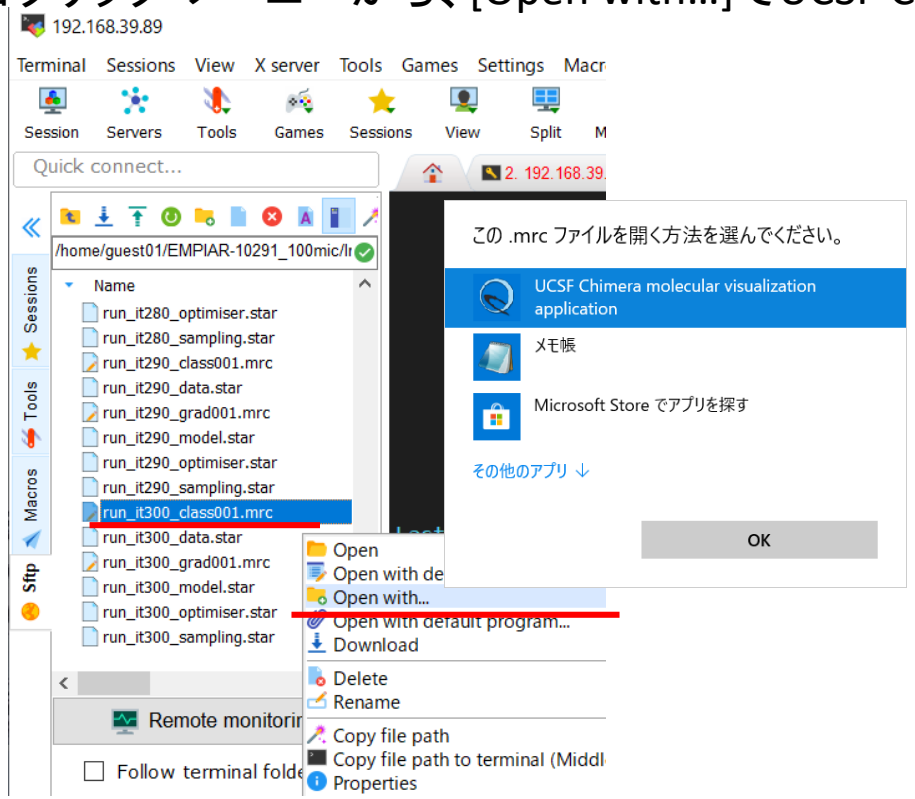
```
011: Class2D/refauto_manu_100mic_64pix
010: Extract/refauto_manu_100mic_64pix
009: AutoPick/manu_100mic_64pix/
```

You can know the resolution from the message from Relion GUI.

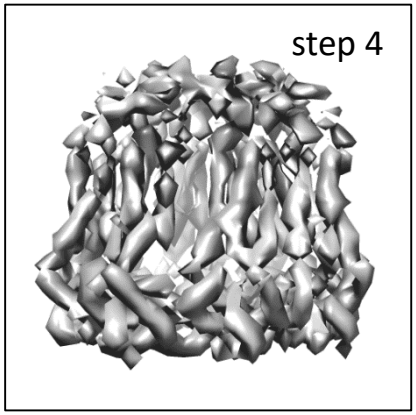
or type a following command:
`tail Refine3D/180pix_100mic/run.out`

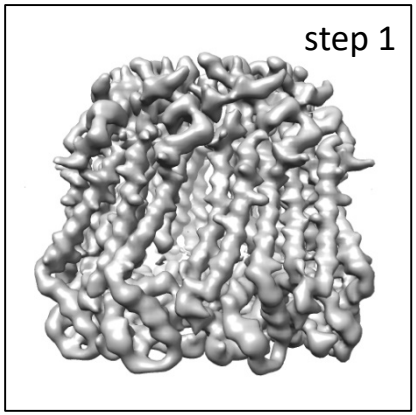
```
 60/  60 sec ..........................................................~~(,_,">
Auto-refine: Refinement has converged, stopping now...
Auto-refine: + Final reconstruction from all particles is saved as: Refine3D/job019/run_class001.mrc
Auto-refine: + Final model parameters are stored in: Refine3D/job019/run_model.star
Auto-refine: + Final data parameters are stored in: Refine3D/job019/run_data.star
Auto-refine: + Final resolution (without masking) is: 4.10667
```

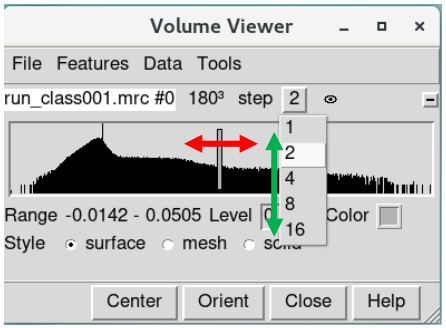
MobaXtermのファイルブラウザで、Refine3D/180pix_100mic/run_class001.mrcを選び、
右クリック・メニューから、[Open with...]でUCSF Chimeraで開いて確認する。



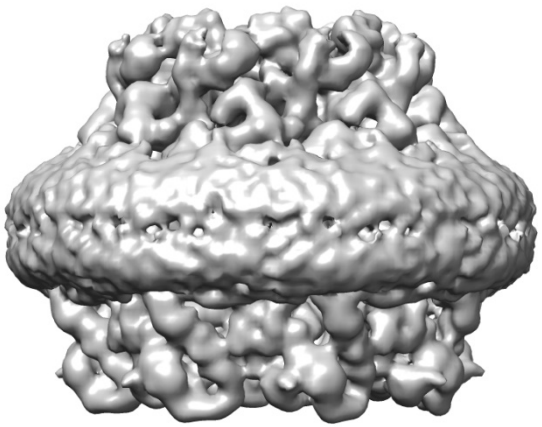
Refine3D/first3dref/run_class001.mrc



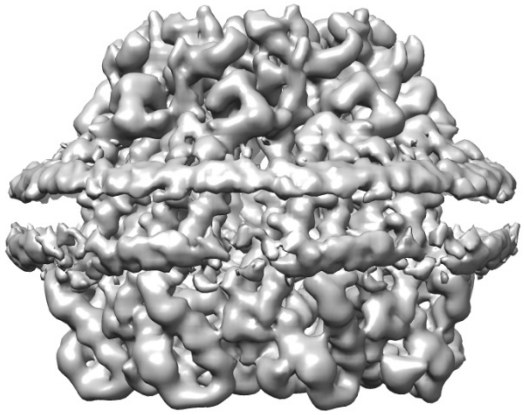




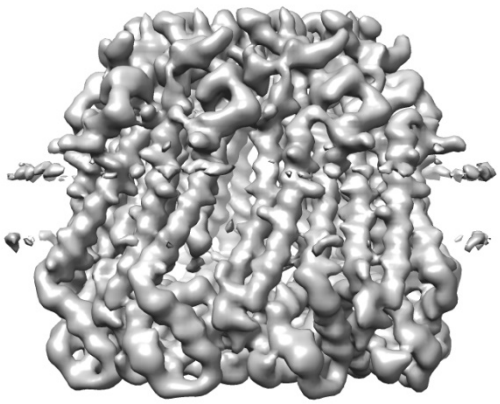
Level threshold for Refined3d map



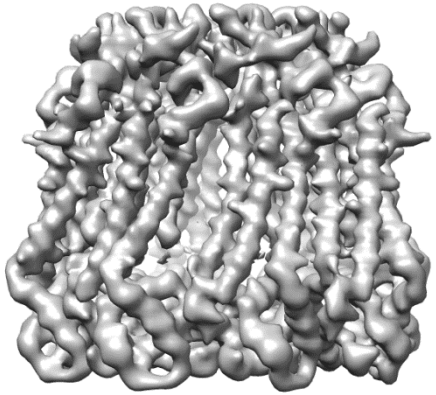
Level = 0.01



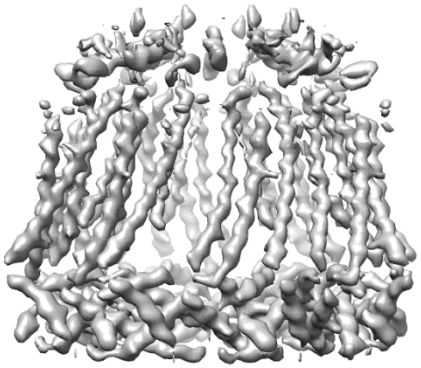
Level = 0.013



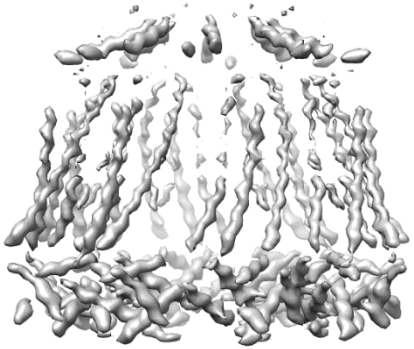
Level = 0.016



Level = 0.018

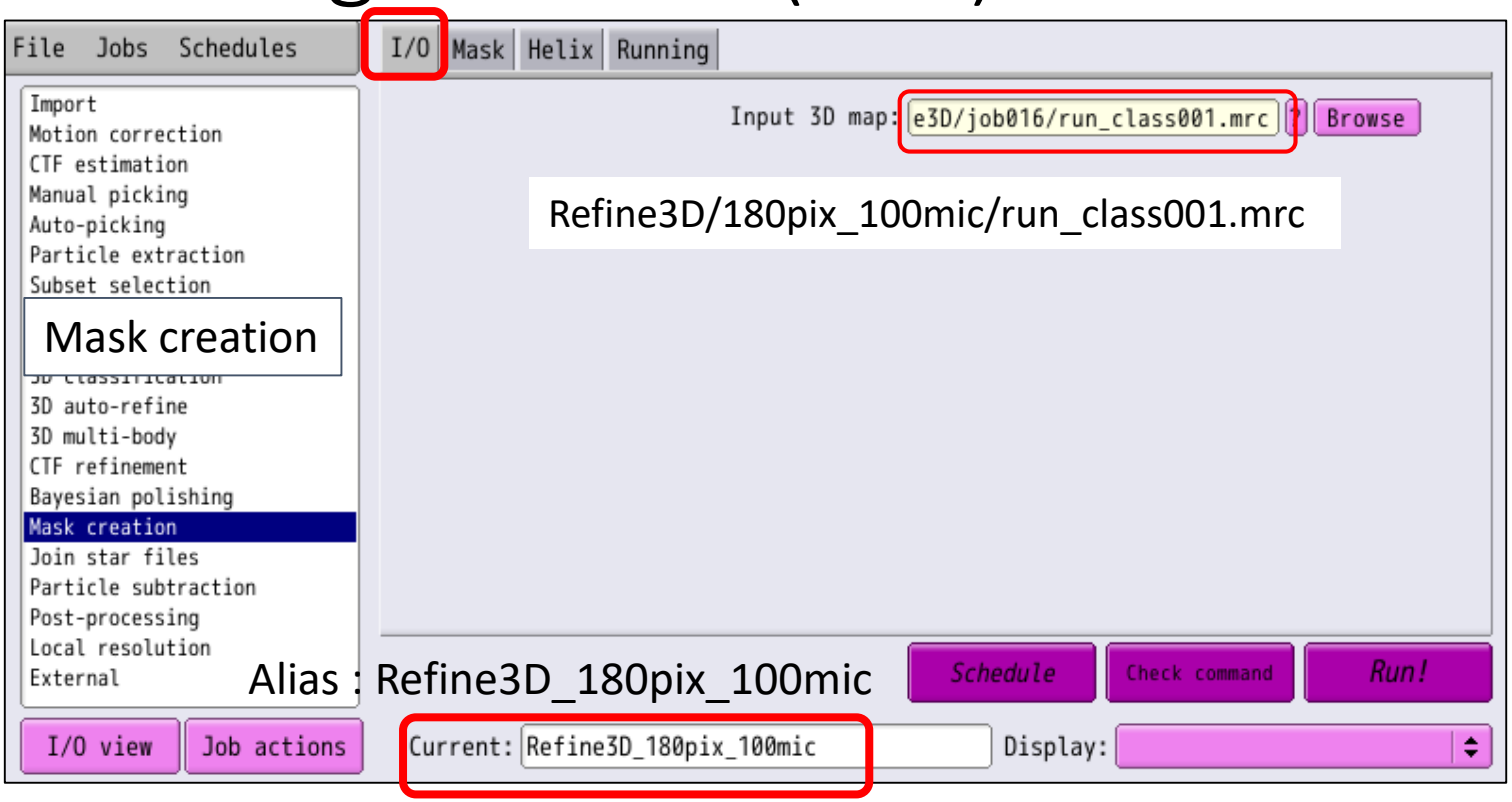


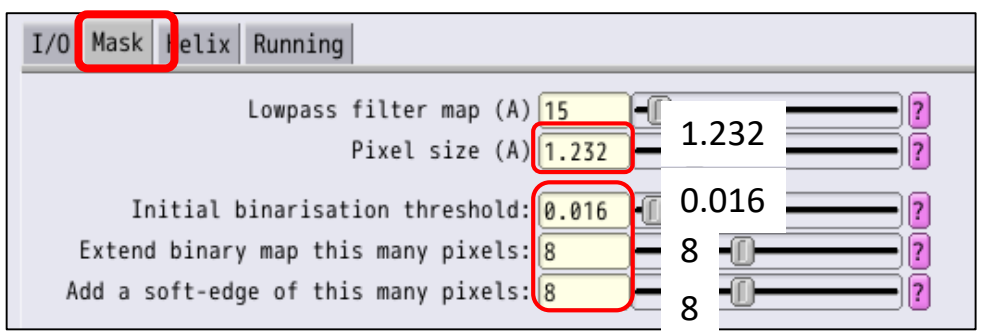
Level = 0.025



Level = 0.035

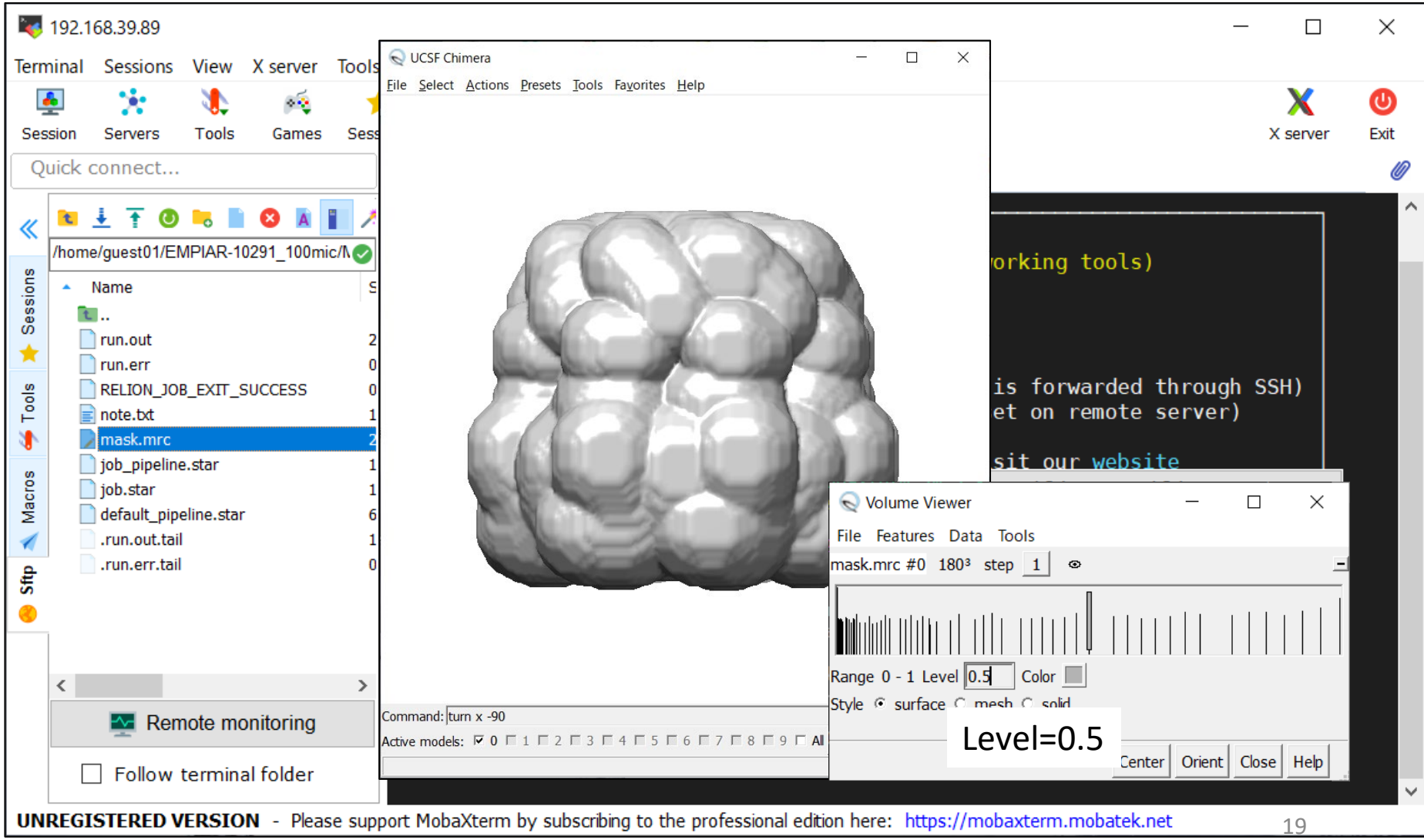
6.1 Making a Mask （1分）





6.1 Making a Mask : Checking mask.mrc

MobaXtermのファイルブラウザで、MaskCreate/Refine3D_180pix_100mic/mask.mrcを選び、右クリック・メニューから、[Open with...]でUCSF Chimeraで開いて確認する。

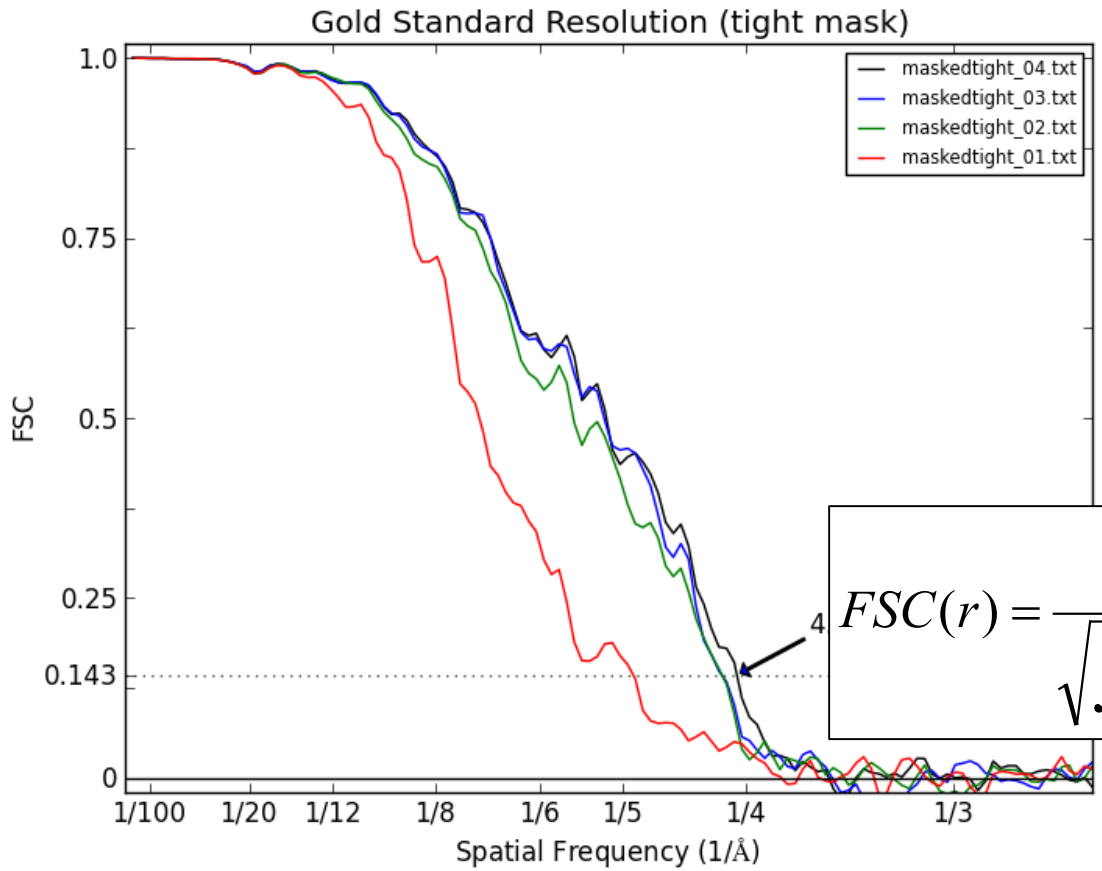


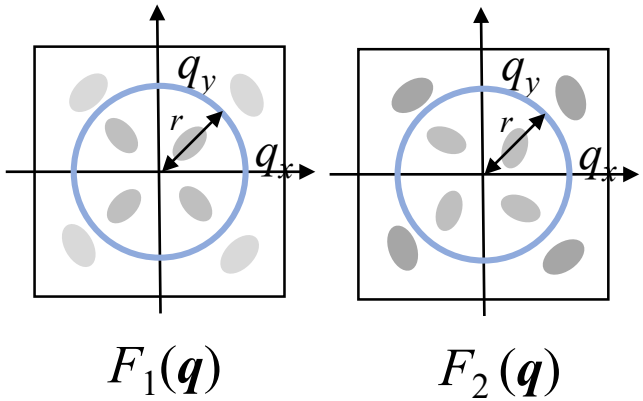
Fourier Shell Correlation (FSC)による解像度

1)2D粒子画像群を二つに分け、それぞれ独立に3Dマップ$f_1(\boldsymbol{x}), f_2(\boldsymbol{x})$を作成する。それをフーリエ変換し$F_1(\boldsymbol{q}), F_2(\boldsymbol{q})$とする。

2)ある周波数 $r=|\boldsymbol{q}|$ の球殻(spherical shell)上で、の相関係数$FSC(r)$を計算する。

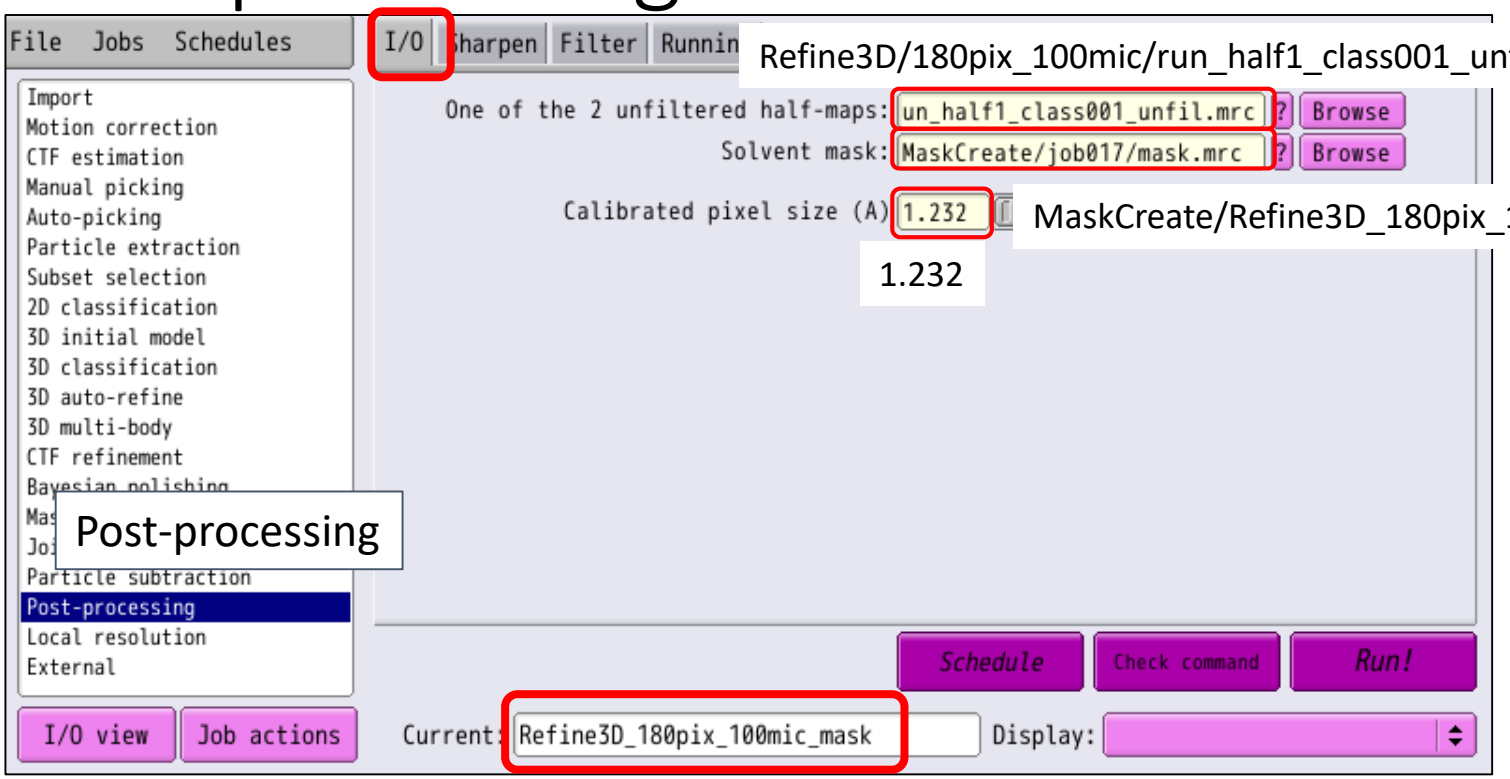
3)相関係数$FSC(r)$を横軸を空間周波数に対して図示。FSC=0.5あるいは0.143(=1/7)のときの周波数の値を解像度とする。



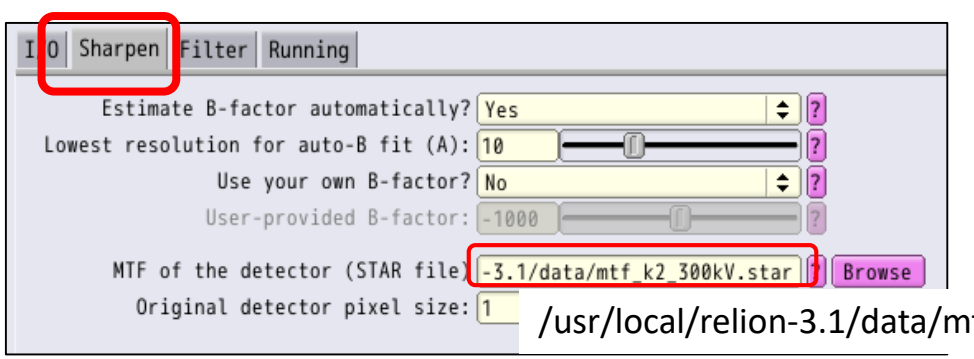


$$FSC(r) = \frac{\int_{|\boldsymbol{q}|=r} \mathrm{Re}\left[F_1(\boldsymbol{q}) \cdot F_2(\boldsymbol{q})^*\right] d\boldsymbol{q}}{\sqrt{\int_{|\boldsymbol{q}|=r} |F_1(\boldsymbol{q})|^2 \, d\boldsymbol{q} \cdot \int_{|\boldsymbol{q}|=r} |F_2(\boldsymbol{q})|^2 \, d\boldsymbol{q}}}$$

6.2 Postprocessing



Alias:Refine3D_180pix_100mic_mask



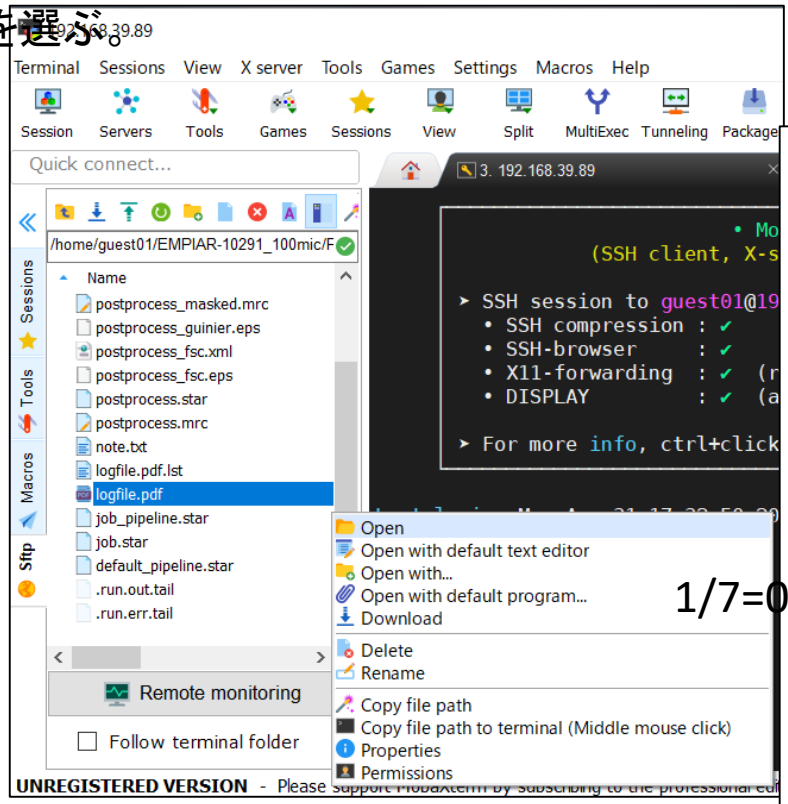
/usr/local/relion-3.1/data/mtf_k2_300kV.star

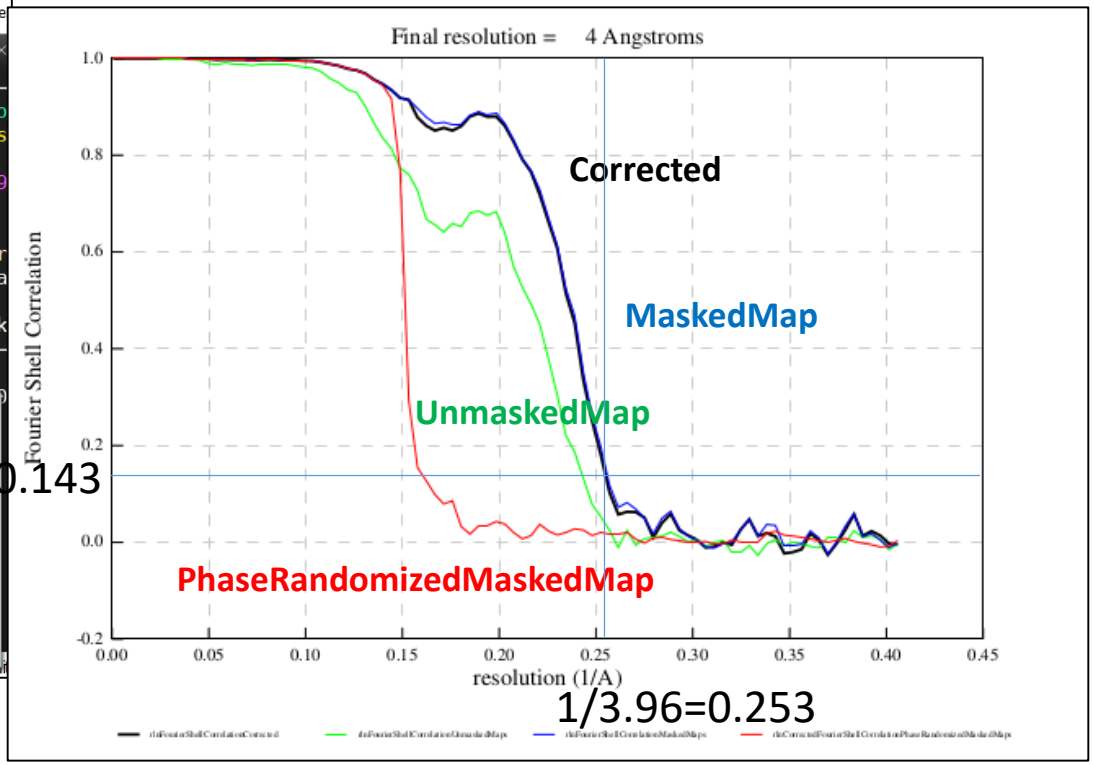
6.2 結果の確認

MobaXtermのファイルブラウザで、
PostProcess/Refine3D_180pix_100mic_mask/logfile.pdf
を選ぶ。

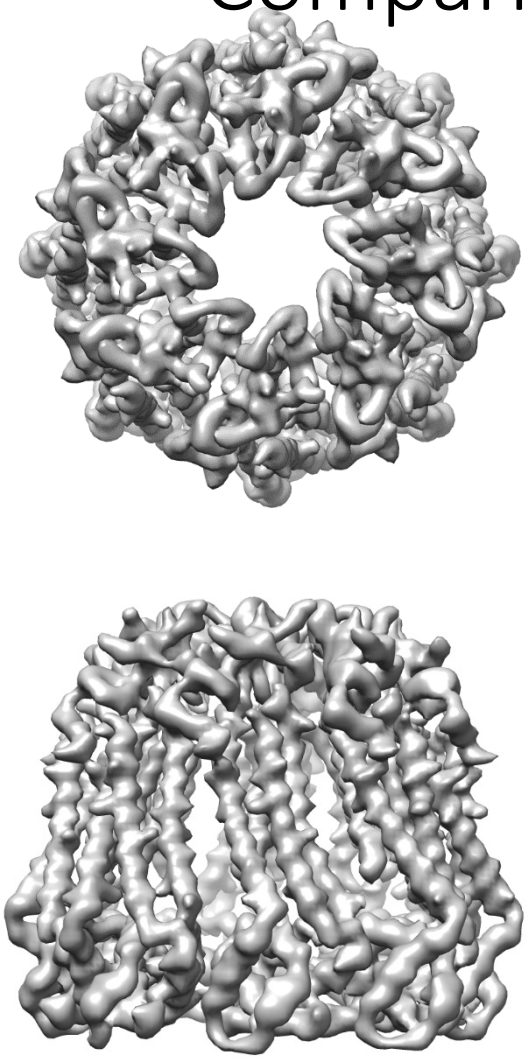
```
  + intercept of fit:          -11.6428
  + correlation of fit:        0.955742
  + apply b-factor of:         -198.981
== Writing output files ...
  + Processed map:             PostProcess/job018/postprocess.mrc
  + Processed masked map:      PostProcess/job018/postprocess_masked.mrc
  + Metadata file:             PostProcess/job018/postprocess.star
  + FINAL RESOLUTION:          3.96
```

+ FINAL RESOLUTION: 3.96

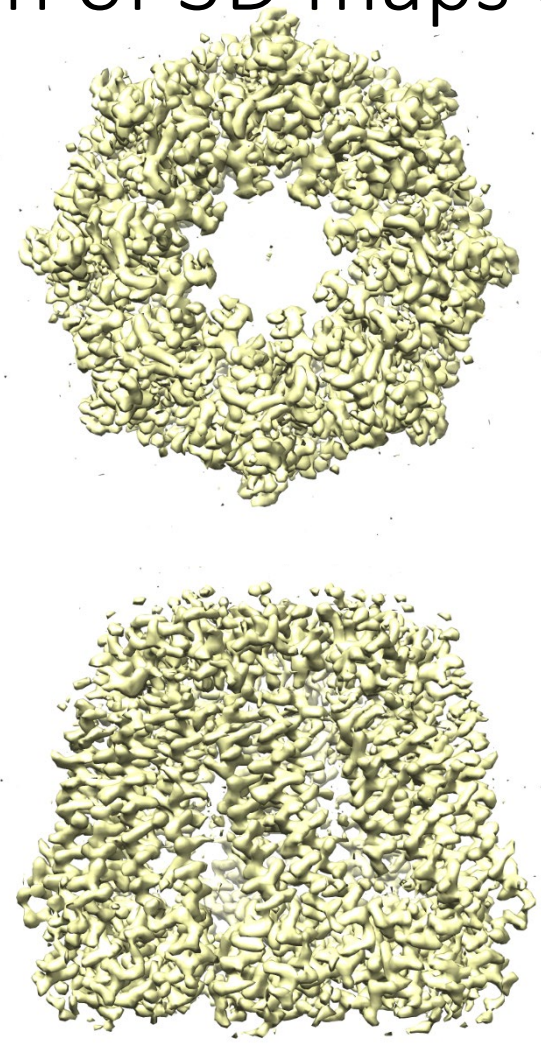




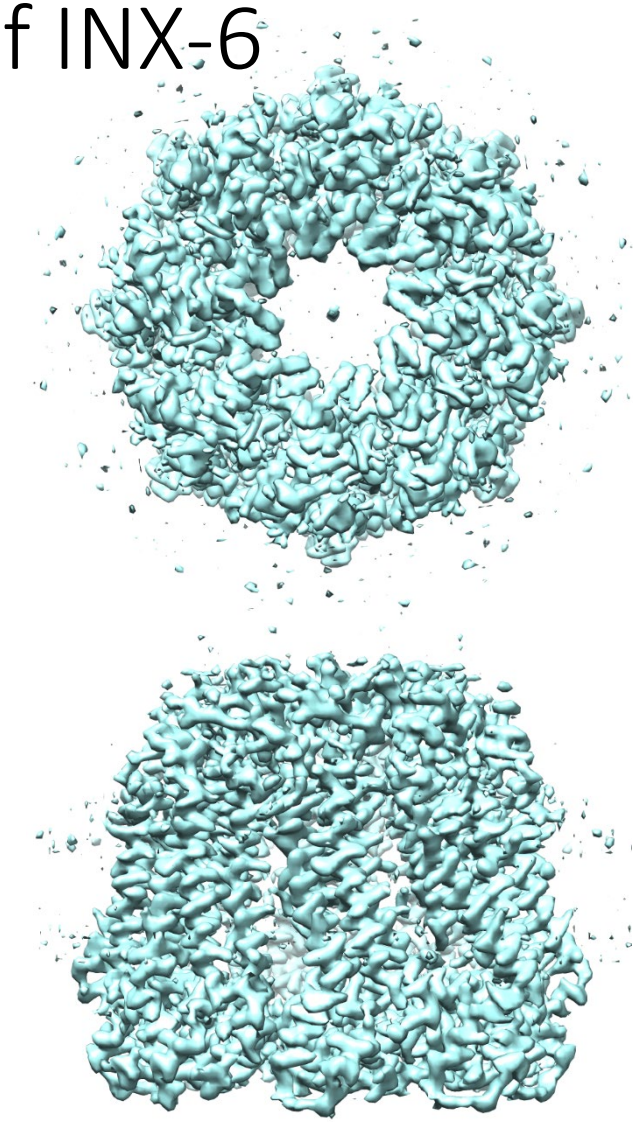
Comparison of 3D maps of INX-6



Refine3D
Resolution: 4.11 Å
from 100 micrographs



Postprocess with mask
Resolution: 3.96 Å
from 100 micrographs



EMD-9973. 180^3 voxels
Resolution: 3.6 Å
From 300 micrographs