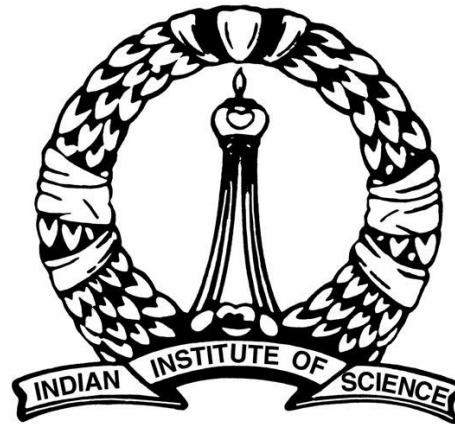


Integrative Structure Determination of Macromolecular Assemblies

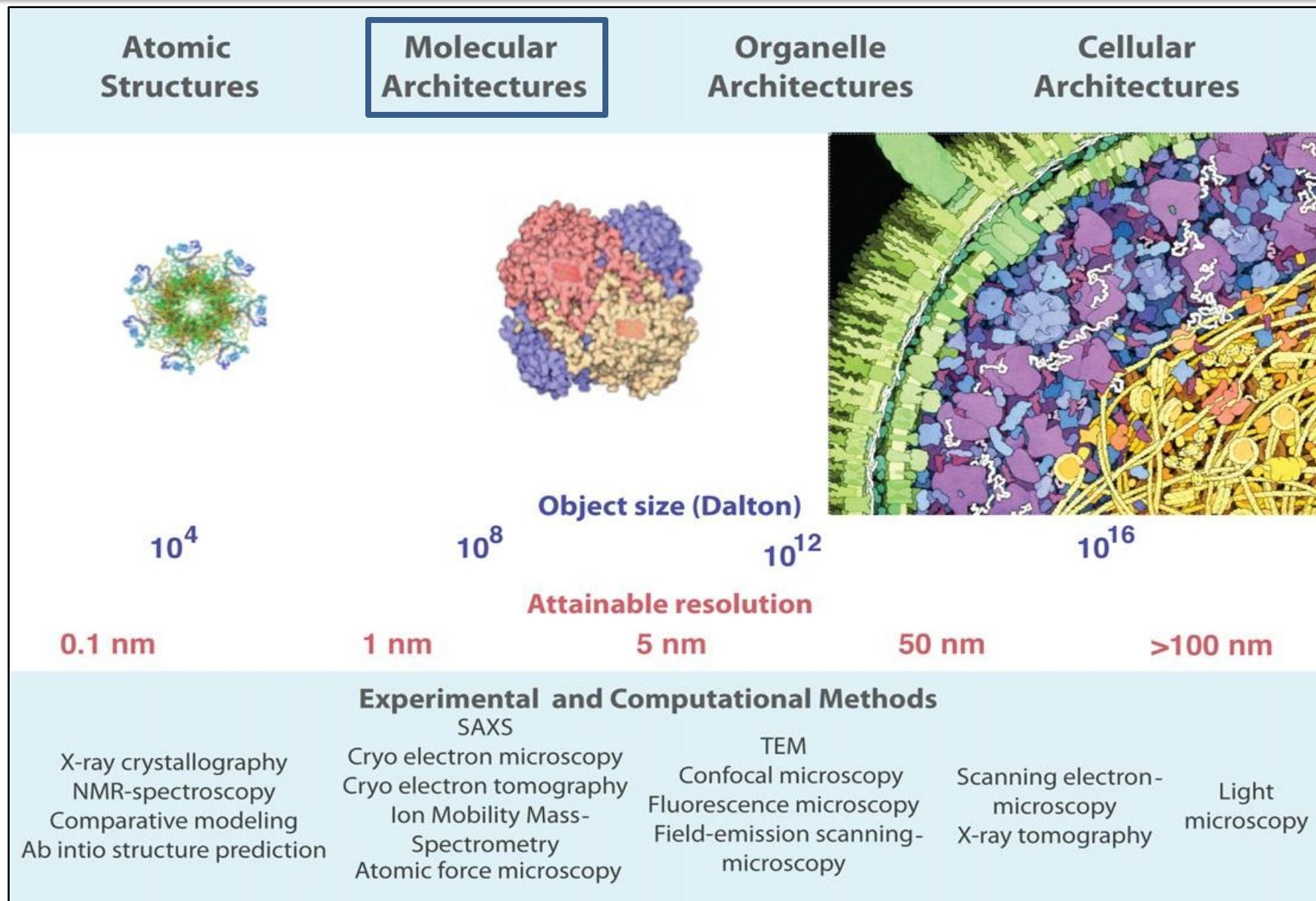


Rakesh Ramachandran

Outline

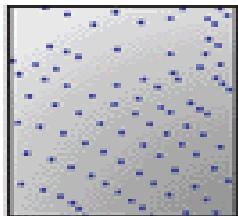
- Introduction
- Cryo-Electron Microscopy based Modeling
- Modeling based on Mass Spectrometry data
- Integrative Structure Determination of Proteasome

Need for Integrative Approach

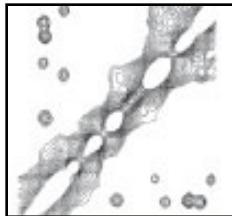


Integrative Structure Determination

High Resolution Information
(Atomic Level)

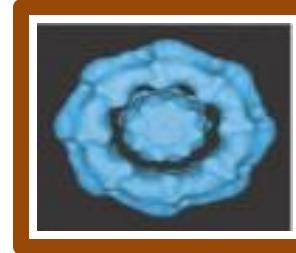


X-ray
Crystallography

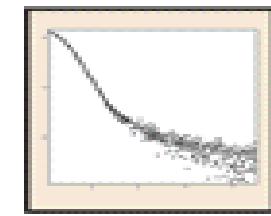


NMR

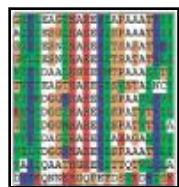
Low Resolution Information
(Gross Shape/Density)



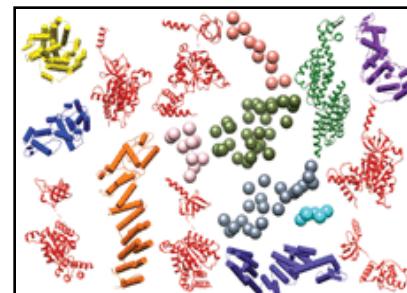
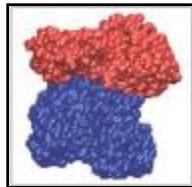
Electron
Microscopy



SAXS

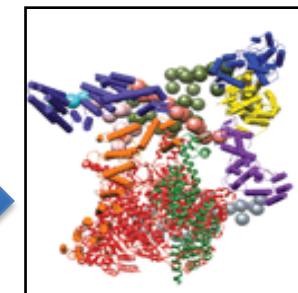


Computational
Approaches

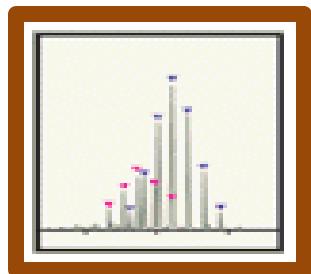


Components and Data at
different resolutions

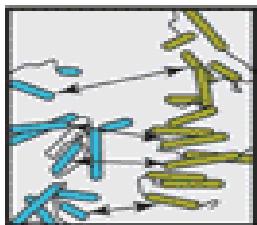
Sampling
and
Analysis



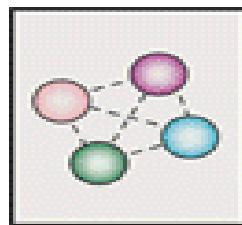
Model



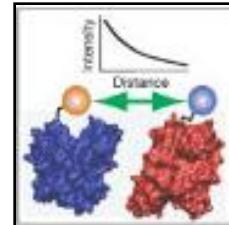
Mass
Spectrometry



Cross-Linking



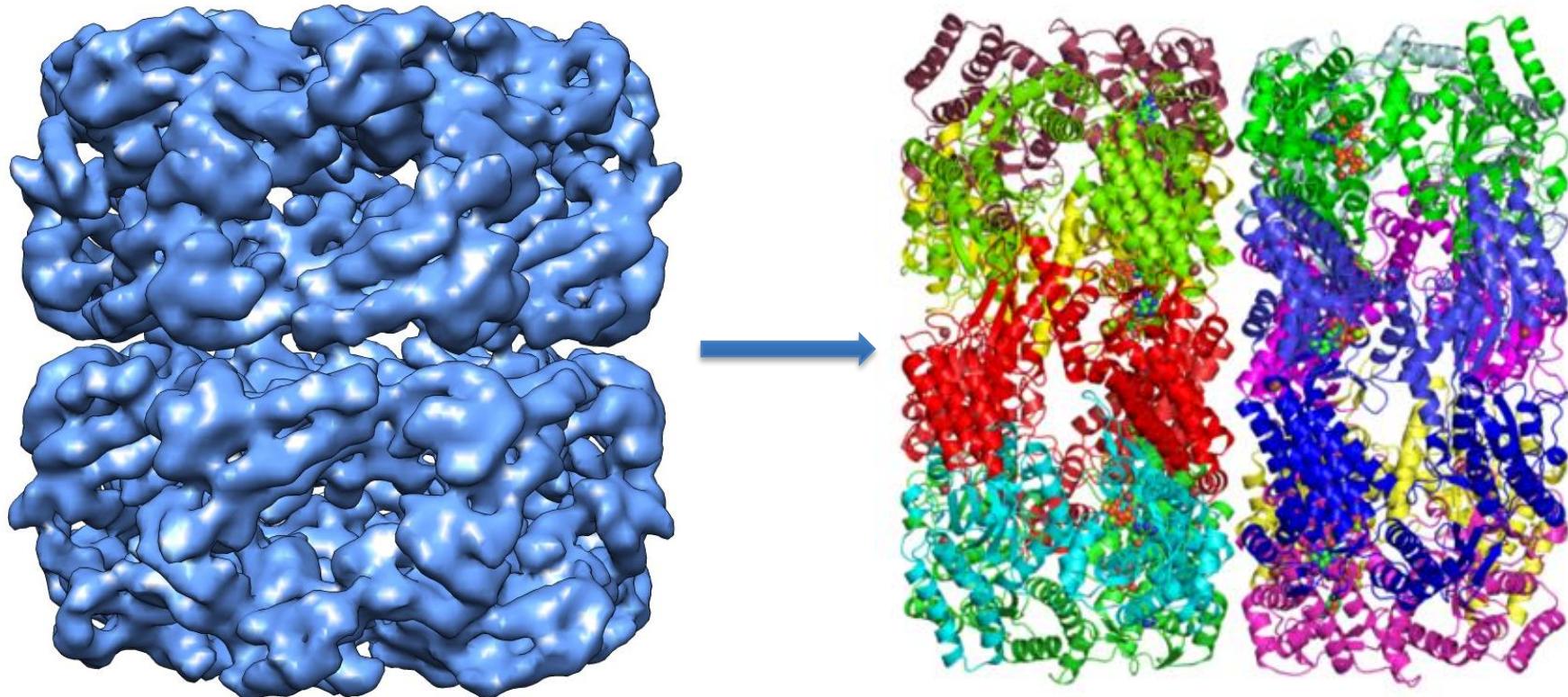
Copurification



FRET

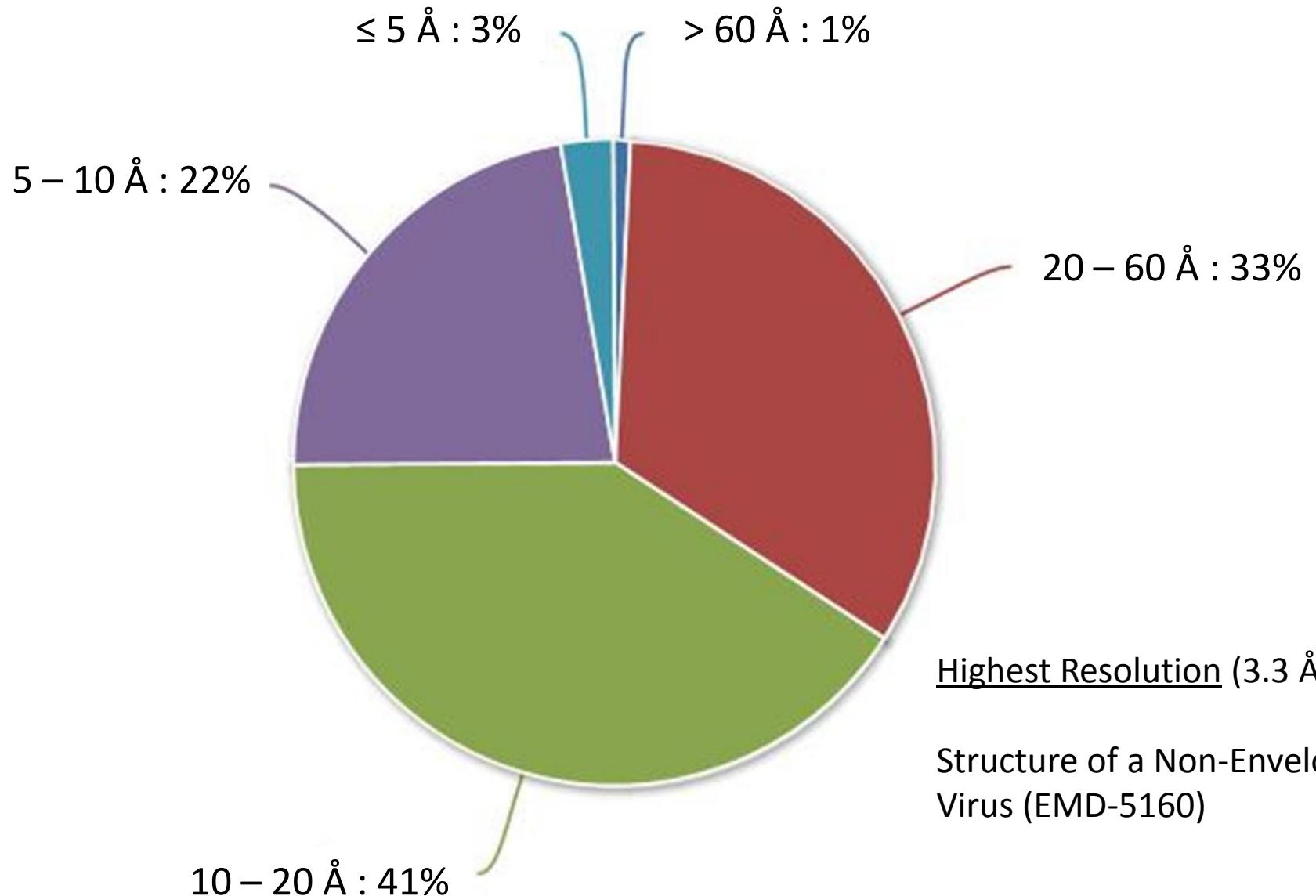
Gross Structural Information
(Subunit proximity, distances
and contacts etc.)

Cryo-Electron Microscopy based Modeling

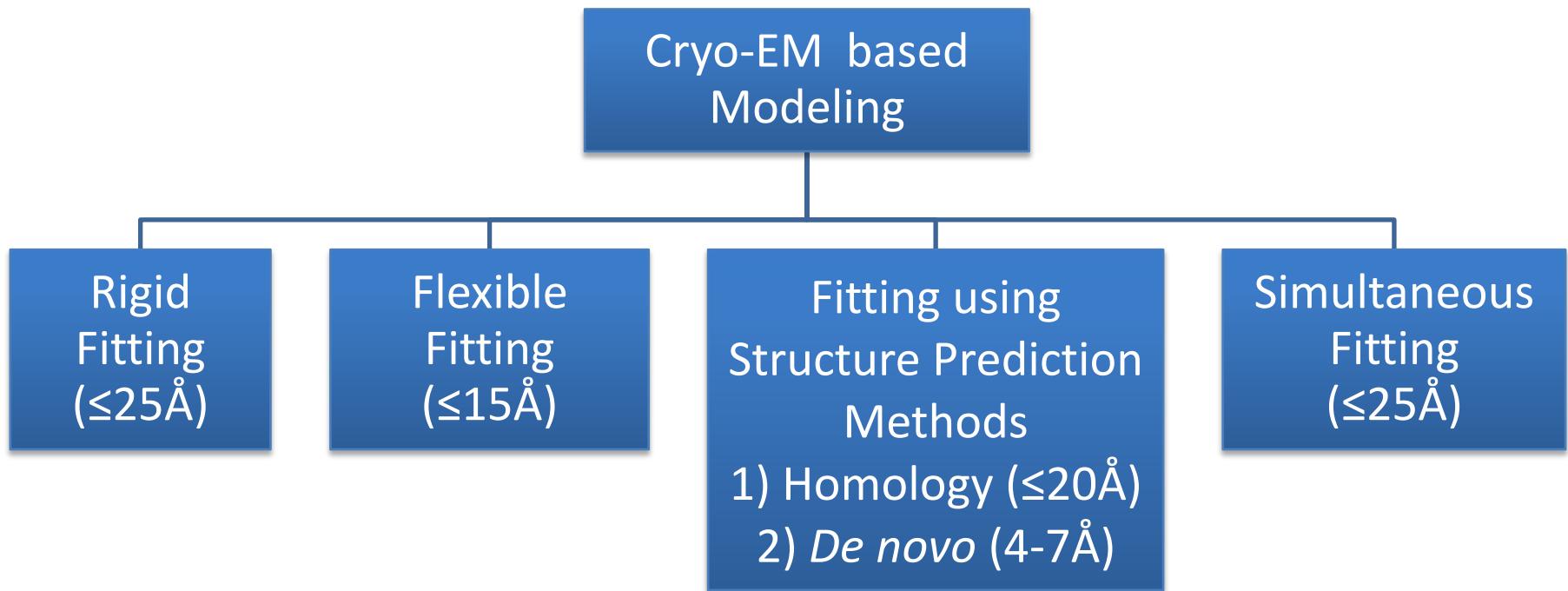


GroEL (EMD-2221)

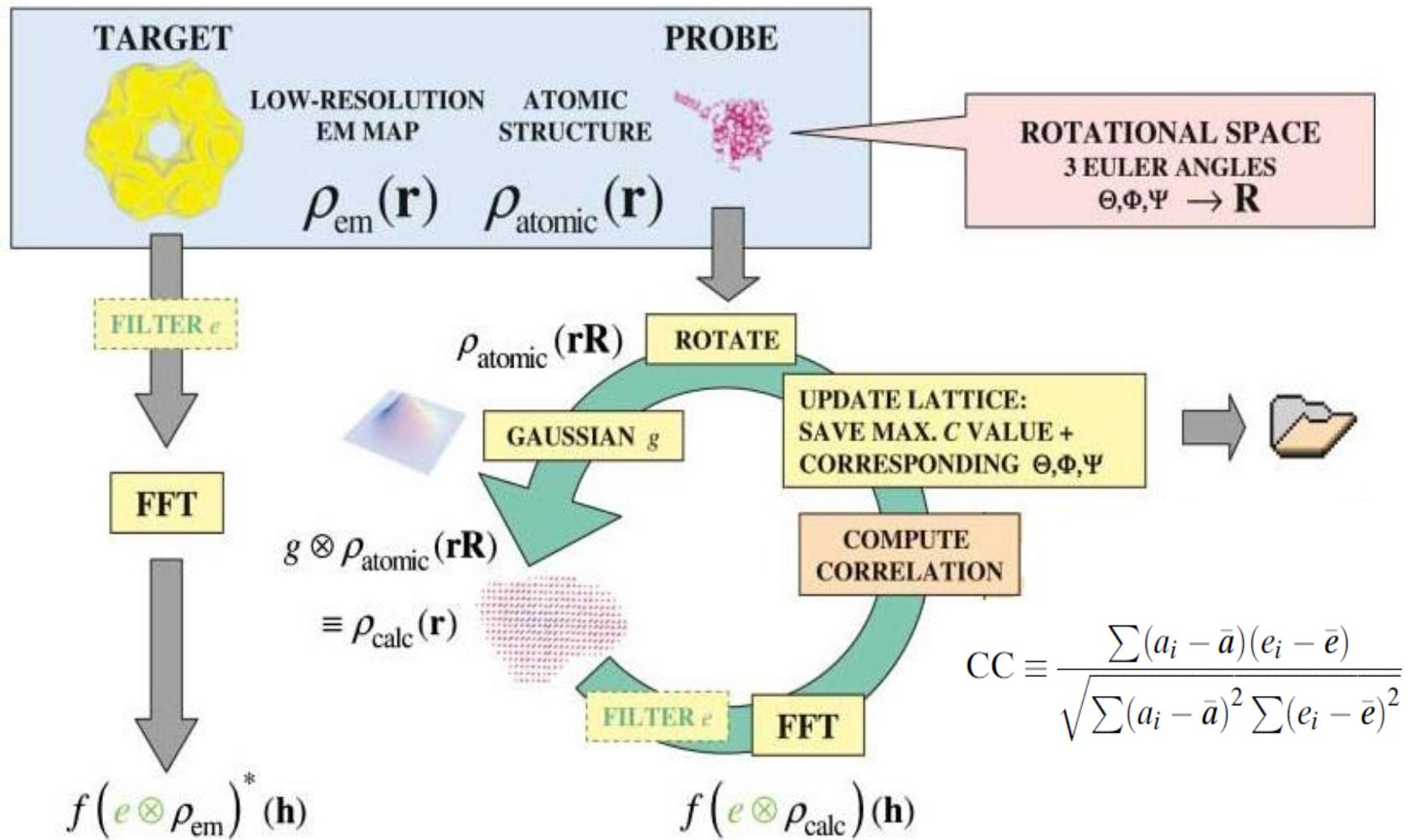
Resolution distribution of EM density maps in EMDB



Cryo-Electron Microscopy Modeling Techniques



Rigid Body Fitting



Flexible Fitting

Energy Function

$$U_{total} = U_{MD} + U_{EM} + U_{SS}$$

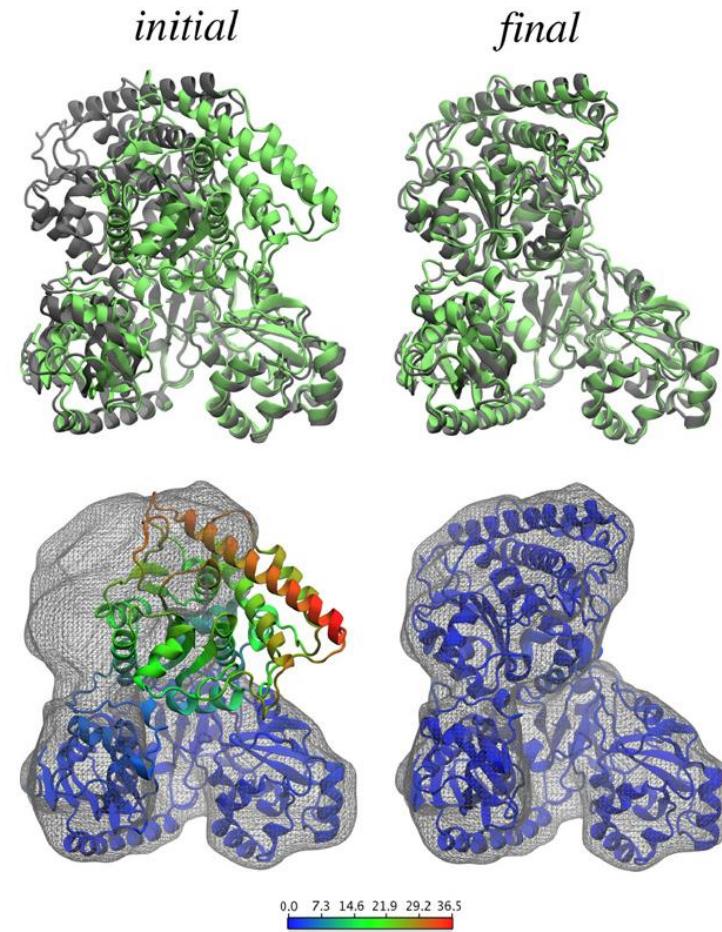
EM Map Energy Term

$$U_{EM}(R) = \sum w_j V_{EM}(r_j)$$

$$V_{EM}(r) = \begin{cases} \xi \left[1 - \frac{\Phi(r) - \Phi_{thr}}{\Phi_{max} - \Phi_{thr}} \right] & \text{if } \Phi(r) \geq \Phi_{thr}, \\ \xi & \text{if } \Phi(r) < \Phi_{thr} \end{cases}$$

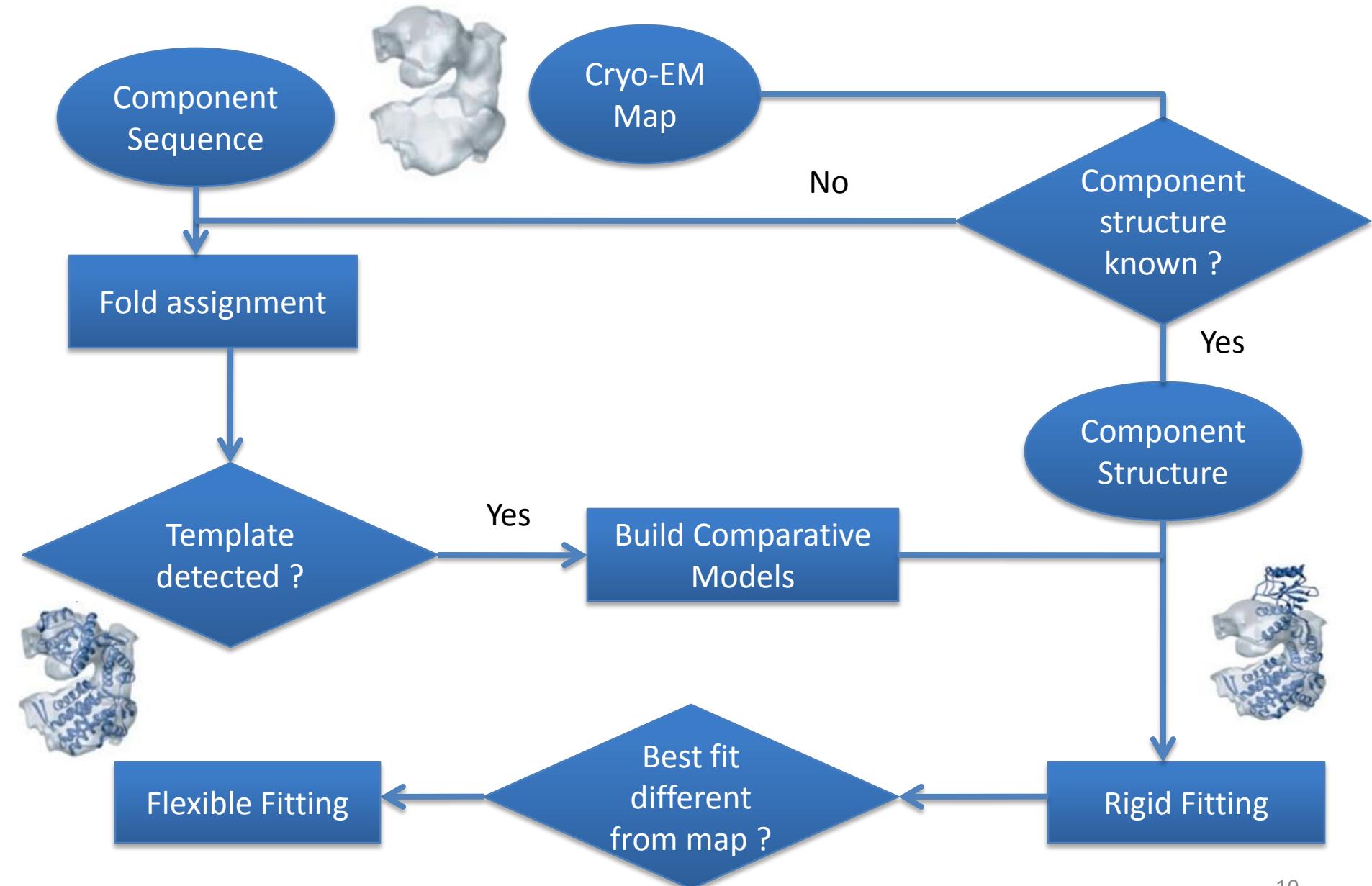
Secondary Structure based Term

$$U_{SS} = \sum_{\mu} k_{\mu} (X_{\mu} - X_{\mu}^0)^2$$

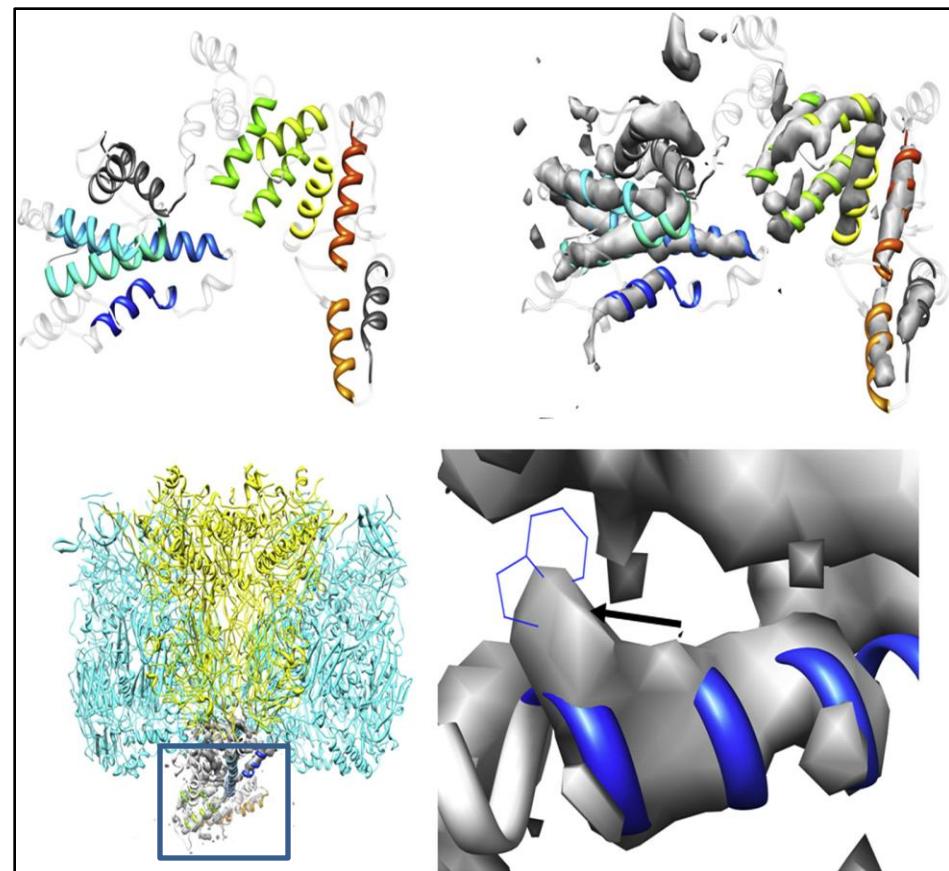
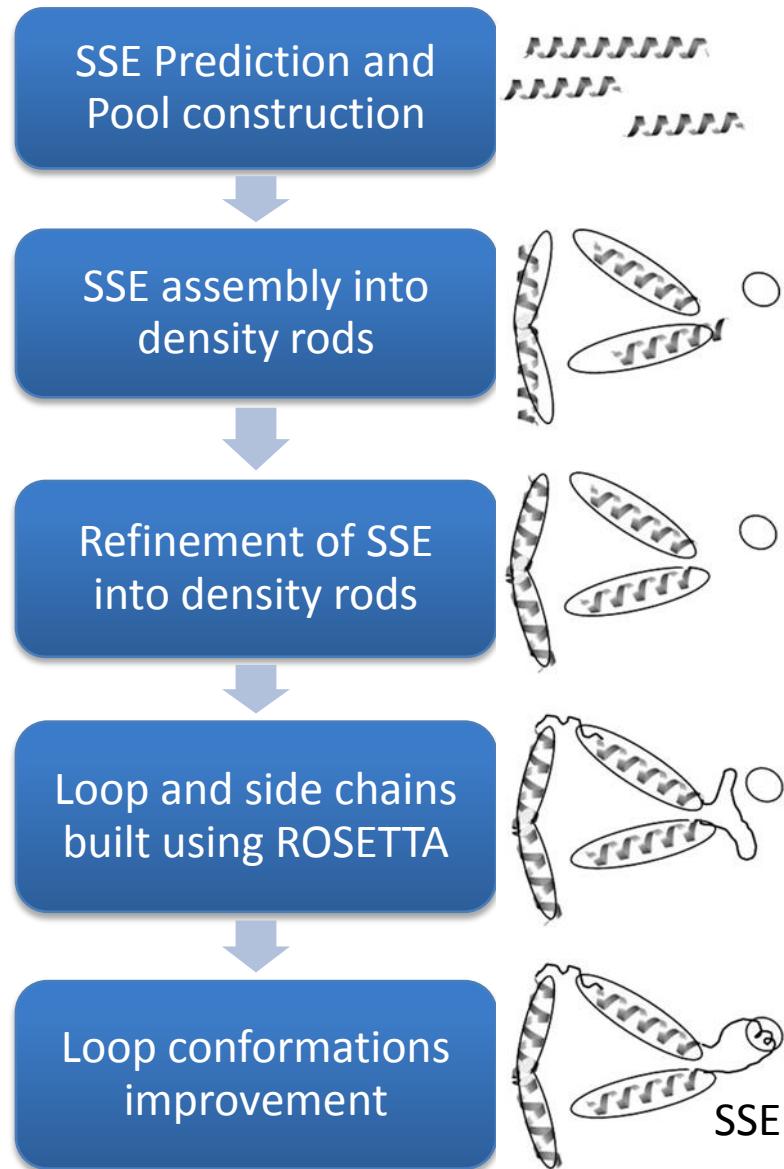


Acetyl-coenzyme A synthase/carbon monoxide dehydrogenase (ACS/CODH)

Homology Modeling and Cryo-EM Fitting



De novo Modeling based on Structure Prediction

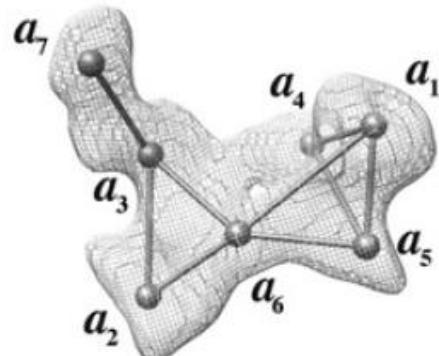


Adenovirus Protein IIIa (Cryo-EM map 6.9 Å)

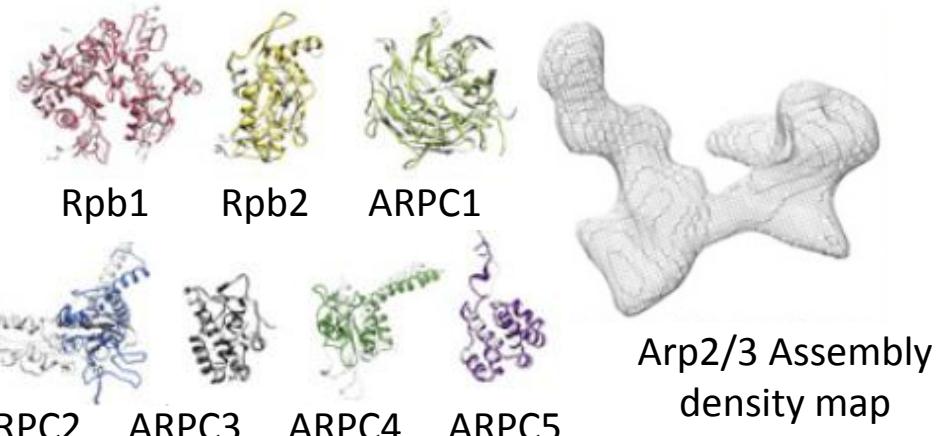
SSE – Secondary Structure Elements

Simultaneous Fitting

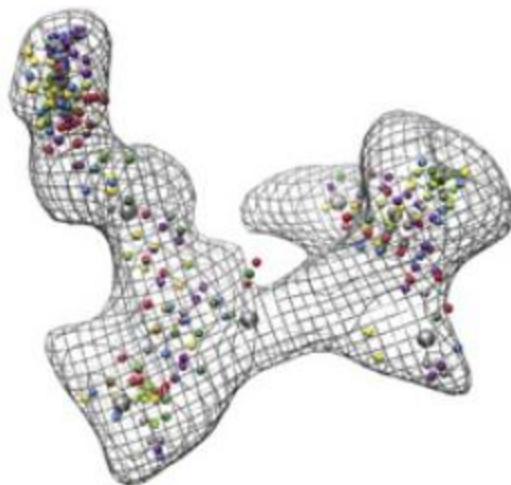
Anchor Graph Construction



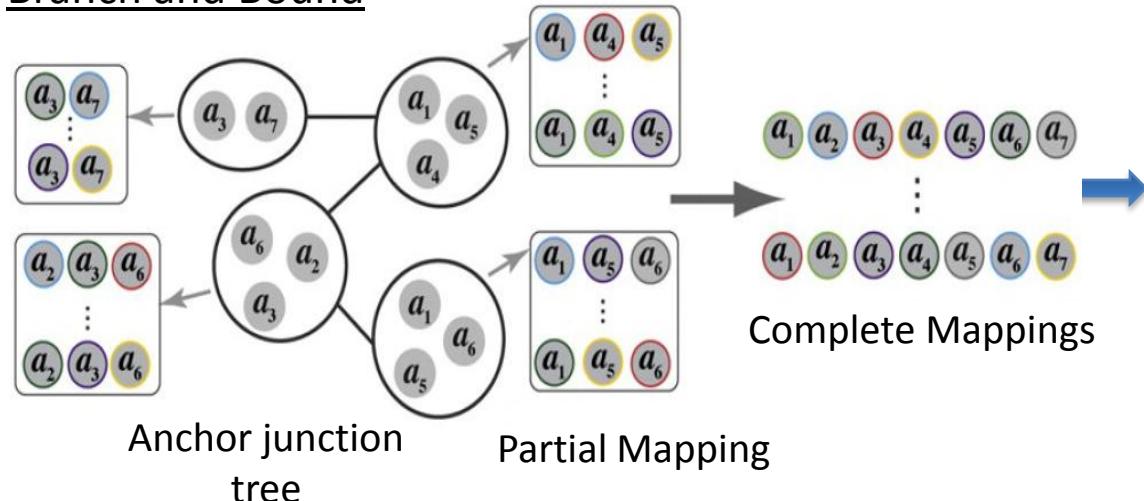
Input



Discretization

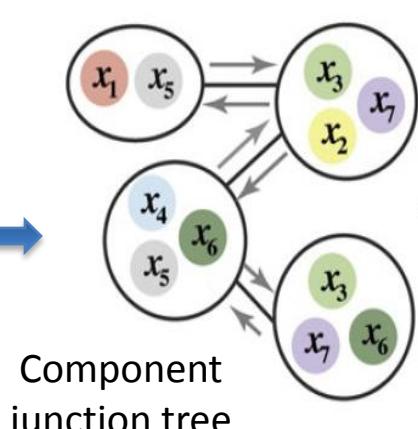


Branch and Bound

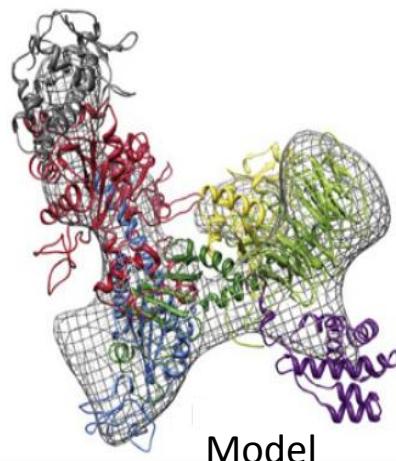


Simultaneous Fitting

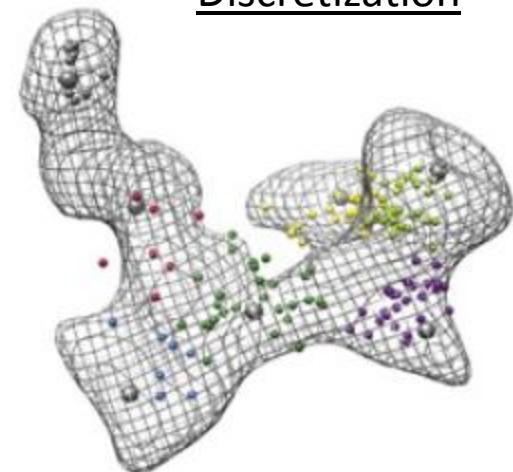
For each selected mapping



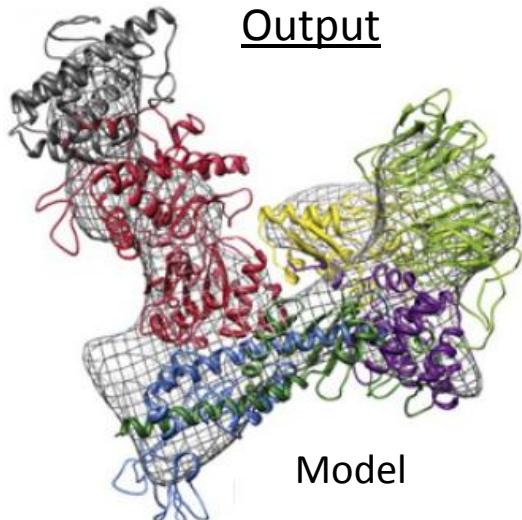
Component junction tree



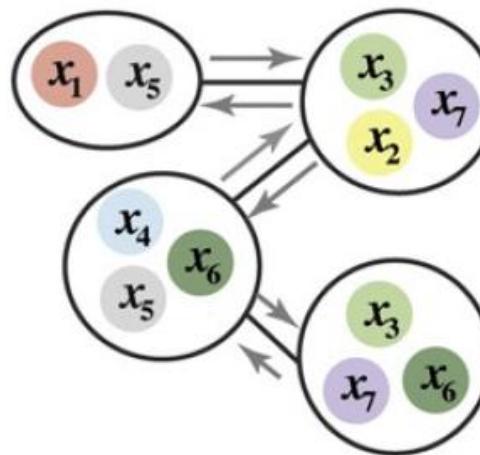
Top 20 ranked Models



Output



Optimization



Summary so far...

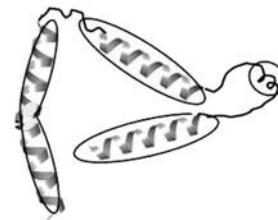
- Cryo-EM modeling techniques provide pseudo atomic models and help in interpreting low resolution density maps.
- The techniques range from simple rigid body fitting to simultaneous fitting techniques.



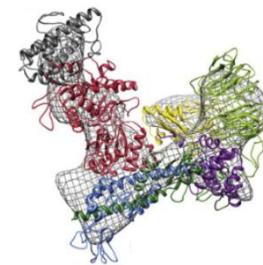
Rigid body



Flexible

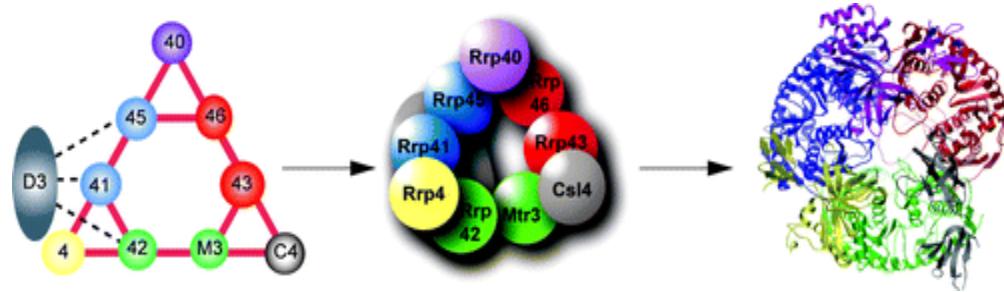


De novo

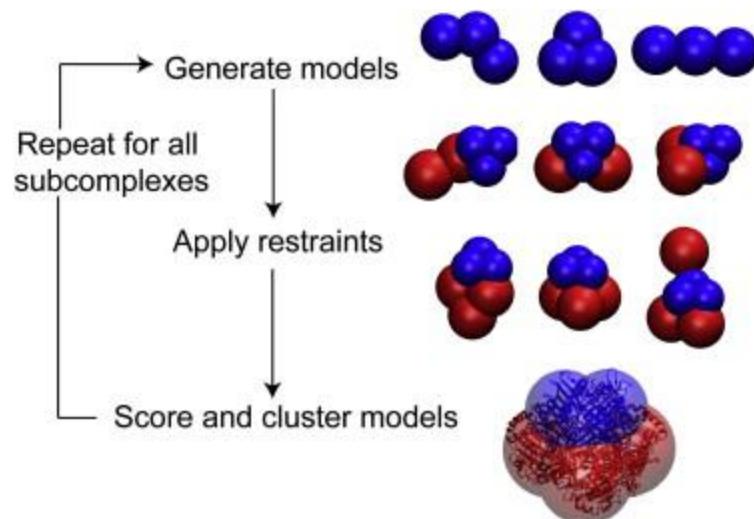


Simultaneous

Modeling based on Mass Spectrometry Data



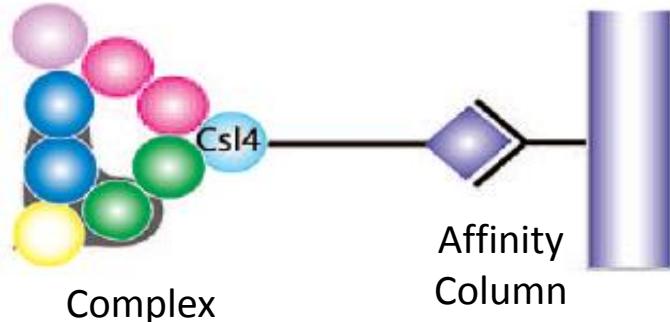
Interaction
Network based
Modeling



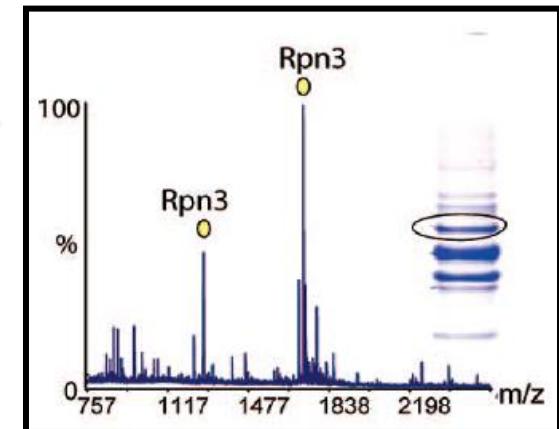
Ion Mobility Mass Spectrometry
(IM - MS) based Modeling

Interaction Network based Modeling

Affinity Purification

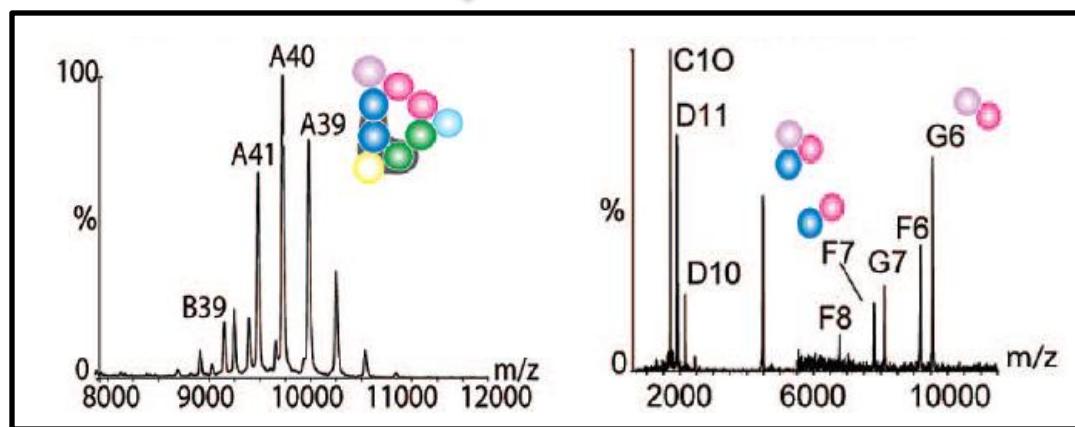


List of Subunits

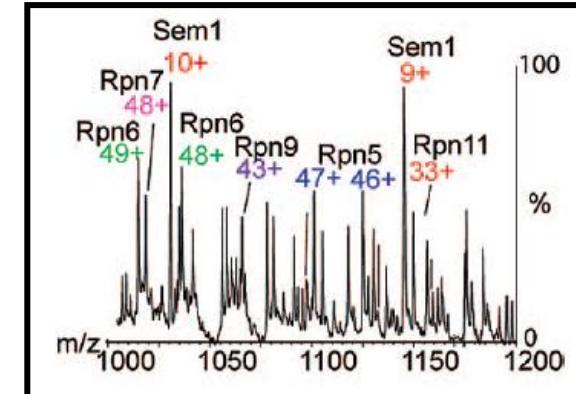


Complex

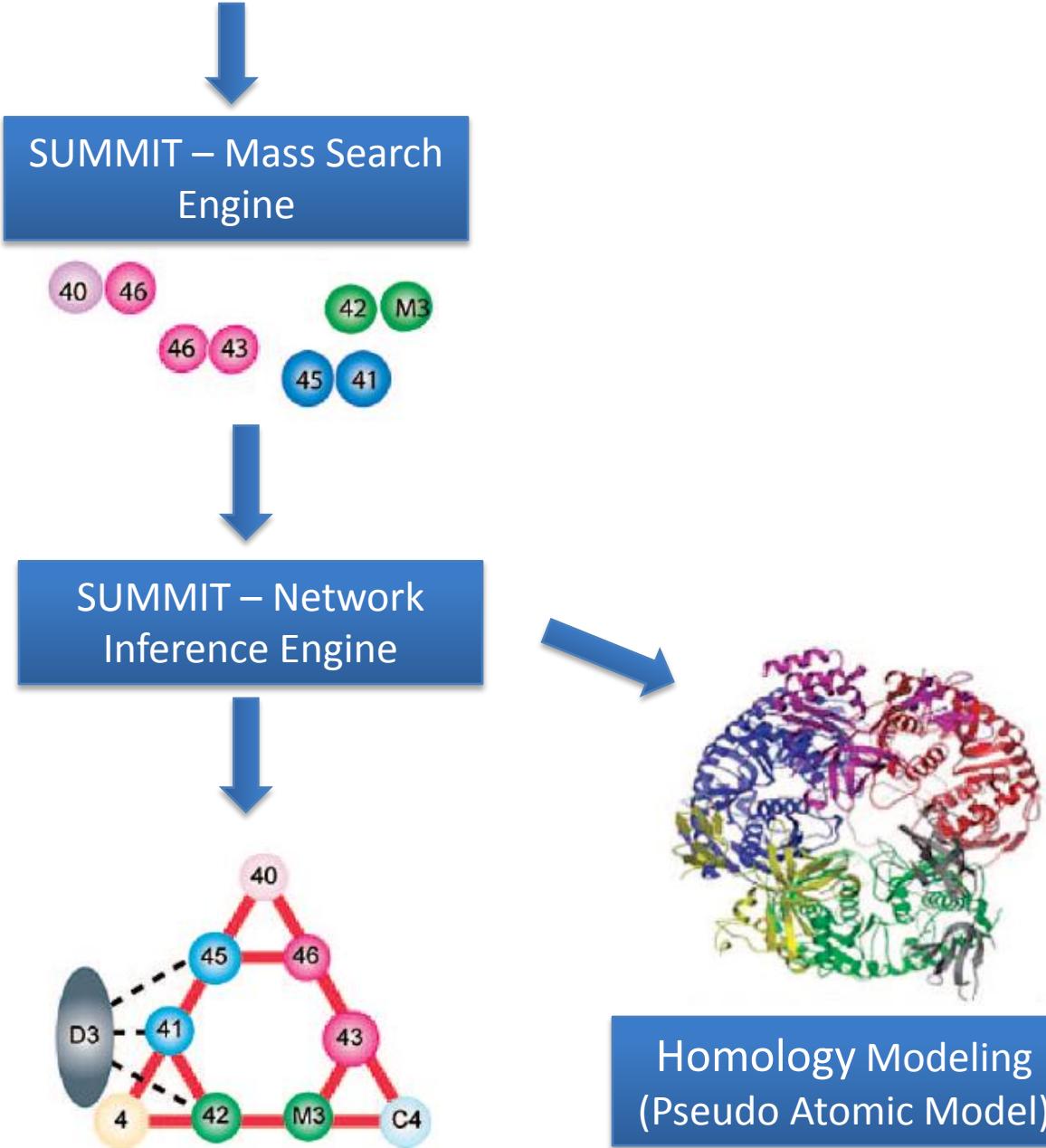
Affinity
Column



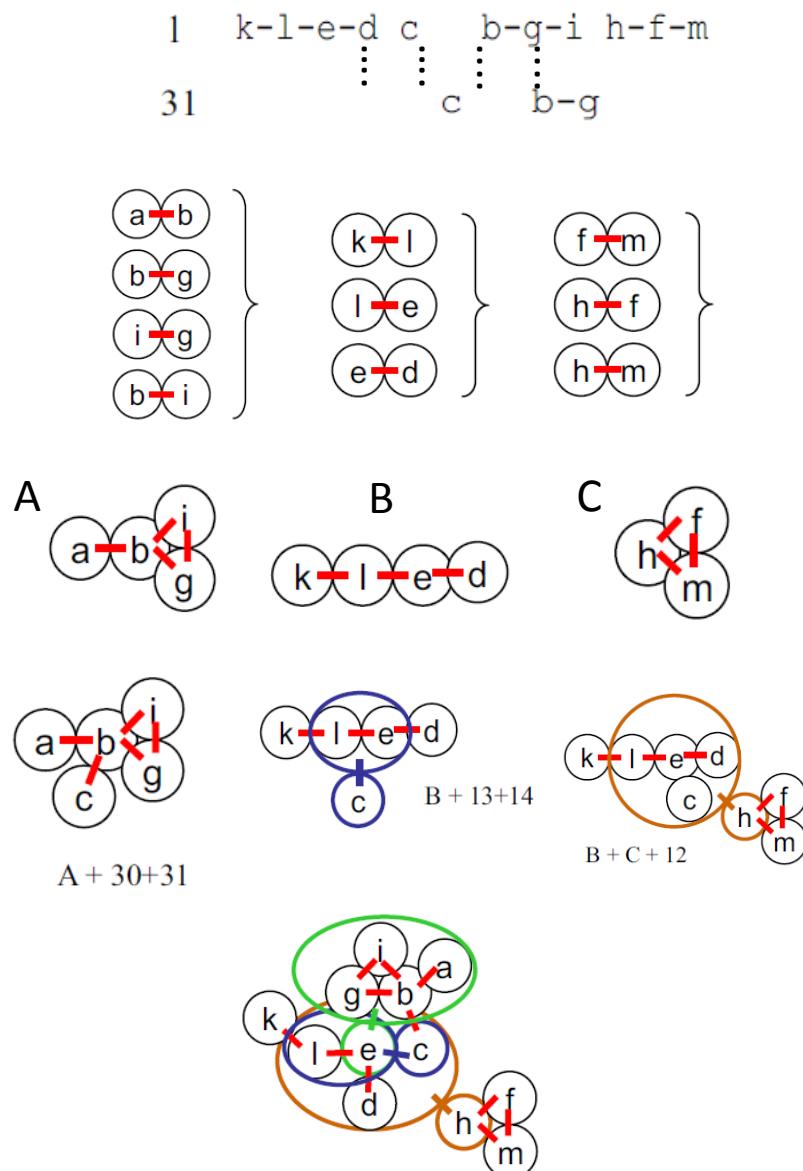
List of Subunit Masses



Establish Stoichiometry of whole complex
and generate subcomplexes



SUMMIT Algorithm



List of subunits and masses of complex and subcomplexes

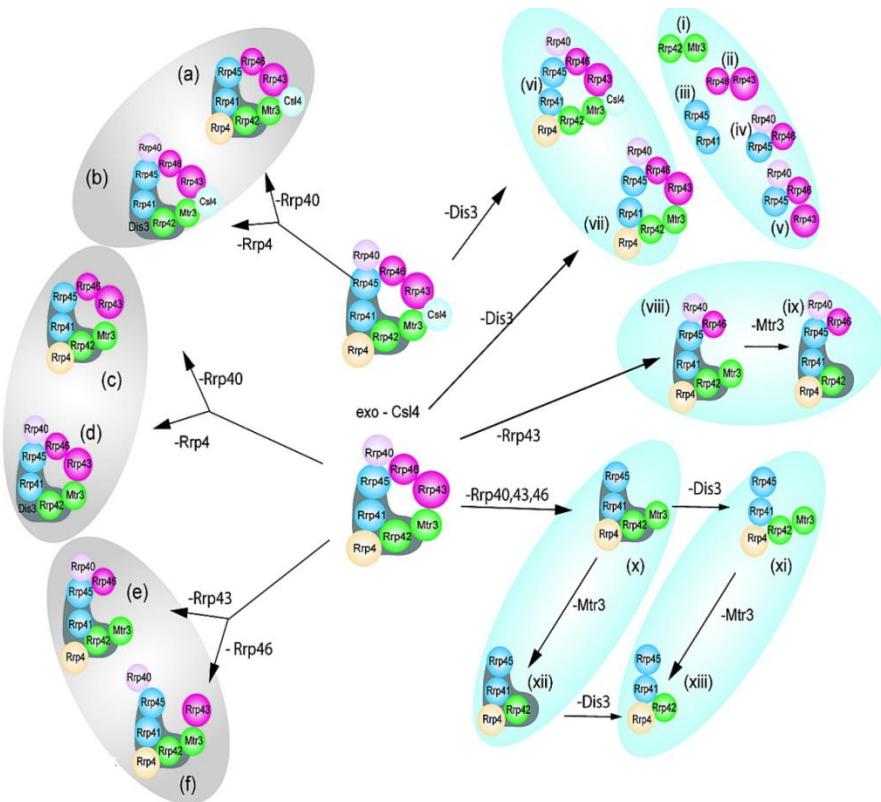
Generate all combinations of subunit pairs

Build modules from the pairs

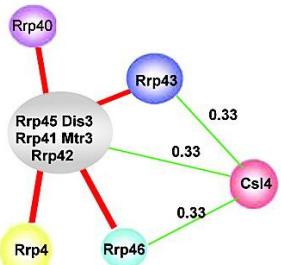
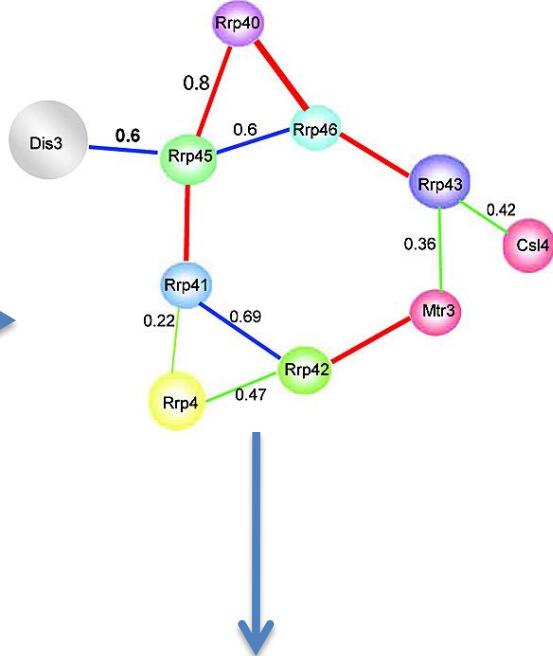
Add new components to the existing modules

Weighted average network

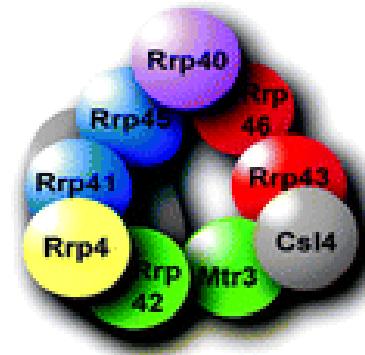
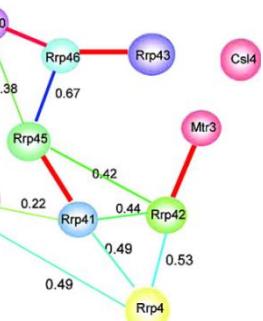
Yeast RNA Exosome Interaction Network



Gas and Solution Phase



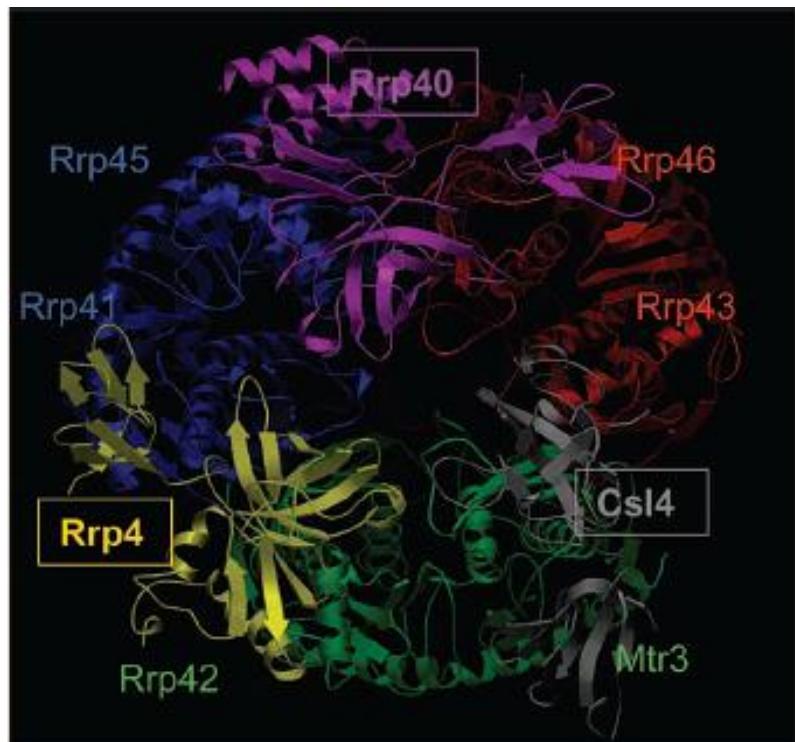
Solution Phase



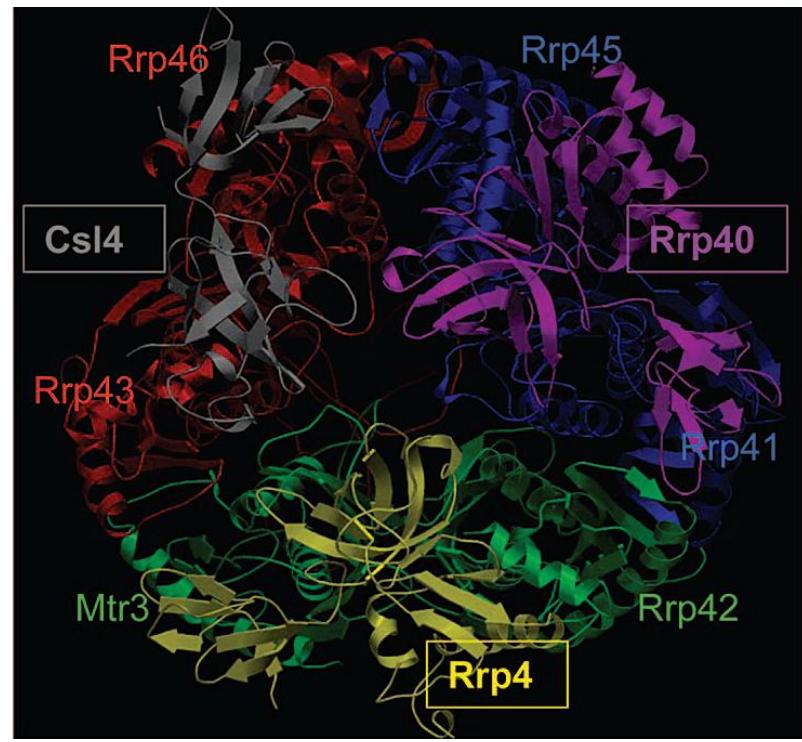
Gas Phase

Homology model based on the Interaction Network

Model A



Model B



Restraint

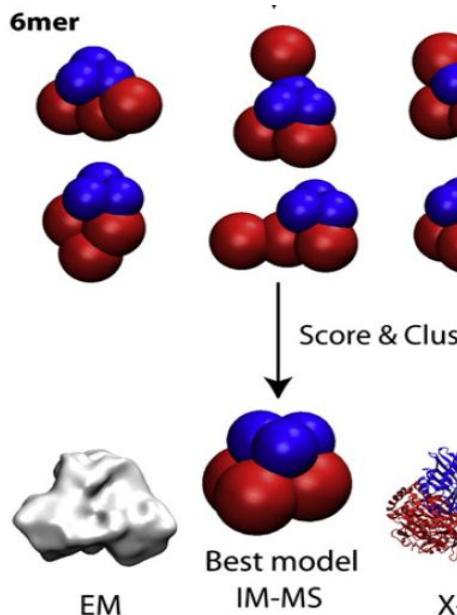
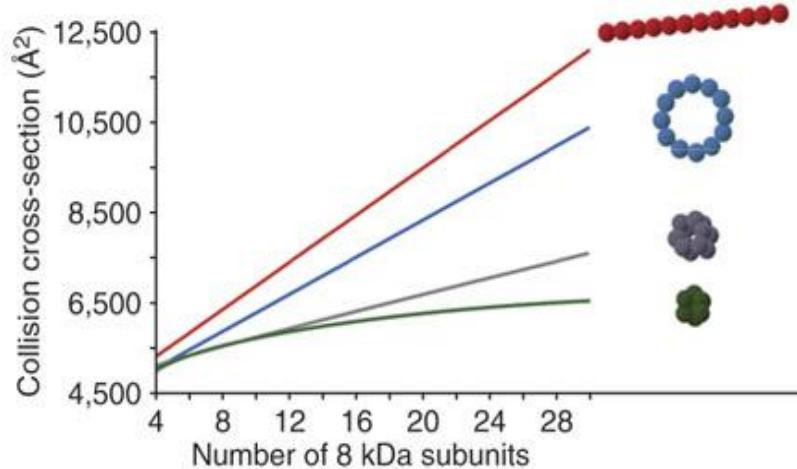
It is defined as a harmonic function which reaches its minimum based on the structural attribute used.

$$\text{Harmonic Function} = k(x - x_0)^2$$

Contact	Angle	Excluded volume
Proximity	Chirality	Shape
Charge	Surface area	Symmetry
Distance	Volume	Localization of particles

Overview of Modeling Strategy using IM-MS Data

Collision cross-section (CCS) vs Assembly Topology



Repeat for
complex and
subcomplexes

Ion Mobility – Mass Spectrometry (IM-MS)

Representation based on CCS

Model Generation

Filter models
 $S_T = S_{CCS} + S_V (+ S_{sym})$

Evaluation to select best model agreeing with S_T

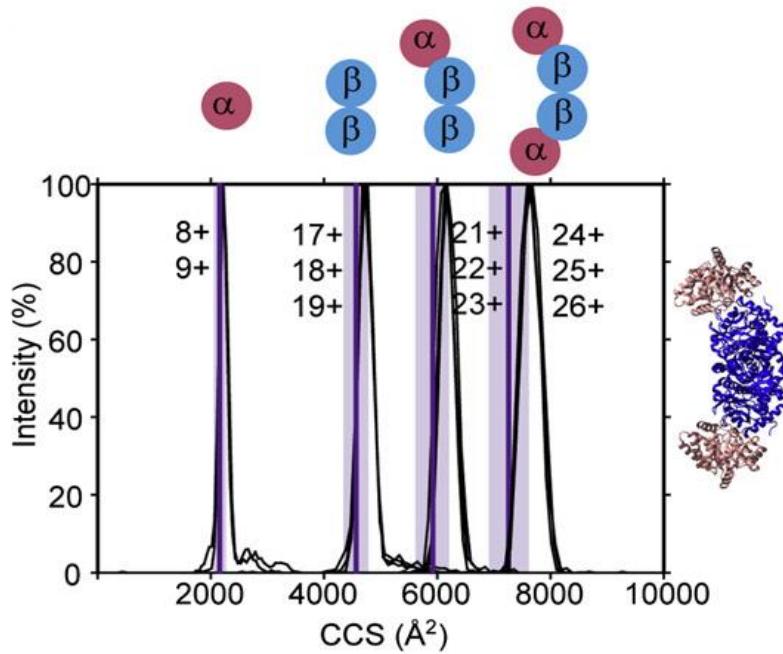
Structure Characterization of Tryptophan Synthase

Volume restraint

$$S_V = \left(\frac{V_{exp} - V_{calc}}{\sigma_{exp}} \right)^2$$

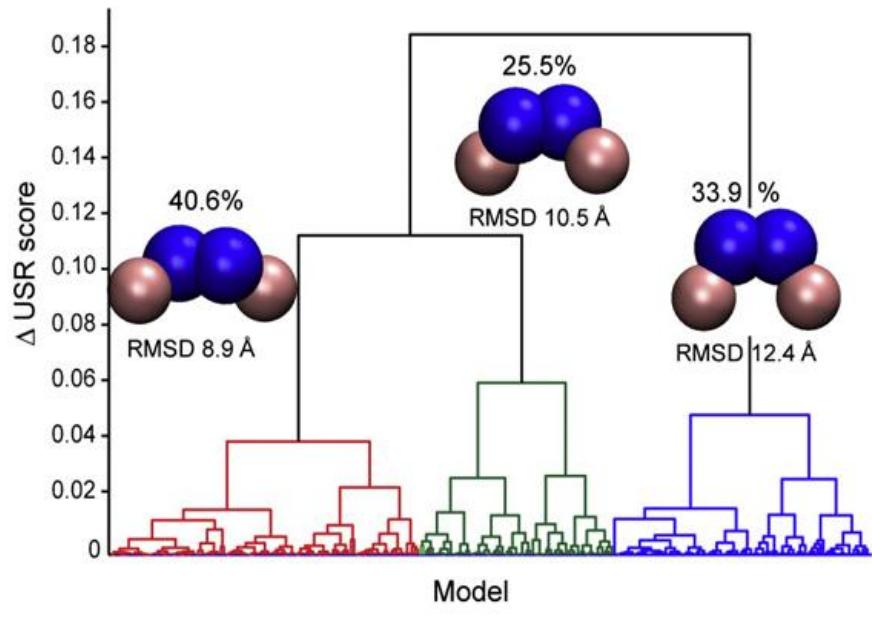
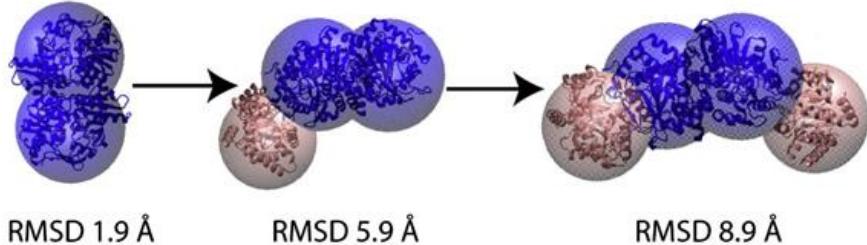
Collision Cross-section restraint

$$S_{CCS} = \left(\frac{CCS_{exp} - CCS_{calc}}{\sigma_{exp}} \right)^2$$



Ion-Mobility Arrival Time Distribution

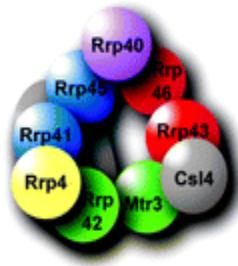
Volume and Collision Cross-section
restraints applied during model generation



Clustering of Models

Summary so far ...

- Validating the model using Mass Spectrometry data (interaction network).
- The topology of complex can be obtained from Ion Mobility Mass Spectrometry.

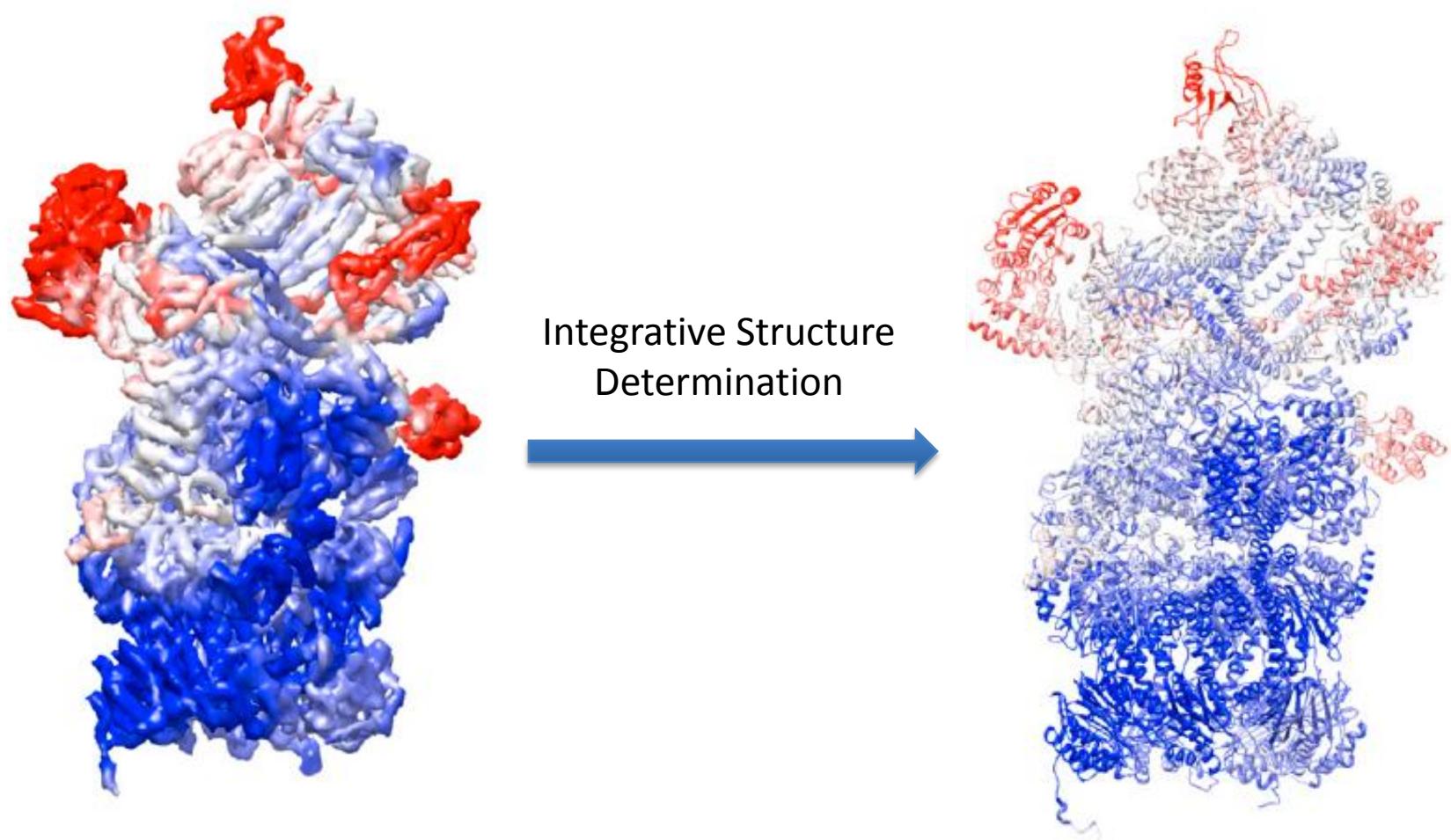


Interaction Network



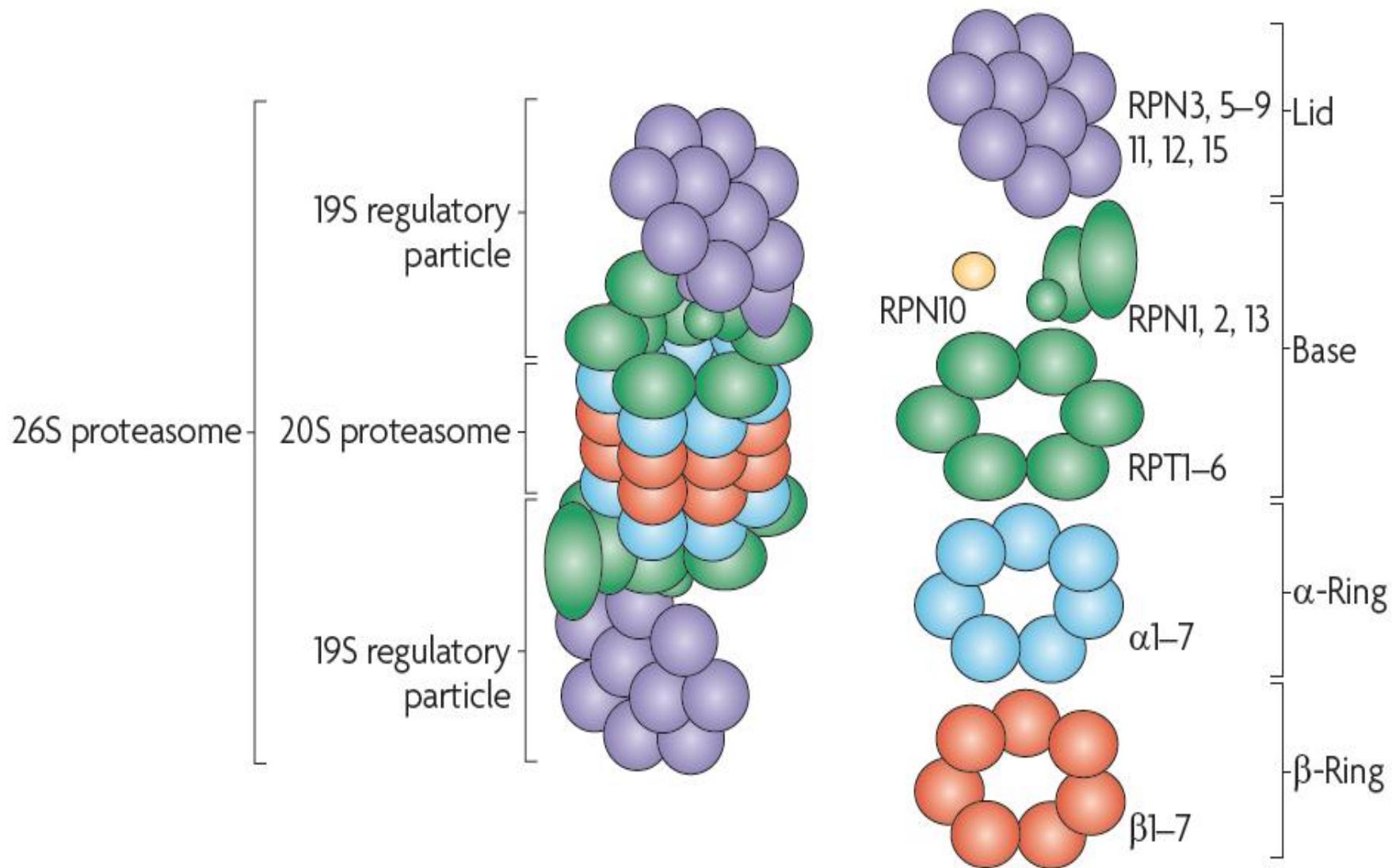
Topology of complex

Integrative Structure Determination of Proteasome

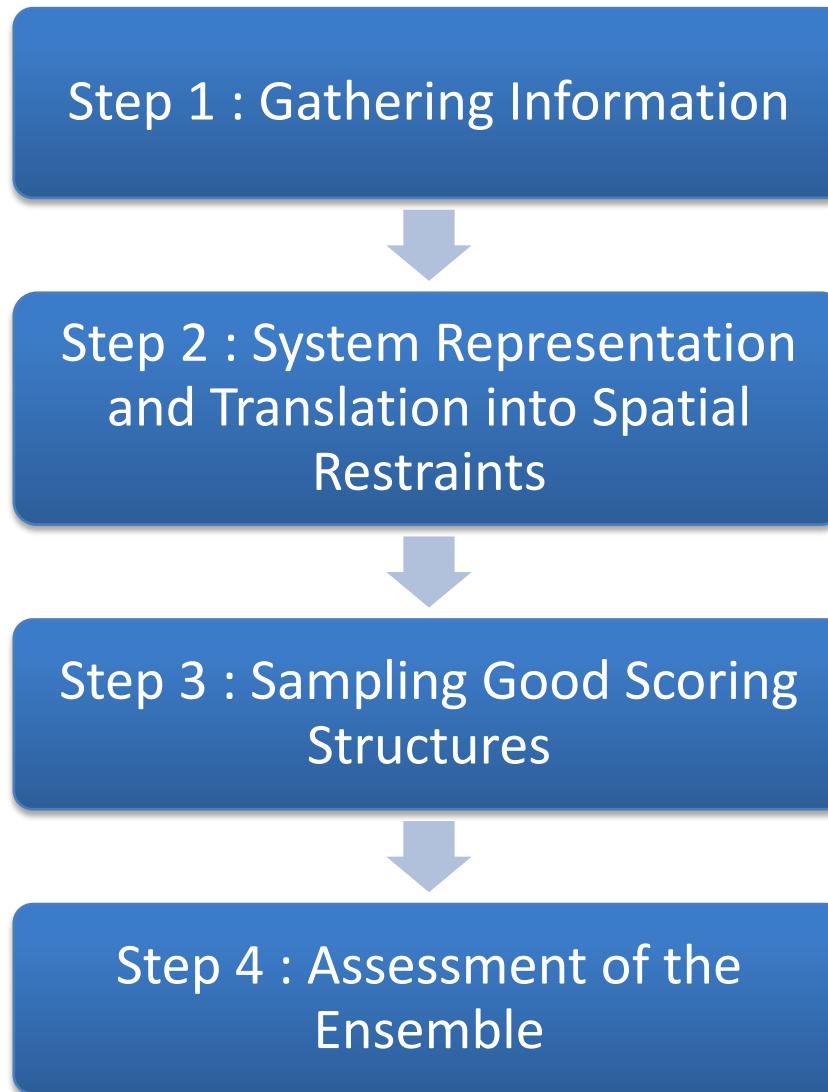


Coloring based on local cross correlation

Architecture of 26S Proteasome



Steps in Integrative Structure Determination



Step 1: Gathering Information

Protein – Protein interactions

- 13 Yeast two Hybrid
- 13 Subunit Cross links and 27 Pull downs

Residue Cross-linking

- 35 Cross links from *Saccharomyces pombe*
- 36 Cross links from *Saccharomyces cerevisiae*

Electron Microscopy

- *Saccharomyces pombe* wild type
- *Saccharomyces cerevisiae* deletion knockout (Rpn10 and Rpn13)

X-ray Crystallography

- 16 protein / domains

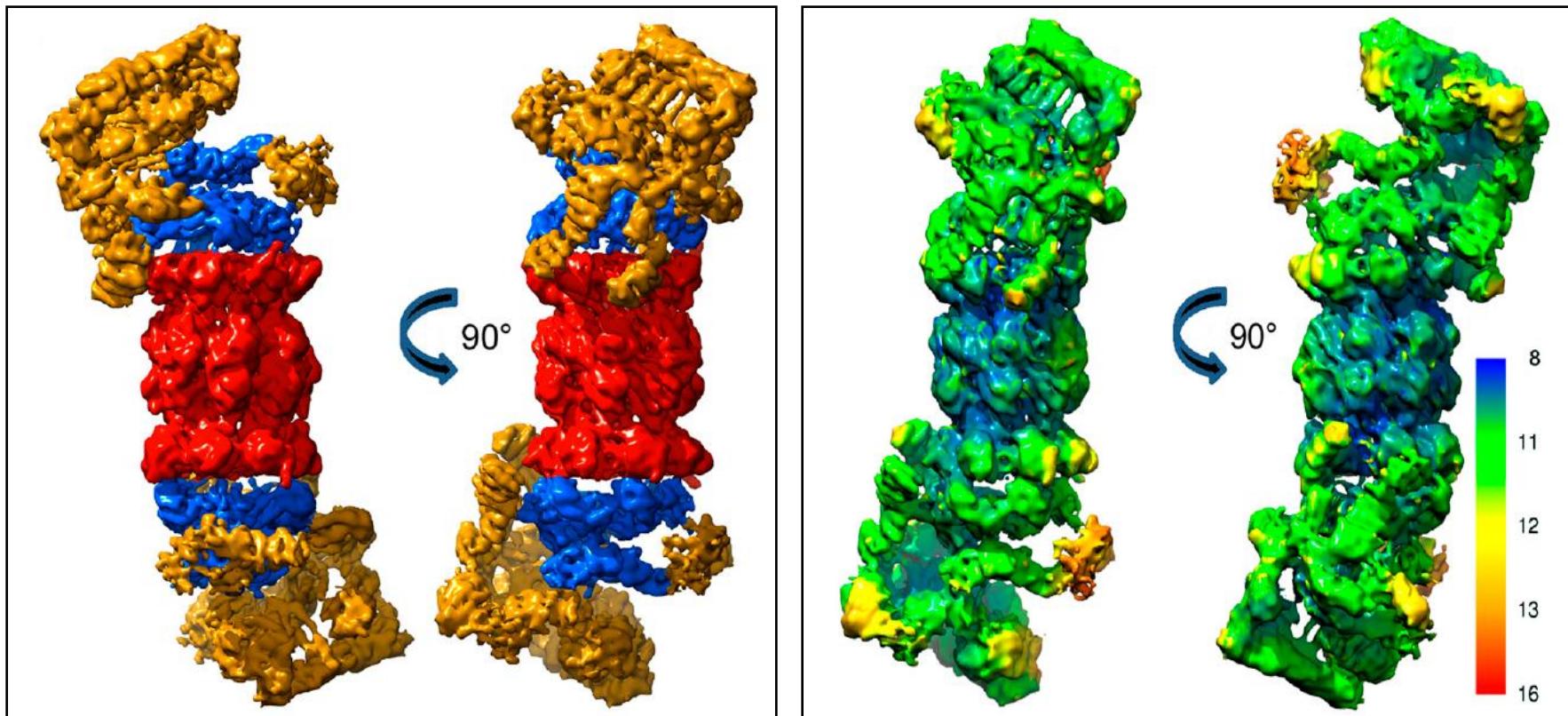
Homology Modeling

- 12 domains

Bioinformatics

- Structural parameters of proteins

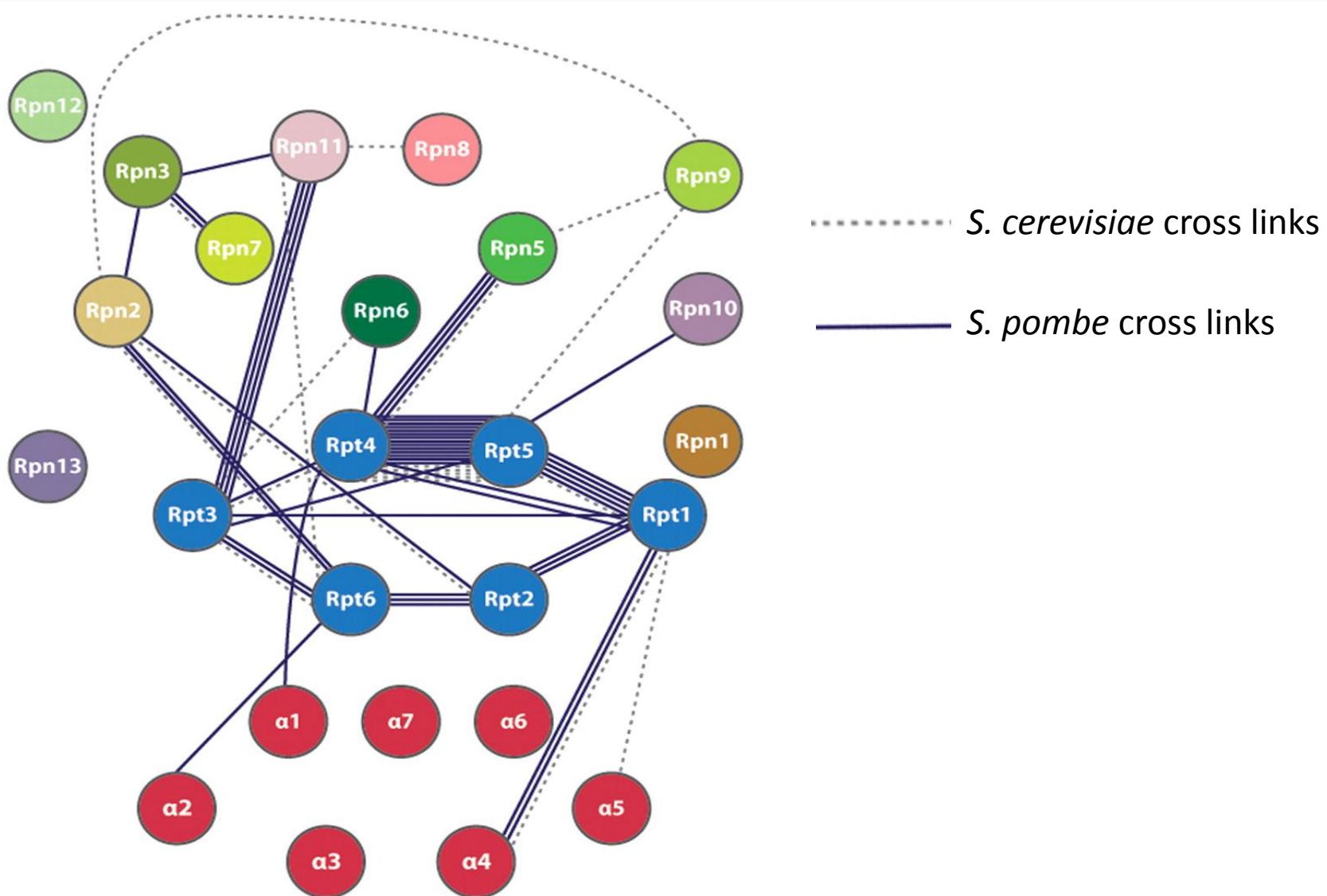
Step 1: 8.4 Å Cryo-EM map of *S. pombe* Proteasome



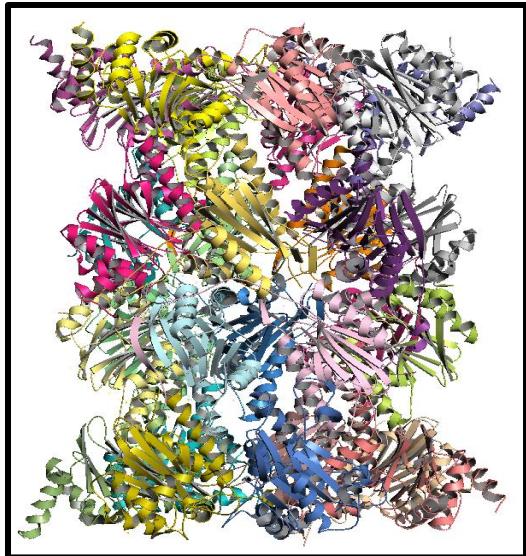
View of the maps around the
pseudo-seven fold axis

Maps colored according to Local Resolution

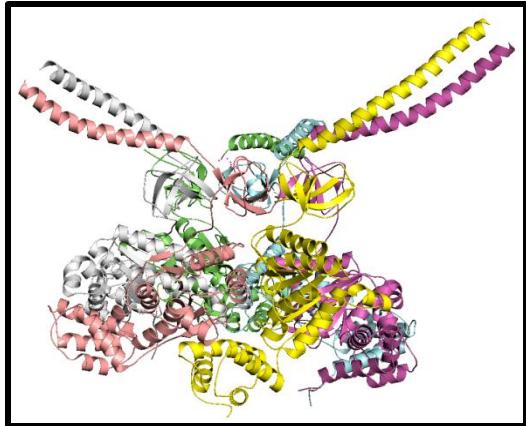
Step 1: Chemical cross-links for *S. pombe* and *S. cerevisiae* 26S proteasomes



Step 1: Structural Coverage of Proteasome

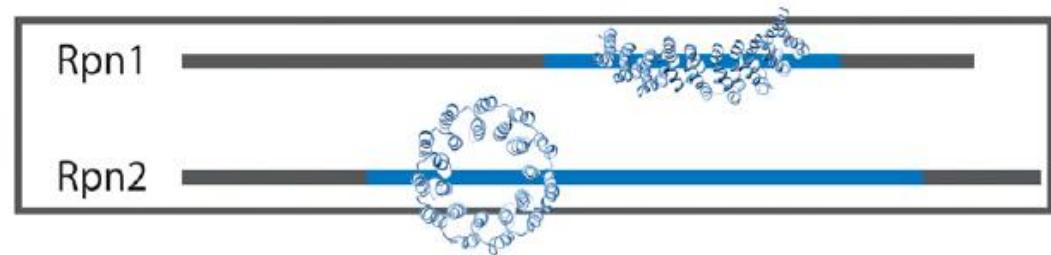


20S Core Proteasome ($\alpha_7 \beta_7$)

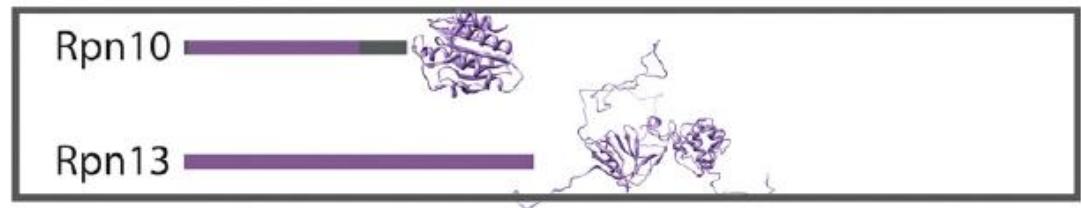


Rpt 1-6 AAA-ATPase modeled
on archaeal homolog PAN

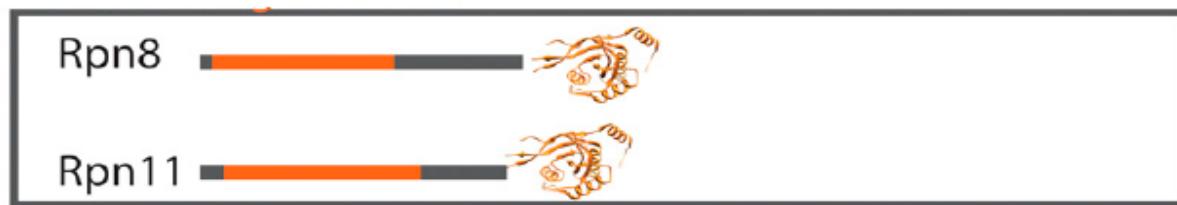
PC Repeat-containing Non-ATPases (Homology Models)



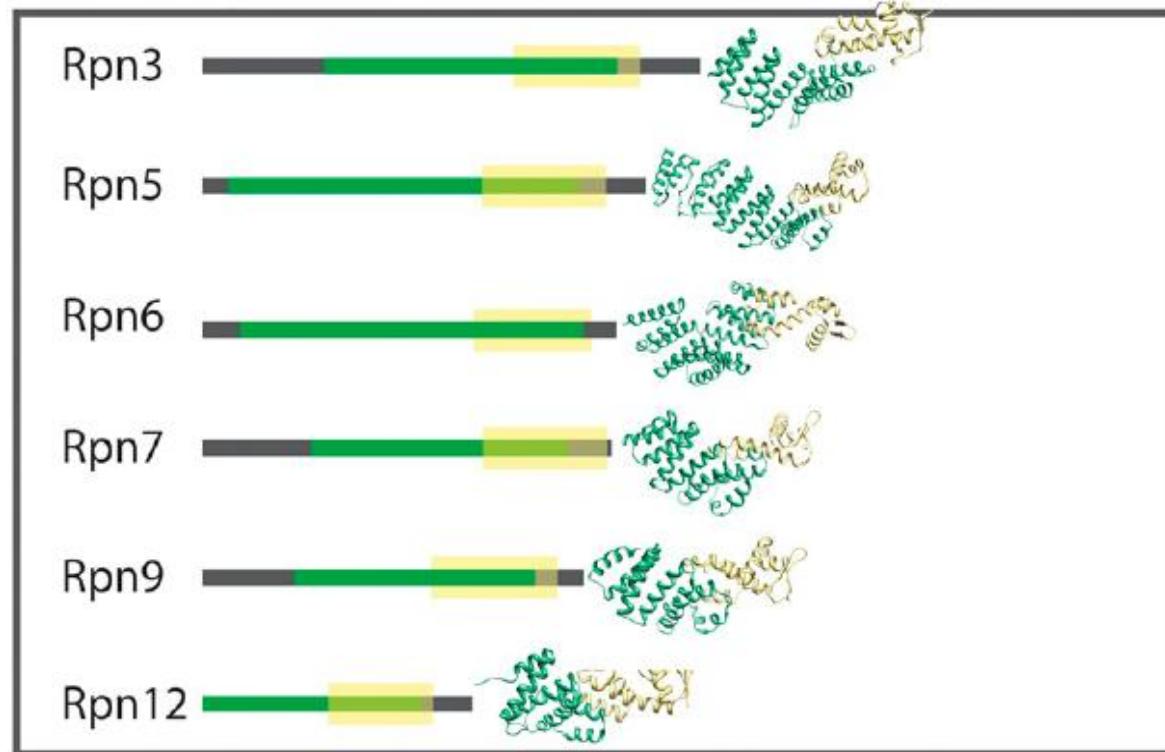
Ubiquitin Receptors (Homology Models)



Step 1: Structural Coverage of Proteasome

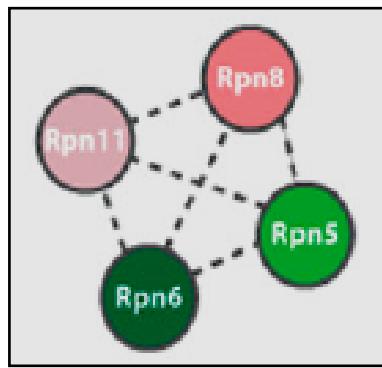


MPN Domain-containing
Non-ATPases
(Homology Models)

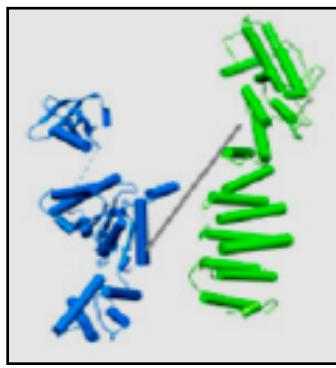


PCI Module-containing
Non-ATPases
(Homology Models)

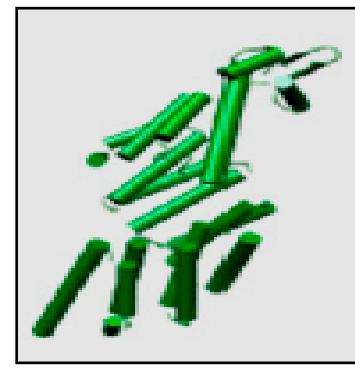
Step 2: System Representation and Spatial Restraints



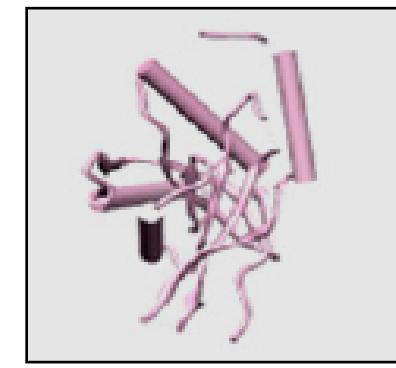
Protein connectivity
from protein – protein
interaction data



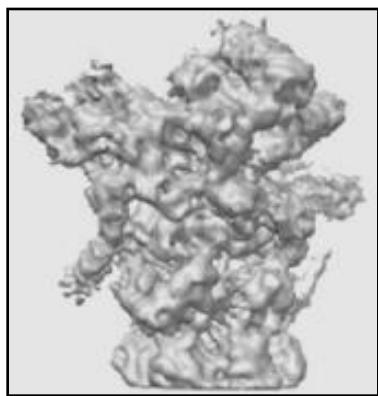
Residue contacts
from cross linking



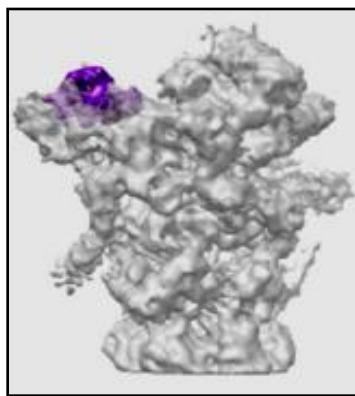
Atomic Structures from
X-ray Crystallography



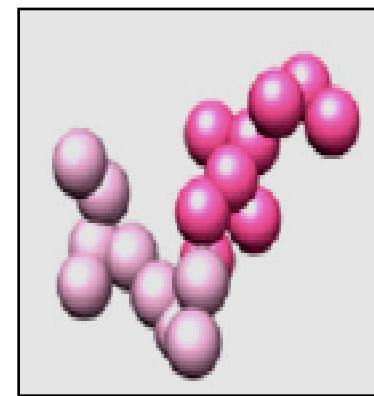
Homology Models



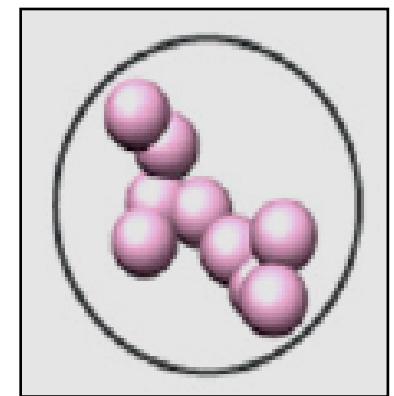
Assembly density
from Cryo-EM



Protein Localization



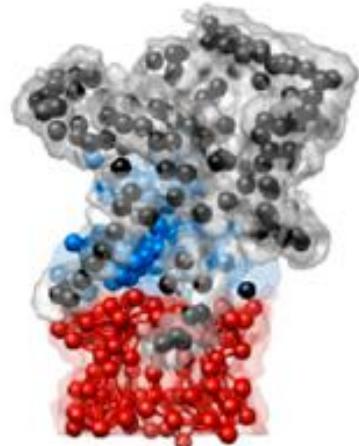
Excluded Volume



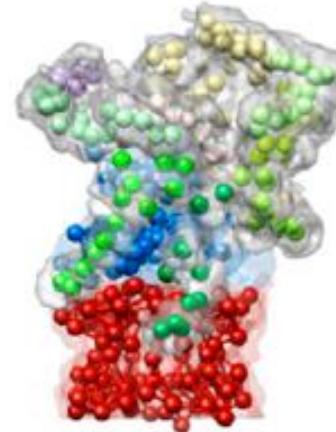
Radius of Gyration

Step 3 : Ensemble Sampling – Subunit Localization

Discretized Density Map

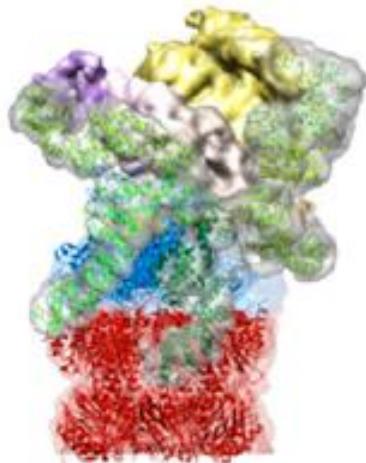


Subunit Localization Step



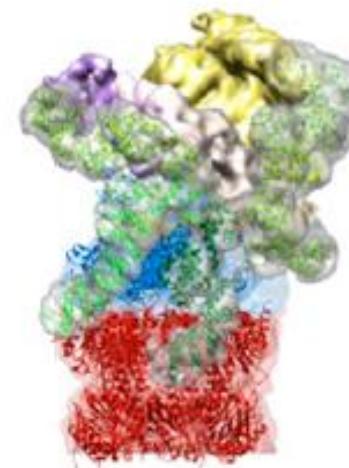
Coarse Sampling
(Subunit Contacts)

Refined Model



Flexible Fitting

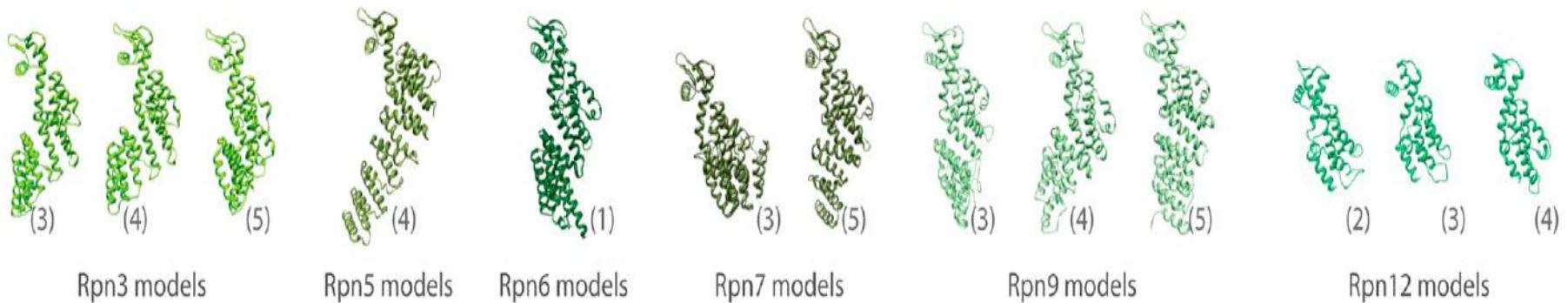
Subunit Configuration



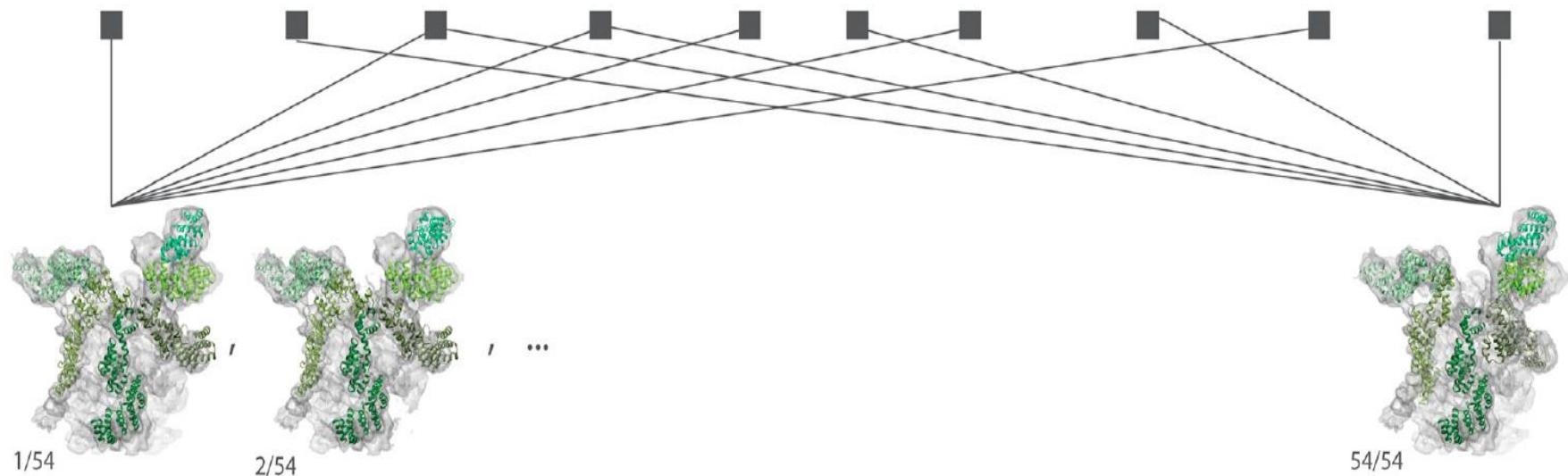
Refinement
(orientation
complementarity)

Step 3 : Subunit Configuration

Models from Homology Modeling

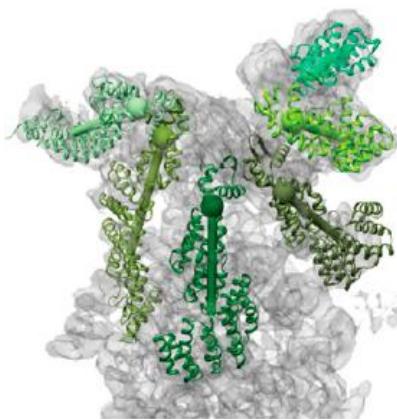


Multiple Subunit Fitting

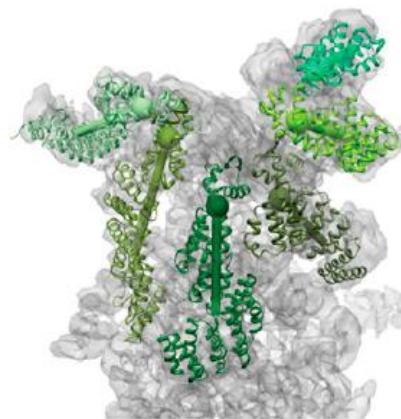


Step 3 : Clustering and Flexible Fitting Step

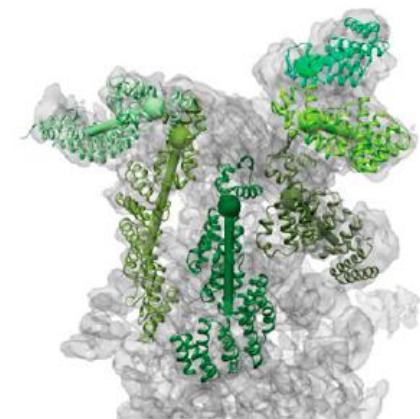
Clustering



Cluster 1

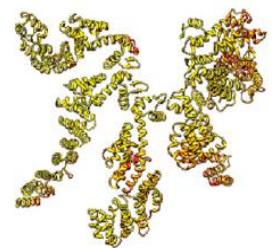
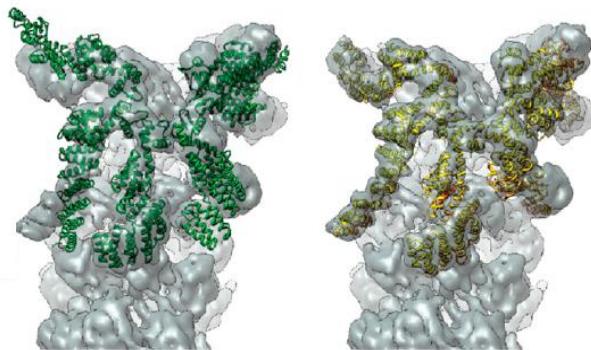


Cluster 2

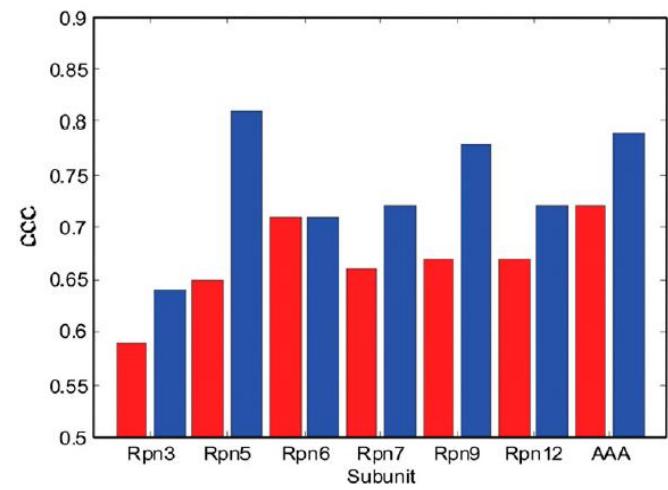


Cluster 3

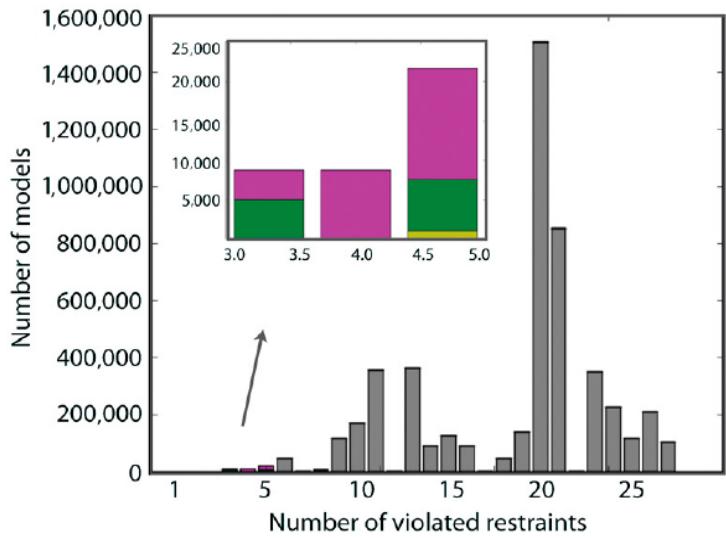
Flexible Fitting



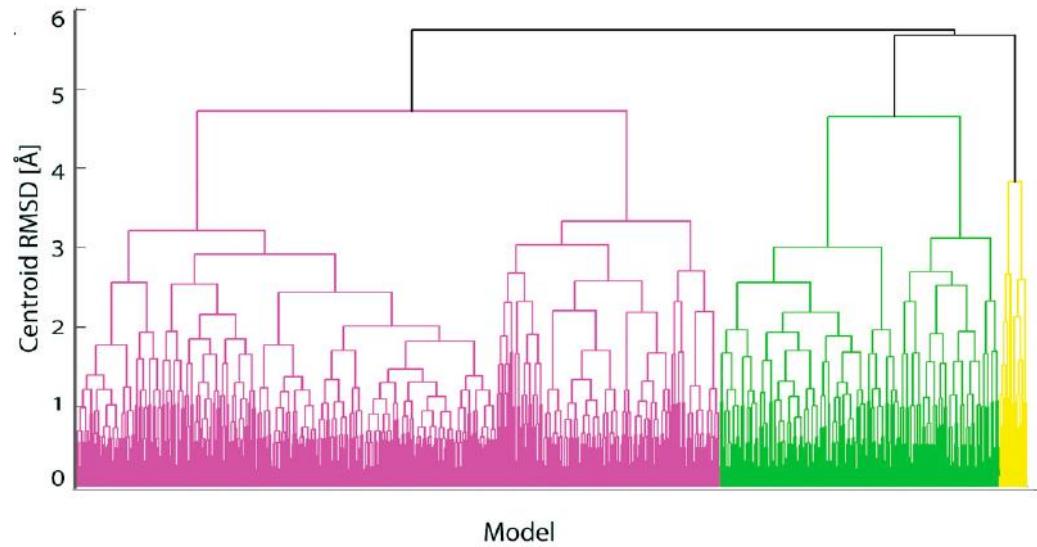
0.5 0.7 0.9



Step 4 : Assessment of Data and Structure



Selecting the models which violate atmost five restraints

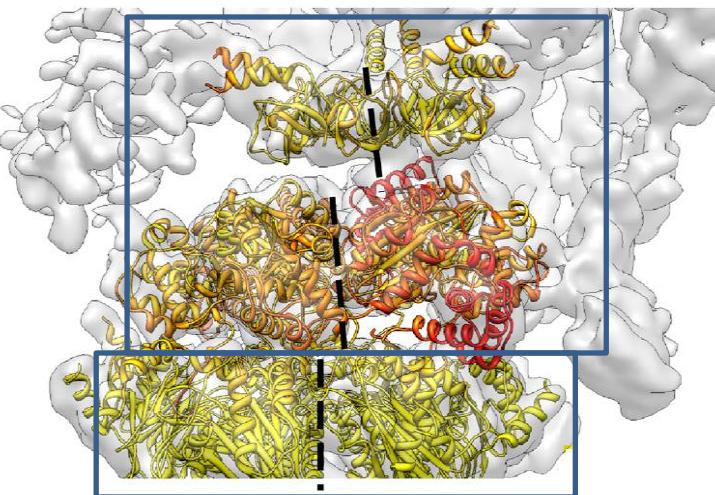


Clustering of the models in the ensemble

Jackknifing to validate the model obtained

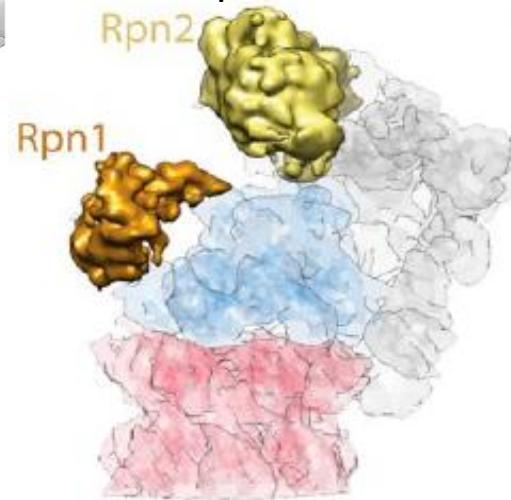
Modular Architecture of 26S Proteasome

Rpt 1-6 Subcomplex

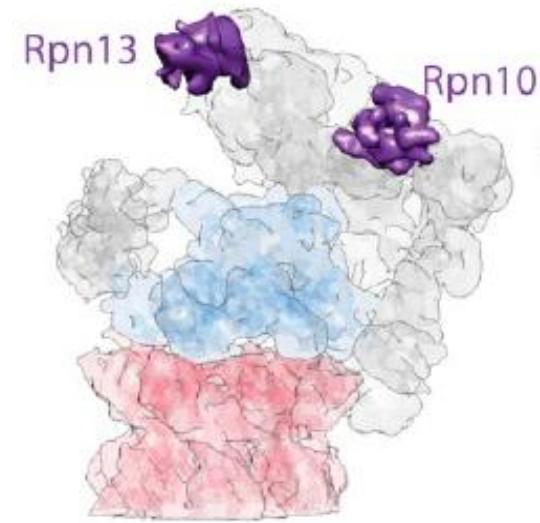


20S Core Proteasome ($\alpha_7 \beta_7$)

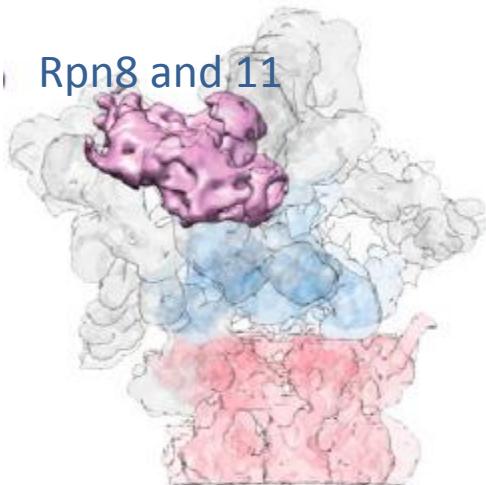
PC-repeat containing proteins



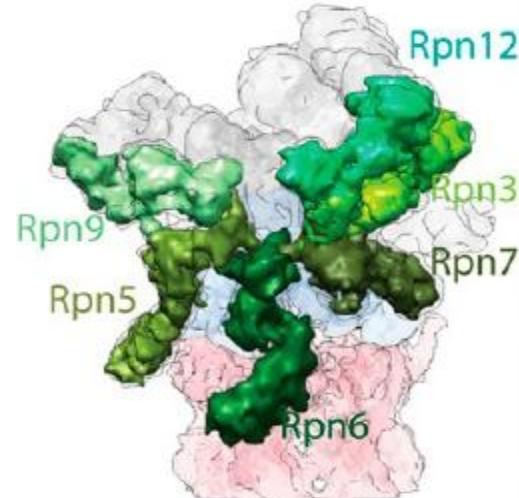
Ubiquitin Receptors



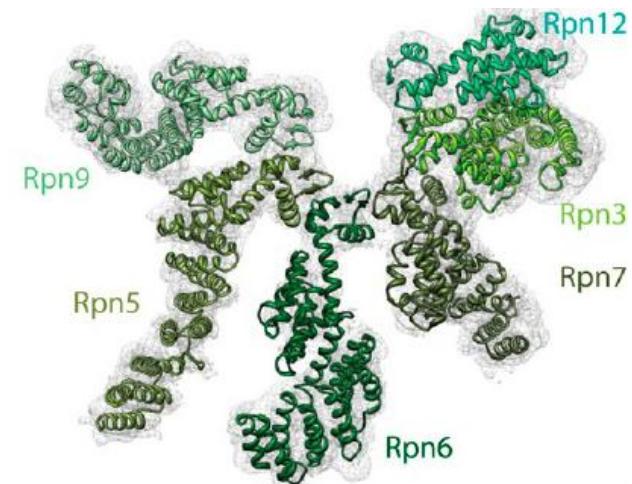
Rpn8 and 11



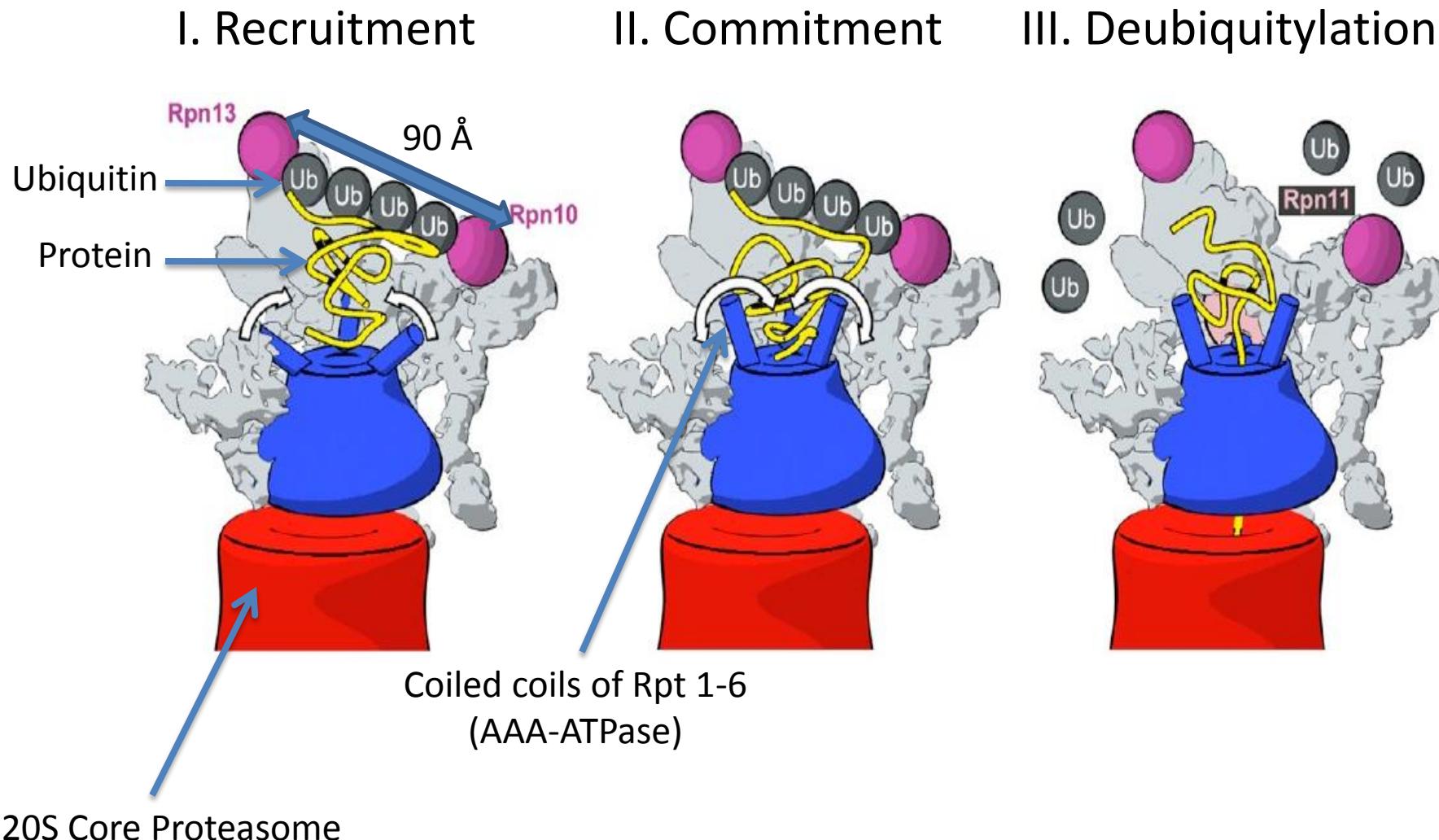
MPN containing proteins



PCI containing proteins in a heterohexamer

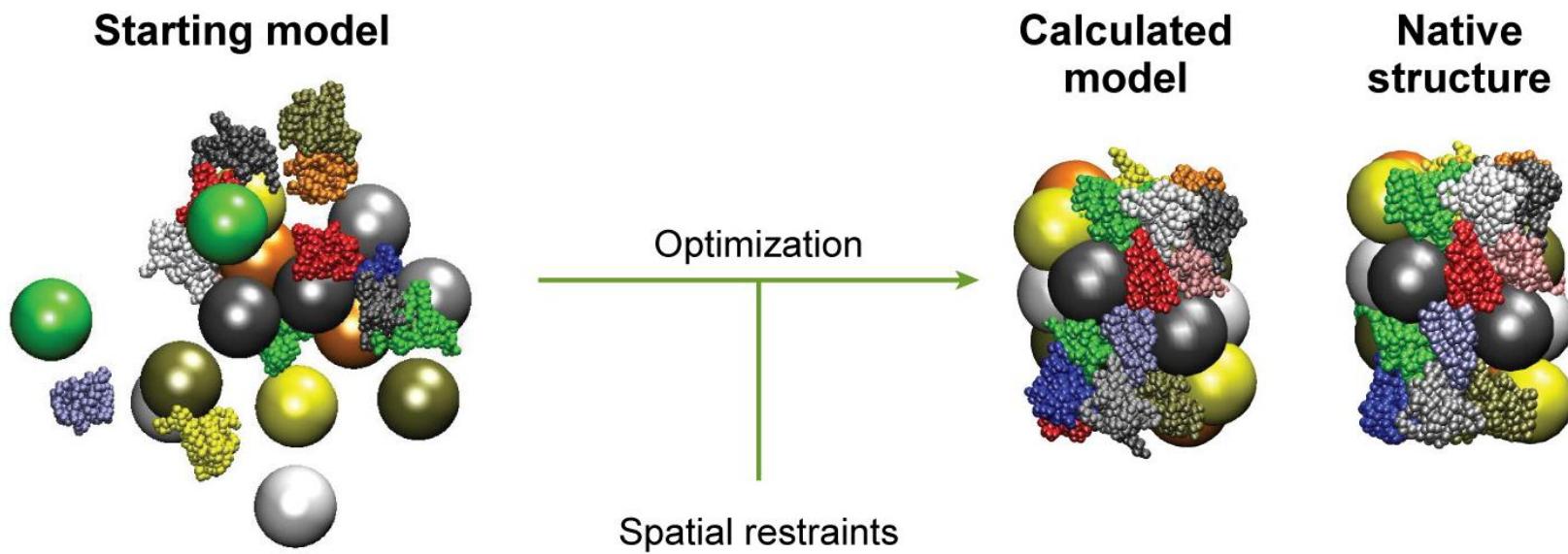


Mechanistic Model for Early Steps of Degradation

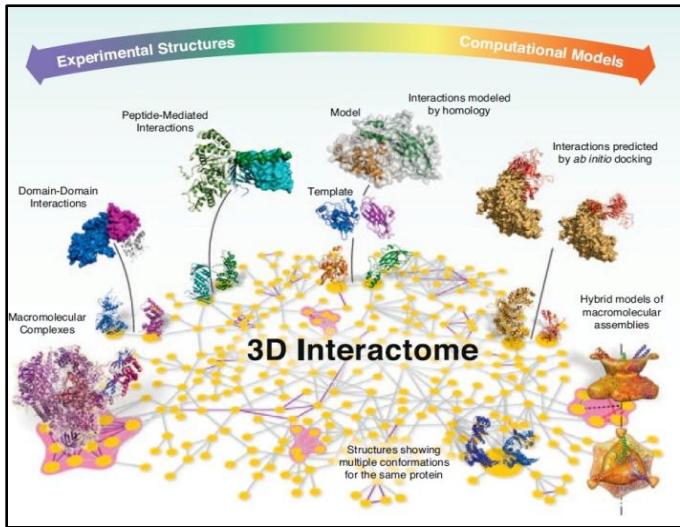


Summary

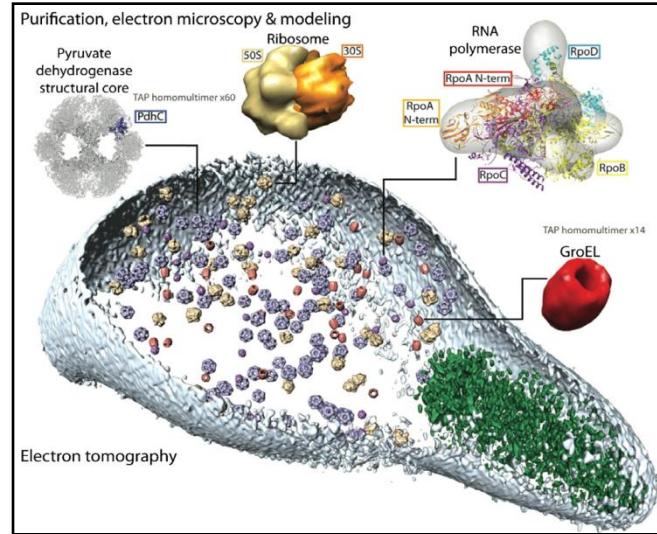
- Integrative Structure Determination computationally combines data from multiple sources to obtain models of assemblies with higher accuracy.



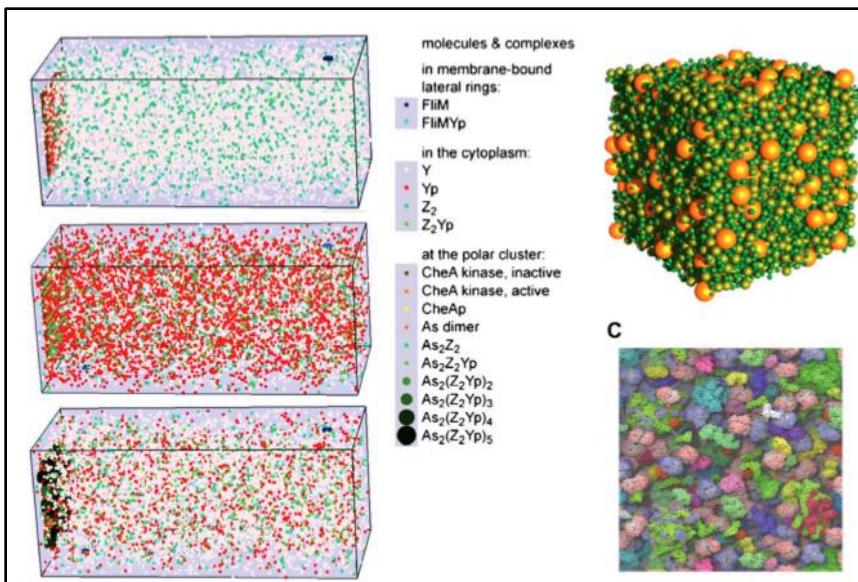
Future Challenges



Building 3D Interactome



Visual Proteomics



Modeling the Dynamic
Proteome Organization

(THANK YOU!!!)

