



Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek

KARLO SKUBE I INGRID TOMLJANOVIĆ

**EKSPRESIJA I IDENTIFIKACIJA NOVIH PROTEINA
DJELOMIČNO NEUREĐENE TERCIJARNE STRUKTURE (IDP) IZ
BILJKE UROČNJAK *ARABIDOPSIS THALIANA* NAKON
IZLAGANJA ABIOTIČKOM STRESU**

Zagreb, 2015.

Ovaj rad izrađen je na Zavodu za molekularnu biologiju Biološkog odsjeka Prirodoslovno-matematičkog fakulteta pod vodstvom doc. dr. sc. Dubravka Pavokovića i predan je na natječaj za dodjelu Rektorove nagrade u akademskoj godini 2014. / 2015.

SADRŽAJ

1. UVOD
 - 1.1. Struktura globularnih proteina
 - 1.2. Nativno neuređeni proteini (IDP)
 - 1.2.1. Važnost i otkriće IDP
 - 1.2.2. Bioinformatički alati za analizu IDP
 - 1.2.3. Eksperimentalne metode istraživanja IDP
 - 1.2.4. Podjela IDP prema mehanizmu djelovanja i staničnoj ulozi
 - 1.3. Abiotski stres u biljaka
 - 1.4. IDP u biljkama pri normalnim i stresnim uvjetima
2. CILJ ISTRAŽIVANJA
3. MATERIJALI I METODE
 - 3.1. Identifikacija IDP u proteomu *A. thaliana*
 - 3.1.1. Priprema biljnog materijala
 - 3.1.2. Izazivanje abiotskog stresa hladnoćom
 - 3.1.3. Izolacija IDP
 - 3.1.4. Ukoncentriravanje proteina membranskom ultrafiltracijom
 - 3.1.5. Određivanje koncentracije proteina
 - 3.1.6. Dvodimenzionalna poliakrilamidna gel elektroforeza (2D-PAGE)
 - 3.1.7. Bojanje proteina srebrom
 - 3.1.8. Denzitometrijska analiza ekspresije proteina alatom GelAnalyzer 2010a
 - 3.1.9. Masena spektrometrija proteina
 - 3.2. Bioinformatička karakterizacija identificiranih proteina
 - 3.2.1. Identifikacija očuvanih domena (*engl.* conserved domain search)
 - 3.2.2. Računalna procjena neuređenosti proteina (*engl.* disorder prediction)
 - 3.2.2.1. PrDOS
 - 3.2.2.2. DISOPRED3
 - 3.2.2.3. MFDp2
 - 3.2.3. BLAST pretraga na zapisima baze podataka proteinskih struktura (PDB)
 - 3.2.4. BLAST pretraga na zapisima baze podataka DisProt
 - 3.2.5. Analiza identificiranih proteina prema rječniku Gene Ontology (GO)
4. REZULTATI
 - 4.1. Identifikacija IDP u proteomu *A. thaliana*
 - 4.1.1. Odvajanje IDP i globularnih proteina
 - 4.1.1.1. Diferencijalna ekspresija IDP
 - 4.1.1.2. Denzitometrijska analiza ekspresije proteina alatom Gel Analyzer 2010a
 - 4.1.2. Identifikacija proteina masenom spektrometrijom
 - 4.2. Bioinformatička analiza identificiranih proteina
 - 4.2.1. Identifikacija očuvanih domena (*engl.* conserved domain search)
 - 4.2.2. Računalna procjena neuređenosti proteina (*engl.* disorder prediction)
 - 4.2.3. BLAST pretraga na zapisima baze podataka proteinskih struktura (PDB)
 - 4.2.4. BLAST pretraga na zapisima baze podataka DisProt
 - 4.2.5. Analiza identificiranih proteina prema rječniku Gene Ontology (GO)
 - 4.2.5.1. Analiza prema biološkom procesu (*engl.* GO: biological process)
 - 4.2.5.2. Analiza prema staničnoj lokalizaciji (*engl.* GO: cellular component)
 - 4.2.5.3. Analiza prema funkciji proteina (*engl.* GO: molecular function)

5. RASPRAVA
6. ZAKLJUČAK
7. ZAHVALE
8. LITERATURA
9. SAŽETAK
10. SUMMARY

1. UVOD

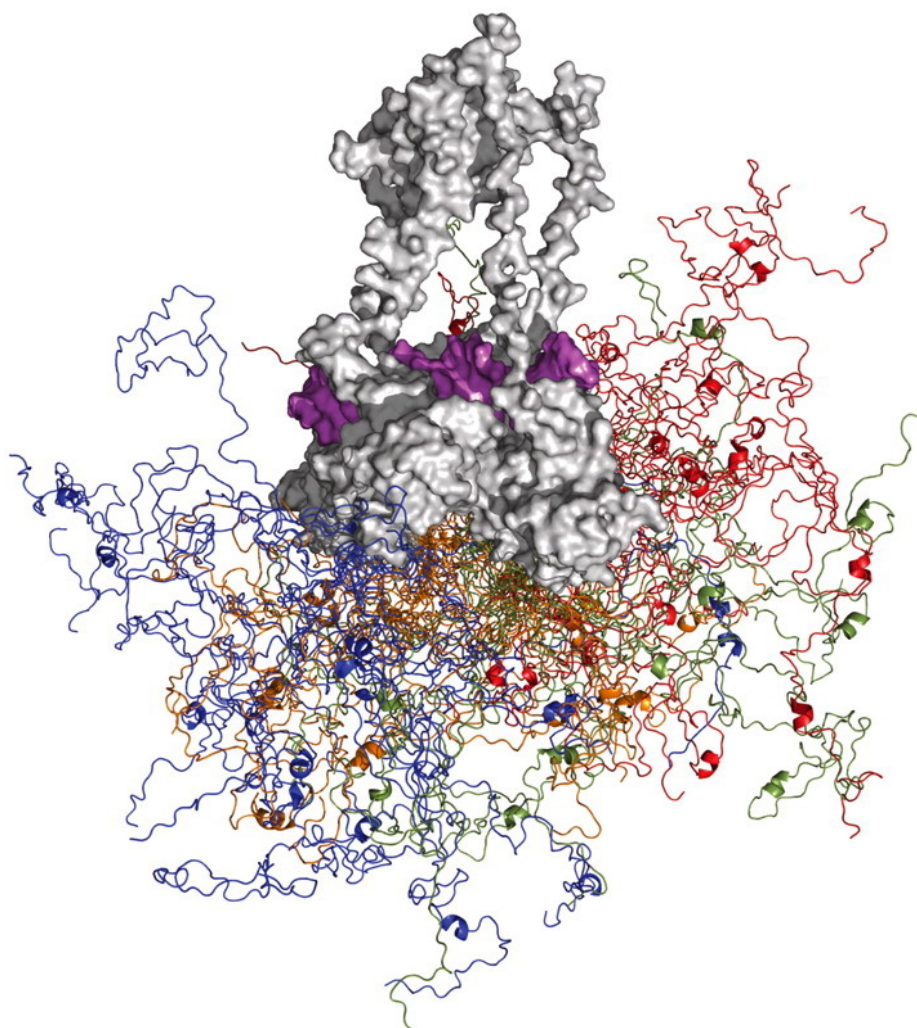
1.1. Struktura globularnih proteina

Proteini su polimeri građeni od aminokiselina vezanih peptidnom vezom. Broj aminokiselina u proteinima znatno varira, no većina ih sadrži manje od 2000 aminokiselina. U svim organizmima proteine gradi 20 aminokiselina u L-konformaciji. Svaka aminokiselina ima amino i karboksilnu skupinu, atom vodika te bočni ogranak vezan na alfa atom ugljika. Bočni ogranci aminokiselina određuju njihova svojstva te se one međusobno razlikuju po veličini, reaktivnosti, hidrofobnosti i naboju (Nelson i Cox, 2013).

Primarnu strukturu proteina čini linearni slijed aminokiselina koje svojim svojstvima definiraju sekundarnu, tercijarnu i kvaternu strukturu. Sekundarna struktura odnosi se na određeni segment polipeptidnog lanca te opisuje lokalnu prostornu raspodjelu atoma proteinske okosnice. Neki od najčešćih elemenata sekundarne strukture su α -zavojnica, β -ploča i β -okret. Tercijarnu strukturu proteina čini sveukupan trodimenzionalni raspored atoma polipeptidnog lanca. Ukoliko je protein sastavljen od više zasebnih polipeptidnih lanaca (podjedinica), pri opisu njihovog položaja u prostoru govorimo o kvaternoj strukturi proteina (Nelson i Cox, 2013).

1.2. Nativno neuređeni proteini (IDP)

Nativno neuređeni proteini i proteinske regije (*engl.* intrinsically disordered proteins/regions, IDP/IDR) pripadaju nedavno priznatoj skupini proteina koji su biološki aktivni usprkos nedostatku jasno definirane trodimenzionalne strukture (Uversky, 2014a). Za razliku od globularnih proteina, IDP u nativnom stanju ne posjeduju jedinstvenu stabilnu strukturu, već fluktuirajući skup strukturnih konformacija (Uversky, 2014b; Csizmok i sur., 2007). Takve skupove konformacija (*engl.* conformational ensembles) odlikuje heterogenost i dinamičnost, tj. izrazita varijabilnost prostornih koordinata atoma polipeptidnog lanca u ovisnosti o vremenu (Uversky, 2014b; Habchi i sur., 2014). IDP najčešće pokazuju manjak stabilne tercijarne, rjeđe i sekundarne strukture, ali nipošto ne i potpuno odsustvo strukturnih elemenata (Tompa, 2012).



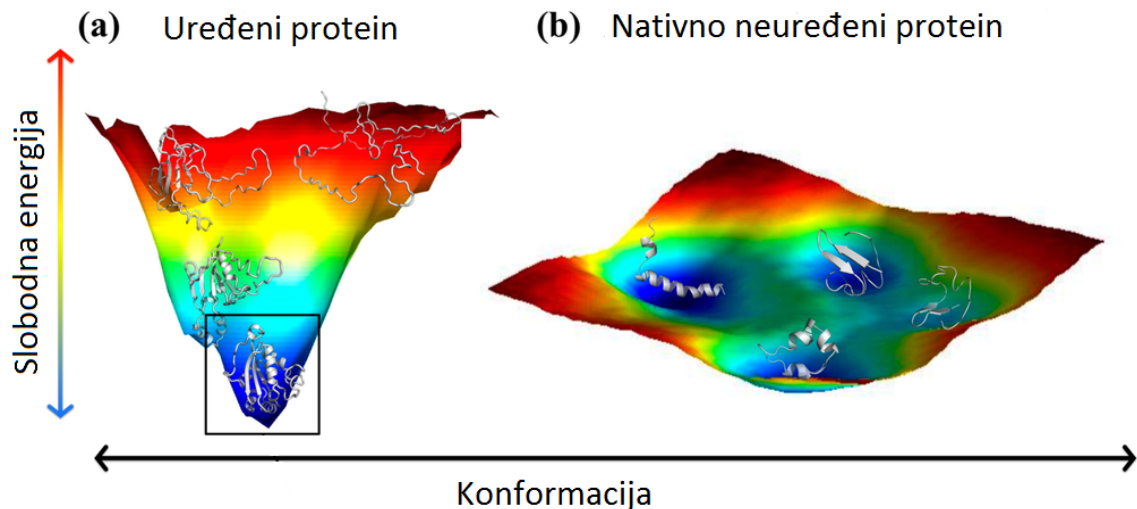
Slika 1. Primjer IDP: struktura tumor supresora p53 i njegove nativno neuređene N-terminalne transaktivacijske domene prikazane kao skup od 20 konformacija (Wells i sur., 2008).

1.2.1. Važnost i otkriće IDP

Postupno otkrivanje proteina čija konformacijska stabilnost nije izravan preduvjet njihovoj funkcionalnosti dovelo je, u proteklih nekoliko desetljeća, do promjene ustaljenih gledišta o vezi strukture s funkcijom proteina te do osnutka novog područja istraživanja (Tompa, 2012). Upravo je nedostatak strukturnog uređenja element koji se smatra ključnim za funkciju IDP (Pietrosemoli i sur., 2013). Za ulogu mnogih proteina nužno je da istodobno sadrže i uređene i neuređene regije (Uversky, 2014b).

Zahvaljujući unaprijeđenim metodama istraživanja, IDP su postali prihvaćeni kao mnogobrojna klasa proteina s važnim staničnim ulogama. Primjerice, važnost IDP očituje se kroz njihovu ulogu u razvoju neurodegenerativnih poremećaja, tumora i kardiovaskularnih bolesti kod ljudi (Uversky i sur., 2014e; Midic i sur., 2009; Uversky i sur., 2008). Neki od IDP koji su često predmet biomedicinskih istraživanja su alpha-synuclein, protein tau, huntingtin i tumor supresori PTEN i p53 (Uversky i sur., 2014c; Malaney i sur., 2013). IDP su atraktivne mete za razvoj novih lijekova (Uversky, 2012; Uversky, 2014b). Istaknut je i niz potencijalnih primjena IDP u biotehnologiji (Uversky, 2014d).

Razlog relativno nedavnom otkriću i priznavanju IDP bilo je prevladavajuće mišljenje da je jasno definirana stabilna struktura nužan preduvjet za specifičnu funkciju nekog proteina. Izravan uzrok gubitka funkcije proteina uslijed denaturacije su deformacije u njegovoj strukturi. Ovaj koncept, tzv. "*structure-function paradigm*," bio je temelj moderne strukturne biologije koji se očituje već u Fischerovom ključ-brava modelu iz 1894. Snažno ga potkrjepljuju mnogobrojne proteinske strukture u Protein Data Bank (PDB) bazi podataka, koja sadrži čak 99274 strukture od 1.3.2015 (Bernstein i sur., 1977). Stoga, prva otkrića funkcionalnih proteina koji svojom prirodom odstupaju od ove paradigme bila su smatrana anomalijama i nerijetko ignorirana od strane šire znanstvene zajednice (Tompa, 2012). No, danas je poznato da mnoge strukture u PDB-u sadrže nativno neuređene regije (IDR) koje se smataju isključivo pri vezanju liganda (Tompa, 2012). Također, upravo se dinamičnošću IDP i IDR razjašnjavaju pojave kao što su poteškoće u kristalizaciji proteina i vrlo česte regije nedostatka elektronske gustoće (*engl.* missing electron density regions) u strukturama riješenim metodom kristalografije X-zraka (Gall i sur., 2007; Dosztanyi i sur., 2009). Ove regije su u tolikoj mjeri mobilne, tj. nedostaje im stabilna fiksna struktura, da ih se ne može adekvatno opisati samo jednim skupom prostornih koordinata (Uversky, 2014b). Stoga, proteini mogu biti neuređeni na različitim mjestima duž polipeptidnog lanca, u različitim količinama te na različitim razinama strukture.

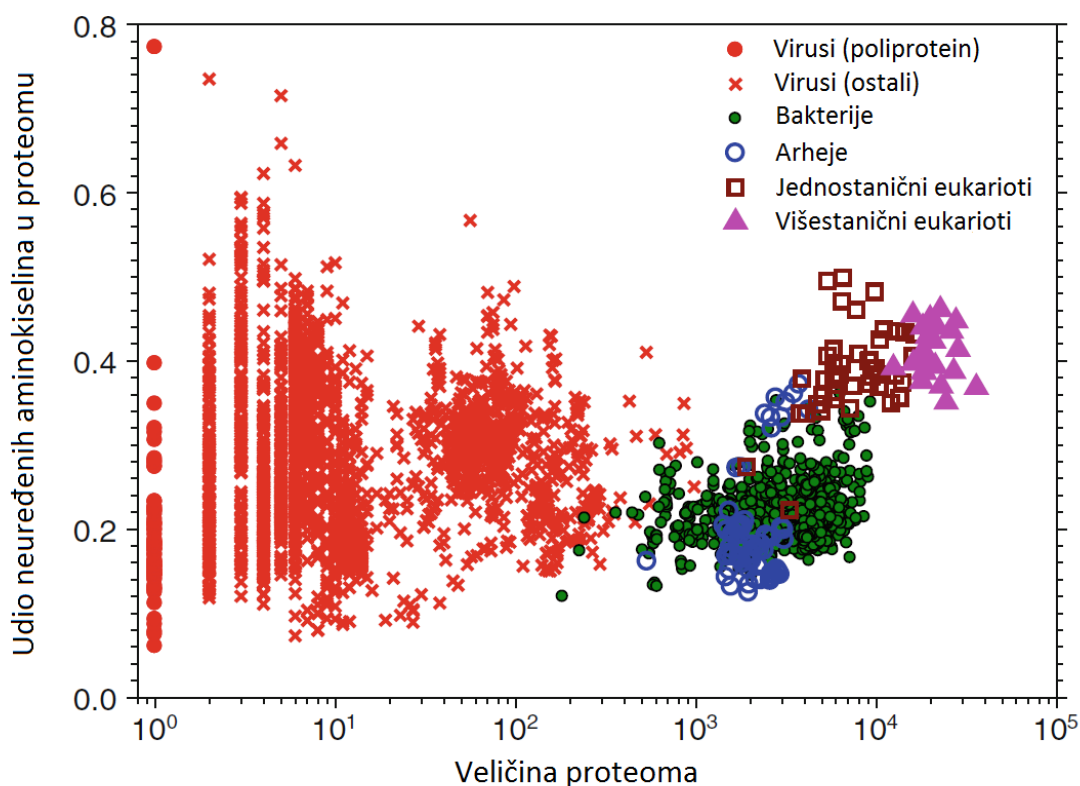


Slika 2. Usporedni shematski prikaz strukturnih konformacija i krajobraza slobodne energije (*engl.* free-energy landscape). Naspram jedne konformacije uređenog proteina (a), IDP ima skup od 4 različite konformacije (b). Prilagođeno iz Burger i sur. (2014).

Pomna eksperimentalna karakterizacija IDP bila je isprva onemogućena time što je većina metoda bila prilagođena analizi globularnih proteina, ali je nekolicina značajnih računalnih studija pružila najraniji uvid u mnoge aspekte IDP koji su kasnije eksperimentalno potvrđeni. Fokus ovih studija bile su mnogobrojnost i uloga IDP u staničnim procesima (detaljno opisana u nastavku uvoda, v. podnaslov 1.2.4.). Već 1998. godine pokazano je da tadašnje izdanje SwissProt baze podataka sadrži 15,000 proteina s neuređenim regijama od 40 ili više uzastopnih aminokiselina (Romero i sur., 1998).

Kasnije studije ukazale su na činjenicu da su IDP brojniji u eukariotskim nego u prokariotskim proteomima. Primjerice, Tompa (2012) navodi da prema konzervativnim procjenama 10-35% prokariotskih i 15-45% eukariotskih proteina sadrži duge neuređene regije, (LDR, *engl.* long disordered regions) od 30 ili više uzastopnih aminokiselina. Proteini se najčešće definiraju kao IDP ako sadrže bar jednu ovakvu LDR. Prema Peng i sur. (2014) eukarioti imaju 25%, a prokarioti 7% LDR u svojim proteinima. Ward i sur. (2004a) tvrde da su LDR prisutne u 2.0% arhejskih, 4.2% bakterijskih i 33.0% eukariotskih proteina.

Dosad najopsežnija računalna studija IDP, provedena na ukupno 3484 proteoma, ukazuje na jasnu granicu u količini neuređenosti između prokariota i eukariota, kao i na oštar porast u neuređenosti pri prijelazu iz prokariotskih u eukariotski tip stanice. Autori predlažu da takav porast broja IDP i IDR odražava mehanizam prilagodbe stanica na povećane potrebe za signalizacijom i regulacijom pri porastu stanične kompleksnosti (Xue i sur., 2012). Postotci neuređenosti po skupinama organizama u navedenim se studijama razlikuju zbog različitih softverskih alata korištenih za procjenu neuređenosti (v. podnaslov 1.2.2.), no u svima je vidljiv trend porasta količine IDP s kompleksnošću stanica.



Slika 3. Sadržaj neuređenosti (*engl.* disorder content: the average fraction of disordered residues) u odnosu na veličinu proteoma. Šest skupina označenih simbolima obuhvaća 3,484 proteoma iz virusa, arheja, bakterija i eukariota. Korišten je procjenitelj neuređenosti PONDR-VSL2B. Prilagođeno iz Xue i sur., 2012.

1.2.2. Bioinformatički alati za analizu IDP

Primjena bioinformatičkih alata u ranim računalnim studijama uvelike je potakla razvoj područja istraživanja nativno neuređenih proteina (Habchi i sur., 2014). Bioinformatički alati su i danas od iznimne važnosti za izučavanje svojstava IDP, a validacijom rezultata i usmjeravanjem eksperimentalnih istraživanja ubrzavaju otkriće novih IDP i IDR (He i sur., 2009). Dosad je razvijeno više od 50 različitih računalnih alata za procjenu nativne neuređenosti proteina (procjenitelji neuređenosti, *engl.* intrinsic disorder predictors) (He i sur., 2009).

Procjenitelji neuređenosti kao ulazne podatke koriste aminokiselinski slijed proteina. Zatim, na temelju određenog kriterija pružaju informaciju o sklonosti polipeptidnog lanca za neuređenošću (*engl.* disorder propensity) na razini pojedinih aminokiselina (He i sur., 2009). U razvoju prvih algoritama za procjenu neuređenosti, kao kriterij za diskriminaciju neuređenih regija iskorištene su posebnosti aminokiselinskog sastava IDP. Analizama primarne strukture pokazano je da su IDP bogati hidrofilnim i polarnim aminokiselinama te glicinom i prolinom, a siromašni aromatskim i hidrofobnim aminokiselinama (Habchi i sur., 2014; Dosztanyi i sur. 2009). Ovakav sastav odgovoran je za nemogućnost potpunog smatanja IDP u stabilne tercijarne strukture, a nekad i sekundarne strukture (Habchi i sur., 2014).

Procjenitelji neuređenosti međusobno se jako razlikuju te je zato pri detaljnoj interpretaciji neuređenosti nekog proteina nužno koristiti više njih ili kvalitetan metaprocjenitelj (metaprediktor), poznavati za što je softver specijaliziran te obratiti pažnju na postojanje konsenzusnih regija u izlaznim podacima. Metaprocjenitelji su često kvalitetniji u procjeni neuređenosti od samostalnih alata. Oni istovremeno kombiniraju nekoliko procjenitelja te traže konsenzus njihovih procjena. Prema načelima i metodama na kojima počivaju, procjenitelje neuređenosti može se svrstati u tri široke kategorije, a može ih se podijeliti i prema duljini neuređenih regija za čije se prepoznavanje koriste.

Procjenitelji koji pripadaju prvoj širokoj skupini, tzv. “propensity-based predictors” oslanjaju se na fizikalno-kemijska svojstva aminokiselina te na skale sklonosti aminokiselina za formiranje određenih strukturnih elemenata, uzimajući pritom u obzir načelo da IDP posjeduju visoki neto naboj i nisku hidrofobnost (Uversky, 2014b). Ovakvi alati su CH plot (Uversky i sur., 2000), FoldIndex (Prilusky i sur., 2005) i GlobPlot (Linding, 2003).

Druga velika skupina alata za procjenu neuređenosti obuhvaća algoritme bazirane na strojnom učenju. Ona uključuje prve razvijene procjenitelje neuređenosti, tj. PONDR skupinu (<http://www.pondr.com/>, 25.4.2015.) (Romero i sur., 1997; Li i sur., 1997; Romero i sur., 2000; Peng i sur., 2006). U ovu skupinu također pripadaju i DisEMBL te DISOPRED2 (Linding i sur., 2003; Ward i sur., 2004b). Ovakvi procjenitelji trenirani su na podacima iz proteinskih baza podataka, primjerice na kratkim regijama nedostatka elektronske gustoće u PDB-u ili na neuređenim proteinima i regijama iz baze DisProt (<http://www.disprot.org/>, 25.4.2015.). DisProt je baza eksperimentalno potvrđenih IDP, čija najnovija verzija sadrži 694 IDP i 1539 IDR (Sickmeier i sur., 2007). Broj zapisa u bazi DisProt je relativno malen naspram rezultata računalnih predikcija brojnosti IDP. Ta činjenica jasno ocrta veliku potrebu za novim visokoprotočnim studijama fokusiranim na eksperimentalnu identifikaciju IDP na proteomskoj razini (Csizmek i sur., 2007).

Treća skupina alata vodi se idejom da se IDP ne smataju potpuno jer ne mogu ostvariti dovoljan broj stabilizirajućih kontakata između svojih konstitutivnih aminokiselina. Primjer ovakvog alata je IUPred, koji procjenjuje ukupan energetski doprinos mogućih intramolekularnih kontakata (Dosztanyi i sur., 2005). Za razliku od prethodne skupine, ovakvi procjenitelji su neovisni o skupovima podataka za strojno učenje koji ne sadrže ravnomjerne količine svih vrsta IDP, a posebno većinski neuređenih IDP. Stoga se IUPred i sl. smatraju najboljima za identifikaciju većinski neuređenih proteina ili dugih neuređenih regija (Dosztanyi i sur., 2009).

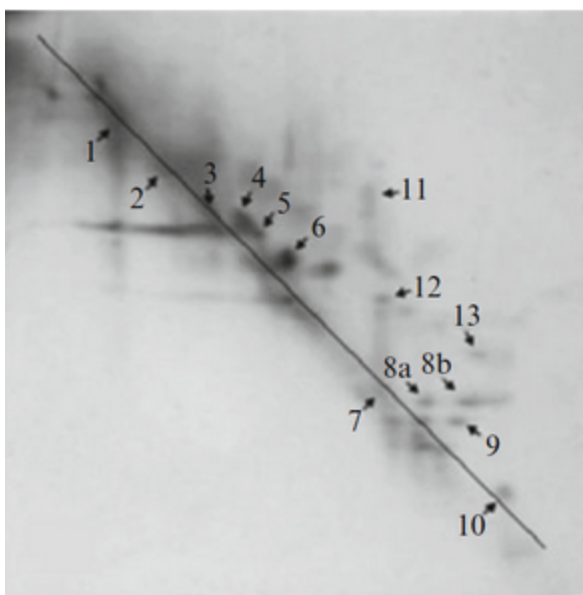
Važno je naglasiti da su procjenitelji neuređenosti uključeni u proces nepristrane evaluacije kvalitete softvera za predviđanje strukture proteina, tj. CASP (Critical Assessment of Structure Prediction, <http://www.predictioncenter.org/index.cgi>, 25.4.2015.). Prema rezultatima analiza provedenih u sklopu CASP10 trenutno najbolji procjenitelji neuređenosti su DISOPRED3, PrDOS i MFDp, detaljnije opisani u podnaslovu 3.2.2. (Monastyrskyy i sur., 2013).

1.2.3. Eksperimentalne metode istraživanja IDP

Proteomske metode identifikacije IDP temelje se na neobičnim svojstvima IDP kao što su otpornost na denaturaciju visokim temperaturama i jakim kiselinama (Habchi i sur., 2014; Tompa, 2010). U ovakvim uvjetima većina globularnih (uređenih) proteina denaturira i precipitira, dok IDP ostaju u otopini (Csizmok i sur., 2007; Tompa, 2010).

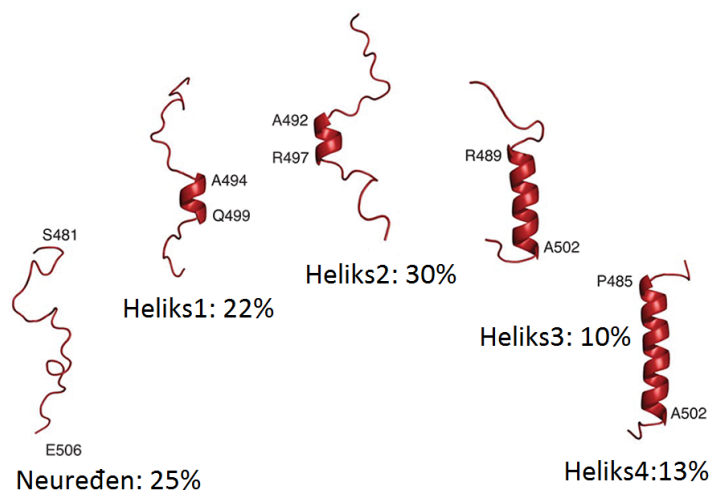
Csizmok i sur. (2006) razvili su proteomsku metodu otkrivanja IDP koja uključuje toplinsku denaturaciju kao prvi korak u pročišćavanju IDP. Daljnje razdvajanje IDP i preostalih termostabilnih globularnih proteina postiže se dvodimenzionalnom poliakrilamidnom gel elektroforezom s 8M urejom. U prvoj dimenziji u nativnom gelu dolazi do razdvajanja proteina prema omjeru naboja i mase. U drugoj dimenziji zbog prisutnosti 8M ureje dolazi do denaturacije i usporavanja globularnih proteina u odnosu na prvu dimenziju, dok IDP ostaju većinom nepromijenjeni te jednako mobilni. Zbog toga se očekuje da će IDP, prešavši jednaku udaljenost u oba gela, biti pozicionirani na dijagonali denaturirajućeg gela (Slika 4). Usporeni globularni proteini smještaju se u područje iznad dijagonale. Identifikacija proteina iz gela vrši se metodom masene spektrometrije.

Primjenom proteomskih metoda IDP su dosad uspješno identificirani u proteomima *S. cerevisiae*, *E. coli*, *D. melanogaster* i fibroblastima NIH3T3 vrste *M. musculus* (Galea i sur., 2009; Csizmok i sur., 2006; Szóllósi i sur., 2008). Sličan je princip korišten u identifikaciji fosforiliranih proteina iz sjemena *A. thaliana* (Irar i sur., 2006).



Slika 4. Separacija i identifikacija IDP u uzorcima *S. cerevisiae* tretiranim toplinom. Proteini su identificirani masenom spektrometrijom. (Csizmok i sur., 2006).

Mnoge metode strukturne biofizike primjenjive su u detekciji nedostatka jedinstvene stabilne strukture proteina. Najistaknutija metoda koja omogućava kvantitativni opis IDP je spektroskopija NMR. NMR pruža informacije o dinamici IDP te heterogenom skupu konformacija kojim se opisuje struktura IDP. No, zbog ograničenja NMR kao što je visoka količina smetnji (*engl.* noise) naspram signala i uprosječnjavanja signala brzoizmjenjujućih događaja, nužno je upotpuniti analize ostalim spektroskopskim metodama, atomističkim simulacijama i dr. metodama računalne biofizike (Habchi i sur., 2014; Burger i sur., 2014).



Slika 5. Skup konformacija dijela neuređene C-terminalne domene virusnog N_{TAIL} proteina prema mjerenjima NMR (Jensen i sur., 2011).

1.2.4. Podjela IDP prema mehanizmu djelovanja i staničnoj ulozi

IDP su isprva bili dijeljeni u diskretne klase po stupnju neuređenosti, no danas se smatra da postoji kontinuirani spektar proteinske strukture koji obuhvaća sve mogućnosti od većinskih IDP do većinski uređenih (globularnih) proteina (van der Lee i sur., 2014; Dyson i Wright, 2005). IDP najčešće dijelimo prema ulozi u staničnim procesima (*engl.* biological process) i prema načinu djelovanja na molekularnoj razini (*engl.* molecular function).



Slika 6. Kontinuum proteinske strukture. Prilagođeno iz Habchi i sur., 2014.

Prema načinu djelovanja IDP svrstavamo u pet skupina (Habchi i sur., 2014). U četiri skupine (*engl.* effectors, scavengers, assemblers, display sites) funkcija se zasniva na molekularnom prepoznavanju i vezanju IDP i liganda pri čemu vezanje liganda izaziva ispravno smatanje IDP (*engl.* folding upon binding) (van der Lee i sur., 2014). No, neki IDP ostaju neuređeni i nakon vezanja, tvoreći tzv. “fuzzy complexes,” kao u slučaju vezanja transkripcijskog faktora CREB na koaktivator CBP (Fuxreiter i Tompa, 2012; Tompa, 2010; van der Lee i sur., 2014). Primjer “effector” proteina je p21^{Cip1}, inhibitor proteina Cdk2 (Habchi i sur., 2014). Funkcija pete skupine, tj. dinamičnih veznih elemenata (*engl.* flexible linkers, entropic chains), ovisi isključivo o fleksibilnosti polipeptidnog lanca, a ne o interakcijama s ligandima (Habchi i sur., 2014).

IDP su najviše uključeni u staničnu regulaciju i signalizaciju (Wright i Dyson, 2014). Iakoucheva i sur. (2002) navode da 70% svih signalnih proteina sadrži IDR. Također, IDP sudjeluju u kontroli staničnog ciklusa te imaju šaperonske uloge (Tompa, 2012; Pietrosemoli i sur., 2013; Kovacs i Tompa, 2012).

Uloga IDP u složenim mrežama proteinskih interakcija kao što su signalne kaskade i mreže transkripcijskih faktora objašnjava se prednostima koje potječu od neuređenosti, tj. konformacijske fleksibilnosti (Pazos i sur., 2013). One uključuju veću brzinu interakcija, kratkotrajno visokospecifično vezanje više partnera s niskim afinitetom te veće interakcijske površine u kompleksima (Uversky, 2013). Funkcionalna mjesta u IDP koja određuju specifičnost vezanja su oznake molekularnog prepoznavanja (*engl.* MoRF, molecular recognition features) i kratki linearni motivi (SLiM) (Pietrosemoli i sur., 2013). Funkcija IDP često je regulirana posttranslacijskim modifikacijama (van der Lee i sur., 2014).

IDP u tkivu modelnog organizma, biljke uročnjak (*Arabidopsis thaliana*) nisu eksperimentalno proučavani na razini proteoma. Irar i sur. (2006) koristili su metodu pročišćavanja proteina na bazi termostabilnosti za identifikaciju fosforiliranih LEA proteina u sjemenkama *A. thaliana*, no cilj tog istraživanja nije bila identifikacija IDP. Prema računalnoj studiji Ward i sur. (2004a) *A. thaliana* sadrži 33.8% LDR od 30 ili više uzastopnih aminokiselina. Nedavna računalna genomska analiza otkrila je prve pojedinosti o ulogama i sadržaju IDP u *A. thaliana* (Pietrosemoli i sur., 2013). Korištenjem procjenitelja neuređenosti DISOPRED pokazano je da čak 57.2% proteina u *A. thaliana* sadrži LDR.

Proteini koji su prema bioinformatičkim procjenama klasificirani kao IDP pokazali su značajnu korelaciju sa biološkim procesima unutar četiri široke kategorije: signalizacija, stanični ciklus, razvoj i odgovor na stres. U analizi bioloških procesa u koje su proteini uključeni koristila se klasifikacija prema pojmovima iz kontroliranog rječnika Gene Ontology (*engl.* Gene Ontology

terms, <http://geneontology.org/>). Gene Ontology (GO) je centralni izvor bioloških informacija nastao iz potrebe za konzistentnim načinom opisivanja genskih produkata (Ashburner i sur., 2000). GO sadrži tri ontologije, tj. kontrolirana rječnika kojima se sustavno opisuju tri glavna aspekta genskih produkata: stanična lokalizacija, molekularna funkcija i biološki proces u koji su uključeni (*engl.* cellular component (CC), molecular function (MF), biological process (BP)). Uz razvijene ontologije, GO na osnovi ovih ontologija također sadrži opširne anotacije genskih produkata iz mnogih baza podataka.

1.3. Abiotski stres u biljaka

Biljke su sesilni organizmi izrazito podložni utjecaju okolišnog stresa. Stres predstavljaju svi abiotski i biotski faktori koji značajno ograničavaju normalan rast i razvoj biljaka te negativno utječu na urod kultiviranih biljaka. U faktore biotskog stresa ubrajamo bakterije, viruse, herbivorne insekte i sl., dok abiotski stres obuhvaća visoke i niske temperature, sušu, teške metale, slanost, mehanički stres i sl. (Dangl i Jones, 2001; Kosová i sur., 2011).

Biljke su razvile brojne mehanizme odgovora na stres zbog kojih dolazi do značajnih promjena u proteomu pri stresnim uvjetima. Istraživanja na razini proteoma izrazito pridonose razjašnjavanju fizioloških mehanizama biljne tolerancije na stres (Kosová i sur., 2011). Zajednički aspekti biljnih odgovora na stresove koji uzrokuju staničnu dehidraciju (npr. hladoća, suša i slanost) uključuju povećanu sintezu nekoliko vrsta šaperona, hidrofilnih COR/LEA (*Cold Regulated/Late Embryogenesis Abundant*) proteina zaštitne uloge, te fotosintetskih proteina OEC sustava (*Oxygen Evolving Complex*) kao što je oxygen-evolving enhancer protein 1 (Kosová i sur., 2011; Amme i sur., 2006; Bae i sur., 2003)

1.4. IDP u biljkama pri normalnim i stresnim uvjetima

Biljke odlikuje visok stupanj fenotipske plastičnosti, tj. sposobnosti za prilagodbu na promjenjive okolišne uvjete. Smatra se da IDP u biljkama pridonose fenotipskoj plastičnosti jer omogućavaju odvijanje vrlo kompleksnih proteinskih interakcija unutar okolišnog odgovora (Pazos i sur., 2013). IDP se u biljaka dijele u 5 velikih skupina: LEA proteini, transkripcijski faktori NAC i bZIP, te proteini obitelji GRAS i CRY (*Cryptochrome*) (Sun i sur., 2013, Pazos i sur., 2013).

Unatoč brojnim bioinformatički predviđenim biljnim IDP, rad na eksperimentalnoj potvrdi je u zaostatku i potrebna su daljnja istraživanja (Marin i Ott, 2014.)

GRAS proteini imaju ulogu u rastu i razvoju biljaka te sudjeluju u provođenju signala. Oni djeluju kao integratori signala, primjerice onih od različitih hormona. Integraciju omogućava njihova neuređena N-terminalna regija koja interagira s nekoliko veznih partnera uz pomoć oznaka molekularnog prepoznavanja, tj. MoRF. C-terminalna regija je uređena i evolucijski visoko očuvana (Pazos i sur., 2013, Sun i sur., 2012). CRY proteini su flavoproteini koji imaju esencijalnu ulogu u svjetlosnoj signalizaciji te utječu na otvaranje supki i produkciju antocijanina u biljaka. Pokazano je da CRY1 protein interagira s negativnim regulatorom fotomorfogeneze, COP1 (*constitutive photomorphogenic 1*) (Yang i sur., 2001). Eksperimentalno je dokazano da su rekombinantne C-terminalne domene biljnih CRY proteina neuređene (Sun i sur., 2013).

Transkripcijski faktori NAC i bZIP te LEA proteini su tri glavne skupine IDP koje pomažu biljkama da podnesu razne vrste abiotskog stresa (Sun i sur., 2013). NAC proteini su mnogobrojna skupina biljnih proteina koji sudjeluju u razvoju, senescenciji, obrani i odgovoru na abiotški stres (Ricachenevsky i sur., 2013, Jensen i sur., 2010). Ovi proteini imaju kraću, uređenu N-terminalnu domenu za vezanje DNA i dulju neuređenu C-terminalnu domenu varijabilnije sekvence s čestom pojavom aminokiselina Pro, Ser i Thr. Ova transkripcijska regulatorna domena veže nekoliko različitih partnera pri čemu se lokalno smata u strukturu alfa-zavojnice. Transkripcijski faktori bZIP također sadrže DNA-vezujuću i neuređenu domenu, a reguliraju gene cirkadijskog ritma, razvoja i odgovora na stres i zračenje (Sun i sur., 2013.).

LEA proteini su najproučavaniji i najbrojniji biljni IDP (Marin i Ott, 2014). Ekspimiraju se u kasnim fazama razvoja sjemenke i konstitutivno u meristemskim tkivima. LEA pomažu u toleranciji biljke na abiotški stres izazvan sušom i niskim temperaturama (Hincha i Thalhammer, 2012, Kovacs i sur., 2008, Candat i sur., 2014). Preko 700 LEA podijeljeno je u tri grupe, a najveća grupa su dehidrini (Sun i sur., 2013). Osim u biljkama, homolozi LEA proteina nalaze se u nekih mikroorganizama i beskralješnjaka (Hincha i Thalhammer, 2012). Klasifikacija LEA nalazi se u LEAPdb bazi podataka (<http://forge.info.univ-angers.fr/~gh/Leadb/index.php>, 25.4.2015., Hunault i sur., 2010). Mehanizmi djelovanja LEA još su uvelike nerazjašnjeni. Primjerice, LEA proteini COR15A i COR15B u *A. thaliana* se pri uvjetima suše iz neuređenog stanja smataju u alfa-zavojnicu, no način djelovanja nije im poznat (Marin i Ott, 2014.). Prema istraživanjima u uvjetima *in vitro*, uloge LEA uključuju šaperonske aktivnosti, vezanje metalnih iona, vode i vezikula te stabilizaciju membrana i proteina u stresnim uvjetima (Boucher i sur., 2010, Sun i sur., 2013, Tolleter i sur., 2010).

2. CILJ ISTRAŽIVANJA

2. CILJ ISTRAŽIVANJA

Nedavno otkriće IDP promijenilo je tradicionalna gledišta o vezi strukture i funkcije proteina. Važnost IDP očituje se u njihovoj mnogobrojnosti i ulogama u staničnim procesima poput signalizacije, regulacije transkripcije i šaperonskih aktivnosti. Poznata je i njihova uključenost u patogenezu bolesti i odgovor organizma na stresne uvjete. Računalne analize pokazuju da su IDP najbrojniji u eukariotima, gdje mogu činiti čak trećinu proteoma. No, ukupan broj eksperimentalno potvrđenih IDP još uvijek je vrlo malen zbog donedavnog nedostatka primjenjivih metoda istraživanja. Eksperimentalna istraživanja na razini proteoma su osobito potrebna u biljaka, gdje su IDP još uvelike neistražena skupina proteina.

Cilj ovog istraživanja je identificirati nove IDP u modelnoj biljci *A. thaliana* pri normalnim i stresnim uvjetima te razjasniti njihove moguće uloge u stanici analizom karakteristika njihove primarne strukture.

Specifični ciljevi istraživanja su:

1. Uzgojiti biljke *A. thaliana* i izazvati kontrolirani abiotski stres tretmanom hladnoće.
2. Izolirati IDP od većine globularnih proteina metodom pročišćavanja toplinskom denaturacijom.
3. Razdvojiti IDP i preostale termostabilne uređene proteine denaturirajućom dvodimenzionalnom elektroforezom s 8M urejom prema Csizmok i sur. (2006).
4. Utvrditi postoje li razlike u ekspresiji IDP i globularnih proteina u normalnim i stresnim uvjetima te ih kvantificirati denzitometrijskom analizom.
5. Identificirati proteine metodom masene spektrometrije.
6. Analizirati primarnu strukturu identificiranih proteina bioinformatičkim alatima za procjenu neuređenosti.
7. Istražiti uloge identificiranih proteina pretragom proteinskih baza podataka i analizom Gene Ontology anotacija za odgovarajući skup gena iz *A. thaliana*.

3. MATERIJALI I METODE

3.1. Identifikacija IDP u proteomu *A. thaliana*

3.1.1. Priprema biljnog materijala

U istraživanju su korištene biljke *A. thaliana* genotipa Oy-0. Sjemenke *A. thaliana* prebačene su u tubicu i dodan je 1 mL 70% etanola uz inkubaciju 1 minutu na tresilici. Nakon kratkog centrifugiranja etanol je otpipetiran, a na sjemenke je dodano 1 mL sterilizacijske otopine (1% Izosan G (w/w), 0.1% mukazol (v/v) u deH₂O). Sjemenke su u otopini inkubirane 10 minuta, kratko centrifugirane (1 minuta na 13000 x g) te 5 puta isprane s sterilnom deH₂O u laminaru s horizontalnim protokom zraka. Sjemenke su platirane u petrijevke sa PG0 medijem (Tablica 1., Negrutiu i sur. 1975) te pohranjene 48 h na +4°C. Petrijevke su potom premještene u komoru rasta gdje su održavane na 22°C sa fotoperiodom od 12 h (80 μmol fotona m⁻²s⁻¹). Presađivanje biljaka u zemlju izvršeno je nakon 14 dana.

Tablica 1. Sastav PG0 hranjivog medija.

Makroelementi	c/mM	Organske komponente	c/μM
NaH ₂ PO ₄ x H ₂ O	1.81	Inozitol	5.5
KCl	8.05	Vitamin B1	0.03
(NH ₄) ₂ SO ₄	3.03	Vitamin B6	0.06
MgSO ₄ x 7H ₂ O	2.03	Nikotinska kiselina	0.08
KNO ₃	19.8		
CaCl ₂ x 2H ₂ O	2.04		
Mikroelementi	c/μM	Regulatori rasta	c/μM
H ₃ BO ₃	10	2.4 - D	0.45
MnSO ₄ x 2H ₂ O	10	BAP	0.44
ZnSO ₄ x 7H ₂ O	3.7	Ostale komponente:	
KI	9.5	Agar-agar	7.5 g/L
Na ₂ MoO ₄ x 2H ₂ O	0.1	Saharoza	30 g/L
CuSO ₄ x 5H ₂ O	0.01		
CoCl ₂ x 6H ₂ O	0.01		
Na ₂ EDTA	100		
FeSO ₄ x 7H ₂ O	100		

3.1.2. Izazivanje abiotskog stresa hladnoćom

Tretman biljaka hladnoćom započet je 33 dana nakon presađivanja u zemlju. Odvojeno je 10 biljaka u kontrolnu skupinu, a 10 biljaka je podvrgnuto tretmanu (Slika 7). Biljke su izložene hladnoći u komorici s ledom tijekom 24 h te potom uzorkovane. Fotoperiod kao i jačina svjetlosti bili su identični kontrolnim biljkama. Pri uzorkovanju rozete su odrezane od ostatka biljke, osušene i izvagane. Nakon vaganja svi uzorci su zamrznuti u tekućem dušiku te pohranjeni u zamrzivač na $-80\text{ }^{\circ}\text{C}$ do izolacije proteina.



Slika 7. Kontrolna skupina biljaka prije uzorkovanja (lijevo) i biljke tijekom abiotskog stresa hladnoće (desno).

3.1.3. Izolacija IDP

Izolacija IDP provedena je prema Csizmok i sur. (2006) kako bi se IDP odvojili od većine globularnih proteina. Tarionik i tučak su ohlađeni tekućim dušikom, uzorci tkiva su smrvljeni u fini prah na koji je dodano 1 mL ekstrakcijskog pufera (50 mM Tris/HCl, 150 mM NaCl, 2 mM DTT, pH 7,5) te polivinilpirolidon (PVP) na vrh spatule. Homogenat je centrifugiran 2 h pri $25000 \times g$ i $4\text{ }^{\circ}\text{C}$ kako bi se uklonili stanični ostaci. Supernatant je prebačen u nove tubice, zagrijavan 10 minuta na $99\text{ }^{\circ}\text{C}$ te ponovno centrifugiran 2 h pri $25000 \times g$ i $4\text{ }^{\circ}\text{C}$.

3.1.4. Ukoncentriravanje proteina membranskom ultrafiltracijom

Nakon izolacije IDP provedeno je ukoncentriravanje membranskom ultrafiltracijom (Amicon 30 kDa, Millipore, Germany). Uzorci su prebačeni u posudice za filtrat spojene na centrikone te centrifugirani 90 minuta na 5500 x g pri 5 °C. Prema potrebi, određeni uzorci su ponovno ukoncentrirani 30 minuta pri 4 °C na 5500 x g. Uzorci su pohranjeni na -20 °C.

3.1.5. Određivanje koncentracije proteina

Koncentracija proteina u ukoncentriranim uzorcima određena je metodom po Bradfordu (1976). Metoda se temelji na vezanju boje Coomassie Brilliant Blue G-250 na proteine te njenoj promjeni boje iz crveno-smeđe u tamno plavu. Prvo je napravljena slijepa proba koja sadrži 20 µL ekstrakcijskog pufera i 1 mL Bradford radne otopine (2,85% (v/v) etanola + 5,28% (v/v) H₃PO₄ + 6% (v/v) Bradford matične otopine)). Uzorci su pripremljeni uzimanjem 20 µL uzorka i 1 mL Bradford radne otopine te kratko izmješani prije mjerenja apsorbancije. Apsorbancija je mjerena pri $\lambda = 595$ nm korištenjem spektrofotometra (UV/VIS Unicam Spectrometer UV 4, UK).

3.1.6. Dvodimenzionalna poliakrilamidna gel elektroforeza (2D-PAGE)

Prva dimenzija elektroforeze provedena je prema Laemmliju (1970) u mini vertikalnom poliakrilamidnom gelu debljine 0,75 mm (Tablica 2). Za elektroforezu je korišten pufer za nativnu elektroforezu (25 mM Tris + 125 mM glicin). Uzorci su razdvajani 15 minuta pri 100 V u gelu za koncentriranje (4%) te 70 minuta pri 190 V u gelu za razdvajanje (12%) dok boja bromfenol plavo nije došla do kraja gela. U svakoj jažici nanešena je ista masa proteina.

Po završetku elektroforeze gel je namakan u otopini 6M ureje i pufera za nativnu elektroforezu 15 minuta.

Tablica 2. Sastav gela za nativnu PAGE

	Gel za razdvajanje (12%)	Gel za koncentriranje (4%)
reH₂O	3,35 mL	3,050 mL
Tris/HCl	1,5 M, pH 8,8; 2,50 mL	0,5 M, pH 6,8; 1,250 mL
AA/Bis (30%)	4,00 mL	0,665 mL
APS (10%)	50 μ L	35 μ L
TEMED	5 μ L	8 μ L

U drugoj dimenziji elektroforeza je provedena u denaturirajućem gelu s 8M urejom (Tablica 3.). Nakon inkubacije u 6 M ureji, nativni gel je izrezan na trakice u širini jažica. Trakice su položene horizontalno na vrh denaturirajućeg vertikalnog gela debljine 1,00 mm i nadsvođene 0,5%-tnom agarozom. Elektroforeza je provedena u puferu za nativnu elektroforezu, na 100 V prvih 15 minuta te na 180 V još 2 sata i 45 minuta. U 50. minuti sustav je prebačen u hladnjak na 4 °C. Po završetku elektroforeze gel se inkubirao u fiksativu (50% etanol + 12% octene kiseline + 0,5 mL formaldehida u 1 L deH₂O).

Tablica 3. Sastav gela za denaturirajuću PAGE s 8 M urejom

	Gel za razdvajanje (12%)
ureja	9,61 g
1,5 M Tris/HCl pH 8,8	5,00 mL
AA/Bis (30%, w/v)	8,00 mL
APS (10%)	100 μ L
TEMED	7 μ L

3.1.7. Bojanje proteina srebrom

Proteini su po završetku denaturirajuće 2D-PAGE i inkubacije u fiksativu obojeni srebrom prema Blum i sur. (1987). Gelovi su isprani u 30%-tnom etanolu tri puta po 10 minuta. Neposredno prije upotrebe pripremljene su otopine:

- otopina 3: 20 mg $\text{Na}_2\text{S}_2\text{O}_3 \times 5\text{H}_2\text{O}$ u 100 mL reH_2O
- otopina 5: 200 mg AgNO_3 + 75 μL HCHO u 100 mL reH_2O
- otopina 7: 6 g Na_2CO_3 + 50 μL HCOH + 2 mL otopine 3 u 98 mL reH_2O

Nakon ispiranja u etanolu, gelovi su inkubirani minutu u otopini 3 te isprani u reH_2O 3 puta po 20 sekundi. Zatim slijedi inkubacija u otopini 5 u trajanju od 20 minuta. U tom koraku dolazi do vezanja srebra na proteine, tj. do impregnacije. Po isteku 20 minuta gelovi se ispiru 2 puta po 20 sekundi u reH_2O te se inkubiraju 3 minute u otopini 7. U ovom koraku dolazi do razvijanja, tj. vizualizacije. Razvijanje se zaustavlja kratkim ispiranjem u reH_2O te inkubacijom gelova u stop otopini (50% etanol + 12% octena kiselina u deH_2O).

Gelovi su skenirani pri rezoluciji od 300 DPI na 48-bitnom skeneru Epson V700.

3.1.8. Densitometrijska analiza ekspresije proteina alatom GelAnalyzer 2010a

Radi kvantifikacije razlike u jačini ekspresije pojedinih proteina napravljena je densitometrijska analiza korištenjem softvera GelAnalyzer 2010a (<http://www.gelanalyzer.com/>, 25.4.2015., dr. Istvan Lazar). Svaka proteinska mrlja od interesa na gelu je označena, detektirana te je napravljena korekcija pozadine. Za daljnje analize i usporedbu ekspresije izmjeren je i zabilježen parametar sirovi volumen (eng. *raw volume*) za svaku proteinsku mrlju.

3.1.9. Masena spektrometrija proteina

Analiza proteina izoliranih iz gela je provedena na spektrometru masa MALDI TOF/TOF 4800 Plus analyzer (Applied Biosystems, USA) (Pavoković i sur., 2011). Za identifikaciju proteina koristio se Protein Pilot v3.5 softver s integriranim Mascot algoritmom za pretraživanje

proteinskih baza podataka. Baza podataka po kojoj se vršilo pretraživanje sekvenci je *Arabidopsis thaliana* NCBIInr.

3.2. Bioinformatička karakterizacija identificiranih proteina

Provedena je opširna bioinformatička analiza proteina identificiranih masenom spektrometrijom. U pripremi skupa podataka za analizu korišteni su resursi dostupni na mrežnim stranicama NCBI (<http://www.ncbi.nlm.nih.gov/>, 25.4.2015., The National Center for Biotechnology Information). Za sve proteine identificirane masenom spektrometrijom pohranjena je aminokiselinska sekvenca u FASTA obliku te zabilježen jedinstveni identifikacijski broj (*engl.* gi number) odgovarajućeg zapisa u NCBIInr proteinskoj bazi podataka. Kada je bilo moguće, korišteni su identifikacijski brojevi zapisa iz NCBI RefSeq baze referentnih sekvenci za *A. thaliana*.

3.2.1. Identifikacija očuvanih domena (*engl.* conserved domain search)

Domene su funkcionalne jedinice proteina koje su često evolucijski očuvane. Prisutnost očuvanih domena u sekvencama identificiranih proteina analizirana je CD-Search alatom, koji predstavlja sučelje za pretragu baze podataka očuvanih domena na NCBI-u (*engl.* Conserved Domain Database). CD-Search alat je dostupan na poveznici <http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi> (25.4.2015.) (Marchler-Bauer i sur., 2014). Iz rezultata pretrage za svaki su protein zabilježeni identifikacijski kodovi specifičnih domena prisutnih u sekvenci, odgovarajući kratki opis domena te podaci o hijerarhijski višim skupinama (*engl.* superfamilies, multidomains) kojima domene pripadaju.

3.2.2. Računalna procjena neuređenosti proteina (*engl. disorder prediction*)

3.2.2.1. PrDOS

PrDOS je jedan od tri korištena računalna alata za procjenu neuređenosti proteina u ovom istraživanju. Prema evaluacijama CASP10 eksperimenta, ovaj se procjenitelj pokazao najkvalitetnijim (Monastyrskyy i sur., 2013). PrDOS na bazi aminokiselinskog slijeda predviđa neuređene regije koristeći dva integrirana procjenitelja (Ishida i Kinoshita, 2007). Izlazni podaci PrDOS-a uključuju vjerojatnost svake pojedine aminokiseline u proteinu da bude neuređena (*engl. disorder probability*). Procjene neuređenosti su vršene unošenjem sekvence svakog proteina u FASTA formatu u sučelje na poveznici <http://prdos.hgc.jp/cgi-bin/top.cgi> (25.4.2015.) pri standardnim postavkama.

3.2.2.2. DISOPRED3

DISOPRED3 je drugi najbolji procjenitelj prema CASP10, a uz procjenu neuređenih regija pruža i informaciju o mogućim veznim mjestima u proteinu (Jones i Cozzetto, 2014). DISOPRED3 algoritam treniran je na skupovima podataka sačinjenih od proteinskih struktura riješenih kristalografijom X-zrakama. U izlaznim podacima sadržane su i dodatne procjene sekundarne strukture na bazi alata PSIPRED. Procjene su vršene analogno postupku za PrDOS za svaki protein. Alati su dostupni u sklopu “PSIPRED Protein Analysis Workbench” servera na poveznici <http://bioinf.cs.ucl.ac.uk/psipred/?disopred=1> (25.4.2015.) (Buchan i sur., 2013).

3.2.2.3. MFDp2

MFDp2 je najnovija verzija procjenitelja MFDp (“Multilayered Fusion-based Disorder predictor”), koji je treći najbolji tip procjene prema evaluacijama CASP10 (Mizianty, Peng i Kurgan, 2013; Mizianty i sur., 2010; Monastyrskyy i sur., 2013). Mrežni server MFDp2 dostupan je na poveznici <http://biomine-ws.ece.ualberta.ca/MFDp2/index.php> (25.4.2015.). Procjene su vršene analogno prethodno opisanim procedurama.

3.2.3. BLAST pretraga na zapisima baze podataka proteinskih struktura (PDB)

Samostalni IDP teško kristaliziraju te su njihove strukture rijetko riješene kristalografijom X-zraka. Kako bismo u našem skupu proteina identificirali one čija je struktura eksperimentalno riješena, izvršili smo pretragu sekvenci po sličnosti korištenjem alata BLAST (Basic Local Alignment Search Tool) sa standardnim parametrima na poveznici <http://blast.ncbi.nlm.nih.gov/Blast.cgi> (25.4.2015.). Korištenjem BLAST-a, aminokiselinska sekvencama svakog od proteina identificiranih masenom spektrometrijom uspoređena je sa sekvencama proteina iz *A. thaliana* čije su strukture pohranjene u PDB-u. Zabilježeni su pozitivni rezultati pretrage (*engl.* hit) uz prag sličnosti od 80% (*engl.* % identity), no pri odabiru pozitivnih rezultata razmatrana je i E-vrijednost.

3.2.4. BLAST pretraga na zapisima baze podataka DisProt

Kako bismo u skupu proteina identificirali one koji se već nalaze u bazi podataka DisProt (baza potvrđenih IDP) ili one koji su značajno slični nekom od zapisa u DisProtu, izvršena je pretraga sekvenci po sličnosti korištenjem Smith-Waterman algoritma na poveznici <http://www.disprot.org/search.php> (25.4.2015.) (Sickmeier i sur., 2007). Za razliku od heurističkog pristupa kod BLAST-a, Smith-Waterman algoritam egzaktno uspoređuje aminokiselinsku sekvencu danog proteina sa svim sekvencama pohranjenim u određenoj bazi podataka. Izvršena je pretraga DisProt baze sa sekvencama svih proteina identificiranih masenom spektrometrijom te duljih od 50 aminokiselina.

3.2.5. Analiza identificiranih proteina prema rječniku Gene Ontology (GO)

U konzorcij GO uključen je niz baza podataka koje su usvojile anotaciju na osnovi kontroliranih rječnika, a među njima je i baza podataka o *A. thaliana*, tj. TAIR (The Arabidopsis Information Resource, <https://www.arabidopsis.org/> (25.4.2015.)). Korištenjem GO Annotations alata na mrežnim stranicama TAIR-a provedena je pretraga anotacija za proteine identificirane masenom spektrometrijom. U pretrazi su korištene oznake lokusa (*engl.* locus identifiers) odgovarajućih gena (Tablica 4.). Analizirane su stanična lokalizacija i molekularna funkcija

proteina te biološki proces u koji su proteini uključeni. Za svaki od ova tri aspekta proteini su klasificirani u nekoliko glavnih kategorija prema krovnim terminima rječnika (GO Slim).

Tablica 4. Oznake lokusa odgovarajućih gena u *A. thaliana* korištene u klasifikaciji proteina prema Gene Ontology pojmovima.

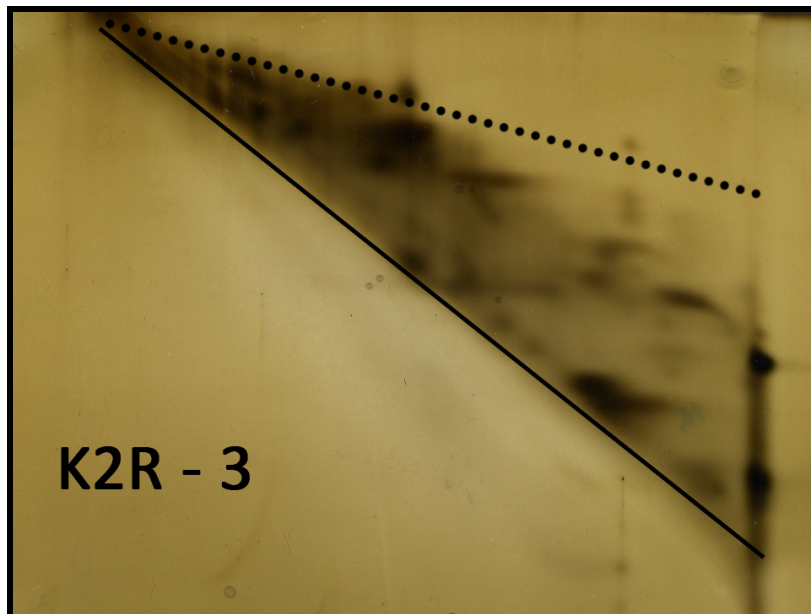
Broj proteinske mrlje	Ime proteina	Oznaka lokusa
1	adenylate kinase family protein [Arabidopsis thaliana]	AT5G35170
2	Pentatricopeptide repeat-containing protein At2g17210	AT2G17210
3	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	AT4G18520
4, 15, 31, 30	uncharacterized protein [Arabidopsis thaliana]	AT5G37667
5, 18, 33, 36	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]	AT1G09940
6	uncharacterized protein [Arabidopsis thaliana]	AT1G51970
7, 9	core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein [Arabidopsis thaliana]	AT5G14550
8, 12, 27	glutamyl-tRNA reductase 1 [Arabidopsis thaliana]	AT1G58290
10	calcium-dependent lipid-binding domain-containing protein [Arabidopsis thaliana]	AT1G04540
11	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	AT2G27610
13	uncharacterized protein [Arabidopsis thaliana]	AT1G34041
14	hypothetical protein [Arabidopsis thaliana]	AT1G51670
16	uncharacterized protein [Arabidopsis thaliana]	AT2G47844
17	uncharacterized protein [Arabidopsis thaliana]	AT2G20515
19, 28	phytochrome kinase substrate 2 [Arabidopsis thaliana]	AT1G14280
20	Serine/threonine-protein kinase HT1; AltName: Full=High leaf temperature protein 1 [Arabidopsis thaliana]	AT1G62400
21	calmodulin 1 [Arabidopsis thaliana]	AT5G37780
22	protein CURVATURE THYLAKOID 1A [Arabidopsis thaliana]	AT4G01150
23	ribonucleoprotein RBP31 [Arabidopsis thaliana]	AT4G24770
24	Chaperonin-like RbcX protein [Arabidopsis thaliana]	AT4G04330
25	COP1-interacting protein-like protein [Arabidopsis thaliana]	AT4G00920
26	uncharacterized protein [Arabidopsis thaliana]	AT4G11670
29, 35	photosystem II subunit P-1 [Arabidopsis thaliana]	AT1G06680
32	embryo defective 1241 [Arabidopsis thaliana]	AT5G17710
34	PS II oxygen-evolving complex 1 [Arabidopsis thaliana]	AT5G66570
37	F-box and associated interaction domains-containing protein [Arabidopsis thaliana]	AT1G32430
38	uncharacterized protein [Arabidopsis thaliana]	AT2G28140
39	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	AT3G15590
40	thioredoxin M2 [Arabidopsis thaliana]	AT4G03520

4. REZULTATI

4.1. Identifikacija IDP u proteomu *A. thaliana*

4.1.1. Odvajanje IDP i globularnih proteina

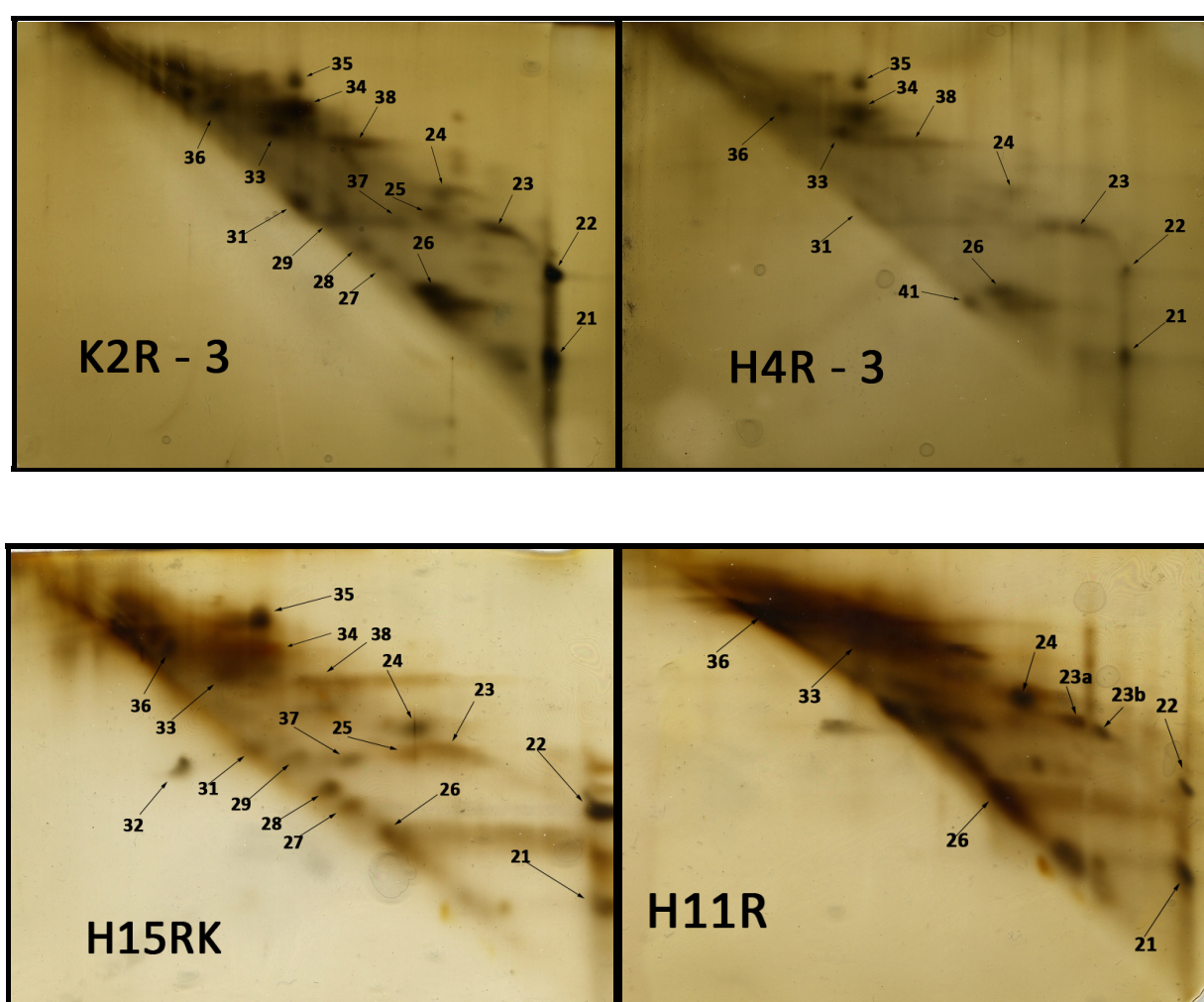
IDP su uspješno izolirani od većine globularnih proteina toplinskom denaturacijom na 99 °C. Kombinirana nativna i denaturirajuća 2D elektroforeza s 8M urejom omogućila je daljnje razdvajanje IDP od globularnih proteina. U drugoj dimenziji globularni proteini denaturiraju zbog 8M ureje, usporavaju u odnosu na prvi gel te se smještaju u područje iznad dijagonale, dok IDP prolaze približno istu udaljenost u oba gela i smještaju se na dijagonali denaturirajućeg gela (Slika 8).



Slika 8. Primjer razlike u migraciji IDP i globularnih proteina u 2D elektroforezi u uzorcima kontrolne skupine *A. thaliana* (K2R-3). U prvoj dimenziji proteini su razdvajani u 12% poliakrilamidnom nativnom gelu, a u drugoj u denaturirajućem 12% poliakrilamidnom gelu s 8M urejom. Gelovi su obojani srebrom. Puna linija označava dijagonalu oko koje su smješteni IDP, a isprekidana linija označava područje smještanja globularnih proteina.

4.1.1.1. Diferencijalna ekspresija IDP

IDP kontrolne skupine biljaka i biljaka izloženih hladnoći su razdvojeni 2D elektroforezom. Nakon bojanja proteina srebrom analizirane su kvalitativne razlike u ekspresiji (Slika 9). U odnosu na kontrolnu skupinu zapažena je ekspresija jednog dodatnog proteina (broj 41 na gelu H4R-3) te izostanak ekspresije deset proteina kod biljaka izloženih hladnoći (proteini označeni brojevima 25, 27-29, 31, 32, 34, 35, 37, 38).



Slika 9. Ekspresija proteina u *A. thaliana*. Proteini su razdvojeni kombiniranom nativnom i denaturirajućom elektroforezom s 8M urejom. K2R, H15RK - kontrolna skupina biljaka, H4R, H11R biljke izlagane hladnoći 24 h. Mase proteina nanjete na gelove iznose: K2R i H4R - 2,4 μ g te H15RK i H11R - 4,1 μ g.

4.1.1.2. Densitometrijska analiza ekspresije proteina alatom Gel Analyzer 2010a

Razlike u jačini ekspresije zajedničkih proteina za kontrolnu skupinu i skupinu izloženu hladnoći tokom 24 sata kvantificirane su alatom GelAnalyzer 2010a (Tablica 5). U biljaka tretiranih hladnoćom došlo je povećanja ekspresije 4 proteina i smanjenja ekspresije jednog proteina u odnosu na kontrolu.

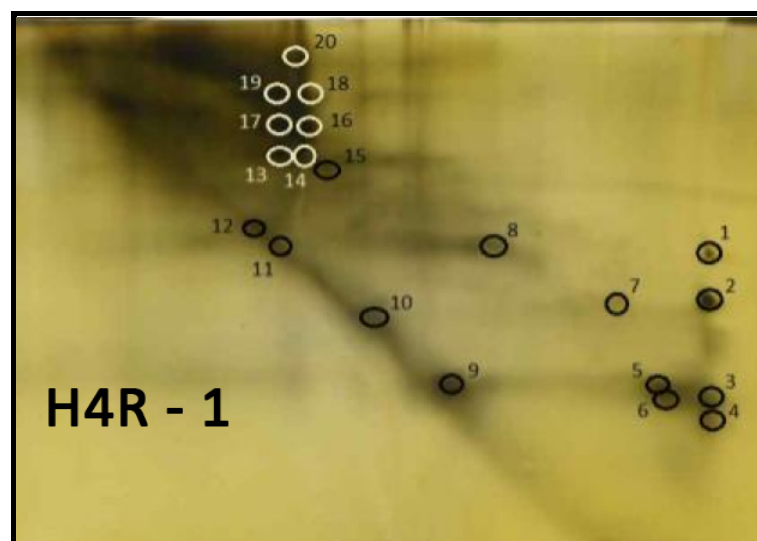
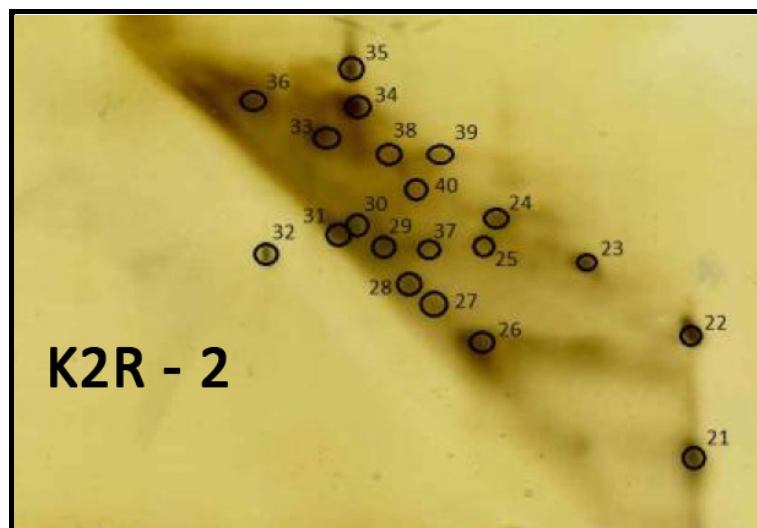
Tablica 5. Densitometrijska analiza proteinskih mrlja dobivena pomoću programa GelAnalyzer 2010a.

Proteinska mrlja	identifikacijski kod proteina	volumen mrlja:		odnos ekspresije prema kontroli
		kontrola	hladnoća	
21	gi 238481447	664	1053	↑
22	gi 18411555	1549	518	↓
23	gi 15233980	417	612	↑
24	gi 18412539	537	1145	↑
25	gi 18411479	53	-	0
26	gi 240255786	383	1151	↑
27	gi 110740768	818	-	0
28	gi 15223704	N/A	-	0
29	gi 186478207	N/A	-	0
31	gi 186527236	241	-	0
32	gi 18418410	814	-	0
33	gi 15218333	N/A	N/A	N/A
34	gi 15240013	368	-	0
35	gi 186478207	1309	-	0
36	gi 15218333	N/A	N/A	N/A
37	gi 15223198	582	-	0
38	gi 42569405	191	-	0

4.1.2. Identifikacija proteina masenom spektrometrijom

Masenom spektrometrijom identificirani su proteini iz 40 proteinskih mrlja na dva elektroforetska gela: kontrola (K2R-R) i hladnoća (H4R-1) (Tablica 6, Slika 10). U 40 proteinskih mrlja pronađeno je ukupno 29 različitih proteina. Identični proteini su identificirani iz proteinskih mrlja pod brojevima:

- 5, 18, 33, 36
- 4, 15, 30, 31
- 7, 9
- 8, 12, 27
- 19, 28
- 29, 35



Slika 10. Preparativni gelovi K2R-2 (gore) i H4R-1 (dolje) korišteni za masenu spektrometriju.

Tablica 6. Proteini identificirani masenom spektrometrijom iz uzoraka K2R-2 i H4R-1.

Proteinska mrlja	Identifikacijski broj	Masa	MASCOT	Ime proteina
1	gi 79329010	64873	38	adenylate kinase family protein [Arabidopsis thaliana]
2	gi 218546775	81549	50	Pentatricopeptide repeat-containing protein At2g17210
3	gi 145340387	63593	37	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]
4	gi 186527236	3818	40	uncharacterized protein [Arabidopsis thaliana]

5	gi 15218333	58269	48	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]
6	gi 15218085	14507	44	uncharacterized protein [Arabidopsis thaliana]
7	gi 42567851	44525	46	core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein [Arabidopsis thaliana]
8	gi 15217924	59436	49	glutamyl-tRNA reductase 1 [Arabidopsis thaliana]
9	gi 42567851	44525	39	core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein [Arabidopsis thaliana]
10	gi 15219742	64820	36	calcium-dependent lipid-binding domain-containing protein [Arabidopsis thaliana]
11	gi 15226200	97122	42	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]
12	gi 15217924	59436	48	glutamyl-tRNA reductase 1 [Arabidopsis thaliana]
13	gi 186488184	5831	22	uncharacterized protein [Arabidopsis thaliana]
14	gi 12321684	72873	39	hypothetical protein [Arabidopsis thaliana]
15	gi 186527236	3818	39	uncharacterized protein [Arabidopsis thaliana]
16	gi 186509018	5159	38	uncharacterized protein [Arabidopsis thaliana]
17	gi 18399379	17171	47	uncharacterized protein [Arabidopsis thaliana]
18	gi 15218333	58269	43	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]
19	gi 15223704	49100	52	phytochrome kinase substrate 2 [Arabidopsis thaliana]
20	gi 97052021	47372	45	Serine/threonine-protein kinase HT1; AltName: Full=High leaf temperature protein 1 [Arabidopsis thaliana]
21	gi 238481447	19799	52	calmodulin 1 [Arabidopsis thaliana]
22	gi 18411555	17686	48	protein CURVATURE THYLAKOID 1A [Arabidopsis thaliana]
23	gi 15233980	18307	57	ribonucleoprotein RBP31 [Arabidopsis thaliana]
24	gi 18412539	20186	41	Chaperonin-like RbcX protein [Arabidopsis thaliana]
25	gi 18411479	34538	42	COP1-interacting protein-like protein [Arabidopsis thaliana]
26	gi 240255786	126619	58	uncharacterized protein [Arabidopsis thaliana]
27	gi 110740768	59436	48	hypothetical protein [Arabidopsis thaliana]
28	gi 15223704	49100	58	phytochrome kinase substrate 2 [Arabidopsis thaliana]
29	gi 186478207	23744	52	photosystem II subunit P-1 [Arabidopsis thaliana]
30	gi 186527236	3818	40	uncharacterized protein [Arabidopsis thaliana]
31	gi 186527236	3818	38	uncharacterized protein [Arabidopsis thaliana]
32	gi 18418410	35472	45	embryo defective 1241 [Arabidopsis thaliana]

33	gi 15218333	58255	62	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]
34	gi 15240013	35137	212	PS II oxygen-evolving complex 1 [Arabidopsis thaliana]
35	gi 186478207	23744	134	photosystem II subunit P-1 [Arabidopsis thaliana]
36	gi 15218333	58269	50	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]
37	gi 15223198	45550	30	F-box and associated interaction domains-containing protein [Arabidopsis thaliana]
38	gi 42569405	20147	42	uncharacterized protein [Arabidopsis thaliana]
39	gi 22331104	69350	42	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]
40	gi 238480210	7921	57	thioredoxin M2 [Arabidopsis thaliana]

4.2. Bioinformatička analiza identificiranih proteina

4.2.1. Identifikacija očuvanih domena (*engl. conserved domain search*)

Alatom CD-Search uspješno su analizirane očuvane domene svih 29 identificiranih proteina (Tablica 7.). Ovi podatci omogućuju detaljniju funkcionalnu karakterizaciju tih proteina.

Tablica 7. Očuvane domene 29 proteina prema izlaznim podacima alata CD-Search.

Proteinska mrlja	Ime proteina	kodovi domena	domena 1 (engl.)	domena 2 (engl.)
------------------	--------------	---------------	------------------	------------------

1	adenylate kinase family protein [Arabidopsis thaliana]	cd01428, pfam09353	Adenylate kinase (ADK)	Domain of unknown function (DUF1995)
2	RecName: Full=Pentatricopeptide repeat-containing protein At2g17210	TIGR00756, pfam01535	pentatricopeptide repeat domain (PPR motif)	PPR repeat
3	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	pfam01535, pfam01535, TIGR00756, pfam01535	pentatricopeptide repeat domain (PPR motif)	PPR repeat
4, 15, 31, 30	uncharacterized protein [Arabidopsis thaliana]	-	-	-
5, 18, 33, 36	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]	cd05213, pfam00745	NADP-binding domain of glutamyl-tRNA reductase	Glutamyl-tRNA reductase dimerization domain
6	uncharacterized protein [Arabidopsis thaliana]	-	-	-
7, 9	core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein [Arabidopsis thaliana]	pfam02485	Core-2/I-Branching enzyme	-
8, 12, 27	glutamyl-tRNA reductase 1 [Arabidopsis thaliana]	cd05213, pfam00745	NADP-binding domain of glutamyl-tRNA reductase	Glutamyl-tRNA reductase dimerization domain
10	calcium-dependent lipid-binding domain-containing protein [Arabidopsis thaliana]	cd04051	C2 domain present in Soybean genes Regulated by Cold 2 (SRC2)-like proteins	-
11	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	pfam01535, pfam01535, pfam01535, TIGR00756	pentatricopeptide repeat domain (PPR motif)	PPR repeat
13	uncharacterized protein [Arabidopsis thaliana]	-	-	-
14	hypothetical protein [Arabidopsis thaliana]	cd01650, pfam13966	RT_nLTR: Non-LTR (long terminal repeat) retrotransposon and non-LTR retrovirus reverse transcriptase (RT)	zinc-binding in reverse transcriptase
16	uncharacterized protein [Arabidopsis thaliana]	-	-	-

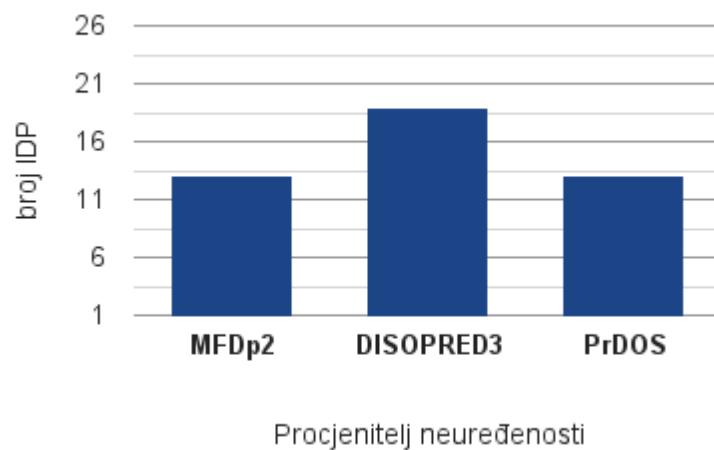
17	uncharacterized protein [Arabidopsis thaliana]	-	-	-
19, 28	phytochrome kinase substrate 2 [Arabidopsis thaliana]	-	-	-
20	RecName: Full=Serine/threonine-protein kinase HT1; AltName: Full=High leaf temperature protein 1 [Arabidopsis thaliana]	cd13999	Catalytic domain of Mitogen-Activated Protein Kinase (MAPK) Kinase-like, Serine/Threonine kinases	-
21	calmodulin 1 [Arabidopsis thaliana]	cd00051, cd00051	EF-hand, calcium binding motif	EF-hand, calcium binding motif
22	protein CURVATURE THYLAKOID 1A [Arabidopsis thaliana]	pfam14159	CAAD domains of cyanobacterial aminoacyl-tRNA synthetase	-
23	ribonucleoprotein RBP31 [Arabidopsis thaliana]	cd12399, cd12399	RNA recognition motif in Helicobacter pylori HP0827 protein and similar	RNA recognition motif in Helicobacter pylori HP0827 protein and similar
24	Chaperonin-like RbcX protein [Arabidopsis thaliana]	*samo super-obitelj domena: pfam02341	* RbcX protein; The RBCX protein has been identified as having a possible chaperone-like function	-
25	COP1-interacting protein-like protein [Arabidopsis thaliana]	-	-	-
26	uncharacterized protein [Arabidopsis thaliana]	pfam05664	Protein of unknown function (DUF810)	-
29, 35	photosystem II subunit P-1 [Arabidopsis thaliana]	PLN00042	photosystem II oxygen-evolving enhancer protein 2	-
32	embryo defective 1241 [Arabidopsis thaliana]	cd00446	nucleotide exchange factor GrpE	-
34	PS II oxygen-evolving complex 1 [Arabidopsis thaliana]	PLN00037	photosystem II oxygen-evolving enhancer protein 1	-
37	F-box and associated interaction domains-containing protein [Arabidopsis thaliana]	smart00256, pfam07734	A Receptor for Ubiquitination Targets	F-box associated
38	uncharacterized protein [Arabidopsis thaliana]	pfam07795	Protein of unknown function (DUF1635)	-

39	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	-	-	-
40	thioredoxin M2 [Arabidopsis thaliana]	cd02947	TRX family	-

4.2.2. Računalna procjena neuređenosti proteina (*engl. disorder prediction*)

U procjeni neuređenosti razmatrano je 26 od 29 proteina. Tri proteina su fragmenti kraći od 51 aminokiseline zbog čega nisu uzeti u obzir. IDP je definiran kao protein koji sadrži LDR (neuređenu regiju dulju od 30 uzastopnih aminokiselina). Izlazni podaci ukazuju da je prema procjenitelju MFDp2 pronađeno 13 IDP, prema procjenitelju DISOPRED3 19 IDP, a prema procjenitelju PrDOS 13 IDP (Slika 11). Za dvanaest proteina svi procjenitelji imaju konsenzus o neuređenosti, a dva od tri procjenitelja imaju konsenzus o neuređenosti 14 proteina (Tablica 8).

Postotak neuređenosti 26 proteina (omjer broja neuređenih i ukupnih aminokiselina u polipeptidnom lancu) grafički je prikazan prema izlaznim podacima procjenitelja MFDp2 (Slika 12). Najveći postotak neuređenosti prema MFDp2 ima protein *phytochrome kinase substrate 2* (broj 19 i 28), koji je u cjelini neuređen.



Slika 11. Usporedba broja proteina klasificiranih kao IDP prema izlaznim podacima tri procjenitelja neuređenosti.

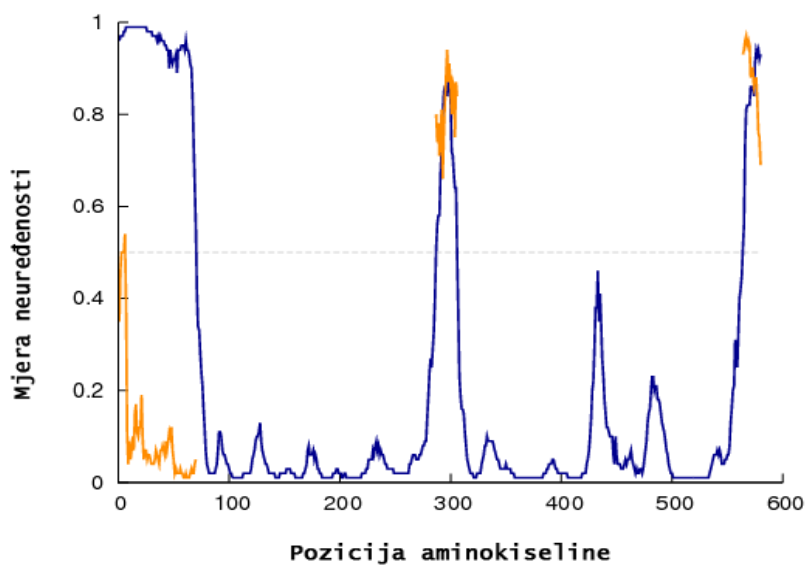
Slika 12. Postotak neuređenosti 26 proteina prema izlaznim podacima alata MFDp2.

Tablica 8. Procjena neuređenosti 26 proteina prema tri bioinformatička procjenitelja MFDp2, DISOPRED3 i PrDOS. IDP je definiran kao protein koji sadrži LDR (dugu neuređenu regiju). + — prisutnost LDR, - — odsutnost LDR. Za konsenzus više procjenitelja: IDP — nativno neuređeni protein, - — nije IDP

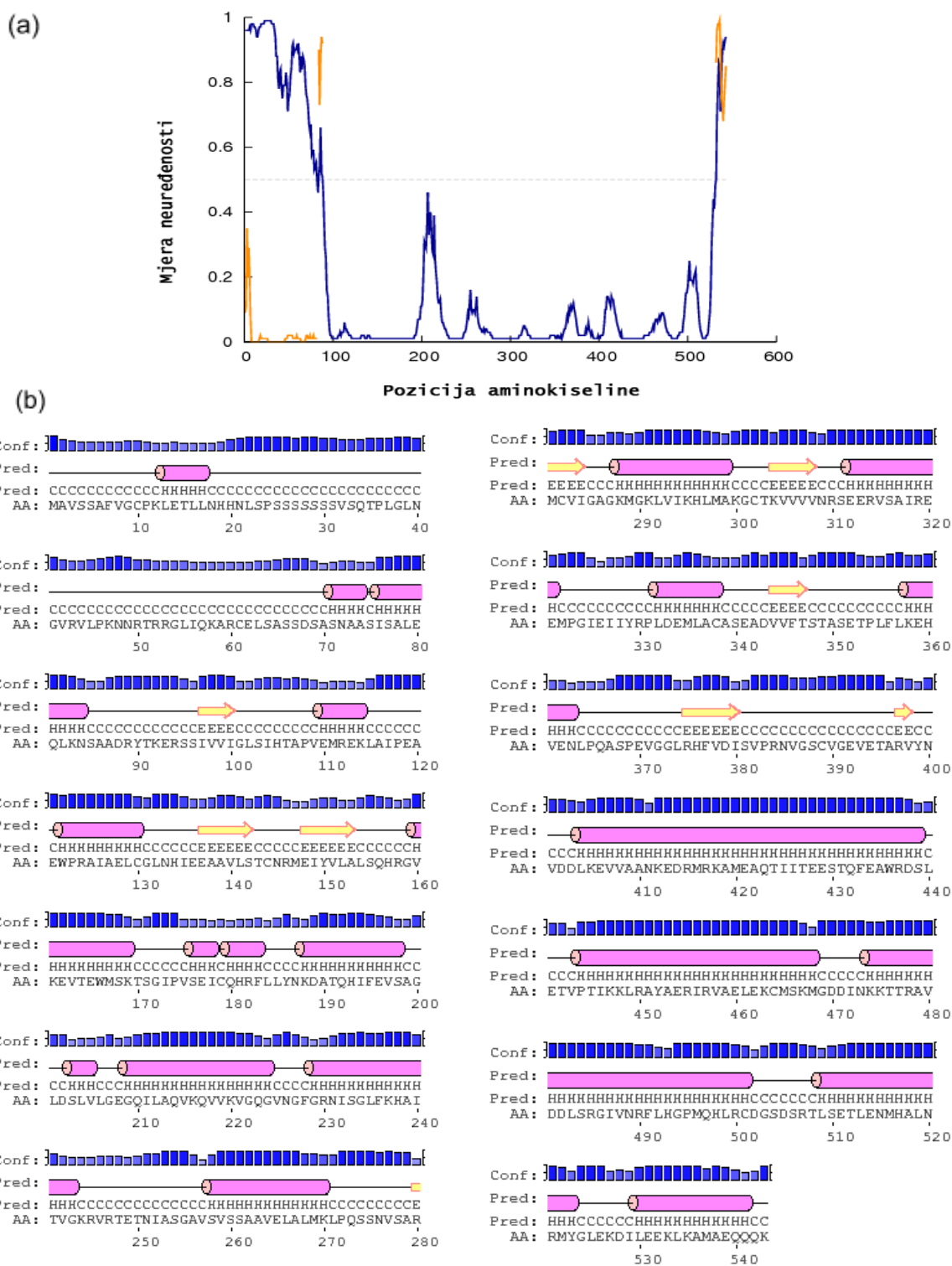
Proteinska mrlja	Ime proteina	MFDp2 LDR	DISOPRED3 LDR	PrDOS LDR	Konsenzus u dva procjenitelja	Konsenzus u tri procjenitelja
1	adenylate kinase family protein [Arabidopsis thaliana]	+	+	+	IDP	IDP
2	Pentatricopeptide repeat-containing protein At2g17210	-	-	-	-	-
3	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	-	+	+	IDP	-
5,18, 33, 36	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]	+	+	+	IDP	IDP
6	uncharacterized protein [Arabidopsis thaliana]	-	-	-	-	-
7, 9	core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein [Arabidopsis thaliana]	-	+	-	-	-
8, 12, 27	glutamyl-tRNA reductase 1 [Arabidopsis thaliana]	+	+	+	IDP	IDP
10	calcium-dependent lipid-binding domain-containing protein [Arabidopsis thaliana]	+	+	+	IDP	IDP
11	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	-	-	-	-	-

14	hypothetical protein [Arabidopsis thaliana]	-	-	-	-	-
17	uncharacterized protein [Arabidopsis thaliana]	-	+	-	-	-
19, 28	phytochrome kinase substrate 2 [Arabidopsis thaliana]	+	+	+	IDP	IDP
20	Serine/threonine-protein kinase HT1; AltName: Full=High leaf temperature protein 1 [Arabidopsis thaliana]	+	+	+	IDP	IDP
21	calmodulin 1 [Arabidopsis thaliana]	-	-	-	-	-
22	protein CURVATURE THYLAKOID 1A [Arabidopsis thaliana]	-	+	-	-	-
23	ribonucleoprotein RBP31 [Arabidopsis thaliana]	+	+	+	IDP	IDP
24	Chaperonin-like RbcX protein [Arabidopsis thaliana]	-	+	-	-	-
25	COP1-interacting protein-like protein [Arabidopsis thaliana]	+	+	+	IDP	IDP
26	uncharacterized protein [Arabidopsis thaliana]	+	+	+	IDP	IDP
29, 35	photosystem II subunit P-1 [Arabidopsis thaliana]	-	+	-	-	-
32	embryo defective 1241 [Arabidopsis thaliana]	+	+	+	IDP	IDP
34	PS II oxygen-evolving complex 1 [Arabidopsis thaliana]	+	+	+	IDP	IDP
37	F-box and associated interaction domains-containing protein [Arabidopsis thaliana]	-	-	-	-	-
38	uncharacterized protein [Arabidopsis thaliana]	+	+	-	IDP	-
39	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	+	+	+	IDP	IDP
40	thioredoxin M2 [Arabidopsis thaliana]	-	-	-	-	-

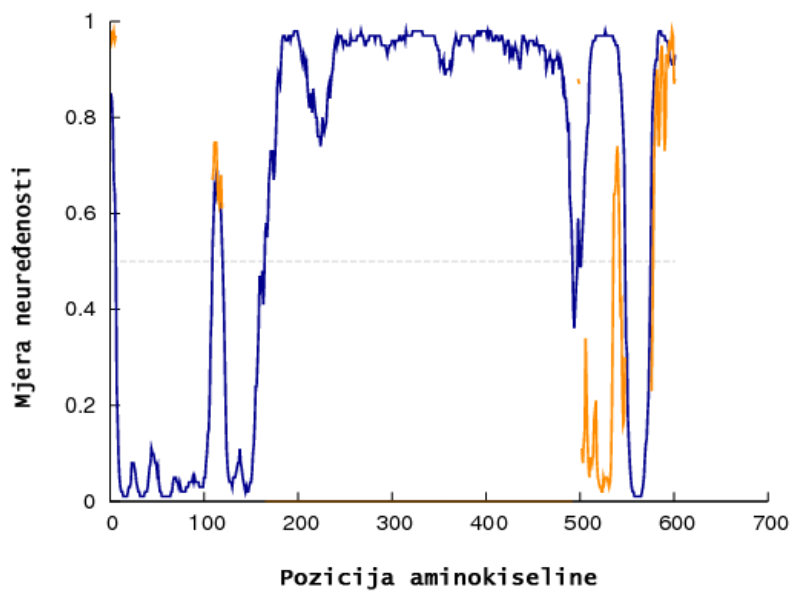
Dvanaest proteina koji su konsenzusom sva tri procjenitelja neuređenosti identificirani kao IDP detaljnije su analizirani. Mjera neuređenosti u aminokiselinskom slijedu prema izlaznim podacima procjenitelja DISOPRED3 uključuje i procjenu neuređenih veznih regija u proteinu (Slike 13-24). Za protein *glutamyl-tRNA reductase 1*, čija je struktura jedina identificirana u bazi podataka PDB, prikazan je i profil sekundarne strukture prema alatu PSIPRED.



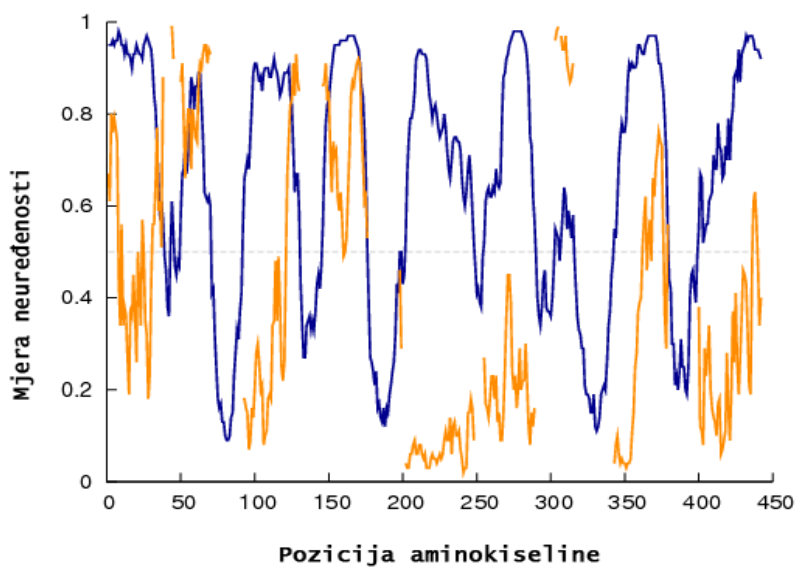
Slika 13. Grafički prikaz profila neuređenosti polipeptidnog lanca za protein “*adenylate kinase family protein*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



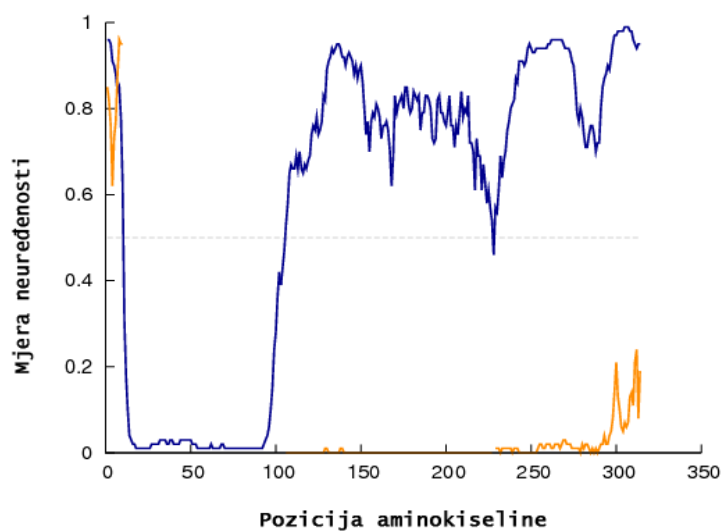
Slika 14. Profil neuređenosti (a) prema alat DISOPRED3 i profil sekundarne strukture (b) prema alat PSIPRED za protein “*glutamyl-tRNA reductase 1*”. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina. PSIPRED: roza boja - alfa-zavojnica, žuta - beta ploča, crna crta - zavojnica (*coil*).



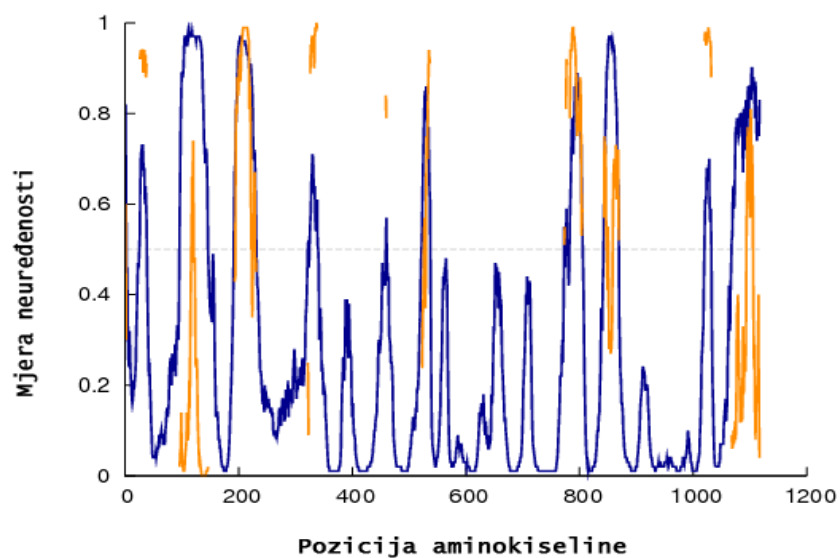
Slika 15. Grafički prikaz profila neuređenosti za protein “*calcium-dependent lipid-binding domain-containing protein*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



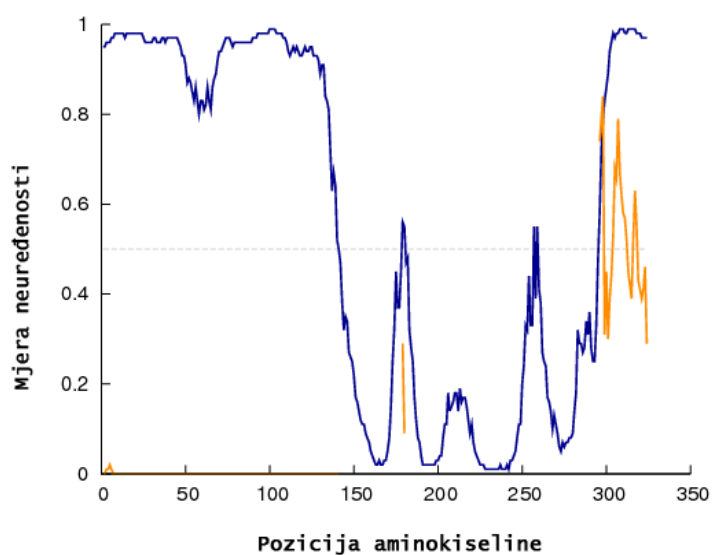
Slika 16. Grafički prikaz profila neuređenosti za protein “*phytochrome kinase substrate 2*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



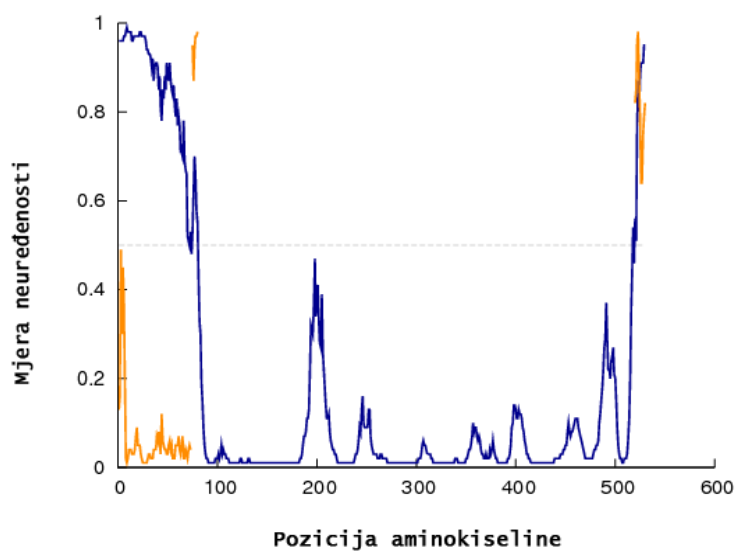
Slika 17. Grafički prikaz profila neuređenosti za protein “*COP1-interacting protein-like protein*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



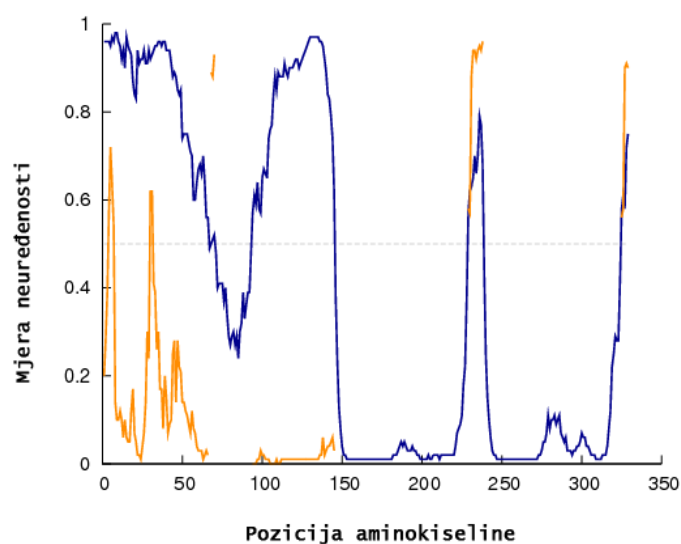
Slika 18. Grafički prikaz profila neuređenosti za protein “*uncharacterized protein*” (mrlja 26) prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



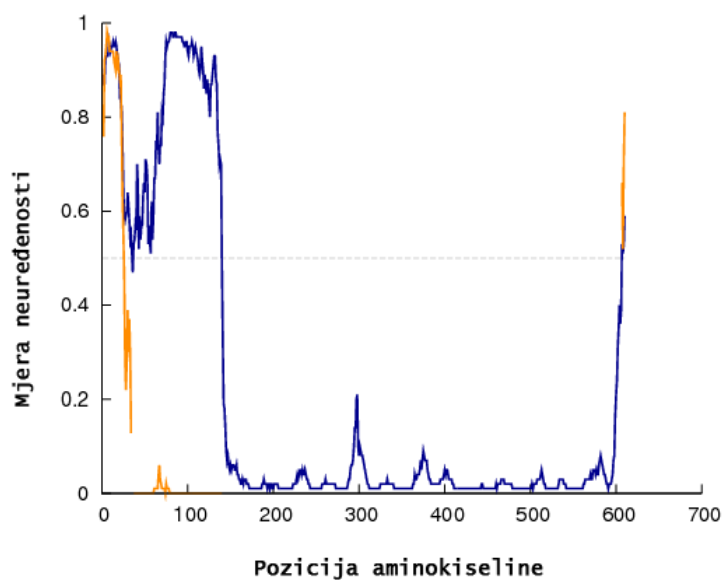
Slika 19. Grafički prikaz profila neuređenosti za protein "*embryo defective 1241*" prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



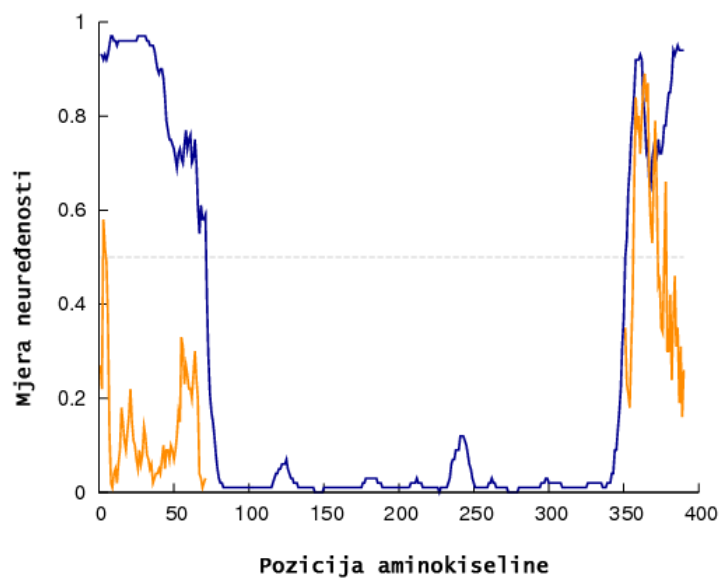
Slika 20. Grafički prikaz profila neuređenosti za protein "*glutamyl-tRNA reductase 2*" prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



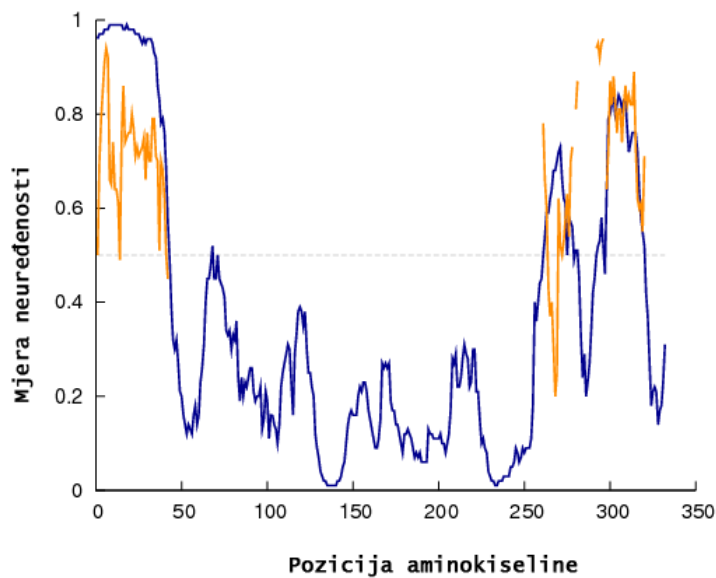
Slika 21. Grafički prikaz profila neuređenosti za protein “*ribonucleoprotein RBP31*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



Slika 22. Grafički prikaz profila neuređenosti za protein “*pentatricopeptide repeat-containing protein*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



Slika 23. Grafički prikaz profila neuređenosti za protein “*serine/threonine-protein kinase HT1*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.



Slika 24. Grafički prikaz profila neuređenosti za protein “*PS II oxygen-evolving complex I*” prema alatu DISOPRED3. Plavom bojom označena je vjerojatnost neuređenosti svake aminokiseline, a žutom bojom procjena neuređenih veznih regija proteina.

4.2.3. BLAST pretraga na zapisima baze podataka proteinskih struktura (PDB)

BLAST pretragom na zapisima PDB baze podataka pronađeno je da 5 od 29 proteina identificiranih masenom spektrometrijom imaju eksperimentalno riješenu strukturu ili veliku sličnost sa nekom riješenom strukturom proteina, pohranjenom u bazi podataka PDB (Tablica 9). Od ovih proteina, tri su IDP prema konsenzusu procjena neuređenosti.

Tablica 9. Rezultati BLAST pretrage na zapisima PDB baze podataka. E-vrijednost označava očekivanje da je za pojedinu pretragu dano preklapanje slučajno. Kada je E-vrijednost vrlo mala, tj. teži u nulu, kao u prikazanim slučajevima, smatra se da rezultat nije slučajan.

Proteinska mrlja	Ime proteina	PDB struktura	Ime strukture	% sličnosti
5, 18, 33, 36	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]	4N7R_A	Chain A, Crystal Structure Of Arabidopsis Glutamyl-trna Reductase In Complex With Its Binding Protein	85%
8, 12, 27	glutamyl-tRNA reductase 1 [Arabidopsis thaliana]	4N7R_A	Chain A, Crystal Structure Of Arabidopsis Glutamyl-trna Reductase In Complex With Its Binding Protein	99%
20	Serine/threonine-protein kinase HT1; AltName: Full=High leaf temperature protein 1 [Arabidopsis thaliana]	3PPZ_A	Chain A, Crystal Structure Of Ctr1 Kinase Domain In Complex With Staurosporine	**37% uz E-vrijednost 1e-56
21	calmodulin 1 [Arabidopsis thaliana]	4AQR_A	Chain A, Crystal Structure Of A Calmodulin In Complex With The Regulatory Domain Of A Plasma-Membrane Ca ²⁺ -Atpase	83%
24	Chaperonin-like RbcX protein [Arabidopsis thaliana]	4GR2_A	Chain A, Structure Of Atrbcx1 From Arabidopsis Thaliana.	98%

**Protein iz mrlje 20 (High leaf temperature protein 1) nije zadovoljio kriterij od 85% sličnosti, no zbog strogosti ovakvog praga sličnosti i iznimno male E-vrijednosti zabilježeno je njegovo poklapanje s PDB strukturom 3PPZ_A.

4.2.4 BLAST pretraga na zapisima baze podataka DisProt

Pretragom pomoću alata BLAST na zapisima baze podataka DisProt pronađeno je da 5 od 29 proteina identificiranih masenom spektrometrijom imaju velik postotak sličnosti sa nekim od prijašnje eksperimentalno potvrđenih IDP (Tablica 10). Od ovih proteina, četiri su IDP prema konsenzusu procjena neuređenosti, dok protein iz mrlje 21 (*calmodulin 1*) nije klasificiran kao IDP.

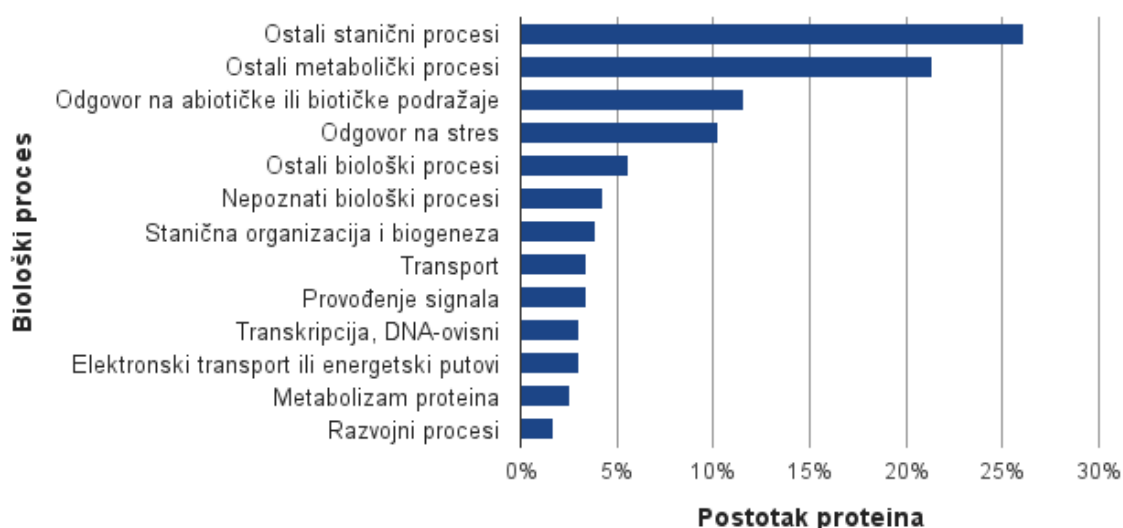
Tablica 10. Rezultati BLAST pretrage na zapisima DisProt baze podataka. Mjere sličnosti “*identity*” i “*similarity*” razlikuju se po tome što “*identity*” označava postotak potpuno identičnih aminokiselina na odgovarajućim pozicijama u aminokiselinskim slijedovima pojedinih proteina.

Proteinska mrlja	Ime proteina	DisProt zapis	E-vrijednost	Sličnost
20	Serine/threonine-protein kinase HT1; AltName: Full=High leaf temperature protein 1 [Arabidopsis thaliana]	DP00171: RAF proto-oncogene serine/threonine-protein kinase	2.10E-28	28.9% identity (64.0% similar)
21	calmodulin 1 [Arabidopsis thaliana]	DP00344: Calmodulin	6.10E-48	76.0% identity (82.9% similar)
23	ribonucleoprotein RBP31 [Arabidopsis thaliana]	DP00324: Heterogeneous nuclear ribonucleoprotein A1 [Isoform A1-B]	8.70E-13	29.6% identity (67.9% similar)
32	embryo defective 1241 [Arabidopsis thaliana]	DP00103: Protein grpE	6.30E-07	26.9% identity (64.8% similar)
34	PS II oxygen-evolving complex 1 [Arabidopsis thaliana]	DP00188: Oxygen-evolving enhancer protein 1, chloroplastic	5.60E-108	76.6% identity (94.0% similar)

4.2.5. Analiza identificiranih proteina prema rječniku Gene Ontology (GO)

4.2.5.1. Analiza prema biološkom procesu (*engl. GO: biological process*)

Analiza anotacija proteina prema rječniku za biološke procese Gene Ontology pokazala je da su “odgovor na abiotičke ili biotičke podražaje” te “odgovor na stres” najzastupljenije kategorije za analizirani skup proteina uz kategorije “ostali stanični procesi” i “ostali metabolički procesi” (Slika 25).



Slika 25. Grafički prikaz zastupljenosti pojedinih Gene Ontology kategorija s obzirom na biološki proces za analizirani skup proteina.

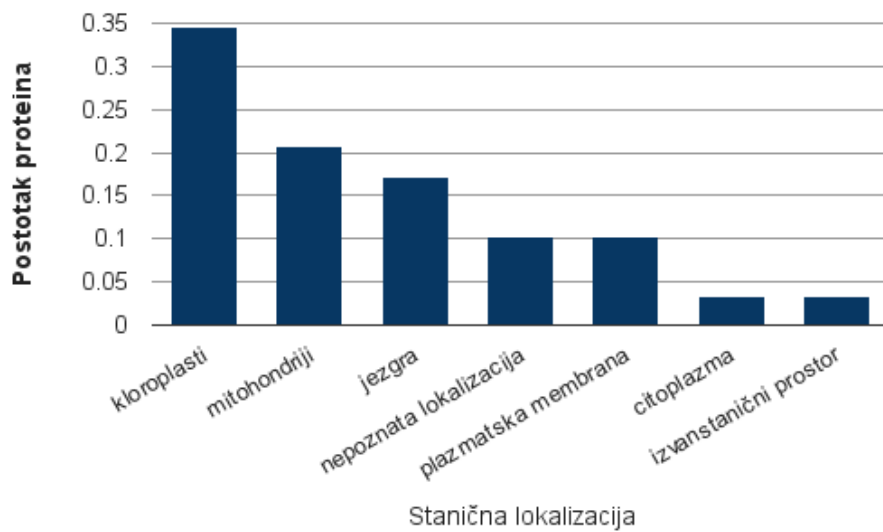
4.2.5.2. Analiza prema staničnoj lokalizaciji (*engl. GO: cellular component*)

Prema specifičnoj staničnoj lokalizaciji identificiranih proteina, najviše ih je iz organela stanične energetike: kloroplasta i mitohondrija te iz jezgre. (Tablica 11, Slika 26).

Tablica 11. Stanična lokalizacija pojedinačnih proteina prema rječniku Gene Ontology.

Proteinska mrlja	Ime proteina	Stanična lokalizacija
1	adenylate kinase family protein [Arabidopsis thaliana]	kloroplast
2	Pentatricopeptide repeat-containing protein At2g17210	mitohondrij
3	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	mitohondrij
4, 15, 31, 30	uncharacterized protein [Arabidopsis thaliana]	nepoznata
5, 18, 33, 36	glutamyl-tRNA reductase 2 [Arabidopsis thaliana]	kloroplast
6	uncharacterized protein [Arabidopsis thaliana]	jezgra
7, 9	core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein [Arabidopsis thaliana]	plazmatska membrana
8, 12, 27	glutamyl-tRNA reductase 1 [Arabidopsis thaliana]	kloroplast
10	calcium-dependent lipid-binding domain-containing protein [Arabidopsis thaliana]	plazmatska membrana
11	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	mitohondrij
13	uncharacterized protein [Arabidopsis thaliana]	nepoznata
14	hypothetical protein [Arabidopsis thaliana]	jezgra
16	uncharacterized protein [Arabidopsis thaliana]	mitohondrij
17	uncharacterized protein [Arabidopsis thaliana]	izvanstanična regija
19, 28	phytochrome kinase substrate 2 [Arabidopsis thaliana]	plazmatska membrana
20	Serine/threonine-protein kinase HT1; AltName: Full=High leaf temperature protein 1 [Arabidopsis thaliana]	jezgra
21	calmodulin 1 [Arabidopsis thaliana]	jezgra
22	protein CURVATURE THYLAKOID 1A [Arabidopsis thaliana]	kloroplast
23	ribonucleoprotein RBP31 [Arabidopsis thaliana]	kloroplast
24	Chaperonin-like RbcX protein [Arabidopsis thaliana]	kloroplast
25	COP1-interacting protein-like protein [Arabidopsis thaliana]	nepoznata
26	uncharacterized protein [Arabidopsis thaliana]	jezgra

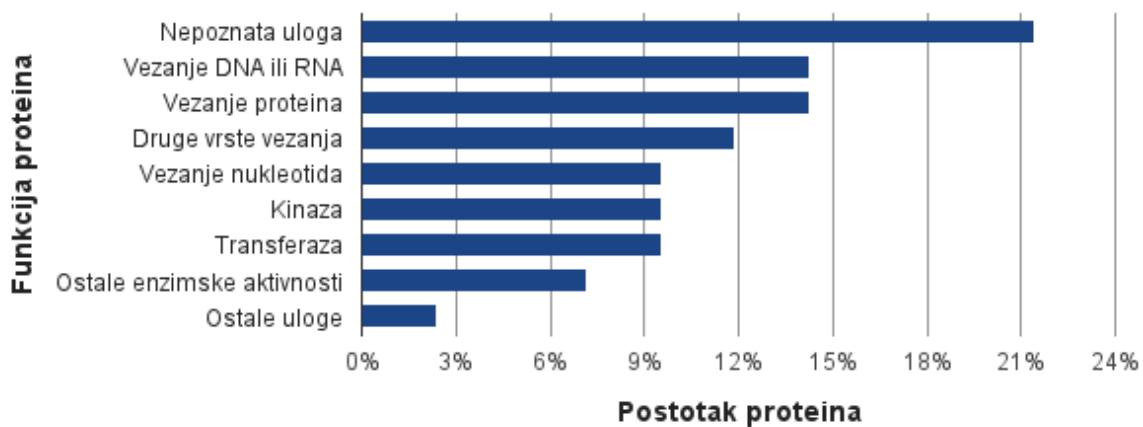
29, 35	photosystem II subunit P-1 [Arabidopsis thaliana]	kloroplast
32	embryo defective 1241 [Arabidopsis thaliana]	kloroplast
34	PS II oxygen-evolving complex 1 [Arabidopsis thaliana]	kloroplast
37	F-box and associated interaction domains-containing protein [Arabidopsis thaliana]	citoplazma
38	uncharacterized protein [Arabidopsis thaliana]	mitohondrij
39	pentatricopeptide repeat-containing protein [Arabidopsis thaliana]	mitohondrij
40	thioredoxin M2 [Arabidopsis thaliana]	kloroplast



Slika 26. Grafički prikaz postotka identificiranih proteina prema specifičnoj staničnoj lokalizaciji na osnovi anotacija u bazi podataka TAIR.

4.2.5.3. Analiza prema funkciji proteina (*engl. GO: molecular function*)

Analiza proteina prema rječniku za molekularnu funkciju Gene Ontology pokazala je da su najzastupljenije kategorije za analizirani skup proteina “nepoznata uloga,” “vezanje DNA ili RNA” i “vezanje proteina” (Slika 27).



Slika 27. Grafički prikaz zastupljenosti pojedinih Gene Ontology kategorija s obzirom na funkciju za analizirani skup proteina.

5. RASPRAVA

5. RASPRAVA

Glavni cilj ovog istraživanja bila je identifikacija i karakterizacija novih IDP u modelnom organizmu *A. thaliana* izloženom kontroliranom abiotskom stresu. Zbog donedavnog manjka prikladnih eksperimentalnih metoda, IDP su dosad istraživani u proteomima tek nekolicine organizama (Galea i sur., 2009; Csizmok i sur., 2006; Szóllósi i sur., 2008). Iako su poznate mnoge važne uloge IDP u biljaka, u literaturi se ne navodi da je dosad provedena studija IDP na razini proteoma biljnog organizma (Marin i Ott, 2014).

Primjenom metode prema Csizmok i sur. (2006) IDP su uspješno izolirani i razdvojeni od globularnih proteina u svim uzorcima. U biljkama tretiranim hladnoćom uočena je ekspresija jednog novog proteina i izostanak ekspresije deset proteina u odnosu na kontrolnu skupinu. Denzitometrijskom analizom utvrđeno je povećanje ekspresije 4 proteina i smanjenje ekspresije jednog proteina u biljkama izlaganim hladnoći u odnosu na kontrolu skupinu. Poznato je da u uvjetima abiotskog stresa dolazi do diferencijalne ekspresije brojnih proteina, kao i do *de novo* sinteze proteina uključenih u odgovor na stres (Kosová i sur., 2011; Amme i sur., 2006).

Identifikacija proteina iz obje skupine biljaka provedena je pomoću masene spektrometrije. Ukupno je identificirano 40 proteinskih mrlja sa dva elektroforetska gela, unutar kojih je pronađeno 29 različitih proteina. Pri denzitometrijskoj analizi diferencijalne ekspresije proteina, na analiziranim gelovima napravljena je korespondencija mrlja sa mrljama identificiranim masenom spektrometrijom samo sa gela sa uzorkom kontrolne skupine (mrlje 20-40), dok je zbog lošeg poklapanja na gelu sa uzorkom biljaka tretiranih hladnoćom korespondencija identificiranih proteina izostavljena (mrlje 1-20). Proteini identificirani sa oba gela razmatrani su u daljnjoj bioinformatičkoj analizi primarne strukture.

Od ukupno 29 identificiranih proteina, 26 je analizirano alatima za procjenu neuređenosti proteina. U analizama su korištena tri najkvalitetnija računalna procjenitelja neuređenosti prema evaluacijama navedenim u CASP10 (PrDOS, DISOPRED3 i MFDp2) kako bi se što točnije utvrdilo u kolikoj su mjeri pronađeni proteini neuređeni (Monastyrskyy i sur., 2013). Prema standardnom postupku u klasifikaciji IDP, razmatrana je prisutnost dugih neuređenih regija (LDR) od 30 ili više uzastopnih aminokiselina, tj. IDP su definirani kao proteini koji sadrže bar jednu takvu LDR (Peng i sur., 2014; Tompa, 2012).

Konsenzusom dva od tri procjenitelja otkriveno je 14, a konsenzusom sva tri procjenitelja dokazano je 12 novih IDP. U dosadašnjim studijama IDP u biljci *A. thaliana* otkriveno je ukupno 24 IDP (Marin i Ott, 2014). Usporedba svih 12 IDP otkrivenih u ovom istraživanju s dosad poznatim IDP iz *A. thaliana* pokazuje da je riječ o potpuno novim nalazima. S druge strane, neki

od IDP prethodno otkrivenih u drugim biljkama po funkciji su slični IDP otkrivenim u ovom radu (Marin i Ott, 2014). Također, preklapanje postoji i sa nalazima proteomskih studija *A. thaliana* u abiotskom stresu čiji primarni cilj nije bilo otkrivanje IDP.

Od 12 ukupno otkrivenih IDP, interpretirana je funkcija za onih 8 za koje je prethodno napravljena odgovarajuća korespondencija prema mrlji na gelu i potom uspješno analizirana diferencijalna ekspresija. Unutar tih 8 IDP, jedan od najneuređenijih bio je *phytochrome kinase substrate 2* (PKS2). Prema *Conserved Domains* bazi podataka, PKS2 ne sadrži niti jednu poznatu očuvanu domenu po kojoj bi se mogla razmatrati njegova funkcija. Procjenitelj DISOPRED3 ukazuje na prisutnost brojnih neuređenih regija u PKS2, a prema procjeni MFDp2 on je 100% neuređen. Riječ je o membranskom proteinu iz obitelji PKS koji interagira s proteinima *phytochrome kinase substrate 1* (PKS1) te fototropinima 1 i 2 (PHOT1 i PHOT2). PKS2 je potreban za normalni fototropizam reguliran proteinima PHOT1 i PHOT2 (Kami i sur., 2013, de Carbonnel i sur., 2010).

Proteini *glutamyl-tRNA reductase 1 i 2* (GluTR1 i 2) su dva izoforma enzima GluTR kodirana genima HEMA1 i 2 u *A. thaliana* (Apitz i sur., 2014). Ovi proteini u stanicama su uključeni u biosintezu klorofila. Oba su klasificirana kao IDP zbog prisutnosti N-terminalne neuređene regije duge stotinjak aminokiselina. No, GluTR je većinski uređen enzim za koji je nađena struktura u bazi podataka PDB, što potvrđuje i prisutnost mrlja GluTR1 i 2 u području iznad dijagonale na denaturirajućim elektroforetskim gelovima. Iako je funkcija GluTR2 poznata, ne treba odbaciti mogućnost da neuređena N-terminalna regija funkcionalno pridonosi njegovoj ulozi u stanici. Prethodno je pokazano da ovaj protein ima ulogu u odgovoru na abiotski stres (Nagai i sur., 2007; Turan i Tripathy, 2014). S druge strane, uočeno je da je ekspresija GluTR1 i 2 u abiotskom stresu iznimno smanjena ili izostaje, kao i nekonzistentan raspored njihovih mrlja na gelovima. Radi kvalitetnije interpretacije ovih proteina u budućim je studijama potrebno ponoviti identifikaciju mrlja s gelova, a postoji i mogućnost istraživanja njihove neuređenosti primjenom metoda kao što je spektroskopija NMR.

Noovotkriveni IDP *ribonucleoprotein RBP31* je kloroplastni ribonukleoprotein (cpRNP) koji ima ulogu u regulaciji razine specifičnih RNA transkripata u kloroplastima (Tillich i sur., 2009). RBP31 je visoko neuređen u dva segmenta u N-terminalnoj polovici (do 150. aminokiselina), dok njegov C-terminalni dio sadrži dvije jednake domene za vezanje RNA (segmenti 151-228 i 245-321). Na profilu neuređenosti jasno je vidljivo da ove domene imaju strukturu te da su povezane kratkom neuređenom regijom sa visokom vjerojatnošću funkcije vezanja. Prisutnost ovih domena objašnjava položaj RBP31 iznad dijagonale denaturirajućeg elektroforetskog gela. Također, primjećena je povišena ekspresija RBP31 u biljkama tretiranim

hladnoćom. U skladu s time, nedavno je pokazano da je upravo uloga RBP31 esencijalna za otpornost kloroplasta na stres izazvan hladnoćom (Kupsch i sur., 2012). Nadalje, u DisProt bazi podataka pronađen je zapis koji je značajno sličan RBP31. Svi cpRNP imaju N-terminalni signalni peptid koji regulira njihov unos u kloroplaste nakon sinteze. Naše analize pokazuju da je ovaj peptid neuređen što je u potpunom skladu s činjenicom da su IDP najčešće uključeni u signalizaciju te dodatno opravdava klasifikaciju RBP31 kao IDP (Wright i Dyson, 2015).

COP1-interacting protein-like protein (CIPLP) sadrži neuređenu C-terminalnu regiju od 250 aminokiselina. Iako za ovaj protein nije pronađena niti jedna poznata domena u *Conserved Domains* bazi podataka, njegov profil neuređenosti pokazuje da sadrži N-terminalnu uređenu regiju od 100 ak. Uočeno je da je CIPLP smješten između dijagonale i područja globularnih proteina na elektroforetskom gelu, a njegova ekspresija je u kontroli vrlo niska. Uloga ovog proteina nije poznata, no CIPLP proteini slični su *COP1-interacting* skupini proteina. Postoji mogućnost da CIPLP također interagira s COP1. COP1 i *COP1-interacting* proteini sudjeluju u ekspresiji gena važnih u biljnom razvitku u ovisnosti o svjetlosti, tzv. fotomorfogenezi (Yamamoto, 1998).

Jedan od najzanimljivijih nalaza IDP je *uncharacterized protein* (UP-DUF810) iz proteinske mrlje 26. Ekspresija UP-DUF810 bila je pojačana u uvjetima abiotskog stresa. Ovaj protein je iznimno dug (1117 aminokiselina) i sadrži domenu nepoznate funkcije DUF810 (*domain of unknown function 810*) dugu skoro 700 aminokiselina i karakterističnu za skupinu potpuno neokarakteriziranih biljnih proteina. Profil neuređenosti UP-DUF810 je kompleksan te sadrži niz neuređenih i uređenih regija, kao i velik broj neuređenih regija s veznom funkcijom. Nedavna bioinformatička studija je usporedbom po sličnosti s poznatim proteinima pokušala razjasniti funkciju osam domena nepoznate funkcije iz Pfam baze podataka, uključujući i DUF810. No, za DUF810 je pokazana samo djelomična sličnost sa jednim proteinom poznate funkcije, Sec6CT iz *S. cerevisiae*, koji sudjeluje u vezanju i fuziji vezikula s plazmatskom membranom (Goonesekere, Shipely i O'Connor, 2010). Jedan od mogućih razloga zašto je domena DUF810 neokarakterizirana jest donedavni manjak metoda prikladnih istraživanju neuređenih proteina i regija. Primjenom spektroskopskih metoda u budućim istraživanjima postoji mogućnost za determinaciju skupa strukturnih konformacija koje posjeduju proteini s ovom neuređenom domenom te bi se na taj način mogao stjeći uvid u njihovu staničnu funkciju.

Profil neuređenosti pokazuje da IDP *embryo defective 1241* (EMB1241) sadrži neuređene N- i C-terminalne regije duge 150 i 50 aminokiselina, te središnju uređenu domenu dugu 150 aminokiselina. EMB1241 je esencijalni dio translokacijskog kompleksa PAM u mitohondrijima. Na elektroforetskom gelu EMB1241 jedini se nalazio ispod dijagonale gela, što ukazuje na to da

je ovaj protein mobilniji u dentaurirajućim uvjetima nego u fiziološkim, te je izravan dokaz njegovog statusa kao IDP. No, izostanak ekspresije EMB1241 u abiotskom stresu nije moguće sa sigurnošću interpretirati zbog nedovoljnog razdvajanja točaka u gelovima te je nužno ponoviti te analize. Prema *Conserved Domains* bazi podataka, EMB1241 sadrži GrpE domenu (*nucleotide exchange factor GrpE*) koja se veže na Hsp70 (*Heat shock protein 70*) omogućavajući njegovu pravilnu funkciju i time prevenciju agregacije proteina u denaturirajućim uvjetima izazvanim abiotskim stresom. Tompa i Kovacs (2012) tvrde da su mnogi IDP šaperoni, što je u skladu sa šaperoninskom funkcijom proteina EMB1241.

PS II oxygen-evolving complex 1 (PSII OEC1) sadrži neuređene N- i C-terminalne regije te središnju uređenu domenu *photosystem II oxygen-evolving enhancer protein 1*. PSII OEC1 ima ulogu u ciklusu oksidacije vode stabilizacijom OEC kompleksa (Offenbacher i sur., 2013). U prijašnjem istraživanju ekvivalent PSII OEC1 (P12359) identificiran je kao IDP u *S. oleracea* i odgovara zapisu DP00188 u DisProt bazi podataka (Lydakis-Simantiris i sur., 1999). No, iako je PSII OEC1 prisutan u mrlji 29 koja se nalazi na dijagonali elektroforetskog gela, on je također identificiran u točki 35 iznad dijagonale te se njegova neuređenost ne može u potpunosti potkrijepiti položajem na denaturirajućem gelu.

Važno je napomenuti da među 4 IDP dokazana konsenzusom procjenitelja, a čija ekspresija nije mogla biti analizirana na razmatranim gelovima, postoje proteini čija je tercijska struktura visoko neuređena. Najneuređeniji od njih je protein *calcium-dependent lipid-binding domain-containing protein*. Prema *Conserved Domains* bazi podataka, ovaj protein sadrži N-terminalnu C2_SRC2 (*C2 Soybean Regulated by Cold 2-like*) domenu (regija 11 - 145). Na grafičkom profilu neuređenosti uočeno je da je jedino ta regija proteina uređena, dok je velik segment proteina neobično visoko neuređen naspram procjena za sve ostale otkrivene IDP. COR (*Cold Regulated*) proteini nisu dovoljno istraženi, no poznato je da jesu IDP te da imaju ulogu u stabilizaciji membrana u uvjetima hladnoće (Thalhammer i Hinch, 2014). Sličnost sekvence ovog proteina s skupim proteina COR objašnjava pojavu njegove ekspresije u biljkama tretiranim hladnoćom kao i otkriće da je IDP.

Uz detaljnu analizu pojedinih proteina, krajnji dokaz da je skup proteina identificiranih u ovom istraživanju obogaćen proteinima koji imaju ulogu u odgovor na abiotski stres su rezultati analiza anotacija odgovarajućih gena u *A. thaliana* prema rječniku Gene Ontology. Ove analize pokazuju i da su mnogi od otkrivenih proteina uključeni u još nepoznate funkcije te funkcije vezanja proteina, DNA i RNA. Veći dio proteina je lokaliziran u organelima gdje sudjeluju u signalizaciji te stabilizaciji proteina i proteinskih kompleksa. Funkcije IDP često se temelje na vezanju različitih vrsta liganda.

U zaključku, rezultati ovog istraživanja u skladu su s dosadašnjim studijama proteoma *A. thaliana* u abiotskom stresu kao i sa studijama IDP u drugim biljkama. S obzirom na vrlo mali broj dosad okarakteriziranih biljnih IDP, 12 noovotkrivenih IDP predstavlja značajan doprinos trenutnom znanju o ulozi neuređenosti proteina u biljaka.

Naše istraživanje prvo je ovakve vrste provedeno u Republici Hrvatskoj. Iako se područje istraživanja IDP razvija velikom brzinom, zbog nedavnog otkrića ovih proteina još uvijek postoji velika potreba za njihovom eksperimentalnom karakterizacijom na razini proteoma. Stoga, kao prvo proteomsko istraživanje IDP na biljnom organizmu ova studija predstavlja značajan pomak za čitavo područje IDP. IDP posjeduju važne uloge u stanicama, a svaki od otkrivenih proteina predstavlja metu budućih istraživanja koja će te uloge razjasniti. Razumijevanje uloga IDP u stanici ključno je za budući razvoj njihovih primjena u biotehnologiji. Želimo nastaviti takvu vrstu istraživanja u Republici Hrvatskoj i time doprinjeti aktualnoj i važnoj tematici.

6. ZAKLJUČAK

6. ZAKLJUČAK

- Tretmanom hladnoće 24h uspješno je izazvan kontrolirani abiotski stres u biljkama *A. thaliana*.
- Korištenjem metode po Csizmok i sur. (2006) uspješno su odvojeni IDP i globularni proteini iz biljnog tkiva.
- Dokazana je razlika u ekspresiji IDP u biljkama podvrgnutim abiotskom stresu u odnosu na kontrolnu skupinu biljaka.
- Masenom spektrometrijom identificirano je ukupno 29 različitih proteina koji uključuju proteine poznatih uloga u abiotskom stresu te proteine potpuno nepoznate uloge.
- Eksperimentalnim i bioinformatičkim analizama aminokiselinskog slijeda proteina identificiranih masenom spektrometrijom prema najstrožim je kriterijima otkriveno 12 potpuno novih IDP.
- Analiza anotacija odgovarajućih gena za identificirane proteine dokazuje njihovu ulogu u odgovoru na abiotski stres.

7. ZAHVALE

Od srca se zahvaljujemo doc. dr. sc. Dubravku Pavokoviću na stručnom vodstvu, strpljenju, pruženom znanju te svom uloženom trudu i vremenu.

Također, zahvaljujemo izv. prof. dr. sc. Biljani Balen, doc. dr. sc. Petri Peharec Štefanić i dr. sc. Petri Cvjetko na svim pruženim savjetima i ugodnom vremenu provedenom u laboratoriju.

8. LITERATURA

- Amme, S. (2006). Proteome analysis of cold stress response in *Arabidopsis thaliana* using DIGE-technology. *Journal of Experimental Botany*, 57(7), pp.1537-1546.
- Apitz, J., Schmied, J., Lehmann, M., Hedtke, B. and Grimm, B. (2014). GluTR2 Complements a hema1 Mutant Lacking Glutamyl-tRNA Reductase 1, but is Differently Regulated at the Post-Translational Level. *Plant and Cell Physiology*, 55(3), pp.645-657.
- Ashburner, M., Ball, C., Blake, J., Botstein, D., Butler, H., Cherry, J., Davis, A., Dolinski, K., Dwight, S., Eppig, J., Harris, M., Hill, D., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J., Richardson, J., Ringwald, M., Rubin, G. and Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nat Genet*, 25(1), pp.25-29.
- Bae, M., Cho, E., Choi, E. and Park, O. (2003). Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. *The Plant Journal*, 36(5), pp.652-663.
- Bernstein, F., Koetzle, T., Williams, G., Meyer, E., Brice, M., Rodgers, J., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977). The protein data bank: A computer-based archival file for macromolecular structures. *Journal of Molecular Biology*, 112(3), pp.535-542.
- Blum, H., Beier, H. and Gross, H. (1987). Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis*, 8(2), pp.93-99.
- BOUCHER, V., BUITINK, J., LIN, X., BOUDET, J., HOEKSTRA, F., HUNDERTMARK, M., RENARD, D. and LEPRINCE, O. (2010). MtPM25 is an atypical hydrophobic late embryogenesis-abundant protein that dissociates cold and desiccation-aggregated proteins. *Plant, Cell & Environment*, 33(3), pp.418-430.
- Buchan, D., Minneci, F., Nugent, T., Bryson, K. and Jones, D. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, 41(W1), pp.W349-W357.
- Burger, V., Gurry, T. and Stultz, C. (2014). Intrinsically Disordered Proteins: Where Computation Meets Experiment. *Polymers*, 6(10), pp.2684-2719.
- Candat, A., Paszkiewicz, G., Neveu, M., Gautier, R., Logan, D., Avelange-Macherel, M. and Macherel, D. (2014). The Ubiquitous Distribution of Late Embryogenesis Abundant Proteins across Cell Compartments in *Arabidopsis* Offers Tailored Protection against Abiotic Stress. *The Plant Cell*, 26(7), pp.3148-3166.
- Csizmok, V., Dosztanyi, Z., Simon, I. and Tompa, P. (2007). Towards Proteomic Approaches for the Identification of Structural Disorder. *CPPS*, 8(2), pp.173-179.
- Csizmok, V., Szöllösi, E., Friedrich, P. and Tompa, P. (2006). A Novel Two-dimensional Electrophoresis Technique for the Identification of Intrinsically Unstructured Proteins. *Molecular & Cellular Proteomics*, 5(2), pp.265-273.

- Dangl, J. and Jones, J. (2001). Plant pathogens and integrated defence responses to infection. *Nature*, 411(6839), pp.826-833.
- de Carbonnel, M., Davis, P., Roelfsema, M., Inoue, S., Schepens, I., Lariguet, P., Geisler, M., Shimazaki, K., Hangarter, R. and Fankhauser, C. (2010). The Arabidopsis PHYTOCHROME KINASE SUBSTRATE2 Protein Is a Phototropin Signaling Element That Regulates Leaf Flattening and Leaf Positioning. *PLANT PHYSIOLOGY*, 152(3), pp.1391-1405.
- Dosztanyi, Z., Csizmok, V., Tompa, P. and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*, 21(16), pp.3433-3434.
- Dosztanyi, Z., Meszaros, B. and Simon, I. (2009). Bioinformatical approaches to characterize intrinsically disordered/unstructured proteins. *Briefings in Bioinformatics*, 11(2), pp.225-243.
- Dyson, H. and Wright, P. (2005). Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology*, 6(3), pp.197-208.
- Fuxreiter, M. and Tompa, P. (2012). Fuzzy Complexes: A More Stochastic View of Protein Function. *Fuzziness*, pp.1-14.
- Galea, C., High, A., Obenauer, J., Mishra, A., Park, C., Punta, M., Schlessinger, A., Ma, J., Rost, B., Slaughter, C. and Kriwacki, R. (2009). Large-Scale Analysis of Thermostable, Mammalian Proteins Provides Insights into the Intrinsically Disordered Proteome. *J. Proteome Res.*, 8(1), pp.211-226.
- Gall, T., Romero, P., Cortese, M., Uversky, V. and Dunker, A. (2007). Intrinsic Disorder in the Protein Data Bank. *Journal of Biomolecular Structure and Dynamics*, 24(4), pp.325-341.
- Gooneseckere, N., Shipely, K. and O'Connor, K. (2010). The challenge of annotating protein sequences: The tale of eight domains of unknown function in Pfam. *Computational Biology and Chemistry*, 34(3), pp.210-214.
- Habchi, J., Tompa, P., Longhi, S. and Uversky, V. (2014). Introducing Protein Intrinsic Disorder. *Chem. Rev.*, 114(13), pp.6561-6588.
- He, B., Wang, K., Liu, Y., Xue, B., Uversky, V. and Dunker, A. (2009). Predicting intrinsic disorder in proteins: an overview. *Cell Res*, 19(8), pp.929-949.
- Hincha, D. and Thalhammer, A. (2012). LEA proteins: IDPs with versatile functions in cellular dehydration tolerance. *Biochemical Society Transactions*, 40(5), pp.1000-1003.
- Hunault, G. and Jaspard, E. (2010). LEAPdb: a database for the late embryogenesis abundant proteins. *BMC Genomics*, 11(1), p.221.

- Iakoucheva, L., Brown, C., Lawson, J., Obradović, Z. and Dunker, A. (2002). Intrinsic Disorder in Cell-signaling and Cancer-associated Proteins. *Journal of Molecular Biology*, 323(3), pp.573-584.
- Irar, S., Oliveira, E., Pagès, M. and Goday, A. (2006). Towards the identification of late-embryogenic-abundant phosphoproteome in Arabidopsis by 2-DE and MS. *PROTEOMICS*, 6(S1), pp.S175-S185.
- Ishida, T. and Kinoshita, K. (2007). PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Research*, 35(Web Server), pp.W460-W464.
- Jensen, M., Communie, G., Ribeiro, E., Martinez, N., Desfosses, A., Salmon, L., Mollica, L., Gabel, F., Jamin, M., Longhi, S., Ruigrok, R. and Blackledge, M. (2011). Intrinsic disorder in measles virus nucleocapsids. *Proceedings of the National Academy of Sciences*, 108(24), pp.9839-9844.
- Jensen, M., Kjaersgaard, T., Nielsen, M., Galberg, P., Petersen, K., O'Shea, C. and Skriver, K. (2010). The Arabidopsis thaliana NAC transcription factor family: structure–function relationships and determinants of ANAC019 stress signalling. *Biochem. J.*, 426(2), pp.183-196.
- Jones, D. and Cozzetto, D. (2014). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. *Bioinformatics*, 31(6), pp.857-863.
- Kami, C., Allenbach, L., Zourelidou, M., Ljung, K., Schütz, F., Isono, E., Watahiki, M., Yamamoto, K., Schwechheimer, C. and Fankhauser, C. (2013). Reduced phototropism in pks mutants may be due to altered auxin-regulated gene expression or reduced lateral auxin transport. *The Plant Journal*, 77(3), pp.393-403.
- Kosová, K., Vítámvás, P., Prášil, I. and Renaut, J. (2011). Plant proteome changes under abiotic stress — Contribution of proteomics studies to understanding plant stress response. *Journal of Proteomics*, 74(8), pp.1301-1322.
- Kovacs, D. and Tompa, P. (2012). Diverse functional manifestations of intrinsic structural disorder in molecular chaperones. *Biochemical Society Transactions*, 40(5), pp.963-968.
- Kovacs, D., Kalmar, E., Torok, Z. and Tompa, P. (2008). Chaperone Activity of ERD10 and ERD14, Two Disordered Stress-Related Plant Proteins. *PLANT PHYSIOLOGY*, 147(1), pp.381-390.
- Kupsch, C., Ruwe, H., Gusewski, S., Tillich, M., Small, I. and Schmitz-Linneweber, C. (2012). Arabidopsis Chloroplast RNA Binding Proteins CP31A and CP29A Associate with Large Transcript Pools and Confer Cold Stress Tolerance by Influencing Multiple Chloroplast RNA Processing Steps. *The Plant Cell*, 24(10), pp.4266-4280.

- LAEMMLI, U. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227(5259), pp.680-685.
- Li X, Romero P, Rani M, Dunker AK, Obradovic Z (1999) Predicting protein disorder for N-, C-, and internal regions. *Genome Inform Ser Workshop Genome Inform 10:30–40*
- Linding, R. (2003). GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Research*, 31(13), pp.3701-3708.
- Linding, R., Jensen, L., Diella, F., Bork, P., Gibson, T. and Russell, R. (2003). Protein Disorder Prediction. *Structure*, 11(11), pp.1453-1459.
- Lydakis-Simantiris, N., Hutchison, R., Betts, S., Barry, B. and Yocum, C. (1999). Manganese Stabilizing Protein of Photosystem II Is a Thermostable, Natively Unfolded Polypeptide †. *Biochemistry*, 38(1), pp.404-414.
- Malaney, P., Pathak, R., Xue, B., Uversky, V. and Davé, V. (2013). Intrinsic Disorder in PTEN and its Interactome Confers Structural Plasticity and Functional Versatility. *Scientific Reports*, 3.
- Marchler-Bauer, A., Derbyshire, M., Gonzales, N., Lu, S., Chitsaz, F., Geer, L., Geer, R., He, J., Gwadz, M., Hurwitz, D., Lanczycki, C., Lu, F., Marchler, G., Song, J., Thanki, N., Wang, Z., Yamashita, R., Zhang, D., Zheng, C. and Bryant, S. (2014). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, 43(D1), pp.D222-D226.
- Marín, M. and Ott, T. (2014). Intrinsic Disorder in Plant Proteins and Phytopathogenic Bacterial Effectors. *Chem. Rev.*, 114(13), pp.6912-6932.
- Midic, U., Oldfield, C., Dunker, A., Obradovic, Z. and Uversky, V. (2009). Protein disorder in the human diseasome: unfoldomics of human genetic diseases. *BMC Genomics*, 10(Suppl 1), p.S12.
- Mizianty, M., Peng, Z. and Kurgan, L. (2013). MFDp2. *Intrinsically Disordered Proteins*, 1(1), p.e24428.
- Mizianty, M., Stach, W., Chen, K., Kedarisetti, K., Disfani, F. and Kurgan, L. (2010). Improved sequence-based prediction of disordered regions with multilayer fusion of multiple information sources. *Bioinformatics*, 26(18), pp.i489-i496.
- Monastyrskyy, B., Kryshchak, A., Moulton, J., Tramontano, A. and Fidelis, K. (2013). Assessment of protein disorder region predictions in CASP10. *Proteins*, 82, pp.127-137.
- Nagai, S., Koide, M., Takahashi, S., Kikuta, A., Aono, M., Sasaki-Sekimoto, Y., Ohta, H., Takamiya, K. and Masuda, T. (2007). Induction of Isoforms of Tetrapyrrole Biosynthetic Enzymes, AtHEMA2 and AtFC1, under Stress Conditions and Their Physiological Functions in Arabidopsis. *PLANT PHYSIOLOGY*, 144(2), pp.1039-1051.
- Negrutiu, I., Beeftink, F. and Jacobs, M. (1975). Arabidopsis thaliana as a model system in somatic cell genetics I. Cell and tissue culture. *Plant Science Letters*, 5(5), pp.293-304.

Offenbacher, A., Polander, B. and Barry, B. (2013). An Intrinsically Disordered Photosystem II Subunit, PsbO, Provides a Structural Template and a Sensor of the Hydrogen-bonding Network in Photosynthetic Water Oxidation. *Journal of Biological Chemistry*, 288(40), pp.29056-29068.

Pavoković, D., Poljuha, D., Horvatić, A., Ljubešić, N., Hagege, D. and Krsnik-Rasol, M. (2011). Morphological and proteomic analyses of sugar beet cultures and identifying putative markers for cell differentiation. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 108(1), pp.111-119.

Pazos, F., Pietrosemoli, N., García-Martín, J. and Solano, R. (2013). Protein intrinsic disorder in plants. *Front. Plant Sci.*, 4.

Peng, K., Radivojac, P., Vucetic, S., Dunker, A. and Obradovic, Z. (2006). Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics*, 7(1), p.208.

Peng, Z., Yan, J., Fan, X., Mizianty, M., Xue, B., Wang, K., Hu, G., Uversky, V. and Kurgan, L. (2014). Exceptionally abundant exceptions: comprehensive characterization of intrinsic disorder in all domains of life. *Cellular and Molecular Life Sciences*, 72(1), pp.137-151.

Pietrosemoli, N., García-Martín, J., Solano, R. and Pazos, F. (2013). Genome-Wide Analysis of Protein Disorder in *Arabidopsis thaliana*: Implications for Plant Environmental Adaptation. *PLoS ONE*, 8(2), p.e55524.

Prilusky, J., Felder, C., Zeev-Ben-Mordehai, T., Rydberg, E., Man, O., Beckmann, J., Silman, I. and Sussman, J. (2005). FoldIndex(C): a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics*, 21(16), pp.3435-3438.

Ricachenevsky, F., Menguer, P. and Sperotto, R. (2013). kNACKing on heaven's door: how important are NAC transcription factors for leaf senescence and Fe/Zn remobilization to seeds?. *Front. Plant Sci.*, 4.

Romero, P., Obradovic, Z., Kissinger, C., Villafranca, J. and Dunker, A. (1997). Identifying disordered regions in proteins from amino acid sequence. *Proceedings of International Conference on Neural Networks (ICNN'97)*.

Romero, P., Obradovic, Z., Kissinger, C.R., Villafranca, J.E., Guilliot, S., Garner, E., and Dunker, A.K. (1998) Thousands of proteins likely to have long disordered regions. *Pacific Symposium on Biocomputing* 3:435-446.

Romero, P., Obradovic, Z., Li, X., Garner, E., Brown, C. and Dunker, A. (2000). Sequence complexity of disordered protein. *Proteins: Structure, Function, and Genetics*, 42(1), pp.38-48.

Sickmeier, M., Hamilton, J., LeGall, T., Vacic, V., Cortese, M., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V., Obradovic, Z. and Dunker, A. (2007). DisProt: the Database of Disordered Proteins. *Nucleic Acids Research*, 35(Database), pp.D786-D793.

Sun, X., Jones, W. and Rikkerink, E. (2012). GRAS proteins: the versatile roles of intrinsically disordered proteins in plant signalling. *Biochem. J.*, 442(1), pp.1-12.

Sun, X., Rikkerink, E., Jones, W. and Uversky, V. (2013). Multifarious Roles of Intrinsic Disorder in Proteins Illustrate Its Broad Impact on Plant Biology. *The Plant Cell*, 25(1), pp.38-55.

Szöllősi, E., Bokor, M., Bodor, A., Perczel, A., Klement, E., Medzihradzsky, K., Tompa, K. and Tompa, P. (2008). Intrinsic Structural Disorder of DF31, a Drosophila Protein of Chromatin Decondensation and Remodeling Activities. *J. Proteome Res.*, 7(6), pp.2291-2299.

Thalhammer, A. and Hinch, D. (2014). A mechanistic model of COR15 protein function in plant freezing tolerance: integration of structural and functional characteristics. *Plant Signaling & Behavior*, 9(12), p.e977722.

Tillich, M., Hardel, S., Kupsch, C., Armbruster, U., Delannoy, E., Gualberto, J., Lehwark, P., Leister, D., Small, I. and Schmitz-Linneweber, C. (2009). Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. *Proceedings of the National Academy of Sciences*, 106(14), pp.6002-6007.

Tolleter, D., Hinch, D. and Macherel, D. (2010). A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1798(10), pp.1926-1933.

Tompa, P. (2010). *Structure and function of intrinsically disordered proteins*. Boca Raton: Chapman & Hall/CRC Press.

Tompa, P. (2012). Intrinsically disordered proteins: a 10-year recap. *Trends in Biochemical Sciences*, 37(12), pp.509-516.

Turan, S. and Tripathy, B. (2014). Salt-stress induced modulation of chlorophyll biosynthesis during de-etiolation of rice seedlings. *Physiol Plantarum*, 153(3), pp.477-491.

Uversky, V. (2012). Intrinsically disordered proteins and novel strategies for drug discovery. *Expert Opin. Drug Discov.*, 7(6), pp.475-488.

Uversky, V. (2013). A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Science*, 22(6), pp.693-724.

Uversky, V. (2014a). Introduction to Intrinsically Disordered Proteins (IDPs). *Chem. Rev.*, 114(13), pp.6557-6560.

Uversky, V. (2014b). Intrinsically Disordered Proteins. *SpringerBriefs in Molecular Science*.

Uversky, V. (2014d). Proteins without unique 3D structures: Biotechnological applications of intrinsically unstable/disordered proteins. *Biotechnology Journal*, 10(3), pp.356-366.

Uversky, V. (2014e). The triple power of D³: Protein intrinsic disorder in degenerative diseases. *Front Biosci*, 19(2), p.181.

Uversky, V., Davé, V., Iakoucheva, L., Malaney, P., Metallo, S., Pathak, R. and Joerger, A. (2014c). Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases. *Chem. Rev.*, 114(13), pp.6844-6879.

Uversky, V., Gillespie, J. and Fink, A. (2000). Why are natively unfolded proteins unstructured under physiologic conditions?. *Proteins: Structure, Function, and Genetics*, 41(3), pp.415-427.

Uversky, V., Oldfield, C. and Dunker, A. (2008). Intrinsically Disordered Proteins in Human Diseases: Introducing the D² Concept. *Annual Review of Biophysics*, 37(1), pp.215-246.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R., Daughdrill, G., Dunker, A., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D., Kim, P., Kriwacki, R., Oldfield, C., Pappu, R., Tompa, P., Uversky, V., Wright, P. and Babu, M. (2014). Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.*, 114(13), pp.6589-6631.

Ward, J., McGuffin, L., Bryson, K., Buxton, B. and Jones, D. (2004b). The DISOPRED server for the prediction of protein disorder. *Bioinformatics*, 20(13), pp.2138-2139.

Ward, J., Sodhi, J., McGuffin, L., Buxton, B. and Jones, D. (2004a). Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *Journal of Molecular Biology*, 337(3), pp.635-645.

Wells, M., Tidow, H., Rutherford, T., Markwick, P., Jensen, M., Mylonas, E., Svergun, D., Blackledge, M. and Fersht, A. (2008). Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proceedings of the National Academy of Sciences*, 105(15), pp.5762-5767.

Wright, P. and Dyson, H. (2014). Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, 16(1), pp.18-29.

Xie, H., Vucetic, S., Iakoucheva, L., Oldfield, C., Dunker, A., Uversky, V. and Obradovic, Z. (2007). Functional Anthology of Intrinsic Disorder. 1. Biological Processes and Functions of Proteins with Long Disordered Regions. *J. Proteome Res.*, 6(5), pp.1882-1898.

Xue, B., Dunker, A. and Uversky, V. (2012). Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *Journal of Biomolecular Structure and Dynamics*, 30(2), pp.137-149.

Yamamoto, Y. (1998). Role of a COP1 Interactive Protein in Mediating Light-Regulated Gene Expression in Arabidopsis. *THE PLANT CELL ONLINE*, 10(7), pp.1083-1094.

Yang, H. (2001). The Signaling Mechanism of Arabidopsis CRY1 Involves Direct Interaction with COP1. *THE PLANT CELL ONLINE*, 13(12), pp.2573-2587.

9. SAŽETAK

KARLO SKUBE I INGRID TOMLJANOVIĆ

EKSPRESIJA I IDENTIFIKACIJA NOVIH PROTEINA DJELOMIČNO NEUREĐENE TERCIJARNE STRUKTURE (IDP) IZ BILJKE UROČNJAK (*ARABIDOPSIS THALIANA*) NAKON IZLAGANJA ABIOTIČKOM STRESU

Nativno neuređeni proteini i proteinske regije (*engl.* intrinsically disordered proteins/regions, IDP/IDR) pripadaju nedavno priznatoj skupini proteina koji su biološki aktivni usprkos nedostatku jasno definirane trodimenzionalne strukture. Za razliku od globularnih proteina, IDP u nativnom stanju ne posjeduju jedinstvenu stabilnu strukturu, već fluktuirajući skup strukturnih konformacija pogodan u staničnoj signalizaciji, stabilizaciji proteina i drugim adapterskim funkcijama. U eukariotima, IDP imaju važnu ulogu u odgovoru na abiotski stres. Abiotski stres u biljaka negativno utječe na rast i prinos, i u interesu čovjeka je da se saznaju molekularni mehanizmi njegovog djelovanja. Kako je otkrivanje IDP i njihove funkcije u biljkama u vrlo ranom stadiju, cilj ovog istraživanja bila je identifikacija i karakterizacija novih IDP u modelnoj biljci *Arabidopsis thaliana* izloženoj abiotskom stresu hladnoće. U izolaciji IDP korišteno je toplinsko razdvajanje pri 99°C pri čemu većina globularnih proteina precipitira. IDP i preostali globularni proteini su dodatno razdvojeni nativnom i denaturirajućom dvodimenzionalnom poliakrilamidnom elektroforezom s 8M urejom. Densitometrijska analiza proteina ukazala je da u biljaka tretiranih hladnoćom deset proteina nedostaje te da je došlo do povećanja ekspresije 4 proteina i smanjenja ekspresije jednog proteina u odnosu na kontrolu. Masenom spektrometrijom identificirano je ukupno 40 proteinskih mrlja, unutar kojih je pronađeno 29 različitih proteina. Primarna struktura identificiranih proteina analizirana je s tri najkvalitetnija bioinformatička alata za procjenu neuređenosti proteina: PrDOS, DISOPRED3 i MFDp2. Prema njihovom konsenzusu 12 proteina je klasificirano kao IDP. Prema rječniku Gene Ontology za većinu identificiranih proteina dokazana je uloga u abiotskom stresu i to vezivanjem na razne vrste bioloških molekula. Njihova lokalizacija upućuje na kloroplaste, mitohondrije i jezgru. S obzirom na vrlo mali broj dosad okarakteriziranih biljnih IDP, 12 noovotkrivenih IDP predstavlja značajan doprinos trenutnom znanju o ulozi neuređenosti proteina u biljaka. Naše istraživanje prvo je ovakve vrste provedeno u Republici Hrvatskoj. Kao prvo proteomsko istraživanje IDP na biljnom organizmu ova studija predstavlja značajan pomak za čitavo područje IDP.

Ključne riječi: adapterske molekule, šaperoni, procjenitelji neuređenosti, stres hladnoća

10. SUMMARY

KARLO SKUBE AND INGRID TOMLJANOVIĆ

EXPRESSION AND IDENTIFICATION OF NOVEL INTRINSICALLY DISORDERED PROTEINS IN *ARABIDOPSIS THALIANA* DURING ABIOTIC STRESS

Intrinsically disordered proteins and protein regions (IDP/IDR) are a recently recognized group of proteins which are biologically active despite their inherent lack of a well-defined three-dimensional structure. Unlike globular proteins, IDPs lack a unique and stable 3D structure in their native state and exist as fluctuating ensembles of conformations which are well-suited for cellular signaling, stabilization and other adapter functions. In eukaryotes, IDPs have important roles in abiotic stress response. Abiotic stress negatively affects plant growth and yield, and it is of global interest to unravel molecular mechanisms of plant abiotic stress responses. Considering that research on plant IDPs is still at a very early stage, the goal of this study was to identify and characterize novel IDPs in the model plant *Arabidopsis thaliana* during cold stress. Cellular extracts were enriched for IDPs by heat denaturation at 99°C which eliminated the majority of globular proteins. IDPs were further separated from remaining thermostable globular proteins using a native/8M urea 2D polyacrylamide gel electrophoresis. Densitometric expression analysis indicated that, when compared to the control group, cold-treated plants lacked expression of 10 proteins, while 4 were upregulated and one was downregulated. Mass spectrometry identified proteins from a total of 40 excised spots, which were shown to contain 29 different proteins. Sequence analysis of identified proteins was conducted using three currently best computational tools for disorder prediction: PrDOS, DISOPRED3 and MFDp2. A total of 12 proteins were classified as IDPs using a consensus-based prediction approach. Gene Ontology analyses demonstrated that the majority of identified proteins had roles in abiotic stress exerted through binding of different partner molecules. Cellular localization analysis of proteins pointed mainly towards chloroplasts, mitochondria and the nucleus. Considering the small number of experimentally characterized IDPs at present, the discovery of 12 new IDPs will greatly contribute to current knowledge on intrinsic disorder in plants. Our study is the first of its kind conducted in Croatia. As the first high-throughput study of IDPs in a plant proteome, our research will have a great impact on the whole IDP field.

Keywords: adapter molecules, chaperones, disorder predictors/disorder prediction software, cold stress