

IFT 3290 hiver 2004

Bioinformatique

Miklós Csűrös

André-Aisenstadt 3253
csuros@iro.umontreal.ca

[http ://www.iro.umontreal.ca/~csuros/IFT3290/](http://www.iro.umontreal.ca/~csuros/IFT3290/)

Préalables :

IFT2010 — Programmation 2

IFT1978 — Probabilités et statistique

Horaire :

cours	mercredi 11 :30–13 :30	AA 1207
cours	vendredi 14 :30–15 :30	AA 1177
démo	vendredi 15 :30-17 :30	AA 1360

Démonstrateur : on verra. . .

Évaluation : 30% TP, 30% intra, 40% final.

Seuil : au moins 40% au meilleur des deux exams.

4 travaux

Diversité et universalité

Diversité de la vie

Universalité au niveau moléculaire : interaction de macromolécules

Protéines et acides nucléiques

Protéines

Fonctions :

- structure de la cellule [p.e collagène form des fibres en tissu conjonctif]
- enzymes : catalysateur de réactions spécifiques
- protéines membranaires
- signalisation [p.e. EGF — facteur de croissance pour l'épiderme]
- ...

Exemple : kinesine

(moteur moléculaire)

©UIUC Theoretical and Computational Biophysics Group

Info sur une protéine

facteur de croissance EGF : voie de signalisation (www)

recherche dans une base de donnée : SWISS-PROT (www)

(Tapez «EGF receptor» au site Web de SWISS-PROT et suivez le lien vers EGFR_HUMAN.)

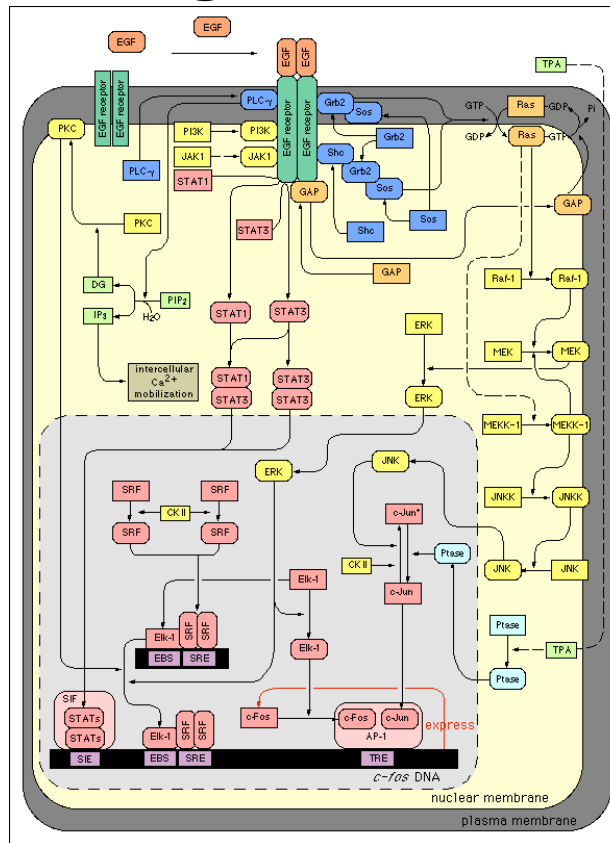
pleine d'information liée : e.g., structure 3D (www)

(Suivez le lien «PDB» dans la ligne «HSSP».)

Séquence

MRPSGTAGAA LLALLAALCP ASRALEEKKV CQGTSNKLTQ
LGTFEDHFLS LQRMFNCEV VLG NLEITYV QRNYDLSFLK
...

Voie de signalisation EGF



SPAD, Kyushu University

IFT 3290, H2004, Miklós Csűrös
Université de Montréal

8

View of SwissProt: P00533

View of SWISS-PROT: A00001

View of UniProtKB: P00533

View of NCBI: NM_000533

View of EMBL: U00001

View of GenBank: F00001

View of RefSeq: NM_000533

View of Ensembl: ENSG00000100001

View of UCSC: uc000533

View of BioGRID: B000533

View of STRING: S000533

View of KEGG: K000533

View of TrEMBL: T000533

View of PDB: 1PXX

View of Pfam: PF000533

View of InterPro: IPR000533

View of SMART: SM000533

View of PROSITE: PS000533

View of BLOCKS: BK000533

View of ProSite: PS000533

View of ProSitePlus: PP000533

View of PRESCAN: PR000533

View of DIP: DP000533

View of ModBase: MB000533

View of Keywords: KW000533

View of Transmembrane: TM000533

View of Glycosylation: GLY000533

View of Repeat: RP000533

View of Signal: SL000533

View of Transferase: TR000533

View of Tyrosine-protein kinase: TYR000533

View of ATP-binding: AB000533

View of Phosphorylation: PH000533

View of Alternative splicing: AS000533

View of Features: FE000533

View of Key: KY000533

View of From: FR000533

View of To: TO000533

View of Length: LG000533

View of Description: DS000533

View of Entry: EN000533

View of Title: TI000533

View of Accession: AC000533

View of Date: DT000533

View of Version: V000533

View of Status: ST000533

View of Type: TP000533

View of Category: CT000533

View of Subcategory: SC000533

View of Species: SP000533

View of Organism: OR000533

View of Strain: SR000533

View of Sex: SX000533

View of Age: AG000533

View of Disease: DI000533

View of Treatment: TR000533

View of Outcome: OC000533

View of Follow-up: FU000533

View of Publication: PU000533

View of Citation: CI000533

View of Reference: RE000533

View of Comment: CM000533

View of Note: NT000533

View of History: HY000533

View of Change: CH000533

View of Revision: RV000533

View of Release: RL000533

View of Update: UP000533

View of Patch: PT000533

View of Hotfix: HF000533

View of Erratum: ER000533

View of Withdrawal: WD000533

View of Deletion: DL000533

View of Duplication: DU000533

View of Inversion: IV000533

View of Translocation: TL000533

View of Amplification: AM000533

View of Loss: LS000533

View of Gain: GN000533

View of Rearrangement: RR000533

View of Fusion: FS000533

View of Split: SP000533

View of Join: JO000533

View of Break: BR000533

View of Connect: CN000533

View of Separate: SE000533

View of Merge: ME000533

View of Divide: DV000533

View of Multiply: MU000533

View of Subtract: SU000533

View of Add: AD000533

View of Remove: RM000533

View of Insert: IN000533

View of Delete: DE000533

View of Replace: RP000533

View of Move: MV000533

View of Copy: CP000533

View of Paste: PT000533

View of Undo: UN000533

View of Redo: RD000533

View of Cut: CU000533

View of Copy: CO000533

View of Paste: PA000533

View of Find: FI000533

View of Replace: RF000533

View of Select: SL000533

View of Deselect: DSL000533

View of Toggle: TG000533

View of Hide: HD000533

View of Show: SH000533

View of Zoom: ZM000533

View of Pan: PN000533

View of Scroll: SC000533

View of Drag: DR000533

View of Drop: DP000533

View of Click: CL000533

View of Double-click: DC000533

View of Right-click: RC000533

View of Middle-click: MC000533

View of Wheel-click: WC000533

View of Touch: TU000533

View of Pinch: PI000533

View of Spread: SP000533

View of Rotate: RO000533

View of Translate: TR000533

View of Scale: SC000533

View of Stretch: ST000533

View of Shrink: SR000533

View of Expand: EX000533

View of Collapse: CP000533

View of ExpandAll: EP000533

View of CollapseAll: CA000533

View of Refresh: RF000533

View of Reload: RL000533

View of Reset: RS000533

View of Restore: RT000533

View of Save: SV000533

View of Load: LD000533

View of Print: PR000533

View of Export: EX000533

View of Import: IM000533

View of Open: OP000533

View of Close: CL000533

View of Save As: SA000533

View of Open Recent: OR000533

View of Recent List: RL000533

View of Favorites: FA000533

View of Bookmarks: BM000533

View of History: HY000533

View of Home: HM000533

View of Stop: ST000533

View of Pause: PA000533

View of Play: PL000533

View of Full Screen: FS000533

View of Exit: EX000533

View of Help: HP000533

View of About: AB000533

View of License: LI000533

View of Privacy Policy: PP000533

View of Terms of Service: TS000533

View of Contact Us: CU000533

View of Feedback: FB000533

View of Support: SU000533

View of FAQ: FQ000533

View of Glossary: GL000533

View of Index: IX000533

View of Search: SR000533

View of Filter: FL000533

View of Sort: SO000533

View of Group: GR000533

View of Unfold: UF000533

View of Fold: FO000533

View of Expand: EX000533

View of Collapse: CP000533

View of Maximize: MA000533

View of Restore Down: RD000533

View of Window: W000533

View of Viewport: VP000533

View of Canvas: CV000533

View of Layer: LY000533

View of Style: ST000533

View of Class: CL000533

View of ID: ID000533

View of Name: NA000533

View of Value: VA000533

View of Unit: UT000533

View of Format: FM000533

View of Align: AL000533

View of Justify: JU000533

View of Left: LE000533

View of Center: CE000533

View of Right: RI000533

View of Justified: JJ000533

View of Paragraph: PG000533

View of Line: LN000533

View of Bulleted: BU000533

View of Numbered: NU000533

View of Decreased Indent: DI000533

View of Increased Indent: II000533

View of Outdent: OU000533

View of Indent: IN000533

View of Decrease Indent: DCI000533

View of Increase Indent: ICI000533

View of Bold: BO000533

View of Italic: IT000533

View of Underline: UL000533

View of Strikeout: SK000533

View of Superscript: SS000533

View of Subscript: SB000533

View of Text Color: TC000533

View of Background Color: BC000533

View of Font Size: FS000533

View of Font Family: FF000533

View of Font Weight: FW000533

View of Font Style: FT000533

View of Font Variant: FV000533

View of Font Ligatures: FLG000533

View of Font Discretionary Ligs: FDL000533

View of Font Embedding: FE000533

View of Font Subsetting: FSub000533

View of Font Rendering: FR000533

View of Font Smoothing: FSm000533

View of Font Antialiasing: FAA000533

View of Font ClearType: FTC000533

View of Font Core fonts: FC000533

View of Font User fonts: FUF000533

View of Font System fonts: FSY000533

View of Font Application fonts: FAF000533

View of Font Personal fonts: FPE000533

View of Font Public fonts: FPU000533

View of Font Private fonts: FPR000533

View of Font Shared fonts: FSH000533

View of Font Localized fonts: FLF000533

View of Font Embedded fonts: FEF000533

View of Font External fonts: FEX000533

View of Font Remote fonts: FRF000533

View of Font Network fonts: FNF000533

View of Font Internet fonts: FIN000533

View of Font Intranet fonts: FIIN000533

View of Font Extranet fonts: FEIN000533

View of Font Global fonts: FG000533

View of Font All fonts: FAL000533

EGF à PDB

Structure Explorer - 1FGK 1/10/03 9:22 AM

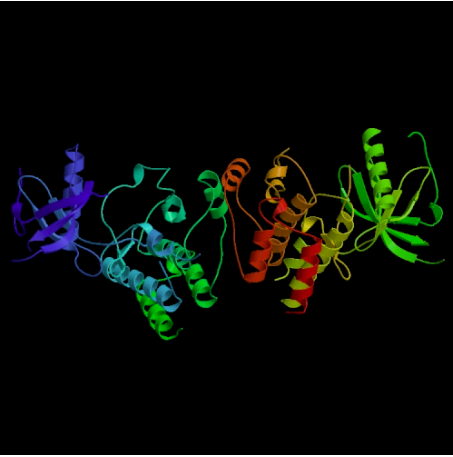
PDB
PROTEIN DATA BANK

Structure Explorer - 1FGK

Title Crystal Structure Of The Tyrosine Kinase Domain Of Fibroblast Growth Factor Receptor 1
Classification Phosphotransferase
Compound Mol. Id: 1; Molecule: Fgf Receptor 1; Chain: A, B; **Fragment:** Tyrosine Kinase Domain, Human Egr1
Residues That Possess Ptk Activity; **Synonym:** Fgfr1K, Fibroblast Growth Factor Receptor 1; E: 2.7.1.112; **Engineered:** Yes; **Mutation:** L457V, C488A, C584S
Exp. Method X-ray Diffraction

[View Structure](#)

[Summary Information](#)
[View Structure](#)
[Download/Display File](#)
[Structural Neighbors](#)
[Geometry](#)
[Other Sources](#)
[Sequence Details](#)
[Structure Factors \(compressed\)](#)



© RSCB

<http://www.rcsb.org/pdb/cgi/structure.cgi?job=graphics&id=1FGK;page=&opt=show&size=500> Page 1 of 1

Séquences protéiques

Protéine : un ou plusieurs polymères d'acides aminés (polymère est une grosse molécule formée de l'union de molécules plus petites)

Alanine (Ala) Glycine (Gly) Méthionine (Met)

Sérine (Ser) Cystéine (Cys) Histidine (His)

Asparagine (Asn) Thréonine (Thr) Acide aspartique (Asp)

Isoleucine (Ilu) Proline (Pro) Valine (Val)

Acide glutamique (Glu) Lysine (Lys) Glutamine (Gln)

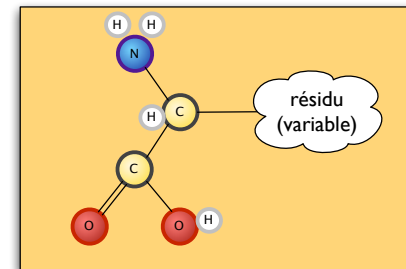
Tryptophane (Trp) Phénylalanine (Phe) Leucine (Leu)

Arginine (Arg) Tyrosine (Tyr)

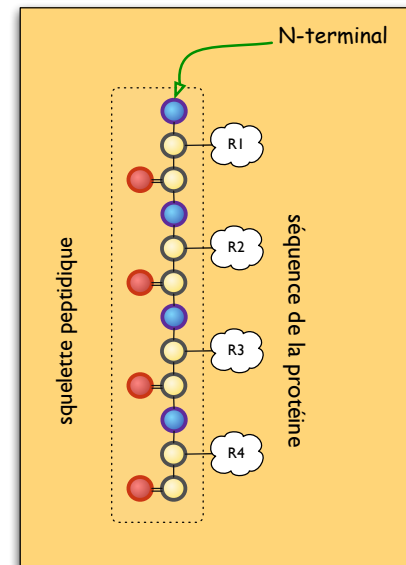
La séquence détermine la structure

⇒ **Problème de prédiction de structure protéique.**

Acides aminés



Acide aminé



Protéine

(Phe, Pro, Met, Gly)

Song Tan, Penn State U

Autres problèmes

Comment est la structure établie ?

⇒ Repliement de protéines

La structure détermine la fonction

⇒ Problème de prédiction de fonction protéique.

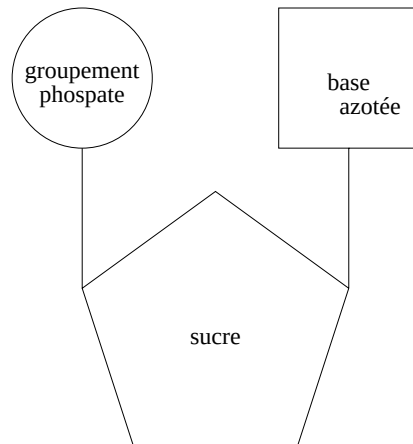
Interaction de protéines

⇒ Analyses de voies métaboliques, chemins de signalisation, réseaux régulatoires

D'où viennent les protéines ?

Encodage en acides nucléiques : ADN et ARN.

ADN : polymère de nucléotides

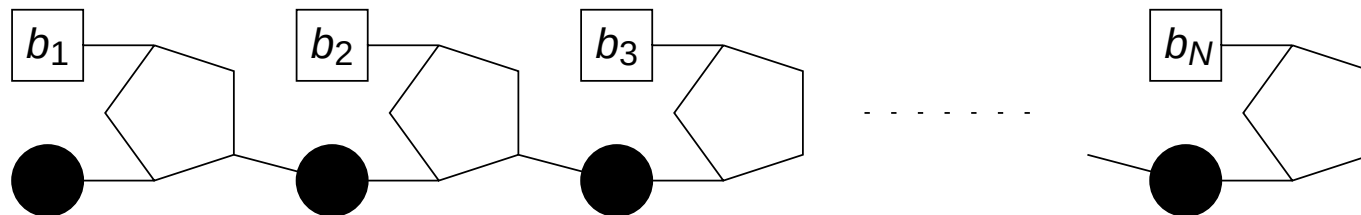


nucléotide :

ADN 2

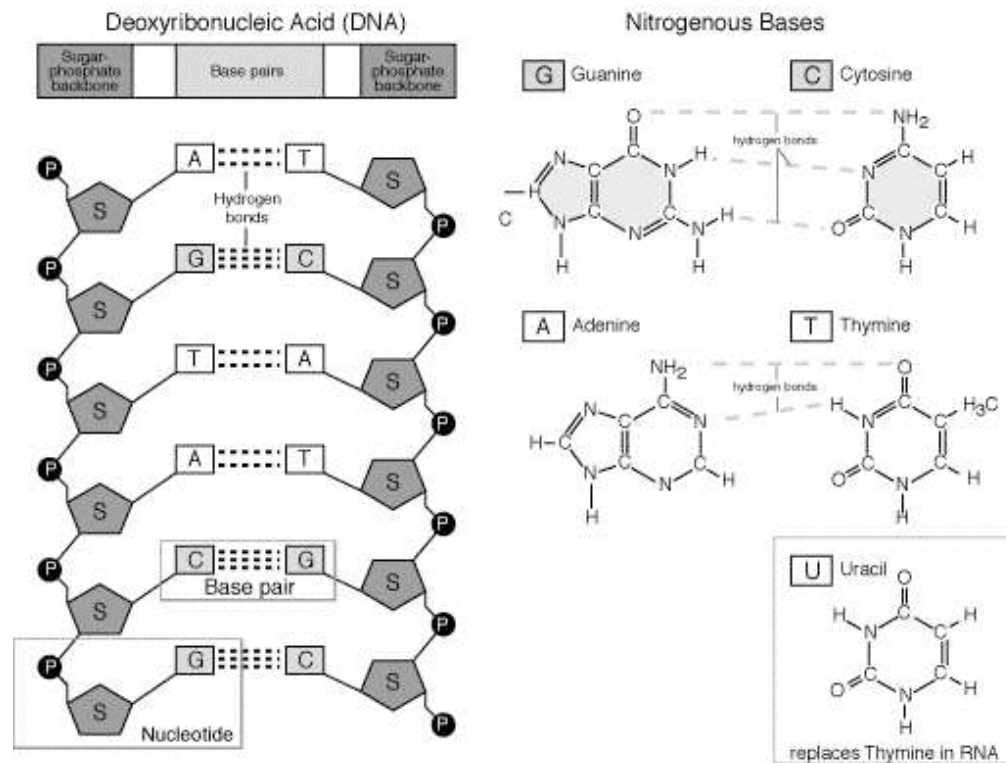
4 bases : adénine, thymine, guanine, et cytosine

chaîne de ADN :



$b_i = A, C, G, T.$

ADN : bases complémentaires



Access Excellence

Watson et Crick : double hélice

1 pas de l'hélice : 10 paires de base, 3.4 nm (forme usuelle)

Doug Lundberg

ADN

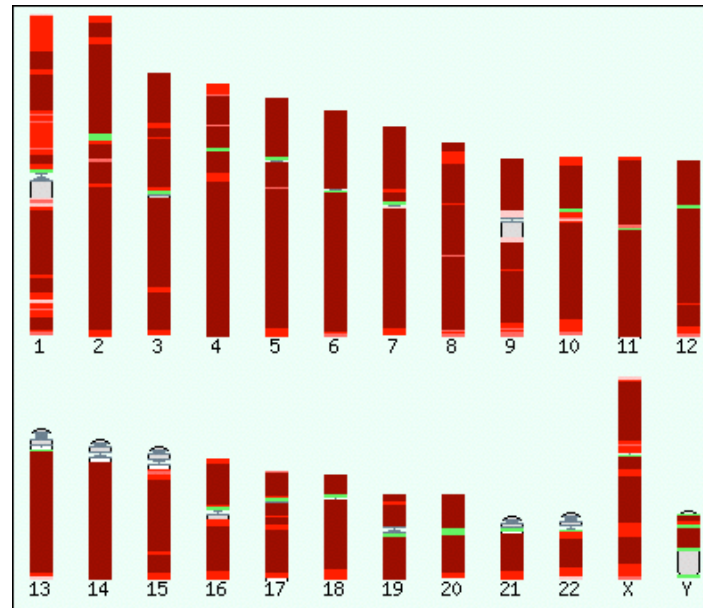
stockage de l'ADN : dans le noyau de la cellule, organisé en chromosomes

1. duplication de l'information en ADN : hérédité
2. duplication de l'information en ARN (qui est traduit en protéines)

Réplication

Lodish et al., Molecular Biology of the Cell, 2002

Chromosomes



22 chromosomes diploïdes et 2 chromosomes sexuels

tailles : $250 \cdot 10^6$ – $23 \cdot 10^6$ paires de base
taille totale du génome humaine : $3 \cdot 10^9$ pb

NCBI

Code génétique

20 acides aminés encodés par 4 nucléotides ?

Encodages par triplets : 64 codons

AAA = Phe	AAG = Phe	AAT = Leu	AAC = Leu
AGA = Ser	AGG = Ser	AGT = Ser	AGC = Ser
ATA = Tyr	ATG = Tyr	ATT = FIN	ATC = FIN
...			

3 triplets d'arrêt, 1 triplet de début (encode aussi Met)

De ADN à protéine

1. transcription : copie du brin informatif à ARN messenger
(ARN : utilise Uracile au lieu de Thymine)

2. traduction : ARNm à protéine (par le ribosome) : acides aminés fournis
par ARN de transfert

Mécanisme universelle !

⇒ Problème de déterminer la séquence de l'ADN.

De ADN à protéine

Lodish et al., Molecular Biology of the Cell, 2002

Gènes

Gène : unité d'hérédité, se traduit à une protéine (en général)

Exons et introns

⇒ Problème de prédiction de gènes.

Expression d'un gène : quantité de ARNm transcrit ⇒ Problème d'analyse de l'expression génétique.

Comparaisons

Comparaisons de séquences : pour déterminer structure, fonction, ...

⇒ **Problème de comparaison de séquences.**

Grandes bases de données

⇒ **Problème de recherche de séquences.**

Évolution

Calcul moléculaire