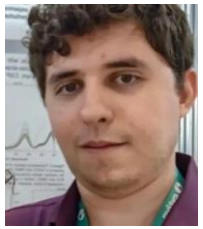
Antonio Oliveira



Murilo Sanches

Larissa Adolfo

Exploring the Energy Landscapes of Intrinsically Disordered Proteins

Vitor Leite

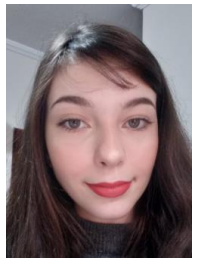
Departamento de Física

São Paulo State University

São José do Rio Preto, SP - Brasil

Ingrid Martins

Rafael Viegas



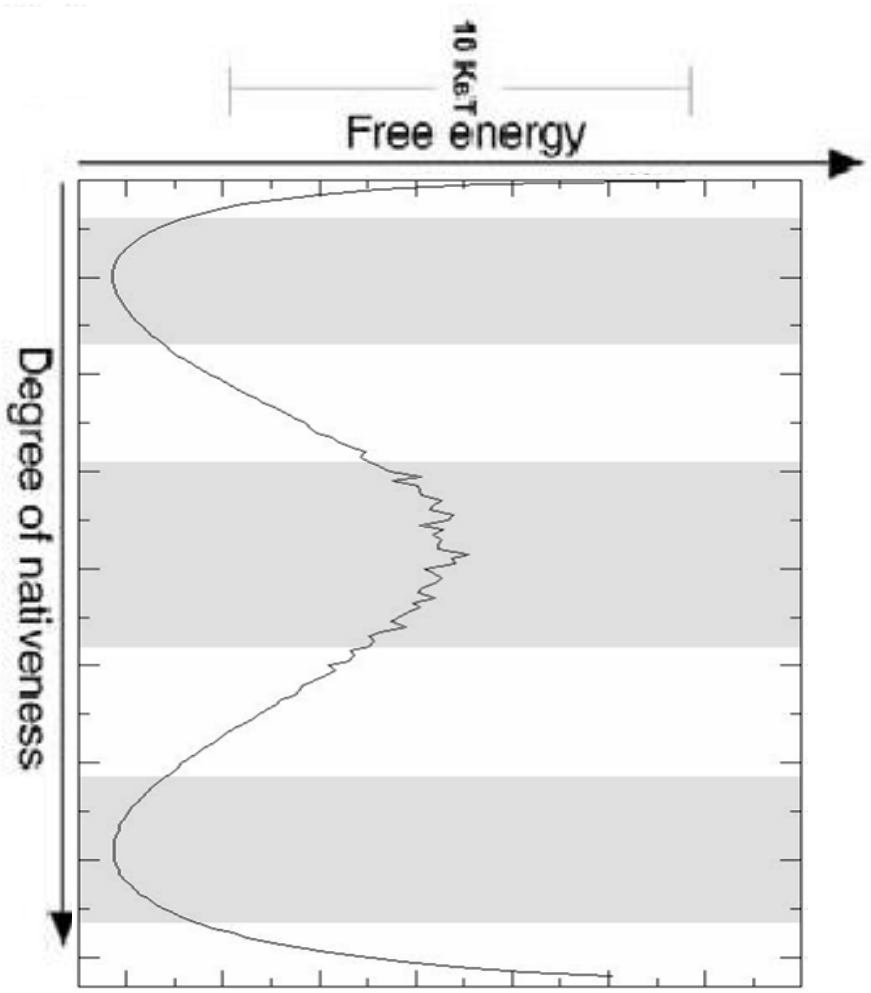
Juliana Camargo

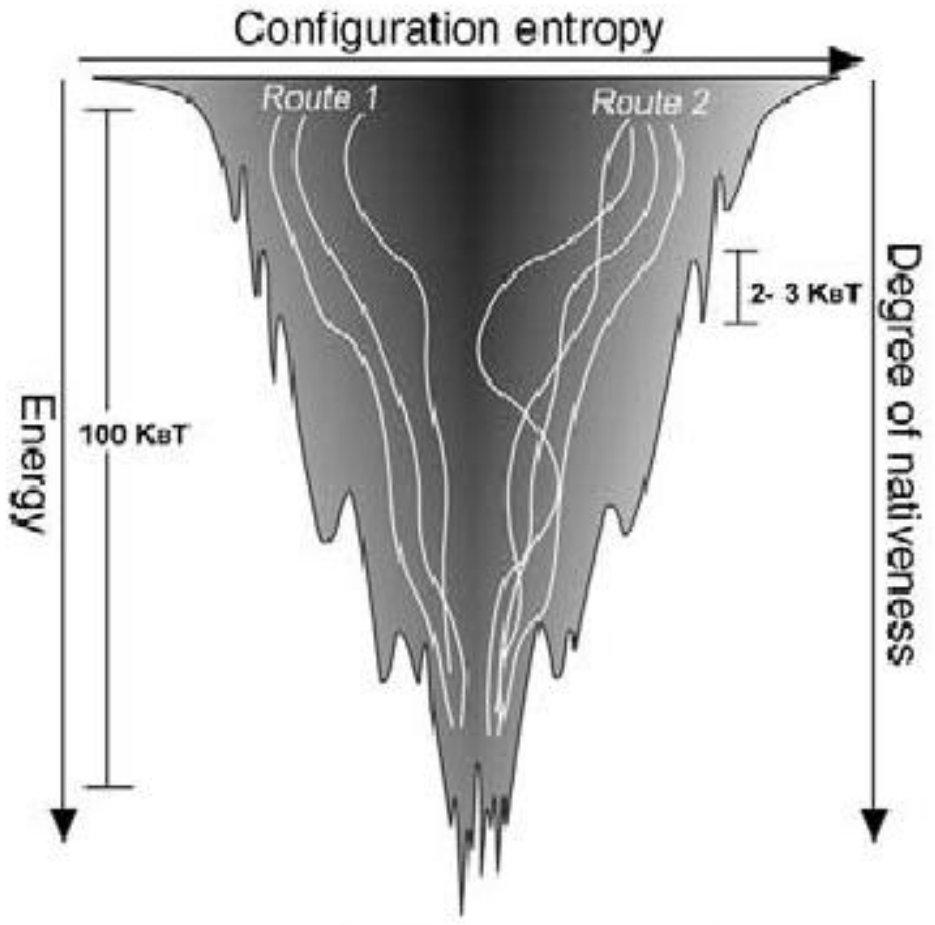
ICTP SAIFR

Current Topics in Molecular Biophysics (CTMB3)

Instituto Principia – SP – 7/10/2024

Protein Folding Energy Landscape – General Picture



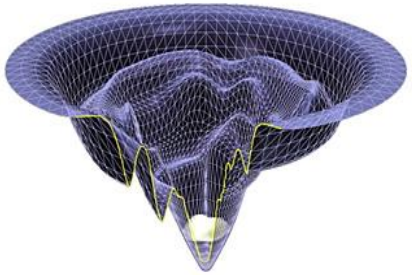


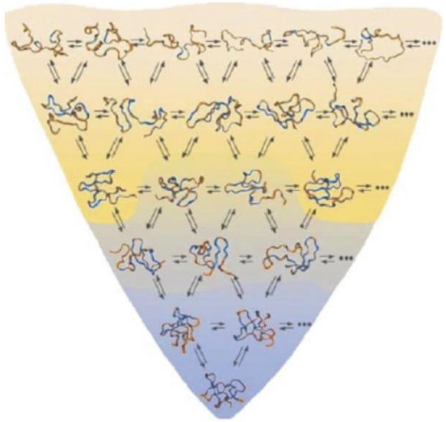
$$F = \overline{E} - T\overline{S}$$

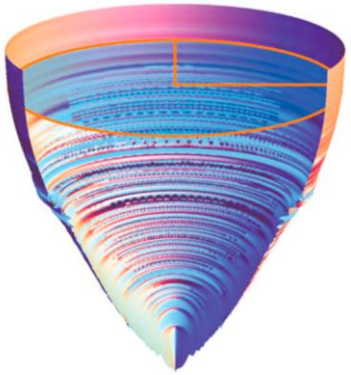
JN Onuchic & PG Wolynes, *Curr. Opin. Struct. Bio.* 2004

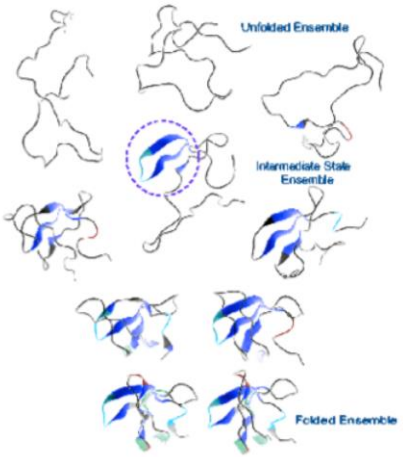
Protein Funnel Visualization

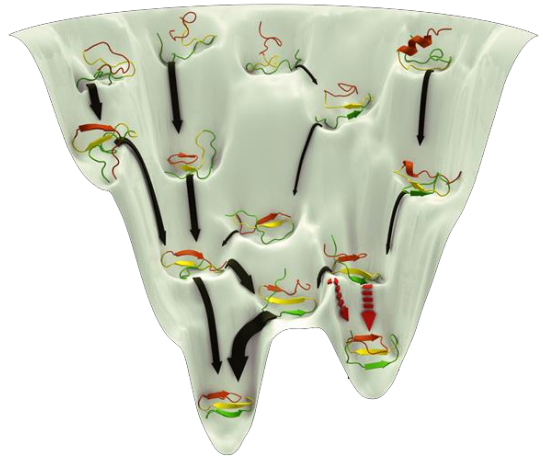
- Beyond 1D representation

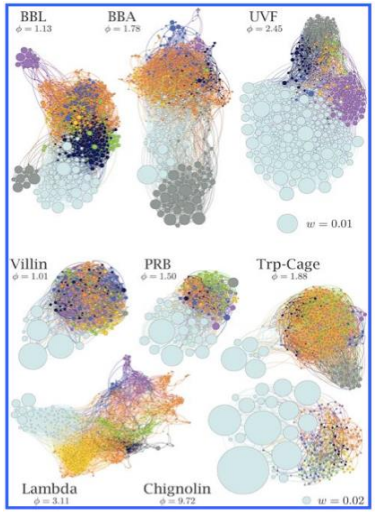


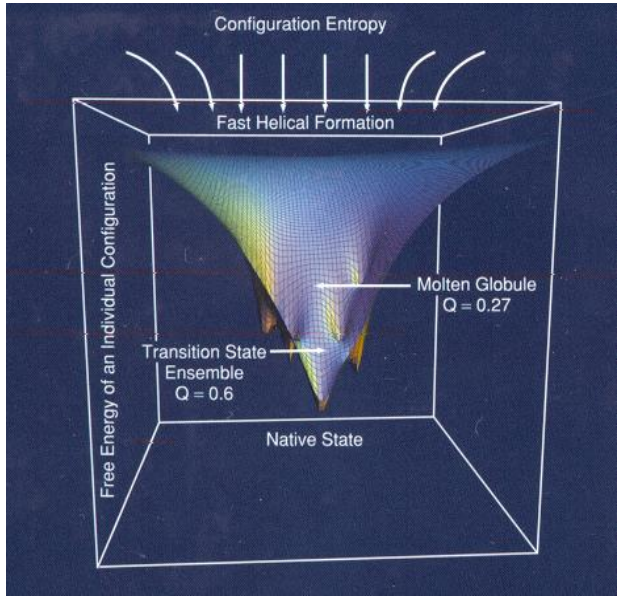


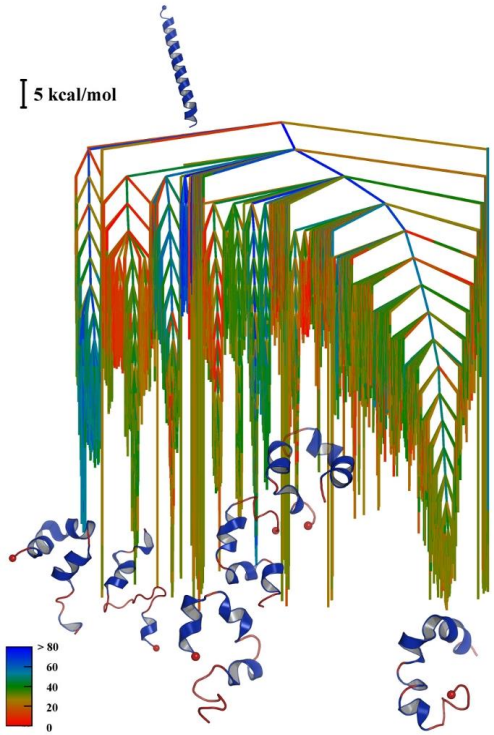












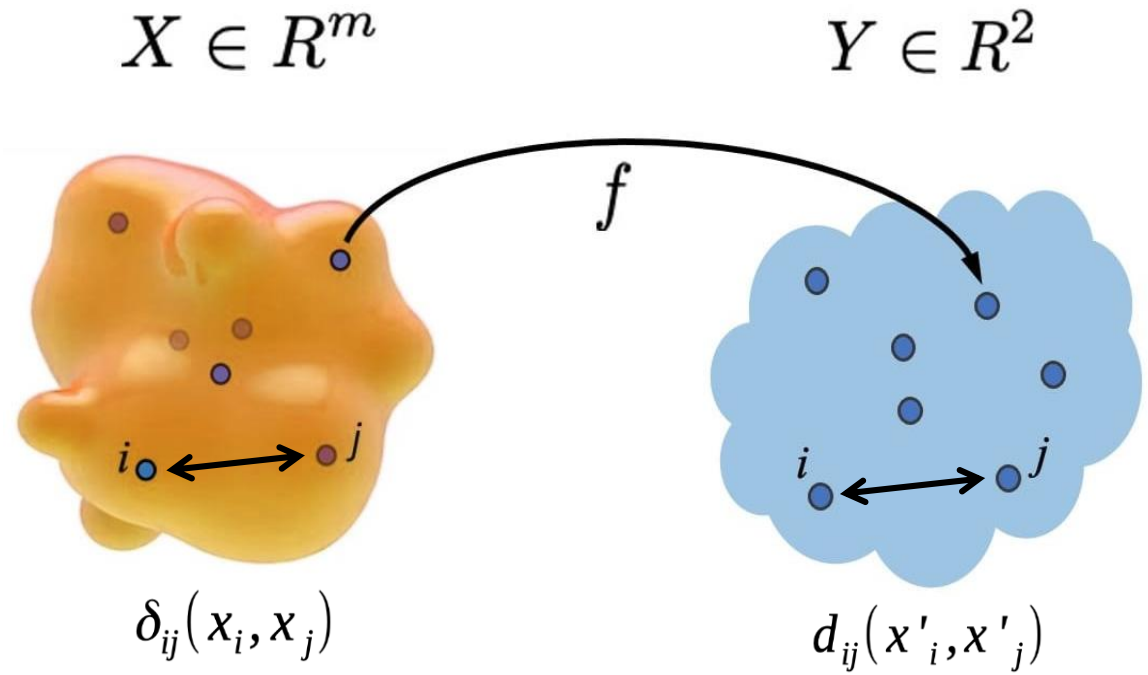
- Mechanism: distribution of transition paths
- Meta-stable states
- Degenerative diseases
  (beta-amyloid Alzheimer's disease)

Energy Landscape Visualization Method (ELViM)

Multidimensional Projection

A.Oliveira Jr., V. Leite, *et al*
JCTC (2019), PLOS One (2014)

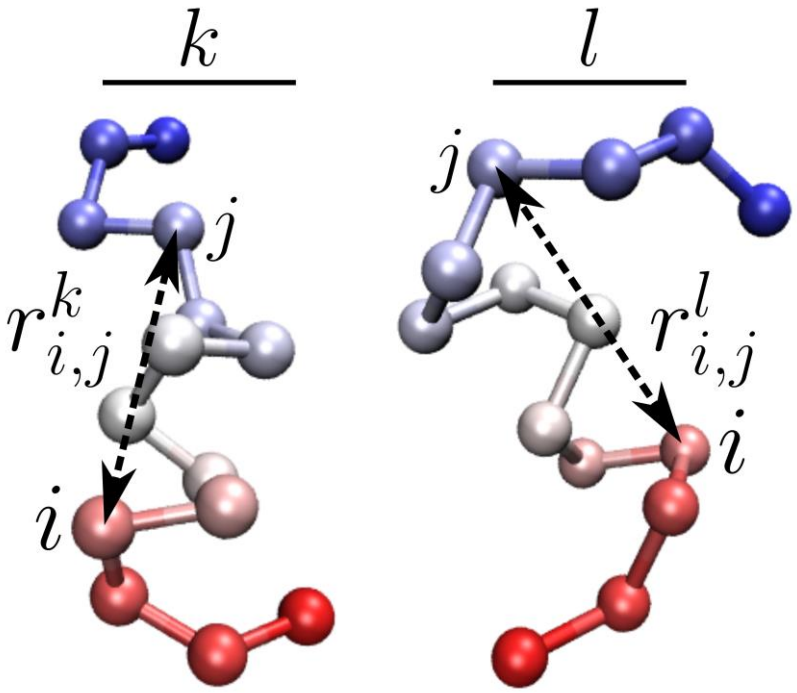
Viegas, Leite, *et al*
JCIM (2024)



$$f : X \to Y \iff |\delta_{ij} - d_{ij}| \approx 0 \qquad \text{for all } (i,j) \text{ pairs}$$

Metrics

- Effective distance between any two configurations
- Based in internal distances between its elements (C-α)



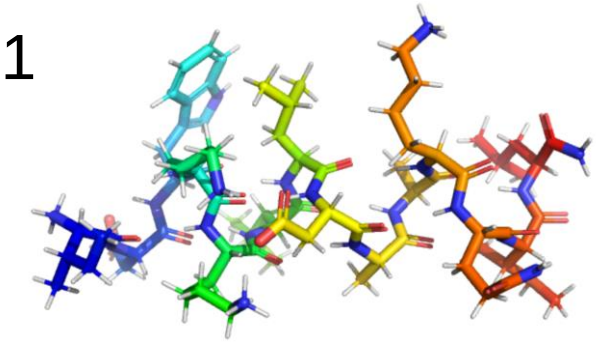
$$q_w^{k,l} = \frac{1}{N_p} \sum_{i,j \in pairs} \exp\left[\frac{-(r_{i,j}^k - r_{i,j}^l)^2}{2\sigma_{i,j}^2}\right]$$

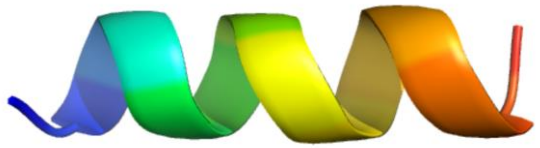
$$\sigma_{i,j} = \sigma_0 |i-j|^\epsilon$$

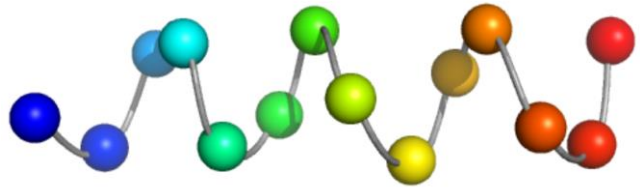
➔ $\delta_{k,l} = 1 - q_w^{k,l}$

Wolynes & Papoian, JACS 2003, PNAS 2004

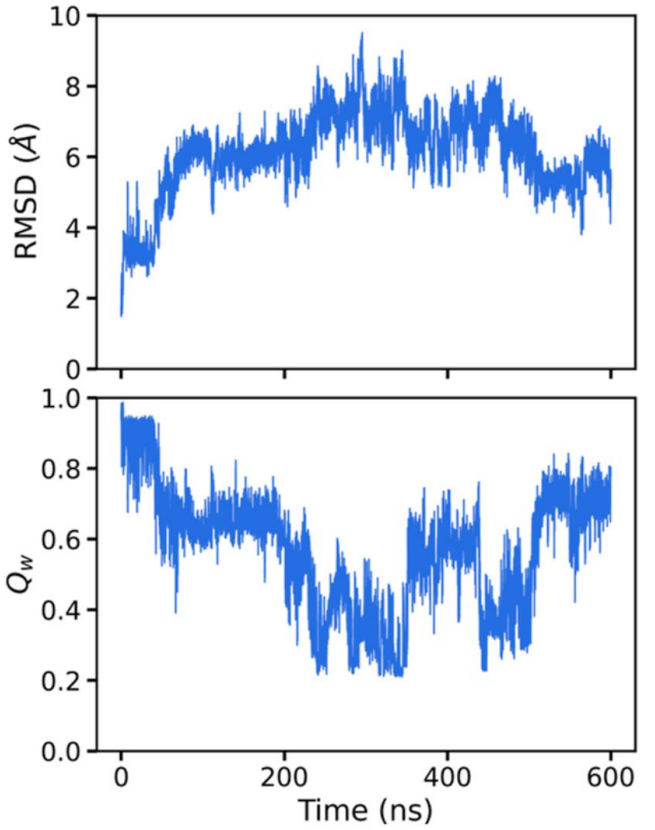
MP1

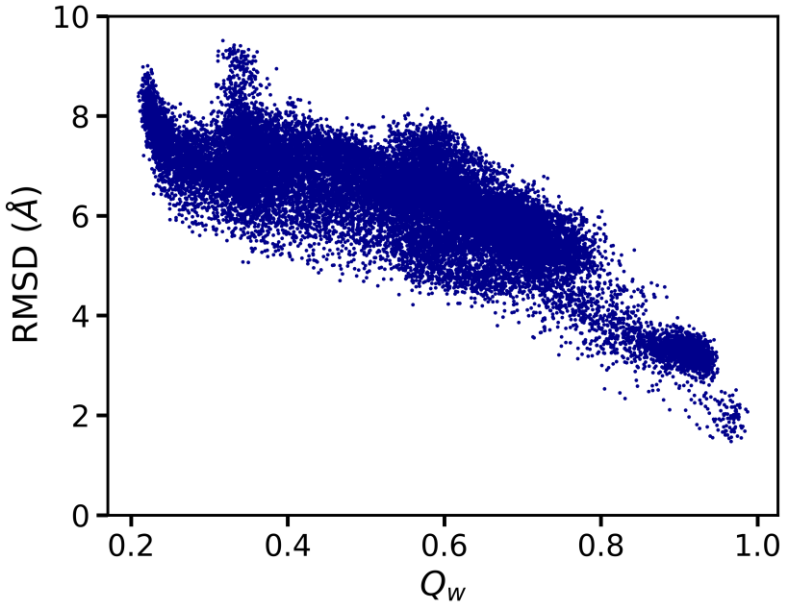


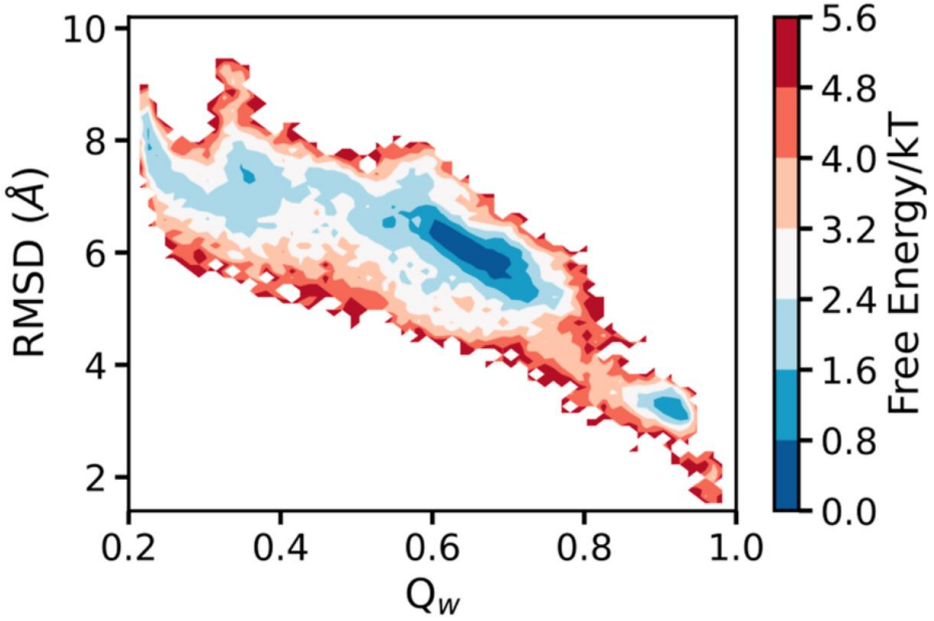




C_α structure-based model

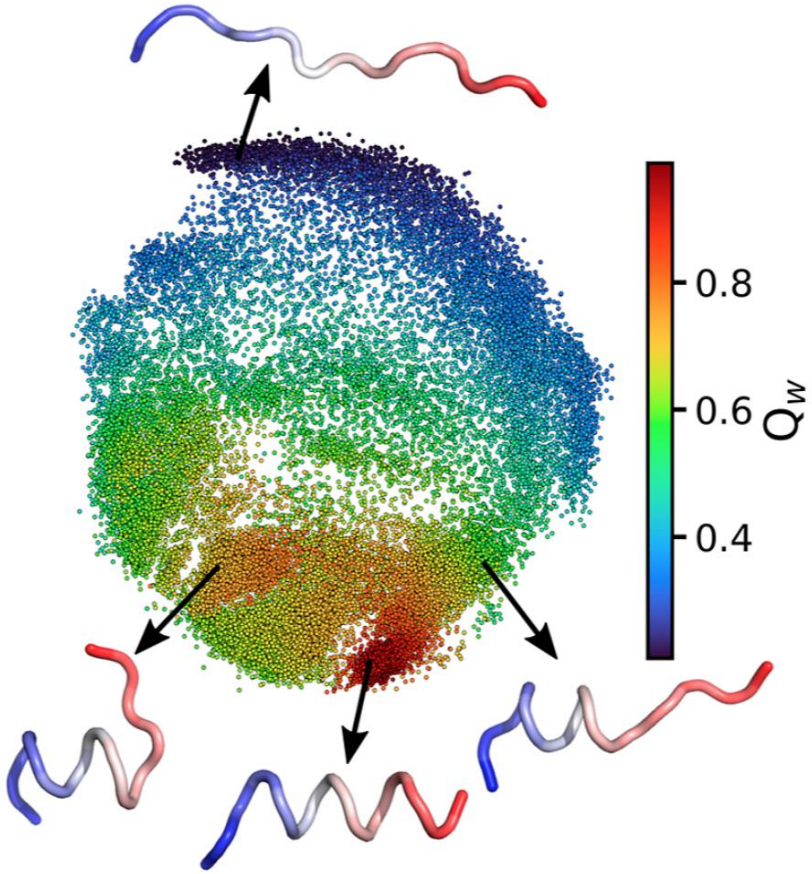


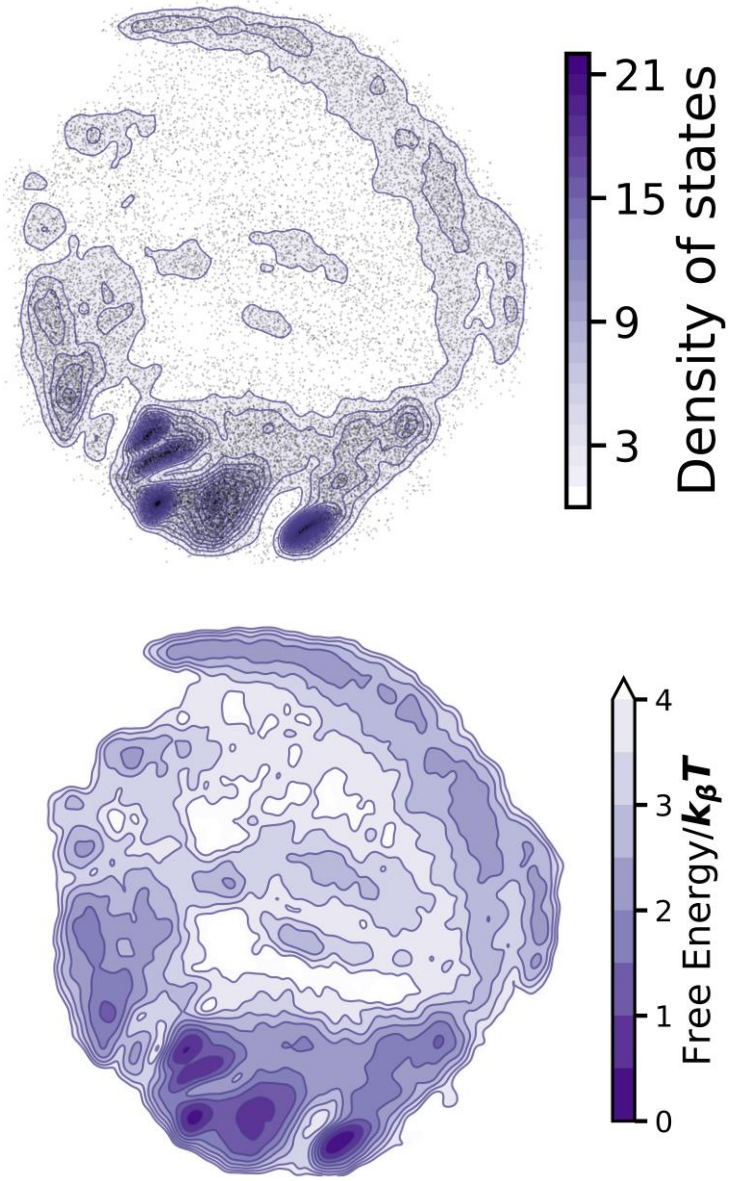


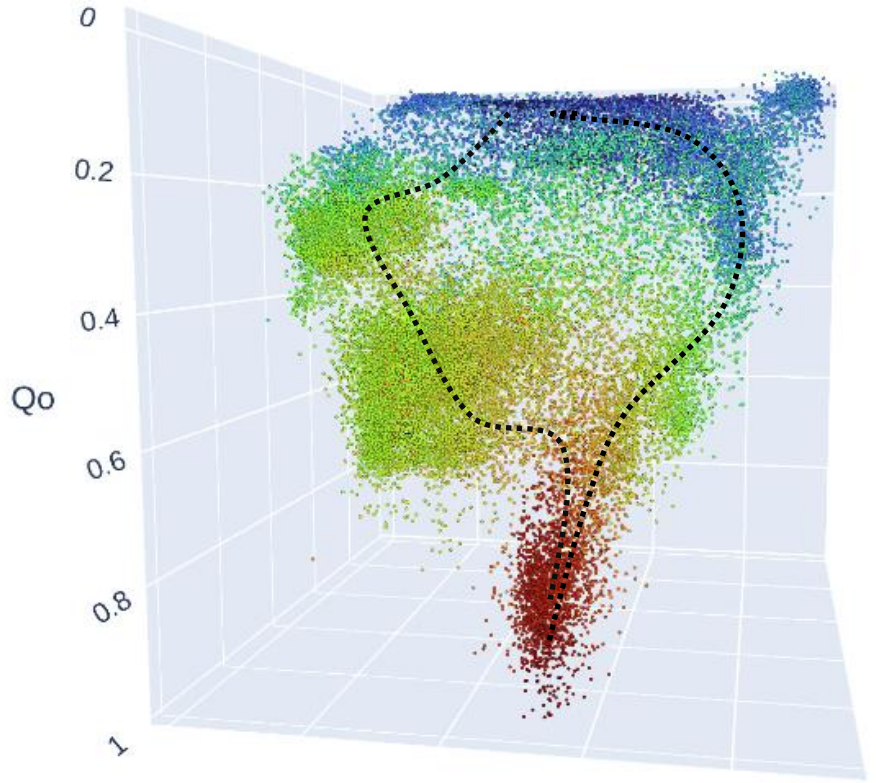


Viegas, Leite, *et al, JCIM (2024)*

ELViM: Density of States, Transition State Ensembles, Folding Routes, Meta Stable States…



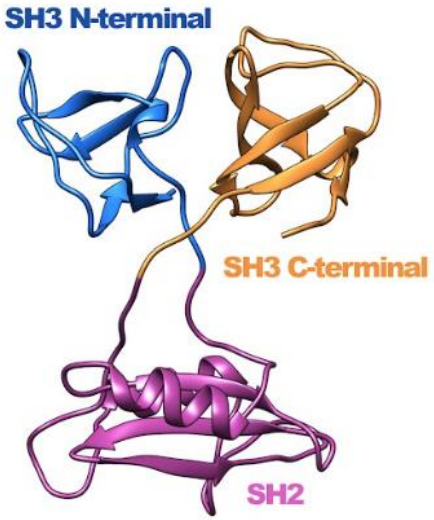


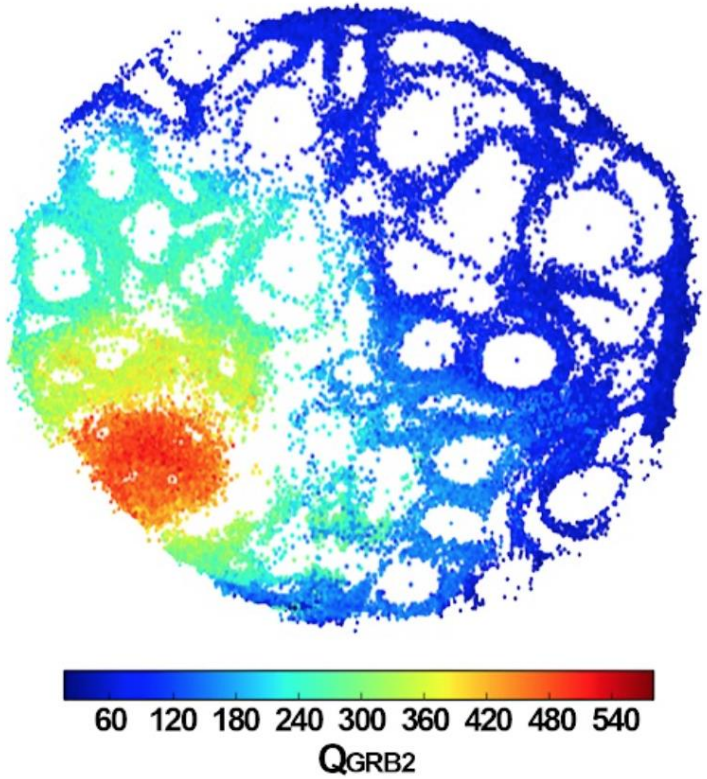


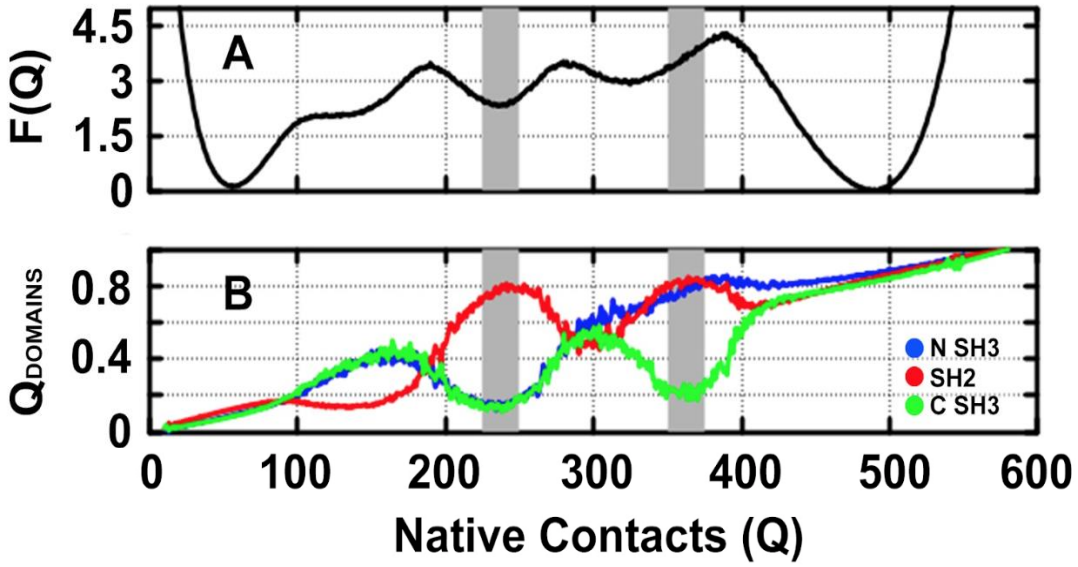
- Reaction-coordinate free method
- No need of reference conformation

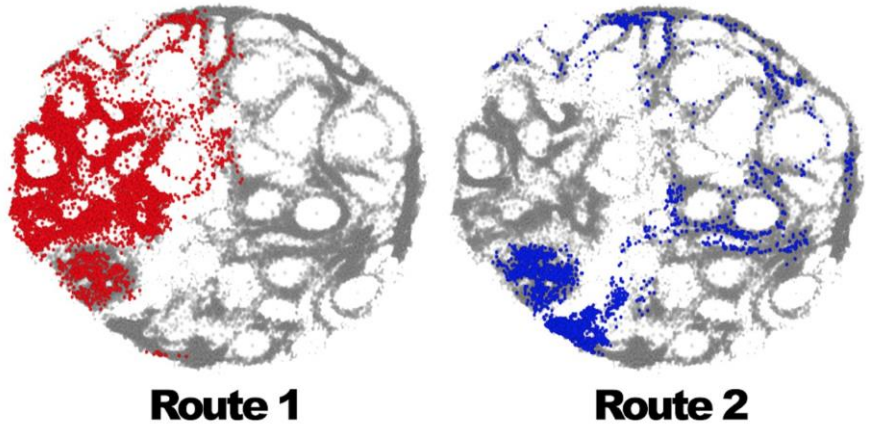
Viegas, Leite, *et al, JCIM (2024)*

GRB2 - Growth-fator receptor-bound protein 2





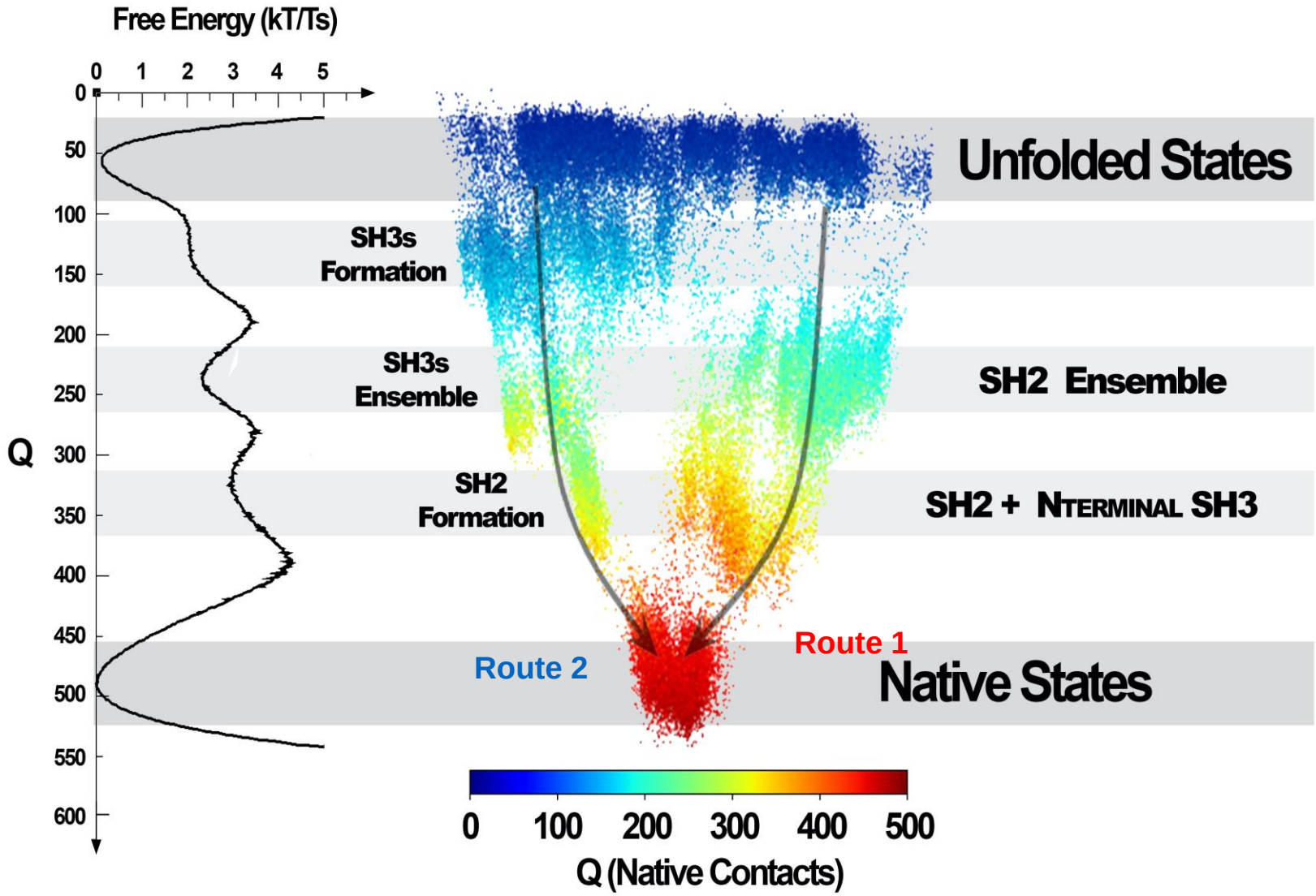




R. Dias, V. Leite *et. al., JCIM 2023*

GRB2 - Growth-fator receptor-bound protein 2

In a 3D view

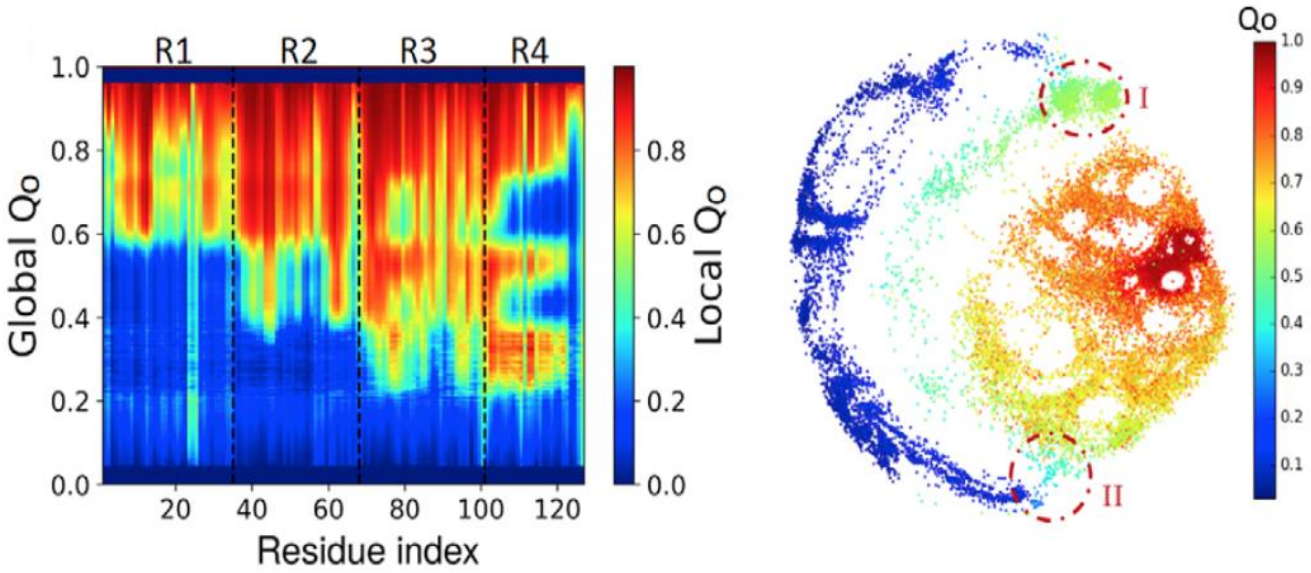


R. Dias, V. Leite *et. al., JCIM 2023*

Resolving the fine structure in the energy landscapes of repeat proteins

Murilo N. Sanches[1], R. Gonzalo Parra[2], Rafael G. Viegas[1,3], Antonio B. Oliveira Jr.[4], Peter G. Wolynes[4], Diego U. Ferreiro[5*] and Vitor B.P. Leite[1*]

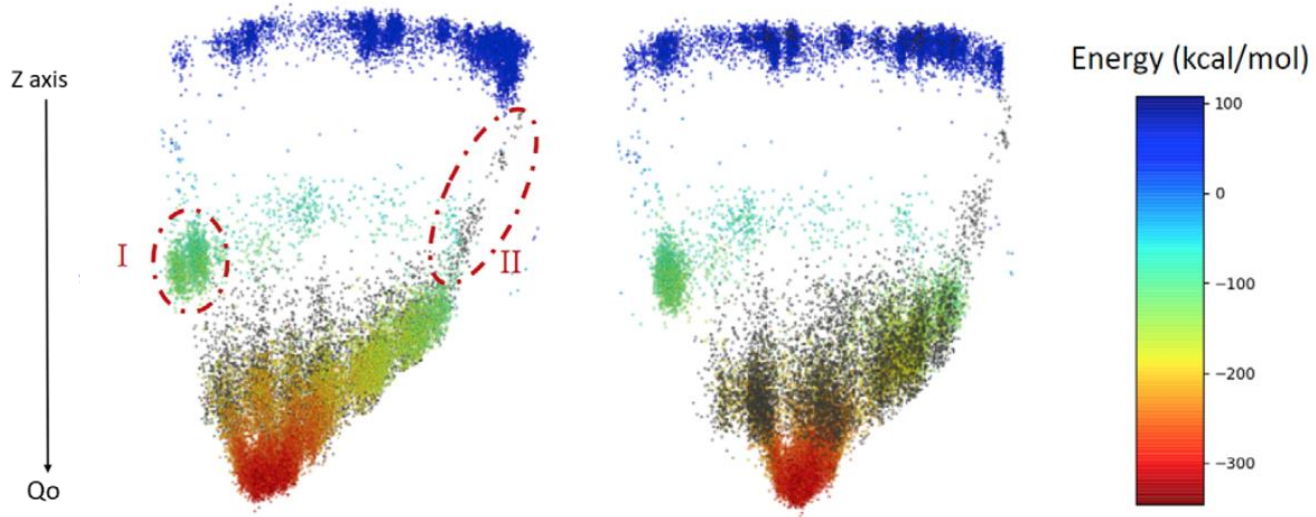
QRB Discovery 2022



3ANK

4ANK

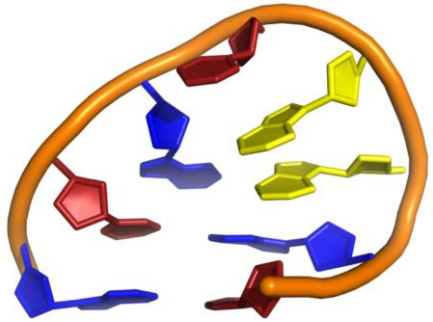
6ANK



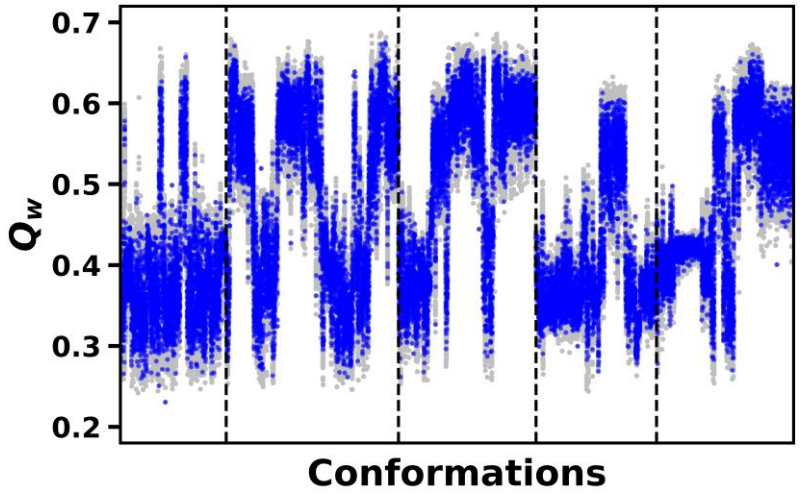
Characterizing the Energy Landscape of an RNA Tetraloop

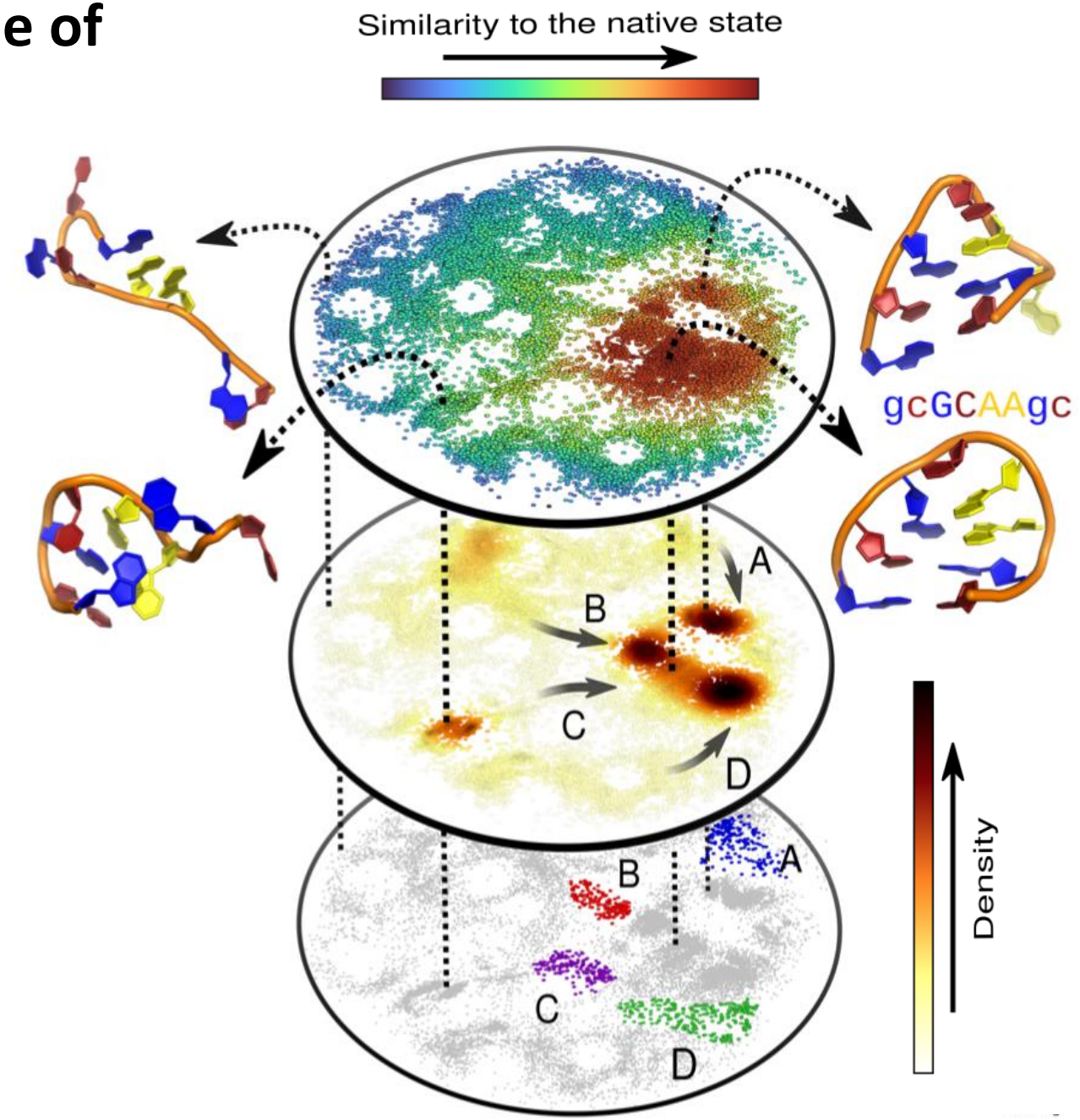
Rafael Viegas, Angel E. Garcia, VBPL, *et. al.*

JCIM 2023

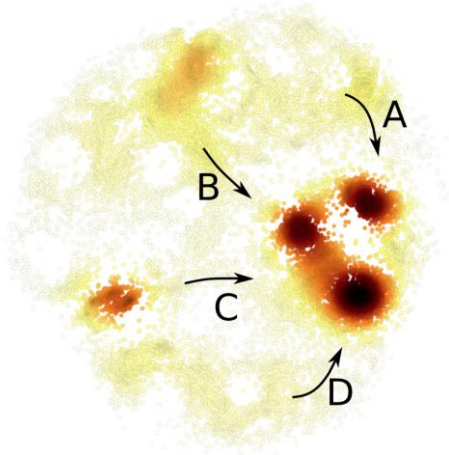


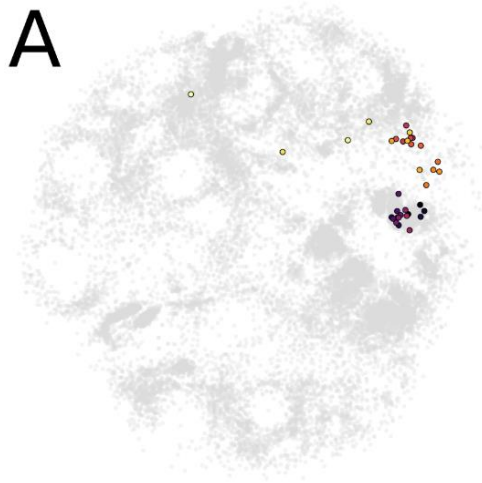
gcGCAAgc

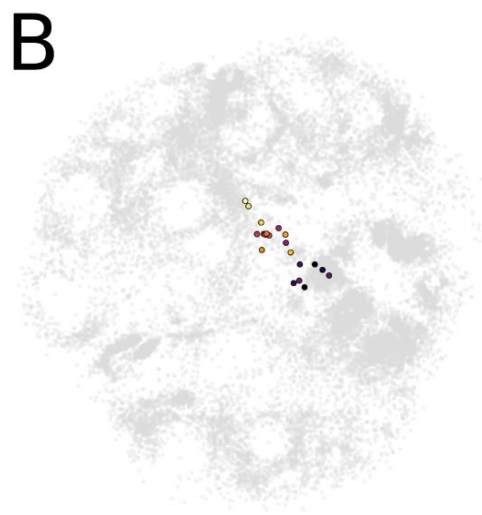


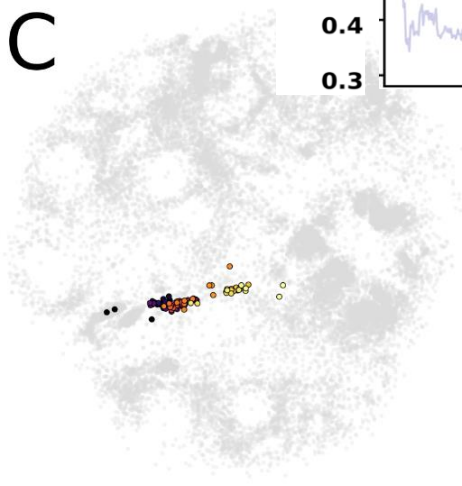


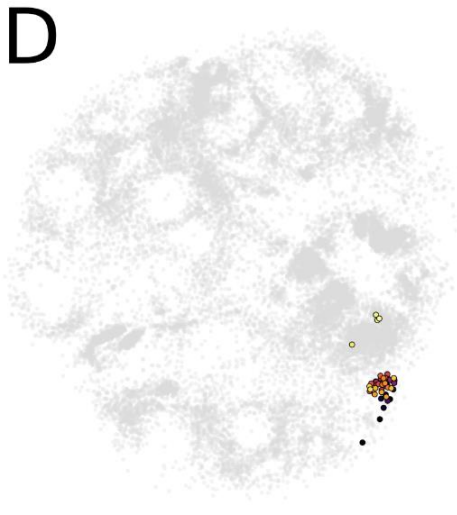
Characterizing the Folding Transition-State Ensembles of the RNA Tetraloop

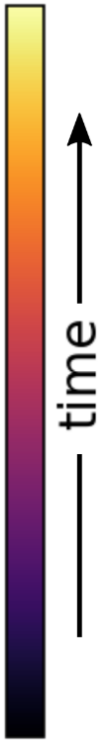


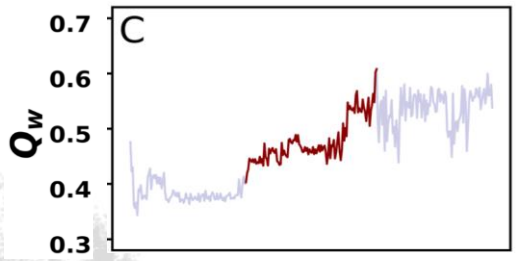


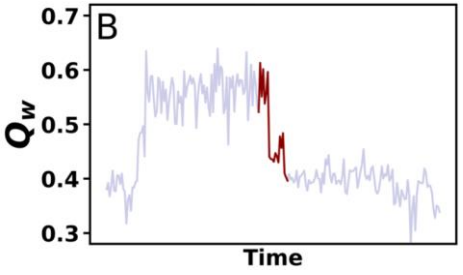


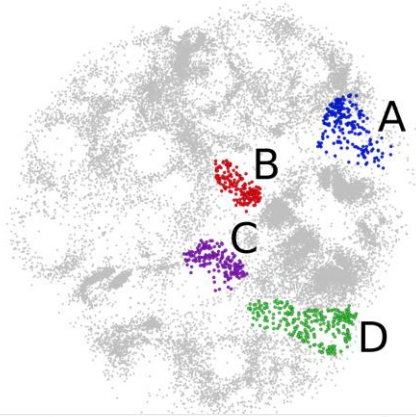


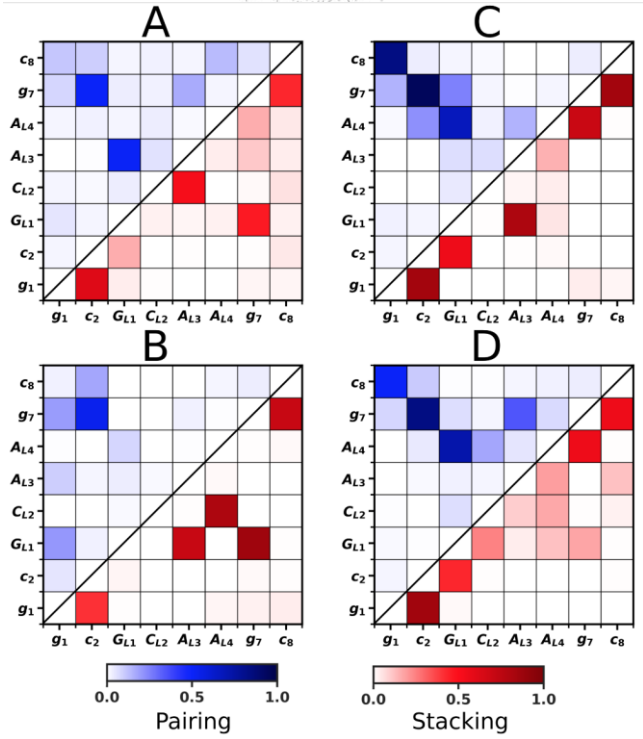


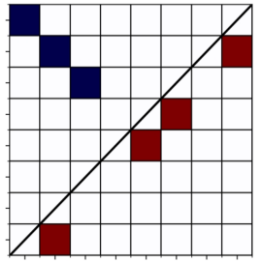












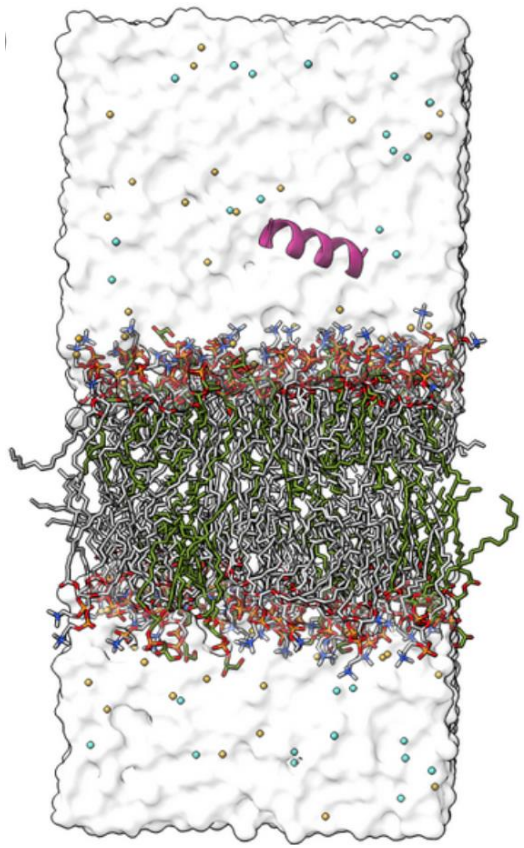
1ZIH
(native)

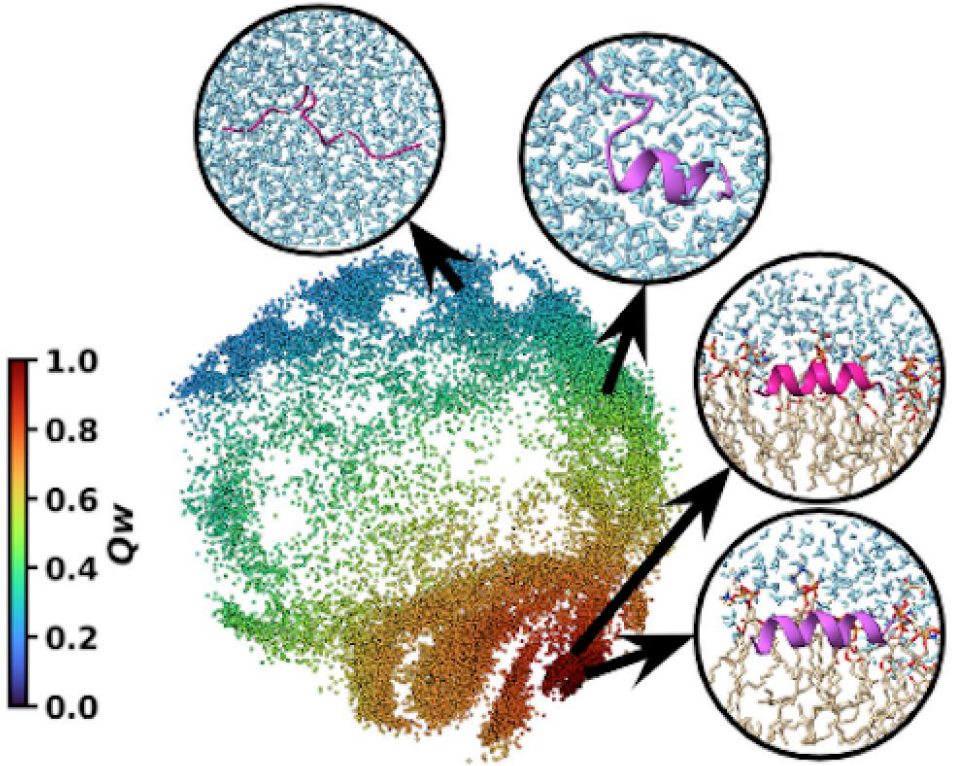
Probing Mastoparan-like Antimicrobial Peptides Interaction with Model Membrane Through Energy Landscape Analysis

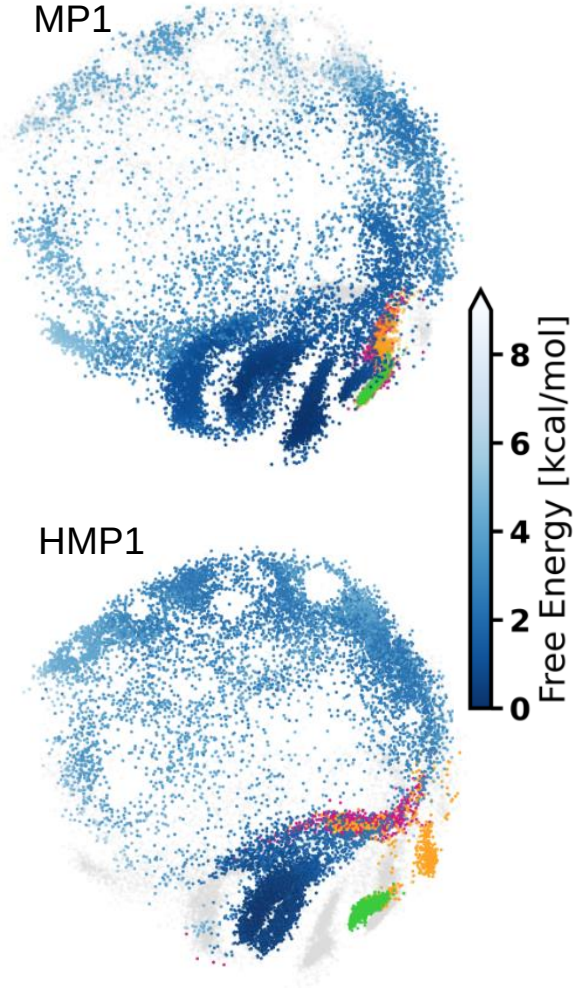
THE JOURNAL OF PHYSICAL CHEMISTRY B 2023

Ingrid B. S. Martins,[§] Rafael G. Viegas,[§] Murilo N. Sanches, Alexandre S. de Araujo, and Vitor B. P. Leite*

MP1 & HMP1

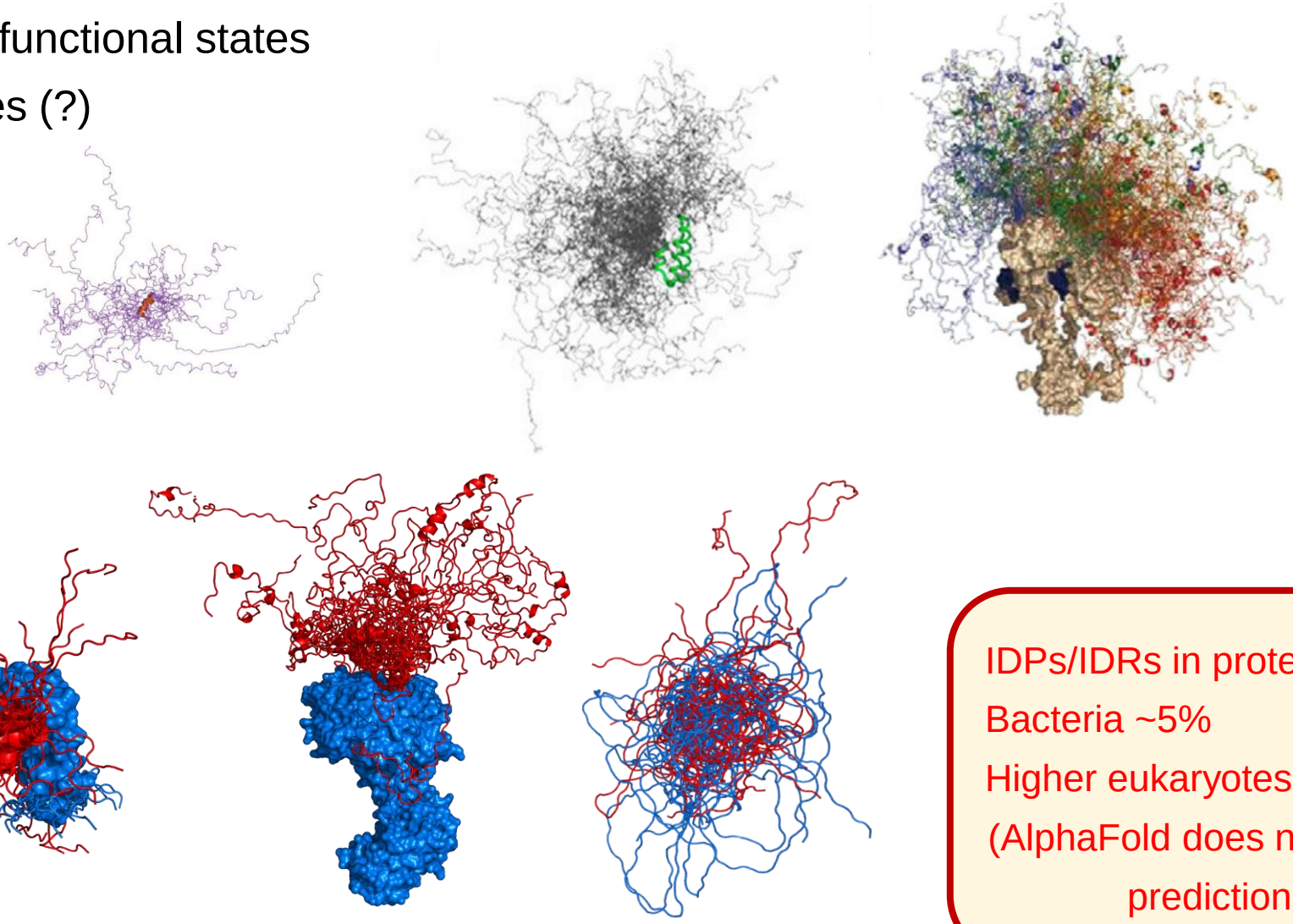






Intrinsically Disordered Proteins (IDPs)

- Representation of functional states
- Energy Landscapes (?)



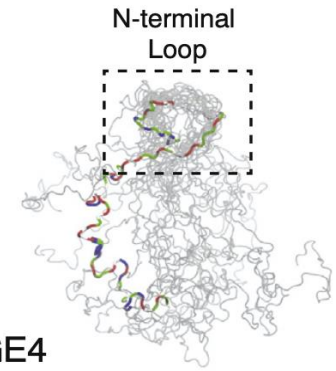
IDPs/IDRs in proteome:
Bacteria ~5%
Higher eukaryotes ~30-50%
(AlphaFold does not provide predictions)

Schuler *et al.*, *Curr. Opin. Struct. Biol.* 60, 66-76 (2020)

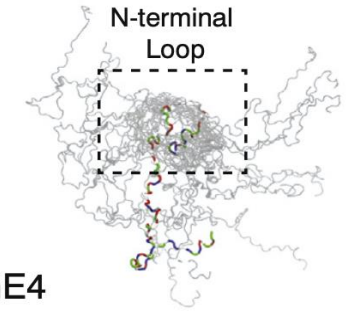
IDP: Prostate-associated gene 4 (PAGE4)

102 AA
Different phosphoforms:
(AWSEM Simulation)

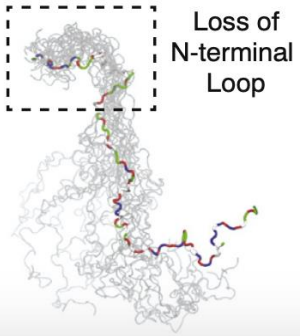
WT-PAGE4
(no phosphor.)



HIPK1-PAGE4
(2 phosphor.)



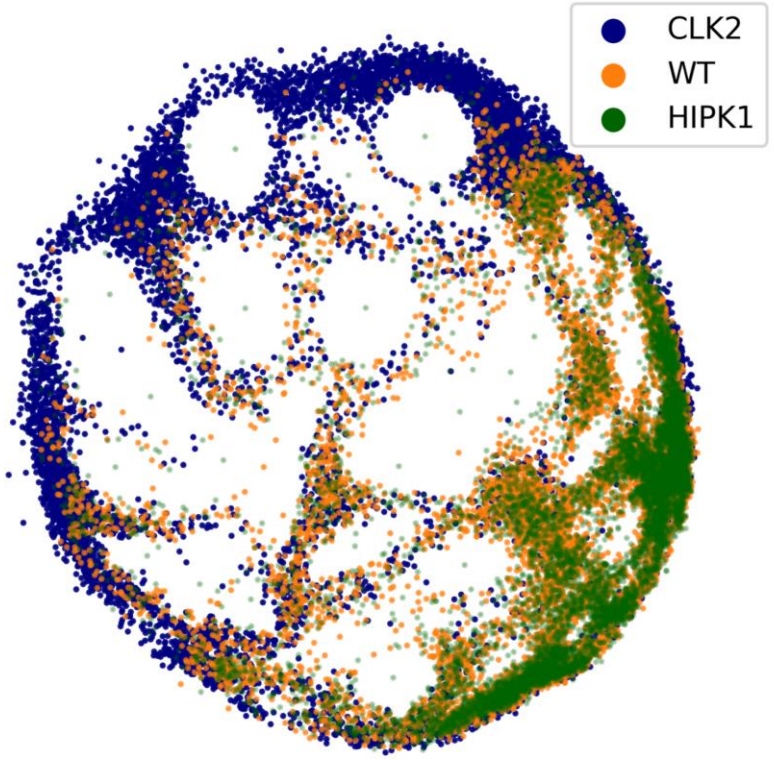
CLK2-PAGE4
(many phosphr.)



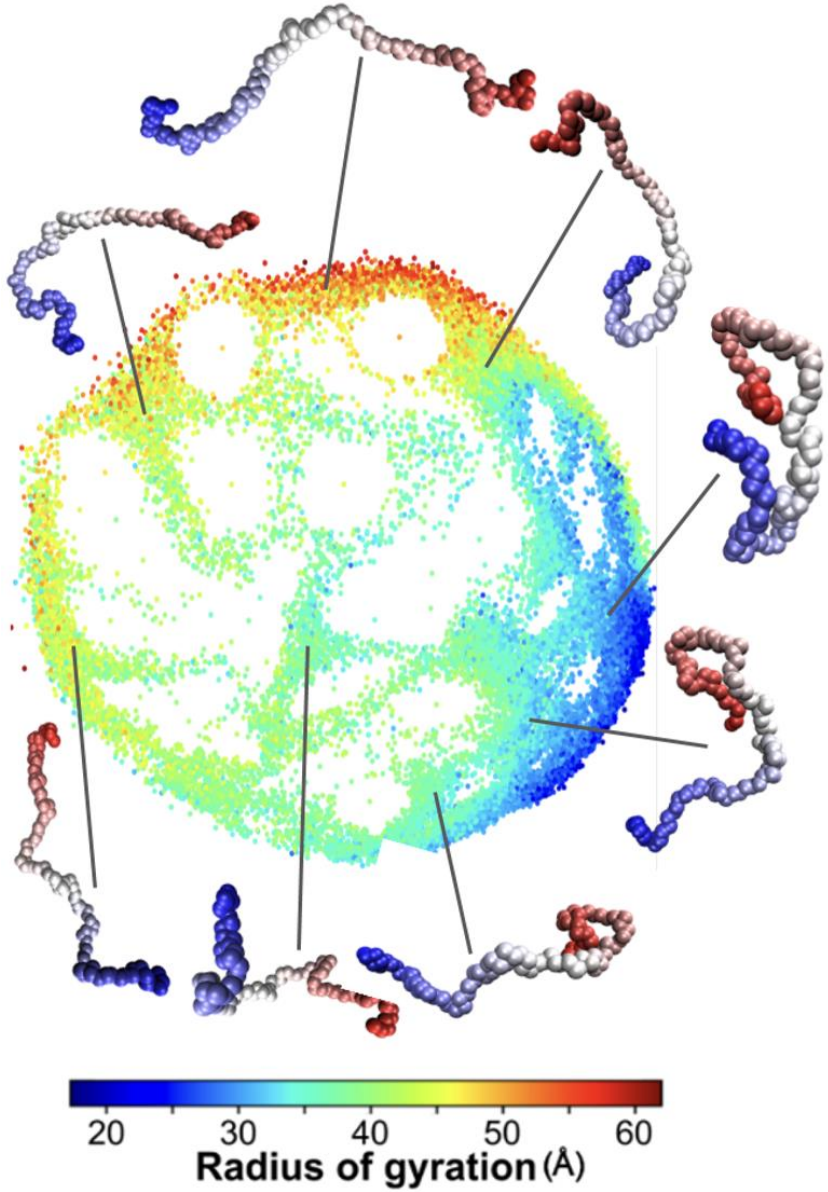
X. Lin et al,
J Mol Biol (2018)

No trivial reaction coordinate

- 10k conformations of each phosphoform (all together)

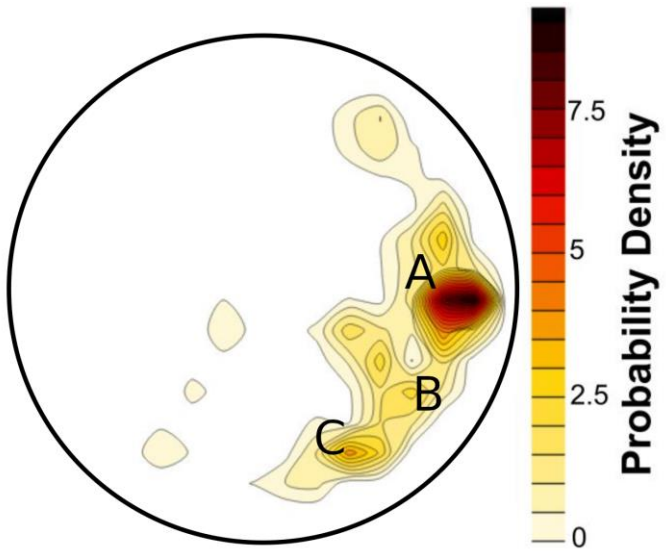


Oliveira, V. Leite, *JCTC* (2021)

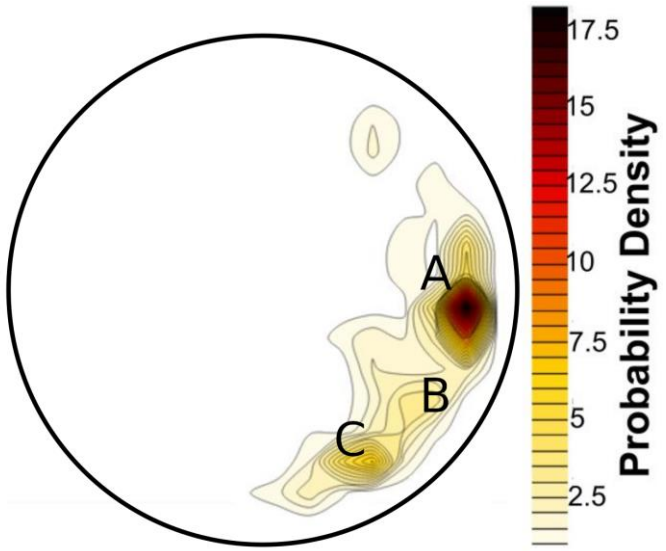


Density of States

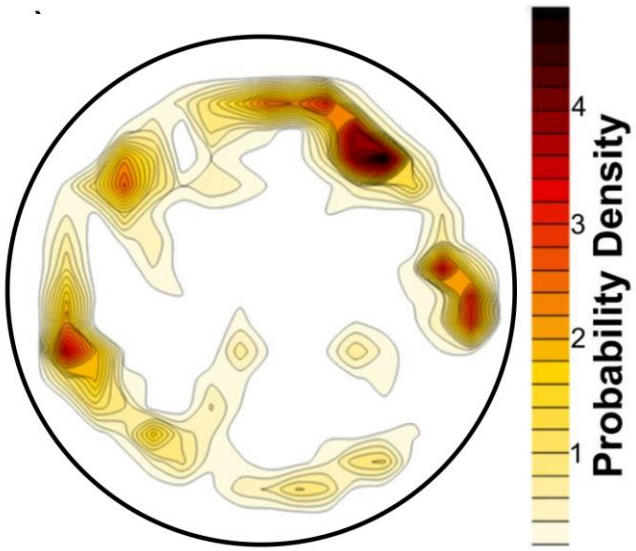
WT



HIPK1



CLK2

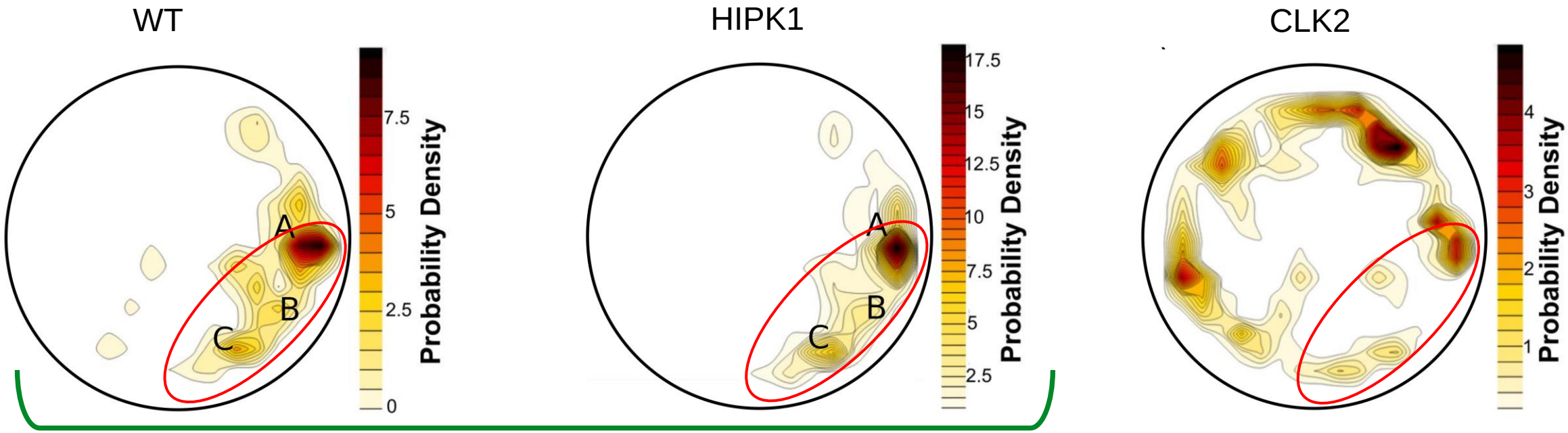


Experimental data

- Thr51 ⇔ C-jun binding site

Conformation ⇔ Function

Type	C-Jun (in vitro)
WT	Binds ++
HIPK1	Binds +
CLK2	No bind

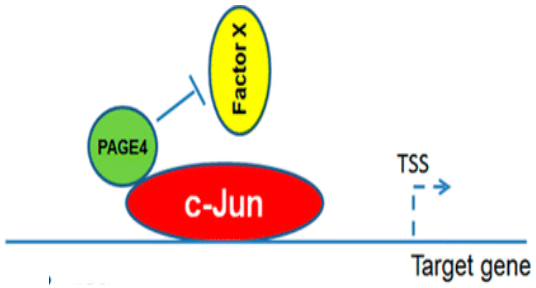


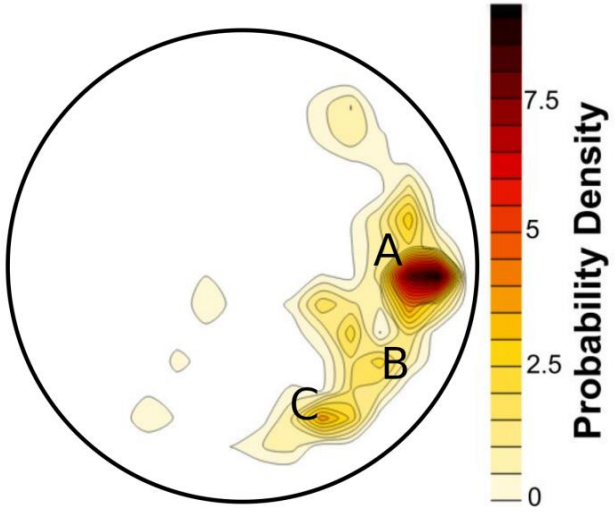
Functional Mechanisms

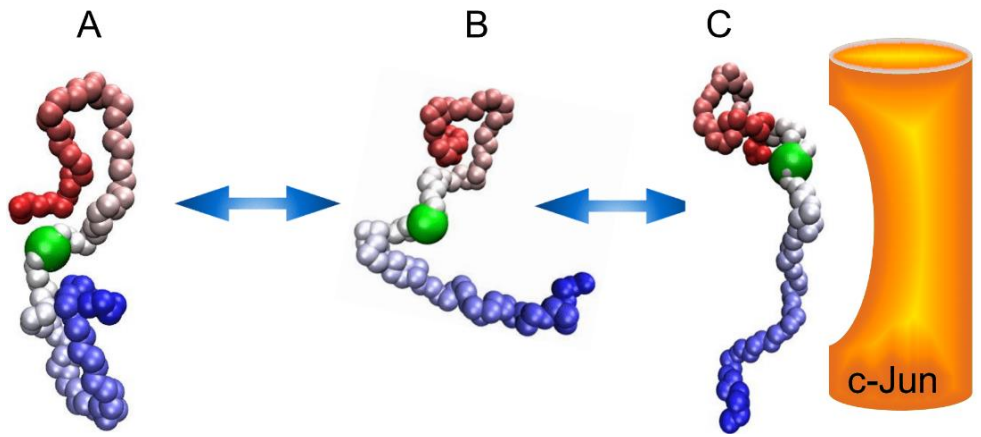
- Fly-casting mechanism

(Wolynes *et al, PNAS* 2000, PNAS 2010)

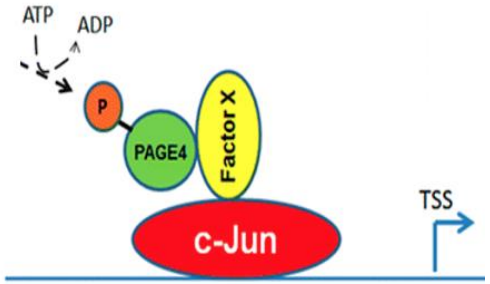
WT

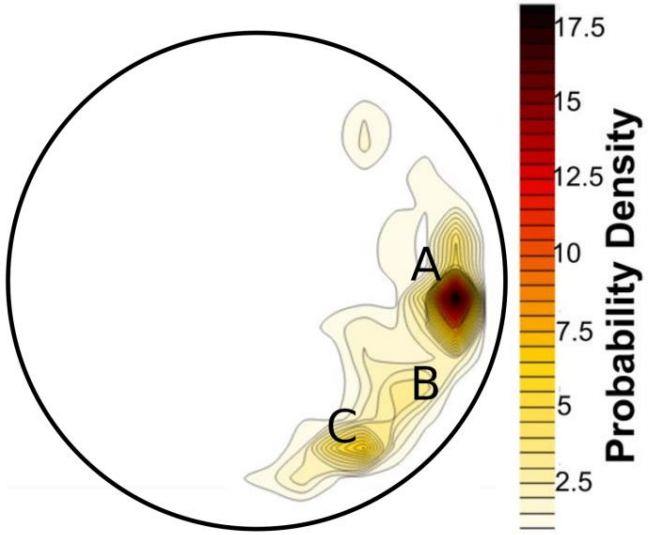


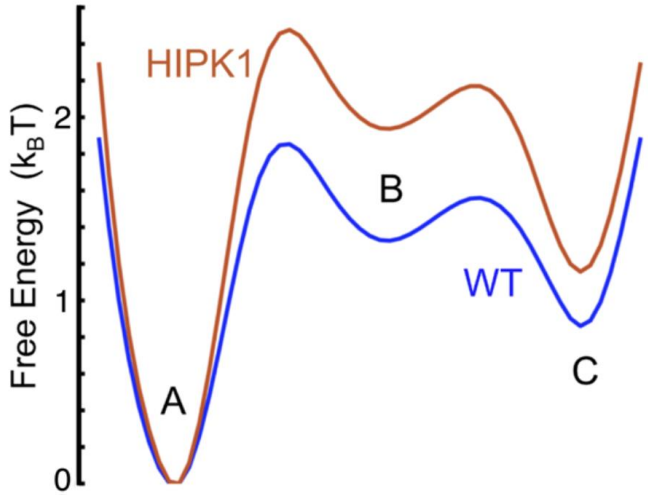




HIPK1







HIPK1 no Fly-casting

Protein Ensemble Database

proteinensemble.org

Lazar, *et al. NAR,* 2020

- PED is an open-access database for the deposition of structural ensembles, including *intrinsically disordered proteins (IDPs)*.
- "Manually curated data of structural ensembles measured with nuclear magnetic resonance spectroscopy, small-angle X-ray scattering, fluorescence resonance energy transfer... "

- Proof of concept: Can we make sense out of these ensembles?

- Fragment of the nuclear pore complex protein (Nup)153 - NUS (1313-1390). Fuertes, *et al. PNAS*, 2017
- - NUL (884-993).
- Sic1 N-terminal targeting domain (1-90). Gomes, *JACS*, 2020
- N-terminal SH3 domain of Drk protein (1-59). Lincoff,. *Comm. Chem*. 2020

Viegas, Martins & Leite, *JCIM* **2024**

Nuclear pore complex protein

Nup153 fragment - NUS (1313-1390)

Fuertes, *et al. PNAS*, 2017

78 residues.

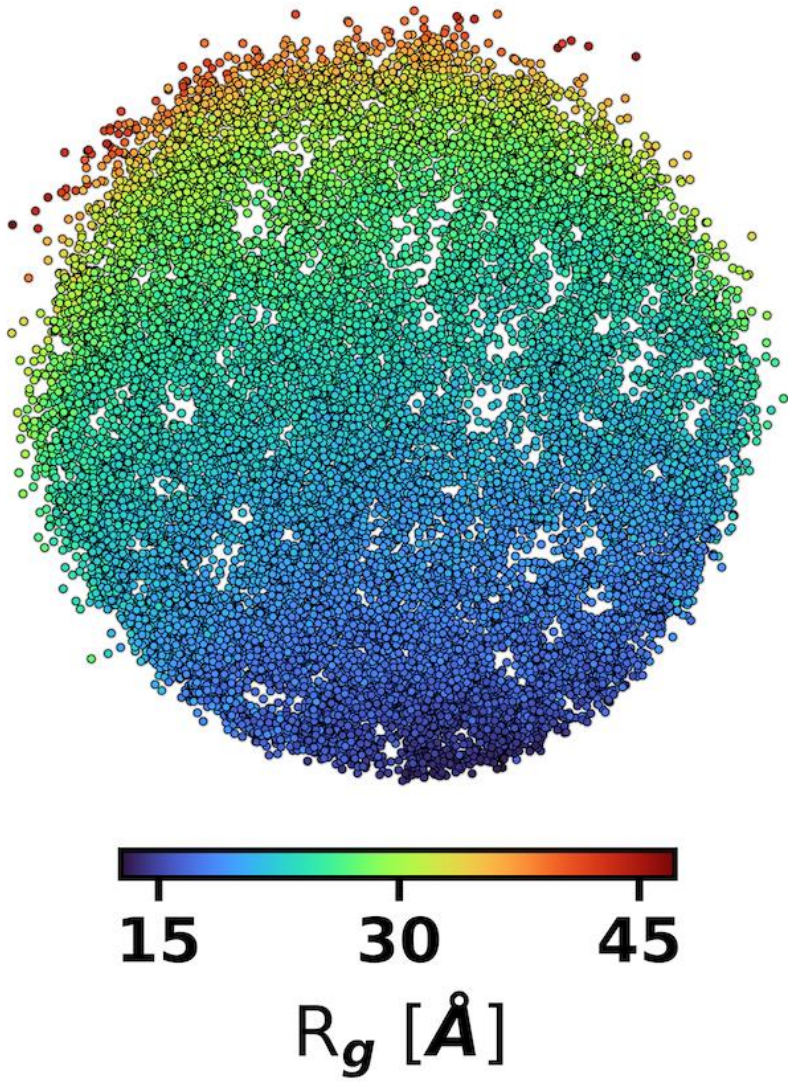
Total of 28,078 structures

PED00149: denatured conditions

- e001: 9482 models, R_g = 23.38
- e002: 9405 models, R_g = 23.78
- e003: 9473 models, R_g = 23.51

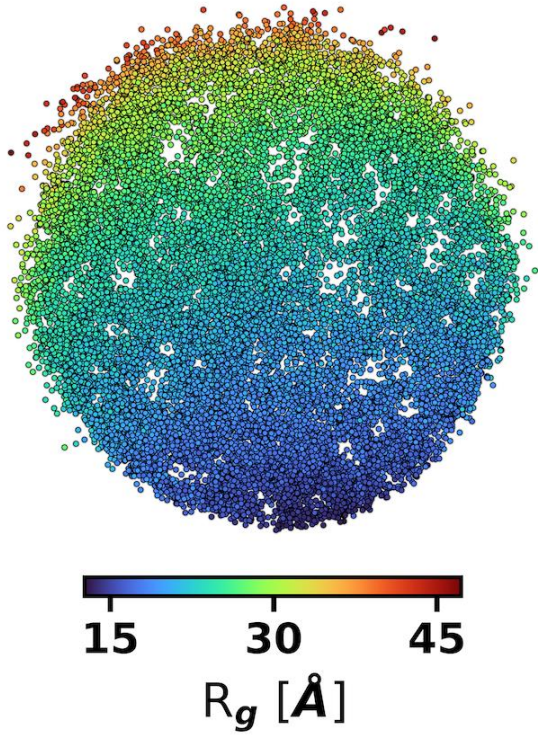
PED00150: native conditions

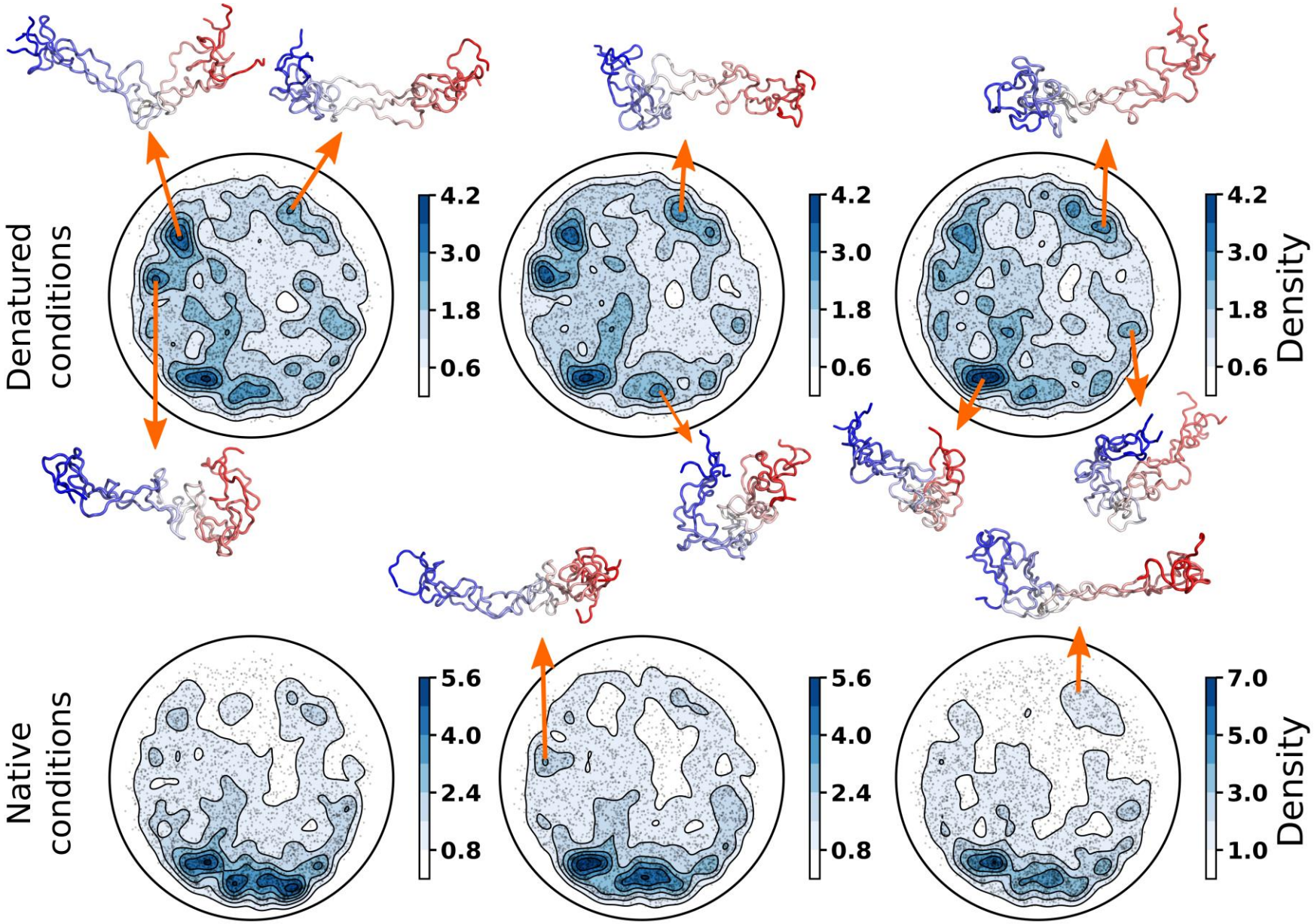
- e001: 9255 models, R_g = 21.45
- e002: 9248 models, R_g = 21.72
- e003: 9277 models, R_g = 21.69



Viegas, *Martins & Leite*, *JCIM* **2024**

Nup153 fragment
NUS (1313-1390)

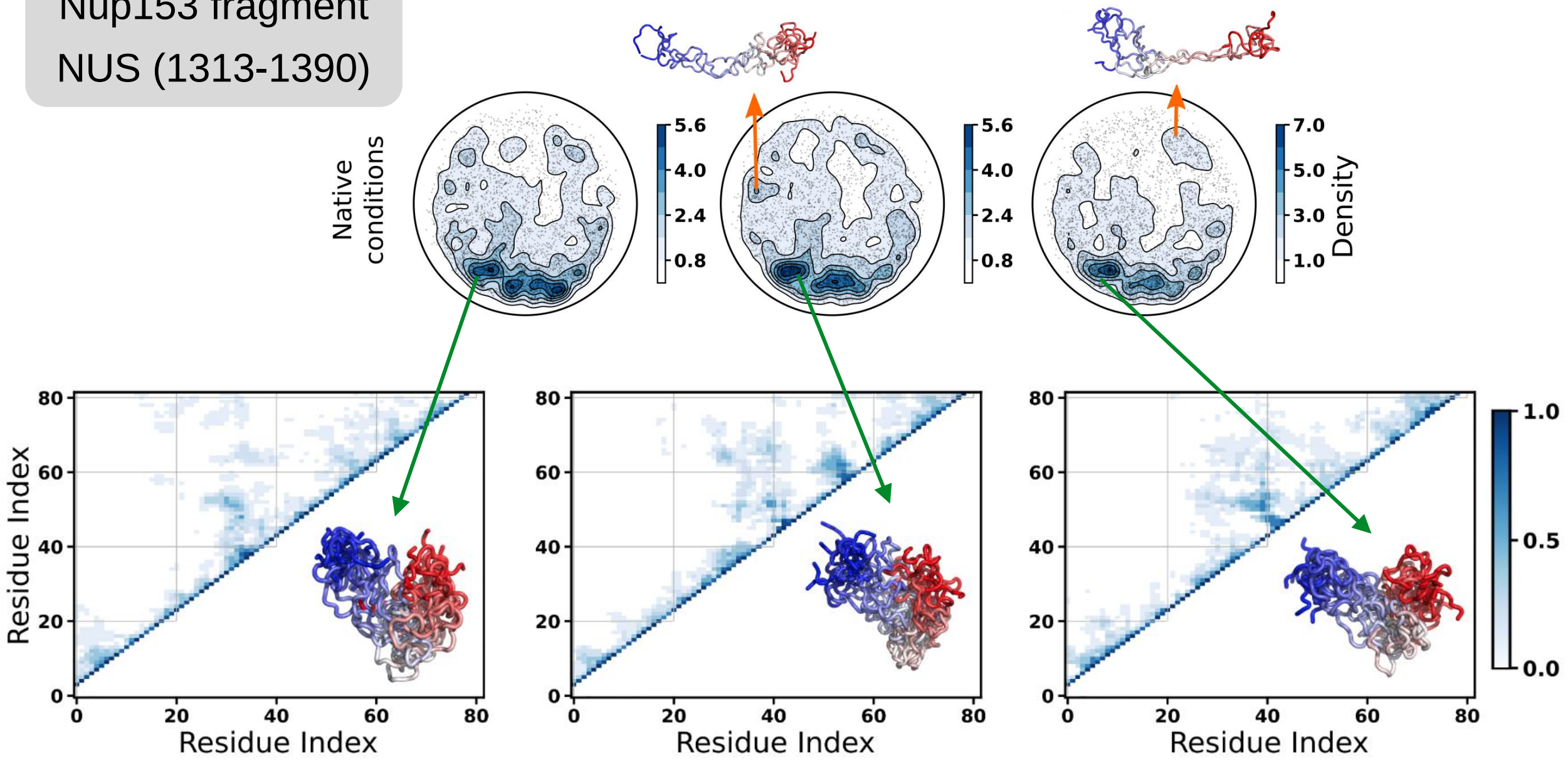




Viegas, *Martins & Leite*, *JCIM* **2024**

Nup153 fragment

NUS (1313-1390)

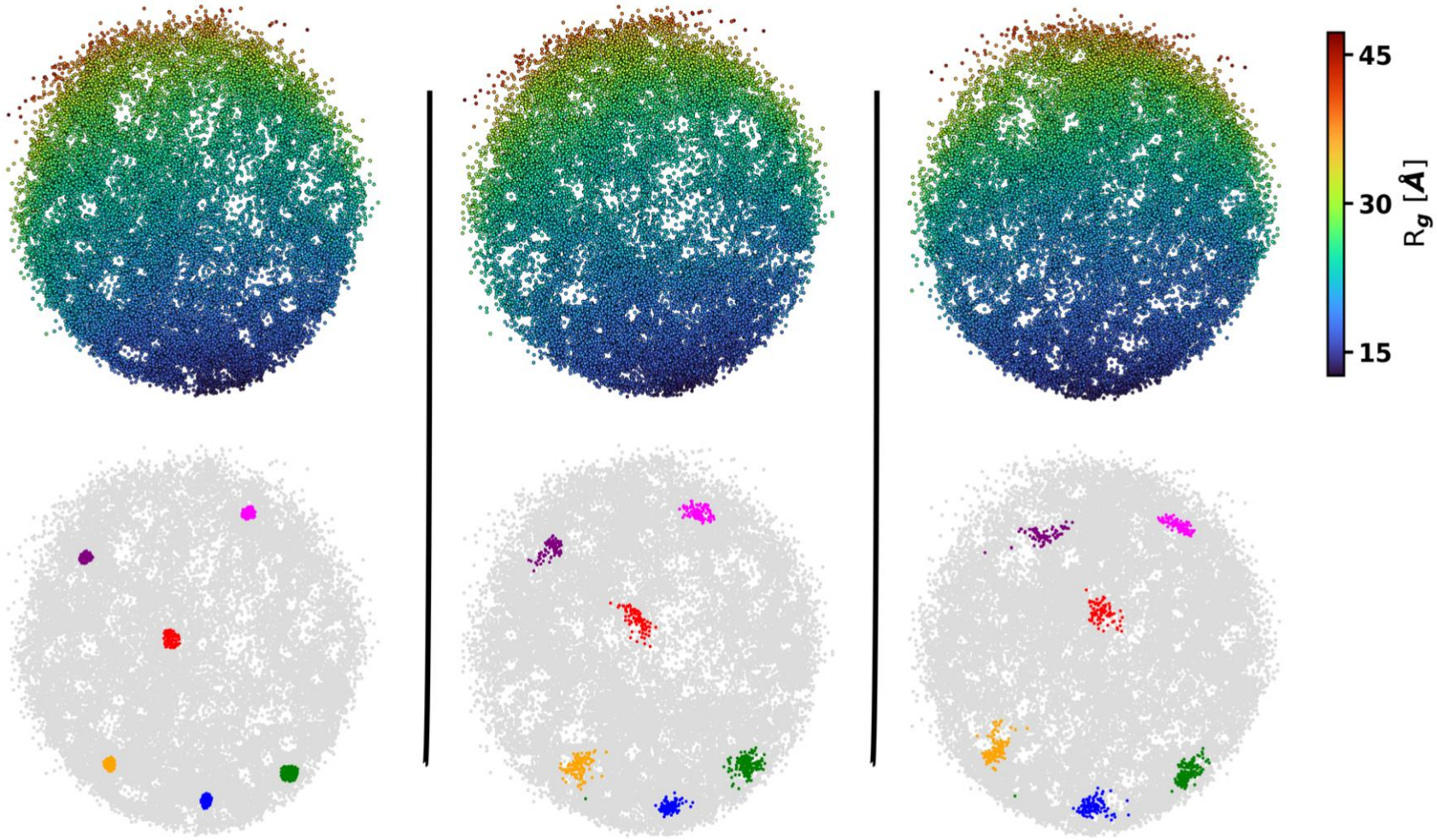


Viegas, *Martins & Leite*, *JCIM* **2024**

Nup153 fragment

NUS (1313-1390)

ELViM Reproducibility



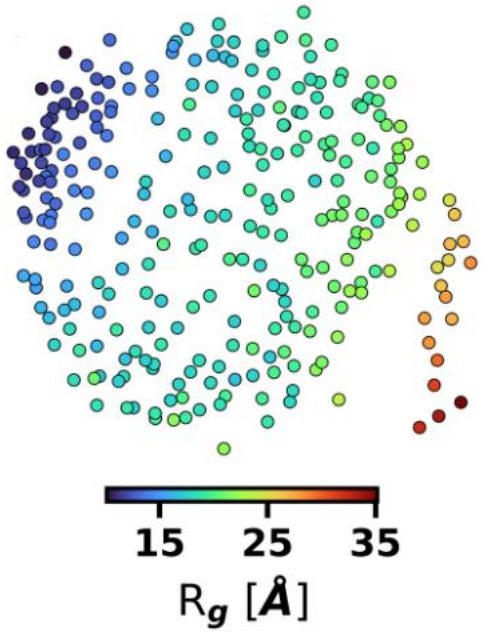
Viegas, *Martins & Leite*, *JCIM* **2024**

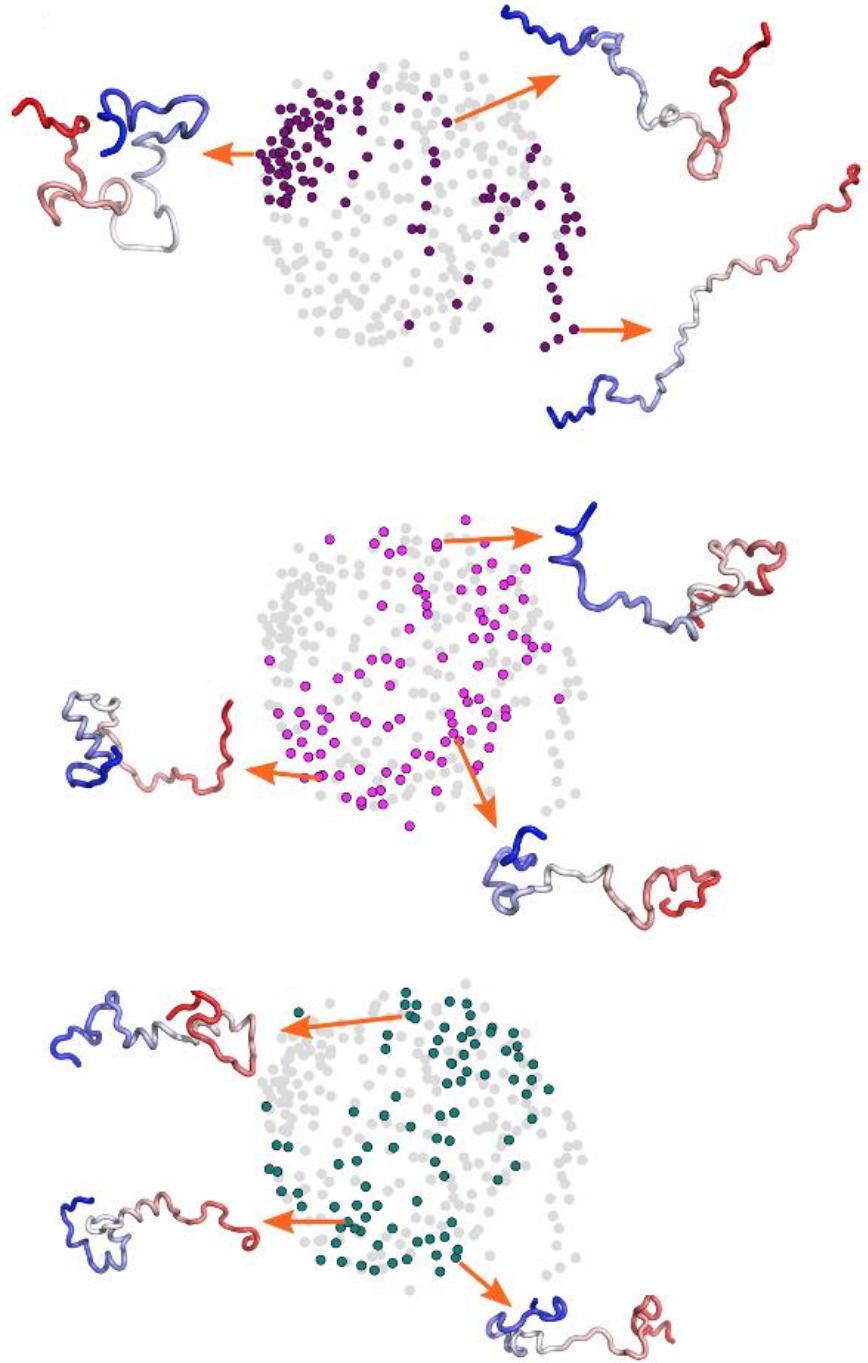
N-terminal SH3 Domain of Drk protein

Lincoff, *et. al., Comm. Chem*. 2020

Consistent with NMR, SAXS and smFRET data.

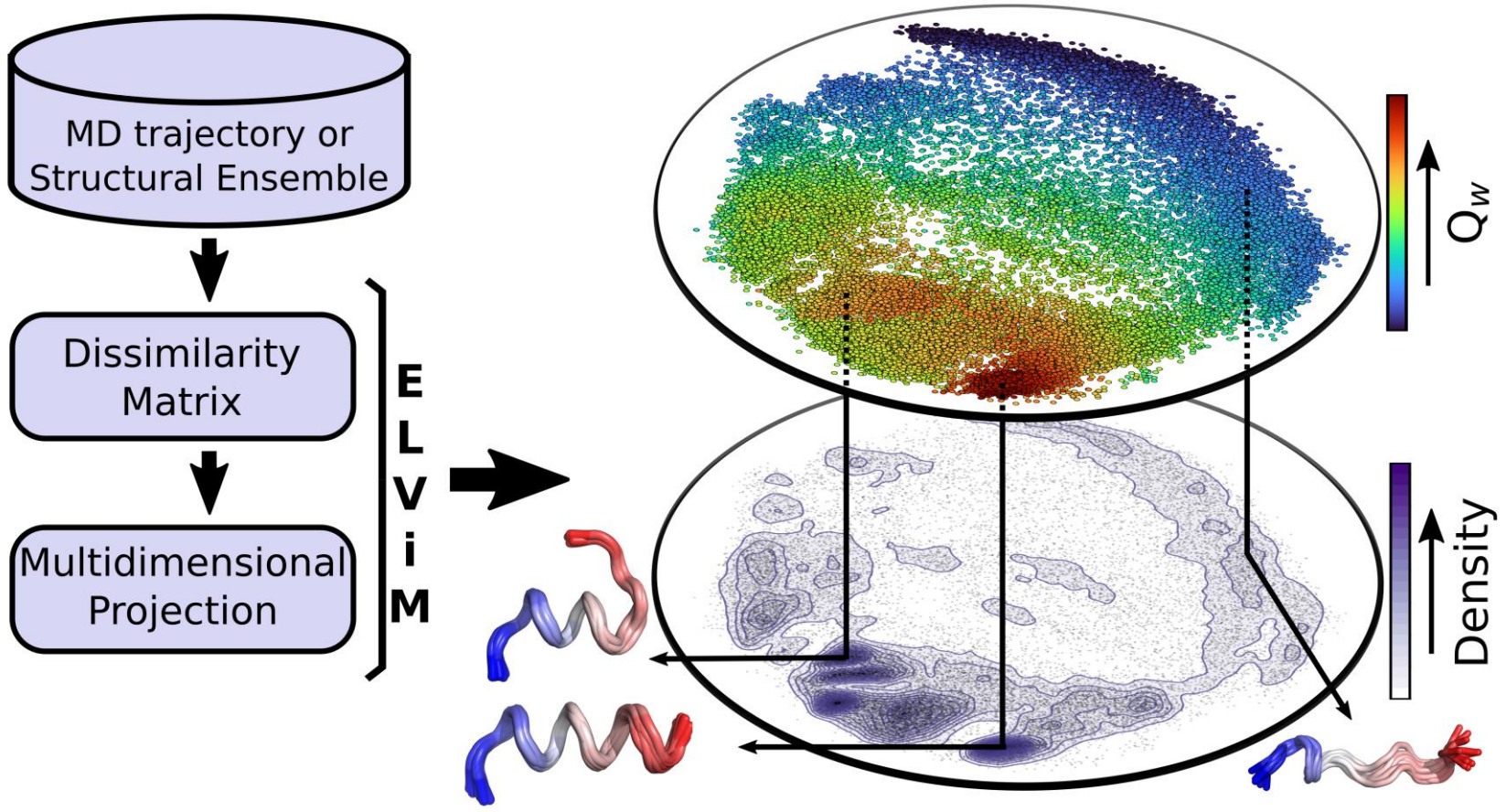
59 residues
Total of 288 conformations





Viegas, Martins & Leite, *JCIM* **2024**

ELViM



Viegas *et. al.*, *JCIM* **2024**

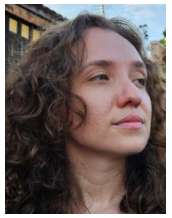
GitHub: https://github.com/VLeiteGroup/ELViM

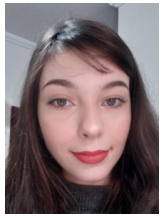
Acknowledgments

The Group
@ IBILCE-UNESP

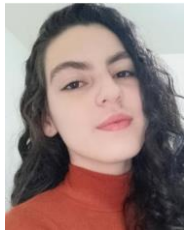

Rafael Viegas


Murilo Sanches

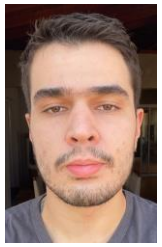

Larissa Adolfo


Juliana Camargo

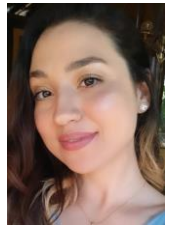

Gustavo Catanoce


Giovana Trevejo


Lucas Rosseti


Vicente Christiano


Antonio Oliveira (@Rice Univ)


Ingrid Martins (@EMS)

@DF-IBILCE:

- Alexandre Suman
- Leandro C. Oliveira
- Raphael Dias

- Angel Garcia, Los Alamos Natl. Lab.
- Diego Ferreiro, Univ. Buenos Aires
- Fernando Paulovich, TU/e, Netherlands
- José N. Onuchic, Rice Univ.
- Peter Wolynes, Rice Univ.
- Prakash Kulkarni, Natl. Cancer Inst., CA
- Susmita Roy, IISER Kolkata, India
- Victor Tsai, CCU, Taiwan

vitor.leite@unesp.br

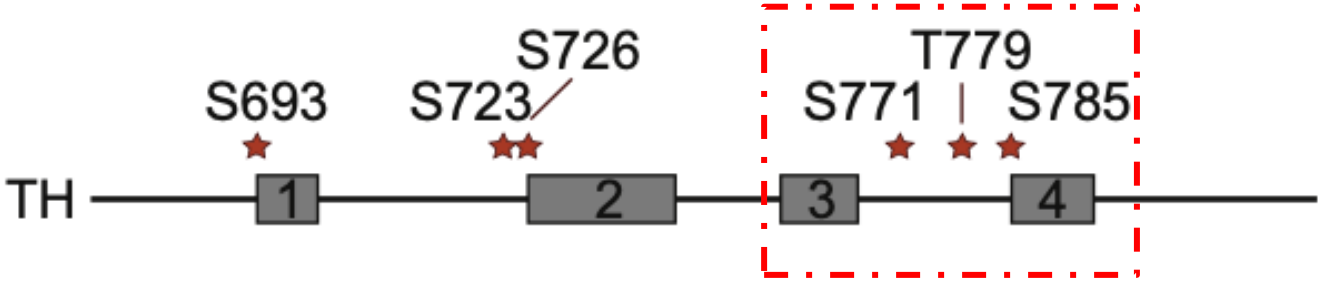
Thank you!

Obrigado!

vitor.leite@unesp.br

Na+/H+ exchanger 1

(NHE1)



R. Hendus-Altenburger
K. Lindorff-Larsen, B.B. Kragelund

Cellular Signalling 37 (2017) 40–51

The disordered distal tail of NHE1 is six-times phosphorylated by the mitogen activated protein kinase 2 (MAPK1, ERK2). Using NMR, they found that two out of those six phosphorylation sites had a stabilizing effect on transient helices.

Molecular Dynamics

- Residues 755 - 796 (TH3 and TH4)
- Phosphorilations (S771, T779, S785)
- Amber03ws force field - GROMACS 2020
- 1 microsecond of simulation
- Phosphorilation parameters added
- 5 replicas for WT &Phosphorilated

ELViM

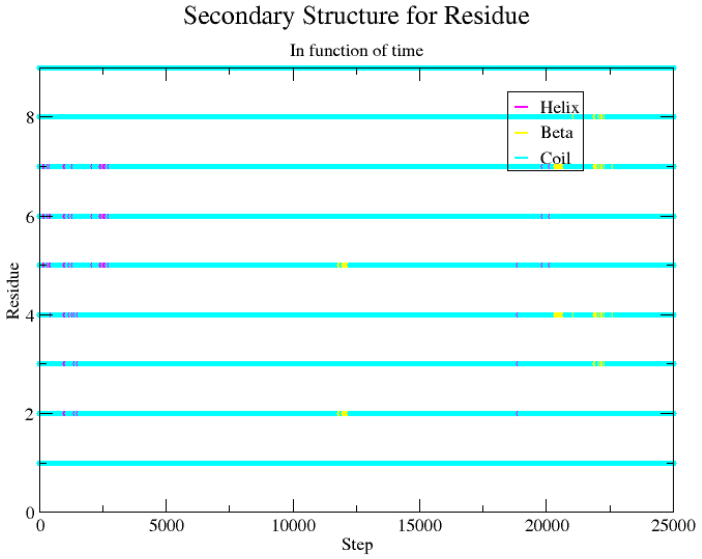
- 12500 structures - 5 replicas of WT
- 12500 structures - 5 replicas of Phosphorilated

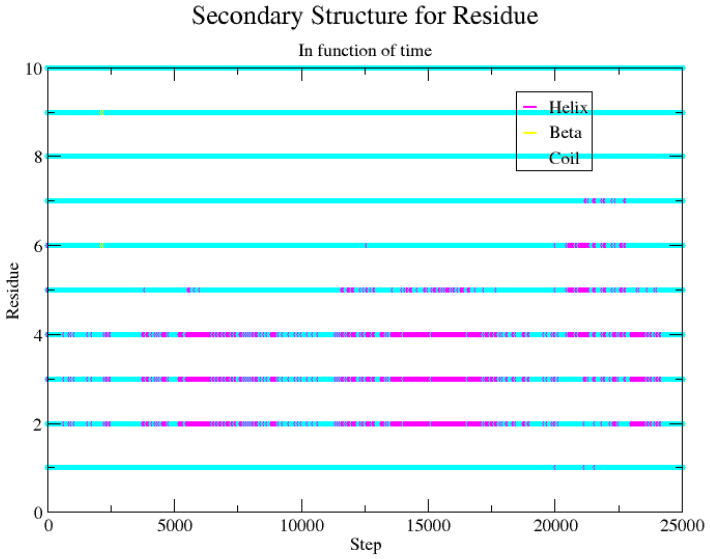
Transient Helix populations

TH3

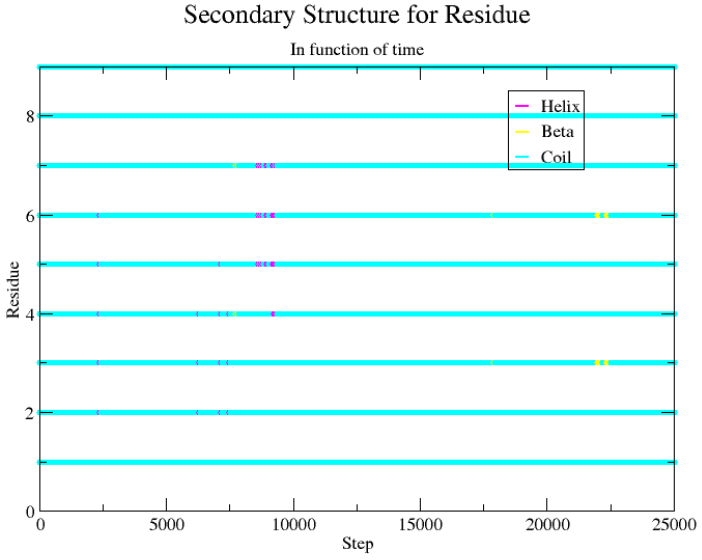
TH4

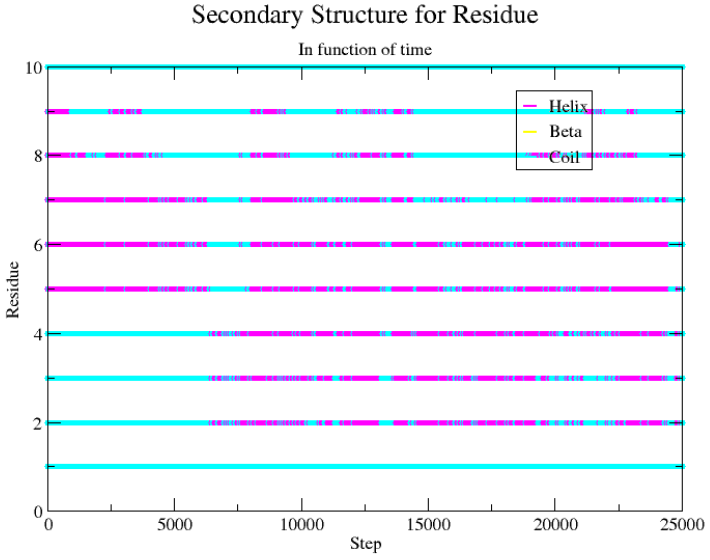
WT

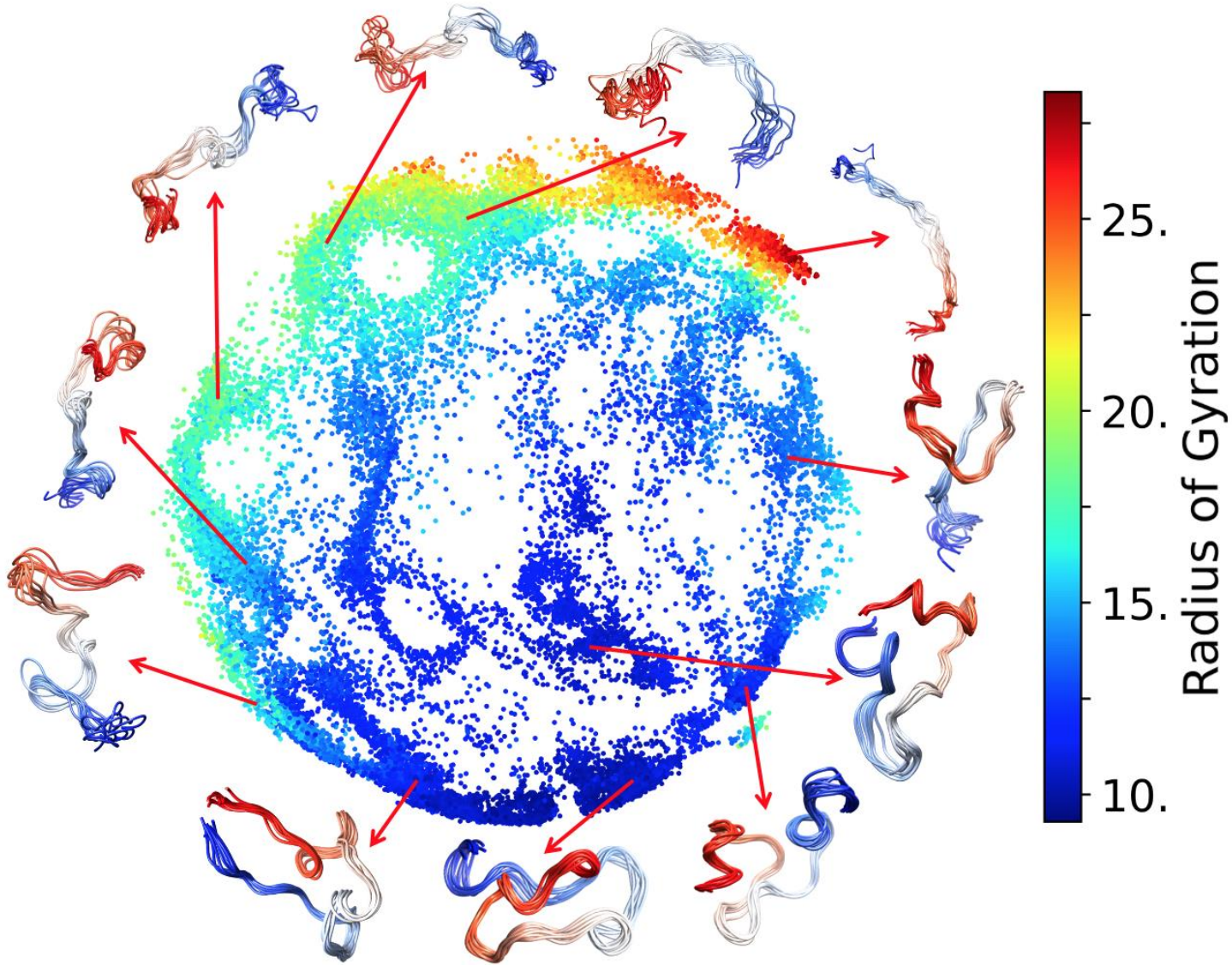


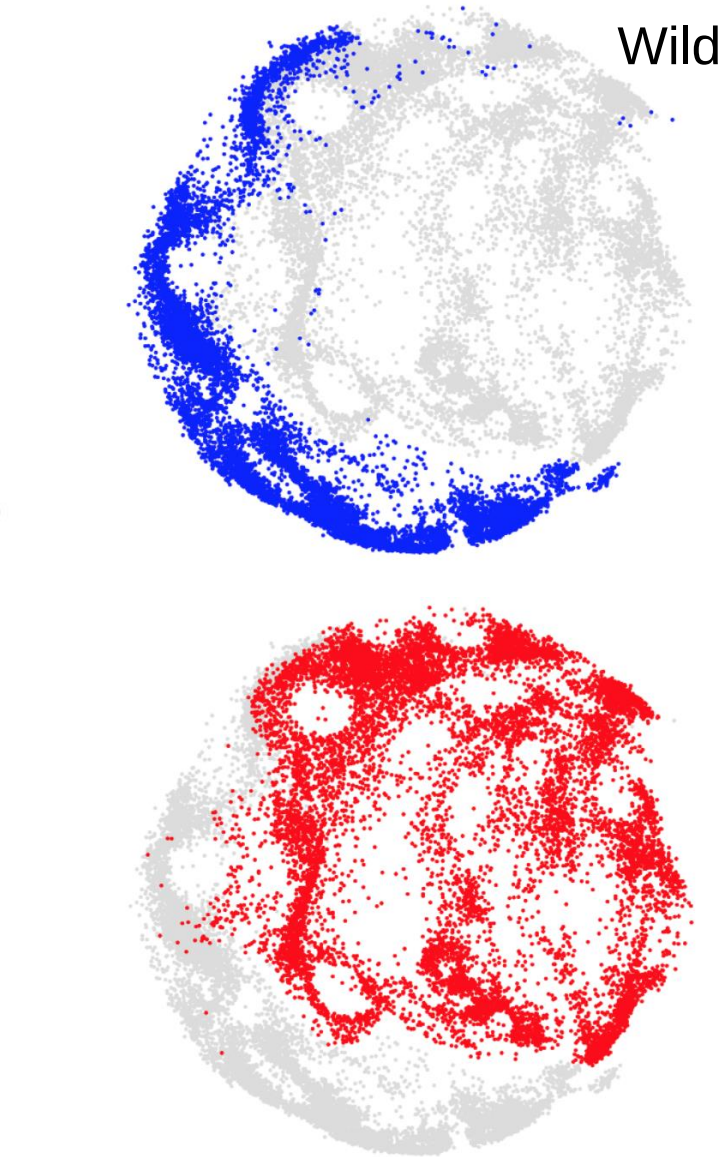


Phospho





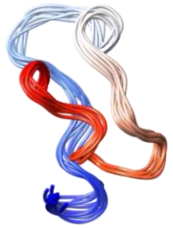


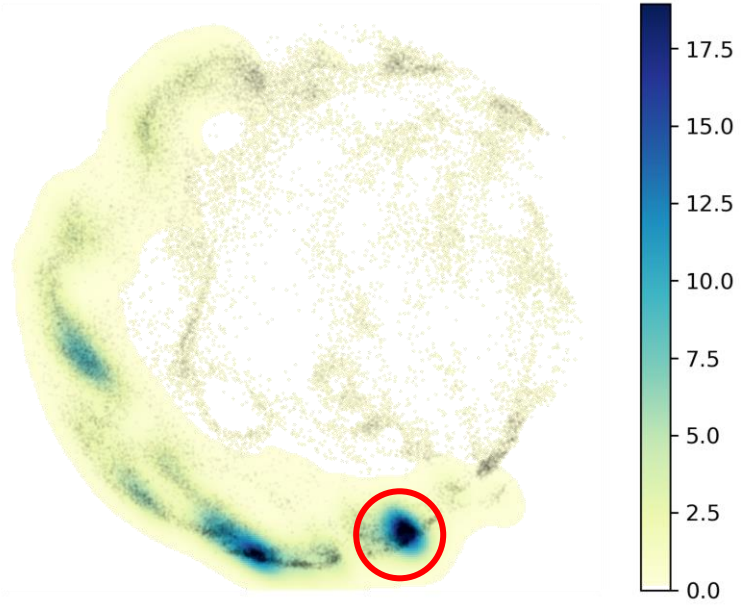


Wild Type

Phosphorylated

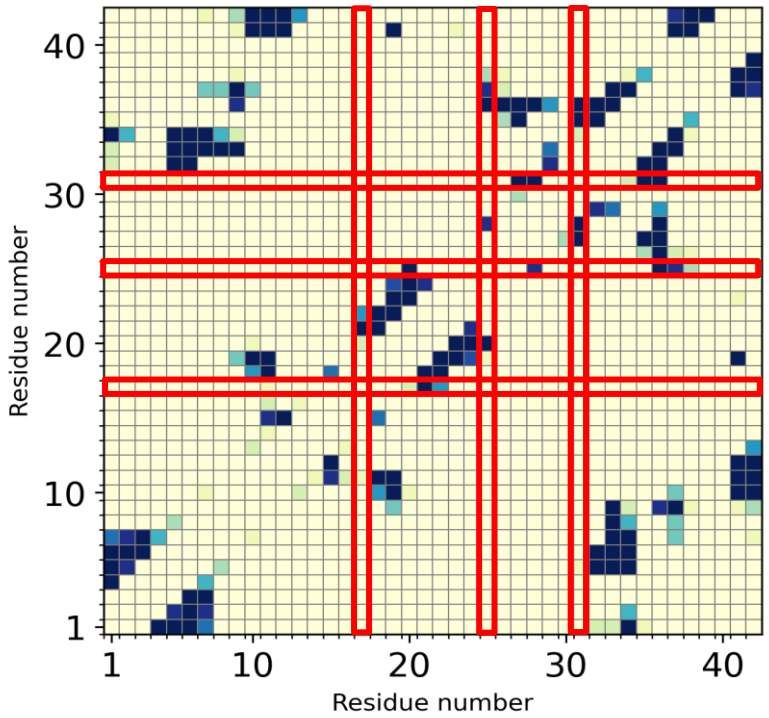
WT



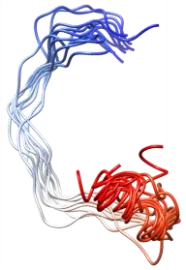


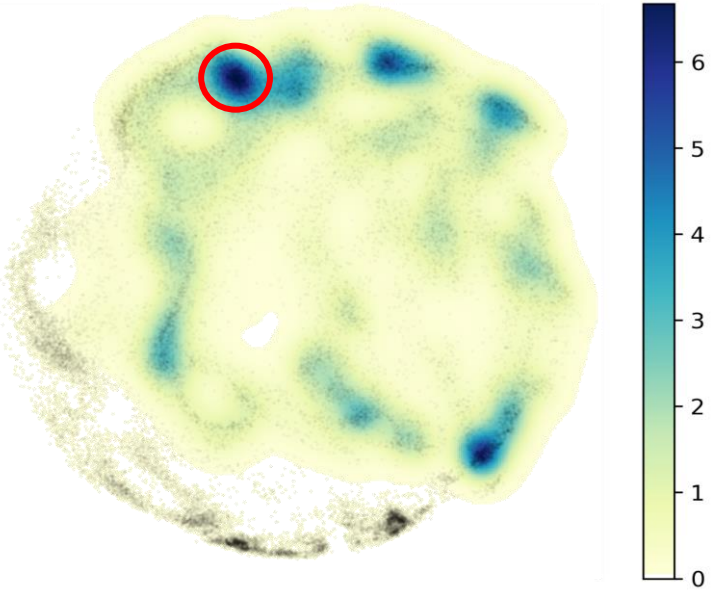
C-Term

N-Term

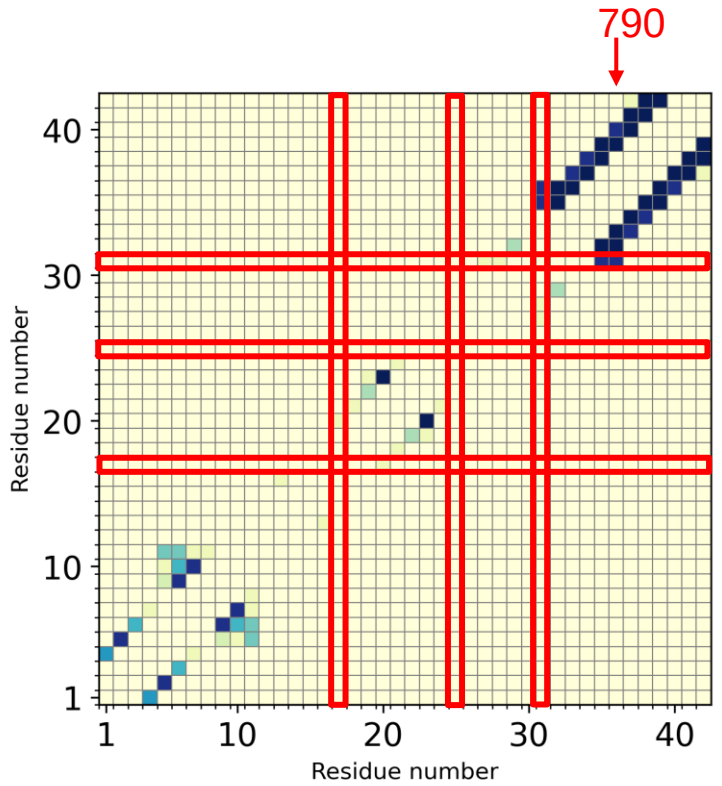


Phospho



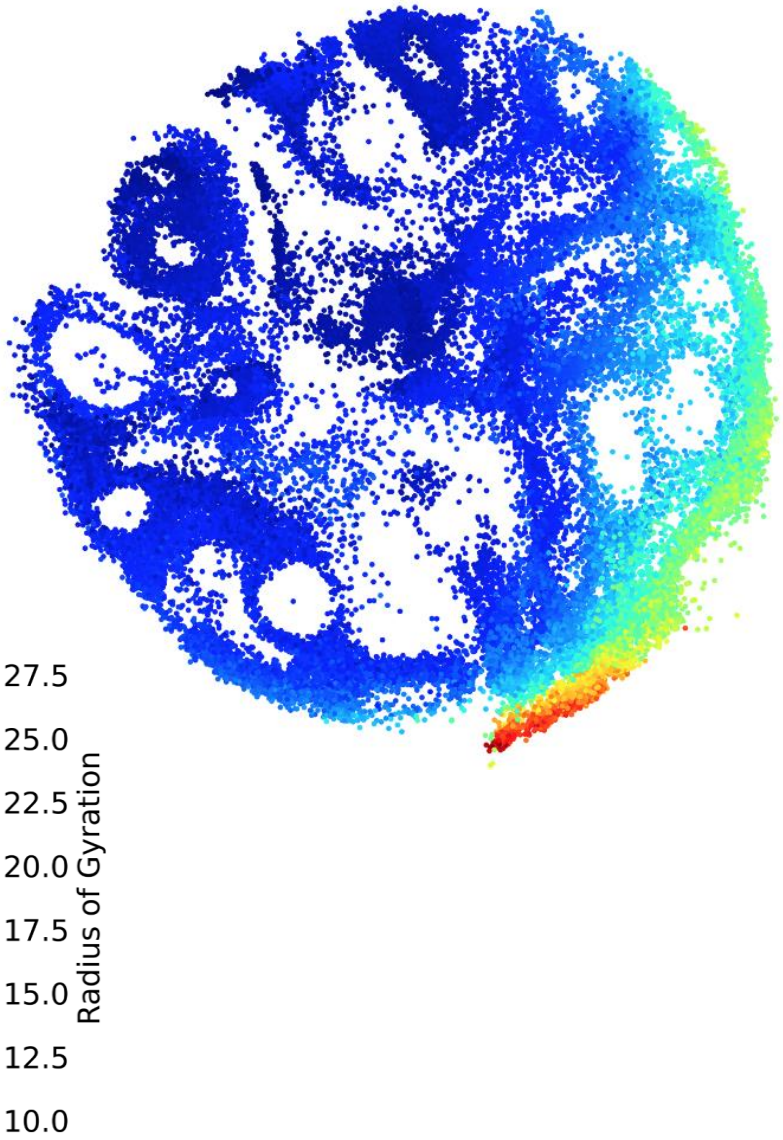


790



MD/ELViM

- WT, WT phosphorilated
- R790V, R790V phosphorilated



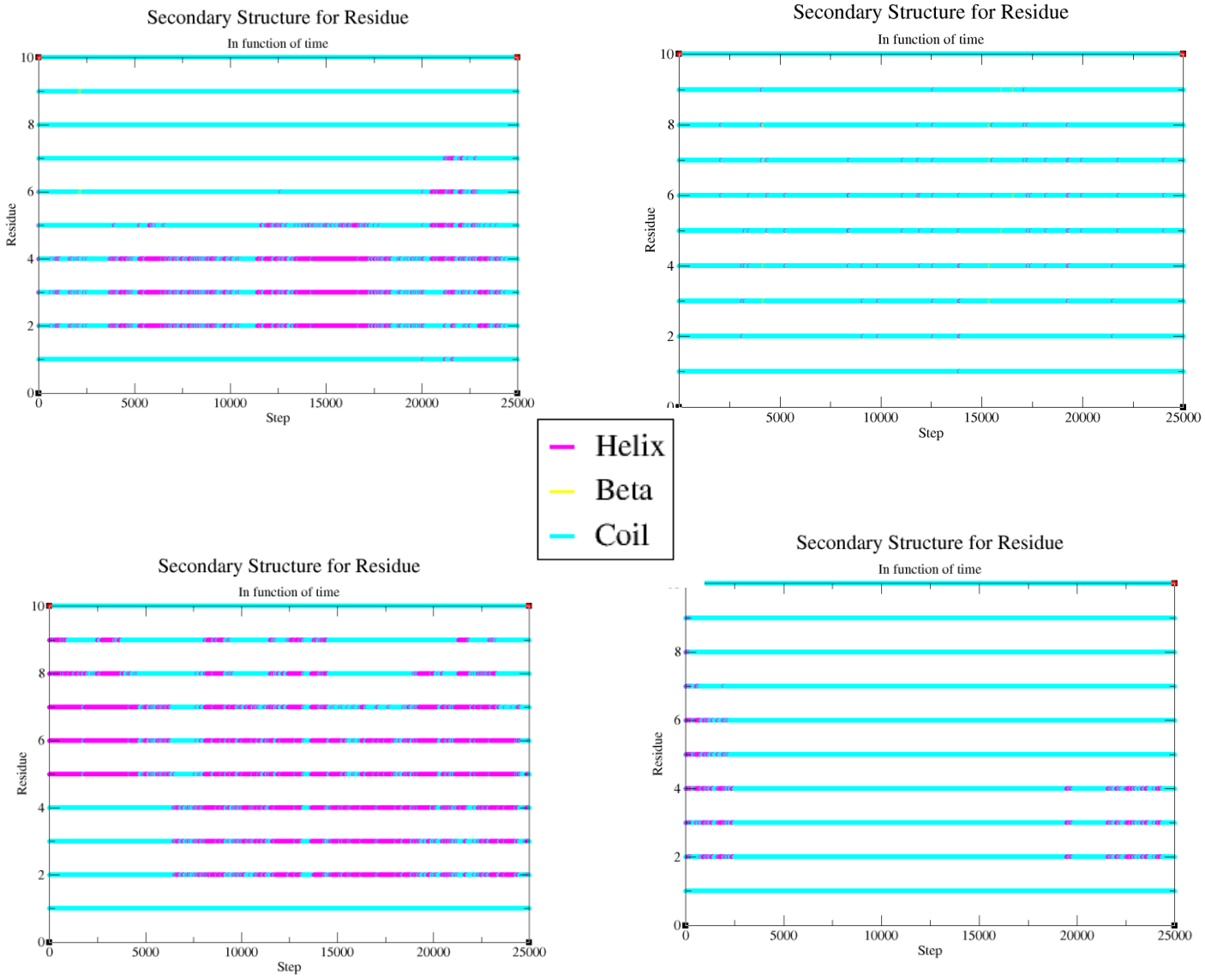
Transient Helix 4

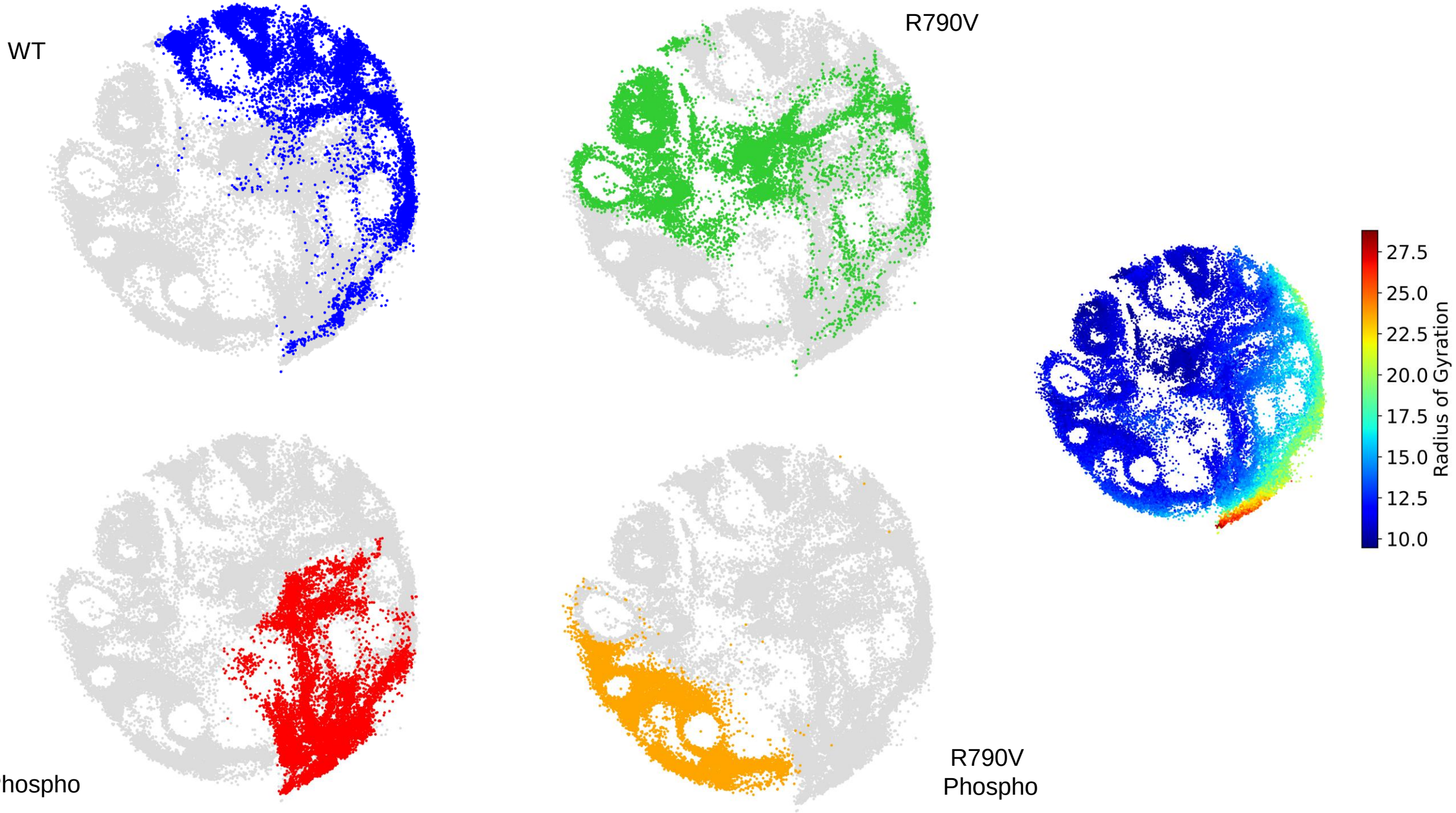
WT

R790V

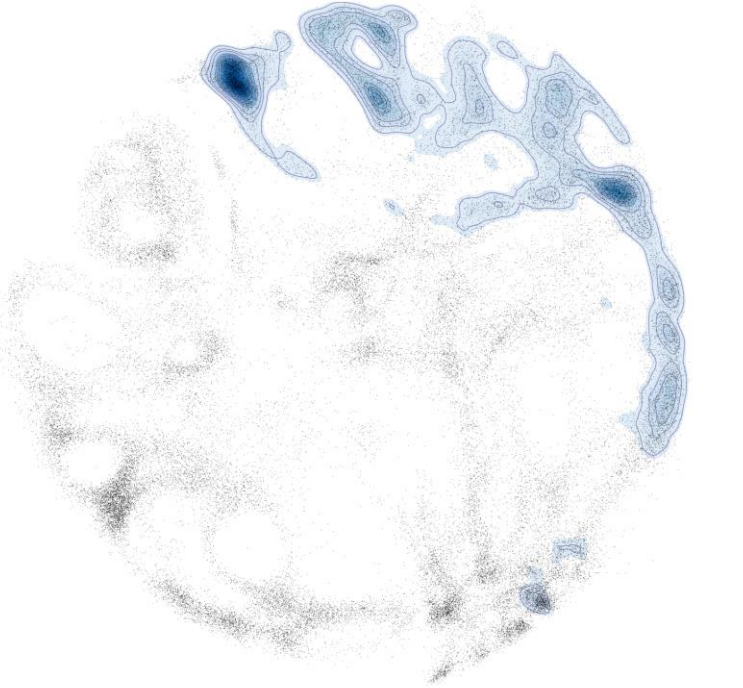
w/o Phospho

with Phospho

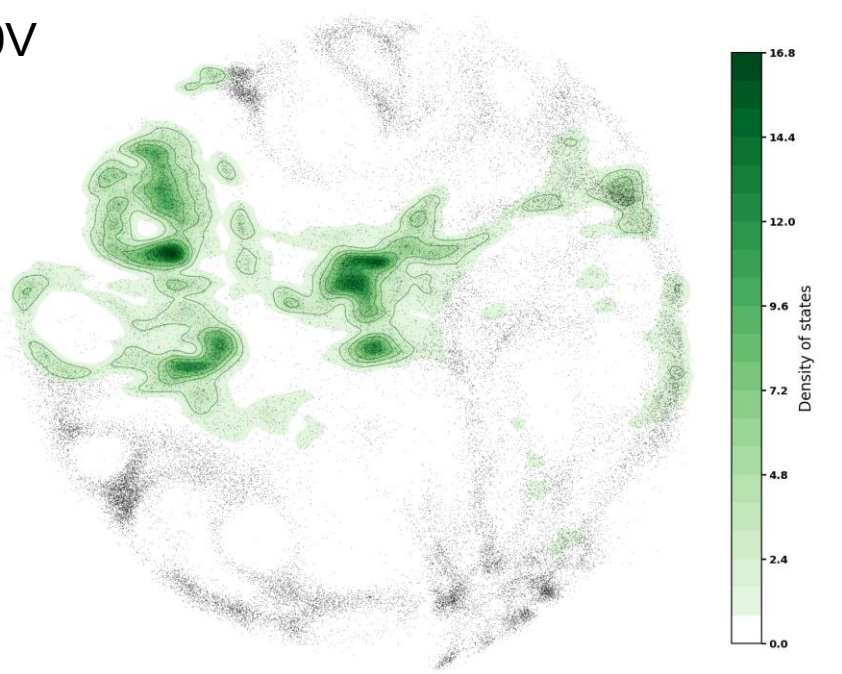



WT
R790V
Phospho
R790V
Phospho
27.5
25.0
22.5
20.0
17.5
15.0
12.5
10.0
Radius of Gyration

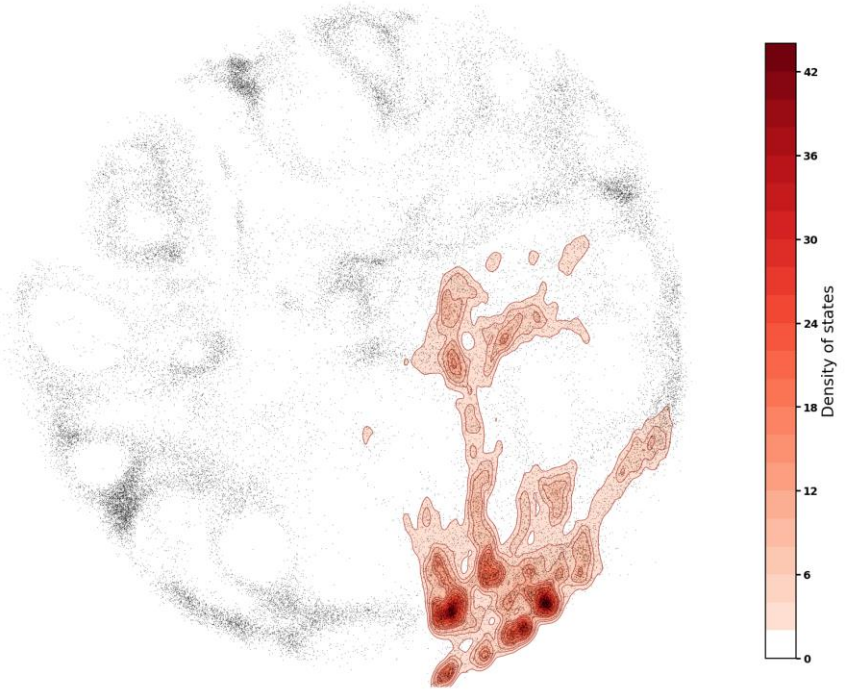
WT



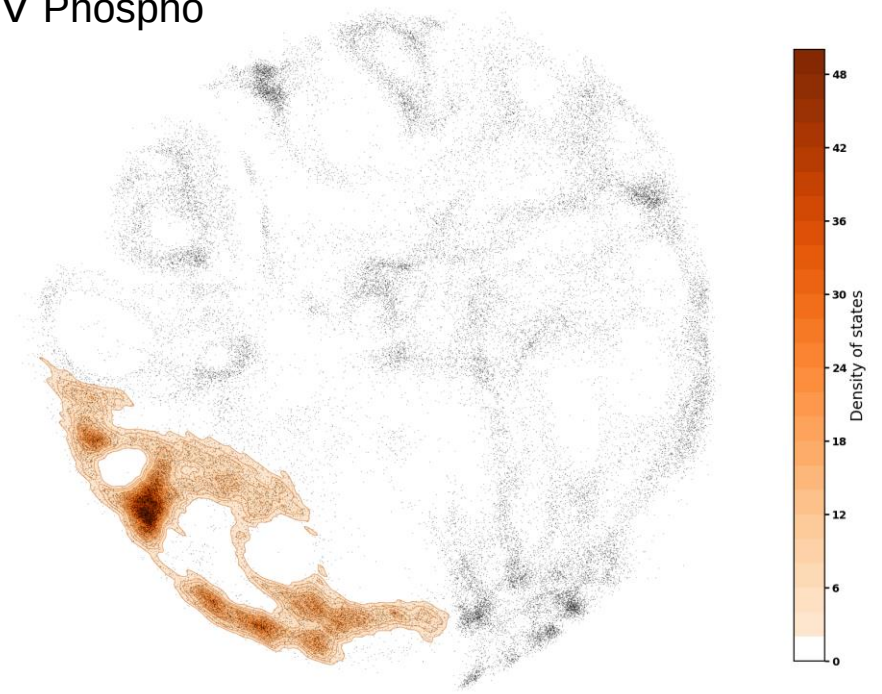
R790V



Phospho

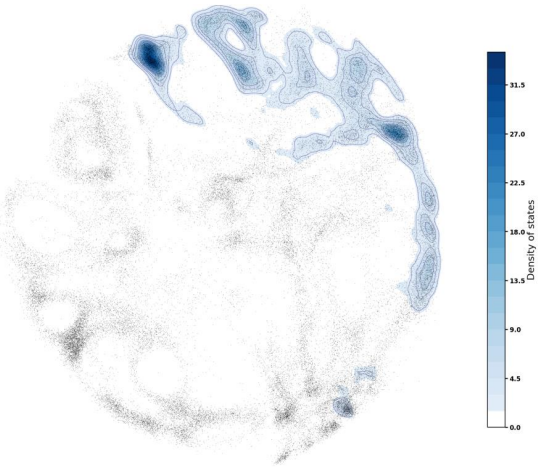


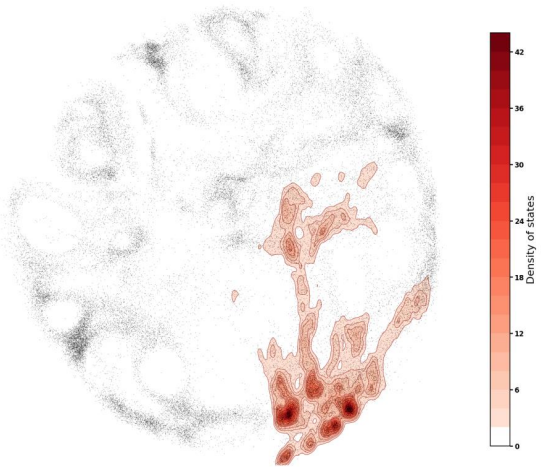
R790V Phospho



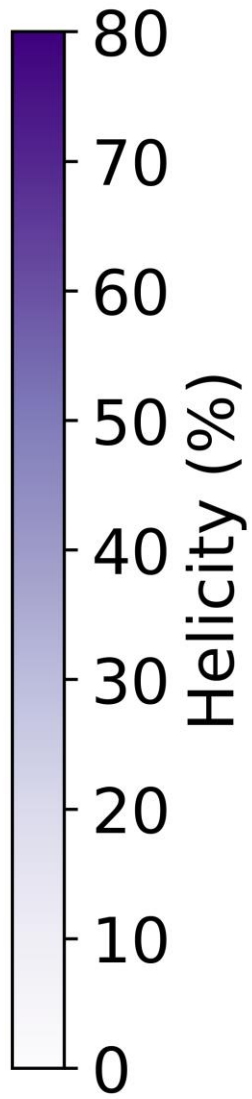
Helicity of Transient Helix 4

WT

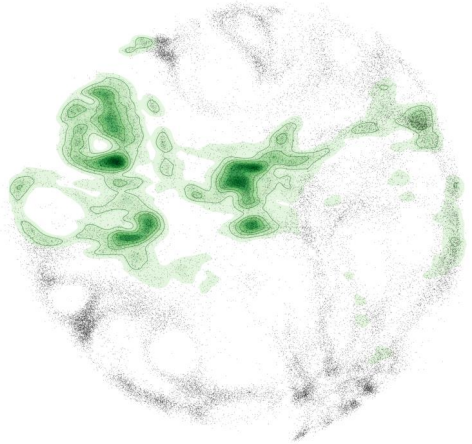


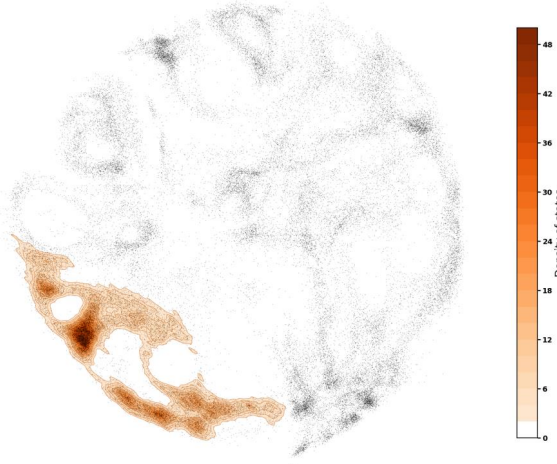


Phospho



R790V





R790V Phospho

Tau Fragments Energy Landscape

w/ **Joan-Emma Shea**
UC Santa Barbara

Tau protein ➔ assembly and stabilization of microtubules ➔

aggregates ➔ accumulation of Tau protein ➔ Alzheimer's disease

- Fragment: 295 – 313 (18 AAs) ➔ Seeds the fibrillization of the full protein
- Mutant P301L ➔ More prone to aggregate

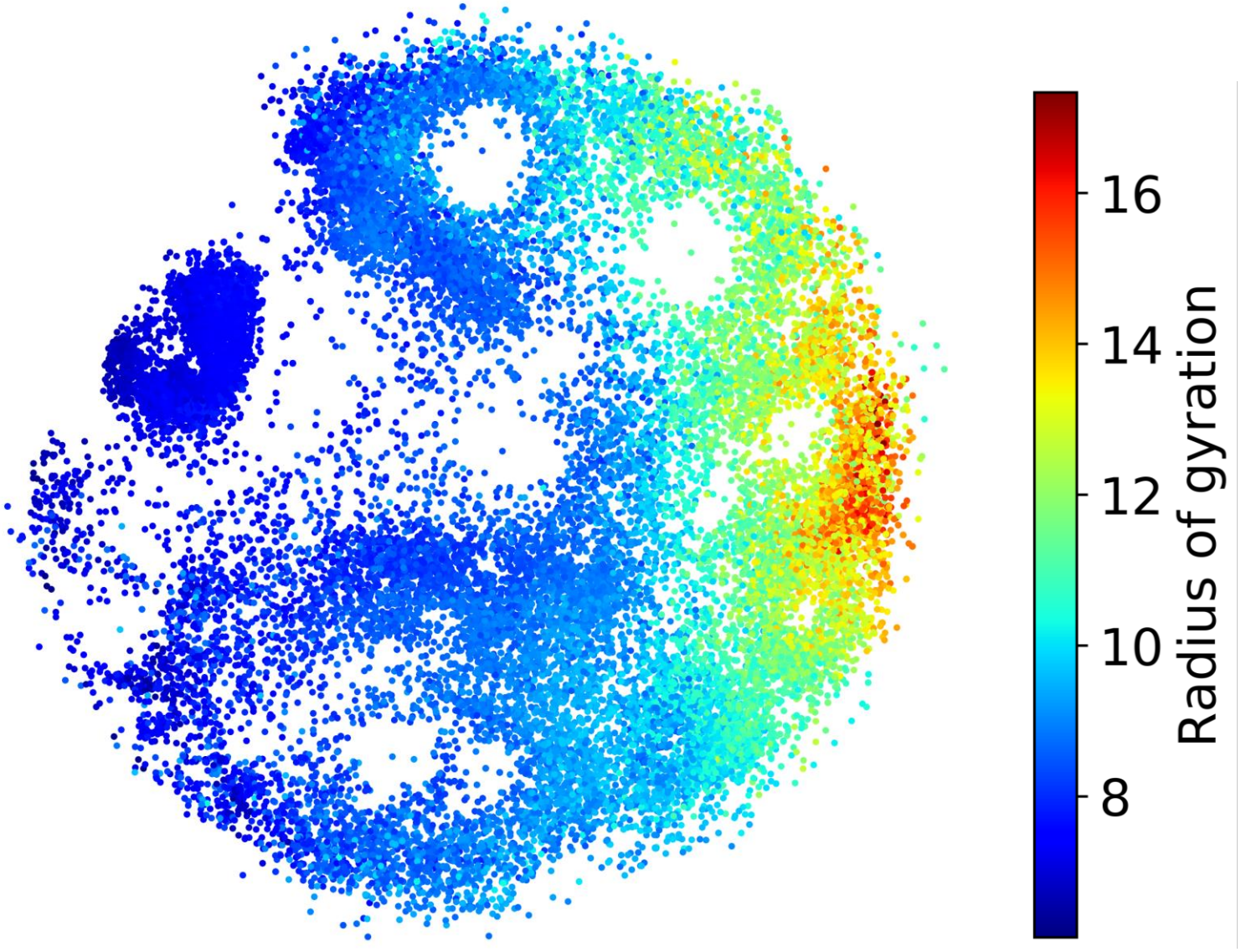
- Amber99sb Simulations
- WT: monomers, dimers & tetramers
- P301L: monomers, dimers & tetramers

➔ **ELViM analysis**

ELViM Projection

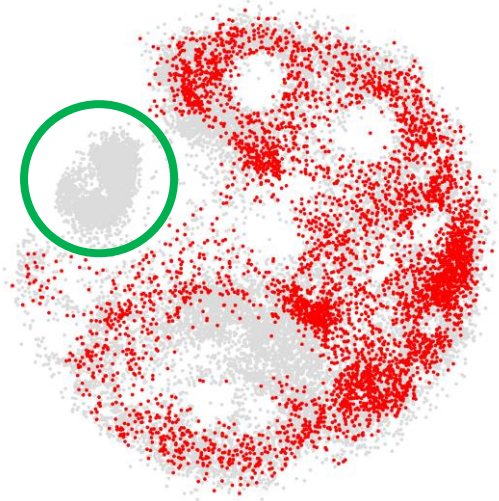
WT & P301L data:

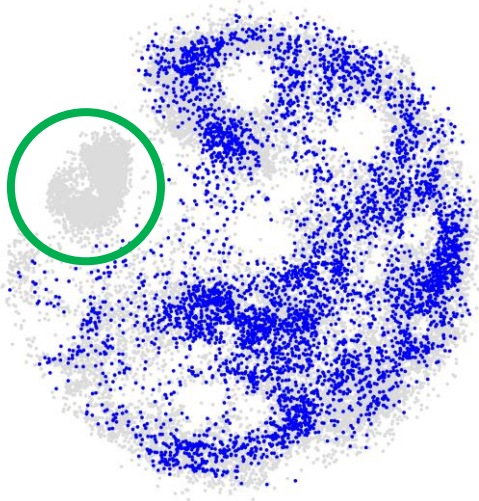
monomers, dimers & tetramers

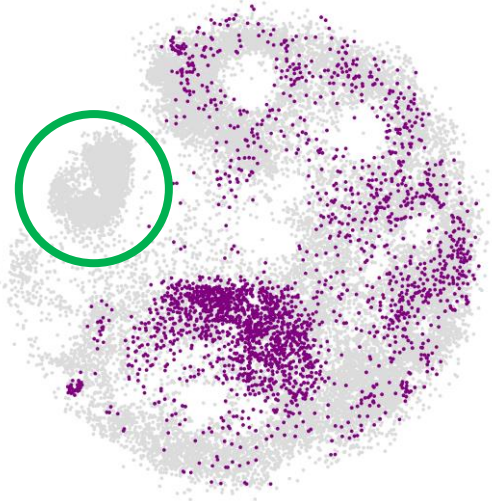


5000 points in each ensemble

Wild Type





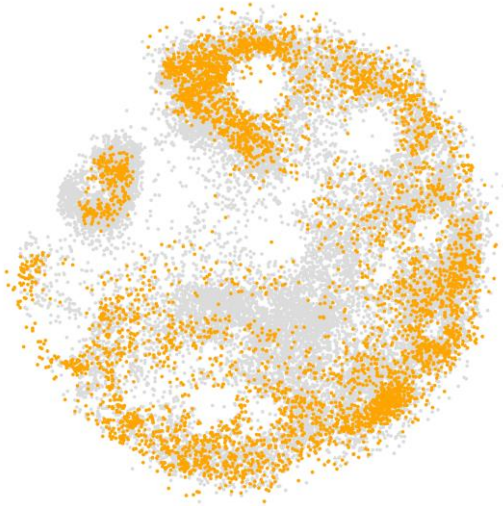


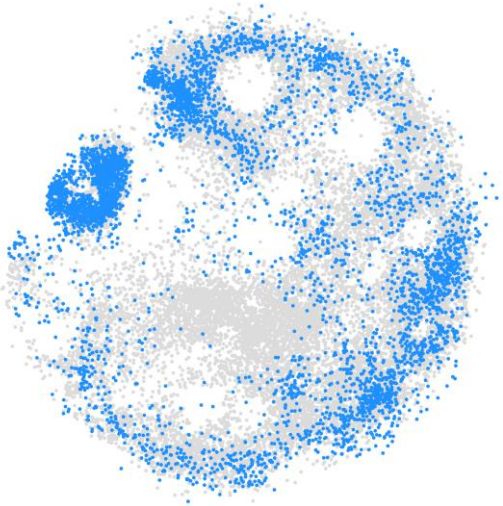
Monomers

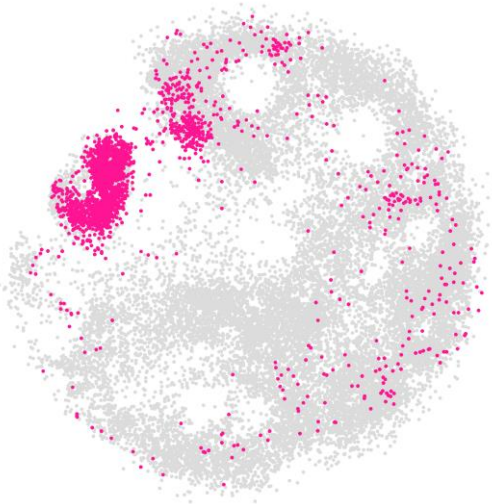
Dimers

Tetramers

P301L





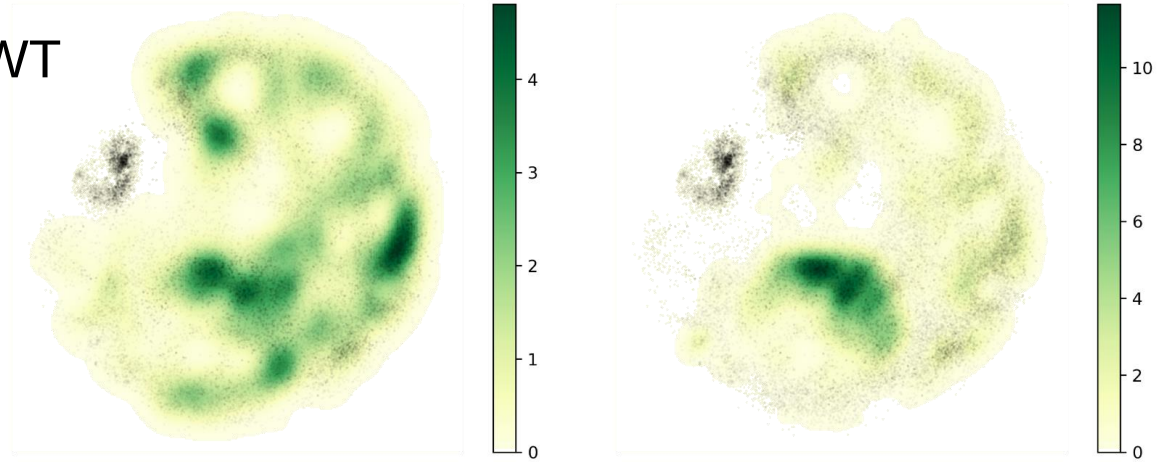


Density of States

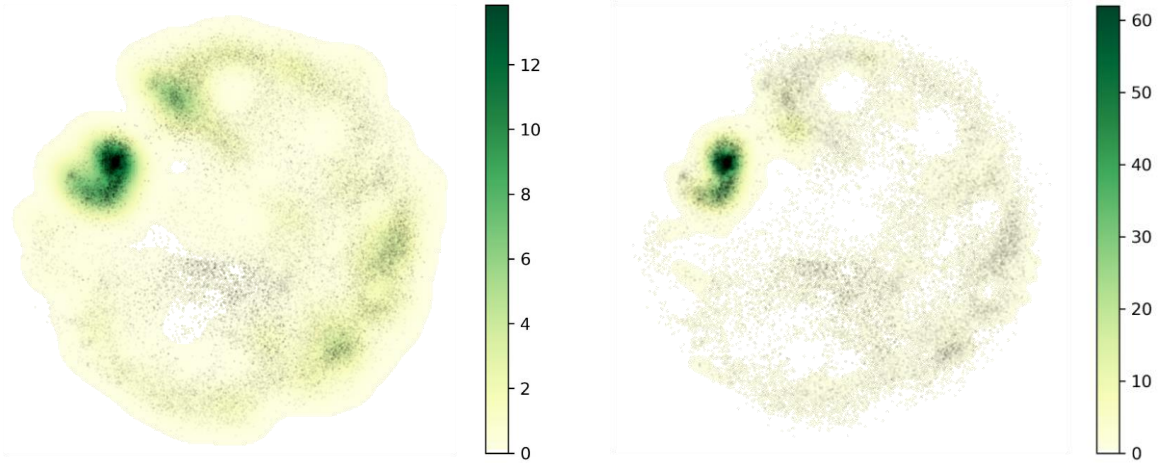
Dimers

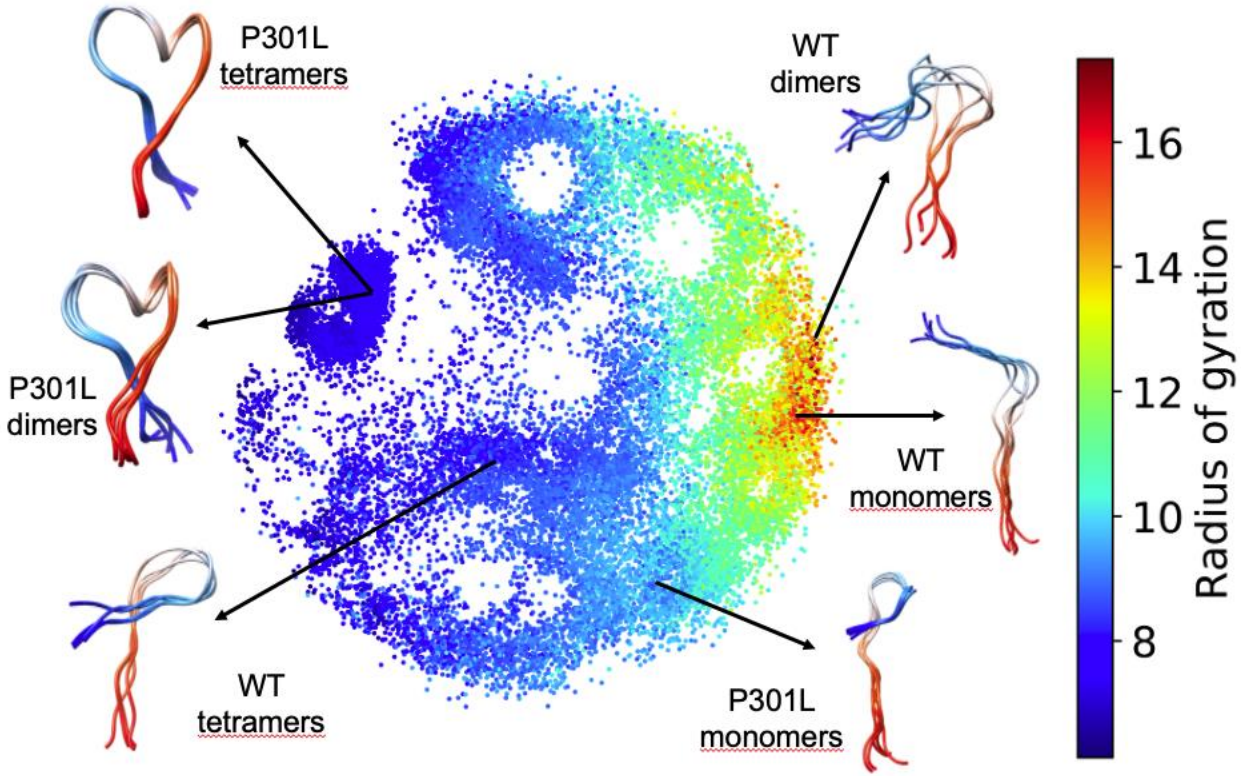
Tetramers

WT

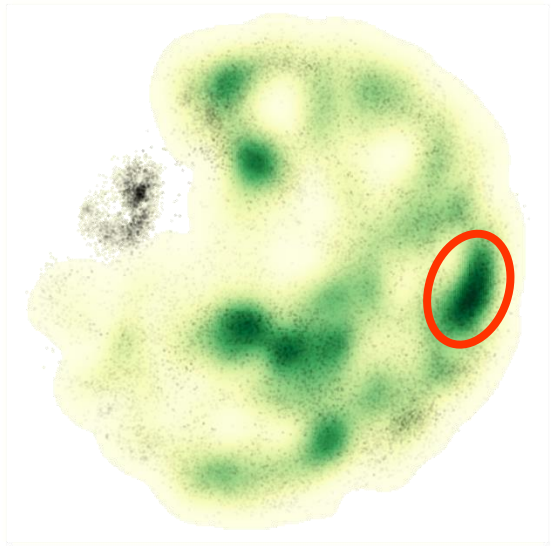


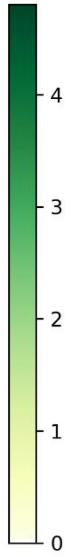
P301L

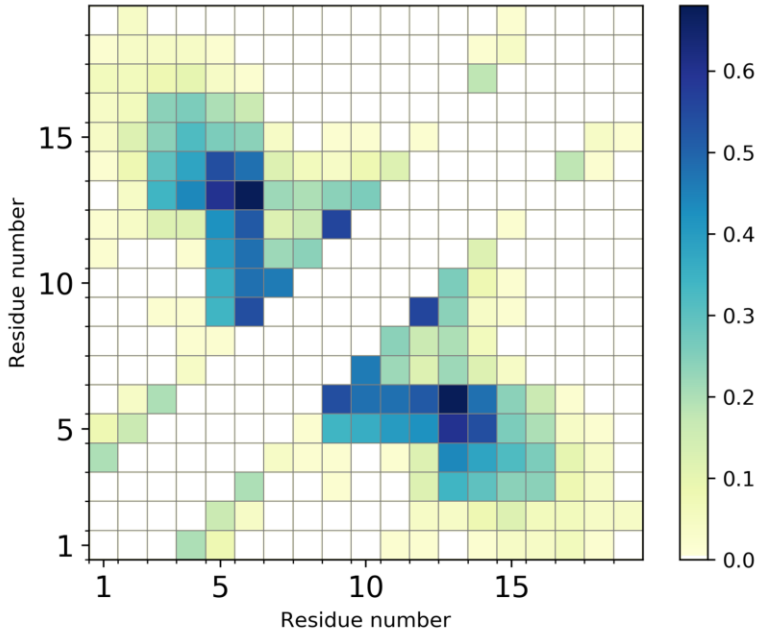


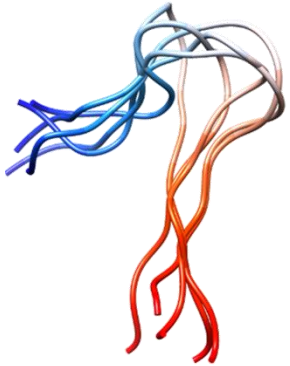


WT dimers

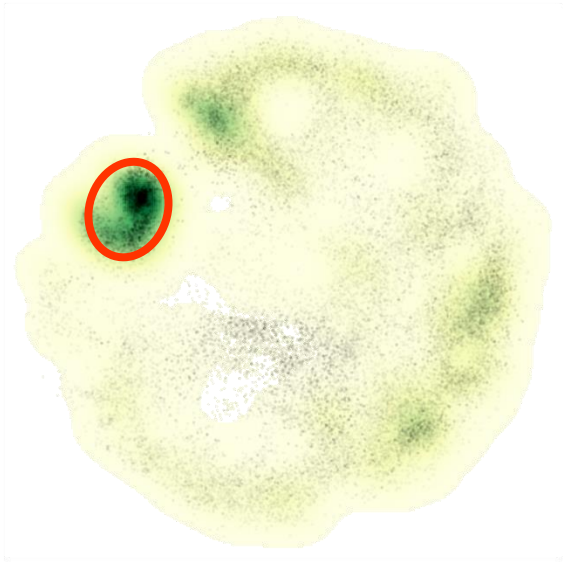


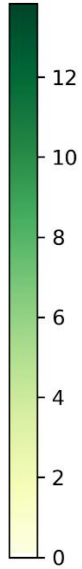


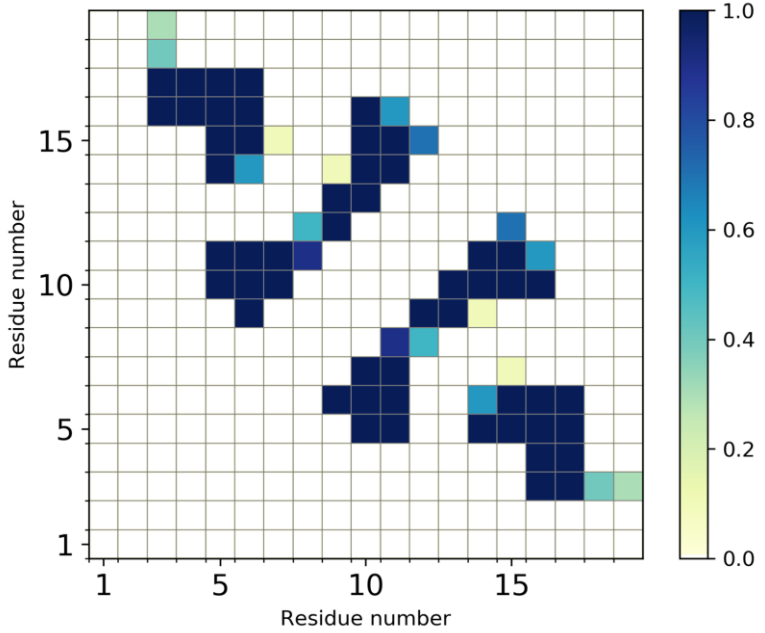


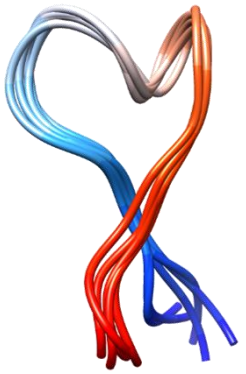


P301L dimers









Conclusion

"Progress in science depends on new techniques, new discoveries and new ideas, probably in that order."

— Sydney Brenner

Energy Landscape Visualization Method (ELViM):

- Reaction-Coordinate free
- Can be used with other sampling methods
- Conformation dependent only
- Can be used for any resolution

  (C_alpha, All-atom, large units)
- Different systems (e.g. RNA, DNA,

  biomolecular assemblies, chromatin)
- Code available and soon available as a

  web-server!!!!!

IDPs:

- Single chain (under different conditions)

  - differential analysis
- Functional mechanisms
- Effective metric complex systems

  - e.g. aggregation
- Oligomers and fiber formation

  - Amyloid-β, Tau, etc
- ~ 30% of proteome are IDPs or IDRs

  proteinensemble.org