

Network of Optimized Dynamic Energy Signatures (NODES) Program Overview

Abhishek Singharoy



NODES Proposers' Day

August 1, 2025





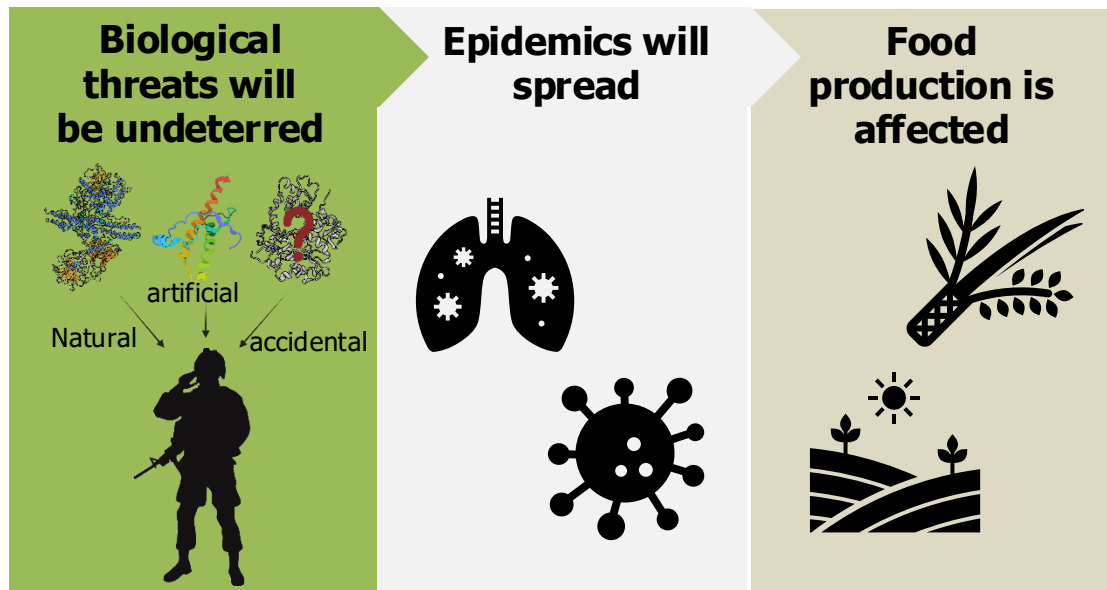
Agenda

Time (EST)		Length	Title	Presenter
12:00	13:00	1:00	Registration	All
13:00	13:05	0:05	Remarks from BTO Leadership	Dr. Mike Koeris (DARPA BTO Director)
13:05	13:35	0:30	NODES Technical Program Overview	Dr. Abhishek Singharoy (NODES Program Manager)
13:35	13:40	0:05	Mathematical Foundations of NODES	Dr. Yannis Kevrekidis (DSO Program Manager)
13:40	13:50	0:10	National Lab Representative (LLNL)	Dr. Jonathan Allen (Informatics Scientist)
13:50	14:00	0:10	National Lab Representative (ORNL)	Dr. Thomas L. Beck (Section Head of Science Engagement)
14:00	14:10	0:10	Ethical, Legal and Social Implications (ELSI)	Ms. Heather Roff (Adaptive Capabilities Office)
14:10	14:25	0:15	Doing Business with DARPA	Ms. Lydia Richards (Contracting Office)
14:25	14:35	0:10	DARPA Connect	Mr. Garrett Wright (Small Business Programs Office)
14:35	14:55	0:20	DARPA/BAA Inbox & Submission	Mr. David Swan III (Solicitation Coordinator)
14:55	15:05	0:10	Break/submit questions to DARPA team	Submit questions in person or by email to NODES@darpa.mil
15:05	16:35	1:30	Lightning Talks - Teaming	Attendees who submitted timely and conforming Lightning Talk attachments



Extensible Modeling and Simulation Platform for Universal Protein Function

Implications of Not Knowing Protein Function

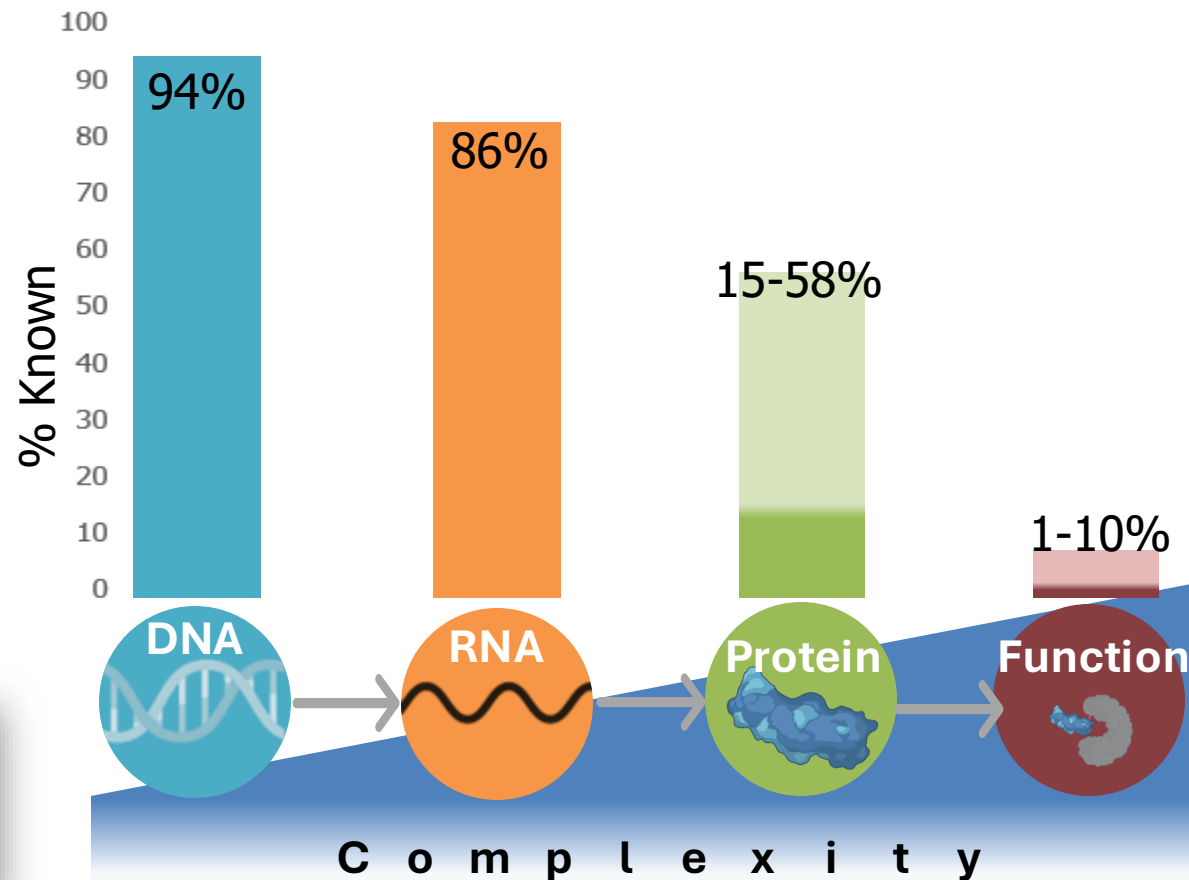


Adversaries will spring surprises willy nilly...

STAT+ | OPINION 66 FIRST OPINION

The U.S. is at risk of losing the proteomics race to China

Unlocking the power of human proteins could make a greater impact on science than the Human Genome Project



DNA: Deoxyribonucleic Acid

RNA: Ribonucleic Acid

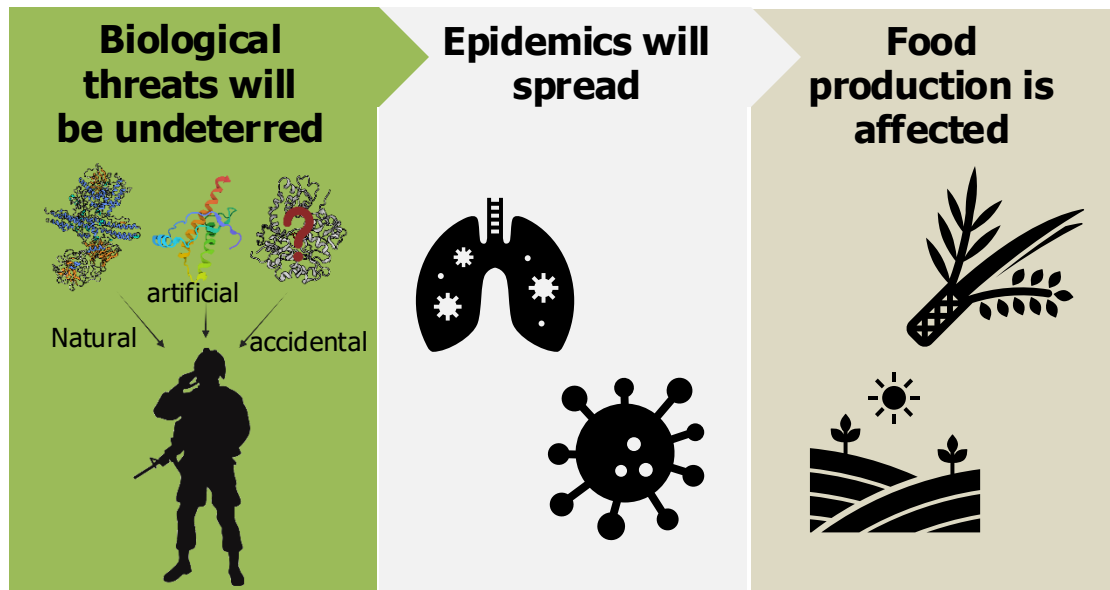
CUI: Controlled Unclassified Information

Vision: Deep technology to predict multiple unknown protein functions in **1 hour** and with **>90% confidence**, available in both **class** and **CUI forms** with a public-facing User Interface



Extensible Modeling and Simulation Platform for Universal Protein Function

Implications of Not Knowing Protein Function

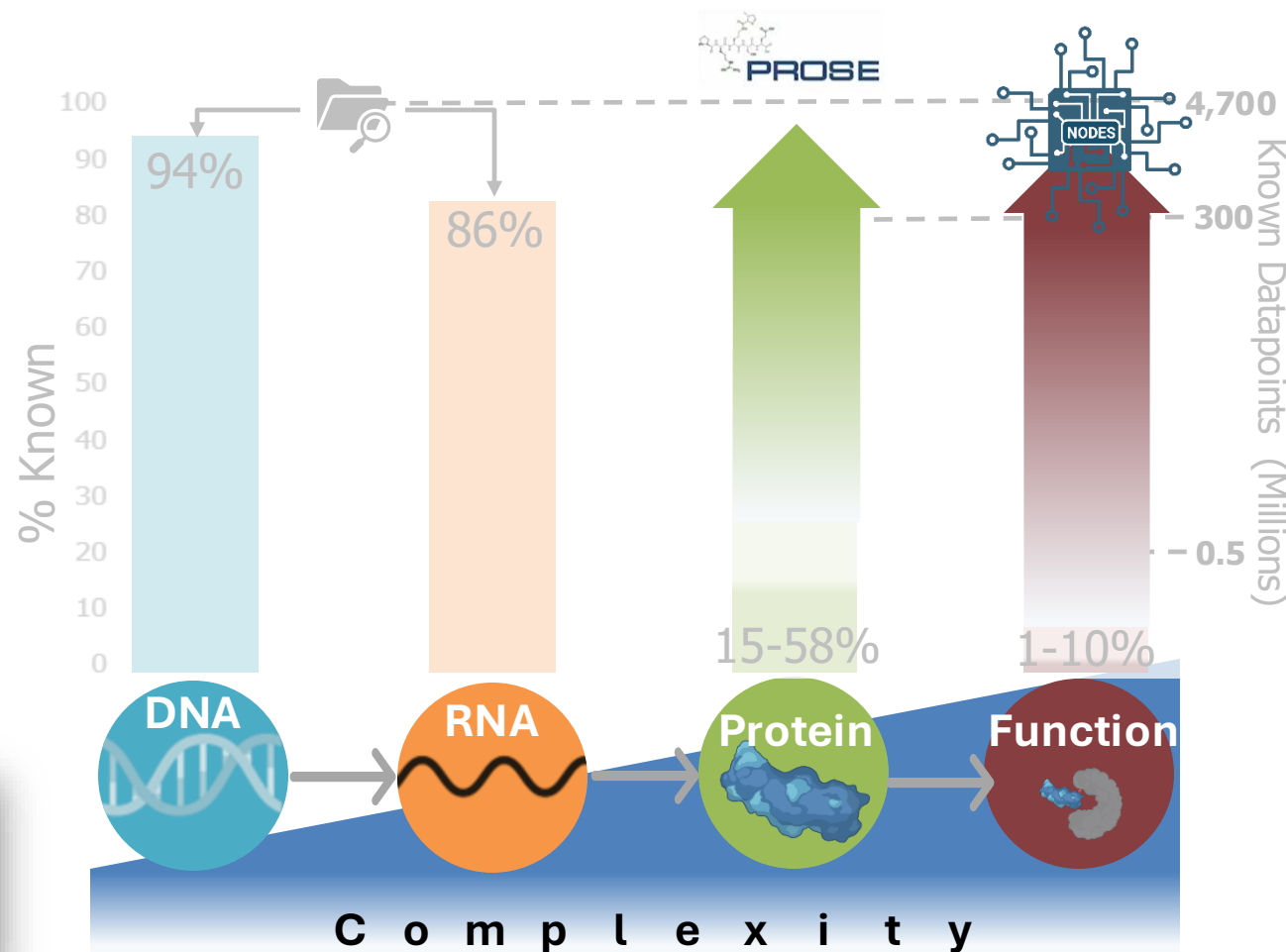


Adversaries will spring surprises willy nilly...

STAT+ | OPINION 66 FIRST OPINION

The U.S. is at risk of losing the proteomics race to China

Unlocking the power of human proteins could make a greater impact on science than the Human Genome Project



DNA: Deoxyribonucleic Acid

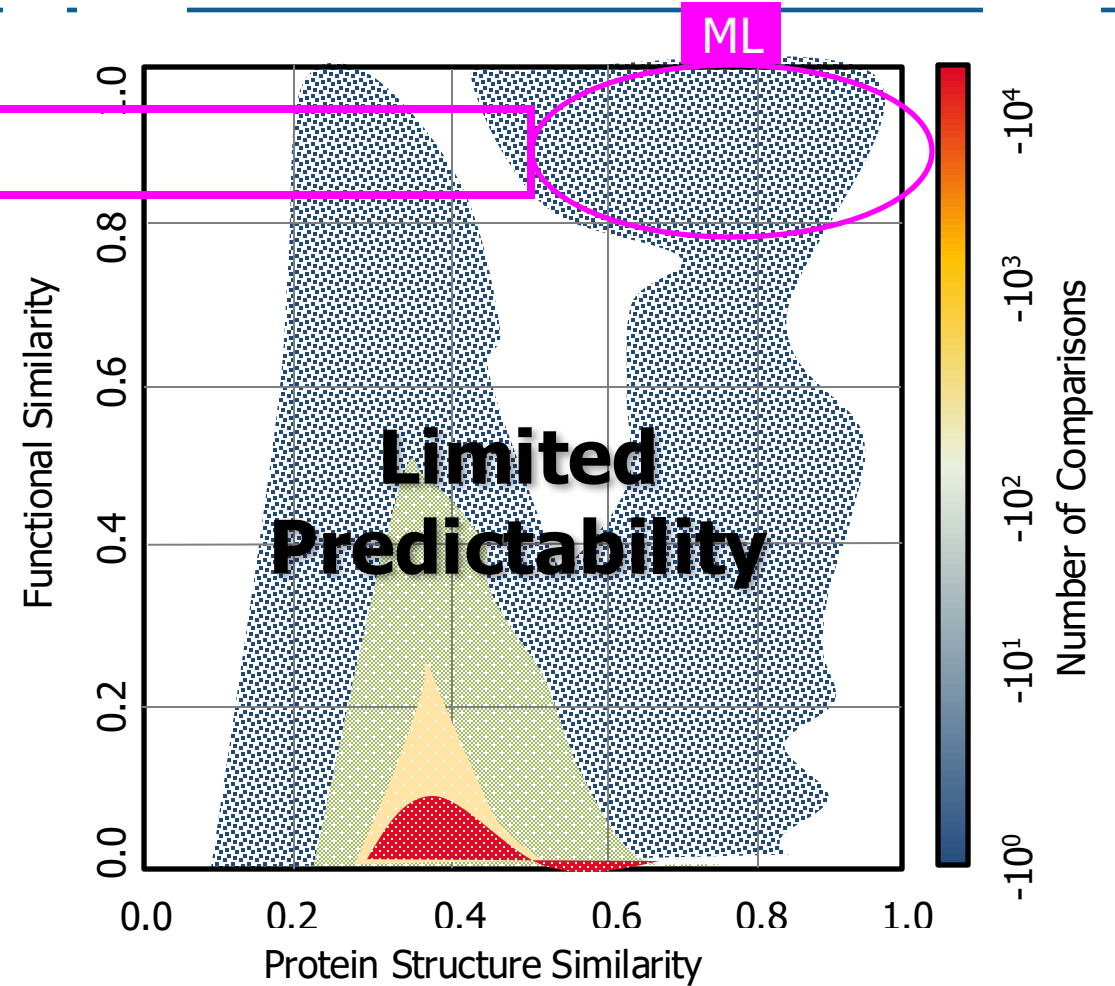
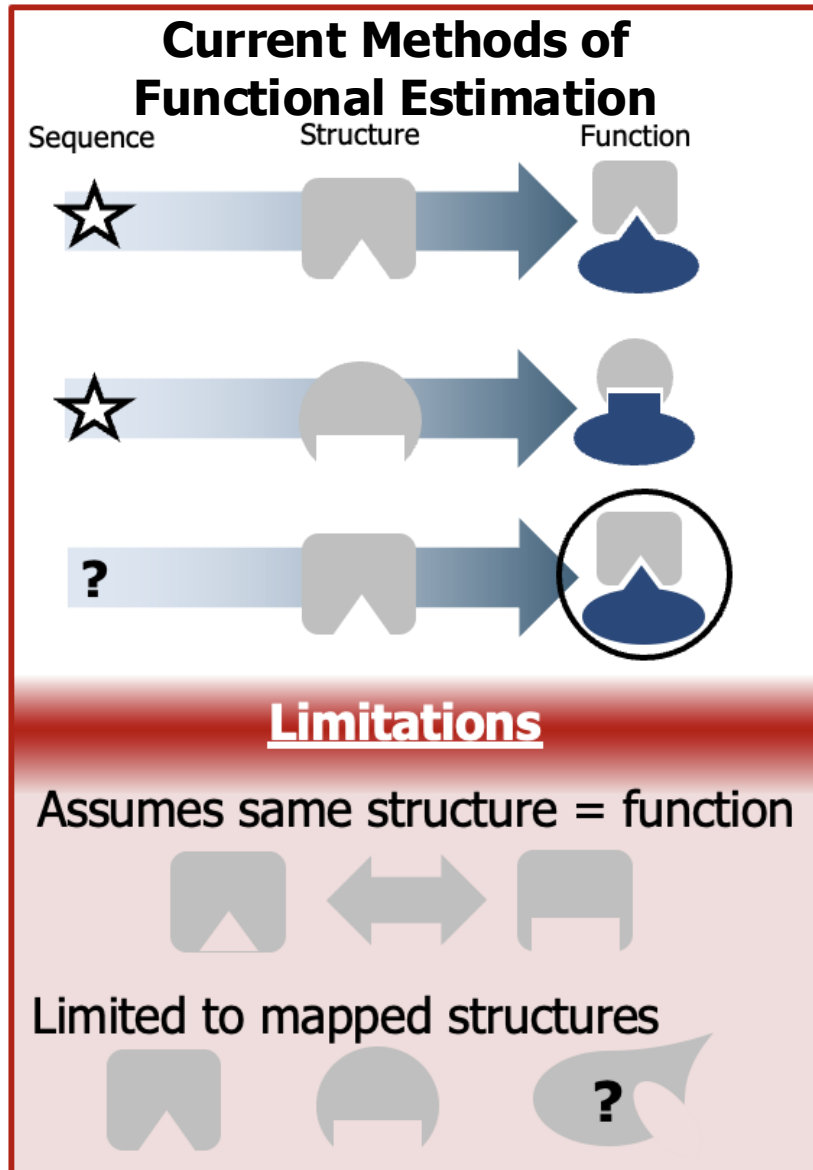
RNA: Ribonucleic Acid

CUI: Controlled Unclassified Information

Vision: Deep technology to predict multiple unknown protein functions in **1 hour** and with **>90% confidence**, available in both **class** and **CUI forms** with a public-facing User Interface



Incomplete Protein and Bioagent Characterization Offers Surprise at Scale

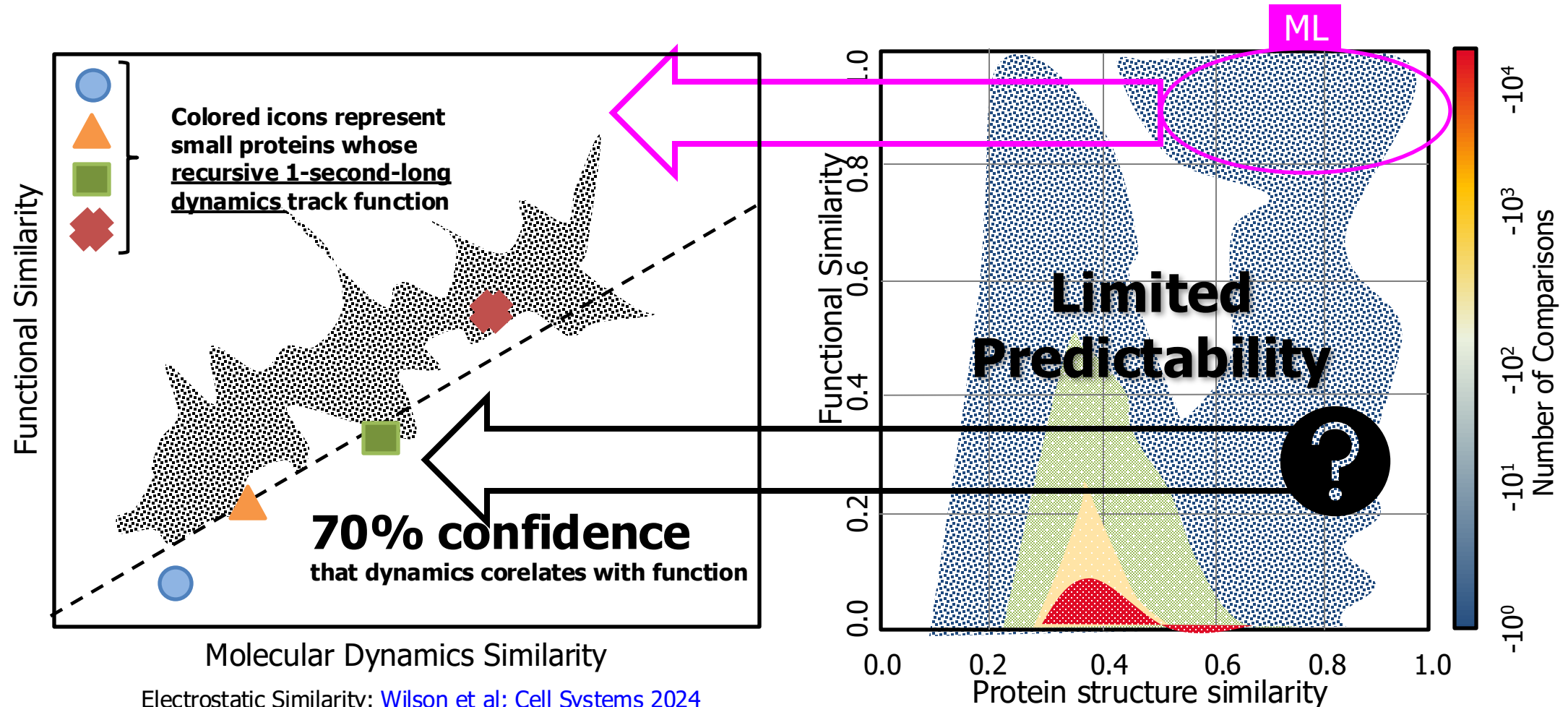


Sequence-Structure-Function relationship: [Leman et al; Nature Communication 2023](#)

Current predictions are limited to structural comparisons with exhausted datasets; limiting function assessment



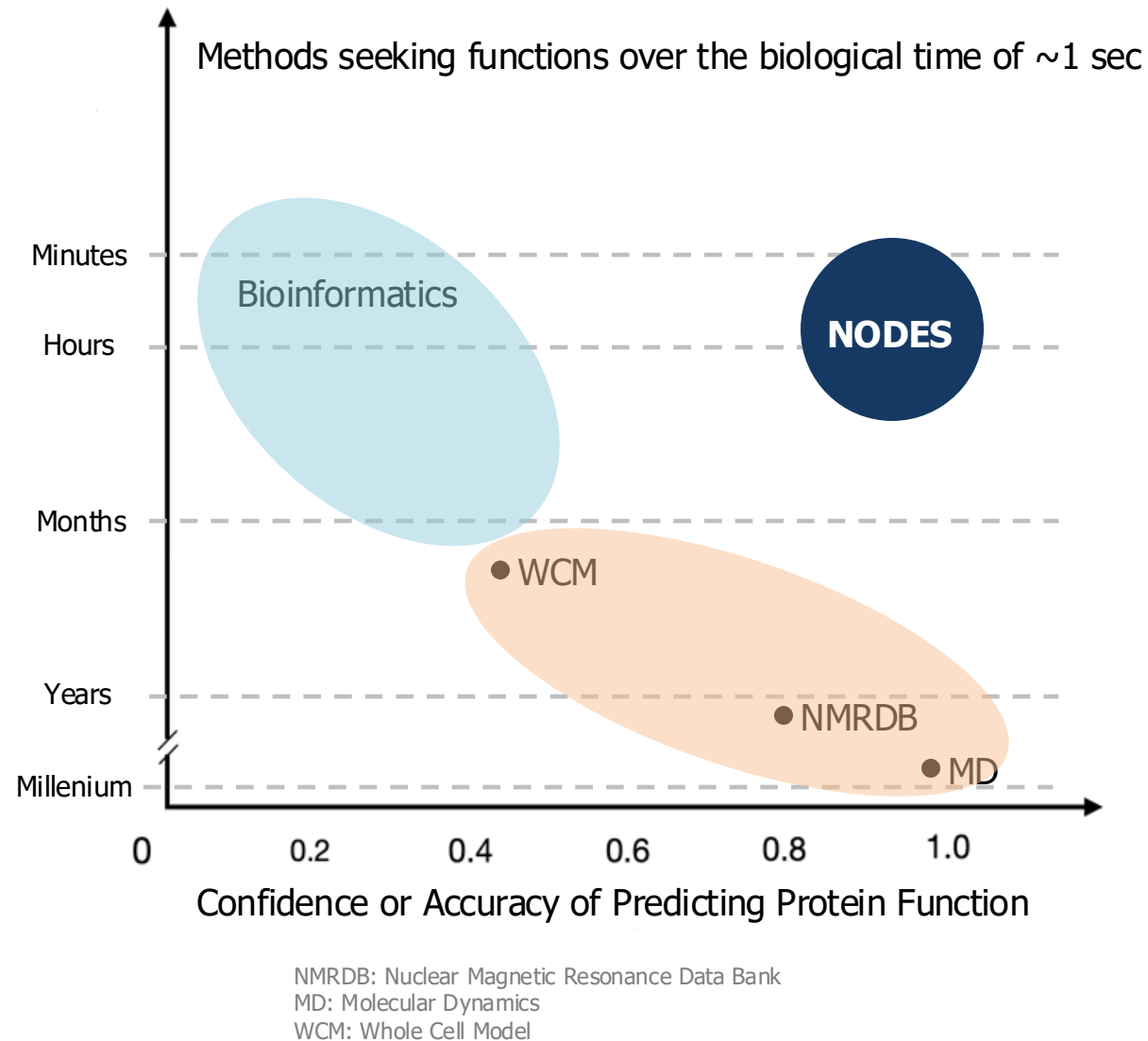
Hypothesis: Movement is the Best Predictor of the Possible Functional Space



Dynamics captures both system as well as the context (surroundings) in any number of dimensions



Current Technologies Cannot Resolve Balance between Accuracy and Speed

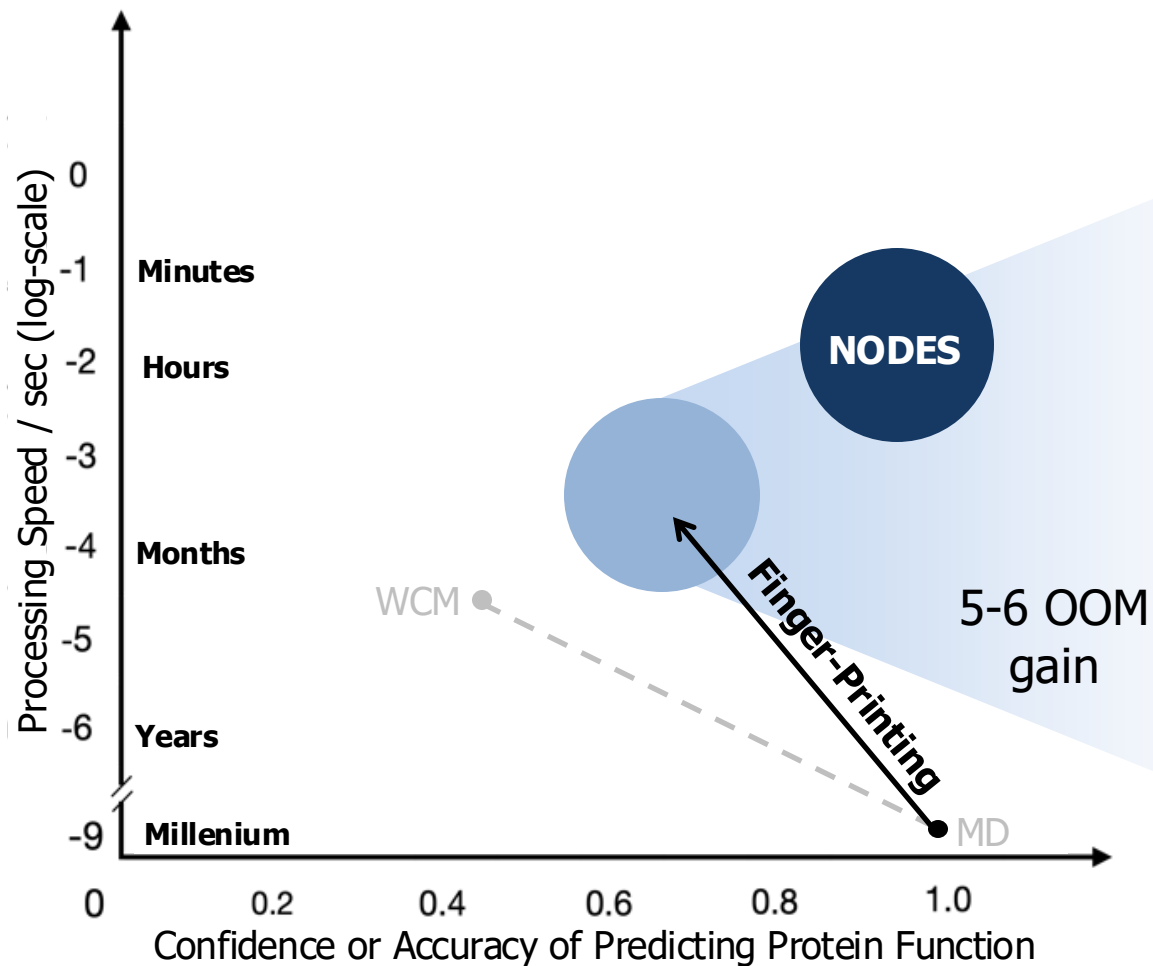


	Acceptable	Ok	Poor		
	Speed	Accuracy	Dynamics	Dimension	Context
PSI-PRED	Acceptable	Poor	Poor	Poor	Poor
AF2/RF	Ok	Poor	Poor	Ok	Poor
EVO/ESM	Ok	Poor	Poor	Acceptable	Ok
NMRDB	Ok	Acceptable	Acceptable	Poor	Ok
WCM	Ok	Ok	Ok	Acceptable	Ok
MD	Poor	Acceptable	Acceptable	Acceptable	Acceptable

Current models utilize stationary structures, have limited accuracy for **capturing molecular dynamics**, and have input size limits



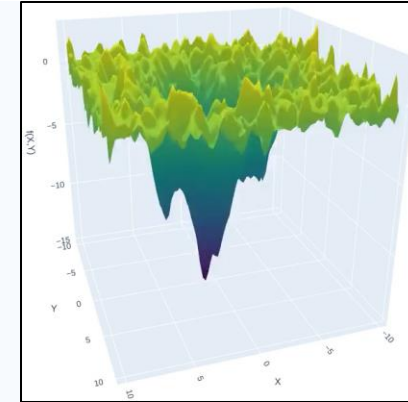
Enabling Tech 1: NODES for Rapid Fingerprinting Protein Function



Speed accelerated by lowering model accuracy to acquire fingerprint; accuracy can be regained through Machine Learning

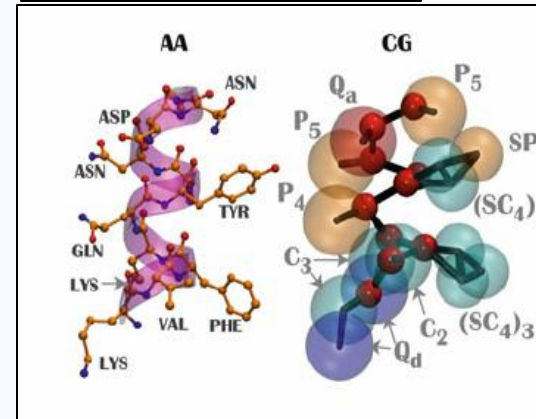
MD: Molecular Dynamic
OOM: Orders of Magnitude

WCM: Whole Cell Model



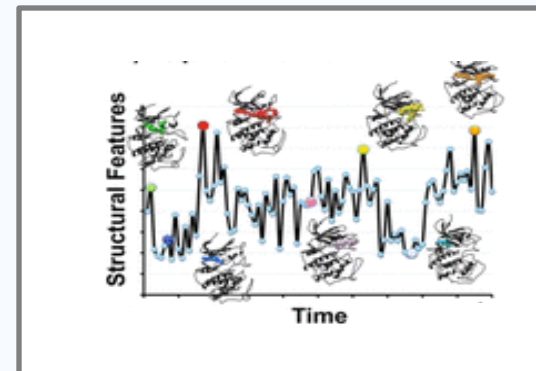
Annealing energy
(decreasing N)

3 OOM gain



Coarse-graining
atoms
(decreasing N)

3 OOM gain

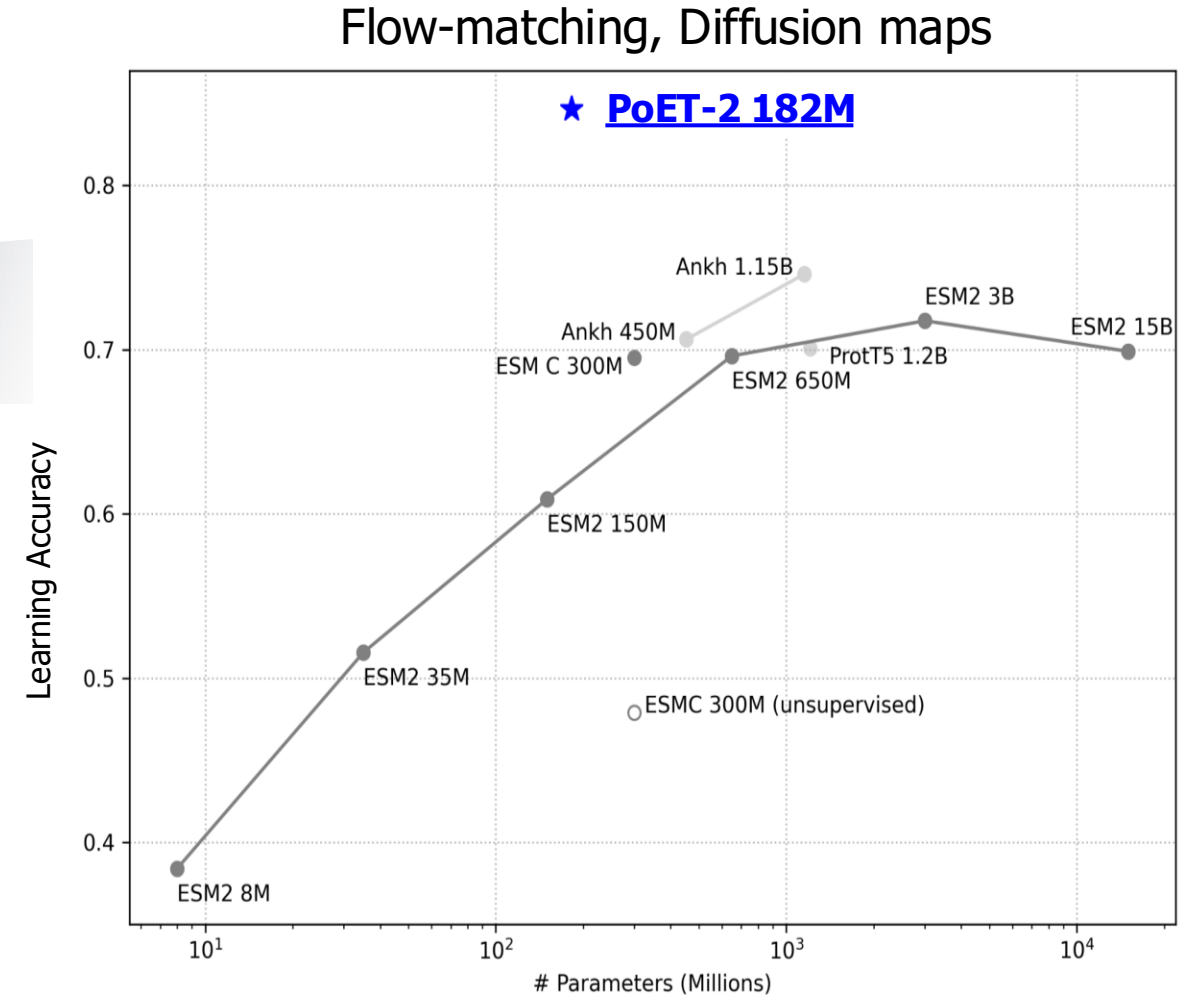
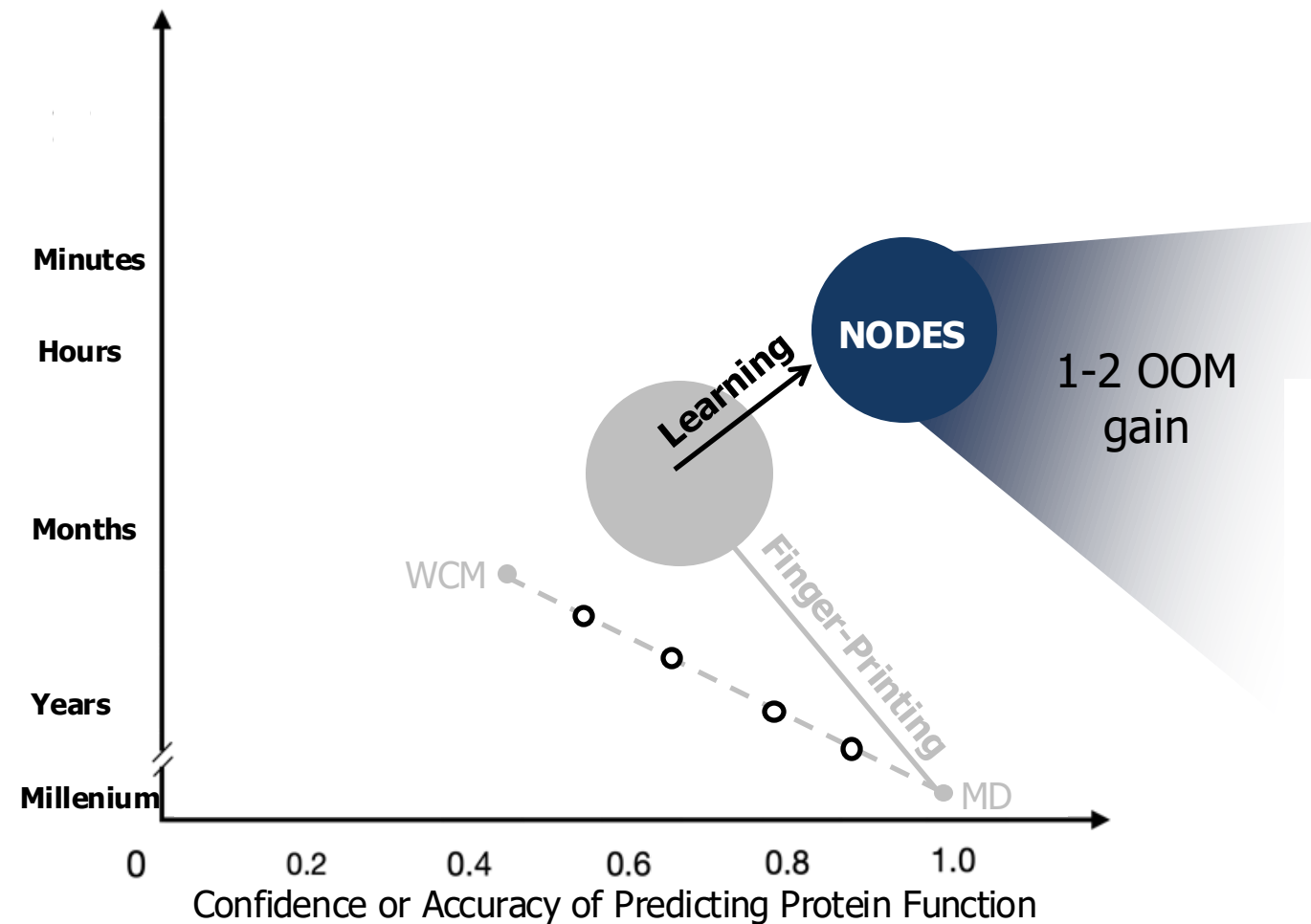


Accelerating time
(increasing Δ)

3 OOM gain



Enabling Tech 2: Recovering Accuracy and Prediction of Bioagent Function



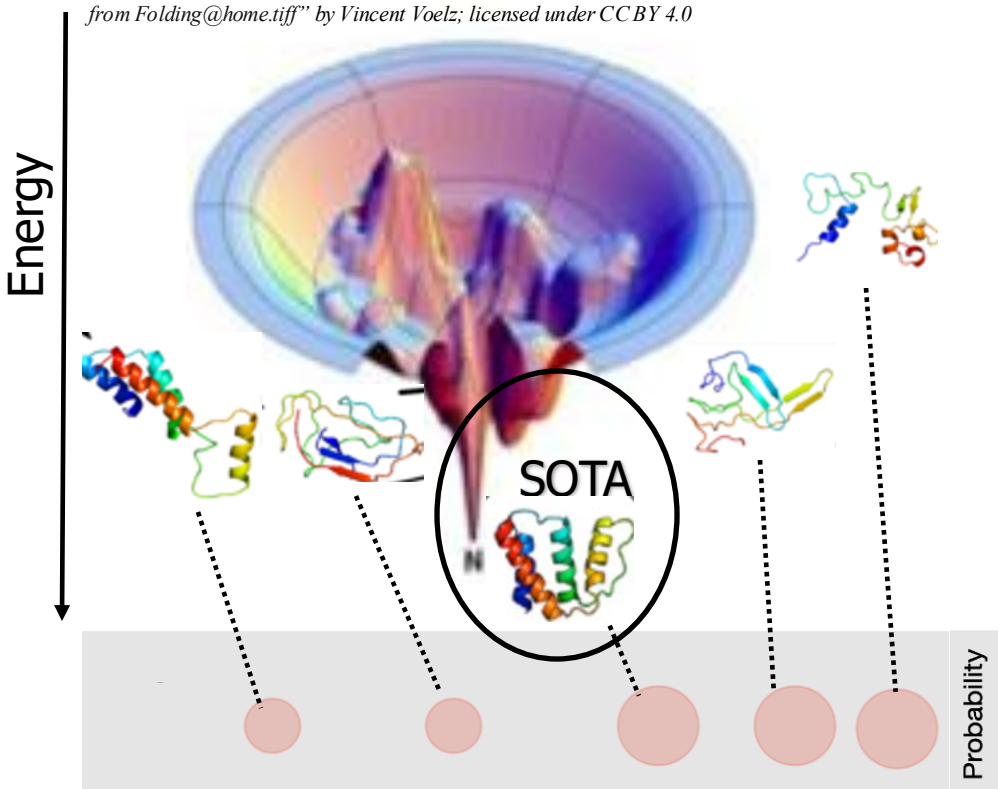
Prediction of Folding Patterns. Yang et al. Sci. Rep. (2013)

This grammar is conserved across protein finger-printing methods, and allows us to bring resolution back



Insight - Dimensionality Reduction Can Come to the Rescue

Modified from "Funnel-shaped energy landscape" by Ken A. Dill; and ACBP protein from "ACBP MSM from Folding@home.tiff" by Vincent Voelz; licensed under CC BY 4.0



(5x8-dimensions)

Network of Optimized Dynamic Energy Signatures

of protein functions: structure, energy, shape, solubility, pH, salt, pressure, temperature...

Proteins spent 99% of lifetime in low-energy states, performing majority of functions. Each state has distinct signatures.

Number of signatures grow with context complexity of functions



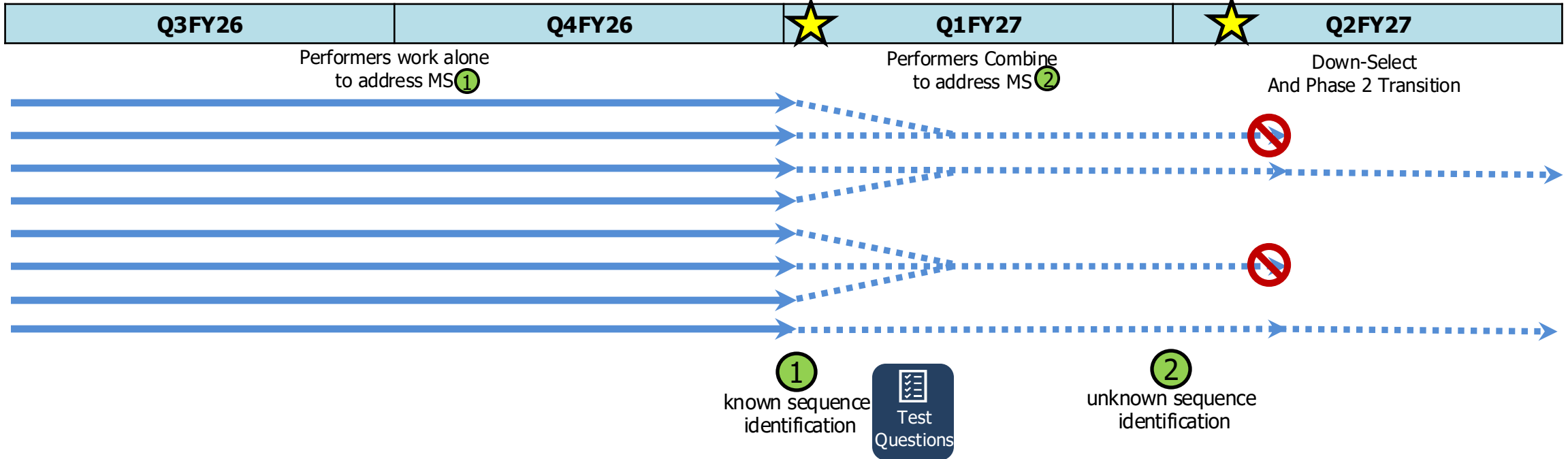
To Do list: [Chan et al; Curr.Opn.Strct.Biol. 2025](#)

DISTRIBUTION A: Approved for Public Release, Distribution Unlimited



NODES Program Structure – Phase I

Phase I (12 months)



T&E: Performer HPC Access (ORNL)

T&E: Model development on Government computational servers (ORNL)

Commercial Partner Engagement & Recruitment

★ PI Meeting / Demo Day



Test and Evaluation (T&E)

- Phase I: performers to undergo two capability demonstrations
 - 6 months – Capability Demonstration 1
 - 9 months – Capability Demonstration 2
 - Performers will work with DARPA-partnered Government partners when conducting CD1 and CD2
 - Performers' codes will need to run on Government partners supercomputers to conduct the tests
 - Output results will be evaluated by DARPA and the Government partner
 - Performers are expected to achieve goals and metrics, which will be shown next



Capability Demonstration 1 (CD1) at 6 months

Test performers' models for the ability to accurately predict function

Test set:

- protein sequences provided by DARPA
- known proteins with published function

Deliverables:

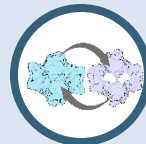
- quantitative results (graphs/data/charts) for the tests, and model's performance

Performers are expected to reach 60% accuracy (notional) of obtained results when compared to published literature, in the following tests:



• **Functional Area 1: folding**

- Thermodynamic weights
- Secondary structural content
- Predicted local distance difference test (pLDDT) scores
- Gene Ontology (GO) annotations on protein function



• **Functional Area 2: binding**

- Binding constants
- Interface interaction maps
- Root Mean Square Deviation to known structures



• **Functional Area 3: allostery**

- Binding constants
- Interface interaction maps
- Root Mean Square Deviation to known structures



Model performance (notional)

- <300 GPU hours per protein
- Total model inference time \leq 1 week



Capability Demonstration 2 (CD2) at 9 months

Test performers' models for the ability to accurately predict function

Test set:

- protein sequences provided by DARPA
- unpublished proteins for which the sequences are unknown to the performers

Deliverables:

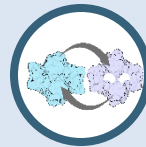
- quantitative results (graphs/data/charts) for the tests, and model's performance

Performers are expected to reach **70% accuracy (notional) of obtained results** when compared to published literature, in the following tests:



• **Functional Area 1: folding**

- Thermodynamic weights
- Secondary structural content
- Predicted local distance difference test (pLDDT) score



• **Functional Area 2: binding**

- Binding constants
- Interface interaction maps
- Root Mean Square Deviation to govt-provided structures (unknown to performers)

• *Note that there will be no testing for Functional Area 3 (FA3).*

Model performance (notional)



- <300 GPU hours per protein
- Total model inference time \leq 1 week



Guidelines

- **Abstracts are required**

- 5-page limit for Phase I
- 1-page limit for Phase II (will not be evaluated)
- Feedback on whether an Oral Proposal Package (OPP) is requested, will be provided by DARPA
- Abstract step allows feedback to proposers on whether a proposed approach or technology fits, and prevents effort spent by proposers on assembling a proposal that may not be successful

- **OPP**

- Specific instructions (including content submission guidelines) will be provided in the invitation to participate
- Oral presentations are expected to be held in-person (encouraged) over the course of 1-2 days in October 2025 in the Washington, DC metro area
- Oral presentations are anticipated to be scheduled for 60 minutes each, allowing for a strictly limited 40-minute presentation time, and up to 20 minutes of questions and answers



Elements of a Successful Technical Proposal

Goals of the project: Develop fast and accurate MD models to probe the limitless space of de novo protein sequences, provide a tool to expedite threat characterization when the nation or warfighters are introduced to an unknown agent, shorten the time to developing Medical Countermeasures (MCMs), as well as support biomedical research by providing expedited ways to understand infections.

- Hypothesis corroborating, extending, or challenging the foundational principles of NODES.
- Narrative supporting the novelty and uniqueness of the proposed simulation as well as of AI/ML approaches for seeking signatures of molecular dynamics. The text shall demonstrate a thoughtful integration of simulation, learning and generation to avoid incremental duplication of current efforts.
- Benchmark for computational efficiency, preferably on national supercomputers or comparable architectures. Both hard and soft scaling are encouraged.
- Estimates of simulation time and number of simulation systems.
- Proposed methods for correlating protein functions with dynamics, as well as methods of generalizing movement beyond single families of proteins.
- Choice of loss functions, features and representations.



Elements of a Successful Technical Proposal (2)

- Preliminary evidence of feasibility of the proposed innovations over the proposed period of Phase I. Publications are encouraged but not required. Demonstration of how prior research was leveraged to maximize the impact and value of the requested funding.
- Use of workflows(s) or figure(s) to depict the essence of the proposed solution.
- Specific plans, including cost, time estimates, and teaming composition to address all functional areas outlined in the program description.
- Identified risks to successful execution and fulfillment of program goals and proposed strategies for mitigating these.
- Description of plans to meet and/or exceed program metrics and milestones; these claims must be justified with literature-based explanations, data, and projections.
- Technology transfer package plan to transfer data, computational models, etc. to the USG team at the end of each phase of the program.



Anticipated Timeline*

*Subject to change in program announcement

- **PS Posting Date:** 31 July 2025

<https://sam.gov/workspace/contract/opp/303f7a218cb44e6193d573ae9a2d83bd/view>

- **Proposers' Day:** 1 August 2025 (today)

- **Frequently Asked Questions Submittal Deadline:** 15 August 2025

Answers will posted here: <https://www.darpa.mil/research/programs/nodes>

- **Abstract Due Date:** 29 August 2025

- **Abstract Feedback Anticipated:** Early-Mid September 2025

- **OPP Due Date:** 7 October 2025

- **Estimated Period of Performance Start:** Jan/Feb 2026



Reminders

- **Submit early**
 - NODES abstracts are **mandatory** and **decisional**
 - Submission of full oral proposal package (OPP) **will only be accepted if requested by DARPA**
 - Adhere to the abstract and OPP guidance and templates provided in the PS
 - Corroborate all technical claims with analytical or experimental evidence
 - Submit only one proposal per lead organization
 - Participate in multiple proposing teams (as subs), if desired
 - Refer to the PS for guidance when in doubt
- **Teaming Profiles** - If you were not able to present a Lightning Talk, we encourage you to fill out the Teaming Profile Template provided to you via email and submit it to NODES@darpa.mil by **COB Tuesday, August 5**

OPP: Oral Proposal Package
PS: Program Solicitation

Questions: **NODES@darpa.mil**



- **Dr. Abhishek Singharoy, Program Manager**
 - Dr. Cristina Russo, Technical SETA
 - Dr. Taylor Engdahl, Technical SETA
 - Dr. Aaron Short, Technical SETA
 - Ms. Shawn Rich, Business Financial Manager (BFM)
 - Ms. Natasha Woods, Program Security



Happy to Answer Any Questions?