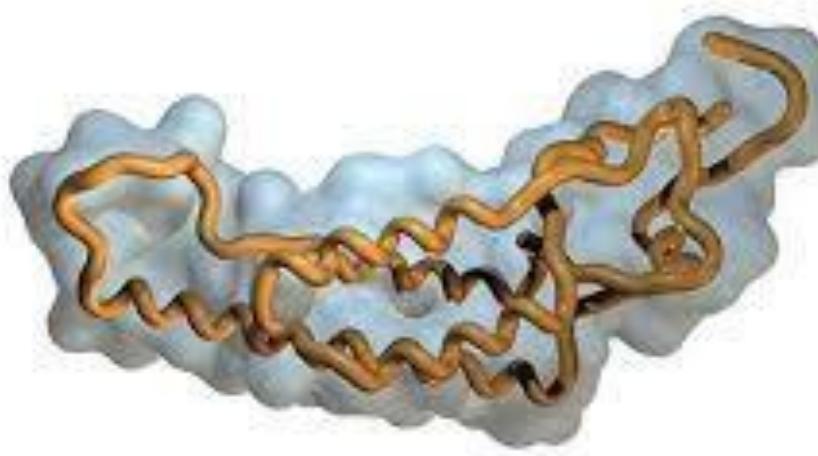


# Estructura de proteínas

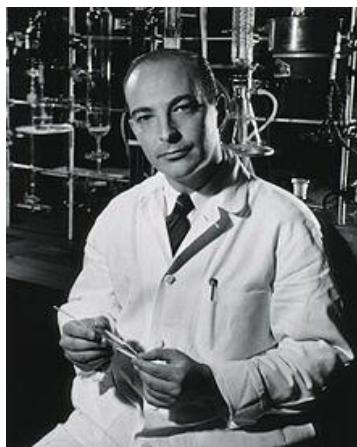


*Ceci n'est pas une protéine.*

Las proteínas son polímeros biológicos lineales compuestos por aminoácidos. Sus tamaños son tan grandes que constituyen un ejemplo de macromoléculas. Pueden estar constituidas por una o varias cadenas iguales o distintas. La secuencia ordenada de aminoácidos está definida por la secuencia del gen que codifica para dicha proteína. Los aminoácidos que componen a las proteínas son mayormente 20, sin embargo en distintos procesos luego de la síntesis de la proteína, distintos aminoácidos puede sufrir modificaciones covalentes que aumentan el repertorio químico de los aminoácidos. Los tamaños de las cadena/s que constituyen a las proteínas es muy variables, yendo de ~15 hasta ~30000.

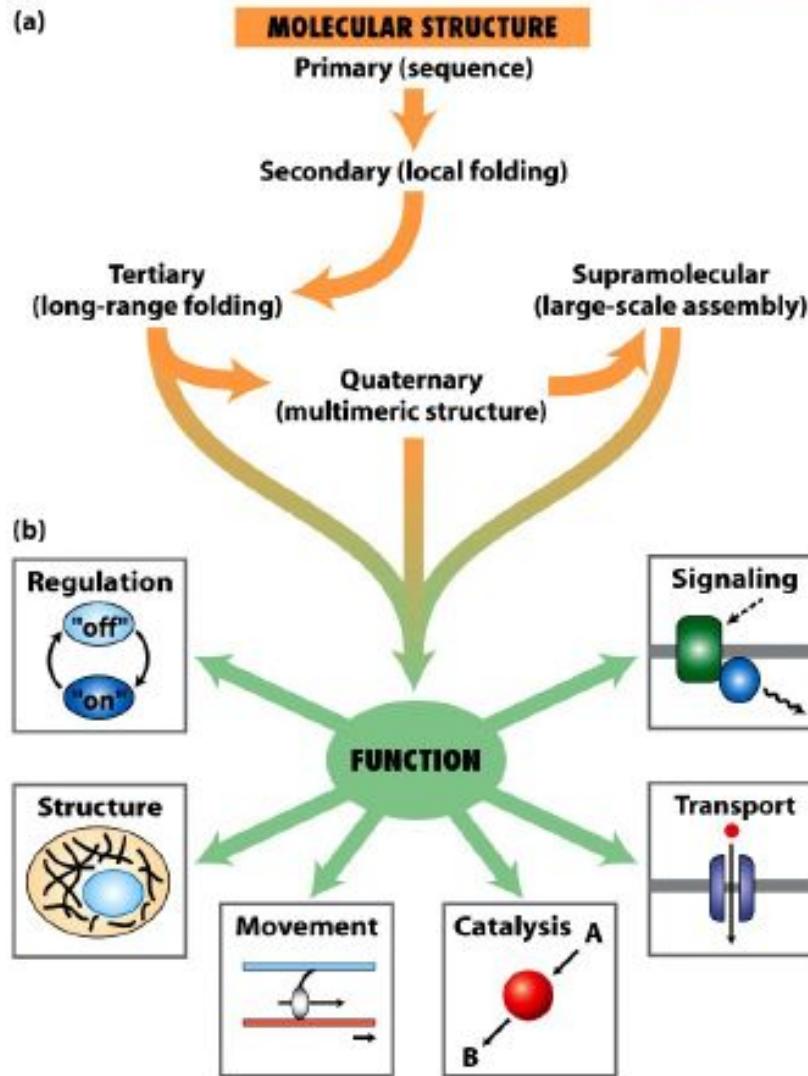
In his book *For the Love of Enzymes* published in 1989 the brilliant biochemist Arthur Kornberg wrote eloquently:

What chemical feature most clearly enables the living cell and organism to function, grow and reproduce? Not the carbohydrate stored as starch in plants or glycogen in animals, nor the depots of fat. It is not the structural proteins that form muscle, elastic tissue, and the skeletal fabric. Nor is it DNA, the genetic material. Despite its glamor, DNA is simply the construction manual that directs the assembly of the cell's proteins. The DNA is itself lifeless, its language cold and austere. What gives the cell its life and personality are enzymes. They govern all body processes; malfunction of even one enzyme can be fatal. Nothing in nature is so tangible and vital to our lives as proteins, and yet so poorly understood and appreciated by all but a few scientists.

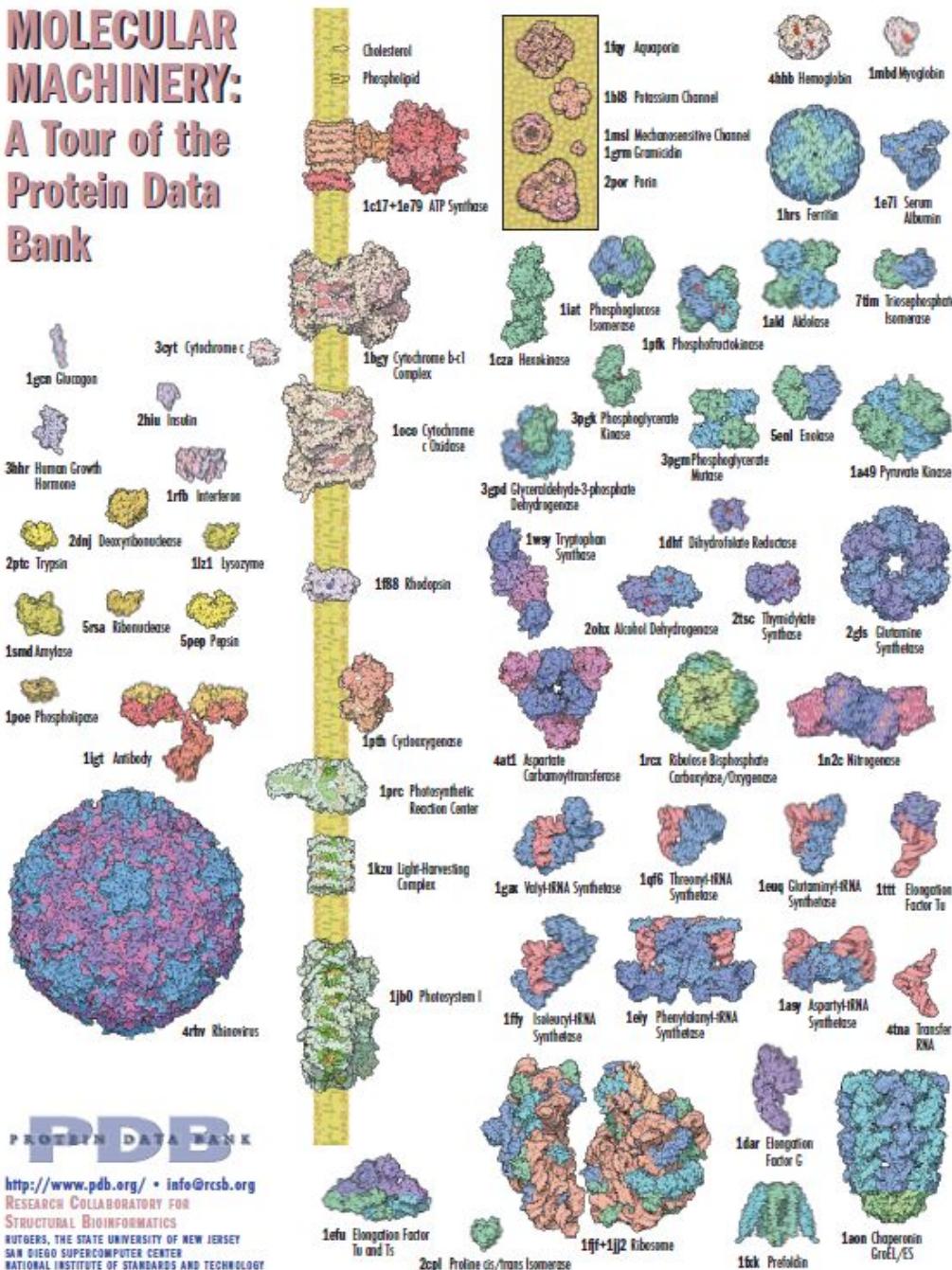


1918-2007

# Formas y funciones de proteínas

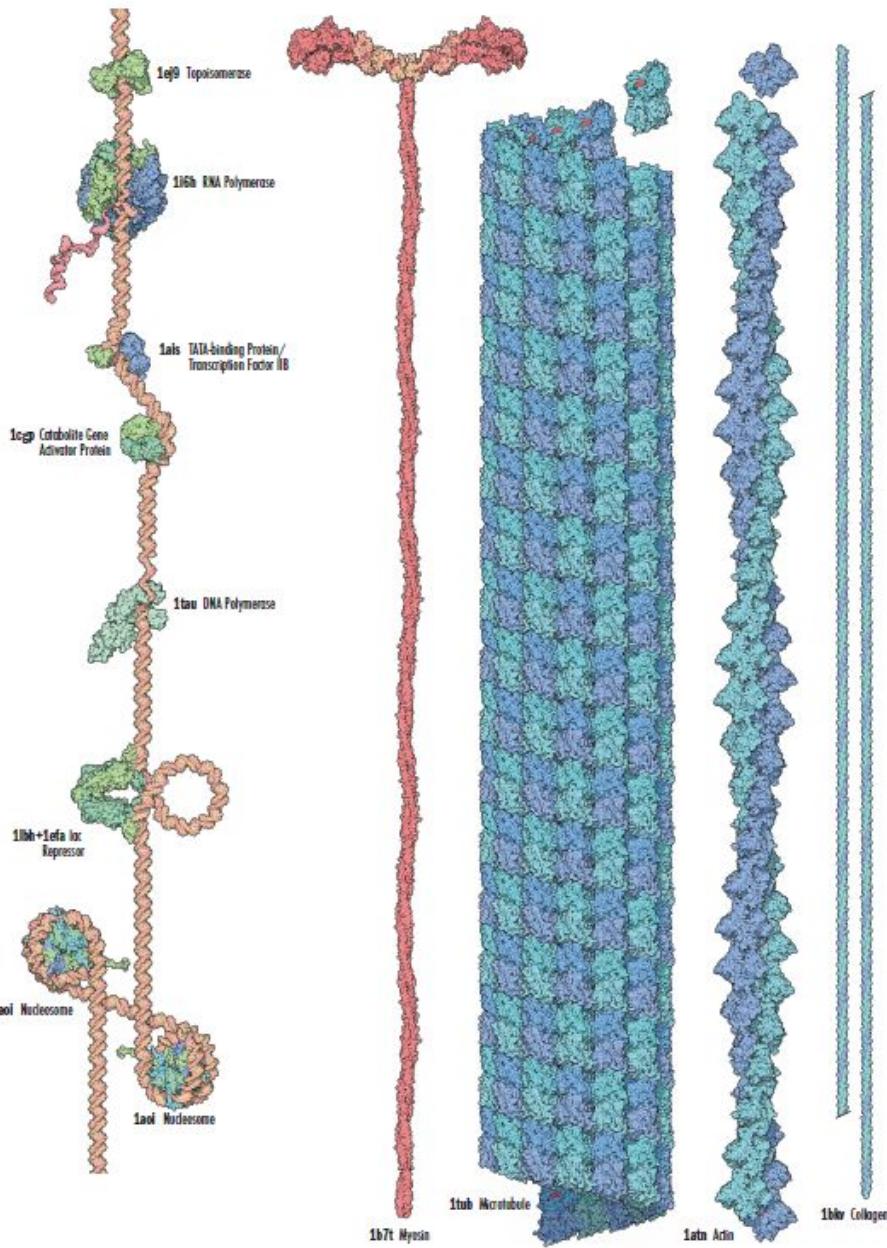


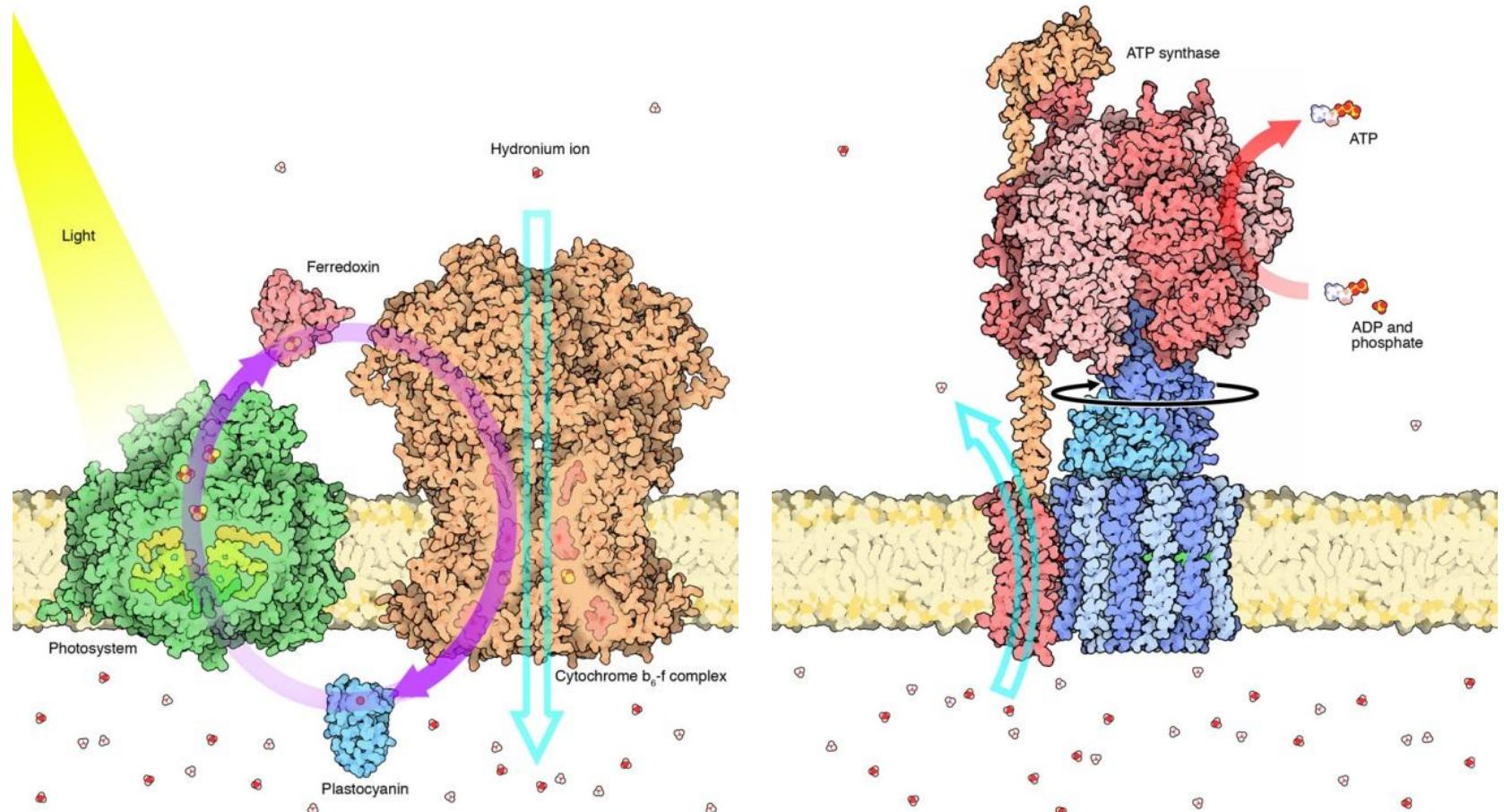
# MOLECULAR MACHINERY: A Tour of the Protein Data Bank



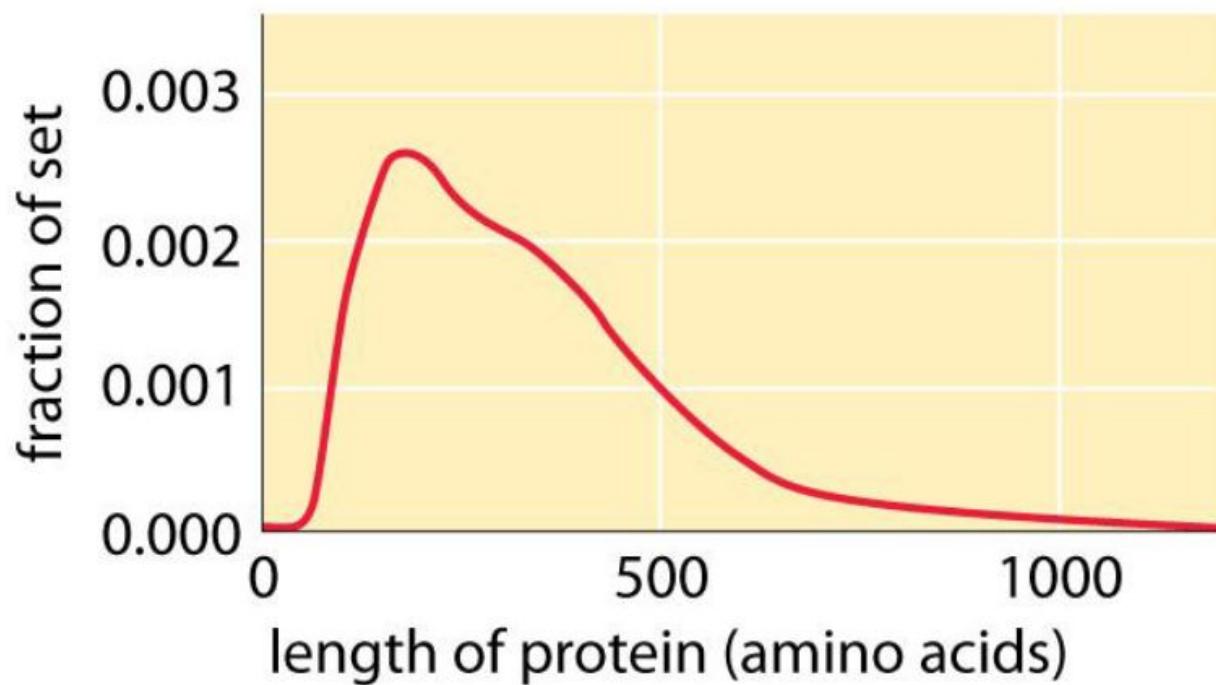
**PDB**  
PROTEIN DATA BANK

<http://www.pdb.org/> • [info@rcsb.org](mailto:info@rcsb.org)  
RESEARCH COLLABORATORY FOR  
STRUCTURAL BIOINFORMATICS  
RUTHERFORD, THE STATE UNIVERSITY OF NEW JERSEY  
SAN DIEGO SUPERCOMPUTER CENTER  
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY





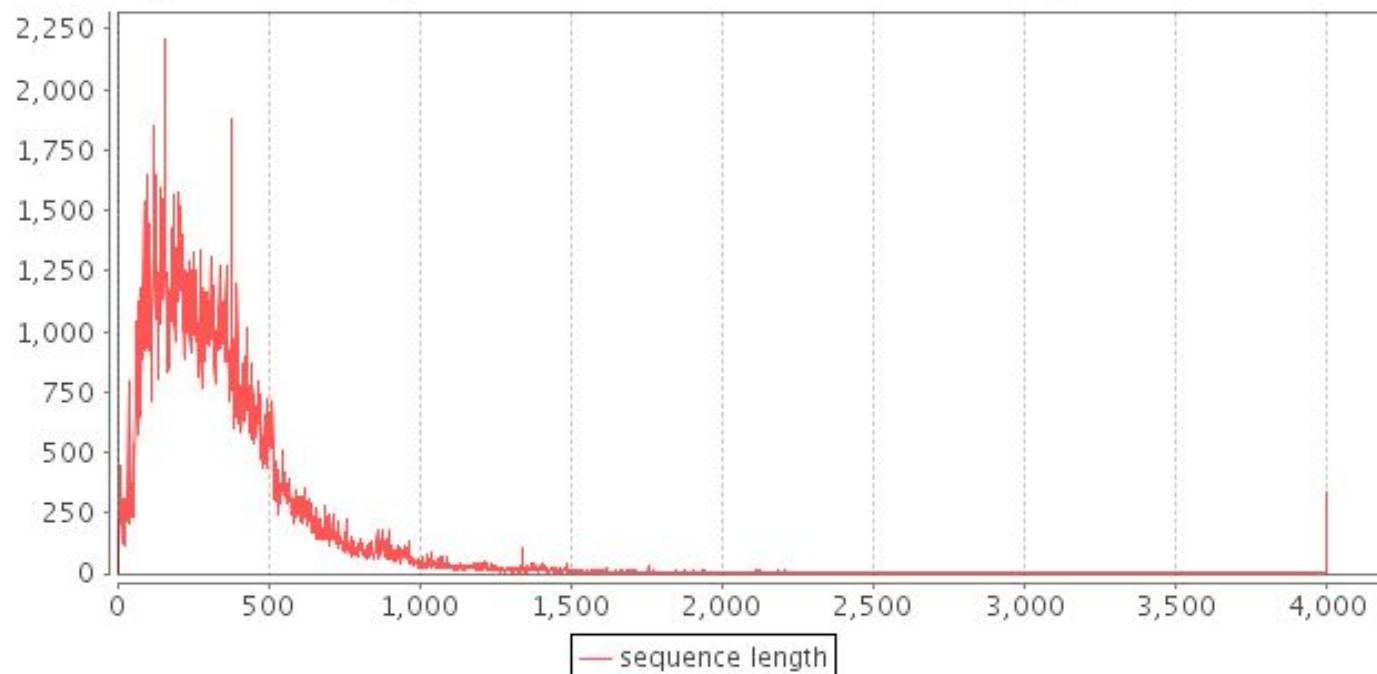
## Tamaños



## Sequence data

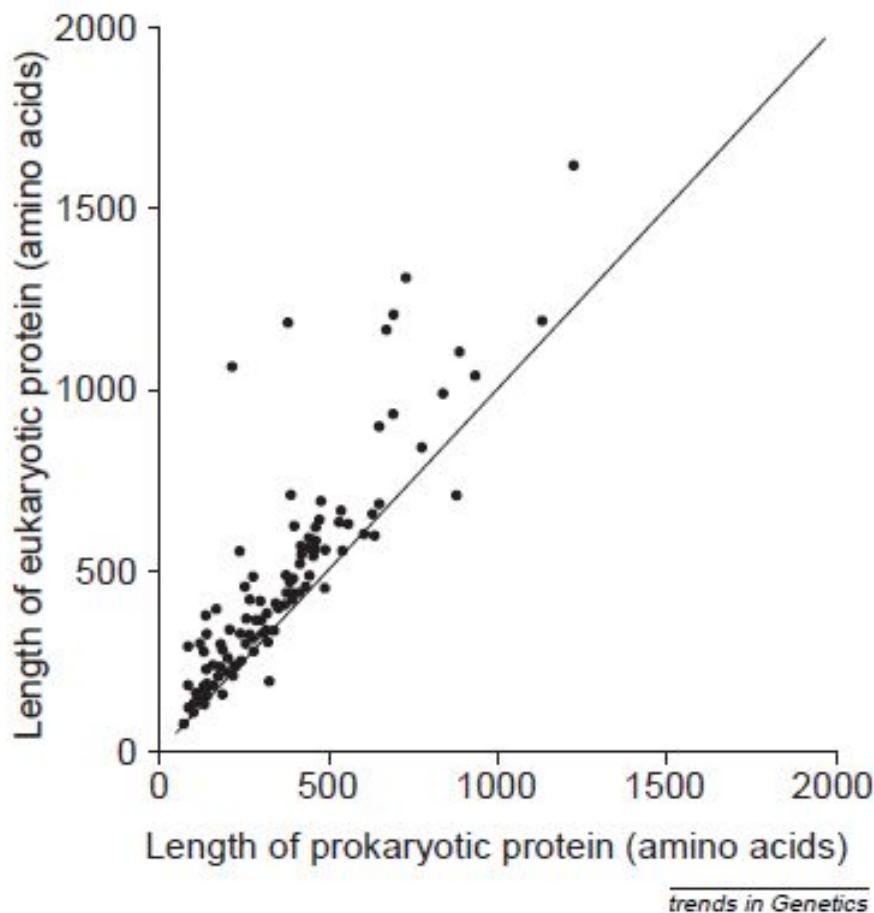
---

### Sequence length distribution in UniProtKB/Swiss-Prot

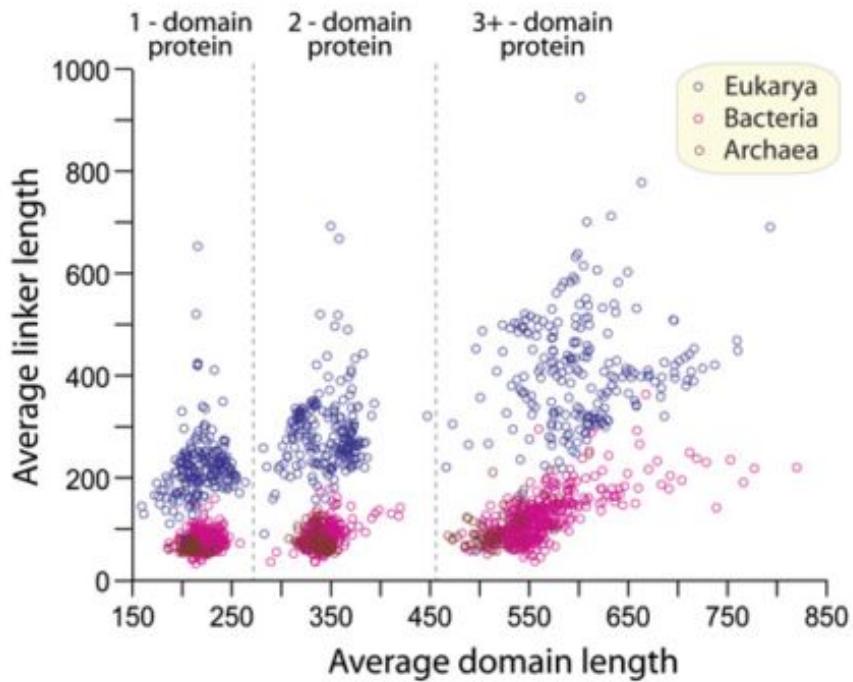


The shortest sequence is P83570 at 2 AA while the longest sequence is A2ASS6 at 35,213 AA

# Longitud en proteínas eucarióticas vs procarióticas



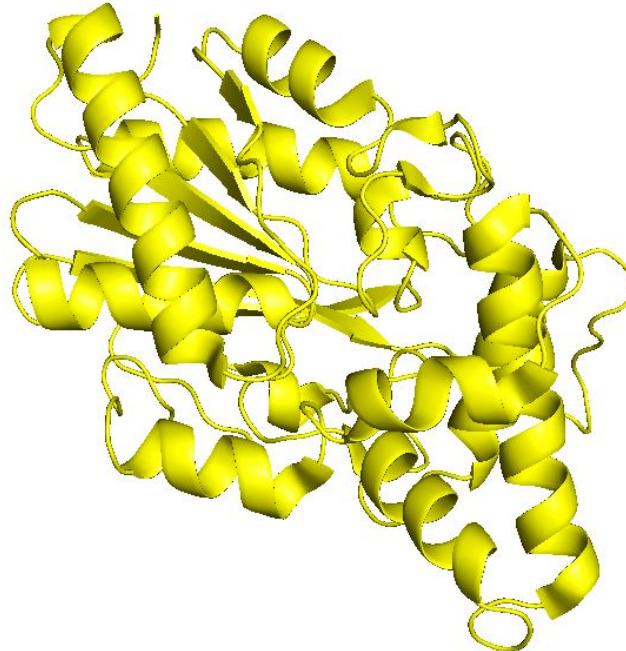
*trends in Genetics*



**Fig. 1.** Plot shows average lengths (amino acid numbers) of domains and linkers in 745 genomes, including 215 Eukarya (blue circles), 478 Bacteria (pink circles), and 52 Archaea (brown circles). Mean values of proteins with different domain numbers within the same genome could be separated well (dash lines) because of the increasing aggregate lengths.

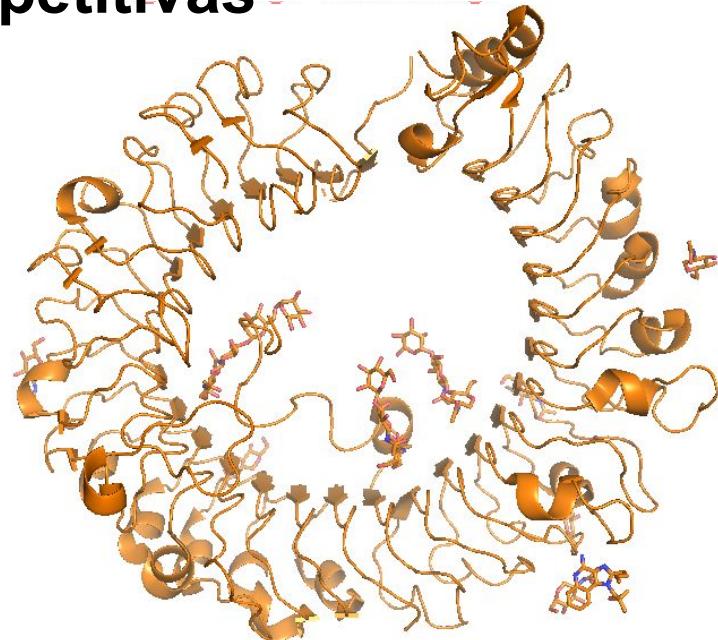
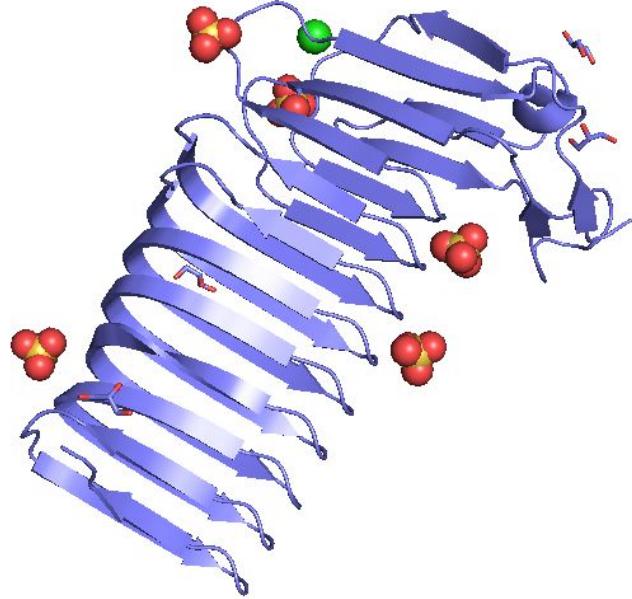
# Tipos de proteínas

## Proteínas globulares



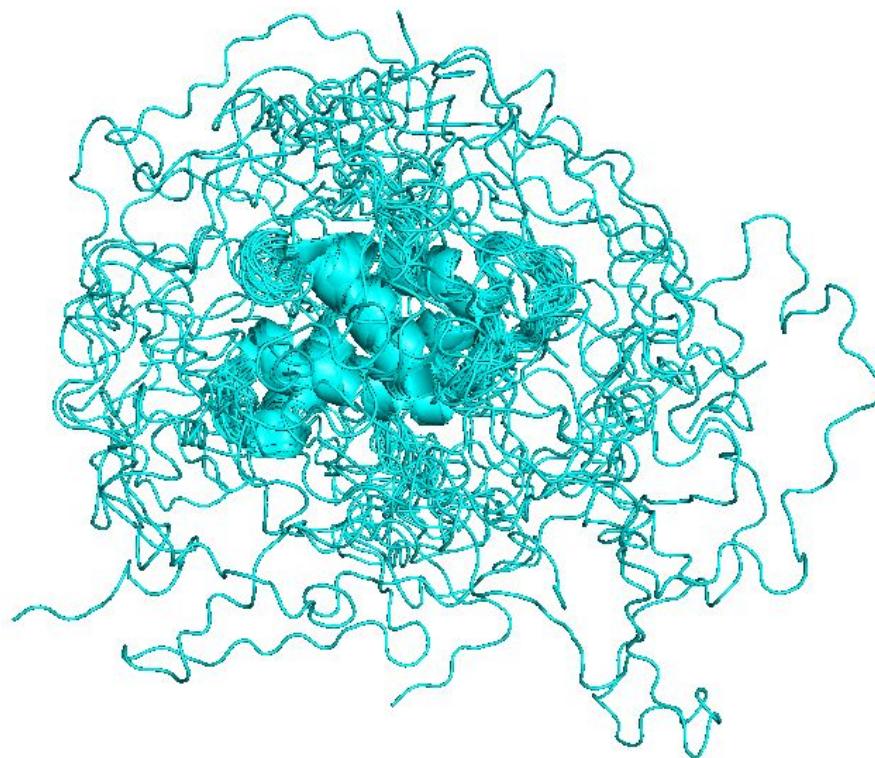
Globulares: alta proporción de estructura secundaria, importante “core” hidrofóbico, gran número de interacciones intra-moleculares

## Proteínas repetitivas



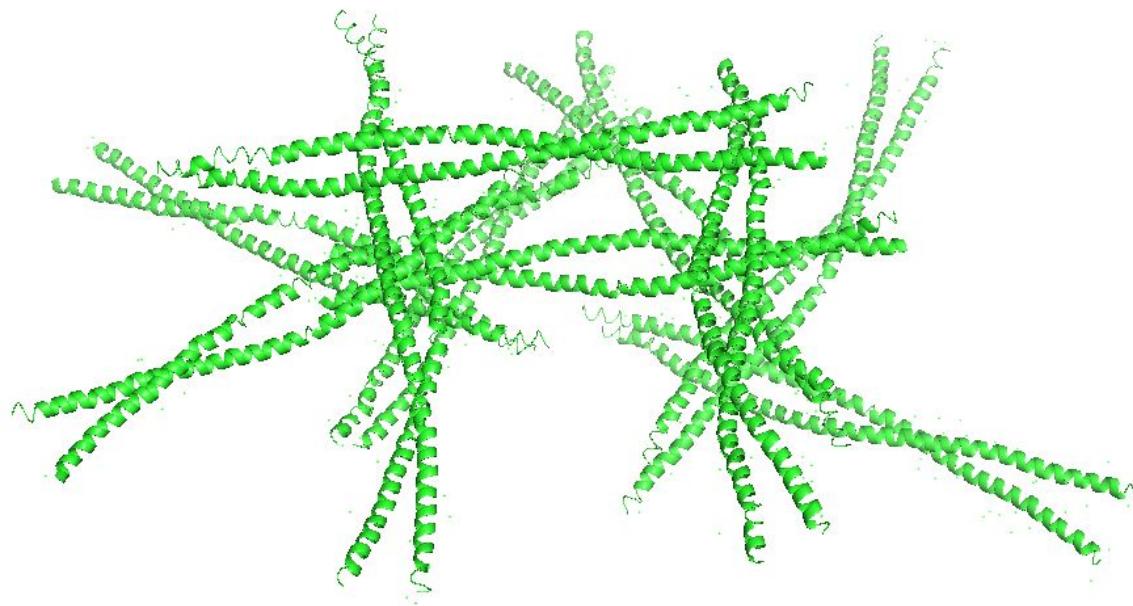
Repetitivas: poseen repeticiones internas a nivel de estructura primaria que se refleja en la estructura secundaria y terciaria.

# Proteínas desordenadas



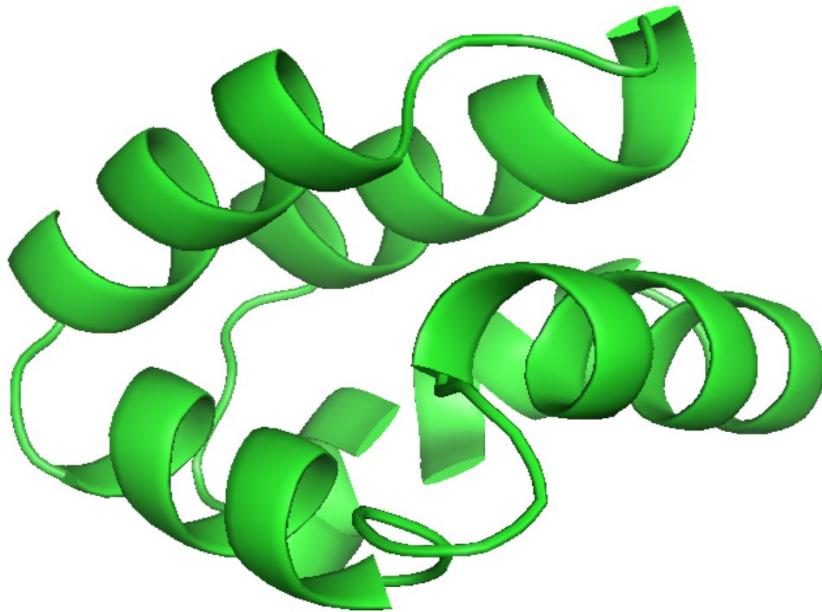
Desordenadas: no posee una estructura terciaria única ni definida. El estado nativo está formado por gran cantidad de confórmeros. Poco porcentaje de estructura secundaria, bajo contenido de contactos

# Proteínas fibrosas



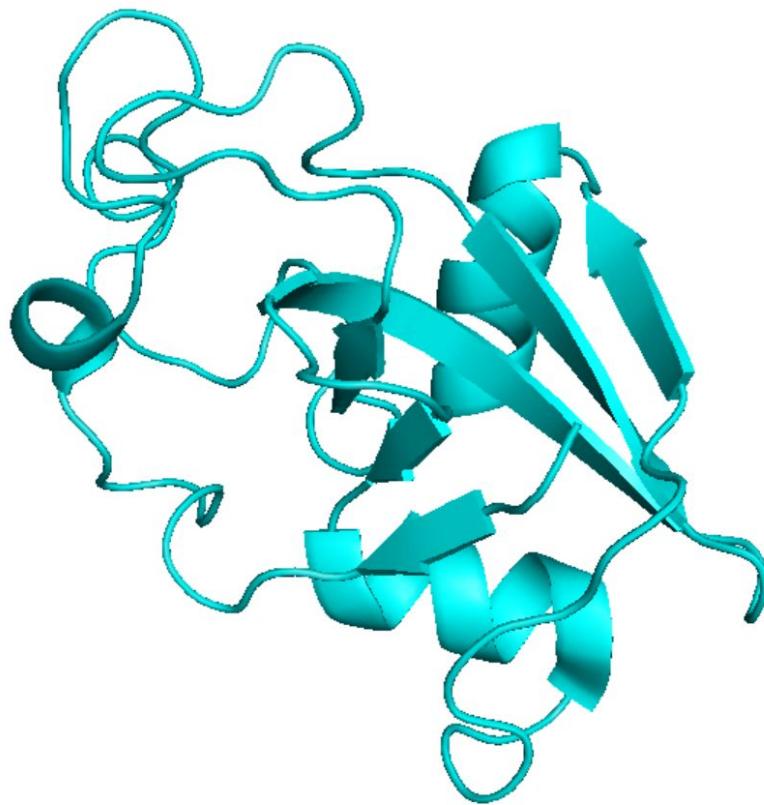
Fibrosas: proteínas que muestran estructura elongada, mayormente alfa hélices o beta plegadas. Asociadas en general a roles estructurales

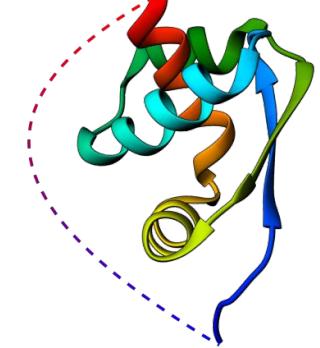
# Proteínas circulares



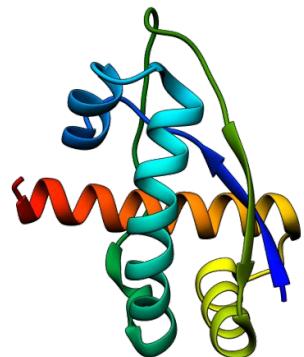
Cíclicas: Las proteínas cíclicas poseen sus terminales N y C unidos. Se encuentran en bacterias, animales y plantas

# Proteínas anudadas

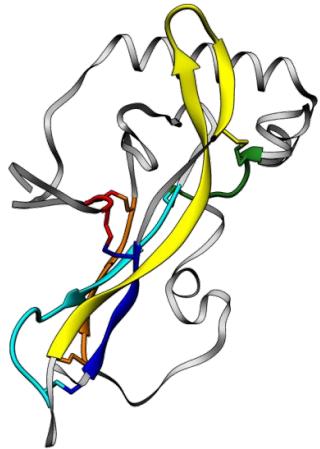




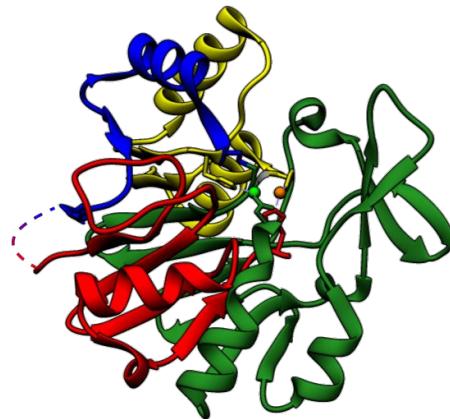
Probabilistic  
knot



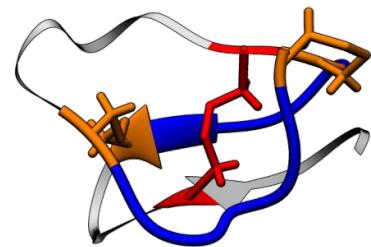
Knotoid



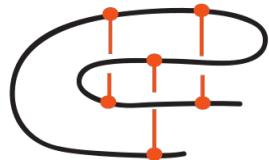
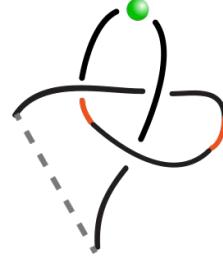
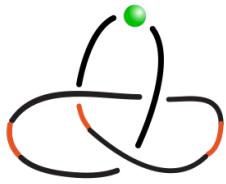
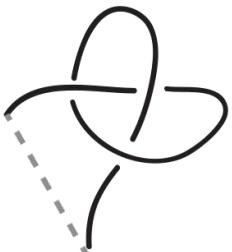
Deterministic knot  
including disulfide



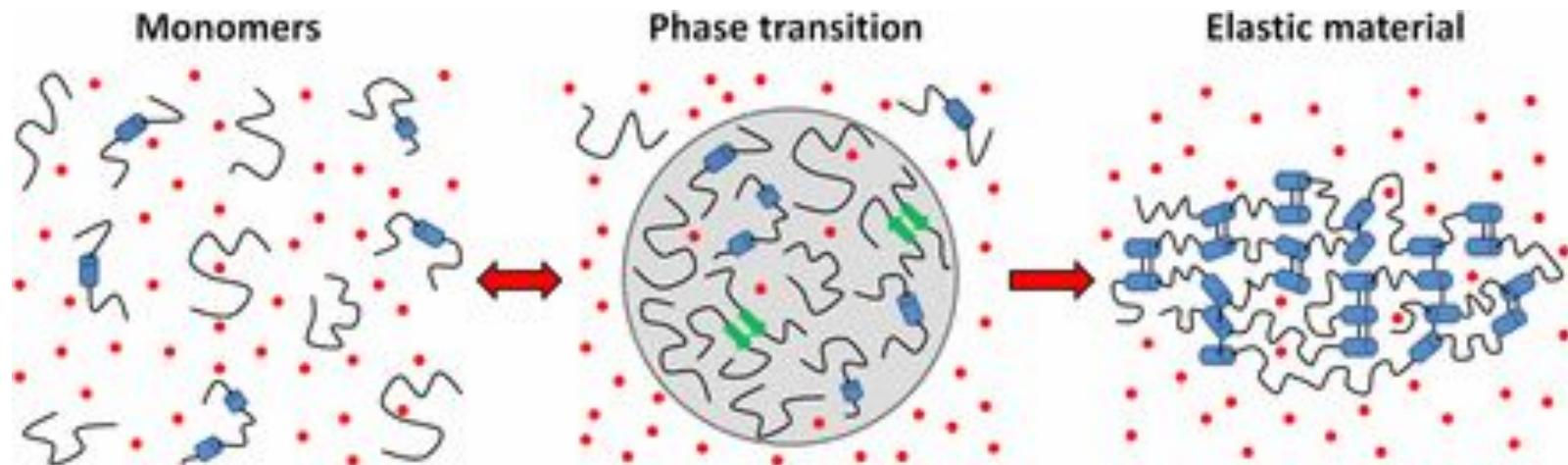
Probabilistic knot  
including disulfide



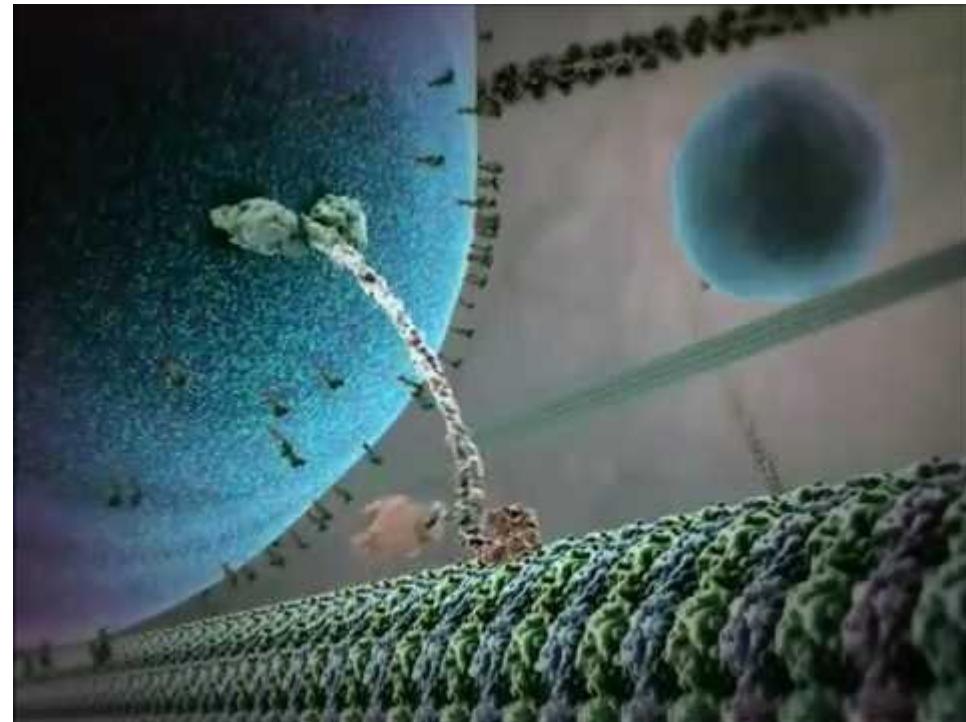
Cysteine  
"knot"



# Phase transition proteins



# Proteínas que funcionan como motores



# Proteínas de Meteoritos???

## Hemolithin: a Meteoritic Protein containing Iron and Lithium.

Malcolm. W. McGeoch<sup>1</sup>, Sergei Dikler<sup>2</sup>  
and Julie E. M. McGeoch<sup>3\*</sup>.

<sup>1</sup> PLEX Corporation, 275 Martine St., Suite 100, Fall River, MA 02723, USA.

<sup>2</sup> Bruker Scientific LLC, 40 Manning Rd, Billerica MA 01821.

<sup>3</sup> Department of Molecular and Cellular Biology, Harvard University, 52 Oxford St., Cambridge MA 02138, USA.

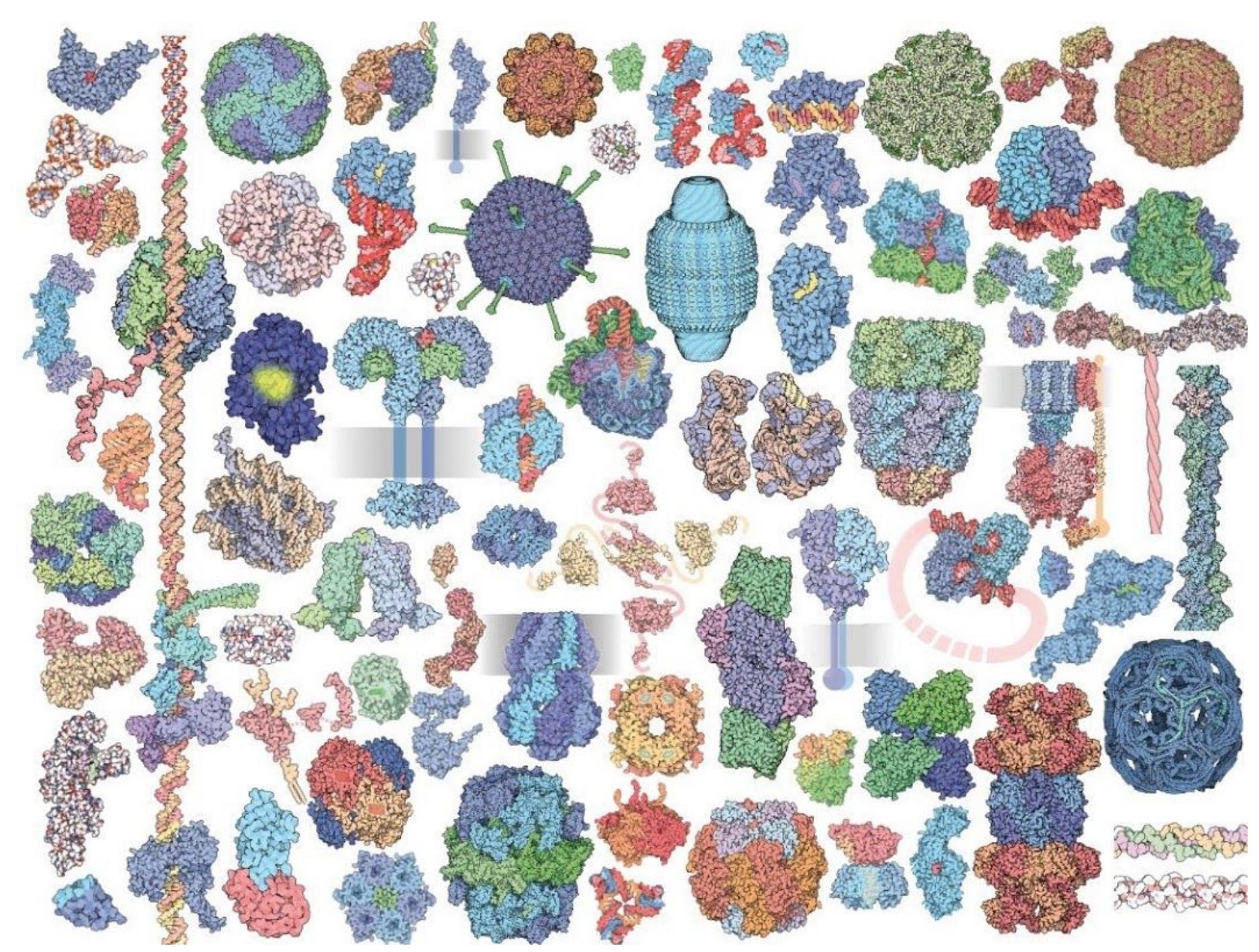
\*Corresponding author. E-mail: mcgeoch@fas.harvard.edu

### ABSTRACT

This paper characterizes the first protein to be discovered in a meteorite. Amino acid polymers previously observed in Acfer 086 and Allende meteorites [1,2] have been further characterized in Acfer 086 via high precision MALDI mass spectrometry to reveal a principal unified structure of molecular weight 2320 Daltons that involves chains of glycine and hydroxy-glycine residues terminated by iron atoms, with additional oxygen and lithium atoms. Signal-to-noise ratios up to 135 have allowed the quantification of iron and lithium in

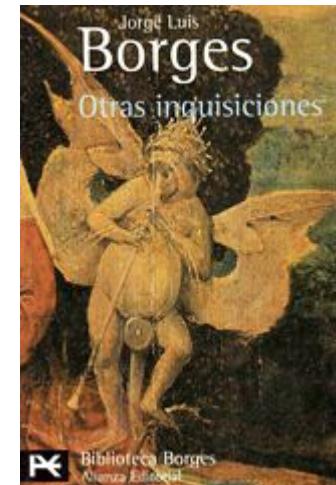
## **Dark proteins**

Las denominadas **Dark proteins** son proteínas de las cuales no se conoce su relación estructura-función. No se las puede relacionar con ninguna otra proteína o ninguna de las propiedades conocidas de las proteínas en general. Son regiones codificantes, que se expresan y traducen. Constituyen el 20-50% de los proteomas. No se sabe qué función, estructura tienen. Serían como la materia oscura del universo se sabe que existen, pero no se sabe qué son.



**“El Emporio celestial de conocimientos benévolos”** es una cierta enciclopedia china mencionada por el escritor argentino Jorge Luis Borges en el ensayo “El idioma analítico de John Wilkins” (1952) en el cual se escribe que los animales se clasifican en:

- (a) pertenecientes al Emperador,
- (b) embalsamados,
- (c) amaestrados,
- (d) lechones,
- (e) sirenas,
- (f) fabulosos,
- (g) perros sueltos,
- (h) incluidos en esta clasificación,
- (i) que se agitan como locos,
- (j) innumerables
- (k) dibujados con un pincel finísimo de pelo de camello,
- (l) etcétera,
- (m) que acaban de romper el jarrón,
- (n) que de lejos parecen moscas.



# Un poco de historia...

*Sur l'existence de la matière albumineuse dans les végétaux*

Antoine Francois de Fourcroy (1789) Annales de Chimie 3, 352.



1755-1809

Estudia 3 tipos de sustancias : albúmina, fibrina y gelatina

“A consistency often thick and flowing, insipid taste, solubility in cold water, precipitation by heat, solubility in alkalis and above all in ammonia, separating itself at the temperature of boiling water from all these liquids in which it is dissolved, passing into putrefaction without acidity, such are the properties which characterize albuminous substance”

*On the composition of some animal substances*  
Gerrit Jan Mulder (1839) Journal für praktische  
Chemie 16, 129



Se concentra en las llamadas “albuminas”: fibrina, albumina de huevo, de suero bovino y de trigo

	Fibrin	Albumin	
		v. Eiern	v. Serum
Kohlenstoff .	<b>54,56</b>	—	<b>54,48</b> — <b>54,84</b>
Wasserstoff .	<b>6,90</b>	—	<b>7,01</b> — <b>7,09</b>
Stickstoff .	<b>15,72</b>	—	<b>15,70</b> — <b>15,83</b>
Sauerstoff .	<b>22,13</b>	—	<b>22,00</b> — <b>21,23</b>
Phosphor .	<b>0,33</b>	—	<b>0,43</b> — <b>0,33</b>
Schwefel .	<b>0,36</b>	—	<b>0,38</b> — <b>0,68</b>

Fig. 1.2 Table of Mulder's analytical results, reproduced from his paper in 1838.<sup>14</sup> The formula calculated by Mulder for fibrin or egg albumin, on the basis of the content of phosphorus and sulphur, was  $C_{400}H_{620}N_{100}O_{120}P_1S_1$ . For serum albumin it was the same, but with two atoms of sulphur.

En 1838 siguiendo una sugerencia de Jacob Berzelius propone el término “proteína” para las sustancias “albuminosas”. Otra consecuencia del estudio de Mulder es que las proteínas son “macromoléculas”.

## El descubrimiento de los aminoácidos

1819 Leucina	1889 Lisina
1820 Glicina	1890 Cisteína
1846 Tirosina	1895 Arginina
1865 Serina	1896 Histidina
1866 Acido Glutamico	1901 Valina
1869 Acido Aspartico	1901 Prolina
1873 Asparagina, Glutamina	1901 Triptofano
1875 Alanina	1903 Isoleucina
1881 Fenilalanina	1922 Metionina
	1936 Treonina

# El enlace peptídico

22 de Septiembre de 1902 en el 74th Annual Meeting of the Gesellschaft der deutschen Naturforschen und Ärzte (Society of German Naturalists and Physicians) in Karlsbad, Bohemia

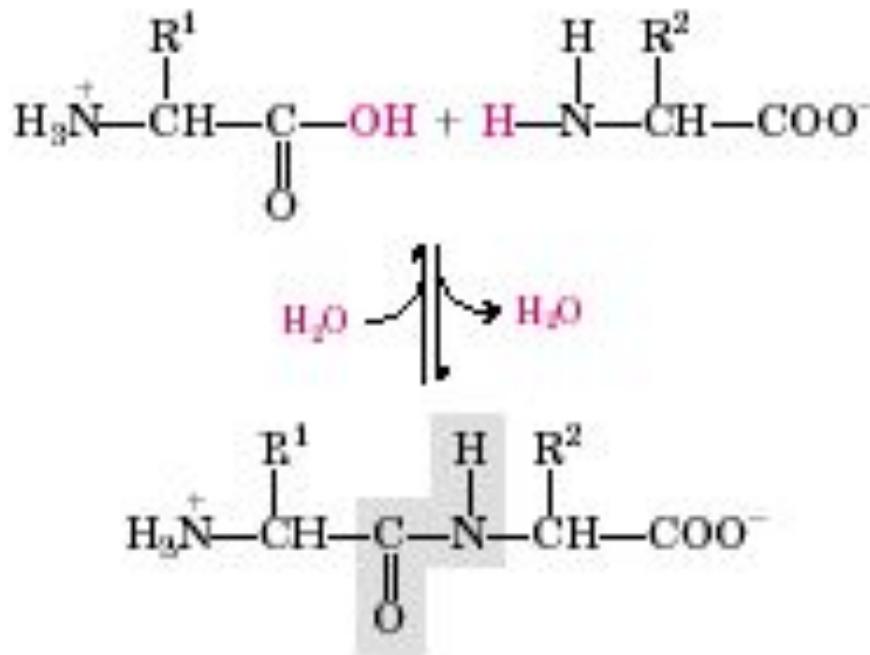


Franz Hofmeister  
1850-1922



Hermann Emil Fischer  
1852-1919

## Indicios acerca de la existencia del enlace peptídico



- (1) La mayoría del N de las proteínas está unido como amina/imina
- (2) Las proteínas dan la reacción de Biuret positiva
- (3) Los aminoácidos pueden unirse en forma sintética para dar oligopéptidos formando un enlace amida
- (4) Las proteínas son hidrolizadas en ácidos fuertes (enlace amida) y por “fermentos”

My entire yearning is directed toward the first synthetic enzyme. If its preparation falls into my lap with the synthesis of a natural protein material, I Will consider my misión fulfilled

Emil Flscher

# On December 12, 1902, Emil Fischer delivered his Nobel Prize lecture in Stockholm, Sweden

*Of the chemical aids in the living organism the ferments—mostly referred to nowadays as enzymes—are so pre-eminent that they may justifiably be claimed to be involved in most of the chemical transformations in the living cell. The examination of the synthetic glucosides has shown that the action of the enzymes depends to a large extent on **the geometrical structure of the molecule to be attacked, that the two must match like lock and key.** Consequently, with their aid, the organism is capable of performing highly specific chemical transformations which can never be accomplished with the customary agents. To equal Nature here, the same means have to be applied, and I therefore foresee the day when physiological chemistry will not only make extensive use of the natural enzymes as agents, **but when it will also prepare synthetic ferment**s for its purposes.*

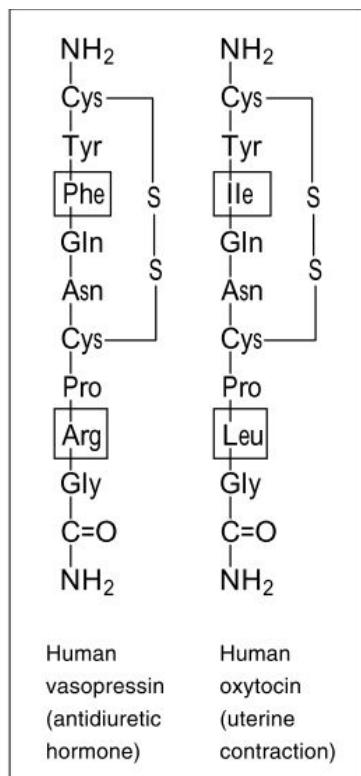
# Fisher y el inicio de la síntesis química de péptidos y proteínas:

1901: glicil-glicina

1906. Reporta la síntesis de un octapéptido

L-Leu-(Gly)3-L-Leu-(Gly)3-L-Leu- (Gly)8-Gly

1907: 15 gly y 3 leu



# Las proteínas son *Macromoléculas* o *agregados*?

Teoría del coloide (1880-1930)

“But also the fundamental chemical and physical investigation of colloids is absolutely essential for the further development of the individual sciences themselves, particularly geology, biology, physiology and botany. The entire life of the cells (in the animal and plant realms) and inorganic nature meet in large part at the interactions between colloids”. Wolfgang Ostwald 1907

# Primeras intuiciones sobre el plegado de las proteínas



*“Amino acids per se and their spatial arrangement within the protein must become the chemical key to understanding of proteins”* Kossel, 1900.

**Ludwig Karl Martin Leonhard Albrecht Kossel**  
**1853-1927**

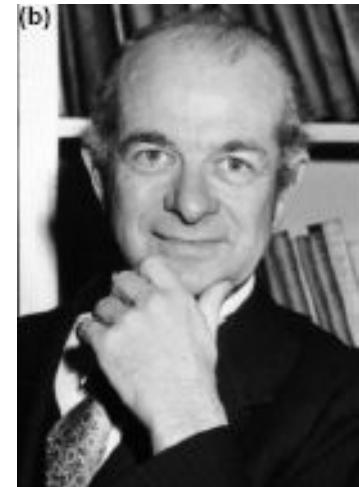
*ON THE STRUCTURE OF NATIVE, DENATURED, AND  
COAGULATED PROTEINS*

BY A. E. MIRSKY\* AND LINUS PAULING

GATES CHEMICAL LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA,  
CALIFORNIA

Communicated June 1, 1936

**“Our conception of a native protein molecule (showing specific properties) is the following. The molecule consists of one polypeptide chain which continues without interruption throughout the molecule (or, in certain cases, of two or more such chains), this chain is folded into a uniquely defined configuration, in which it is held by hydrogen bonds between the peptide nitrogen and oxygen atoms and also between the free amino and carboxyl groups of the diamino and dicarboxyl amino acids residues”**



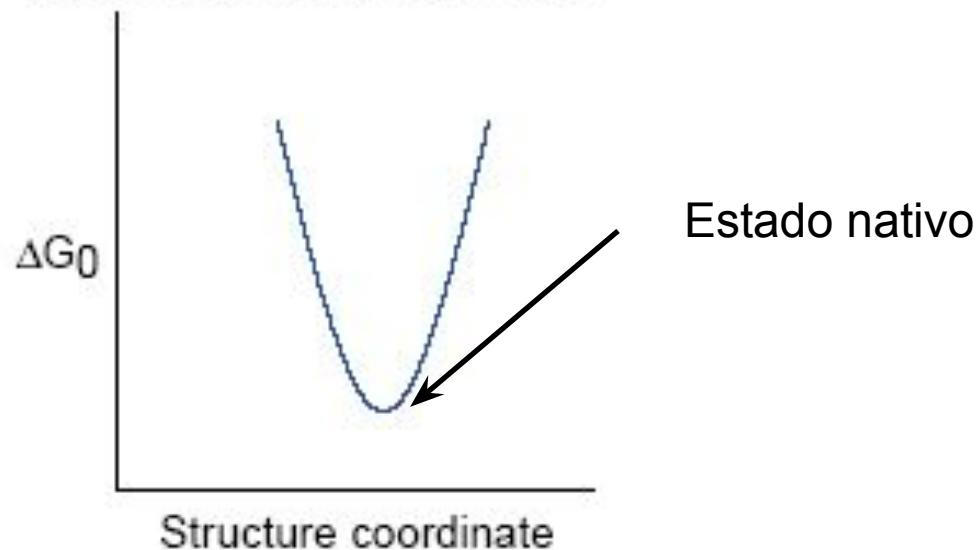
**“The characteristic specific properties of native proteins we attribute to their uniquely defined configurations”**

**“The denature protein molecule we consider to be characterized by the absence of a uniquely defined configuration”**

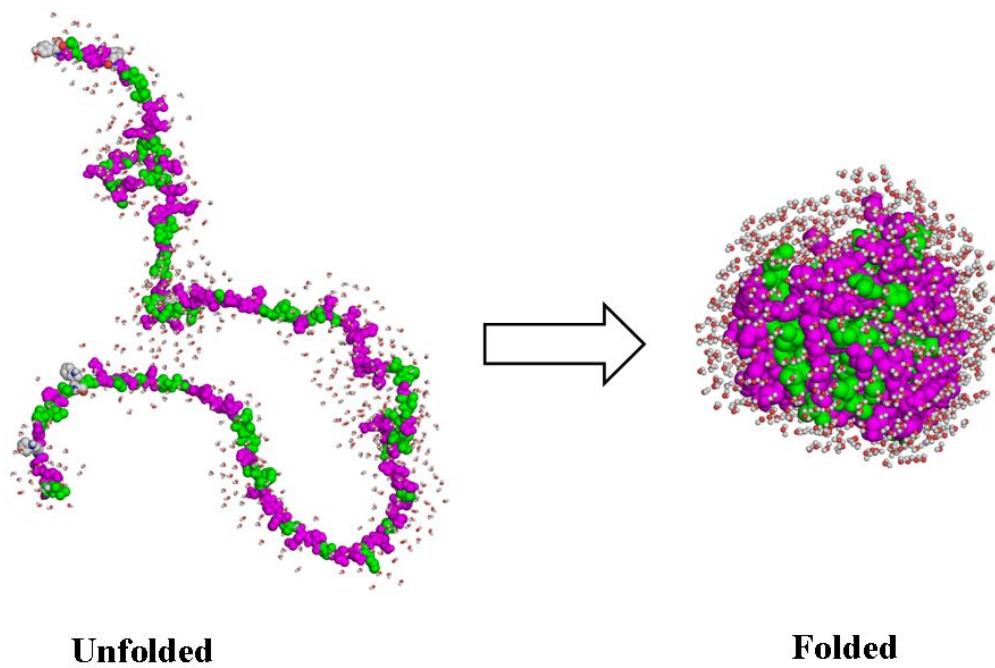
L Pauling

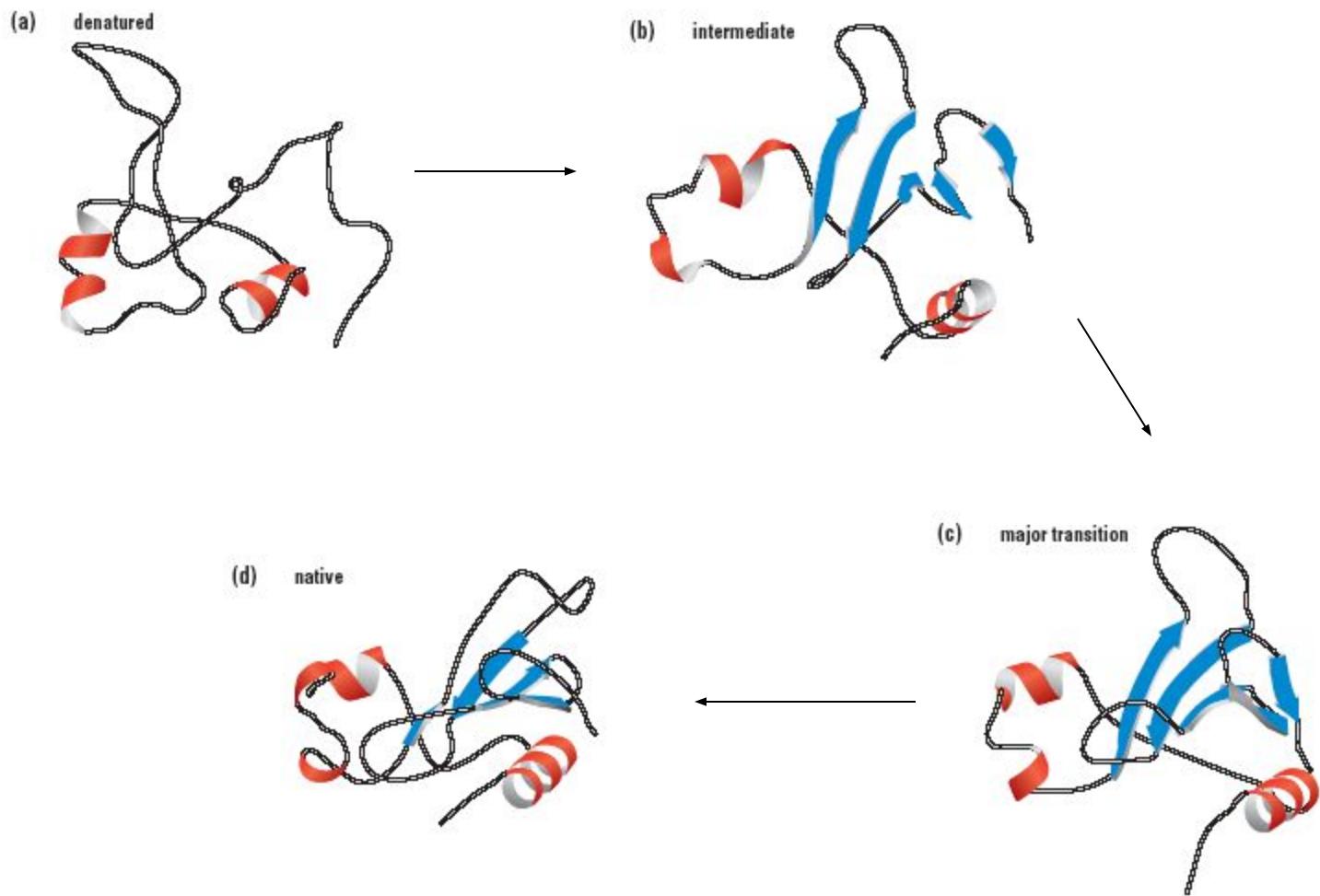
(a) ‘Simplistic view’

(i) Smooth energy landscape



Las proteínas para funcionar se pliegan (la mayoría!)



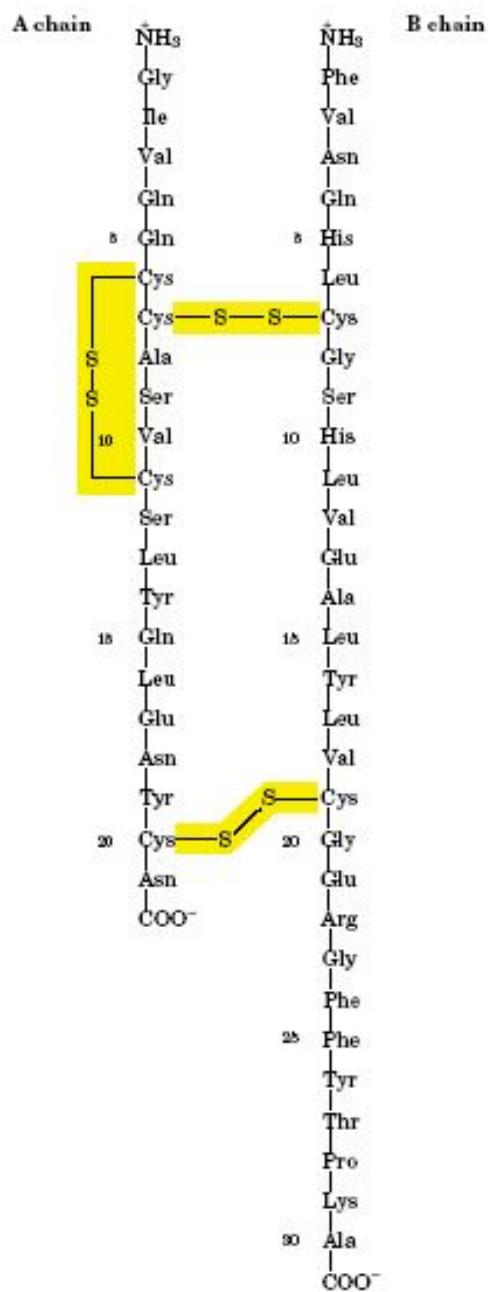


# Secuenciación: Estructura primaria



Frederick Sanger  
1918-2013

Nobel Prize (Chemistry, 1958)  
Nobel Prize (Chemistry, 1980)





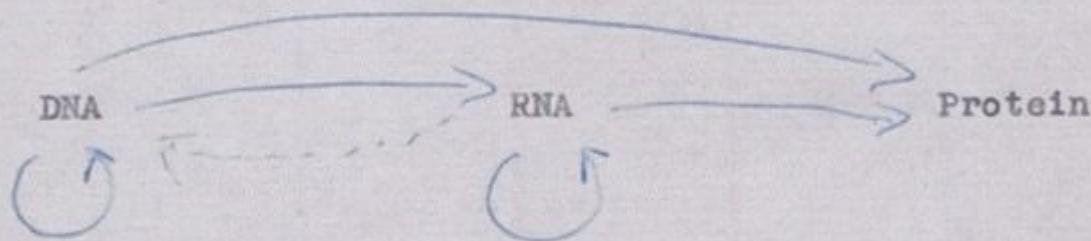
"The more likely hypothesis is that the folding is simply a function of the order of amino acids, provided that it take place as the newly formed chain comes off the template"

F H C Crick, On Protein Synthesis, Symp. Soc. Exptl. Biol.,  
Vol.12,  
pp.138–163, 1958.

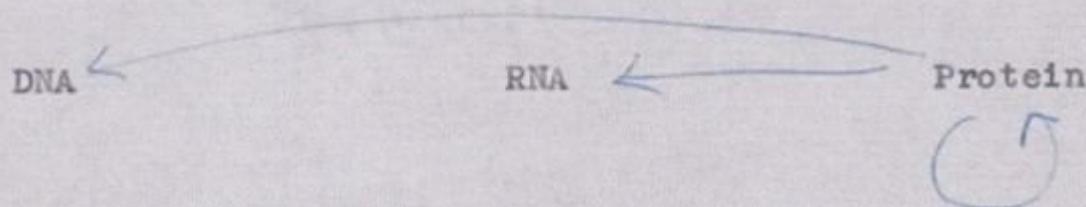
# Dogma central de la biología

The Central Dogma: "Once information has got into a protein it can't get out again". Information here means the sequence of the amino acid residues, or other sequences related to it.

That is, we may be able to have



but never



where the arrows show the transfer of information.

# Experimento de Anfinsen (1956-1957)

## Teoría “termodinámica”



Christian B. Anfinsen  
(1916-1995)

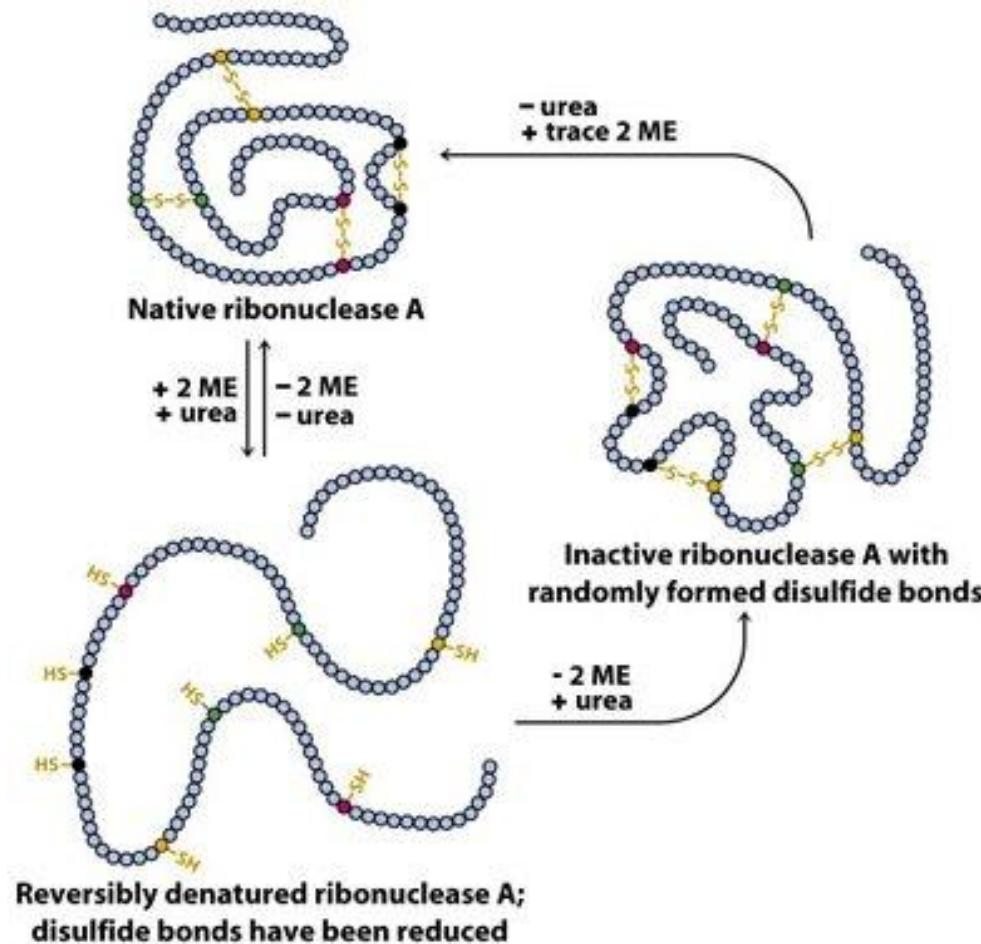
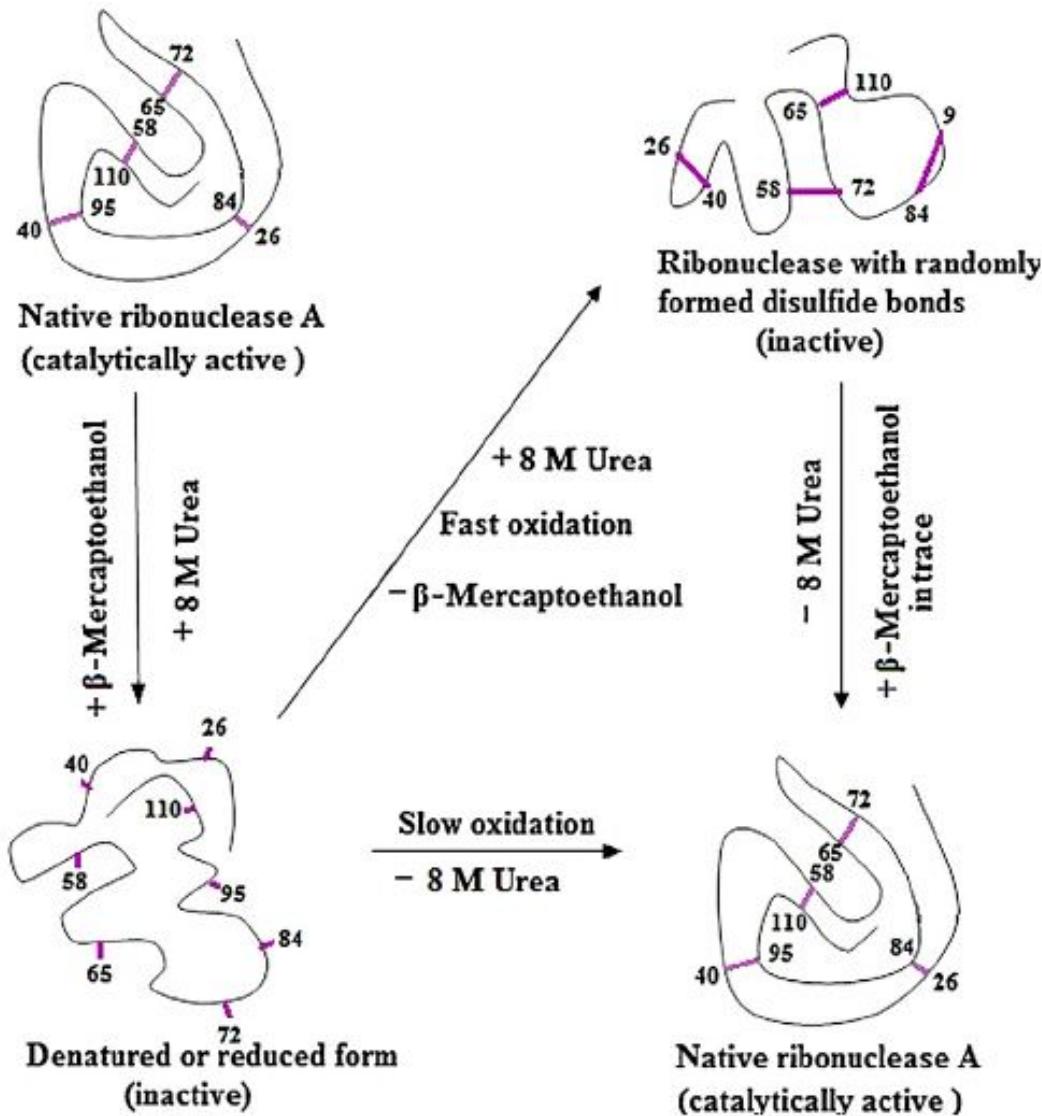


Figure 4-29 Principles of Biochemistry, 4/e  
© 2006 Pearson Prentice Hall, Inc.



# Hipótesis sobre la renaturalización

- a. Los 4 puentes disulfuros nativos se formaban relativamente rápido y en un solo paso
- b. Los 4 puentes disulfuros nativos se formaban secuencialmente, dando lugar a intermediarios inactivos (con 1 puente, 2 puentes, etc)
- c. Los 4 puentes disulfuros nativos se formaban siguiendo a o b pero la proteína debía acomodar además su estructura secundaria para alcanzar su estado nativo
- d. Los puentes disulfuro se formaban al azar (en distintas combinaciones de las 8 Cys intervenientes) dando intermediarios inactivos (con puentes disulfuro no-nativos). Sin embargo, en presencia de trazas de beta mercaptoetanol, los puentes se rompían y volvían a formarse. En el tiempo, la enorme mayoría de los puentes disulfuro eran nativos

THE KINETICS OF FORMATION OF NATIVE RIBONUCLEASE  
DURING OXIDATION OF THE REDUCED POLYPEPTIDE CHAIN

By C. B. ANFINSEN, E. HABER,\* M. SELA,† AND F. H. WHITE, JR.

LABORATORY OF CELLULAR PHYSIOLOGY AND METABOLISM, NATIONAL HEART INSTITUTE,  
NATIONAL INSTITUTES OF HEALTH

Communicated by John T. Edsall, July 31, 1961

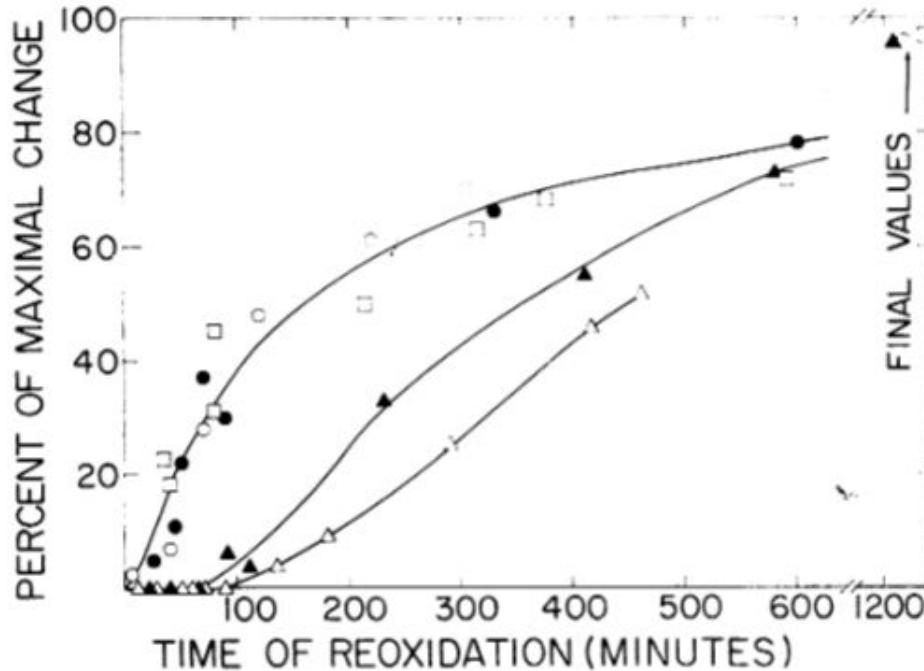


FIG. 1.—Changes, during the oxidation of reduced ribonuclease, in SH groups as followed by titration with *p*-chloromercuribenzoate (●) and by reaction with radioactive iodoacetate (○), in optical rotation (□), and in enzymatic activity as measured against ribonucleic acid (▲) and against uridylic-2',3'-cyclic phosphate (△).

## Reductive Cleavage of Disulfide Bridges in Ribonuclease

Michael Sela, Frederick H. White, Jr., and Christian B. Anfinsen

Laboratory of Cellular Physiology and Metabolism, National Heart Institute,  
National Institutes of Health, Bethesda, Maryland

"When completely reduced and fully inactivated enzyme (1mg/ml) was subjected to oxidation by air bubbling at room temperature, for 68 hours, 0.01M phosphate, pH7>8, ribonuclease activity reappeared to the extent of circa 12-19% of specific activity"

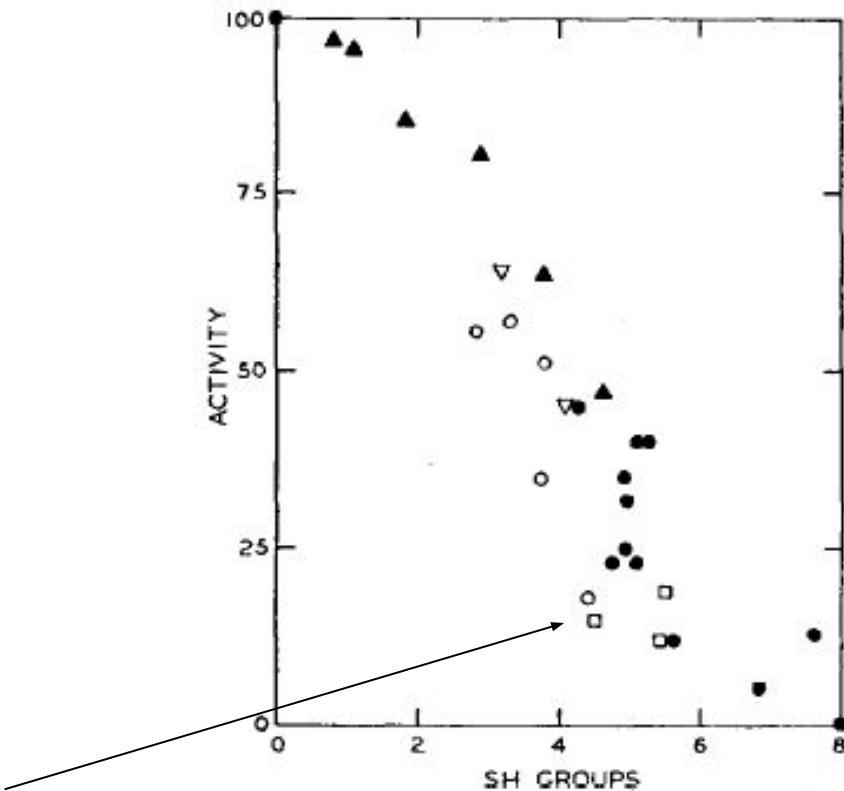


Fig. 1. Activity of ribonuclease at various stages of reduction (expressed as percentage of the specific activity of native ribonuclease) as a function of the number of moles of sulphhydryl per mole of enzyme. ▲, Reduction in absence of urea; ●, reduction in 8M urea; □, reoxidation of fully reduced, inactive ribonuclease; ○, reoxidation of samples containing more than six sulphhydryl groups per average molecule; ▽, reoxidation of samples containing about four sulphhydryl groups per average molecule.

## Regeneration of Native Secondary and Tertiary Structures by Air Oxidation of Reduced Ribonuclease

FREDERICK H. WHITE, JR.

Laboratory of Cellular Physiology and Metabolism, National Heart Institute, National Institutes of Health, United States Public Health Service, Bethesda 14, Maryland

WITH A NOTE BY J. BELLO, D. HARKER AND E. DE JARNETTE

1356

(Received for publication, October 31, 1960)

Vol. 236, No. 5

reduction of RNases A and B. Therefore, the increase in B component is the result of its higher yield on oxidation.

Chromatography of reduced RNase was not considered feasible because of the possibility of oxidation on the column. Reduced CAM RNase was eluted from the column as shown in Fig. 1. Two maxima were produced which appeared to coincide with those of the A and B components of native and Reox RNases, although a considerable spreading occurred. Reduced CM RNase, not shown in the figure, came off of the column as a single peak near the front, before gradient elution was started.

Preliminary experiments (23) revealed the activity of Reox RNase A to be 80 to 100% of the specific activity of native RNase A. In a further investigation, the activities of Reox RNases A and B were compared with that of RNase A, derived from an oxidation control experiment which was performed by subjecting native RNase to the same conditions of air oxidation, desalting, and chromatography as for reduced RNase. With ribonucleic acid, uridine-2',3'-cyclic phosphate, and cytidine-2',3'-cyclic phosphate as substrates, the activities of the Reox RNases were the same as that of the control RNase A within an experimental error of  $\pm 10\%$ .

The isoionic points of these peaks were determined by measurement of the pH values of their solutions after deionization on mixed bed columns of Dowex 50 and Nalcite SAR-10. For the

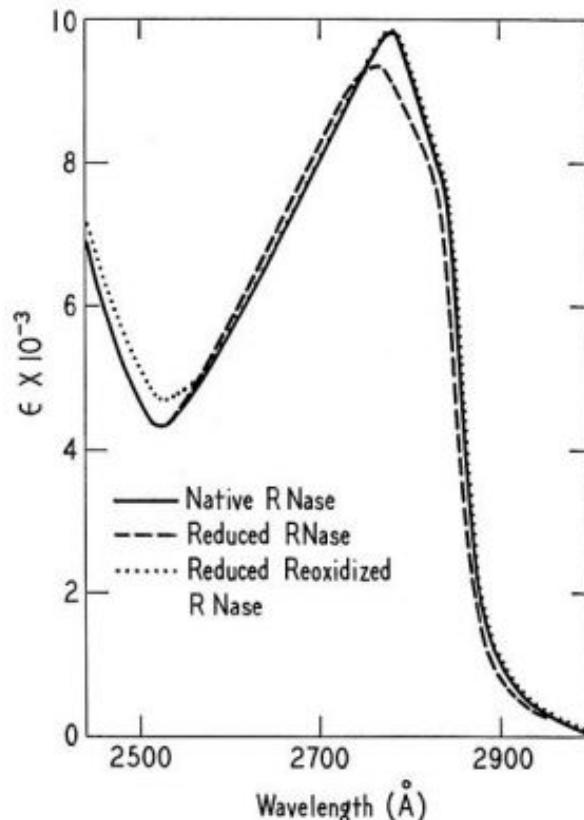
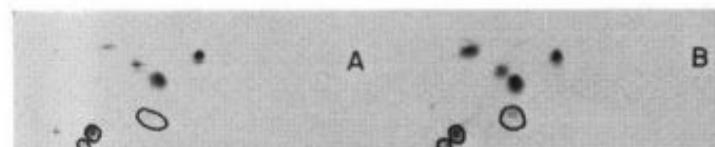


FIG. 3. Ultraviolet spectra of native RNase A, reduced RNase A, and Reox RNase A.

## Principles that Govern the Folding of Protein Chains

Christian B. Anfinsen

exists under conditions similar to those for which it was selected—the so-called physiological state.

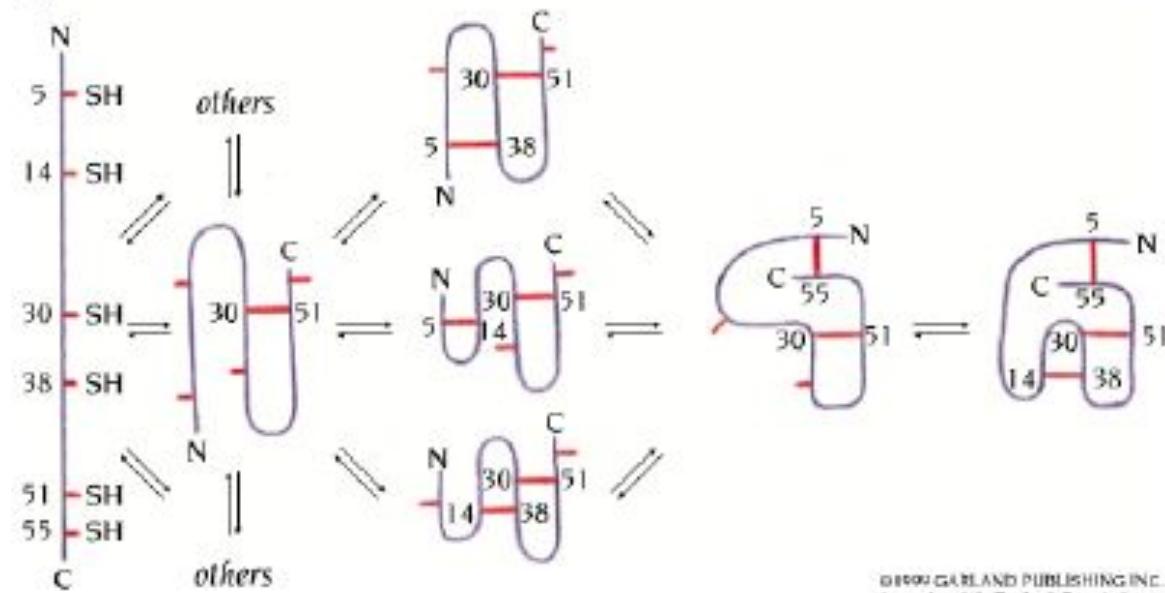
After several years of study on ribonuclease molecule it became clear to us, and to many others in the field of protein conformation, that proteins devoid of restrictive disulfide bonds and other covalent cross-linkages would make more convenient models for study of the thermodynamic and kinetic aspects of the nucleation, and subsequent pathways, of polypeptide chain folding.

The three-dimensional structure of a native protein in its normal physiological milieu (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and others) is the one in which the Gibbs free energy of the whole system is lowest; that is, that the native conformation is determined by the totality of interatomic interactions, and hence by the amino acid sequence, in a given environment....In terms of natural selection through the 'design' of macromolecules during evolution, this idea emphasize[s] the fact that a protein molecule only makes stable, structural sense when it exists under conditions similar to those for which it was selected—the so-called physiological state. (Anfinsen 1973)

## Derivaciones de los experimentos de Anfinsen

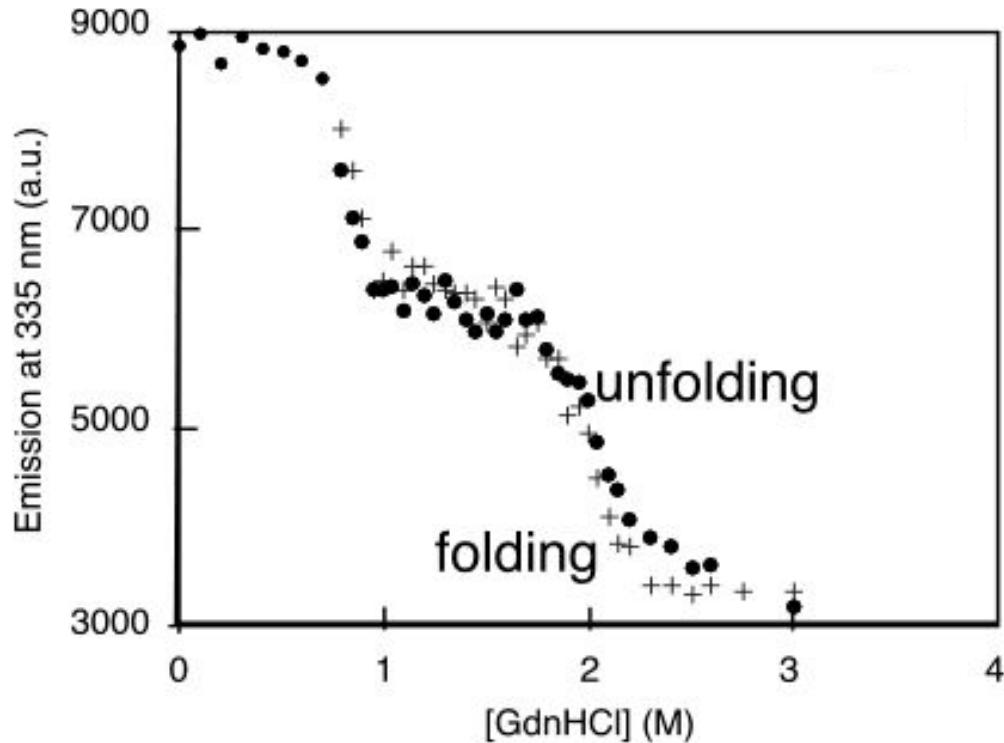
- 1 Las proteínas se pueden plegar y desplegar repetidas veces
- 2 La “fuerza impulsora” es termodinámica
- 3 Toda la información para el plegado de una proteína está en su secuencia
- 4 El estado nativo de las proteínas está en un mínimo de energía

(b)



©1997 GARLAND PUBLISHING INC.  
A member of the Taylor & Francis Group

# 1 Las proteínas se pueden plegar y desplegar repetidas veces



## 2 La “fuerza impulsora” es termodinámica

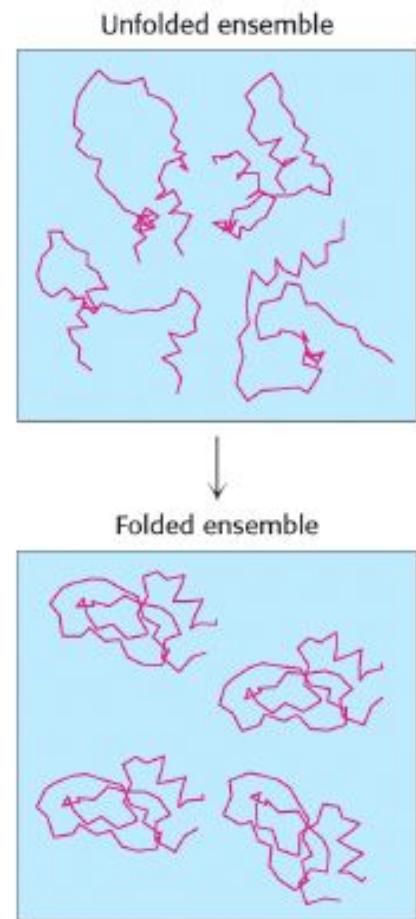
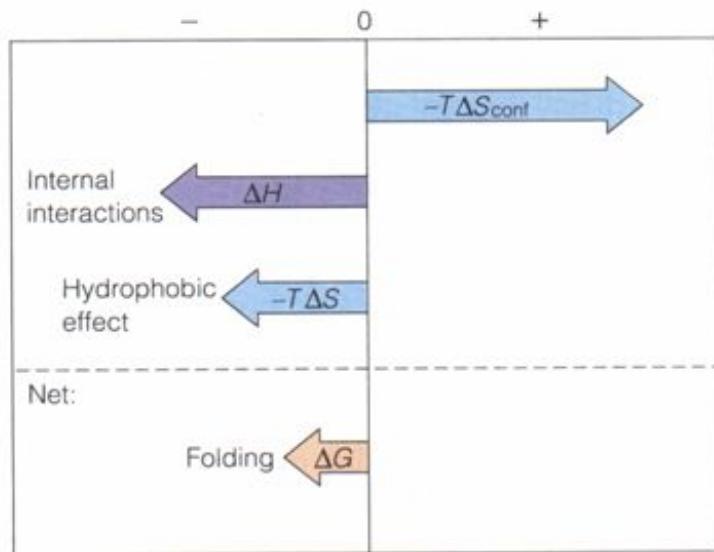
$$\Delta G = \Delta H - T\Delta S$$

**H** q-q, dipolo-dipolo, q-dipolo, q-dipolo inducido, dipolo-dipolo inducido, dipolo inducido-dipolo inducido, cation-π, π-π

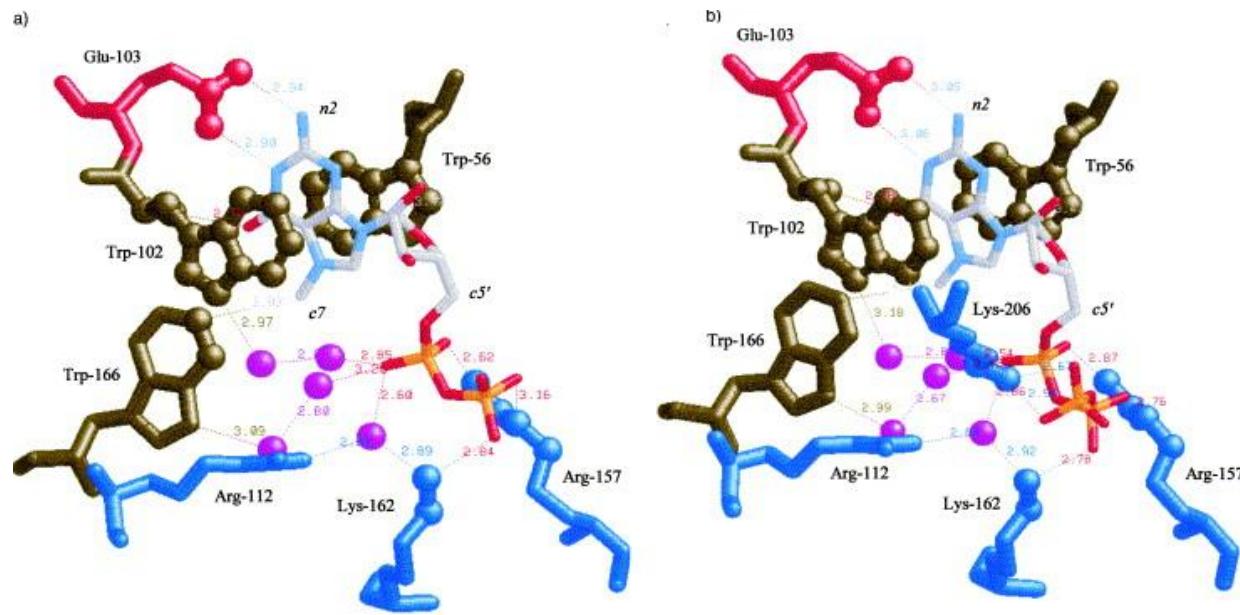
**S** Efecto hidrofóbico

**“Biology is dominated by the chemistry of the noncovalent bond”**

Alan Fersht, 1992



# Los contactos intermoleculares son clave para estabilizar a la proteína



# Plegamiento de proteínas

## Teoría cinética

*Para describir la posición de un átomo en el espacio uno necesita definir 3 coordenadas*

*Para una proteína de 100 aminoácidos necesitamos 6000 coordenadas (consideramos que un AA contiene en promedio 20 átomos)*

*Si cada una de esas coordenadas puede tomar al menos 2 valores distintos tendríamos que el espacio conformacional de la proteína sería de  $2^{6000}$*

*Si cada transición entre cada una de éstas conformaciones distintas requiere de un tiempo de por ejemplo  $110^{-13}$  segundos....*

*...plegar una proteína requeriría*

*$2^{5987}$  segundos!!!*

O sea la edad del universo!!!!

“Paradoja de Levintal”

*Extrait du Journal de Chimie Physique*, 1968, **65** n° 1, p. 44.

## ARE THERE PATHWAYS FOR PROTEIN FOLDING ?

by CYRUS LEVINTHAL

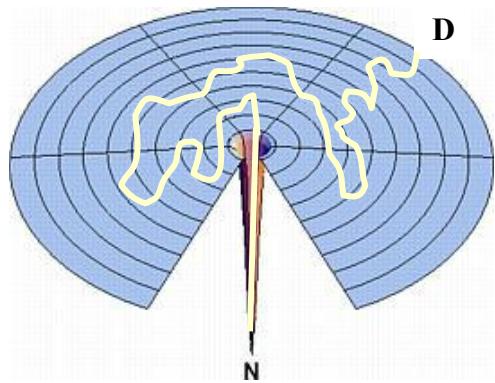
[*Massachusetts Institute of Technology, Department of Biology Cambridge, Massachusetts.*]

---

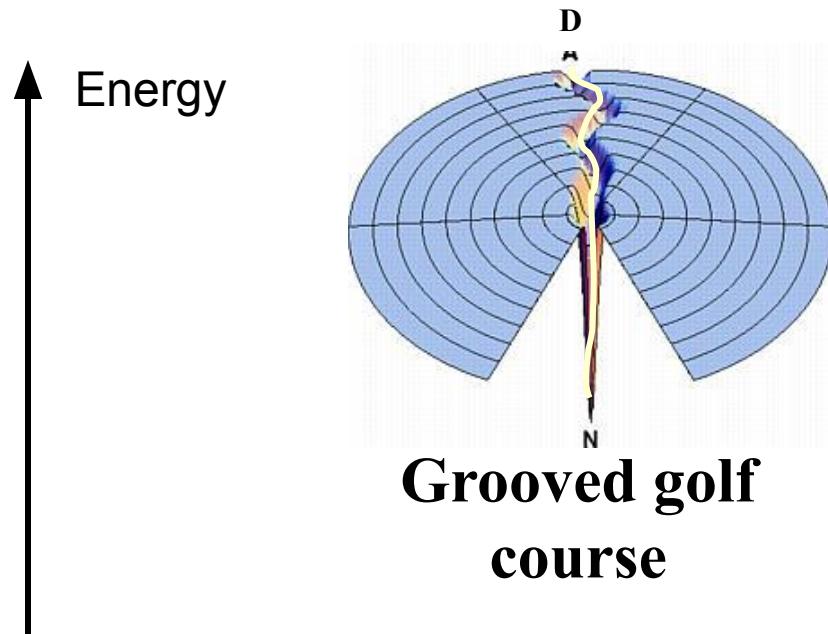
### SUMMARY

sation more likely. Thus, a pathway of folding means that there exist a well-defined sequence of events

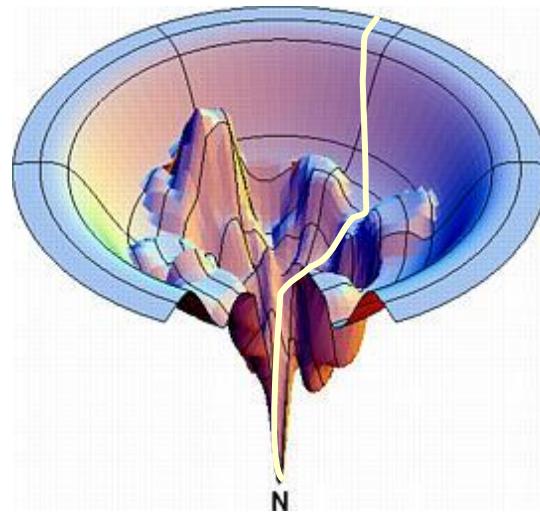
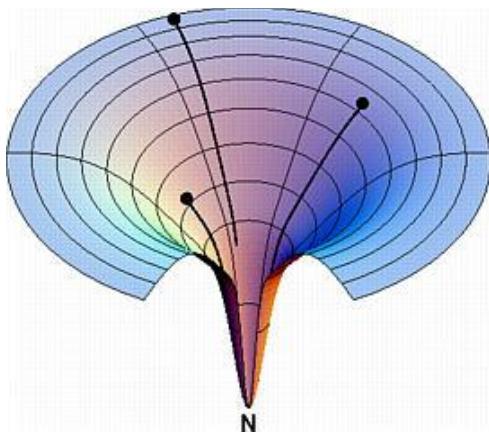
# El problema del golfista ciego



Paradoja de levinthal:  
el campo de golf plano



Paradoja de levinthal:  
El campo de golf con canales o caminos



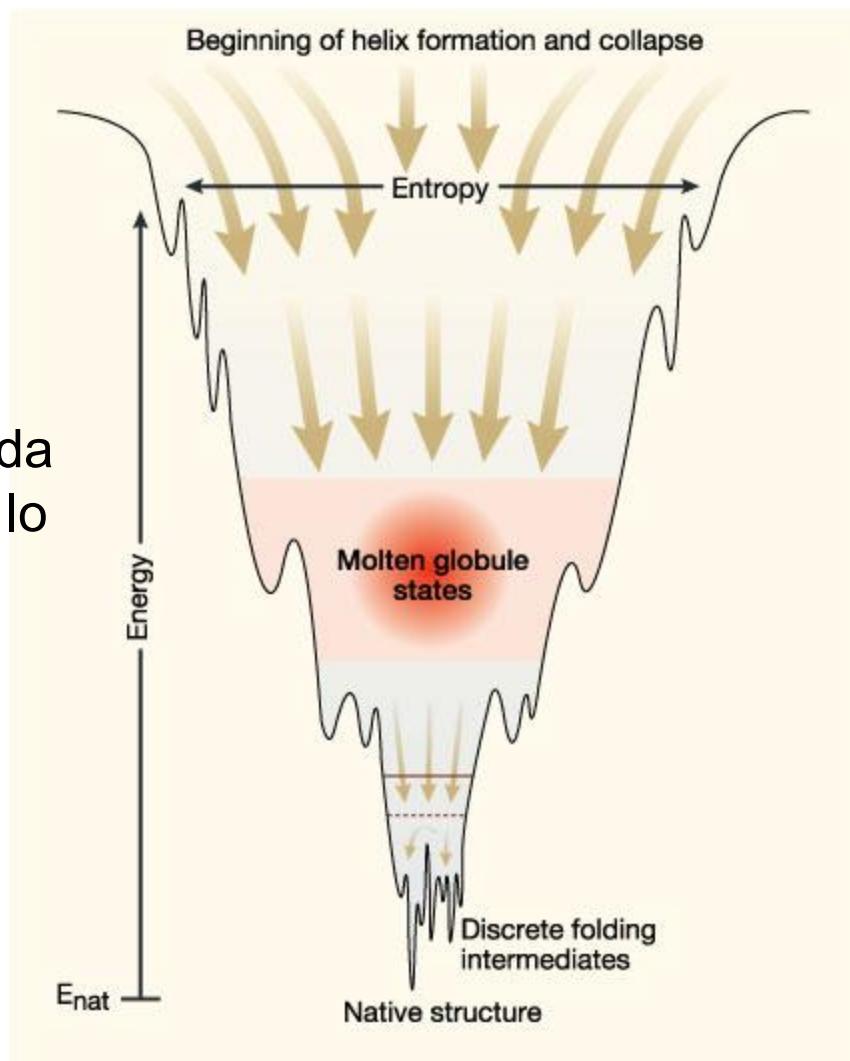
Solución del camino.

El fondo del embudo es el mínimo termodinámico.

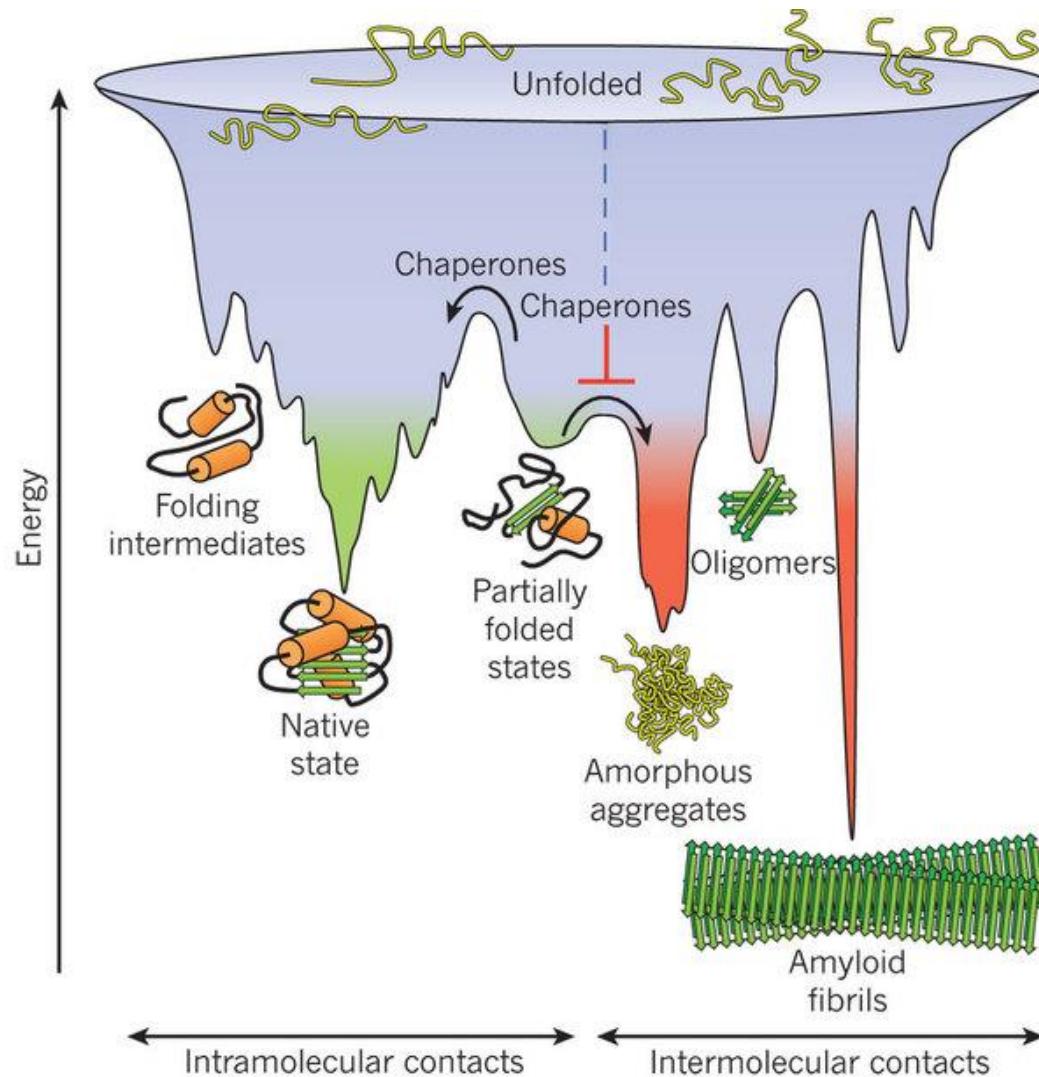
Varios diferentes caminos cinéticos alcanzan el fondo

Más realístico, embudo áspero

El espacio conformacional se reduce cada vez que se produce un buen contacto o lo que se llama un contacto “nativo”



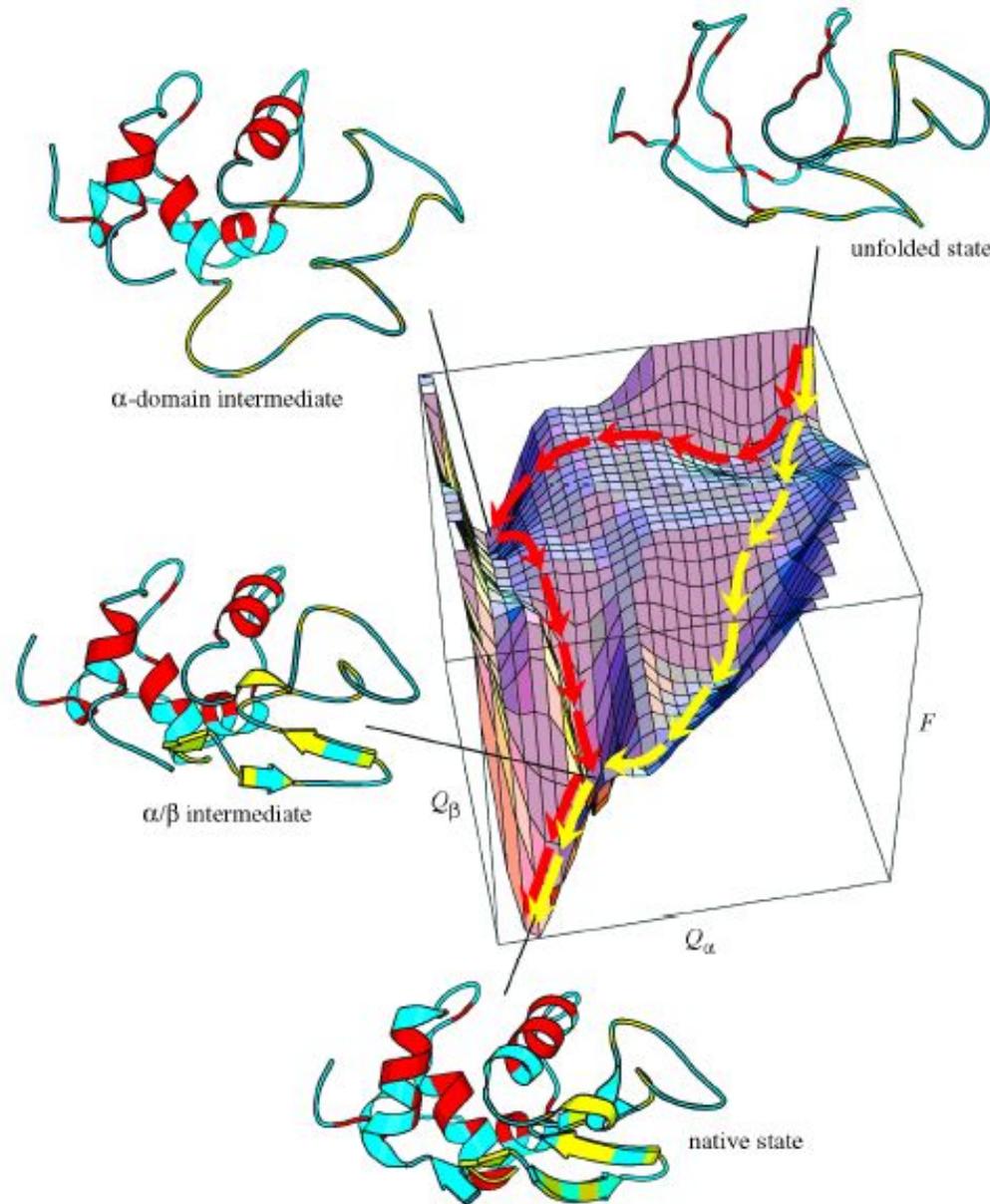
## 3 y 4 consecuencias del experimento de Anfinsen



# Enfermedades vinculadas a problemas en el plegamiento

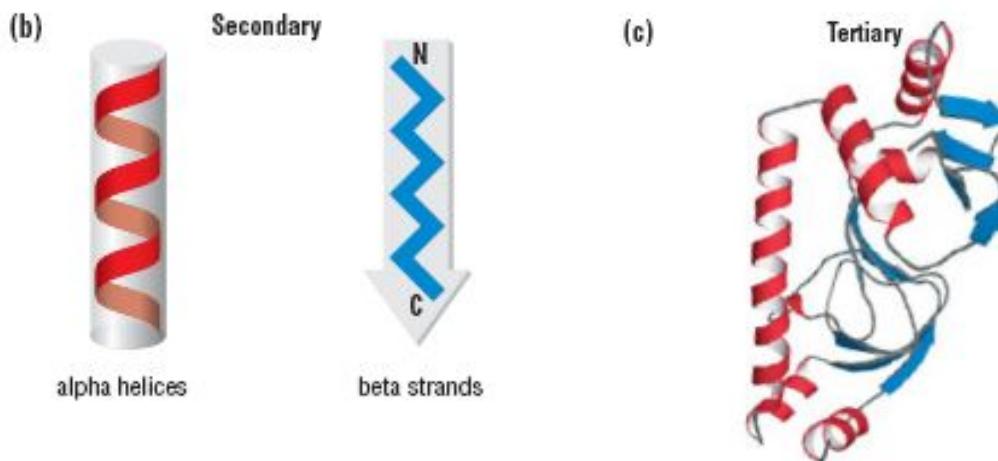
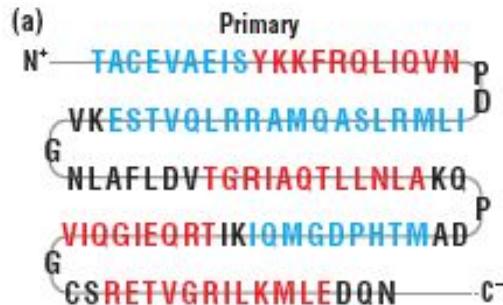
Table 1 | **Selected members of the family of amyloid diseases**

Clinical Syndrome	Fibril Component
Alzheimer's disease	A $\beta$ peptide, 1–42, 1–43
Spongiform encephalopathies	Full-length prion or fragments
Primary systemic amyloidosis	Intact light chain or fragments
Secondary systemic amyloidosis	76-residue fragment of amyloid A protein
Familial amyloidotic polyneuropathy I	Transthyretin variants and fragments
Senile systemic amyloidosis	Wild-type transthyretin and fragments
Hereditary cerebral amyloid angiopathy	Fragment of cystatin-C
Haemodialysis-related amyloidosis	$\beta_2$ -microglobulin
Familial amyloidotic polyneuropathy II	Fragments of apolipoprotein A-1
Finnish hereditary amyloidosis	71-residue fragment of gelsolin
Type II diabetes	Fragment of islet-associated polypeptide
Medullary carcinoma of the thyroid	Fragments of calcitonin
Atrial amyloidosis	Atrial natriuretic factor
Lysozyme amyloidosis	Full-length lysozyme variants
Insulin-related amyloid	Full-length insulin
Fibrinogen $\alpha$ -chain amyloidosis	Fibrinogen $\alpha$ -chain variants



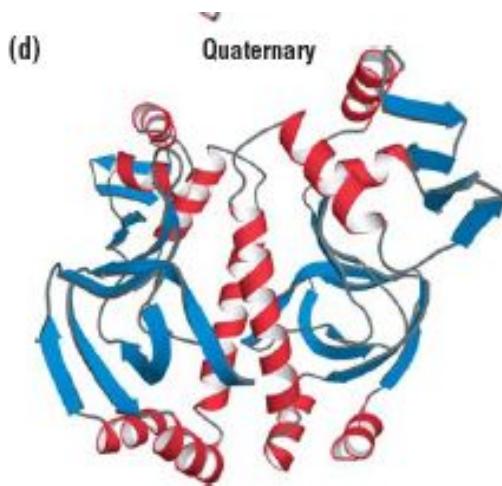
# Anatomía de una proteína

## Jerarquías estructurales

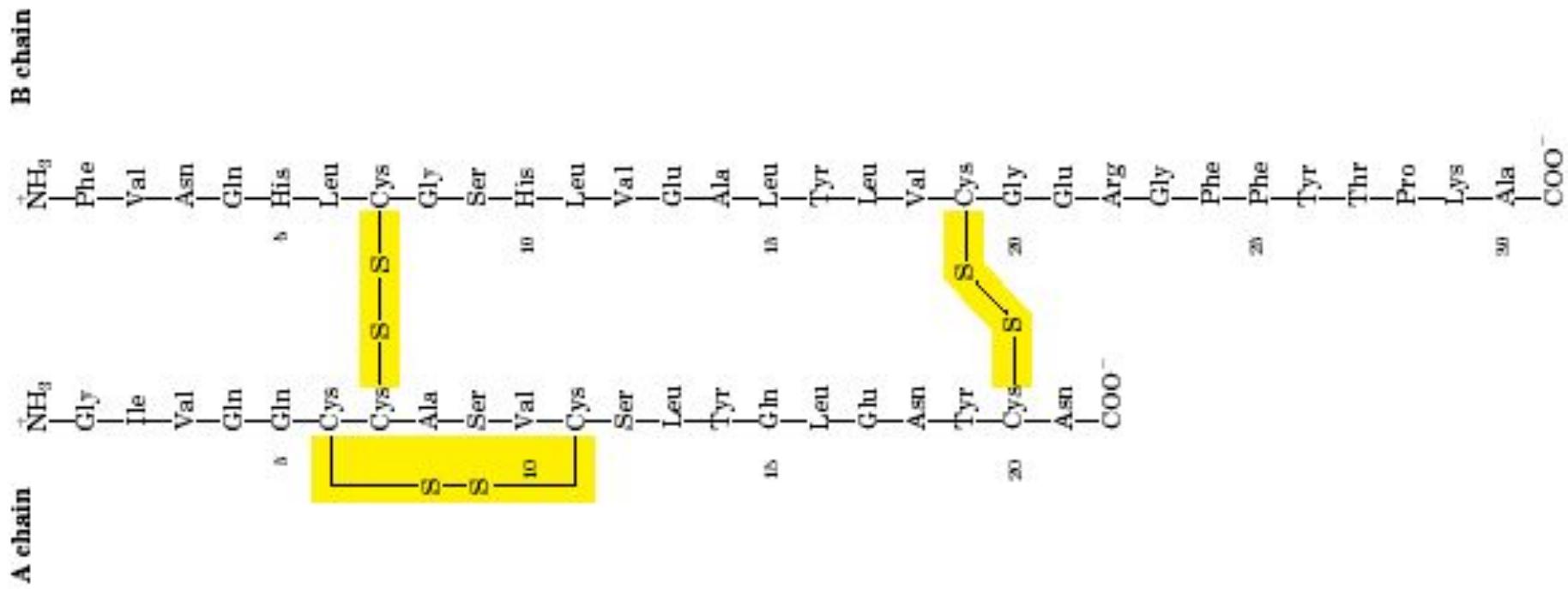


Frontispiece, Kaj Ulrich Linderstrøm-Lang in 1952.

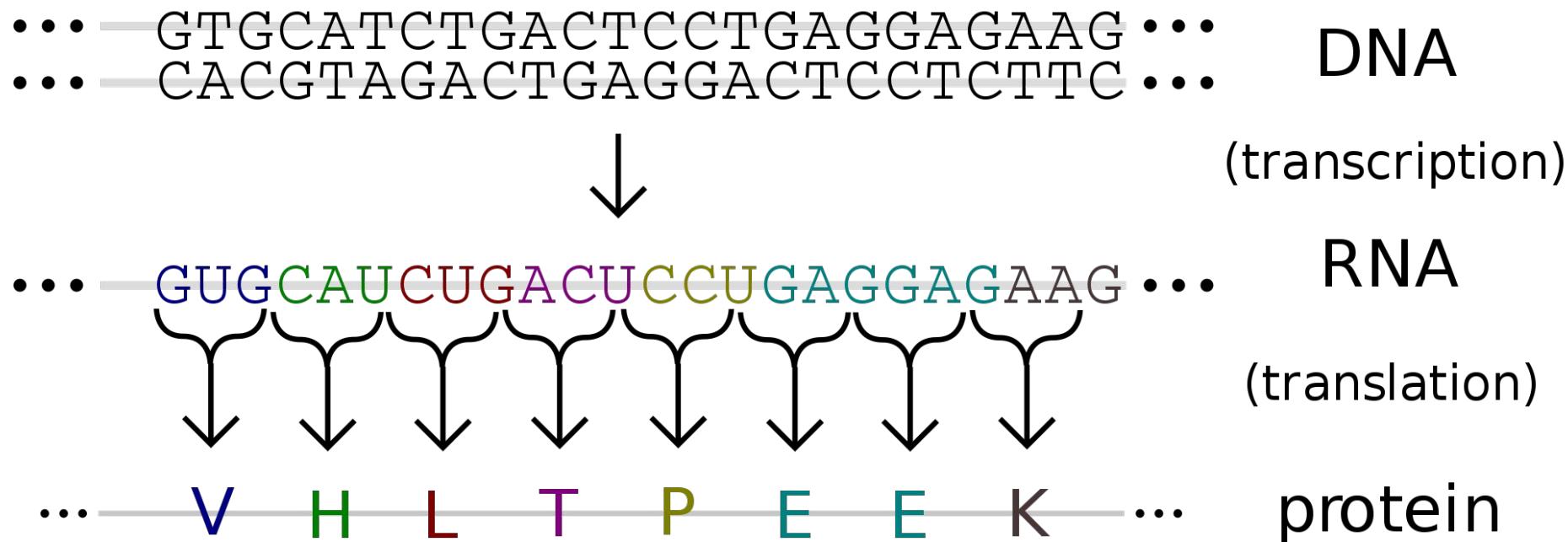
Linderstrøm-Lang 1952



## Estructura primaria: secuencia ordenada de aminoácidos desde el extremo amino terminal al carboxilo terminal



La estructura primaria está definida por la secuencia del gen que codifica a la proteína



# Variabilidad Secuencial: se observa que las secuencias de la misma proteína en distintos organismos pueden ser distintas

	*	100		*	120		*	140		*	160	
LPXA_ECOLI :	TRVEIGDRNRRIRESVTIHRGTVQGGGLTKVGSDNLLMINAHIAHDCTVGNRCILANNATLAGHVSVDDFAIIGGMTAVHQF											: 162:
LPXA_HELPY :	YELIIIGEDNLIREFCMINPGTEGGIKKTLIGDKNLLMAYVHVAHDCVIGSHCILANGVTLAGHIEIGDYVNIGGLTAIHQF											: 157:
LPXA_HELPJ :	YELMVGEDNLIREFCMINPGTEGGIKKTLIGDKNLLMAYVHVAHDCVIGSHCILANGVTLAGHIEIGDYVNIGGLTAIHQF											: 157:
AAAP77779 :	KELIIIGDDNLIREFTTLNPAGTAGGRGKTIIGDRNLFMAYVHIAHDCTVGNDCILANNATLGGHVELGDYVNIGGLTIVHQF											: 155:
LPXA_ECO57 :	TRVEIGDRNRRIRESVTIHRGTVQGGGLTKVGSDNLLMINAHIAHDCTVGNRCILANNATLAGHVSVDDFAIIGGMAAVHQF											: 162:
LPXA_SALTY :	TRVEIGDRNRRIRESVTIHRGTVQGGGLTKVGSDNLLMINAHVAHDCTVGNRCILANNATLAGHVSVDDFAIIGGMTAVHQF											: 162:
LPXA_SALT1 :	TRVEIGDRNRRIRESVTIHRGTVQGGGLTKVGSDNLLMINAHVAHDCTVGNRCILANNATLAGHVSVDDFAIIGGMTAVHQF											: 162:
LPXA_YEREN :	TRVEIGDRNRRIRESVSIHRGTVQGGGLSKVGSDNLLMINAHIAHDCTIIGDRCILANNATLGGHVEIDDFAIIGGMAAVHQF											: 162:
LPXA_YERPE :	TRVEVGDRNRRIRESVTIHRGTTQGGGVTKVGCDNLLMVNTVVAHDCVIGNRCILANNAALGGHVIEDDYAIIIGGMAAVHQF											: 162:
LPXA_PROMI :	TQVIIGDRNLIRESVTIHRGTTQGGNITKIGNDNLLMINTHVAHDCIIIGDRCIIIANNGTLGGHVTLGDYVIIGGMSAVHQF											: 162:
CAD83354 :	TRIEIGNYNQIRESVTIHRGTVQGGQVTKIGNSNLFMINVHIAHDCTIIGNNCIMANNVTLGGHVVKVDDYTIIGGMAAVHQF											: 162:
LPXA_WIGBR :	TKVYIGDRNKIRENSTIHRGTVQSNKITKIGNDNLEMVNVHIAHDCTVIENNCCVMANVTLGGHVVKIGNHVVIIGGMAAVHQF											: 162:
LPXA_PASMU :	TRTIIGDRNRRIRESVTIHRGTAQGGSVTVIGDDNLLMVNVHVAHDCRIKNRCILANNATLAGHVELDDFVIVGGMSAIHQF											: 162:
LPXA_HAEIN :	TKTIIGNSNKIREHVTIHRGTIQGCGVTAIGNNNLLMINVHVAHDCQIKNNCILANNATLAGHVELDDFVIVGGMSAIHQF											: 162:
	t 6G N IRE 6hrGT qg 3 6G NLLM n H6AHDC 6g C66ANn tL GH6 6 1 6GG6 a6HQF											

Actividad biológica

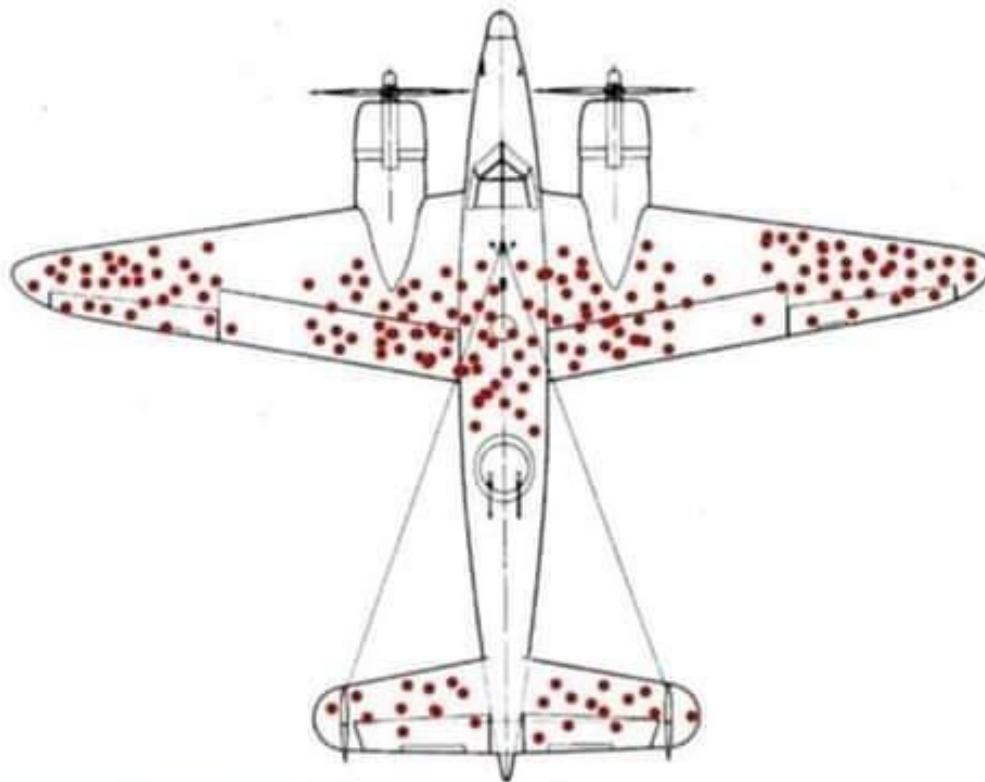
Función

Cinética

Termodinámica

Historia

Los alineamientos de proteínas son como los aviones que llegaban a destino durante la Segunda Guerra Mundial



El estudio de la variabilidad secuencial dio las bases  
para el desarrollo de la evolución molecular

*J. Theoret. Biol.* (1965) **8**, 357–366

## Molecules as Documents of Evolutionary History

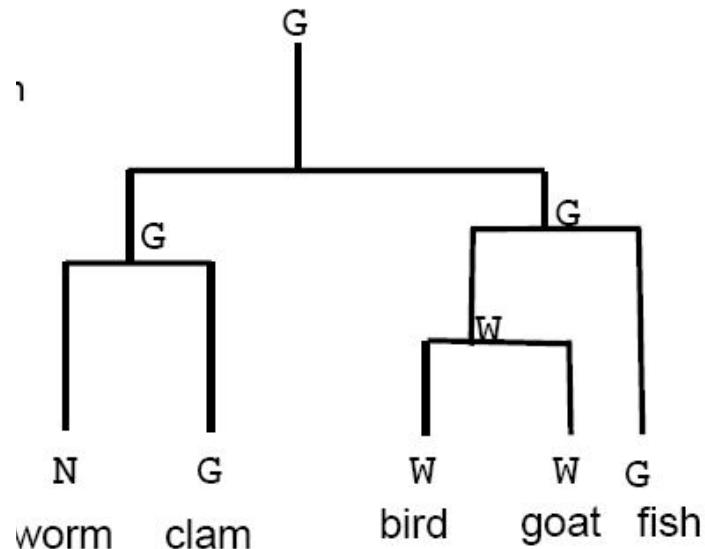
EMILE ZUCKERKANDL AND LINUS PAULING

*Gates and Crellin Laboratories of Chemistry,  
California Institute of Technology, Pasadena, California, U.S.A.*

(Received 17 September 1964)

Las posiciones alineadas son homólogas entre sí, esto quiere decir que comparten un ancestro común durante la evolución

Worm	ALL <b>N</b> KDRTFGY
Clam	ALV <b>G</b> KDTTFGY
Bird	AIV <b>W</b> KDHYLTY
Goat	PPQ <b>W</b> KDHYVCV
Fish	PIQ <b>G</b> RTYVDCC



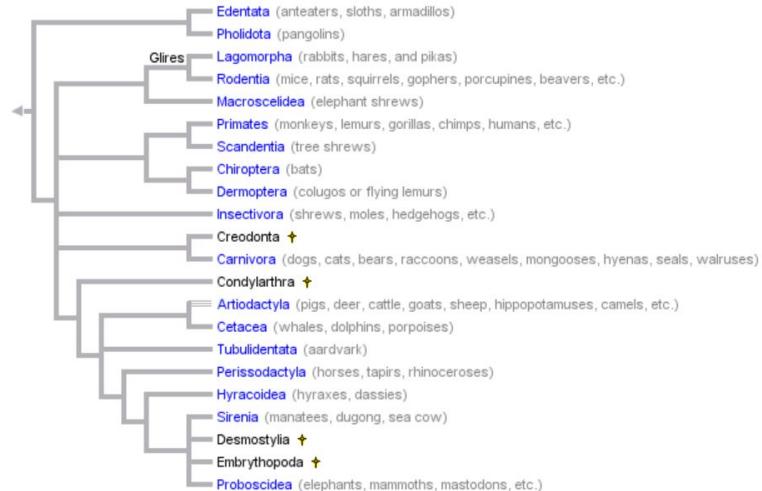
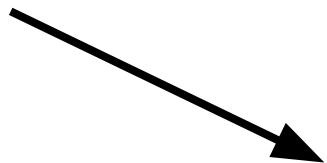
# De la misma forma esto se extiende a un conjunto de secuencias homólogas que podrían ser de esta forma ortólogas

```

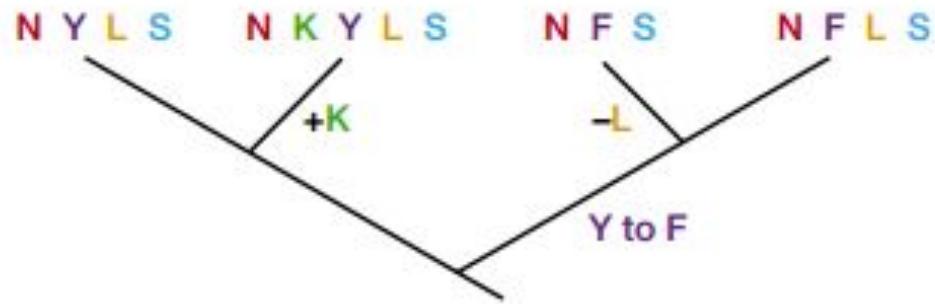
| * 20 * 40 * 60 * 80 * 100 * : 83
lial_A : FVFEYRTELRRRDTQVELKAKRDEALAKRRNF : -CEIPNPTCNCDEMCPGCGSIVKIRPCLISRHPPPEIINIRRA-LIPIFVSLGKTCDFCSRFIDRBA : 102
lw5_B : MNQGTVNNSNDIHKGINSNNVENSQCAQGAAPIILSRKCPPIENIRRA-LIPIFVSLGKTCDFCSRFIDRBA : 73
3fey_C : GTVNWSVEIHKGINSNNLSEQCAQGAAPIILSRKCPPIENIRRA-LIPIFVSLGKTCDFCSRFIDRBA : 70
3ull_B : PGG----VITSMEMIFKSPFCQISLQKERRILLSRENPPIEVISTPVMAMVVEILKEKENCYLOBBA : 70
4bl8_A : GRRRRRKRR-----RSLEPMIGCVYDNNNQIPEATQRKLLISLERSPEEVIL-QSTVVRERQVQIASEDFCFDTPA : 74
4b8j_A : DEPRMNEGVPSSEIDPKQICATTKRLLILSRERNPPIEVIL-KTUVVGEEVILR-SPHTIVOBBA : 64
4rxh_B : PMAAGTW-----SCLANNLRLPILISQNPPIPVV-VSPPVQGLLDDDFPFRIDRBA : 58
4tnm_A : VITRMEMIFDSDIDLCLATEKPERKLLISREPSPEEVINTPVVVDIPEELKRNENCTIOBBA : 67
4uad_A : TSEADLGNASDNQGICLISVGAARLLLSDRNPPIDILKS-CIIPLVHODERDDNPSIOBBA : 66
4uae_A : GTVNWSNDIHKGINSNNVENSQCAQGAAPIILSRKCPPIENIRRA-LIPIFVSLGKTCDFCSRFIDRBA : 70
4wv6_A : WalTNIASG s qT 66 gAEF F6 ll 3 6 EQA6WALGN6aGD RD V6 6 PEL Rn tw EsN CR k p p

```

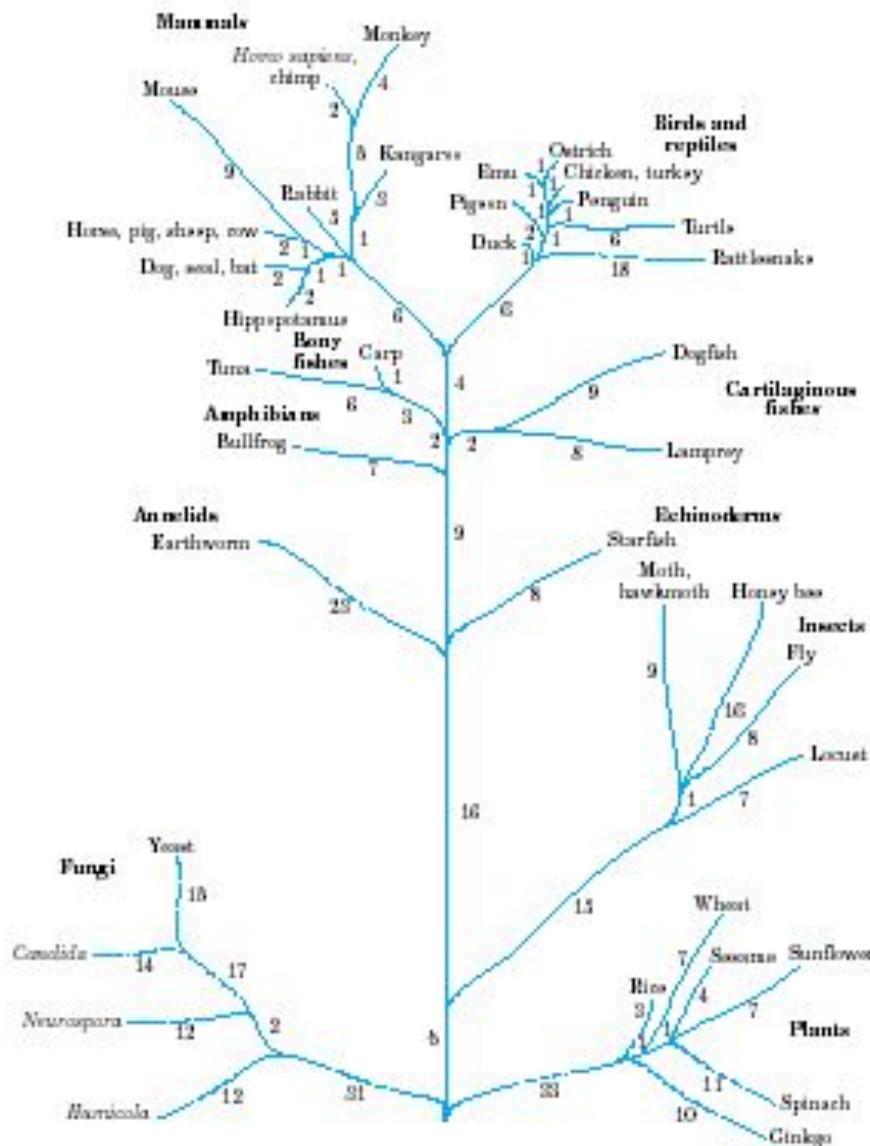
6 s Q6 at Rk6LS e PPI 66 g66 v f6 6QE A



seqA	N	•	F	L	S
seqB	N	•	F	-	S
seqC	N	K	Y	L	S
seqD	N	•	Y	L	S



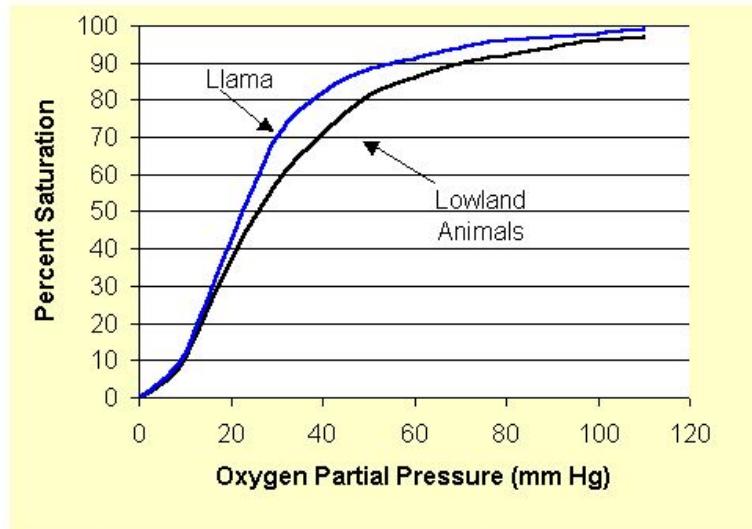
# Los árboles filogenéticos describen el proceso evolutivo de organismos o proteínas



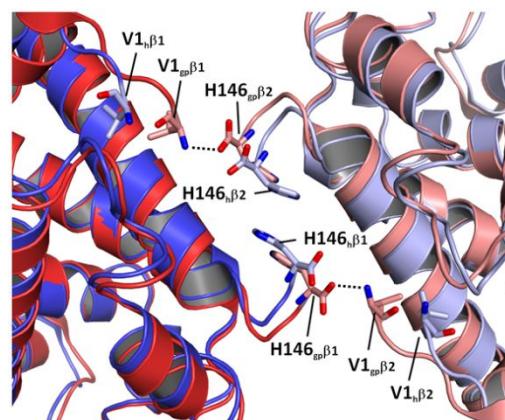
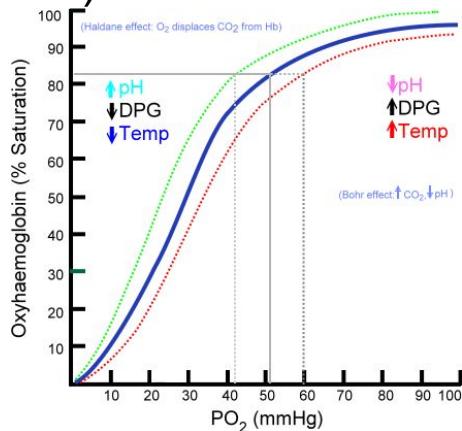
# Ortólogos con igual función y adaptación: evolución de la Hb



Llama, alpaca, vicuña viven a 5000mts de altura



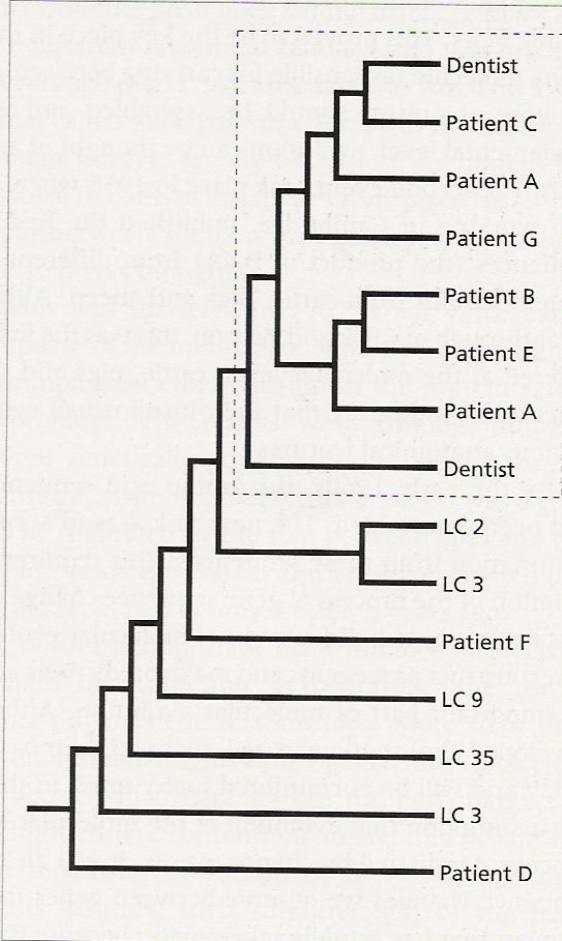
En general se estabiliza el estado R, se desestabiliza el estado T, baja afinidad por 2,3 difosfoglicerato (disminuyendo el número de contactos con este efector)



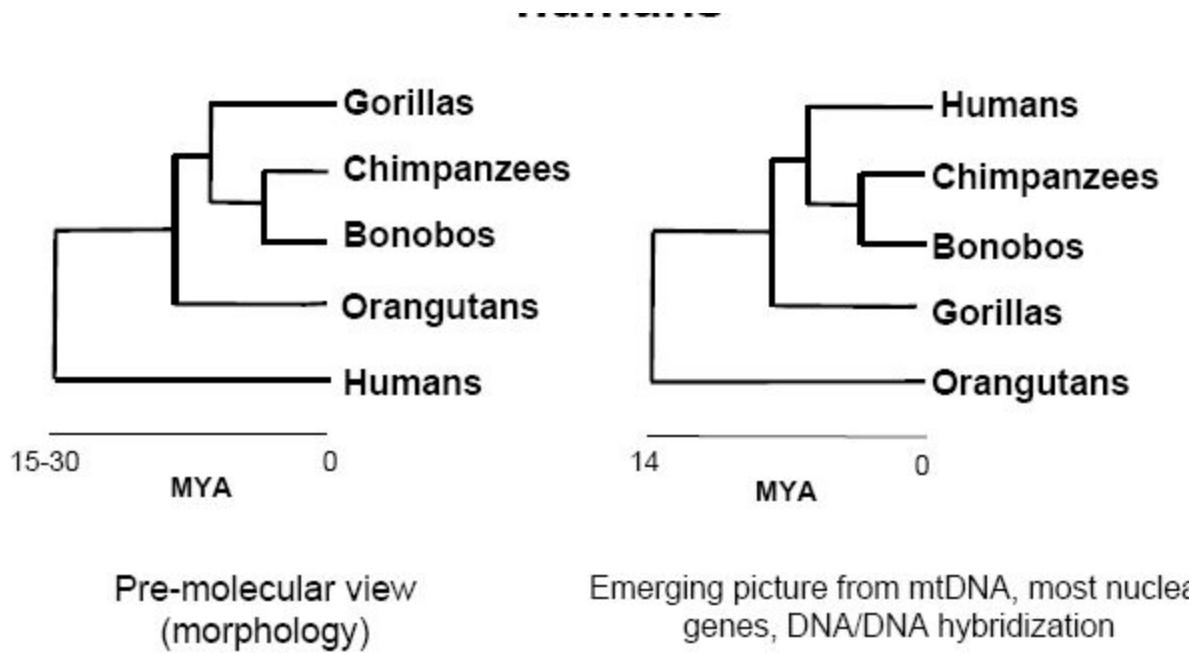
Pairet B, Jaenicke E (2010)  
Structure of the Altitude  
Adapted Hemoglobin of  
Guinea Pig in the R-State.

# El caso del dentista asesino

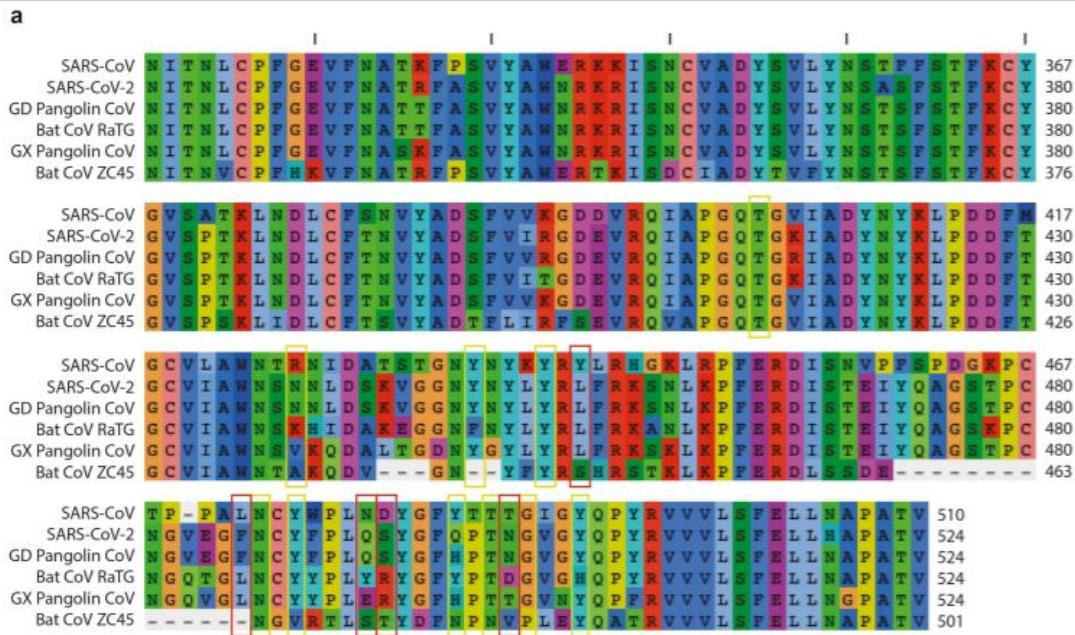
**Fig. 1.2** The case of the Florida dentist. Each branch represents the sequence from part of the envelope (*env*) gene of HIV-1. Viral sequences were obtained from the dentist and seven of his former patients (labelled A to G), also infected with the virus. Five of these patients (A, B, C, E and G), have sequences very closely related to those of the dentist (boxed), suggesting that he infected them. Two of his other former patients (D and F) had other risk factors for HIV infection and their viruses are separated from the dentist by sequences taken from local controls (LC)—HIV-infected individuals living within a 90-mile radius of the dentist's surgery. Because HIV-1 is so variable, two different sequences are included for the dentist and patient A. Data taken from Ou *et al.* (1992).



# Evolución del hombre: morfología vs datos moleculares

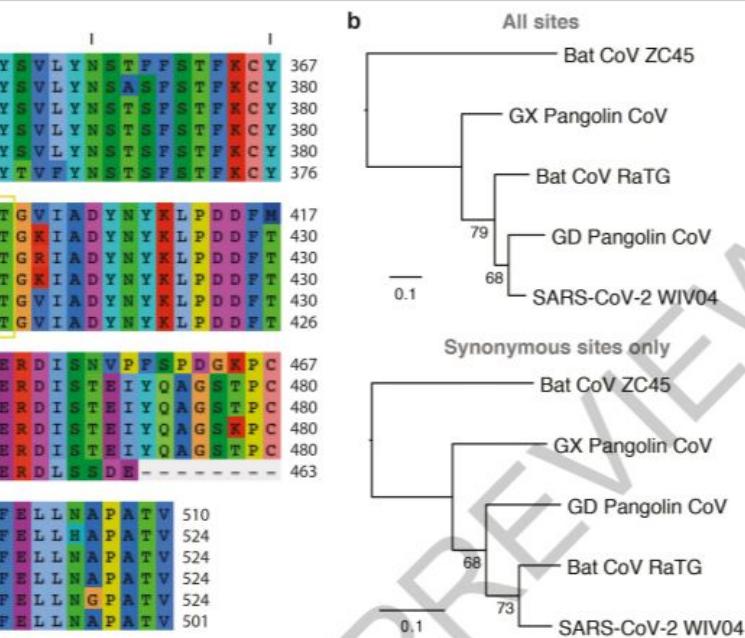


# Desmentir la creación de un nuevo virus pagado por Bill Gates...



**Fig. 3 | Analysis of the receptor-binding domain (RBD) sequence.**

(A) Sequence alignment showing the RBD in human, pangolin and bat coronaviruses. The five critical residues for binding between SARS-CoV RBD and human ACE2 protein are indicated in red boxes, and ACE2-contacting residues are indicated with yellow boxes, following Wan et al.<sup>9</sup>. Note that in Guangdong pangolin sequence, the codon positions coding for amino acids



337 proline, 420 aspartic acid, 499 proline and 519 asparagine have ambiguous nucleotide compositions, resulting in possible alternative amino acids at these sites (threonine, glycine, threonine and lysine, respectively). GD: Guangdong, GX: Guangxi. (B) Phylogenetic trees of the SARS-CoV-2 related lineage estimated from the entire RBD region (upper) and synonymous sites only (lower). Branch supports obtained from 1,000 bootstrap replicates are shown.

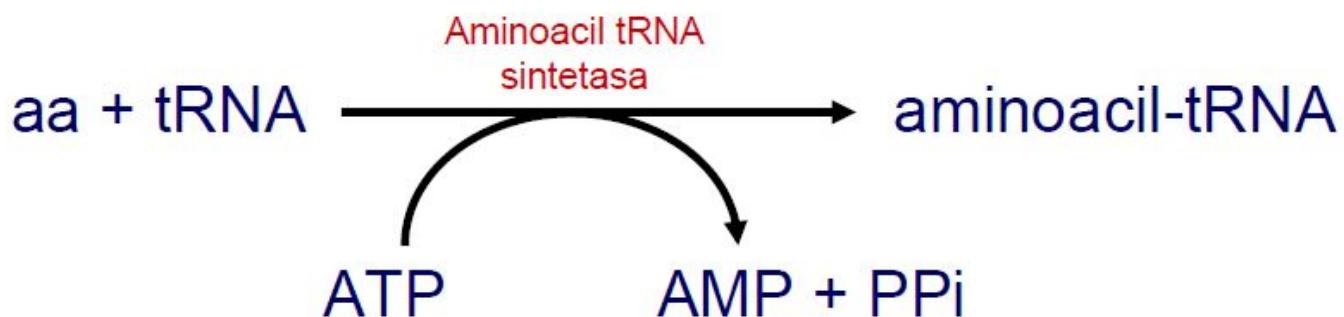
## Enlace peptídico

Enlace no favorecido termodinámicamente

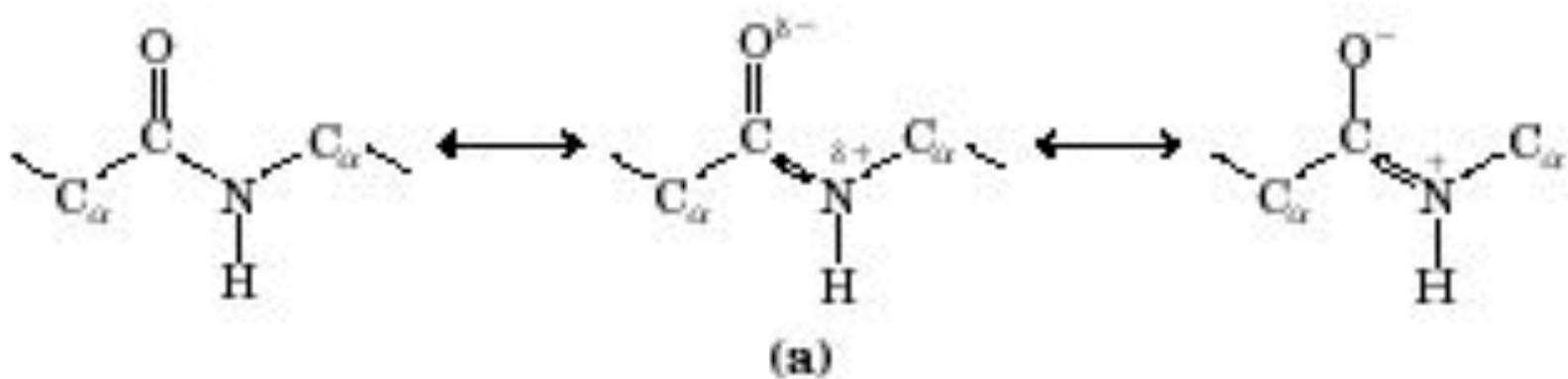
$$\Delta G = 10 \text{ KJ/mol}$$

La reacción sin catalizar es muy lenta

En la célula la formación del enlace peptídico está acoplado a la hidrólisis de ATP



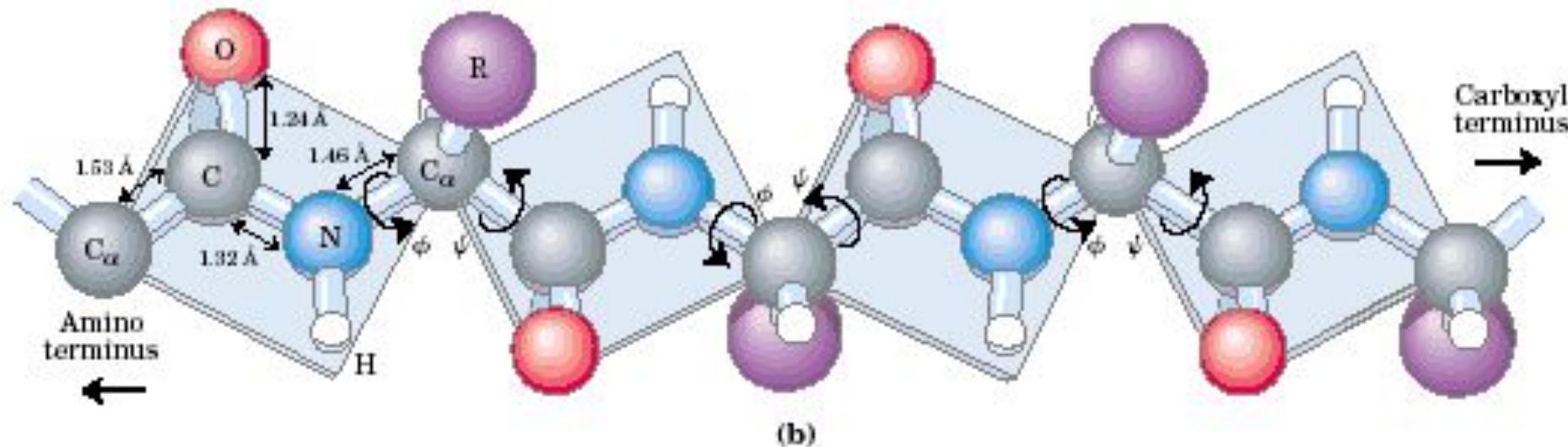
# Enlace peptídico



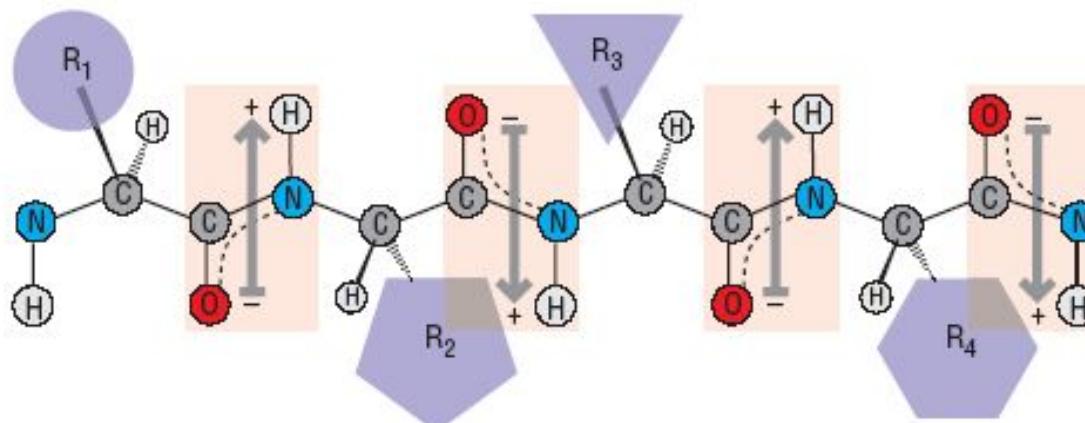
Propiedades fisicoquímicas:

1. Plano
2. Dipolo
3. Configuración trans

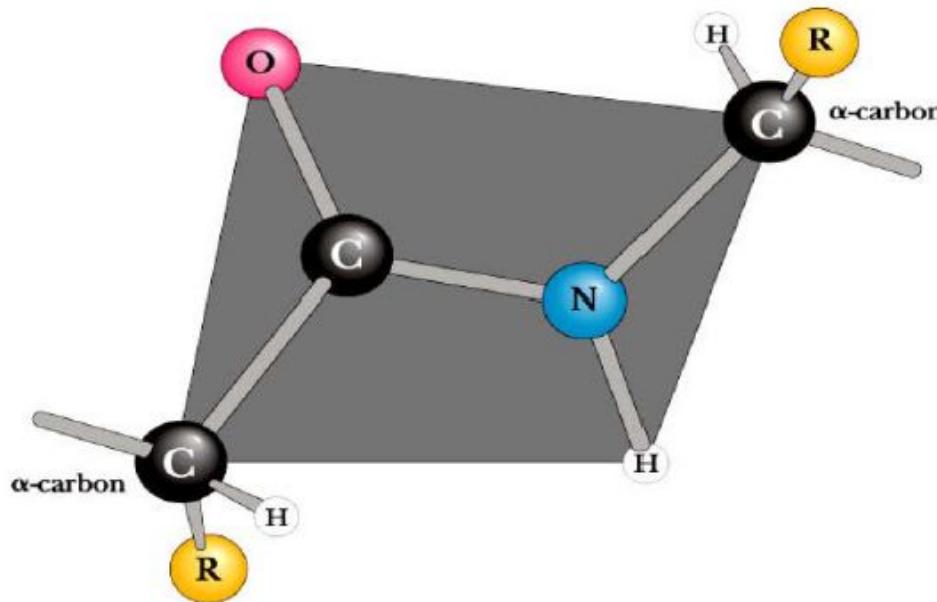
## Consecuencias de la resonancia



(b)



- Geometría esencialmente plana. Los átomos que participan en el enlace (C, O, N, H) están en el mismo plano



## Restricciones estéricas

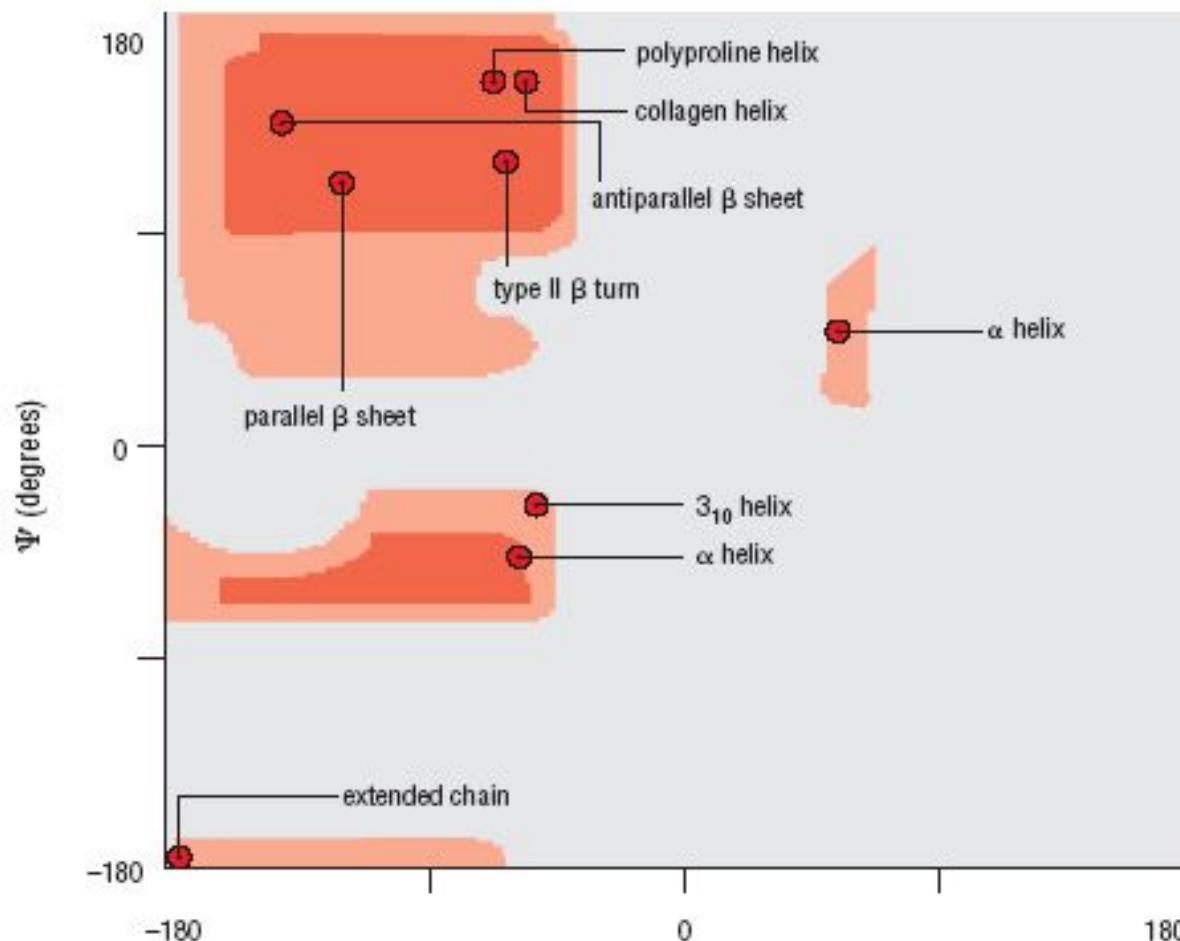
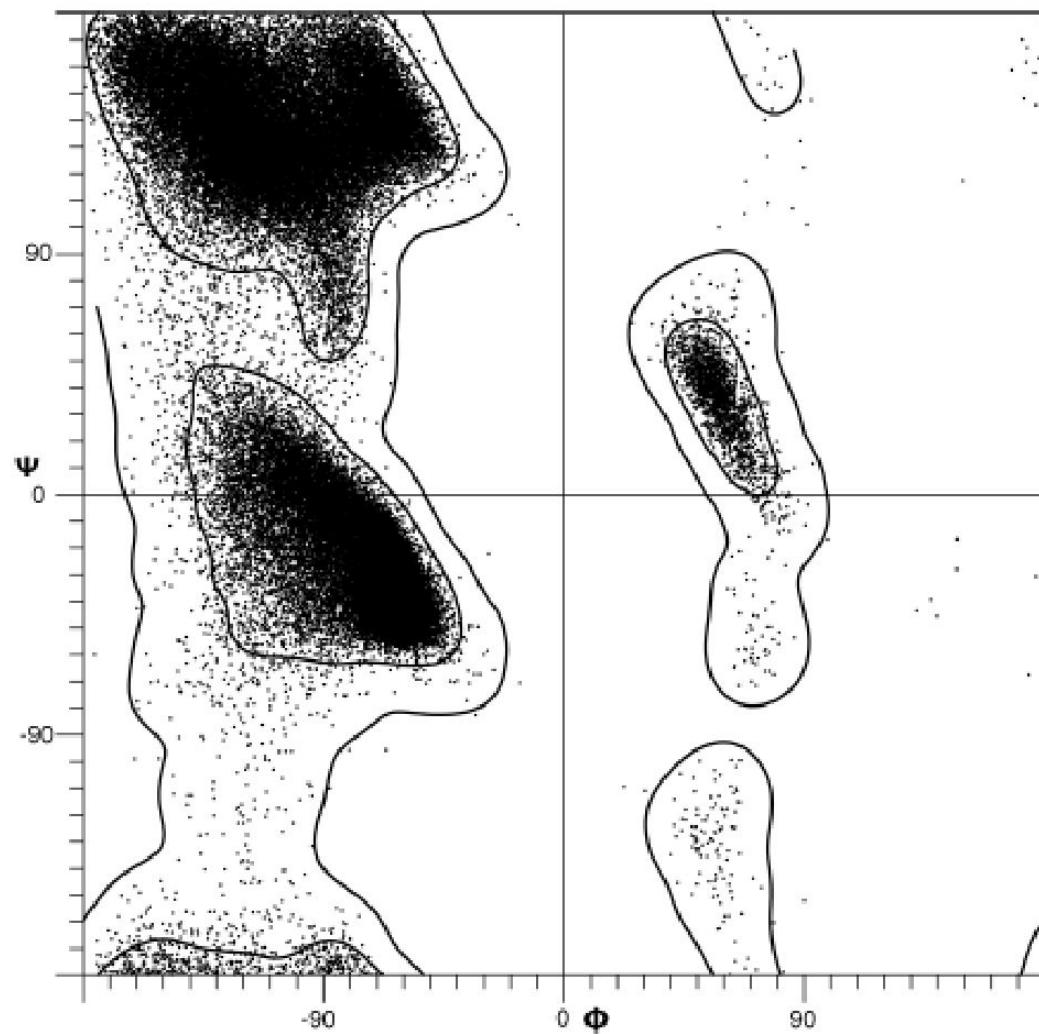
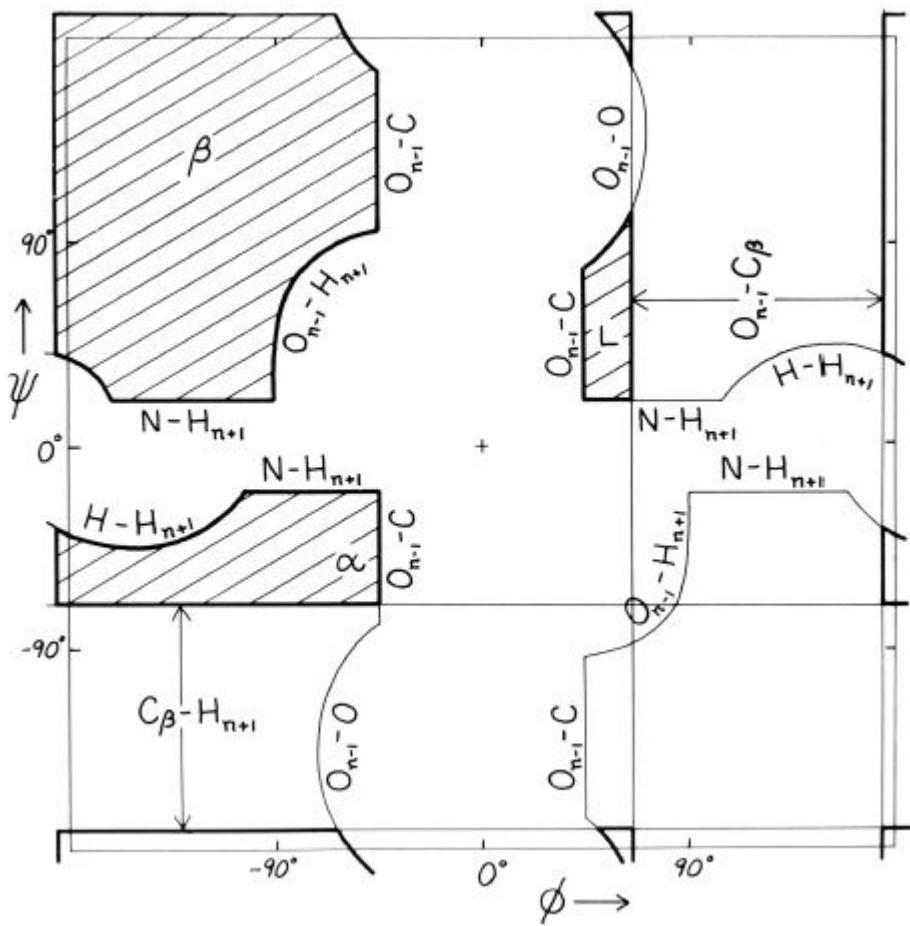
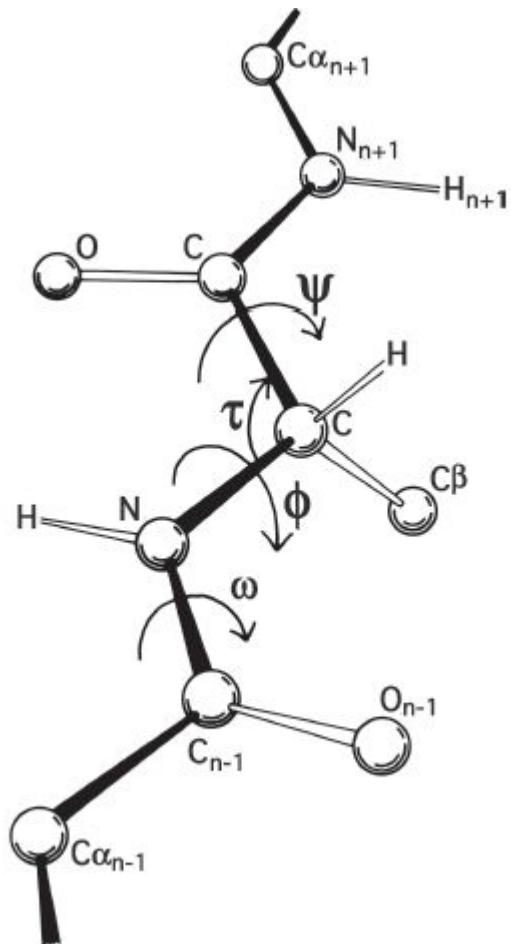


Diagrama de Ramachandran

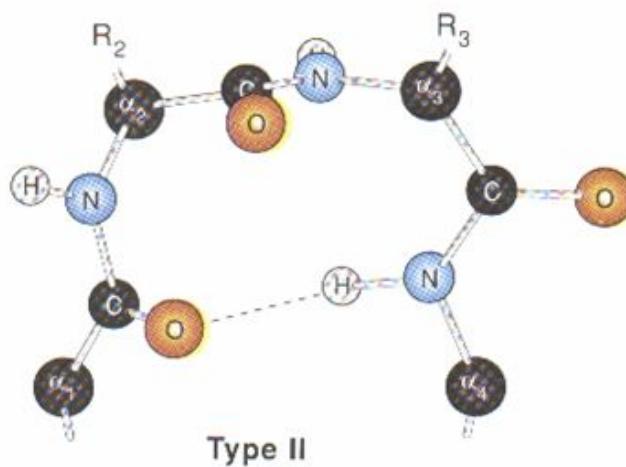
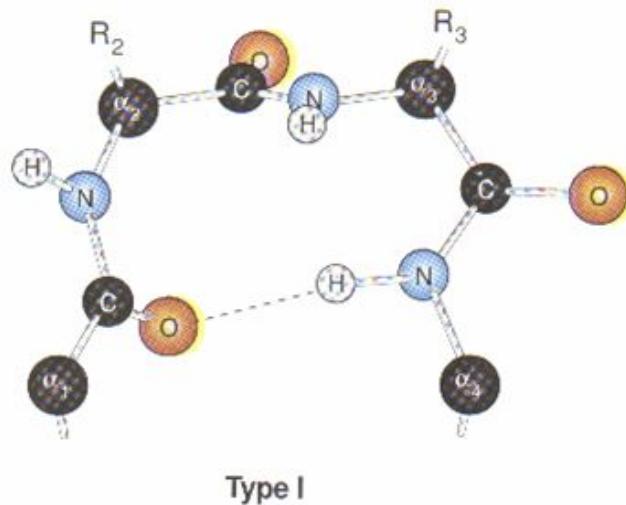


# Restricciones estéricas

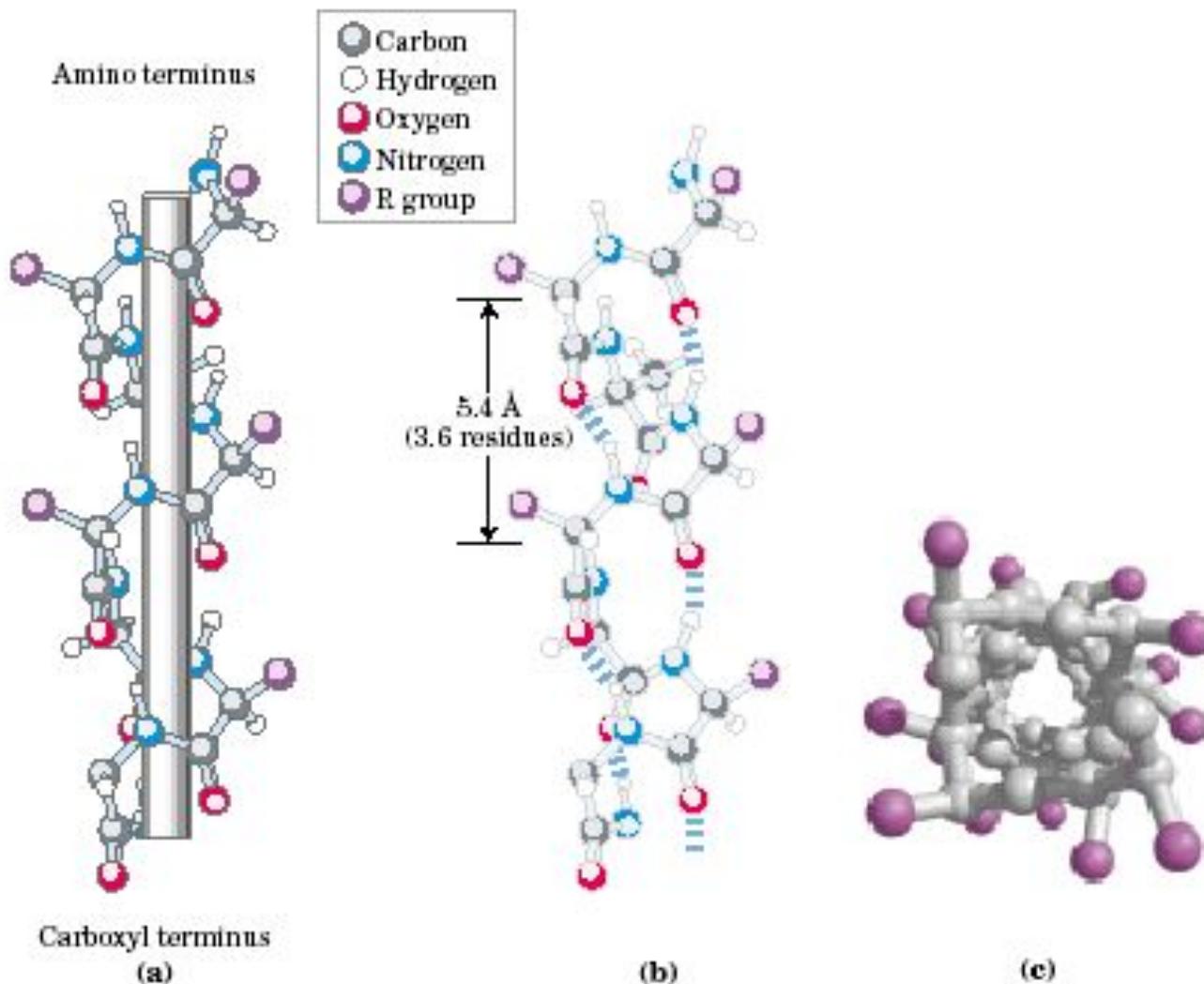


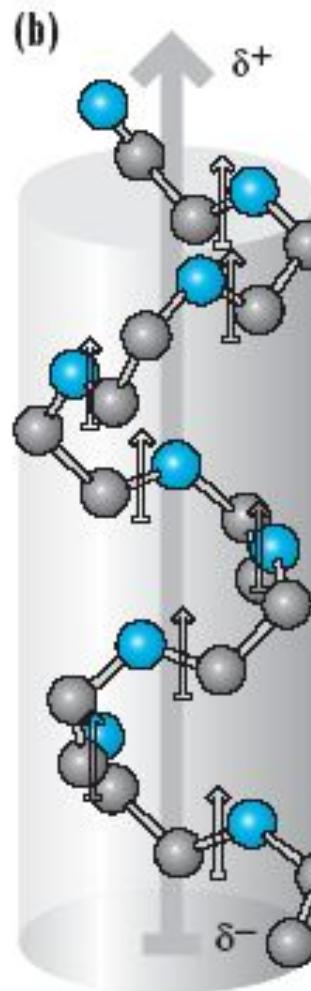
**Estructura secundaria:** Arreglo local de la estructura primaria estabilizado por puentes de hidrógeno

Bends (beta turns, reverse turns or hairpin): estructuras secundarias simples



# Estructura Secundaria: hélice alfa





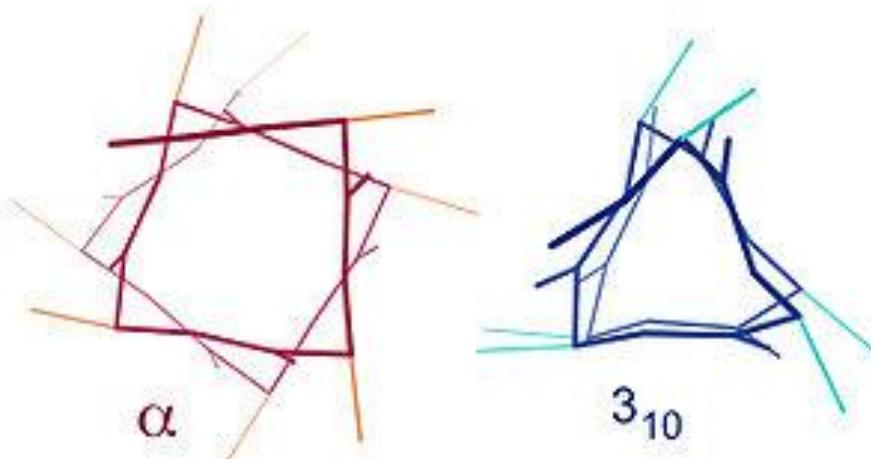
NH<sub>2</sub> terminal

Momento dipolar

COOH terminal

### Average Conformational Parameters of Helical Elements

Conformation	Phi	Psi	Omega	Residues per turn	Translation per residue
Alpha helix	-57	-47	180	3.6	1.5
3-10 helix	-49	-26	180	3.0	2.0
Pi-helix	57	-70	180	4.4	1.15
Polyproline I	-83	+158	0	3.33	1.9
Polyproline II	-78	+149	180	3.0	3.12
Polyproline III	-80	+150	180	3.0	3.1



# THE STRUCTURE OF PROTEINS: TWO HYDROGEN-BONDED HELICAL CONFIGURATIONS OF THE POLYPEPTIDE CHAIN

BY LINUS PAULING, ROBERT B. COREY, AND H. R. BRANSON\*

GATES AND CRELLIN LABORATORIES OF CHEMISTRY,  
CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA†

Communicated February 28, 1951

## THE PLEATED SHEET, A NEW LAYER CONFIGURATION OF POLYPEPTIDE CHAINS

BY LINUS PAULING AND ROBERT B. COREY

GATES AND CRELLIN LABORATORIES OF CHEMISTRY,\* CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA

Communicated March 31, 1951

For many years it has been assumed that in silk fibroin, stretched hair and muscle, and other proteins with the  $\beta$ -keratin structure the polypeptide chains are extended to nearly their maximum length, about 3.6 Å per residue, and during the last decade it has been assumed also that the chains form lateral hydrogen bonds with adjacent chains, which have the opposite orientation. A hydrogen-bonded layer of this sort is represented diagrammatically in figure 1.<sup>1-4</sup>

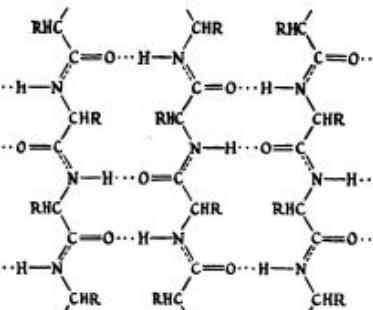


FIGURE 1

Diagrammatic representation of a hydrogen-bonded layer structure of polypeptide chains with alternate chains oppositely oriented.

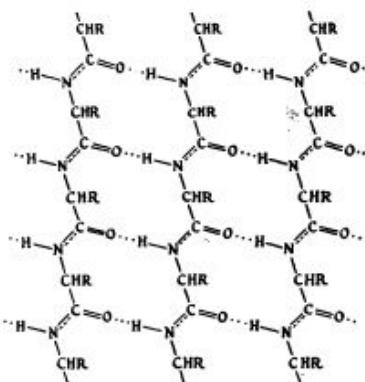


FIGURE 2

Diagrammatic representation of a hydrogen-bonded layer structure of polypeptide chains with all chains similarly oriented (the pleated sheet).



FIGURE 2  
The helix with 3.7 residues per turn.

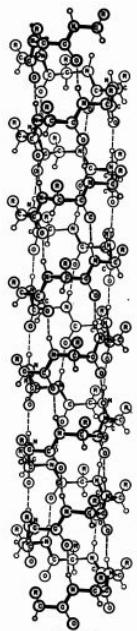


FIGURE 3  
The helix with 5.1 residues per turn.

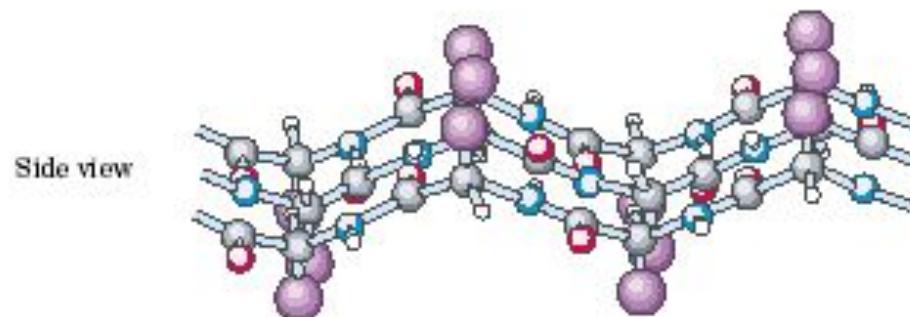
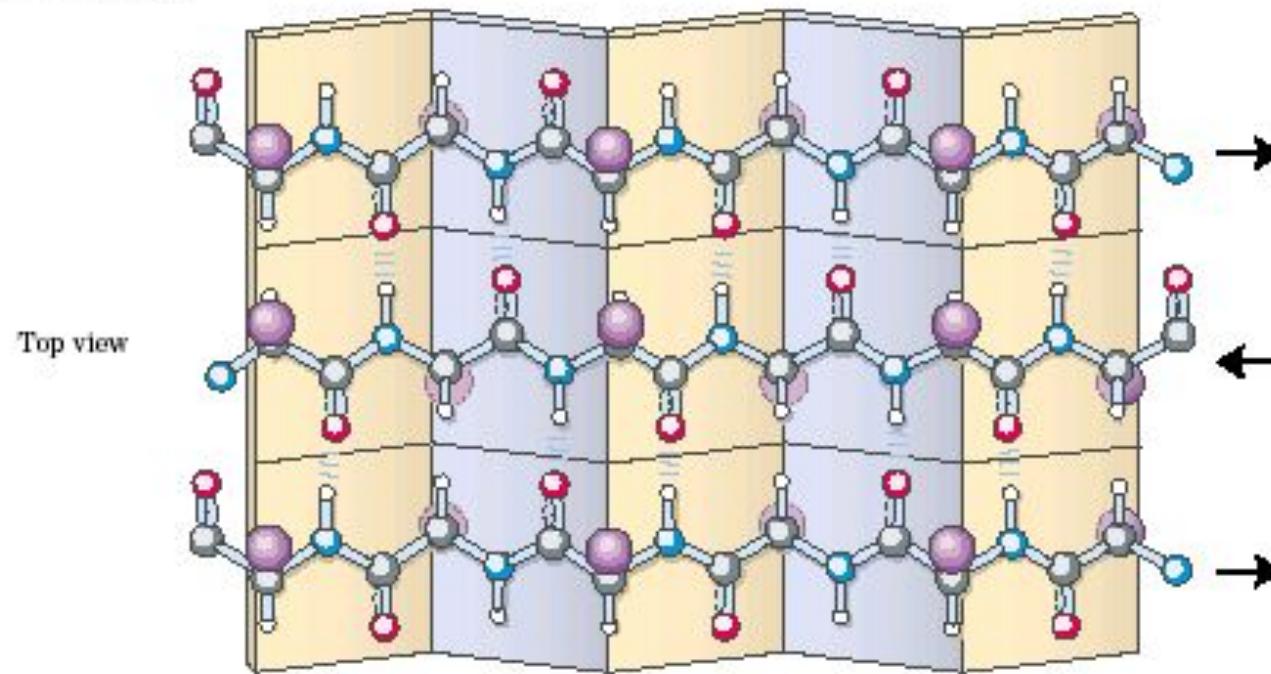


Casa natal de Linus Pauling USA



# Estructura Secundaria: hoja beta plegada

(a) Antiparallel



Período: 7 Å

(b) Parallel

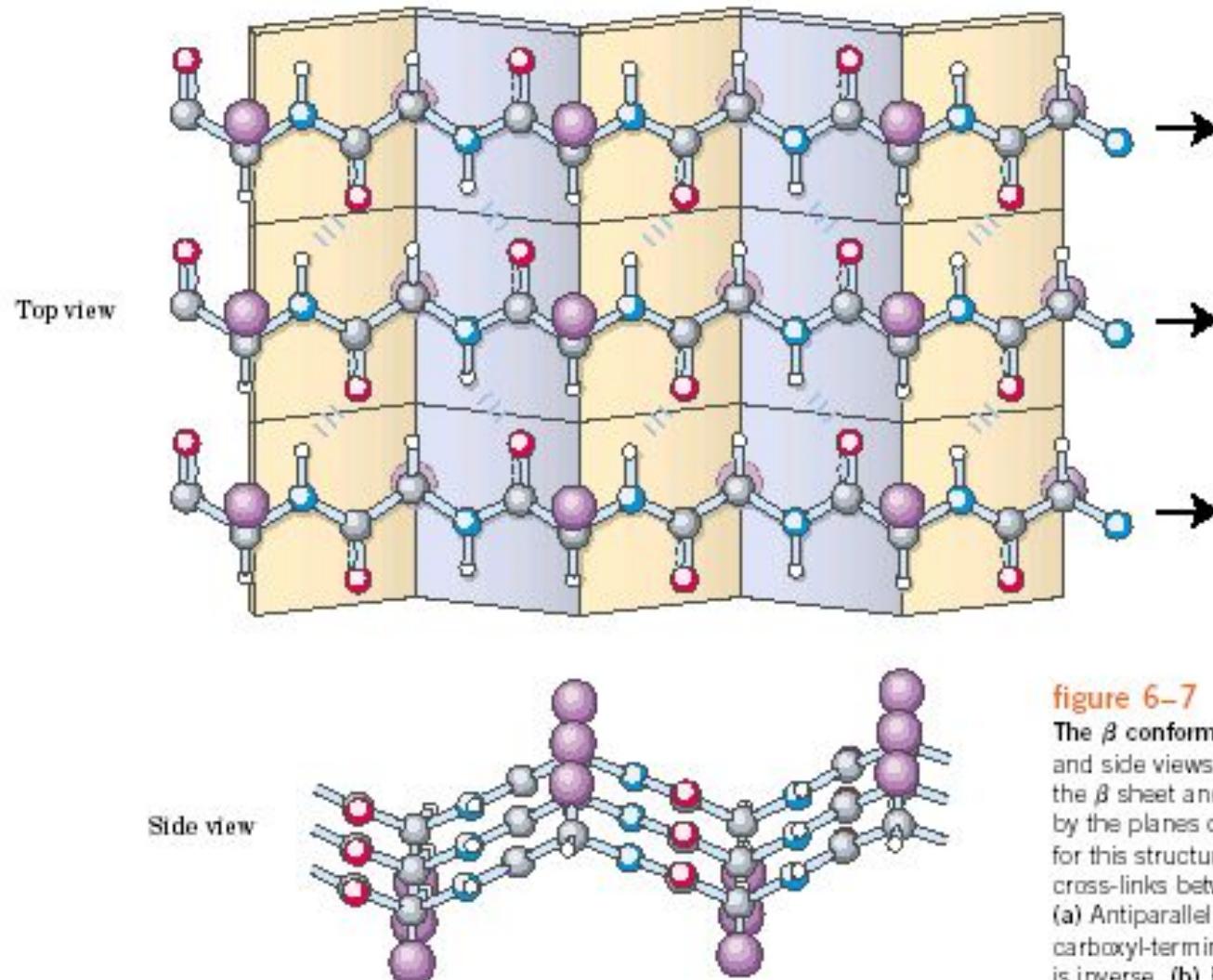
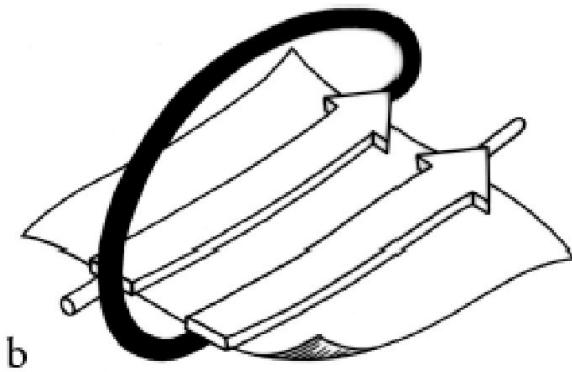
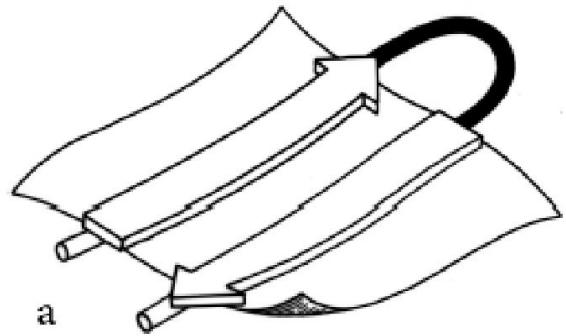


figure 6–7

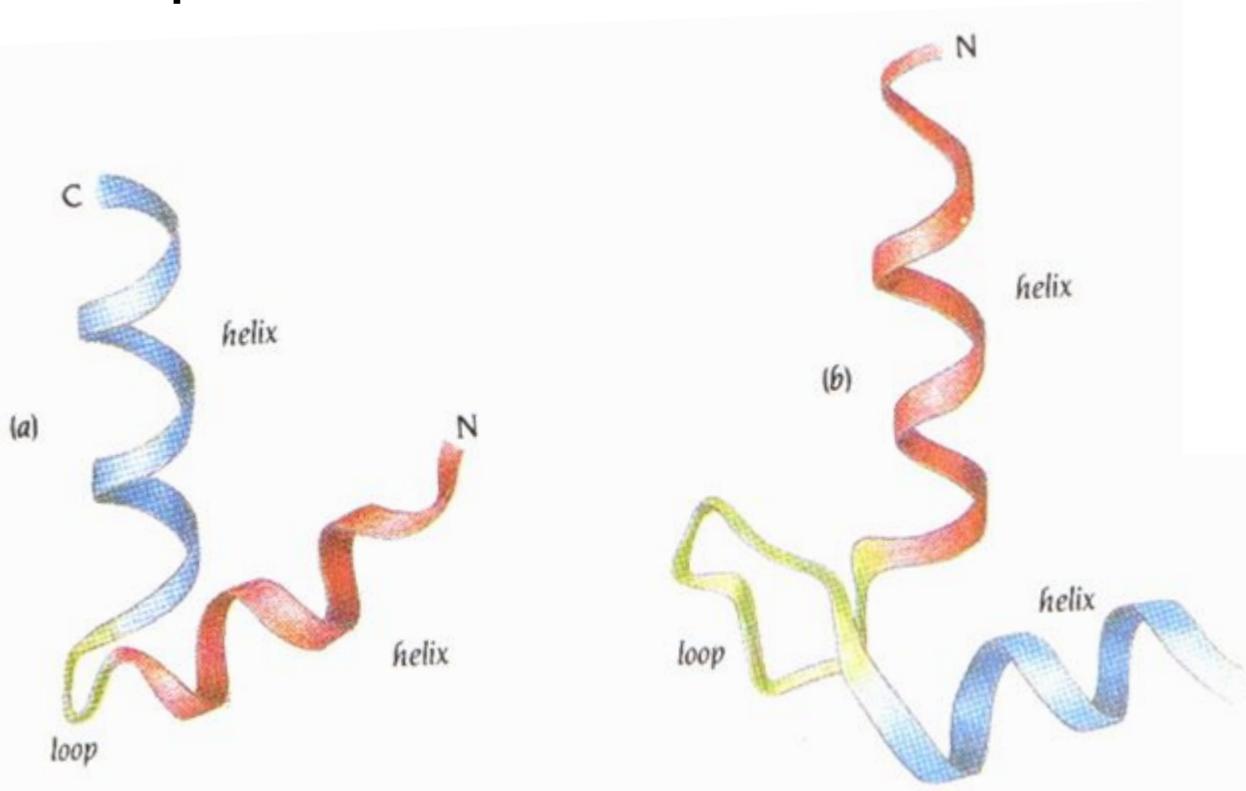
The  $\beta$  conformation of polypeptide chains. These top and side views reveal the R groups extending out from the  $\beta$  sheet and emphasize the pleated shape described by the planes of the peptide bonds. (An alternate name for this structure is  $\beta$ -pleated sheet.) Hydrogen-bond cross-links between adjacent chains are also shown. (a) Antiparallel  $\beta$  sheet, in which the amino-terminal to carboxyl-terminal orientation of adjacent chains (arrows) is inverse. (b) Parallel  $\beta$  sheet.

Período: 6.5 Å

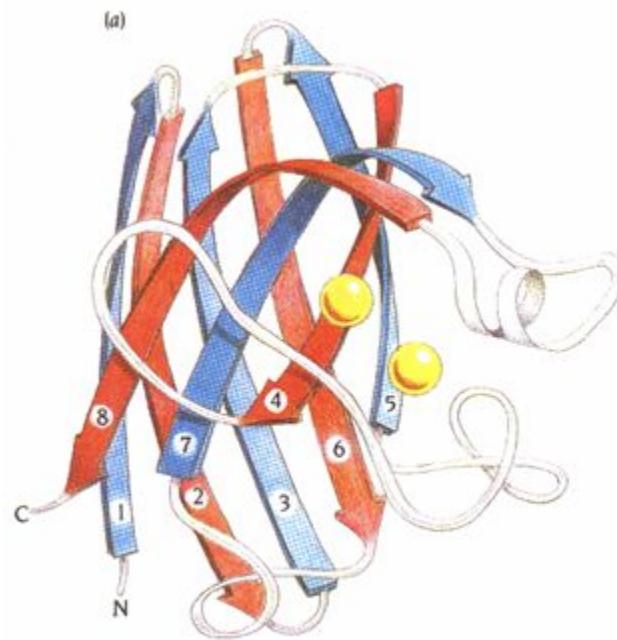
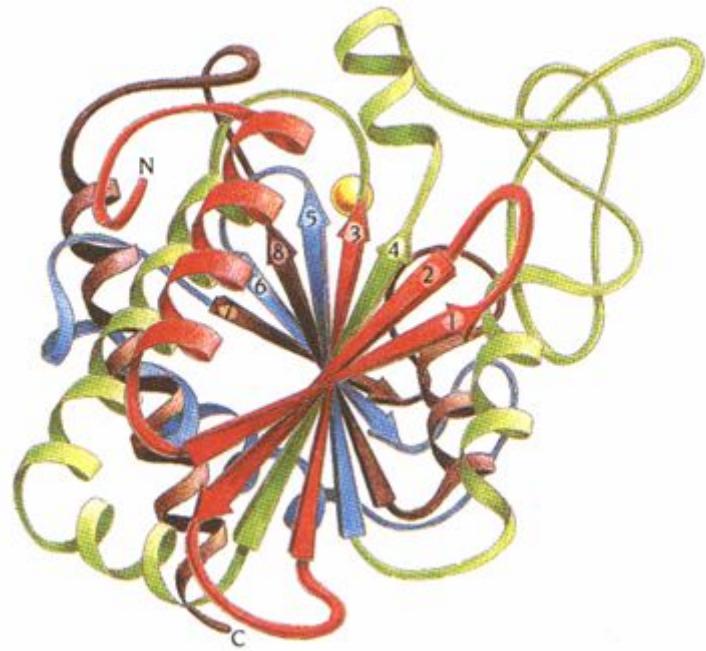


# Conexiones entre elementos de estructura secundaria

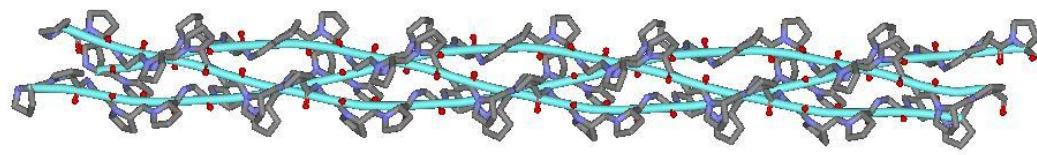
## Loops



# Estructura terciaria: arreglo espacial de los elementos de estructura secundaria

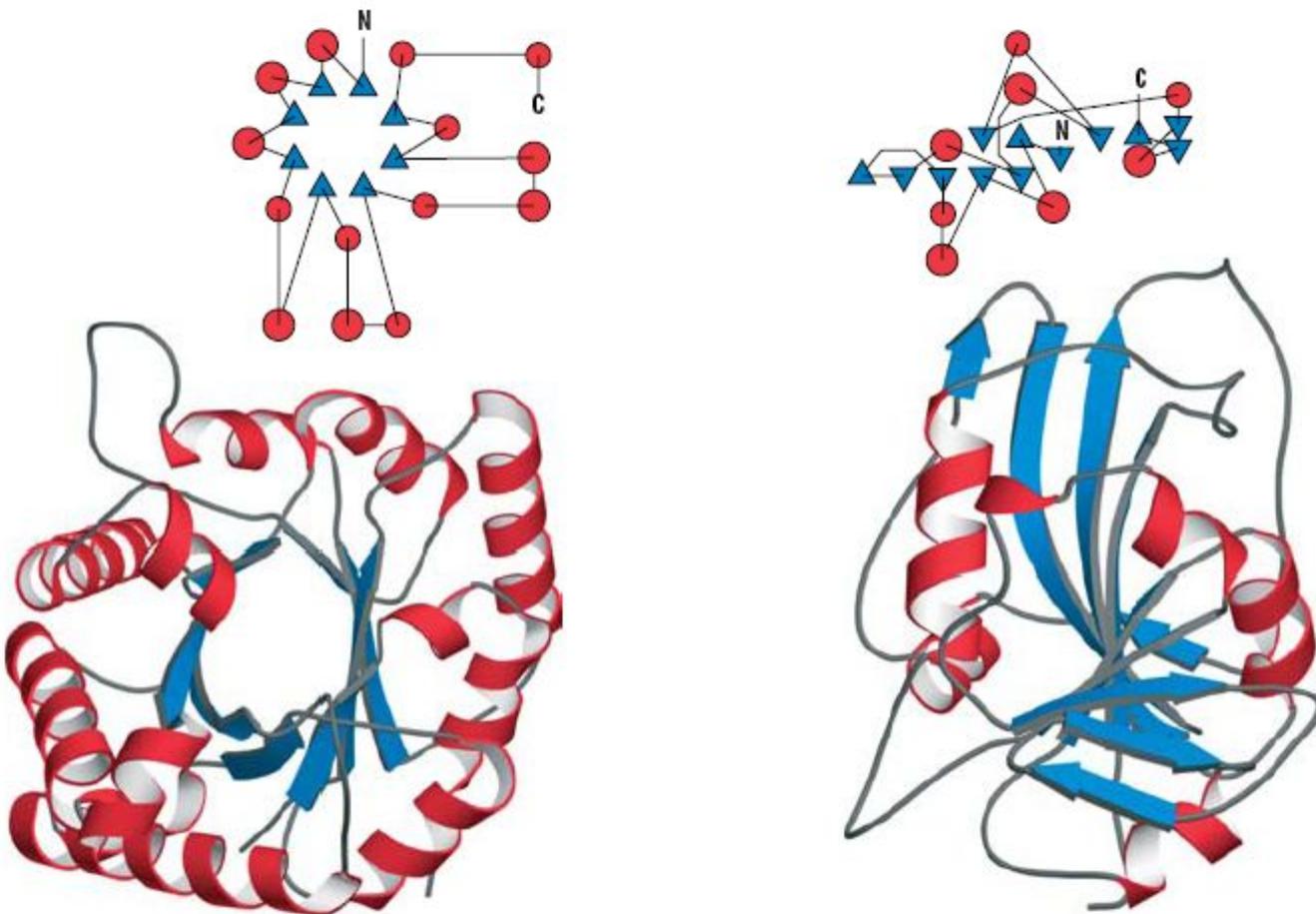


Globulares



fibrilares

**Topología:** forma global de la proteína teniendo en cuenta la conectividad entre los elementos de estructura secundaria

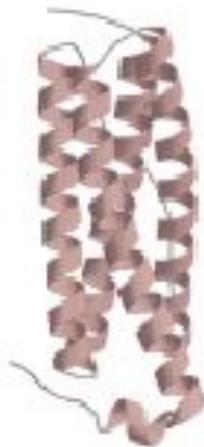


## Clasificación de los plegamientos según su tipo de estructura secundaria

All  $\alpha$



1a06  
Serum albumin  
Serum albumin  
Serum albumin  
Serum albumin  
Serum albumin  
Human (*Homo sapiens*)



1bcf  
Ferritin-like  
Ferritin-like  
Ferritin  
Bacterioferritin (cytochrome  $b_1$ )  
*Escherichia coli*



1gat  
 $\alpha/\alpha$  toroid  
Glycosyltransferases of the superhelical fold  
Glucoamylase  
Glucoamylase  
*Aspergillus awamori*, variant x100

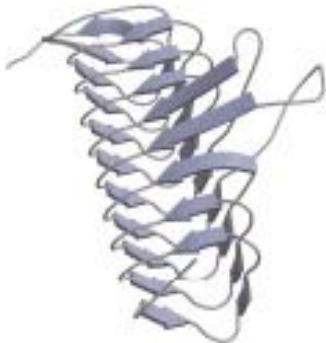


1enh  
DNA-binding 3-helical bundle  
Homeodomain-like  
Homeodomain  
*engrailed* Homeodomain  
*Drosophila melanogaster*

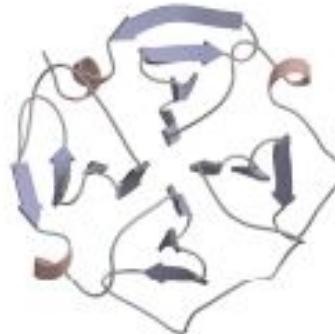
All  $\beta$



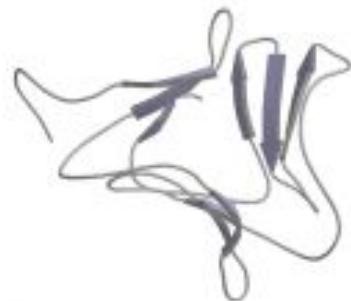
- 1hoc  
 $\alpha$ -Amylase inhibitor
- $\alpha$ -Amylase inhibitor
- $\alpha$ -Amylase inhibitor
- HOE-467A
- Streptomyces tendae 4158



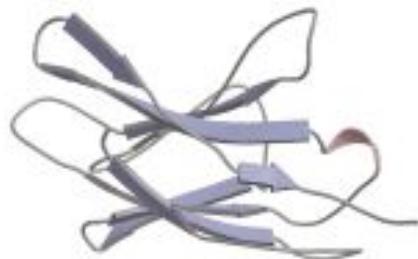
- 1lxa  
Single-stranded left-handed  $\beta$  helix
- Trimeric LpxA-like enzymes
- UDP N-acetylglucosamine acyltransferase
- UDP N-acetylglucosamine acyltransferase
- Escherichia coli



- 1pxr  
Four-bladed  $\beta$  propeller
- Hemopexin-like domain
- Hemopexin-like domain
- Collagenase-3 (MMP-13), carboxyl-terminal domain
- Human (*Homo sapiens*)

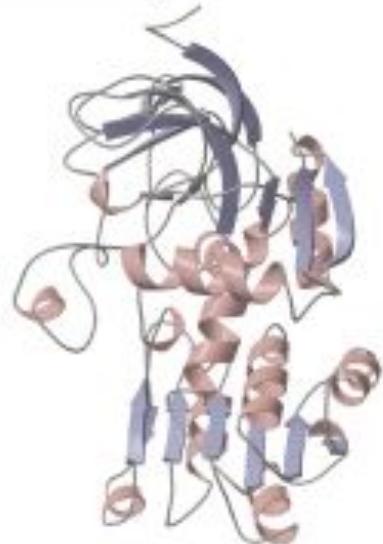


- 1jpc  
 $\beta$ -Prism II
- $\alpha$ -D-Mannose-specific plant lectins
- $\alpha$ -D-Mannose-specific plant lectins
- Lectin (agglutinin)
- Snowdrop (*Galanthus nivalis*)

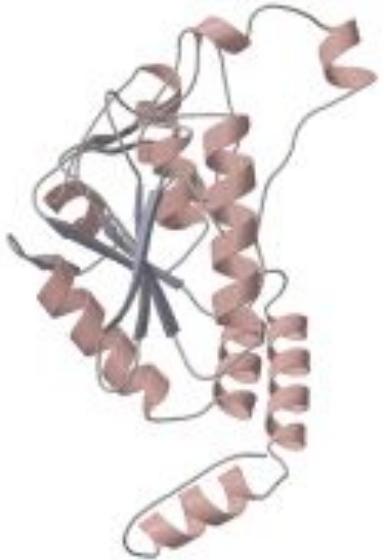


- 1cd8  
Immunoglobulin-like  $\beta$  sandwich
- Immunoglobulin
- Antibody variable domain-like
- CD8
- Human (*Homo sapiens*)

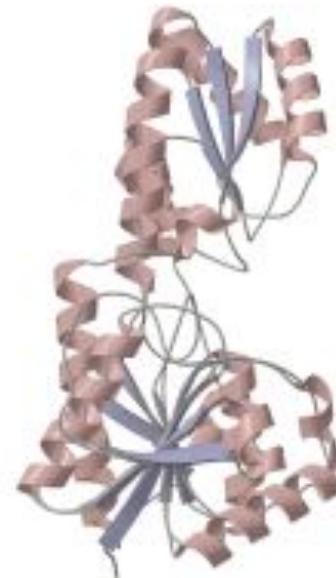
$\alpha\beta$



- 1deh  
NAD(P)-binding Rossmann-fold domains  
NAD(P)-binding Rossmann-fold domains
- Alcohol/glucose dehydrogenases,  
carboxyl-terminal domain
- Alcohol dehydrogenase
- Human (*Homo sapiens*)



- 1dub  
Cretonase-like  
Cretonase-like  
Cretonase-like
- Enoyl-CoA hydratase
- Rat (*Rattus norvegicus*)



- 1pfk  
Phosphofructokinase  
Phosphofructokinase  
Phosphofructokinase  
Phosphofructokinase
- Escherichia coli*

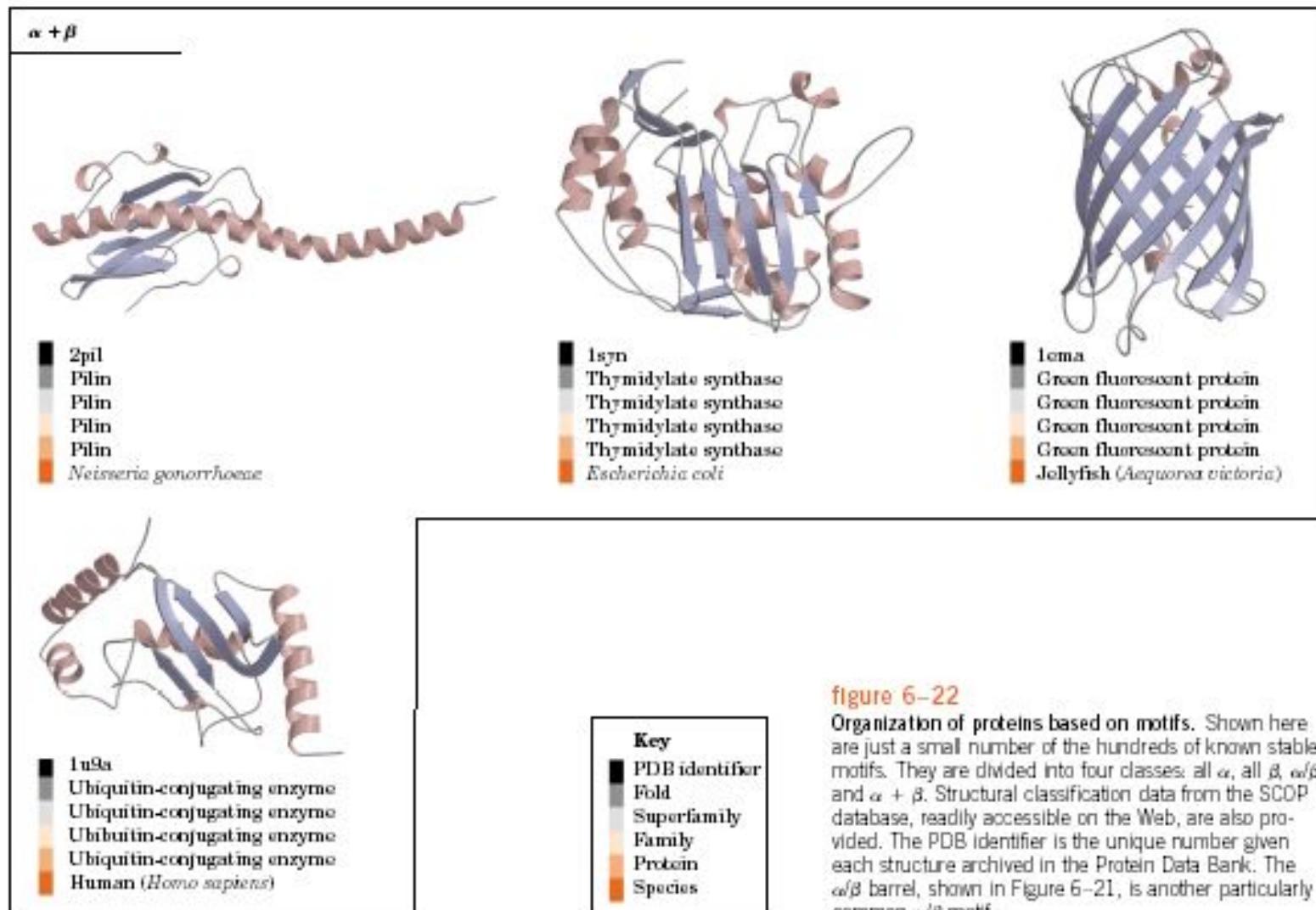
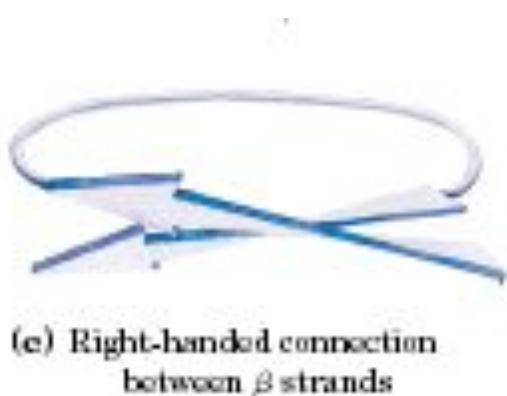


figure 6–22

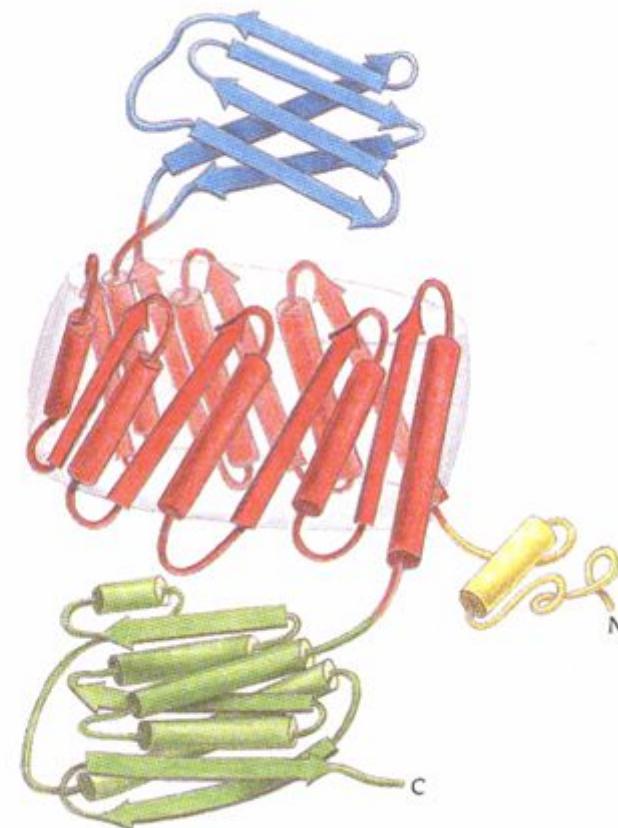
Organization of proteins based on motifs. Shown here are just a small number of the hundreds of known stable motifs. They are divided into four classes: all  $\alpha$ , all  $\beta$ ,  $\alpha/\beta$ , and  $\alpha + \beta$ . Structural classification data from the SCOP database, readily accessible on the Web, are also provided. The PDB identifier is the unique number given each structure archived in the Protein Data Bank. The  $\alpha/\beta$  barrel, shown in Figure 6–21, is another particularly common  $\alpha/\beta$  motif.

# Algunos aspectos sobre la estructura terciaria

## Motivos estructurales



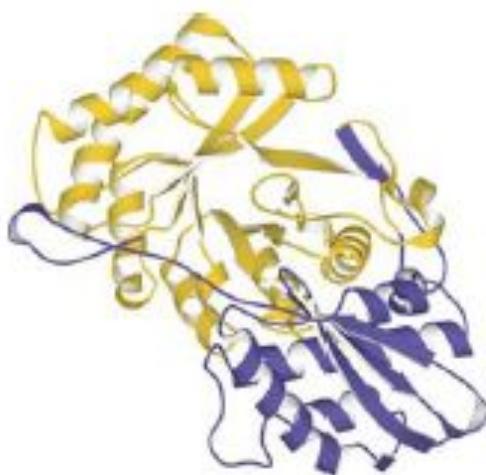
# Dominios

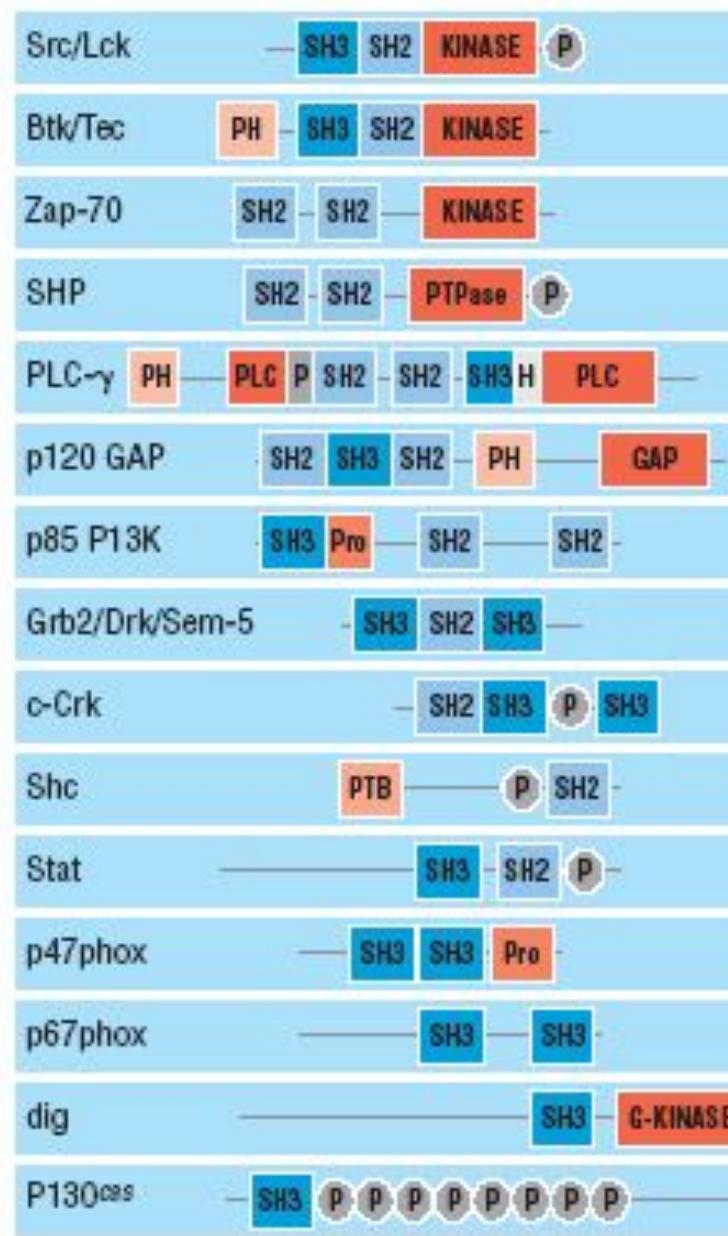


(a)

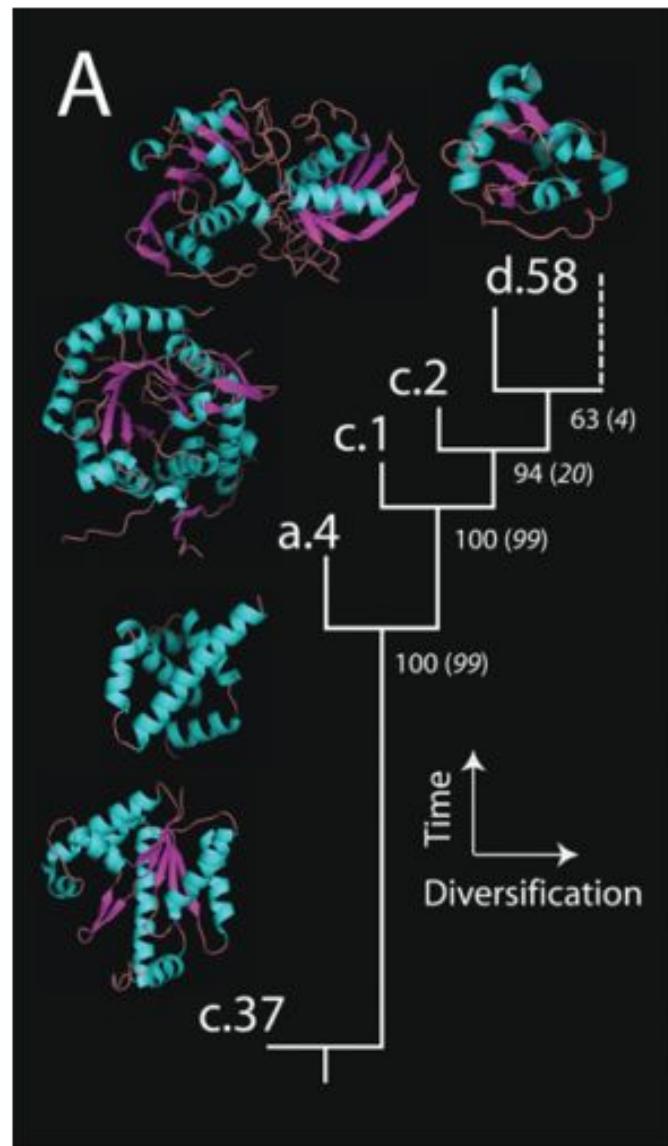


(b)

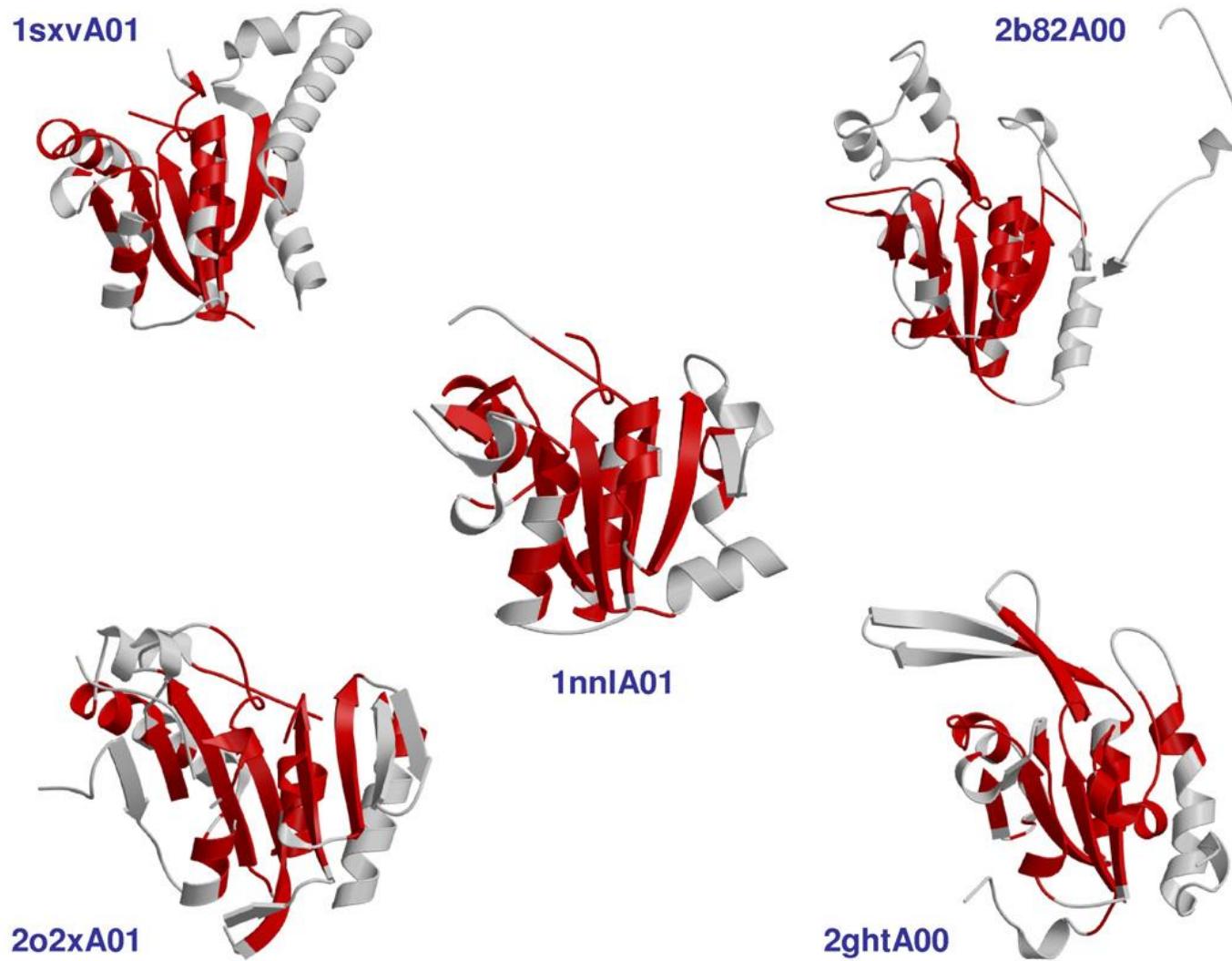




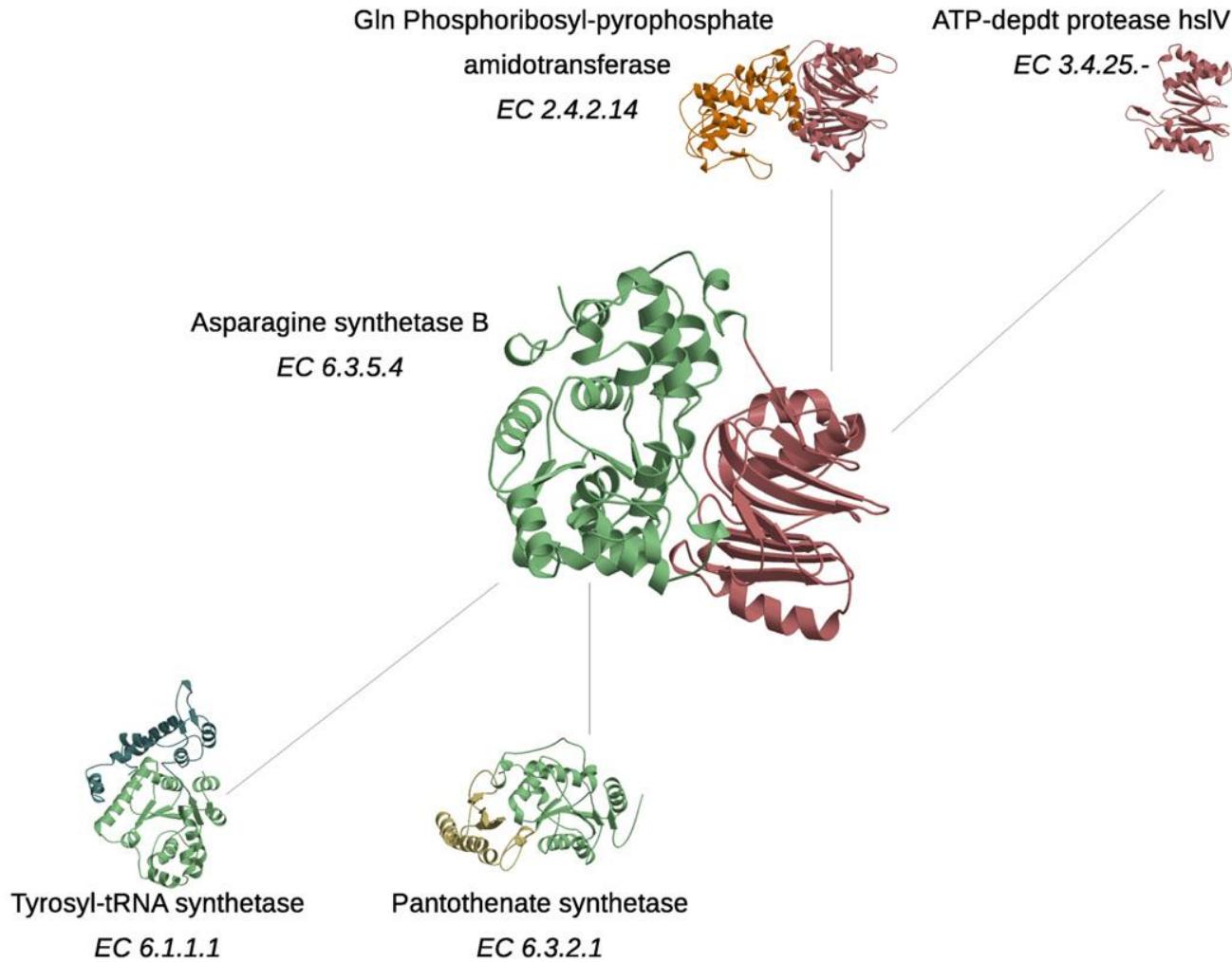
# Divergencia estructural



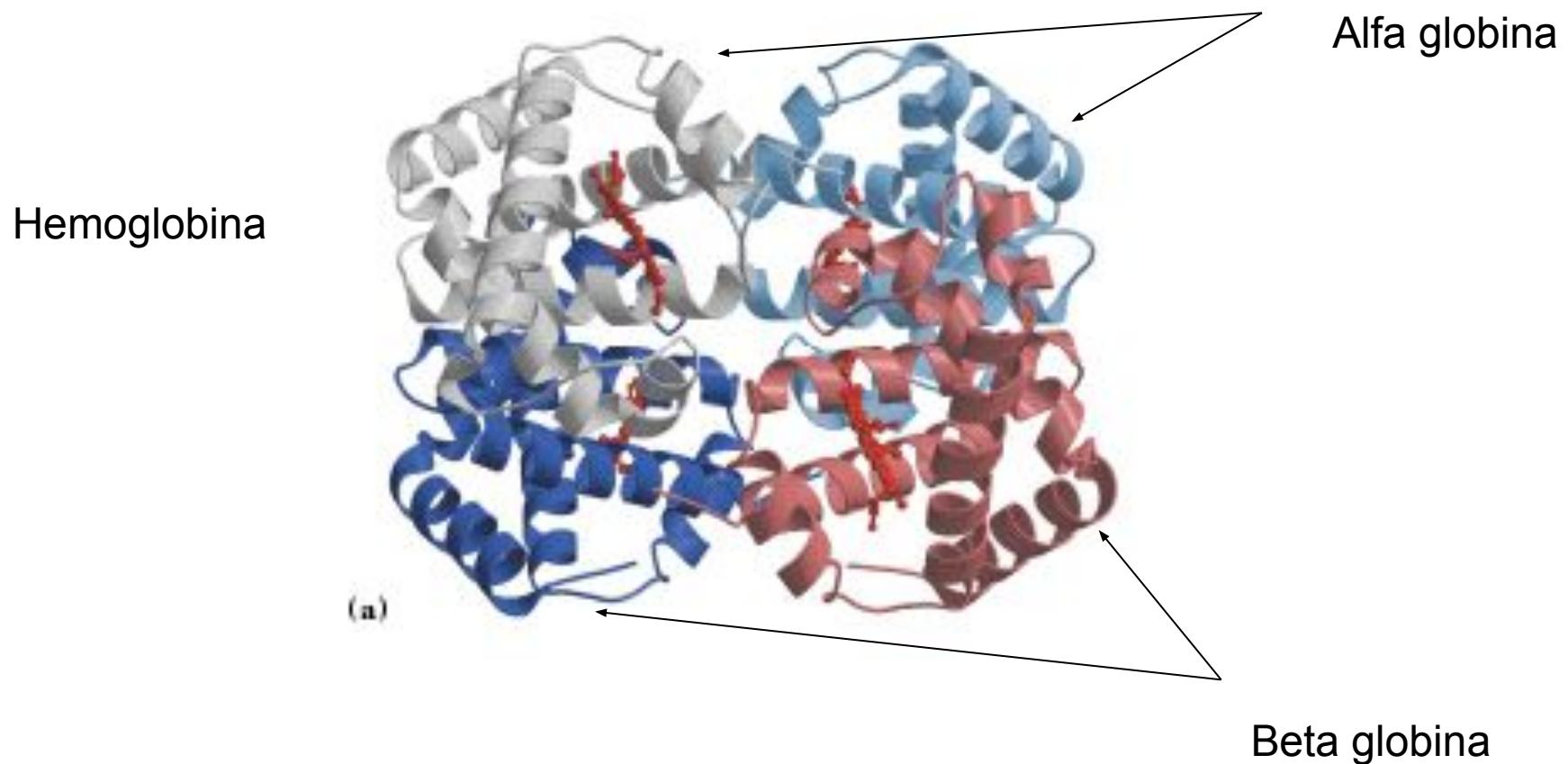
# Embellishment: evolución de proteínas y adaptación funcional



# Actividad, función y dominios

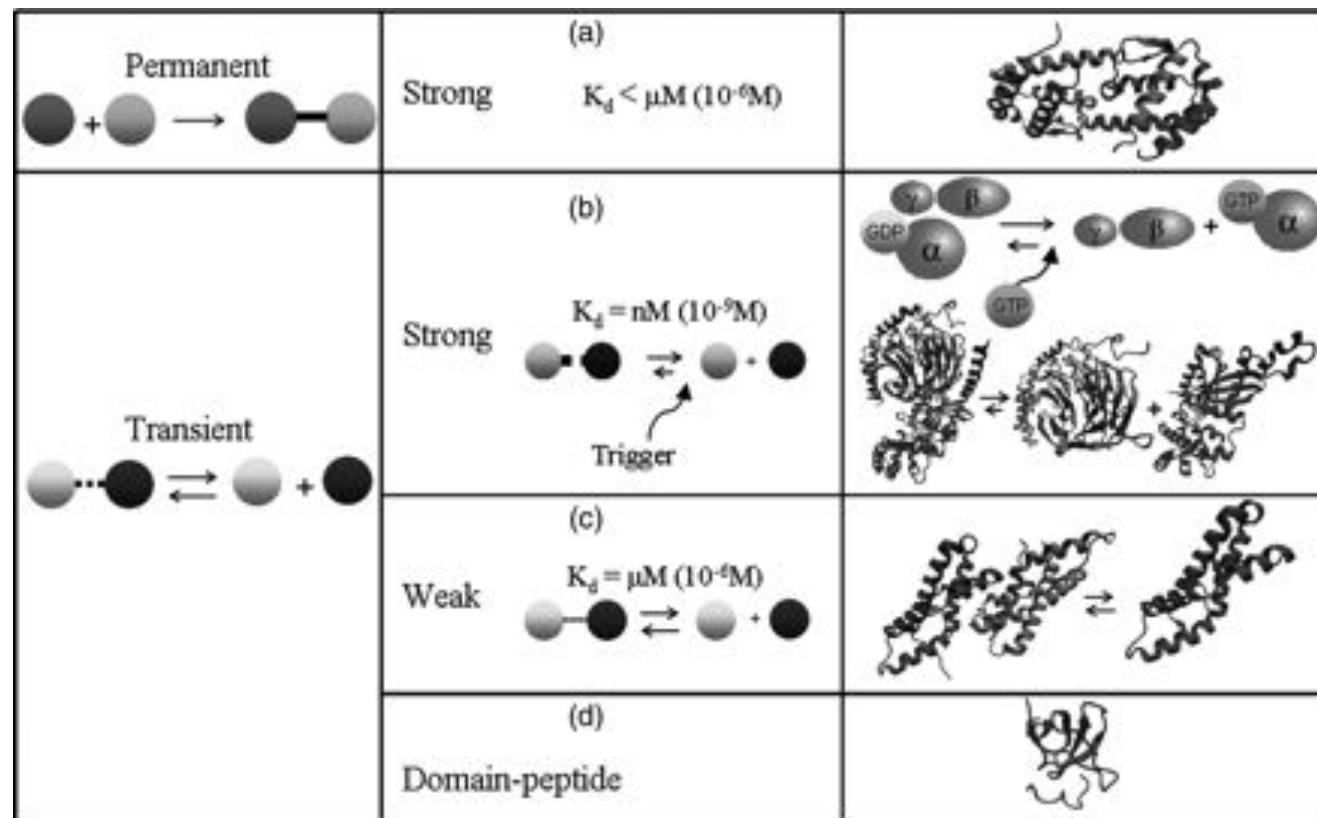


**Estructura cuaternaria:** estructura derivada de la interacción entre distintas cadenas de proteínas.

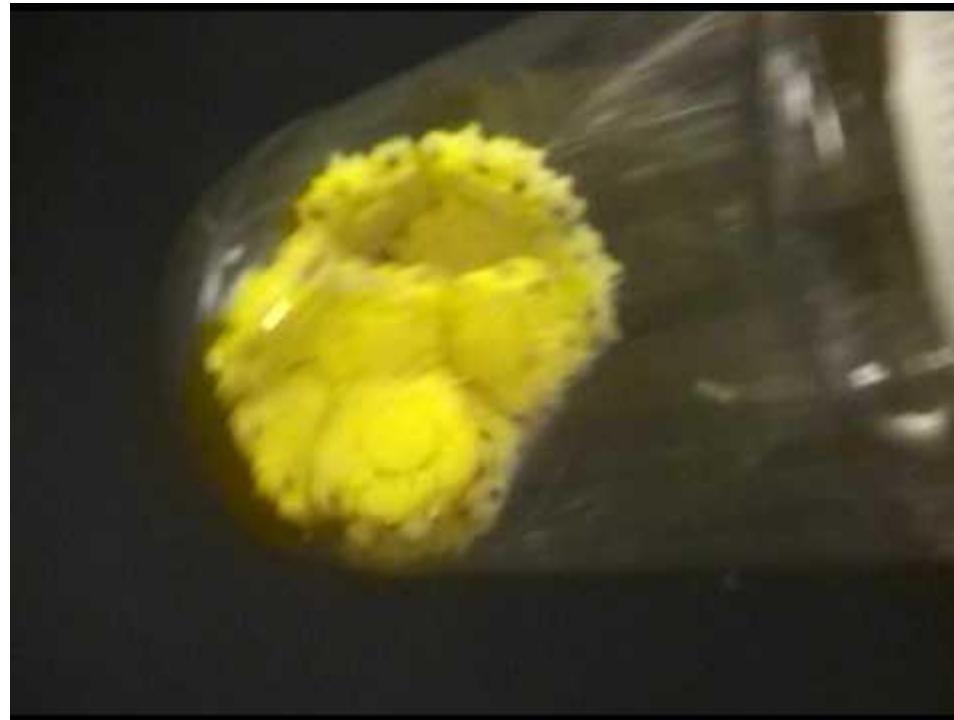


Oligomero: estructura cuaternaria donde las distintas cadenas no tienen función biológica en forma independiente

Complejos transientes: estructura cuaternaria donde las distintas cadenas se unen para formar una estructura transitoria, cumplir una determinada función y luego disociarse. Las cadenas (proteínas) que forman este complejo tienen funciones biológicas independientes al estado asociado.



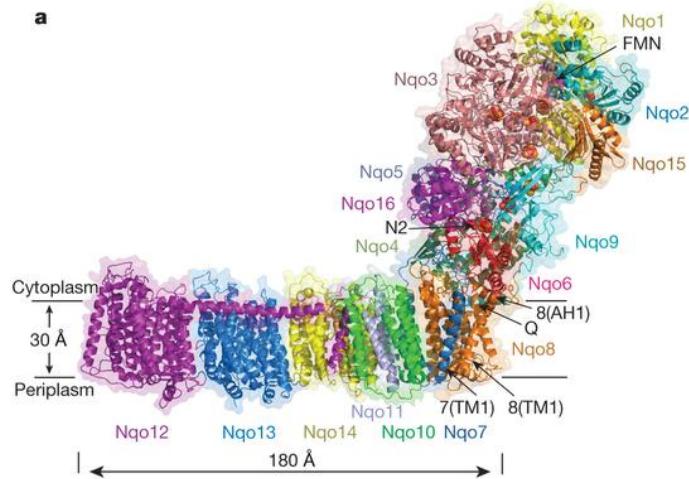
Las proteínas oligoméricas cumplen con el experimento de Anfinsen?



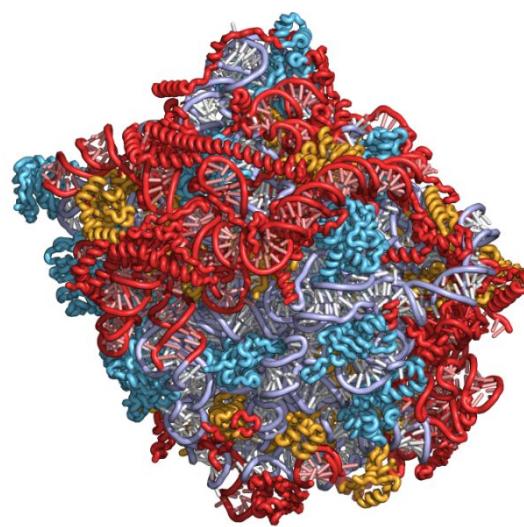
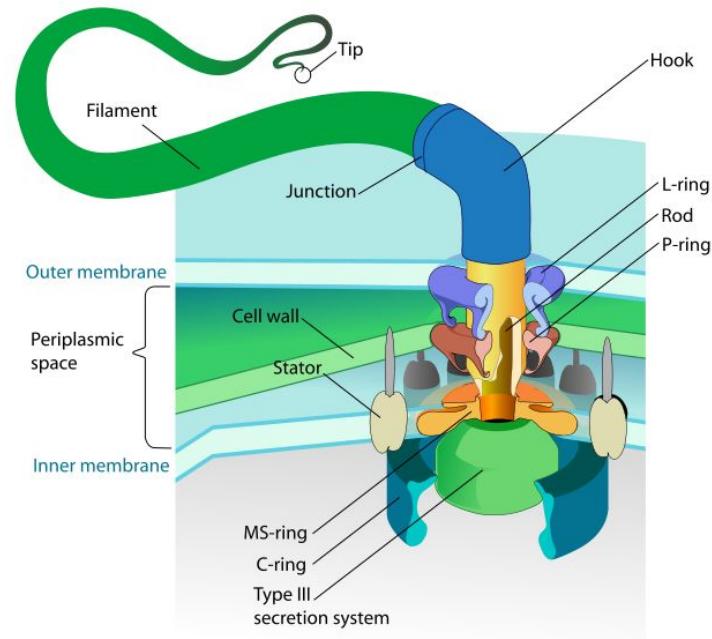
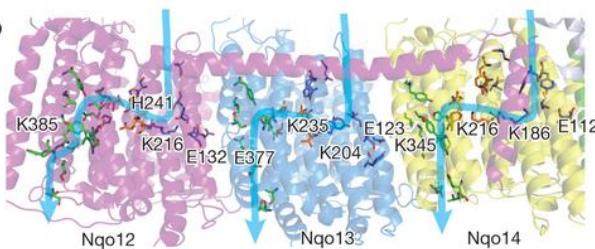
<https://www.youtube.com/watch?v=X-8MP7g8XOE>

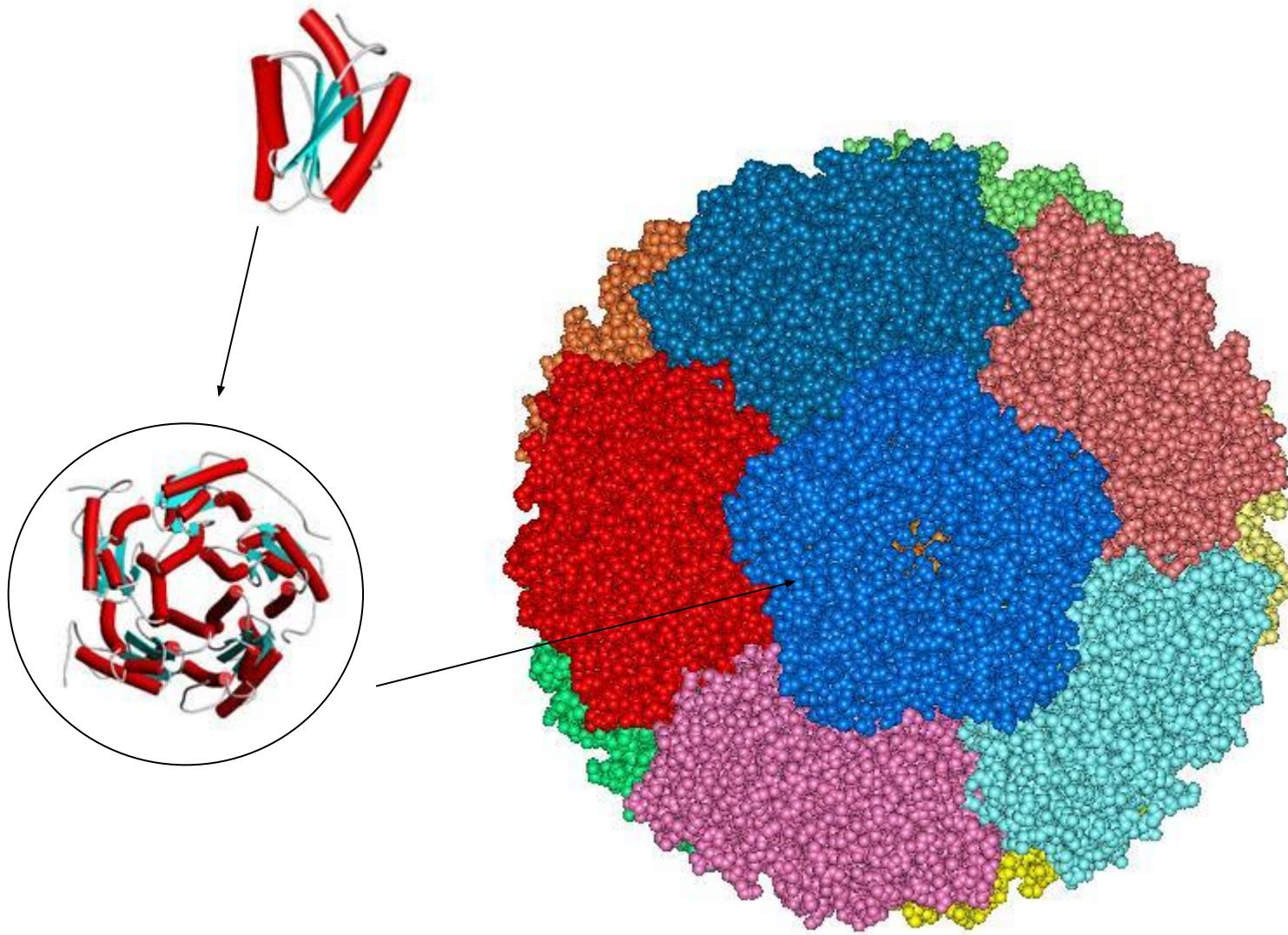
# Complejos supramacromoleculares

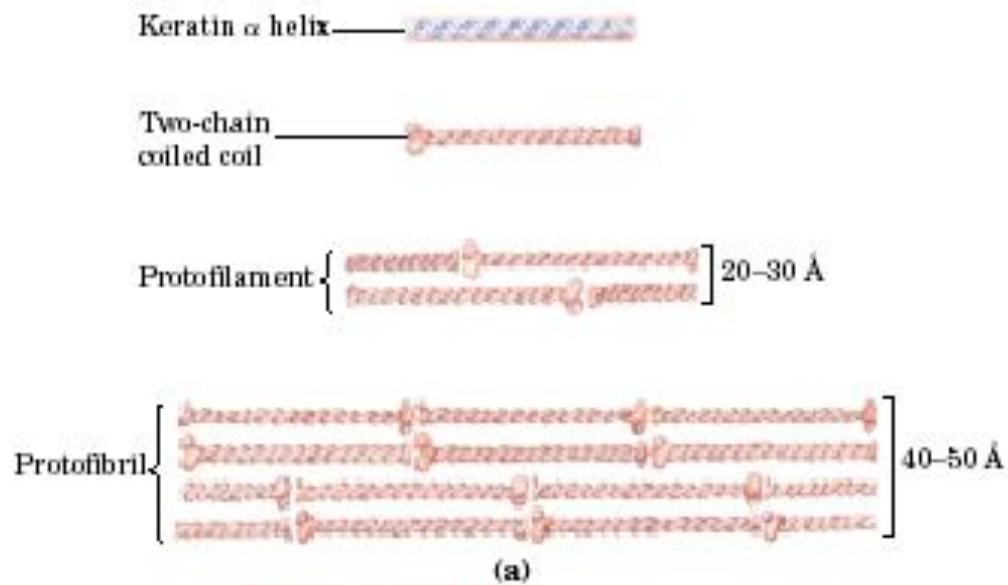
a



b

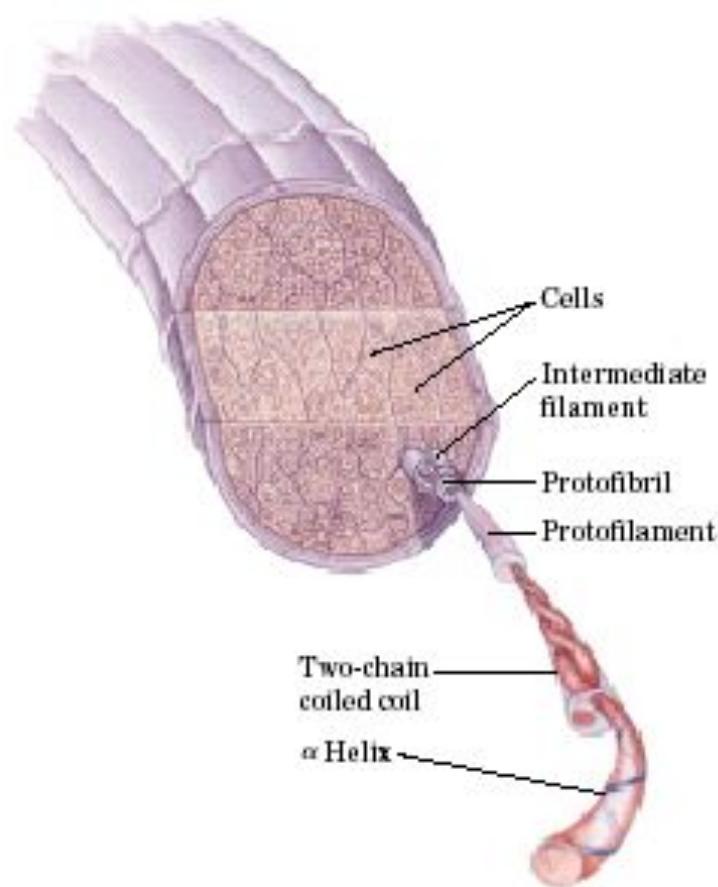






**figure 6–11**

**Structure of hair.** (a) Hair  $\alpha$ -keratin is an elongated  $\alpha$  helix with somewhat thicker elements near the amino and carboxyl termini. Pairs of these helices are interwound in a left-handed sense to form two-chain coiled coils. These then combine in higher-order structures called protofilaments and protofibrils. About four protofibrils—32 strands of  $\alpha$ -keratin altogether—combine to form an intermediate filament. The individual two-chain coiled coils in the various substructures also appear to be interwound, but the handedness of the interwinding and other structural details are unknown. (b) A hair is an array of many  $\alpha$ -keratin filaments, made up of the substructures shown in (a).



**Cross section of a hair**  
(b)

# Función vs Actividad

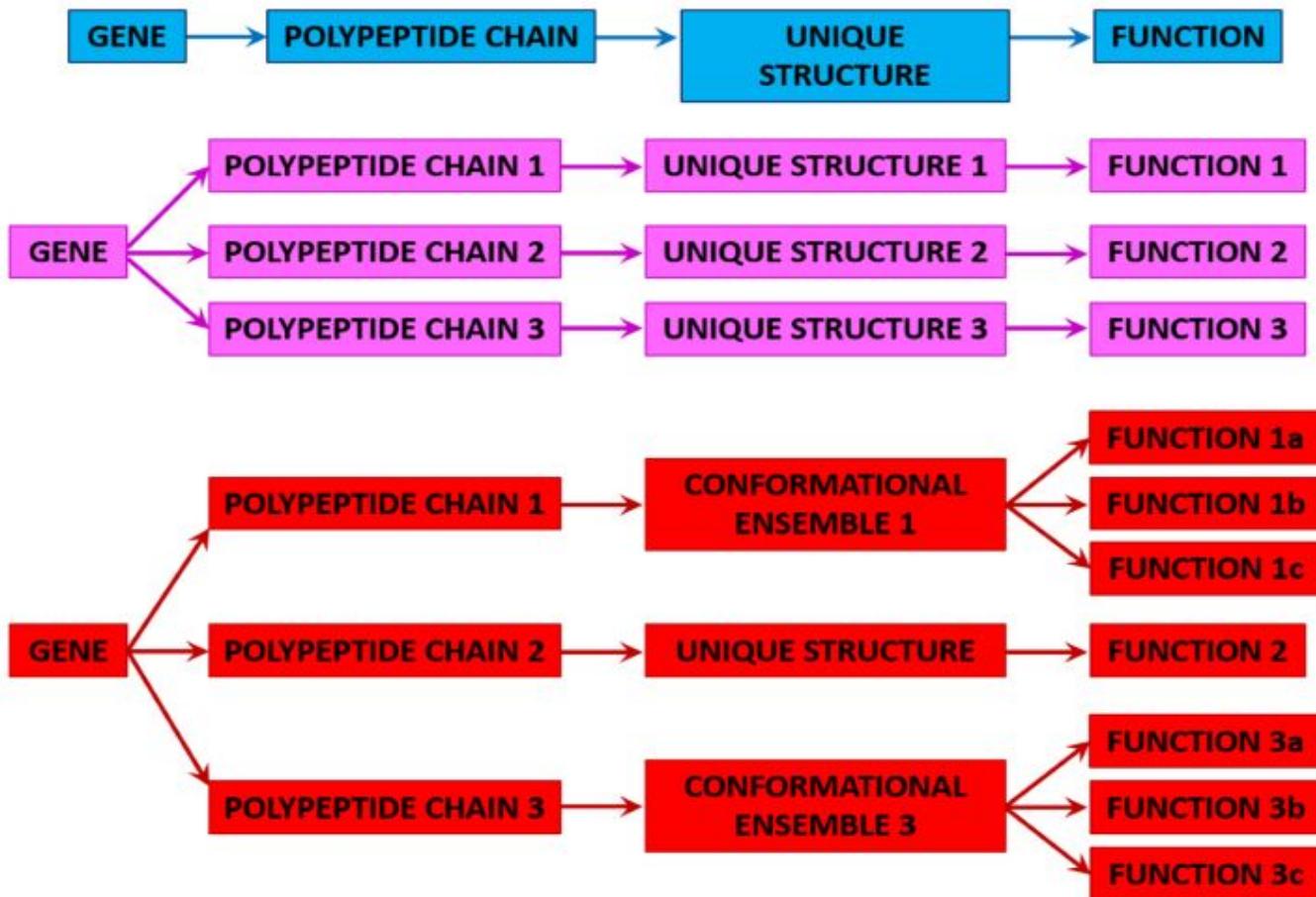
No tenemos que confundir el nombre del producto del gen, con la función o habilidad que tenga la proteína

Ej. empresarial

**Producto génico:** secretaria

**Funciones:** tipeo de datos, sacar fotocopias, hacer café, organizar eventos..

Hemoglobina y Mioglobina tienen la misma “actividad”: unir oxígeno  
Pero las dos tienen “funciones” muy distintas.



 Gene Ontology Consortium

Home Documentation Downloads User stories Community Tools About Contact us

## Search GO data

terms and gene products

Search

## Enrichment analysis (beta)

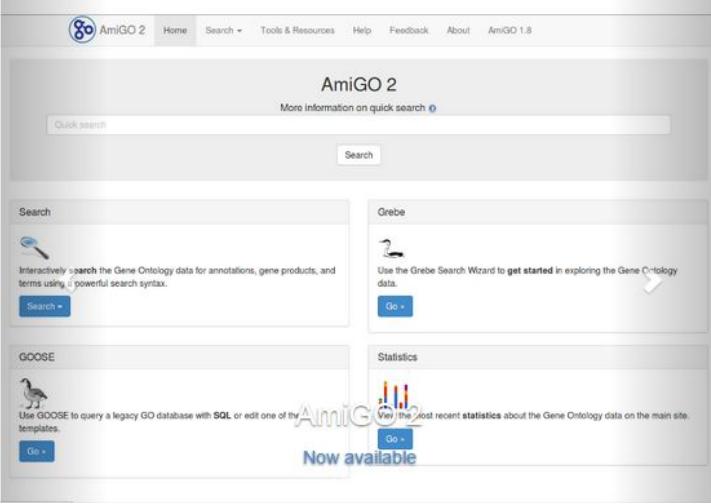
Your gene IDs here...

biological process ▾

H. sapiens ▾

Submit

# Gene Ontology Consortium



The screenshot shows the AmiGO 2 homepage. At the top, there's a navigation bar with links for Home, Search, Tools & Resources, Help, Feedback, About, and AmiGO 1.8. Below the navigation is a search bar with a "Quick search" placeholder and a "Search" button. To the right of the search bar is a "Grebe" section featuring a search icon and a link to the Grebe Search Wizard. Further down are sections for "GOOSE" (with a search icon) and "Statistics" (with a bar chart icon). A banner at the bottom reads "Now available".

Search 

## Highlighted GO term

Representing "phases" in GO biological process

The GOC has recently introduced a new term **biological phase** (GO:0044848), as a direct subclass of biological process. This class represents a distinct period or stage during which biological processes can occur.

more

## Random FAQs

- I have a question about gene or

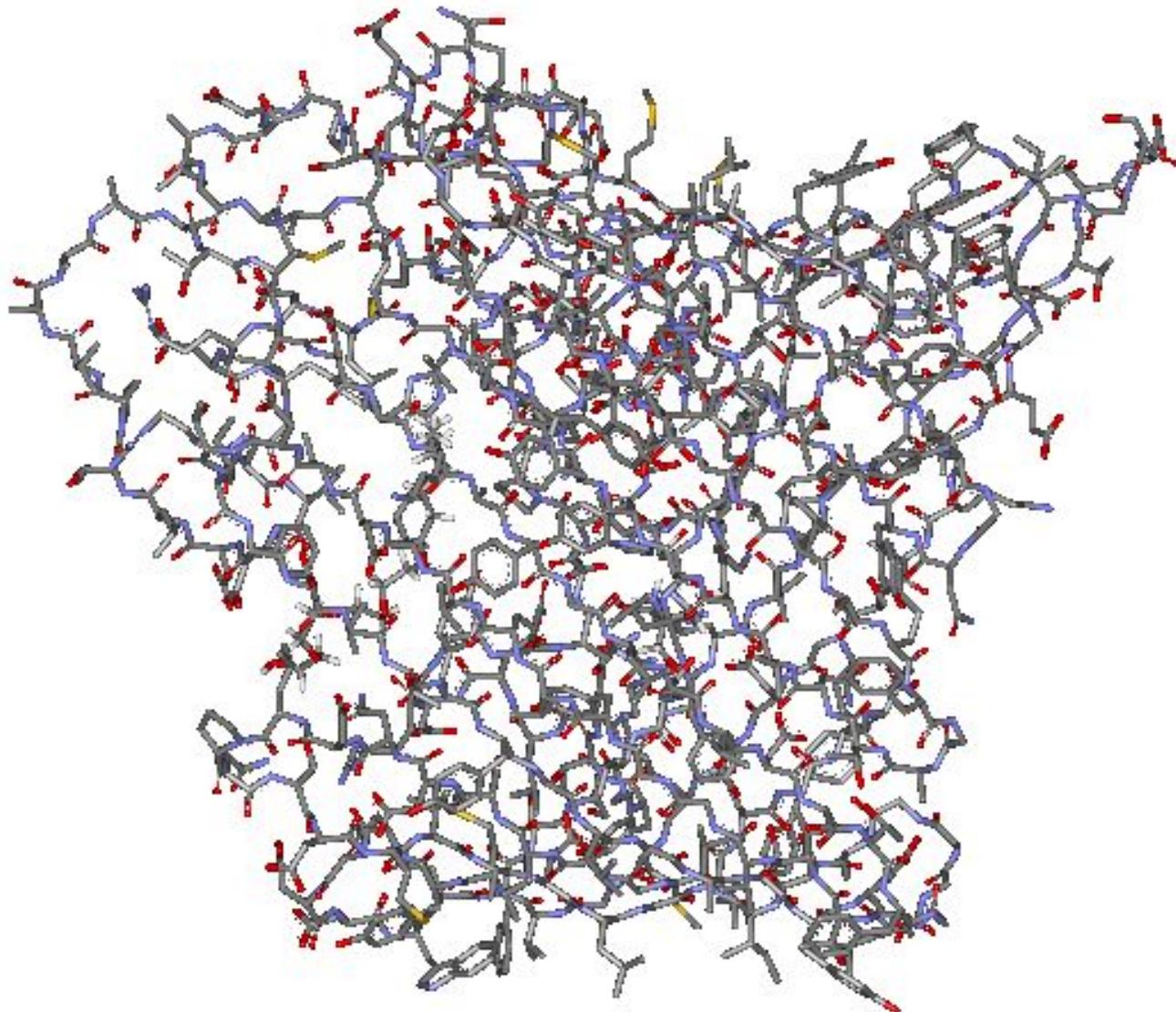
## GO terms: Molecular procesos, molecular function y celular component

Database	Gene Product ID	Symbol	GO Identifier	GO Term Name
Process				
UniProtKB	P69905	HBA1	<a href="#">GO:0006810</a>	transport
UniProtKB	P69905	HBA1	<a href="#">GO:0006898</a>	receptor-mediated endocytosis
UniProtKB	P69905	HBA1	<a href="#">GO:0010942</a>	positive regulation of cell death
UniProtKB	P69905	HBA1	<a href="#">GO:0015671</a>	oxygen transport
UniProtKB	P69905	HBA1	<a href="#">GO:0015671</a>	oxygen transport
UniProtKB	P69905	HBA1	<a href="#">GO:0015671</a>	oxygen transport
UniProtKB	P69905	HBA1	<a href="#">GO:0015701</a>	bicarbonate transport
UniProtKB	P69905	HBA1	<a href="#">GO:0015701</a>	bicarbonate transport
UniProtKB	P69905	HBA1	<a href="#">GO:0015701</a>	bicarbonate transport
UniProtKB	P69905	HBA1	<a href="#">GO:0042542</a>	response to hydrogen peroxide
UniProtKB	P69905	HBA1	<a href="#">GO:0042744</a>	hydrogen peroxide catabolic process
UniProtKB	P69905	HBA1	<a href="#">GO:0044281</a>	small molecule metabolic process
UniProtKB	P69905	HBA1	<a href="#">GO:0051291</a>	protein heterooligomerization
UniProtKB	P69905	HBA1	<a href="#">GO:0055114</a>	oxidation-reduction process

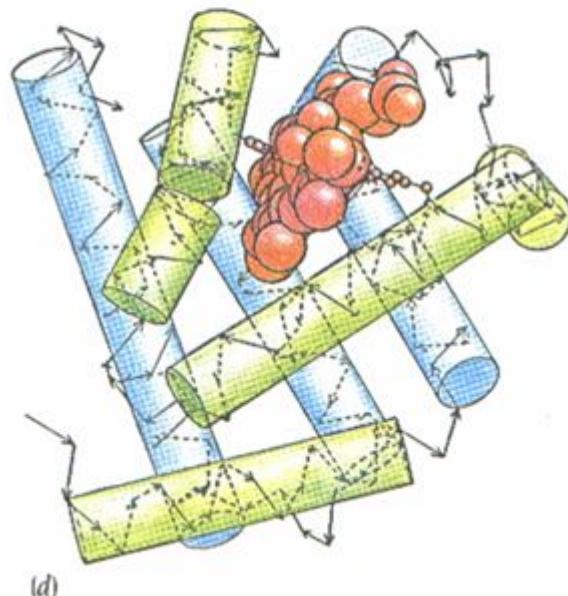
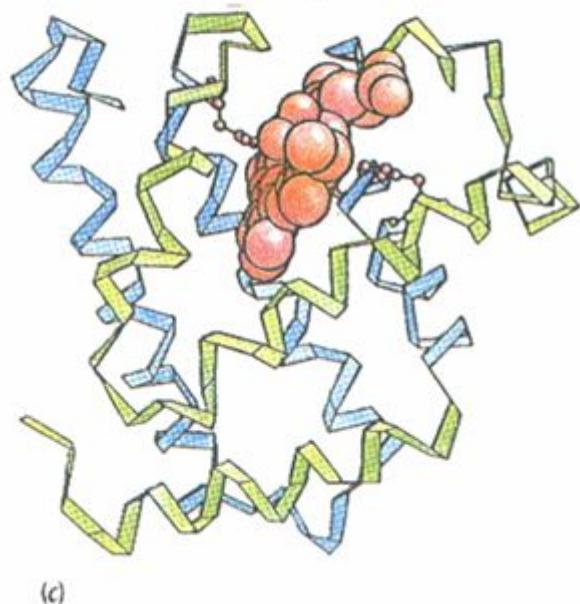
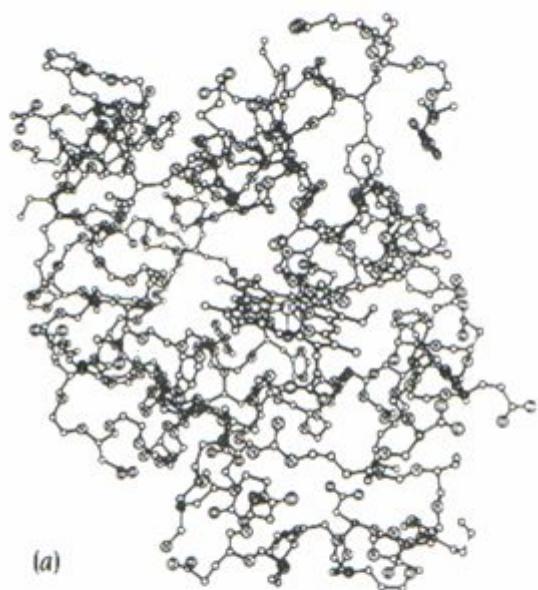
Function				
UniProtKB	P69905	HBA1	<a href="#">GO:0005344</a>	oxygen transporter activity
UniProtKB	P69905	HBA1	<a href="#">GO:0005506</a>	iron ion binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0005515</a>	protein binding
UniProtKB	P69905	HBA1	<a href="#">GO:0019825</a>	oxygen binding
UniProtKB	P69905	HBA1	<a href="#">GO:0020037</a>	heme binding
UniProtKB	P69905	HBA1	<a href="#">GO:0046872</a>	metal ion binding
UniProtKB	P69905	HBA1	<a href="#">GO:0004601</a>	peroxidase activity
UniProtKB	P69905	HBA1	<a href="#">GO:0031720</a>	haptoglobin binding

				Component
UniProtKB	P69905	HBA1	<a href="#">GO:0005576</a>	extracellular region
UniProtKB	P69905	HBA1	<a href="#">GO:0005576</a>	extracellular region
UniProtKB	P69905	HBA1	<a href="#">GO:0005576</a>	extracellular region
UniProtKB	P69905	HBA1	<a href="#">GO:0005576</a>	extracellular region
UniProtKB	P69905	HBA1	<a href="#">GO:0005576</a>	extracellular region
UniProtKB	P69905	HBA1	<a href="#">GO:0005829</a>	cytosol
UniProtKB	P69905	HBA1	<a href="#">GO:0005829</a>	cytosol
UniProtKB	P69905	HBA1	<a href="#">GO:0005833</a>	hemoglobin complex
UniProtKB	P69905	HBA1	<a href="#">GO:0005833</a>	hemoglobin complex
UniProtKB	P69905	HBA1	<a href="#">GO:0005833</a>	hemoglobin complex
UniProtKB	P69905	HBA1	<a href="#">GO:0016020</a>	membrane
UniProtKB	P69905	HBA1	<a href="#">GO:0022627</a>	cytosolic small ribosomal subunit
UniProtKB	P69905	HBA1	<a href="#">GO:0031838</a>	haptoglobin-hemoglobin complex
UniProtKB	P69905	HBA1	<a href="#">GO:0070062</a>	extracellular exosome
UniProtKB	P69905	HBA1	<a href="#">GO:0070062</a>	extracellular exosome
UniProtKB	P69905	HBA1	<a href="#">GO:0070062</a>	extracellular exosome
UniProtKB	P69905	HBA1	<a href="#">GO:0070062</a>	extracellular exosome
UniProtKB	P69905	HBA1	<a href="#">GO:0071682</a>	endocytic vesicle lumen
UniProtKB	P69905	HBA1	<a href="#">GO:0072562</a>	blood microparticle

# Relación estructura-función



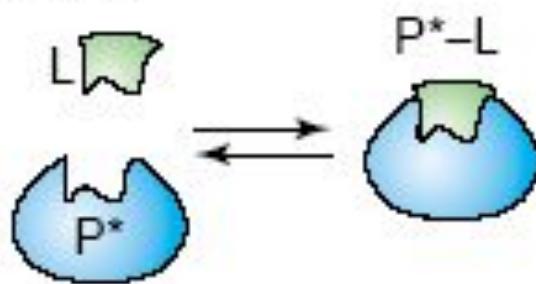
**Figure 2.9** (a) Two-dimensional representation of the structure of myoglobin showing all atoms by straight lines. (b) Stereoscopic representation of the surface of myoglobin. (c) Two-dimensional representation of the structure of myoglobin displaying all atoms by spheres. (d) Three-dimensional ribbon diagram of the structure of myoglobin showing different degrees of backbone flexibility.



# Relación Estructura-Función

## Teoría “Llave-cerradura”

(ii) Lock and key

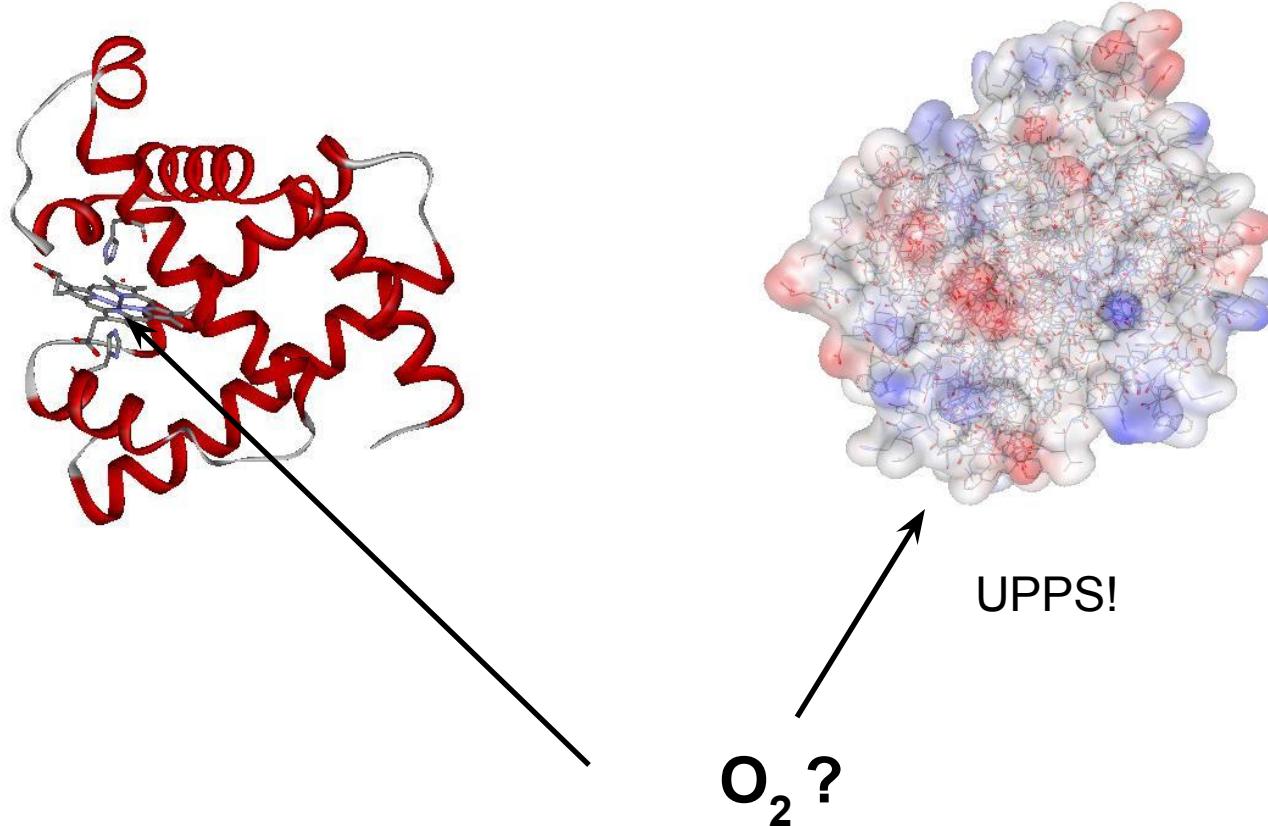


*The restricted effects of the enzymes may therefore be explained by the assumption that the approach of the molecules that cause the chemical process can occur only in the case of a similar geometric shape.*

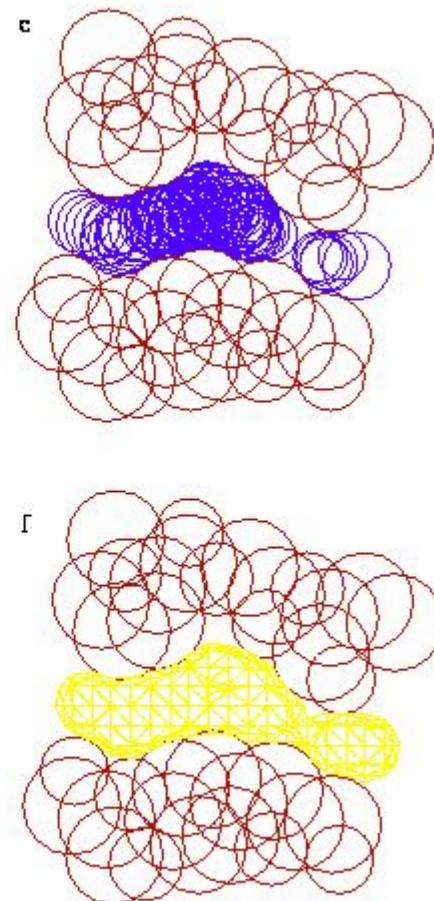
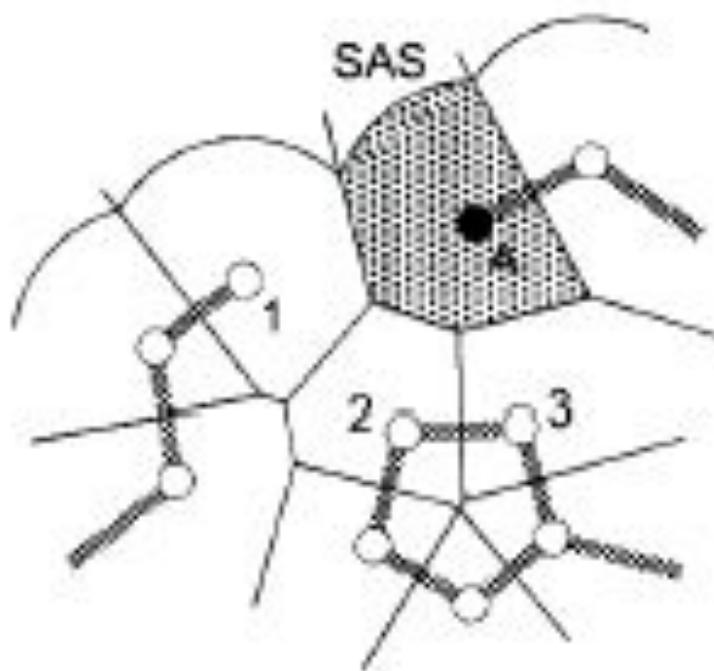
*To use a picture, I would like to say that enzyme and glucoside have to fit to each other like a lock and key in order to exert a chemical effect on each other.*

E. Fischer, Ber., 27 (1894) 3189–3232.

Cómo interacciona el oxígeno con el hemo en la mioglobina?

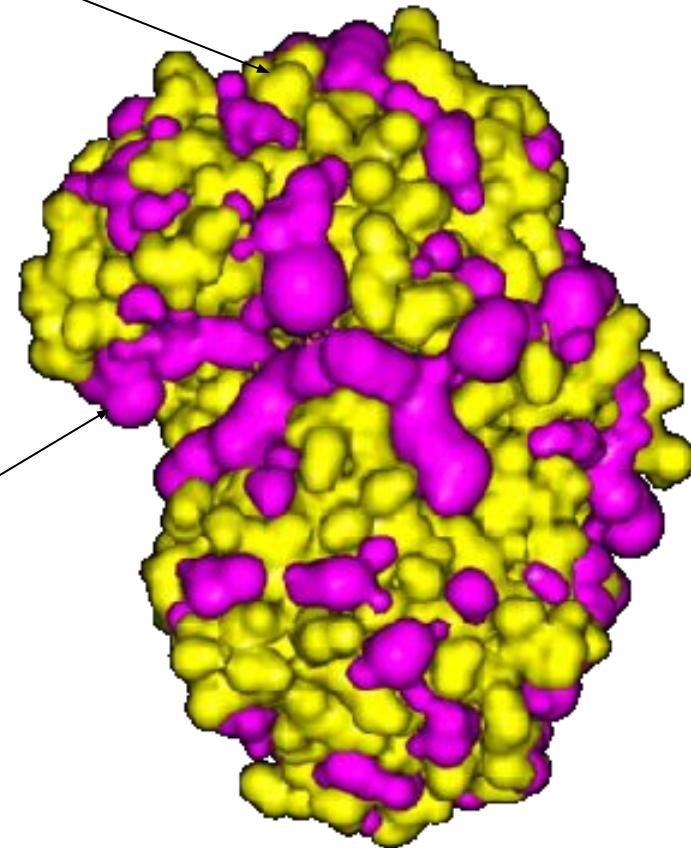


# Superficies y cavidades



Superficie

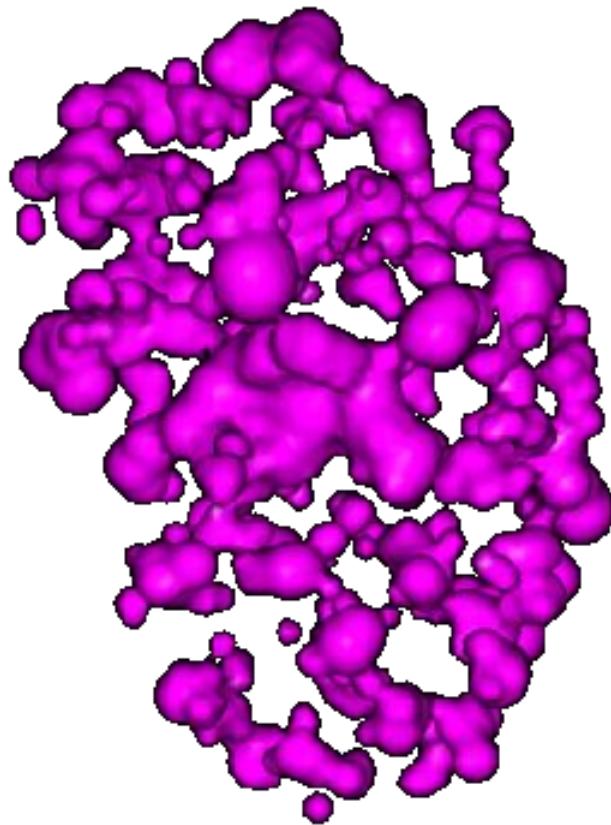
(amarillo)



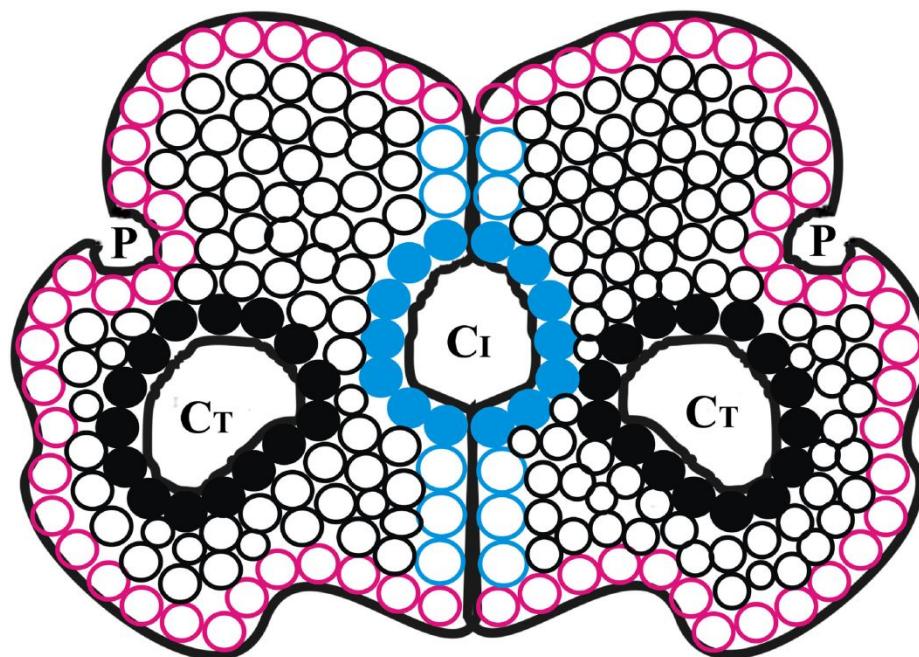
Surcos

(fucsia)

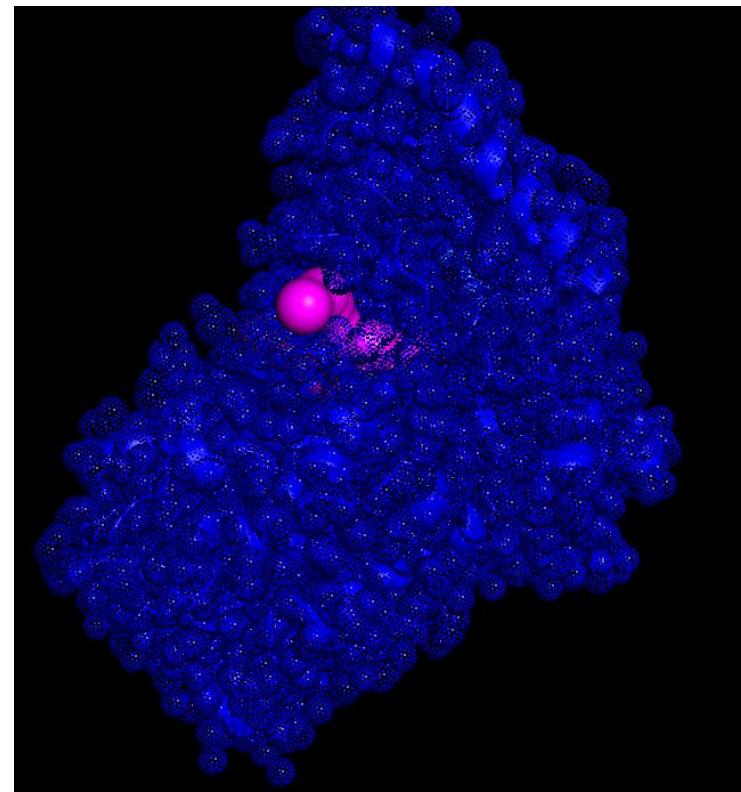
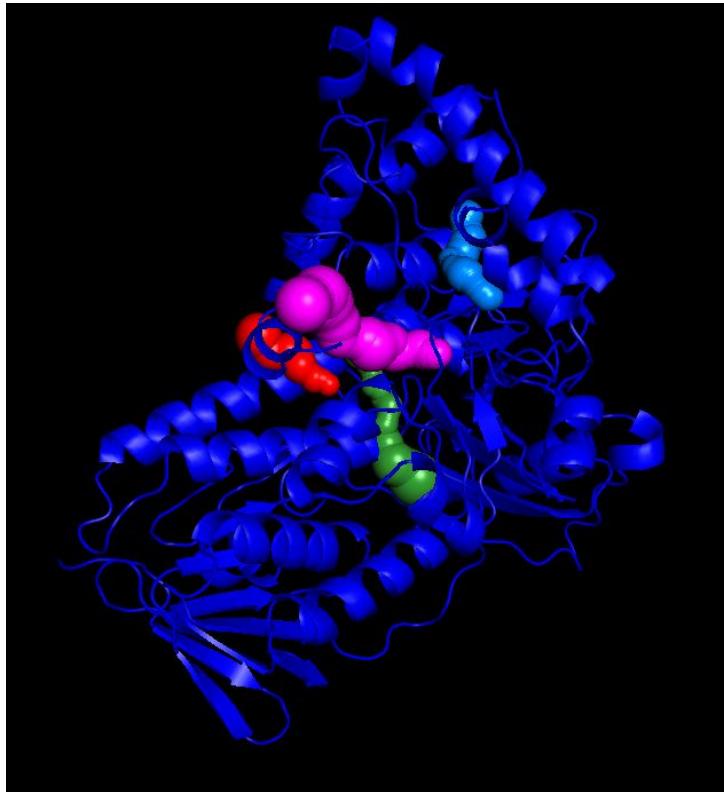
**Los surcos conducen a cavidades...**



Voids= Son cavidades que no tienen contacto con el solvente



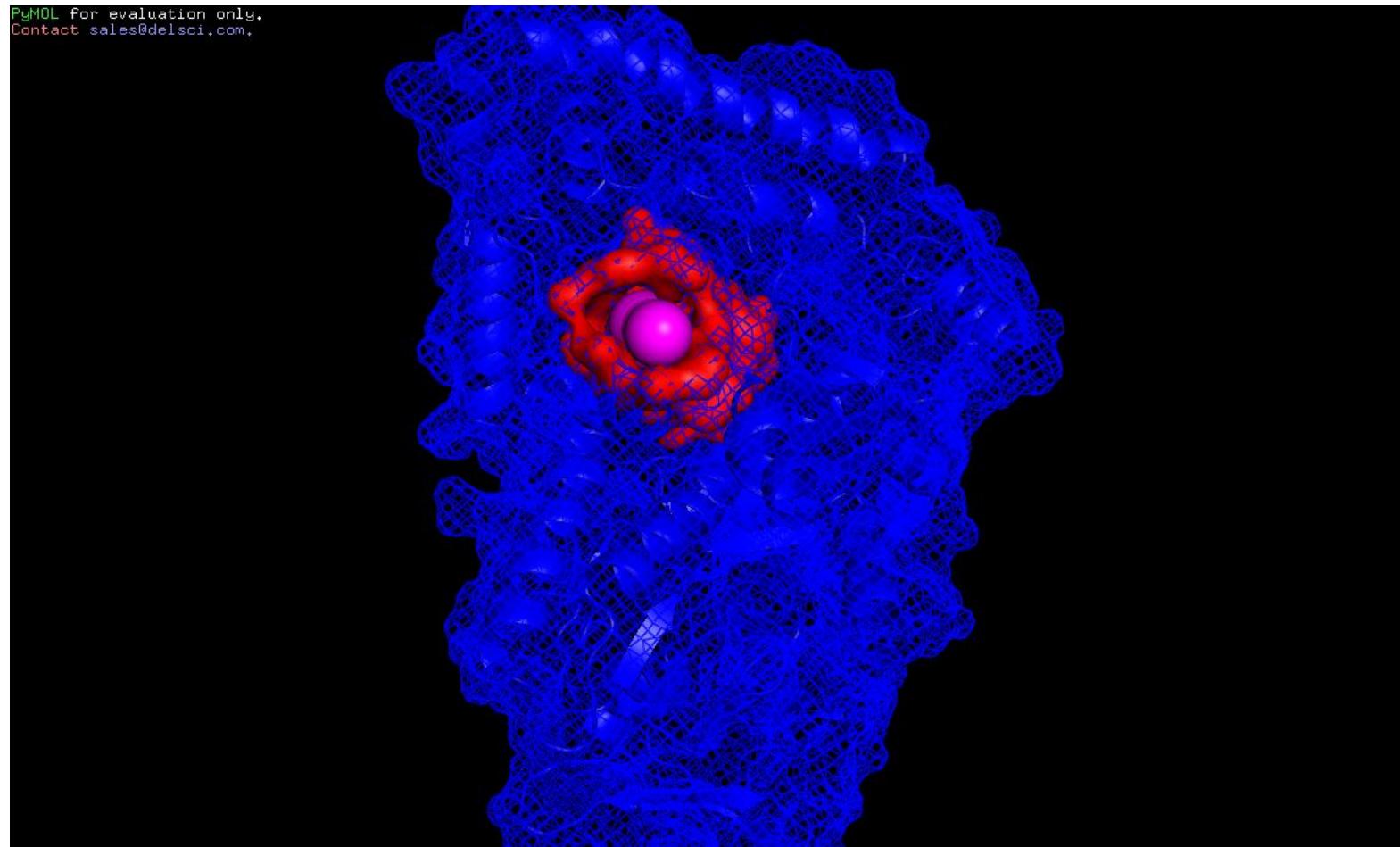
**Túneles:** conductos que comunican la superficie con el interior de la proteína. Pueden también conectar distintas cavidades.



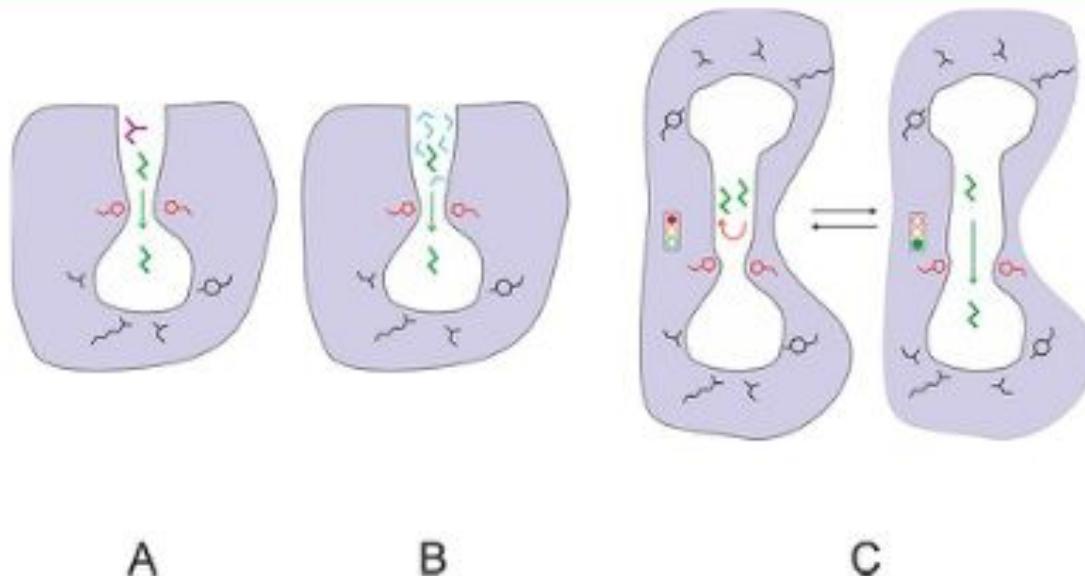
## Función de los túneles

1. Facilitan el acceso de los sustratos desde el solvente al sitio activo (en cavidad) de una enzima/proteína
2. Impiden que otros posibles ligandos (menos específicos) entre al sitio activo de la enzima y generen productos no activos
3. Previenen la liberación de productos tóxicos a la célula (como intermediarios de reacción en una reacción catalizada por una enzima)
4. Facilitan las reacciones donde el agua podría impedirla
5. Sincronizan reacciones complejas (por ejemplo aquellas que requieran cofactores y/o más de un sustrato)

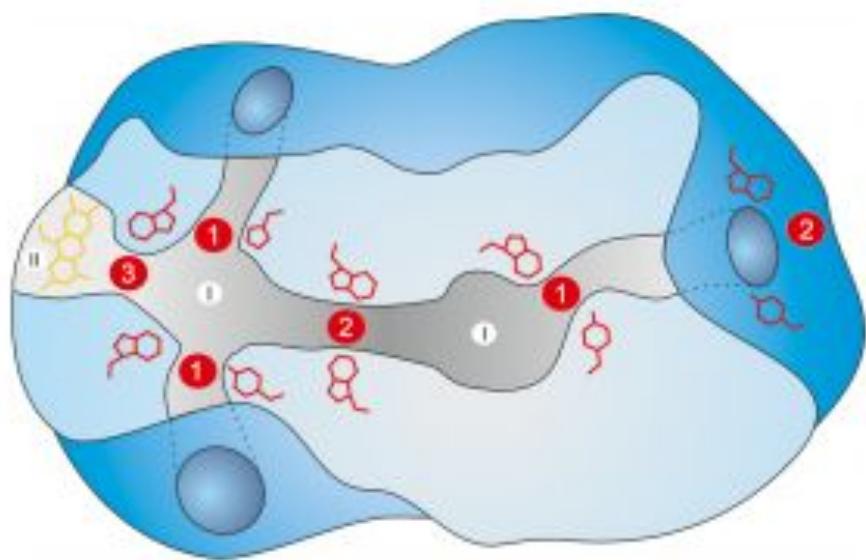
**Cavities:** las cavidades pueden o no estar asociadas a túneles. En general pueden alojar el sitio activo de la proteína



## Gate keepers: controlling the in/out transit in a protein



**Figure 2.** Schematic illustration of the molecular functions of protein gates: (A) control of substrate access, (B) control of solvent access, (C) control and synchronization of reactions. Protein is represented by the area colored in gray, active site cavity by the area in white, gating residues by red lines, substrate molecules by green or violet lines, and water molecules by blue lines.

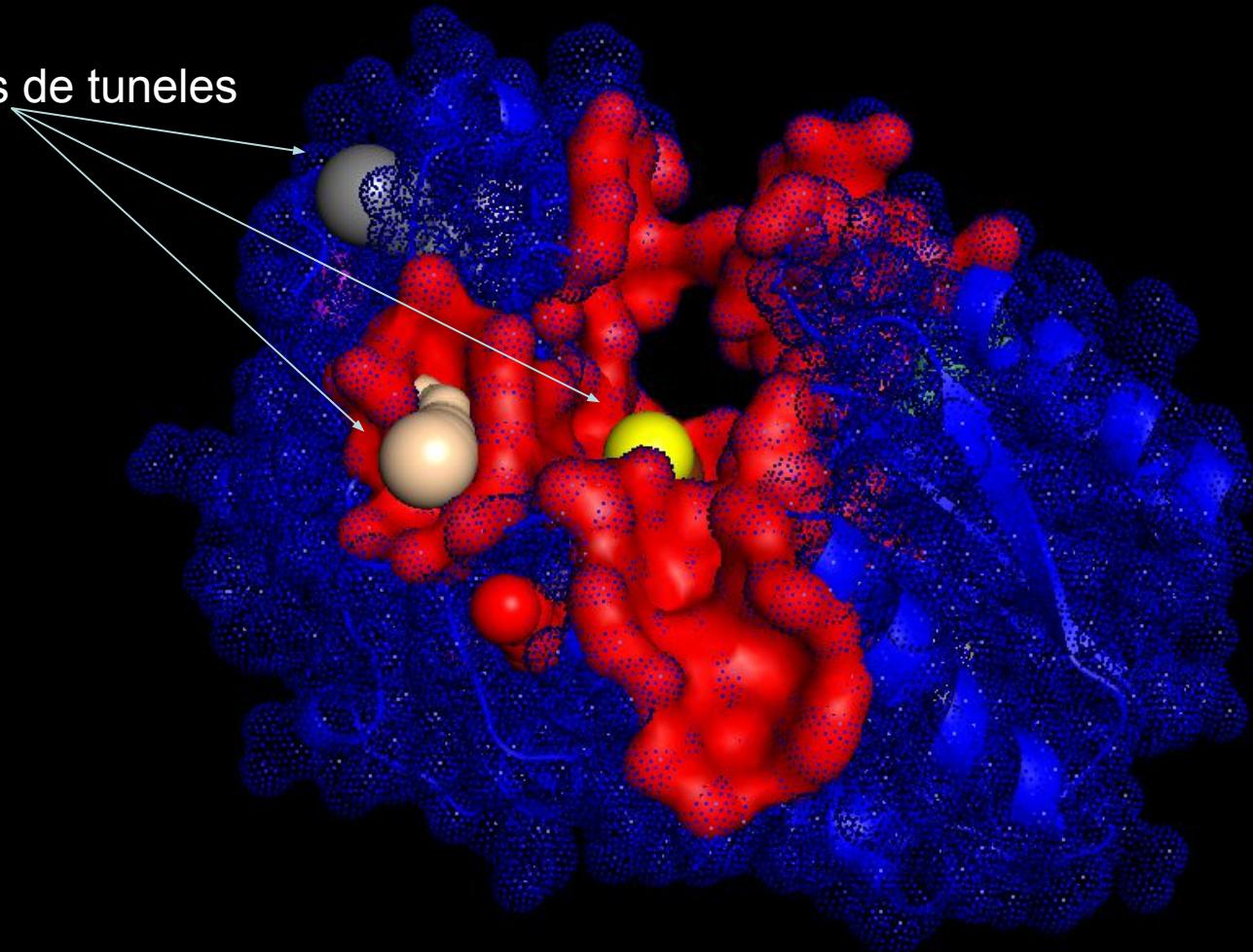


**Figure 6.** Locations of gates within a protein structure. Schematic representation of an enzyme with two active sites connected by a tunnel (I), a cofactor cavity (II), and multiple access tunnels. Gating residues in red may be located at the entrance to the active site (1), at the entrance or the bottleneck of the tunnel (2), and between the active site cavity and the cofactor cavity (3).

Pockets: regiones especiales de la superficie de una proteína, en general asociadas a la unión de ligandos

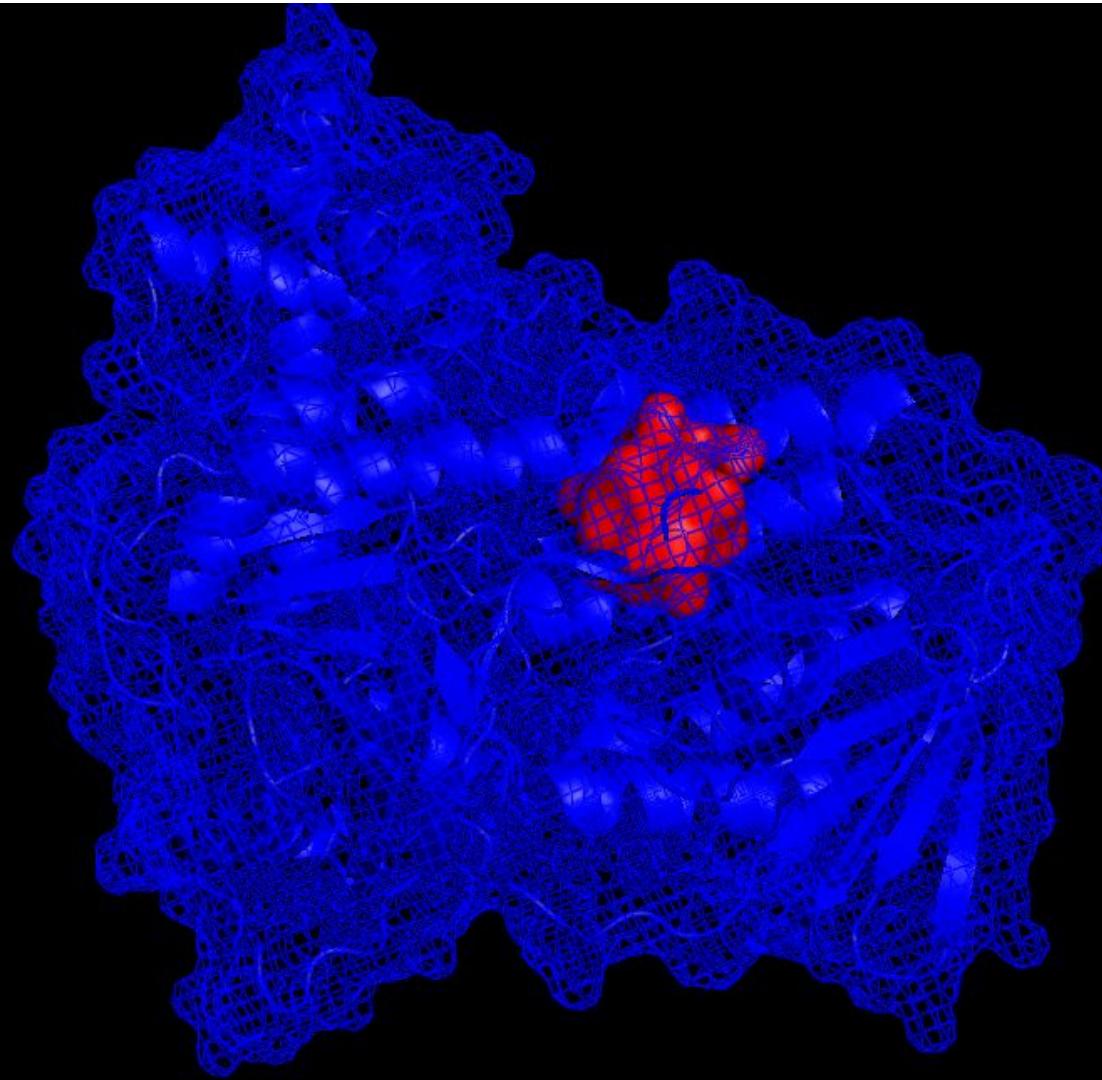
PyMOL for evaluation only.  
Contact sales@delscicom.

Entradas de tuneles



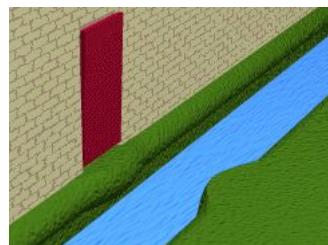
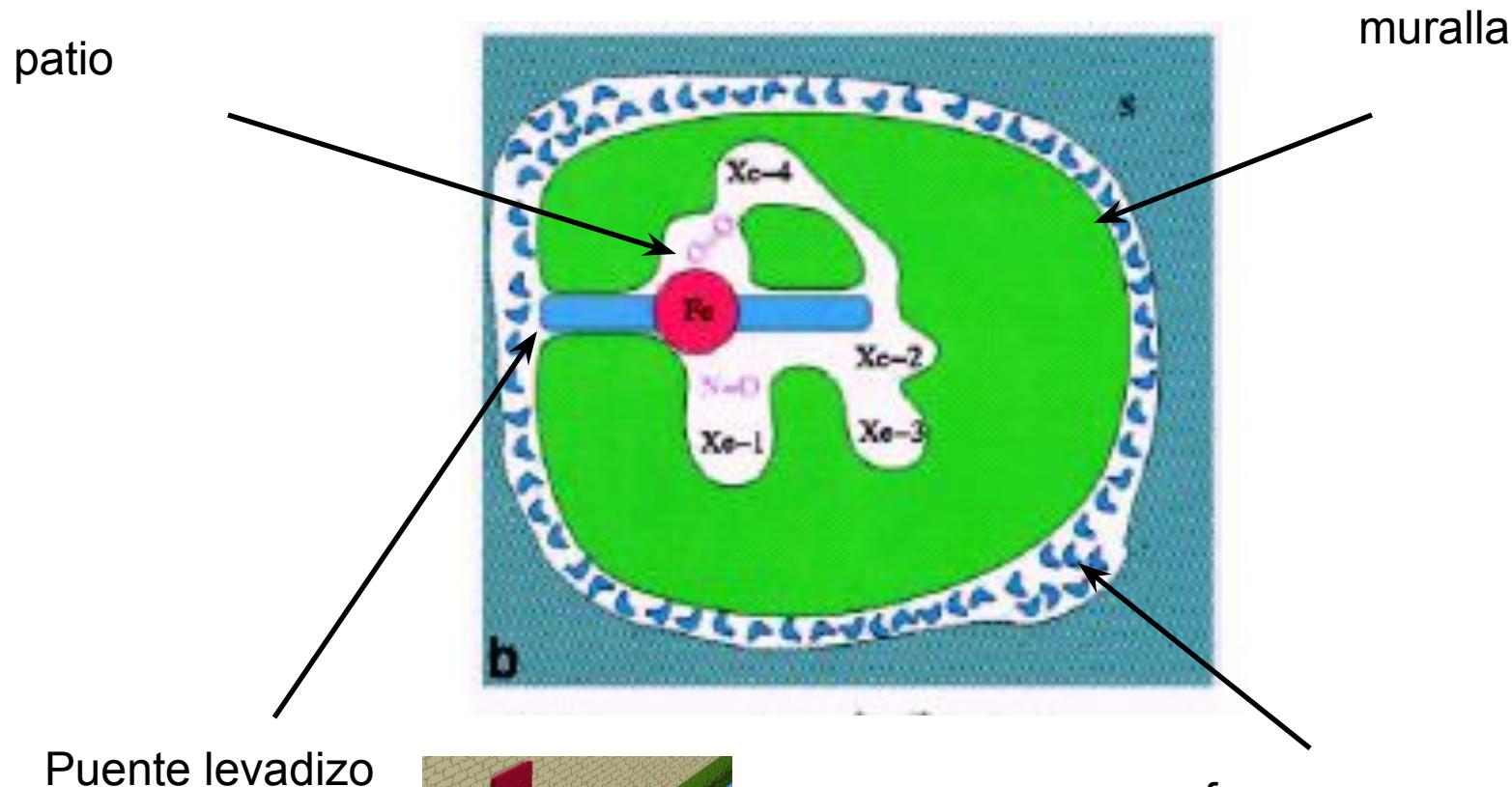
# Voids: cavidades que no tienen acceso al exterior

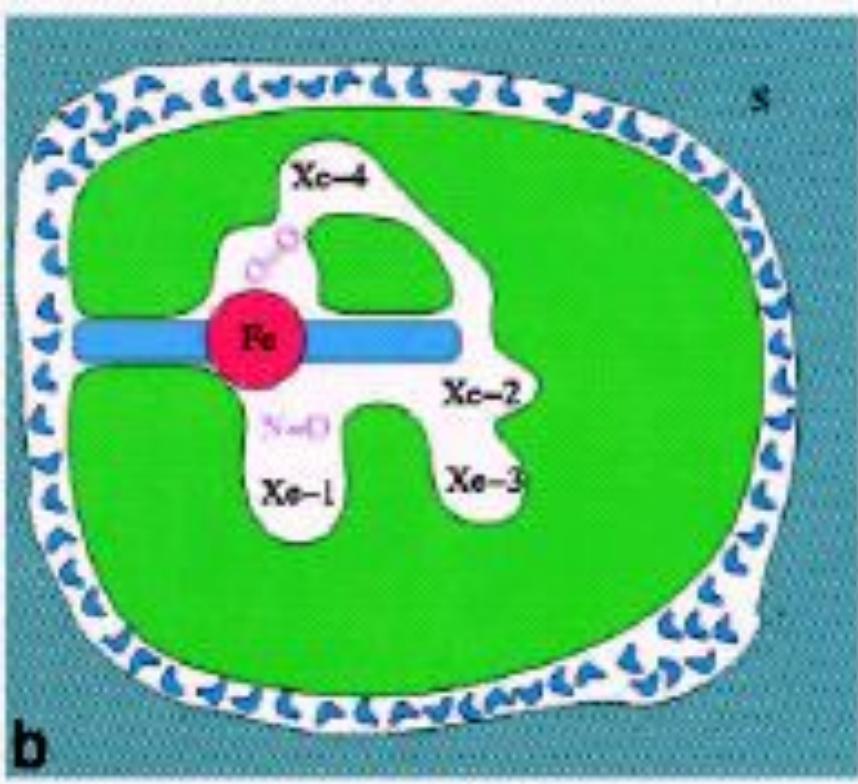
PyMOL for evaluation only.  
Contact sales@delsci.com.

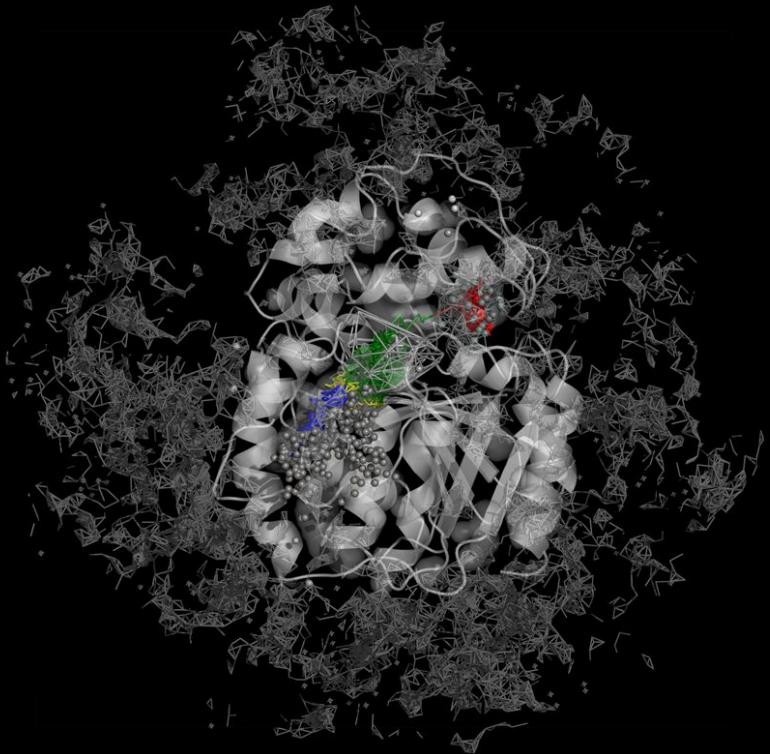


# Proteínas como castillos:

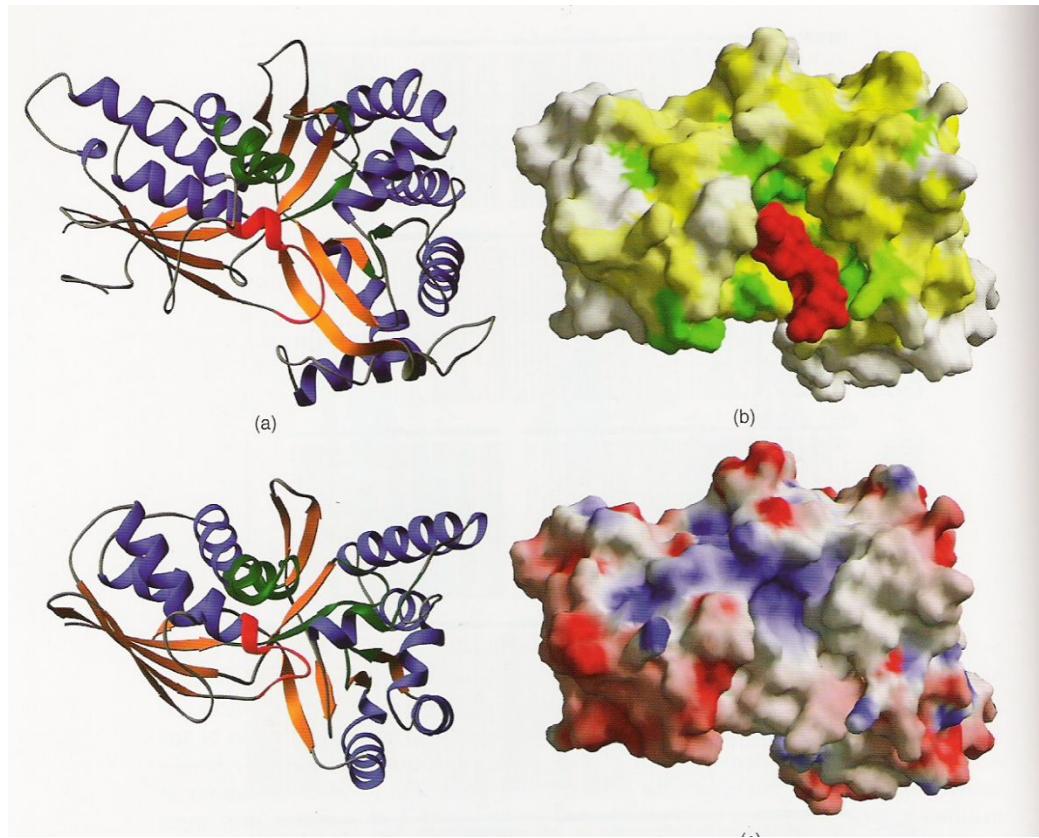
las proteínas tienen superficies, cavidades, surcos y entradas y salidas para el tránsito de ligandos (sustratos, productos, moduladores, etc)



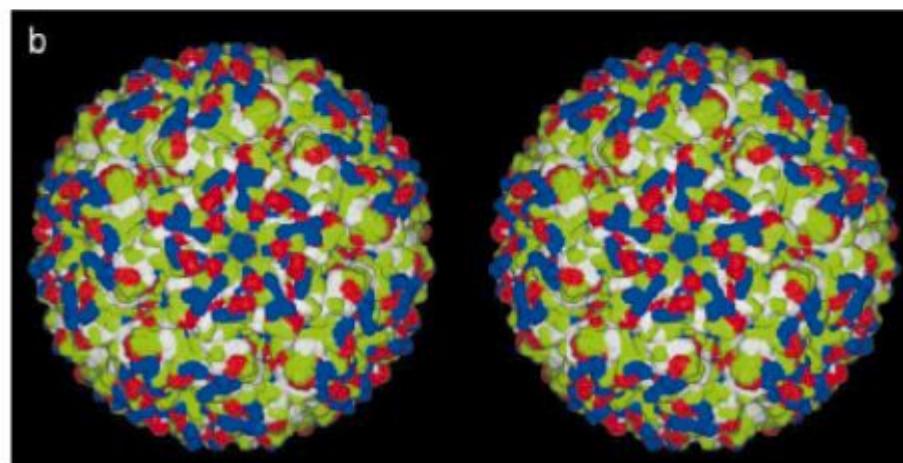
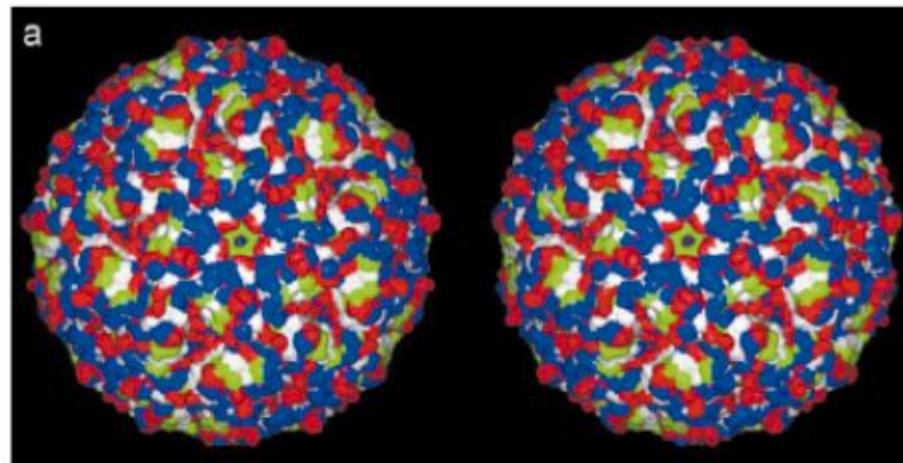




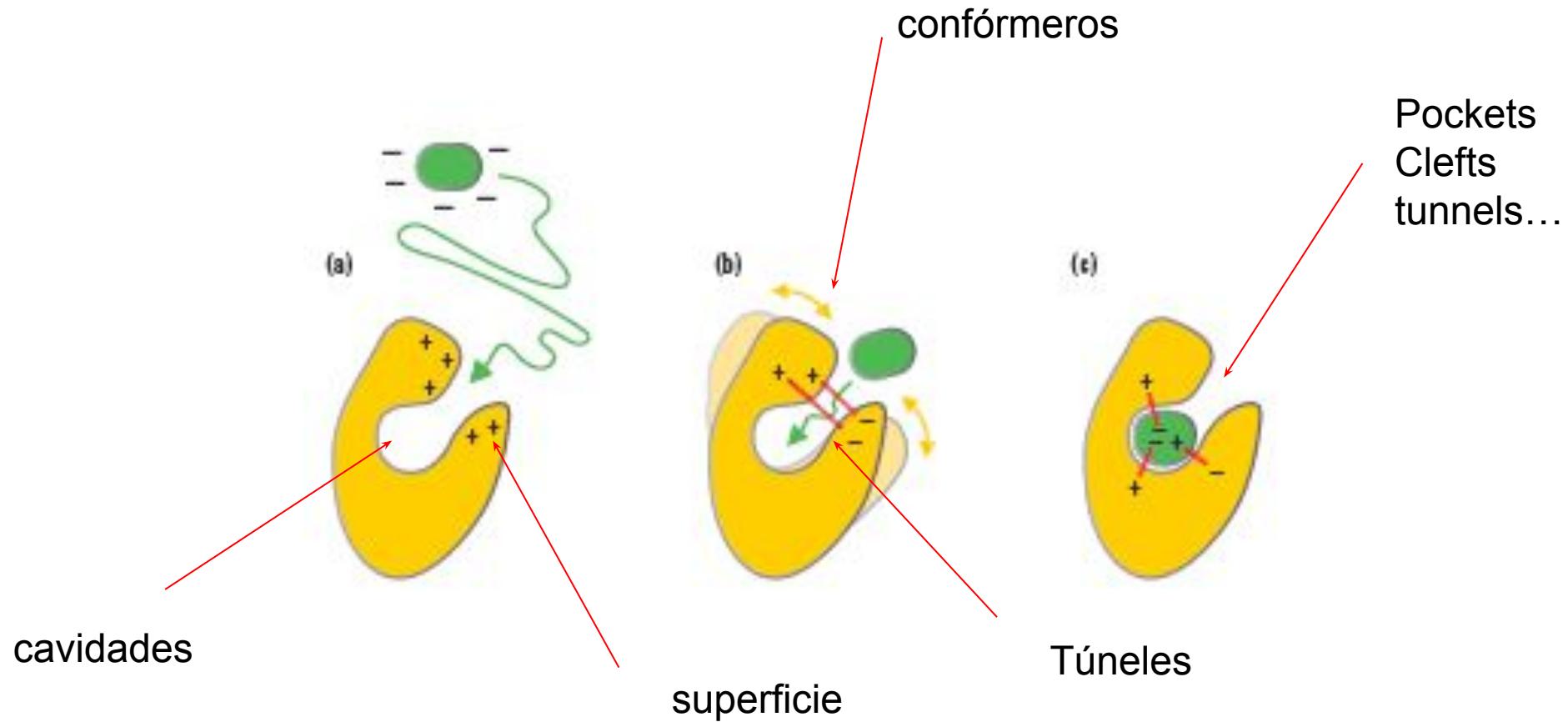
AQUA -  $\overline{Z}$  - DUCT by  $\overline{T}$  -  $\overline{Z}$  - G

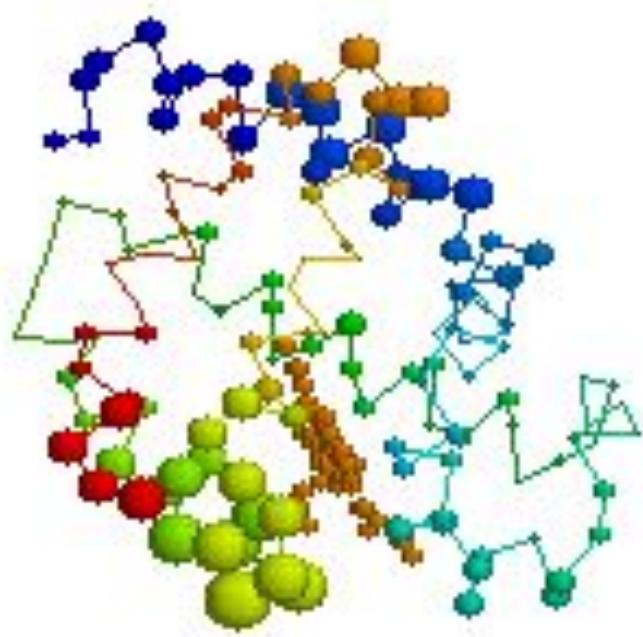


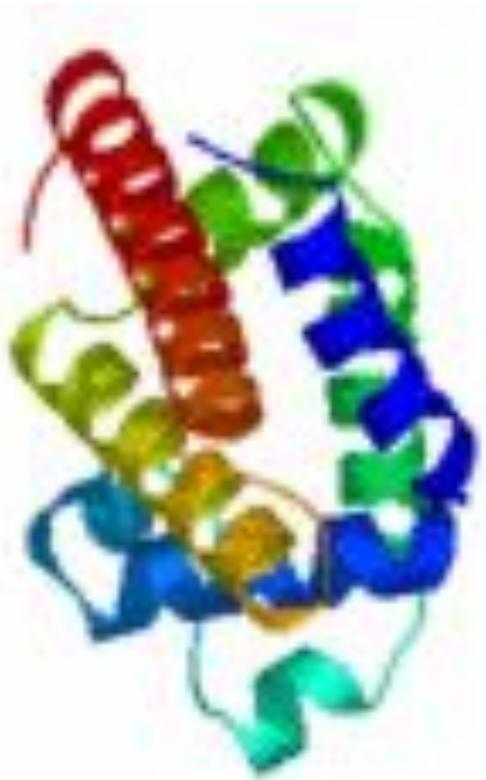
## Comparación de superficies en la misma proteína de organismos psicrófilos y termófilos



## La actividad y función de una proteína depende de varios pasos



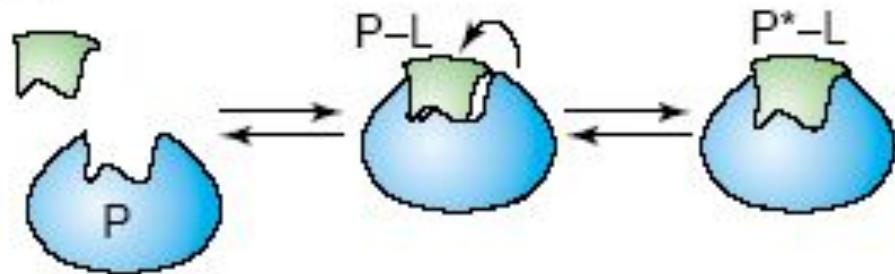




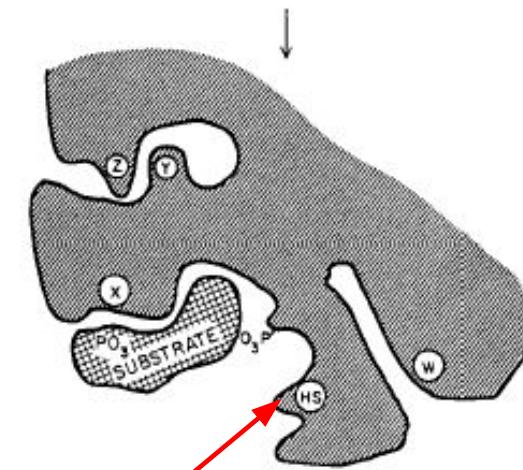
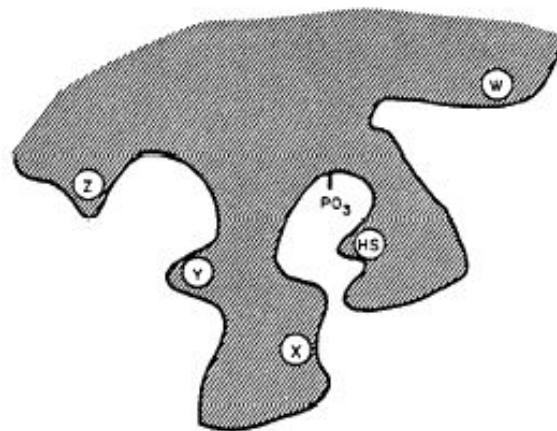
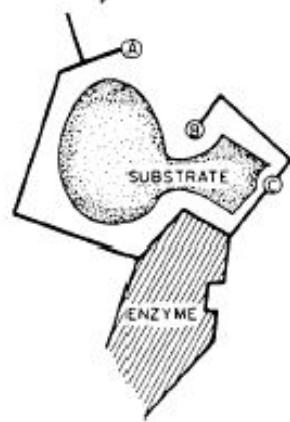
# Relación Estructura-Función

## Teoría del Ajuste inducido

(iii) Induced fit



D. Koshland,  
1920-2007



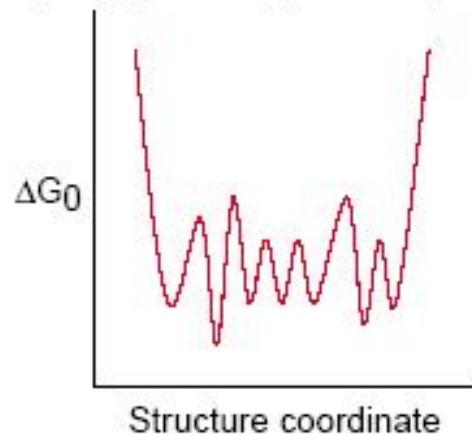
Con y sin sustrato había una diferencia en los SH titulables!

# Relación Estructura Función

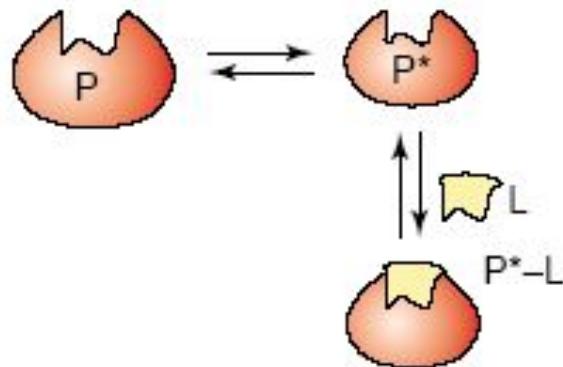
Diversidad Conformatacional  
en el estado nativo: teoría  
del pre-equilibrio

(b) 'New view'

(i) Rugged energy landscape



(ii) Pre-equilibrium



Monod, Wyman and Changeux, 1965

Fred Karush, 1950



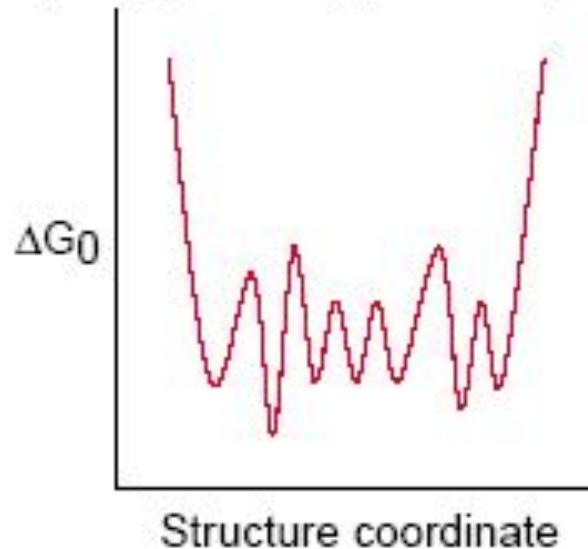
© The Nobel Foundation

J Monod  
1910-1976

## Visión Actual: el fondo del embudo es RUGOSO!

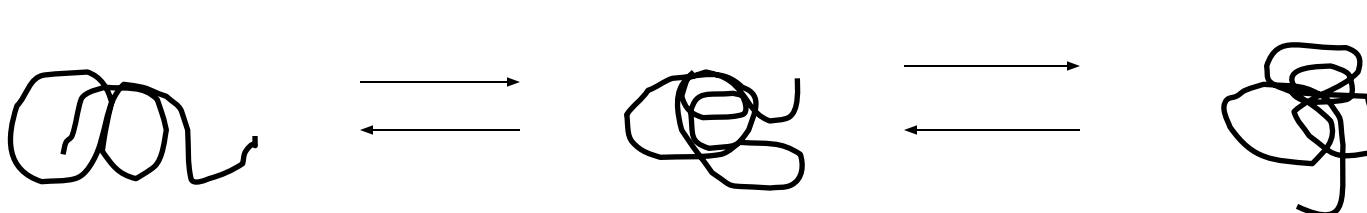
(b) 'New view'

(i) Rugged energy landscape

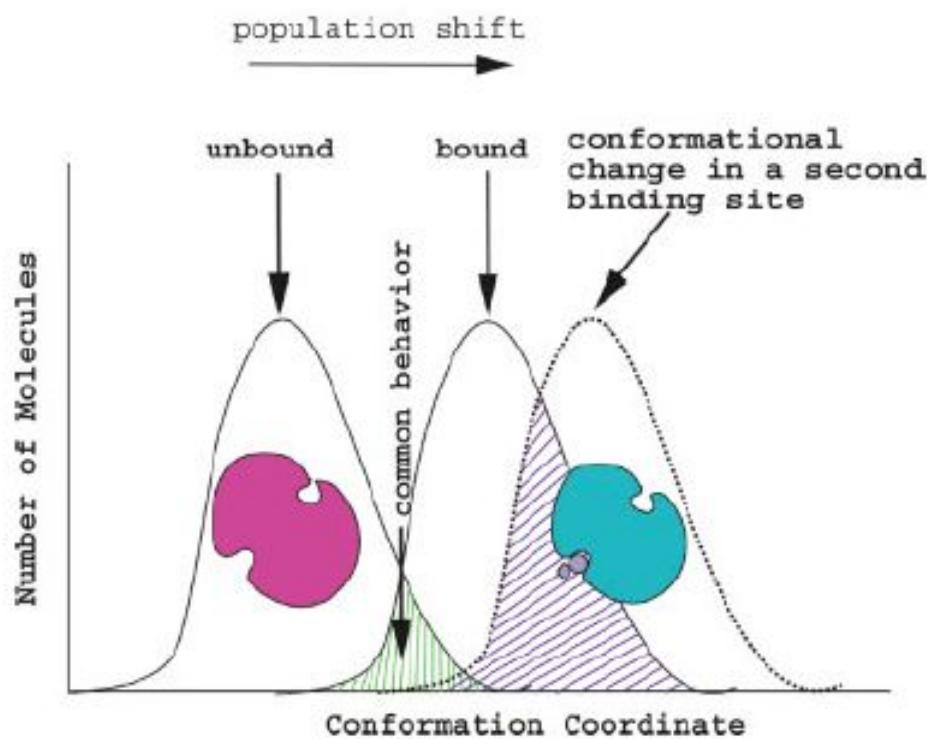


$$\frac{P_{unfolded}}{P_{folded}} = e^{-\Delta \Delta G_{fold} / kT}$$

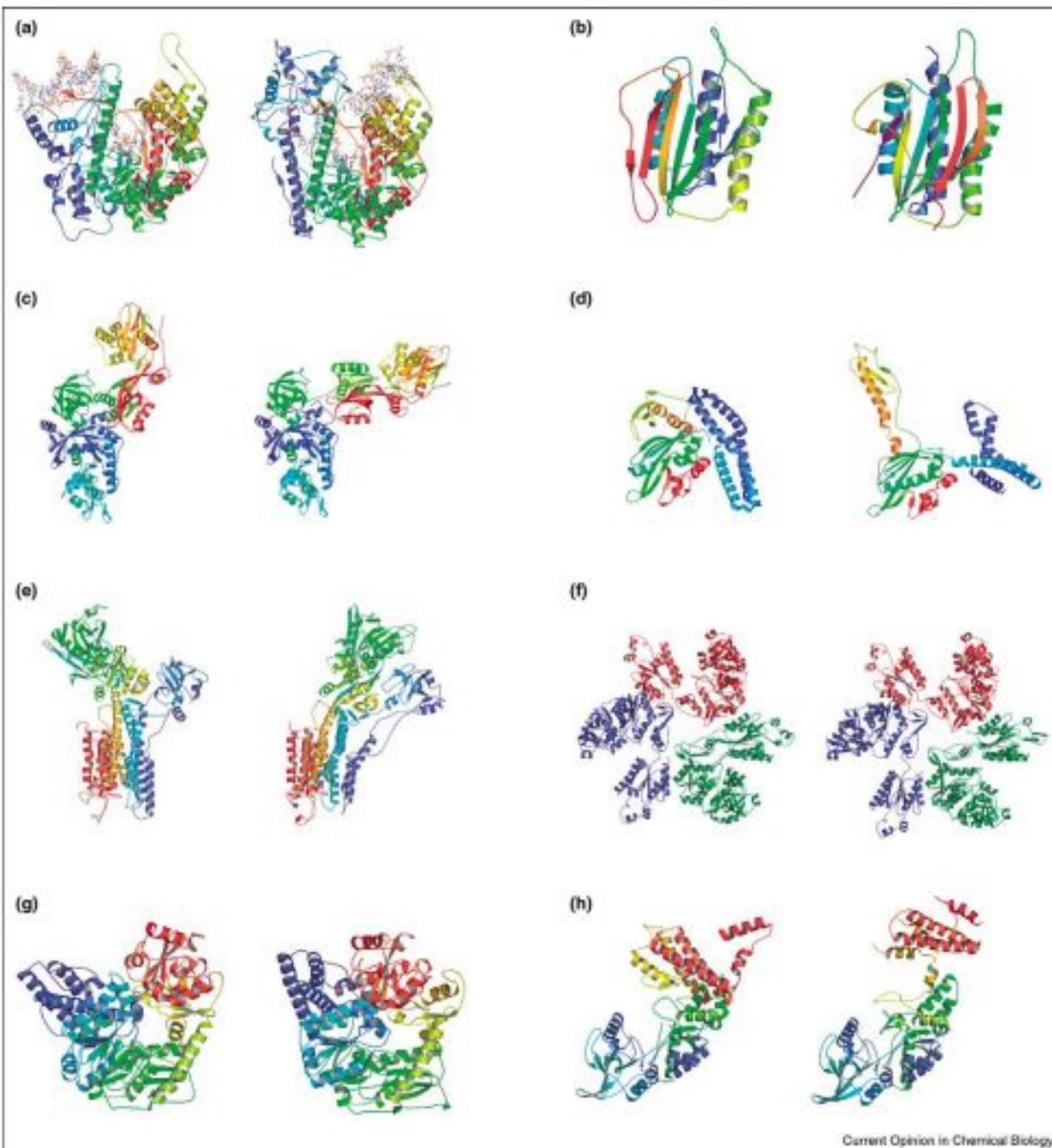
Eq Boltzman nos da la distribución de los confórmeros en el equilibrio



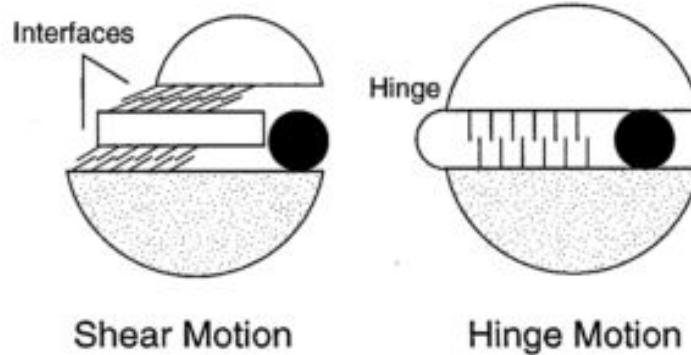
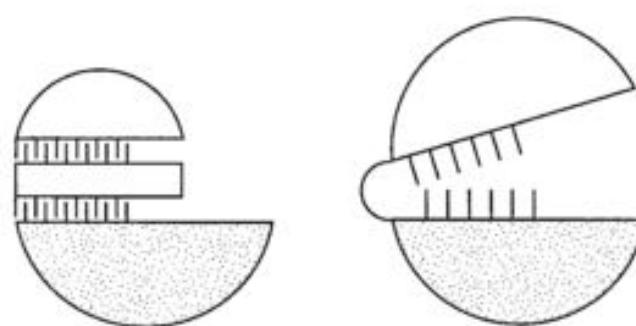
**Estado nativo:** el estado nativo de una proteína no es único y se lo describe como un conjunto de confórmeros



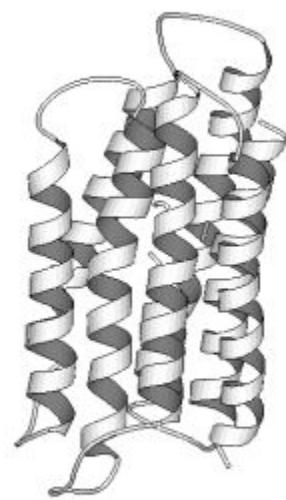
# Cambios conformacionales en proteínas



## Hinge and shear motions



# Shear or Hinge?



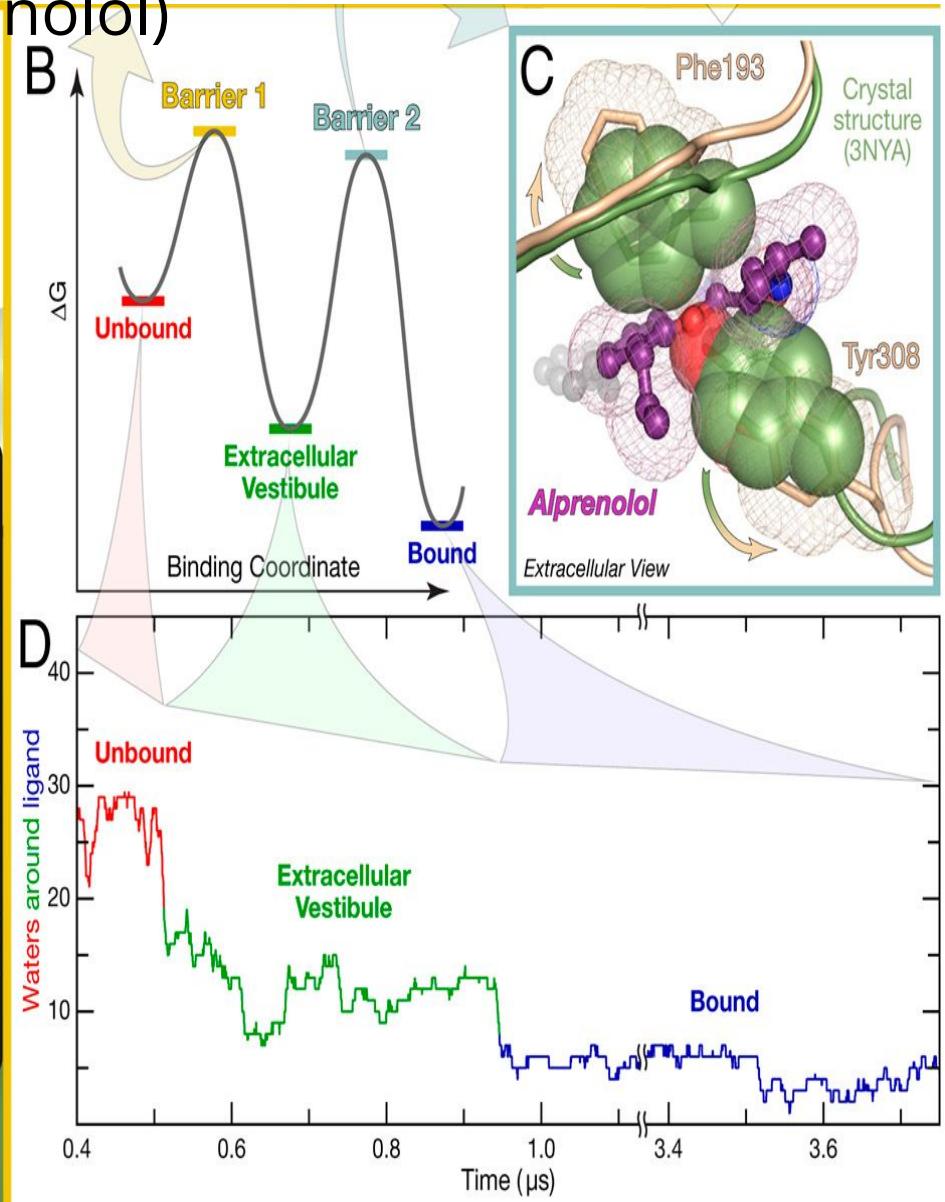
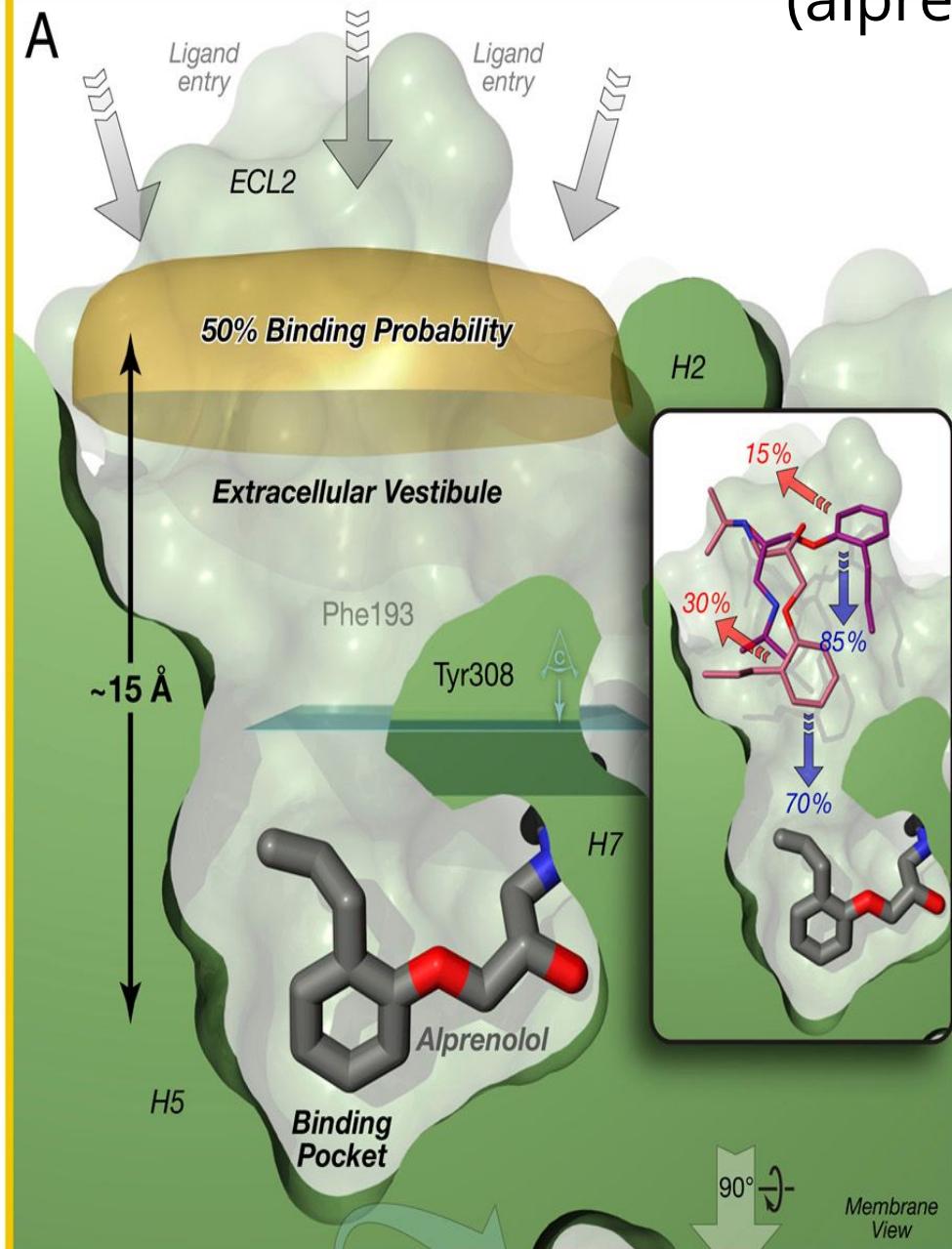
## Shear or Hinge?



# Shear or Hinge?

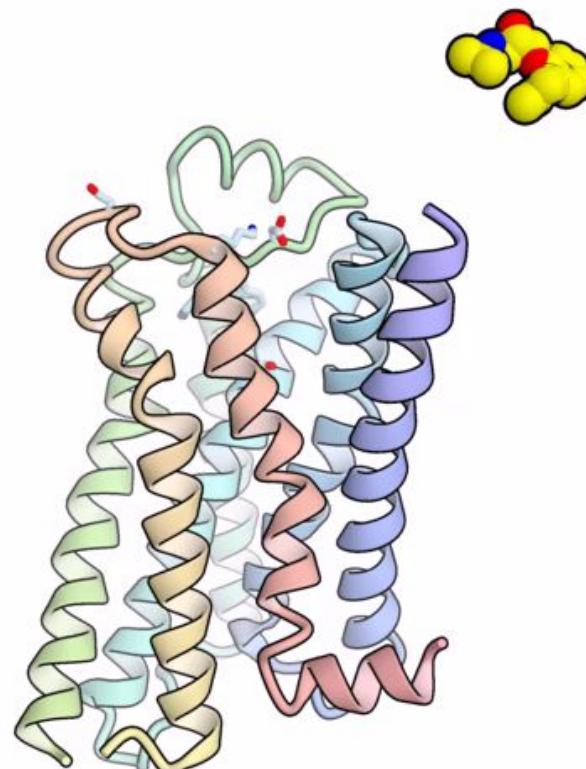


# Unión de un ligando a su receptor (alprenolol)



# Unión de la proteína G al receptor

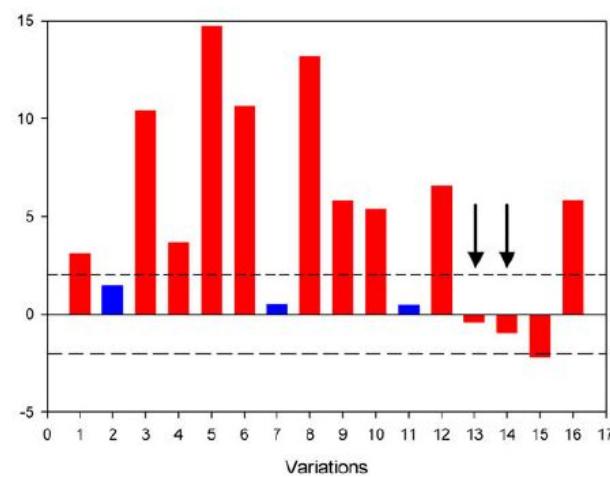
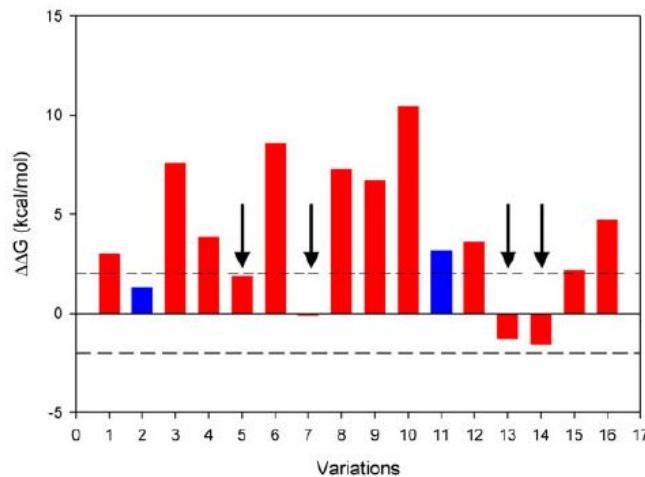
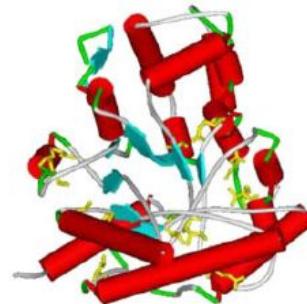
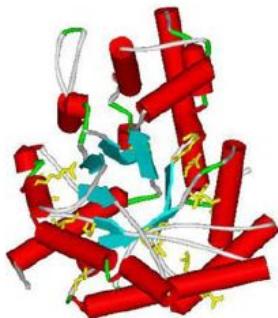
0.00 us



Dror et al., PNAS 2011

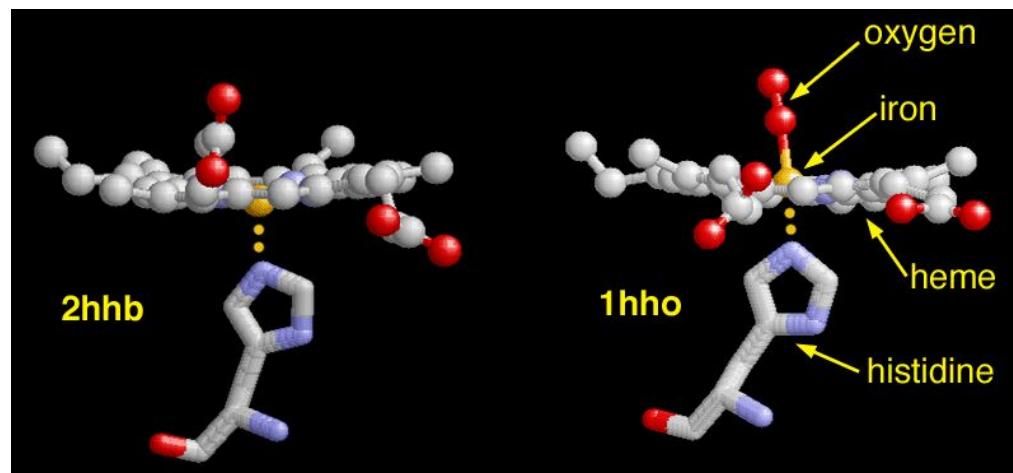
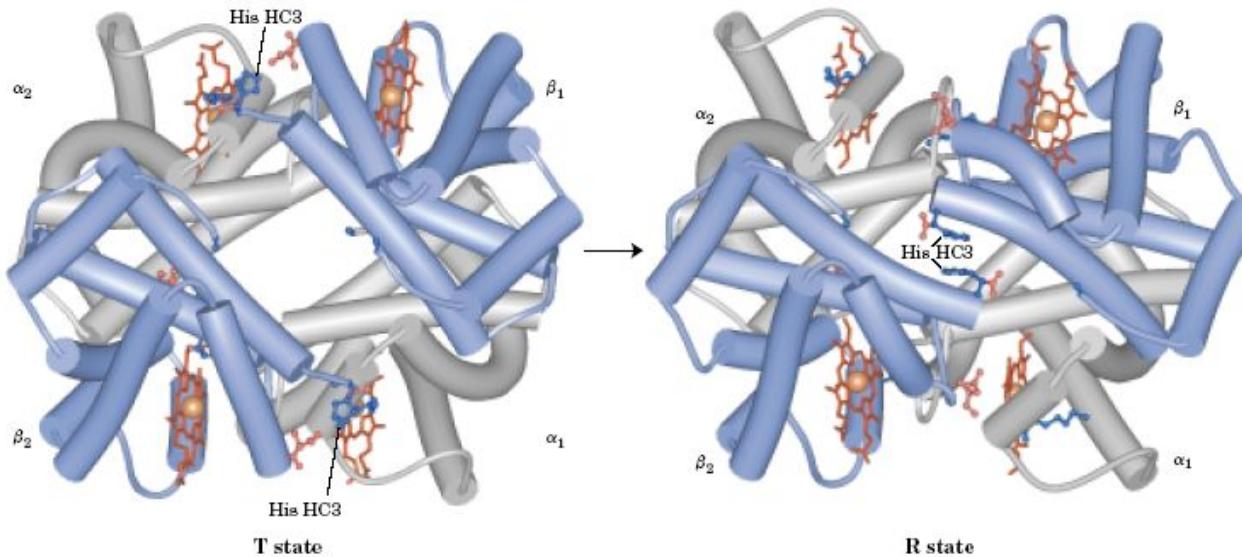
*Existe una primera asociación en el “vestíbulo” a 15 Å del sitio activo.  
La subsecuente llegada al sitio activo requiere la deformación de la  
proteína y el pasaje de la droga por un fino canal*

# Diversidad conformacional y efecto de las mutaciones en una proteína



# Hemoglobina

## 4 cadenas polipeptídicas



# Das Gleichgewicht zwischen Hämoglobin und Sauerstoff.

Von

Felix Haurowitz<sup>1).</sup>.

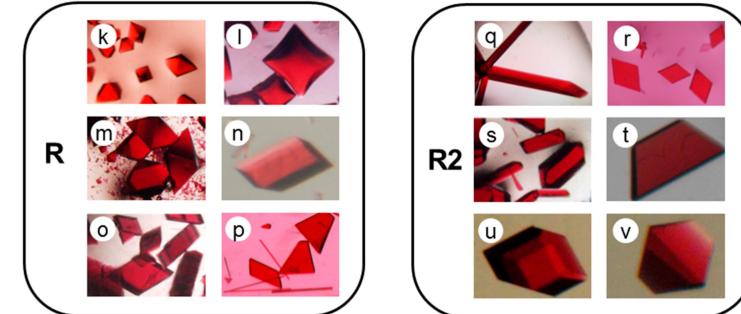
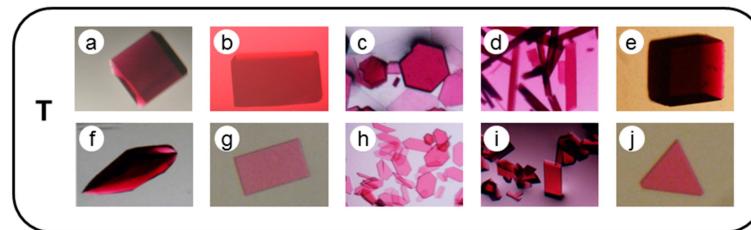
Mit 2 Figuren im Text.

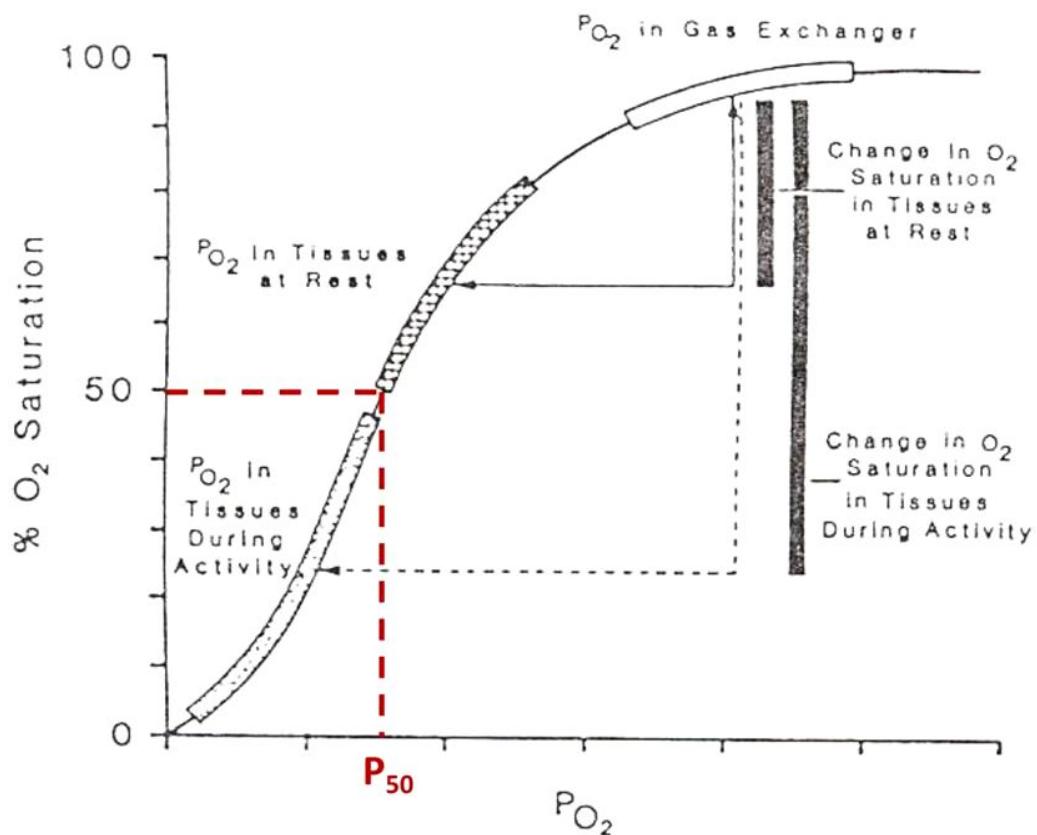
(Aus dem Medizin.-chemischen Institut der Deutschen Universität in Prag.)

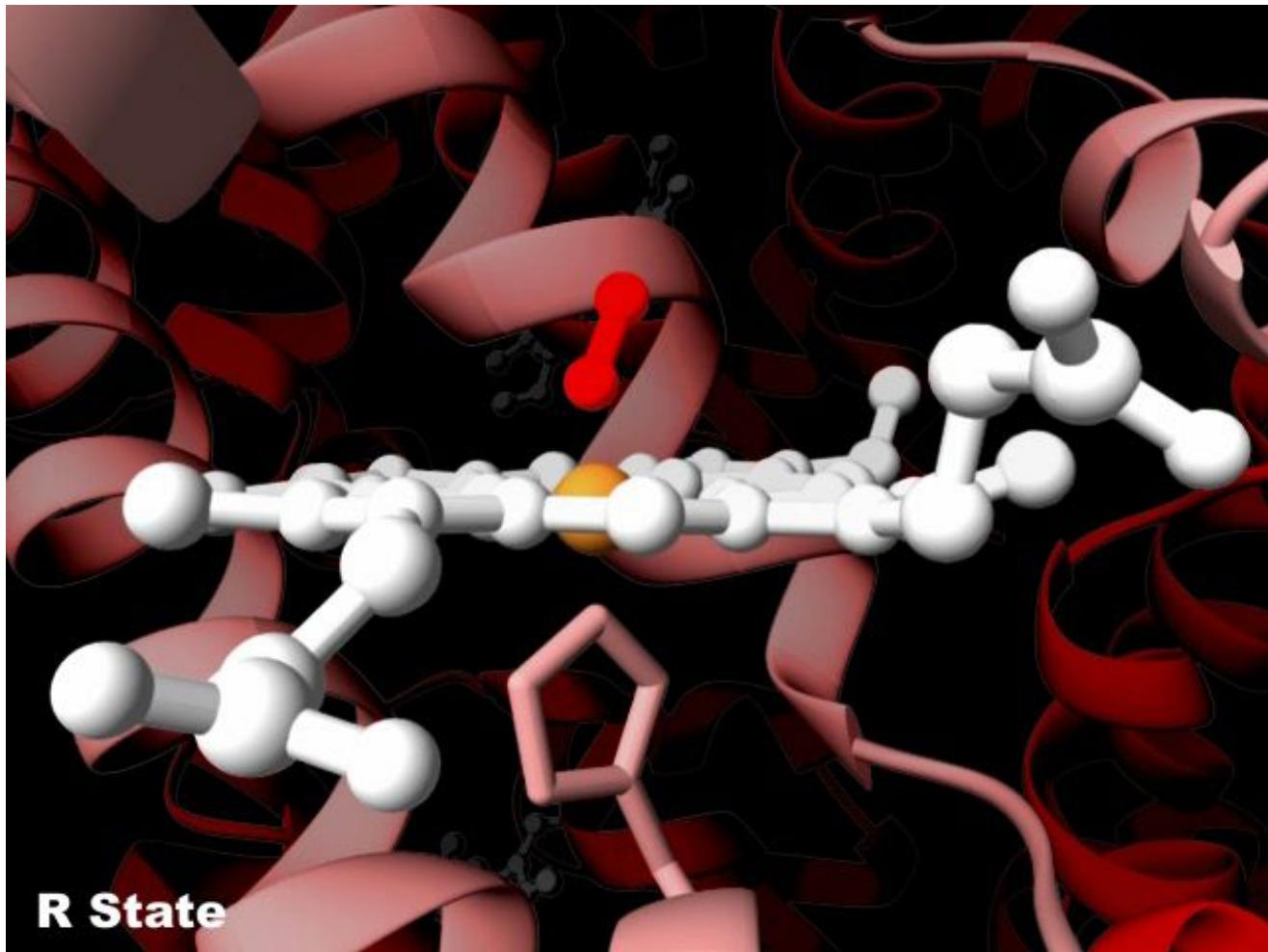
(Der Schriftleitung zugegangen am 27. Juni 1938.)

## 1. Theorien über die Vereinigung von Hämoglobin mit Sauerstoff.

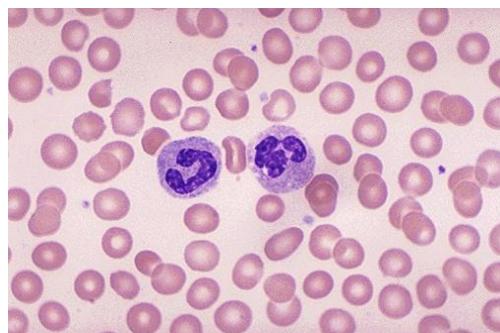
Die Reaktion zwischen Hämoglobin und Sauerstoff verläuft sicher nicht nach der einfachen Gleichung  $Hb + O_2 = HbO_2$ ; denn an Stelle der dabei zu erwartenden hyperbolischen Dissoziationskurve erhält man eine S-förmige Kurve, welche einer Reaktion höherer Ordnung entspricht.



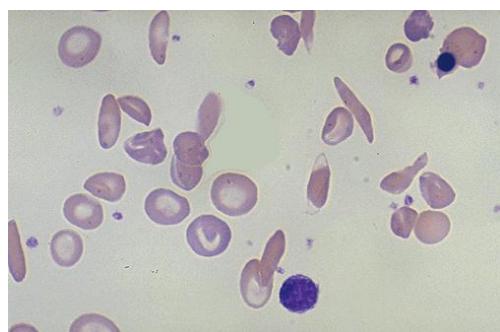




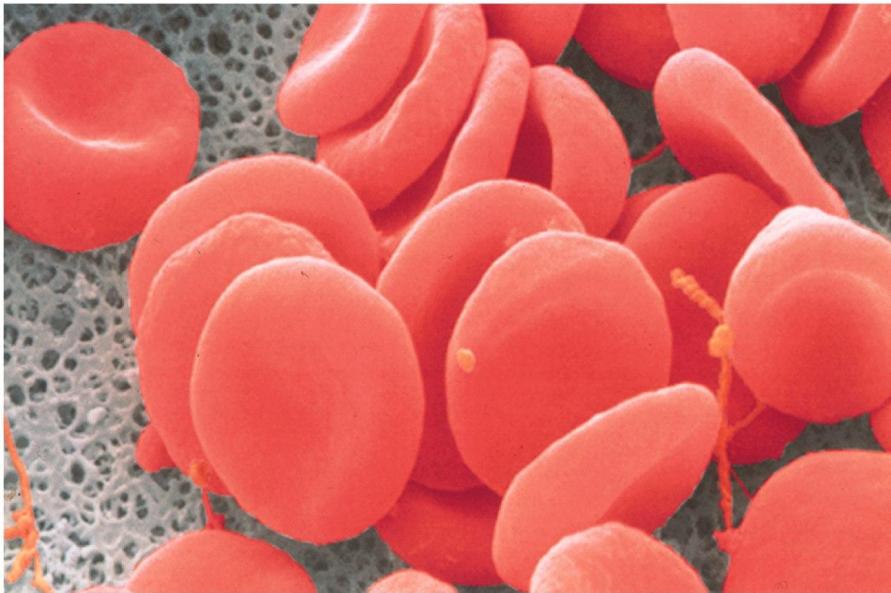
# Hemoglobina S



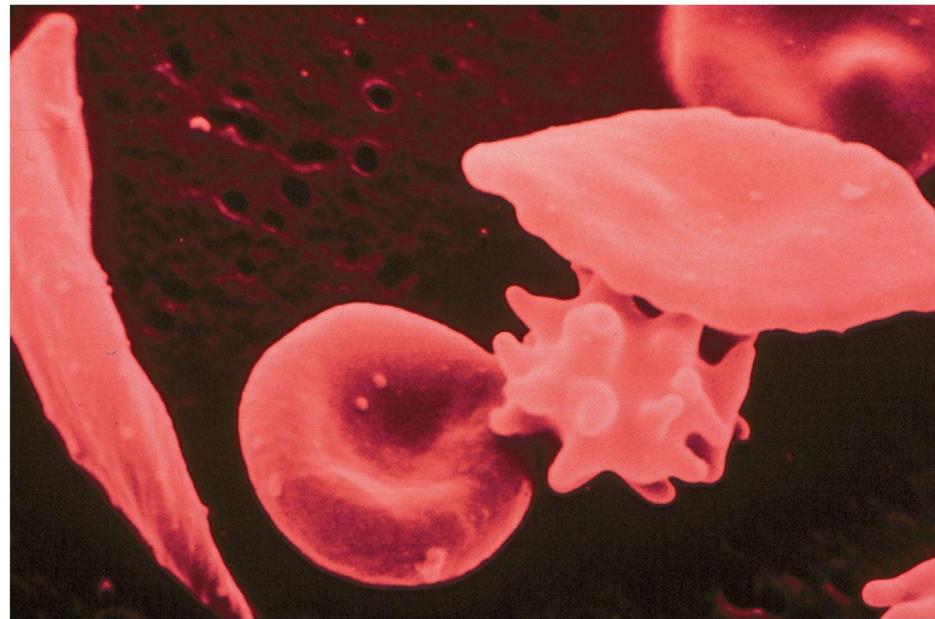
Sangre normal HbA



Sangre drepanocítica  
HbS



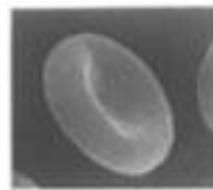
2  $\mu\text{m}$



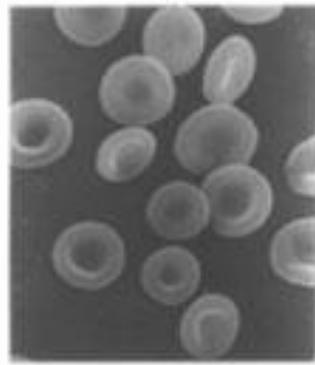
Linus Pauling sugirió en 1949 que la anemia drepanocítica era una enfermedad molecular.

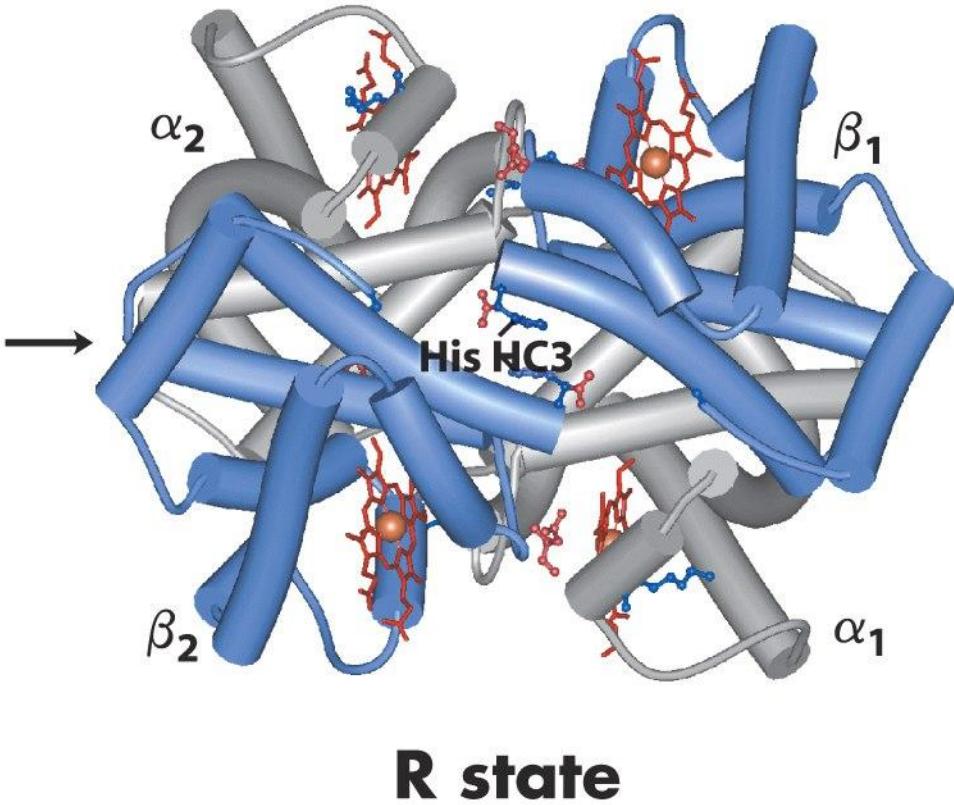
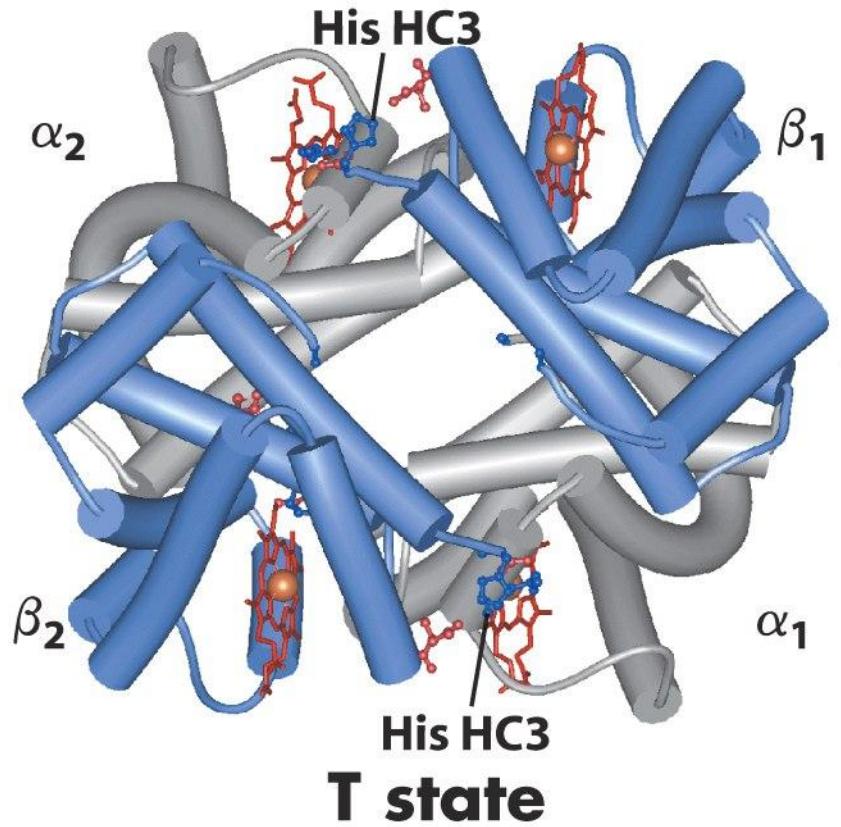
Posteriormente se demostró que la HbS es una variante mutante de la HbA, donde el Glu 6 de la cadena beta se sustituye por una Val.

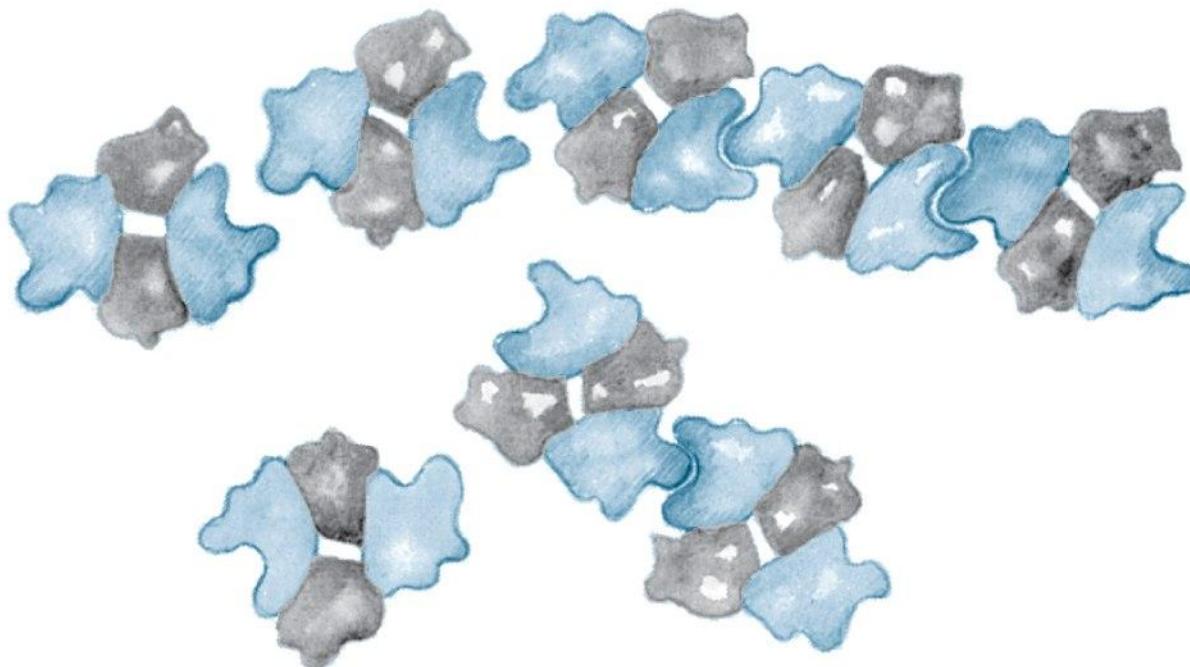
Normal Red Blood Cell → Sickle Red Blood Cell



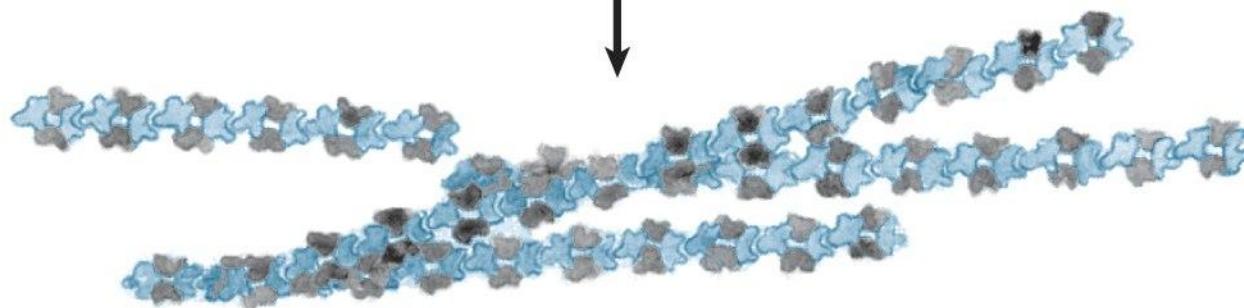
OXY-STATE    ↔    DEOXY-STATE



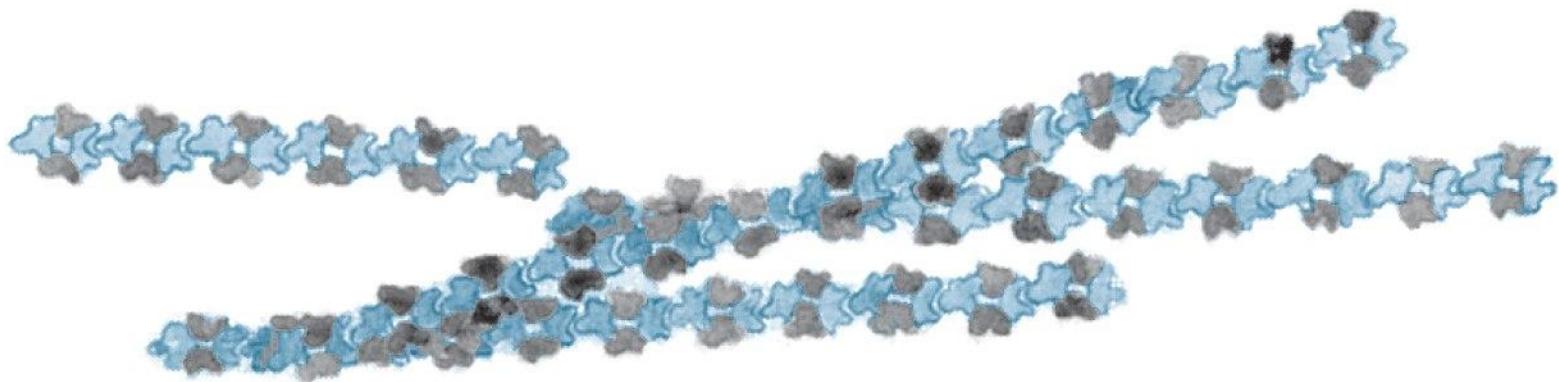




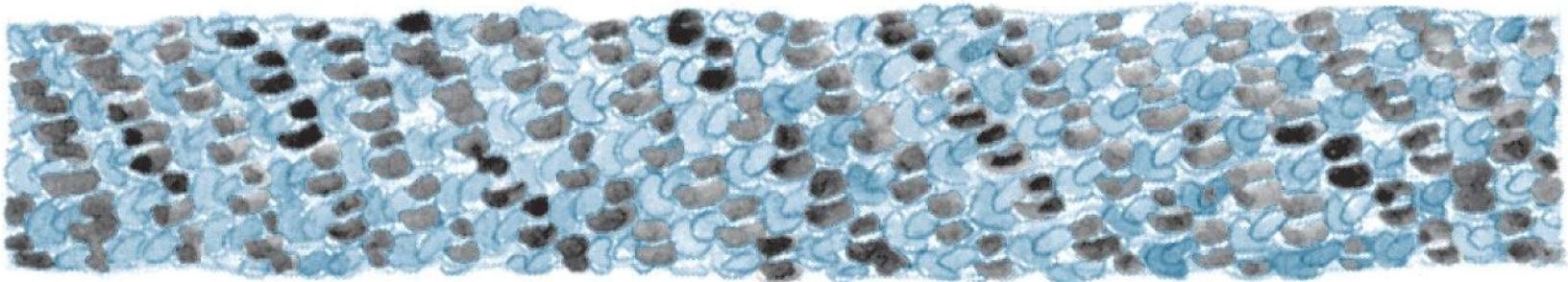
**Interaction between molecules**



**Strand formation**

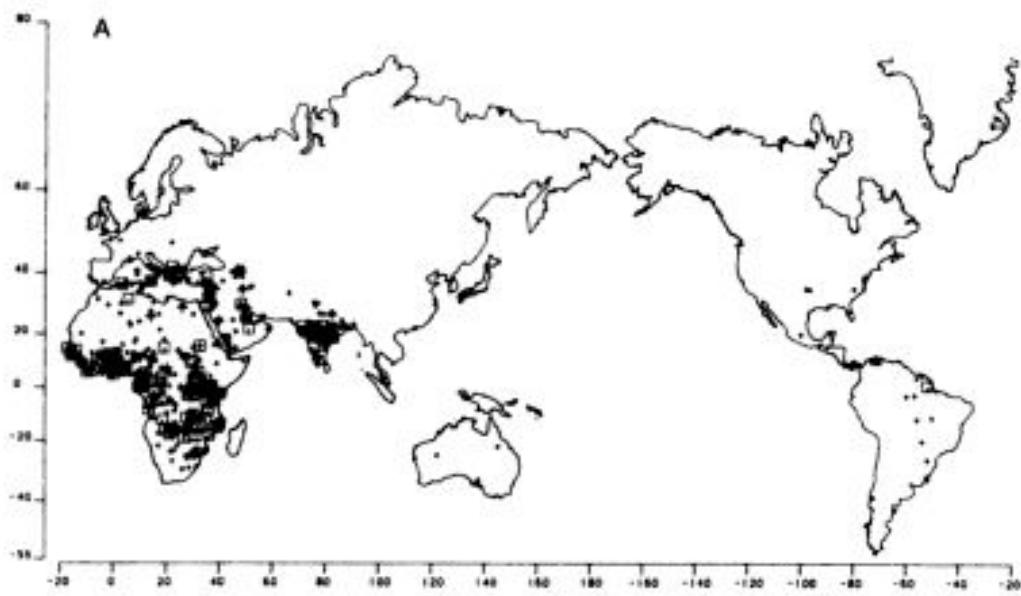


**Strand formation**

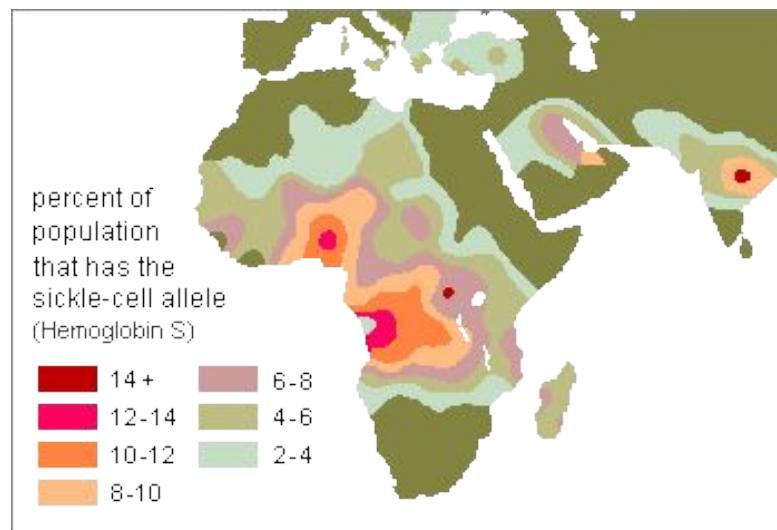
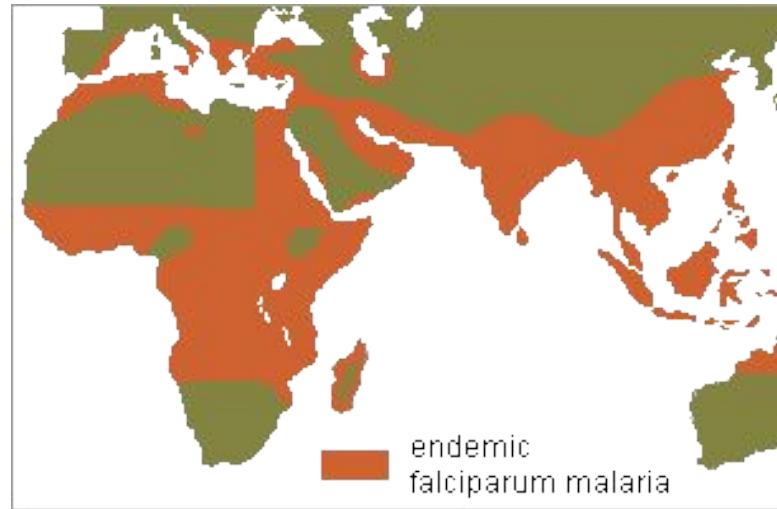


**Alignment and crystallization  
(fiber formation)**

# Distribución mundial de Hb S

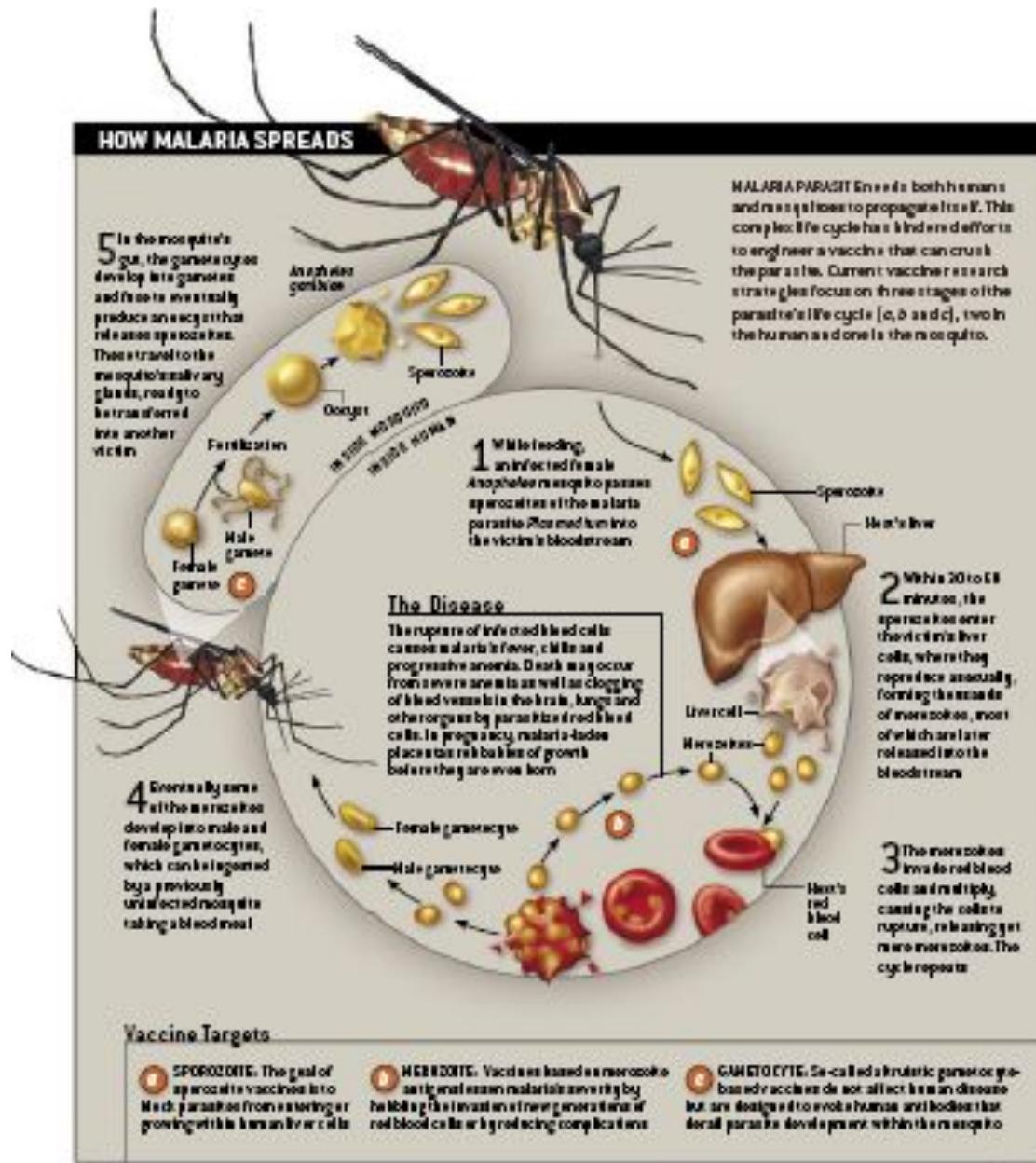


**Fig. 2.14.1.A** Geographic distribution of hemoglobin S available data points in the world. Symbols are described in table 1.14.1.



## 3 distintos orígenes de la mutación



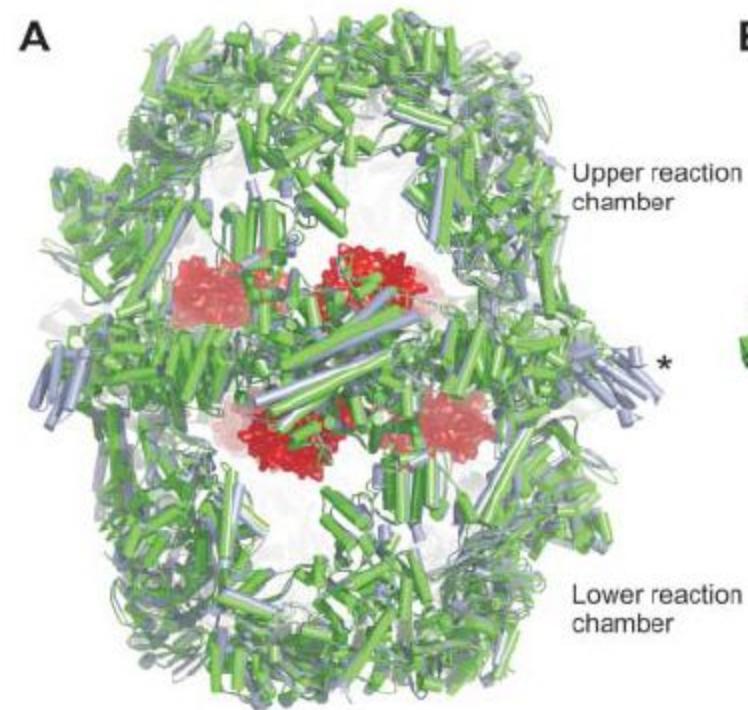
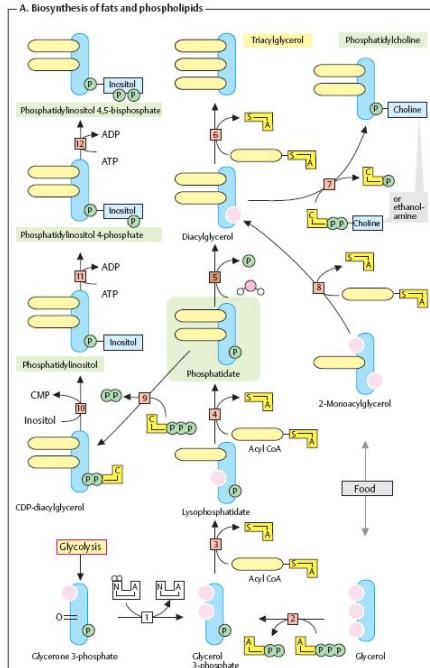


“Nada tiene sentido en biología si no es a la luz de la teoría de la evolución”

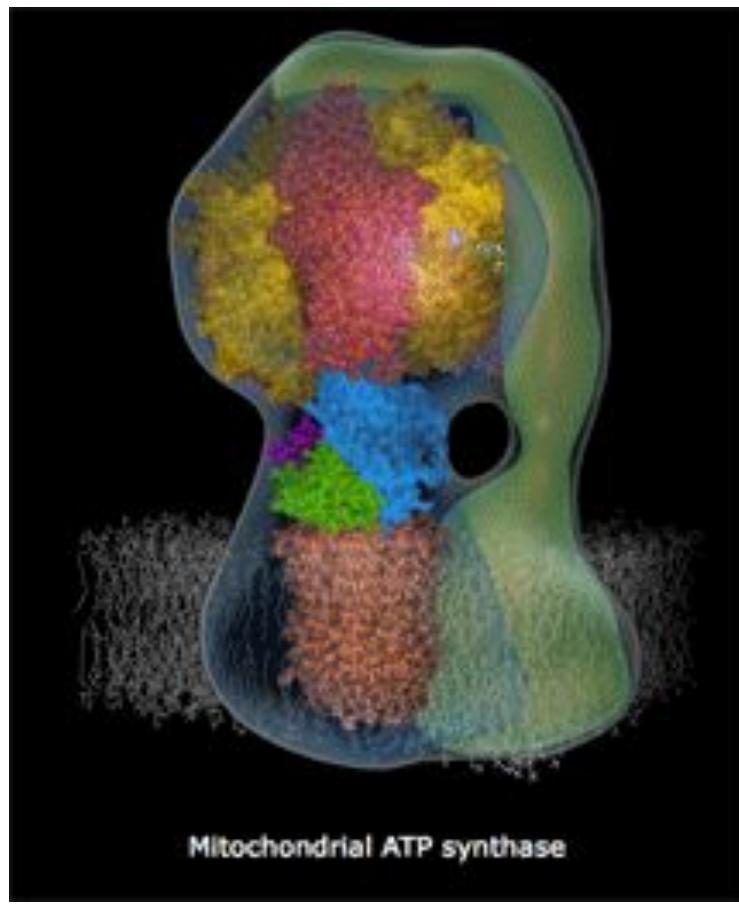
T. Dobzhansky

# Complejos supramacromoleculares

## Por ejemplo: FAS (fatty acid synthase)

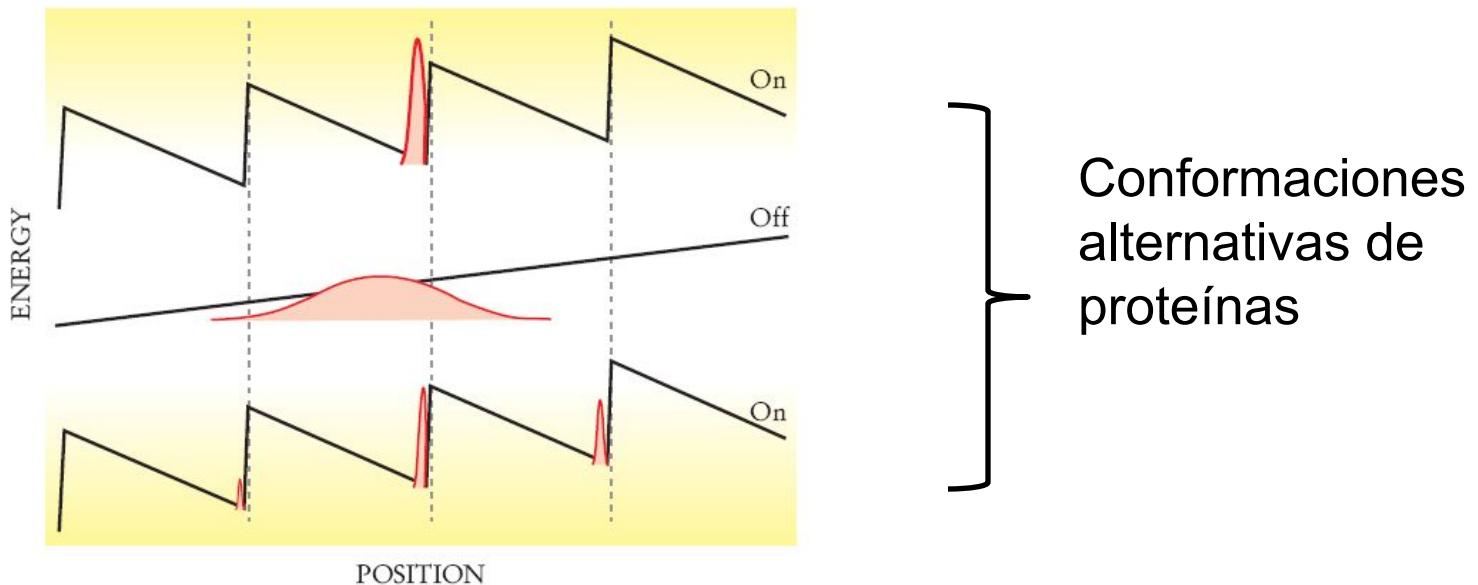


Complejos mitocondriales para la fosforilación oxidativa  
Fotosistemas I y II  
Degradación de proteínas en el proteasoma



Mitochondrial ATP synthase

## Movimiento browniano con “dirección”

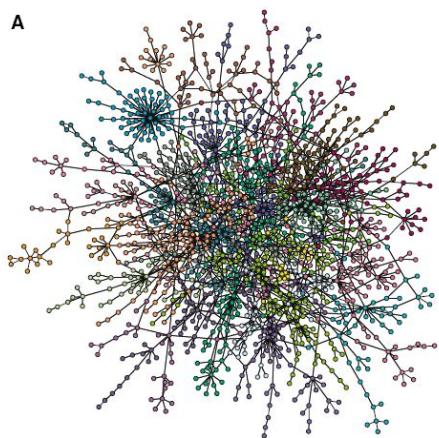




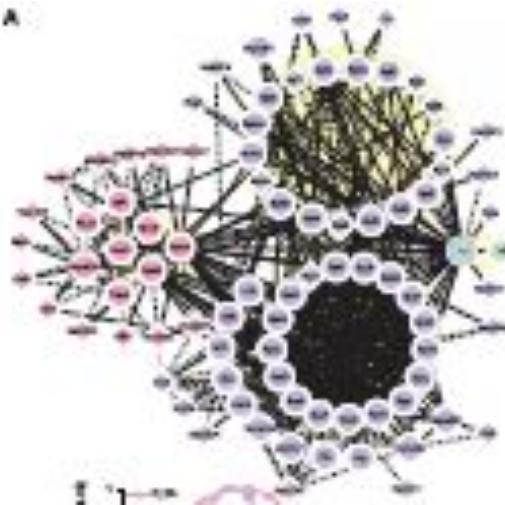
An artistic representation of a molecular ratchet mechanism that shows how Brownian motion is harnessed to spin the gear unidirectionally.

Credit: Astumian, R. Dean. "Making molecules into motors." *Scientific American* (2007): 57-64. (online)

A



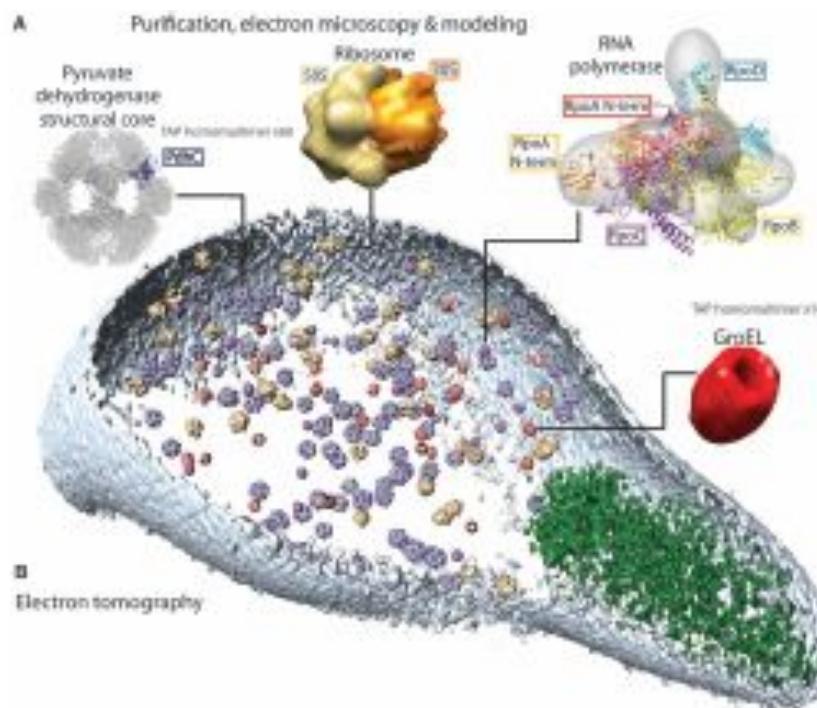
A



Módulos “funcionales”:  
movilidad,  
traducción, energía, etc

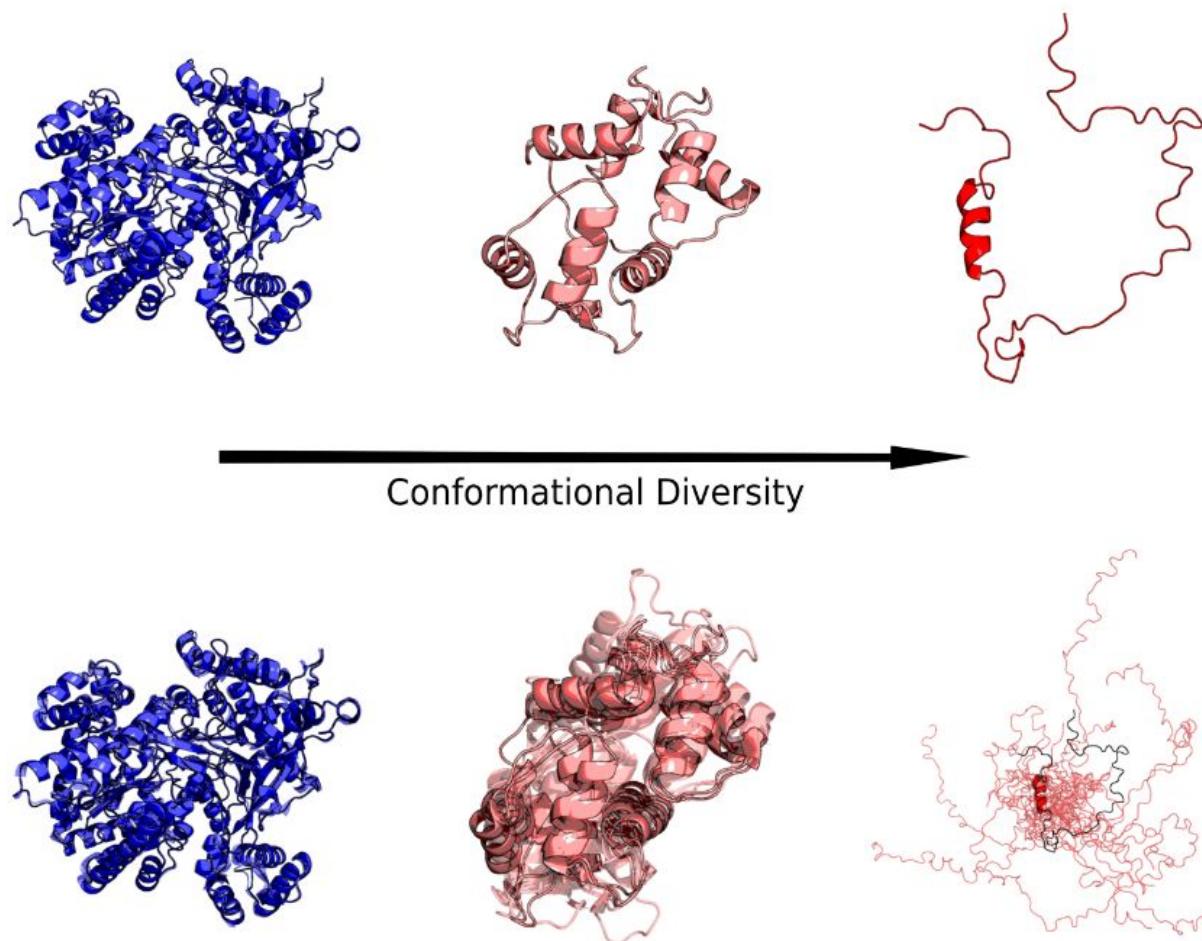
Patrón de interacción  
de un proteoma

A



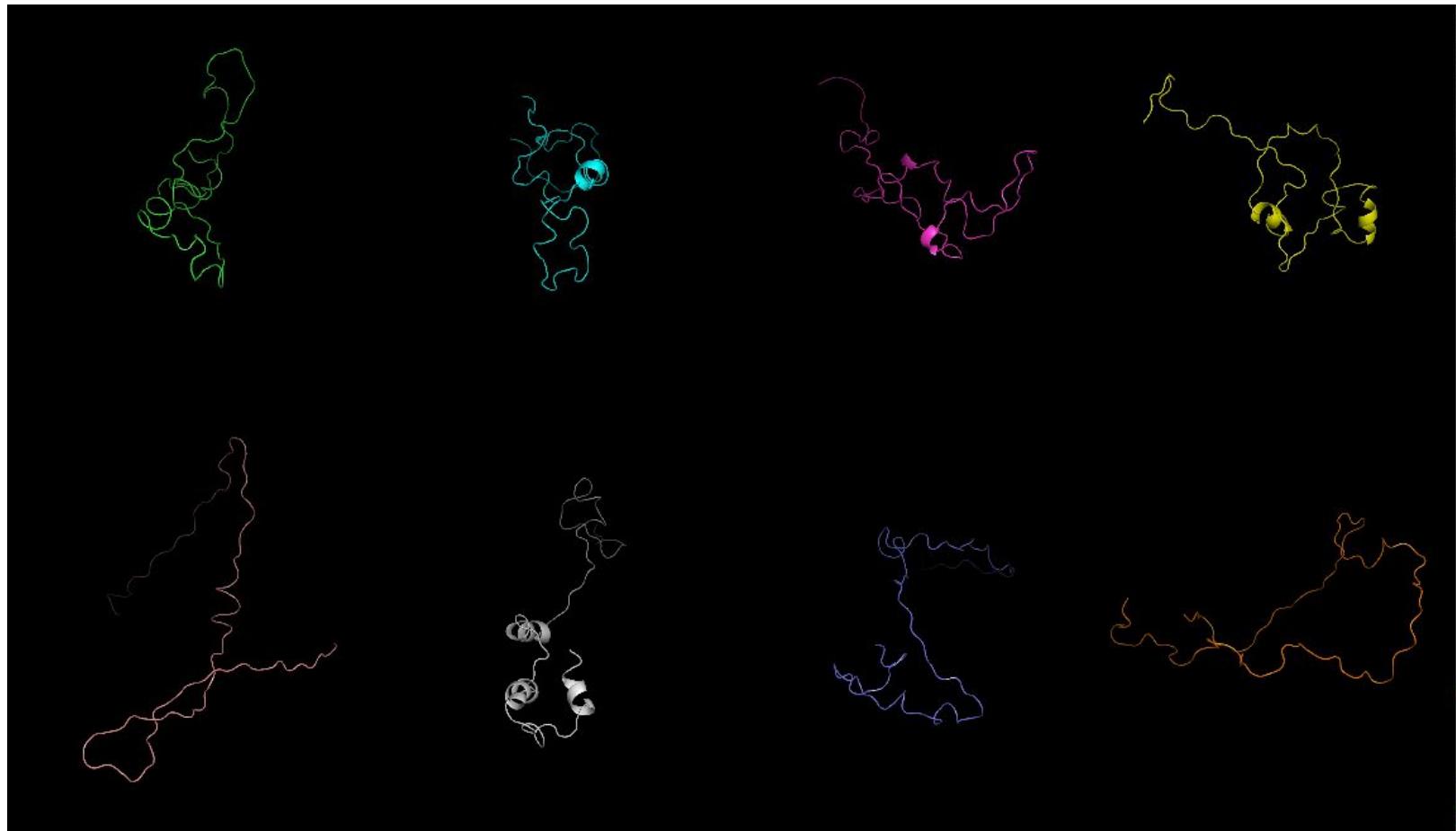
Organismo

Existe un continuo en función de la flexibilidad?

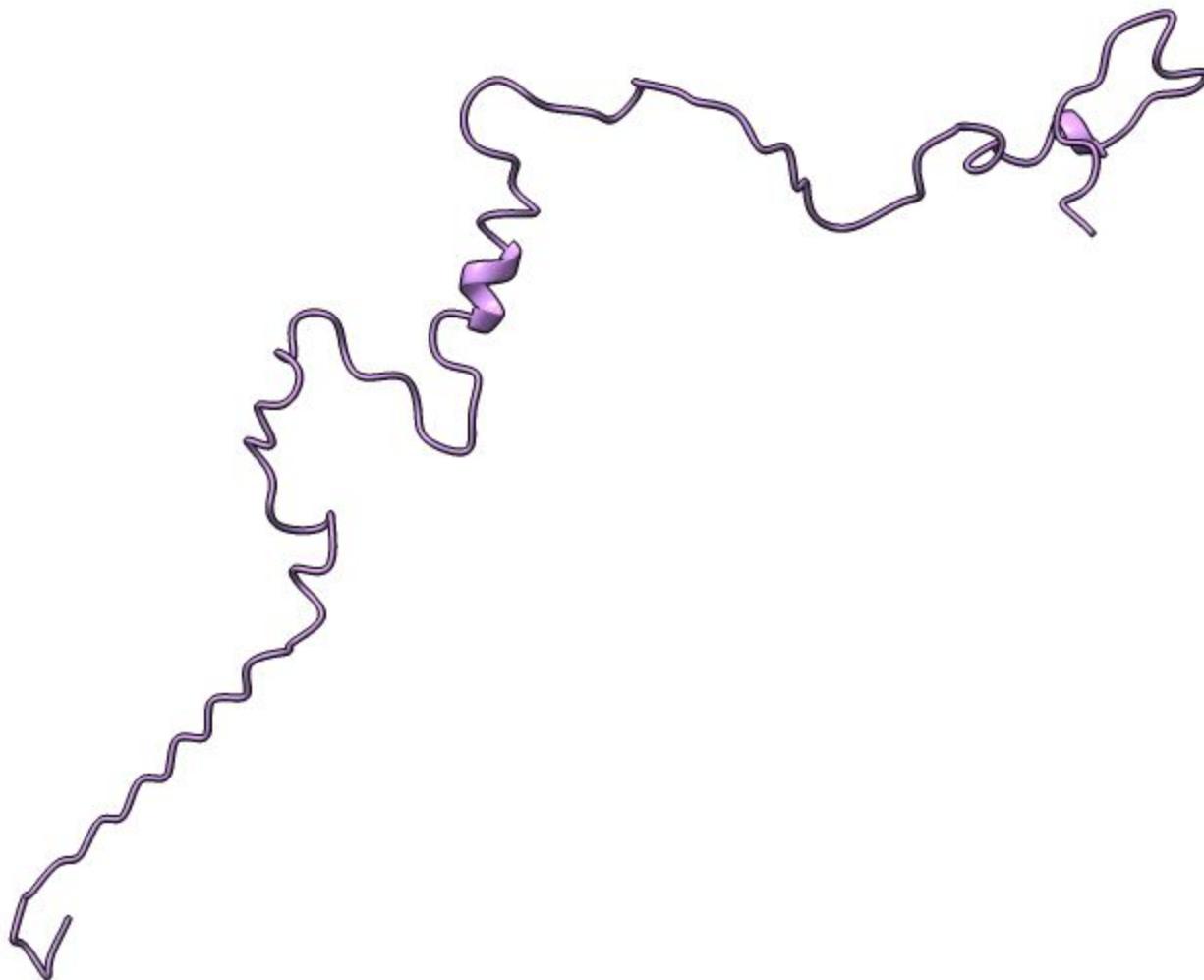


# Proteínas desordenadas

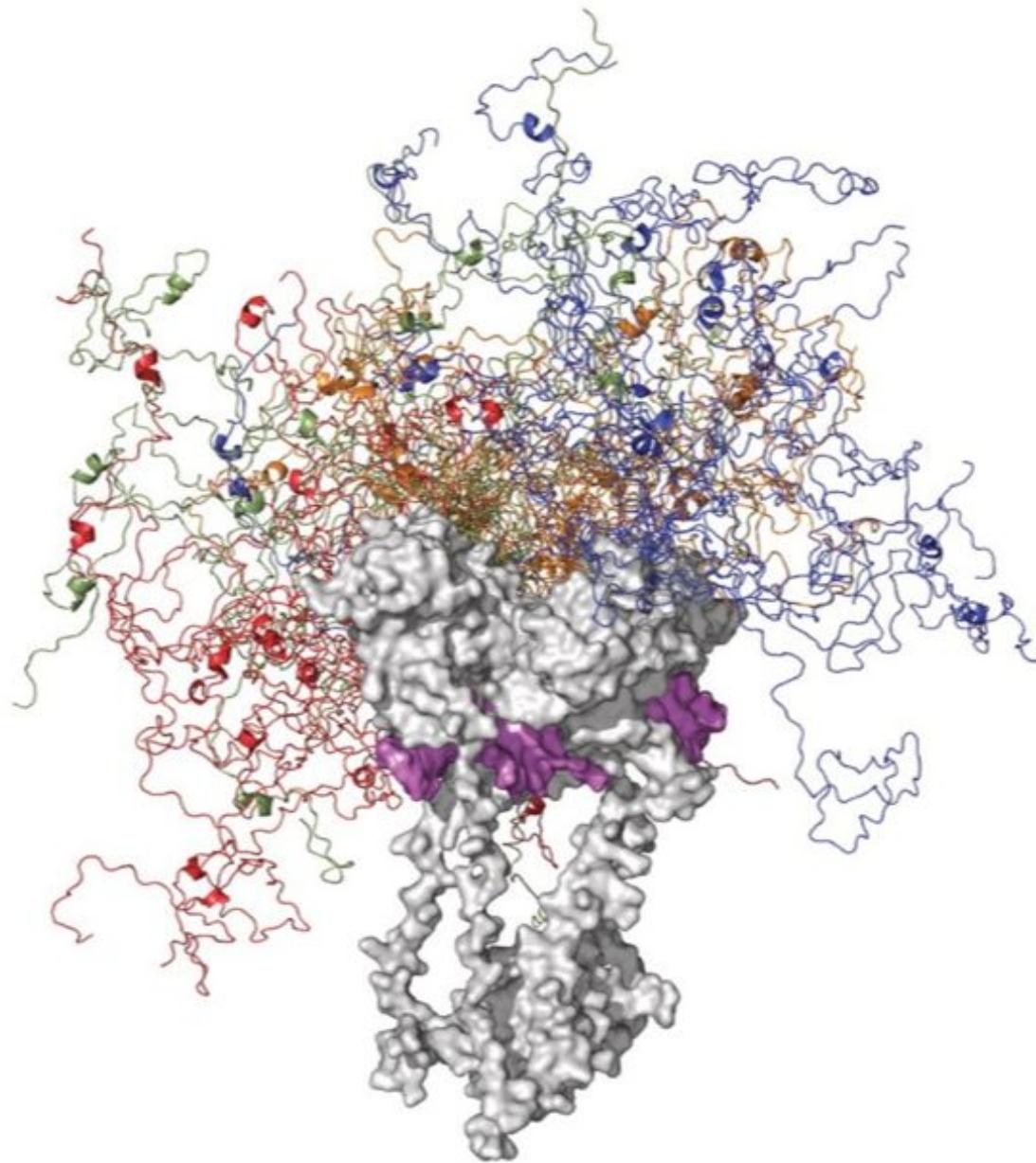
Desafían el “dogma” central de la biología?

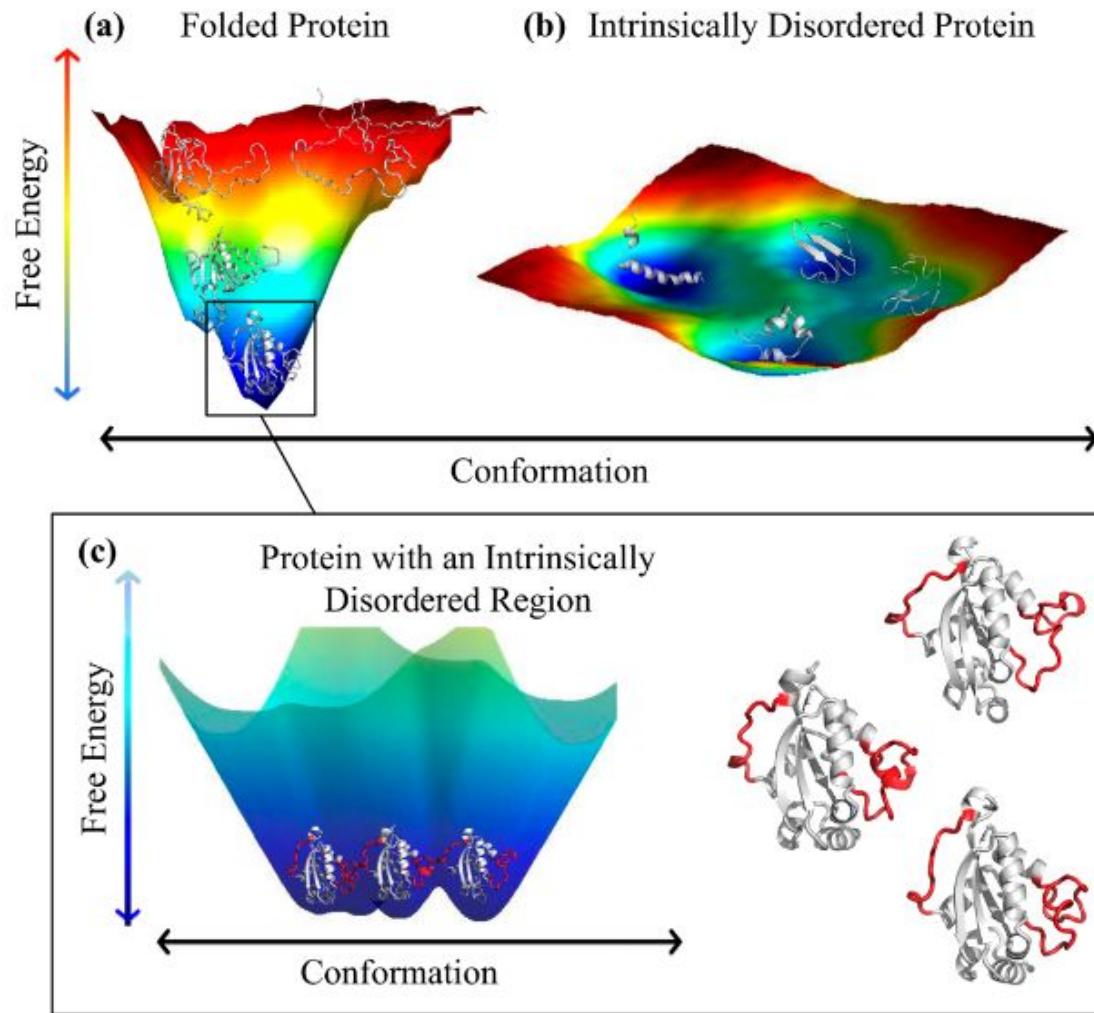


## Un confórmero de una IDP



# Un ensemble de una IDP





# PED<sup>3</sup>: Protein Ensemble Database

The database of conformational ensembles describing flexible proteins

Select experimental data type:  NMR  SAXS  BOTH  ANY

[Search](#)

 [Go!](#)

## Welcome to the Protein Ensemble Database!

### What is an IDP?

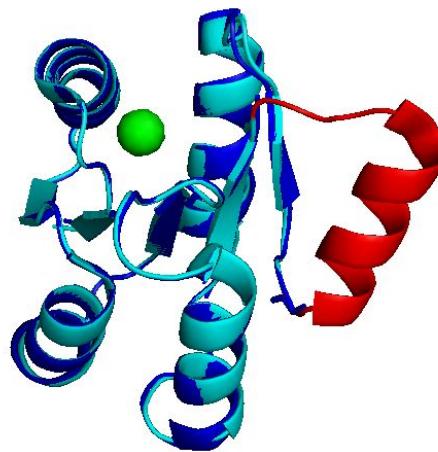
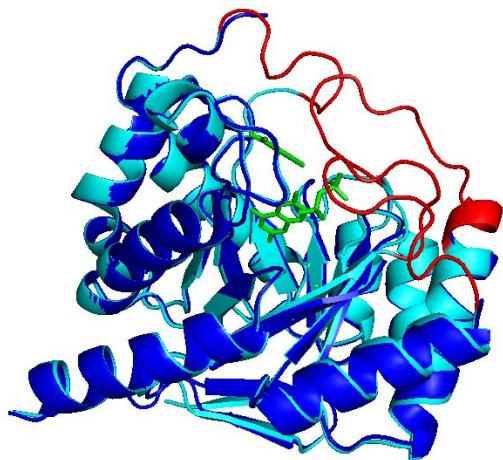
Structural heterogeneity is intimate to the functioning of many proteins and thus describing a protein with a single native structure is often insufficient to elucidate its function. In particular, intrinsically disordered proteins (IDPs) can only be approached by solution techniques and described as structural ensembles. The same is true for multidomain proteins that have disordered linkers. Apparently, the ensemble representations of these proteins carry essential function-related information, yet they have not been available until now.

[Read more about IDPs...](#)

### What is an ensemble?

The goal of PED is to serve as an openly accessible database for the deposition of structural information on IDP- and denatured protein ensembles based on Nuclear Magnetic Resonance (NMR) and Small-angle X-ray Scattering (SAXS) data. We are also hosting purely computational models, typically from Molecular Dynamics (MD) simulations. The deposition of structural coordinates as well as primary data can be used for evaluating and recalculating the ensembles, thus supporting the evolution of new modeling methods leading to much improved skills of connecting "unstructure" with function.

El desorden implica ausencia de  
estructura



# Desorden, taxonomía y porcentajes

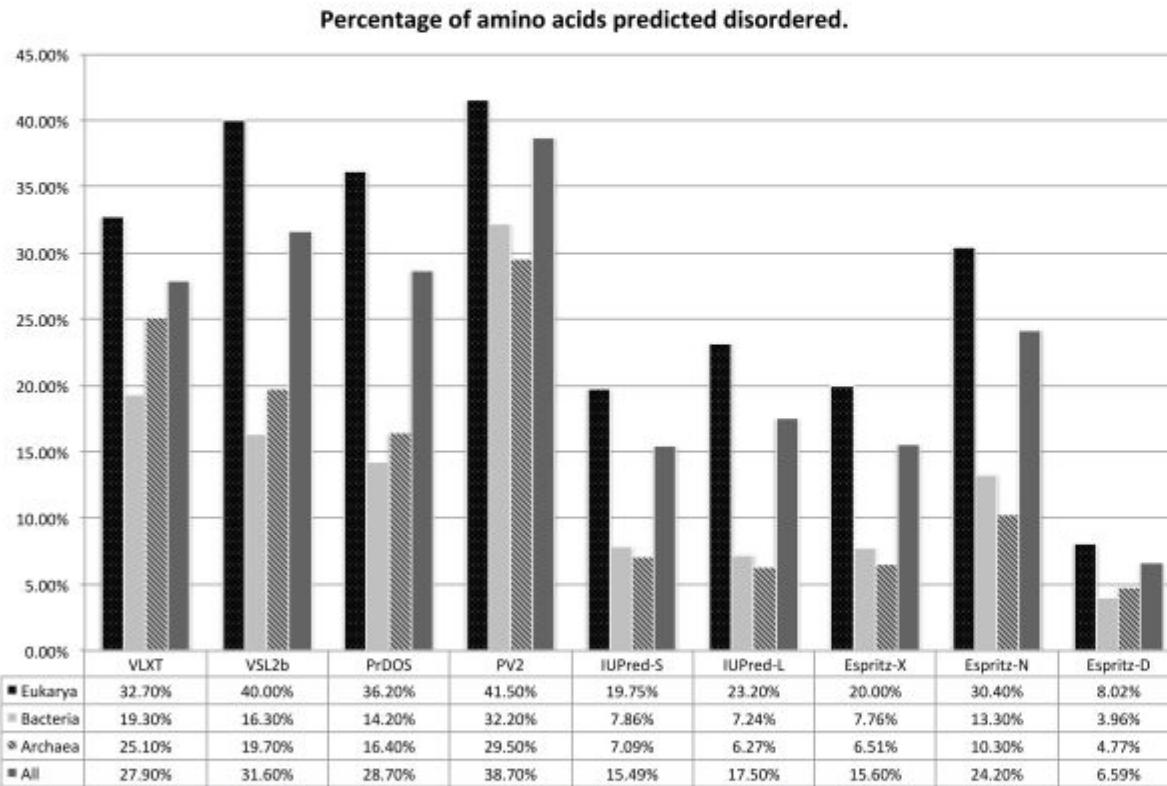
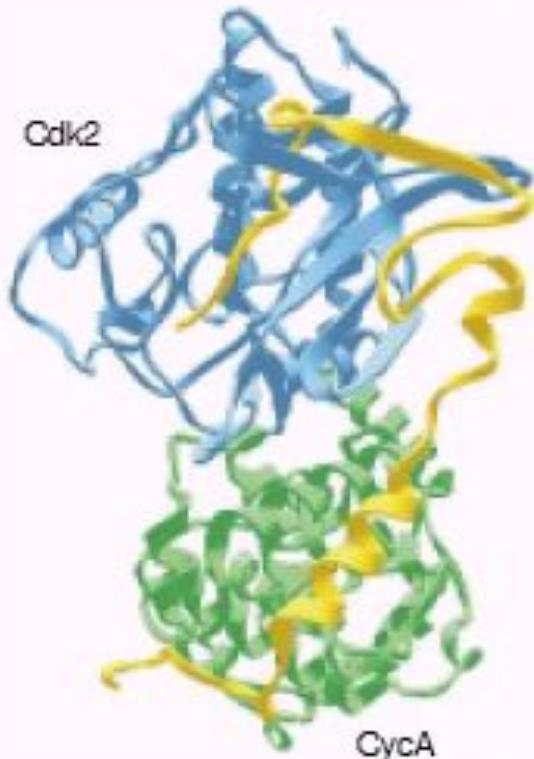


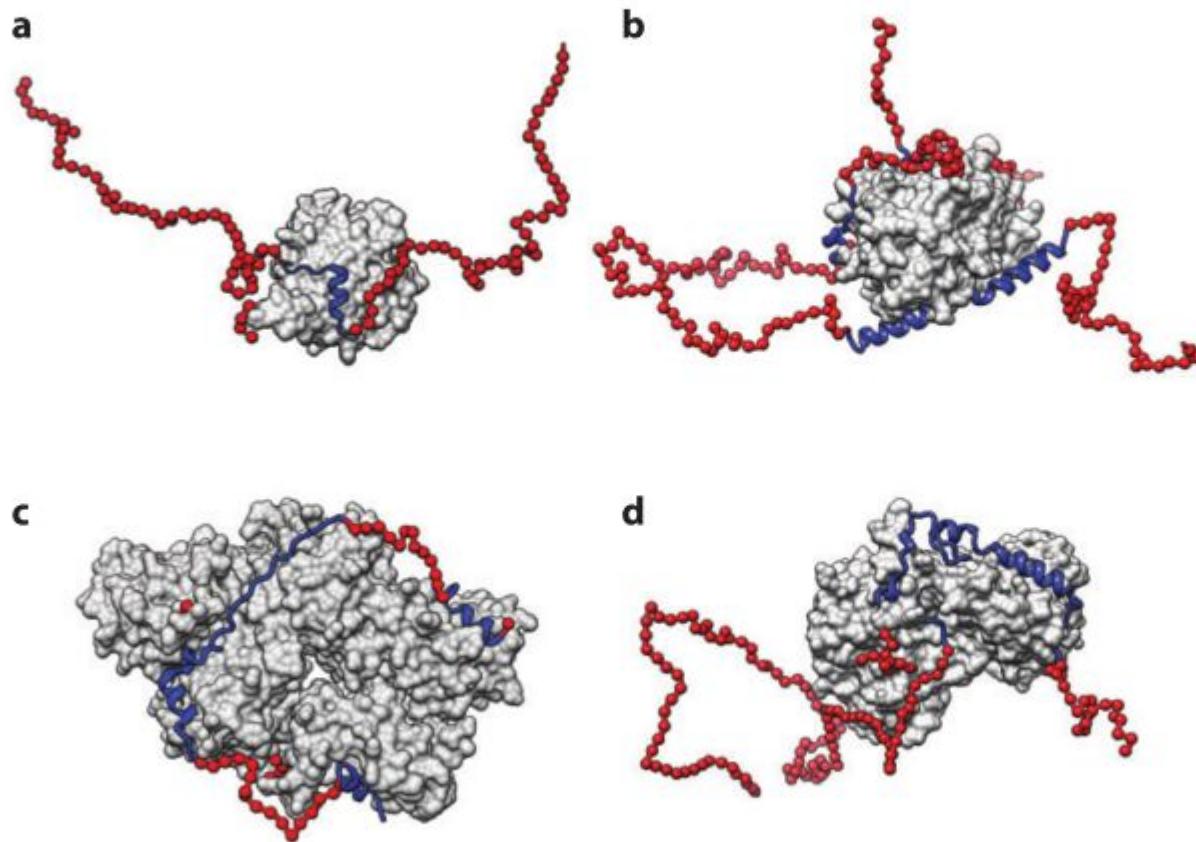
Figure 4. A bar chart grouped by prediction method of global percentage disorder predicted per domain of cellular life. The *X* axis shows results per domain grouped by predictor, the *Y* axis shows the percentage of all amino acid residues for a given domain of life predicted disordered by a given method.

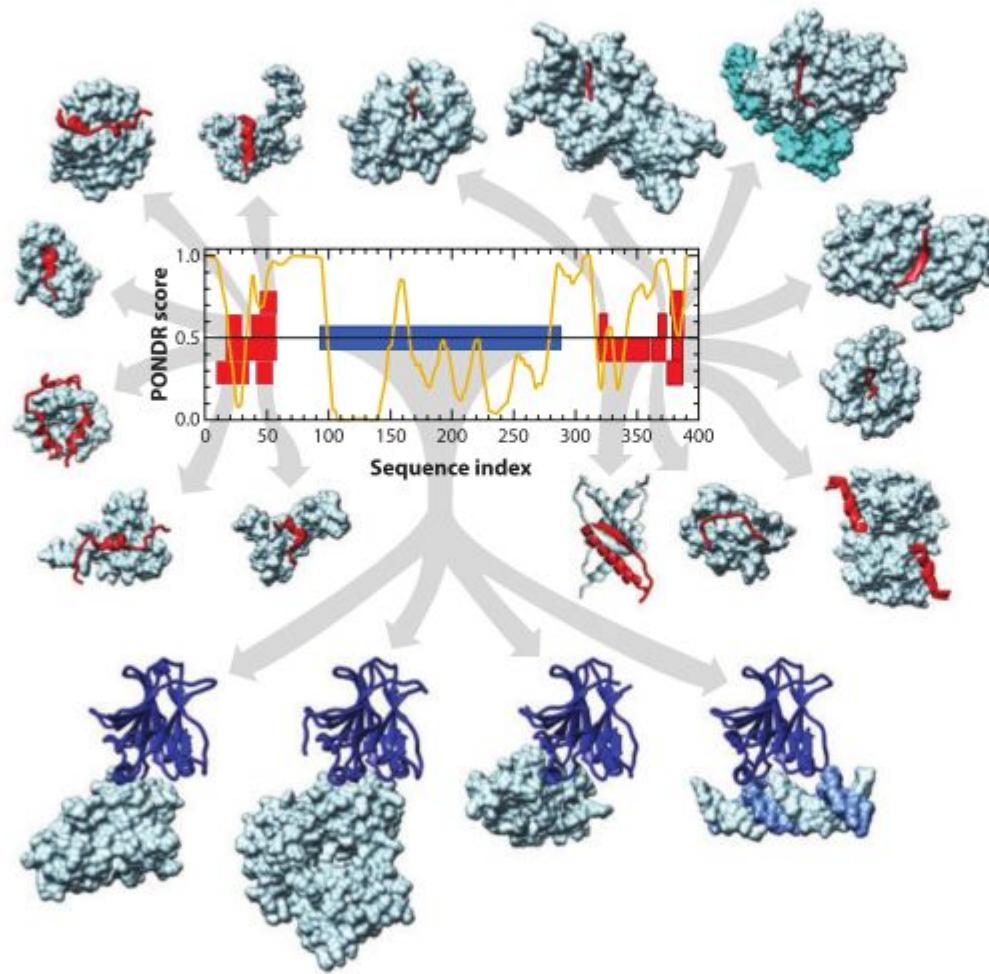
(a)

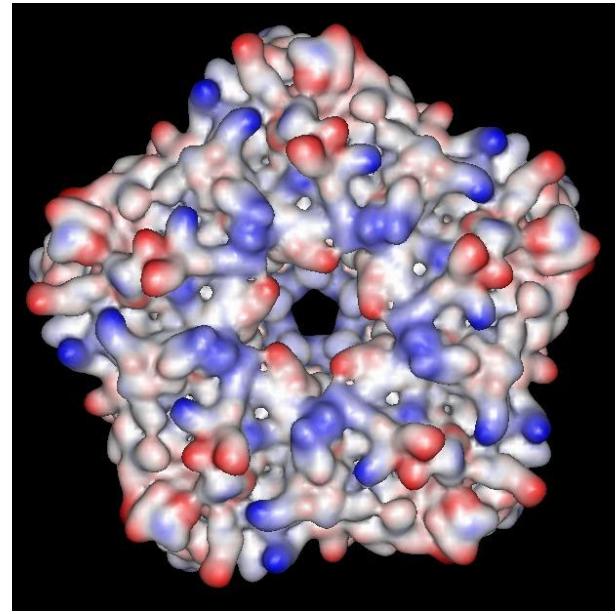
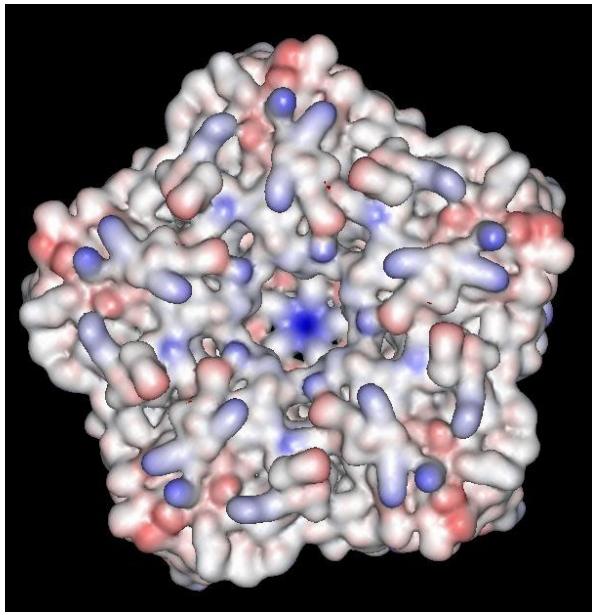
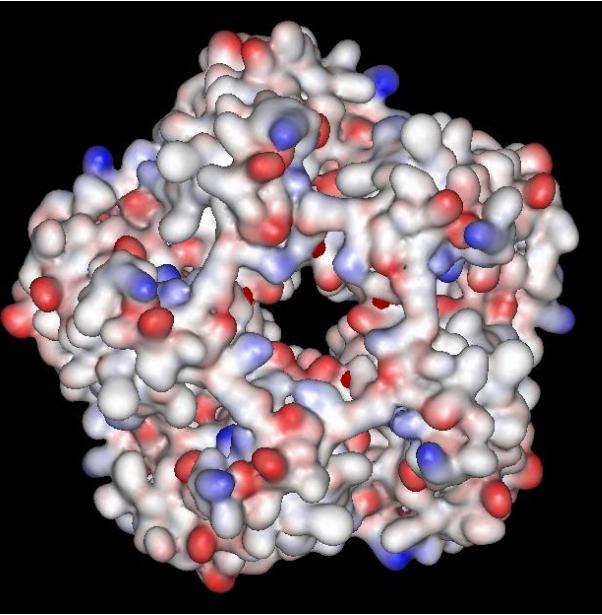


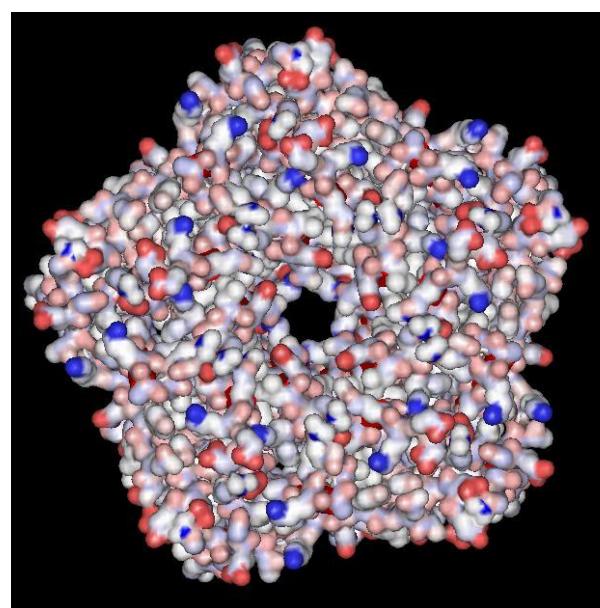
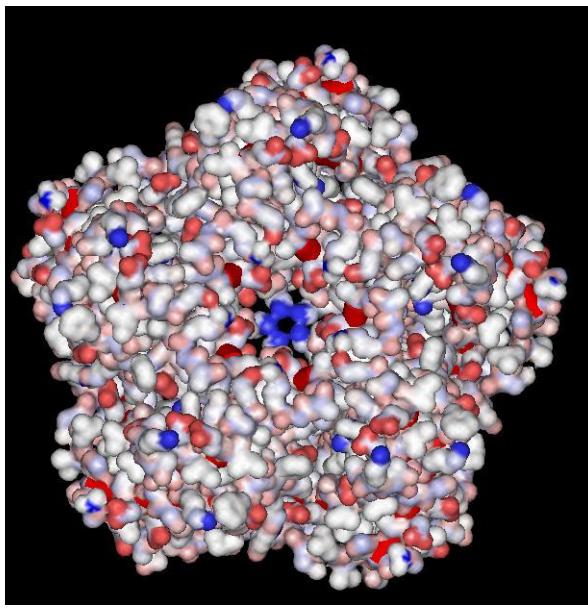
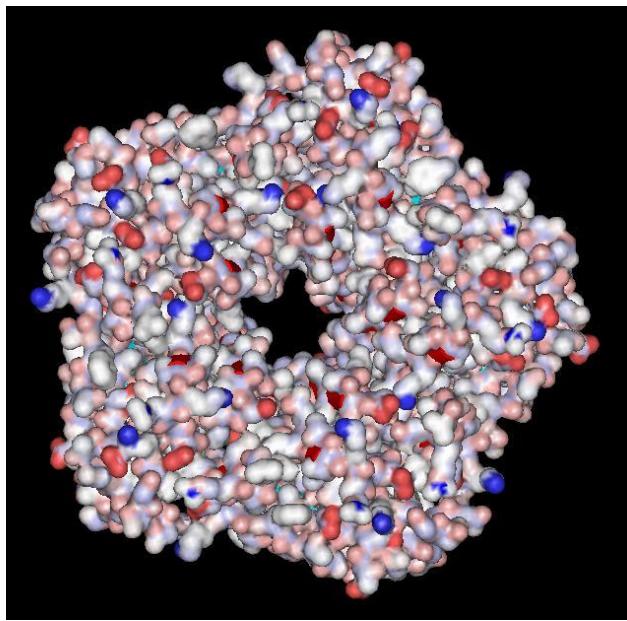
(b)











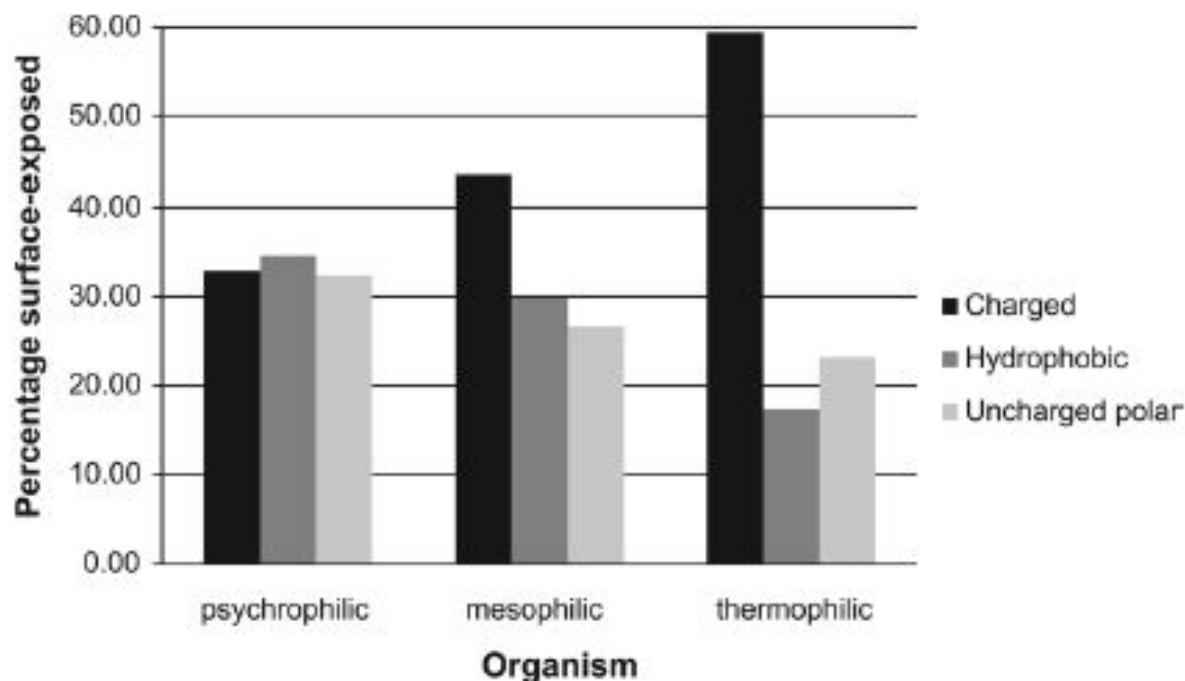


FIG. 1. Charged, polar uncharged, and hydrophobic surface-exposed percentage corresponding to psychrophilic, mesophilic, and hyperthermophilic LSs. This graphic shows the differences in surface amino acid distribution related to the different noncovalent bonds stabilizing LS.