

Az *in vivo* fehérjeaggregáció útvonalait befolyásoló tényezők vizsgálata *Escherichia coliban*

Györkei Ádám

Ph.D. értekezés tézisei



Témavezető: Dr. Papp Balázs és Dr. Kintses Bálint
Biológia Doktori Iskola
ELKH Szegedi Biológiai Kutatóközpont, Biokémia Intézet
Szegedi Tudományegyetem, Természettudományi és
Informatikai Kar
Szeged, 2022

Összefoglalás

Az elmúlt évtizedekben széles körben elfogadott volt a nézet, miszerint a fehérje aggregátumok belsejében natív szerkezettől és funkciótól mentes makromolekulák találhatóak. Ez köszönhető volt a fehérje aggregátumok felfedezésének körülményeinek, valamint a humán szempontból fontos betegségekben betöltött fontos szerepüknek. Az elmúlt években azonban egyre több olyan eredmény látott napvilágot, amelyek szerint a fehérje aggregációnak fontos fiziológiás szerepe lehet mind baktériumokban, mind eukariótákban. Később arra is fény derült, hogy léteznek olyan aggregátumok, amelyekben a fehérjék megőrzik natív, vagy natív-szerű szerkezetüket és aktivitásukat. Egyes fehérjék aggregátumainak kimerítő vizsgálata ellenére nem történtek olyan szisztematikus mérések, amelyek rendszerszinten megmutatták volna, hogy mely tényezők befolyásolják a képződésüket.

A doktori munkám során szisztematikusan vizsgáltam a különböző *in vivo* aggregációs útvonalakat *E. coli*-ban fehérje overexpresszió hatására egy mikroszkópia alapú, gépi tanulással segített módszerrel. Ennek segítségével igazoltuk a két lényegesen különböző aggregációs fenotípus jelenlétét a sejtekben. Továbbá proteom szinten azonosítottuk azokat a fehérjéket, amelyek a folding folyamatát megelőzően gyorsan vagy azt követően lassabban léptek aggregátumokba. A két aggregációs típus gyakorisága összemérhető proteom szinten, azonban jelentős különbségeket látunk a sejten belüli lokalizációjuk tekintetében. Míg a sötét aggregátumokat képző fehérjék jellemzően a membránban lokalizálódnak

fiziológias körülmények között, a citoplazma aggregálódó fehérjéinek nagy része a lassú aggregációja következtében a folding folyamatát követően csapódik ki. Annak érdekében, hogy azonosíthassuk azokat a fehérje paramétereket, amelyek meghatározzák a két aggregációs típus képződését, összeállítottunk egy szisztematikus adatsort a proteom legtöbb tagjának tulajdonságairól. Azt találtuk, hogy a natív konformációjuk elérése előtt aggregálódó fehérjékkel szemben a fluoreszcensen aggregálódó proteinek gyorsabb foldingkinetikát és kisebb sebességű fehérjeaggregációt mutatnak. A dajkafehérje-rendszerrel, különösen a DnaK fehérjével mutatott interakció szintén segít a fehérjéknek elérni a natív szerkezetüket a transzlációs környezetben az aggregációt megelőzően. Eredményeink arra utalnak, hogy a gyors folding és a lassan bekövetkező aggregáció lehetővé teszi egy fehérje számára, hogy elérje natív konformációját az aggregáció bekövetkezése előtt. Összességében ezek az eredmények arra engednek következtetni, hogy az aktív, natív vagy natív szerű szerkezettel bíró aggregátumok létrejötte az inaktívakkal szemben nagy mértékben az aggregáció és a folding közti kinetikai kompetíció kimenetelének következménye (de Groot és Ventura, 2006). Több bizonyíték is alátámasztja azt, hogy ez a folyamat normál expresszió és genetikai kontextus mellett is jelentőséggel bír. Például a dajkafehérje-rendszer alapvető szereppel bír, normál expressziós szint mellett, az újonnan szintetizálódott és a foldingon átesett fehérjék foldingjában is (Kim et al., 2013). Továbbá a DnaK dajkafehérje klienseinek is két típusa ismert annak függvényében, hogy a chaperon a transzláció során segít a natív konformáció elérésében, vagy ezt követően segít megtartani a szerkezetet (Calloni et al., 2012). Míg a transzlációs apparátus környezete feldúsul

aggregációra érzékeny naszcens polipeptidekben, több fehérje, amely elérte natív szerkezetét, szintén képes aggregálódni, amennyiben koncentrációja átlépi a kritikus oldhatósági határát; ezt a jelenséget szuperszaturációnak hívjuk (Ciryam et al., 2015). A szuperszaturációt mutató fehérjékről kiderült, hogy feldúsulnak neurodegeneratív betegségekben szerepet játszó proteinek között, amely arra utal, hogy szerepük van ezekben a patológiákban és feltehetően az öregedéssel és a stresszhez köthetőekben is. Szükséges további kutatások elvégzése annak érdekében, hogy megérthessük a változó környezeti tényezők és az overexpresszió hatását az aggregációs útvonalra.

Ismert, hogy a fiziológias körülmények között abundáns fehérjék rendkívül alacsony felszíni ragadósággal bírnak azért, hogy az elkerüljék a nemspecifikus interakciókat, amelyek zavart okozhat az ő és más fehérjék működésében (Levy et al., 2012). Eredményeink rávilágítottak, hogy a magas felszíni ragadósággal rendelkező fehérjék szintén alulreprezentáltak az oldható csoportban, arra utalva, hogy a nemspecifikus interakciók csökkentése mögött az aggregáció elkerülése lehet az egyik legfontosabb hajtóerő. Ezen felül azonban azt is láthatjuk, hogy a sötét aggregálódó fehérjék marginálisan magasabb felszíni ragadósággal bírnak a fluoreszcensen aggregálódó társaikhoz képest. Mivel a fehérjék belsejében nem látunk ilyen jellegű különbséget, ezarra utalhat, hogy bizonyos sötét aggregálódó fehérjék esetén a felszíni ragadóság is hozzájárulhat az aggregáció folyamatához. Ez úgy következhet be, hogy az aggregáció az előtt következik be, hogy a fúziós GFP partner elérné natív szerkezetét (ismerve a kísérletekben tapasztalt fenotípust), de azt követően, hogy a naszcens fehérje szerkezete összeomlott volna egy natív szerű, kompakt

szerkezetű, amely felszíne a natív fehérjéhez már részben hasonlatos (Hartl és Hayer-Hartl, 2009; Kim et al., 2013). További vizsgálatok minden bizonnyal további információkkal szolgálhatnak arra nézve, hogy milyen sorrendben következnek be az aggregáció folyamat lépései ezeknél a fehérjéknél.

Munkánk a fehérjerendezetlenség aggregációban játszott szerepével kapcsolatban is új információkkal szolgált. Először is bemutattuk, hogy az *E. coli* fehérjék in vivo szolubilitását nagy mértékben segíti a fehérjerendezetlenség, és ez a hatás független a felszíni ragadóságtól. Továbbá, mivel a fehérjerendezetlenség és a szolubilitás közti korreláció független a töltéstől és a hidrofobicitástól, meg tudjuk erősíteni a nézetet, amely szerint a fehérjerendezetlenség a szerkezeti rugalmasság biztosítása mellett járul hozzá az oldhatósághoz (Santner et al., 2014; Simone et al., 2012). A rendezetlen fehérje régiókról ismert, hogy egyfajta entrópiikus sörteként viselkednek, valamint oldhatóság szempontjából kedvező felszínt biztosítanak. Fontos megjegyezni, hogy az oldható fehérjék több rendezetlen régiót is tartalmaznak, tovább erősítve az entrópiikus hatások szerepét. Továbbá, bár a két aggregációs osztály képviselői nem különböznek a rendezetlen régiók gyakoriságában, a sötét aggregálódó fehérjék több rendezetlen aminosavat tartalmaznak ezeken a régiókon kívül. További vizsgálatok szükségesek, amelyekből megtudhatjuk a rendezetlen aminosavak és régiók egyedi hozzájárulását az oldhatóság fenntartásához és az aggregációhoz. Másodsorban eredményeink arra is rávilágítanak, hogy a természetes szelekció hogyan befolyásolja a fehérjerendezetlenség mértékét proteom szinten. Ugyan ismert, hogy a magasan expresszált fehérjék nagyobb mértékű rendezetlenséggel bírnak

*E. coli*ban (Tartaglia et al., 2009), az emögött álló mechanizmusok eddig ismeretlenek maradtak. Eredményeinkből arra következtethetünk, hogy a megnövekedett fehérjerendezetlenség a magasan expresszált fehérjék esetében arra szolgál, hogy elkerüljék az aggregációt fiziológiás körülmények között. Így tehát mind az alacsony felszíni ragadósság, mind a fehérjerendezetlenség arra szolgál, hogy az abundáns fehérjék elkerüljék az aggregációt *E. coli*ban, valamint feltehetően más baktériumokban is. Harmadsorban pedig jelentős különbséget láthatunk a prokarióták és eukarióták esetében a rendezetlenség aggregációra gyakorolt hatásában. Az eukariótákról közismert, hogy jóval több rendezetlen fehérjét tartalmaznak és több rendezetlen szegmenst is. Ez köszönhető a jelátvitelben és egyéb változatos, átmeneti fehérje interakciókat igénylő sejtfolyamatokban betöltött szerepüknek (Peng et al., 2015). Ezzel egybevé, hogy igazolódott a rendezetlen régiók szerepe a stresszindukált aggregációs folyamatokban (Määttä et al., 2020) ember esetében. Továbbá szemben az *E. coli*val, a magasan expresszált élesztő fehérjék szignifikánsan alacsonyabb mennyiségű rendezetlenséget mutatnak, feltehetően a lineáris motívumok közvetítette nemspecifikus interakciók elkerülése érdekében (Macossay-Castillo et al., 2019). Ez a különbség elvben adódhat a fehérjék tulajdonságaiból, a sejtes környezetből vagy a kettő kombinációjából. Azonban azt látjuk, hogy az *E. coli*ban expresszált humán fehérjék is mutatják a rendezetlenség és az oldhatóság közti pozitív korrelációt, amely arra utal, hogy a különbség a sejtes környezetből adódik. A különbségek mögött rejlő mechanizmusok feltárása jelentős mértékben hozzájárulhat a fehérje-aggregáció folyamatának és a rendezetlenség szerepének

megértéséhez, valamint fontos mérföldkőként szolgálhatna jövőbeni nagy áteresztőképességű vizsgálatokhoz.

Eredményeink a biotechnológiai ipar számára is fontos implikációkkal bírhatnak. Új fehérjetermelési megoldások és rendezőelvek kerülhetnek előtérbe (Rinas et al., 2017; Wu et al., 2011), amennyiben megoldottá válna a funkcionális fehérjék kinyerése az aggregátumokból. A natív szerű enzimek felhasználhatóak immobilizált biokatalizátorként (Rinas et al., 2017; Wu et al., 2011) ipari környezetben. Az aggregátumok vívívóanyagként is szolgálhatnak terápiás felhasználások esetében is, mivel nagyobb stabilitással és eltarthatósággal bírnak. Azt látjuk, hogy minden enzimosztály képviselői jellemzően fluoreszcens aggregátumot képeznek, amely a széles körű felhasználhatóságot sugallnak ipari alkalmazásokban. Metodológiai szempontból a nagy áteresztőképességű mikroszkópia skálázható, hatékony és robusztus módszernek bizonyult az aktív fehérjéket tartalmazó aggregátumok azonosításában ipari célokra. Fontos megjegyezni, hogy bizonyos fehérjék az aktív aggregátumokban idővel inaktív, hibásan feltekeredett formává alakulnak (Elia et al., 2017), ezért a jövőbeli vizsgálatoknak

nagy hangsúlyt kell helyezniük azokra a hatásokra, amelyek segítségével egy fehérje megőrizheti aktivitását, továbbá arra, hogy mekkora az aktív és inaktív fehérjék pontos részaránya az aggregátumok belsejében. A nagy áteresztőképességű mikroszkópiás és gépi tanulási módszerünk, kiegészítve nagy áteresztőképességű fehérjetisztítási (Jäger et al., 2020) és funkcionális vizsgálati eljárásokkal (Huang et al., 2015) fontos részét képezheti az aktív fehérjeaggregátumokban rejlő lehetőségek kiaknázásának.

Célkitűzések:

- Az egyes aggregátum típusok *in vivo* körülmények közötti gyakoriságának vizsgálata, ennek összevetése korábbi *in vitro* eredményekkel.
- Annak megállapítása, hogy a fehérjék *in vivo* körülmények között valóban az oldhatósági határuk közeli koncentrációban fordulnak-e elő a „Life on the edge” hipotézisnek megfelelően.
- A nagy áteresztőképességű mikroszkópia és gépi tanulás módszer hatékonyságának megbecslése *in vivo* aggregáció mérések alkalmazása során.
- Az egyes aggregátumtípusok kialakulásával összefüggő fehérjeparaméterek azonosítása, az aggregáció mögöttes mechanizmusainak megértése.
- A kinetikai kompetíció szerepének vizsgálata az aggregáció típusának meghatározásában.
- Az *in vivo* szolubilitást elősegítő paraméterek azonosítása, különös tekintettel a fehérje szerkezeti és felszíni tulajdonságaira.
- Annak vizsgálata, hogy általánosíthatóak-e eredményeink a heterológ fehérjék expressziójának esetére.

Alkalmazott módszerek:

- Sejttenyésztés (ASKA könyvtár)
- Nagy áteresztőképességű fluoreszcens egysejt-mikroszkópia
- Képelemzés és fenotípusklasszifikáció gépi tanulással
- Génexpresszió kísérletes ellenőrzése: gélelektroforézis és western blot
- Fehérje térszerkezetek bioinformatikai analízise (Areaimol, CCP4)
- Webcrawler (Perl)

- Statisztikai próbák és regressziós analízisek (R)
- Adatbázisok létrehozása és kezelése (R és perl)

A tézis legfontosabb eredményei:

1.) Túltermelés hatására a fehérjék jelentős része képez aggregátumot (kb. 70%) *in vivo* körülmények között. Az aggregátumok közel megegyező mennyiségben tartalmaznak natív-szerű struktúrájú vagy hibásan feltekeredett fehérjéket. Ugyanakkor egy szintén szignifikáns fehérje csoport oldható marad nagymértékű túltermelés ellenére is (kb. 30%), kétségbe vonva az állítást, miszerint a fehérjék döntő többsége az oldhatósági limit körüli koncentrációval bír a sejtekben.

2.) Az aggregációra hajlamos fehérjék köre jelentős átfedést mutat a korábbi *in vitro* eredményekkel, ugyanakkor árnyalja is azokat. A validációs kísérletek eredményeinek tükrében kijelenthető, hogy a nagy áteresztőképességű mikroszkópia és a gépi tanulás alkalmas a fehérje aggregáció gyors és hatékony vizsgálatára, valamint információval szolgál a fehérje folding állapotáról is.

3.) Az aggregátumok létrejöttét befolyásoló tényezők közül számos korábban is ismert volt, amelyek nagy részét a mi adataink is megerősítették. Ilyenek például a natív expressziós szint, a negatív töltés, továbbá a hidrofobicitás. Ezen felül összefüggésbe hozható a fehérjék oldhatósága további tulajdonságokkal, mint a felszíni ragadósság által mediált nemkívánatos interakciók alacsony volta és a magas fogú fehérjerendezetlenség. Ezek korábban ellentmondásos módon jelentek

meg az irodalomban, azonban *E. coli* esetében egyértelmű hatás igazolódott.

4.) A képződő fehérje aggregátumok típusát legnagyobb mértékben a fehérje folding és az aggregáció sebességének viszonya befolyásolja. Ez egybevág a kinetikai kompetíció hipotézisével, amelyet megerősítettünk in vivo körülmények között proteom szinten is. Ezen felül a felszíni ragadósság is bírhat bizonyos szereppel a foldingon részben átesett fehérjék esetében.

5.) Az oldhatóságot elősegítő tényezők közül a fehérje rendezetlenség pozitív szerepét is megerősítettük a heterológ módon expresszált fehérjék esetén. Ez arra utal, hogy a fehérje rendezetlenség ellentmondásos szerepe a különböző organizmus-specifikus hatások eredménye lehet.

A disszertációhoz kapcsolódó publikációk:

1. Györkei, Ádám; Daruka, Lejla; Balogh, Dávid; Őszi, Erika; Magyar, Zoltán; Szappanos, Balázs; Fekete, Gergely; Fuxreiter, Mónika; Horváth, Péter; Pál, Csaba; Kintsés, Bálint; Papp Balázs Proteome-wide landscape of solubility limits in a bacterial cell

SCIENTIFIC REPORTS (elfogadott közlemény, publikálás alatt, 2022)

2. Kintsés, Bálint; Méhi, Orsolya; Ari, Eszter; Számel, Mónika; Györkei, Ádám; Jangir, Pramod K; Nagy, István; Pál, Ferenc; Fekete, Gergely; Tengölics, Roland et al. Phylogenetic barriers to horizontal transfer of antimicrobial peptide resistance genes in the human gut microbiota.

NATURE MICROBIOLOGY 4 : 3 pp. 447-458. , 12 p. (2019) MTMT:

[30435652] IF: 9,68

Referenciák:

- Calloni, G., Chen, T., Schermann, S.M., Chang, H., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-hartl, M., and Hartl, F.U. (2012). DnaK functions as a central hub in the E. coli chaperone network. *CellReports* 1, 251–264.
- Ciryam, P., Kundra, R., Morimoto, R.I., Dobson, C.M., and Vendruscolo, M. (2015). Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases. *Trends Pharmacol. Sci.* 36, 72–77.
- Elia, F., Cantini, F., Chiti, F., Dobson, C.M., and Bemporad, F. (2017). Direct Conversion of an Enzyme from Native-like to Amyloid-like Aggregates within Inclusion Bodies. *Biophys. J.* 112, 2540–2551.
- de Groot, N.S., and Ventura, S. (2006). Protein activity in bacterial inclusion bodies correlates with predicted aggregation rates. *J. Biotechnol.* 125, 110–113.
- Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* 16, 574–581.
- Huang, H., Pandya, C., Liu, C., Al-obaidi, N.F., Wang, M., Zheng, L., and Toews, S. (2015). Panoramic view of a superfamily of phosphatases through substrate profiling. *PNAS* 1974–1983.
- Jäger, V.D., Lamm, R., Küsters, K., Ölçücü, G., Oldiges, M., Jaeger, K., Büchs, J., and Krauss, U. (2020). Catalytically-active inclusion bodies for biotechnology — general concepts , optimization, and application. *Appl. Microbiol. Biotechnol.* 104, 7313–7329.
- Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-hartl, M., and Hartl, F.U. (2013). Molecular Chaperone Functions in Protein Folding and Proteostasis.
- Levy, E.D., De, S., and Teichmann, S.A. (2012). Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. *Proc. Natl. Acad. Sci.* 109, 20461–20466.
- Määttä, T.A., Rettel, M., Helm, D., Stein, F., and Savitski, M.M. (2020). Aggregation and Disaggregation Features of the Human Proteome. *Mol Syst Biol* 16.
- Macossay-Castillo, M., Marvelli, G., Guharoy, M., Jain, A., Kihara, D., Tompa, P., and Wodak, S.J. (2019). The Balancing Act of Intrinsically Disordered Proteins : Enabling Functional Diversity while Minimizing Promiscuity. *J. Mol. Biol.* 431, 1650–1670.
- Peng, Z., Yan, J., Fan, X., Mizianty, M.J., Xue, B., Wang, K., Hu, G., Uversky, V.N., and Kurgan, L. (2015). Exceptionally abundant exceptions : comprehensive characterization of intrinsic disorder in all domains of life. *Cell. Mol. Life Sci.* 72, 137–151.
- Rinas, U., Garcia-fruitós, E., Corchero, J.L., Vázquez, E., Seras-franzoso, J., and Villaverde, A. (2017). Bacterial Inclusion Bodies : Discovering Their Better Half. *Trends Biochem. Sci.* 42, 726–737.

Santner, A.A., Croy, H.C., Vasanwala, H.F., Uversky, N.V., Van, J.Y.-Y., and Dunker, A.K. (2014). Sweeping away protein aggregation with entropic bristles: Intrinsically disordered protein fusions enhance soluble expression. *Biochemistry* 51, 7250–7262.

Simone, A. De, Kitchen, C., Kwan, A.H., Sunde, M., Dobson, C.M., and Frenkel, D. (2012). Intrinsic disorder modulates protein self-assembly and aggregation. *PNAS* 109, 6951–6956.

Tartaglia, G.G., and Vendruscolo, M. (2009). Correlation between mRNA expression levels and protein aggregation propensities in subcellular localisations. *Mol. Biosyst.* 5, 1873–1876.

Wu, W., Xing, L., Zhou, B., and Lin, Z. (2011). Active protein aggregates induced by terminally attached self-assembling peptide ELK16 in *Escherichia coli*. *Microb. Cell Fact.* 42, 1–8.