

SVEUČILIŠTE U ZAGREBU  
**FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA**  
ZAVOD ZA ELEKTRONIČKE SUSTAVE I OBRADBU  
INFORMACIJA

DIPLOMSKI RAD br.1047

**Predviđanje proteinskih interakcija  
rastavom proteinskih kompleksa**

Ivo Majić

Zagreb, rujan 2006.

*Posebno hvala Mili Šikiću za strpljivost, razumijevanje i mnoge ideje koje je podijelio sa mnom u vrijeme pripreme i izrade ovog rada.*

## **Sažetak**

Proteini su skupina temeljnih bioloških molekula, a svoje funkcije najčešće obavljaju kroz međusobnu interakciju. Kroz prošlost su evoluirale i od manjih su spajanjem nastajale veće. Poznavanjem temeljnih svojstava evolucije može se pretpostaviti i koje su nastale od kojih. Prirodan je zaključak da su proteini koji su u prošlosti evoluirali te se spojili u jedan i danas u stanju interagirati jedan s drugim. Taj zaključak vodi na predviđanje proteinskih interakcija. Kroz ovaj rad je načinjena baza podataka tako predviđenih interakcija te je pokazano da se one značajno podudaraju sa stvarnima.

## Sadržaj

1. Uvod .....	1
2. Osnovne značajke biologije i kemije proteina .....	3
3. Baze podataka o proteinima .....	6
3.1. Protein Data Bank (PDB) .....	7
4. Ciljevi rada i primijenjene metode .....	9
4.1. Baza stabla nastanka .....	9
4.2. Baza interakcija .....	13
5. Rezultati .....	16
6. Web sučelje i grafički prikaz mreže interakcija .....	20
6.1. Izgled grafičkog prikaza proteinskih interakcija .....	21
6.2. Rad grafičkog prikaza i njegova izvedba .....	23
6.3. Osvježavanje sadržaja .....	25
6.4. Grafički objekt .....	27
6.5. Točka, Linija, Tekst .....	29
6.6. Kontejner Grafičkih Objekata .....	31
7. Zaključak .....	34
8. Literatura .....	35
9. Dodatak A : Sadržaj priloženog optičkog medija .....	36

## **Popis oznaka i kratica**

eng                engleski

PDB              od eng. naziva Protein Data Bank

JSP                od eng. naziva Java Servlet Pages

## Popis tablica

Tablica 1. Popis aminokiselina i njihovih skraćenih oznaka.....	3
Tablica 2. Osnovni brožčani pokazatelji baze stabla nastanka i baze interakcija ..	16
Tablica 3. Rezultati usporedbe ispitnog skupa i baze stabla nastanka .....	18
Tablica 4: Popis mogućih akcija korisnika i odgovarajućih reakcija grafičkog prikaza .....	22
Tablica 5. Popis članova apstraktnog razreda <code>GrafickiObjekt</code> .....	28
Tablica 6. Popis članova razreda <code>Točka</code> (dodano u odnosu na razred <code>GrafickiObjekt</code> ) .....	29
Tablica 7. Popis članova razreda <code>Linija</code> (dodano u odnosu na razred <code>GrafickiObjekt</code> ) .....	30
Tablica 8. Popis članova razreda <code>Tekst</code> (dodano u odnosu na razred <code>GrafickiObjekt</code> ) .....	30
Tablica 9: Popis članova razreda <code>KontejnerGrafickihObjekata</code> .....	32

## Popis slika

Slika 1. Građa aminokiselina i njihovo vezanje u proteine.....	4
Slika 2. Struktura baze stabla nastanka.....	10
Slika 3. Algoritam određivanja stabla nastanka i popunjavanja istoimene baze podataka.....	12
Slika 4. Struktura baze interakcija .....	13
Slika 5. Algoritam generiranja proteinskih interakcija iz baze stabla nastanka .....	14
Slika 6. Izgled web sučelja.....	20
Slika 7. Izgled interaktivnog grafičkog prikaza proteinskih interakcija.....	21
Slika 8. Organizacija temeljnih objekata grafičkog prikaza mreže proteinskih interakcija .....	23

## 1. Uvod

Proteini su jedna od temeljnih skupina bioloških molekula. Prisutni su u svim funkcijama unutar stanice; oni su gradivne molekule, katalizatori reakcija, signalne molekule, molekule zadužene za motoriku, regulaciju i td. U posljednje vrijeme svjedoci smo snažnog razvoja biologije i genetike koja je usko povezana sa strukturom i interakcijama proteina.

Kroz proteklih tridesetak godina, kako su računala postajala snažnija, tako su i ona sve više uključivana u ovo područje istraživanja. Korištena su za izrade različitih računalnih modela, ali su značajno ubrzala i pojednostavila poslove poput organizacije i pretraživanja baza podataka. Zbog toga su se, među ostalim, pojavile i mnoge baze različitih informacija o proteinima.

Istraživanja proteina mogu se ugrubo podijeliti na istraživanja strukture, istraživanja proteinskih interakcija te istraživanja funkcije proteina. Ovaj diplomski rad nastao je kao nastavak rada na znanstveno-istraživačkom projektu, a pripada u područje istraživanja proteinskih interakcija.

Specifičan je po tome što predviđa interakcije korištenjem uobičajenih pretpostavki vezanih uz evoluciju proteina.

Prema dostupnim informacijama, takav projekt zasad nigdje nije proveden, a rezultati mogu značajno olakšati istraživanje u području proteinskih interakcija.

Rad je podijeljen u sedam poglavlja. U drugom se opisuje biološka i kemijska pozadina djelovanja proteina i nastanka proteinskih interakcija, a u trećem su iznesene informacije o javno dostupnim bazama podataka i njihovim svojstvima koja su iskorištena u ovom radu.

Okosnicu rada čini četvrto poglavlje koje detaljno opisuje ciljeve i metode koje su korištene za njihovo postizanje.

Kako stvar ne bi ostala na pukim pretpostavkama, dobiveni rezultati verificirani su korištenjem provjerenih informacija o laboratorijski utvrđenim proteinskim interakcijama, što je opisano petim poglavljem.



Dobiveni podaci su uobličeni u web sučelje kako bi bili dostupni svima koji su za njih zainteresirani. Najvažniji detalji njegove izvedbe opisani su u šestom poglavlju.

Sedmo poglavlje je zaključak i predstavlja kratak presjek kroz najvažnije rezultate.

## 2. Osnovne značajke biologije i kemije proteina

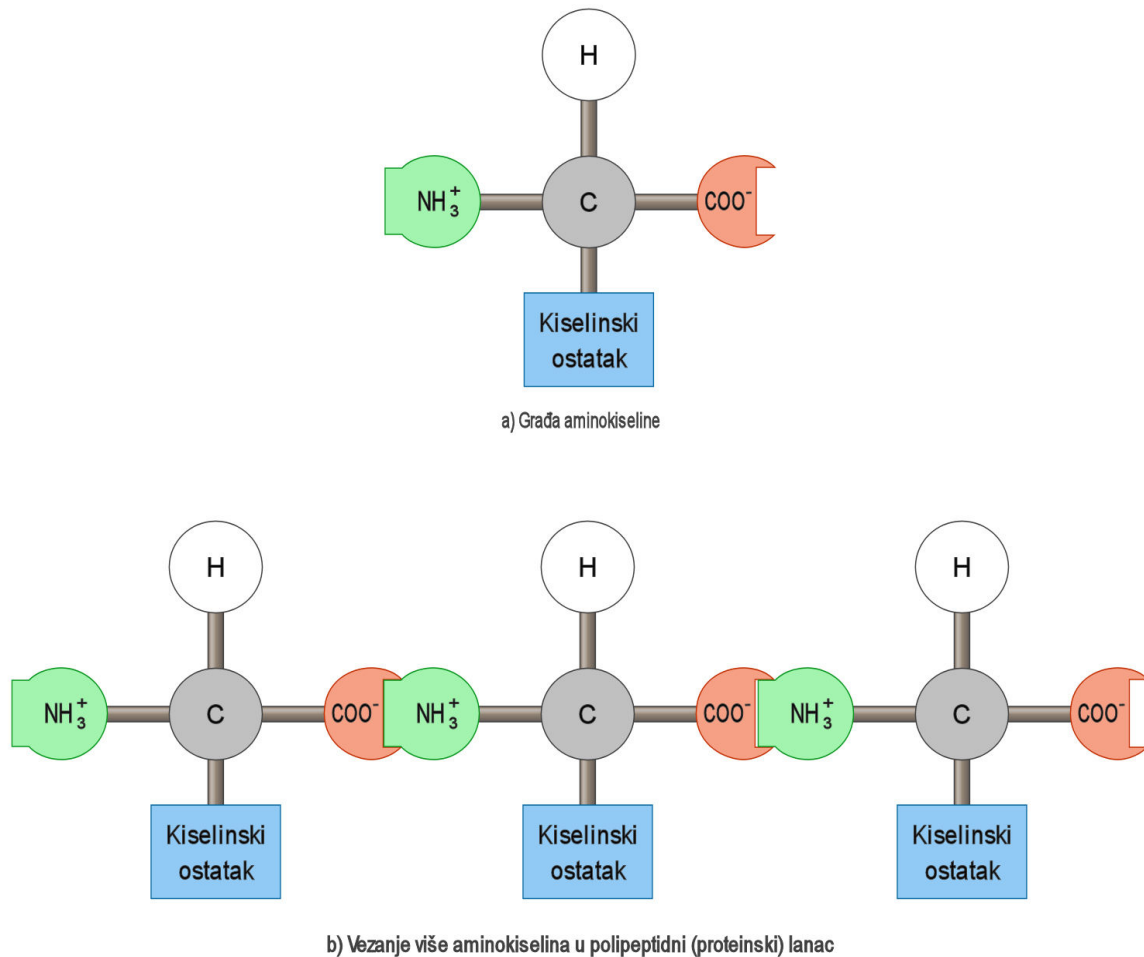
Proteini su biološke makromolekule i osnovni građevni elementi stanice. Sastavljeni su od kemijskih spojeva koji se zovu aminokiseline, a zanimljiva je činjenica da su svi dosad pronađeni proteini kombinacija samo dvadeset vrsta aminokiselina (Tablica 1) [8].

Tablica 1. Popis aminokiselina i njihovih skraćenih oznaka

Aminokiselina	Troslovna skraćenica	Jednoslovna skraćenica
Alanin	Ala	A
Valin	Val	V
Isoleucin	Ile	I
Leucin	Leu	L
Metionin	Met	M
Fenilalanin	Phe	F
Tirozin	Tyr	Y
Triptofan	Trp	W
Lizin	Lys	L
Arginin	Arg	R
Histidin	His	H
Glutamat	Glu	E
Aspargin	Asn	N
Glutamin	Gln	Q
Aspartat	Asp	D
Serin	Ser	S
Treonin	Thr	T
Cistein	Cys	C
Glicin	Gly	G
Prolin	Pro	P

Povezivanje aminokiselina moguće je zbog njihove posebne građe (Slika 1). Na prvom dijelu slike shematski je prikazan izgled molekule aminokiseline. Jasno

su vidljiva četiri osnovna dijela: središnje postavljen tzv.  $\alpha$  atom ugljika, peptidna ( $\text{NH}_3$ ) skupina, karboksilna ( $\text{COO}$ ) skupina i kiselinski ostatak. Kiselinski ostatak je dio koji uvjetuje ponašanje aminokiseline i to je jedini dio po kome se one međusobno razlikuju.



**Slika 1. Građa aminokiselina i njihovo vežanje u proteine**

Drugi dio slike pokazuje način stvaranja polipeptidnih lanaca iz pojedinih aminokiselina. Vidljivo je da se aminokiseline povezuju stvaranjem peptidne veze između peptidne skupine jedne i karboksilne skupine druge aminokiseline, te da opisani mehanizam spajanja omogućuje povezivanje lanaca proizvoljne duljine. Kombiniranjem jednog ili više polipeptidnih lanaca nastaju proteini.

Kemijski spojevi sastavljeni od više istovjetnih gradivnih komponenti zovu se polimeri. Ako se sastoje samo od dvije, trada su dimeri, od tri trimeri i td. Kod proteina, osim pojma polimera, postoji još pojam homomera i heteromera. Prvi

označava proteine koji se sastoje od međusobno jednakih lanaca, dok drugi označava one sastavljene od međusobno različitih.

Strukturu proteina uobičajeno je promatrati na četiri razine [8]:

- Primarna razina obuhvaća poredak aminokiselina u polipeptidnom lancu.
- Sekundarna razina dodatno u obzir uzima lokalno svijanje pojedinih dijelova polipeptidnog lanca. Uzrok svijanja je relativna blizina susjednih kiselinskih ostataka koji se međusobno privlače ili odbijaju svojim elektrostatskim poljima.
- Tercijarna razina, za razliku od sekundarne, razmatra i utjecaj međusobno udaljenih aminokiselina na svijanja polipeptidnog lanca.
- Kvartarna razina u obzir uzima i međudjelovanje pojedinih polipeptidnih lanaca.

Proteini u živom svijetu rijetko kada djeluju samostalno. Puno je češći primjer njihove međusobne interakcije, kao i interakcije proteina s drugim molekulama.

Razmatranje međudjelovanja dvaju proteinskih lanaca ili međudjelovanja dvaju proteina analizom svih mogućih interakcija između svih prisutnih atoma previše je složen zadatak i za najbolja računala današnjice. Čak i kad to ne bi bilo tako, nemoguće je simulirati još i interakciju obiju molekula s okolinom [2]. Zbog toga se u analizi proteinskih interakcija često rabe različiti modeli koji uspijevaju predstaviti neka svojstva promatranog procesa, ali ipak određene njegove karakteristike ostaju izvan njihova domašaja, što je razlog smanjenju pouzdanosti.

Uobičajen postupak analize proteinskih interakcija, dakle, uključuje pretpostavljanje rezultata na temelju izrađenog modela, te, zatim, njihovo verificiranje nekom vrstom izravnog laboratorijskog pokusa.

### 3. Baze podataka o proteinima

Danas je na Internetu dostupno obilje informacija vezanih uz proteine. One su podijeljene u nekoliko desetaka baza podataka, pri čemu dolazi do značajne redundancije. Spomenute baze mogu se grubo podijeliti na baze proteina i baze proteinskih interakcija, a cilj je ovog poglavlja spomenuti one među njima koje su bile neophodne za izradu ovog rada.

Baze proteina sadrže informacije o strukturi i funkciji pojedinih proteinskih lanaca. Najpoznatije među njima su

- Protein Data Bank (PDB) [5]
- Protein Information Resource (PIR)
- Swiss-Prot [7]

Baze proteinskih interakcija, kao što im sam naziv kaže opisuju interakcije među pojedinim proteinskim lancima, pri tome stavljajući u drugi plan samu strukturu pojedinih lanaca. Najznačajnije takve baze su:

- Database of Interacting Proteins (DIP)
- Muenich information center for Protein Sequences Protein Protein Interactions (MIPS PPI)

Kod baza proteinskih interakcija potrebno je istaknuti vrlo često postojanje više skupova podataka, od kojih je jedan skup pouzdanih podataka dobivenih nekim oblikom laboratorijskog pokusa, dok su preostali skupovi daleko manje pouzdani (ali, najčešće, i daleko veći) jer su dobiveni korištenjem nekog računalnog modela.

Za ovaj rad najznačajnija je PDB baza podataka te će biti detaljnije opisana slijedećim poglavljem.

### **3.1. Protein Data Bank (PDB)**

Ova baza zamišljena je kao jedinstveno mjesto na kojem će se nalaziti i obrađivati informacije o 3-D strukturi velikih molekula, kao što su proteini i nukleinske kiseline [5]. Ta struktura određuje se kristalografijom X-zrakama (eng. X-ray crystallography), nuklearnom magnetskom rezonancom (NMR) i elektronskim mikroskopiranjem. Budući da se ova baza, osim u akademske svrhe, koristi i kao izvor informacija u farmaceutskoj i biotehnološkim industrijama, pripada u skupinu najbolje održavanih.

Ondje pohranjeni podaci dostupni su javno i podijeljeni su u tzv. .pdb datoteke, od kojih svaka opisuje strukturu jedne makromolekule.

Datoteke sadrže preciznu informaciju o prostornim koordinatama svakog od atoma koji sačinjavaju pojedinu molekulu, pri čemu su atomi razvrstani u razne organizacijske jedinice, ovisno o vrsti molekule u kojoj se nalaze. To, primjerice, znači da će se za jedan atom nekog proteinskog lanca u odgovarajućoj datoteci naći njegove koordinate, naziv aminokiseline čiji je on dio, kao i njen položaj unutar lanca. Osim toga, svaka datoteka sadržava i informaciju o znanstvenom članku prema kome je formirana, ali i sve ostale informacije koje su neophodne za jednoznačan opis kemijske strukture odgovarajuće molekule.

Omogućeno je nekoliko načina pretraživanja od kojih su najznačajnije pretraga po PDB identifikatoru, potpuna pretraga tekstova dokumenata i pretraga po autoru datoteke. Osim toga, korištenjem sučelja ove baze moguće je izraditi i 3-D prikaz molekulske strukture.

Unatoč svemu navedenom, PDB baza je bila od presudne važnosti za izradu ovog diplomskog rada zbog jedne njene nespomenute značajke. Naime, ona u svakoj .pdb datoteci sadrži i sve potrebne informacije o proteinskim lancima analiziranog proteina. Tu se prvenstveno misli na slijedeće:

- Broj proteinskih lanaca unutar proteina
- Redoslijed aminokiselina u svakom od lanaca

Osim navedenog, kao dio PDB baze egzistira i tablica u kojoj su poznati proteinski lanci svih poznatih proteina podijeljeni u skupine (eng. *clusters*) [6]. U svakom od njih se nalaze lanci koji su barem 90% slični jedan drugome. Osim toga, svi lanci unutar jedne skupine sortirani su od „najboljeg“ prema „najgorem“ [6], što znači da se prvi lanac može smatrati predstavnikom cijele skupine, a svaki preostali se nalazi bliže vrhu popisa, što je sličniji njemu.

## 4. Ciljevi rada i primijenjene metode

Temelj cjelokupnog diplomskog rada slijedeće su pretpostavke:

- Kroz evoluciju života složeni proteini su nastajali od jednostavnijih
- Nakon nastanka složenog proteina njegovi sastavni dijelovi su i dalje egzistirali i u zasebnom obliku
- Mutacije koje su se događale nisu u bitnom izmijenile izvoran oblik proteina

Uzmemo li navedeno u obzir, logično je pretpostaviti da su proteini koji su se kroz evoluciju spojili u veći protein sposobni ostvariti međusobnu interakciju. Ova činjenica je iskorištena za predviđanje proteinskih interakcija.

Rješavanje diplomskog zadatka rada izvedeno je u nekoliko koraka:

1. Formiranje tzv. „baze stabla nastanka“ koja za svaki protein nastoji pretpostaviti proteine koji su ga u prošlosti sastavili.
2. Formiranje baze podataka o proteinskim interakcijama i dodavanje interakcija pretpostavljenih iz nastanka proteina
3. Izrada web sučelja prema objema bazama

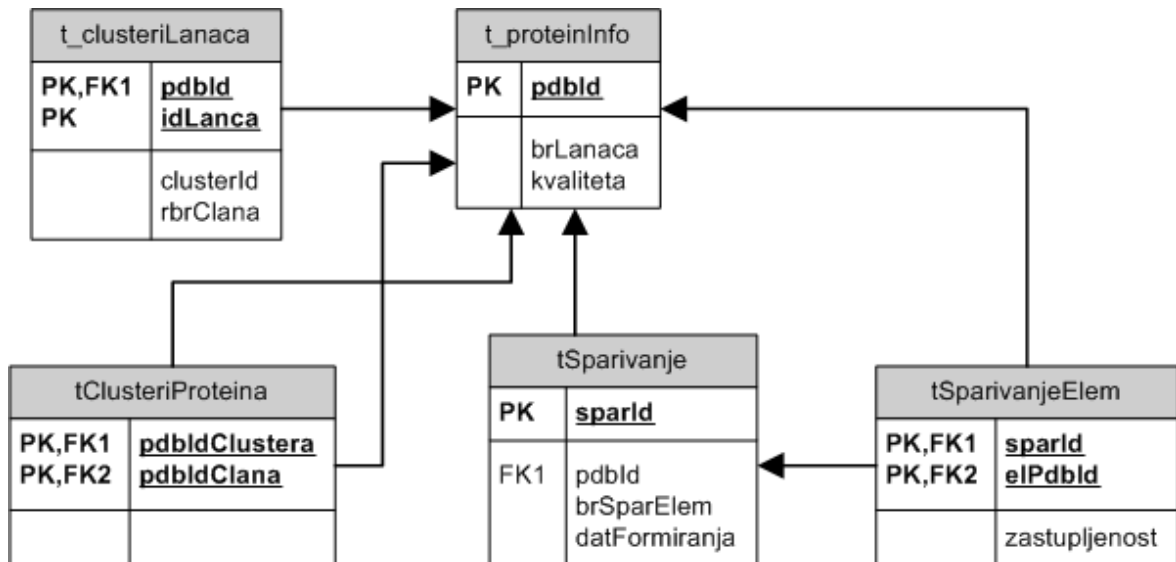
### 4.1. Baza stabla nastanka

Baza stabla nastanka je baza podataka koja opisuje evoluciju proteina. Ona nije zamišljena kao skup dokazanih informacija zbog nemogućnosti njegova stvaranja, već je zamišljena kao skup najboljih pretpostavki vezanih uz ovu tematiku.

Namjena joj je određivanje potencijalnih načina nastanka što većeg broja PDB proteina. Pri tome se pretpostavljalo da je zadani protein, ukoliko je složen, uvijek nastao kombinacijom dvaju ili više jednostavnih.

Struktura baze prikazana je na slijedećoj slici (Slika 2):





Slika 2. Struktura baze stabla nastanka

Tablice `t_clusteriLanaca` i `t_proteinInfo` popunjene su podacima izravno preuzetim iz PDB baze podataka.

Tablica `t_clusteriLanaca` sadržava podatke o skupinama proteinskih lanaca. Proteini se identificiraju svojim PDB identifikatorom (`pdbid`), a njihovi lanci i dodatnom jednoslovnom oznakom (`idLanca`). Za svaki lanac čuva se identifikator skupine kojoj pripada (`clusterId`) te broj koji ga smješta na određenu poziciju unutar nje (`rbrClana`). Skupine označavaju grupe međusobno sličnih proteinskih lanaca, a lanci koji su bliže početku grupe, tj. imaju manji broj `rbrClana` smatraju se „boljima“, odnosno sličnijima proteinskom lancu prema kome je skupina formirana. U kontekstu ovog rada svi proteinski lanci koji pripadaju istoj skupini smatrat će se jednakima.

Tablica `t_proteinInfo` sadržava popis PDB proteina te, za svaki protein (`pdbid`), njegovu duljinu (`brLanaca`) izraženu brojem lanaca koji se u njemu nalaze.

Nakon popunjavanja `t_clusteriLanaca` i `t_proteinInfo` popunjavane su i preostale tablice, za što je bio primijenjen poseban algoritam. Cilj je bio za što više složenijih proteina dobiti rastave na jednostavnije. Rastave su pohranjivane u tablicu `tSparivanje`, gdje je svaka rastava dobila svoj identifikator (`sparId`). Svaki od proteina koji je sudjelovao u rastavi (`elPdbId`) evidentiran je u tablicu

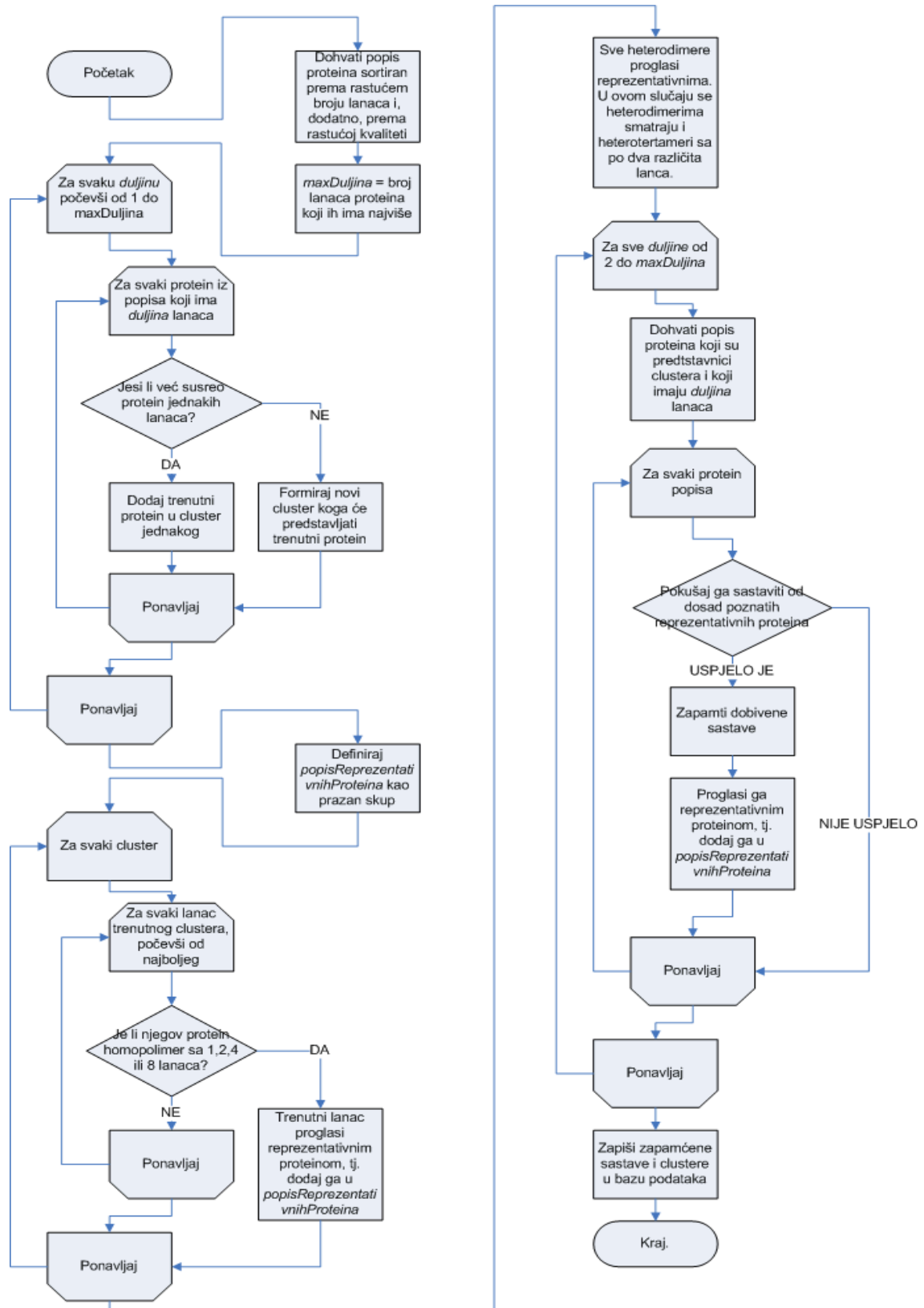
`tSparivanjeElem`. Ovdje je potrebno primijetiti da jedan protein u rastavi može biti prisutan i više puta, zbog čega je zastupljenost evidentirana posebnim stupcem tablice `tSparivanjeElem` (zastupljenost).

Tijekom izrade baze stabla nastanka pokazalo se da postoje grupe proteina koji su slični, tj. koji su sastavljeni od jednakog broja jednakih lanaca. Kako su oni u kontekstu ovog rada jednaki, nepotrebno je analizirati svakog od njih posebno. Zbog toga je uvedena tablica `tClusteriProteina`. Ona pohranjuje informacije o sličnim proteinima tako da, za svaki protein (`pdbIdClana`) za koga se utvrdi sličnost sa već analiziranim (`pdbIdClustera`), uvede po jedan zapis.

Algoritam kojim su popunjavane tablice `tSparivanje`, `tSparivanjeElem` i `tClusteriProteina` prikazan je slikom Slika 3.

Temeljna ideja pri njegovom kreiranju bila je početi od monomera te od njih kombinirati dimere. Nakon toga, od monomera i dimera trimere, zatim tetramere i td. Kako je struktura većine proteina u PDB bazi podataka dobivena postupkom tzv. kristalografije X-zrakama, potrebno je istaknuti jednu njenu značajku koja ima velik utjecaj na rezultate ovog rada. Radi se o pojavi „umnožavanja“ rezultata. Primjerice, ukoliko se kristalografiji podvrgne protein koji ima jedan lanac, lako je kao rezultat dobiti homodimer, homotetramer ili čak homooktamer. Jednako tako, u slučaju heterodimera, moguće je kao rezultat, primjerice, dobit tetramer sa po dva jednaka lanca. Ipak, pojava ovakvih rezultata može se povezati sa smanjenom kvalitetom izvođenja postupka kristalografije, zbog čega se u postupku određivanja reprezentativnih monomera koriste slijedeća pravila:

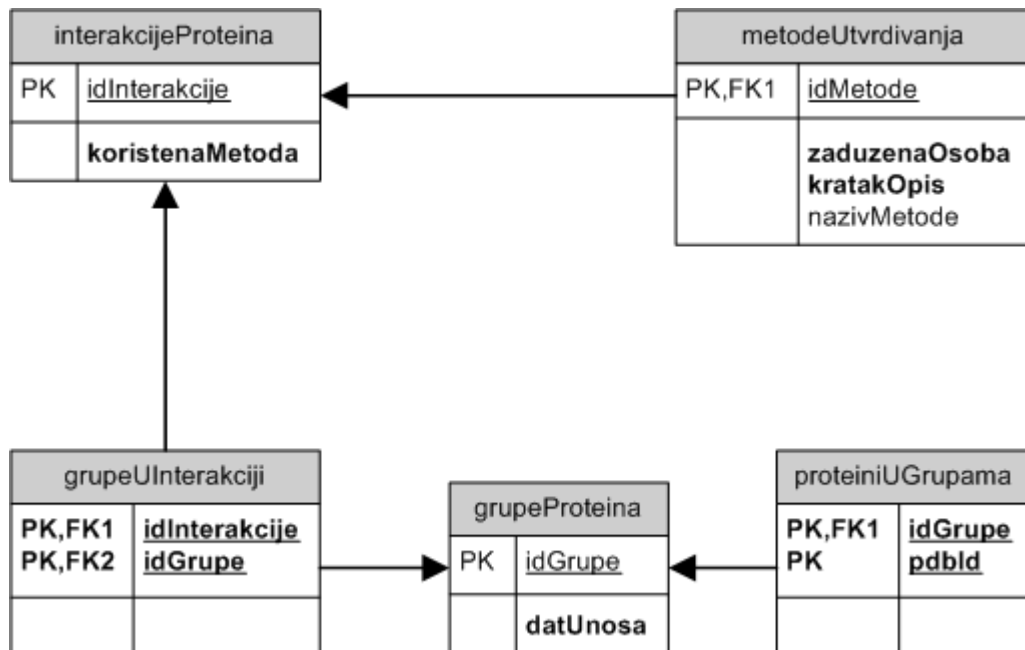
1. Ukoliko neka skupina lanaca u sebi sadrži lance koji pripadaju monomeru, tada se najbolji od tih monomera proglašava reprezentativnim.
2. Ukoliko skupina nema monomera, tada se reprezentativnim proglašava najbolji između lanaca koji pripadaju homodimerima, homotetramerima ili homooktamerima.
3. Ukoliko skupina nema ni takvih lanaca, ostaje bez predstavnika.



Slika 3. Algoritam određivanja stabla nastanka i popunjavanja istoimene baze podataka

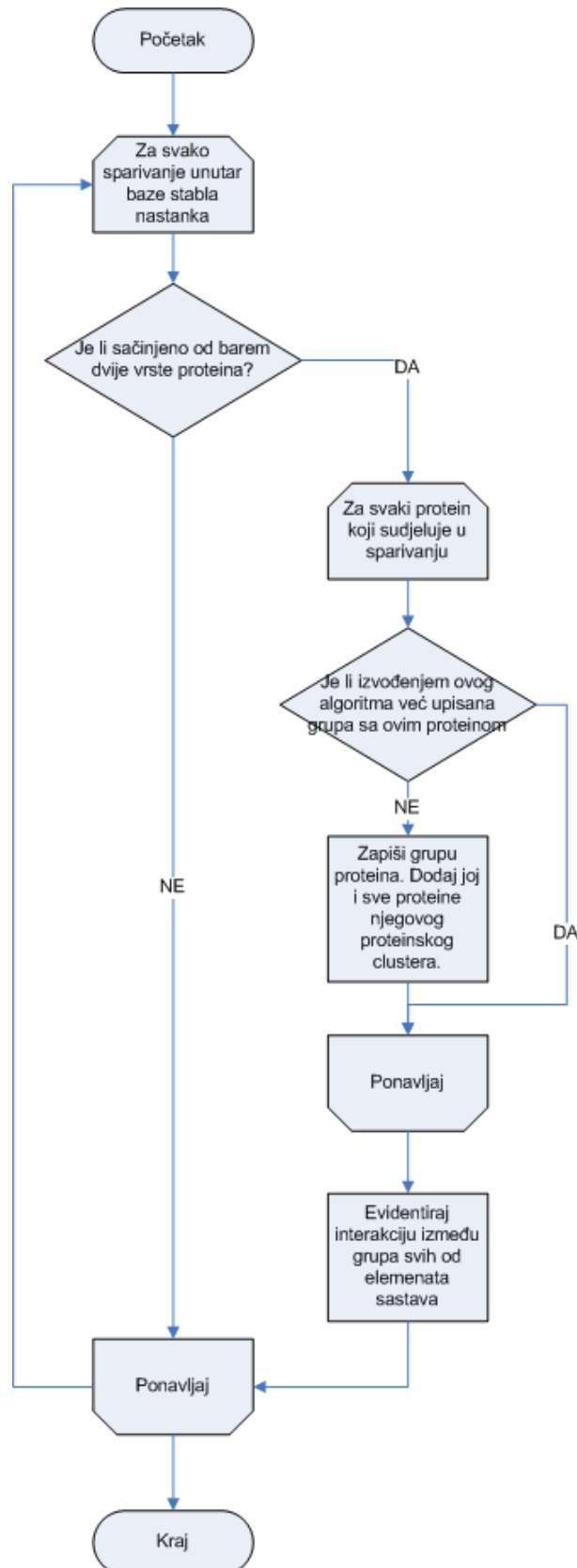
## 4.2. Baza interakcija

Baza interakcija je baza podataka namijenjena pohrani informacija o proteinskim interakcijama. Tu se prvenstveno misli na pohranu informacije o tome da neki proteini međusobno interagiraju. Baza je koncipirana s namjerom da u budućnosti bude ispunjavana podacima dobivenima i drugim istraživanjima. Njena struktura prikazana je na slijedećoj slici (Slika 4):



Slika 4. Struktura baze interakcija

Tablicama `grupeProteina` i `proteiniUGrupama` opisuju se grupe proteina i njihovi sastavni elementi (element može biti `<PDB identifikator>` ili `<PDB identifikator>:<lanci>`, pri čemu potonje označava samo dio `<PDB identifikator>` proteina. Uvođenje grupa potrebno je iz sličnog razloga kao i u bazi stabla nastanka. Naime, i ovdje se pokazuje da često postoji više proteina koji se u interakcijama ponašaju jednako. Kako bi se izbjeglo umnažanje broja evidentiranih interakcija, u bazu se zasebno evidentiraju grupe proteina te tada postoji potreba za samo jednim zapisom o interakciji.



Slika 5. Algoritam generiranja proteinskih interakcija iz baze stabla nastanka

Interakcije se definiraju između dviju ili više grupa, a evidentiraju se u tablicama interakcije proteina i grupe interakcije. Pri tome postojanje interakcije između grupe A i grupe B znači da interagira bilo koja kombinacija jednog proteina iz A i jednog proteina iz B.

Tablica metode utvrđivanja osigurava razlikovanje pojedinih metoda utvrđivanja postojanja interakcije. Primjerice, sve interakcije pretpostavljene ovim radom imat će jednu metodu utvrđivanja jer su utvrđene na jednak način.

Kao što je objašnjeno na početku ovog poglavlja, ideja rada je, na temelju stabla nastanka, pretpostaviti interakcije među proteinima. Algoritam kojim se to čini prikazan je na slici Slika 5.

U osnovi, algoritam svaku rastavu u bazi stabla nastanka proglašava interakcijom u bazi interakcija. Ukoliko je rastava realizirana s više od dva sastavna proteina, tada će i interakcija biti definirana kao interakcija više proteina.

## 5. Rezultati

Najvažniji brožčani pokazatelji baze stabla nastanka i baze interakcija dani su u slijedećoj tablici (Tablica 2):

Tablica 2. Osnovni brožčani pokazatelji baze stabla nastanka i baze interakcija

<b>Baza stabla nastanka</b>	
Ukupan broj analiziranih proteina:	33336
Broj proteinskih grupa:	896
Broj proteina u proteinskim grupama:	4882
Broj sparivanja:	13796
Broj proteina koji imaju barem jedno sparivanje:	16698
<b>Baza interakcija</b>	
Broj proteina:	2286
Broj proteinskih grupa:	1202
Broj interakcija između proteinskih grupa:	1774
Prosječan broj grupa po interakciji:	2.4047
Prosječan broj proteina po grupi:	1.9018
Procjena stvarnog broja interakcija:	64735

Kako bi se pokazala vjerodostojnost dobivenih podataka iskorišten je skup ispitnih podataka objavljen u [1]. Radi se o popisu tzv. „sigurnih“ interakcija potvrđenih nekim oblikom laboratorijskog pokusa. Jedina njegova nepraktičnost je u tome što se radi o popisu interagirajućih proteinskih segmenata, što će reći da u njemu, primjerice, stoji da međusobno interagiraju lanci A i B proteina  $p_{dbA}$  sa lancem C proteina  $p_{dbB}$  te njihovom interakcijom nastaju lanci D, E i F proteina  $p_{dbC}$ .

Kada se uspoređuju rezultati dobiveni ovim radom s onima iz ispitnog skupa, za svaki promatrani protein može se dobiti slijedeće:

- Jednak rezultat. Protein je u ispitnom skupu sastavljen od jednakih proteina kao i u bazi stabla nastanka, odnosno, preciznije, u bazi stabla nastanka postoji barem jedna rastava proteina jednaka onoj u ispitnom skupu.
- Polovičan rezultat. Za ovakav protein u bazi stabla nastanka ne postoji rastava opisana u prethodnom slučaju, ali postoji barem jedna rastava čiji je točno jedan protein dio rastave istog proteina u ispitnom skupu.
- Različit rezultat. U ovom slučaju u bazi stabla nastanka ne postoji rastava ni sa jednim sastavnim proteinom koji je dio ekvivalentne rastave u ispitnom skupu.

U slučaju polovičnog rezultata potrebno je razmotriti još tri dodatna slučaja vezana uz sastavni protein po kome se rastave baze stabla nastanka i ispitnog skupa međusobno razlikuju.

1. Bolji rezultat. U ovom slučaju je protein predložen bazom stabla nastanka „bolji“ ili jednako dobar onom koga daje ispitni skup. „Bolji“ od dva proteina je onaj komu je prosječna pozicija lanca unutar skupine lanaca veća.
2. Lošiji rezultat. Ovaj slučaj je obrat prethodnog, a uključuje i situaciju u kojoj baza stabla nastanka nema protein koji bi bio ekvivalentan onom u ispitnom skupu.

Kod različitog rezultata rastave baze stabla nastanka i ispitnog skupa se međusobno razlikuju za oba sastavna proteina pa se definicije pojmova „bolji“ i „lošiji rezultat“ redefiniiraju i proširuju.

1. Bolji rezultat. Oba sastavna proteina su „bolja“ kod baze stabla nastanka.
2. Neutralan rezultat. Jedan sastavni protein je „bolji“ kod baze stabla nastanka.



3. Lošiji rezultat. Oba sastavna proteina su „bolja“ kod ispitnog skupa. Ovo uključuje i situaciju u kojoj baza stabla nastanka nema niti jednu rastavu promatranog proteina.

Rezultati usporedbe prikazani su u slijedećoj tablici (Tablica 3):

Tablica 3. Rezultati usporedbe ispitnog skupa i baze stabla nastanka

Vrsta rezultata	Broj proteina koji ima takav rezultat	% od ukupnog broja proteina
<b>Jednak rezultat</b>	17	20
<b>Polovičan rezultat</b>	38	45
Bolji rezultat	30	36
Lošiji rezultat	8	9
<b>Različit rezultat</b>	29	35
Bolji rezultat	15	18
Neutralan rezultat	14	17
Lošiji rezultat	0	0
<b>Ukupno</b>	84	100

Bilo bi prirodno očekivati da se ovakve informacije traže u usporedbi s bazom interakcija, ali to nije učinjeno na taj način jer u bazi stabla nastanka, za razliku od baze interakcija, postoje rastave koja se sastoje od samo jednog proteina. One su nastale jer se u postupku generiranja rastava, ukoliko protein nije mogao biti u potpunosti sastavljen, formirao rastav koji sadržava samo one sastavne proteine koji su se mogli dokučiti, a umjesto svakog nedokučivog proteinskog lanca sadržava znak \*.

Uzme li se u obzir način na koji su definirane grupe rezultata, ispravni se mogu smatrati oni iz grupe jednakih, te bolji iz grupa polovičnih, odnosno različitih.

Na koncu ovog poglavlja, budući da u ispravne rezultate ulazi 65 od ukupno 84 ispitna proteina (77%), može se zaključiti kako dobivene baze podataka

sadržavaju pretpostavljene proteinske interakcije koje se u značajnoj mjeri podudaraju sa stvarnima.

## 6. Web sučelje i grafički prikaz mreže interakcija

Jedan od značajnih dijelova ovog diplomskog rada je i web sučelje prema stvorenim bazama podataka. Samo sučelje izrađeno je korištenjem JSP (akronim od eng. Java Servlet Pages) tehnologije, a izgled mu je prikazan na slijedećoj slici (Slika 6):

**Sveučilište u Zagrebu**  
**Fakultet elektrotehnike i računarstva**  
**Zavod za elektroničke sustave i obradbu informacija**

**Baza podataka o nastanku i interakcijama proteina**

PDB ID:   
 Traži

**Protein: 1BKF**

O proteinu    Stablo nastanka    **Interakcije**

**Popis interakcija u kojima protein susjeluje:**

ID interakcije (metoda utvrđivanja)	ID grupe	Proteini grupe
5738 (Metoda stabla nastanka)	8735	1AUE:A
	8736	1BKF
5850 (Metoda stabla nastanka)	7812	1FKI 1D7I 1BL4 1QPF 1D7J 1D6O 1QPL 1EYM 1D7H 1A7X
	8736	1BKF
5875 (Metoda stabla nastanka)	8736	1BKF
	8906	1AUI
6797 (Metoda stabla nastanka)	8736	1BKF
	9472	1J4R
6799 (Metoda stabla nastanka)	9473	1VJY
	7812	1BL4 1QPF 1D7J 1D6O 1QPL 1EYM 1D7H 1A7X
6800 (Metoda stabla nastanka)	8736	1BKF
	9473	1VJY

Dizajn i izrada: Ivo Majić, 2006

Slika 6. Izgled web sučelja

Od njegovih funkcija potrebno je istaći slijedeće:

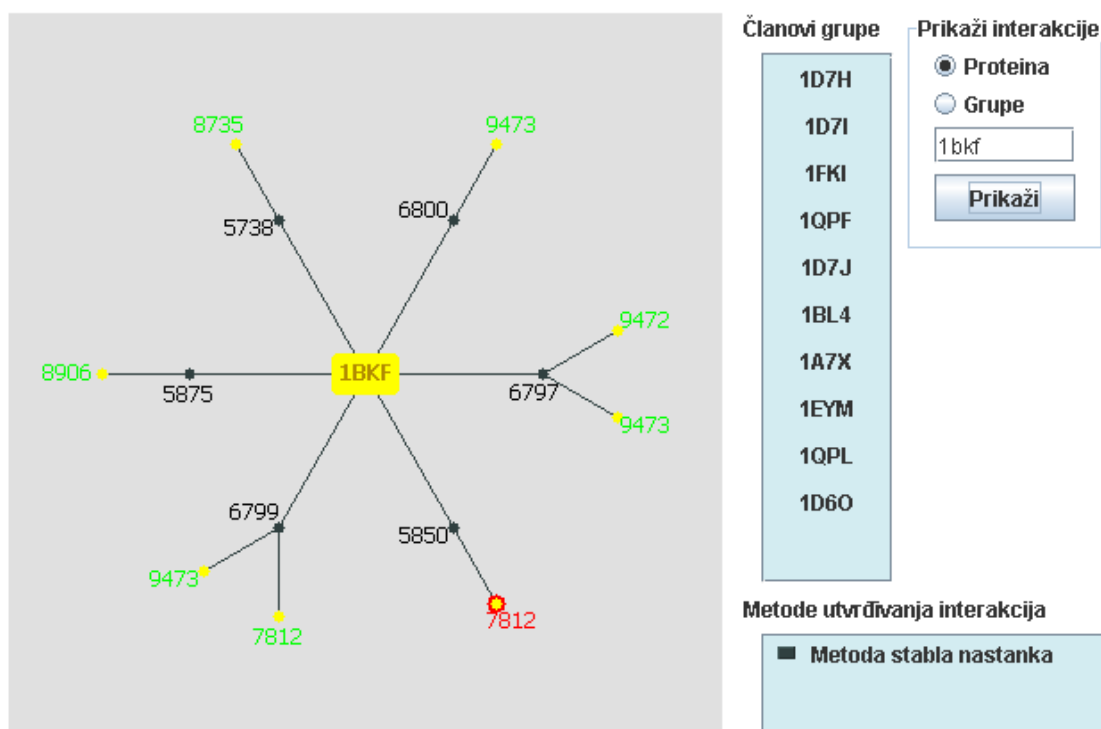
- mogućnost pretraživanja baza prema PDB identifikatoru proteina
- tablični prikaz poznatih interakcija (Slika 6. Izgled web sučelja), podataka iz baze stabla nastanka, kao i općenitih informacija za svaki protein, uz reference na druge baze podataka.

- prikaz statističkih informacija o popunjenosti baza podataka
- prikaz uputa za korisnike
- interaktivan grafički prikaz mreže proteinskih interakcija

Programerski je najzahtjevniji dio rada interaktivan grafički prikaz pa će o njemu više riječi biti u nastavku.

### 6.1. Izgled grafičkog prikaza proteinskih interakcija

Grafički prikaz je program napisan programskim jezikom Java i izveden u obliku Java Applet-a. Izvorno je namijenjen izvođenju u okviru web sučelja, ali zbog svoje realizacije lako može biti uklopljen u proizvoljnu statičku aplikaciju. Njegovi izgled prikazan je slijedećom slikom (Slika 7):



Slika 7. Izgled interaktivnog grafičkog prikaza proteinskih interakcija

Na samom prikazu vidljive su slijedeće komponente:

- Dva popisa: jedan koji sadržava popis proteina trenutno selektirane grupe (u daljnjem tekstu lista proteina) i drugi koji sadržava popis korištenih metoda (u daljnjem tekstu lista metoda).
- Slika interakcija. Predstavljena je grafom sastavljenim od tri vrste elemenata: točke, linije i teksta. Svaka točka na slici ima pripadajući tekst, a postoje dvije vrste točaka: male okrugle i velika četvrtasta (u sredini). Element koji ona opisuje naziva se stožerni element (na slici to je protein 1BKF). Točka grupe je svaka od točaka koje se nalaze na obodu grafa (na slici je selektirana jedna takva točka – točka grupe 7812), a ona predstavlja grupu proteina koja sudjeluje u interakciji. Točka interakcije je točka koja povezuje stožerni element i grupe svake od interakcija u kojima on sudjeluje. Uz nju je napisan broj - identifikator pripadne interakcije.
- Okvir za izravan izbor proteina ili grupe proteina nalazi se uz desni rub prikaza i korisniku omogućuje izravan izbor proizvoljnog proteina, odnosno proizvoljne grupe proteina.

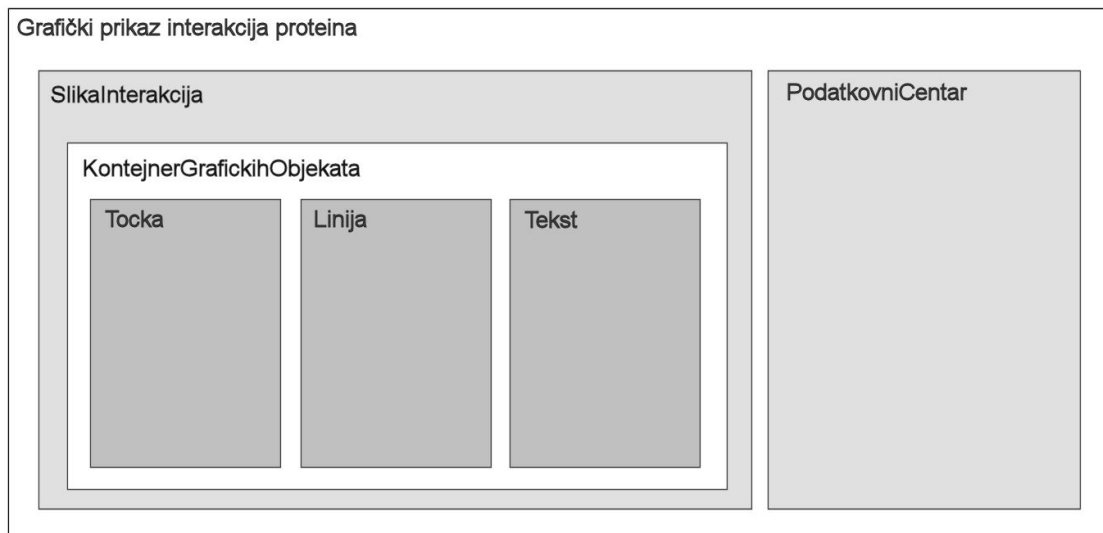
Grafički prikaz osmišljen je kao web komponenta, zbog čega se sve korisničke akcije nastojalo realizirati mišem. Popis akcija i odgovarajućih reakcija prikaza dan je slijedećom tablicom:

**Tablica 4: Popis mogućih akcija korisnika i odgovarajućih reakcija grafičkog prikaza**

Akcija	Reakcija
Klik	<p><u>Selektiranje objekta</u>. Ukoliko je klik izveden nad linijom, točkom ili tekstom selektiraju se svi elementi koji pripadaju u istu selekcijsku skupinu. U primjeru na slici neovisno je da li korisnik klikne na tekst „7812“ ili na pripadnu točku. U oba slučaja selektirat će se oba grafička objekta, jer zajedno čine jednu selekcijsku skupinu.</p> <p>Ukoliko se selektira selekcijska skupina koja označava grupu proteina (primjerice, kao na slici) selektiranje prati i odgovarajuće ispunjavanje liste proteina.</p>
Dvostruki klik	<p><u>Izbor novog stožernog elementa</u>. Ukoliko se dvostruki klik izvede nad nazivom ili točkom proteina ili grupe proteina, tada dotični protein, odnosno dotična grupa postaju stožerni elementi.</p> <p>Isto vrijedi i u slučaju kada se dvostruki klik izvede nad sadržajem liste proteina.</p>

## 6.2. Rad grafičkog prikaza i njegova izvedba

Grafički prikaz realiziran je primjenom objektno orijentirane programske paradigme i kao takav sačinjen je od objekata. Nekoliko njih je ključno te će biti detaljnije opisano u nastavku. Slijedeća slika donosi njihovu osnovnu organizaciju (Slika 8):



**Slika 8. Organizacija temeljnih objekata grafičkog prikaza mreže proteinskih interakcija**

`SlikaInterakcija` je kontrola izvedena iz kontrole `JPanel` i služi iscrtavanju grafičkog prikaza proteinskih interakcija, a svu komunikaciju prema trima osnovnim objektima (točka, linija i tekst) obavlja putem kontejnera `KontejnerGrafickihObjekata`.

Njegova je uloga povezivanje osnovnih grafičkih objekata u smislenu cjelinu. To je, posebice, nužno za slijedeće tri funkcionalnosti:

- Grupna selekcija. Prilikom selekcije, primjerice, teksta, nužno je selektirati i objekt na koji se taj tekst odnosi. Da bi to moglo biti učinjeno selektirani objekt mora biti „svjestan“ objekata u svojoj okolini koji ovise o njegovom svojstvu selektiranosti. Iluzija takve svjesnosti za ispunjenje ove zadaće postiže se kroz `KontejnerGrafickihObjekata` u kome su pohranjene informacije o tzv. „seleksijskim grupama“. Kada neki objekt iz grupe kontejneru prijavi svoje stanje selektiranosti, svim se objektima grupe

pošalje naredba za selekciju, dok se svima ostalima pošalje naredba za deselekciju.

- Pokretanje akcije. Ova funkcionalnost usko je povezana s grupnom selekcijom, ali uključuje dodatan problem. Primjerice, kada se izvodi dvoklik na objekt koji nije tekst, potrebno je „znati“ koji je tekst (identifikator) „povezan“ s tim objektom, kako bi se sadržaj teksta mogao iskoristiti za promjenu trenutno prikazanih podataka. Povezanost s grupnom selekcijom vidljiva je u jednakosti između „seleksijskih“ i „akcijskih“ grupa, a razlika je u tome što je kod akcija za svaku grupu potrebno dobiti i njen identifikator. To je postignuto dodavanjem reference za svaku od „seleksijskih“ grupa unutar kontejnera `KontejnerGrafickihObjekata`. Referenca pokazuje na grafički objekt čiji se identifikator smatra identifikatorom grupe.
- Iscrtavanje objekta. Kod iscrtavanja nužno je voditi računa o redoslijedu. Zbog toga je ono riješeno u slojevima. Svaki grafički objekt „zna“ u kojem sloju (slojevima) treba iscrtati koji svoj dio, a kontejner ih proziva jednog po jednog za svaki sloj u kome imaju što za iscrtati, sve dok svi ne iscrtaju sve. Prvi sloj u kome se iscrtava ima identifikator 0, drugi 1 i td, sve do posljednjega koji ima identifikator 100. Pseudokod algoritma iscrtavanja prikazan je slijedećim odsječkom:

```

slijSloj = 0;
dok (slijSloj <= 100) {
    trenSloj = slijSloj;
    slijSloj = 100;

    za svaki (GrafickiObjekt o) {
        trenSlijSloj = o.iscrtaj(trenSloj);
        ako (trenSlijSloj >= 0 i
            trenSlijSloj < slijSloj) {
            slijSloj = trenSlijSloj;
        }
    }
}

```

PodatkovniCentar je komponenta kojoj je glavna i jedina zadaća komunikacija s bazom podataka. Sve komponente prikaza koje trebaju dohvaćati podatke iz baze to čine putem podatkovnog centra.

### 6.3. Osvježavanje sadržaja

Prilikom osvježavanja sadržaja grafičkog prikaza, poziva se jedna od inačica metode `osvjeziPodatke()` unutar centra `PodatkovniCentar`. Ona inicira dohvat novih podataka iz baze, pravilno popunjava unutarnje memorijske strukture Centra, obavještava objekt `SlikaInterakcija`, te objekte zadužene za čuvanje sadržaja liste proteina i liste korištenih metoda o potrebi za ažuriranjem prikaza. To čini pozivom metode `podaciOsvjezeni()` kod objekta `SlikaInterakcija`, odnosno metoda `promijeniPodatke()` kod svake od lista.

Unutar metode `SlikaInterakcija.podaciOsvjezeni()` provodi se slijedeći postupak:



1. Određuju se središnja točka *SlikeInterakcija* (*sredTocka*) i najmanja udaljenost od središnje točke do ruba slike (*maxRadijus*).
2. Dohvaća se stožerni element (*stozerniElement*), kreiraju se objekti:

```
to = Tocka (sredTocka);
te = Tekst (stozerniElement, to, POR_SREDINA);
```

3. Objekti *to* i *te* se pohranjuju u kontejner te se formira nova selekcijska grupa koju čine ta dva objekta. Identifikator grupe je objekt *te*.
4. Dohvaća se broj interakcija stožernog elementa (*brInterakcija*) i određuje veličina  $prirastKuta = 2 * \pi / brInterakcija$ .
5. Dohvaćaju se sve interakcije, numeriraju se brojevima počevši od 0 (*interRbr*), te se za svaku ponavljaju koraci 6.-15.?
6. Dohvaća se identifikator interakcije (*interId*), te se određuju koordinate točke interakcije (*interTocka*) kao koordinate točke koja je  $maxRadijus/2$  udaljena od točke *sredTocka*, a prema osi x zatvara kut od  $kut = interRbr * prirastKuta$ .
7. Kreiraju se objekti:

```
toInter = Tocka (interTocka);
teInter = Tekst (interId, toInter, POR_PO_STRANI);
liInter = Linija(to, toInter);
```

8. Objekti *toInter*, *liInter* i *teInter* dodaju se u kontejner, te se formira nova selekcijska grupa u koju se dodaju sva tri. Identifikator novostvorene selekcijske grupe je *teInter*.
9. Dohvaća se broj grupa unutar trenutne interakcije (*brGrupa*), te se definiraju veličine  $pocetniKutGrupa = kut - \pi/4$  i  $prirastKutaGrupe = \pi/2 / (brGrupa - 1)$ .

10. Dohvaćaju se sve grupe trenutne interakcije, osim grupe kojoj pripada trenutni stožerni element. Dohvaćene grupe se numeriraju počevši sa brojem 0 (*rbrGrupe*). Za svaku grupu ponavljaju se koraci 11.-15.?

11. Dohvaća se identifikator grupe (*idGrupe*), te se određuju koordinate točke grupe (*grupaTocka*) kao koordinate točke koja je  $\text{maxRadijus}/4$  udaljena od točke *interTocka*, a prema osi x zatvara kut od  $\text{kutGrupe} = \text{pocetniKutGrupa} + \text{rbrGrupe} * \text{prirastKutaGrupe}$ .

12. Kreiraju se slijedeći objekti:

```
toGrupa = Tocka (grupaTocka);
teGrupa = Tekst (idGrupe, toGrupa, POR_PO_STRANI);
liGrupa = Linija(toInter, toGrupa);
```

13. Objekti *toGrupa*, *liInter* i *teGrupa* dodaju se u kontejner.

14. Objekti *toGrupa* i *teGrupa* dodaju se u novu selekcijsku skupinu. Njen identifikator je *teGrupa*.

15. Objekt *liGrupa* dodaje se u selekcijsku skupinu *teInter*.

Prilikom osvježavanja podataka *SlikaInterakcija* puni *KontejnerGrafickihObjekata* novim objektima koristeći se metodama `dodajObjekt()`. Prethodno se svi postojeći objekti kontejnera uklanjaju pozivom metode `ukloniSveObjekte()`.

#### 6.4. Grafički objekt

Tri su osnovna elementa grafičkog prikaza. To su:

- Točka
- Linija
- Tekst

Svaki od njih ima pripadni (istoimeni) razred izveden iz osnovnog apstraktnog razreda `GrafickiObjekt` (sadržaj apstraktnog razreda prikazan je tablicom Tablica 5).

Tablica 5. Popis članova apstraktnog razreda `GrafickiObjekt`

Član	Opis
<code>Color dohvatiBoju ()</code>	Dohvaća boju kojom se iscrtava objekt.
<code>Color dohvatiBojuPozadine ()</code>	Dohvaća boju pozadine korištenu za iscrtavanje objekta.
<code>int dohvatiReferentniSloj ()</code>	Dohvaća redni broj sloja u kome se objekt iscrtava. Ako se objekt iscrtava u više slojeva, rezultat je redni broj prvog sloja u kome objekt iscrtava neki svoj element.
<code>Point dohvatiReferentnuTocku ()</code>	Dohvaća referentnu točku objekta. Referentna točka je točka u odnosu na koju se definiraju sve veličine u objektu.
<code>postaviBoju (Color boja)</code>	Postavlja boju objekta.
<code>postaviBojuPozadine (Color boja)</code>	Postavlja boju pozadine.
<code>postaviReferentniSloj (int sloj)</code>	Postavlja referentni sloj. Ako se objekt iscrtava u više slojeva, zadržat će se odnos između slojeva. To znači da će se kod objekta koji se iscrtava u 1. i 2. sloju, ukoliko se referentni sloj postavi na 4, iscrtavanje odvijati u 4. i 5. sloju.
<code>postaviReferentnuTocku (Point tocka)</code>	Postavlja referentnu točku objekta.
<code>int iscrtaj (Graphics2D g, int sloj)</code>	Iscrtava dio objekta koji se nalazi u zadanom sloju. Ukoliko je time završeno iscrtavanje, vraća se -1. Ako iscrtavanje nije završeno, vraća se redni broj sloja u kome se još nešto treba do crtati.
<code>registrirajKontejner (KontejnerGrafickihObjekata k)</code>	Ovu metodu poziva kontejner kako bi objekt bio „svjestan“ kome pripada. Grafički objekt u jednom trenutku može pripadati samo jednom kontejneru. Ponovni poziv ove metode rezultira zanemarivanjem prethodno registriranog kontejnera.
<code>obradiAkcijuMisa (MouseEvent akcija)</code>	Obrađuje proslijeđenu akciju miša i po potrebi obavještava kontejner o vlastitom

	selektiranju ili izvršenoj akciji pokretanja.
<b>postaviSelektiranost</b> ( <b>boolean</b> <i>selektiran</i> )	Postavlja stanje selektiranosti grafičkog objekta.
<b>postaviIdentifikator</b> ( <b>Object</b> <i>identifikator</i> )	Postavlja identifikator grafičkog objekta.
<b>boolean</b> <b>dohvatiSelektiranost()</b>	Dohvaća stanje selektiranosti grafičkog objekta.
<b>Object</b> <b>dohvatiIdentifikator()</b>	Dohvaća identifikator grafičkog objekta.

### 6.5. Točka, Linija, Tekst

Točka, Linija i Tekst tri su osnovna razreda izvedena iz razreda `GrafickiObjekt`. Izmjene njihovog sadržaja u odnosu na `GrafickiObjekt` prikazane su slijedećim tablicama (Tablica 6, Tablica 7, Tablica 8):

Tablica 6. Popis članova razreda `Točka` (dodano u odnosu na razred `GrafickiObjekt`)

Član	Opis
<b>Točka</b> ( <i>Point tocka</i> )	Kreira točku pretpostavljene boje i pretpostavljenog oblika, bez naslova.
<b>registrirajIncidentnuLiniju</b> ( <i>Linija l</i> )	Ovu metodu poziva <code>Linija</code> u trenutku kada joj se kao jedna od krajnjih točaka zada trenutna <code>Točka</code> . Na taj način <code>Točka</code> postaje svjesna linija koje „izlaze“ iz nje. Potrebno je istaći da <code>Točka</code> može imati više incidentnih linija.
<b>registrirajTekst</b> ( <i>Tekst naslov</i> )	Ovu metodu poziva <code>Tekst</code> u trenutku kada mu se kao točka koju naslovljava zada trenutna <code>Točka</code> . Ovim mehanizmom točka postaje „svjesna“ svog naslova.
<b>tekstIzmijenjen()</b>	Ovom metodom <code>Tekst</code> obavještava razred <code>Točka</code> da je promijenio svoje postavke. U ovoj metodi <code>Točka</code> provjerava njegove postavke i eventualno se prilagođuje novim zahtjevima.

Tablica 7. Popis članova razreda `Linija` (dodano u odnosu na razred `GrafickiObjekt`)

Član	Opis
<code>Linija</code> ( <i>Točka prva</i> , <i>Točka druga</i> )	Kreira liniju pretpostavljene boje između točaka <i>prva</i> i <i>druga</i> .

Tablica 8. Popis članova razreda `Tekst` (dodano u odnosu na razred `GrafickiObjekt`)

Član	Opis
<code>int</code> <code>POR_SREDINA</code>	Konstanta poravnanja. Specificira poravnanje u kome se tekst nalazi točno u sredini točke. Točka je tada četvrtastog oblika i prilagođava se dimenzijama teksta.
<code>int</code> <code>POR_PO_STRANI</code>	Konstanta poravnanja. Specificira poravnanje u kome se tekst nalazi u okolici točke na automatski određenoj najboljoj poziciji.
<code>Tekst</code> ( <i>Object tekst</i> , <i>Točka refTočka</i> , <i>int poravnanje</i> , <i>Graphics2D grafika</i> )	Kreira objekt <code>Tekst</code> pretpostavljenog fonta, boje i veličine. Proslijeđeni parametar <i>tekst</i> koristi se kao tekstni sadržaj, a u grafičkom se prikazu ispisuje vrijednost koju vraća njegovo <code>toString()</code> svojstvo. <i>refTočka</i> označava točku uz koju se tekst veže, a za vrijeme izvođenja ovog konstruktora poziva se njena <code>registrijTekst(this)</code> metoda. Ovisno o postavkama parametra <i>poravnanje</i> , novostvoreni objekt se poravnava prema objektu <i>refTočka</i> , a u određenim slučajevima inicira i njenu izmjenu. <i>grafika</i> je parametar koji se koristi za određivanje dimenzija iscrtanih objekta, a potrebno ga je postaviti na vrijednost dobivenu korištenjem <code>getGraphics()</code> metode komponente na kojoj se vrši iscrtavanje.
<code>int</code> <code>dohvatiDimenzijuX()</code>	Dohvaća širinu teksta.
<code>int</code> <code>dohvatiDimenzijuY()</code>	Dohvaća duljinu teksta.
<code>int</code> <code>dohvatiPoravnanje()</code>	Dohvaća trenutno poravnanje teksta.
<code>postaviFont(Font f)</code>	Postavlja font kojim će tekst biti ispisivan.
<code>Font</code> <code>dohvatiFont()</code>	Dohvaća font za ispis teksta.

Osim navedenog u tablicama, potrebno je dodatno razmotriti mehanizam kojim se usklađuju točka i njen naslov...

Tekst u svom konstruktoru dobiva objekt `Tocka` kojeg naslovljava. U tom istom konstruktoru on poziva metodu `tocka.registrirajTekst(this)` koja pak interno poziva `tocka.tekstIzmijenjen(this)`. Izvođenjem potonje metode `tocka` dolazi do informacija o poravnanju teksta, njegovim dimenzijama i njegovoj referentnoj točki. Ukoliko je poravnanje postavljeno na `POR_SREDINA`, točka će se iscrtati kao pravokutnik nešto većih dimenzija od dimenzija teksta. Središnja točka tog pravokutnika odgovarat će referentnoj točki objekta `tocka`. U to vrijeme objekt `Tekst` ispisivat će se također sa referentnom točkom (koju je preuzeo iz objekta `tocka`) u središtu.

U slučaju poravnanja `POR_PO_STRANI` objekt `Tocka` ne mora prilagođavati svoj izgled objektu `Tekst`, već se iscrtava u obliku kružića, objekt `Tekst` sam se poravnava se u odnosu na referentnu točku objekta `Tocka`.

## 6.6. Kontejner Grafičkih Objekata

`KontejnerGrafickihObjekata` je razred namijenjen organizaciji grafičkih objekata u smislenu cjelinu.

Njegova je uloga detaljnije opisana u poglavlju 6.2 (Rad grafičkog prikaza), a odnosi se na slijedeće tri funkcionalnosti:

- Grupnu selekciju,
- Pokretanje akcije i
- Iscrtavanje objekata.

Za efikasno zadovoljavanje nabrojениh funkcija potrebno je osigurati slijedeće mogućnosti:

- Određivanje identifikatora selekcijske grupe iz identifikatora jednog od njenih članova.
- Određivanje svih članova selekcijske grupe iz njenog identifikatora
- Postojanje liste svih grafičkih objekata

Osim navedenog, u obzir je potrebno uzeti i činjenicu da je moguća pojava dviju jednakih grupa proteina u dvije (ili više) različitih interakcija, zbog čega takvu grupu nije moguće identificirati putem njenog prirodnog identifikatora jer se u jednom trenutku na grafičkom prikazu može naći više instanci jednake grupe.

Kako bi se riješio taj problem svim grafičkim objektima unutar `KontejneraGrafickihObjekata` uvedeni su novi cjelobrojni identifikatori. Prvi grafički objekt koji dođe u kontejner ima identifikator 0, drugi 1 i td. Prilikom pražnjenja kontejnera identifikatori se ponovno počinju dodjeljivati od 0. Budući da je potrebno osigurati mogućnost prevođenja cjelobrojnog identifikatora u objekt koga identificira, lista grafičkih objekata (*listaObjekata*) načinjena je u obliku efikasno pretražive (po identifikatoru) liste parova (identifikator, objekt).

Problem određivanja članova grupe iz njenog identifikatora riješen je uvođenjem liste *listaSelGrupa*. Ona je načinjena u obliku efikasno pretražive (po identifikatoru) liste parova (identifikator, skup objekata grupe). Na koncu, obrnuto pretraživanje (traženje identifikatora za nekog člana selekcijske grupe) zahtjeva uvođenje dodatne liste koja svaki član grupe povezuje s identifikatorom grupe (*listaClanovaSelGrupa*).

Sučelje objekta `KontejnerGrafickihObjekata` prikazano je u slijedećoj tablici (Tablica 9):

**Tablica 9: Popis članova razreda `KontejnerGrafickihObjekata`**

Član	Opis
<code>KontejnerGrafickihObjekata()</code>	Konstruktor.
Integer <code>dodajObjekt</code> ( <code>GrafickiObjekt o</code> )	Dodaje objekt kontejneru. Nakon dodavanja kontejner poziva metodu <code>o.registrirajKontejner(this)</code> kako bi objekt postao „svjestan“ svoje pripadnosti kontejneru. Metoda vraća identifikator dodijeljen objektu.
<code>ukloniSveObjekte()</code>	Prazni kontejner.
<code>selektiranObjekt</code> ( <code>GrafickiObjekt o</code> )	Ovu metodu poziva objekt u trenutku kada ga korisnik selektira. Time obavještava kontejner da je potrebno

	selektirati preostale objekte iz iste selekcijske skupine.
<b>pokrenutObjekt</b> (GrafickiObjekt <i>o</i> )	Ovu metodu poziva objekt u trenutku kada korisnik nad njime izvede akciju pokretanja (primjerice dvoklik mišem). Time obavještava kontejner da je potrebno pokrenuti akciju pridruženu selekcijskoj skupini kojoj objekt pripada.
<b>obradiAkcijuMisa</b> (MouseEvent <i>akcija</i> )	Ovu metodu poziva objekt <i>SlikaInterakcija</i> u trenutku kada korisnik izvede neku akciju mišem iznad nje. Kontejner u ovoj metodi poziva istoimenu metodu svakog od svojih sastavnih grafičkih objekata.
<b>dodajSelekcijskuGrupu</b> (Tekst <i>identifikator</i> )	Dodaje novu selekcijsku grupu kontejneru. Identifikator selekcijske grupe može biti samo objekt tipa <i>Tekst</i> .
<b>dodajElementUSelekcijskuGrupu</b> (GrafickiObjekt <i>o</i> , Tekst <i>idSelGrupe</i> )	Dodaje novi element u postojeću selekcijsku grupu. Ako grupa ne postoji dodavanje se ignorira.
<b>iscrtajPrikaz</b> (Graphics <i>g</i> )	Is crtava sve svoje objekte na objekt <i>g</i> .



## 7. Zaključak

Model predviđanja proteinskih interakcija razvijen u sklopu ovog diplomskog rada u svojoj je srži vrlo jednostavan. Počiva na temeljnim pretpostavkama evolucije proteina, ali i živog svijeta općenito.

Tu, prije svega, mislim na slijedeće:

- Kroz evoluciju života složeni proteini su nastajali od jednostavnijih
- Nakon nastanka složenog proteina njegovi sastavni dijelovi su i dalje egzistirali i u zasebnom obliku
- Mutacije koje su se događale nisu u bitnom izmijenile izvoran oblik proteina

Ove pretpostavke, iako bi ih se moglo nazvati trivijalnim, dovode do skupa proteinskih interakcija koje su temeljan rezultat ovog diplomskog rada. Njihova podudarnost s interakcijama dokazanim u laboratoriju je iznimno velika, što ne samo da potvrđuje vjerodostojnost postavljene teorije, nego dobivenim rezultatima daje i određenu vrijednost.

Iako se na temelju njih ne može deciderano izjaviti: „Da, ova dva proteina uistinu interagiraju!“, smatram da oni mogu predstavljati i predstavljaju vrijedan skup ulaznih podataka za biologe i kemičare koji pomoću njih mogu, u potrazi za proteinskim interakcijama, barem suziti početni izbor kandidata.

---

Ivo Majić

## 8. Literatura

- [1] Chen Rong, et al U: "A Protein-Protein Docking Benchmark", *PROTEINS: Structures, Function, and Genetics*, vol. 52, pp. 88–91, 2003.
- [2] Eric Alm, Adam P. Alkin U: "Biological networks", *Current Opinion in Structural Biology*, vol. 13, pp. 193–202, 2003.
- [3] Irene M.A. Nooren, Janet M. Thornton, U: "Diversity of protein-protein interactions", *The EMBO Journal*, vol. 22, no. 14, pp. 3486–3492, 2003.
- [4] Shoshana J. Wodak, Jeffrey Castura, Chris Orsi U: "Integrative bioinformatics: making sense of the networks", *Elsevier*, vol. 1, no. 2, pp. 179–187, 2003.
- [5] H.M. Berman, et al. U: "The Protein Data Bank", *Nucleic Acids Research*, vol. 28, pp. 235–242, 2000.
- [6] S Interneta: "The Protein Data Bank", <http://www.rcsb.org/pdb/>, 3. rujna 2006.
- [7] S Interneta: "Uniprot Knowledge Base – User Manual", <http://www.expasy.org/sprot/userman.html>, 3. rujna 2006.
- [8] Lodish et al: "Molecular cell biology 5th edition", *W.H. Freeman and company*, 2003.

## 9. Dodatak A : Sadržaj priloženog optičkog medija

Priloženi optički medij ima slijedeću strukturu direktorija. U zagradama je naveden sadržaj svakog od njih.

- Diplomski (tekst diplomskog rada u PDF formatu)
- Podaci (kopija sadržaja svih baza podataka korištenih za izradu diplomskog rada sa pripadnim skriptnim datotekama za instalaciju)
- Baze (izvoran kod potreban za generiranje baze stabla nastanka i baze interakcija)
- Web (izvoran kod web sučelja)
- Dokumentacija (JavaDoc dokumentacija koda iz prethodna tri direktorija, u HTML obliku)

Upute za korištenje sadržaja svakog od nabrojanih direktorija nalaze se u svakom od njih u tekstualnoj datoteci „CITAJ.ME“.