

# XXXI. OTDK Biológia Szekció

## A rendezetlenség szerepe a korai STOP kodonnal rendelkező mRNS-ek degradációjában

készítette:

**Ács Veronika**

Biológus MSc  
I. évfolyam

témavezető:

dr. Kalmár Lajos, tudományos főmunkatárs  
MTA TTK Enzimológiai Intézet

EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
TERMÉSZETTUDOMÁNYI KAR  
BIOLÓGIAI INTÉZET



Budapest, 2012

# 1. Tartalomjegyzék

|                                                                                                                                      |    |
|--------------------------------------------------------------------------------------------------------------------------------------|----|
| 1. Tartalomjegyzék .....                                                                                                             | 2  |
| 2. Bevezetés .....                                                                                                                   | 2  |
| 3. Célkitűzések .....                                                                                                                | 6  |
| 4. Módszerek .....                                                                                                                   | 7  |
| 4.1 Különböző fajok NMD-fehérjéinek összegyűjtése, homológia keresés .....                                                           | 7  |
| 4.2 A különböző fajok NMD-fehérjéinek összehasonlítása, szerkezeti feltérképezése bioinformatikai módszerekkel .....                 | 7  |
| 4.3 A rendezetlen szakaszok kötéseinek és szerepének feltárása az emlős NMD-fehérjékben bioinformatikai módszerek segítségével ..... | 8  |
| 5. Eredmények és megvitatásuk .....                                                                                                  | 8  |
| 5.1 Különböző fajok NMD-fehérjéinek összegyűjtése, homológia keresés .....                                                           | 8  |
| 5.2 A választott fajok NMD-fehérjéinek összehasonlítása .....                                                                        | 9  |
| 5.3 A rendezetlen szakaszok kötéseinek és szerepének feltárása az emlős NMD-fehérjékben bioinformatikai módszerek segítségével ..... | 10 |
| 6. Összefoglalás .....                                                                                                               | 17 |
| 7. Köszönetnyilvánítás .....                                                                                                         | 18 |
| 8. Irodalomjegyzék .....                                                                                                             | 18 |
| 9. Függelék .....                                                                                                                    | 20 |

## 2. Bevezetés

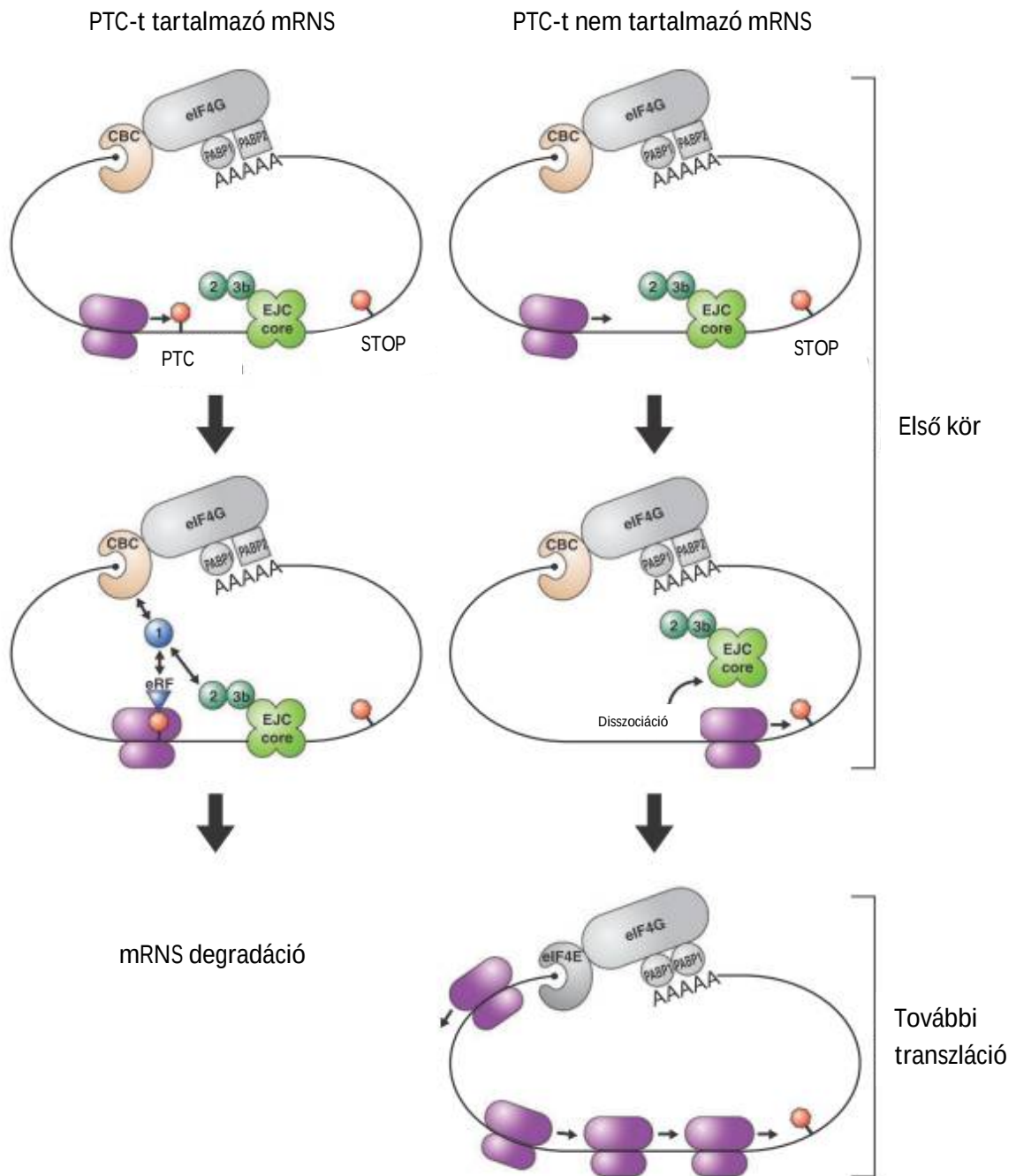
A *Nonsense-mediated mRNA-decay* (NMD) eukariótákra jellemző, a sejt citoplazmájában végbemenő poszttranszkripcionális ellenőrző folyamat, amely a korai STOP kodonnal (PTC) rendelkező mRNS-eket ismeri fel és degradálja. Az ilyen mRNS-ek – amelyek többek között a DNS véletlenszerű *nonsense* mutációjával, *frameshift* mutációval, vagy a *splicing* során keletkezhetnek – a transzlációt követően csonka fehérjéket eredményeznének, amik egyáltalán nem, vagy épp nem megfelelően látnák el működésüket a sejt folyamataiban. Az NMD tehát a csonkolt és így a sejt számára szükségtelen fehérjéket kódoló mRNS-ek kiszűrésére és lebontására evolválódott mechanizmus. Eukariótákban az NMD kulcsfehérjéi, az úgynevezett *up-frameshift* (UPF) fehérjék evolúciósan konzerváltak és a gombáktól egészen az emlősökig minden élőlényben nélkülözhetetlenek a folyamathoz. Az UPF1, UPF2 és UPF3 trimert képezve indítják be az mRNS degradációt úgy, hogy

elsőként ők kapcsolódnak a korai STOP kodonnal elakadt riboszómához. Soksejtűek esetében a *Suppressor with Morphogenetic effect on Genitalia* (SMG) kinázok (SMG1, SMG5, SMG6 és SMG7) is fontos szerepet játszanak a mechanizmusban, egyrészt az UPF1 foszforilációját szabályozzák, másrészt az mRNS elvágásával annak lebontását indítják el [1].

A folyamat egyik kulcskérdése az, hogy vajon a sejt hogyan képes megkülönböztetni a normál és a korai STOP kodonnal rendelkező mRNS-eket egymástól. Az NMD-nek többféle mechanizmusa is ismert, amelyek elsősorban a felismerés alapjében különböznek [2]. Gombákban és egyes alacsonyabb rendű állatokban, így például férgekben az úgynevezett *downstream sequence element* (DSE) segítségével ismeri fel a sejt a korai STOP kodonnal rendelkező mRNS-eket. A DSE ismétlődő szakaszokat is tartalmazó szekvencia, amely a normális STOP kodon előtt, vagyis attól *upstream* helyezkedik el. A riboszóma a leolvasás során felismeri a DSE szekvenciát és eltávolítja róla az őt specifikusan kötő faktorokat, így például az irodalomban már leírt Hrp1-et. PTC-vel rendelkező mRNS esetén azonban az elakadt riboszómától *downstream* helyezkedik el a DSE, így a transzláció során el nem távolított DSE-kötő faktorok kapcsolódhatnak az UPF fehérjékkel, ami egyelőre egy még fel nem tárt mechanizmussal elindítja a mRNS degradációját [3].

Egy másik, már alaposabban vizsgált NMD mechanizmus a PTC következtében létrejövő extrém hosszú 3'-UTR régiókat ismeri fel. A kiterjedtebb 3'-UTR régió ugyanis növeli a riboszóma és az mRNS polyA-végét kötő *polyadenin-binding* fehérjék (PABP) közti távolságot, amelyek így kevésbé tudnak egymással kapcsolódni. A PABP-k helyett ekkor az UPF faktorok kötnek be a riboszómához kapcsolt eukarióta *releasing factor 3*-hoz (eRF3), amelyek beindítják az NMD folyamatát. A mechanizmus ezen típusa rovarokban és növényekben egyaránt megtalálható [2, 4].

A harmadik fajta mechanizmus az intronos, vagy *exon junction complex* (EJC)-alapú NMD, amely emlősökre és növényekre egyaránt jellemző. Az RNS-ből a splicing során kivágódó intronok helyébe, pontosabban az exon-exon határoktól 20-24 bázispárral *upstream* a spliceoszóma fehérékomplexeket, EJC-ket helyez el, amelyek a korábbi intronok helyét jelölik. Ezek később alkalmasak lesznek arra, hogy segítségével a sejt eldöntse, ép mRNS átírását végzi-e vagy sem. Az intronalapú NMD-ben számos fehérje vesz részt, amelyek három fő komplexet alkotnak a folyamatban [2] (1. ábra).



**1. ábra** A transzláció első körének vázlata emlősökben korai STOP kodonnal rendelkező, illetve PTC-t nem tartalmazó mRNS esetében. Az UPF2 és UPF3 fehérjék az EJC-hez kötődve vesznek részt a folyamatban, az NMD kiváltásának iniciátora az UPF1-UPF2 interakció. Korai STOP kodonnal rendelkező mRNS-ben (baloldali ábra) legalább egy EJC található a PTC-től *downstream*, aminek következtében UPF1 kapcsolódik a riboszómához, majd az EJC komplexhez. Ebben a transzláció első körében az mRNS 5'-végén található CBC komplex segíti. PTC-t nem tartalmazó mRNS esetében (jobboldali ábra) az EJC-k mind *upstream* helyezkednek el a STOP kodonhoz képest, így azokat az átírást végző riboszóma szétszereli. A transzláció első körét követően a CBC komplex és a nukleáris PABP fehérjék (PABP2) is lecserélődnek, és az így épnek jelölt mRNS további transzlációkon mehet keresztül [1].

Az első az EJC komplex, amelyet négy magfehérje épít fel: az Y14-MAGOH heterodimer, az eIF4A3 és az MLN51. A riboszóma-komplex a *releasing factor 1*-et és 3-at (eRF1 és eRF3) foglalja magába a riboszóma alegységek mellett. A harmadik, úgynevezett *cap-binding complex* (CBC) – PABP komplex az 5' véget védő NCBP1-NCBP2 heterodimerből, az mRNS polyA végét kötő PABP-kból, valamint a CBC-t és PABP-t összekötő eIF4G fehérjéből épül fel.

PTC-t nem tartalmazó mRNS esetében a STOP kodonnál megálló riboszómához az eRF3 fehérje kapcsolódik, mely után a mRNS 3'-végén levő PABP fehérjével lép interakcióba, és ezzel zárul a transzlációs kör. A transzláció során az EJC komplexeket a riboszóma-asszociált WIBG nevű fehérje szereli szét a két komplex kölcsönhatásakor [5]. Amennyiben a PTC legalább 50-55 nukleotidnyi távolságra található az mRNS-en maradt EJC-től *upstream*, az NMD rendszer felismeri és degradálja az aberrált mRNS-t. Ekkor a megnövekedett távolság miatt a PABP helyett az UPF1 fehérje és a hozzá asszociált SMG1 kináz köt be a riboszóma-komplexhez, amelyben az NCBP1 is segíti UPF1-et. Az így kialakult eRF1-eRF3-UPF1-SMG1 komplexet SURF komplexnek nevezzük. Ezt követően megtörténik a riboszómán levő UPF1 és EJC komplex kapcsolódása, amelyben az EJC-hez asszociált UPF3 és UPF2 szolgálnak hídként. Ennek hatására megtörténik az UPF1 SMG1 kináz általi foszforilációja, mely a degradációs folyamat indítójele. Az SMG1 kináz aktivitását SMG8 és SMG9 fehérjék szabályozzák. A foszforilált UPF1 defoszforilációjában és az mRNS degradációban az SMG5 és SMG7 fehérjék, esetenként pedig az SMG6 játszik fontos szerepet. Előbbi deadenilációt és decappinget okoz, vagyis az mRNS két végét támadja, utóbbi pedig nukleáz aktivitásának köszönhetően képes az mRNS elvágására, majd a további degradációs folyamatok kiváltására [6].

A transzláció első köre emlősökben egyben az mRNS épségének ellenőrzését is jelenti. Ennek során az EJC-k lekerülnek az RNS-ről és szétszerelődnek, a nukleáris PABP-k (PABP2) teljesen lecserélődnek a citoplazmatikus PABP-kre (PABP1) a 3'-végen, valamint a CBC komplex szerepét is más fehérje (eIF4E) veszi át. Mindezek következtében a sejt képes megkülönböztetni a már ellenőrzött és így folyamatosan transzlálható mRNS-eket a még új, első körre váró mRNS-ektől. Újabb kutatások kiderítették, hogy az első transzlációs körben képződő fehérjék szabályozott lebontásra kerülnek, rövid peptidláncokká emésztődnek és részt vesznek az MHC I-hez kötött antigén bemutatásban. E mechanizmusnak köszönhetően biztosított, hogy a csonkolt fehérjéket kódoló mRNS-ekről még egy működő kópia se keletkezzen [7].

### 3. Célkitűzések

Habár az NMD mechanizmusáról egyre több információ áll rendelkezésünkre, továbbra is felmerül a kérdés, hogy hogyan is képes a folyamat a távolság mérésére, a sokszor igen nagy és különböző távolságok áthidalására, és ezáltal az mRNS sorsáról való döntéshozatalra, illetve hogy milyen különbségeket találhatunk a három mechanizmus összehasonlítása esetén. Mind a DSE, mind a 3'-UTR alapú, mind pedig az intronalapú NMD számos, egymástól nagy távolságban elhelyezkedő fehérje kapcsolódásán, komplexek kommunikációján alapul. Az intronalapú NMD-hez legalább három nagy távolságot áthidaló kölcsönhatás létrejöttére van szükség: a CBC-SURF kapcsolatra, a SURF-EJC interakcióra, valamint a foszforilált UPF1-et tartalmazó komplex és a degradációs fehérjék egymásra találására. Továbbá a folyamat hatékonysága érdekében szükség van a komplexek dinamikus szétesésére és újraszerveződésére is. Ehhez a flexibilitáshoz vélhetően nélkülözhetetlenek az olyan, a komplexekből hosszan kinyúló távolságmérő fehérjeszakaszok, amelyek nagy hidrodinamikai sugara előnyt jelent az olykor igen nagy távolságok áthidalásában is (*fly casting* jelenség) [8]. Ezek a tulajdonságok a rendezetlen fehérjékre, fehérjeszakaszokra jellemzőek, amelyek nem rendelkeznek egy jól leírható harmadlagos szerkezettel, a kötésük gyakran gyengébb és kevésbé specifikus a szerkezettel rendelkező doménekhez képest, a fehérjék gyors egymásra találásában azonban nagy előnyt jelentenek. Emellett egyes NMD fehérjék rendezetlensége a komplexek összetartását, negyedleges szerkezetének stabilizálását biztosíthatja, vagyis „állványzatként” is funkcionálhat a komplexen belül [9].

Jelen munka célja volt a bevezetésben említett különböző típusú NMD-mechanizmusokban bizonyítottan részt vevő, az irodalomból már ismert fehérjék, illetve azok lehetséges homológjainak szekvenciális, szerkezeti és funkcionális összehasonlítása bioinformatikai módszerekkel. Ehhez öt, a témában már jól jellemzett és különböző NMD mechanizmusokkal rendelkező példafaj (*Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Homo sapiens*) a folyamatban részt vevő fehérjéit kívántuk összehasonlítani, hogy így feltárjuk a háromféle NMD folyamat, és egyben az öt vizsgált példafaj fehérjei közti hasonlóságokat és különbségeket. A vizsgálat során kiemelten fontosnak tartottuk a rendezetlenség mértékének, a rendezetlen szakaszok hosszának és fehérjén belüli elhelyezkedésének összevetését a már bizonyított és a bioinformatikai módszerekkel jóslott homológok között. Különös hangsúlyt fektettünk az irodalomban legjobban jellemzett EJC-alapú NMD mechanizmus fehérjeinek vizsgálatára,

illetve azok rendezetlen szakaszainak megismerésére, a rendezetlen szakaszok lehetséges funkciójának feltásárára az NMD folyamatán belül.

## 4. Módszerek

### 4.1 Különböző fajok NMD-fehérjéinek összegyűjtése, homológia keresés

Munkánk során öt, az NMD témájában jól publikált példafaj (*Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Homo sapiens*) a folyamatban részt vevő fehérjéit gyűjtöttük össze az eddig a témában megjelent irodalom, valamint a UniProt adatbázis (<http://www.uniprot.org>) segítségével. Lehetséges homológokat a választott fajokban – amelyeket az irodalomban eddig még nem írtak le, de feltételezhetően léteznek és megtalálhatóak az adatbázisban – a BLAST program segítségével kerestünk a humán NMD-fehérjéket alapul véve. A szekvenciális hasonlóság mellett az egyes domének meglétét vagy hiányát, illetve azoknak a fehérjén belüli elhelyezkedését is figyelembe vettük a homológok keresésénél. Ehhez a Pfam adatbázist (<http://pfam.sanger.ac.uk>) használtuk.

### 4.2 A különböző fajok NMD-fehérjéinek összehasonlítása, szerkezeti feltérképezése bioinformatikai módszerekkel

Az előző pontban leírtak alapján talált fehérjék összehasonlításához az általunk fontosnak vélt adatokat gyűjtöttük össze. Minden egyes fehérjéhez hozzárendeltük a hosszúságát, a jóslott rendezetlen szakaszait, a bennük található esetleges *coiled-coil* és egyszálú töltött alfa-hélix (*Charged Single Alpha Helix*, CSAH) [10] régiókat – hogy kiszűrjük a nem valódi rendezetlen részeket – valamint a fehérjékhez tartozó rendezetlenségi mértéket is. A rendezetlen szakaszokat az IUPred (<http://iupred.enzim.hu/>) [11] és a PONDR-VSL2 [12] programok segítségével gyűjtöttük össze, amelynek során kizárólag a huszonöt aminosavnál hosszabb rendezetlen szekvencia részeket vettük figyelembe az egybefüggő rendezetlen szakaszok jóslásánál. A potenciális *coiled-coil* régiókat a COILS ([http://www.ch.embnet.org/software/COILS\\_form.html](http://www.ch.embnet.org/software/COILS_form.html)) [13], a CSAH régiókat pedig a CSAH szerver segítségével (<http://himalia.chem.elte.hu/csah/index.html>) szűrtük ki. A rendezetlenségi mértéket IUPred predikció után saját készítésű programmal számoltunk.

### 4.3 A rendezetlen szakaszok kötéseinek és szerepének feltárása az emlős NMD-fehérjékben bioinformatikai módszerek segítségével

A már kigyűjtött rendezetlen régiók a fehérje működésében és magában az NMD folyamatában betöltött szerepét elsősorban a témában eddig megjelent irodalom, a UniProt adatbázis, illetve a Pfam segítségével határoztuk meg. Ennek során kikerestük az NMD-ben részt vevő fehérjék rendezetlen kötőhelyeit, figyelembe véve ezen szakaszok egyéb lehetséges feladatait is, mint például a távolságmérést, illetve a komplexek összetartását, tehát az „állványzatként” való működésüket. A kísérletes úton bizonyított rendezetlen kötőmotívumok mellett irányadóként figyelembe vettük a szekvencia alapján kötőhelyet jósoló ANCHOR algoritmus eredményeit is (<http://anchor.enzim.hu>) [14].

## 5. Eredmények és megvitatásuk

### 5.1 Különböző fajok NMD-fehérjeinek összegyűjtése, homológia keresés

Az NMD folyamatában részt vevő fő fehérjék bioinformatikai módszerekkel történő keresése során sikerült az összes, publikációkban is leírt fehérjét megtalálnunk. Emellett az irodalomban eddig le nem írt vagy az adott fajban feltételezhetően megtalálható, de még azonosítatlan fehérjék közül is sikerült néhány lehetséges homológot találnunk. Az egyes példafajokhoz tartozó fehérjék megléte, illetve hiánya nagymértékben összeegyeztethető a korábban ismertetett NMD mechanizmusokkal. Az egyes fajokhoz tartozó főbb NMD-fehérjéket és a köztük levő hasonlóság mértékét táblázatba foglaltuk (Táblázat). Az összes általunk vizsgált fehérjét sikerült megtalálnunk *Homo sapiens*-ben, amelyben az EJC alapú NMD működik. *A. thaliana*-ra az EJC és a 3'-UTR alapú NMD egyaránt jellemző. Az SMG fehérjék közül benne csupán az SMG7-et sikerült megtalálnunk, ami egyezik az irodalomban eddig leírtakkal. *D. melanogaster*-ben 3'-UTR-es NMD működik. Az mRNS degradációt benne az SMG6 végzi [6], hozzá tartozó SMG7-et eddig az irodalomban nem írtak le. Az EJC-t alkotó négy fehérje benne is megtalálható, arra azonban nincs bizonyíték, hogy össze is szerelődnek és együtt komplexet alkotnak [15]. *C. elegans*-ban a vártnak megfelelően az UPF-eket, az mRNS végeire jellemző fehérjéket és az SMG kinázokat sikerült megtalálnunk, míg *S. cerevisiae*-ben az SMG-eket eddig nem írták le és mi sem találtunk lehetséges homológokat kutatásunk során.

## 5.2 A választott fajok NMD-fehérjéinek összehasonlítása

A választott öt példafaj NMD-ben részt vevő főbb fehérjéinek összehasonlítása során az alábbi táblázatban szereplő eredményeket kaptuk (további részletekért lásd a Függelék):

**Táblázat:** A táblázat a kutatásunk során összegyűjtött, az öt példafajhoz tartozó NMD-ben részt vevő főbb fehérjét tartalmazza. A zárójelen kívüli számok az adott fehérje hosszát, a zárójelen belüliek pedig az általunk jósolt rendezetlenségi mértéket jelölik (0-1 közötti tartományban). A '-' jel az általunk az adott fajban nem megtalált fehérjére utal. Zöld színnel a humán fehérjéhez nagymértékben hasonló, rózsaszínnel a hozzá csak kismértékben hasonlító, fehérrel pedig a humán fehérjéhez csekély mértékben hasonlító fehérjét jelöltük. A hasonlóság mértékének megállapításához a szekvenciális hasonlóságot, a domének meglétét és elhelyezkedését, a jósolt rendezetlen szakaszokat, a jósolt *coiled-coil* régiókat és a rendezetlenség összmértékét vettük figyelembe.

|               | <i>H. sapiens</i> | <i>D. melanogaster</i> | <i>A. thaliana</i> | <i>C. elegans</i> | <i>S. cerevisiae</i> |
|---------------|-------------------|------------------------|--------------------|-------------------|----------------------|
| <b>UPF1</b>   | 1129 (0,23)       | 1180 (0,22)            | 1254 (0,31)        | 1069 (0,24)       | 971 (0,07)           |
| <b>UPF2</b>   | 1272 (0,43)       | 1241 (0,36)            | 1029 (0,43)        | 1142 (0,32)       | 1089(0,20)           |
| <b>UPF3</b>   | 483 (0,76)        | 492 (0,73)             | 482 (0,74)         | 368 (0,55)        | 387 (0,49)           |
| <b>EIF4A3</b> | 411 (0,05)        | 399 (0,03)             | 414 (0,04)         | -                 | -                    |
| <b>MLN51</b>  | 703 (0,98)        | 782 (0,98)             | 605 (0,91)         | -                 | -                    |
| <b>Y14</b>    | 174 (0,59)        | 165 (0,32)             | 191 (0,56)         | -                 | -                    |
| <b>MAGOH</b>  | 146 (0,01)        | 147 (0,01)             | 150 (0,13)         | -                 | -                    |
| <b>SMG1</b>   | 3657 (0,10)       | 3218 (0,05)            | -                  | 2322 (0,03)       | -                    |
| <b>SMG5</b>   | 1016 (0,23)       | 1177 (0,25)            | -                  | 549 (0,01)        | -                    |
| <b>SMG6</b>   | 1419 (0,47)       | 948 (0,16)             | -                  | 1241 (0,44)       | -                    |
| <b>SMG7</b>   | 1137 (0,51)       | -                      | 1059 (0,29)        | 458 (0,17)        | -                    |
| <b>NCBP1</b>  | 790 (0,07)        | 800 (0,08)             | 848 (0,08)         | 798 (0,08)        | 861 (0,02)           |
| <b>NCBP2</b>  | 156 (0,04)        | 154 (0,03)             | 257 (0,51)         | 154 (0,11)        | 208 (0,25)           |
| <b>PABP1</b>  | 636 (0,32)        | 634 (0,25)             | 671 (0,38)         | 646 (0,35)        | 577 (0,30)           |
| <b>PABP2</b>  | 306 (0,62)        | 224 (0,45)             | 629 (0,37)         | -                 | -                    |
| <b>eRF1</b>   | 437 (0,25)        | 438 (0,04)             | 436 (0,04)         | -                 | -                    |
| <b>eRF3</b>   | 628 (0,33)        | 619 (0,3)              | 543 (0,22)         | -                 | -                    |
| <b>WIBG</b>   | 204 (0,93)        | 207 (0,99)             | 204 (0,98)         | -                 | -                    |
| <b>RNPS1</b>  | 305 (1)           | 374 (0,96)             | 414 (0,95)         | -                 | -                    |
| <b>PNRC2</b>  | 139 (0,83)        | -                      | -                  | -                 | -                    |

Míg az UPF1 mind az öt példafajban nagymértékű szekvenciális és szerkezeti hasonlóságot mutat, addig az UPF2 és UPF3 fehérjék a különböző fajokban már kevésbé hasonlítanak a humán UPF2-höz és UPF3-hoz. A fő különbséget esetünkben főként a szekvenciális eltérés adja, a domének és az általunk jósolt rendezetlen, illetve *coiled-coil* régiók az összehasonlításuk során nagymértékben átfedtek egymással. *H. sapiens*ben (a gerincesekre jellemzően) az UPF2 fehérje N-terminális részén található egy mintegy 70 aminosavból felépülő potenciális CSAH régió, amely a többi faj UPF2-fehérjéjéből hiányzik. Ez a stabil egyszálú alfa-hélix szerkezet a távolságtartásban, dimerizációban játszhat fontos szerepet.

Az EJC-t alkotó fehérjék közül a teljes mértékben rendezetlen MLN51 mind növényekben, mind pedig rovarokban csekély szekvenciális hasonlóságot mutat a humán MLN51-gyel. A másik három EJC-magfehérje, vagyis az Y14, a MAGOH és az EIF4A3 *A. thaliana*-ban és *D. melanogaster*-ben aminosavsorrendjét és szerkezetét tekintve is nagymértékben hasonlítanak a nekik megfelelő humán fehérjékhez.

A példafajokban megtalált SMG-k viszonylag magas rendezettségük ellenére csupán kis hasonlóságot mutatnak a humán SMG fehérjékkel. Az eltérést esetükben inkább szekvenciális különbségek adják, a doménjeik ugyanis nagymértékben hasonlítanak egymásra a különböző fajokban.

Bár az NCBP2 fehérjék hosszúságukban és rendezetlenségük mértékében egyes esetekben eltérnek a humán fehérjétől, doménjeik és szekvenciabeli hasonlóságuk igen nagy. A példafajok NCBP1 fehérjéje bár hosszúságukat, doménjeiket és rendezetlen szakaszaikat tekintve – a rovarok NCBP1 fehérjéjét kivéve – is hasonlítanak, azonban szekvenciálisan nagymértékben eltérnek a humántól.

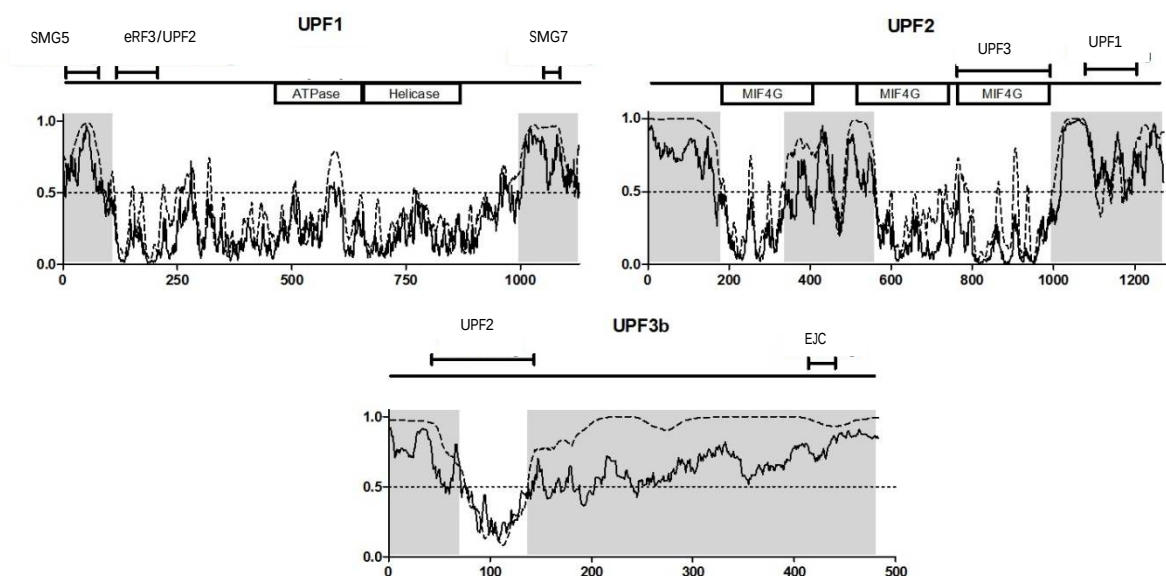
A példafajok citoplazmatikus PABP1 fehérjéi minden tekintetben nagymértékű hasonlóságot mutatnak. A nukleáris PABP2-t nem találtuk meg féregben és élesztőben, a növényi PABP2 pedig mind szekvenciájában, mind pedig szerkezetében eltér a humántól. *D. melanogaster*-ben a fehérje doménjei, valamint az aminosavsorrendje is valamelyest hasonlítanak a humán PABP2-höz.

A növényi és a rovarokra jellemző eRF fehérjék és a WIBG is minden tekintetben nagymértékű hasonlóságot mutat a nekik megfelelő humán fehérjékkel. Az UPF2-höz és EJC-hez kapcsolódó teljes mértékben rendezetlen RNPS1 esetében már kevésbé konzervált a szekvencia.

### **5.3 A rendezetlen szakaszok kötéseinek és szerepének feltárása az emlős NMD-fehérjékben bioinformatikai módszerek segítségével**

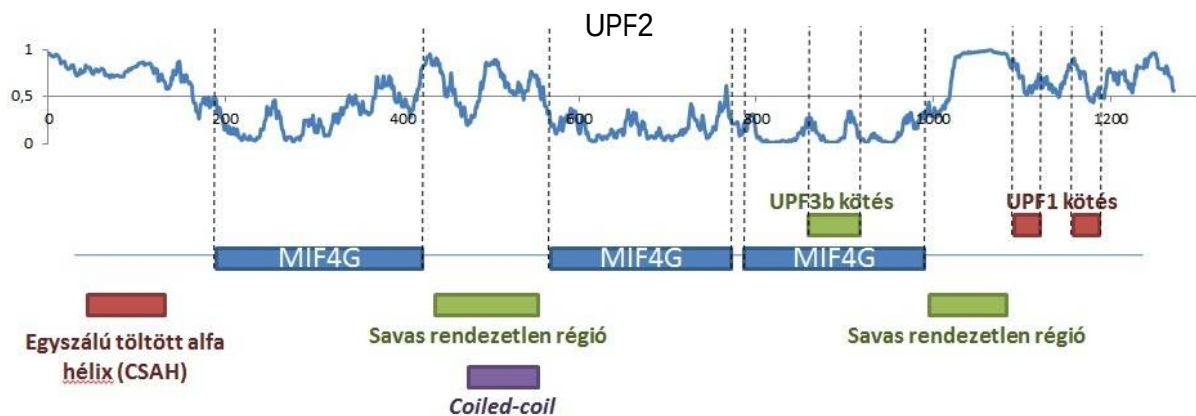
A fenti vizsgálatokat követően az EJC alapú NMD-re, azon belül pedig az emlős NMD-fehérjék rendezetlen szakaszaira fektettük a hangsúlyt. Arra voltunk kíváncsiak, hogy az emlős NMD-ben az egyes fehérjék rendezetlen szakaszai milyen funkciót láthatnak el a folyamatban, illetve hogy pontosan milyen kötésekben vesznek részt (2-7. ábrák). Az emlős NMD folyamatában számos olyan fehérjét találtunk, amelyben viszonylag nagy a rendezetlenség mértéke és amelyben a rendezetlen szakaszok fontos szerepet töltenek be komplexeken belüli, illetve komplexek közötti kötések kialakításában.

Az emlős UPF1 rendezetlen N-terminusával köti SMG5-öt, szintén rendezetlen C-terminális végével pedig SMG7-et [6]. Mindkét SMG egy rendezett szakasszal kapcsolódik hozzá. Az UPF1 ATP-áz és helikáz aktivitással is rendelkezik, mindkét funkcióját rendezett szakaszok látják el a fehérjén belül. Szintén rendezett rész képes kapcsolódni eRF3-hoz, illetve ugyanezen szakasz UPF2-t is köti (2. ábra). A két fehérje kompetál egymással UPF1 kötéseért [16]. Irodalmi adatok alapján UPF1 az NCBP1-et is köti a 419-700-ig tartó szakaszán [17] (az ábrán nincs feltüntetve). A COILS prediktor erre a részre, valamint NCBP1 UPF1-et kötő szakaszára *coiled-coil* régió jelenlétét jósolja mindkét fehérjében.



**2. ábra** Az NMD kulcsfehérjéi, az UPF-ek rendezetlenségi diagramjai. A folytonos vonal az IUPred, a szaggatott pedig a PONDR-VSL2 program alapján készült. A rendezetlen régiók szürke háttérrel vannak feltüntetve. Az emlős UPF1 rendezetlen N-terminusával köti SMG5-öt, szintén rendezetlen C-terminális végével pedig SMG7-et. Rendezett szakaszával képes kapcsolódni eRF3-hoz, illetve ugyanezen szakasz UPF2-t is köti. Az UPF3 nagymértékben rendezetlen fehérje, amely viszonylag rövid rendezett régiójával köti UPF2-t. Az EJC-n belül pedig az Y14-MAGOH komplexet köti rendezetlen C-terminusával.

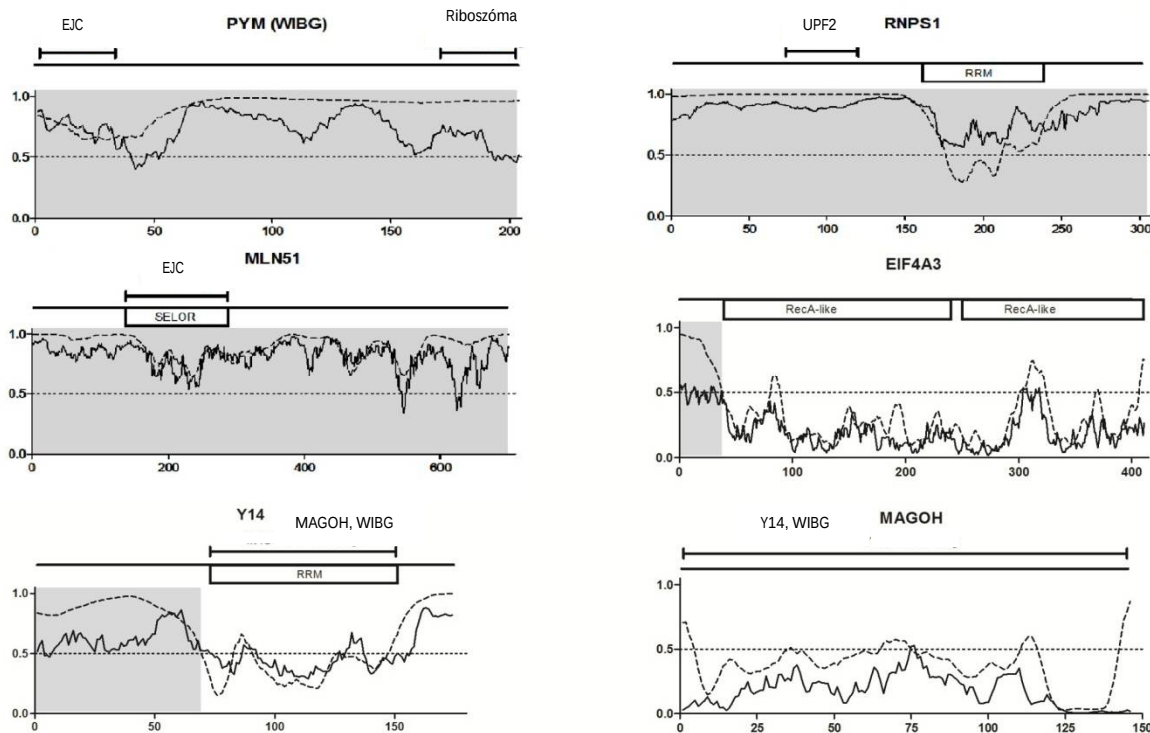
UPF2 szintén rendelkezik hosszabb rendezetlen szakaszokkal, ezenkívül a fehérje N-terminális végére a CSAH szerver CSAH domén jelenlétét jósolja. Három *Middle domain of eukaryotic initiation factor 4G* (MIF4G) domén is található UPF2-ben, amelyekre az alfa-helikális ismétlődések jellemzőek. Az első kettőt egy hosszabb rendezetlen szakasz választja el egymástól. A harmadik MIF4G domén felelős az UPF3-kötésért [18], míg UPF1-et a fehérje a C-terminusán levő, alapvetően rendezetlen régióval köti (2. és 3. ábra). Ez a régió a kötés során egy hurokkal elválasztott alfa-helikális és béta-hajtű szerkezetet vesz fel [16].



**3. ábra** A humán UPF2 fehérje részletes rendezetlenségi diagramja, amely az IUPred segítségével készült. A jóslott rendezetlen régiókat a vonal (0,5-es érték) feletti szakaszok mutatják. A fehérje a 487-559-ig tartó szakaszán a UniProt adatbázis és a COILS szervert is *coiled-coil* tartalmaz.

UPF3 nagymértékben rendezetlen fehérje, amely viszonylag rövid rendezett régiójával köti UPF2-t [18]. Az EJC-n belül pedig az Y14-MAGOH komplexet köti rendezetlen C-terminusával. MAGOH és Y14 rendezett szakaszokkal kötődnek UPF3-hoz (2. ábra). UPF3 még a sejtmagban kapcsolódik az EJC komplexhez, UPF2 pedig a citoplazmában kerül rá a komplexre.

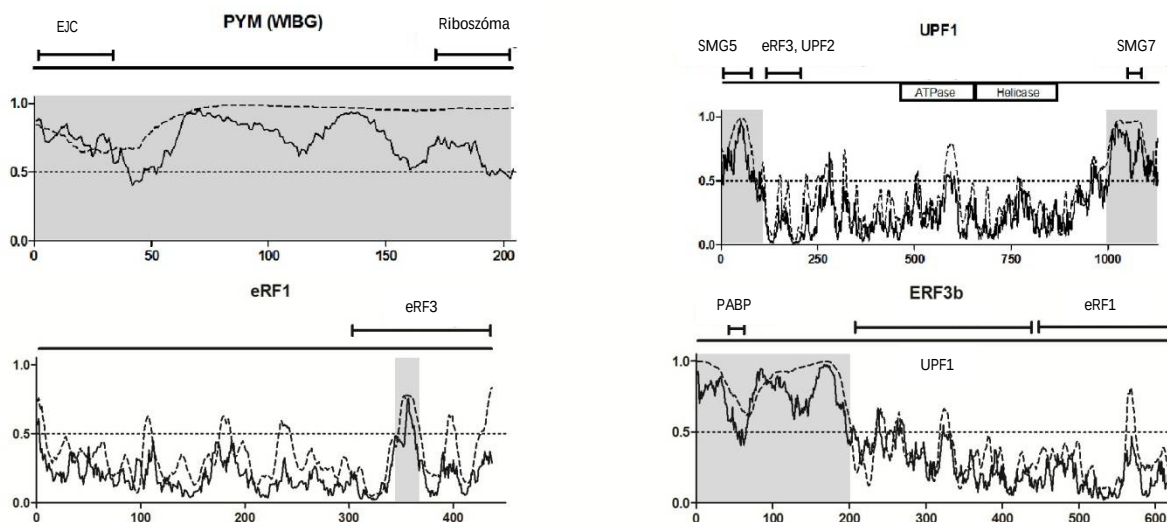
Az EJC egyik magfehérjéje, a MAGOH egy csaknem teljes mértékben rendezett fehérje, míg Y14 egy igen hosszú N-terminális rendezetlen régióval rendelkezik, amit egy terjedelmes rendezett szakasz követ (4. ábra). Ez utóbbi felelős az Y14-gyel, a WIBG-vel és MAGOH-val való kapcsolódásáért egyaránt. A kölcsönhatásban WIBG rendezetlen N-terminális vége vesz részt, miközben szintén rendezetlen C-terminusa a riboszóma kisebbik alegységéhez köt [19] (4. ábra). Alapvetően a riboszómához kapcsolódik. Az ő feladata indukálni az EJC komplex szétszerelődését, amennyiben a riboszóma közel kerül az EJC-hez. Ehhez vélhetően nélkülözhetetlen WIBG nagymértékű rendezetlensége, hiszen így a fehérje könnyebben képes létrehozni az EJC-vel való kapcsolatot és kiváltani a komplex fehérjék disszociációját.



**4. ábra** Az EJC komplex fehérjéinek rendezetlenségi diagramjai. MAGOH egy csaknem teljes mértékben rendezett fehérje, míg Y14 egy igen hosszú N-terminális rendezetlen régióval rendelkezik. A rendezetlen MLN51 az egész EJC komplexet áthidalja.

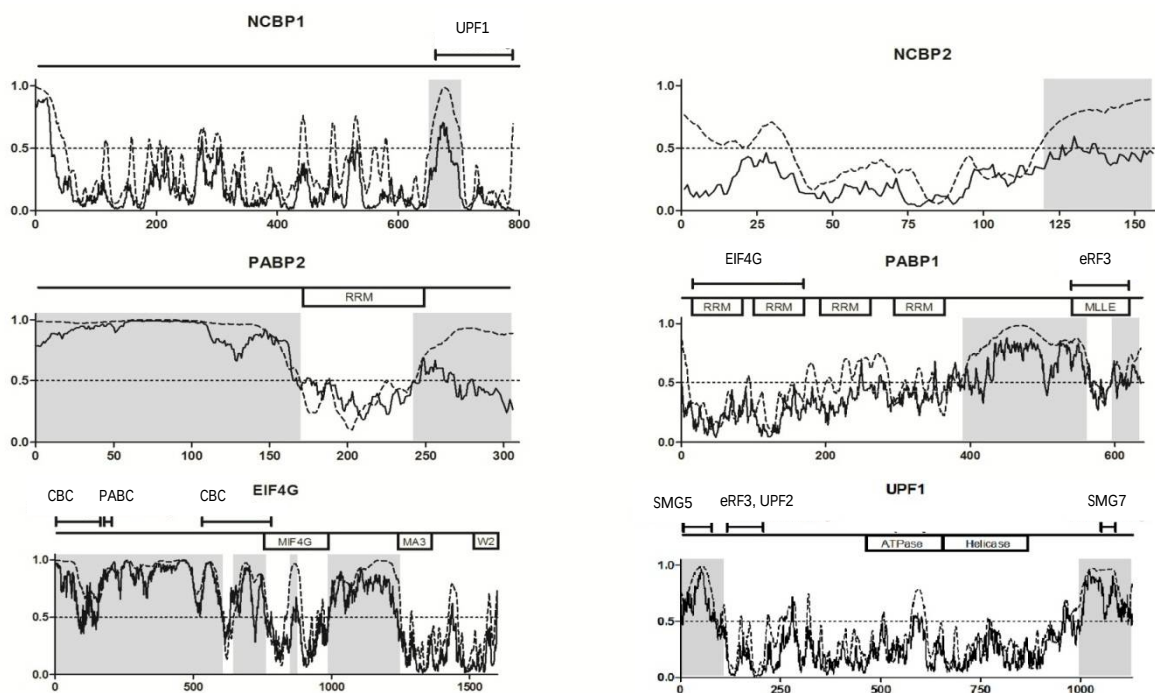
Vélhetően az EJC „állványzataként” funkcionál a szinte teljesen rendezetlen MLN51, amely az egész komplexet áthidalva tartja össze a fehérjéket [20] (4. ábra), míg az eIF4A3 nagyrészt rendezett és csak egy igen rövid rendezetlen régiója van az N-terminális végén. A csaknem teljesen rendezetlen RNPS1 az irodalom szerint UPF2-höz kapcsolódik és feltehetően UPF2 EJC-hez való kötődését segíti elő eddig ismeretlen módon [21] (4. ábra).

A riboszómához asszociált eRF1 nagyrészt rendezett, az eRF3-kötésben azonban rövid C-terminális rendezetlen régiója is részt vesz a szerkezettel rendelkező szakaszok mellett. Az eRF3 középső rendezett része UPF1-et, míg rendezett C-terminális vége eRF1-et képes kötni. Rendezetlen N-terminális vége pedig a PABP-vel képes kölcsönhatni (5. ábra). Ez a kölcsönhatás rendkívül fontos az NMD megakadályozásában, így az mRNS transzlációjának terminációjában is. Ha ugyanis eRF3 és PABP1 között nem jön létre kapcsolat a korai STOP kodon és így a nagy távolság következtében, UPF1 köthet eRF3-hoz. Amennyiben egy EJC is található a riboszómától *downstream*, létrejön az UPF1-UPF2 kapcsolat és elindul az NMD, illetve a mRNS degradációjának folyamata. Az eRF3-PABP1 kapcsolatban tehát igen fontos szerepet kapnak a nagy távolságokat is áthidalni képes rendezetlen szakaszok (5., 6. és 8. ábrák). Az EJC disszociációját kiváltó WIBG rendezetlen C-terminálisán köti a riboszóma komplexet.



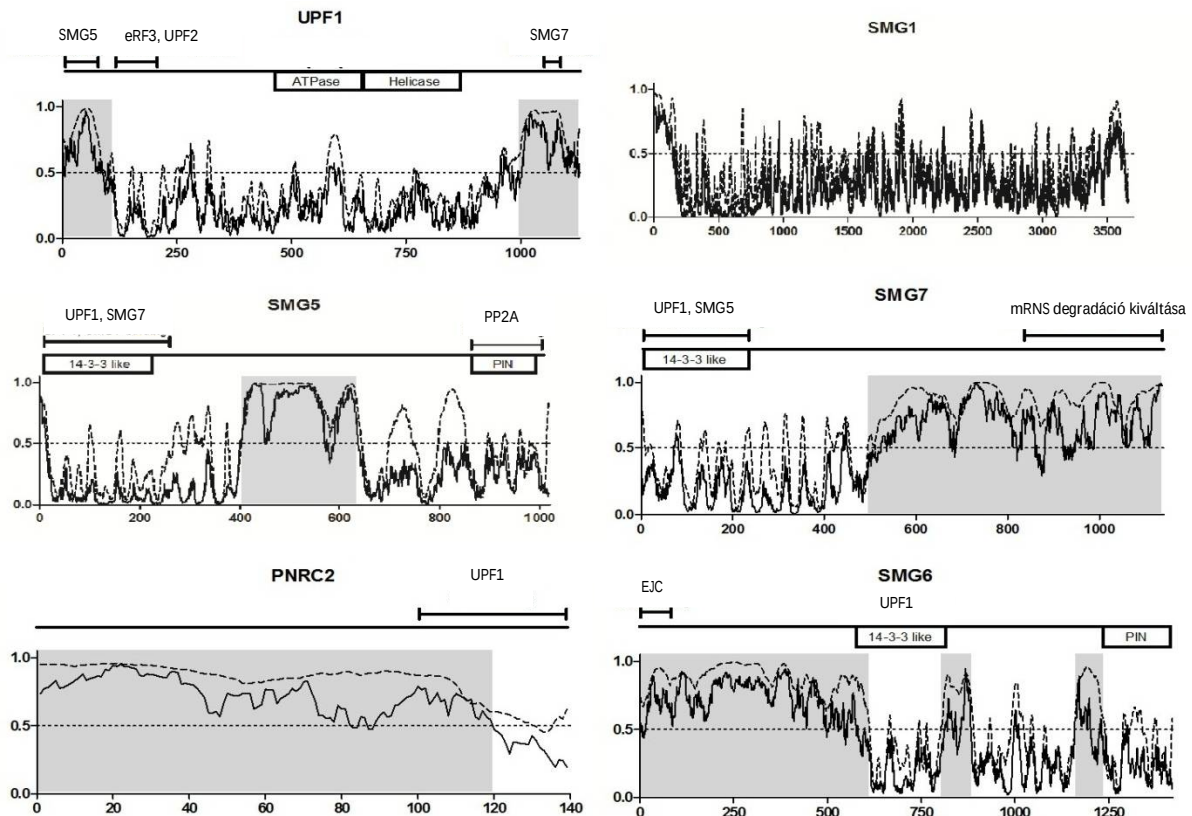
**5. ábra** A riboszóma komplex fehérjéinek rendezetlenségi diagramjai. A riboszómához asszociált eRF1 nagyrészt rendezett, az eRF3-kötésben azonban rövid C-terminális rendezetlen régiója is részt vesz. eRF3 középső rendezett része UPF1-et, míg C-terminális vége eRF1-et képes kötni. Rendezetlen N-terminális vége pedig a PABP-vel képes kölcsönhatni. WIBG rendezetlen C-terminusa a riboszóma kisebbik alegységéhez köt, míg szintén rendezetlen N-terminusa EJC-vel képes kapcsolódni.

A citoplazmatikus PABP1 N-terminálisán négy RNS-kötő RRM motívum található, ebből az első kettő egyben az eIF4G kötéséért is felelős. Az RRM-eket és a szintén rendezett eRF3-kötő régiót egy hosszabb rendezetlen szakasz választja el egymástól [23] (6. ábra). Ez vélhetően fontos szerepet játszik abban, hogy eRF3 és PABP1 a köztük levő esetlegesen nagy távolság ellenére is kapcsolódhasson egymással. EIF4G-ben is számos rendezetlen domén található, amelynek fontos szerepe lehet a CBC-EIF4G-PABP komplex összetartásában. Ilyen például a két CBC-kötő hely a fehérjén, a köztük található PABP-t kötő régió, valamint a MIF4G domén egy része (6. ábra). NCBP2 nagyrészt rendezett, de C-terminális végén egy hosszabb rendezetlen régió is található [23]. NCBP1 szintén nagyon rendezett, de ezt a rendezettséget 664-790 között egy rendezetlen hurok töri meg [17, 24] (6. ábra), amely feltehetően az UPF1-kötésért felelős. A COILS server *coiled-coil* szerkezet kialakítására hajlamos régió jelenlétét jósolta erre a szakaszra.



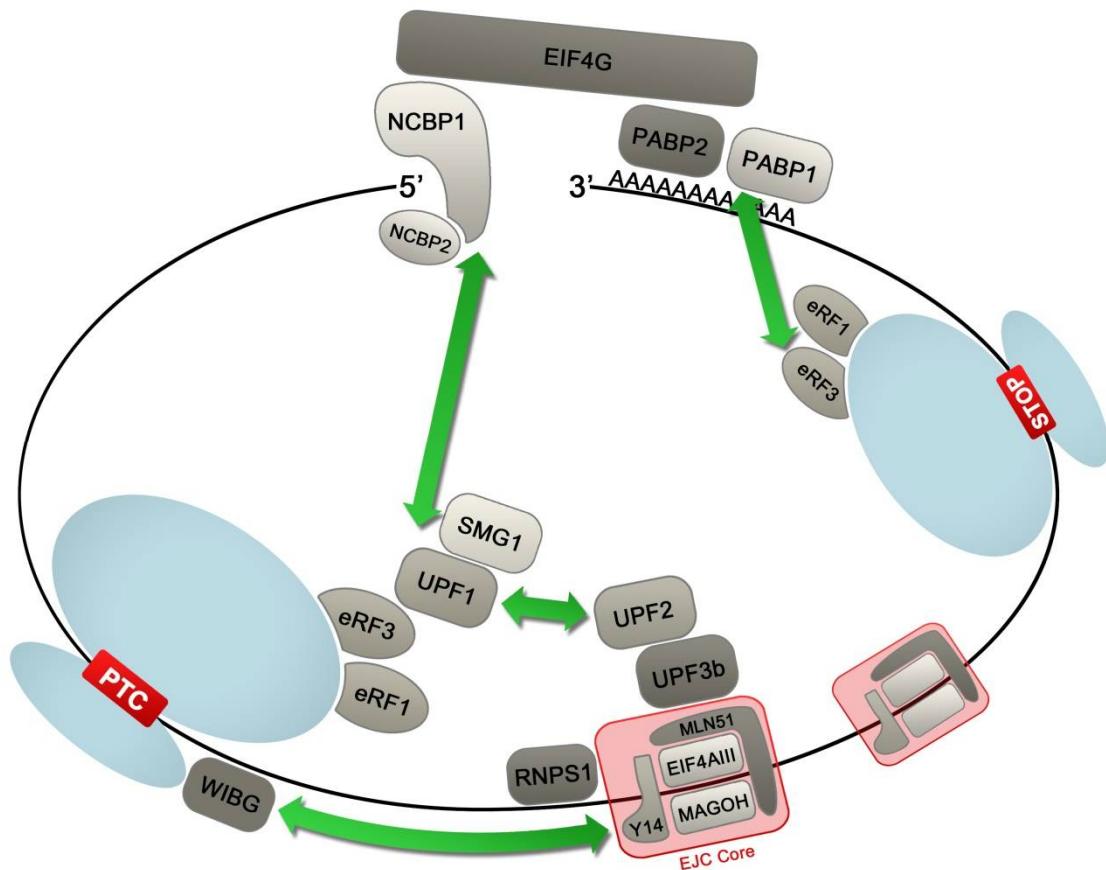
**6. ábra** A CBC-eIF4G-PABP komplex fehérjének rendezetlenségi diagramjai. EIF4G-ben számos rendezetlen domén található, ilyen például a két CBC-kötő hely a fehérjén, a köztük található PABP-t kötő régió, valamint a MIF4G domén egy része. NCBP1 szintén nagyon rendezett, de ezt a rendezettséget 664-790 között egy rendezetlen szakasz töri meg, amely az UPF1-kötésért felelős. PABP1 N-terminálisán az RRM-eket és a szintén rendezett eRF3-kötő régiót egy hosszabb rendezetlen szakasz választja el egymástól.

Bár az SMG1 nagymértékben rendezett, a többi UPF1-hez kapcsolódó SMG-k között találni hosszabb rendezetlen szakaszokat. Az SMG5 rendezett N-terminusával kapcsolódik UPF1-hez és SMG7-hez, míg szintén rendezett C-terminusán levő *PIN-like* doménjével a PP2A foszfatáz komplexet köti (7. ábra). A két domént egy hosszabb rendezetlen régió kapcsolja össze. SMG7 rendezett N-terminális végével UPF1-et és SMG5-öt köti, rendezetlen második felével pedig feltehetően a decapping és deadeniláció folyamatát váltja ki az mRNA degradációban. SMG6 szintén rendelkezik hosszabb-rövidebb rendezetlen szakaszokkal, UPF1-hez egy rendezett régiójával köt [25], C-terminális végén pedig SMG5-höz hasonlóan ő is rendelkezik *PIN-like* doménnel, amely SMG5-tel ellentétben nem csupán foszfodiészteráz, de endonukleáz aktivitással is bír [6] (7. ábra). A szintén nagymértékben rendezetlen PNRC2 feladata az NMD-t kiváltó korai komplex összekötése a *decapping* rendszerrel [22] (7. ábra).



**7. ábra** Az SMG kinázok és a degradáció kiváltását elősegítő PNRC2 rendezetlenségi diagramjai. SMG5, 6 és 7 is tartalmaz hosszabb rendezetlen szakaszokat, amik doméneket választanak el egymástól, vagy maguk is kötésben vesznek részt. PNRC2 csaknem teljes mértékben szerkezet nélküli fehérje.

Látható tehát, hogy az intronalapú NMD-ben három nagy fehérjekomplex és számos, a komplexekhez asszociált, valamint köztük hídként funkcionáló fehérje vesz részt. Ezen fehérjemolekulák jelentős része tartalmaz hosszabb, vélhetően igen fontos működést ellátó rendezetlen szakaszokat. UPF1 például mindkét rendezetlen végén SMG kinázokhoz képes kötni, UPF2 doménjeit rendezetlen szakaszok választják el egymástól, miközben UPF1-gyel való kötéséért is alapvetően egy szerkezet nélküli szakasz felelős, EIF4G és MLN51 pedig „állványzatként” tartják össze a hozzájuk kapcsolódó fehérjéket, stb. Az NMD ellen, így az mRNS translációjának végbemenetele érdekében kapcsolódik PABP1 és eRF3, amelyek kötését rendezetlen szakaszuk segíti elő. WIBG, az EJC-et szétszerelő, így szintén az NMD ellen dolgozó fehérje, valamint RNPS1 teljesen rendezetlen fehérjeként vesznek részt a komplexek kommunikációjában. Az alábbi összefoglaló ábra a nagy hidrodinamikai sugárral rendelkező rendezetlen fehérjék, illetve fehérjeszakaszok szerepét igyekszik bemutatni az emlős NMD-ben, azon belül is a komplexek közötti kommunikáció létrehozásában:



**8. ábra** Az EJC alapú NMD vázlatja és emlősökben. Az NMD target (PTC-t tartalmazó) és normál mRNS-en a komplexek között létrejövő ismert kölcsönhatásokat zöld nyilak jelölik. A fehérjék a szürke különböző árnyalataival vannak feltüntetve, amely a rendezetlenség mértékét jelöli a rendezett világosszürkétől a teljesen rendezetlen sötétszürkéig.

## 6. Összefoglalás

A rendezetlenség fontos szerepet játszik az NMD folyamatában részt vevő fehérjék kapcsolódásában. A transzlálódó mRNS-hez kötődő komplexek között gyakran igen nagy távolságok áthidalására van szükség (a humán mRNS mérettartománya: 237-101520 bp, az exonok hosszának mérettartománya 2-22753 bp), amihez a komplexekből „kinyúlni képes” hosszú rendezetlen szakaszok nélkülözhetetlenek látszanak. Munkánk során igyekeztünk átfogó képet adni a rendezetlenség és az emlős NMD folyamatában szerepet játszó komplexek kapcsolódásáról, különös tekintettel kiemelve a hídként szerepet játszó fehérjék/fehérjeszakaszok szerepét (8. ábra).

Látható, hogy a rendezetlenség mértéke viszonylag magas az NMD folyamatában részt vevő fehérjék körében és a komplexek közötti kölcsönhatásban főként a rendezetlen fehérjék

vesznek részt. Ezt bizonyítja a 2-8. ábrákon feltüntetett számos, nagyrészt rendezetlen fehérje, amelyek komplexek összetartásában (pl.: MLN51, EIF4G) és egymásra találásában (pl.: UPF2, UPF3, WIBG) játszanak fontos szerepet.

Elmondható továbbá, hogy a többi példafajt tekintve az alacsonyabb rendű élőlények megfelelő fehérjei hasonló, vagy a legtöbb esetben valamivel kisebb mértékű rendezetlenséggel jellemezhetőek, de a szerkezet nélküli szakaszok esetükben is megtalálhatóak és vélhetően fontos szerepet játszanak az NMD folyamatában.

Munkánk során megállapítottuk, hogy az NMD rendkívül bonyolult és számos fehérje jelenlétét igénylő folyamatában a rendezetlenség kulcsfontosságú jelenség, amelynek léte megoldást kínál olyan biológiai problémákra is, mint például az mRNS különböző részei között létrejövő és nagy távolságokat áthidaló sokfehérjés komplexek jelenlétének és gyors kapcsolódásának problémája.

## 7. Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőmnek, dr. Kalmár Lajosnak, valamint dr. Tompa Péternek a TDK munkám elkészítéséhez nyújtott szakmai segítségéért, mindenkori támogatásáért és biztatásáért. Továbbá szeretnék köszönetet mondani dr. Silhavy Dánielnek is a témában nyújtott szakmai segítségéért, tanácsaiért.

## 8. Irodalomjegyzék

1. Chang, Y.F., J.S. Imam, and M.F. Wilkinson, *The nonsense-mediated decay RNA surveillance pathway*. Annu Rev Biochem, 2007. **76**: p. 51-74.
2. Brogna, S. and J. Wen, *Nonsense-mediated mRNA decay (NMD) mechanisms*. Nat Struct Mol Biol, 2009. **16**(2): p. 107-13.
3. Gonzalez, C.I., et al., *Nonsense-mediated mRNA decay in Saccharomyces cerevisiae*. Gene, 2001. **274**(1-2): p. 15-25.
4. Kerenyi, Z., et al., *Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay*. EMBO J, 2008. **27**(11): p. 1585-95.
5. Gehring, N.H., et al., *Disassembly of exon junction complexes by PYM*. Cell, 2009. **137**(3): p. 536-48.
6. Nicholson, P. and O. Muhlemann, *Cutting the nonsense: the degradation of PTC-containing mRNAs*. Biochem Soc Trans, 2010. **38**(6): p. 1615-20.
7. Apcher, S., et al., *Major source of antigenic peptides for the MHC class I pathway is produced during the pioneer round of mRNA translation*. Proc Natl Acad Sci U S A, 2011. **108**(28): p. 11572-7.
8. Shoemaker, B.A., J.J. Portman, and P.G. Wolynes, *Speeding molecular recognition by using the folding funnel: the fly-casting mechanism*. Proc Natl Acad Sci U S A, 2000. **97**(16): p. 8868-73.

9. Balazs, A., et al., *High levels of structural disorder in scaffold proteins as exemplified by a novel neuronal protein, CASK-interactive protein1*. FEBS J, 2009. **276**(14): p. 3744-56.
10. Suveges, D., et al., *Charged single alpha-helix: a versatile protein structural motif*. Proteins, 2009. **74**(4): p. 905-16.
11. Dosztanyi, Z., et al., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content*. Bioinformatics, 2005. **21**(16): p. 3433-4.
12. Peng, K., et al., *Length-dependent prediction of protein intrinsic disorder*. BMC Bioinformatics, 2006. **7**: p. 208.
13. Lupas, A., M. Van Dyke, and J. Stock, *Predicting coiled coils from protein sequences*. Science, 1991. **252**(5010): p. 1162-4.
14. Dosztanyi, Z., B. Meszaros, and I. Simon, *ANCHOR: web server for predicting protein binding regions in disordered proteins*. Bioinformatics, 2009. **25**(20): p. 2745-6.
15. Gatfield, D., et al., *Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways*. EMBO J, 2003. **22**(15): p. 3960-70.
16. Clerici, M., et al., *Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2*. EMBO J, 2009.
17. Hwang, J., et al., *UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps*. Mol Cell, 2010. **39**(3): p. 396-409.
18. Kadlec, J., E. Izaurralde, and S. Cusack, *The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3*. Nat Struct Mol Biol, 2004. **11**(4): p. 330-7.
19. Bono, F., et al., *Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex*. EMBO Rep, 2004. **5**(3): p. 304-10.
20. Andersen, C.B., et al., *Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA*. Science, 2006. **313**(5795): p. 1968-72.
21. Gehring, N.H., et al., *Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements*. Mol Cell, 2005. **20**(1): p. 65-75.
22. Cho, H., K.M. Kim, and Y.K. Kim, *Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex*. Mol Cell, 2009. **33**(1): p. 75-86.
23. Mangus, D.A., M.C. Evans, and A. Jacobson, *Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression*. Genome Biol, 2003. **4**(7): p. 223.
24. Mazza, C., et al., *Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex*. EMBO J, 2002. **21**(20): p. 5548-57.
25. Chiu, S.Y., et al., *Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1*. RNA, 2003. **9**(1): p. 77-87.

## 9. Függelék

|                 | Faj<br>(fehérje UniProt<br>azonosítója) | Fehérje<br>hossza<br>(aminosav) | Jószolt<br><i>coiled-<br/>coil</i>                                                     | Rendezetlen<br>régiók<br>(>=25)                | Rendezetlenség<br>mértéke |
|-----------------|-----------------------------------------|---------------------------------|----------------------------------------------------------------------------------------|------------------------------------------------|---------------------------|
| <b>UPF-ek</b>   |                                         |                                 |                                                                                        |                                                |                           |
| UPF1/RENT1/NAM7 |                                         |                                 |                                                                                        |                                                |                           |
|                 | <i>H. sapiens</i><br>(Q92900)           | 1129                            | 580-<br>620?                                                                           | 6-77, 1003-<br>1126                            | 0,23                      |
|                 | <i>D. melanogaster</i><br>(Q9VYS3)      | 1180                            | 570-600                                                                                | 1060-1100                                      | 0,22                      |
|                 | <i>A. thaliana</i><br>(Q9FJR0)          | 1254                            | 590-<br>620,<br>328-348                                                                | 0-110, 960-<br>1240                            | 0,31                      |
|                 | <i>C. elegans</i><br>(O76512)           | 1069                            | 1035-<br>1065                                                                          | 0-86, 948-<br>1069                             | 0,24                      |
|                 | <i>S. cerevisiae</i><br>(P30771)        | 971                             | 930-970                                                                                | 1-30                                           | 0,07                      |
| UPF2/RENT2/NMD2 |                                         |                                 |                                                                                        |                                                |                           |
|                 | <i>H. sapiens</i><br>(Q9HAU5)           | 1272                            | 50-150,<br>375-<br>420?,<br>490-<br>560,<br>986-<br>1006,<br>1200-<br>1240             | 0-160, 420-<br>460, 488-<br>556, 1020-<br>1172 | 0,43                      |
|                 | <i>D. melanogaster</i><br>(Q9W3L1)      | 1241                            | 20-80,<br>1170-<br>1200                                                                | 0-40, 420-<br>475, 970-<br>1130                | 0,36                      |
|                 | <i>A. thaliana</i><br>(O80955)?         | 1029                            | 80-150,<br>250-300                                                                     | 150-300,<br>740-840,<br>980-1020               | 0,43                      |
|                 | <i>C. elegans</i><br>(Q95XD2)           | 1142                            | 20-50,<br>274-<br>301,<br>305-<br>335,<br>593-<br>616,<br>920-<br>980,<br>1060<br>1130 | 372-421,<br>916-997,<br>1008-1142              | 0,32                      |
|                 | <i>S. cerevisiae</i><br>(P38798)        | 1089                            | 243-<br>270,                                                                           | 819-994                                        | 0,2                       |

|                                          |                                    |     |                  |                         |      |
|------------------------------------------|------------------------------------|-----|------------------|-------------------------|------|
|                                          |                                    |     | 510-550, 948-984 |                         |      |
| UPF3B/RENT3B/<br>UPF3X                   |                                    |     |                  |                         |      |
|                                          | <i>H. sapiens</i><br>(Q9BZI7)      | 483 | 200-253, 345-410 | 190-492                 | 0,76 |
|                                          | <i>D. melanogaster</i><br>(Q9W1H3) | 492 | 175-225          | 0-55, 247-483           | 0,73 |
|                                          | <i>A. thaliana</i><br>(Q9FVW4)?    | 482 | -                | 160-480                 | 0,74 |
|                                          | <i>C. elegans</i><br>(Q8WQF3)      | 368 | 170-210, 220-275 | 178-368                 | 0,55 |
|                                          | <i>S. cerevisiae</i><br>(P48412)   | 387 | 270-292          | 0-74, 263-321, 336-387  | 0,5  |
| <b>EJC fő fehérjei</b>                   |                                    |     |                  |                         |      |
| EIF4A3                                   | <i>H. sapiens</i><br>(P38919)      | 411 | -                | -                       | 0,05 |
|                                          | <i>D. melanogaster</i><br>(Q9VHS8) | 399 | -                | -                       | 0,03 |
|                                          | <i>A. thaliana</i><br>(Q9CAI7)     | 414 | -                | -                       | 0,04 |
| MLN51/BTZ/CASC3                          |                                    |     |                  |                         |      |
|                                          | <i>H. sapiens</i><br>(O15234)      | 703 | 90-130           | 0-703                   | 0,98 |
|                                          | <i>D. melanogaster</i><br>(C8VV40) | 782 | 510-560          | 0-557, 563-632, 644-782 | 0,98 |
|                                          | <i>A. thaliana</i><br>(Q93ZJ9)     | 605 | -                | 0-605                   | 0,91 |
| Y14/RBM8A                                |                                    |     |                  |                         |      |
|                                          | <i>H. sapiens</i><br>(Q9Y5S9)      | 174 | -                | 7-73                    | 0,6  |
|                                          | <i>D. melanogaster</i><br>(Q9V535) | 165 | 43-64, 160-165   | 34-65                   | 0,32 |
|                                          | <i>A. thaliana</i><br>(F4I9J7)     | 202 | -                | 0-80                    | 0,56 |
| MAGOH/MGN                                |                                    |     |                  |                         |      |
|                                          | <i>H. sapiens</i><br>(P61326)      | 146 | -                | -                       | 0,01 |
|                                          | <i>D. melanogaster</i><br>(P49028) | 147 | -                | -                       | 0,01 |
|                                          | <i>A. thaliana</i><br>(O23676)     | 150 | -                | -                       | 0,13 |
| <b>Riboszóma-asszociált<br/>fehérjék</b> |                                    |     |                  |                         |      |
| eRF1                                     |                                    |     |                  |                         |      |
|                                          | <i>H. sapiens</i>                  | 437 | -                | -                       | 0,25 |

|                          |                                    |     |                         |                   |       |
|--------------------------|------------------------------------|-----|-------------------------|-------------------|-------|
|                          | (P62495)                           |     |                         |                   |       |
|                          | <i>D. melanogaster</i><br>(Q9VPH7) | 438 | -                       | -                 | 0,05  |
|                          | <i>A. thaliana</i><br>(Q39097)     | 436 | -                       | -                 | 0,039 |
| eRF3                     |                                    |     |                         |                   |       |
|                          | <i>H. sapiens</i><br>(Q8IYD1)      | 628 | -                       | 0-52, 63-195      | 0,33  |
| <b>PABP-CBC kompleks</b> |                                    |     |                         |                   |       |
| PABP2/PAB2/<br>PABPN1    |                                    |     |                         |                   |       |
|                          | <i>H. sapiens</i><br>(Q86U42)      | 306 | 110-150                 | 0-165             | 0,62  |
|                          | <i>D. melanogaster</i><br>(Q7KNF2) | 224 | 14-34?,<br>45-75        | 0-80              | 0,45  |
|                          | <i>A. thaliana</i><br>(P42731)     | 629 | 280-330                 | 420-540           | 0,37  |
| PABP1/PAB1/<br>PABPC1    |                                    |     |                         |                   |       |
|                          | <i>H. sapiens</i><br>(P11940)      | 636 | 250-300                 | 421-566           | 0,32  |
|                          | <i>D. melanogaster</i><br>(P21187) | 634 | 256-<br>282,<br>600-630 | 447-560           | 0,25  |
|                          | <i>A. thaliana</i><br>(Q9FXA2)?    | 671 | 289-309                 | 409-605           | 0,38  |
|                          | <i>C. elegans</i><br>(Q9U302)      | 646 | 289-309                 | 417-580           | 0,35  |
|                          | <i>S. cerevisiae</i><br>(P04147)   | 577 | 294-<br>314,<br>556-576 | 0-37, 447-<br>577 | 0,3   |
| NCBP1/CBP80              |                                    |     |                         |                   |       |
|                          | <i>H. sapiens</i><br>(Q09161)      | 790 | 645-<br>676,<br>685-712 | 0-26              | 0,07  |
|                          | <i>D. melanogaster</i><br>(Q7K4N3) | 800 | 188-<br>208,<br>640-690 | 660-700           | 0,08  |
|                          | <i>A. thaliana</i><br>(Q9SIU2)     | 848 | 630-690                 | -                 | 0,08  |
|                          | <i>C. elegans</i><br>(O01763)      | 798 | 633-<br>674,<br>677-711 | 660-693           | 0,08  |
|                          | <i>S. cerevisiae</i><br>(P34160)   | 861 | 603-623                 | -                 | 0,02  |
| NCBP2/CBP20              |                                    |     |                         |                   |       |
|                          | <i>H. sapiens</i><br>(P52298)      | 156 | -                       | -                 | 0,04  |
|                          | <i>D. melanogaster</i>             | 154 | -                       | -                 | 0,03  |

|                                |                                    |      |                                           |                                           |      |
|--------------------------------|------------------------------------|------|-------------------------------------------|-------------------------------------------|------|
|                                | (Q9V3L6)                           |      |                                           |                                           |      |
|                                | <i>A. thaliana</i><br>(Q9XFD1)     | 257  | -                                         | 160-257                                   | 0,51 |
|                                | <i>C. elegans</i><br>(Q93594)      | 154  | -                                         | -                                         | 0,11 |
|                                | <i>S. cerevisiae</i><br>(Q08920)   | 208  | -                                         | 168-208                                   | 0,25 |
| EIF4G1                         |                                    |      |                                           |                                           |      |
|                                | <i>H. sapiens</i><br>(Q04637)      | 1599 | 450-483,860-893                           | 0-84, 107-605, 633-715, 730-759, 989-1238 | 0,6  |
|                                | <i>D. melanogaster</i><br>(O61380) | 1666 | -                                         | 0-750, 1000-1286                          | 0,59 |
|                                | <i>A. thaliana</i><br>(Q76E23)     | 1727 | 1176-1215                                 | 0-913, 956-1102, 1189-1216, 1320-1540     | 0,75 |
|                                | <i>S. cerevisiae</i><br>(P39935)   | 952  | 192-230, 301-321                          | 0-416, 473-605, 849-952                   | 0,73 |
| <b>EJC-asszociált fehérjék</b> |                                    |      |                                           |                                           |      |
| RNPS1                          |                                    |      |                                           |                                           |      |
|                                | <i>H. sapiens</i><br>(Q15287)      | 305  | -                                         | 0-305                                     | 1    |
|                                | <i>D. melanogaster</i><br>(Q9VHC0) | 374  | -                                         | 0-135, 260-374                            | 0,96 |
|                                | <i>A. thaliana</i><br>(Q9SEE9)?    | 414  | -                                         | 0-100, 180-414                            | 0,95 |
| WIBG/PYM                       |                                    |      |                                           |                                           |      |
|                                | <i>H. sapiens</i><br>(Q9BRP8)      | 204  | 80-120, 147-177                           | 0-204                                     | 0,93 |
|                                | <i>D. melanogaster</i><br>(P82804) | 207  | 60-90, 150-180                            | 2-207                                     | 0,99 |
|                                | <i>A. thaliana</i><br>(Q9LPZ4)     | 204  | 80-120, 150-170, 182-200                  | 0-207                                     | 0,98 |
| <b>SMG fehérjék</b>            |                                    |      |                                           |                                           |      |
| SMG1                           | <i>H. sapiens</i><br>(Q96Q15)      | 3657 | 189-209, 1970-1997, 2500-2520?, 3232-3252 | 0-140, 3517-3597                          | 0,1  |

|      |                                    |      |                                                               |                       |      |
|------|------------------------------------|------|---------------------------------------------------------------|-----------------------|------|
|      | <i>D. melanogaster</i><br>(Q70PP2) | 3218 | 1118-<br>1138,<br>2282-<br>2302,<br>2876-<br>2903             | 0-60                  | 0,05 |
|      | <i>C. elegans</i><br>(O01510)      | 2322 | 245-<br>285,<br>1540-<br>1570<br>,2257-<br>2277               | -                     | 0,03 |
| SMG5 |                                    |      |                                                               |                       |      |
|      | <i>H. sapiens</i><br>(Q9UPR3)      | 1016 | 798-841                                                       | 400-637,<br>1037-1513 | 0,23 |
|      | <i>D. melanogaster</i><br>(Q9V414) | 1177 | 25-75,<br>420-<br>450,<br>620-<br>650,<br>1040-<br>1080       | 460-660,<br>860-960   | 0,25 |
|      | <i>C. elegans</i><br>(Q94994)      | 549  | 395-<br>425,<br>480-525                                       | -                     | 0,01 |
| SMG6 |                                    |      |                                                               |                       |      |
|      | <i>H. sapiens</i><br>(Q86US8)      | 1419 | 150-<br>200,<br>570-<br>600,<br>803-<br>823,<br>1202-<br>1222 | 16-497                | 0,47 |
|      | <i>D. melanogaster</i><br>(Q9VBZ6) | 948  | 362-382                                                       | 40-90, 110-<br>150    | 0,16 |
|      | <i>C. elegans</i><br>(Q9BL68)      | 1241 | 263-<br>297,<br>333-<br>353,<br>450-500                       | 0-491                 | 0,44 |
| SMG7 |                                    |      |                                                               |                       |      |
|      | <i>H. sapiens</i><br>(Q92540)      | 1137 | -                                                             | 537-819,<br>988-1137  | 0,51 |
|      | <i>A. thaliana</i><br>(A9QM73)     | 1059 | -                                                             | 760-1059              | 0,29 |
|      | <i>C. elegans</i><br>(Q9Y053)      | 458  | -                                                             | 384-458               | 0,17 |
| SMG8 |                                    |      |                                                               |                       |      |
|      | <i>H. sapiens</i>                  | 991  | -                                                             | 16-46, 650-           | 0,25 |

|       |                                    |     |                         |                    |      |
|-------|------------------------------------|-----|-------------------------|--------------------|------|
|       | (Q8ND04)                           |     |                         | 740                |      |
|       | <i>D. melanogaster</i><br>(Q9VKQ6) | 944 | 570-600                 | 650-680            | 0,15 |
|       | <i>C. elegans</i><br>(O62301)      | 873 | 200-<br>230,<br>510-540 | 520-610            | 0,17 |
| SMG9  |                                    |     |                         |                    |      |
|       | <i>H. sapiens</i><br>(Q9H0W8)      | 520 | -                       | 0-180              | 0,42 |
|       | <i>D. melanogaster</i><br>(O46080) | 487 | 42-70                   | 0-115, 138-<br>165 | 0,3  |
|       | <i>C. elegans</i><br>(Q9XWJ1)      | 385 | 295-340                 | 0-40               | 0,15 |
| PNRC2 |                                    |     |                         |                    |      |
|       | <i>H. sapiens</i><br>(Q9NPJ4)      | 139 | 15-45                   | 0-120              | 0,83 |

## Nyilatkozat a pályázó egyéni hozzájárulásáról a pályaműben leírt eredményekhez

Alulírott Ács Veronika a dolgozatban leírt munkafolyamatok közül a:

- NMD-ben részt vevő fehérjék kigyűjtését a témában megjelent publikációk és a UniProt adatbázis alapján
- Homológok keresését
- Rendezetlenségi és *coiled-coil* predikciókat
- NMD-fehérjék szekvencia- és szerkezetelemzését, valamint összehasonlítását a példafajoknál
- Rendezetlen szakaszok és funkciók összevetését humán NMD-fehérjékben, amelyben a Pfam és UniProt adatbázisokat vettem alapul
- Az ANCHOR predikciókat

végeztem el.



aláírás

Budapest, 2012. december 21.