



Protein Biophysics, Biochemistry
& Bioinformatics Lab

From protein folding to function to evolution to design: Computational Research at PB³ Lab

César A. Ramírez-Sarmiento

Protein Biophysics, Biochemistry and Bioinformatics Lab

Institute for Biological and Medical Engineering
Pontificia Universidad Católica de Chile

Millennium Institute for Integrative Biology



Our research group and funding sources have significantly grown since the formation of our laboratory in 2016



César Ramírez-Sarmiento

Principal Investigator



Protein Biophysics,
Biochemistry
& Bioinformatics Lab



Maira Rivera
Postdoctoral
Fellow



Pablo Galaz-
Davison
PhD student



Nicolás
Núñez
PhD student



Felipe Gatica
Undergraduate



Amparo Núñez
Research
Assistant



Paula Blazquez
PhD student



Ignacio
Retamal-Farfán
Master student



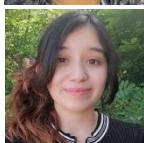
Aransa Griñen
Undergraduate



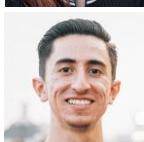
Javiera Reyes
PhD student



J. Alejandro
Molina
PhD student



Cyndi Tabilo
Master student



Felipe
Engelberger
Undergraduate



FONDECYT:

1201684 (CRS)
3190731 (MR)

Internacional Cooperation:

ANID-FAPESP 2019/13259-9
(Universidade de Sao Paulo)
ANID-CONCYTEC covbio0012
(U. Peruana Cayetano Heredia)



Millennium Institute for Integrative Biology
ICN17_022



<https://pb3.sitios.ing.uc.cl>

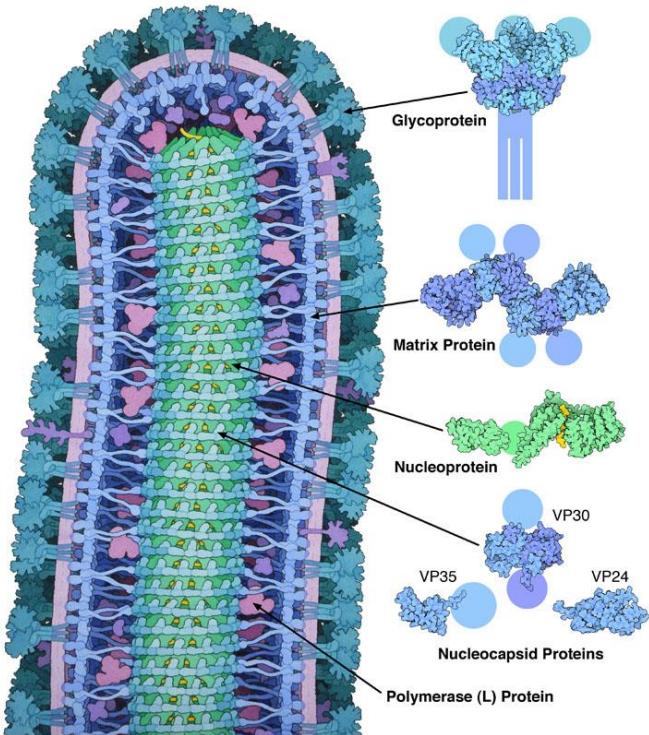


@PB3Lab

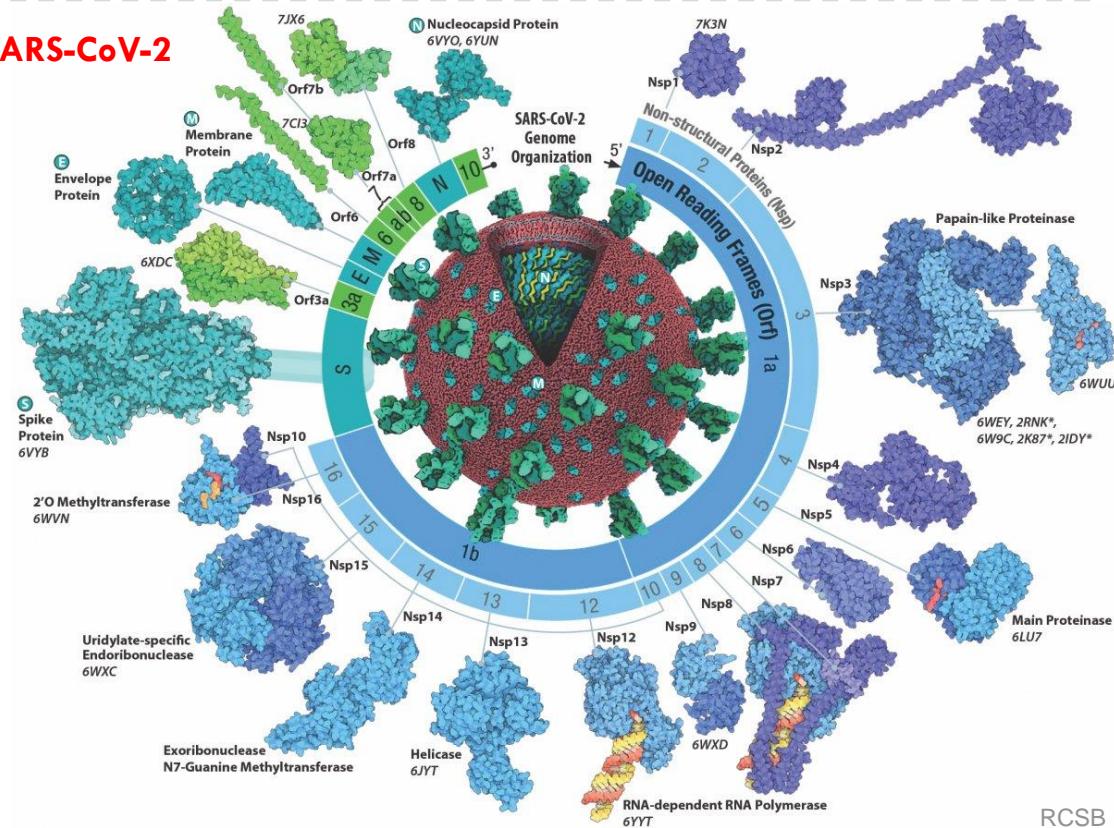
pb3@ing.puc.cl

Proteins are the workhorse molecules of life, physically taking part in essentially every biological function, structure and activity

Ebola virus

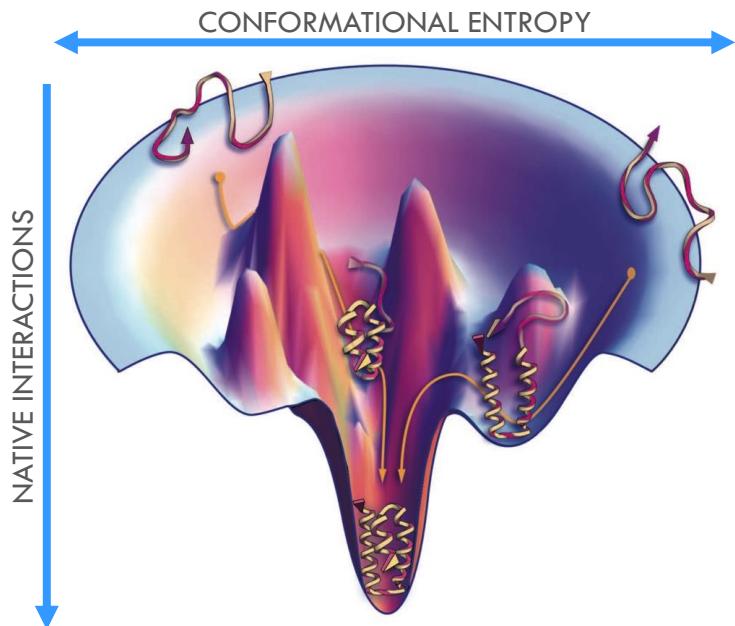


SARS-CoV-2

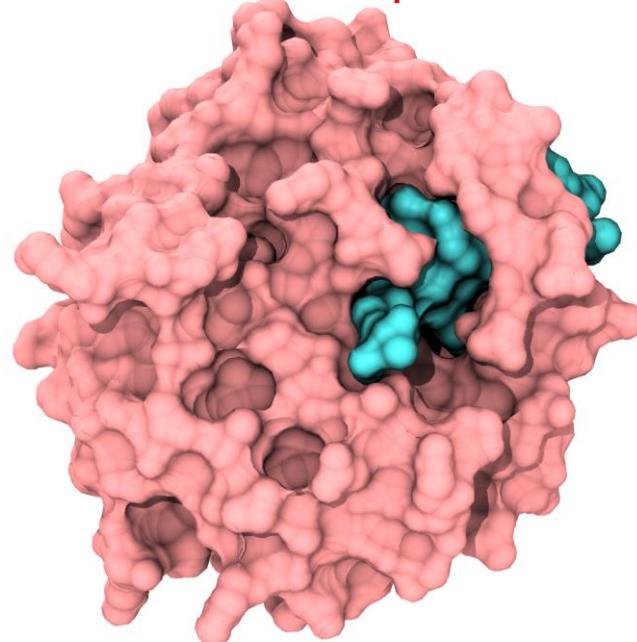


Our research is focused on protein folding and function, which we believe is fundamental for a comprehensive understanding of cellular response and evolution

Protein folding as a biophysical problem



Protein function as the consequence of folding

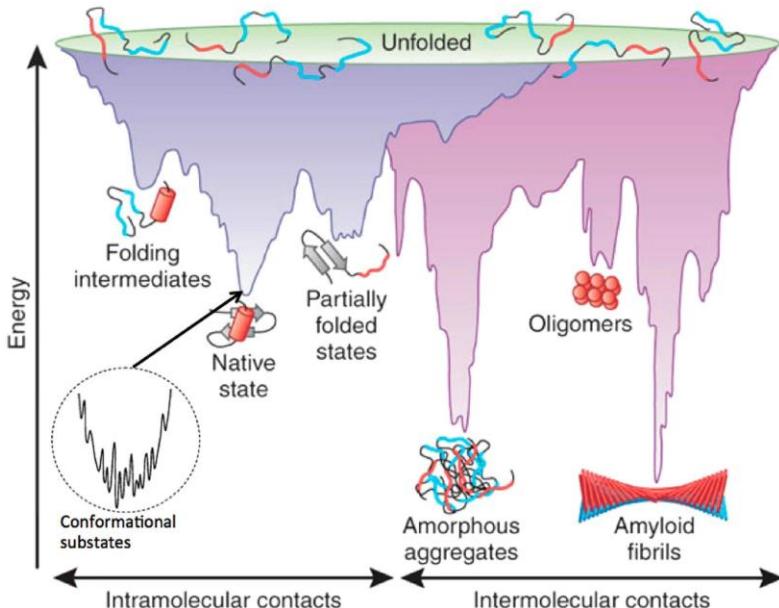


These themes are mostly considered as fundamental research, however...

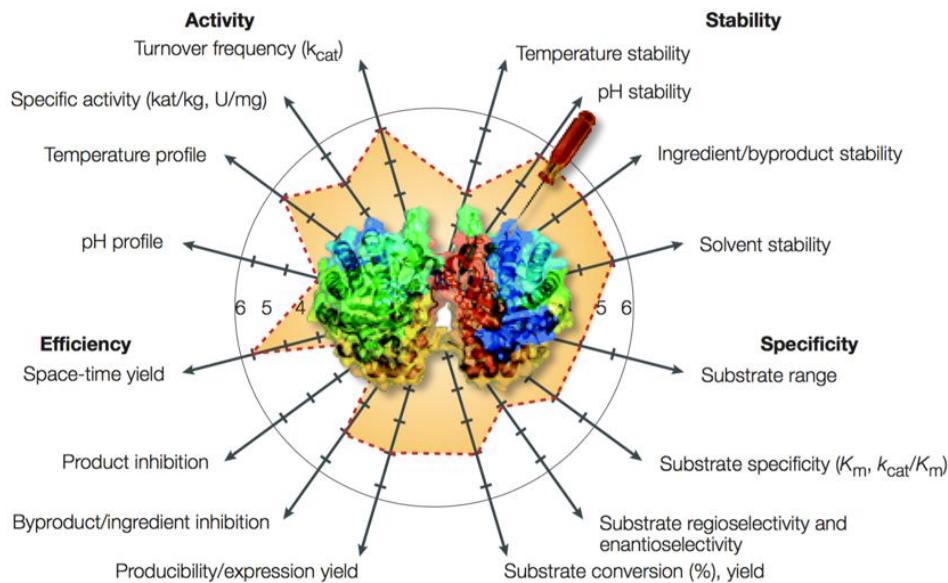
Bryngelson JD & Wolynes PG (1987) | PNAS **84**, 7524–7528

Our research is focused on protein folding and function, which we believe is fundamental for a comprehensive understanding of cellular response and evolution

Protein folding as a biomedical problem



Protein function as a biotechnological tool



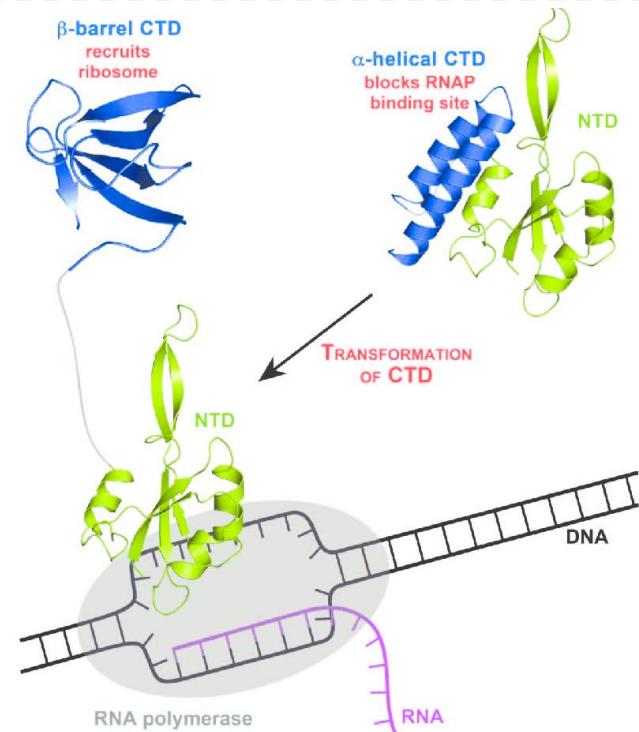
...we know that folding can directly be involved in several diseases and that enzymes with novel activities can be discovered and/or engineered for specific biotechnological needs.

Our research focuses on exploring the landscapes of protein folding and function by integrating computational and experimental strategies



Landscapes of protein function

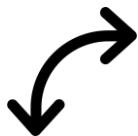
Hallmarks of the PET-degrading activity of PETase



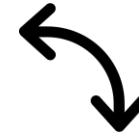
Landscapes of protein folding
Fold-switch of the transcription factor RfaH

We answer some of these questions through a myriad computational methods for the analysis of protein sequence, structure, function, evolution and dynamics

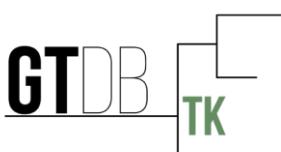
MOLECULAR DOCKING (FUNCTION)



Rosetta
Commons



PROTEIN SEQUENCE ANALYSIS



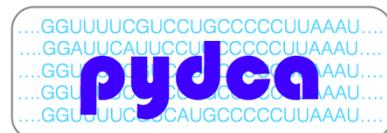
PROTEIN STRUCTURE MODELLING



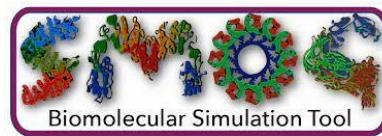
Rosetta
Commons



COEVOLUTIONARY ANALYSIS



MOLECULAR DYNAMICS



AMBER MD

GROMACS
FAST. FLEXIBLE. FREE.



DISCOVERING & CHARACTERIZING CURRENT AND NOVEL
PET HYDROLASES

PET is one of the most used plastics in packaging, but its chemical and mechanical recycling faces several issues.



PET contribution to
total plastic
production

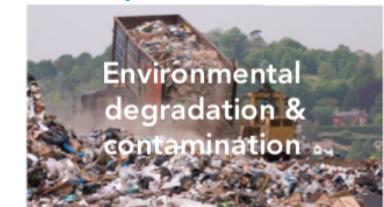


PET amounts for

20%

of the total of plastics used for
packaging

~3% ductility after 3 cycles
Afterwards, used only in textiles



Costly and may also be
highly toxic

Garcia JM (2016) | Chem 1, 813-815

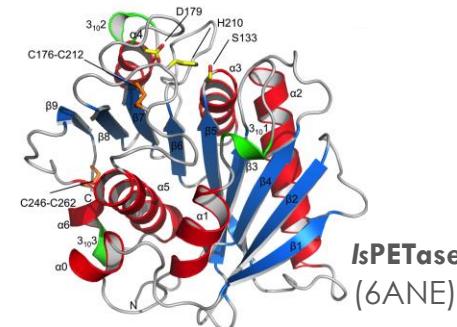
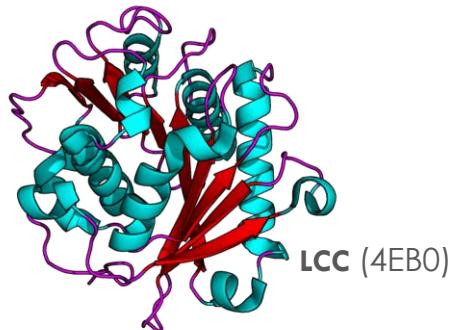
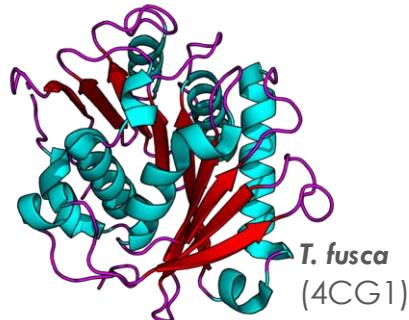
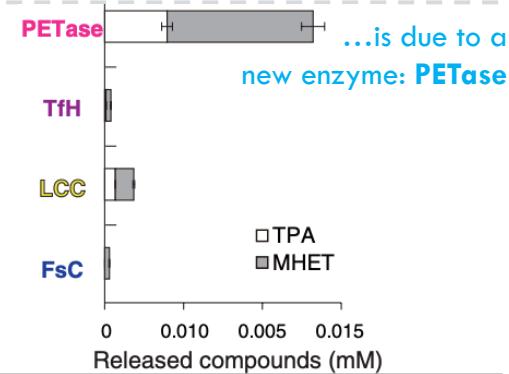
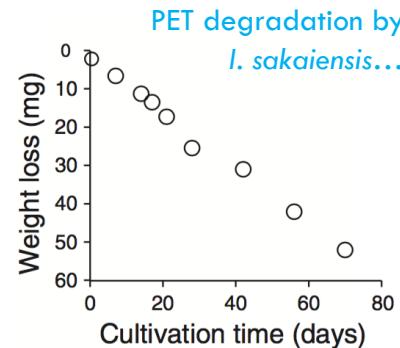
Geyer R et al (2017) | Sci Adv 13, e1700782

Jambeck JR et al (2015) | Science 347, 768-771

Several PET hydrolases have been discovered, but only a few are active at room temperature and the molecular mechanisms enabling these adaptations are unknown

250

Samples collected from a plastic-rich landfill in Japan

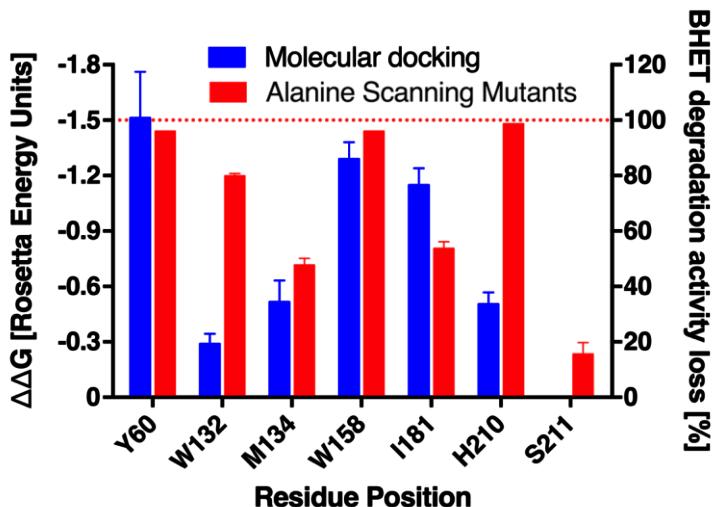
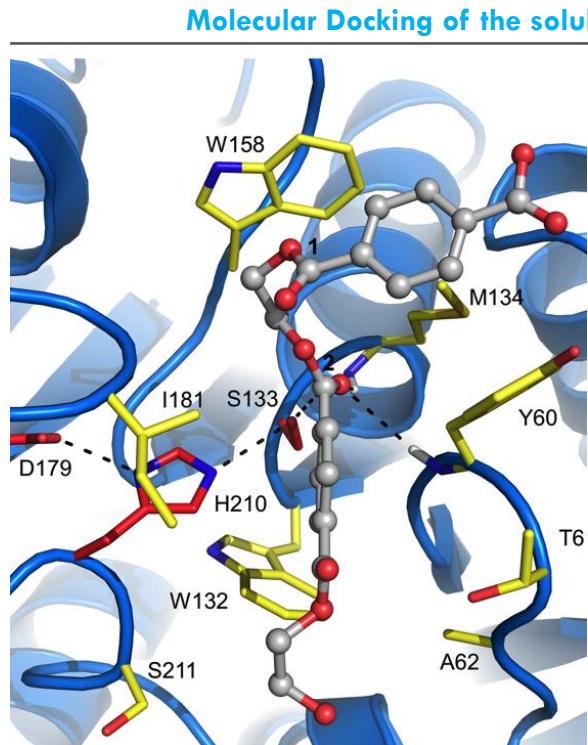
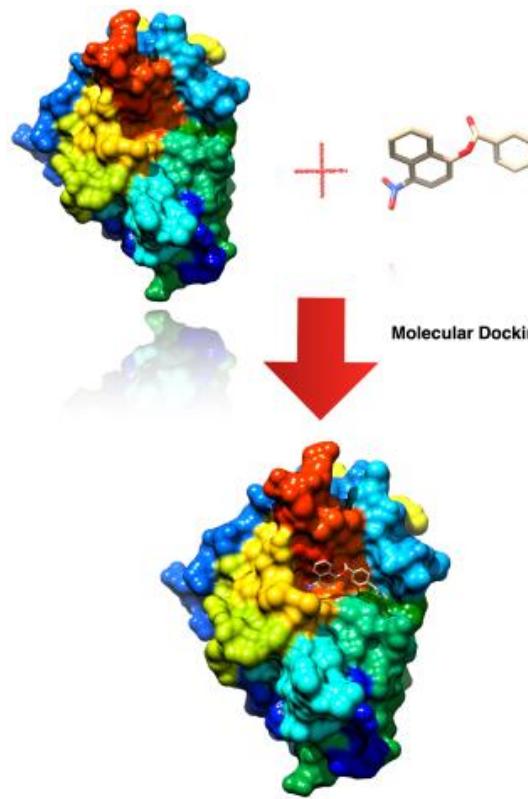


Are there any other cold-active PET hydrolases? What are the hallmarks of their temperature adaptations?

Fecker T, Galaz-Davison P et al (2018) | *Biophys J* **114**, 1302–1312

Yoshida S et al (2016) | *Science* **351**, 1196–1199

By solving the structure of *IsPETase* and combining it with molecular docking of a substrate analog, we accurately determined active site residues involved in its activity



Significant agreement between per-residue binding energies and **BHET degradation of active site mutants.**

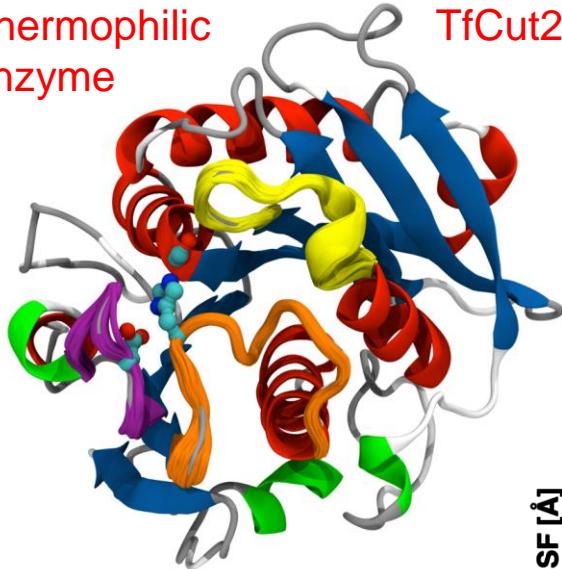
Joo S et al (2018) | *Nat Comm* **9**, 382

Fecker T, Galaz-Davison P et al (2018) | *Biophys J* **114**, 1302–1312

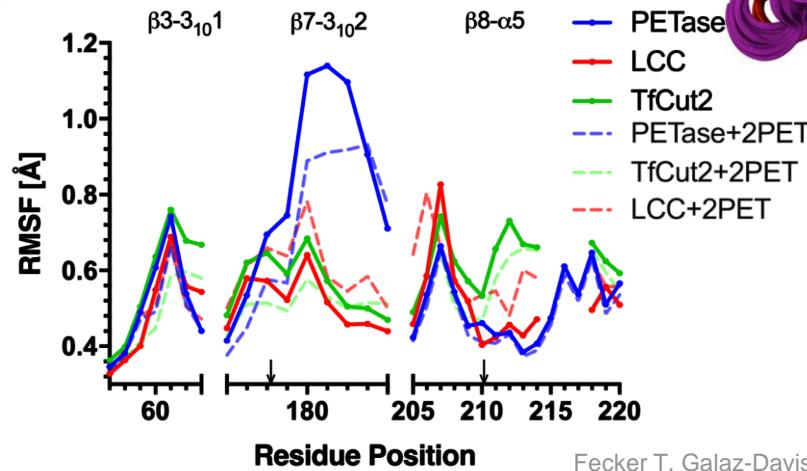
Molecular dynamics simulations at different temperatures demonstrate that the higher flexibility of the active site of PETase is a hallmark of its activity at 40 °C

Thermophilic
enzyme

TfCut2

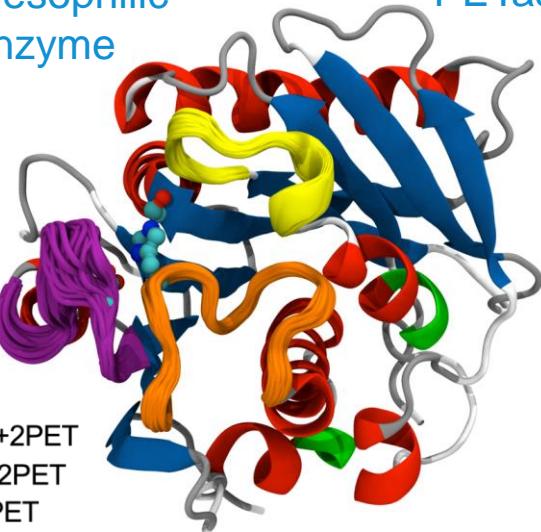


MD simulations of
4x100 ns
in explicit solvent at
298 K and 1 atm



Mesophilic
enzyme

PETase

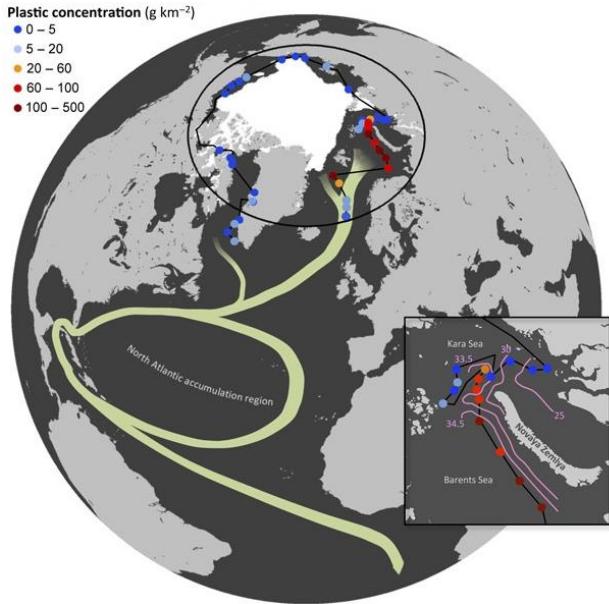


Reasoning that *IsPETase* is not the only PET hydrolase active at room temperature, we aimed to discover novel Arctic and Antarctic PET hydrolases such as Mors1

Plastics have reached

Polar regions

due to the ocean currents...



Searching Antarctic PETase homologs in UNIPROT

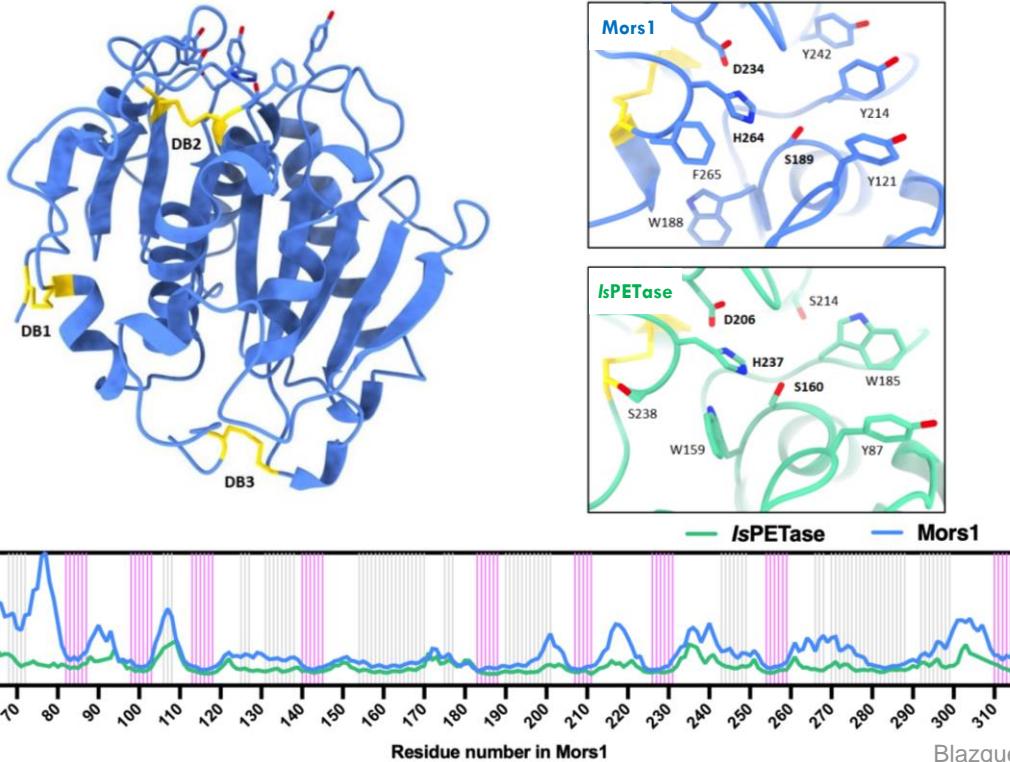
	TfCut2	TT	α_1	β_1	TTT	β_2	TT	β_3	TT	η_1		
TfCut2	1	ANPVERGPNPDTALLEARS	GPF	SVSEE	NVSRL	SASGF	GCGCTIYYPREN...	NTY	GAVAI	SPGYTGTEA		
LCC	36	SNPYORGPNPTRSALTAD	GPF	SVATY	TVSRSL	SVSGFGGGVIVYYPTGTS..	LT	FGGIA	MSPGYTADAS			
IsPETase	29	TNPYVARGPNPATAASLEASAGPFTVRSFTVSRP..	GPFTVRSFTVSRP..	..SGYGA	CTVYYPTNAG..	GT	VGAIAIVPGY	TARQS				
OaCut	45	DCEFTRGPNPTPSSLEAST	GPYSVATR	SVA..	SVSGFGGGCTLHYPTNTT..	GT	MGAIAAVVPGF	LLQES				
Mors1	59	DCIADSKITAVALS	GPFSIRT	KRISQSAK	GFGGCTIHYP	PTNASGC	LLGAI	AVVPGF	LLQES			
	TfCut2	e	α_2	β_4	TT	α_3	TTT	α_4	β_5			
TfCut2	66	SIAWLGERIASHG	FVVITID	DTITLDQPD	SRAEQLNA	AALNHM...	INRASS	STVRSRIDSSRLA	VMGHS			
LCC	101	SIAWLGRRLASHG	FVVVLV	NTNSRFD	YPDPSRASQ	LAQNLNYL...	RTSSP	SAVRARLDANR	LAVALAGHS			
IsPETase	93	SIKWWCPRLASHG	FVVITID	NTNSTLDQ	PPSSRSQQMA	RQVASLNLTSS	SP	YGVDTGM	VGWWS			
OaCut	110	SIDFWGPKLASHG	FVVVITI	ISANGFD	DPASRATOL	GRALDYV	INQNSGSNSPI	SGMVDTTR	GLGVGWS			
Mors1	127	SIKWWCPRLASWGF	VVITIN	TNSIYDDPD	SRAAQLNA	DNMIA....	DDTVGSMIDPKR	RLGAIGWS				
	TfCut2	eeeeeee	α_5	β_6	TT	α_6	TT	α_7	β_8	TT		
TfCut2	131	MGGGCSLRLA	SQRPD	KAAPL	TPWHLN.	KNWSS	VRVPTLI	IGADLDTI	TPVLT	HARPFYINS		
LCC	166	MGGGCTLR	IAEQNPSLK	AAVPL	TPWHTD.	KTFN.	TSVPVLIV	GAEADTV	VAPVVSQHAI	PFYQNL		
IsPETase	161	MGGGCSLISA	ANNPSLK	AAAPQAP	WDSS.	TNFSS	VTVP	LI	FACE	ENDSIA		
OaCut	178	MGGGGCA	LOASG.	DRLSAAI	PIAPWNQGGNR	QDQIETPT	LV	ACEN	DVVASVN	SPYFNIR		
Mors1	190	MGGGCAL	KLATERST	VRAIMPL	APYHD..	KSYGEV	KTP	TLV	DIETKKY	ANAFYKNAI		
	TfCut2	eeeeeee	β_8	TT	η_2	eeeeeee	α_7	eeeeeee	η_3	α_8	β_9	TT
TfCut2	198	KAYDEL	DGAT	THFAPN...	IPNKT	IIGKYSV	AWLKR	FVDNDTRY	TQFLC	PGPRDGLF	GEVEEYR.	
LCC	232	KVXVEL	LDNA	SHFSCAN	...SNNA	AIISVYT	ISWMKLW	VWDNT	TRYRQFLC	CNVNDPALS.	DFRTNN.	
IsPETase	227	KQFLE	INGGSH	SCAN	SGNSNQAL	LGKKGVA	AMKRFMD	NDTRY	STFACE	ENPNSTRVSD	FRTAN.	
OaCut	245	KAYDEL	INGGSH	FCANDGGSIGGL	LGKKGVA	WMKRFIDND	TRYDAFLC	GPDH.	AANRSVSEYR.	DT	CNY	
Mors1	254	KMKVEV	NNNGSH	FCPSYR.	FNEIL	LLSKPGI	AMQRYIN	NDTRE	DKFLC	ANENYSKS	SPRISAYDYKD	

Cózar A et al (2017) | Sci Adv 3(4), e1600582

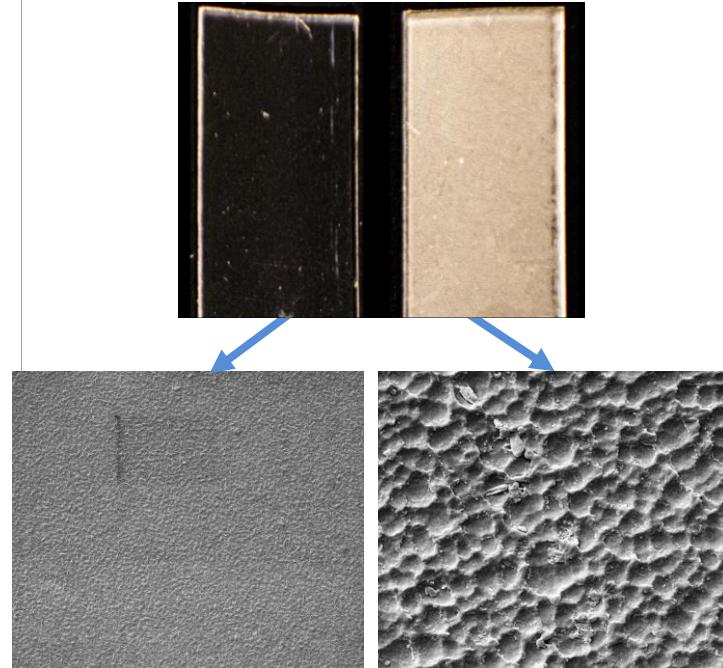
Blazquez-Sánchez P et al (2021) | Appl Environ Microbiol 88(1), AEM0184221

Reasoning that *IsPETase* is not the only PET hydrolase active at room temperature, we aimed to discover novel Arctic and Antarctic PET hydrolases such as Mors1

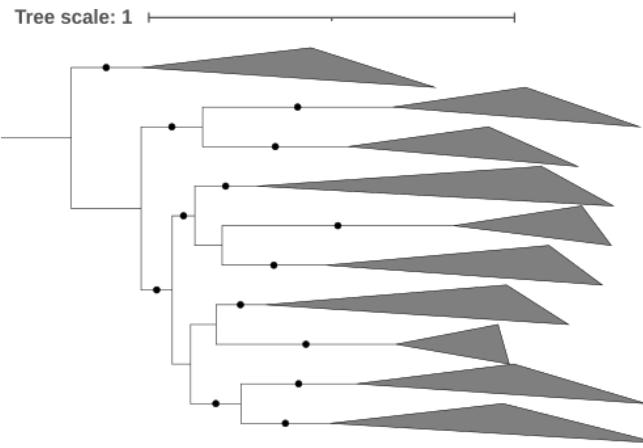
Comparative modelling and MD simulations of Mors1



PET film degradation assays



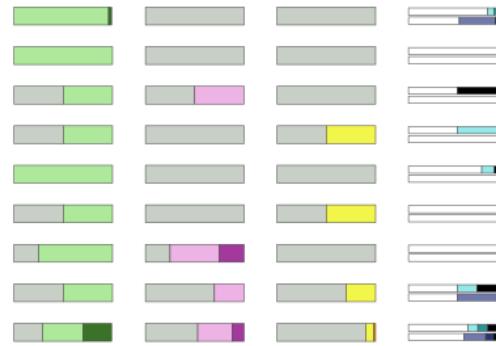
Based on these findings, we successfully performed metagenomic searches of PET hydrolysis and consumption metabolic pathways in polar marine organisms



PETase
none
PETase
>1 PETase

MHETase
none
MHETase
>1 MHETase

TPA degradation pathway
None
At least 1 protein
Full pathway



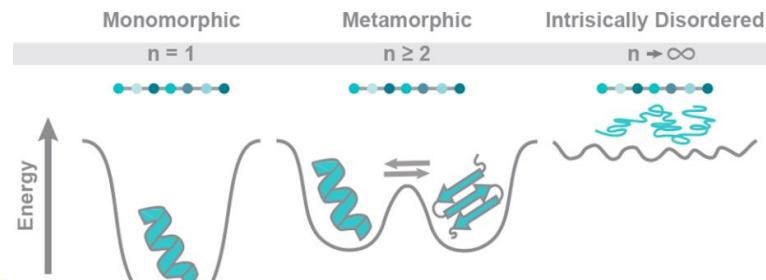
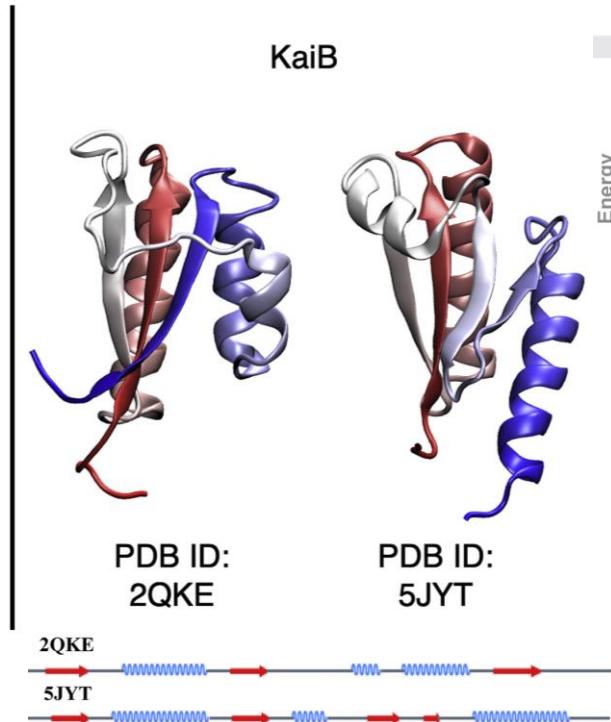
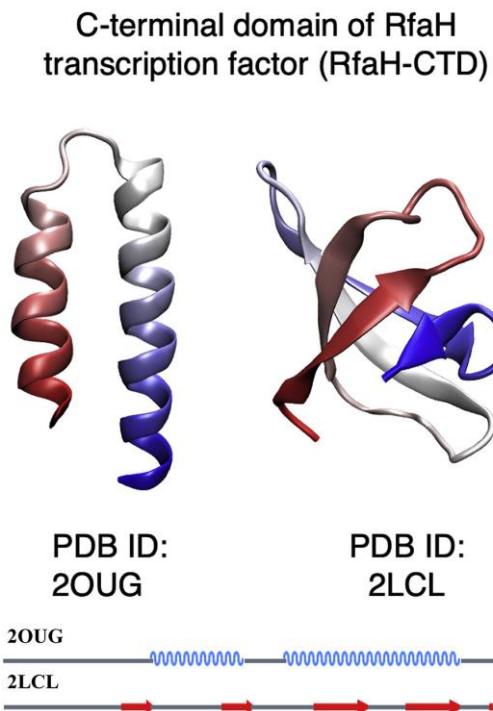
Antarctic distribution
None
Surface
Deep
Surface and deep

Arctic Distribution
None
Surface
Deep
Surface and deep

WANDERING THE REFOLDING LANDSCAPES OF
METAMORPHIC PROTEINS



Metamorphic proteins challenge the sequence-structure paradigms of protein folding by switching between two different native states with different functions



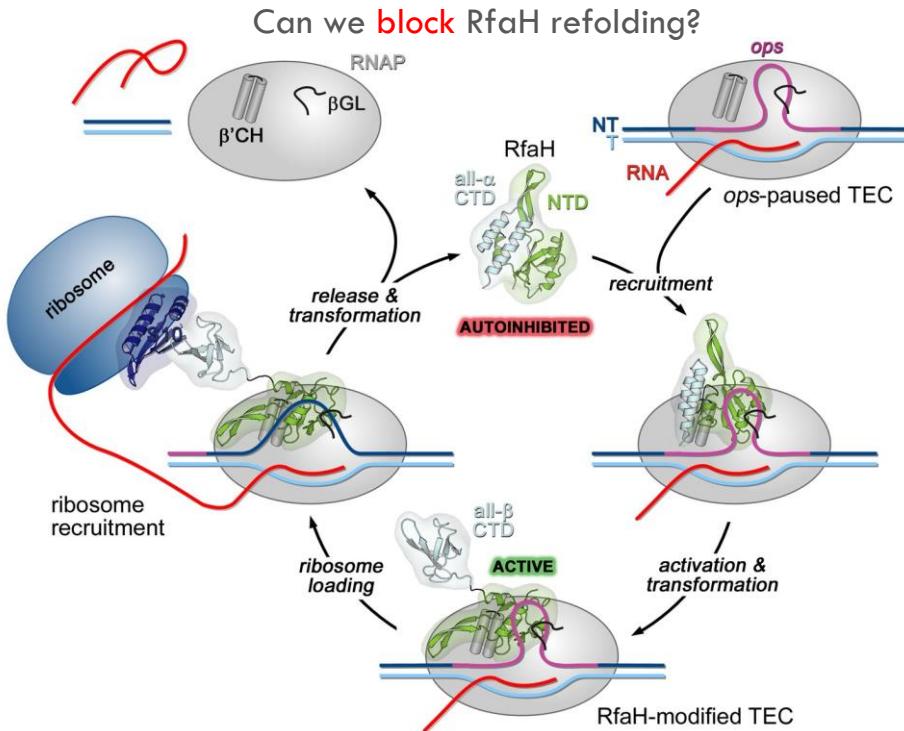
Bioinformatic estimations suggest that up to

4%

of the PDB are metamorphic candidates

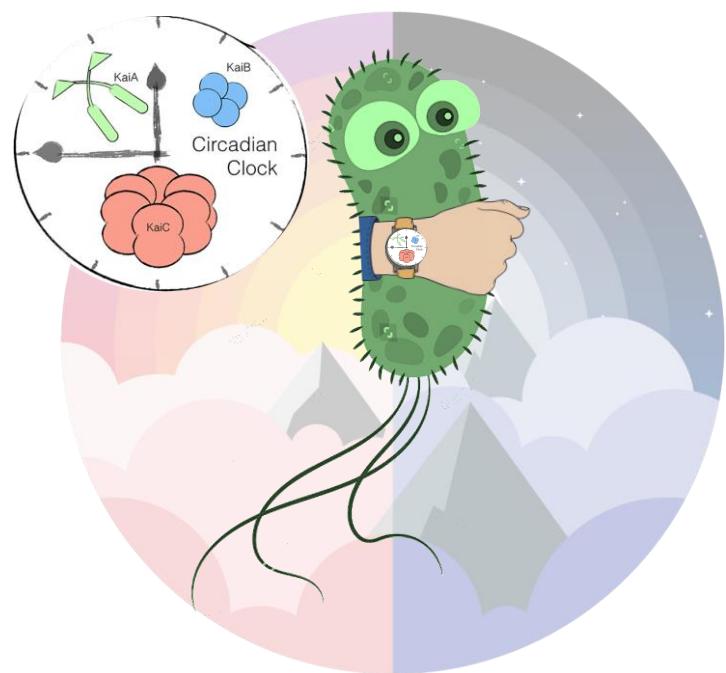
Dishman AF & Volkman BF (2018) | ACS Chem Biol 13, 1438–1446
Porter LL & Looger LL (2018) | PNAS 115, 5968–5973
Kulkarni P et al (2018) | Prot Sci 27, 1557–1567
Chen N et al (2020) | Biophys J 119, 1380–1390

We employ MD simulations, bioinformatics and experimental strategies to unveil the fold-switch mechanisms of RfaH & KaiB and their emergence during protein evolution



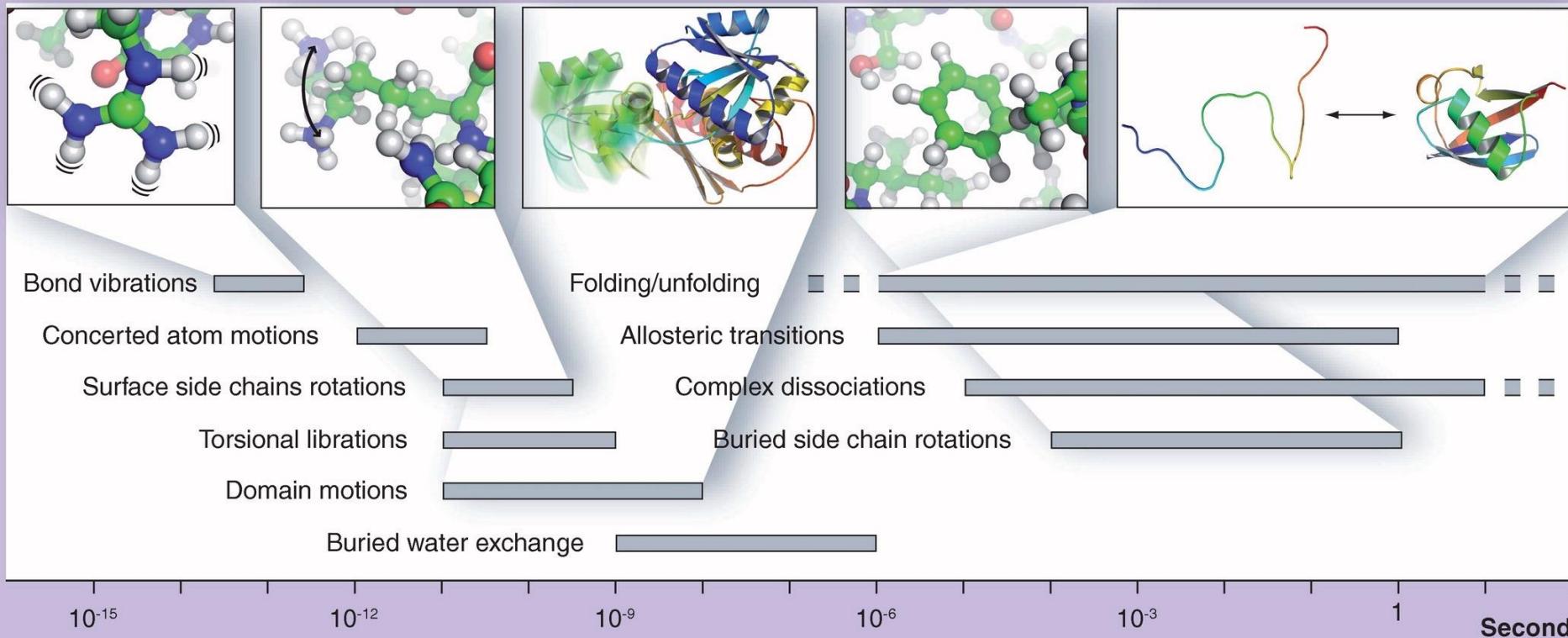
Can we design our own metamorphic proteins?

Can we alter the periodicity of the Kai clock?



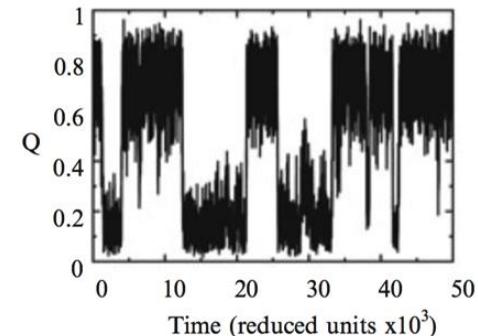
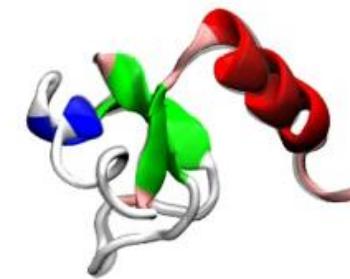
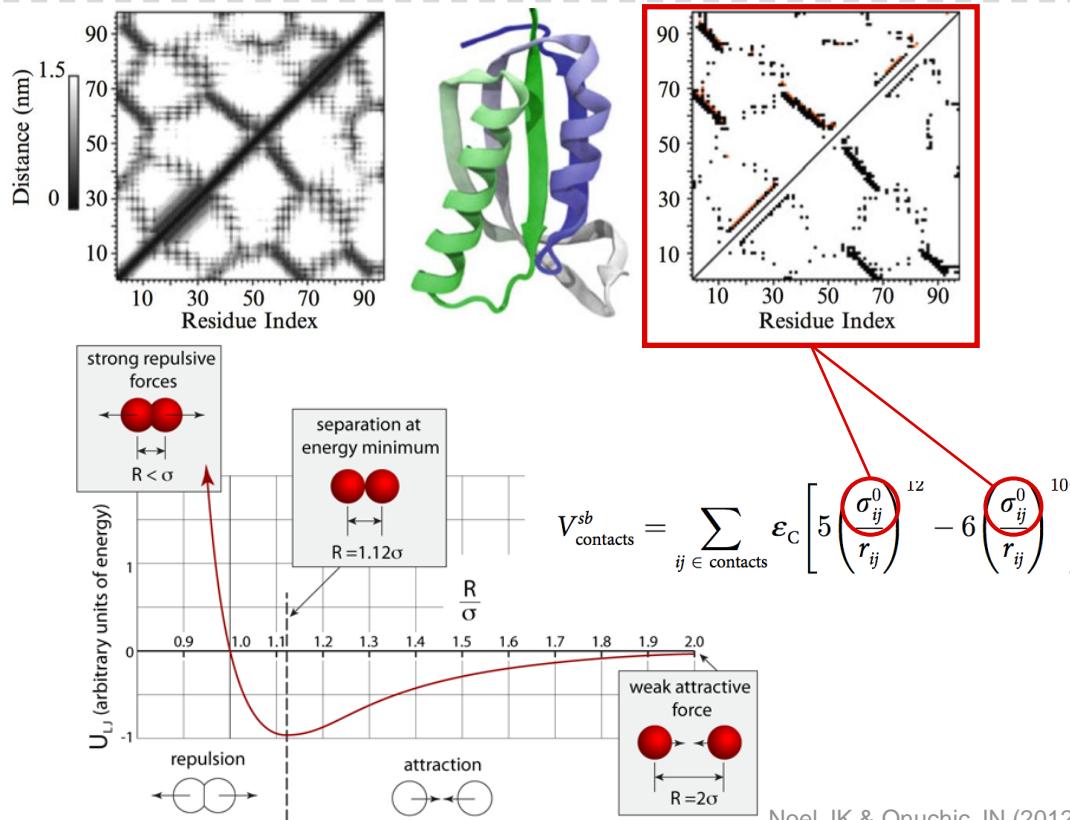
Zuber PK et al (2018) | eLife 7, e36349
Tseng R et al (2017) | Science 355, 1174–1180

Unfortunately, conventional MD simulations do not enable sampling of large-scale transitions within the timescales of protein folding. What can we do?



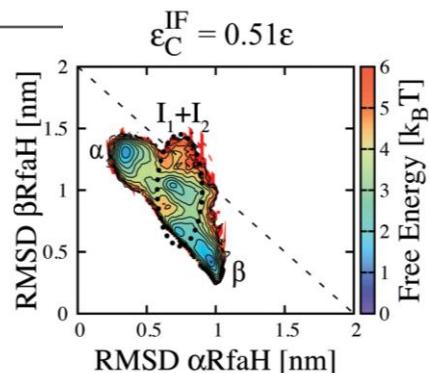
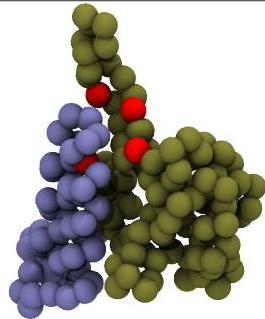
Teilum K et al (2009) | Cell Mol Life Sci 66, 2231-2247

The topology of the native state of a given protein can be explicitly used in MD simulations to efficiently explore protein folding with reduced computational costs

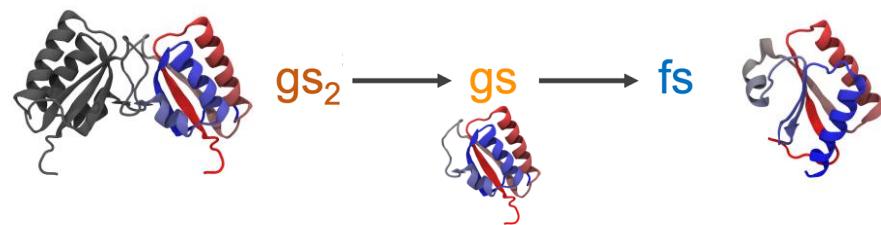
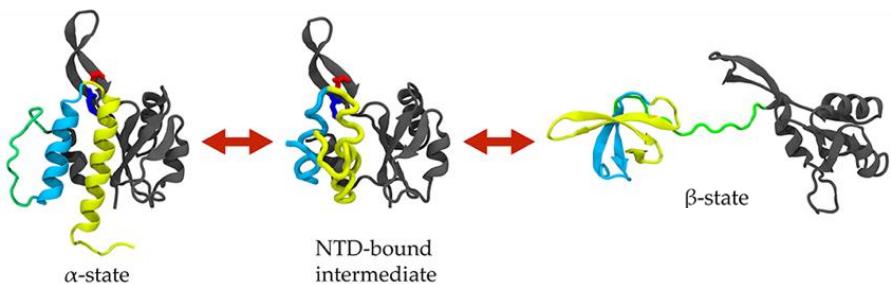
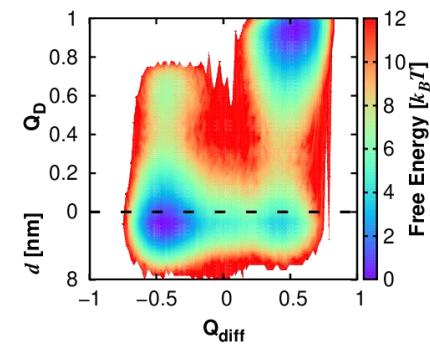
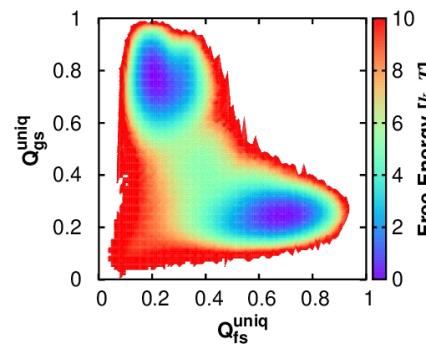


These simulation models allow us to explore the refolding pathways of these proteins and identify kinetic traps that could be the target of protein design or drug discovery

Fold-switch landscape of RfaH



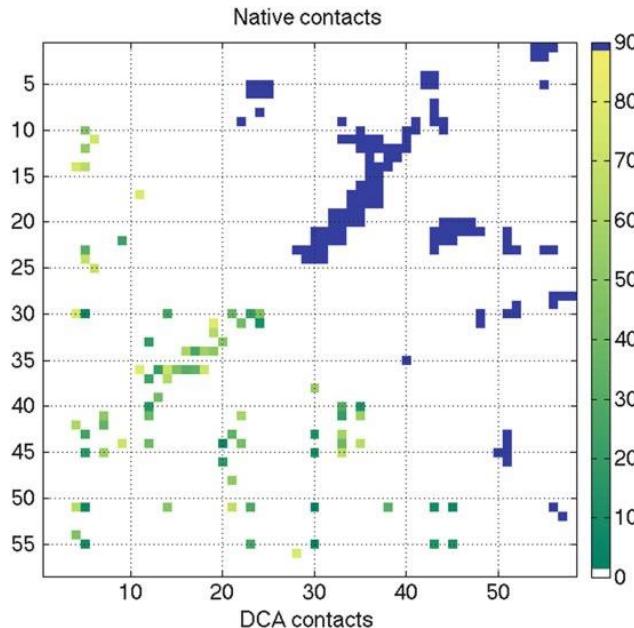
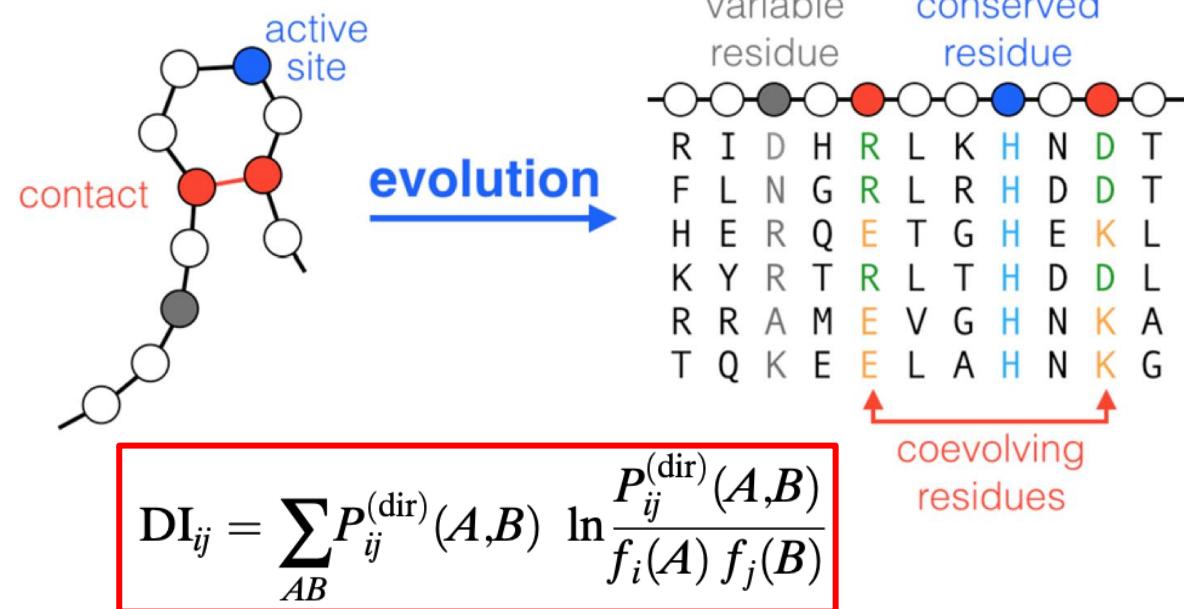
Fold-switch landscape of KaiB



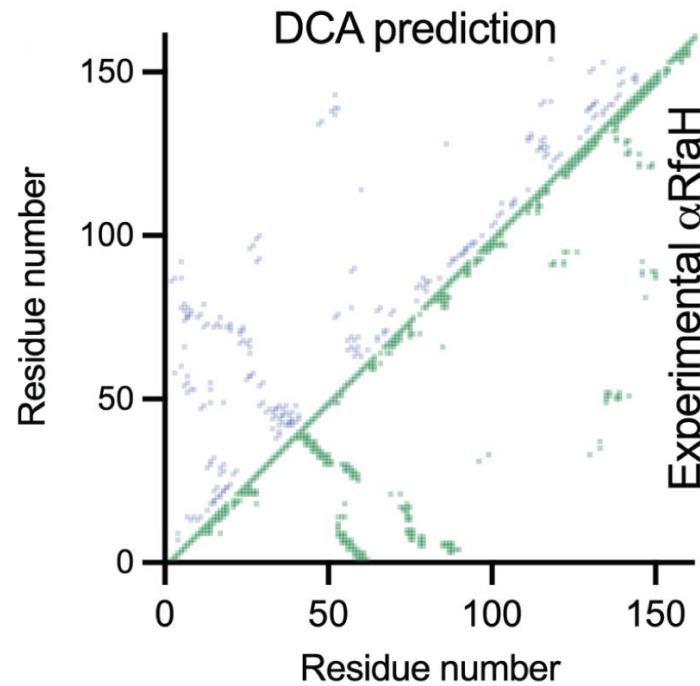
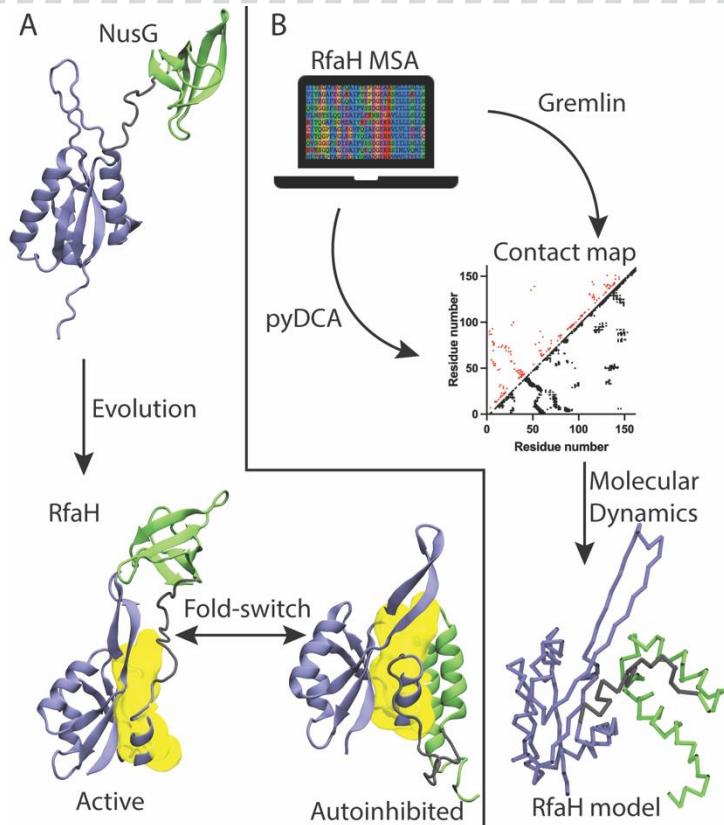
Ramirez-Sarmiento CA et al (2015) | PLOS Comput Biol 11, e1004379

Rivera M et al (2022) | Biophys J, resubmitted

We can then determine residue pairs in direct physical contact via statistical analysis of coevolving residues in MSAs using maximum entropy principles and Potts models



Using genomic and metagenomic sequence databases and secondary structure predictions, we were able to maximize the prediction of native contacts for RfaH



SOUNDS EXCITING FOR YOU?

WANNA LEARN MORE?

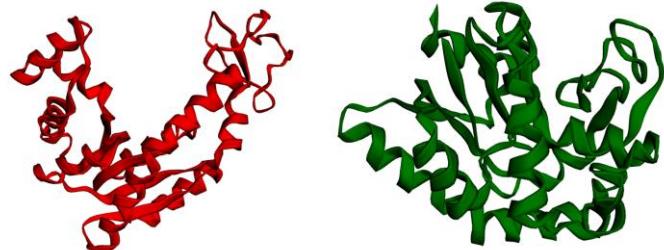
We have also developed open-source cloud-computing tutorials on molecular modelling and simulation of biomolecules for execution on Google Colab

Lab.10 / IBM3202 – Conformational changes using Structure-Based Models

Theoretical aspects

The **energy landscape theory** and the **principle of minimum frustration** in protein folding have provided the theoretical basis for the generation of simplified models to simulate the pathways of protein folding of different proteins. Noteworthy, recent work has demonstrated their utility for simulating other functionally relevant phenomena, such as **protein misfolding and conformational changes** associated to biological function. Most of these applications have been generated through savvy and careful combinations of the native bonded and non-bonded terms from **two or more structures deposited in the PDB** in two or more conformational states (i.e. open and closed conformations, alpha and beta states, etc).

```
A
import py3Dmol
#First we assign the py3Dmol.view as a two-panel viewer
view=py3Dmol.view(800,400,viewergrid=(1,2))
#Here we set the background color as white
view.setBackgroundColor('white')
#The following lines are used to add the addModel class
#to read the open state structure
view.addModel(open('4ake_clean.txt', 'r').read(),'pdb',viewer=(0,0))
#Here we set the visualization style and color
view.setStyle({'chain':'A'},{'cartoon': {'color':'red'}},viewer=(0,0))
#Now we do the same for the closed state structure
view.addModel(open('1ake_clean.txt', 'r').read(),'pdb',viewer=(0,1))
#Here we set the visualization style and color
view.setStyle({'chain':'A'},{'cartoon': {'color':'green'}},viewer=(0,1))
#Here we center the molecule for its visualization
view.zoomTo()
#And finally visualize the structures using the command below
view.show()
```



1/24/2022

Table 1. Overview of the Content and Required Software to Execute the Cloud-Based Tutorials

Tutorial	Description	Software	Based on Previous Tutorial?
Lab.00	Installing software on Google Colab for IBM3202 tutorials	pyRosetta, ^a GROMACS, ^b SBM-enhanced GROMACS ^c	No
Lab.01	Warm-up on Colab and brief review of biomolecular databases	Biopython, ^d py3Dmol, ^e NGL Viewer ^f	No
Lab.02	Visualizing and comparing molecular structures in Google Colab using py3Dmol	Biopython, ^d py3Dmol, ^e NGL Viewer ^f	No
Lab.03	Phylogenetic analysis using biopython and RAxML	Biopython, ^d miniconda, ^g MAFFT, ^h MSAViewer, ⁱ ModelTest-ng, ^j RAxML-NG ^k	No
Lab.04	Comparative modeling using MODELLER	Biopython, ^d py3Dmol, ^e MSAViewer, ⁱ MODELLER ^l	Yes
Lab.05	Membrane protein modeling using PyRosetta	pyRosetta, ^a py3Dmol ^e	Yes
Lab.06	Molecular docking on Autodock	Biopython, ^d py3Dmol, ^e miniconda, ^g Open Babel, ^m PDB2PQR, ⁿ MGLTools, ^o Autodock Vina ^p	Yes
Lab.07	Molecular dynamics on GROMACS	GROMACS, ^b Biopython, ^d py3Dmol, ^e NGL Viewer ^f	Yes
Lab.08	Trajectory analysis using MDAnalysis	py3Dmol, ^e MDAnalysis ^q	Yes
Lab.09	Folding simulations using structure-based models	SMOG2, SBM-enhanced GROMACS, ^c Biopython, ^d py3Dmol, ^e NGL Viewer ^f	No
Lab.10	Conformational changes using structure-based models	SMOG2, SBM-enhanced GROMACS, ^c Biopython, ^d py3Dmol, ^e NGL Viewer ^f	No
Lab.11	Prediction of interactions from the coevolutionary analysis of sequence information	Biopython, ^d py3Dmol, ^e infernal, ^r pyDCA ^s	Yes
Lab.12	Protein folding <i>ab initio</i> using Rosetta	pyRosetta, ^a Biopython, ^d py3Dmol ^e	Yes

pb3lab/ibm3202

Google Colab Tutorials for IBM3202

1 Contributor 0 Issues 51 Stars 30 Forks



Felipe Engelberger
Undergrad student

Engelberger F et al (2021) | J Chem Educ 98(5) 1801–1807

cesar.ramirez@uc.cl

26



Protein Biophysics, Biochemistry
& Bioinformatics Lab

From protein folding to function to evolution to design: Computational Research at PB³ Lab

César A. Ramírez-Sarmiento

Protein Biophysics, Biochemistry and Bioinformatics Lab

Institute for Biological and Medical Engineering
Pontificia Universidad Católica de Chile

Millennium Institute for Integrative Biology