

Modeling the Nup84 Complex with IMP

Seth Axen and Hubert Nethercott
BMI 206, Fall 2014

Structural Characterization by Cross-linking Reveals the Detailed Architecture of a Coatomer-related Heptameric Module from the Nuclear Pore Complex*

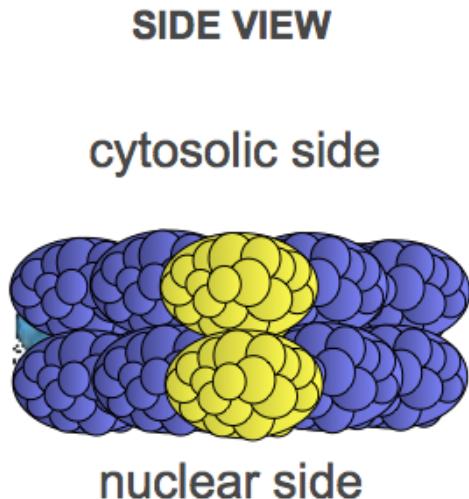
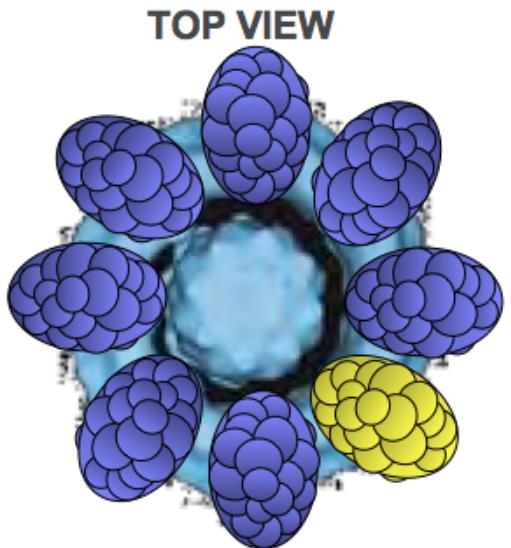
Yi Shi‡§, Javier Fernandez-Martinez§¶, Elina Tjioe§||, Riccardo Pellarin§||,
Seung Joong Kim§||, Rosemary Williams¶, Dina Schneidman-Duhovny||, Andrej Sali||**,
Michael P. Rout¶**, and Brian T. Chait‡**

Most cellular processes are orchestrated by macromolecular complexes. However, structural elucidation of these endogenous complexes can be challenging because they frequently contain large numbers of proteins, are compositionally and morphologically heterogeneous, can be dynamic, and are often of low abundance in the cell. Here, we present a strategy for the structural characterization of such complexes that has at its center chemical cross-linking with mass spectrometric readout. In this strategy, we isolate the endogenous complexes using a highly optimized sample preparation protocol and generate a comprehensive, high-quality cross-linking dataset using two complementary cross-linking reagents. We then determine the structure of the complex using a refined integrat-

flexibility of the assembly. These additional details further support and augment the protocoatomer hypothesis, which proposes an evolutionary relationship between vesicle coating complexes and the NPC, and indicates a conserved mechanism by which the NPC is anchored in the nuclear envelope. *Molecular & Cellular Proteomics* 13: 10.1074/mcp.M114.041673, 2927–2943, 2014.

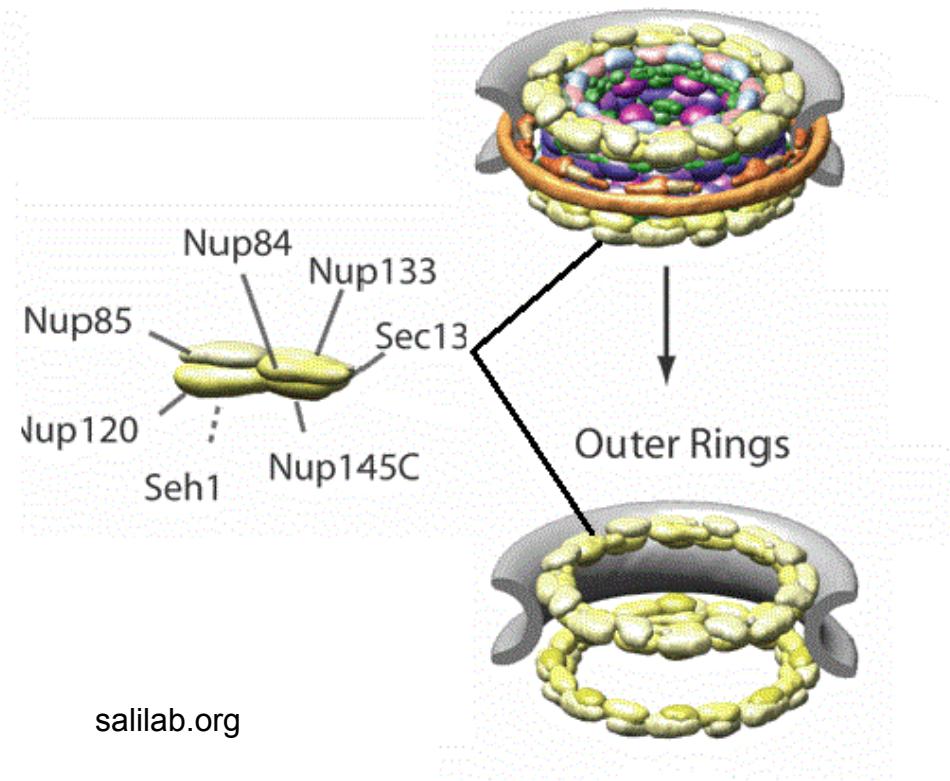
Macromolecular complexes are the building blocks that drive virtually all cellular and biological processes. In each eukaryotic cell, there exist many hundreds of these protein complexes (1–3), the majority of which are still poorly under-

Nuclear pore complex



In *S. cerevisiae*
~480 Nups in NPC:
-8 half-spokes x 2 “sides”
-30 Nups per half-spoke

Heptameric Nup84 complex



-transcriptional organization

-telomere tethering in nuclear periphery (dsDNA repair)

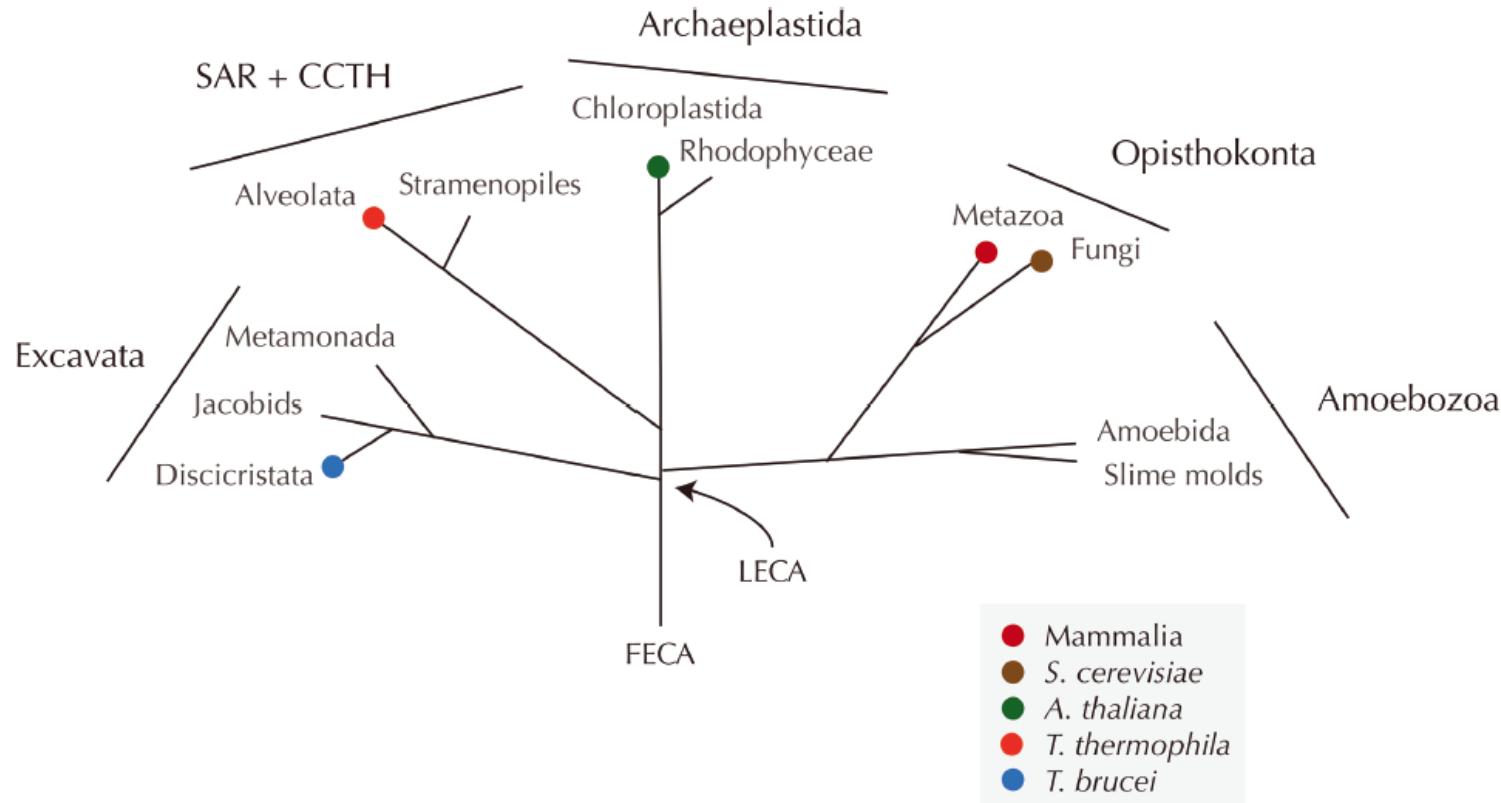
-nuclear trafficking

2 Goals of the Paper:

1. Present a workflow for integrative modeling of Nup84 complex and other macromolecular structures.
2. Test and present data in support of the protocoatomer hypothesis

Protocoatomer Hypothesis (Devos et al., 2004)

The Nup heptamers serve as “membrane curving modules” and are functionally and structurally similar to Vesicle-coating complexes (VCC) clathrin, COPI, and COPII

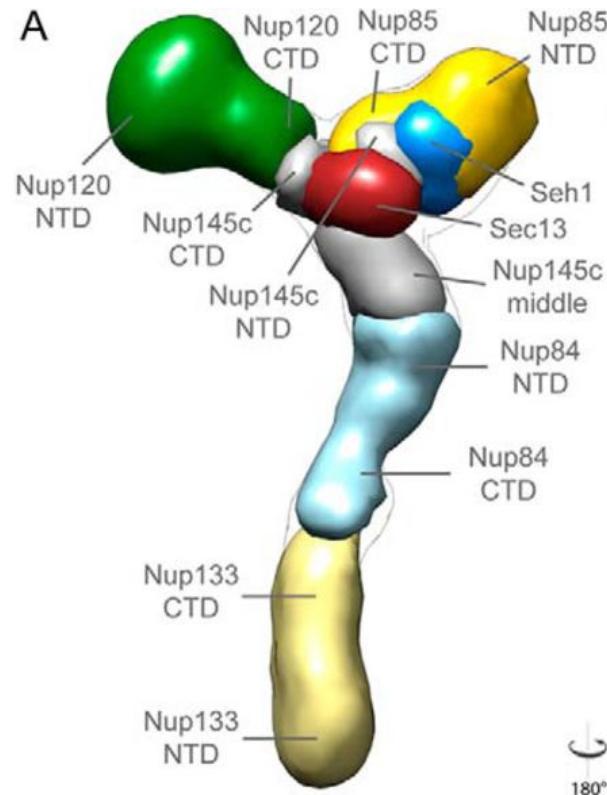


MC Field et al., 2014

Results: Protocoatomer hypothesis

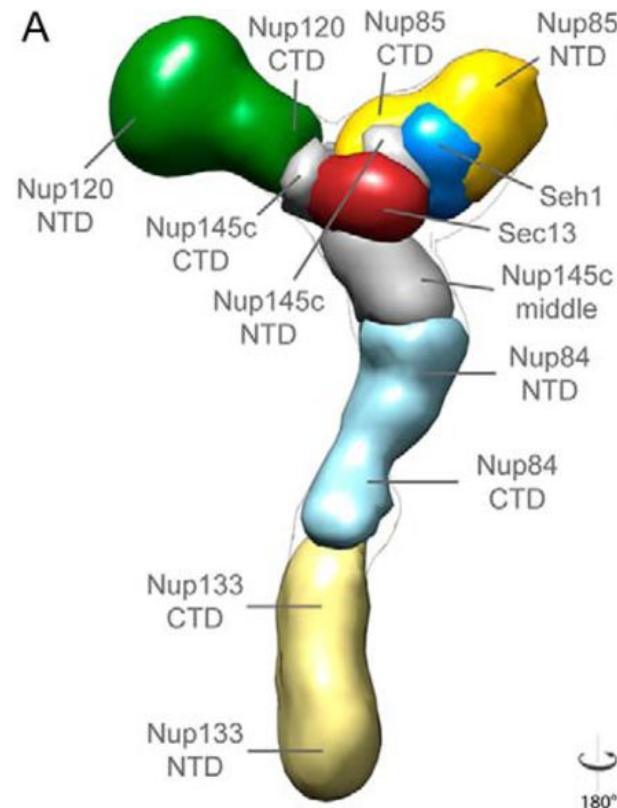
-This method detects an ALPS motif in Nup133: senses membrane curvature; then anchors itself to pore membrane

-C-terminal trimeric interaction between Nup145c, Nup120, and Nup85 in the hub (similar to C-terminal interaction of the clathrin triskelion)



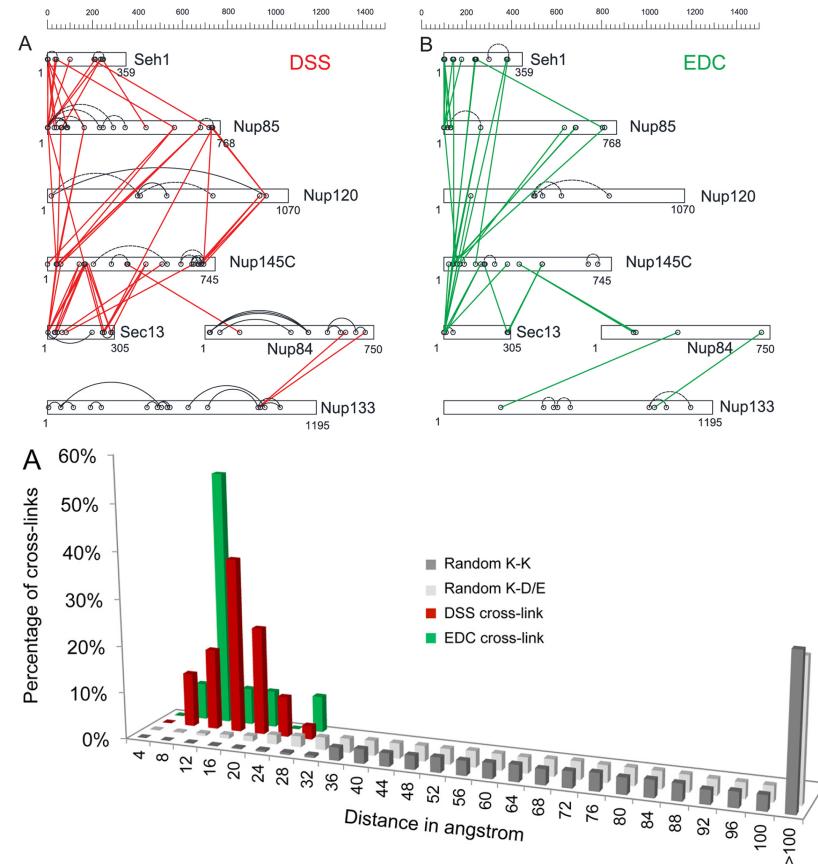
Results: more detail in hub region

-Heterodimeric interaction
between beta-propeller proteins
Sec 13 and Seh1 (similar to COPI
and COPII assembly)



Cross-linking/LC-MS

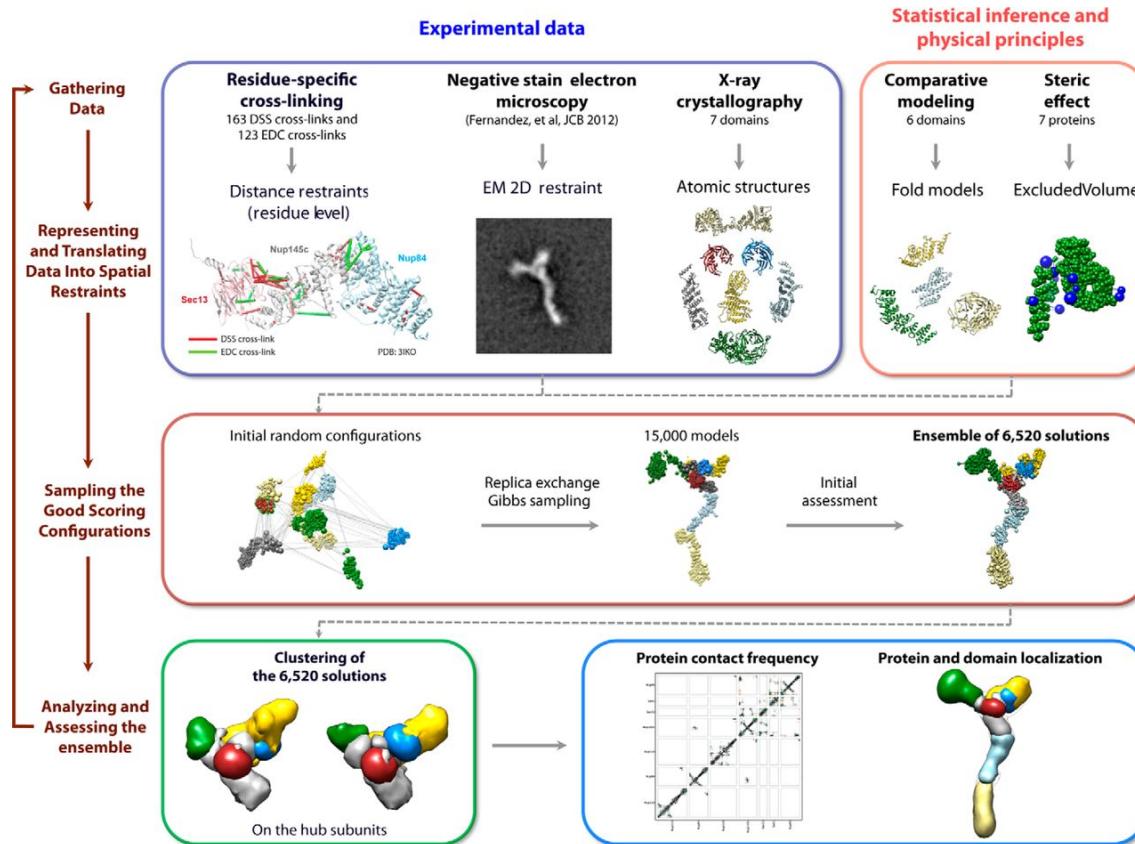
1. DSS (11.4-30 Å spacer) and EDC (0-17 Å)
2. Crosslinks analyzed by LC-MS
 - 163 unique DSS and 104 EDC cross-links identified



Spatial Restraints

- EM two-dimensional restraint
- restraints from the cross-linking data
- excluded volume restraint: representative volume of each bead
- sequence connectivity restraint: “flex room” of 4x radius between two beads

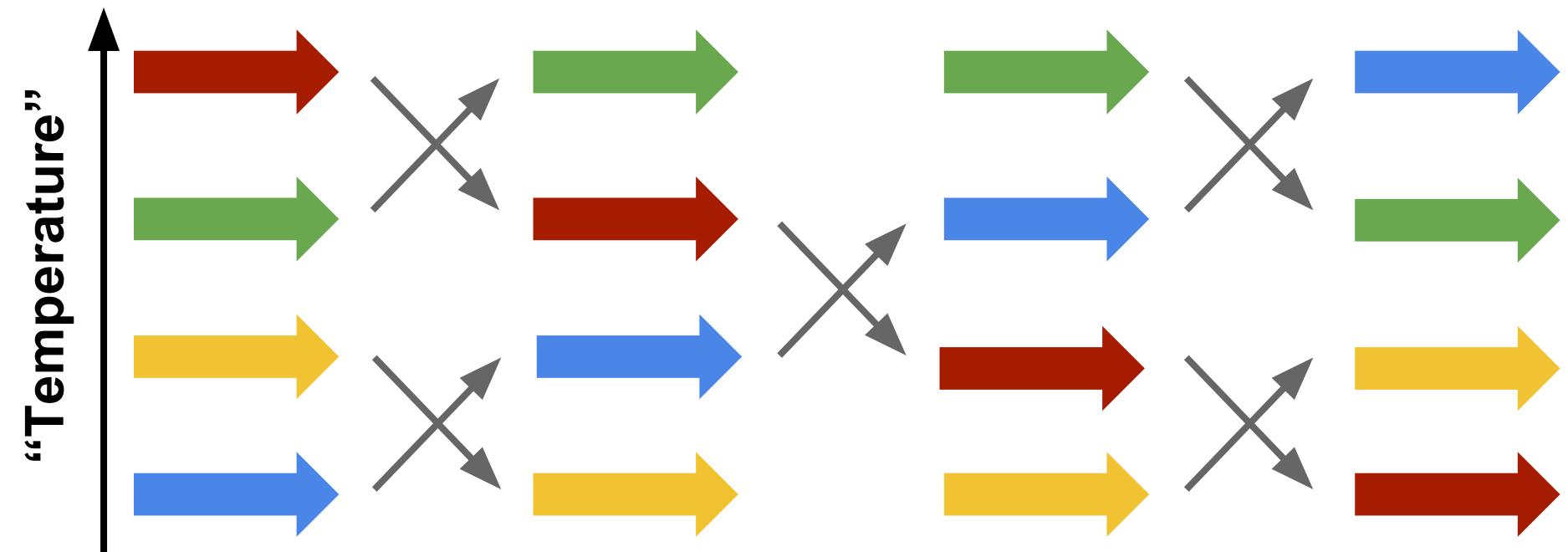
Workflow



Running IMP

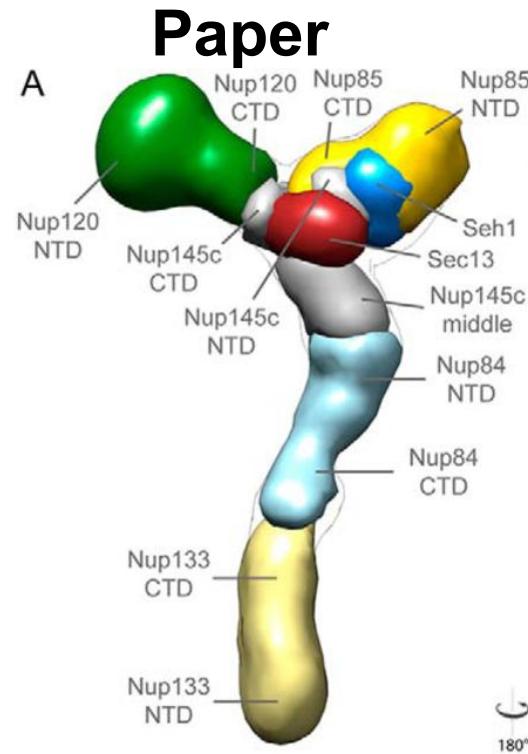
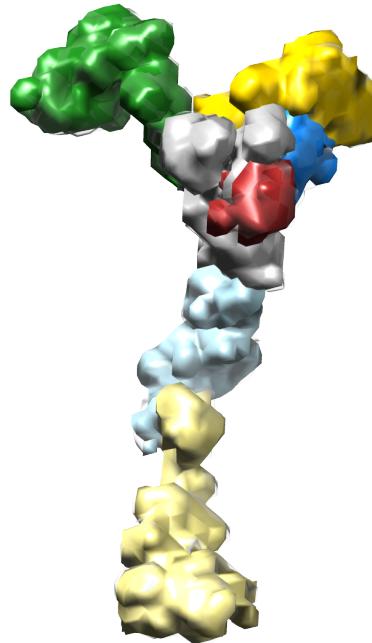
- Removed crystallographic interface constraints (for efficiency)
- Replica exchange over 4-6 nodes
- Three Replicates with 317-682 structures each (total 1680)

Replica Exchange Improves Sampling



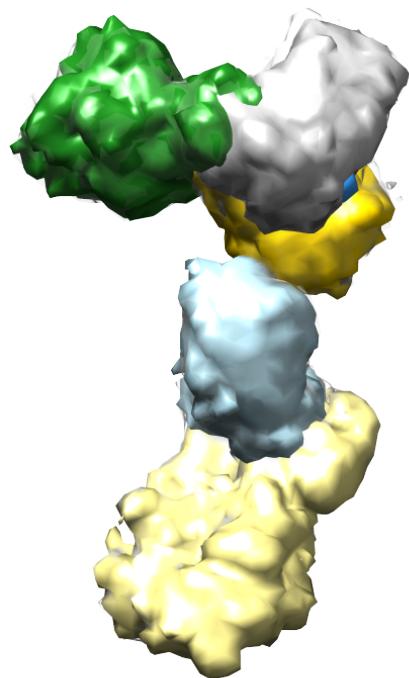
Top 4 Solutions Resemble Paper Ensemble

Run 1 (top 4/681)

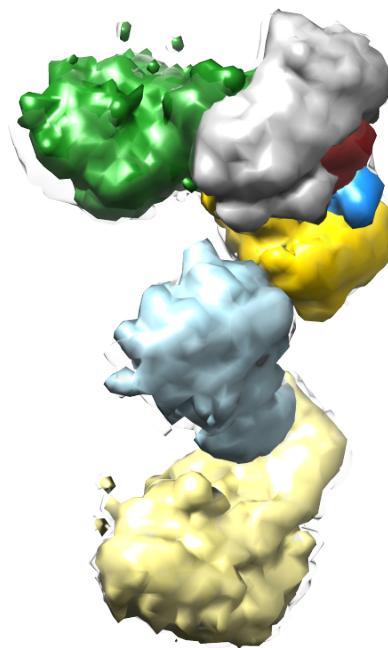


Top 100 Run 1 Solutions Clustered (K-means)

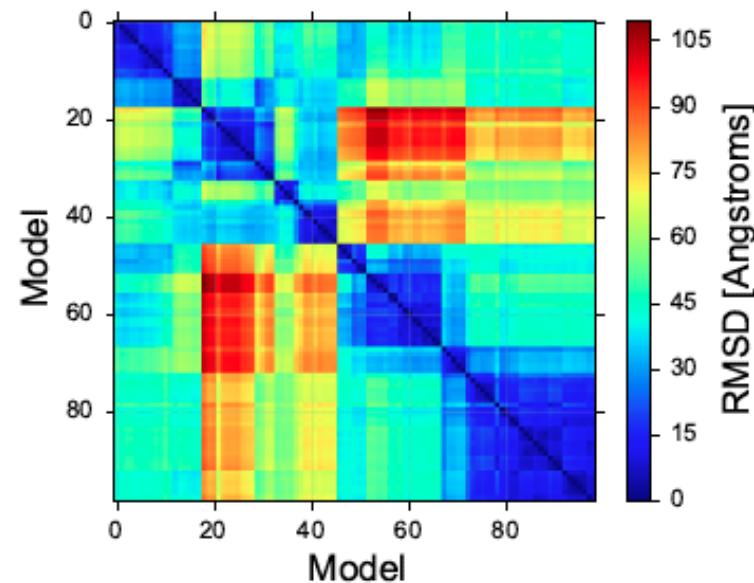
Cluster 1 (65/100)



Cluster 2 (35/100)

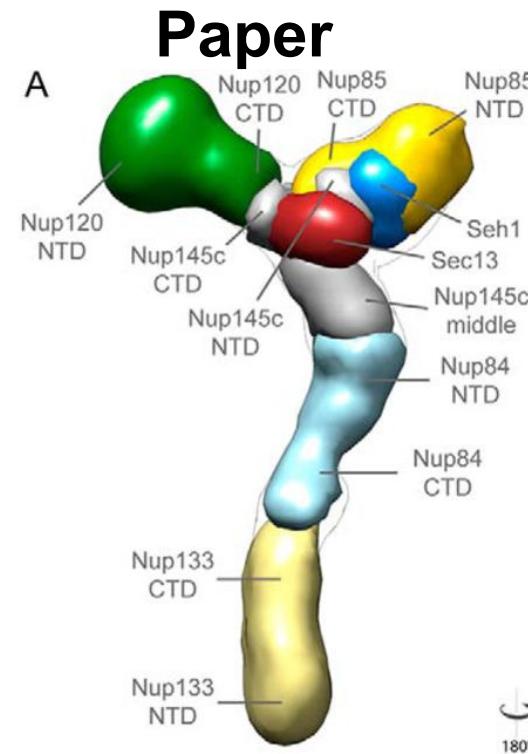
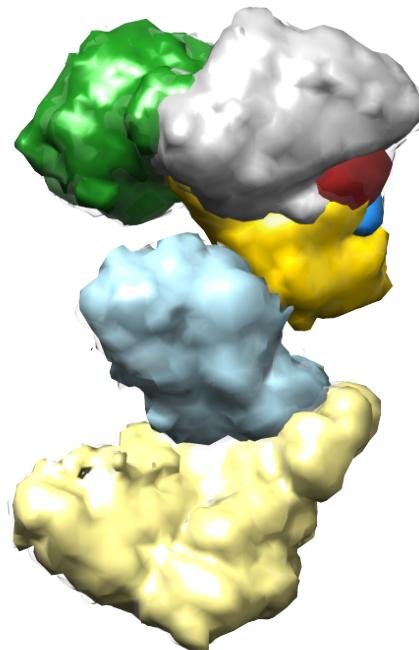


RMSD Matrix



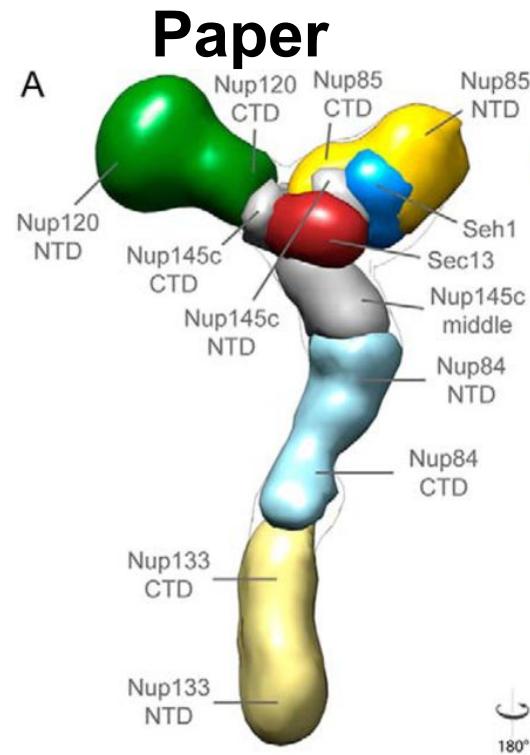
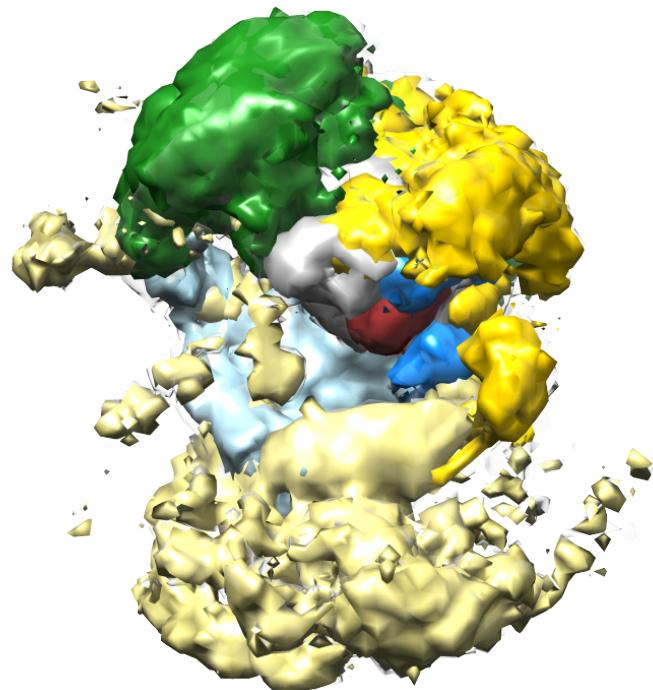
Top 250 Run 1 Solutions Look Less Like Paper

Density (top 250/681)



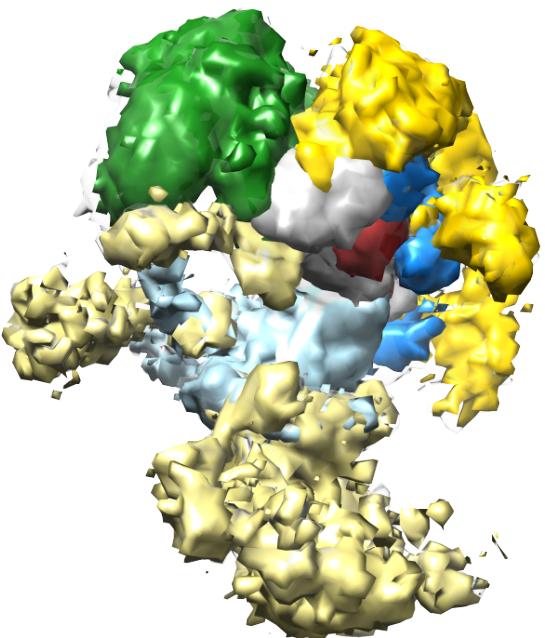
But it could be worse (Run 2)

Density (top 250/317)

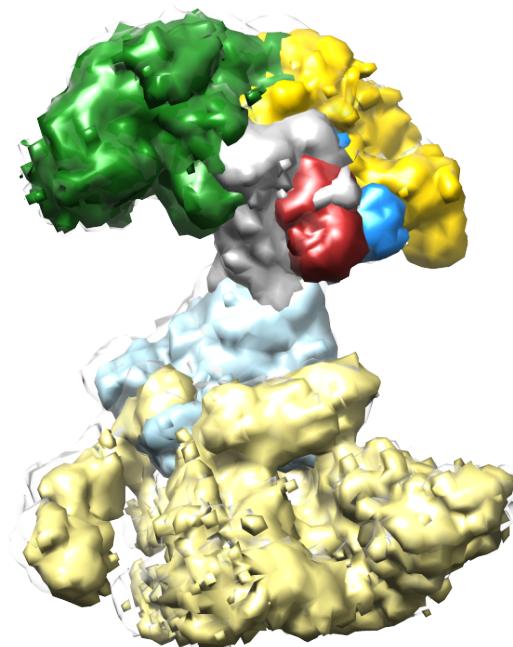


Top 100 Run 2 Solutions Clustered (K-means)

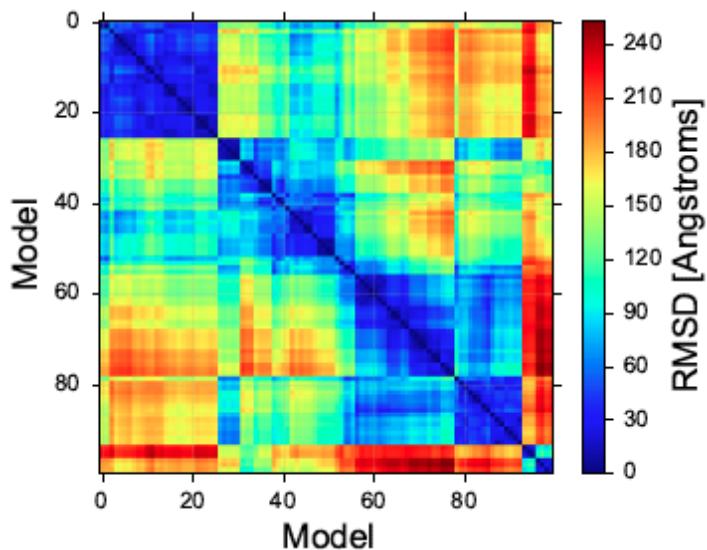
Cluster 1 (56/100)



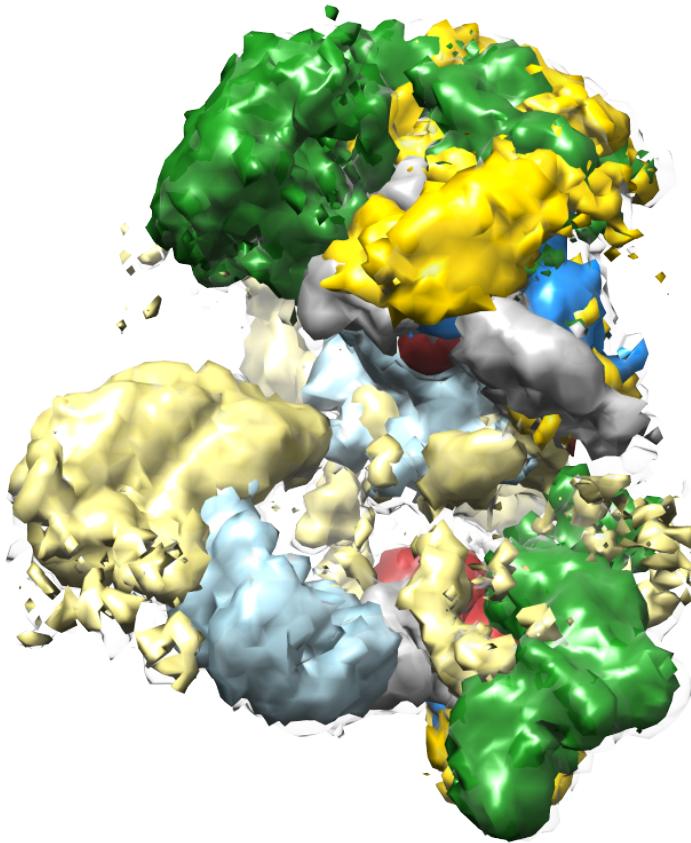
Cluster 2 (44/100)



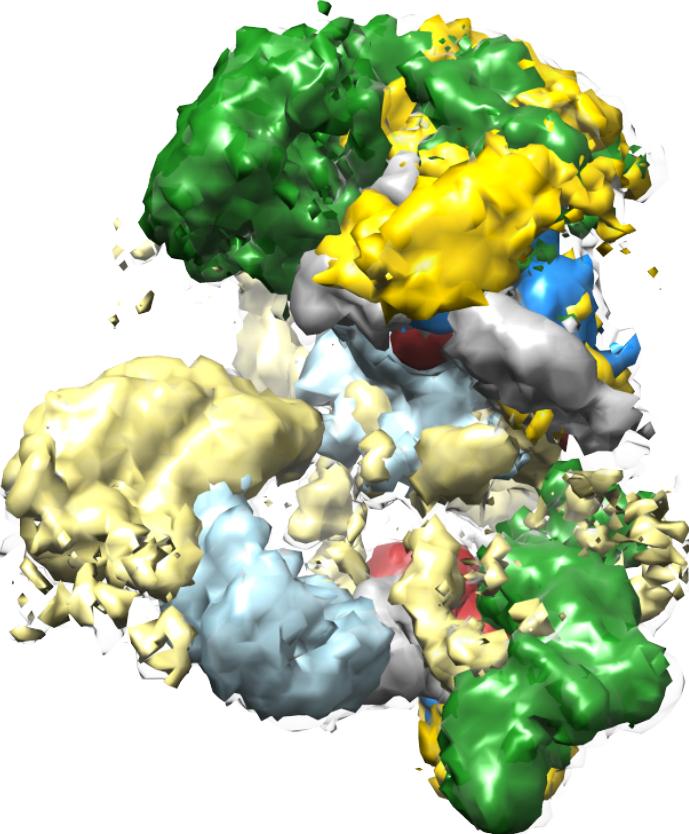
RMSD Matrix



All replicate runs merged (top 250/1680)



Why so poor?

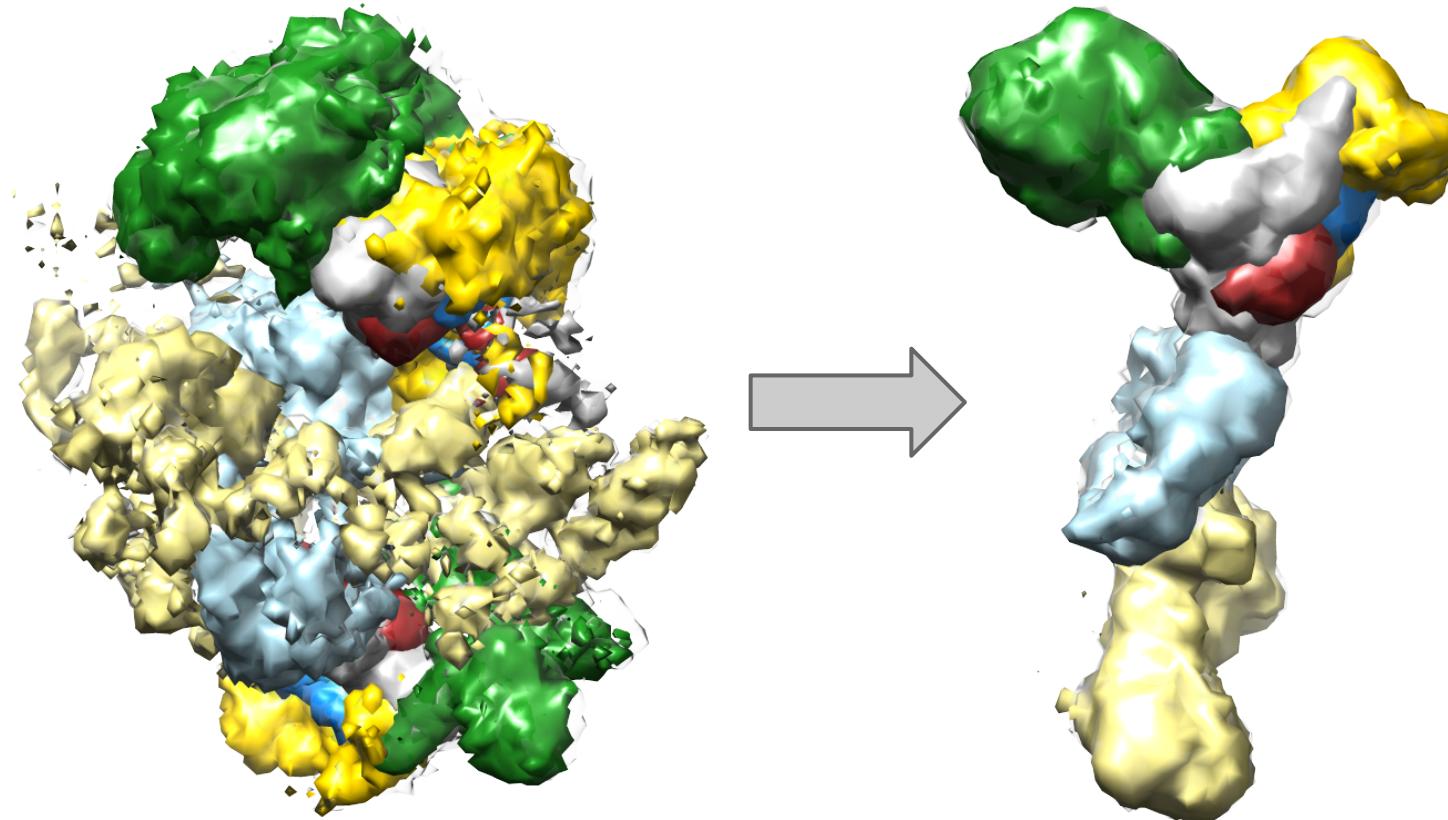


- 2D EM constraints a y-shaped silhouette, but not orientation
- IMP Clustering only uses Root Mean Square Deviation (RMSD)

$$RMSD(\mathbf{A}, \mathbf{B}) = \sqrt{\frac{1}{N} \sum_i^N d(a_i, b_i)^2}$$

- Provided scripts don't align solutions

Impact of Alignment (top 100/1680)



RMSD vs dRMSD

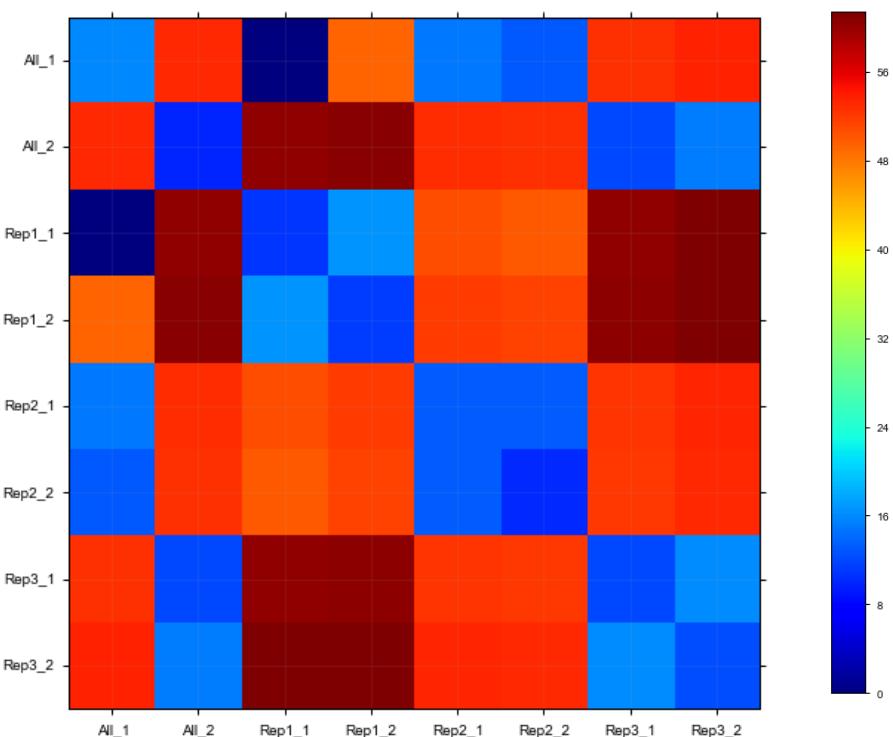
$$RMSD(\mathbf{A}, \mathbf{B}) = \sqrt{\frac{1}{N} \sum_i^N d(a_i, b_i)^2}$$

$$dRMSD(\mathbf{A}, \mathbf{B}) = \sqrt{\frac{2}{N(N-1)} \sum_{i < j} [d(a_i, a_j) - d(b_i, b_j)]^2}$$

- Requires 3D alignment of structures
- Allows subsequent density calculations
- Does not require 3D alignment
- Can only be used for sets of structures
- Allows calculation of cross-precision

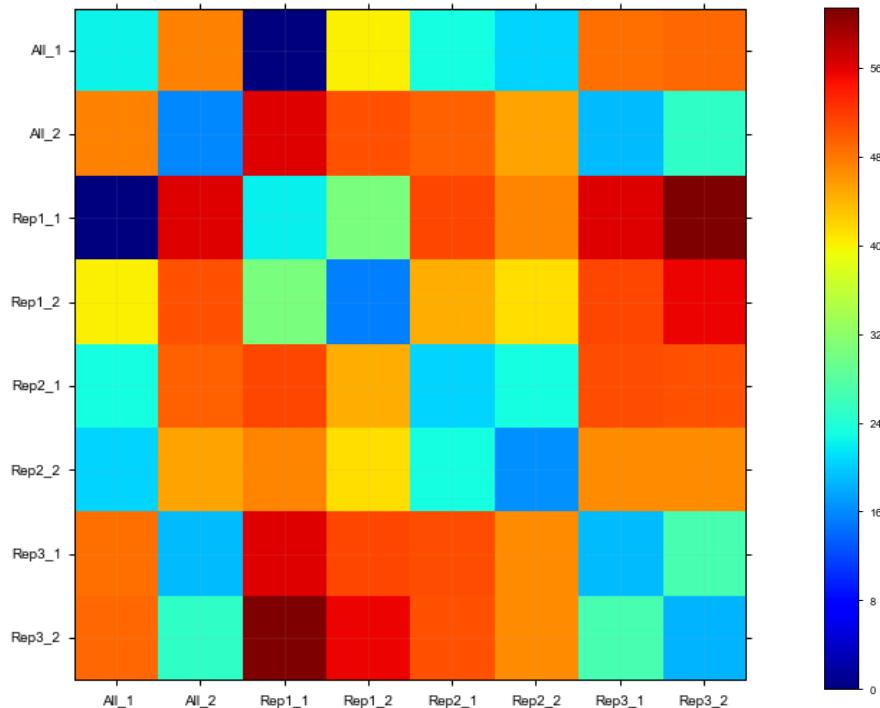
dRMSD Precision of Nup84 Complex

- Precision: average pairwise dRMSD within/between all solutions in a group
- If sufficient sampling (convergence) and good clustering, expect a one-to-one correspondence between clusters from different runs



dRMSD Precision of Nup84 “Hub”

- Precision: average pairwise dRMSD within/between all solutions in a group
- If sufficient sampling (convergence) and good clustering, expect a one-to-one correspondence between clusters from different runs

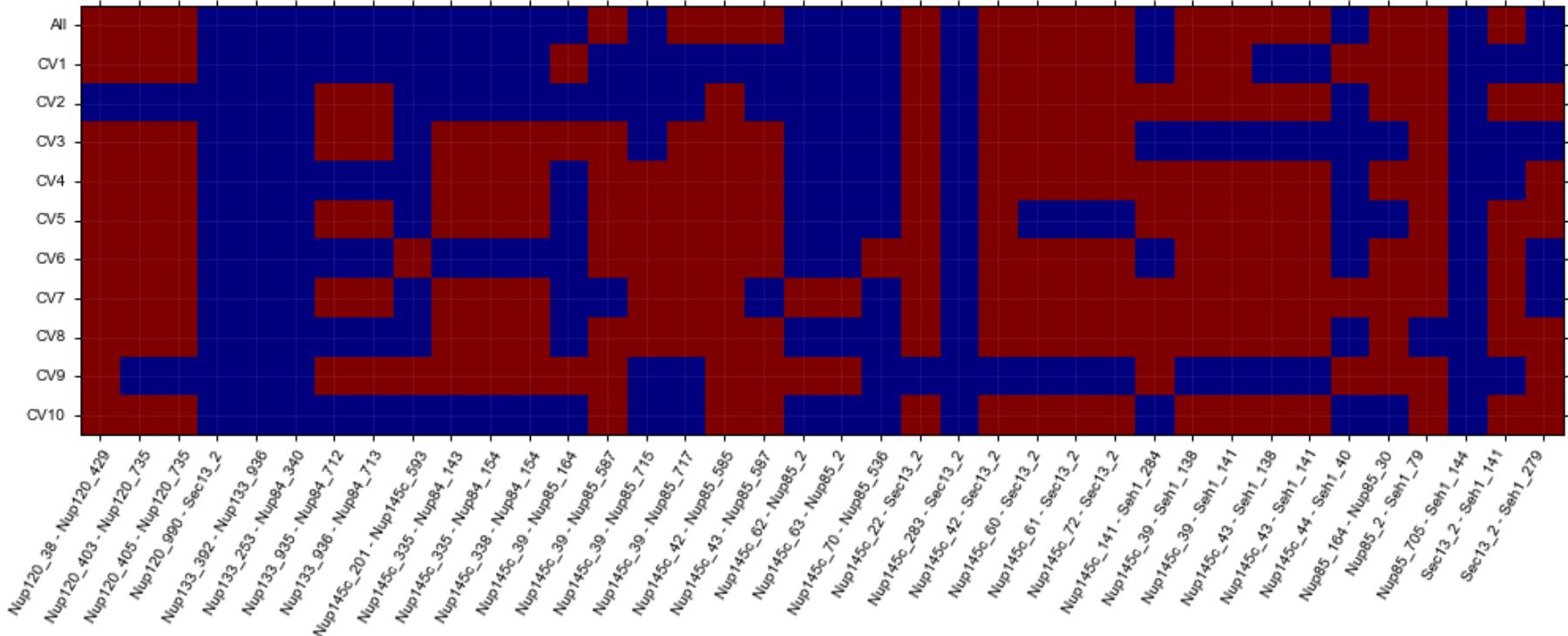


Results: Cross-Validation of Cross-Linkers

- 15% of cross-links not satisfied by over half of top 250 solutions
- Divided cross-links into 10 random sets
- Generated >250 solutions
- Checked if cross-links satisfied

Results: Cross-Validation of Cross-Linkers

On average 96% of left-out cross-links are satisfied



Future Directions

- Include 3D EMs as constraint (Kampmann M, Blobel G, 2007)
- Use frequency of observing cross-link to weight more common cross-links more highly
- Generate more solutions
- Align all solutions
- Cross-validation/learning curve by subsetting solutions
- Learning curve

Acknowledgments

Katie Pollard

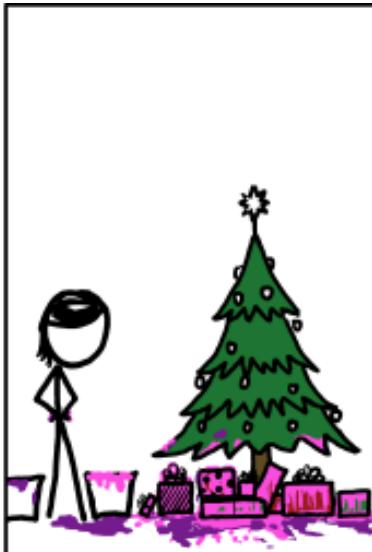
Andrej Sali

Elina Tijoe

SJ Kim

Ben Webb

Happy Holidays!!



xkcd