

# Takanori Nakane, Dari Kimanius, Erik Lindahl, and Sjors HW Scheres. 2018. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. eLife 7, (June 2018), e36861

## Abstract

- 提出RELION中的multi-body refinement;
- 目标是解决异质性问题;
- 建模:
  - 使用多个(用户自定义数量)刚体的组合来刻画异质性;
  - 使用分别的focused refinement和裁剪来计算各个刚体的参数;
  - 使用PCA等方法分析最重要的运动;
- 在两套测试数据ribosome和spliceosome上进行了测试。

## Introduction

异质性是蛋白结构的经典问题。在本文以前,作者列举了如下一些处理方案:

- Discrete classification;
- Normal mode analysis;
- Manifold embedding / learning。

作者提出使用multi-body来建模异质性,因为在一些生物结构中发现了这种现象。例如: yeast mitochondrial ribosome, Plasmodium falciparum cytoplasmic ribosome。使用multi-body通常会对每个body进行focused/masked refinement, 也会结合partial signal subtraction来使用。

作者提到这种方法的一个弊端是每个body需要足够大, 否则无法精确对齐。同时提到最近有一个WarpCraft工作, 结合NMA来生成伪原子模型并约束不同body之间的motion从而可以解析出更小的区域的工作。

## Method

建模:

$$X_i = \text{CTF}_i \left( \sum_{b=1}^B \mathbf{P}_{\phi_b} V_b \right) + N_i$$

其中

- $X_i$ 为观测的二维投影；
- $\text{CTF}_i$ 为 $X_i$ 的CTF参数；
- $B$ 为刚体数量；
- $V_b$ 为第 $b$ 个刚体的密度图；
- $\phi_b$ 为第 $b$ 的刚体（相对于整体）的位置参数；
- $\mathbf{P}_{\phi_b}$ 为投影算子；
- $N_i$ 为噪声。

算法：Figure 1。逐个刚体迭代。

实现：

- `relion_refine` 或 `relion_refine_mpi` 中给出multi-body选项；
- 刚体由输入的所在区域的mask给出；
- mask有一定的soft edge，可以overlap，但是refine第 $b$ 个刚体时排除掉第 $b$ 个以外部分的信号；
- `relion_flex_analyse` 中给出异质性的分析；
- 使用PCA对相对运动（包含3个Euler角和3个平移）进行分析，找到主成分后插值为一条 $M$ 个点的轨迹从而获得动态。

## Results

- Ribosome bound to the drug emetine
  - EMPIAR-10028；
  - 刚体：LSU + SSU + head；
  - 分辨率：总体3.2A，LSU 3.2A => 3.1A，SSU 3.3A => 3.2A，head 4.0A => 3.7A；
  - B-因子：-61A<sup>2</sup>不变。
  - 动态：Figure 2，Video 1，Video 2。
- Spliceosomal B-complex from yeast
  - EMPIAR-10180；
  - 刚体：core + foot + helicase + SF3b；
  - 分辨率：总体4.3A，core 3.9A => 3.8A，foot 4.2A => 4.0A，helicase 4.6A => 4.5A，SF3b 9.2A => 5.1A => 4.4A；

## Discussion

- Multi-body refinement的第一个好处是可以解析异质性存在时不同区域的结构；
  - 但是作者提到了如果使用focused refinement，则spliceosome的SF3b可以解析到3.9Å，明显好于multi-body refinement的4.4Å。
  - 不过作者断言，multi-body refinement更自动化，不需要更多expertise，并且进行multi-body refinement之后仍然可以继续focused alignment；
- Multi-body refinement的另外一个好处是可以用PCA分析动态，生成movie。
  - 但是作者同样指出这些movie的中间帧并不是“well-defined”的，只是may be useful in exploring biologically relevant molecular motions。
- Multi-body refinement的一个缺陷是可能无法处理小蛋白，作者指出本文中最小的单元也有290kDa。
  - 作者指出WarpCraft有可能是一个有潜力的方向。