

## Válaszok Dr. Csabai István bírálatára

Nagyon szépen köszönöm Dr. Csabai Istvánnak az MTA doktori dolgozatomra adott pozitív és támogató bírálatát és a bírálatra fordított energiát.

A feltett kérdésekre az alábbiakban válaszolok:

*Kérdés: A 10. oldalon megjegyzi, hogy „a statisztikai potenciálok ... általában felülmúlják a fizika alapú energiafüggvényeket”. Mi ennek a magyarázata? Mi hiányzik a fizikai képből? A statisztikai potenciáloknak lehetőséget adnak-e valamilyen fizikai interpretációra?*

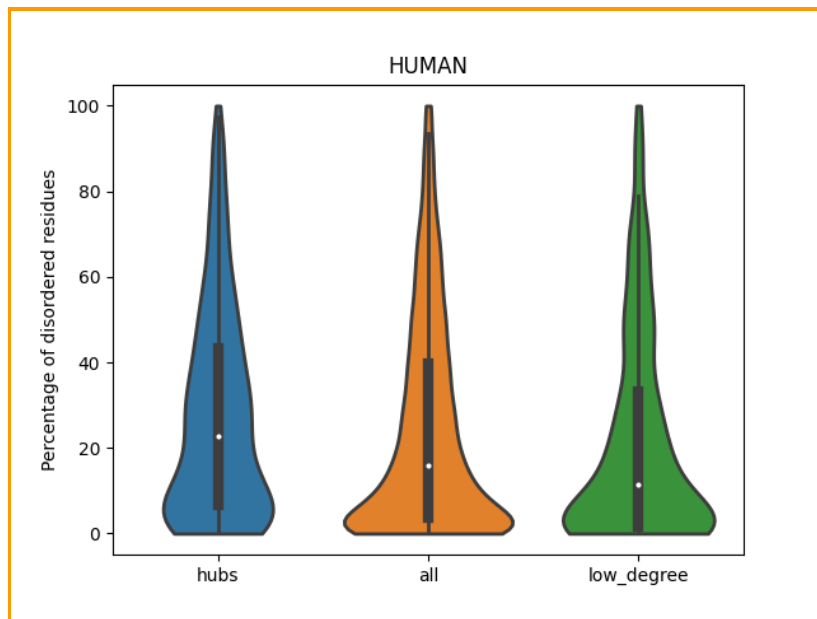
A statisztikus potenciálok robusztusságának kulcsa a hidrofóbicitás. Ezt a jelenséget, ami a domináns tényező a globuláris fehérjék stabilitásában, a statisztikus potenciálok viszonylag jól tudják modellezni, a klasszikus fizikai alapú energiafüggvények viszont nem. A statisztikus potenciáloknak nagyon sok formája és számítási módja van. Az általunk használt megközelítés az egyes aminosav párokhoz rendel egy-egy értéket Thomas és Dill 2001-es cikkében leírt algoritmus alapján [Thomas,P.D. and Dill,K.A. (1996) PNAS, 93, 11628]. Ennek a megközelítésnek a lényege, hogy azokat a paramétereket keresi egy iteratív eljárás során, amelyek maximalizálják a natív szerkezetek valószínűségét az alternatív szerkezetekhez (konformációkhoz) képest. Ezért ebben az esetben a statisztikai potenciálok a valós fizikai energiafüggvények közelítésének tekinthetők az adott formalizmus és alternatív szerkezetek esetén.

*Kérdés: A fehérje hálózatokban gyakran valóban sok homológ fehérje jelenik meg egy-egy csomópont körül, mely a „hagyományos” gén-duplikációs modellt támasztja alá. Mennyiben módosíthatja az eredményeket, ha erre korrigálunk?*

A génduplikáció eredményeképpen valóban létrejöhetnek új kapcsolatok a fehérjék között, azonban ennek nem ez az egyetlen módja. Nagyon gyakori, hogy teljesen új fehérjék, illetve fehérje régiók jelennek meg, és ezek az új régiók gyakran részben vagy egészében rendezetlenek. Ez nagymértékben hozzájárulhat új kölcsönhatások kialakulásához. 2006-ban megjelent munkánk fő hipotézise az volt, hogy ez a mechanizmus az egyik fő motorja a fehérje kölcsönhatási hálózatokban központi szereppel bíró csomópontok megjelenésének.

A gén-duplikációra korrigálhatunk, ha redundancia szűrést végzünk a fehérjékre, és a szűrt kölcsönhatási hálózaton vizsgáljuk a hub fehérjék és a rendezetlenség kapcsolatát. Az alábbi ábrán látható, hogy az általános eredmények nem változnak, például, hogy a hub fehérjékben magasabb a rendezetlen aminosavak aránya (1. ábra).

A számításokhoz a BioGrid adatbázist [Stark,C., *et al.* (2006). Nucleic Acids Res., 34, D535] használtam (version 4.4.228), és a humán fehérjék kölcsönhatási hálózatát szűrtem a CD-HIT algoritmussal [Fu,L., *et al.* (2012) Bioinformatics, 28, 3150] 0.4-es szekvencia azonossági küszöbértékkel. A rendezetlen aminosavak arányát az IUPred módszerünkkel határoztam meg.



1. ábra. A rendezetlen aminosavak arányának megoszlása a humán proteom kölcsönhatási hálózatában a hub fehérjékre (a partnerek számát tekintve a legfelső 1000), az összes fehérjére, és az alacsony fokszámú fehérjékre (a partnerek számát tekintve a legalsó 1000).

*Kérdés: Említi a disszertációban, hogy a rendezetlen kötőhelyekhez jóval nehezebb gyógyszermolekulákat tervezni, mint a jól ismert szerkezetekhez. Milyen módszereket lehet egyáltalán alkalmazni? Hozhat-e valamilyen módszer áttörést a közeljövőben?*

A rendezetlen fehérjéket célzó gyógyszermolekulák azonosítása egyelőre gyerekcipőben jár, bár vannak ígéretes eredmények is ezen a területen. Az eddig vizsgálatok alapján sikerült rendezetlen szegmenshez kötődő specifikus kismolekulát azonosítani, többek között a p53, a c-MYC, a p27 illetve az androgén receptor fehérje esetében [Saurabh,S., *et al.* (2023) ACS Omega, 8, 9729]. Eddig leginkább kísérletes screenelés segítségével azonosítottak rendezetlen fehérjékhez kötődő kismolekulákat. További kísérletekkel lehet jellemezni a konkrét kötőrégiót és a kötődés hatását a rendezetlen fehérje konformációs tulajdonságaira. Érdekes módon, a kötődés mechanizmusa nagyon eltérő lehet: a kismolekula kötődése blokkolhatja a rendezetlen fehérje kötődését a partner fehérjéhez, de okozhatja a rendezetlen fehérje konformációs sokaságának kiszélesedését vagy leszűkülését is, ezáltal megváltoztatva a partnerfehérjével való kölcsönhatás termodinamikai paramétereit. Az androgén receptor esetében három kötőrégió keresztül kapcsolódik a ligand [Zhu,J., Salvatella,X. and Robustelli,P. (2022). Nat. Commun., 13, 6390.]. A c-MYC esetében a flexibilis ligand több dinamikus régió keresztül hatott kölcsön [Jin,F., *et al.* (2013) PLoS Comput. Biol., 9, e1003249].

Általánosságban, a kísérletes és számítógépes módszerek kombinációjára van szükség a kismolekula és a rendezetlen fehérje közötti kölcsönhatás jellegének megértéséhez illetve a gyógyszermolekula optimalizálásához. Az elmúlt években jelentős előrelépés történt a molekuladinamikai számítások terén, mind az erőter függvények mind a mintavételezés tekintetében. Azonban ezek a módszerek még mindig nagyon lassúak és nem eléggé megbízhatóak. Felmerült a gépi tanulás és kvantummechanikai számítások kombinálása is,

ezzel próbálkozott a New Equilibrium Biosciences cég is. A nyelvi modellek a fehérjék és a kismolekulák szintjén is segíthetnek potenciális ligandok azonosításában [Bordin,N., *et al.* (2023) Trends Biochem. Sci., 48, 345.; Moret,M., *et al.* (2023) Nat. Commun., 14, 114.]. Összességében, bár a gépi a tanulás a rendezetlen fehérjékre történő gyógyszertervezés területén is hozhat előrelépést, ehhez sokkal több kísérletes adatra és több vizsgált rendszerre lenne szükség, ahhoz hogy megértsük a kismolekulák rendezetlen fehérjékhez való kötődésének molekuláris alapelveit, illetve, hogy mennyire lehetnek ezek a kölcsönhatások specifikusak sejtes környezetben.

*Kérdés: A vizsgálatokban felhasznált COSMIC adatbázis szomatikus mutációkat tartalmaz. Vizsgálták-e, hogy azok az örökletes mutációk melyek növelik a rák kockázatát milyen kapcsolatban vannak a rendezetlen régiókkal?*

Ezt a kérdést konkrétan nem vizsgáltuk. Feltételezésem szerint a rákhoz kapcsolódó örökletes mutációk is nagyobb arányban fordulnak elő rendezett részekben általánosságban, de vannak példák az ellenkezőjére is. Az egyik szép példa az MLH1 fehérje, ami több, Lynch szindrómához kapcsolódó mutációt tartalmaz. Ezek közül egy mutáció (Y379C), ami vizsgálataink alapján evolúciósan rendkívül konzervált [Pajkos,M., Zeke,A. and Dosztányi,Z. (2020) Biomolecules, 10. 1115], rendezetlen részre esik.

*Általános kérdés: Az utóbbi években az AlphaFold és hasonló nagy gépi tanulás alapú módszerek jelentős átalakulást hoztak a szerkezetbecslő módszerek terén. Milyen változásokat hozhat ez a rendezetlen fehérjék vonatkozásában?*

A gépi tanulásos módszerek már eddig is jelentősen átalakították a rendezetlen fehérjék predikcióját. A legújabb rendezetlenség predikciós módszerek egyre inkább mélytanulásos eljárásokra épülnek [Kurgan,L., *et al.* (2023). Nat. Protoc., 18, 3157]. Az AlphaFold2 módszer a rendezetlen régiók predikciójában is az egyik legjobb módszernek bizonyult [Conte,A.D., *et al.* (2023) Proteins, 91, 1925]. Szintén nagyon fontos előrelépést hozhatnak a fehérje nyelvi modelleken alapuló eljárások [Bordin,N., *et al.* (2023) Trends Biochem. Sci., 48, 345.]. Ezeknek a módszereknek a nagy előnye, hogy nem szükséges hozzájuk többszörös szekvencia illesztés, ezáltal gyorsabbak és nem befolyásolja működésüket, hogy ha egy fehérjéhez kevesebb hasonló szekvenciát lehet találni, ami nagyon gyakori a rendezetlen fehérjék esetén. A gépi tanuláson alapuló térszerkezet predikciós módszerek új lendületet adtak a fehérje tervezés területének is [Ferruz,N., *et al.* (2023) Comput. Struct. Biotechnol. J., 21, 238]. Ugyanakkor fontos megjegyezni, hogy az AlphaFold2 nem oldotta meg a fehérje folding problémáját, és a fehérjék dinamikai és termodinamikai tulajdonságainak jellemzése továbbra is megoldásra vár. Bár az AlphaFold2 legújabb verziója képes komplexek térszerkezetének becslésére is, tapasztalatunk szerint csak korlátozottan tudja a rendezetlen fehérjék kölcsönhatásait modellezni. Továbbra is alapvető probléma a rendezetlen fehérjék esetében a konformációs sokaság jellemzése. Szemben a globuláris fehérjék térszerkezetének becslésével, ezen a területen a módszerek betanításához és teszteléséhez szükséges adatok és a módszerek hatékonyságának jellemzésére szolgáló metrikák jóval kevésbé kiforrottak, ami megnehezíti az áttörés elérését. Összességében azonban, figyelembe véve azt a forradalmi átalakulást, amit a gépi tanulásos módszerek hoztak az elmúlt pár évben, nagyon nehéz ezen a téren előrejelezni, hogy milyen változások várhatóak a rendezetlen fehérjék vonatkozásában.

Ismételten köszönöm Dr. Csabai Istvánnak a bírálatot és kérdéseket. Kérem fenti válaszaim szíves elfogadását, és az MTA doktora cím odaítélésének támogatását!

Budapest, 2023. December 5.

Dr. Dosztányi Zsuzsanna