

Nom Assignatura (20431) Biologia Estructural

Titulació/estudi: Biologia humana

Curs: 4t

Trimestre: 2n

Nombre de crèdits ECTS: 6,5

Hores dedicació estudiant: 66 hores presencials

Llengua o llengües de la docència: anglés

Professorat: Baldomero Oliva, Núria Centeno-Boada, Jordi Villà, i Jana Selent

1. Presentació de l'assignatura

Amb aquesta assignatura es pretén que l'alumne adquireixi coneixements pràctics i teòrics sobre l'estructura proteica i de les biomacromolècules, així com dels mètodes emprats per al seu estudi i caracterització. La matèria que integrarà el present curs comprèn:

- La introducció a les bases biofísiques de sistemes moleculars.
- Els principis estructurals dels biopolímers: proteïnes i DNA.
- La determinació de l'estructura tridimensional de biomol.lècules.
- La relació estructura/funció de les proteïnes.
- La simulació molecular de proteïnes i DNA.

2. Competències a assolir

1) Reconeixement dels patrons estructurals de biomolècules i fàrmacs i relació amb la seva funció biològica. L'estudiant ha de demostrar comprensió dels descriptors físicoquímics d'estructura: termes d'energia potencial, solubilitat, acidesa, hidrofobicitat

2) Relacionar l'estructura tridimensional de les biomolècules amb la seva funció biològica

L'estudiant ha de demostrar comprensió de:

- Estabilitat de les proteïnes i determinants de la seva funció. Fonaments del plegament de proteïnes globulars i de membrana.
- Relació entre seqüència, estructura, i funció: flexibilitat i similitud global i local de la seqüència, conservació tridimensional de centres actius, conservació d'interaccions amb lligands i altres proteïnes.
- Bases i aplicacions del modelat per homologia.

L'estudiant ha de ser capaç de:

- Reconèixer les classes de plegament i els dominis principals de les proteïnes.
- Identificar els residus conservats en una estructura i descriure la seva possible funció estructural.
- Analitzar estructuralment la interacció entre proteïnes o entre proteïnes i altres compostos.

3) Manejar el programari bàsic que permet tractar les dades d'estructures i seqüències de biomolècules.

L'estudiant ha de demostrar comprensió de:

- Les bases de dades mes importants actualment en la recerca en biomedicina.
- Patrons seqüencials, matrius de pesos posicionals, models probabilistes de les seqüències i models ocults de Markov.
- Bases teòriques del modelat de proteïnes per homologia i per mètodes *ab initio*

L'estudiant ha de ser capaç de:

- Usar el programa RASMOL per a visualitzar i analitzar estructures de proteïnes.
- Identificar els elements d'un fitxer del Protein Data Bank (PDB), extreure'n la seqüència de proteïna i dividir una estructura en les seves cadenes.
- Calcular l'estructura secundària d'una proteïna amb DSSP i comparar-la amb una predicció.
- Fer recerques de similitud utilitzant BLAST, FASTA i PSI-BLAST fent servir diverses bases de dades i matrius de pesos; i en particular cerca de motius en seqüències de DNA: llocs d'splicing, motius promotors.
- Obtenir un alineament múltiple de seqüències amb CLUSTAL o T-COFFEE
- Accedir i manegar els repositoris d'informació genòmica i de proteïnes: ENSEMBL, GENOME BROWSER, NCBI, SWISSPROT, SCOP, CATH, PDB, SMART i PFAM.

4) Conèixer els mètodes d'anàlisi experimental d'estructures i sistemes biològics

L'estudiant ha de demostrar comprensió de:

- Els raigs X; la seva obtenció i propietats. Aplicacions: anàlisi d'un cristall de proteïna.

Fonaments de la ressonància magnètica nuclear (RMN). RMN bidimensional (aplicacions a la determinació d'estructura de biomolècules).

3. Continguts

TEMARI TEÒRIC

PART I: INTRODUCCIÓ

Tema 1. Introducció a l'assignatura.

Disseny de l'assignatura. Distribució de les pràctiques i requeriments. Descripció del treball de biologia estructural. Sistema de puntuació.

PART II: PRINCIPIS DE BIOLOGIA ESTRUCTURAL

Tema 2. Proteïnes: cadena polipeptídica, estructura secundària, terciària i cuaternària.

Seqüència i informació codificada. Blocs conservadors de residus. Homologia i relació seqüencial. Entropia i solvatació. Principis energètics: conceptes de força i treball. Efecte de clatració i entropia del medi solvent. Principis termodinàmics. Principis generals del plegament de proteïnes globulars: nucli hidrofòbic i elements d'estructura secundària. Definició de l'estructura secundària. L'espai de Ramachandran. L'hèlix alfa i el full beta. Estructura supersecundària i connexions d'estructures secundàries (llaços). Empaquetament d'alfa-hèlixs. El domini tot alfa. Domini α/β : el plegament Rossmann i el barril TIM. Domini $\alpha+\beta$. Domini tot beta: barril i super-barril de β meandres, el sandvitx *greek-key* i el *jelly-roll*.

PART III: ESPAI CONFORMACIONAL

Tema 3. Espai conformacional i dinàmica molecular

Camps de força moleculars. Utilitzar la funció d'energia potencial per trobar estructures estables de models tridimensionals de proteïnes mitjançant processos de minimització. L'espai de fase, l'espai conformacional i la funció de partició. Entropia (S), entalpia (H) i energia lliure (G). Adonar-nos que els sistemes són flexibles i a usar la dinàmica molecular per explorar aquesta flexibilitat. Comprendre i aplicar mètodes computacionals com el potencial de força mitja o la teoria de perturbacions de l'energia lliure per relacionar la descripció atòmica (microscòpica) amb l'informació mesurable al laboratori (macroscòpic) sobre processos bioquímics. Algoritmes de diferències finites (Verlet). *Simulated Annealing*. Efecte del solvent en simulacions de proteïnes. Condicions de contorn i tractament del camp electrostàtic.

PART IV: DETERMINACIÓ DE L'ESTRUCTURA

Tema 4a. Cristal·lografia de proteïnes i difracció de raigs X

Nocions de mecànica ondulatòria. El principi de la superposició d'ones. Sèries de Fourier i la transformada de Fourier. Difracció de raigs X en cristalls. Expressió de la intensitat de reflexió. La llei de Bragg. Mapes de densitat electrònica. Factors tèrmics i estructurals de vibració. El problema de les fases. Reemplaçament isomorf únic i múltiple difracció anòmala. Refinament estructural i reemplaçament molecular. Microscòpia electrònica d'alta resolució (HREM). Combinació de mètodes experimentals: HREM + X-ray. Espectres de massa

Tema 4b. Ressonància magnètica nuclear

El fenomen de la relaxació i l'espín. La ressonància magnètica nuclear: la inducció nuclear, el camp pulsat de radiofreqüència (RF) i el *free-induction decay* (FID). Nuclear Overhauser Enhancement (NOE). Espectroscòpia de RMN bidimensional. Determinació de l'estructura tridimensional de proteïnes i àcids nucleics mitjançant restriccions de distància.

Tema 5. Modelatge Comparatiu

La classificació de proteïnes i les relacions evolutives de la seqüència. Definició d'homologia. Mètodes de superposició d'estructures terciàries. Caracterització dels centres actius i centres funcionals. Models ocults de Markov. PFAM i SMART. Aliniament de seqüències. Selecció del motlle. Detecció de problemes en l'aliniament. Regions variables i regions conservades. Classificació de llacós. Construcció del model i de la bastida d'una proteïna. Optimització de llacós i cadenes laterals.

Tema 6. Predicció de plegament

El Teorema del plegament invers. Potencials estadístics. Xarxes neuronals i predicció d'estructura secundària (THREADER i PSIPRED). Mètodes de predicció de plegament i threading. Inferència per funció (3DPSSM i Modlink+). Aliniament d'estructura secundària (TOPITS). Mètodes ab initio i *mini-threading* (ROSETTA).

Tema 7. Estructures en Biologia de Sistemes

Partició de proteïnes en dominis. Interaccions entre cadenes i entre dominis. Predicció d'interaccions físiques basades en dominis. Complexes transitius i permanents. Altres prediccions de relacions entre gens i proteïnes. Sistemes de comunicació i senyalització (fosforilacions). Estudi de xarxes d'interaccions: Interactoma. Transferència d'informació.

PART V: ESTRUCTURA/FUNCIÓ

Presentació de 15 treballs a escollir entre les següents temàtiques:

Estructura del DNA i RNA

La doble hèlix de DNA. Els solcs major i menor de la doble hèlix. Formes B, A i Z de DNA. La triple hèlix de DNA. Reconeixement de bases de DNA. El nucleosoma i la heterocromatina. Estructura de t-RNA. Estructura de la RNAsaP. Estructura del Ribozima. Predicció de l'estructura secundària de RNA.

Reconeixement del DNA. Factors de transcripció

Estructura del "lambda Cro" i la seva interacció amb DNA: predicció per modelatge de la interacció. Motius *helix-turn-helix*, *zinc-fingers* i *leucine-zippers*. La transcripció del DNA i els factors de transcripció: famílies amb motius *zinc-finger*. Homeodominis. *Leucine-zippers*.

Maquinària de replicació: DNA polimerases, helicases, topoisomerases, etc.

Replicació del DNA. Estructura de la DNA polimerasa I i del fragment Klenow. Estructura d'alguns enzims de restricció i proteïnes estabilitzadores de la cadena senzilla de DNA (SSB). Topoisomerases tipus I i tipus II i l'alliberació de l'estrès torsional del DNA. Transcriptasa inversa. La recombinació del DNA: complex *RuvA-RuvB-Holliday junction*.

Maquinària de traducció: RNA polimerasa, Aa-t-RNA sintetasa, el ribosoma

Estructura de la RNA polimerasa de *E. coli*. Estructura de les aminoacil-tRNA sintetases: el domini catalític i el domini que identifica l'anticodó. El ribosoma: subunitats 30S i 50S del ribosoma procariota. Encaixament de les proteïnes del ribosoma en l'estructura de la seqüència 16S de rRNA. La combinació de la difracció de raigs X i la microscòpia electrònica d'alta resolució.

Enzims amb motius d'unió a nucleòtid

Els processos d'oxidoreducció i els grups prostètics nucleotídics NAD, FAD i FMN. El domini d'unió de NAD: el plegament de Rossmann i la hipòtesi de Rossmann. Transferència

estereoespecífica del protó de NADH. El motiu d'unió de nucleòtid NAD-binding. Fusió de gens entre el motiu FMN-binding del barril α/β i el citocrom b5. Els motius Walker A i B. Quinases: CDK, MAPK, Tyr-quinases i Ser-quinases.

Exemples d'enzims catalítics: proteases

La catalització mitjançant enzims. L'estat de transició, el centre actiu i l'encaixament enzim-substrat. Centres específics d'unió a substrat (*subsites*). Les serin-proteases: aminoàcids de la tríada catalítica i de l'*oxyanion hole*. Metal·loproteases: carboxipeptidases.

Proteïnes del sistema immunitari

Recombinació somàtica implicada en la varietat dels anticossos. Dominis conservats i variables i les regions determinants de la complementarietat. El complex antigen-anticòs. Estructura dels complexos d'histocompatibilitat (MHC): *human lymphocyte antigen* (HLA). Polimorfismes de la cadena pesada de MHC de classe I.

Proteïnes de membrana

Dificultats en la cristal·lització de les proteïnes de membrana. Hèlixs transmembrana i gràfics d'hidropaticitat de l'estructura primària. Bacteriorodopsina. Centre de reacció fotosintètica de *Rhodospseudomonas viridis*. Reacció de fotosíntesi, transferència d'energia i efecte túnel. Estructura de les proteïnes b-transmembrana: porines.

Famílies de receptors i proteïnes G

Proteïnes G i receptors de factors de creixement. El factor de creixement epidèrmic (EGF) i el seu receptor. L'oncogen *v-erb B*. Els receptors d'insulina, EGF i PDGF. Transducció del senyal: tirosina-quinases. L'amplificació del senyal. Estructura de l'adenilat ciclase. Estructura d'una forma truncada de *ras p21* (cH-ras p21). El motiu d'unió a GTP: mutacions implicades en l'activació oncogènica.

Encaix proteïna-proteïna, proteïna-DNA

Comparació complementària de superfícies. Programes automàtics d'encaix proteïna-proteïna i proteïna-DNA. Combinació dels mètodes d'encaix rígid de molècules i de mètodes de simulació molecular.

Estructura del RNA

Tipus d'enllaços, empaquetament i ponts d'hidrogen. Estructura secundària. Predicció de l'estructura secundària. Estructura del t-RNA: El T-loop i els U-turns. Ubicació de l'anticodó. Formacions d'hibrids. Estructura del Ribozima. Estructura de la RNAsa P.

Capsides de virus

Simetria icosaèdrica del virus esfèrics. Utilització de més d'una cadena polipeptídica en un virus complex. La càpside del picorna virus. Virus animals i vegetals.

Spliceosoma

Components U1, snRNA, snRNP U1A. Unitat de l'anell Sm. Proteïna C. Factors auxiliars A2F. Estructura del U2 i U5

Xaperones

Plegament ajudat per proteïnes. Hsp 60, Hsp70, Hsp 90, Hsp100, GroEL i GroES. Dominis d'unió a ATP.

Sistema d'ubiquitinació. Proteosoma

Sistema d'ubiquitinació i eliminació de proteïnes desplegadas. Estructura del proteosoma. Estructura de les Ubiquitines i dels ring finger. Formació de complexos.

Nucleosoma

Empaquetament del DNA. Heterocromatina. Histones H1, H2, H3 i H4

TEMARI PRÀCTIC I DE SEMINARIS

El programa teòric es complementa amb les pràctiques/seminaris següents (sessions de dues hores). L'ordre de presentació es aproximadament l'ordre natural de presentació al calendari:

Seminari 1. Introducció a linux i als programes gràfics

Iniciació Linux: comandes principals, espais de disc. Navegació en directoris. Editors de text.

Visualització i anàlisi de biomolècules: Com es un fitxer del PDB. Visualització amb Rasmol i VMD d'estructures del PDB. Estudi de complexes de proteïnes. Separació en cadenes. Exploració de interaccions entre proteïnes i entre molècules: Càlcul de distàncies i reconeixement de la interfície i analitzar ponts d'hidrogen. Ús del DSSP.

Seminari 2. Caracterització de motius, estructura secundària i terciària.

Caracterització de motius d'estructura supersecundària (b-hairpin, a-hairpin, b-a-b). Estudi dels tipus de plegament: tot alfa, tot beta, alfa/beta, i alfa+beta. Estudi detallat de la classificació de dominis i bases de dades d'estructures.

Seminari 3 Extracció d'informació d'estructures de proteïnes.

Conceptes bàsics de la determinació per difracció de raigs X d'estructura de proteïnes i biomolècules. El PDB: definicions i formats. Lectura d'un fitxer de PDB.

Seminari 4 Optimització de l'energia i simulació per dinàmica molecular

Principis d'optimització energètica. Funció d'energia, camp de forces, mètodes de minimització (Newton-Rapson, Conjugate Gradient, Steepest Descent). Optimització de l'energia d'una proteïna.

Pràctica 1. Optimització de l'energia i simulació per dinàmica molecular.

Estudi del sistema (proteïna) construcció de la topologia i optimització de l'estructura. Simulació per dinàmica molecular del model dissenyat per homologia. Anàlisi i avaluació dels resultats de l'optimització energètica. Anàlisi de la trajectòria: Fluctuacions, RMSD, modes essencials i ponts d'hidrogen.

Seminari 5. Cristal·lografia de proteïnes

Gincama: Principis de cristal·lografia de proteïnes. Com funciona la transformada de Fourier. Resolució del problema de les fases. Mapa de densitat electrònica. Obtenció d'una estructura a partir de l'espectre de difracció.

Pràctica 2. Alineament de seqüències i cerca d'homologia remota

Part 1: Alineaments múltiples mitjançant ClustalW. Position Specific Substitution Matrix (PSSM).

Mètode de cerca mitjançant Psi-Blast.

Part 2: Models ocults de Markov (HMMER). Bases de dades PFAM.

Pràctica 3. Superposició estructural i caracterització de plegaments

Superposició de proteïnes dirigida i càlcul de RMSD. Caracterització del tipus de plegament de diverses estructures problema utilitzant els programes d'alineament estructural (STAMP) i els programes gràfics. Superposició d'estructures.

Pràctica 4. Modelatge comparatiu

Part 1: Cerca de proteïnes homòlogues a la proteïna problema i alineament múltiple per ClustalW.

Part 2: Alineament estructural de les proteïnes homòlogues. Alineament múltiple per Hidden Markov.

Part 3: Modelatge mitjançant MODELLER.

Part 4: Comparació estructural i caracterització del model entre les estructures homòlogues amb les quals aquest es va generar i la forma optimitzada del model. Utilització de PROCHECK.

Part 5: Optimització i validació estereoquímica d'un model.

Pràctica 5. Predicció de plegament (*threading*)

Part 1: Anàlisi de potencial estadístics mitjançant Prosa II. Estudi estructural mitjançant Prosall dels models generats per homologia

Part 2: Estudi predictiu de l'estructura d'una proteïna mitjançant algorismes de plegament invers, *threading*. Utilització de FUGUE, LOOPP i 3DPSSM a internet.

Part 3: Predicció d'estructura secundària per Internet i mètodes de predicció mitjançant Machine Learning (Psi-Pred). Millorar un model d'homologia

Seminari 6. Encaix de proteïnes (Docking)

Repàs de superfícies de potencial i funcions de puntuació (scoring). Reconeixement i predicció de llocs d'interacció (binding site)

Pràctica 6. Encaix de proteïnes (Docking)

Càlculs de docking de proteïnes amb PyDock. Selecció d'estructures de complexos binaris. Macrocomplexos.

4. Avaluació

L'avaluació del rendiment acadèmic es farà segons el barem següent (sobre un total de 10 punts):

Preguntes d'elecció múltiple: 1,5 punts

Preguntes d'assaig: 2.5 punts

Examen de pràctiques: 3 punts

Treball de presentació: 3 punts

Qualsevol tipus de còpia dels apartats d'avaluació implica la no-superació de l'assignatura.

Per aprovar l'assignatura s'exigeix un 40% de nota de cada parcial per promediar, i un 5 sobre 10 del global un cop les notes han estat escalades.

5. Bibliografia i recursos didàctics

5.1. Bibliografia bàsica

BRANDEN, Carl; TOOZE, John. Introduction to Protein Structure. 2a. ed. Garland Publishing, 1999.

CANTOR; SCHIMMEL. Biophysical Chemistry. WH Freeman & Co, 1980.

DAUNE, M. Molecular Biophysics. Oxford: University Press, 1999.

DRENTH. Principles of Protein X-ray Crystallography. Nova York: Springer-Verlag Inc, 1998.

LEACH, A. Molecular Modelling: Principles and Applications. 2a. ed. Harlow: Pearson Education, 2001.

FINKELSTEIN, ALEXEI V. Protein physics : a course of lectures / Alexei V. Finkelstein, Oleg B. Ptitsyn AmsterdamBoston : Academic Press, cop. 2002.

PHILIP E. BOURNE Structural bioinformatics / edited by Philip E. Bourne, Helge Weissig Hoboken, N.J. : Wiley-Liss, c2003

TAYLOR, W. R. (WILLIE R.). Protein geometry, classification, topology and symmetry : a computational analysis of structure / William R. Taylor and András Aszódi Bristol : Institute of Physics Pub., cop. 2005

5.2. Bibliografia complementària

ULLO LANGE ET AL. Introduction to peptides and proteins / by Ullo Langel, Benjamin F. Cravatt, Astrid Graslund, N.G.H. von Heijne, Matjaz Zorko (2010)

A. PANCHENKO, T. PRZYTYCKA Protein-protein interactions and networks : computing methods for identification, analysis and prediction. London: Springer, 2008

PETER TOMPA & ALAN FERSHT Structure and function of intrinsically disordered proteins / by Peter Tompa ; Alan Fersht Publicación. 2010

K. NAJARIAN Systems biology and bioinformatics. CRC Press, 2009

JAKE Y. CHEN & STEFANO LONARDI Biological data mining . 2009

SHELDON J. PARK Protein engineering and design. CRC Press, 2009

GENTLEMAN, ROBERT, R programming for bioinformatics. Boca Raton, Fla. : CRC Press, c2009

VALIENTE FERUGLIO, GABRIEL. Combinatorial pattern matching algorithms in computational biology using Perl and R. Chapman & Hall/CRC, 2009

BURKOWSKI, F. J. Structural bioinformatics : an algorithmic approach. London : Chapman & Hall/CRC, c2009

CREIGHTON, THOMAS E. The biophysical chemistry of nucleic acids & proteins. Helvetian Press, 2010

IRENA ROTERMAN Structure-function relation in proteins / Núm. estándar 978-81-7895-409-7 b14339729

5.3. Recursos didàctics

<http://sbi.imim.es/courses.php>

6. Metodologia

text

7. Programació d'activitats

text