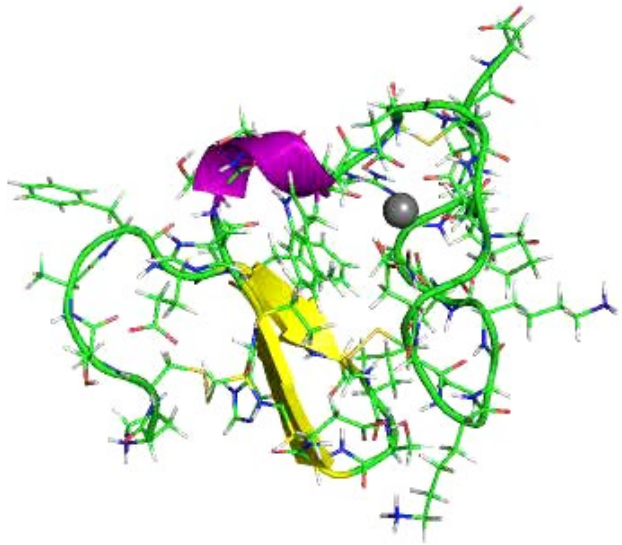
DAY 3 (5 July, 2023) 15:00/17:00 Molecular Dynamics simulations and stability calculation Javier Sancho

- MD simulations
- Observing unfolding: relaxation MD
- Calculation of ΔH and ΔCp
- Calculation of ΔG
- Biophysical insight
- Calculation of ΔΔG

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

MD simulation

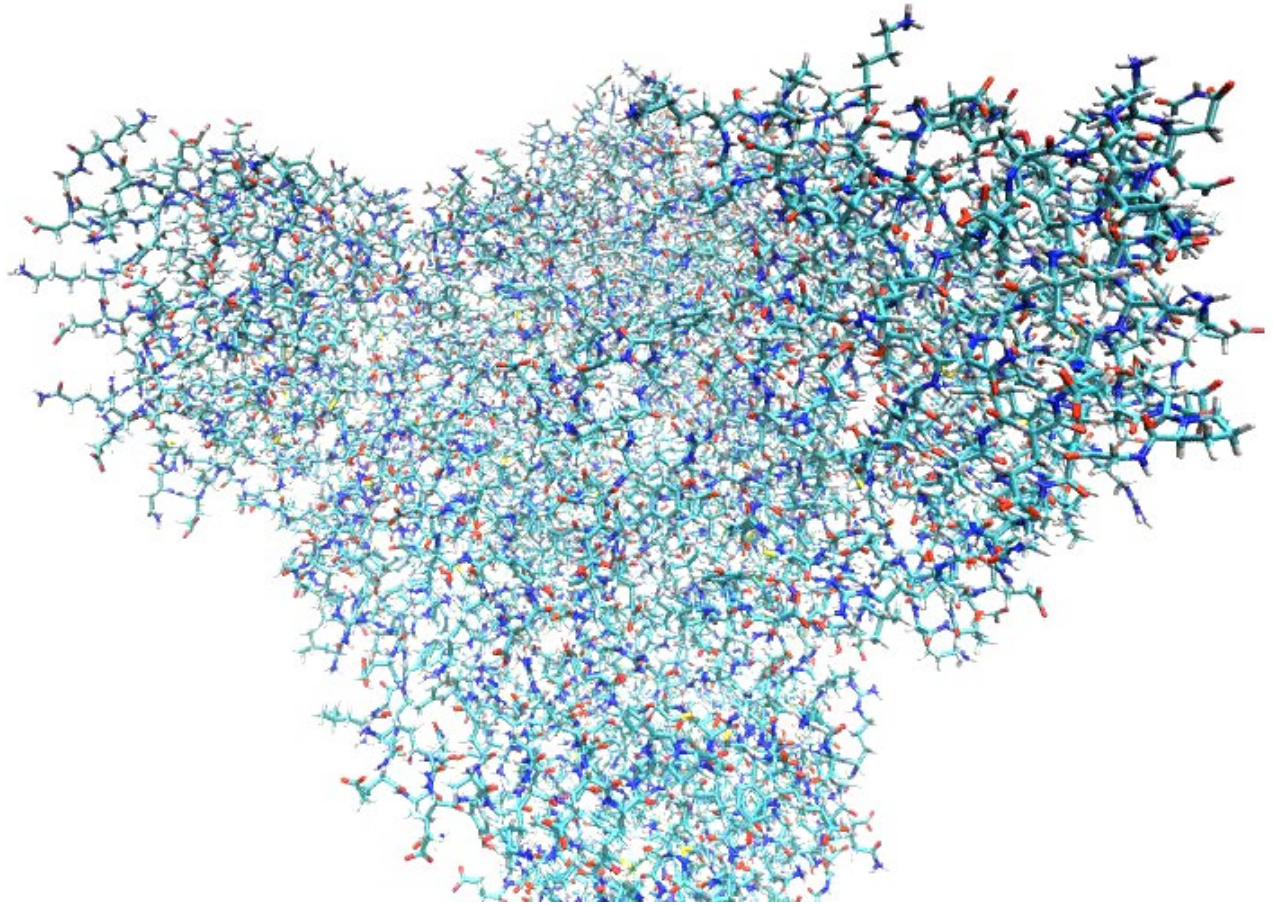


LR5 from LDL receptor (wt)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

MD simulation

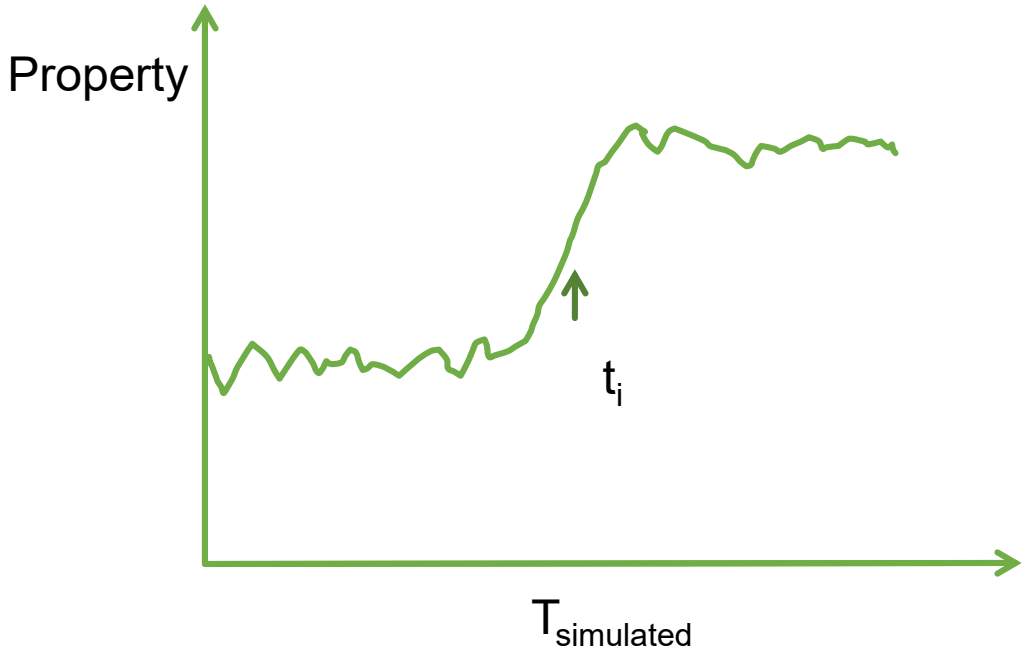


MCAD (K304E)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Observing unfolding: relaxation MD



A change is modeled in the WT structure (usually an AA substitution) and the variant protein is simulated to observe how/if its conformation relaxes into something diferent (e.g. it unfolds)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Observing unfolding: relaxation MD

Variation

- RMSD / TM-score / RMSDist
- RMSF

Structure

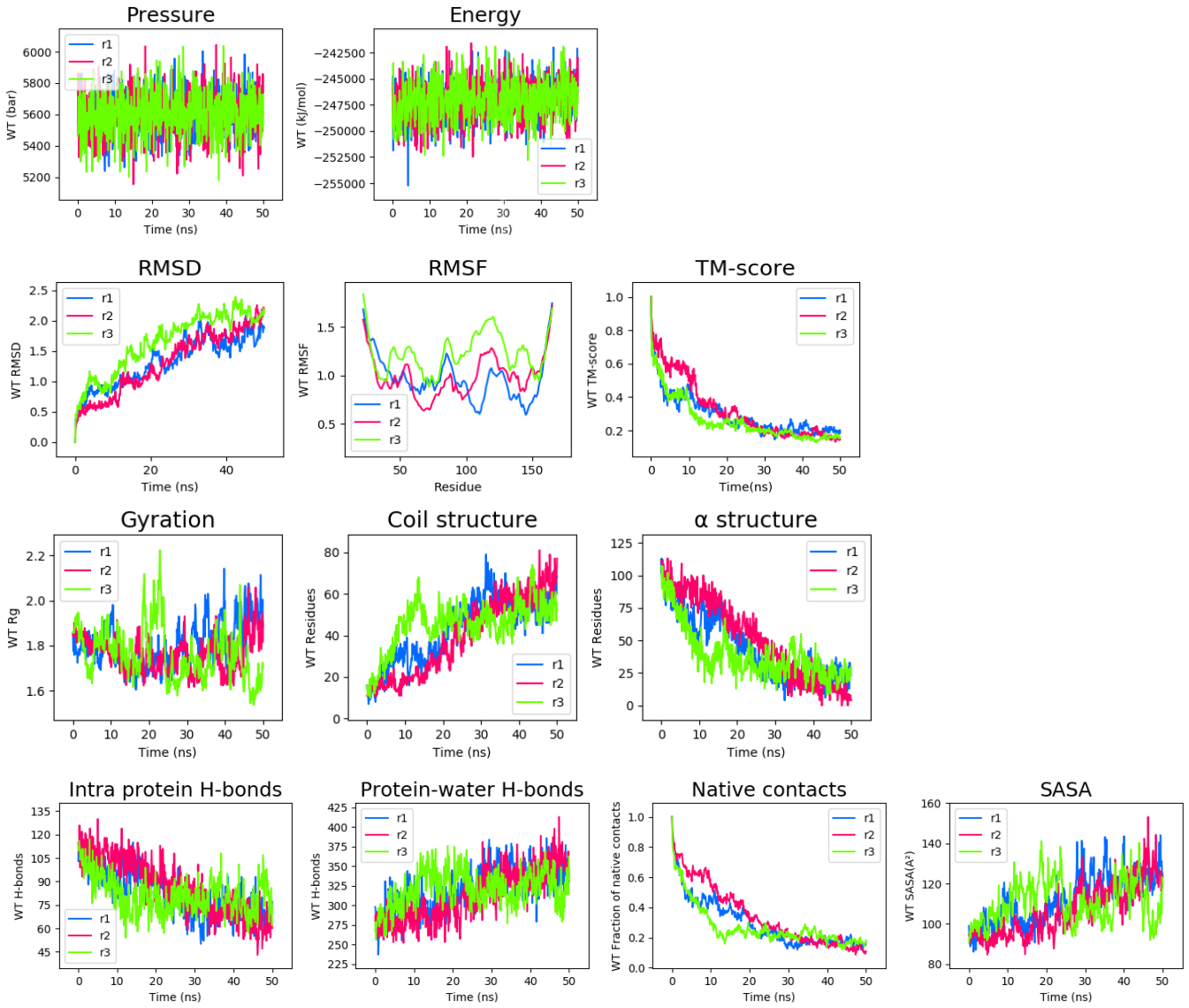
- Radius of gyration
- SASA
- Secondary structure

Interactions

- Hydrogen bonds
- Native contacts

Energetics

- Energy
- Pressure

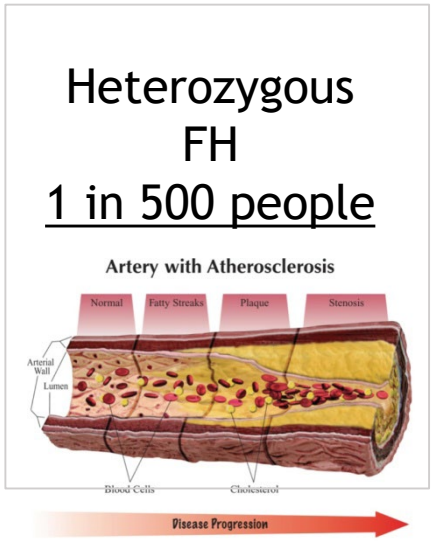


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

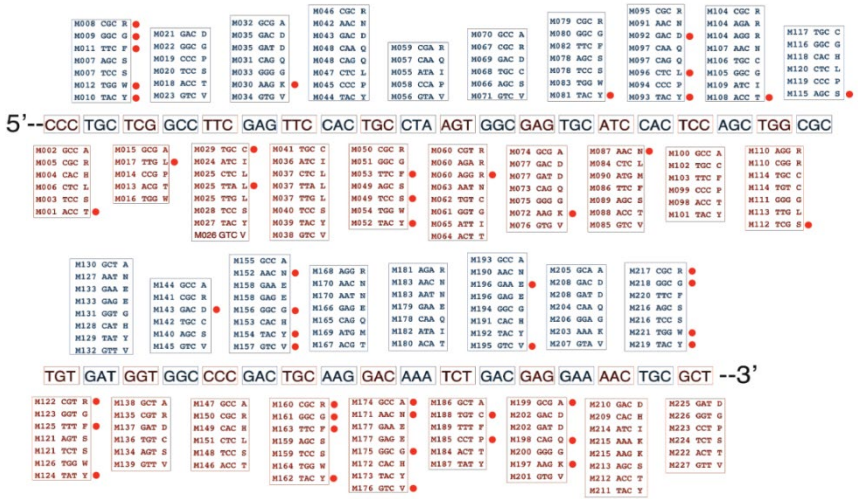
MOSBRI
Molecular-Scale Biophysics Research Infrastructure

Observing unfolding: relaxation MD

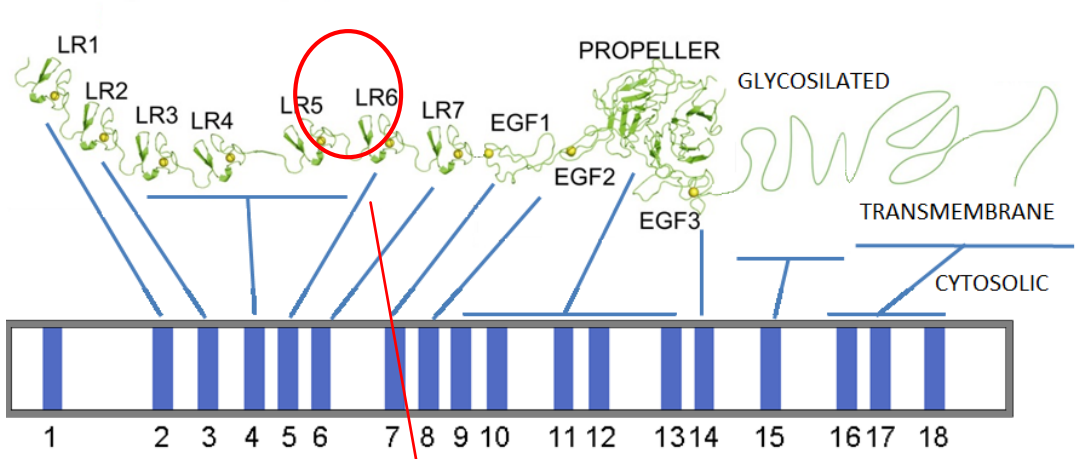
Heterozygous
FH
1 in 500 people

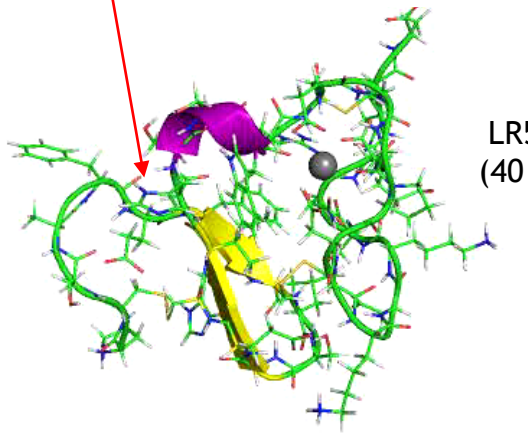


227 possible AA point mutations arising from SNPs in LR5



LDL receptor





LR5 module
(40 residues)

Computational diagnosis of protein conformational disease: Short MD simulations reveal a fast unfolding of r-LDL mutants that cause Familial Hipercolesterolemia. S. Cuesta-López, F. Falo & J. Sancho. *Proteins.* 66:87-95 (2007).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

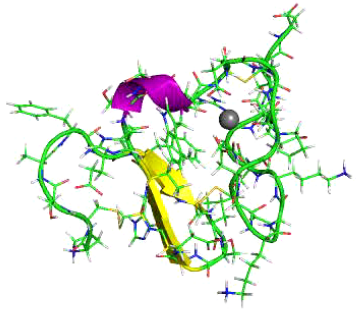
MOSBRI Molecular-Scale Biophysics Research Infrastructure

Observing unfolding: relaxation MD

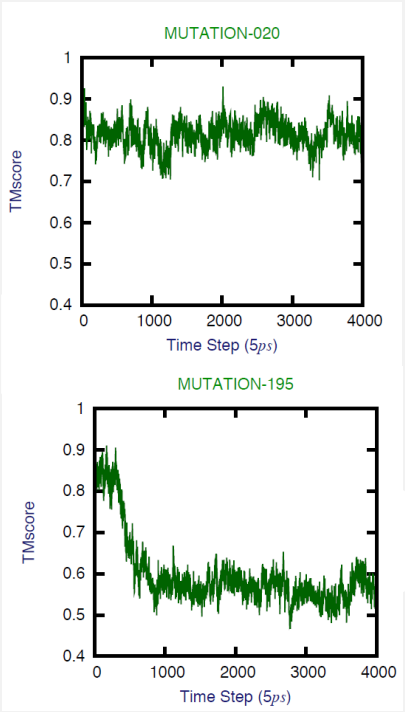
System configuration for Molecular Dynamics

- Mutants generated from crystal structure (PDB: 1ajj) with SCWRL for rotamer selection
- Solvation in a cubic box (~ 5500 TIP3 molecules) and neutralization with Na^+ counterions (blue spheres)
- Preparation phase: 3 ns of step-descending minimization/equilibration steps
- Production phase: 20 ns of Langevin MD simulations using NAMD and CHARMM periodic boundary conditions and PME

A total of 6μs simulation (865 khours of CPU time run in the MareNostrum Computer during 3 years).



INDIVIDUAL TM-SCORES



Principal component analysis (PCA)
Frecuency (low, high) of conformations along the simulation

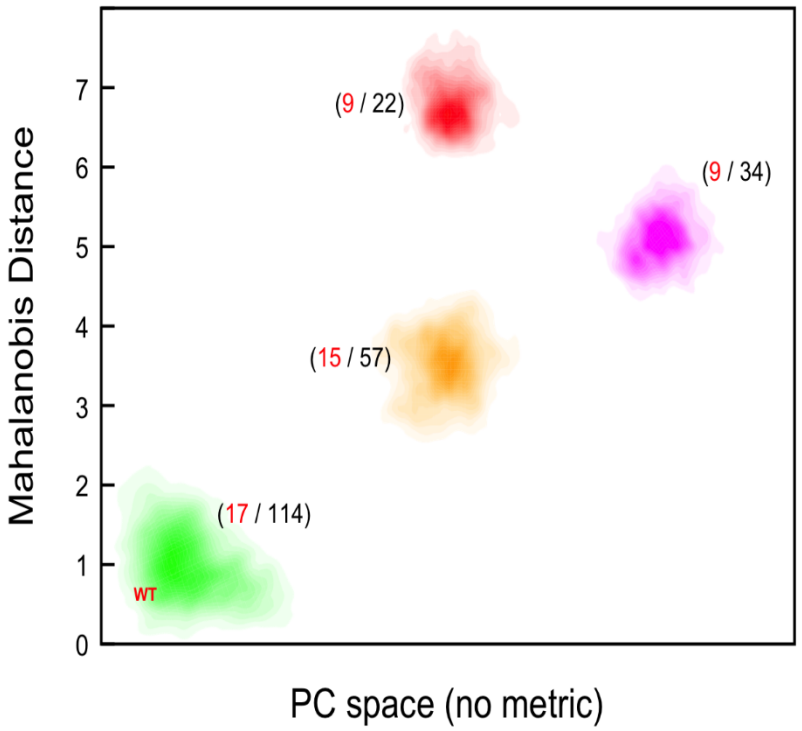
STABLE MUTATION **UNSTABLE MUTATION**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

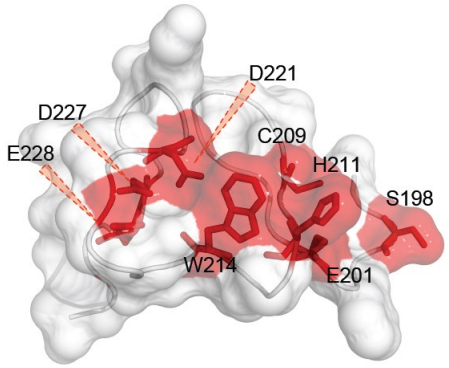
Observing unfolding: relaxation MD

Clustering of 227 mutants from a meta trajectory and distribution of 50 known FH mutations



wt → mt	wt → mt	Index	PMUT	CONDEL	Cluster	Phenotype
GCC→TCC	A199{178}S	M020	N	N	●	N
GCC→GAC	A199{178}D	M021	N	N	●	N
GCC→GGC	A199{178}G	M022	N	N	●	N
GCC→GTC	A199{178}V	M023	N	N	●	D
TTC→ATC	F200{179}I	M024	N	N	●	N
TTC→CTC *TTC→TTA TTC→TTG	F200{179}L	M025	N	N	●	D
TTC→GTC	F200{179}V	M026	N	N	●	D
TTC→TAC	F200{179}Y	M027	N	N	●	N
TTC→TCC	F200{179}S	M028	N	N	●	N
*TTC→TGC	F200{179}C	M029	N	N	●	D
*GAG→AAG	⊕E201{180}K	M030	N	N	●	*D*
GAG→CAG	⊕E201{180}Q	M031	N	D	●	*D*
GAG→GCG	⊕E201{180}A	M032	N	D	●	*D*
GAG→GGG	⊕E201{180}G	M033	N	D	●	*D*
GAG→GTG	⊕E201{180}V	M034	N	D	●	*D*
GAG→GAT GAG→GAC	⊕E201{180}D	M035	N	D	●	*D*
TTC→ATC	F202{181}I	M036	N	D	●	D
TTC→TTG TTC→TTA TTC→CTC	F202{181}L	M037	N	N	●	D

Residues that bear 16 out of the 17 non-destabilizing known mutations



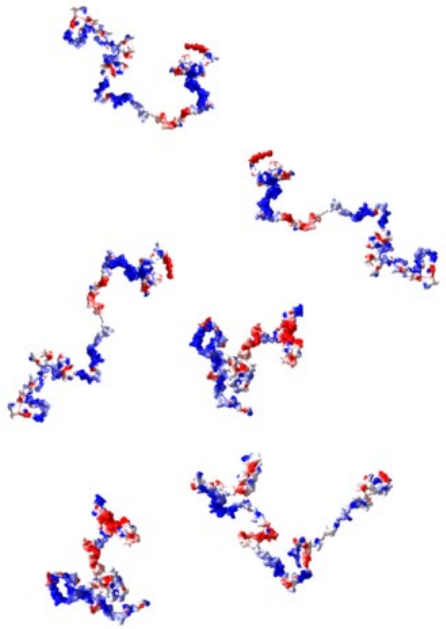
Computational diagnosis of protein conformational disease: Short MD simulations reveal a fast unfolding of r-LDL mutants that cause Familial Hipercolesterolemia. S. Cuesta-López, F. Falo & J. Sancho. *Proteins.* 66:87-95 (2007).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

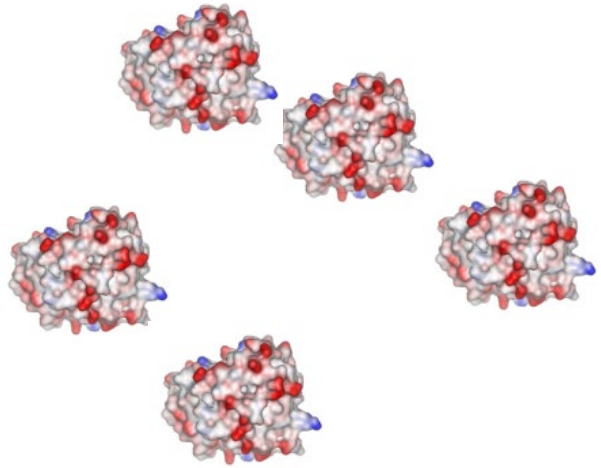
MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔH and ΔCp

H_u



$\Delta H,\ \Delta C_p,\ \Delta S,\ \Delta G$

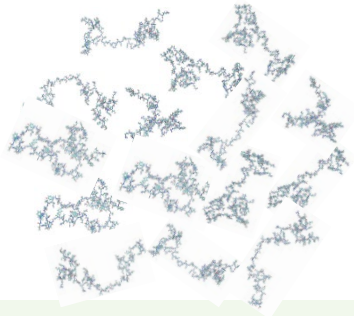


H_f

```
          10         20         30         40         50         60
MANLGCWMLV LFVATWSDLG LCKKRPKPGG WNTGGSRYPG QGSPGGNRYP PQGGGGWGQP

          70         80         90        100
HGGGWGQPHG GGWGQPHGGG WGQPHGGGWG QGGGTHSQWN KPSKP
```

ProtSA



Unfolded ensemble

http://webapps.bifi.es/

RCSB PDB PROTEIN DATA BANK

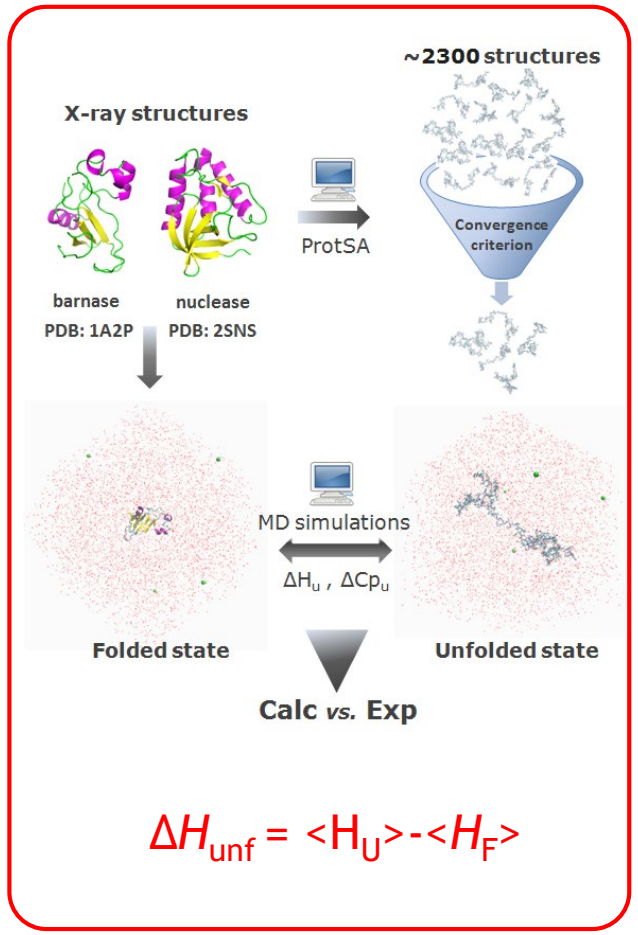
Native state

X-ray structure, PDB

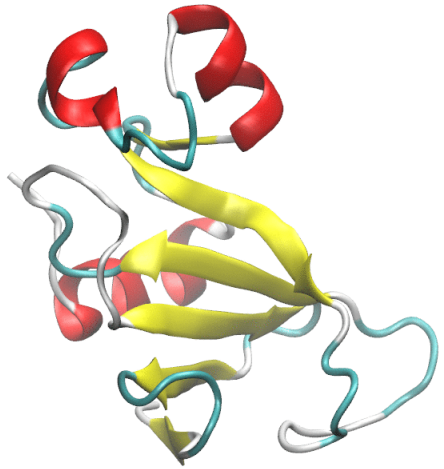
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

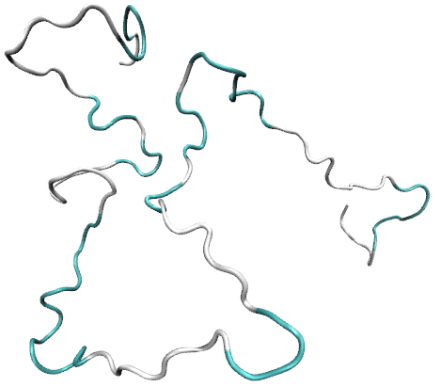
MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔH and ΔCp



$\Delta H_{unf} = \langle H_U \rangle - \langle H_F \rangle$



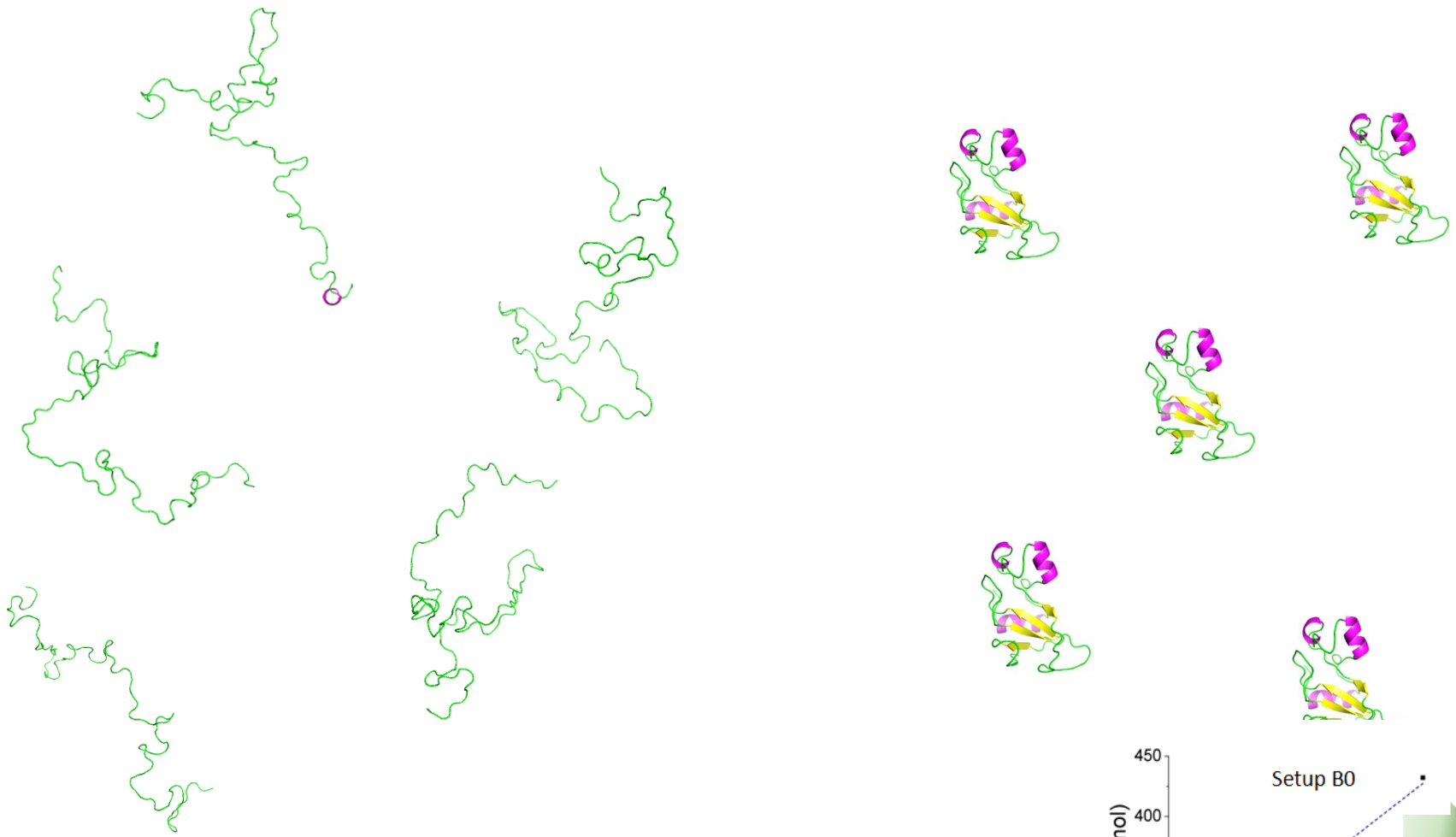


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

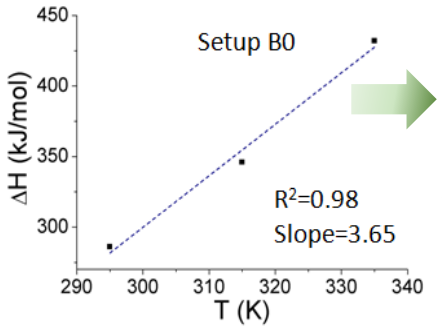
Calculation of ΔH and ΔCp

H_u H_f



$\Delta H_{unf} = H_u - H_f$

$\Delta Cp_{unf} = \delta \Delta H_{unf} / \delta T$



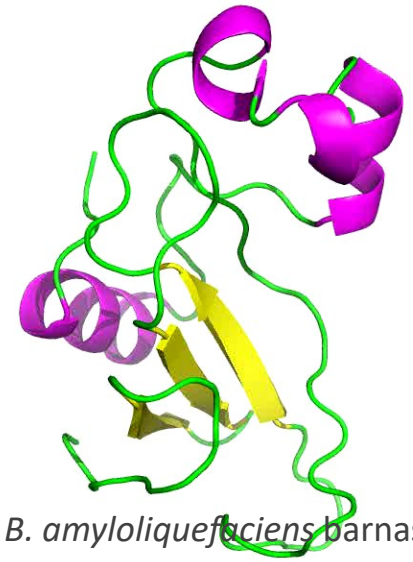
ΔC_p

Accurate Calculation of Barnase and SNase Folding Energetics Using Short Molecular Dynamics Simulations and an Atomistic Model of the Unfolded Ensemble: Evaluation of Force Fields and Water Models
Galano-Frutos & Sancho, J. *Chem. Inf. Model.* 59:4350-4360 (2019)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 1010048

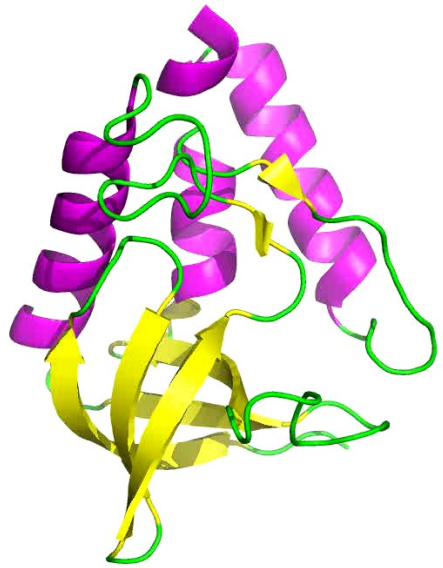
MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔH and ΔCp



B. amyloliquefaciens barnase

PDB: **1A2P**



S. aureus nuclease

PDB: **2SNS**

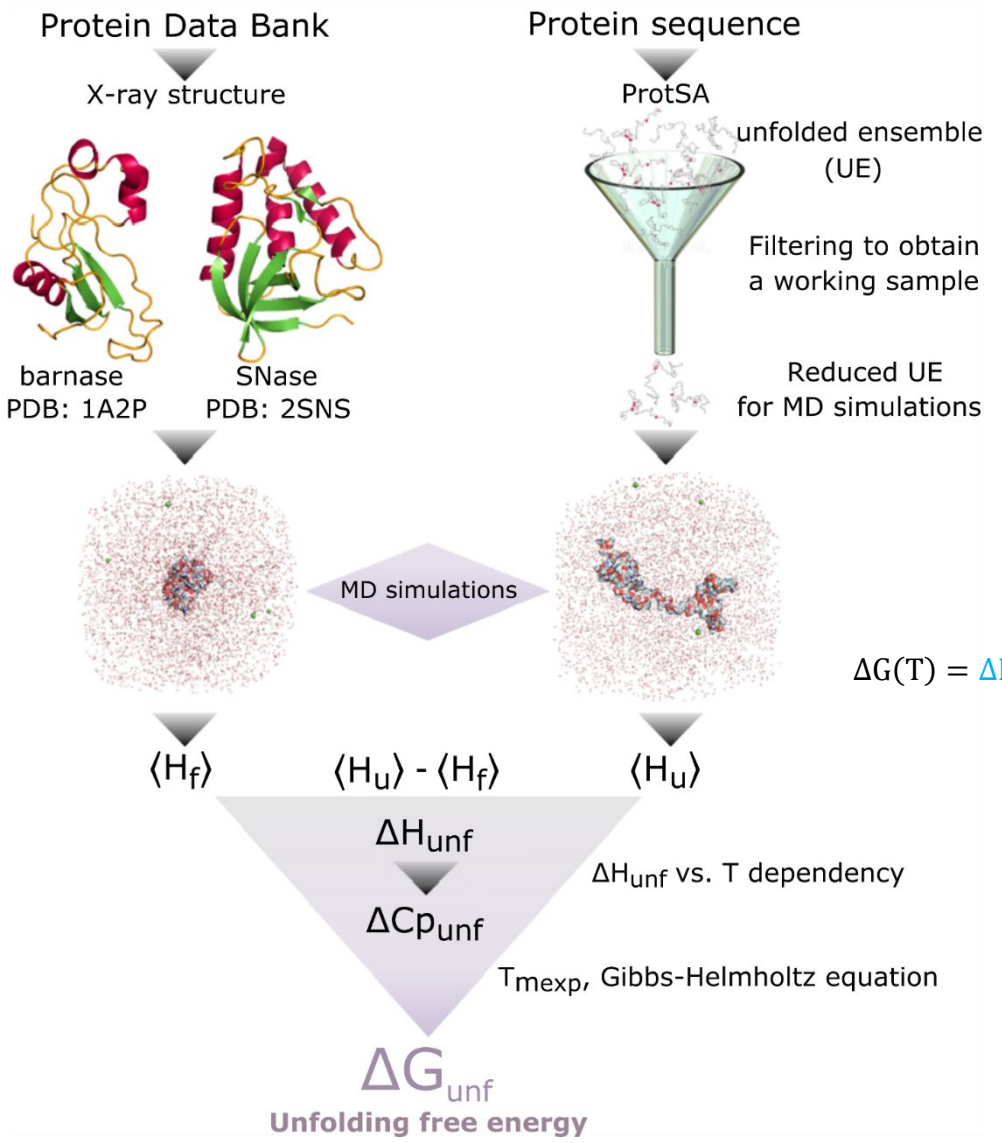
		ΔH_{unf} (kJ/mol)		ΔCp_{unf} (kJ/molK)	
Protein	Temp (K)	Experimental	Calculated	Experimental	Calculated
Barnase	315	427 (±24)	346±27	5.3 ± 0.7	3.7±0.4
Nuclease	317	248 ± 25	268±33	6.1 ± 2.0	7.4±0.1

Accurate Calculation of Barnase and SNase Folding Energetics Using Short Molecular Dynamics Simulations and an Atomistic Model of the Unfolded Ensemble: Evaluation of Force Fields and Water Models
Galano-Frutos & Sancho, J. *Chem. Inf. Model.* 59:4350-4360 (2019)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔG



Gibbs- Helmholtz equation

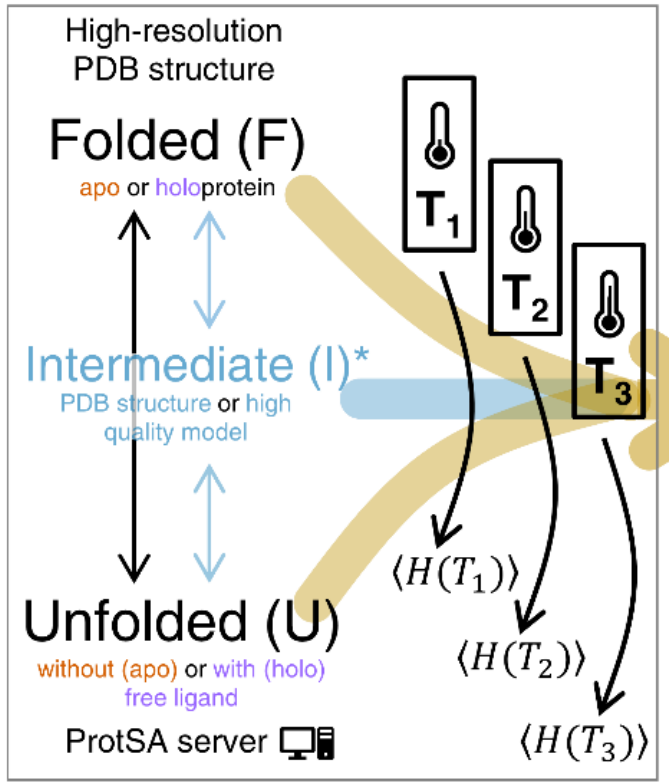
$$\Delta G(T) = \Delta H_{Tm} \times (1 - T/T_m) - \Delta Cp \times [T_m - T + T \times \ln(T/T_m)]$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

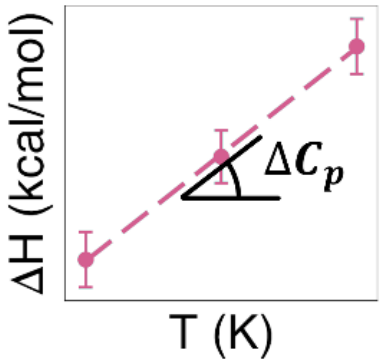
Calculation of ΔG

$$\Delta G = \Delta H_{Tm} \times (1 - T/T_m) - \Delta C_p \times [T_m - T + T \times \ln(T/T_m)]$$



Folding states

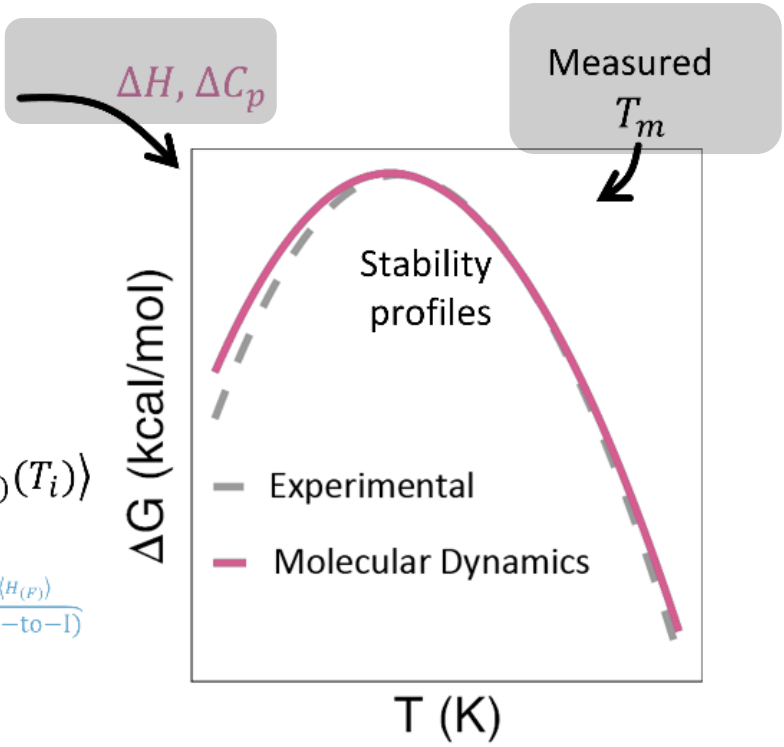
Molecular Dynamics



$$\Delta H_{unf}(T_i) = \langle H_{(U)}(T_i) \rangle - \langle H_{(F)}(T_i) \rangle$$

$$^{*}\Delta H_{unf}(T_i) = \frac{\langle H_{(U)} \rangle - \langle H_{(I)} \rangle}{\Delta H_{unf}(\text{I-to-U})} + \frac{\langle H_{(I)} \rangle - \langle H_{(F)} \rangle}{\Delta H_{unf}(\text{F-to-I})}$$

ΔCp: Linear relationship ΔH vs. T



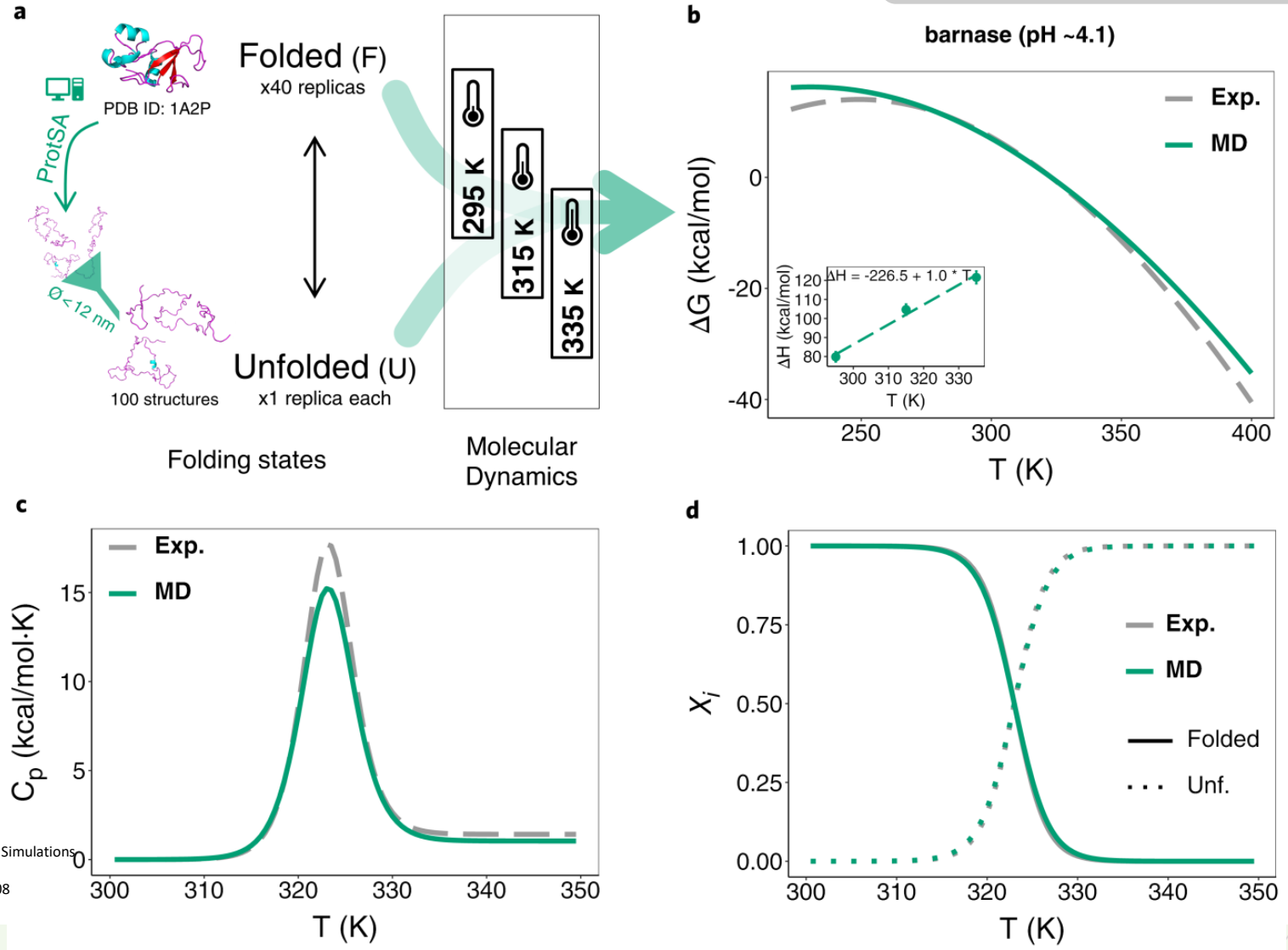
Integrated Gibbs-Helmholtz equation for ΔG(T)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔG

BARNASE (two-state)



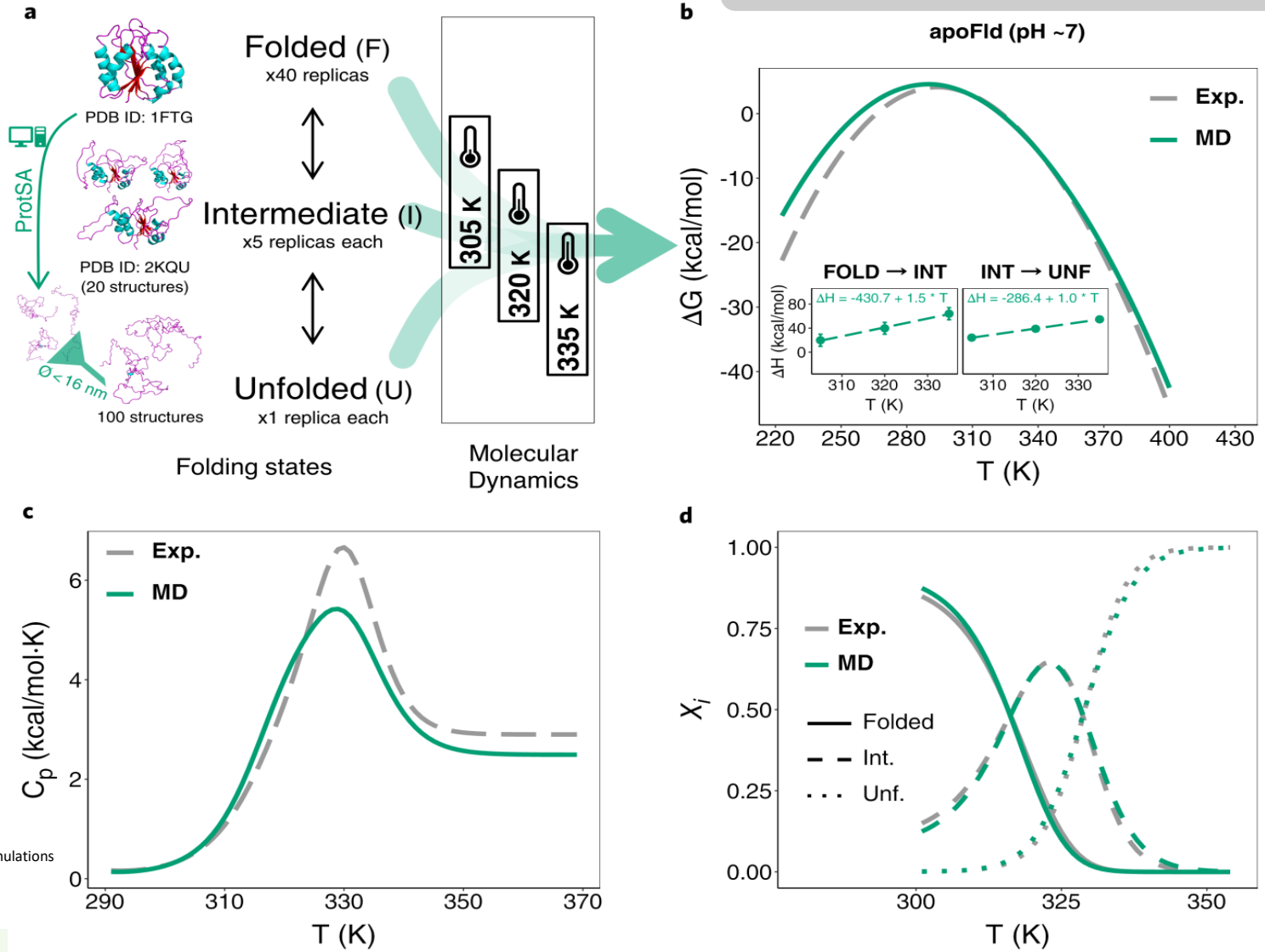
Calculation of Protein Folding Thermodynamics using Molecular Dynamics Simulations
Juan José Galano-Frutos, Francho Nerín-Fonz, Javier Sancho
bioRxiv 2023.01.21.525008; doi: https://doi.org/10.1101/2023.01.21.525008

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔG

APOFLAVODOXIN (three-state)

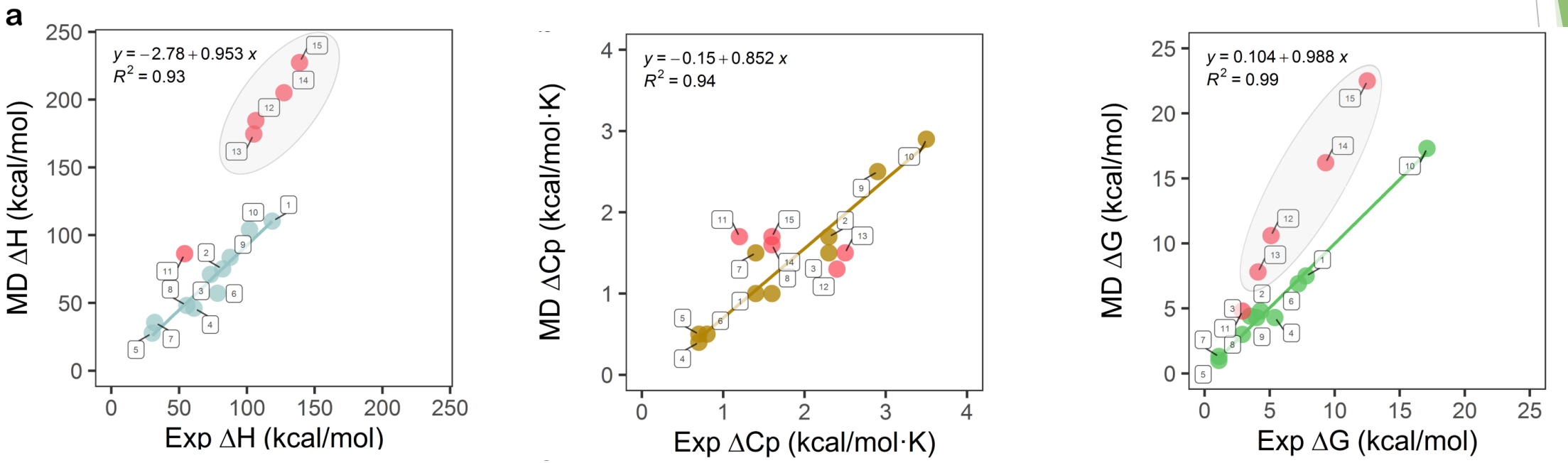


Calculation of Protein Folding Thermodynamics using Molecular Dynamics Simulations
Juan José Galano-Frutos, Francho Nerín-Fonz, Javier Sancho
bioRxiv 2023.01.21.525008; doi: https://doi.org/10.1101/2023.01.21.525008

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔG



Calculation of Protein Folding Thermodynamics using Molecular Dynamics Simulations
Juan José Galano-Frutos, Francho Nerín-Fonz, Javier Sancho
bioRxiv 2023.01.21.525008; doi: https://doi.org/10.1101/2023.01.21.525008

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔG

Roads to ΔG from MD simulation

A- From the equilibrium constant K

$\Delta G = -kT \times \ln(K_{unf})$

B- From ΔH *and* ΔS

$\Delta G = \Delta H - T \times \Delta S$

ΔS

C- From ΔH, ΔC_p and T_m

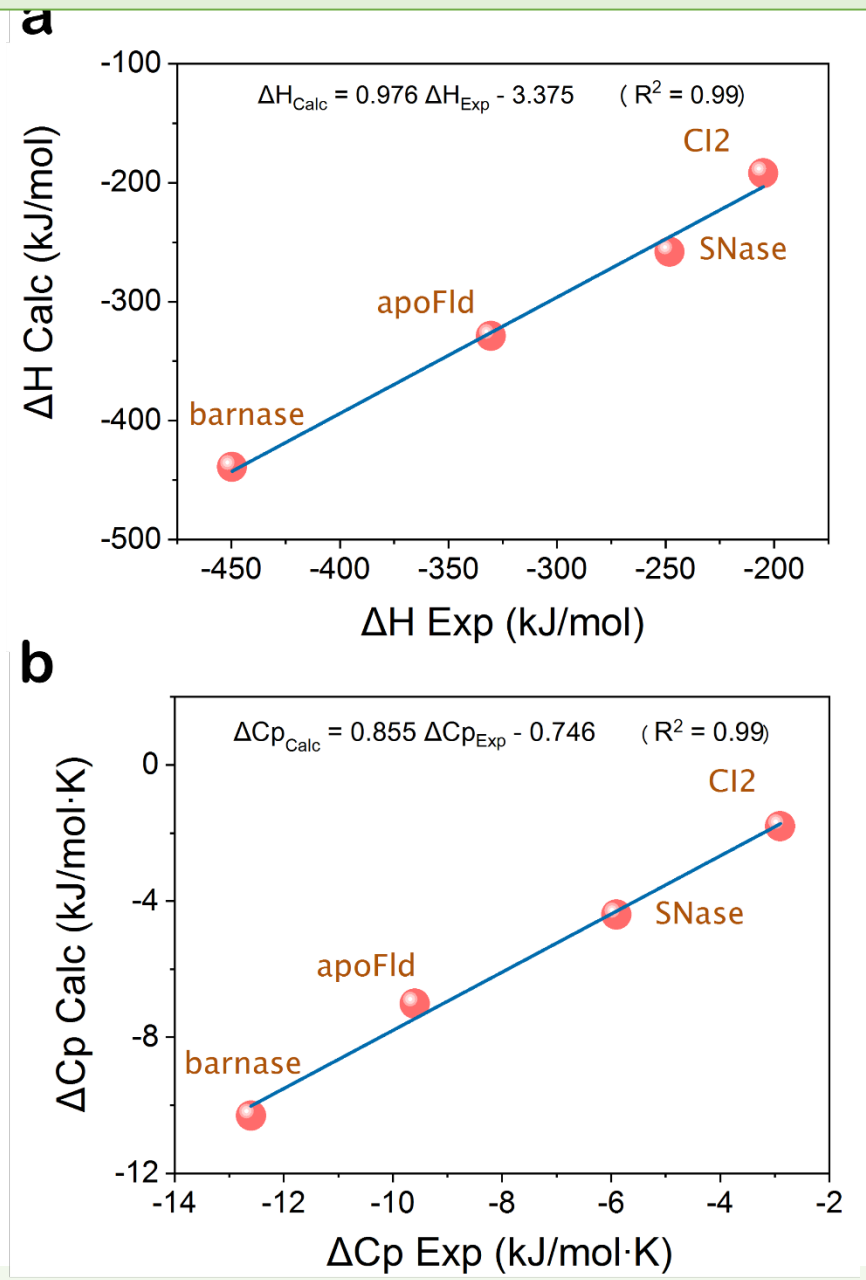
$\Delta G = \Delta H_{Tm} \times (1 - T/T_m) - \Delta C_p \times [T_m - T + T \times \ln(T/T_m)]$

T_m

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Biophysical insight

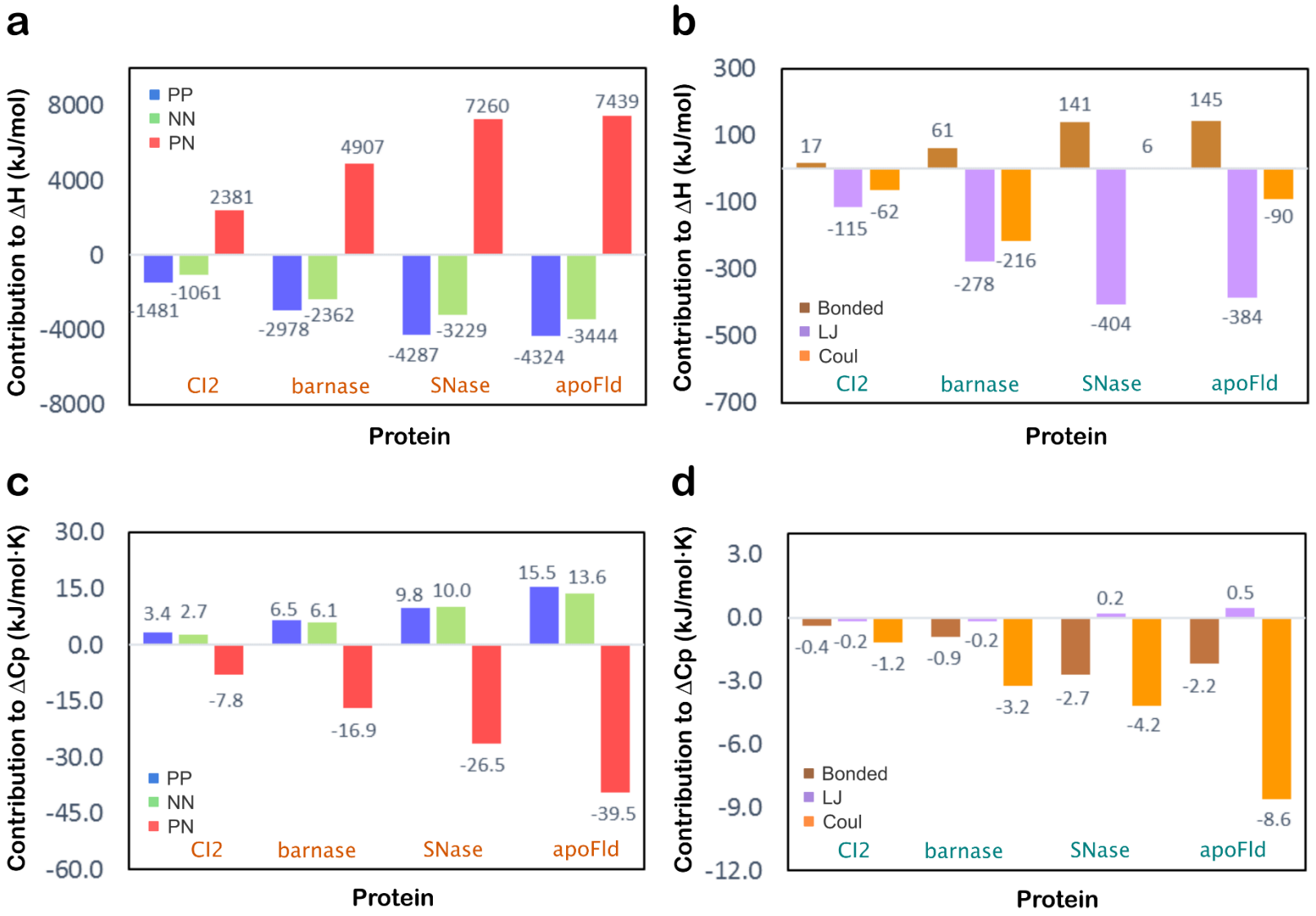


Molecular Interactions and Forces that Make Proteins Stable: A Quantitative Inventory from Atomistic Molecular Dynamics Simulations
Juan José Galano-Frutos, Javier Sancho
bioRxiv 2023.01.23.525230; doi: https://doi.org/10.1101/2023.01.23.525230

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

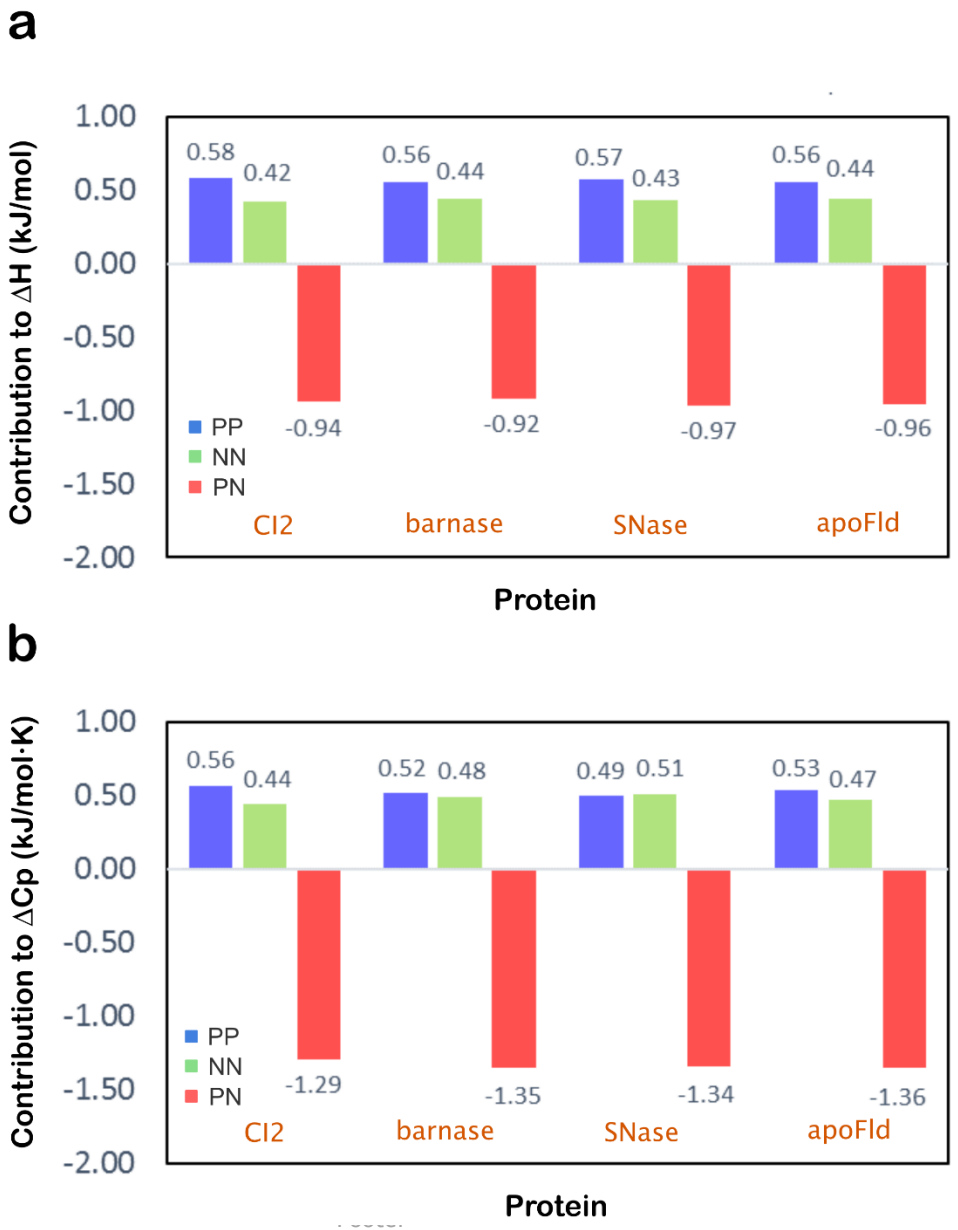
Biophysical insight



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

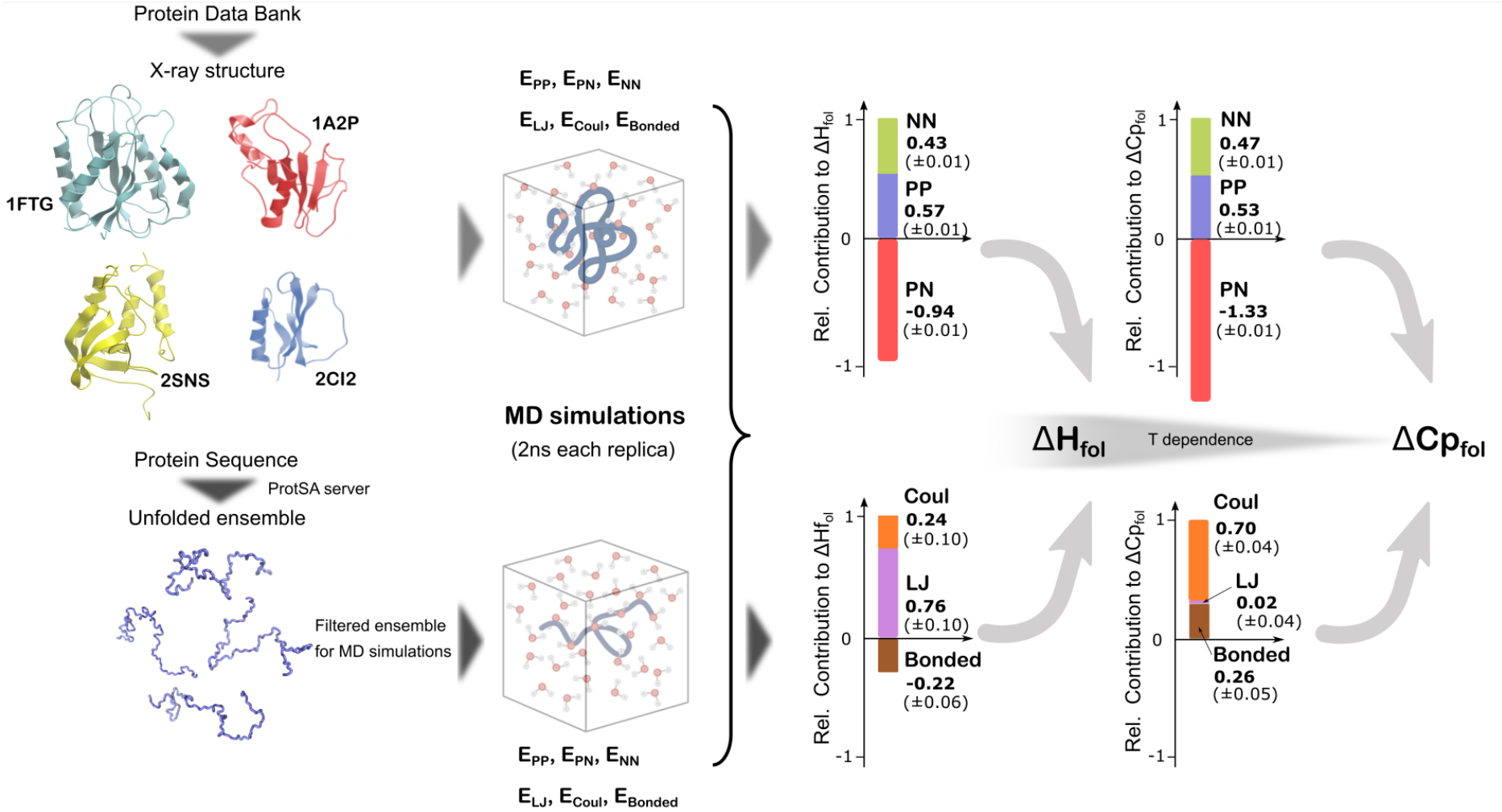
Biophysical insight



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

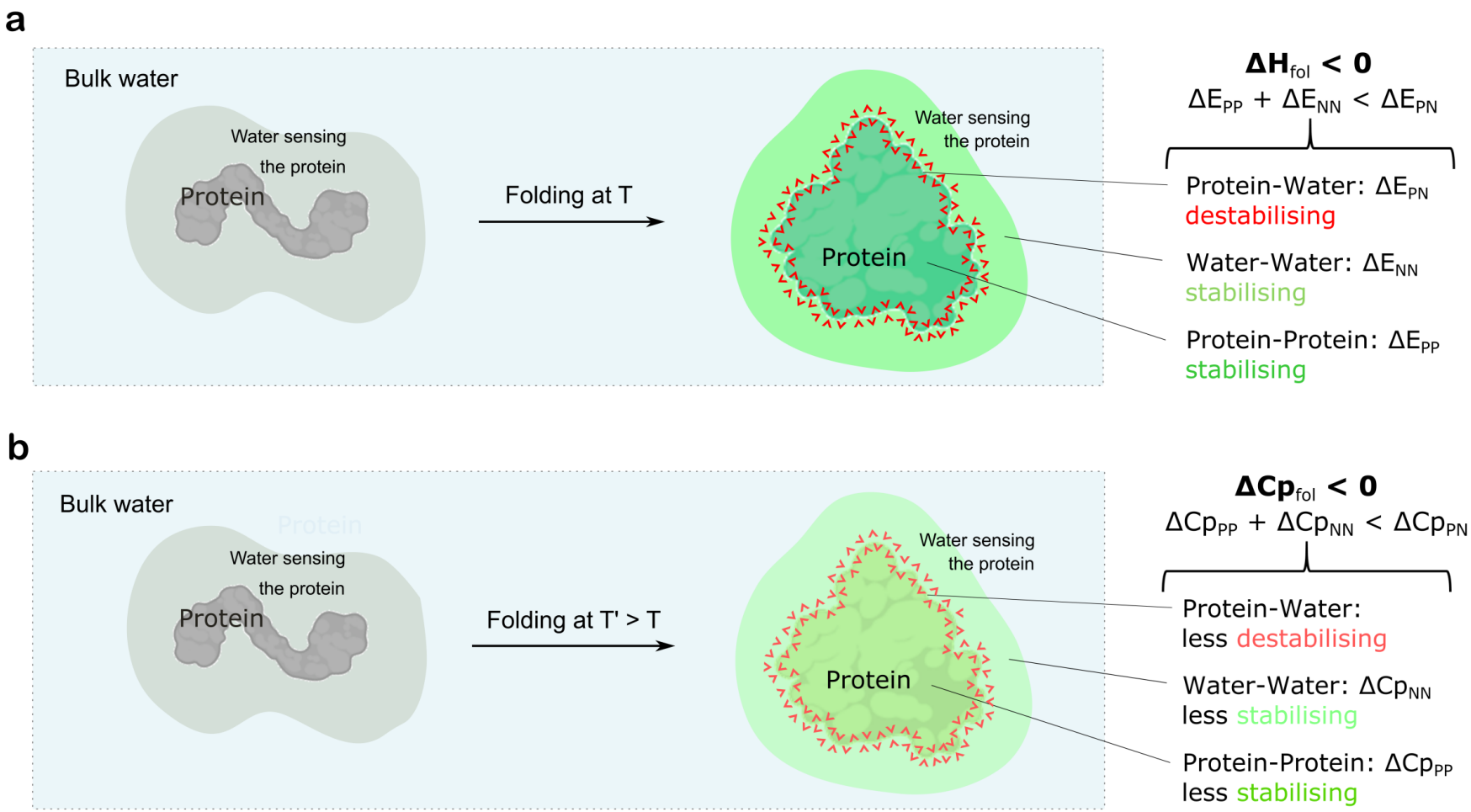
Biophysical insight



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Biophysical insight



Protein folding in water is driven by a strengthening of van der Waals and coulombic interactions, at the cost of introducing some strain in the folded polypeptide.

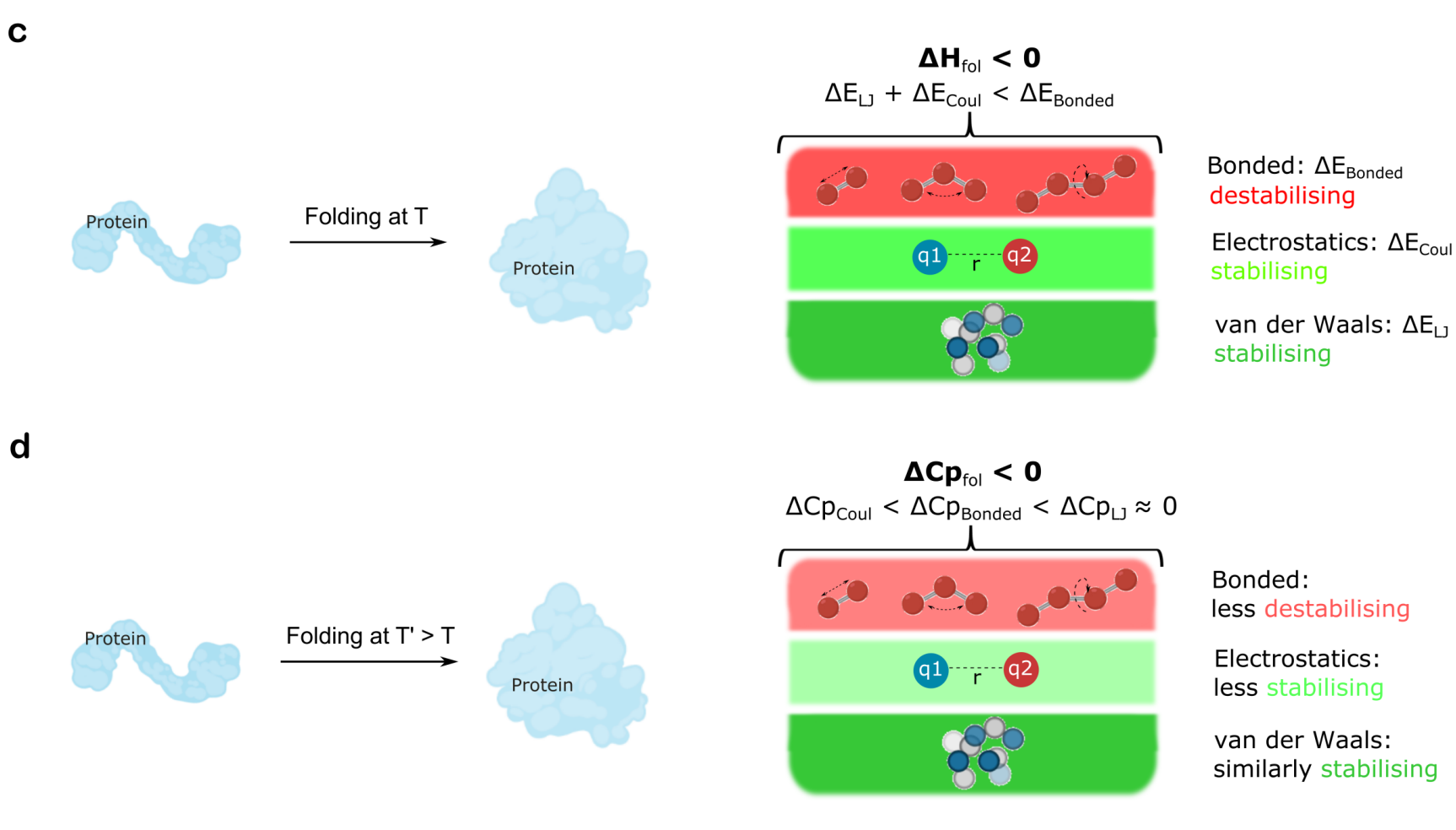
A half of the protein-solvent interactions that are lost when the protein folds are compensated by the new interactions that are established between solvent molecules.

•Molecular Interactions and Forces that Make Proteins Stable: A Quantitative Inventory from Atomistic Molecular Dynamics Simulations
Juan José Galano-Frutos, Javier Sancho
bioRxiv 2023.01.23.525230; doi: https://doi.org/10.1101/2023.01.23.525230

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Biophysical insight



Increasing temperature weakens both the stabilizing and destabilizing molecular contributions to ΔH_{fol} but it weakens more steeply the destabilizing protein-solvent interactions. The overall consequence is that enthalpy stabilizes the folded state more at higher temperatures.

The sign of the heat capacity change is specifically determined by protein-solvent interactions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔΔG

1) From ΔG (ΔΔG= ΔG-ΔG')

A- From the equilibrium constant *K*

$\Delta G = -kT \times \ln(K_{unf})$

B- From *ΔH and ΔS*

$\Delta G = \Delta H - T \times \Delta S$

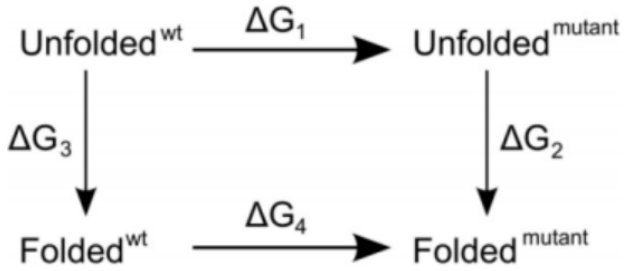
C- From *ΔH*, ΔC_p and T_m

$\Delta G = \Delta H_{Tm} \times (1 - T/T_m) - \Delta C_p \times [T_m - T + T \times \ln(T/T_m)]$

2) From free energy calculations

3) From unfolding simulations

2) From free energy calculations



$\Delta\Delta G = \Delta G_3 - \Delta G_2 = \Delta G_1 - \Delta G_4$

Experimentally, we would determine ΔG_3 and ΔG_2. The difference $\Delta\Delta G = \Delta G_3 - \Delta G_2$ is then the difference in folding free energy between wild type and mutant.

In the simulations, this would be too demanding. However, we can relatively accurately obtain the mutation free energies ΔG_1 and ΔG_4 from "alchemical" mutations in which we morph one amino acid into another and compute the according free energy change in the folded state (ΔG_4 and in the unfolded one (ΔG_1).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔΔG

1) From ΔG (ΔΔG= ΔG-ΔG')

A- From the equilibrium constant *K*

$\Delta G = -kT \times \ln(K_{unf})$

B- From *ΔH and ΔS*

$\Delta G = \Delta H - T \times \Delta S$

C- From *ΔH*, ΔC_p and T_m

$\Delta G = \Delta H_{Tm} \times (1 - T/T_m) - \Delta C_p \times [T_m - T + T \times \ln(T/T_m)]$

2) From free energy calculations

3) From unfolding simulations

3) From unfolding simulations

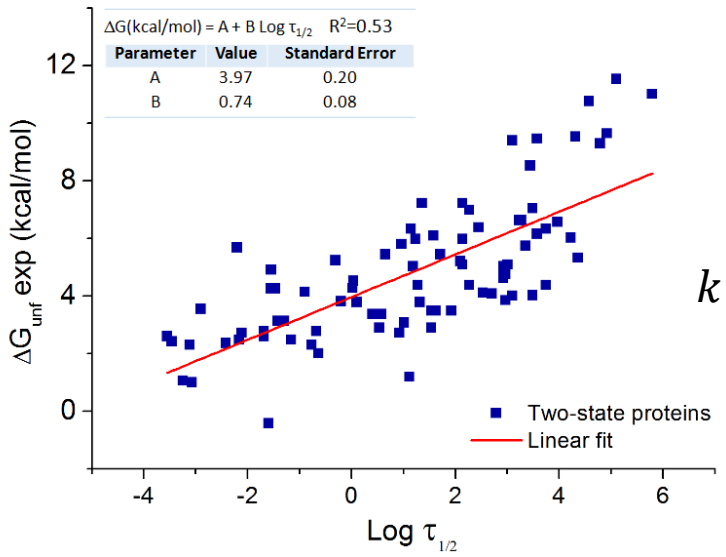
$$\Delta G_{unf} = a + b \times \log \frac{\ln 2}{-\frac{\ln(1-\chi_{unf})}{\tau_{unf}} \times \frac{298}{T} \times e^{\frac{\Delta G^{\ddagger}}{R} \times \left(\frac{1}{T} - \frac{1}{298}\right)}},$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

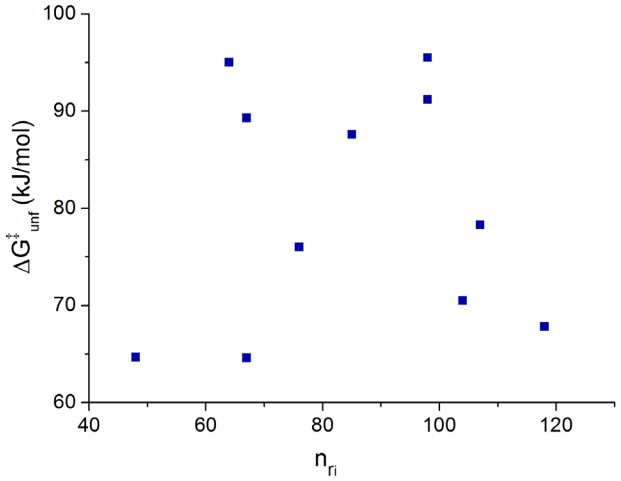
Calculation of ΔΔG

- ΔG linearly correlates to the log of half-life of unfolding at 298 K of 89 2-state proteins



- Eyring-Kramers equation for kinetic constants' dependence on T
- No correlation with protein length
- Average value for estimation (82.5±12.1 kJ/mol)

$$k = \frac{k_B T}{h} \times e^{-\left(\frac{\Delta G^{\ddagger}}{RT}\right)}$$



$$\Delta G_{unf} = a + b \times \log \frac{\ln 2}{-\frac{\ln(1-\chi_{unf})}{\tau_{unf}} \times \frac{298}{T} \times e^{\frac{\Delta G^{\ddagger}}{R} \times \left(\frac{1}{T} - \frac{1}{298}\right)}},$$

Galano-Frutos et al, *Briefings in Bioinformatics*, **22**:3–19 (2021)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

Calculation of ΔΔG

Variation

- RMSD / TM-score / RMSDist
- RMSF

Structure

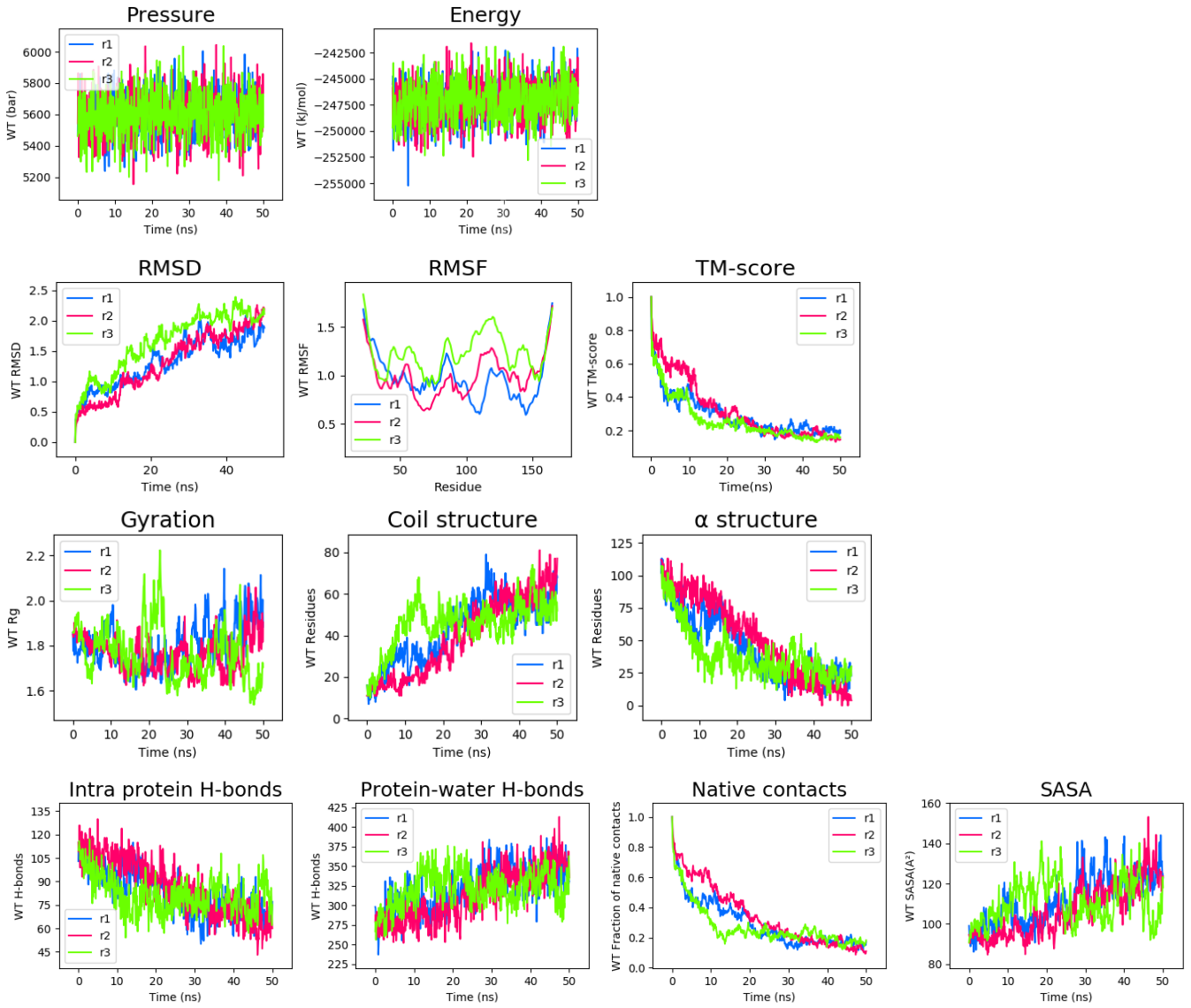
- Radius of gyration
- SASA
- Secondary structure

Interactions

- Hydrogen bonds
- Native contacts

Energetics

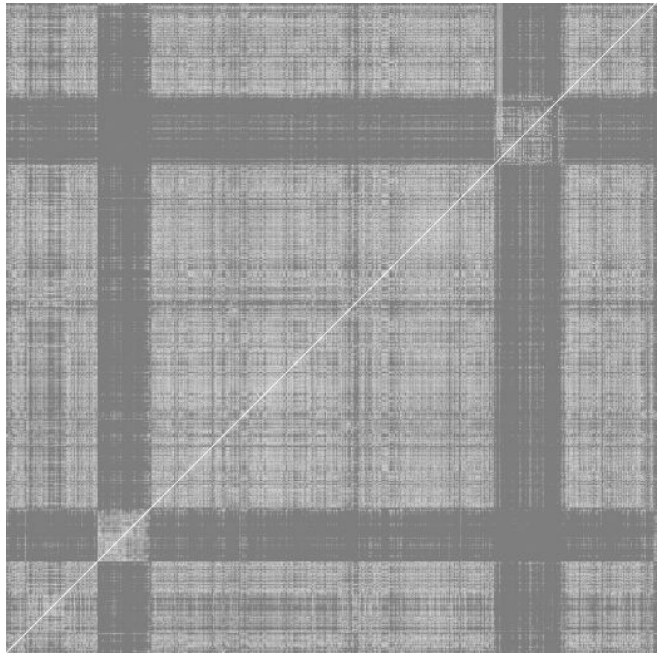
- Energy
- Pressure



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

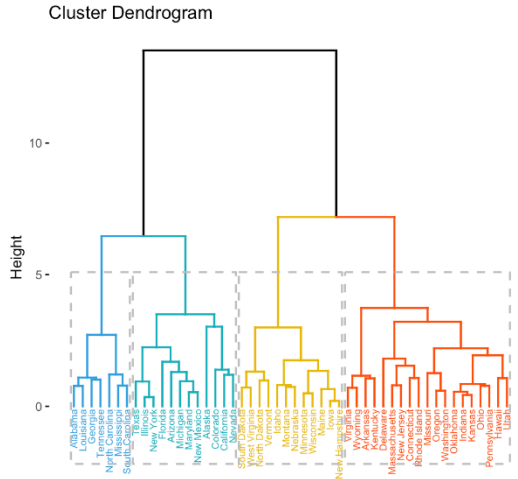
Calculation of ΔΔG

- 2D-RMSD compares all frames of one (or more) trajectory between them, resulting in a matrix.



2D-RMSD matrix

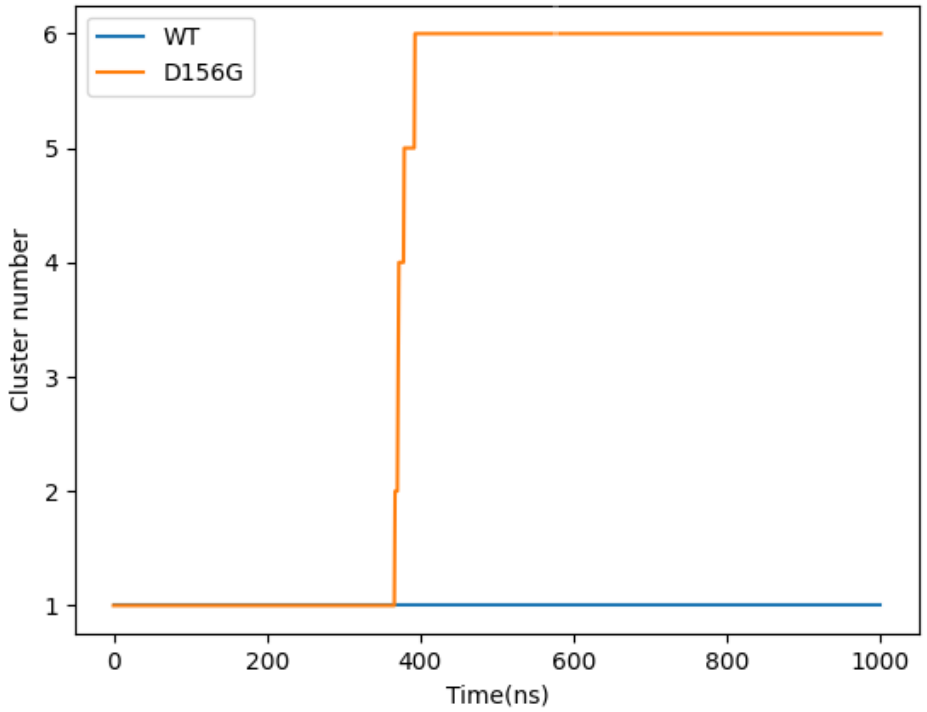
- Agglomerative clustering: Each object is a leaf on a tree, it joins branches
- We use 2D-RMSD as a distance matrix
- We plot which cluster the simulation is in at each moment

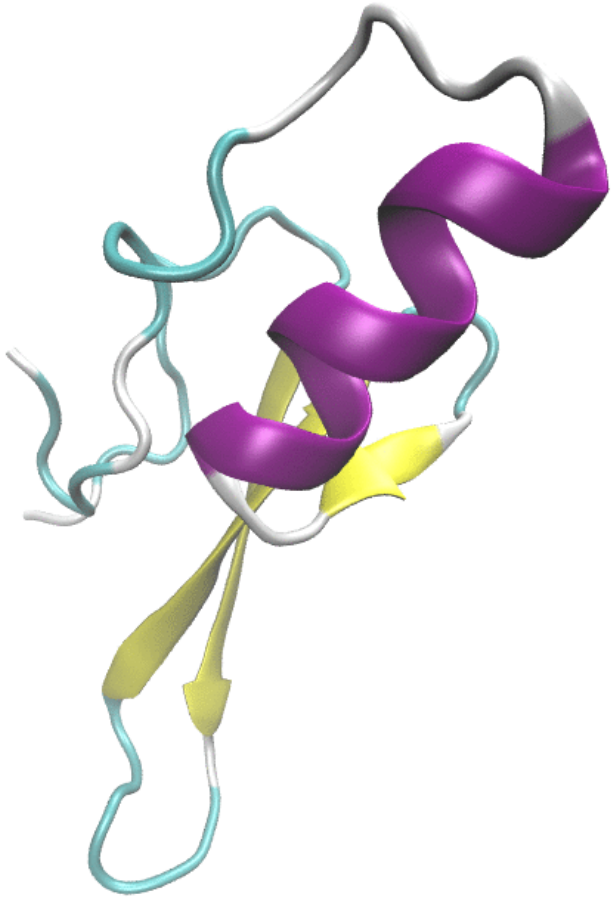


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

Calculation of ΔΔG

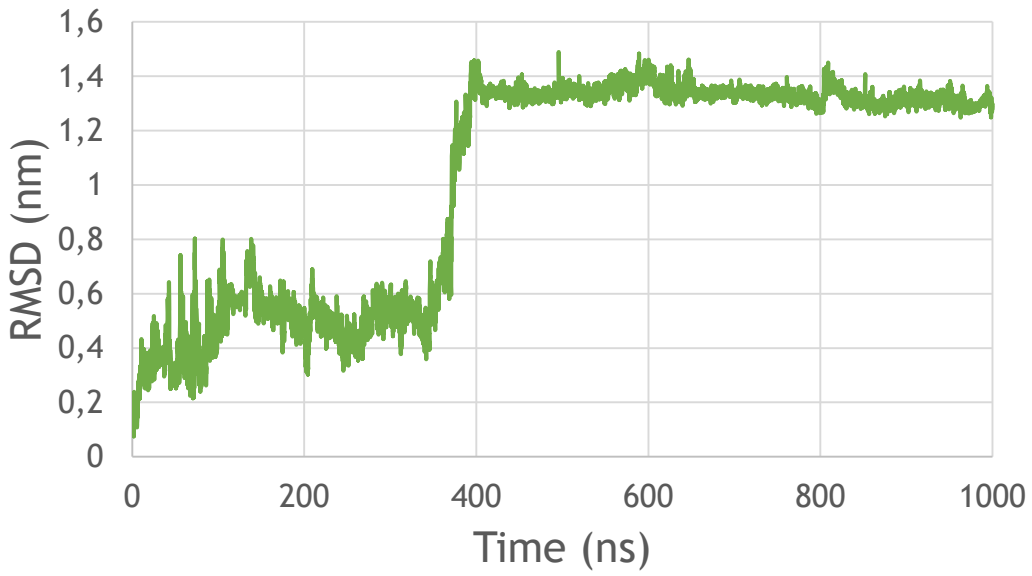




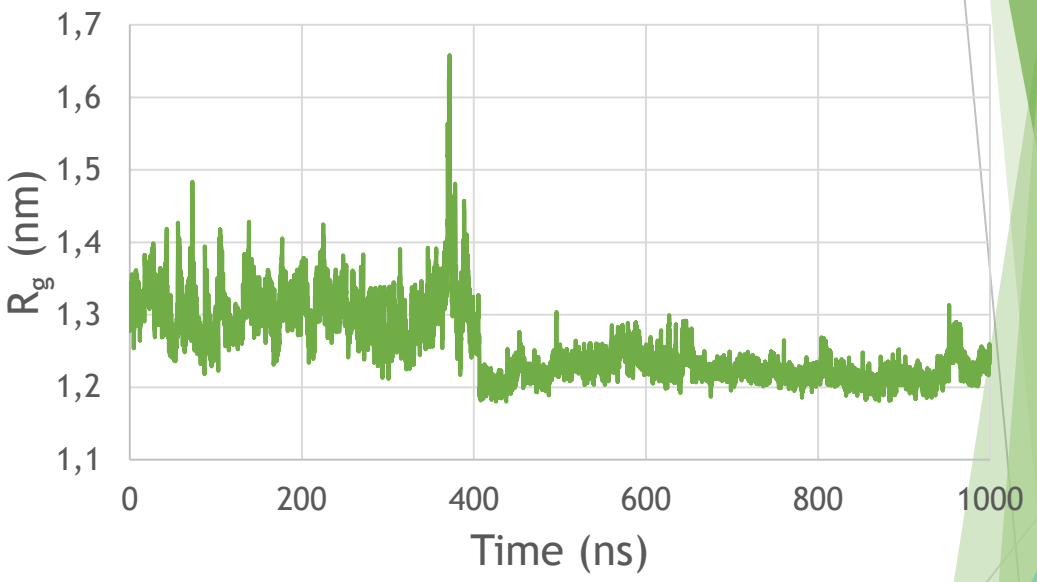
This project has received funding from the European Union's Horizon 2020 research and innovatio

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

RMSD



Radius of gyration



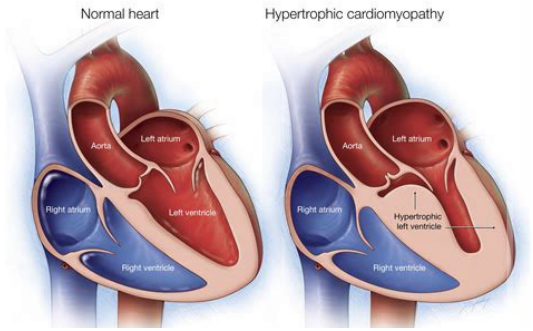
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure

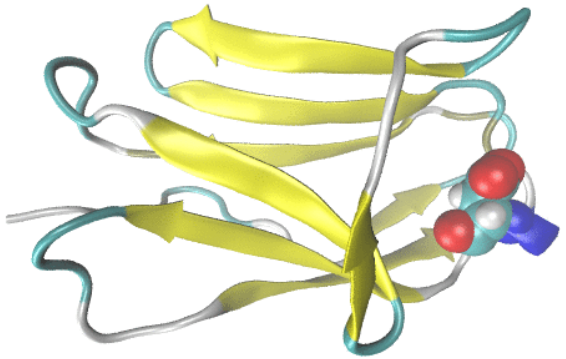
Calculation of ΔΔG

- MYBPC3 is related to hypertrophic cardiomyopathy
- Some variants could not be expressed
- Hypothesis: Too destabilizing expressed

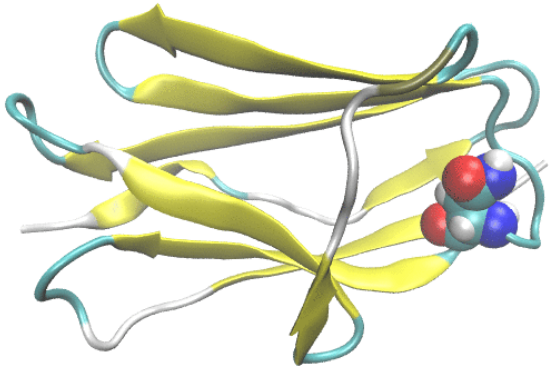
Suay-Corredera C, Pricolo MR, Herrero-Galán E, Velázquez-Carreras D, Sánchez-Ortiz D, García-Giustiniani D, Delgado J, Galano-Frutos JJ, García-Cebollada H, Vilches S, Domínguez F, Molina MS, Barriales-Villa R, Frisso G, Sancho J, Serrano L, García-Pavía P, Monserrat L, Alegre-Cebollada J. Protein haploinsufficiency drivers identify MYBPC3 variants that cause hypertrophic cardiomyopathy. J Biol Chem. 2021 297:100854.



wt



N610N

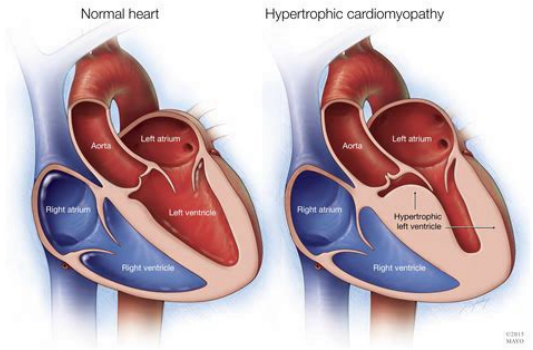


$$\Delta G_{unf} = a + b \times log\left(\frac{ln2}{-\frac{\ln(1-\chi)}{\tau} \times \frac{298}{T} \times e^{\Delta G^{\ddagger}\left(\frac{1}{RT}-\frac{1}{298T}\right)}}\right)$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

Calculation of ΔΔG

- MYBPC3 is related to hypertrophic cardiomyopathy
- Some variants could not be expressed
- Hypothesis: Too destabilizing to be expressed



Frutos JJ, García-Cebollada H, Vilches S, Domínguez F, Molina MS, Barriales-Villa R, Frisso G, Sancho J, Serrano L, García-Pavía P, Monserrat L, Alegre-Cebollada J. Protein haploinsufficiency drivers identify MYBPC3 variants that cause hypertrophic cardiomyopathy. J Biol Chem. 2021 297:100854.

Table 5.38. Summary of the relaxation Molecular Dynamics simulations in cMyBP-C (Partly adapted from [141])

Structure	Domain-Variant	Rep	WT Unfolding time (ns)	WT Unfolded fraction (time, ns)	WT ΔG_u (kcal/mol)	Variant Unfolding time (ns)	Variant Unfolded fraction (time, ns)	Variant ΔG_u (kcal/mol)	$\Delta\Delta G^a$ (kcal/mol)
PDB ID: 3CX2	C1-P187R	1	>1000	<0.01 (1000)	>3.92±0.33	>1000	<0.01 (1000)	>3.92±0.33	NA
		2	>1000			>1000			
		3	>1000			>1000			
PDB ID: 2MQ0	C3-V471E	1	>1000	<0.01 (1000)	>3.92±0.33	>1000	0.33 (1000)	2.74±0.38	>1.18±0.71
		2	>1000			830			
		3	>1000			>1000			
Homology model	C4-D610N	1	>1000	0.33 (1000)	2.74±0.38	520	>0.99 (520)	<1.74±0.47	>1.00±0.85
		2	>1000			430			
		3	750			180			
PDB ID: 1GXE	C5-P645L	1	250	>0.99 (250)	<1.51±0.50	0	>0.99 (10)	<0.47±0.60	~1.04±1.1
		2	50			10			
		3	50			0			
Homology model	C9-R1138H (His^+)	1	>1000	<0.01 (1000)	>3.92±0.33	320	>0.99 (860)	<1.90±0.46	>2.02±0.79
		2	>1000			320			
		3	>1000			860			
	C9-R1138H (His^0)	1	>1000	<0.01 (1000)	>3.92±0.33	>1000	0.66 (1000)	2.42±0.41	>1.5±0.74
		2	>1000			630			
		3	>1000			860			
	C10-G1206D	1	>1000	<0.01 (1000)	>3.92±0.33	400	>0.99 (400)	<1.66±0.48	>2.26±0.81
		2	>1000			10			
		3	>1000			250			
Homology model	C10-T1237P	1	>1000	<0.01 (1000)	>3.92±0.33	>1000	0.33 (1000)	2.74±0.38	>1.18±0.71
		2	>1000			>1000			
		3	>1000			620			
	C10-Y1251H	1	>1000	0.33 (1000)	2.74±0.38	100	0.66 (1000)	2.42±0.41	0.32 ±0.81
		2	570			630			
		3	>1000			>1000			
	C10-L1268P	1	>1000	0.33 (1000)	2.74±0.38	190	>0.99 (200)	<1.44±0.50	>1.30±0.88
		2	540			200			
		3	>1000			30			

[a] Positive ΔΔG values are indicative of mutant-induced domain destabilization at 298 K.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI Molecular-Scale Biophysics Research Infrastructure

DAY 3 (5 July, 2023) 15:00/17:00 Molecular Dynamics simulations and stability calculation

- MD simulations
- Observing unfolding: relaxation MD
- Calculation of ΔH and ΔCp
- Calculation of ΔG
- Biophysical insight
- Calculation of ΔΔG

- More questions?

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004806

MOSBRI
Molecular-Scale Biophysics Research Infrastructure