

Razvoj metaprediktora za utvrđivanje neuređenosti proteina

Master rad iz oblasti
Bioinformatika
Matematički fakultet

Una Stanković
una_stankovic@yahoo.com

18. jul 2018.

Sažetak

U ovom radu biće predstavljen proces kreiranja metaprediktora za analizu neuređenosti proteina. Najpre će biti objašnjeni svi pojmovi vezani za biološke osnove ovog problema, a potom i sam razvojni proces iz informatičkog ugla. Sav softver primenjen pri izradi rada je korišćen isključivo u akademske svrhe.

Sadržaj

1	Uvod	2
2	Biološke osnove	2
2.1	Proteini	2
2.1.1	Funkcije i osobine proteina	2
2.1.2	Struktura proteina	3
2.1.3	Savijanje proteina	7
2.1.4	Denaturacija proteina	7
2.1.5	Neuređenost proteina	8
3	Prediktori	8
3.1	SPOT-Disorder prediktor	8
4	Zaključak	8
5	Literatura	8
	Literatura	8

1 Uvod

Neuređenost proteina

2 Biološke osnove

U ovoj sekciji biće ukratko predstavljene biološke osnove neophodne za razumevanje rada i motivacije koja stoji iza određenih njegovih elemenata.

Najpre, biće opisano šta su proteini koje su njihove osnovne funkcije i kakva im je struktura. Potom, biće opisana svaka od struktura ponao-sob, uz priložen grafički prikaz istih. Na kraju, posebno će biti opisani neuređeni proteini, njihova uloga i uzroci koji mogu dovesti do njihove pojave.

2.1 Proteini

Proteini (grč. *protos* - "zauzimam prvo mesto") su najvažniji sastojak žive materije i utiču na brojnost i raznolikost živih bića. Specifičnost proteina je tolika da svaka biljna i životinjska vrsta ima svoje proteine, a kod viših organizama oni se razlikuju i na individualnom nivou. Broj proteina u živim bićima je ogroman, a kao primer uzmimo *E.coli* sa 3000 i čoveka sa 5 miliona proteina. [7]

Proteini i peptidi su linearni polimeri izgrađeni od 22 L – *aminokiseline*¹ koje se javljaju u prirodi i povezani su peptidnim vezama. [2] Poznavanje i razumevanje funkcija proteina je esencijalno u istraživanju bilo kog biološkog procesa, sa posebnim naglaskom na oboljenja ljudi, s obzirom da se mnoga od njih mogu pojaviti zbog funkcionalnih mutacija. [4]

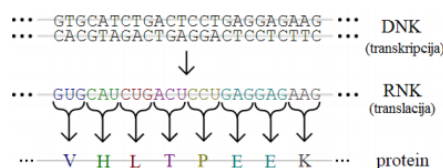
Svaki molekul proteina nastaje u ćeliji živog organizma. Redosled aminokiselina u proteinskom lancu određen je redosledom nukleotida u deoksiribonuleinskoj kiselini (DNK). Informacija o broju, vrsti i redosledu aminokiselina zapisana je na delovima DNK koji se nazivaju geni. Svakoj trojci nukleotida u genu na osnovu genetskog koda jedinstveno je pridružena po jedna aminokiselina. U procesu genske ekspresije, posredstvom glasničke (eng. messenger) ribonukleinske kiseline (RNK) i transportne (eng. transfer) RNK, enkodirana informacija iz DNK se prevodi u niz aminokiselina u proteinskom lancu. [4] Proces sinteze proteina predstavlja *centralnu dogmu molekularne biologije*, čiji se grafički prikaz može videti na slici 1.

2.1.1 Funkcije i osobine proteina

Proteini su biološki najaktivniji molekuli sa velikim brojem esencijalnih uloga koje se dele na:

- dinamičke, od kojih su najvažnije:
 1. transportna - prenos molekula (kiseonika, gvožđa, lipida) i hormona od mesta sinteze do mesta delovanja,
 2. biološka - regulacija metaboličkih procesa u ćeliji, kontrola i regulacija transkripcije gena i translacija,

¹ L – *aminokiseline* su one sa levom prostornom konfiguracijom, analogno, postoje i D – *aminokiseline*, sa desnom



Slika 1: Prikaz centralne dogme molekularne biologije [4]

3. katalizatorska - biološka katalizacija ²,
 4. zaštitna - keratin, koagulacija krvi,
 5. održavanje zapremine tečnosti u organizmu,
- strukturne
 1. obezbeđivanje čvrstine i elastičnosti organa,
 2. davanje oblika organizmu,
 3. izgradnja strukturnih elemenata ćelije i
 4. bitna uloga u kontraktilnim i pokretnim elementima organizma.

Postoji nekoliko osobina koje karakterišu proteine, to su:

- izgradnja kompleksnih jedinjenja sa različitim supstancama po principu strukturne komplementarnosti i
- visoka osetljivost na različite agense koji ih denaturišu ³. Neki od agenasa su: visoka temperatura, pritisak, mehaničko tretiranje, dejstvo kiselina, baza, organskih rastvarača, materija, itd. [7]

2.1.2 Struktura proteina

Proteini imaju specifičnu trodimenzionalnu strukturu koja se formira presavijanjem polipeptidnog lanca na različite načine. Unutrašnjost takve strukture ima visoku gustinu, pa takav lanac ne dopušta promene u sastavu i zahteva prisustvo aminokiselina tačno određene veličine. Uobičajena raspodela aminokiselina u proteinima je daleko od ravnomerne. Neke aminokiseline se javljaju mnogo češće od ostalih, tako se, na primer, leucin pojavljuje devet puta više od triptofana.

Niska aminokiselina sadrži sve potrebne informacije kako bi se formirala trodimenzionalna struktura proteina, za koju se smatra da je najstabilnija. Proteinsku strukturu održavaju različite vrste kovalentnih i nekovalentnih interakcija između hemijskih skupina, npr. vodonične, jonske, elektrostatičke, dipolne, itd. Nabiranjem i uvijanjem lanaca nastaju različiti oblici proteina: vlaknasti, globularni ili eliptični. [3] Ako mutacija dovede do toga da aminokiselina sa malim bočnim lancem bude zamenjena aminokiselinom sa velikim, pojaviće se problem sa formiranjem trodimenzionalne strukture. Ako bi se, pak, velika aminokiselina zamenila sa malom pojavio bi se prazan prostor, što bi moglo dovesti do destabilizacije molekula proteina.

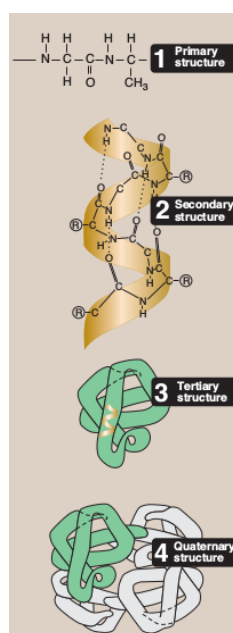
²Katalizacija predstavlja proces povećavanja brzina reakcija

³Denaturizacija proteina je proces koji izaziva promene u strukturi proteina menjajući i njihovo fiziološko dejstvo.

U molekulima proteina postoji hijerarhijska strukturalna organizacija u četiri nivoa:

1. primarna,
2. sekundarna,
3. tercijarna i
4. kvaternarna. [7]

Na slici 2 se može videti opšti prikaz mogućih struktura proteina, a na drugoj 3 šematski prikaz.

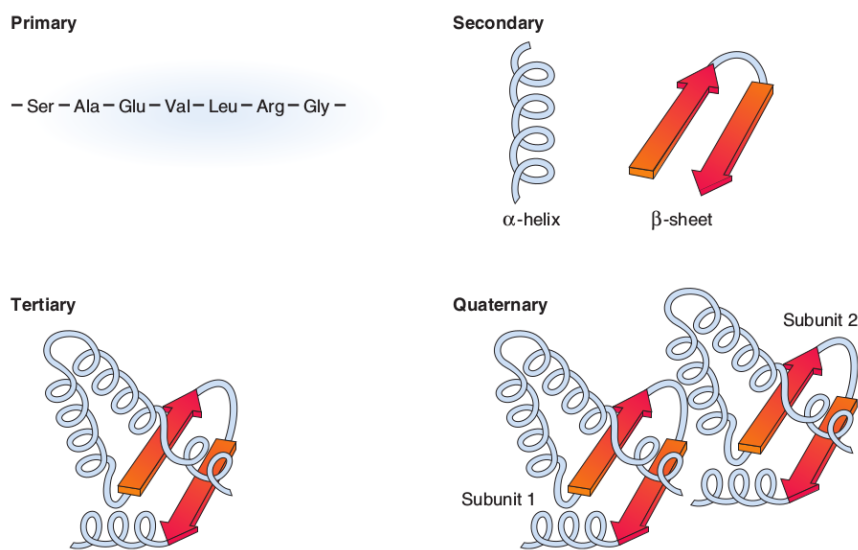


Slika 2: Prikaz struktura proteina [6]

Primarna struktura Predstavlja sâmu sekvencu aminokiselina⁴. Ona je od ključnog značaja za funkciju proteina zbog interakcija između bočnih lanaca aminokiselina koji određuju trodimenzionalnu strukturu. Proteini koji imaju sličnu sekvencu aminokiselina su *homologi*, a poređenje sekvenci među takvim proteinima može ukazati na genetsku relaciju između različitih vrsta.

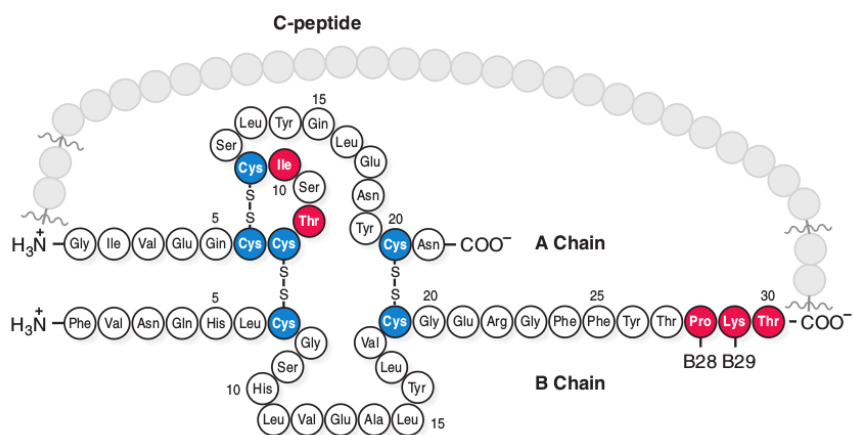
Mnoge genetske bolesti rezultuju u proteinima sa abnormalnim redosledom aminokiselina, što uzrokuje nepravilno presavijanje i gubitak ili nemogućnost normalnog funkcionisanja. Ukoliko su nam poznate strukture normalnih i mutiranih proteina te informacije možemo iskoristiti za dijagnostifikovanje ili proučavanje bolesti. [6] Primarna struktura će sa mutacijama, u najmanju ruku, izmeniti unos, brisanje ili menjanje aminokiselina. Promene u primarnoj strukturi mogu imati uticaja i na više

⁴Redosled kojim su aminokiseline poredane u nekom polipeptidu se zove sekvenca aminokiselina. [7]



Slika 3: Šematski prikaz struktura proteina [5]

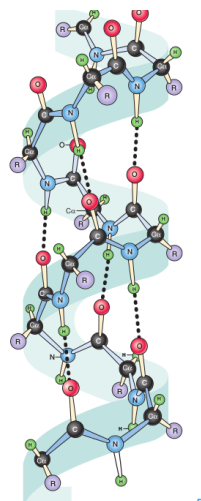
nivoe proteinskih struktura. Takve promene često dovode do lošeg presavijanja proteina i mogu dovesti do njegovog gubitka funkcije. [1] Prikaz izgleda primarne strukture na primeru insulina kod čoveka se vidi na slici 4.



Slika 4: Prikaz primarne strukture [5]

Sekundarna struktura Označava pravilno pojavljivanje ponavljane prostornog rasporeda primarne strukture u jednoj dimenziji. [3] Ovu strukturu čini nekoliko različitih oblika od kojih su najčešći α -heliks i β -presavijena ploča ili β -struktura, a čest je i tzv. β -okret.

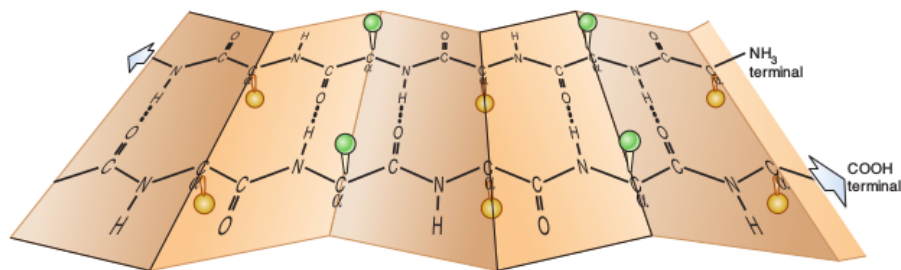
α -heliks - tip sekundarne strukture kod kog se gusto pakovani polipeptidni lanac spiralno uvrće. Karakteriše se brojem peptidnih jedinica po okretu i rastojanjem između dva okreta. Spada pod energetske najstabilnije, a time, i najstabilnije strukture proteina. Heliks mogu obrazovati i *L*- i *D*- aminokiseline. Postoje dva tipa heliksa: levostrani i desnostrani. Prikaz izgleda α -heliksa se vidi na slici 5.



Slika 5: Prikaz α -heliksa [5]

β -struktura - Za razliku od α -heliksa sastoji se od dva ili više peptidnih lanaca ili segmenata polipeptidnih lanaca, a obrazuje se kada ovakvi tipovi lanca povežu uzdužno. Postoje dva tipa β -strukture: paralelna i antiparalelna. [7]

Prikaz izgleda β -strukture se vidi na slici 6. **β -okreti** - obrću pravac



Slika 6: Prikaz β -strukture [5]

polipeptidnog lanca praveći kompaktan globularan oblik. [6]

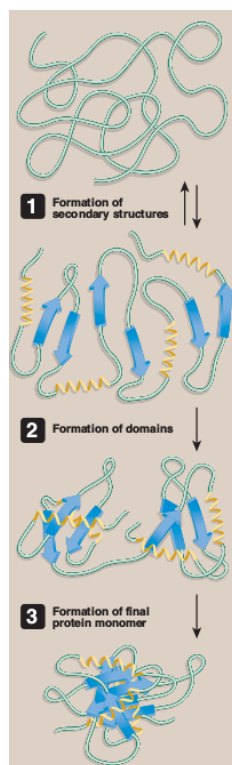
Tercijarna struktura Podrazumeva unutarmolekularno slaganje polipeptidnog lanca u kompaktnu trodimenzionalnu strukturu specifičnog oblika. [3] Takva trodimenzionalna konformacija nastaje prostornim organizovanjem polipeptidnog lanca koji već ima sekundarnu strukturu. Na taj način, ostaci aminokiselina koji su udaljeni u primarnoj strukturi do-

laze blizu. Proteini koji imaju ovakvu strukturu su globularni i kompaktni sa velikom gustinom u središtu. [7]

Kvaternarna struktura Predstavlja agregaciju više peptidnih lanaca u molekulu proteina. [3] Mnogi proteini, posebno oni velike mase, izgrađeni su od nekoliko polipeptidnih lanaca. Svaka takva komponenta naziva se *podjedinica* ili *protomer*. Oni mogu biti identični⁵ ili se razlikovati prema strukturi. Ovakav raspored dovodi do brzog i efikasnog transfera substrata od jednog aktivnog centra enzima do drugog. [7]

2.1.3 Savijanje proteina

Interakcije između lanaca aminokiselina, koji se nalaze sa strane, određuju kako se dugački polipeptidni lanac presavija u trodimenzionalni oblik funkcionalnog proteina. Presavijanje proteina koje se događa u ćeliji traje od nekoliko sekundi do nekoliko minuta. Na slici 7 se može videti opšti prikaz savijanja proteina.



Slika 7: Prikaz savijanja proteina [6]

2.1.4 Denaturacija proteina

Denaturisanje proteina rezultuje u odvijanju i dezorganizaciji proteinskih sekundarne i tercijarne strukture. Pod idealnim uslovima, denaturisanje proteina može biti *reverzibilno*. To znači da bi se protein, pri

⁵Tada takve proteine nazivamo *oligomerima*

prestanku delovanja agenasa, vratio u normalno stanje. Međutim, većina proteina ostaje trajno *neuređena*. O neuređenosti proteina biće više reči u nastavku.

Jedno od objašnjenja zašto se protein ne vraća u originalno stanje se sastoji u tome da protein počinje sa savijanjem pre nego što se izvrši sinteza celog lanca. Osim toga, specijalizovana grupa pomoćnih proteina (engl. *chaperones*) su neophodni za pravilno savijanje mnogih vrsta proteina. Ovi pomoćni proteini interaguju sa polipeptidima u nekoliko faza tokom procesa savijanja, imaju ulogu u tome da održavaju protein nesavijenim dok sinteza nije gotova ili imaju ulogu katalizatora. [6] Loše savijanje proteina može dovesti do različitih bolesti kao što su: amiloidna bolest ili Prionova bolest.

2.1.5 Neuređenost protiena

3 Prediktori

Neuređenost proteina se može utvrditi eksperimentalno, u laboratoriji, ili pomoću prediktora za automatsko utvrđivanje neuređenosti na osnovu primarne sekvence proteina. Prediktora za utvrđivanje neuređenosti danas ima toliko da dolazi do razvoja *metaprediktora*. Oni kombinuju, na različite načine, izlaze iz nekoliko prediktora.

Neki od prediktora, koji će biti korišćeni i detaljnije opisani u nastavku, su:

- SPINE-D/SPOT-Disorder
- PONDR
- s2D
- IUPred
- ESpritz
- SEG
- Disopred2

3.1 SPOT-Disorder prediktor

4 Zaključak

5 Literatura

Literatura

- [1] Bradford A. Jameson Denise R. Ferrier. *Lippincott's Illustrated Reviews Flash Cards*. Wolters Kluwer Health, 2001 Market Street, Philadelphia, PA 19103, 2015.
- [2] Gerhard Michal Dietmar Schomburg. *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2012.
- [3] Ivana Čepelak Dubravka Čvorišćec. *Štrausova medicinska biokemija*. Medicinska naklada, 2009.

- [4] Jovana J. Kovačević. Strukturna predikcija funkcije proteina i odnos funkcionalnih kategorija proteina i njihove neuređenosti, 2015. on-line na: <http://www.math.rs/files/DoktoratJK2015.pdf>.
- [5] Michael Lieberman. *Biochemistry, molecular biology, and genetics*. — 6th edition. Lippincott Williams & Wilkins, a Wolters Kluwer business, 351 West Camden Street, Baltimore, MD 21201, Two Commerce Square, 2001 Market Street, Philadelphia, PA 19103, 2014.
- [6] Denise R. Ferrier Richard A. Harvey. *Lippincott Illustrated Reviews: Biochemistry, 5th edition*. Lippincott Williams & Wilkins, a Wolters Kluwer business, 351 West Camden Street, Baltimore, MD 21201, Two Commerce Square, 2001 Market Street, Philadelphia, PA 19103, 2011.
- [7] Vesna Spasojević-Kalimanovska Slavica Spasić, Zorana Jelić-Ivanović. *Opšta biohemija*. 2002.