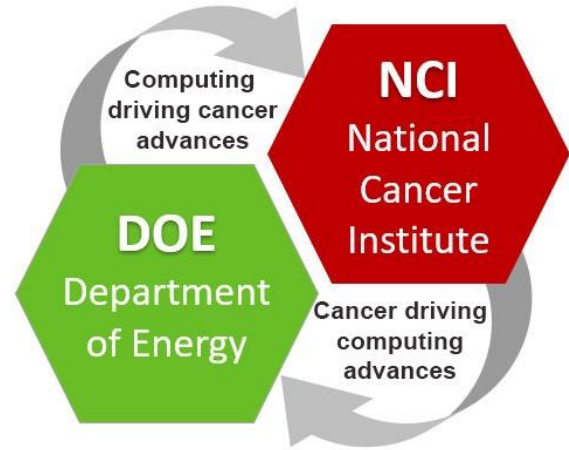


National Cancer Institute-Department of Energy Collaboration

A partnership to accelerate cancer research through predictive modeling and advanced computing



Initiatives Supported NSCI and PMI

NIH NATIONAL CANCER INSTITUTE

Argonne NATIONAL LABORATORY

OAK RIDGE NATIONAL LABORATORY

Lawrence Livermore NATIONAL LABORATORY

Los Alamos NATIONAL LABORATORY EST. 1943

Frederick National Laboratory for Cancer Research

sponsored by the National Cancer Institute



MOSSAIC: Modeling Outcomes using Surveillance data and Scalable Artificial Intelligence for Cancer

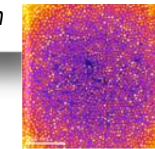


Clinical Domain – Precision oncology surveillance
Near real-time SEER reporting using state-of-the art Transformer language models
Improved clinical trial selection and feasibility assessment

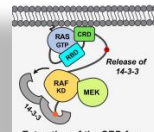
IMPROVE: Innovative Methodologies and New Data for Predictive Oncology Model Evaluation

Drug Response	Drug 1	Drug 2	Drug 3	Drug 4
WITHIN	Green	Red	Green	Red
WITHOUT	Red	Green	Red	Green

Pre-clinical Domain – Improved predictive models
Comparing and improving deep learning models of tumor drug response
Improved experimental design

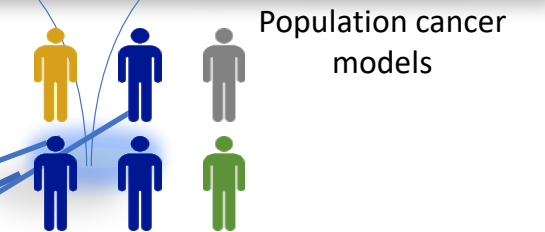


ADMIRRAL: AI-Driven Multi-scale Investigation of RAS/RAF Activation Lifecycle



Molecular Domain – Multiscale biological simulations
Machine learning guided exploration of protein dynamics
Prediction of protein domain movement with molecular resolution

Multiple distinct cancer models



Single mechanism

U.S. DEPARTMENT OF ENERGY

NIH NATIONAL CANCER INSTITUTE

Crosscut: CANDLE experimental technology

2023 R&D100 Award Winner!



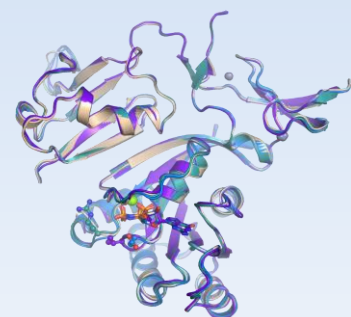
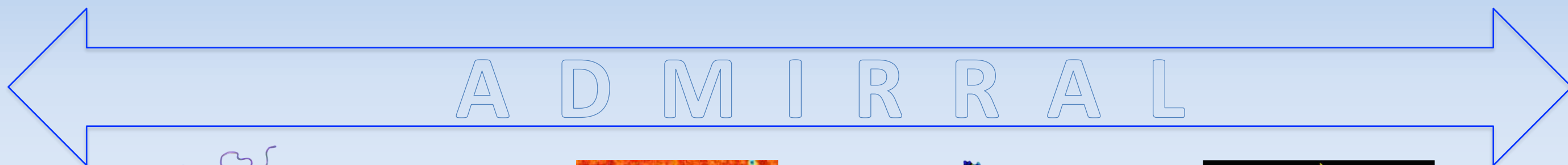
The ADMIRRAL Project

AI-Driven Multiscale Investigation of Ras-Raf Activation Lifecycle

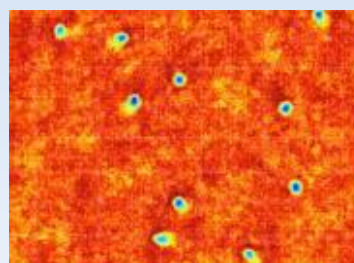
FNLAC Meeting
March 11, 2014

Dwight V. Nissley
Frederick National Laboratory for Cancer Research
US National Institutes for Health

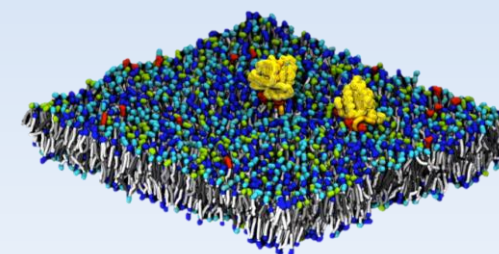
Frederick H. Streitz
Lawrence Livermore National Laboratory
US Department of Energy



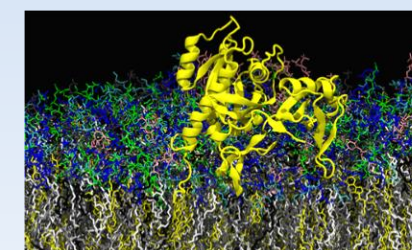
Protein Folding



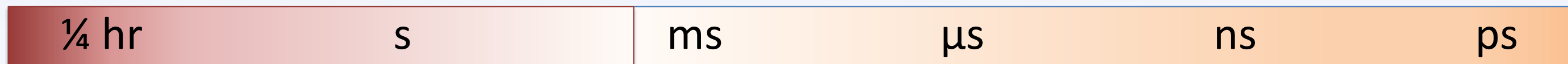
DDFT model



Coarse Grain MD



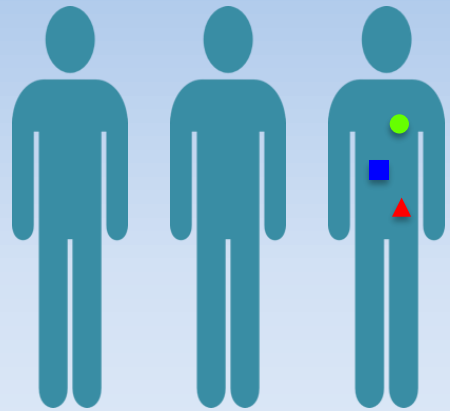
All-Atom MD



NCI-DOE Collaboration - ADMIRRAL

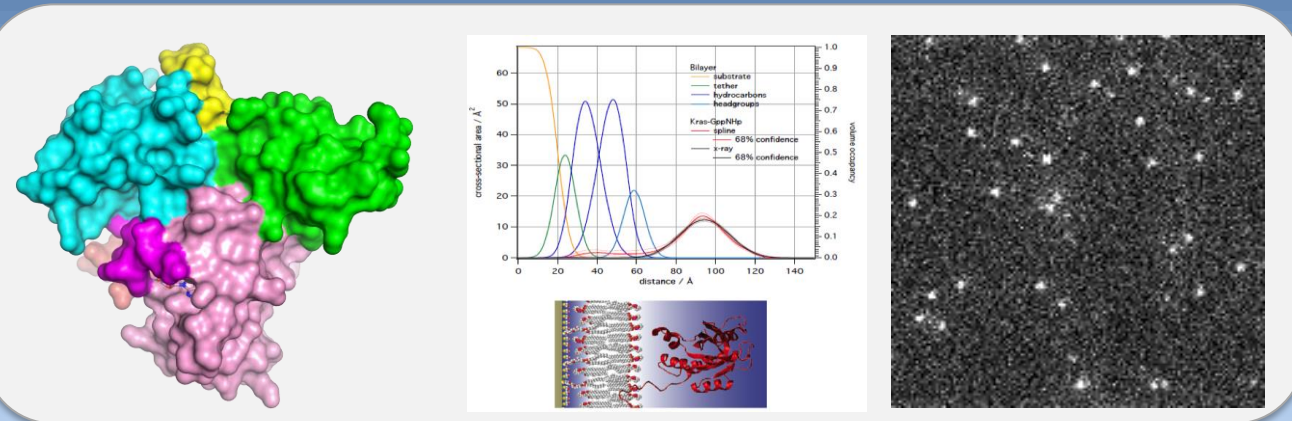


20-30% of cancers have mutated RAS

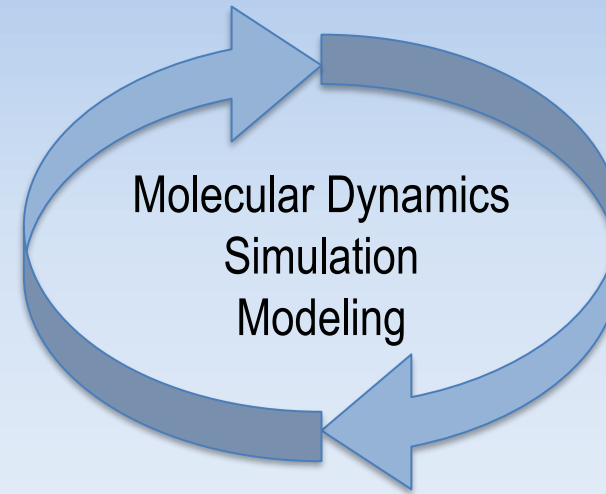


Current therapies ineffective against RAS-driven cancer

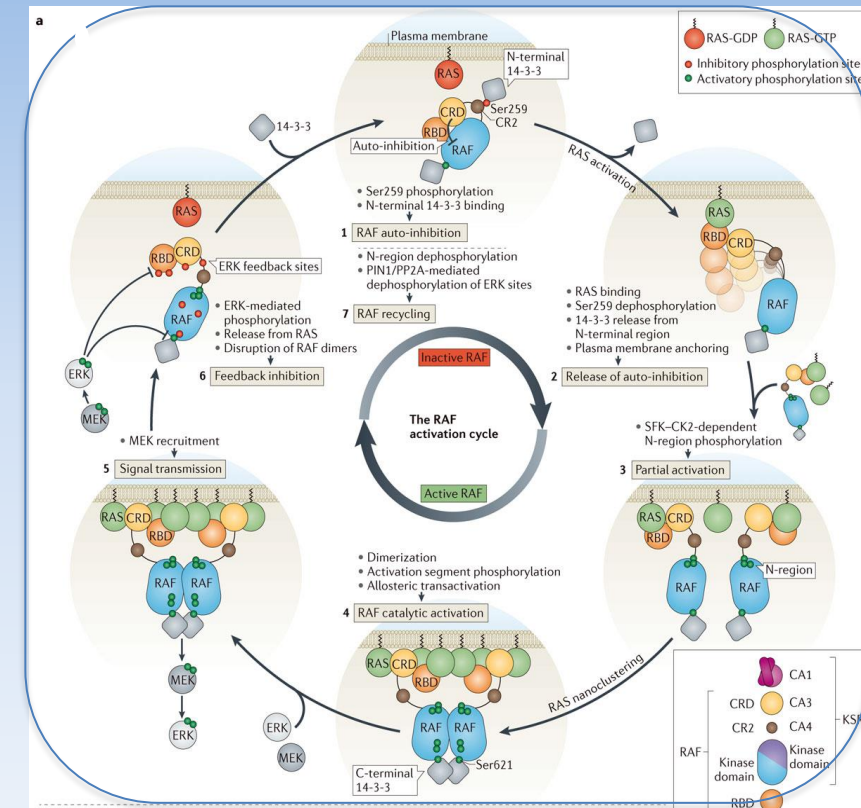
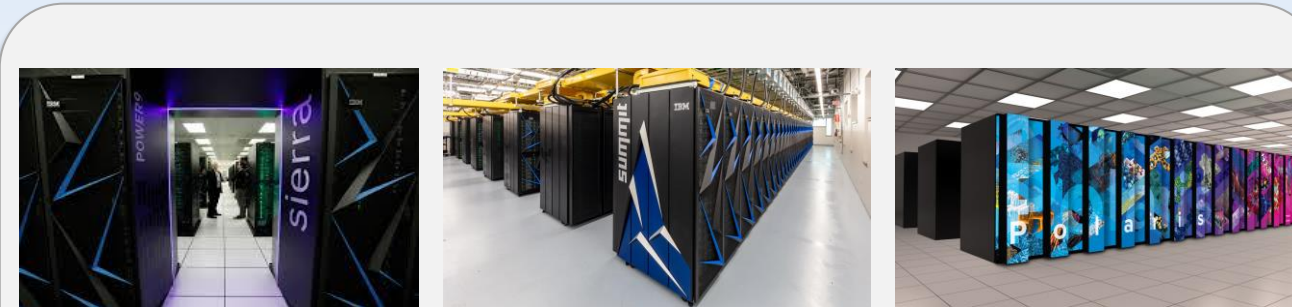
Biological and mechanistic insights to facilitate discovery of therapeutics



FNLCR



LLNL



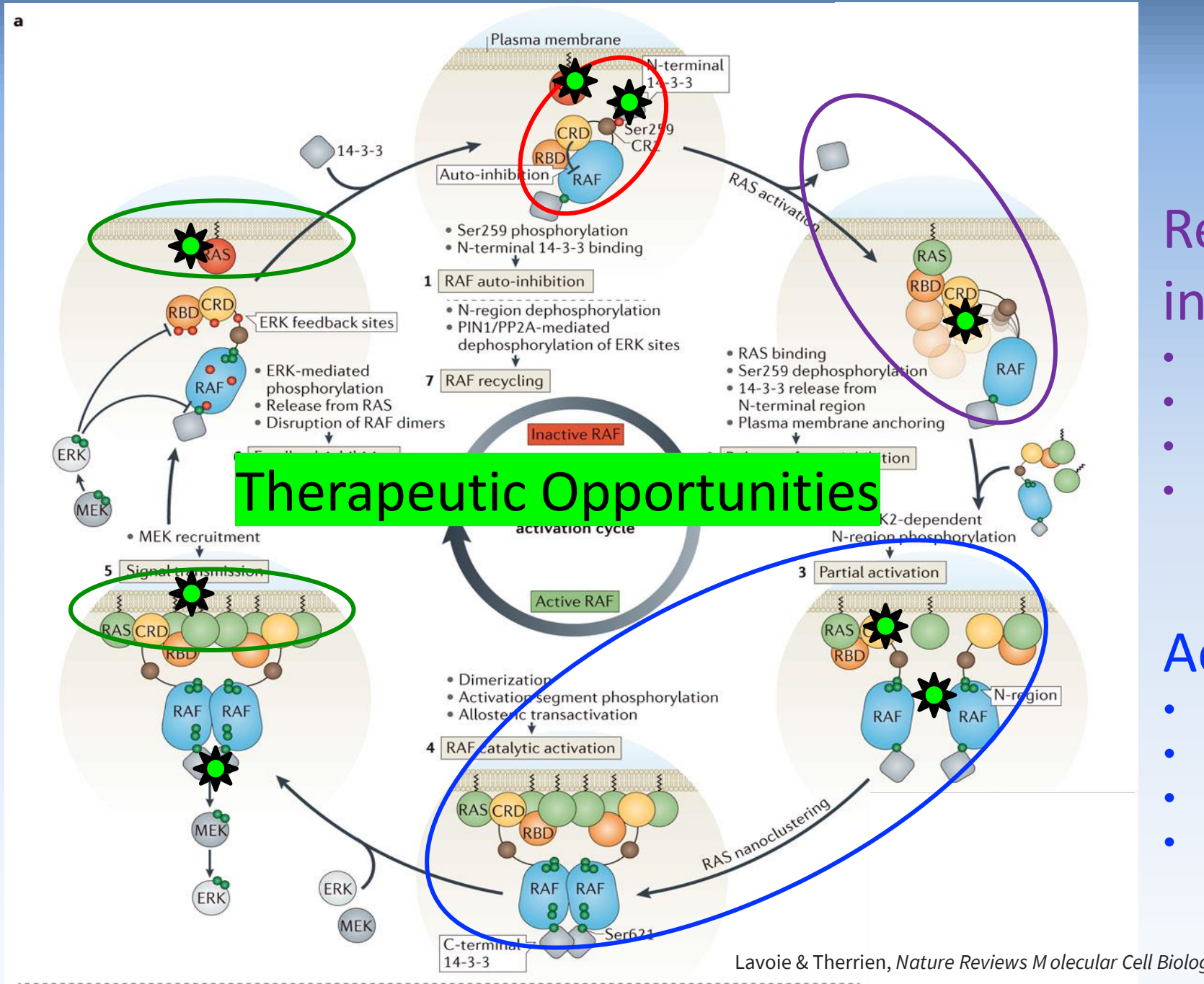
Ras - Raf Activation Lifecycle

Auto-inhibited RAF



KRAS
membrane

Signaling
platform



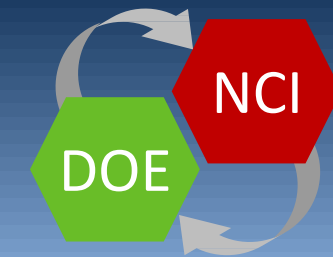
Release of Auto-inhibition

- Dephosphorylation,
- 14-3-3 release/rearrangement
- structural changes
- Engagement of RAS

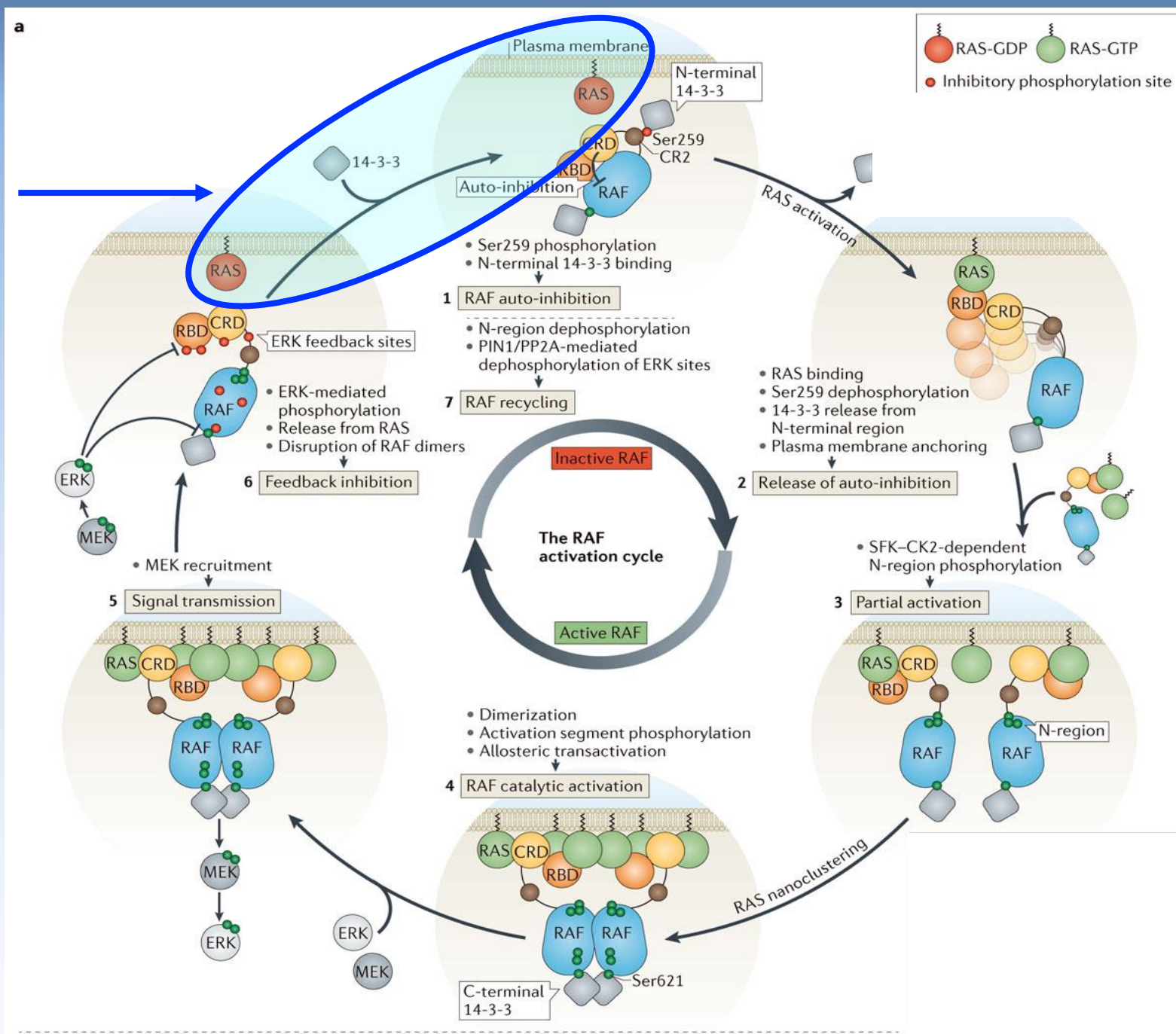
Activation

- CRD membrane engagement
- RAS clustering
- Activating phosphorylation
- Kinase domain dimerization

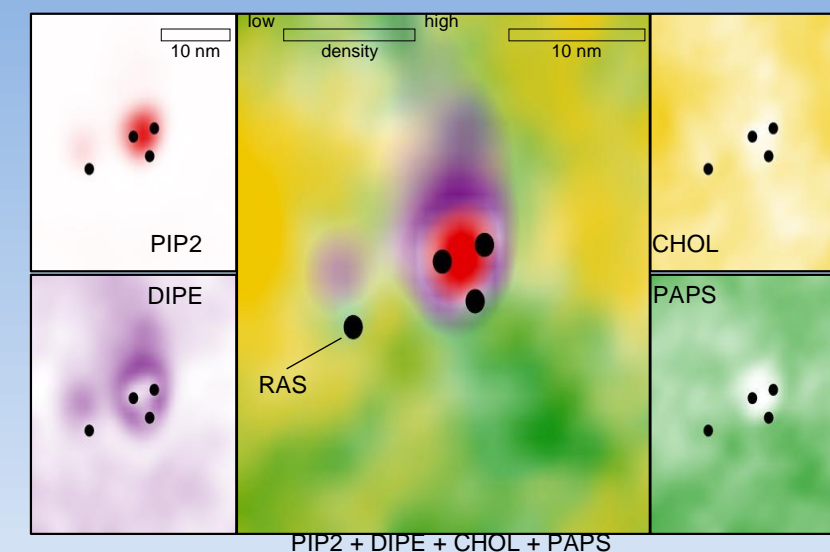
Ras - Raf Activation Lifecycle



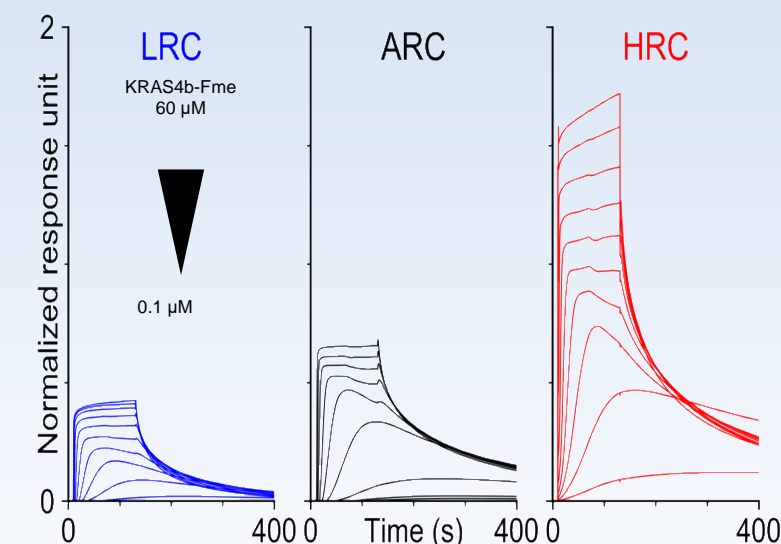
Pilot 2



Lipid fingerprints associated with RAS clusters



RAS binding correlates with fingerprint



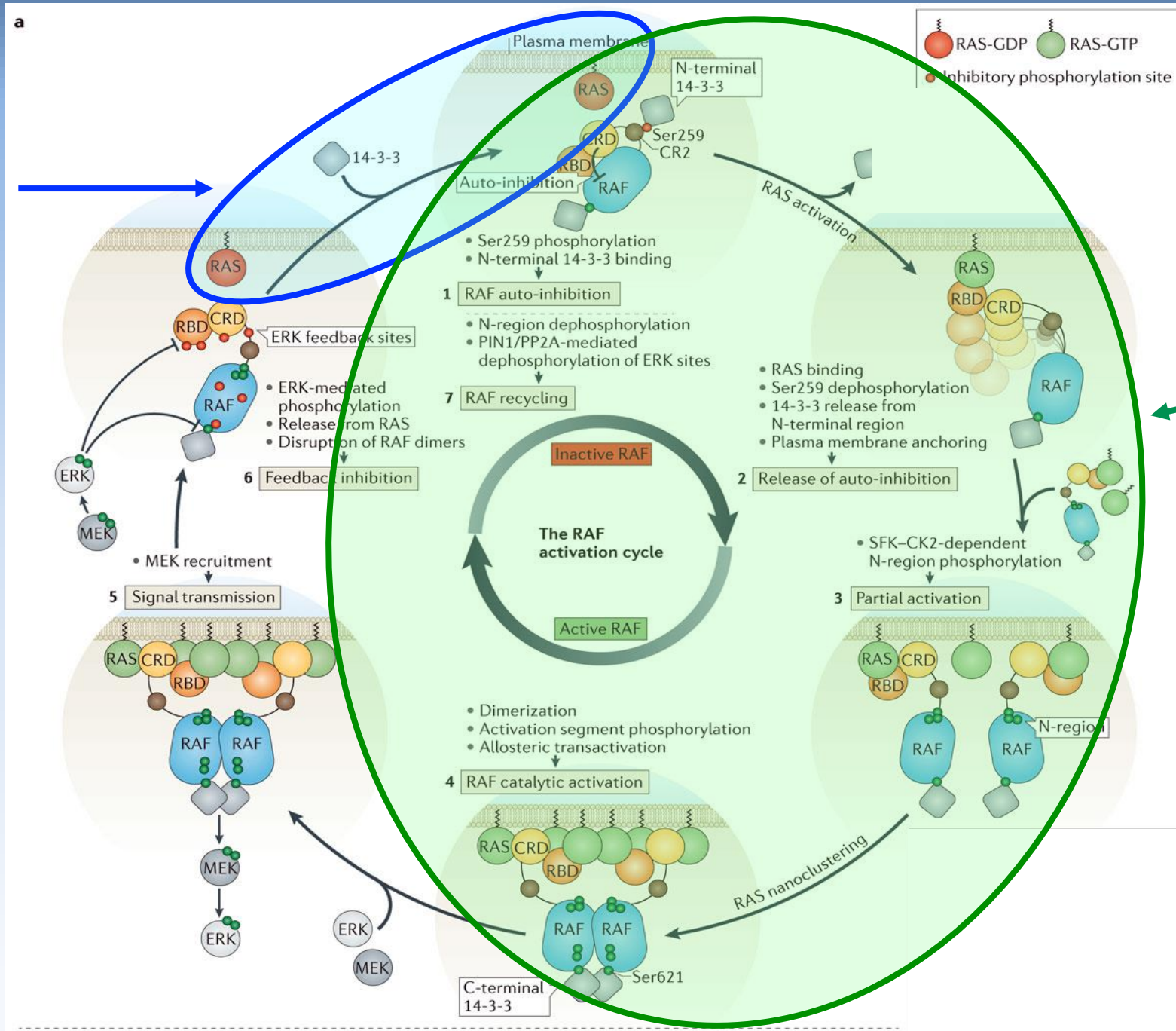
Ras - Raf Activation Lifecycle



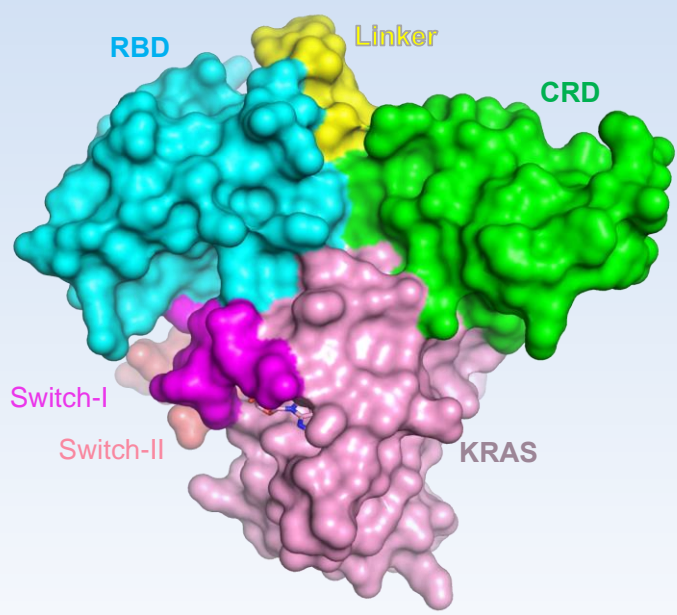
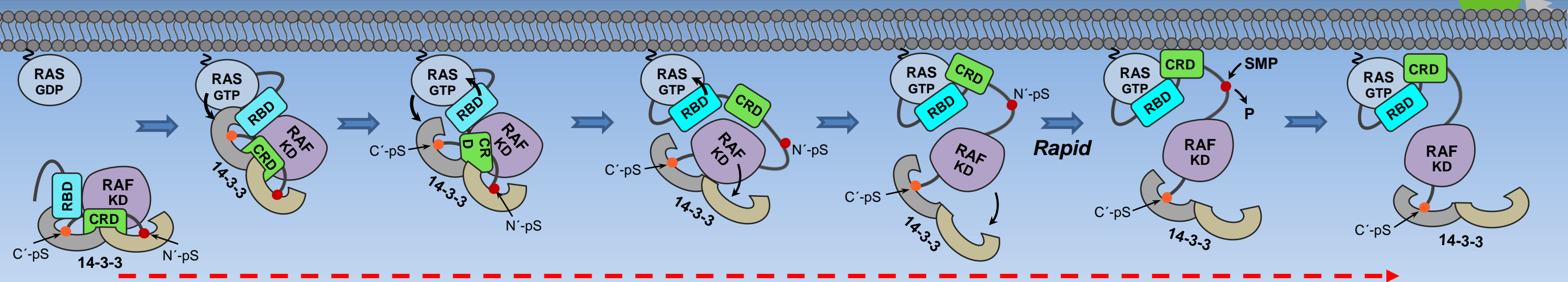
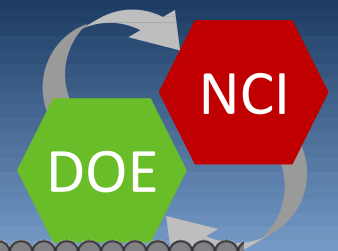
Pilot 2

ADMIRRAL

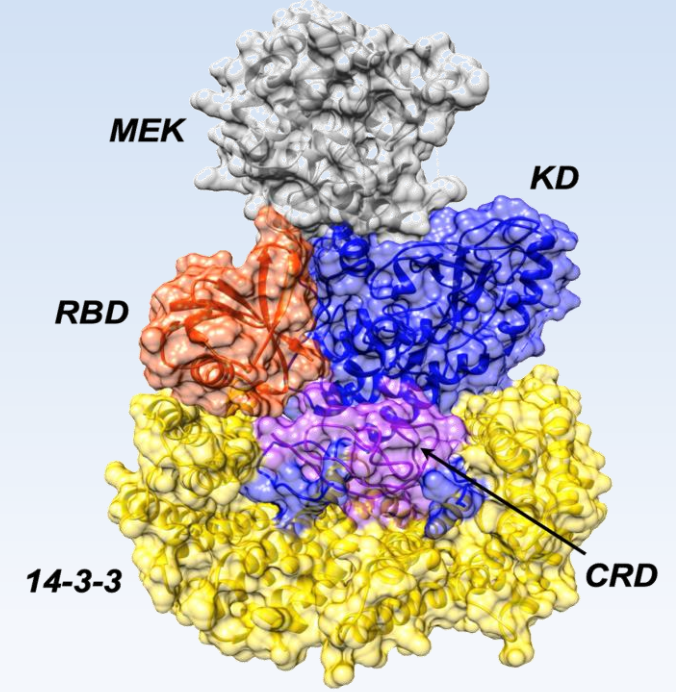
AI-Driven Multi-scale Investigation of Ras-RAF Activation Lifecycle



ADMIRRAL Project: RAF Activation



Autoinhibited BRAF-MEK1-14-3-3 complex

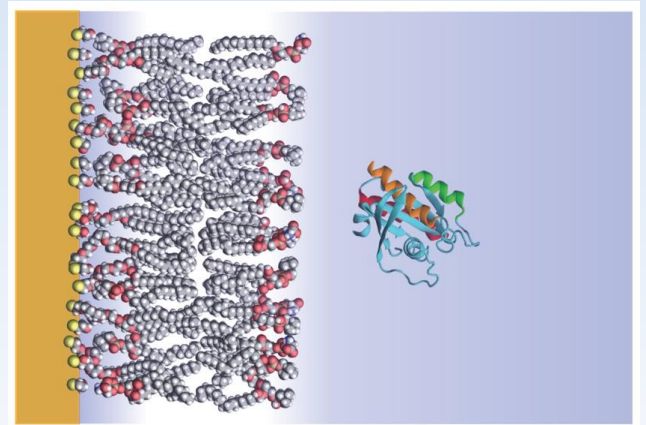
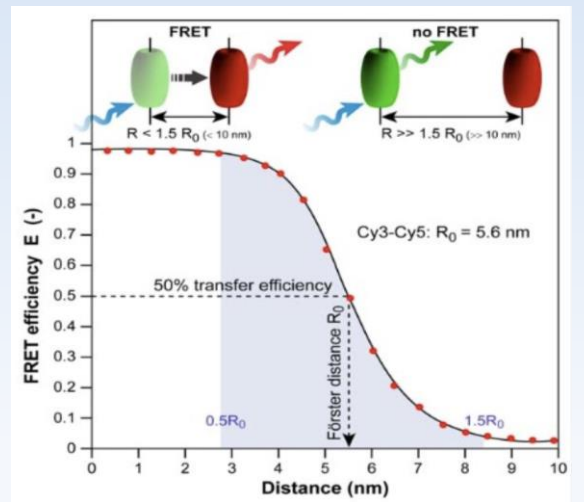
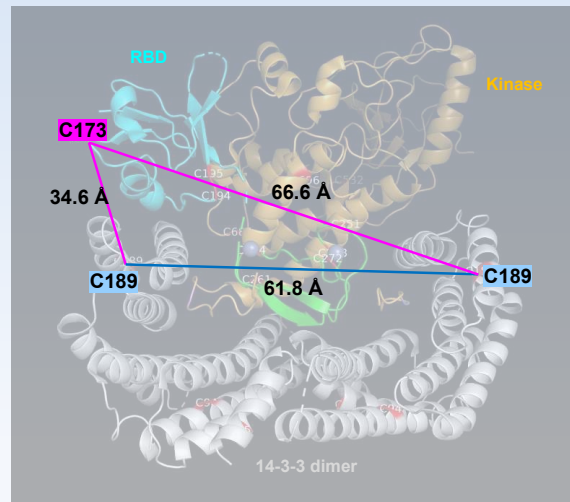
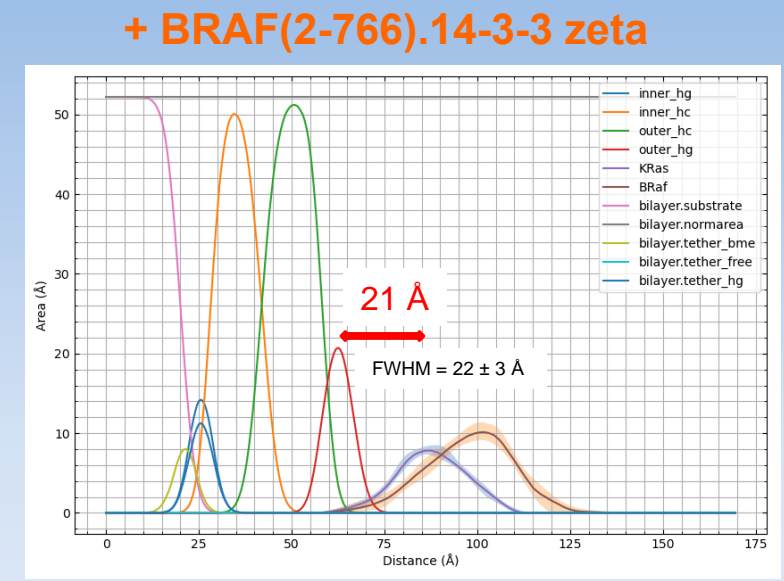
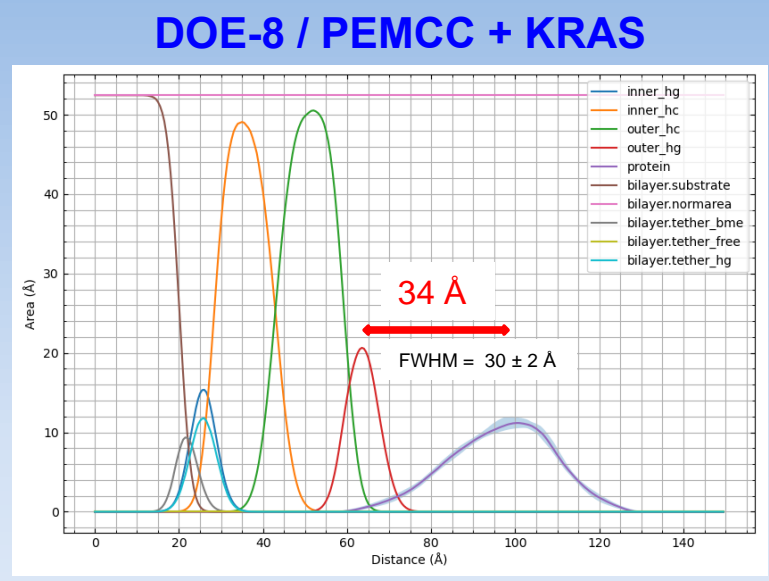
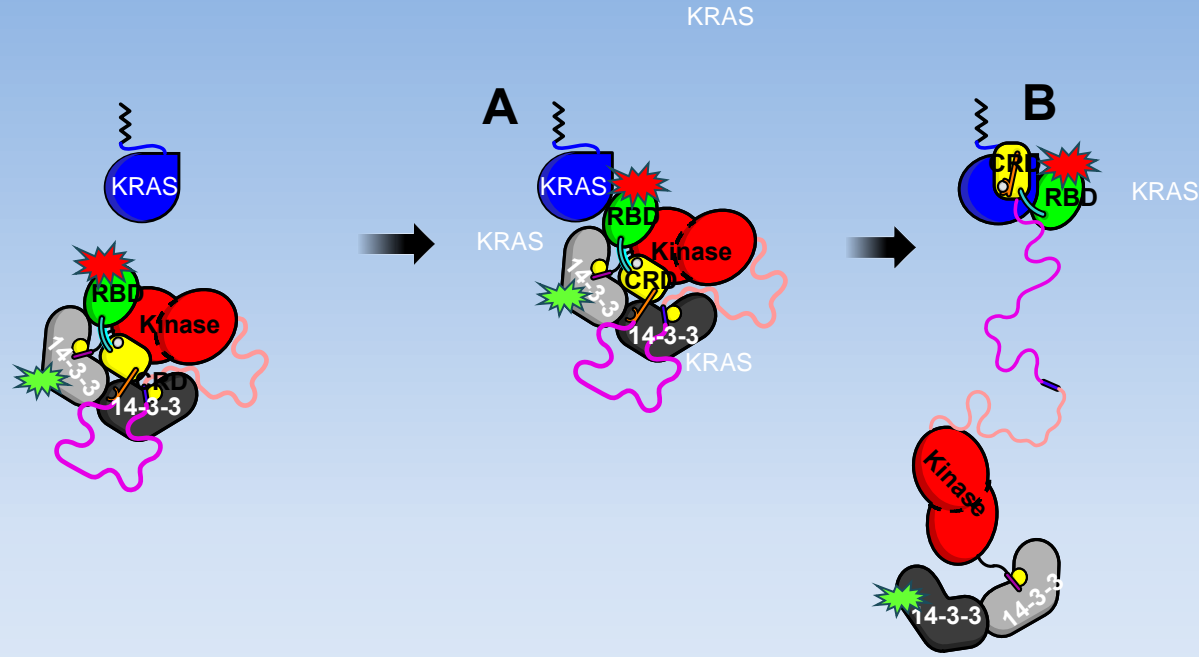


ADMIRRAL Experimental Approaches



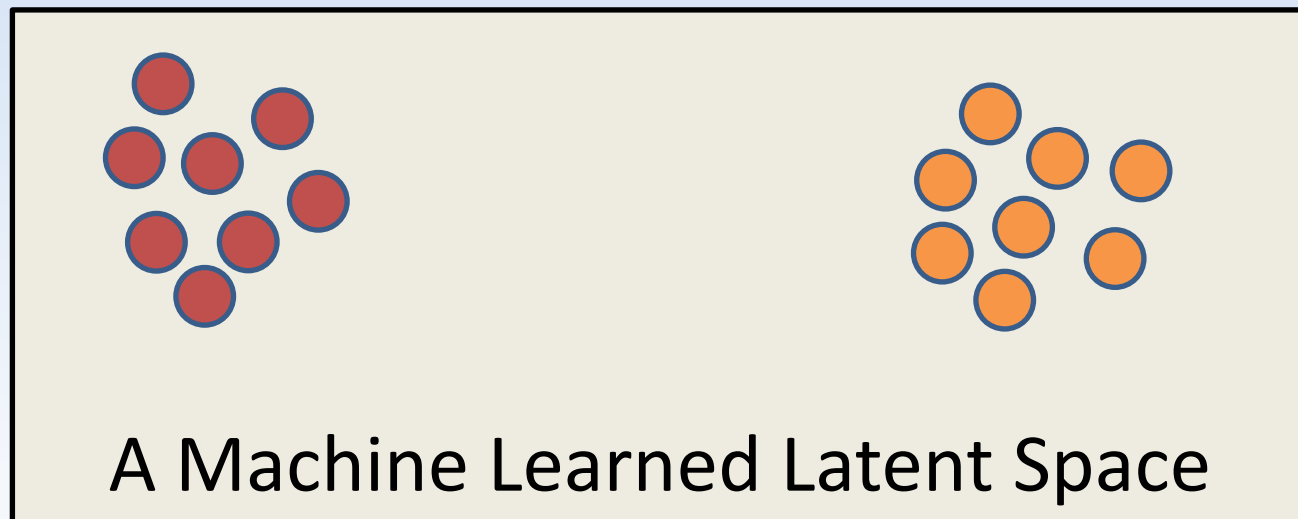
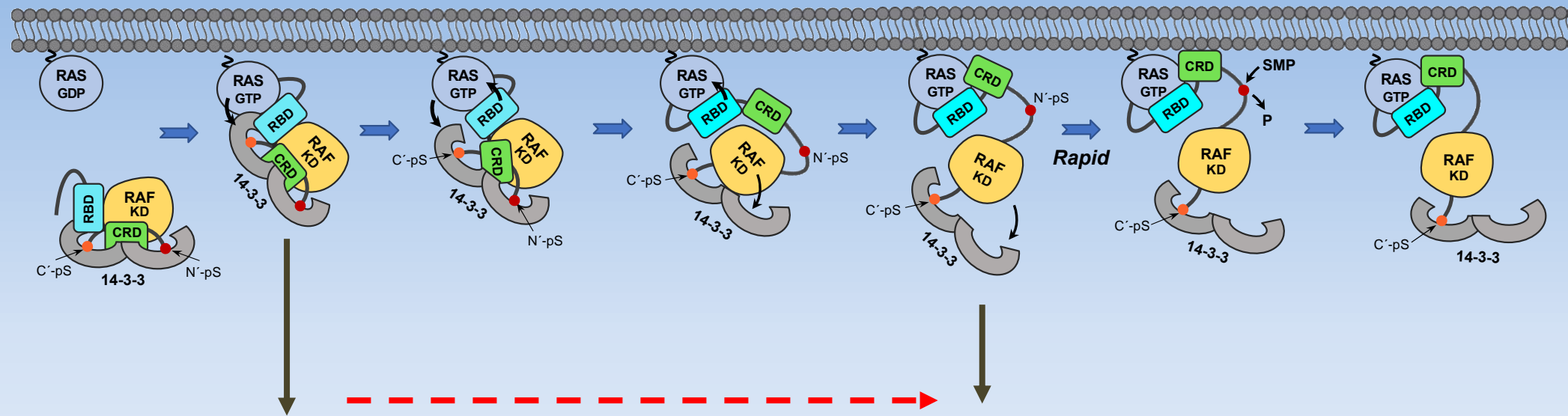
smFRET to monitor conformational changes

Neutron Reflectivity: Structure of BRAF.1433 at the membrane

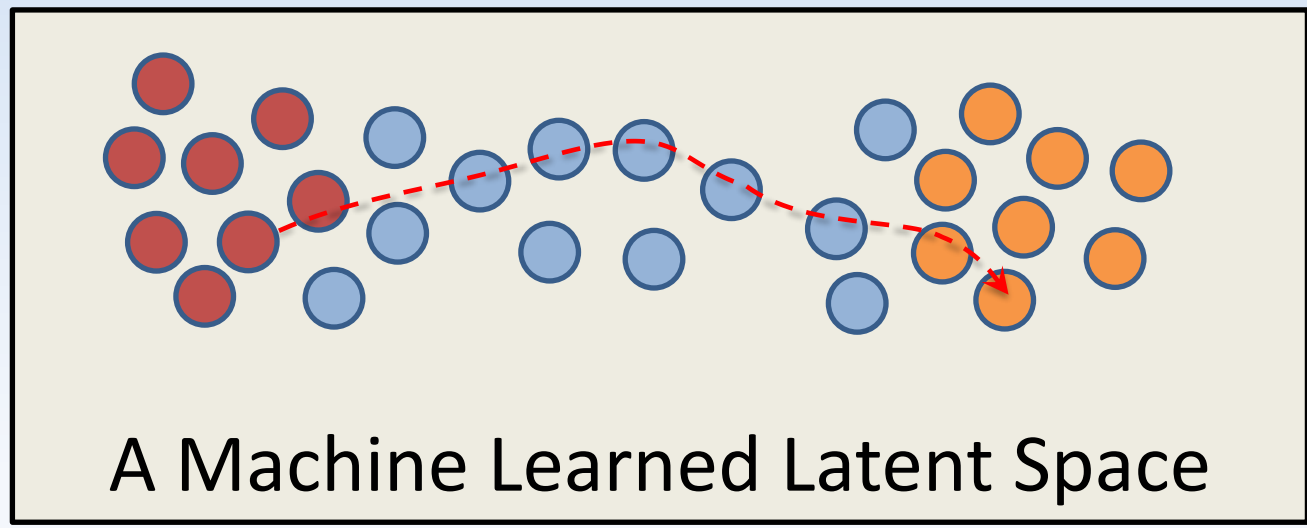
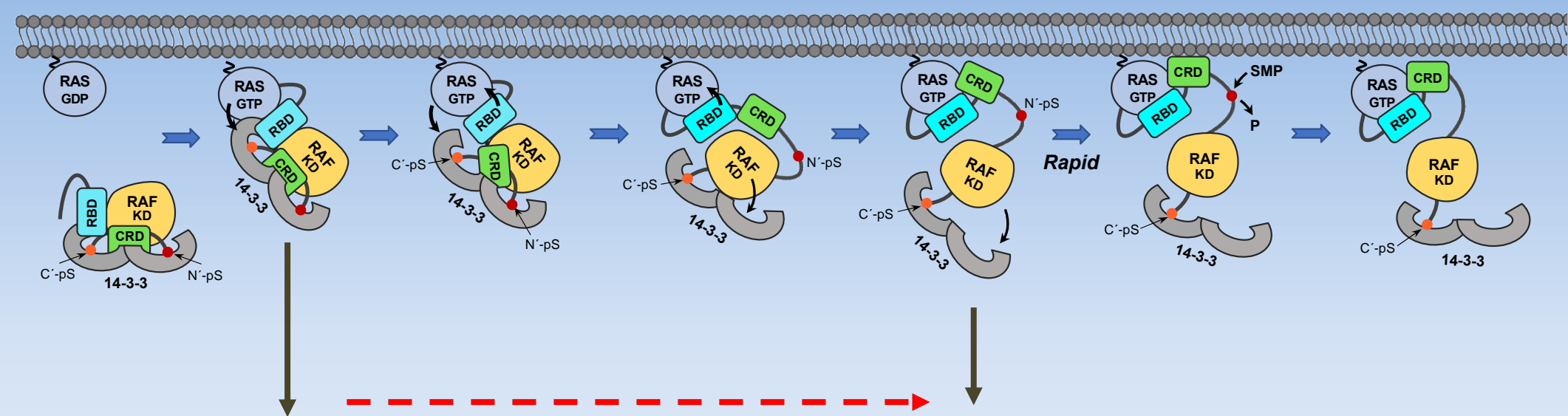


NR data fitting in progress

Use ML to develop suitable latent spaces that capture evolution of protein structures



Iteratively sample-and-retrain a latent space to find a bridge between configurations



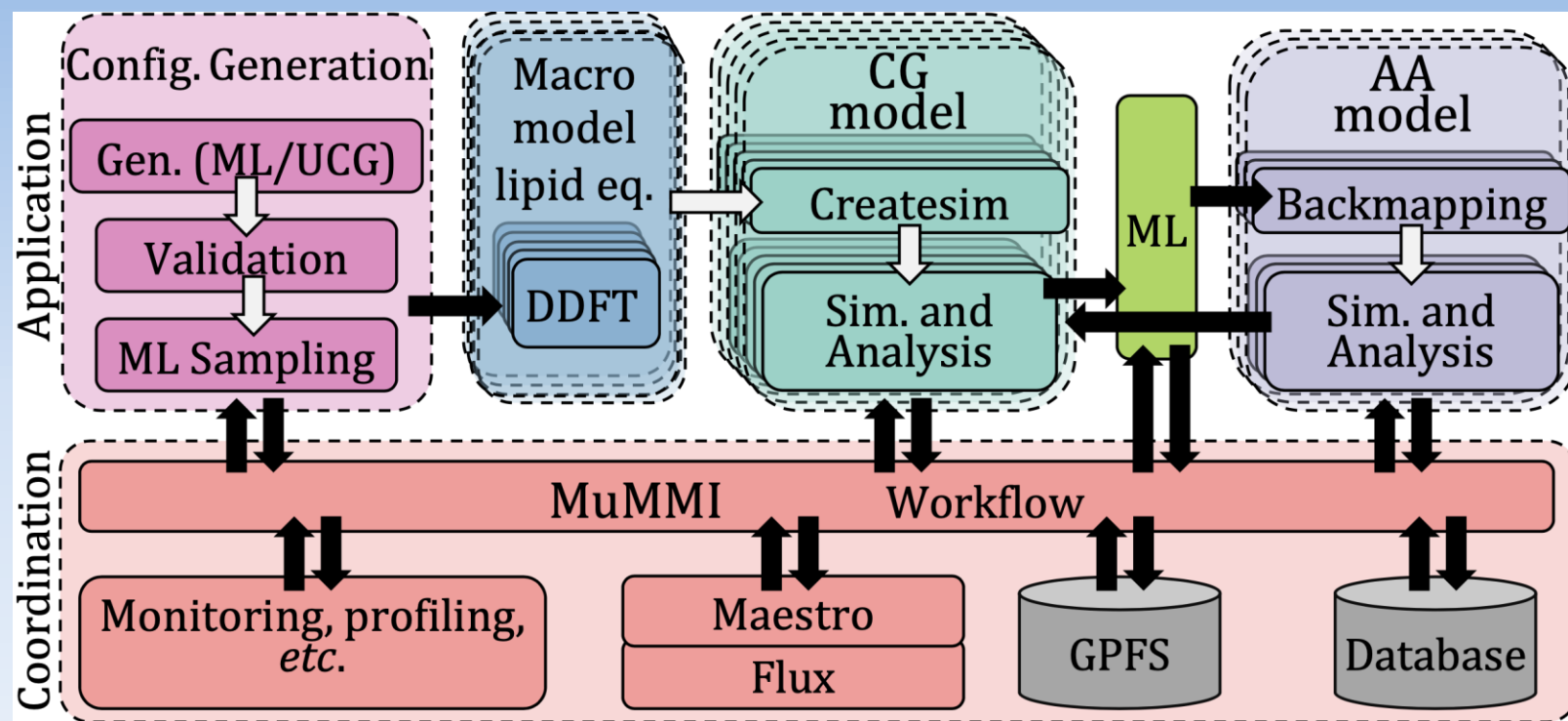
```

    graph TD
      A[Run [U]CG simulations] --> B[Create/update latent space]
      B --> C[Sample new configs to complete the trajectory]
      C --> D[Run [U]CG simulations]
      D --> E{Trajectory complete?}
      E -- No --> B
      E -- Yes --> F{Trajectory valid?}
      F -- Yes --> G[Task Complete]
      F -- No --> C
  
```


MuMMI* 2.0 - managing and executing a workflow at exascale



The MuMMI 2.0 workflow manages ML-generated selection of hypotheses and subsequent validation



- The coarsest scale samples protein conformational changes
- Macro model equilibrates lipid configurations
- CG model validates hypothesis
- AA model validates CG model

Builds on award-winning development of MuMMI during Pilot 2

* Multi-scale Machine-learned Modeling Infrastructure

Outreach/Capability Transfer



MuMMI Framework <https://github.com/mummi-framework>

Includes mummi-core and mummi-ras codebases

ddcMD <https://github.com/LLNL/ddcMD>

A fully GPU-accelerated molecular dynamics program for the Martini force field

MemSurfer <https://github.com/LLNL/MemSurfer>

MemSurfer is a tool to compute and analyze membrane surfaces found in a wide variety of large-scale molecular simulations.

MARTINI parameters <http://cgmartini.nl/index.php/force-field-parameters/lipids>

Refined lipid parameters that reproduce experimental phase diagram have been uploaded to MARTINI website for public use

Maestro <https://github.com/LLNL/maestrowf>

Maestro Workflow Conductor is a Python tool and library for specifying and automating multi-step computational workflows both locally and on supercomputers.

Flux <https://flux-framework.github.io>

Flux is a next-generation resource and job management framework

DataBroker <https://github.com/IBM/data-broker>

The Data Broker (DBR) is a distributed, in-memory container of key-value stores enabling applications in a workflow to exchange data through one or more shared namespaces.

Dynim <https://github.com/LLNL/dynim>

Framework for dynamic diversity sampling used for coupling pairs of scales in MuMMI

TarIdx <https://github.com/LLNL/pytaridx>

Framework for creating robust archives of data to allow MuMMI to handle up to billions of files.

Campaign 1 data

https://modac.cancer.gov/assetDetails?dme_data_id=NCI-DME-MS01-1069927 & https://modac.cancer.gov/assetDetails?dme_data_id=NCI-DME-MS01-10699461

<https://bbs.llnl.gov/RAS-lipid-dependent-dynamics-data.html>

Over 8M files (> 350TB) of data transferred to Predictive Oncology Model and Data Clearinghouse

Acknowledgements



FNLCR

Andy Stephen
Que Van
Suzanne Sandin
Erik Larsen
Sophie Krahnke
Tommy Turbyville
Rebika Shrestha
Pedro Andrade Bonilla
Rodrigo Caceres
Dom Esposito
Matt Drew
Simon Messing
Jenna Hull
Kelly Snead
Jen Mehalko
Billy Burgan
Dhirendra Simanshu
Daniel Bonsor

Caroline Dehart
Robert D'Ippolito
Grace Scheidemantle
Dwight Nissley

NCI

Debbie Morrison
David Durant
Dan Ritt

UCSF

Frank McCormick

NMRFAM

Marco Tonelli

MCW

Jason Sidabras

NIST, CMU

Frank Heinrich

LLNL

Fikret Aydin
Harsh Bhatia
Timo Bremer
Tim Carpenter
Tim Hsu
Joseph Chavez
Brian Van Essen
Jeremy Tempkin
Jim Glosli
Helgi Ingólfsson
Konstantia Georgouli
Mark Andrew Heimann
Loic Eric Pottier
Felice Lightstone
Christa Manning
Joseph Moon
Francesco Di Natale
Tomas Ooppelstrup

Claudio Santiago
Xiaohua Zhang
Fred Streitz

SJSU

Liam Stanton

Cornell

Robert Stephany

Extension to additional biology

