

Comparaison

et

**Prédiction
de
structures3D**

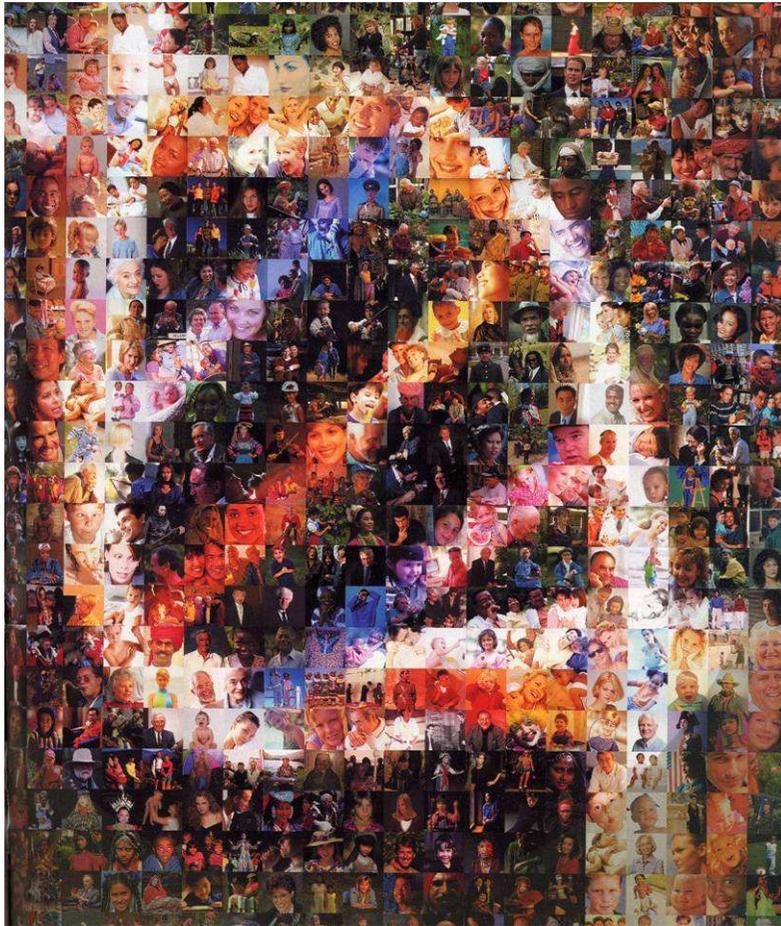


Prédiction de la structure 3D des protéines

- * La structure 3D de la protéine est centrale à la génomique, car plus encore que la séquence, c'est elle qui détermine sa fonction. La structure va déterminer les sites actifs ou les épitopes d'antigénicité.
- * Prédire la structure 3D d'une protéine, c'est donc approcher la fonction potentielle qu'elle assure dans la cellule.
- * Outre leur importance fonctionnelle, les structures 2D et 3D présentent un intérêt cognitif majeur car elles sont mieux conservées que les séquences au cours de l'évolution.

Le séquençage du génome humain

Point de départ du vaste programme de génomique structurale



~ 3 milliards de paires de bases

~ 32 000 gènes différents

Mais n x 100 000 protéines

Nature, vol 409, 15 février 2001

Génome :

Ensemble du matériel génétique (patrimoine héréditaire) d'un individu ou d'une espèce. Il est constitué de molécules d'acides nucléiques (ADN ou ARN). Les gènes, c'est-à-dire les parties d'ADN porteuses d'une information génétique, ne constituent qu'une partie du génome.

Chez les virus, le génome viral est contenu dans une molécule d'ADN simple ou double brin ou une molécule d'ARN.

Chez les procaryotes, on distingue :

le génome chromosomique (une molécule d'ADN circulaire).

le génome extra chromosomique (plasmides, épisomes).

Chez les eucaryotes, on distingue :

le génome nucléaire (du noyau)

le génome mitochondrial (des mitochondries), transmis par la mère.

le génome chloroplastique (des chloroplastes), chez les plantes.

Transcriptome :

Ensemble des ARN messagers transcrits à partir du génome.

Protéome :

Terme inventé en 1995. Le protéome est le "complément protéique total du génome", c'est à dire l'ensemble des protéines exprimé par le génome d'une espèce donnée. Il assure le développement, la croissance et le fonctionnement de la cellule (donc de l'organisme).

Métabolome :

Ensemble des métabolites produits par les cellules dans des conditions spécifiques.

!! RAPPEL !!

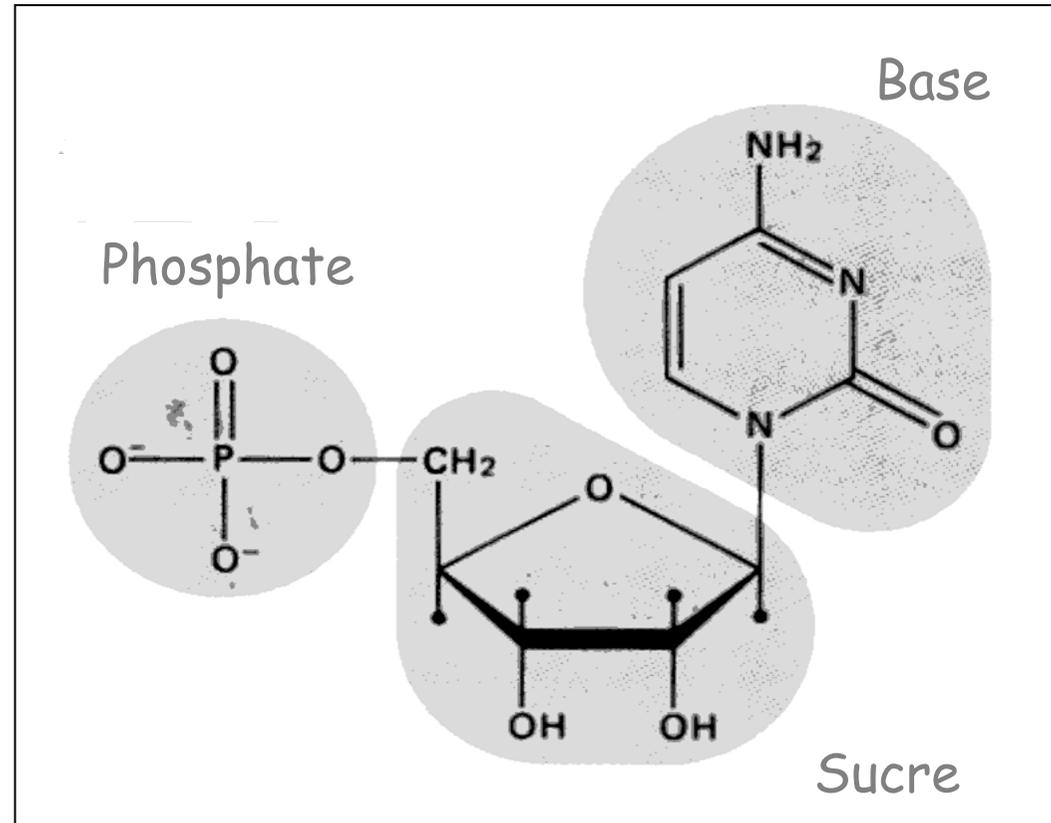
Nucléotides :

Base azotée : A, C, G, T, U

Sucre à 5 atomes de C
= PENTOSE

désoxyribose ou ribose

Groupe(s) phosphate(s)



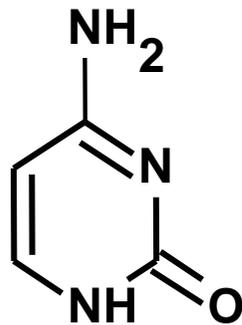
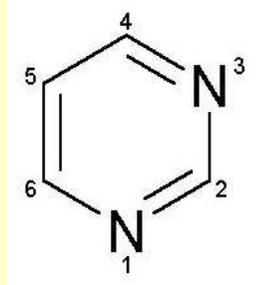
ADN : Acides DésoxyriboNucléiques

ARN : Acides RiboNucléiques

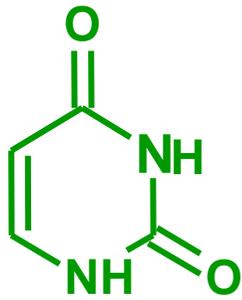
!! RAPPEL !!

Les bases azotées :

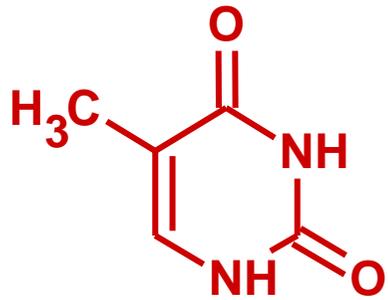
Pyrimidines



Cytosine (C)

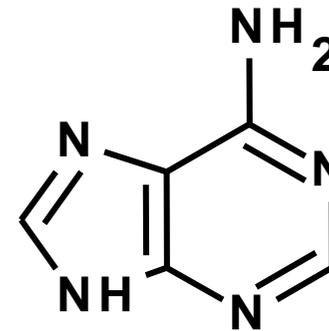
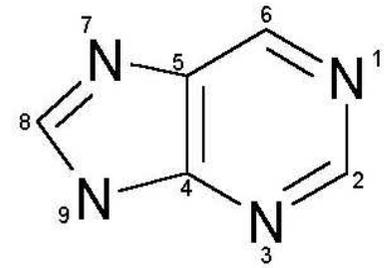


Uracile (U)

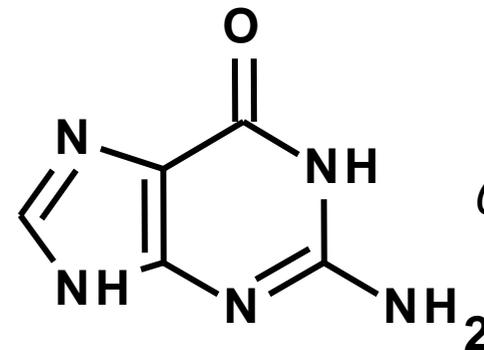


Thymine (T)

Purines



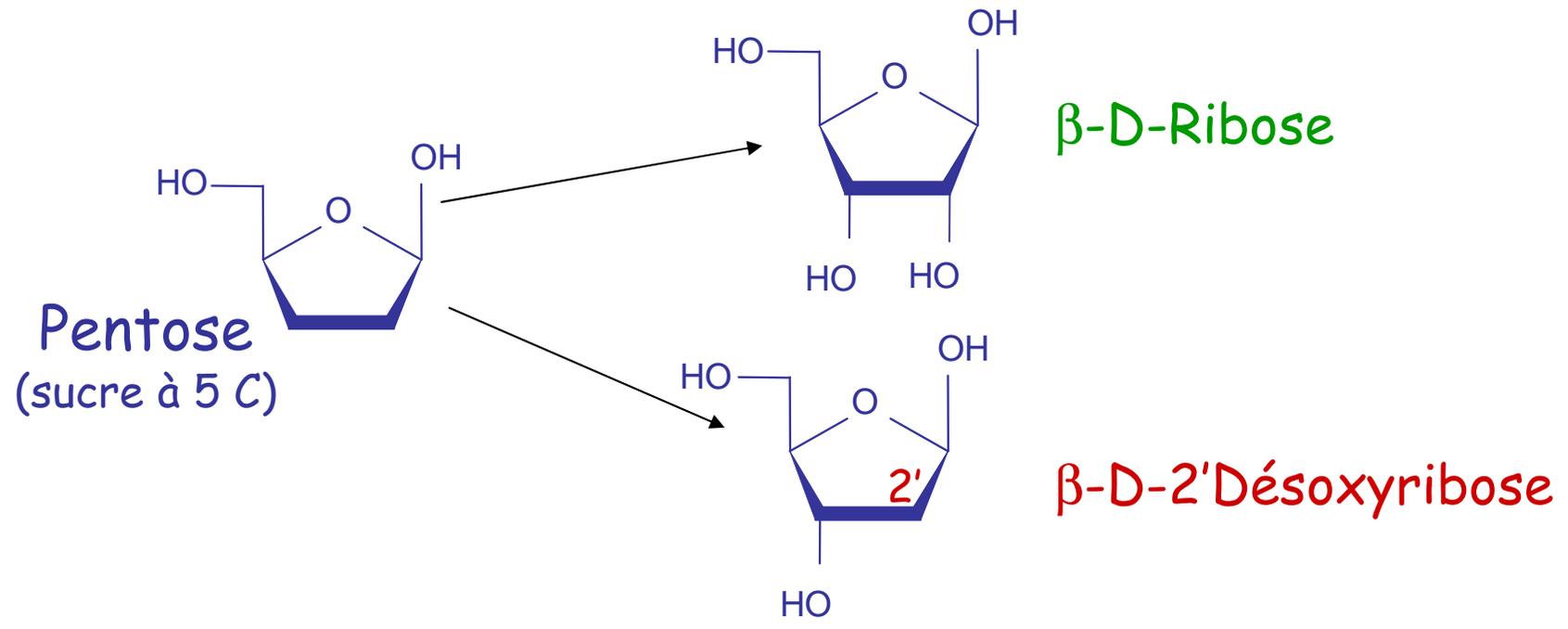
Adénine (A)



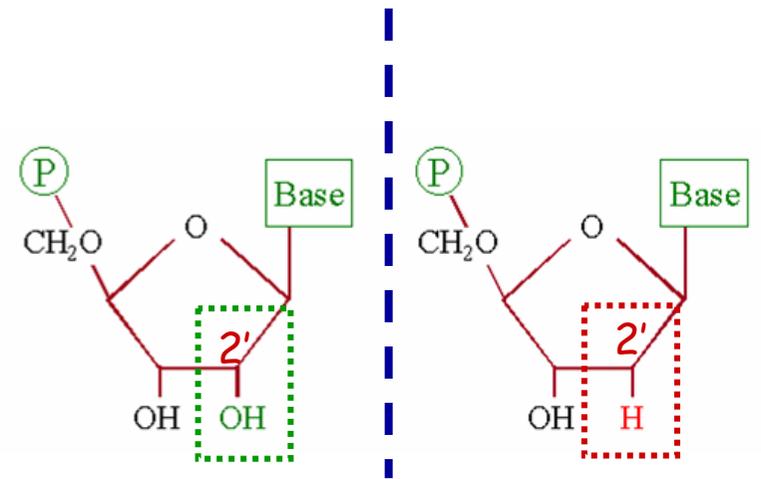
Guanine (G)

!! RAPPEL !!

Les sucres :



Ribose
Ribonucléotide
ARN



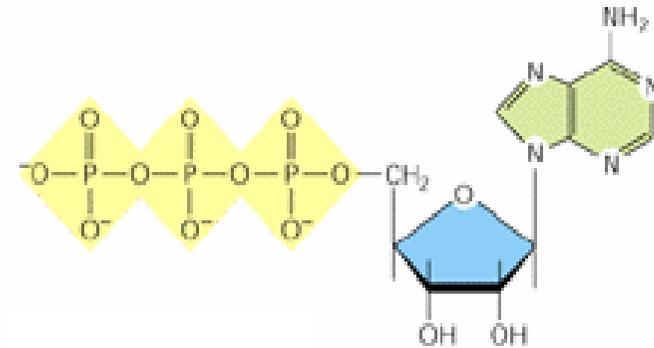
Désoxyribose
Désoxyribonucléotide
ADN

!! RAPPEL !!

Nomenclature :

Base + Sucre = Nucléoside

Base + Sucre + Phosphate = Nucléotide



Bases	Nucléosides	Abbr.	Nucléotides	Abbr.
Adénine	Adénosine	A	Adénosine Mono, Di ou Tri-Phosphate	AMP, ADP, ATP
	désoxyAdénosine	dA	désoxyAdénosine Mono, Di ou Tri-Phosphate	dAMP, dADP, dATP
Guanine	Guanosine	G	Guanosine Mono, Di ou Tri-Phosphate	GMP, GDP, GTP
	désoxyGuanosine	dG	désoxyGuanosine Mono, Di ou Tri-Phosphate	dGMP, dGDP, dGTP
Cytosine	Cytidine	C	Cytidine Mono, Di ou Tri-Phosphate	CMP, CDP, CTP
	désoxyCytidine	dC	désoxyCytidine Mono, Di ou Tri-Phosphate	dCMP, dCDP, dCTP
Thymine	désoxyThymidine	dT	désoxyThymidine Mono, Di ou Tri-Phosphate	dTMP, dTDP, dTTP
Uracile	Uridine	U	Uridine Mono, Di ou Tri-Phosphate	UMP, UDP, UTP

!! RAPPEL !!

La synthèse protéique

ARNm $\xrightarrow{\text{Traduction}}$ Protéine

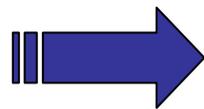
Code génétique

4 bases $\xrightarrow{\hspace{1.5cm}}$ 20 aa

CODON = 3 bases $\xrightarrow{\hspace{1.5cm}}$ 1 aa

$4^3 = 64$ codons $\xrightarrow{\hspace{1.5cm}}$ 20 aa

plusieurs codons $\xrightarrow{\hspace{1.5cm}}$ 1 même aa



Dégénérescence du
code génétique

!! RAPPEL !!

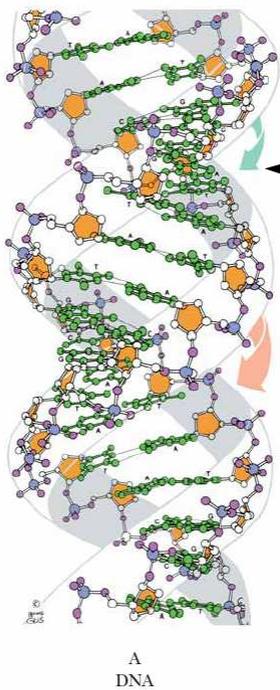
Le code génétique

Code dégénéré :

64 codons (4x4x4)
pour 20 acides aminés

Un acide aminé avec
plusieurs codons

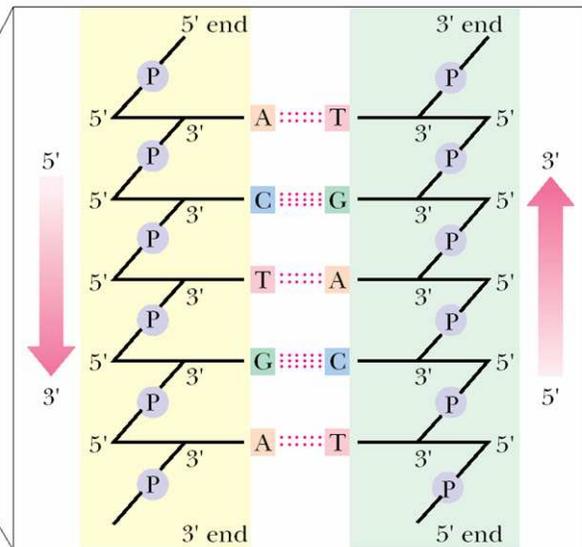
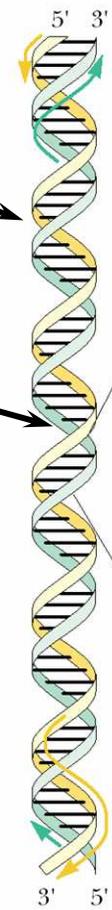
	T	C	A	G	
T	TTT Phe TTC Phe TTA Leu TTG Leu	TCT Ser TCC Ser TCA Ser TCG Ser	TAT Tyr TAC Tyr TAA Stop TAG Stop	TGT Cys TGC Cys TGA Stop <u>TGG Trp</u>	T C A G
C	CTT Leu CTC Leu CTA Leu CTG Leu	CCT Pro CCC Pro CCA Pro CCG Pro	CAT His CAC His CAA Gln CAG Gln	CGT Arg CGC Arg CGA Arg CGG Arg	T C A G
A	ATT Ile ATC Ile ATA Ile <u>ATG Met/start</u>	ACT Thr ACC Thr ACA Thr ACG Thr	AAT Asn AAC Asn AAA Lys AAG Lys	AGT Ser AGC Ser AGA Arg AGG Arg	T C A G
G	GTT Val GTC Val GTA Val GTG Val	GCT Ala GCC Ala GCA Ala GCG Ala	GAT Asp GAC Asp GAA Glu GAG Glu	GGT Gly GGC Gly GGA Gly GGG Gly	T C A G



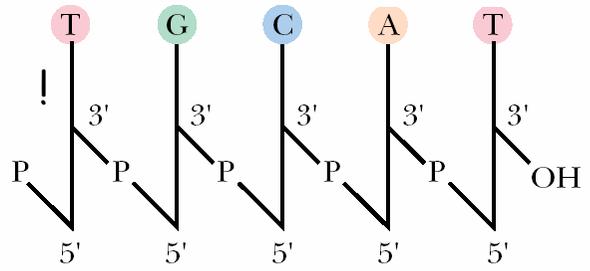
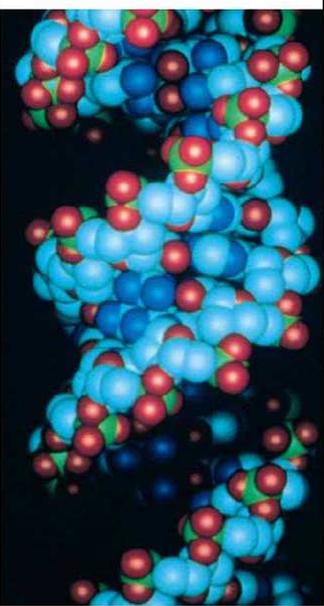
!! RAPPEL !!

Sillon mineur

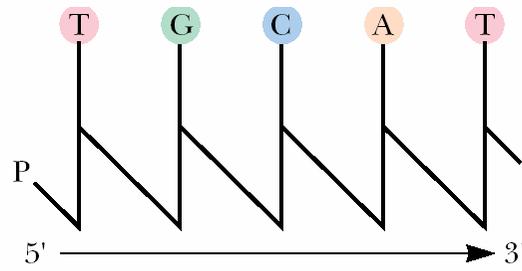
Sillon majeur



Segment of unwound double helix illustrating the antiparallel orientation of the complementary strands



or simply



Sens de lecture 5' vers 3' !!!!

Génomique :

Science qui étudie les génomes. La génomique est l'étude et l'analyse exhaustive et multidisciplinaire des génomes. Elle vise à dresser l'inventaire de l'ensemble des gènes d'un organisme à les localiser sur les chromosomes et à caractériser leur séquence ainsi qu'à étudier leur fonction.

La génomique se développe en deux volets complémentaires. Le premier volet correspond à l'**analyse structurale** (structure physique et organisation) du génome et du protéome : elle s'appuie sur l'analyse des séquences (issue du séquençage) et sur les données de la génomique structurale. Le second volet correspond à la **génomique fonctionnelle** (ou post-génomique), qui étudie la fonction des gènes, leur mode de régulation et leurs interactions.

La génomique, dans son acceptation la plus large, englobe l'analyse de la structure physique du génome et du protéome (qui s'appuie sur l'analyse des séquences et les données de la génomique structurale) et la génomique fonctionnelle (ou post-génomique).

Génomique fonctionnelle (post-génomique) :

Etude et analyse directe du transcriptome et du protéome : elle vise à déterminer la fonction des gènes à partir de leurs produits d'expression (ARN et protéines), ainsi qu'à étudier leur mode de régulation et leurs interactions.

Génomique structurale (protéomique structurale) :

Désigne l'étude du protéome (ensemble des protéines identifiées à partir des génomes) du point de vue de la structure tridimensionnelle : étude des familles structurales protéiques, des repliements structuraux et de la relation structure/fonction.

Protéomique :

Regroupe les recherches (détection, séparation, identification, étude) permettant d'accéder directement aux protéines (protéome), de déterminer leurs activités, leurs fonctions et d'analyser leurs interactions et leurs modifications au cours du temps.

!! RAPPEL !!

Du gène à la protéine

* La première séquence

1953, Frederick Sanger, Séquence de l'insuline

2003, 900.000 séquences protéiques entrées, (5.217.756

2003, 20.000.000 séquences nucléiques, (26.000.000 (10,

* ADN, ARN, Protéine

Transcription (ADN \longrightarrow ARN)

Traduction (ARN \longrightarrow Protéine)

Génie génétique, modifications des gènes et production de protéines

* Importance de la séquence

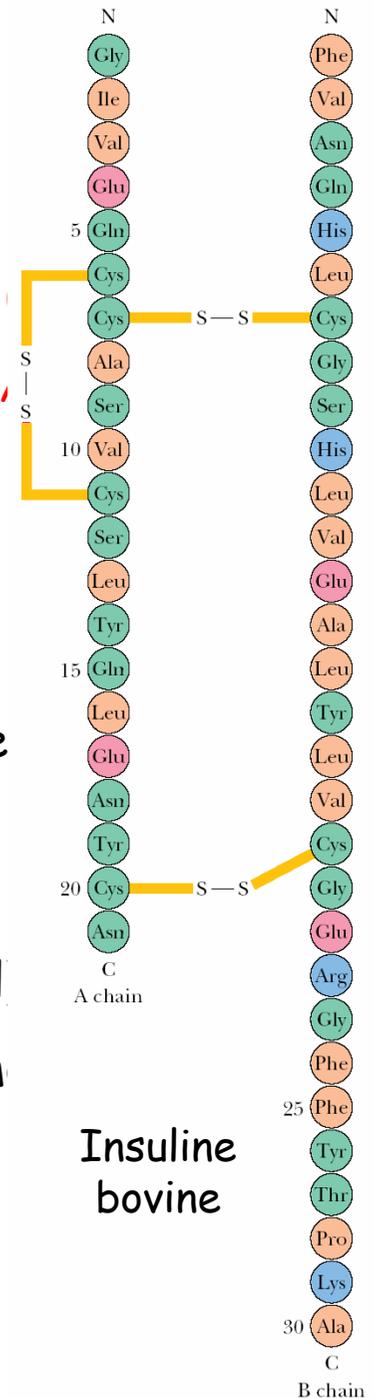
Orientation de la séquence (NH₂ terminal, COOH terminal)

Compréhension de mécanismes d'action, de la pathologie m

Reflet de l'évolution et de l'histoire des protéines.

* Bio-Informatique et Protéine

Prédiction de structures 2D et 3D



!! RAPPEL !!

Rôles des protéines :

* Catalyse enzymatique :

- Classe de protéine la plus importante (plus de 3000)
- Enzymes, molécules biologiques agissant comme des catalyseurs.
- Transformation intra ou extracellulaire de composés...
- Nom en "ase"

Exemples : # DNAPolymérase, impliquée dans la réplication et la réparation de l'ADN

Protéases, impliquée dans la dégradation des protéines.

Amylases, impliquées dans la transformation de l'amidon en sucres.

alcool déshydrogénase, transformation de l'éthanol sanguin en acétaldéhyde...

* Protéines régulatrices :

- Ne participe pas directement dans la réaction, mais régulent l'activité d'autres protéines

Exemples : # Insuline, hormone de régulation du métabolisme du glucose

Répresseur Lac, gestion de la dégradation procaryote du lactose...

!! RAPPEL !!

Rôles des protéines :

* Protéines de transport :

- Transport de métabolites spécifiques

Exemples : # Hémoglobine, transport du O₂ des poumons aux tissus

Sérum Albumine, transport des *acides gras*

Canaux transmembranaires pour le transport spécifique de métabolites,...

* Protéines de réserves :

- Source d'élément pour le développement d'un organisme

Exemples : # Besoin en AZOTE, ovalbumine, Caséine (mais aussi en C, H, O, S)

Besoin en Fer, la Ferritine...

* Protéines contractiles et de mobilité :

- Capacité de mouvement (*mobilité et division cellulaire, contraction musculaire*)
- Edifices filamenteux multiprotéiques (*actine, myosine*),(*tubuline*)...

!! RAPPEL !!

Rôles des protéines :

* Protéines de structure :

- Maintien des structures biologiques intracellulaires (*cytosquelette d'actine*) ou extracellulaires (*matrice extracellulaire, collagène,...*)
- Structure extracellulaire (*cheveux, ongles, cornes grâce à l' α -kératine, toile d'araignée avec la β -kératine...*)

* Protéines de protection et *d'agression* :

- Rôles biologiques actifs dans la défense de l'organisme contre une attaque extérieure (*Immunoglobulines ou anticorps*)
- Réparation de l'organismes (*coagulation sanguine, fibrinogène et thrombine*), adaptation au froid (*protéine antigel*)...
- *Attaque ou défense (Protéines lytiques ou neurotoxique de venin de serpent ou d'abeille...)*

Génomique structurale

Dans la logique du séquençage des génomes qui identifie les séquences d'ADN, la génomique structurale s'attèle à déterminer la structure 3D des protéines codées par le génome.

Moyens

- Programme de grande échelle:

Structural Genomics Initiatives
PSI : Protein Structure Initiative

- Etudes pilotes aux Etats-Unis, en Europe et en Asie

But

Déterminer, par cristallographie et RMN, la structure 3D d'un ensemble de 10000 à 20000 protéines sélectionnées (représentatives de l'ensemble des familles protéiques structurales) afin d'avoir une meilleure connaissance et compréhension des repliements structuraux types et de la relation structure-fonction.

Stockage

Les données atomiques produites viendront enrichir la PDB, banque de données de structures des protéines.

Critères des protéines "cibles" à déterminer

un intérêt structural

- Elle doit être représentative d'une structure type/
une famille protéique \equiv un groupe de séquences protéiques présentant une similarité minimale de 25-35%

Note: La similarité de séquence induit une similarité de structure

un intérêt fonctionnel

- Elle doit pouvoir ouvrir la voie à une fonction ou une catégorie fonctionnelle nouvelle (et mettre ainsi en évidence, par comparaison de structure, des protéines de fonction similaire).

un intérêt biologique

- Elle doit pouvoir constituer le point de départ d'un programme expérimental (mutagénèse dirigée, étude de liaisons avec des ligands, essais enzymatiques, études d'interaction protéine-protéine ...).

un intérêt biomédical et économique

La démarche des centres de génomique structurale

Amplification, expression et purification de la protéine

- Amplification par PCR des ADNc d'intérêt
- Clonage de chacune de ces séquences dans un vecteur d'expression approprié
- Séquençage du gène cloné pour vérifier l'amplification de la séquence codante
- Obtention de la protéine issue de l'expression de l'ADNc à un taux suffisamment élevé
- Confirmation de l'identité de la protéine exprimée
- Purification de la protéine,

Préparation de la protéine

- Définition des conditions convenables de cristallisation ou de solution destinées à la spectroscopie RMN

Détermination expérimentale de la structure 3D

- par cristallographie aux rayons X
- par RMN

Exploitation des données structurales

- Calcul de modèles structuraux protéiques comparatifs à partir de la nouvelle structure obtenue.
- Etablissement d'inférences fonctionnelles (annotation) à partir des modèles dérivés de la nouvelle structure.

Approche de la bioinformatique

La bioinformatique peut approcher la structure des protéines de plusieurs manières différentes :

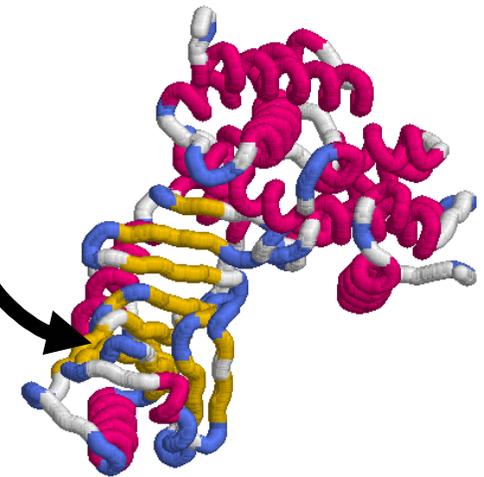
- ★ en comparant les séquences protéiques assurant des fonctions voisines dans différents organismes
- ★ en classant les grands types de repliements
- ★ en identifiant des sous-structures et des domaines
- ★ en prédisant et en modélisant une structure 3D compatible avec différents paramètres constatés ou imposés.

Ce qui lui permet ainsi de rationaliser le choix des protéines à étudier (et des structures "types" à déterminer par les méthodes expérimentales), lesquelles enrichiront les bases de données (séquences, structures et propriétés biologiques).



Une question

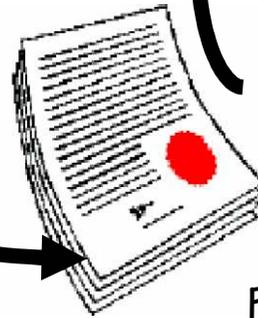
Processus de calcul



Une structure ?



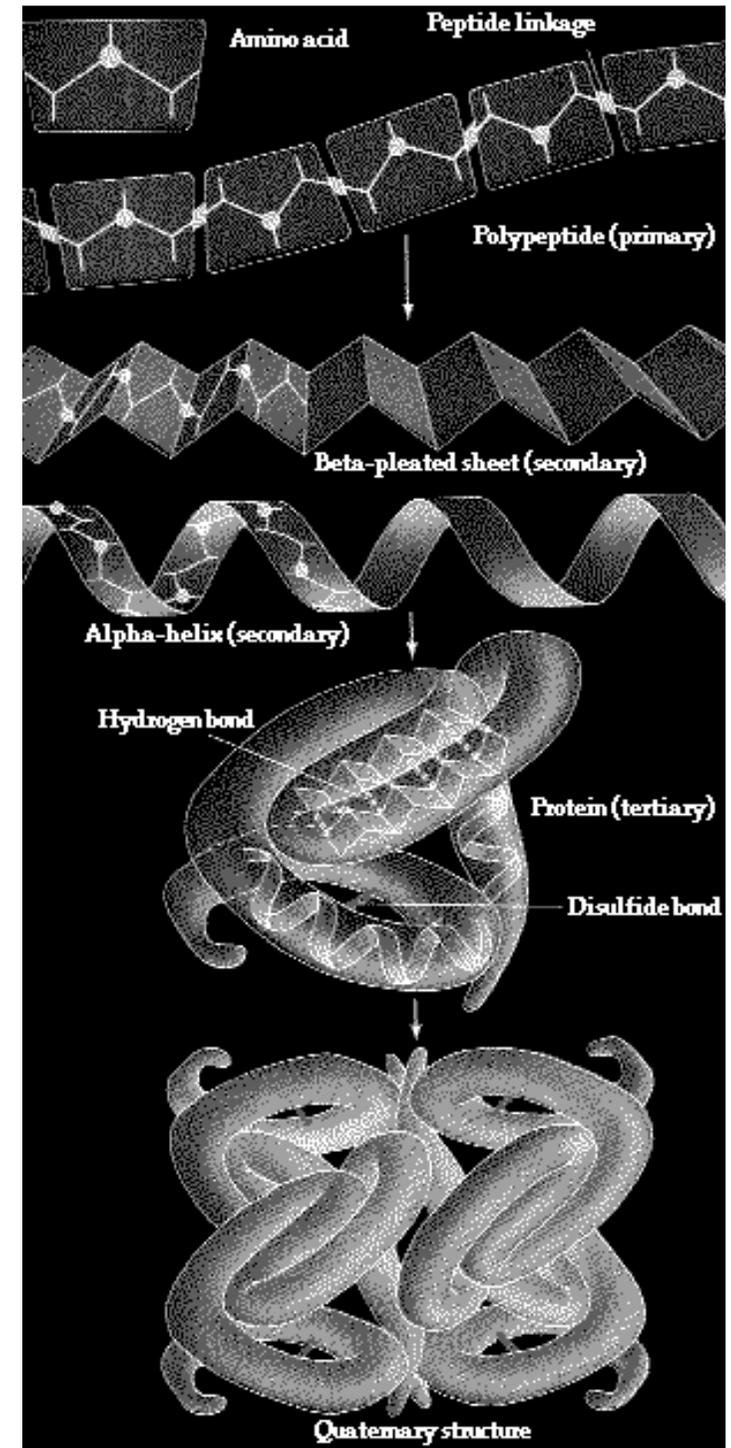
Banques de données



Fichiers

!! RAPPEL !! Niveaux d'organisation

- Primaire : Séquence en acides aminés,
Présence de ponts disulfure.
- Secondaire : Arrangement spatial des acides aminés proches dans la séquence
Hélices α , feuillets β , boucles...
- Tertiaire : Arrangement spatial de structures II^{re},
Structure 3D de la protéine
- Quaternaire : Arrangement spatial des différentes sous-unités constituant une protéine à plusieurs chaînes peptidiques



!! RAPPEL !!

Les acides aminés

Généralités

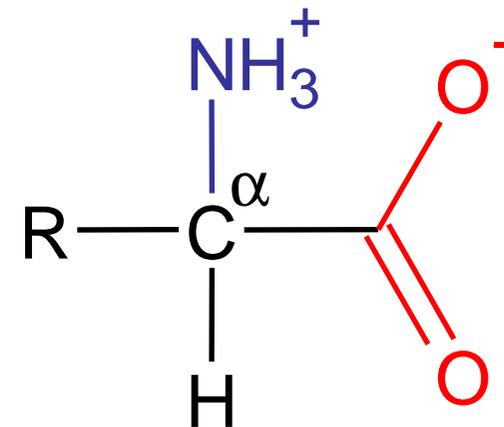
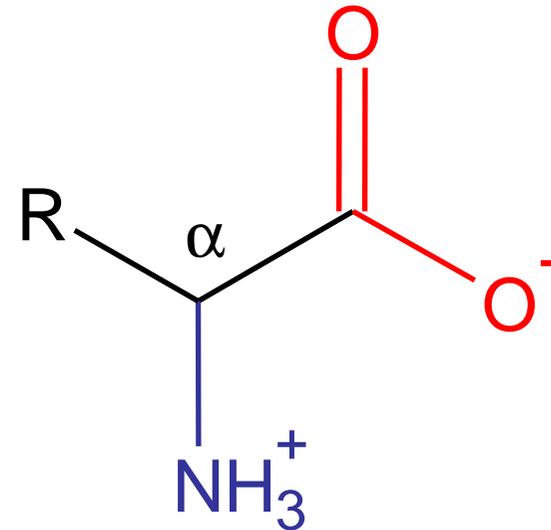
20 Acides aminés

Unités structurales de base des protéines

Une fonction acide (COOH)

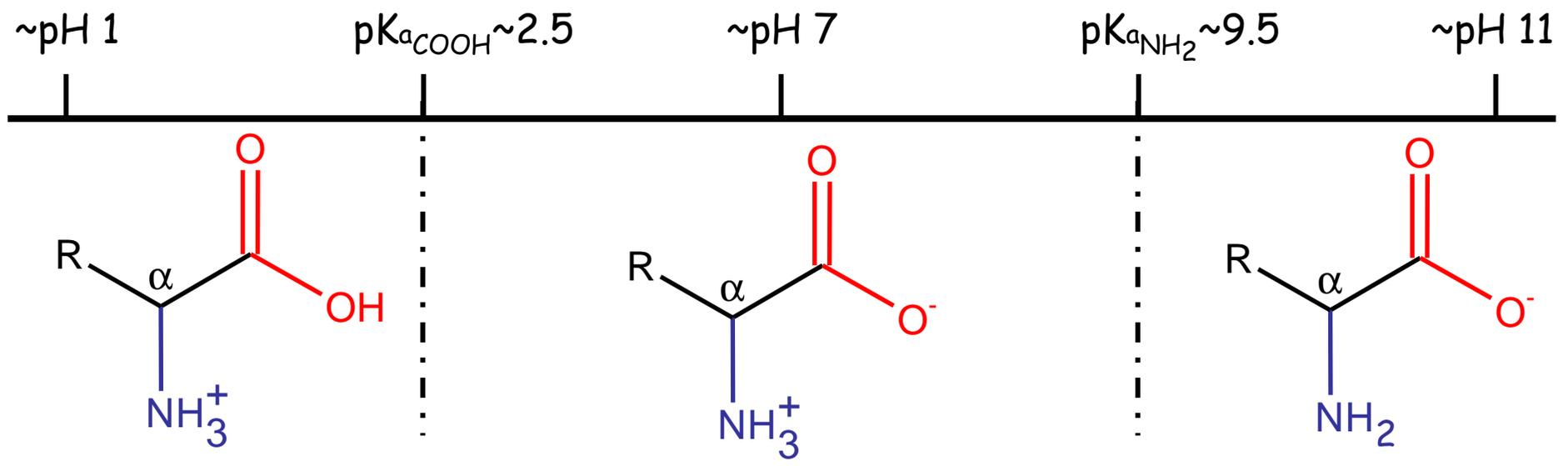
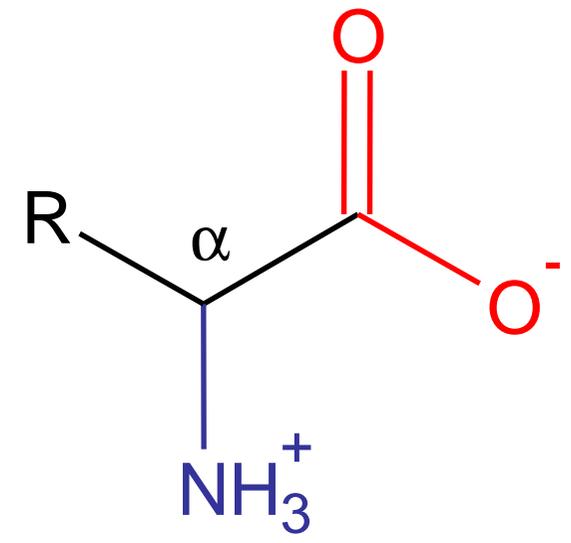
Une fonction basique (NH₂)

Une chaîne latérale variable (R)



!! RAPPEL !!

Forme zwitterionique qui évolue en fonction du pH



!! RAPPEL !!

Amino Acid	α -COOH pK_a	α -NH ₃ ⁺ pK_a	R group pK_a
Alanine	2.4	9.7	
Arginine	2.2	9.0	12.5
Asparagine	2.0	8.8	
Aspartic acid	2.1	9.8	3.9
Cysteine	1.7	10.8	8.3
Glutamic acid	2.2	9.7	4.3
Glutamine	2.2	9.1	
Glycine	2.3	9.6	
Histidine	1.8	9.2	6.0
Isoleucine	2.4	9.7	
Leucine	2.4	9.6	
Lysine	2.2	9.0	10.5
Methionine	2.3	9.2	
Phenylalanine	1.8	9.1	
Proline	2.1	10.6	
Serine	2.2	9.2	~13
Threonine	2.6	10.4	~13
Tryptophan	2.4	9.4	
Tyrosine	2.2	9.1	10.1
Valine	2.3	9.6	

!! RAPPEL !!

Calcul du pI des acides aminés, peptides et protéines

pI, pH auquel la charge nette des acides aminés, peptides ou protéines est nulle

Pour les acides aminés :

pI = moyenne des 2 pKa entourant la forme neutre de l'acide aminé

Pour les peptides :

Regarder la charge nette en fonction du pH

Pour les grands peptides et les protéines :
Prédiction du pI, site dédié à la détermination en fonction de la séquence en acides aminés

!! RAPPEL !!

Classification biochimique des 20 acides aminés

Plusieurs facteurs de variabilité :

Taille (longueurs des chaînes, encombrement stérique...)

Forme (encombrement stérique)

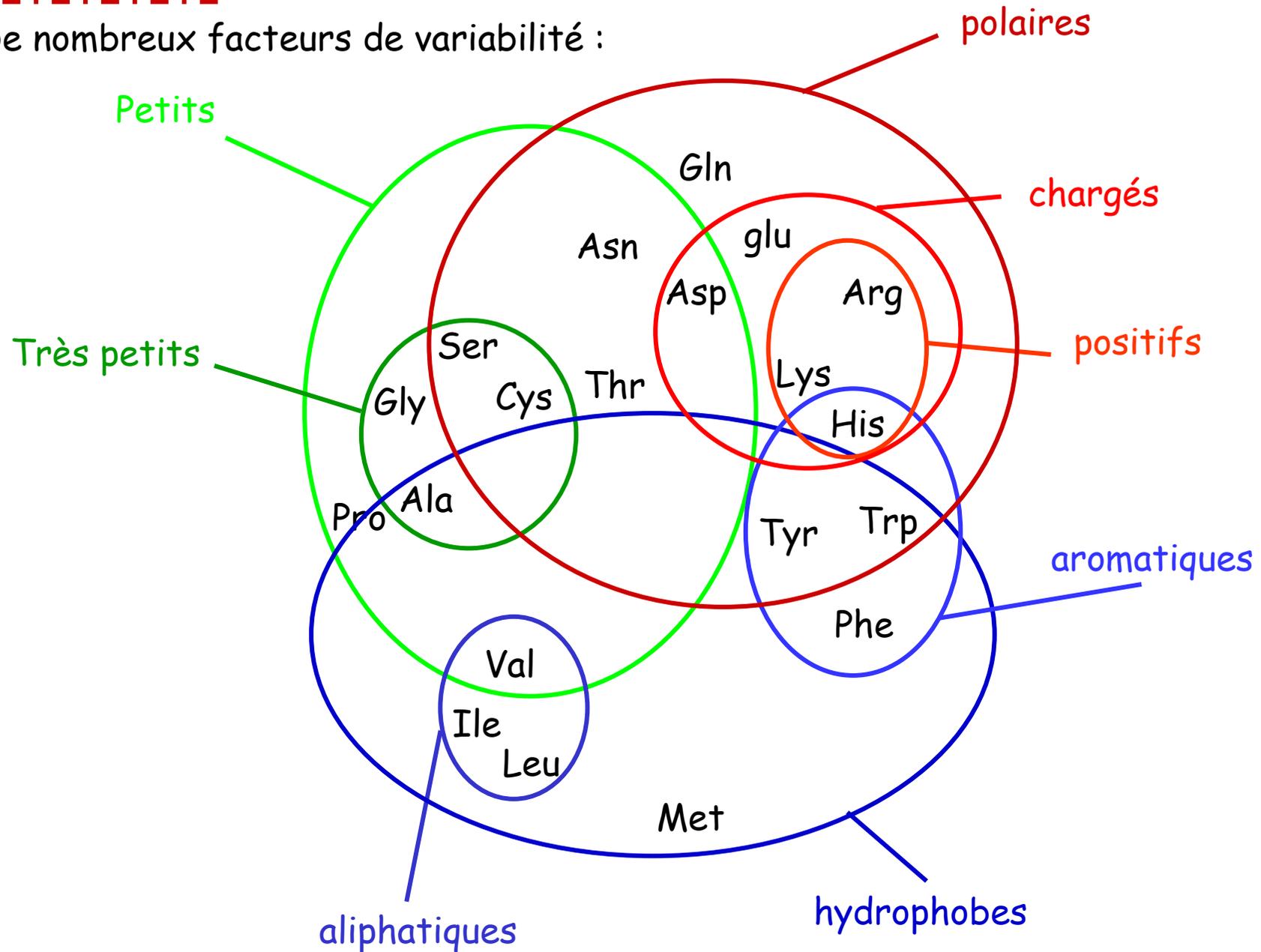
Affinité aux solvants (polarité, hydrophobicité vs. hydrophilicité)

Charges (acide ou base, dipolaire...)

Réactivité chimique

!! RAPPEL !! Classification des 20 acides aminés :

De nombreux facteurs de variabilité :



Prédiction de topologie transmembranaires : Echelle de Kyte et doolittle (1982)

Recherche d'hélices α transmembranaires hydrophobes :
régions de 20 à 30 acides aminés

Etablissement d'un profil hydrophobicité de la protéine grâce à
un classement des acides aminés en fonction de leur
hydrophobicité (Kyte et Doolittle)

Coefficient de
partage d'un acide
aminé entre un
solvant polaire et
apolaire

Ile: 4.500
Val: 4.200
Leu: 3.800
Phe: 2.800
Cys: 2.500
Met: 1.900
Ala: 1.800

Gly: -0.400
Thr: -0.700
Ser: -0.800
Trp: -0.900
Tyr: -1.300
Pro: -1.600

His: -3.200
Asp: -3.500
Glu: -3.500
Asn: -3.500
Gln: -3.500
Lys: -3.900
Arg: -4.500

!! RAPPEL !!

Peptides et protéines

Généralités

* Bio-molécules abondantes et diversifiées dans la cellule, 50% du poids sec de la cellule.

* Rôles variés et multiples

* Degré d'organisation :

Acide aminé, dipeptides (2 acides aminés)...

oligo et poly-peptides (jusqu'à une centaine d'Aa...)

protéines (une ou plusieurs chaînes polypeptidiques)

* Le Dalton, masse d'un atome d'H, protéine de 1500 g.mol⁻¹, masse de 1500 Da ou 1.5 kDa

* Localisations cellulaires et formes des protéines : solubles (fibreuse ou globulaire)

membranaires (périphériques, trans)

!! RAPPEL !!

Peptides et protéines



Insulin



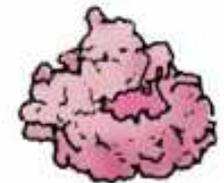
Cytochrome *c*



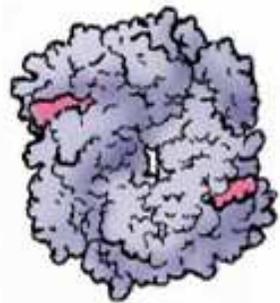
Ribonuclease



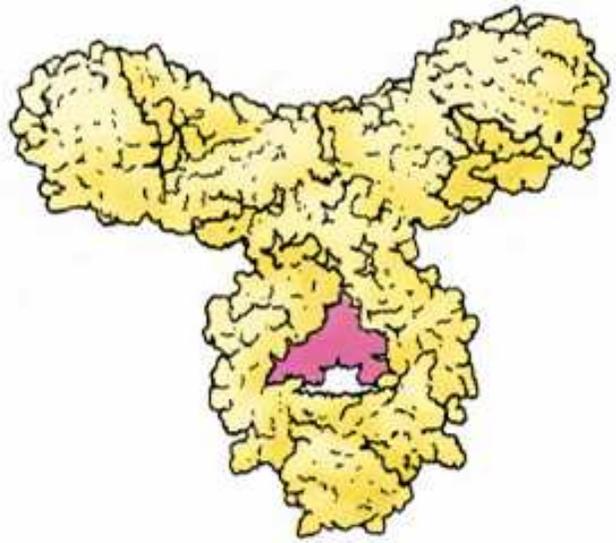
Lysozyme



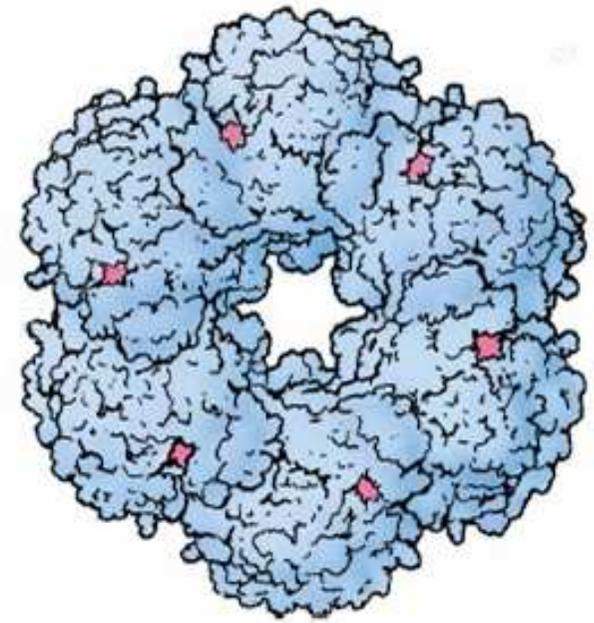
Myoglobin



Hemoglobin



Immunoglobulin



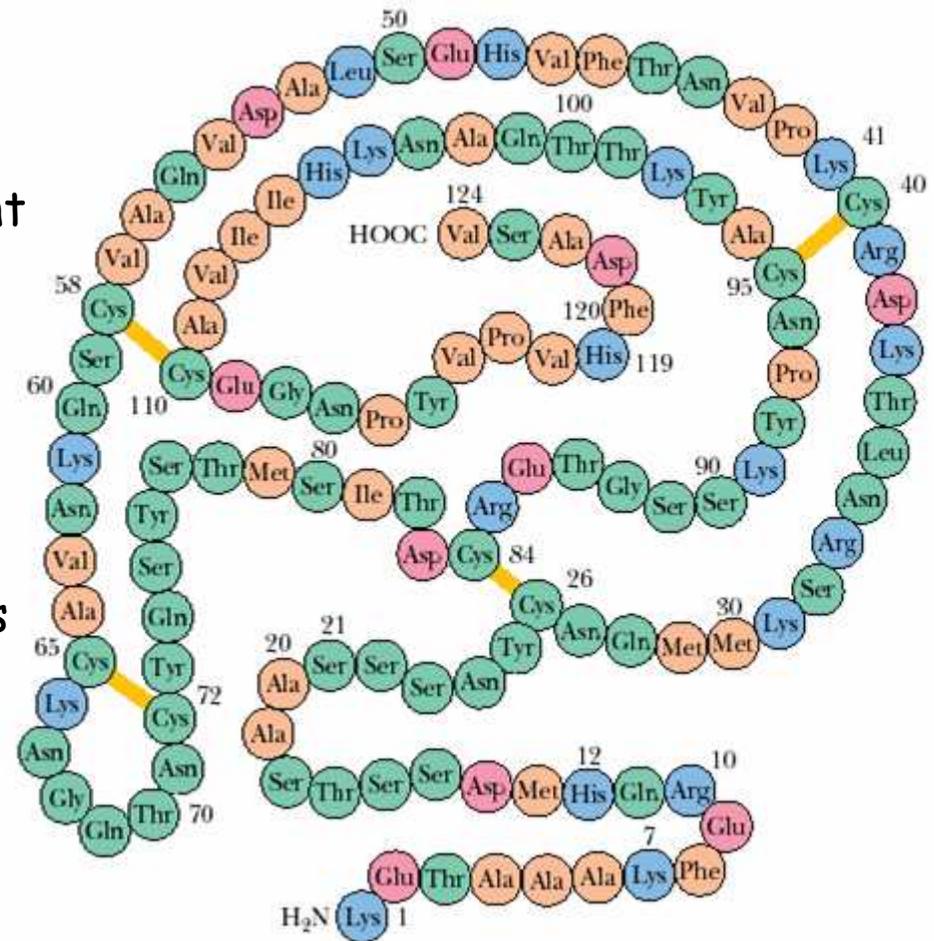
Glutamine synthetase

5,0 nm

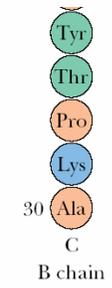
Structures Primaires : La séquence peptidique



- * Chaîne linéaire sans ramification
- * Une seule direction
- * Les deux extrémités sont chimiquement différentes
- * Début de la chaîne NH₂ terminal
- * Fin de la chaîne, COOH terminal
- * Les extrémités peuvent être modifiées suite à la synthèse
- * Peptide AGCR différent du peptide RCGA
- * Présence de ponts disulfures
- * Masse moyenne d'un résidu : 110 Da



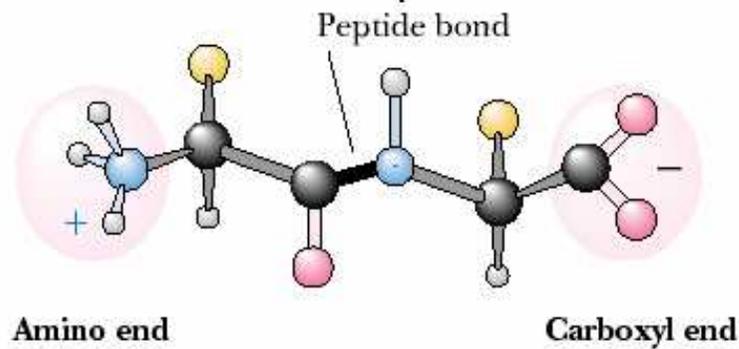
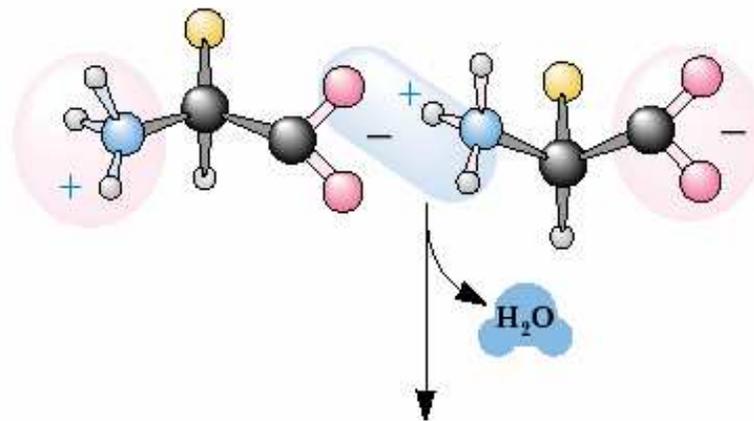
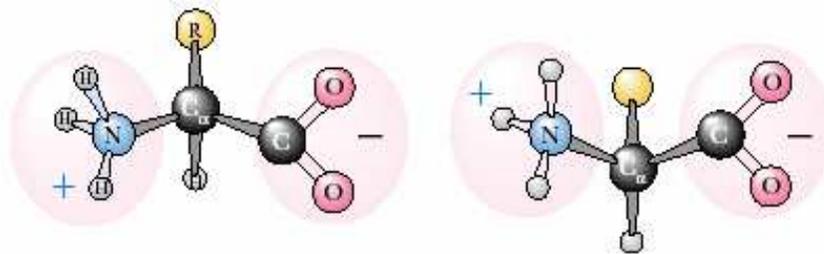
Ribonucléase pancréatique de bovine



!! RAPPEL !!

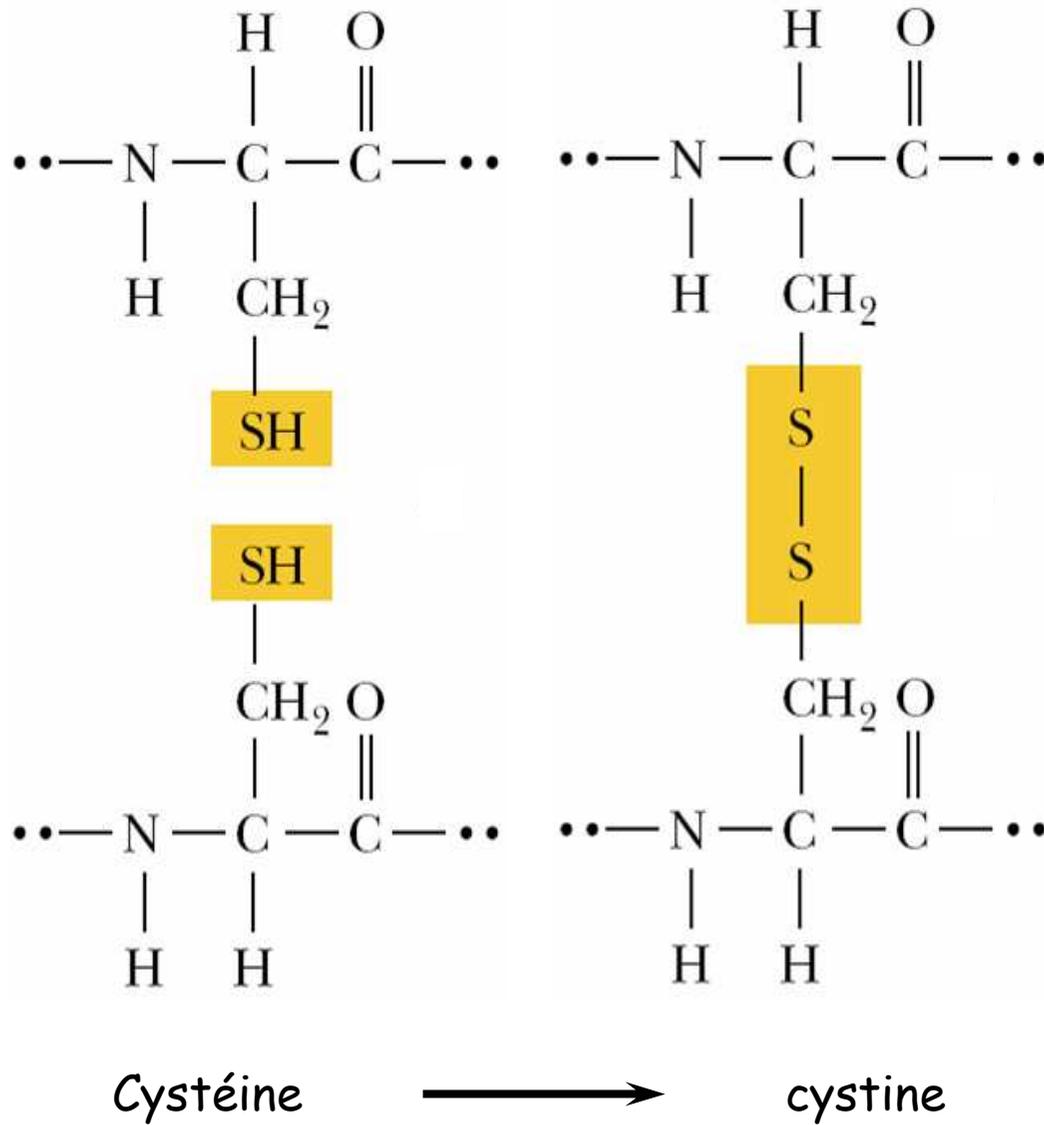
!! RAPPEL !!

La liaison peptide



!! RAPPEL !!

Ponts disulfures



Représentation des séquences primaires

Convention d'écriture : NH_2 terminal à gauche, COOH terminal à droite

Bio-Informatique, code à une lettre

Format FASTA :

Première ligne, >le titre

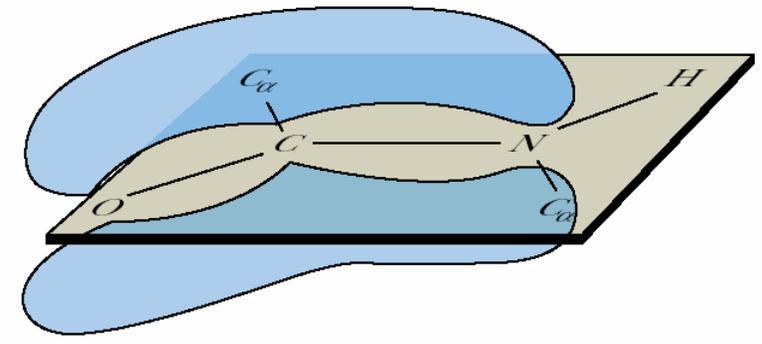
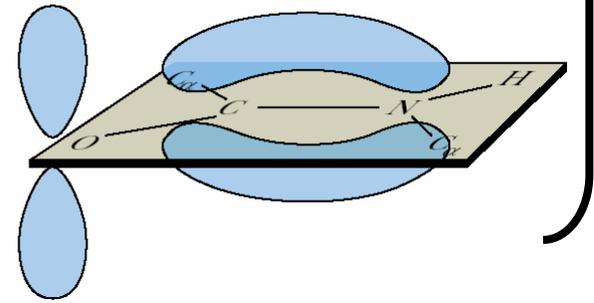
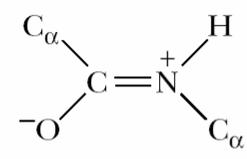
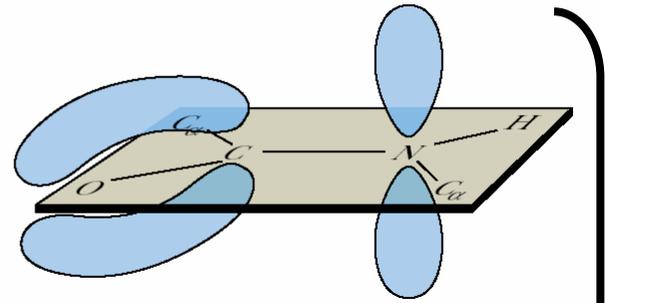
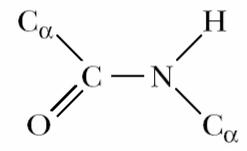
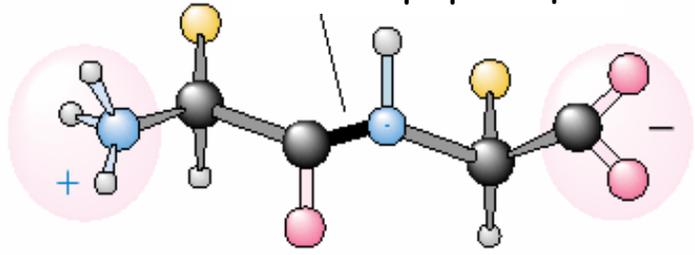
Lignes suivantes, la séquence des acides aminés (code à une lettre) du N vers le C term

```
>sp|P08924|GLB1_LUMTE Globin I, (Globin D) - Lumbricus terrestris  
ECLVTEGLKVKLQWASAFGHAHQRVAFGLELWKGILREHPEIKAPFSRVRGDNIYSPQFG  
AHSQRVLSGLDITISMLDTPDMLAAQLAHLKVQHVERNLKPEFFDIFLKHLLHVLGDRLG  
THFDFGAWHDCVDQIIDGIKDI
```


!! RAPPEL !!

La liaison peptidique forme un plan

Liaison peptidique



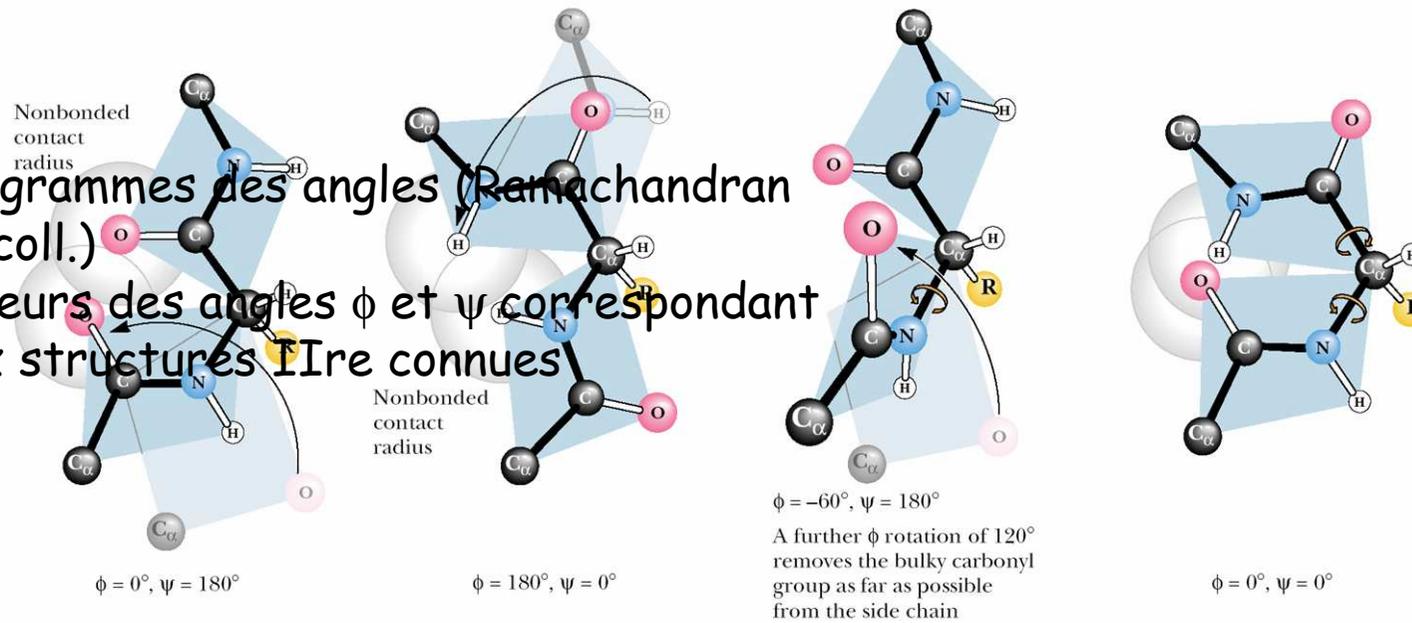
!! RAPPEL !!

Les angles ϕ et ψ

Rotation autour du carbone α des plans des liaisons peptidiques

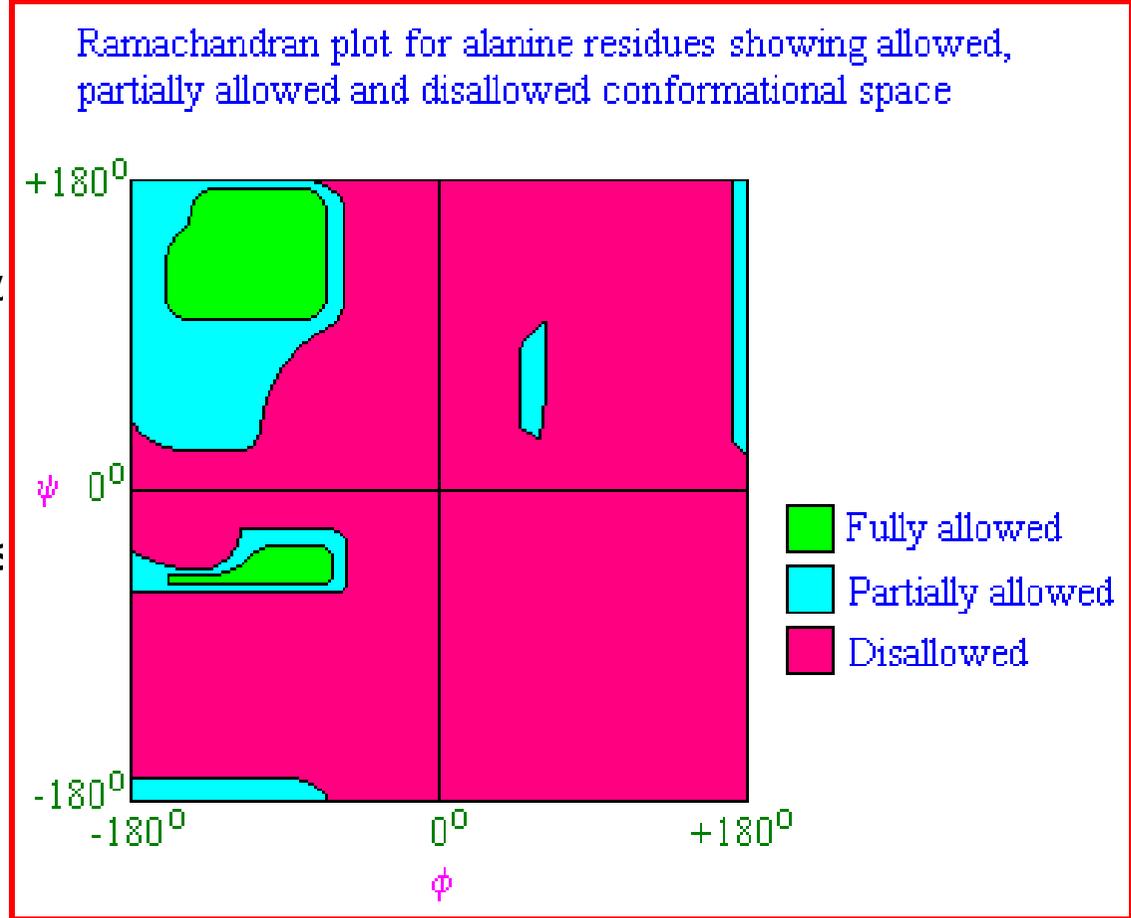
Il existe des valeurs de ϕ et ψ interdites

Diagrammes des angles (Ramachandran et coll.)
Valeurs des angles ϕ et ψ correspondant aux structures IIire connues



!! RAPPEL !!

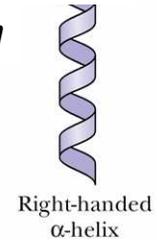
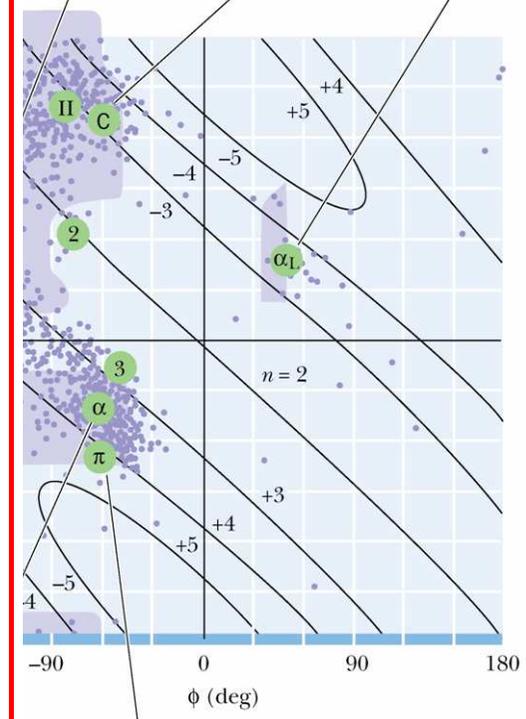
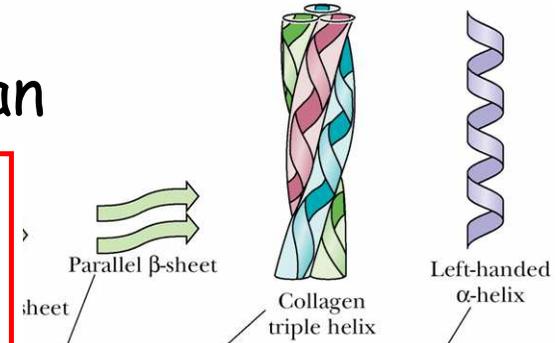
Diagrammes de Ramachandran



Hélice α

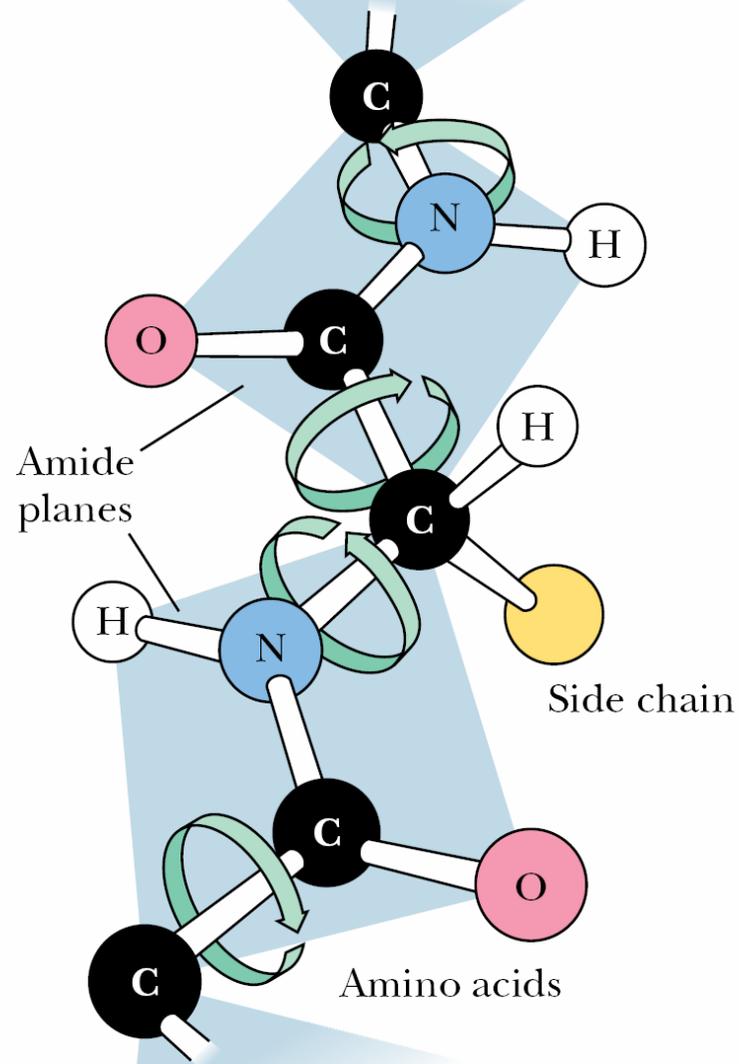
Feuillets

<http://www.fst.rdg.ac.uk/courses/fs916/lect1/lect1.htm>



!! RAPPEL !!

(c)



!! RAPPEL !!

Structures II^{res} :

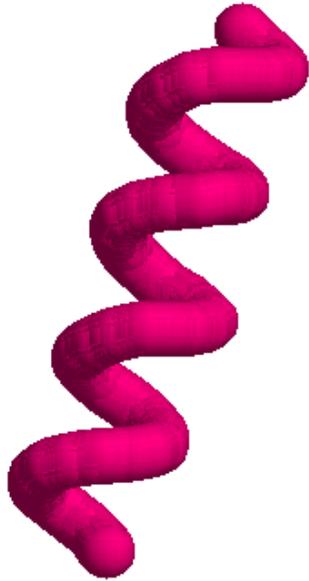
Arrangement spatial des acides aminés proches dans la séquence

Hélices α , feuillets β , boucles...

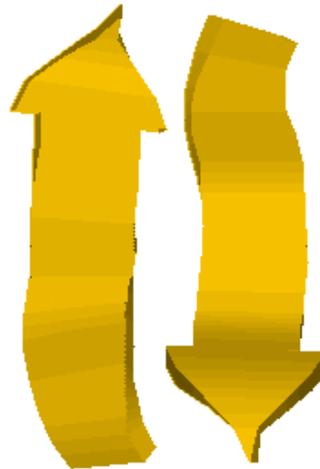
!! RAPPEL !! Rappels: Structure des protéines

La structure secondaire

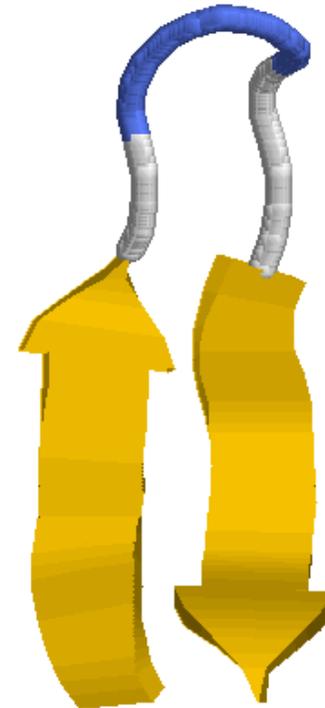
Elle correspond aux repliements (formations périodiques) qu'adoptent des portions partielles d'une protéine donnée.



Hélice α

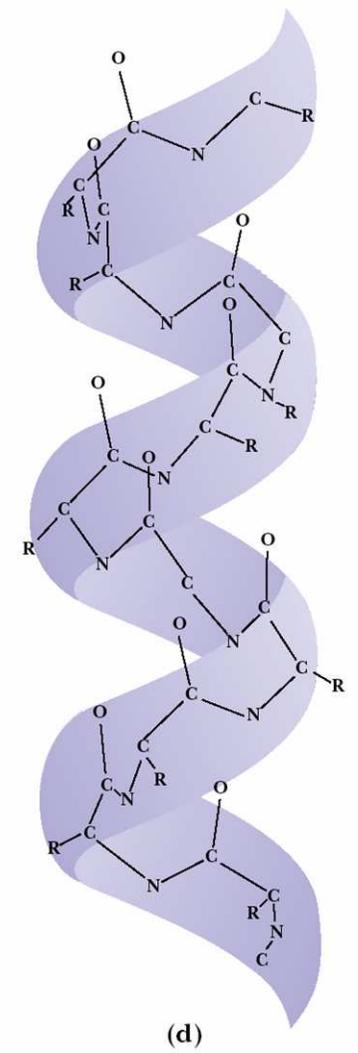
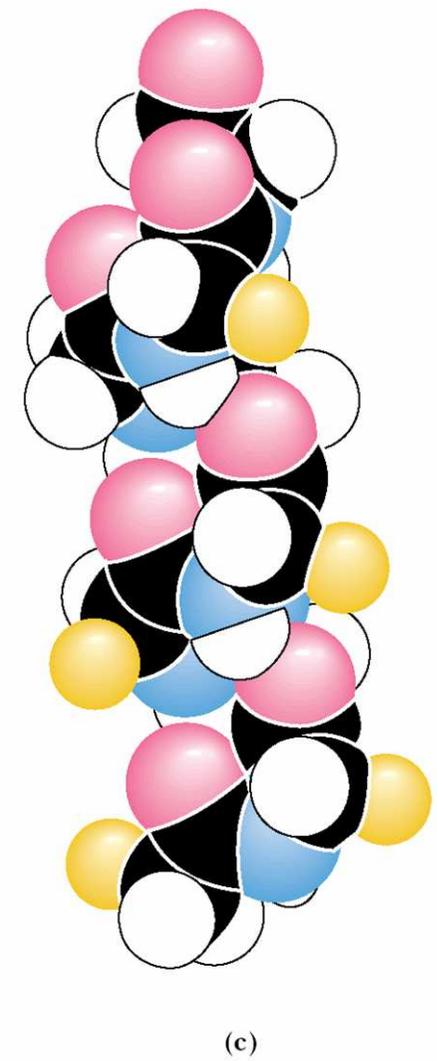
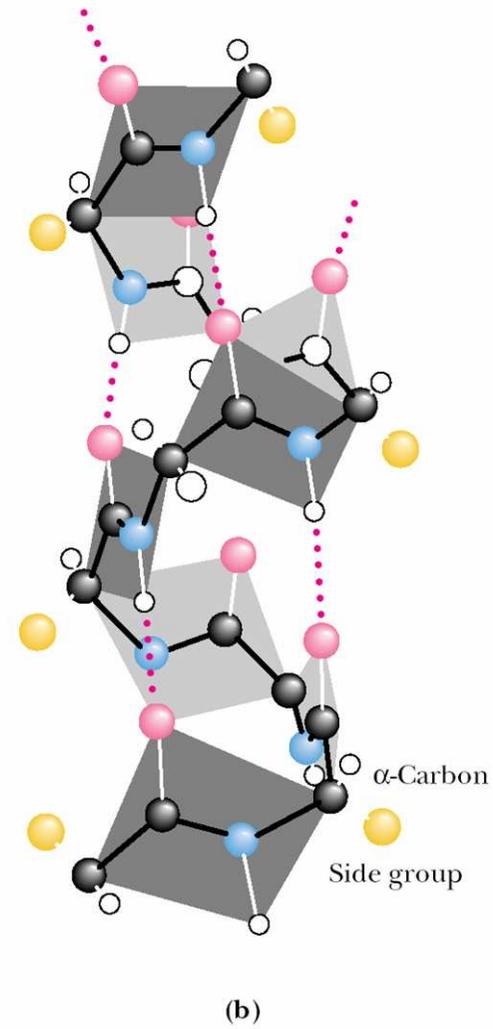
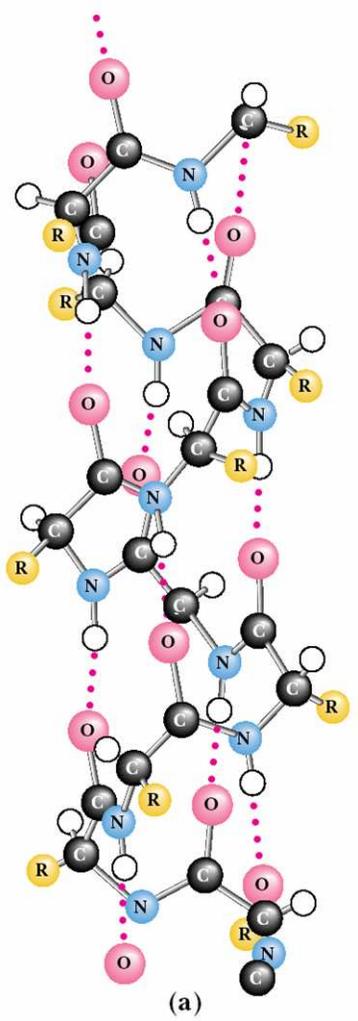


Brin β



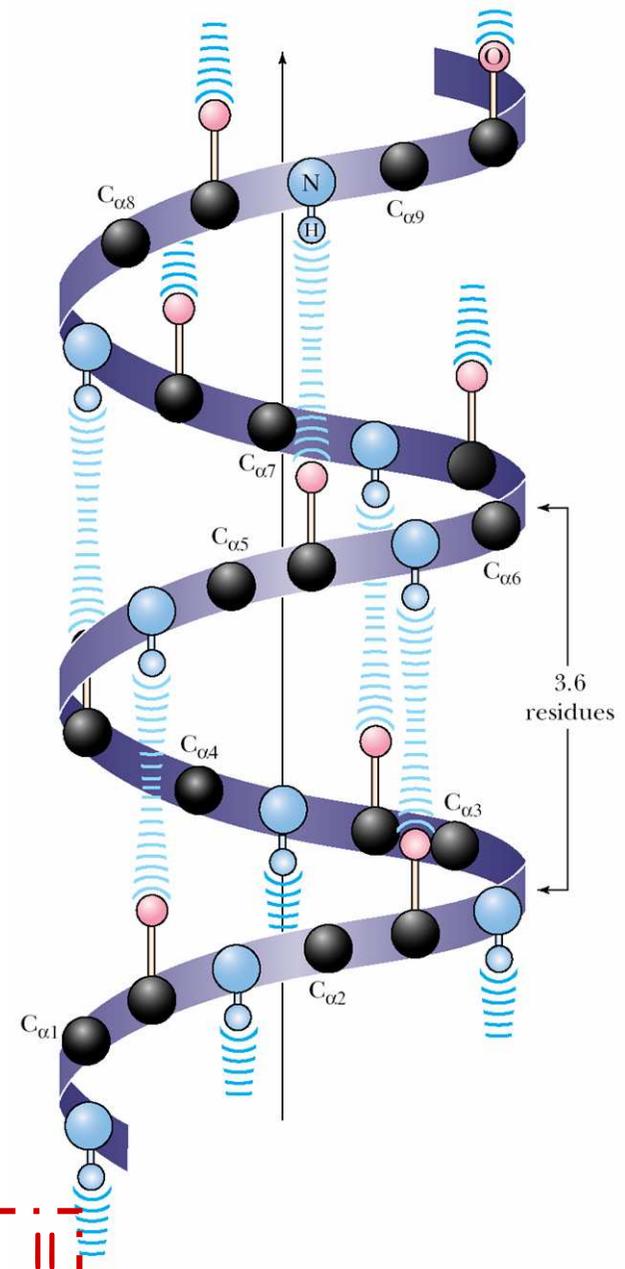
Boucle (*turn*)

L'hélice α **!! RAPPEL !!**



Caractéristiques de l'hélice α :

- * Structure proposée par Pauling & Corey, 1951 (observation par diffraction rayon X six ans plus tard).
- * Structure en bâtonnet, la chaîne polypeptidique à l'intérieur, les chaînes latérales vers l'extérieur. Enroulement droit
- * Stabilisation par des liaisons H entre tous les groupements C=O (n) et N-H (n+4)
- * Translation de 1.5\AA le long de l'axe et rotation de 100° d'un résidu à l'autre
- * 3.6 résidus par tour, soit un pas d'hélice de 5.4\AA ($3.6 \text{ résidus} \times 1.5\text{\AA}$)



!! RAPPEL !!

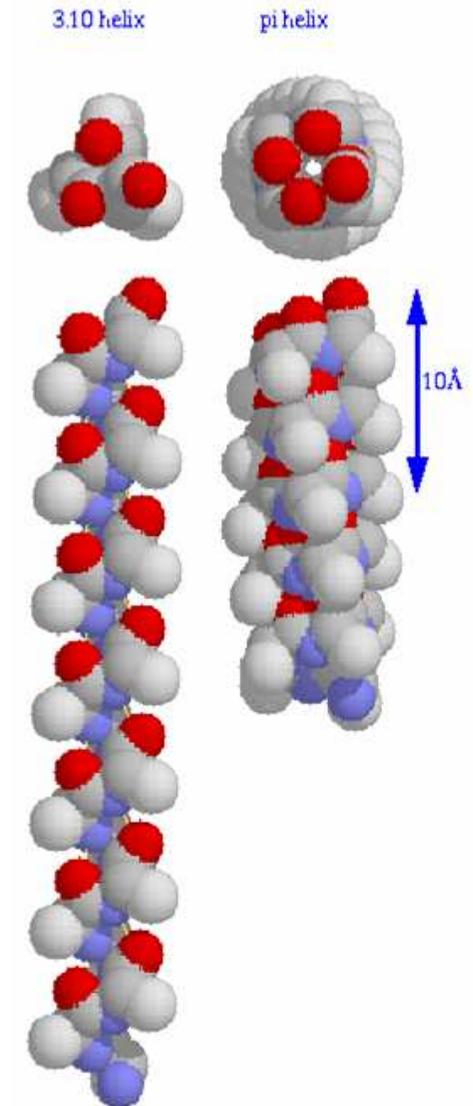
Deux autres structures hélicoïdales extrêmement rares :

!! RAPPEL !!

Hélice 3_{10} : pas de 6\AA , 3 résidus par tour

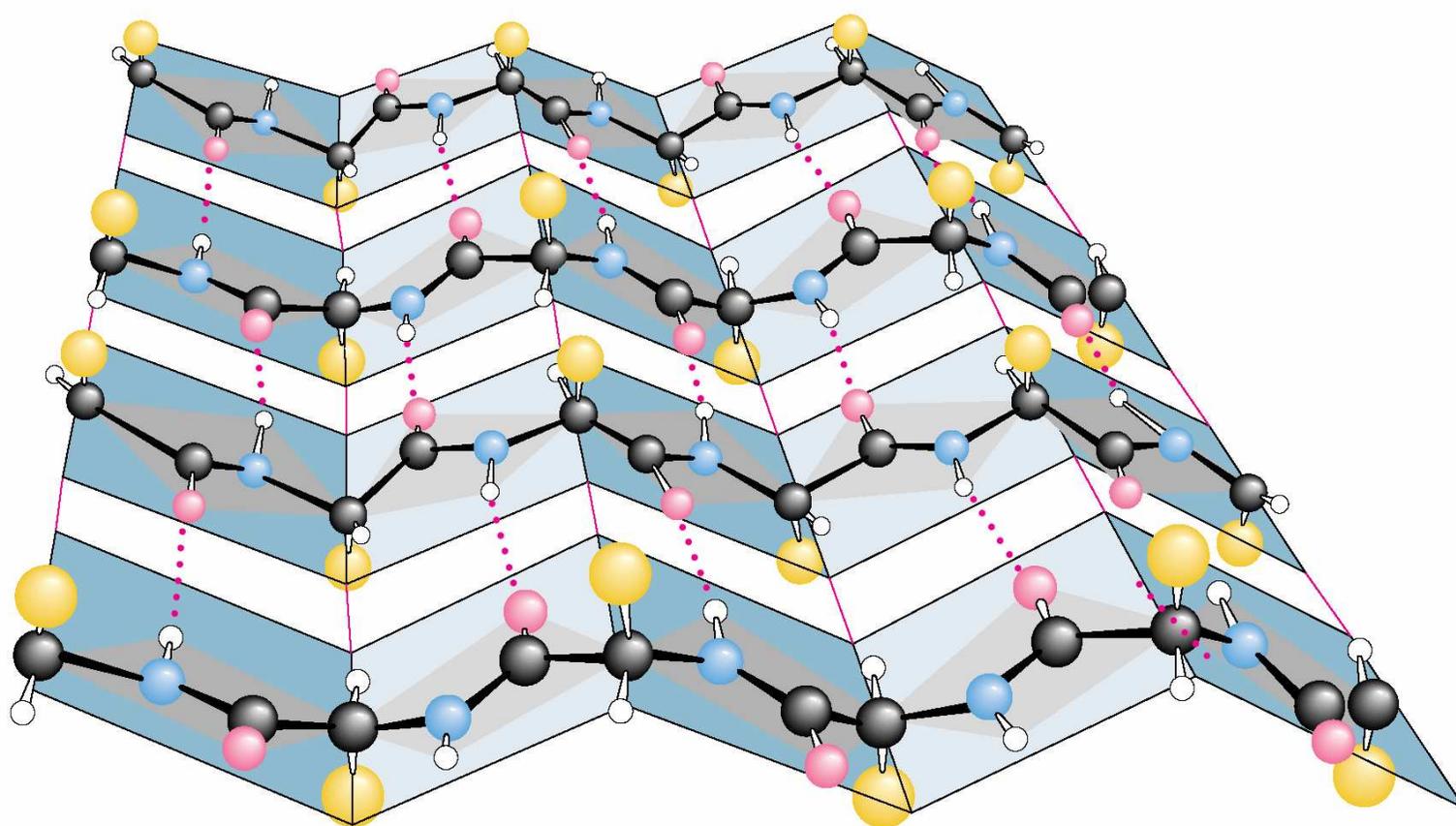
Hélice π : pas de 5.2\AA , 4.4 résidus par tour,

(Hélice α : pas de 5.4\AA , 3.6 résidus par tour)



Le feuillet β :

!! RAPPEL !!



!! RAPPEL !!

Caractéristiques du feuillet β :

* Egalement proposé par Pauling & Corey, 1951

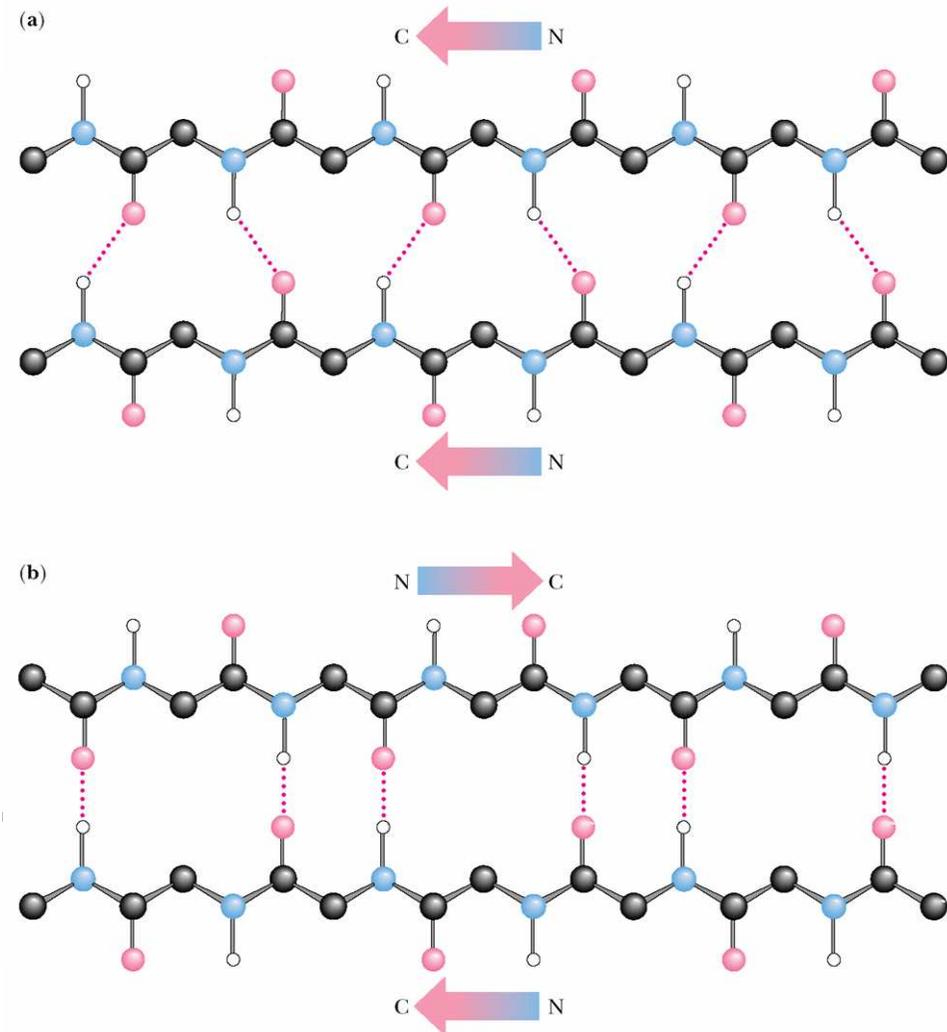
* Chaîne polypeptidique étendue, distance entre 2 acides aminés adjacents, 3.5\AA , (1.5\AA dans l'hélice α)

* Unité répétitive de 7\AA

* Stabilisation par liaison H entre C=O et N-H de deux brins différents

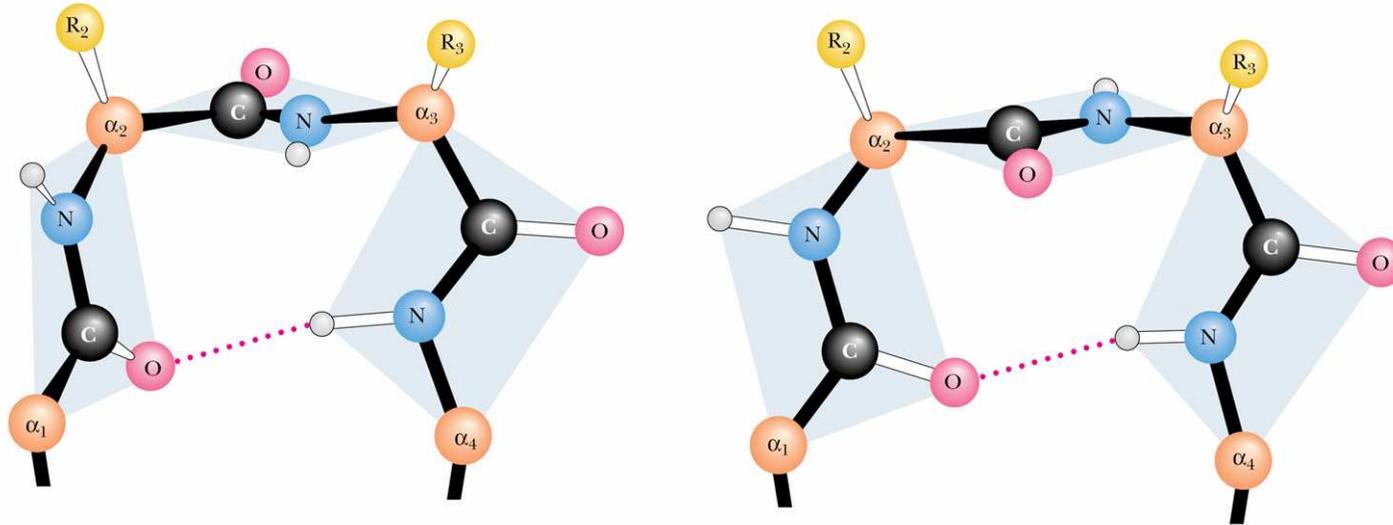
* Les $C\alpha$ alternativement de part et d'autre du plan médian, chaînes latérales de part et d'autre du feuillet et sur des côtés opposés.

* Feuilletes parallèles ou anti-parallèles



!! RAPPEL !!

Boucle β ou le tour β (bend or turn)



- * Favorisés par la présence de *Gly* ou de *Pro* dans la boucle.
- * Structure rencontrée dans les feuilletts b anti-parallèles

Structure des protéines

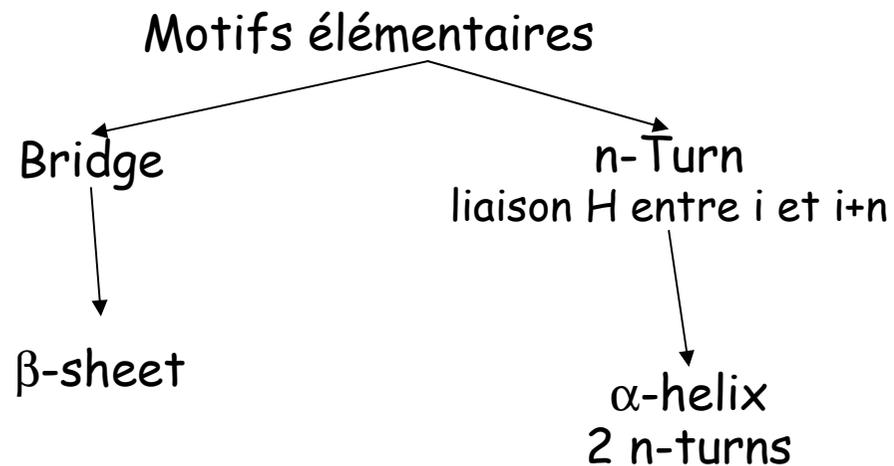
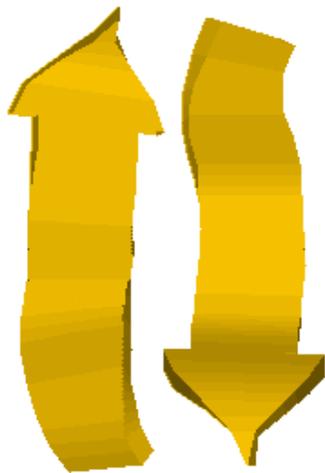
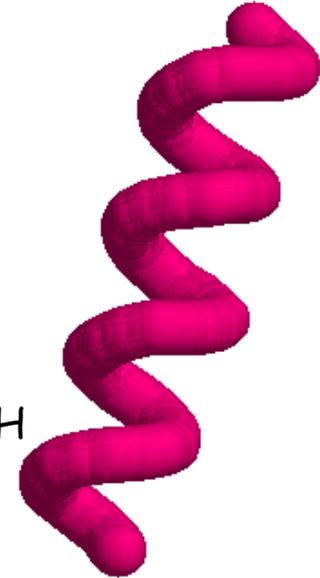
DSSP: Dictionary of Protein Secondary Structure
Kabsch et Sander (1983), *Biopolymers*, 22, 2577

Principe (basé sur les structures tridimensionnelles)

algorithme de reconstruction de structure secondaire

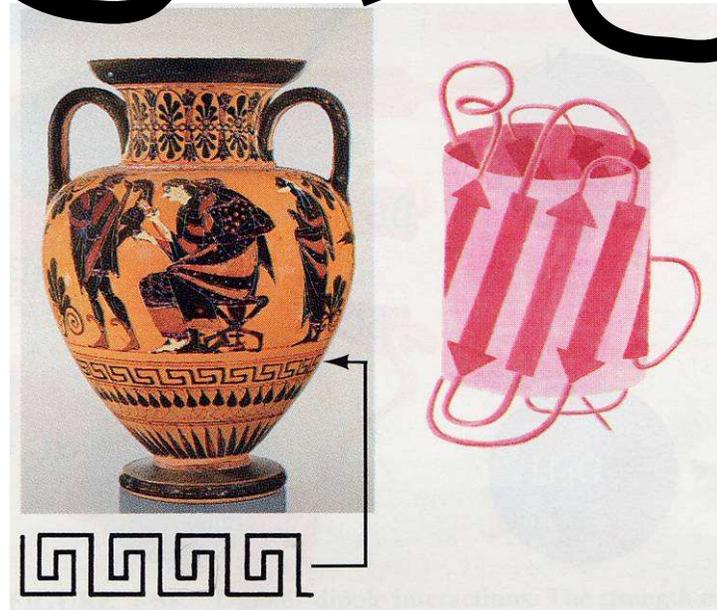
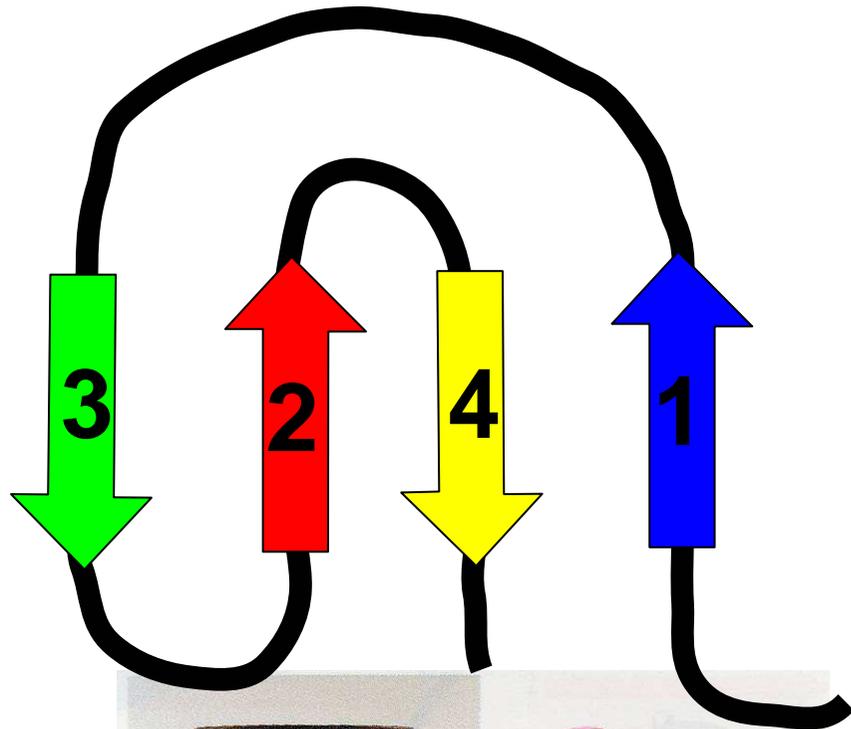
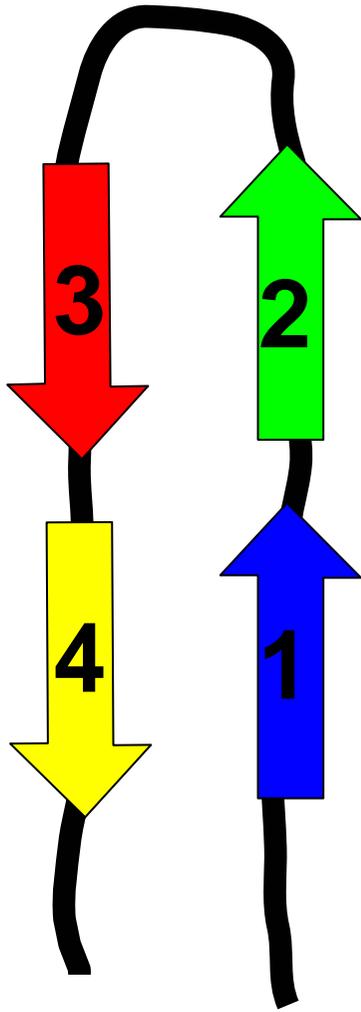
Calcul des liaisons hydrogène entre CO et NH dans la structure

Classification des structures secondaires en fonction des liaisons H

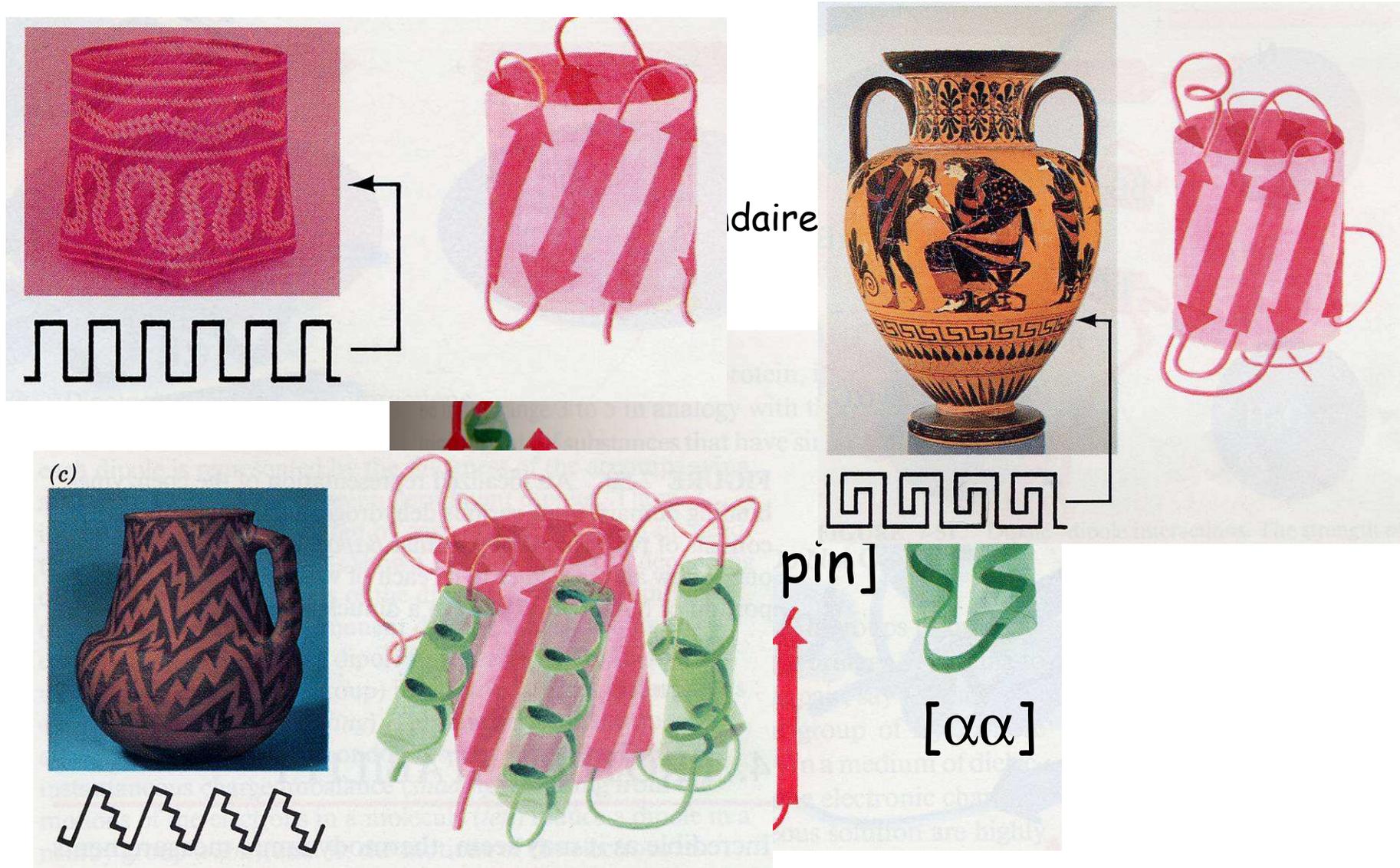


!! RAPPEL !! Rappels: Structure des protéines

La structure suprasecondaire



!! RAPPEL !! NB : les structures supra secondaires



Motifs tonneaux β : ressemblance avec des motifs de vases grecs

!! RAPPEL !! Rappels: Structure des protéines

La structure tertiaire

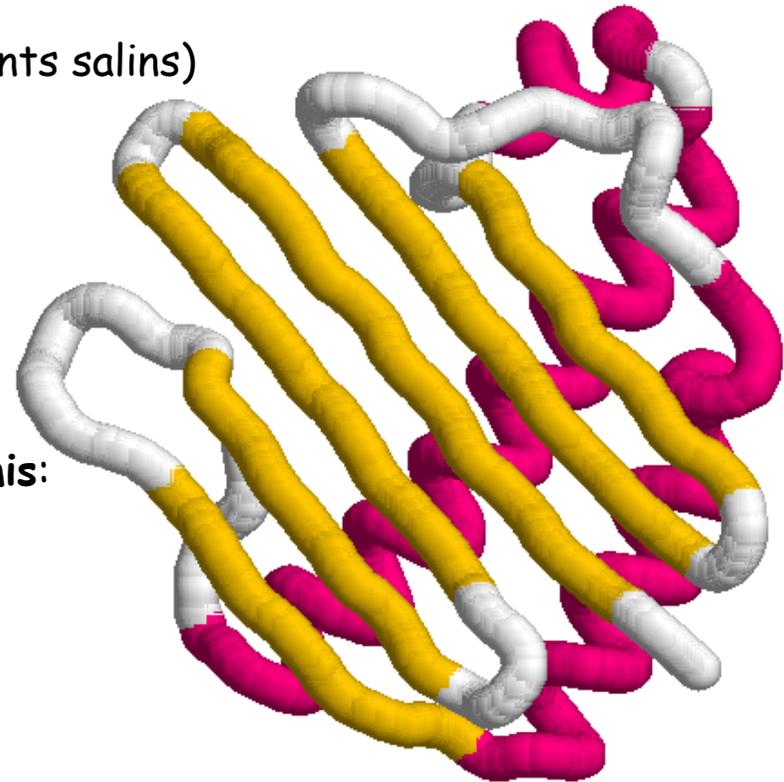
Conformation 3D thermodynamiquement stable qu'adoptent les différents éléments de la structure 2D entre eux pour former la protéine (ou une des sous-unités d'une protéine plus complexe)

La conformation native d'une protéine dépend à la fois de sa séquence et du milieu dans lequel elle est solubilisée.

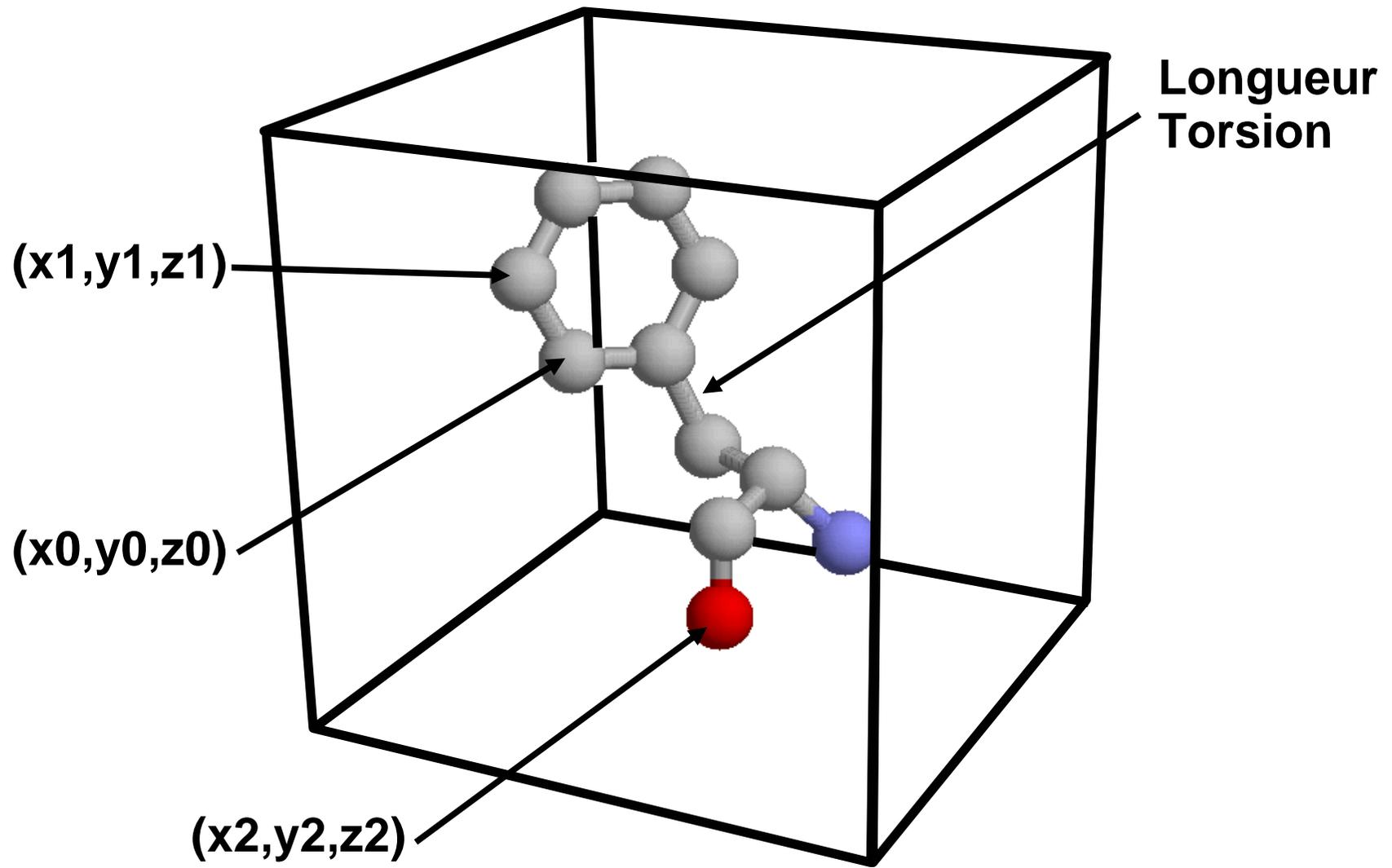
- ★ des liaisons non covalentes (les liaisons H ou les ponts salins)
- ★ des liaisons covalentes (des ponts disulfures)

Le **repliement 3D** (= "fold") est le meilleur **compromis**:

- l'enfouissement des résidus hydrophobes
- les possibilités de rotation autour des liaisons chimiques.



Rappels: Structure des protéines



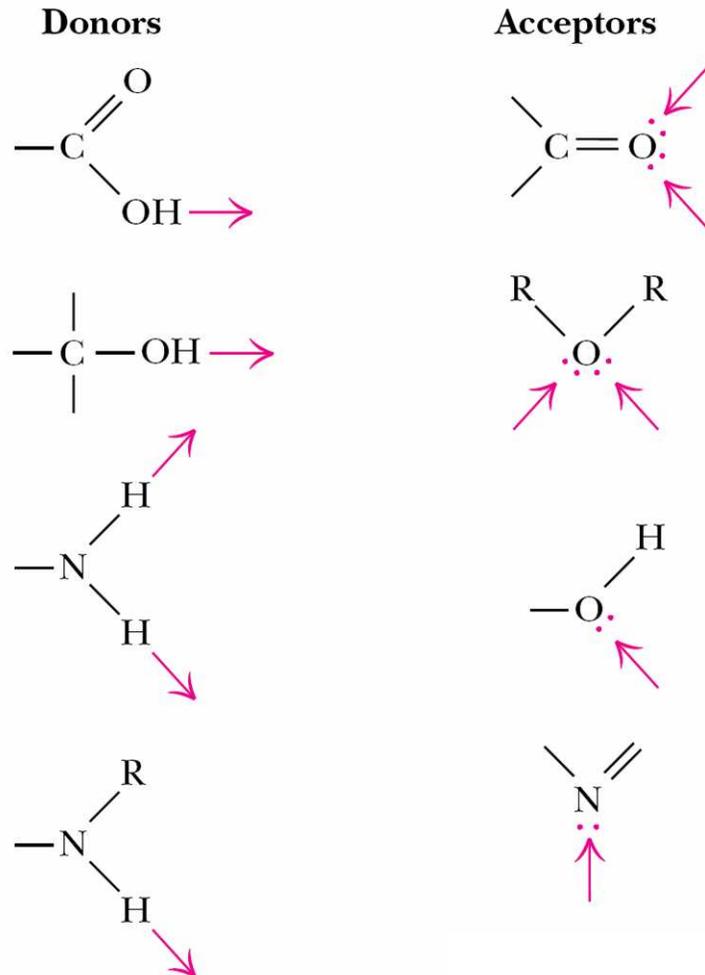
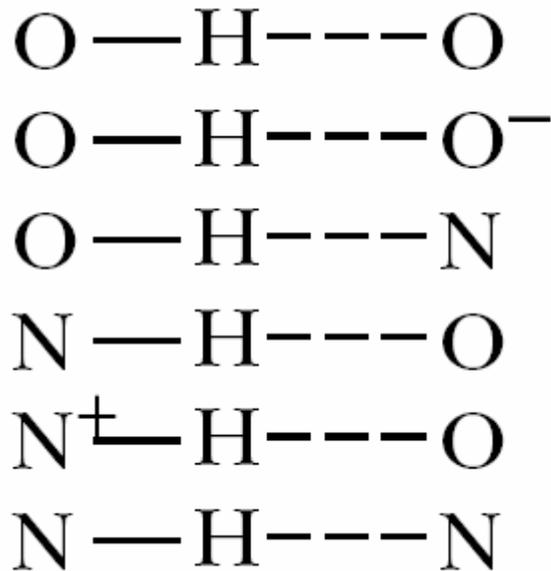
!! RAPPEL !!

Structures tertiaires

« All of the information necessary for folding the peptide chain into its "native" structure is contained in the amino acid sequence of the peptide. » (Anfinsen, 1960s)

Forces mises en jeu :

Liaisons Hydrogènes, dans la structure II^{re},
 $C=O \cdots H-N$ mais aussi entre les chaînes latérales
 pour stabiliser l'agencement en structure III^{re}.

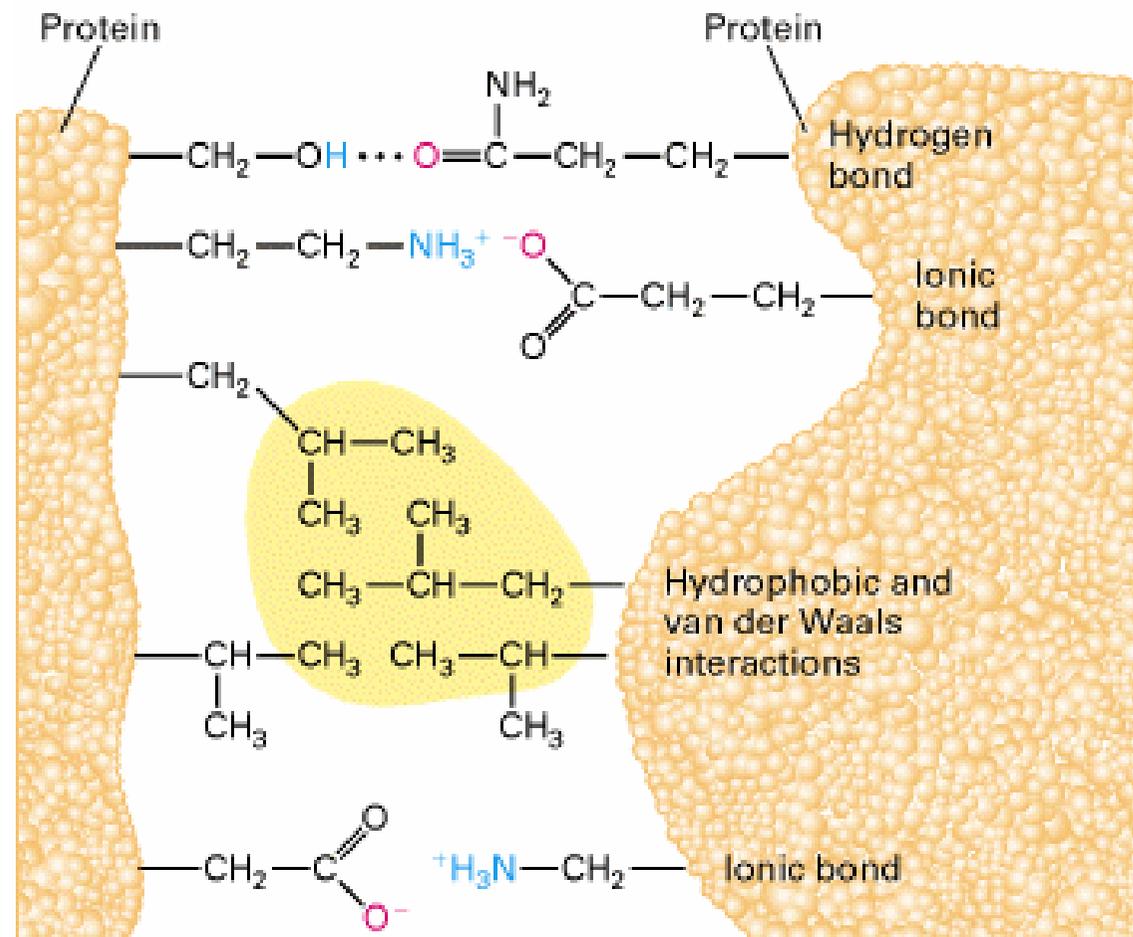


!! RAPPEL !!

Structures tertiaires

Forces mises en jeu :

Interactions hydrophobes, énergiquement favorable d'enfourer les acides aminés non polaires et d'exposer vers l'extérieur les chaînes latérales des acides aminés polaires, vrai pour tous les types de protéines !!!

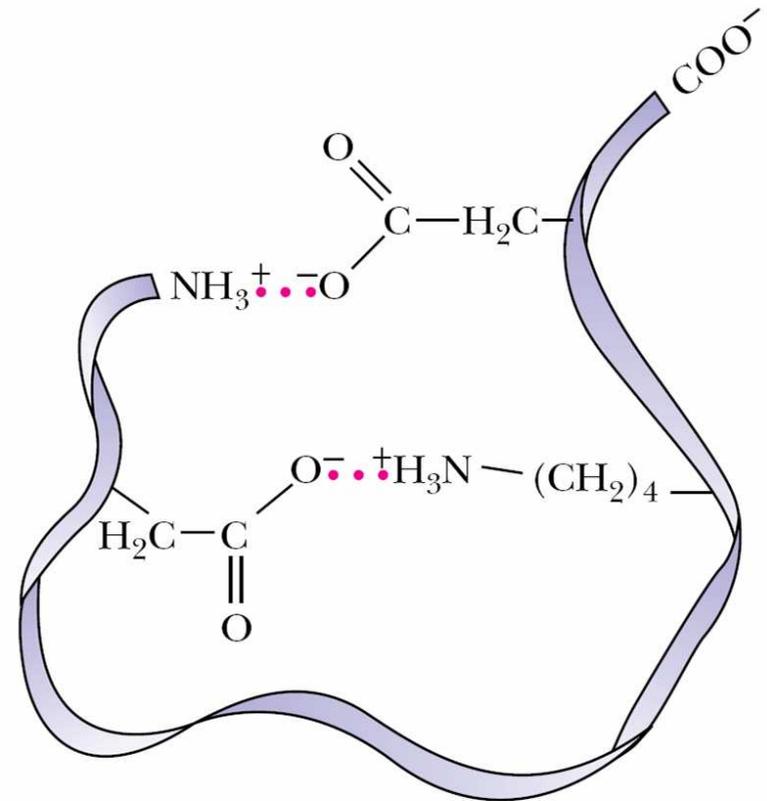


!! RAPPEL !!

Structures tertiaires

Forces mises en jeu :

Interactions électrostatiques, attractives ou répulsives en fonction des charges du pH, intervention des chaînes latérales des acides aminés chargés (basiques et acides)

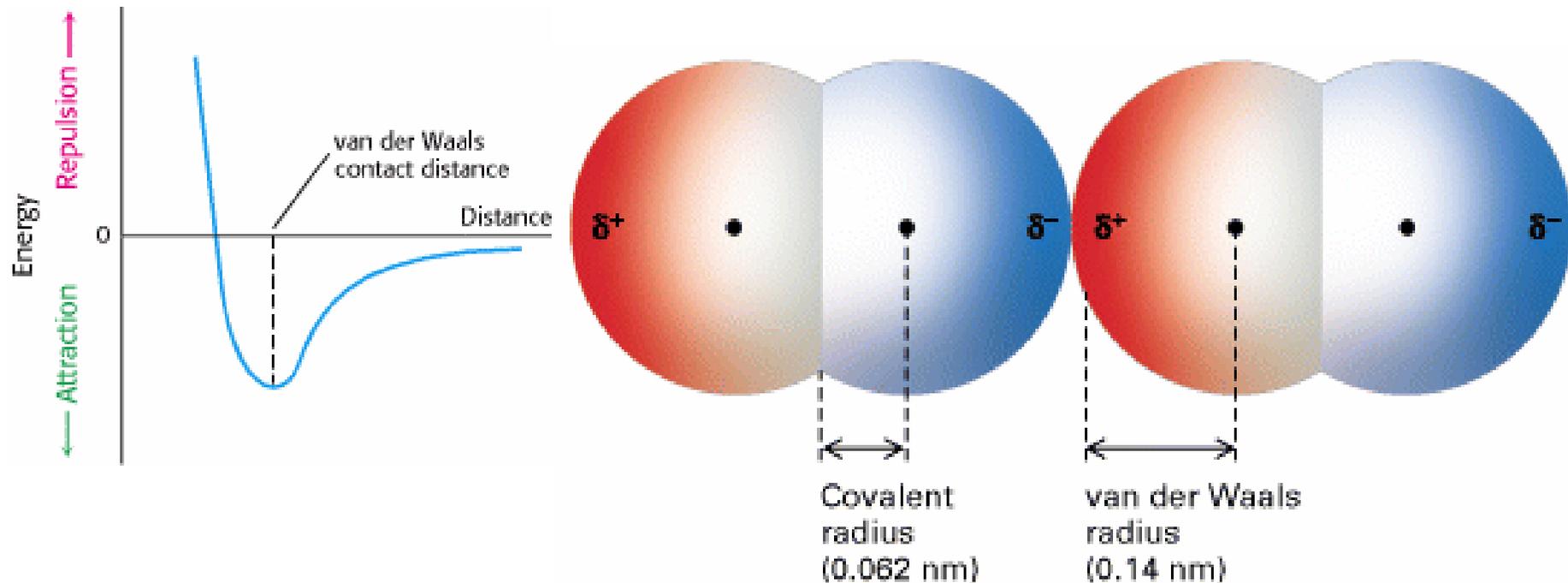


!! RAPPEL !!

Structures tertiaires

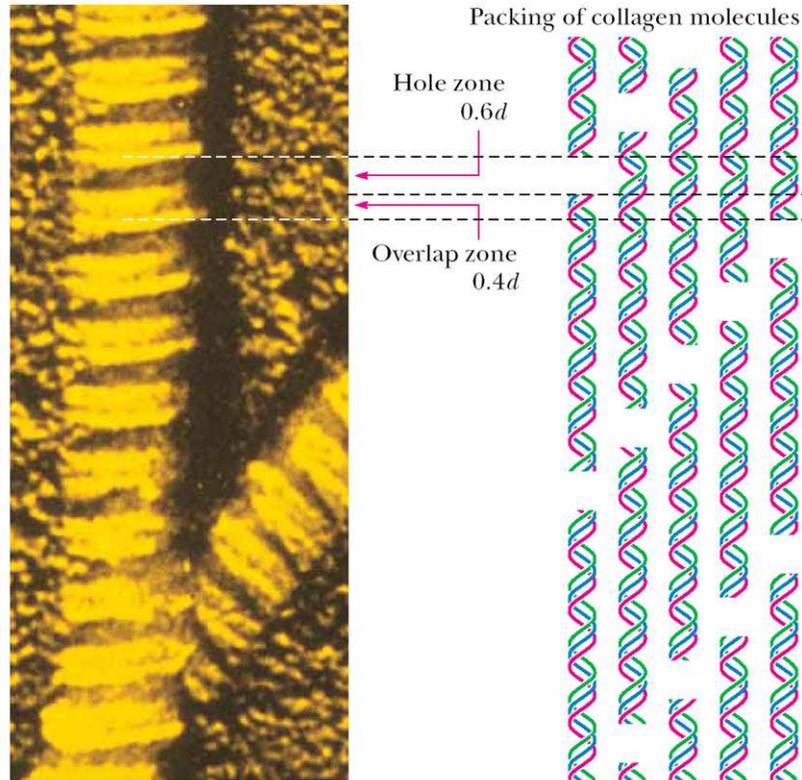
Forces mises en jeu :

Forces de Van der Waals, interactions dipôles/dipôles attractives ou répulsives, de faibles énergies (0.4 à 4.0 kJ.mol⁻¹), rôle à ne pas sous estimer !!!



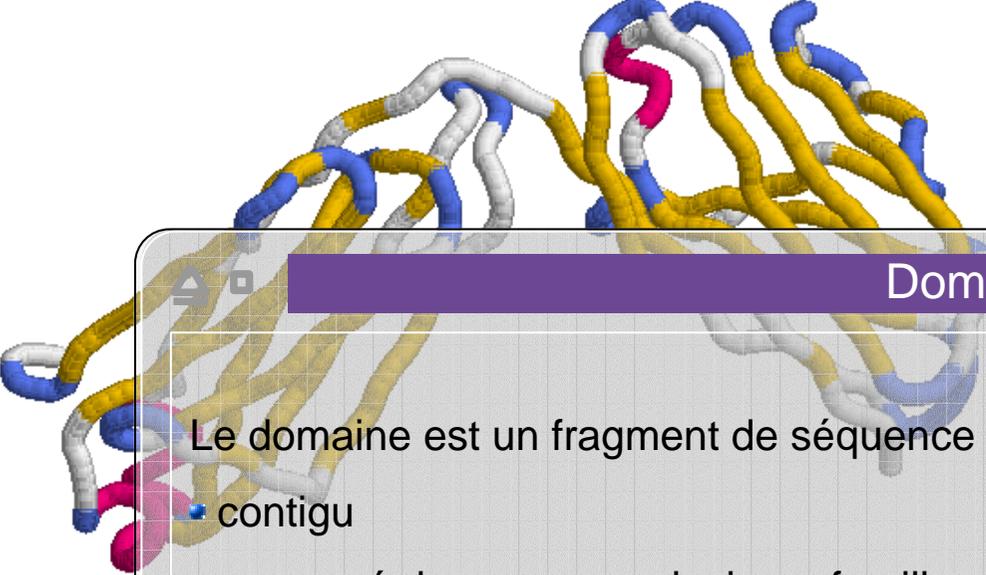
!! RAPPEL !! Structures tertiaires : du "2D" au 3D

Protéines solubles fibreuses (essentiellement un rôle de protéines de structure, le collagène...)



Collagène I

!! RAPPEL !! Rappels: Structure des protéines



Domaine

Le domaine est un fragment de séquence (bloc):

- contigu
- conservé dans une ou plusieurs familles de protéines
- se replie indépendamment (structure secondaire spécifique).

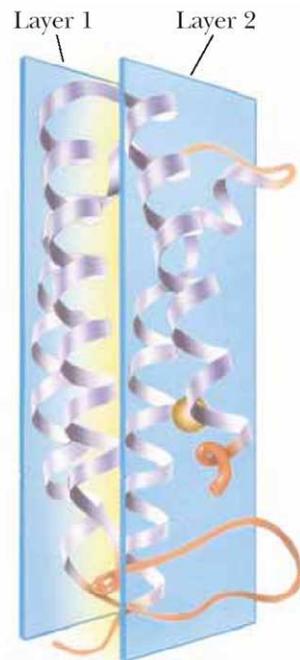
Il peut être dupliqué et réutilisé par des protéines de fonctions différentes (gènes "mosaïque").

The image shows a 3D ribbon diagram of a protein structure. A specific region of the protein is highlighted in yellow and blue, representing a domain. A semi-transparent window titled 'Domaine' is overlaid on the image, containing text that defines a domain as a contiguous fragment of sequence that is conserved in one or more protein families and folds independently. It also notes that domains can be duplicated and reused in proteins with different functions, often found in mosaic genes.

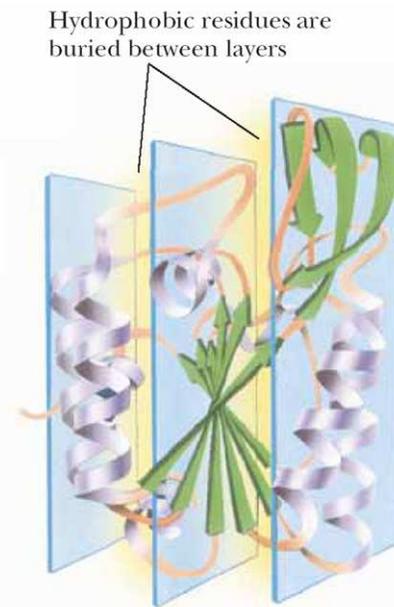
!! RAPPEL !! Structures tertiaires : du "2D" au 3D

Protéines solubles globulaires : Association hélices α , feuillets β , boucles, régions non structurées pour former une protéine fonctionnelle

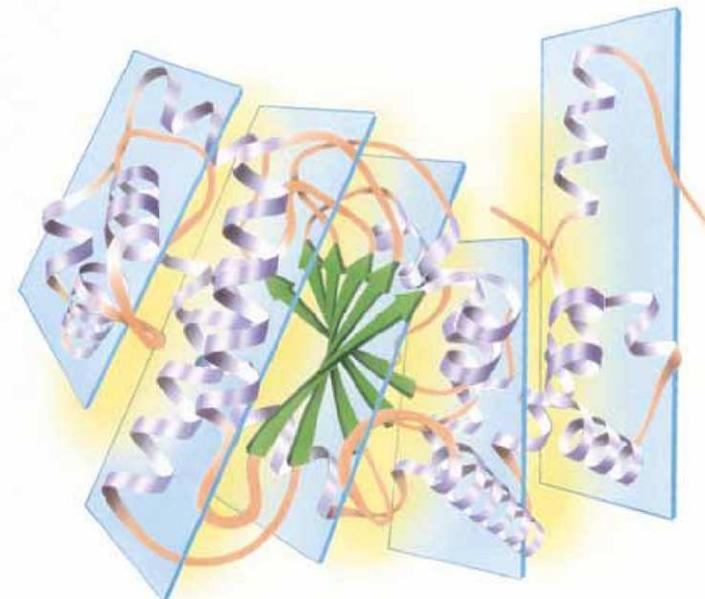
Formation de domaines (100 à 200 résidus, 2 ou plusieurs couches de structures II^{re}, isolement du cœur hydrophobe de la protéine



(a) Cytochrome c'



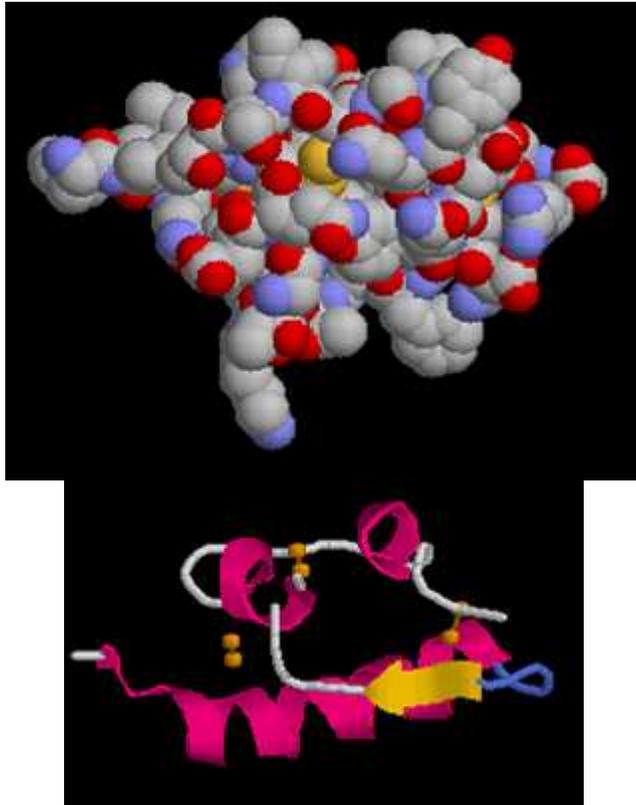
(b) Phosphoglycerate kinase (Domain 2)



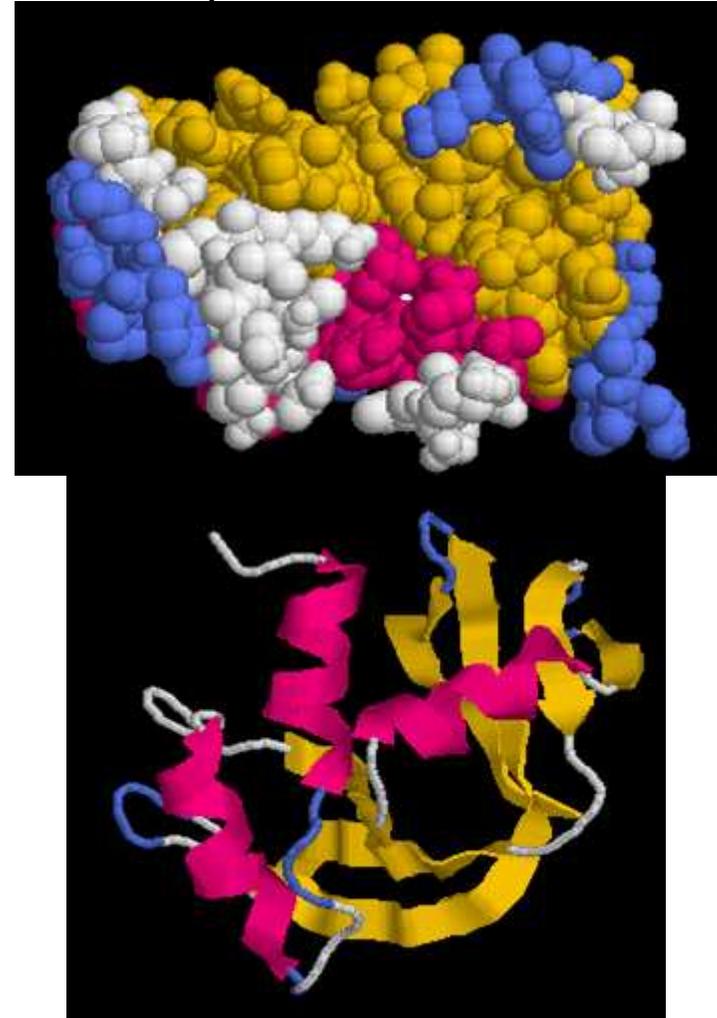
(c) Phosphorylase (Domain 2)

Structures III^{re} : du "2D" au 3D

Protéines solubles globulaires : Association hélices α , feuillets β , boucles, régions non structurées pour former une protéine fonctionnelle.



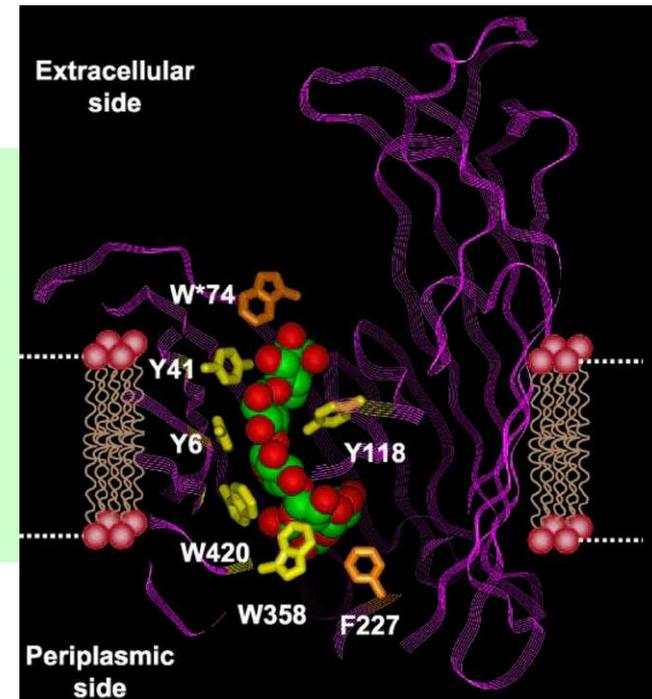
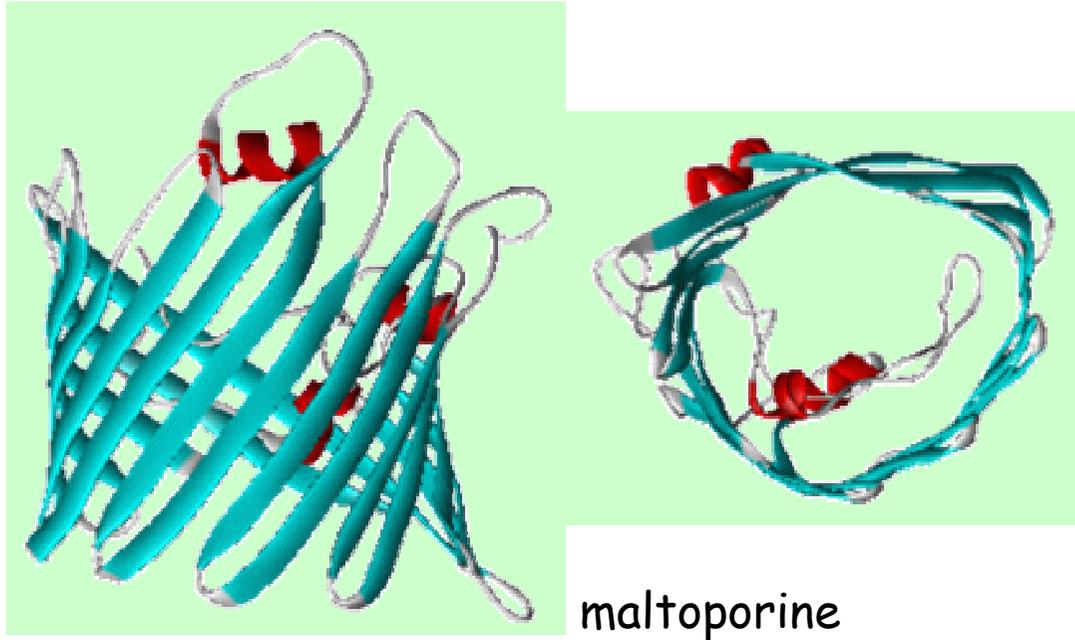
Insuline



Ribonucléase pancréatique

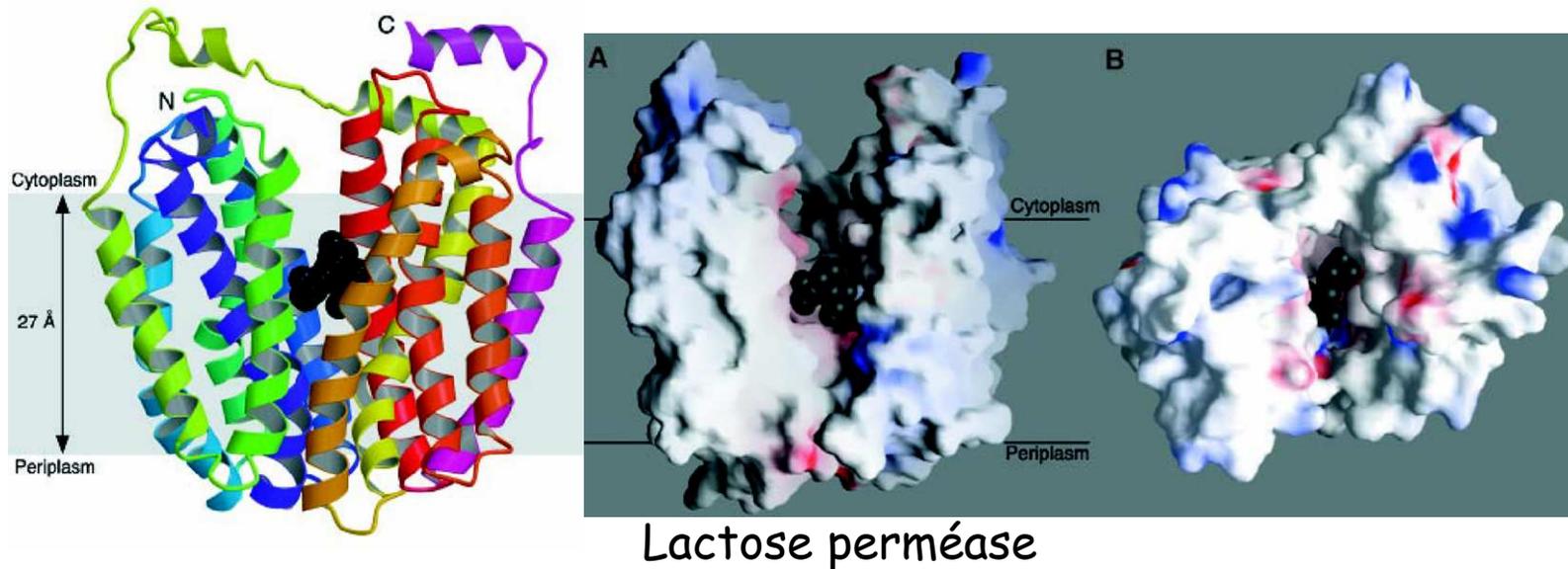
!! RAPPEL !! Structures tertiaires : du "2D" au 3D

Protéines membranaires : ancrées dans la bicouche lipidique, présence de domaines hydrophobes en contact avec les lipides, feuillets β , domaines extracellulaires et intracellulaires



!! RAPPEL !! Structures tertiaires : du "2D" au 3D

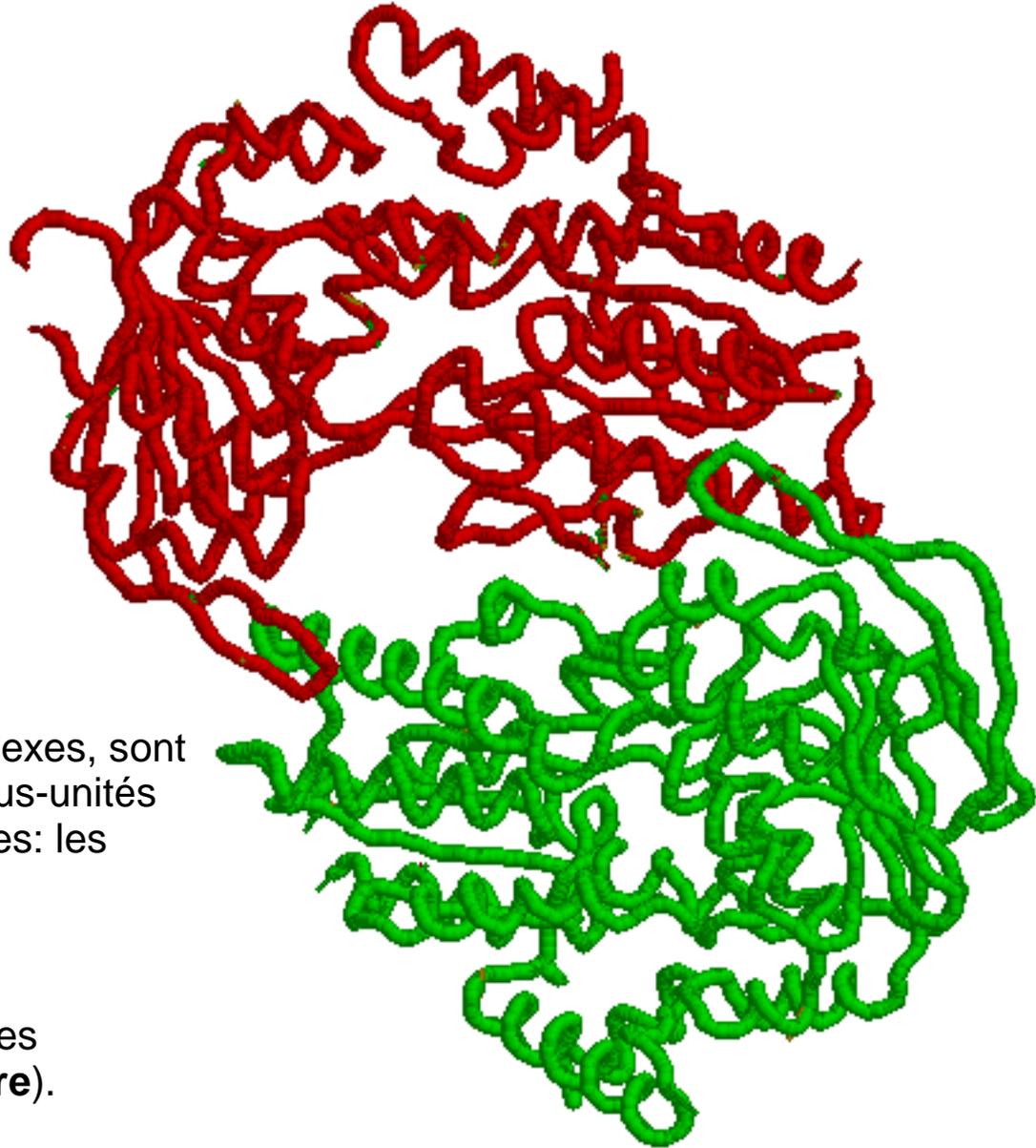
Protéines membranaires : ancrées dans la bicouche lipidique, présence de domaines hydrophobes en contact avec les lipides, hélices α , faisceaux d'hélices α , domaines extracellulaires et intracellulaires



!! RAPPEL !! Rappels: Structure des protéines

La structure quaternaire

- Certaines protéines, complexes, sont constituées de plusieurs sous-unités (plusieurs chaînes protéiques: les **monomères**).
- L'arrangement spatial de ces différentes unités (**oligomère**).



!! RAPPEL !!

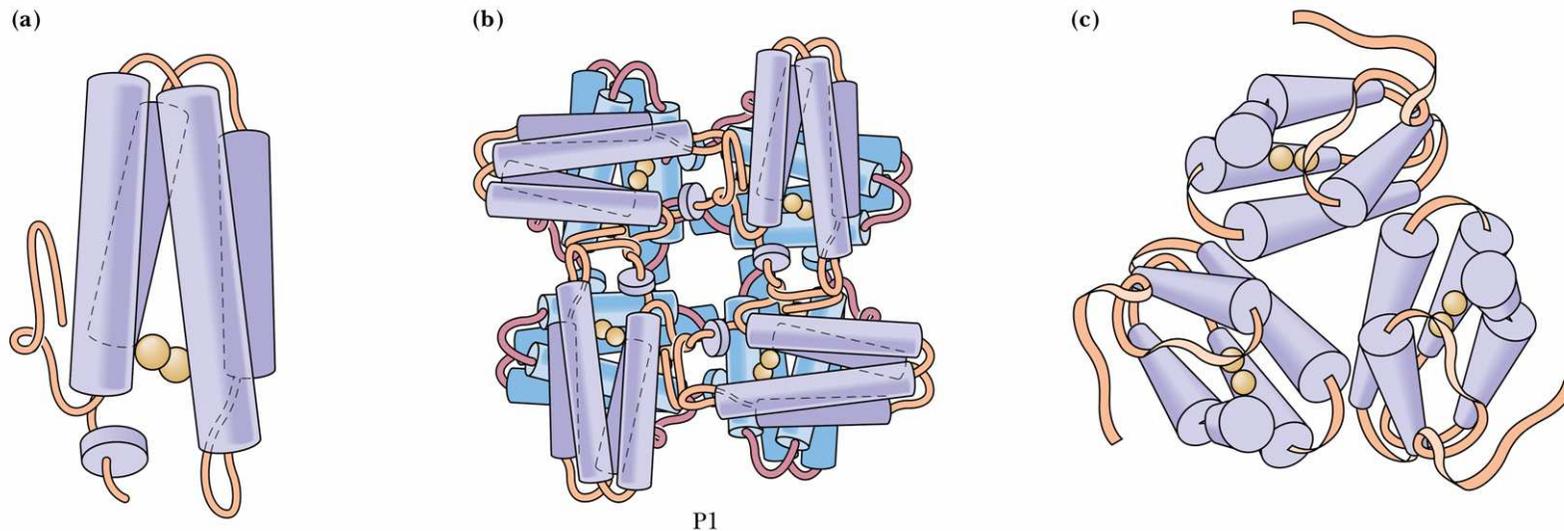
Structure IVre

Etats oligomériques de l'hémérythrine dans divers vers marin

(a) hémérythrine monomérique de *Thermiste zostericola*

(b) hémérythrine octamérique de *Phascolopsis gouldii*

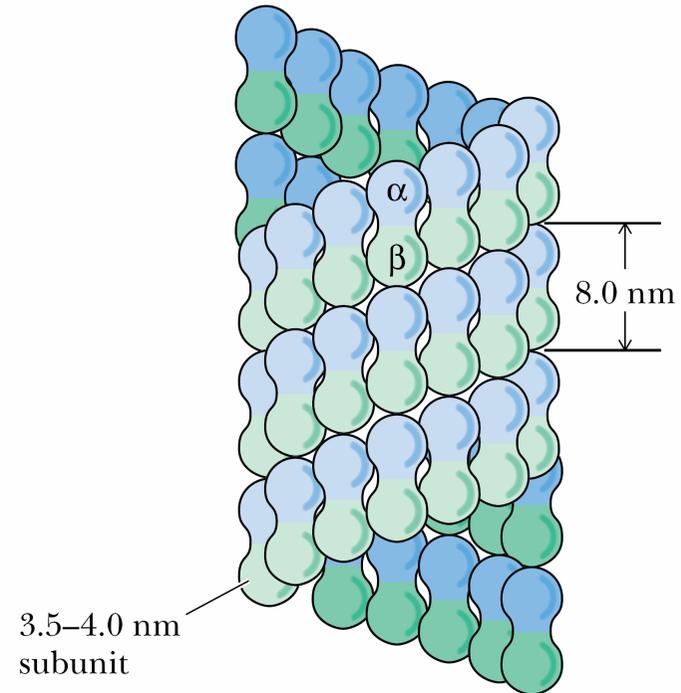
(c) hémérythrine trimérique de *Siphonosoma*.



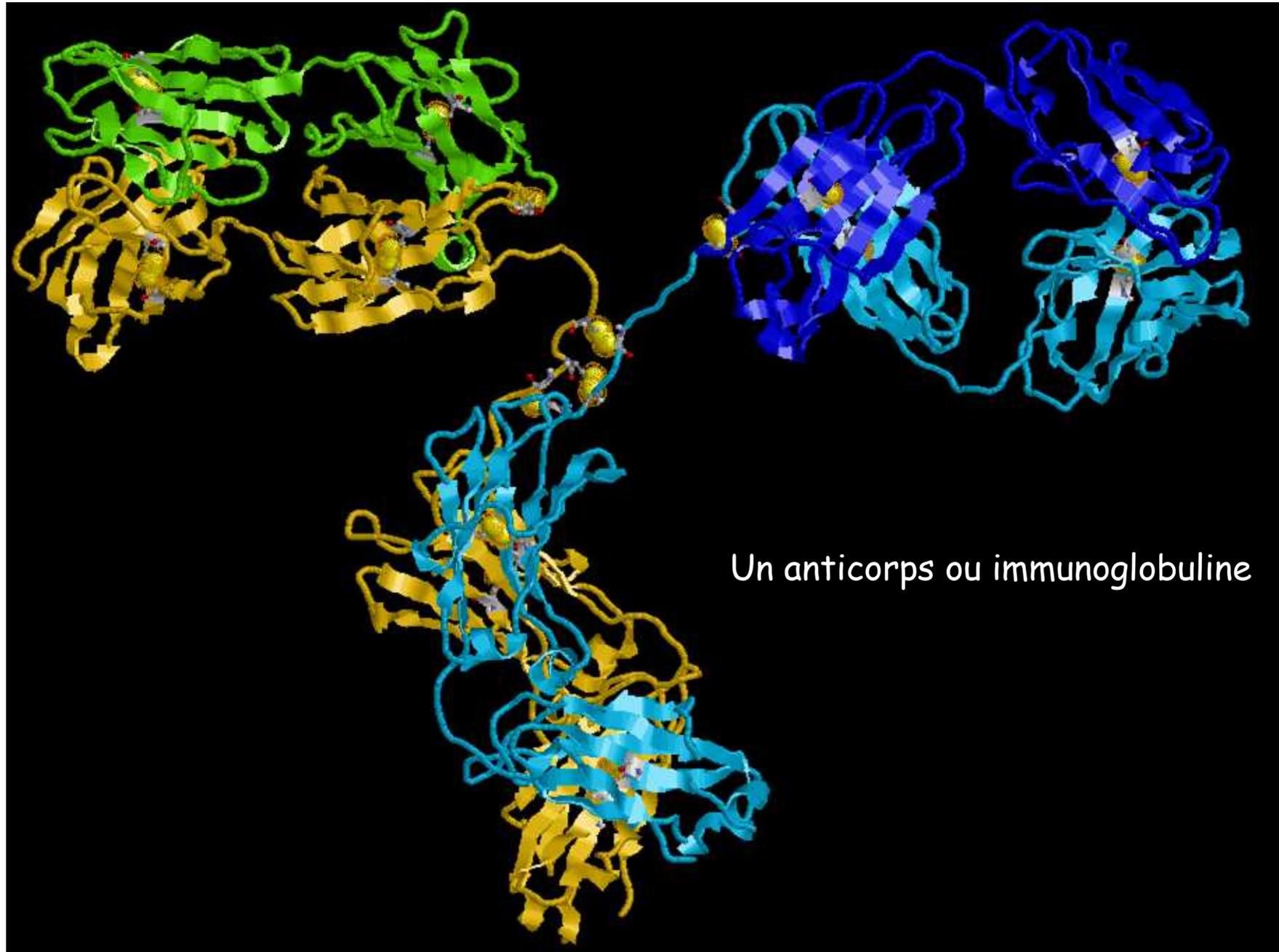
!! RAPPEL !!

Structure IVre

Microtubule (monomères α et β tubuline)

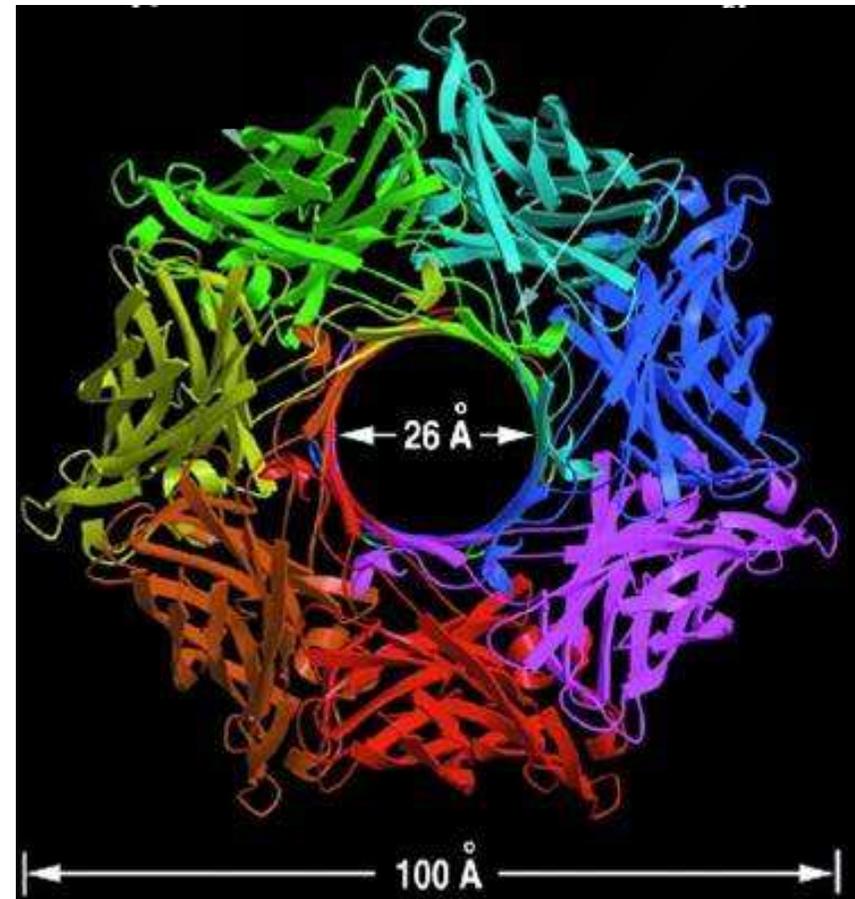
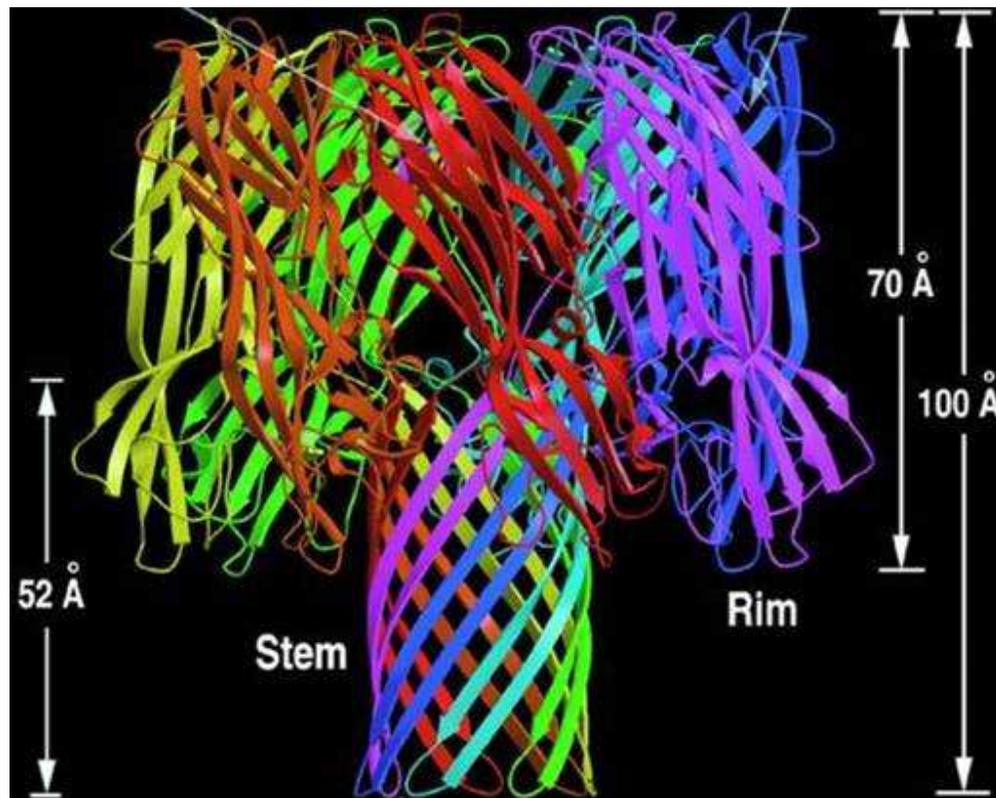


Structure IV^{re}



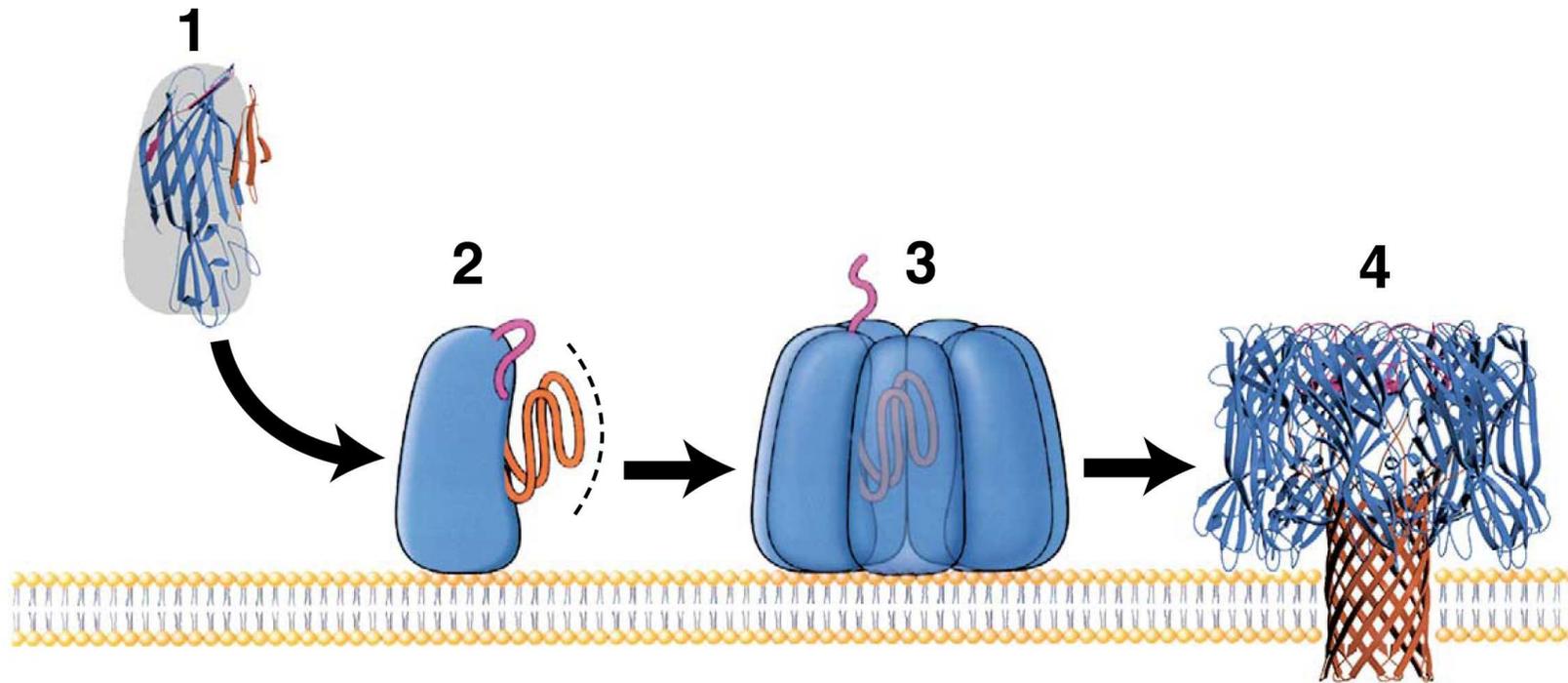
Un anticorps ou immunoglobuline

Structure IV^{re} de protéine membranaire à tonneaux β



L' α -hémolysine

Structure IV^{re} de protéine membranaire à tonneaux β



Structure IV^{re} de protéine membranaire à hélices α

