

電子顕微鏡解析

川端 猛

大阪大学 蛋白質研究所

寄附研究部門准教授

計算生命科学の基礎 IV

2017年10月25日(水) 17:00-18:30

今日の話題

- 最近の電子顕微鏡をめぐる状況
- 電子顕微鏡の仕組み
- 単粒子解析
- 単粒子解析の画像処理の例
- 電顕の3Dマップからの原子モデルの構築法

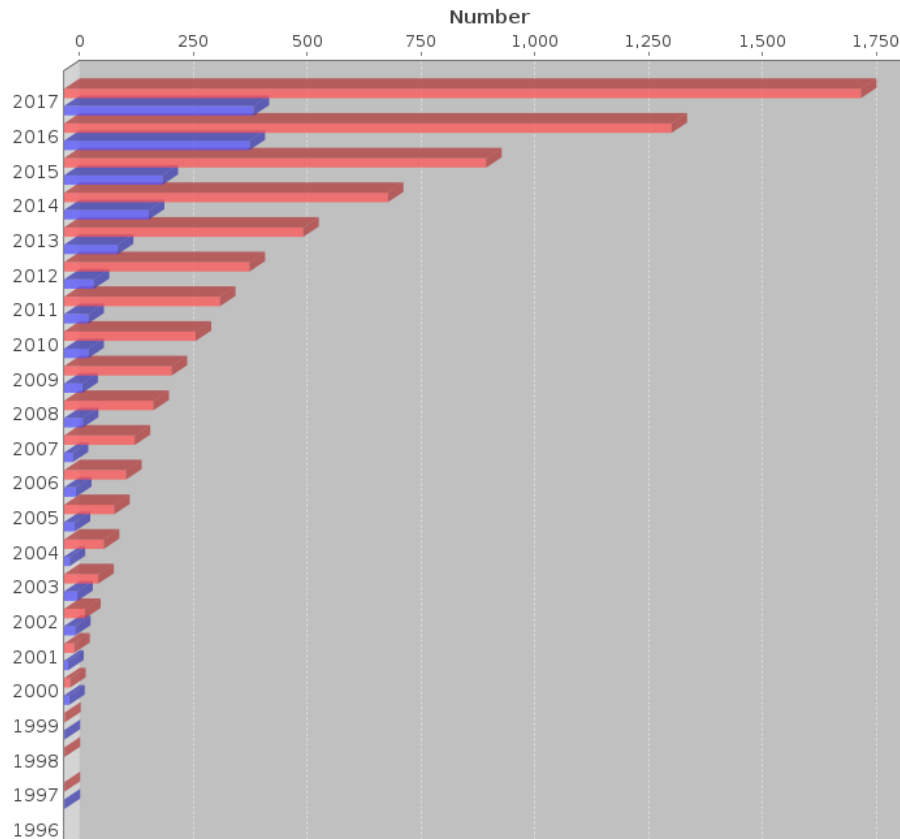
最近の電子顕微鏡をめぐる状況

PDBデータの実験手法の内訳

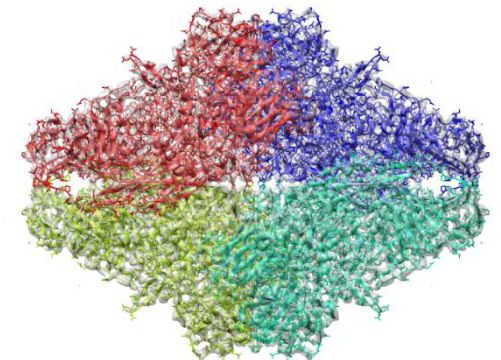
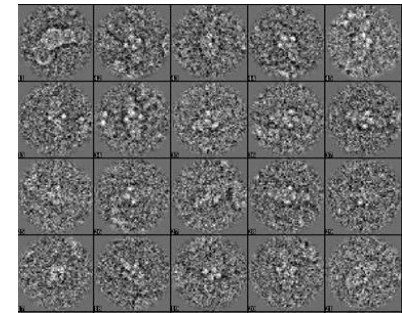
実験手法	X-ray	NMR	Electron Microscopy	Hybrid	Others	Total
エントリ数	120180	11998	1741	110	222	134251
比(%)	89.5%	8.9%	1.3%	0.08%	0.17%	100%

Yearly Growth of Structures Solved By Electron Microscopy

number of structures can be viewed by hovering mouse over the bar



RCSB-PDB, 2017/10/18



2017年のノーベル化学賞

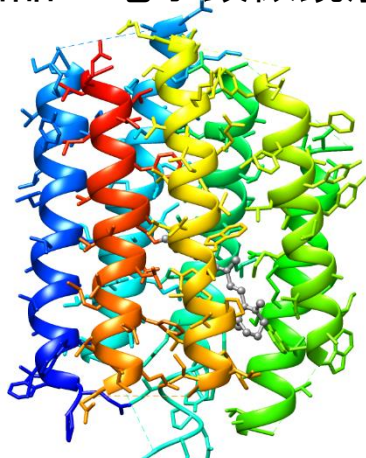
Developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution

溶液中の生体分子の高解像度の構造決定のための低温電子顕微鏡の技術の開発

Richard Henderson

(1945-, MRC Laboratory of Molecular Biology)

バクテリアの紫膜を用いた
バクテリオロドプシンの二次元
結晶の電子顕微鏡解析



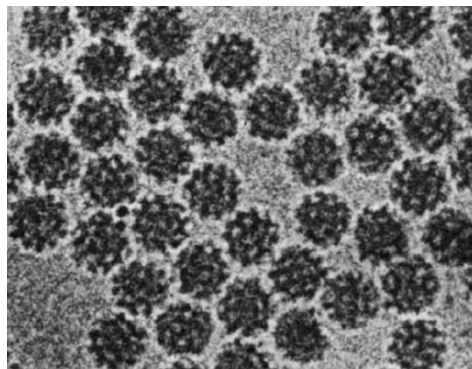
PDBcode:1brd (1990)

ELECTRON CRYSTALLOGRAPHY (3.5 Å)

Jacques Dubochet

(1942-, EMBL, Heidelberg)

クライオ電子顕微鏡の試料のための非晶質(アモルファス)氷の生成法を初めて確立

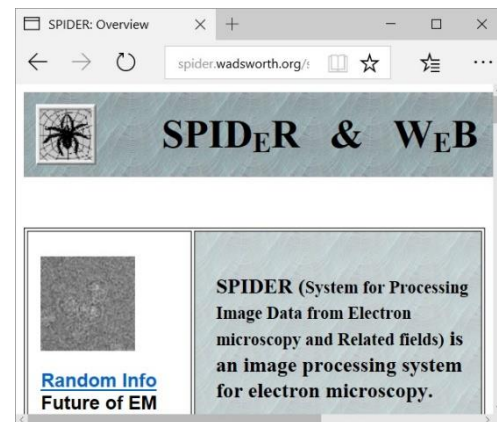


Nature 308, 32 - 36 (1984);

Joachim Frank

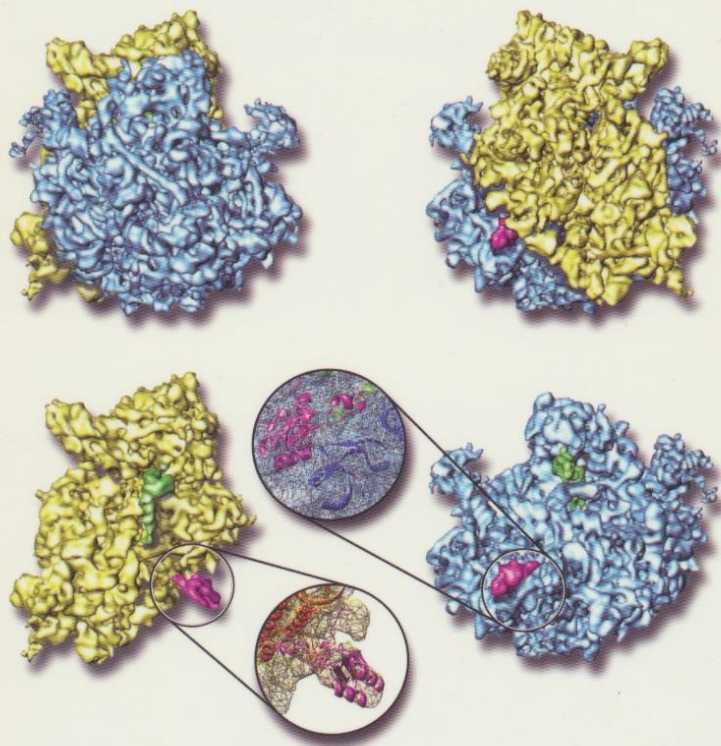
(1940-, Columbia university)

単粒子解析のための理論とプログラムの開発



Frank, J., Shimkin, B., Dowse, H.
SPIDER—A modular software system for electron image processing.
Ultramicroscopy, 6, 343-357, (1981)

Three-Dimensional Electron Microscopy of Macromolecular Assemblies



JOACHIM FRANK

to Takeshi :

I hope you will have
fun with single-particle
reconstruction !

Joachim Frank

Okazaki

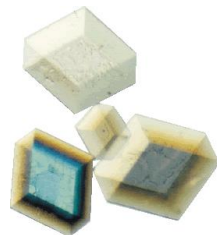
March 31 '08

Takeshi Kawabata

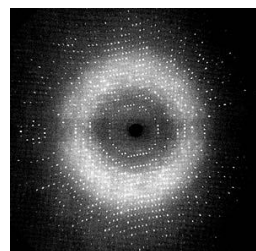
立体構造の決定法

X線結晶解析

大量発現
精製
結晶化

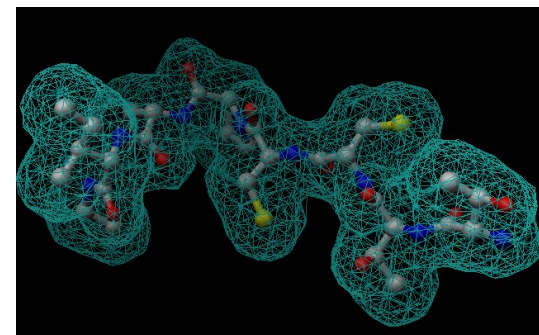


X線回折強度の測定



3次元
電子密度マップ

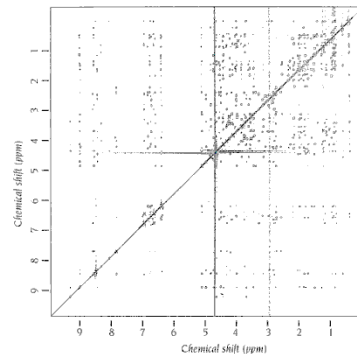
原子モデルの構築
原子モデルの精密化



NMR(核磁気共鳴法)

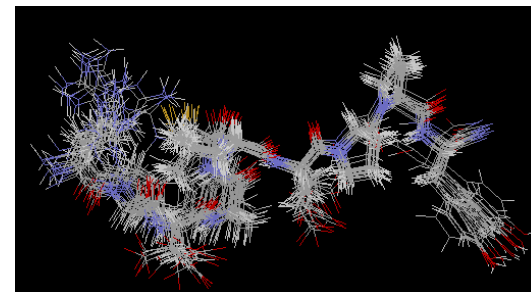
大量発現
精製

核磁気共鳴解析
(NOE解析, 帰属と距離拘束の抽出)



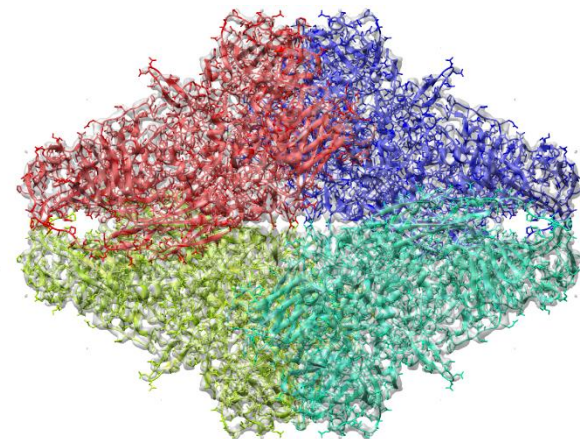
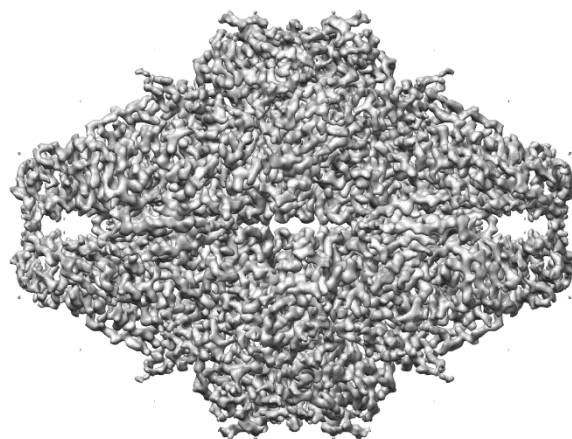
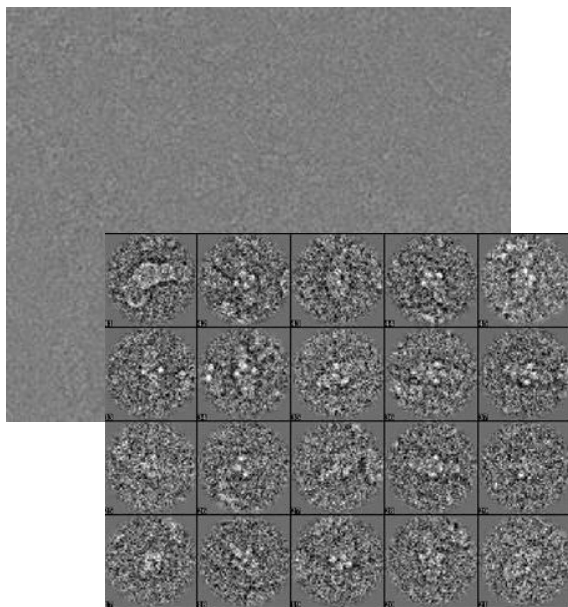
原子間
距離拘束

原子モデルの構築
原子モデルの精密化



- (1)多数分子の平均構造の観察(精製が重要。大きな分子、複合体はより難しくなる)
- (2)発現・精製・結晶化のステップで、各タンパク質によって実験条件の調整が必須
- (3)実験データの収集・原子モデル構築の段階で計算機の支援が不可欠

電子顕微鏡(単粒子解析) による立体構造データの取得



2D画像群

→
3次元再構成

3D画像

→
原子モデリング

原子モデル

EMPIAR-10013
402.2GB

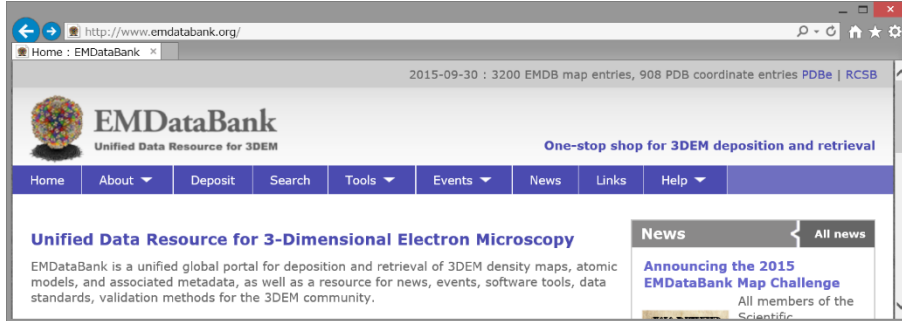
emd_5995.map
150 MB

pdb3j7h.ent
2.6 MB

EMDataBank

3DEMの密度マップを集めたデータベース

Project team

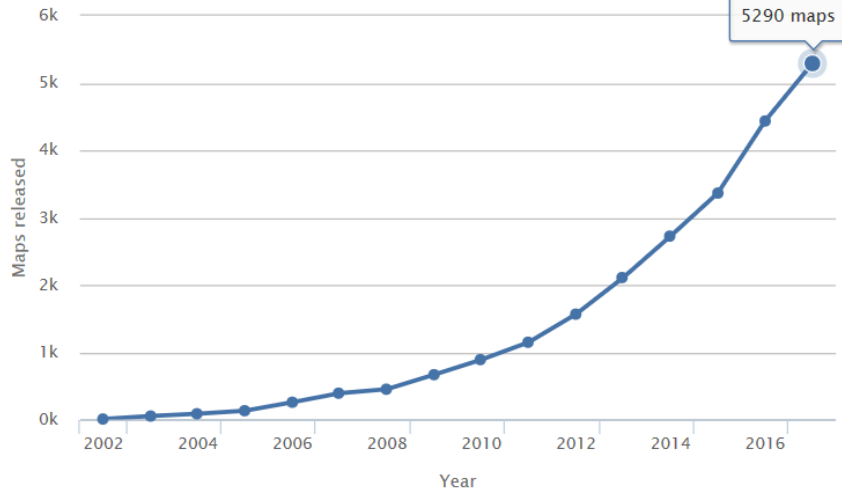


Gerard Kleywegt (PDBe), Helen Berman (RCSB-PDB) and Wah Chiu (Baylor)



Ingvar Lagerstedt (PDBe), Ardan Patwardhan (PDBe), Cathy Lawson (RCSB PDB), Eduardo Sanz Garcia (PDBe)

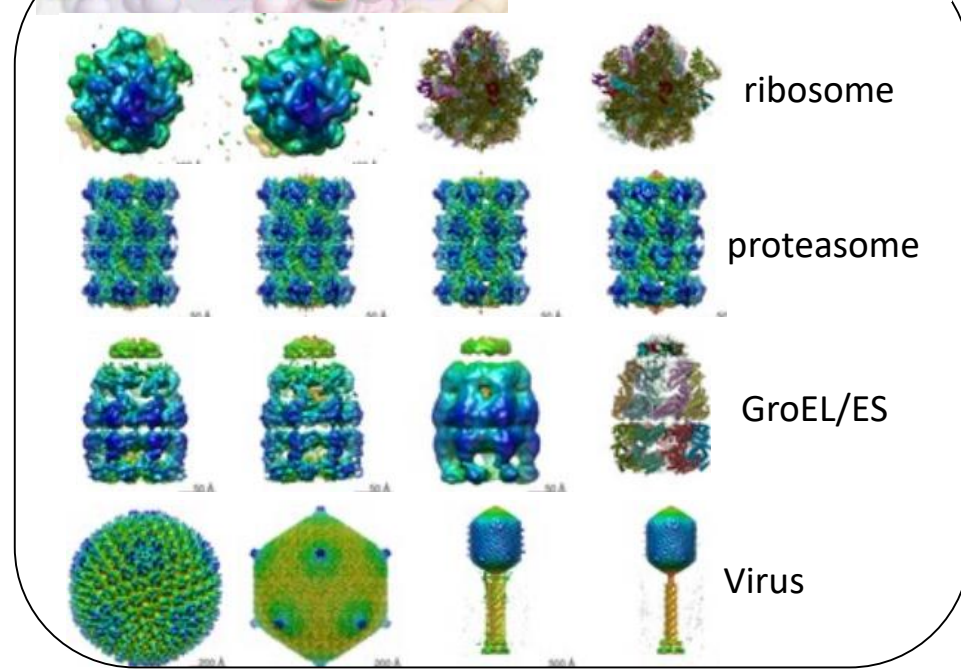
Cumulative number of maps released



5290個の EMDB のエントリー (2017/10/25) :



EMDBのデータをわかりやすくまとめたサーバ(PDBj, 鈴木)



EMPIAR (Electron Microscopy Public Image Archive)

<https://www.ebi.ac.uk/pdbe/emdb/empiar/>

EMDBに登録された3Dマップの元となった2D画像群を集めたデータベース



Electron Microscopy Public Image Archive

EMPIAR home Deposition REST API FAQ About EMPIAR

EMPIAR, the Electron Microscopy Public Image Archive, allows you to upload, and download and reprocess the thousands of images deposited in the archive.

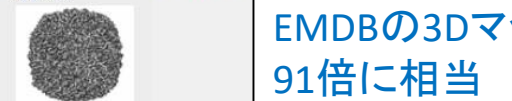
Deposit your data in EMPIAR to share it with the scientific community.

Browse and download EMPIAR datasets using the table below.

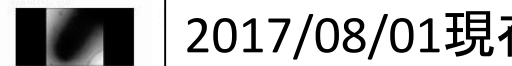
Show 50 entries

Dataset Title

EMPIAR-10122 Mixed under-/over-focused data collected by VPP-Cs-corrector coupled EM.



EMPIAR-10115 Tilt-series of e. coli carrying the ple7 p... YFP-MreB hyper-overexpressed by ind...



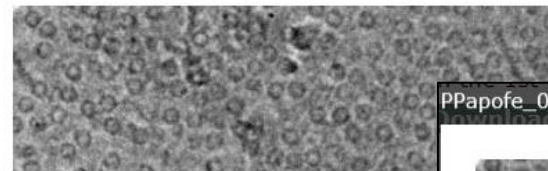
EMPIAR-10114

EMPIAR-10122

Mixed under-/over-focused data collected by VPP-Cs-corrector coupled EM.

Publication:

PPapofe_0003.mrc



Related EMD entry:

Deposited:

Released:

Last modified:

Category:

Image format:

No. of images or tilt series:

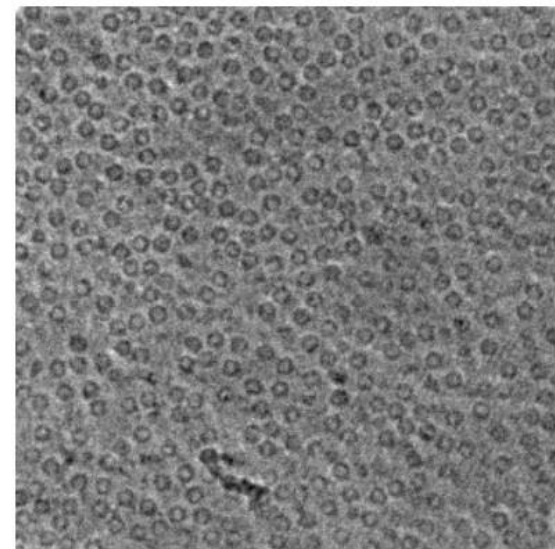
PDBの原子モデルの総量(33GB)の141400倍に相当
EMDBの3Dマップの総量(516 GB)の91倍に相当

2017/08/01現在
データエントリー数:87

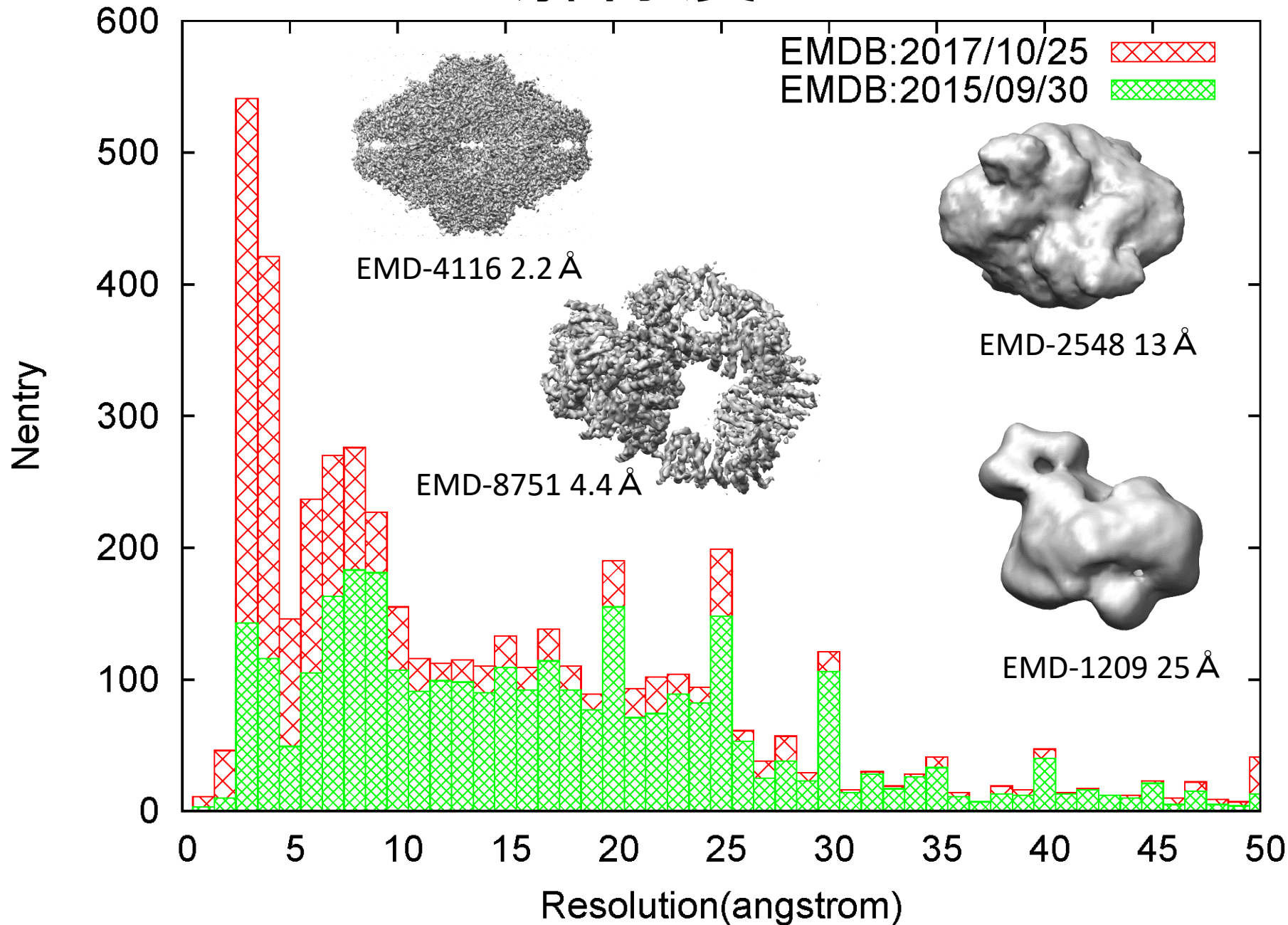
全データサイズ: **47.2 TB**

最大のエントリー: EMPIAR-10061 : 12.4 TB

PPapofe_0004.mrc



EMDBの解像度



高解像度の電顕密度マップの出現

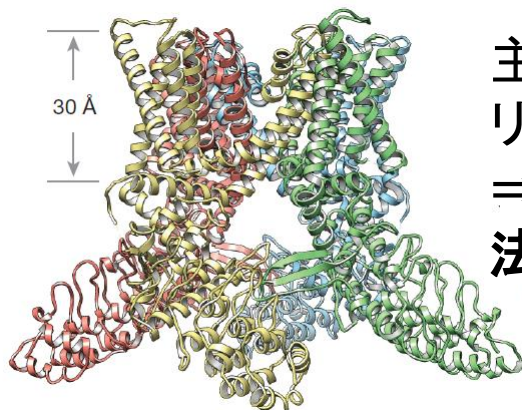
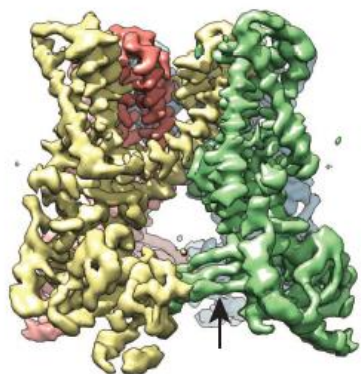
The revolution will not be crystallized. Nature 525, 172–174

2013年にUCSFのYifan Chengのグループが、**3.4 Å**の解像度の膜タンパク質のマップを発表

Liao et al. (2013) Nature, **504**,107.

TRPV1 (transient receptor potential channel)

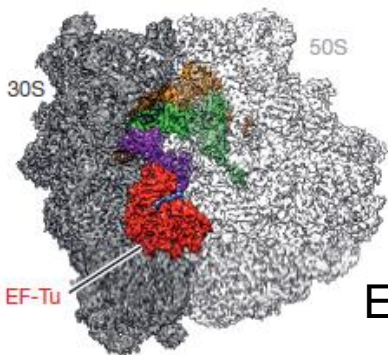
Cao et al. (2013) Nature, **504**,113.



主鎖構造だけでなく、側鎖の形状も低分子リガンド(カプサイシン)も見える?

⇒ X線結晶解析の密度マップと同様な手法でデノボの原子モデルが作成できる。

EMD-5776, 5778, 5779, PDB-3j5q,3j5r, 3j5p

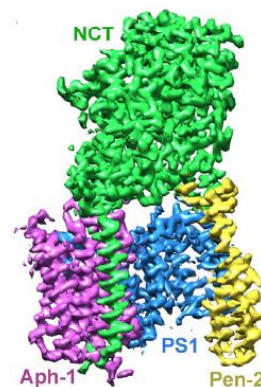


E.Coli ribosome
EF-TU complex

2.9 Å

EMD-2847, PDB-5afl

Fischer et al. (2015) Nature, **520**,567.



Human γ -secretase

3.4 Å

EMD-3061, PDB-5a63

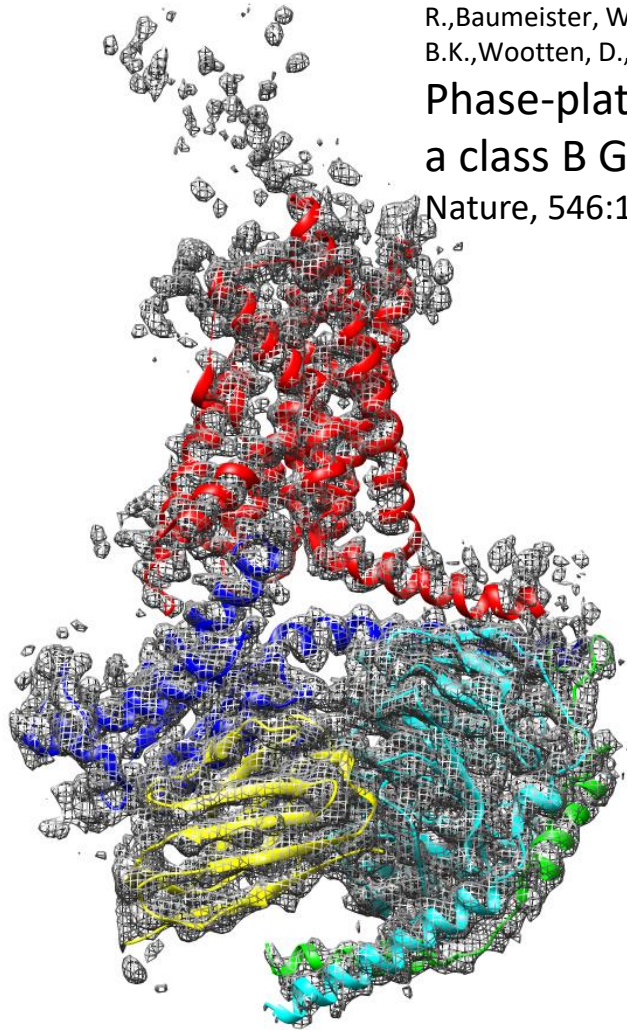
Bai et al. (2015) Nature, **525**,212.

最近報告された構造

Liang, Y.L., Khoshouei, M., Radjainia, M., Zhang, Y., Glukhova, A., Tarrasch, J., Thal, D.M., Furness, S.G.B., Christopoulos, G., Coudrat, T., Danev, R., Baumeister, W., Miller, L.J., Christopoulos, A., Kobilka, B.K., Wootten, D., Skiniotis, G., Sexton, P.M.

Phase-plate cryo-EM structure of a class B GPCR-G-protein complex.

Nature, 546:118-123, 2017



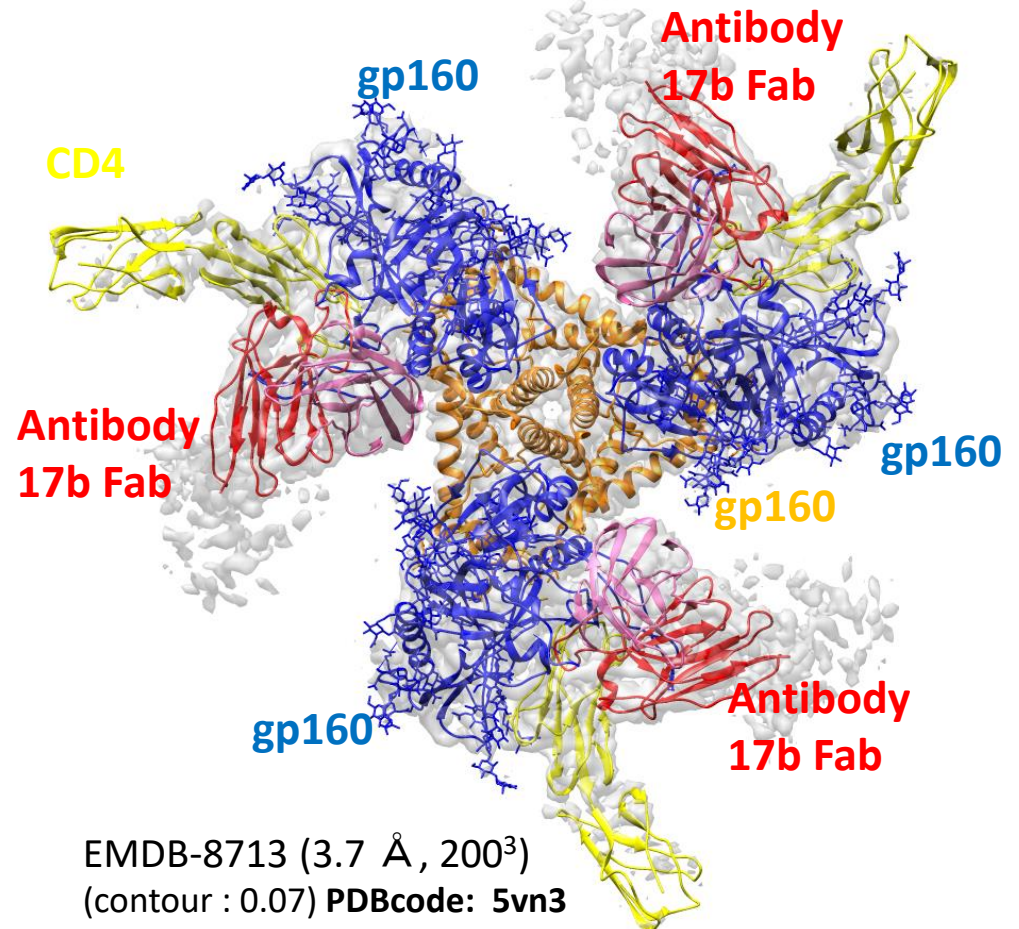
EMDB-8623 (4.1 Å, 200³)

(contour Level: 0.05) PDBcode: **5uz7**

Ozorowski, G., Pallesen, J., de Val, N., Lyumkis, D., Cottrell, C.A., Torres, J.L., Copps, J., Stanfield, R.L., Cupo, A., Pugach, P., Moore, J.P., Wilson, I.A., Ward, A.B.

Open and closed structures reveal allosteric and pliability in the HIV-1 envelope spike.

Nature, 547:360-363, 2017



EMDB-8713 (3.7 Å, 200³)

(contour : 0.07) PDBcode: **5vn3**

Map ChallengeとModel Challenge

http://challenges.emdatabank.org/

EMDataBank Validation Challenges

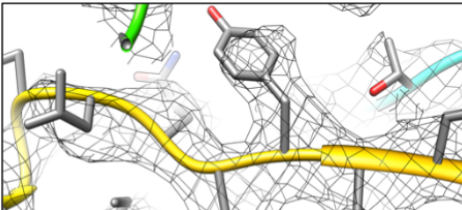
ABOUT MAP CHALLENGE MODEL CHALLENGE NEWS

News

Announcing the 2nd EMDataBank Model Challenge
published Wed, 10/14/2015 - 09:06

EMDataBank is pleased to announce the 2015/2016 **Model Challenge**.

All members of the Scientific Community--at all levels of experience--are invited to participate as



Model Challenge
Registration for the 2015/2016 Model Challenge is now open. Please visit the [model challenge page](#) for more details.

Map Challenge
Registration and Submissions are open for the 2015 Map Challenge. Please visit the [map challenge page](#) for more details.

WANTED

Challengers
Assessors

create/assess moderate-to-high resolution 3DEM reconstructions

2015 EMDataBank MAP CHALLENGE
challenges.emdatabank.org

Map Challenge : 2次元画像群から、3次元密度マップを作成するコンテスト









Model Challenge : 3次元密度マップから原子モデルを作成するコンテスト

※どちらも解像度がよい(<5 Å)マップがターゲット

Model Challenge Targets

The eight targets chosen by the model challenge committee are shown below. All are based on recently reported 3DEM structures. The target order is according to unique molecular weight, from 19 kDa (TMV) at left to 2500 kDa (ribosome) at right. For challenge rules and additional information about each target see the [model challenge guide](#).

Target Map Download: You can use [this rsync script](#) to download target maps from one of three wwPDB ftp sites. Alternately, you can download individual maps from EMDB atlas pages (click on EMDB entry link in the table below, select "download" tab).

	1. Tobacco Mosaic Virus	2. T20S Proteasome	3. GroEL	4. TRPV1 Channel	5. Brome Mosaic Virus	6. β-Galactosidase	7. γ-Secretase	8. 70S Ribosome
target								
Map "A" EMDB entry Primary Citation Reported Resolution (Å)	EMD-2842 Fromm et al 3.3	EMD-5623 Li et al 4.1	EMD-6422 -- 4.1	EMD-5778 Liao et al 3.3	EMD-6000 Wang et al 3.8	EMD-5995 Bartesaghi et al 3.2	EMD-2677 Lu et al 4.5	EMD-2847 Fischer et al 2.9
Map "B" EMDB entry Primary Citation Reported Resolution (Å)		EMD-6287 Campbell et al 2.8				EMD-2984 Bartesaghi et al 2.2	EMD-3061 Bai et al 3.4	EMD-6316 Li et al 3.6
Imposed Map Symmetry	Helical	Dihedral (D7)	Dihedral (D7)	Cyclic (C4)	Icosahedral (I)	Dihedral (D2)	None (C1)	None (C1)
Sample MW (MDa)	--	0.7	0.8	0.3	4.6	0.47	0.17	2.5
Unique MW (kDa)	19	50	56	80	80	120	170	2500

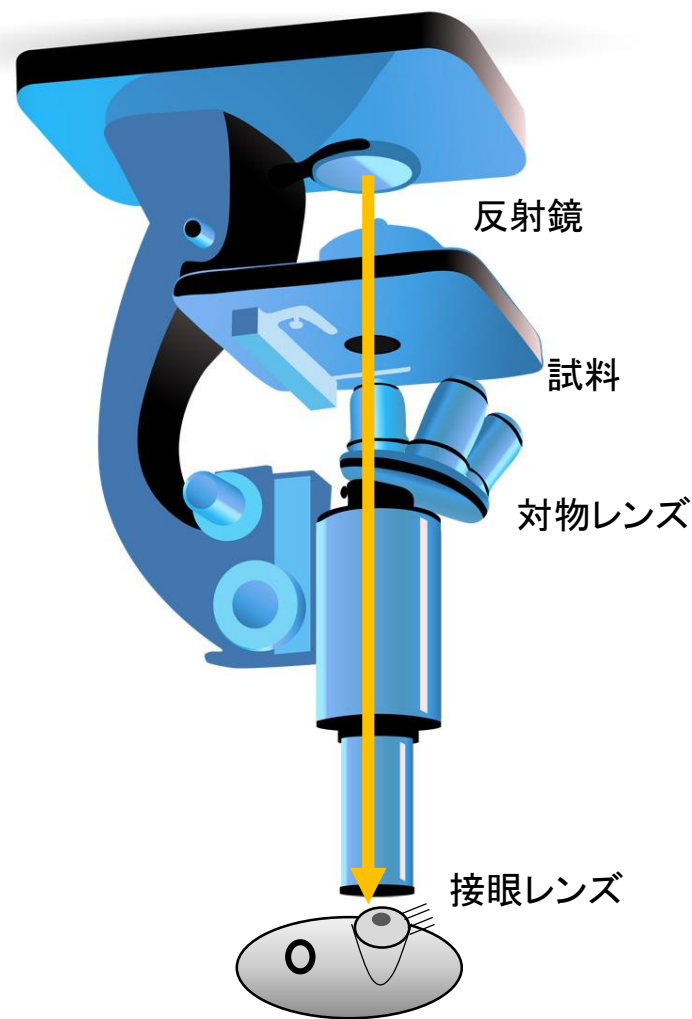
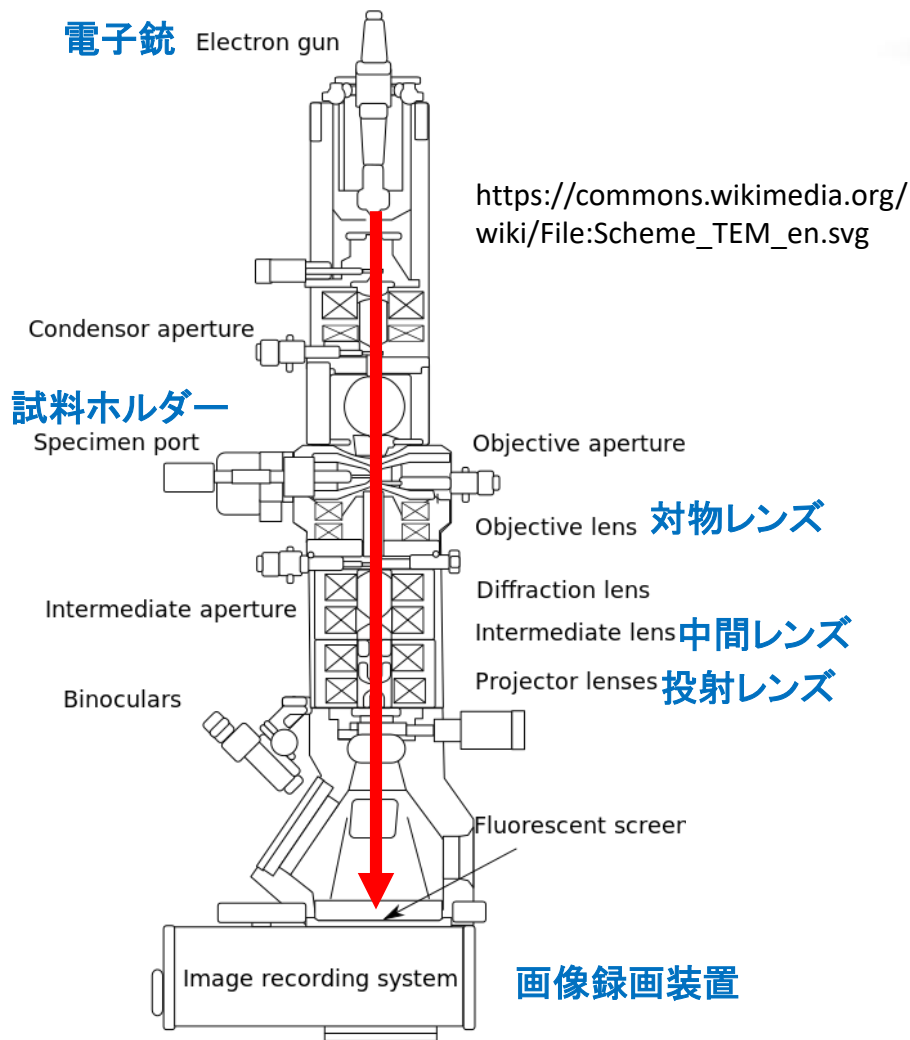
Deadline for submitting model is June 17, 2016 (21:00 UTC) => June 18 (6:00 JPN)

低温電子顕微鏡の仕組み

電子顕微鏡と光学顕微鏡

電子顕微鏡(透過型)

光学顕微鏡



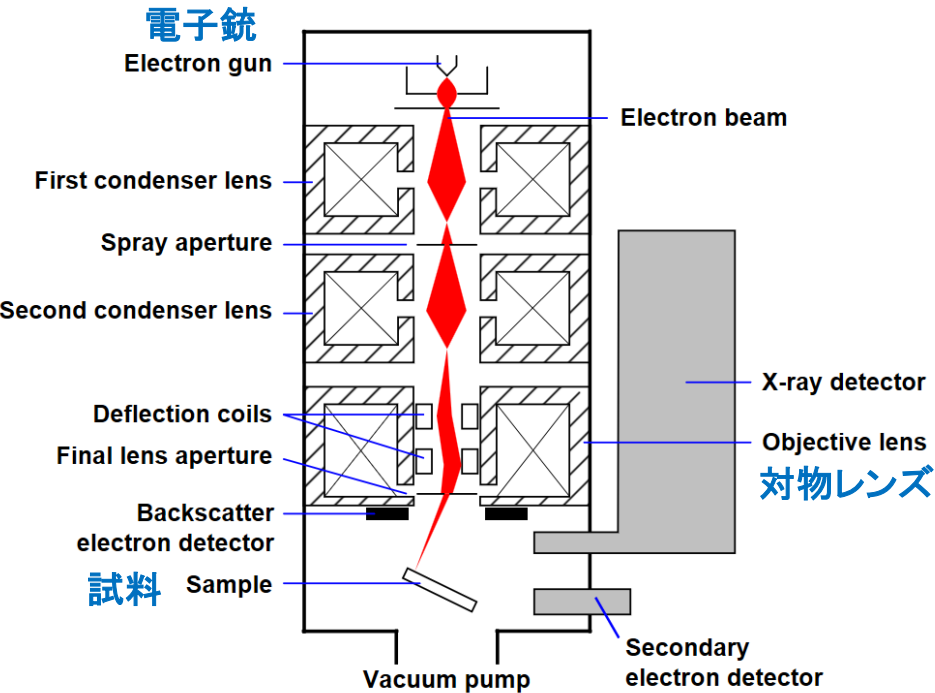
電子線を使用。経路は真空。磁界レンズを使用。蛍光スクリーンと銀塩写真フィルムか直接電子検出器で像を検出

可視光線を使用。経路は空气中。光学レンズを使用。肉眼やCCDカメラで像を観察

走査型(SEM)と透過型(TEM)顕微鏡

走査型電子顕微鏡

Scanning electron microscopy

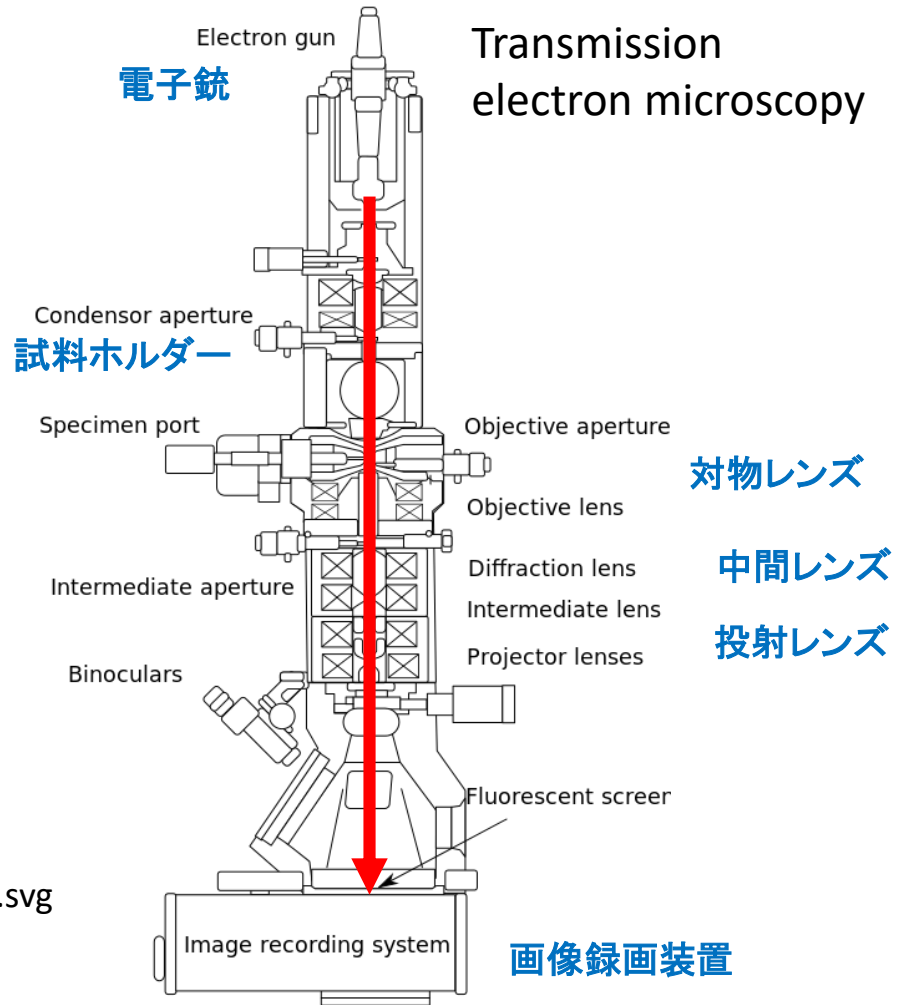


[https://commons.wikimedia.org/wiki/File:Schema_MEB_\(en\).svg](https://commons.wikimedia.org/wiki/File:Schema_MEB_(en).svg)

細く絞った電子線を試料に走査(scan)する。
試料からの二次電子の放出量を検出器で計測。
試料の表面を観察する。

透過型電子顕微鏡

Transmission electron microscopy



https://commons.wikimedia.org/wiki/File:Scheme_TEM_en.svg

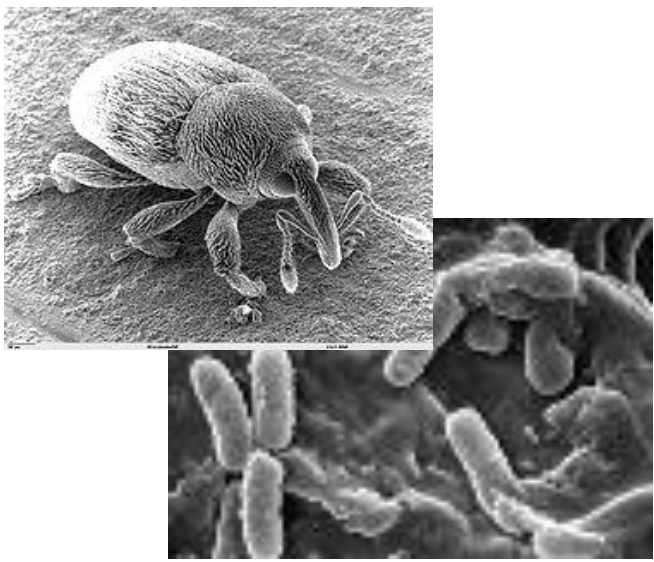
電子線を試料に照射し、透過した電子線の強弱などを像として検出。

走査型(SEM)と透過型(TEM)顕微鏡

走査型(SEM)



日本電子(株)社製
JSM-7900F 走査電子顕微鏡



透過型(TEM)



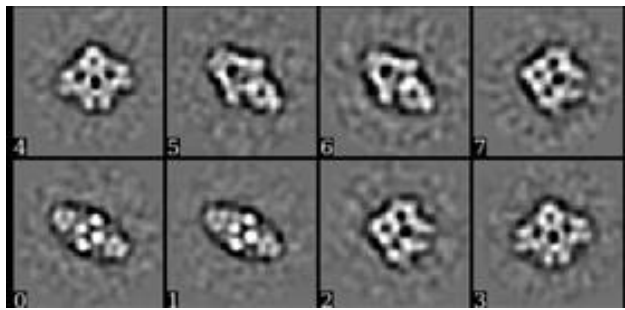
日本電子(株)社製
JEM-2200FS



[https://commons.wikimedia.org/wiki/File:A_\(H5N1\)_virion,_a_type_of_bird_flu_virus_which_is_a_subtype_of_avian_influenza_A.jpg](https://commons.wikimedia.org/wiki/File:A_(H5N1)_virion,_a_type_of_bird_flu_virus_which_is_a_subtype_of_avian_influenza_A.jpg)



Thermo Fischer (FEI)社製
Titan Krios G2



Beta-galactosidase (EMPIAR-10012)

クライオ電子顕微鏡の試料

電顕は試料を真空にさらすため、水分を含んだ生体試料はそのままでは計測できない。

ネガティブ染色法 (negative staining):

酢酸ウラニル溶液を試料に加え、生体分子に重原子を付着させる方法

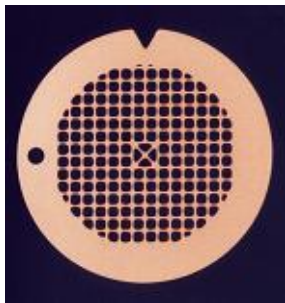
アモルファス氷で包埋する方法: 水を含んだ試料を急速に冷却させ、生体分子がアモルファス状態の氷(非晶質氷)に埋まった状態にして観察する方法

※氷の状態の水ではなく、急冷したアモルファス状態の水であることが重要。

氷は規則正しい結晶状態であるため、強く結晶回折して、生体分子を見えなくしてしまう。

クライオ電子顕微鏡 (Cryo Electron Microscopy)

アモルファス氷で包埋した試料を、冷却した電子顕微鏡で観察する方法のこと



グリッド

直径3-5mmほどの銅板
多数の小さな穴
微小な穴のあいた
カーボン膜を貼って使用

グリッドの上に、生体分子試料をかけて、
液体エタン+液体ヘリウムで急速冷却して
試料を作成する。

生体分子

アモルファス氷

カーボン膜

2013年以降の電顕の 高解像度化の要因

1. 直接電子検出器(Direct Electron Detector)の導入： これまでは電子を光に変えてCCDカメラで検出していた。DEDは電子を直接CMOSセンサーで検出する。
2. 動画撮影の導入： 毎秒数百枚のフレームを動画で撮影。動画から、電子線照射による試料のぶれを補正した静止画像を作成する。
3. 自動試料搬送装置(auto loader)などの自動データ取得機能： 無人状態で長時間動作させることで大量のデータの取得が可能となった。
4. ボルタ位相板(Volta phase plate)の導入： カーボン膜を電子線の経路上に置くことで、高いコントラストを実現
5. 高性能な計算資源・ソフトウェアの改良： 大容量のハードディスク、メモリの導入、クラスタマシンやGPUによる高い計算力。ベイズ推定など見通しのよい理論に基づいたソフトウェアの開発。

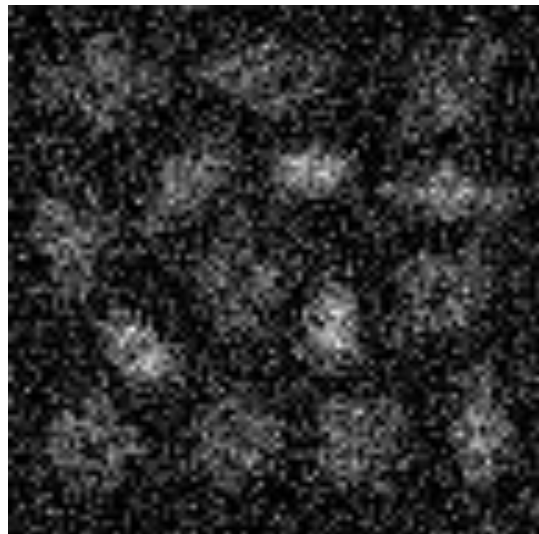
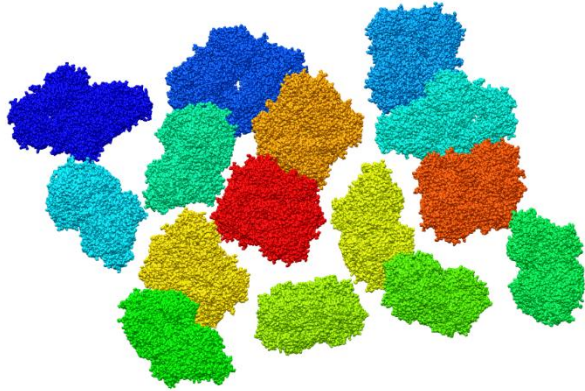
電顕による構造解析の特徴

- 結晶化は不要。しかし試料の濃度などの調整は必須。
- 膜タンパク質でも適用可(detergent, nanodisc)
- 比較的大きな分子に向いている(リボゾーム、GroEL)
- 最近ではGPCR, ヘモグロビンほどの大きさでも高解像度で観察可能（ミオグロビンは不可らしい...）
- 原則的に、標本中の分子が全て同じ形状をしていることを仮定
- 近年のソフト(Relion)では、一つの標本から数個の形状を提案することも可能になってきている。
- 高解像度といっても3-4 Åぐらい(?)
- 低分子化合物が見える場合もある。しかし、結合部位はわかるが、精密なポーズは難しい？
- 抗体の結合部位を調べるのに有効？

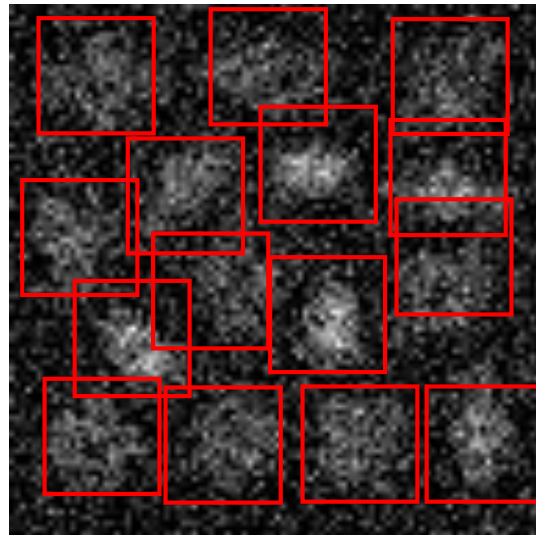
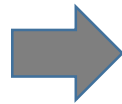
单粒子解析

単粒子解析(single particle analysis)

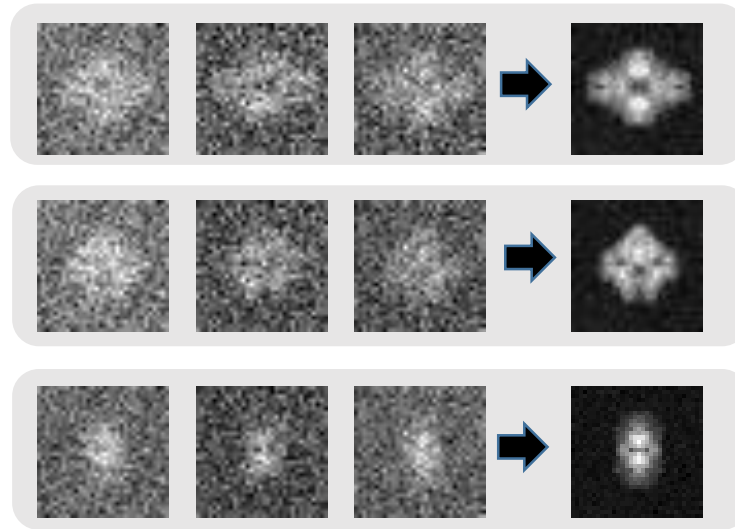
包埋された
様々な向きの生体分子



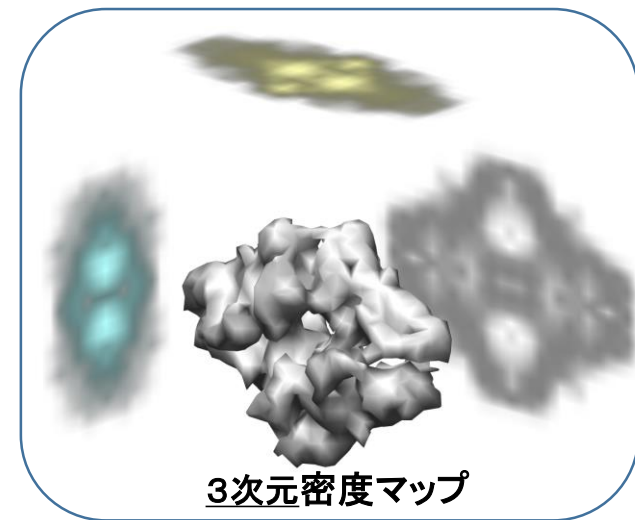
電顕像(micrograph)



分子像の切り出し(picking, boxing)



2D画像の分類
と平均化

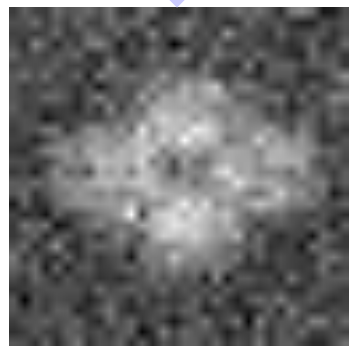


3次元密度マップ

3次元再構成

画像分類と平均化によるノイズ除去

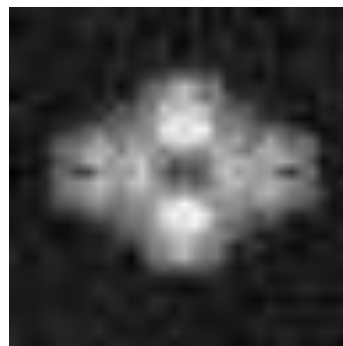
原子モデル(PDB:3j7h)を5 Åの解像度で格子幅7 Åで32x32の画像に投影。SNR=0.5のノイズ。



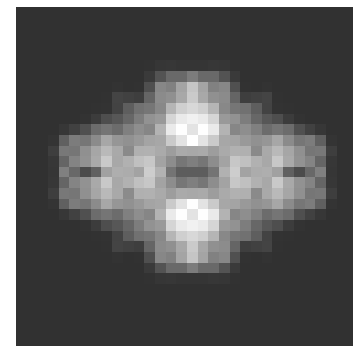
5枚の平均



10枚の平均



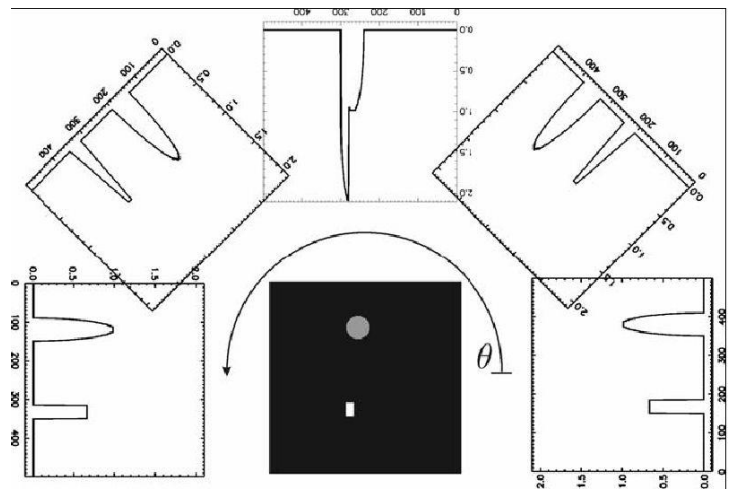
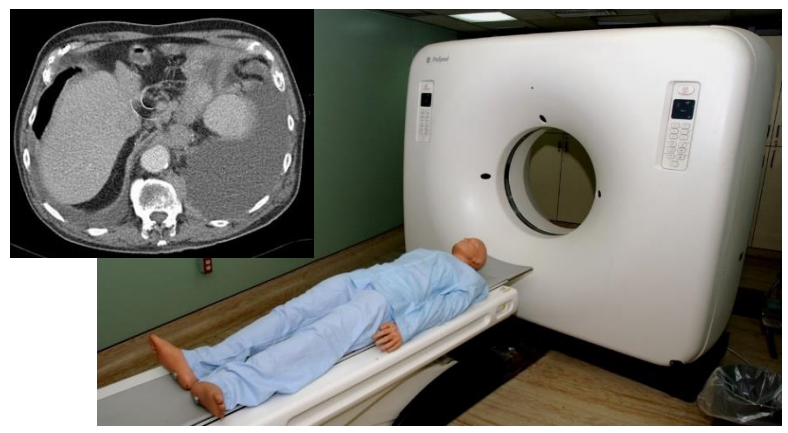
100枚の平均



原シグナル画像

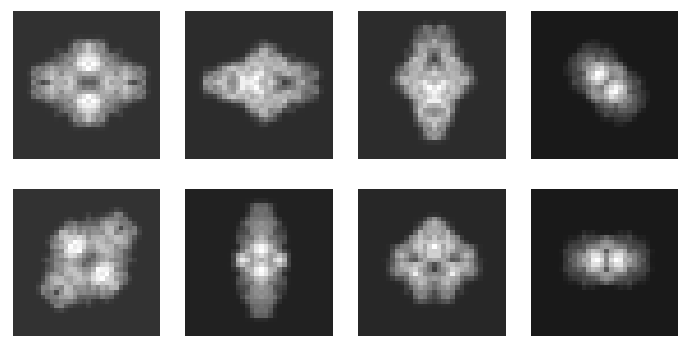
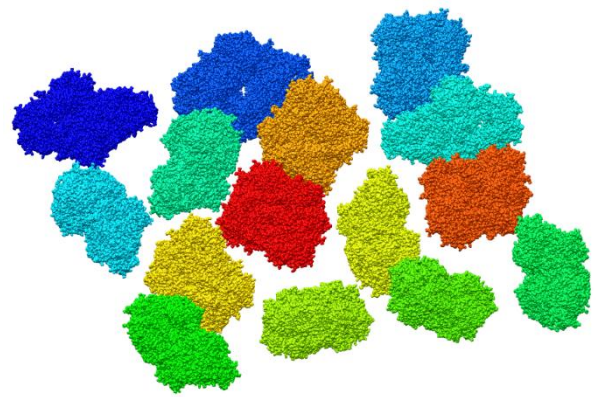
高次元再構成の問題

コンピュータ断層撮影



1D投影像から2D画像(断層像)を求める
カメラを被写体のまわりに回転させて撮影
投影方向は撮影時にわかっている

単粒子解析



2D投影像から3D画像を求める
多数の同一の分子が様々な方向で包埋
各粒子画像の投影方向を推定する必要

再構成のアルゴリズム

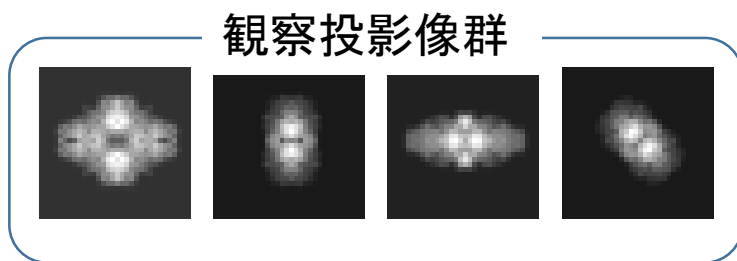
各投影画像の投影方向が既知の場合

(1)実空間で行うアルゴリズム: 逆投影法

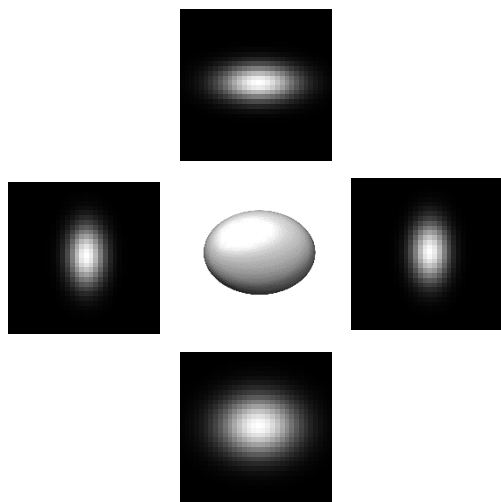
(2)周波数空間で行うアルゴリズム: フーリエ再構成法

投影画像の投影方向が未知の場合

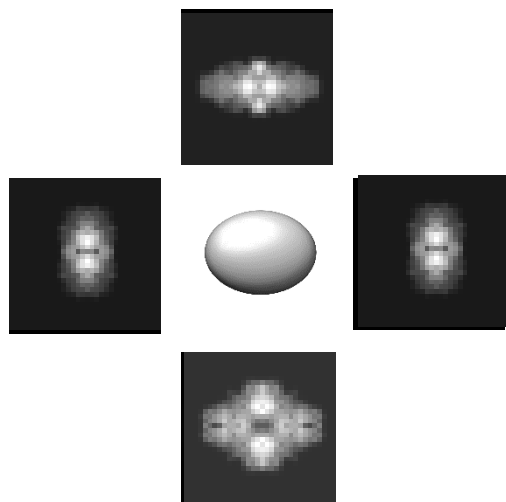
 初期3D画像を導入。以下の(1)-(3)を繰り返す



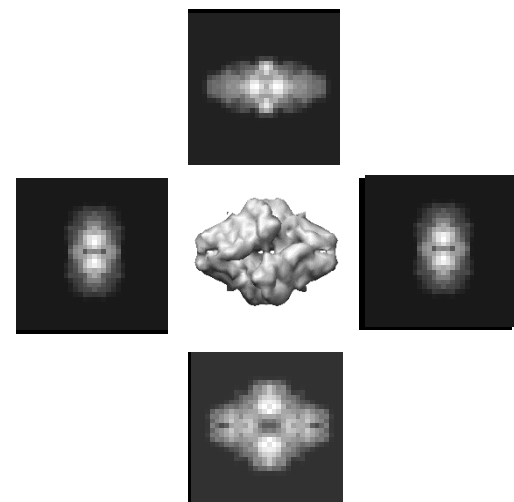
(1)3D画像から疑似投影像群を作成



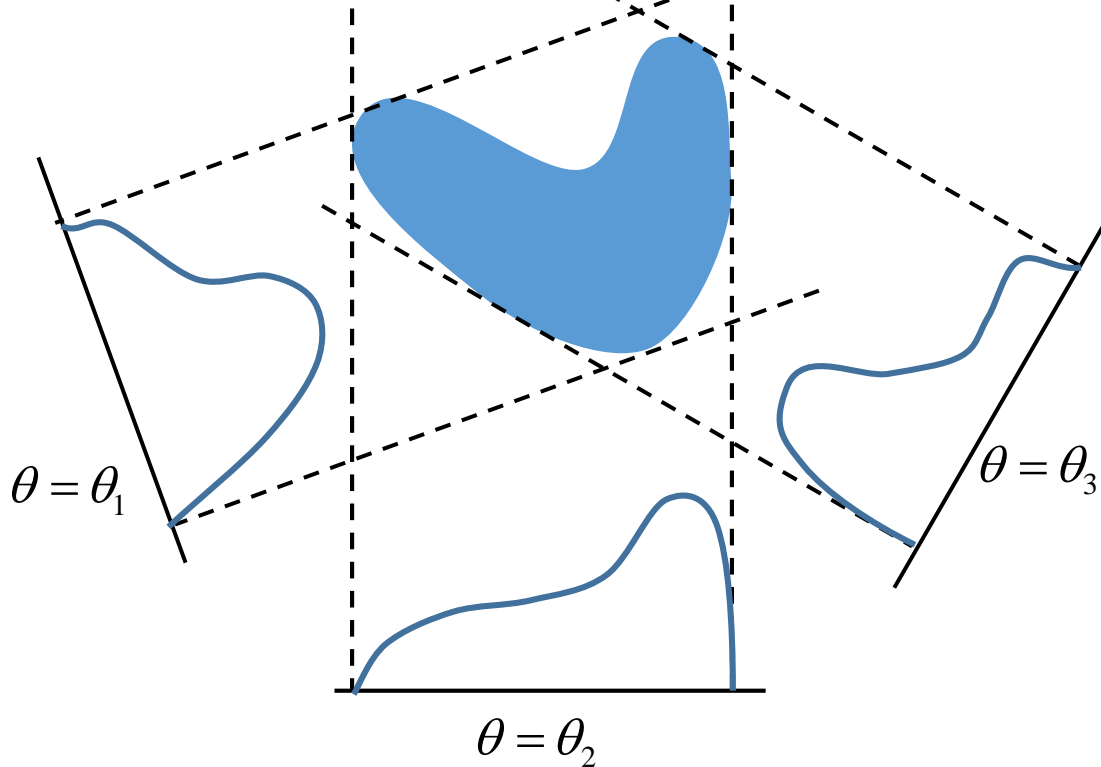
(2)各観察投影像に最も近い疑似投影像を選ぶ
⇒ 投影方向の推定



(3)推定投影方向を用いて3D画像を再構成。更新。



逆投影法: Back projection



$p_i(\mathbf{r}^{(i)})$: i -番目の投影像

$$\mathbf{r}^{(i)} = \{x^{(i)}, y^{(i)}\}$$

: i -番目の投影像のXY局所座標

$$b_i(\mathbf{r}^{(i)}, z^{(i)}) = p_i(\mathbf{r}^{(i)})t(z^{(i)})$$

: i -番目の投影像の「逆投影体」
("back projection body")

$$t(z^{(i)}) = \begin{cases} 1 & -D/2 \leq z \leq D/2 \\ 0 & \text{otherwise} \end{cases}$$

: "top hat" function

ナイーブな逆投影法では、低周波成分が強調された3D画像が再構成されてしまう。

=> 投影画像に高周波成分を強調するフィルタをかけてから逆投影をする

重み付き逆投影法 (weighted back projection)

フィルタ補正逆投影法 (filtered back projection)

$$\bar{\rho}(\mathbf{r}) = \sum_{i=1}^N b_i(\mathbf{r}^{(i)}, z^{(i)})$$

: 再構成された3D画像

フーリエ再構成法 Fourier reconstruction

投影断面定理 Projection-slice theorem

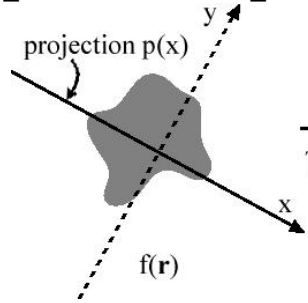
$$p(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$F(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-2\pi i(xk_x + yk_y)] dx dy$$

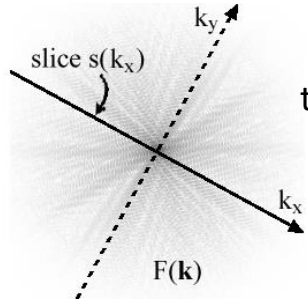
$$F(k_x, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-2\pi i x k_x] dx dy = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x, y) dy \right] \exp[-2\pi i x k_x] dx = \int_{-\infty}^{\infty} p(x) \exp[-2\pi i x k_x] dx$$

2D: $\int_{-\infty}^{\infty} p(x) \exp[-2\pi i x k_x] dx = F(k_x, 0)$

y軸に投影した1D投影像をフーリエ変換すると、2Dの周波数空間での $k_y=0$ の直線になる。



Fourier Transform

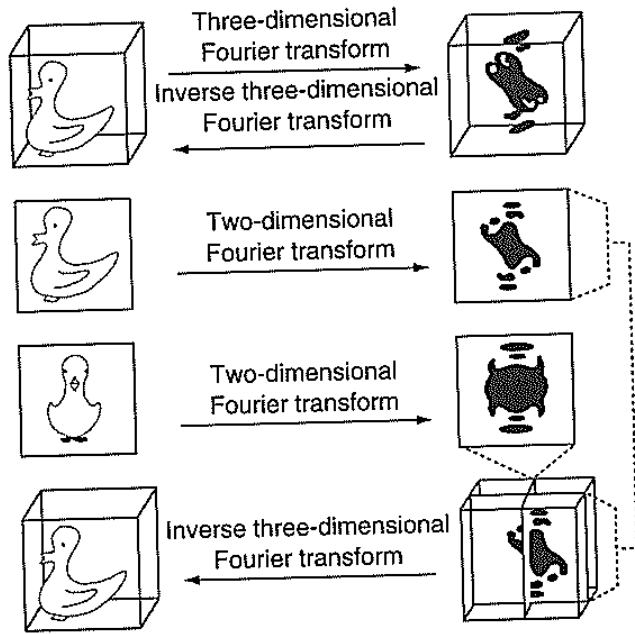


taken from wiki

3D: $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) \exp[-2\pi i(xk_x + yk_y)] dx dy = F(k_x, k_y, 0)$

Fourier transform of 2D projection is the slice plane the Fourier 3D map.

z軸に投影した2D投影像をフーリエ変換すると、3Dの周波数空間 $k_z=0$ の平面になる。



taken from Frank (2006), p 196

画像解析のソフトウェア

統計情報 - 3次元再構成 ソフトウェア		EMDBの統計情報 (EM navigatorによる)	
3次元再構成 ソフトウェア	▲	データエントリ数	
n/a	2401		
RELION	1269		
SPIDER	1155		
EMAN	1013		
IMOD	435		
IMAGIC	405		
FREALIGN	377		
XMIPP	284		
SPARX	228		

プログラム	開発グループ	コメント
RELION	Sjor Scheres (MRC)	ベイズ推定(MAP EM)を採用。2D分類と3D構築。現在の一流誌の論文の多くはこのソフトを使っている。
SPIDER	Joachim Frank (Columbia Univ.)	ノーベル賞受賞者が開発を続けているレガシーなソフトウェアパッケージ。
EMAN2	Steven Ludtke (Baylor College of Medicine)	Relion出現前は、最もよく使われていたパッケージだったが、最近ではRelionに圧倒されている。
FREALIGN	Nikolaus Grigorieff (Howard Hughes Medical Institute)	3D refinementに特化したソフト。

実際の画像解析の例

EMAN2のチュートリアルから

EMAN2 <http://blake.bcm.edu/emanwiki/EMAN2>

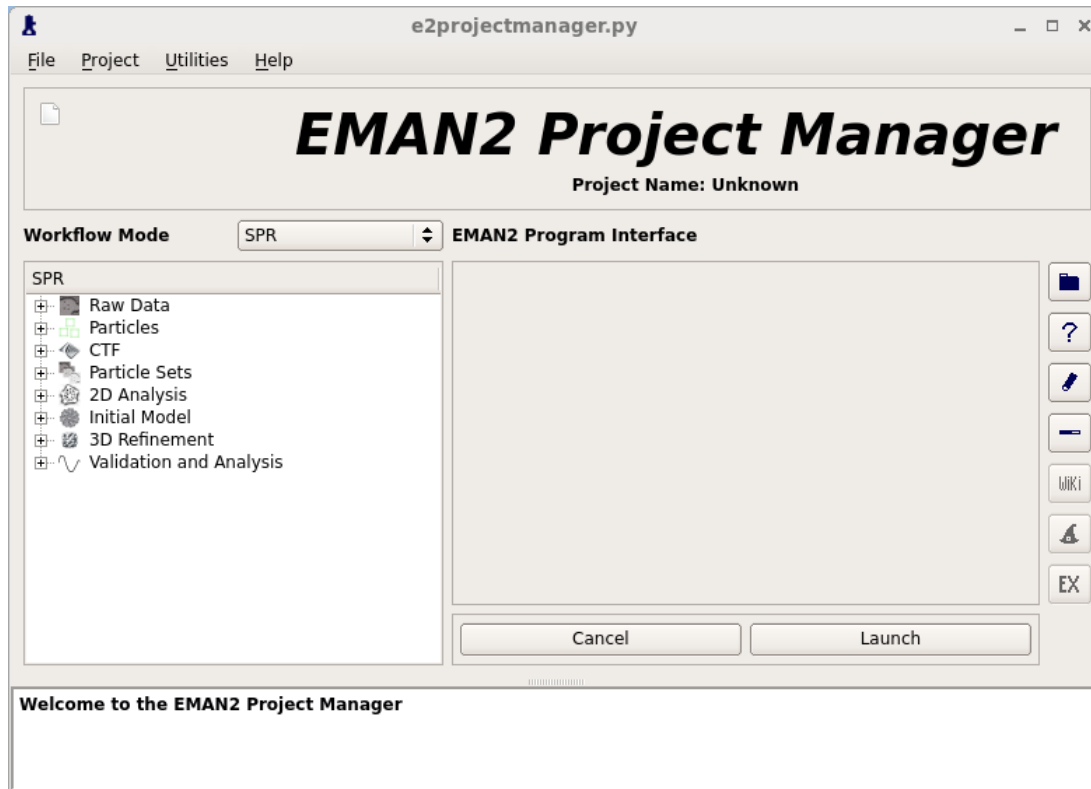
Stephen Ludtke のグループ (Baylor College of Medicine, TX,US)で開発

C++. で書かれ、Pythonでラップ

多くのpythonのスク립トが提供 : **e2proc2d.py**, **e2proc3d.py**, **e2boxer.py**, ...

GUI付きのスク립トもある (such as e2boxer.py, e2display.py, e2projectmanger.py...).

並列化は、マルチスレッドとMPI による並列化計算が可能。GPUには未対応。



一連の作業を管理する
ウィザードタイプのGUIスク립ト
e2projectmanager.py
も提供されている。

使用する電顕画像データ

- EMAN2のチュートリアルから取得

<http://blake.bcm.edu/emanwiki/EMAN2/Tutorials>

・原論文 : Bartesaghi A, Matthies D, Banerjee S, Merk A, Subramaniam S Structure of β -galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy, PNAS.(2014) **111**, 11709-11714; **EMDB-5995**; PDBcode:**3j7h**).

- ・原論文で用いた画像データはEMPIARデータベースからダウンロードできる。

EMPIAR-10012: micrographs (single image) 509 images in MRC format, 7420 x 7676.

Total file size :**108.0 GB**

EMPIAR-10013: micrographs (multiple image) 509 images in MRC format, 7420 x 7676 ,

Frames per image 38, Total file size: **402.2GB**

- ・チュートリアルのデータは、練習用にデータを小さくしている。

動画は扱わない

509枚の画像 => 96枚の画像

画像の解像度 7420 x 7626 pixels => 3838 x 3710 pixels

96 images in HDF format: such as, BGal_000021.hdf, Bgal_000025.hdf, ...,

Bgal_000496.hdf. Each HDF file has 54 MB, 3838 x 3710 pixels.

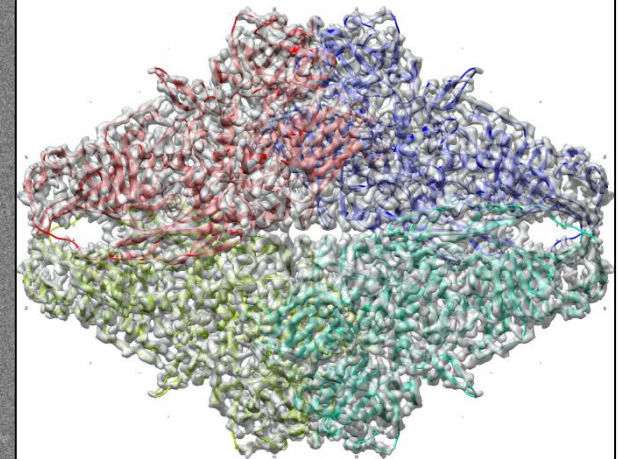
Total file size is 54MB x 96 = **5 GB**.

低温電子顕微鏡の画像の例



BGal_000021.hdf
3838 x 3710 pixels.

最終的な3次元マップ
(emd_5995, PDB:3j7h)



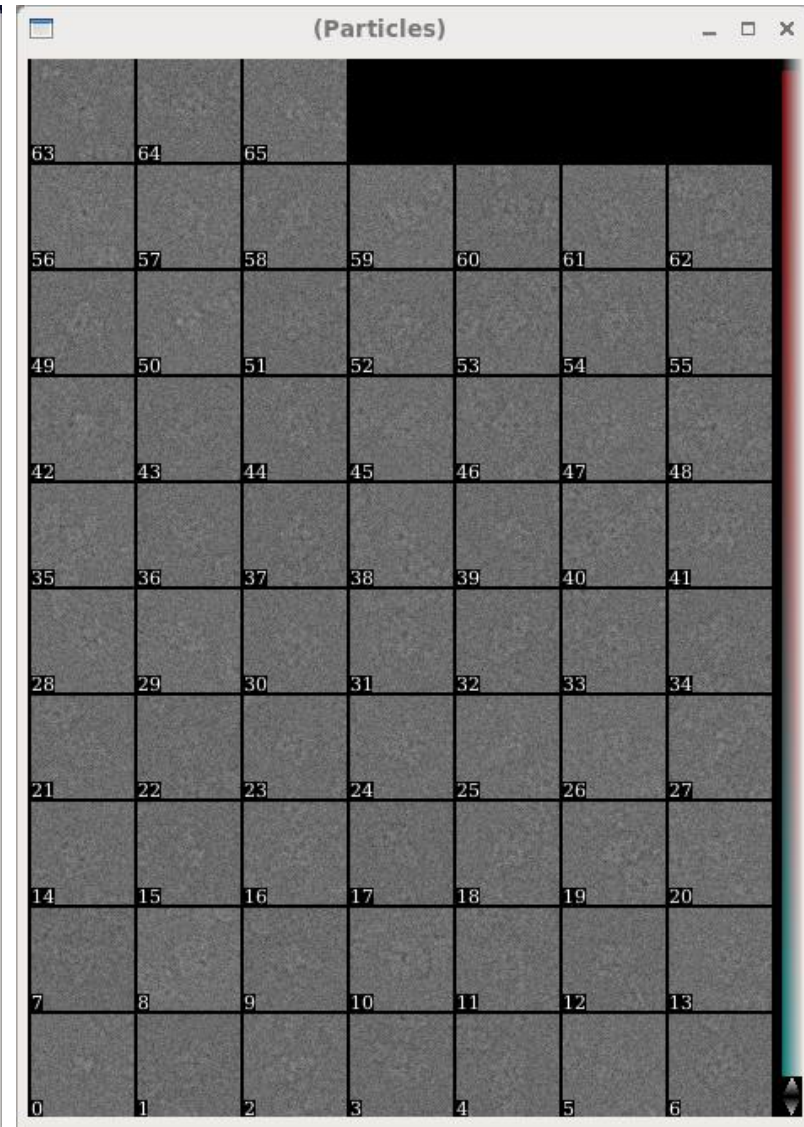
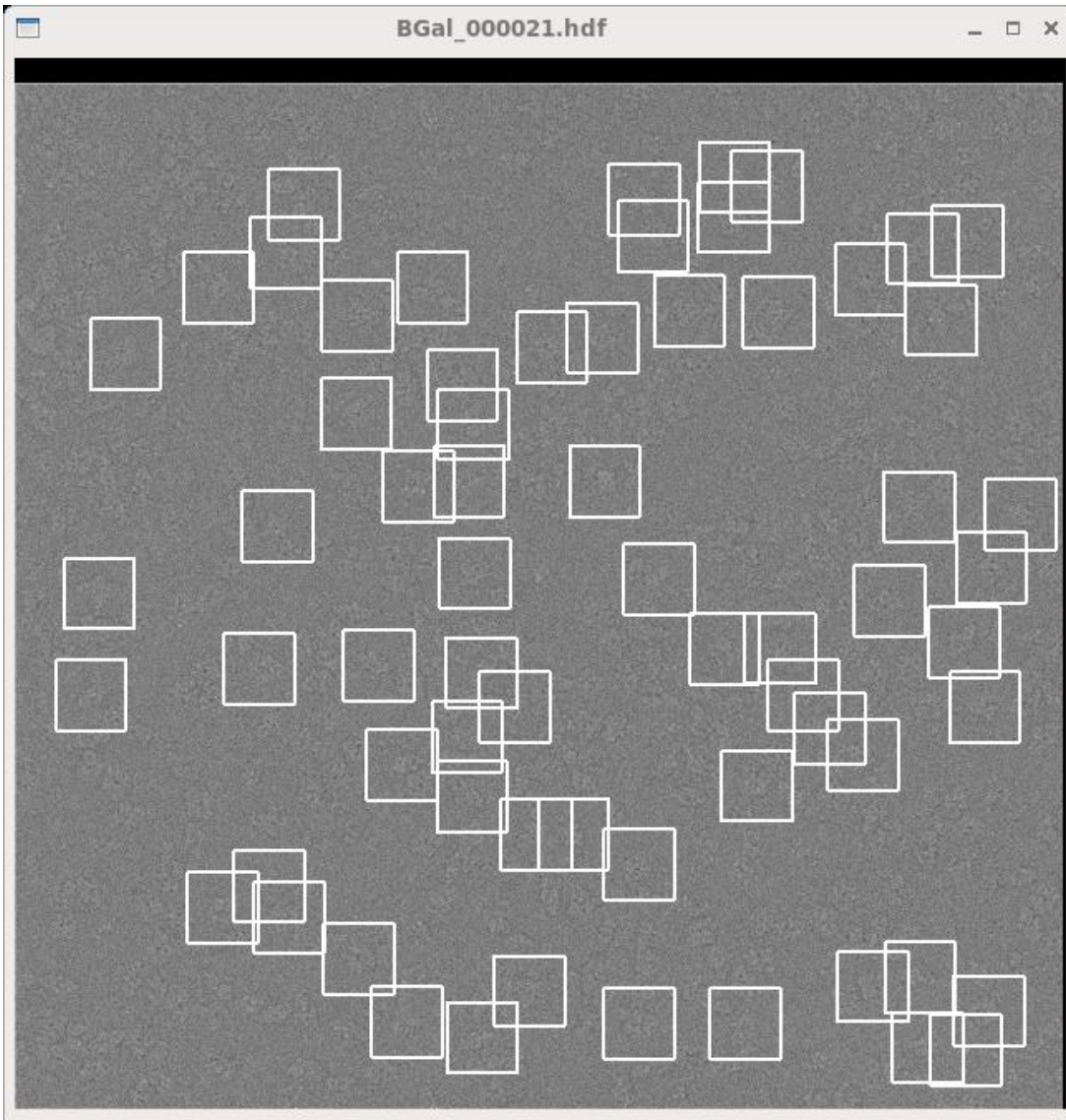
EMAN2の処理の手続き

Motion Correction (動画補正)

計算時間は4コアのデスクトップマシンによる
(Dell Precision T1700, Xeon E3-1220 v3, 3.10GHz, 4 threads).

1. Prepare your project
2. Launch e2projectmanager.py
3. Import and estimate image defocus [1min]
4. Extracting/boxing particles from micrographs 粒子画像の抽出
5. CTF Correction 位相コントラスト伝達関数による補正 [4min]
6. Building sets
7. Generating reference free class average [15min] 2D画像分類と平均画像
8. Eliminate bad particles
9. Select good class averages
10. Make initial models [1min] 初期密度3Dマップの生成
11. Building sets
12. 3D refinement 1 [12 min] 3Dマップの精密化 1
13. 3D refinement 2 [284 min = 4hour 44min] 3Dマップの精密化 1
14. Eliminate bad particles (optional)
15. Final refinement [673 min = 11hour 13min] 3Dマップの精密化 3

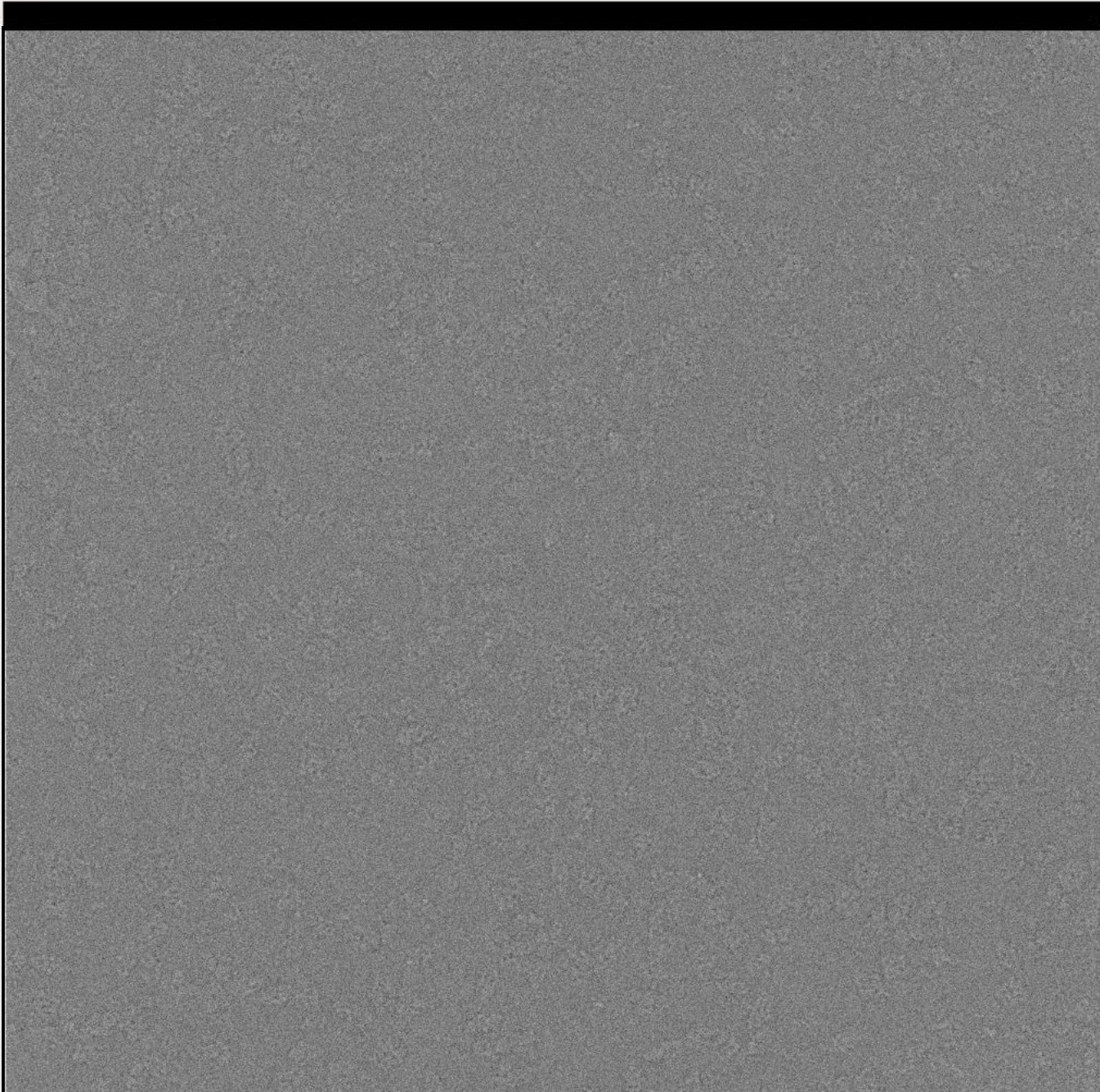
粒子画像の抽出 (picking, boxing)

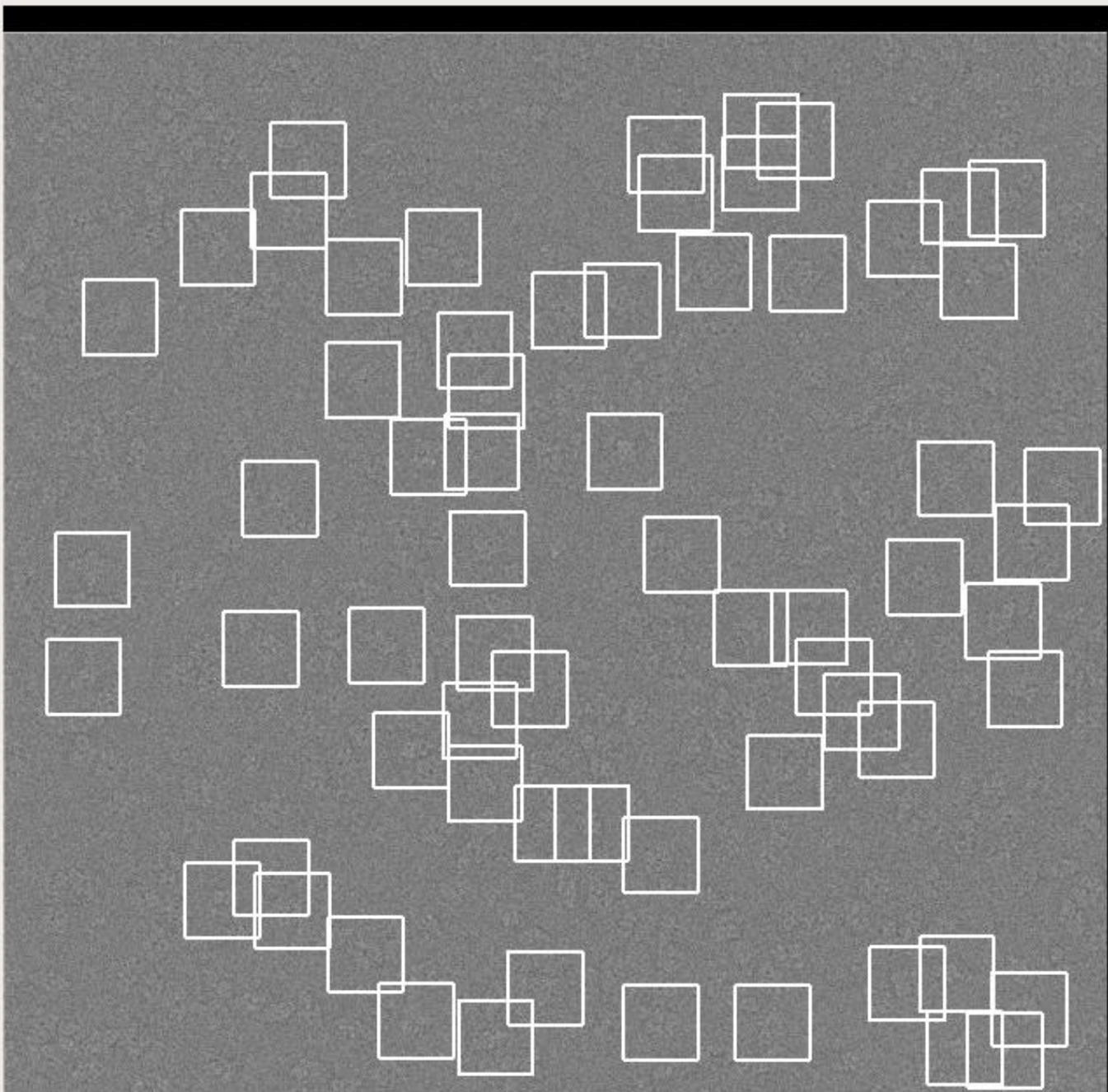


手作業で1000個程度抽出してから、残りは自動抽出する方法が一般的。最終的には数万~数十万個の粒子を使用する。チュートリアルデータでは、抽出済みのboxファイルが提供



BGal_000021.hdf





位相コントラスト伝達関数 (CTF) による補正

CTF mathematically describes how aberrations in a transmission electron microscope modify the image of a sample. This sets the resolution of High-Resolution Transmission Electron Microscopy (HRTEM), also known as phase contrast TEM. By considering the recorded image as a CTF-degraded true object, describing the CTF allows the true object to be reverse-engineered. This is typically denoted CTF-correction, and is vital to obtain high resolution structures in 3D electron microscopy, especially cryo-electron microscopy. Its equivalent in light-based optics, is the optical transfer function.

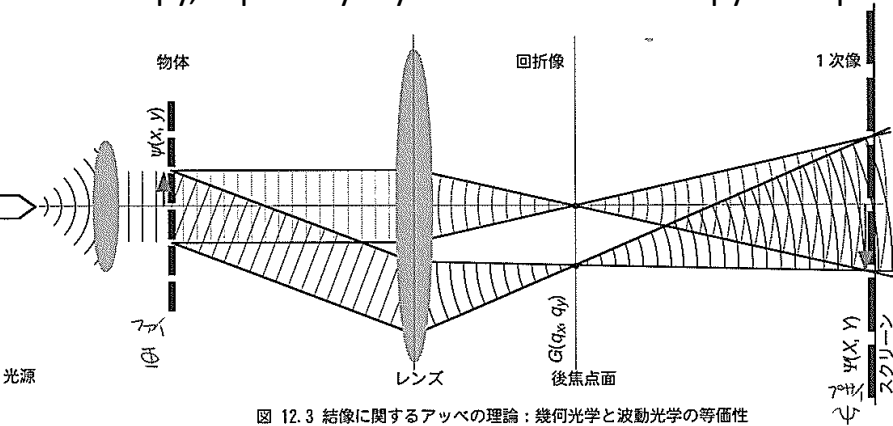
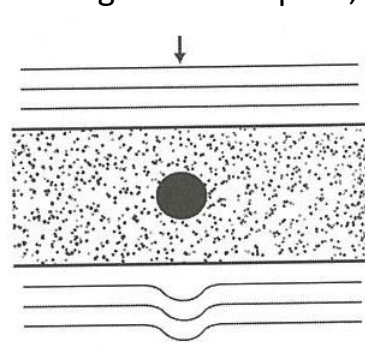
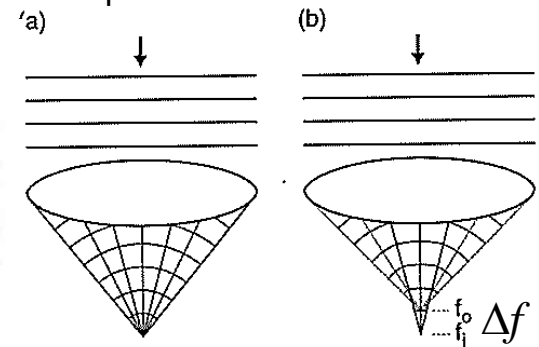


図 12.3 結像に関するアッペの理論：幾何光学と波動光学の等価性



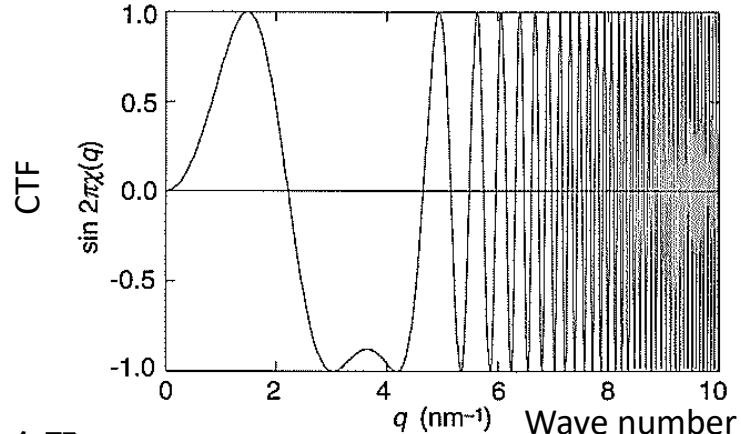
A phase object and its effect on an incoming plane wave.



Frank (2006) :defocus

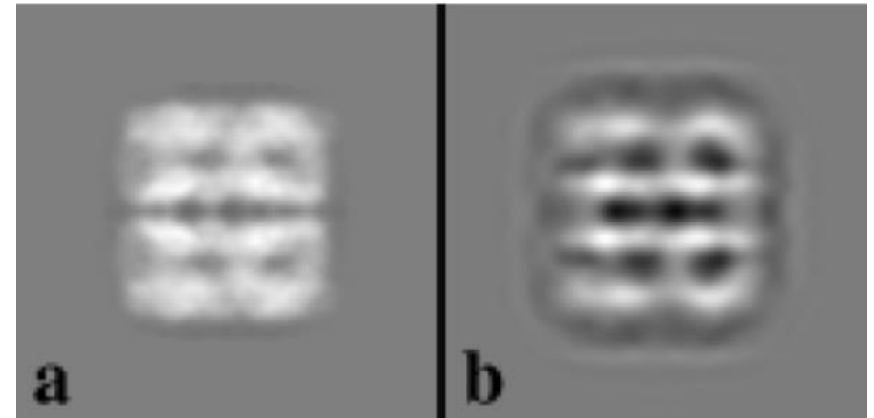
The effect of spherical aberration
 C_s : spherical aberration constant

CTF: $\sin(2\pi\chi(q)) = \sin(\pi\lambda\Delta f q^2 - \pi/2\lambda^3 C_s q^4)$
 λ : wave length



今野(2003)

図 12.13 コントラスト伝達関数の例
 $(\lambda=2.51\text{pm}, C_s=1.2\text{mm}, \Delta f=100\text{nm})$



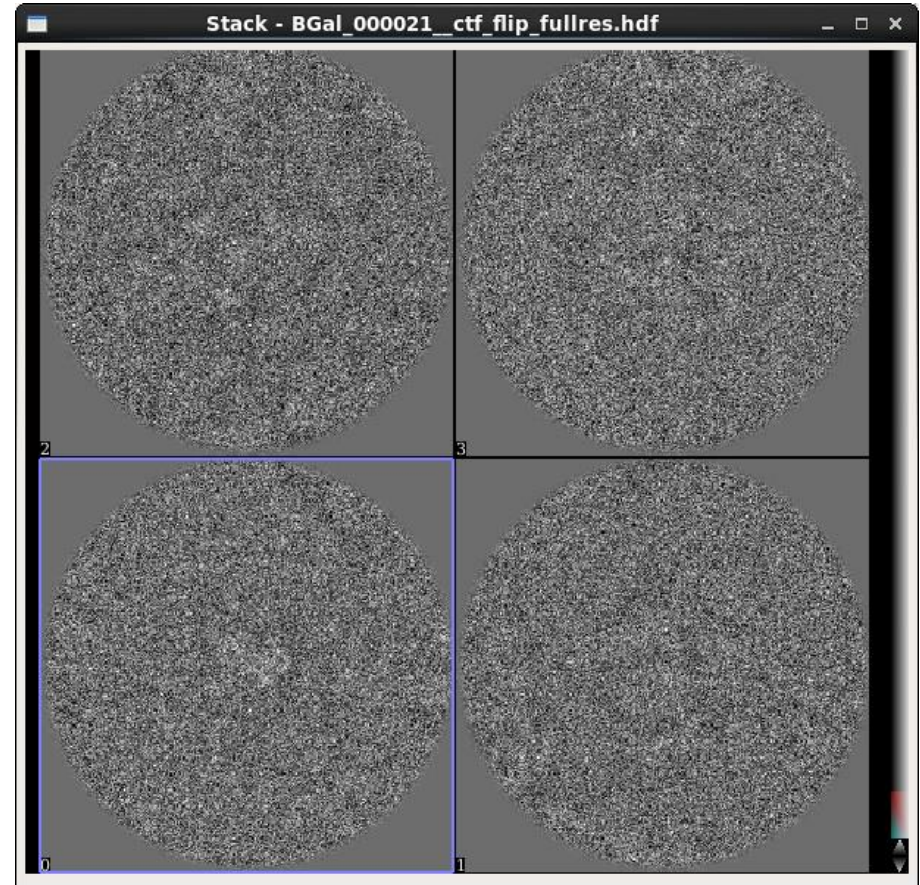
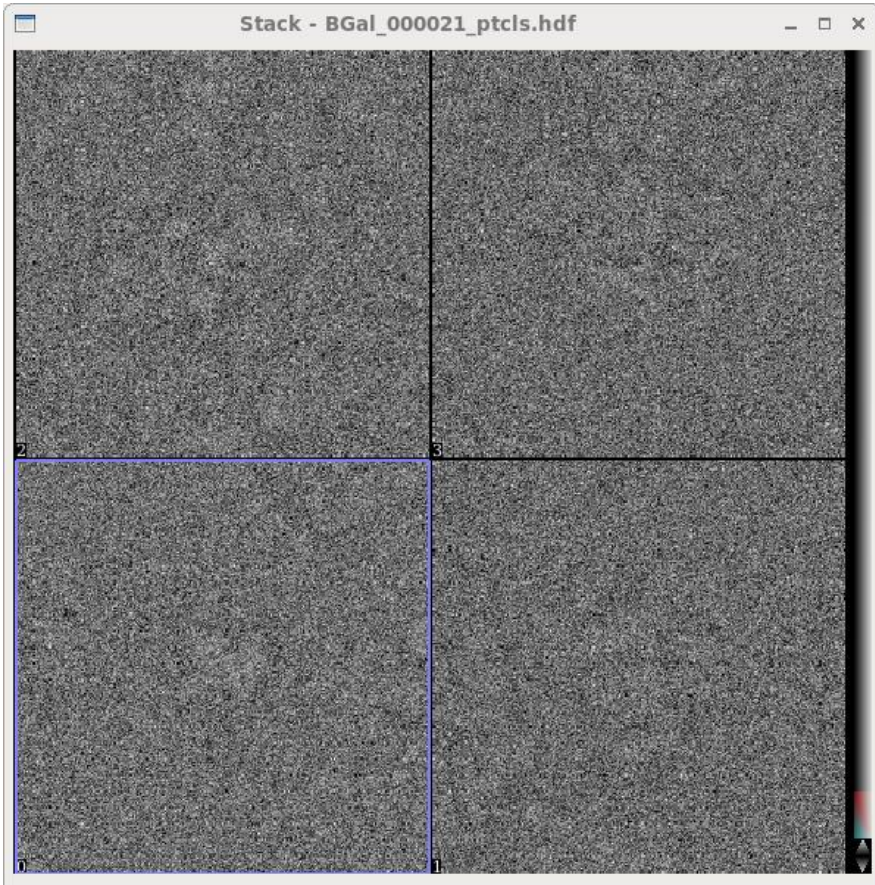
True projection of GroEL
 (Ludtke and Chiu)

Same particle CTF and
 envelop functions applied.

Example of CTF corrections

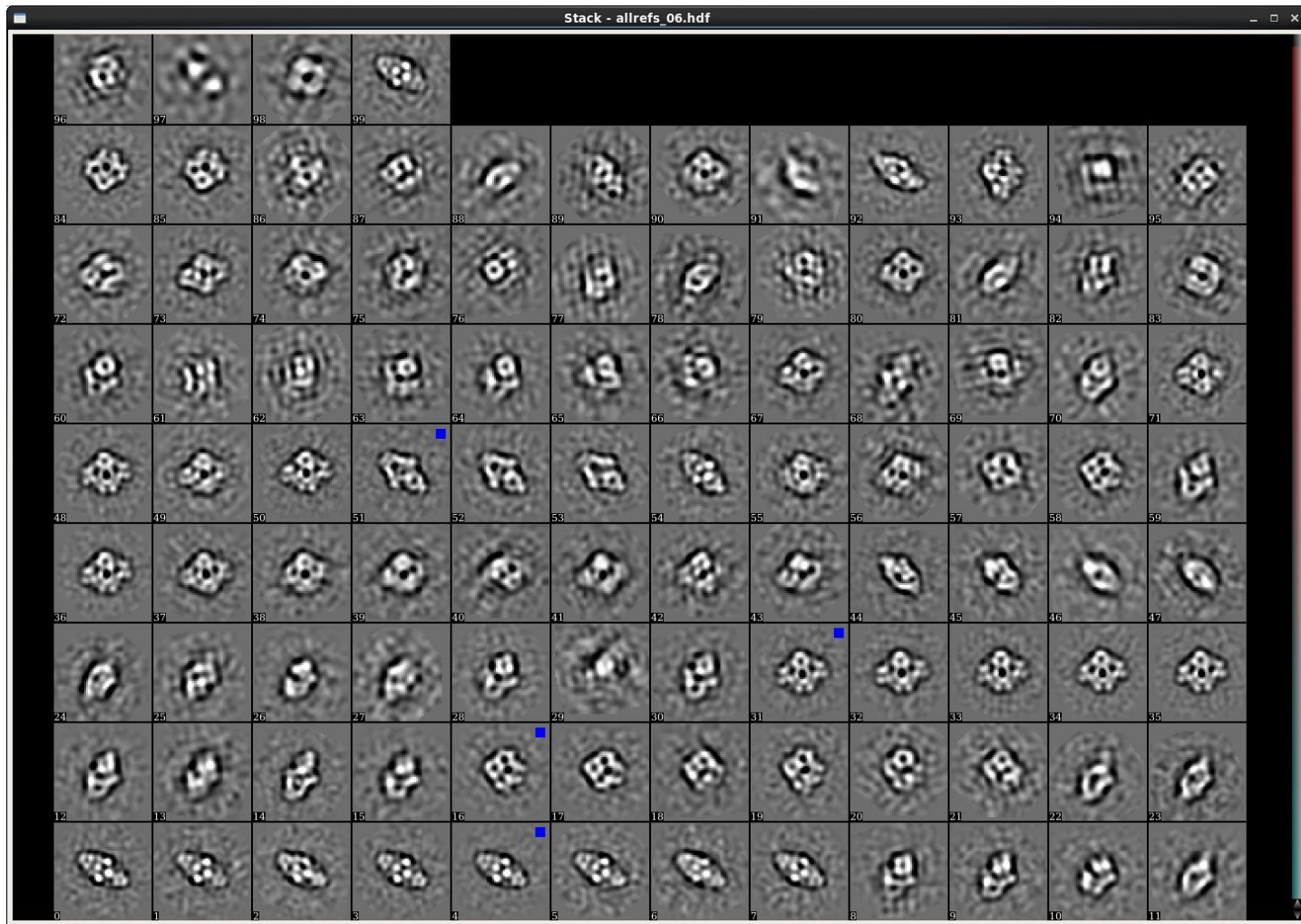
Before correction

After correction

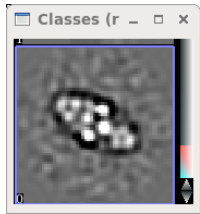


Step 7. 2D画像の分類と平均像の作成 (e2refine2d.py)

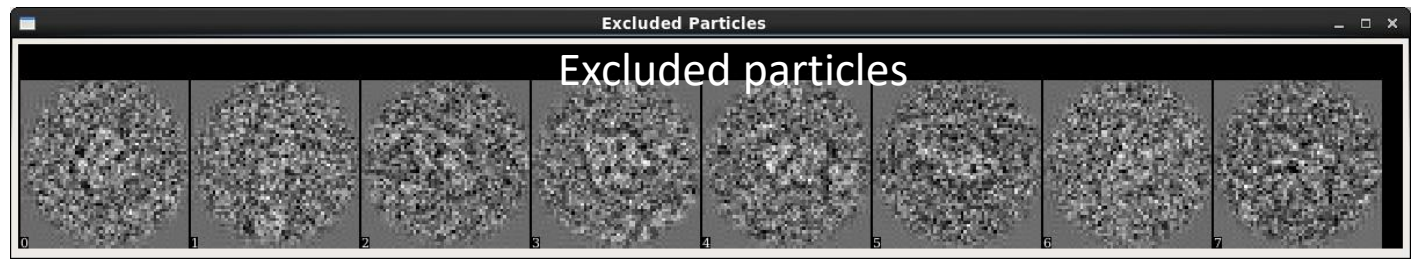
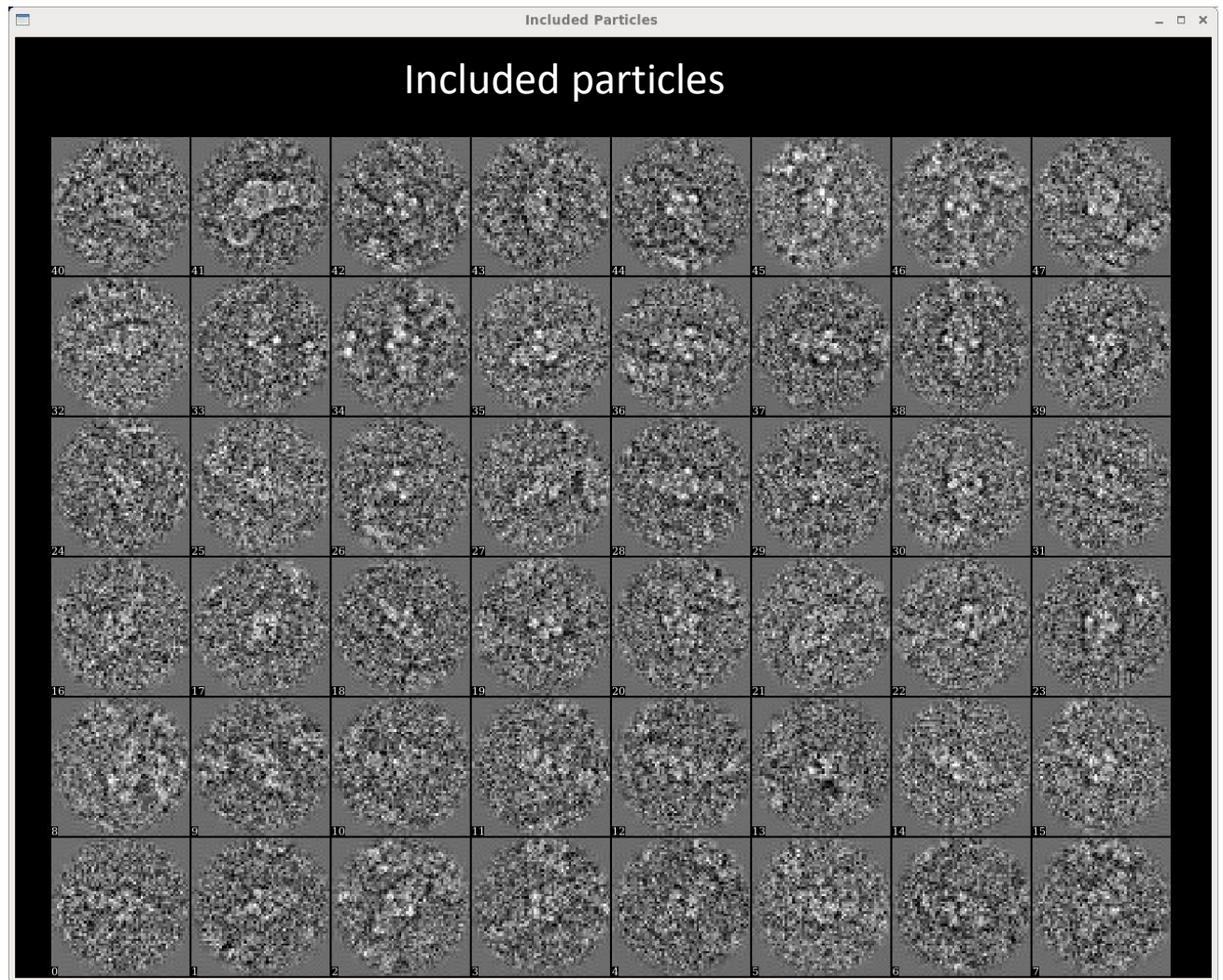
Input : all__ctf_flip_lp14.lst (5513 images with 56 x 56 pixels), Ncls :100



Example of “good” classes (class 0)



Averaged image



EMAN2の2D画像分類の算法(e2refine2d.py)

Iterative MSA (Multivariate Statistical Analysis) -based reference-free classification algorithm.

MSA and K-means for Rotational/translational invariant of 2D images

1. Initialize the iterative process by making some initial guesses at class-averages. These are ***invariant-based***, meaning even with MSA, this initial classification is not exceptionally good.
 - a. Generate ***rotational/translational invariants*** for each particle (input_fp)
 - b. Perform MSA on the invariants to define an orthogonal subspace representing the most important differences among the classes (input_fp_basis) (MSA may be PCA ?)
 - c. Reproject the particles into the MSA subspace using `--nbasis` vectors (input_fp_basis_proj)
 - d. Classify the particles into `--ncls` classes using ***K-means*** (classmx_00)
 - e. Iterative class-averaging of the particles in each class to produce a set of initial averages (classes_init)

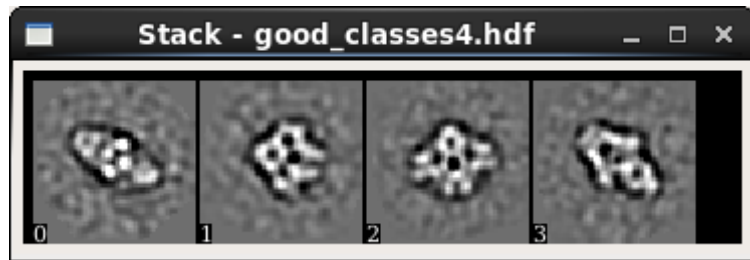
MSA and K-means for aligned 2D images

2. Align the current class-averages to each other, and sort them by similarity, keeping them centered (allrefs_YY) (Note that YY starts with 01 and is incremented after each iteration)
3. Perform MSA on the (aligned) class-averages. Again, this represents largest differences, but now performed on images, not invariants. (basis_YY)
4. Select a subset of `--naliref` averages to use as alignment references for this iteration (aliref_YY)
5. Align each particle to each of the reference averages from the last step. Keep the orientation corresponding to the best-matching reference. (simmx_YY)
6. Project aligned particles using reference MSA vectors from basis_YY (input_YY_proj)
7. K-means classification of input_YY_proj (classmx_YY)
8. New iterative class-averages (classes_YY)
9. Loop back to step 2 until `--iter` loops are complete

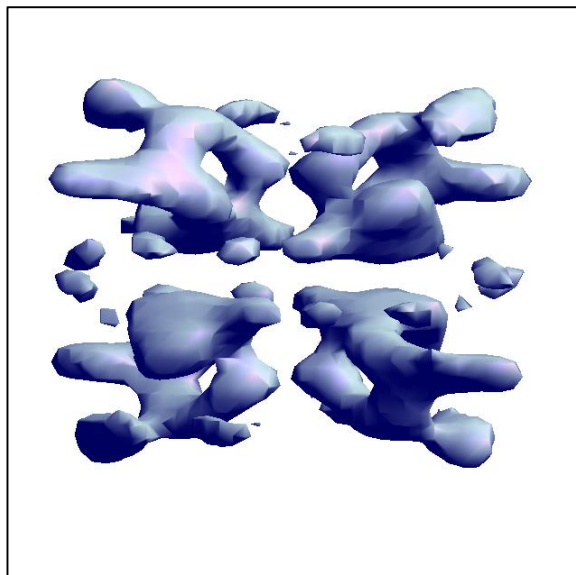
初期モデルの作成 (e2initialmodel.py)

※D2の対称性の拘束を用いる.

4つの代表画像から作成した場合

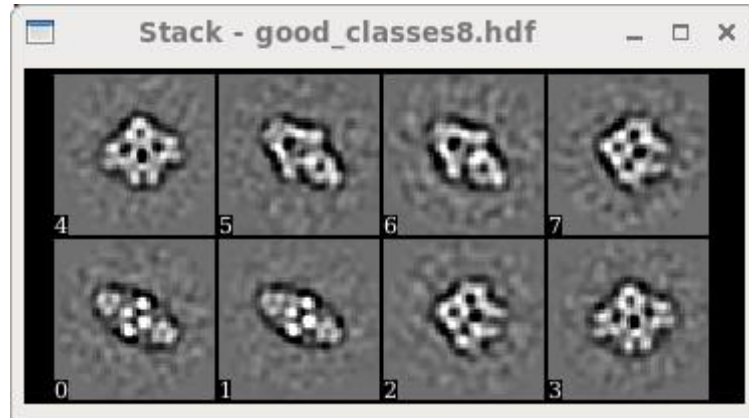


56 x 56

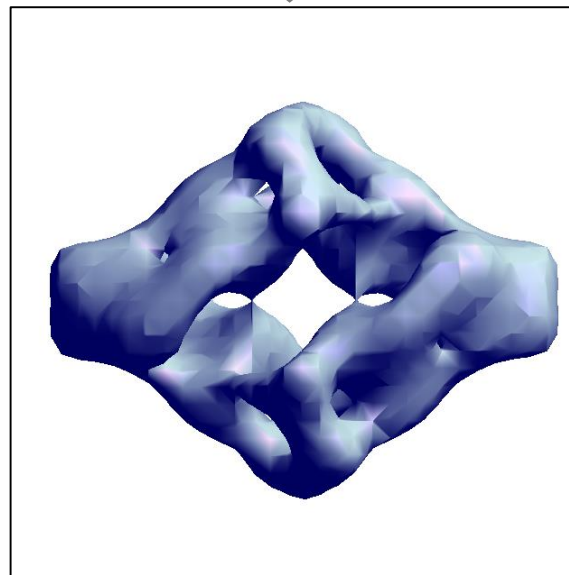


56 x 56 x 56

8つの代表画像から作成した場合



56 x 56



56 x 56 x 56

3Dマップの精密化 1 (e2refine_easy.py)

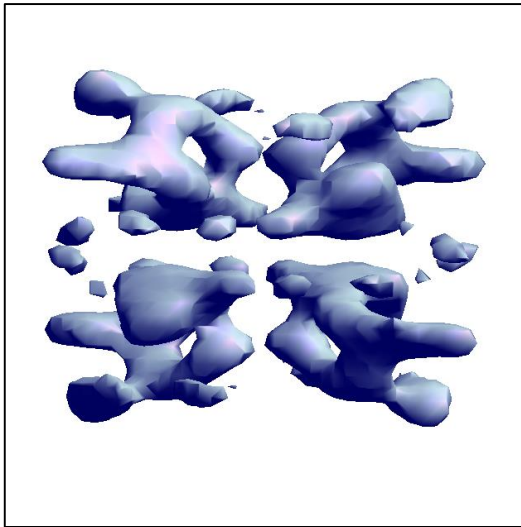
Input : sets/all-bad1__ctf_flip_lp14.lst (5513 images with 56 x 56)

Initial 3D map is taken from the step 10 (initial model).

targetres : 18 Å

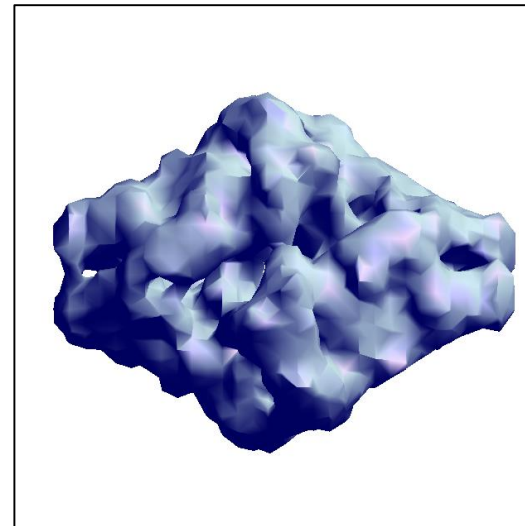
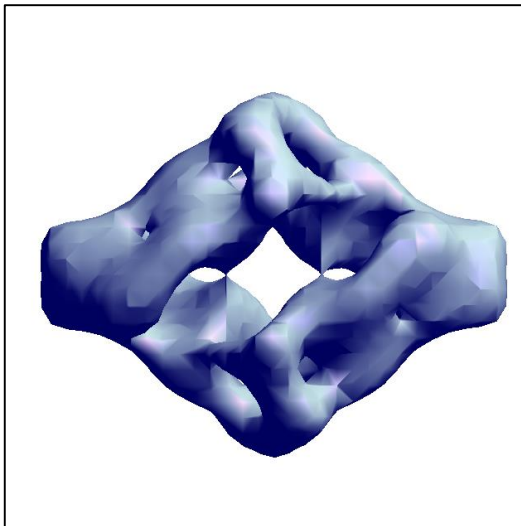
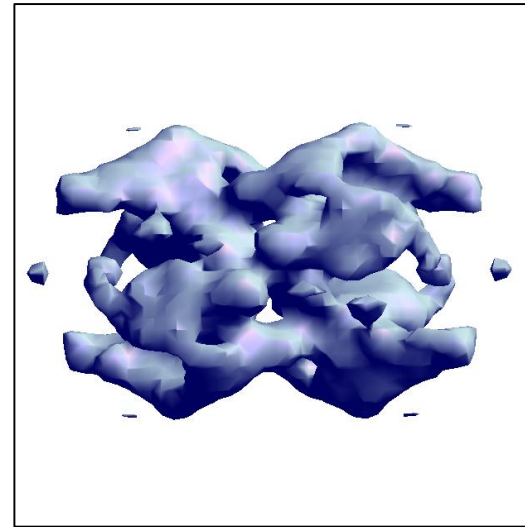
Initial model

56 x 56 x 56



refined model1

56 x 56 x 56



e2refine_easy.py
Includes
re-classification
of 2D images.

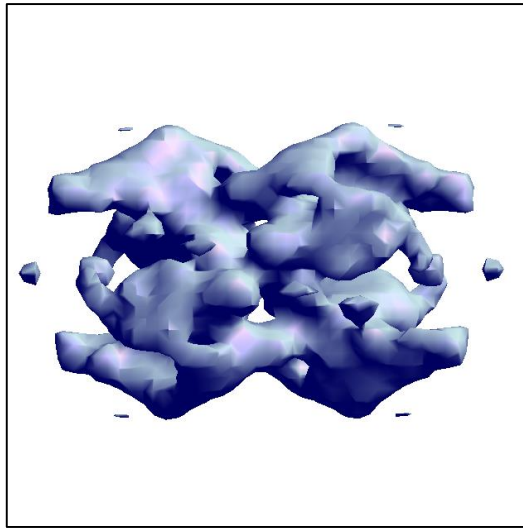
3D マップの精密化2 (e2refine_easy.py)

Input : sets/all-bad1__ctf_flip_lp5.lst (5513 images with 128 x 128)

Initial 3D map is taken from the step 12 (3D refinement1). targetres : 6 Å

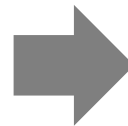
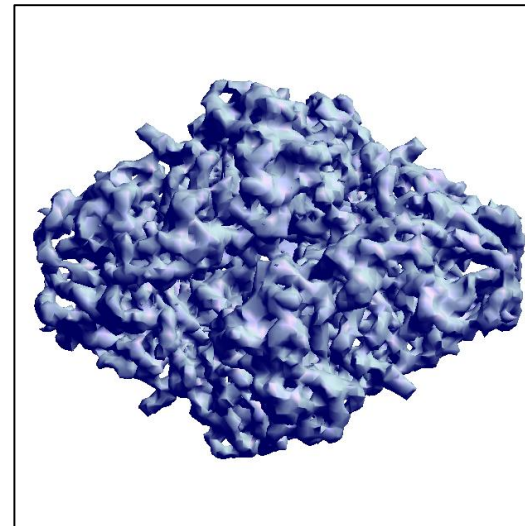
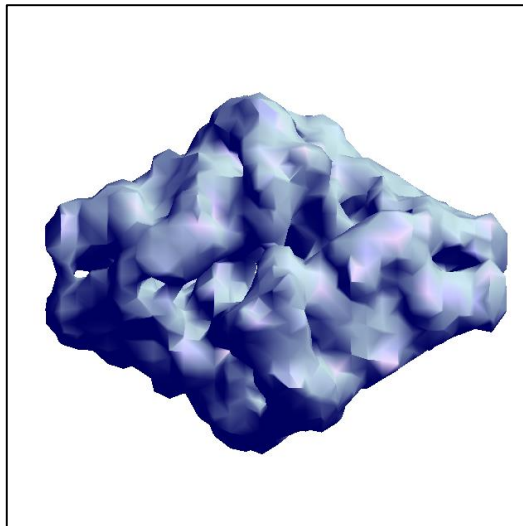
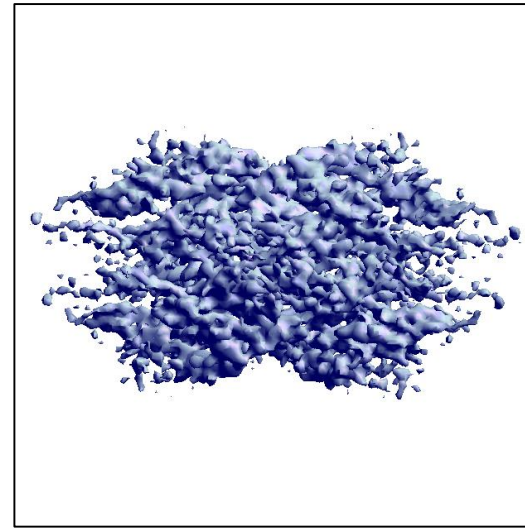
refined model1

56 x 56 x 56



refined model2

128 x 128 x 128



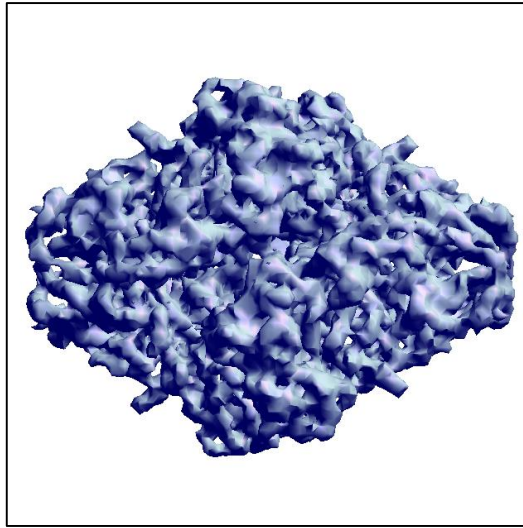
最終精密化(e2refine_easy.py)

Input : sets/all-bad1__ctf_flip_fullres.lst (5513 images with 256 x 256)

Initial 3D map is taken form the step 13 (3D refinement2). targetres : 4Å

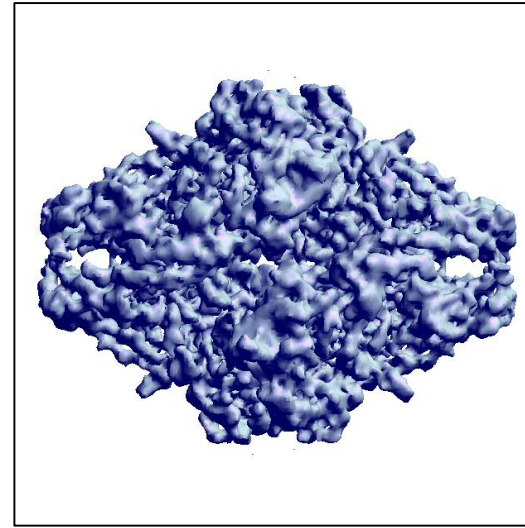
refined model2

128 x 128 x 128

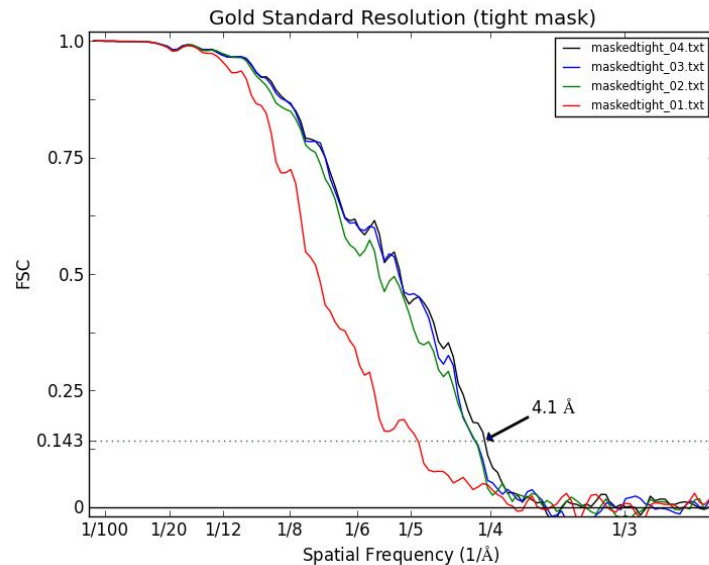


Final refinement

256 x 256 x 256

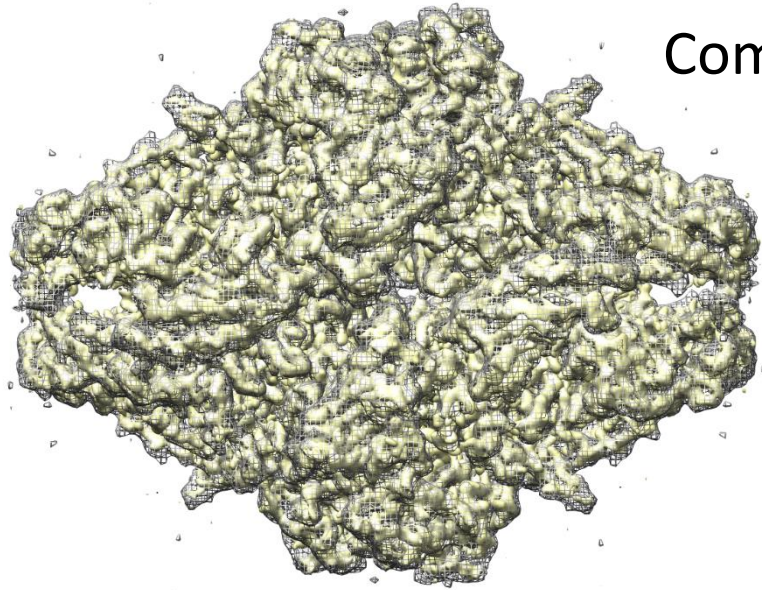


FSC : Fourier Shell Correlation



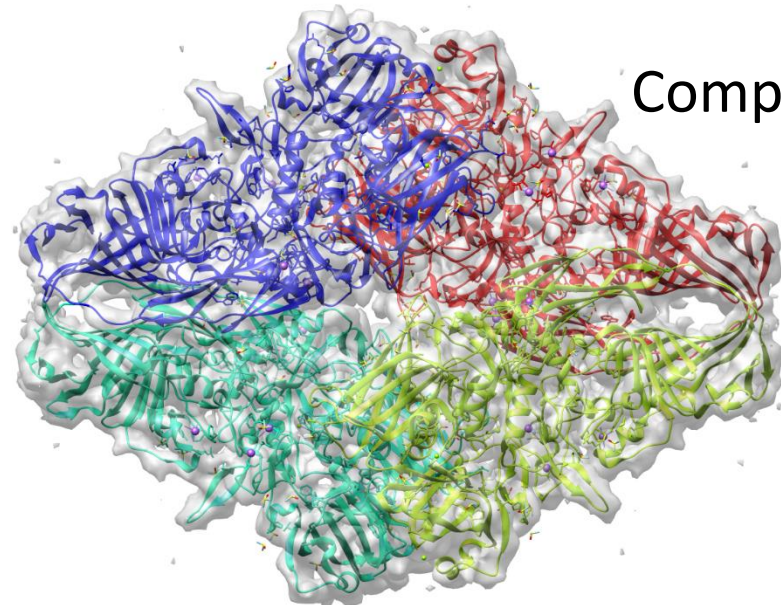
Resolution
: 4.1 Å

他のマップやX線構造との比較



Comparison with EMD-5995

Corr.Coeff. = 0.9175



Comparison with X-ray crystal structure (1dp0)

Corr.Coeff. = 0.7426

(HOHs were removed.
Reso.For Sim.Map=4.0 Å)

RELION

MRCのSjor Scheres のグループが開発

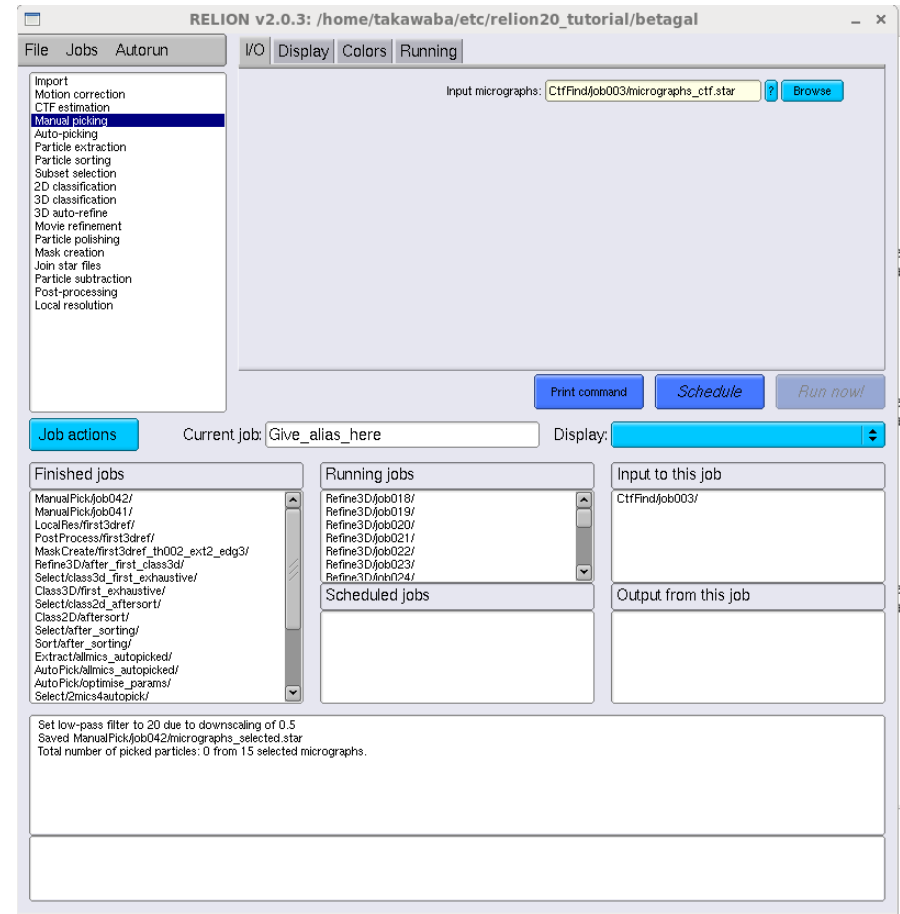
2D分類と3D再構成のプログラムを提供。
動画補正、CTF補正などは、GUIから別のプログラムを呼び出すことで対応。

2D分類と3D再構成では、ベイズ推定の枠組みを採用。最大事後確率推定(Maximum a posteriori estimation)のEMアルゴリズム。

3次元再構成では、周波数領域で行う。入力には、各2D画像を用い、2D平均画像は使わない。2D平均像は粒子抽出に用いる。

一つのサンプル(観察2D画像群)から、複数の3Dマップを推定することができる。

並列計算は、マルチコア(MPI), マルチスレッドに加え、ver 2から、GPUにも対応。

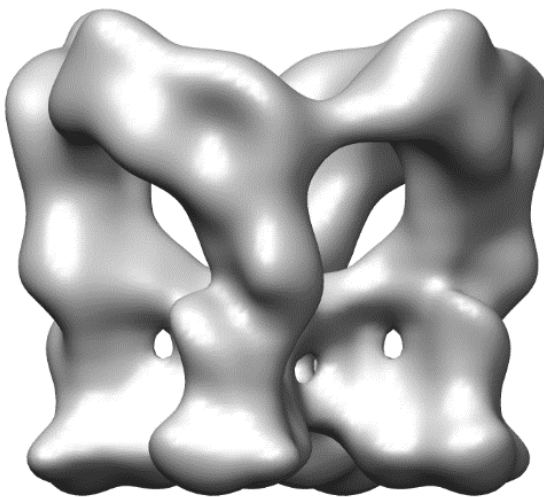


クラスタマシンでなくても
GPU2枚付きの1台のデスクトップ機で
現実的な時間で終了できる !

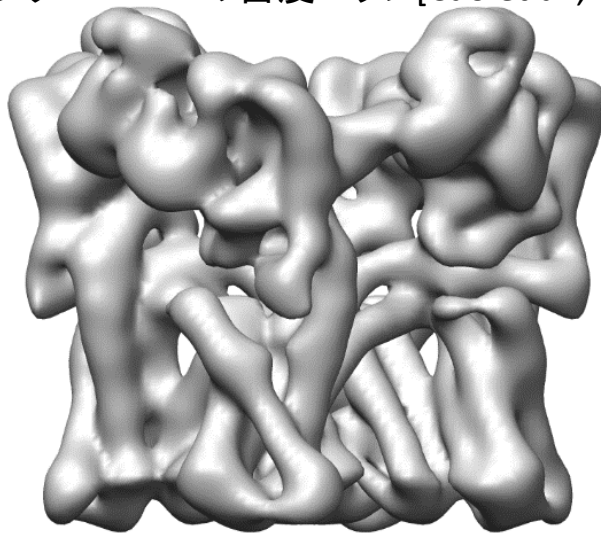
電顕の3D密度マップからの 原子モデルの構築法

マップの解像度による形状特徴の違い

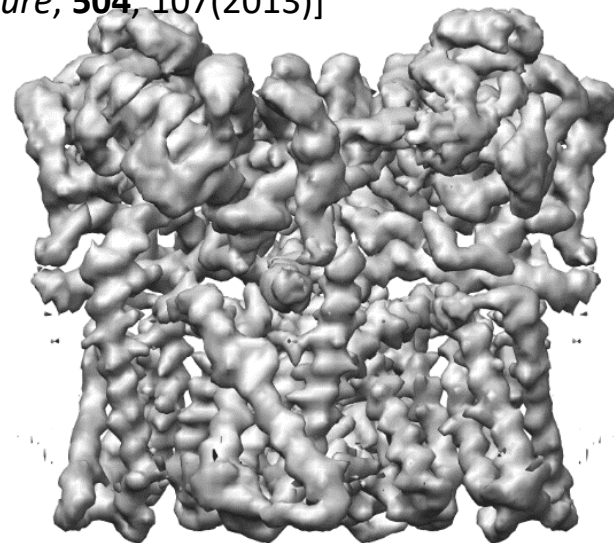
カプサイシンレセプターTRPV1の密度マップ[Cao et al., *Nature*, 504, 107(2013)]



解像度10 Å (疑似データ)
ドメイン



解像度5 Å (疑似データ)
2次構造(特にαヘリックス)



解像度3.275 Å (EMD-5778)
側鎖の一部(特にαヘリックス)

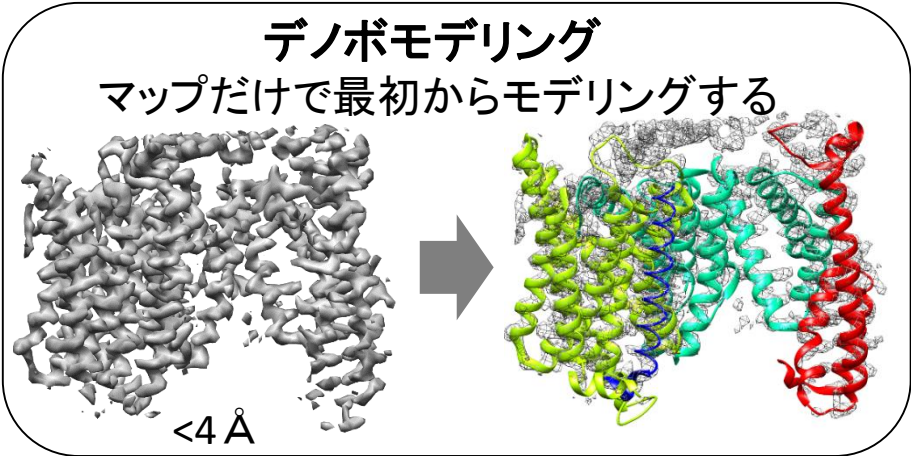
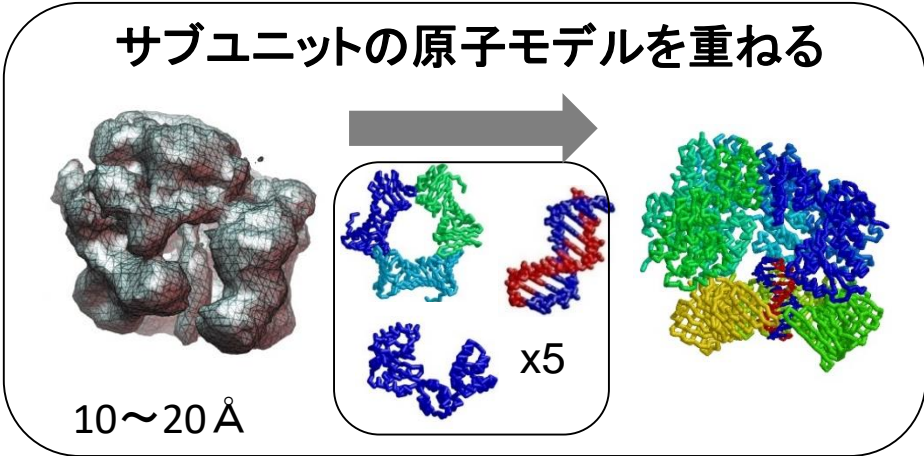


密度マップから構築した
原子モデル(PDB:3j5p)

モデリングの二つのアプローチ

低解像度の場合

高解像度の場合

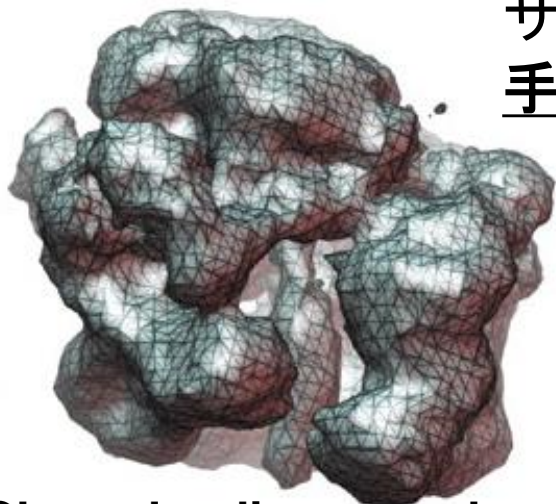


	低解像度	高解像度
解像度の目安	5 Å 以上	4 Å 以下
方針	サブユニットの原子モデル※を置く。それを変形・精密化	マップから原子モデルを構築。それを精密化
手作業で配置・モデリング	<i>UCSF Chimera</i>	<i>COOT</i>
自動で配置・モデリング	<i>SITUS, IMP, gmfit</i>	<i>ROSETTA, PathWalker</i>
変形・精密化	<i>MDFF (NAMD)</i>	<i>REFMAC, phenix_real_space_refine</i>

※ホモロジーモデリングによる予測構造もよく用いられる。

低解像度のモデリングの例

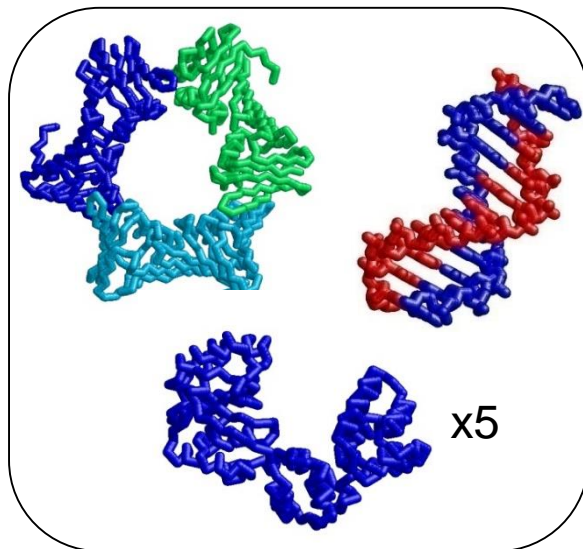
サブユニットの原子モデルを
手作業で密度マップに重ねる



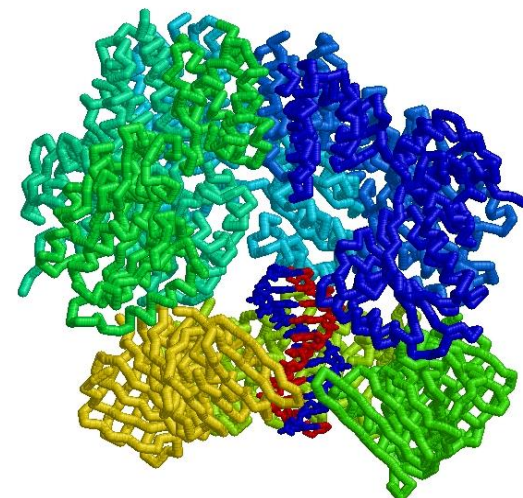
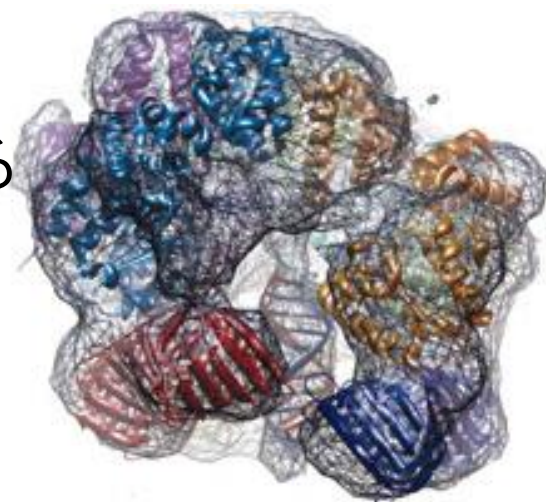
Clamp-loading complex

電顕による複合体の密度マップ

解像度:12Å



X線結晶解析で解かれた
サブユニットの原子モデル



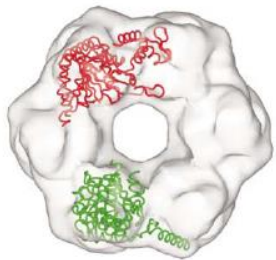
複合体の原子モデル

低解像度の密度マップ



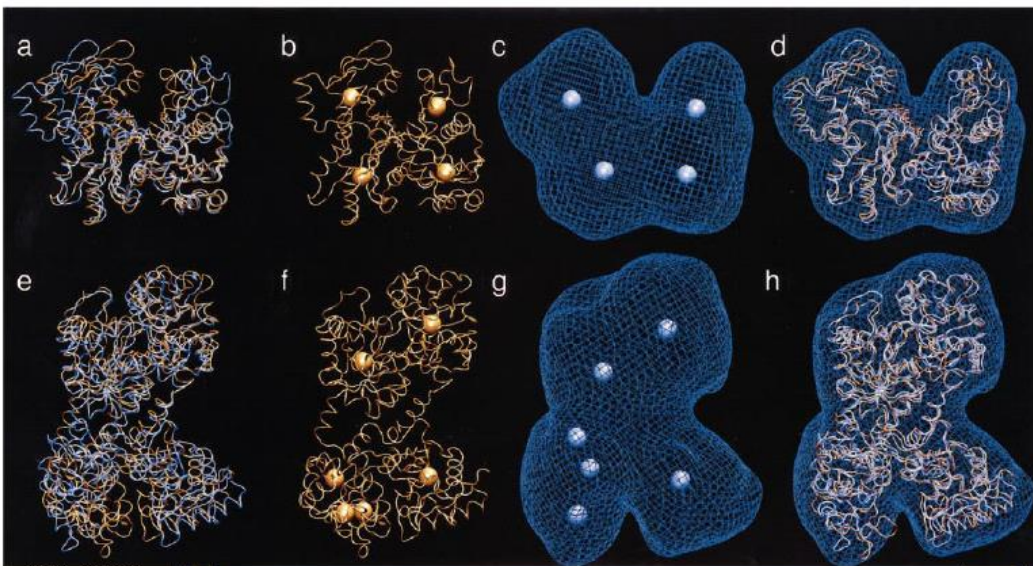
原子レベルの解像度のモデル

Miyata T, Suzuki H, Oyama T, Mayanagi K, Ishino Y, Morikawa K. (2005). Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. PNAS, 102, 13795-13800.



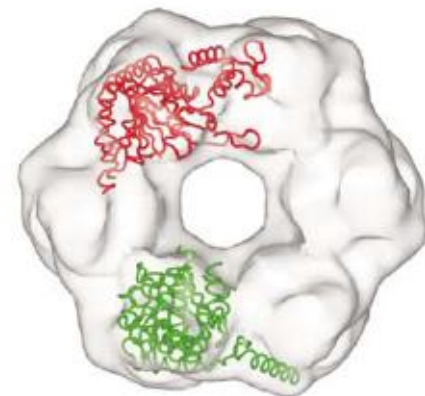
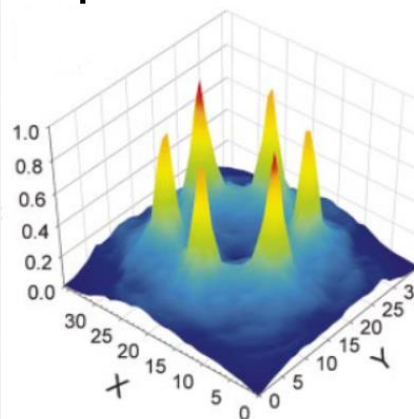
Program SITUS

The popular software for docking of atomic models into 3D density map
<http://situs.biomachina.org/>



Points for atomic model Points for 3D density Superimposed structure

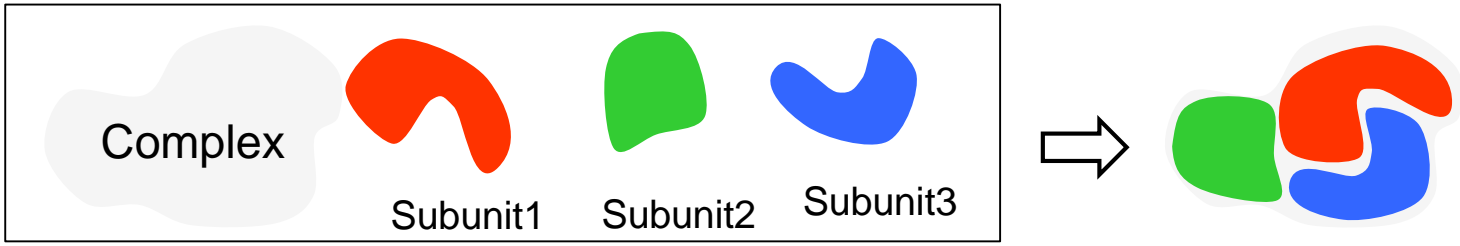
qdock : Representative 3D points are generated using **Vector Quantization**. The superimposing problem becomes the matching between two groups of 3D points.



colores - Exhaustive One-At-A-Time 6D Search .colores is a rigid-body search tool, suitable for one-at-a-time fitting of single subunit structures, which are not necessarily expected to account for the full map. The fitting procedure consists of 3 steps: (A) An exhaustive rigid body search on a discrete 6D lattice (B) an automatic ("black box") peak detection based on the correlation scores on the lattice (C) the final off-lattice refinement of solutions to the nearest maximum of the correlation).

ハイブリッド・モデリング

様々な部分的な実験情報を統合してモデリングを行うこと



いくつかの実験情報を満たすように最適化

マップへのフィット

Subunits fit density map of the complex



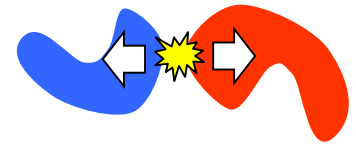
対称性

symmetric configuration of the subunits.



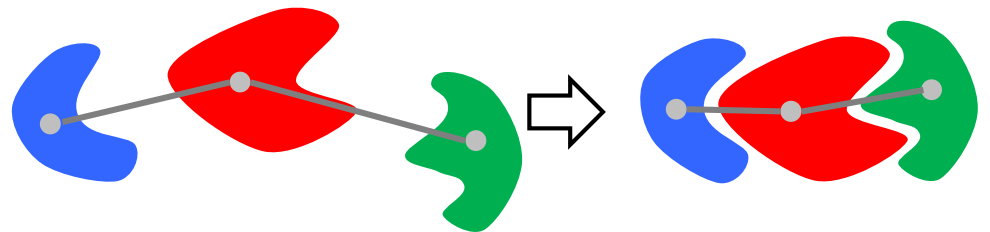
反発

repulsion among the subunits.

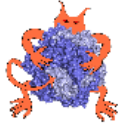


近接性

Proximity (contact/interaction) of the subunits obtained from the experiment (chemical cross-linking, mass spectroscopy..)



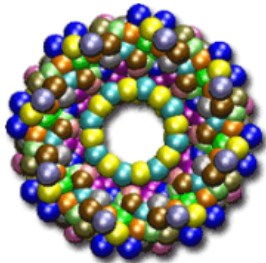
IMP: Integrative Modeling Platform



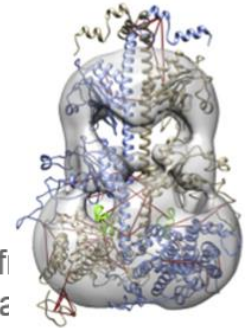
Developed by Andrej Sali Lab

[home](#) | [about](#) | [news](#) | [download](#) | [doc](#) | [source](#) | [systems](#) | [tests](#) | [bugs](#) | [contact](#)

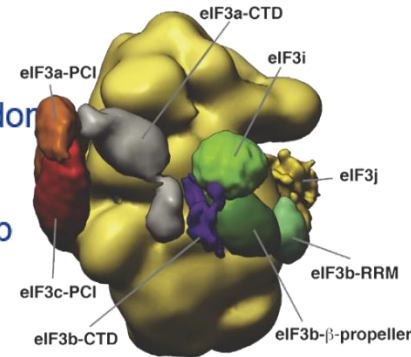
IMP, the Integrative Modeling Platform



IMP's broad goal is to contribute to a comprehensive structural characterization of biomolecules ranging in size and complexity from peptides to large macromolecular assemblies, by integrating data from biochemical and biophysical experiments. IMP provides an open source C++ and Python toolbox for solving complex modeling problems, and a number of applications for tackling some common problems in a user-friendly way. IMP



- FoXS: fitting structures against SAXS data
- MultiFoXS: determination of multi-state models
- Integrative docking: combining experimental data with computational pairwise docking
- Conformational sampling: sampling antibody conformations using the rapidly exploring random walk
- EMageFit: solution of a complex structure using subunit structures and EM class averages
- MultiFit: solution of a complex structure using subunit structures and a cryo-EM density map
- CNMultiFit: solution of symmetric complexes using a cryo-EM density map
- SAXSMerge: merging multiple SAXS profiles



gmfit: fitting using Gaussian Mixture Model(GMM)

Atomic model for Subunit1

$$f_k(\mathbf{r}) = w_k \sum_m^{M_k} w_{k,m} \phi_{k,m}(\mathbf{r})$$

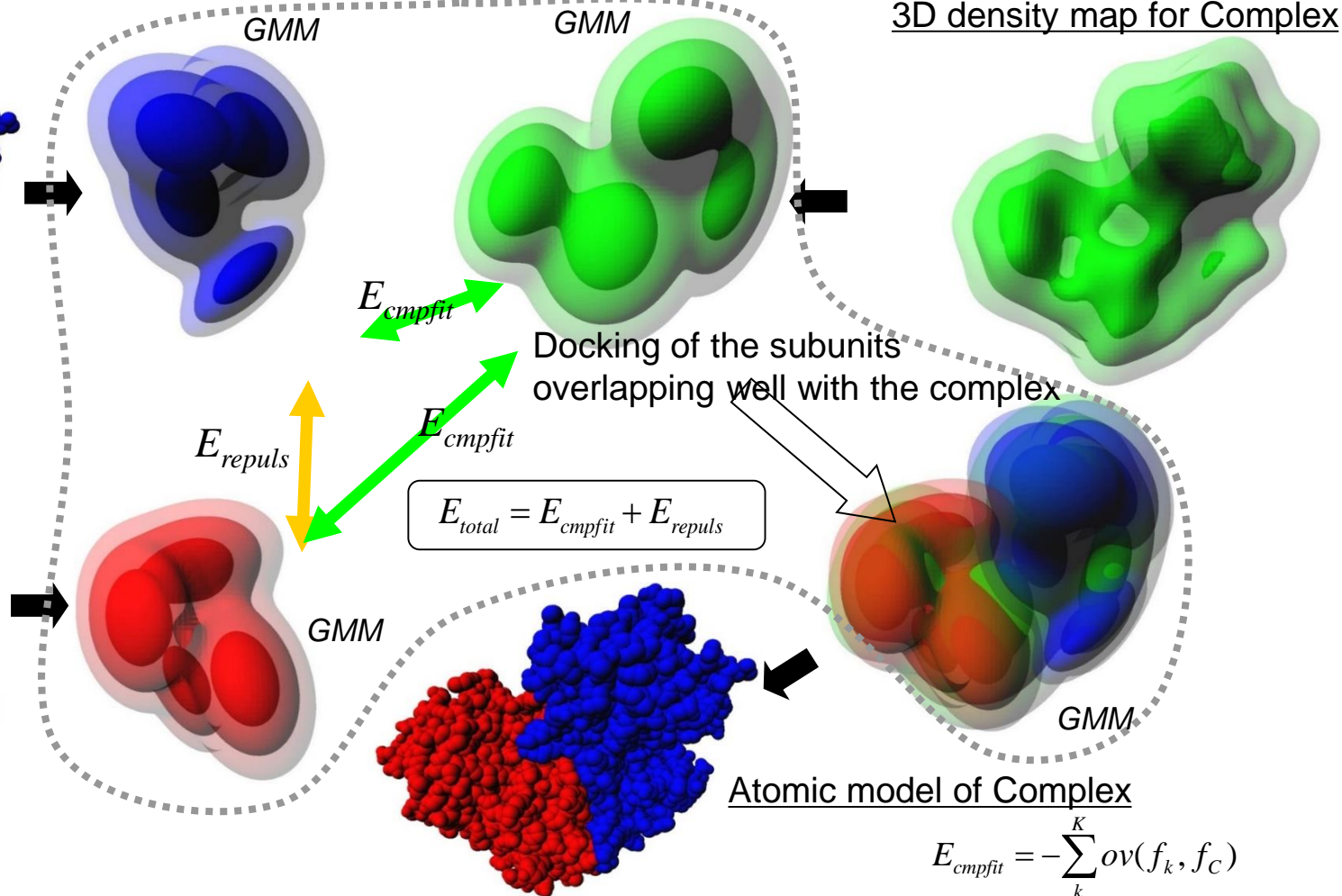
$$f_C(\mathbf{r}) = \sum_i^N w_i \phi_i(\mathbf{r})$$

3D density map for Complex

Atomic model for Subunit 2

Cyclin-dependent kinase2 (1finA)

Cyclin A (1finB)



$$E_{total} = E_{cmpfit} + E_{repuls}$$

Atomic model of Complex

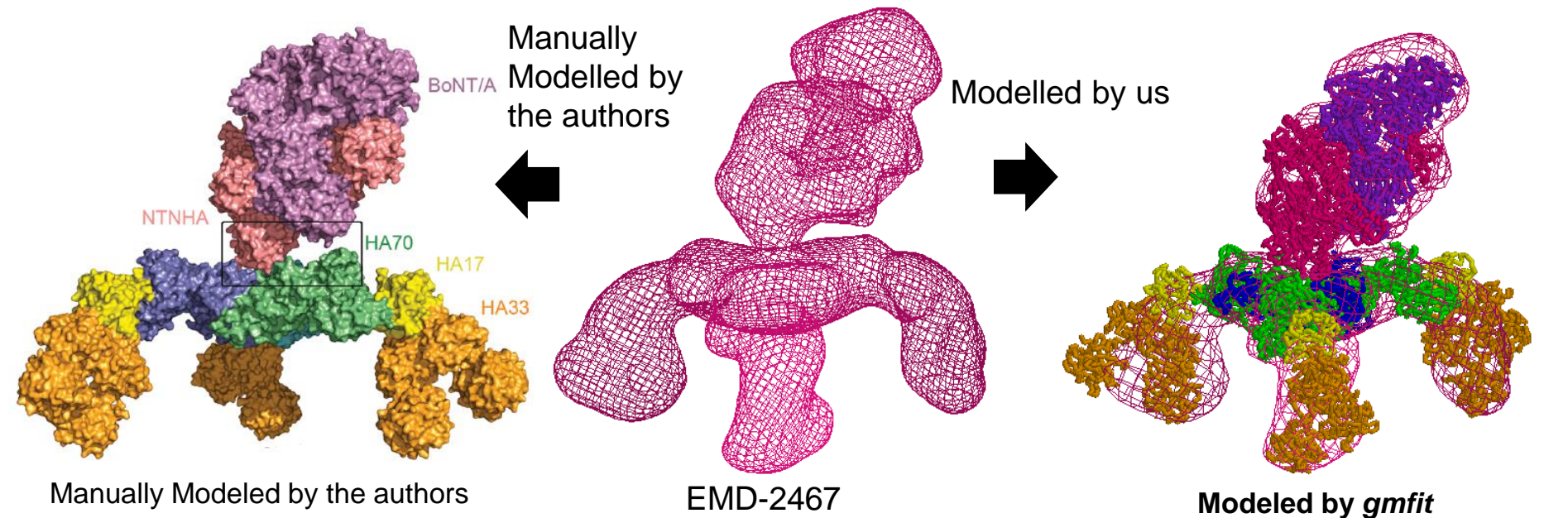
Overlap between two gdfs $ov(\phi_A, \phi_B)$ can be obtained analytically.

$$ov(\phi_A, \phi_B) = \int_{-\infty}^{\infty} \phi_A(\mathbf{r}) \phi_B(\mathbf{r}) d\mathbf{r} = \frac{1}{(2\pi)^{3/2} |\Sigma_A + \Sigma_B|^{1/2}} \exp\left[-\frac{1}{2} (\boldsymbol{\mu}_A - \boldsymbol{\mu}_B)^T (\Sigma_A + \Sigma_B)^{-1} (\boldsymbol{\mu}_A - \boldsymbol{\mu}_B)\right]$$

$$E_{cmpfit} = -\sum_k^K ov(f_k, f_C)$$

$$E_{repuls} = \sum_k^K \sum_{k'>k}^K ov(f_k, f_{k'})$$

Fitting Test: Botulinum Toxin (emd-2467)



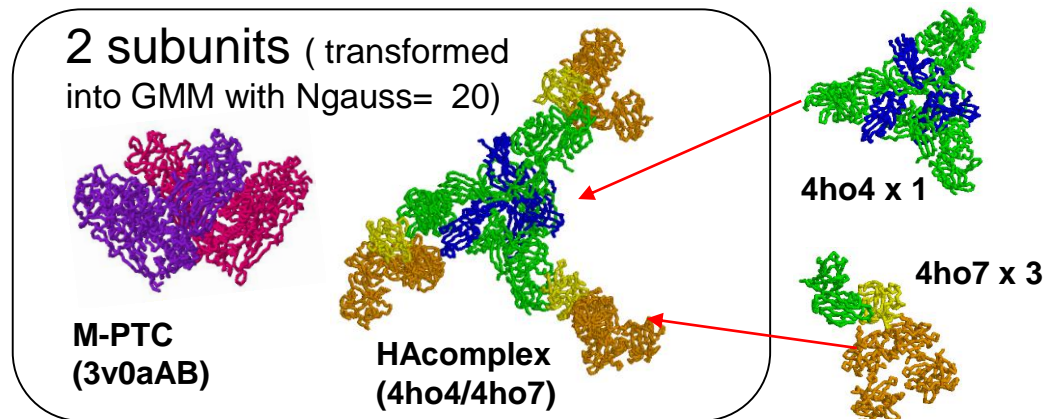
Letet al, (2013), Structure of a Bimodular Botulinum Neurotoxin Complex Provides Insights into Its Oral Toxicity. *PLOS PATHOGENS*, Volume 9, Issue 10, e1003690

Prediction with the **Segmentation-Fitting** initialization. NI=1000, NS=100, 0.779089 **Ngauss for complex = 20**
Ngauss for subunit = 20

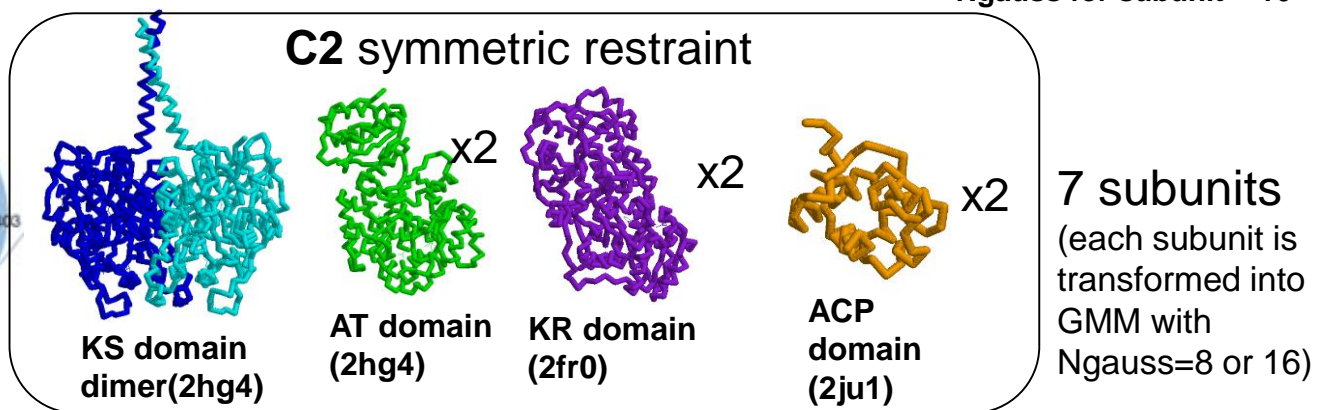
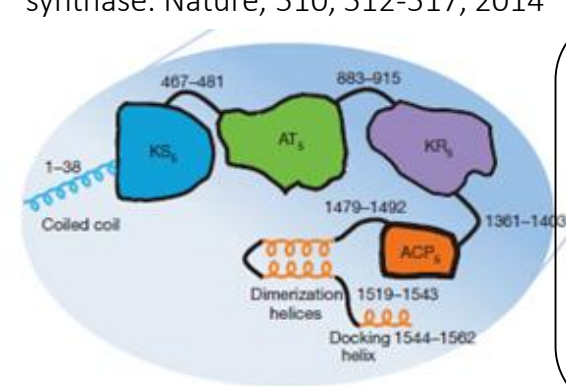
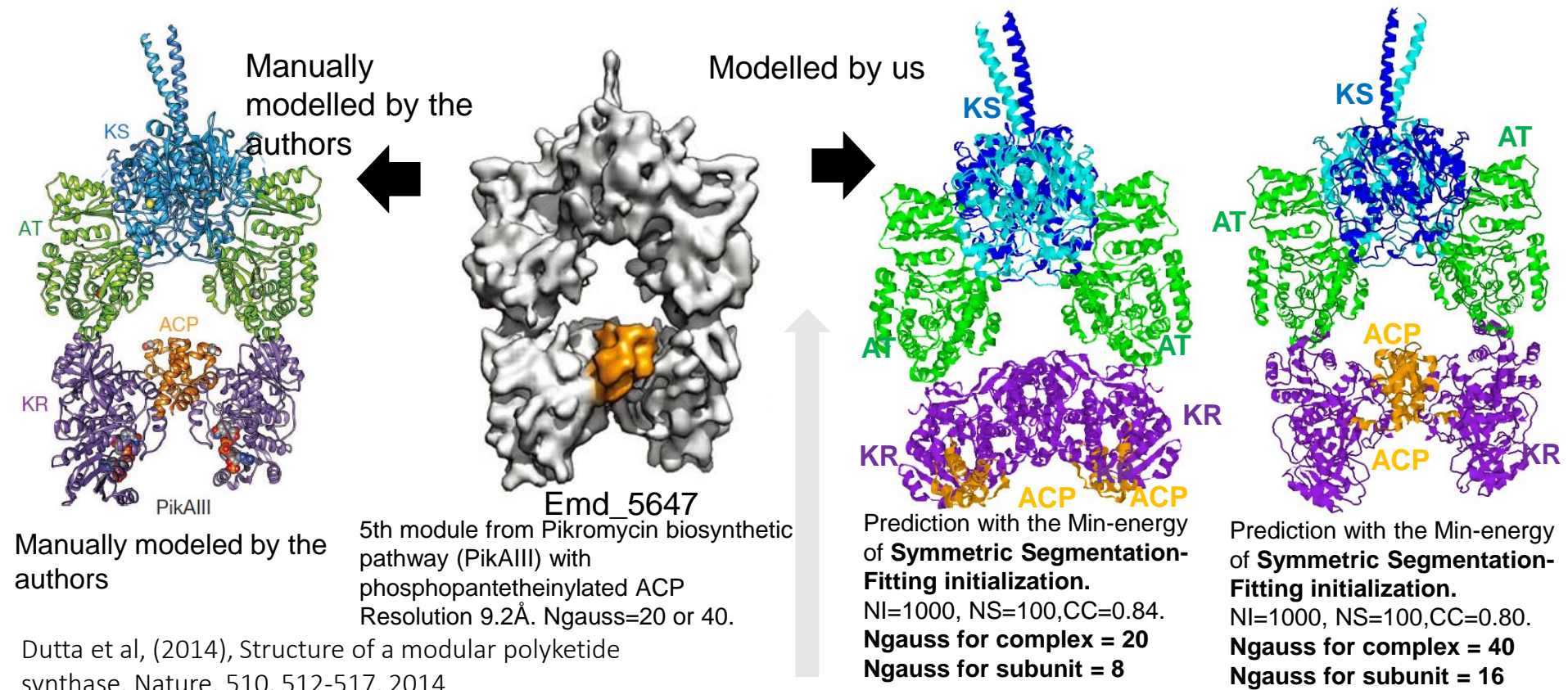
Composed of 14 chains:

M-PTC: BoNT x 1, NTN-HA x 1

HAcomplex: HA70 x 3, HA17 x 3, HA33 x 6

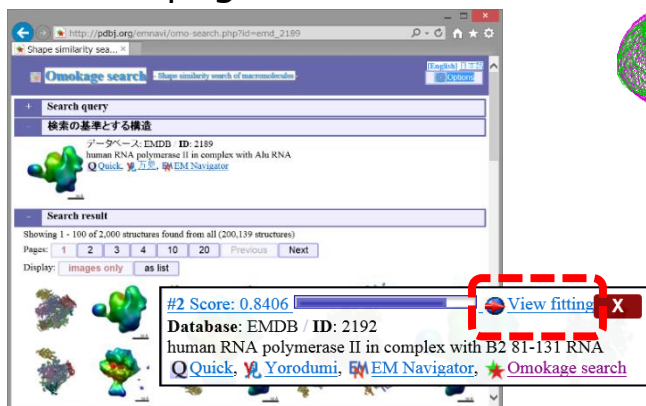


Fitting Test: polyketide synthase (emd-5647)

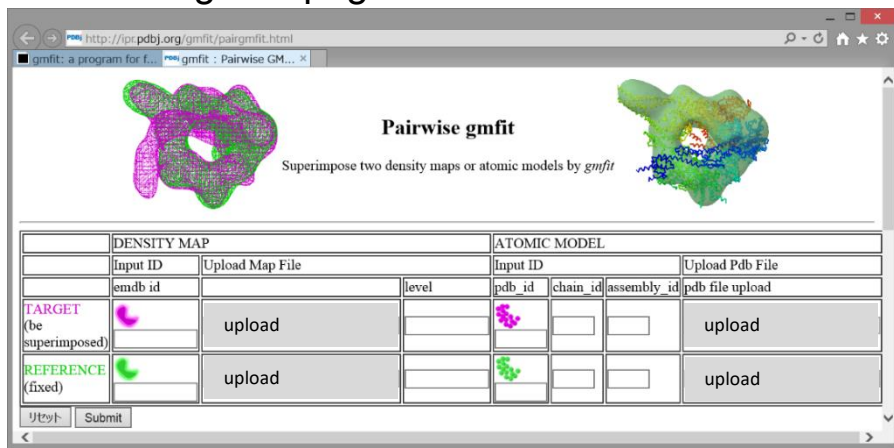


gmfit のWEBサービス

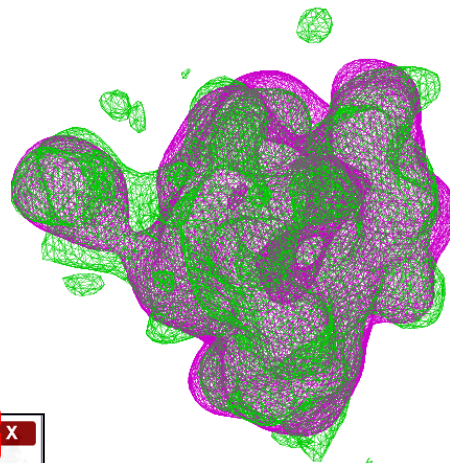
Fitting assemblies can be obtained by **gmfit** program by clicking the link "**View fitting**" in the **Omokage search** page.



If you googled by "**gmfit**", you can find the "Pairwise gmfit" page.



You can input two **emdb_id / pdb_id** or **upload your own map / pdb files** for the fitting.



Map vs. map

EMDB 2190

EMDB 1883

(RNA polymerase II)

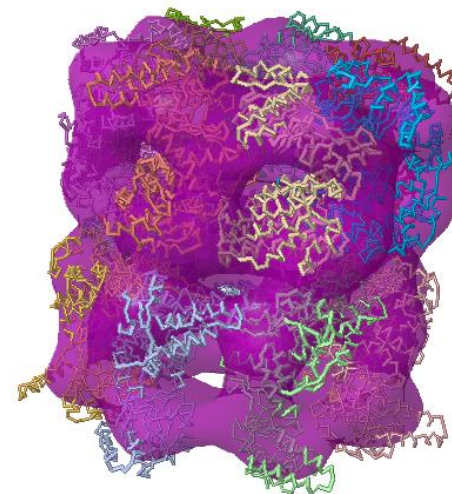


Map vs. atomic model

EMDB 1095

PDB 4aas

(GroEL)

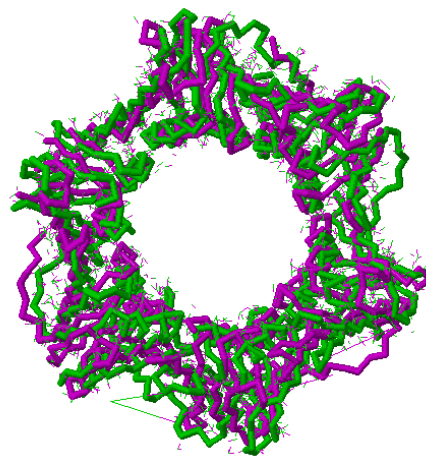


Atomic model vs. atomic model

PDB 3ifv

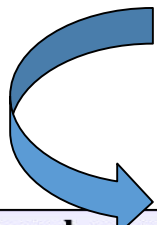
PDB 4k3l

(DNA clamps, dimer vs. trimer)



「Omokage検索」のサービス

Omokage search
1-2 min



Omokage search - Shape similarity search of macromolecules - [English] 日本語 menu

Search query: Query is EMDB-2190.

Registered data structure in databanks

Upload your original/modified data

Query structure data:

ID of EMDB, PDB or SASBDB: **e2190** Search ?

Database: EMDB / ID: 2190

Movie Structure viewer

human RNA polymerase II in complex with AluRA RNA

Human RNA polymerase II

As the query, **EMDB_ID**, **PDB_ID** and **SASBDB_ID** are available.

Users can upload their own model or map file as query.

Search result

Showing 1 - 100 of 2,000 structures found from all (212,387 structures)

Pages: 1 2 3 4 10 20 Previous Next

Display: images only as list download CSV file (for)

Re-rank top 50 data by *gmfit* Corr. Coeff. (beta) Submit

EMDB-2191 human RNA polymerase II in complex with B2 RNA

PDB-5c44 transcribing RNA Polymerase II complex

PDB-4bbr RNA polymerase II-TFIIB complex

EMDB-2192 human RNA polymerase II in complex with B2 81-131 RNA

PDB-4a3i RNA Polymerase II binary complex with DNA

PDB-4bxx RNA polymerase II complexes with Bye1, a chromatin-binding PHF3 DIDO homologue

PDB-1y1w Complete RNA Polymerase II elongation complex

PDB-2b8k 12-subunit RNA Polymerase II

PDB-3h0g RNA Polymerase II from *Schizosaccharomyces pombe*

PDB-3j0k RNA polymerase II within the human VP16-Mediator-pol II-TFIIF assembly

Many maps and atomic models of RNA polymerase II are found.

Pairwise fitt

Map vs Map comparison

2190 by *gmfit*

Current [JAVA] Char [HTML5]

gmfit

id: 2190 : 0.977

Pairwise fitt

Map vs Atom comparison

2190 by *gmfit*

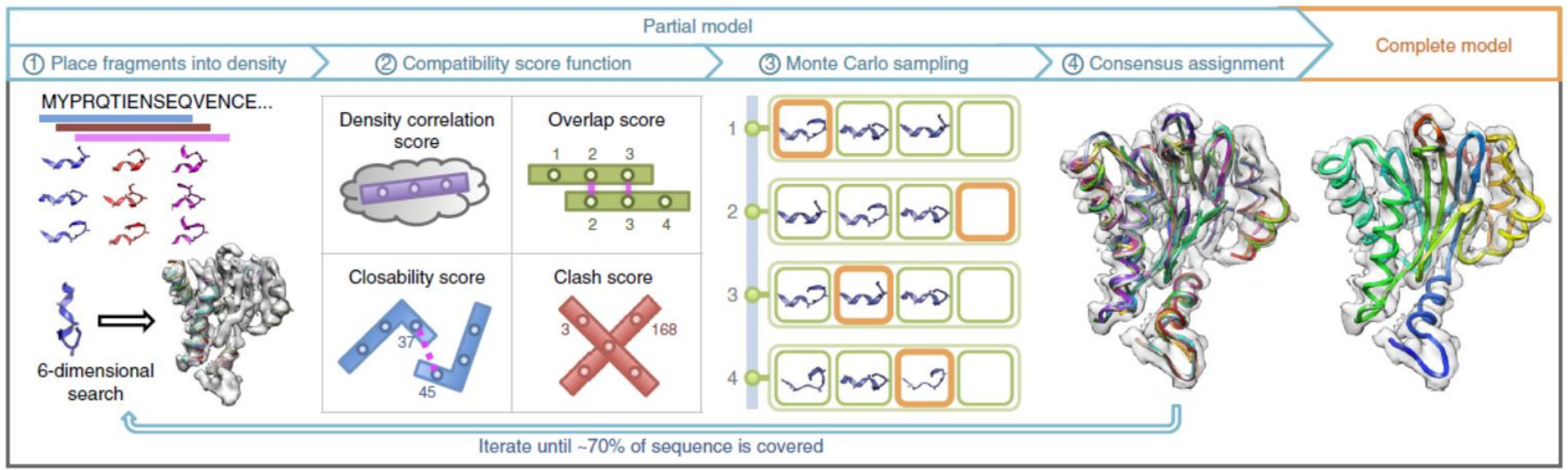
Current [JAVA] C to:[HTML5]

gmfit

id: 2190 : 0.902

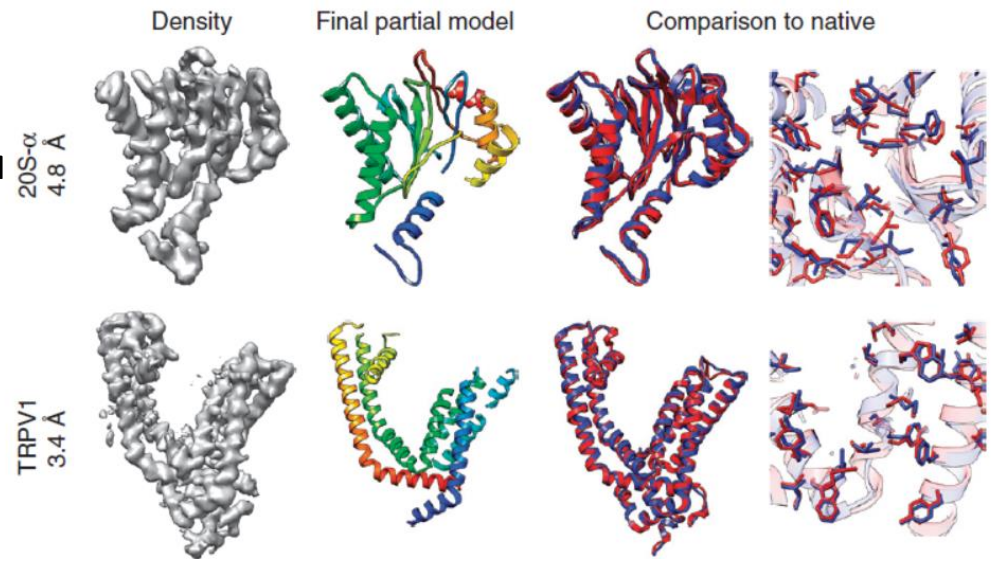
デノボモデリング: ROSETTA

ROSETTAを用いたモデリングの手続き



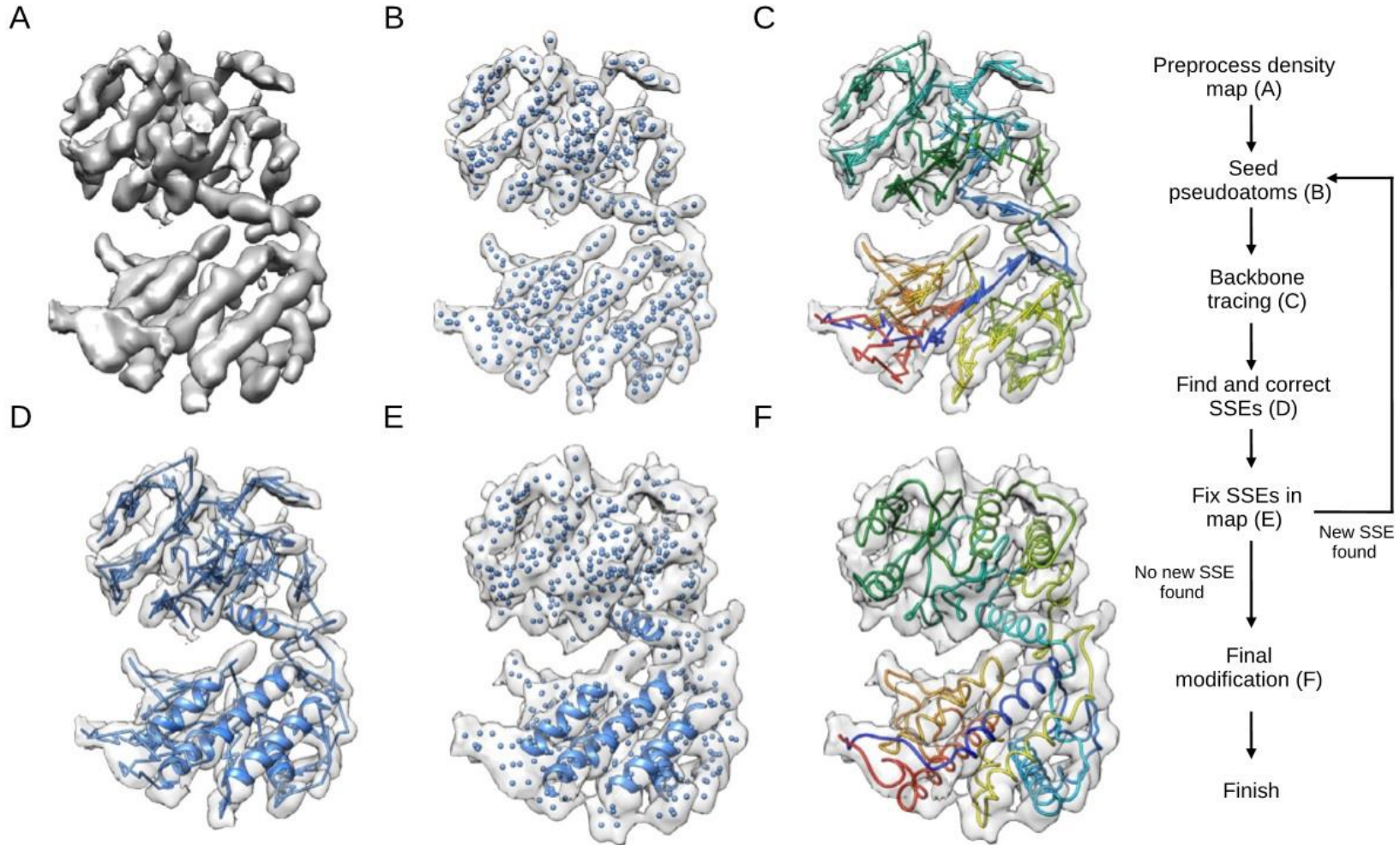
Ray Yu-Ruei Wang Mikhail Kudryashev Xueming Li
 Edward H Egelman Marek Basler Yifan Cheng David
 Baker Frank DiMaio. De novo protein structure
 determination from near-atomic-resolution cryo-EM
 maps, Nature Methods 12, 335–338 (2015)

密度マップを、ドメイン程度の
 小さい領域に切り出す(セグメン
 テーション)が必須



デノボモデリング: PathWalker

アミノ酸に相当する点群を生成、それをできるだけタンパク質らしく鎖で結んでいく方法

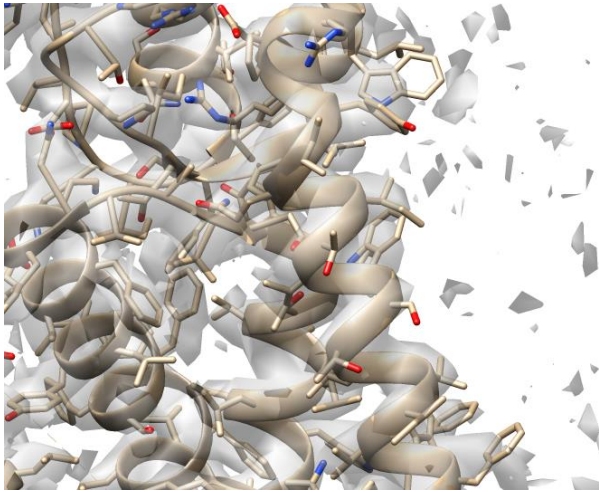


密度マップを、ドメイン程度の小さい領域に切り出す(セグメンテーション)が必須

Muyuan Chen^{a, c}, Philip R. Baldwin^{b, c}, Steven J. Ludtke^c, Matthew L. Baker^c, De Novo modeling in cryo-EM density maps with Pathwalking, *J Str Biol*, 196,289–298 (2016)

電顕による原子モデルを使うための 注意点

1. 原子モデルの元となった密度マップを重ねて表示し、妥当なモデリングか確認する必要。(PDBに対応するEMDB IDが記載されている。表示はUCSF Chimeraを推奨。[tool]->[Volume Data]->[Volume Viewer]で簡単にしきい値を変えて表示できる。)



原子モデルのどの部分が、きちんと対応する密度があるか、また、原子モデルが置いていないマップの部分があるかなどが確認できる。

EMD-3061, PDB-5a63

2. X線結晶解析に比べて、一つのマップの部分において実質的な解像度が異なる場合がある。例えば、分子表面と分子の内側、膜貫通領域と可溶性領域で様相が異なることに注意。

3. 3~4 Åの解像度のマップでは、おおまかなアミノ酸部位、結合部位などはわかるが、側鎖や結合分子の配座はそれほど正確ではないことに注意。