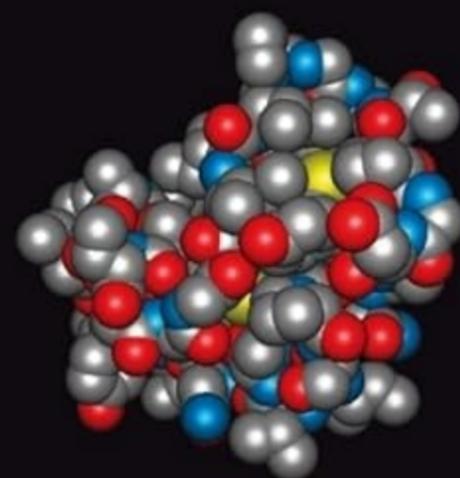
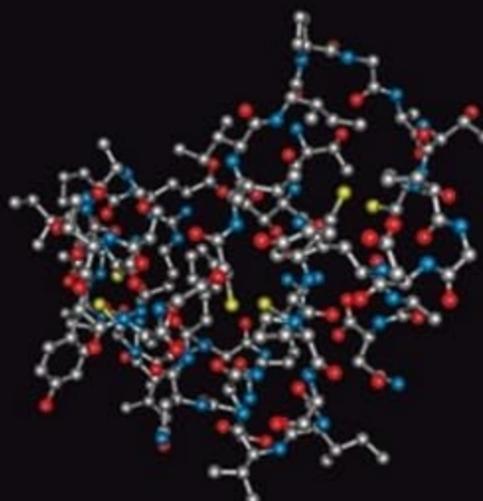
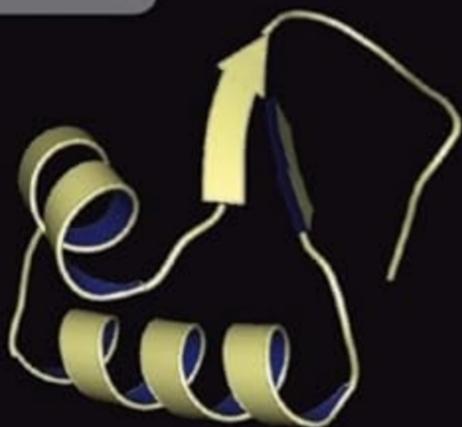
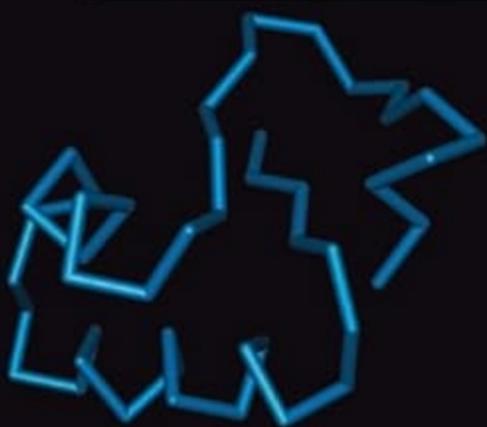


**Gilbert Deléage
Manolo Gouy**



Bioinformatique

Cours et applications

2^e édition

Licence 3
Master
Écoles d'ingénieurs

RESSOURCES



NUMÉRIQUES

DUNOD

TABLE DES MATIÈRES

Comment utiliser cet ouvrage	VI
Avant-propos	IX
Chapitre 1 • La composition en acides aminés	1
1.1 Acides aminés et séquence	1
1.2 Informations déduites de la composition en acides aminés	4
Chapitre 2 • Bases de données pour données de bases	7
2.1 Les banques de données généralistes	7
2.2 Une entrée SWISS-PROT	14
2.3 Les interrogations Entrez, ACNUC, SRS	17
Chapitre 3 • La comparaison de deux séquences	21
3.1 Matrice de points	21
3.2 Matrice de substitution	26
Chapitre 4 • Recherche dans les banques	33
4.1 Score de similitude entre séquences	33
4.2 Recherche globale ou locale	36
4.3 FASTA	37
4.4 BLAST	41
Chapitre 5 • Alignement de séquences	47
5.1 Introduction	47
5.2 Comparaison de protéines homologues (algorithme global)	49
5.3 Meilleur chevauchement entre séquences (algorithme local)	52
5.4 Alignements multiples	54
5.5 Représentation « logo »	57
Chapitre 6 • Bases théoriques de la phylogénie moléculaire	59
6.1 Arbres phylogénétiques	59
6.1.1 Arbres racinés et arbres non racinés	61
6.1.2 Le format Newick d'arbres phylogénétiques	62
6.2 Arbre des espèces - arbres de gènes	63
6.2.1 Nombre d'arbres binaires possibles	64
6.3 Modèle markovien de l'évolution moléculaire	65

Table des matières

6.3.1	Matrice de transition	67
6.3.2	Quelques modèles nucléotidiques de Markov	68
6.3.3	Longueur d'une branche	70
6.3.4	Modélisation de la variation des taux d'évolution entre sites	70
6.4	Choix des sites	72
6.5	Matrices de taux de substitution entre séquences protéiques	73
6.6	Distances évolutives entre paires de séquences	74
Chapitre 7 - Algorithmes pour la phylogénie moléculaire		77
7.1	Parcimonie	78
7.1.1	Algorithme	78
7.1.2	Heuristiques	82
7.1.3	Propriétés	82
7.1.4	Implémentations	83
7.1.5	Longueurs de branches des arbres de parcimonie	83
7.1.6	Traitement des indels	84
7.2	Méthodes de distances	85
7.2.1	Méthode d'évolution minimale	86
7.2.2	Méthode Neighbor-Joining	86
7.3	Maximum de vraisemblance	89
7.3.1	Modèle probabiliste utilisé au maximum de vraisemblance	90
7.3.2	Calcul de la vraisemblance	90
7.3.3	L'algorithme de Felsenstein	91
7.3.4	Prise en compte de la variabilité des vitesses d'évolution entre sites	92
7.3.5	Optimisation de la vraisemblance	93
7.3.6	Implémentation	94
7.4	Estimation de la fiabilité d'un arbre par bootstrap	94
7.5	Choix des méthodes de calcul d'arbres	97
Chapitre 8 - Recherche de fonctions		99
8.1	Définitions	99
8.2	Détection de signatures de séquence (PROSITE)	100
8.3	Recherche de fonction avec pondération par la fréquence	103
8.4	Méthodes à base de profils	106
Chapitre 9 - Profils physico-chimiques		111
9.1	Pourquoi les profils physico-chimiques ?	111
9.2	Hydrophobie-paramètres-construction du profil - interprétation	111
9.3	Amphiphilie	114
9.4	Accessibilité au solvant	115
Chapitre 10 - Prédictions de structures secondaires		117
10.1	Méthode « statistique empirique »	120
10.2	Méthode information directionnelle (GOR)	123
10.3	Méthode de recherche des plus proches voisins (NNM)	127

10.4 Méthode auto-optimisée (SOPM)	130
10.5 Méthode auto-optimisée avec alignements (SOPMA)	131
10.6 Méthodes neuronales	132
10.7 Autres méthodes	134
10.7.1 Méthode statistique discriminante (DSC)	134
10.7.2 Méthode neuronale (PREDATOR)	134
10.7.3 Méthode hiérarchisée réseaux de neurones (HNN)	134
10.7.4 Méthodes utilisant les chaînes de Markov	135
10.7.5 Combinaison de méthodes	135
10.8 Critères de qualité prédictive	137
Chapitre 11 - Prédiction de structures 3D	139
11.1 Principe des méthodes de détermination expérimentale	139
11.2 Le format PDB	140
11.3 Les différents modes de représentations	142
11.4 Classification de structures 3D	145
11.5 Comparaison de structures 3D	146
11.6 Énergétique moléculaire	148
11.7 Optimisation de structures 3D	151
11.8 Modélisation de structures 3D	152
11.8.1 Les méthodes d'enfilage des repliements (<i>threading</i>)	153
11.8.2 Modélisation par homologie	154
11.8.3 Les alphabets structuraux	163
11.8.4 Les méthodes de novo	166
Chapitre 12 - Détection de sites 3D dans les protéines	169
12.1 Problématique	169
12.2 Méthode SuMO	170
Cas pratique d'analyse de séquences	173
Cas pratique de modélisation moléculaire de protéine par homologie	183
Conclusion	191
Bibliographie	193
Glossaire	199
Index	201