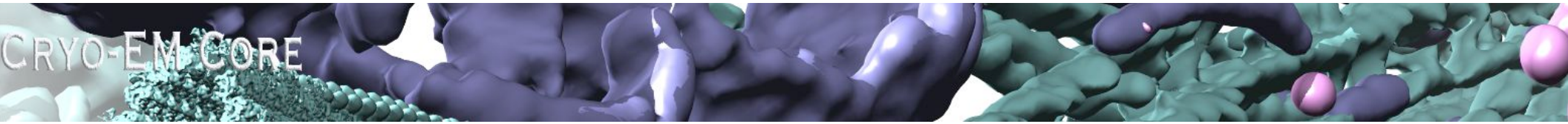


Conformational Variability in CryoEM using Deep Learning

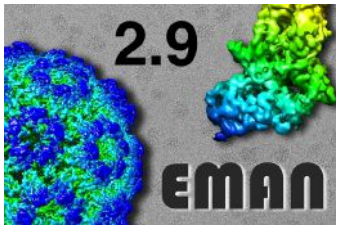
Steve Ludtke

Baylor College of Medicine
Charles C. Bell Jr., Professor of Structural Biology
Biochemistry and Molecular Biology
Deputy Director, Advanced Technical Cores
Director, CryoEM/CryoET Core

https://youtu.be/HcxoaEEWjII
(longer talk at IPAM)

Baylor College of Medicine

ADVANCED TECHNOLOGY CORES



Baylor College of Medicine

VERNA & MARRS McLEAN DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

Structural Biology

- NMR
  - local structure, specific distances, local dynamics
- X-ray Crystallography
  - high resolution structures, often non-native

Alphafold →

ok... it's pretty decent. Positive and negative impacts?

- CryoEM → **Still needed for info on complexes & flexibility/dynamics**
  - intermediate - high resolution structure, in-vitro flexibility
- CryoET + CryoFIB → **Still needed to observe native structures, observe native assembly, etc.**
  - low - intermediate resolution structure, 3-D variability, cellular context
- Super-resolution Fluorescence (dynamic localization/co-localization)

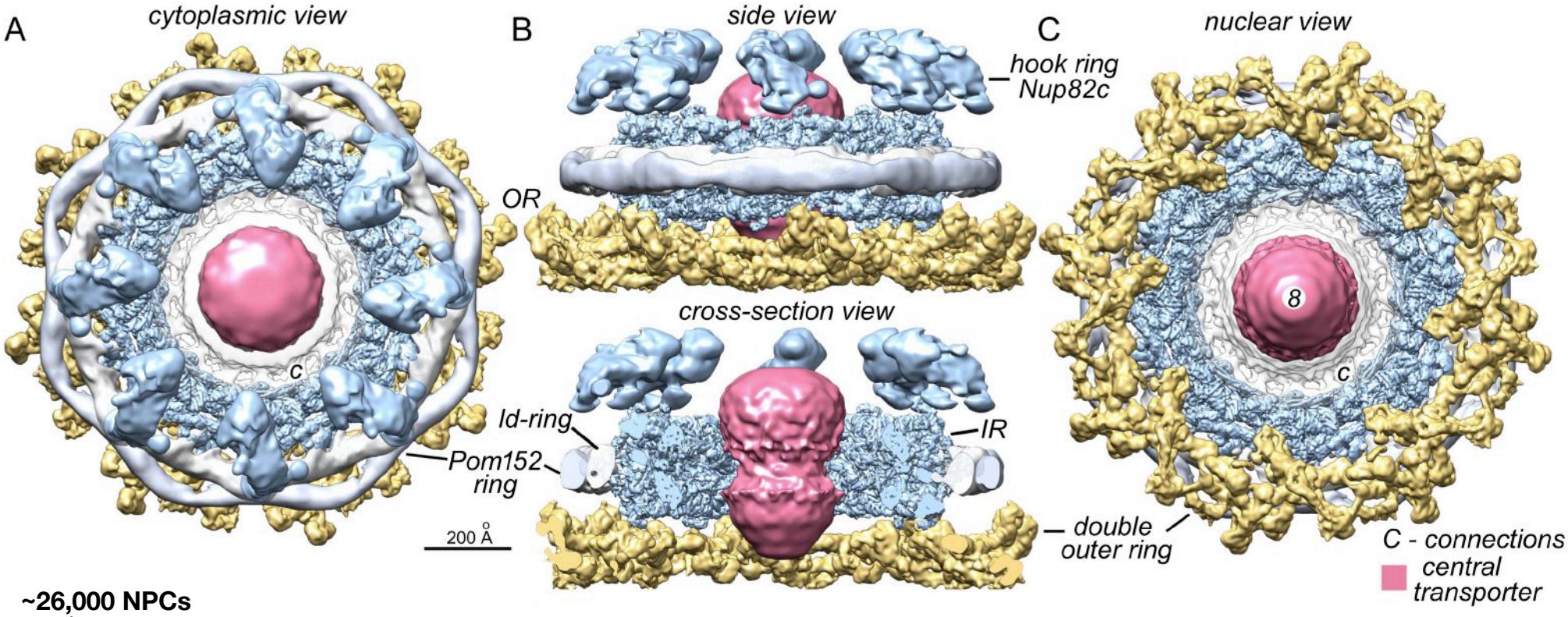
Comprehensive structure and functional adaptations of the yeast nuclear pore complex

Christopher W. Akey,[1,18,*] Digvijay Singh,[2,18] Christna Ouch,[1,3,18] Ignacia Echeverria,[4,12,18] Ilona Nudelman,[5] Joseph M. Varberg,[9] Zulin Yu,[9] Fei Fang,[6] Yi Shi,[6] Junjie Wang,[7] Daniel Salzberg,[4] Kangkang Song,[3] Chen Xu,[3] James C. Gumbart,[8] Sergey Suslov,[2] Jay Unruh,[9] Sue L. Jaspersen,[9,10] Brian T. Chait,[7] Andrej Sali,[4,13,14] Javier Fernandez-Martinez,[5,16,17,*] Steven J. Ludtke,[11,*] Elizabeth Villa,[2,15,*] and Michael P. Rout[5,19,*]

[1]Department of Physiology and Biophysics, Boston University School of Medicine, 700 Albany Street, Boston, MA 02118, USA
[2]Section of Molecular Biology, Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA
[3]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, MA 01605, USA
[4]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA, USA
[5]Laboratory of Cellular and Structural Biology, The Rockefeller University, New York, NY 10065, USA
[6]Department of Cell Biology, University of Pittsburgh, Pittsburgh, PA, USA
[7]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, NY, USA
[8]School of Physics, Georgia Institute of Technology, Atlanta, GA 30332, USA
[9]Stowers Institute for Medical Research, Kansas City, MO, USA
[10]Department of Molecular and Integrative Physiology, University of Kansas Medical Center, Kansas City, KS, USA
[11]Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, 1 Baylor Plaza, Houston, Texas 77030, USA
[12]Department of Cellular and Molecular Pharmacology, San Francisco, San Francisco, CA 94158, USA
[13]Quantitative Biosciences Institute, University of California San Francisco, San Francisco, CA 94158, USA
[14]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA 94158, USA
[15]Howard Hughes Medical Institute, University of California, San Diego, La Jolla, CA 92093, USA
[16]Present address: Ikerbasque, Basque Foundation for Science, 48013 Bilbao, Spain
[17]Present address: Instituto Biofisika (UPV/EHU, CSIC), University of the Basque Country, 48940 Leioa, Spain
[18]These authors contributed equally
[19]Lead contact

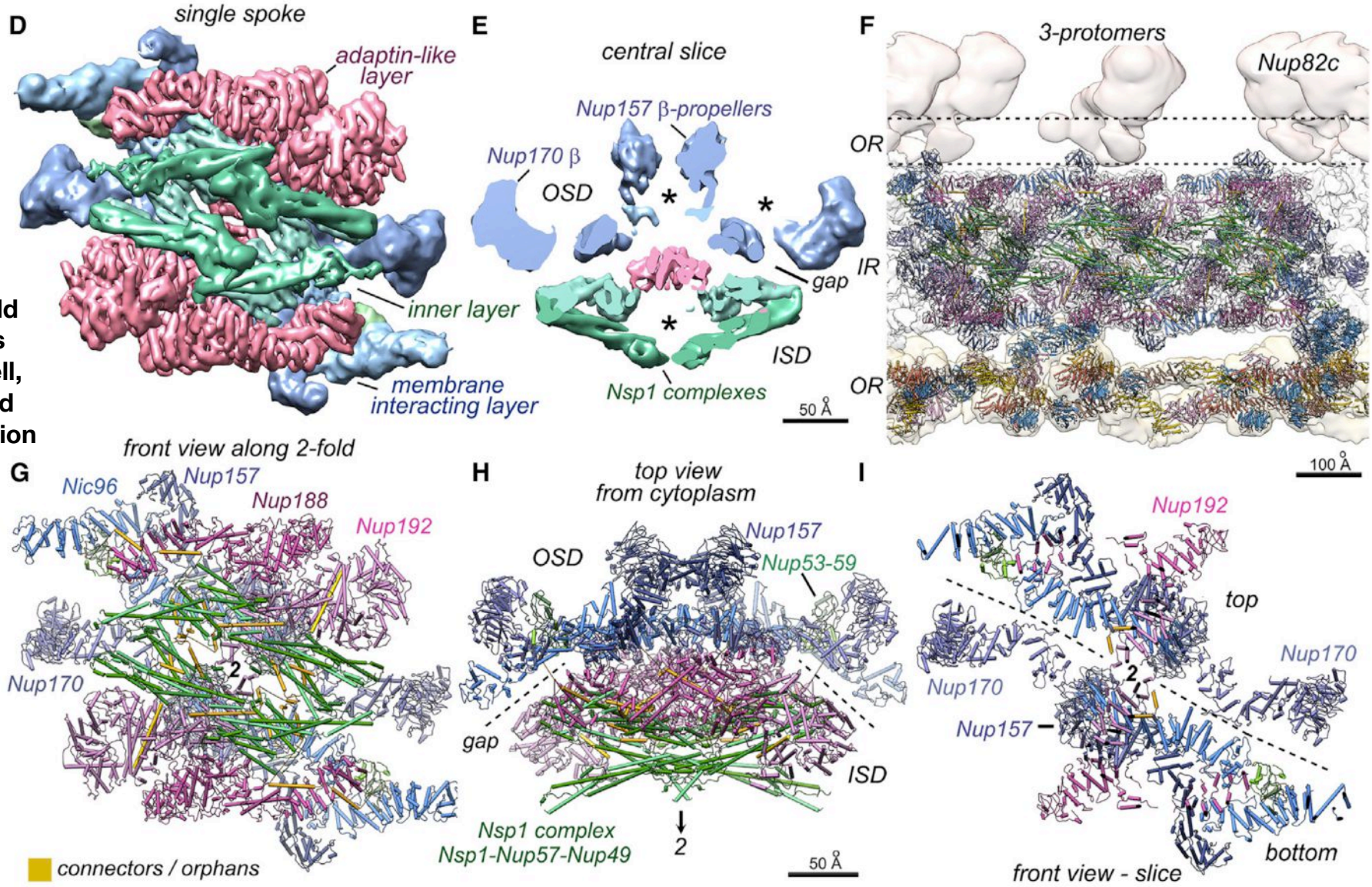
Cell. (2022) 185, 361-378.e25.

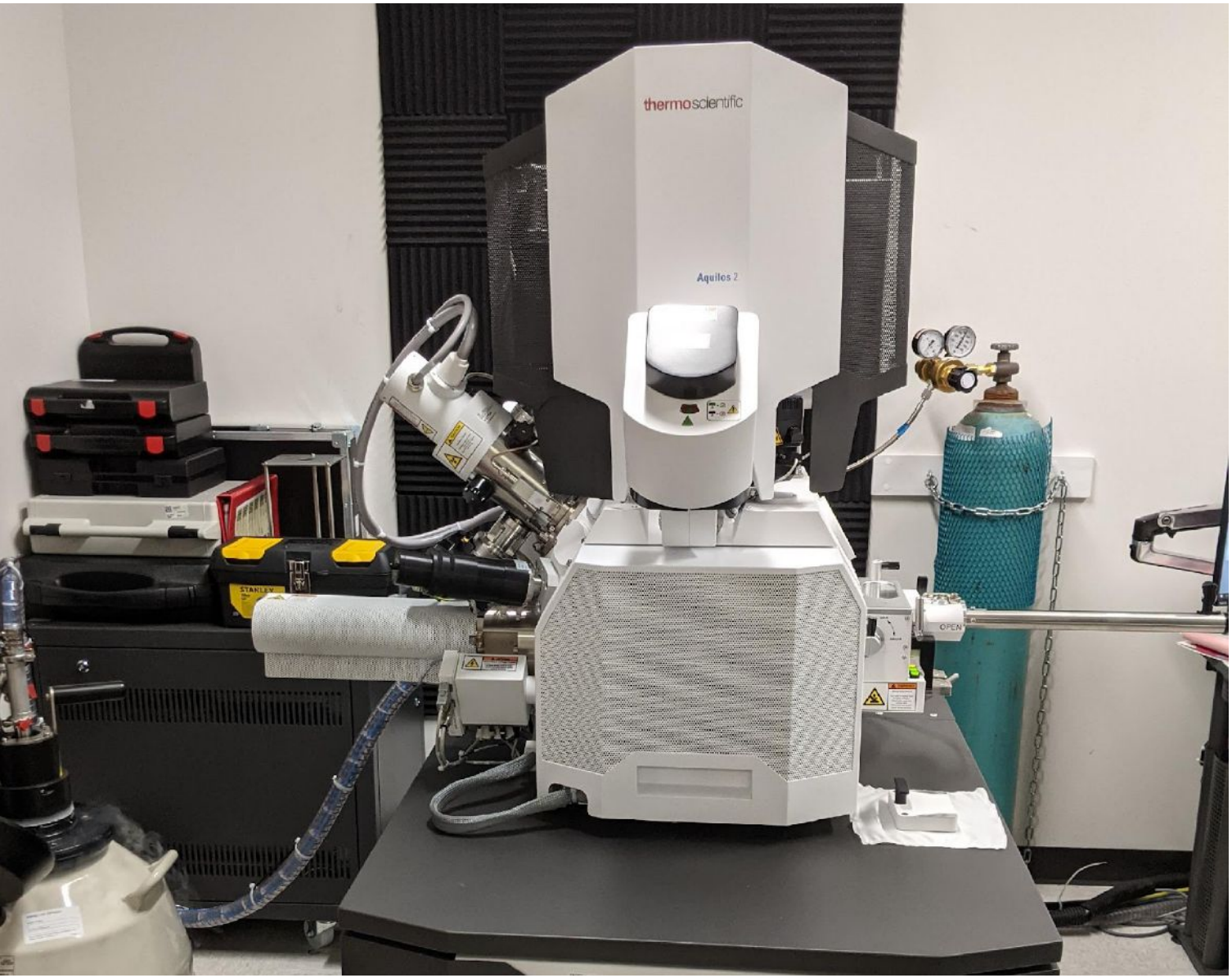
Single Particle CryoEM



~26,000 NPCs
7-11 Å resolution

Alphafold models dock well, but need modification





ThermoFisher Aquilos 2
Cryo FIB/SEM
with EasyLift

FIB = Focused Ion Beam
SEM = Scanning Electron Microscope

Used for:
- cutting 100-500 nm cellular lamella
- tissue lift-out
- automatic slice-and-view

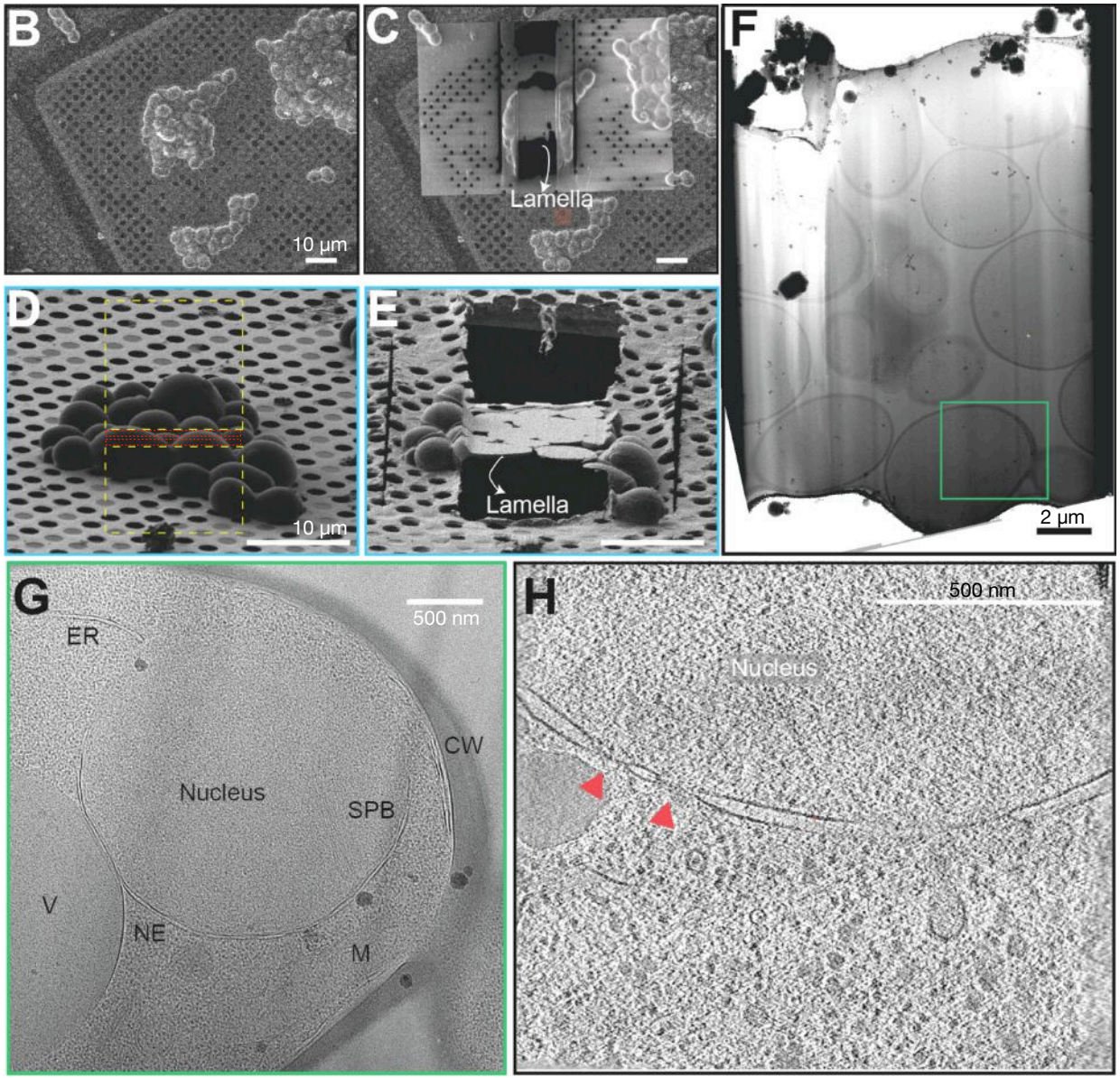
- Cryo-SEM imaging
- Platinum sputter coater
- Platinum GIS deposition
- Gallium FIB

To be purchased this year:
- iFLM (widefield fluorescence)*
- EM-ICE high pressure freezer

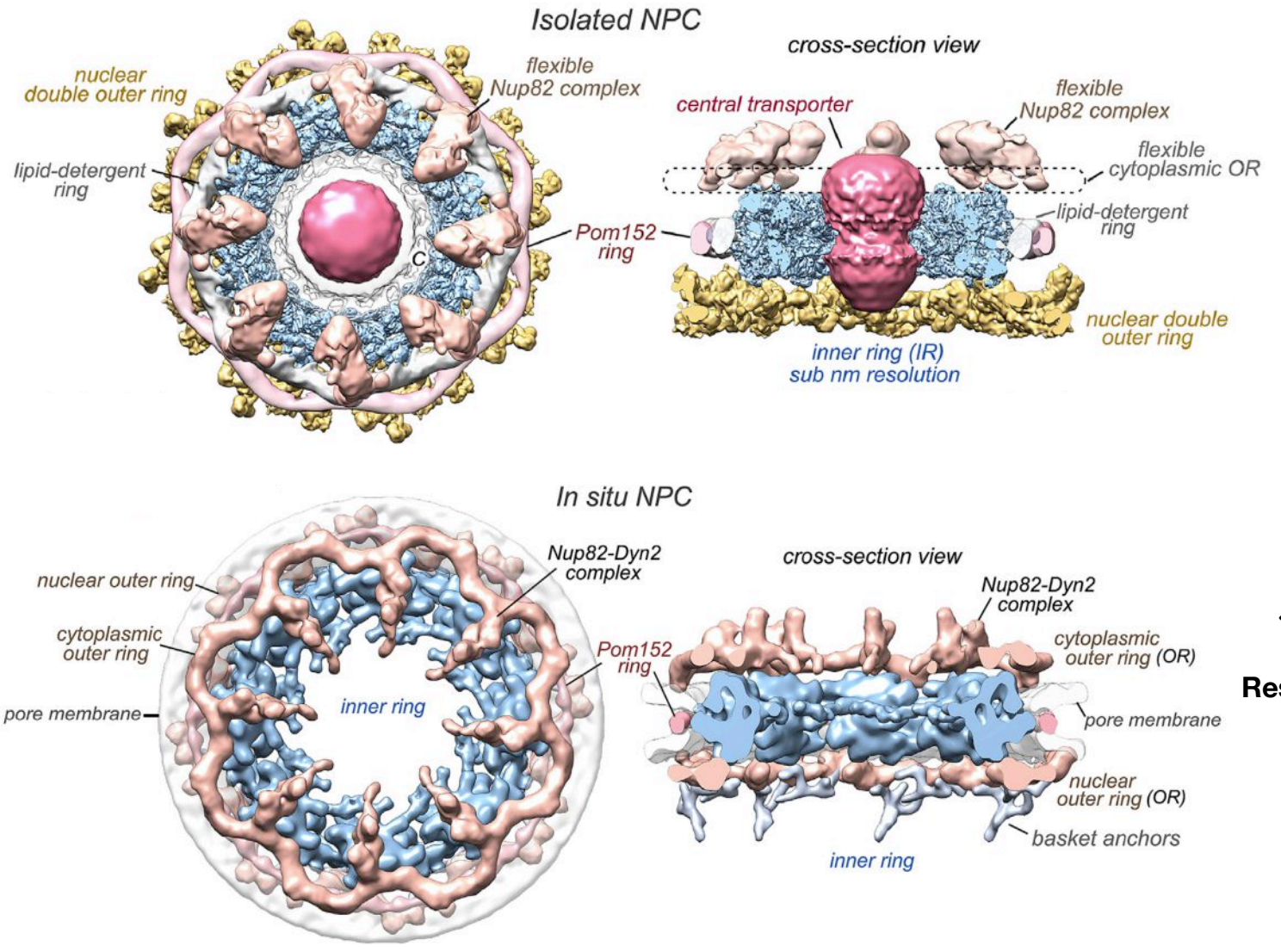
S. cerevisiae Nuclear Pore Complex

~52 MDa, ~30 unique Nups

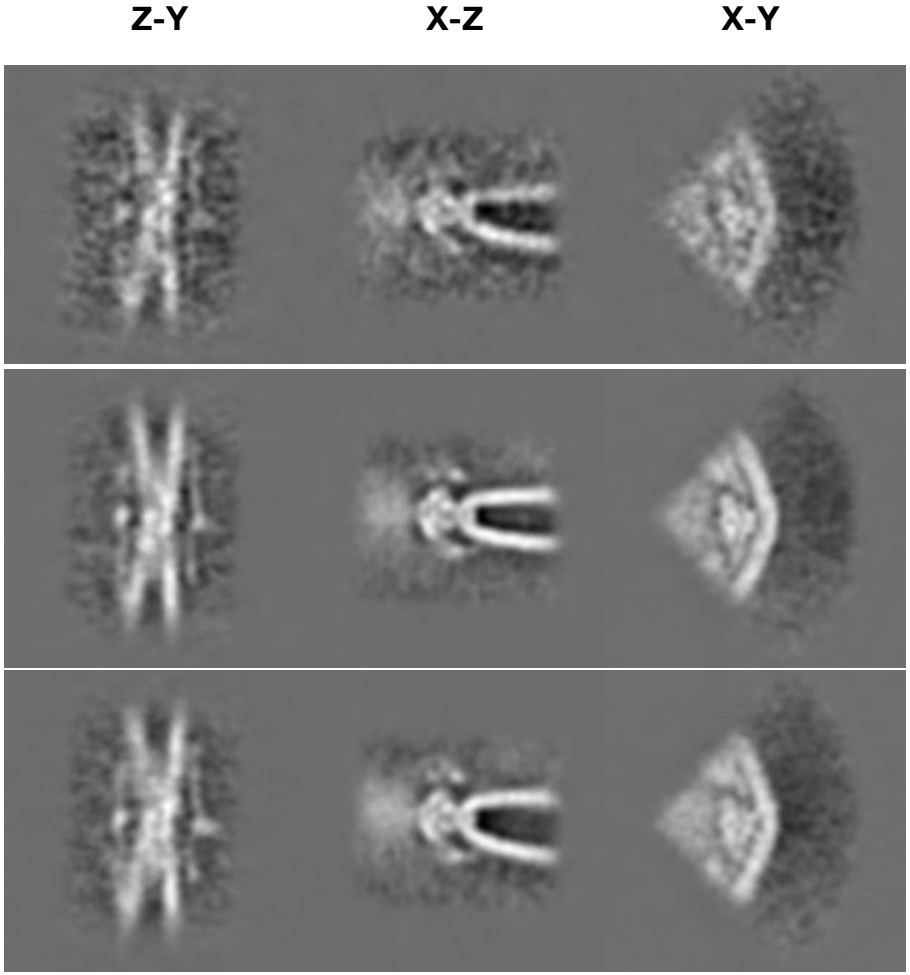
~550 total proteins in complex



Digvijay Singh
Villa Lab

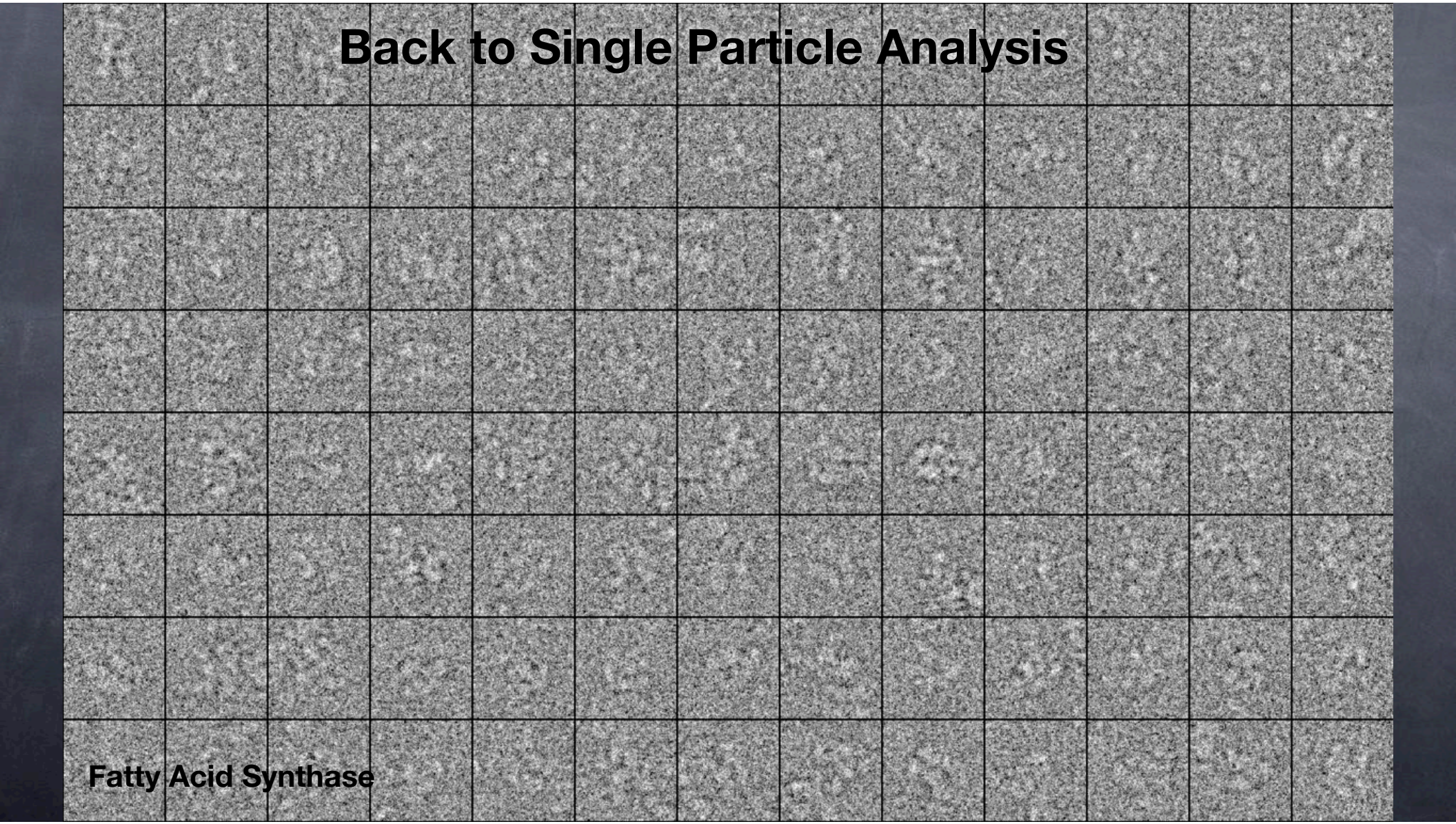


~1000 Tomograms
~1000 NPCs
Resolution gets "stuck" at ~30 Å



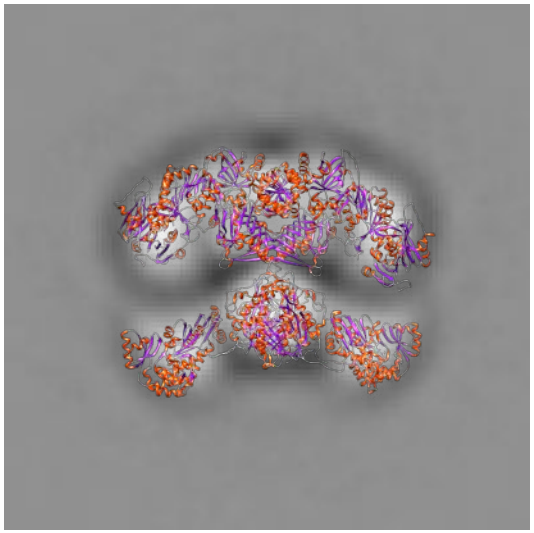
3 Orthogonal Views
Traditional Classification (subtomogram)

Back to Single Particle Analysis



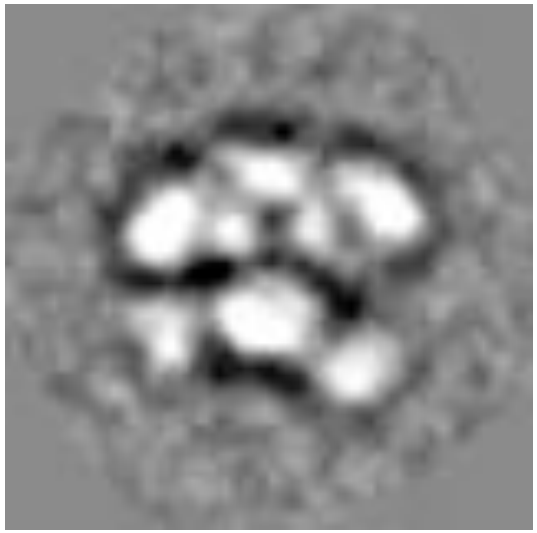
Fatty Acid Synthase

Single View CryoEM Average



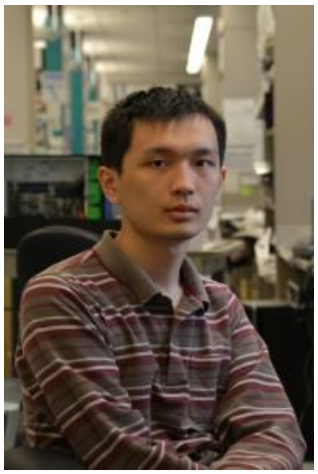
150 Å

Fatty Acid Synthase, ~30 Å motion



150 Å

GMM Acknowledgements

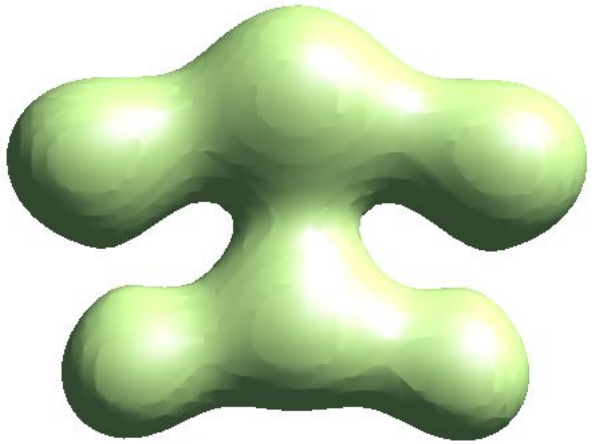


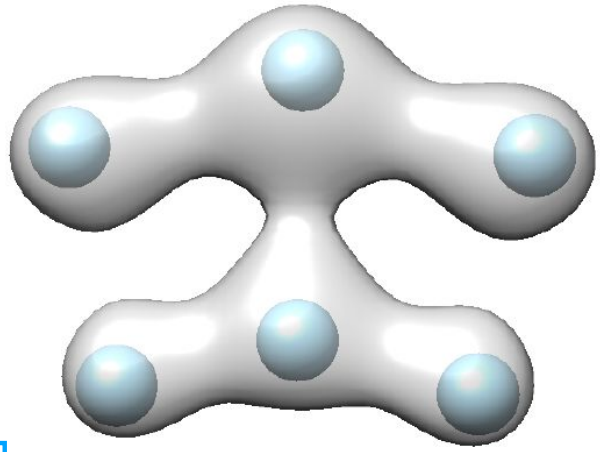
Muyuan Chen

We thank the NIH for its support: R01GM080139.

Chen, M. and Ludtke, S. J., Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM. (2021). *Nature Meth*. 18, 930-936.

Gaussian representation





$$map(x) = \sum A_j e^{\frac{-(x-p_j)^2}{\sigma_j^2}}$$

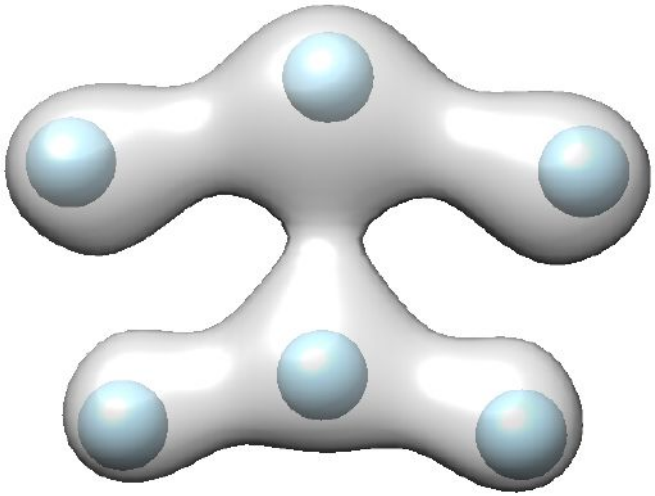
3D Position vector

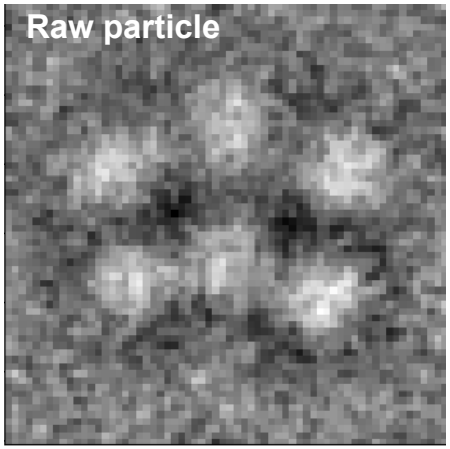
Amplitude

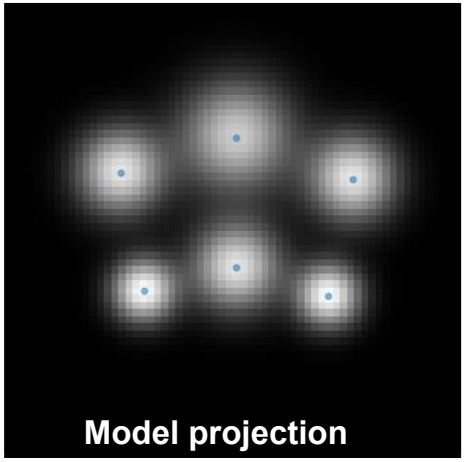
Sigma

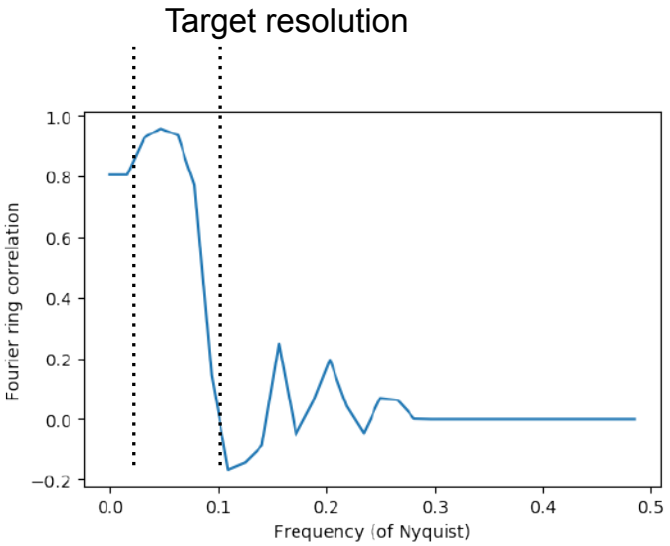
- Adjustable complexity
- Easier to model continuous motion

Particle-projection comparison



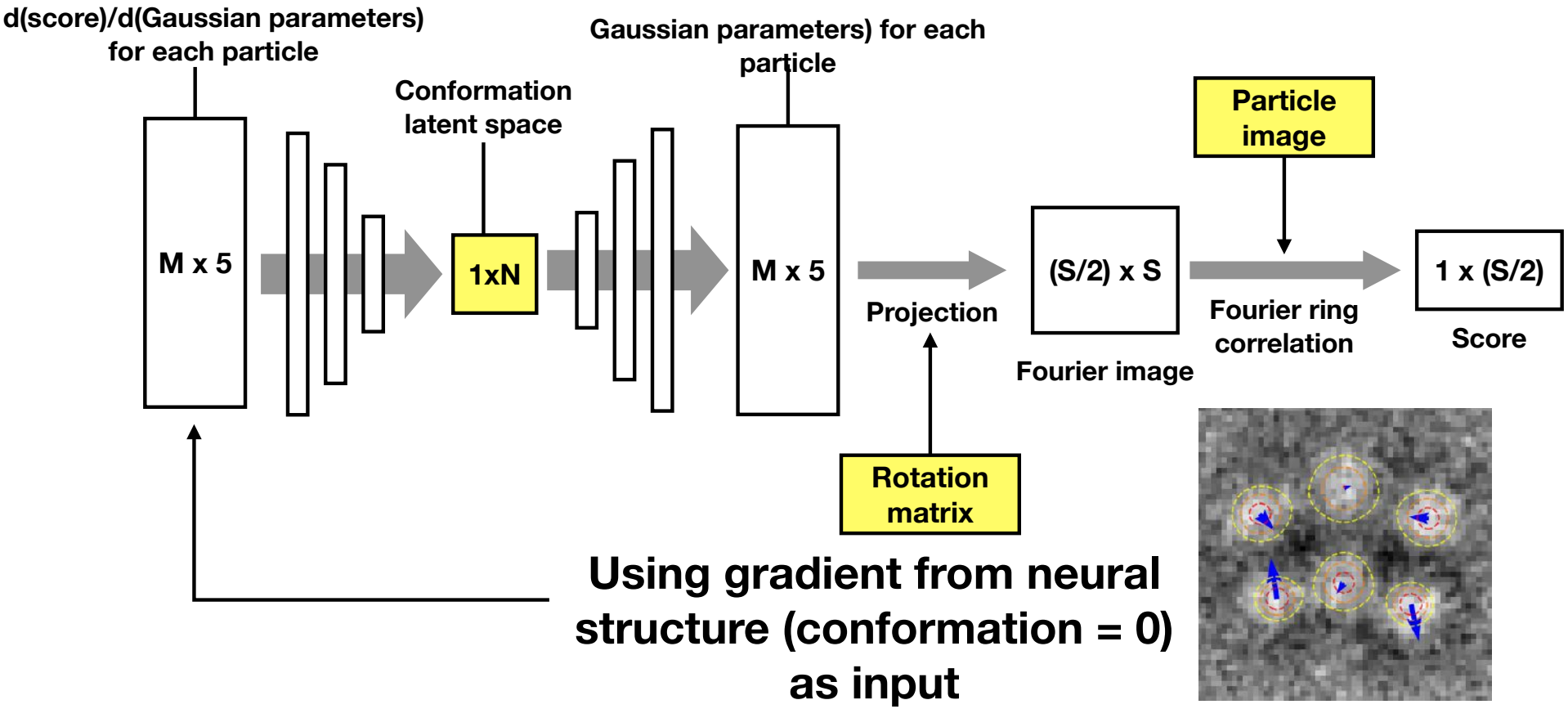




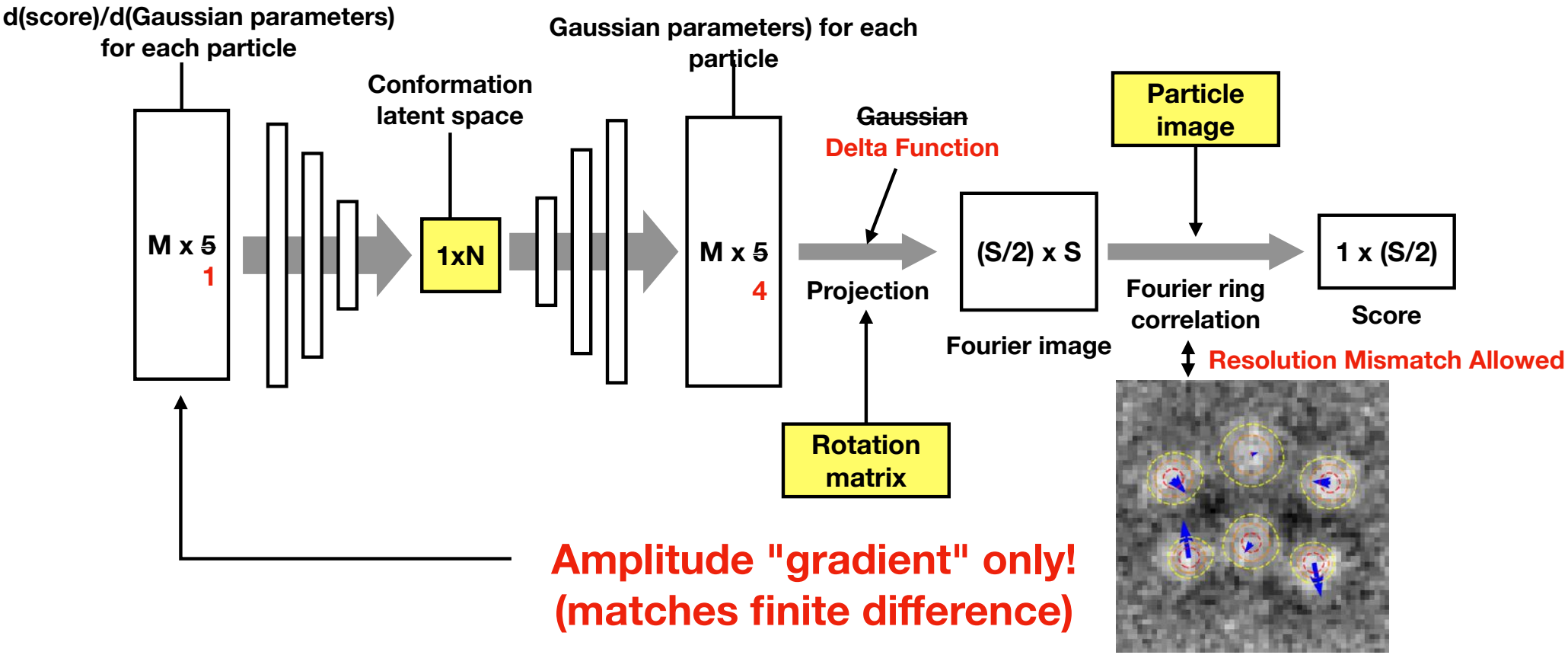


Particle-projection Fourier ring correlation (FRC)

Deep Learning



Deep Learning



IP3R Acknowledgements

Medical School UTHealth at Houston

Irina Serysheva's Lab:

Guizhen Fan
Mariah Baker
Alexander Seryshev
Raj Sharma
Venkata Mallampalli

Matthew Baker

Baylor College of Medicine (Houston)

Steven Ludtke's Lab:

Muyuan Chen

Zhao Wang

University of Rochester (Rochester, NY)

David Yule's Lab:

Lara Terry
Vikas Arige

KU Leuven (Leuven, Belgium)

Geert Bultynck's Lab:

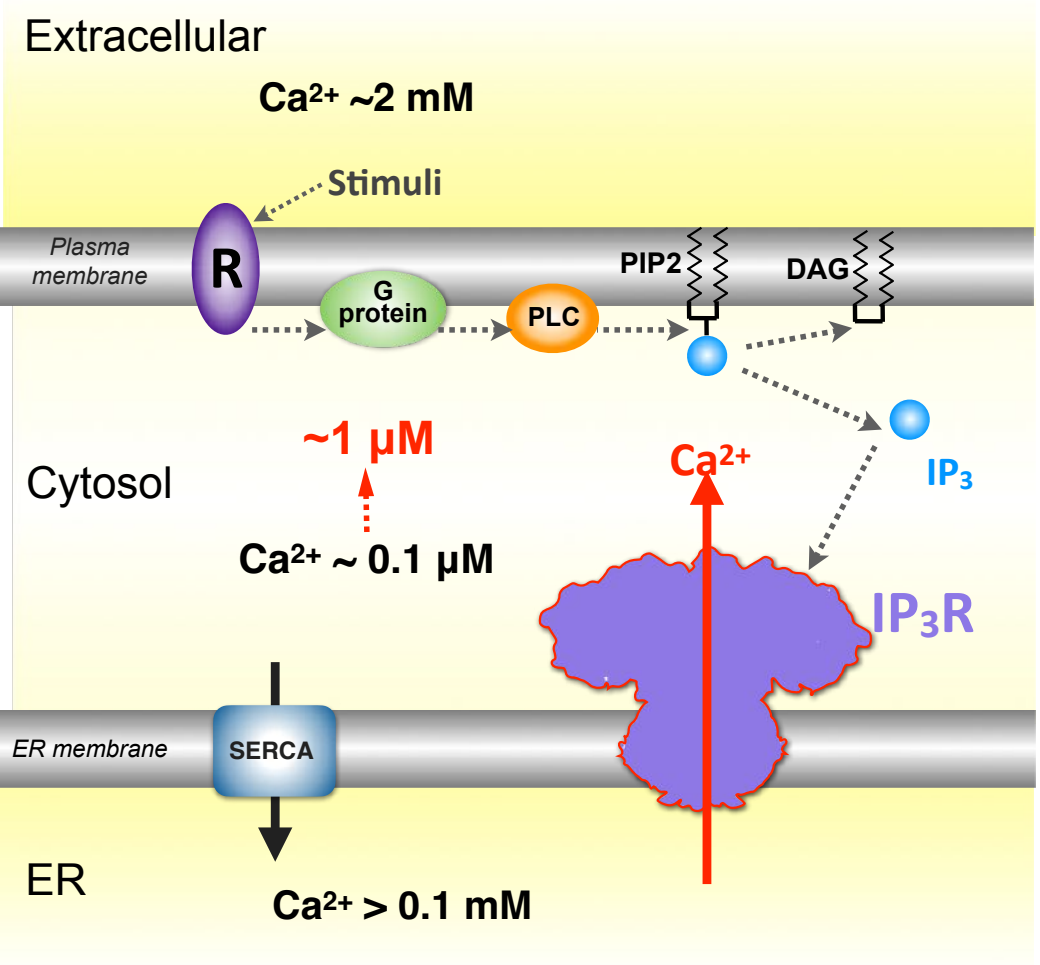
Hristina Ivanova
Ian de Ridder

UTHealth Cryo-EM Core Facility: uthealth.corefacilities.org

Supported by NIGMS, NINDS, NSF, CPRIT, AHA & Welch Foundation

Inositol 1,4,5 - Trisphosphate Receptors

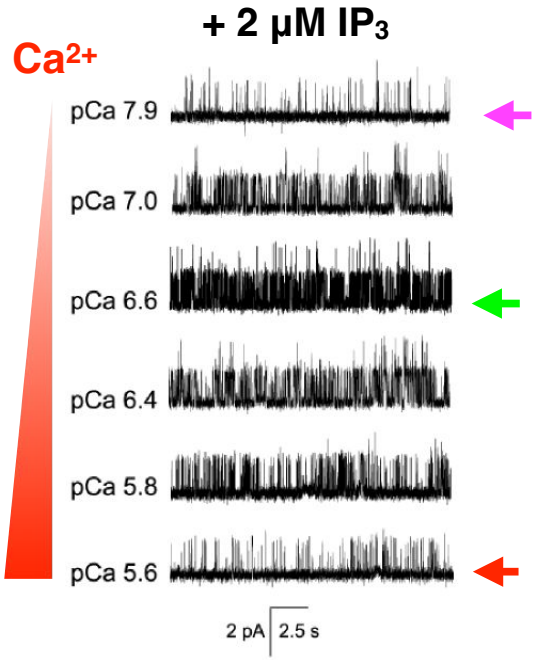
IP_3-gated Ca^{2+} Release Channels

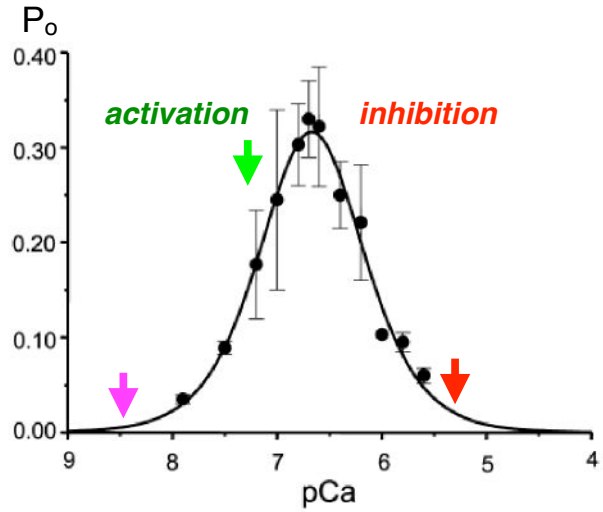


- Expressed in virtually **all eukaryotic cells**
- **Intracellular ion channels**
- Response to **many extracellular stimuli** (hormones, growth factors, neurotransmitters, neurotrophins, odorants, light, and etc.)
- Ligand-gated ion channels: **primary ligands - IP_3 & Ca^{2+}**
- Associate ***in vivo*** with **multiple modulatory proteins** (>100)

Slide courtesy of Irina Serysheva

Ca^{2+}- Dependent Activation of IP_3R1

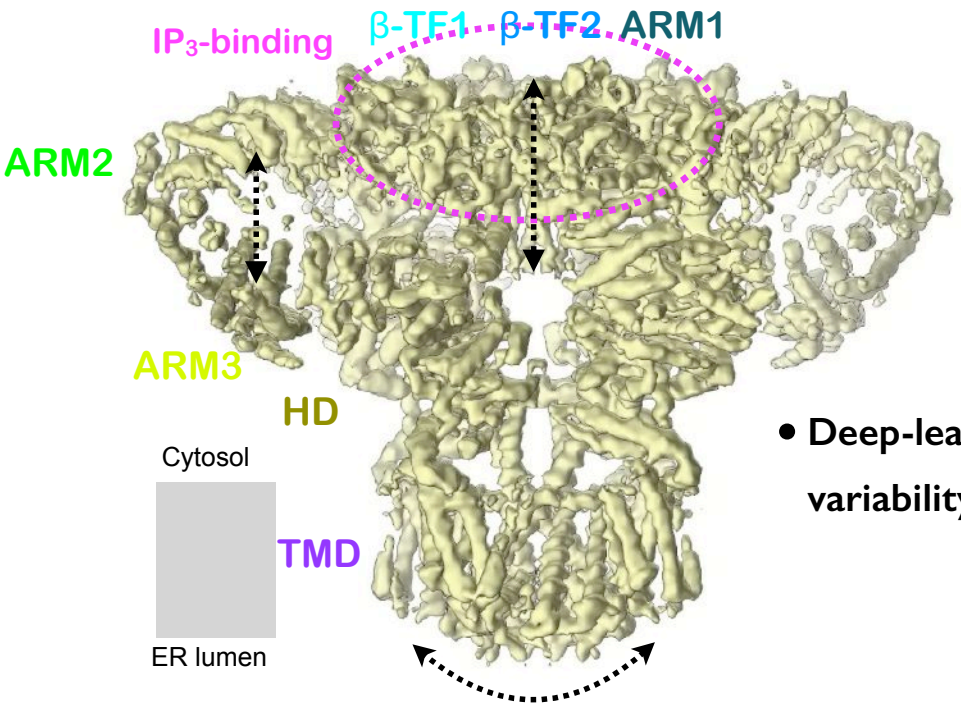


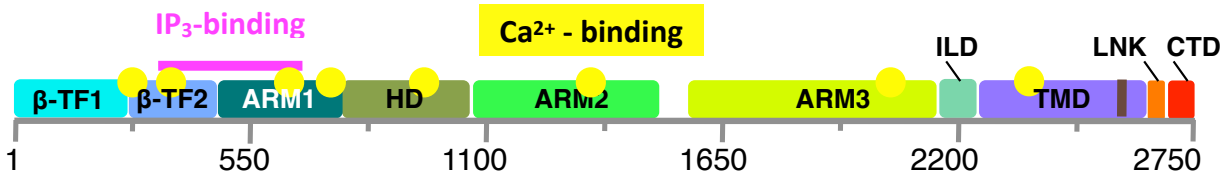


Bezprozvanny, Watras & Ehrlich, *Nature* 1991

Slide courtesy of Irina Serysheva

Intrinsically Dynamic IP_3R1 Structure





- Deep-learning based analysis of structural variability from 2D cryo-EM images!

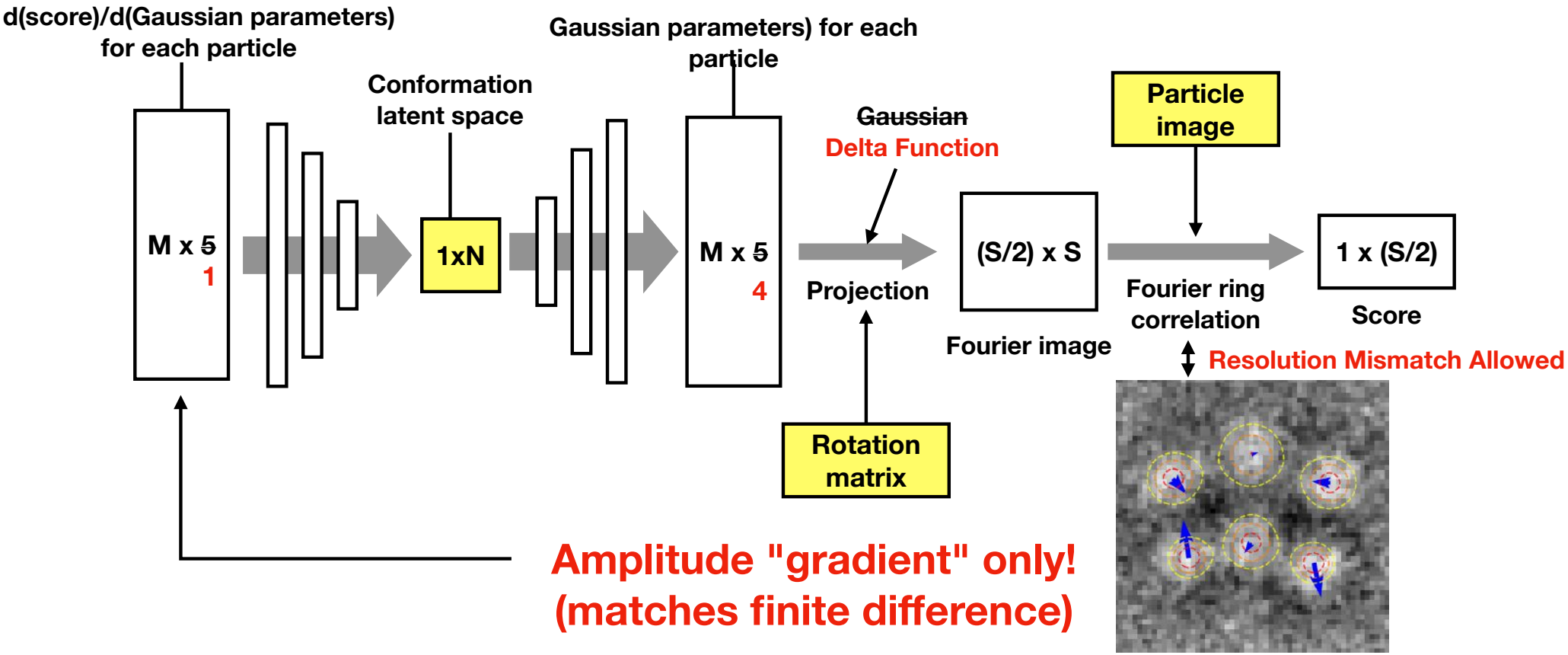
Particles

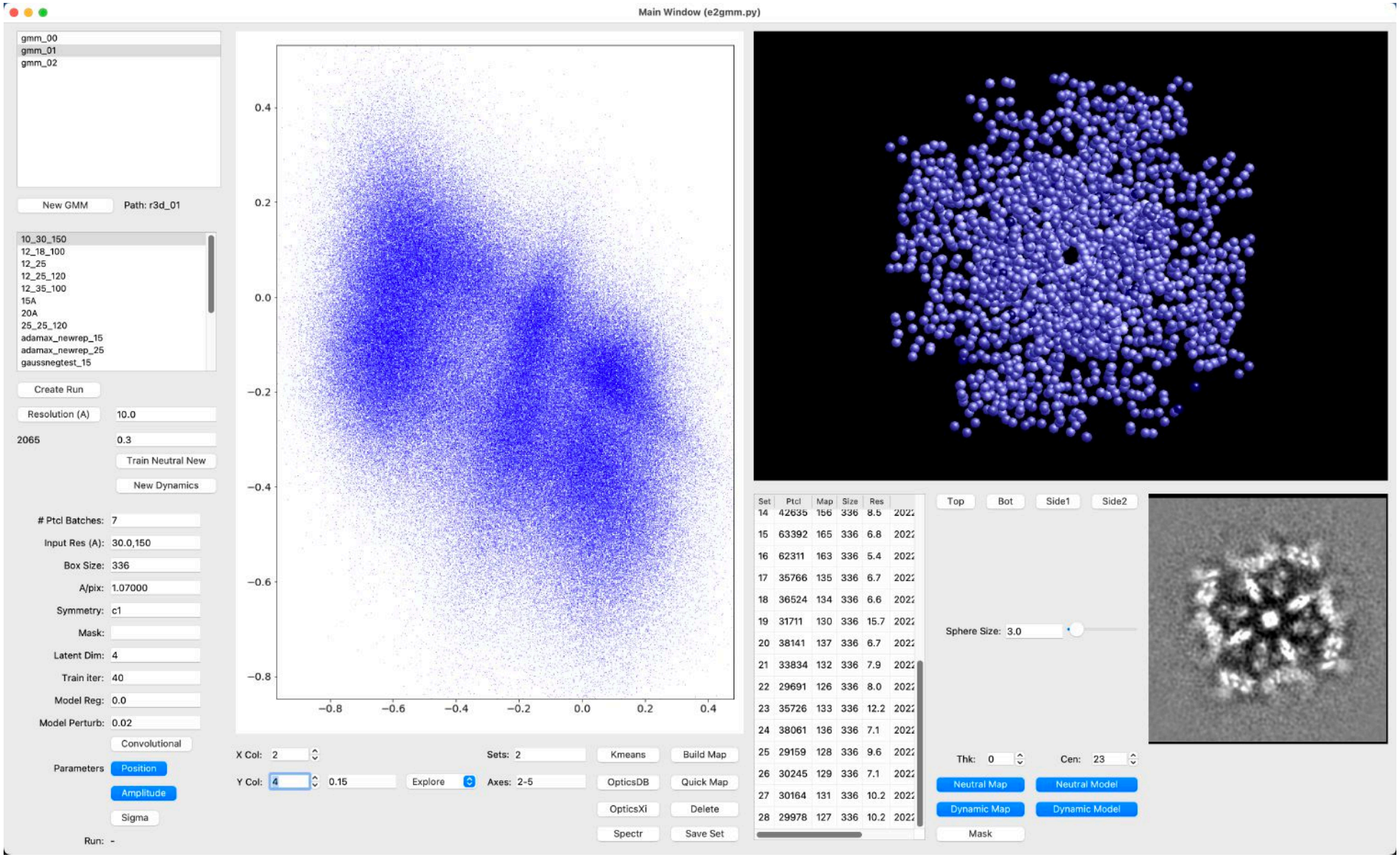
- 133k apo
- 133k Ca+IP3+ATP
- 133k high Ca

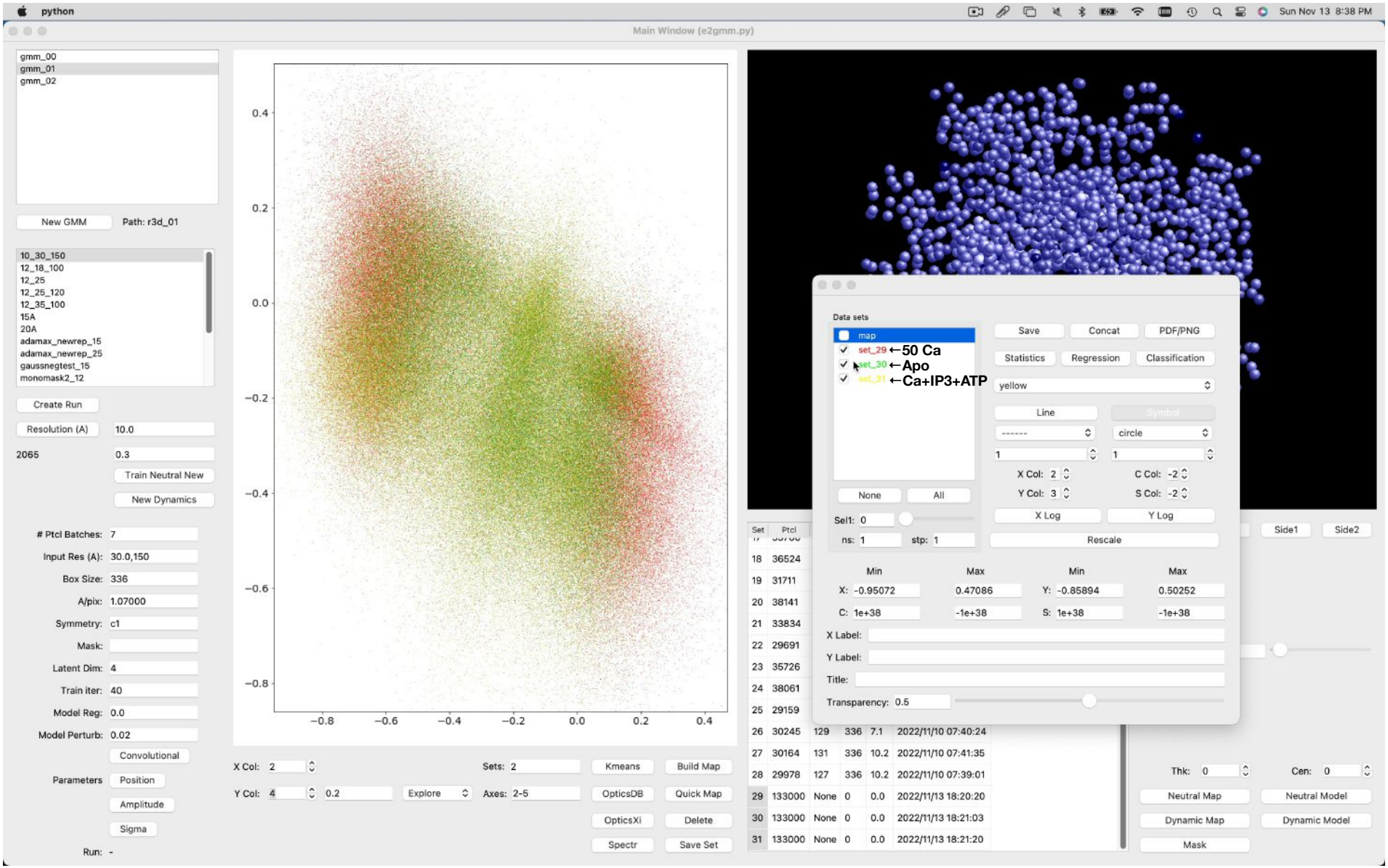
2022, under review

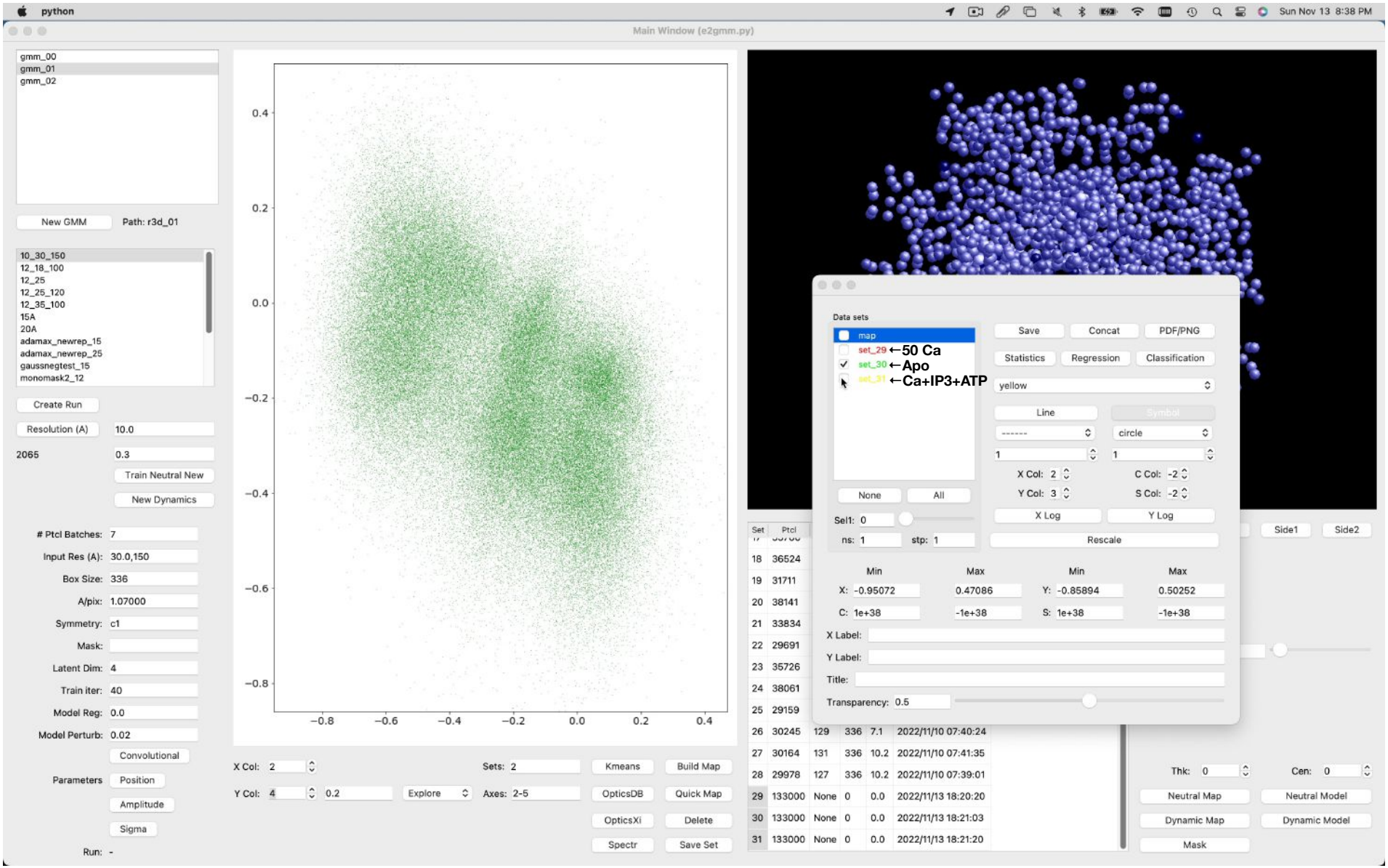
Slide courtesy of Irina Serysheva

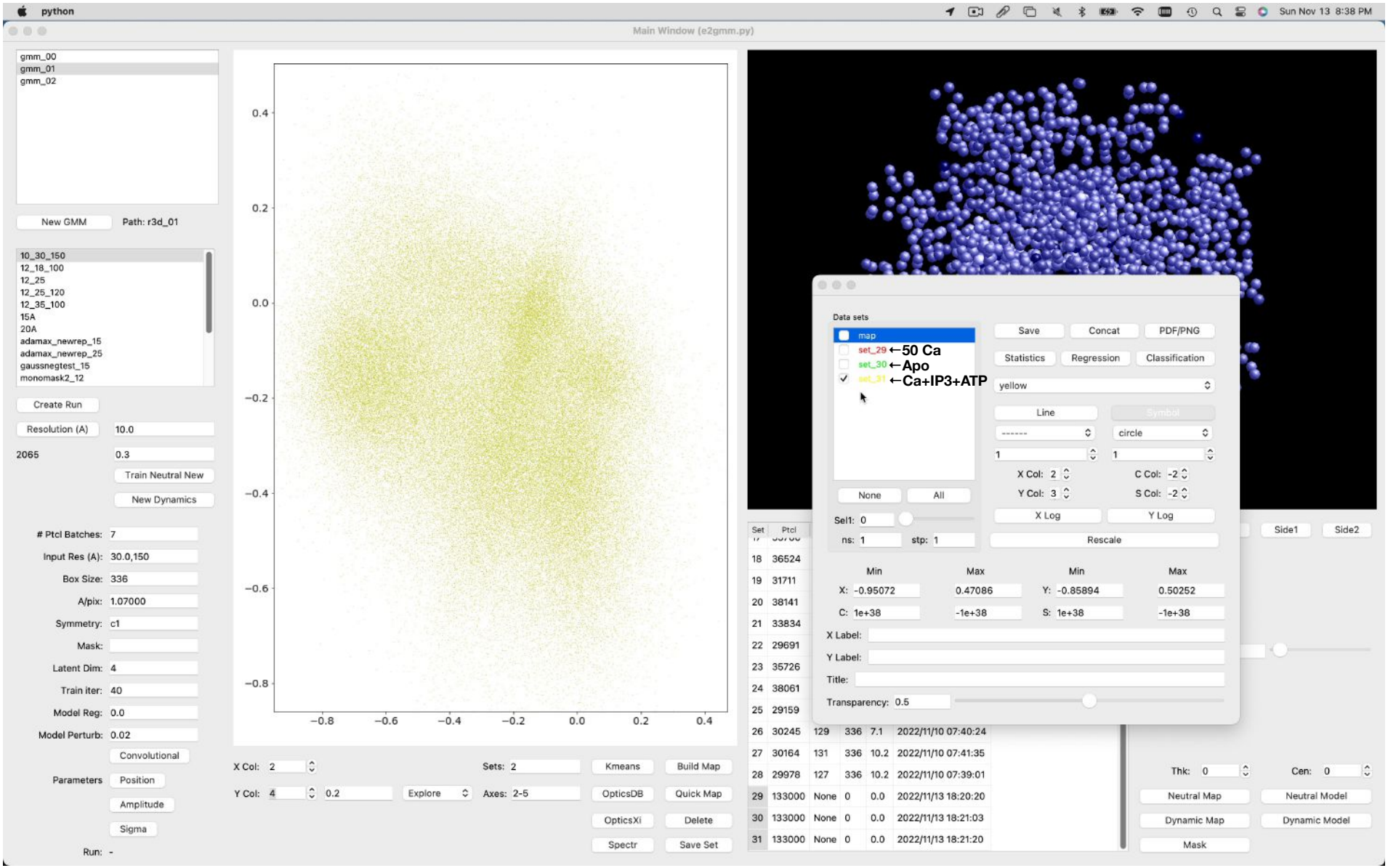
Deep Learning

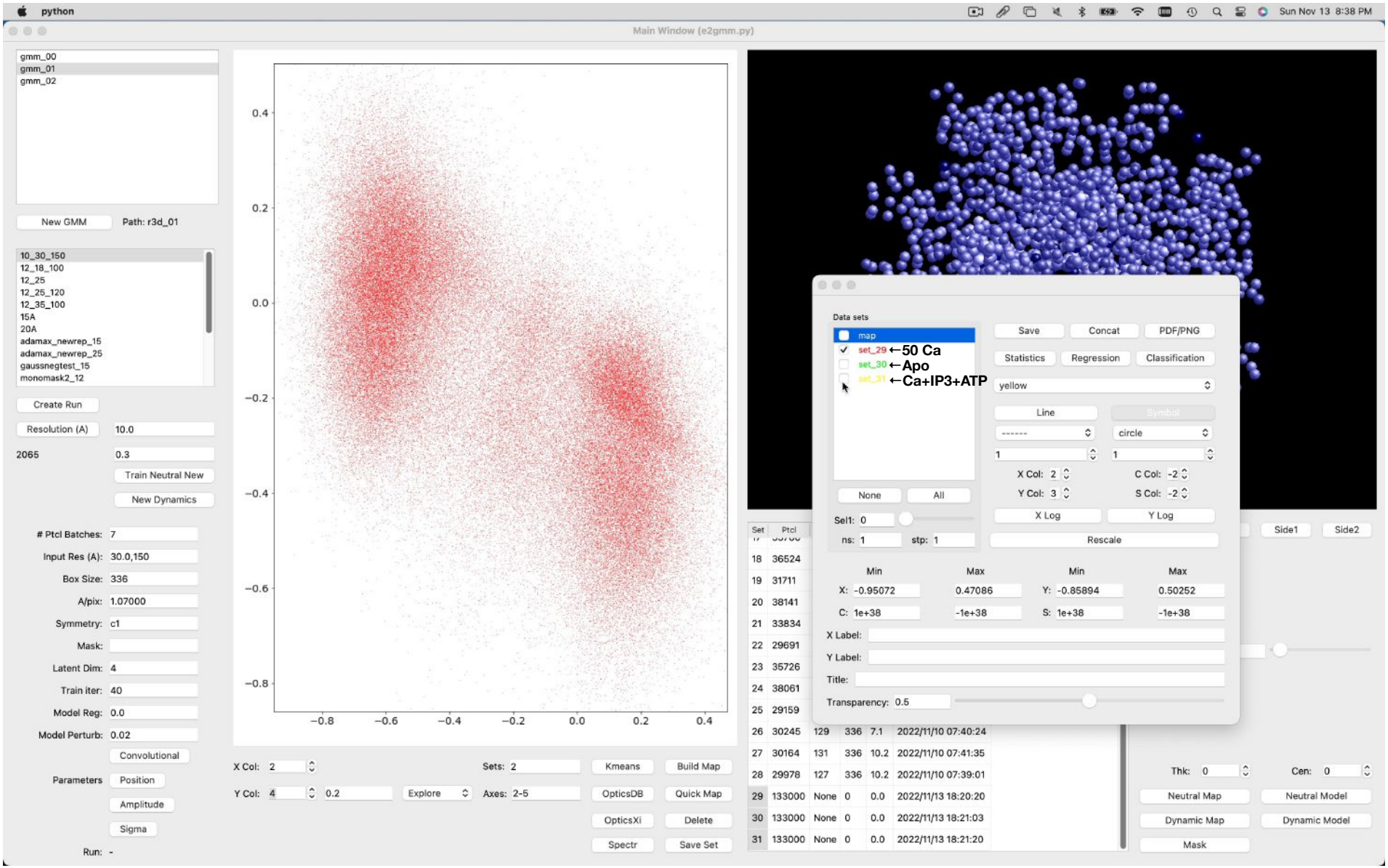


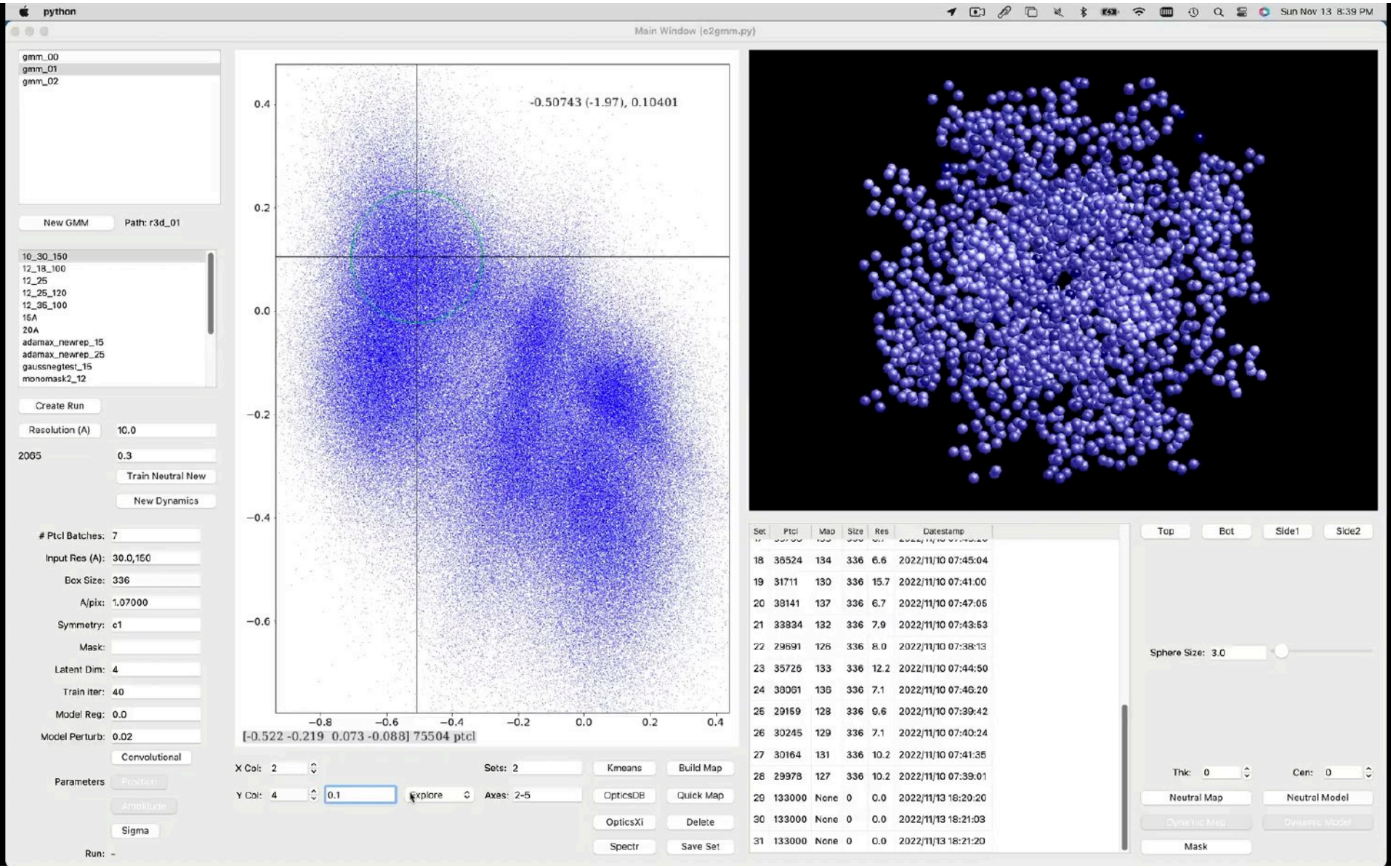

Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
25_25_120
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches:
7
Input Res (A):
30.0,150
Box Size:
336
A/pix:
1.07000
Symmetry:
c1
Mask:
Latent Dim:
4
Train iter:
40
Model Reg:
0.0
Model Perturb:
0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run: -
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
X Col: 2
Y Col: 4
0.15
Explore
Sets: 2
Axes: 2-5
Kmeans
Build Map
OpticsDB
Quick Map
OpticsXi
Delete
Spectr
Save Set
Set Ptcl Map Size Res
14 42635 156 336 8.5 2022
15 63392 165 336 6.8 2022
16 62311 163 336 5.4 2022
17 35766 135 336 6.7 2022
18 36524 134 336 6.6 2022
19 31711 130 336 15.7 2022
20 38141 137 336 6.7 2022
21 33834 132 336 7.9 2022
22 29691 126 336 8.0 2022
23 35726 133 336 12.2 2022
24 38061 136 336 7.1 2022
25 29159 128 336 9.6 2022
26 30245 129 336 7.1 2022
27 30164 131 336 10.2 2022
28 29978 127 336 10.2 2022
Top
Bot
Side1
Side2
Sphere Size: 3.0
Thk: 0
Cen: 23
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask

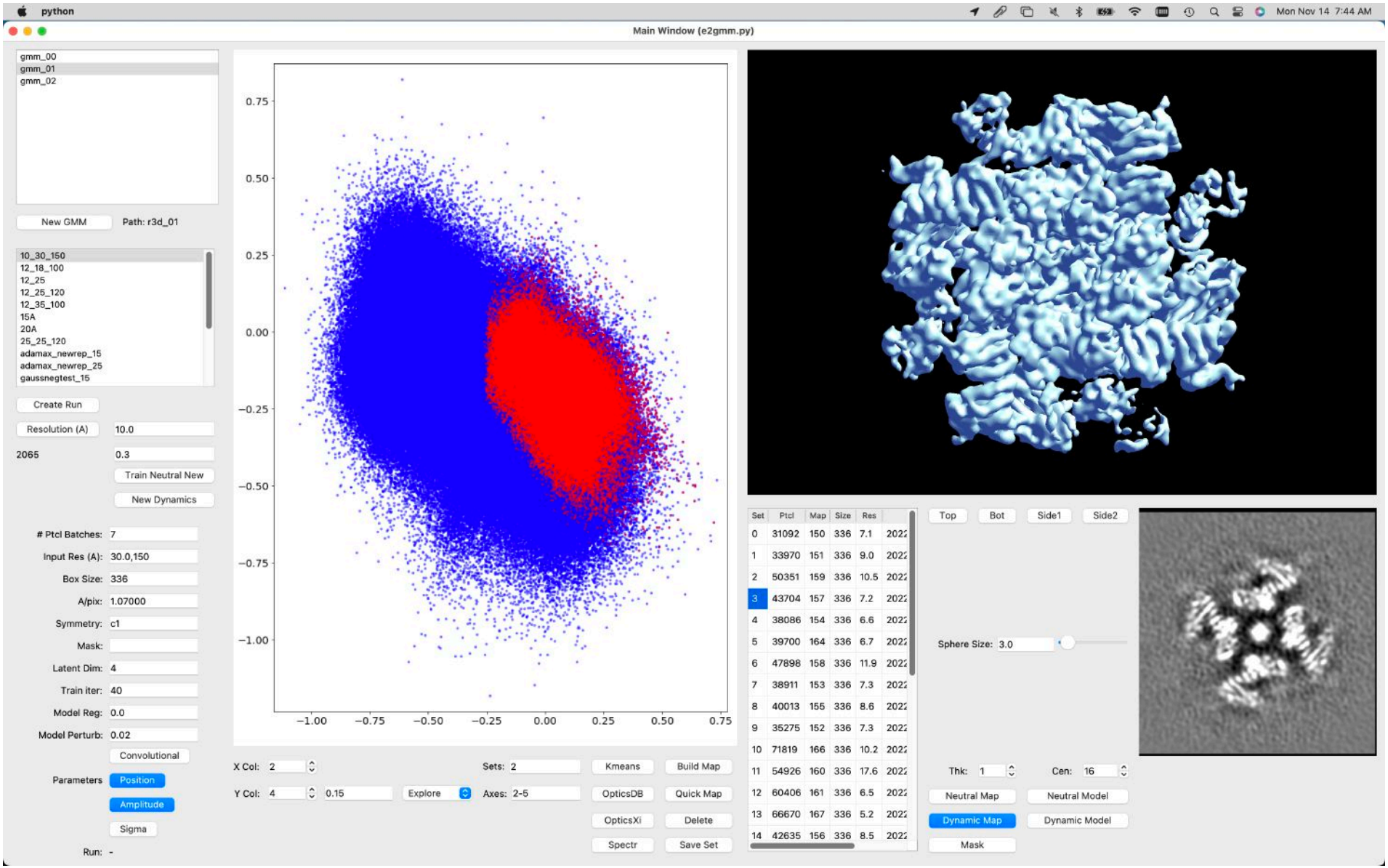

python
Sun Nov 13 8:38 PM
Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
monomask2_12
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches: 7
Input Res (A): 30.0,150
Box Size: 336
A/pix: 1.07000
Symmetry: c1
Mask:
Latent Dim: 4
Train iter: 40
Model Reg: 0.0
Model Perturb: 0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run: -
X Col: 2
Y Col: 4
0.2
Explore
Sets: 2
Axes: 2-5
Kmeans
Build Map
OpticsDB
Quick Map
OpticsXi
Delete
Spectr
Save Set
Data sets
map
set_29 ←50 Ca
set_30 ←Apo
set_31 ←Ca+IP3+ATP
Save
Concat
PDF/PNG
Statistics
Regression
Classification
yellow
Line
Symbol
circle
X Col: 2
Y Col: 3
C Col: -2
S Col: -2
None
All
Sel1: 0
ns: 1
stp: 1
X Log
Y Log
Rescale
Min
Max
X: -0.95072
0.47086
Y: -0.85894
0.50252
C: 1e+38
-1e+38
S: 1e+38
-1e+38
X Label:
Y Label:
Title:
Transparency: 0.5
Side1
Side2
Thk: 0
Cen: 0
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask
Set Ptcl
18 36524
19 31711
20 38141
21 33834
22 29691
23 35726
24 38061
25 29159
26 30245 129 336 7.1 2022/11/10 07:40:24
27 30164 131 336 10.2 2022/11/10 07:41:35
28 29978 127 336 10.2 2022/11/10 07:39:01
29 133000 None 0 0.0 2022/11/13 18:20:20
30 133000 None 0 0.0 2022/11/13 18:21:03
31 133000 None 0 0.0 2022/11/13 18:21:20

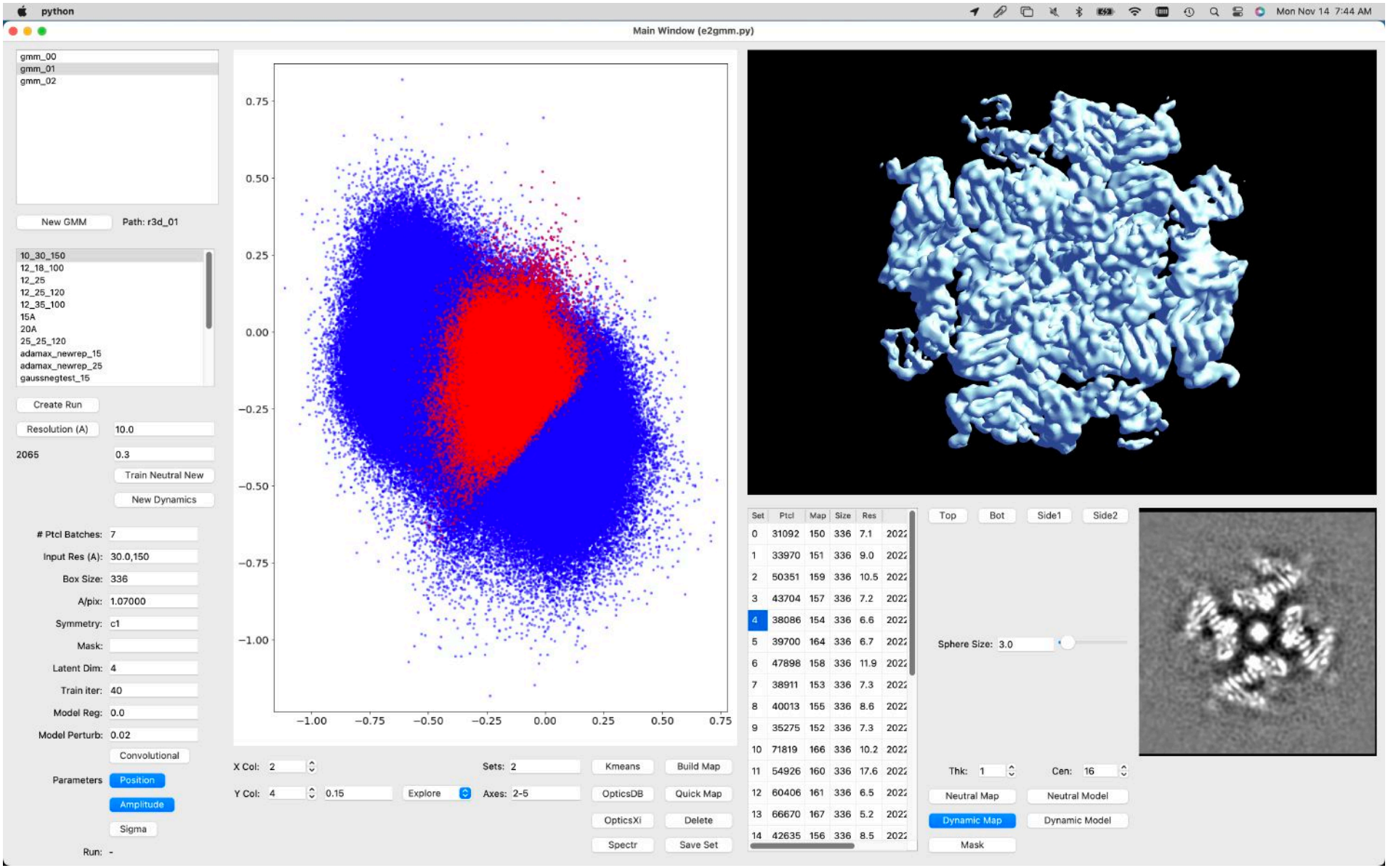

python
Sun Nov 13 8:38 PM
Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
monomask2_12
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches: 7
Input Res (A): 30.0,150
Box Size: 336
A/pix: 1.07000
Symmetry: c1
Mask:
Latent Dim: 4
Train iter: 40
Model Reg: 0.0
Model Perturb: 0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run: -
X Col: 2
Y Col: 4
0.2
Explore
Sets: 2
Axes: 2-5
Kmeans
OpticsDB
OpticsXi
Spectr
Build Map
Quick Map
Delete
Save Set
Data sets
map
set_29 ←50 Ca
set_30 ←Apo
set_31 ←Ca+IP3+ATP
None
All
Sel1: 0
ns: 1
stp: 1
Save
Concat
PDF/PNG
Statistics
Regression
Classification
yellow
Line
Symbol
circle
X Col: 2
Y Col: 3
C Col: -2
S Col: -2
X Log
Y Log
Rescale
Min
Max
X: -0.95072
0.47086
C: 1e+38
-1e+38
Y: -0.85894
0.50252
S: 1e+38
-1e+38
X Label:
Y Label:
Title:
Transparency: 0.5
Side1
Side2
Thk: 0
Cen: 0
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask
Set Ptcl
18 36524
19 31711
20 38141
21 33834
22 29691
23 35726
24 38061
25 29159
26 30245 129 336 7.1 2022/11/10 07:40:24
27 30164 131 336 10.2 2022/11/10 07:41:35
28 29978 127 336 10.2 2022/11/10 07:39:01
29 133000 None 0 0.0 2022/11/13 18:20:20
30 133000 None 0 0.0 2022/11/13 18:21:03
31 133000 None 0 0.0 2022/11/13 18:21:20


python
Sun Nov 13 8:38 PM
Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
monomask2_12
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches: 7
Input Res (A): 30.0,150
Box Size: 336
A/pix: 1.07000
Symmetry: c1
Mask:
Latent Dim: 4
Train iter: 40
Model Reg: 0.0
Model Perturb: 0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run: -
X Col: 2
Y Col: 4
0.2
Explore
Sets: 2
Axes: 2-5
Kmeans
OpticsDB
OpticsXi
Spectr
Build Map
Quick Map
Delete
Save Set
Data sets
map
set_29 ←50 Ca
set_30 ←Apo
set_31 ←Ca+IP3+ATP
None
All
Sel1: 0
ns: 1
stp: 1
Save
Concat
PDF/PNG
Statistics
Regression
Classification
yellow
Line
Symbol
circle
X Col: 2
Y Col: 3
C Col: -2
S Col: -2
X Log
Y Log
Rescale
Min
Max
X: -0.95072
0.47086
C: 1e+38
-1e+38
Y: -0.85894
0.50252
S: 1e+38
-1e+38
X Label:
Y Label:
Title:
Transparency: 0.5
Side1
Side2
Thk: 0
Cen: 0
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask
Set Ptcl
18 36524
19 31711
20 38141
21 33834
22 29691
23 35726
24 38061
25 29159
26 30245 129 336 7.1 2022/11/10 07:40:24
27 30164 131 336 10.2 2022/11/10 07:41:35
28 29978 127 336 10.2 2022/11/10 07:39:01
29 133000 None 0 0.0 2022/11/13 18:20:20
30 133000 None 0 0.0 2022/11/13 18:21:03
31 133000 None 0 0.0 2022/11/13 18:21:20


python
Sun Nov 13 8:38 PM
Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
monomask2_12
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches: 7
Input Res (A): 30.0,150
Box Size: 336
A/pix: 1.07000
Symmetry: c1
Mask:
Latent Dim: 4
Train iter: 40
Model Reg: 0.0
Model Perturb: 0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run: -
X Col: 2
Y Col: 4
0.2
Explore
Sets: 2
Axes: 2-5
Kmeans
OpticsDB
OpticsXi
Spectr
Build Map
Quick Map
Delete
Save Set
Data sets
map
set_29 ←50 Ca
set_30 ←Apo
set_31 ←Ca+IP3+ATP
None
All
Sel1: 0
ns: 1
stp: 1
Save
Concat
PDF/PNG
Statistics
Regression
Classification
yellow
Line
Symbol
circle
X Col: 2
Y Col: 3
C Col: -2
S Col: -2
X Log
Y Log
Rescale
Min
Max
X: -0.95072
0.47086
C: 1e+38
-1e+38
Y: -0.85894
0.50252
S: 1e+38
-1e+38
X Label:
Y Label:
Title:
Transparency: 0.5
Side1
Side2
Thk: 0
Cen: 0
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask
Set Ptcl
18 36524
19 31711
20 38141
21 33834
22 29691
23 35726
24 38061
25 29159
26 30245 129 336 7.1 2022/11/10 07:40:24
27 30164 131 336 10.2 2022/11/10 07:41:35
28 29978 127 336 10.2 2022/11/10 07:39:01
29 133000 None 0 0.0 2022/11/13 18:20:20
30 133000 None 0 0.0 2022/11/13 18:21:03
31 133000 None 0 0.0 2022/11/13 18:21:20

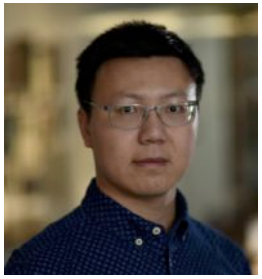

python
Sun Nov 13 8:39 PM
Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
monomask2_12
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches: 7
Input Res (A): 30.0,150
Box Size: 336
A/pix: 1.07000
Symmetry: c1
Mask:
Latent Dim: 4
Train iter: 40
Model Reg: 0.0
Model Perturb: 0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run: -
-0.50743 (-1.97), 0.10401
[-0.522 -0.219 0.073 -0.088] 75504 ptcl
X Col: 2
Y Col: 4
0.1
Explore
Sets: 2
Axes: 2-5
Kmeans
OpticsDB
OpticsXi
Spectr
Build Map
Quick Map
Delete
Save Set
Set Ptcl Map Size Res Datestamp
18 36524 134 336 6.6 2022/11/10 07:45:04
19 31711 130 336 15.7 2022/11/10 07:41:00
20 38141 137 336 6.7 2022/11/10 07:47:05
21 33834 132 336 7.9 2022/11/10 07:43:53
22 29691 126 336 8.0 2022/11/10 07:38:13
23 35726 133 336 12.2 2022/11/10 07:44:50
24 38061 136 336 7.1 2022/11/10 07:46:20
25 29159 128 336 9.6 2022/11/10 07:39:42
26 30245 129 336 7.1 2022/11/10 07:40:24
27 30164 131 336 10.2 2022/11/10 07:41:35
28 29978 127 336 10.2 2022/11/10 07:39:01
29 133000 None 0 0.0 2022/11/13 18:20:20
30 133000 None 0 0.0 2022/11/13 18:21:03
31 133000 None 0 0.0 2022/11/13 18:21:20
Top
Bot
Side1
Side2
Sphere Size: 3.0
Thk: 0
Cen: 0
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask


python
Mon Nov 14 7:44 AM
Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
25_25_120
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches:
7
Input Res (A):
30.0,150
Box Size:
336
A/pix:
1.07000
Symmetry:
c1
Mask:
Latent Dim:
4
Train iter:
40
Model Reg:
0.0
Model Perturb:
0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run:
-
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
X Col:
2
Y Col:
4
0.15
Explore
Sets:
2
Axes:
2-5
Kmeans
OpticsDB
OpticsXi
Spectr
Build Map
Quick Map
Delete
Save Set
Set
Ptcl
Map
Size
Res
0 31092 150 336 7.1 2022
1 33970 151 336 9.0 2022
2 50351 159 336 10.5 2022
3 43704 157 336 7.2 2022
4 38086 154 336 6.6 2022
5 39700 164 336 6.7 2022
6 47898 158 336 11.9 2022
7 38911 153 336 7.3 2022
8 40013 155 336 8.6 2022
9 35275 152 336 7.3 2022
10 71819 166 336 10.2 2022
11 54926 160 336 17.6 2022
12 60406 161 336 6.5 2022
13 66670 167 336 5.2 2022
14 42635 156 336 8.5 2022
Top
Bot
Side1
Side2
Sphere Size:
3.0
Thk:
1
Cen:
16
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask


python
Mon Nov 14 7:44 AM
Main Window (e2gmm.py)
gmm_00
gmm_01
gmm_02
New GMM
Path: r3d_01
10_30_150
12_18_100
12_25
12_25_120
12_35_100
15A
20A
25_25_120
adamax_newrep_15
adamax_newrep_25
gaussnegtest_15
Create Run
Resolution (A)
10.0
2065
0.3
Train Neutral New
New Dynamics
Ptcl Batches:
7
Input Res (A):
30.0,150
Box Size:
336
A/pix:
1.07000
Symmetry:
c1
Mask:
Latent Dim:
4
Train iter:
40
Model Reg:
0.0
Model Perturb:
0.02
Convolutional
Parameters
Position
Amplitude
Sigma
Run:
-
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
X Col:
2
Y Col:
4
0.15
Explore
Sets:
2
Axes:
2-5
Kmeans
OpticsDB
OpticsXi
Spectr
Build Map
Quick Map
Delete
Save Set
Set
Ptcl
Map
Size
Res
0 31092 150 336 7.1 2022
1 33970 151 336 9.0 2022
2 50351 159 336 10.5 2022
3 43704 157 336 7.2 2022
4 38086 154 336 6.6 2022
5 39700 164 336 6.7 2022
6 47898 158 336 11.9 2022
7 38911 153 336 7.3 2022
8 40013 155 336 8.6 2022
9 35275 152 336 7.3 2022
10 71819 166 336 10.2 2022
11 54926 160 336 17.6 2022
12 60406 161 336 6.5 2022
13 66670 167 336 5.2 2022
14 42635 156 336 8.5 2022
Top
Bot
Side1
Side2
Sphere Size:
3.0
Thk:
1
Cen:
16
Neutral Map
Neutral Model
Dynamic Map
Dynamic Model
Mask

Acknowledgements

Steve Ludtke



Zhao Wang

Beckman Foundation: 2021FIB-41

CPRIT Core Award: RP190602

NIH: R01GM080139

Core Staff

Sneka Raveendran

Isaac Forrester

Sana Qureshi

Robyn Leidel

also:
Erik Anderson
Phil Baldwin
Zhili Yu
Adam Fluty