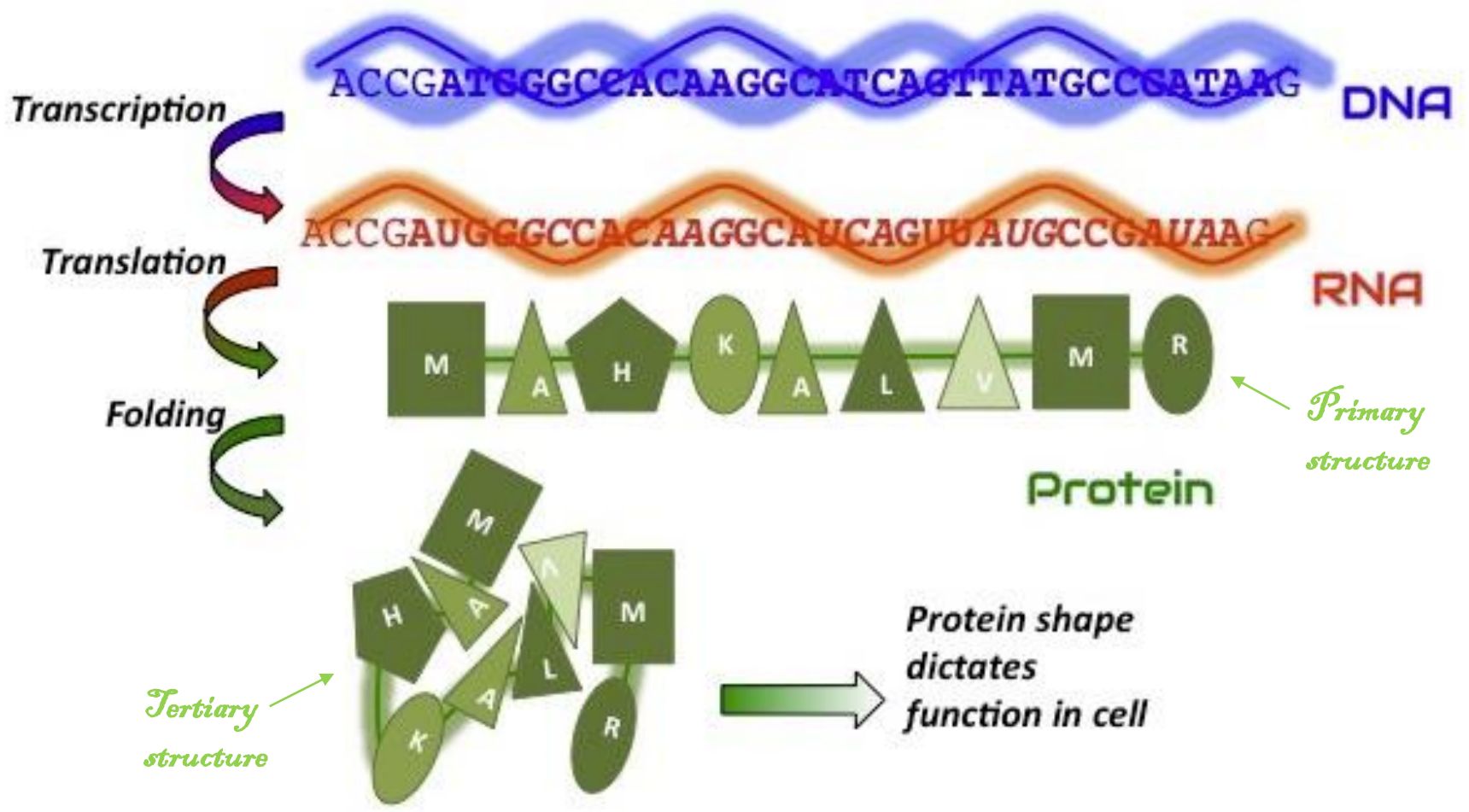
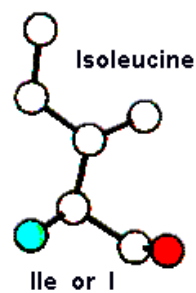
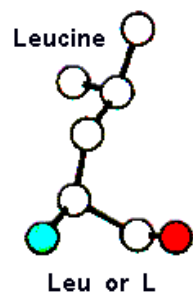
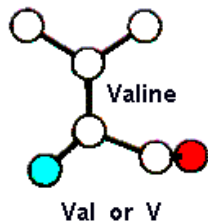
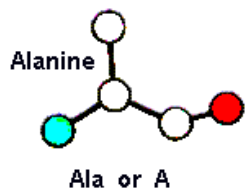


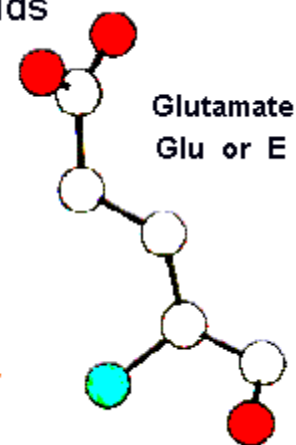
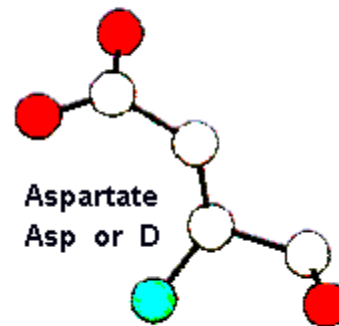
A DOGMA CENTRAL DA BIOLOGIA



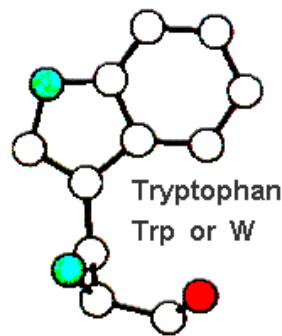
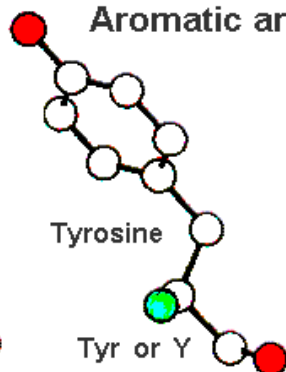
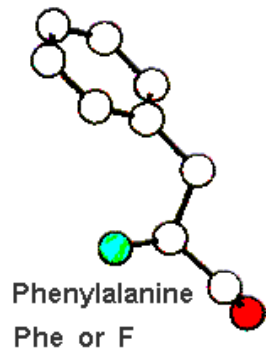
Aliphatic amino acids



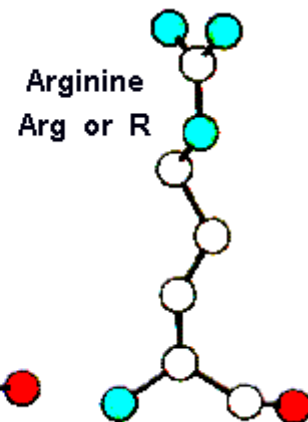
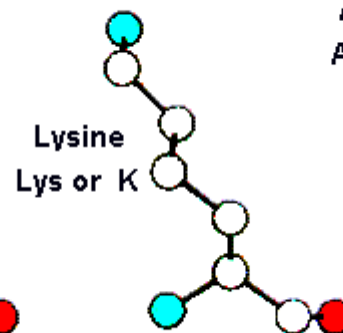
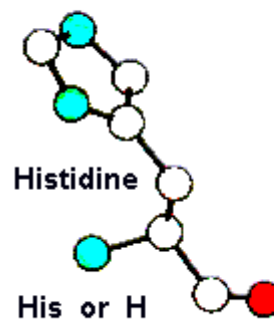
Acidic amino acids



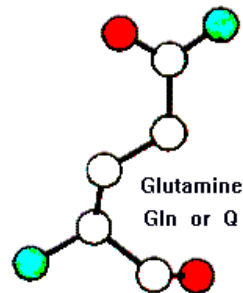
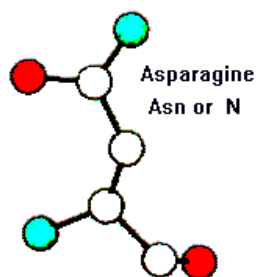
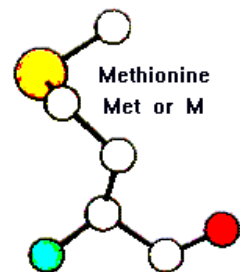
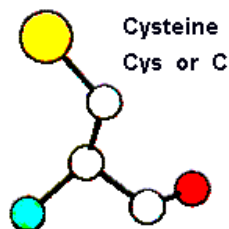
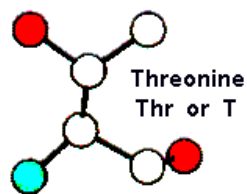
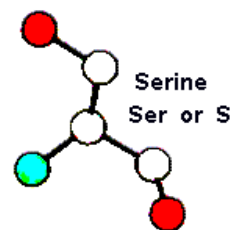
Aromatic amino acids



Basic amino acids



Neutral-polar amino acids



Conformationally important amino acids

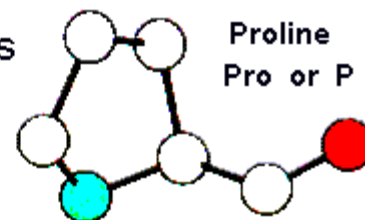
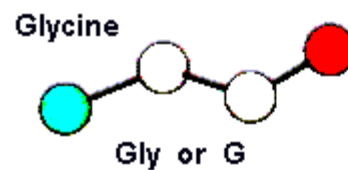


table 5-2

Dados Moleculares Sobre Algumas Proteínas

	Peso Molecular	Num. Resíduos	Num. Cadeias
Cytochrome <i>c</i> (human)	13,000	104	1
Ribonuclease A (bovine pancreas)	13,700	124	1
Lysozyme (egg white)	13,930	129	1
Myoglobin (equine heart)	16,890	153	1
Chymotrypsin (bovine pancreas)	21,600	241	3
Chymotrypsinogen (bovine)	22,000	245	1
Hemoglobin (human)	64,500	574	4
Serum albumin (human)	68,500	609	1
Hexokinase (yeast)	102,000	972	2
RNA polymerase (<i>E. coli</i>)	450,000	4,158	5
Apolipoprotein B (human)	513,000	4,536	1
Glutamine synthetase (<i>E. coli</i>)	619,000	5,628	12
Titin (human)	2,993,000	26,926	1

Composição de aminoácidos de duas proteínas

table 5-3

Amino acid	Number of residues per molecule of protein	
	Bovine cytochrome <i>c</i>	Bovine chymotrypsinogen
Ala	6	22
Arg	2	4
Asn	5	15
Asp	3	8
Cys	2	10
Gln	3	10
Glu	9	5
Gly	14	23
His	3	2
Ile	6	10
Leu	6	19
Lys	18	14
Met	2	2
Phe	4	6
Pro	4	9
Ser	1	28
Thr	8	23
Trp	1	8
Tyr	4	4
Val	3	23
Total	104	245

*Note that standard procedures for the acid hydrolysis of proteins convert Asn and Gln to Asp and Glu, respectively. In addition, Trp is destroyed. Special procedures must be employed to determine the amounts of these amino acids.

table 5–6

Pontos Isoelétricos (pI) de Algumas Proteínas

Protein	pI
Pepsin	~1.0
Egg albumin	4.6
Serum albumin	4.9
Urease	5.0
β -Lactoglobulin	5.2
Hemoglobin	6.8
Myoglobin	7.0
Chymotrypsinogen	9.5
Cytochrome <i>c</i>	10.7
Lysozyme	11.0

Estrutura 3D de proteínas

Estrutura primária

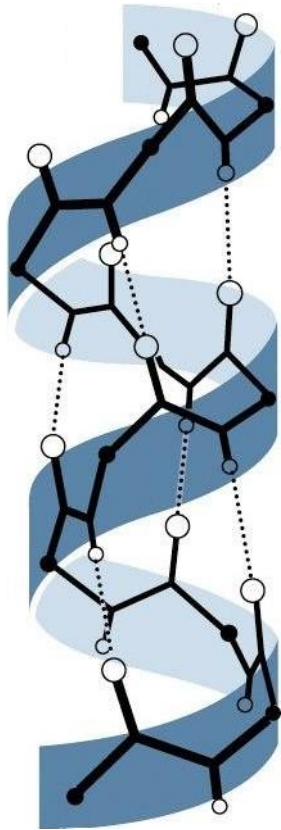
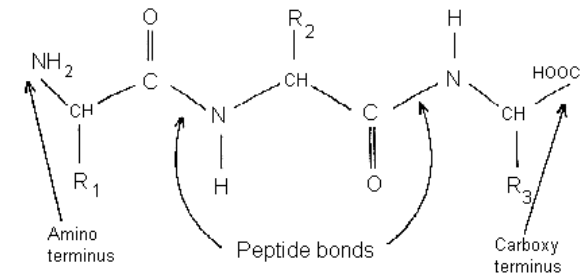
Estrutura secundária

Estrutura terciária

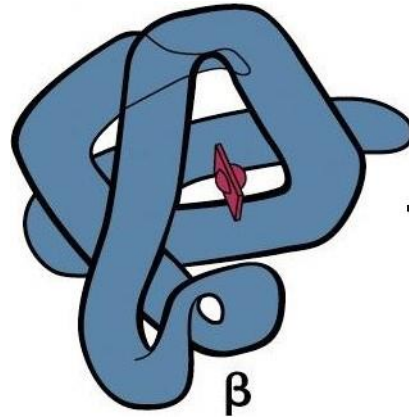
Estrutura quaternária

Enovelamento de proteínas

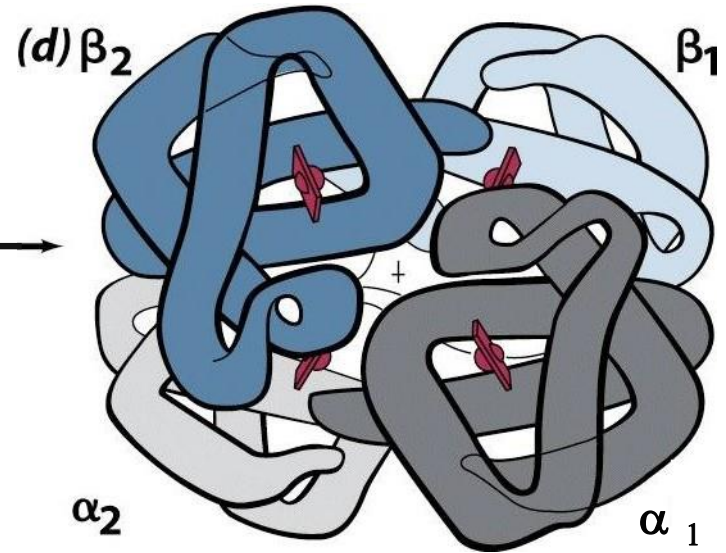
... Lys - Ala - His - Gly - Lys - Lys - Val - Leu - Gly - Ala ...
 Estrutura primária = estrutura covalente
 = sequência de aminoácidos na cadeia polipeptídica



Estrutura secundária
 (estruturas regulares de trechos de sequência: hélice, fita, volta, “cadeia randômica”)

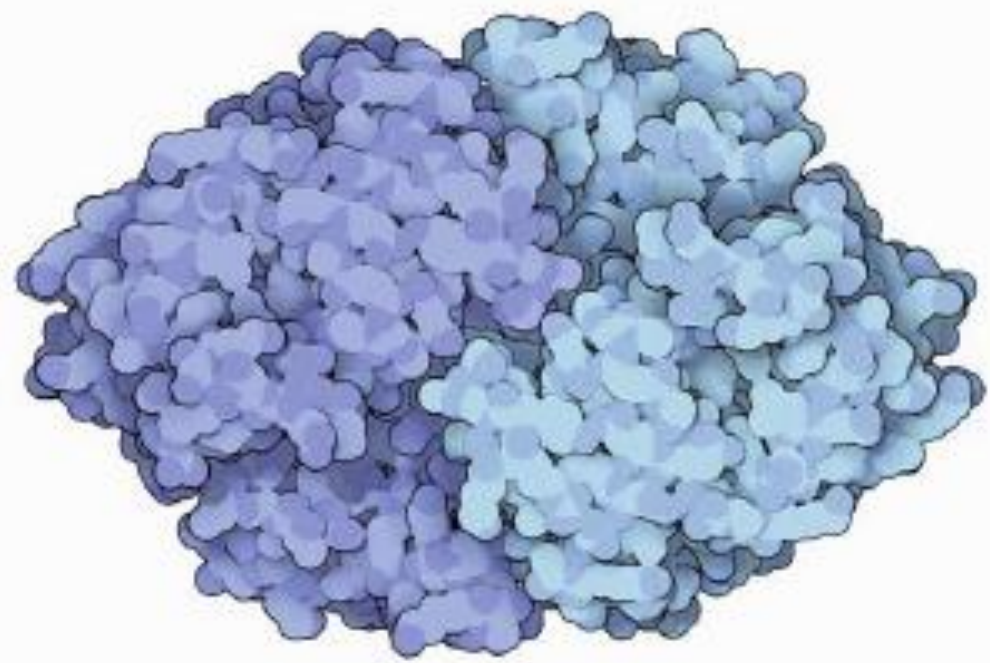


Estrutura terciária:
 enovelamento/**topologia**
 de uma cadeia polipeptídica

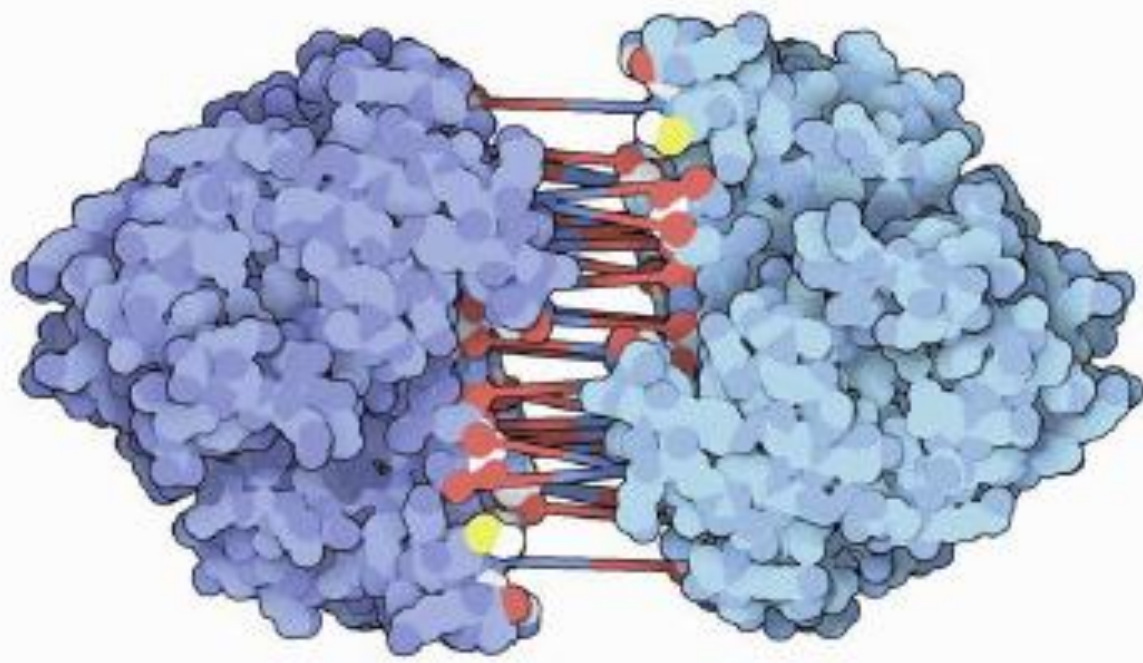


Estrutura quaternária:
 Associação de 2 ou + **cadeias/subunidades**
 ie: as quatro cadeias ($\alpha_2\beta_2$) se juntando para formar a proteína funcional de hemoglobina

Estruturas secundárias, terciária e quaternária são mantidas principalmente por meio de interações não-covalentes: pontes de hidrogênio, interações van der Waals, interações carga-carga e dipolo-dipolo e o efeito hidrofóbico.



Exceção: pontes dissulfetos entre cadeias laterais de cisteína



dipole interactions

1 r^2 to r^4 8

ionic interactions

3 r^1 10

van der Waals

0.5 r^6 1.5

covalent C,N,O bonds

40 fixed distance 130

RT

hydrogen bond

0.6 1.5 2.4 to 5 Å 7

covalent S-S bond

60 65

0.5

1

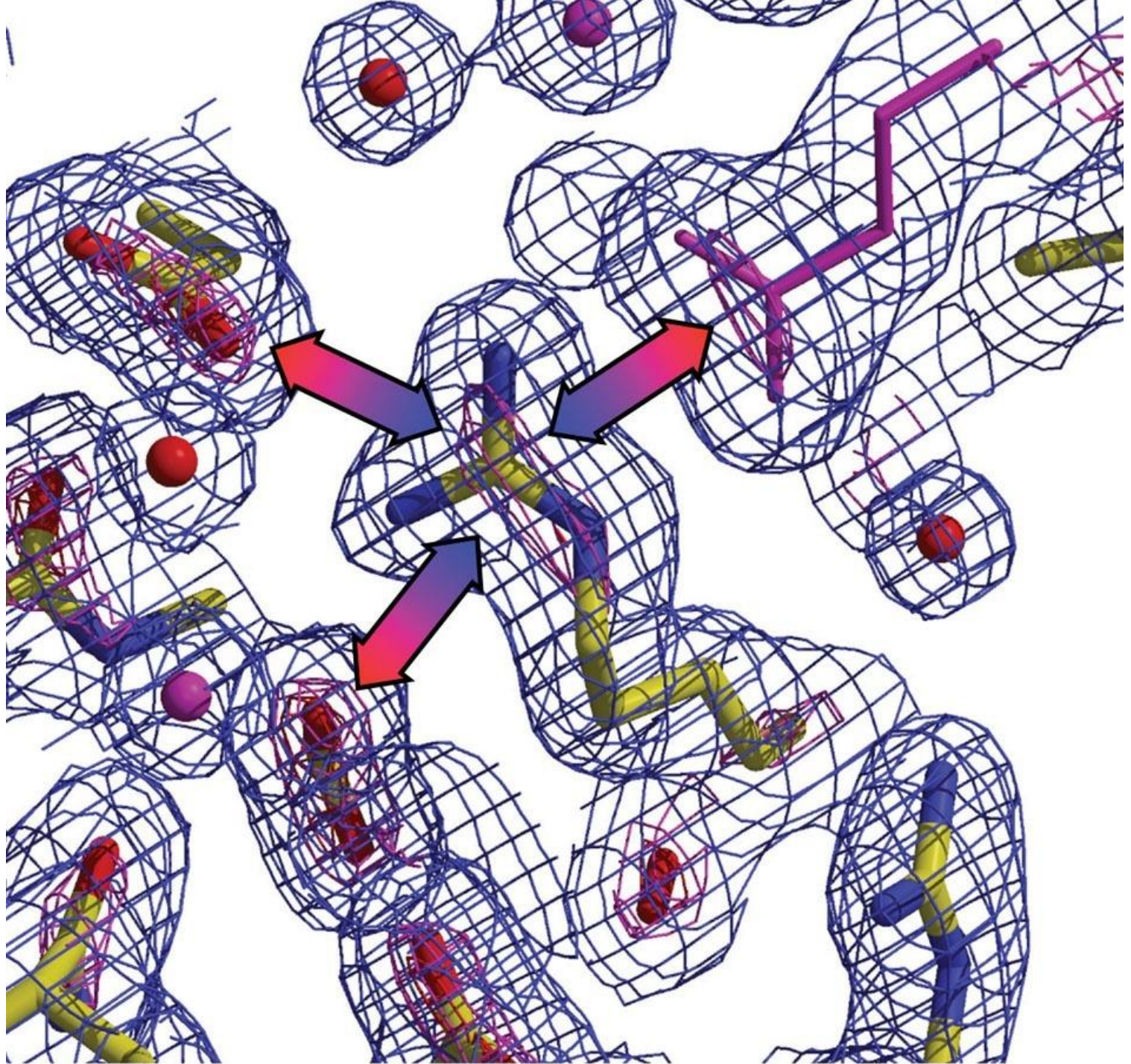
5

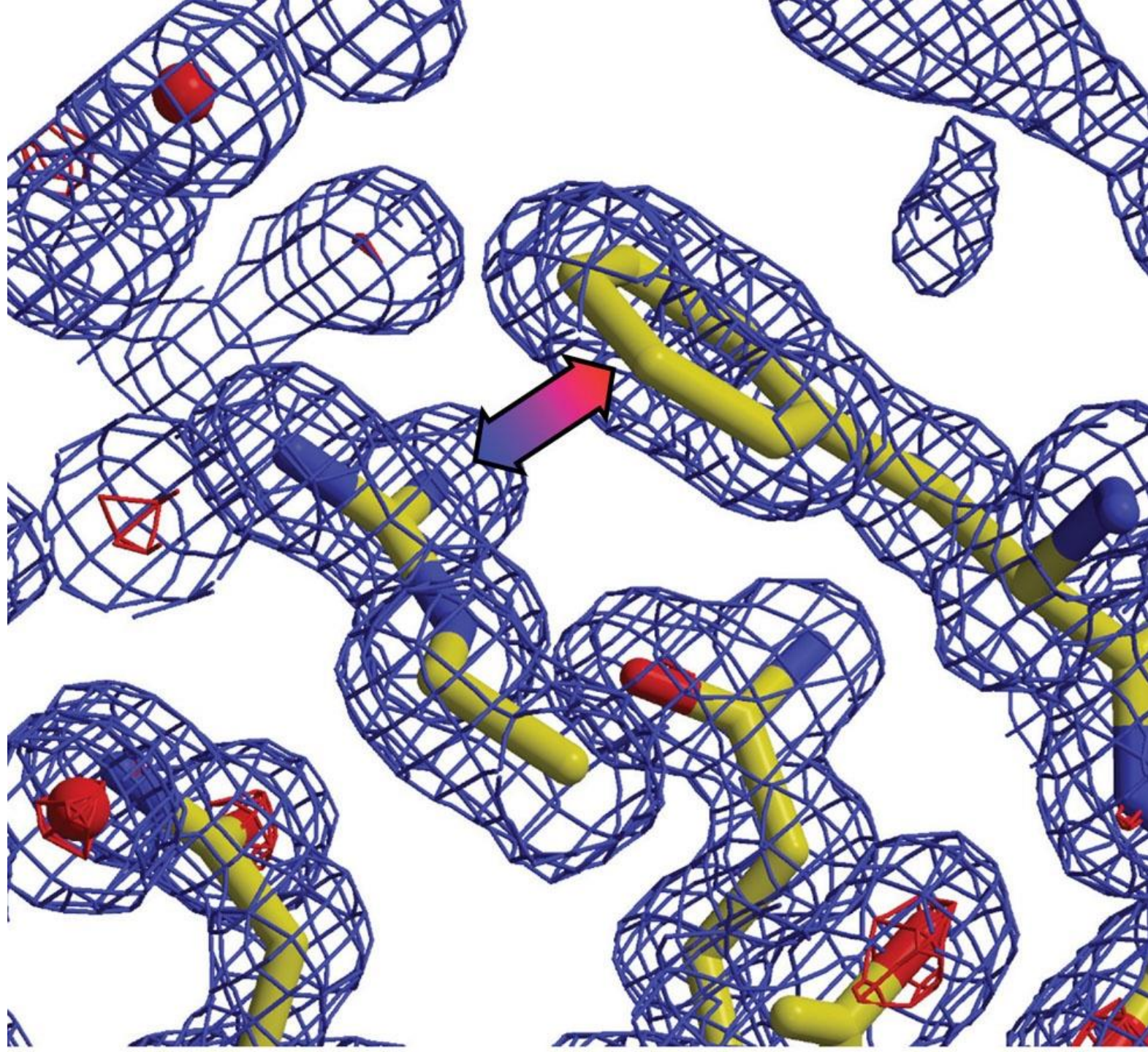
10

50

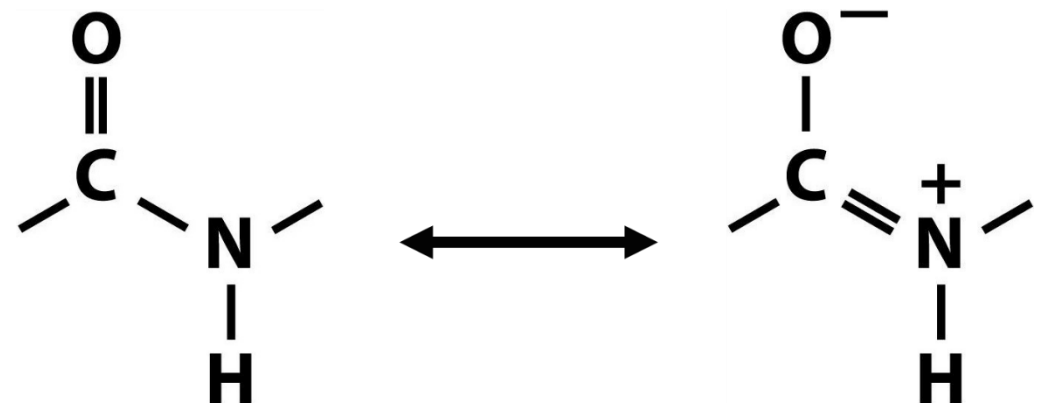
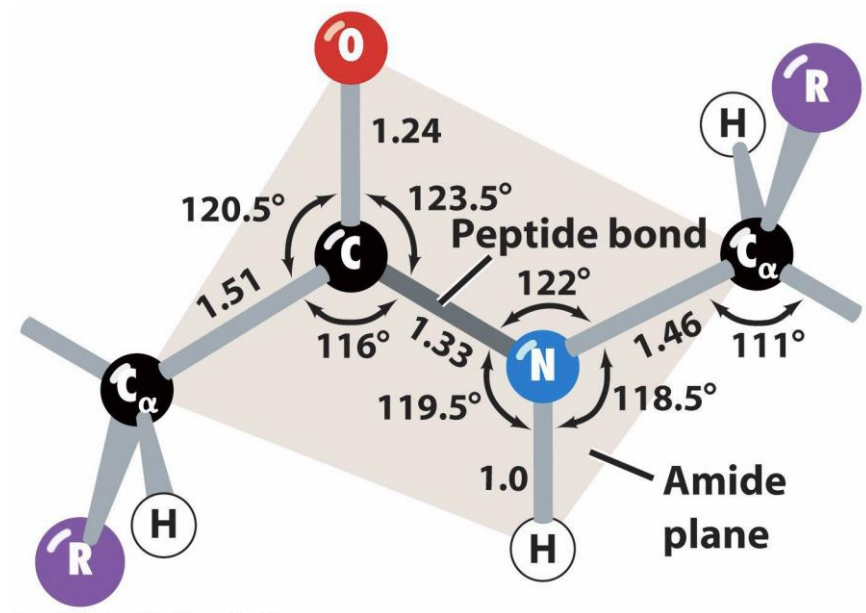
100

Bond energy (-kcal/mol)



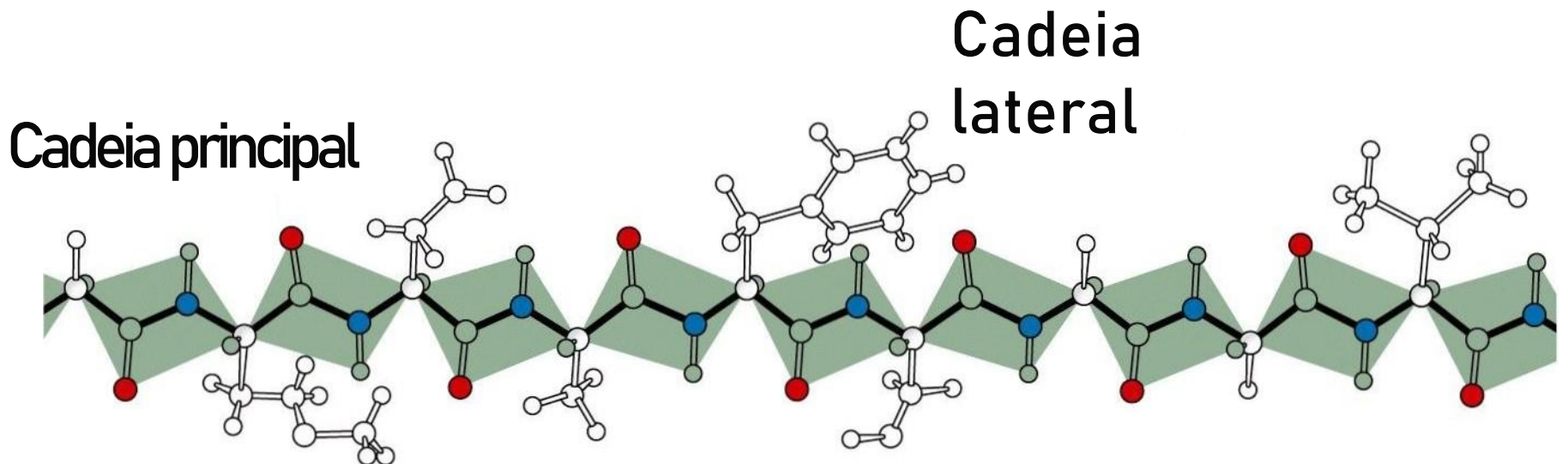


A ligação peptídica (conecta dois resíduos adjacentes na cadeia principal)



Deslocalização eletrônica (ressonância) na ligação peptídica resulta numa ligação C-N com caráter parcialmente dupla e uma alta barreira energética de rotação

A ligação
peptídica
(conecta dois
resíduos
adjacentes na
cadeia principal)



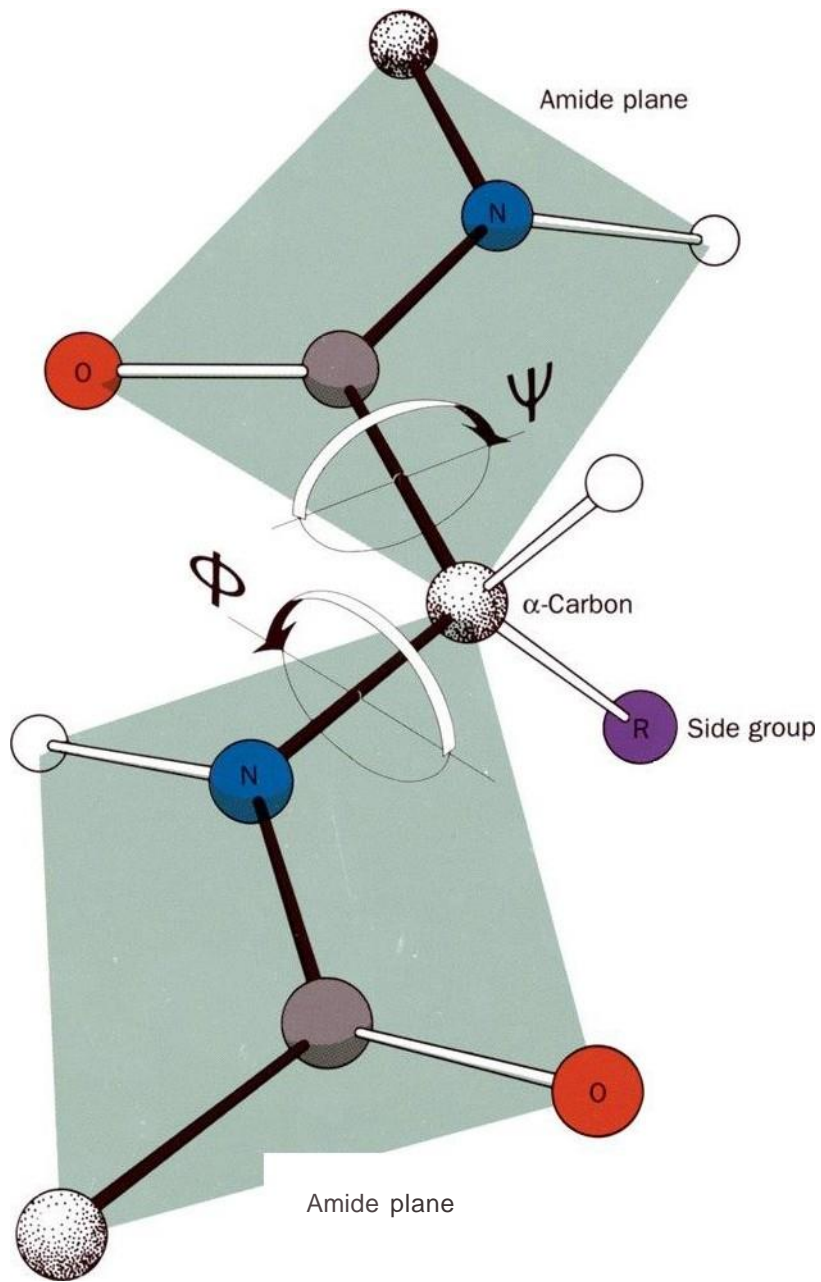


TABLE 8-1 van der Waals Distances for Interatomic Contacts

Contact Type	Normally Allowed (Å)	Outer Limit (Å)
H...H	2.0	1.9
H...O	2.4	2.2
H...N	2.4	2.2
H...C	2.4	2.2
O...O	2.7	2.6
O...N	2.7	2.6
O...C	2.8	2.7
N...N	2.7	2.6
N...C	2.9	2.8
C...C	3.0	2.9
C...CH ₂	3.2	3.0
CH ₂ ...CH ₂	3.2	3.0

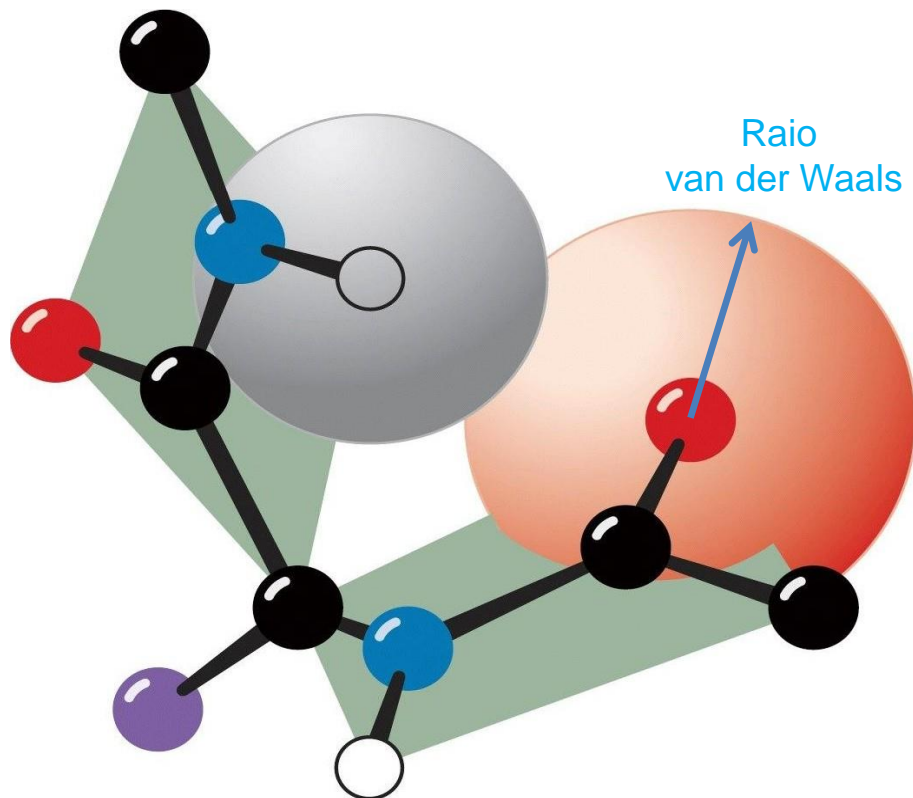


Diagrama Ramachandran

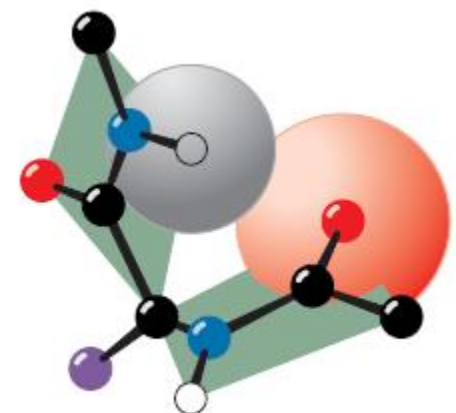
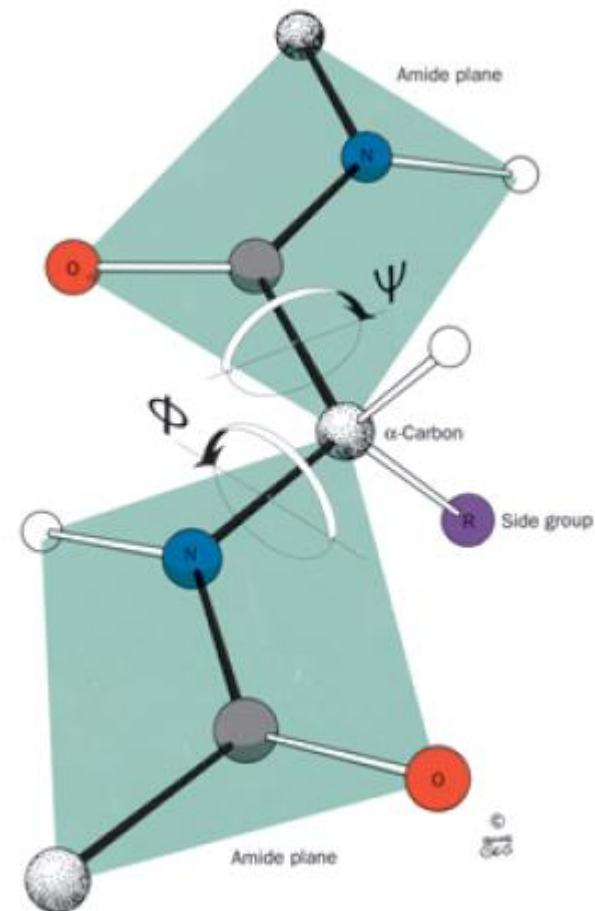
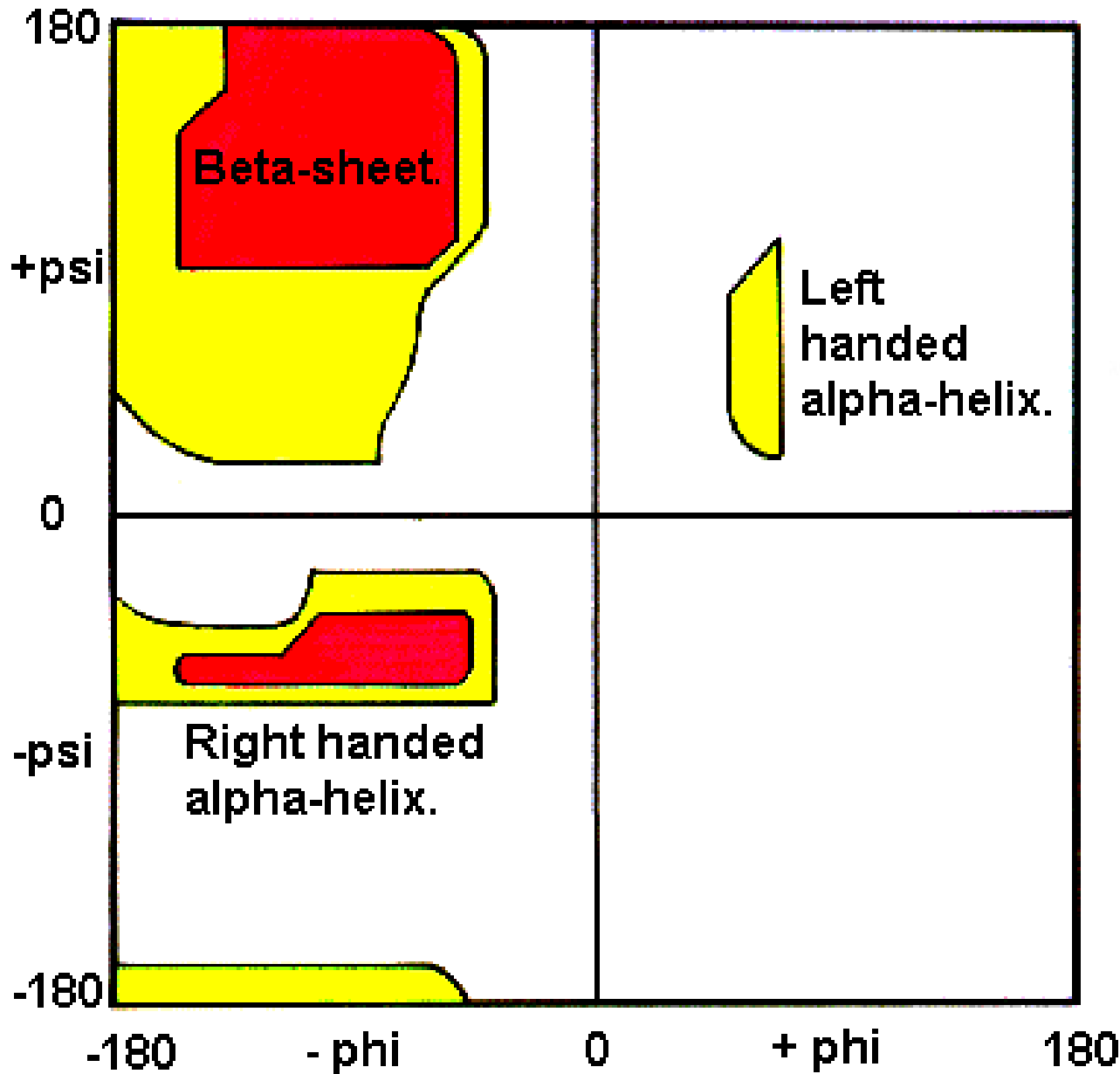


FIGURE 8-6 Steric interference between adjacent residues.

Diagrama Ramachandran

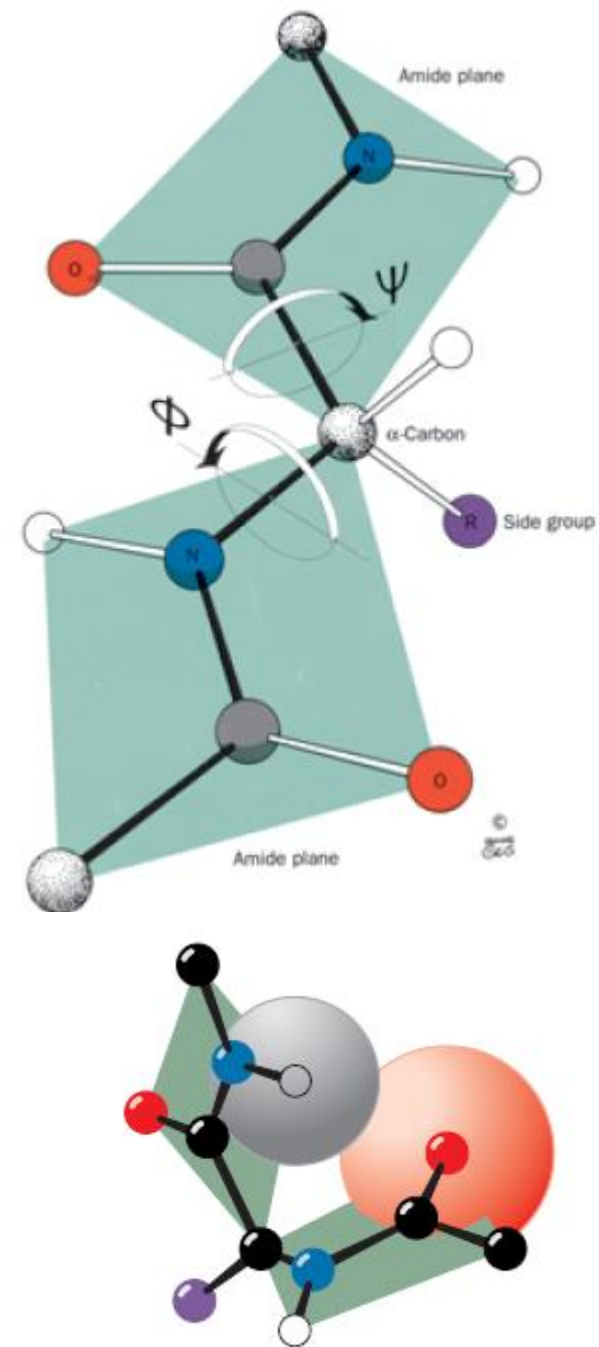
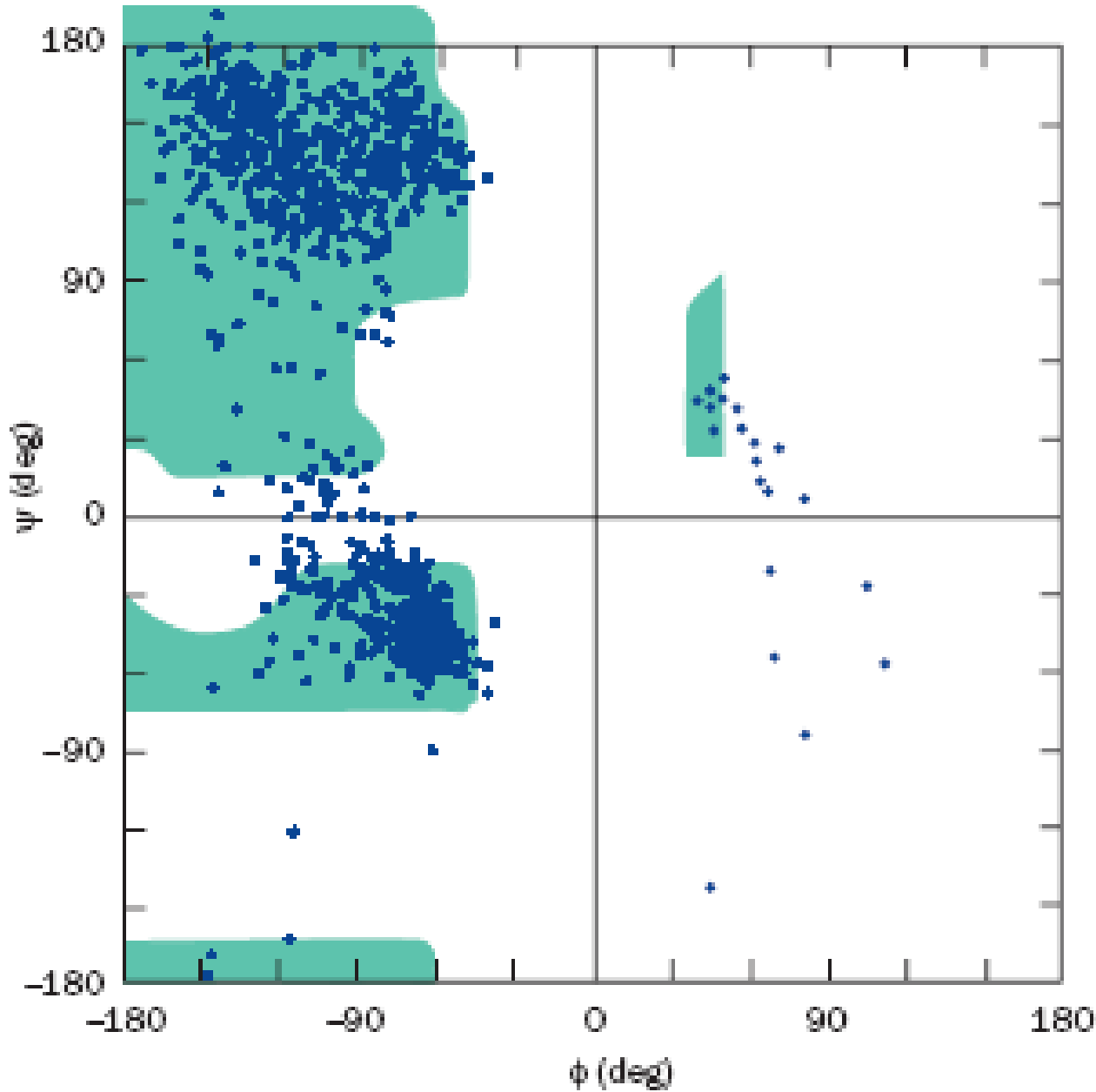
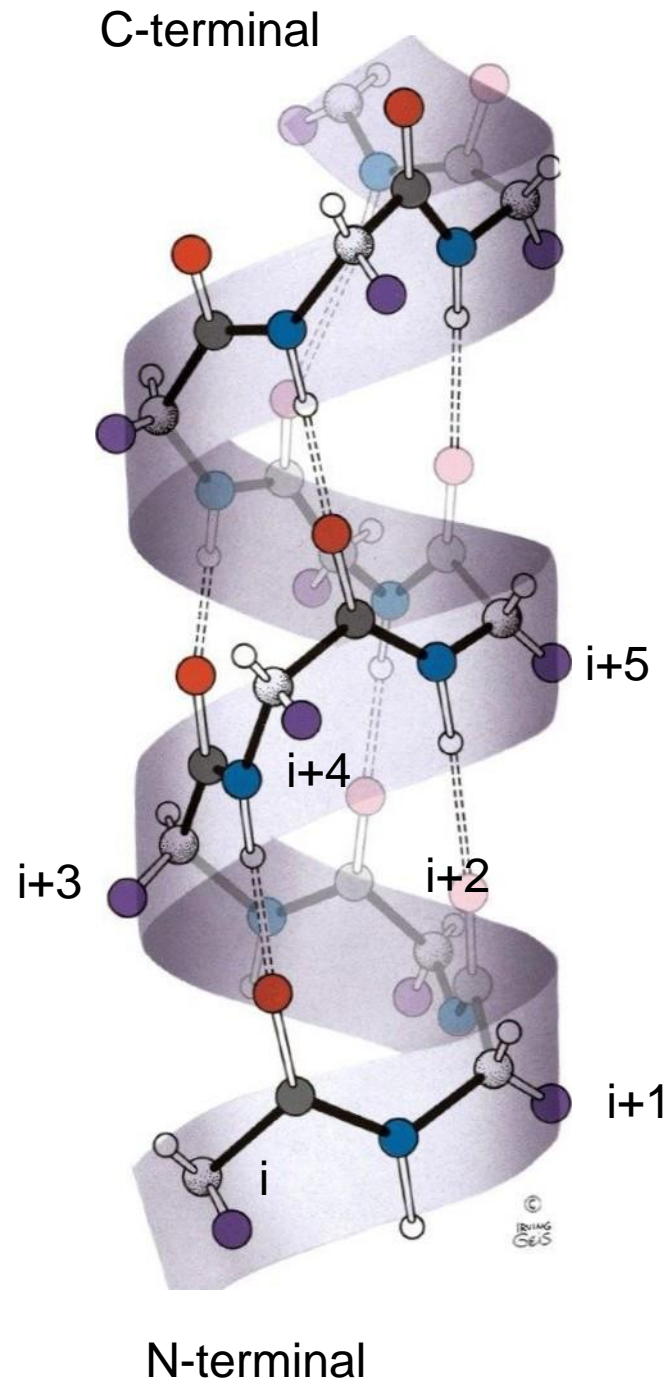


FIGURE 8-6 Steric interference between adjacent residues.

Alfa-hélice de mão direita

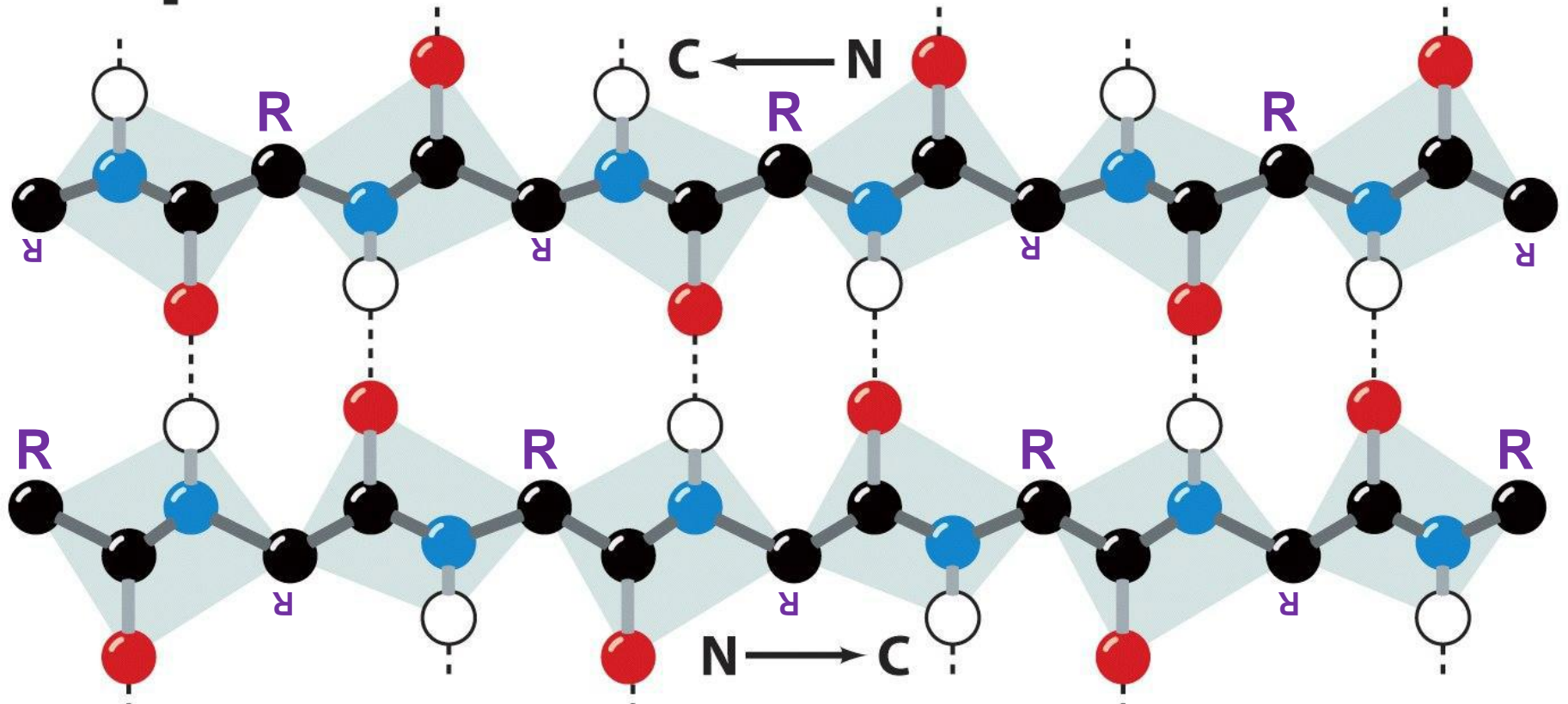
- Estabilizada por pontes de hidrogênio entre grupos carbonila e grupos NH-amídicas em posições i e $i+4$
- Cadeias laterais (roxas) apontam radialmente por fora
- Esta estrutura satisfaz os requerimentos de formação de pontes de hidrogênio da cadeia principal em ambientes aquosos e hidrofóbicos
- Esta estrutura pode acomodar quase qualquer sequência (porem, alguns aminoácidos como gly e pro são raramente encontrados em alfa-hélices)



Fitas e folhas beta

- Folha estabilizada por pontes de hidrogênio ente grupos NH-amídicos e carbonilas em fitas beta adjacentes
- Cadeias laterias (R) apontam alternadamente por cima e por baixo do plano da folha.
- Esta estrutura satisfaz os requerimentos de formação de pontes de hidrogênio da cadeia principal em ambientes aquosos e hidrofóbicos
- Esta estrutura pode acomodar quase qualquer sequência, porem pro e gly são menos representados

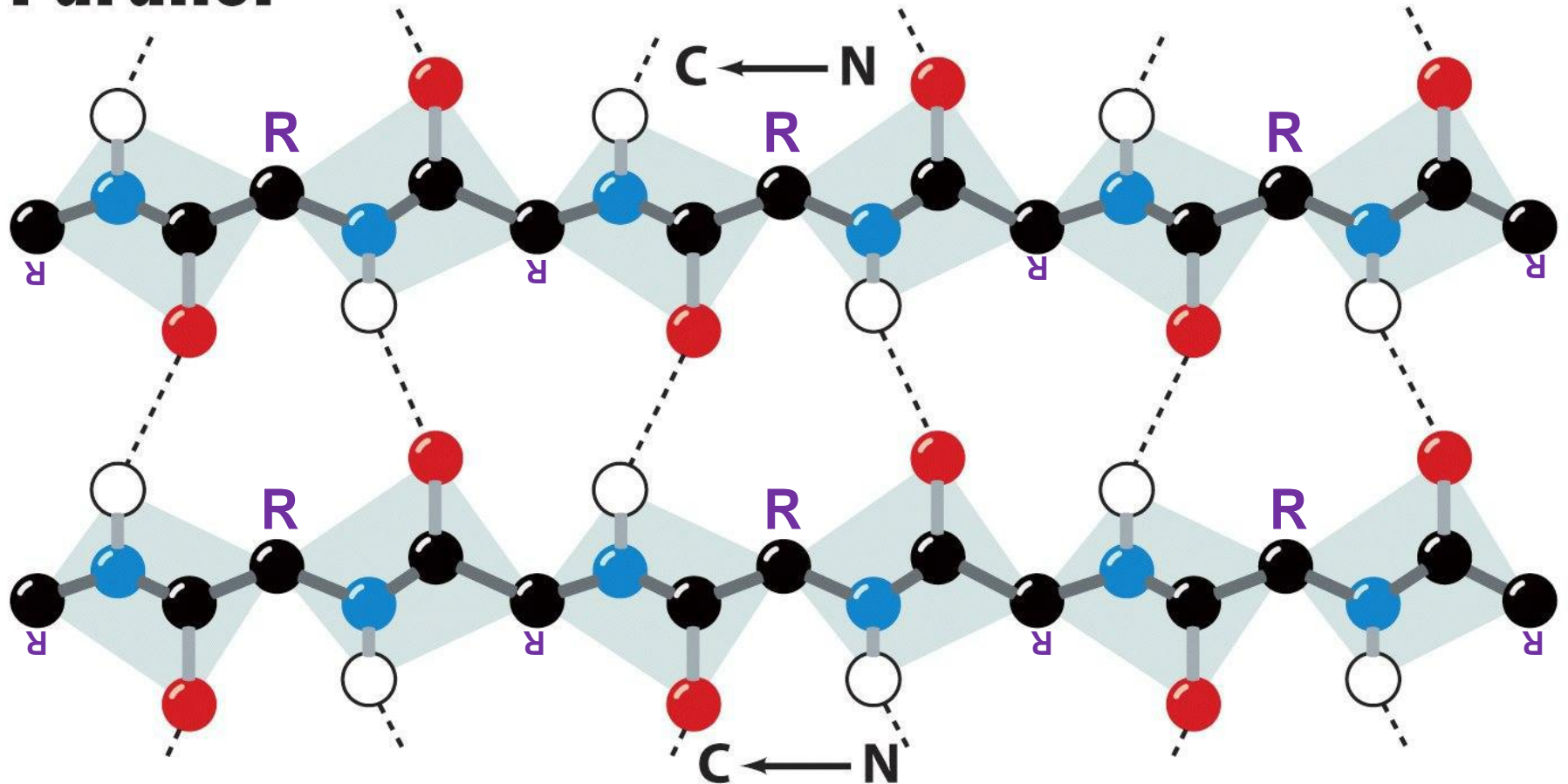
Antiparallel



Fitas e folhas beta

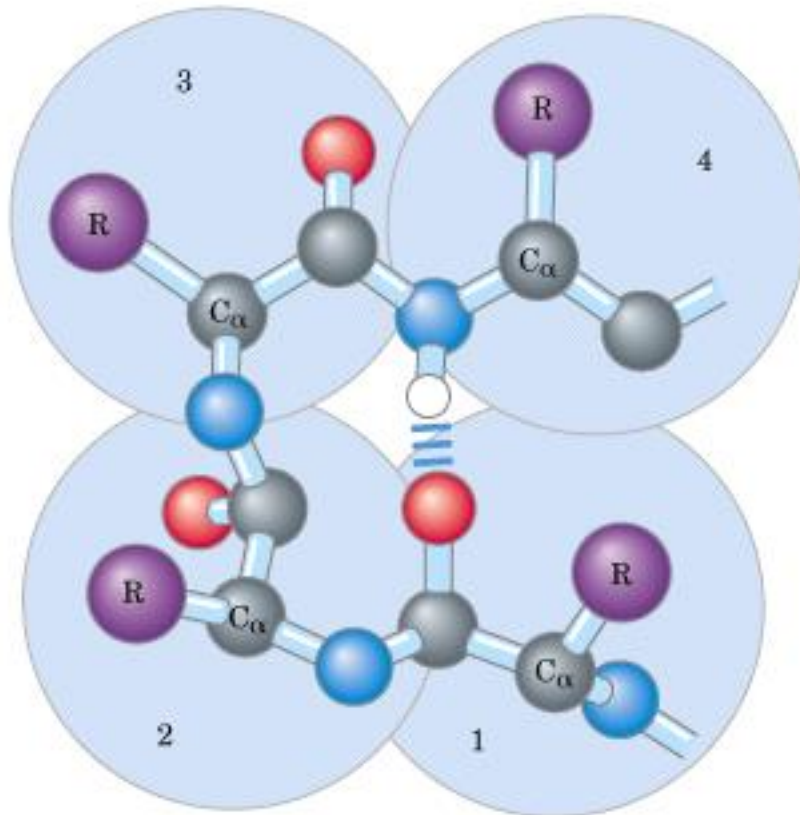
- Folha estabilizada por pontes de hidrogênio ente grupos NH-amídicos e carbonilas em fitas beta adjacentes
- Cadeias laterias (R) apontam alternadamente por cima e por baixo do plano da folha.
- Esta estrutura satisfaz os requerimentos de formação de pontes de hidrogênio da cadeia principal em ambientes aquosos e hidrofóbicos
- Esta estrutura pode acomodar quase qualquer sequência , porem pro e gly são menos representados

Parallel

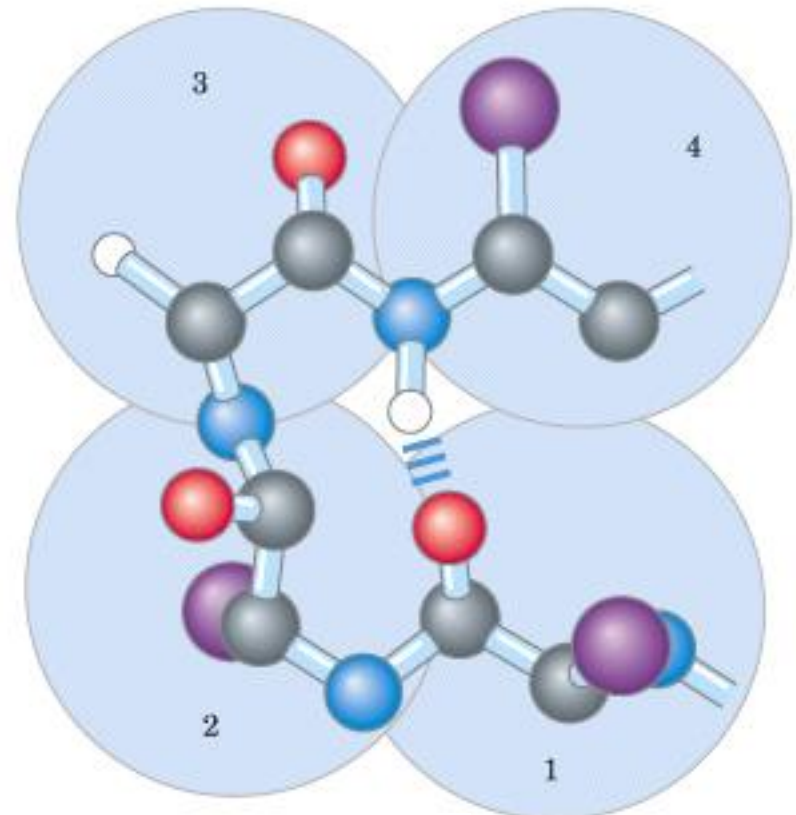


VOLTAS

Notar ligação de hidrogênio entre os grupos peptídicos do 1° e 4° resíduo. É comum encontrar Pro e Gly nestas estruturas.



Tipo I



Tipo II

Beta voltas

Estrutura de queratina de cabelo.

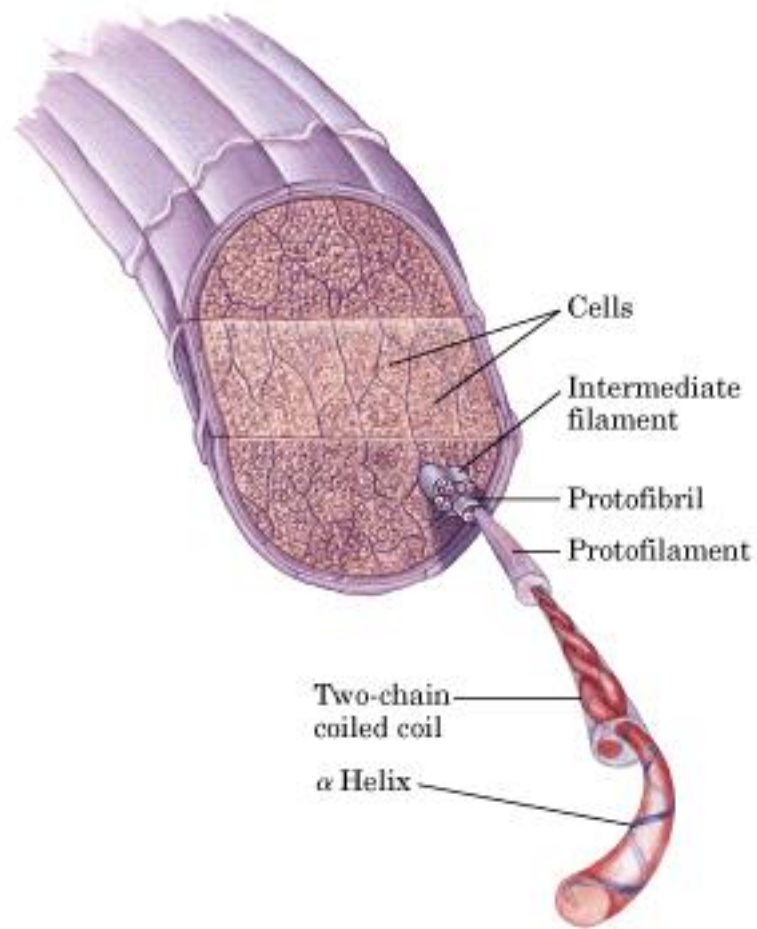
Keratin α helix — 

Two-chain coiled coil — 

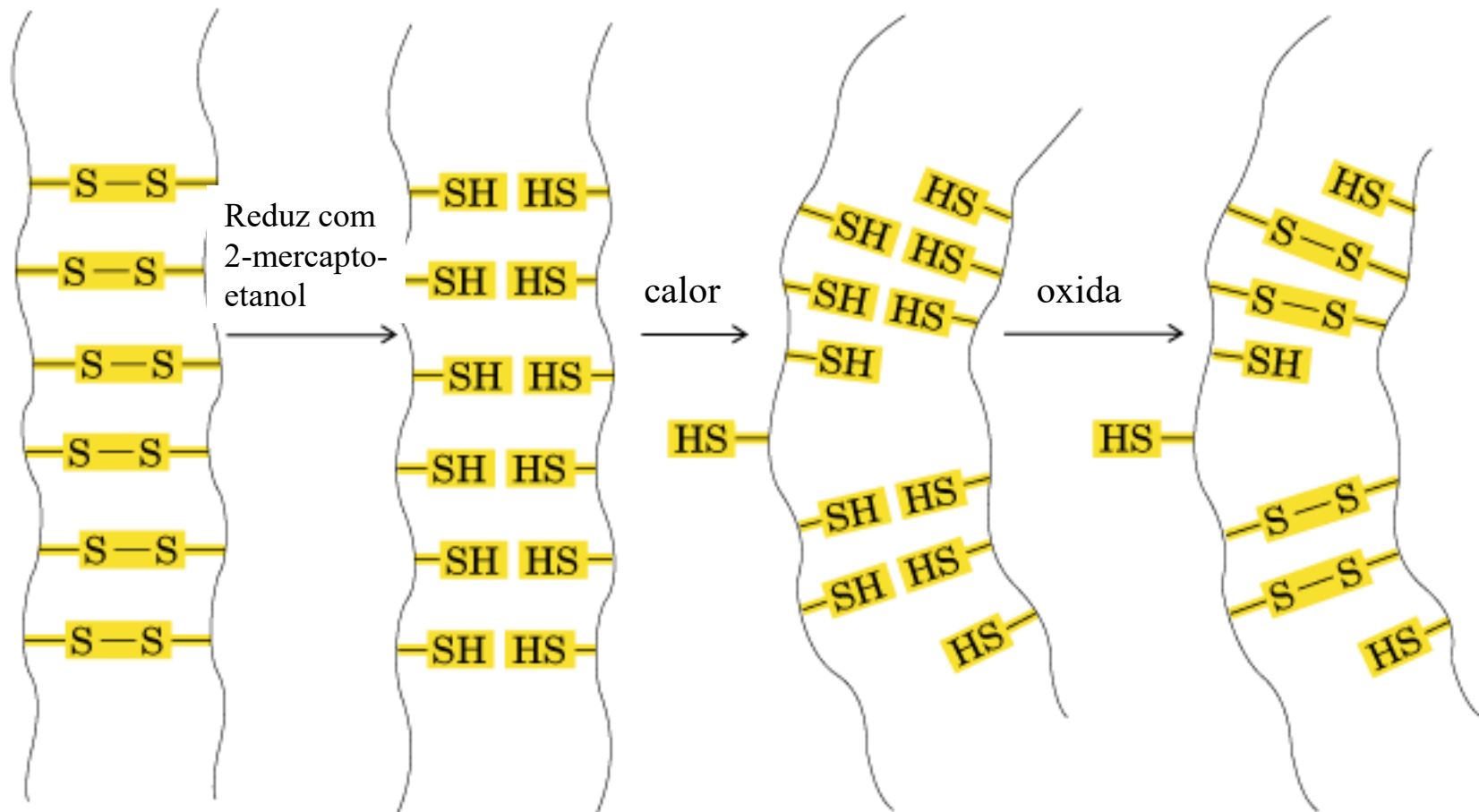
Protofilament {  } 20–30 Å

Protofibril {  } 40–50 Å

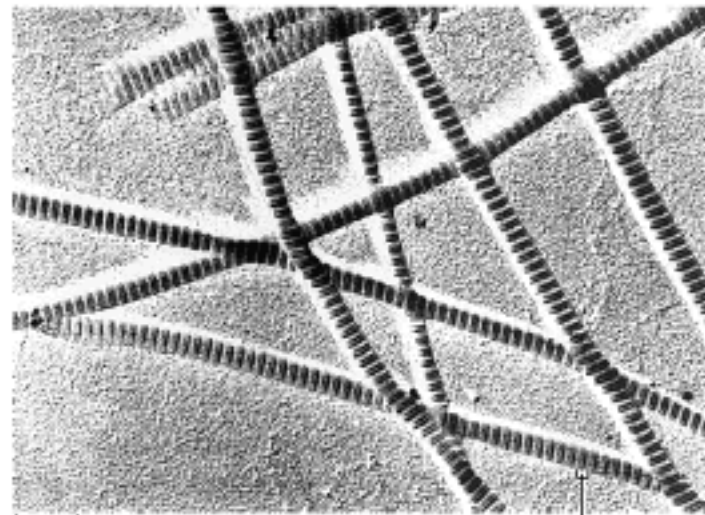
(a)



Cross section of a hair
(b)



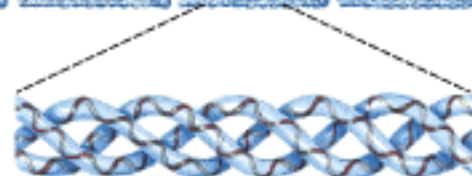
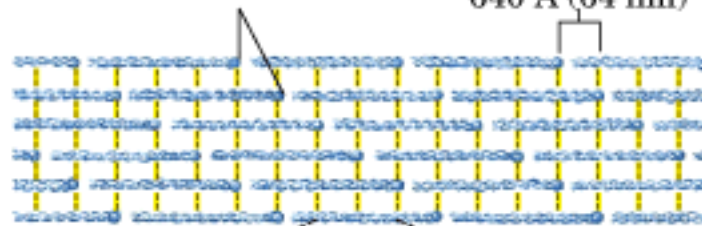
Colágeno: componente de tecido conectivo (tendão, cartilagem, matriz orgânica de osso)



250
nm

Heads of collagen
molecules

Cross-striations
640 Å (64 nm)



Section of collagen
molecule

Tripla hélice

Colágeno: um triplo-hélice de cadeias com 3 aminoácidos por volta.
Helicoidal com sentido de mão esquerda.
Super-helicoidal com sentido de mão direita.
Composição: 35% Gly, 11% Ala, 21% Pro ou hidroxí-Pro.
Com repetições Gly-X-Pro ou Gly-X-hidroxí-Pro



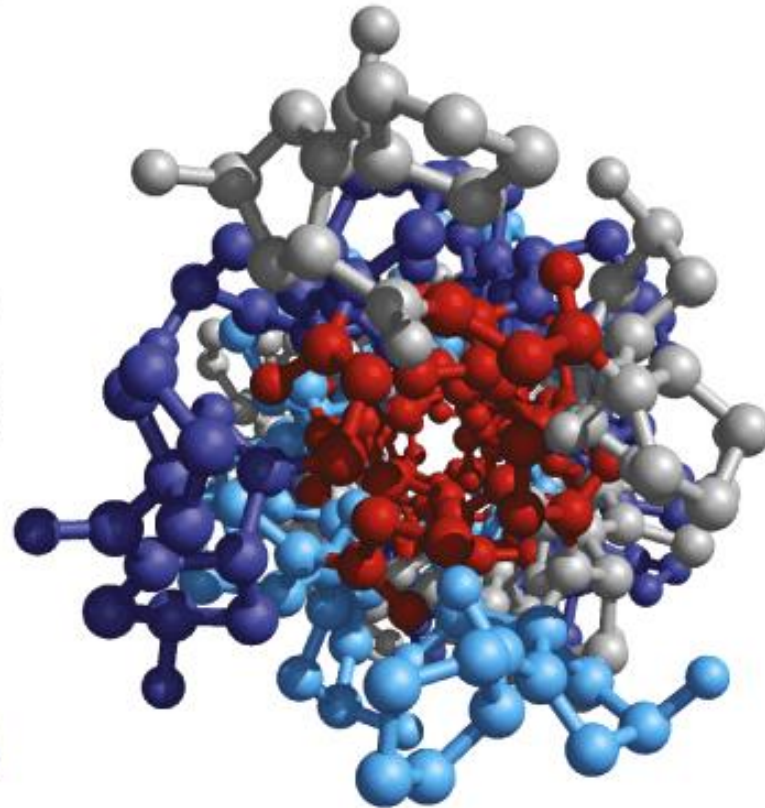
(a)



(b)

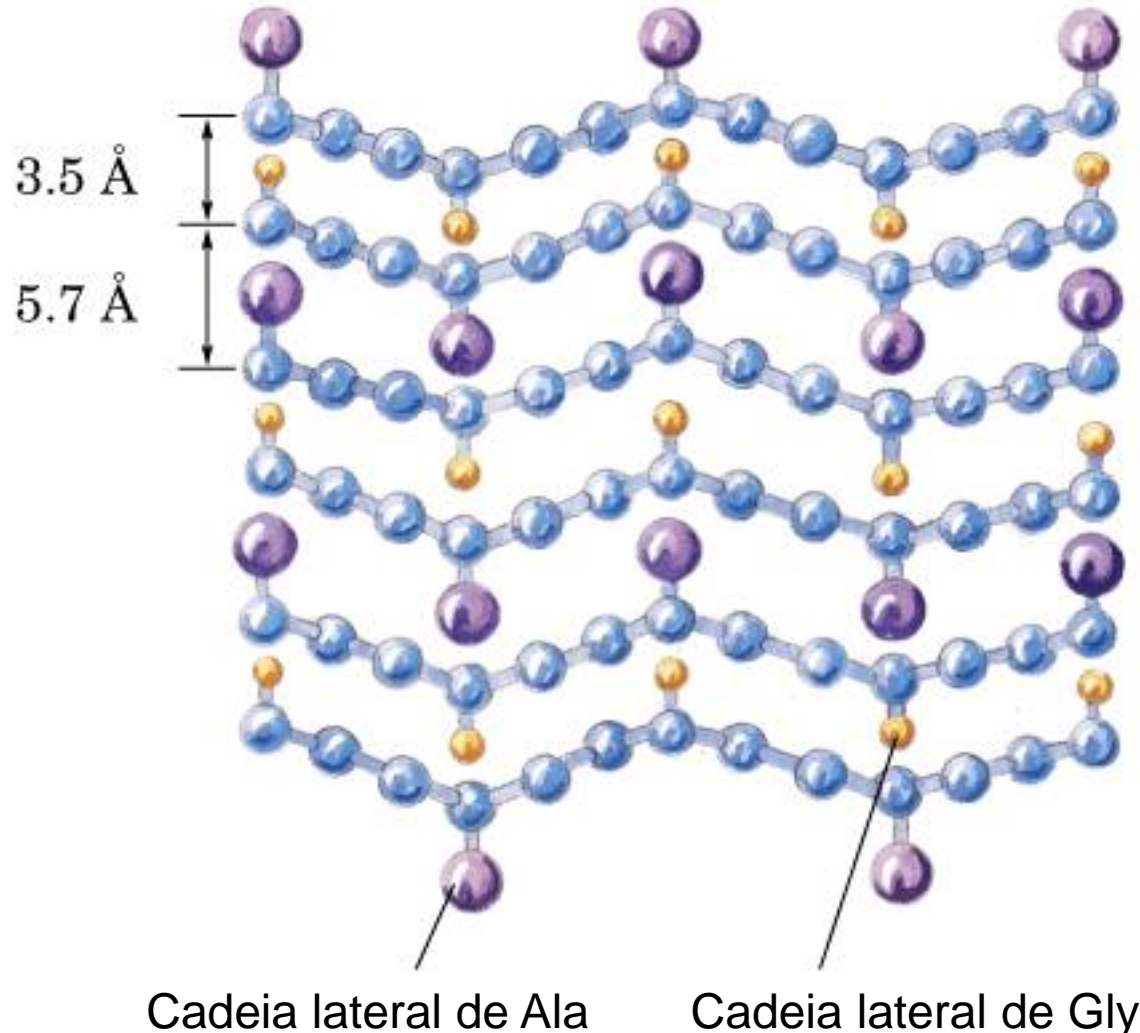


(c)



(d)

Estrutura de fibroína de bicho de seda: estrutura folha beta antiparalela rica em Gly e Ala.



Proteins. 2001 Aug 1;44(2):119-22.

Silk fibroin: structural implications of a remarkable amino acid sequence.

Zhou, Confalonieri, Jacquet, Perasso, Li, Janin

Abstract

The amino acid sequence of the heavy chain of *Bombyx mori* silk fibroin was derived from the gene sequence. The 5,263-residue (391-kDa) polypeptide chain comprises **12 low-complexity "crystalline" domains made up of Gly-X repeats and covering 94% of the sequence; X is Ala in 65%, Ser in 23%, and Tyr in 9% of the repeats.** The remainder includes a nonrepetitive 151-residue header sequence, 11 nearly identical copies of a 43-residue spacer sequence, and a 58-residue C-terminal sequence. The header sequence is homologous to the N-terminal sequence of other fibroins with a completely different crystalline region. In *Bombyx mori*, each crystalline domain is made up of subdomains of approximately 70 residues, which in most cases begin with repeats of the GAGAGS hexapeptide and terminate with the GAAS tetrapeptide. Within the subdomains, the Gly-X alternance is strict, **which strongly supports the classic Pauling-Corey model, in which beta-sheets pack on each other in alternating layers of Gly/Gly and X/X contacts.**

Dimensões de albumina sérica humana (585 aminoácidos, 64500 kDa)

β Conformation
 $2,000 \times 5 \text{ \AA}$

α Helix
 $900 \times 11 \text{ \AA}$

Native globular form
 $130 \times 30 \text{ \AA}$

Estrutura de Proteínas Globulares e Estrutura Terciária

Interações fracas importantes para manter a estrutura de proteínas em solução aquosa

Pontes de Hidrogênio

- entre grupos neutros
- entre ligações peptídicas

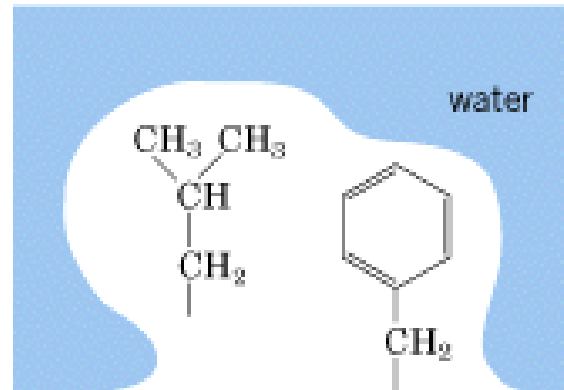


Interações Iônicas e Dipolares

- atração
- repulsão



Interações hidrofóbicos



Interações van der Waals

Qualquer par de átomos próximos

dipole interactions

1 r^2 to r^4 8

ionic interactions

3 r^1 10

van der Waals

0.5 r^6 1.5

covalent C,N,O bonds

40 fixed distance 130

RT

hydrogen bond

0.6 1.5 2.4 to 5 Å 7

covalent S-S bond

60 65

0.5

1

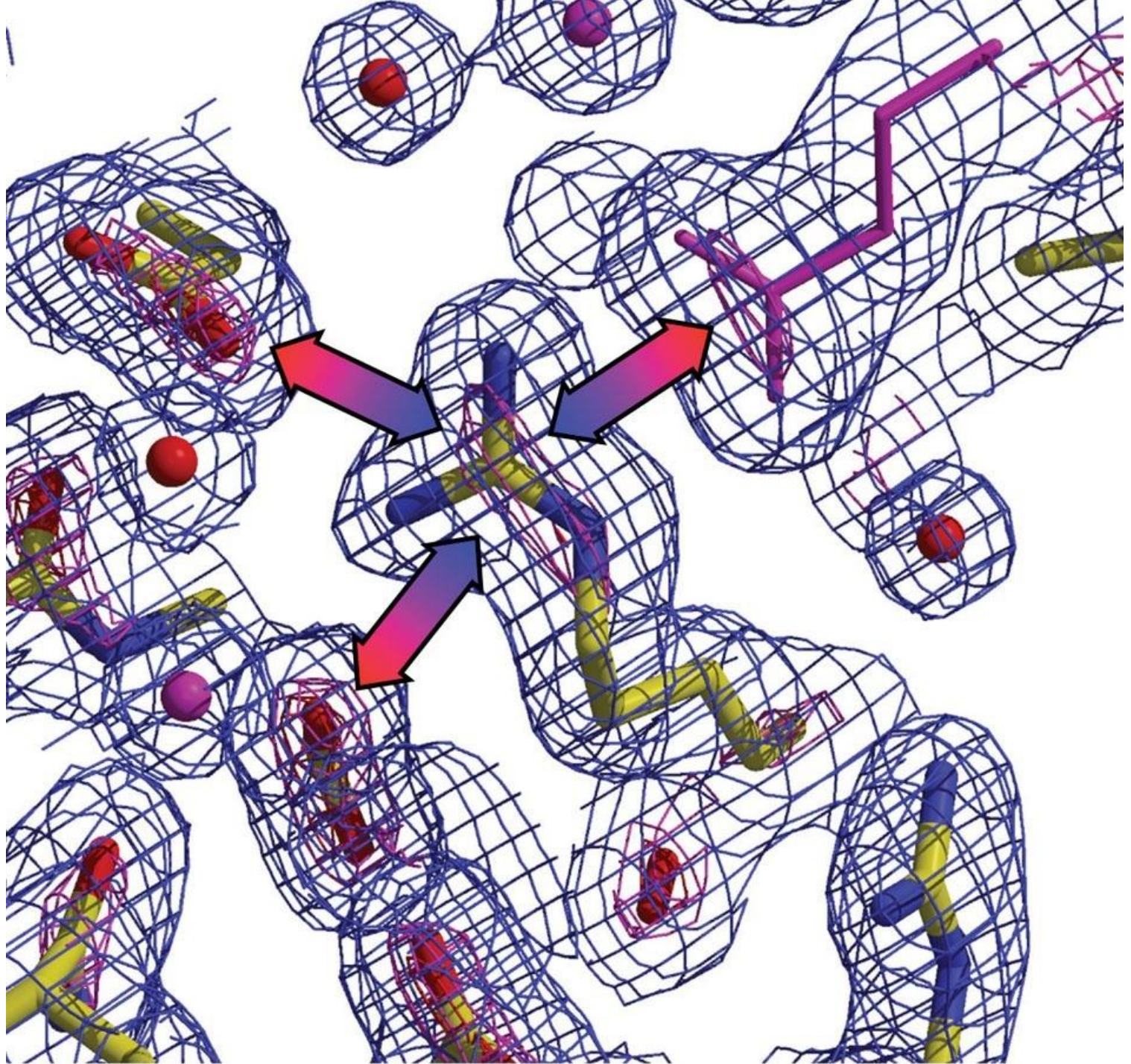
5

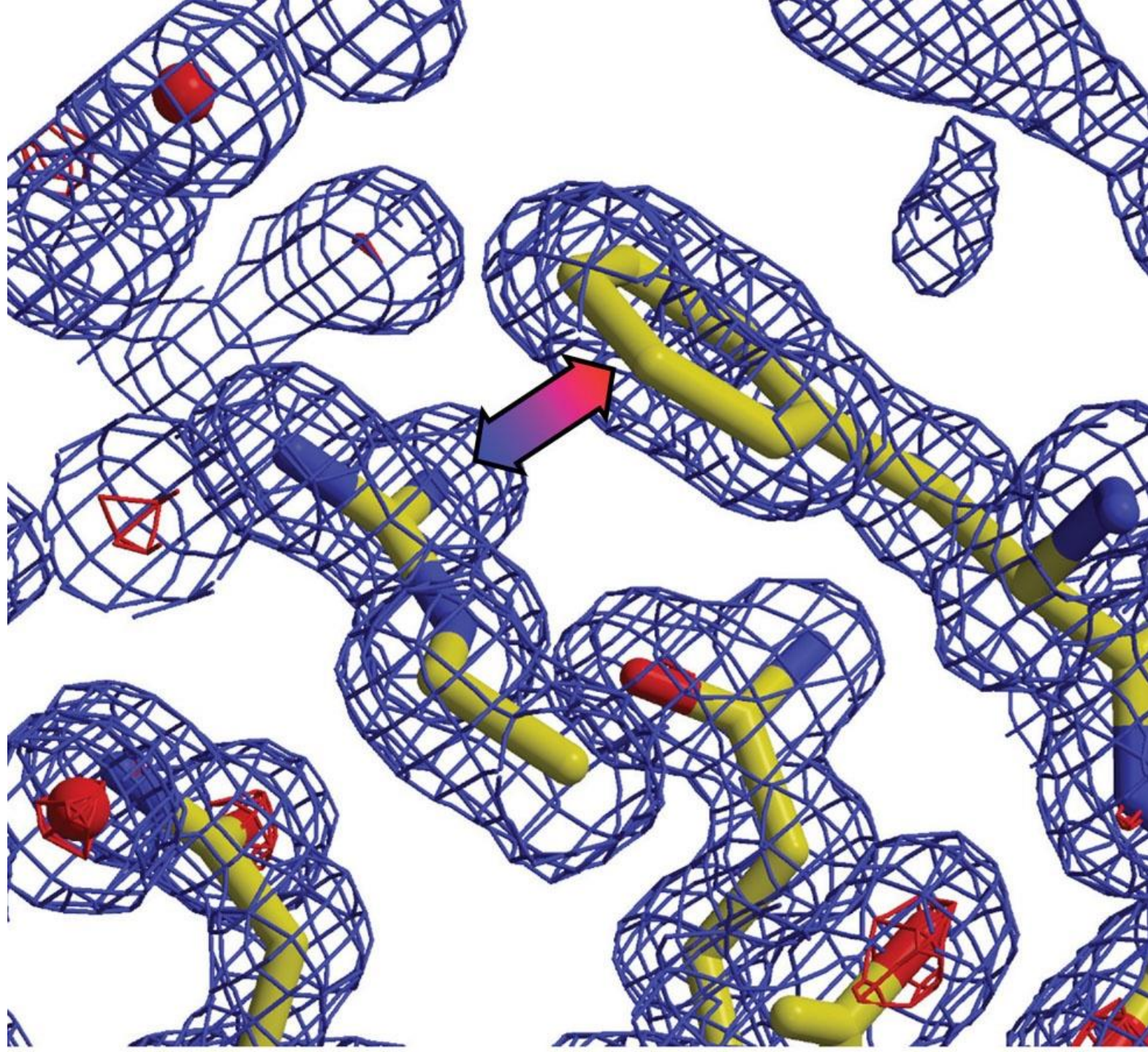
10

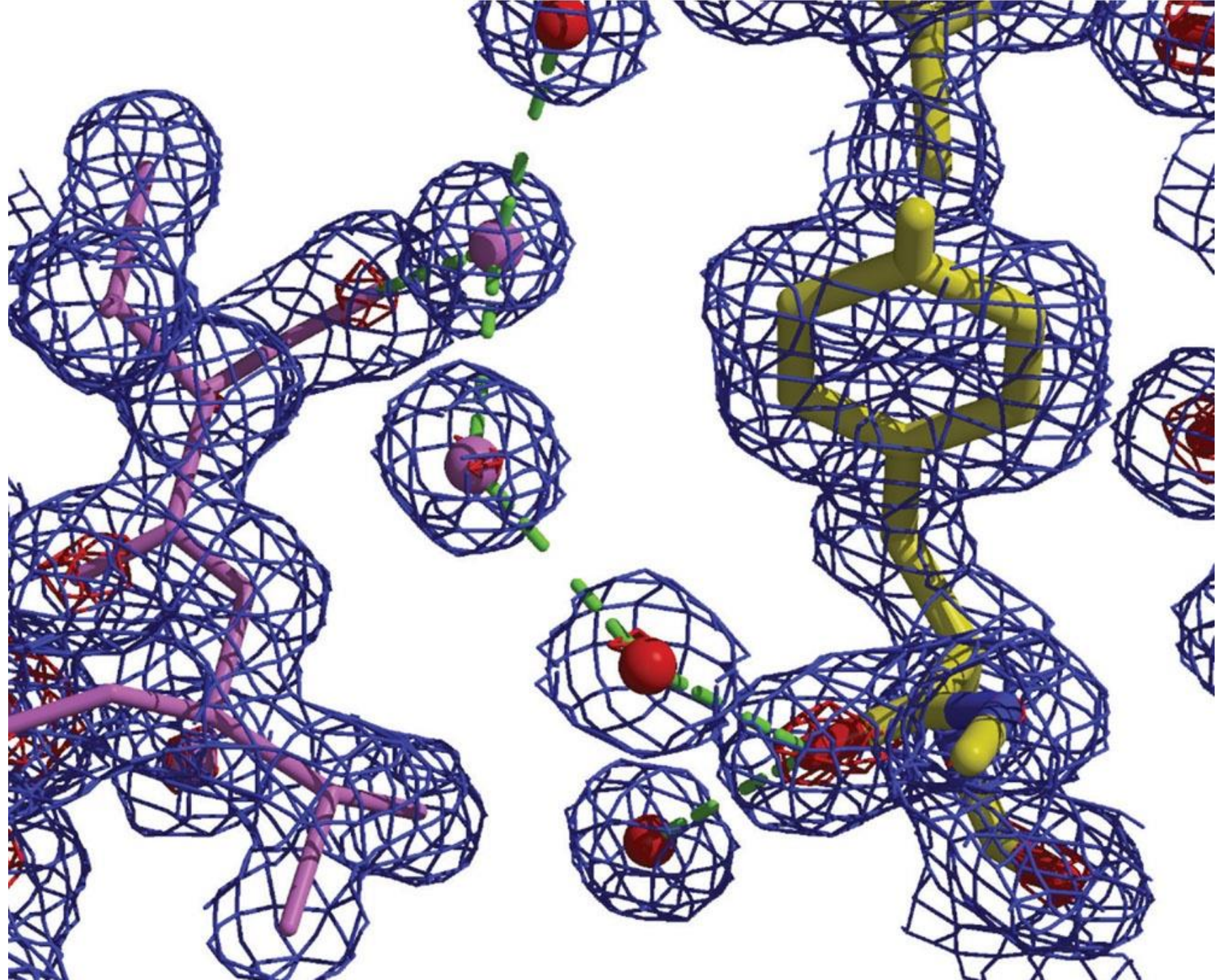
50

100

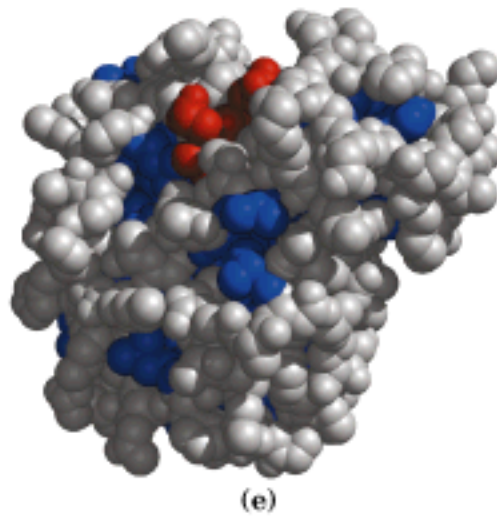
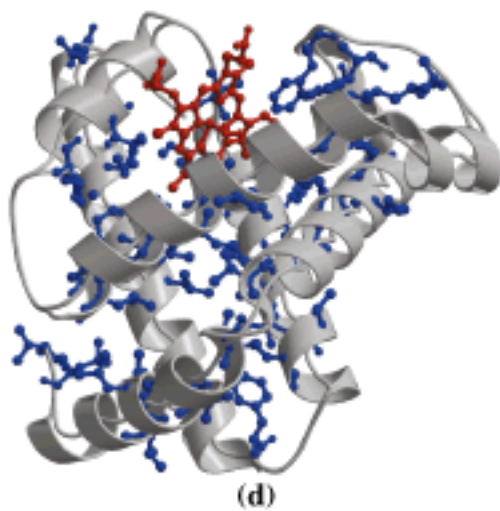
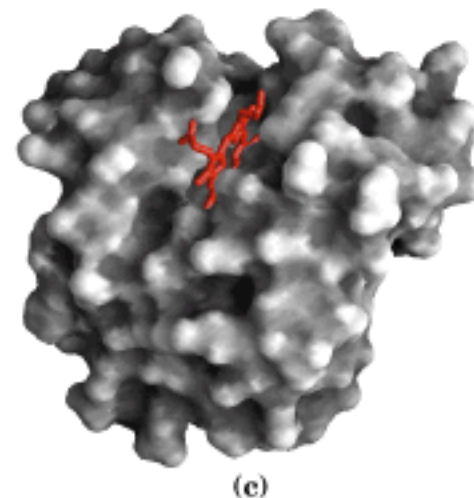
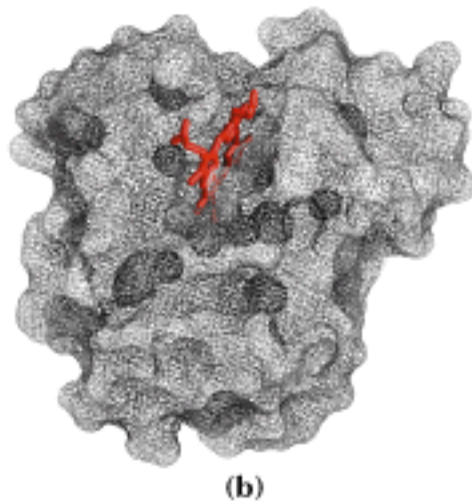
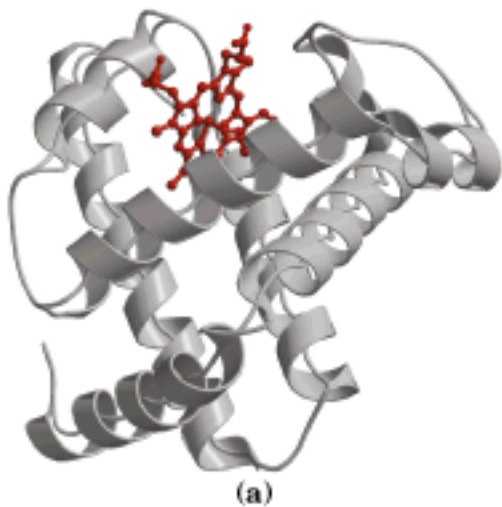
Bond energy (-kcal/mol)







Diferentes maneiras de representar a estrutura 3D de uma proteína (ex: mioglobina de baleia)

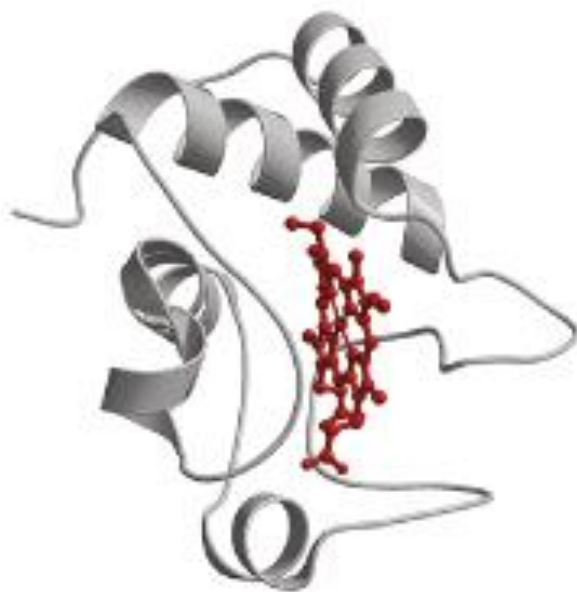
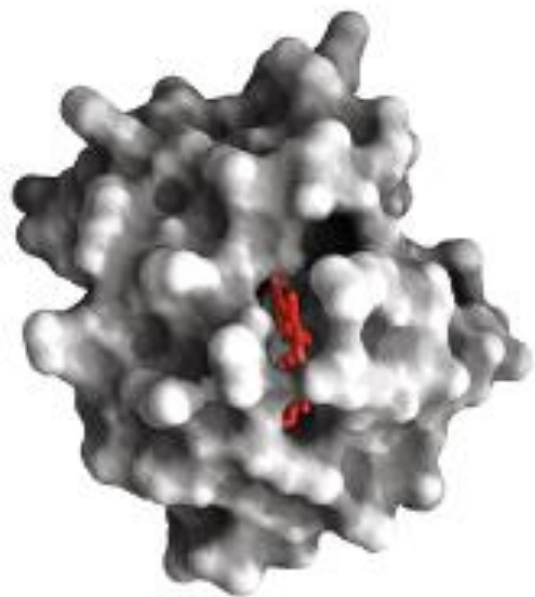


Topografia
de superfície

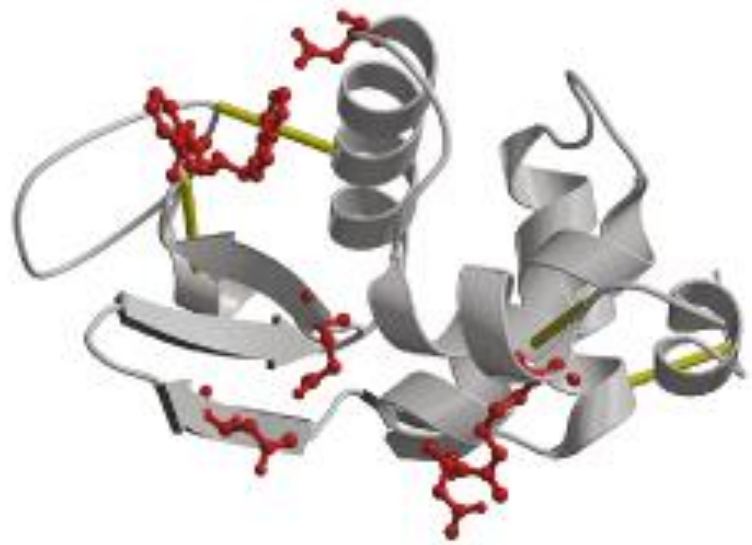
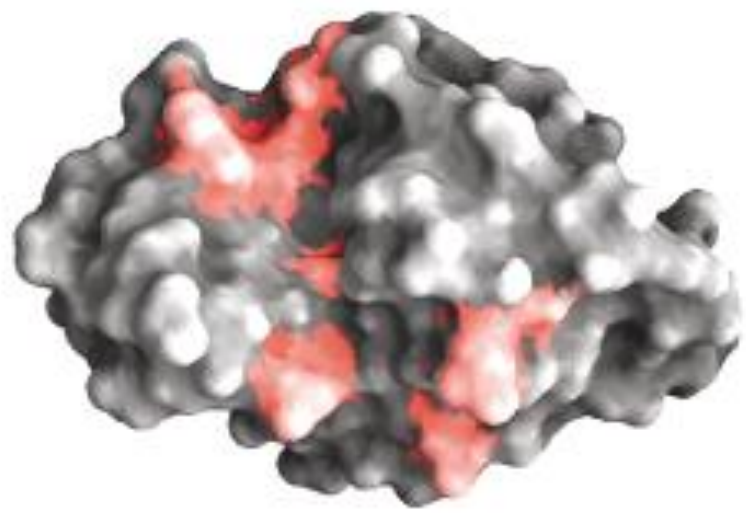
Fita

modelo "space-filling"

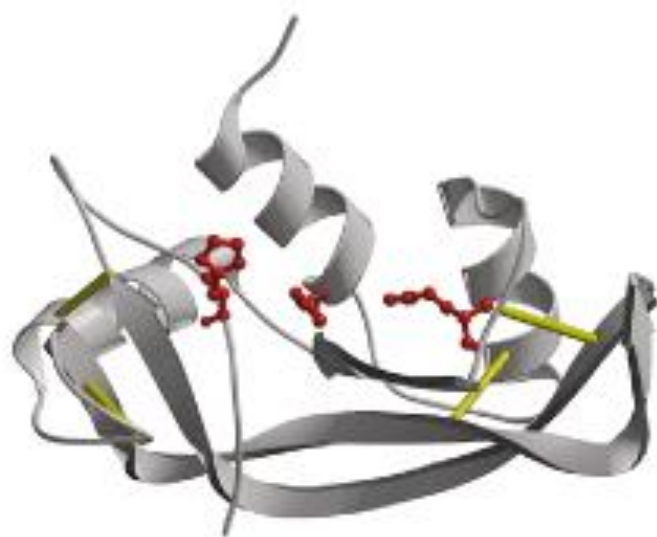
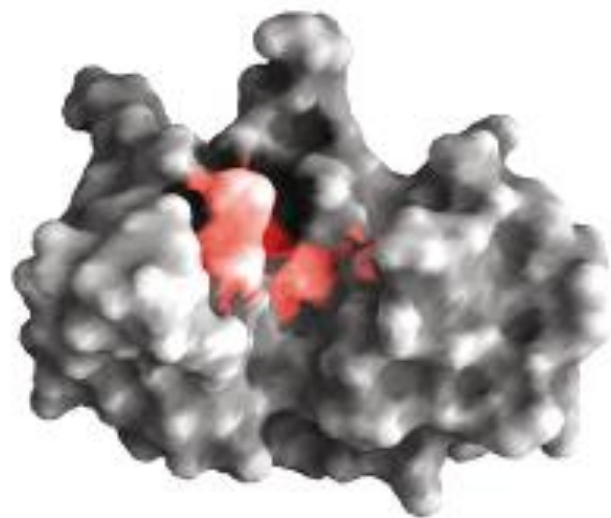
Notar localização dos aminoácidos hidrofóbicos em azul



Cytochrome c

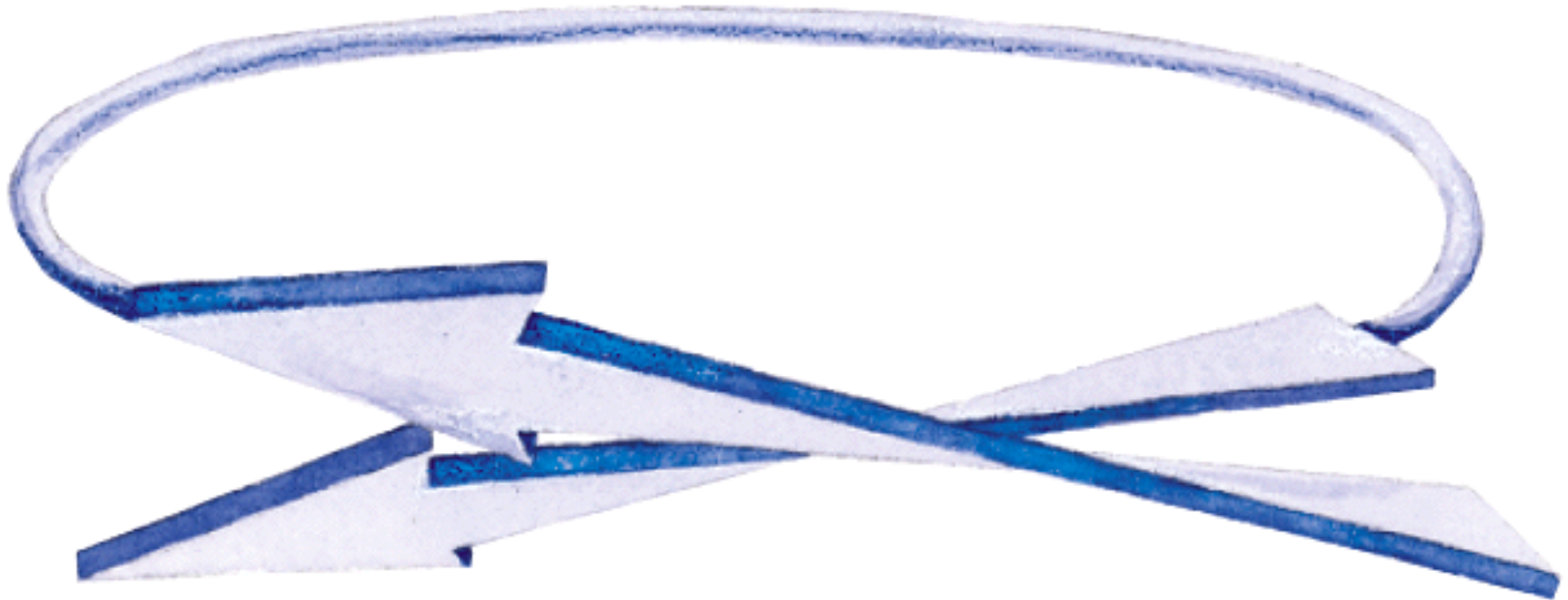


Lysozyme

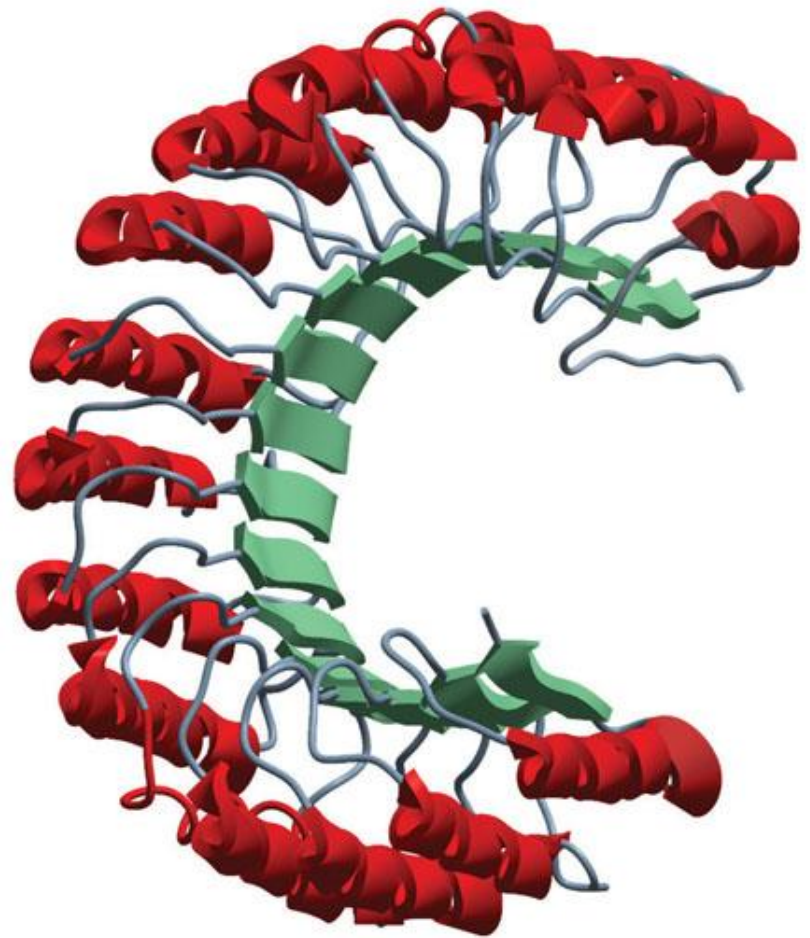
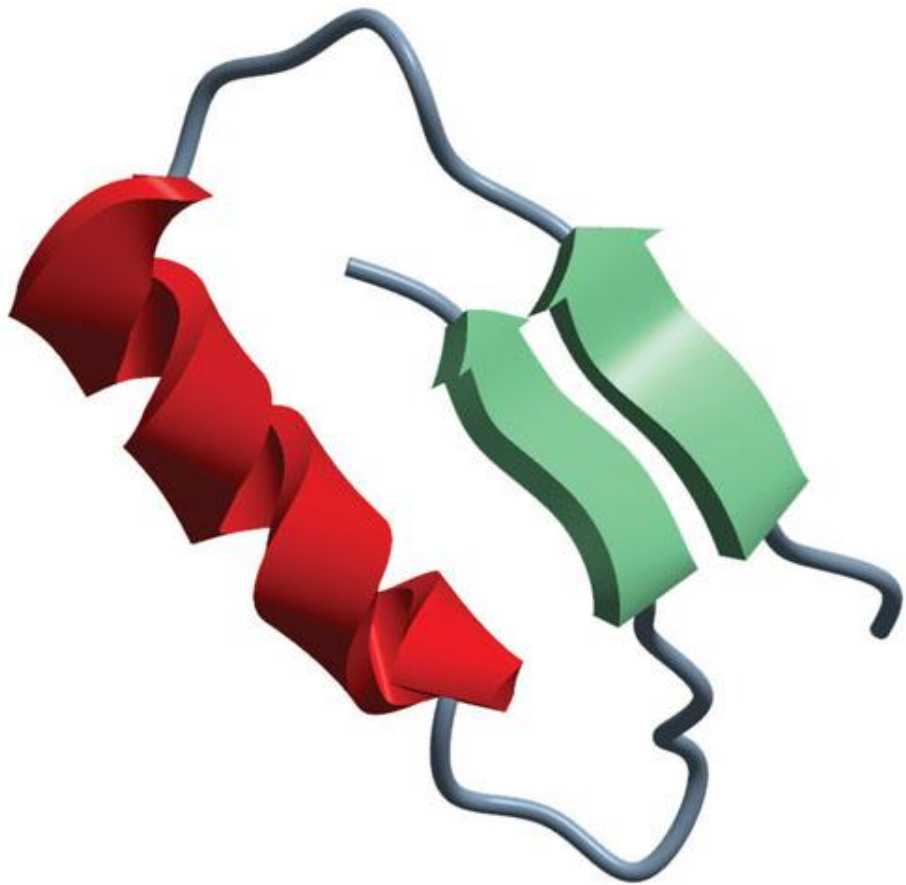


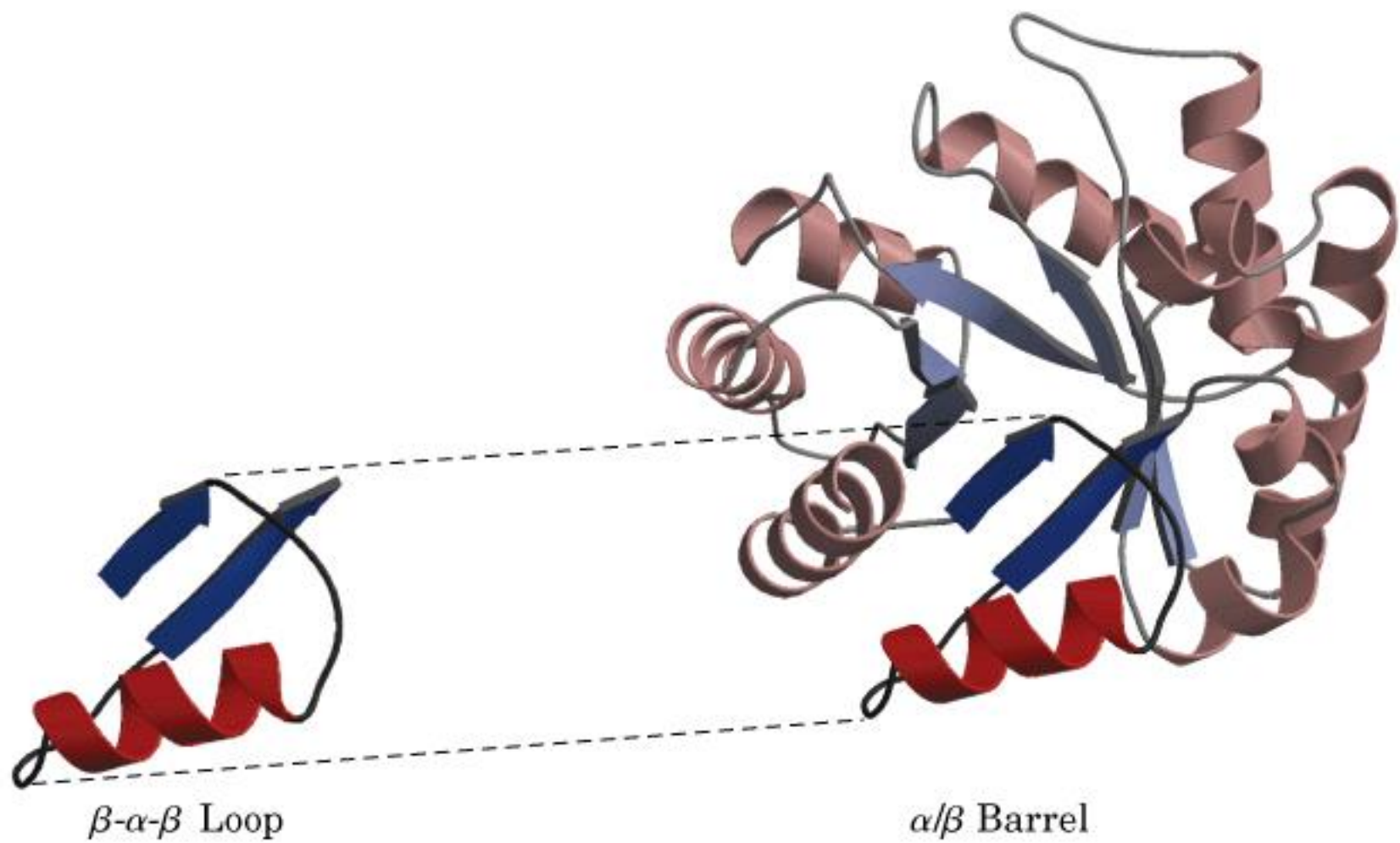
Ribonuclease

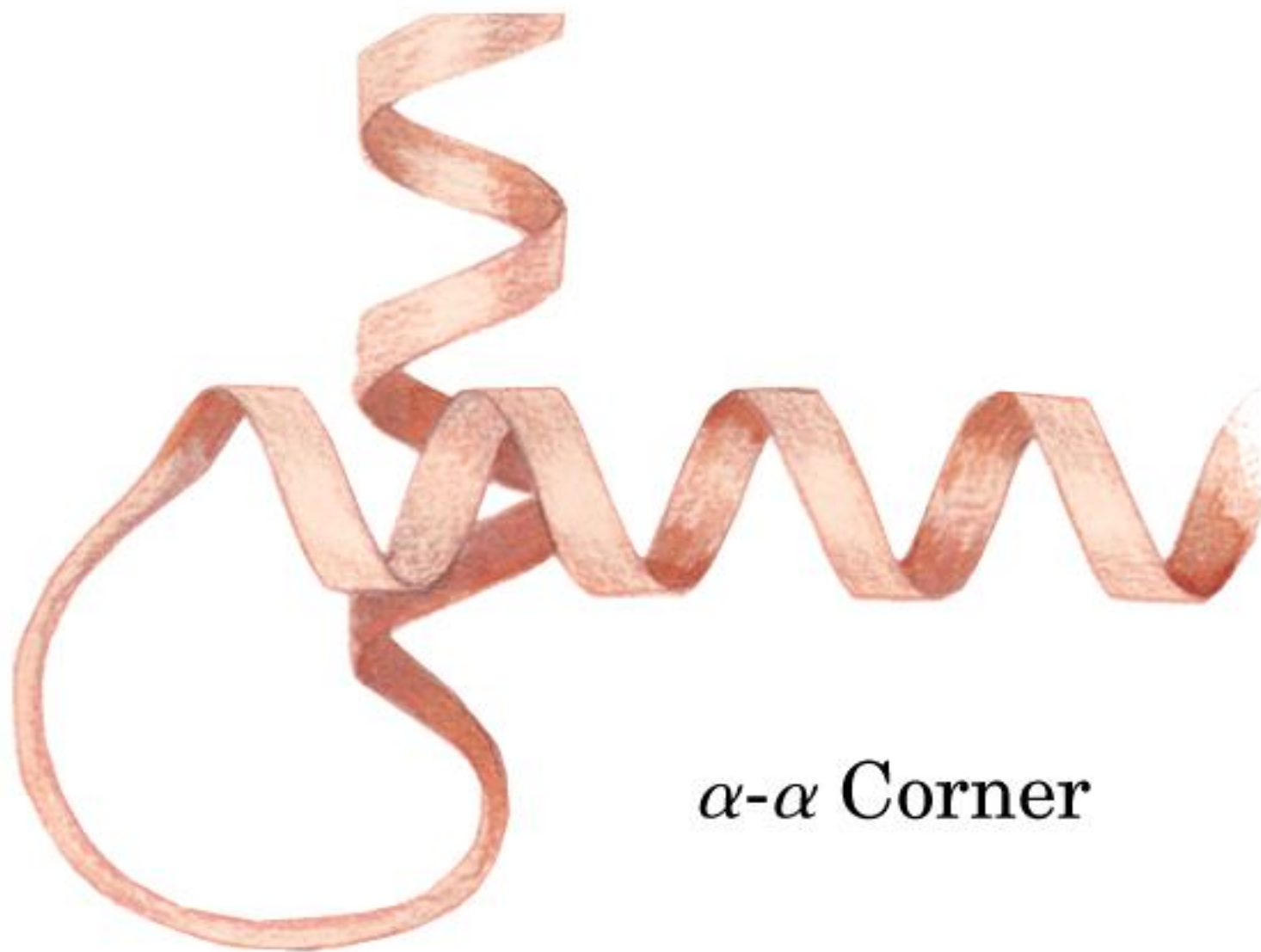
Estruturas super-secundárias



Conexões entre fitas beta
(mão direita)

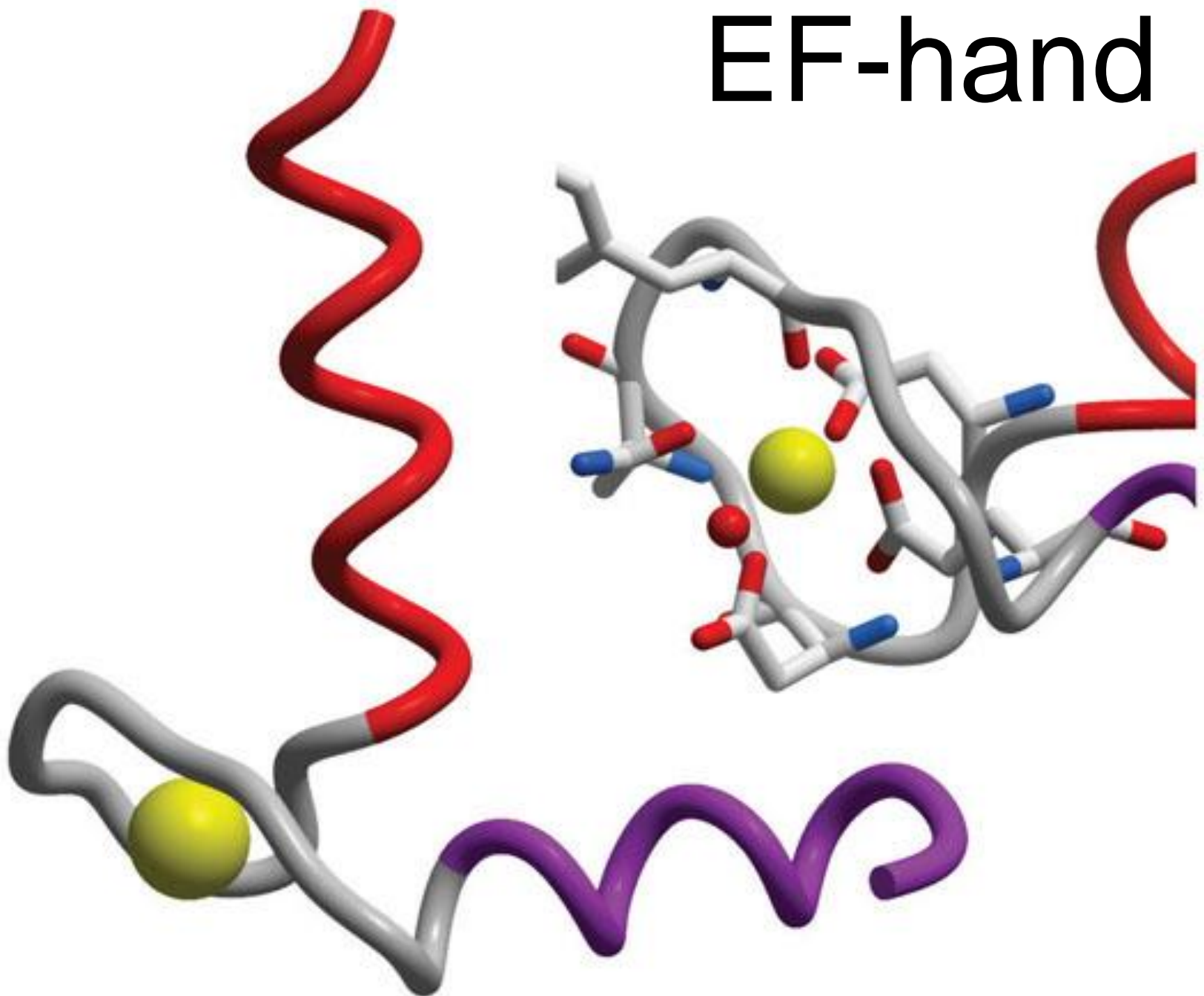






α - α Corner

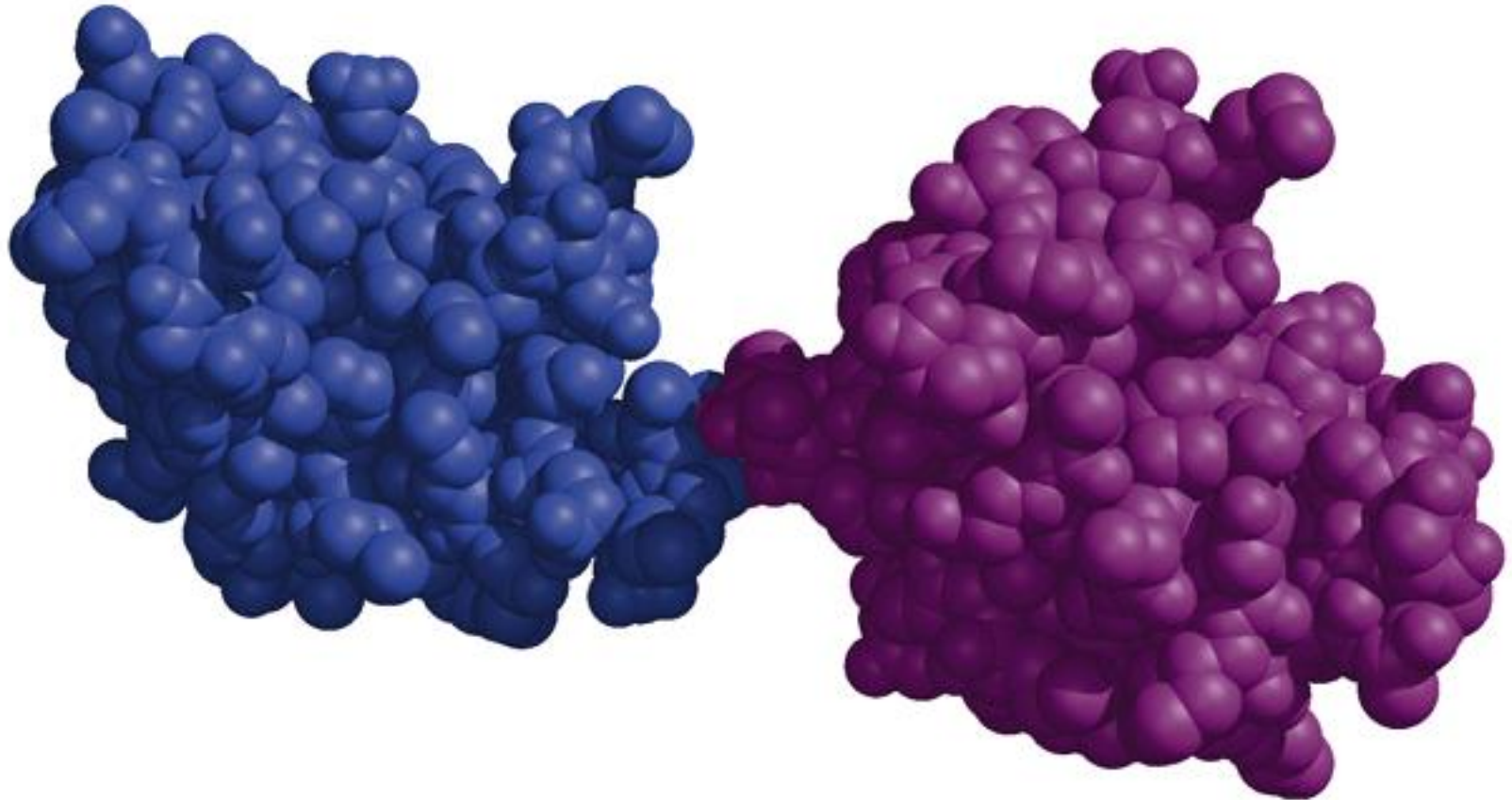
EF-hand





(d) β Barrel

**Domínios: Regiões “independentes”
de estrutura terciária dentro de uma proteína.**



CATH: Protein Structure Classification

<http://www.cathdb.info/>



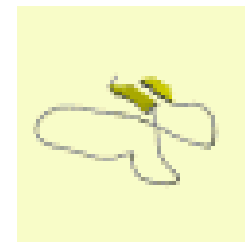
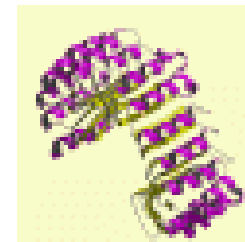
Sistemática de classificar
Estruturas Protéicas.

C = Classe

A = Arquitetura

T = Topologia

H = Homologia

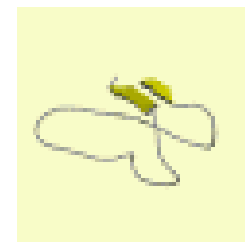
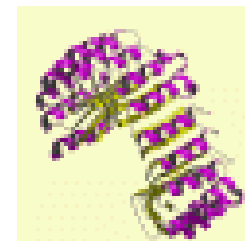


CATH: Protein Structure Classification

<http://www.cathdb.info/>

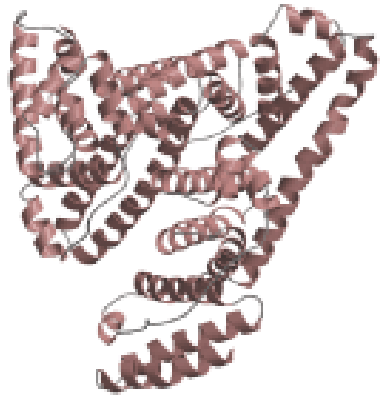


- **Classe**
 - Mais alto nível de classificação derivada de **conteúdo de estrutura secundária**.
- Há 4 classes:
- Tudo Alfa
 - Tudo Beta
 - Alfa-Beta
 - Poucas Estruturas Secundárias.

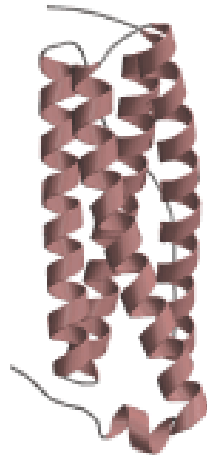


Tudo alfa

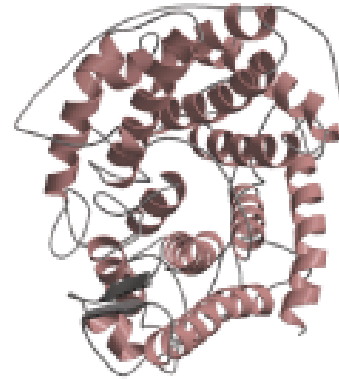
All α



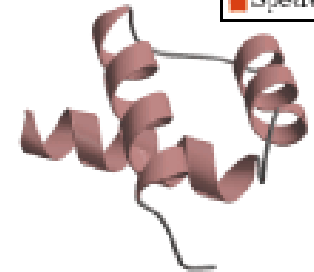
1a06
 Serum albumin
 Serum albumin
 Serum albumin
 Serum albumin
 Human (*Homo sapiens*)



1bcf
 Ferritin-like
 Ferritin-like
 Ferritin
 Bacterioferritin (cytochrome b_5)
Escherichia coli



1gai
 α/α toroid
 Glycosyltransferases of the
 superhelical fold
 Glucoamylase
 Glucoamylase
Aspergillus awamori, variant x100



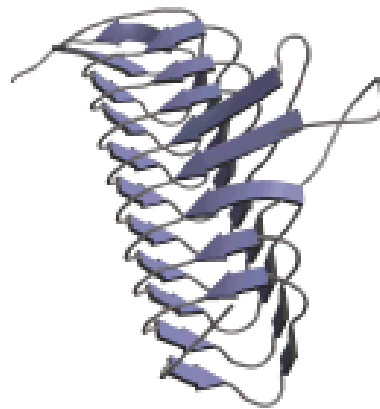
1enh
 DNA-binding 3-helical bundle
 Homeodomain-like
 Homeodomain
engrailed Homeodomain
Drosophila melanogaster

Key	
■	PDB Identifier
■	Fold
■	Superfamily
■	Family
■	Protein
■	Species

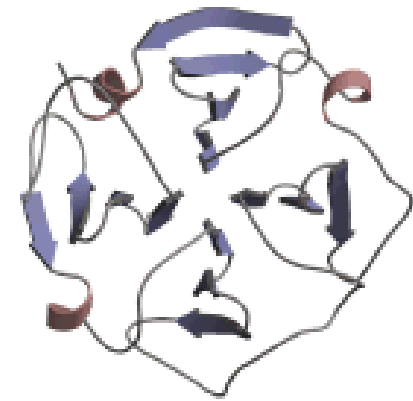
Tudo beta



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



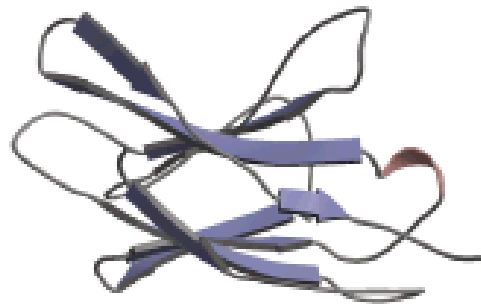
1lxe
 Single-stranded left-handed β helix
 Trimeric LpxA-like enzymes
 UDP *N*-acetylglucosamine acyltransferase
 UDP *N*-acetylglucosamine acyltransferase
Escherichia coli



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*)

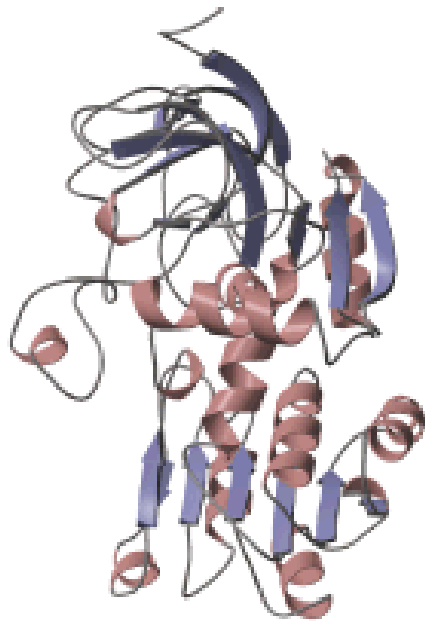


1cd8
 Immunoglobulin-like β sandwich
 Immunoglobulin
 Antibody variable domain-like
 CDS
 Human (*Homo sapiens*)

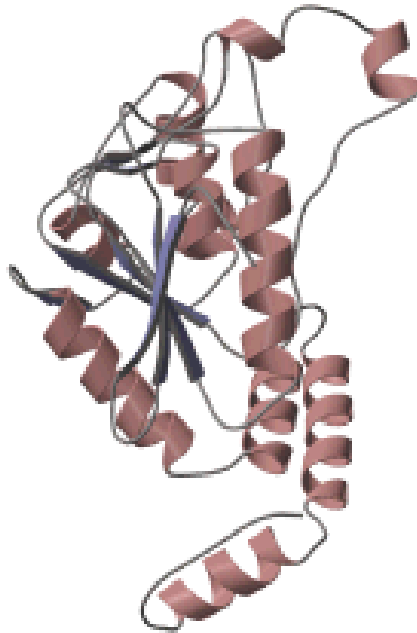
Key	
█	PDB Identifier
█	Fold
█	Superfamily
█	Family
█	Protein
█	Species

Alfa/beta

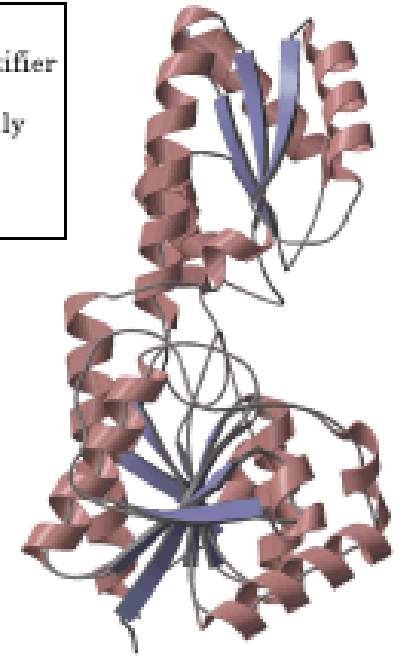
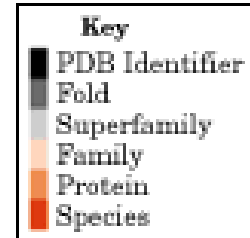
α/β



- 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*

CATH: Protein Structure Classification

<http://www.cathdb.info/>

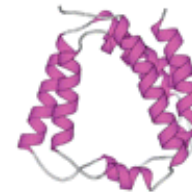


Arquitetura

Descreve as **orientações das estruturas secundárias** mas ignora a conectividade entre elas.



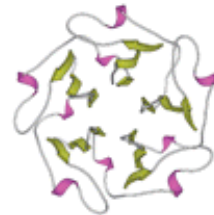
α Horseshoe (1lshA02)



α solenoid (1pprM01)



$\alpha\alpha$ barrel (1fce001)



5-bladed β propeller (1t12A00)



3-layer ($\beta\beta\beta$) sandwich (1tg7A02)



$\alpha\beta$ super-roll (1ewfA01)



3-layer ($\beta\alpha\beta$) sandwich (1j5uA01)



$\alpha\beta$ prism (1g6sA01)

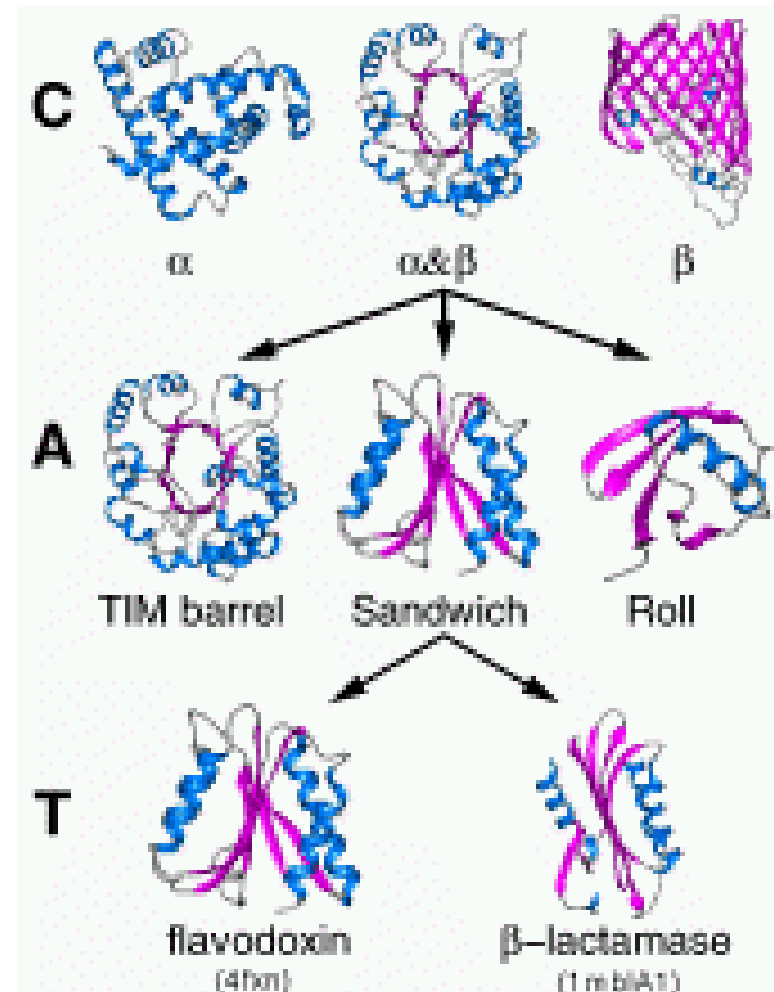
CATH: Protein Structure Classification

<http://www.cathdb.info/>



Topologia (Fold Group)

Neste nível na hierarquia, as estruturas são agrupados se eles tem a **mesma arquitetura e a mesma conectividade (topologia)** no domínio central (cerne). Domínios com a mesma topologia podem variar devido a “decorações estruturais” adicionadas ao cerne comum.



CATH: Protein Structure Classification

<http://www.cathdb.info/>



Superfamília de Homólogos (Homologous Superfamily)

Este nível agrupa proteínas com similaridade nas suas sequencias e por isso acredita que compartilham um ancestral comum. Por isso são chamados de “homólogos”.

Species Sequence Alignment of Myoglobin

```
Mouse      MGLSDGEWQLVVLNVWGKVEADLAGHGQEVLIIGLFKTHPETLDKFDKFKNLKSEEDMKGSE 60
Rat        MGLSDGEWQMVLNIWGKVEGDLAGHGQEVLIISLFKAHPETLEKFDKFKNLKSEEMKSSE 60
Human      MGLSDGEWQLVVLNVWGKVEADIPGHGQEVLIIRLFKGHHPETLEKFDKFKHLKSEDEMKA 60
Pig        MGLSDGEWQLVVLNVWGKVEADVAGHGQEVLIIRLFKGHHPETLEKFDKFKHLKSEDEMKA 60
Cow        MGLSDGEWQLVVLNAWGKVEADVAGHGQEVLIIRLFTGHPETLEKFDKFKHLKTEAEMKA 60
Zebrafish  --MAD--HDLVLKCGAVEADYAANGGEVLNRLFKKEYPDTLKLFPKFSGIS-QGDLAGSP 55
           ::*   :***: ** *. * ..:* **  *. :*:* * * *. :. : : *

Mouse      DLKKHGCTVLTALGTILKKKGQHAAEIQPLAQSHATKHKIPVKYLEFISEIIIEVLKKRH 120
Rat        DLKKHGCTVLTALGTILKKKGQHAAEIQPLAQSHATKHKIPVKYLEFISEVIIQVLKKRY 120
Human      DLKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIPVKYLEFISECIIQVLQSKH 120
Pig        DLKKHGNTVLTALGGILKKKGHHEAELTPLAQSHATKHKIPVKYLEFISEAIIQVLQSKH 120
Cow        DLKKHGNTVLTALGGILKKKGHHEAEVKHLAESHANKHKIPVKYLEFISDAIIHVLHAKH 120
Zebrafish  AVAAHGATVTKKLGELLKAKGDHAALLKPLANTHANIHKVALNRFRLITEVLVKVMAEKA 115
           :  ** *. * * : * * * * * :  *:*:* * *::: : : * : : : * :

Mouse      SGDFGADAQGAMSKALELFRNDIAAKYKELGFQG 154
Rat        SGDFGADAQGAMSKALELFRNDIAAKYKELGFQG 154
Human      PGDFGADAQGAMNKALELFRKDMASNYKELGFQG 154
Pig        PGDFGADAQGAMSKALELFRNDMAAKYKELGFQG 154
Cow        PSDFGADAQAAMSKALELFRNDMAAQYKVLGFHG 154
Zebrafish  G--LDAAGQGalRRVMDAVIDGIDGYYKEIGFAG 147
           :.* *. * : : : . * : . * * : * * *
```

CATH: Protein Structure Classification

<http://www.cathdb.info/>



CATH v3.4 contém 104,238 sequencias de proteínas diferentes.

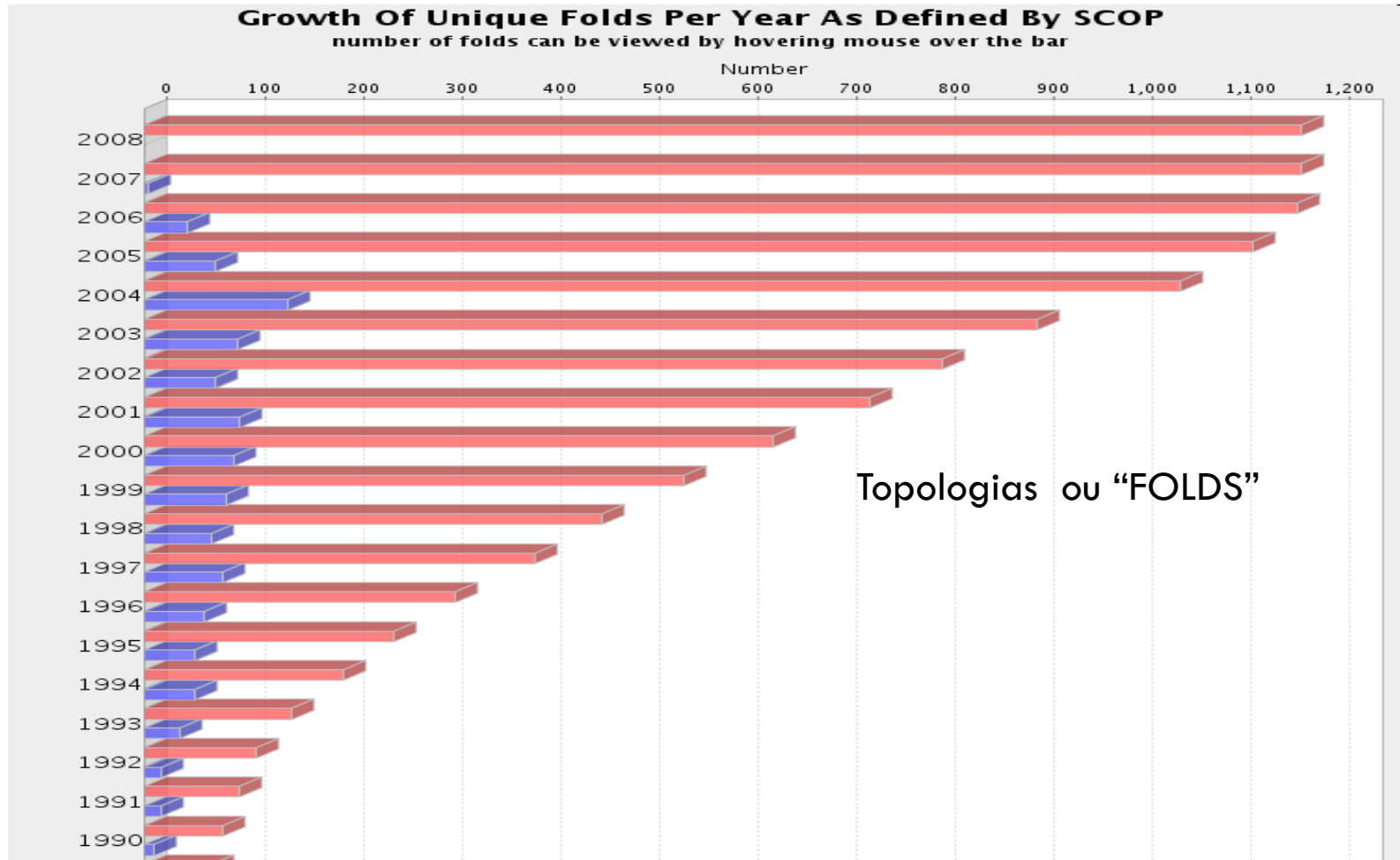
Domínios: 152,920

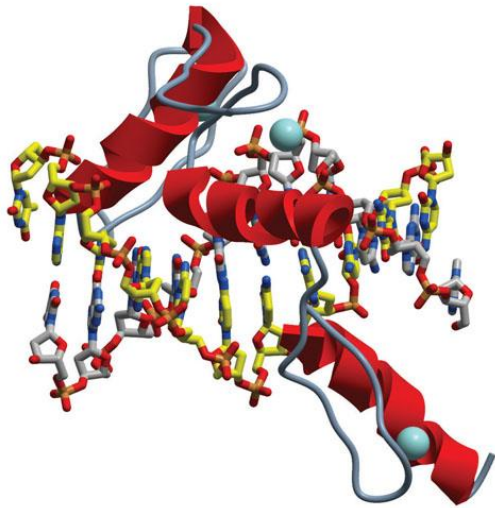
Famílias de Sequencias: 11,330

Superfamílias: 2,549

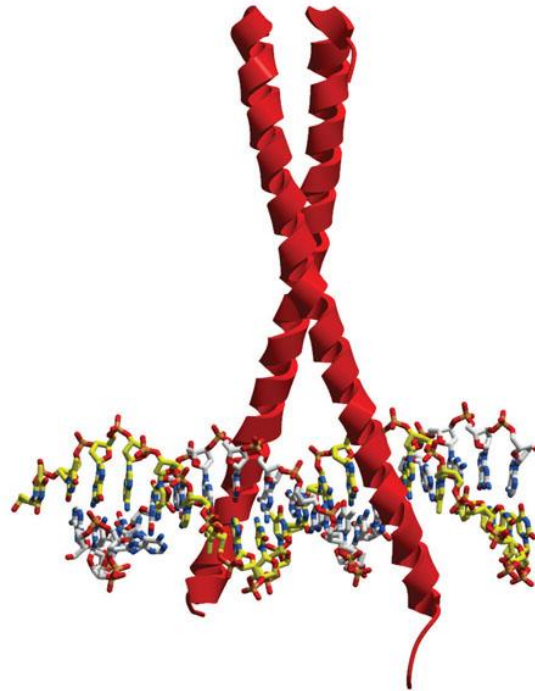
Topologias (Folds): 1,282

Número limitado de Topologias

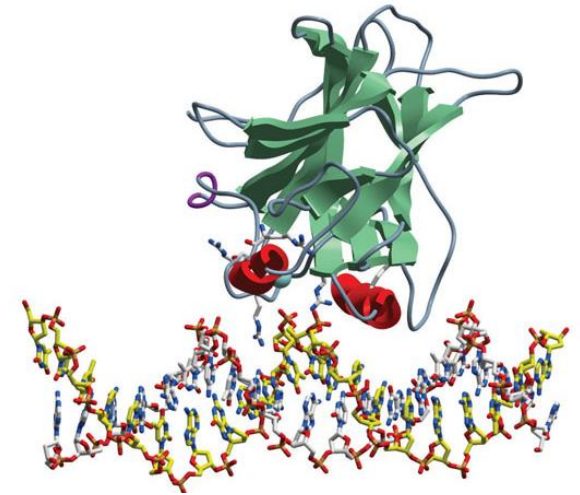




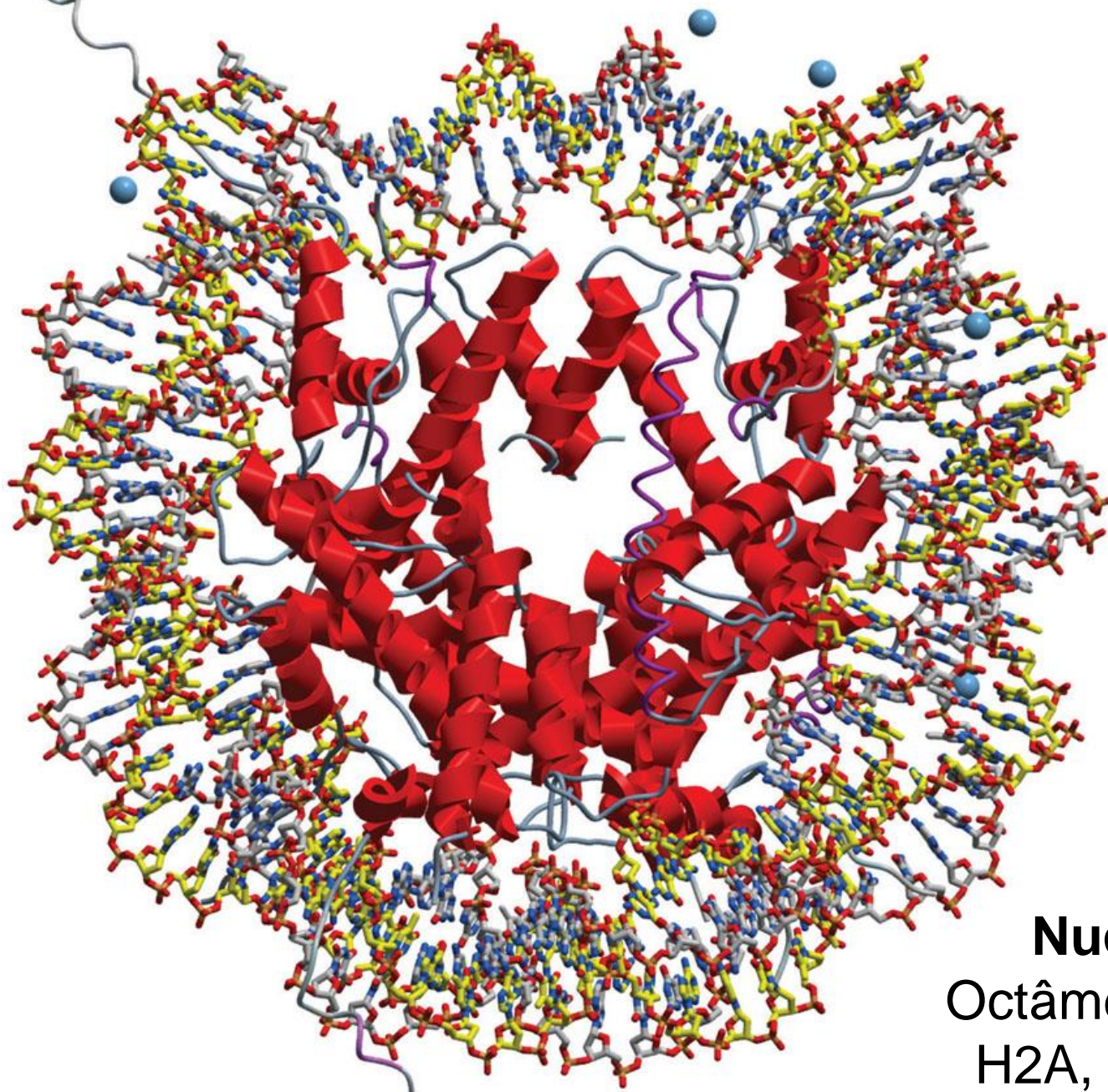
Dedo de zinco



Zipper
de
leucina

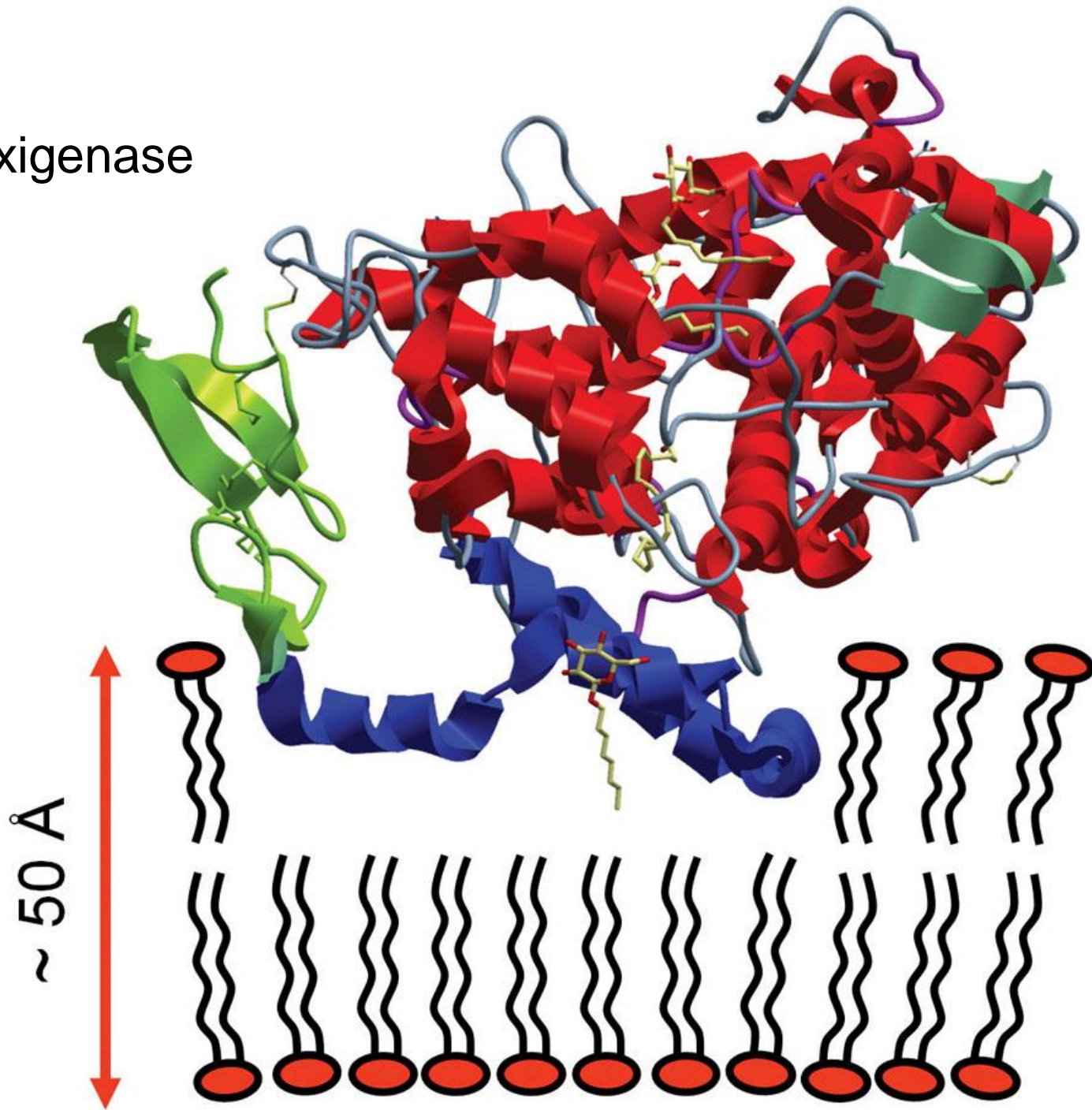


Supressor
de tumor
p53

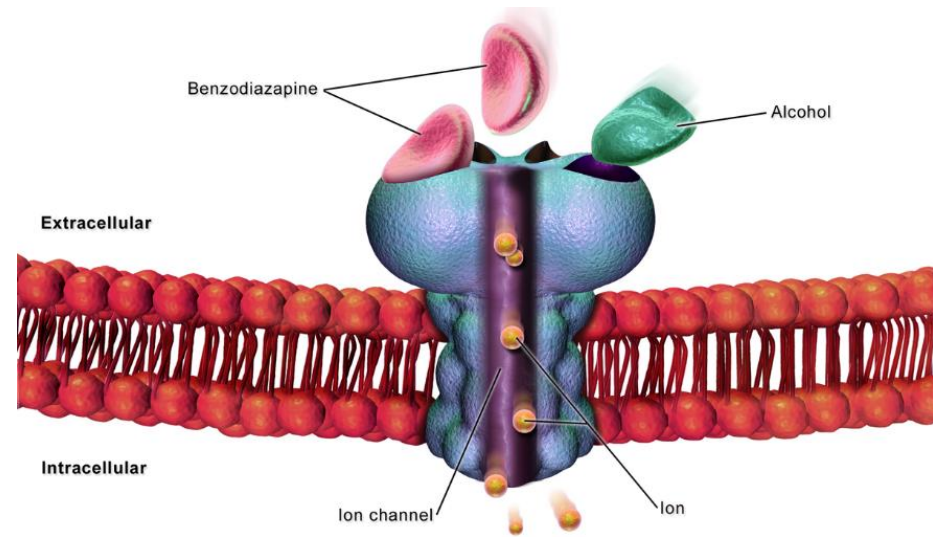
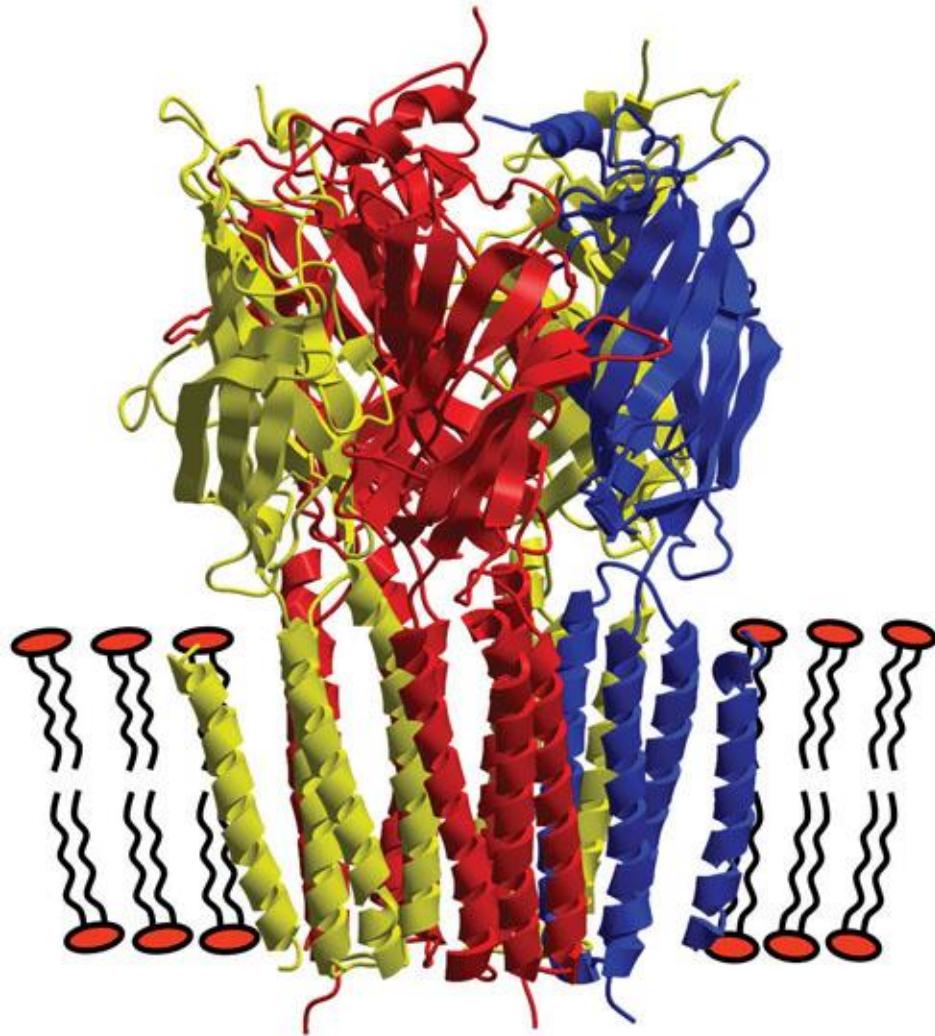


Nucleossomo
Octâmero de histonas
H2A, H2B, H3 e H4

COX2
Cyclooxygenase



Neuroreceptor de GABA_A (ácido gama-aminobutírico)



Primary Structure = sequence of amino acids

3-letter code

Lys-Thr-Tyr-Phe-Pro-His-Phe-Asp-Leu-Ser-His-**Gly** ...

1-letter code

KTYFPHF**D**LSH**G**

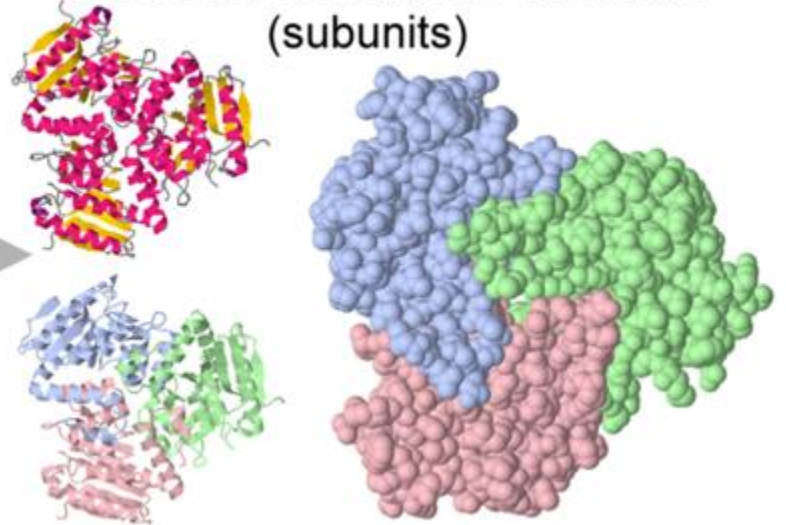
Secondary Structure = **alpha helices**, **beta strands**



Tertiary Structure = fold helices and strands into domains



Quaternary Structure (Biological Units) = functional assemblies of chains (subunits)



Estrutura quaternária descreve a maneira pela qual diferentes cadeias polipeptídicas se associam para formar complexos multiméricos ... **Uma outra maneira de gerar proteínas com múltiplos domínios.**

A estrutura quaternária frequentemente representa a forma ativa da proteína.

Ela permite comunicação entre uma subunidade e outra (os fenômenos de alosteria e cooperatividade)

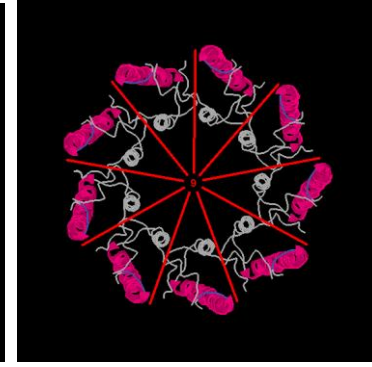
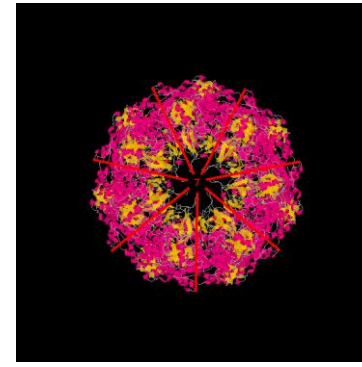
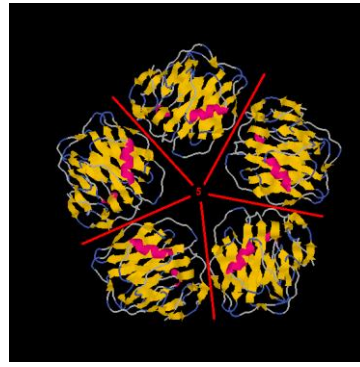
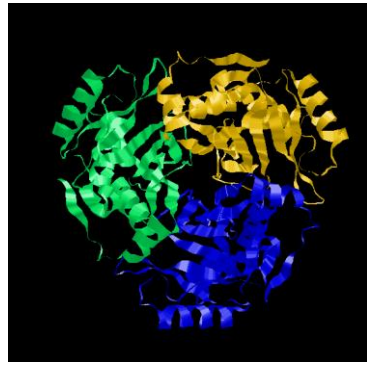
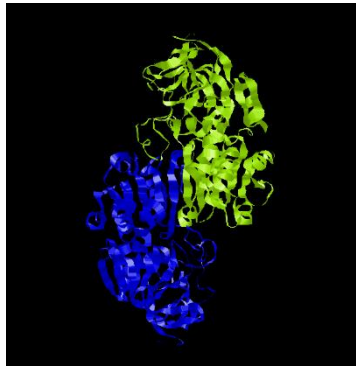
Hetero-multímeros (ie hemoglobina $\alpha_2\beta_2$)

Homo-multímeros (ie GroEL α_{14})

Complexos multiméricos (especialmente homo-multiméricos) frequentemente exibem simetria estrutural

Simetria cíclica

N identical unidades idênticas com posições relacionadas por um eixo de rotação de ordem n (C_n)



Simetria C_2 :

Horse Liver Alcohol
Dehydrogenase

Simetria C_3 :

Chloramphenicol
Acetyltransferase

Simetria C_5 :

Serum amyloid
P component

Simetria C_7 :

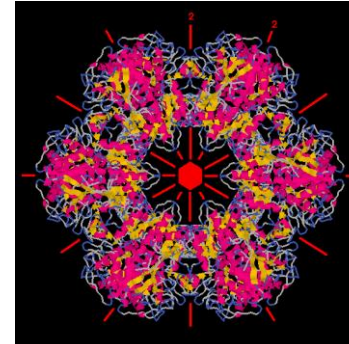
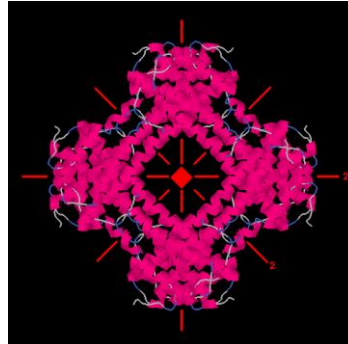
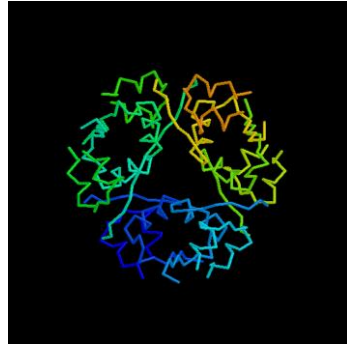
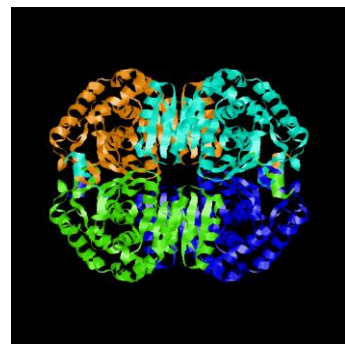
GroEL-GroES-
ADP complex

Simetria C_9 :

Light Harvesting
Complex

Simetria dihedral (D_N)

$2 \times N$ unidades idênticas relacionadas por um eixo C_N e N eixos C_2 .

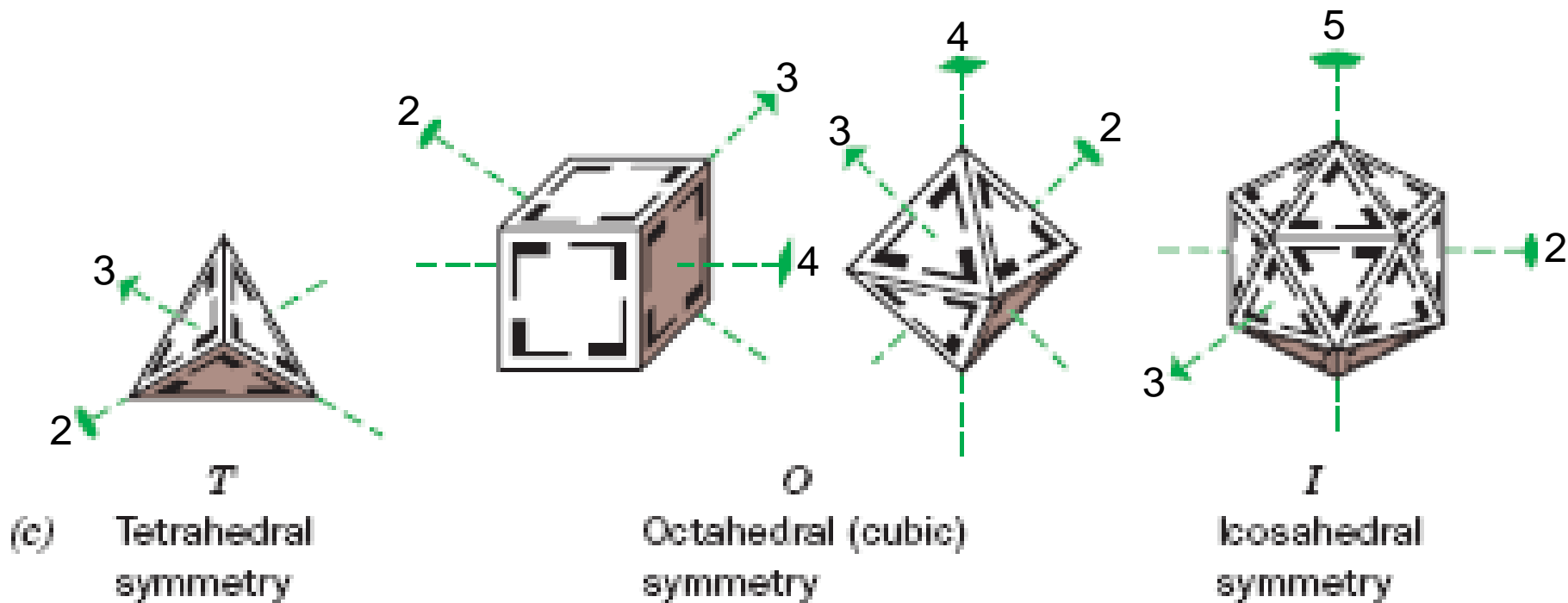


D_2 : Dogfish Lactate
Dehydrogenase

D_3 : Insulin

D_4 : Hemerythrin

D_6 : Glutamine Synthetase



Outros objetos simétricos

- Tetrahédron (*T*),
- Octahédron (*O*),
- Icosohédron (*I*),

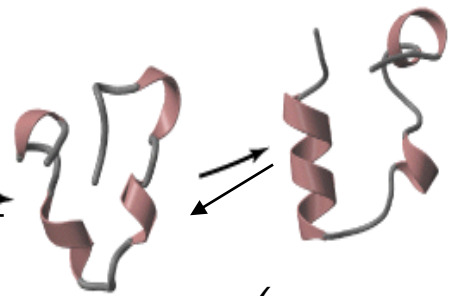
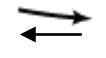
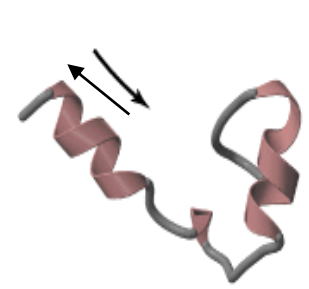
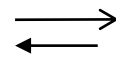
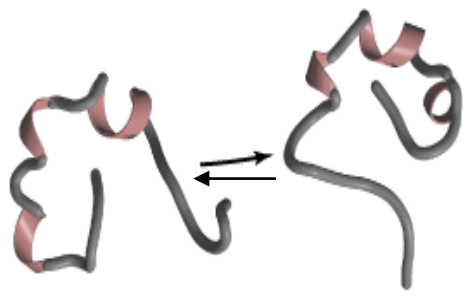
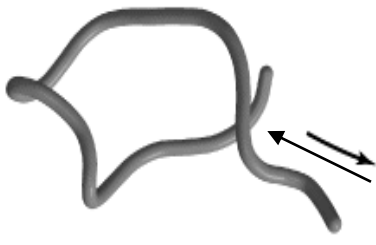
Têm 12, 24, e 60 posições equivalentes, respectivamente

Enovelamento e Desenovelamento de Proteínas

*Cadeia com
mais entropia*

Enovelamento ou Renaturação →

*Cadeia com
menos entropia*

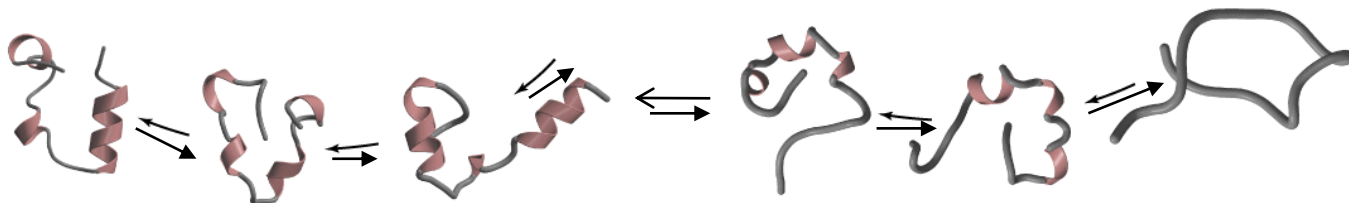
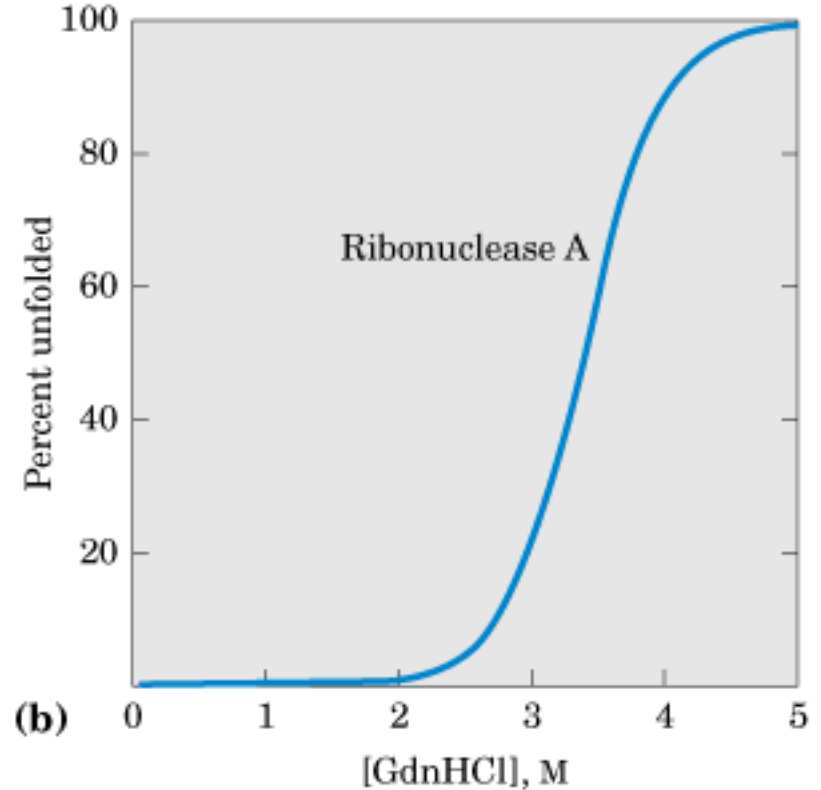
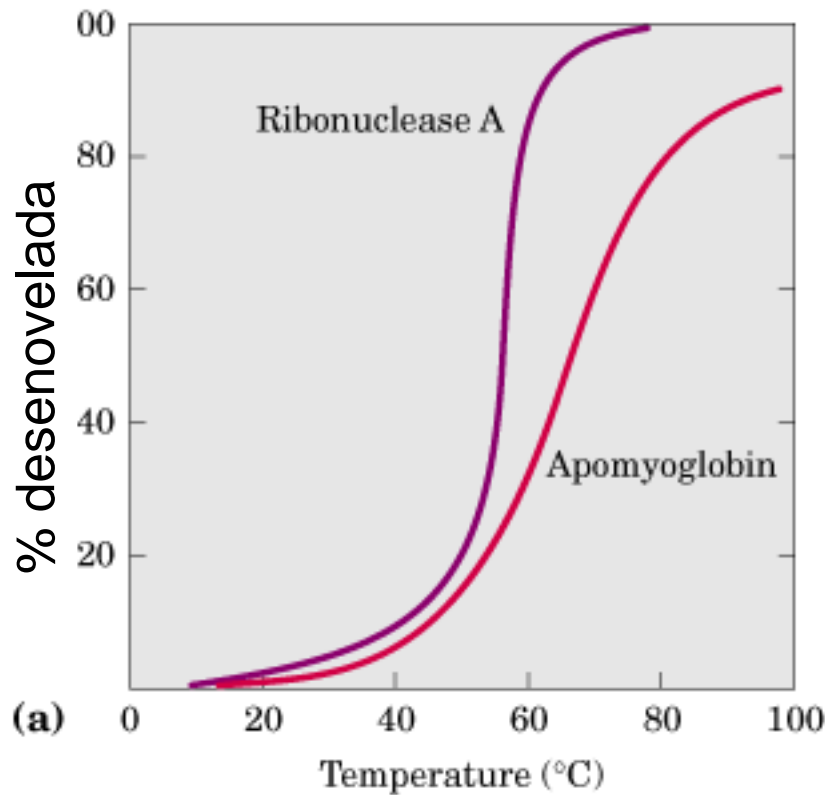


*(muitas
conformações
possíveis)*

*(poucas
conformações)*

← Desnaturação ou Desenovelamento

Desnaturação ou Desenovelamento → ← Renaturação ou Reenovelamento

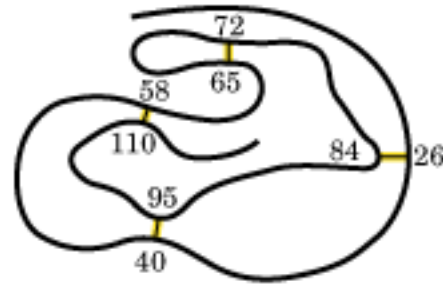


Desnaturação ou Desenovelamento →

Enovelamento “Espontâneo” de Proteínas

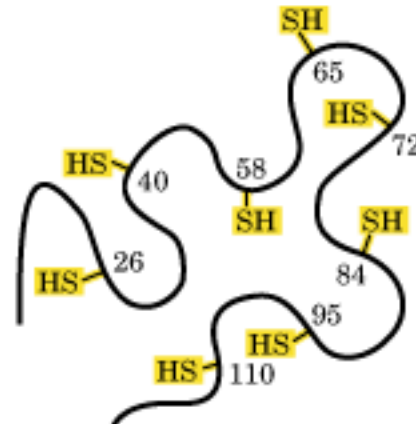
Expto de
Christian Anfinsen
com RNaseA
e sua
“Hipótese
Termodinâmica”

Prêmio Nobel
em 1972



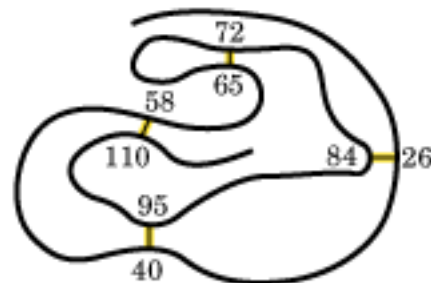
Estado nativo
Cataliticamente ativo

+uréia
+2-mercaptoetanol

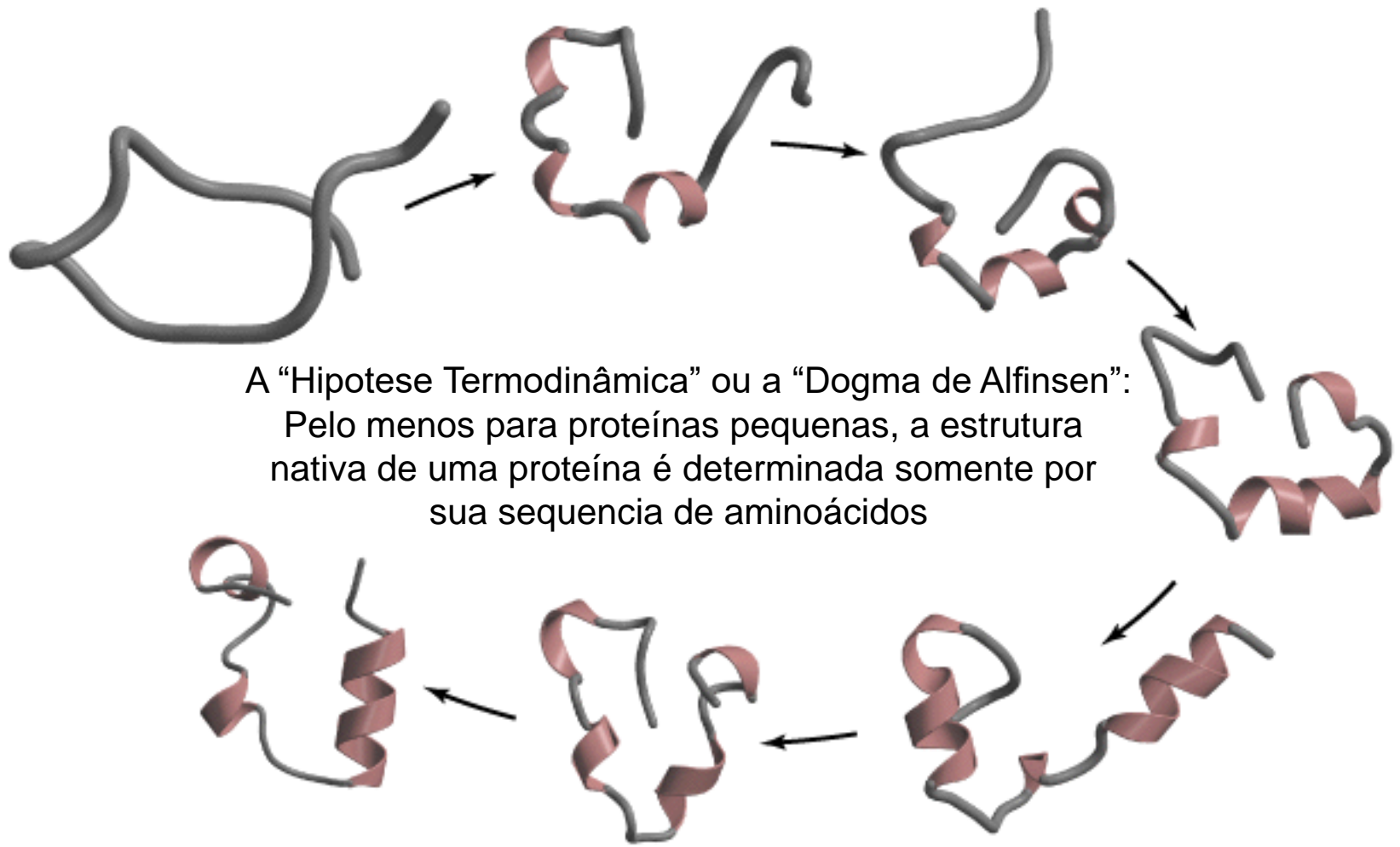


Estado desenovelado
Cataliticamente inativo
Pontes dissulfetos reduzidos

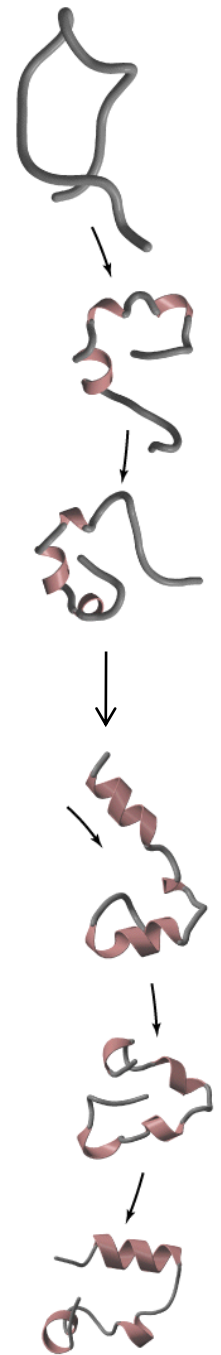
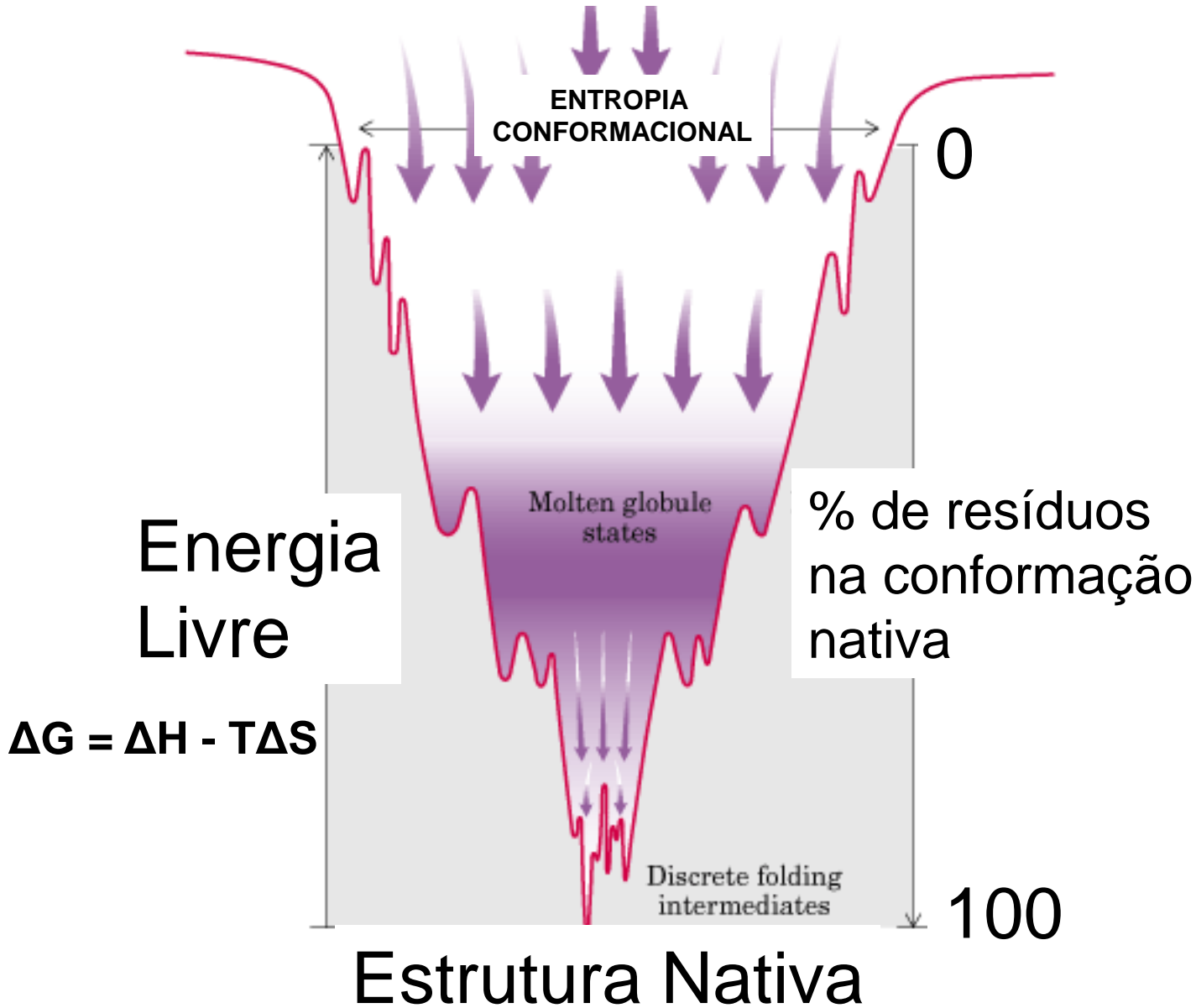
remover uréia e 2-mercaptoetanol
Expor a oxigênio



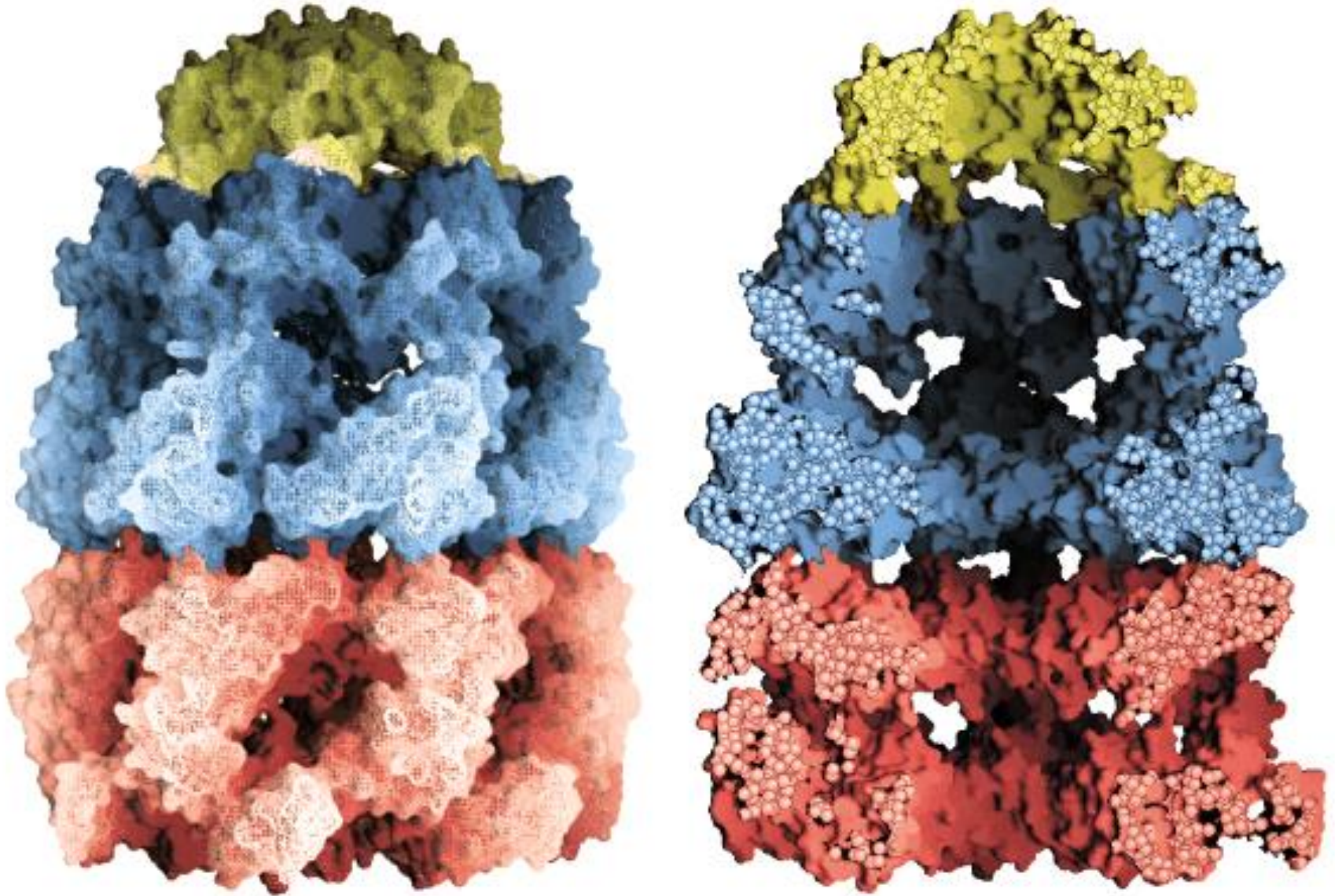
Estado nativo
Cataliticamente ativo
Cisteínas oxidadas
Pontes dissulfetos reproduzidos

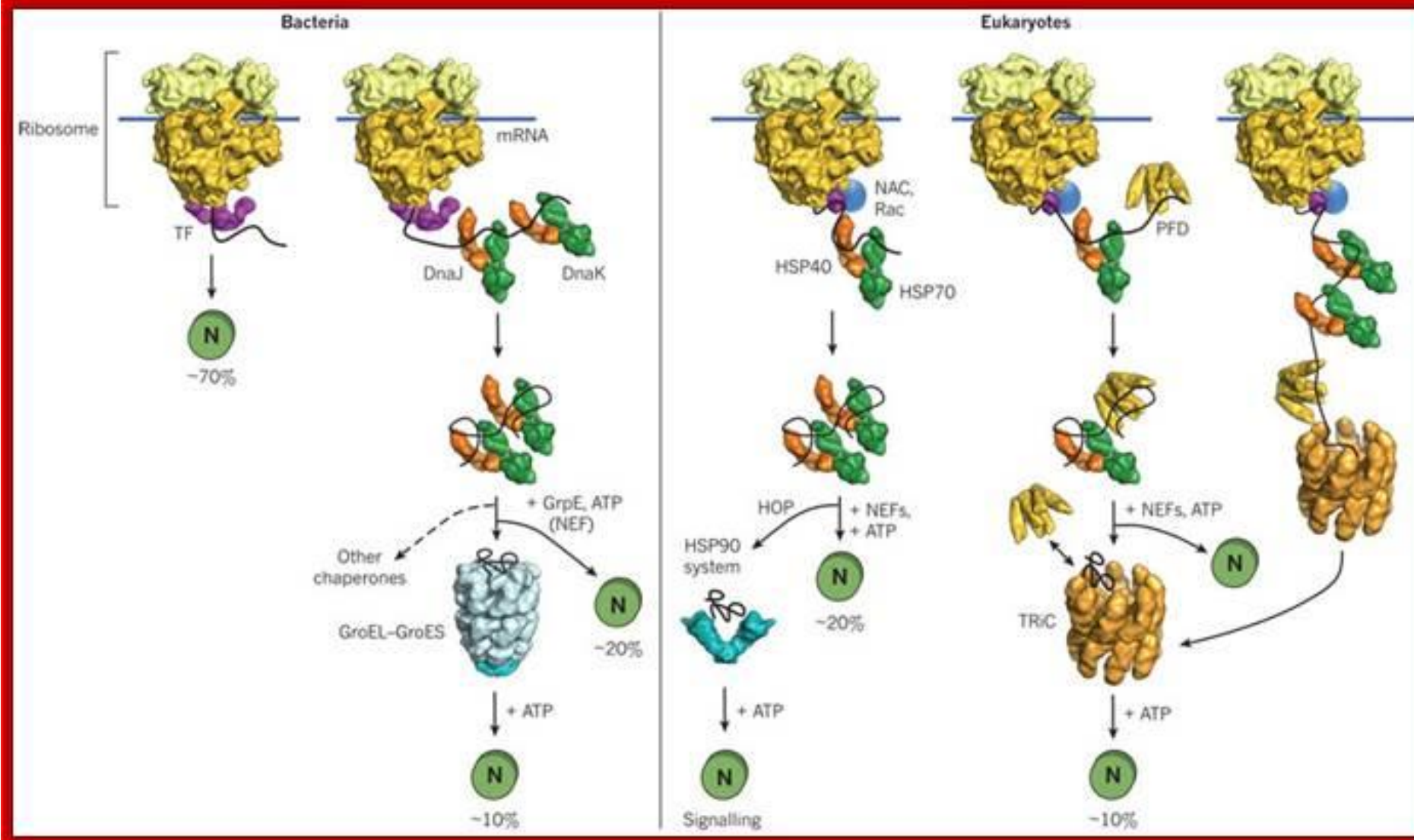


Início de formação de hélices e fitas “enovelamento” da cadeia



Chaperona molecular bacteriana (GroEL/GroES)





DnaK and DnaJ are involved in guiding the protein chain to fold; if it fails proteins are guided to chaperonins for folding
 Nature **475**, 324-332 (2011) doi:10.1038/nature10317;