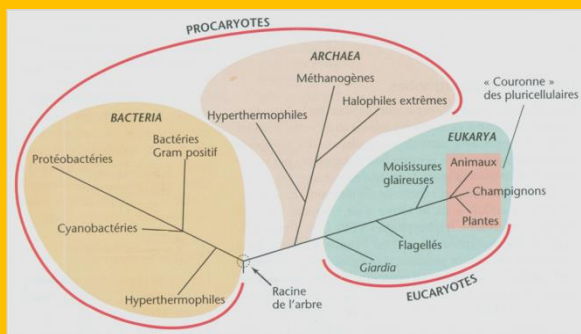
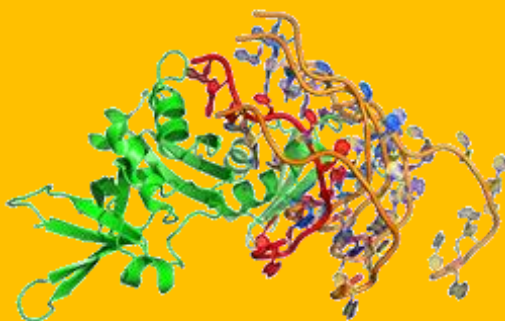


Université Frères Mentouri Constantine 1  
Institut de la Nutrition, de l'Alimentation et des Technologies Agro-alimentaires (INATAA)  
1<sup>e</sup> année Master Biotechnologie alimentaire



# COURS DE BIOINFORMATIQUE

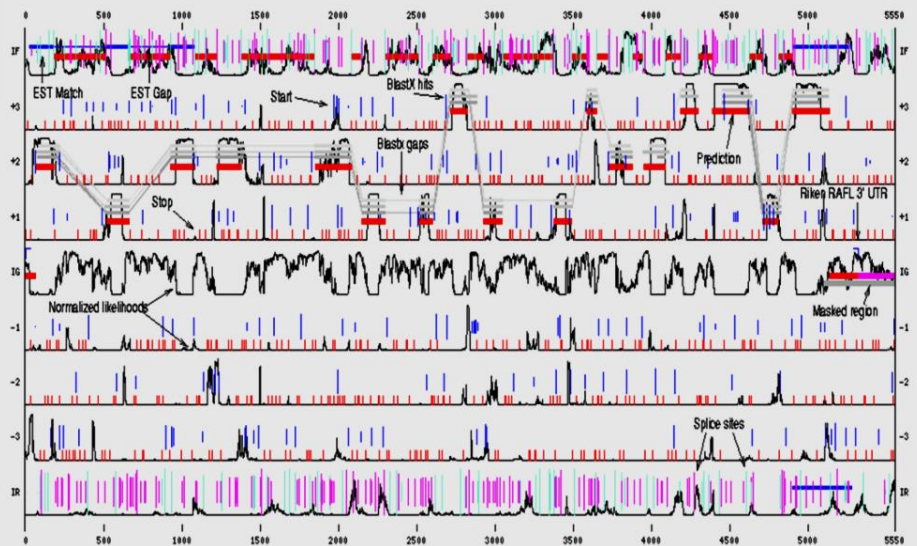
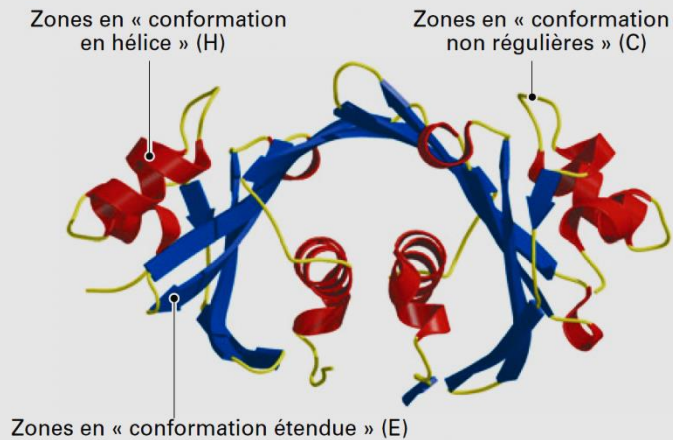
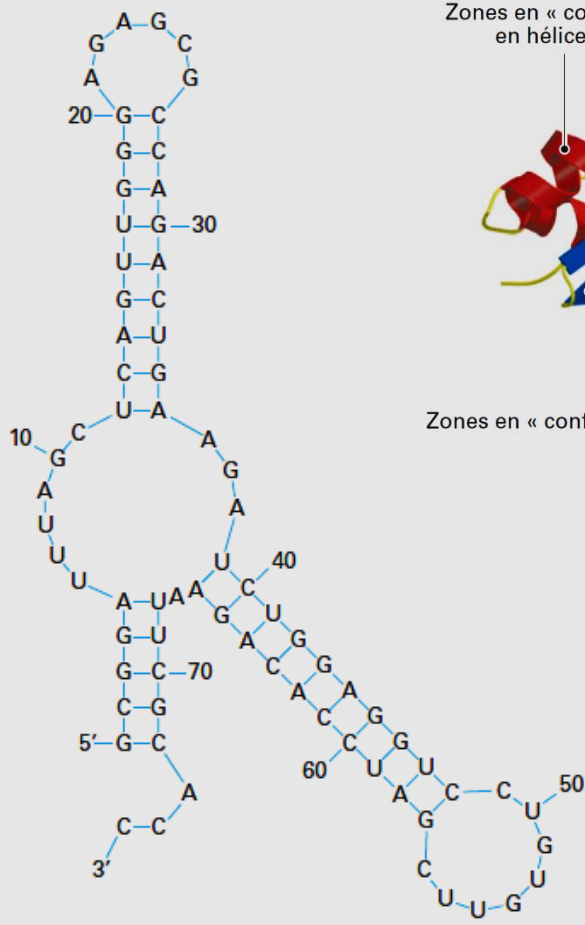


# CHAPITRE IV

## ANNOTATION DES SÉQUENCES

>ARN

```
GCGGAUUUAGCUCAGUUGGAGAGCGCCAGACUGAAGAUCUGGA  
GGUCCUGUGUUCGAUCCACAGAAUUCGCACCA
```



**L'annotation est la recherche d'informations  
(position, structure, fonction) sur une  
séquence (nucléique, protéique) inconnue**

## ANNOTATION

## DÉFINITION

✓ **Annotation structurale (Where)**

Permet de positionner les gènes et leurs produits :

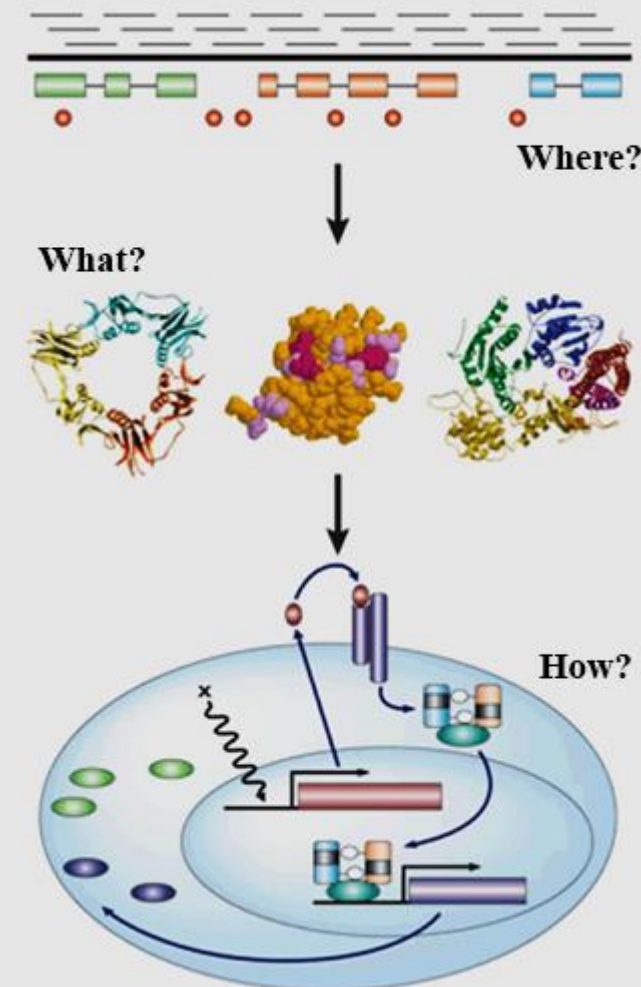
- ☐ éléments répétés (éléments transposables, satellites, etc.)
- ☐ gènes des ARN stables (ARNr, ARNt...)
- ☐ gènes protéiques
- ☐ régions régulatrices, etc.

✓ **Annotation fonctionnelle (What)**

Permet de prédire les fonctions et produits potentiels des gènes préalablement identifiés, leurs structures 2D et 3D (ARN et protéines) ainsi que leurs positions cellulaires (pour les protéines).

✓ **Annotation relationnelle (How)**

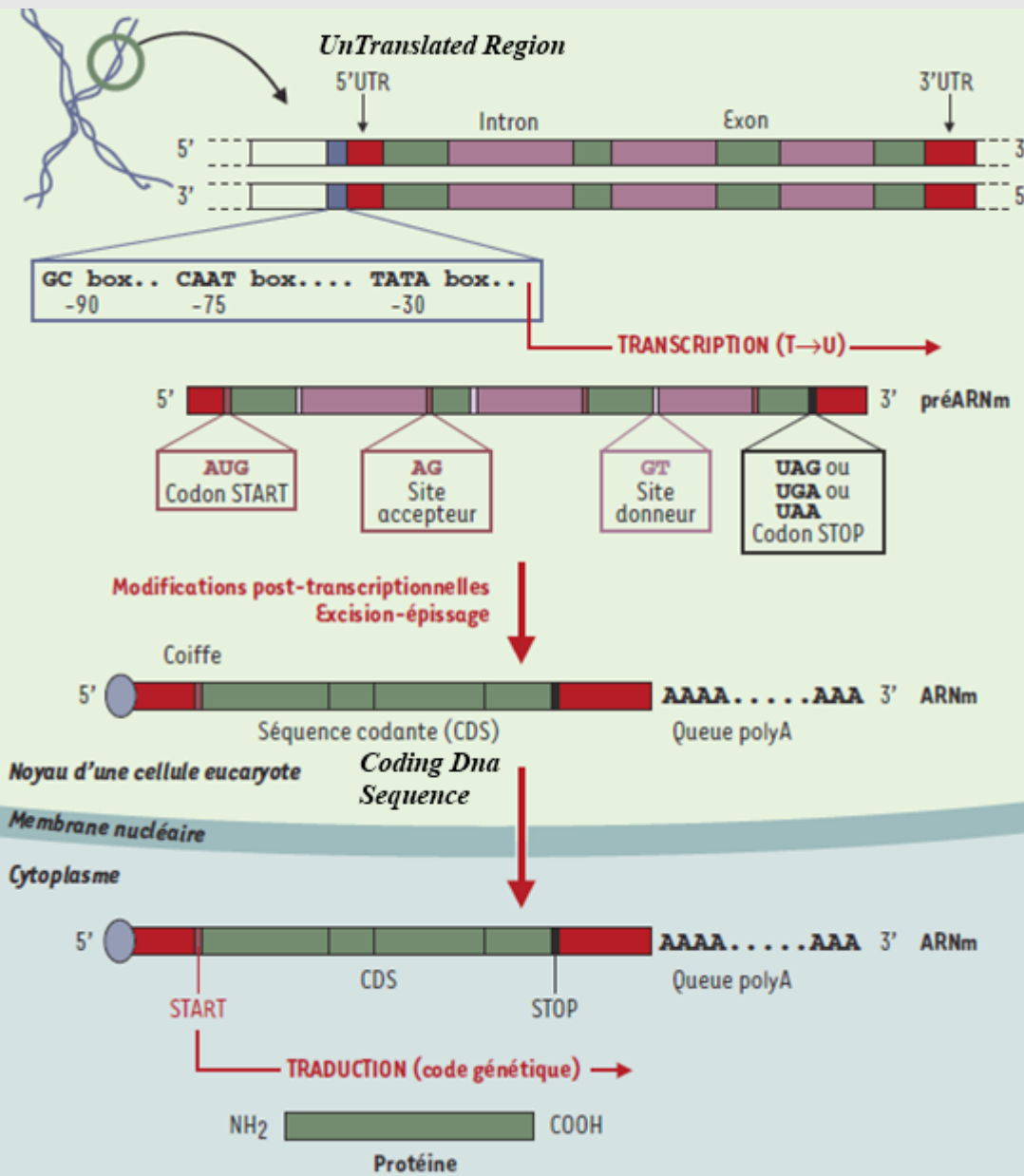
Permet de déterminer les interactions que les objets biologiques préalablement identifiés sont susceptibles d'entretenir (familles de gènes, réseaux de régulation, réseaux métaboliques, etc.).





## ANNOTATION

## PRINCIPE

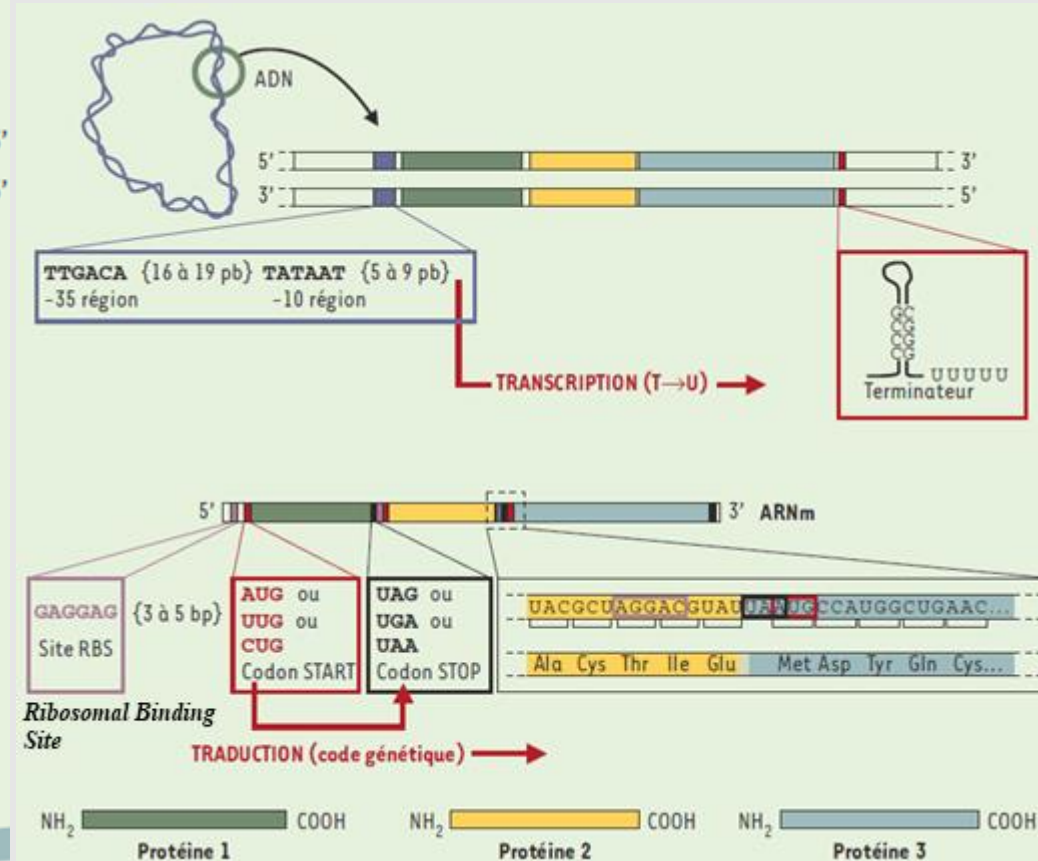


EUCARYOTES

VS

PROCARYOTES

3



L'annotation est basée essentiellement sur la reconnaissance des différentes structures d'un gène (motifs, *ORF* "Open Reading Frame, etc.).

## 1. MÉTHODE EXPÉRIMENTALE

Par alignement de la séquence de l'ADN génomique ( $ADN_g$ ) avec la séquence de l'ADN<sub>c</sub> complet isolé du même organisme ;

## 2. MÉTHODES INTRINSÈQUES (*Ab initio*)

Par l'emploi de modèles statistiques capables de localiser des séquences codantes et non codantes à partir d'exemples (séquences de gènes modèles) ;

## 3. MÉTHODES EXTRINSÈQUES (COMPARATIVES)

- ✓ Par la recherche des ORF's (*Open Reading Phase*) et des motifs;
- ✓ Par comparaison aux banques de données biologiques (avec les outils BLAST, FASTA) à la recherche des séquences d'ARNm et de protéines qui ressemblent à la séquence étudiée ;

## 4. MÉTHODES INTÉGRATIVES

Par la combinaison des trois dernières approches.

## ANNOTATION

## OUTILS

Exemple de logiciels d'annotation:

- ✓ Pour les séquences nucléiques
  - GeneFinder
  - Artemis
  - Geneious
- ✓ Pour les séquences protéiques
  - InterproScan

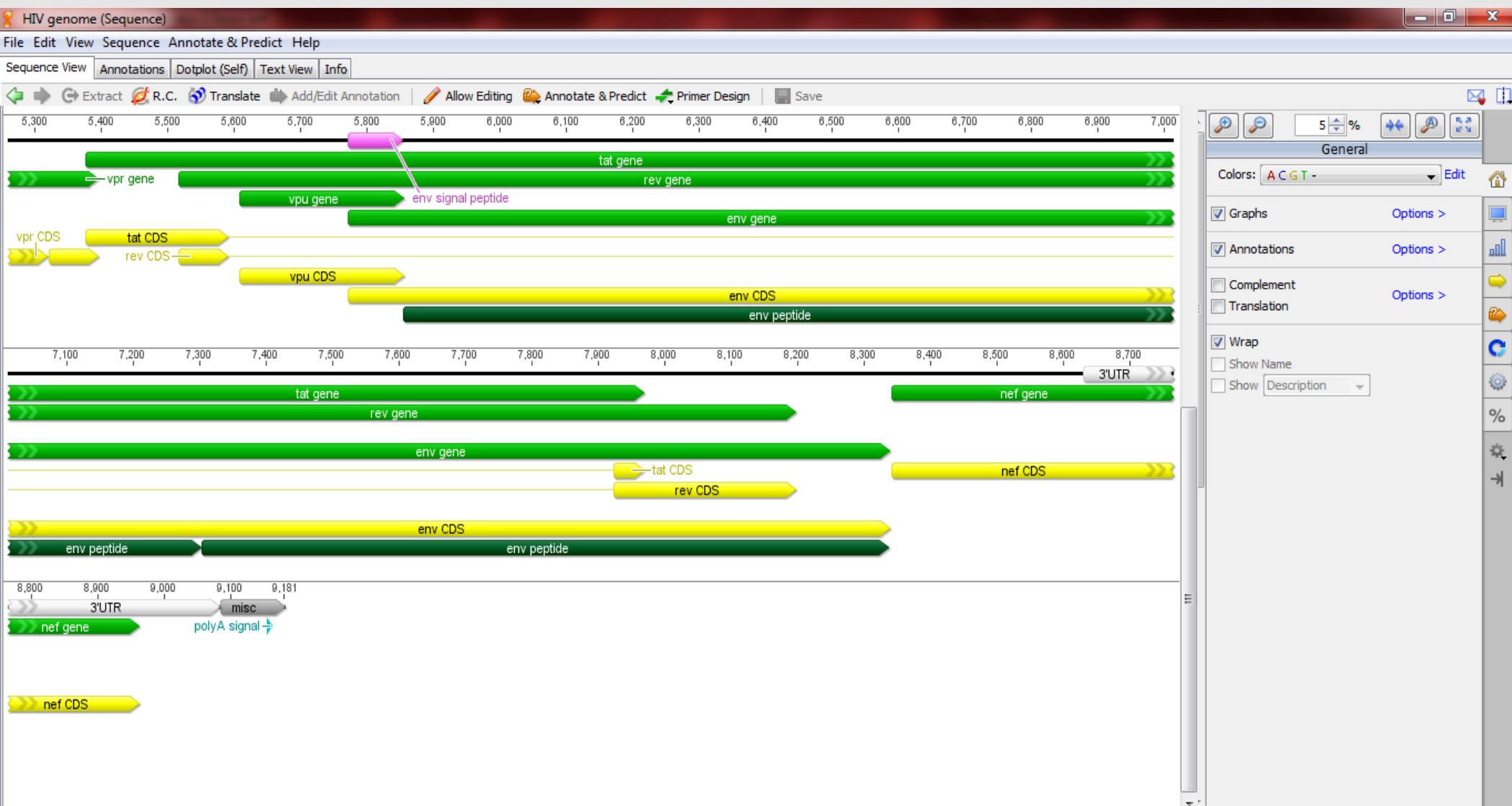
Exemple de banque de séquences annotées:

- ✓ Pour les séquences nucléiques
  - Ensembl
  - Genatlas
- ✓ Pour les séquences protéiques
  - Prosite
  - Pfam
  - Interpro

## ANNOTATION



## OUTILS





## ANNOTATION

# ARTEMIS



## OUTILS



- ✓ La fonction des protéines est une conséquence indirecte de leur séquence en acides aminés, qui passe par le repliement tridimensionnel de cette chaîne en une forme compacte, spécifique et stable ;
- ✓ La prédiction de la structure des protéines (secondaire, tertiaire, quaternaire) est un problème très ardu, essentiellement à cause de 02 problèmes:
  - contrairement aux nucléotides, il n'existe pas de règles de complémentarité strictes entre les acides aminés ;
  - les chaînes formées par les acides aminés sont beaucoup plus flexibles et disposent de beaucoup plus de degrés de liberté que les brins d'acides nucléiques d'où une grande variété de structures.

- ✓ Afin de résoudre ces problèmes, différentes méthodes algorithmiques et statistiques ont été développées, chacune ayant abouti au développement de logiciels de prédiction plus ou moins efficaces:

☐ **Prédiction des structures secondaires:**

- **PSIPRED, SSPPRO, PROF\_King**

☐ **Recherche d'informations:**

- **Ponts disulfure: DIANA, GDAP ;**
- **Homologie: 3DJury, HHPred**
- **Fonction: PFP, ProtFun**

☐ **Prédiction de la structure tertiaire (3D)**

- **SWISS-MODEL (Modélisation par homologie), GenTHREADER**  
(Modélisation par enfilage), HMMStr/Rosetta (Modélisation *Ab initio*)

☐ **Evaluation des modèles: Eval123D, Procheck, Verify3D**