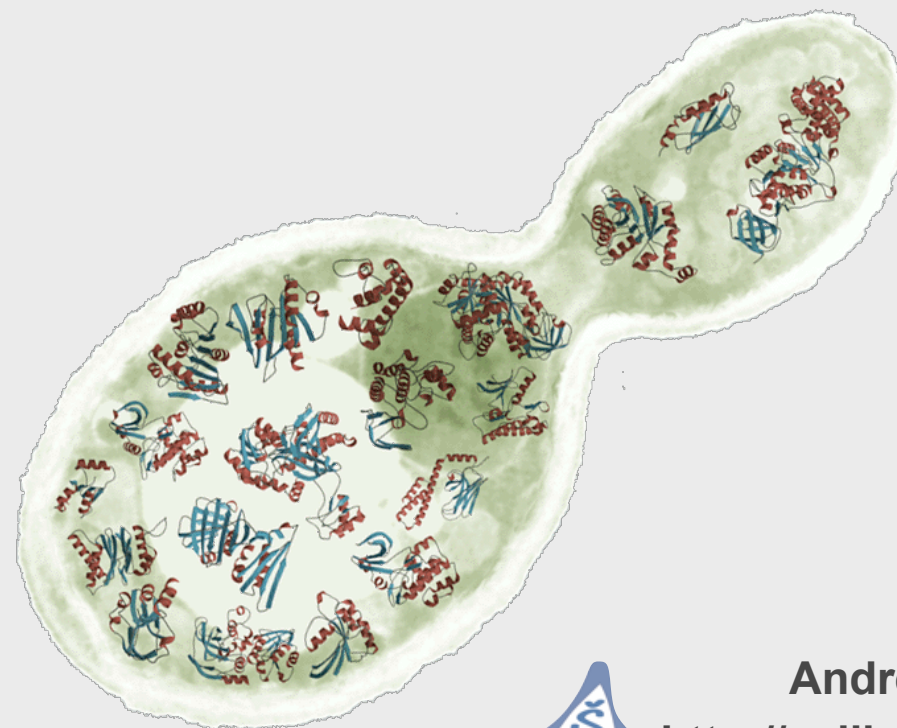
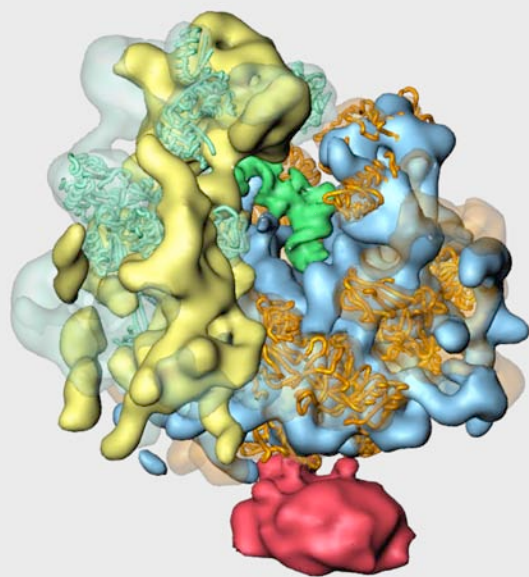
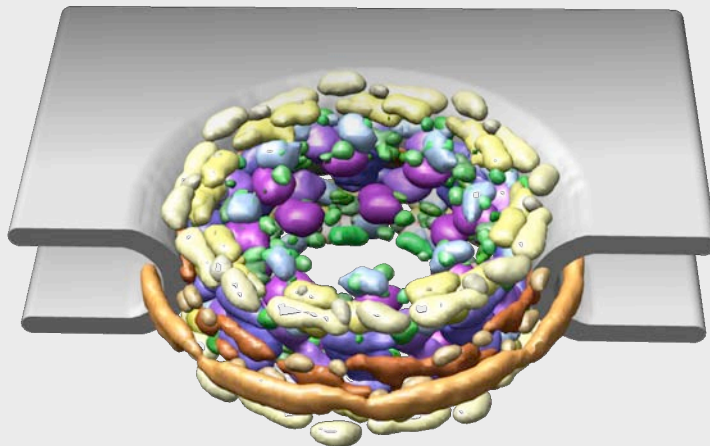


# Integrative structure determination at pseudo-atomic resolution



Andrej Sali  
<http://salilab.org/>

UCSF

qb3  
ucb ucsc-ucsf

Department of Bioengineering and Therapeutic Sciences  
Department of Pharmaceutical Chemistry  
California Institute for Quantitative Biosciences  
University of California at San Francisco

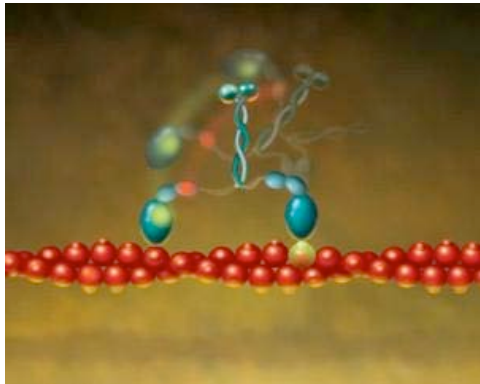
# Topics

1. Introduction to integrative (hybrid) structure determination
2. Comparative model building
3. Predicting accuracy of atomic models
4. Iterative sequence-structure alignment and model building
5. Electron microscopy
6. Small angle x-ray scattering
7. Proteomics
8. Concluding Remarks

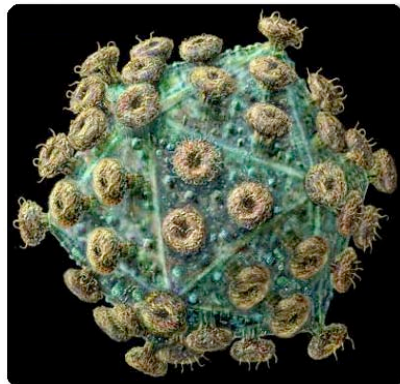
# Immediate Goal

Maximize **accuracy**, **resolution**, **completeness**, and **efficiency** of the structural coverage of proteins and their assemblies (static structures).

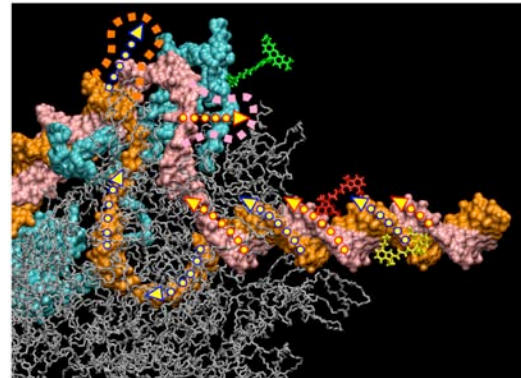
Motivation: Structures will allow us to understand how machines work, how they evolved, how they can be controlled, modified, and perhaps even designed.



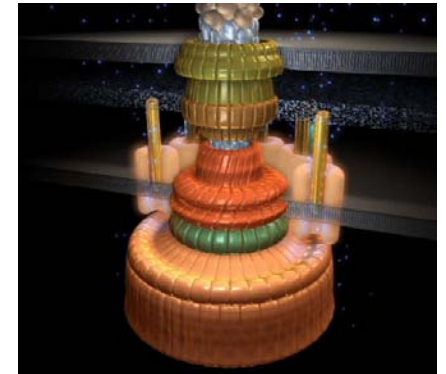
kinesin



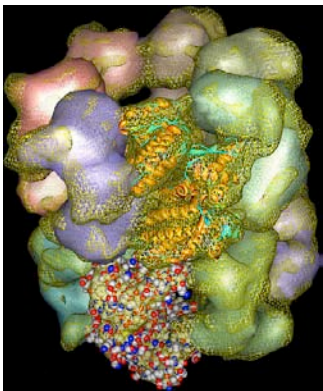
virus



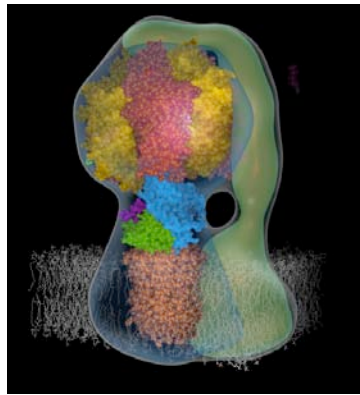
tRNA synthetase



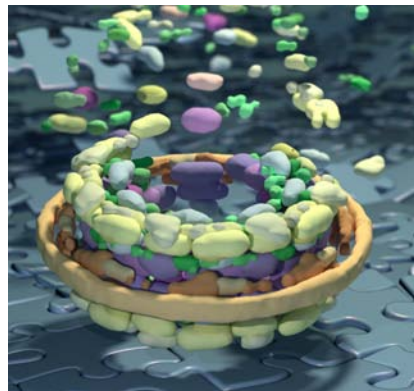
flagellar motor



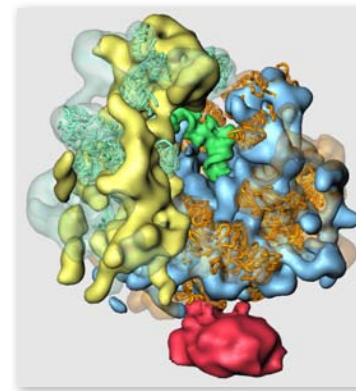
GroEL chaperonin



ATP synthase



nuclear pore complex

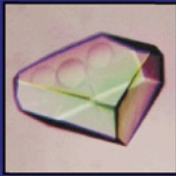
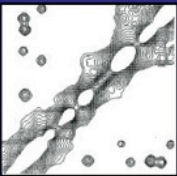

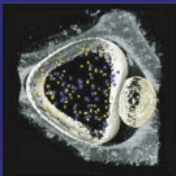
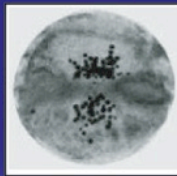
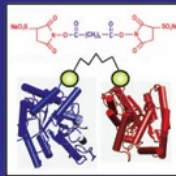



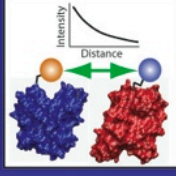
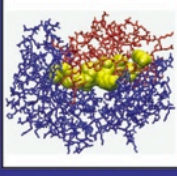
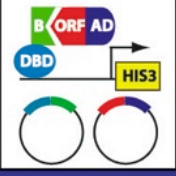


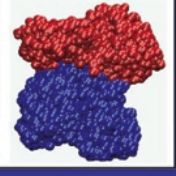
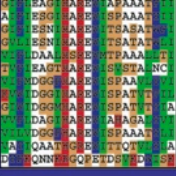
ribosome

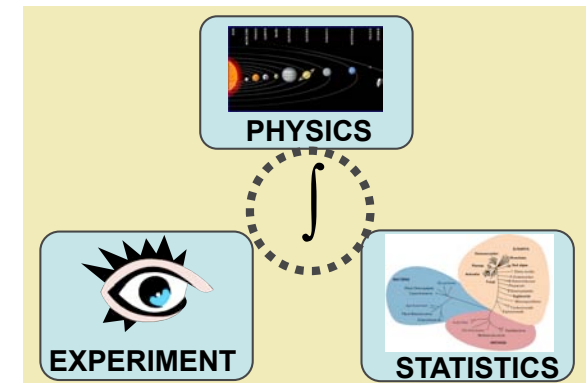
There are thousands of biologically relevant macromolecular complexes whose structures are yet to be characterized, involved in a few hundred core biological processes.

# Integrative determination of macromolecular structures for maximizing accuracy, resolution, completeness, and efficiency

Use structural information from any source: measurement, first principles, rules; resolution: low or high resolution to obtain the set of all models that are consistent with it.

						
<b>X-ray crystallography</b>	<b>NMR spectroscopy</b>	<b>2D &amp; single particle electron microscopy</b>	<b>electron tomography</b>	<b>immuno-electron microscopy</b>	<b>chemical cross-linking</b>	<b>affinity purification mass spectroscopy</b>
subunit structure	subunit structure				subunit structure	
subunit shape	subunit shape	subunit shape	subunit shape			
subunit-subunit contact	subunit-subunit contact	subunit-subunit contact	subunit-subunit contact		subunit-subunit contact	subunit-subunit contact
subunit proximity	subunit proximity	subunit proximity	subunit proximity	subunit proximity	subunit proximity	subunit proximity
subunit stoichiometry	subunit stoichiometry					
assembly symmetry	assembly symmetry	assembly symmetry	assembly symmetry	assembly symmetry		
assembly shape	assembly shape	assembly shape	assembly shape			
assembly structure	assembly structure					

						
<b>FRET</b>	<b>site-directed mutagenesis</b>	<b>yeast two-hybrid system</b>	<b>gene/protein arrays</b>	<b>protein structure prediction</b>	<b>computational docking</b>	<b>bioinformatics</b>
				subunit structure		
				subunit shape		
subunit-subunit contact	subunit-subunit contact	subunit-subunit contact	subunit-subunit contact		subunit-subunit contact	subunit-subunit contact
subunit proximity		subunit proximity	subunit proximity			



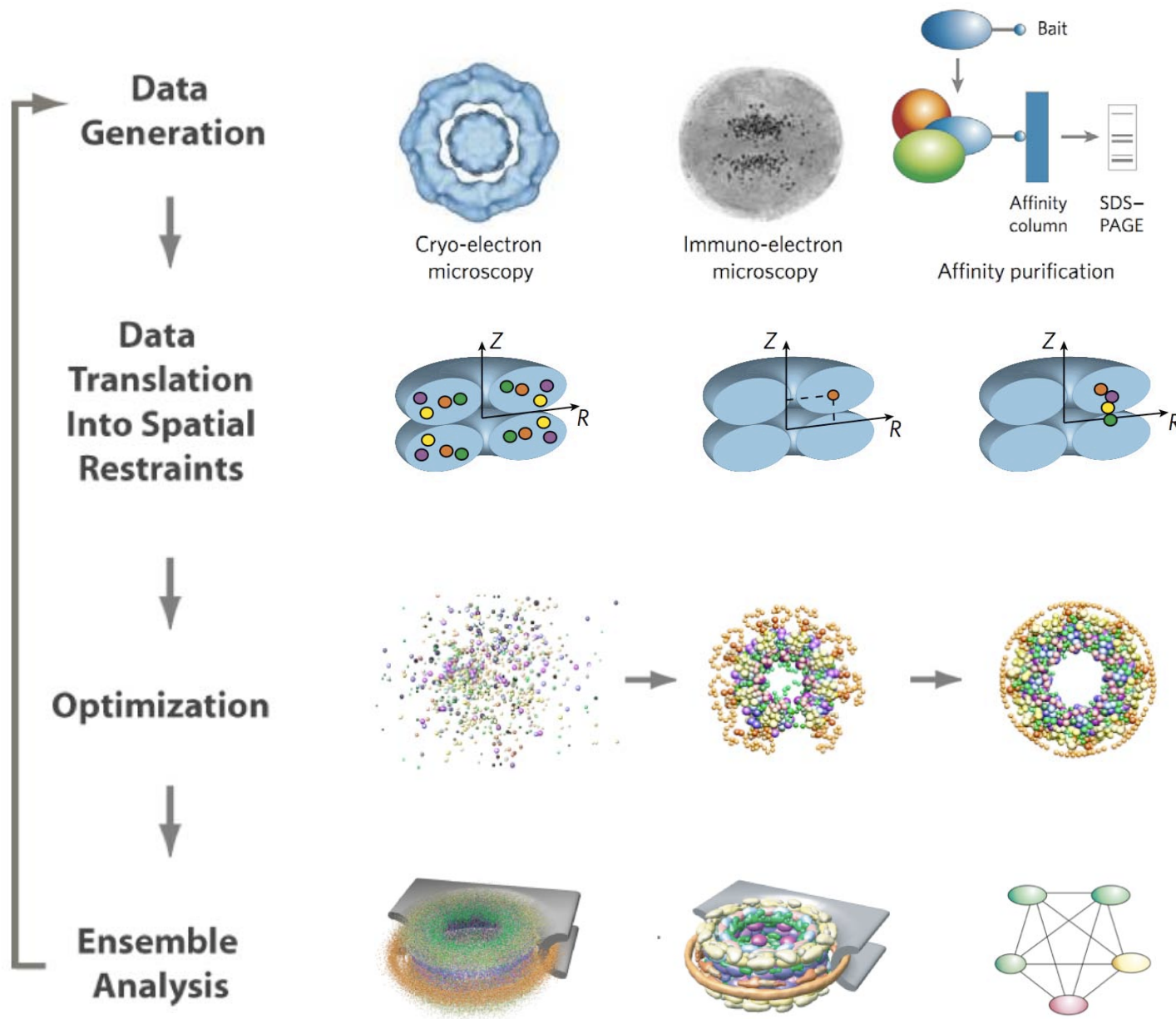
Sali, Earnest, Glaeser, Baumeister. From words to literature in structural proteomics. *Nature* 422, 216-225, 2003.

# Integrative determination of macromolecular structures

Alber *et al.* *Nature* **450**, 683-694, 2007.

Robinson, Sali, Baumeister. *Nature* **450**, 974-982, 2007.

Alber, Foerster, Korkein, Topf, Sali. *Annual Reviews in Biochemistry* **77**, 11.1–11.35, 2008.

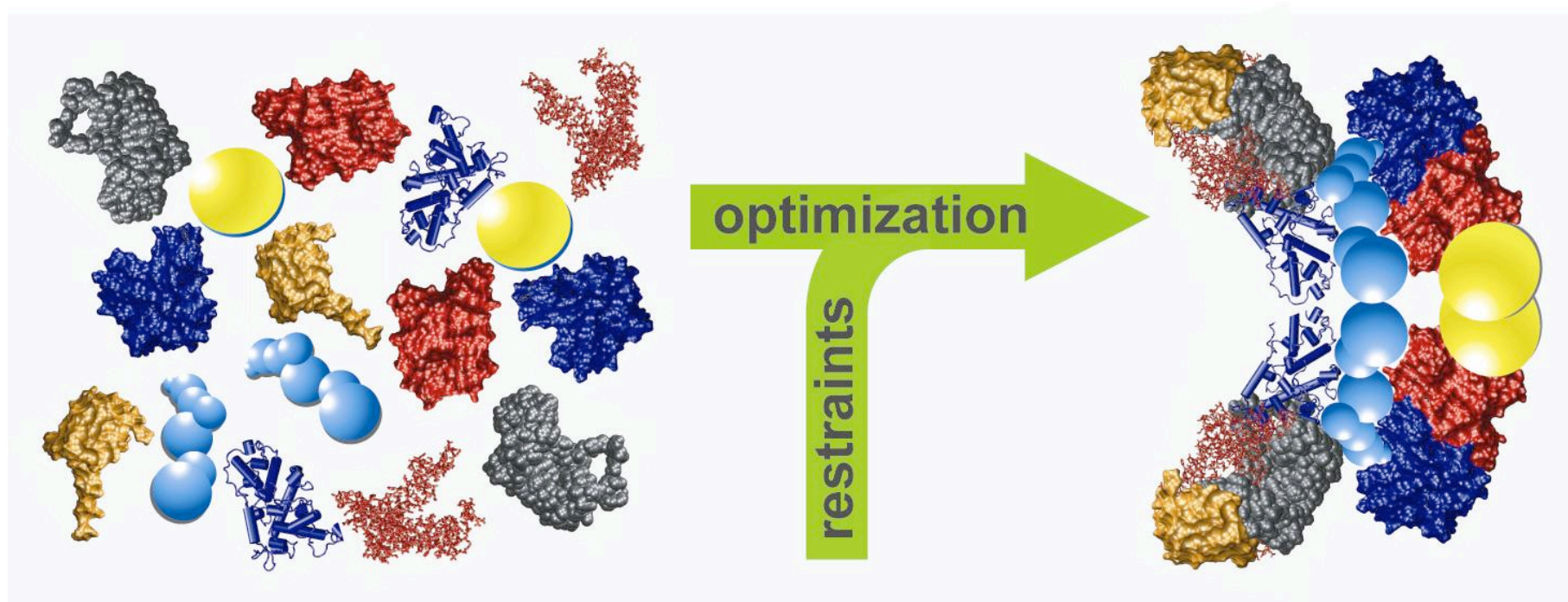


# Characterizing Structures by Satisfaction of Spatial Restraints

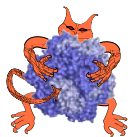
1. Representation of a system.
2. Scoring function (spatial restraints).
3. Optimization / sampling.

There is nothing but points and restraints on them. We seek joint pdf for  $\mathbf{R}$ , given information  $\mathbf{I}$ :

$$P(\mathbf{R} / \mathbf{I}) \approx \prod_i p_i(\mathbf{r}_i / \mathbf{I}_i)$$

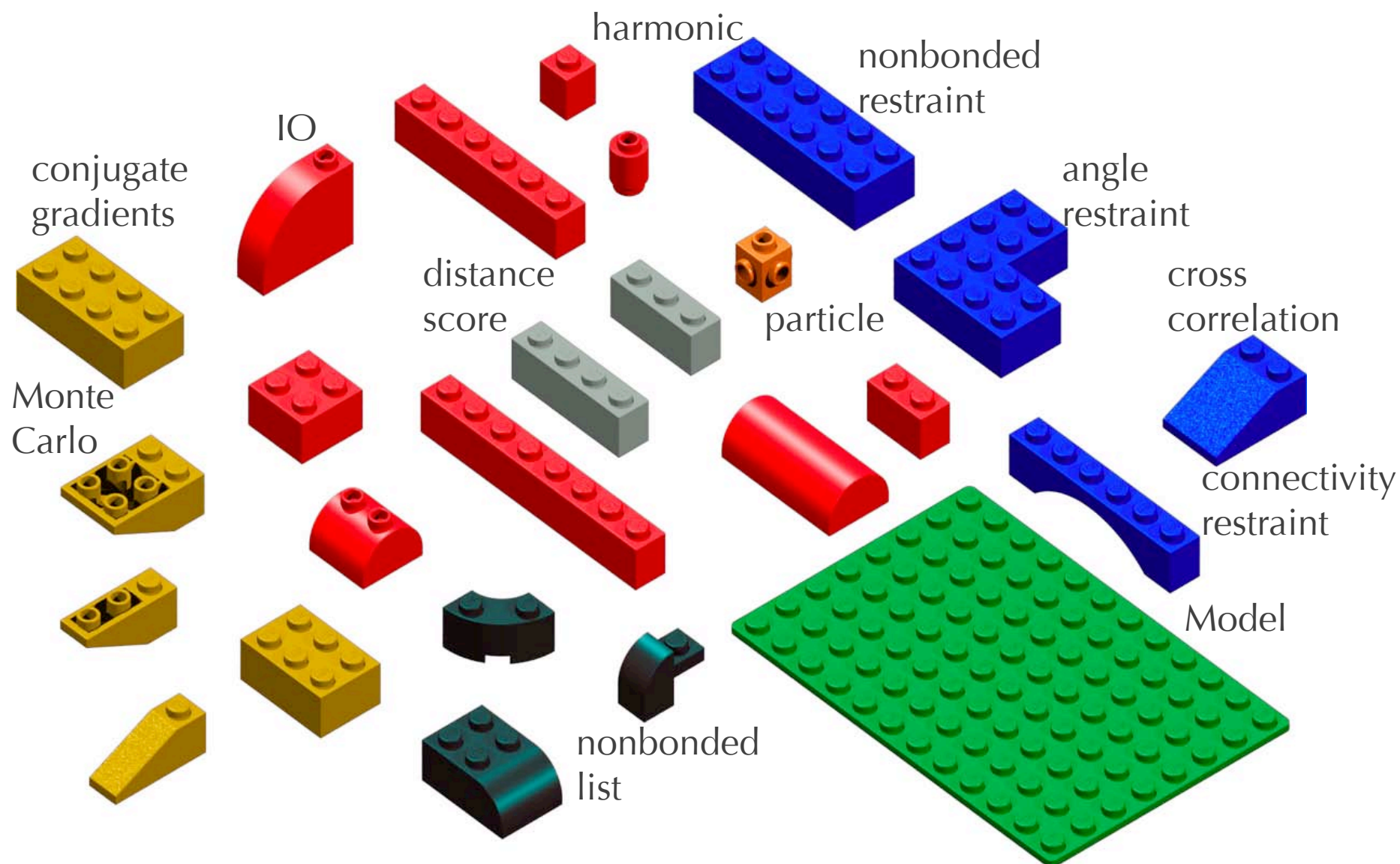


*Integrative Modeling Platform (IMP):* <http://salilab.org/imp>

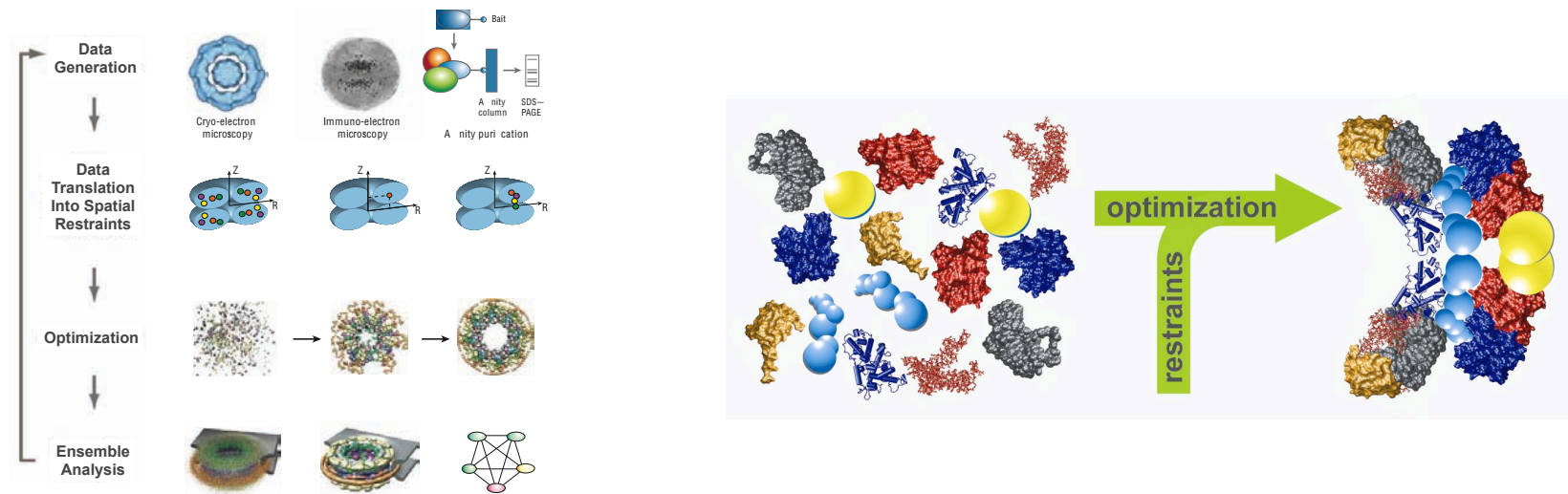


# Integrative Modeling Platform (IMP): Building blocks for structural modeling

<http://salilab.org/imp/>



# Why Integrative Modeling?



1. Benefits from the **synergy** among the input data, maximizing accuracy, resolution, completeness, and efficiency of structure characterization.
2. Finds “**all**” models consistent with the data, not just one.
3. Facilitates **assessing** the data and results in terms of precision and accuracy.
4. Provides feedback to **guide** future experiments (eg, “what if”, ...).



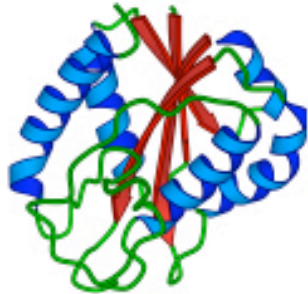
# Topics

1. Introduction to integrative (hybrid) structure determination
- 2. Comparative model building**
3. Predicting accuracy of atomic models
4. Iterative sequence-structure alignment and model building
5. Electron microscopy
6. Small angle x-ray scattering
7. Proteomics
8. Concluding Remarks

# Principles of protein structure

D. Baker & A. Sali. *Science* **294**, 93-97, 2001.

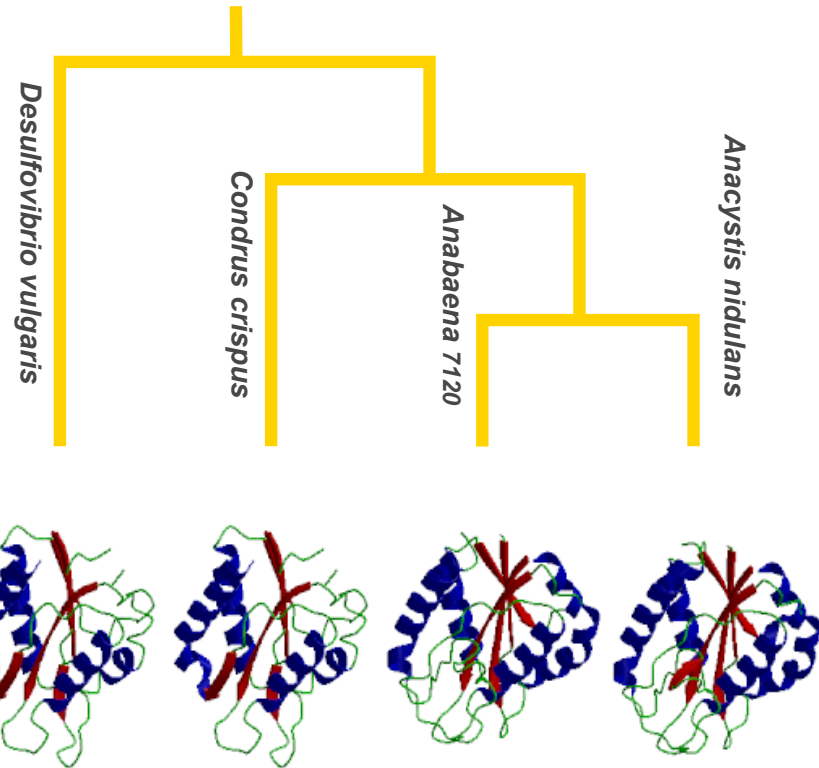
GFCHIKAYTRLIMVG...



## Folding

(physics)

***Ab initio* prediction**

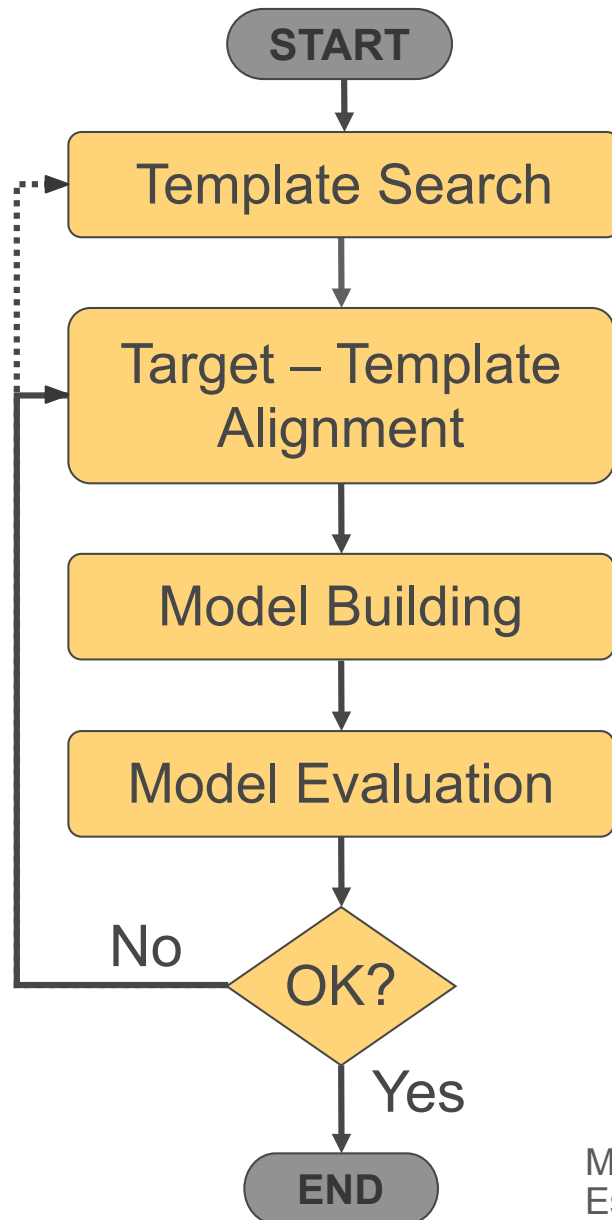


## Evolution

(“statistical” rules)

**Threading  
Comparative Modeling**

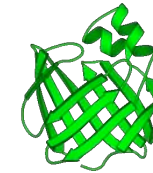
# Steps in Comparative Protein Structure Modeling



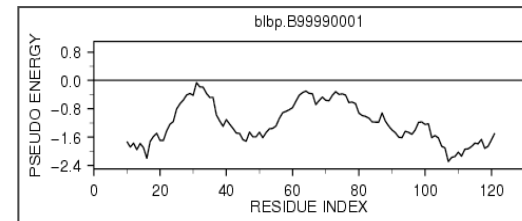
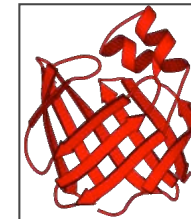
**TARGET**

ASILPKRLFGNCEQTSDEGLK  
IERTPLVPHISAQNVCLKIDD  
VPERLIPERASFQWMNDK

**TEMPLATE**



ASILPKRLFGNCEQTSDEGLK IERTPLVPHISAQNVCLKIDDVPERLIPE  
MSVIPKRLYGNCEQTSEEAI RIEDSPIV---TADLVCLKIDEIPERLVGE

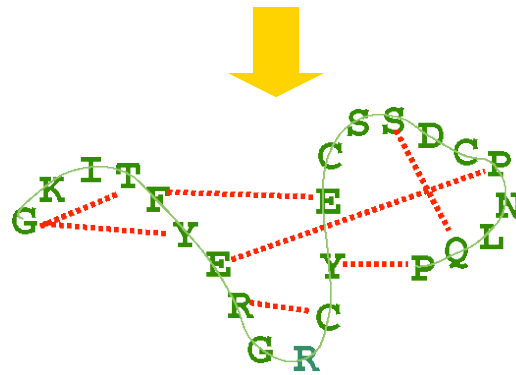


Marti-Renom *et al.* *Ann. Rev. Biophys. Biomolec. Struct.* **29**, 291, 2000.  
Eswar *et al.* Comparative Protein Structure Modeling Using MODELLER.  
Current Protocols in Protein Science, 2007.  
Schwede *et al.* Outcome of a workshop on applications of protein models in  
biomedical research. *Structure* 17, 151-159, 2009.

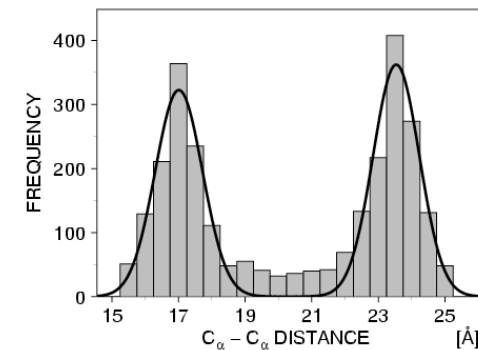
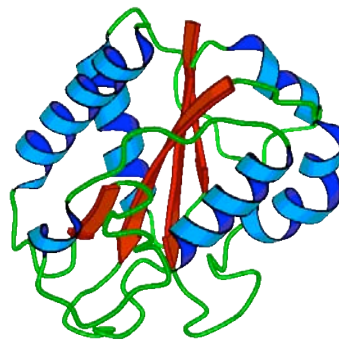
# Comparative modeling by satisfaction of spatial restraints MODELLER

3D GKITFYERGFQGHCYESDC-NLQP...  
SEQ GKITFYERG---RCYESDCPNLQP...

## 1. Extract spatial restraints



## 2. Satisfy spatial restraints



$$P(\mathbf{R} / \mathbf{I}) = \prod_i p_i(\mathbf{r}_i / \mathbf{I}_i)$$

A. Šali & T. Blundell. *J. Mol. Biol.* **234**, 779, 1993.  
J.P. Overington & A. Šali. *Prot. Sci.* **3**, 1582, 1994.  
A. Fiser, R. Do & A. Šali, *Prot. Sci.*, **9**, 1753, 2000.

<http://salilab.org/>

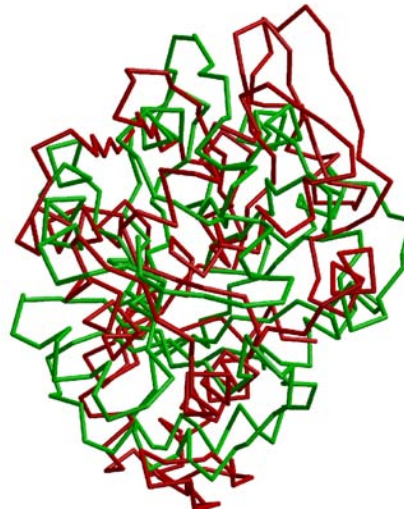
# Typical errors in comparative models

MODEL

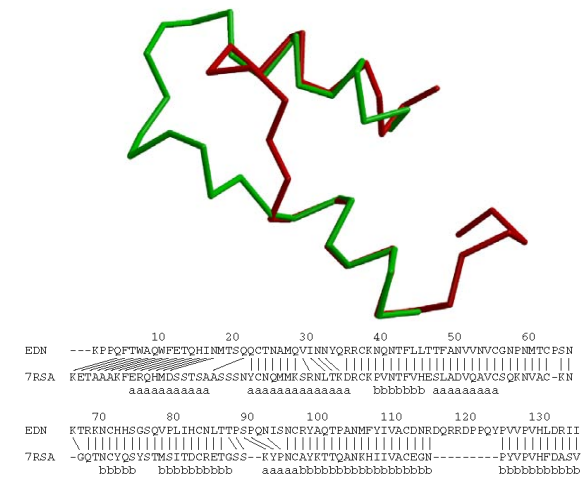
X-RAY

TEMPLATE

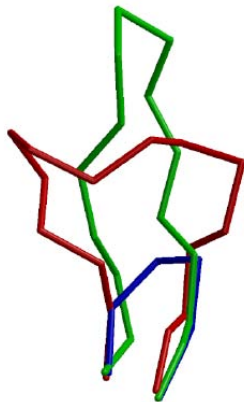
## Incorrect template



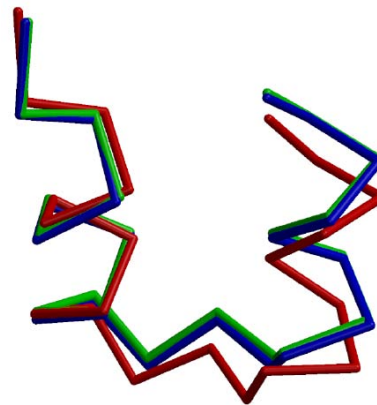
## Misalignment



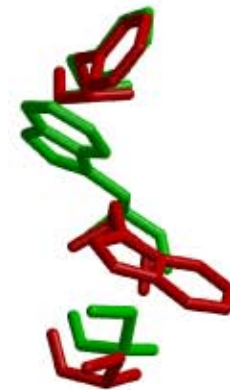
## Region without a template



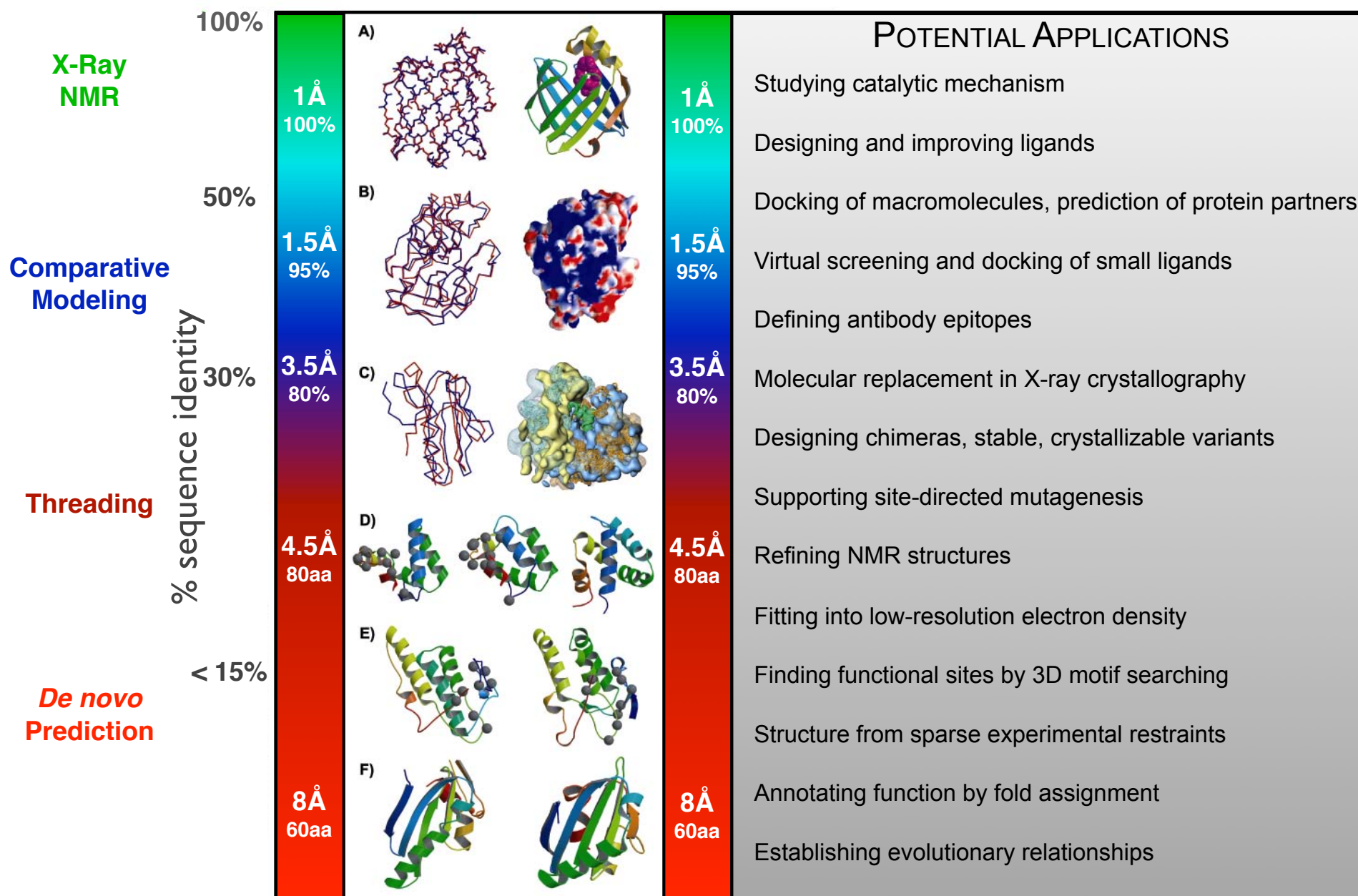
## Distortion/shifts in aligned regions



## Sidechain packing



# Model accuracy determines utility



# Topics

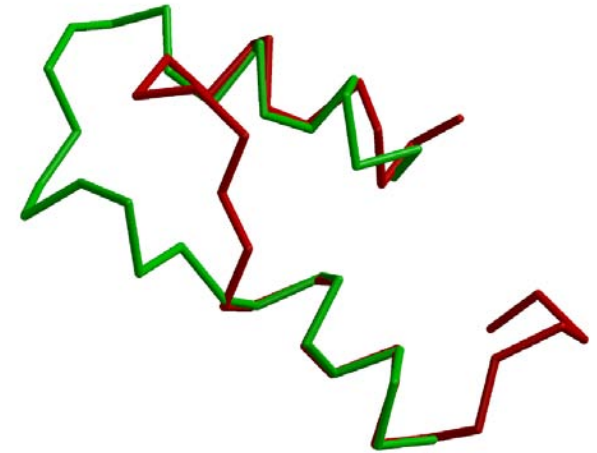
1. Introduction to integrative (hybrid) structure determination
2. Comparative model building
3. Predicting accuracy of atomic models
- 4. Iterative sequence-structure alignment and model building**
5. Electron microscopy
6. Small angle x-ray scattering
7. Proteomics
8. Concluding Remarks

# Minimizing errors in sequence-structure alignment

```

      10      20      30      40      50      60
EDN  ---KPPQFTWAQWFETQHINMTSQQCTNAMQVINNYQRRCKNQNTFLLTTFANVVNVCGNPNMTCPSN
7RSA KETAAAKFERQHMDSTSAAASSNYCNQMMKSRNLTKDRCKPVNTFVHESLADVQAVCSQKNVAC-KN
      aaaaaaaaaa aaaaaaaaaaaaaa bbbbbbb aaaaaaaaaa

      70      80      90      100     110     120     130
EDN  KTRKNCHHSGSQVPLIHCNLTTPSPQNISNCRYAQTFANMFYIVACDNRDQRRDPQYPVVPVHLDRII
7RSA -GQTNCYQSYSTMSITDCRETGSS--KYPNCAYKTTQANKHIIVACEGN-----PYVPVHFDA SV
      bbbbb  bbbbbbbbbbb aaaaabbbbbbbbbbbbbbbbbbb bbbbbbbbbbb
  
```

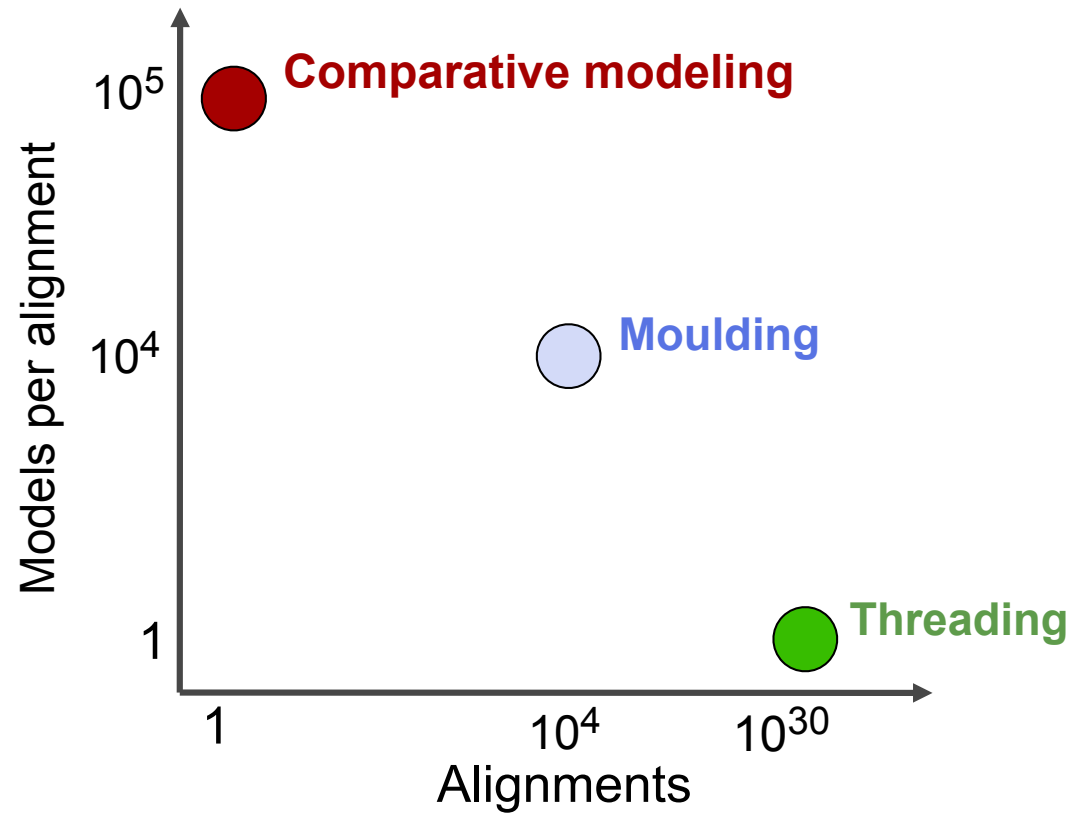
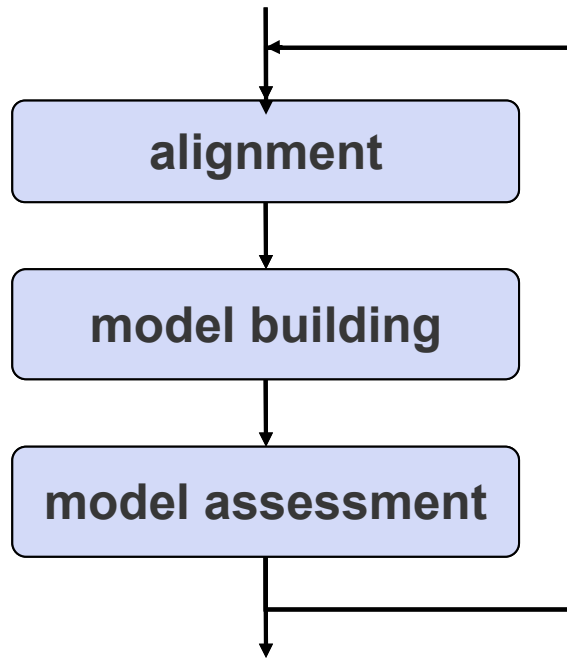


- Complex gap penalty functions.
- Multiple sequence profiles.
- Hidden Markov Models.
- Threading.



# Moulding: iterative alignment, model building, model assessment

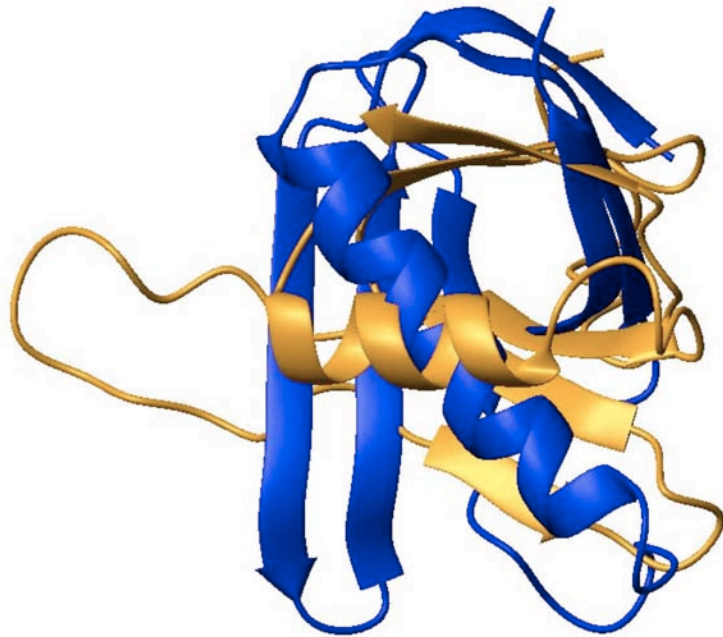
B. John, A. Sali. *Nucl. Acids Res.* **31**, 1982-1992, 2003.  
D. Eramian, B. Webb.



# Application to a difficult modeling case

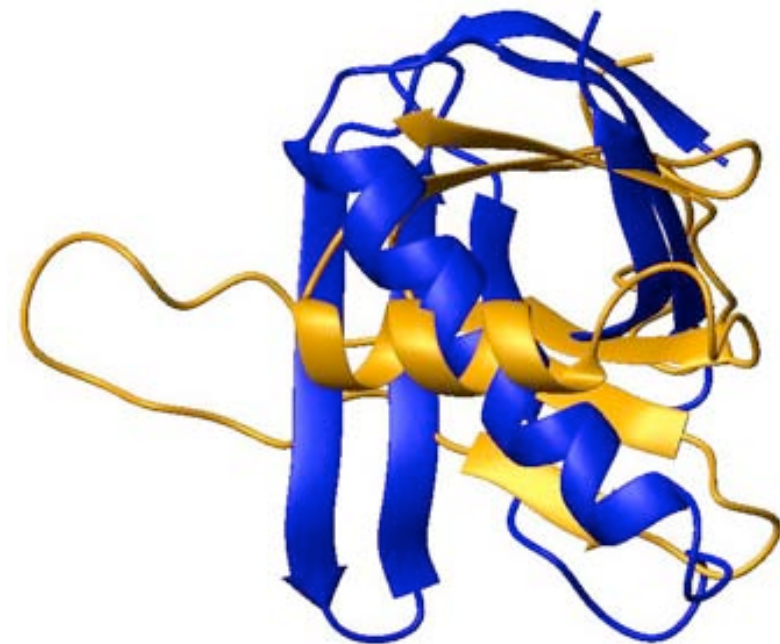
1BOV-1LTS (4.4% sequence identity)

initial



C $\alpha$  RMSD 10.1 Å

final



C $\alpha$  RMSD 3.6 Å

**1lts structure**

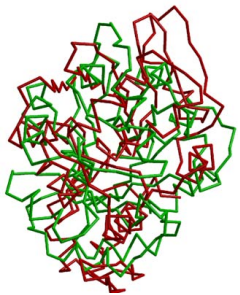
**1lts model**

# Topics

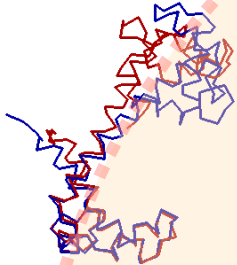
1. Introduction to integrative (hybrid) structure determination
2. Comparative model building
3. Predicting accuracy of atomic models
4. Iterative sequence-structure alignment and model building
5. **Electron microscopy**
6. Small angle x-ray scattering
7. Proteomics
8. Concluding Remarks

# Comparative modeling and fitting into EM density

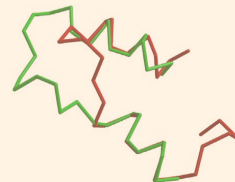
Incorrect templates



Rigid-body movements



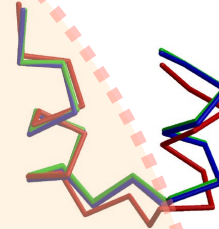
Misalignments



Regions without a template



Distortion and shifts of aligned regions



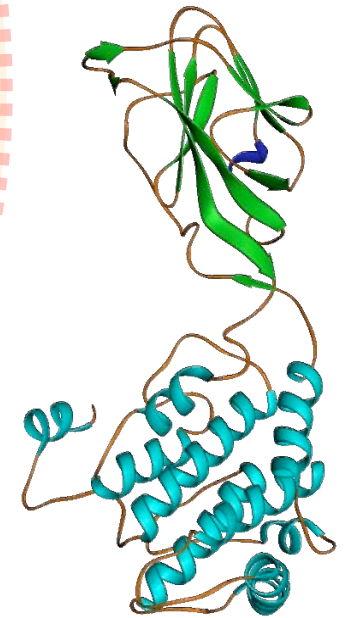
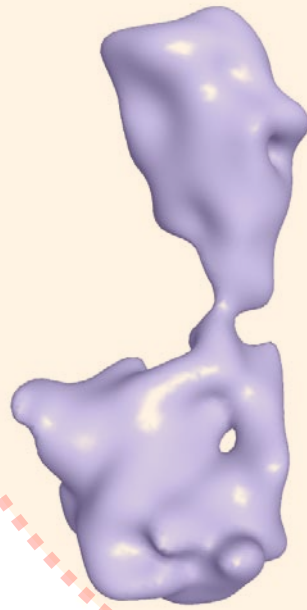
Sidechain packing



20Å

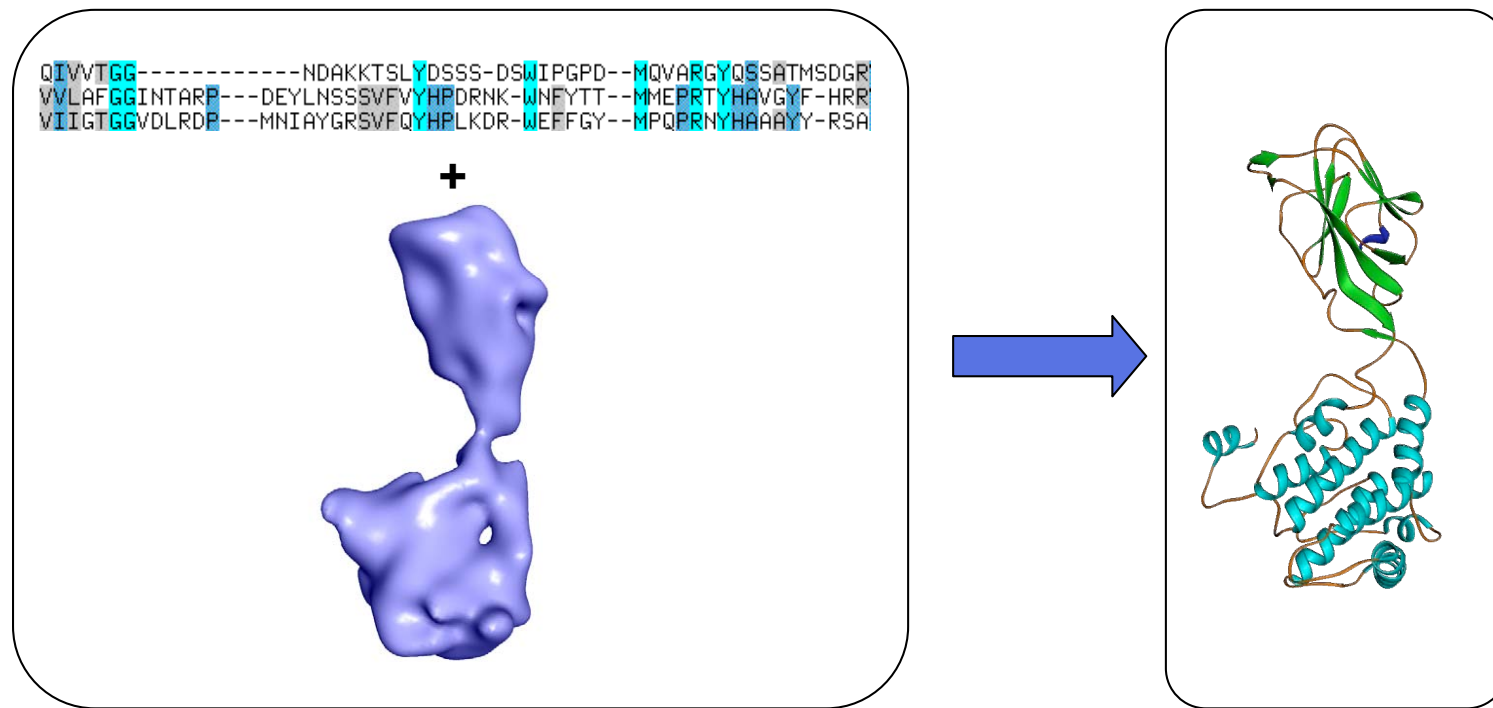
10Å

2Å



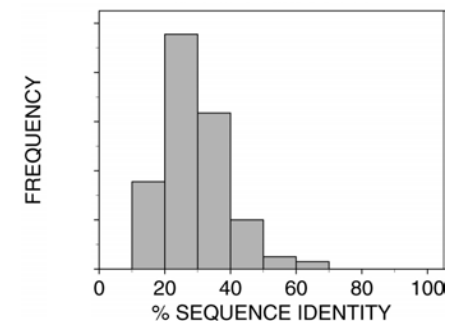
# Comparative modeling and fitting into EM density

Improve comparative modeling by fitting models into the target EM density map;  
Improve fitting into an EM density map by simultaneous model building.



## Motivation:

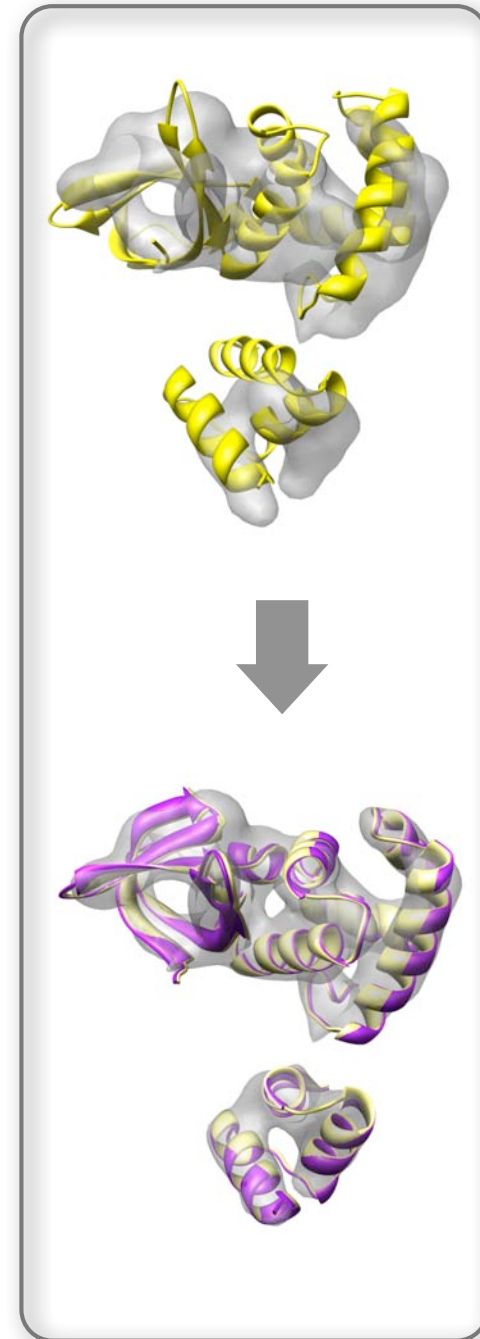
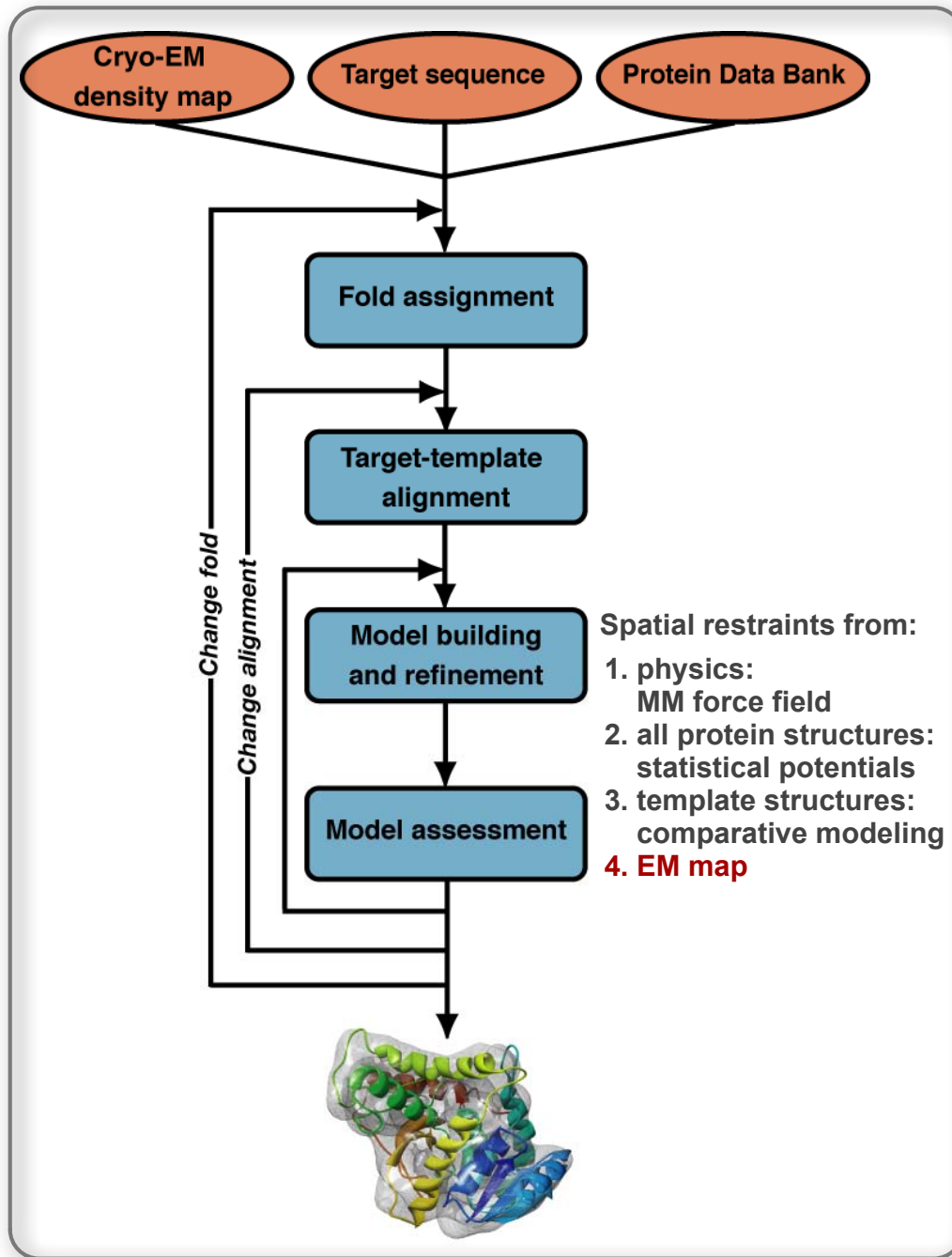
- Number of known structures in PDB: ~65,000
  - Number of known sequences modeled by CM: ~2,200,000
- Pieper *et al*, Nucl. Acids Res., 2009.



# Protein structure modeling in an EM map

Topf, Baker, John, Chiu, Sali. *J. Struct. Biol.*, 2004.  
Topf & Sali, *Curr. Opin. Str. Biol.*, 2005.

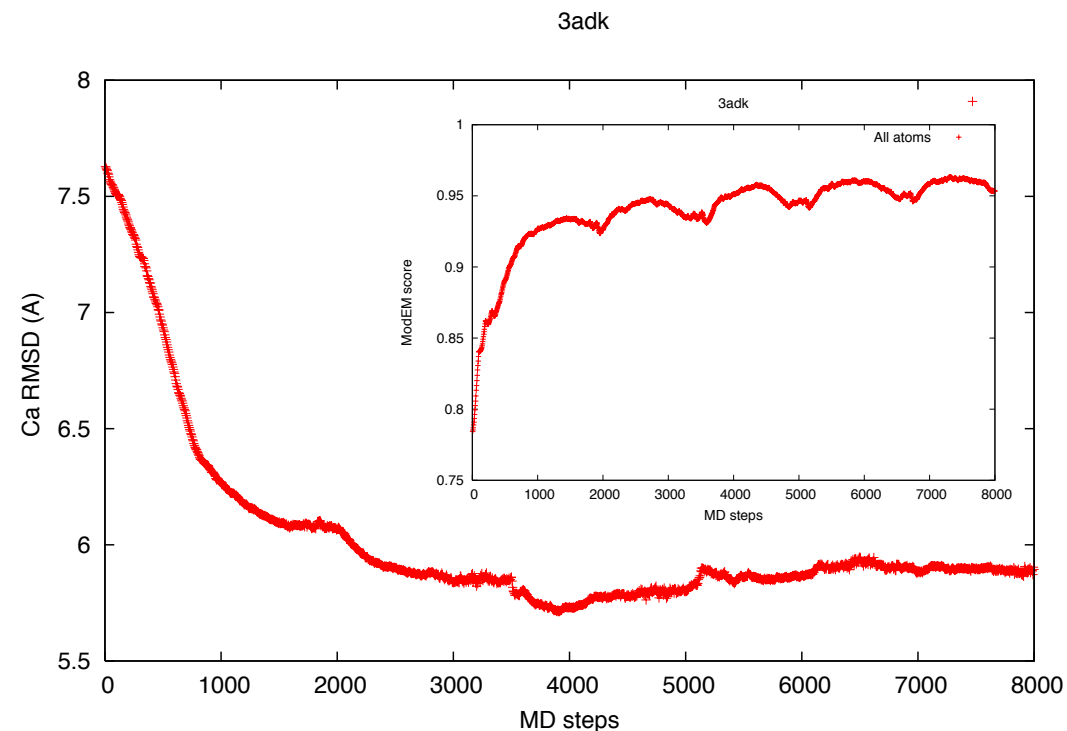
Topf, Baker, Marti-Renom, Chiu & Sali. *J. Mol. Biol.*, 2006.  
Topf, Lasker, Webb, Wolfson, Chiu & Sali. *Structure*, 2008.



# Sample refinement of 1adk

Topf, Lasker, Webb, Wolfson, Chiu & Sali. *Structure*, 2008.

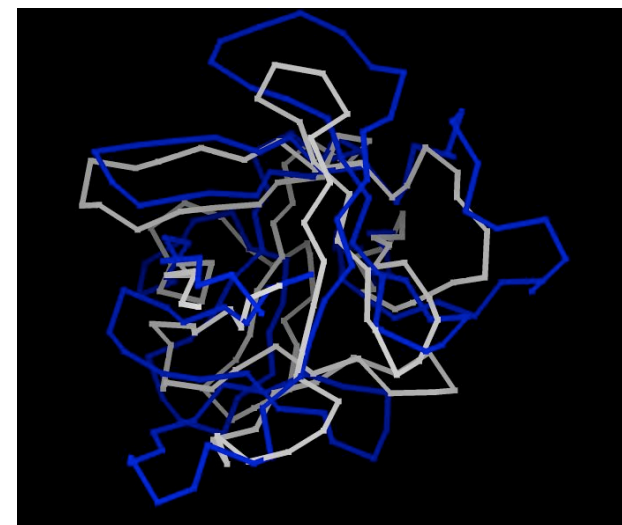
- EM map (10 Å) from native structure;
- secondary structure segments as rigid bodies, loops flexible;
- scoring function consisting only of model-map correlation coefficient, soft-sphere atom overlap, stereochemistry;
- optimization by a combination of “molecular dynamics” with simulated annealing and conjugate gradients minimization.



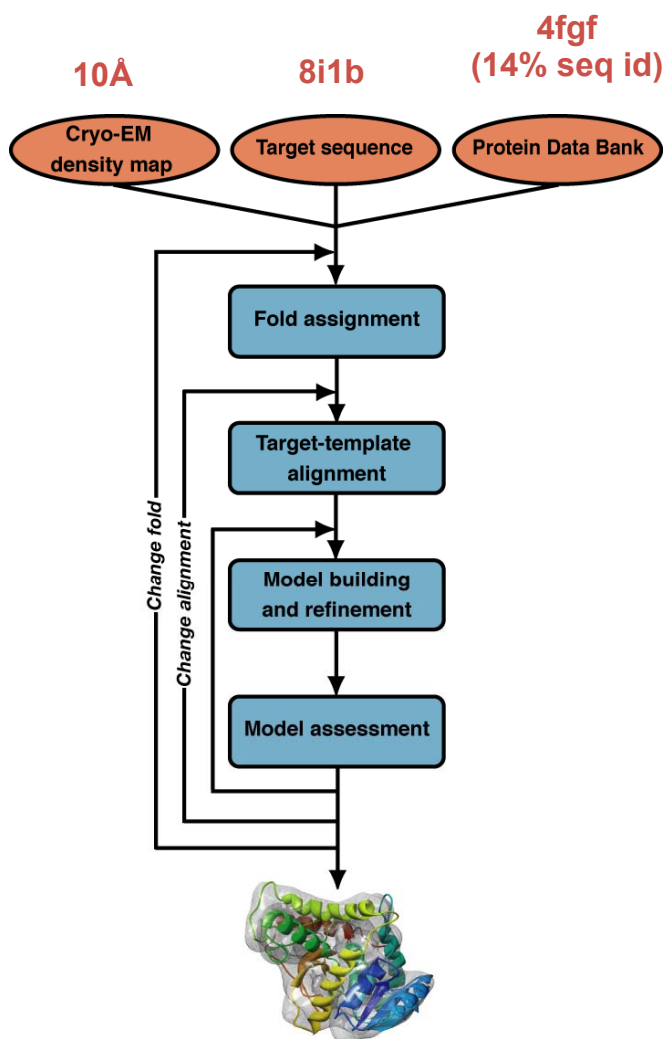
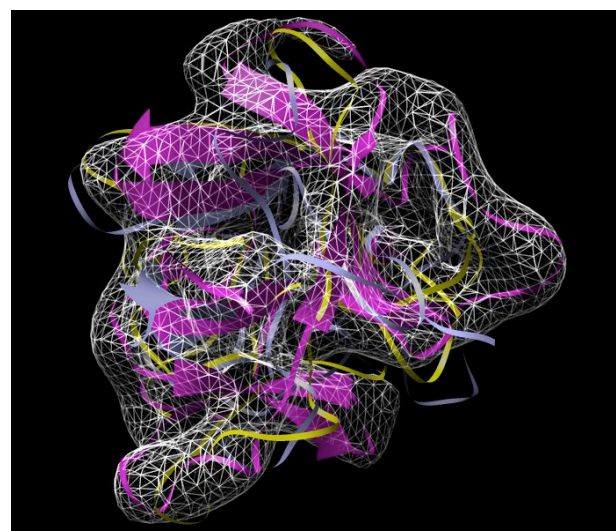
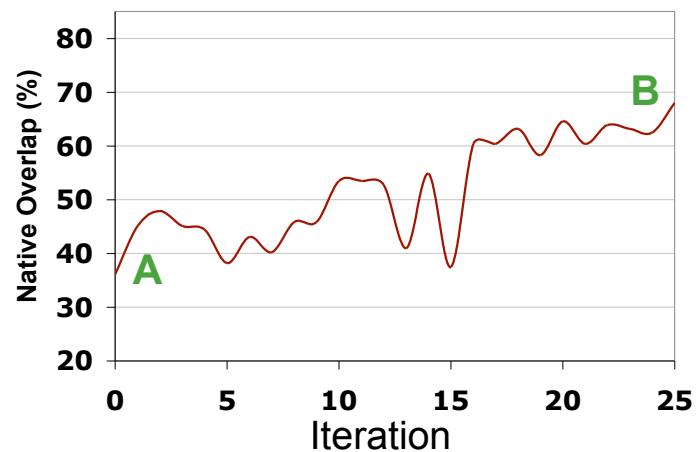
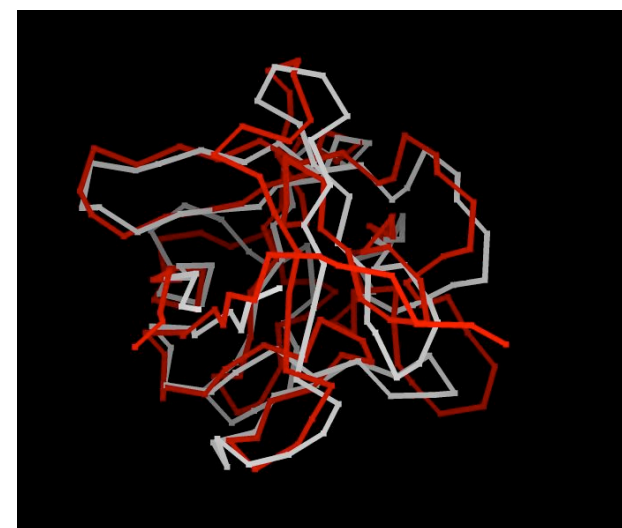
# Moulding into EM maps

Topf, Baker, Marti-Renom, Chiu & Sali. *J. Mol. Biol.*, 2006.

**A.** 37% of  $C_{\alpha}$  within 5Å



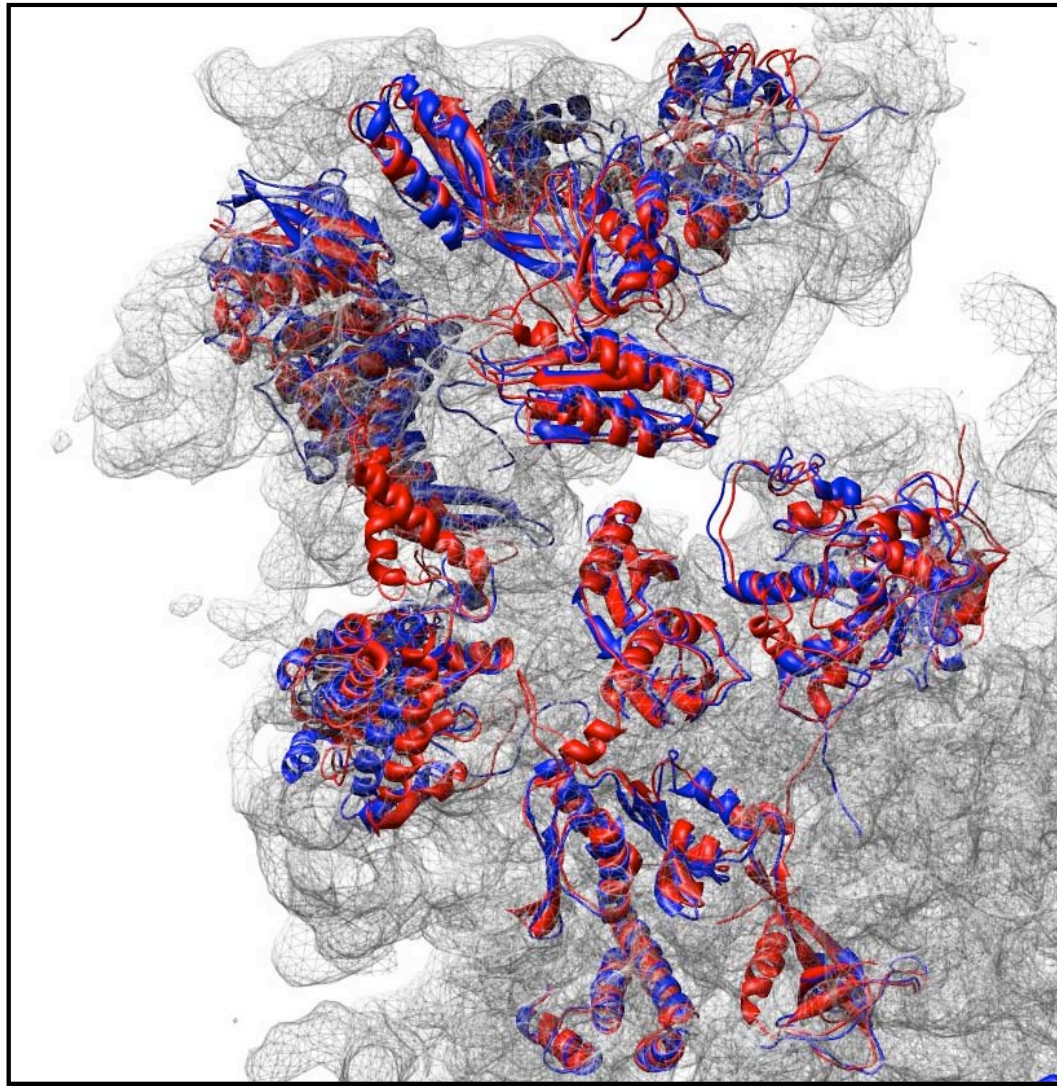
**B.** 69% of  $C_{\alpha}$  within 5Å





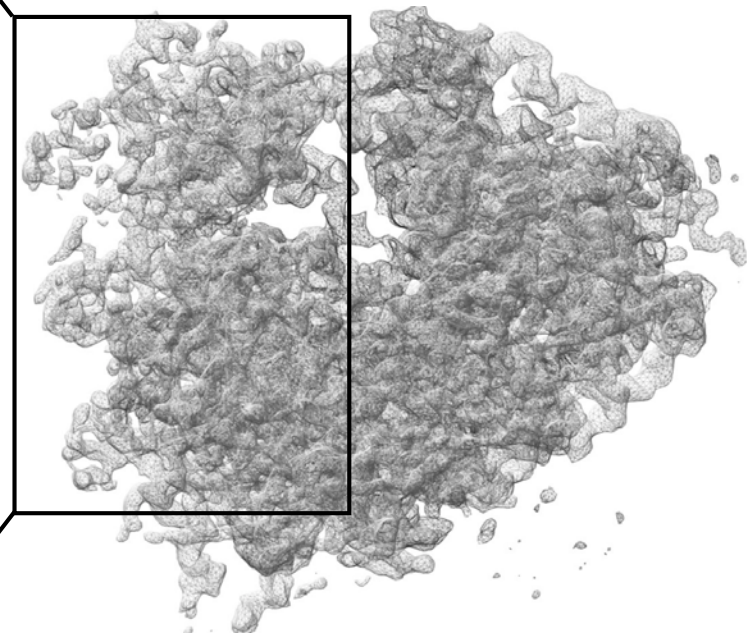
# Dog Ribosome at 8.7 Å

Chandramouli, Topf, Menetret, Eswar, Gutell, Sali, Akey. *Structure*, 2008.



Added value:

- Unique insertions/deletions in proteins
- Unique protein-RNA contacts
- Mammalian-only ribosomal proteins



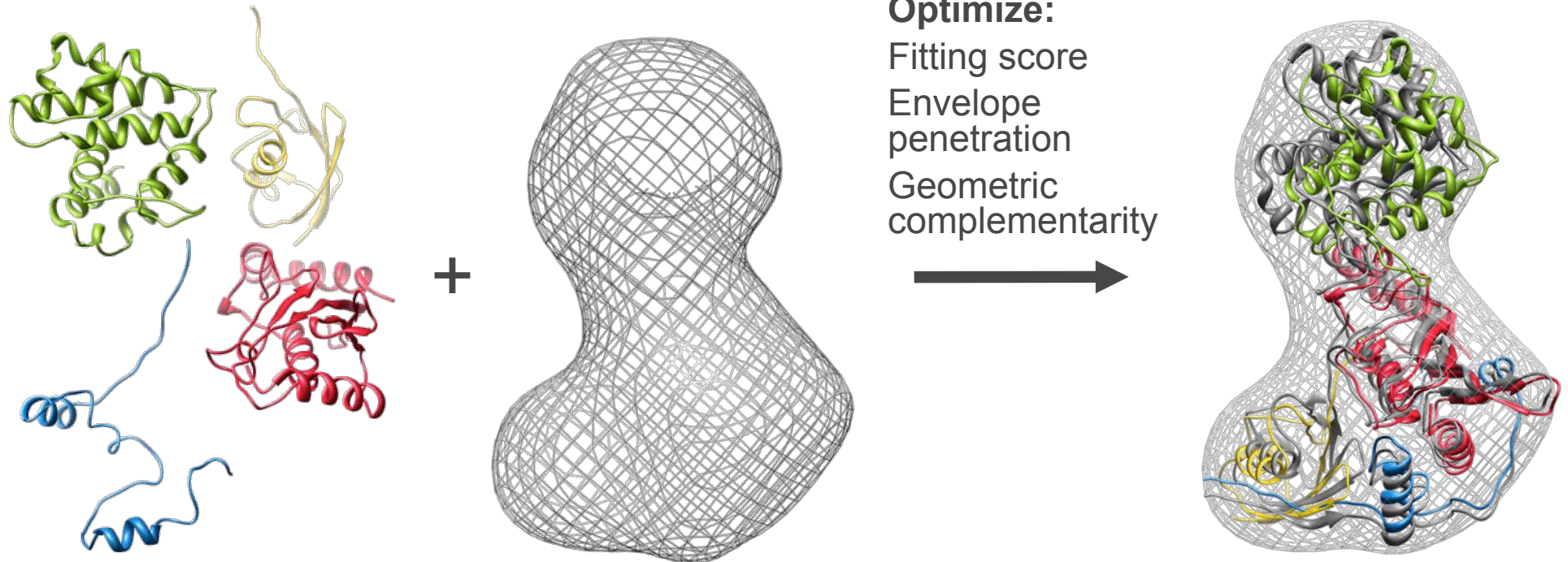
*Thermus Thermophilus* 30S ribosomal subunit (proteins - red; RNA - yellow)

Homology models of the mammalian ribosomal proteins (blue)



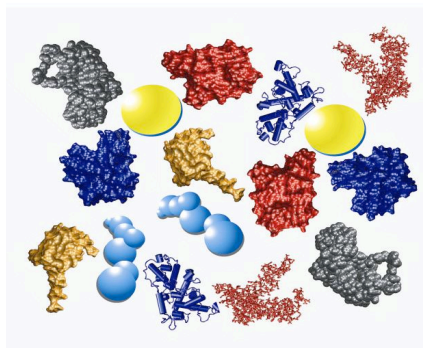
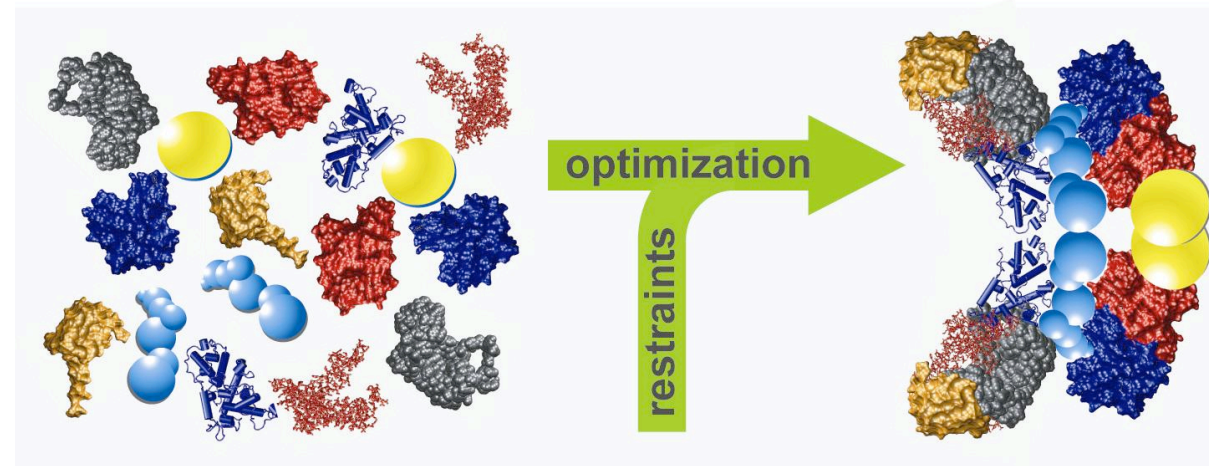
# MultiFit / DOMINO: Fitting of multiple components into EM density maps

Lasker, Topf, Sali, and Wolfson, *J. Mol. Biol.* 388, 180-194, 2009.

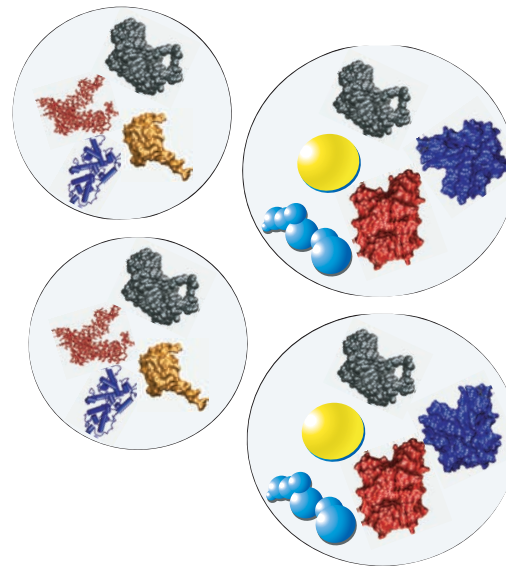


# Approach

1. Representation of a system.
2. Scoring function (spatial restraints)
3. **(Combinatorial) Optimization**

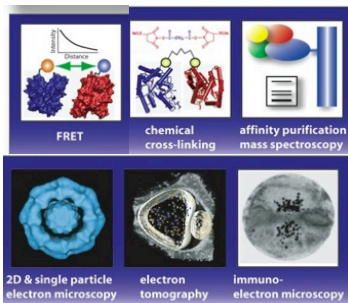
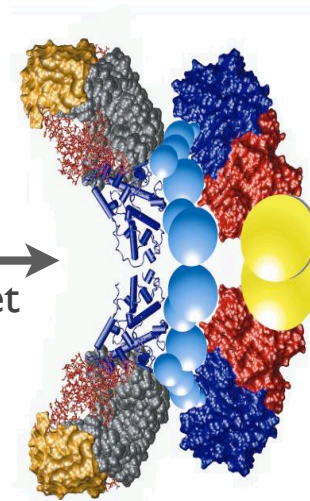


Determine subsets for sampling from input data



Generate a discrete sampling space for each subset

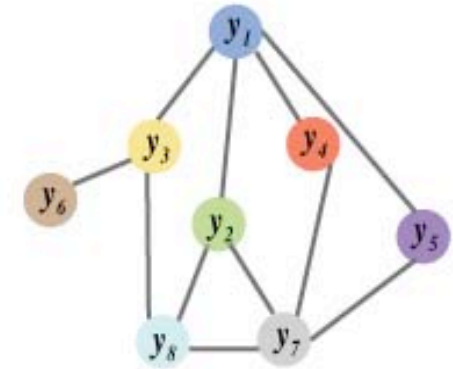
Gather subset solutions



# Divide-and-Conquer

1. **Represent** the scoring function as a graph.

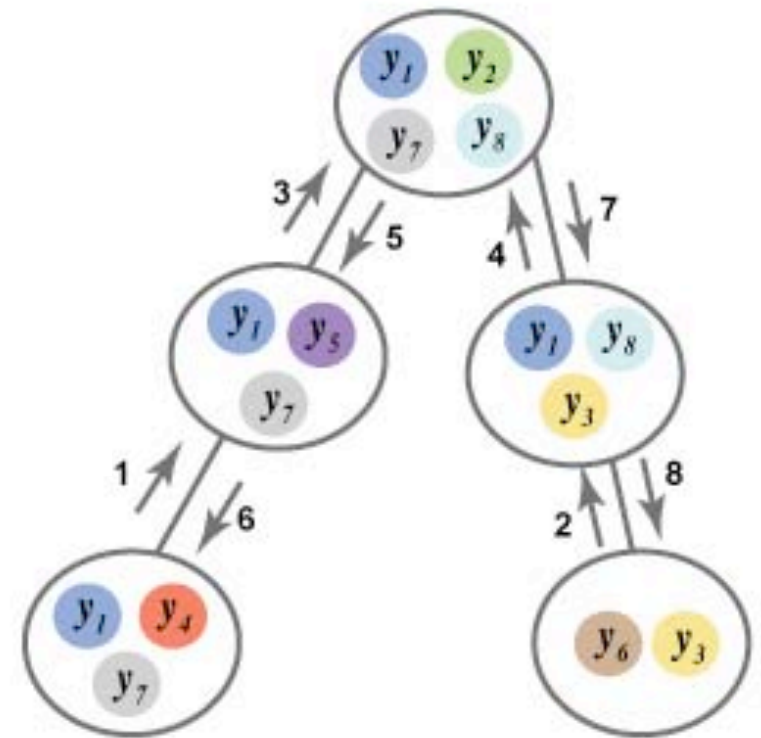
$$F(y_1, \dots, y_8) = \alpha_2(y_2) + \alpha_6(y_6) + \alpha_7(y_7) \\ + \beta_{1,2}(y_1, y_2) + \beta_{1,3}(y_1, y_3) + \beta_{1,4}(y_1, y_4) + \beta_{1,5}(y_1, y_5) \\ + \beta_{2,7}(y_2, y_7) + \beta_{2,8}(y_2, y_8) + \beta_{3,6}(y_3, y_6) + \beta_{3,8}(y_3, y_8) \\ + \beta_{4,7}(y_4, y_7) + \beta_{5,7}(y_5, y_7) + \beta_{7,8}(y_7, y_8)$$



2. **Decompose** the set of variables into relatively decoupled subsets (a junction tree algorithm).

3. **Optimize** each subset independently by a traditional optimizer, to get the optimal and a number of suboptimal solutions.

4. **Gather** subset solutions into the best possible global solutions (message passing algorithms; eg, belief-propagation).

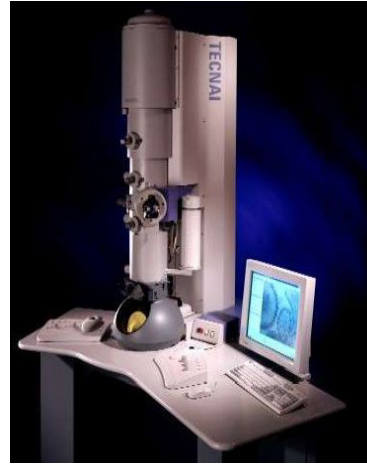


# 3D-EM process

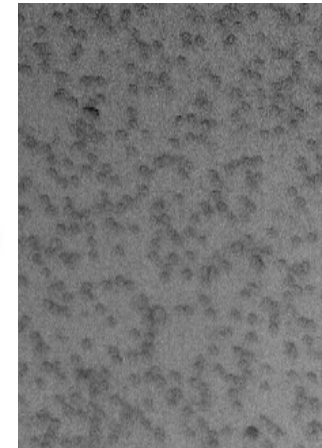
Specimen



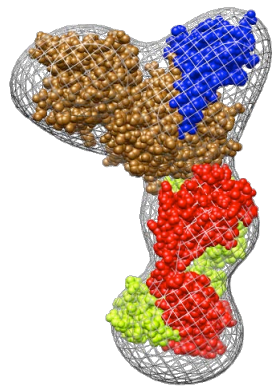
Microscope



Digitized  
micrograph



Particle selection,  
alignment,  
classification,  
averaging



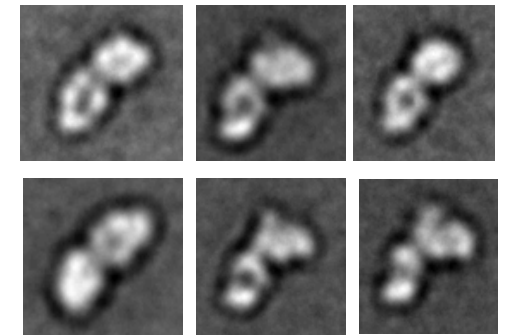
Model

Fitting (& modeling)



3D map

3D reconstruction



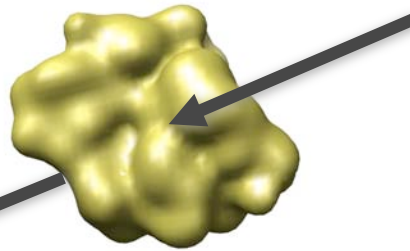
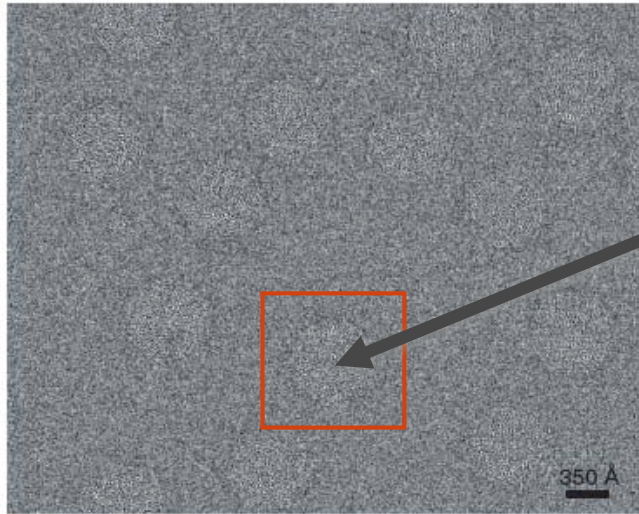
2D classes  
(averages)

Additional restraints

**J. Velazquez-Muriel**

# Scoring: Comparison of EM image and model

J. Velazquez-Muriel, K. Lasker, D. Schneidman



**Difference between an image and model projection (P. Penczek):**

$$term = \frac{1}{2} \sum_{i=1}^N \|P(\alpha_i, y) - \mathbf{d}_i\|^2 = \frac{1}{2} \sum_{i=1}^N \|P(\alpha_i, D(m)) - \mathbf{d}_i\|^2$$

- Each particle (class) image  $\mathbf{d}_i$  is compared with the corresponding projection  $P(\alpha_i, y)$  of the down-sampled model  $m$ ,  $D(m)$ .
- $\alpha_i$  is the vector with the 5 required parameters to fully define the projection of a model (*i.e.*, the three orientation angles and the two translation distances).
- Assuming conformational and configurational homogeneity, though generalization may be possible.
- Can be easily extended to tilt series of images to improve data-to-parameter ratio.

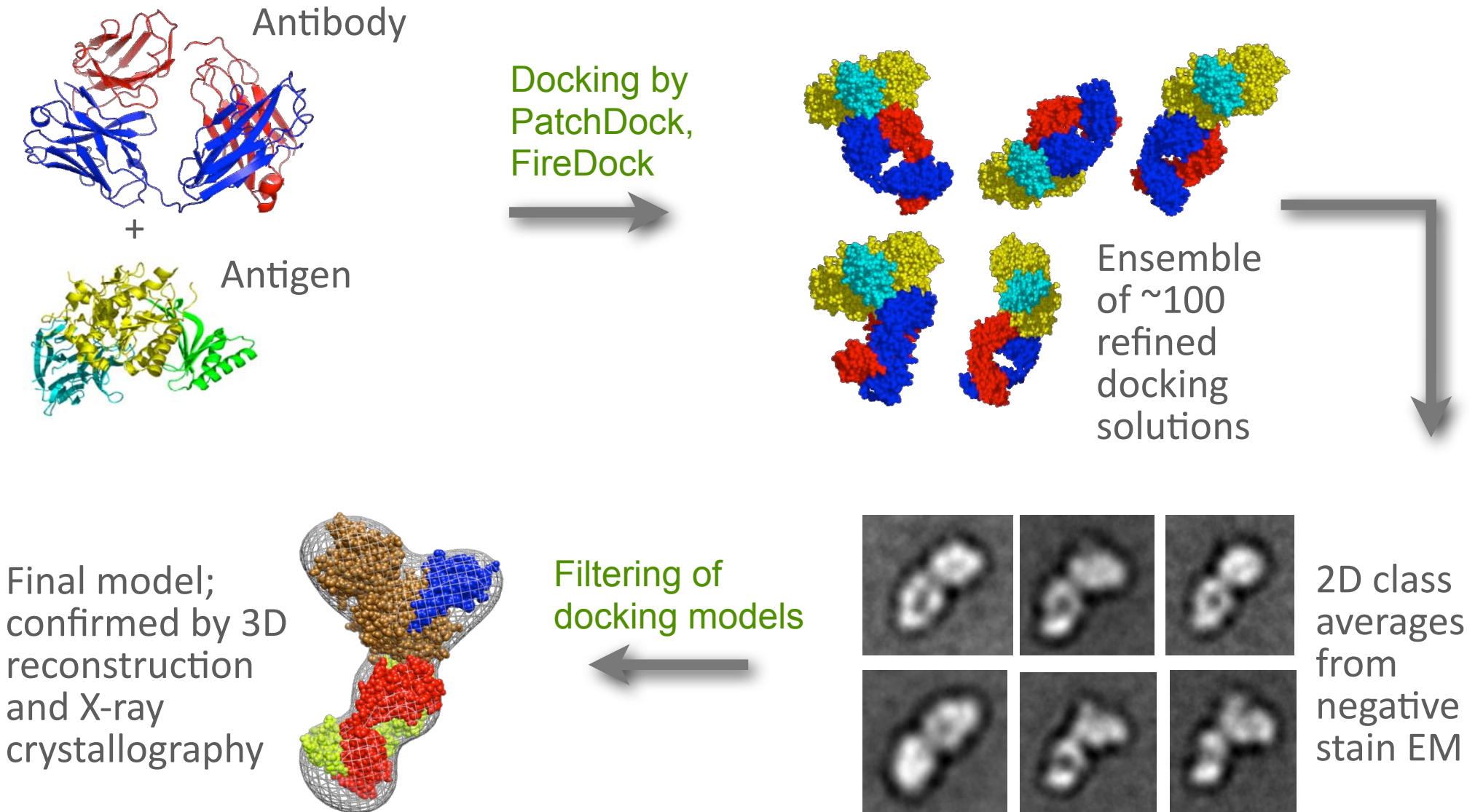
# Filtering: Select models that fit all 2D class averages

For each model:

- Generate ~400 model projections, by enumerating rotation and tilt angles.
- For each projection:
  - For each class average:
    - Find optimal rotation and translation for overlapping model projection with class average, using cross-correlation / convolution algorithm (Penczek and Frank; *Ultramicroscopy*, 1992)
    - Refine the approximate solution for the five degrees of freedom by a simplex (Nelder-Mead) algorithm.

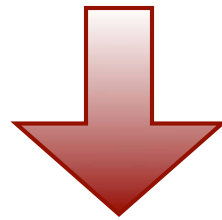
# Application to an antigen - antibody complex

D. Schneidman, J. Velazquez; with A. Rajpal, P. Strop, A. Rossi (Pfizer); A. Avila-Sakar, M. Liao, H. Kim, Y. Cheng (UCSF); K. Krukenberg, D. Agard (UCSF); S. Sobhanifar, V. Dötsch (U. Frankfurt)





**Filtering:** Select models that fit  
all 2D class averages

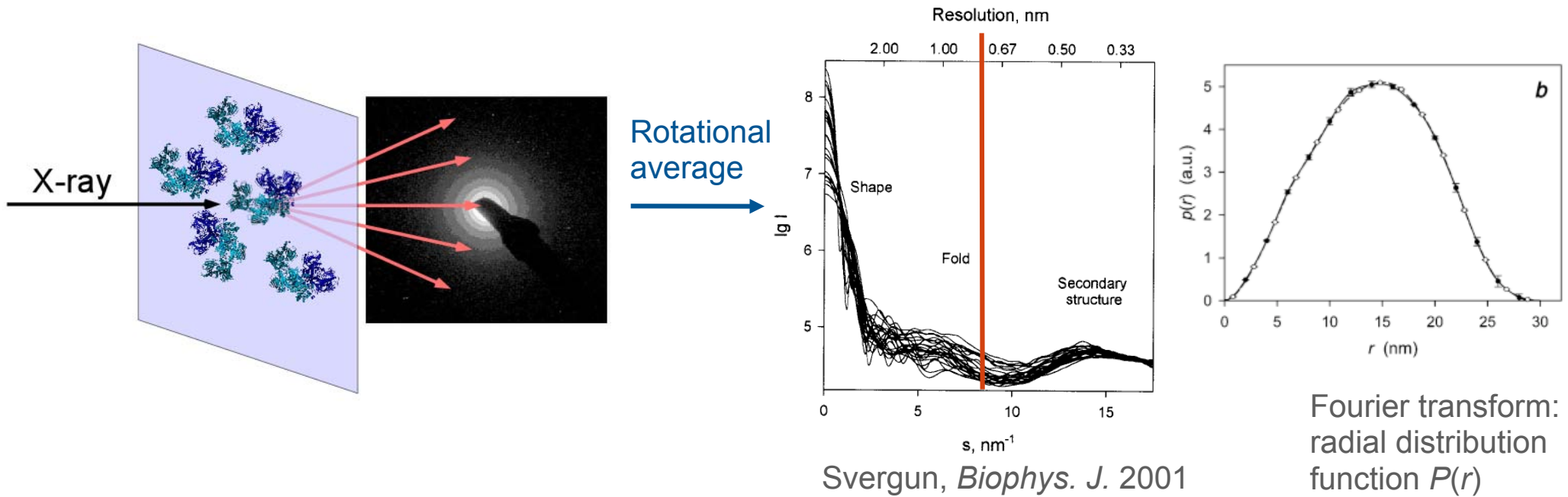


**Optimization:** Generate models that fit  
all 2D class averages

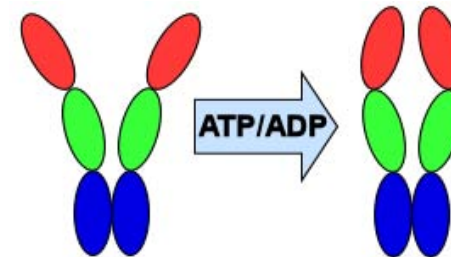
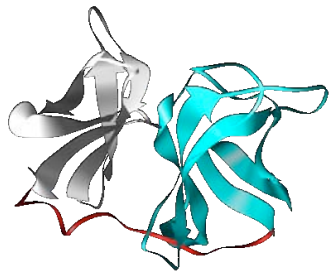
# Topics

1. Introduction to integrative (hybrid) structure determination
2. Comparative model building
3. Predicting accuracy of atomic models
4. Iterative sequence-structure alignment and model building
5. Electron microscopy
- 6. Small angle x-ray scattering**
7. Proteomics
8. Concluding Remarks

# Small angle X-ray scattering (SAXS)

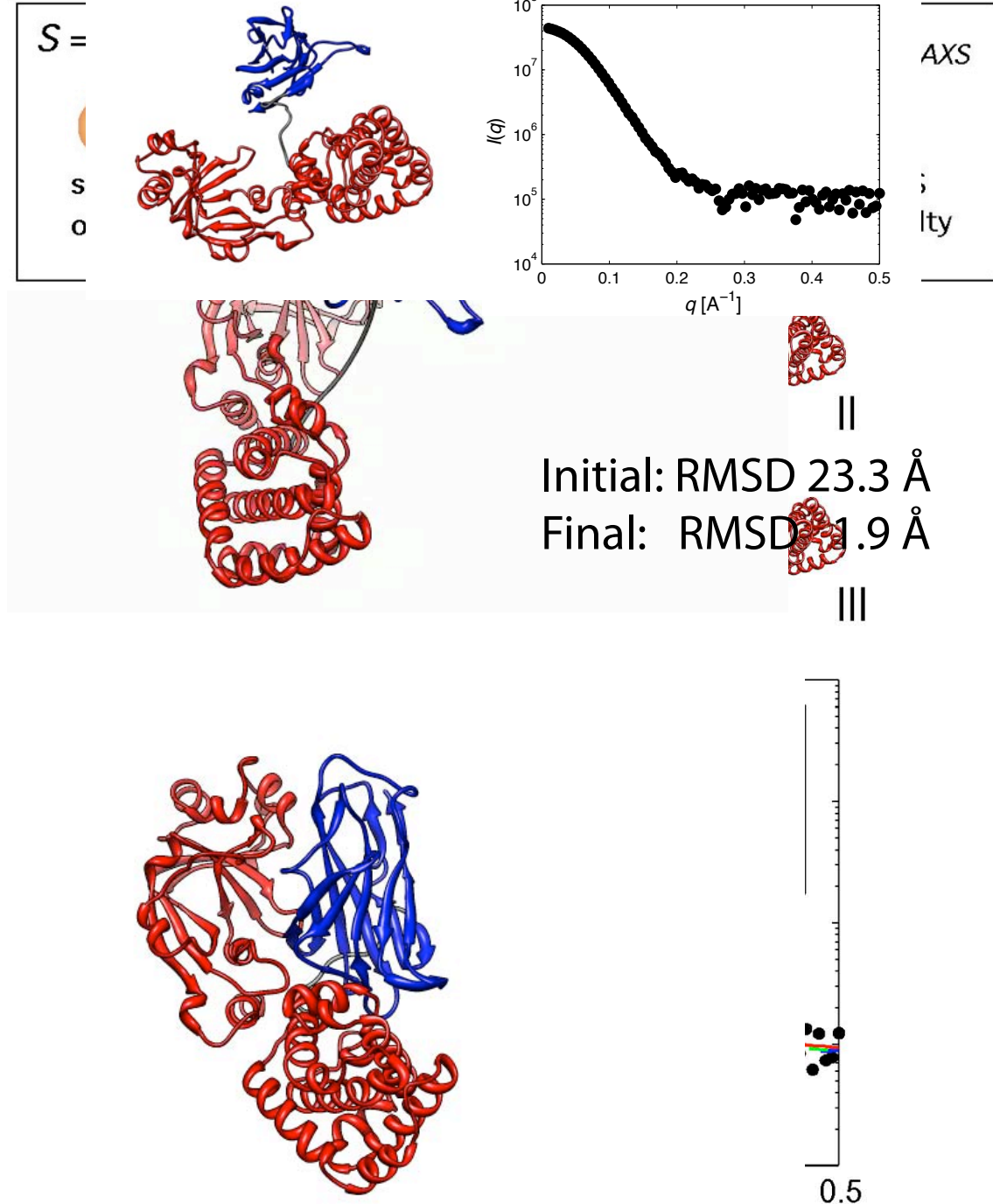
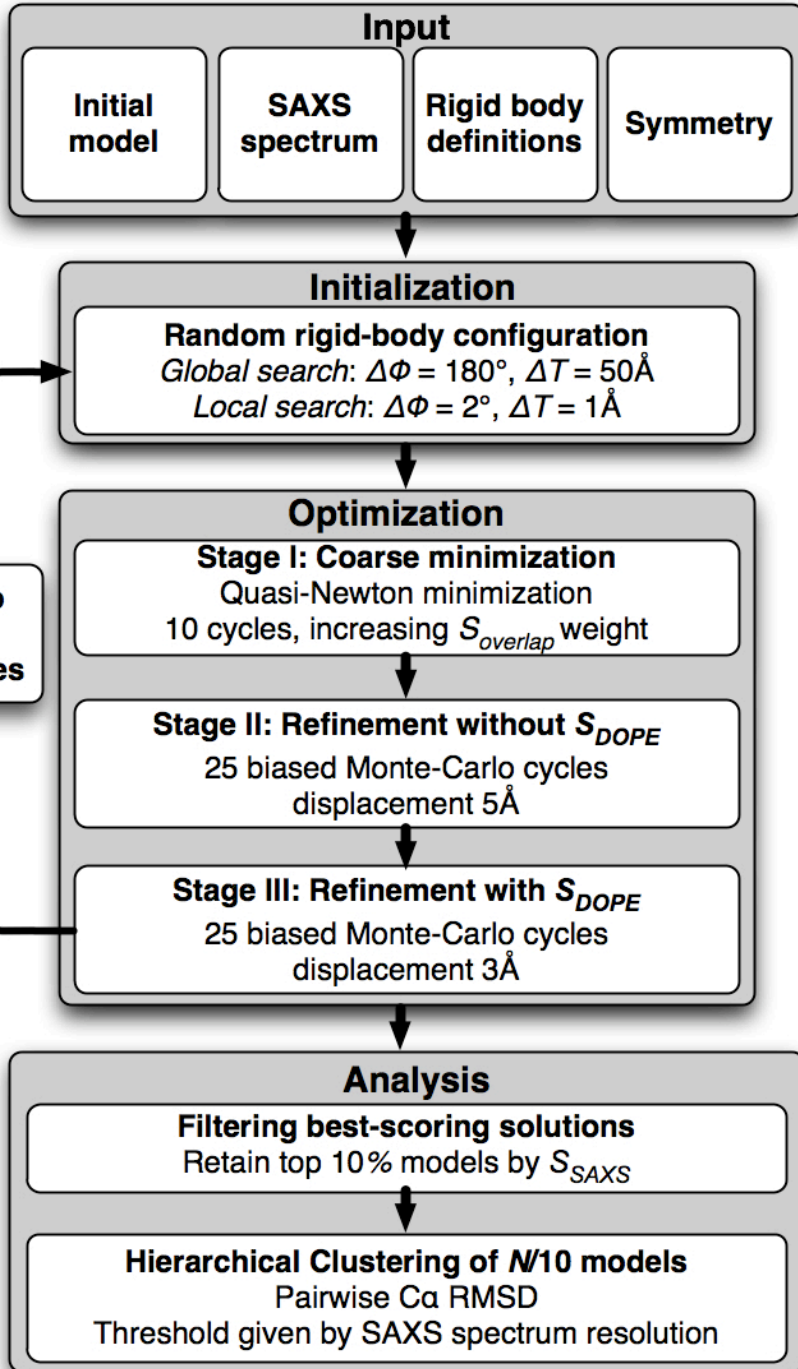


- Limited information content of a SAXS spectrum
- Integration with additional data
- Quaternary structure
- Changes in quaternary structure

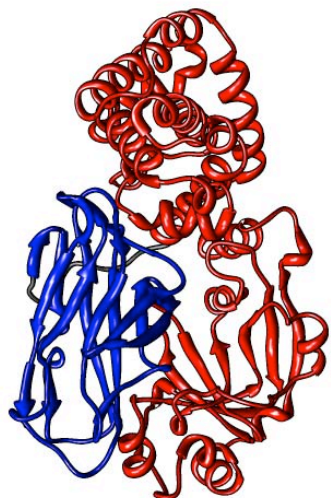


# Protocol

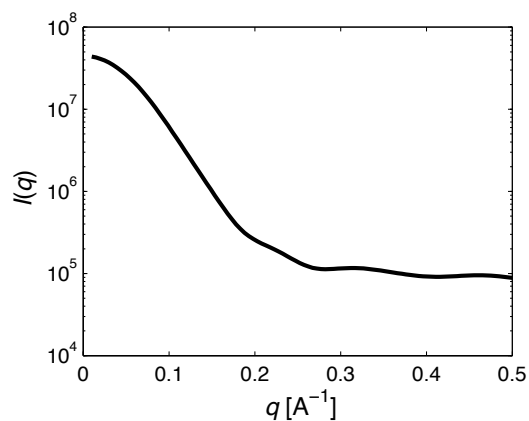
F. Förster, B. Webb, K.A. Krukenberg, H. Tsuruta, D.A. Agard, A. Sali. *J. Mol. Biol.* 382, 1089-1106, 2008.



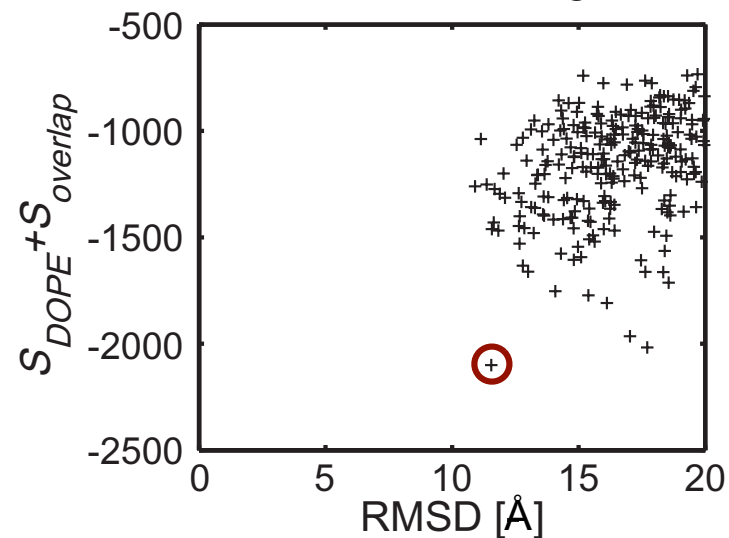
# Benefit of integration of SAXS with modeling



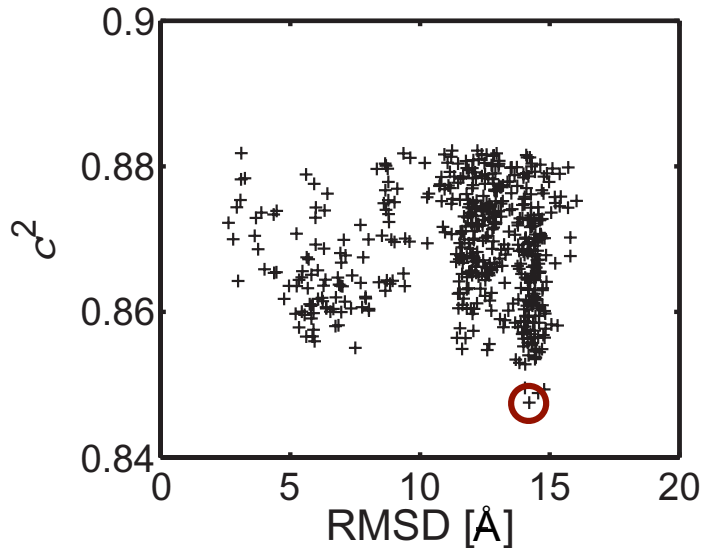
Diphtheria toxin (1mdt)



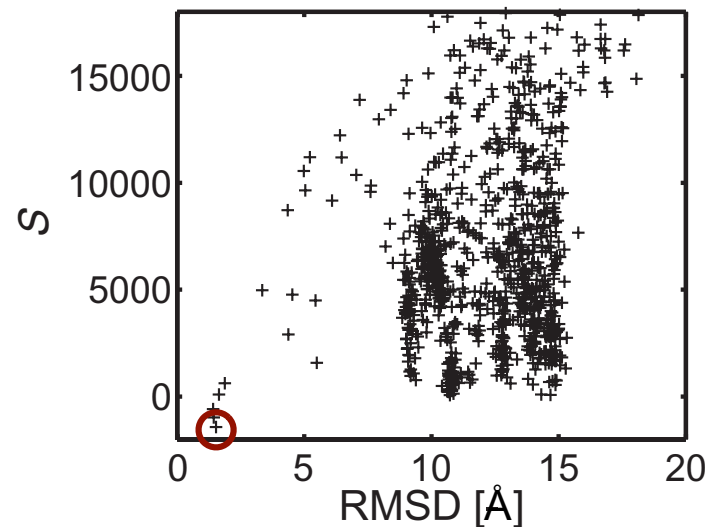
Molecular docking **fails**



SAXS on its own **fails**



Combined scoring **succeeds**

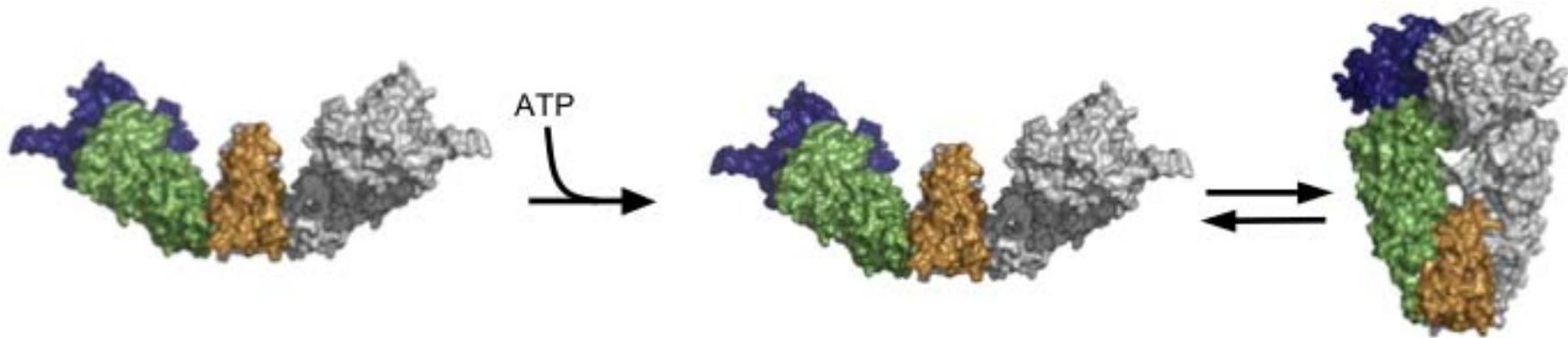


# Summary of SAXS method

- Atomic models can be determined that are consistent with given SAXS data and additional restraints.
- Integration of information increases accuracy.
- Configurations can be sampled “exhaustively” for up to 4 domains.
- Configuration accuracy depends on rigid body accuracy ( $\sim 3 \text{ \AA}$  C $\alpha$  RMSD necessary).
- Integration of further information is possible.

# SAXS maps Hsp90 states

K.A. Krukenberg, F. Förster, L. Rice, A. Sali, D.A. Agard, *Structure* **16**, 755-765, 2008.



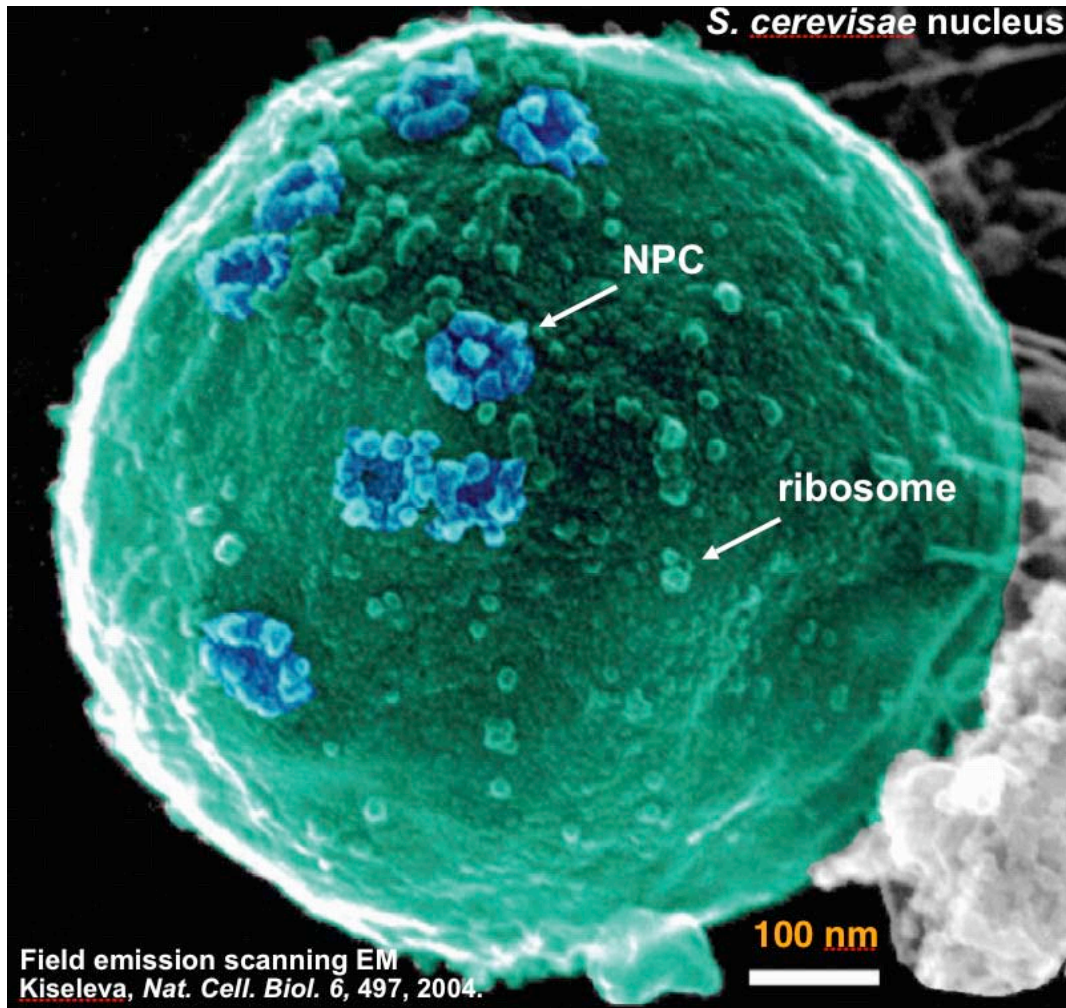
- Crystallographic structures of opened and closed states are probably inaccurate representations of solution states.
- The apo structure of *E. coli* Hsp90 is wide open.
- *E. coli* ATP-Hsp90 is in equilibrium between the wide-opened and closed states.

# Topics

1. Introduction to integrative (hybrid) structure determination
2. Comparative model building
3. Predicting accuracy of atomic models
4. Iterative sequence-structure alignment and model building
5. Electron microscopy
6. Small angle x-ray scattering
- 7. Proteomics**
8. Concluding Remarks



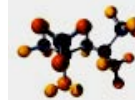
# Nuclear Pore Complex (NPC)



Consists of broadly conserved **nucleoporins** (nups).  
50 MDa complex: ~480 proteins of 30 different types.  
**Mediates** all known nuclear **transport**, via cognate transport factors (karyoferins or kaps)

1. Structure
2. Evolution
3. Mechanism of transport
4. Mechanism of assembly
5. Interactions with other systems

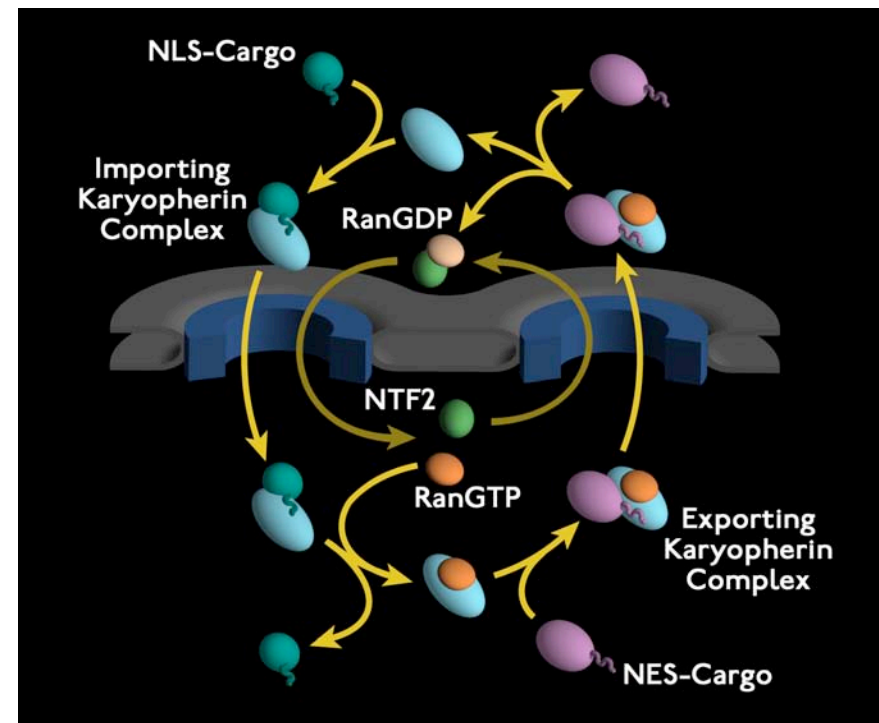
A large collaborative effort with Mike Rout and Brian Chait at Rockefeller University, also involving many other collaborators (Acknowledgments).



NCDIR

National Center for Dynamic  
Interactome Research

NIH TCNP

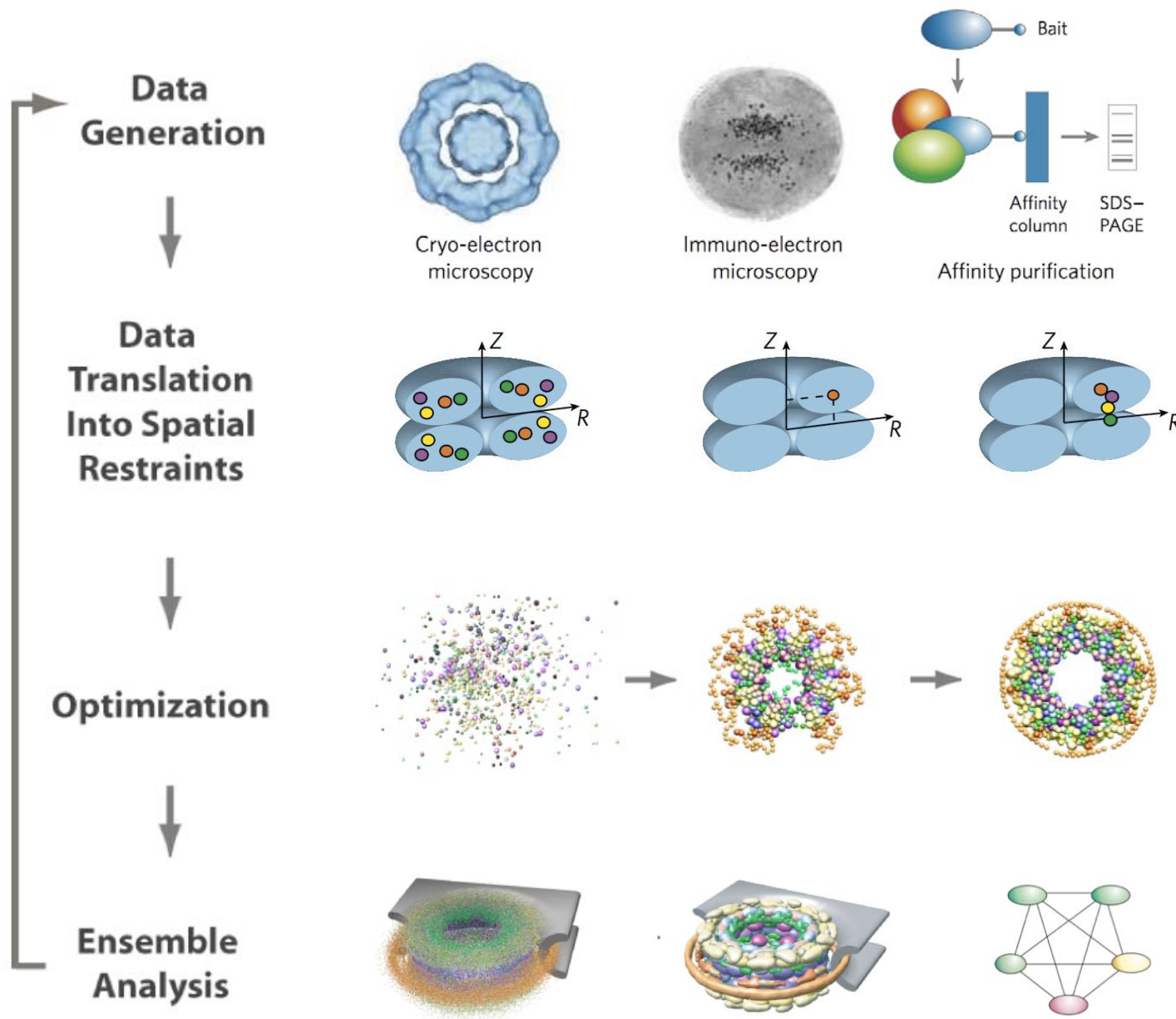


# Integrative determination of macromolecular structures

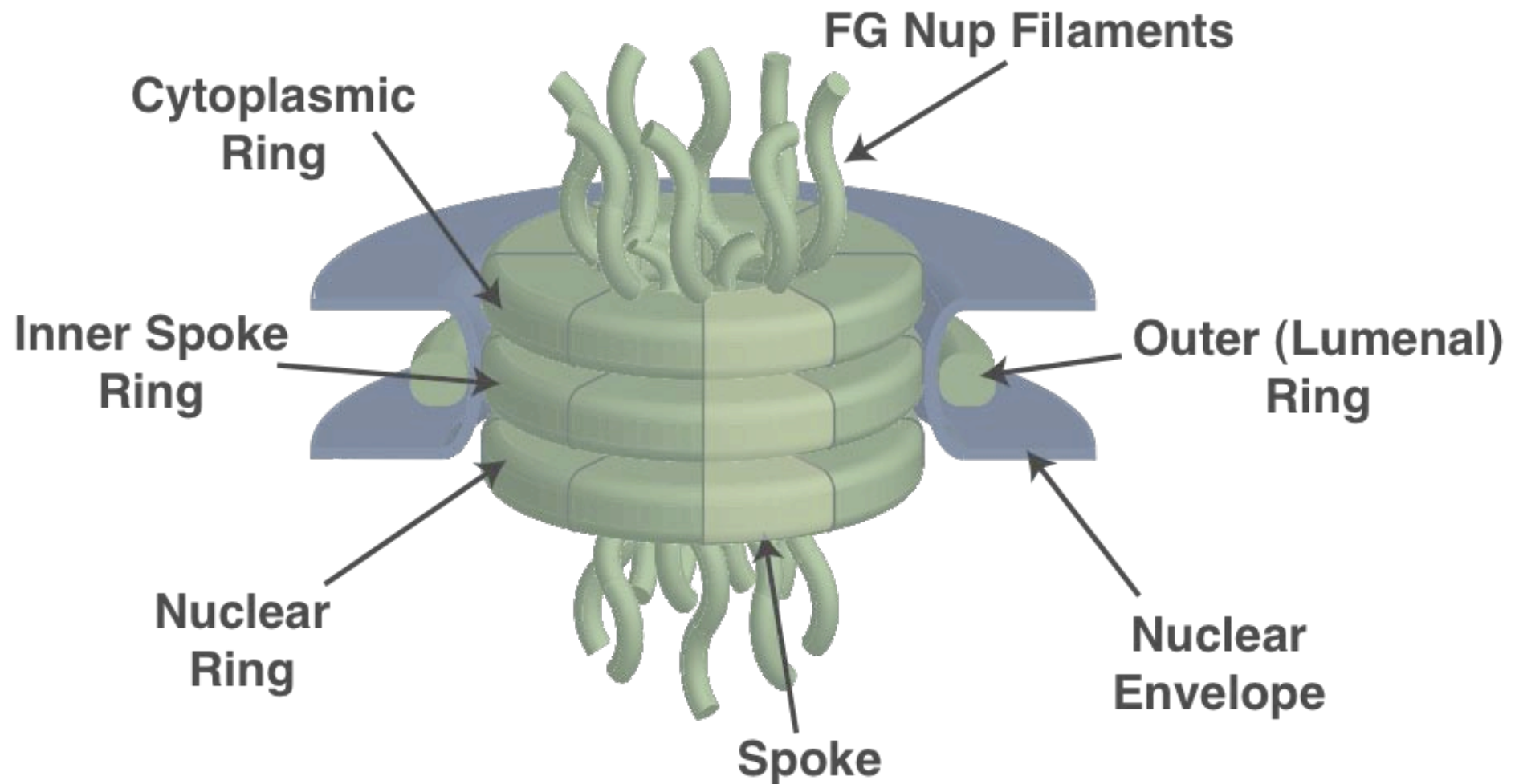
Alber *et al.* *Nature* **450**, 683-694, 2007.

Robinson, Sali, Baumeister. *Nature* **450**, 974-982, 2007.

Alber, Foerster, Korke, Topf, Sali. *Annual Reviews in Biochemistry* **77**, 11.1–11.35, 2008.



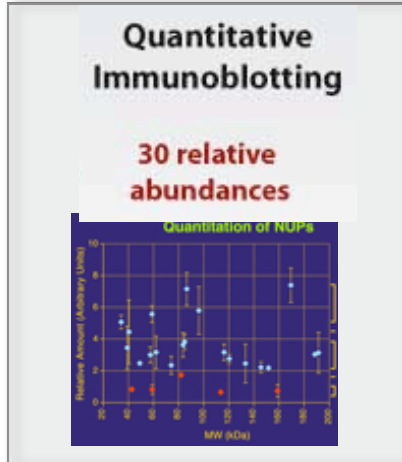
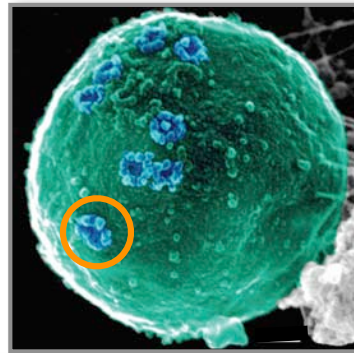
# What was known about the NPC structure?



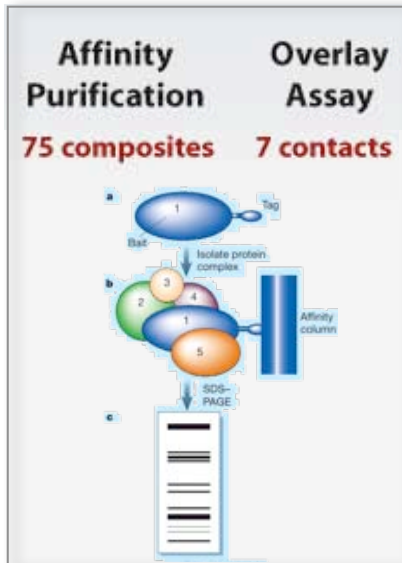
R. Milligan, W. Baumeister, O. Medalia, G. Blobel, E. Hurt, U. Aebi, T. Schwartz, M. Stewart, C. Akey, M. Rout, ...

# Configuration of 456 proteins in the Nuclear Pore Complex

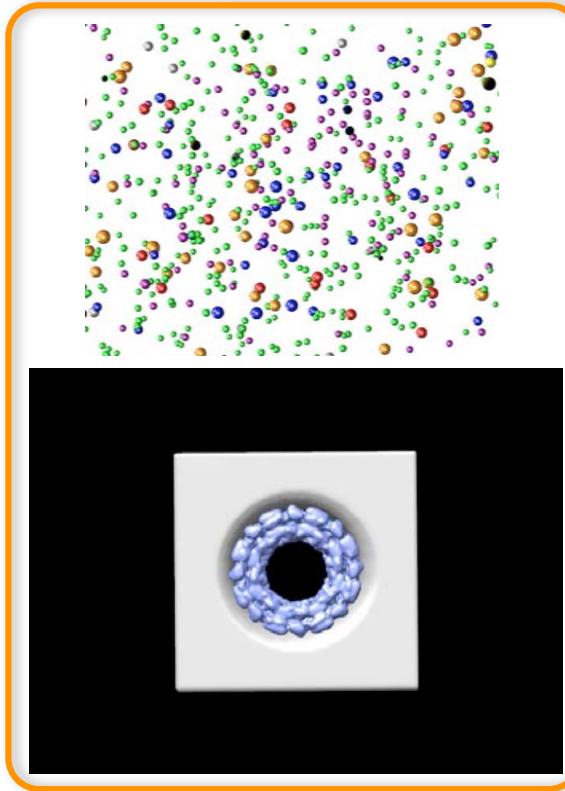
Alber *et al.* *Nature* 450, 684-694, 2007.  
Alber *et al.* *Nature* 450, 695-702, 2007.



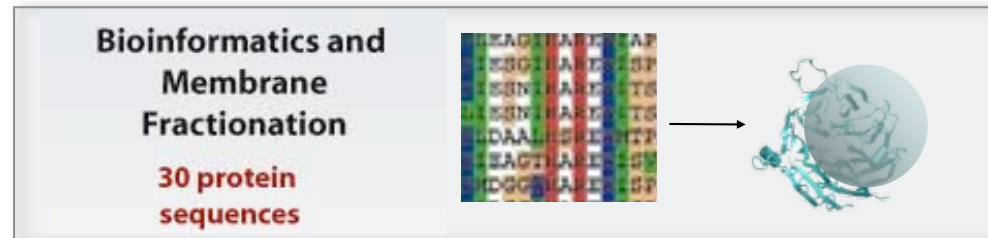
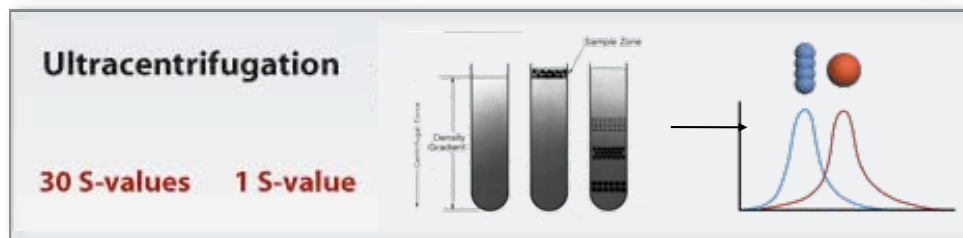
**Protein Stoichiometry**



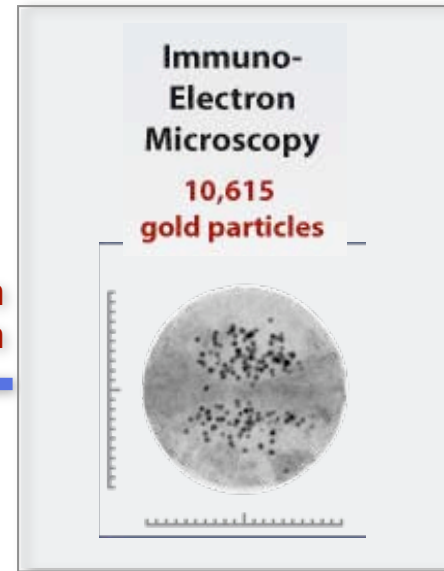
**Protein-protein Proximities**



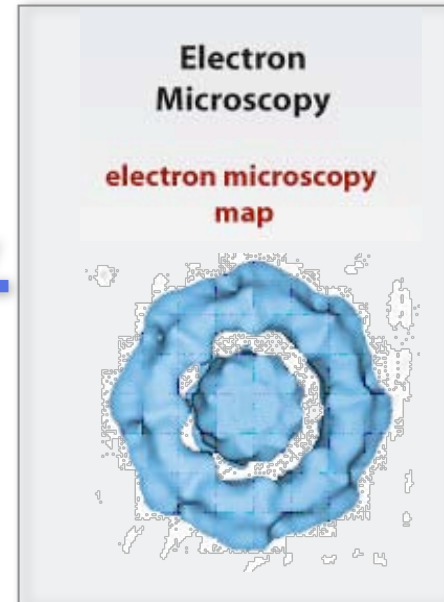
**Protein Shape**



**Protein Localization**

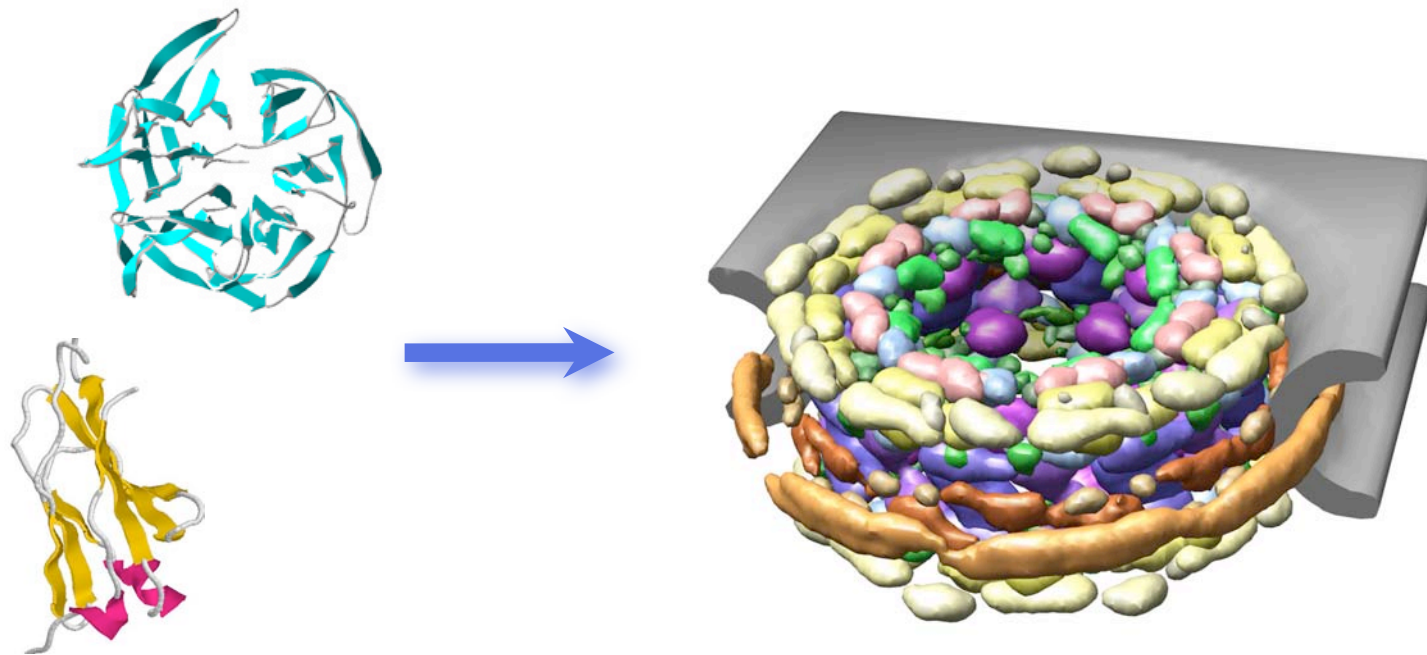
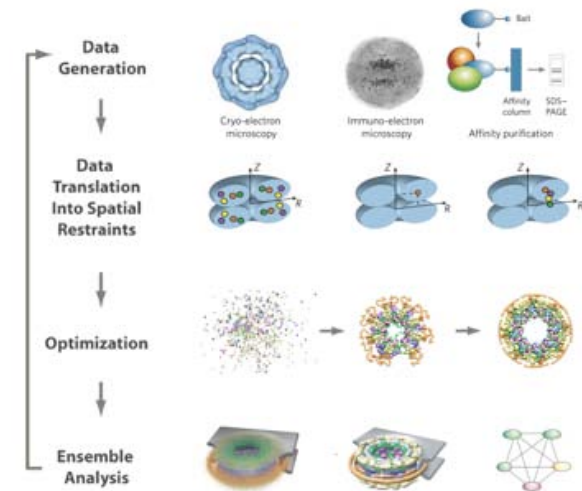


**Symmetry**

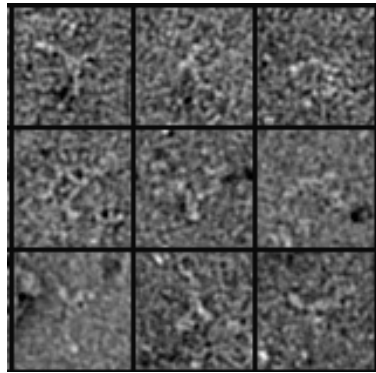
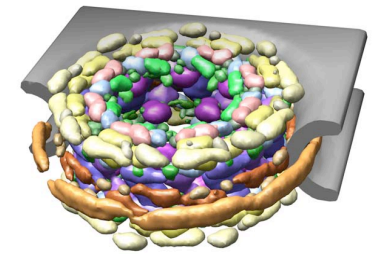


# Towards a higher resolution structure of the NPC

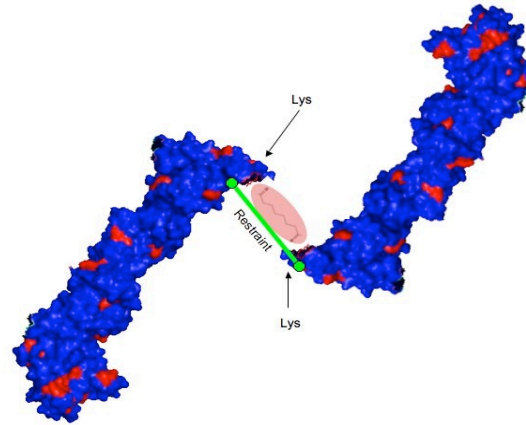
Characterize structures of the individual subunits, then fit them into the current low-resolution structure, aided by additional experimental information.



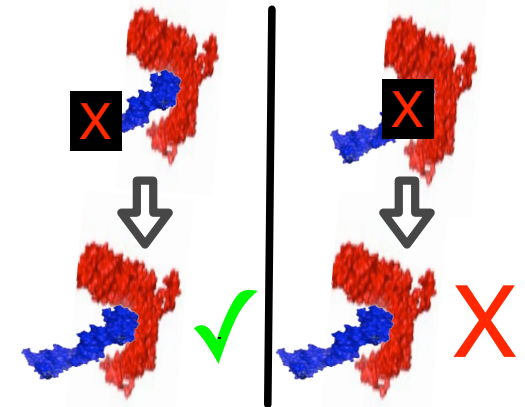
# Determining a pseudo-atomic structure of the Nuclear Pore Complex?



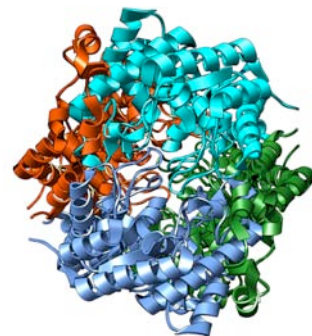
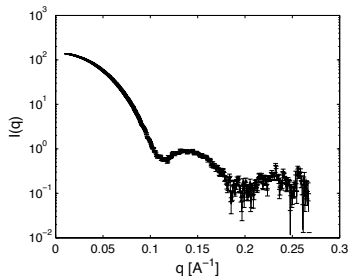
**EM density fitting**  
D. Stokes



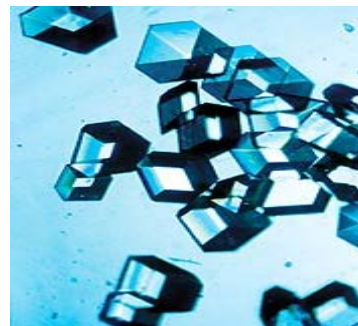
**Chemical crosslinking**  
B. Chait, M. Rout



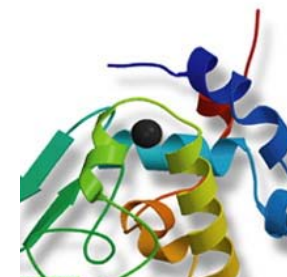
**Domain-deletion pullouts**  
M. Rout



**SAXS data**  
NYSGXRC  
H. Tsuruta, SSRL



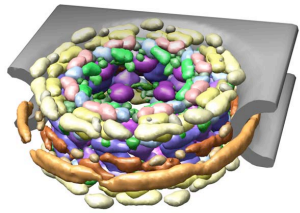
**X-ray crystallography**  
R. Stroud, CSMP  
S. Burley, NYSGXRC



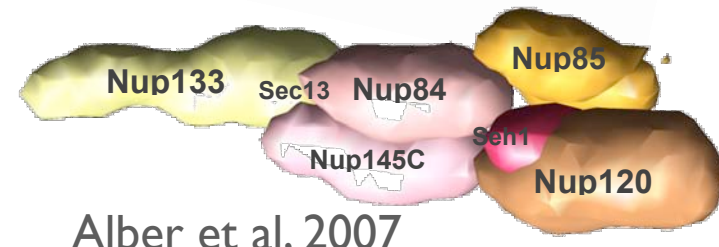
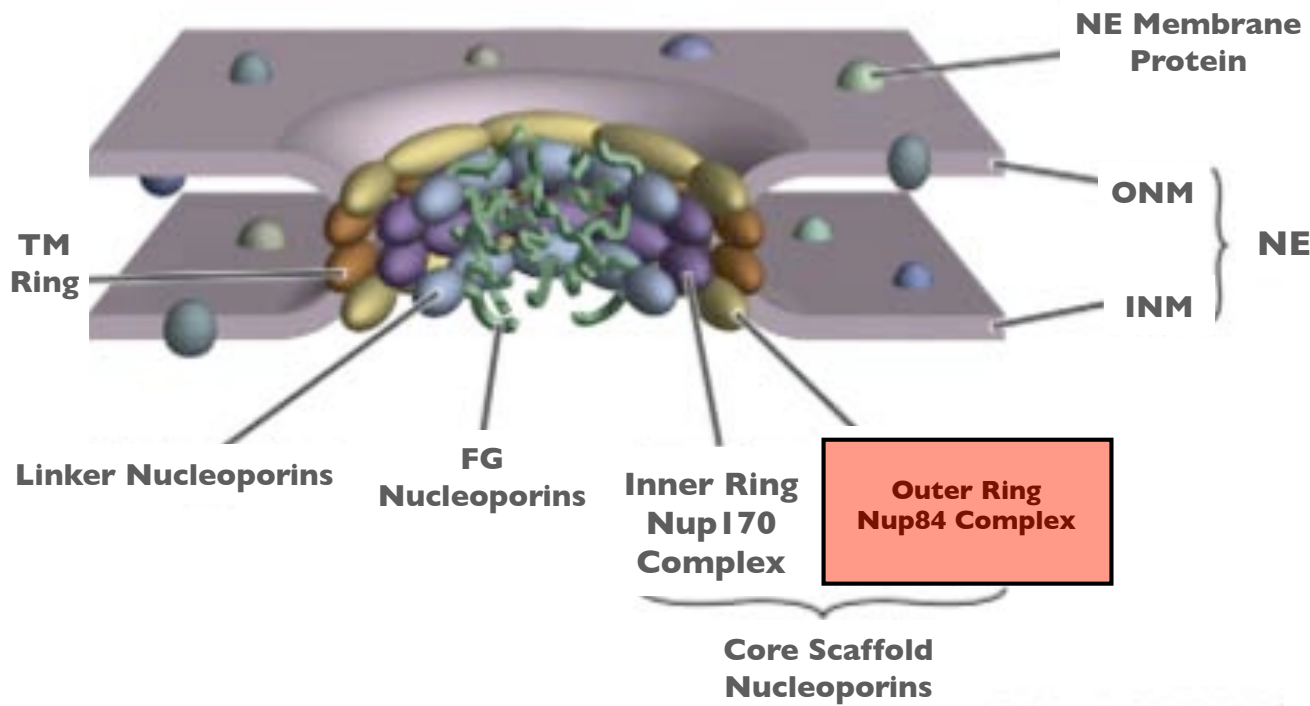
**Homology modeling**  
**molecular docking**  
A. Sali

```

A I L V G S M P R R D G M E R K D L L K A N V K I F K C G A
V E V C P V D C F Y E S P N F L V I H P D E C I D C A L C E R
M A C K P E C P V N I Q G S - - L Y A D A N S I D C S
C - - I A C G A C K P E C P V N I Q G S - - L Y A I D A D S
    
```



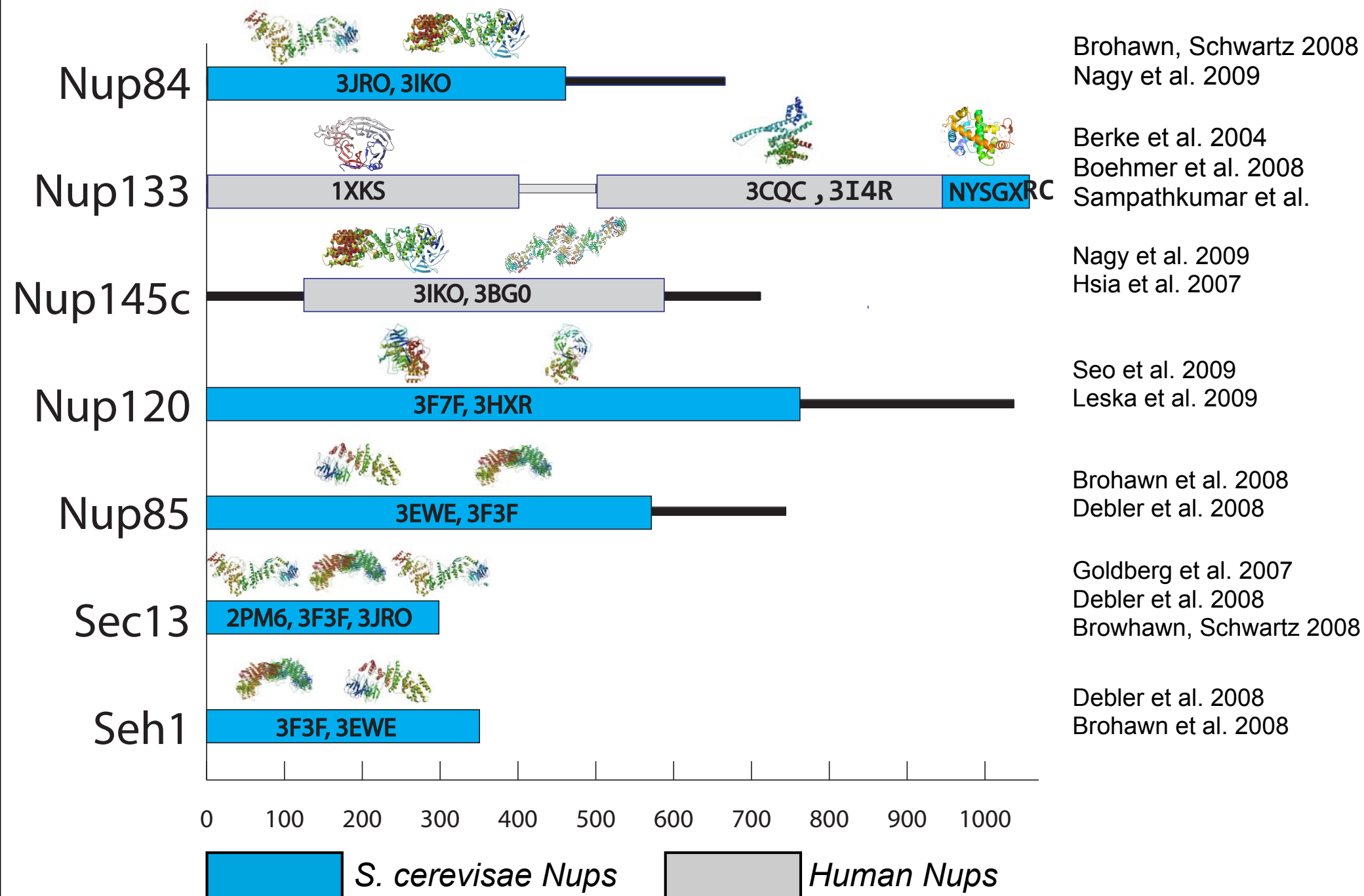
# The Nup84 complex in the NPC



Alber et al, 2007

- 7-protein complex
- Forms the two outer rings of the NPC
- Present in 16 copies in the NPC
- Proteins share a common ancestor with vesicle coating complexes

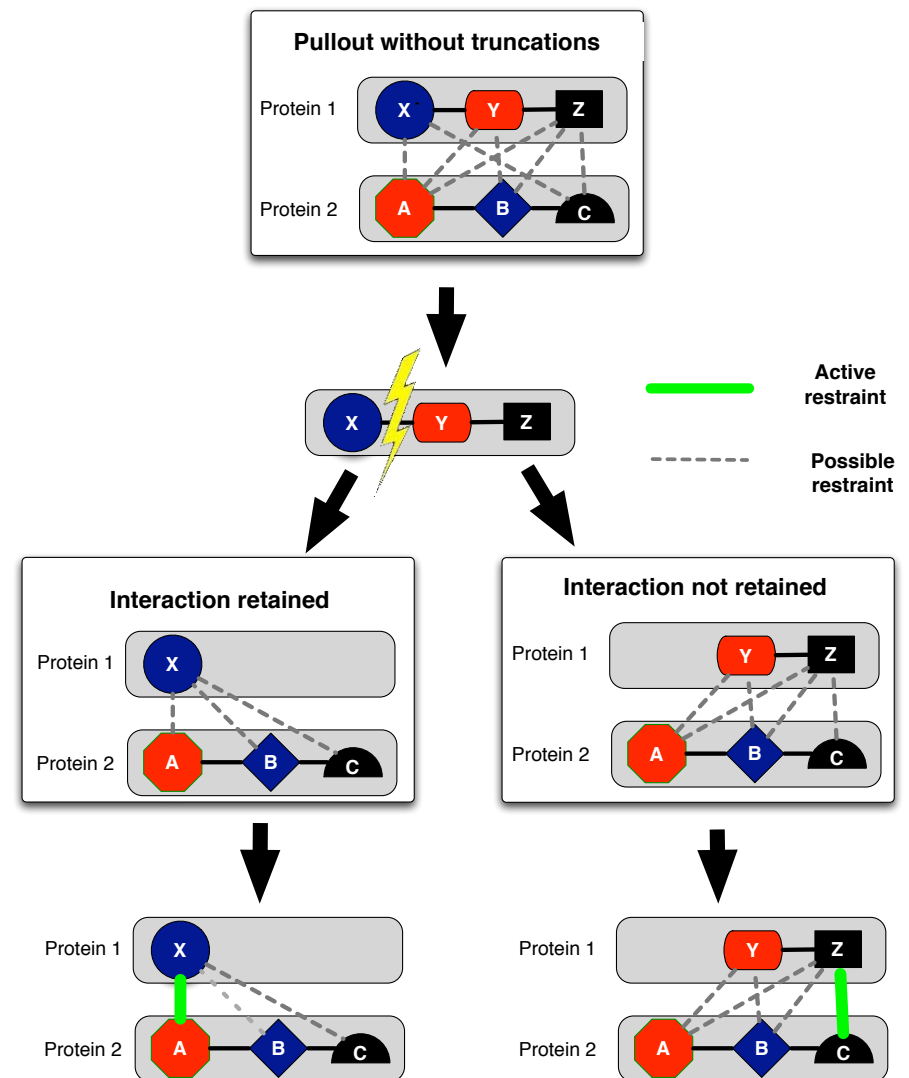
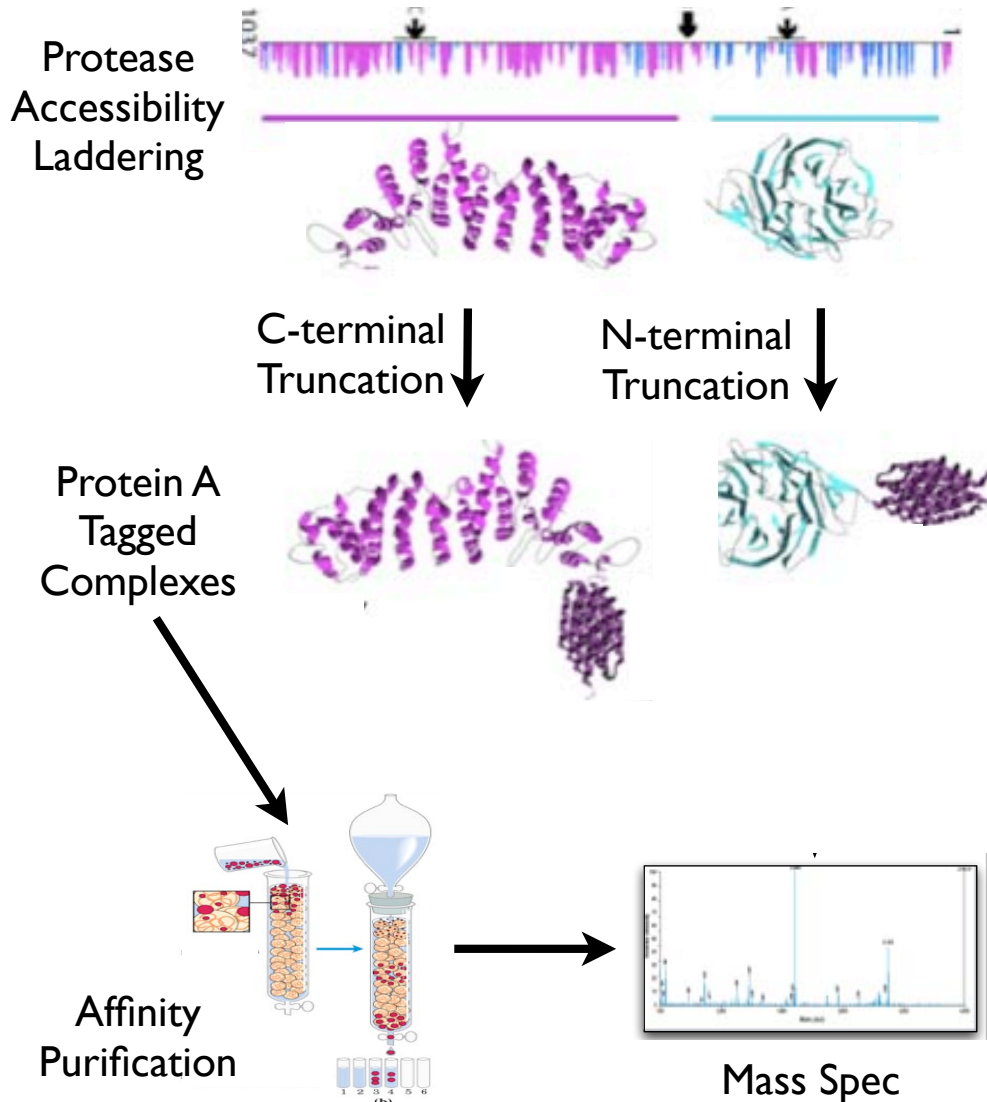
# Structural Coverage of the Nup84 Complex





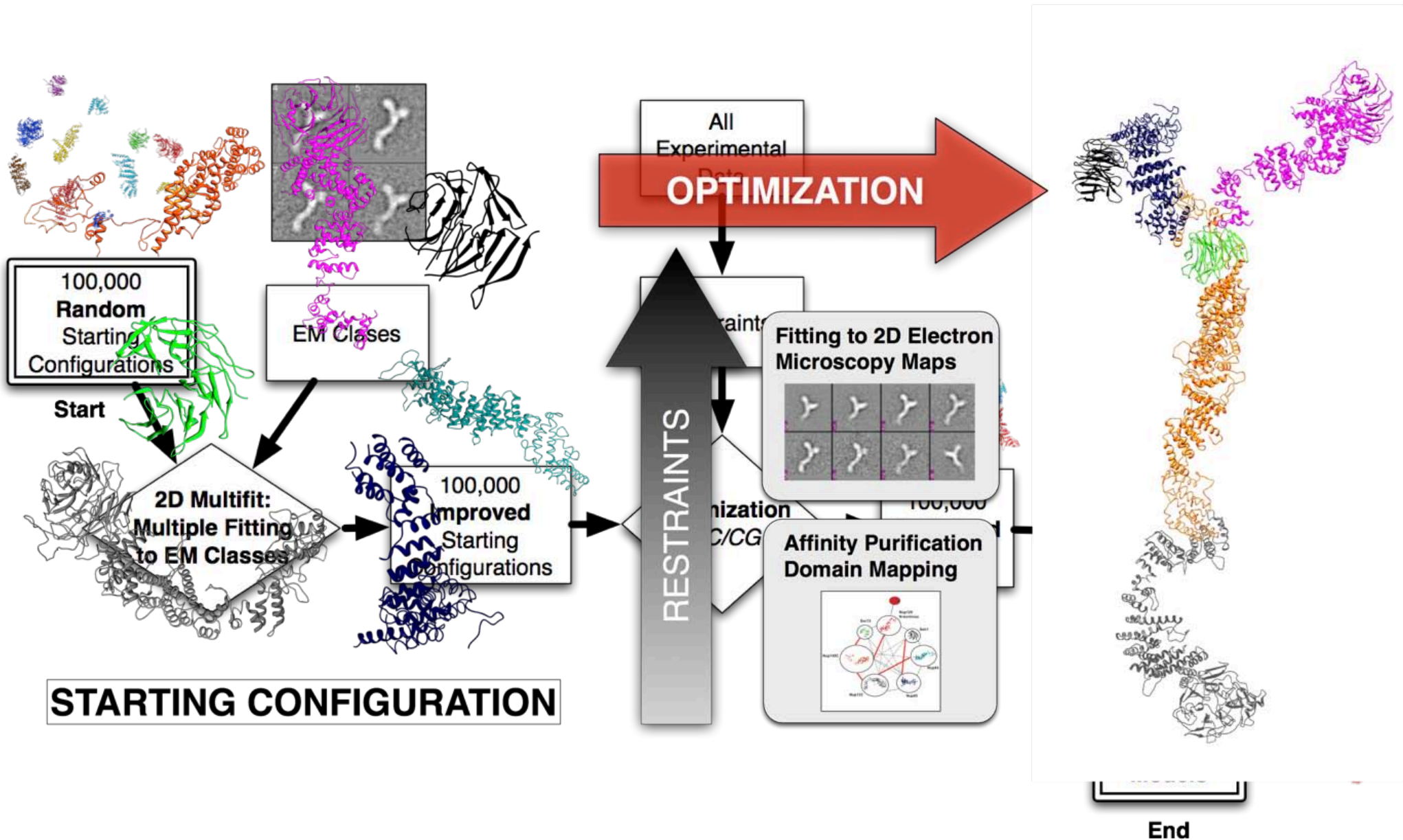
# Domain Mapping to identify interacting domains: From position to orientation

Domain-Mapped Affinity Purification → Domain Connectivity Restraints

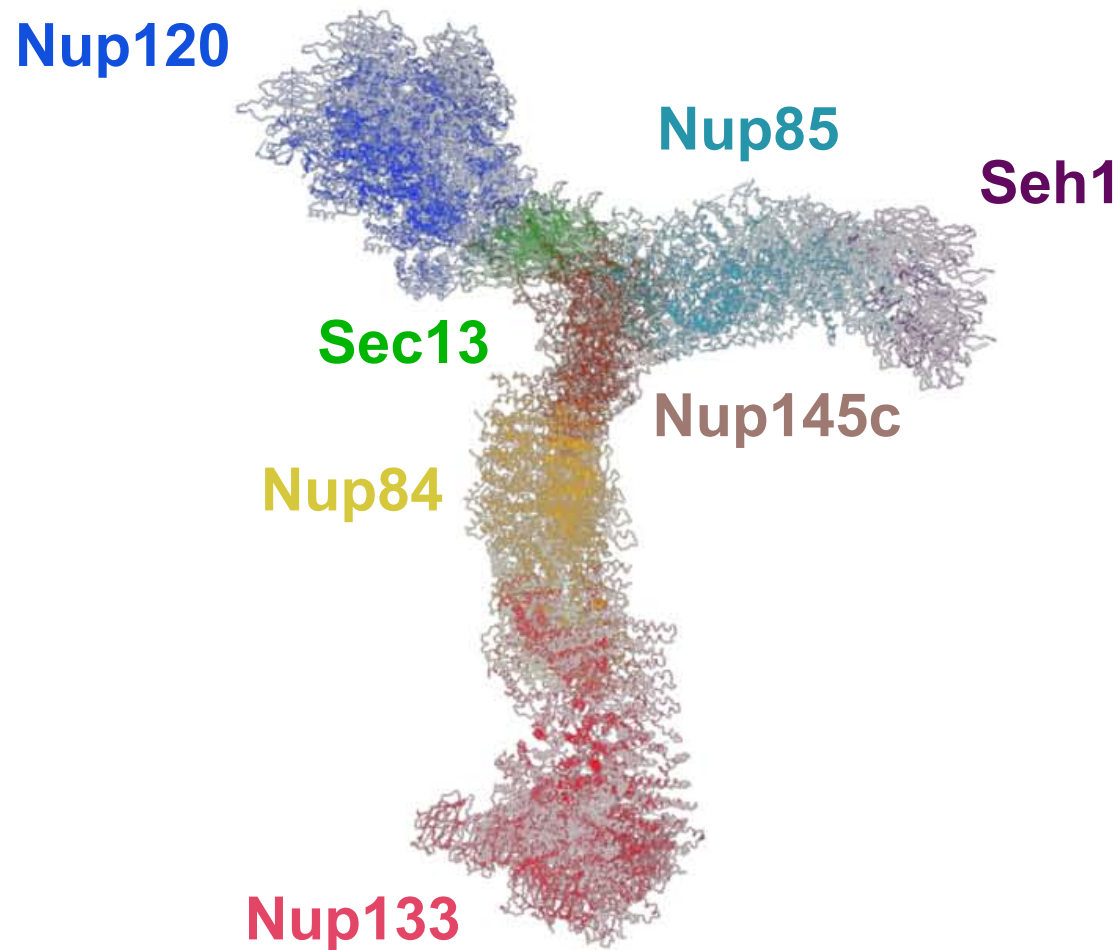




# Nup84 complex: Optimization



# Nup84 complex: Ensemble of good scoring solutions



- ~5000 good scoring structures (representative sample shown)
- All restraints are satisfied (2D-EM, domain deletion, ...)
- Domain-domain orientations are resolved uniquely.

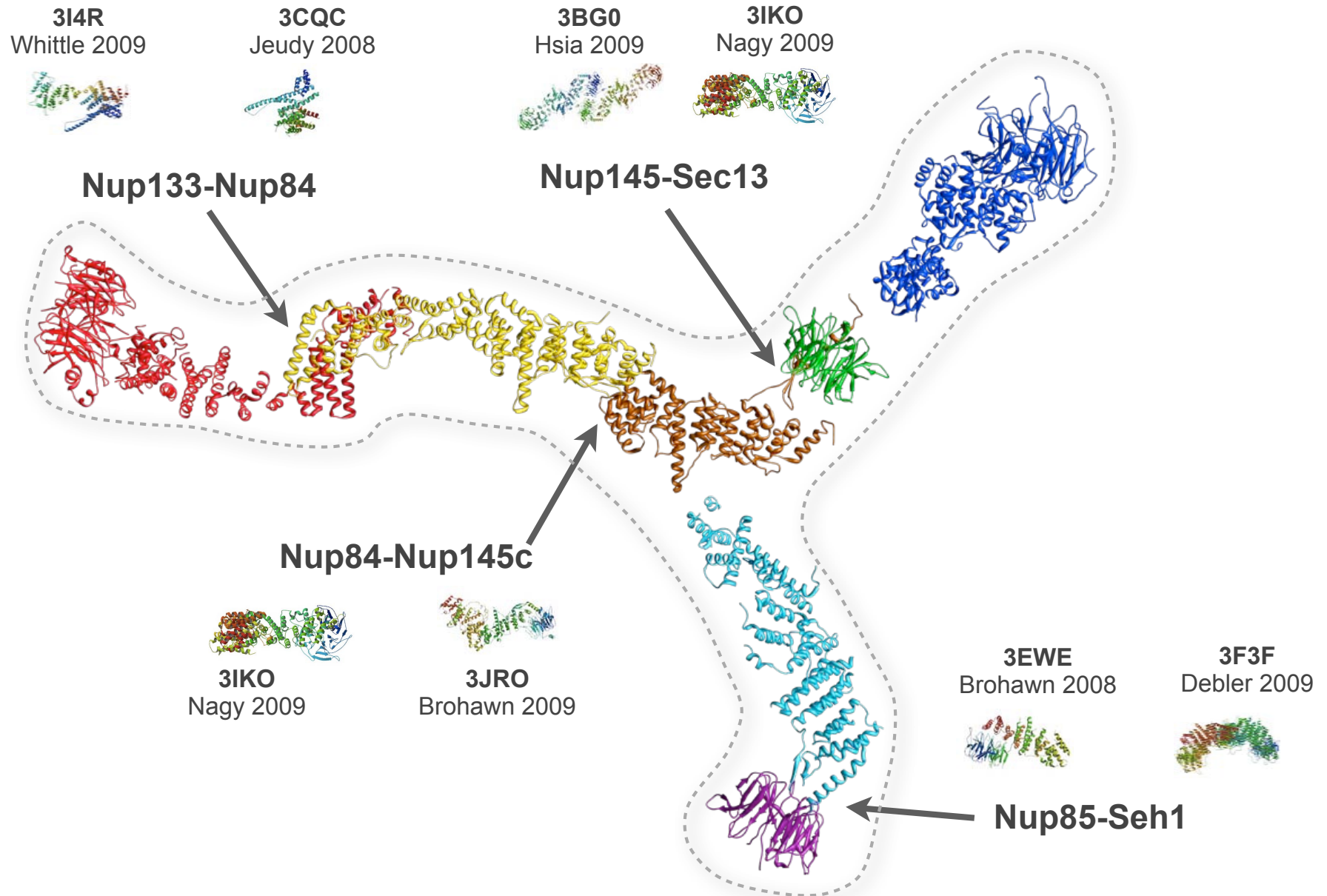
# How accurate is a model?

## Assessing the well-scoring models

*Alber et al. Nature 450, 695-702, 2007.*

1. Self-consistency of independent experimental data.
2. Structural similarity among the configurations in the ensemble that satisfy the input restraints.
3. Simulations where a native structure is assumed, corresponding restraints simulated from it, and the resulting calculated structure compared with the assumed native structure.
4. Patterns emerging from a mapping of independent and unused data on the structure that are unlikely to occur by chance.
5. Experimental spatial data that were not used in the calculation of the structure.

# Model is consistent with crystal contacts in PDB files with pairs of subunits



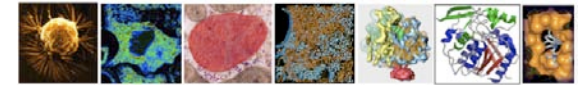
# In Conclusion

The goal is a comprehensive description of the multitude of interactions between molecular entities, which in turn is a prerequisite for the discovery of general structural principles that underlie all cellular processes.



This goal will be achieved by a **formal** integration of **experiment**, **physics**, and **statistical inference**, spanning all relevant size and time scales.

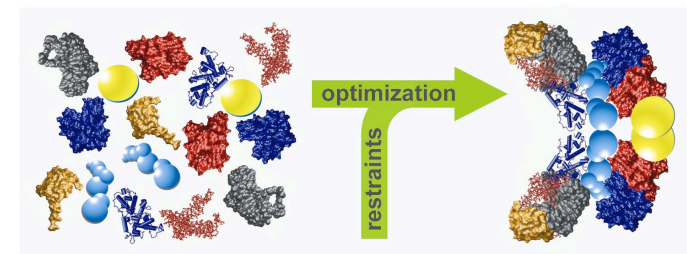
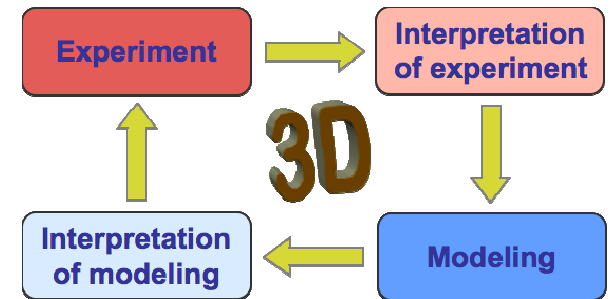
**Need for a formal 3D model as a reflection of all data that all contributors are continually referring to and updating.**



X-ray crystallography	NMR spectroscopy	2D & single particle electron microscopy	electron tomography	immuno-electron microscopy	chemical cross-linking	affinity purification mass spectroscopy
subunit structure	subunit structure	subunit shape	subunit shape	subunit structure	subunit structure	
subunit shape	subunit shape	subunit-subunit contact	subunit-subunit contact	subunit-subunit contact	subunit-subunit contact	subunit-subunit contact
subunit-subunit contact	subunit-subunit contact	subunit proximity	subunit proximity	subunit proximity	subunit proximity	subunit proximity
subunit proximity	subunit proximity	subunit stoichiometry	subunit stoichiometry	subunit stoichiometry	subunit stoichiometry	subunit stoichiometry
subunit stoichiometry	subunit stoichiometry	assembly symmetry	assembly symmetry	assembly symmetry	assembly symmetry	assembly symmetry
assembly symmetry	assembly symmetry	assembly shape	assembly shape	assembly shape	assembly shape	assembly shape
assembly shape	assembly shape	assembly structure	assembly structure	assembly structure	assembly structure	assembly structure
assembly structure	assembly structure					

FRET	site-directed mutagenesis	yeast two-hybrid system	gene/protein arrays	protein structure prediction	computational docking	bioinformatics
subunit-subunit contact	subunit-subunit contact	subunit-subunit contact	subunit-subunit contact	subunit structure	subunit-subunit contact	subunit-subunit contact
subunit-subunit contact	subunit-subunit contact	subunit proximity	subunit proximity	subunit shape	subunit-subunit contact	subunit-subunit contact



Sali, Earnest, Glaeser, Baumeister. From words to literature in structural proteomics. *Nature* 422, 216-225, 2003.  
 Robinson, Sali, Baumeister. The molecular sociology of the cell. *Nature* 450, 974-982, 2007.  
 Alber, Foerster, Korkin, Topf, Sali. *Annual Reviews in Biochemistry* 77, 11.1-11.35, 2008.  
 Russell, Lasker, Velazquez, Phillips, Schneidman, Sali. *Current Opinion in Cell Biology* 21, 1-12, 2009

# Acknowledgments

(<http://salilab.org>)

## QB3 @ UCSF

**Keren Lasker**

**Ben Webb**

Jeremy Phillips

Daniel Russel

Javier Velazquez-Muriel

Seung Joong Kim

Dina Schneidman

Elina Tijoe

## Alumni:

Frank Alber (USC)

Damien Devos (EMBL)

Maya Topf (Birkbeck College)

Narayanan Eswar (E. Lilly)

Marc Marti-Renom (Valencia)

Mike Kim (Mechanical Zoo)

Dmitry Korkin (UM, Columbia)

Fred Davis (HHMI)

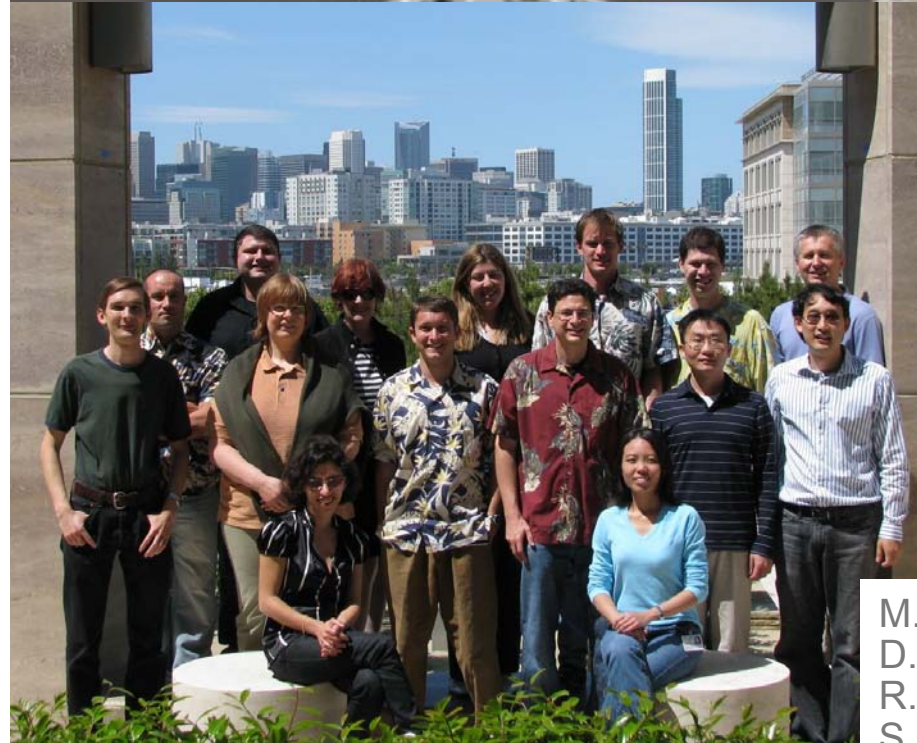
M. Madhusudhan (Singapore)

D. Eramian (UCSF)

Min-Yi Shen (Applied Biosys)

Friederich Foerster (MPI)

Bret Peterson (Google)



## Rockefeller University

**Mike Rout**

Javier Fernandez

Jody Franke

Jaclyn Novatt

Loren Hough

John LaCava

**Brian Chait**

Jeff DeGrasse

M. Field, Cambridge University

D. Stokes, NYSBC

R. Diaz-Avalos, NYSBC

S. Burley, NYSGXRC

P. Sampathkumar, NYSGXRC

H. Tusruta,, SLAC

A. Martel, SLAC

J. Tainer, LBNL

R. Stroud, CSMP

**Wah Chiu (Baylor)**

**Chris Akey (BU)**

**Joachim Frank (Columbia U)**

**John Aitchison (ISB)**

**David Agard (UCSF)**

**Wolfgang Baumeister (MIPS)**

**Haim Wolfson (TAU)**

NIH, NSF

The Sandler Family Foundation

Human Frontiers Science Program

IBM, Intel, Hewlett-Packard, NetApps

Pfizer, Structural Genomix

Pharmaceuticals, Mike Homer,

Ron Conway