



Acknowledgments

- This research was performed as a JHPCN exploratory theme (Project: EX23504 and EX24505).
- We are grateful for the support from the Information Technology Center at Nagoya University via "Collaborative Research Project on Computer Science with High-Performance Computing".
- Development and experimental computation were partially performed at the Research Center for Computational Science, Okazaki, Japan (Project: 22-IMS-C181).
- The presenter is grateful for the job opportunity by KAKENHI 20H05932 (Transformative Research Areas (A) "Multifaceted Proteins: Expanding and Transformative Protein World").

背景

- タンパク質はアミノ酸が重合した高分子であり、固有の立体構造をもつ。立体構造を知ることは個々のタンパク質の機能を理解する上で重要である。
- 個々のタンパク質の立体構造を観察するだけでなく、さまざまなタンパク質の立体構造を比較できると、さらに理解が深まる。

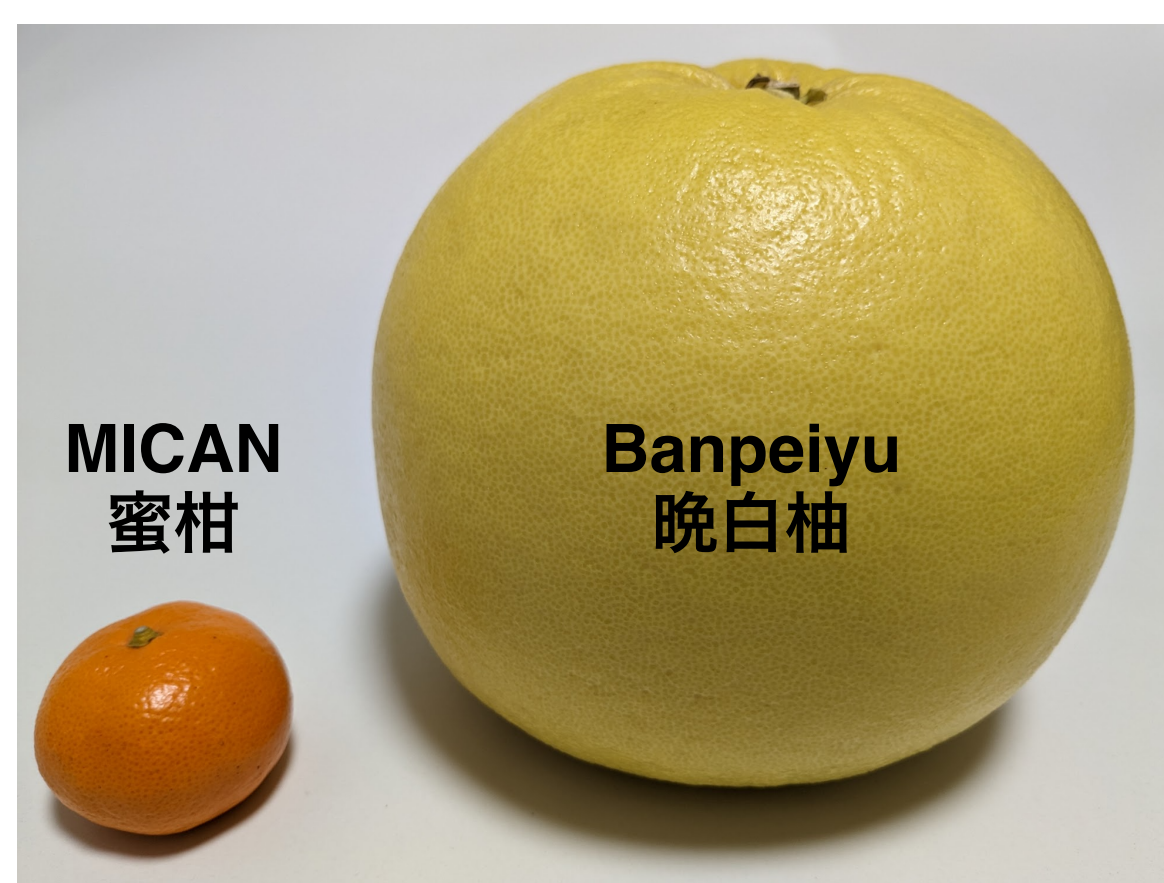
大きな立体構造の登場

- これまで構造比較の対象は100から1,000アミノ酸残基程度の構造を対象に行うことが多かった。
- 近年、低温電子顕微鏡による構造決定手法の進歩により10,000残基以上の巨大な立体構造が日常的に決定されるようになった。

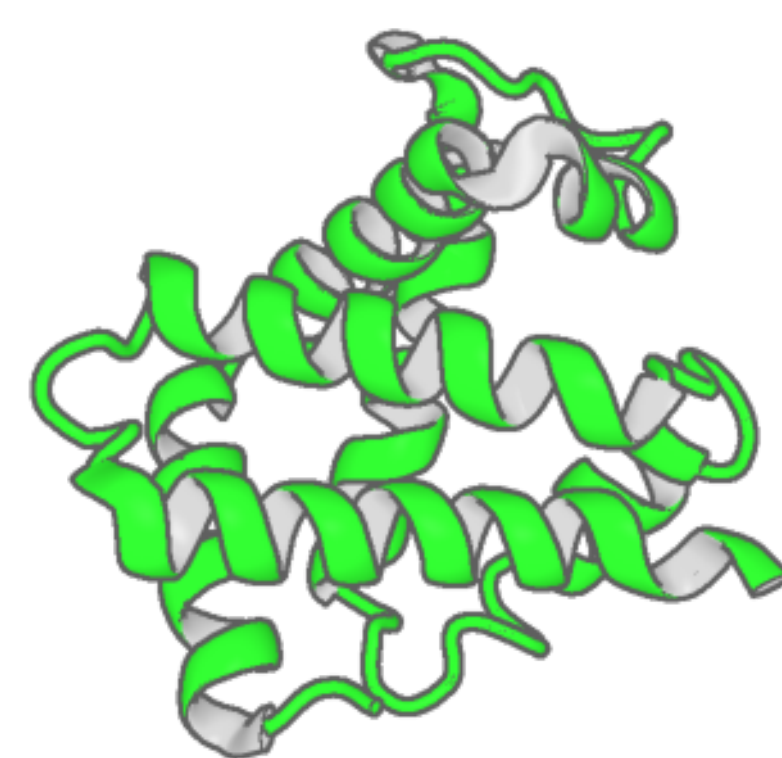
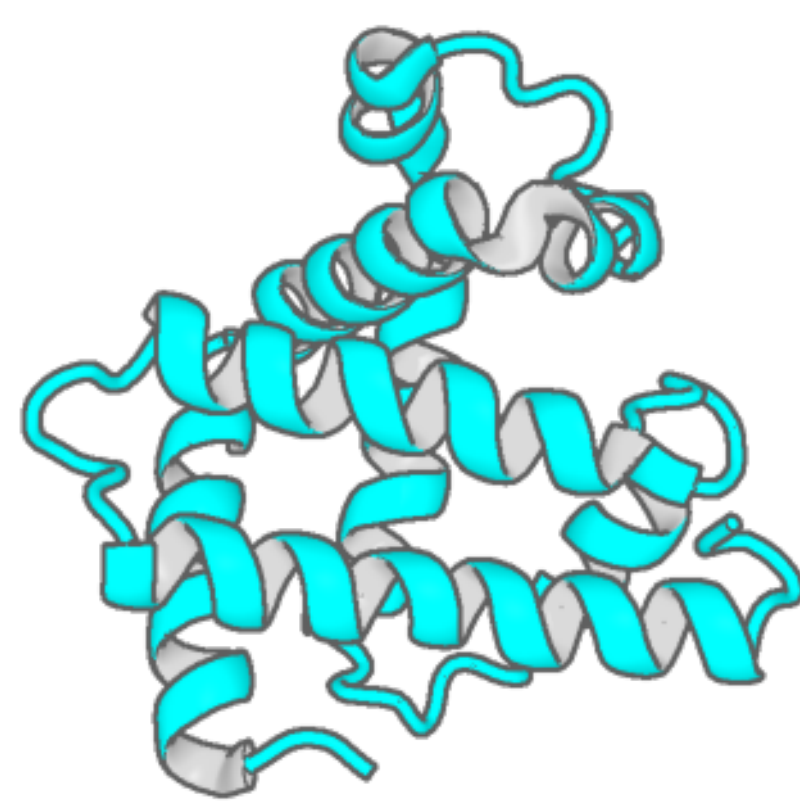
本研究の目的

- 巨大な立体構造に対応した立体構造比較手法を開発し、その性能を評価する。

0. 蜜柑と晩白柚



1. 背景：タンパク質の立体構造と構造比較

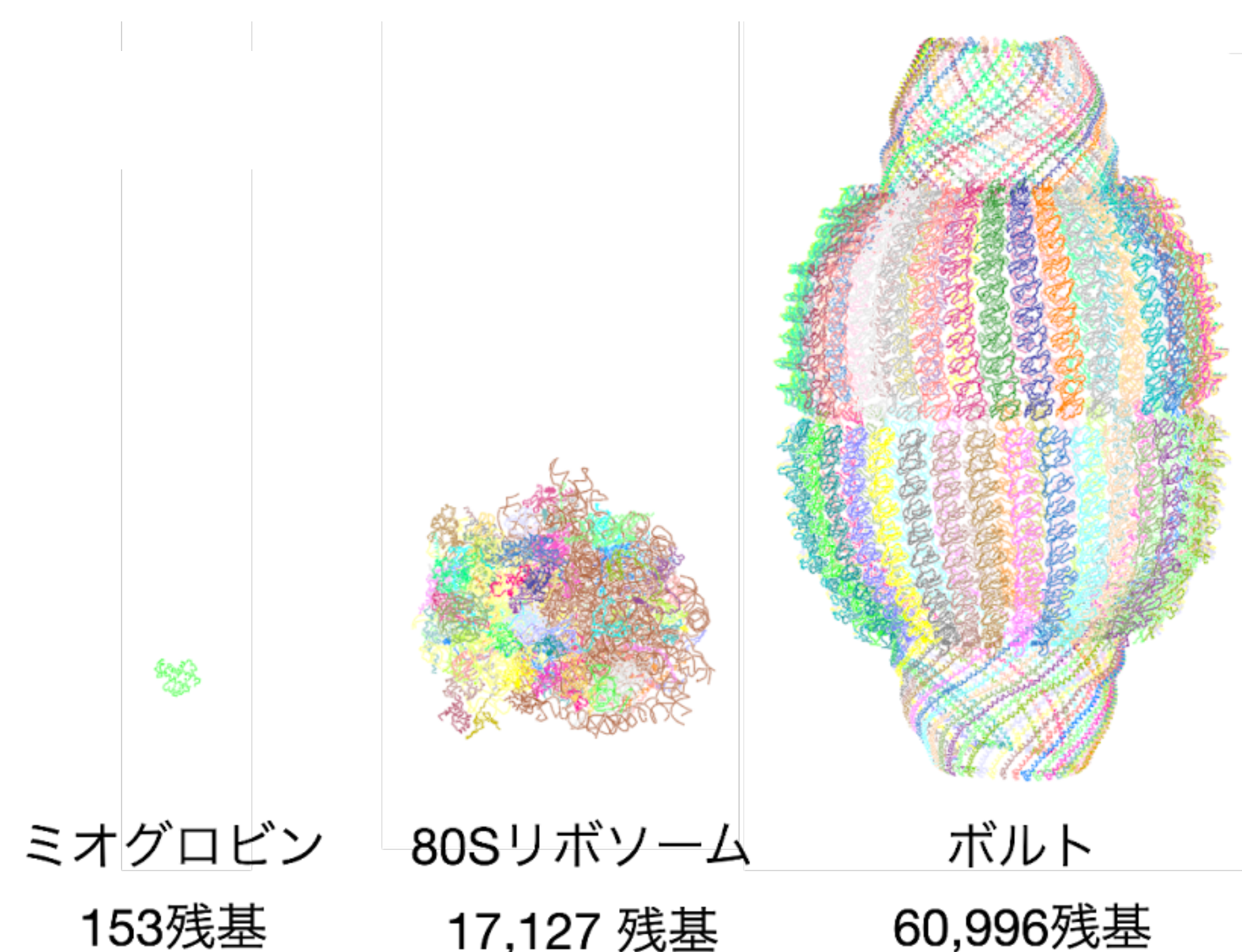


ミオグロビン (1MBN) ヘモグロビン (2DHB)

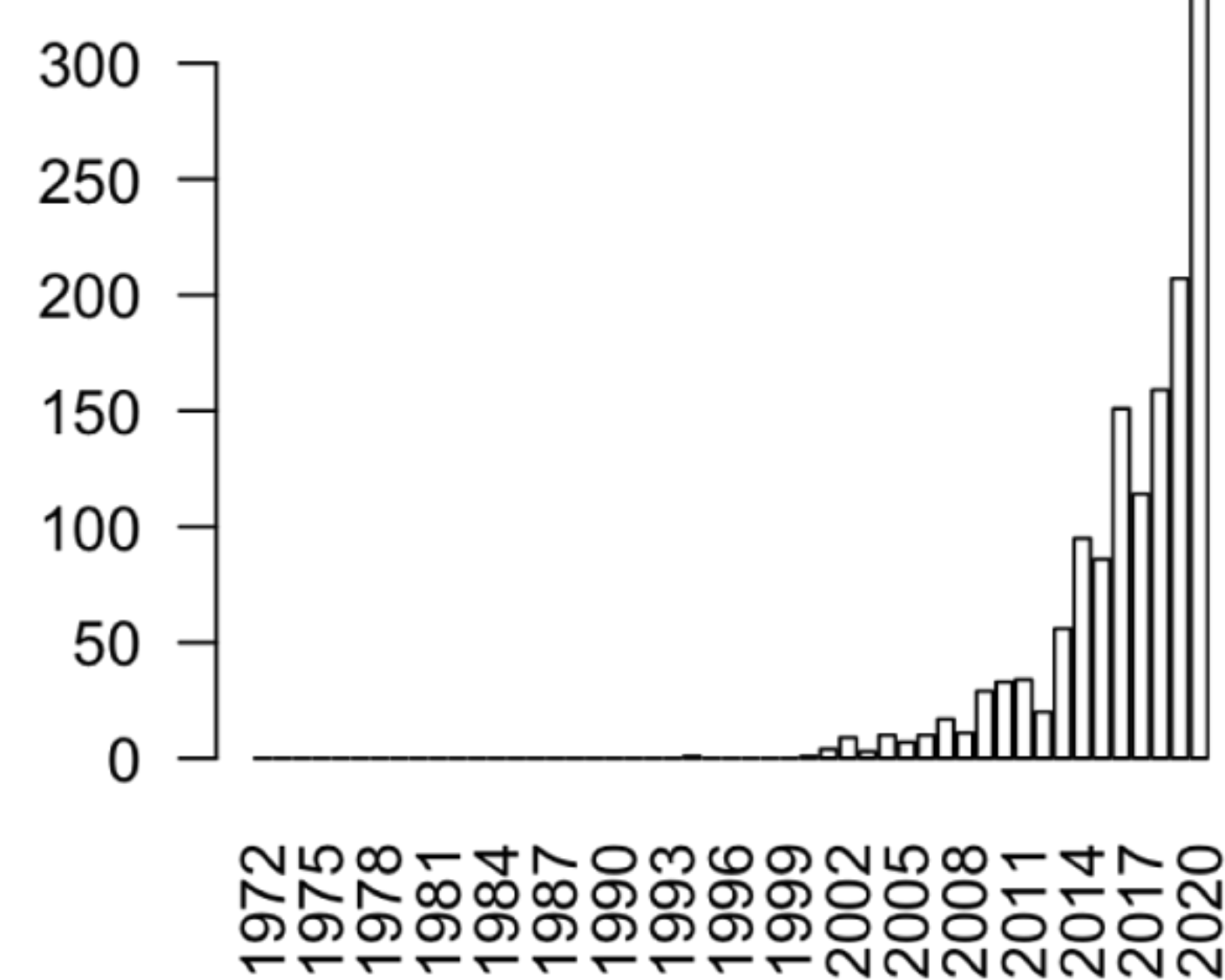
構造比較で何がわかるか？

- 異なるタンパク質同士の比較 -> 機能推定
- 同じタンパク質同士の比較 -> ダイナミクス

2. 巨大複合体の構造比較：巨大な立体構造が急激に増加



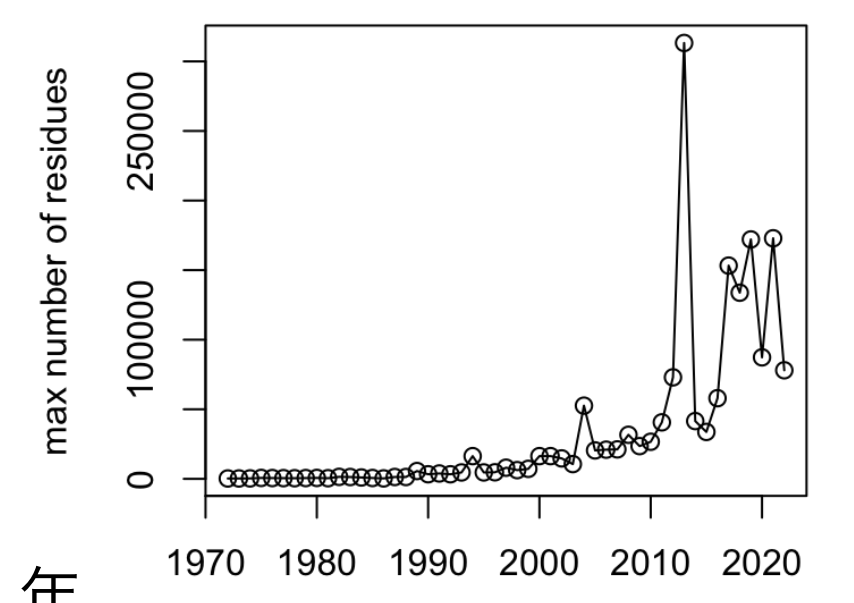
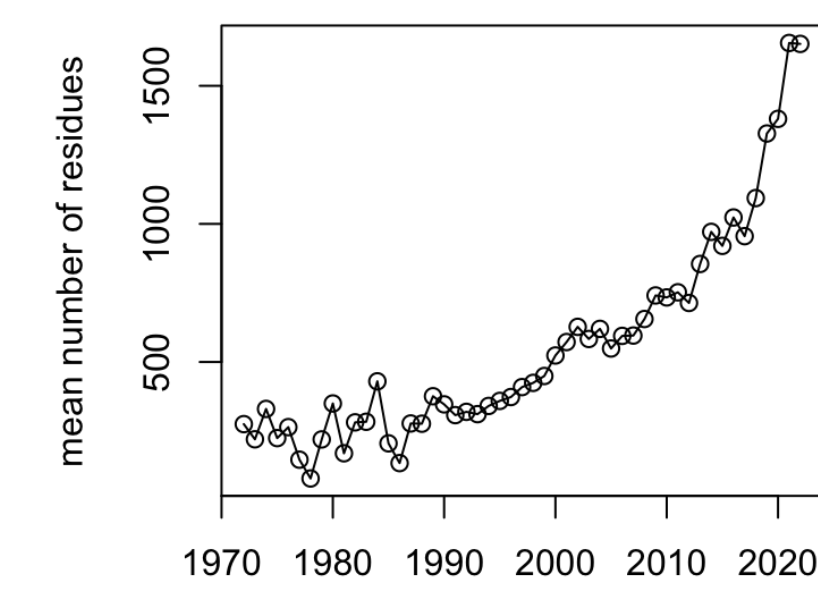
10,000残基以上を持つ巨大構造の年別登録件数推移



1 構造に含まれる残基数

平均

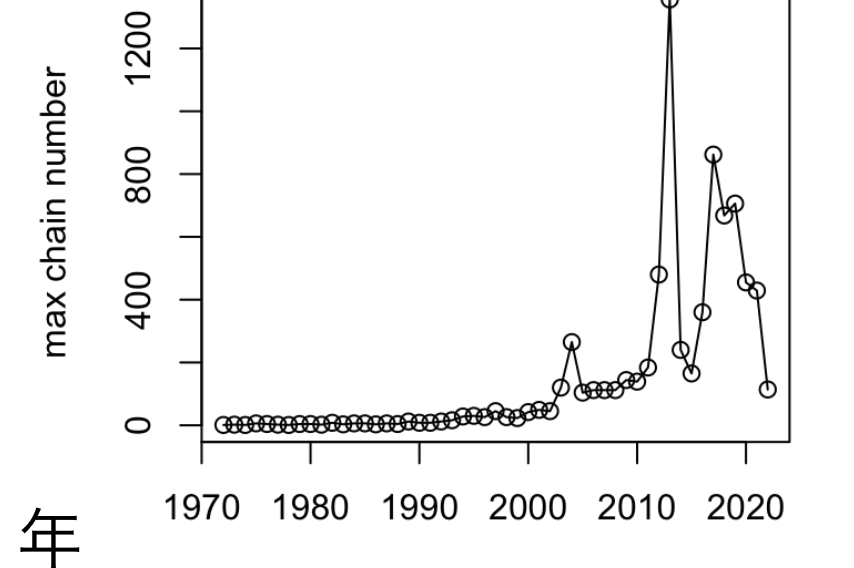
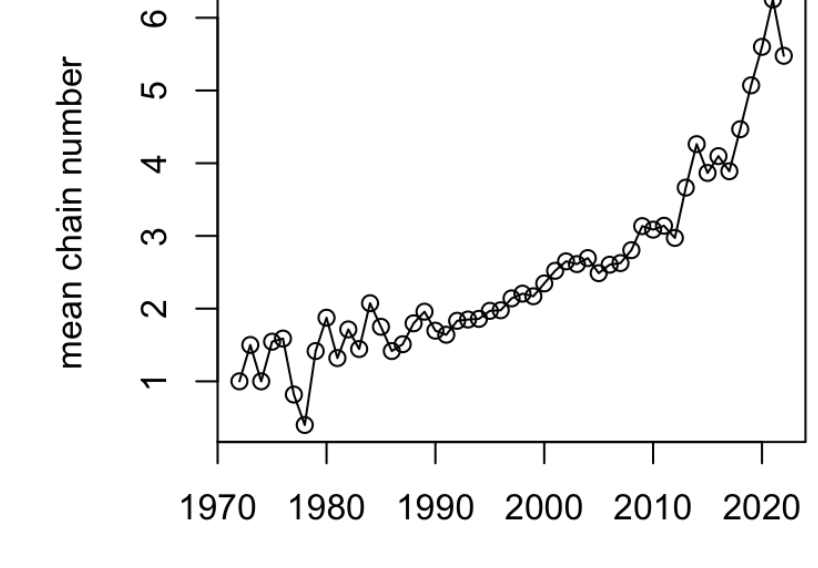
最大値



1 構造に含まれるポリマー鎖数

平均

最大値



3. MICAN-BANPEIYOU (晩白柚) の開発

開発方針

- 比較アルゴリズムは既存のもの利用
- アルゴリズムのボトルネック解消と並列化を実施

MICAN (Minami et al. 2013 2015)

- タンパク質複合体に対するノンシーケンシャル構造比較

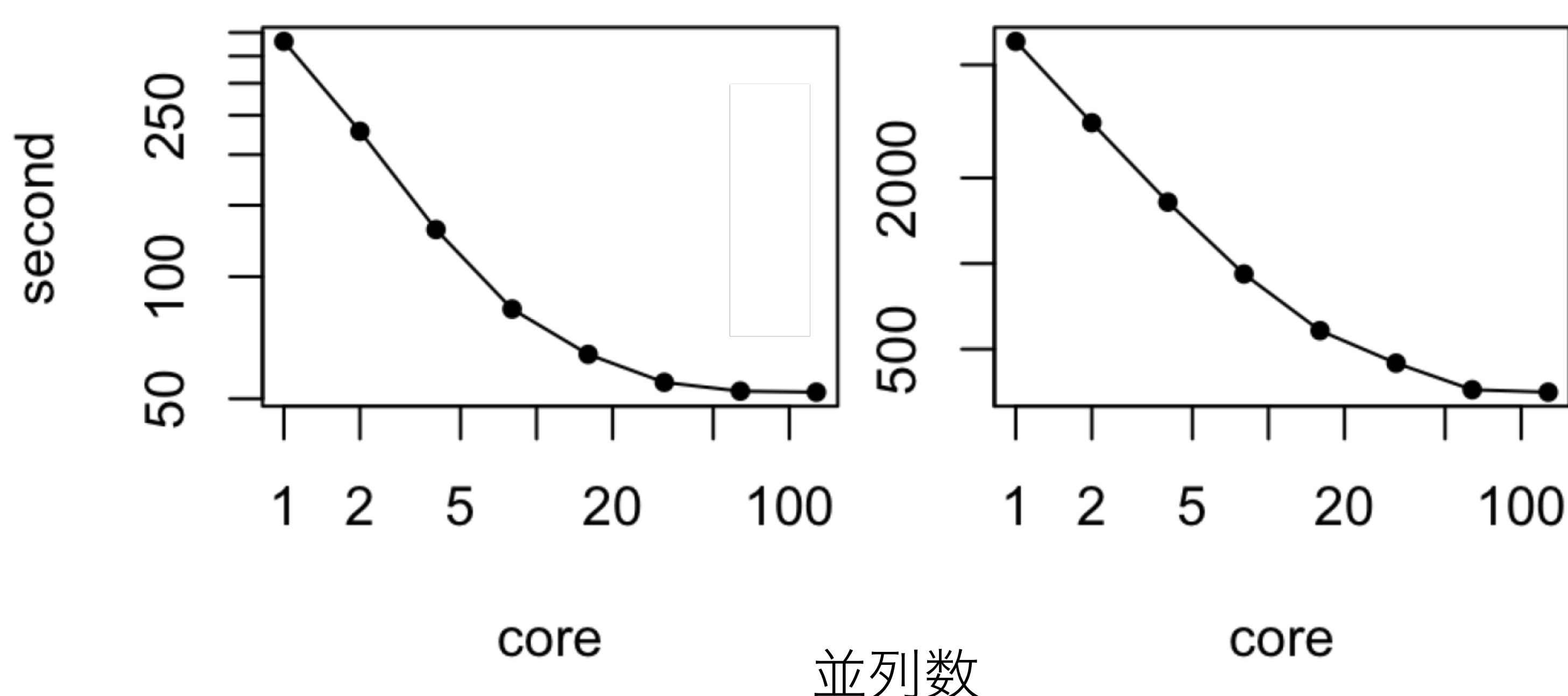
MICAN-Banpeiyou (Sakuma et al. in preparation)

- 核酸・タンパク質複合体に対するノンシーケンシャル構造比較

並列化による速度変化

3,800残基 vs 4,100残基

11,500残基 vs 12,500残基



4. 巨大な構造比較の実例

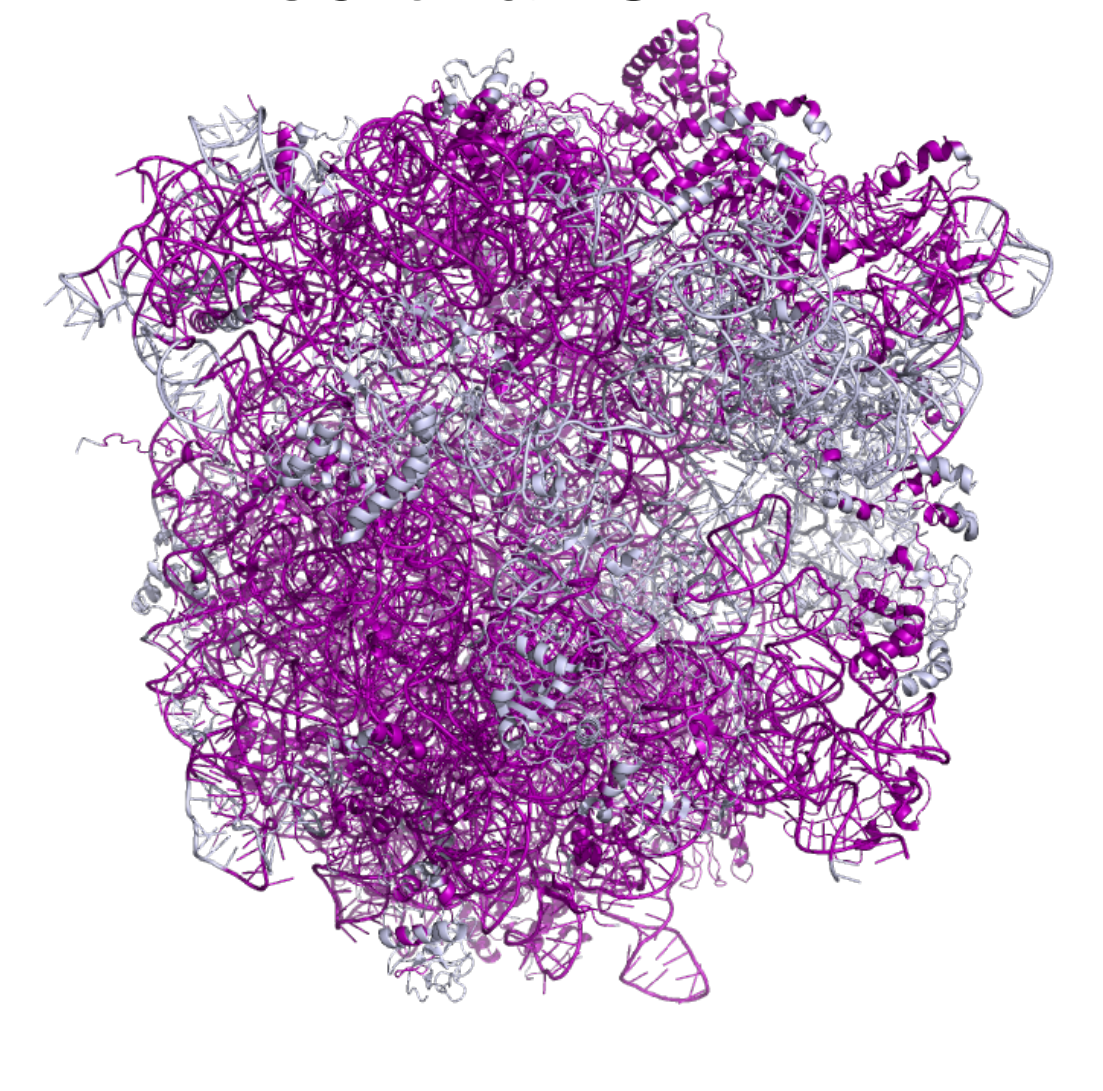
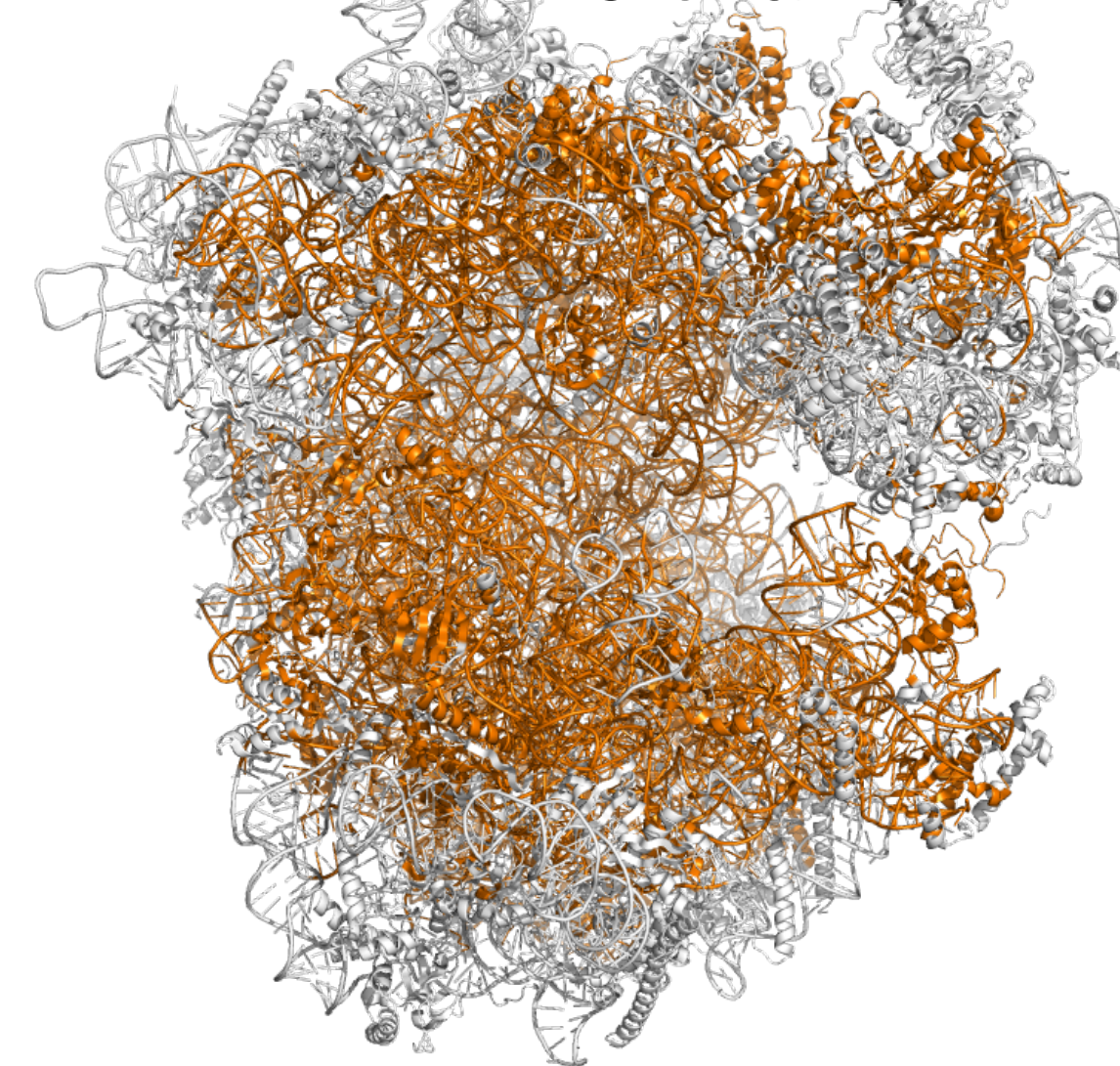
実例：真核リボソーム vs 原核リボソーム

4v88

4v5g

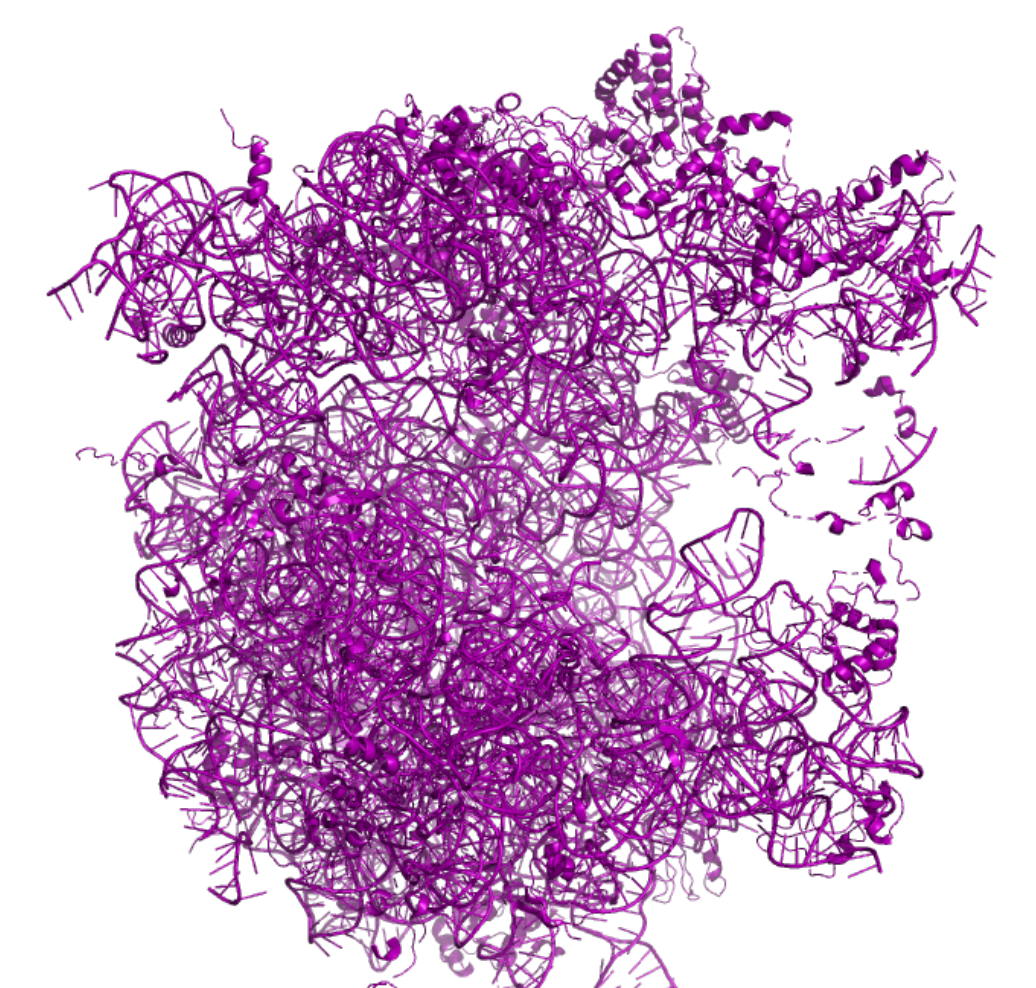
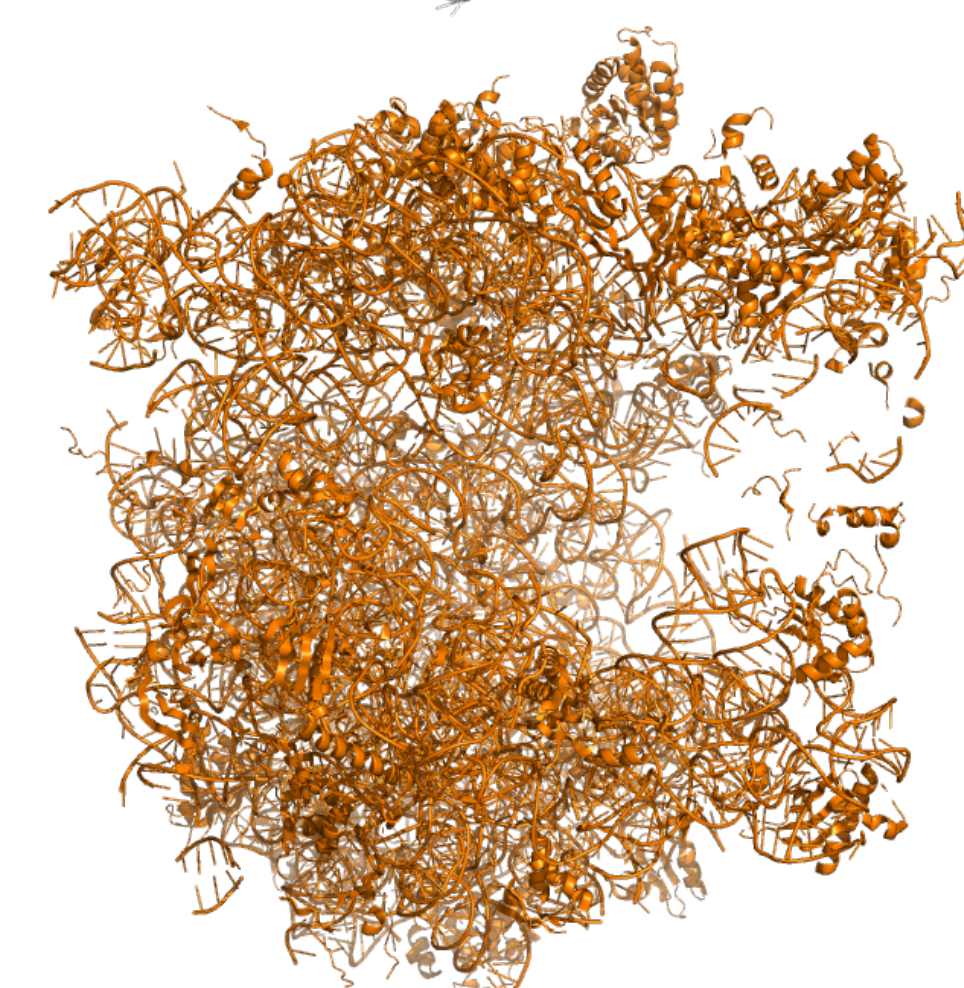
11298 amino-acids
5209 nucleotides
79 chains

6490 amino-acids
4766 nucleotides
59 chains



全体

アライメント部分



TM-score=0.393, Coverage= 46.2% TM-score=0.576, Coverage= 67.7%

sTMscore(Nali) = 0.803

Ca/C3' RMSD = 4.581 Å

Sequence identity = 34.5%

C3score = 0.942