

Prédiction de la Structure des Protéines

fonction des protéines

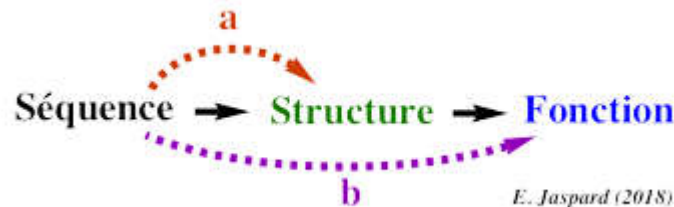
Les protéines peuvent avoir plusieurs fonctions:

- Motricité**
- Catalytique (enzyme)**
- Reconnaissance moléculaire (anticorps)**
- Liaison et fixation**
- Transport (hémoglobine)**
- Structurale (collagène)**

La relation structure - fonction des protéines

b. La relation structure - fonction des protéines

L'enchaînement séquence => structure => fonction traduit qu'une séquence en acides aminés contient l'information nécessaire pour que la chaîne polypeptidique se replie dans la conformation native qui confère à la protéine (à l'enzyme) sa fonction biologique (son activité enzymatique).



La structure des protéines se décompose en 4 parties

La structure primaire.

La structure primaire est la séquence en acide aminé.
Elle est codée par le génome.

La structure secondaire

La chaîne polypeptidique adopte des repliements préférentiels. Hélice ou feuillet.

La structure tertiaire

La chaîne polypeptidique se replie en domaines plus ou moins globulaires

La structure quaternaire

Plusieurs chaînes polypeptidiques (différentes ou non) s'associent.

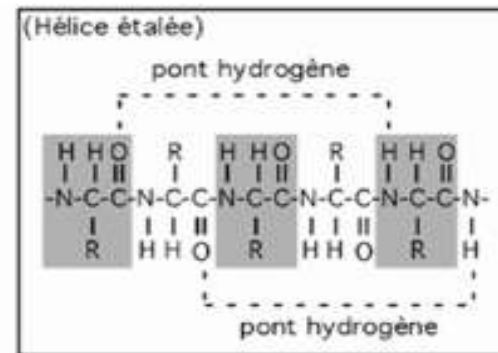
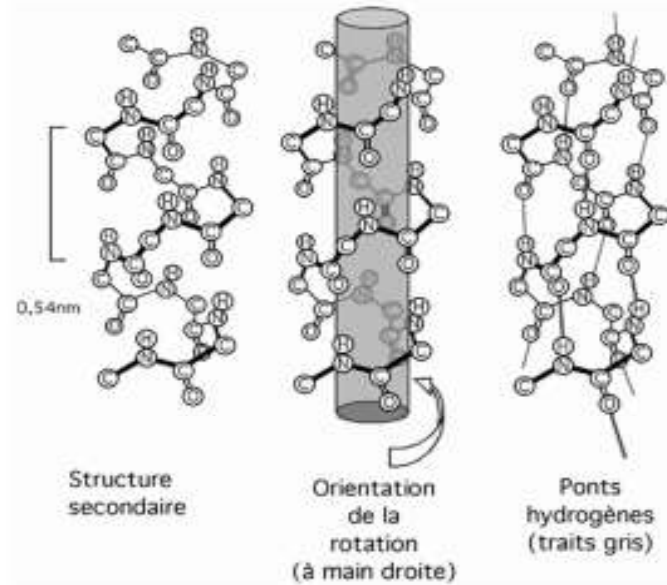
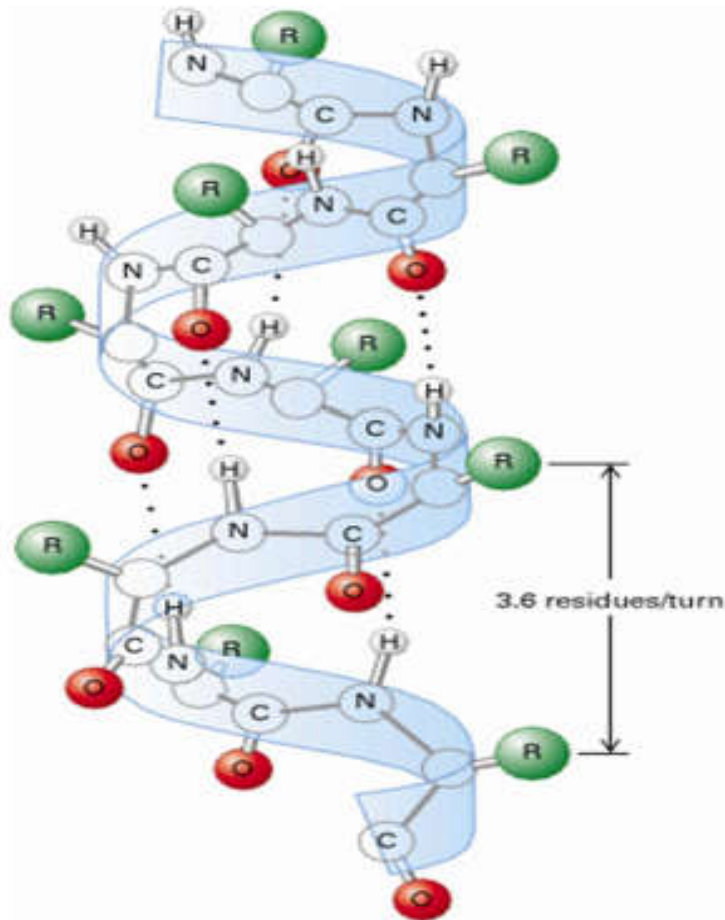
Oligomères, homo-oligomère

Parfois, il faut aussi l'ajout d'autres groupes prosthétiques. (Cofacteurs, métaux,...).

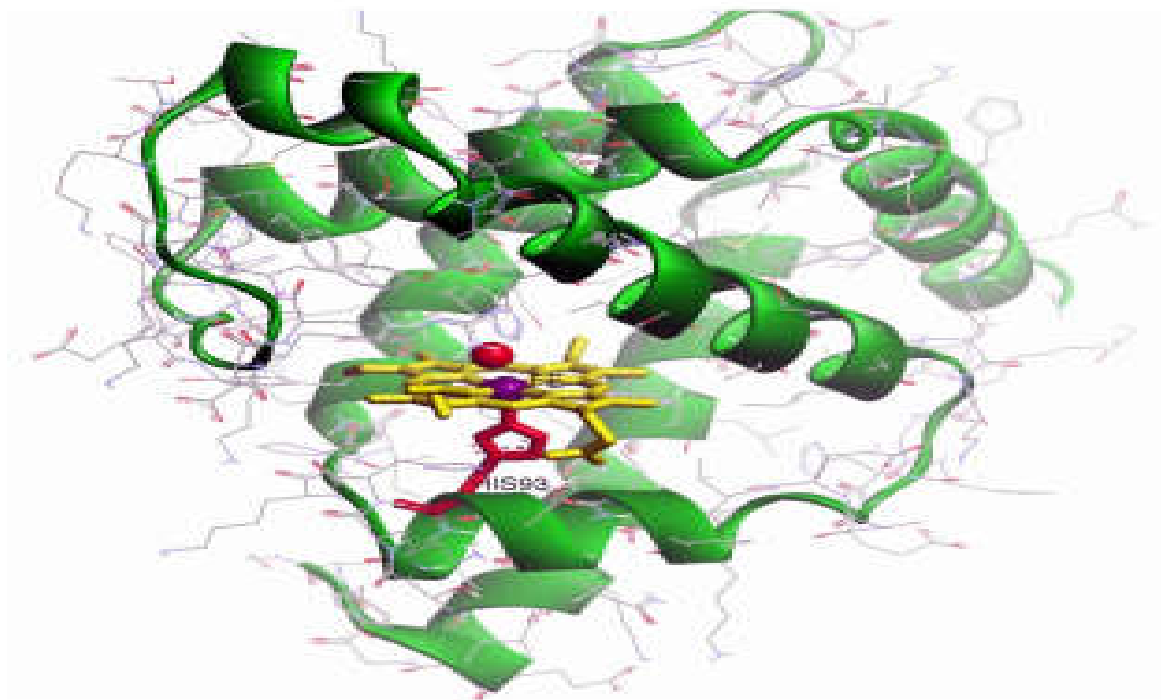
Pour garder l'intégrité de la fonction, il faut que le repliement soit correct.

Structures secondaires

1. L'hélice alpha



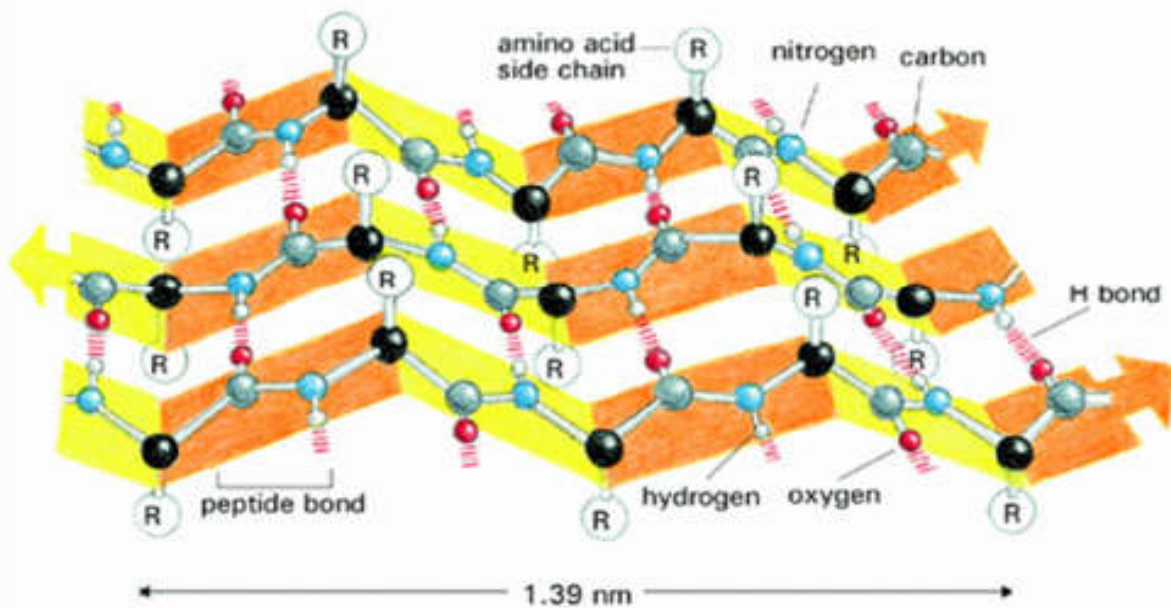
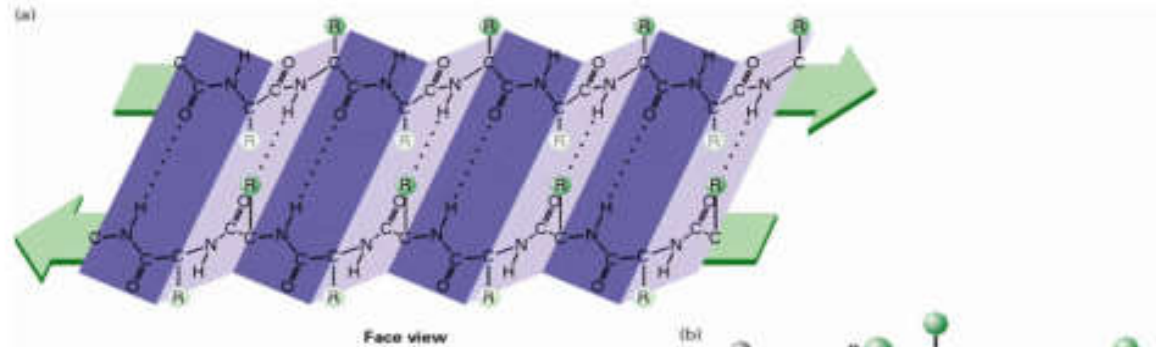
Exemple de protéines riches en hélices alpha



Myoglobine

Structures secondaires

2. Brins bêta et le feuillet bêta



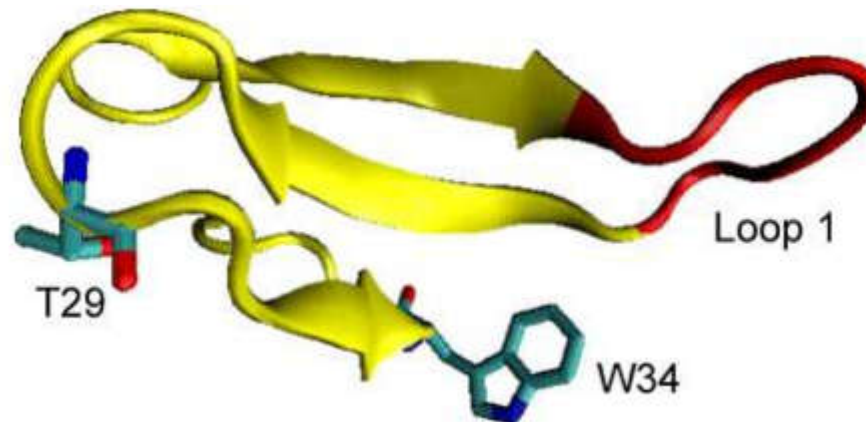
Structures secondaires

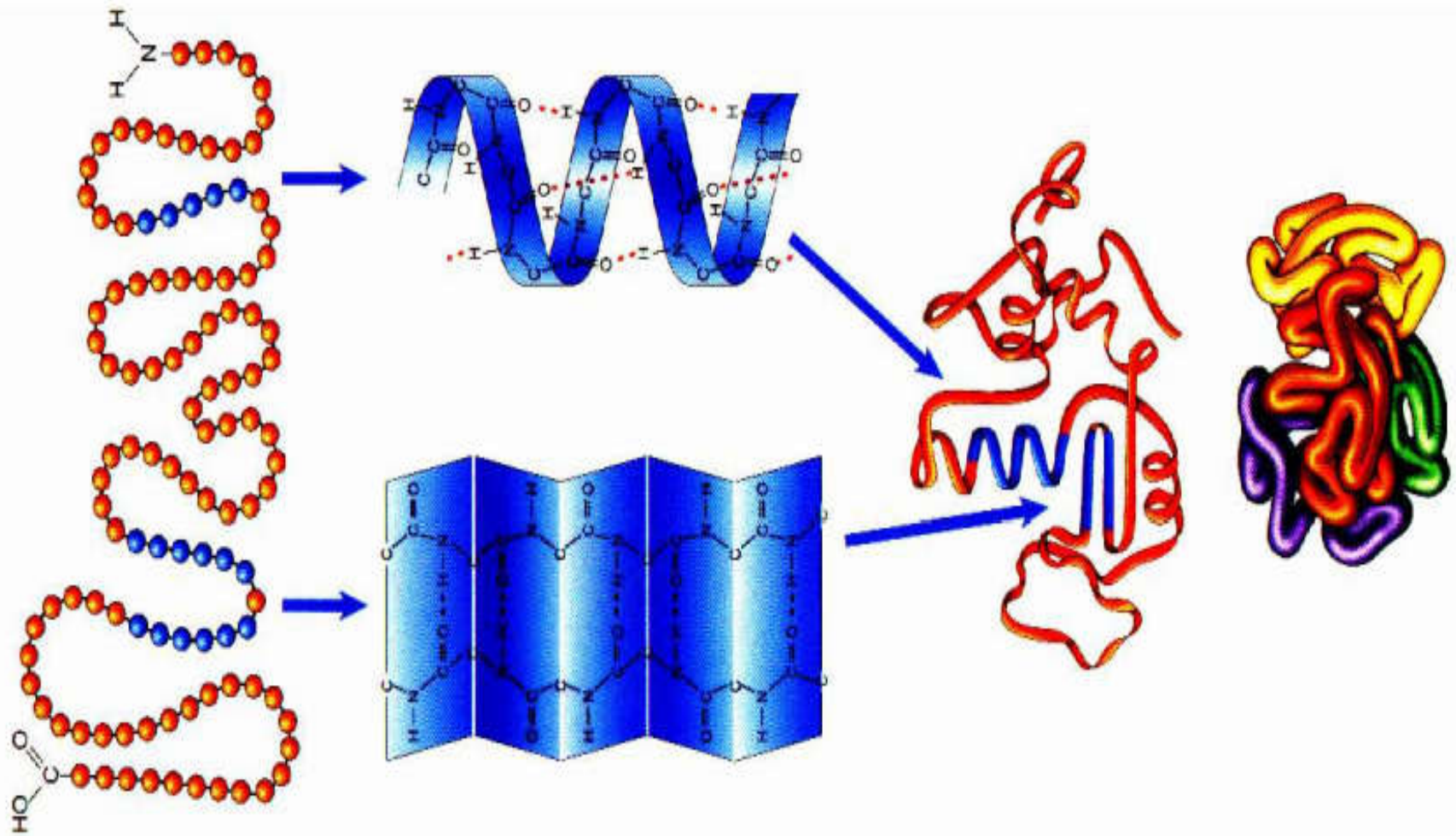
3. Les coudes (Turns)

Relient plusieurs segments successifs de feuillet antiparallèles . très souvent à la surface de la protéine.

Caractéristiques:

- 2 à 4 résidus
- Connectent deux brins β antiparallèles
- Différents types classés selon les angles phi et psi des différents résidus





Structure
primaire

Structure
secondaire

Structure
tertiaire

Structure
quaternaire

Prédiction bioinformatique des structures secondaires

▷ Approche statistique : *Méthode de Chou Fasman*

La méthode de Chou-Fasman est une technique empirique probabiliste pour la prédiction de structures secondaires des protéines. Elle a été développée dans les années 70.

La méthode de Chou-Fasman est basée sur l'analyse des fréquences de chaque acide aminé dans les trois structures secondaires (Hélice alpha, feuille Bêta, et coudes). Ces fréquences sont extraitées depuis les structures de protéine résolues avec Cristallographie et rayon X. Pour chaque acide aminé il y a un ensemble de paramètres calculé à partir de ces fréquences pour chaque structure secondaire.

Prédiction bioinformatique des structures secondaires

▷ Approche statistique : *Méthode de Chou Fasman*

Exploiter la répartition non uniforme des acides aminés dans les structures d'hélice α ou de feuillet β .

- ▷ Pour chaque acide aminé sont déterminés empiriquement des coefficients reflétant la propension de l'acide aminé pour une hélice α , un feuillet β ...
- ▷ $P(h)$ propension à figurer dans une hélice α , $P(f)$ propension à figurer dans un feuillet β

Des tables d'occurrence des acides aminés dans les différents états structuraux ont été établies. Chou et Fasman ont analysé les 29 structures 3D connues en 1974 et ils ont affecté chaque acide aminé à un état structural : hélice, brin, apériodique ou coude.

Acide aminé	Nombre	hélice	feuille	apériodique	coude
Arg	142	53	26	63	40
Cys	94	25	22	47	33
Gln	162	68	35	59	47
Tyr	184	48	53	83	62
Val	357	144	119	94	53
Total	4741	1798	930	2013	1400

- Sur la base de cette table, Chou et Fasman ont établi la table des paramètres conformationnels $P(\alpha)$, $P(\beta)$ et $P(\text{coude})$ en calculant les rapports des fréquences.

Exemple pour Tyr : $P(\beta) (\text{Tyr}) = [(53 / 184) / (930 / 4741)] = 147$.

Nom	$P(\alpha)$	$P(\beta)$	$P(\text{turn})$
Alanine	142	83	66
Arginine	98	93	95
Aspartic Acid	101	54	146
Asparagine	67	89	156
Cysteine	70	119	119
Glutamic Acid	151	037	74
Glutamine	111	110	98
Glycine	57	75	156
Histidine	100	87	95
Isoleucine	108	160	47
Leucine	121	130	59
Lysine	114	74	101
Methionine	145	105	60
Phenylalanine	113	138	60
Proline	57	55	152
Serine	77	75	143
Threonine	83	119	96
Tryptophan	108	137	96
Tyrosine	69	147	114
Valine	106	170	50

Algorithme de Chou-Fasman

Règles pour une hélice α

- ▷ Identifier les régions de 6 résidus consécutifs pour lesquelles 2/3 des résidus satisfont $P(h) > 100$
- ▷ étendre ces régions dans les deux directions, jusqu'à atteindre 4 résidus consécutifs de $P(h)$ moyen < 100
- ▷ les régions pour lesquelles $P(h) > P(f)$ sont déclarées hélice α .

Règles pour un feuillet β

- ▷ Identifier les régions de 5 résidus consécutifs pour lesquelles trois résidus au moins satisfont $P(f) > 100$
- ▷ étendre ces régions
- ▷ les régions pour lesquelles $P(f) > 105$ et $P(f) > P(h)$ sont déclarées feuillets β .