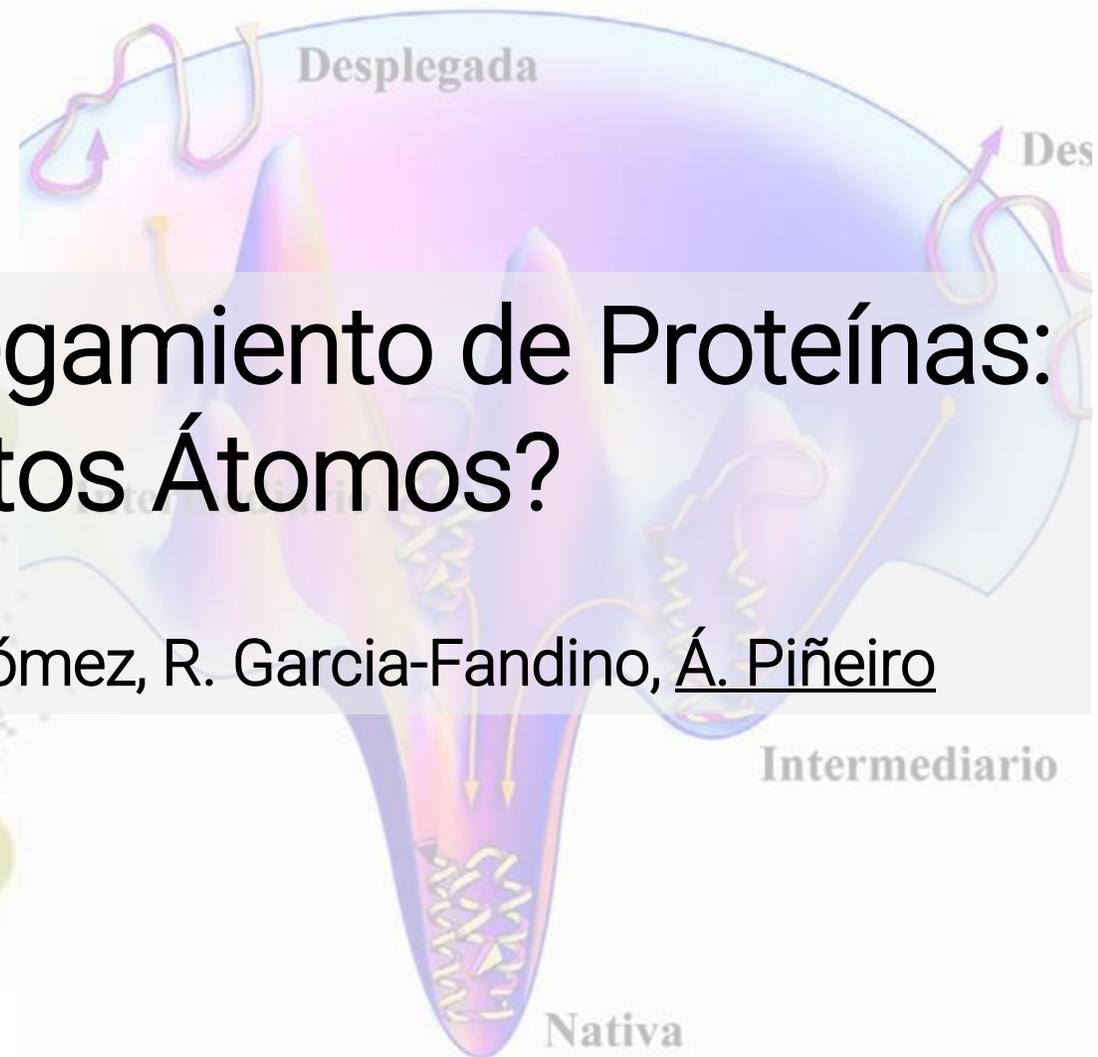


# Computación Cuántica en Plegamiento de Proteínas: ¿Suficientes Qubits para Tantos Átomos?

D. Conde-Torres, M. Mussa-Juane, D. Faílde, A. Gómez, R. Garcia-Fandino, Á. Piñeiro



# ¿ Quiénes somos y qué hacemos?



**Rebeca García**  
CSO



**Ángel Piñeiro**  
CEO



**Amelia Anderson**  
PhD student CyTx

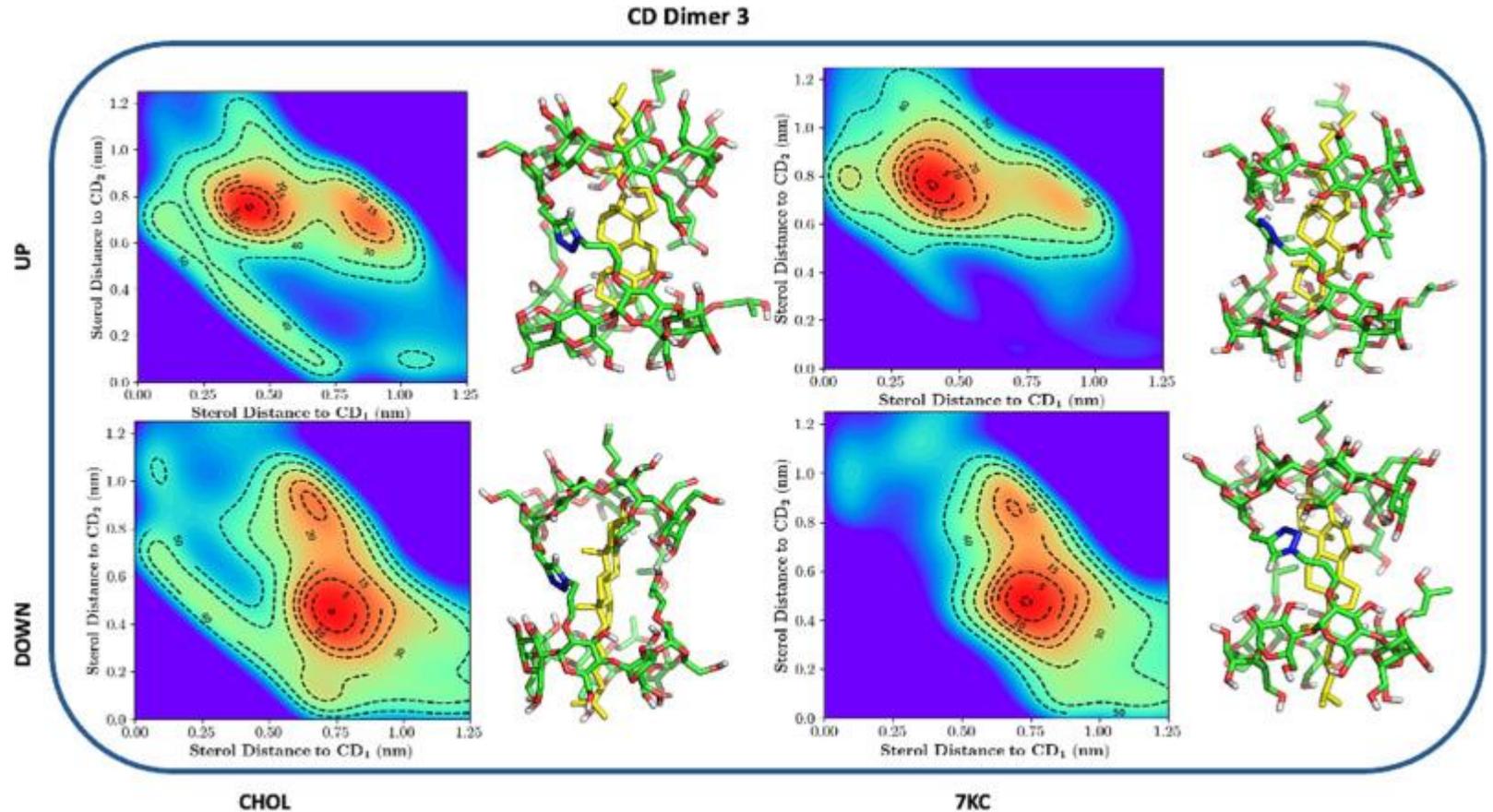
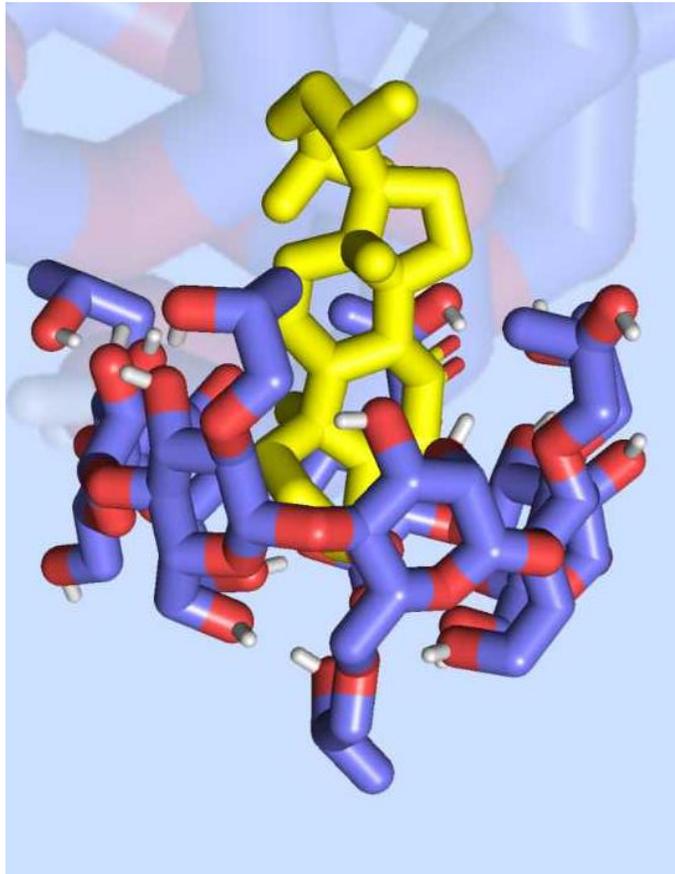


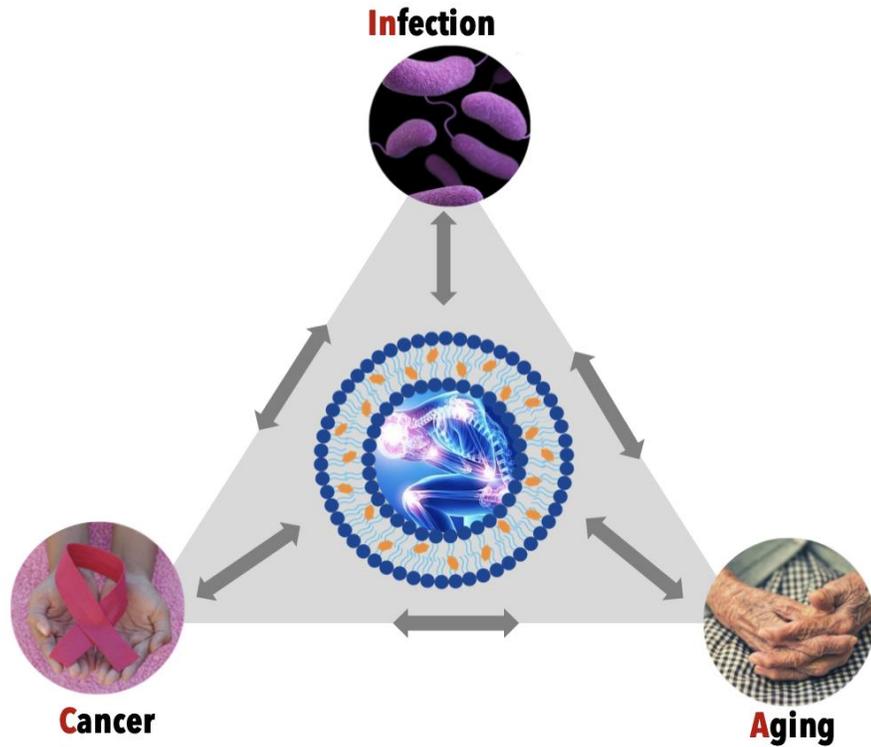
**Alexandre Blanco**  
PhD student DI Xunta



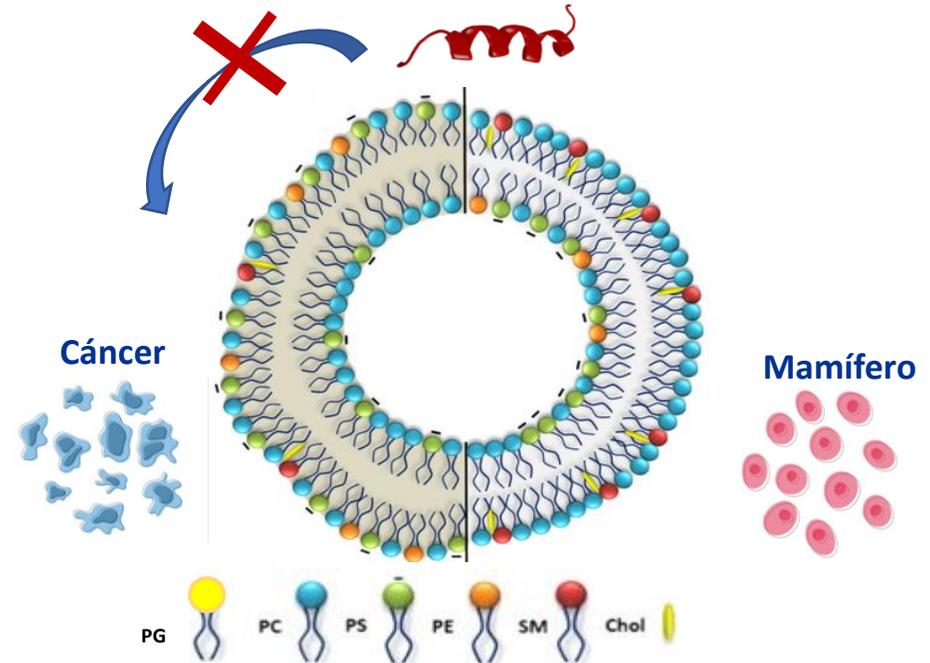
**Fabián Suárez**  
PhD student DI Xunta

**Caracterización de interacciones moleculares a partir de simulaciones computacionales → diseño/optimización de moléculas**



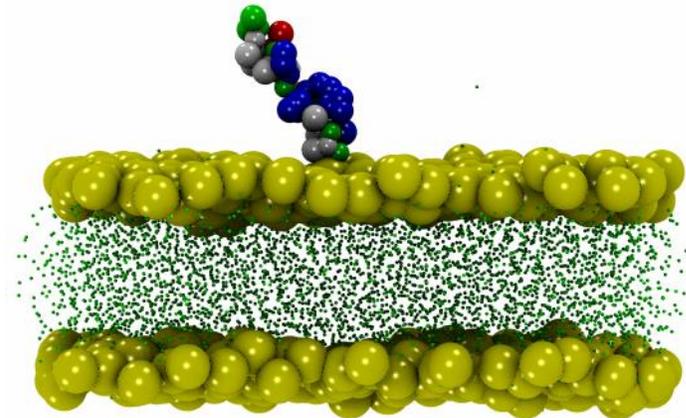
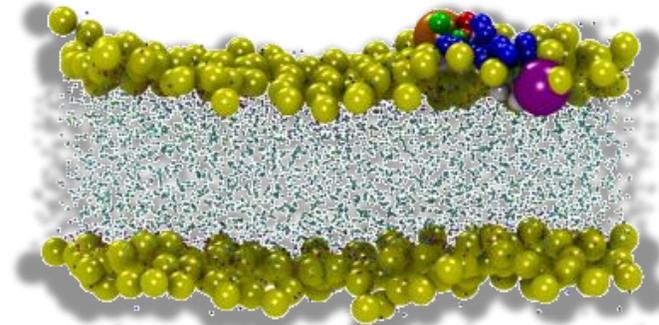
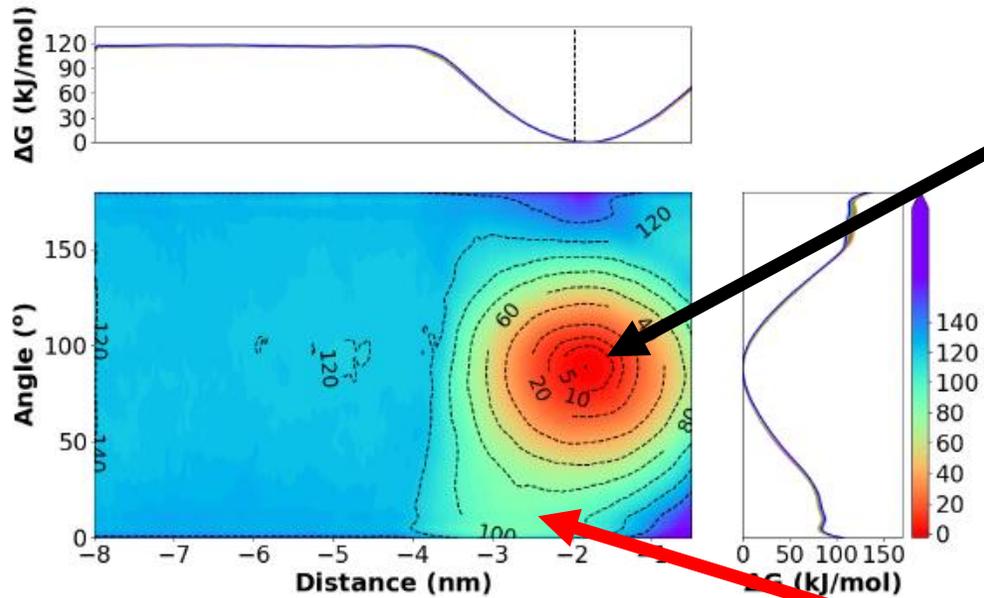


**Triángulo de CAln**  
(**C**ancer-**A**ging-**I**nfection)



**Problema: estructura**  
**de los péptidos**

## Aproximación(es)



nature

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [articles](#) > article

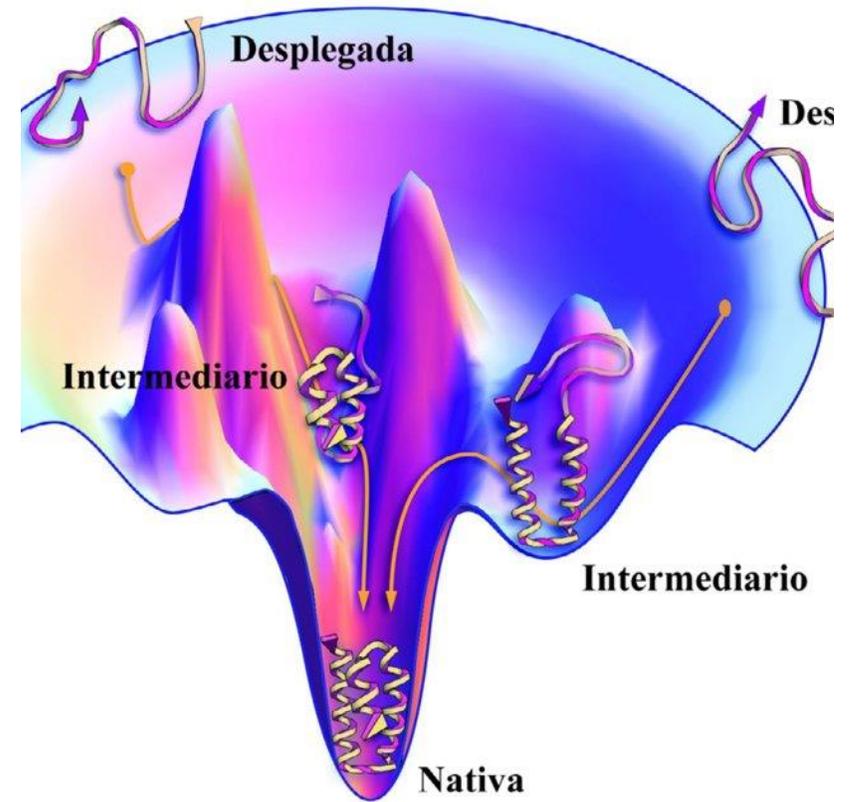
Article | [Open access](#) | [Published: 15 July 2021](#)

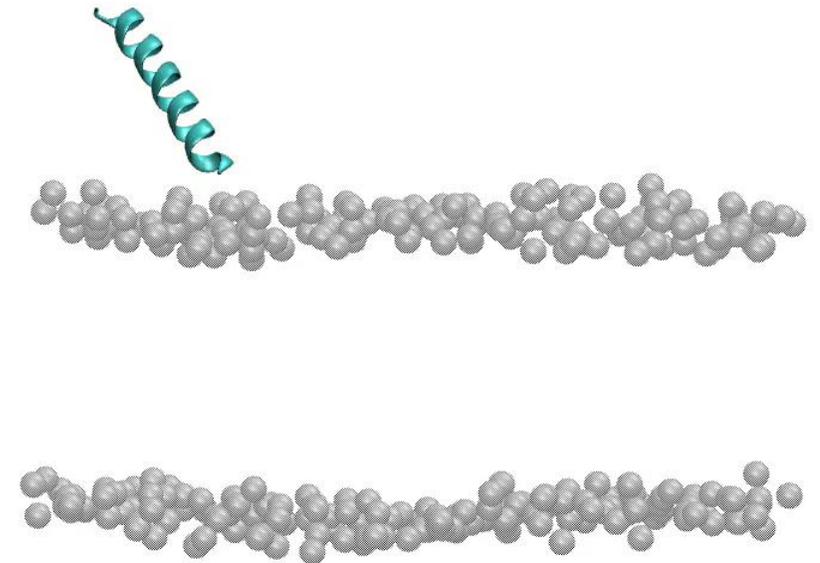
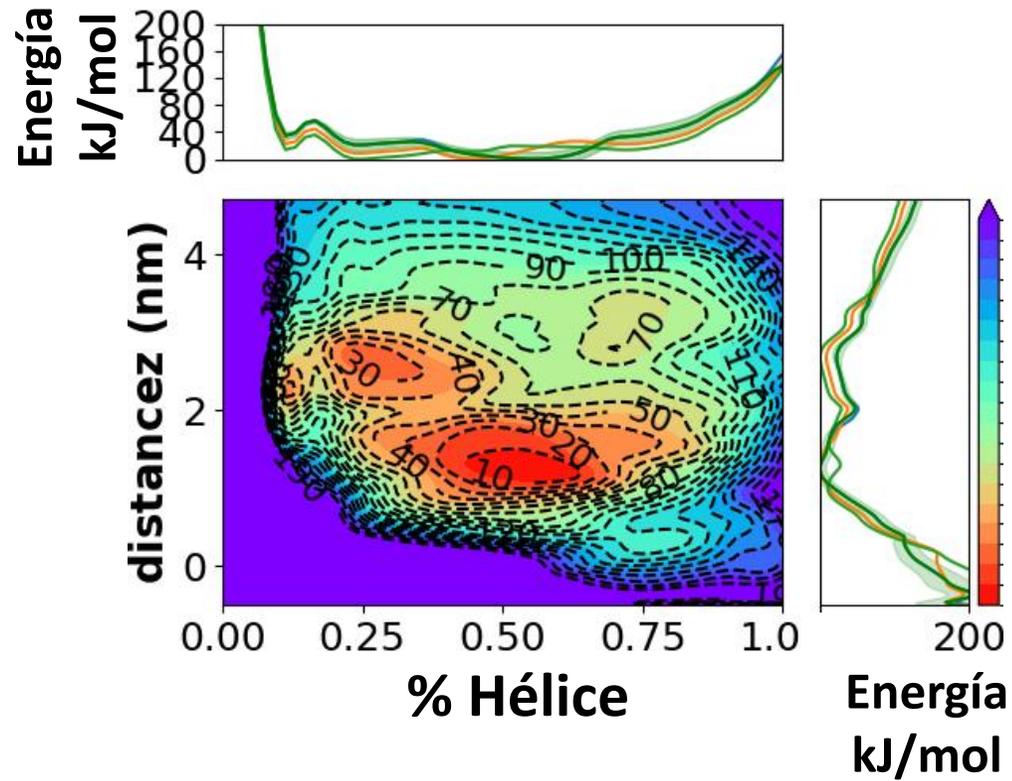
## Highly accurate protein structure prediction with AlphaFold

[John Jumper](#) , [Richard Evans](#), [Alexander Pritzel](#), [Tim Green](#), [Michael Figurnov](#), [Olaf Ronneberger](#), [Kathryn Tunyasuvunakool](#), [Russ Bates](#), [Augustin Žídek](#), [Anna Potapenko](#), [Alex Bridgland](#), [Clemens Meyer](#), [Simon A. A. Kohl](#), [Andrew J. Ballard](#), [Andrew Cowie](#), [Bernardino Romera-Paredes](#), [Stanislav Nikolov](#), [Rishub Jain](#), [Jonas Adler](#), [Trevor Back](#), [Stig Petersen](#), [David Reiman](#), [Ellen Clancy](#), [Michal Zielinski](#), [Martin Steinegger](#), [Michalina Pacholska](#), [Tamas Berghammer](#), [Sebastian Bodenstern](#), [David Silver](#), [Oriol Vinyals](#), [Andrew W. Senior](#), [Koray Kavukcuoglu](#), [Pushmeet Kohli](#) & [Demis Hassabis](#)  [— Show fewer authors](#)

*Nature* **596**, 583–589 (2021) | [Cite this article](#)

**1.29m** Accesses | **10k** Citations | **3528** Altmetric | [Metrics](#)





[https://github.com/qiskit-community/qiskit-research/blob/main/docs/protein\\_folding/protein\\_folding.ipynb](https://github.com/qiskit-community/qiskit-research/blob/main/docs/protein_folding/protein_folding.ipynb)

npj | quantum information

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [npj quantum information](#) > [articles](#) > [article](#)

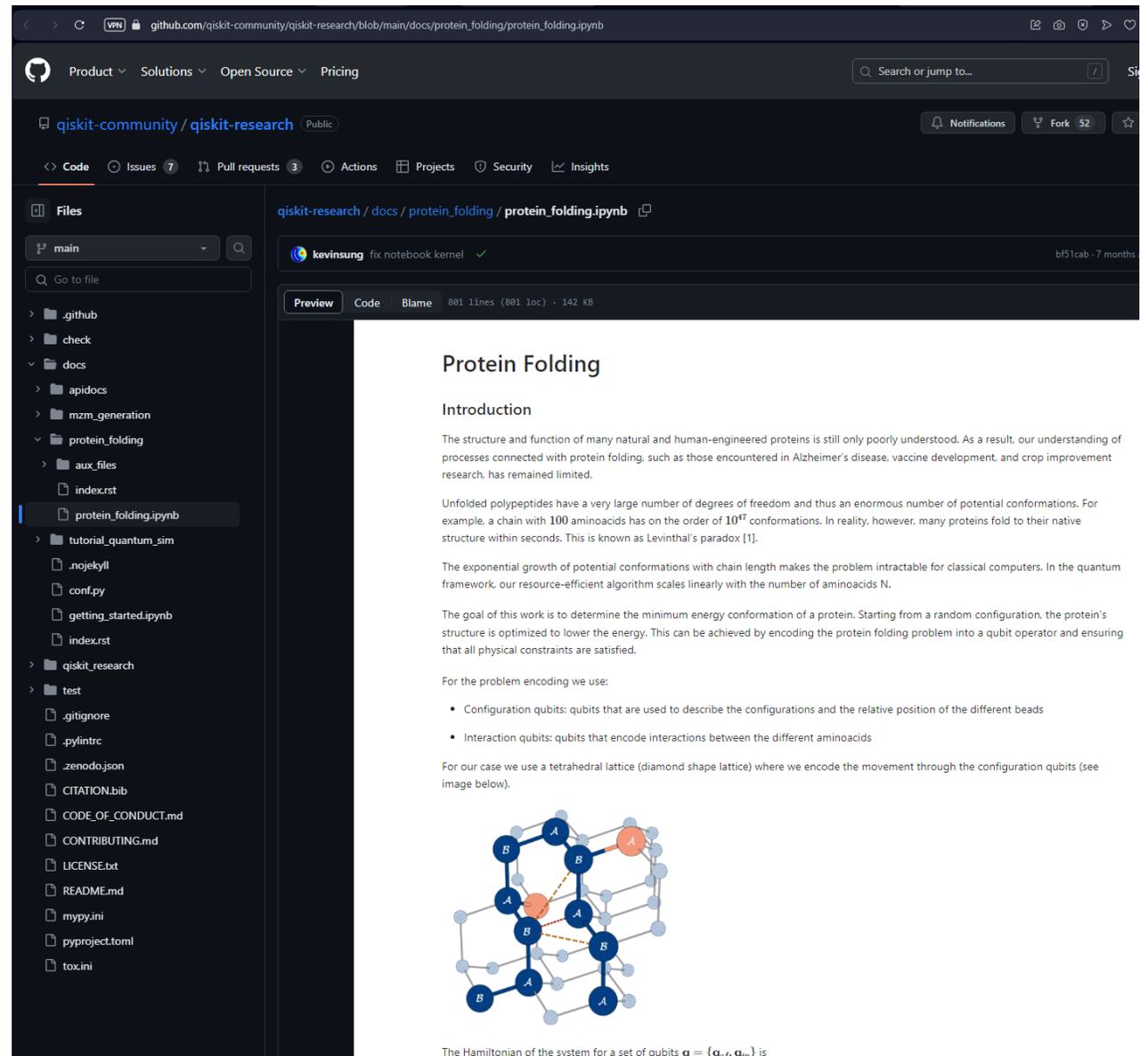
Article | [Open access](#) | [Published: 17 February 2021](#)

## Resource-efficient quantum algorithm for protein folding

[Anton Robert](#), [Panagiotis Kl. Barkoutsos](#), [Stefan Woerner](#) & [Ivano Tavernelli](#) 

[npj Quantum Information](#) **7**, Article number: 38 (2021) | [Cite this article](#)

**23k** Accesses | **49** Citations | **54** Altmetric | [Metrics](#)



Product ▾ Solutions ▾ Open Source ▾ Pricing

qiskit-community / qiskit-research Public

Code Issues 7 Pull requests 3 Actions Projects Security Insights

Files

- main
- Go to file
- .github
- check
- docs
  - apidocs
  - mzm\_generation
  - protein\_folding
    - aux\_files
    - index.rst
    - protein\_folding.ipynb
  - tutorial\_quantum\_sim
- qiskit\_research
  - test
    - .gitignore
    - .pylintrc
    - .zenodo.json
    - CITATION.bib
    - CODE\_OF\_CONDUCT.md
    - CONTRIBUTING.md
    - LICENSE.txt
    - README.md
    - mypy.ini
    - pyproject.toml
    - tox.ini

kevincsung fix notebook kernel ✓ b51cab · 7 months

Preview Code Blame 881 lines (881 loc) · 142 KB

### Protein Folding

#### Introduction

The structure and function of many natural and human-engineered proteins is still only poorly understood. As a result, our understanding of processes connected with protein folding, such as those encountered in Alzheimer's disease, vaccine development, and crop improvement research, has remained limited.

Unfolded polypeptides have a very large number of degrees of freedom and thus an enormous number of potential conformations. For example, a chain with 100 aminoacids has on the order of  $10^{47}$  conformations. In reality, however, many proteins fold to their native structure within seconds. This is known as Levinthal's paradox [1].

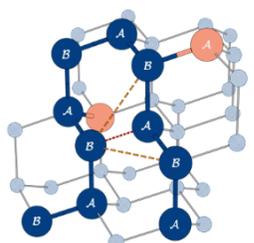
The exponential growth of potential conformations with chain length makes the problem intractable for classical computers. In the quantum framework, our resource-efficient algorithm scales linearly with the number of aminoacids  $N$ .

The goal of this work is to determine the minimum energy conformation of a protein. Starting from a random configuration, the protein's structure is optimized to lower the energy. This can be achieved by encoding the protein folding problem into a qubit operator and ensuring that all physical constraints are satisfied.

For the problem encoding we use:

- Configuration qubits: qubits that are used to describe the configurations and the relative position of the different beads
- Interaction qubits: qubits that encode interactions between the different aminoacids

For our case we use a tetrahedral lattice (diamond shape lattice) where we encode the movement through the configuration qubits (see image below).



The Hamiltonian of the system for a set of qubits  $\mathbf{a} = \{a_1, a_2, \dots\}$  is

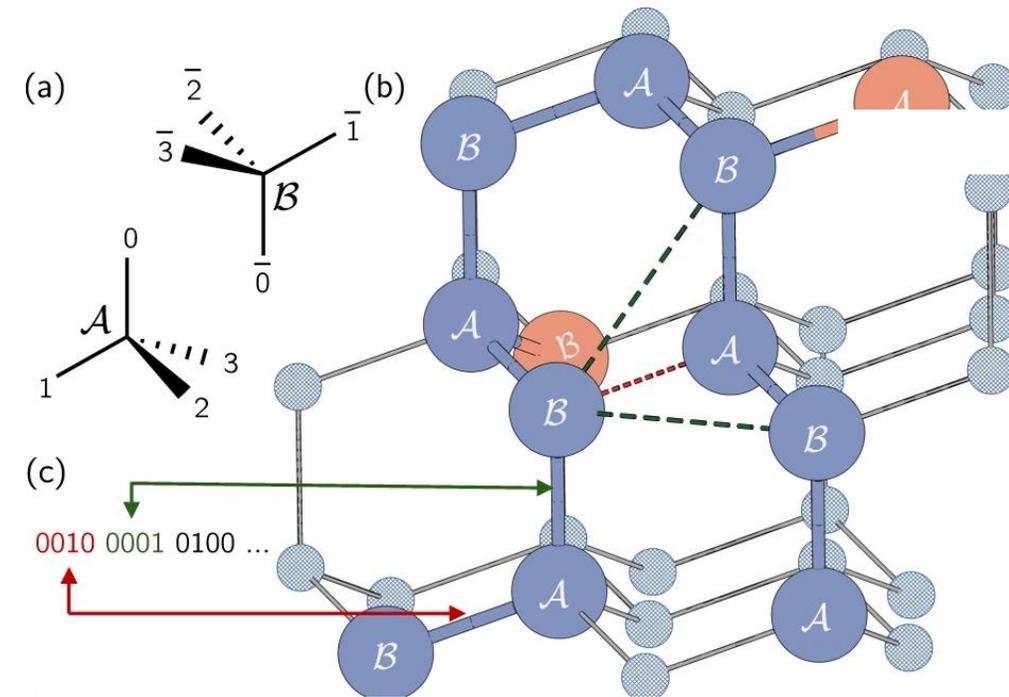
## The configuration qubits

As for the previous models in literature, a polymer configuration is grown on the lattice by adding the different beads one after the other and encoding, in the qubit register the different "turn"  $t_i$  that defines the position of the bead  $i + 1$  relatively to the previous bead  $i$ . Using a tetrahedral lattice, we distinguish two sets of non-equivalent lattice points  $\mathcal{A}$  and  $\mathcal{B}$  (see Fig. 1). At the  $\mathcal{A}$  sites, the polymer can only grow along the directions  $t_i \in \{0, 1, 2, 3\}$  while at site  $\mathcal{B}$  the possible directions are  $t_i \in \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$ . Along the sequence, the  $\mathcal{A}$  and  $\mathcal{B}$  sites are alternated so that we can use the convention that  $\mathcal{A}$  (respectively  $\mathcal{B}$ ) sites correspond to even (odd)  $i$ . Without loss of generality, the first two turns can be set to  $t_1 = \bar{1}$  and  $t_2 = 0$  due to symmetry degeneracy. To encode the turns, we assign one qubit per axis  $t_i = q_{4i-3}q_{4i-2}q_{4i-1}q_{4i}$  (Fig. 1(c)). Therefore, the total number of qubits required to encode a conformation  $\mathbf{q}_{cf}$  corresponds to  $N_{cf} = 4(N - 3)$ .

## The interaction qubits

To describe the interactions, we introduce a new qubit register  $\mathbf{q}_{in}$ , composed of  $q_{i,j}^{(l)}$  for each  $l^{\text{th}}$  nearest neighbor ( $l$ -NN) interaction on the lattice (see red and green dashed lines for  $l = 1$  and  $l = 2$  in Fig. 1, b) between beads  $i$  and  $j$ .

pairwise interaction energies  $\epsilon_{i,j}^{(l)}$  between the beads at distance  $l$  can be arbitrarily defined to reproduce a fold of interest or it can be adapted from pre-existing models, like the one proposed by Miyazawa and Jernigan (MJ) for 1-NN interactions<sup>17</sup>.



$$\mathbf{q} = \{\mathbf{q}_{cf}, \mathbf{q}_{in}\},$$

$$H(\mathbf{q}) = H_{gc}(\mathbf{q}_{cf}) + H_{ch}(\mathbf{q}_{cf}) + H_{in}(\mathbf{q})$$

## JMB

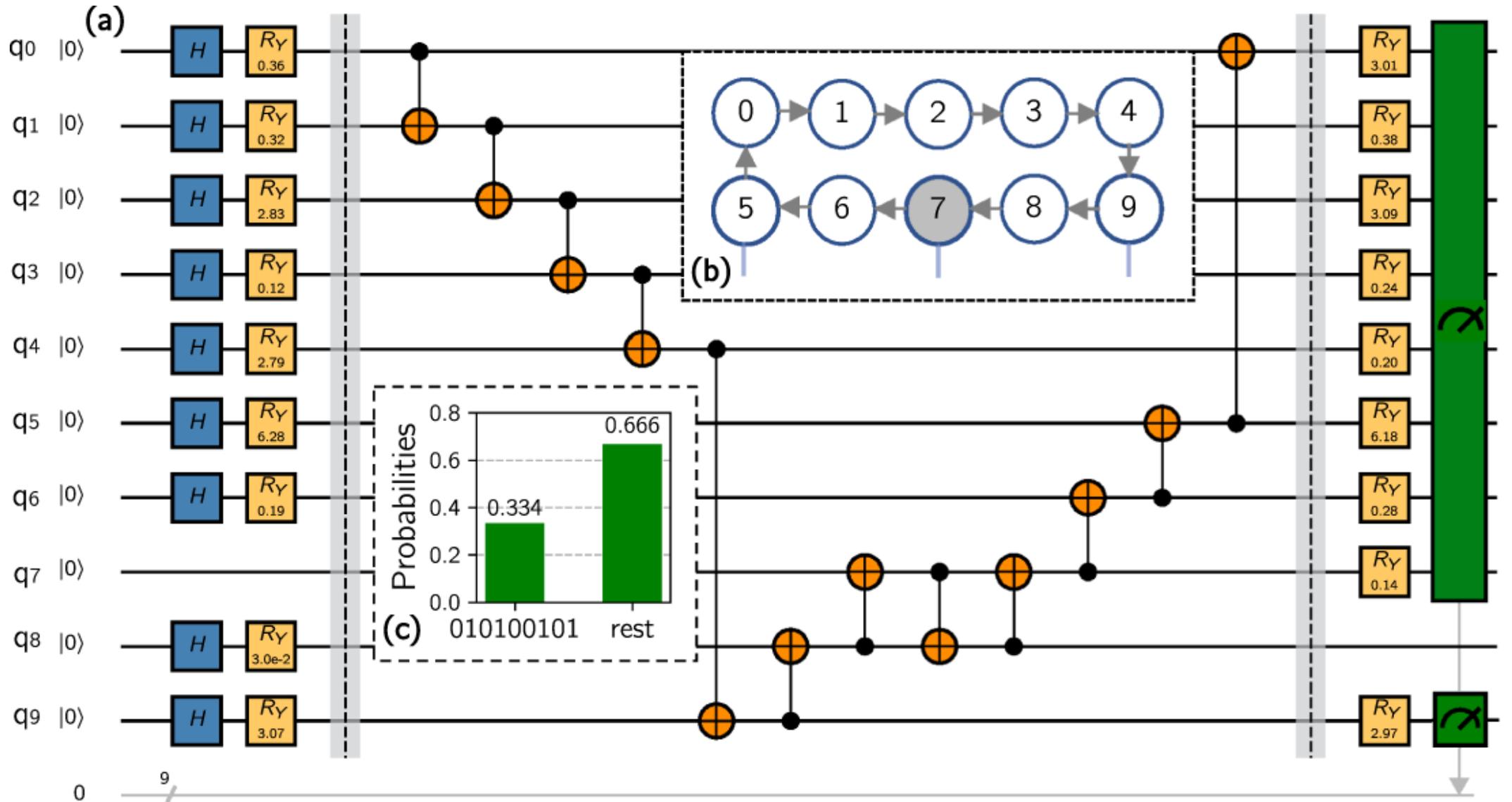
*J. Mol. Biol.* (1996) 256, 623–644

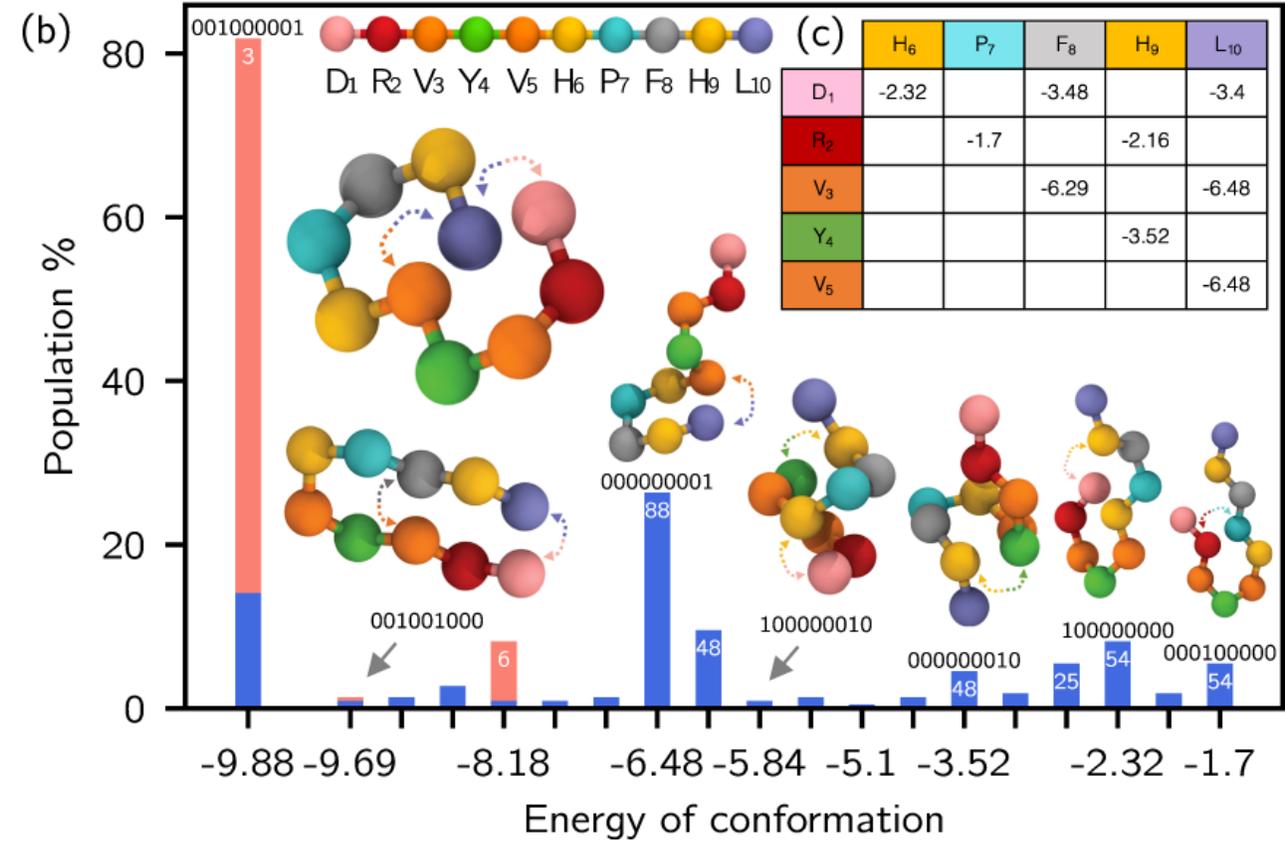
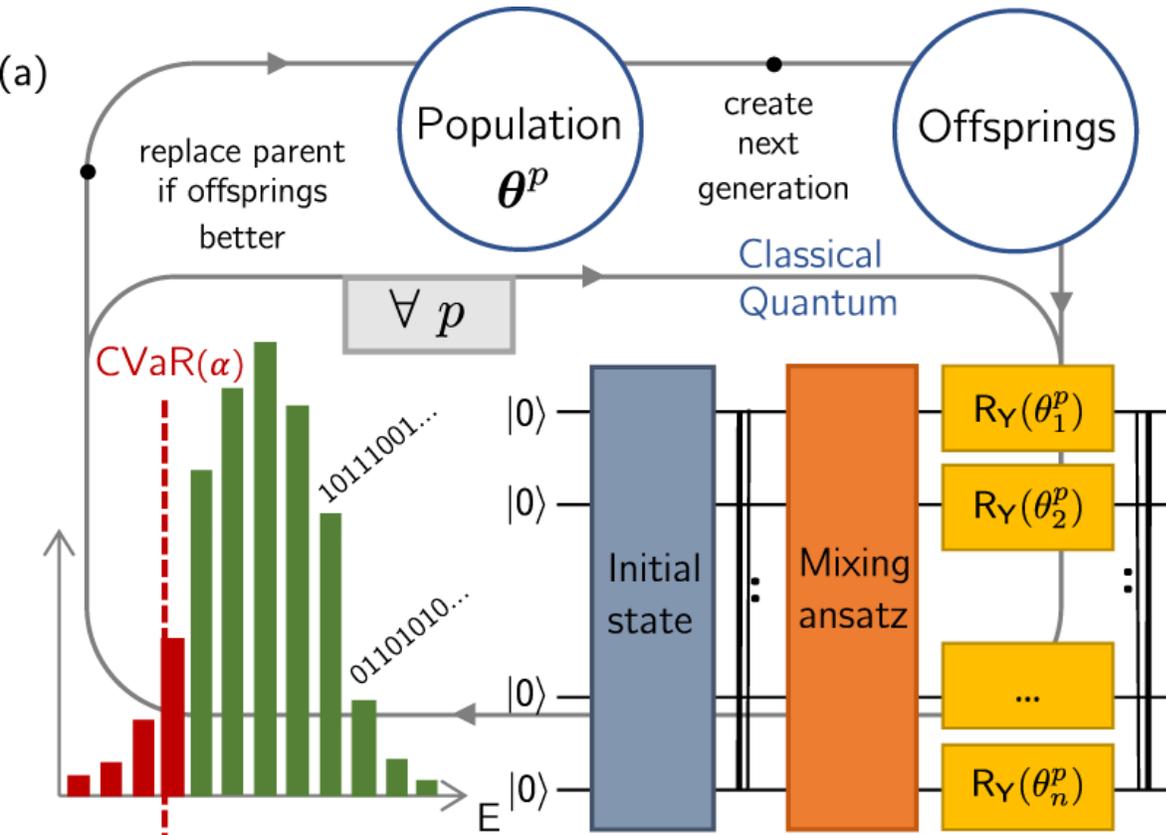
### Residue–Residue Potentials with a Favorable Contact Pair Term and an Unfavorable High Packing Density Term, for Simulation and Threading

Sanzo Miyazawa<sup>1</sup> and Robert L. Jernigan<sup>2\*</sup>

$\epsilon_{ij}$

	CYS	MET	PHE	ILE	LEU	VAL	TRP	TYR	ALA	GLY	THR	SER	GLN	ASN	GLU	ASP	HIS	ARG	LYS	PRO
CYS	-5.44	-5.05	-5.63	-5.03	-5.03	-4.46	-4.76	-3.89	-3.38	-3.16	-2.88	-2.86	-2.73	-2.59	-2.08	-2.66	-3.63	-2.70	-1.54	-2.92
MET	0.70	-6.06	-6.68	-6.33	-6.01	-5.52	-6.37	-4.92	-3.99	-3.75	-3.73	-3.55	-3.17	-3.50	-3.19	-2.90	-3.31	-3.49	-3.11	-4.11
PHE	0.52	-0.22	-6.85	-6.39	-6.26	-5.75	-6.02	-4.95	-4.36	-3.72	-3.76	-3.56	-3.30	-3.55	-3.51	-3.31	-4.61	-3.54	-2.83	-3.73
ILE	0.80	-0.18	0.14	-6.22	-6.17	-5.58	-5.64	-4.63	-4.41	-3.65	-3.74	-3.43	-3.22	-2.99	-3.23	-2.91	-3.76	-3.33	-2.70	-3.47
LEU	0.59	-0.09	0.06	-0.16	-5.79	-5.38	-5.50	-4.26	-3.96	-3.43	-3.43	-3.16	-3.09	-2.99	-2.91	-2.59	-3.84	-3.15	-2.63	-3.06
VAL	0.73	-0.02	0.14	0.00	-0.01	-4.94	-5.05	-4.05	-3.62	-3.06	-2.95	-2.79	-2.67	-2.36	-2.56	-2.25	-3.38	-2.78	-1.95	-2.96
TRP	0.67	-0.63	0.12	0.19	0.11	0.13	-5.42	-4.44	-3.93	-3.37	-3.31	-2.95	-3.16	-3.11	-2.94	-2.91	-4.02	-3.56	-2.49	-3.66
TYR	0.60	-0.12	0.25	0.25	0.41	0.19	0.04	-3.55	-2.85	-2.50	-2.48	-2.30	-2.53	-2.47	-2.42	-2.25	-3.33	-2.75	-2.01	-2.80
ALA	0.59	0.29	0.31	-0.05	0.19	0.10	0.03	0.18	-2.51	-2.15	-2.15	-1.89	-1.70	-1.44	-1.51	-1.57	-2.09	-1.50	-1.10	-1.81
GLY	0.64	0.37	0.79	0.55	0.56	0.50	0.43	0.36	0.19	-2.17	-2.03	-1.70	-1.54	-1.56	-1.22	-1.62	-1.94	-1.68	-0.84	-1.72
THR	0.70	0.16	0.52	0.23	0.33	0.38	0.26	0.15	-0.04	-0.09	-1.72	-1.59	-1.59	-1.51	-1.45	-1.66	-2.31	-1.97	-1.02	-1.66
SER	0.61	0.22	0.61	0.42	0.48	0.42	0.50	0.21	0.11	0.13	0.00	-1.48	-1.37	-1.31	-1.48	-1.46	-1.94	-1.22	-0.83	-1.35
GLN	0.43	0.30	0.56	0.33	0.25	0.24	-0.01	-0.31	0.00	-0.01	-0.29	-0.18	-0.89	-1.36	-1.33	-1.26	-1.85	-1.85	-1.02	-1.73
ASN	0.93	0.33	0.67	0.91	0.70	0.91	0.39	0.10	0.61	0.32	0.14	0.23	-0.13	-1.59	-1.43	-1.33	-2.01	-1.41	-0.91	-1.43
GLU	1.23	0.43	0.50	0.47	0.58	0.49	0.36	-0.06	0.34	0.45	0.00	-0.15	-0.30	-0.04	-1.18	-1.23	-2.27	-2.07	-1.60	-1.40
ASP	0.54	0.61	0.59	0.68	0.79	0.71	0.28	0.01	0.16	-0.05	-0.32	-0.23	-0.33	-0.06	-0.16	-0.96	-2.14	-1.98	-1.32	-1.19
HIS	0.48	1.11	0.21	0.75	0.44	0.48	0.08	-0.17	0.55	0.53	-0.06	0.19	-0.02	0.18	-0.29	-0.26	-2.78	-2.12	-1.09	-2.17
ARG	0.71	0.23	0.58	0.48	0.43	0.38	-0.16	-0.28	0.44	0.10	-0.42	0.21	-0.72	0.07	-0.79	-0.80	-0.04	-1.39	-0.06	-1.85
LYS	1.11	-0.15	0.53	0.34	0.20	0.45	0.15	-0.31	0.08	0.18	-0.23	-0.15	-0.65	-0.19	-1.08	-0.90	0.24	0.57	0.13	-0.67
PRO	0.40	-0.49	0.29	0.23	0.42	0.10	-0.36	-0.43	0.03	-0.04	-0.21	-0.02	-0.69	-0.04	-0.22	-0.11	-0.19	-0.56	-0.15	-1.18



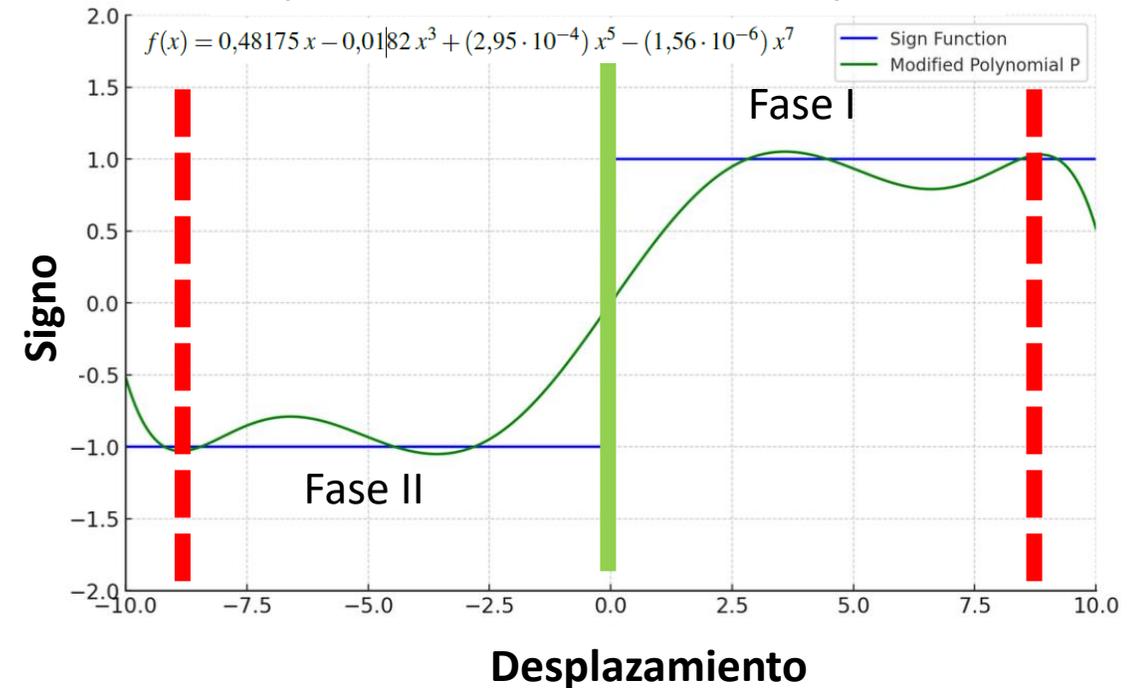


Fauchère, J., and Pliska, V. **1983**. Hydrophobic parameters of amino-acid side chains from the partitioning of N-acetyl-amino-acid amides. *Eur. J. Med. Chem.* 8: 369–375

Nomenclatura, grado de hidrofobicidad y comportamiento en disolución de los aminoácidos naturales

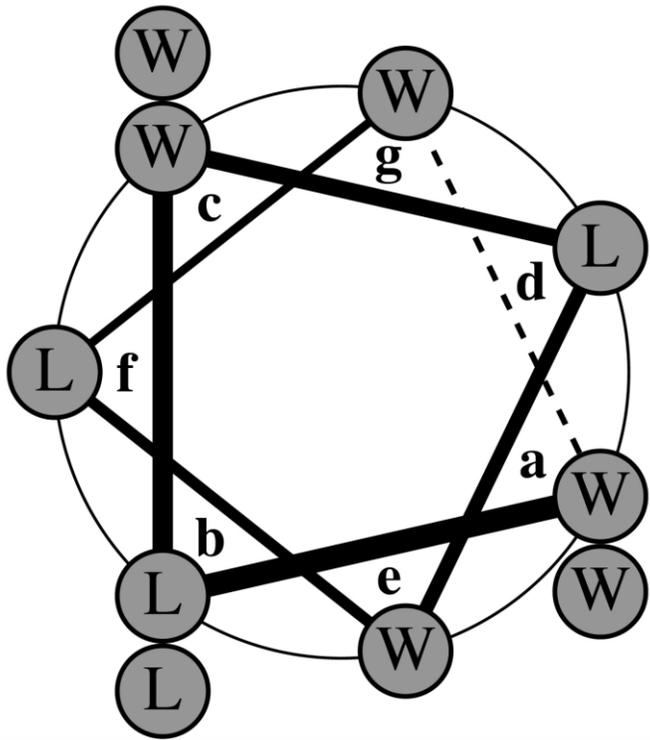
Código de 3 letras	Código de 1 letra	Hidrofobicidad	Comportamiento
ASP	D	-0.77	ácido
GLU	E	-0.64	ácido
LYS	K	-0.99	básico
ARG	R	-1.01	básico
HIS	H	0.13	básico
HISH	H	0.13	básico
GLY	G	0.00	hidrofóbico
ALA	A	0.31	hidrofóbico
VAL	V	1.22	hidrofóbico
LEU	L	1.70	hidrofóbico
ILE	I	1.80	hidrofóbico
PRO	P	0.72	hidrofóbico
PHE	F	1.79	hidrofóbico
MET	M	1.23	hidrofóbico
TRP	W	2.25	hidrofóbico
SER	S	-0.04	polar
THR	T	0.26	polar
CYS	C	1.54	polar
TYR	Y	0.96	polar
ASN	N	-0.60	polar
GLN	Q	-0.22	polar

## Aproximación de Chebyshev

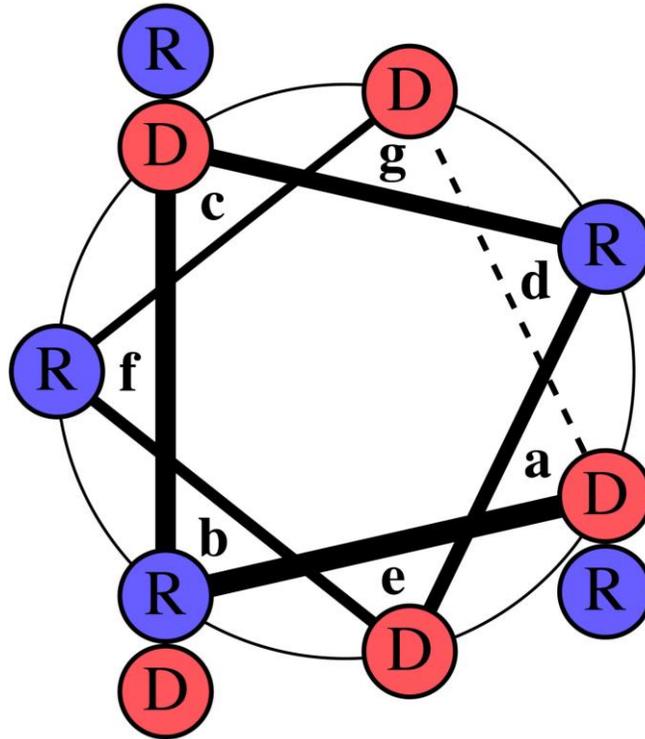


$$H(\mathbf{q}) = H_{gc}(\mathbf{q}_{cf}) + H_{ch}(\mathbf{q}_{cf}) + H_{in}(\mathbf{q}) + H_{fase}(\mathbf{q})$$

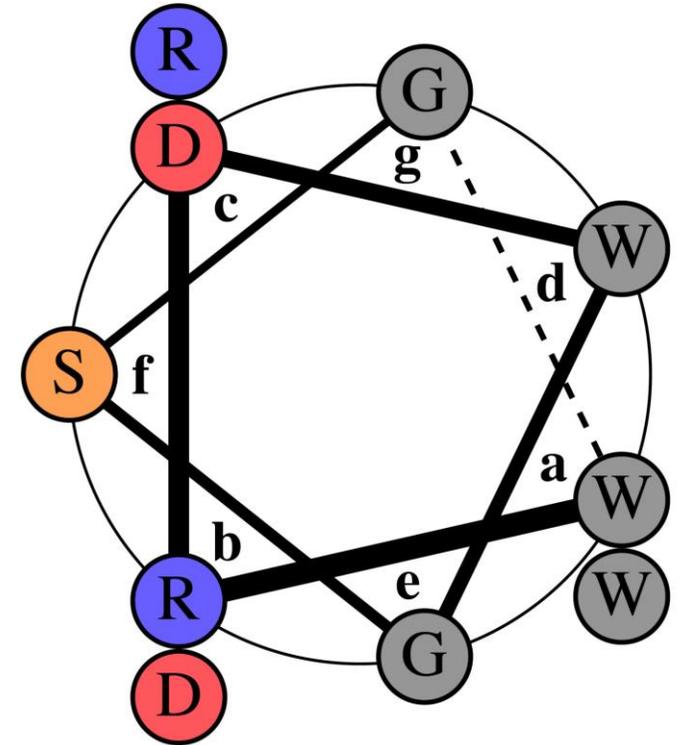
$$H_{fase}(\mathbf{q}) = \sum (\Delta_{polaridad} \cdot signo \cdot K)_a$$



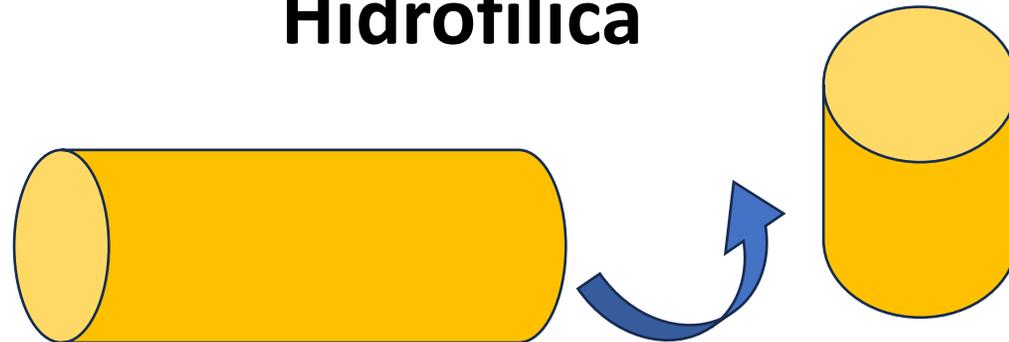
**Hidrofóbica**

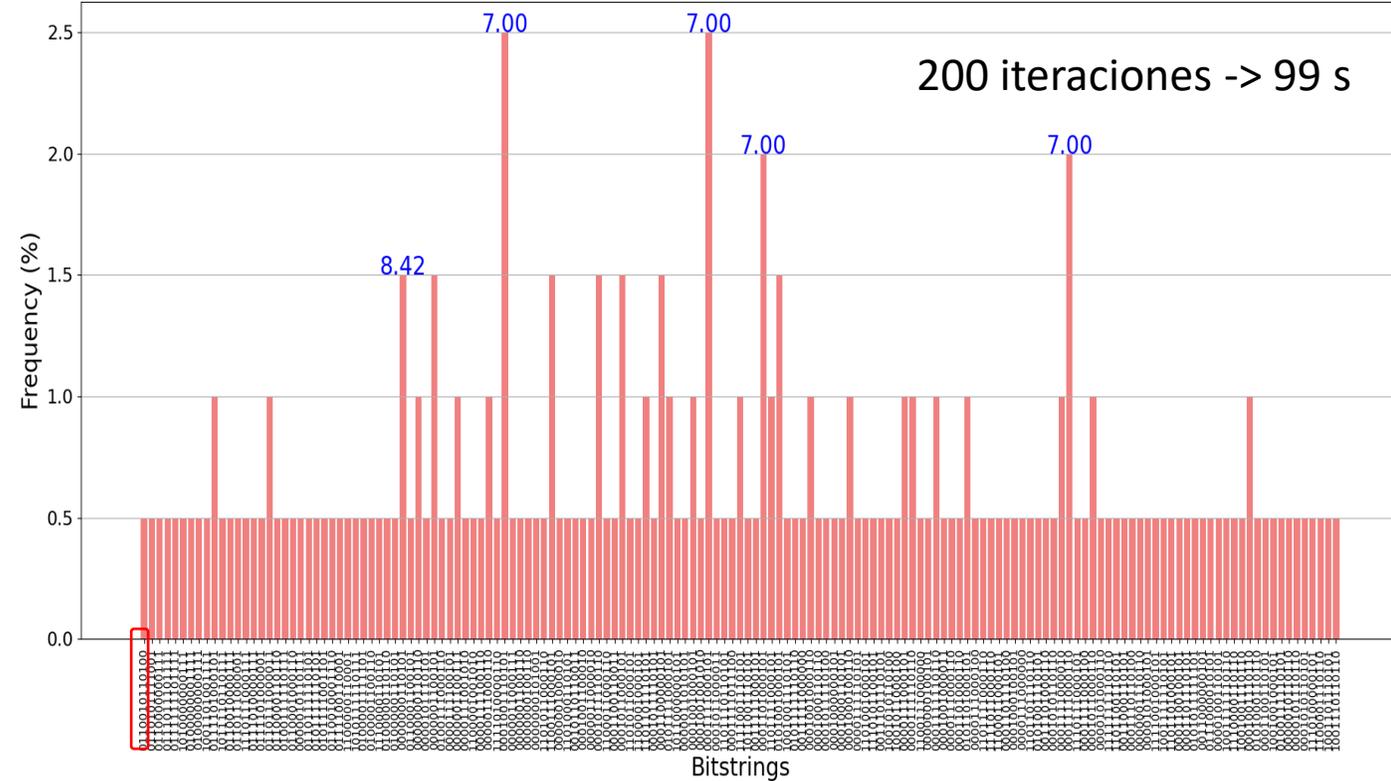
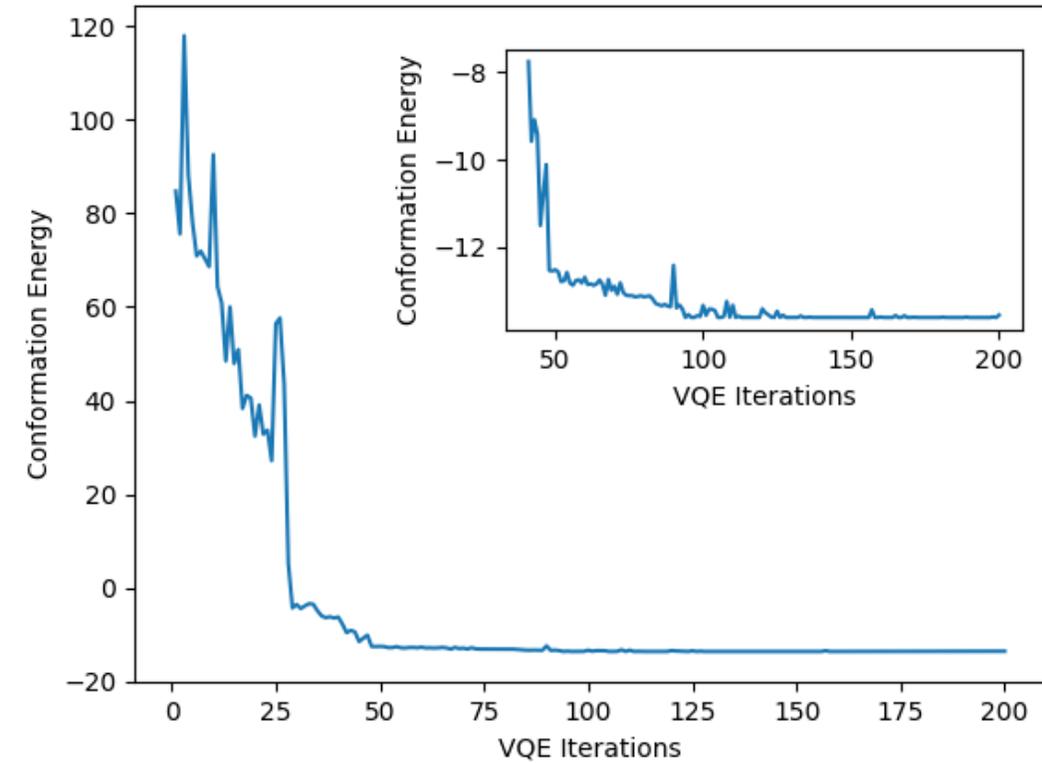


**Hidrofílica**

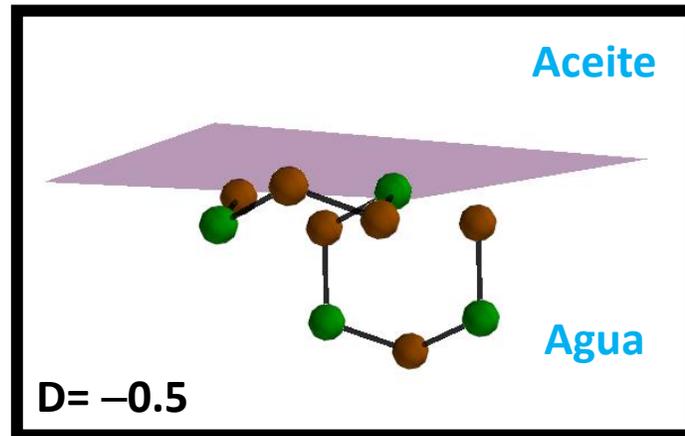
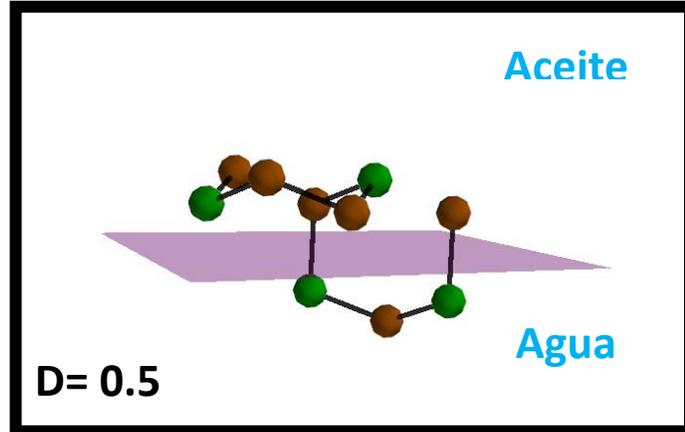


**Anfifílica**

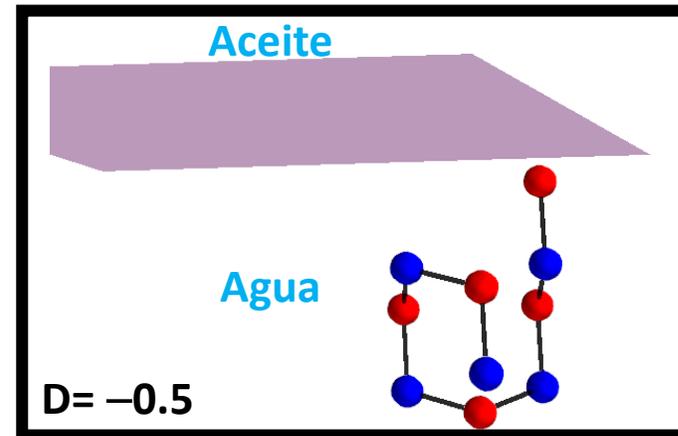
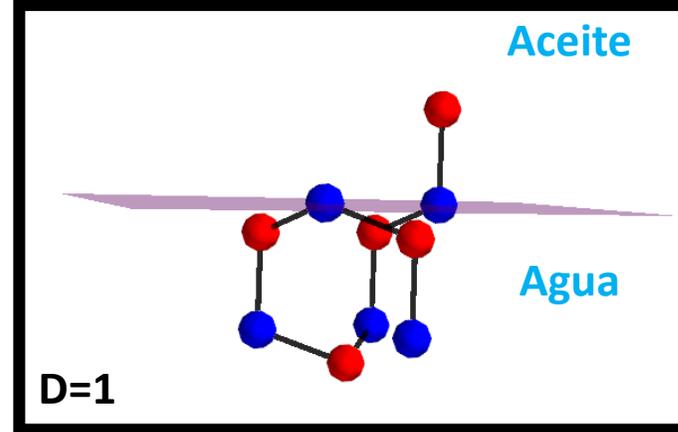




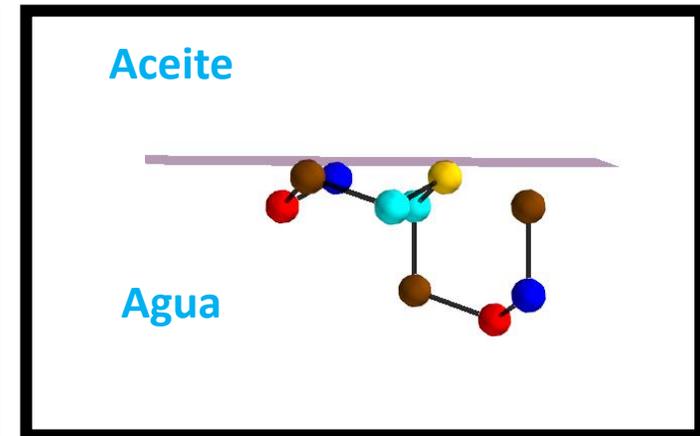
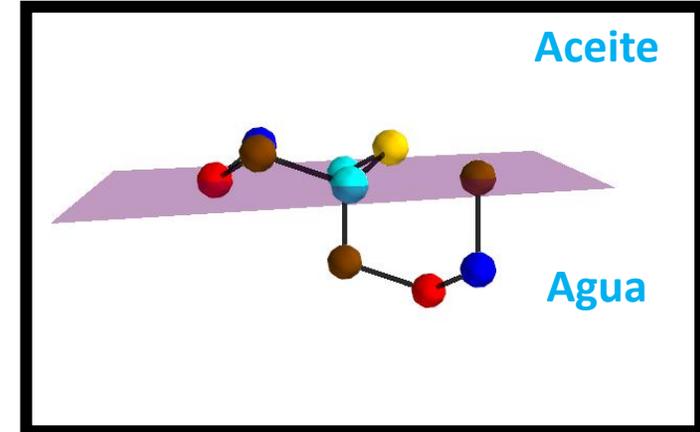
## Hidrofóbica



## Hidrofílica



## Anfifílica



- W (Trp)
- L (Leu)
- R (Arg)
- D (Asp)
- G (Gly)
- S (Ser)

- **Ventaja cuántica** en problemas de plegamiento de péptidos → **eficiencia de muestreo**, principalmente cuando crece el número de aminoácidos ya que el espacio de posibles soluciones crece exponencialmente.
- **Hemos conseguido implementar el efecto de una interfase** que separa dos medios de diferente polaridad, para modelar la superficie de una membrana.
- Nuestra **implementación de la interfase es eficiente**, ya que no precisa aumentar el número de qubits.
- El modelo tiene **muchas aproximaciones** que hacen que los resultados todavía no sean útiles desde el punto de vista práctico, principalmente por el **reducido número de partículas** que se pueden modelar y por la **aproximación discreta** de las posiciones de estas partículas.
- **Muchas oportunidades de optimización** utilizando nuevos algoritmos de computación cuántica y grandes expectativas en mejora de hardware.

- **Daniel Conde-Torres**
- **Mariamo Mussa-Juane**
- **Daniel Faílde**
- **Andrés Gómez**
- **Rebeca García Fandiño**



A iniciativa do Polo de Tecnoloxías Cuánticas de Galicia conta con financiamento de:

## Fondos REACT EU



AXENCIA  
GALEGA DE  
INNOVACIÓN



UNIÓN EUROPEA



Despregamento dunha infraestrutura baseada en tecnoloxías cuánticas da información que permita impulsar a I+D+i en Galicia.

Apoiar a transición cara a unha economía dixital.

Operación financiada pola Unión Europea, a través do FONDO EUROPEO DE DESENVOLVEMENTO REXIONAL (FEDER), como parte da resposta da Unión á pandemia da COVID-19.

PROGRAMA OPERATIVO  
FEDER GALICIA  
2014-2020

*Unha maneira de facer Europa*