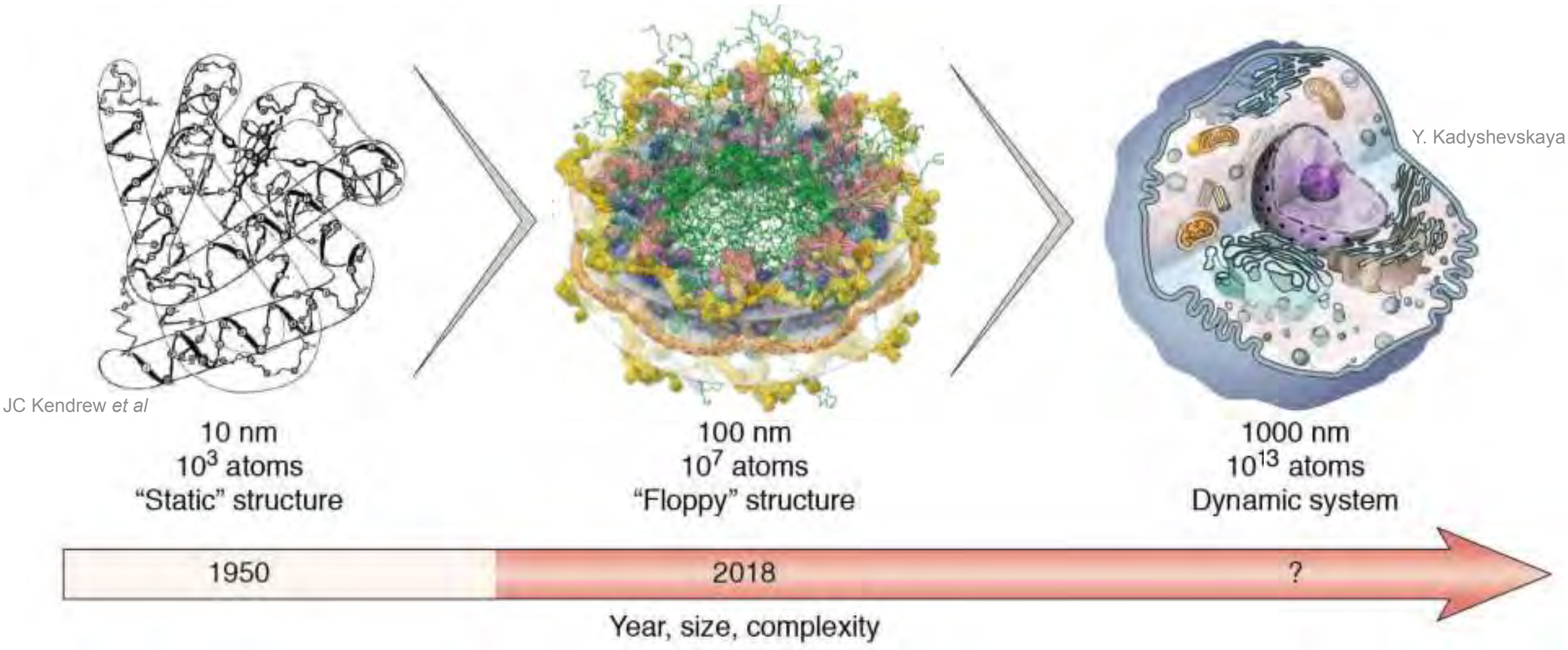
Spatiotemporal modeling of complex biological systems: From data integration to model integration

Andrej Sali
http://salilab.org/

Department of Bioengineering and Therapeutic Sciences
Department of Pharmaceutical Chemistry
California Institute for Quantitative Biosciences
University of California, San Francisco
Research Collaboratory for Structural Bioinformatics

Our goal

(R)Evolution of structural biology: Progress towards the cell model



Y. Kadyshevskaya

Sali. *JBC*, 2021

Our goal:

Develop, apply, and disseminate integrative spatiotemporal modeling methods to facilitate biological and biomedical discovery.

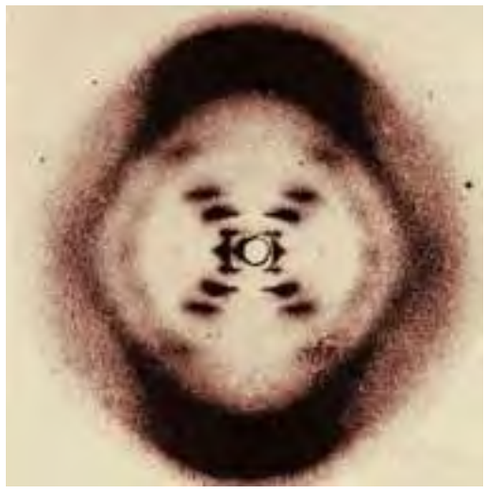
Our approach

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

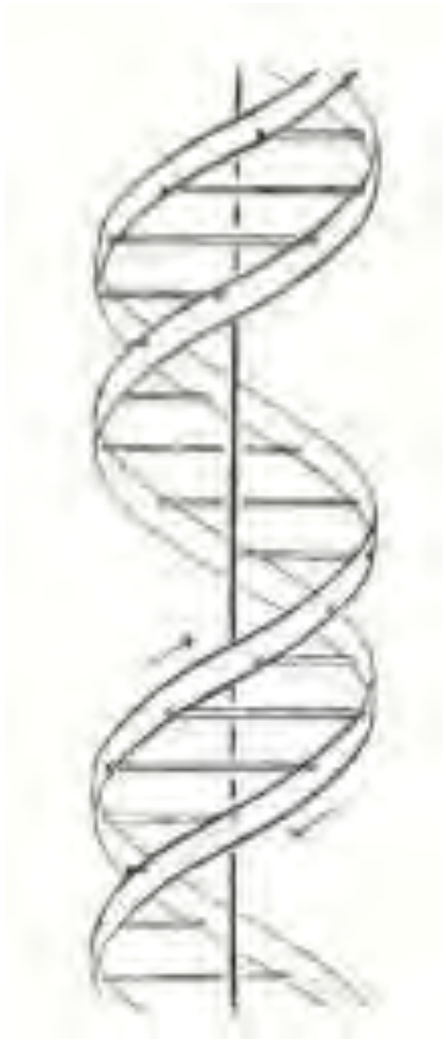
A Structure for Deoxyribose Nucleic Acid

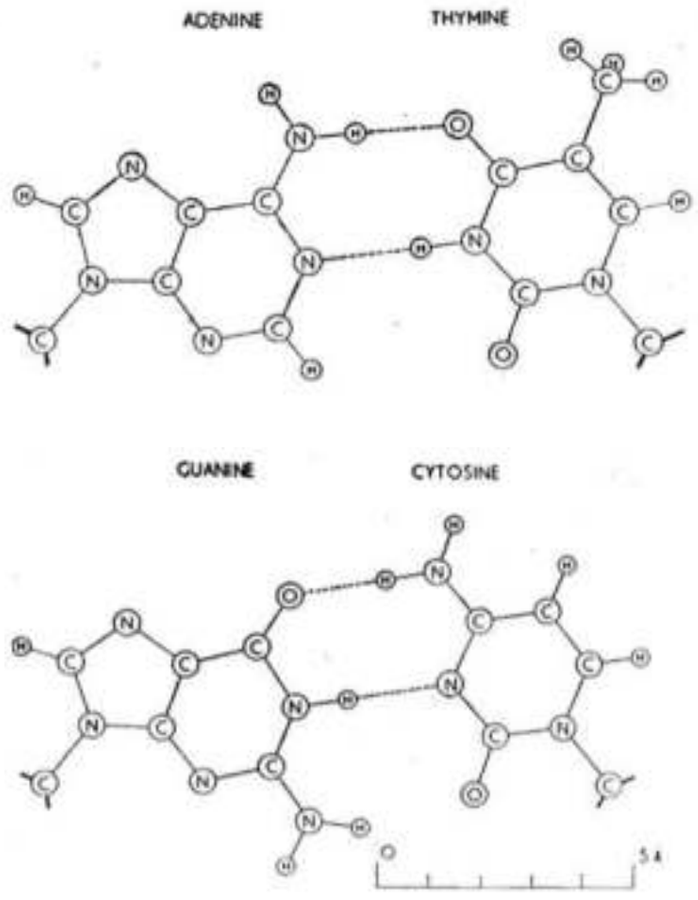
J. D. WATSON
F. H. C. CRICK

No. 4356 April 25, 1953
NATURE



X-ray diffraction





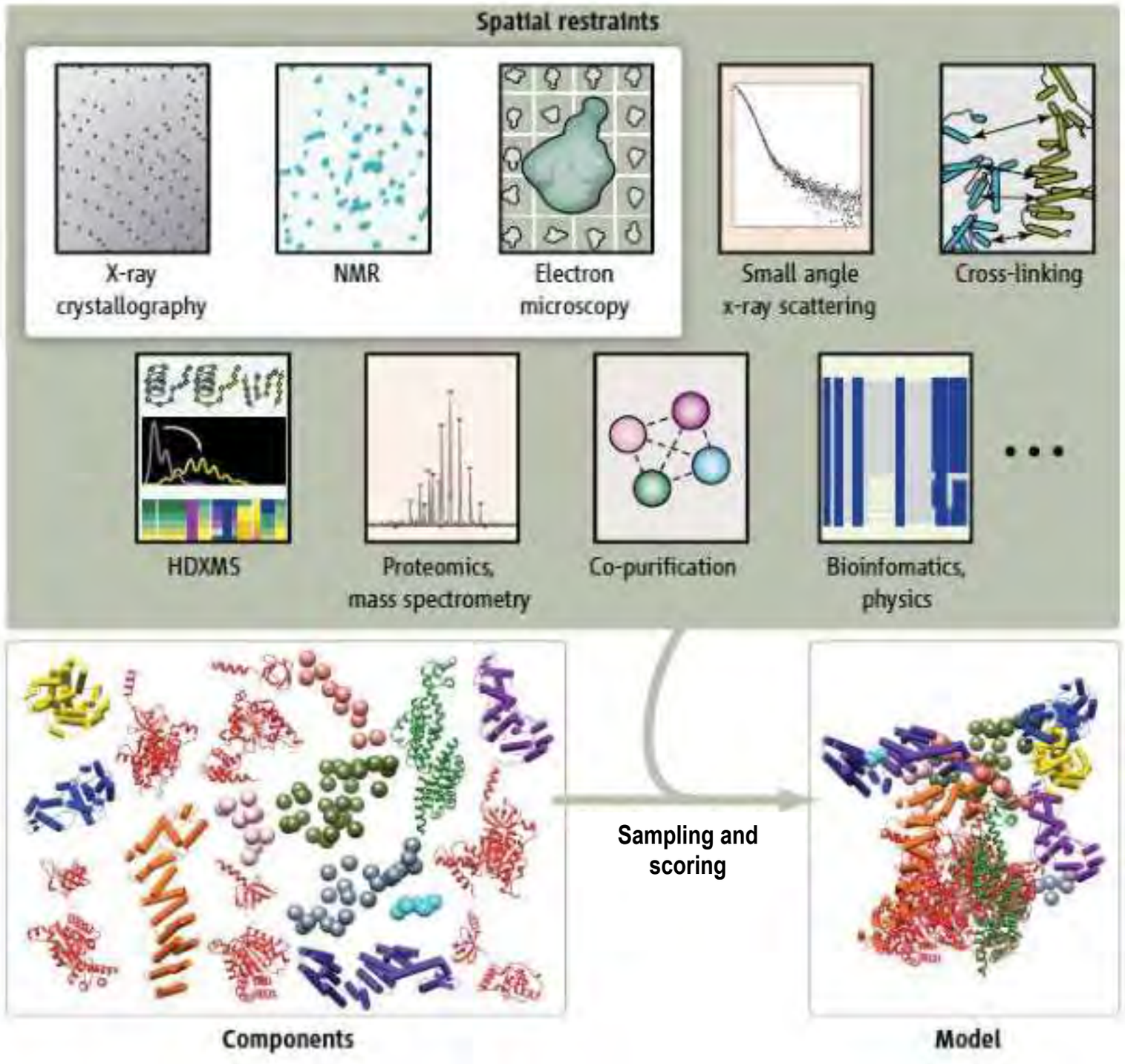
Composition
Stoichiometry
Chemical complementarity

To understand and modulate cellular processes, we need their models.

These models are best generated by considering all available information.

Integrative modeling

- Uses multiple types of information (experiments, physical theories, statistical inferences, prior models).
- Maximizes accuracy, precision, completeness, and efficiency of modeling.
- Finds all models for which computed data match the experimental data within an acceptable threshold.
- Bayesian formulation; different model representations.



Sali *et al*. *Nature* **422**, 216-225, 2003.
Alber *et al*. *Nature* **450**, 683-694, 2007
Robinson *et al*. *Nature* **450**, 974-982, 2007
Alber *et al*. *Ann.Rev.Biochem.* **77**, 11.1–11.35, 2008
Russel *et al*. *PLoS Biology* **10**, 2012
Ward *et al*. *Science* **339**, 913-915, 2013
Schneidman *et al*. *Curr.Opin.Str.Biol*., 96-104, 2014.
Sali *et al*. *Structure* **23**, 1156-1167, 2015.
Berman *et al, Structure* **27**, 92-102, 2019.
Rout & Sali. *Cell* **177**, 1384-1403, 2019.
Sali. *JBC*, 2021

$$p(M|D,I) \propto p(D|M,I) \cdot p(M|I)$$

A model is built iteratively, contributes continuously.

While it may be hard to live with generalization, it is inconceivable to live without it.
Peter Gay, Schnitzler's Century (2002).

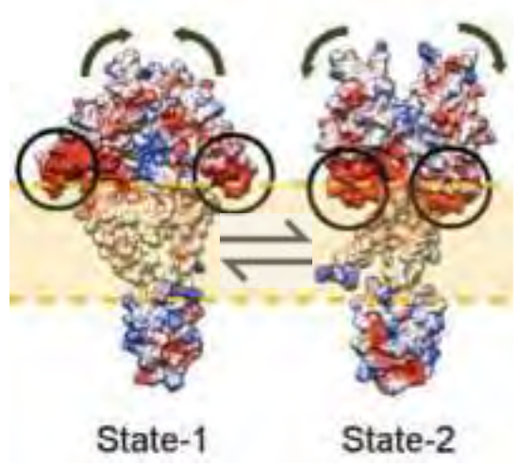
Integrative modeling is applicable to different model representations

Schneidman *et al*, *Bioinformatics,* 2013

Single static
atomic structure

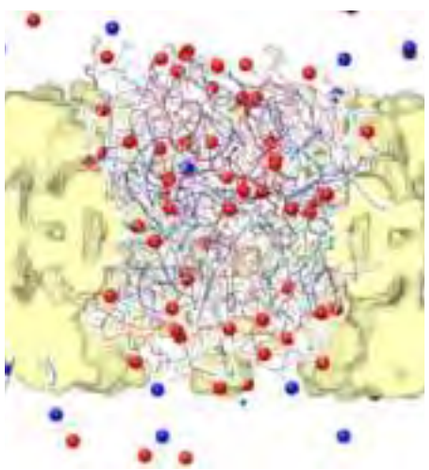
Kim *et al*, *Nature*, 2018

Single static
coarse-grained structure



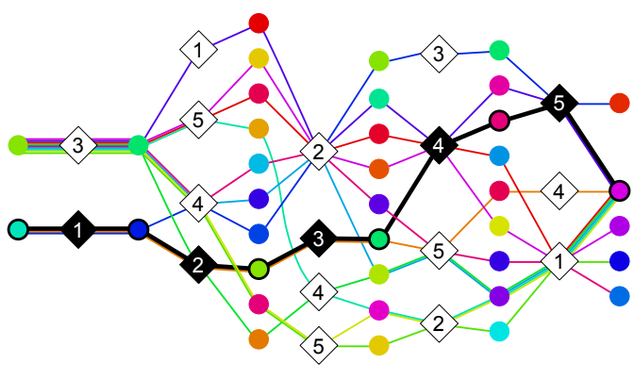
Molnar *et al, Structure*, 2014

Multi-state structure



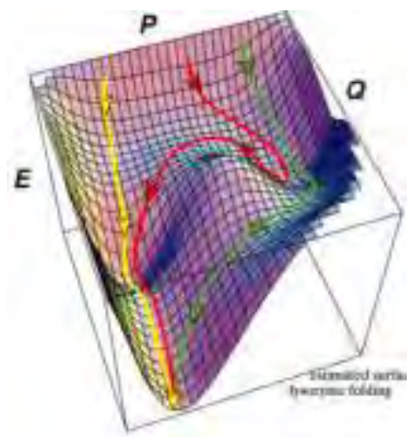
Raveh *et al*

Dynamic process



Calhoun *et al*, *eLife,* 2018

Molecular network



Chemmama *et al*

Energy landscape
(structure, stability, kinetics)

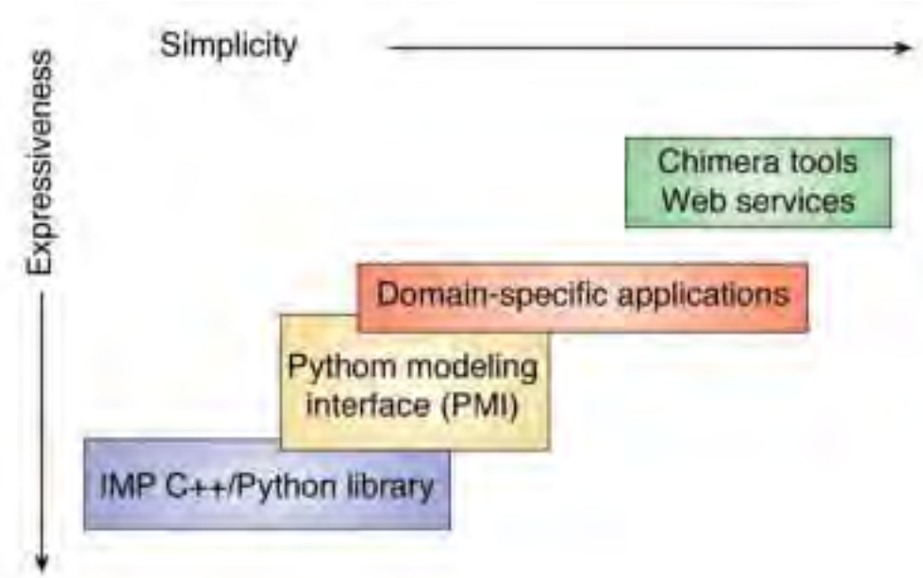
Integrative Modeling Platform (IMP)

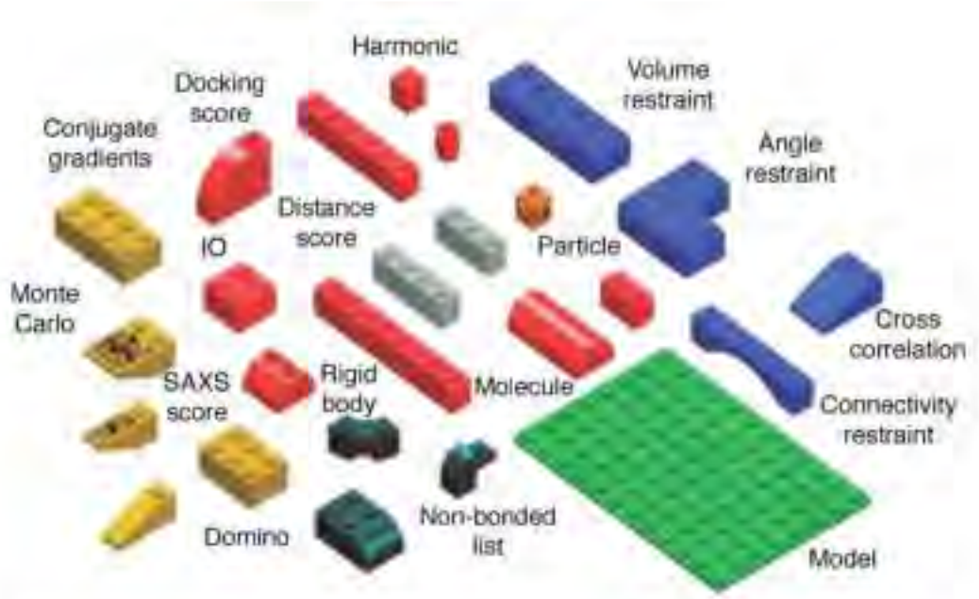
http://integrativemodeling.org

D. Russel, K. Lasker, B. Webb, J. Velazquez-Muriel, E. Tjioe, D. Schneidman, F. Alber, B. Peterson, A. Sali, PLoS Biol, 2012.

R. Pellarin, M. Bonomi, B. Raveh, S. Calhoun, C. Greenberg, G.Dong, S.J. Kim, I. Chemmama, D. Saltzberg, S. Viswanath, S. Axen, G. Sai, I. Echeverria, R. Ramachandran, J. Tempkin, T. Sanyal, A. Palar, M. Hancock, D. Mondal

Open source, versions, documentation, wiki, examples, mailing lists, unit testing, bug tracking, ...



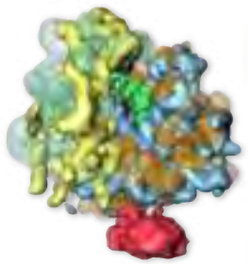


Representation:	Scoring:	Sampling:	Analysis:
Spatiotemporal Resolution Atomic [121] Coarse-grained [4] Multi-scale [4] Ridig bodies [124] Symmetry [4,45] Single state [34] Multiple states [86,126] *Molecular network* [128]	EM Particle image 2D class average [36,122] 3D density map [123] Immuno-EM [17] Cross-linking with MS [33] Native MS [125] Cys-scanning mutagenesis [86] FRET spectroscopy [45,127] Small-angle scattering [121,129] H/D exchange by MS [130] Second harmonic generation Point mutations [131] Genetic interactions [68] Affinity co-purification [17] Molecular mechanics force field [121] Statistical potentials [134] Homology-derived restraints [121,135] Ambiguous restraints [17,19] Bayesian likelihoods [5,123]	*Optimization* Simplex [36] Steepest descent Conjugate gradients *Sampling* Monte Carlo [123] Simulated annealing Replica exchange [126] Molecular dynamics [121] Brownian dynamics [4] Gibbs sampling [4] *Enumeration* Dynamic programming Divide-and-conquer enumeration [29,132] Rapidly-exploring random tree [66,133]	Data violation [4] Model clustering [9] Model validation [9]

Sali. *JBC*, 2021

Example integrative structures from our lab

Models allow us to understand how biological systems work, how they evolved, how they can be controlled, modified, and designed.



Ribosomes
Frank, Akey

PCSK9-Fab
Cheng, Agard, Pons

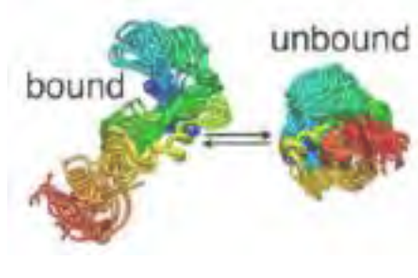
Actin
Chiu

TRiC/CCC
Frydman, Chiu

DNA-PKcs
Blundell

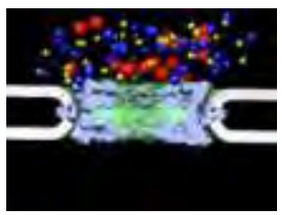
RyR channel
Serysheva, Chiu

Hsp90 landscape
Agard



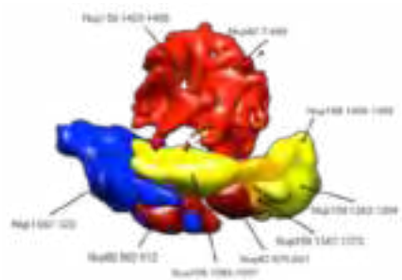
Substrate folding by Hsp90
Agard

Nuclear Pore Complex
Rout, Chait



NPC transport
Rout, Chait, Aitchison, Chook, Cowburn

Nup84 complex
Rout, Chait

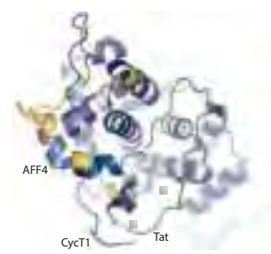


Nup82 complex
Rout, Chait

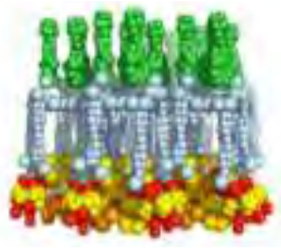
Nup133
Rout, Chait

SEA complex
Rout, Chait, Dokudovskaya

PDE6
Chu

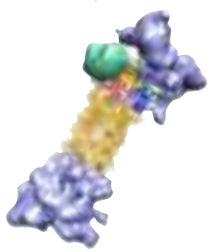


Tat:AFF4:P-TEFb:TAR
Hurley



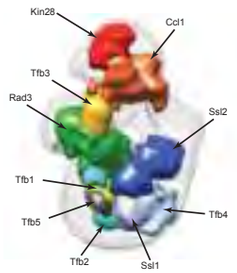
Spindle Pole Body
Davis, Muller

Microtubule nucleation
Agard



26 Proteasome
Baumeister

PhoQ His kinase
DeGrado



TFIIH
Ranish

MET antibodies
Chen, Vieth

40S-eIF1-eIF3
Aebersold,Ban

Prion aggregation
Prusiner, Degrado

Ecm29-proteasome Huang

Exosome
Rout, Chait

Mediator
Kornberg,Burlingame

Complement C3(H2O)
Rappsilber

Pol II (G)
Roeder

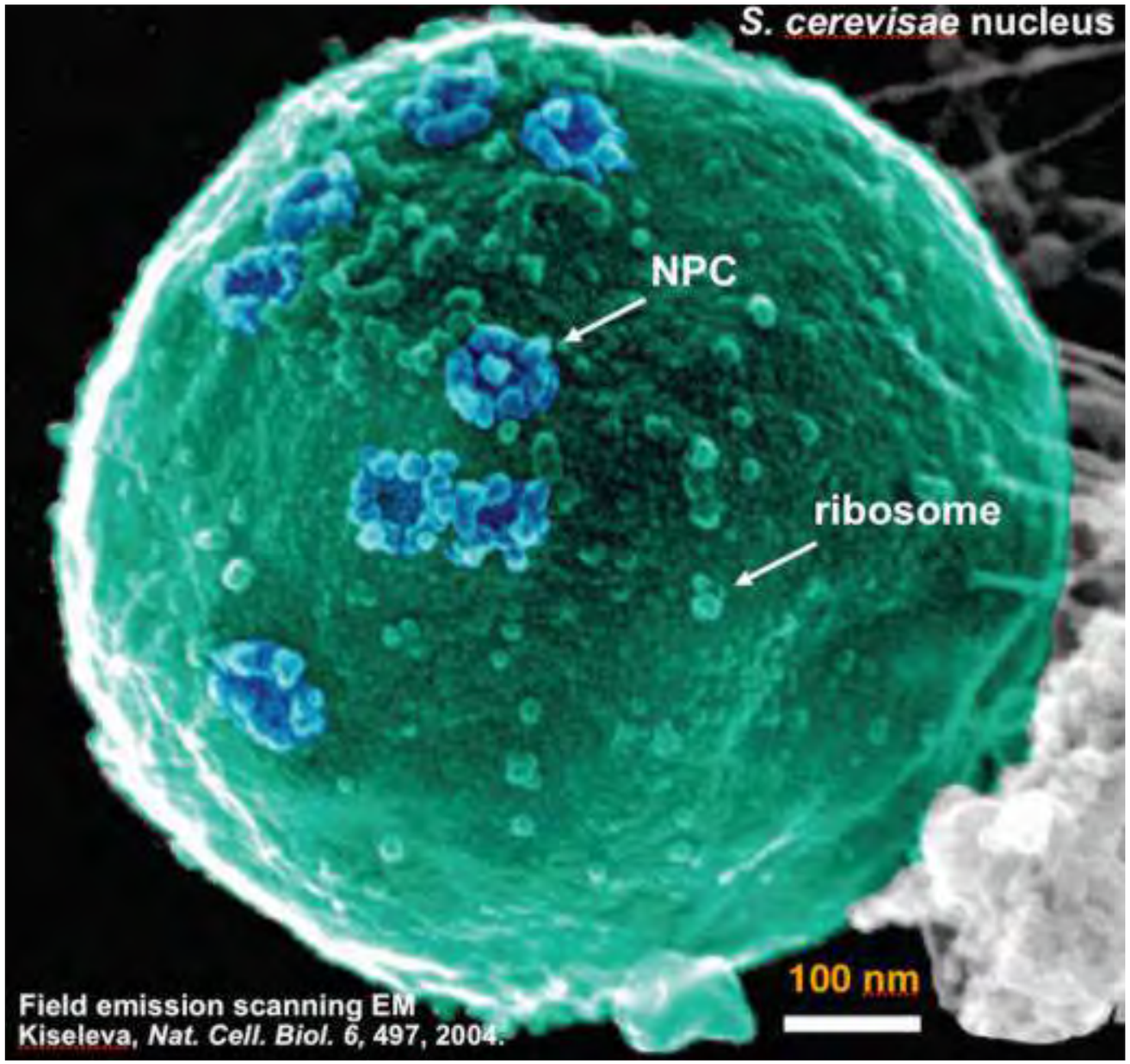
Pom152 ring complex
Rout

Rvb1/Rvb2/Ino80
Narlikar

pMHCII-TCR
Aswad

An example application

Nuclear Pore Complex (NPC)



Consists of broadly conserved nucleoporins (nups).

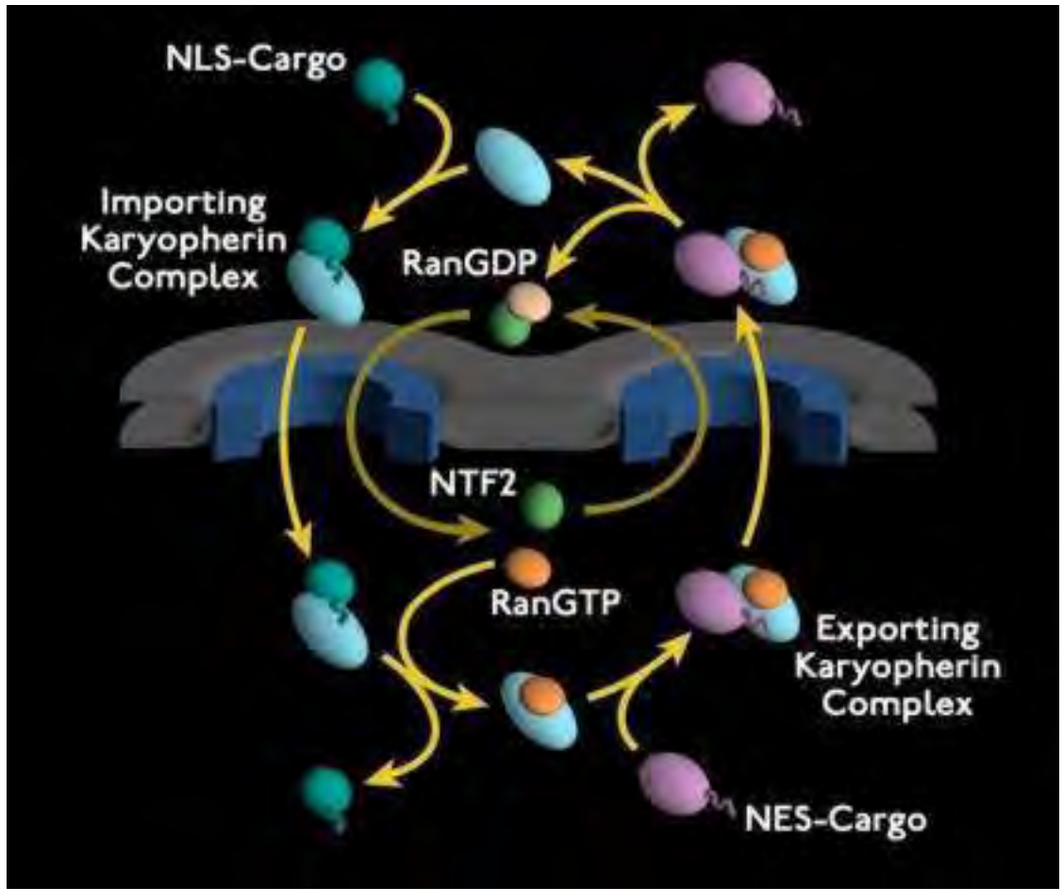
50 MDa complex: ~500 proteins of ~30 different types.

Mediates nucleocytoplasmic transport of proteins and RNA, *via* cognate transport factors (karyoferins or kaps).

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

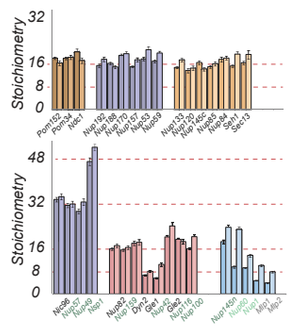
A large collaborative effort with Mike Rout and Brian Chait at Rockefeller University, also involving many other collaborators (Acknowledgments).

NCDIR *National Center for Dynamic Interactome Research* NIH TCNP



Integrative structure determination of the complete yeast Nuclear Pore Complex at ~9 Å precision

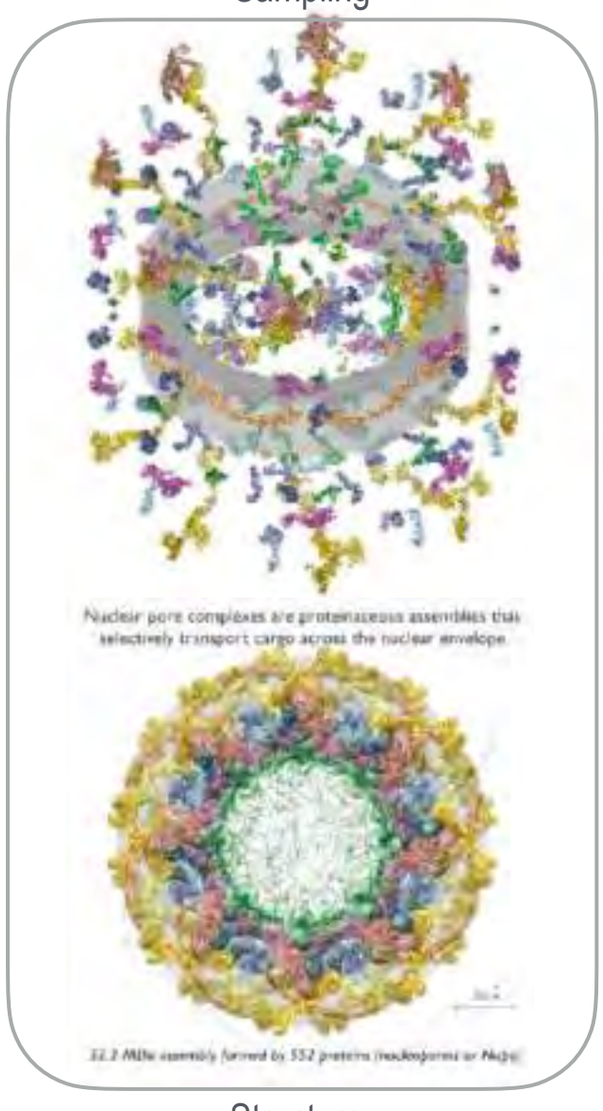
SJ Kim, J Fernandez-Martinez, I Nudelman, Y Shi, W Zhang, B Raveh, T Herricks, BD Slaughter, J Hogan, P Upla, IE Chemmama, R Pellarin, I Echeverria, M Shivaraju, AS Chaudhury, J Wang, R Williams, JR Unruh, CH Greenberg, EY Jacobs, Z Yu, MJ de la Cruz, R Mironska, DL Stokes, JD Aitchison, MF Jarrold, JL Gerton, SJ Ludtke, CW Akey, BT Chait, A Sali, MP Rout. *Nature*, 2018.



Quantitative mass spectrometry

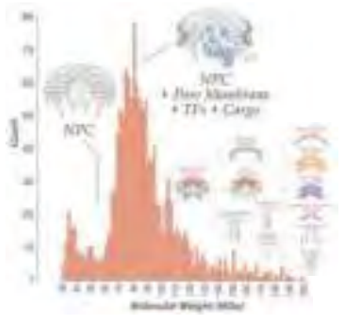
Stoichiometry
552 Nups, 32 types

Sampling

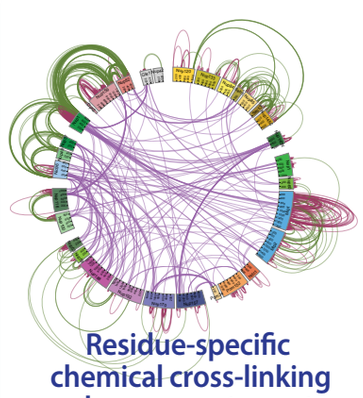


Structure

NPC mass
52.3 MDa



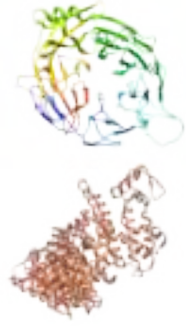
Label free and native mass spectrometry



Residue-specific chemical cross-linking and mass spectrometry

Distances
3,077 residue pairs

Atomic structures
50 domains
3 sub-complexes

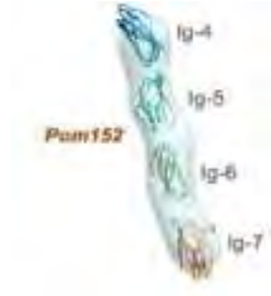


X-ray, NMR, comparative modeling, and integrative structure determination

Cryo-electron tomography

Density map
20~25 Å (inner-ring)
28 Å (overall)

Size and shape
147 segments



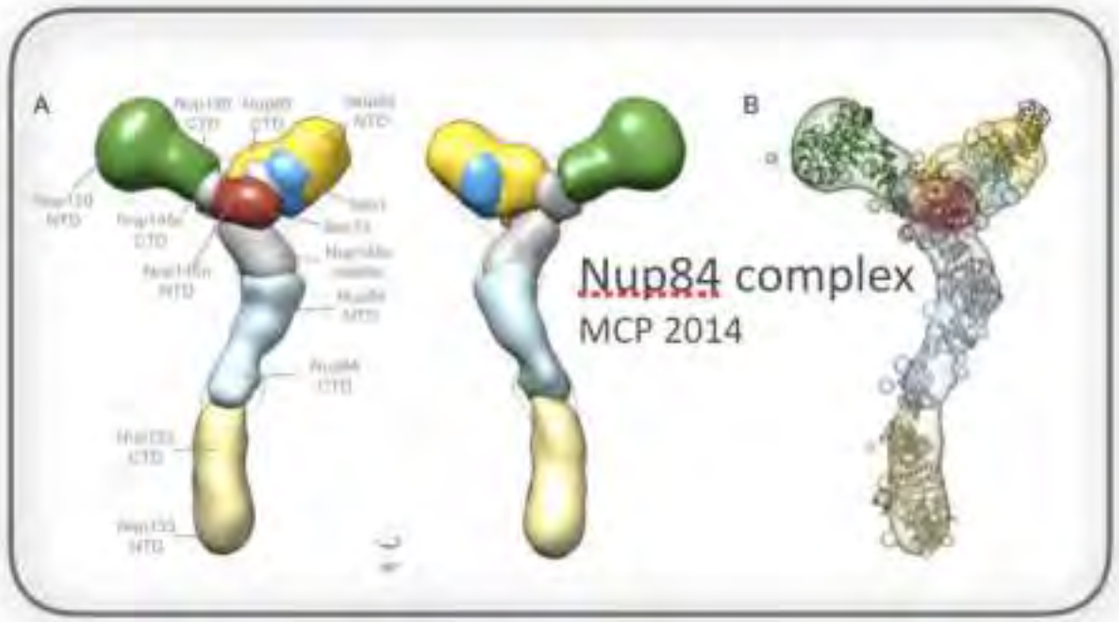
Small angle X-ray scattering

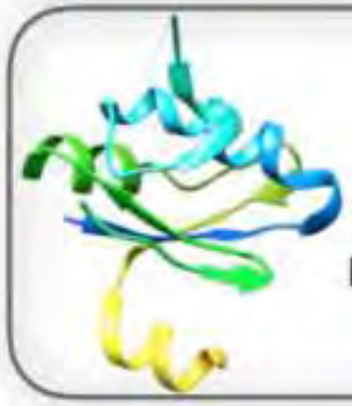
Representation of NPC components

Domains, proteins, sub-complexes: X-ray and NMR structures, integrative structures, comparative models.
Multi-scaling, rigid / flexible: to facilitate imposing spatial restraints and sampling efficiency.

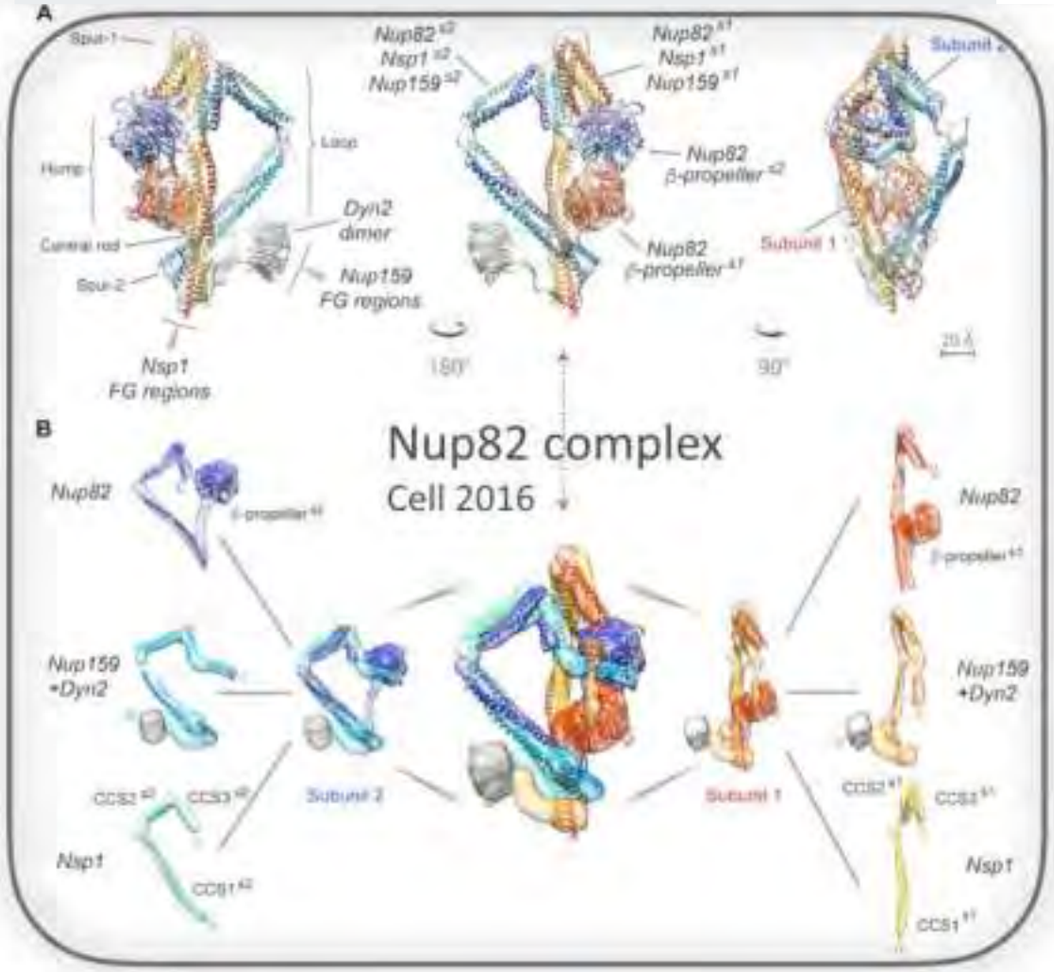
Nup192 / Nup188
Structure 2013

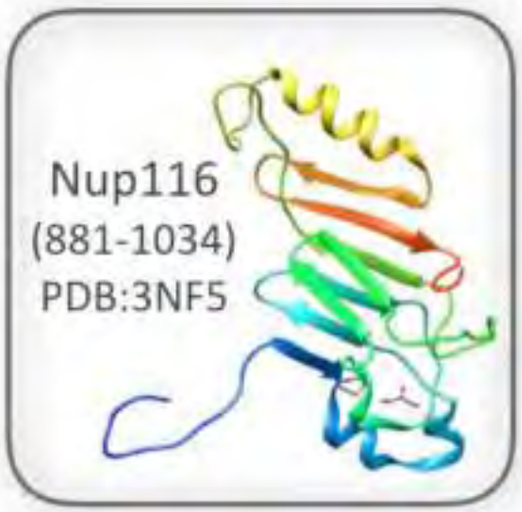
Nup2
(497-615)
PDB:3N7C



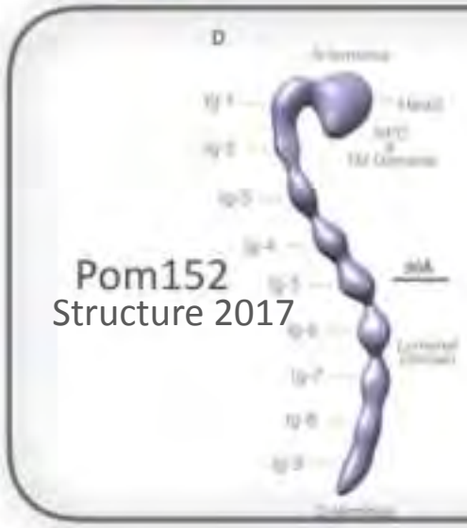


Nup53
(261-383)
PDB:3P3D





Nup116
(881-1034)
PDB:3NF5

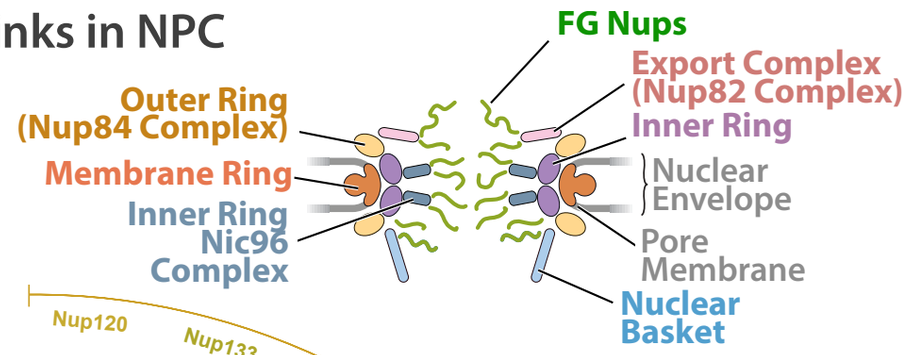


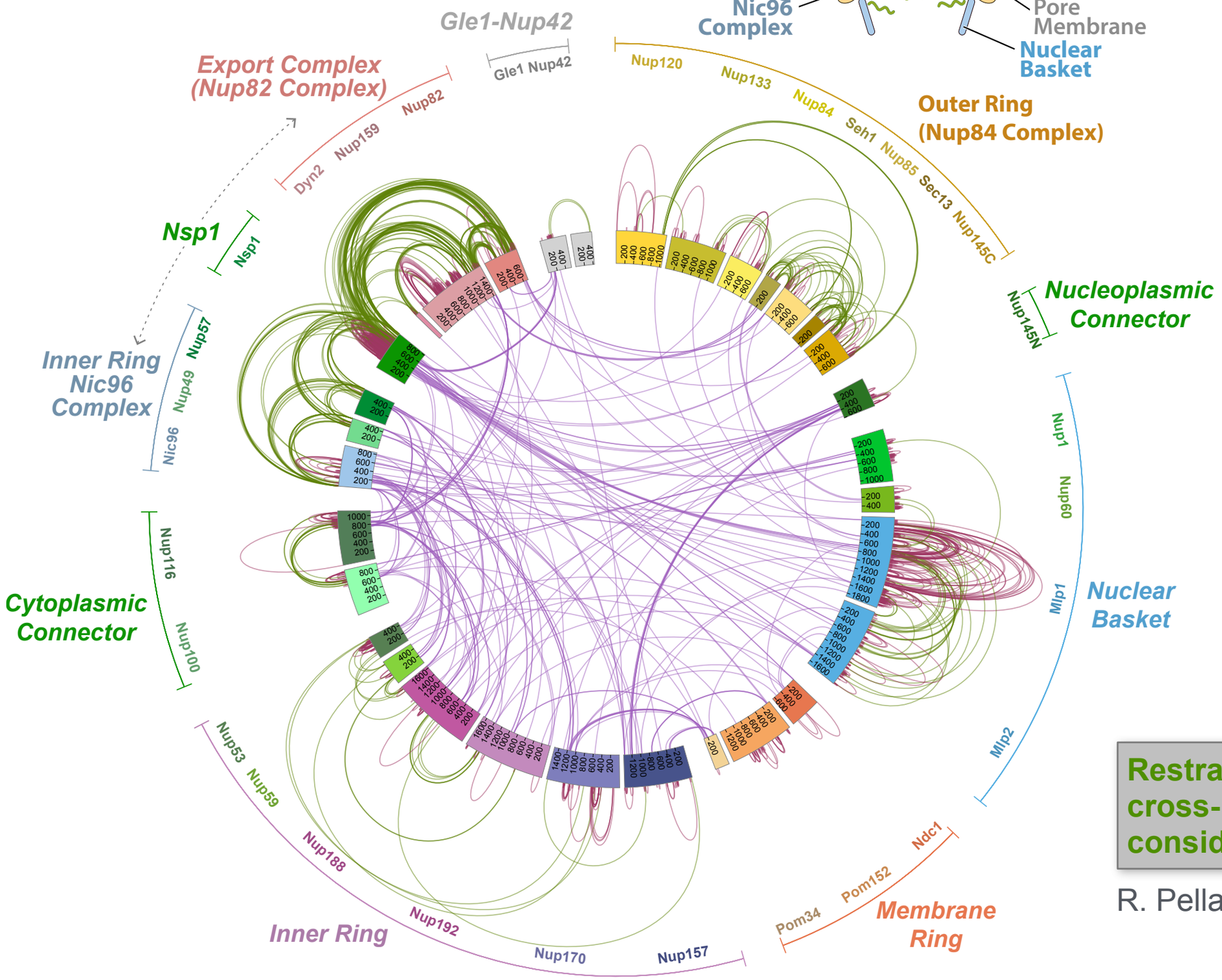
Pom152
Structure 2017

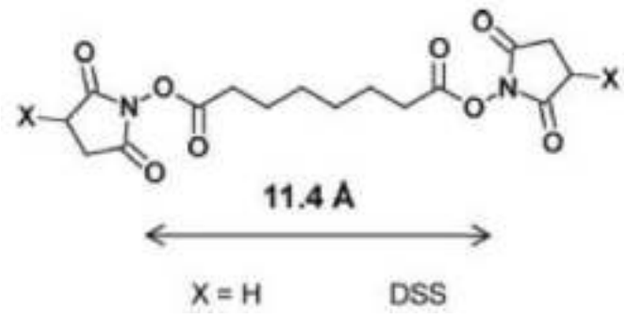
Chemical cross-links by mass spectrometry

Circos Plot of the 3,077 Chemical Cross-links in NPC

Inter-molecular (between modules)
Inter-molecular (within module)
Intra-molecular







Followed by limited proteolysis and MS detection of cross-linked Lys residue pairs.

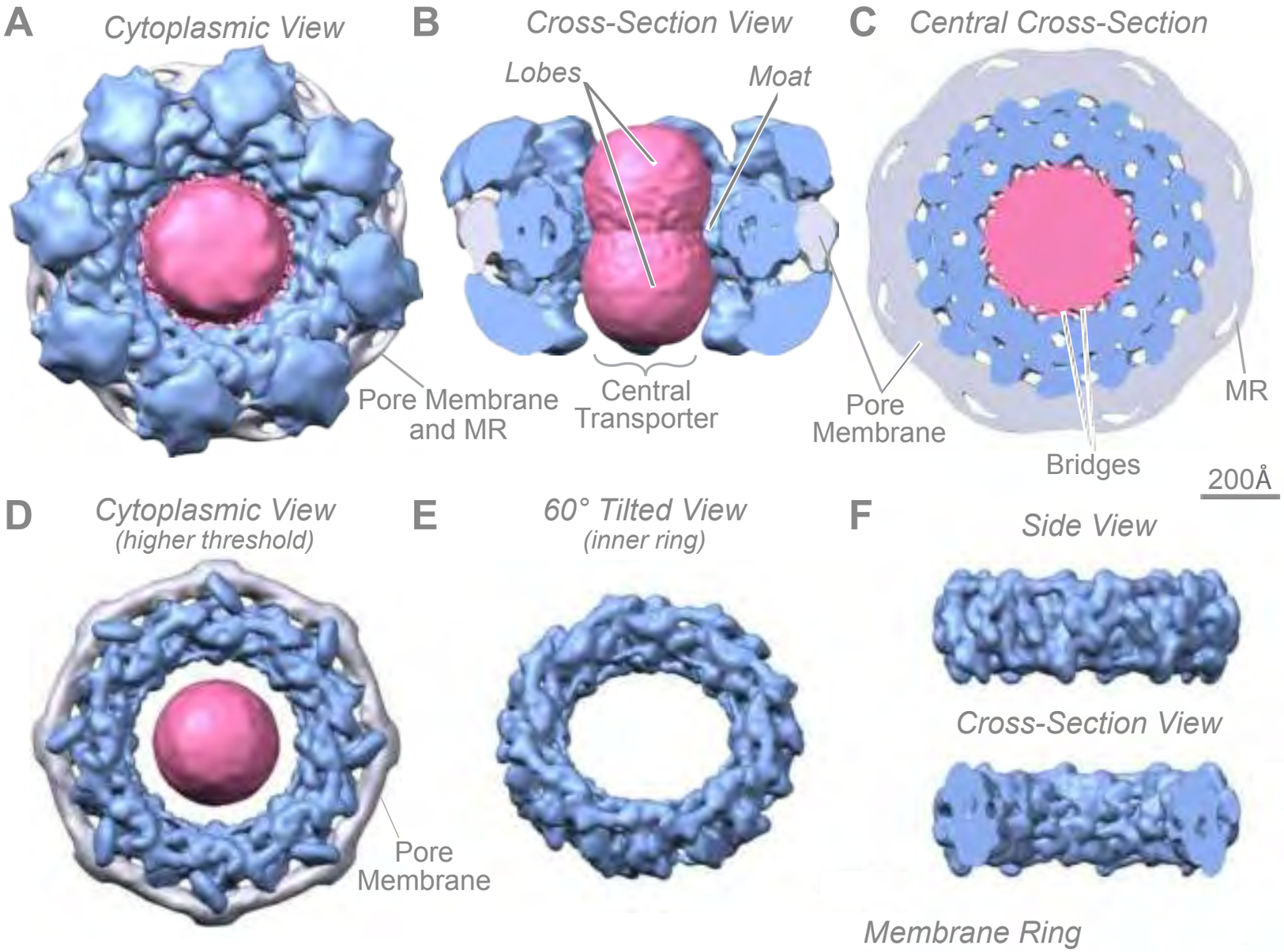
Restraints: upper distance bounds on cross-linked atoms or beads; Bayesian, considering assignment ambiguity.

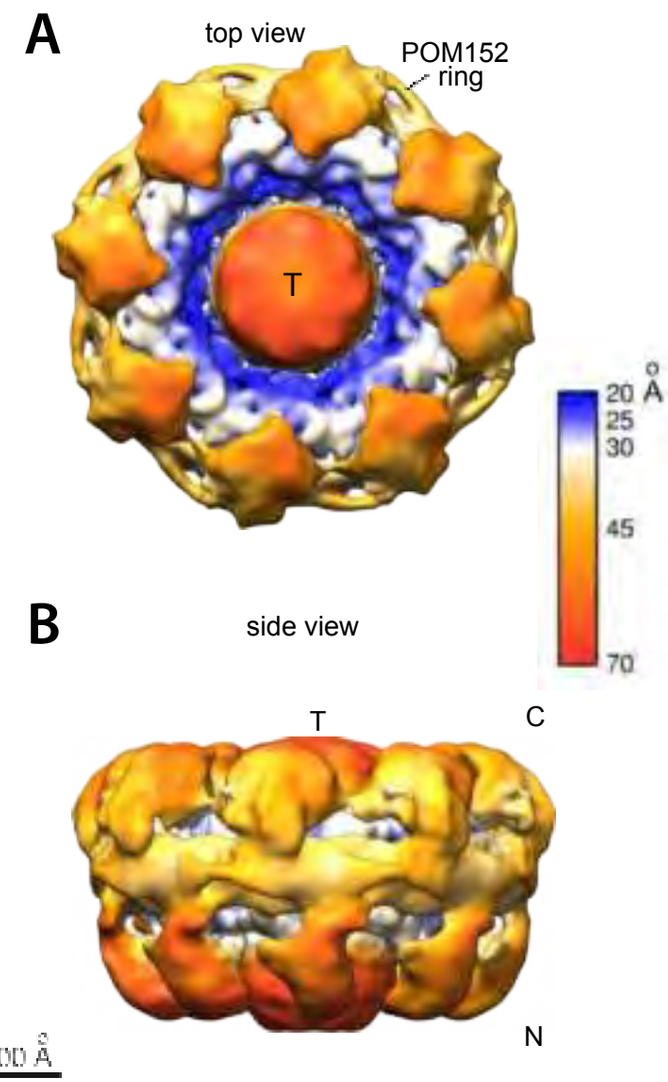
R. Pellarin, M. Bonomi

10.5281/zenodo.1149746

J. Fernandez, Y. Shi, I. Nudelman, W Zhang, R Williams, J Wang, M. Rout, B. Chait

Density map by cryo-electron tomography with sub-tomogram and C_8-symmetry averaging





Restraint: Cross-correlation between a model and the map, each represented by a Gaussian Mixture Model (~2000 Gaussians each).

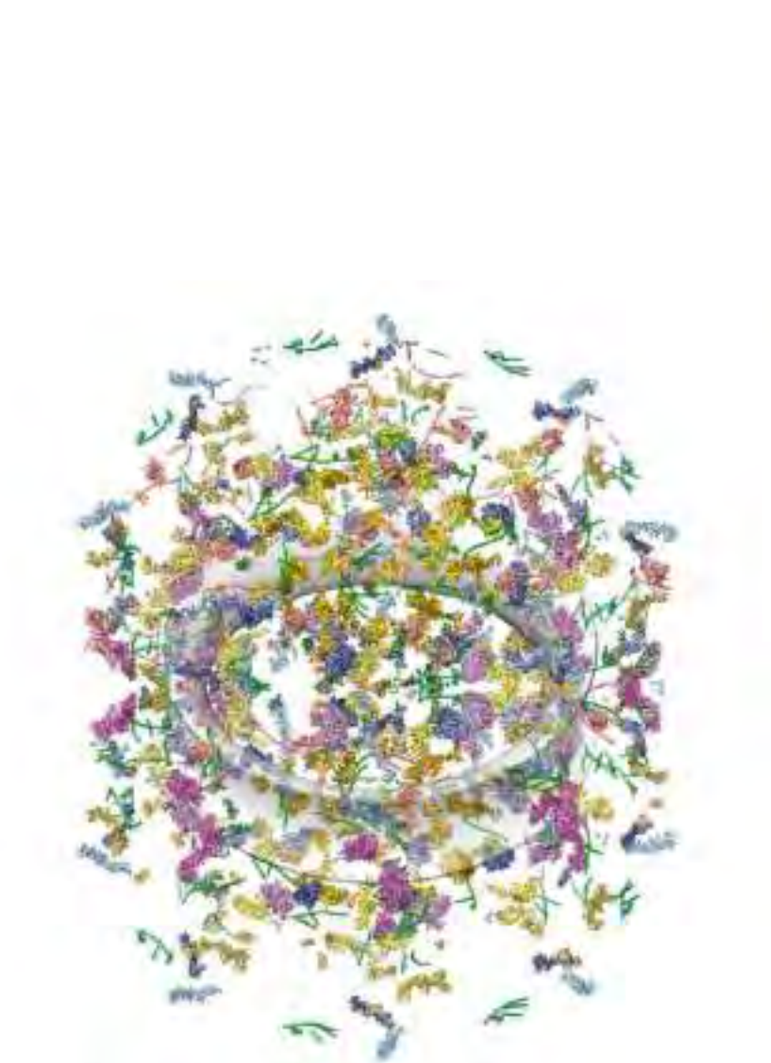
R. Pellarin, M. Bonomi

Resolution of ~28 Å
(~20Å in inner ring)

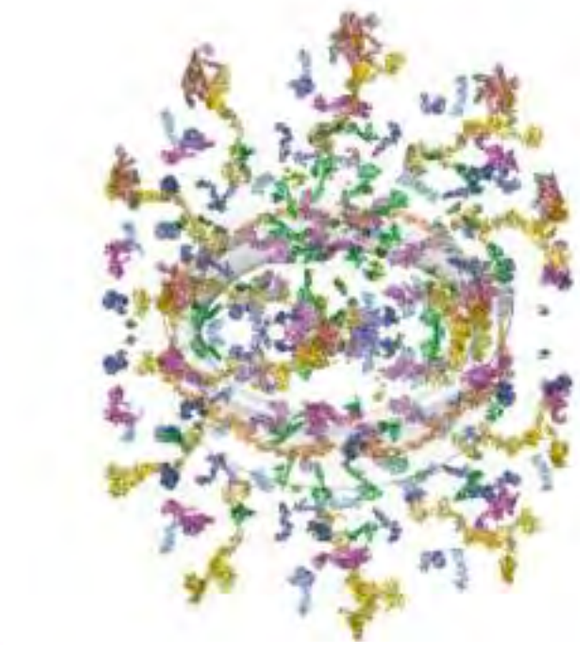
EMPIAR-10155, EMD-7321

I. Nudelman, S. Ludtke, C. Akey

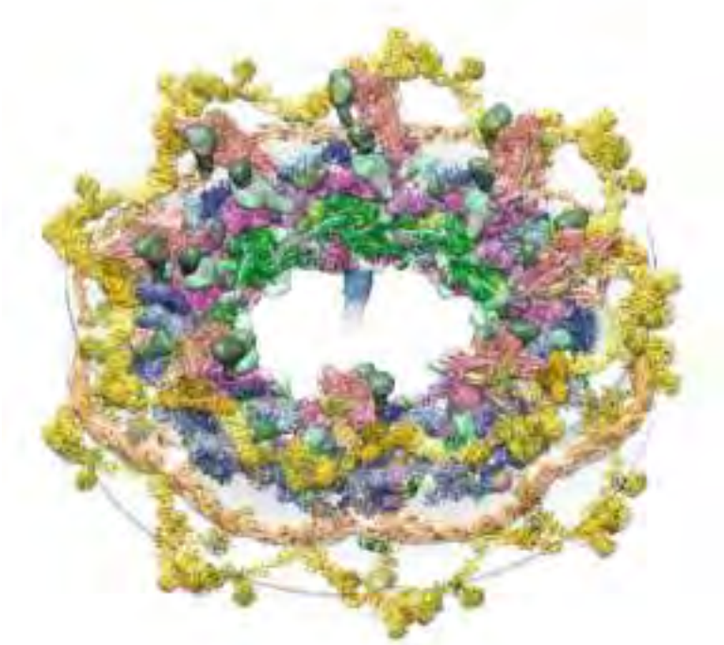
Structural sampling



Random initial structures



Optimization

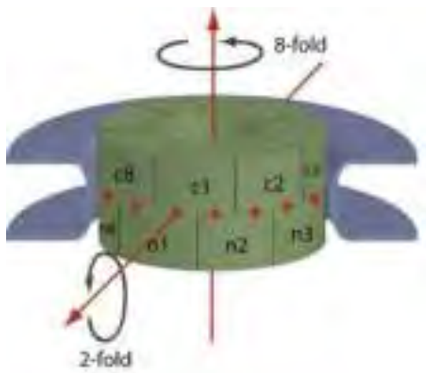


Model precision of ~9 Å (integration!)

Ensemble of optimized structures

Model precision > Sampling precision

Constrain sampling to a unit cell, while fully considering inter-unit relationships:



Gibbs Sampling Monte Carlo enhanced by Replica Exchange (Sugita & Okamoto, 1999).

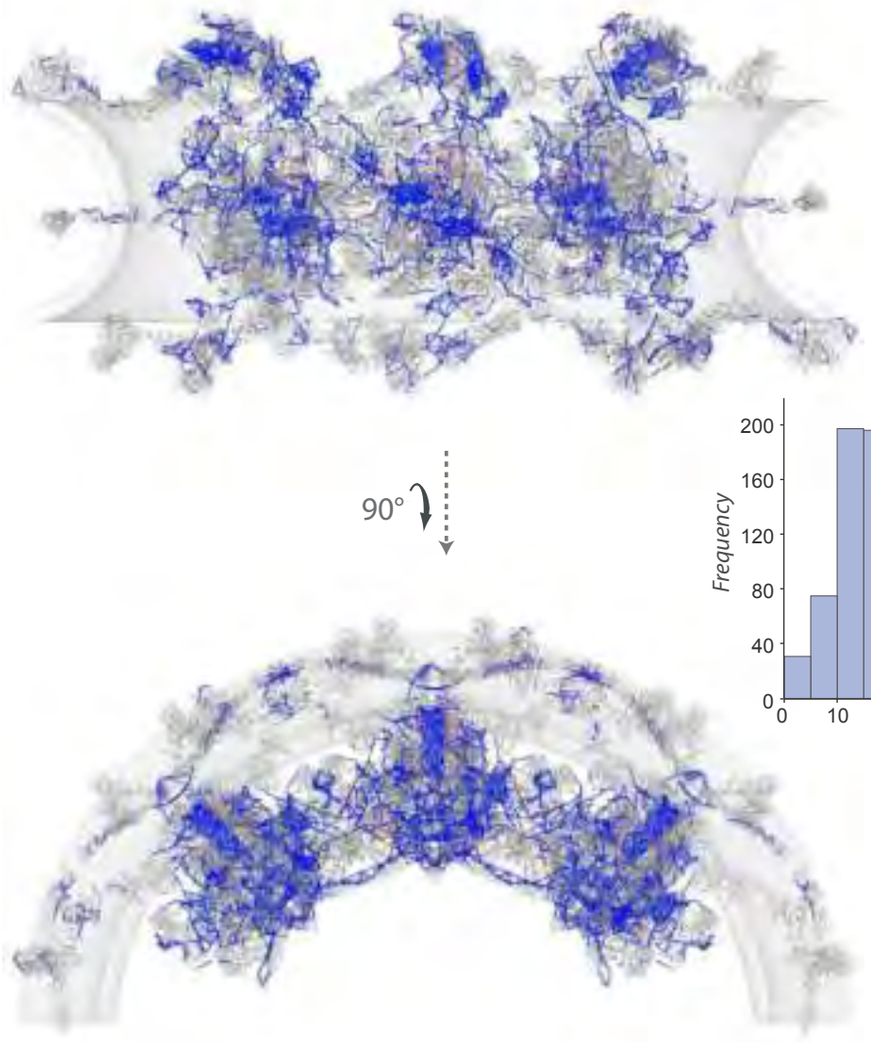
Sampled variables in a unit cell (~10^7 MC moves):

- Rigid body positions and orientations (small random translations & rotations)
- Bead positions (small random translations)

Validation: Model fits data used for modeling

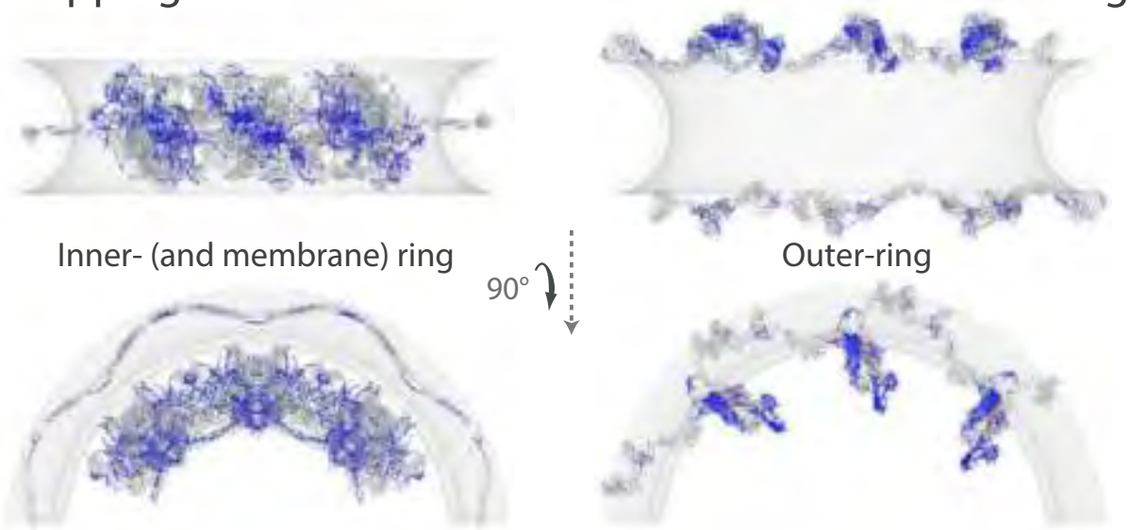
Structure satisfies most chemical cross-links.

E Satisfaction of the Cross-links on the entire NPC Structure



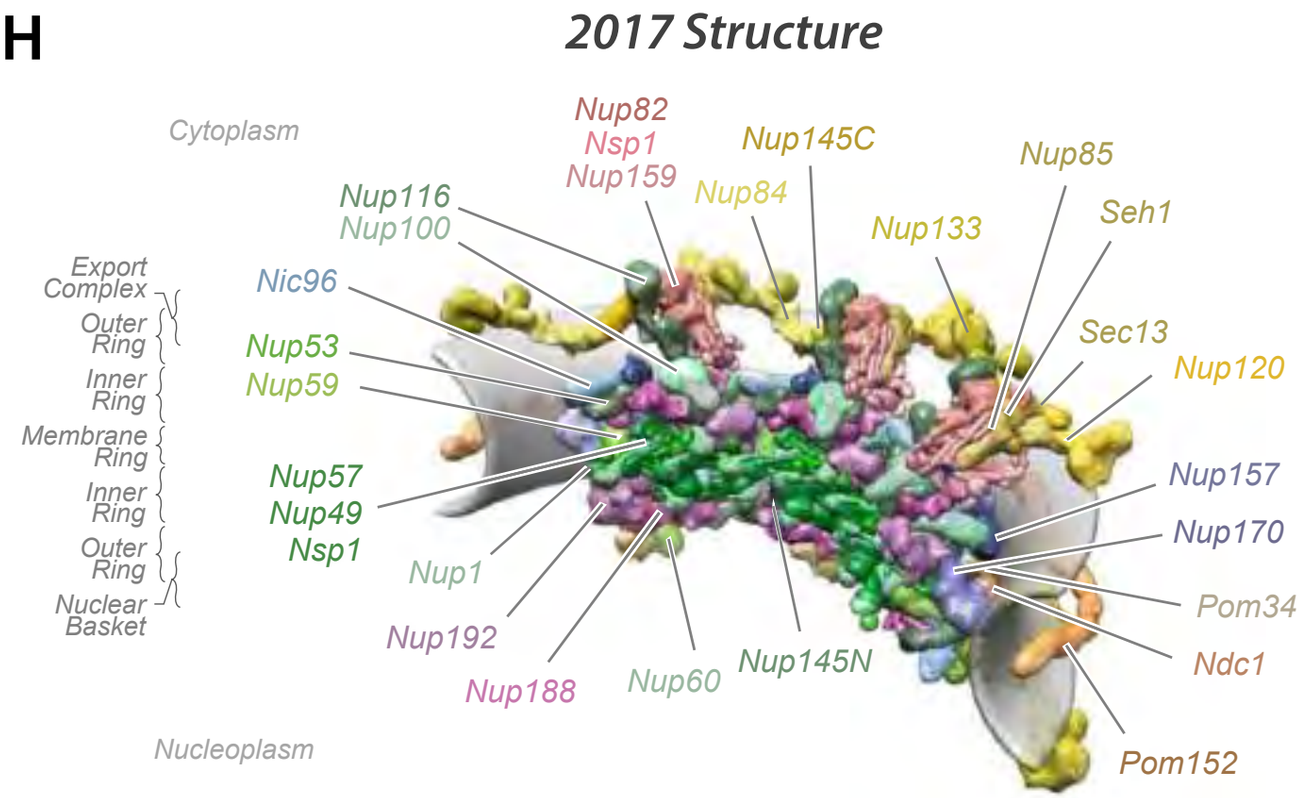
F Mapping of the Cross-links onto the Connector Nups

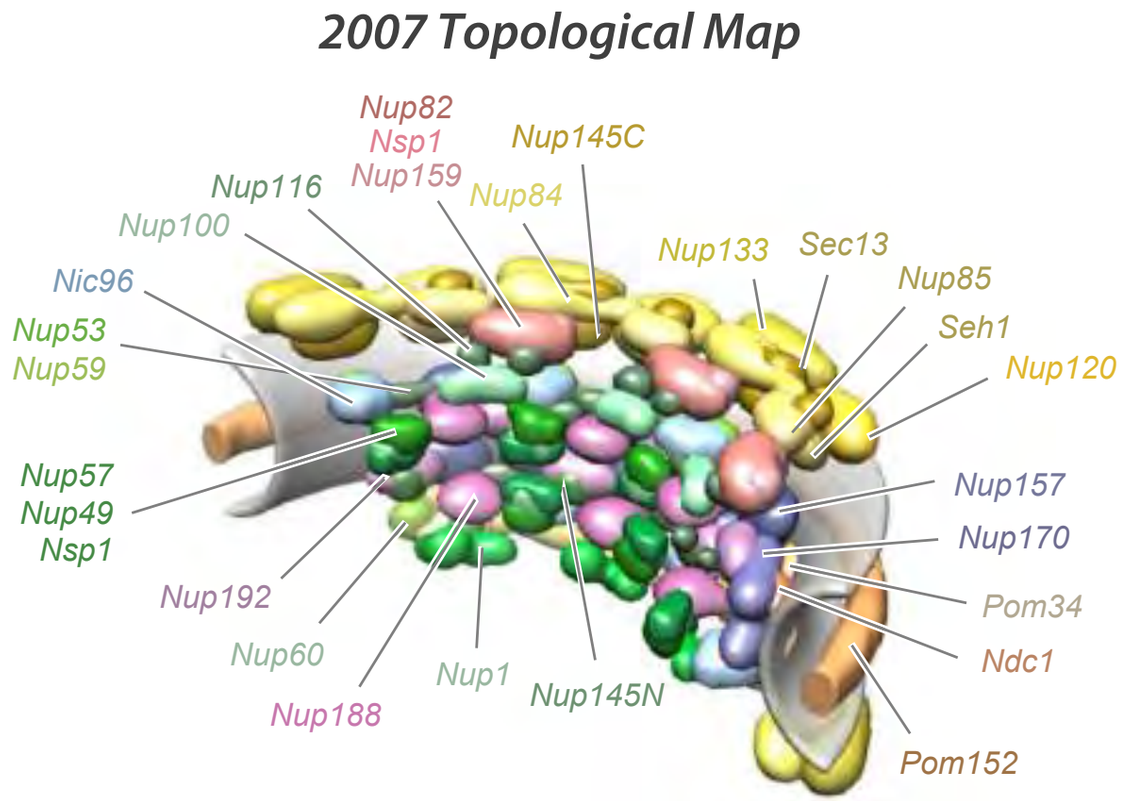
G Mapping of the Cross-links onto the inner- and outer-rings



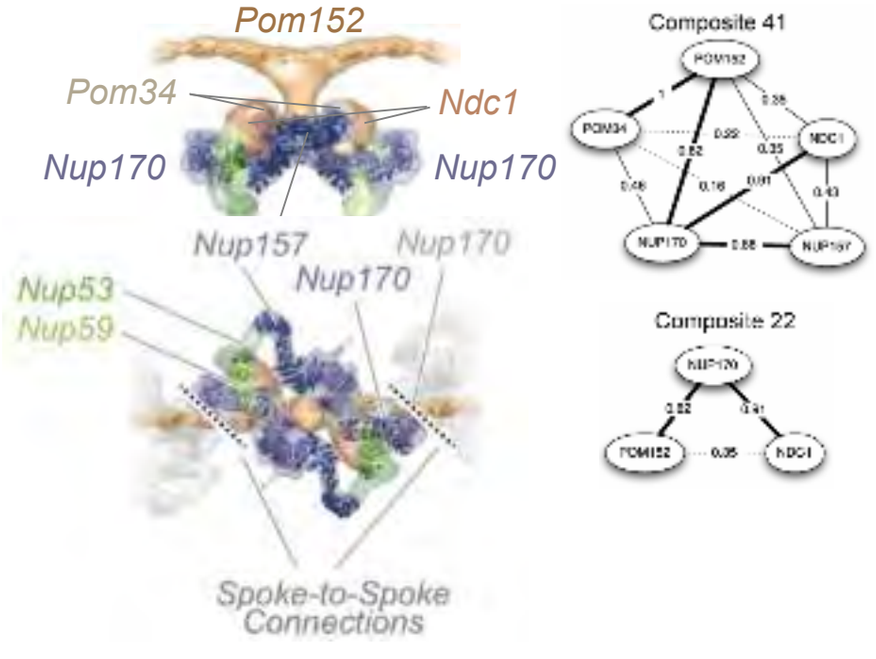
Validation: Model fits data not used for modeling

H

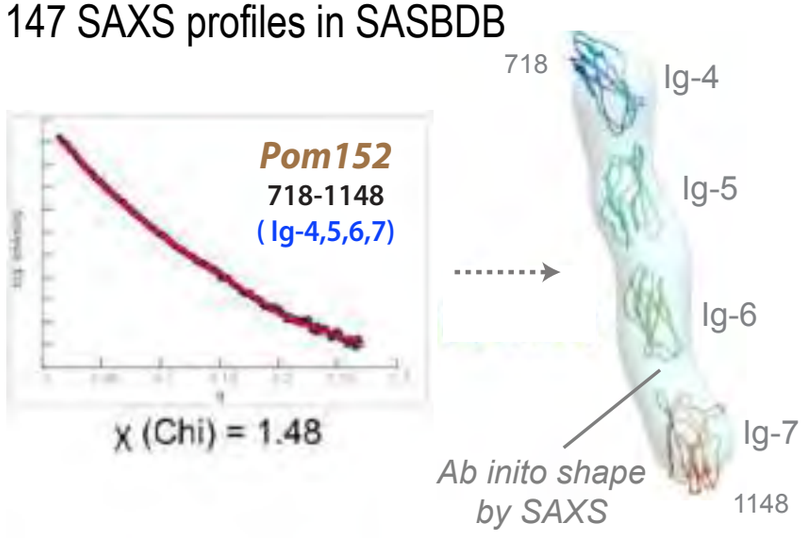




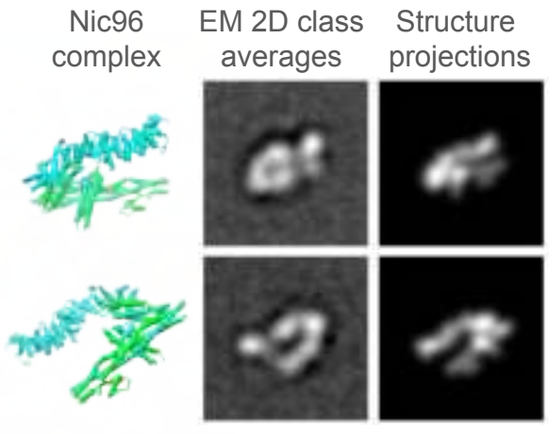
I Satisfaction of affinity purification and overlay assays data (composites)



J Satisfaction of SAXS data



K Satisfaction of EM 2D



Architecture of the yeast NPC

FG Repeats:
Central Transporter

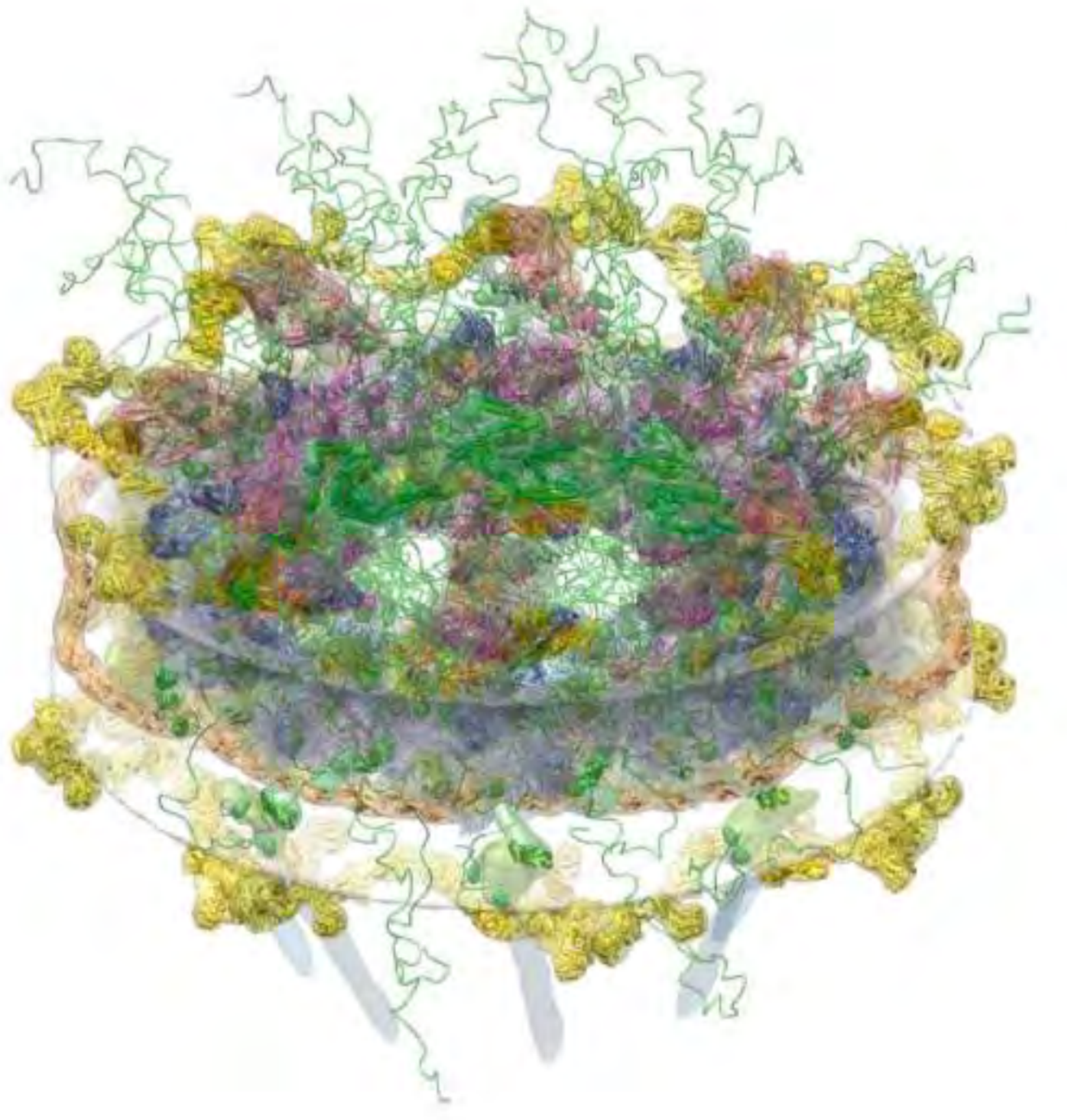
FG Repeat Anchors

Outer Rings

Inner Rings

Membrane Rings

Cytoplasmic Export Platform
&
Nuclear Basket

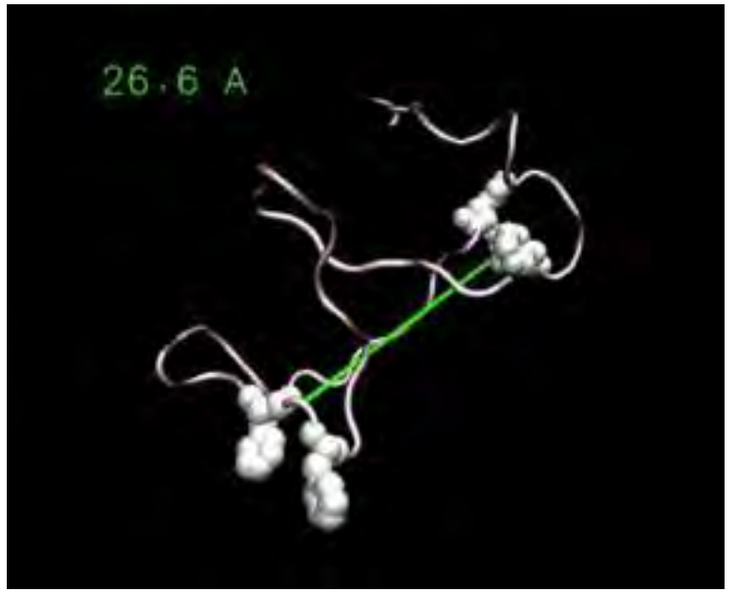


~9 Å Resolution

Transport through the NPC

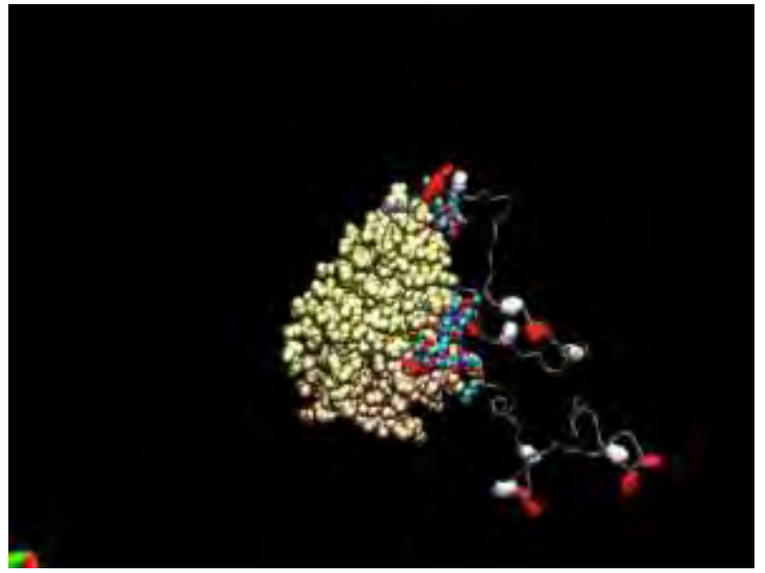
w/ M. Rout, D. Cowburn

All-atom MD simulations of multiple FG repeats



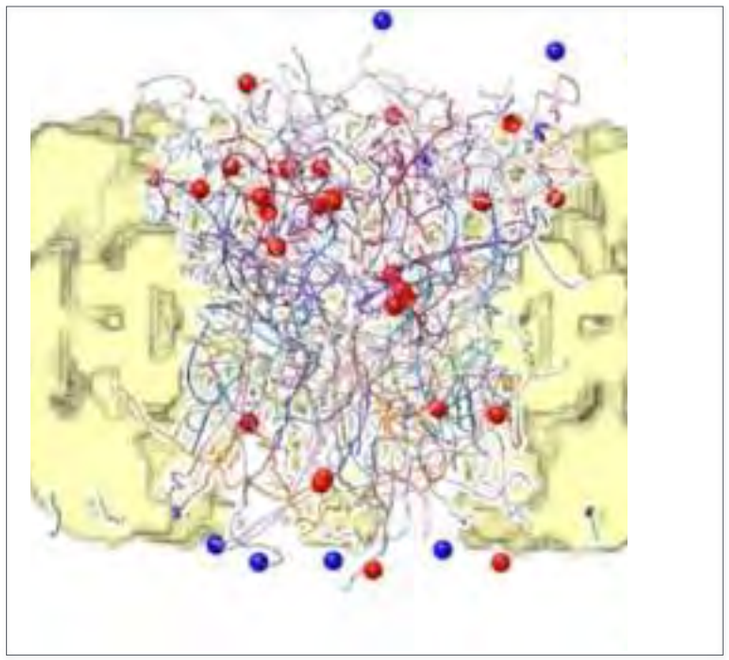
Raveh *et al, PNAS* 2016

All-atom simulation of $FSFG_6$:NTF2 interaction on the Anton supercomputer



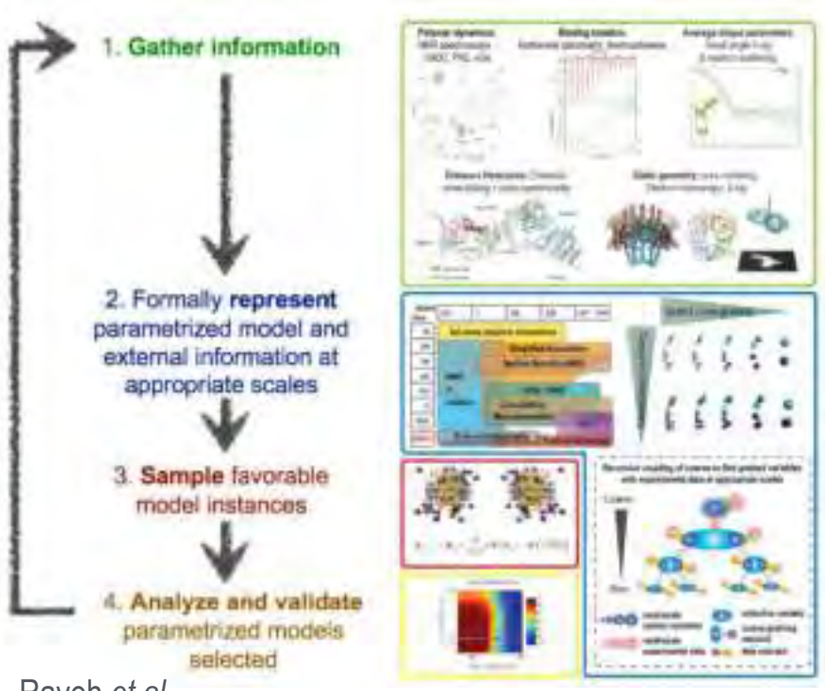
Raveh *et al, PNAS* 2016

Coarse-grained BD simulations of NPC transport



Timney, Raveh *et al, JCB 2016*

Integrative modeling of transport through the NPC



Raveh *et al*

Example method development

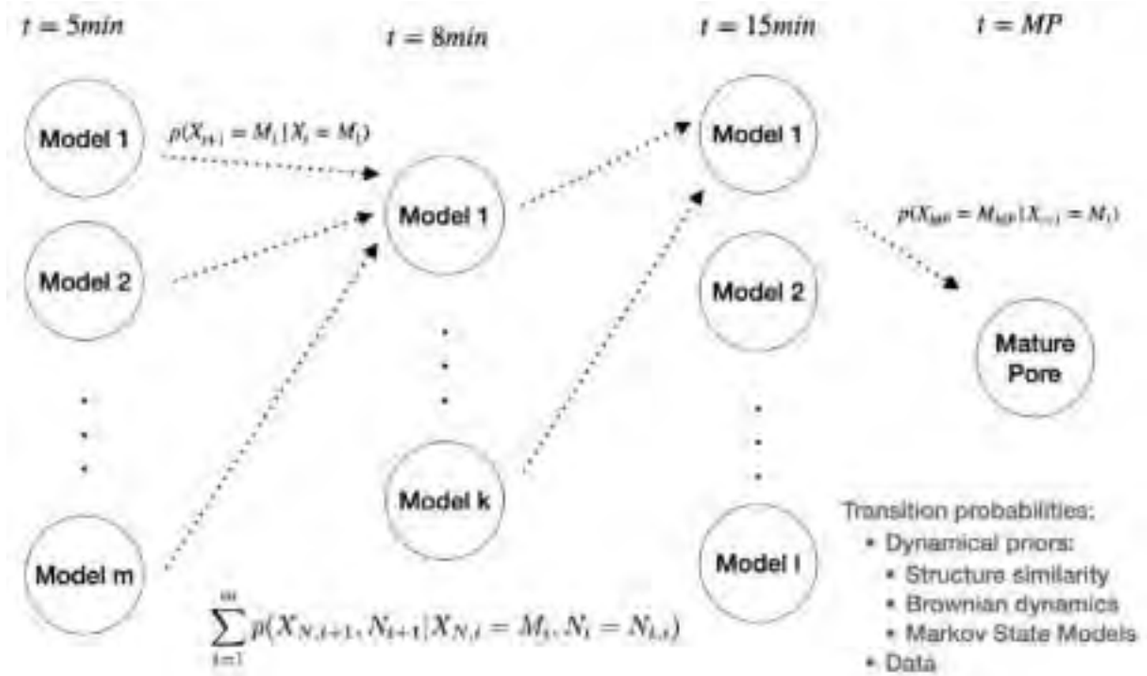
Integrative spatiotemporal modeling of dynamic processes: Assembly of the NPC

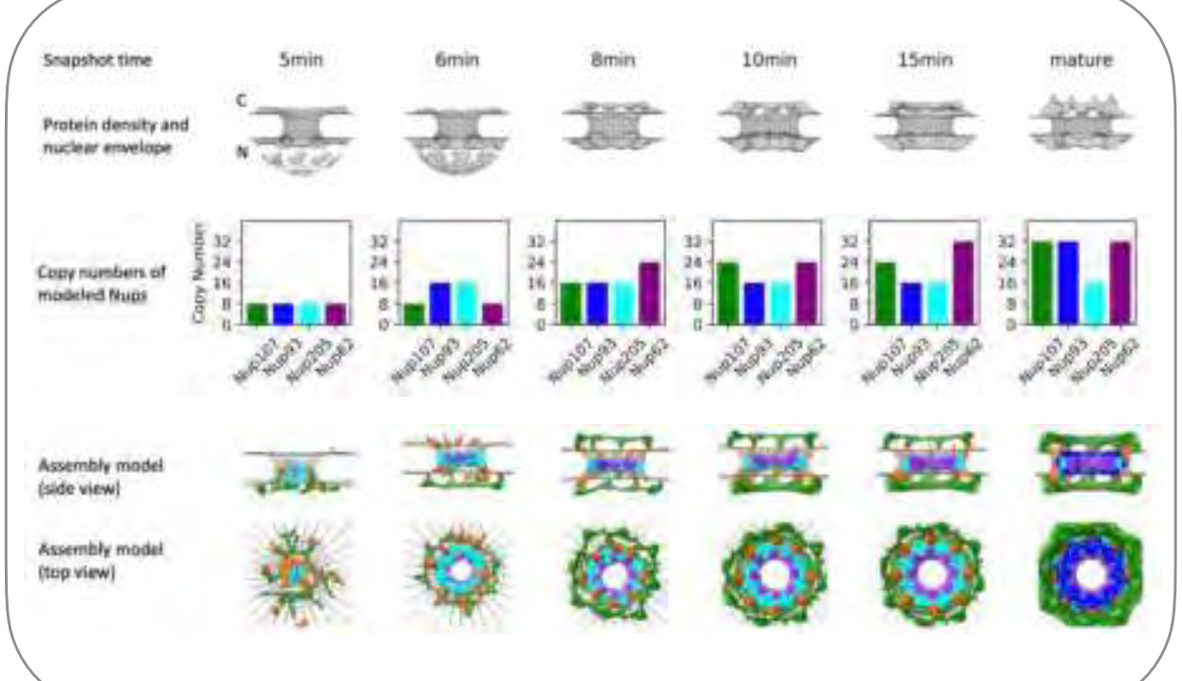
1) Model several snapshots using integrative modeling of static structures

2) Connect snapshots into trajectories

Data Source J. Ellenberg *et al*	Contribution to scoring function J. Tempkin
Atomic resolution model of fully assembled NPC structure	Harmonic restraints on native protein-protein contacts
Live-cell correlated electron tomography derived protein densities	Restraint on global overlap between proposed structural model and observed densities
Live-cell correlated electron tomography derived NE shape	Spatial constraints on the size and shape of the nuclear envelope environment
Quantitative fluorescence data	Restraints on Nup copy numbers over time

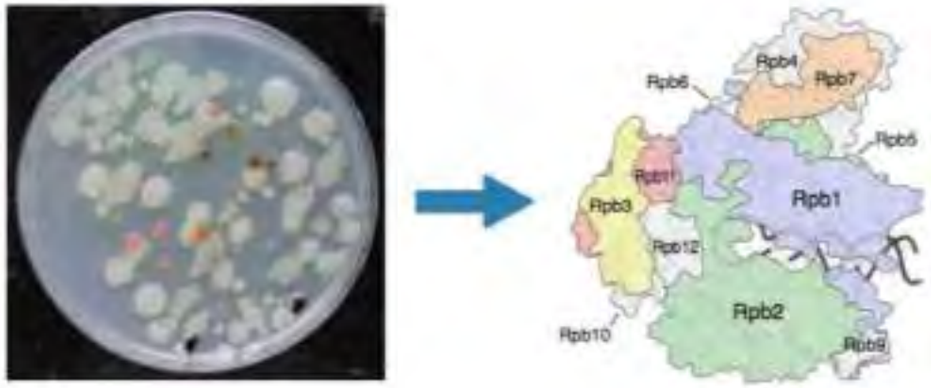
All data-derived contributions are multiplied to yield a time-dependent Bayesian model of $P(X_{N,t}, N_t)$.





Jeremy Tempkin

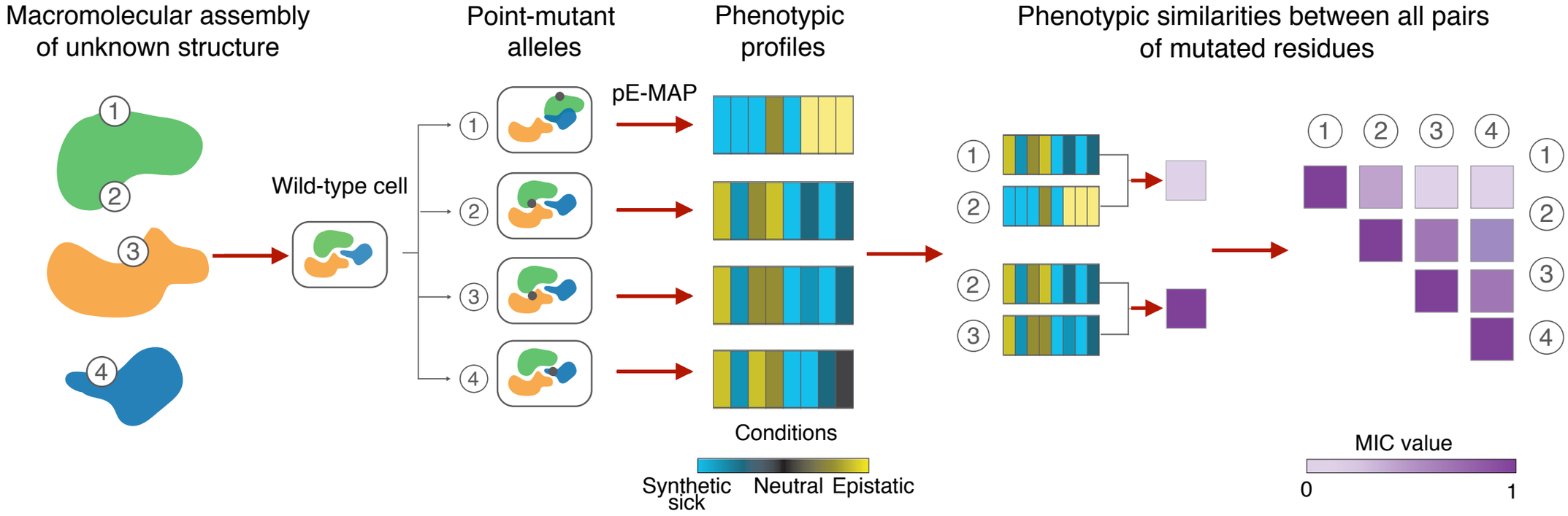
w/ Jan Ellenberg, Shotaro Otsuka *et al*, EMBL

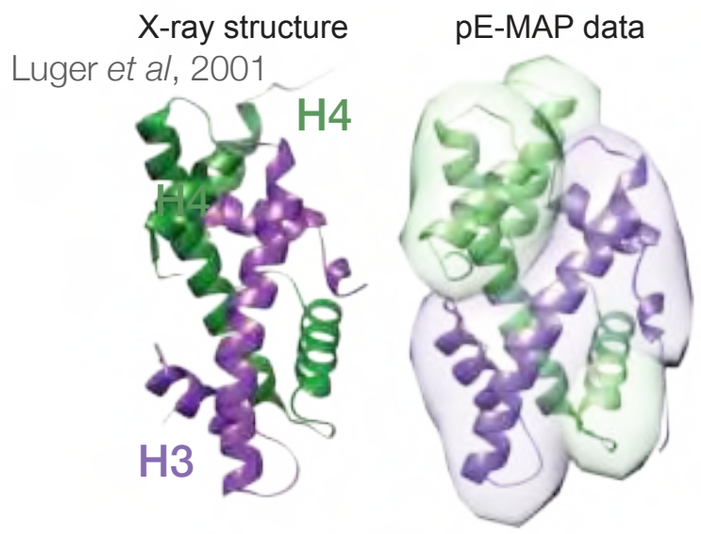


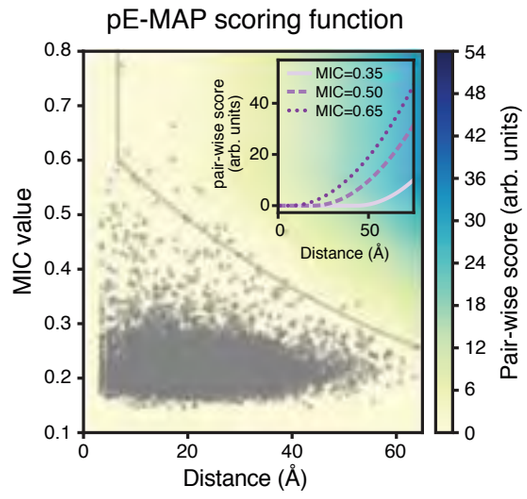
Genetic interaction mapping informs integrative structure determination of protein complexes

Hannes Braberg[1,2,*], Ignacia Echeverria[1,2,3,*], Stefan Bohn[1,2,4,*,†,‡], Peter Cimermancic[5,*,§], Anthony Shiver[5,¶], Richard Alexander[1], Jiewei Xu[1,2,4], Michael Shales[1,2], Raghuvar Dronamraju[6], Shuangying Jiang[7], Gajendradhar Dwivedi[6,#], Derek Bogdanoff[5], Kaitlin K. Chaung[8], Ruth Hüttenhain[1,2,4], Shuyi Wang[1], David Mavor[2,3], Riccardo Pellarin[3,**], Dina Schneidman[1], Joel S. Bader[10], James S. Fraser[2,3], John Morris[11], James E. Haber[8], Brian D. Strahl[6], Carol A. Gross[12], Junbiao Dai[7], Jef D. Boeke[13,14,15,16,††], Andrej Sali[2,3,11,††], Nevan J. Krogan[1,2,4,17,††]

Science 11 Dec 2020:
Vol. 370, Issue 6522, eaaz4910







Community

Integrative Methods Task Force Workshop

Facilitate computing, validating, archiving, and disseminating integrative structures.

WORLDWIDE PDB
PROTEIN DATA BANK

Andrej Sali, Helen M. Berman, Torsten Schwede, Jill Trewhella, Gerard Kleywegt, Stephen K. Burley, John Markley, Haruki Nakamura, Paul Adams, Alexandre Bonvin, Wah Chiu, Tom Ferrin, Kay Grünewald, Aleksandras Gutmanas, Richard Henderson, Gerhard Hummer, Kenji Iwasaki, Graham Johnson, Cathy Lawson, Frank di Maio, Jens Meiler, Marc Marti-Renom, Guy Montelione, Michael Nilges, Ruth Nussinov, Ardan Patwardhan, Matteo dal Peraro, Juri Rappsilber, Randy Read, Helen Saibil, Gunnar Schröder, Charles Schwieters, Claus Seidel, Dmitri Svergun, Maya Topf, Eldon Ulrich, Sameer Velankar, and John D. Westbrook. *Structure* **23**, 1156-1167, 2015.

First Integrative Methods Task Force Workshop was held at the European Bioinformatics Institute in Hinxton, UK, on October 6 and 7, 2014 (organized by G. Kleywegt *et al*):

What should be archived?
How should integrative models be represented?
How should the data and integrative models be validated?
How should the data and models be archived?
What information should accompany the publication of integrative models?

Data working group

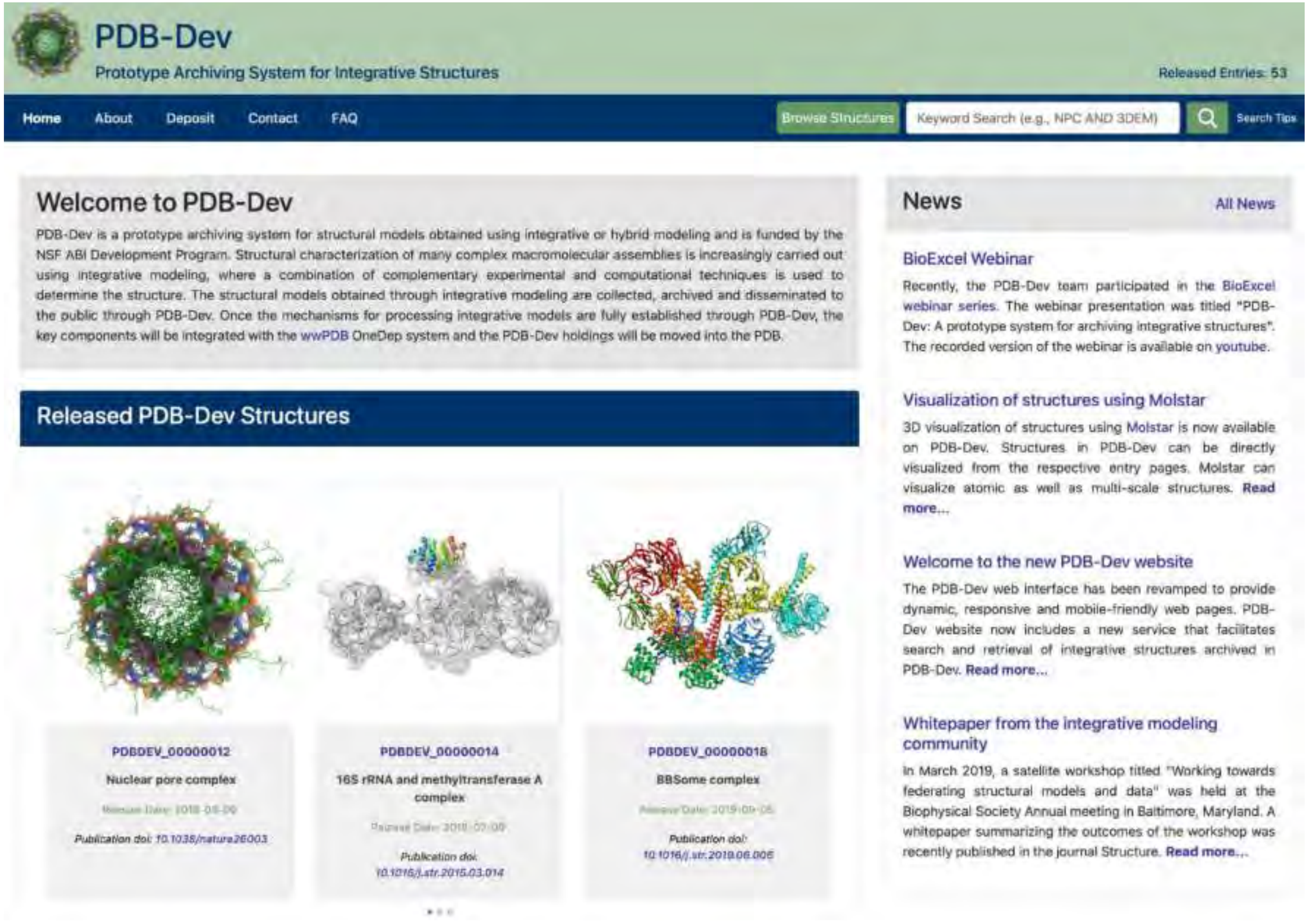
H. Berman, J. Trewhella

Model working group

A. Bonvin, F. Dimaio, G. Hummer, J. Meiler, E. Tajkhorshid, T. Schwede, A. Sali

Nascent wwPDB archive for integrative structures

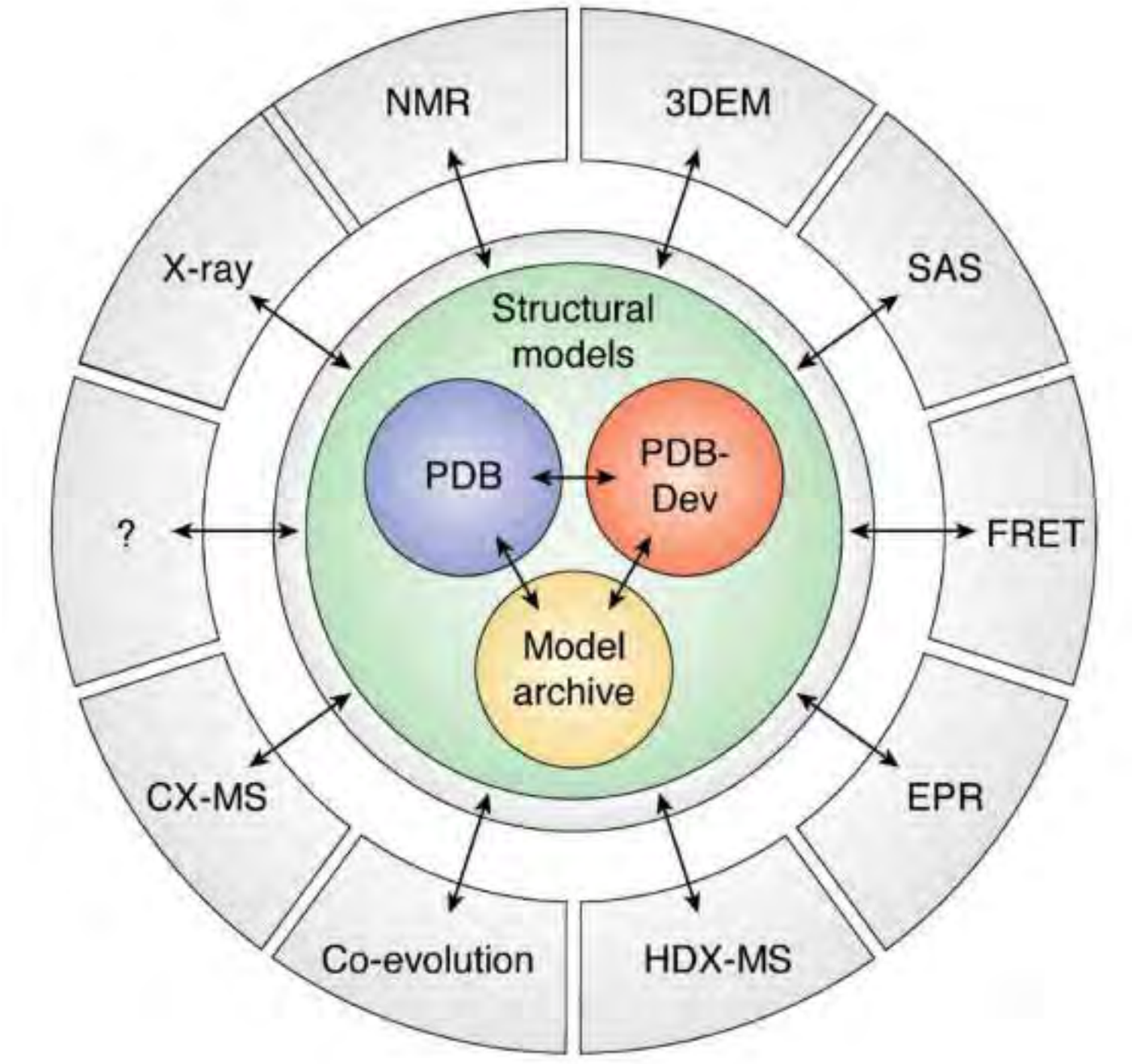
https://pdb-dev.wwpdb.org



SK Burley, G Kurisu, JL Markley, H Nakamura, S Velankar, HM Berman, A Sali, T Schwede, J Trewhella.
PDB-Dev: A Prototype System for Depositing Integrative/Hybrid Structural Models, *Structure* 25, 1317-1318, 2017.

B. Vallat, B. Webb, J. Westbrook, A. Sali, H. Berman, *Structure,* 2018.

Federated wwPDB archive for integrative structures and data



wwPDB, Helen Berman

Cell mapping

The Pancreatic β-Cell Consortium:

Spatiotemporal multi-scale whole cell mapping to better understand β-cell biology

"structure without function is a corpse … function without structure is a ghost", SA Wainwright

USC

Berkeley
UNIVERSITY OF CALIFORNIA

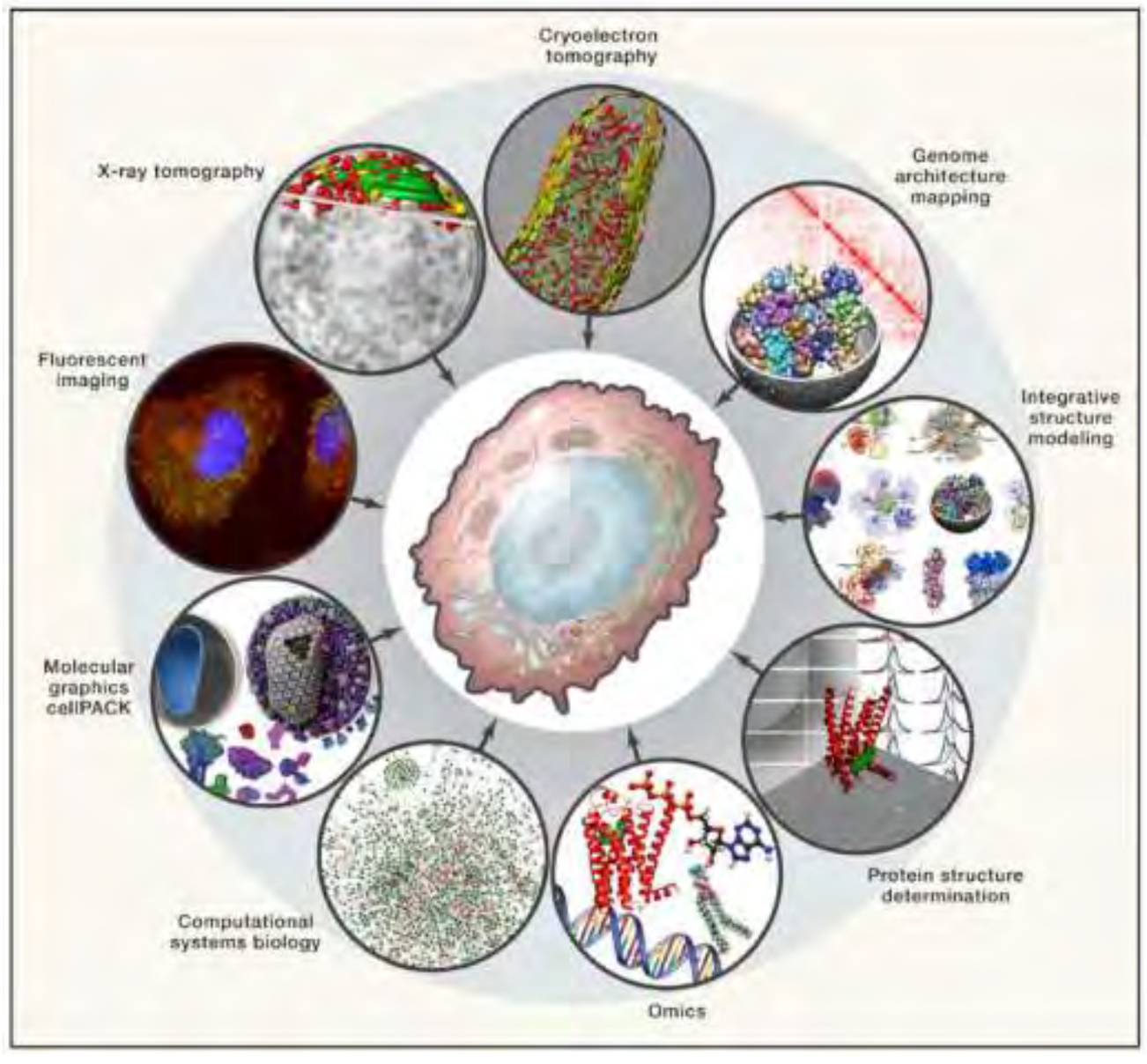
UCSF
University of California
San Francisco

The Scripps Research Institute

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Université de Montréal

iHuman INSTITUTE



- Multi-disciplinary team:
  - Biologists
  - Chemists
  - Engineers
  - Physicians
  - Computational scientists
  - Artists
- Global, open-source community
- Provide the basis for next generation drug and cell-therapy design

www.pbcconsortium.org

others welcome

Singla *et al*. *Cell* **173**, 11-19, 2018

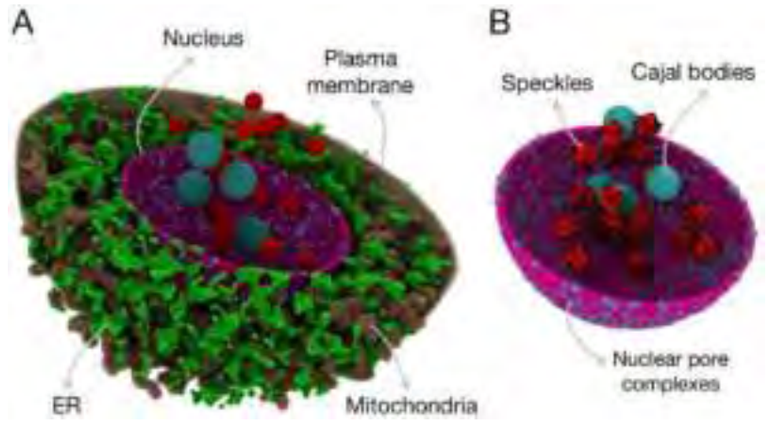
Some approaches to whole cell modeling

A Whole-Cell Computational Model Predicts Phenotype from Genotype

Jonathan R. Karr,[1,4] Jayodita C. Sanghvi,[2,4] Derek N. Macklin,[2] Miriam V. Gutschow,[2] Jared M. Jacobs,[2] Benjamin Bolival, Jr.,[2] Nacyra Assad-Garcia,[3] John I. Glass,[3] and Markus W. Covert[2,*]

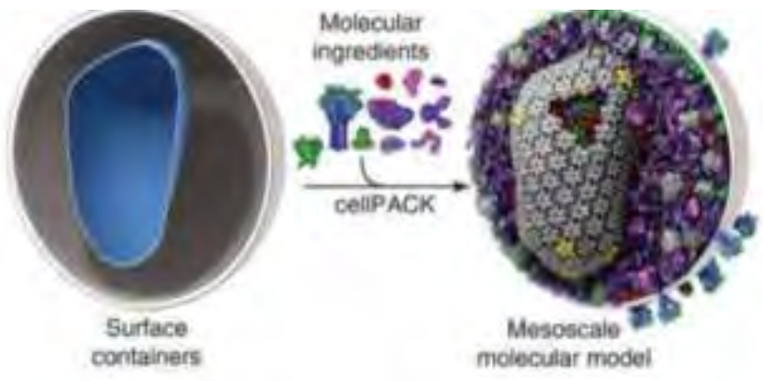
An *in-silico* human cell model reveals the influence of spatial organization on RNA splicing

Zhaleh Ghaemi[1,2*], Joseph R. Peterson[1], Martin Gruebele[1,2,3], Zaida Luthey-Schulten[1,2,3]



cellPACK: a virtual mesoscope to model and visualize structural systems biology

Graham T Johnson[1–3], Ludovic Autin[1], Mostafa Al-Alusi[1], David S Goodsell[1], Michel F Sanner[1] & Arthur J Olson[1]



...

A model of the cell must be useful and feasible

- Spatiotemporal
- Multiscale (multiple representations)
- Harmonized across representations
- Integrative (based on all information)
- Reflect uncertainty
- Accurate, precise, complete, general
- Rationalize data
- Predict data
- Guide experiments
- Allow modulating cell
- Iteratively improvable

Singla, McClary, White, Alber, Sali, Stevens, *Cell **173***, 11-19, 2018

Idea for modeling the whole cell: Bayesian metamodeling

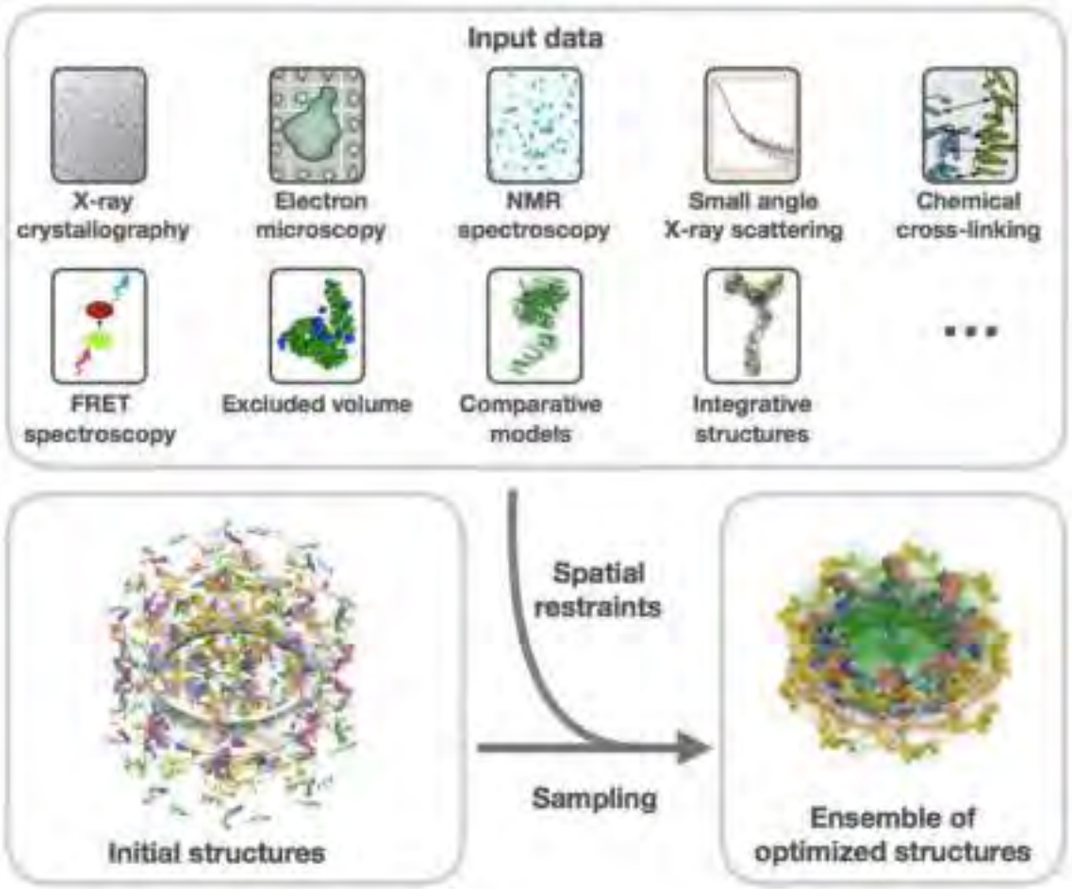
Divide-and-conquer modeling:

1. Construct all useful and feasible types of models for any aspects of any part of the cell, using existing modeling methods, including significantly various types of integrative modeling based on varied types of experimental data, physical theories, statistical inferences, and prior models.

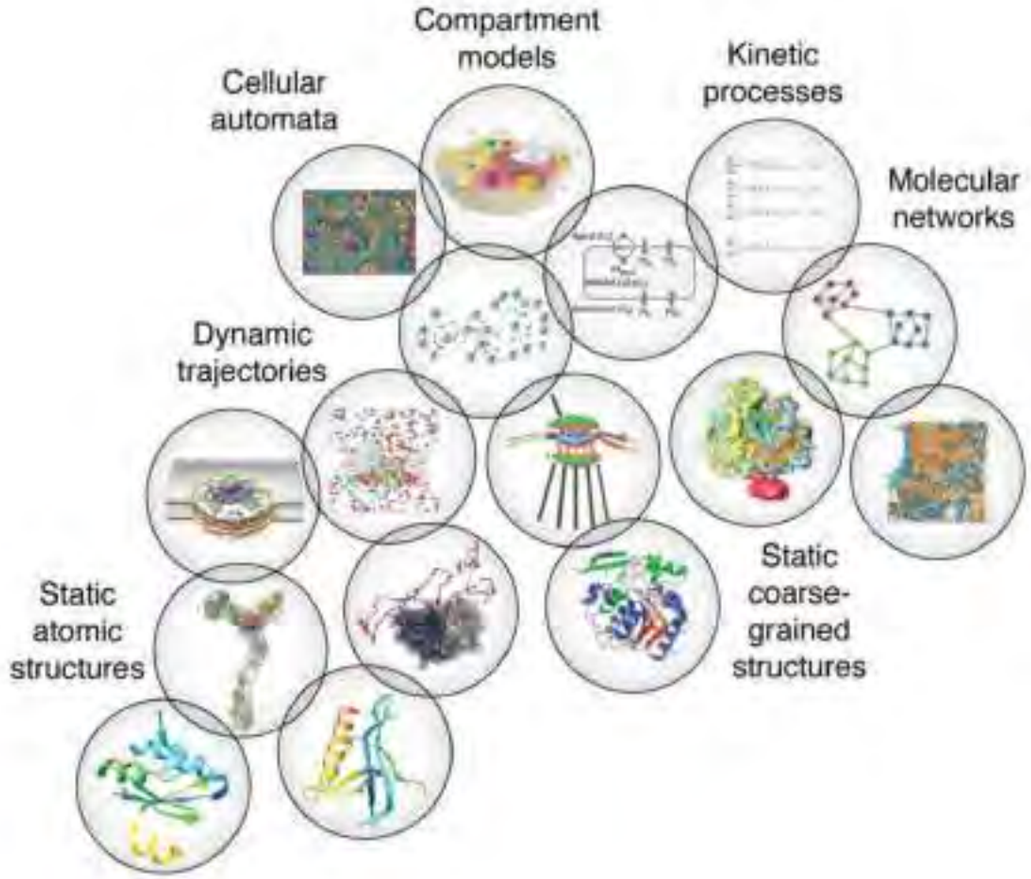
2. Harmonize these models with each other, by **“metamodeling”**.

Meta-modeling divide-and-conquers integrative modeling, to make it applicable to the entire cell

Integrative modeling:
Integrating heterogeneous types of information into a single model



Meta-modeling:
Integrating heterogeneous types of models into a single meta-model

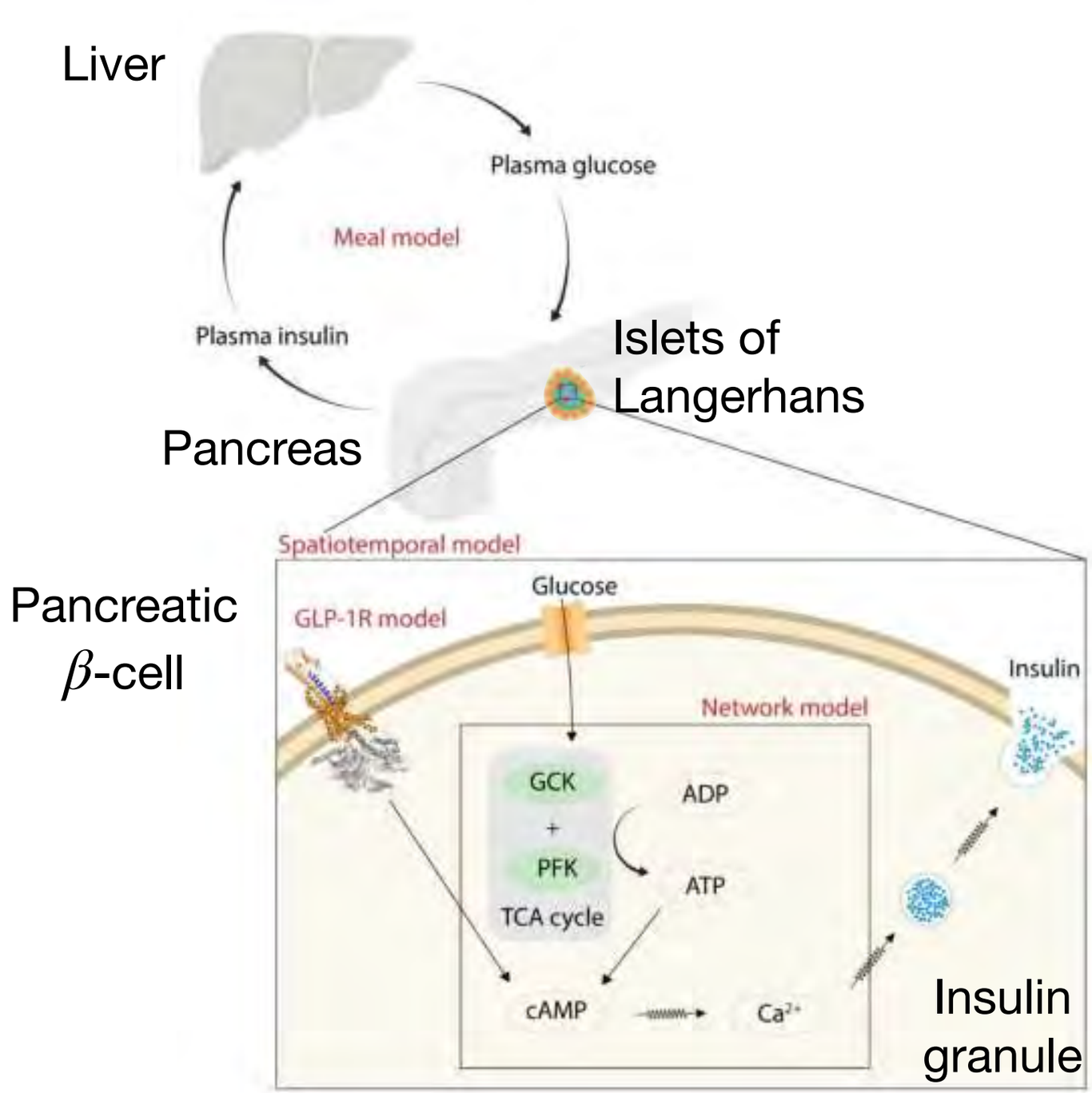


$$p(M|D,I) \propto p(D|M,I) \cdot p(M|I)$$

posterior likelihood prior

M model
D data
I prior information

Testbed system for modeling an entire cell (and beyond): Glucose-stimulated insulin secretion (GSIS) in pancreatic β-cells



Pancreatic β-cell

Bayesian metamodeling of biological systems: From data integration to model integration

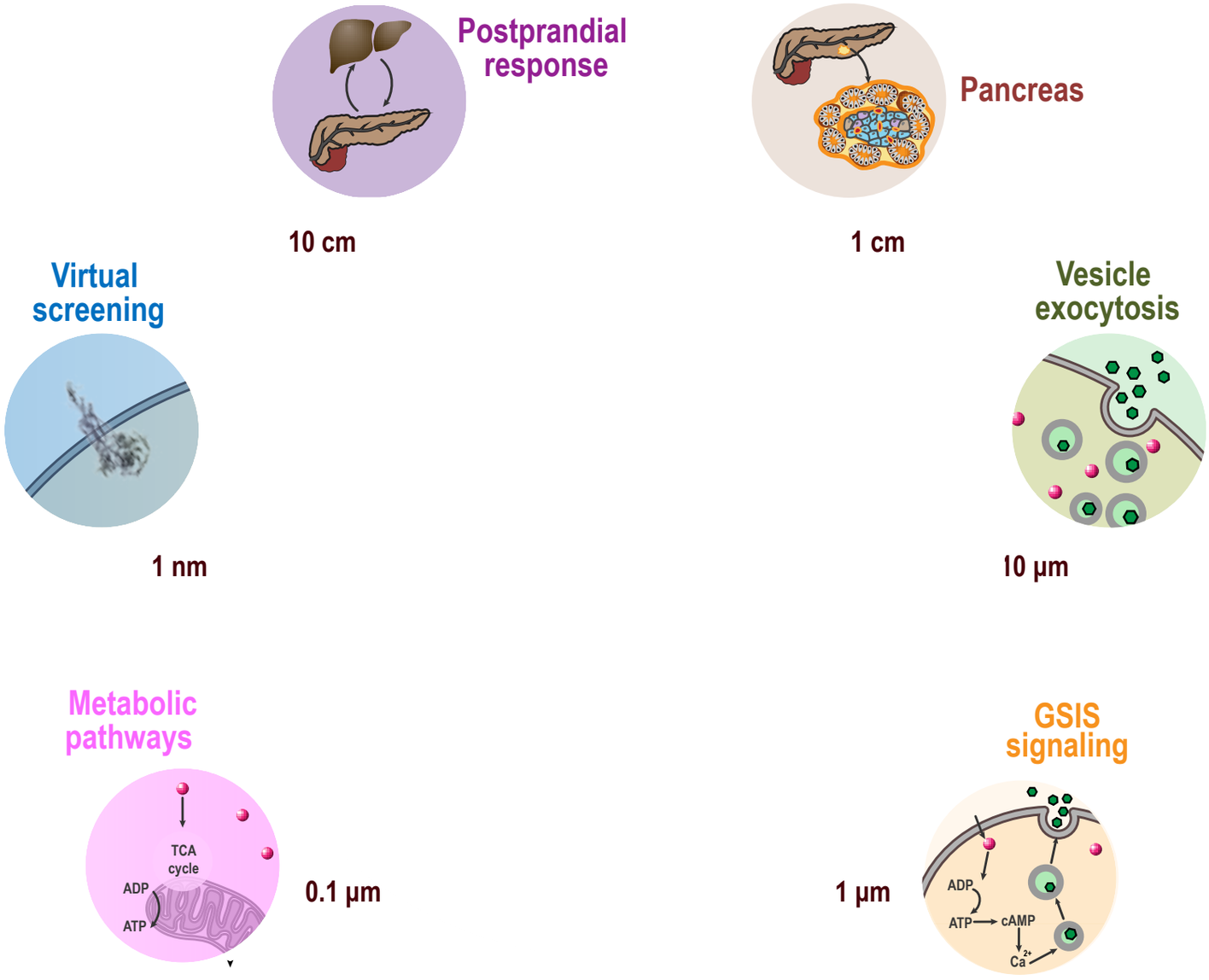
Input: Individual models

Different aspects of system
Different scales
Different representations
Different data

Output:

Metamodel of the entire system
Harmonized input models

"Whole" picture



Bayesian metamodeling: Overview of how

Translate each model into a common representation by constructing a probabilistic *surrogate model* describing statistical relations among model variables.

Identify and statistically couple related variables across probabilistic surrogate models, potentially referring to common reference variables.

Update variables of surrogate models followed by updating input models in the context of all other models.

Bayesian metamodeling: Input models

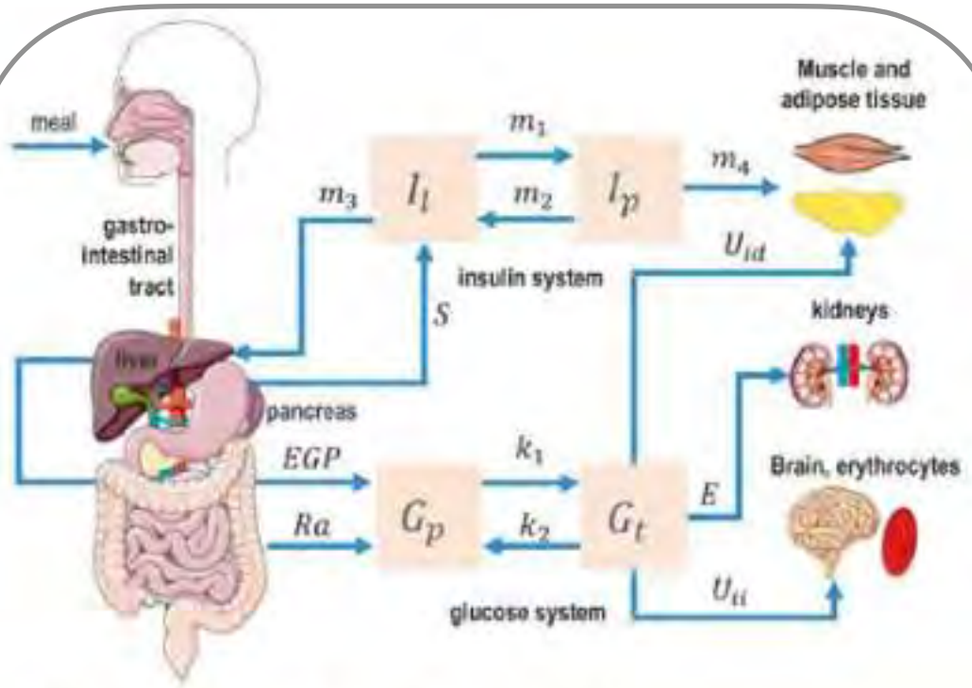
Input models

Conversion

Coupling

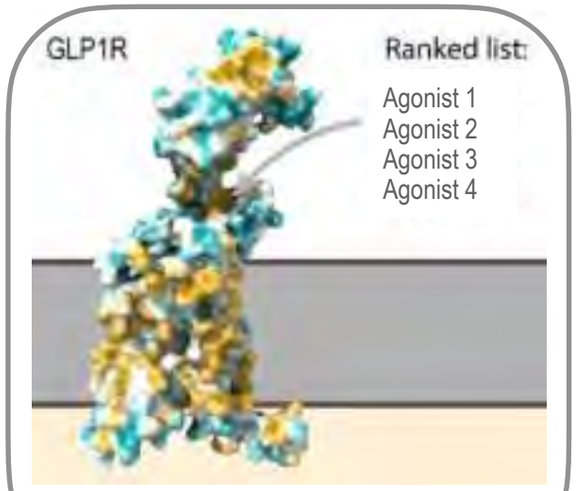
Back-propagation

Output

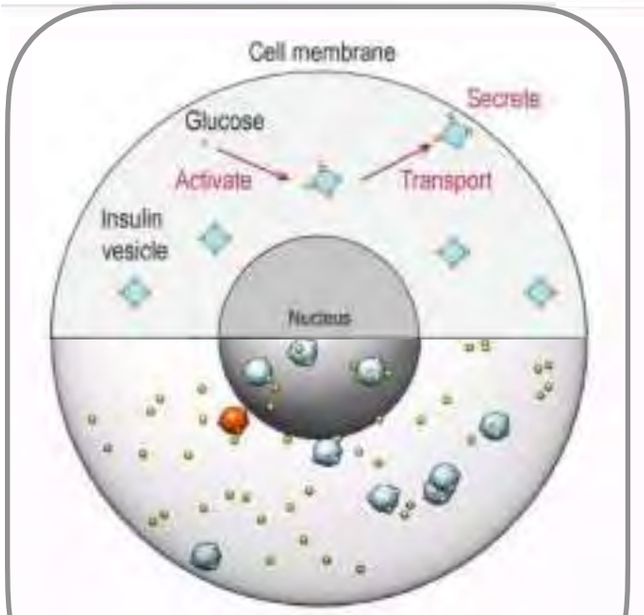


$$\dot{G}_p(t) = EGP(t) + Ra(t) - U_{ii}(t) - E(t) - k_1 G_p(t) + k_2 G_t(t)$$
$$\dot{G}_t(t) = U_{id}(t) + k_1 G_p(t) - k_2 G_t(t)$$
$$\dot{I}_l(t) = -[m_1 + m_3(t)] \cdot I_l(t) + m_2 I_p(t) + S(t)$$
$$\dot{I}_p(t) = -(m_2 + m_4) \cdot I_p(t) + m_1 I_l(t)$$
$$\vdots$$

postprandial response



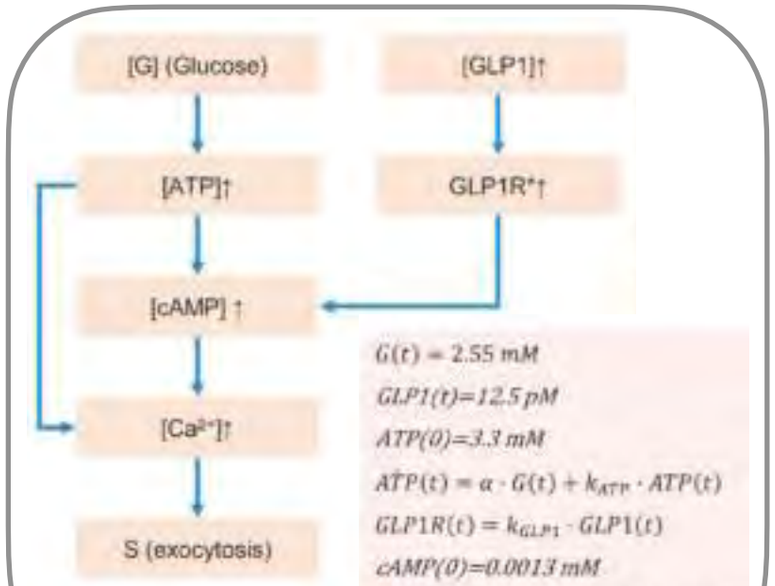
virtual screening



vesicle exocytosis

Level	GLP1 [pM]
low	400
medium	600
medium-high	1200
high	1600

GLP1 data

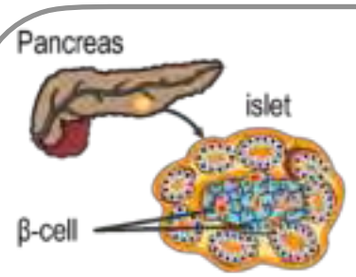


$$G(t) = 2.55\ mM$$
$$GLP1(t) = 12.5\ pM$$
$$ATP(0) = 3.3\ mM$$
$$\dot{ATP}(t) = \alpha \cdot G(t) + k_{ATP} \cdot ATP(t)$$
$$GLP1R(t) = k_{GLP1} \cdot GLP1(t)$$
$$cAMP(0) = 0.0013\ mM$$

GSIS signaling

Time [min]	ΔG_e [mM min^{-1}]
0	0.047
1	0.057
2	0.069

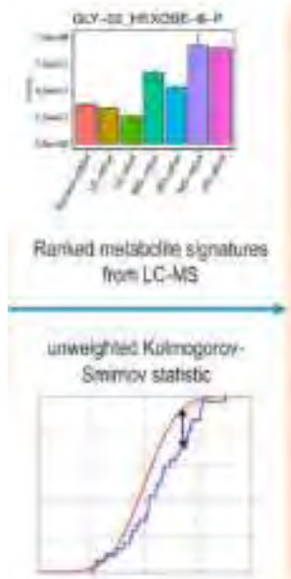
GI data



$$S_{pancreas} = N_i S_{islet}$$
$$S_{islet} = N_c S_{cell}$$

pancreas

Different conditions



Statistical enrichment score

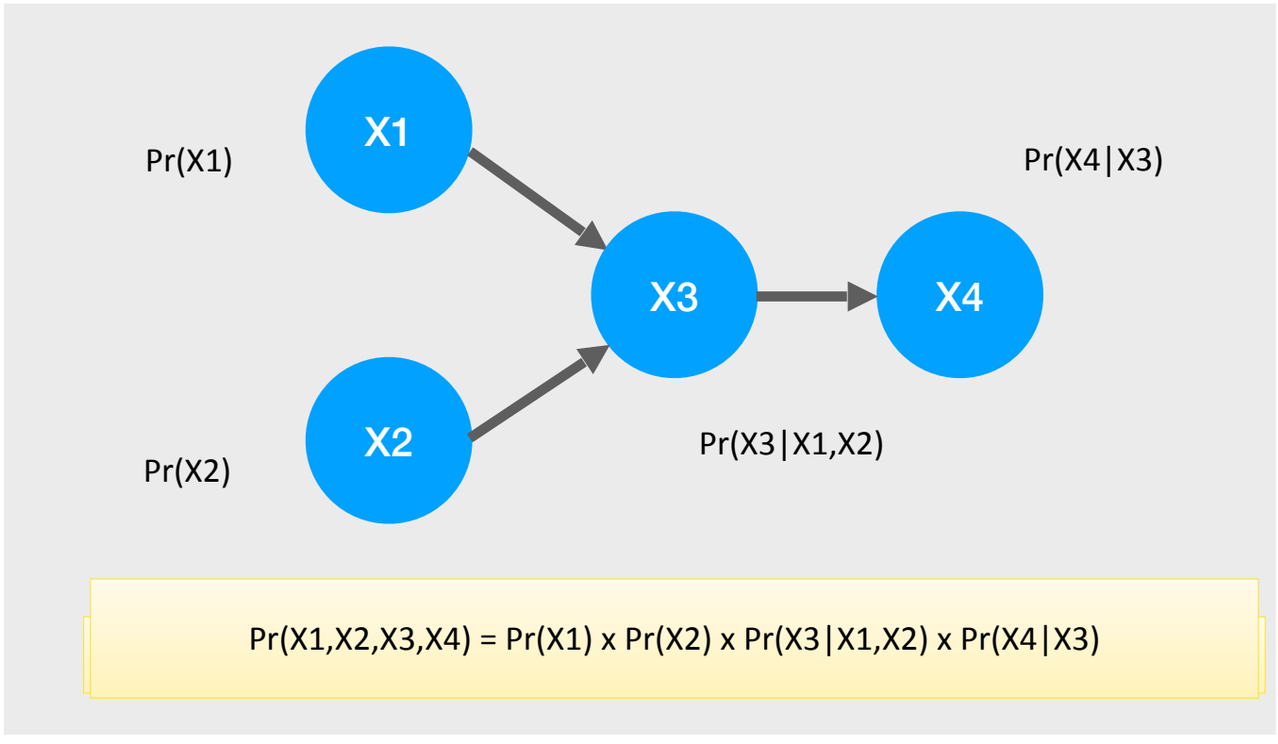
KEGG pathway	Score
Nicotinate and Nicotinamide	-1.474
Histidine	-1.212
Cysteine and Methionine	-1.060
ABC Transporters	-0.960
Glutathione	-0.900
Aminoacyl tRNA biosynthesis	-0.820
Glyoxylate and dicarboxylate	-0.767
Arginine and Proline	-0.660
Alanine, aspartate, and glutamate	-0.626
Pantothenate and CoA biosynthesis	1.455
Glycolysis and gluconeogenesis	1.462
Neuroactive ligand-receptor interaction	1.246
Pyrimidine	1.193
Purine	1.120
Pyruvate	1.024
Beta alanine	0.927
Pentose phosphate pathway	0.893
Lysine degradation	0.837
Citrate cycle (TCA)	0.830
Taste transduction	0.759
Proximal tubule bicarbonate reclamation	0.715
Taurine and hypotaurine	0.648
Glycine, serine and threonine	0.520

insulin metabolism

Bayesian metamodeling: Conversion

Need: A universal representation for any type of model.

Answer: Probabilistic Graphical Model (*eg*, Bayesian networks).

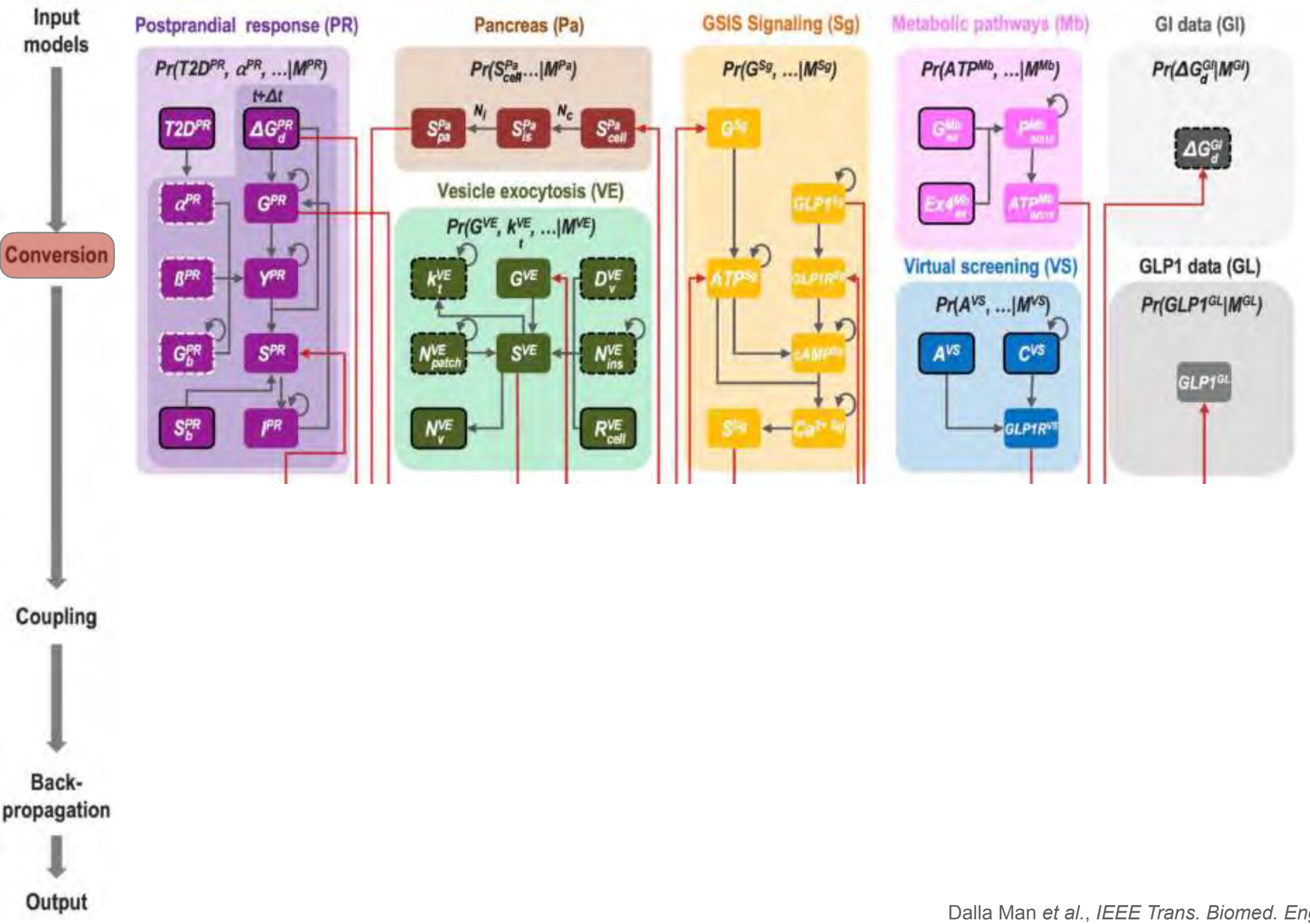


D. Koller, N. Friedman, F. Bach, *Probabilistic Graphical Models: Principles and Techniques* (MIT Press, 2009).

BNET in MatLab

Bayesian metamodeling: Conversion

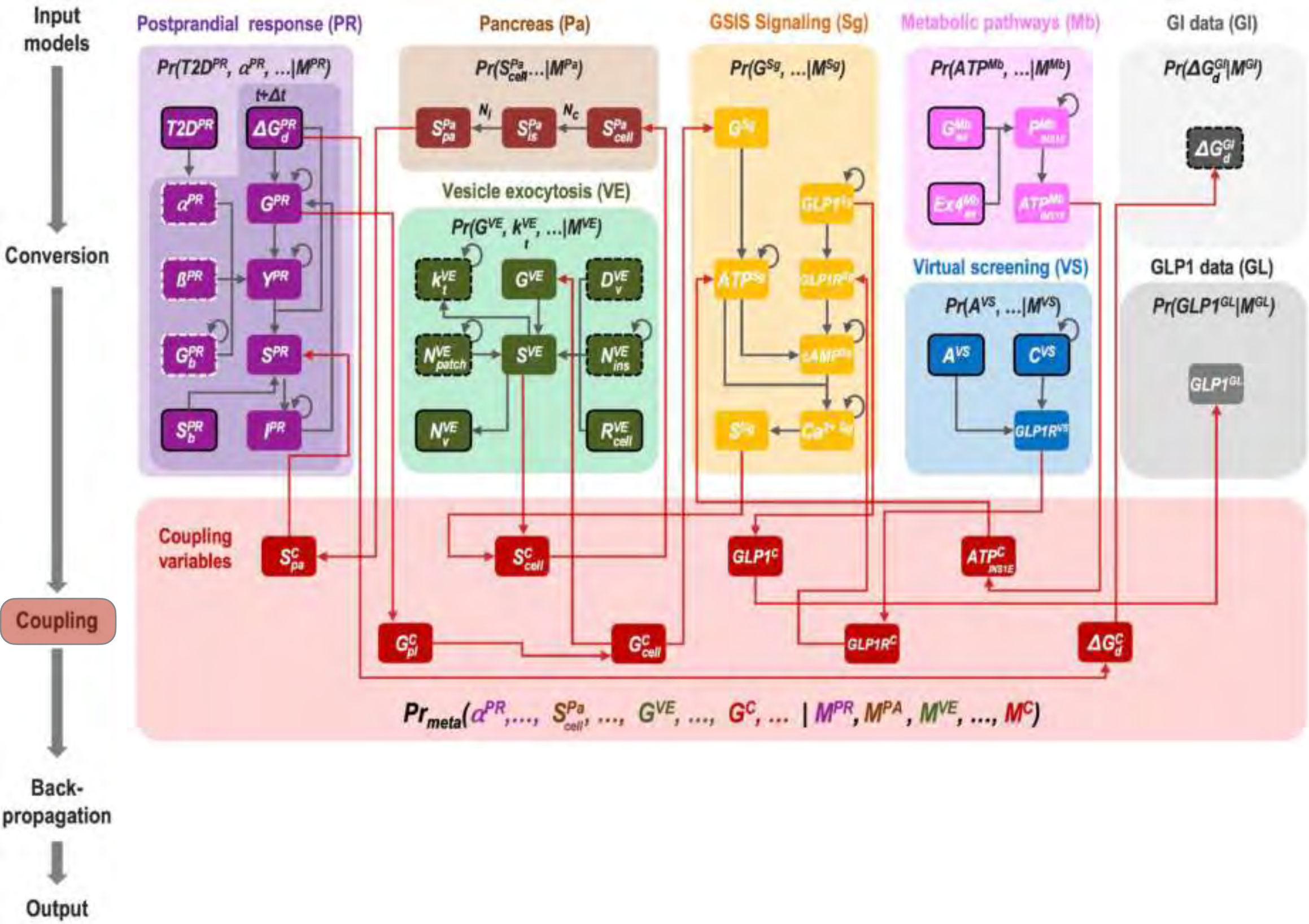
Learn a *probabilistic surrogate model* describing statistical relations among each model variables.



Dalla Man *et al.*, *IEEE Trans. Biomed. Eng.*, 2007

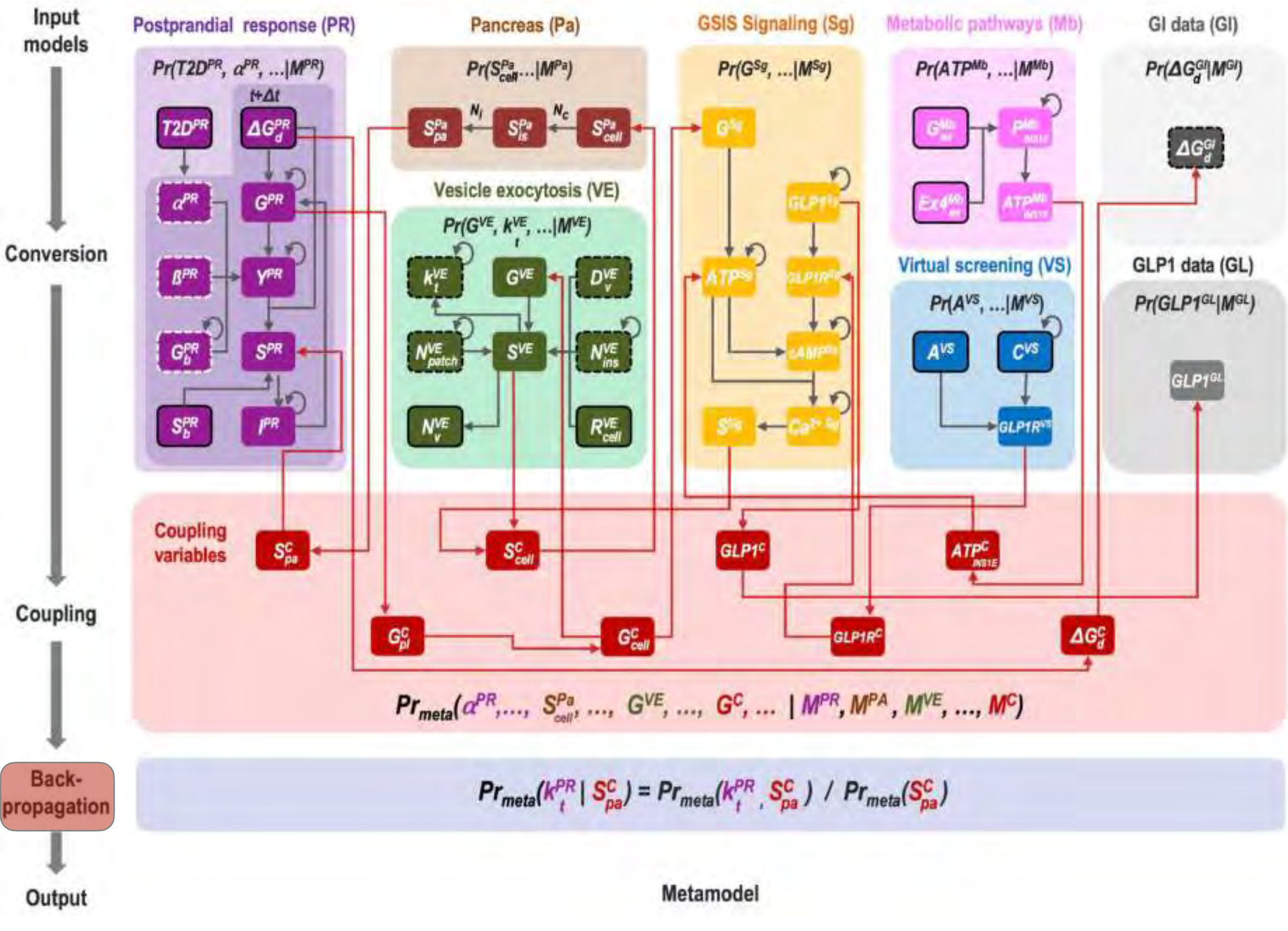
Bayesian metamodeling: Coupling

Identify and statistically couple related variables across probabilistic surrogate models.



Bayesian metamodeling: Backpropagation

Update variables of input models in the context of all other models.



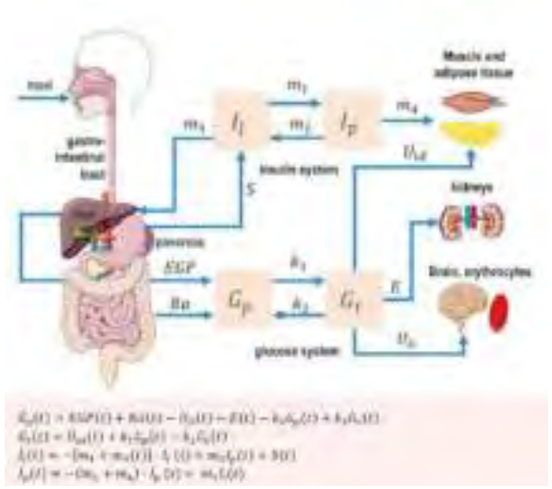
Advantages of Bayesian metamodeling

- **Modularity**
  iteratively expand with new models
- **Multiscale modeling**
  different scales bridges through statistical correlations
- **Efficiency**
  models constructed and computed in parallel
- **Collaboration**
  sharing of resources, data, models, expertise
- **Completeness**
  different aspects straightforwardly integrated
- **Contexualization**
  information flow between variables of different models
- **Bayesian formalism**
  uncertainty estimates
- **Accuracy and precision**
  assessed and asymptotically improving
- **Conflict resolution**
  identify and resolve conflicts among models

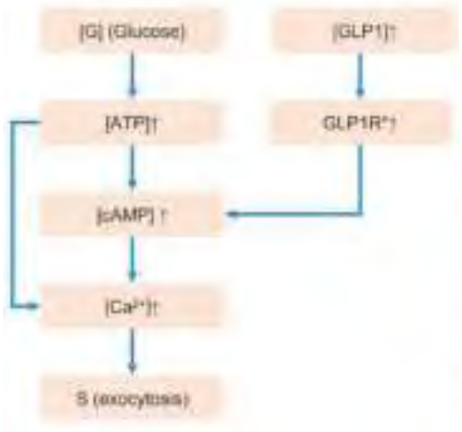
Contextualization: information flow between variables from different models, representations, and scales

Metamodeling captures incretin effect (*ie*, the effect of elevated GLP1 concentrations on postprandial insulin and glucose levels; Drucker *et al*. *Lancet,* 2006).

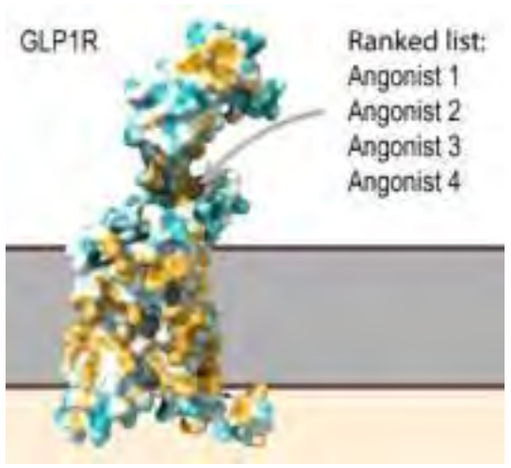
Postprandial response (body; ODE)
Excludes GLP1

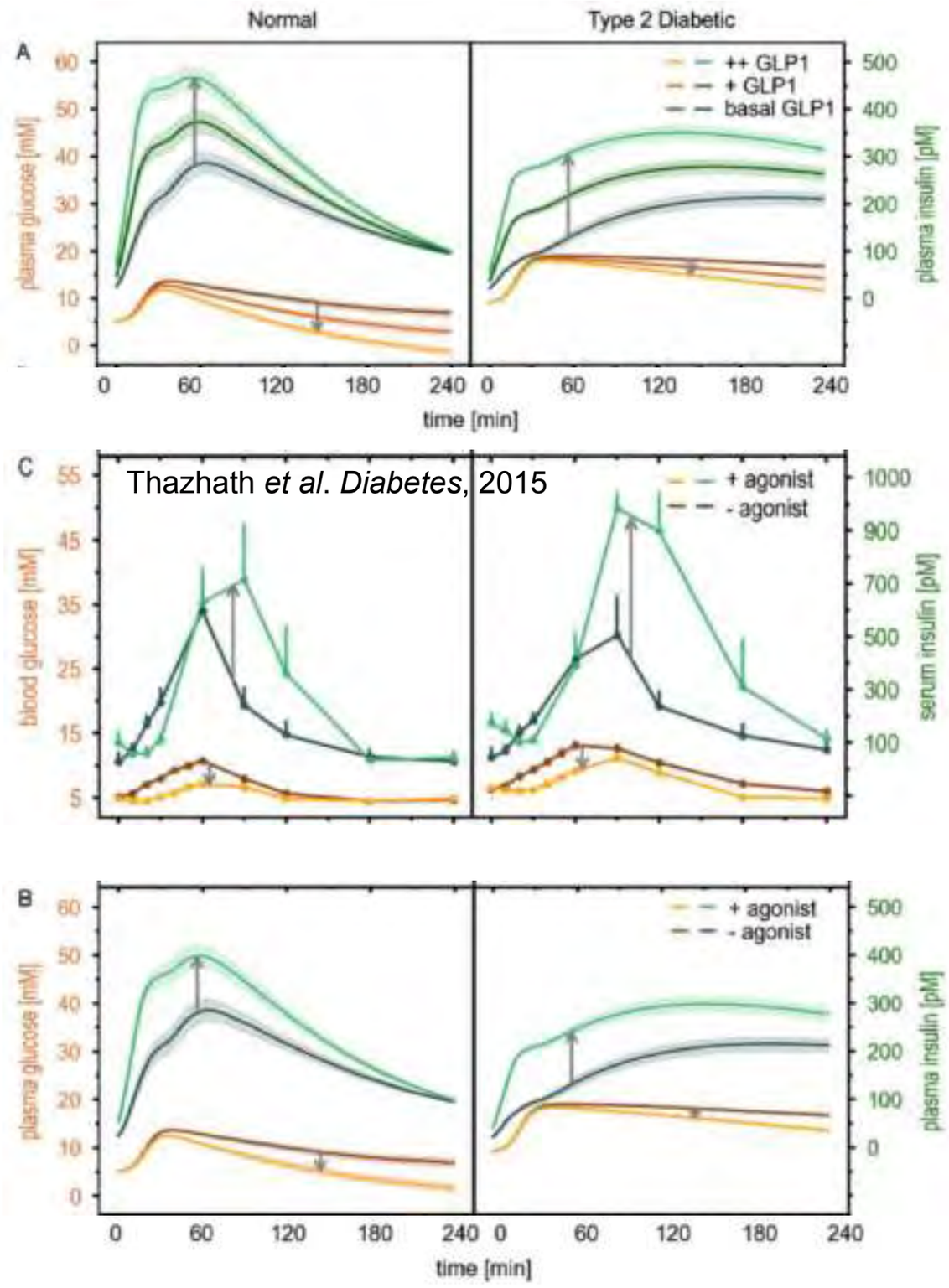


Signaling (cell; ODE)
Includes GLP1



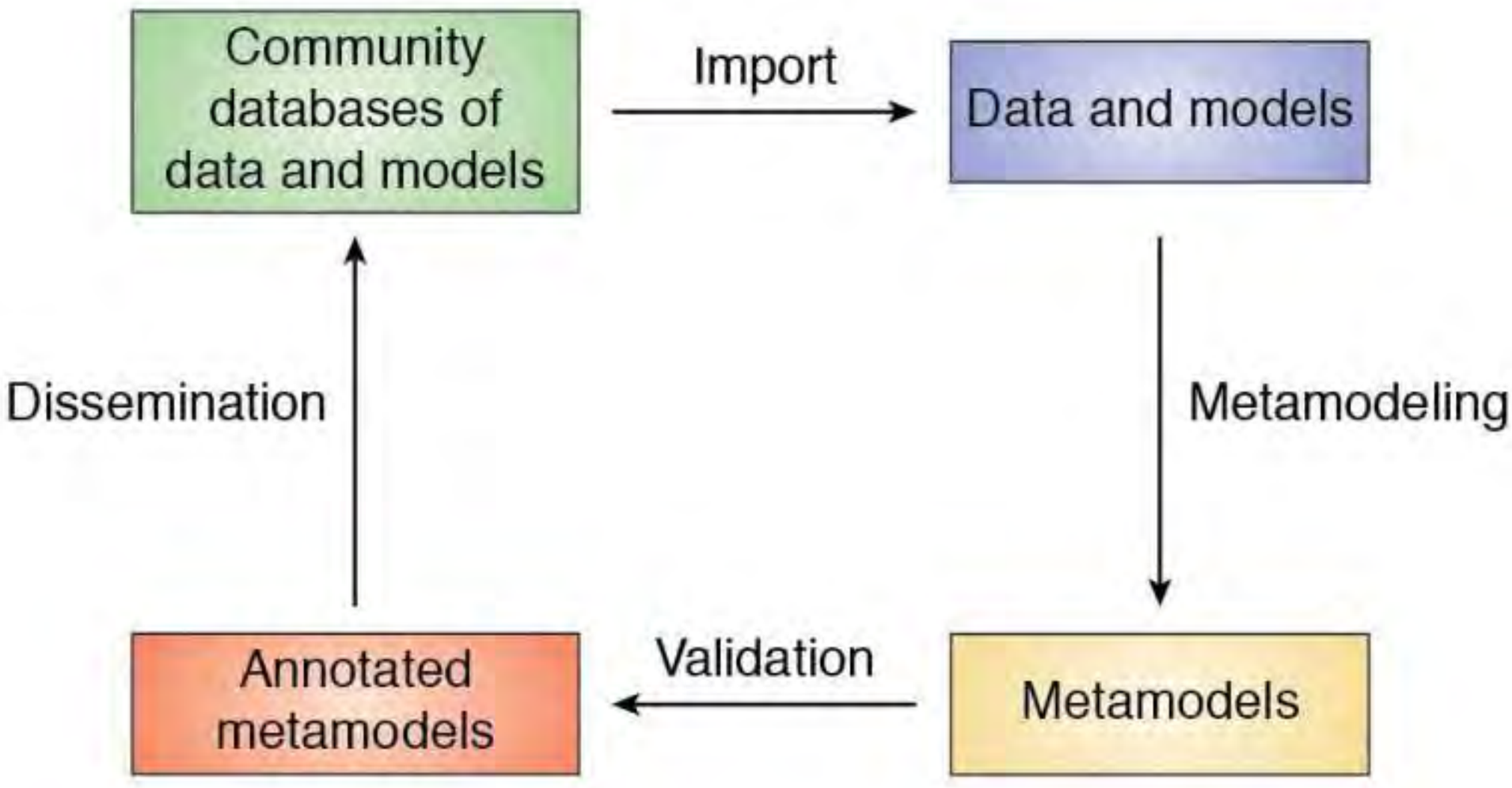
GLP1R activation (atomic; spatial)
Includes GLP1 agonists





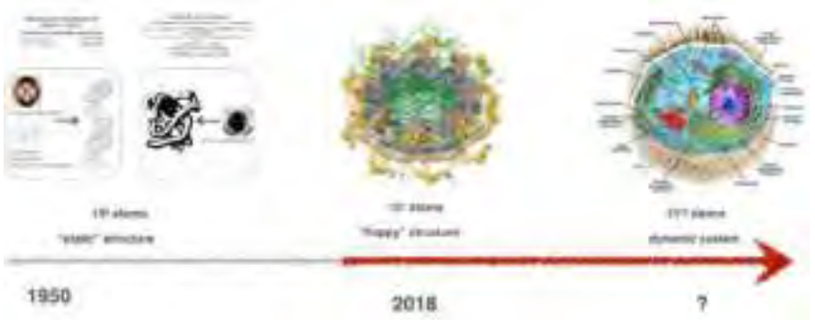
Vision for metamodeling for cell mapping

Our goal is to develop a comprehensive cyber infrastructure platform for modeling the cell. This platform will be critical for establishing an effective socio-technical ecosystem that will allow scientists to experiment, collaborate, and disseminate reproducible experimental data and cell models.



Sali. *JBC*, 2021

w/ Carl Kesselman, Kate White, Brinda Vallat, Helen Berman, PBCC

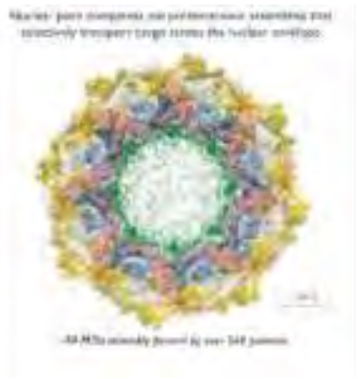


Summary

- Integrative modeling is a general approach for modeling based on various types of data.

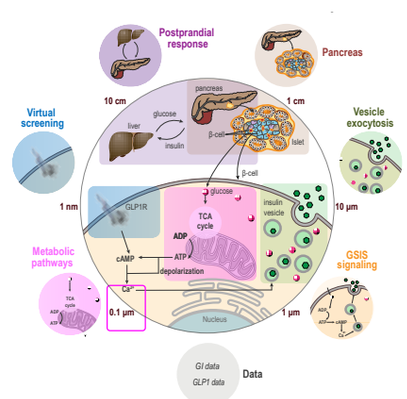
$$p(M|D,I) \propto p(D|M,I) \cdot p(M|I)$$

- Application of integrative modeling can describe structures of large biomolecular systems, especially those whose component structures are already available in the Protein Data Bank.



- Integrative structure models and data are best disseminated through the worldwide Protein Data Bank.

- Bayesian metamodeling is a modular framework for integrative modeling through model integration, potentially applicable to mapping the cell.



Acknowledgements: Integrative biology is a team sport

Our group @ UCSF

Tanmoy Sanyal
Jeremy Tempkin
Ben Webb
Ignacia Echeverria
Ilan Chemmama
Daniel Saltzberg
Seth Axen
Sai Ganesan
Thomas Peulen
Aji Palar
Matthew Hancock
Dibyendu Mondal

Our group (former)

Barak Raveh (Hebrew Univ)
Kala Pilla
Seung Joong Kim
Riccardo Pellarin
Charles Greenberg
Shruthi Viswanath
Max Bonomi
Kala Pilla
Nikita Chopra
Leah Kressin
Javier Velazquez Muriel
Daniel Russel
Keren Lasker
Sara Calhoun
Peter Cimermancic
Dina Schneidman
Elina Tjioe
GQ Dong
Riccardo Pellarin
Frank Alber
Bret Peterson
Friedrich Foerster
Mike Kim
Maya Topf

iHuman Institute at ShanghaiTech

Liping Sun
ChenXi Wang
Jihui Zhao
Weimin Li
Angdi Li

Nuclear Pore Complex

Mike Rout (Rockefeller Univ)
J Fernandez-Martinez
I Nudelman
EY Jacobs
AS Chaudhury
R Mironska
Brian Chait (Rockefeller Univ)
Y Shi
W Zhang
R Williams
J Wang
John Aitchison (ISB)
T Herricks
Chris Akey (BU)
Steven Ludtke (Baylor)
David Stokes (NYSBC)
P Upla
JL Gerton (Stowers Inst)
BD Slaughter
M Shivaraju
JR Unruh
Martin Jarrold (Indiana Univ)
J Hogan
Jason de la Cruz (HHMI)
Z Yu
David Cowburn (AECOM)
Steven Almo (AECOM)

wwPDB Hybrid/Integrative Methods Task Force

Stephen Burley
Sameer Velankar
Gerard Kleywegt
Haruki Nakamura
Gensi Kurisu
John Markley

Data Working Group:
Helen Berman
John Westbrook
Brinda Vallat
Jill Trewhella

Model Working Group:
Torsten Schwede
Jens Meiler
Frank Dimao
Gerhard Hummer
Emad Tajkhorshid
Alexandre Bonvin

Tom Ferrin
Tom Goddard

Pancreatic β-Cell Consortium

pbcconsortium.org

Ray Stevens (USC, iHuman)
Kyle McClary
Kate White (USC)*
Jitin Singla
Nick Graham (USC)
Donqing Zheng
Peter Butler (UCLA)
Carl Kesselman (USC)
Frank Alber (UCLA)
Helen Berman (USC, Rutgers U)
Brinda Vallat (Rutgers U)

Nuclear Pore Complex @EMBL

Jan Ellenberg
Shotaro Otsuka
Antonio Politi
Arina Rybina
Oeyvind Oedegard
Vilma Jimenez Sabinina

Integrative Modeling Platform (IMP)
open source, salilab.org/imp

Funding
NIH, NSF
EMBL (Baden-Württemberg Stiftung)