Tisztelt Hallgatók!

A tavalyi ZH betekintésen ígértem a hallgatóknak, egy áttekintő, magyarázó összefoglalót az IDPs két előadáshoz, mert többen túl soknak, szerteágazónak, és nehezen tanulhatónak érezték a diasort. Azt gondoltam, hogy ezt most automatikusan megosztom önökkel itt az előadás anyagok között.

Üdvözlettel:

Angéla

Rendezetlen fehérjék bioinformatikája – összegzés és kiegészítő írott magyarázat.

Igyekeztem az anyag szerkezetét összefoglalni, de ugyanakkor összefüggő szöveget is írni ott, ahol a tavalyi ZH-k alapján úgy éreztem, hogy hiányos a megértés. A kifejtős kérdésekre adott rosszul megfogalmazott, sokszor értelmetlen válaszok azt tükrözték számomra, hogy nehézséget okoz a tanultak valódi megértése: a diák vázlatosak (természetesen), ezért az órai magyarázatok is kellenek, és esetleg a nagyon más előképzettséggel rendelkező hallgatóknak még az sem elég.

1. **Elméleti bevezetés a rendezetlen fehérjék világába**

Az ezredforduló körül alapozódott meg, hogy a fehérjék nem feltétlenül rendelkeznek jól definiált térszerkezettel, helyette a polipeptid lánc flexibilis, mozgékony és számos konformációt felvehet. Felismerték, hogy az ilyen rendezetlen fehérjék is lehetnek natívak, és rendelkezhetnek „normális” sőt fontos sejtbeli funkcióval.

A kezdeti felismerésben szerepet játszottak meglepő egyedi kísérletes eredmények egyes fehérjék kémiai és fizikai tulajdonságaira vonatkozóan. A jelenség jelentősségét viszont csak a széleskörű fehérjeszekvencia és térszerkezeti adatok elérhetősége/elemzése után lehetett igazán felmérni (bioinformatika szerepe!):

1. A sok fehérjeszekvencia adatban, ami mára elérhető, feltűnően nagyobb arányban vannak alacsony komplexitású fehérjék/régiók, mint a szerkezeti adatbázisokban, ezek nagy része rendezetlen (lehet). Vannak ugyan olyan szerkezettel bíró fehérjék/fehérje domének, amik szintén alacsony komplexitást mutatnak a szekvenciájukban, így pl. a coiled-coil hélixet (két hosszú hélix egymás köré csavarodva) tartalmazó fehérjék, vagy a kollagén típusúak. A rendezetlen fehérjék sem feltétlenül mind alacsony komplexitásúak, de jellemző egyfajta torzulás az aminosav összetételükben (több a hidrofil vagy töltött aminosav, kevesebb a hidrofób, vagy a diszulfid hidak kialakítására képes cisztein.)
2. Ha a szekvencia adatbázis tényleges fehérje szekvenciáit randomizáljuk (természetesen számítógépesen), akkor ezek az alacsonykomplexitású régiók sokkal kevésbé lesznek jelen, ami azt sugallja, hogy tényleges meglévő alacsony komplexitású régiók többsége nem esetlegesen keletkezett, hanem az evolúcióban pozitív szelekció alá esett, azaz funkcióval rendelkező fehérje/fehérje szakasz lehet.

A rendezetlen fehérjékre jellemző tulajdonságok mind logikusan a rendezetlenségre magára visszavezethetők:

**Sajátos aminosav összetétel**: vizet kell szeretniük ahhoz (hidrofil és töltött aminosav jöhet, hidrofób kerülendő), hogy a sejtbeli vizes közegnek kitett állapotban lehessenek, és hidrofób mag nélkül viszonylag szabadon „tekergőzzenek a lében”.

**Oldószernek kitett**: ugyanaz, mint fent

**Flexibilis, mobilis:** ugyanúgy a fenti kép

**Szerkezeti sokaság**: egy fehérjemolekula időben dinamikusan „mozog” a különböző konformációs állapotok között, amennyiben a tisztított fehérje oldatát vizsgáljuk, a többféle konformáció egyszerre van jelen (egyik molekula így, a másik épp máshogy stb).

**Nagy hidrodinamikai térfogat**: nyilván nagyobbnak látszik, ha mindenfelé „csápol”, mintha „összegömbölyödve” meghúzná magát.

**Hőstabilitás**: natív módon is olyan, mint a denaturált „normális” fehérjelánc, de mivel nincs hidrofób magja, nem fog aggregálódni.

**Kémiailag ellenálló:** só, pH vátozást jól tűri, mert nincsenek olyan szerkezetet stabilizáló kölcsönhatások az aminosavjai között, amiket megzavarva denaturálódna a fehérje…

**Proteázoknak kitett:** mivel kitekert és oldószernek kitett, gyakorlatilag a teljes fehérjelánc elérhető és könnyen hozzáférhető a proteázok számára. (A szerkezettel bíró fehérjéknél először csak a felszínen kitett hasítóhelyekhez fér hozzá a proteáz, és csak amikor a gombolyag esetleg szétesik a külső bevágások miatt, akkor tud tovább emészteni…)

**Mobilitási anomáliák bizonyos kísérletekben:** SDS-PAGE (itt a többi fehérje is egyfajta kitekert állapotban vándorol a gélben, ezért a rendezetlenek eltérő mobilitása itt elsősorban az eltérő aminosav összetételre, eltérő töltéseloszlásra vezethető vissza, ami miatt az SDS-t valahogy máshogy köti, mint a többi fehérje), SEC/gélszűrés (itt a nagyobb hidrodinamikai térfogata miatt nagyobbnak látszik, és korábban eluálódik, mint azt a mérete indokolná), SAXS, DLS (szintén a nagyobb hidrodinamikai térfogat miatt nagyobbnak látszik).

A rendezetlenség funkcionális sokszínűségére és jelentőségére jó összefoglaló a 14. Dia

Látható, hogy a rendezetlen fehérjékben keletkező mutációk/hiányok/szabályozásbeli defektusok kapcsolatba hozhatók elsősorban a rákos és a neurodegeneratív betegségekkel.

Régen úgy gondolták, hogy a szekvencia egyértelműen meghatározza a 3D szerkezetet, ami pedig egyértelműen meghatározza a funkciót. Ma ezt a paradigmát ki kell bővíteni. A szekvencia többé-kevésbé egyértelműen meghatározza a 3D szerkezetett vagy éppen a rendezetlenséget, ez a szerkezet, vagy éppen a rendezettség hiánya meghatározza/lehetővé teszi a specifikus funkciót.

Végülis a rendezetlenség is tekinthető egyfajta szerkezeti állapotnak, sőt az egész dolgot hasznos egy kontinuitásban szemlélni: definiált 3D szerkezettől a rendezetlenségig, közben pedig a molten globule és a pre-molten globule állapotokkal. A teljesen merev 3D szerkezet amúgy sem igazán létezik, mindig van a fehérjének egy kis mozgása „lélegzése”, az enzimek aktív centrumai körül nagyobb mozgások is gyakran előfordulnak, hiszen a reakció kiindulási, köztes és végtermékei nem ugyanazok, a kötődésük is feltétlez legalább minimális változást az őket megkötő aktív hely szerkezetében. Több fehérje/fehérjeszakasz esetében a másodlagos szerkezeti elemek ugyan jól definiáltak, de ezek egymáshoz képest elmozdulhatnak (molten globule). Vannak olyan fehérjék is, amik rendezetlenek, de egyes szakaszai hajlamosak bizonyos mértékig másodlagos szerkezetet felvenni (pl. az esetek 20%-ában alfa helikális lehet, ezt nevezzük szerkezeti propenzitásnak), ezek tipikusan molekuláris felismerésben szerepet játszó szakaszok. És természetesen akadnak teljesen rendezetlen fehérjék/fehérje szakaszok is.

Ezeken kívül van még egy másik aspektus is: Hogy viselkedik egy rendezetlen fehérje, ha egy kötő partneréhez kötődik (nem szabad állapotban)

1. A rendezetlen fehérjék olykor feltekerednek, ha egy másik fehérjepartnerükhöz kötődnek (induced folding, indukált feltekeredés).
2. Lehetséges, hogy a rendezetlen fehérjében bizonyos %-ban jelen volt egy másodlagos szerkezeti elem, ami a kötődésben szerepet játszik. Ilyenkor ez a konformáció kiszelektálódik a komplexben… (conformational selection).
3. Az is létezik, hogy két rendezetlen, vagy egy rendezett és egy rendezetlen fehérje komplexében sem lesz teljes a rendezettség, a kötött állapotban is rendezetlen marad (fuzzy komplexek)

Milyen kísérletekből tudhatjuk, hogy egy fehérje rendezetlen?

1. Tuti nem fog kristályosodni önmagában, sőt a hosszabb rendezetlen szakaszok zavarják az amúgy rendezett fehérje kristályosodását is. Ha rövidebb szakaszok rendezetlenek, és sikerül a fehérjét kristályosítani, és róla röntgenszerkezetet meghatározni, akkor abban a szerkezetben nem fog látszani (vagy nagyon rossz B faktorral) a rendezetlen rész. Ezért a PDB szerkezeti adatbázisban biztosan nem lesznek benne a teljesen rendezetlen szakaszok. (A rendezetlenség nem egyenlő a szabálytalan szerkezettel).
2. Egyéb szerkezetvizsgálati módszerekben egyértelműen kimutatható a rendezetlenség: SAXS, NMR, CD, MS.
3. Fizikai tulajdonságok méréséből: hőstabilitás, proteolitikus szenzitivitás, anomális mobilitás (SEC, SDS-PAGE), de ezek a mérések önmagukban nem elegendőek a rendezetlenség „bizonyításához”.
4. **Adatbázisok**

Itt azt javaslom, hogy először fussák végig a diákat, és figyeljenek a felépítésre, amit itt alább össze is foglalok. Elsősorban a bevezető, és az elméleti megalapozó diák, ill. az összefoglaló diák érdekesek. Ha ez megvan, akkor érdemes belemenni a részletekbe, leginkább úgy, hogy felmennek a netre, és kinyitják az adott adatbázist, és belekukkantanak, hogy mi mindent tartalmaz, és mi a megjelenése stb. Az ilyen tapasztalat után sokkal érthetőbb és megjegyezhetőbb lesz az anyag. Ne vesszenek el az apró részletekben!

22. dia: milyen típusú adatbázisok vannak a rendezetlenséggel kapcsolatban?

23. dia: Mire jók az adatbázisok?

**Szekvencia adatbázisok**: DisProt, MobiDB, IDEAL; bevezető és összefoglaló diák, összehasonlító dia: 49

**Szerkezeti adatbázisok**: PED (Protein Ensemble Database): szabad állapotú rendezetlen fehérjék szerkezeti sokaságai, FuzDB: rendezetlen fehérjék rendezetlen fuzzy komplexeinek adatbázisa, összefoglaló: 63. Dia

**Interakciós adatbázisok**: ELM (Eukaryotic Linear Motifs, ELM-ek gyűjteménye), DIBS (Database of Disordered Binding Sites, rendezetlen – globuláris interakciók), MFIB (Mutual Folding Induced by Binding, redezetlen-rendezetlen fehérje interkciók úgy, hogy a kötődés hatására kölcsönösen rendeződnek); összefoglaló dia: 82!

**Ebben a részben is van néhány elméleti megalapozás,** pl az ELM-ek és a MORF-ok fogalma (65. Dia), és összehasonlítása, amit sokan nem vettek az első körben – pedig ezt tényleg érdemes megjegyezni, mert 20 év múlva is így lesz, az adatbázisok meg sokkal változóbbak…

Tehát az ELM és a MORF is a rendezetlen fehérjék funkcionális jellemzően felismerő szakaszai. Tehát ezeken keresztül tud megvalósulni számos specifikus kölcsönhatás. Mi a fő különbség? Az ELMek lineáris motívumok, ami azt jelenti, hogy a rövid aminosav szekvencia valamelyest konzervált (leírható reguláris kifejezéssel, konszenzus szekvenciával), és a felismerő tulajdonság a szekvenciában van „kódolva”. Ezzel szemben a MORFok nem feltétlenül konzerváltak szekvenciájukban, de szerkezeti propenzitással bíró szakaszok (kicsit hosszabbak is mint az ELMek), ami alapját képezi egy specifikus kölcsönhatás kialakulásának a rendezetlen fehérje és a rendezett partnere között.

1. **Rendezetlenség prediktorok**

Itt is azt javaslom, hogy először fussák végig, figyelve a mondanivaló szerkezetére. (Amúgy azért nagyon jól van ám megszerkesztve ez a két óra!!!) Aztán próbálják megérteni a részleteket. Ne magoljanak – az fájdalmas, és nem sokat ér.

Az első diák: 4. 5. Kellenek.

A módszerek:

1. **Egyszerű statisztikai**
	1. AAindex (7. Dia)
	2. DisEMBL (1 paraméter: PDB szerkezetekben a „rendezetlenség” annotáció (3 érték lehet…))
	3. Uversky plot (2 paraméter: hidrofobicitás és töltés)
	4. FoldIndex (pozíció specifikus „Uversky plot”)
2. **Gépi tanulás**: tanuló adatszett (ismert rendezett és rendezetlen fehérjék sok ismert paraméterrel) -> sokféle paraméter az input -> (neural network, hidden layer) -> output: rendezett vagy rendezetlen az adott fehérje; a tanuló szetten ellenőrizhető/pontozható az eredmény, több körben az az „analízis mód” fog kiszelektálódni a rendszerben, ami a legjobb pontszámokat adja a tanuló szetten. Aztán ez használható ismeretlen fehérjéken is prediktorként.
	1. PDONR
	2. ESPRITZ
	3. Support Vector Machine
3. Szerkezeti megközelítés
	1. IUPRED

Jó összefoglaló dia a 47.

Itt is van egy kvázi kitérő, amit többen belekevertek a ZH-ban egyes válaszokba, de nem volt egyértelmű, hogy értik, miről szólt. A „két dolog, amit inkább ne tegyünk” című diákkal kapcsolatban: itt két olyan prediktort mutat, ami egyoldalú megközelítésen alapul, és ezért csak NAGYON limitált a relevanciája. Az egyik a SEG, ami tulajdonképpen egy szekvencia analizáló eszköz, ami képes a szekvencia szakaszok komplexitását számolni (alapvetően a ZH-ban is kérdezett képlet alapján). Ez teljesen oké, azzal kell vigyázni, ha pusztán az alacsony komplexitásból azt a hibás következtetést vonjuk le, hogy az biztosan rendezetlen (ui. lehet pl. coiled coil, vagy kollagénszerű is), vagy főleg az ellenkező következtetést, ha magas komplexitású, az biztosan rendezett. Szóval ezeket így egyértelműen állítani nem lehet. A másik ilyen prediktor a NORSp, ami másodlagos szerkezeti elemeket jósol, és ha semmilyent sem tud jósolni adott szakaszra azt rendezetlennek tekinti. Ez azért nem igazán jó megközelítés, mert (1) szabálytalan szakasz is lehet rendezett, azaz definiált atomi koordinátákkal leírható, nem flexibilis; (2) a rendezetlen fehérjékben is lehetnek szerkezeti propenzitással bíró szakaszok, amikre az ilyen predikciók is adnak másodlagos szerkezeti becslést.

1. **Pár alapfogalom**, ha nem tiszta, tessék utánaolvasni – az előző ZH-ból úgy tűnt, hogy van némi keveredés a fejekben… Ezek a következők:

**Fehérje domén** = önálló feltekeredési egység -> a térszerkezetben is elkülönüló szerkezeti egység az adott fehérjén belül (nem egyenlő és nem keverendő az alegységgel)

**Alegység** = összetett, azaz több polipeptid láncból álló fehérjék esetén értelmezhető, egy alegység az egy polipeptidlánc az összetett fehérjén belül. Az összetett fehérjékben az alegységek között szorosabb kapcsolat van, mint az ún. fehérjekomplexek esetében a fehérjék között, bár a határ a kettő között nem mindig egyértelmű.

**elsődleges „szerkezet”** = szekvencia, aminosav sorrend

**másodlagos szerkezet** =szerkezeti elemek: hélix, béta-szál, turn…,

**harmadlagos szerkezet** = a polipeptidlánc 3D térszerkezete,

**negyedleges szerkezet** = több alegységből álló fehérjék 3D szerkezete,

(csak a teljesség kedvéért ***szuperharmadlagos szerkezet*** = több doménes, de egy alegységes és részben rendezetlen fehérjék esetén a polipeptid lánc 3D szerkezete, a doméneknek egymáshoz viszonyított helyzete esetleges intramolekuláris interakciója, amit a flexibilis linkerek megengednek/elősegítenek; ez dinamikus szerkezet általában),

**szerkezeti sokaság** = rendezetlen fehérjék különböző lokális E minimumokat jelentő konformációs állapotainak összessége, amik között az egyes fehérje molekulák aktuális konformációi váltakoznak/mozognak.

**Molten globule (olvadt csepp):** a fehérje olyan szerkezeti állapota, amikor a másodlagos szerkezeti elemek viszonylag jól definiáltak, de azok egymáshoz való térbeli viszonya már sokkal kevésbé. Ilyen állapot lehet egy globuláris fehérje denaturációja során átmeneti állapotként, de vannak olyan fehérjék/fehérje domének, amik natív állapotban tartósan ilyenek.