Máster Universitario en Bioinformática

Proteómica y Bioinformática Estructural

Curso académico 2024-2025

Dra. Magdalena Nikolaeva Koleva
magdalena.nikolaeva@professor.universidadviu.com

24/09/2024

De:
Planeta Formación y Universidades

Sesión 3

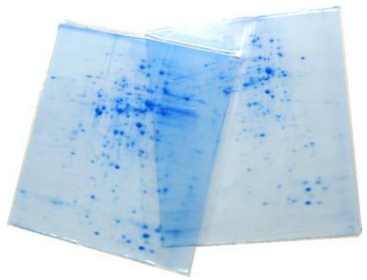
Proteómica estructural

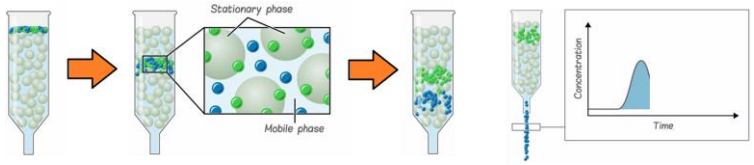
Universidad Internacional de Valencia

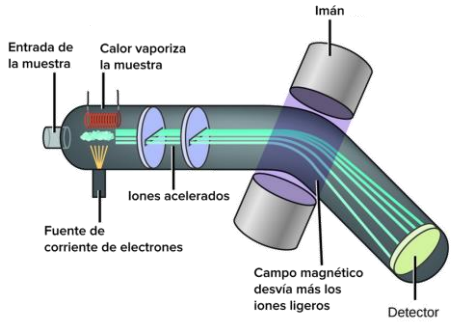
24/09/2024

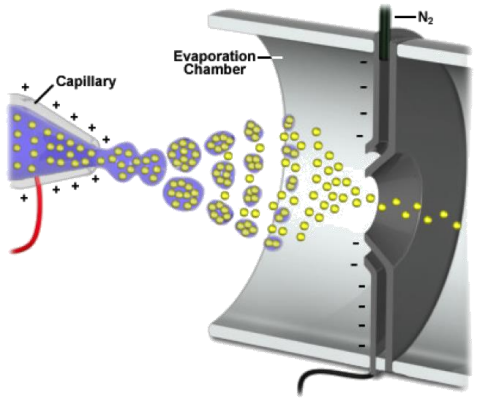
De:
Planeta Formación y Universidades

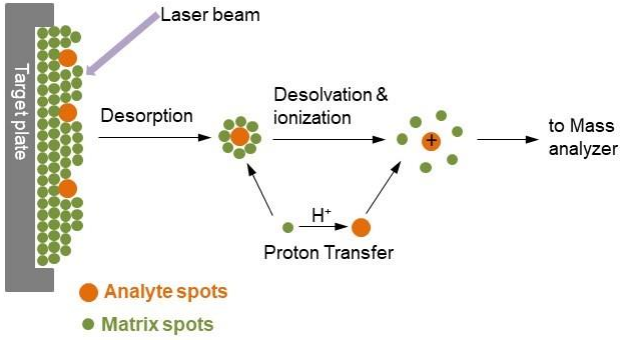
¿Qué vimos en la anterior sesión?

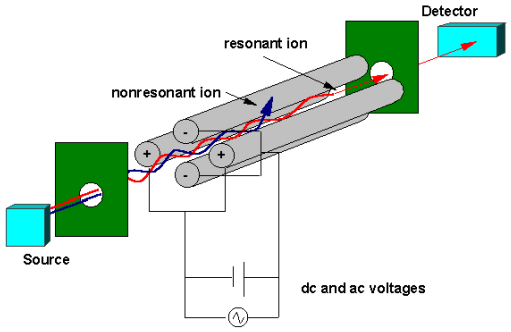


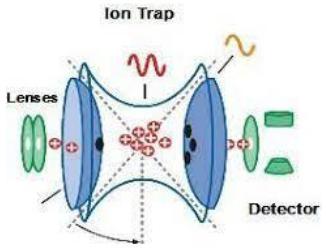


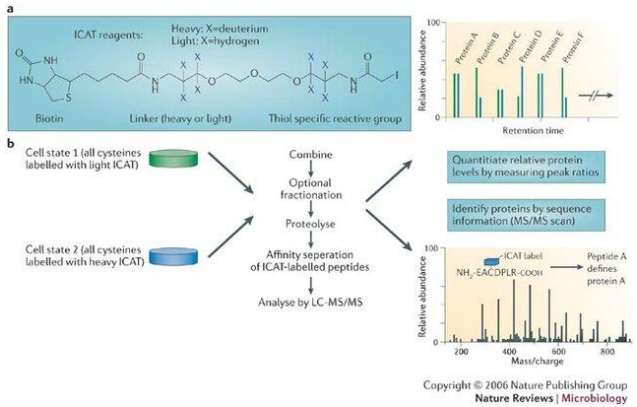












Temario - Contenidos

Tema 3. Proteómica estructural
3.1. Aspectos estructurales de las proteínas: niveles de organización
3.2. Métodos de caracterización estructural en proteómica
3.3. Servidores y bases de datos en proteómica estructural
3.4. Bioinformática estructural

Temario - Contenidos

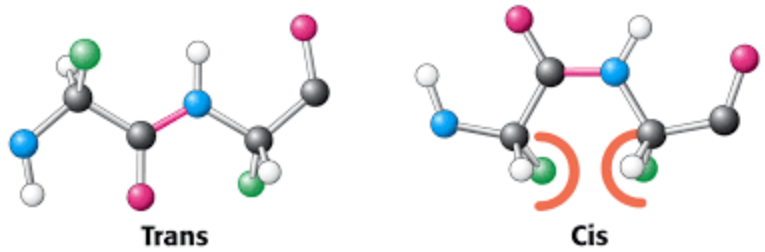
Tema 3. Proteómica estructural
3.1. Aspectos estructurales de las proteínas: niveles de organización
3.2. Métodos de caracterización estructural en proteómica
3.3. Servidores y bases de datos en proteómica estructural
3.4. Bioinformática estructural

3.1. Aspectos estructurales de las proteínas

Cadena polipeptídica:

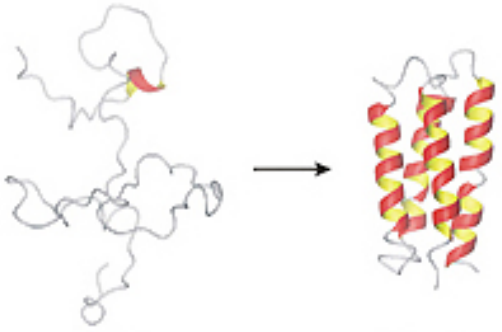
Presenta flexibilidad lo que le permite adoptar diferentes conformaciones, aunque con ciertas restricciones:

- Planaridad del enlace peptídico
- Choques estéricos de cadenas laterales



Plegamiento proteico: conformación global estable (3D) esencial para su función

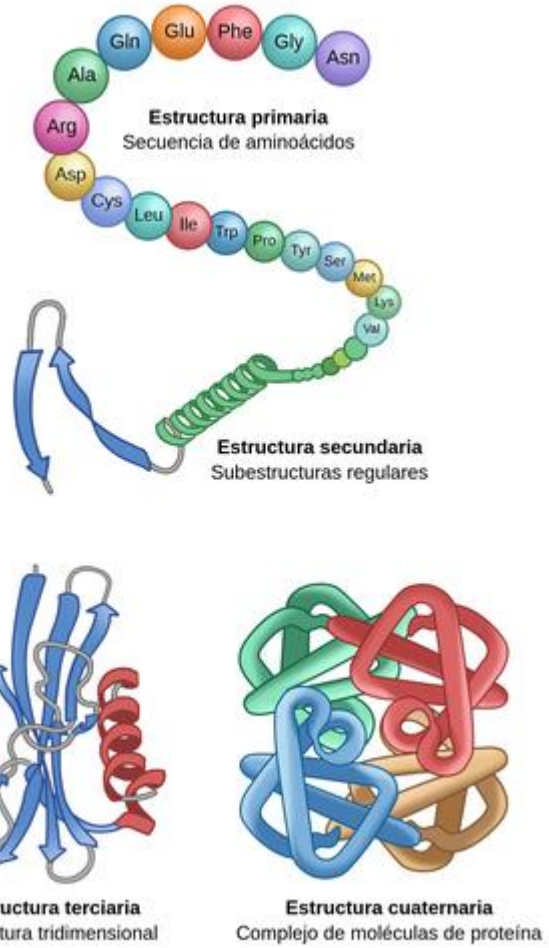
Desplegada:
Sin función



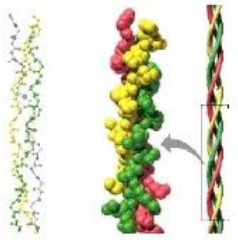
Plegada:
Con función

3.1. Aspectos estructurales de las proteínas

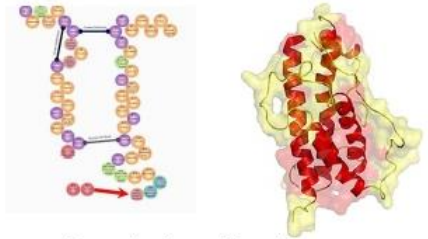
Estructura de las proteínas:



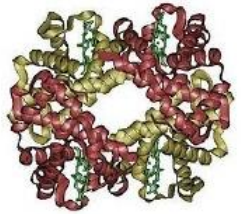
Plegamiento: estructura tridimensional única que determina la función específica de una proteína



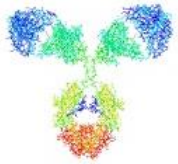
•Estructural (colágeno y queratina)



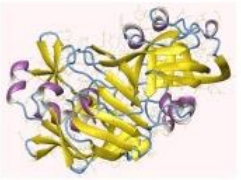
Reguladora (insulina y hormona del crecimiento)



•Transportadora (hemoglobina),



•Inmunológica (anticuerpos),

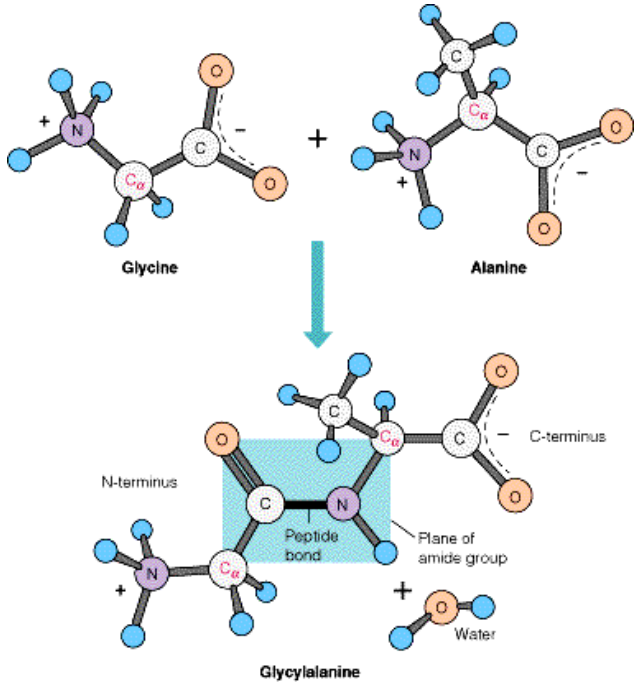


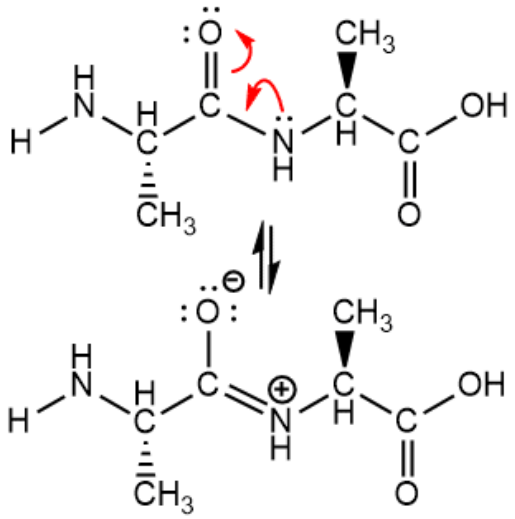
•Enzimática (pepsina),

3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura primaria:
  - Secuencia de amino ácidos
  - Determina el plegamiento proteico
  - Enlace peptídico – planaridad, deslocalización electrónica





3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura primaria:
  - Ángulos de torsión (phi – N-Cα, psi – Cα-C)
    - Rotación libre
    - Determinante de conformaciones
    - Importancia cadena lateral (R) – restricciones de estructura química
  - Ángulo enlace peptídico (omega – N-C)

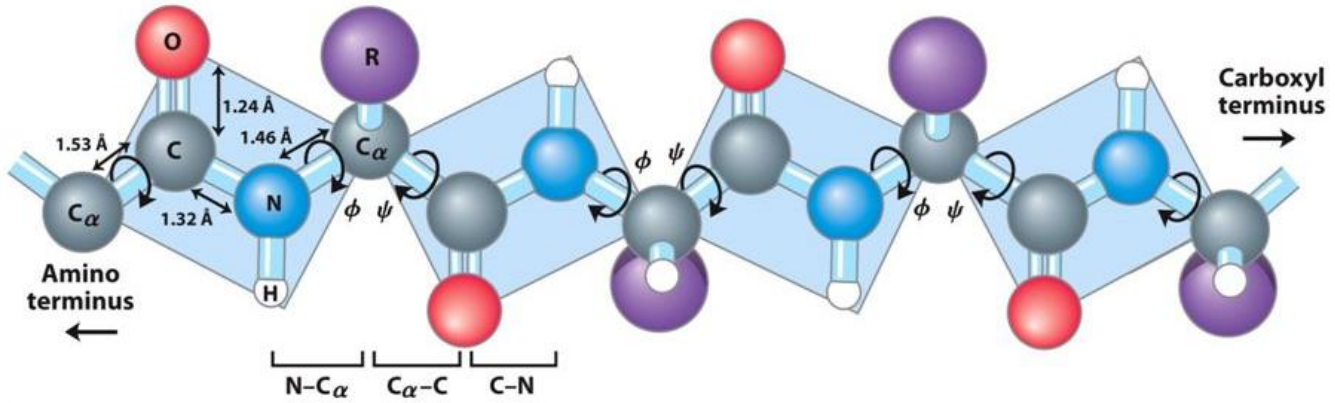
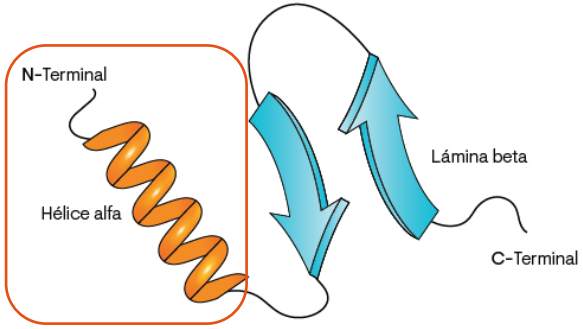


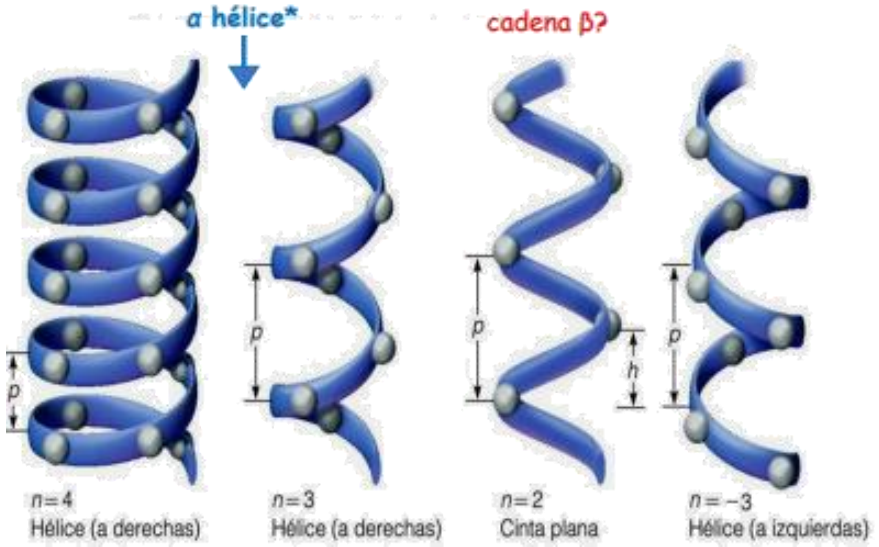
Figure 4-2b
Lehninger Principles of Biochemistry, Sixth Edition

3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura secundaria:
  - Hélices: p=paso de hélice, n=nº de residuos/vuelta
    - hélice π: n=4.4 residuos/vuelta, hueca por dentro, inestable
    - α-hélice: n=3.6 residuos/vuelta, p=3.6x0.15nm/res=0.56nm
    - 30% de todas las hélices
    - Los enlaces peptídicos se estabilizan por 2 puentes-H paralelos a la hélice, aumentando su estabilidad
    - Dependencia de pH
    - Inestabilidad en los extremos (solo 1 puente-H)

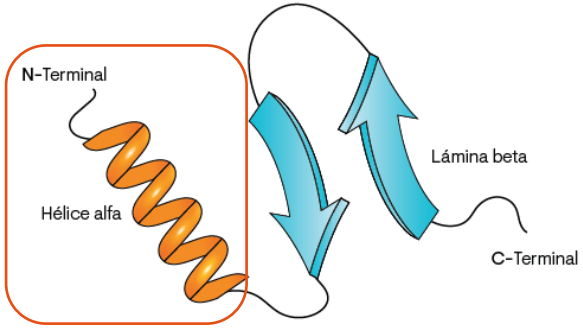


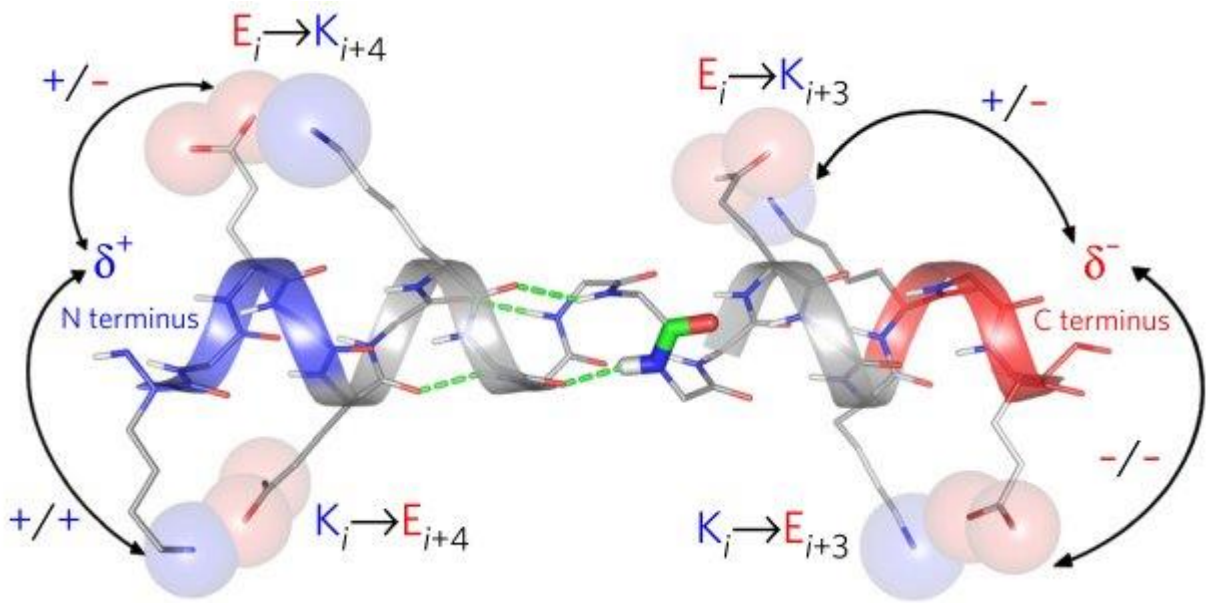


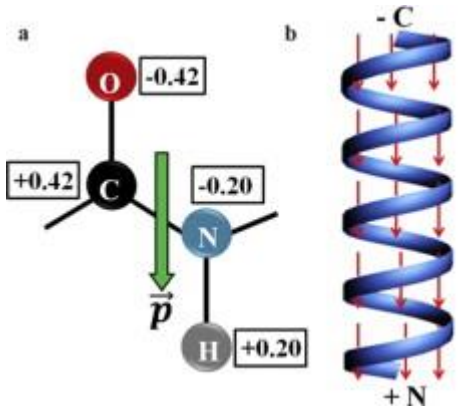
3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura secundaria:
  - Hélices: p=paso de hélice, n=nº de residuos/vuelta
    - α-hélice: n=3.6 residuos/vuelta, p=3.6x0.15nm/res=0.56nm
    - Momento dipolar asociado a la hélice



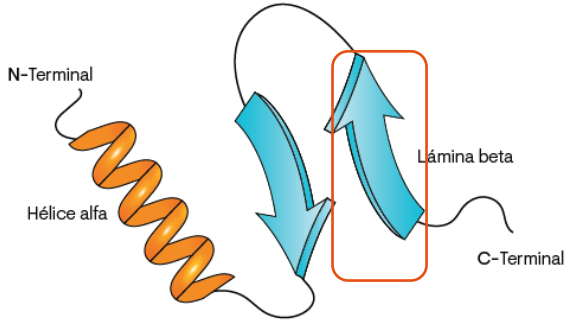


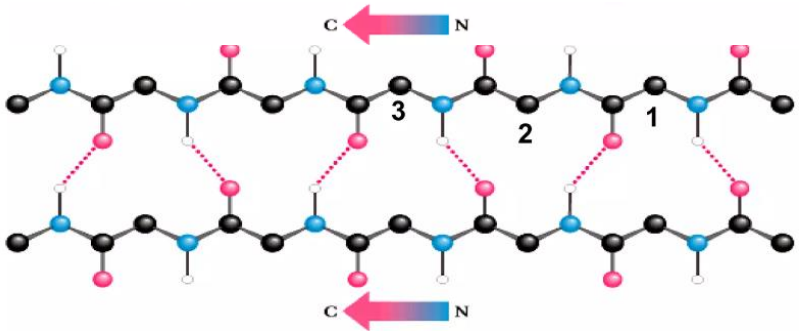


3.1. Aspectos estructurales de las proteínas

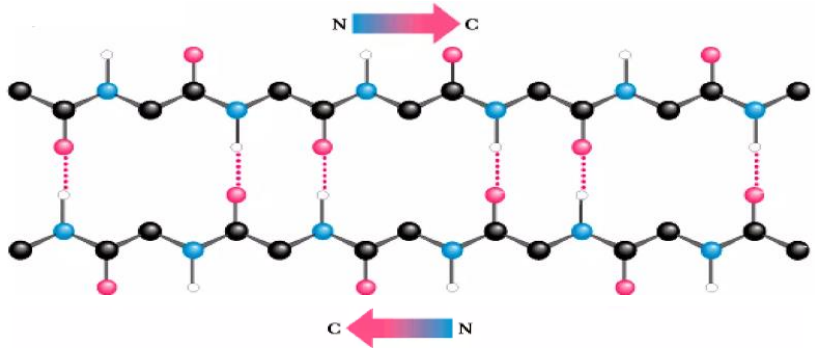
Estructura de las proteínas:

- Estructura secundaria:
  - Láminas-β: conformación extendida
    - Agrupación de 2-3 cadenas aprovechando puentes-H amino/carboxi
    - Paralelas, antiparalelas – diferencias puentes-H
    - Momentos dipolares distintos entre paralelas (suman) y antiparalelas (anulan).
    - Curvatura





Paralela

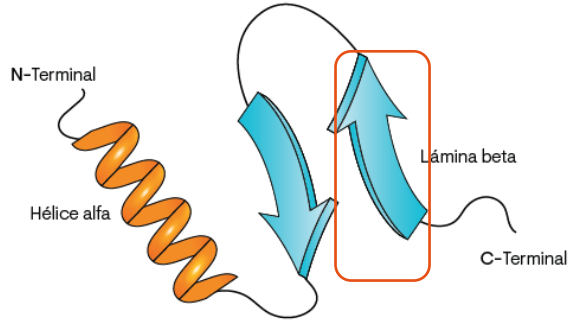


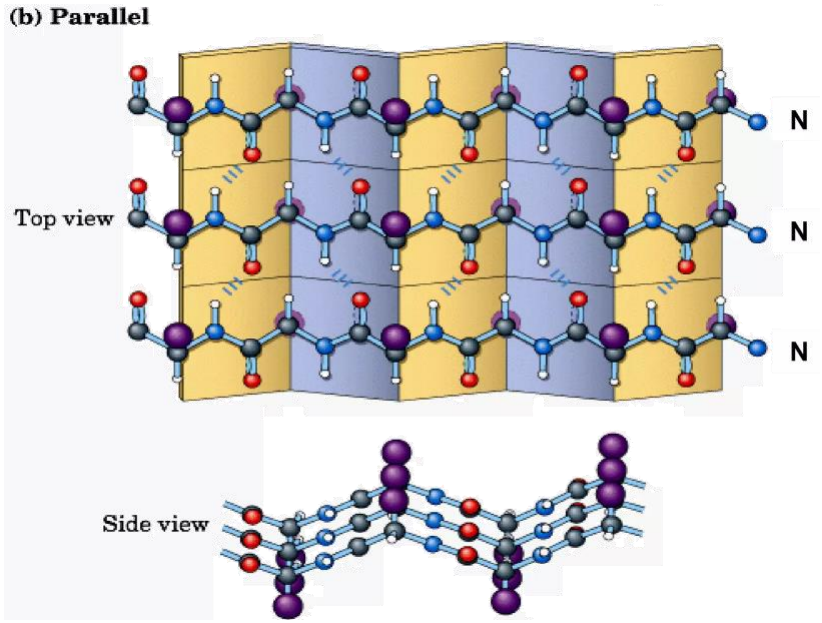
Antiparalela

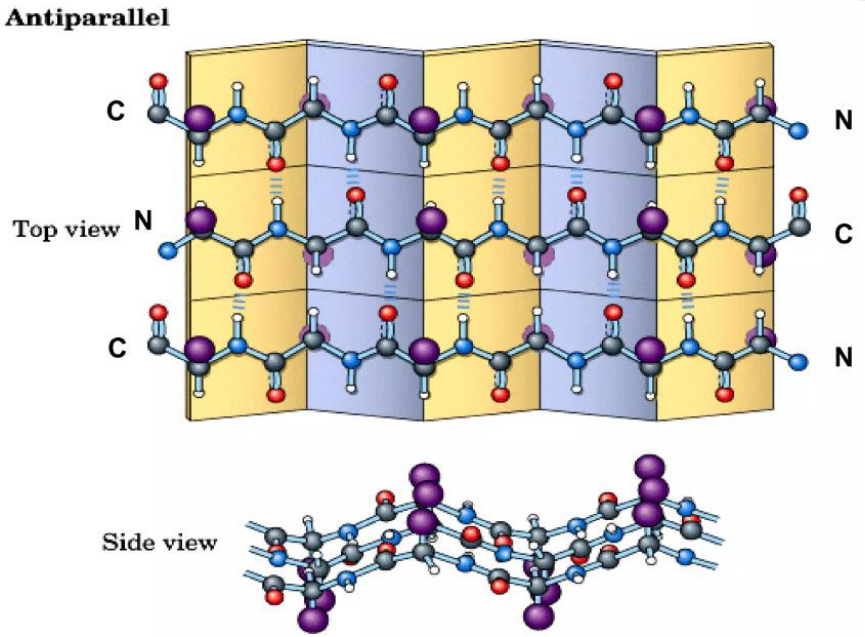
3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura secundaria:
  - Láminas-β - Hojas-β



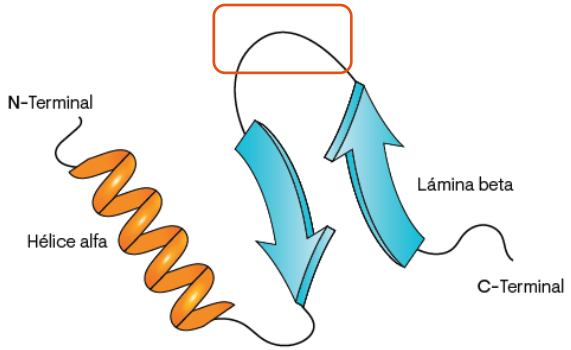


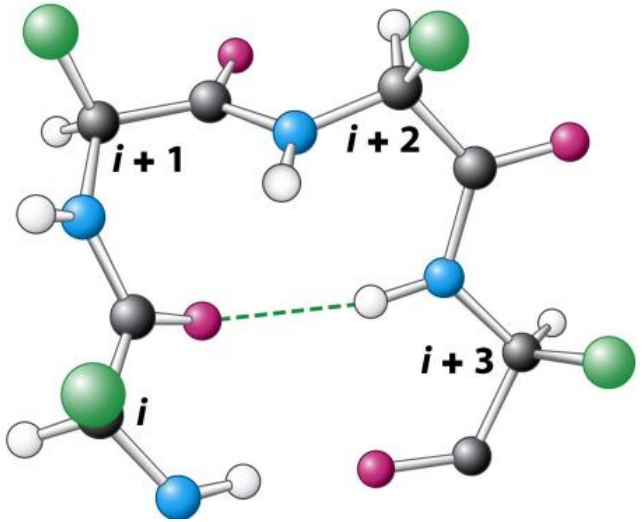


3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura secundaria:
  - Giros-β: elementos de conexión entre hélices y/o láminas.
    - Determinan cambios drásticos de dirección
    - Estabilizado por puente-H entre i e i+3.
    - Pro, Gly



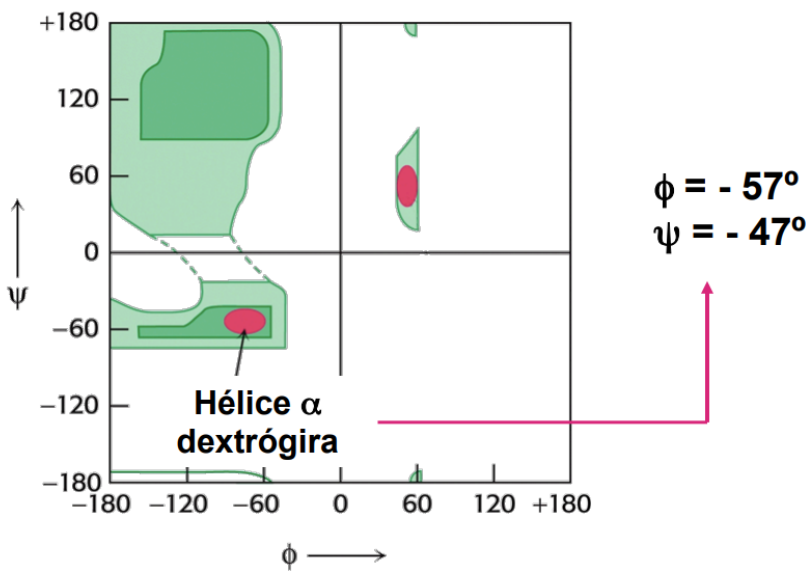


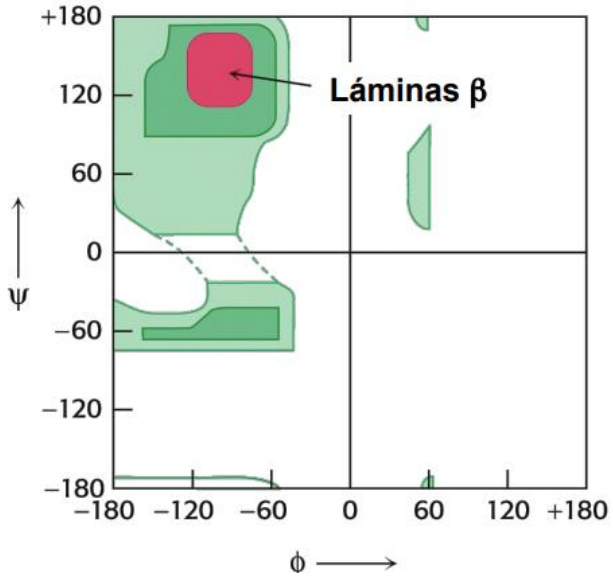
3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura secundaria:
  - Diagrama de Ramachandran:

Representación gráfica de los valores de ángulos de torsión de los residuos proteicos.
Identifica estructuras mal resueltas.





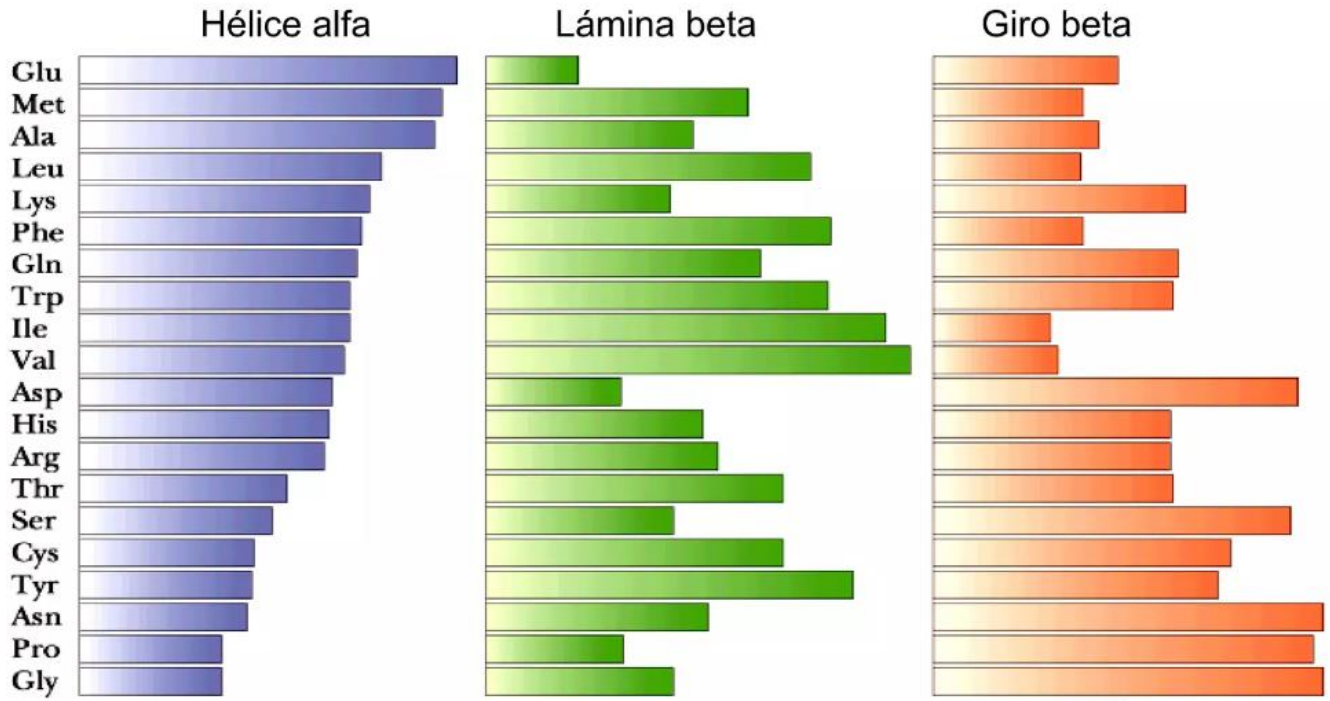
Paralela: ψ = +113º, ϕ = -119º

Antiparalela: ψ = +135º, ϕ = -139º

3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

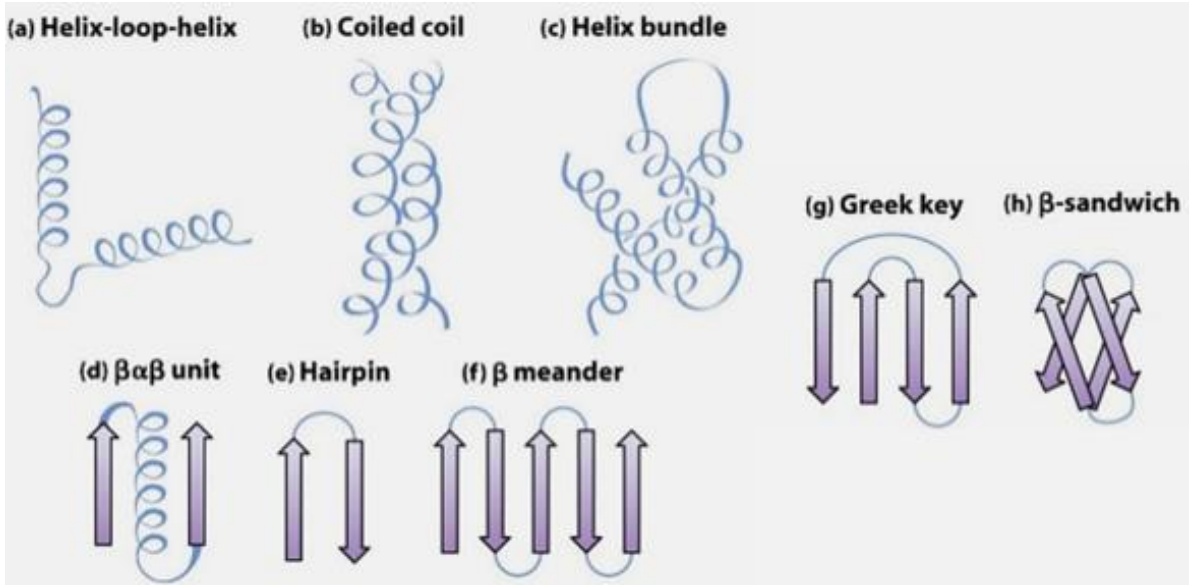
- Estructura secundaria:
  - Frecuencia relativa de aparición de aminoácidos en estructuras secundarias:



3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

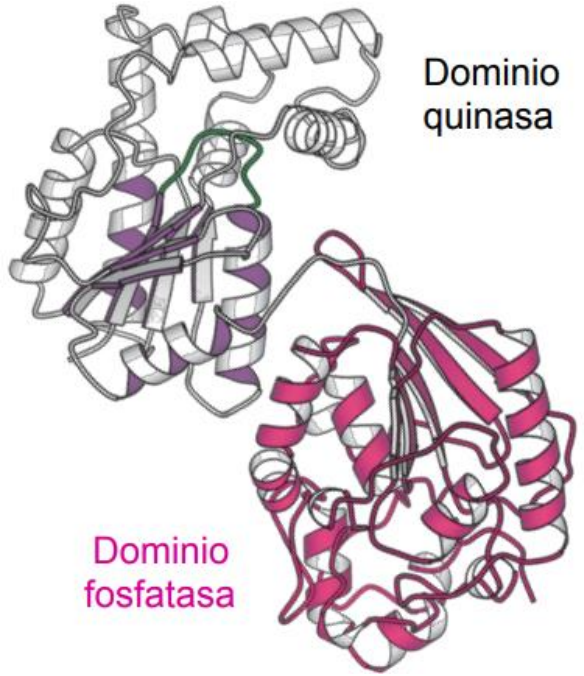
- Estructura terciaria:
  - Supersecundaria – motivos o unidades de plegamiento: ordenación de los elementos de estructura secundaria



3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

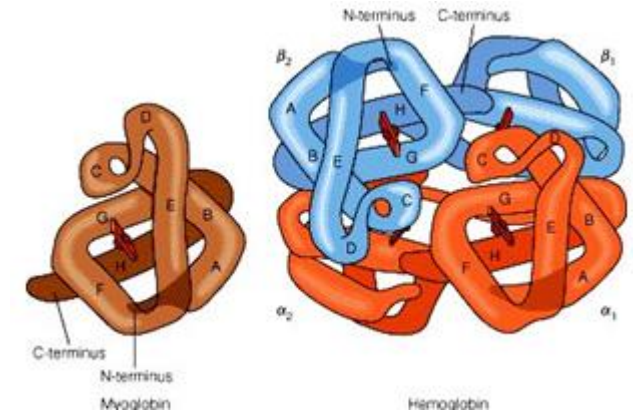
- Estructura terciaria:
  - Dominios: partes de la cadena polipeptídica estables de manera independiente, determinante de aspectos funcionales. Secuencia – estructura – función
    - Efecto hidrofóbico determinante en el plegamiento.
    - Fuerzas de Van der Waals en el núcleo.
    - Puentes de hidrógeno entre cadenas.
    - Pares iónicos entre residuos cargados de la superficie.
    - Puentes disulfuro que estabilizan la estructura 3D.

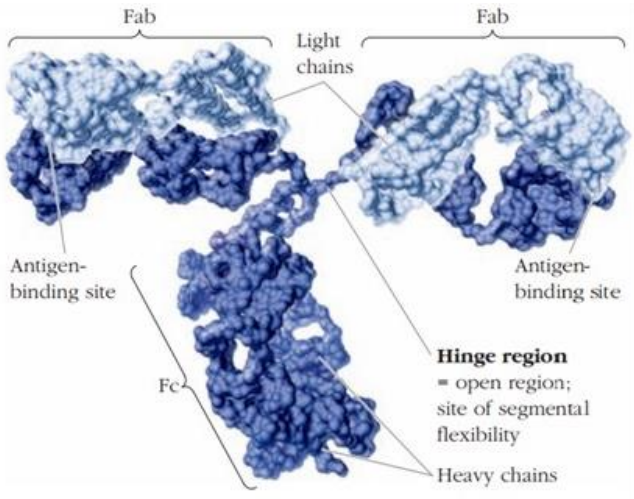


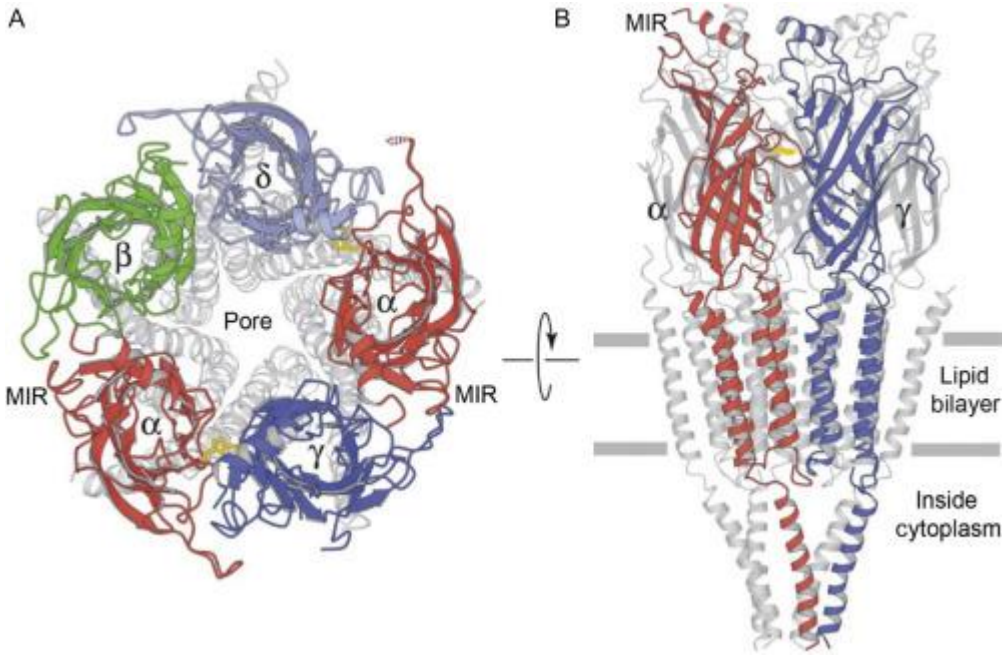
3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Estructura cuaternaria: disposición de varias cadenas polipeptídicas
  - Dominio único vs multidominio (homo/hetero), proteínas de membrana



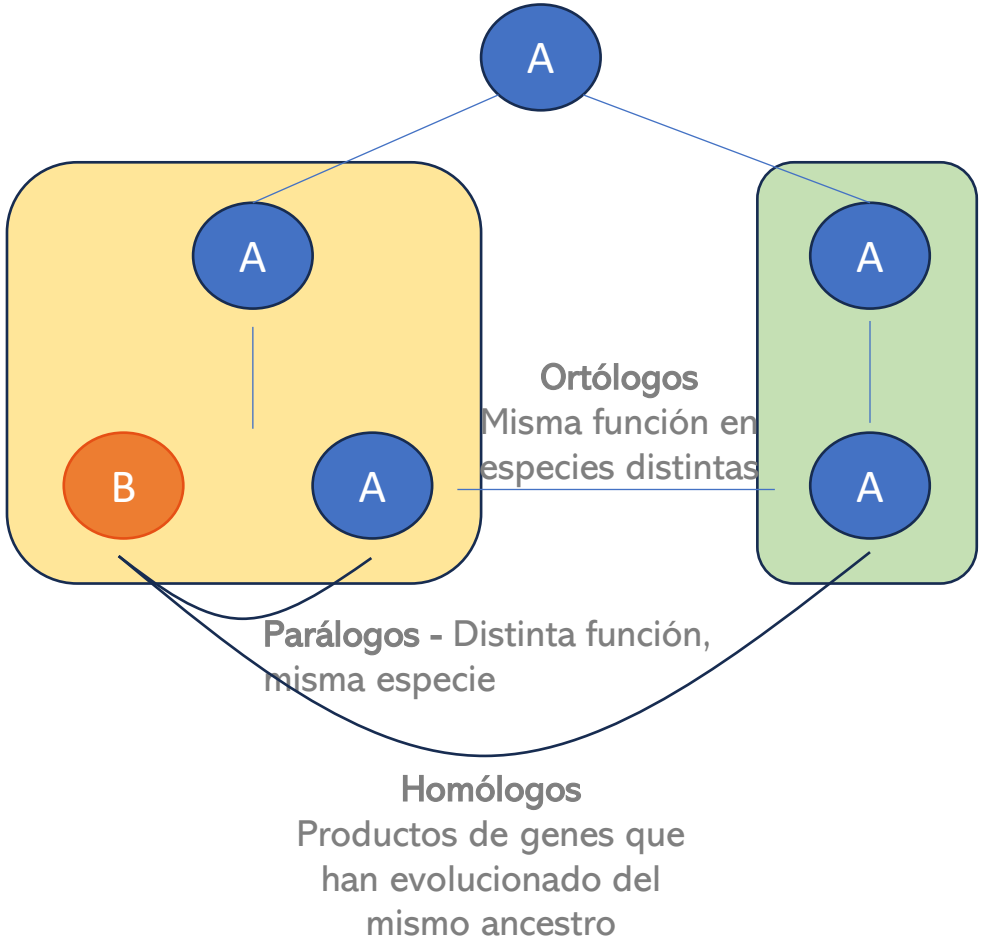




3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

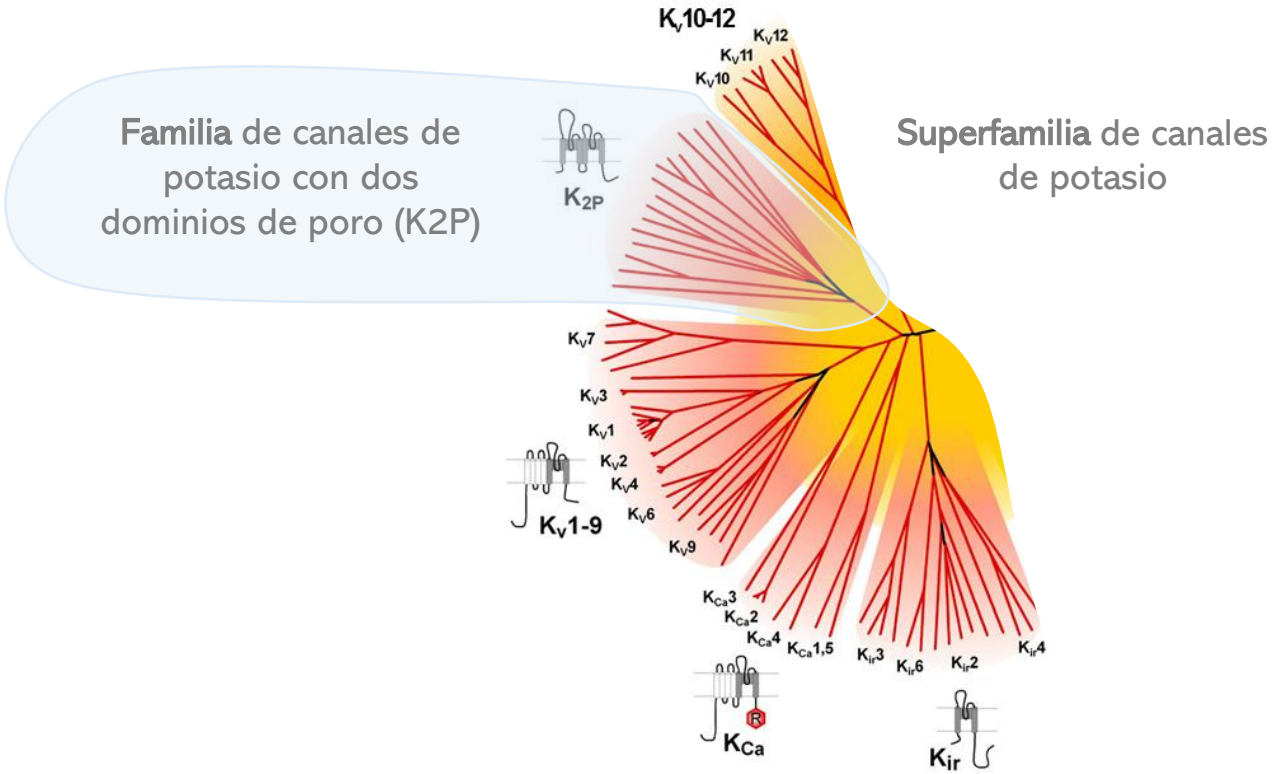
- Homología de proteínas: Homólogos - Ortólogos/Parálogos



3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

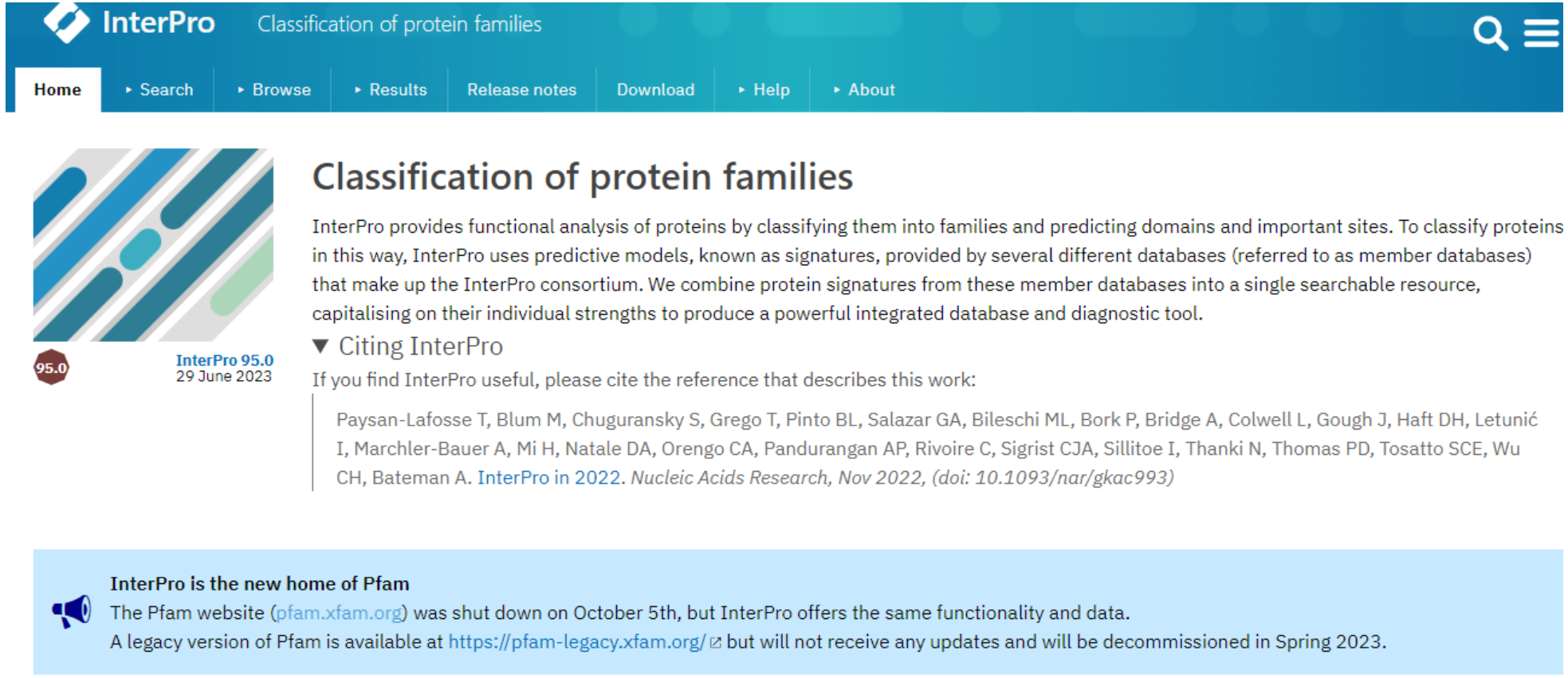
- Familias proteicas: son proteínas relacionadas por homología con una identidad de secuencia >30%, estructura y función similares
- Superfamilias proteicas: proteínas o dominios que tienen baja identidad de secuencia, pero presentan aspectos estructurales y funcionales similares.



3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

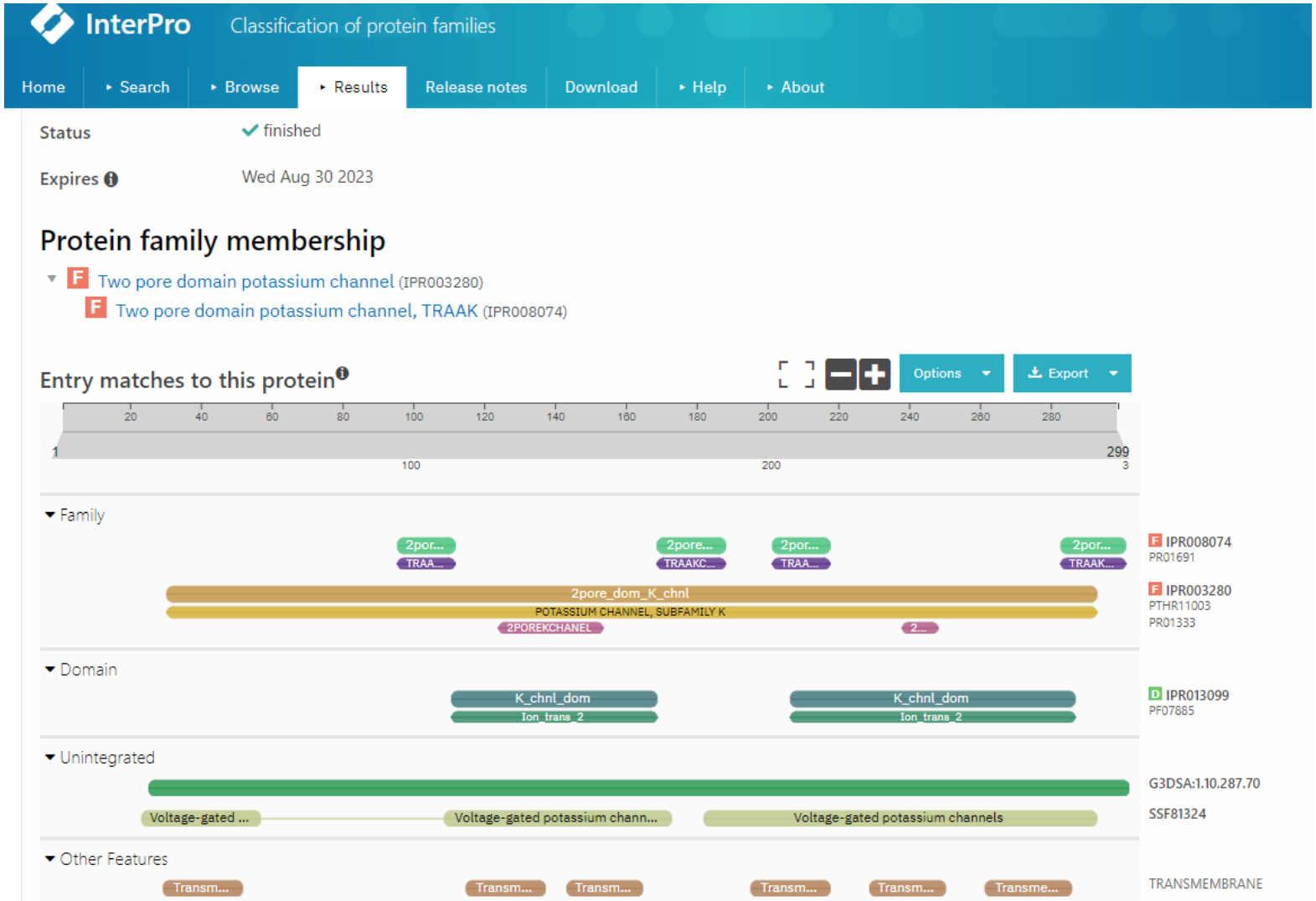
- Recursos informáticos: Pfam → InterPro (www.ebi.ac.uk/interpro)



3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Recursos informáticos: Pfam → InterPro



3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Recursos informáticos: PROSITE (prosite.expasy.org)

Database of protein domains, families and functional sites

New SARS-CoV-2 relevant PROSITE motifs

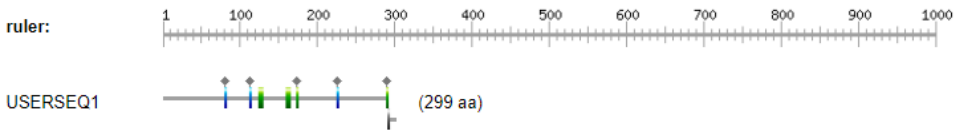
PROSITE consists of documentation entries describing protein domains, families and functional sites as well as associated patterns and profiles to identify them [More... / References / Commercial users].
PROSITE is complemented by ProRule , a collection of rules based on profiles and patterns, which increases the discriminatory power of profiles and patterns by providing additional information about functionally and/or structurally critical amino acids [More...].

Legend:

disulfide bridge | active site | other 'ranges' | other sites

Please note that the graphical representations of domains displayed hereafter are for illustrative purposes only, and that their colors and shapes are not intended to indicate homology or shared function.
For more information about how these graphical representations are constructed, go to https://prosite.expasy.org/mydomains/.

hits by patterns with a high probability of occurrence or by user-defined patterns: [8 hits (by 4 distinct patterns) on 1 sequence]



PS00006 **CK2_PHOSPHO_SITE** *Casein kinase II phosphorylation site* :

80 - 83: SdqE
Predicted feature:
MOD_RES 80 Phosphoserine [condition: S]

112 - 115: SawD
Predicted feature:
MOD_RES 112 Phosphoserine [condition: S]

224 - 227: SklE
Predicted feature:
MOD_RES 224 Phosphoserine [condition: S]

PS00008 **MYRISTYL** *N-myristoylation site* :

124 - 129: GTiiTT

158 - 163: GIllAG

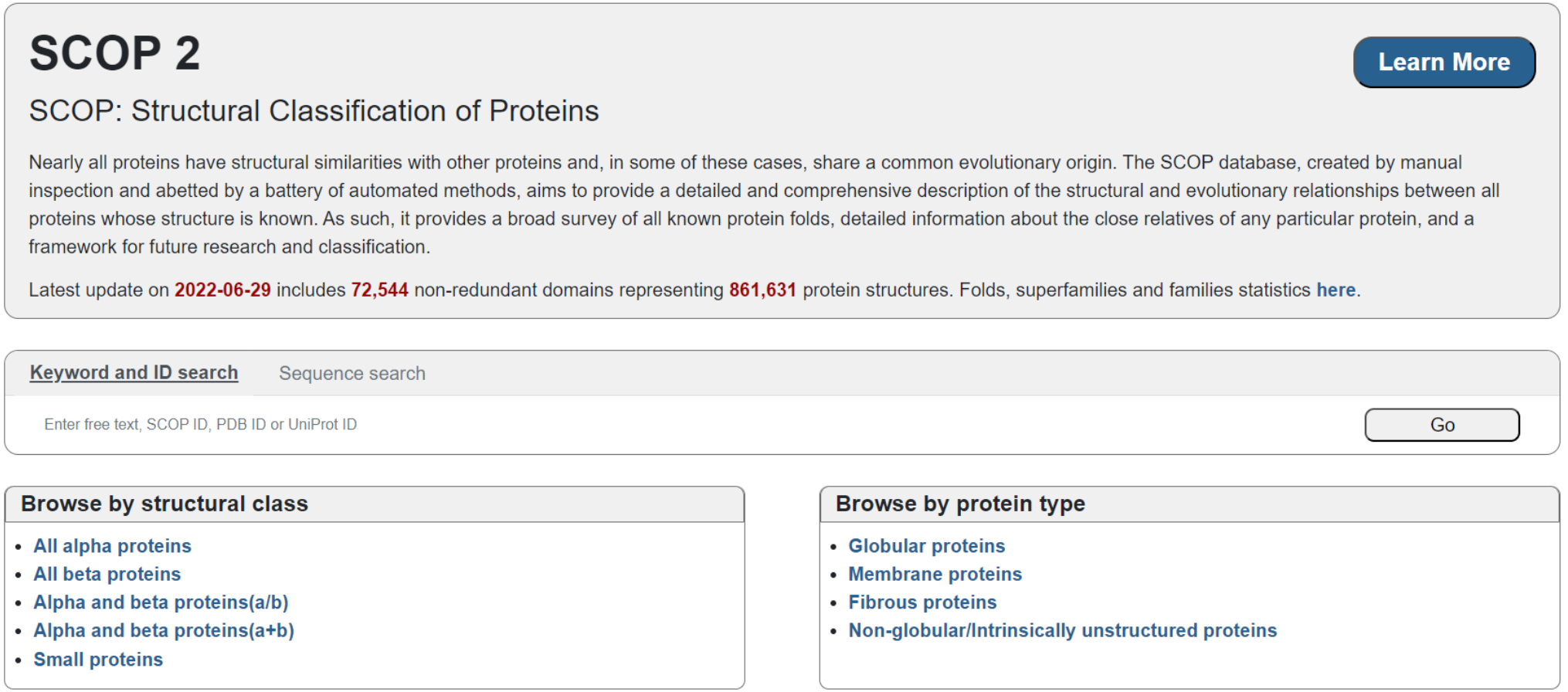
PS00005 **PKC_PHOSPHO_SITE** *Protein kinase C phosphorylation site* :

171 - 173: SlR
Predicted feature:
MOD_RES 171 Phosphoserine [condition: S]

3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

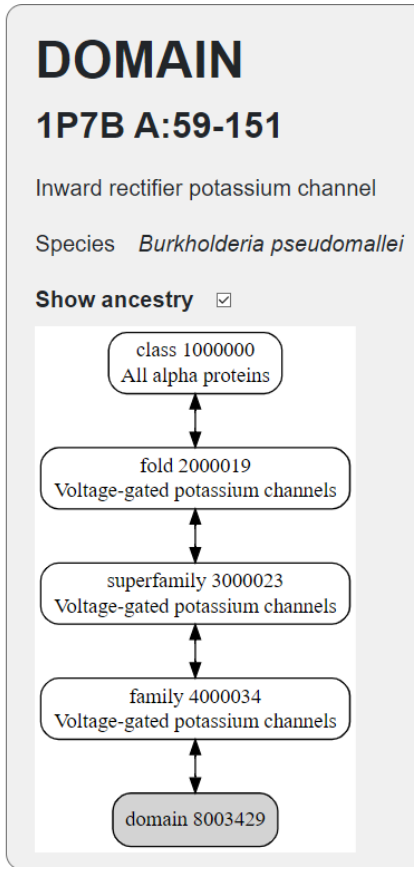
- Recursos informáticos: SCOP (https://scop.mrc-lmb.cam.ac.uk)

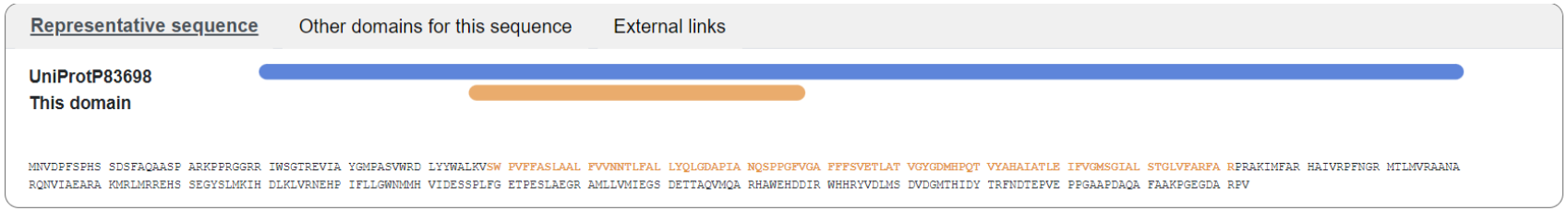


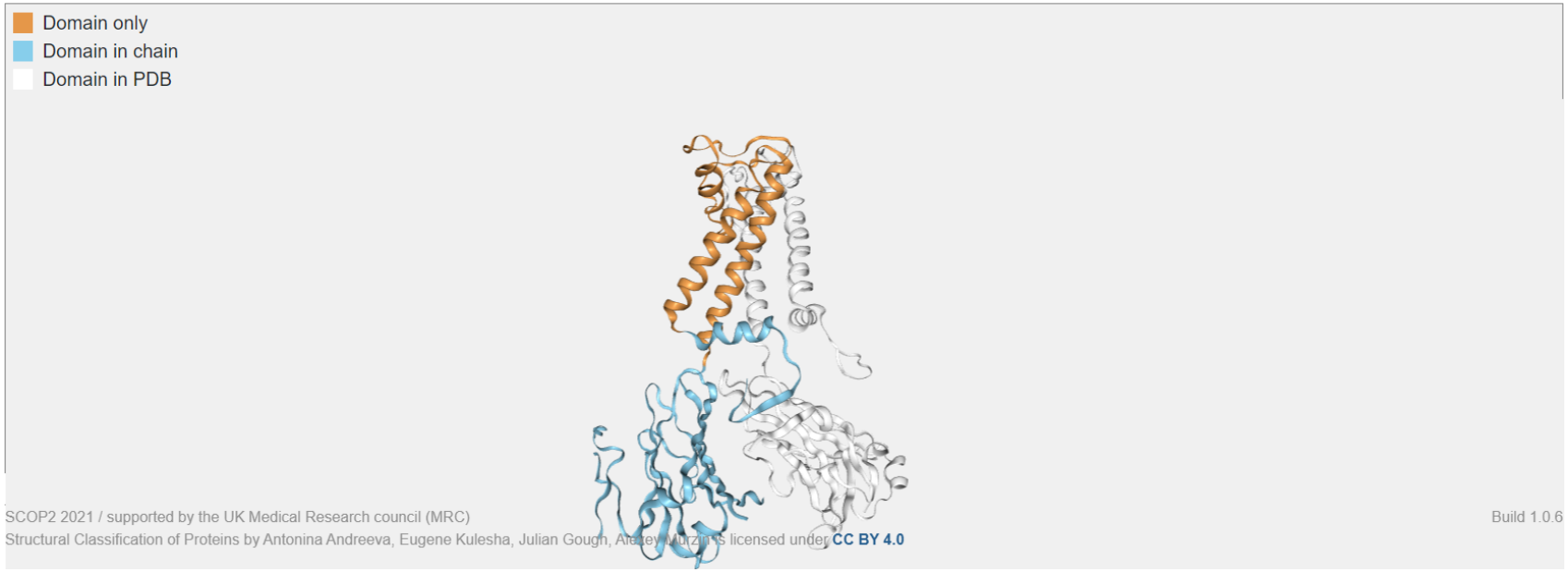
3.1. Aspectos estructurales de las proteínas

Estructura de las proteínas:

- Recursos informáticos: SCOP (https://scop.mrc-lmb.cam.ac.uk)







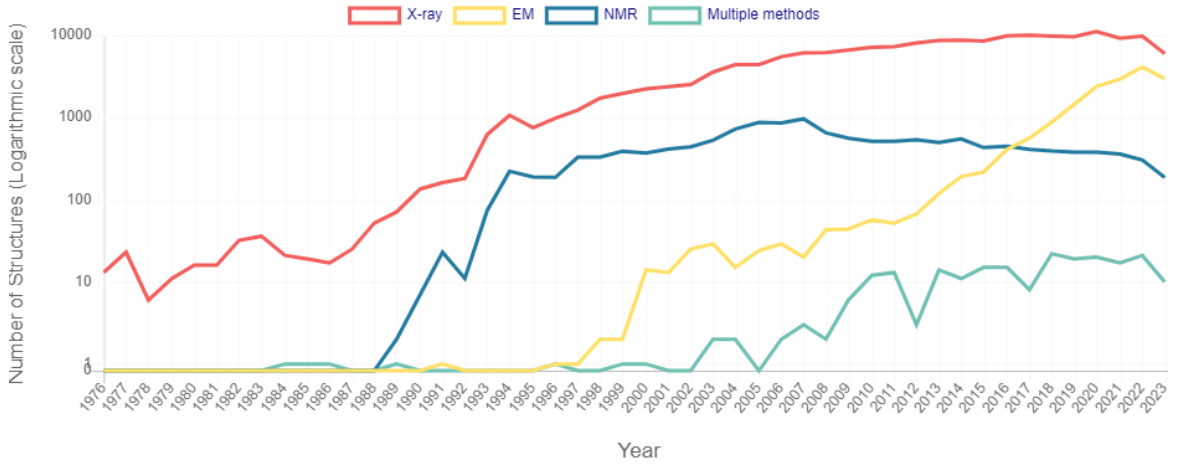
Temario - Contenidos

Tema 3. Proteómica estructural
3.1. Aspectos estructurales de las proteínas: niveles de organización
3.2. Métodos de caracterización estructural en proteómica
3.3. Servidores y bases de datos en proteómica estructural
3.4. Bioinformática estructural

3.2. Métodos de caracterización estructural

Determinación estructural experimental:

- Cristalografía de rayos X (X-ray)
  - Difracción de rayos X por solidos cristalinos
  - Tipo y posición de átomos
  - Sin límite de tamaño
  - Mejores resoluciones
  - Requisito: cristalizar, cristales de calidad



- Resonancia magnética nuclear (NMR)
  - Un campo magnético externo hace que ciertos núcleos absorban radiofrecuencias selectivas. La energía absorbida induce una transición en el espín nuclear, observándose en el espectro RMN.
  - Ensamblados conformacionales
  - Limitaciones de tamaño (depende de potencia del campo magnético)
- Crio-electromicroscopía (Cryo-EM)
  - Estudio a temperatura criogénica evitando artefactos
  - Tomografía crioelectrónica para reconstruir en 3D imágenes 2D anguladas.
  - Resolución similar X-ray
  - Número de estructuras resueltas aumentando

3.2. Métodos de caracterización estructural

Determinación estructural experimental:

- Archivos PDB (ProteinDataBank)

```
HEADER    TRANSPORT PROTEIN                       22-JAN-19   6NR4
TITLE     CRYO-EM STRUCTURE OF THE TRPM8 ION CHANNEL WITH LOW OCCUPANCY ICILIN,
TITLE    2 PI(4,5)P2, AND CALCIUM
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: TRANSIENT RECEPTOR POTENTIAL CATION CHANNEL SUBFAMILY M
COMPND   3 MEMBER 8;
COMPND   4 CHAIN: A, B, C, D;
COMPND   5 ENGINEERED: YES;
COMPND   6 MUTATION: YES
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: FICEDULA ALBICOLLIS;
SOURCE   3 ORGANISM_COMMON: COLLARED FLYCATCHER;
SOURCE   4 ORGANISM_TAXID: 59894;
SOURCE   5 GENE: TRPM8;
SOURCE   6 EXPRESSION_SYSTEM: HOMO SAPIENS;
SOURCE   7 EXPRESSION_SYSTEM_TAXID: 9606
KEYWDS    ION CHANNEL, TRP CHANNEL, TRPM CHANNEL, TRPM8 CHANNEL, COLD SENSING,
KEYWDS   2 LIPID SENSING, MENTHOL, ICILIN, WS-12, PI(4, 5)P2, CALCIUM-PERMEABLE
KEYWDS   3 CHANNEL, COOLING AGENT, TRANSPORT PROTEIN
EXPDTA    ELECTRON MICROSCOPY
AUTHOR    Y.YIN,S.C.LE,A.L.HSU,M.J.BORGNIA,H.YANG,S.-Y.LEE
REVDAT   4   18-DEC-19 6NR4    1       SCALE
REVDAT   3   13-MAR-19 6NR4    1       JRNL
REVDAT   2   27-FEB-19 6NR4    1       KEYWDS
REVDAT   1   20-FEB-19 6NR4    0
JRNL        AUTH   Y.YIN,S.C.LE,A.L.HSU,M.J.BORGNIA,H.YANG,S.Y.LEE
JRNL        TITL   STRUCTURAL BASIS OF COOLING AGENT AND LIPID SENSING BY THE
JRNL        TITL 2 COLD-ACTIVATED TRPM8 CHANNEL.
JRNL        REF    SCIENCE                       V. 363       2019
JRNL        REFN                   ESSN 1095-9203
JRNL        PMID   30733385
JRNL        DOI    10.1126/SCIENCE.AAV9334
```

```
REMARK 247 ELECTRON MICROSCOPY
REMARK 247  THE COORDINATES IN THIS ENTRY WERE GENERATED FROM ELECTRON
REMARK 247  MICROSCOPY DATA. PROTEIN DATA BANK CONVENTIONS REQUIRE
REMARK 247  THAT CRYST1 AND SCALE RECORDS BE INCLUDED, BUT THE VALUES
REMARK 247  ON THESE RECORDS ARE MEANINGLESS EXCEPT FOR THE CALCULATION
REMARK 247  OF THE STRUCTURE FACTORS.
REMARK 300
REMARK 300 BIOMOLECULE: 1
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY. THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 350
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW.  BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE: 1
REMARK 350 AUTHOR DETERMINED BIOLOGICAL UNIT: TETRAMERIC
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B, C, D
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
```

```
ATOM   5129  ND2 ASN A 972     153.460 134.596 175.301  1.00 44.88           N
ATOM   5130  N   LEU A 973     148.369 134.918 173.838  1.00 48.30           N
ATOM   5131  CA  LEU A 973     146.951 134.575 173.773  1.00 48.30           C
ATOM   5132  C   LEU A 973     146.317 135.117 172.500  1.00 48.30           C
ATOM   5133  O   LEU A 973     145.555 134.418 171.822  1.00 48.30           O
ATOM   5134  CB  LEU A 973     146.231 135.124 175.011  1.00 48.30           C
ATOM   5135  N   LEU A 974     146.641 136.359 172.146  1.00 49.12           N
ATOM   5136  CA  LEU A 974     145.985 136.950 170.990  1.00 49.12           C
ATOM   5137  C   LEU A 974     146.613 136.486 169.683  1.00 49.12           C
ATOM   5138  O   LEU A 974     145.979 136.584 168.627  1.00 49.12           O
ATOM   5139  CB  LEU A 974     145.997 138.465 171.121  1.00 49.12           C
```

3.2. Métodos de caracterización estructural

Determinación estructural a escala proteómica: mejora en tecnología y automatización

- PSI – Protein Structure Initiative – hasta año 2015
- SGC – Structural Genomics Consortium – 1500 proteínas resueltas

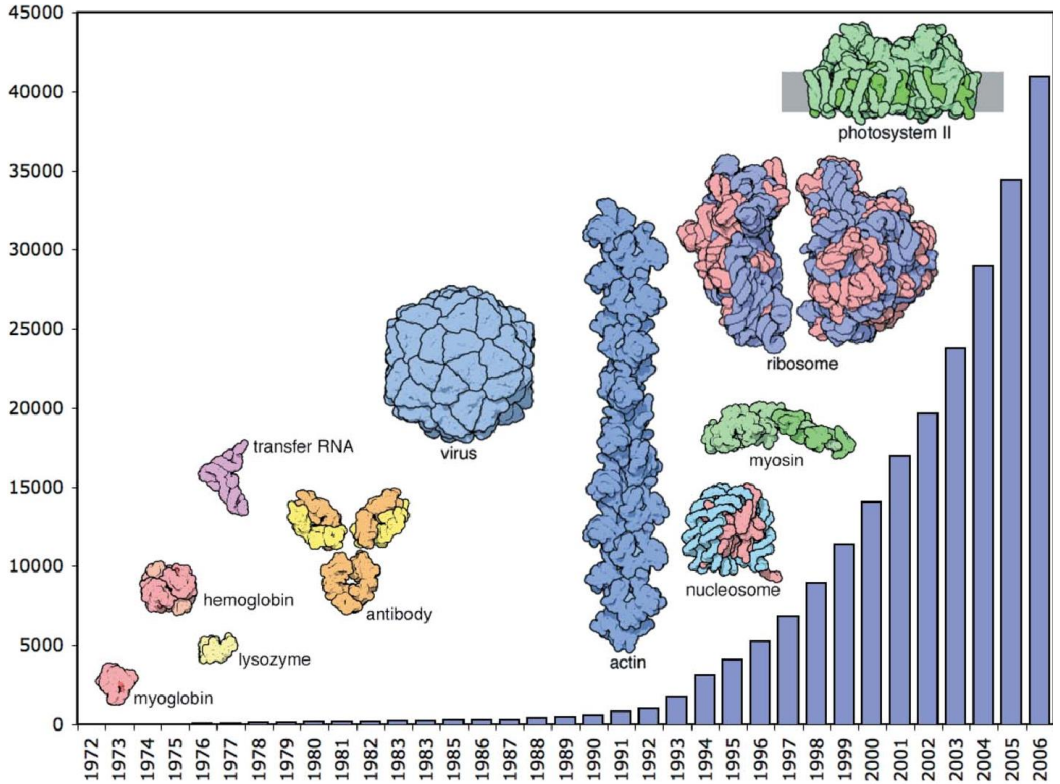


Figure 1

The Structural Genomics Consortium (SGC) is a global public-private partnership dedicated to open science.

2003
Established as a charity in the UK and co-founded by global pharmaceutical organizations, government agencies and international foundations

$400M+
In R&D funding to support early-stage projects since inception

8 Industry Partners
Takeda, janssen, Bayer, Boehringer Ingelheim, Pfizer, MERCK, Bristol Myers Squibb, Genentech

6 Research Sites
University of Toronto, The University of North Carolina at Chapel Hill, Karolinska Institutet, McGill University, Goethe Universität Frankfurt am Main, UCL

Major
Funding Supporters Include
wellcome, imi innovative medicines initiative, NIH, Bill & Melinda Gates foundation, CIHR IRSC, Genome Canada, Mitacs

330+
Partnerships

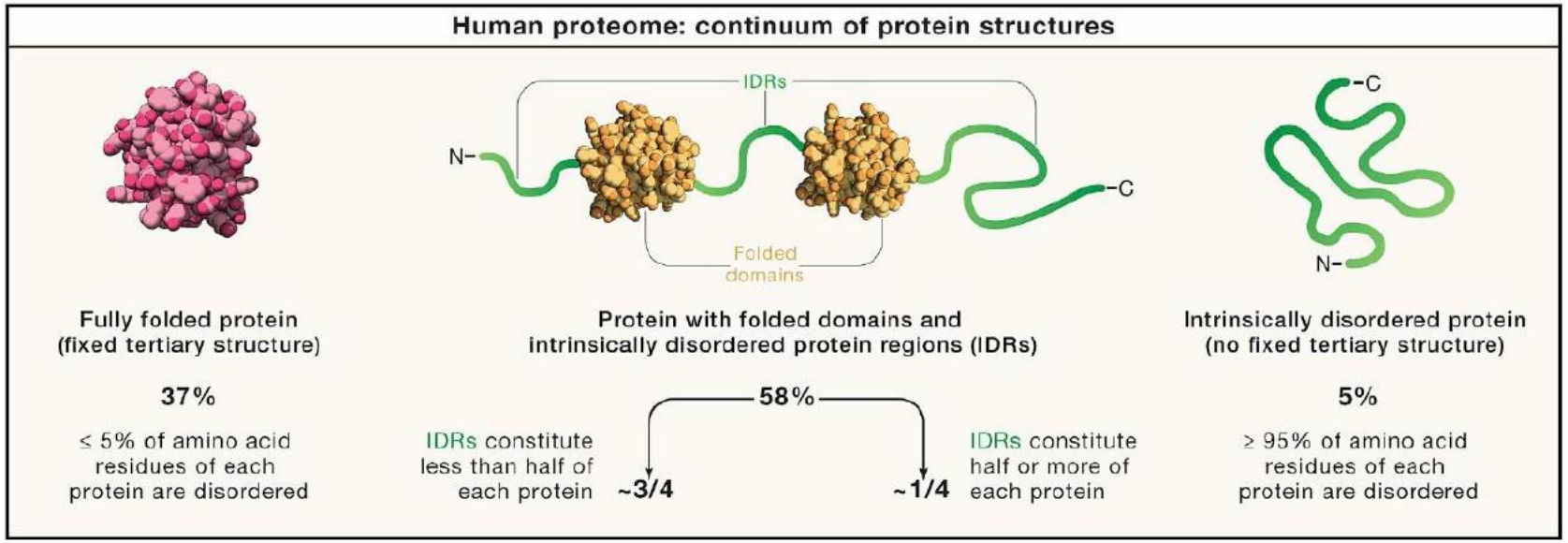
25+
Publications with industry partners each year

3.2. Métodos de caracterización estructural

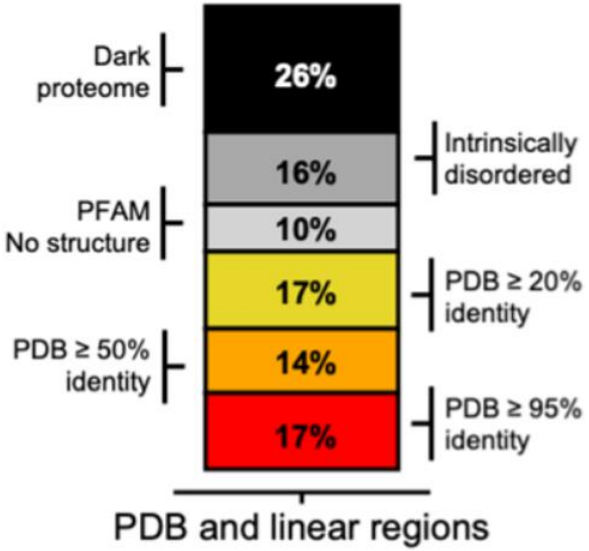
El proteoma estructural humano:

Se estima que existen unas **20.000 proteínas humanas** (no redundantes)

En los últimos **10 años**, los proyectos de proteómica han identificado unas **18.000 proteínas** (!!)



Tsang et al. 2020 Cell 183, 1742

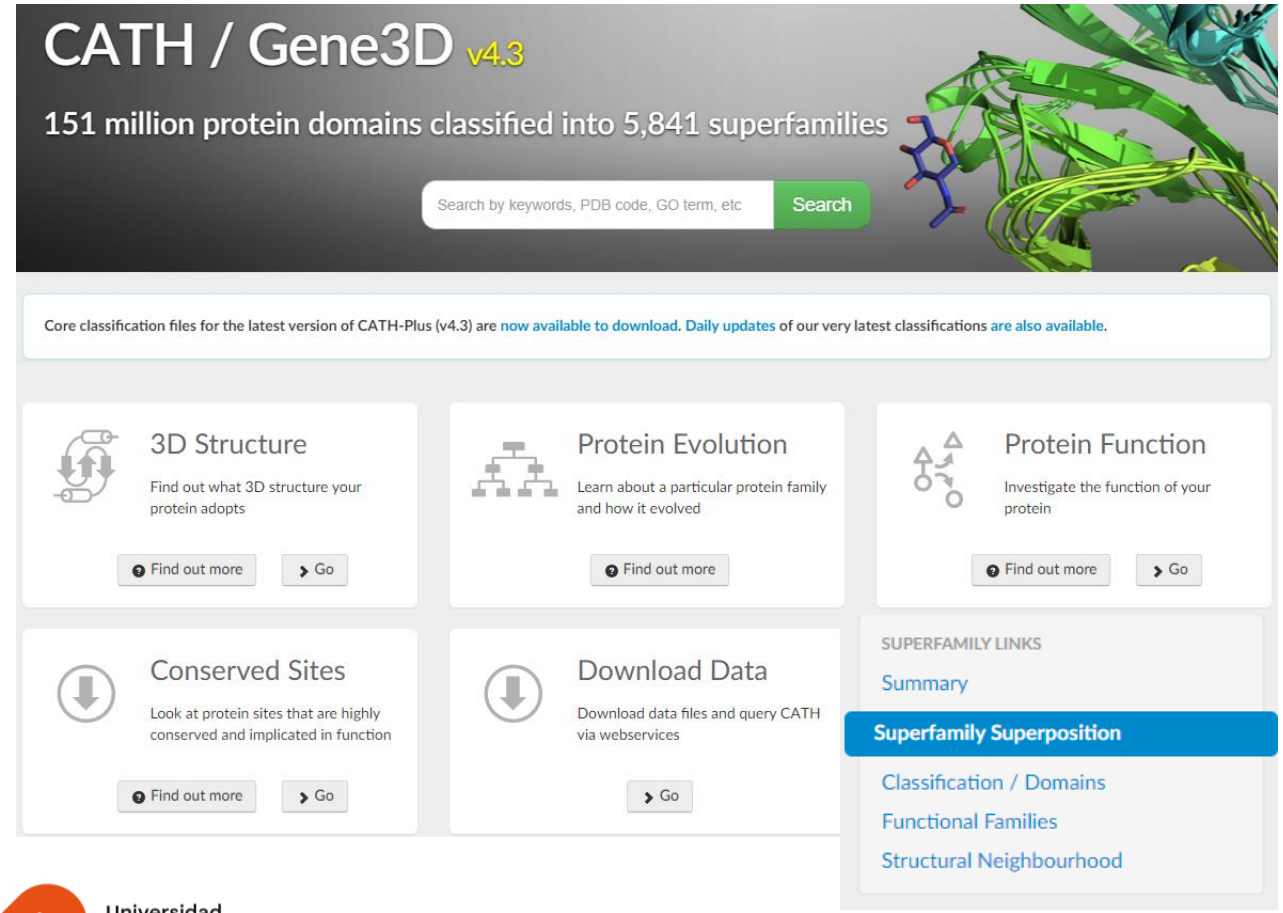


Temario - Contenidos

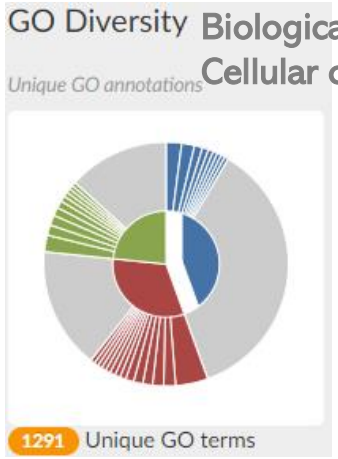
Tema 3. Proteómica estructural
3.1. Aspectos estructurales de las proteínas: niveles de organización
3.2. Métodos de caracterización estructural en proteómica
3.3. Servidores y bases de datos en proteómica estructural
3.4. Bioinformática estructural

3.3. Servidores y bases de datos en proteómica estructural

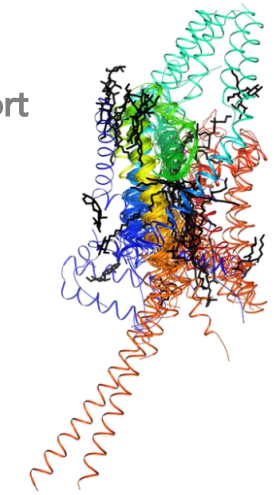
- **Protein Data Bank (PDB)**
- **CATH (www.cathdb.info)**



ID	Function Family (FunFam) Name	Total Sequences	Enzyme?	Structure?	Structural Representative	PDB Sites?	Alignment Diversity (0-100)
1	Sodium channel protein	2092			-	-	84.7
2	Potassium voltage-gated channel subfamily a member	1878		3D	4jtdH03	-	96.2

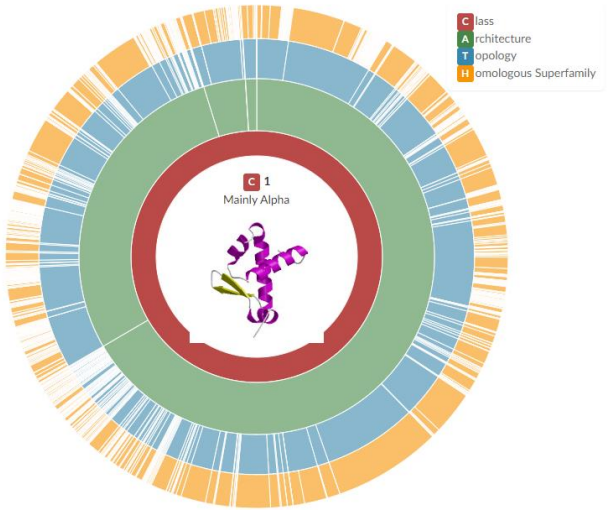


Molecular function: protein binding
Biological process: potassium ion transport
Cellular component: plasma membrane



CATH Classification

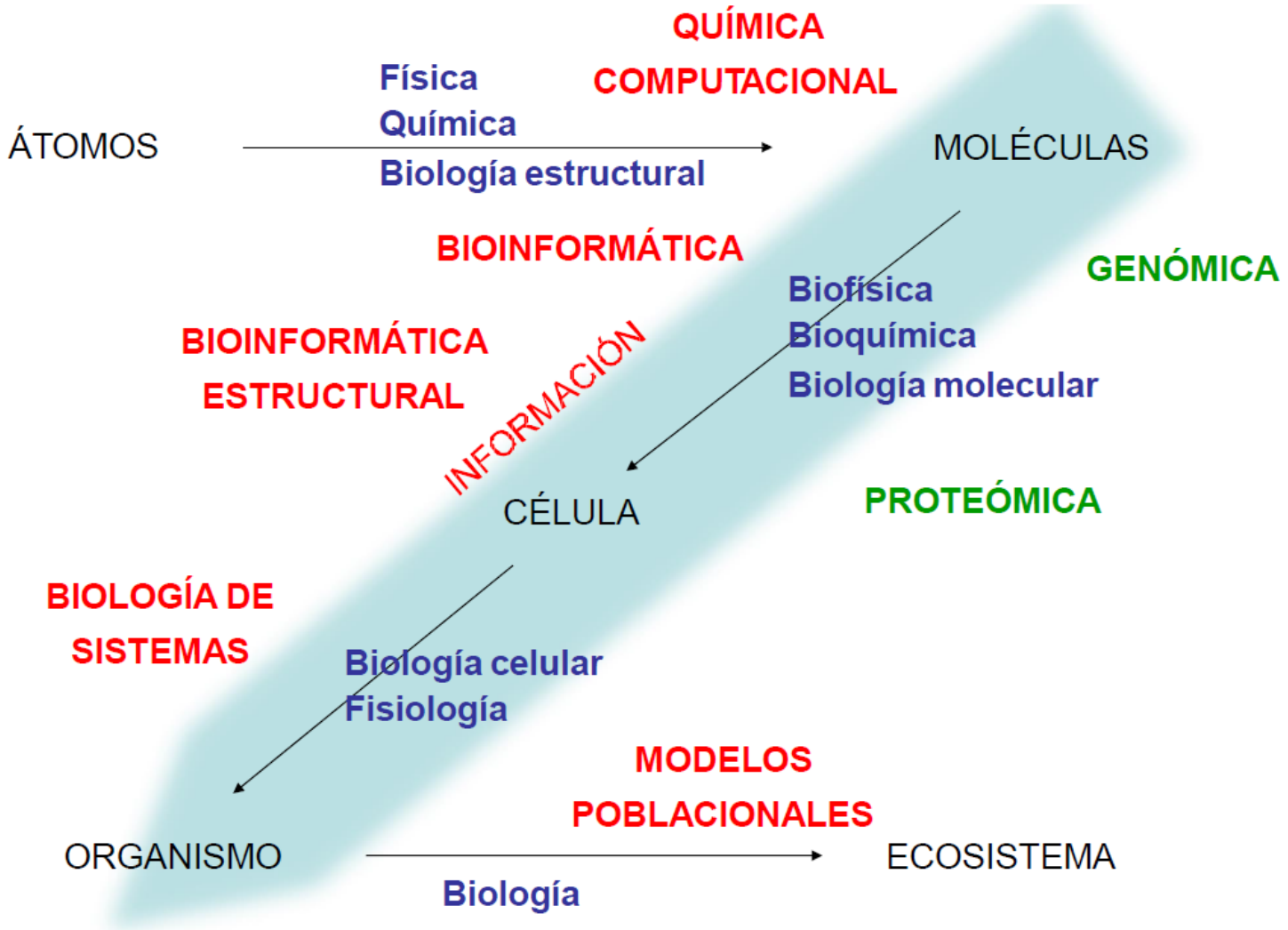
Level	CATH Code	Description
C	1	Mainly Alpha
A	1.10	Orthogonal Bundle
T	1.10.287	Helix Hairpins
H	1.10.287.70	



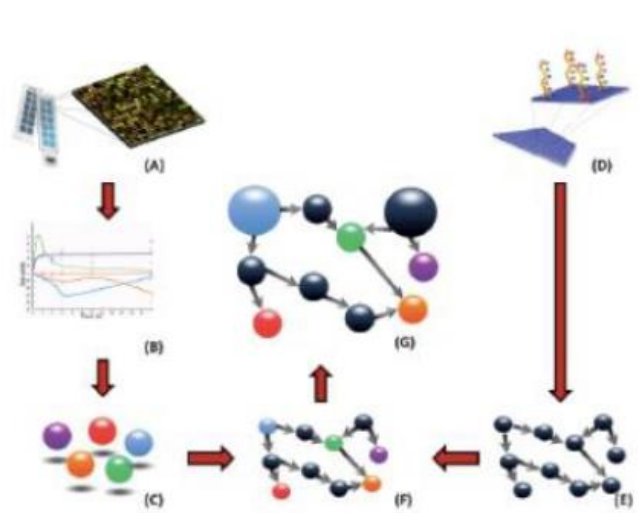
Temario - Contenidos

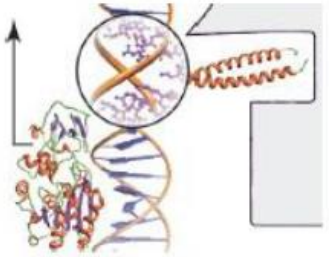
Tema 3. Proteómica estructural
3.1. Aspectos estructurales de las proteínas: niveles de organización
3.2. Métodos de caracterización estructural en proteómica
3.3. Servidores y bases de datos en proteómica estructural
3.4. Bioinformática estructural

3.4. Bioinformática estructural



3.4. Bioinformática estructural

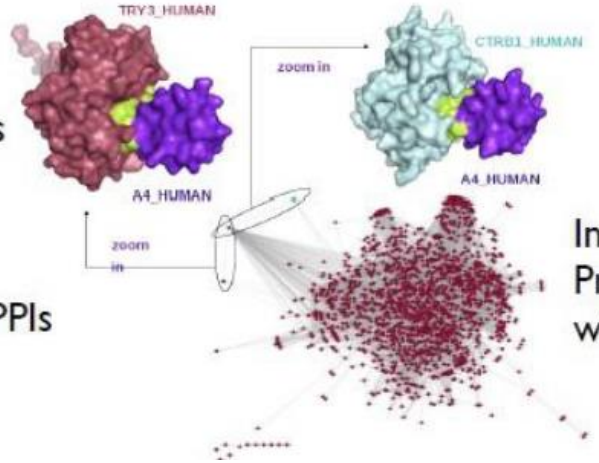




Predict and model Protein-DNA interactions
Predict Position Weight Matrices

Predict gene regulatory network
Predict signalling cascades activating genes
Predict phosphorylation cascades (Kinase-substrate)

Predict the function of proteins
Model the structure of proteins
Predict protein-protein interactions
Score reliability of PPIs
Model the structure of PPIs
Predict the strength of PPIs (Ka)
Predict changes in the network of PPIs

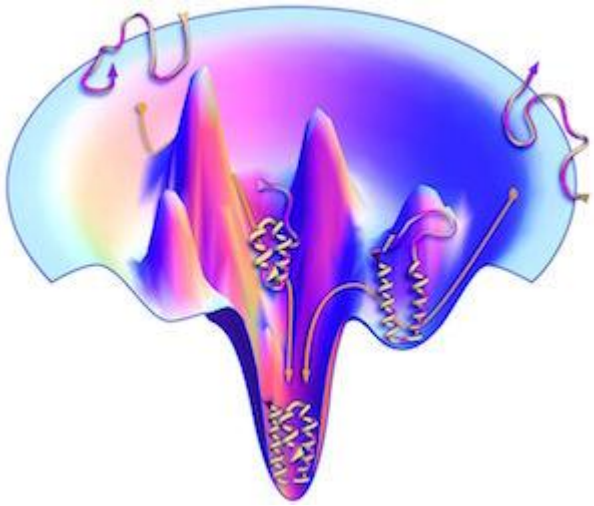


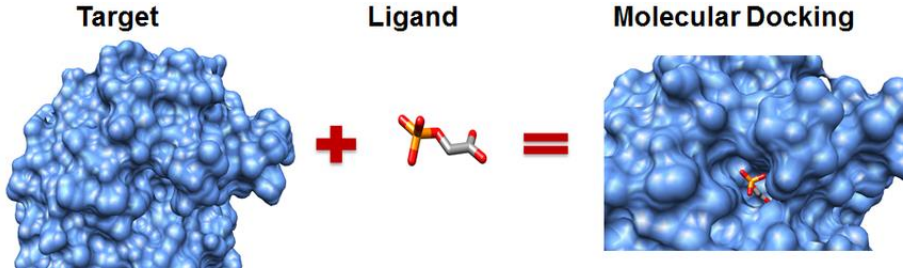
Integration of databases on PPIs
Predict target genes associated with phenotypes

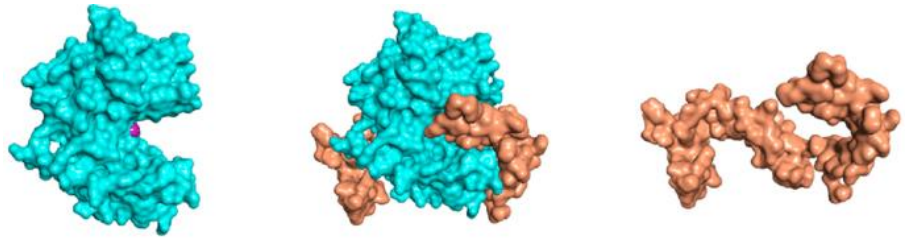
3.4. Bioinformática estructural

Bioinformática estructural

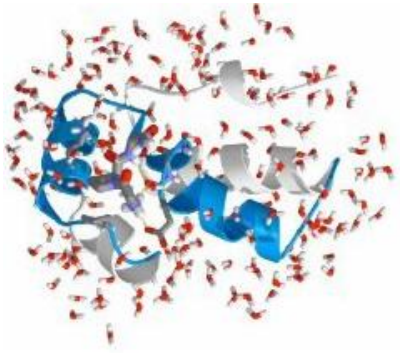
- Plegamiento de proteínas *(folding)*
  - *Ab initio*
  - *Homology modelling*
- Acoplamiento molecular *(docking)*
  - Proteína-proteína
  - Proteína-péptido
  - Proteína-ligando químico
  - Proteína-ácido nucleico







- Dinámica molecular



Universidad
Internacional
de Valencia

universidadviu.com

De:
Planeta Formación y Universidades