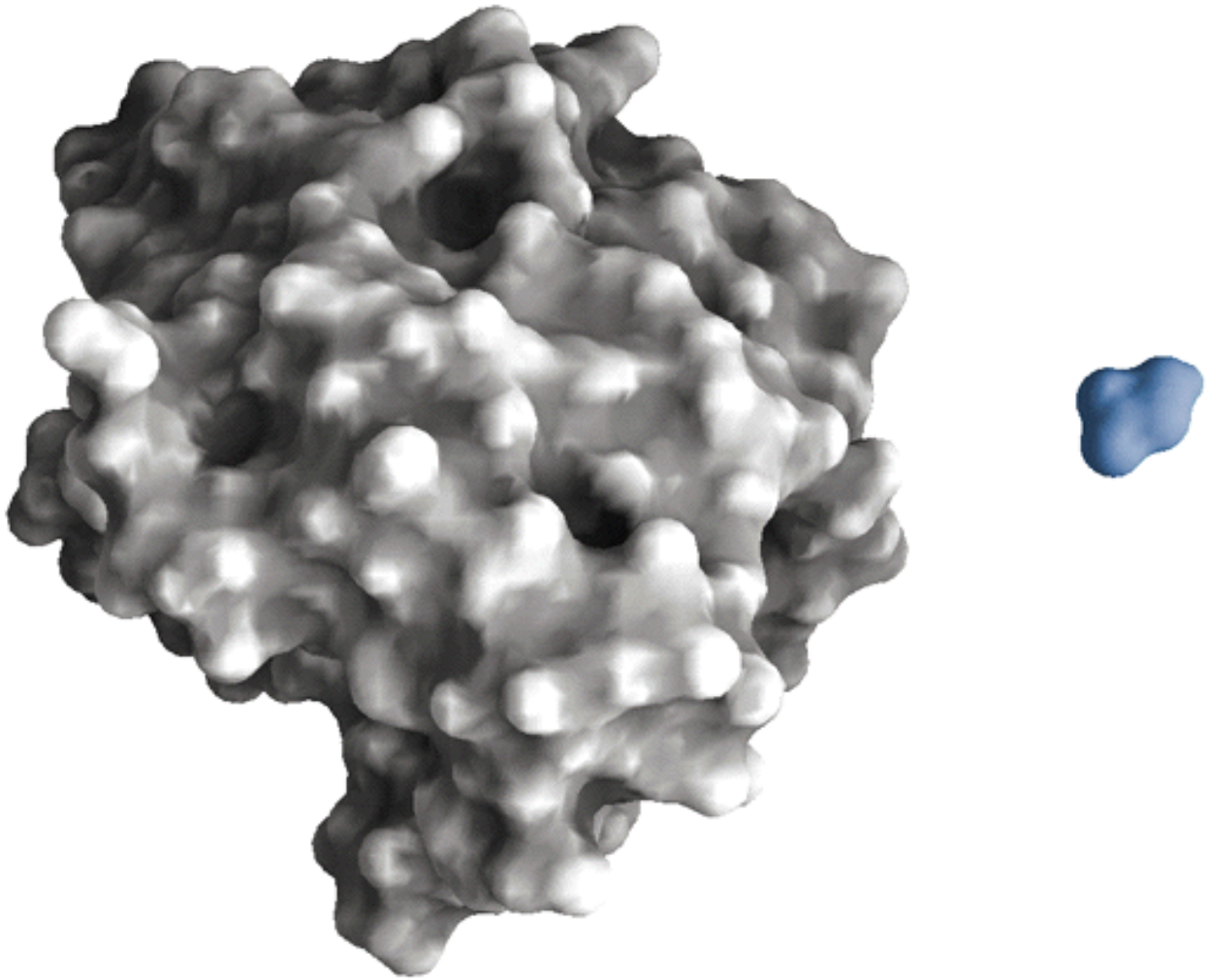
What are the proteins? (second part)

Dr. Abel MORENO, Institute of Chemistry, UNAM. MEXICO.

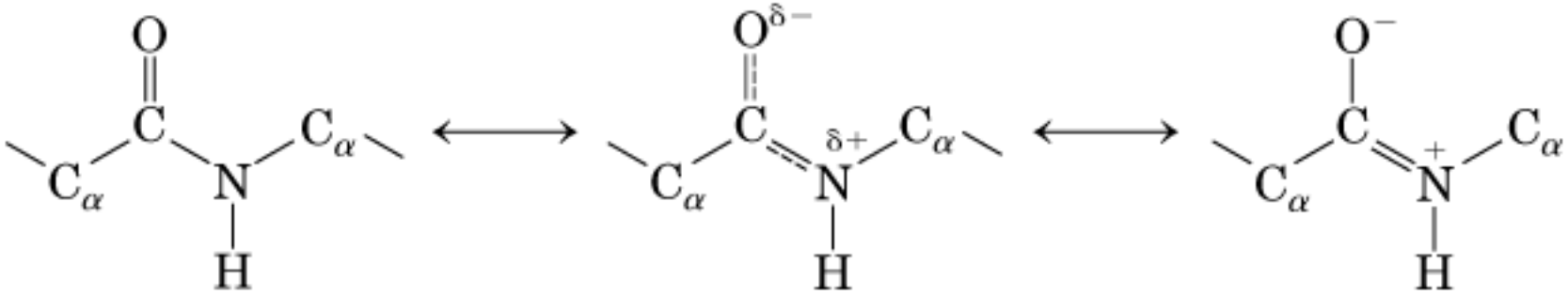
E-mail address: carcamo@unam.mx or abel.moreno@mac.com

Three-dimensional structure of proteins

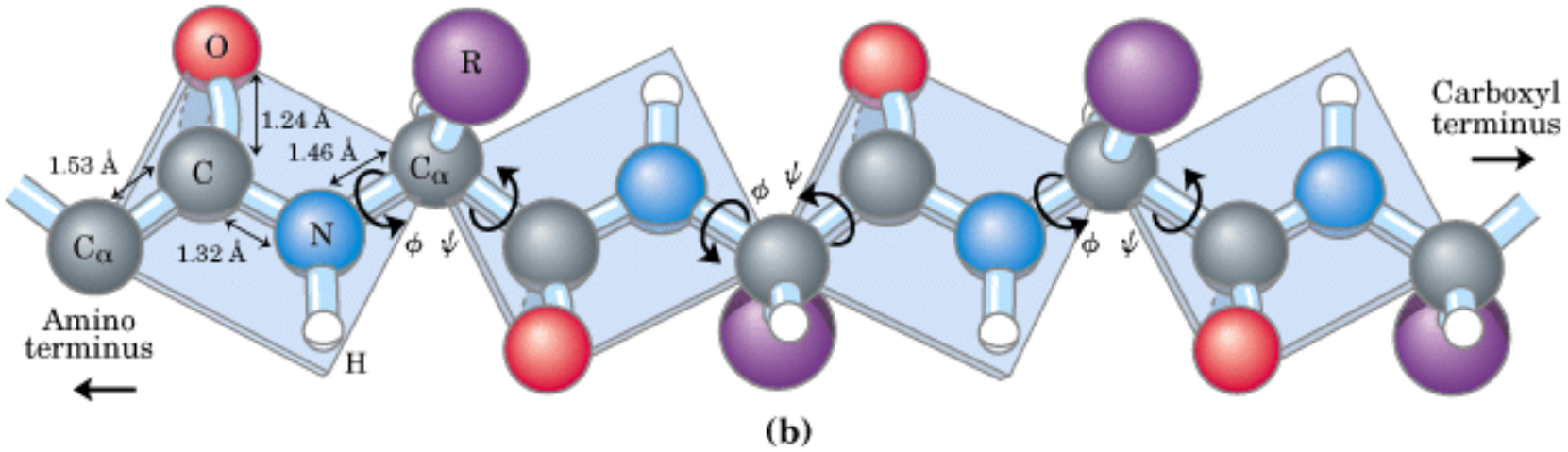


Estructura tridimensional de la enzima quimotripsina, una proteína globular. En azul se muestra la molécula de glicina. Las estructuras de proteínas se almacenan en el Protein Data Bank o PDB: **http://www.rcsb.org/pdb**

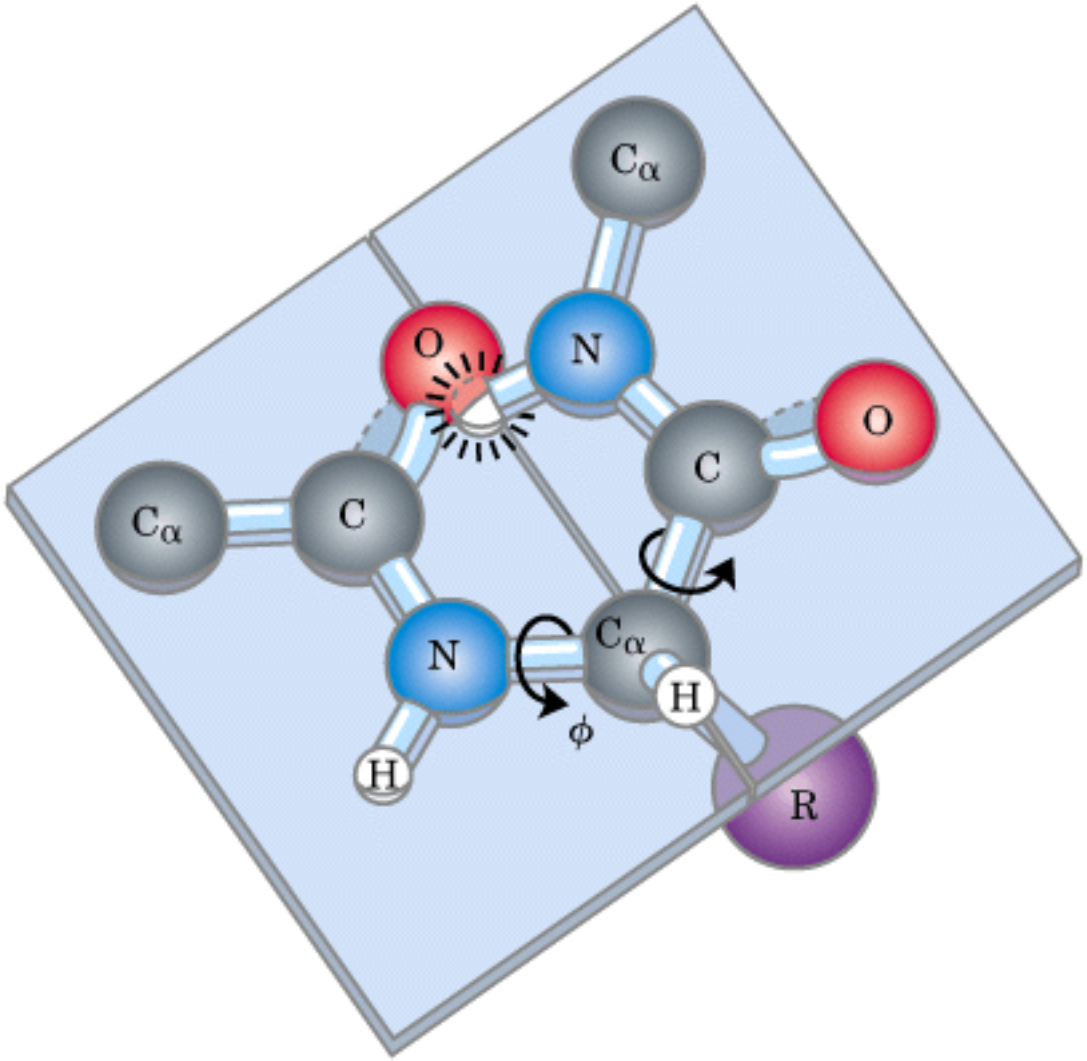
The carbonyl oxygen has a partial negative charge and the amide nitrogen a partial positive charge, setting up a small electric dipole. Virtually all peptide bonds in proteins occur in this trans configuration; an exception is noted in Figure 6–8b.



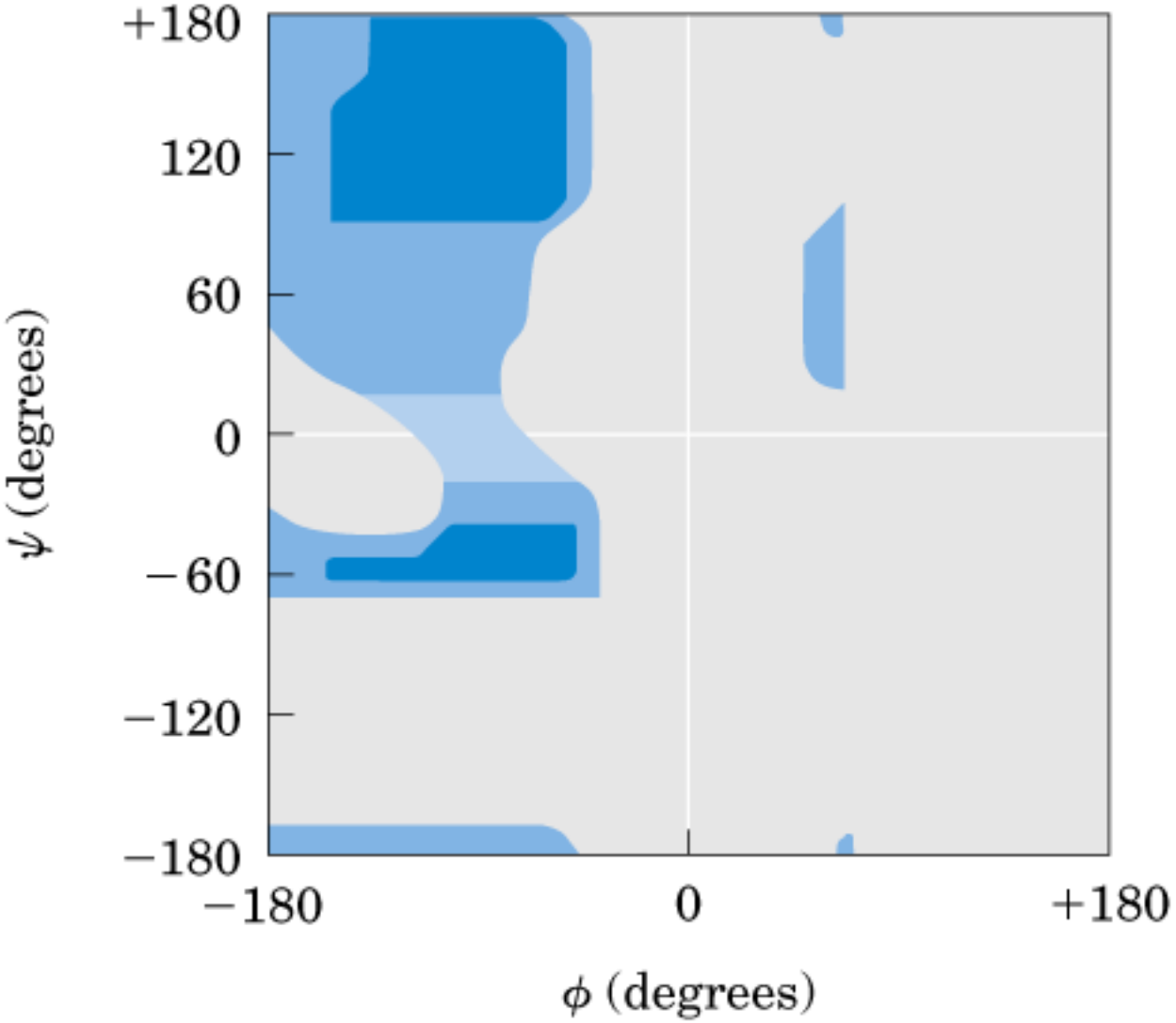
(a)


O
R
1.24 Å
1.53 Å
1.46 Å
C
Cα
N
Cα
1.32 Å
ϕ
ψ
ϕ
ψ
ϕ
ψ
Carboxyl
terminus
Amino
terminus
H

(b)

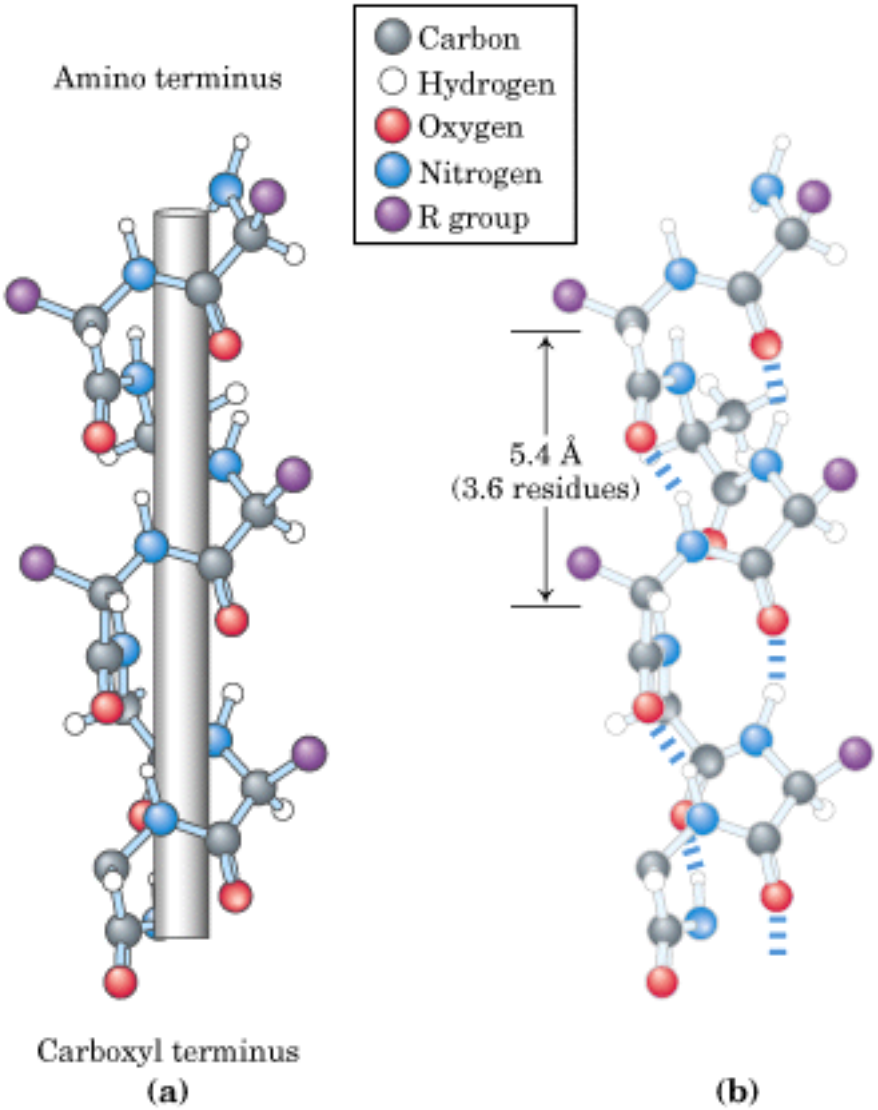


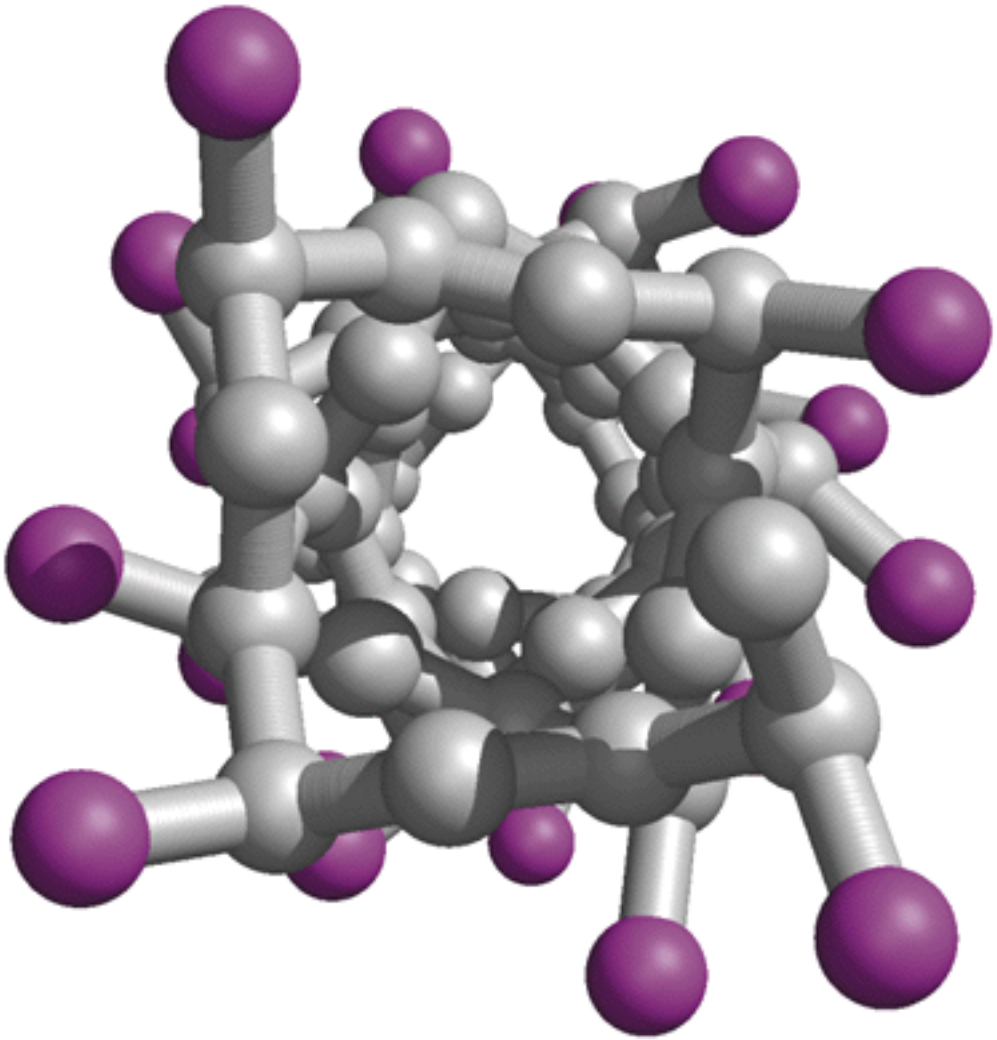
(c)



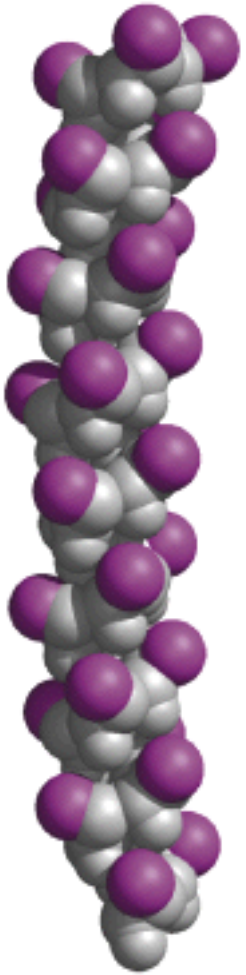
Representación de Ramachandran para el aminoácido L-Ala.

La hélice alfa en proteínas y diferentes aspectos de su estructura.

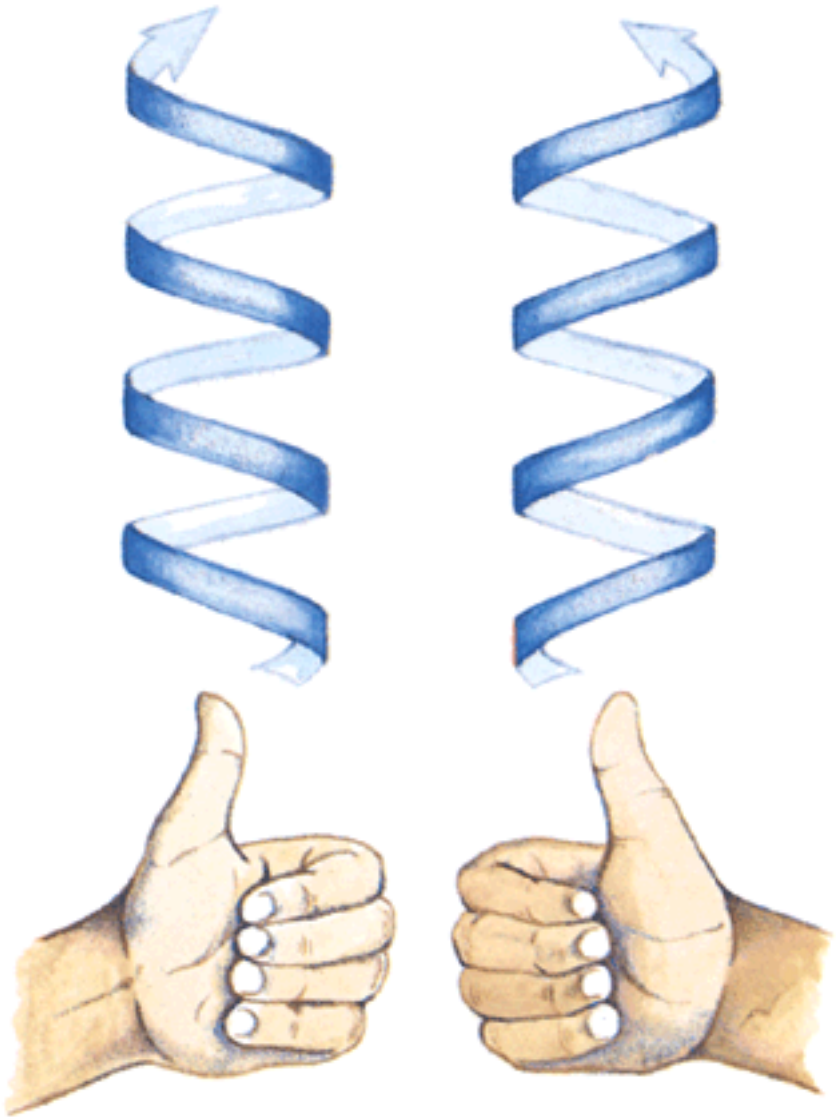


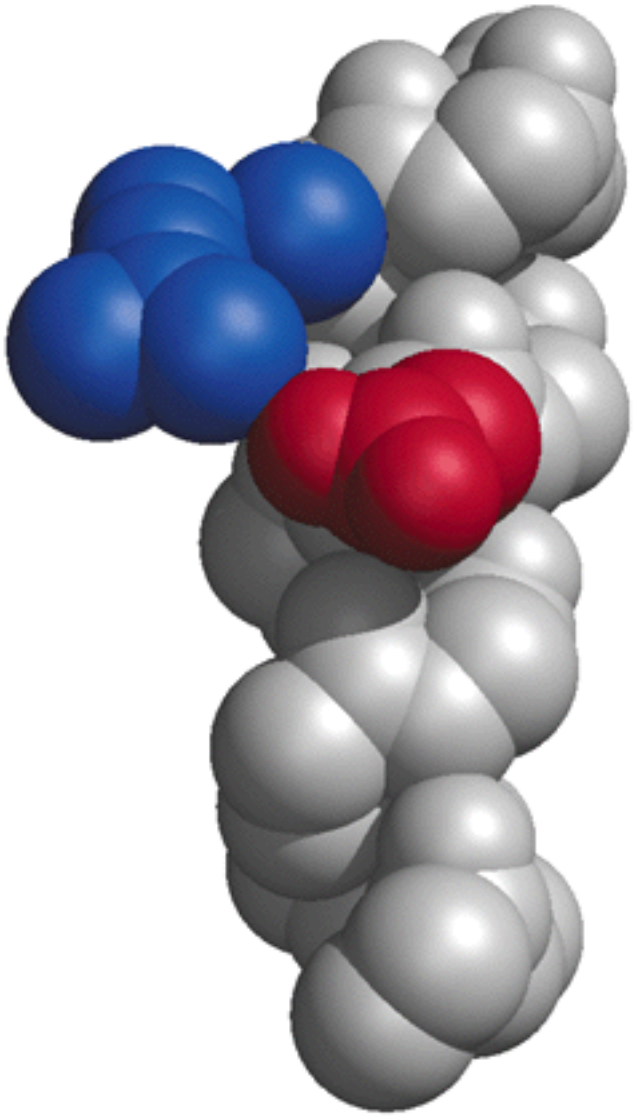


(c)

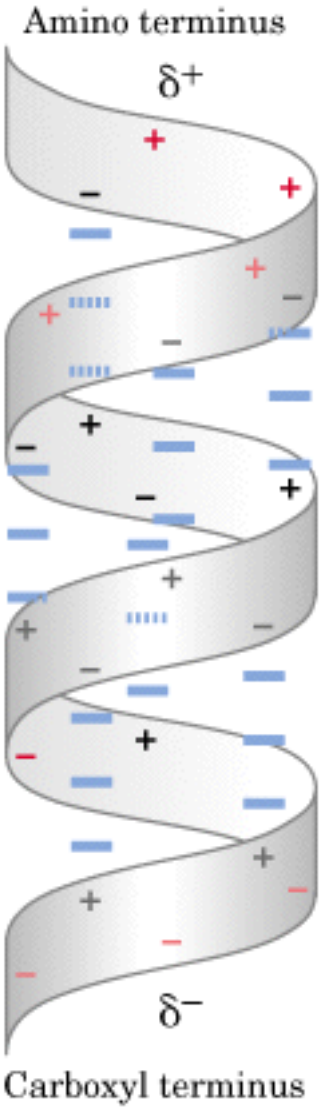


(d)

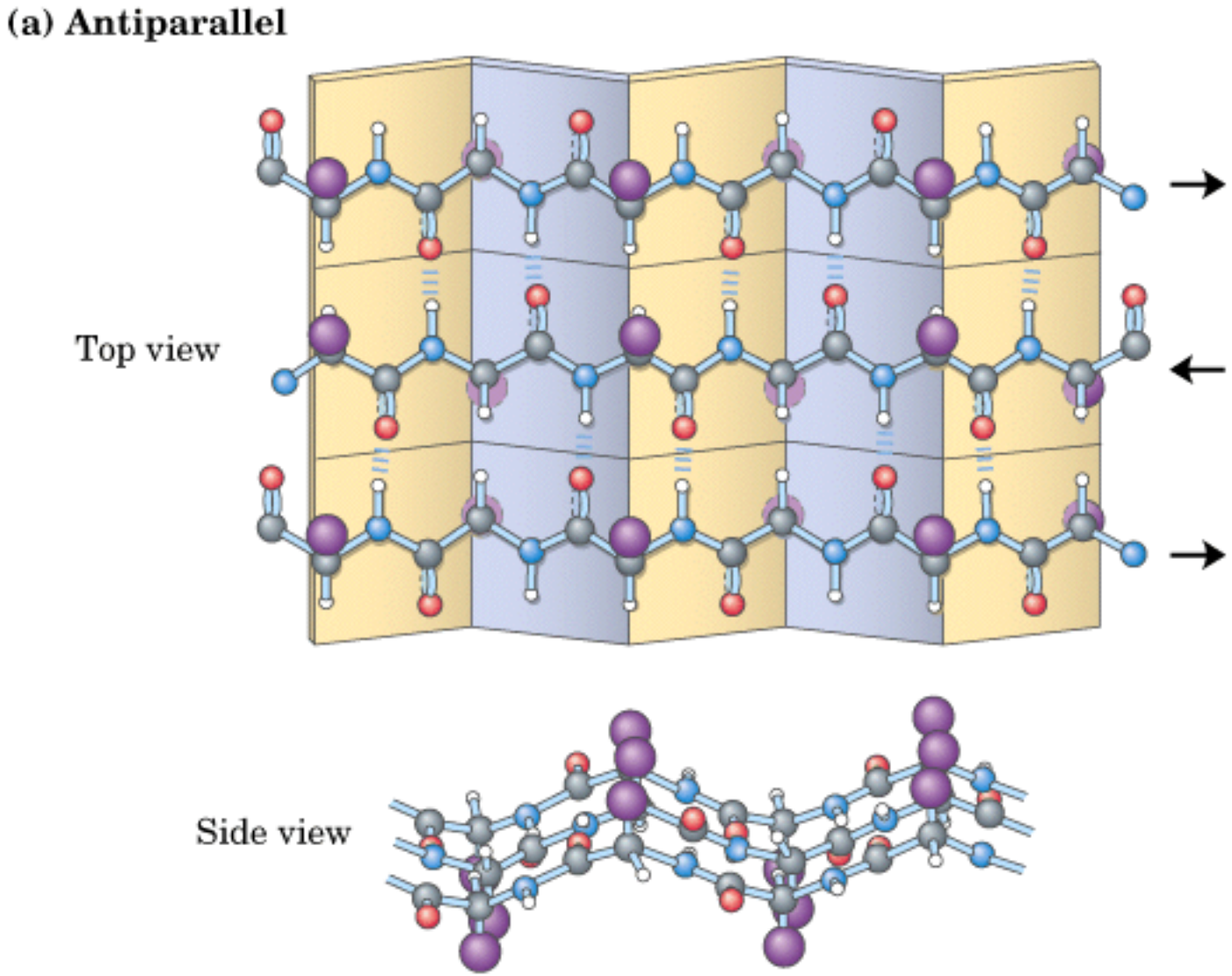




Interacciones entre los grupos R de aminoácidos separados por tres residuos de una hélice alfa.



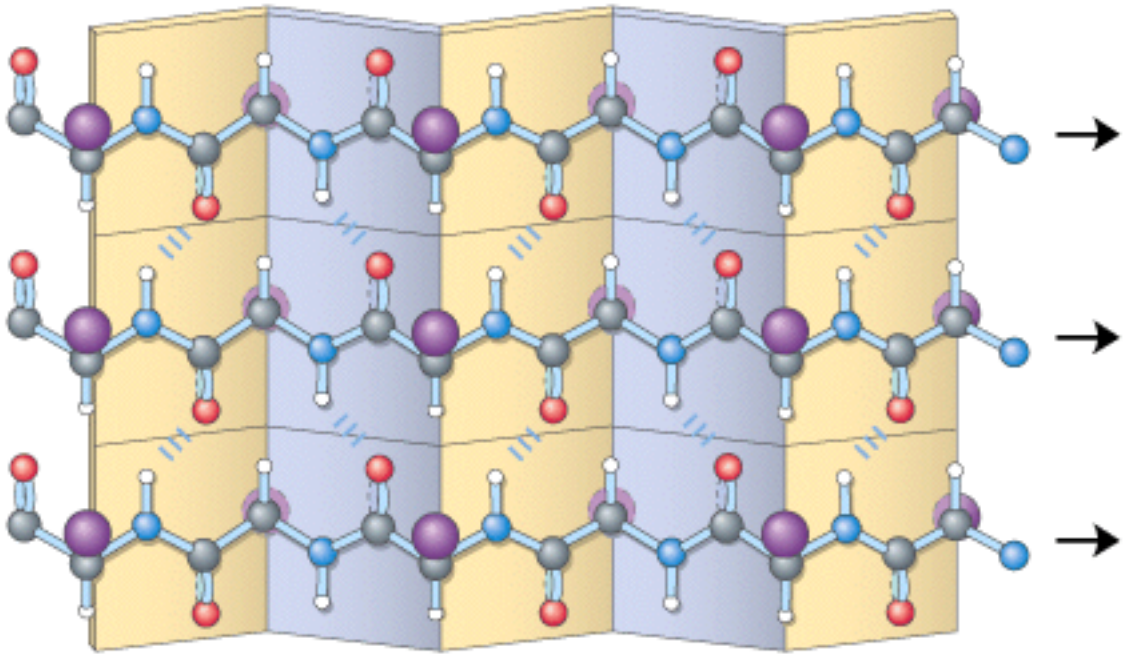
Dipolo de una hélice alfa. Este muestra el dipolo eléctrico del enlace peptídico.



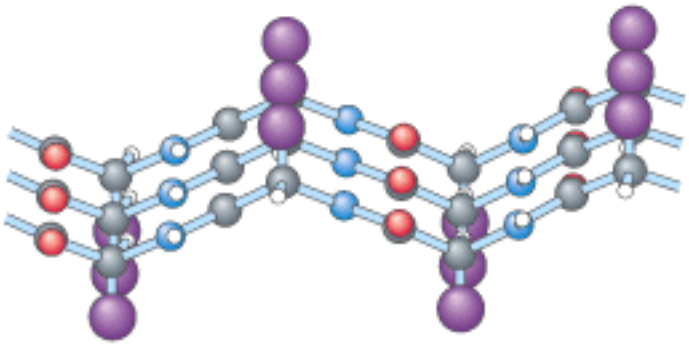
La conformación de hojas beta en cadenas polipeptídicas.

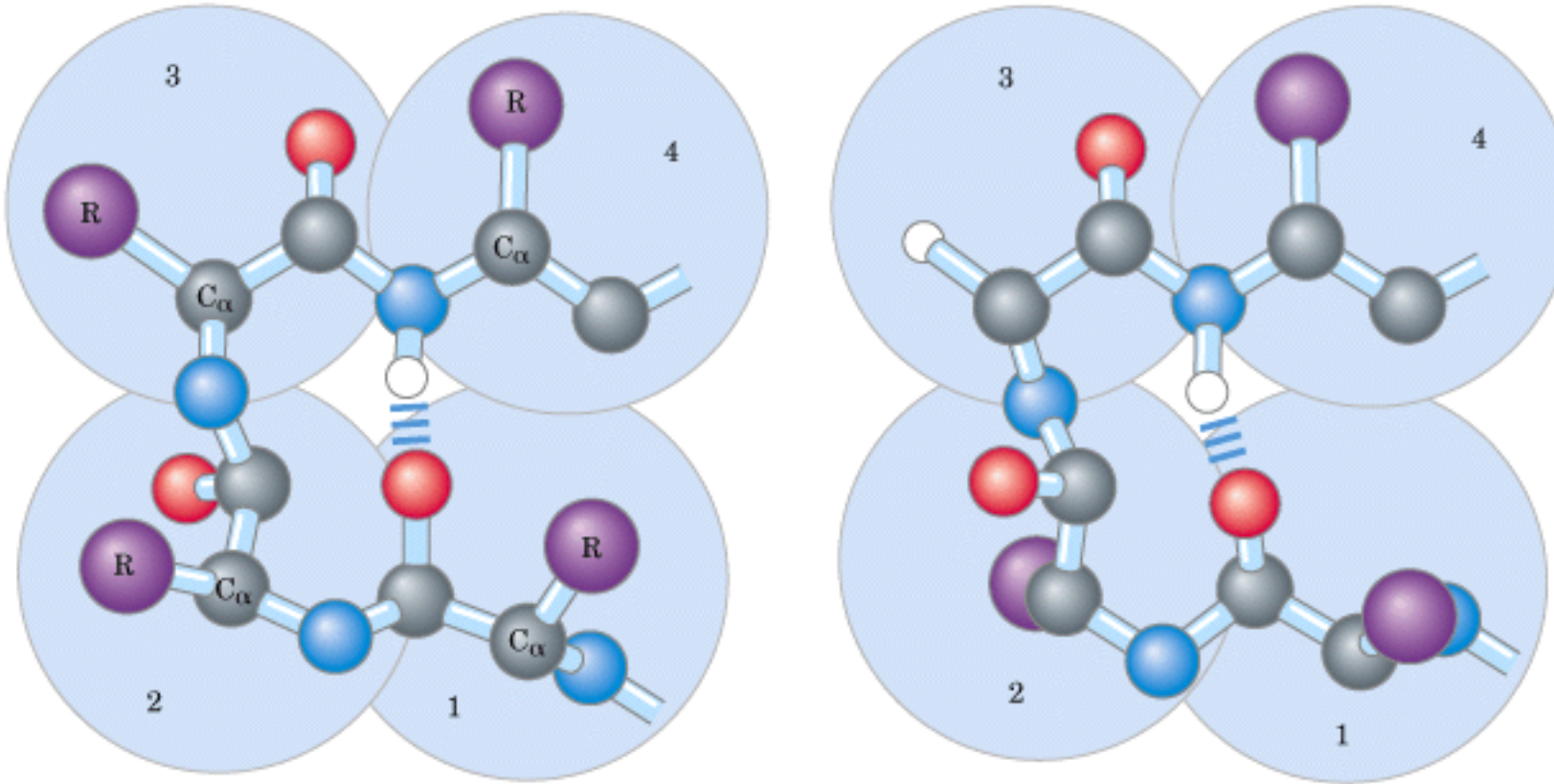
(b) Parallel

Top view



Side view



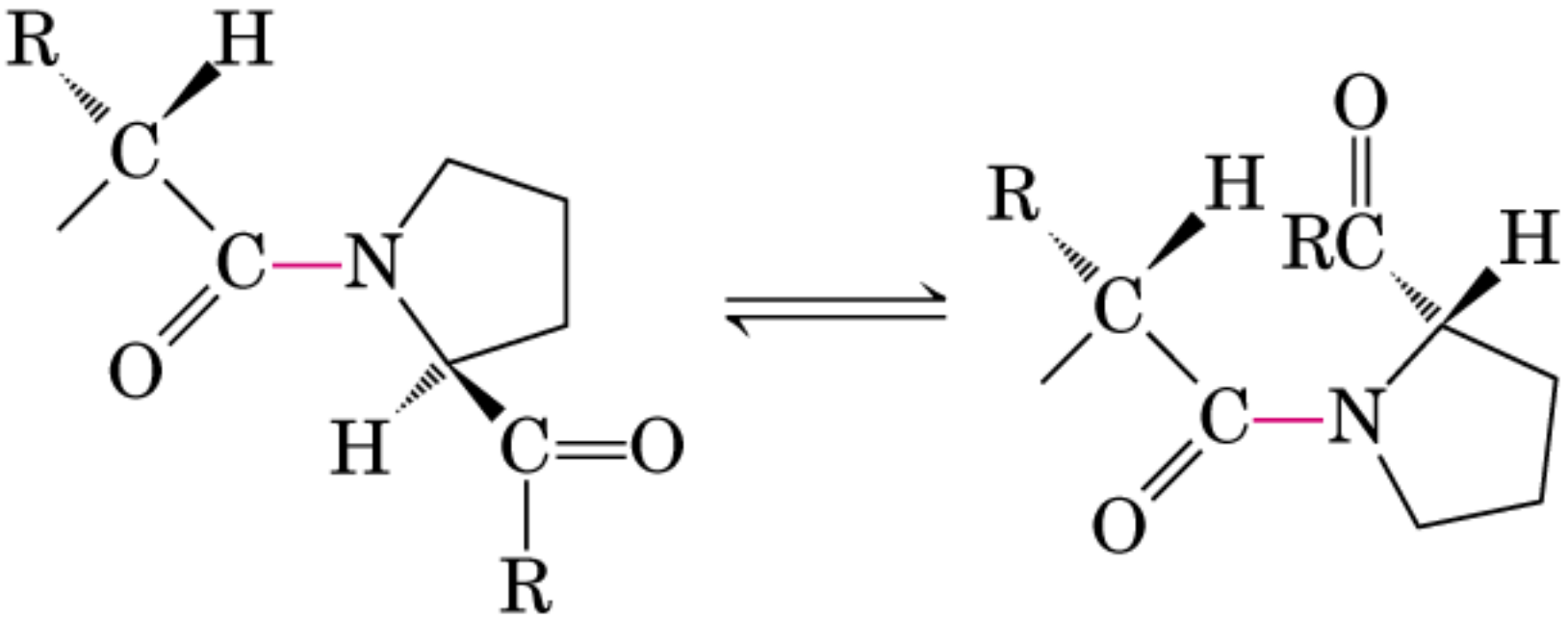


β **Turns**
(a)

La estructura de los giros de las hojas beta.

> 99.95% (excepto prolina)

6% en imino de prolina

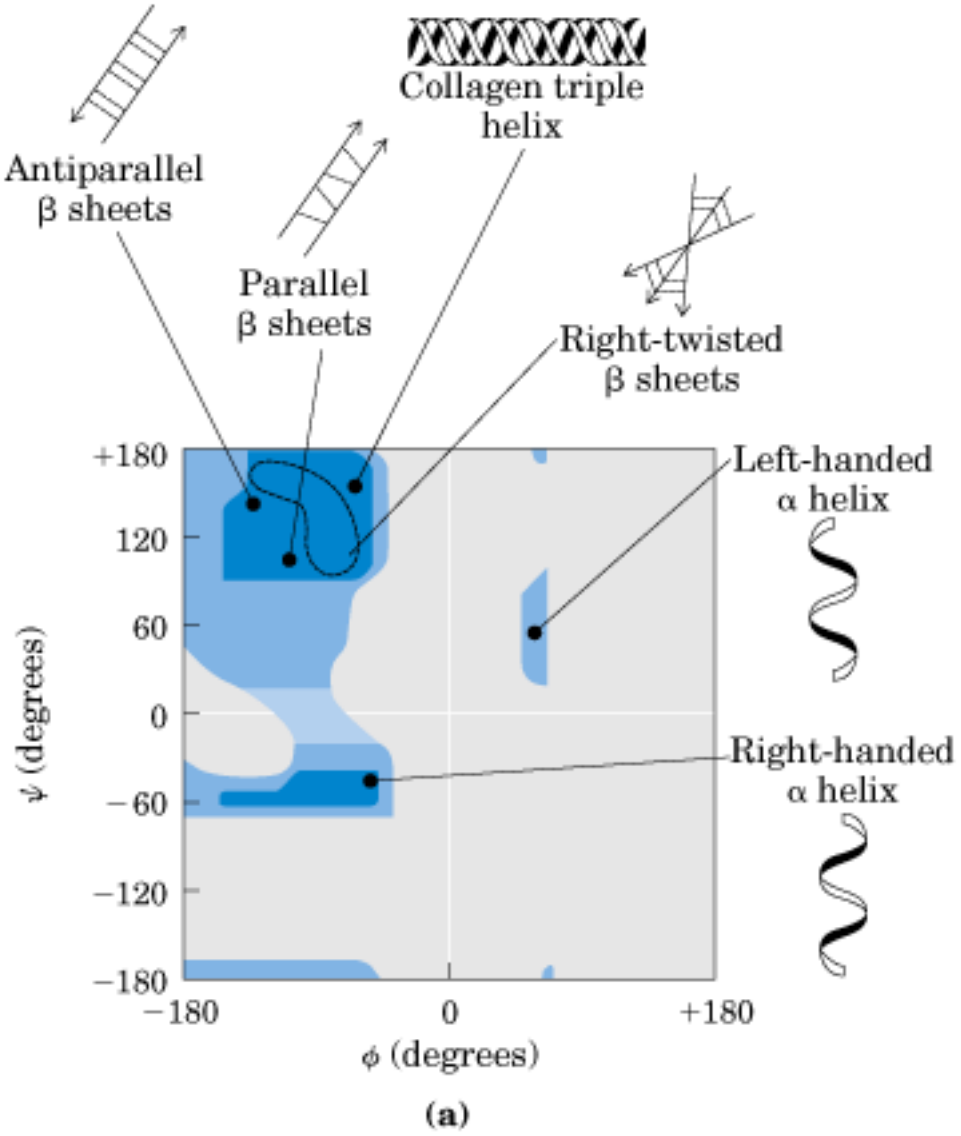


trans

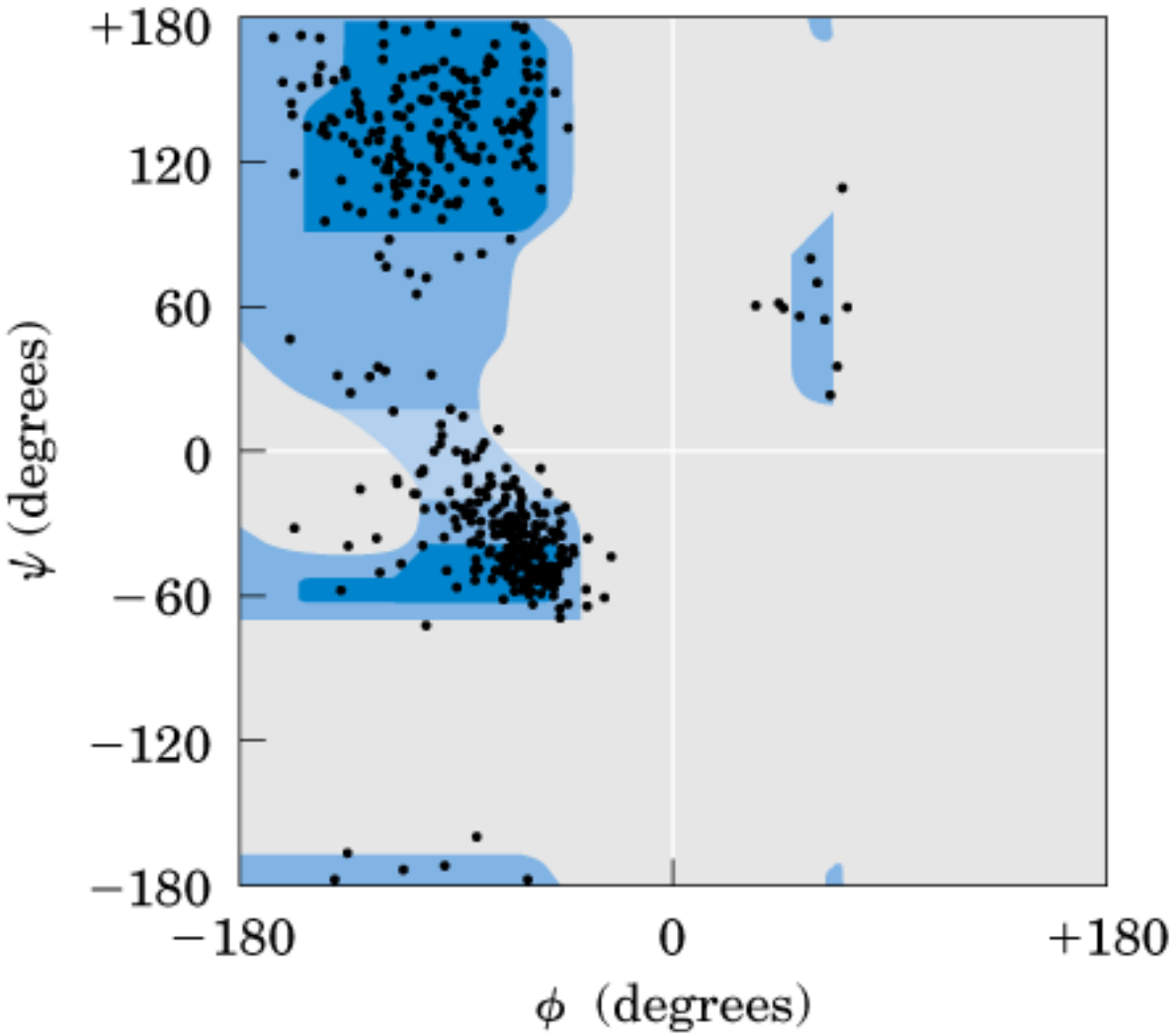
cis

Proline isomers

(b)

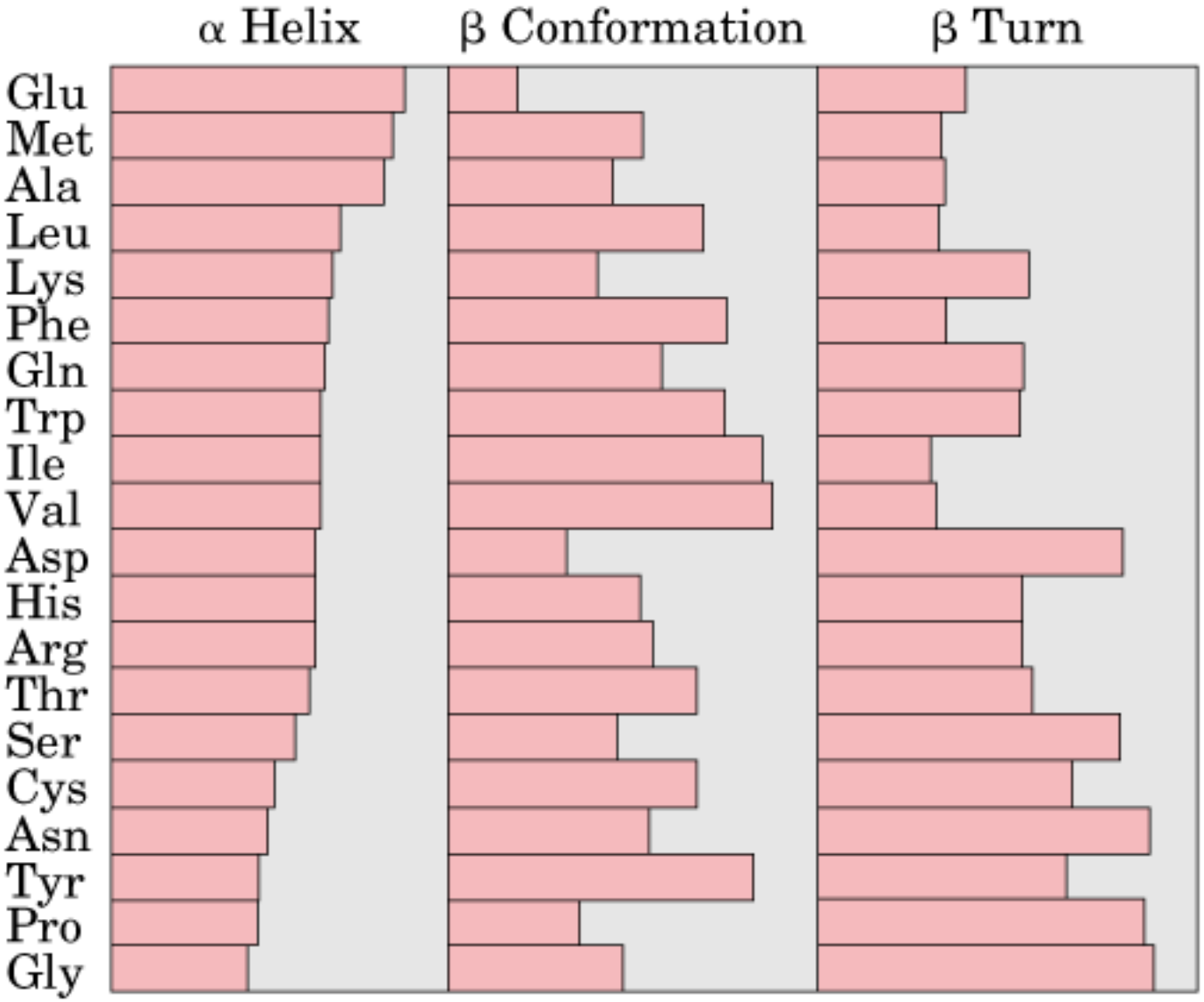


Representación de Ramachandran de diferentes estructuras y conformaciones observadas en proteínas



(b)

Los valores de ϕ y ψ de todos los residuos de aminoácidos con excepción de Gly de la enzima piruvato cinasa.

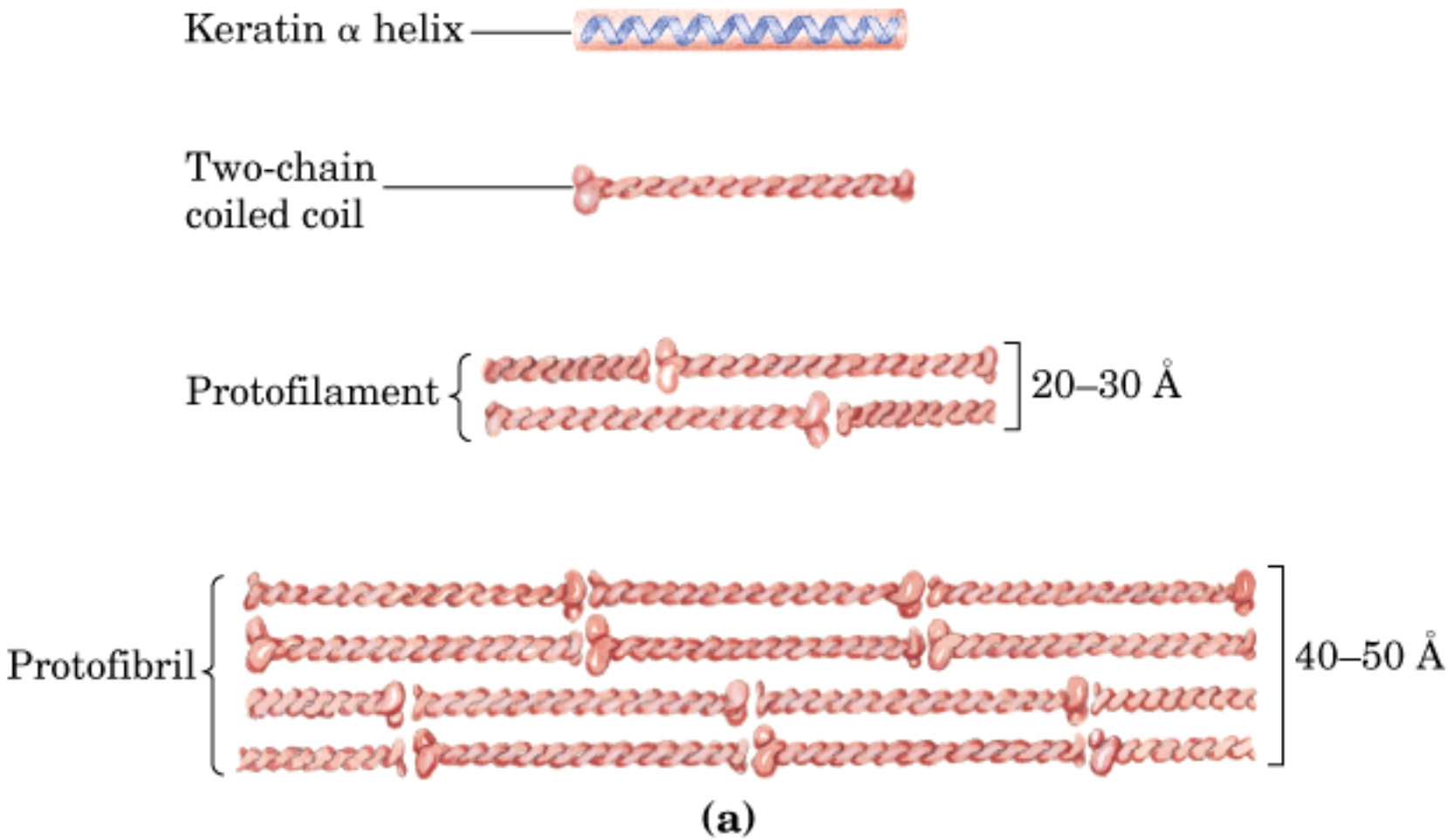


Probabilidades relativas de que un aminoácido se halle en los tres tipos de estructuras secundarias comunes de proteínas.

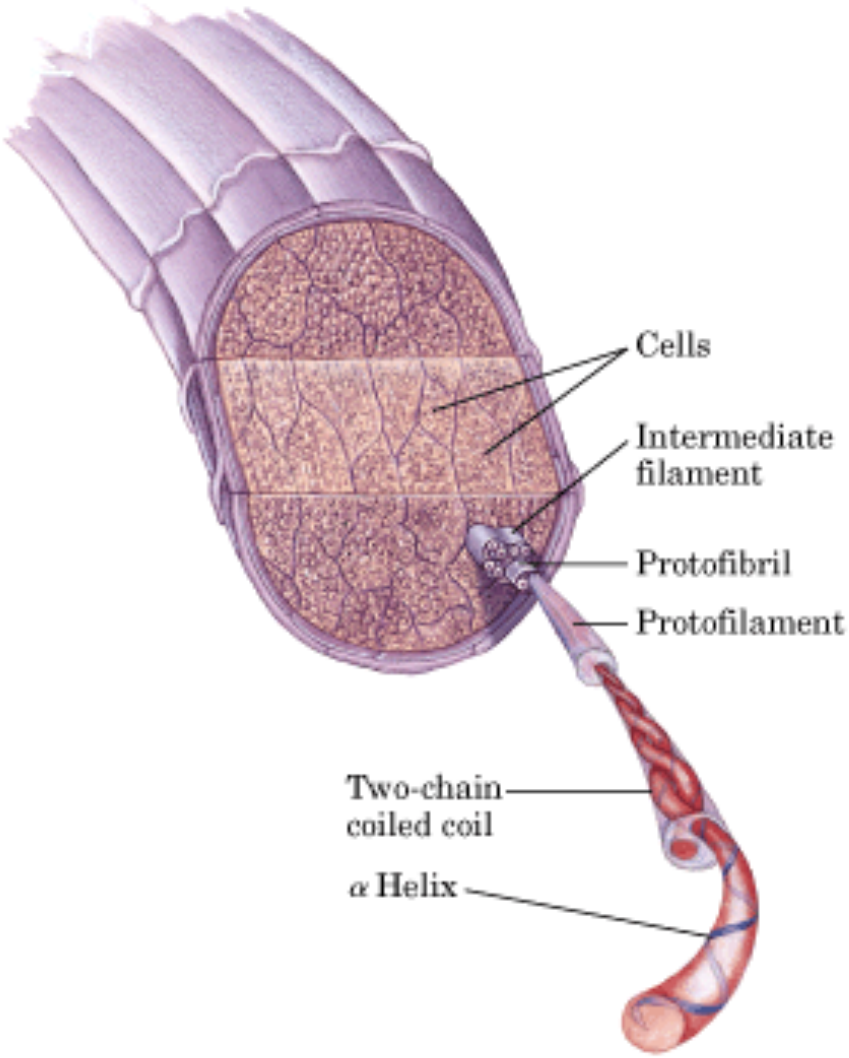
table 6–1

Secondary Structures and Properties of Fibrous Proteins

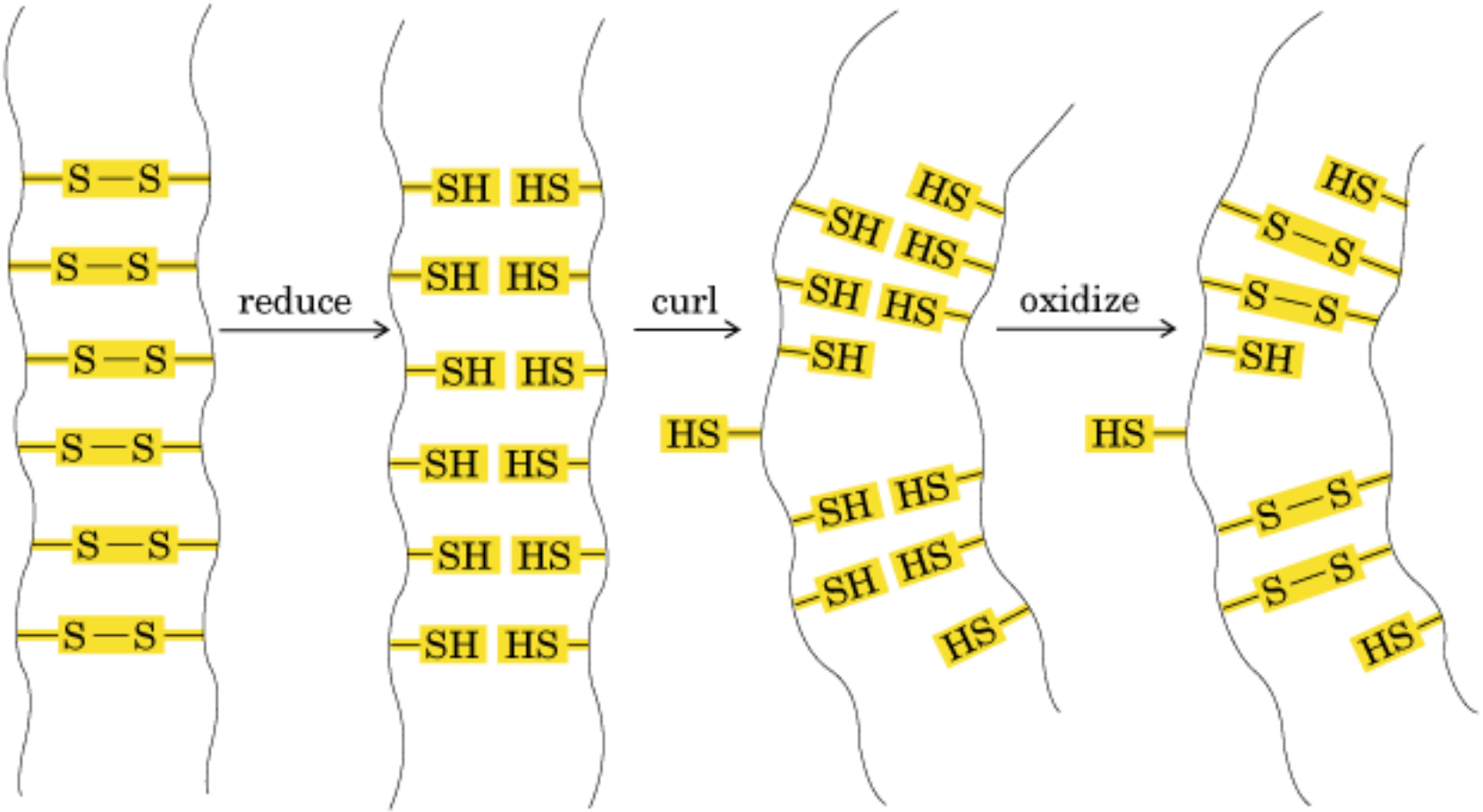
Structure	Characteristics	Examples of occurrence
α Helix, cross-linked by disulfide bonds	Tough, insoluble protective structures of varying hardness and flexibility	α-Keratin of hair, feathers, and nails
β Conformation	Soft, flexible filaments	Silk fibroin
Collagen triple helix	High tensile strength, without stretch	Collagen of tendons, bone matrix



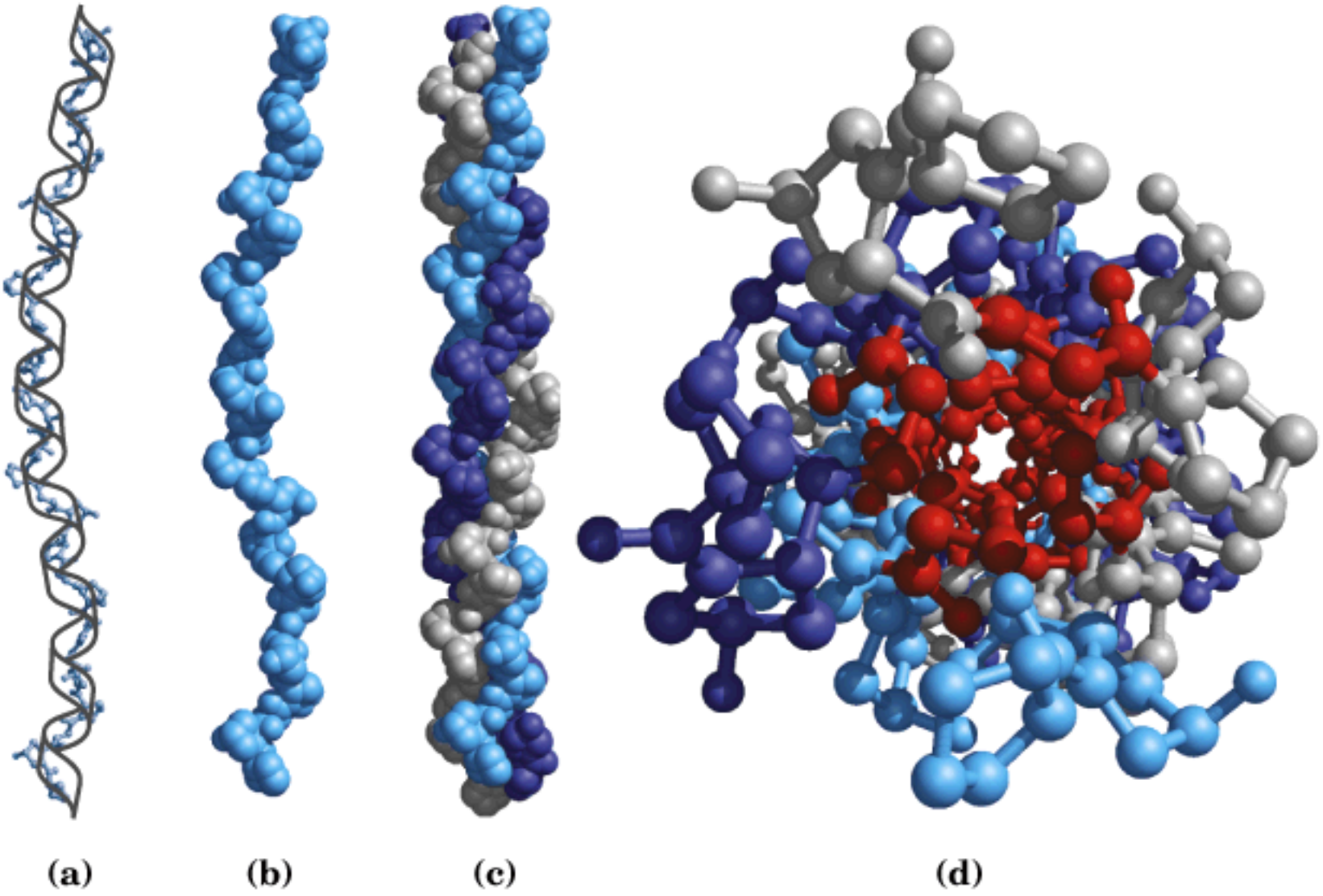
Estructura del cabello: (a) es la alfa-queratina.



Cross section of a hair
(b)

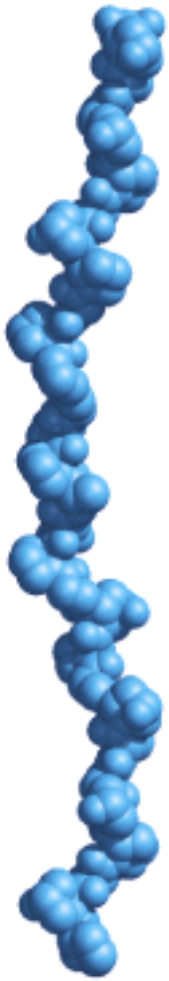


La ondulación permanente del cabello, un ejemplo de ingeniería bioquímica.

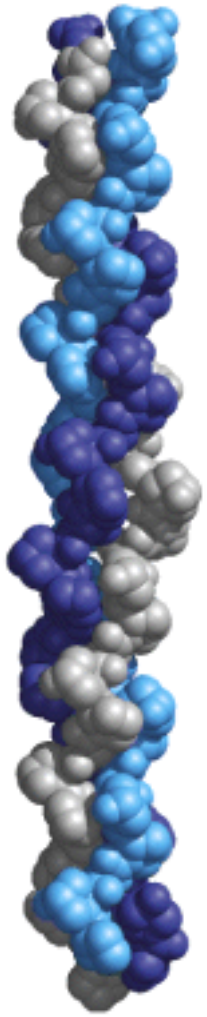


Estructuras de colágeno, proteína que se encuentra abundantemente en la piel y en los huesos.

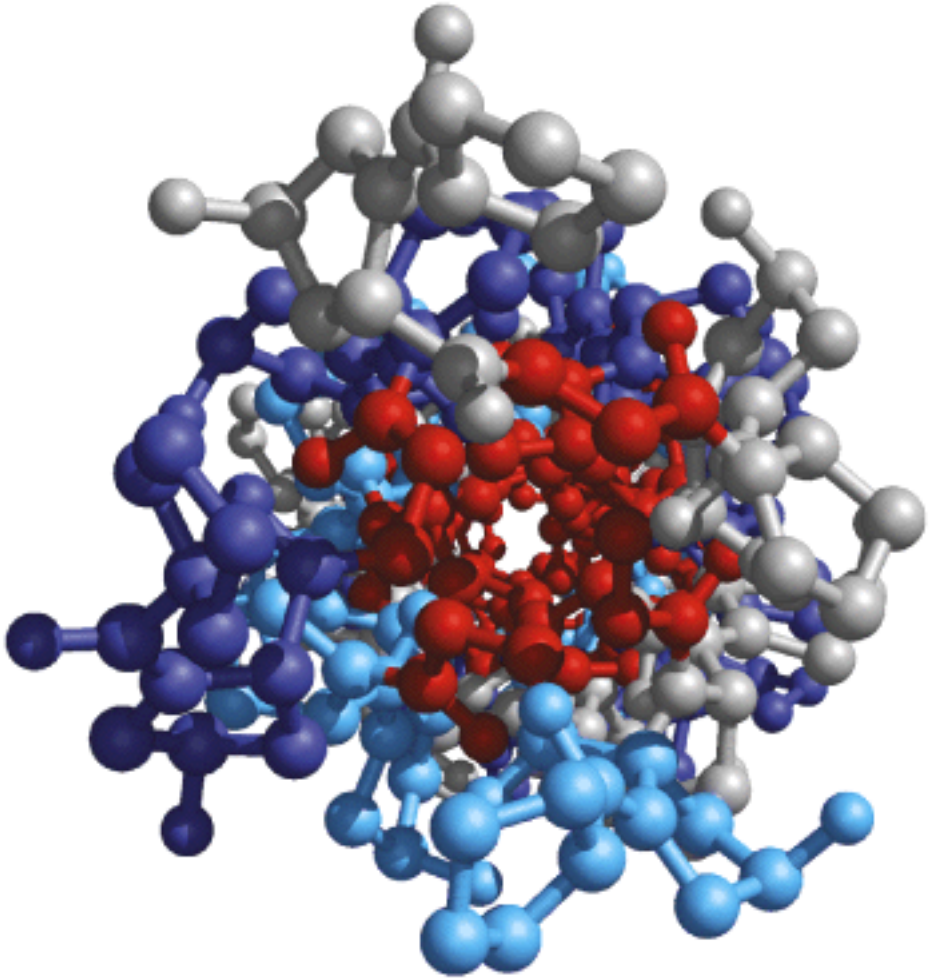
(a)



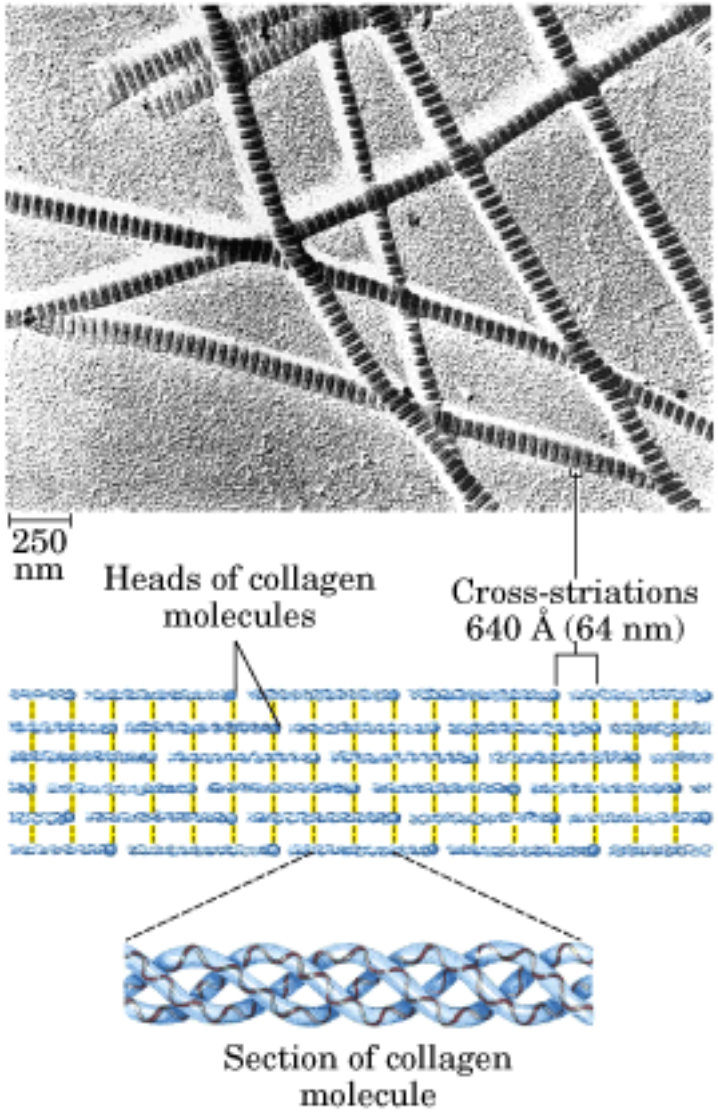
(b)



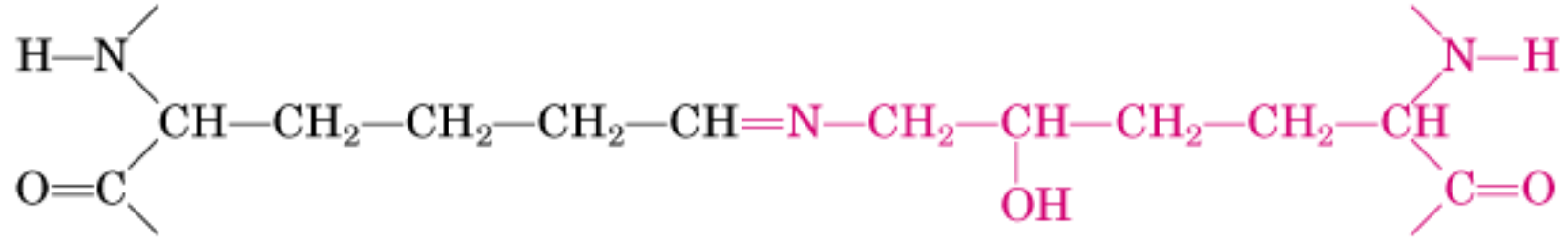
(c)



(d)



Estructuras de las fibrillas del colágeno



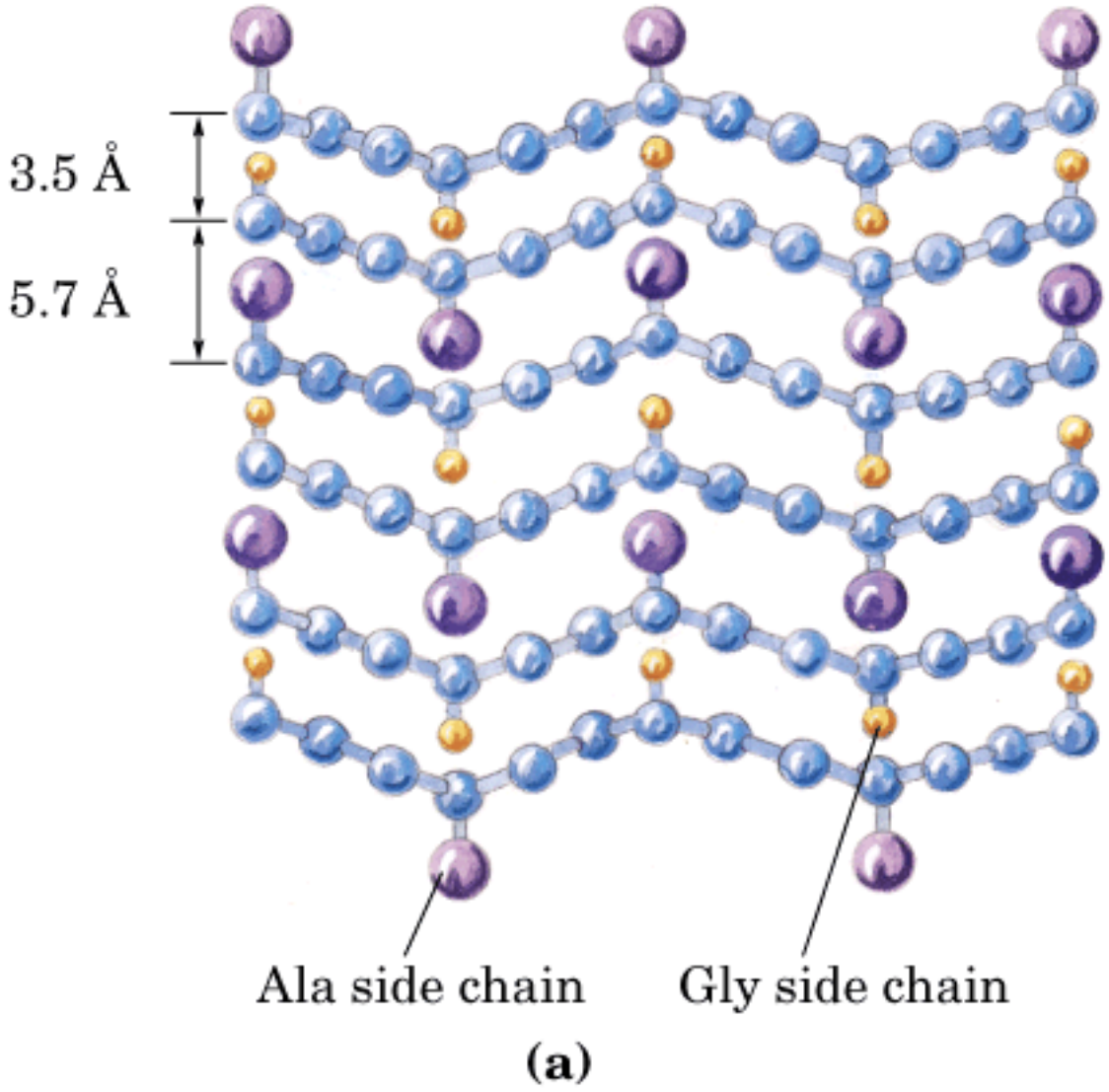
Polypeptide chain

Lys residue minus ϵ-amino group (norleucine)

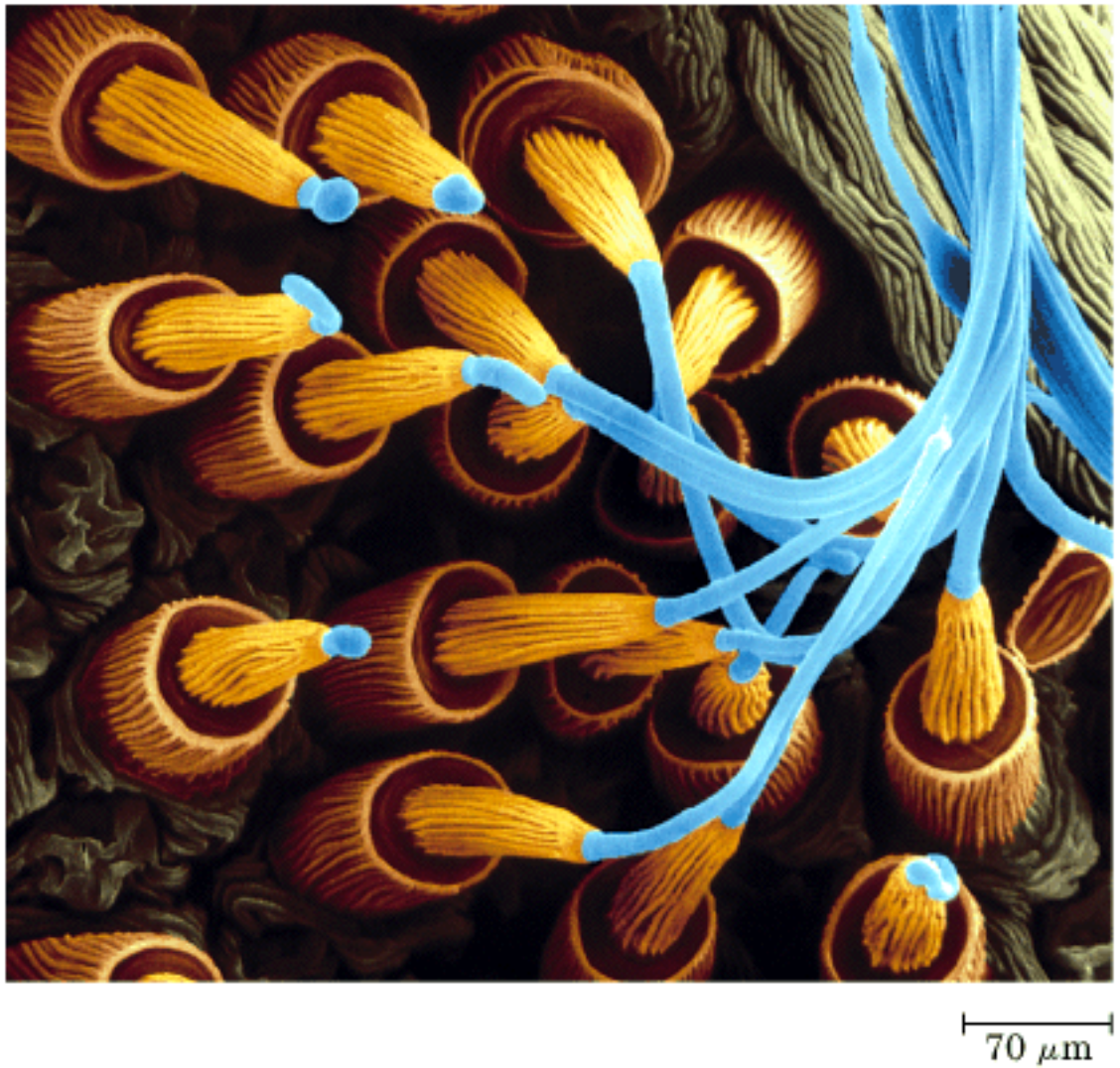
HyLys residue

Polypeptide chain

Dehydrohydroxylysinonorleucine

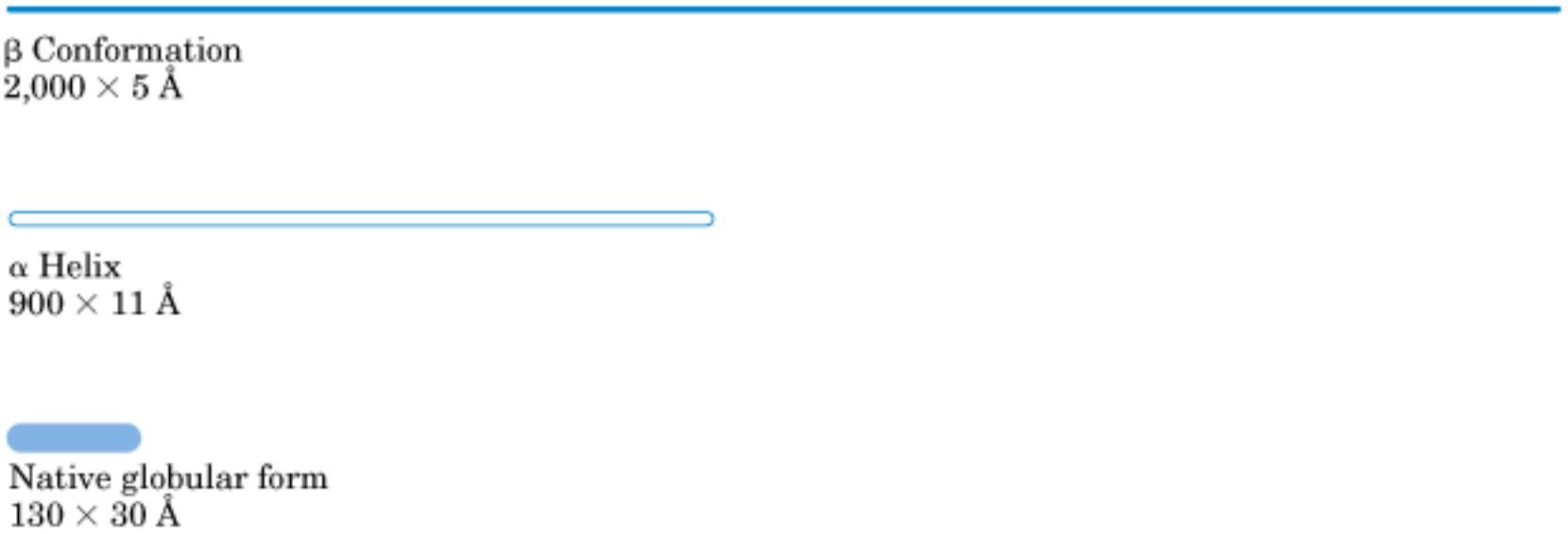


Estructura de la seda, constituida por una proteína llamada fibroína.

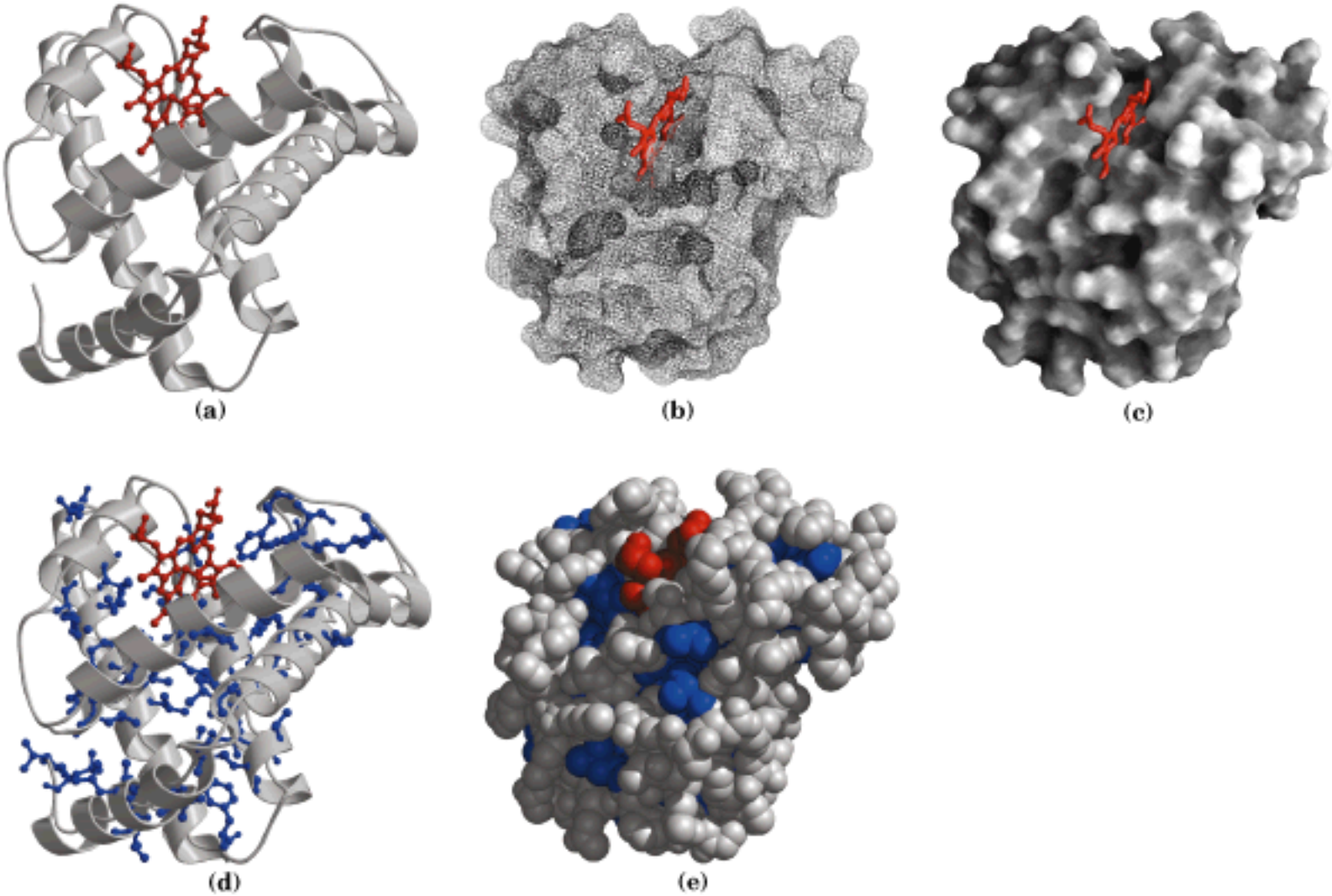


(b)

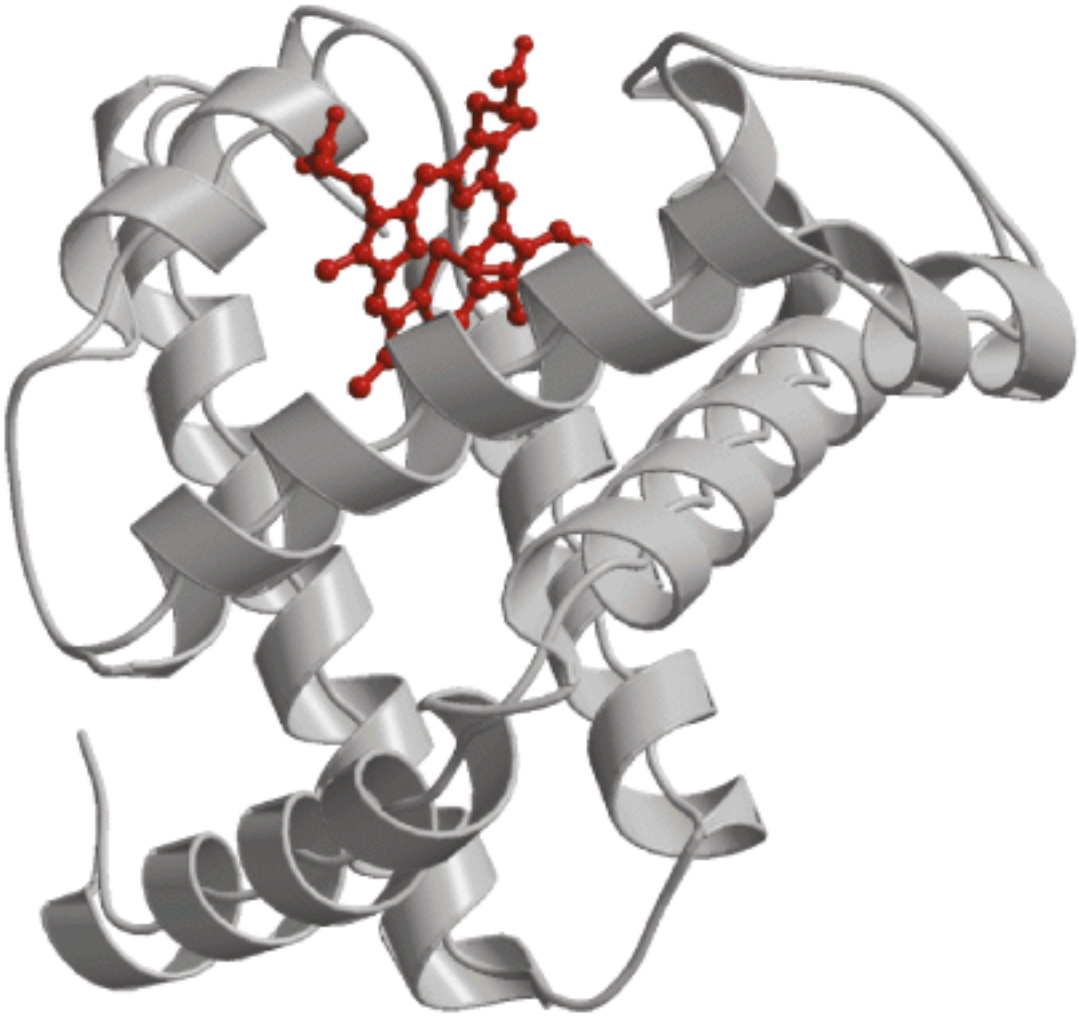
Fibroína (fibras azules) saliendo de las fúsulas de las arañas.



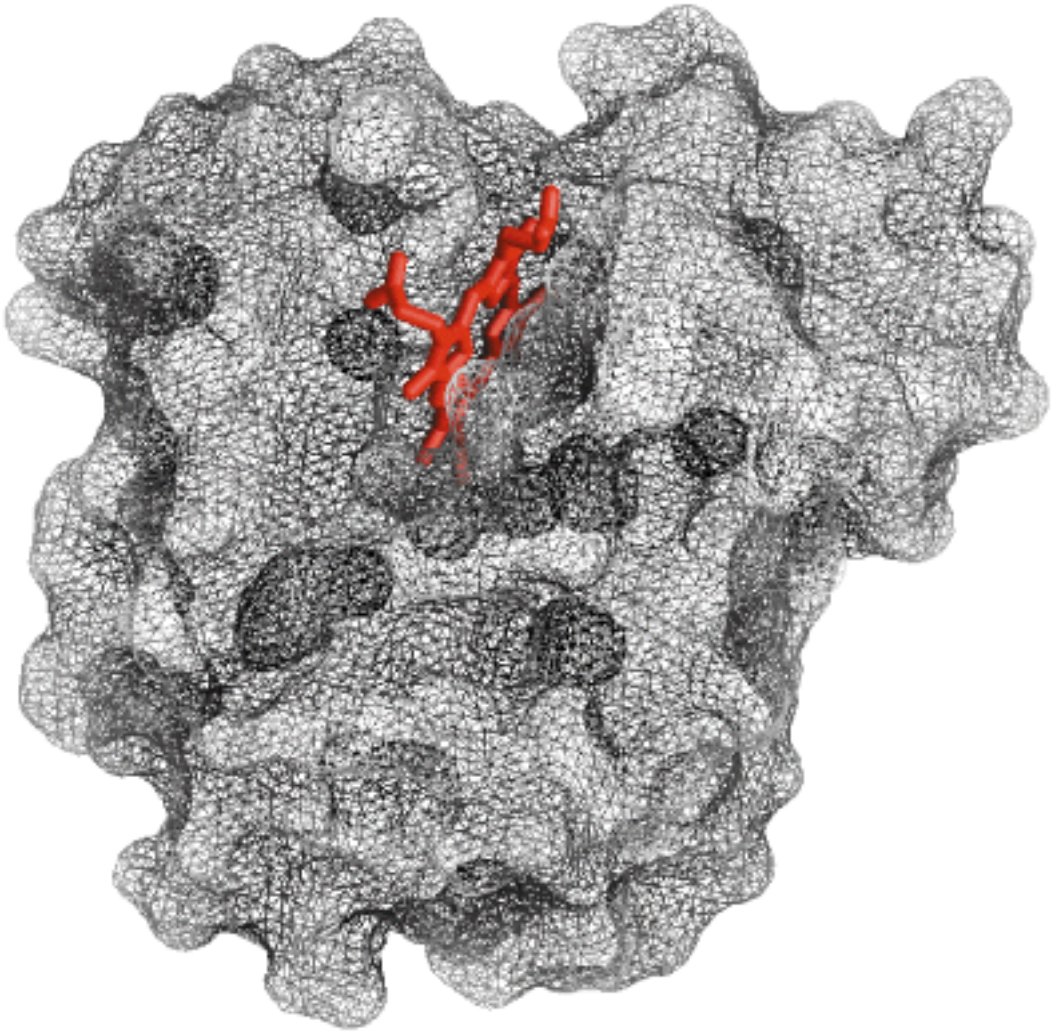
Albúmina Sérica Humana (M_r 64500)



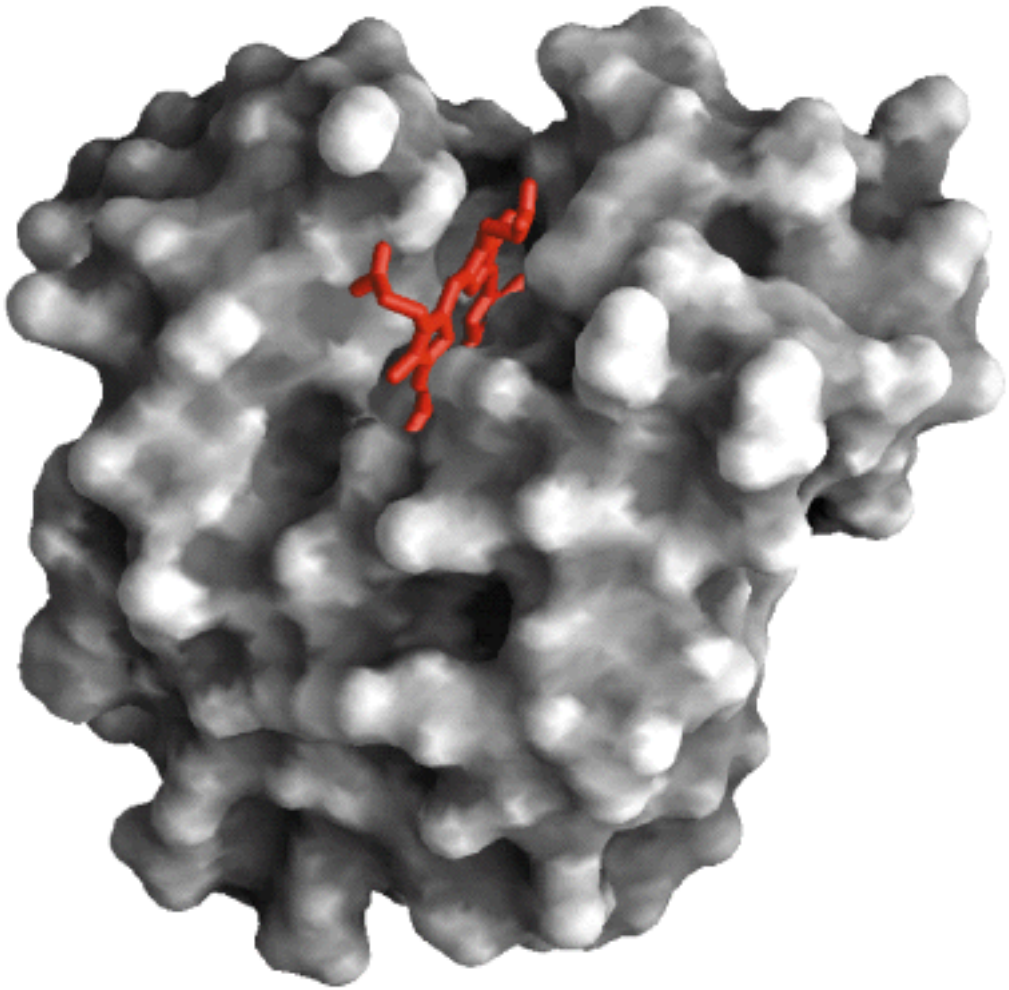
Estructura 3D de la mioglobina del cachalote.



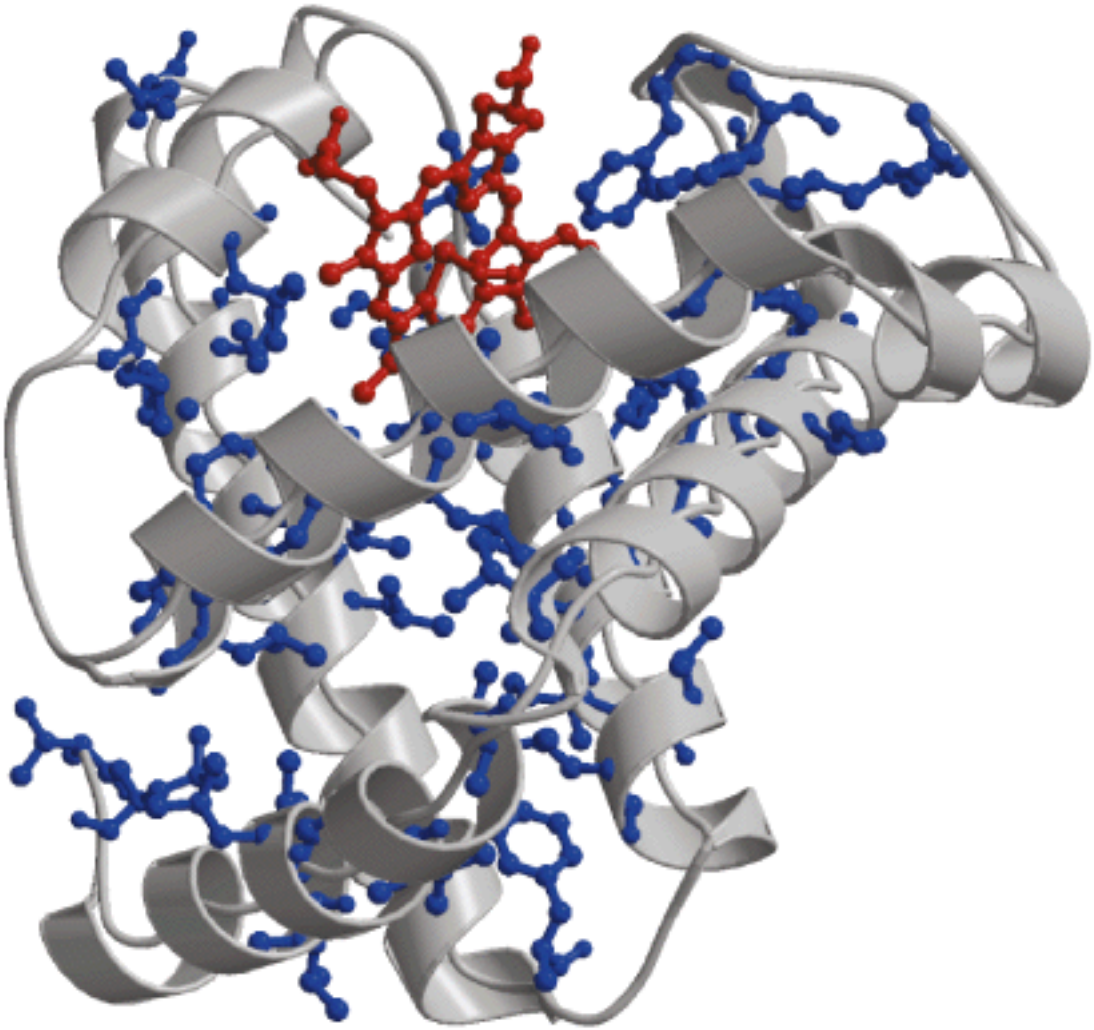
(a)



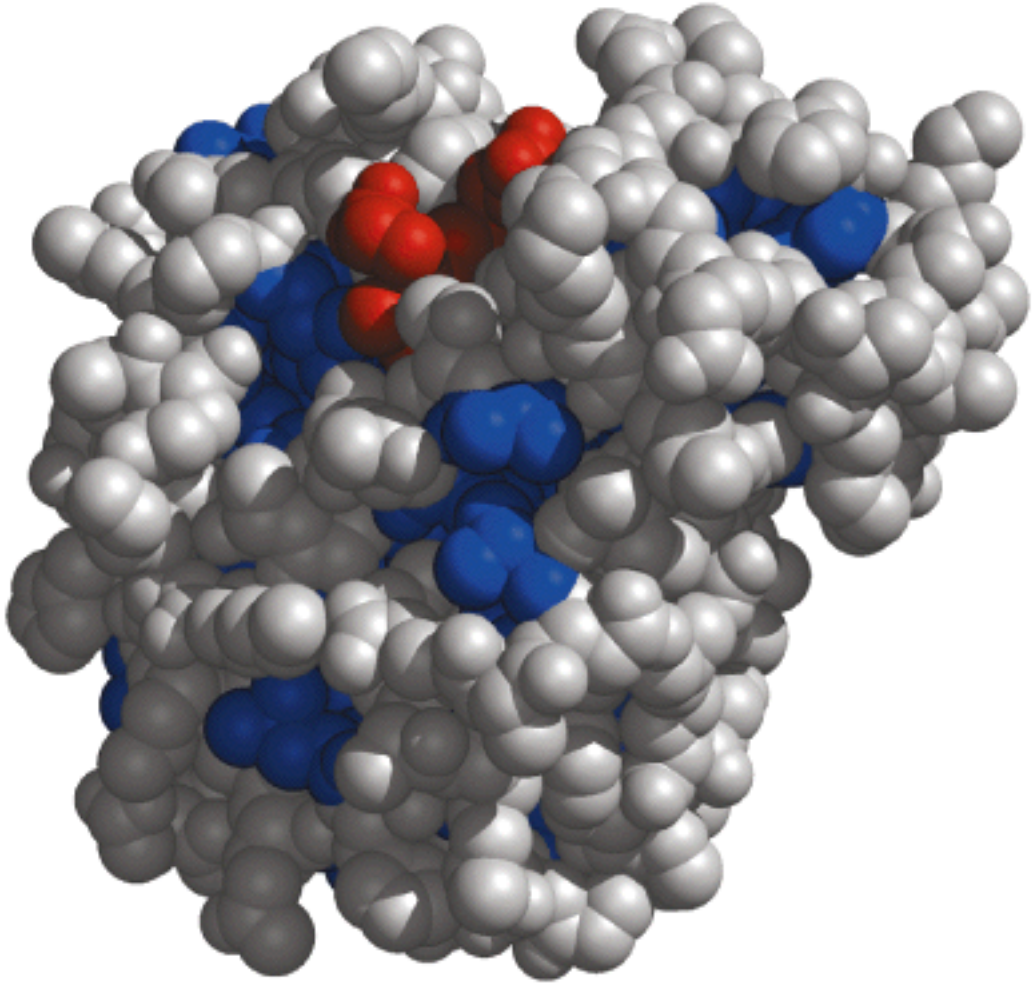
(b)



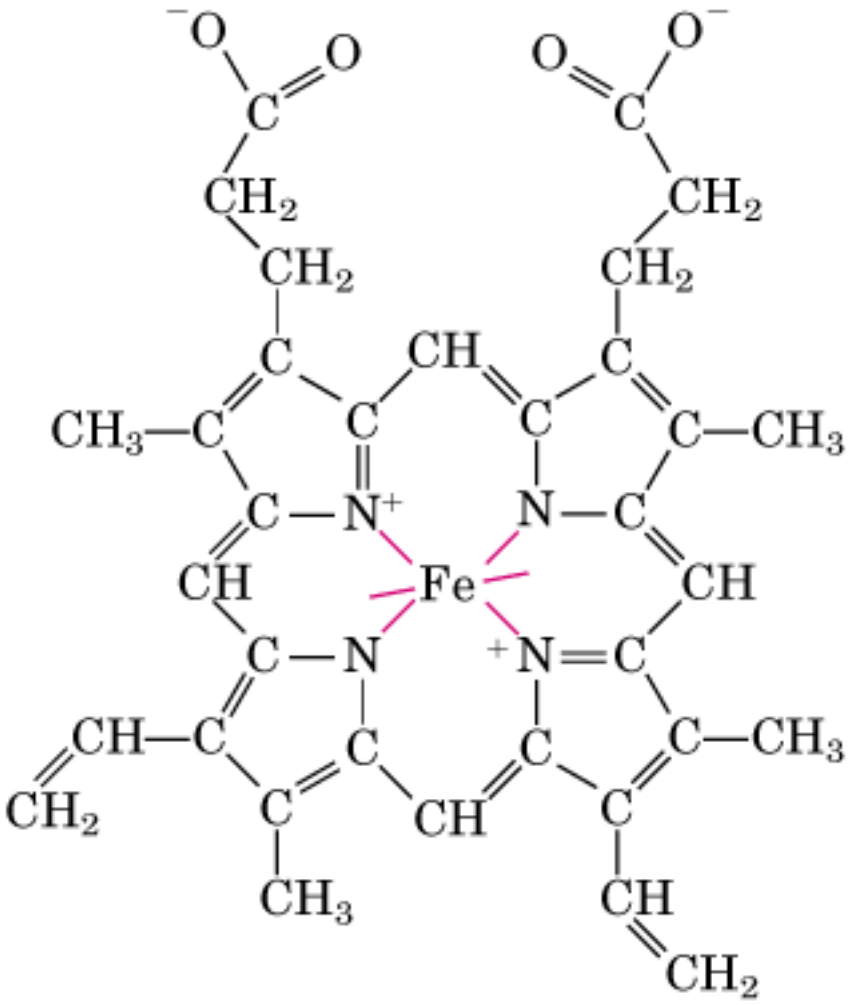
(c)



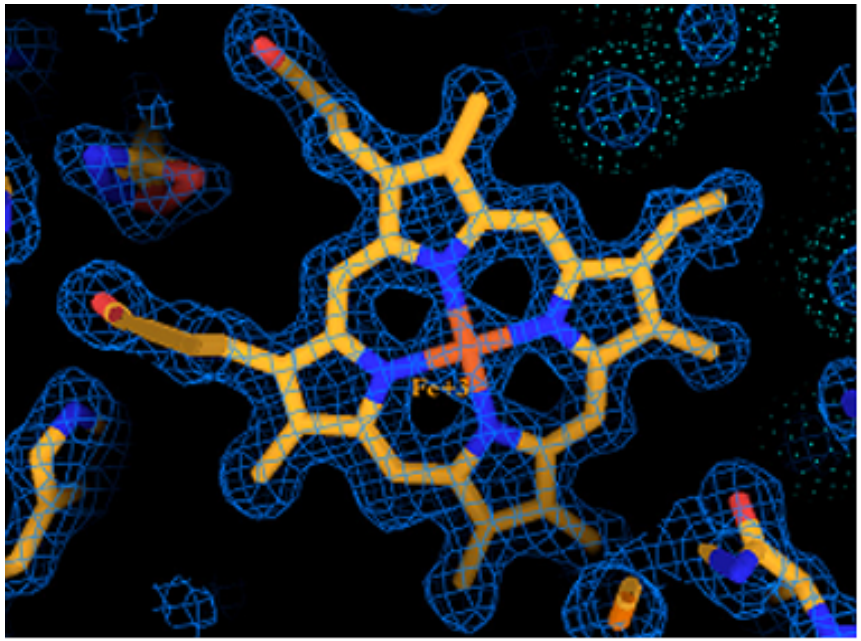
(d)



(e)

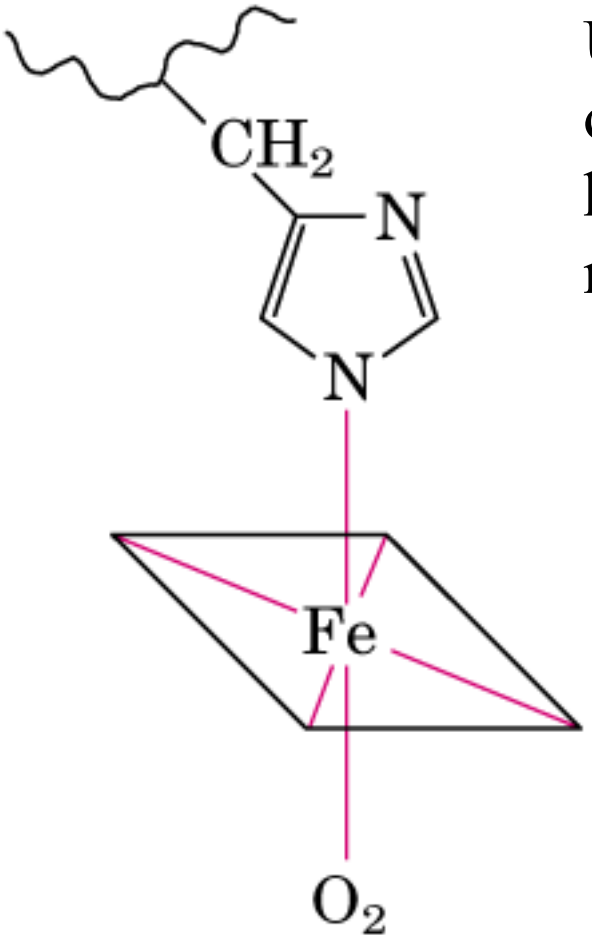


(a)



Grupo hemo en el Cyt C de corazón de bovino.

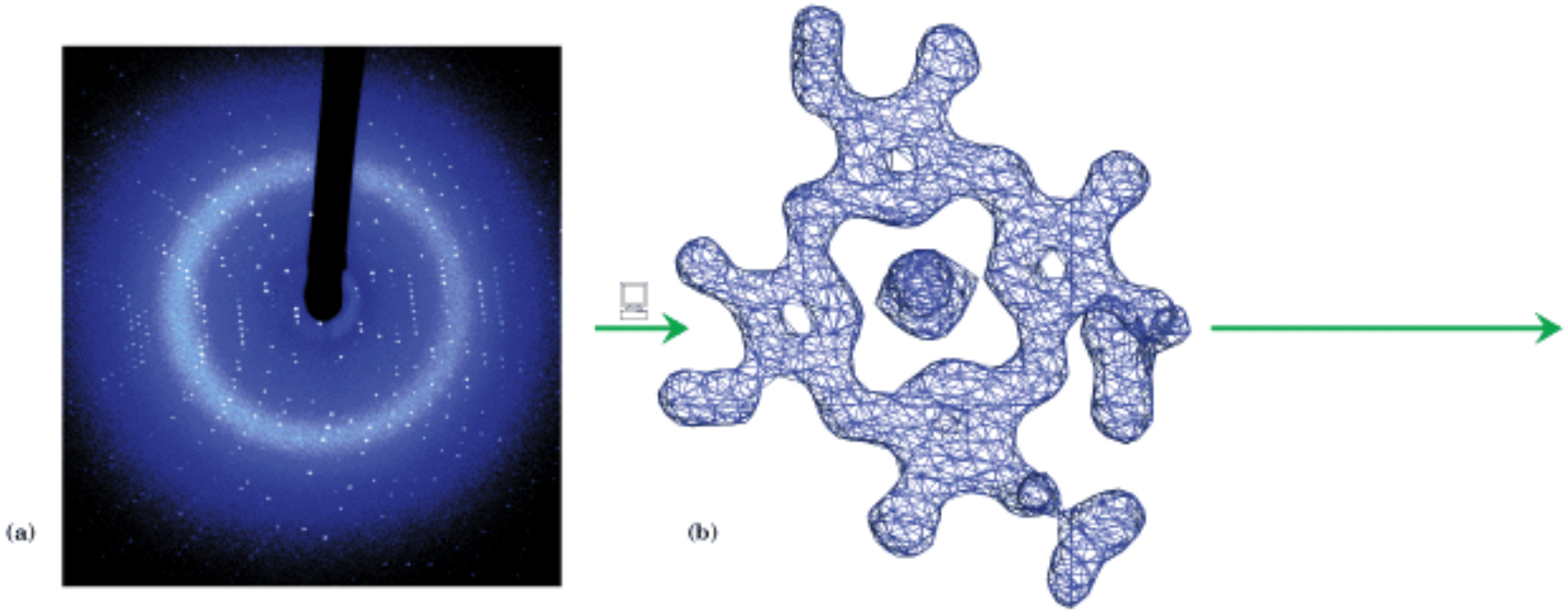
El grupo hemo en proteínas

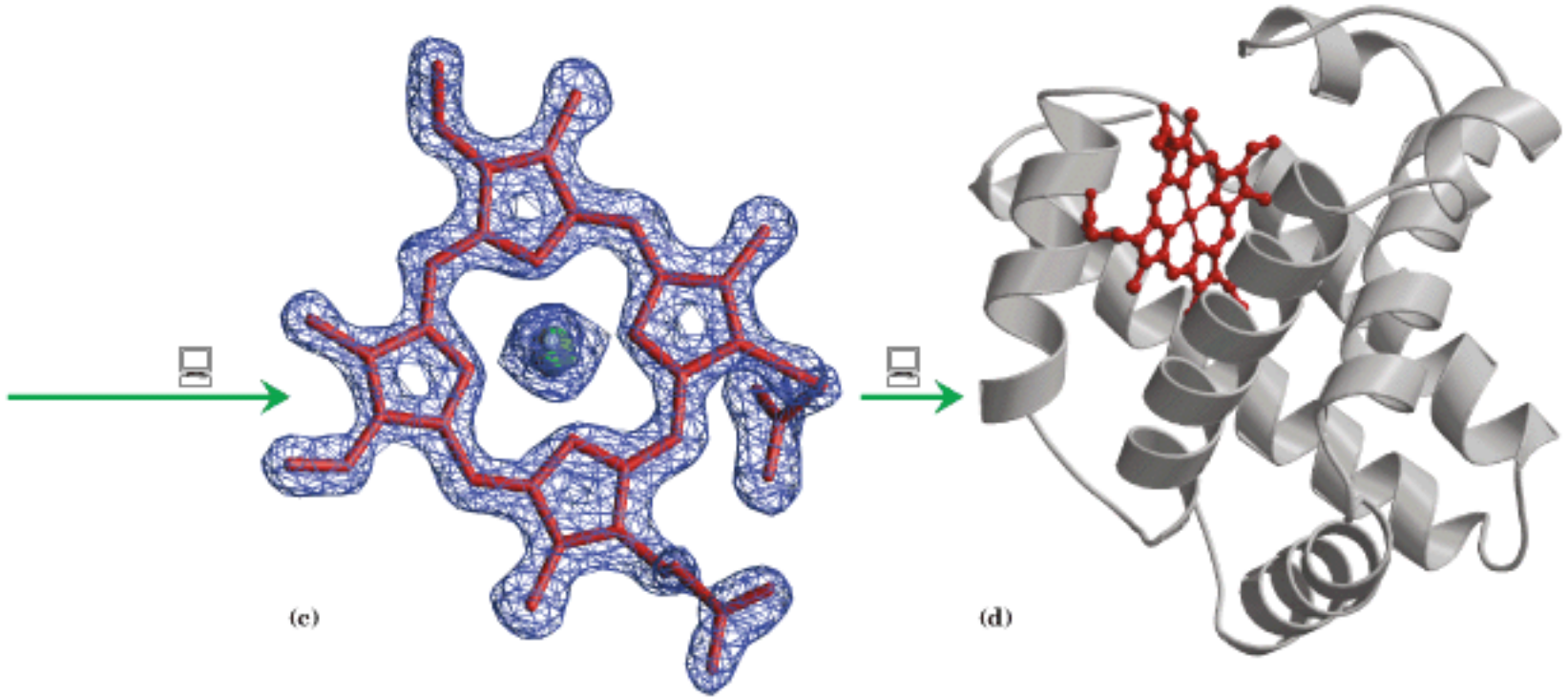


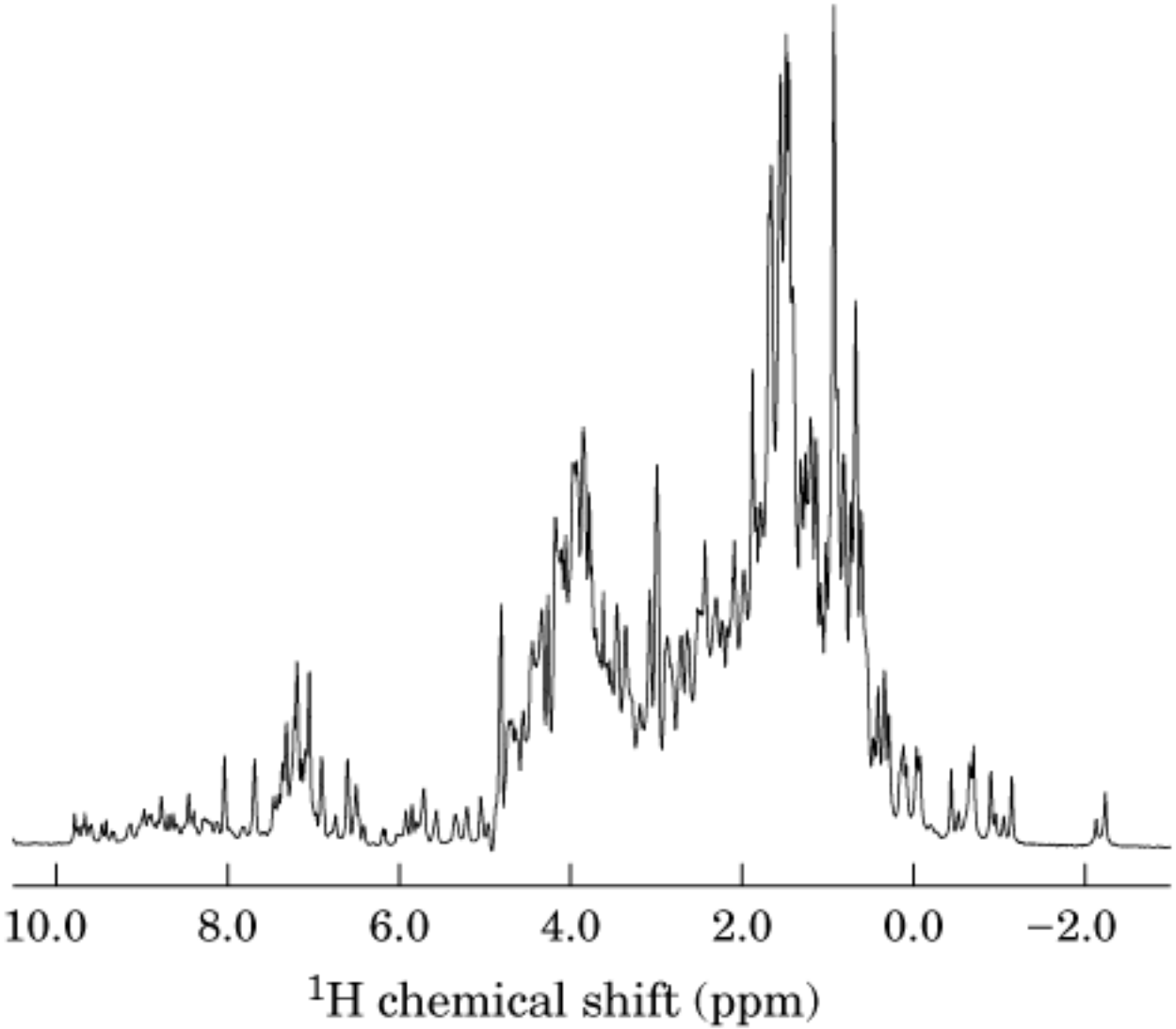
(b)

Unión de Fe y oxigeno en la hemoglobina y la mioglobina.

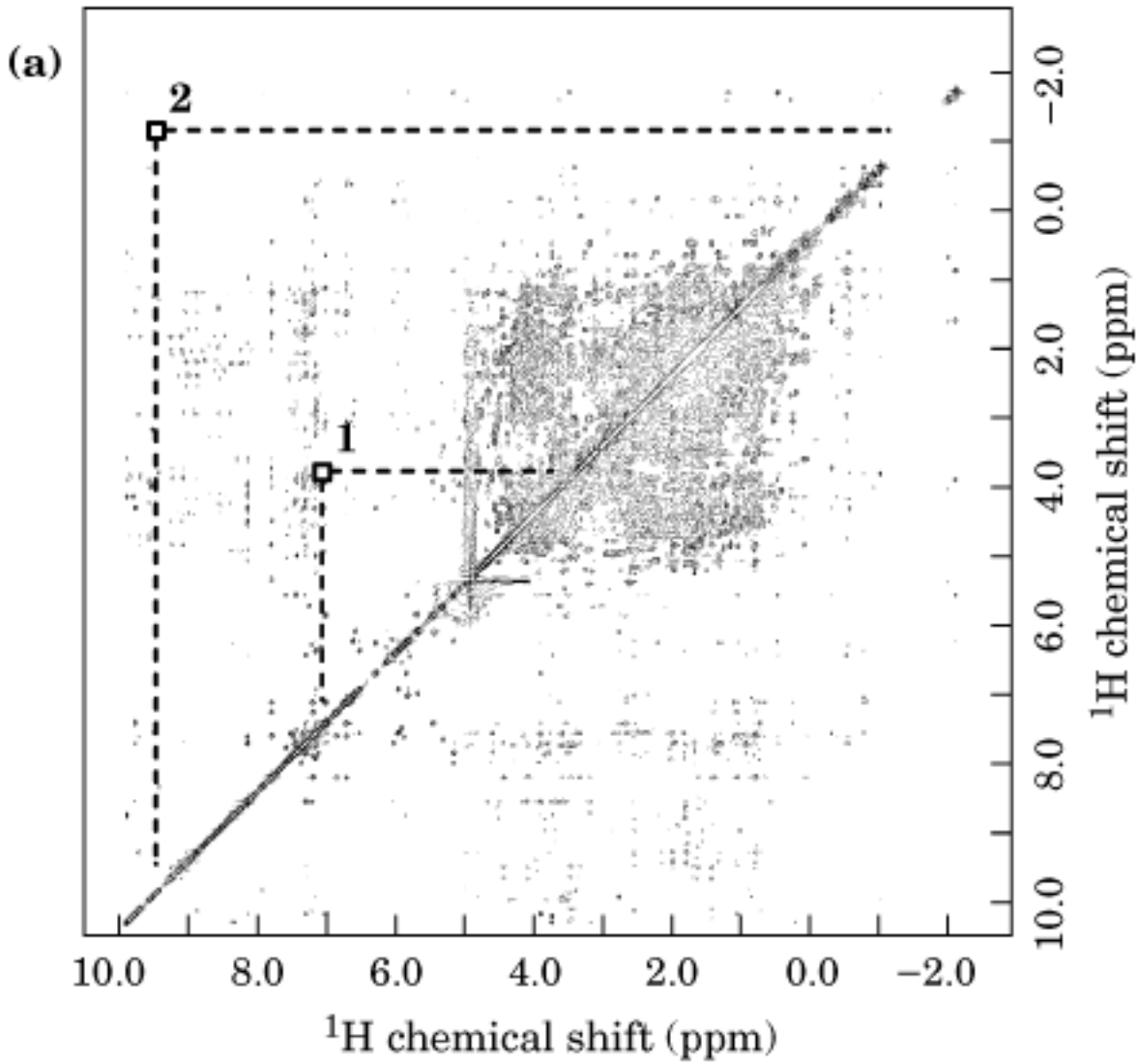
Métodos para determinar la estructura 3D de proteínas



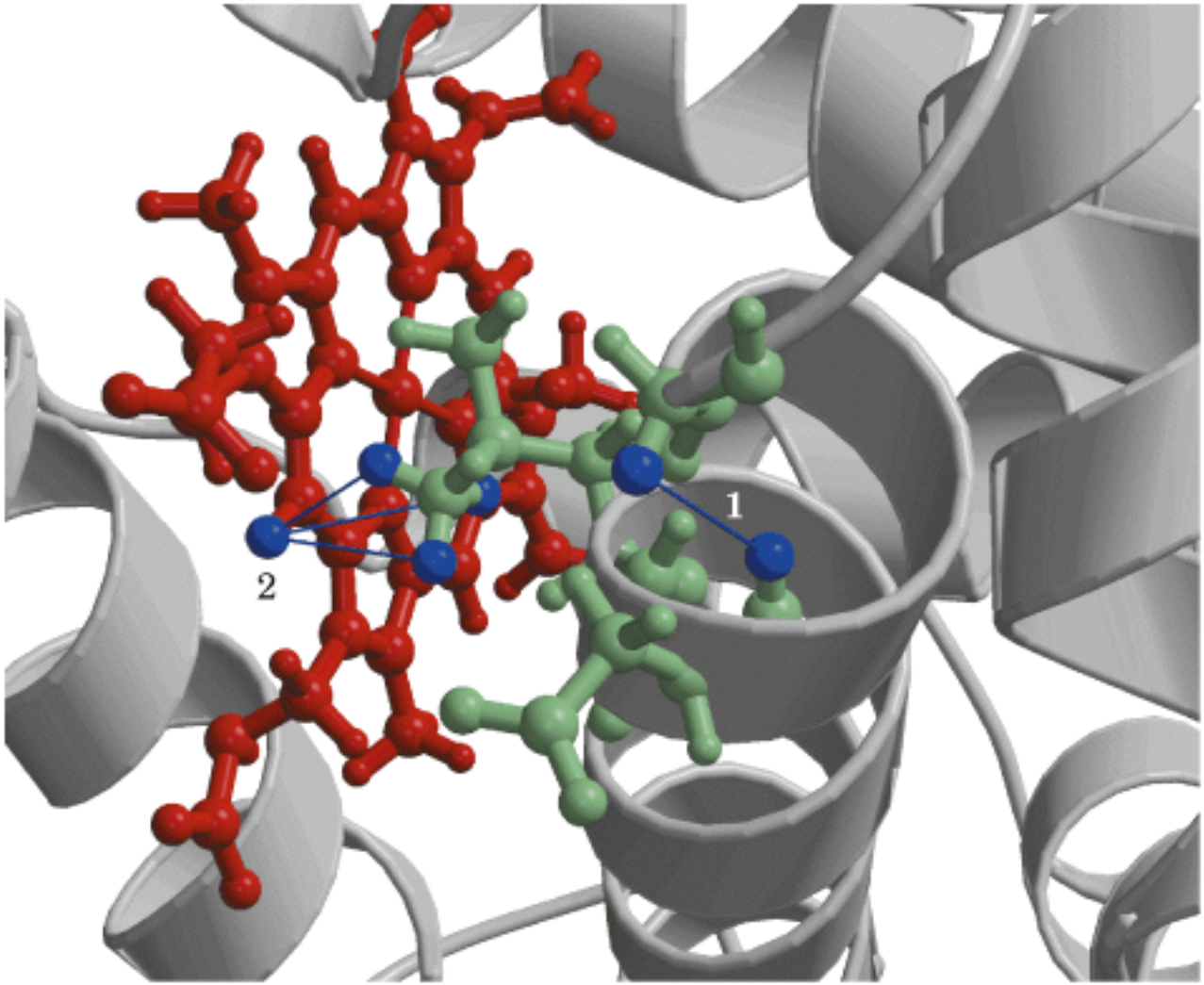

(c)
(d)



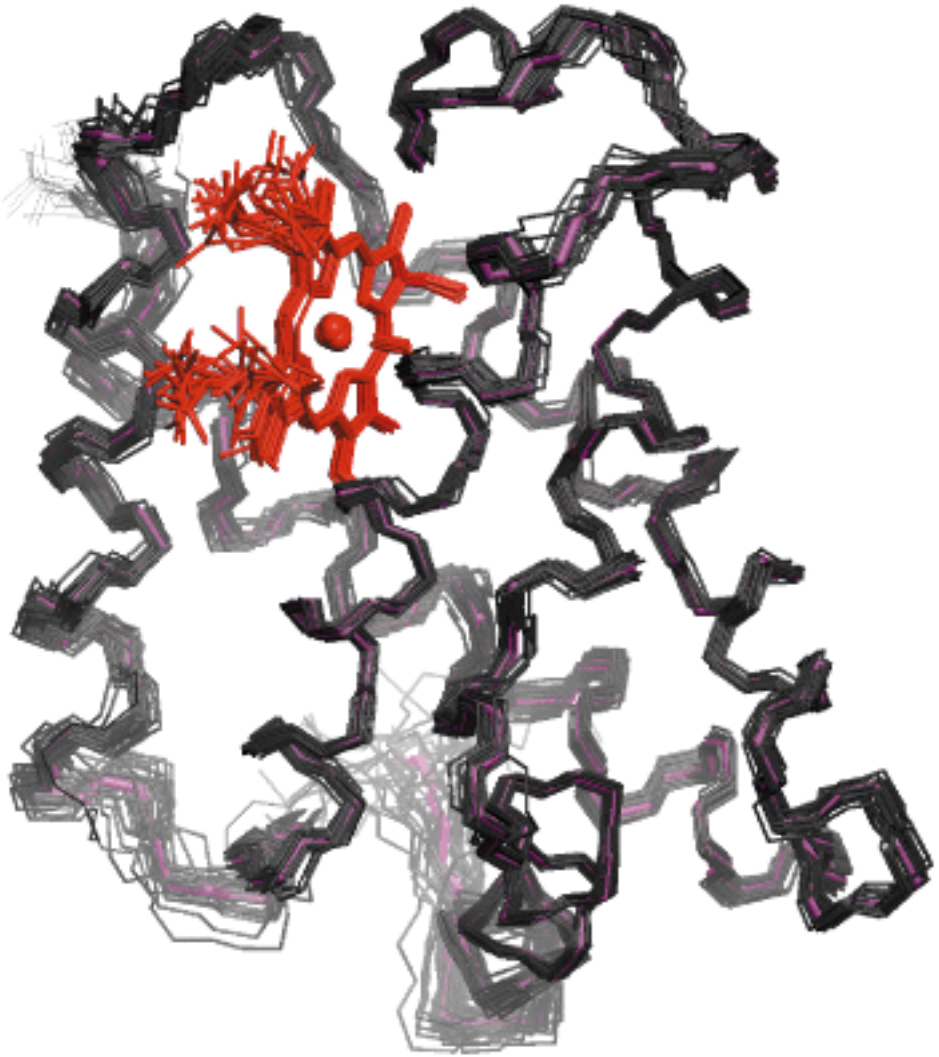
Espectro monodimensional del RMN



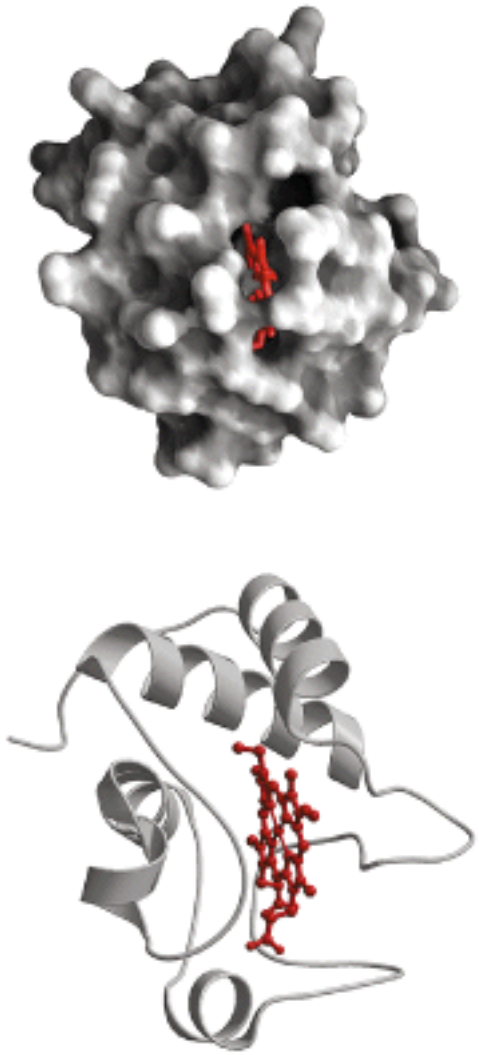
Espectro 2D de RMN para proteínas.



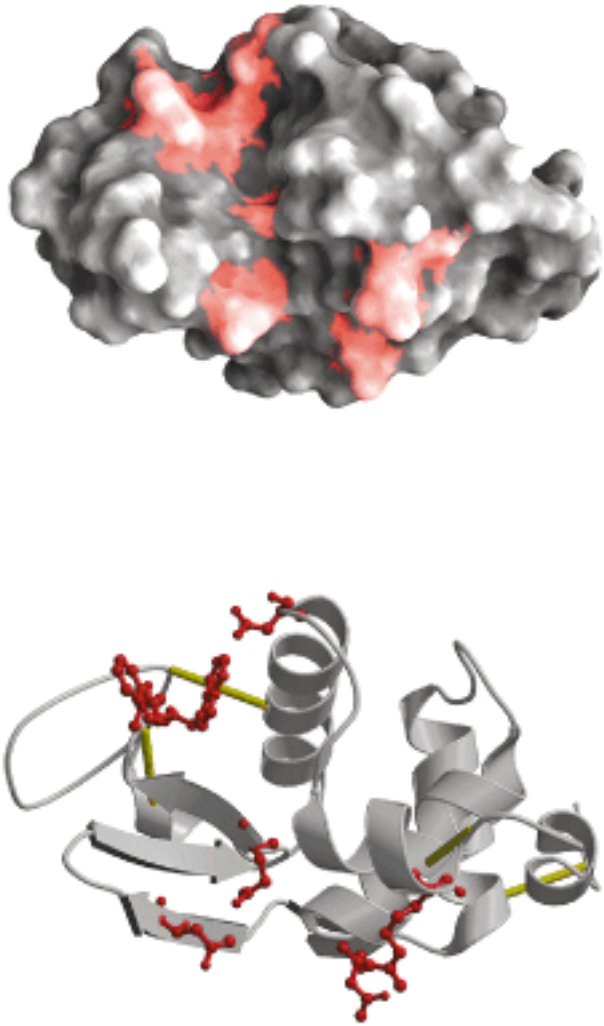
(b)



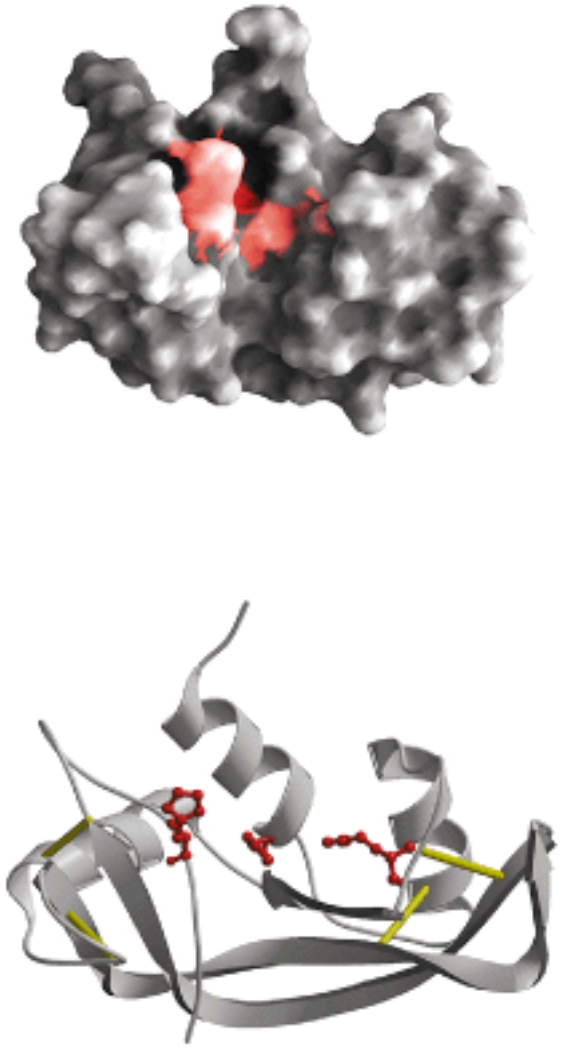
(c)



Cytochrome *c*



Lysozyme



Ribonuclease

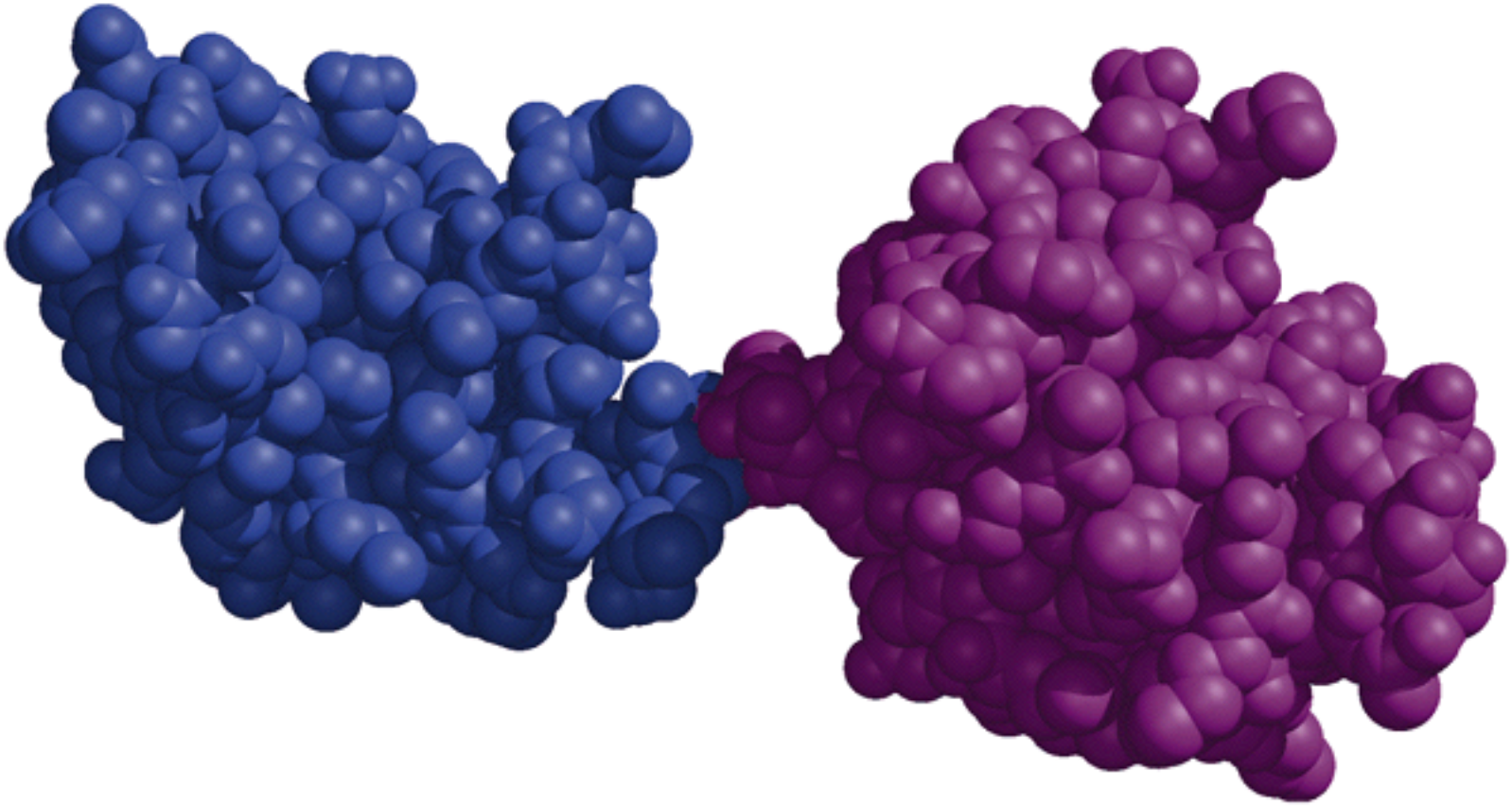
table 6–2

Approximate Amounts of α Helix and β Conformation in Some Single-Chain Proteins*

	Residues (%)	
Protein (total residues)	**α Helix**	**β Conformation**
Chymotrypsin (247)	14	45
Ribonuclease (124)	26	35
Carboxypeptidase (307)	38	17
Cytochrome *c* (104)	39	0
Lysozyme (129)	40	12
Myoglobin (153)	78	0

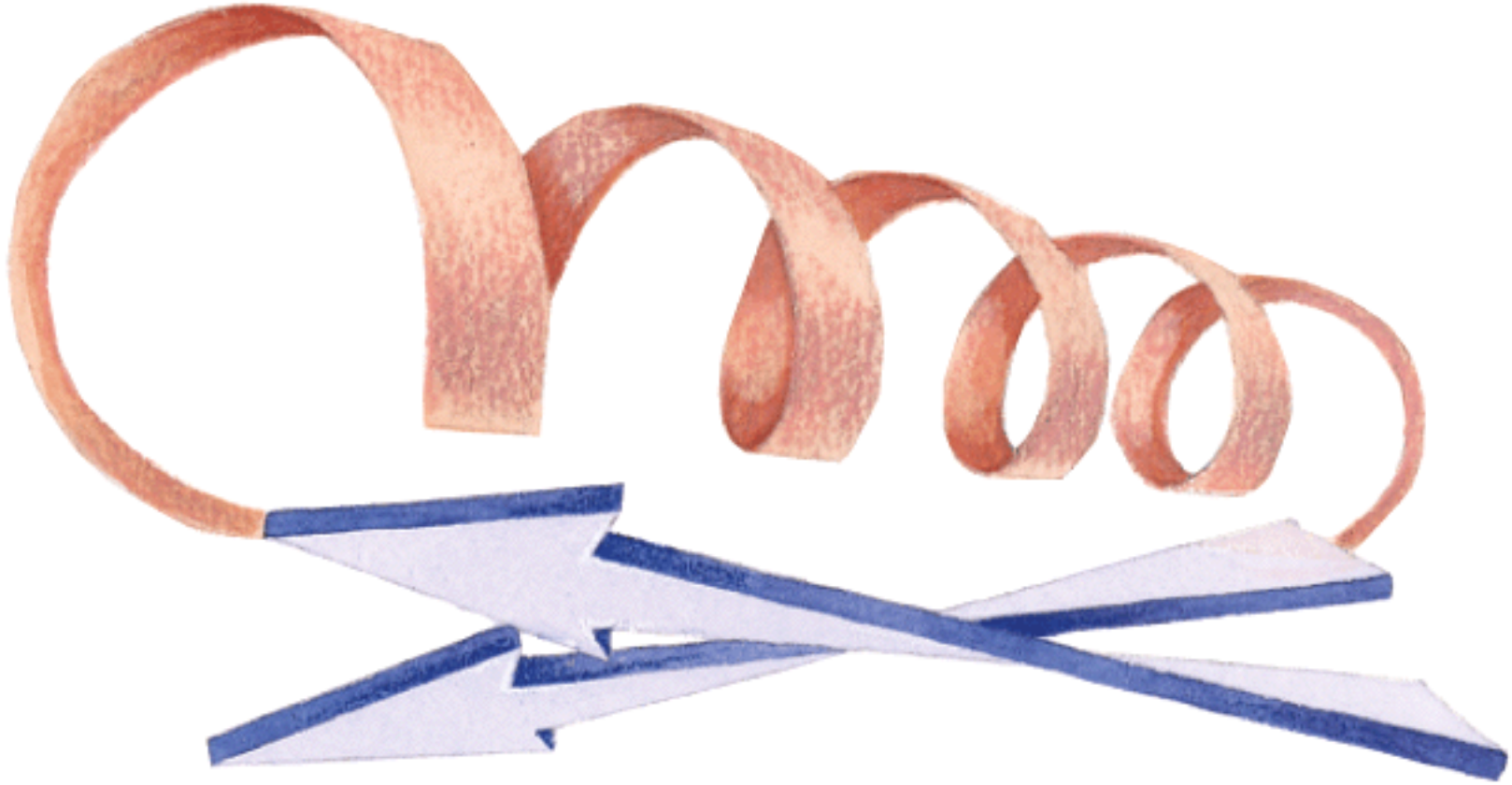
Source: Data from Cantor, C.R. & Schimmel, P.R. (1980) *Biophysical Chemistry,* Part I: *The Conformation of Biological Macromolecules,* p. 100, W.H. Freeman and Company, New York.

*Portions of the polypeptide chains that are not accounted for by α helix or β conformation consist of bends and irregularly coiled or extended stretches. Segments of α helix and β conformation sometimes deviate slightly from their normal dimensions and geometry.

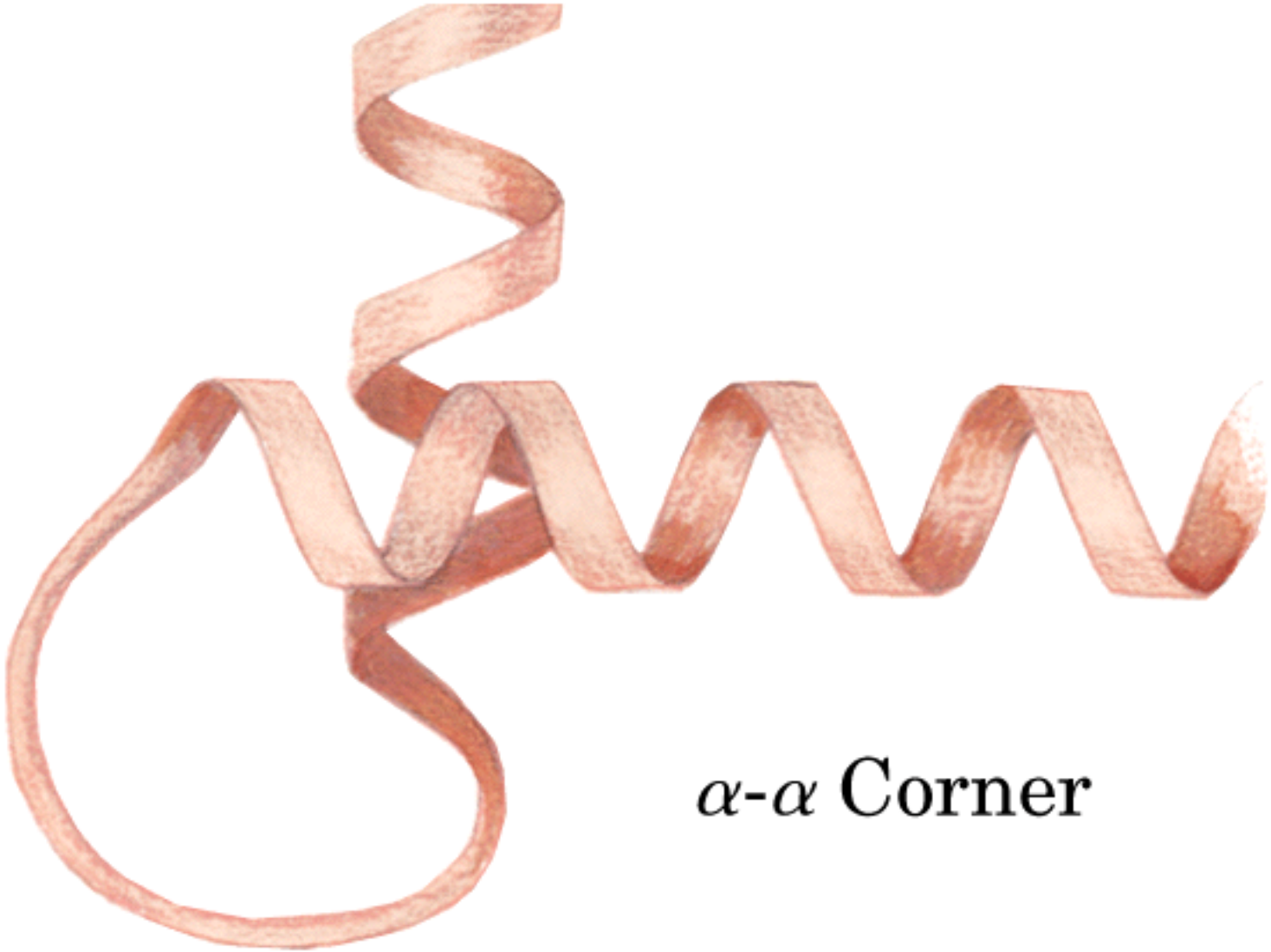


Dominios estructurales en el polipéptido troponina C

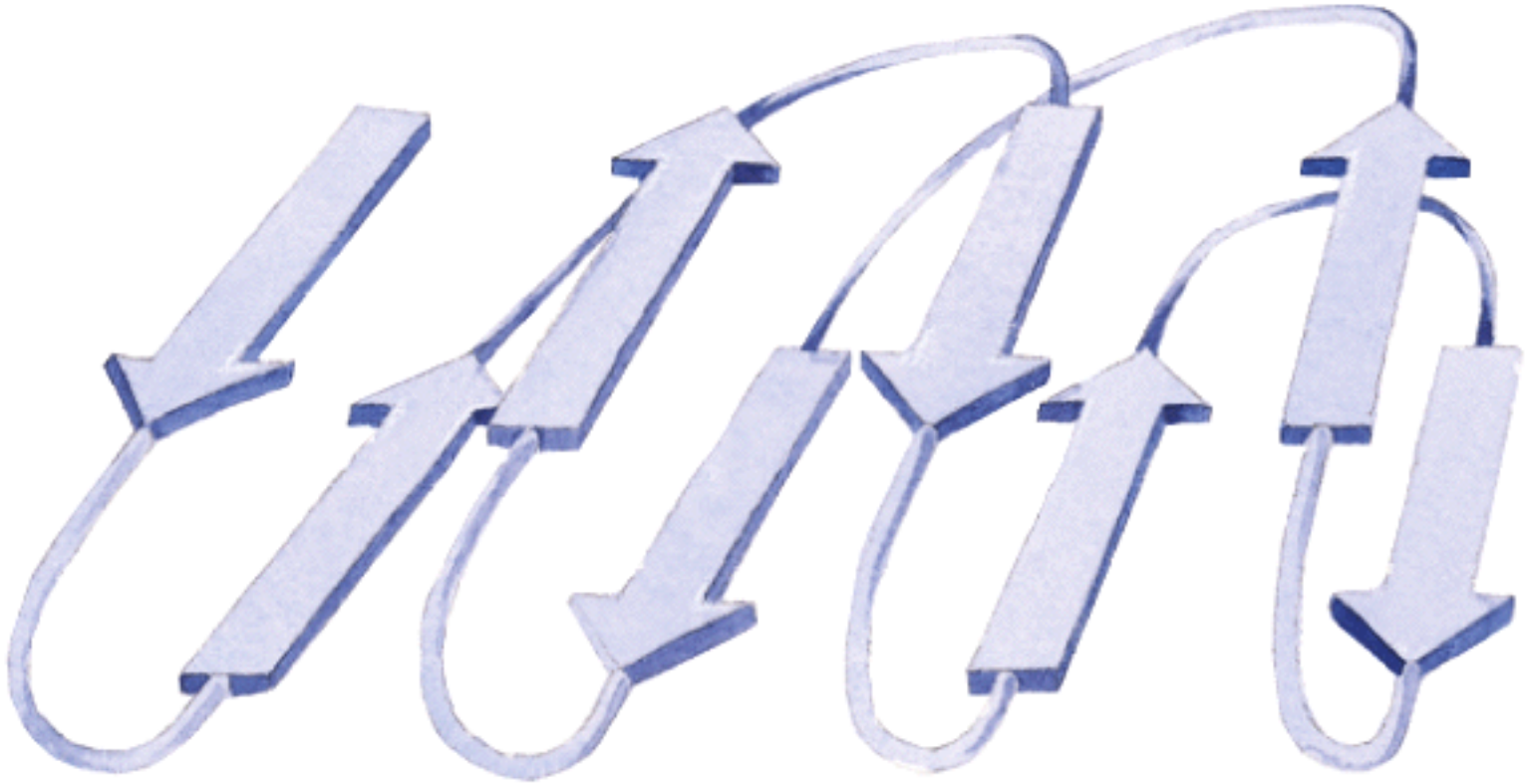
Patrones estables de plegamiento en proteinas



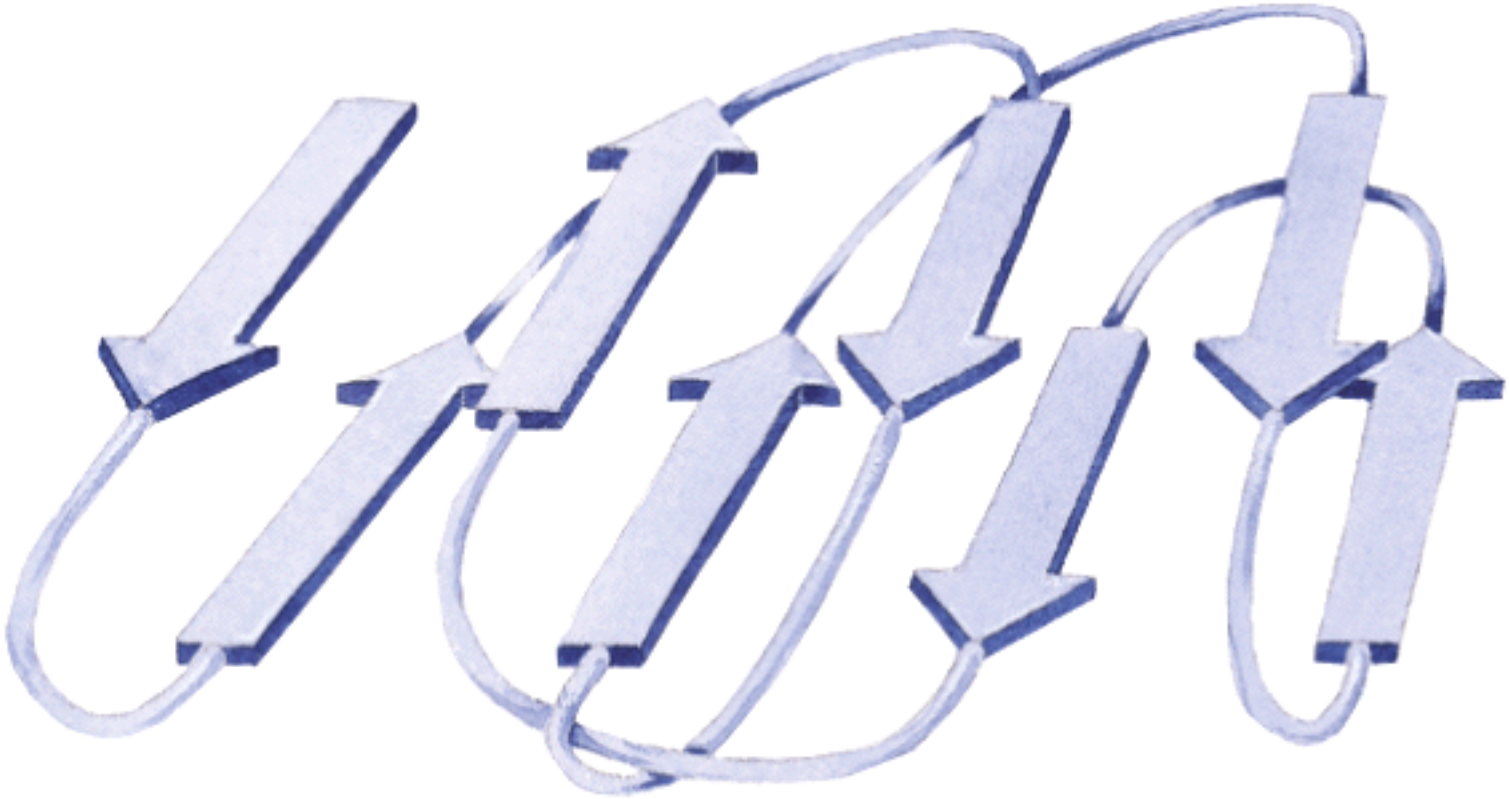
(a) β-α-β Loop



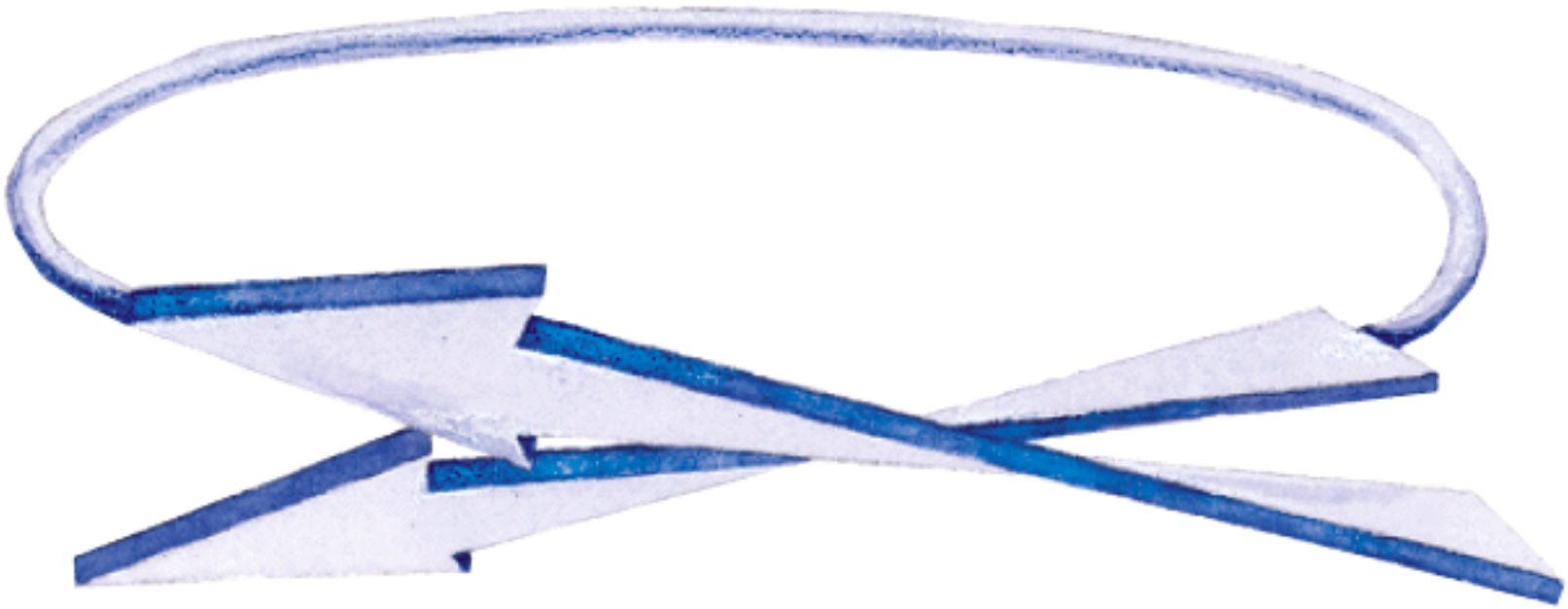
α-α Corner



(b) Typical connections in an all-β motif



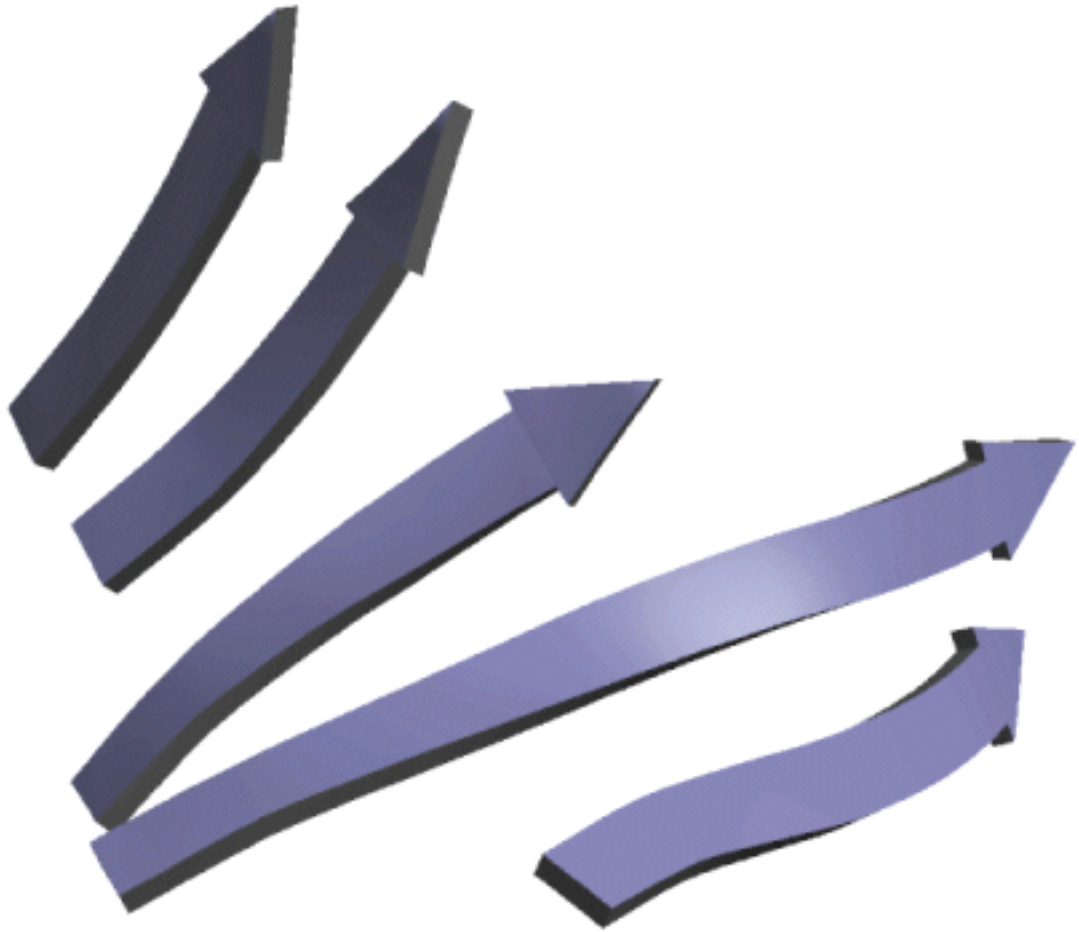
Crossover connection
(not observed)



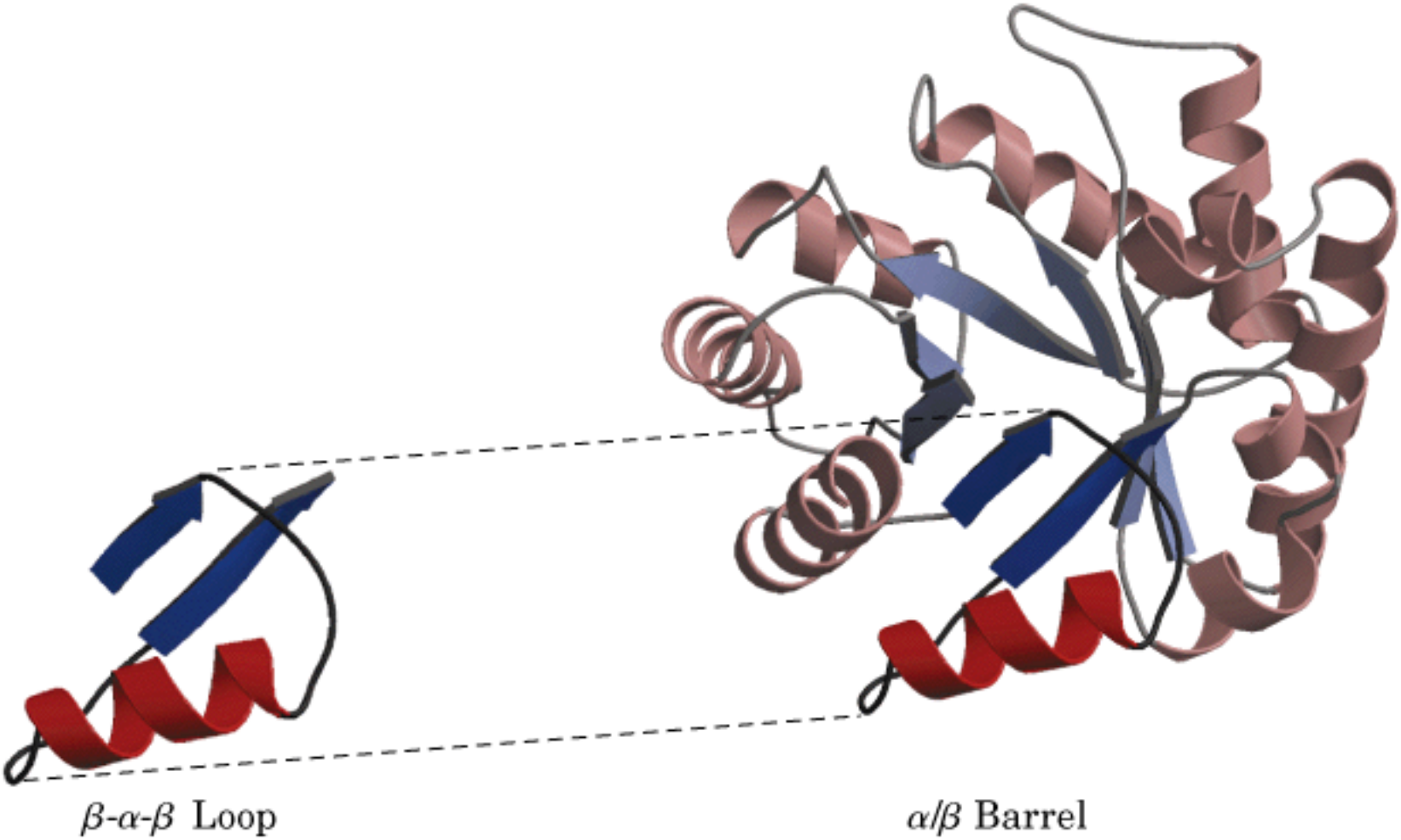
(c) Right-handed connection between β strands

Left-handed connection
between β strands
(very rare)

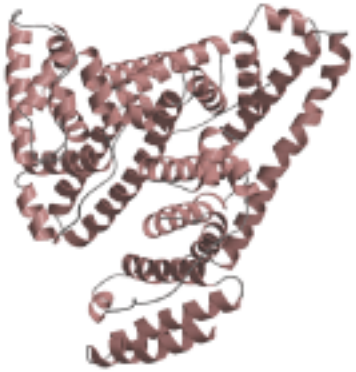
(d) β Barrel



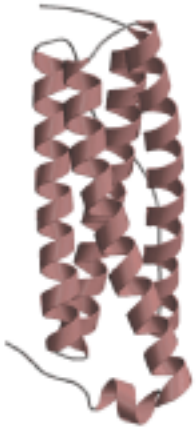
Twisted β sheet

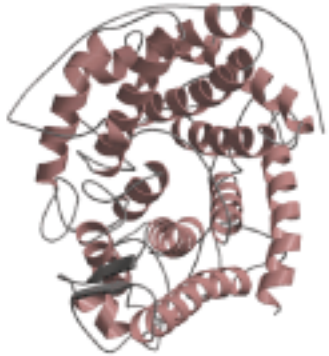

β-α-β Loop
α/β Barrel

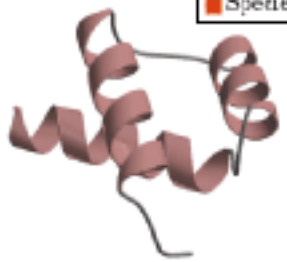
All α



1a06	1bcf	1gai	1enh
Serum albumin	Ferritin-like	α/α toroid	DNA-binding 3-helical bundle
Serum albumin	Ferritin-like	Glycosyltransferases of the superhelical fold	Homeodomain-like
Serum albumin	Ferritin	Glucoamylase	Homeodomain
Serum albumin	Bacterioferritin (cytochrome b_1)	Glucoamylase	*engrailed* Homeodomain
Human (*Homo sapiens*)	*Escherichia coli*	*Aspergillus awamori*, variant x100	*Drosophila melanogaster*

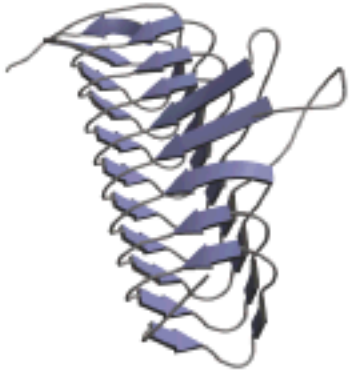




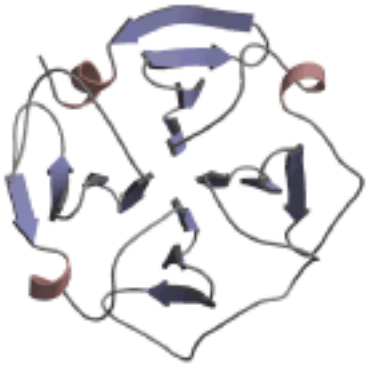


All β

1hoe
α-Amylase inhibitor
α-Amylase inhibitor
α-Amylase inhibitor
HOE-467A
Streptomyces tendae 4158



1lxe
Single-stranded left-handed β helix
Trimeric LpxA-like enzymes
UDP *N*-acetylglucosamine acyltransferase
UDP *N*-acetylglucosamine acyltransferase
Escherichia coli

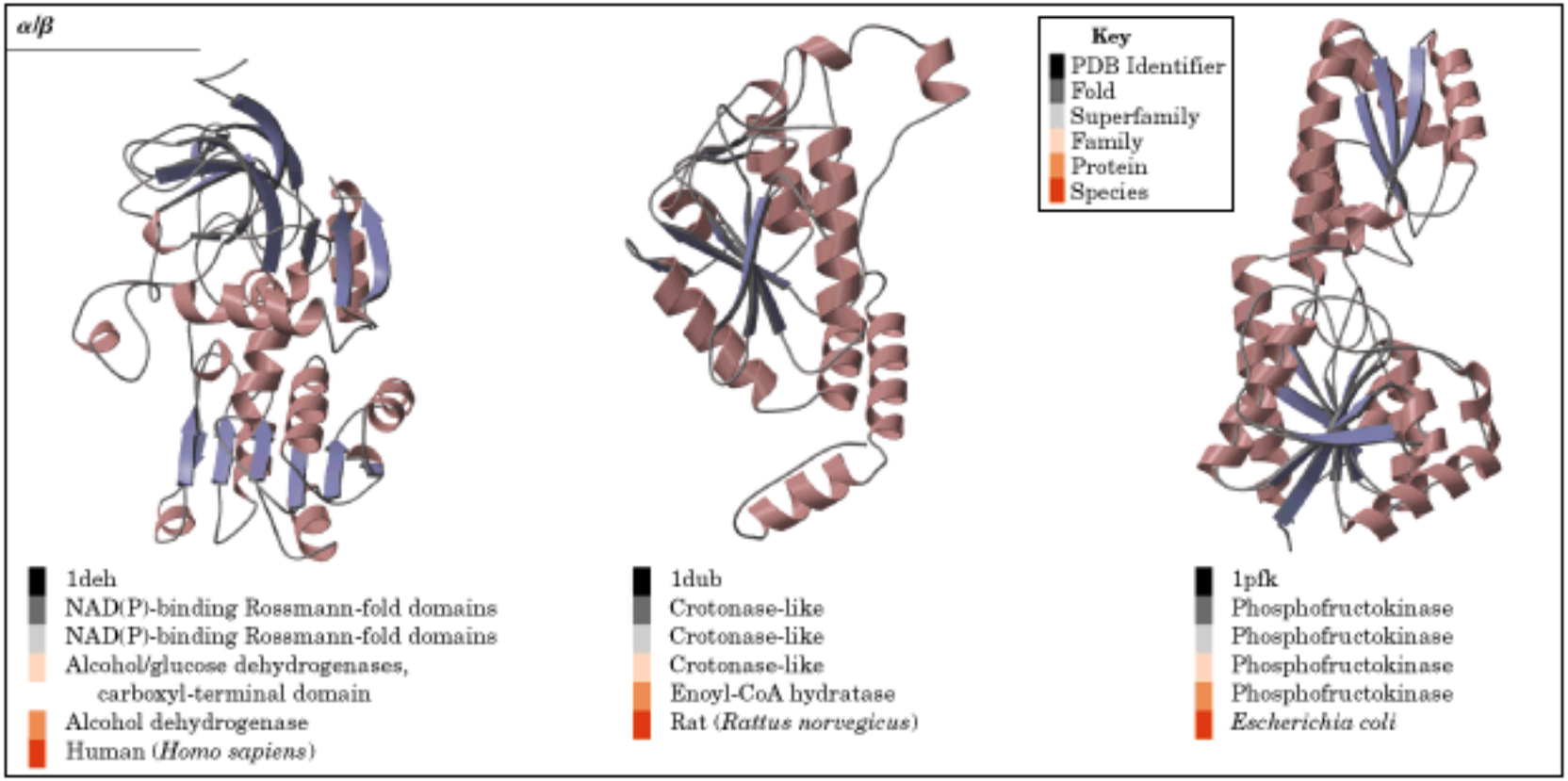


1pex
Four-bladed β propeller
Hemopexin-like domain
Hemopexin-like domain
Collagenase-3 (MMP-13), carboxyl-terminal domain
Human (*Homo sapiens*)

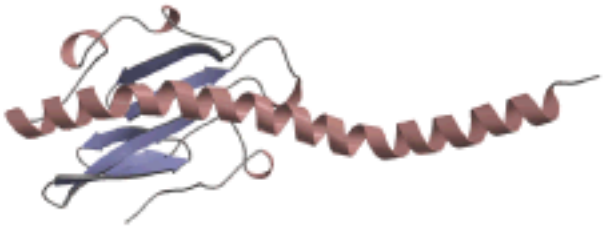
1jpc
β-Prism II
α-D-Mannose-specific plant lectins
α-D-Mannose-specific plant lectins
Lectin (agglutinin)
Snowdrop (*Galanthus nivalis*)

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

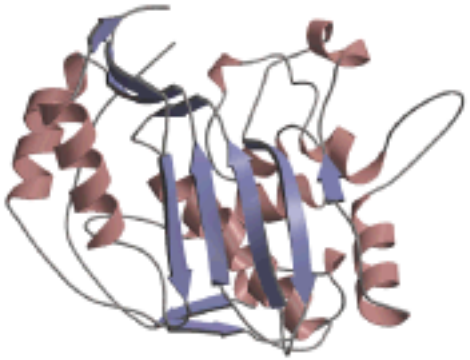
Key
PDB Identifier
Fold
Superfamily
Family
Protein
Species


α/β
Key
PDB Identifier
Fold
Superfamily
Family
Protein
Species
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
Crotonase-like
Crotonase-like
Crotonase-like
Enoyl-CoA hydratase
Rat (Rattus norvegicus)
1pfk
Phosphofructokinase
Phosphofructokinase
Phosphofructokinase
Phosphofructokinase
Escherichia coli

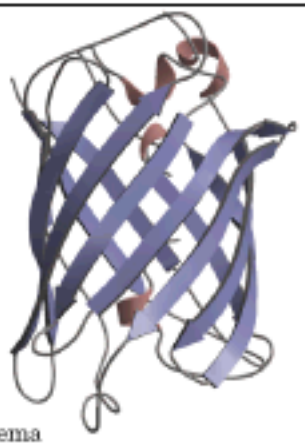
$\alpha + \beta$



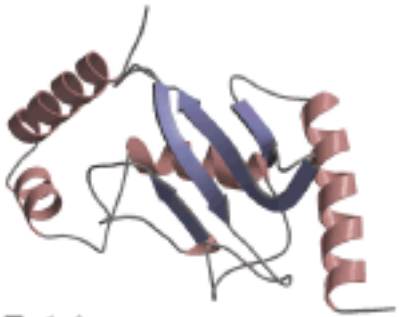
2pil
Pilin
Pilin
Pilin
Pilin
Neisseria gonorrhoeae



1syn
Thymidylate synthase
Thymidylate synthase
Thymidylate synthase
Thymidylate synthase
Escherichia coli

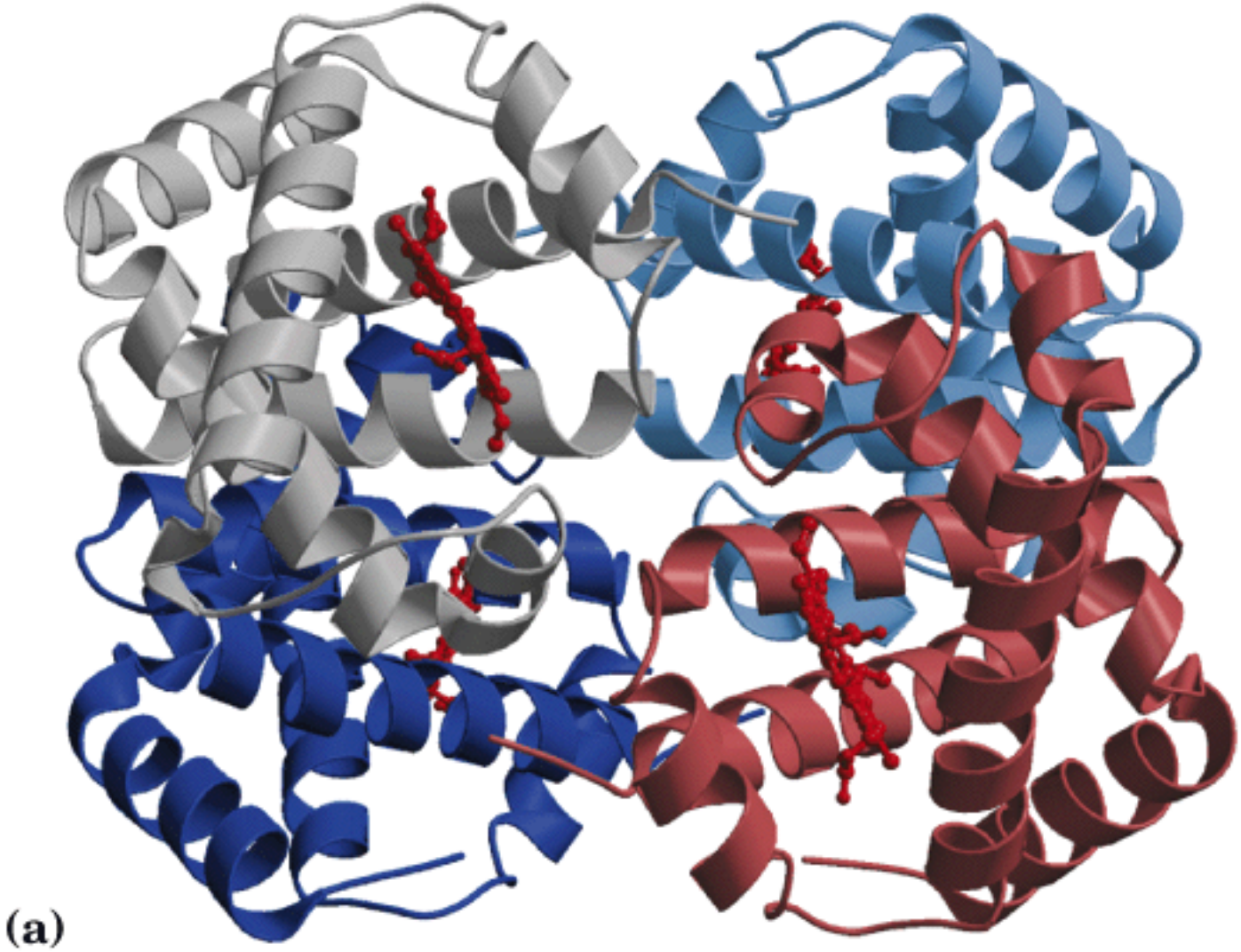


1ema
Green fluorescent protein
Green fluorescent protein
Green fluorescent protein
Green fluorescent protein
Jellyfish (*Aequorea victoria*)

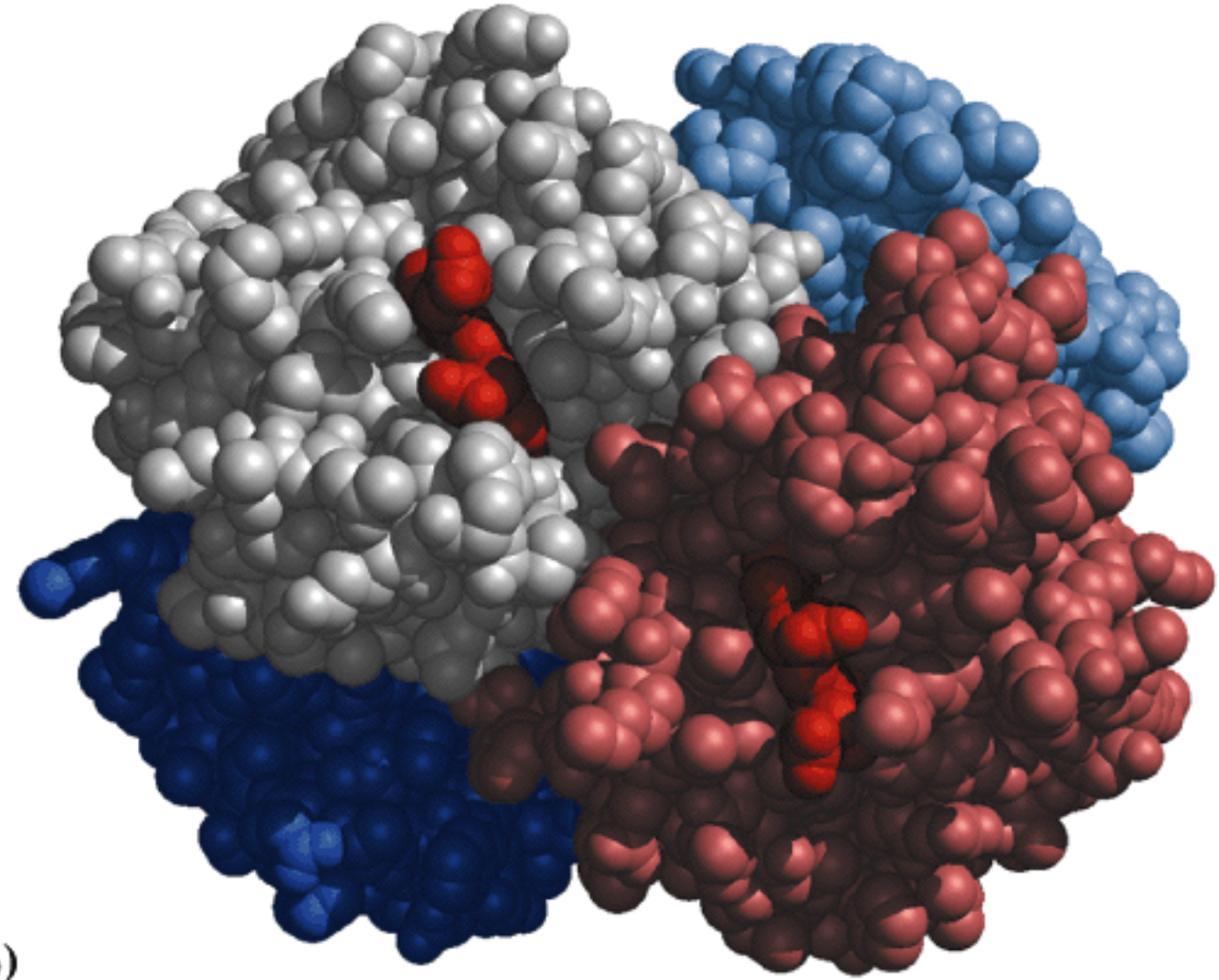


1u9a
Ubiquitin-conjugating enzyme
Ubiquitin-conjugating enzyme
Ubibuitin-conjugating enzyme
Ubiquitin-conjugating enzyme
Human (*Homo sapiens*)

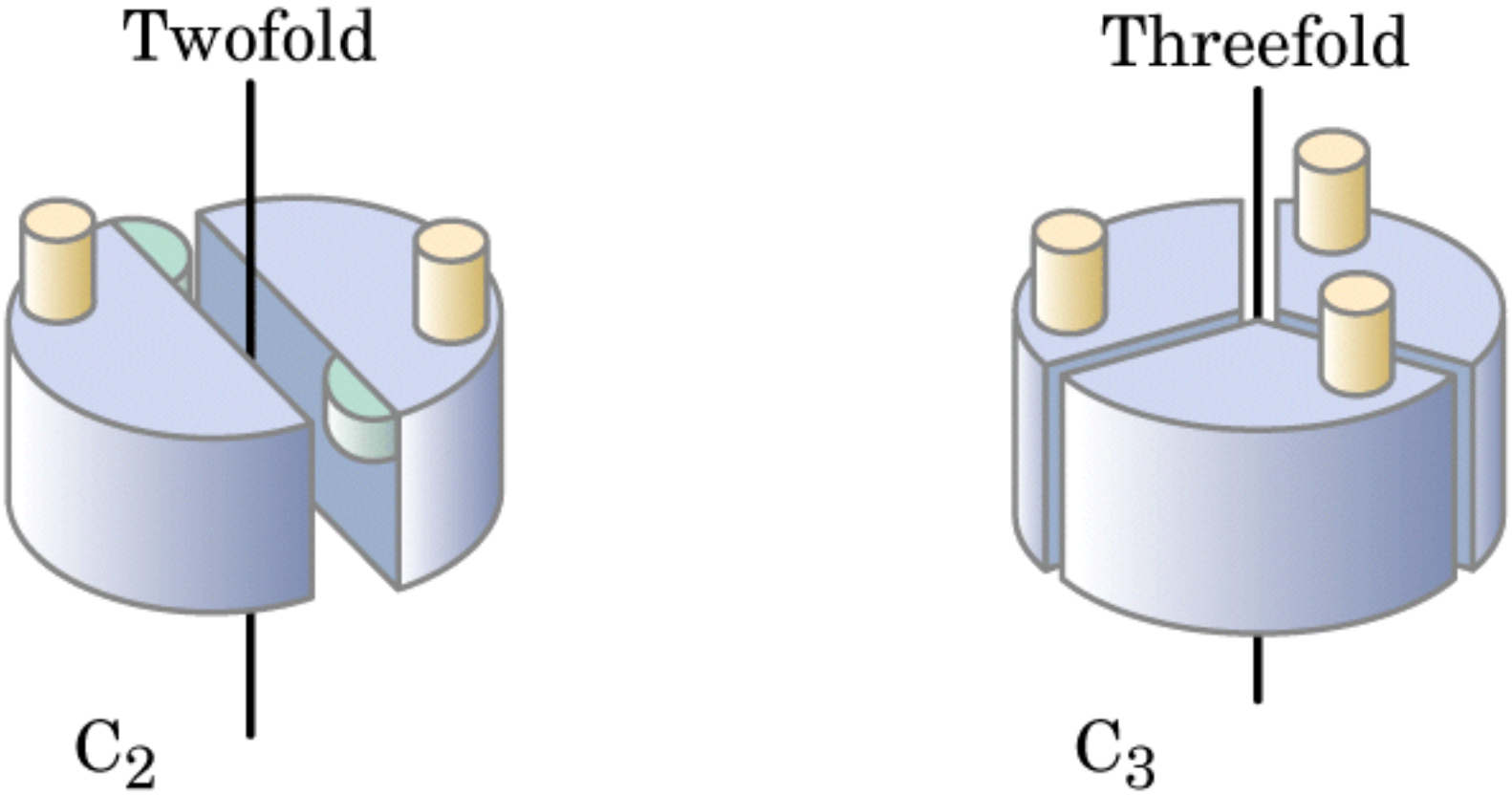
Key
PDB identifier
Fold
Superfamily
Family
Protein
Species



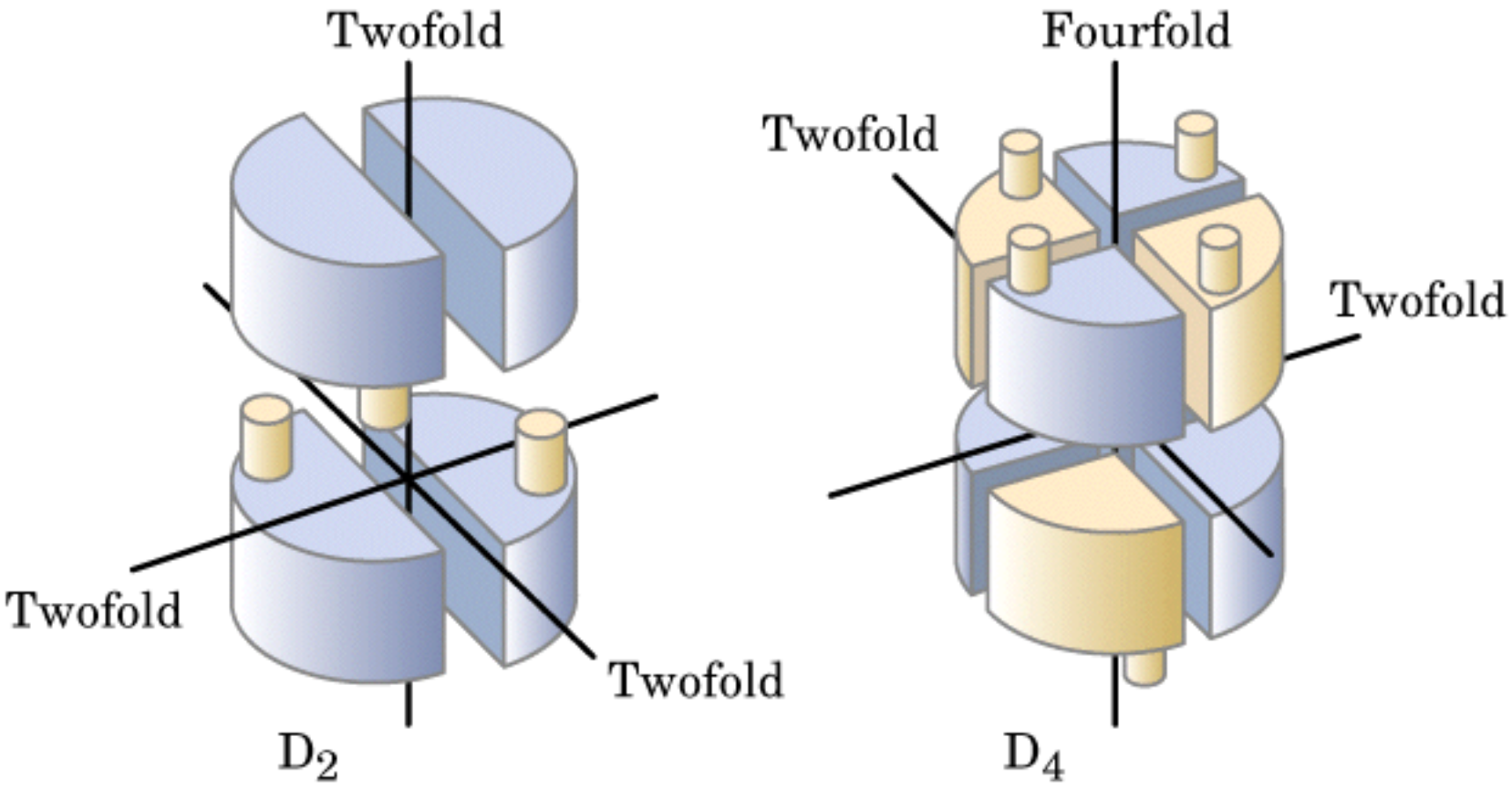
Estructura cuaternaria de la desoxihemoglonbina



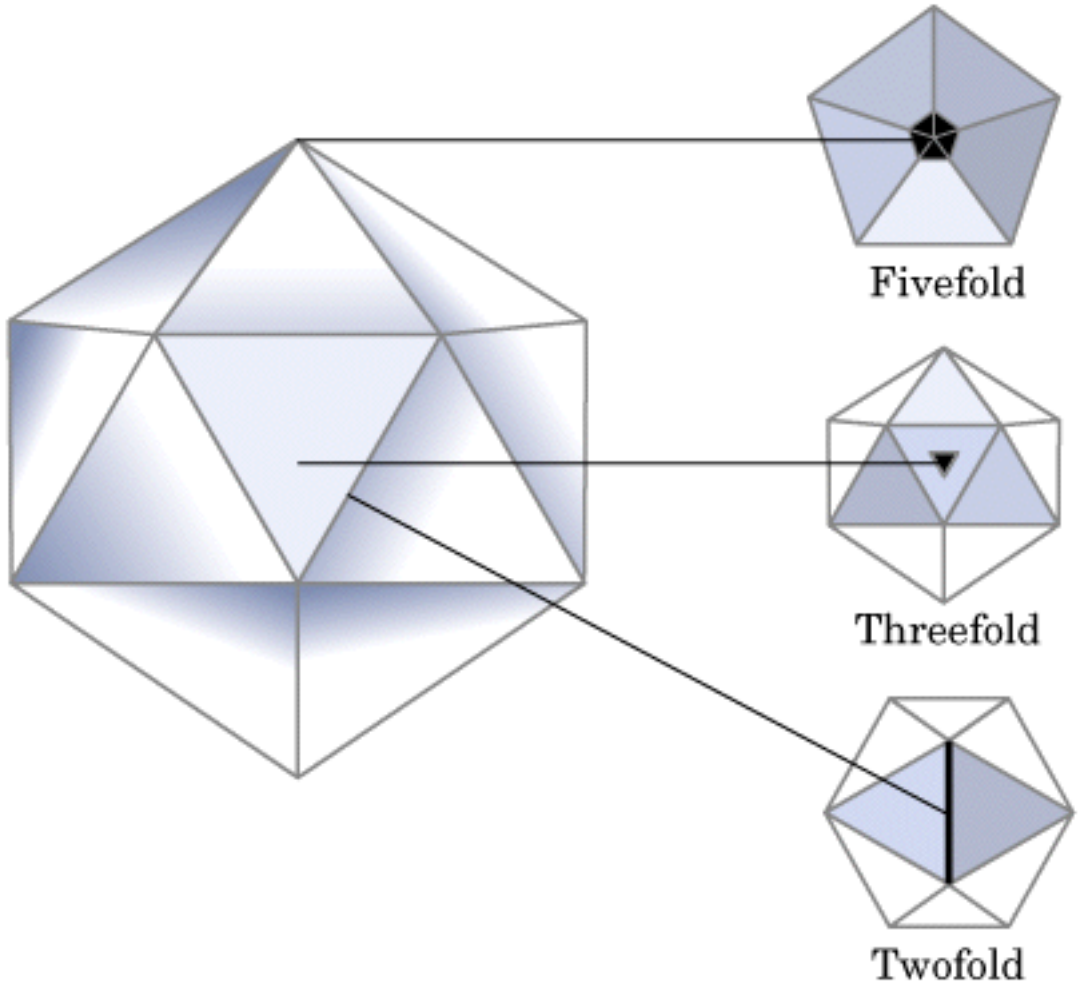
(b)



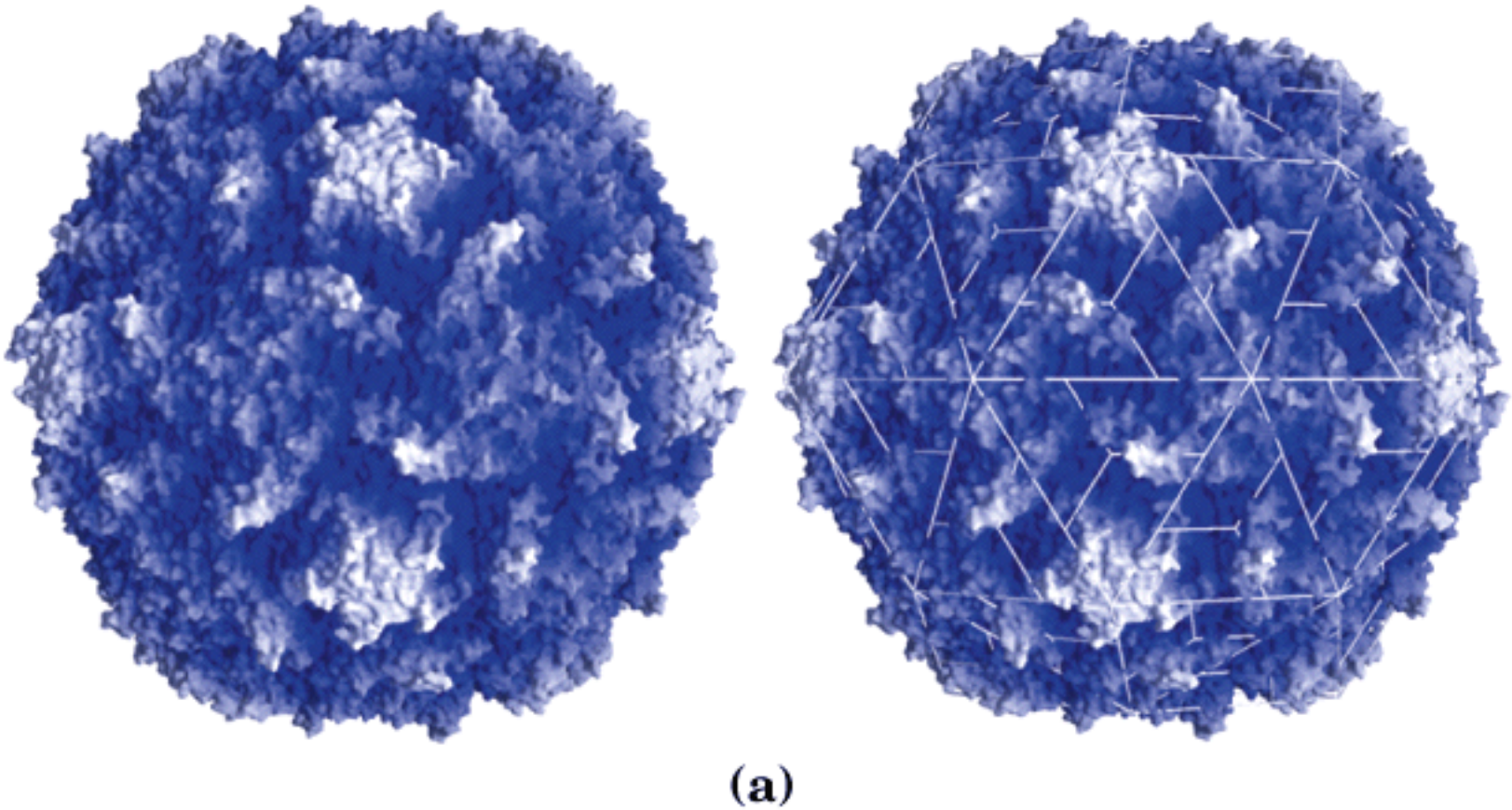
Two types of cyclic symmetry
(a)



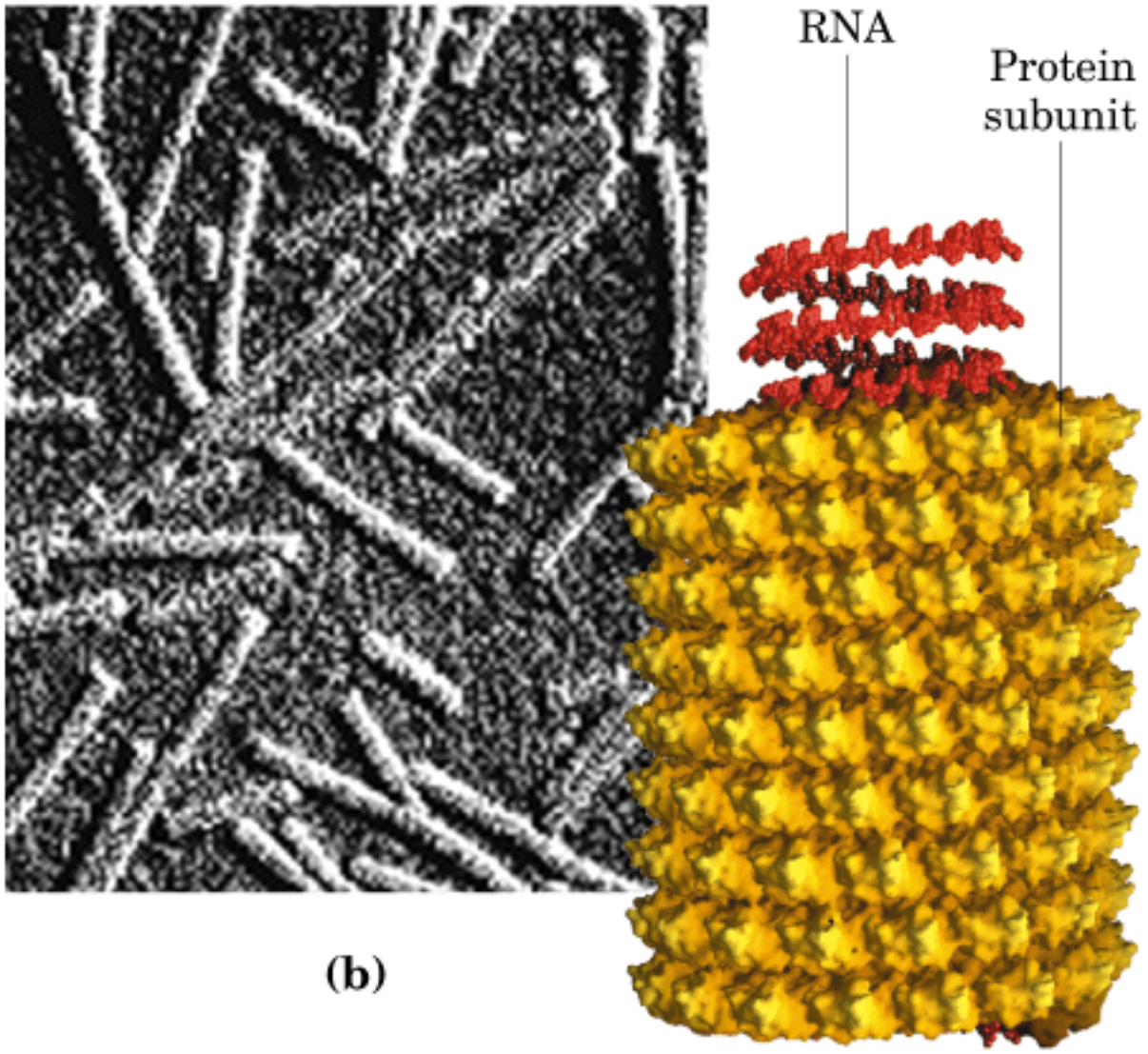
Two types of dihedral symmetry
(b)



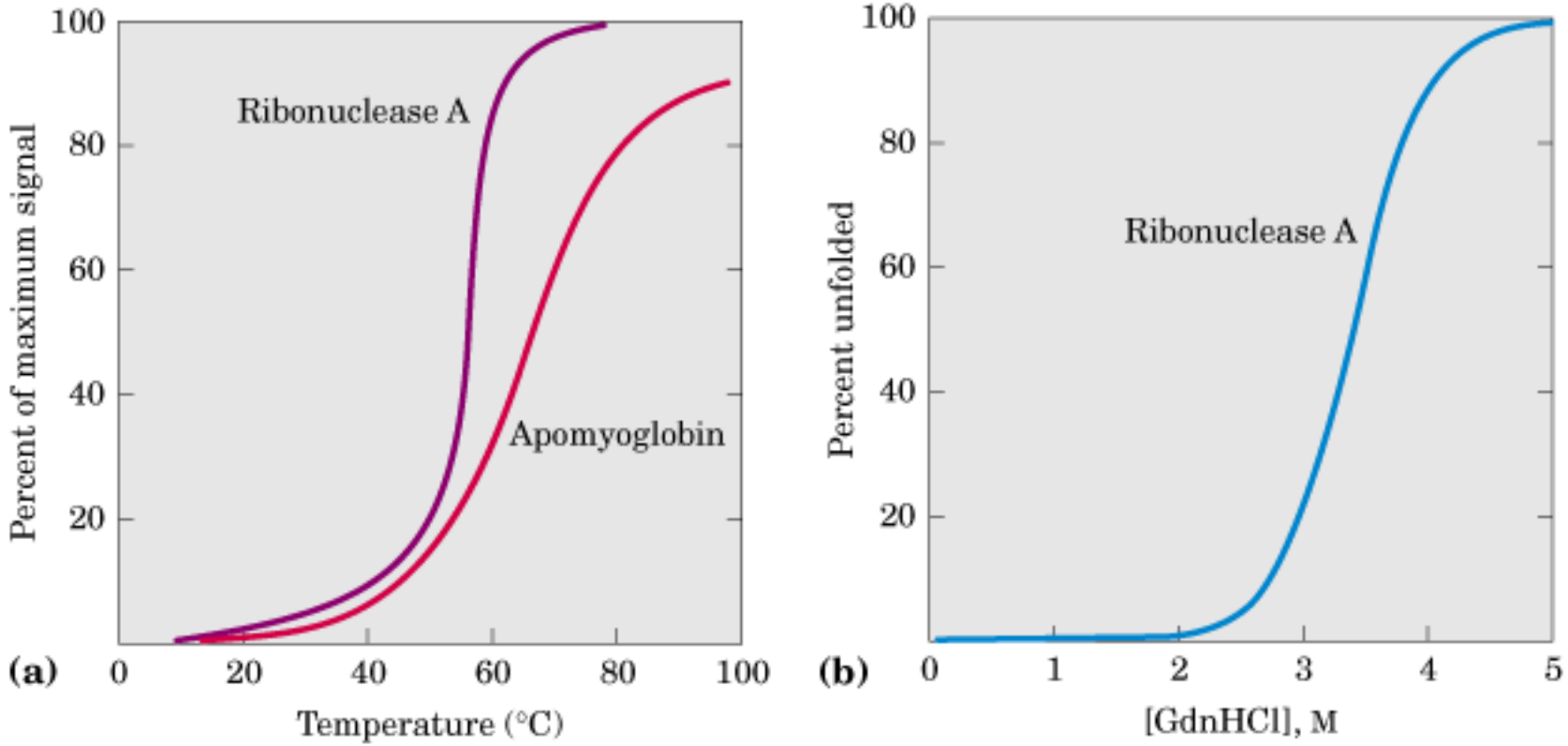
Icosahedral symmetry
(c)



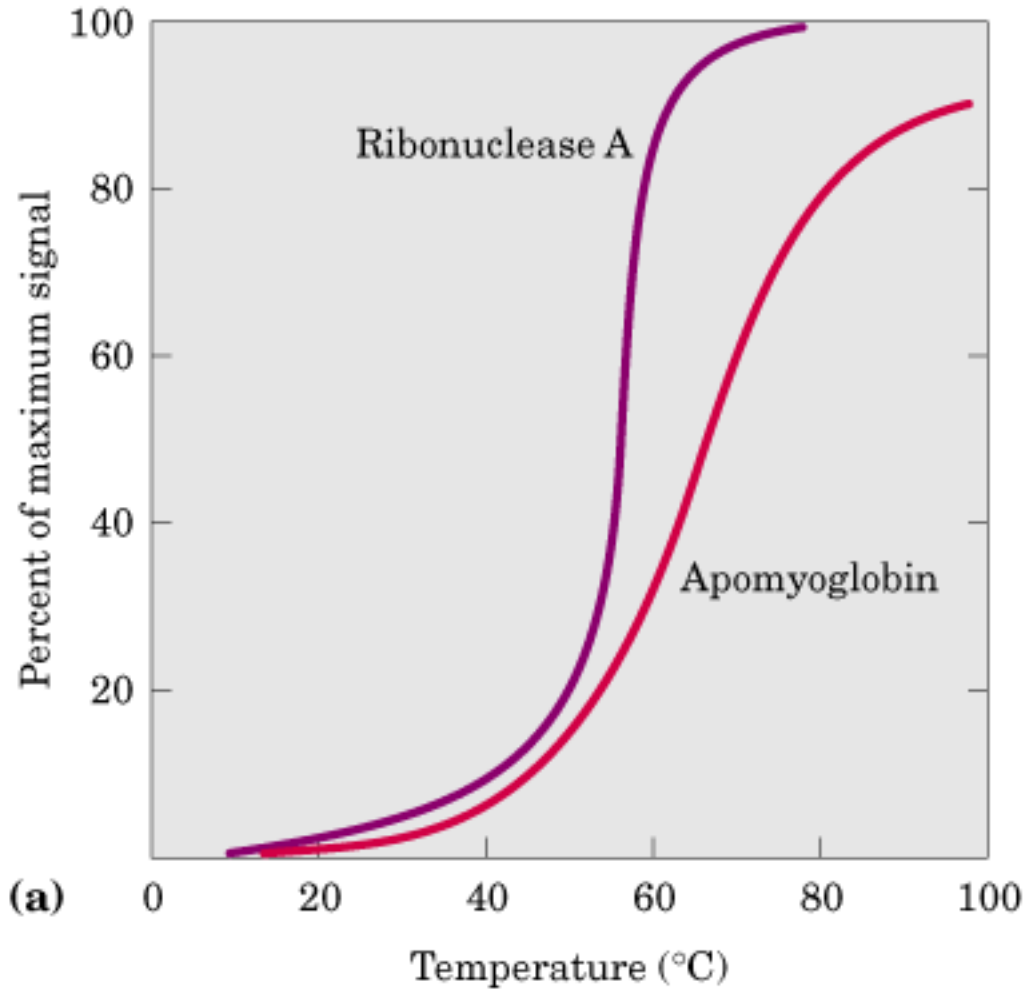
Cápsides de virus (virus de la polio, PDB ID 2PLV)

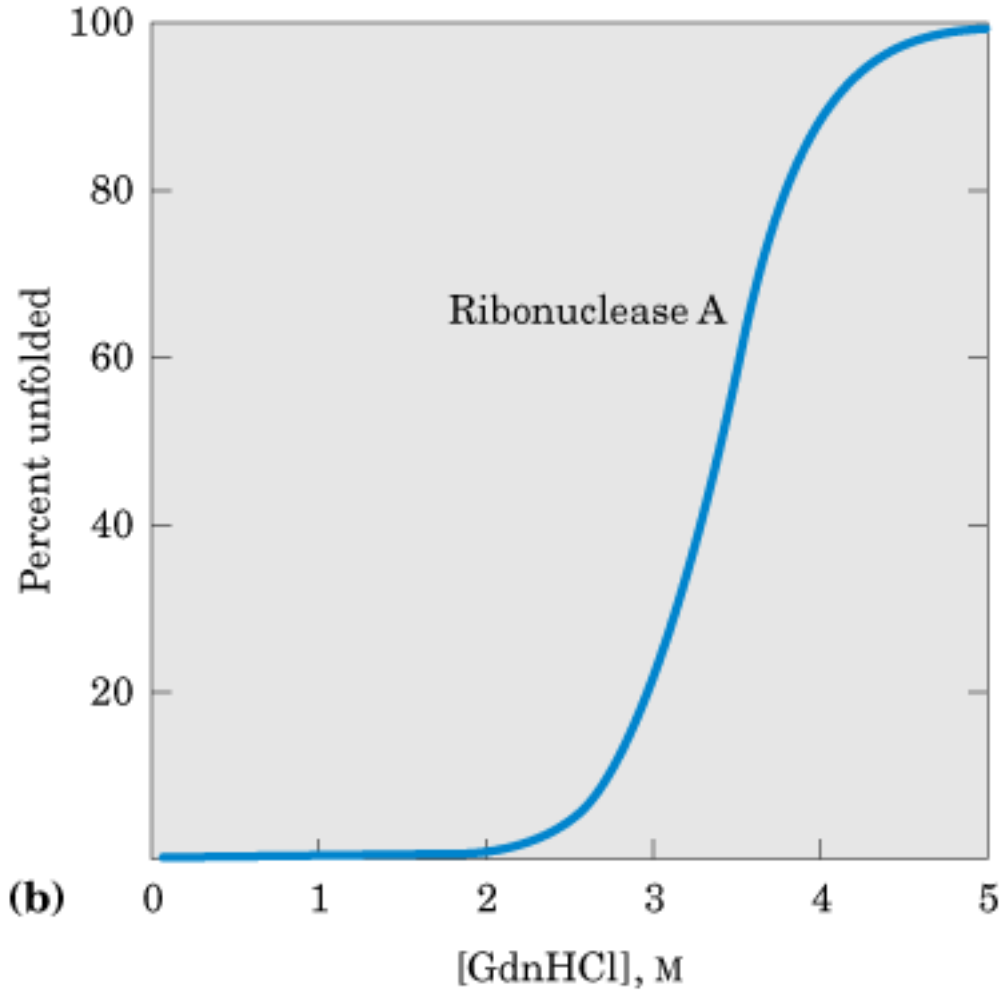


Virus del mosaico del tabaco (PDB ID 1VTM)

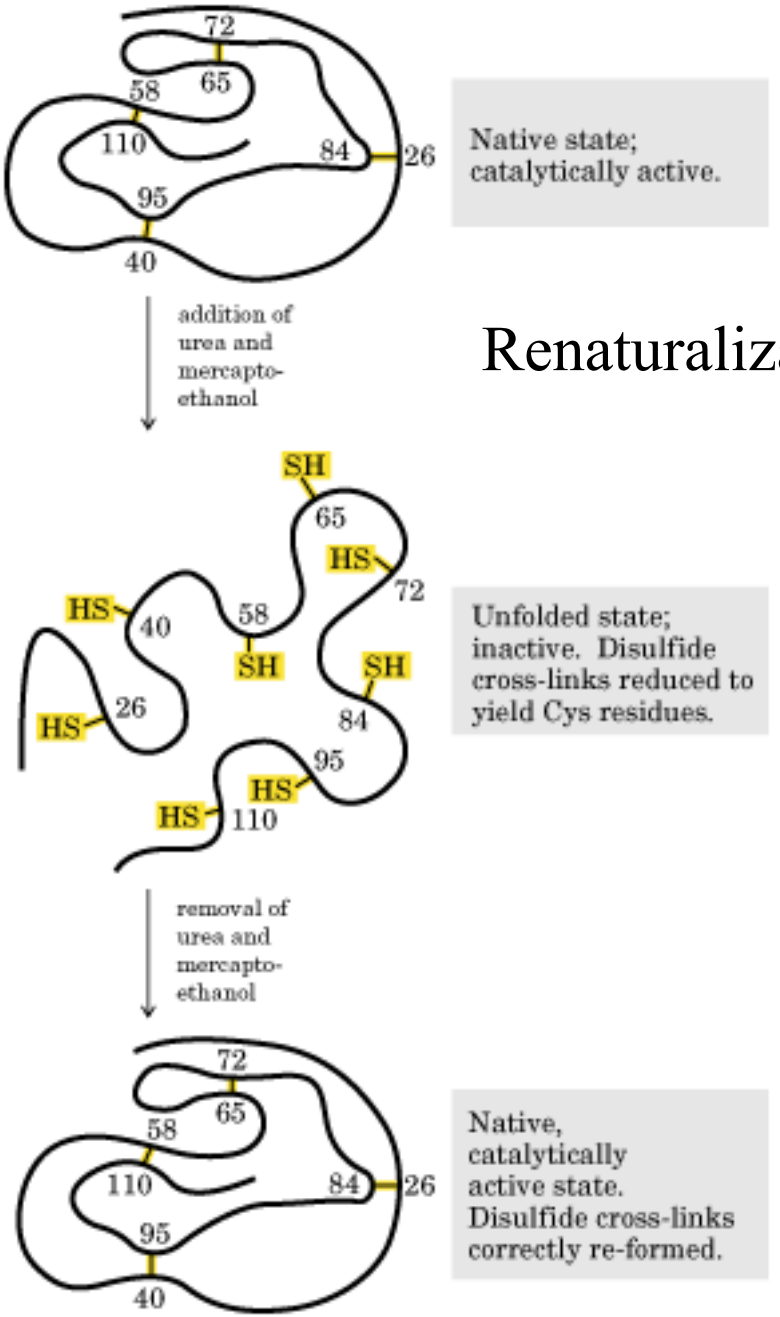


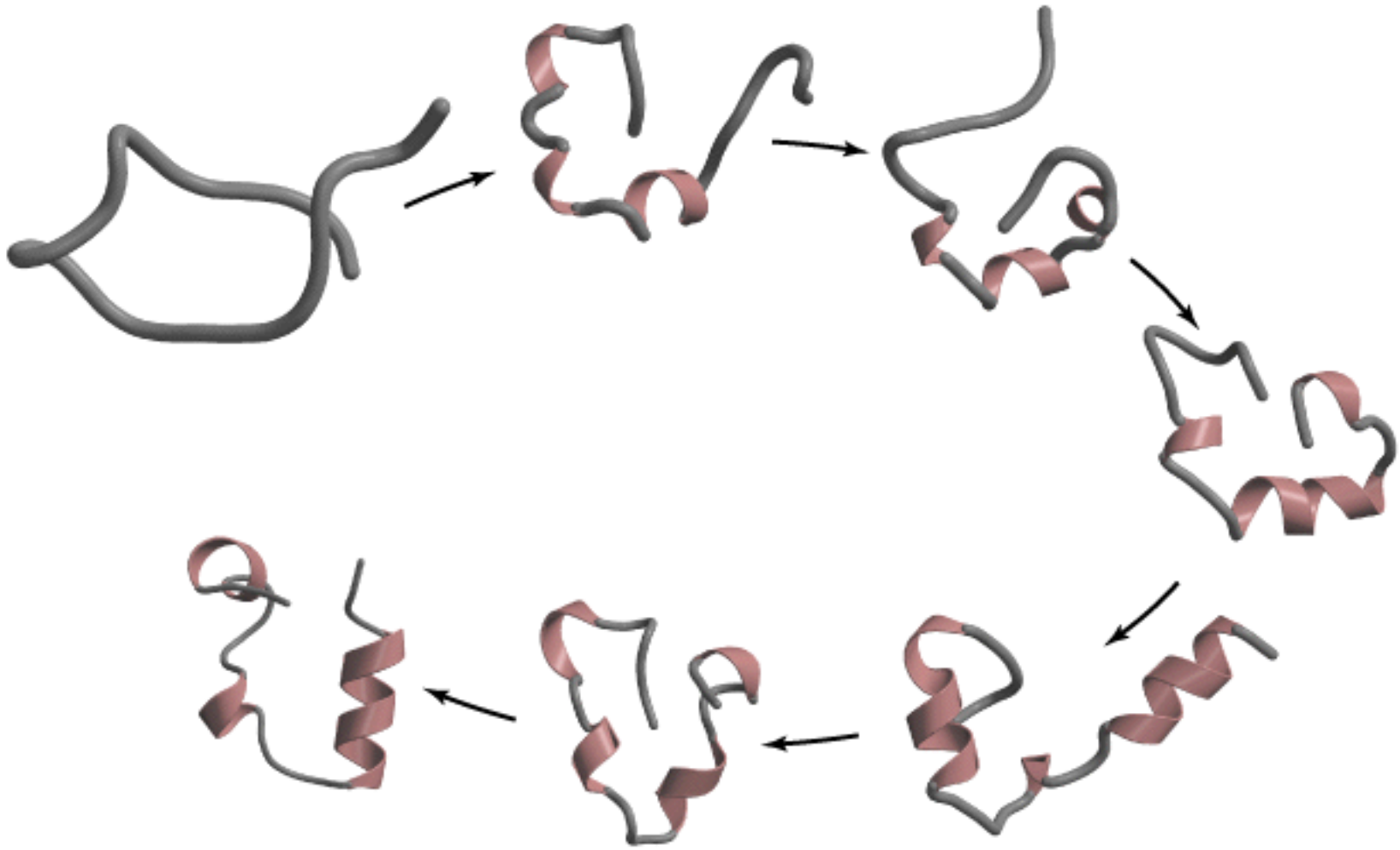
Desnaturalización de proteínas.


100
80
60
40
20
Percent of maximum signal
Ribonuclease A
Apomyoglobin
(a)
0
20
40
60
80
100
Temperature (°C)

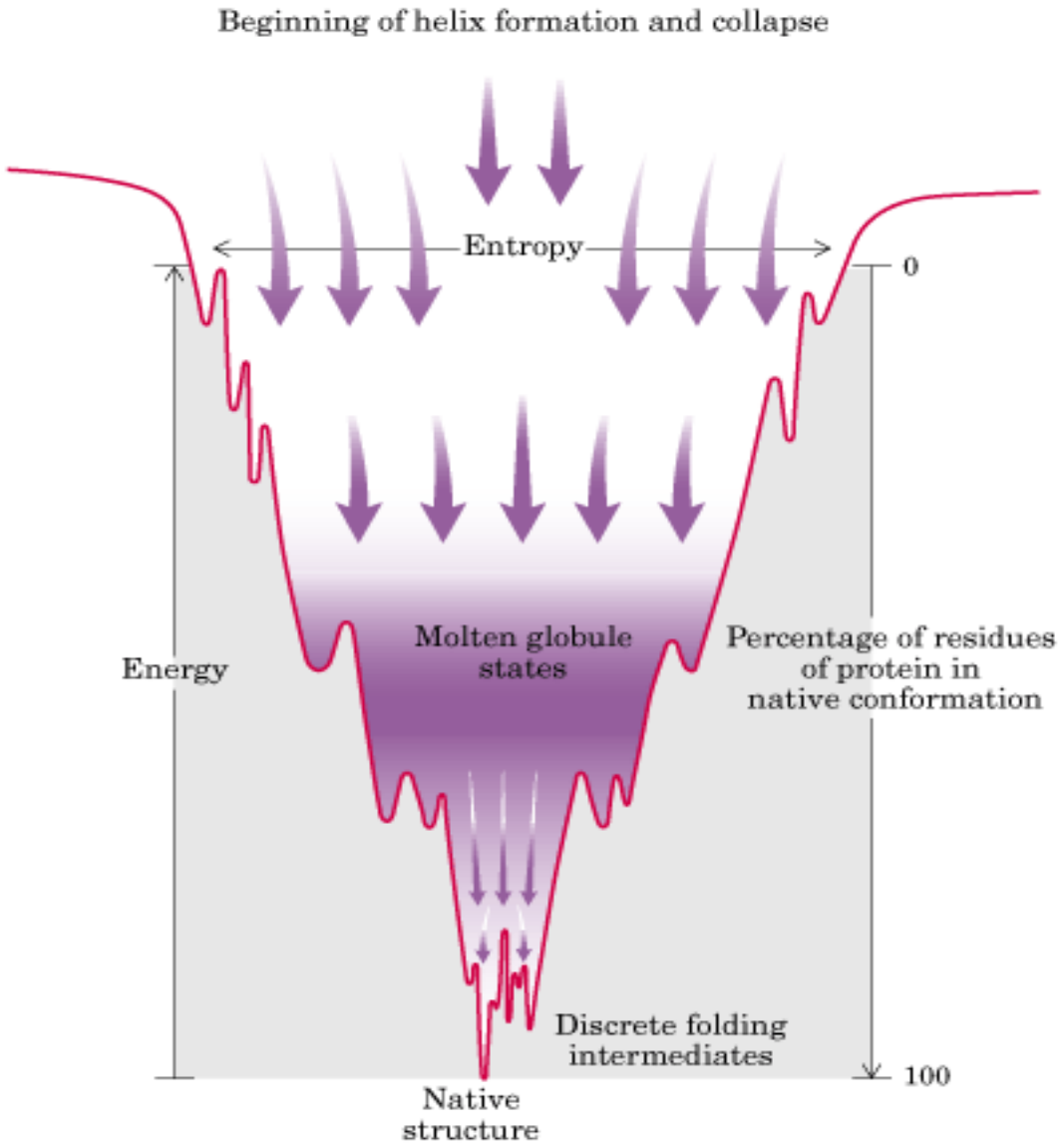

100
80
60
40
20
Percent unfolded
Ribonuclease A
(b)
0
1
2
3
4
5
[GdnHCl], M

Renaturalización de proteínas

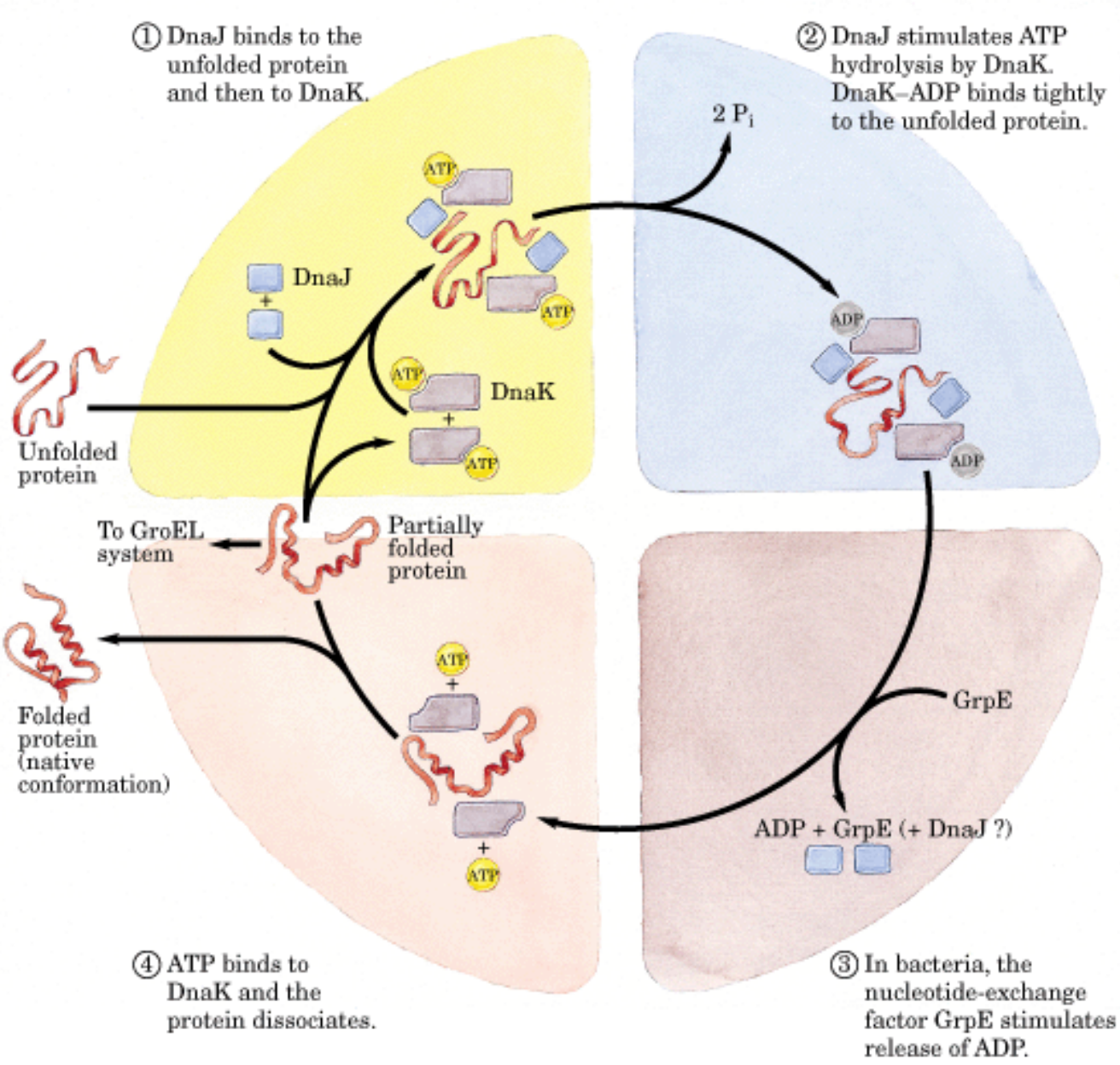




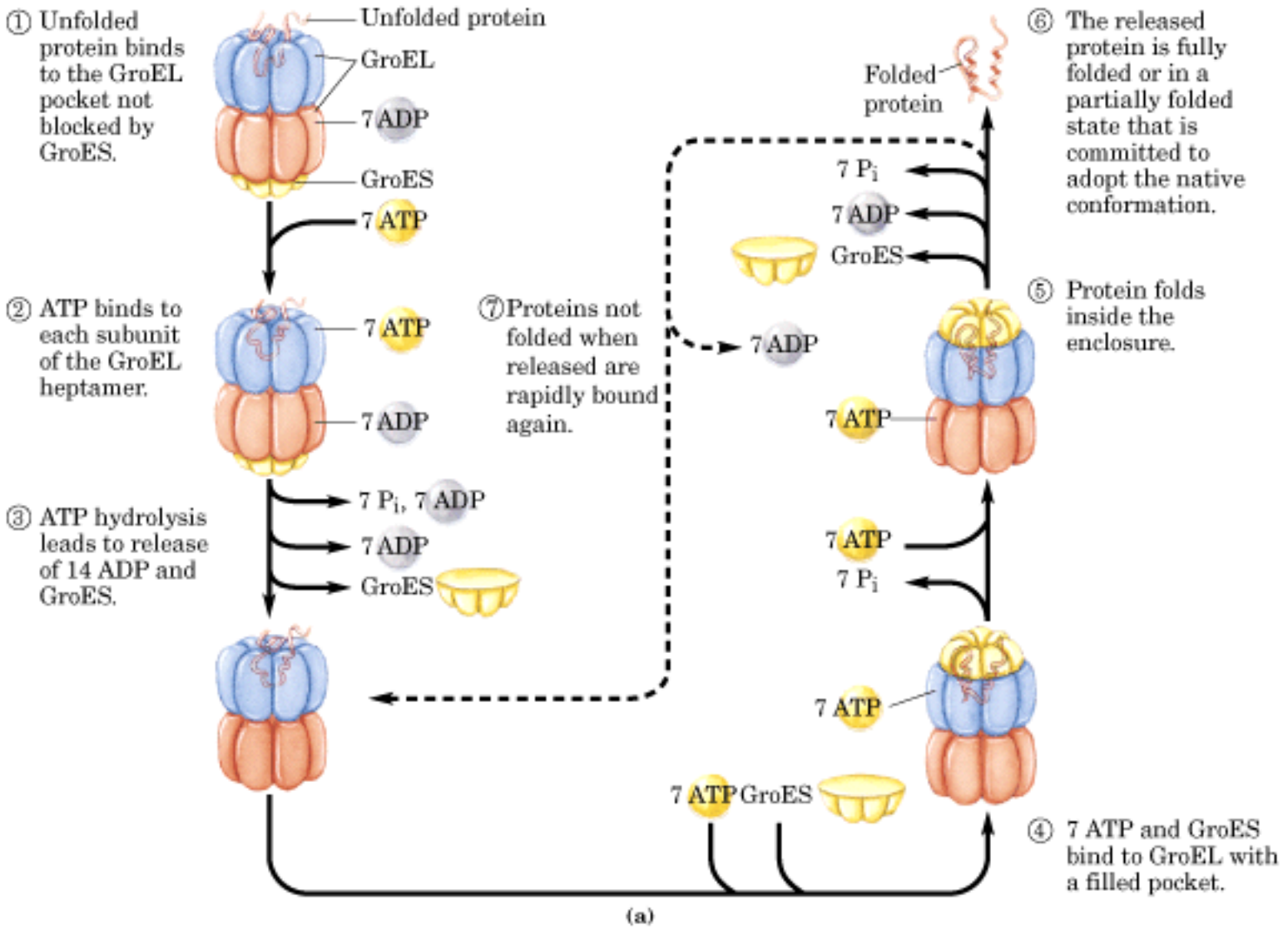
Una ruta de plegamiento simulada.



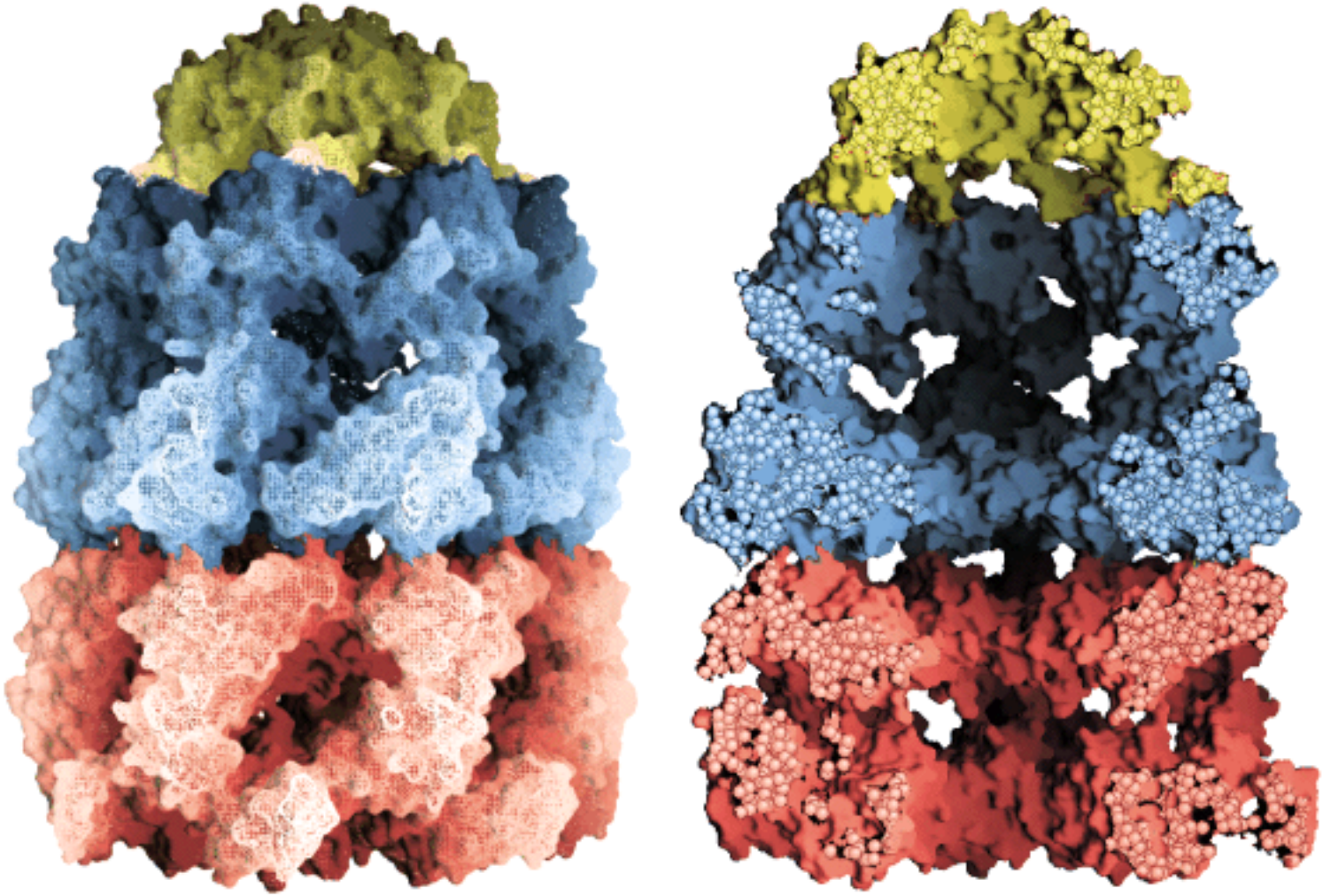
Termodinámica del plegamiento de una proteína representada como un embudo de energía libre.



Las chaperonas en el plegamiento de proteínas

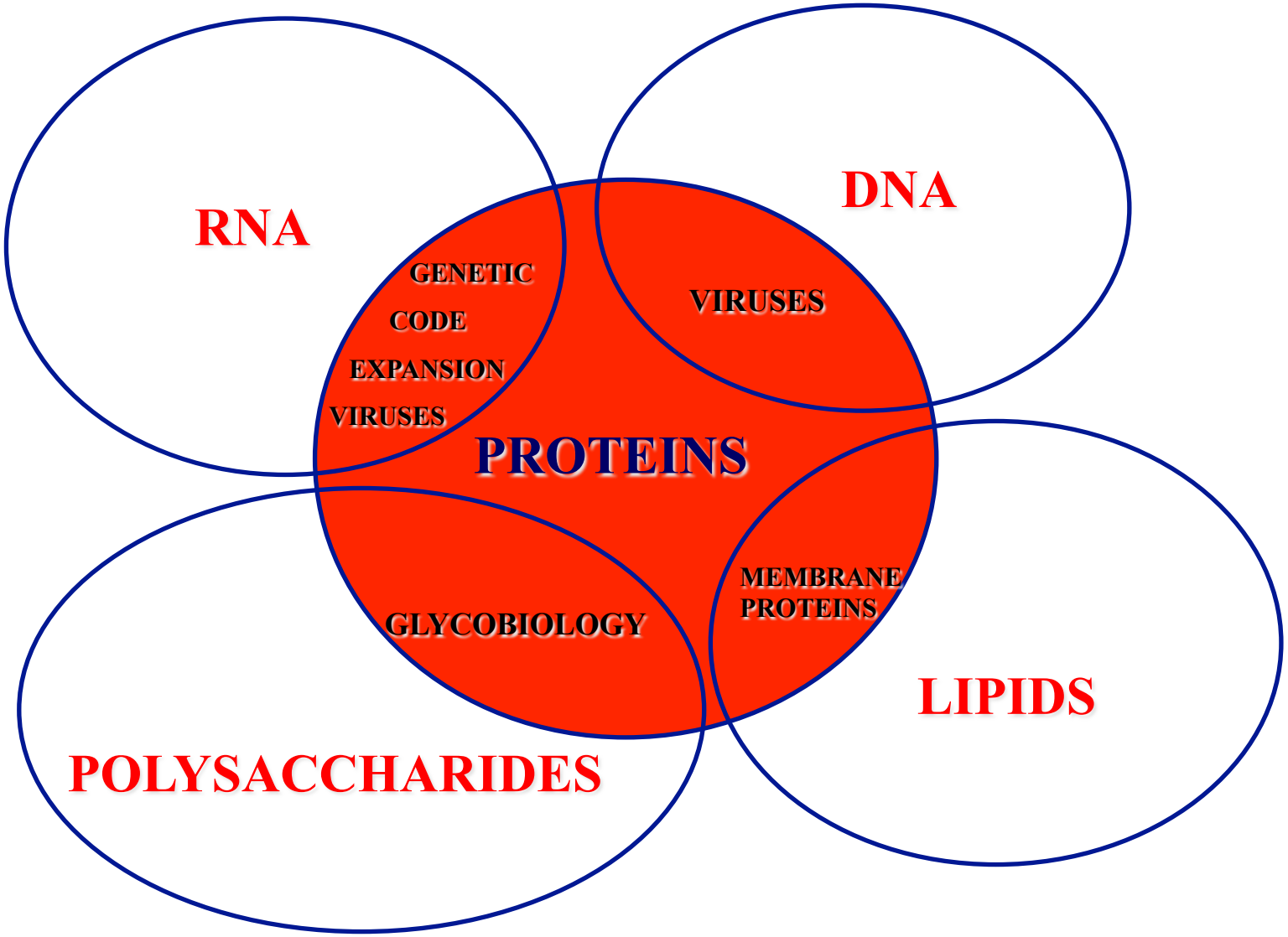


Chaperoninas en el plegamiento de proteinas.



(b)

Imagen de las chaperonas GroEL/GroES, el corte muestra como se pueden unir a otras proteinas.


RNA
DNA
GENETIC
CODE
EXPANSION
VIRUSES
VIRUSES
PROTEINS
MEMBRANE
PROTEINS
GLYCOBIOLOGY
LIPIDS
POLYSACCHARIDES

THREE DIMENSIONAL STRUCTURE FOR BIOLOGICAL SYSTEMS

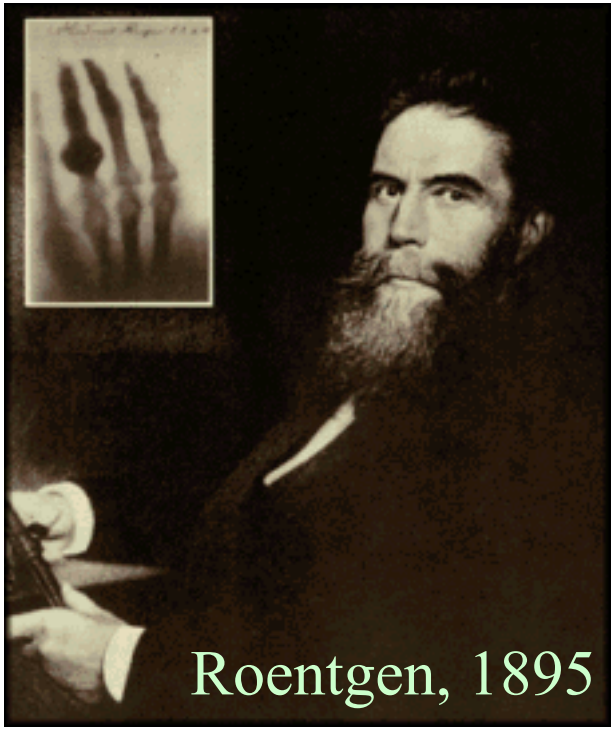
- **1. BY X-RAY DIFFRACTION (CRYSTALS)**
- **2. IN SOLUTION (MW LIMITATIONS)**
- **3. VIA MODELING USING THE PROTEIN DATABASES**
- **4.CRYO ELECTRON MICROSCOPY**

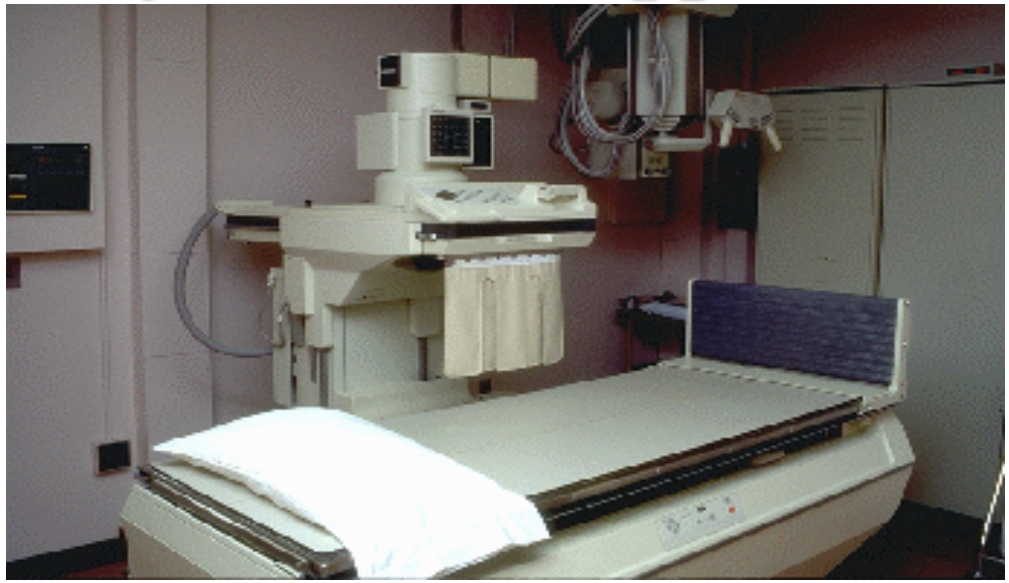
Protein X-ray crystallography is an experimental technique that exploits the fact that X-rays are diffracted by crystals. It is not an imaging technique.

The knowledge of accurate molecular structures is a prerequisite for rational drug design, and for structure based functional studies to aid the development of effective therapeutic agents and drugs.

Crystallography can reliably provide the answer to many structure related questions, from global folds to atomic details of bonding. In contrast to NMR (which is a spectroscopic method), no size limitation exists for the molecule or complex to be studied.

X-rays and their applications





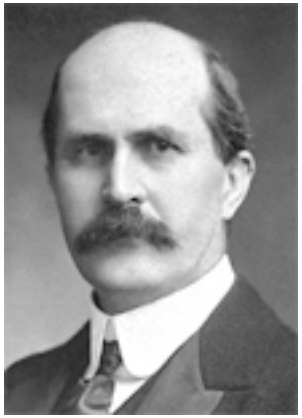
Radiology Unit: for use in diagnosis and treatment of disease

Picture radiography showing a overlapping images.

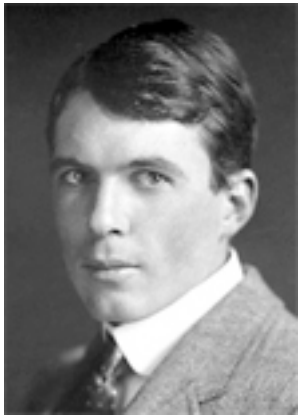
THE DREAM OF A PROTEIN PROTEIN CRYSTALLOGRAPHER…

Max von Laue
(1879-1960)

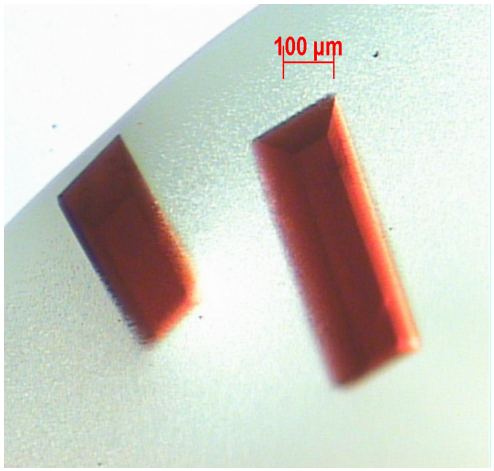
Dorothy Hodgkin
(1910-1994)

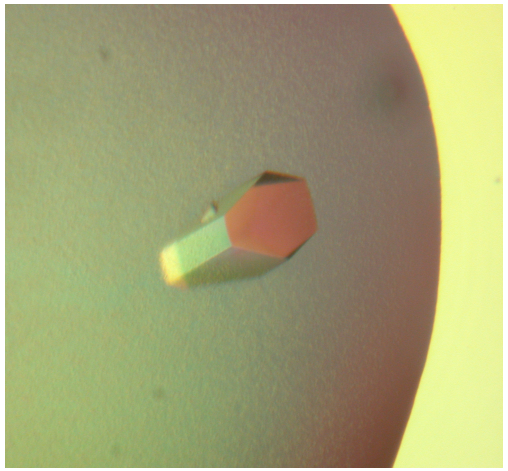


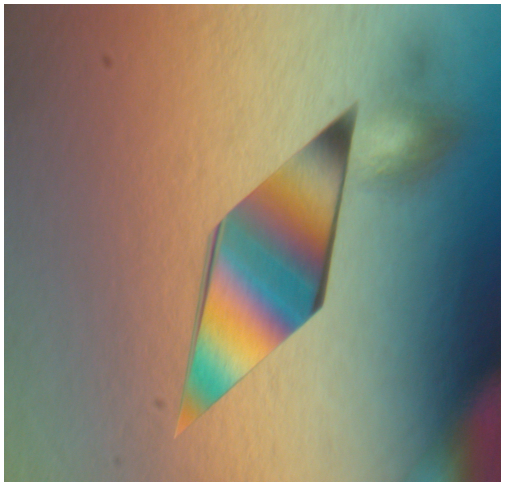
William H. Bragg
(1862-1942)

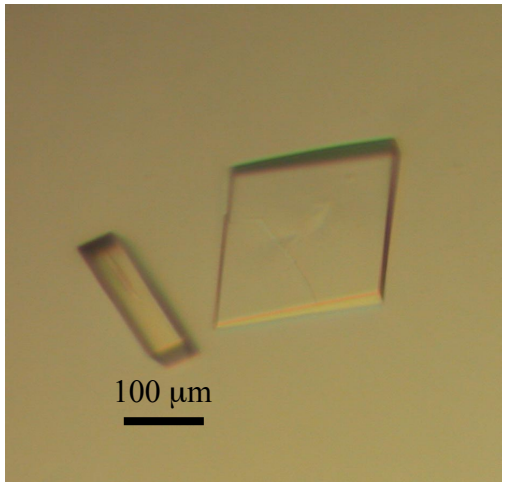


William L. Bragg
(1890 – 1971)



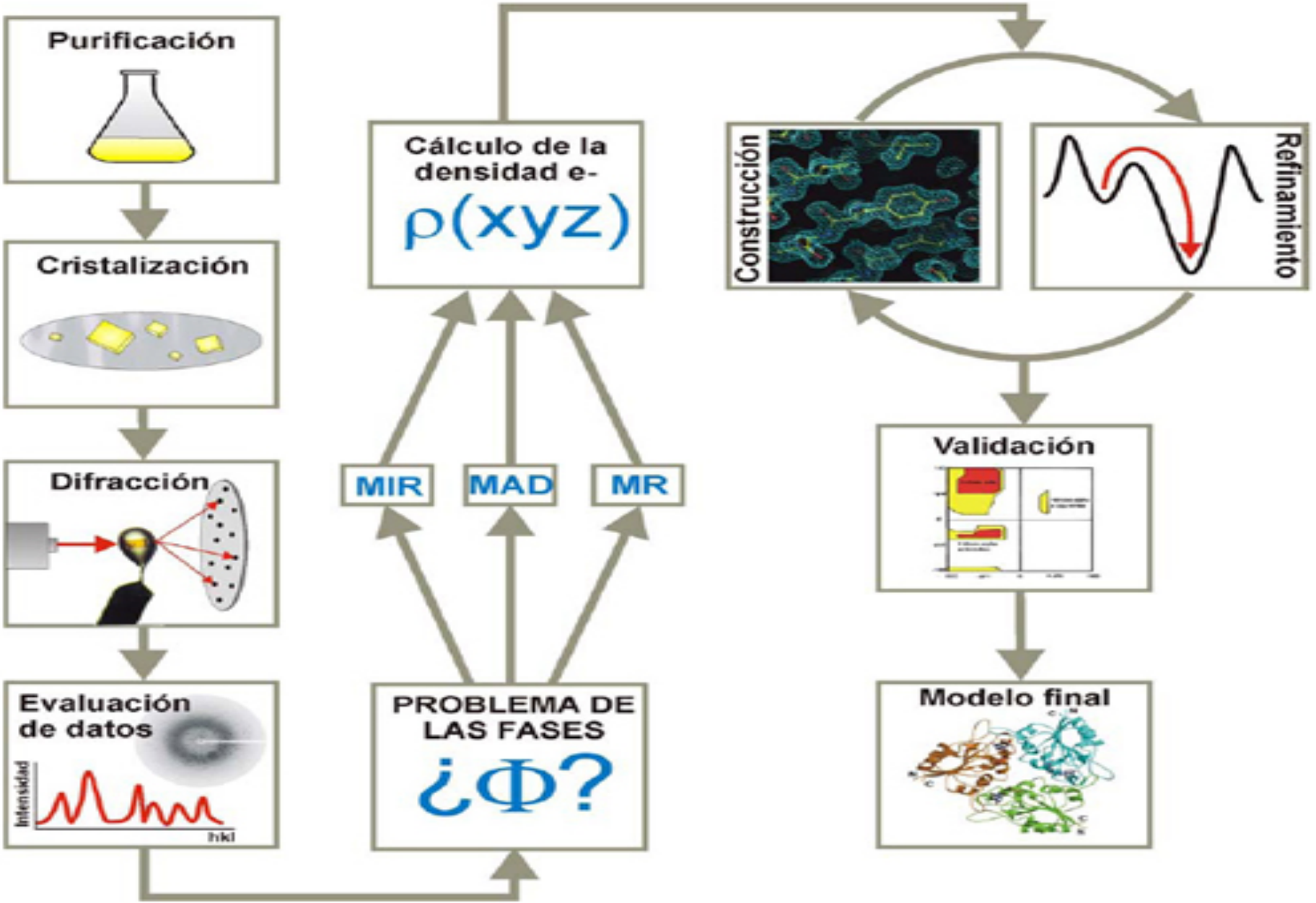






Nowadays: more than 80% of proteins, large complexes, DNA, RNA crystals are usually obtained in droplets ranging from 10 to 1000 nl. Average size of the droplets is about 200 nl.

GENERAL SCHEME FOR 3D STRUCTURE RESOLUTION



CONCLUSIONS

- Biomacromolecular Crystallography has become an important science to be reinforced by the Biotechnology to make new recombinant proteins available through the knowledge of the DNA, RNA technology.

- The knowledge of the 3D structure of biological macromolecules is a necessary condition to explain biological mechanisms or biochemical pathways. The bottleneck is still the crystallization step for protein crystallography research focused on crystallizing difficult targets.

- The Biotechnology of the New Millennium needs to meet different sciences like, Medicine, Materials Science, Physics, Chemistry to spread widely its power to the development of new strategic products biologically or genetically controlled by human beings.

Acknowledgements

Prof. Dr. Gen Sazaki, Tohoku University. Sendai (JAPAN)

Prof. Kazuo Nakajima. IMR-Tohoku University (JAPAN).

Dr. Nurit Mirkin, UNAM (Mexico)

Prof. Richard Giegé, Dr. Bernard Lorber and Dr. Claude Sauter from the IBMC-CNRS, France.

Prof. Dr. Sir Tom L. Blundell, Dr. Dima Chirgadze, Dr. Sachin Surade, Mr. Takashi Ochi. University of Cambridge, UK.

Dr. Daniel Nietlispach. NMR Facility, University of Cambridge, UK.