

ZÁRÓBESZÁMOLÓ

A pályázat címe: Új, orvosbiológiai szempontból fontos moduláris fehérjék azonosítása, szerkezeti és funkcionális jellemzése

OTKA nyilvántartási száma: 49 890

A kutatás célja, a kutatási program ismertetése:

A gerincesekre jellemző moduláris fehérjék kitüntetett orvosbiológiai jelentősége miatt fontos ezeknek a fehérjéknek az azonosítása, biológiai szerepük tisztázása és az ismeretek gyógyászati hasznosítása. Ennek megfelelően célul tűztük ki, hogy:

1. Összehasonlító genomikai elemzésekkel azonosítjuk a gerincesekre jellemző moduláris fehérjéket. Ezek közül kiválasztjuk a ma még ismeretlen funkciójú fehérjéket és bioinformatikai, szerkezeti biológiai és funkcionális genomikai módszerekkel vizsgáljuk biológiai szerepüket és orvosbiológiai jelentőségüket.

2. Folytatjuk a fenti megközelítés alapján általunk korábban azonosított multidomén fehérjék (WFIKKN1 és WFIKKN2) és új modultípusok (WIF-, LCCL- és OLF-modulok) szerkezeti és funkcionális jellemzését.

Elért eredmények:

1. Folyamatosan végezzük a különböző genomok összehasonlító bioinformatikai elemzését evolúciobiológiai összefüggések és új fehérjék, fehérje domének azonosítása céljából.

a.) Ennek során a multidomén fehérjéket felépítő domének mobilitására vonatkozó statisztikai elemzéseket végeztünk. Megállapítottuk, hogy a domének mobilitását legjobban azon lokális környezetek (tripletek) száma jellemzi, amelyekben az adott domén előfordul. Összehasonlítva a domének mobilitását különböző evolúciós csoportokban megállapítható, hogy metazoaokban nagyobb a mobilis domének száma mint prokariótákban vagy más eukariótákban. A különböző metazoa fehérjék összehasonlítása kimutatta, hogy az extracelluláris fehérjéket felépítő domének a legmobilisabbak. Az elemzésből az is kiderült, hogy összefüggés van a domének mérete és mobilitása között: kisebb domének gyakrabban

fordulnak elő különböző multidomén fehérjékben. (Tordai H , Nagy A , Farkas K , Bányai L and Patthy L (2005) **Modules, multidomain proteins and organismic complexity.FEBS J 272, 5064-78.**)

b.) A génazonosítási módszerek megbízhatósága rohamosan fejlődik, ezért új gének azonosításának esélye lényegesen csökkent. Ugyanakkor, súlyos problémát jelent, hogy az azonosított gének jelentős hányadáról bizonyosodik be, hogy a bioinformatikai módszerekkel megjósolt szerkezetük téves. A génpredikciós módszerek bizonytalanságai így komoly problémákat okoznak a (tévesen) megjósolt gének/fehérjék expresszióját szabályozó genomikai elemek meghatározásában, funkciójuk további vizsgálatában.

Csatlakozva, az Európai Unió FP6 programja által finanszírozott BioSapiens konzorcium és az ELTE által koordinált eScience RET munkájához, az *in silico* génpredikciós eljárásokkal rosszul predikált gének azonosítását végezzük. Ennek keretében módszert dolgoztunk ki olyan fehérjék azonosítására, melyek szerkezete ellentmond alapvető fehérjeszerkezeti törvényszerűségeknek. A módszer alkalmas abnormális fehérjeformák azonosítására.

Segítségével kimutattuk, hogy a fehérjekódoló génekről alternatív splicing révén keletkező mRNS izoformák jelentős hányada "életképtelen" fehérjét kódol, így ezeknek az izoformáknak nincs fiziológia szerepe (Tress et al., 2007 **The implications of alternative splicing in the ENCODE protein complement.Proc Natl Acad Sci U S A 104, 5495-500**).

A módszer jelentőségét elsősorban abban látjuk, hogy lehetőséget nyújt a génpredikciós eljárásokkal rosszul prediktált (abnormális fehérjét kódoló) gének azonosítására és így a génpredikció minőségellenőrzésére (Nagy A, Hegyi H, Farkas K, Tordai H, Kozma E, Banyai L, Patthy L (2008) **Quality control of gene prediction. In: Modern Genome Annotation. The Biosapiens Network. Dmitrij Frishman and Alfonso Valencia (eds), Springer**). Az elemzés menete a <http://mispred.enzim.hu> honlapon megtalálható.

c.) A *Drosophila* Teq és a humán neurotripszin multidomén proteázok szekvenciájának összehasonlító vizsgálatával cáfoltuk Didelot és mtsai.(*Science* 313, 851, 2006)) állítását, mely szerint a két fehérje ortológ és így a *Drosophila* Teq funkciójának jellemzése közelebb visz a neurotripszin neurobiológiai funkciójának megértéséhez. (Sonderegger and Patthy, (2007) **Comment on "Tequila, a neurotrypsin ortholog, regulates long-term memory formation in Drosophila". Science 316, 1698**).

2. Kutatási programunkban vállaltuk, hogy vizsgáljuk az általunk korábban bioinformatikai módszerekkel azonosított WFIKKN1 és WFIKKN2 multidomén fehérjék és a

közelmúltban általunk definiált néhány új modultípus, a WIF-, LCCL- és OLF-modulok szerkezeti és funkcionális jellemzését

a.) A WFIKKN2 fehérjéről J. Hill és munkatársai korábban kimutatták, hogy kötődik a miosztatinhoz és gátolja annak aktivitását. A WFIKKN2 és a miosztatin közötti kölcsönhatás szerkezeti feltételeinek tisztázása céljából *Pichia pastoris* szekréción expressziós rendszerben előállítottuk a WFIKKN2 fehérje különálló WAP, FS, KU2 és NTR doménjeit, valamint a WAP+FS és KU2+NTR két-doménes konstrukciókat. Mivel a *Pichia pastoris* expressziós rendszerben a teljes WFIKKN2 fehérjét nem sikerült expresszálni egy új expressziós rendszer meghonosítása vált szükségessé. Sikeresen alkalmaztuk a *Drosophila melanogaster* S2 sejtes expressziós rendszert a teljes WFIKKN2 szekréciónjára.

Módszert dolgoztunk ki a miosztatin próregióját tartalmazó fehérje bakteriális expressziójára és refoldálására.

Felületi plazmonrezonancia (SPR, Biacore) mérésekkel jellemeztük a WFIKKN2 és doménjei valamint a miosztatin és miosztatin prodomén közötti kölcsönhatást és meghatároztuk a kölcsönhatás kialakításában részt vevő doméneket (Kondás és mtsai, kéziratban).

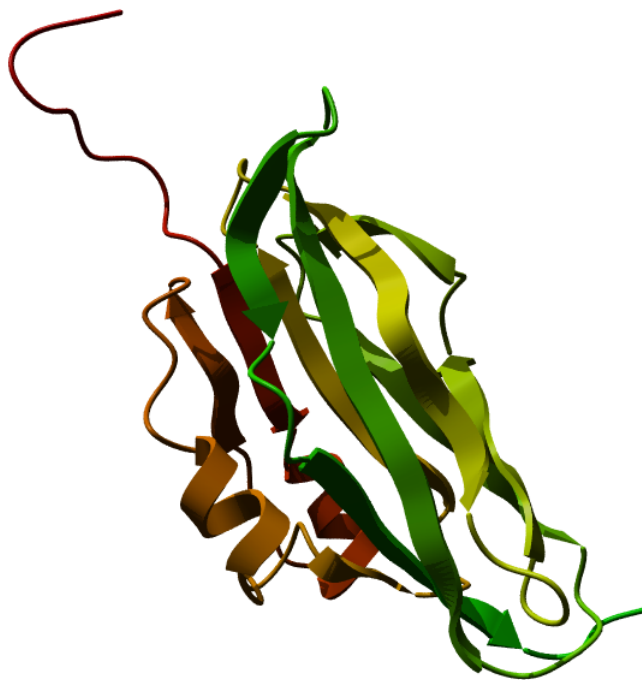
b.) A WFIKKN1 fehérje szerkezetének és biológiai szerepének jellemzése érdekében - a Gottfried Ötting (Australian National University, Canberra, Australia) által vezetett NMR spektroszkópiás munkacsoporttal együttműködve - meghatároztuk a WFIKKN1 második Kunitz-típusú proteázinhibitor doménjének térszerkezetét (**Liepinsh E, Nagy A, Trexler M, Patthy L, Otting G. (2006) Second Kunitz-type protease inhibitor domain of the human WFIKKN1 protein. J Biomol NMR. 35(1):73-78.**). A térszerkezetet más Kunitz-típusú proteázinhibitor domének térszerkezetével, illetve proteáz-proteázinhibitor komplexek ismert térszerkezetével összevetve megállapítható, hogy a WFIKKN1 KUN2 domén 22. pozíciójában levő Trp aminosav oldallánca gátolja egy, a proteázokra és inhibitoraikra jellemző erős kölcsönhatás kialakulását. Korábbi enzimkinetikai méréseink is azt mutatták, hogy a WFIKKN1 KUN2 domén a BPTI-hez viszonyítva több nagyságrenddel gyengébben gátolja a tripszint. Tehát valószínű, hogy a gátlás gyengeségét a 22. pozícióban levő Trp oldallánc kedvezőtlen konformációja okozza.

A KUN2 fehérjével végzett kísérletek megingatják azt a korábbi – a fehérje doménösszetételén alapult – elképzelésünket, hogy a WFIKKN1 fehérje funkciója különböző proteázok gátlása. Valószínűnek tűnik, hogy a WFIKKN1 fehérje a homológ fehérjéhez hasonlóan valamely TGF-beta növekedési faktor családba tartozó fehérje aktivitásának

szabályozásában játszik szerepet. A funkció meghatározása céljából *Drosophila melanogaster* S2 sejt expressziós rendszerben előállítottuk a teljes WFIKKN1 fehérjét és *Pichia pastoris* szekréción rendszerben pedig a különálló WAP, FS, KU2 és NTR doméneket, valamint a WAP+FS és KU2+NTR két domént tartalmazó fehérjéket.

SPR mérésekkel kimutattuk, hogy a WFIKKN1 fehérje, a WFIKKN2-höz hasonlóan, kötődik a miosztatinhoz és a miosztatin prodoménjéhez. Ez a megfigyelés arra utal, hogy a WFIKKN1, a WFIKKN2-hoz hasonlóan, a TGF- β család tagjainak kötése révén játszik szerepet az izomfejlődés szabályozásában.

c.) Gottfried Ötting munkacsoportjával együttműködve meghatározzuk a Wnt jelátvitelben részt vevő WIF-1 fehérje WIF doménjének térszerkezetét. (Liepinsh E, Banyai L, Patthy L, Otting G. (2006) NMR structure of the WIF domain of the human Wnt-inhibitory factor-1. *J Mol Biol.* 357(3):942-950). A domén térszerkezetét az immunoglobulin domének szerkezetére emlékeztető 8 beta redő által alkotott szendvics szerkezet és két alfa helix jellemzi (1. ábra). A fehérje rendkívül erősen köti a refoldálás elősegítésére alkalmazott detergens alkil oldalláncát. Feltételeztük, hogy ez a “kötőhely” a WNT fehérjék palmitoil csoportjának kötésére szolgál.



1. ábra A WIF domén térszerkezete (PDB ID: 2D3J)

d.) A belső fülben nagy mennyiségben expresszálódó cochlin fehérje biológiai szerepének, kölcsönható partnereinek azonosítása céljából többféle expressziós vektorba

klónoztuk a cochlint kódoló cDNS-t. Bakteriális expresszió esetén a képződött fehérje inklúziós testet képez, amelyből nem sikerült oldható, natív szerkezetű fehérjét előállítani. *Pichia pastoris* expressziós rendszerben és *Drosophila melanogaster* S2 sejtekben fehérje szekréciónak nem tapasztaltunk. A cochlin előállítása emlős sejtes expressziós rendszerben folyamatban van.

Szerkezet-funkció vizsgálatok nemcsak a teljes fehérjével végezhetők, hanem a fehérjét alkotó doménekkel is. A cochlin LCCL doménjének bakteriális és *Pichia pastoris* expressziós rendszerben történő előállítását korábban megoldottuk. Az elmúlt évben *Pichia pastoris* expressziós rendszerben expresszáltuk a két vWFA domént tartalmazó fehérjét, a fehérje szerkezeti és funkcionális jellemzése folyamatban van.

A kutatási téma további lehetséges irányai:

1. Annak érdekében, hogy az összes génpredikciós hibát azonosítani tudjunk, folytatjuk további hibaaazonosítási módszerek kidolgozását.
2. A nagyszámú hiba kijavítása érdekében eljárást dolgozunk ki a hibásan jószolt gének helyes szerkezetének automatikus meghatározására.
3. A WFIKKN fehérjék és miosztatin közötti kölcsönhatás jellemzését követően megvizsgáljuk, hogy ezek a fehérjék szerepet játszanak-e más, a TGF-beta növekedési családba tartozó fehérjék aktivitásának szabályozásában.
4. Vizsgáljuk a Wnt jelátvitelben részt vevő, a Wnt fehérjék aktivitását gátló WIF-1 fehérje WIF doménje és a Wnt fehérjék közötti kölcsönhatás szerkezeti feltételeit.
5. Előállítjuk a cochlin fehérjét. Jellemezzük a cochlin és doménjei valamint a feltételezett partner, a Col2A típusú kollagén közötti kölcsönhatást.

A pályázathoz kapcsolódó közlemények:

Patthy L (2007) A general theory of gene sharing. *Nat Genet* 39, 701.

Sonderegger P and **Patthy L** (2007) Comment on "Tequila, a neurotrypsin ortholog, regulates long-term memory formation in *Drosophila*". *Science* 316, 1698; author reply 1698.

Tress ML, Martelli PL, Frankish A, Reeves GA, Wesselink JJ, Yeats C, Olason PL, Albrecht M, **Hegyí H**, Giorgetti A, Raimondo D, Lagarde J, Laskowski RA, Lopez G, Sadowski MI, Watson JD, Fariselli P, Rossi I, **Nagy A**, Kai W, Storling Z, Orsini M, Assenov Y, Blankenburg H, Huthmacher C, Ramirez F, Schlicker A,

Denoed F, Jones P, Kerrien S, Orchard S, Antonarakis SE, Reymond A, Birney E, Brunak S, Casadio R, Guigo R, Harrow J, Hermjakob H, Jones DT, Lengauer T, Orengo CA, **Patthy L**, Thornton JM, Tramontano A and Valencia A (2007) The implications of alternative splicing in the ENCODE protein complement. Proc Natl Acad Sci U S A 104, 5495-500.

Liepinsh E , **Bányai L** , **Patthy L** and Otting G (2006) NMR Structure of the WIF Domain of the Human Wnt-Inhibitory Factor-1. J Mol Biol 357, 942-50.

Liepinsh E , **Nagy A** , **Trexler M** , **Patthy L** and Otting G (2006) Second Kunitz-type protease inhibitor domain of the human WFIKKN1 protein. J Biomol NMR 35, 73-8.

Tordai H , **Nagy A** , **Farkas K** , **Bányai L** and **Patthy L** (2005) Modules, multidomain proteins and organismic complexity. FEBS J 272, 5064-78.

Könyv:

Patthy L. (2007) Protein evolution. Second edition, Blackwell Publishing Limited.

Könyvfejezetek:

Patthy L. (2007) The Evolution of Proteome Complexity and Diversity. In: Evolutionary Genomics and Proteomics. Editors: Mark Pagel and Andrew Pomiankowski, Sinauer Associates, Inc.

Patthy L. (2007) Exons and Protein Modules. Encyclopedia of Life Sciences, John Wiley & Sons

Patthy L (2006) Evolutionary genetics. Concepts and case studies. Evolution of multidomain proteins. pp: 211-221. Eds: Fox CW and Wolf JB, Oxford University Press