NCI-DOE Cancer Initiative: Ras Biology in Membranes

Towards Predictive Biology Through High Performance Computing

Frederick Streitz
Lawrence Livermore National Laboratory
Dwight Nissley
Frederick National Laboratory

Frederick National Laboratory Advisory Council

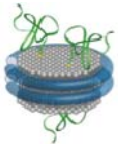
Lawrence Livermore National Laboratory

May 8, 2017

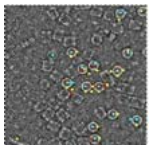
This work was performed under the auspices of the U.S. Department of Energy by Argonne National Laboratory under Contract DE-AC02-06-CH11357, Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344, Los Alamos National Laboratory under Contract DE-AC5206NA25396, and Oak Ridge National Laboratory under Contract DE-AC05-00OR22725.

Cancer Moonshot Pilot 2

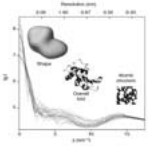
RAS activation experiments (FNLCR)



Experiments on nanodisc



CryoEM imaging



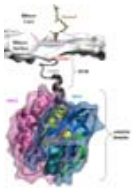
X-ray/neutron scattering

Multi-modal experimental data, image reconstruction, analytics

Protein structure databases

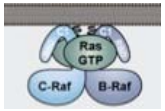
Predictive simulation and analysis of RAS

Phase Field model of lipid membrane



Granular RAS membrane interaction simulations

Atomic resolution RAS-RAF interaction

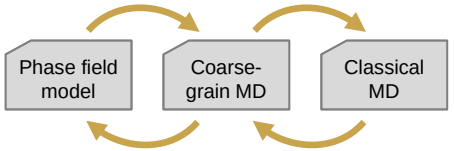


RAS Activation

Adaptive sampling molecular dynamics simulation codes

Adaptive time stepping

Adaptive spatial resolution



High-fidelity subgrid modeling

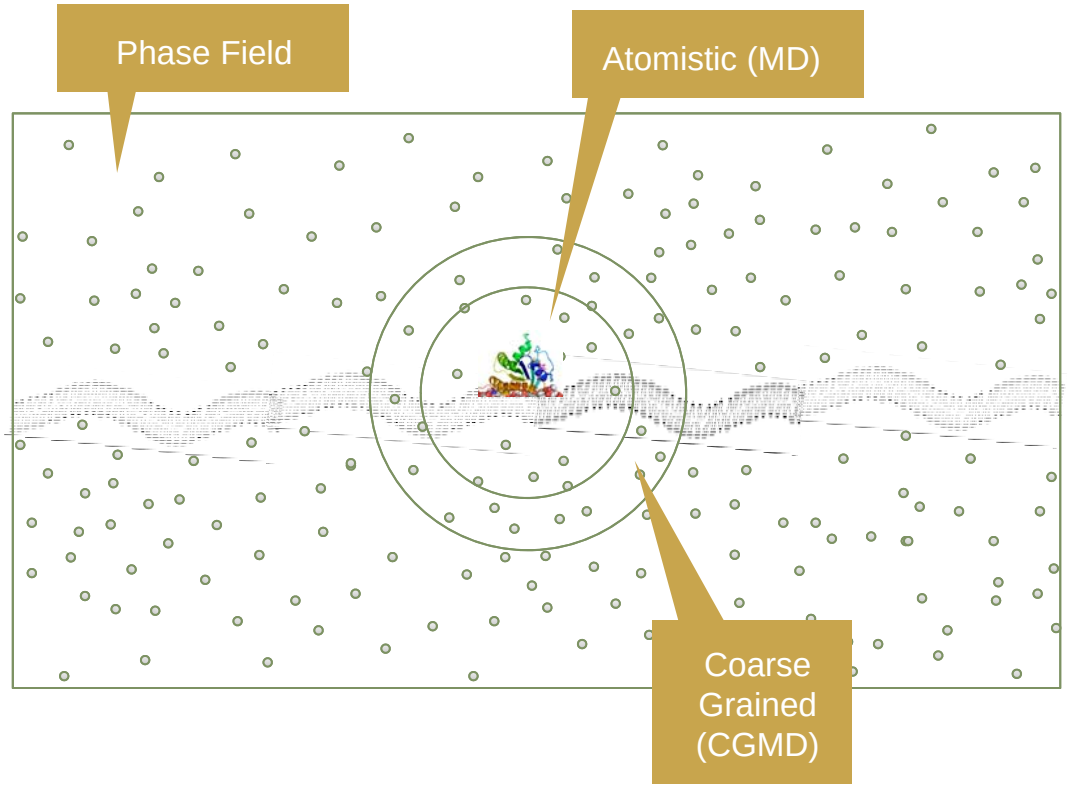
Machine learning guided dynamic validation

Unsupervised deep feature learning

Mechanistic network models

Uncertainty quantification

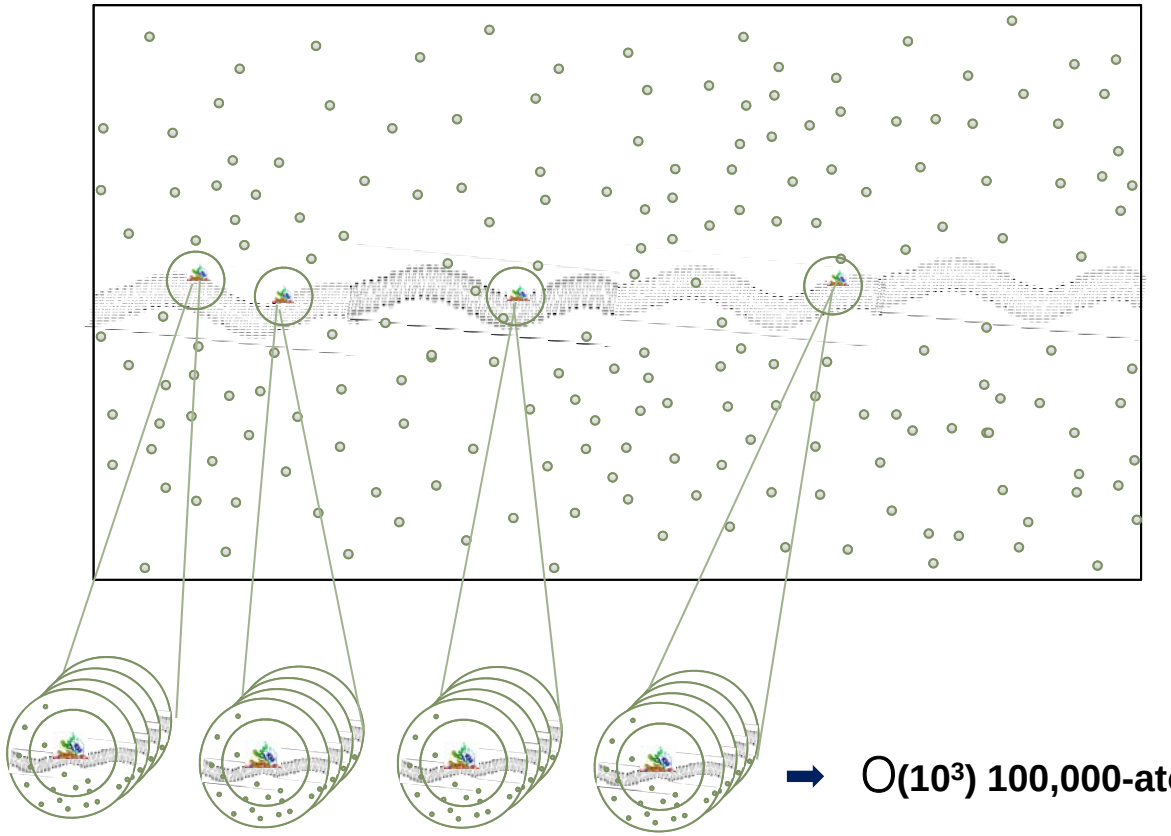
Adaptive resolution MD/CGMD coupled with phase field



- Model complex (many lipid) bilayer with phase field to capture structure and topology
- Model Ras on membrane using full atomistic resolution
- Use CGMD as ”glue” to connect different models

Connecting MD and CGMD with continuum-scale phase field models will access biologically relevant time and length scales

Simulation of full system will incorporate a large number of smaller simulations



- 10-100 μm lipid patches
- Dynamic membrane
- Hundreds of Ras proteins
  - Mutant and wild-type
  - Many conformations
  - Many environments

Investigate diffusion and aggregation in of Ras in context of specific membrane properties

➡ **O(10³) 100,000-atom simulations**

Close collaboration of experimentalists and theorists to build predictive model

Simulations to build model | **Experimental data to inform model**

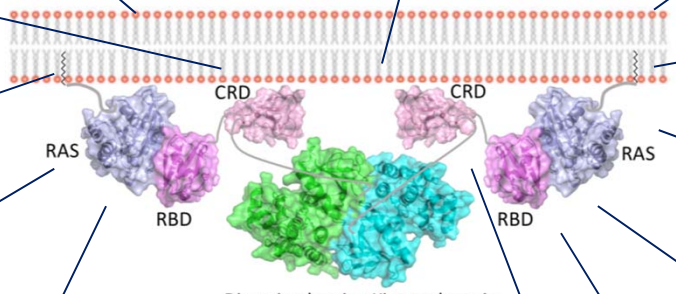
Lipid content: RAS/HVR binding by SPR, alpha assays in nanodiscs, liposomes, imaging in GVUs, lipidomics, SANS (possibly with contrast variation)

Lipid domains: Confocal microscopy RAS/HVR localization in GUVs, Calibrate coarse-grained (CG) simulations with all-atom (AA) Simulations, Calculate free energies of domains

RAS/HVR-membrane binding: SPR in liposomes, biophysical measurements, SANS (with contrast variation), AA and free-energy calculations of RAS/HVR binding to constrain CG parameters, free energies to inform phase field

RAF-membrane affinity: SPR in liposomes, biophysical measurements, MD simulations to identify regions of interest that interact with membrane

HVR structure/dynamics: crystallization, CD, MD of HVR in multi-component lipid platform to inform mobility in phase field model



Farnesyl dynamics: *solid state NMR,* AA and CG simulations of farnesyl in membranes and lipid bilayers informs phase field model

RAS/HVR multimeric state: BRET, step photobleaching, PALM, AA and CG MD of KRAS/HV R on nanodisc and multi-component lipid platform

RAS/HVR mobility & dynamics: single particle tracking, FCS, CG simulations of farnesylated HVR and RAS on nanodiscs and membranes, use to constrain phase field coupling

RAS-RBD binding: SPR, ITC, alpha assays in nanodiscs, TIRF, SANS (possibly with contrast variation), compare with AA simulations and constrain CG simualtions

RAS activity & structure: GTPase, GTP off-rate, crystallization, NMR, cryo-EM?, SANS, AA MD simulations constrain CG parameters

RAS-RBD structure: crystallization, NMR, AA simulations to constrain CG parameters

RAF activation: *dimerization, phosphorylation state(s),* long time-scale CG simulations and kinetic estimation, multi-scale simulations multi-scale simulations of RAS/RAF dynamics on membrane

RBD-CRD and CRAF structure: crystallization, NMR, cryo-EM, CG simulations validated against AA simulations

Simulations of KRAS have started in more biologically relevant lipid environments

Completed coarse-grained (CG) simulations of

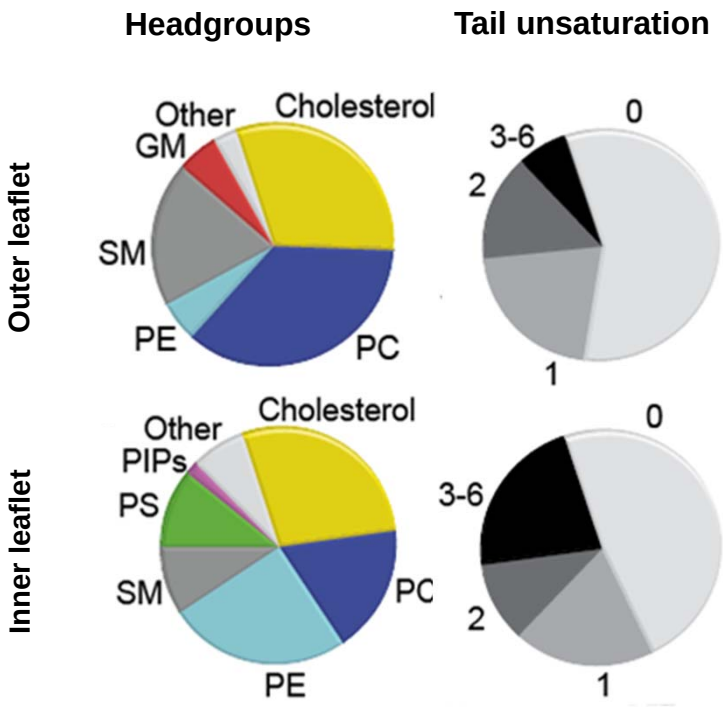
- Average mammalian plasma membrane with 63 distinct lipid types
- Working on improving CG parameters for specific lipid types to be consistent with all-atom (AA) simulations of lipids
- Investigating “simple” average plasma membrane [only 18 lipid types]
- Looking into tissue specific lipid compositions

Initial CGMD of KRAS proteins in complex human average plasma membrane

- 64 Kras4b in 70 nm x 70 nm membrane
- HVR in alpha helix conformation
- Inserted in inner plasma membrane leaflet

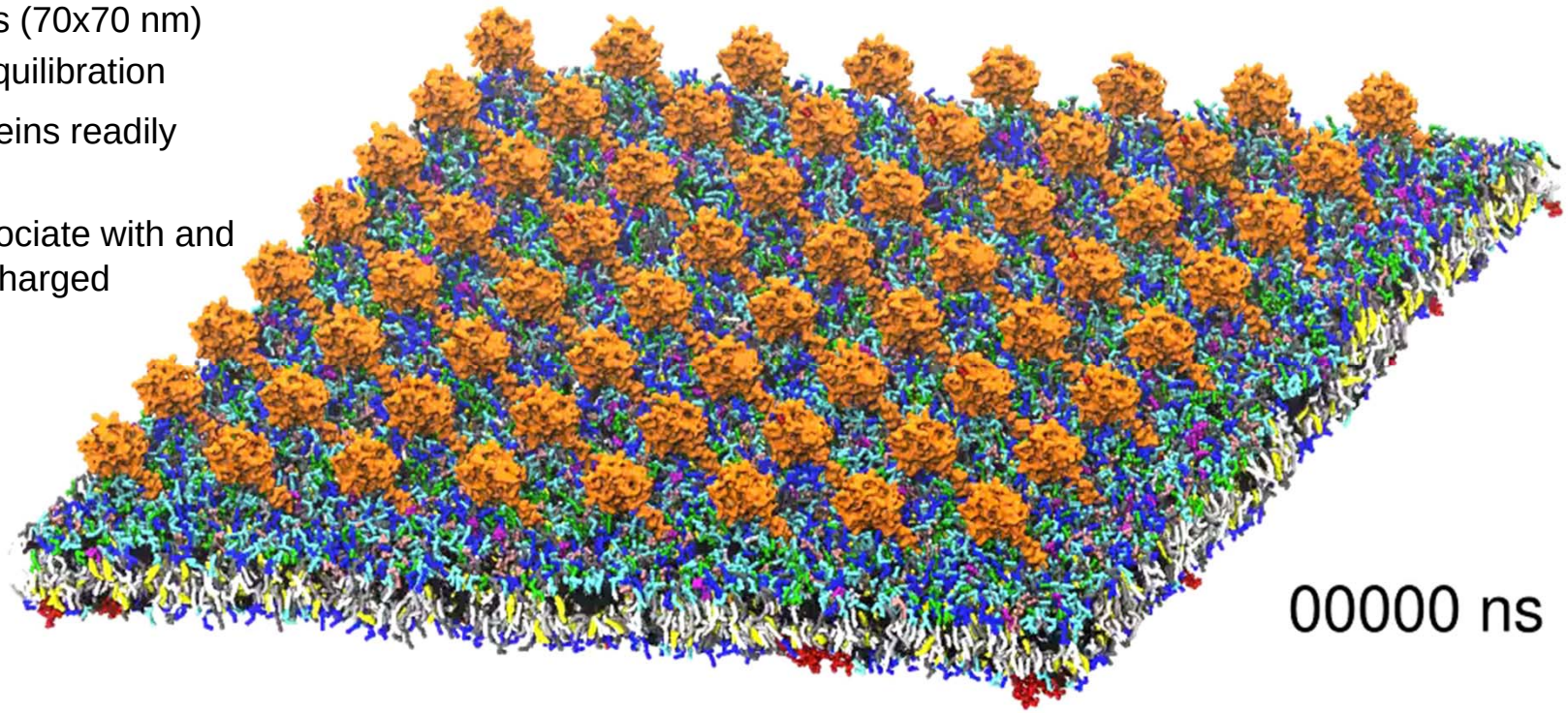
Ingólfsson H.I., M.N. Melo, F. van Eerden, C. Arnarez, C.A. Lopez, T.A. Wassenaar, X. Periole, A.H. de Vries, D.P. Tieleman and S.J. Marrink. 2014. Lipid organization of the plasma membrane. *J Am Chem Soc*, 136:14554-14559

Distribution of lipids in average plasma membrane



KRAS4b in mammalian plasma membrane

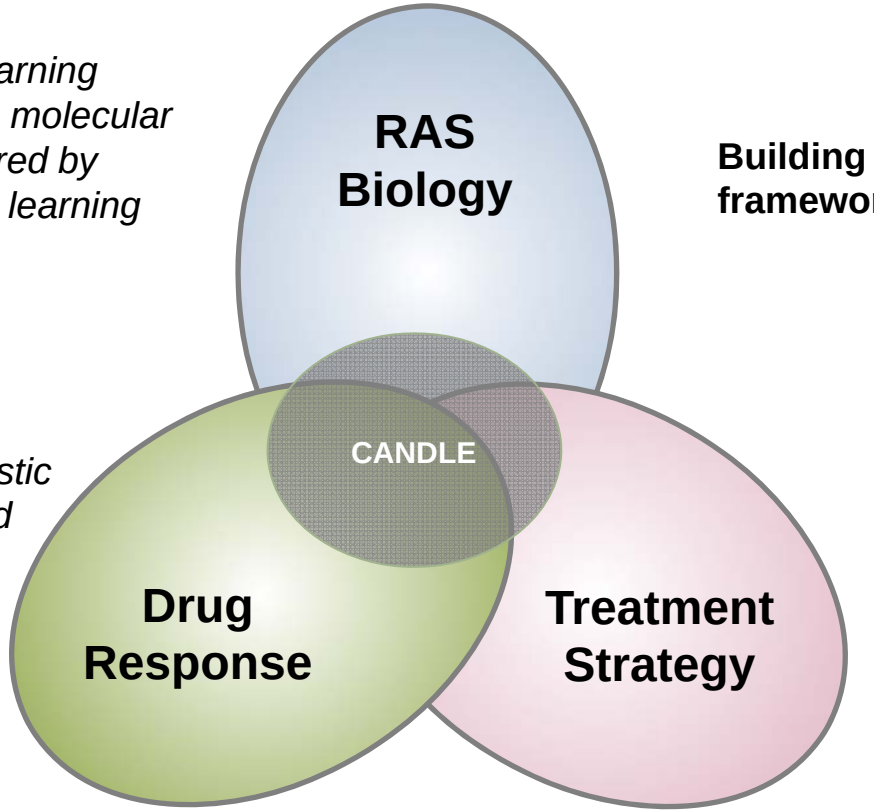
- 20,000 lipids (70x70 nm)
- 40 μs pre-equilibration
- 64 Ras proteins readily cluster
- Rapidly associate with and aggregate charged lipids in the membrane



Helgi Ingólfsson, LLNL

CANDLE: Cancer Distributed Learning Environment

Unsupervised learning inside multiscale molecular simulations steered by semi-supervised learning

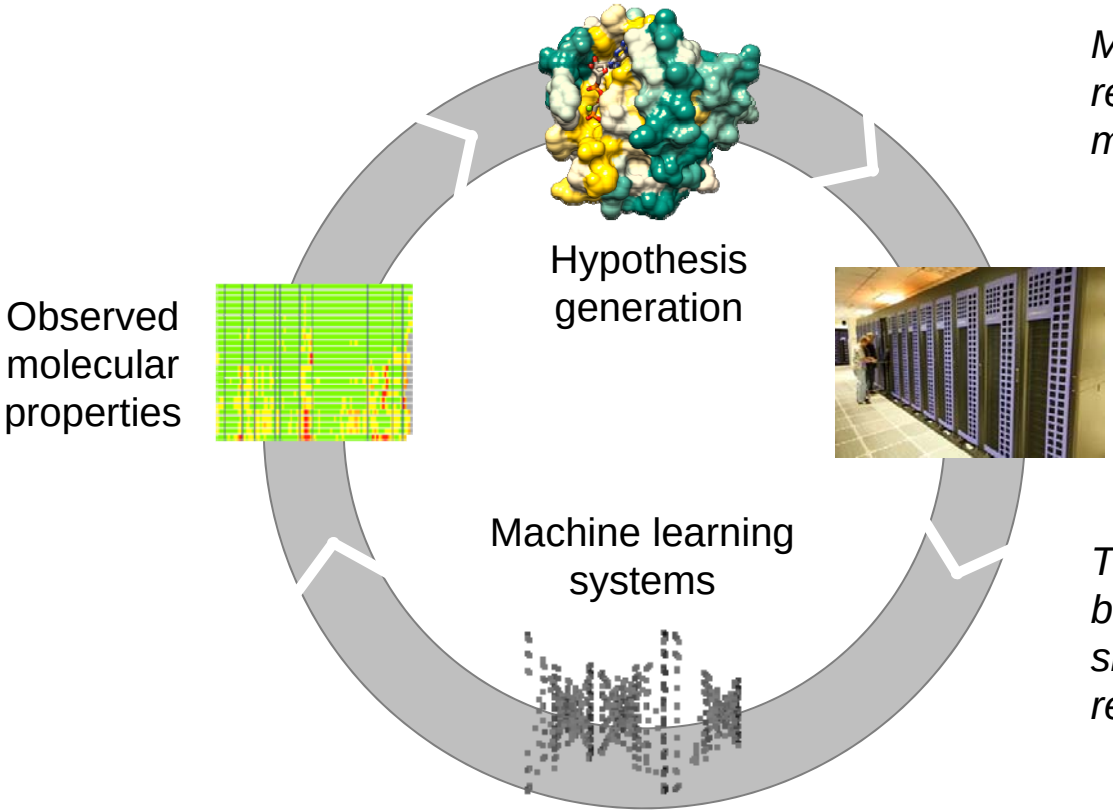


Building a machine learning framework for the NCI-DOE pilots

Supervised learning augmented by stochastic pathway modeling and experimental design

Semi-supervised learning, scalable data analysis and agent based simulations on population scale data

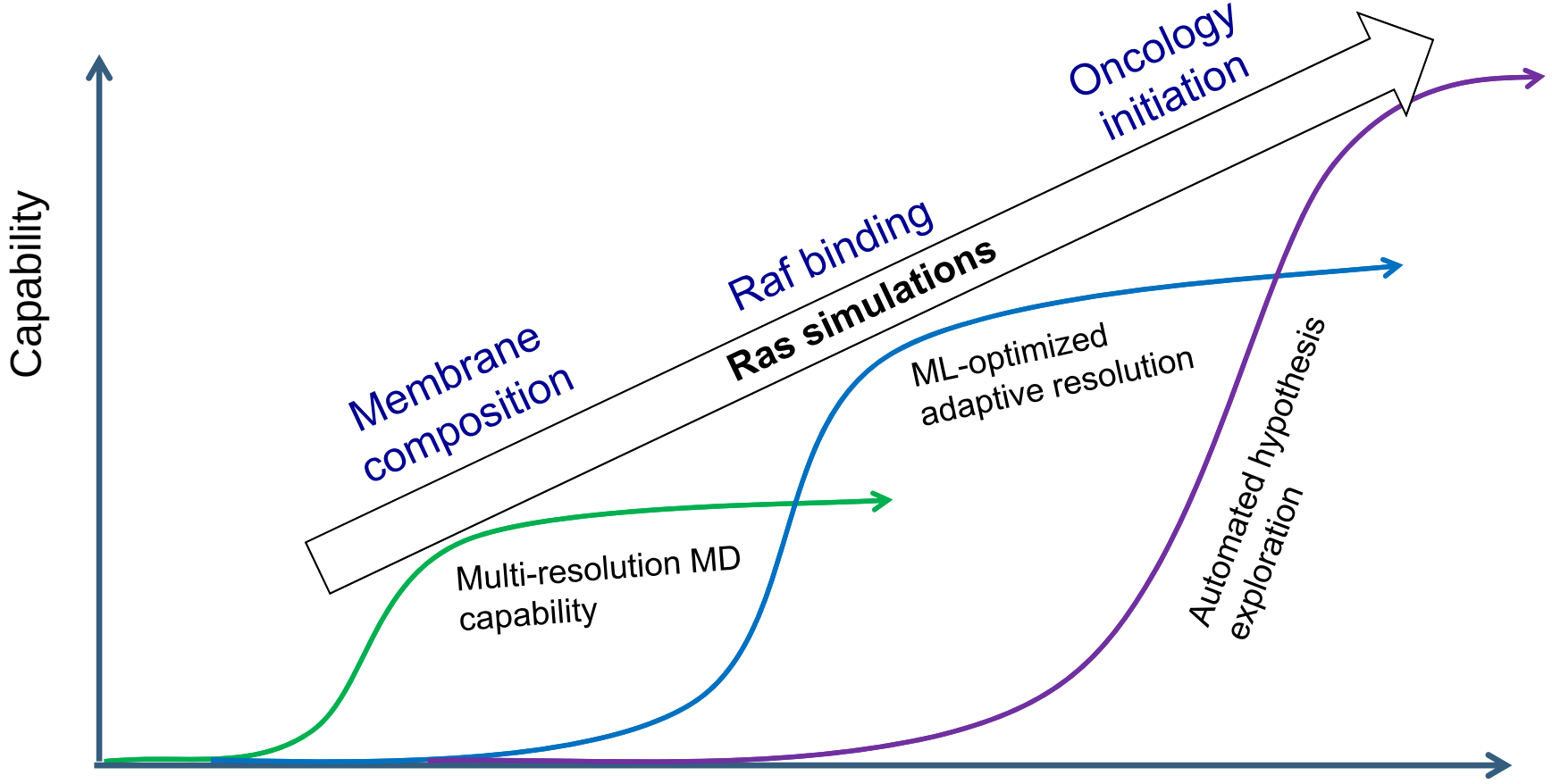
Machine learning enables a new dynamic validation approach to high-fidelity simulation



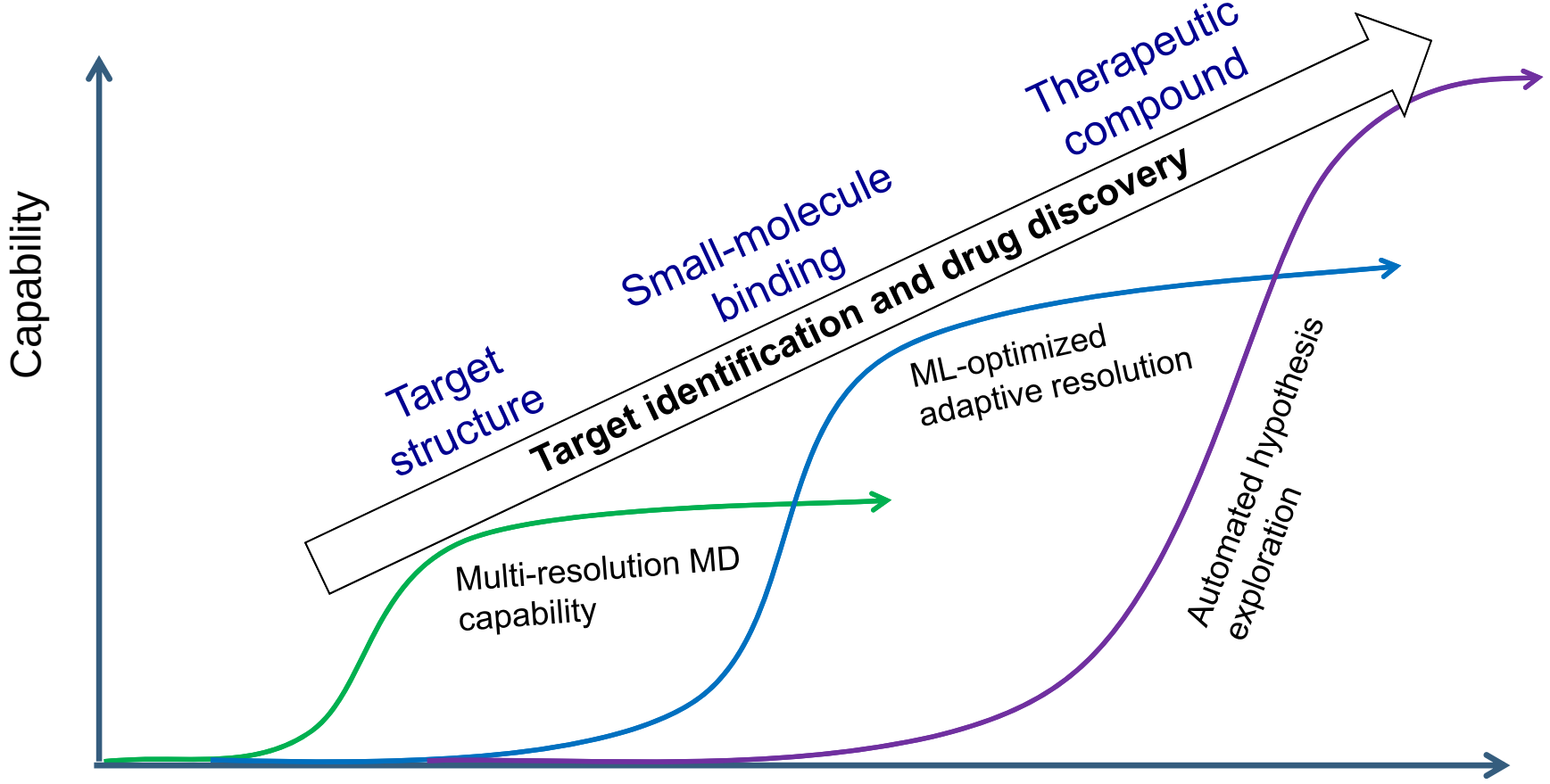
Machine learning systems uncover relationships between computational models and experimental data

This "active learning" approach can be used to optimize solutions with significant reduction in compute requirements

Project will build understanding on computational advances



New computational capability will be broadly applicable to NCI and DOE missions



Multi-disciplinary team from FNLCR, LLNL, LANL, ORNL and ANL

FNLCR / NCI Team: Frantz Jean-Francois, **Frank McCormick**, Dhirendra Simanshu, Eric Stahlberg, **Andy Stephen**, Tommy Turbyville, Debanjan Goswami

Argonne National Lab: *Prasanna Balaprakash,* ***Tom Brettin****, FangFang Xia*

Oak Ridge National Lab: , Pratul K. Agarwal, Debsindhu Bhowmik, ***Arvind Ramanathan,*** Blake A. Wilson, Christopher B. Stanley

Los Alamos National Lab: Angel Garcia, Christoph Jungans, Cesar Lopez, Chris Neale, Danny Perez, Sandrasegaram Gnanakaran, Tim Travers, Art Voter

Lawrence Livermore National Lab: Harsh Bhatia, Barry Belmont, Tim Carpenter, Francesco Di Natale, **Jim Glosli**, Helgi Ingolfsson, Piyush Karande, **Felice Lightstone**, Tomas Oppelstrup, Liam Stanton, Michael Surh, *Sachin Talathi, Brian Van Essen,* Yue Yang, Xiaohua Zhang