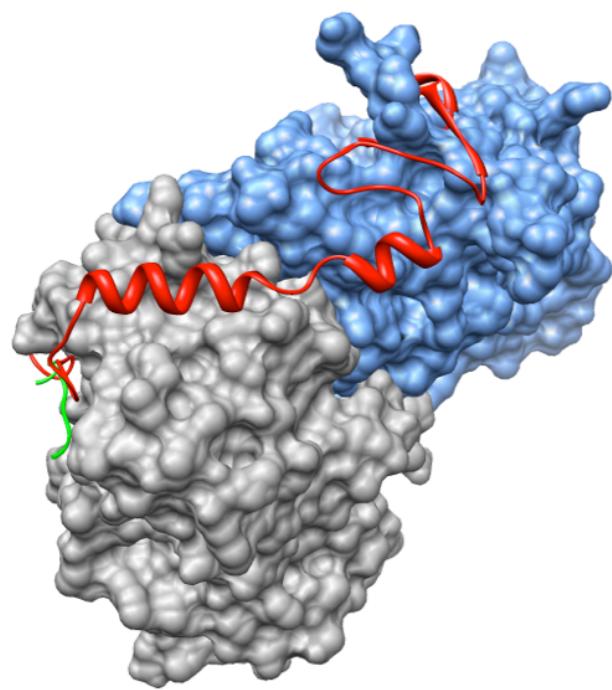
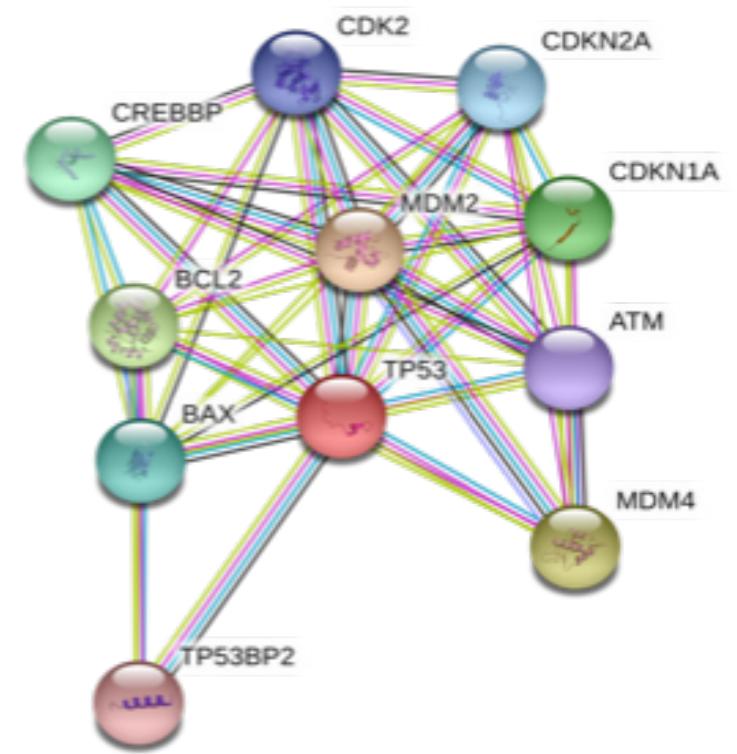


## VISUALIZANDO ALINEAMIENTOS MULTIPLES DE SECUENCIAS Y ESTRUCTURAS DE PROTEINAS

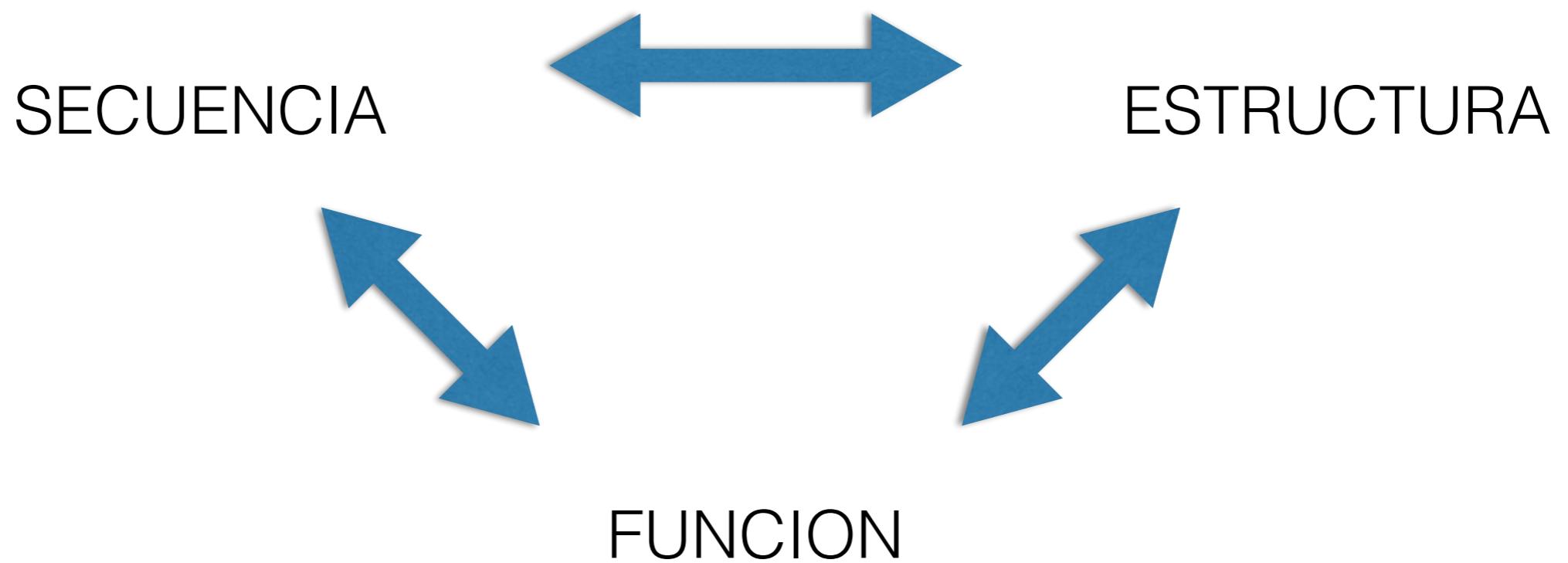
Prof. Lucía Chemes  
 Instituto de Investigaciones Biotecnológicas  
 Universidad Nacional de San Martín



Guías de TP:  
 Lucía Chemes  
 Juliana Glavina

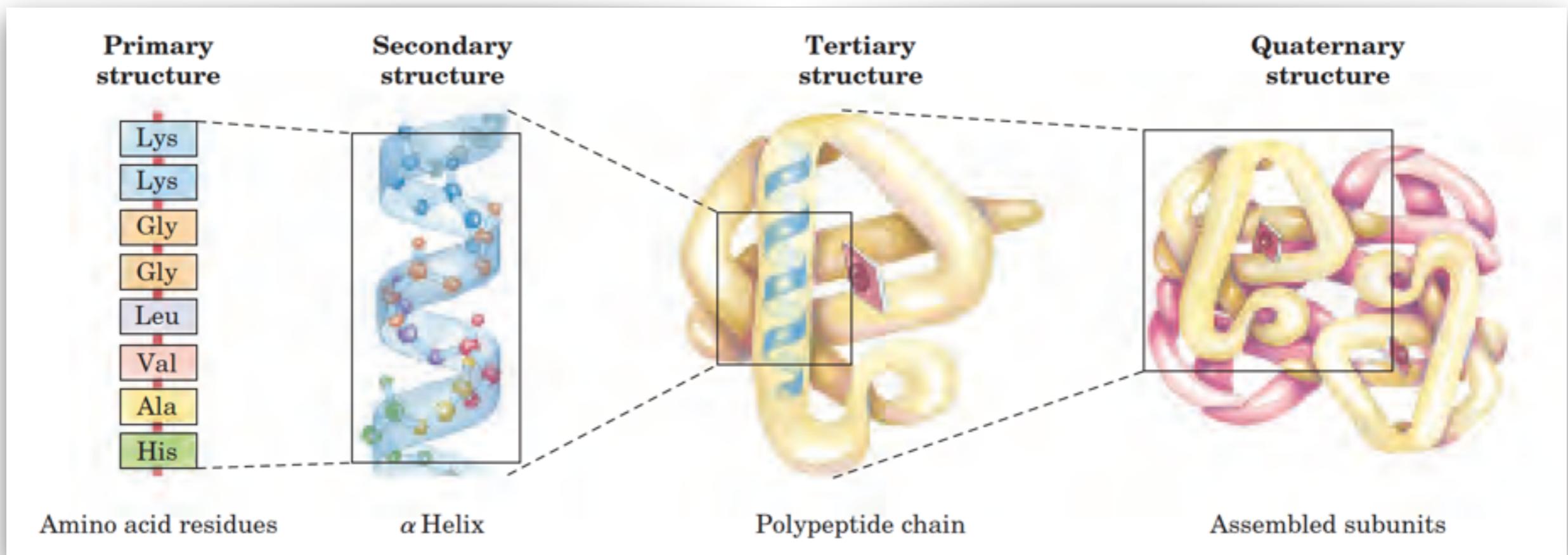


# PROTEINAS

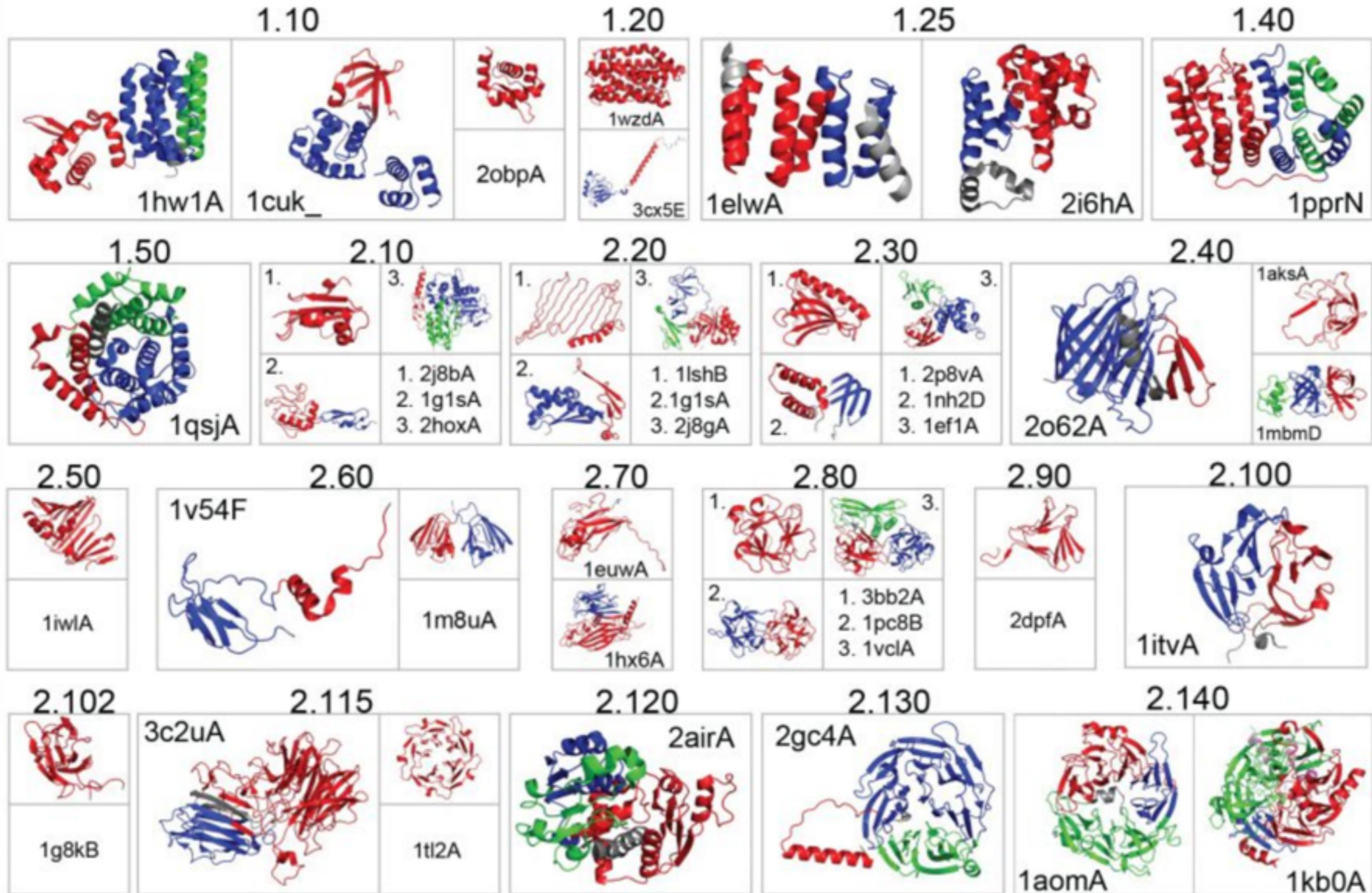


**Las secuencias y las estructuras de las proteínas evolucionan y la información contenida en ambas nos dice mucho acerca su función**

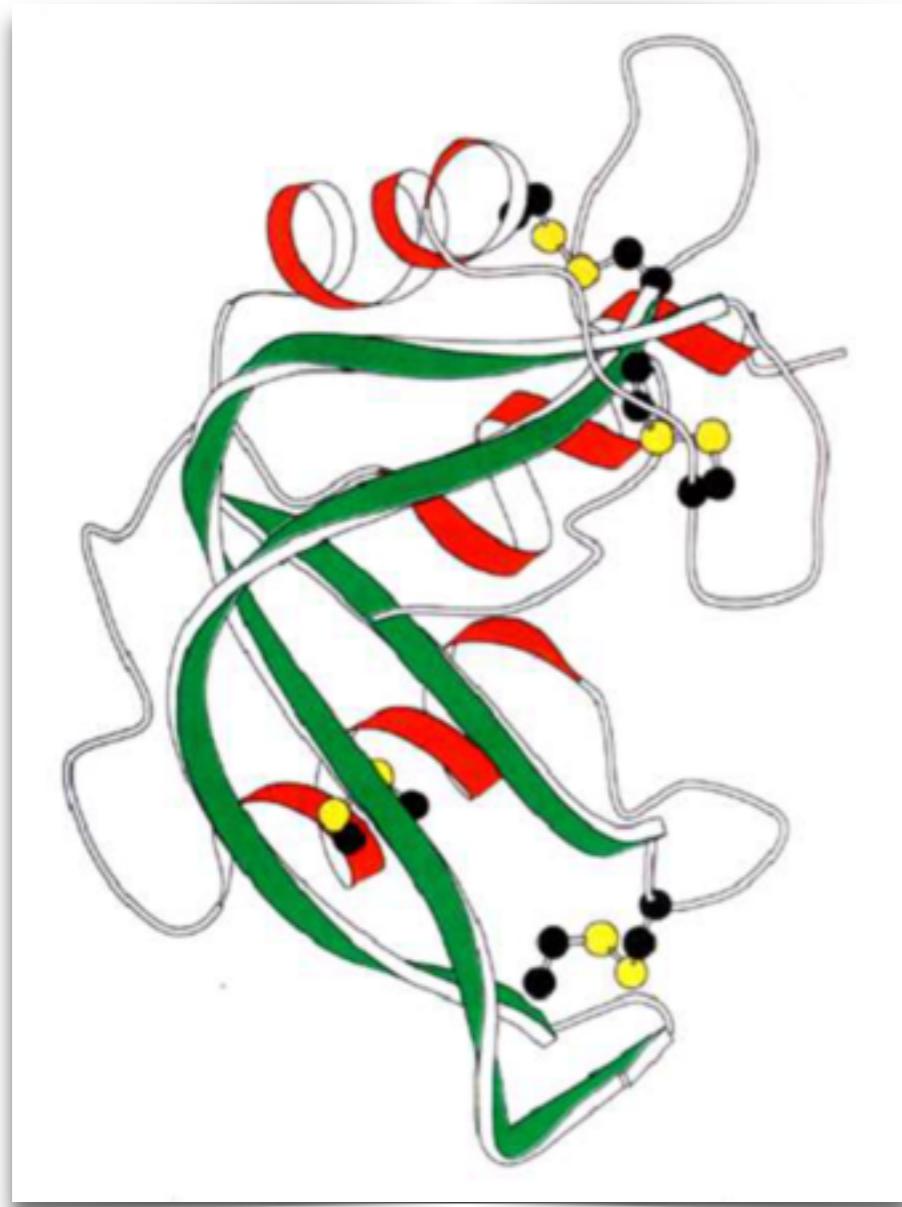
# UNA (MUY) BREVE INTRODUCCION A ESTRUCTURA Y ARQUITECTURA PROTEICA



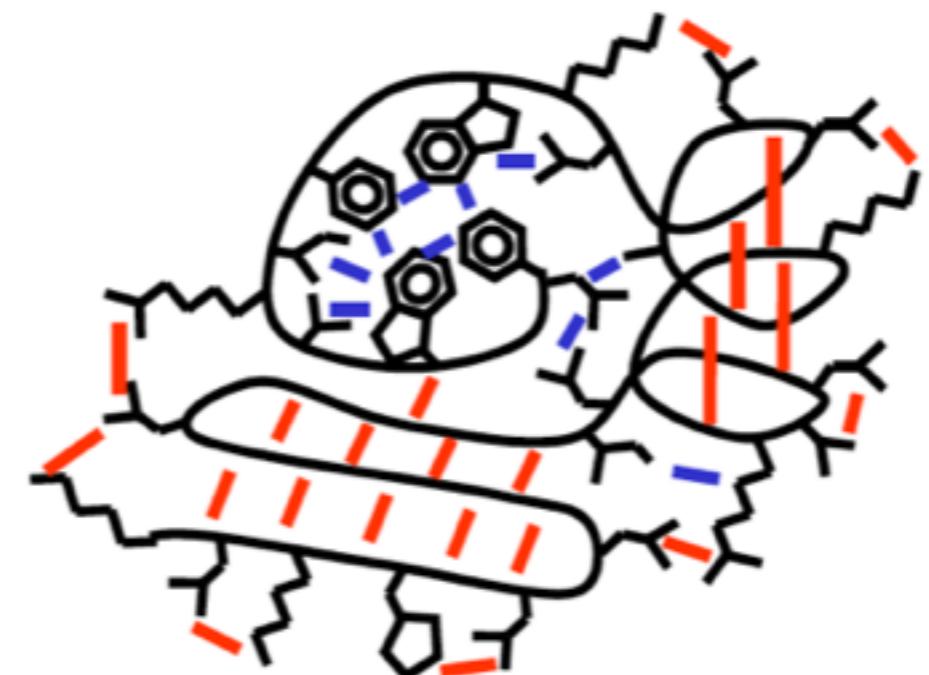
# EXISTE UN UNIVERSO DE DOMINIOS (pero creemos que es limitado)



## dominios plegados



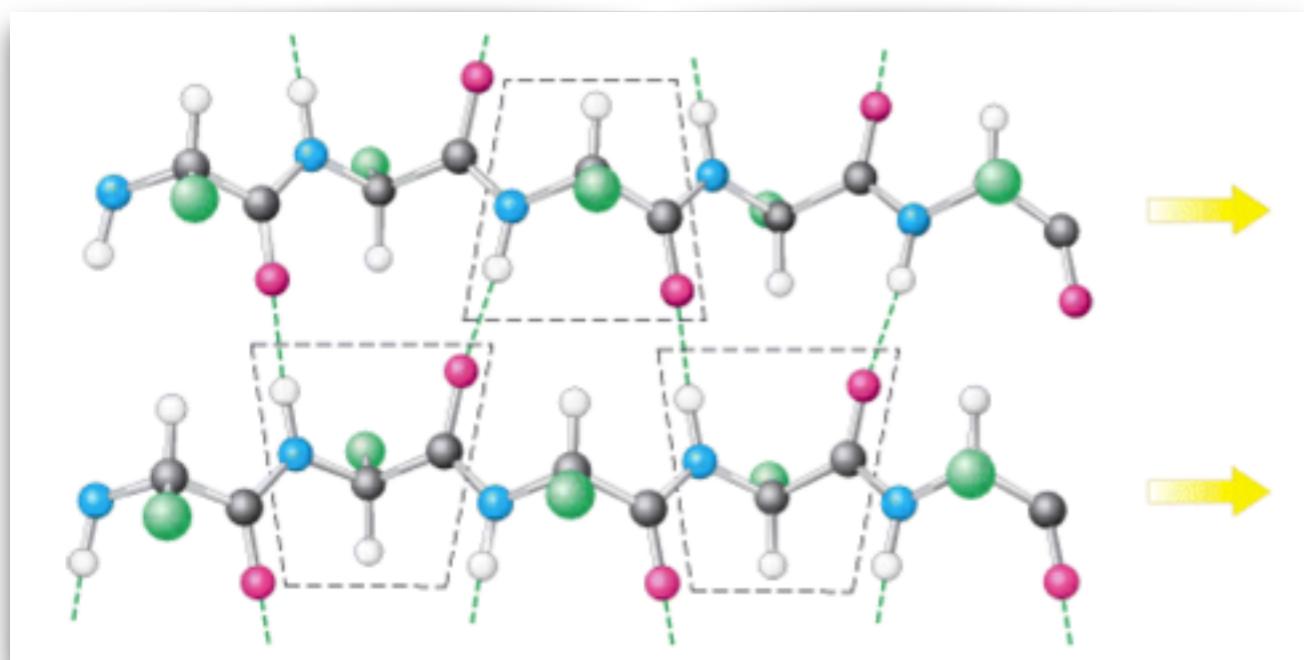
cómo se mantienen  
Estructurados?



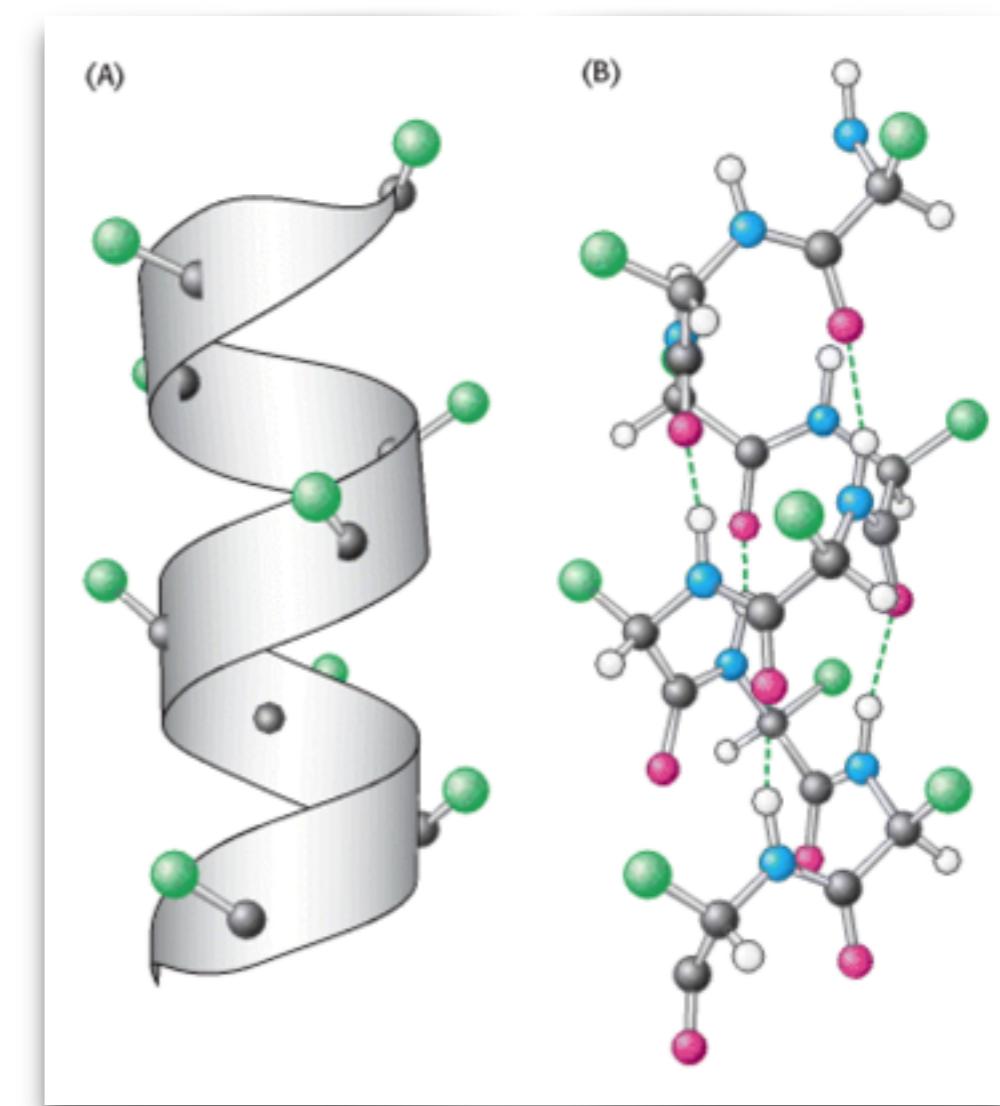
- Interacciones hidrofóbicas
- puentes de hidrógeno

# DOMINIOS GLOBULARES

HOJA BETA



HELICE ALFA



AMBAS ESTABILIZADAS POR PUENTES H



cadena lateral



C



N



H



O

# ARQUITECTURA MODULAR DE LAS PROTEINAS



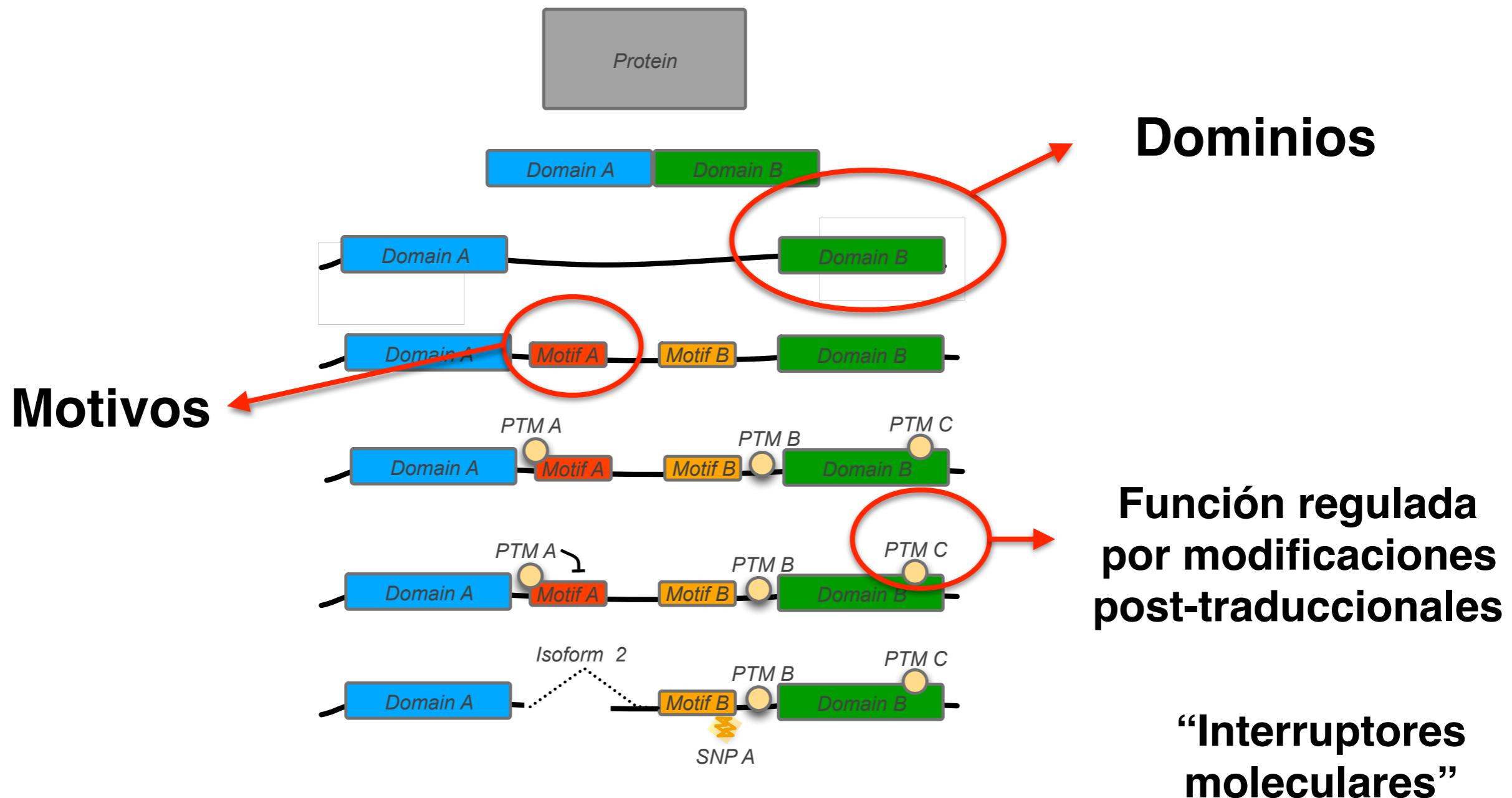
**Podemos describir a las proteínas como un conjunto de módulos**

**Qué es un módulo?**

*una unidad dentro de una proteína que, mediante la interacción con otras moléculas, codifica para una determinada función regulatoria, enzimática o de reconocimiento molecular*

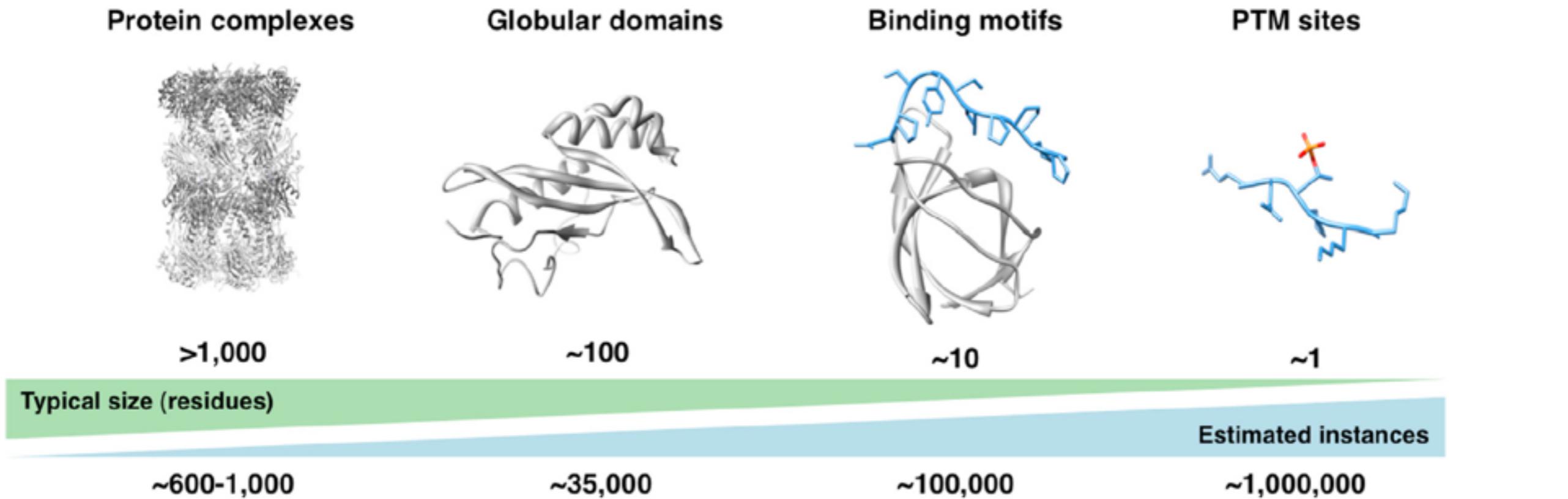
# ARQUITECTURA MODULAR DE LAS PROTEINAS

*Qué tipos de módulos encontramos en proteínas?*



# Qué tipos de módulos existen?

## Protein Module



**Dominios**  
*~1000 superfamilias*

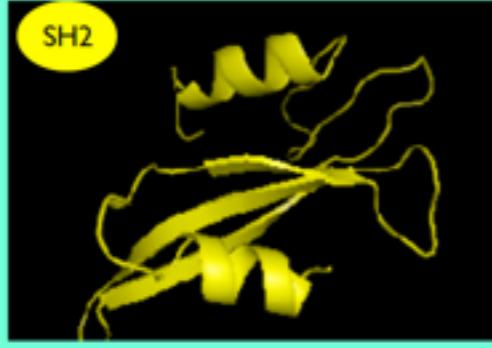
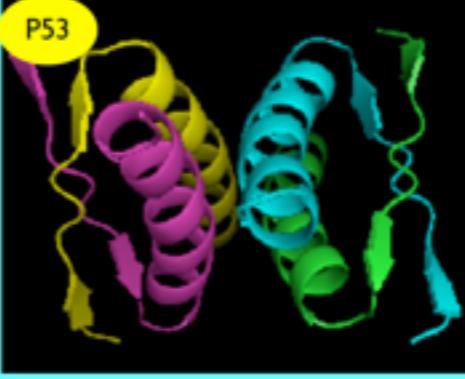
**Motivos**

**Sitios de  
modificación**

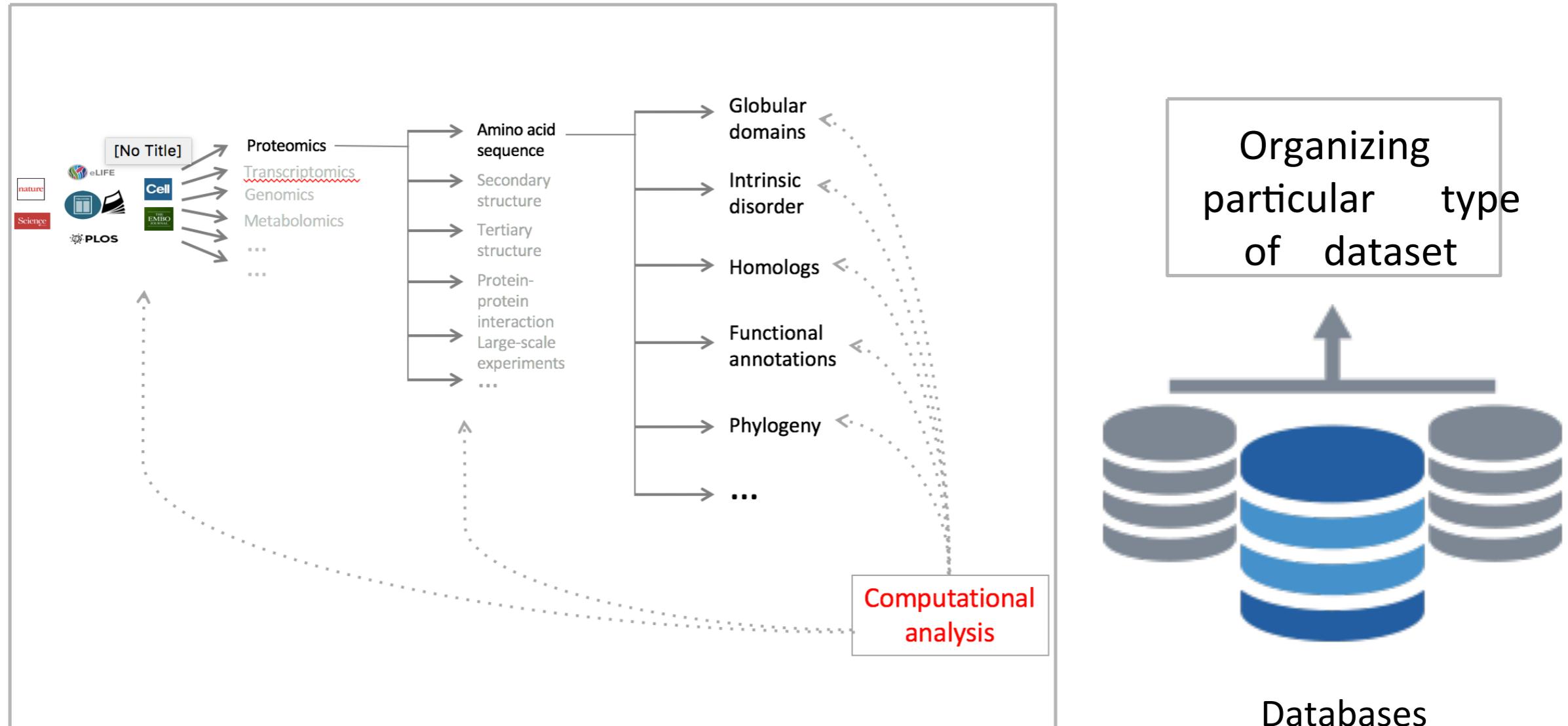
**A million peptide motifs for the molecular biologist**  
Tompa P, Davey NE, Gibson TJ and Babu MM  
Mol Cell (2014)

Fosforilación  
Glicosilación  
Ubiquitinación Etc...

# Qué tipos de módulos existen?

Protein Architecture Modules in Cell Regulation			
Globular ~70%	Natively Disordered ~30%		
	Mutual Fit	Induced Fit	Linear Motif
 SH2 Kinase Phosphatase Acetylase Deacetylase SH3 SH2 PH PDZ Bromo ...	 P53 Coiled Coil Collagen Helix P53 tetramerisation T4 Endonuclease VII HLH DBD ACTR/NCBD ...	 SARA SARA > Smad2 Tcf > beta-Catenin Hif1-alpha > CBP-TAZ Cited2 > CBP-TAZ P27kip1 > CDK ERM C-tail ...	 YPWM NLS / NES / PTSI / KDEL / YPWM / EHI / WRPW / LXXLL / NPF / DPW / RGD ...
<i>Effectors of regulation</i>	<i>Passive components involved in building regulated, often highly dynamic, complexes</i>		

# Qué herramientas bioinformáticas tenemos para investigar la estructura y función de las proteínas?



## Bases de datos de referencia de proteínas

- Uniprot
- Interpro

## Bases de datos especializadas

- PFAM (dominios)
- RCSB-PDB (estructuras)

# Herramientas para el estudio de secuencias proteicas: Análisis de alineamientos múltiples de secuencia

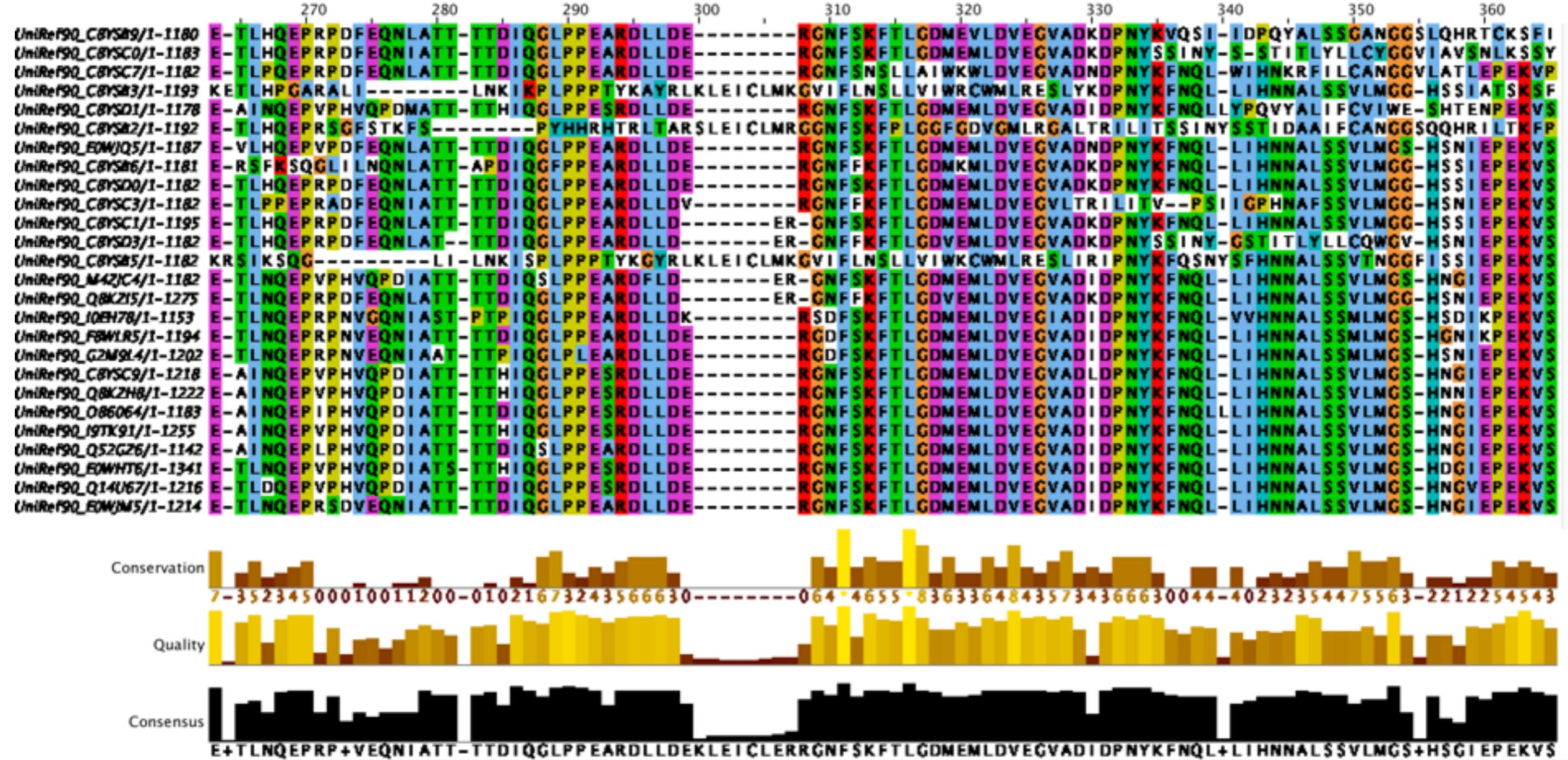
*Qué información se encuentra contenida en un MSA?*



- El patrón de conservación nos permite identificar módulos funcionales
- El patrón de sustitución del MSA contiene información sobre la trayectoria evolutiva de las proteínas que es muy valiosa para analizar su estructura y función

A nivel programático esto permite generar PSSMs y profiles

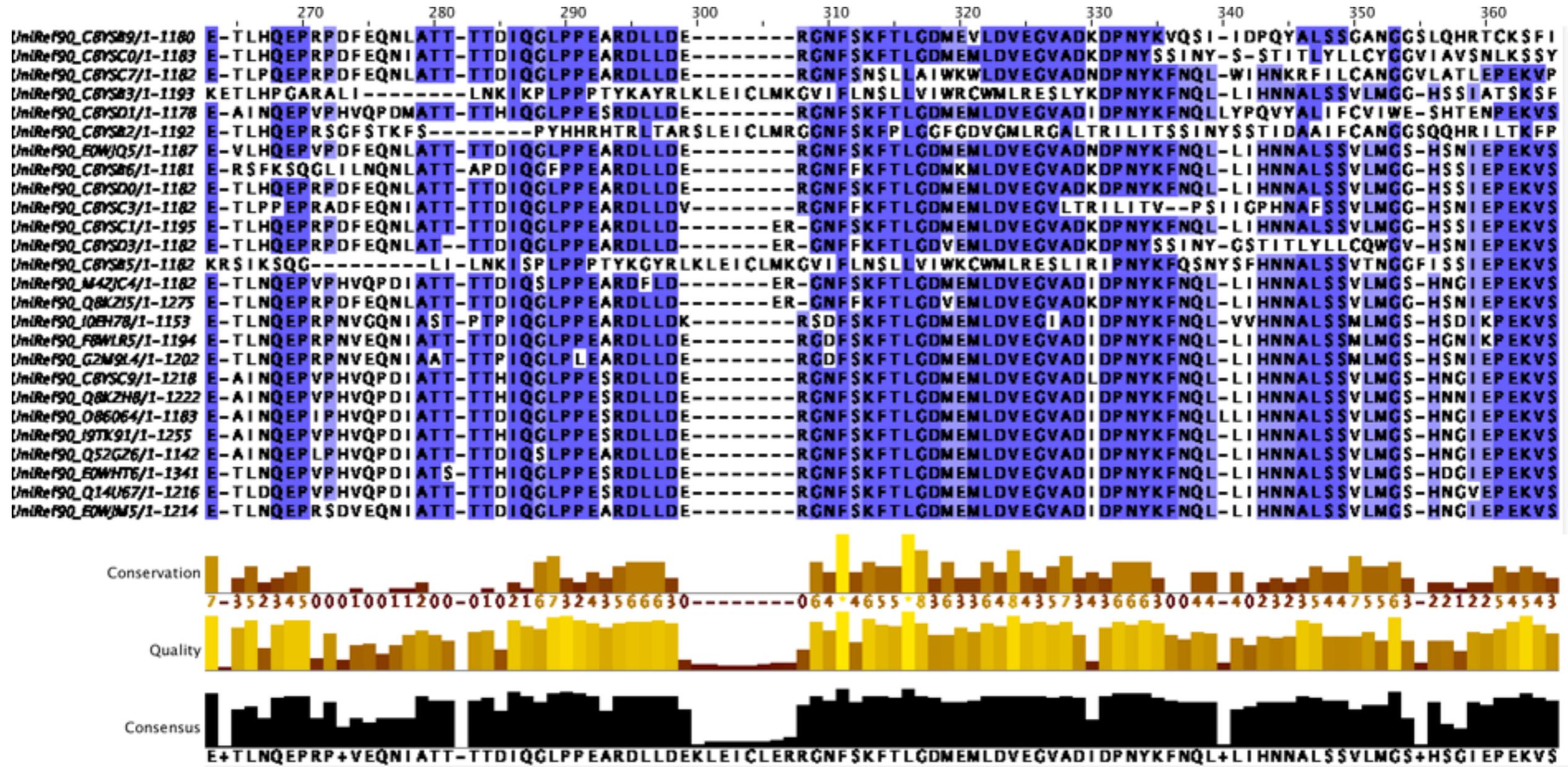
# Qué vemos en un MSA?



Identificación de regiones conservadas y variables  
 límites de dominios  
 variación en la composición aminoacídica de diferentes regiones



# Qué vemos en un MSA?



variación en el % identidad de diferentes regiones

# Herramientas para el estudio de estructuras proteicas: CRISTALOGRAFIA DE RAYOS X

CONSTITUYE LA TECNICA DE MAYOR RESOLUCION CONOCIDA  
PARA ANALIZAR LA ESTRUCTURA DE BIOMOLECULAS

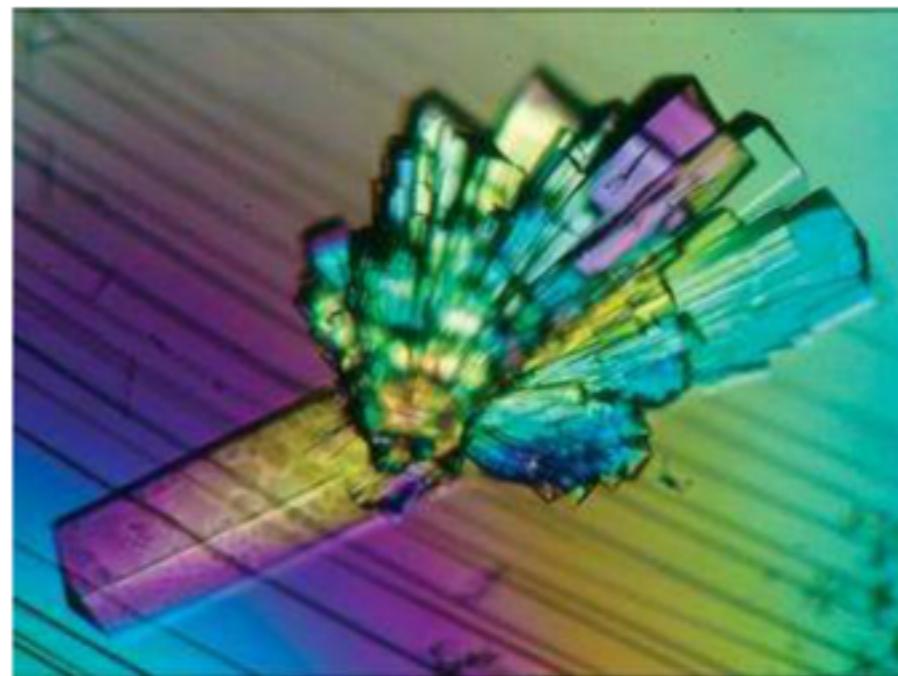
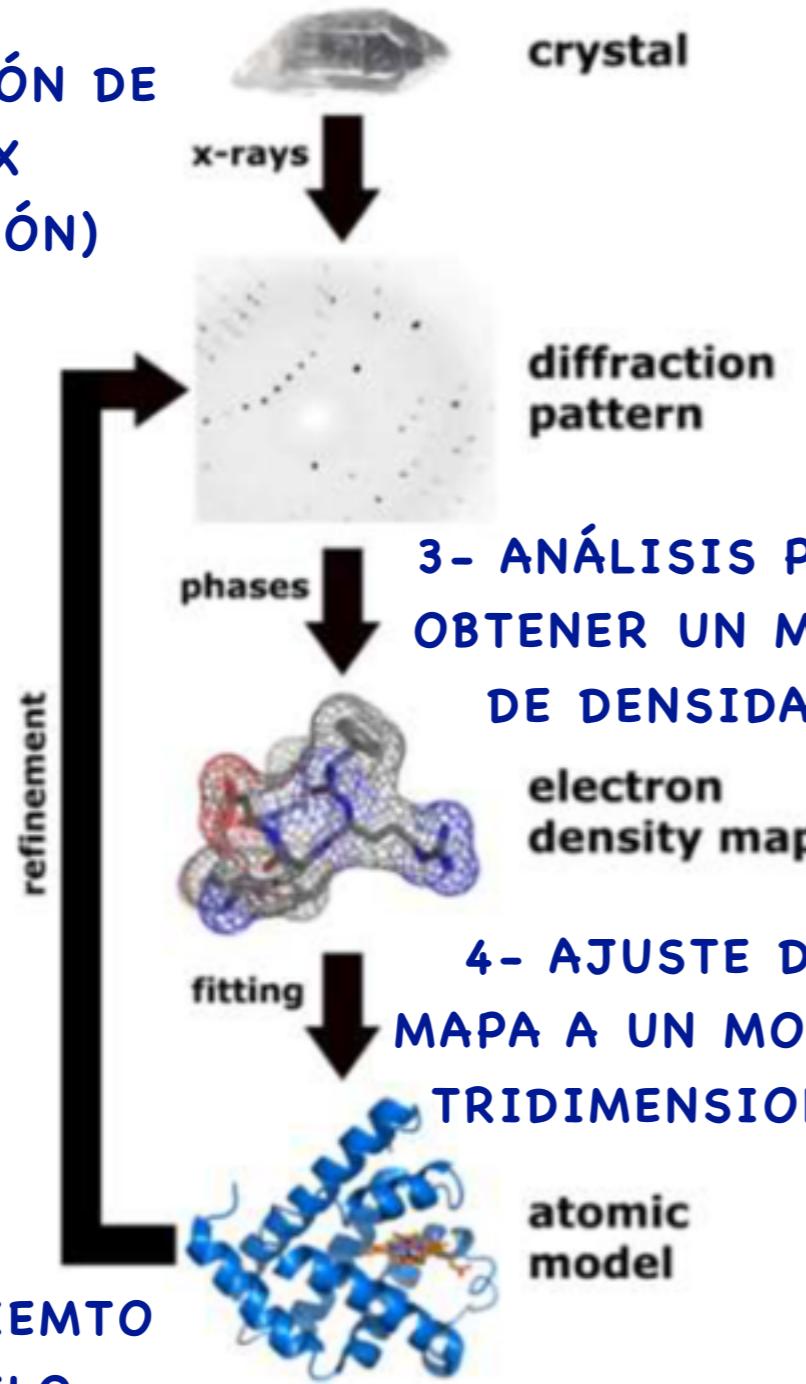
MUCHAS MOLÉCULAS ORIENTADAS IGUAL  
LAS PROTEINAS DEBEN ENCONTRARSE EN UNA FORMA  
CRISTALINA, ALTAMENTE ORDENADA

UNA FUENTE DE RAYOS X  
SU LONGITUD DE ONDA (1.54 Å) ES DEL ORDEN DE LA DISTANCIA  
DE UN ENLACE, POR LO QUE DIFRACTA A TRAVÉS DE LA PROTEÍNA  
UN DETECTOR

# BASES DE LA TECNICA

1- OBTENER UN CRISTAL DE PROTEÍNA

2- DIFRACCIÓN DE  
RAYOS X  
(SINCROTRÓN)

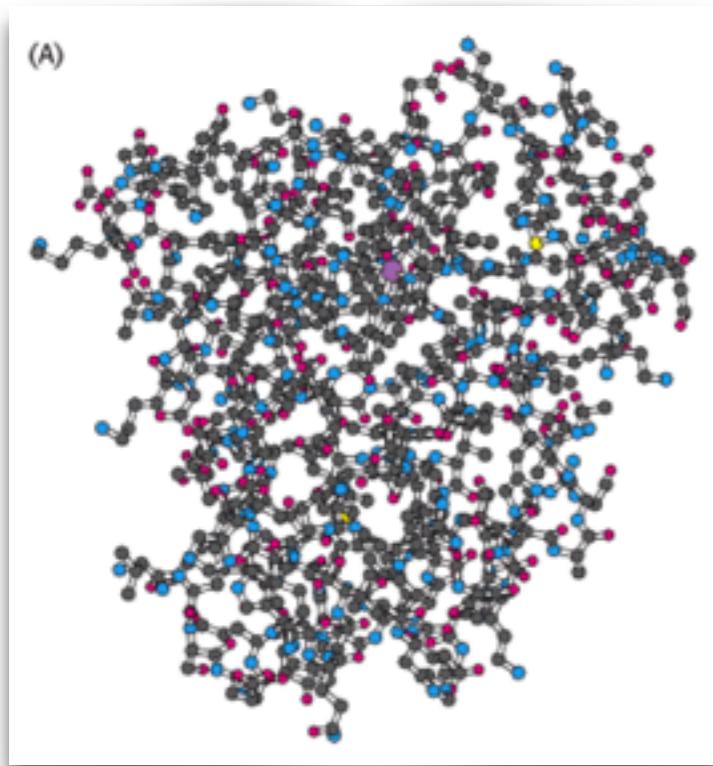


5- REFINAMIENTO DEL MODELO



# Qué vemos en la estructura de una proteína?

*depende del nivel de representación....*

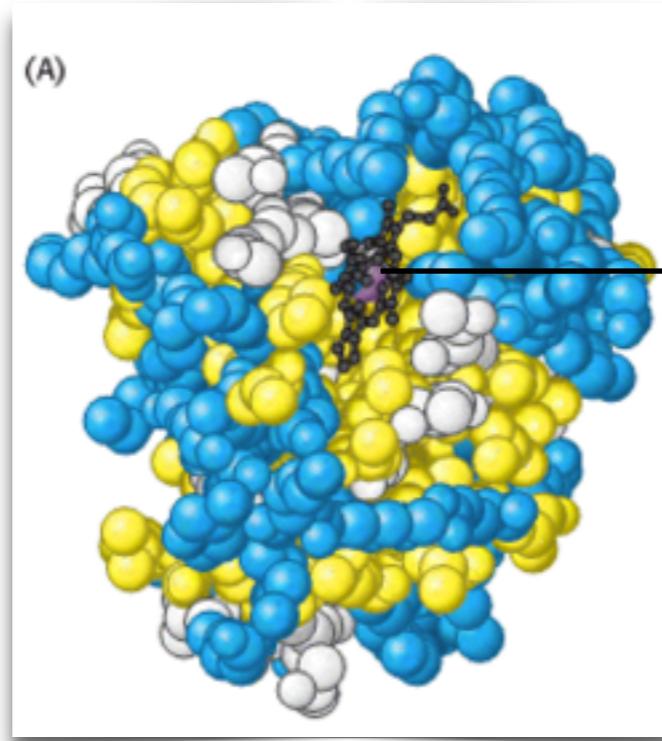


átomos

**MIOGLOBINA**



estructura secundaria



Superficie

→ unión de ligandos

# EVOLUCIÓN DE DOMINIOS DE UNIÓN A OXIGENO

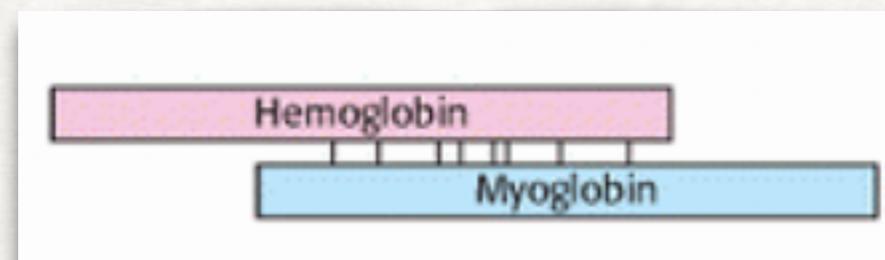
Human hemoglobin ( $\alpha$  chain)

VLSPADKTNVKAAGKVGAGAHAGEYGAELERMFLSFPTTKTYFPHFDLSHG  
SAQVKGHGKKVADALTNAVAHVDDMPNALSDLHAHKLRVDPVNFKLLS  
HCLLVTLAAHLPAEFTPAVHASLDKFLASVSTVLTSKYR

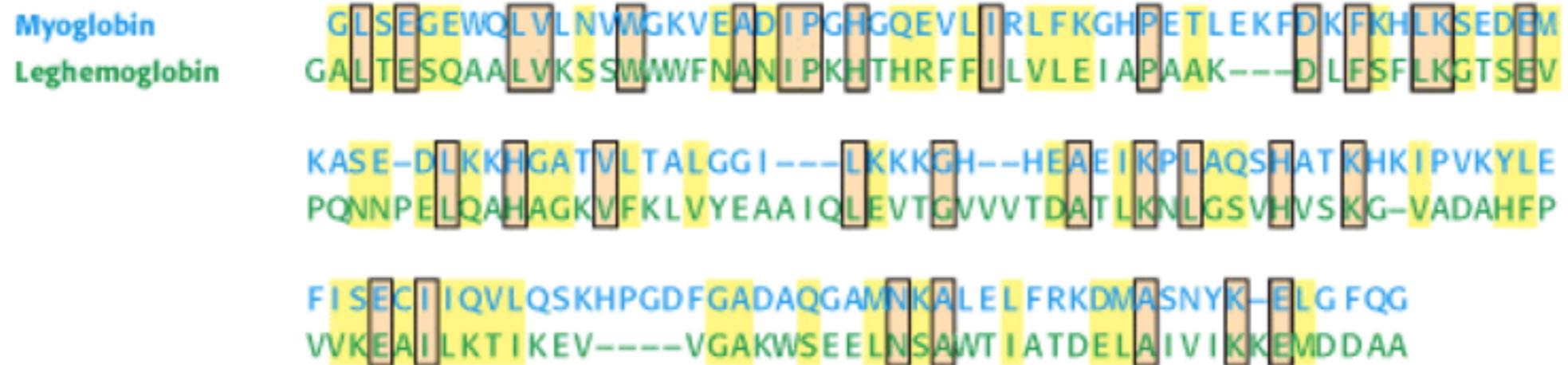
Human myoglobin

GLSDGEWQLVNWGKVEADI PGHQEVLI RLFKGHPETLEKFDKFHLKS  
EDEMKASEDLKKHGATVL TALGGILKKKGHEAEIKPLAQSHATKHKIPVK  
YLEFISECII QVLQSKHPGDFGADAQGMNKALELFRKDMASNYKELGFQG

Secuencias de Hemoglobina y Mioglobina humanas



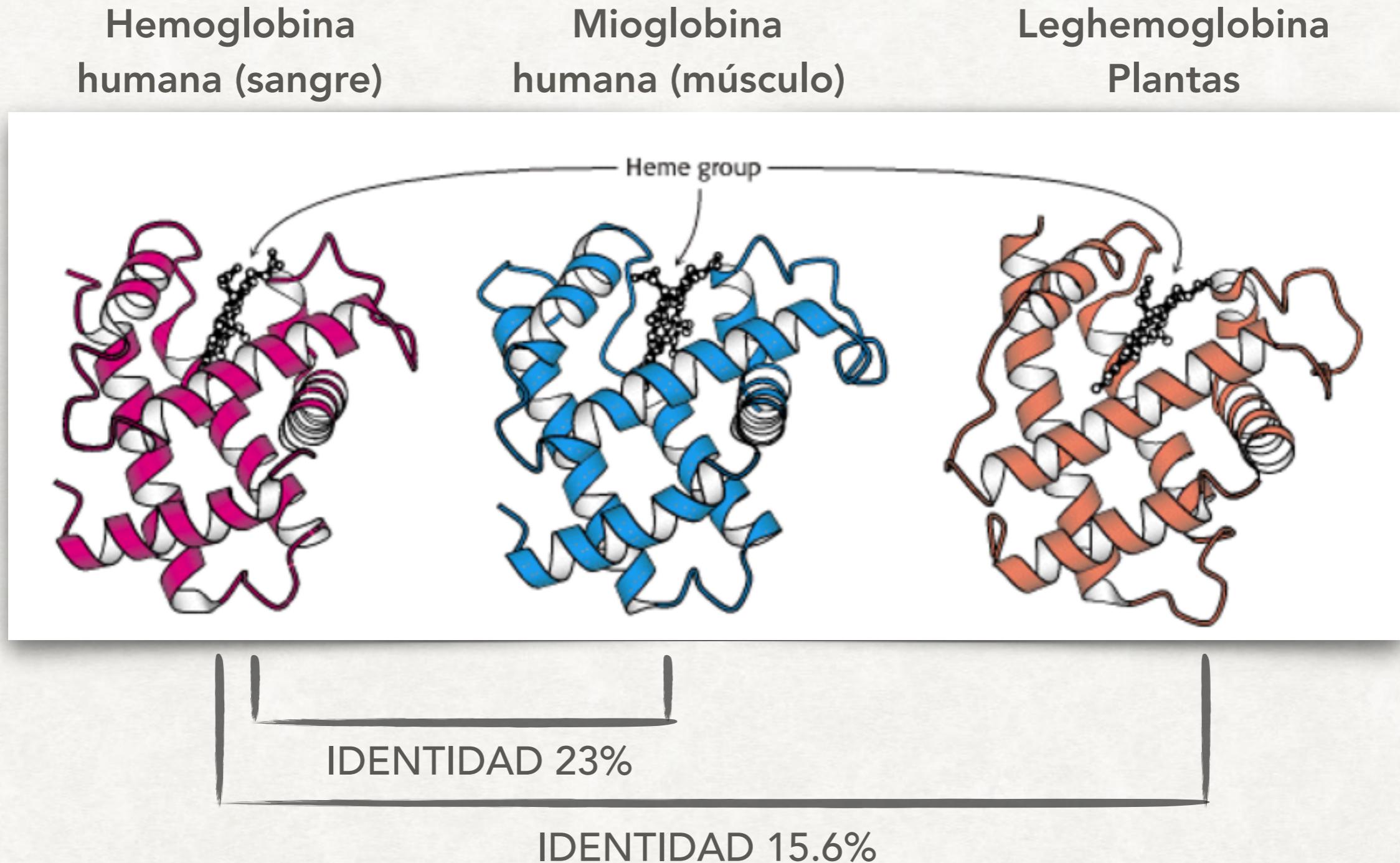
Alineamiento  
**23% IDENTIDAD**



NARANJA: posiciones idénticas

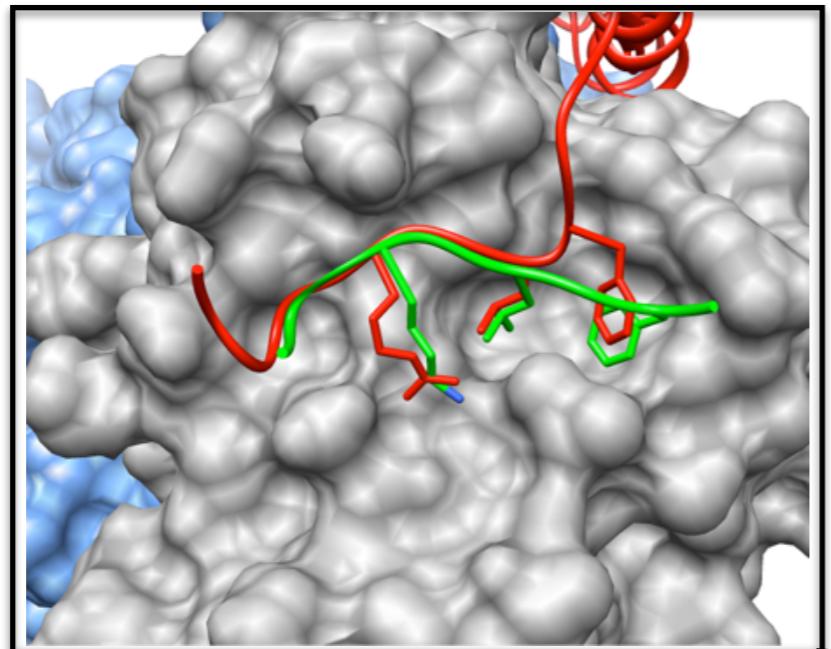
AMARILLO: sustituciones conservativas

# LA ESTRUCTURA SE CONSERVA MAS QUE LA SECUENCIA



Una vez que un nuevo dominio funcional aparece, su estructura se conserva más que su secuencia. Esto indica que existe una alta presión selectiva sobre la estructura

# Chimera: Visualización estructural de dominios, sitios activos y motivos de unión en proteínas



**motivos de unión**

*pueden ser representados por expresiones regulares o por PSSMs*

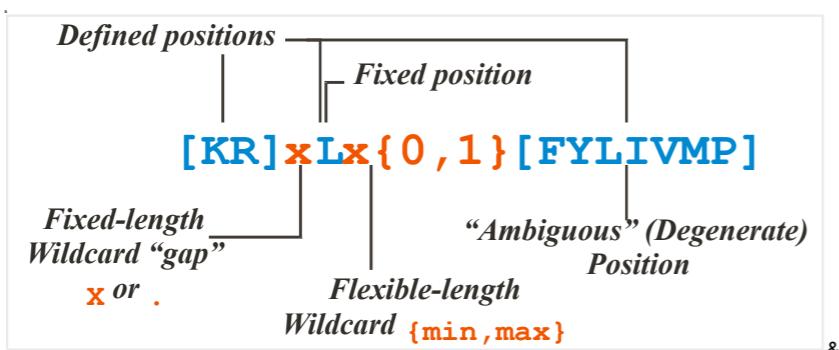
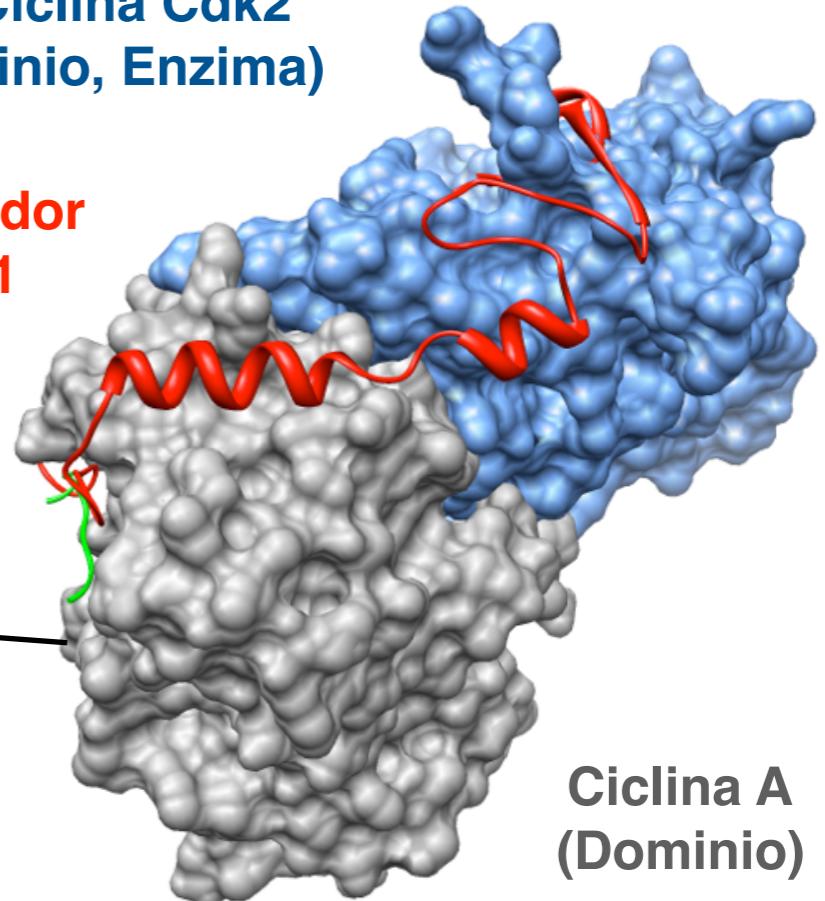


Figure 8.3. Anatomy of a SLiM.

## Ejemplo LOGO PSSMs

**Quinasa dependiente  
de Ciclina Cdk2  
(Dominio, Enzima)**

**Inhibidor  
p21**



**Ciclina A  
(Dominio)**

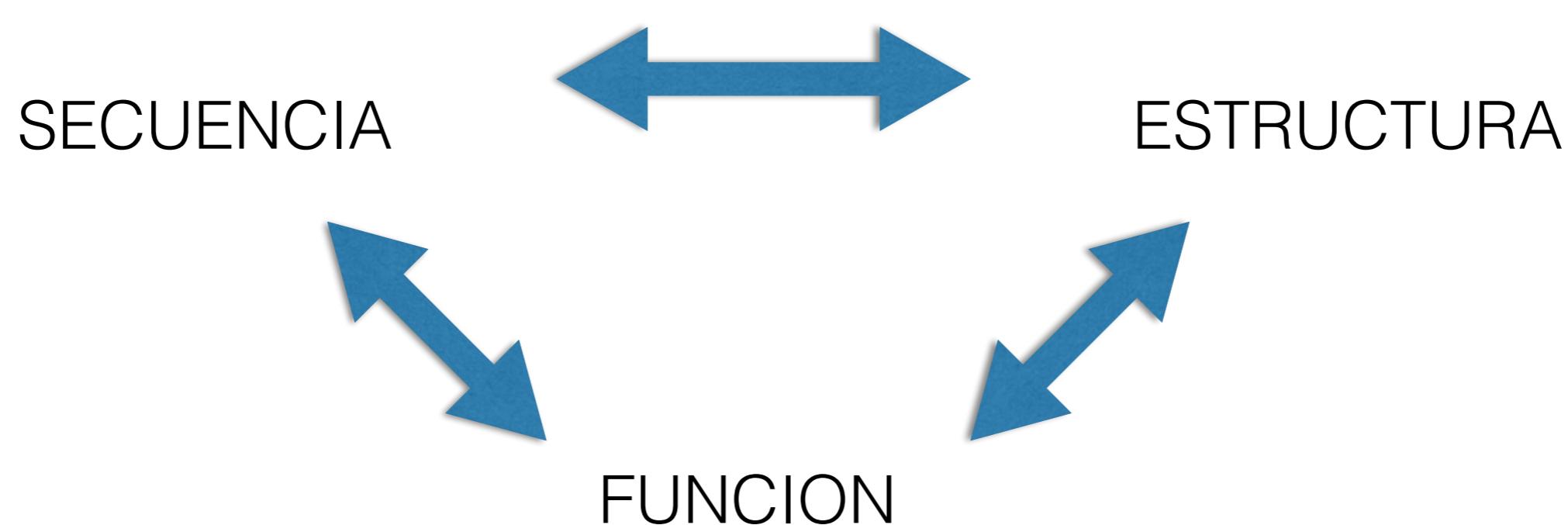
## Ejemplos...

- Señal “KDEL” de retención en RER
- Motivos de unión a Integrina “RGD”
- Motivos de unión a Ciclinas “RxL”

Hoy vamos a usar dos programas para investigar la secuencia y estructura de proteínas

***Jalview***  
***MSAs***

***Chimera***  
***Estructuras***



*apuntando a elucidar las relaciones  
secuencia-estructura-función de proteínas*