

Istraživanje podataka u bioinformatički

Skripta

Mirjana Maljković Ružičić

Matematički fakultet, Univerzitet u Beogradu

Školska godina: 2025/2026

Uvod

Ova skripta je namenjena za predmet Istraživanje podataka u bioinformatici. U skripti su opisane biološke osnove potrebno za bavljenje bioinformatikom. Poglavlja su nastala dopunom materijala dr Miloša Beljanskog.

Sadržaj

1	Osnove biologije	1
1.1	Ćelije i organizmi	1
1.2	Klasifikacija organizama	3
1.3	Građa ćelije	8
1.4	Atomi, interakcije između atoma, molekuli	11
1.5	Makromolekuli	15
1.6	Strukture bioloških polimera	16
1.7	Interakcija sa drugim molekulima	22
1.8	Nukleotidni kod (IUPAC)	24
1.8.1	Osnovni nukleotidi	24
1.8.2	Dvosmisleni (degenerisani) simboli	24
1.8.3	Isključujući simboli	24
1.8.4	Ostalo	24
1.9	Konvencije za pisanje DNK sekvenci	24
1.9.1	Polaritet sekvence	25
1.9.2	Bazni parovi	25
1.9.3	Oligonukleotidi	25
1.9.4	Ponavljajući oligonukleotidi	25
1.9.5	Dvostruko-lančani ponavljajući polimeri	25
1.9.6	Napomene	26
2	Uvod u genomiku	27
2.1	Osnovni pojmovi - genotip, fenotip, gen	27

2.2	Genom, sadržaj genoma	30
2.2.1	Genom	30
2.2.2	GC sadržaj genoma	34
2.2.3	Organizacija genoma kod prokariotskih organizama	35
2.2.4	Organizacija genoma kod eukariotskih organizama	35
2.2.5	Haplotip i haplogrupa	36
2.3	Mutacije	37
2.3.1	Hromozomske mutacije	37
2.3.2	Mutacije na nivou DNA	39
2.3.3	Pojedinačni nukleotidni polimorfizam – SNP	39
2.3.4	Insercije i delecije (IN/DEL)	40
2.3.5	Reparacija (popravka genetskog materijala) DNA	41
2.4	Genska ekspresija	41
2.4.1	Centralna dogma molekulske biologije	42
2.4.2	Replikacija	42
2.5	Nekodirajuće i kodirajuće RNA	45
2.6	Regulacija genske ekspresije	47
2.6.1	Transkripcija i translacija	47
2.6.2	Ekspresija gena kod eukariota i prokariota	48
2.6.3	Sazrevanje (procesovanje) eukariotske mRNA	51
2.7	Genetski kod	52
2.8	Hromozomi	52
2.8.1	Organizacija prokariotskih hromozoma	52
2.8.2	Organizacija eukariotskog hromozoma	54
2.8.3	Ploidija	61
2.9	Virusi	62
2.9.1	Korona virusi i SARS-2 CoV	64
2.10	Određivanje primarne strukture (sekvencionisanje) DNA	66
3	Kratki uvod u hemiju i biologiju belančevina	69
3.1	Amino kiseline	69
3.1.1	Amino kiseline koje ulaze u sastav proteina	69
3.1.2	Fizličko hemijske osobine amino kiselina	70
3.1.3	Polimerizacija amino kiselina - peptidna veza	80
3.1.4	Diedralni uglovi	81
3.2	Nivoi organizacije polipeptidnih lanaca	82
3.2.1	Primarna struktura proteina	82
3.2.2	Sekundarna struktura proteina	87
3.2.3	Aminokiseline i sekundarne strukture	91
3.2.4	Lokalna proteinska struktura – strukturalni alfabet	93
3.2.5	Supersekundarna struktura, motiv, domen	93
3.2.6	Tercijalna i kvaternarna struktura proteina	97
3.2.7	Protein Data Bank (PDB)	99

3.2.8	Metamorfni i neuređeni proteini	99
3.2.9	Globularni proteini	103
3.3	Interakcije između proteina	103
3.3.1	Međumolekulske interakcije	103
3.3.2	Funkcija proteina	104
3.4	Klasifikacija proteina	107
3.4.1	Strukturna klasifikacija proteina	108
3.5	Eksperimentalne metode za određivanje strukture proteina	109
3.5.1	Metoda zasnovana na difrakciji X-zraka	110
3.5.2	2D Nuklearno-magnetna rezonantna spektroskopija	111

Osnove biologije

1.1 Čelije i organizmi

Život se, najkraće i najopštije, može definisati kao visoko (samo)organizovani kompleksan sistem, koji se sastoji od molekula i atoma (jona).

Američka svemirska agencija (NASA) je definisala život kao: *samoodrživi hemijski sistem sposoban za Darvinijsku evoluciju*. Savremena minimalna definicija života je [8]: *Život je samoodrživi hemijski sistem van ravnoteže, sposoban za obradu, transformaciju i akumulaciju informacije prikupljene iz svoje sredine*. Na slici 1.1 se mogu videti osnovne karakteristike živih bića.

Čelija je osnovna strukturalna i funkcionalna (biološka) jedinica svih do sada poznatih organizama.

Organizam ili jedinka, je složeni sistem organa i može biti jednoćelijski ili višećelijski. Jednoćelijski organizmi se sastoje od subcelularnih (strukture unutar ćelije) komponenti, tj. *organela*. Ta jedna ćelija mora da obavlja sve osnovne životne funkcije, kao što su ishrana, rast, razmnožavanje ili kretanje. Primeri jednoćelijskih organizama su bakterije, amebe, neke alge, kvasci.

Višećelijski organizmi mogu biti kolonijalni ili sa diferentovanim ćelijama. Kod kolonijalnih višećelijskih organizam ćelije su grupisane u koloniju, ali su slične i relativno nezavisne, tj. svaka ćelija može da funkcioniše skoro samostalno. Primer kolonijalnog organizma je kolonijalna alga, Volvox. Diferentovani višećelijski organizmi su organizmi kod kojih su ćelije specijalizovane za obavljanje različitih funkcija u organizmu. Na primer, nervne ćelije prenose nervne impulse, mišićne ćelije omogućavaju kretanje, a krvne ćelije prenose kiseonik. Kod složenijih višećelijskih organizama diferentovane ćelije formiraju tkiva, tkiva grade organe, a organi zajedno čine organske sisteme. Primeri diferentovanih višećelijskih organizama su: čovek, životinje ili biljke.

Prema svojoj organizaciji ćelije se manifestuju u dva osnovna tipa:

- prokariotske ćelije čiji je genetički (informacioni) materijal organizovan kao **nukleoid** i nije odvojen membranom od ostatka ćelije i
- eukariotske čiji je genetički materijal organizovan kao **nukleus** odvojen dvostrukom membranom od ostatka ćelije.

Eukariotska ćelija se sastoji od:



Slika 1.1: Osnovne karakteristike živih bića

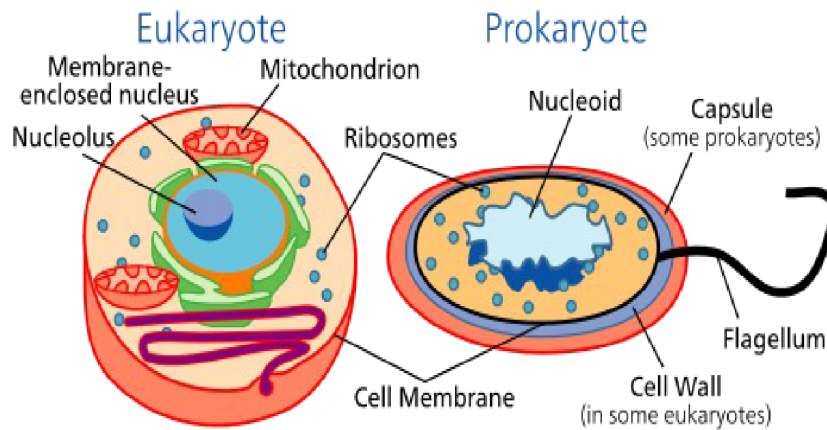
- jedra - organela koja sadrži genetski materijal ćelije u obliku DNK. Jedro kontroliše aktivnosti ćelije, uključujući sintezu proteina, rast i deobu ćelije.
- jedarca (nucleolus) - struktura unutar jedra u kojoj se sintetiše ribozomska RNK (rRNK) i formiraju podjedinice ribozoma.
- mitohondrija - organele u kojima se odvija ćelijsko disanje i proizvodnja energije.
- ribozomi - strukture u kojima se vrši sinteza proteina. Oni prevode genetsku informaciju sa RNK u proteinske molekule.
- ćelijska membrana je polupropustljiva struktura koja okružuje ćeliju i reguliše ulazak i izlazak supstanci između ćelije i spoljašnje sredine.

Prokariotska ćelija sadrži:

- nukleoid, region u ćeliji gde se nalazi DNK prokariota.
- ribosome koji u prokariotskim ćelijama imaju istu osnovnu funkciju kao i kod eukariota — sintezu proteina.
- ćelijsku membranu koja reguliše transport supstanci u ćeliju i iz nje i učestvuje u nekim metaboličkim procesima.
- ćelijski zid koji daje ćeliji oblik i pruža mehaničku zaštitu.

- kapsulu, zaštitni sloj koji se nalazi oko ćelijskog zida kod nekih bakterija. Ona štiti ćeliju od nepovoljnih uslova i može povećati sposobnost bakterije da se veže za površine.
- flagelum, tj. bič je struktura koja omogućava kretanje ćelije u tečnoj sredini.

Na slici 1.2 se mogu videti delovi eukariotske i prokariotske ćelije, a u tabeli 1.1 poredenje njihovih karakteristika.



Slika 1.2: Eukariotska vs prokariotska ćelija

1.2 Klasifikacija organizama

Biološka sistematika organizama se bavi proučavanjem, razgraničavanjem, upoređivanjem, hijerarhijom i filogenijom organizama, ili preciznije, obuhvata

- dodeljivanje imena,
- njihovo opisivanje,
- njihovo uzorkovanje i čuvanje,
- klasifikaciju i određivanje kriterijuma za nju,
- izučavanje njihove evolucije i
- mehanizama adaptacije.

Često se uz termin biološka sistematika koristi i termin biološka taksonomija. Međutim, taksonomija je uži pojam i bavi se tačkama od (a) do (d), tj. taksonomija se pretežno bavi definisanjem taksona i dodeljivanjem taksonomskih rangova (kategorija) određenim taksonima. Najviši taksonomski rang organizma u taksonomiji je **superkraljevstvo** ili **carstvo** (lat. superregnum, eng. superkingdom, takođe i eng. domain – domen). Niže taksonomske kategorije su: **tip** ili **razdeo** (lat. phylum, eng. division), **klasa** (lat. classis), **red** (lat. ordo), **porodica** ili **familija** (lat. familia), **rod** (lat. genus) i **vrsta** (lat. species). Na slici 1.3 se može videti grafički prikaz hijerarhije taksonomskih rangova organizama sa primerom taksonomske klasifikacije čoveka.

Vrsta se definiše kao grupa populacija (populacija – skup jedinki koje se mogu međusobno ukrštati, tj. daju fertile (plodne) potomke i žive na istom geografskom prostoru, tj. nisu rasplodno izolovane), koje mogu biti rasplodno izolovane, ali se mogu međusobno ukrštati, tj. daju fertile potomke.

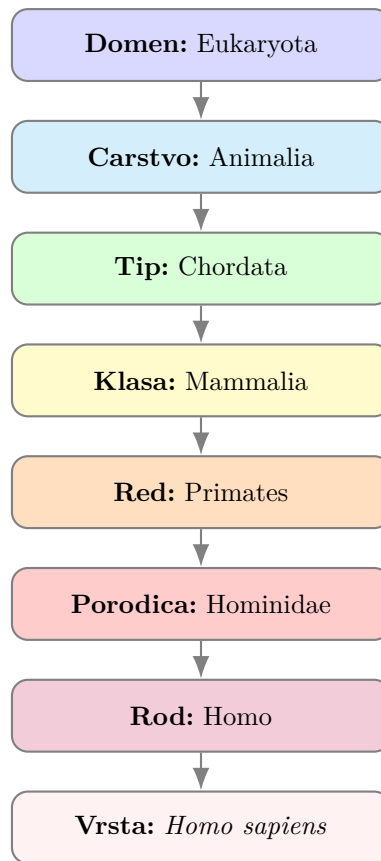
Karakteristika	Prokarioti	Eukarioti
Tipični organizmi	bakterije, arheje	protisti, gljive, biljke, životinje
Tipična veličina	~1–5 μm	~10–100 μm
Tip jedra	nukleoidni region; nema pravog jedra	pravo jedro sa dvostrukom membranom
DNK	obično kružna	linearni molekuli (hromozomi) sa histonskim proteinima
Sinteza RNK i proteina	odvija se u citoplazmi	RNK se sintetise u jedru, proteini u citoplazmi
Ribozomi	50S i 30S	60S i 40S
Citoplazmatska struktura	veoma malo unutrašnjih struktura	visoko organizovana endomembranama i citoskeletom
Kretanje ćelije	flagela građena od flagelina	flagela i cilije sa mikrotubulama; lamelipodije i filopodije sa aktinom
Mitohondrije	nema	od jedne do nekoliko hiljada (neke ćelije ih nemaju)
Hloroplasti	nema	prisutni kod algi i biljaka
Organizacija	obično jednoćelijski organizmi	jednoćelijski, kolonijalni ili višćelijski organizmi sa specijalizovanim ćelijama
Deoba ćelije	binarna fisija (jednostavna deoba)	mitoza (fisija ili pupanje) i mejoza

Tabela 1.1: Poređenje karakteristika prokariotskih i eukariotskih ćelija

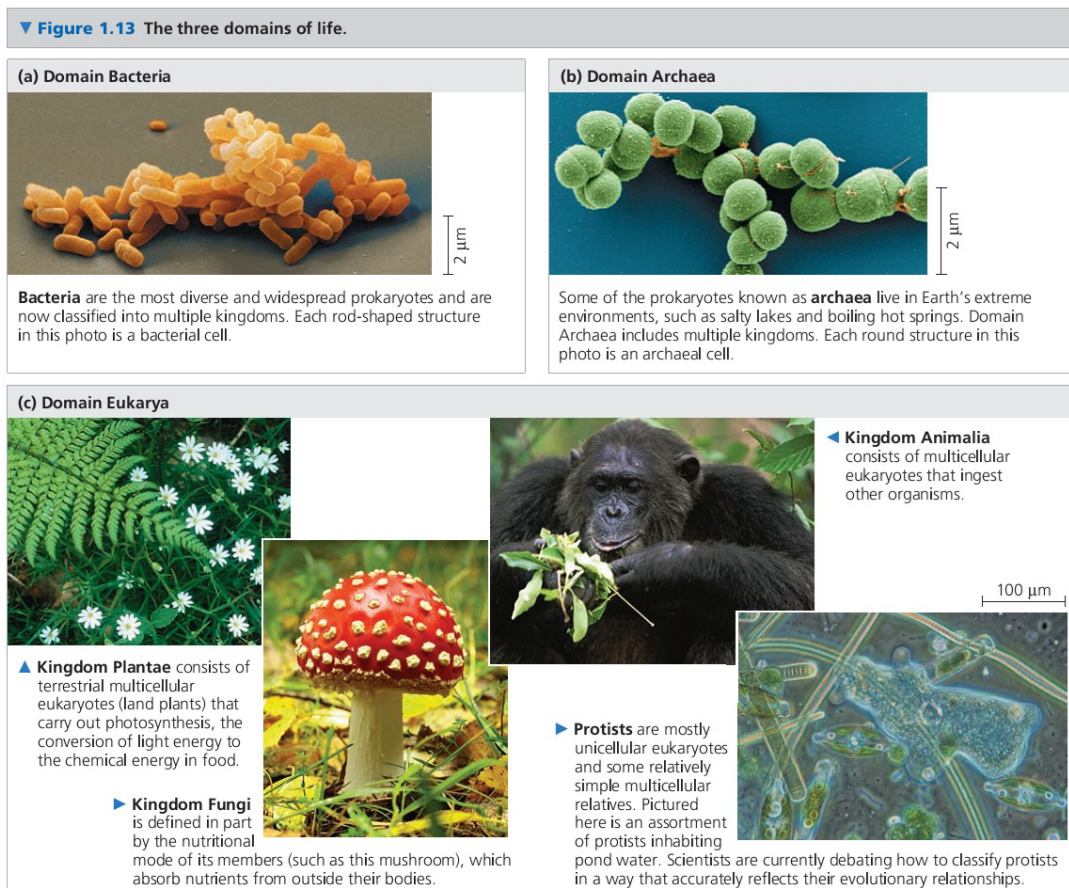
U upotrebi su najmanje tri alternativne klasifikacije organizama, na najvišim taksonomskim rangovima:

1. podela na dva carstva (superdomain): prokariote (Prokaryota ili Monera) i eukariote (Eukaryota), (u ovom slučaju termin pro- i eu-kariote označava pored tipa ćelijske organizacije i taksonomsku kategoriju).
2. podela na 6 kraljevstva: arhebakterije (Archaea, gr.ar), eubakterije (Eubacteria), protiste (lat. Protista), gljive (lat. Fungus), biljke (lat. Planta) i životinje (lat. Animalia).
3. podela na tri carstva (domain, superregnum) (slika 1.4: Arhee (Archea), Bakterije (Bacteria) i Eukarie (Eukarya).

Podela živih organizama na tri carstva izvršena je na osnovu poređenja sekvenci ribozomske RNK. Molekularna analiza pokazala je da se organizmi mogu grupisati u tri osnovne evolutivne linije koje se značajno razlikuju po genetskim i biohemijskim osobinama.



Slika 1.3: Hijerarhijski prikaz taksonomskih rangova organizama sa primerom taksonomske klasifikacije čoveka



Slika 1.4: Podela na tri carstva: Arhee, Bakterije i Eukarie

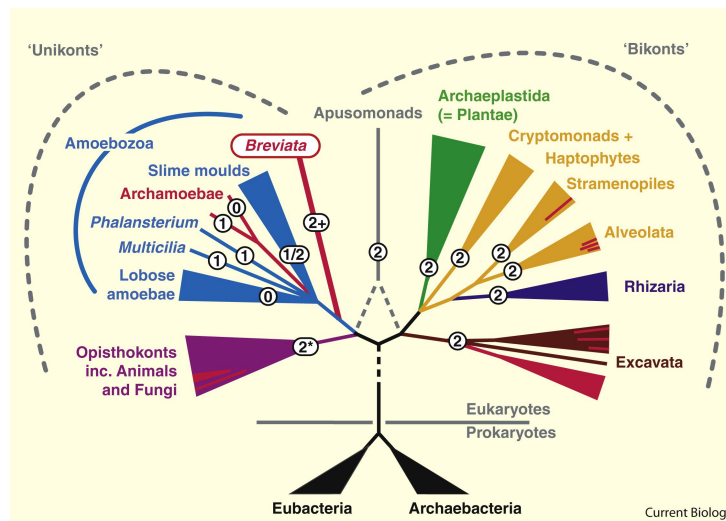
U okviru Eukaria su kraljevstva: protista, gljiva, biljka i životinja.

Eukaria se može podeliti u dva poddomena prema broju flagela (bičeva) kod ćelija predaka i nekim genetskim karakteristikama:

1. Unikonte (Unikonta – jedna flagela ili bič, (lat. uni - jedan + flagellum-a – bič)): razne Protiste, Gljive i Životinje i
2. Bikonte (Bikonta (lat. bi – dva) dve flagele: razne Protiste i Biljke.

Ova podela ponajbolje odslikava evolucione odnose, ali nije opšteprihvaćena.

Slika 1.5 prikazuje filogenetsko stablo eukariota i podelu eukariotskih organizama na dve velike evolutivne grupe: unikonte (Unikonts) i bikonte (Bikonts).



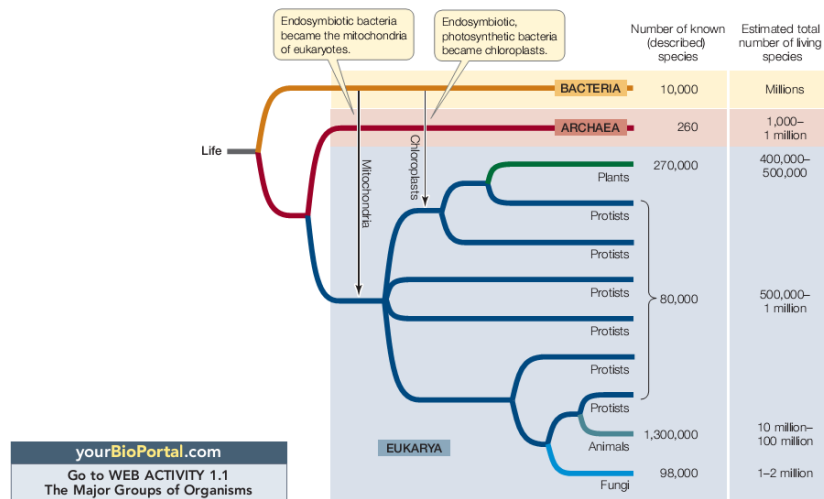
Slika 1.5: Filogenetsko stablo eukariota i podela eukariotskih organizama na dve unikonte i bikonte

Binominalna nomenklatura je formalni sistem kojim se imenuju i kalsifikuju organizmi. Svako ime se sastoji od dva dela, koji upotrebljavaju latinsku gramatičku formu, koji imenuju rod i vrstu, npr. *Homo sapiens*, *Escherichia coli*, itd.

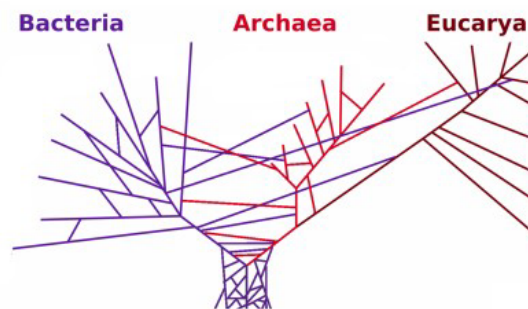
Drvo života (slika 1.6) predstavlja grafički prikaz evolutivnih odnosa između svih živih organizama na Zemlji. U ovom prikazu organizmi su raspoređeni u obliku granajućeg stabla koje potiče od zajedničkog pretka svih živih bića. Svaka grana na stablu predstavlja određenu evolutivnu liniju organizama koja se tokom vremena razdvajala i razvijala u različite vrste.

Uobičajeno se evolucija organizama prikazuje kao drvo života, gde se geni prenose vertikalno, odnosno sa roditelja na potomke tokom generacija. Međutim, genetski materijal se može prenositi i horizontalno, između različitih organizama koji nisu u direktnom odnosu predak–potomak. Slika 1.7 prikazuje horizontalni genski transfer, odnosno prenos genetskog materijala između različitih organizama koji nisu u odnosu predak–potomak. Ovaj proces je naročito čest kod bakterija i pokazuje da evolucija života nije uvek strogo hijerarhijska kao drvo, već može imati i mrežastu strukturu zbog razmene gena između različitih linija organizama.

Habitat predstavlja prirodno okruženje u kome određena vrsta organizma živi i u kome pronalazi uslove potrebne za opstanak, kao što su voda, hranljive materije, odgovarajuća temperatura i hemijski uslovi. Na slici 1.8 su prikazani veoma različiti habitati na Zemlji. Slika ilustruje da se život može pronaći u veoma različitim i često ekstremnim uslovima, što pokazuje veliku



Slika 1.6: Drvo života

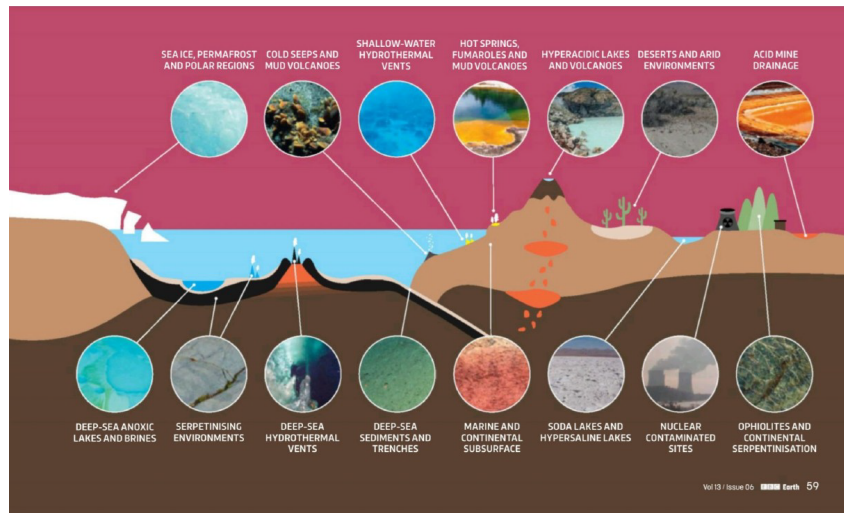


Slika 1.7: Horizontalno genski transfer

prilagodljivost organizama na različite tipove staništa. Ekstremofilni mikroorganizmi su organizmi (najčešće bakterije i arheje) koji mogu da žive i razmnožavaju se u ekstremnim uslovima okoline u kojima većina drugih organizama ne može opstati. Takvi uslovi mogu uključivati veoma visoke ili niske temperature, ekstremne pH vrednosti, visok salinitet, veliki pritisak ili intenzivno zračenje.

U prirodi organizmi ne žive izolovano, već su međusobno povezani različitim oblicima interakcija. Ovi odnosi mogu značajno uticati na opstanak, razmnožavanje i evoluciju vrsta. Interakcije između organizama uključuju i iskorišćavanje jednog organizma od strane drugog. Poseban oblik takvih odnosa predstavlja simbioza. U zavisnosti od toga da li jedna ili obe vrste imaju korist ili štetu od tog odnosa, razlikuju se različiti oblici simbiotskih odnosa, među kojima su parazitizam, mutualizam i komensalizam.

Parazitizam (grč. para = „pored“ i situs = „hrana“) je vrsta odnosa između vrsta u kome jedna vrsta (parazit) živi na račun druge (domaćina) (pr. krpelj, Plasmodium izazivač malarije). Mutualizam (izvedeno od: mutual= obostran, međuzavisan) je tip ekoloških interakcija između



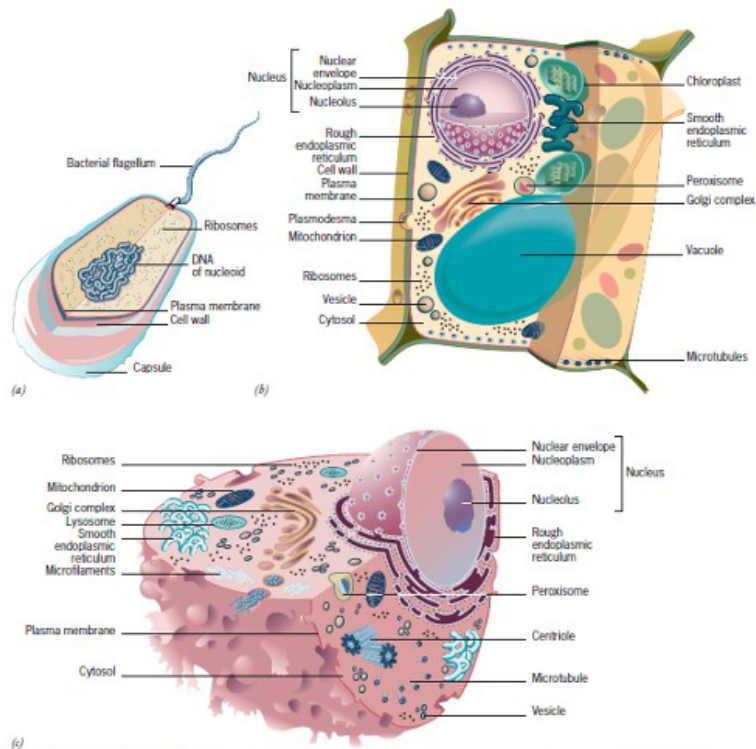
Slika 1.8: Habitati organizama

jedinki različitih vrsta kojim obe jedinke povećavaju svoju adaptivnu vrednost. Pritom, jedinke nisu tokom celog života upućene jedna na drugu (npr. pčele i biljke cvetnice, biljke i gljive mikorize). Komensalizam (izvedeno od lat. cum=zajedno, i lat. mensa=sto, trpeza) je oblik ekoloških interakcija između vrsta koja na jednu vrstu (zvanu komensal) utiče pozitivno, dok za drugu vrstu ne postoje ni pozitivne ni negativne posledice interakcije (npr. spore, plodovi ili semena se zakače za životinju-prenosioca i tako budu rasejana).

1.3 Građa ćelije

Unutrašnjost ćelije čini protoplazma, koja je od spoljašnje sredine odvojena ćelijskom membranom. Membrana omogućava kontrolisanu razmenu supstanci između ćelije i okoline i time održava stabilne uslove neophodne za odvijanje životnih procesa. Unutar ćelije nalaze se različite ćelijske komponente, koje predstavljaju visokoorganizovane makromolekulske komplekse sastavljene od proteina, lipida, nukleinskih kiselina i drugih molekula. Ove strukture obavljaju specifične i specijalizovane funkcije i omogućavaju normalno funkcionisanje ćelije. U ćeliji se mogu razlikovati različite vrste struktura: biološke membrane, globularne i fibrilarne komponente. Biološke membrane formiraju granice između pojedinih delova ćelije i učestvuju u transportu molekula i ćelijskoj komunikaciji. Globularne strukture su kompaktni makromolekulski kompleksi, dok fibrilarne strukture formiraju vlaknaste elemente koji učestvuju u održavanju oblika ćelije i u različitim procesima kretanja i transporta unutar ćelije. Među membranskim subcelularnim komponentama nalaze se ćelijska membrana, jedro, mitohondrije i hloroplasti. Globularne komponente uključuju ribozomalne subjedinice i proteazome, dok fibrilarne strukture čine elementi citoskeleta, kao što su mikrofilamenti izgrađeni od aktina, intermedijarni filamenti i mikrotubule izgrađene od tubulina. Ćelija predstavlja složen i visoko organizovan sistem u kome različite strukture međusobno saraduju i omogućavaju održavanje životnih funkcija.

Na slici 1.9 je dat šematski prikazi generalizovane (a) bakterijske, (b) biljne i (c) životinjske ćelije. U tabeli 1.2 je dat pregled nekih delova prokariotske ćelije, a u tabeli 1.3 eukariotske ćelije.



Slika 1.9: Generalizovana građa ćelije

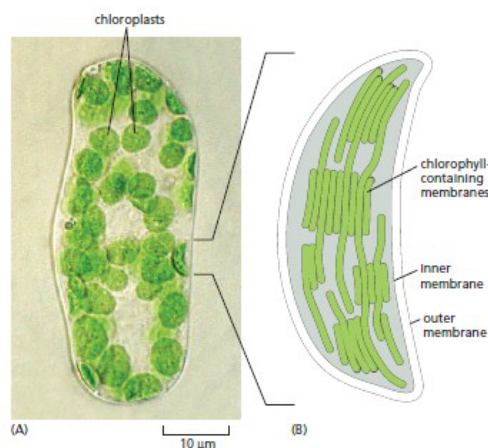
Organel	Glavna funkcija	Struktura	Organizmi	Napomene
Hloroplast (plastid)	fotosinteza, hvatanje energije sunčeve svetlosti	kompartment sa dvostrukom membranom	biljke, protisti	sadrži sopstvene gene; pretpostavlja se da je nastao endosimbiozom
Endoplazmatični retikulum	sinteza i savijanje novih proteina (grubi ER), sinteza lipida (glatki ER)	kompartment sa jednom membranom	svi eukarioti	grubi ER je prekriven ribozomima; glatki ER ima tubularnu strukturu
Goldijev aparat	sortiranje, pakovanje i modifikacija proteina	kompartment sa jednom membranom	svi eukarioti	cis-strana je okrenuta ka ER; trans-strana je udaljena od ER
Mitohondrije	proizvodnja energije oksidacijom organskih molekula i sintezom ATP	kompartment sa dvostrukom membranom	većina eukariota	sadrže sopstvenu DNK; verovatno nastale endosimbiozom
Vakuola	skladištenje, transport i održavanje homeostaze	kompartment sa jednom membranom	eukarioti	posebno velika u biljnim ćelijama
Jedro	održavanje DNK, kontrola aktivnosti ćelije, transkripcija RNK	kompartment sa dvostrukom membranom	svi eukarioti	sadrži najveći deo genoma ćelije

Tabela 1.3: Glavne eukariotske organele

Organela / makromolekul	Glavna funkcija	Struktura	Organizmi
Karboksizom	fiksacija ugljenika	proteinski omotač (kompartiment)	neke bakterije
Hlorozom	fotosinteza	kompleks za sakupljanje svetlosti	zelene sumporne bakterije
Flagelum	kretanje u spoljašnjoj sredini	proteinski filament	neki prokarioti i eukarioti
Magnetozenom	magnetna orijentacija	neorganski kristal u lipidnoj membrani	magnetotaktične bakterije
Nukleoid	održavanje DNK, transkripcija u RNK	DNK-protein kompleks	prokarioti
Plazmid	razmena genetskog materijala (DNK)	kružna DNK	neke bakterije
Ribozom (70S)	translacija RNK u proteine	RNK-protein kompleks	bakterije i arheje
Tilakoid	fotosinteza	fotosistemski proteini i pigmenti	uglavnom cijanobakterije
Mezosomi	funkcije slične Goldijevom aparatu i centriolama	mala nepravilno oblikovana organela koja sadrži ribosome	prisutni u većini prokariotskih ćelija

Tabela 1.2: Prokariotske organele i ćelijske komponente

Hloroplasti hvataju energiju sunčeve svetlosti u biljnim ćelijama i nekim jednoćelijskim eukariotima. Na slici 1.10 je prikazana (A) jedna ćelija izolovana iz lista cvetnice, posmatrana svetlosnim mikroskopom, na kojoj se vide zeleni hloroplasti. (B) šematski prikaz jednog hloroplasta koji pokazuje složeni sistem unutrašnjih membrana koje sadrže molekule hlorofila pomoću kojih se apsorbira svetlost.



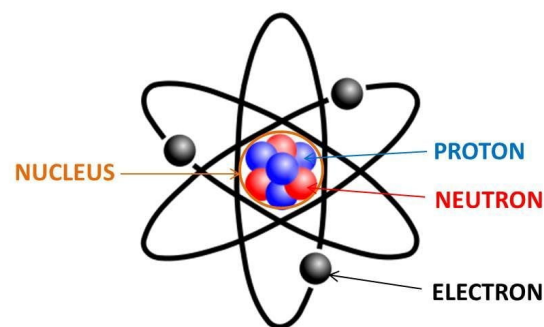
Slika 1.10: Hloroplasti

1.4 Atomi, interakcije između atoma, molekuli

Atomi su osnovne jedinice materije i predstavljaju gradivne blokove svih hemijskih elemenata.

Osnovne čestice atoma su protoni, neutroni i elektroni (slika 1.11), koje zajedno čine strukturu svakog atoma. Protoni imaju pozitivan električni naboj i nalaze se u jezgru atoma, zajedno sa neutronima koji nemaju električni naboj (neutralni su). Jezgro sadrži gotovo svu masu atoma. Elektroni imaju negativan naboj i kreću se oko jezgra u određenim energetske nivoima ili elektronskim ljuskama. Broj protona određuje hemijski element, dok raspored elektrona utiče na hemijska svojstva i način na koji atom reaguje sa drugim atomima.

Valentni elektroni su elektroni koji se nalaze u najudaljenijoj (spoljašnjoj) elektronskoj ljusci atoma. Oni su najvažniji jer učestvuju u hemijskim reakcijama i određuju kako će se atom vezivati sa drugim atomima.



Slika 1.11: Čestice atoma

U periodnom sistemu elemenata (slika 1.12) su organizovani prema rastućem atomskom broju, odnosno broju protona u jezgru. Raspored elemenata u redovima, **periodama**, i kolonama, **grupama**, odražava njihove slične hemijske i fizičke osobine. Atomi elemenata iz iste grupe imaju sličan broj valentnih elektrona, zbog čega pokazuju sličnu reaktivnost, dok se kroz periode postepeno menjaju osobine kao što su atomski radijus, elektronegativnost i energija jonizacije. Periodni sistem tako omogućava predviđanje ponašanja elemenata i njihovih međusobnih interakcija, što je od ključnog značaja za razumevanje hemijskih reakcija i strukture materije. Donji broj uz ime atoma se naziva *redni ili atomski broj* i predstavlja broj protona, a gornji broj je *maseni broj* i jednak je zbiru broja protona i neutrona u jezgru. Atomi jednog elementa se mogu među sobom razlikovati po masi iako svi imaju isti broj protona. Razlika u masi potiče od različitog broja neutrona u jezgru, a takvi atomi se nazivaju *izotopi*.

Elektronski omotač se sastoji od tačno određenog broja elektrona, koji se velikom brzinom kreću u prostoru oko atomskog jezgra. Atomi su električno neutralne čestice, a to znači da su izgrađeni od istog broja elektrona i protona.

Atomska orbitala (slika 1.13) predstavlja deo prostora u atomu gde je najveća verovatnoća nalaženja elektrona (elektronska gustina). Ne može se tačno reći gde je elektron, ali se može reći gde se najčešće nalazi. Veličina i oblik ovog prostora zavise od toga o kojoj se orbiti ($n=1, 2, 3, \dots$) i orbitali (podorbite, $l=0$ -s, 1-p, 2-d, \dots) radi. Svaka s-podorbita sadrži jednu s-orbitalu, svaka p-podorbita sadrži tri p-orbitale, d-podorbita sadrži pet d-orbitala, a svaka f-podorbita sadrži sedam f-orbitala. U svakoj orbitali staju po dva elektrona, tako da se na osnovu atomskog

Hauptgruppe	I	II	III	IV	V	VI	VII	VIII	
Periode	1	¹ ₁ H Wasserstoff						⁴ ₂ He Helium	
	2	⁷ ₃ Li Lithium	⁹ ₄ Be Beryllium	¹¹ ₅ B Bor	¹² ₆ C Kohlenstoff	¹⁴ ₇ N Stickstoff	¹⁶ ₈ O Sauerstoff	¹⁹ ₉ F Fluor	²⁰ ₁₀ Ne Neon
	3	²³ ₁₁ Na Natrium	²⁴ ₁₂ Mg Magnesium	²⁷ ₁₃ Al Aluminium	²⁸ ₁₄ Si Silicium	³¹ ₁₅ P Phosphor	³² ₁₆ S Schwefel	³⁵ ₁₇ Cl Chlor	⁴⁰ ₁₈ Ar Argon
	4	³⁹ ₁₉ K Kalium	⁴⁰ ₂₀ Ca Calcium	⁶⁹ ₃₁ Ga Gallium	⁷⁴ ₃₂ Ge Germanium	⁷⁵ ₃₃ As Arsen	⁸⁰ ₃₄ Se Selen	⁷⁹ ₃₅ Br Brom	⁸⁴ ₃₆ Kr Krypton
	5	⁸⁵ ₃₇ Rb Rubidium	⁸⁸ ₃₈ Sr Strontium	¹¹⁵ ₄₉ In Indium	¹²⁰ ₅₀ Sn Zinn	¹²¹ ₅₁ Sb Antimon	¹³⁰ ₅₂ Te Tellur	¹²⁷ ₅₃ I Iod	¹³² ₅₄ Xe Xenon
	6	¹³³ ₅₅ Cs Cäsium	¹³⁸ ₅₆ Ba Barium	²⁰⁵ ₈₁ Tl Thallium	²⁰⁸ ₈₂ Pb Blei	²⁰⁹ ₈₃ Bi Wismut	^{(209)*} ₈₄ Po Polonium	^{(210)*} ₈₅ At Astat	^{(222)*} ₈₆ Rn Radon
	7	^{(223)*} ₈₇ Fr Francium	^{(226)*} ₈₈ Ra Radium						

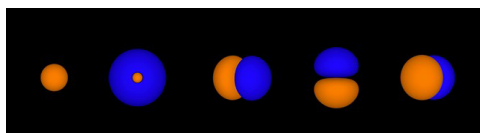
Metalle
 Halbmetalle
 Nichtmetalle

Slika 1.12: Deo periodnog sistem elemenata

broja može odrediti raspored elektrona po orbitalama.

Vrste orbitala po obliku:

- s orbitala ($l=0$) ima oblik sfere i prima 2 elektrona
- p orbitala ($l=1$) ima oblik tege i ima 3 orbitale i prima 6 elektrona
- d orbitala ($l=2$) ima 5 orbitala i prima 10 elektrona
- f orbitala ($l=3$) ima 7 orbita i prima 14 elektrona



Slika 1.13: Atomske orbitale

Hibridizacija je koncept mešanja atomskih orbitala pri čemu se formiraju nove hibridne orbitale. Hibridne orbitale su korisne za objašnjenje i opisivanje geometrijskog oblika molekula. Pri mešanju s i p orbitala, dolazi do pomeranja centra naelektrisanja elektronskog oblaka u odnosu na s orbitalu. Hibridizacija prema broju modifikovanih atomskih orbitala u ugljenikovom atomu može biti:

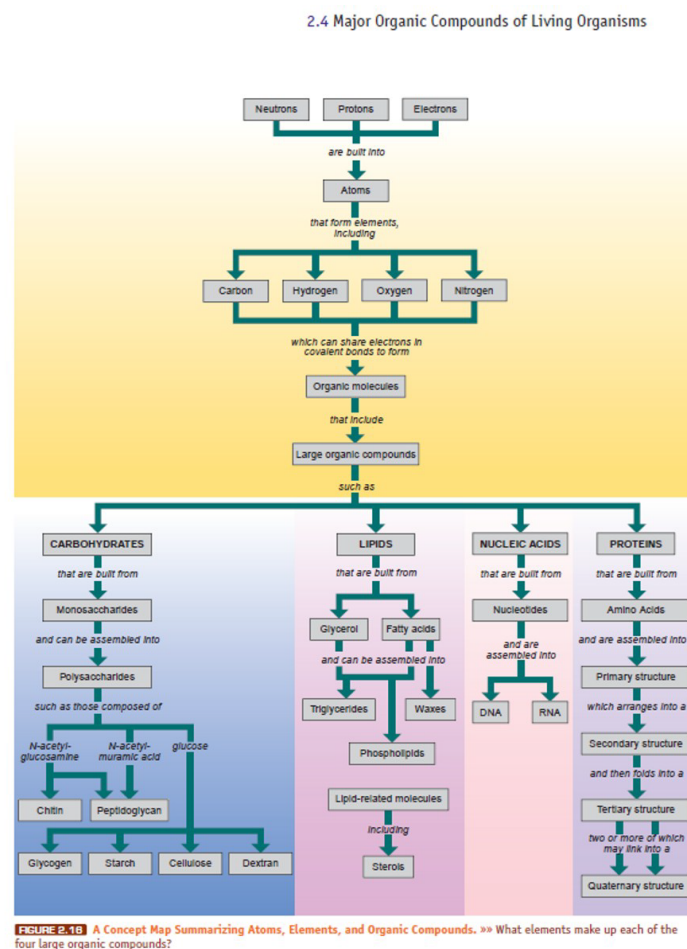
- tetraedarska hibridizacija - sp^3 (zasićena organska jedinjenja)
- trigonalna - sp^2 (nezasićena organska jedinjenja sa dvogubom vezom)
- digonalna - sp (nezasićena organska jedinjenja sa trogubom vezom)

U sp^2 hibridizaciji (slika 1.14) mešaju se $1s$ i $2p$ i nastaju $3sp^2$ orbitale, dok $1p$ ostaje. Zauzimaju oblik trougla u jednoj ravni sa uglovima od 120° .

Tabela 1.15) prikazuje za atome C, N i O koliko veza svaki od njih može da napravi u zavisnosti od hibridizacije. Za svaki atom je dato

- Hybridized orbitals - tip hibridizacije (sp^3 , sp^2 , sp)
- Number of bonds - maksimalan broj veza

hemijske elemente. Najvažniji za žive organizme su ugljenik, vodonik, kiseonik i azot. Atomi međusobno dele elektrone i formiraju kovalentne veze, čime nastaju organski molekuli. Kovalentna veza je hemijska veza koja nastaje kada dva atoma dele jedan ili više parova elektrona. Svaki atom doprinosi jednim elektronom. Postoje polarne i nepolarne kovalentne veze. Kod nepolarne kovalentne veze elektroni su ravnomerno podeljeni (npr. H_2), a kod polarna jedan atom više „vuče“ elektrone (npr. H_2O). Od organskih molekula formiraju se velike organske supstance, koje se dele u četiri glavne grupe: ugljeni hidrati, lipidi, nukleinske kiseline i proteini. Ugljeni hidrati su izgrađeni od monosaharida koji se povezuju u polisaharide kao što su glikogen, škrob i celuloza. Lipidi nastaju od glicerola i masnih kiselina i formiraju trigliceride, voskove, fosfolipide i sterole. Nukleinske kiseline su građene od nukleotida i uključuju DNK i RNK. Proteini su sastavljeni od aminokiselina koje se organizuju u primarnu, sekundarnu, tercijarnu i kvaternarnu strukturu. Slika jasno ilustruje hijerarhiju od osnovnih čestica do kompleksnih biomolekula i njihovu međusobnu povezanost u živim sistemima.



Slika 1.17: Glavne organske komponente živih organizama

1.5 Makromolekuli

Makromolekuli se sastoje od velikog broja atoma (više od 1000) i imaju velike molekulske mase ($M_r > 500$ do nekoliko miliona). Dele se na neorganske i organske. Neorganski makromolekuli

su najčešće kristalne prirode (atomi su raspoređeni u pravilnu, uređenu mrežu) i uglavnom nisu zasnovani na organskoj ugljeničnoj strukturi karakterističnoj za živi svet. Primeri neorganskih makromolekula su dijamant, grafit, pesak (kristal alumosilikata), itd).

Organski makromolekuli su najčešće polimerne prirode. Polimeri (*polus* – mnogo, više + *meros* – deo) nastaju kovalentnim povezivanjem manjih strukturnih (pod)jedinica ili mera, odnosno monomera koje se ponavljaju. Na primer polietilen (materijal od kog se često pravi plastika) se sastoji od mnogo molekula etilena, a polinukleotid od nukleotid.

Oligomeri (*oligos* – nekoliko), sastoje se od manjeg broja monomernih jedinica (broj nije precizan, obično 10–100). Monomerne jedinice mogu biti iste hemijske prirode (homopolimeri, npr.: etilen) ili se mogu manje ili više razlikovati (hetero- ili kopolimeri, npr.: amino kiseline), ali je tip veza između njih isti (npr.: peptidna veza između različitih amino kiselina – polipeptidi). Mogu biti:

- **veštački** – sintetisani (npr.: razne vrste plastike: polivinilhlorid (PVC), veštačka guma)
- prirodni – biološki – **biopolimeri**:
 - polisaharidi (polimeri šećera, npr.: celuloza, hitin (nalazi se u oklopu insekata i zidovima gljiva)),
 - nukleinske kiseline (polimeri nukleotida, npr.: DNK, RNK),
 - lipidi (polimeri metilenskih „mostova“, tj. $-CH_2-$ grupa, npr.: masti, voskovi, steroli, neki vitamini),
 - belančevine (ili proteini, polimeri amino kiselina).

Po svojoj strukturi makromolekuli mogu biti:

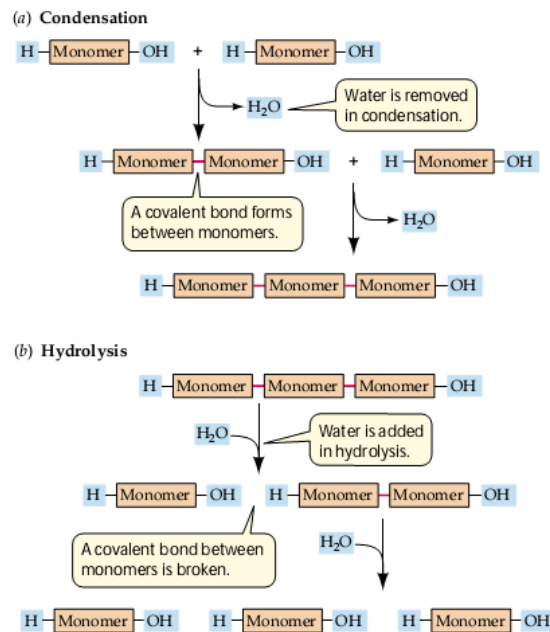
- linearni (npr.: nukleinske kiseline, proteini), ili
- razgranati, tj. pored glavnog lanca mogu imati i bočne. Najvažniji biološki polimeri su: polisaharidi, polinukleotidi, polipeptidi i lipidi.

Procesi ključni za izgradnju i razgradnju bioloških makromolekula u živim organizmima su kondenzacija i hidroliza (slika 1.18). **Kondenzacija** je reakcija u kojoj se monomeri povezuju u polimere uz izdvajanje molekula vode. Sa jednog monomera uklanja se vodonik (H), a sa drugog hidroksilna grupa (OH), koji zajedno formiraju H_2O , dok se između monomera uspostavlja kovalentna veza i nastaje duži lanac. Suprotan proces je **hidroliza**, pri kojoj se dodavanjem vode razgrađuju polimeri na pojedinačne monomere. Molekul vode se razdvaja na H i OH, koji se vezuju za fragmente molekula i dovode do prekida kovalentne veze između monomera.

1.6 Strukture bioloških polimera

Konfiguracija molekula je trajna geometrija koja je rezultat prostornog rasporeda njegovih veza. To je trajni oblik molekula koji može da se promeni samo prekidanjem hemijskih veza, ne može samo rotacijom.

Sposobnost istog skupa atoma da formira dva ili više molekula različitih konfiguracija je *stereoizomerija*. Upotrebljena kao lekovi, jedinjenja različite konfiguracije obično imaju različitu fiziološku aktivnost, uključujući željeni farmakološki efekat, toksikologiju i metabolizam.



3.3 Condensation and Hydrolysis of Polymers (a) A condensation reaction links monomers into polymers. (b) A hydrolysis reaction digests polymers into individual monomers.

Slika 1.18: Izgradnja i razgradnja polimera

Konfiguracija se razlikuje od **hemijske konformacije**, oblika koji se postiže **rotacijom** veze.

U hemiji, **konformacioni izomerizam** je oblik stereoizomerizma u kome se izomeri mogu međusobno konvertovati isključivo rotacijama oko formalno jednostrukih (sigma) veza. Takvi izomeri se generalno nazivaju **konformacionim izomerima** ili **konformerima** i, posebno, **rotamerima**.

Slika 1.19 prikazuje podelu izomera, odnosno različite načine na koje molekuli mogu imati istu hemijsku formulu, ali različitu strukturu ili prostorni raspored atoma. Izomeri su važni za biološke polimere (proteine, DNK, polisaharide), jer njihov oblik direktno utiče na funkciju. Dve glavne grupe izomera su konstitucionim i stereoizomeri. Konstitucionim (strukturni) izomeri se razlikuju po tome kako su atomi povezani. Na primeru sa slike isti su atomi molekula, ali fluor (F) je vezan za različit C atom. Stereoizomeri (prostorni izomeri) imaju istu povezanost atoma ali se razlikuju u prostornom rasporedu.

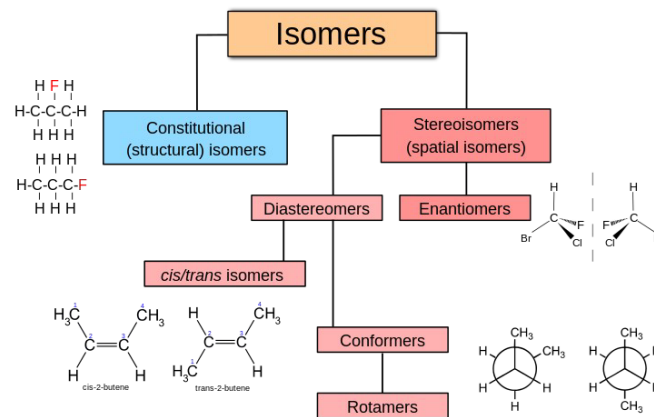
Vrste stereozomera su

- Enantiomeri - ogledalske slike molekula (kao leva i desna ruka)
- Diastereomeri nisu ogledalske slike ali imaju različit raspored u prostoru. U njih spadaju
 - cis/trans izomeri koji se pojavljuju kod dvostruke veze (C=C)
 - Konformeri koji nastaju rotacijom oko jednostruke veze, pri čemu se veza ne prekida. Rotameri su posebna vrsta konformera koji predstavljaju različite rotacione položaje.

Primarna struktura definiše **vrstu, redosled i broj** monomernih jedinica u polimernom lancu i ona je posledica primarnih (kovalentnih) interakcija između pojedinih mera (npr. u slučaju proteina **peptidna veza** između amino kiselina) u biopolimeru. U bioinformatički primarna

Klasa	Molekulska masa (Da)	Veličina (nm)	Broj monomera	Tip monomera
Oligomeri Oligonukleotidi Oligopeptidi Oligosaharidi	10^2 – 10^4	~ 2	$< 10^2$	Nukleotidi α -aminokiseline Monosaharidi
Polimeri Nukleinske kiseline Proteini Polisaharidi	10^4 – 10^7	4–7	10^2 – 10^4	Nukleotidi α -aminokiseline Monosaharidi
Supramolekulski sklopovi Nukleoproteini Proteinski kompleksi Glikoproteini Biomembrane	10^5 – 10^{12}	20–100	10 – 10^2	Nukleinske kiseline, proteini Proteini Proteini, glikani Lipidi, proteini, glikani

Tabela 1.4: Klase biomakromolekula i njihovo poređenje

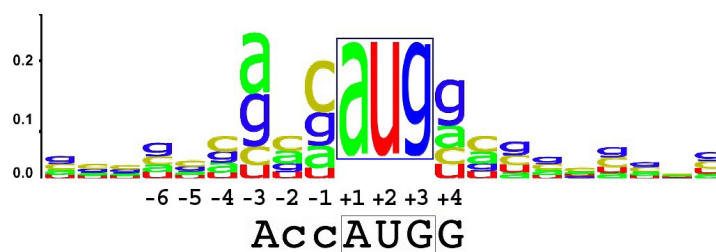


Slika 1.19: Izomeri

struktura se definiše **niskom** slova koja predstavljaju akronime trivijalnih imena monomernih jedinica datog polimera (npr. ACTG..., gde je A – Adenin, C – Citozin, itd., u smeru 5' → 3' (pravac rasta i čitanja DNK/RNK lanca), u slučaju nukleinskih kiselina; GAV..., G – Glicin, A – Alanin, itd., u smeru N-term → C-term., u slučaju proteina). Na nivou primarne strukture možemo definisati **sekvencione motive**, podniske koje imaju specifične biološke funkcije (npr. Shine-Dalgarno podnisika ili sekvenca (kratka sekvenca nukleotida u mRNA kod bakterija koja služi kao signal za početak sinteze proteina): AGGAGG, u slučaju nukleinskih kiselina; Zn++ finger motif (cinkov prst - strukturni motiv u proteinima koji vezuje jon cinka i omogućava proteinu da se pravilno savije i veže za DNK) CXX(XX)CXXXXXXXXXXXXHXXXH, u slučaju proteina, gde je C cistein, a H histidin). **Konsenzus sekvence** su konzervativni sekvencioni motivi. One su rezultat višestrukog poređenja sekvenci („multiple sequence alignment“) i sastoje se od podniski sa konzerviranim i varijabilnim monomernim jedinicama, npr. kod nukleinskih kiselina, u slučaju konsenzus sekvence A[CT]NAYR, A je stalno prisutno u svim poređenim sekvencama, [CT] označava C ili T, N bilo koji nukleotid, A bilo koji nukleotid osim A, Y je C

ili T, a R je A ili G. Podsećanje (A – Adenin, C – Citozin, T – Timin, G – Gvanin). U slučaju ovako predstavljene sekvence [CT] ne daje podatak o zastupljenosti pojedinih nukleotida; to se postiže grafičkim predstavljanjem tzv. sekvenca logoom.

Na slici 1.20 je sekvenca logo (sequence logo) i koristi se da prikaže koliko su određeni nukleotidi zastupljeni i važni na svakoj poziciji u sekvenci. Na dnu je sekvenca AccAUGG, a iznad svake pozicije nalaze se oznake nukleotida čija visina određuje koliko je taj nukleotid čest i važan na toj poziciji. Više slova na jednoj poziciji označava da postoji varijabilnost, tj. više nukleotida se pojavljuje na toj poziciji. Jedno veliko slovo označava da je ta pozicija konzervirana, tj. skoro uvek se pojavljuje isti nukleotid. AUG je start kodon (početak sinteze proteina).

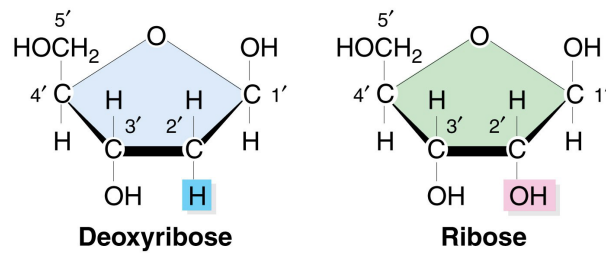


Slika 1.20: Sekvenca logo. Zastupljenost nukleotida oko inicijacionog kodona kod ljudskih mRNK molekula. (http://en.wikipedia.org/wiki/Sequence_logo)

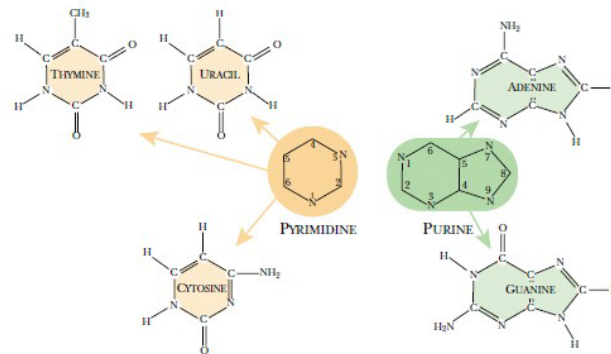
Sekundarnu strukturu definišu dihedralni uglovi konformeri (φ , ψ i ω peptidna veza kod polipeptida, α , β , γ , δ , ϵ , ζ fosfodiesteranski lanac kod polinukleotida), koji čine „kičmu“ polimera. Sekundarna struktura je posledica sekundarnih interakcija (formiranja nekovalentnih, vodoničnih veza) između atoma koji čine kičmu biopolimera (npr. u slučaju proteina sekundarna struktura i njen tip (zavojnice, trake, okreti, petlje) su definisani vodoničnim vezama amidne i karboksilne grupe koje čine kičmu molekula (peptidni lanac)). **Tercijarnu strukturu** definišu položaji svih atoma u prostoru, tj. njihova međusobna rastojanja i njihovi međusobni dihedralni uglovi. **Kvaternerna struktura** predstavlja više bioloških makromolekula međusobno povezanih nekovalentnim interakcijama (npr. hemoglobin se sastoji od dve α i dve β subjedinice, tj. α -globina i β -globina).

Nukleinske kiseline (DNK i RNK) su biološki makromolekuli polimerne prirode. Sastoje se od osnovnih jedinica koje se ponavljaju: **nukleotida**. Nukleotidi se sastoje od šećera **pentoze**, **riboze** u RNK ili **dezoksiriboze** u DNK (slika 1.21), **fosforne kiseline** i azotnih baza – **pirimidina** (citozin C, timin T i uracil U kod RNK) i **purina** (adenin A i guanin G) (slika 1.22). Azotne baze su N-glikozidnom vezom povezane sa C1' atomom u petočlanom prstenu šećera. Ovakvo jedinjenje se naziva nukleozid. Fosforna kiselina je povezana estarskom vezom za C5' atom šećera pentoze; ovakvo jedinjenje se naziva **nukleotid**. Formiranjem fosfodiesterarskih veza između dva nukleotida (na pozicijama C5' i C3') u procesu polimerizacije gradi se **polinukleotidni lanac**. Primarna struktura polinukleotidnog lanca je redosled nukleotida koji se „čita“ u smeru $5' \rightarrow 3'$. Dva polinukleotidna lanca se mogu međusobno povezati nekovalentnim interakcijama (vodonične veze) u **dvostruku zavojnicu**, tako što se purinska baza A sparuje sa pirimidinskom bazom T (u slučaju RNK U), i purinska baza G sa pirimidinskom bazom C.

Slika 1.23 ilustruje Watson–Crick model sparivanja baza u DNK, gde se adenin sparuje sa timinom formirajući dve vodonične veze, dok se gvanin sparuje sa citozinom formirajući



Slika 1.21: Pentoze



The four bases of DNA are adenine, guanine, cytosine and thymine. In RNA, uracil replaces thymine. Pyrimidine bases contain one-ring structures, whereas purine bases contain two-ring structures.

Slika 1.22: Baze nukleinskih kiselina

tri vodonične veze. Uprkos razlici u broju veza, oba para imaju gotovo identičnu prostornu dimenziju zahvaljujući sparivanju purina sa pirimidinom, što omogućava uniformnu strukturu dvostruke zavojnice. Geometrija i raspored vodoničnih veza određuju specifičnost i stabilnost ovih interakcija.

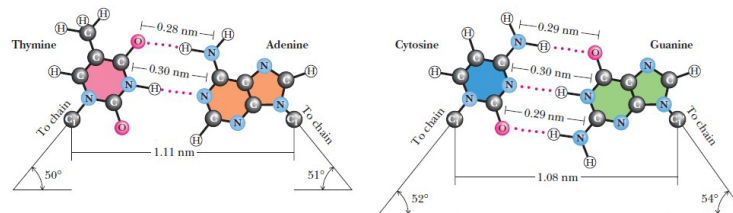
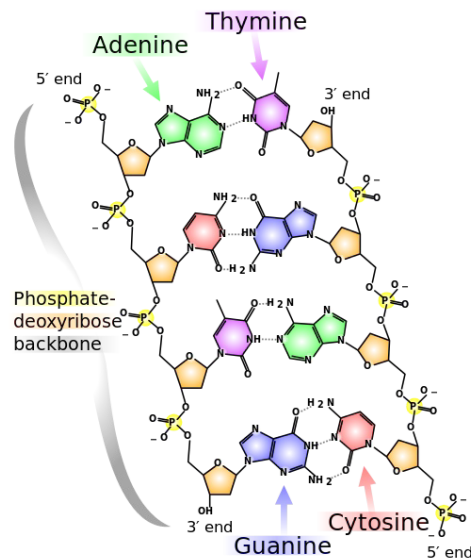


FIGURE 10.17 The Watson-Crick base pairs A:T and G:C.

Slika 1.23: Watson-Crick sparivanja baza A-T i G-C

Slika 1.24 prikazuje deo DNK molekula sastavljen od dva antiparalelna polinukleotidna lanca povezana vodoničnim vezama između komplementarnih baza. Šećerno-fosfatna kičma čini spoljašnji deo molekula, dok su baze usmerene ka unutrašnjosti i sparuju se prema pravilima Watson-Crick modela: adenin sa timinom (dve vodonične veze) i gvanin sa citozinom (tri vodonične veze). Nukleotidi u lancu povezani su fosfodiesterским vezama između 3' i 5' ugljenikovih atoma, a lanac ima definisan smer $5' \rightarrow 3'$.

Slika 1.25 prikazuje Watson-Crick sparivanje baza i hemijsku osnovu formiranja vodoničnih veza između komplementarnih baza. Adenin i timin formiraju dve vodonične veze, dok gvanin i citozin formiraju tri, što doprinosi većoj stabilnosti G-C parova. Elektrostatičke mape prikazuju



Slika 1.24: Prikaz jednog segmenta DNK

raspodelu parcijalnih naelektrisanja (δ^+ i δ^-), koja omogućava specifično i selektivno sparivanje baza, zasnovano na komplementarnosti donorskih i akceptorskih mesta za vodonične veze.

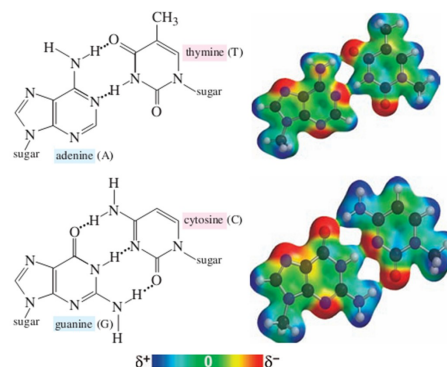


Figure 24-18 The two kinds of hydrogen-bonded base-pairing that occur in DNA. In RNA, adenine forms a base pair with uracil instead of with thymine.

Slika 1.25: Sparivanje baza u DNK

Na slici 1.26 su prikazani strukturni parametri dvostruke spirale DNK. Watson–Crick dvostruka spirala sastoji se od približno 10,5 baznih parova po jednom zavoju. Pošto 360° predstavlja jedan zavoj, svaki susedni bazni par rotira za oko 34.3° . Visina jednog zavoja spirale iznosi 35.7 \AA . \AA (angstrom, Ångström) je jedinica za dužinu koja se koristi za opisivanje veličine atoma i molekula. $1 \text{ \AA} = 10^{-10} \text{ m} = 0.1 \text{ nm}$ Rastojanje između susednih baznih parova duž ose iznosi 3.4 \AA . Bazni parovi su blago nagnuti (6°) u odnosu na horizontalnu ravan. Prečnik spirale iznosi oko 20 \AA . Na slici su prikazani i veliki i mali žleb, koji imaju važnu ulogu u vezivanju proteina.

Sekundarna struktura nukleinskih kiselina obuhvata različite prostorne organizacije koje nastaju kao posledica intramolekulskog sparivanja baza (slika 1.27). Pored klasične dvostruke spirale (dupleksa), javljaju se i jednolančani regioni, ukosnice (hairpin strukture), ispupčenja (bulges), unutrašnje petlje i razgranate strukture (junctions). Ove strukture su posebno izražene kod RNK i imaju važnu ulogu u stabilnosti molekula, kao i u njegovoj funkciji i interakciji sa

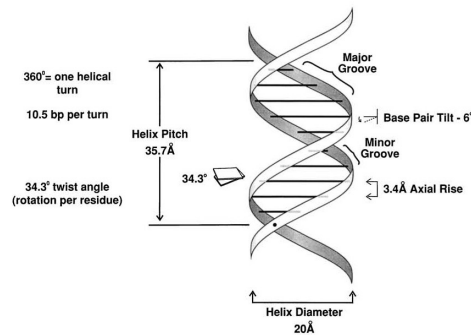
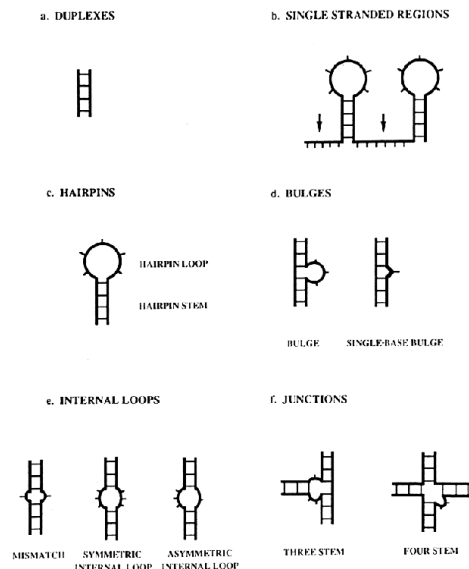


Figure 1.12 The DNA double helix in solution: structural parameters. The Watson-Crick double helix is composed of about 10.5 base pairs per helical turn. Since 360° constitutes one helical turn, there would be a 34.3° twist angle or rotation per residue between adjacent base pairs (see Table 1.4). The helix pitch or length per helical turn is 35.7 Å. The axial rise or distance between two planar base pairs is 3.4 Å. The base pair tilt or deviation from the horizontal plane of the bases is about -6° . The helix diameter or the width of the helix is about 20 Å. Note the positions of the minor groove and the major groove.

Slika 1.26: Dvostruka spirala DNK

proteinima.

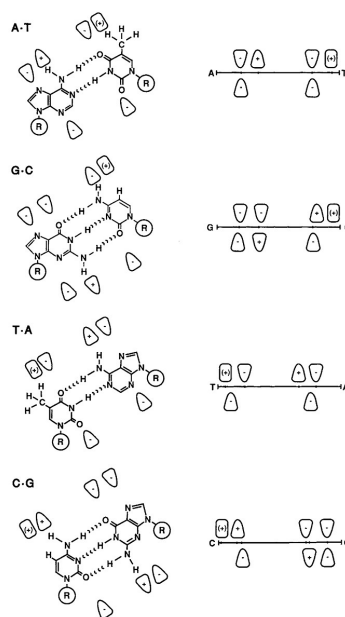


Slika 1.27: Helikoidne strukture kod nukleinskih kiselina

1.7 Interakcija sa drugim molekulima

Na slici 1.28 su prikazana mesta za formiranje vodoničnih veza u velikom i malom žlebu. Bazni parovi A · T, T · A, G · C i C · G prikazani su sa potencijalnim mestima za formiranje vodoničnih veza dostupnim proteinima kako u velikom tako i u malom žlebu. Pozicije donora vodoničnih veza označene su trouglom sa znakom „+“, usmeren od vodonika donora. Akceptori vodoničnih veza označeni su trouglom sa znakom „-“, usmeren ka centru bogatom elektronima. Mesta koja ne učestvuju u vodoničnim vezama, ali mogu učestvovati u elektrostatičkim interakcijama, prikazana su pravougaonicima. Na desnoj strani slike je dat linearni prikaz redosleda potencijalnih mesta za formiranje vodoničnih veza u velikom žlebu (gore) i malom žlebu (dole) za svaki bazni par. Svaki od četiri bazna para poseduje jedinstven obrazac potencijala za vezivanje vodoničnih veza. Definisana sekvenca DNK određuje jedinstvenu trodimenzionalnu površinu proteina koja je

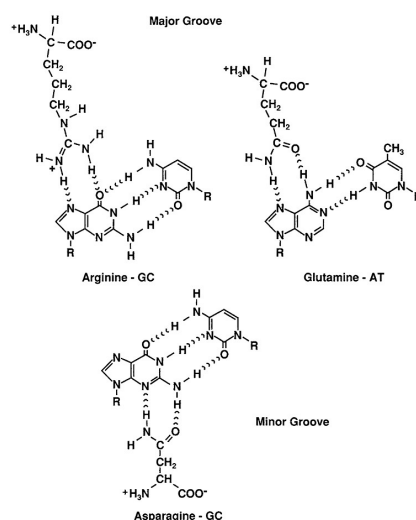
potrebna za interakciju specifičnu za sekvencu.



Slika 1.28: Mesta za formiranje vodoničnih veza u velikom i malom žlebu u

Proteini mogu da čitaju DNK bez razdvajanja lanaca.

Slika 1.29 prikazuje kako proteini prepoznaju DNK sekvencu putem vodoničnih veza između aminokiselina i azotnih baza u velikom i malom žlebu. U velikom žlebu baze imaju jedinstven raspored donorskih i akceptorskih mesta, što omogućava specifično prepoznavanje baznih parova (npr. arginin-G · C, glutamin-A · T). U malom žlebu informacije su manje specifične, pa je prepoznavanje ograničeno. Specifično prepoznavanje DNK od strane proteina putem vodoničnih veza u velikom žlebu omogućava regulaciju genske ekspresije i pravilno funkcionisanje ćelije.



Slika 1.29: Protein-DNK vezivanje

1.8 Nukleotidni kod (IUPAC)

Međunarodna unija za čistu i primenjenu hemiju (IUPAC) definiše sledeće skraćenice za nukleotide:

1.8.1 Osnovni nukleotidi

- A – Adenin
- C – Citozin
- G – Gvanin
- T – Timin
- U – Uracil

1.8.2 Dvosmisleni (degenerisani) simboli

- R – A ili G (purin)
- Y – C ili T (U) (pirimidin)
- M – A ili C (amino)
- K – G ili T (U) (keto)
- S – C ili G (jake veze, 3 vodonične veze)
- W – A ili T (U) (slabe veze, 2 vodonične veze)

1.8.3 Isključujući simboli

- B – C ili G ili T (U) (nije A)
- D – A ili G ili T (U) (nije C)
- H – A ili C ili T (U) (nije G)
- V – A ili C ili G (nije T (U))

1.8.4 Ostalo

- N – A ili C ili G ili T (U) (bilo koji nukleotid)
- . ili - – praznina (gap)

1.9 Konvencije za pisanje DNK sekvenci

DNK sekvence se zapisuju prema standardizovanim pravilima koja definišu način predstavljanja polariteta, baznih parova, oligonukleotida i ponavljajućih sekvenci.

1.9.1 Polaritet sekvence

Sekvenca DNK se piše sleva nadesno u smeru $5' \rightarrow 3'$. Na primer, sekvenca GGAATTCC predstavlja zapis:



1.9.2 Bazni parovi

Bazni parovi označavaju se tačkom između baza, čime se ukazuje na njihovo sparivanje putem vodoničnih veza:

- A · T
- G · C

1.9.3 Oligonukleotidi

Oligonukleotidi predstavljaju kratke sekvence nukleotida i mogu se zapisivati na dva načina:

- pomoću crtice između baza
- kao uzastopna slova

Primeri:

- **Dinukleotidi:** A–T, G–C ili AT, GC
- **Trinukleotidi:** G–A–T ili GAT

1.9.4 Ponavljajući oligonukleotidi

Ponavljajuće jedinice u DNK polimerima označavaju se pomoću notacije $\text{poly}(x)$, gde x predstavlja bazu ili niz baza koji se ponavlja.

- Mononukleotid: $\text{poly}(A)$
- Dinukleotid: $\text{poly}(AT)$
- Trinukleotid: $\text{poly}(GAT)$

1.9.5 Dvostruko-lančani ponavljajući polimeri

Kod dvostruko-lančanih DNK molekula, komplementarni lanci se prikazuju zajedno, pri čemu se njihovo sparivanje označava tačkom. Sekvence se i dalje pišu u smeru $5' \rightarrow 3'$.

- Mononukleotid: $\text{poly}(A) \cdot \text{poly}(T)$
- Dinukleotid: $\text{poly}(AT) \cdot \text{poly}(AT)$
- Trinukleotid: $\text{poly}(GAT) \cdot \text{poly}(ATC)$

1.9.6 Napomene

- Da bi se razlikovala DNK od RNK, može se koristiti oznaka d (deoksi), npr. dA-T ili dG-dC. Kada je jasno da se radi o DNK, ova oznaka se često izostavlja.
- Alternativni zapisi uključuju: poly(dA-dT) i poly(dA) · poly(dT).

Standardizovan zapis DNK sekvenci omogućava jasno i jednoznačno predstavljanje strukture i organizacije nukleinskih kiselina, uključujući smer lanca, komplementarnost baza i ponavljajuće obrasce u sekvenci.

Sledeće skraćenice se obično koriste za opisivanje dužine D/RNK molekula:

- bp = base pair(s) — jedna bp odgovara približno 3.4 Å (340 pm) dužine lanca (1Å=10⁻⁸m, 1pm=10⁻¹⁰m)
- kb (= kbp) = kilo base pairs = 1,000 bp
- Mb = mega base pairs = 1,000,000 bp
- Gb = giga base pairs = 1,000,000,000 bp.

Za slučaj jednolančane DNK/RNK koriste se jedinice nukleotida, skraćeno nt (ili knt, Mnt, Gnt), jer nisu uparene. Da bi se razlikovale jedinica računarske memorije i baza, za bazne parove, mogu se koristiti oznake kbp, Mbp, Gbp, itd. Centimorgan se takođe često koristi za rastojanje duž hromozoma, ali broj baznih parova kojima on odgovara uveliko varira. U ljudskom genomu centimorgan ima oko 1 milion parova baza.

Slika 1.30 prikazuje James Watson i Francis Crick sa modelom DNK iz 1953. godine, koji predstavlja jedno od najznačajnijih otkrića u istoriji biologije. Njihov model dvostruke spirale pokazao je da je DNK sastavljena od dva antiparalelna lanca povezana komplementarnim baznim parovima (A-T i G-C), čime je objašnjen način čuvanja i prenošenja genetičke informacije. Otkriće strukture DNK postavilo je temelje moderne molekularne biologije, genetike i biotehnologije.

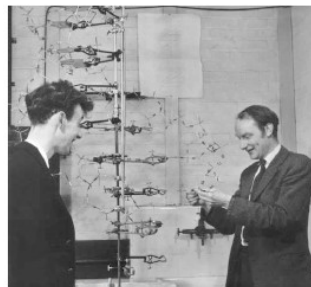


FIGURE 4.01 Watson and Crick in the 1950s

James Watson (b.1928) at left and Francis Crick (b.1916), with their model of part of a DNA molecule in 1953. Courtesy of: A. Barrington Brown, Science Photo Library.

Slika 1.30: Watson i Crick 1953.

Uvod u genomiku

2.1 Osnovni pojmovi - genotip, fenotip, gen

Genotip je genska konstitucija nekog organizma koja može da se odnosi na:

- jedan par alela, što predstavlja uži smisao genotipa;
- celovitu naslednu osnovu (sve gene koje taj organizam poseduje), što predstavlja širi smisao genotipa.

Aleli ili alelomorfi su različiti oblici jednog istog gena koji se nalaze na homolognim hromozomima. Npr. ako jedan hromozom nosi gen za boju očiju (npr. braon), a drugi za istu osobinu ali drugačiju varijantu (plava), oni su homologni jer nose isti gen na istom mestu, ali različite alele.

Key Concept

- Genotip – skup gena koje nosi pojedinačna ćelija ili organizam [1].
- Genotip je opis genetičkog sastava organizma [2].

Fenotip (gr. phainein – pokazati + typos – tip, vrsta) je skup svih osobina jednog organizma koje su nastale zajedničkim delovanjem genotipa i uslova sredine u kojima se dati organizam razvija (genotip + sredina -> fenotip). To su vidljive osobine organizma, kao što je npr. morfologija, razviće, ponašanje, karakteristična biohemija ili fiziologija. Fenotip se, osim u širem smislu, može posmatrati slično kao i genotip, i u užem smislu, kao jedna osobina koju određuje neki genotip. Na slici 2.1 se mogu videti fenotipske varijacije unutar iste vrste. za šare i boje školjke *Donax variabilis*. Različite školjke vrste *Donax variabilis* se razlikuju po boji (svetle, tamne, šarene) i šarama (linije, tačke, kombinacije). Deo razlika dolazi iz gena, a deo iz uticaja okoline (temperatura, hrana, uslovi života. . .).

Key Concept

- Fenotip: Uočljiva karakteristike ćelije ili organizma [1].
- Fenotip: Uočljive karakteristike ćelije ili organizma [2].

Gen je molekulska (fizička i funkcionalna) jedinica nasleđivanja svih organizama. Geni



Slika 2.1: Fenotipske varijacije (šare i boje) školjke (beskičmenjak, mekušac) *Donax variabilis*

se sastoje od poruke, informacije u obliku linearne niske nukleotida DNA, koja se u procesu nasleđivanja prenosi iz generacije u generaciju. Geni su nanizani duž i čine hromozom. Gen za određeno svojstvo uvek se nalazi na istom mestu na hromozomu koje se naziva *genski lokus*.

Key Concept

- Gen – region DNK koji kontroliše određenu naslednu osobinu organizma i obično je odgovoran za određivanje jednog proteina ili RNK molekula [1].
- Gen – jedinica nasleđivanja koja sadrži uputstva koja određuju osobine ili fenotip organizma; u molekularnom smislu, segment DNK koji usmerava proizvodnju proteina ili funkcionalnog RNK molekula [2].

Savremene definicije gena su:

- *A gene is a locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions, and or other functional sequence regions [5, 6]*
 - Gen je locirajući (prepoznatljiv) region genomske sekvence koji odgovara jedinici nasleđivanja i koji je povezan sa regulatornim regionima, transkribovanim regionima i/ili drugim funkcionalnim regionima sekvence
- *A gene is a DNA sequence (whose component segments do not necessarily need to be physically contiguous) that specifies one or more sequence-related RNAs/proteins that are both evoked by genetic regulatory networks (GRNs) and participate as elements in GRNs, often with in-direct effects, or as outputs of GRNs, the latter yielding more direct phenotypic effects [7].*
 - Gen je DNK sekvenca (čiji sastavni delovi ne moraju nužno biti fizički povezani) koja određuje jednu ili više međusobno povezanih RNK/proteina, koji su aktivirani genetičkim regulatornim mrežama (GRN) i učestvuju kao elementi u tim mrežama, često sa indirektnim efektima, ili kao njihovi izlazi, pri čemu ovi poslednji daju

direktnije fenotipske efekte.

Gen ima strukturalnu i funkcionalnu definiciju:

- **strukturalna:** segment DNA (hromozoma) koji se prepisuje (transkribuje) u RNA.
- **funkcionalna:** segment DNA koji obavlja neku funkciju. Jedinica nasleđivanja odgovorna za određenu fenotipsku karakteristiku organizma.

Prema funkciji koju obavljaju, gene možemo podeliti na strukturne i regulatorne.

Strukturni geni su:

- geni koji nose šifru za protein (prepisuju se na mRNA, (protein) kodirajući geni/DNA/RNA) i
- geni koji se samo prepisuju u RNA (tRNA, rRNA)(ne-kodirajući geni-DNA-RNA).

Regulatorni geni (ne-kodirajući geni/DNA) ili **regulatorne sekvence/podniske**, se ne prepisuju već se za njih vezuju molekuli koji regulišu određene procese u organizmu. Znanja o regulatornim genima su još uvek nedovoljna. Pripadaju im geni koji:

- učestvuju u replikaciji,
- učestvuju u kretanju hromozoma u deobi,
- kontrolišu krosing-over i telomere koje doprinose stabilnosti hromozoma. Krosing-over je proces u kome dolazi do razmene delova DNK između homolognih hromozoma. Telomere su krajnji delovi hromozoma koji štite krajeve hromozoma i genetsku informaciju tokom replikacije i sprečavaju spajanje hromozoma.

Hromozom je kompleks DNA i proteina sa sledećim osobinama koje se razlikuju kod organizama:

- veličina. Kod bakterija je 100.000 bp (bazni parovi), a kod biljaka do 3.750.000.000 bp,
- broj. Bakterija ima 1, a biljke do 1200; i
- topologija (linearni – prokariotski i eukariotski organizmi, cirkularni – prokariotski organizmi) varira od vrste do vrste.

Ploidija predstavlja broj kompletnih setova hromozoma u ćeliji. Jedan set (n) sadrži po jedan primerak svakog hromozoma karakterističnog za vrstu. Razlikuju se

- Haploidni (n) – ćelije koje sadrže jedan set hromozoma. Primer: gamete (polne ćelije); prokarioti se najčešće smatraju haploidnim.
- Diploidni ($2n$) – ćelije koje sadrže dva seta hromozoma, odnosno parove homologih hromozoma (jedan poreklom od majke, drugi od oca). Primer: telesne ćelije većine eukariota.
- Poliploidni ($>2n$) – ćelije koje sadrže više od dva seta hromozoma (npr. $3n$, $4n$, $6n$, $12n$). Poliploidija je česta kod biljaka i može doprineti većoj veličini ćelija i evolutivnim promenama.

Kod bakterija se obično govori o haploidnosti, ali neke vrste mogu imati više kopija genoma u jednoj ćeliji, što ne odgovara u potpunosti klasičnoj poliploidiji kod eukariota.

2.2 Genom, sadržaj genoma

2.2.1 Genom

Genom je skup gena koje sadrži jedna haploidna (haploid - uobičajen broj hromozoma neke vrste) ćelija.

Reč genom se kod eukariotskih organizama može odnositi na DNA unutar jedra (jedarni genom) ali i na genome organela – mitohondrijalni i hloroplastni genom. Organele poput mitohondrija i hloroplasta poseduju sopstveni genom i mogu se deliti nezavisno od jedra, ali njihova funkcija i regulacija zavise od jedarnog genoma. Kod prokariota termin genom može obuhvatati i nehromozomske genetičke elemente – plazmide, epizome, transpozome. Takođe može se govoriti o genomu pojedinačne ćelije, organizma, vrste, ili u slučaju virusa, o viralnom genomu (tačnije pseudogenomu).

Key Concept

- Genom – ukupna genetička informacija koju nosi ćelija ili organizam. [1].
- Genom – ukupna genetička informacija koju nose svi hromozomi jedne ćelije ili organizma. [2].

Bioinformatics Note

Hans Winkler, 1920: *I propose the expression Genom for the haploid chromosome set, which, together with the pertinent protoplasm, specifies the material foundations of the species ...*

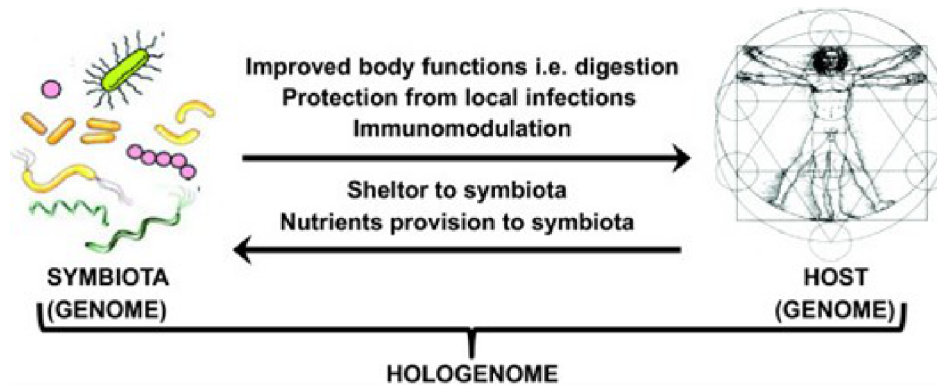
- Prevod: Predlažem izraz „genom“ za haploidni skup hromozoma, koji zajedno sa odgovarajućom protoplazmom određuje materijalne osnove vrste.

Organizam (npr. čovek) nije sam, već funkcioniše zajedno sa svojim mikroorganizmima kao jedna celina. **Holobiom** predstavlja zajednicu organizama koja se sastoji od domaćina i svih mikroorganizama koji žive u njemu ili na njemu. Simbiota označava sve organizme koji žive u simbiozi sa domaćinom. Mikrobiom čine svi mikroorganizmi koji žive u nekom okruženju ili organizmu, zajedno sa njihovim genima.

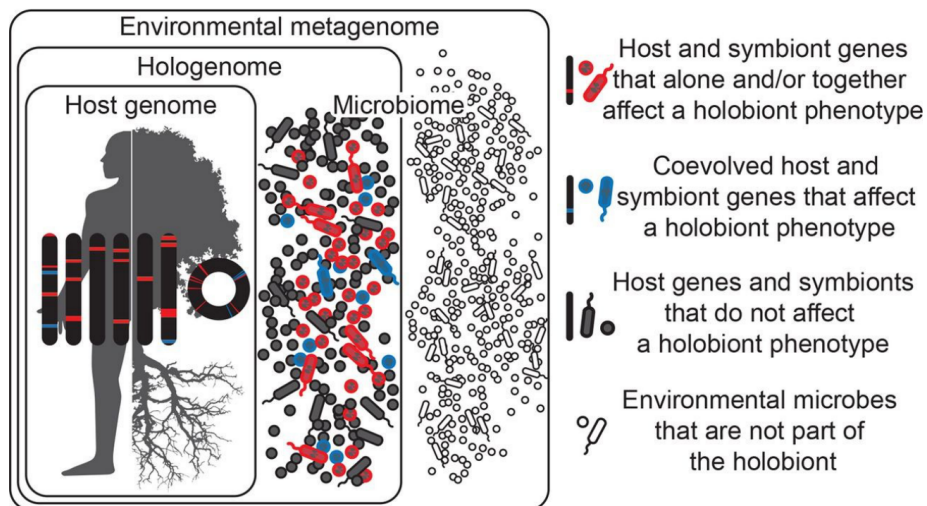
Hologenom predstavlja ukupan skup genoma holobiomta, koji obuhvata genom domaćina i genome svih mikroorganizama i virusa koji žive u njemu ili na njemu, i zajedno čine funkcionalnu celinu. Holobiom i hologenom su predstavljeni na slici 2.2

Na slici 2.3 je prikazan holobiom, tj. skup organizama koji se sastoji od domaćina i vrsta (mikroorganizmi, virusi) koje žive u zajednici sa njim – bionta (u njemu i oko njega). Sadržaj bionta (simbiota) se može menjati tokom vremena, prenositi vertikalno ili horizontalno, može biti koristan ili štetan po domaćina.

Epigenetičko (gr. epi – iznad, okolo) nasleđivanje predstavlja nasledne promene u fenotipu ili na genu, bez promena u sekvencama DNA (mutacije). Takve promene mogu biti vidljive nakon nekoliko generacija ćelija, ili čak više generacija živih bića. Epigenetički mehanizmi uključuju metilacija DNA (hemijska modifikacija DNK pri kojoj se na određene nukleotide (najčešće citozin)



Slika 2.2: Prikaz hologenoma. Izvor: https://www.researchgate.net/figure/The-concept-of-hologenome-Hologenome-consist-of-genome-of-holobiont-which-includes-host_fig1_236928826



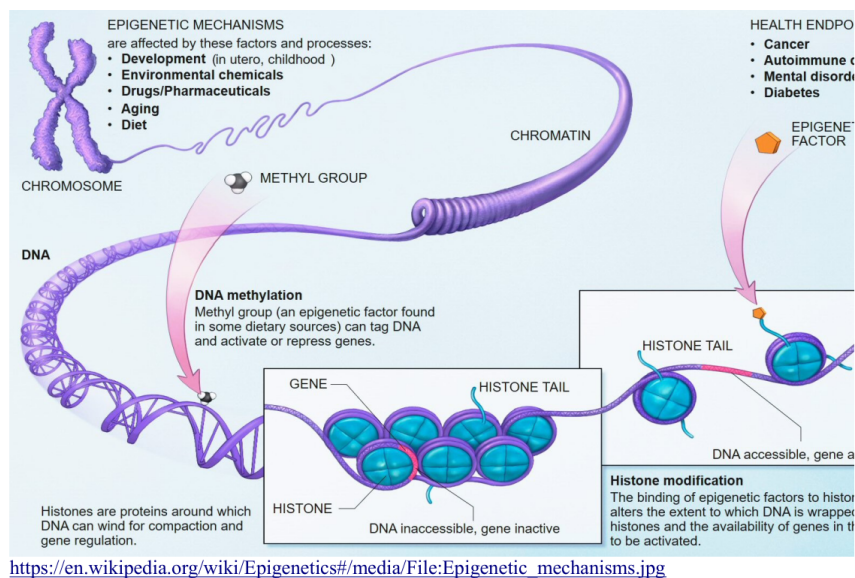
<https://msystems.asm.org/content/1/2/e00028-16>

Slika 2.3: Holobionti. Izvor <https://msystems.asm.org/content/1/2/e00028-16>

dodaje metil grupa ($-CH_3$)), procese na nivou RNA (posebno RNA interferencija - biološki proces kojim se specifični geni „utišavaju“ (inhibira se njihova ekspresija)), obrade proteina, inaktivacija polnih hromozoma, itd. Epigeni faktori mogu biti od velikog značaja u procesima starenja ili malignih transformacija.

Proces prikazan na slici 2.4 pokazuje kako epigenetički mehanizmi regulišu aktivnost gena promenom dostupnosti DNK. Kada je DNK čvrsto upakovana oko histona (kondenzovan hromatin), ona je nedostupna i gen je neaktivan. Suprotno tome, kada je hromatin opušten, DNK je dostupna i gen može biti aktivan i transkribovan u RNK. Metilacija DNK obično dovodi do utišavanja gena, dok modifikacije histona mogu povećati ili smanjiti ekspresiju gena. Na ove procese utiču spoljašnji faktori kao što su ishrana, stres i hemikalije iz okoline. Kao rezultat, isti gen može biti uključen ili isključen u različitim ćelijama ili uslovima, bez promene same DNK sekvence.

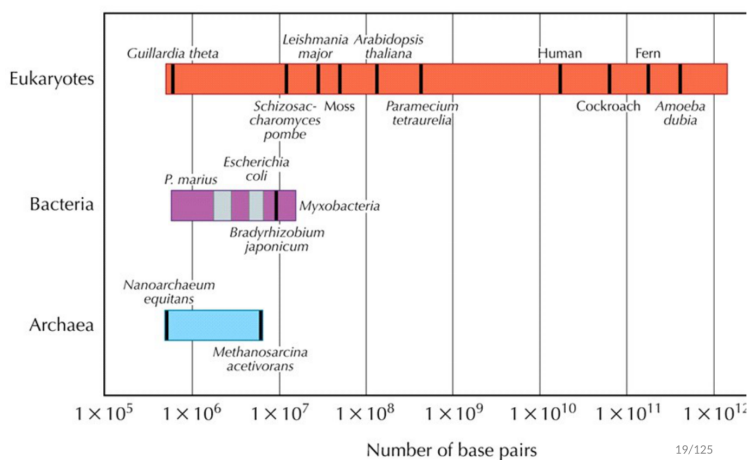
Sadržaj haploidnog genoma je označen terminom **genomski sastav** (sadržaj), (eng. genome composition) i uključuje veličinu genoma, sadržaj neponavljajućih i ponavljajućih sekvenci



Slika 2.4: Epigenetičko nasleđivanje. Izvor https://en.wikipedia.org/wiki/Epigenetics#/media/File:Epigenetic_mechanisms.jpg

(podniski) DNA kao i sadržaj genskih ((protein) kodirajućih/nekodirajućih) i međugenskih sekvenci.

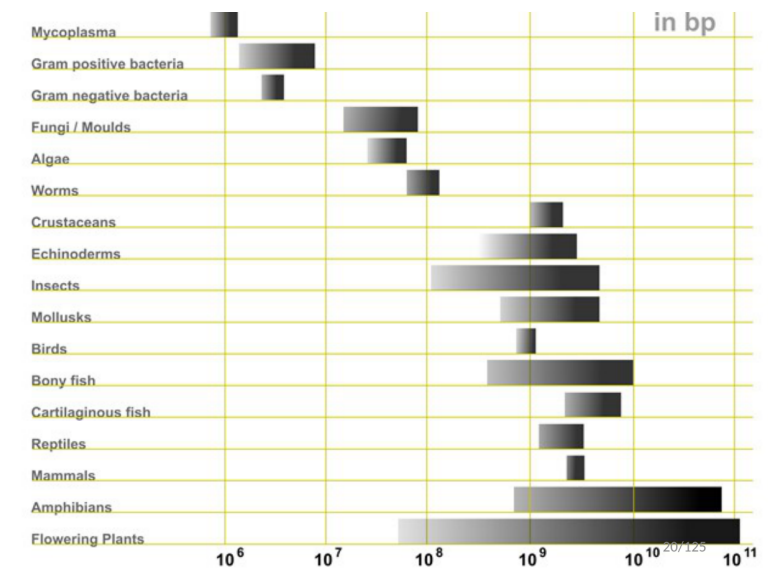
Veličina genoma (eng. genome size) predstavlja (ukupan) sadržaj DNA jednog genoma izražen kao masa (u pikogramima, pg, 10^{-12} gr), ili kao broj baznih parovima (u megabazama, Mbp, 1 pg = 978 Mbp). Kod diploidnih organizama koristi se i termin C-vrednost (eng. C-value). Složenost organizma nije srazmerna veličini genoma, neki jednoćelijski ili biljni organizmi imaju genome veće od ljudskog genoma (veća količina DNA), (Junk DNA?? ili C-value enigma). Veličina genoma može da varira i unutar vrste (individualne varijacije, pored pojedinačnih nukleotidnih zamena (SNP - single nucleotide polymorphysm) i pojedinačnih insercija i delecija (IN/DEL - insertion/deletion), kod humanih genoma su, npr. pronađene IN/DEL varijacije od 360 Mbp. Na slici 2.5 su prikazane veličine genoma za arheje, bakterije i eukariote.



Slika 2.5: Veličina genoma

Na slici 2.6 pre prikazana veličina genoma za razne organizme data kao ukupan broj baznih

parova (haploidan). Na slici 2.7 je prikazan broj ćelija i broj gena za razne organizme.



Slika 2.6: Veličina genoma za razne organizme

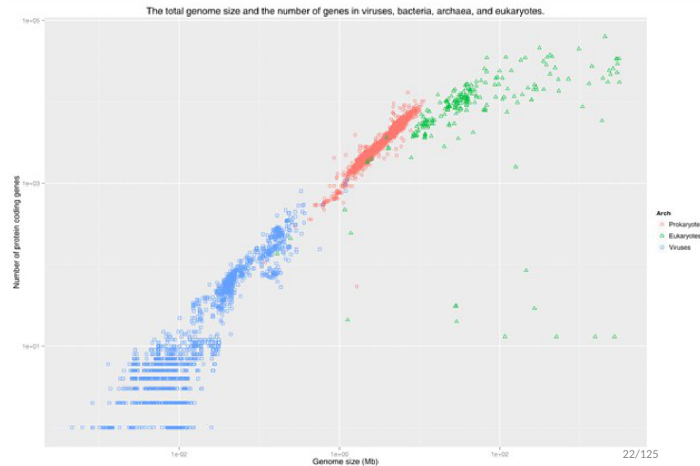
Organism	Number of Cells In Adult*	Number of Genes
<i>Mycobacterium genitalium</i>	1	523
Pathogenic bacterium		
<i>Methanococcus jannaschii</i>	1	1,800
Archaeal methanogen		
<i>Escherichia coli</i> K12	1	4,400
Intestinal bacterium		
<i>Saccharomyces cerevisiae</i>	1	6,000
Baker's yeast (eukaryote)		
<i>Caenorhabditis elegans</i>	959	19,000
Nematode worm		
<i>Drosophila melanogaster</i>	10 ⁴	13,500
Fruit fly		
<i>Arabidopsis thaliana</i>	10 ⁷	27,000
Flowering plant		
<i>Fugu rubripes</i>	10 ¹²	38,000 (est.)
Pufferfish		
<i>Homo sapiens</i>	10 ¹⁴	20,500 (est.)
Human		

The first four of the nine organisms in the table are single-celled microbes; the last six are eukaryotes; the last five are multicellular, four of which are animals; the final two are vertebrates. Although pufferfish and humans have roughly the same number of genes, the pufferfish genome, at 0.365 billion nucleotide pairs, is only one-eighth the size of the human genome.

*Numbers for *Arabidopsis thaliana*, the pufferfish, and human are "order-of-magnitude" rough estimates. 21/125

Slika 2.7: Broj ćelija i broj gena za razne organizme.

Slika 2.8 prikazuje odnos između veličine genoma i broja gena kod različitih grupa organizama (virusi, prokarioti i eukarioti). Uočava se da kod prokariota postoji relativno direktna veza između veličine genoma i broja gena, dok kod eukariota ta veza nije proporcionalna jer veliki deo genoma čine nekodirajuće sekvence. Ova slika ilustruje da veći genom ne znači nužno i veći broj gena, što je povezano sa tzv. C-value paradoksom. C-value paradoks predstavlja pojavu da veličina genoma (količina DNK u ćeliji) nije u korelaciji sa složenošću organizma.



Slika 2.8: Veličina genoma u funkciji broja gena data za viruse (plavo), prokariotske organizme (Arhee i Bakterije) (crveno) i eukariotske organizme (Eukarije) (zeleno).

2.2.2 GC sadržaj genoma

GC-sadržaj je udeo azotnih baza u DNA molekulu. Može se odnositi na specifični fragment DNA ili RNA molekula, ili na ceo genom. GC sadržaj se obično izražava u procentima, a u nekim slučajevima kao odnos. Procentni GC-sadržaj je:

$$\frac{G + C}{A + T + G + C} \times 100\%$$

dok je AT/GC odnos:

$$\frac{A + T}{G + C}$$

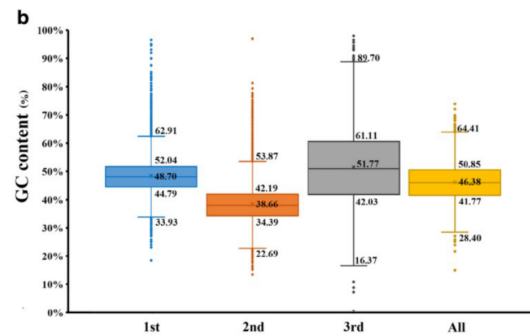
Smatra se da je GC sadržaj nekog organizma određen složenom kombinacijom selektivnih i neutralnih procesa koji rezultiraju u pomeranju baznog sadržaja genoma ka GC ili AT. Kod bakterija, npr. GC sadržaj genoma varira u opsegu od 25-75%. GC parovi međusobno intereaguju sa 3 vodonične veze, a AT sa 2 te se GC sadržaj odražava na stabilnost DNA i sekundarnu strukturu mRNA, kao i na temperaturu topljenja prilikom PCR eksperimenta.

Slika 2.9 prikazuje raspodelu GC sadržaja u genomu organizma kroz različite pozicije kodona (prva, druga i treća), kao i ukupni genom. Uočava se da GC sadržaj nije isti na svim pozicijama, što ukazuje na različite evolutivne pritiske i funkcionalne uloge pojedinih delova gena. Posebno je značajna treća pozicija kodona, koja često pokazuje najveću varijabilnost.

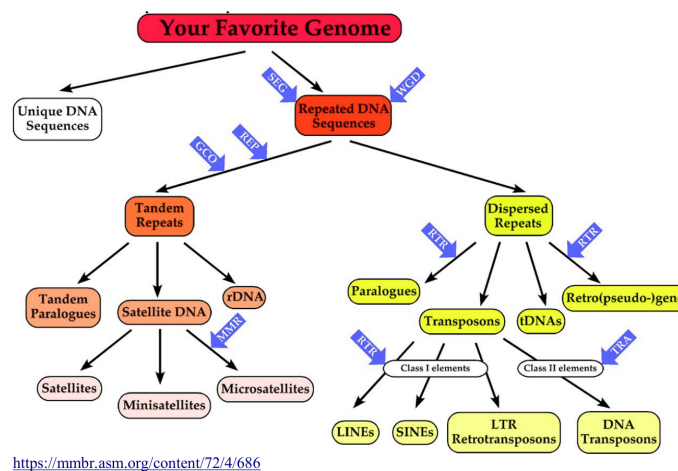
Ponavljajuće niske (eng. repeats) su podniske nukleinskih kiselina (DNA, RNA) koje se ponavljaju dva ili više puta unutar date nukleinske kiseline ili genoma. Nukleotidni sadržaj, dužina ponovka, broj ponovaka i raspored unutar nukleinske kiseline mogu da variraju. Postoje 4 tipa ponovaka: direktni (npr. ATCG — ATCG) i reversni (npr. ATCG — GCTA), komplementarni (npr. ATCG — TAGC, A—T, C—G) i nekomplementarni.

Slika 2.10 prikazuje različite tipove ponovaka u genomu eukariotskih organizama. Ponovci mogu biti raspoređeni na različite načine i čine značajan deo genoma, posebno kod eukariota. Njihova uloga može biti strukturna, regulatorna ili evolutivna.

Guan et al. BMC Genomics (2018) 19:542
Base composition of *H. manillensis*. (b) Box plot indicates GC content variation in different codon positions (1st, 2nd and 3rd represent the first, second and third codon position, respectively) and overall genome (All).



Slika 2.9: Raspodela GC sadržaja u genomu organizma. Izvor: Guan et al. BMC Genomics (2018) 19:542



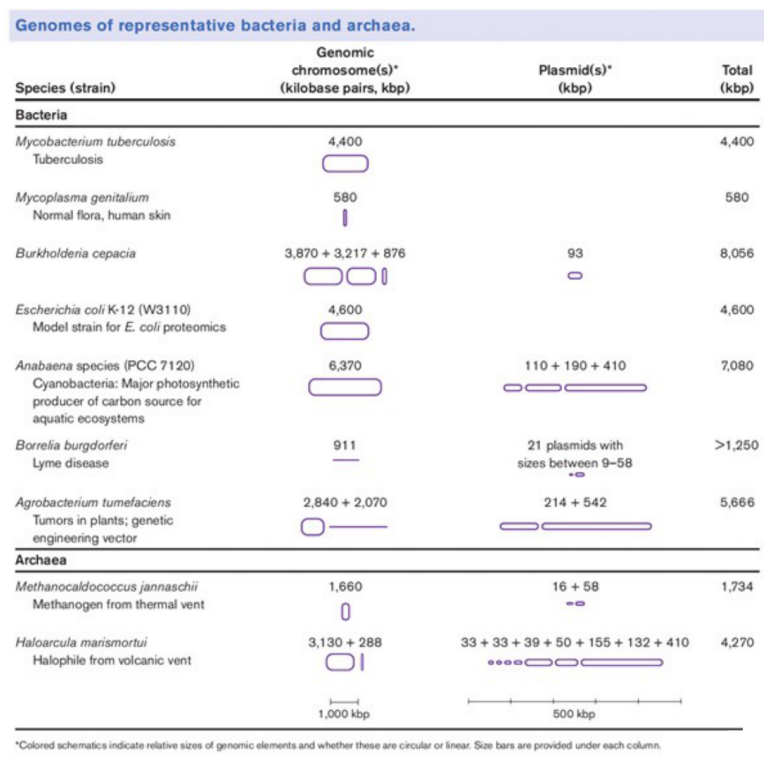
Slika 2.10: Ponovci kod eukariotskih organizama

2.2.3 Organizacija genoma kod prokariotskih organizama

Plazmidi su mali cirkularni (ili linearni) ekstrahromozomalni molekuli DNA prisutni u bakterijama, rede arheama i nekim eukarijama (npr. pivski kvasac). Plazmidi se replikuju nezavisno od hromozoma, mogu se integrisati u hromozom, prenose se u okviru iste, ali i različitih vrsta. Mogu doprineti preživljavanju, odnosno selektivnoj prednosti organizma nosioca plazmida. Slika 2.11 prikazuje osnovnu organizaciju genoma kod prokariotskih organizama, gde je genetski materijal najčešće organizovan u jedan cirkularni hromozom, uz moguće prisustvo plazmida. Genom je kompaktan, sa malim udelom nekodirajućih sekvenci, a geni su često organizovani u operone (funkcionalne jedinice DNK kod prokariota koje se sastoje od više gena koji se zajedno regulišu i transkribuju kao jedna celina). Ovakva struktura omogućava efikasnu regulaciju i brzu adaptaciju organizma.

2.2.4 Organizacija genoma kod eukariotskih organizama

Slika 2.12 prikazuje organizaciju genoma kod eukariotskih organizama, gde je DNK raspoređena u više linearnih hromozoma i povezana sa histonima u obliku hromatina. Genom je složeniji u odnosu na prokariote i sadrži veliki udeo nekodirajućih sekvenci, uključujući introne i repetitivnu DNK. Ovakva organizacija omogućava složenu regulaciju ekspresije gena i veću funkcionalnu



Slika 2.11: Organizacija genoma kod prokariotskih organizama (Arhea i Bakterija)

raznovrsnost.

2.2.5 Haplotip i haplogrupa

Haplotip („jednostruk, sam, jednostavan“) u genetici je kombinacija alela (DNA sekvenci) na susednim lokacijama (lokusima, mestima) na hromozomu koji se prenose zajedno. Haplotip može da bude jedan lokus, nekoliko lokusa, ili celokupni hromozom u zavisnosti od broja rekombinacija koje su se odigrale između datog seta lokusa. Haplotip (u molekularskoj biologiji) se definiše kao set jednonukleotidnih polimorfizama (SNP-ova) u jednom hromozomu iz hromozomskog para koji su statistički asociirani. Smatra se da te asocijacije, i identifikacija nekoliko alela haplotipnog bloka, mogu nedvosmisleno da identifikuju sva druga polimorfna mesta u datom regionu. Takva informacija je veoma korisna u ispitivanju genetičke osnove bolesti, i bila je istraživana za ljudsku vrstu u okviru Međunarodnog Hap-Map projekta.

Specijalizovana primena termina haplotip je za označavanje gena koji se *ne rekombinuju* i *nemaju svoj par na drugom hromozomu*. Kod čoveka takvi geni su locirani na Y hromozomu (Y-DNA) i mitohondrijalnoj DNA (mtDNA) i nazivaju se **haplogrupama**. U molekularskoj evoluciji, **haplogrupa** je grupa sličnih haplotipova koji imaju zajedničko poreklo na šta ukazuju isti SNP-ovi u oba haplotipa. Haplogrupe se označavaju slovima alfabeta, npr. *Y-DNA I* je među Srbima pojedinačno najzastupljenija haplogrupa.

Analize gena na Y hromozomu i gena mtDNA, zajedno sa utvrđivanjem vremena i mesta njihovog ispoljavanja, koriste se u analizi populacione dinamike (u slučaju ljudske populacije izvođenje zaključaka o mestima i periodima seoba današnjeg stanovništva). Y hromozom sa genima koji se ne rekombinuju (gen se nasleđuje kao celina i prenosi se nepromenjen kroz

TABLE 4.03 Components of the Eukaryotic Genome	
(Numbers of copies given is for the human genome.)	
<u>Unique sequences</u>	
Protein encoding genes—comprising upstream regulatory region, exons and introns	
Genes encoding non-translated RNA (snRNA, snoRNA, 7SL RNA, telomerase RNA, Xist RNA, a variety of small regulatory RNAs)	
Non-repetitive intragenic non-coding DNA	
<u>Interspersed Repetitive DNA</u>	
Pseudogenes	
Short Interspersed Elements (SINEs)	
Alu element (300 bp)	~1,000,000 copies
MIR families (average ~130 bp) (mammalian-wide interspersed repeat)	~400,000 copies
Long Interspersed Elements (LINEs)	
LINE-1 family (average ~800 bp)	~200,000–500,000 copies
LINE-2 family (average ~250 bp)	~270,000 copies
Retrovirus like elements (500–1300 bp)	~250,000 copies
DNA transposons (variable; average ~250 bp)	~200,000 copies
<u>Tandem Repetitive DNA</u>	
Ribosomal RNA genes	5 clusters of about 50 tandem repeats on 5 different chromosomes
Transfer RNA genes	multiple copies plus several pseudogenes
Telomere sequences	several kb of a 6 bp tandem repeat
Mini-satellites (= VNTRs)	blocks of 0.1 to 20 kbp of short tandem repeats (5–50 bp), most located close to telomeres
Centromere sequence (α -satellite DNA)	171 bp repeat, binds centromere proteins
Satellite DNA	blocks of 100 kbp or longer of tandem repeats of 20 to 200 bp, most located close to centromeres
Mega-satellite DNA	blocks of 100 kbp or longer of tandem repeats of 1 to 5 kbp, various locations

Slika 2.12: Organizacija genoma kod eukariotskih organizama

generacije (osim ako ne dođe do mutacije)) imaju samo muškarci - prenosi se sa oca na sina. Mitohondrijalni geni ne podležu rekombinaciji (mtDNA može da se rekombinuje, ali sa kopijama DNA unutar istog mitohondriona) i nasleđuju se samo od majke, u retkim slučajevima i od oca, ali je tonedovoljno istraženo).

2.3 Mutacije

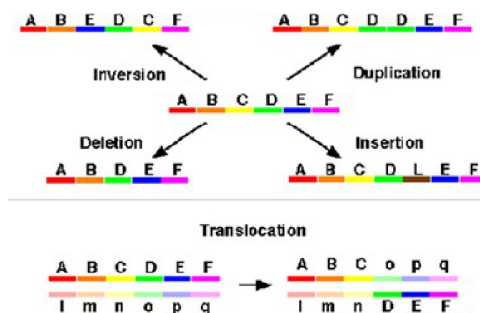
2.3.1 Hromozomske mutacije

Slika 2.13 prikazuje osnovne tipove hromozomskih mutacija: deleciju (gubitak dela hromozoma), duplikaciju (udvostručavanje segmenta), inserciju (dodavanje segmenta), inverziju (obrnuto redosled segmenata) i translokaciju (premeštanje segmenta između hromozoma). Ove promene utiču na strukturu i raspored gena, što može imati značajne posledice na funkciju organizma i pojavu bolesti.

Hromozomske mutacije zahvataju hromozome i vidljive su pod mikroskopom. Razlikuju se numeričke i strukturne promene u pojedinim parovima homologa u hromozomskoj garnituri.

*Numeričke hromozomske mutacije predstavljaju pojedine oblike **aneuploidije**:*

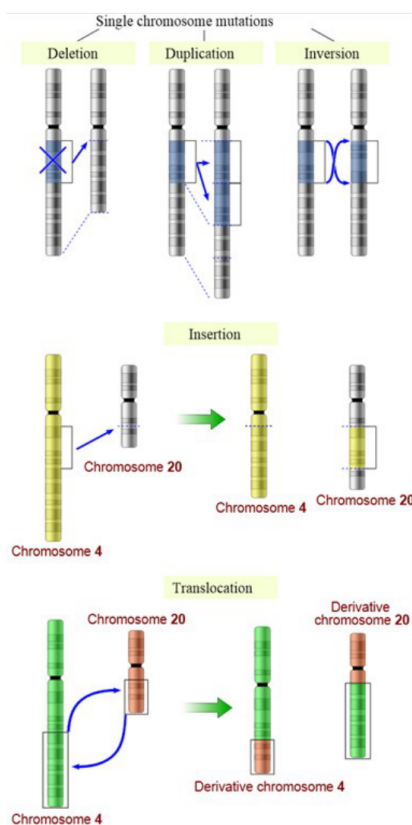
- nulisomiju (nedostatak oba homologa),
- monosomiju (nedostatak jednog homologa),
- trisomiju (jedan homolog više) i
- teorijski moguće stepene polisomije (kvadrisonija itd.).



Slika 2.13: Hromozomske mutacije

. Izvor <http://kmbiology.weebly.com/mutations---notes.html>

Slika 2.14 ilustruje različite tipove hromozomskih mutacija i aneuploidije, kao što su monosomija i trisomija. Prikazane su promene u strukturi hromozoma, kao i promene u broju hromozoma, što može dovesti do genetičkih poremećaja.



Slika 2.14: Hromozomske mutacije i aneuploidija

Genomske mutacije imaju za posledicu promenu normalnog diploidnog hromozomskog broja ($2n$) za n , tj. za jedan ili više haploidnih setova hromozoma (euploidija: haploidija, triploidija, poliploidija). Prema poreklu dodatnih garnitura, poliploidija može biti autopoliploidija (uvišestručavanje sopstvenog genoma) ili alopoliploidija (heteropoliploidija: dodavanje stranih genoma putem hibridizacije).

Strukturne hromozomske mutacije se pak ispoljavaju kao nedostatak (delecija – deficija) ili višak (duplikacija) jednog dela nekog hromozoma. U ovoj kategoriji mutacija relativno su česte i pojave nenormalnog rasporeda pojedinih delova unutar hromozoma (transpozicija) ili

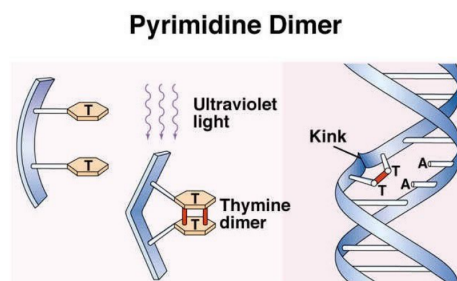
njihovog premeštanja na neki drugi hromozom (translokacija), te obrtanja pojedinih segmenata hromozoma (za 180° – inverzija), što rezultira obrnutim rasporedom lokusa na mutiranoj sekvenci.

Plazmatske mutacije (plazma-mutacije, ekstra-nuklearne mutacije) posebna su kategorija mutacija. Za razliku od genskih, hromozomskih i genomskih, koje se događaju u jedru, one odigravaju se u citoplazmatskim nosiocima genetičkog materijala (kao npr. u mitohondrijalnoj ili plastidnoj DNA).

2.3.2 Mutacije na nivou DNA

Mutacija je promena redosleda nukleotida u genomu organizma, ekstrahromozomalne DNA (plazmida) ili virusa.

Slika 2.15 prikazuje formiranje pirimidinskog dimera u DNK pod uticajem UV zračenja, gde dolazi do povezivanja susednih baza (najčešće timina) i deformacije DNK lanca. Ovakva oštećenja mogu ometati replikaciju i transkripciju, a ćelija ih popravlja mehanizmima reparacije DNK.



Slika 2.15: Mutacije na nivou DNA – pirimidinski dimer <https://sites.google.com/site/uvdnadamagerepair/uv-effects-on-dna/cpd-s>

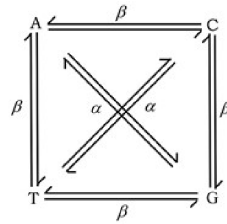
2.3.3 Pojedinačni nukleotidni polimorfizam – SNP

Jednonukleotidni polimorfizam (eng. single-nucleotide polymorphism, SNP) podrazumeva pojavu zamene jednog nukleotida sa nekim drugim nukleotidom. Pod **SNP** varijacijom se podrazumeva ako *više od 1% organizama populacije nema isti nukleotid* na istom mestu u nisci DNA. Ako se SNP nalazi unutar nekog gena, za taj gen se kaže da ima više alela (podsećanje: aleli su različiti oblici jednog istog (polimorfnog) gena). U humanoju populaciji je do sada otkriveno preko 10^6 SNP-ova.

Pojedinačna nukleotidna varijanta (SNV) je pojava zamene jednog nukleotida nekim drugim nukleotidom, *bez obzira na učestalost njegove pojave*. Ako se zamena pojavi u nekoj protein kodirajućoj niski (eng. open reading frame, ORF) i zavisno od toga na kojoj poziciji u kodonu se pojavi nukleotidna zamena, mogu rezultirati

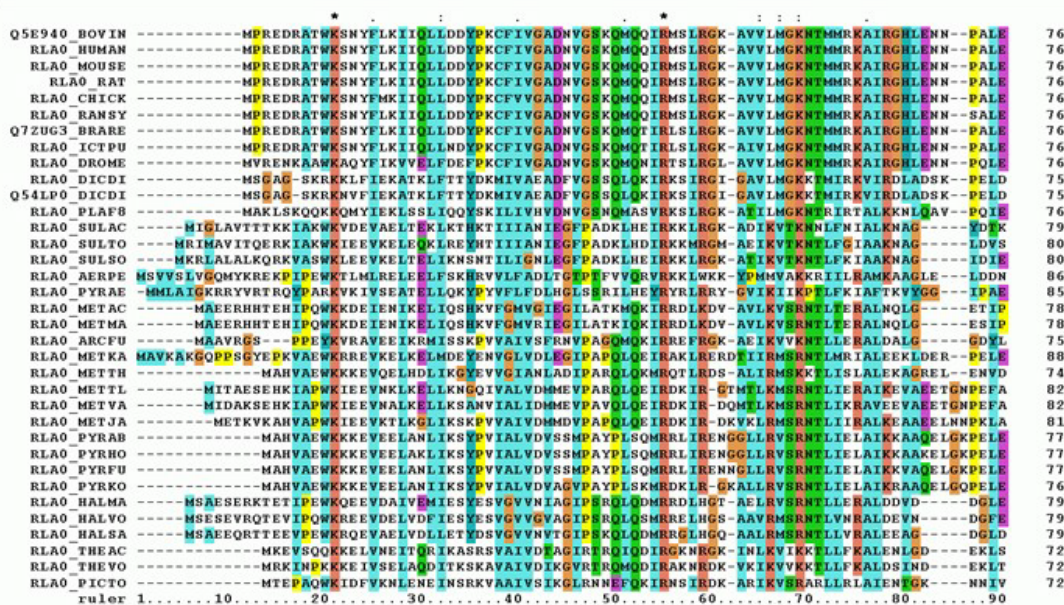
- sinonimnim zamenama (najčešće na 3. ili 1. poziciji u kodonu), što znači da se ne menja aminokiselina (CU- sinonimni kodoni), ili
- nesinonimnim zamenama (2. pozicija u kodonu), što znači da se menja aminokiselina (eng. missense mutation), ili
- pojavljuje stop kodon (eng. nonsense mutation), ili,
- ukoliko je promenjen stop kodon u “sense” kodon – polipeptidni lanac se produžava.

Slika 2.16 prikazuje tipove tačkastih mutacija: tranzicije (zamena purin–purin ili pirimidin–pirimidin) i transverzije (purin–pirimidin).



Slika 2.16: Tačkaste mutacije – tranzicije i transverzije. Izvor: https://en.wikipedia.org/wiki/Point_mutation

Slika 2.17 prikazuje različite efekte tačkastih mutacija na proteinski proizvod: silent (bez promene aminokiseline), missense (promena aminokiseline) i nonsense (nastanak stop kodona).



Slika 2.17: Tipovi tačkastih mutacija – silent, missense, nonsense

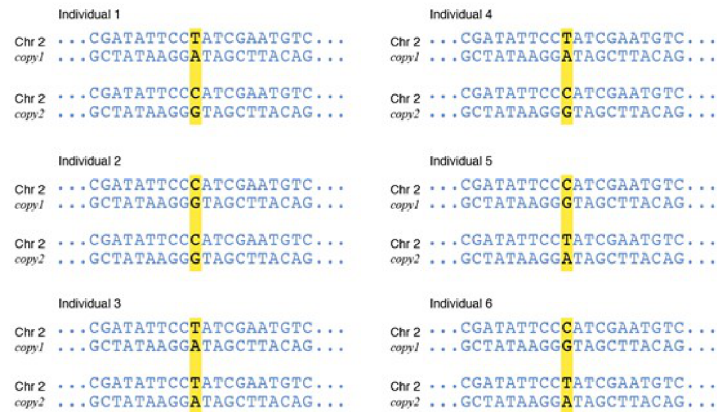
Slika 2.18 prikazuje SNP varijacije između različitih individua, gde se pojedinačni nukleotidi razlikuju na istim pozicijama u DNK sekvenci. Ove varijacije predstavljaju osnovu genetičke raznolikosti i koriste se u populacionim i medicinskim istraživanjima.

Slika 2.19 prikazuje učestalost sinonimnih i nesinonimnih SNP mutacija u zavisnosti od pozicije u kodonu. Najčešće su sinonimne promene na trećoj poziciji kodona, dok promene na drugoj poziciji češće dovode do promene aminokiseline i imaju veći funkcionalni značaj.

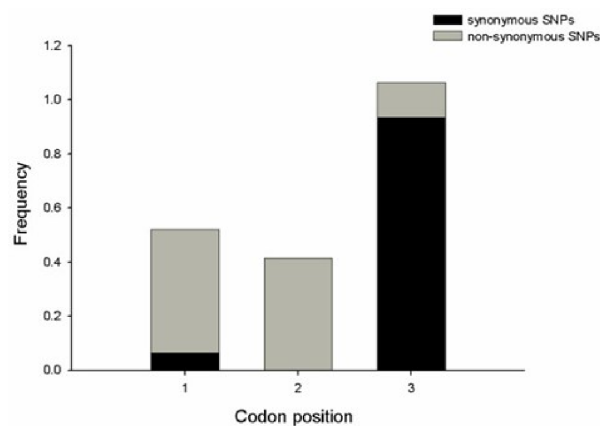
Slika 2.20 prikazuje klasifikaciju SNP mutacija na one koje se nalaze u nekodirajućim i kodirajućim regionima, kao i dalje podele na sinonimne i nesinonimne.

2.3.4 Insercije i delecije (IN/DEL)

Insercije i delecije nukleotida u DNK sekvenci mogu dovesti do promena u dužini sekvence. Ove mutacije mogu imati značajne posledice na strukturu i funkciju proteina.



Slika 2.18: Pojedinačni nukleotidni polimorfizam (SNP)

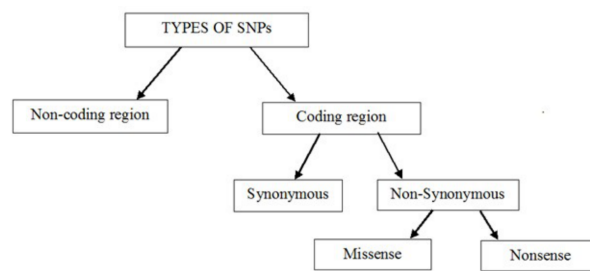
Slika 2.19: Učestalost sinonimnih i nesinonimnih SNP-ova u zavisnosti od pozicije u kodonu. Izvor [BMCGenomics2006, 7:298](#)

Mutacija pomeranja okvira čitanja genetskog koda (eng. frameshift mutation) (slika 2.21) je uzrokovana insercijom ili delecijom jednog ili dva susedna nukleotida (ukoliko su u pitanju tri susedna nukleotida onda se radi o deleciji jedne aminokiseline u polipeptidnom lancu), što rezultira pogrešnim redosledom aminokiselina u polipeptidnom lancu.

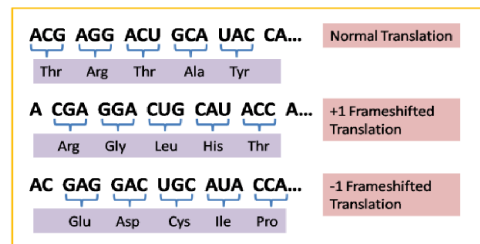
2.3.5 Reparacija (popravka genetskog materijala) DNA

Popravka DNA je kolekcija procesa kojima ćelija identifikuje i koriguje oštećenje DNA molekula koji kodiraju njen genom. Kod ljudskih ćelija, normalne metaboličke aktivnosti i faktori sredine kao što su UV svetlo i radijacija mogu da uzrokuju DNA oštećenje, proizvodeći i do 10^6 individualnih molekulskih oštećenja po ćeliji na dan. Oštećenje na jednom lancu, npr. formiranje pirimidinskog dimera, pod dejstvom UV zračenja se popravljva „ekscizionim reparom“, tako što se oštećeni nukleotidi uklanjaju i zamenjuju neoštećenim koji su komplementarni onima iz neoštećenog lanca.

2.4 Genska ekspresija



Slika 2.20: Tipovi SNPova



Slika 2.21: Mutacija pomeranja okvira čitanja genetskog koda (“frameshift mutation”)

2.4.1 Centralna dogma molekulske biologije

Slika 2.22 prikazuje Fransisa Krika i njegovo prvo formulisanje centralne dogme molekulske biologije. Centralna dogma opisuje tok genetičke informacije od DNK ka RNK i dalje ka proteinima. Ovaj koncept predstavlja osnovu razumevanja genske ekspresije i načina na koji se genetička informacija pretvara u funkcionalne molekule.

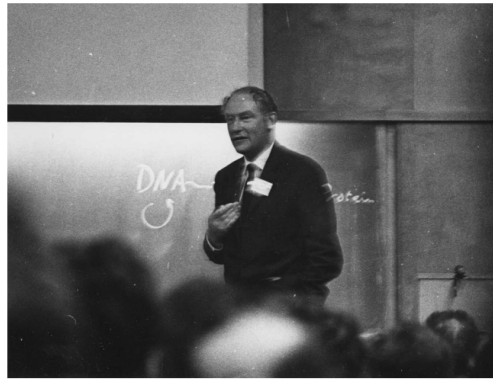
Slika 2.23 prikazuje centralnu dogmu molekulske biologije, odnosno tok informacije DNA → RNA → protein. DNA se **transkribuje** u RNA, a RNA se prevodi u protein tokom **translacije**. Ovaj proces je ključan za ispoljavanje genetičke informacije u ćeliji.

2.4.2 Replikacija

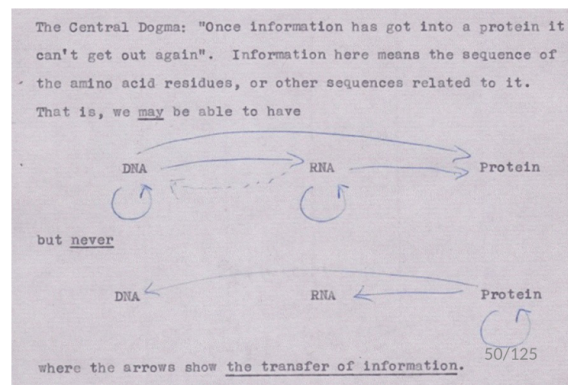
Replikacija (replika = kopija) je složen biološki proces dupliranja molekula DNA pri kome od jednog nastaju dva potpuno identična molekula DNA zahvaljujući principu komplementarnosti vezivanja naspramnih baza u polinukleotidnim lancima. Kada se lanci DNA razmotaju obrazuje se replikaciona viljuška (ima oblik slova Y, slika 2.24). Svaki novonastali molekul DNA sadrži po jedan stari i jedan novi lanac DNA, pa se zbog toga kaže da je replikacija **semikonzervativan** proces (lat. semi = polu; konzervativan = očuvan).

Replikacija se i kod prokariota i kod eukariota vrši *bidirekciono*, što znači da se od mesta gde počinje vrši istovremeno u oba smera i to isključivo u 5'→3' smeru. Pritom, replikacija komplementarnog lanca se sa jedne strane replikacione viljuške vrši kontinuirano (eng. leading strand – vodeći lanac) dok se sa druge strane (eng. lagging strand – prateći lanac, lanac koji zaostaje) sintetišu tzv. Okazaki fragmenti koji se pod dejstvom ligaza spajaju u lanac. Kod prokariota Okazaki fragmenti su dužine 1.000–2.000 nukleotida, dok su kod eukariota preko 10 puta manji.

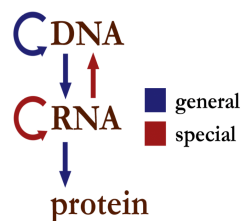
Slika 2.25 prikazuje proces replikacije DNA na nivou replikacione viljuške, gde dolazi do razdvajanja dvostrukog DNA lanca pod dejstvom enzima helikaze. Na mestu razdvajanja formiraju



Crick speaking at the 1963 Cold Spring Harbor Symposium.



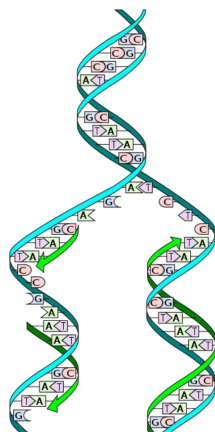
Slika 2.22: Krik govori na simpozijumu u Cold Spring Harboru 1963. godine. Prvi nacrt centralne dogme koji je Krik napravio potiče iz neobjavljene beleške iz 1956. godine. Izvor: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2003243>



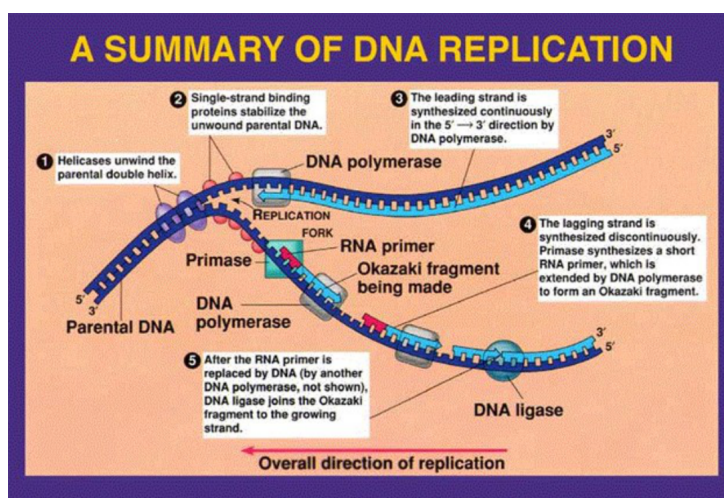
Slika 2.23: Centralna dogma molekulske biologije

se dva jednolančana DNA lanca koji služe kao matrice za sintezu novih lanaca. Sinteza novog DNA lanca odvija se isključivo u smeru $5' \rightarrow 3'$, zbog čega se vodeći lanac sintetiše kontinuirano u pravcu kretanja replikacione viljuške. Nasuprot tome, prateći lanac se sintetiše diskontinuirano u obliku Okazaki fragmenata, koji se kasnije povezuju enzimom DNK ligazom u jedinstven lanac. U procesu učestvuju i drugi enzimi, poput primaze koja sintetiše kratke RNK prajmere i DNK polimeraze koja produžava lanac, čime se obezbeđuje tačno i efikasno umnožavanje genetičkog materijala.

Kod prokariota, čija je DNA prstenasta (kružna), replikacija počinje (eng. origin of replication – oriC) na samo jednom (bakterije) ili više (arheje) mesta i odvija se bidirekciono. Kod bakterija se, npr. mesto (lokus) početka replikacije sastoji od tri funkcionalna elementa: DNA ponovaka, AT regiona i mesta koja prepoznaju regulatorni proteini. Proces počinje u regionu oriC, gde se najpre vezuju regulatorni proteini koji prepoznaju specifične sekvence u tom delu DNA. Njihovo



Slika 2.24: Replikacija

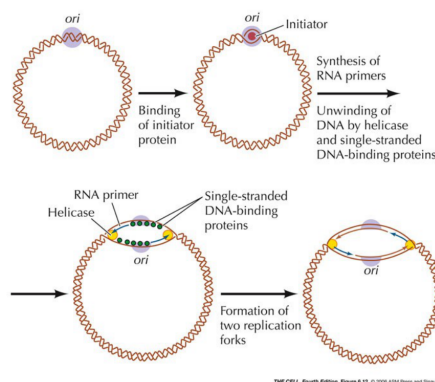


Slika 2.25: DNA replikacija

vezivanje destabilizuje susedni AT region, koji se lakše razdvaja nego GC region jer AT parovi imaju manje vodoničnih veza. Time dolazi do lokalnog otvaranja dvostrukog lanca i formiranja početnog mehura replikacije. Kada se DNA otvori, iz tog mesta polaze dve replikacione viljuške koje se kreću u suprotnim smerovima oko kružnog hromozoma, zbog čega se kaže da je replikacija bidirekciona. Na svakoj viljušci se sintetisuje novi lanci DNA koristeći postojeće lance kao matrice. Kako je hromozom kružan, dve viljuške na kraju obiđu hromozom i sastanu se na suprotnoj strani u odnosu na mesto početka.

Na slici 2.26 je suštinski prikazano da replikacija prokariotskog hromozoma nije nasumična, već strogo organizovana: prvo se prepoznaje oriC, zatim se otvara DNA, pa se formiraju dve viljuške i započinje umnožavanje genetičkog materijala u oba smera. Uloga DNA ponovaka je da obezbede mesta za vezivanje inicijacionih proteina, AT region olakšava početno razdvajanje lanaca, a regulatorni proteini kontrolišu kada i gde će proces početi.

Replikacija DNA u eukariotama, čija je DNA linearna, započinje istovremeno na mnogo mesta duž hromozoma i teče istovremeno. Na taj način replikacija ukupne hromozomske DNA se u eukariotskim ćelijama završava za nekoliko sati, iako je brzina ugrađivanja nukleotida znatno sporija nego kod prokariota. Kod linearnih hromozoma se prilikom svake replikacije gubi po oko



Slika 2.26: Replikacija prokariotskog hromozoma (DNA)

100 – 200 nukleotida, kao posledica orijentacije (3'-5') i diskontinualne sinteze pratećeg lanca. Na krajevima eukariotskih hromozoma se nalaze telomere (télos – 'kraj', méros – 'deo'), kratki (specifični za vrste, kod čoveka npr.: CCCTAA) ponovci nukleotida. Ove nekodirajuće niske štite hromozom od skraćanja prilikom replikacije. Enzim telomeraza dodaje ponavljajuće niske i tako sprečava skraćenje hromozoma.

U tabeli 2.1 su upoređene DNA replikacije kružnih i linearnih hromozoma kod prokariotskih i eukariotskih DNA.

Prokariotska replikacija DNA	Eukariotska replikacija DNA
Odvija se u citoplazmi	Odvija se u jedru
Samo jedno mesto početka replikacije po DNK molekulu	Više mesta početka replikacije u svakom hromozomu
Mesto početka replikacije je dugo oko 100–200 ili više nukleotida	Svako mesto početka replikacije ima oko 150 nukleotida
Replikacija se odvija na jednom mestu po hromozomu	Replikacija se odvija na više mesta istovremeno
Ima samo jedno mesto početka replikacije	Ima više mesta početka replikacije
Inicijaciju vrše proteini DnaA i DnaB	Inicijaciju vrši Origin Recognition Complex
Potreban je topoizomeraza	Potreban je topoizomeraza
Replikacija je veoma brza	Replikacija je veoma spora

Tabela 2.1: Poređenje replikacije DNA kod prokariota i eukariota

2.5 Nekodirajuće i kodirajuće RNA

Ribonukleinska kiselina (RNA) predstavlja ključni molekul u realizaciji genetičke informacije u ćeliji. Iako se dugo smatralo da RNA ima prvenstveno ulogu posrednika između DNA i proteina, savremena istraživanja su pokazala da RNA obavlja širok spektar funkcija u ćeliji. RNA molekuli učestvuju u procesima transkripcije, translacije, regulacije genske ekspresije, kao i u katalitičkim i strukturnim ulogama.

Na osnovu funkcije, RNA se može podeliti na **kodirajuće** i **nekodirajuće** RNA. Kodirajuće RNA su RNVA molekuli koji nose informaciju za sintezu proteina. Najvažniji predstavnik je

mRNA, koja prenosi genetičku informaciju sa DNA do ribozoma, gde se na osnovu nje sintetišu proteini u procesu translacije. Nekodirajuće RNA su RNA molekuli koji se ne prevode u proteine, već imaju različite regulatorne, strukturne i katalitičke funkcije u ćeliji. U ovu grupu spadaju, na primer, tRNA i rRNA (učestvuju u sintezi proteina), kao i brojne regulatorne RNA (miRNA, siRNA, snRNA, lncRNA) koje kontrolišu ekspresiju gena, obradu RNA i druge procese.

Za klasifikaciju nekodirajuće RNA se koristi:

- topologija: nukleusne, citoplazmatske;
- veličina: male (kratke), velike (dugačke);
- funkcija: tRNA, rRNA, RNA splajsovanje, regulacija genske ekspresije, kataliza.

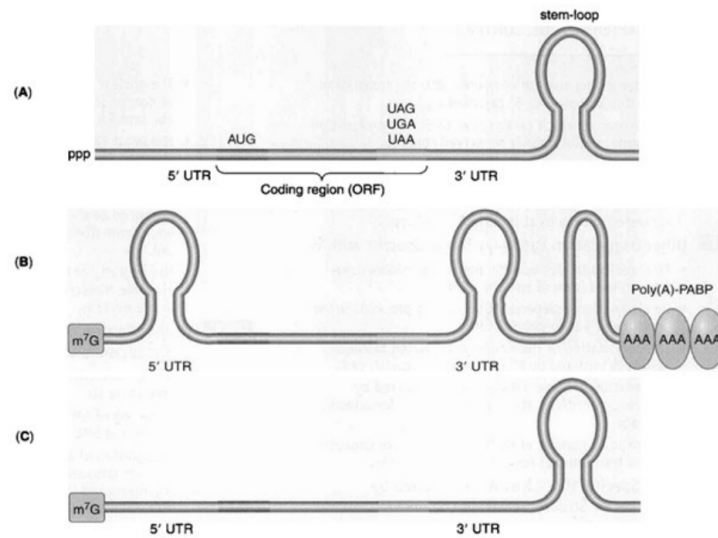
Ekstraćelijske RNA (exRNA) su razne vrste RNA (mRNA, tRNA, miRNA, siRNA, lncRNA) izolovane van ćelija, u obliku proteinskih ili lipoproteinskih kompleksa, ili unutar ekstraćelijskih vezikula. Uloga im nije, još uvek, potpuno razjašnjena, i smatra se da, između ostalog, učestvuju u međućelijskoj komunikaciji, a mnoge od njih se mogu povezati sa raznim patološkim stanjima.

Veće klase nekodirajuće RNA su prikazane u tabeli 2.2

Naziv	Funkcija
Ribozomska RNA (rRNA)	Čini glavni deo ribozoma i učestvuje u sintezi polipeptidnih lanaca
Transportna RNA (tRNA)	Prenosi aminokiseline do ribozoma i prepoznaje kodone na mRNA
Mala nuklearna RNA (snRNA)	Učestvuje u obradi molekula glasničke RNA u jedru eukariotskih ćelija (takode se naziva snRNA ili „snurps“)
Vodička RNA (guide RNA)	Učestvuje u obradi RNA ili DNA u nekim organizmima
Regulatorna RNA	Učestvuje u regulaciji genske ekspresije vezivanjem za proteine ili DNA ili za druge RNA molekule
Antisens RNA	Učestvuje u regulaciji genske ekspresije sparivanjem baza sa mRNA
Prepoznajuća RNA (recognition RNA)	Deo nekoliko enzima (npr. telomerase); omogućava im da prepoznaju određene kratke DNA sekvence
Ribozimi	RNA molekuli sa enzimskom aktivnošću

Tabela 2.2: Glavne klase nekodirajućih RNA i njihove funkcije

Slika 2.27 prikazuje strukturu mRNA kod prokariota i eukariota i njihove glavne razlike. Na delu (A) prikazana je bakterijska mRNA koja nema kapu ni poli(A) rep, ali može sadržati sekundarne strukture (eng. stem-loop) koje učestvuju u stabilnosti i regulaciji. Takođe je označen ORF (kodirajući region) između start (AUG) i stop kodona. Na delu (B) prikazana je eukariotska mRNA koja ima karakterističnu 5' kapu (m^7G) i 3' poli(A) rep, koji povećavaju stabilnost mRNA i omogućavaju efikasnu translaciju. Prisustvo proteina vezanih za poli(A) rep dodatno stabilizuje molekul. Na delu (C) prikazan je poseban slučaj – histonska mRNA kod eukariota, koja nema poli(A) rep, već završava specifičnom stem-loop strukturom, što predstavlja alternativni mehanizam regulacije.



Slika 2.27: Osobine prokariotskih i eukariotskih mRNA.

2.6 Regulacija genske ekspresije

2.6.1 Transkripcija i translacija

Genetička informacija u ćeliji se prenosi po principu:



Ovaj tok predstavlja osnovu ekspresije gena i omogućava realizaciju genetičkog materijala u funkcionalne molekule.

Transkripcija

Transkripcija je proces u kome se informacija sa DNA prepisuje u RNA. Ovaj proces katalizuje enzim RNA polimeraza, koji koristi jedan lanac DNA kao matricu za sintezu komplementarne RNA. Katalizovati znači ubrzavati ili omogućavati hemijsku reakciju bez sopstvenog trošenja u tom procesu. Enzim je biološki molekul (najčešće protein) koji katalizuje hemijske reakcije u ćeliji, odnosno ubrzava njihovo odvijanje bez sopstvenog trošenja.

Proces transkripcije obuhvata tri osnovne faze:

- **Inicijacija** – RNA polimeraza se vezuje za promotorski region DNA (deo DNA koji se nalazi na početku gena) i započinje razdvajanje dvostrukog lanca.
- **Elongacija** – enzim se kreće duž DNA i ugrađuje ribonukleotide po principu komplementarnosti (A-U, T-A, G-C, C-G).
- **Terminacija** – sinteza RNA se završava kada enzim naiđe na terminacionu sekvencu.

Rezultat transkripcije je RNA molekul, najčešće mRNA, koji sadrži informaciju za sintezu proteina.

Obrada RNA kod eukariota

Kod eukariota, primarni RNA transkript (pre-mRNA) prolazi kroz nekoliko faza obrade pre nego što postane funkcionalna mRNA:

- **5' kapiranje** – dodavanje modifikovanog nukleotida (m^7G kapa) na 5' kraj RNA, što omogućava stabilnost i prepoznavanje od strane ribozoma.
- **RNA splajsovanje** – uklanjanje introna (nekodirajućih sekvenci) i spajanje egzona (kodirajućih sekvenci).
- **Poliadenilacija** – dodavanje poli(A) repa na 3' kraj RNA, čime se povećava stabilnost i efikasnost translacije.

Nakon obrade, zrela mRNA se transportuje iz jedra u citoplazmu.

Translacija

Translacija je proces u kome se informacija sa mRNA prevodi u redosled aminokiselina, čime nastaje protein. Ovaj proces se odvija na ribozomima u citoplazmi.

Translacija se odvija u tri faze, kao i transkripcija:

- **Inicijacija** – ribozom se vezuje za mRNA i prepoznaje start kodon (najčešće AUG).
- **Elongacija** – tRNA donosi aminokiseline koje se povezuju peptidnim vezama, formirajući polipeptidni lanac.
- **Terminacija** – proces se završava dolaskom do stop kodona (UAA, UAG, UGA), nakon čega se protein oslobađa.

2.6.2 Ekspresija gena kod eukariota i prokariota

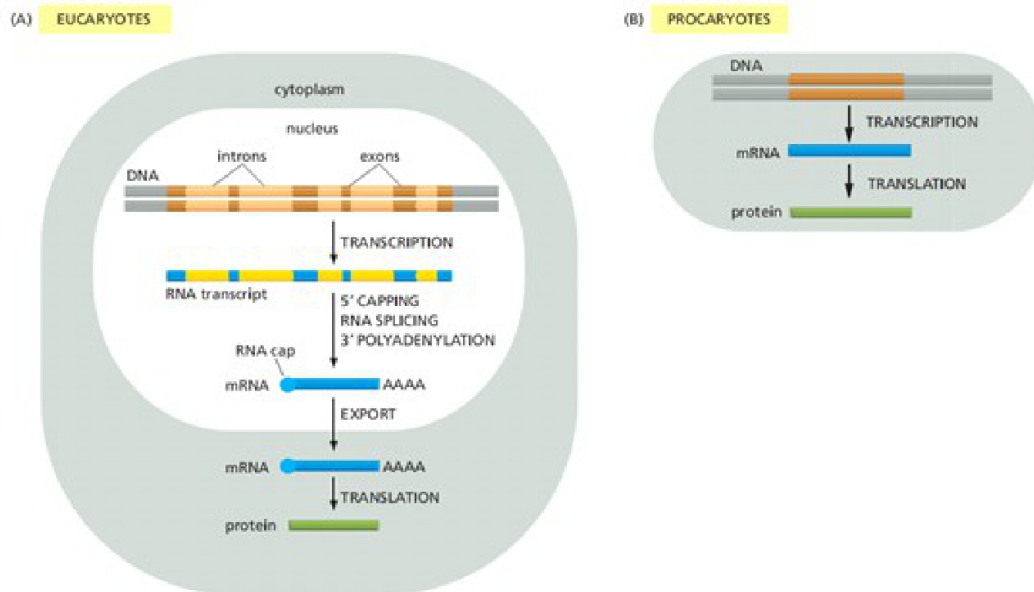
Kod eukariota, procesi transkripcije i translacije su prostorno i vremenski odvojeni. Transkripcija se odvija u jedru, nakon čega sledi obrada RNA i transport mRNA u citoplazmu gde se odvija translacija. Ova složenost omogućava višestruke nivoe regulacije genske ekspresije.

Kod prokariota, ne postoji jedro, pa se transkripcija i translacija odvijaju u istoj ćelijskoj sredini. mRNA kod prokariota ne prolazi kroz složenu obradu kao kod eukariota, što omogućava bržu, ali manje regulisanu ekspresiju gena.

Regulacija genske ekspresije je proces kojim ćelija kontroliše ispoljavanje gena kako bi prilagodila sintezu proteina svojim potrebama.

Slika 2.28 prikazuje osnovne razlike u ekspresiji gena između eukariota i prokariota. Kod eukariota, transkripcija se odvija u jedru, gde nastaje primarni RNA transkript koji sadrži introne i eksone, nakon čega sledi obrada RNA (5' kapiranje, splajsovanje i poliadenilacija). Zrela mRNA se zatim transportuje u citoplazmu gde se odvija translacija u protein. Nasuprot tome, kod prokariota nema jedra, pa se transkripcija i translacija odvijaju istovremeno u citoplazmi, bez dodatne obrade RNA. Ova razlika omogućava složeniju regulaciju genske ekspresije kod eukariota u poređenju sa prokariotima.

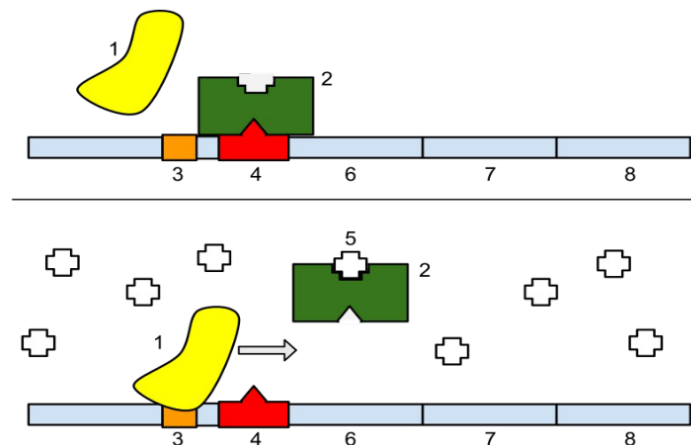
Operon je funkcionalna jedinica regulacije genetičke aktivnosti koja se sastoji iz promotora, operatora i jednog ili više inducibilnih strukturnih gena koji se nalaze pod kontrolom regulatornih



Slika 2.28: Poređenje ekspresije gena kod eukariota i prokariota

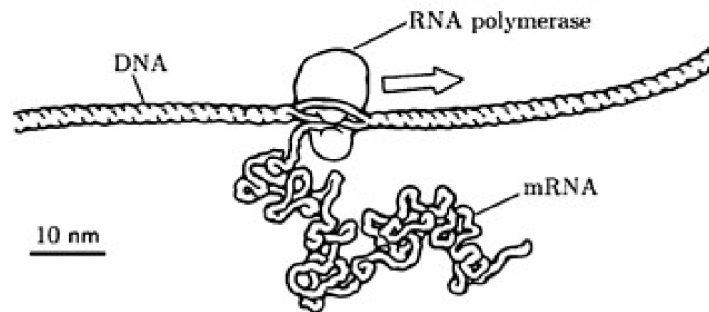
elemenata. Prisustvo operona je pokazano kod prokariota i kod nekih eukaria (nematoda). Svi geni nisu organizovani u operone, konstitutivni geni (geni koji se stalno izražavaju (uvek aktivni) u ćeliji, nezavisno od spoljašnjih uslova) ne podležu represiji aktivnosti, (mada su svi geni u izvesnoj meri podložni regulaciji aktivnosti).

Na slici 2.29 je prikazana regulacija laktoznog operona (negativna kontrola, indukcija). Na slici su prikazani: (1) RNA polimeraza, (2) represor, (3) promoter, (4) operator, (5) laktoza (induktor), (6–8) kodirajuće sekvence lacZ, lacY, lacA. U gornjem delu slike: lac geni su isključeni, nema laktoze (5) koja inhibira (isključuje) represor (2) koji blokira polimerazu (1) da se veže za promotersku sekvencu (3) i započne transkripciju mRNA (6–8). U donjem delu slike: geni su otključani jer laktoza (5) inhibira represor (2), te je RNA polimeraza (1) vezana za promoter (3) i počinje transkripcija (6–8).



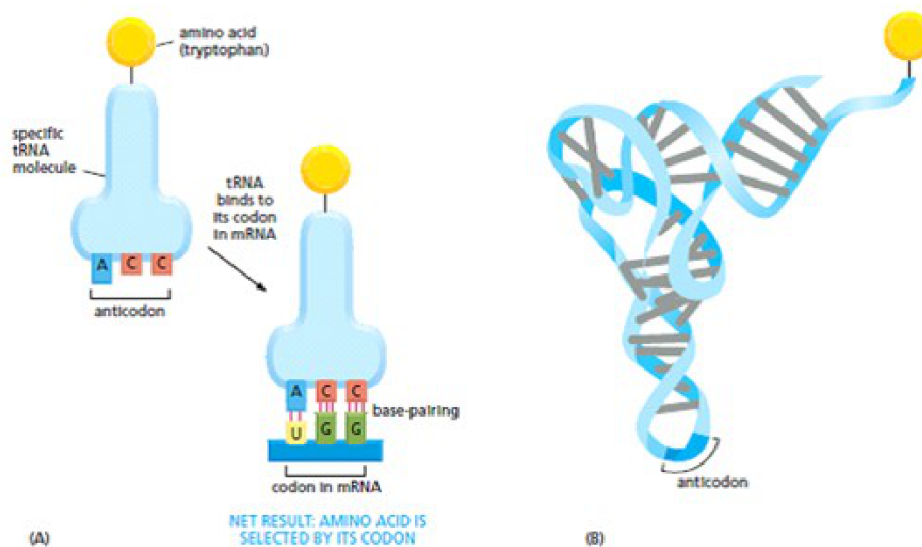
Slika 2.29: Regulacija laktoznog operona (negativna kontrola, indukcija)

Slika 2.30 prikazuje proces transkripcije u kome enzim RNA polimeraza klizi duž DNA molekula i istovremeno sintetiše mRNA. DNA služi kao matrica, dok RNA polimeraza „čita“ njen niz nukleotida i ugrađuje komplementarne ribonukleotide, čime nastaje lanac mRNA. Prikazano je kako se novonastala mRNA odvaja od DNA i formira zavijenu strukturu iza enzima. Ovaj proces je kontinuiran i odvija se u smeru kretanja RNA polimeraze duž DNA.



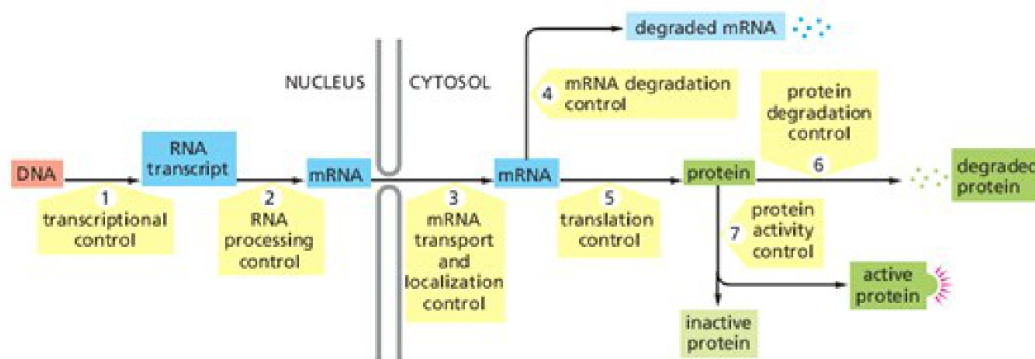
Slika 2.30: Transkripcija DNA u glasničku RNA

Slika 2.31 prikazuje ulogu transportne RNA (tRNA) u procesu translacije i način na koji se određuje redosled aminokiselina u proteinu. Na delu (A) prikazano je kako specifična tRNA nosi određenu aminokiselinu (npr. triptofan) i poseduje antikodon koji je komplementaran kodonu na mRNA. Sparivanjem antikodona tRNA i kodona mRNA obezbeđuje se da se ugradi odgovarajuća aminokiselina, čime se genetička informacija pravilno prevodi u protein. Na delu (B) prikazana je 3D struktura tRNA molekula, koja ima karakterističan oblik savijene „L“ strukture. Ova prostorna organizacija omogućava funkcionalno razdvajanje dva ključna dela molekula: na jednom kraju (3' kraj) nalazi se mesto za vezivanje aminokiseline, dok se na drugom kraju nalazi antikodon koji prepoznaje kodon na mRNA. Ovakav raspored omogućava tRNA da istovremeno komunicira sa ribozomom, mRNA i odgovarajućom aminokiselinom, čime igra ključnu ulogu u preciznosti i efikasnosti translacije.



Slika 2.31: Uloga tRNA u translaciji i prepoznavanje kodona putem antikodona

2.6.3 Sazrevanje (procesovanje) eukariotske mRNA



Slika 2.32: Nivoi regulacije genske ekspresije kod eukariota

Slika 2.32 prikazuje različite nivoe na kojima se vrši regulacija genske ekspresije kod eukariotskih ćelija, od DNA do aktivnog proteina.

(1) **Transkripcioni nivo:** Kontrola započinje na nivou transkripcije, gde se određuje da li će se DNA prepisati u RNA. Ovo je najvažniji nivo regulacije za većinu gena.

(2) **Procesiranje RNA:** Primarni RNA transkript prolazi kroz obradu (splajsovanje, 5' kapiranje i poliadenilacija), čime nastaje zrela mRNA.

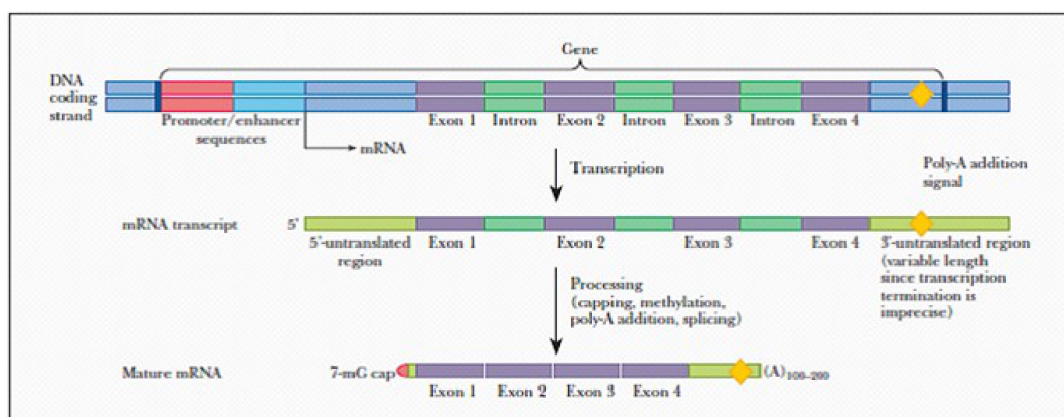
(3) **Transport i lokalizacija mRNA:** mRNA se transportuje iz jedra u citoplazmu i može se usmeriti ka određenim delovima ćelije.

(4) **Stabilnost i razgradnja mRNA:** Količina mRNA u ćeliji reguliše se njenom stabilnošću i brzinom razgradnje.

(5) **Translacioni nivo:** Kontroluje se efikasnost prevođenja mRNA u protein na ribozomima.

(6) **Posttranslacioni nivo:** Nakon sinteze, protein može biti aktiviran, modifikovan ili razgrađen, čime se dodatno reguliše njegova funkcija.

Ovi nivoi regulacije omogućavaju ćeliji preciznu kontrolu nad količinom i aktivnošću proteina, čime se obezbeđuje pravilno funkcionisanje i prilagođavanje uslovima sredine.



Slika 2.33: Organizacija gena kod eukariota i obrada mRNA

Slika 2.33 prikazuje organizaciju gena kod eukariota i proces nastanka zrele mRNA. Na vrhu je prikazana DNA sekvenca koja sadrži regulatorne regione (promoter/enhancer) i naizmenične

delove egzona i introna. Tokom transkripcije nastaje primarni mRNA transkript (pre-mRNA) koji sadrži i kodirajuće i nekodirajuće sekvence, uključujući 5' i 3' UTR regione.

U procesu obrade RNA dolazi do 5' kapiranja, splajsovanja i poliadenilacije. Splajsovanjem se uklanjaju introni i spajaju egzoni u kontinuiranu sekvencu, dok se na 5' kraj dodaje kapa, a na 3' kraj poli(A) rep, što povećava stabilnost i omogućava translaciju.

Kao rezultat nastaje zrela mRNA koja sadrži samo egzonske sekvence i spremna je za transport u citoplazmu i sintezu proteina. Ovaj proces omogućava preciznu obradu genetičke informacije pre njenog prevođenja u funkcionalni protein.

2.7 Genetski kod

Key Concept

- **Kod** je uvek skup pravila koja uspostavljaju korespondenciju (ili preslikavanje) između dva nezavisna sveta [3].
- *Kod* je mali skup proizvoljnih pravila izabran iz potencijalno neograničenog broja, kako bi se obezbedila specifična korespondencija između dva nezavisna sveta. [4]

Genetski kod predstavlja skup pravila po kojima se informacija zapisana u nukleotidnom nizu DNA ili mRNA prevodi u redosled aminokiselina u proteinu. Ova pravila definišu kako se niz od tri nukleotida, nazvan kodon, povezuje sa određenom aminokiselinom ili signalom za početak i završetak sinteze proteina.

Genetski kod je tripletan (slika 2.34), što znači da se čita u grupama od po tri nukleotida, pri čemu svaki kodon određuje jednu aminokiselinu. Takođe je redundantan, jer više različitih kodona može kodirati istu aminokiselinu, ali je istovremeno i jednoznačan, jer jedan kodon uvek određuje samo jednu aminokiselinu.

Kod je gotovo univerzalan kod svih organizama, što ukazuje na zajedničko evolutivno poreklo živih bića. Postoje i specifični kodoni: start kodon (AUG) koji označava početak translacije i kodira metionin, kao i stop kodoni (UAA, UAG, UGA) koji označavaju kraj sinteze proteina.

2.8 Hromozomi

2.8.1 Organizacija prokariotskih hromozoma

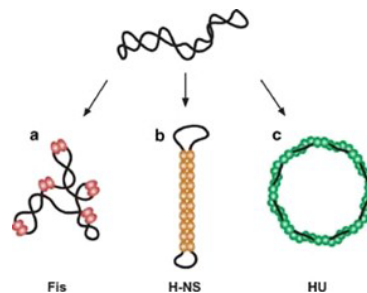
Genetički materijal prokariotskih organizama nije ograničen od ostatka ćelije dvostrukom membranom, kao kod eukariotskih ćelija i označen je terminom **nukleoid**. Sastoji se od jednog ili više hromozoma (DNA molekula), koji mogu biti linearni ili, češće, cirkularni i zastupljeni u jednoj ili više kopija (ploidia).

Hromozomi su složeni kompleksi koji se sastoje od DNA i pridruženih proteina (u slučaju eukariotskih organizama - histona i nehistskih proteina, a kod prokariotskih organizama - histonima sličnih proteina). Jedan molekul DNA čini jedan hromozom.

		Second base				
		U	C	A	G	
First base	U	UUU } Phenylalanine F UUC } UUA } Leucine L UUG }	UCU } Serine S UCC } UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	U C A G
	C	CUU } Leucine L CUC } CUA } CUG }	CCU } Proline P CCC } CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } Arginine R CGC } CGA } CGG }	U C A G
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } Threonine T ACC } ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }	U C A G
	G	GUU } Valine V GUC } GUA } GUG }	GCU } Alanine A GCC } GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } Glycine G GGC } GGA } GGG }	U C A G

Slika 2.34: Kodoni

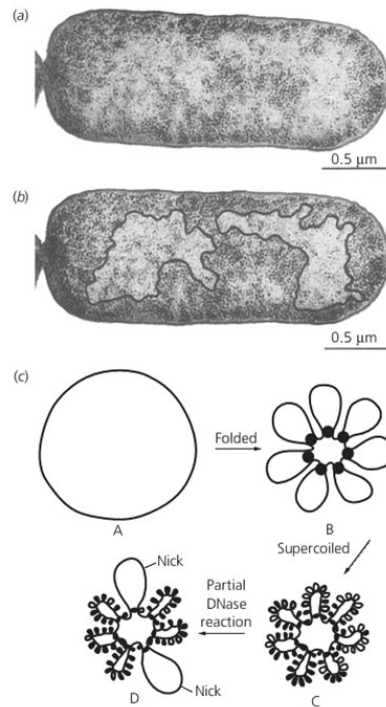
Slika 2.35 prikazuje načine organizacije i pakovanja DNA u prokariotskom hromozomu (nukleoidu), gde ključnu ulogu imaju histonima slični proteini. Protein Fis (Factor for inversion stimulation) omogućava formiranje petlji u DNA, čime doprinosi njenoj lokalnoj reorganizaciji i savijanju. H-NS (Histone-like Nucleoid Structuring protein) povezuje udaljene segmente DNA (tzv. „bridging“) i često učestvuje u utišavanju gena, naročito onih koji su uneti horizontalnim transferom. HU (histone-like) protein ima ulogu u savijanju i stabilizaciji DNA, omogućavajući njeno kompaktno pakovanje u kružnu strukturu nukleoida. Iako prokarioti nemaju prave histone kao eukarioti, ovi proteini obezbeđuju efikasnu organizaciju genoma, njegovu stabilnost i kontrolu genske ekspresije.



Slika 2.35: Fis, H-NS i HU nukleoid asocirani protein

Slika 2.36 prikazuje organizaciju nukleoida u bakterijskoj ćeliji i različite nivoe pakovanja DNA. Na delu (a) vidi se elektronski mikrograf tankog preseka ćelije *E. coli*, gde se uočava homogena unutrašnjost bez jasno izdvojenog jedra, što je karakteristično za prokariote. Na delu (b) prikazana je ista ćelija, ali su sada jasno označene granice nukleoida – regiona u kome se nalazi bakterijski hromozom, koji nije odvojen membranom, ali je funkcionalno organizovan. Donji deo slike (c) prikazuje model organizacije bakterijskog hromozoma kroz više nivoa strukture. U poddelu A prikazana je potpuno razmotana (nekompaktna) DNA, koja bi u tom obliku zauzimala znatno veći prostor u ćeliji. U poddelu B DNA je organizovana u petlje (domene) zahvaljujući

interakcijama sa proteinima, čime se postiže prvi nivo sabijanja. U poddelu C prikazano je dodatno uvijanje DNA u obliku superzavijanja (eng. supercoiling), što omogućava još veću kompaktnost i stabilnost hromozoma. U poddelu D prikazana je situacija kada dolazi do prekida (eng. nick) u jednom lancu DNA unutar domena, što dovodi do relaksacije superzavijanja i gubitka prethodno postignute kompaktnosti. Ovaj model ilustruje kako kombinacija organizacije u domene i superzavijanja omogućava efikasno pakovanje i funkcionalnu organizaciju bakterijskog genoma.



Slika 2.36: Organizacija prokariotskih hromozoma

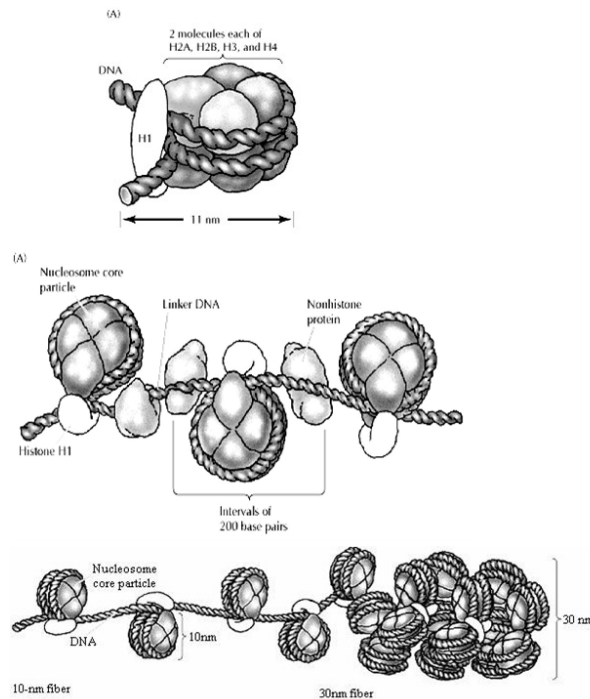
2.8.2 Organizacija eukariotskog hromozoma

Eukariotske ćelije imaju složenu unutarćelijsku građu, koja se sastoji od odeljaka (subcelularnih komponenti). Jedro (nucleus) je odvojeno od ostatka ćelije dvostrukom membranom. U interfaznoj ćeliji (koja nije u procesu deobe), hromozomi su gusto spakovani u jedru i organizovani u hromozomske teritorije. U procesu ćelijske deobe, hromozomi se razdvajaju i postaju lako uočljivi.

Organizacija genetskog materijala u eukariotskoj ćeliji odvija se hijerarhijski: linearna DNA se obmotava oko histonskih proteina formirajući **nukleozome**, nizovi nukleozoma grade **hromatin**, koji se daljim sabijanjem i organizacijom kondenzuje u **hromozome**.

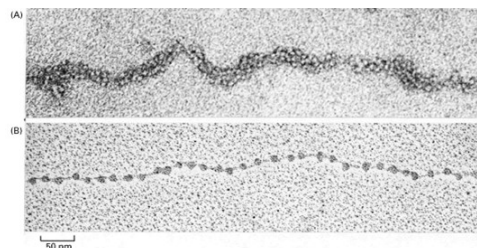
Slika 2.37 prikazuje organizaciju eukariotskog hromatina na nivou nukleozoma. Nukleozom predstavlja osnovnu strukturnu jedinicu i sastoji se od histonskog oktamera (po dve kopije proteina H2A, H2B, H3 i H4) oko kojeg je obmotana DNA. Histon H1 učestvuje u stabilizaciji strukture vezujući DNA na ulazu i izlazu iz nukleozoma. Na višem nivou organizacije, nukleozomi su povezani segmentima tzv. linker DNA, formirajući strukturu koja podseća na „perle na koncu“. Ova organizacija omogućava efikasno pakovanje DNA u jedru, ali i regulaciju genske ekspresije

putem interakcija sa histonskim i nehistonim proteinima.



Slika 2.37: Organizacija eukariotskog hromozoma - građa nukleozoma

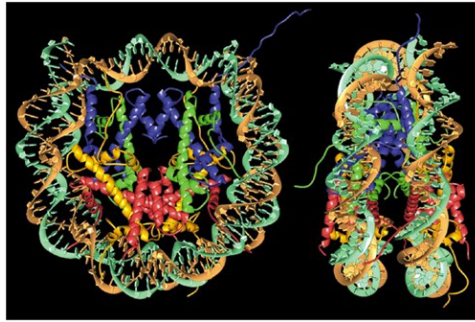
Slika 2.38 prikazuje hromatin u različitim stepenima organizacije. Na slici (B) jasno se uočava struktura „perli na koncu“, gde su pojedinačni nukleozomi povezani linker DNA, što odgovara manje sabijenom, transkripciono aktivnom hromatinu. Na slici (A) prikazana je kompaktnija struktura u kojoj nukleozomi nisu jasno razdvojeni, što odgovara višem stepenu sabijanja hromatina i smanjenoj dostupnosti DNA za gensku ekspresiju.



Slika 2.38: Organizacija eukariotskog hromozoma

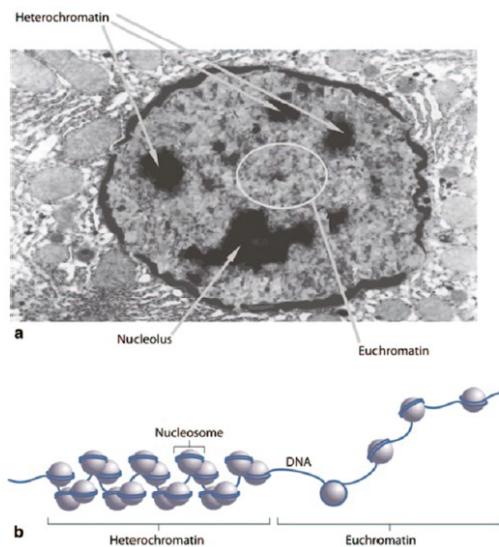
Slika 2.39 prikazuje trodimenzionalnu strukturu nukleozoma. DNA je obmotana oko histonskog oktamera. Prikaz odozgo ilustruje centralni raspored histona i kružno obmotavanje DNA, dok bočni prikaz pokazuje cilindričnu organizaciju nukleozoma i putanju DNA pri ulasku i izlasku iz strukture.

Interfazni hromozomi predstavljaju hromozome u ćeliji tokom interfaze (period između deoba), kada su u dekoncentrovanom stanju i organizovani u obliku hromatina. U ovom obliku nisu vidljivi kao zasebne strukture, već kao difuzna mreža u jedru, što omogućava aktivnu gensku ekspresiju i replikaciju DNA. Stepenu sabijenosti varira između **euhromatina**, koji je transkripciono aktivan, i **heterohromatina**, koji je kondenzovan i uglavnom neaktivan.



Slika 2.39: 3D organizacija eukariotskog hromozoma

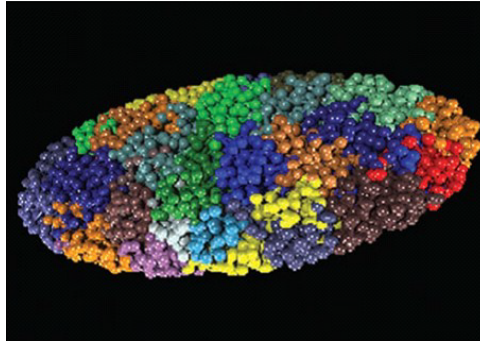
Slika 2.40 prikazuje razliku između euhromatina i heterohromatina u jedru eukariotske ćelije. Na elektronskom mikrografu tamnije oblasti, uglavnom raspoređene uz periferiju jedra, predstavljaju heterohromatin, koji je visoko kondenzovan i transkripciono neaktivan. Svetlije oblasti predstavljaju euhromatin, koji je manje sabijen i u kome se odvija aktivna genska ekspresija. U jedru se takođe uočava jedarce, kao mesto sinteze rRNA. Šematski prikaz pokazuje da je heterohromatin karakterisan gustim pakovanjem nukleozoma, dok euhromatin ima opušteniju strukturu sa razmaknutim nukleozomima, što omogućava pristup transkripcionim faktorima i aktivnost gena.



Slika 2.40: Interfazni hromozomi

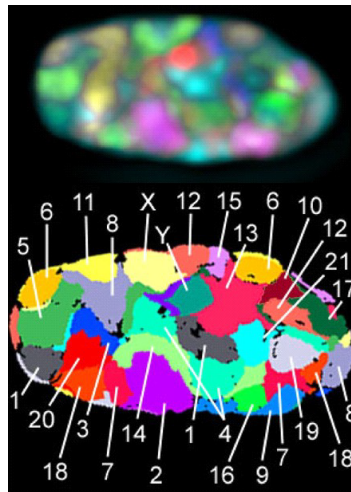
Slika 2.41 prikazuje trodimenzionalnu organizaciju hromozoma u jedru humane ćelije tokom interfaze. Svaki hromozom zauzima sopstvenu prostornu oblast, poznatu kao hromozomska teritorija ili mapa, što ukazuje na to da raspored genetskog materijala u jedru nije nasumičan. Hromozomi su organizovani tako da omogućavaju efikasnu regulaciju genske ekspresije i međusobne interakcije. Ovakva prostorna organizacija ima važnu ulogu u funkciji genoma i kontroli ćelijskih procesa.

Slika 2.42 prikazuje raspored hromozoma u jedru ćelije tokom interfaze, gde je svaki hromozom obeležen različitom bojom. Gornji deo predstavlja eksperimentalni fluorescentni snimak, dok donji deo prikazuje analizirani model sa jasno izdvojenim hromozomima označenim brojevima.



Slika 2.41: Mapa hromozoma

Uočava se da svaki hromozom zauzima specifičnu prostornu oblast unutar jedra, poznatu kao hromozomska mapa.

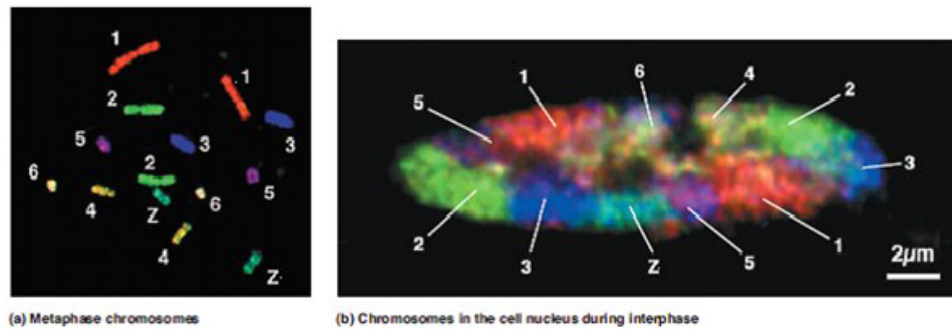


Slika 2.42: Mapa hromozoma

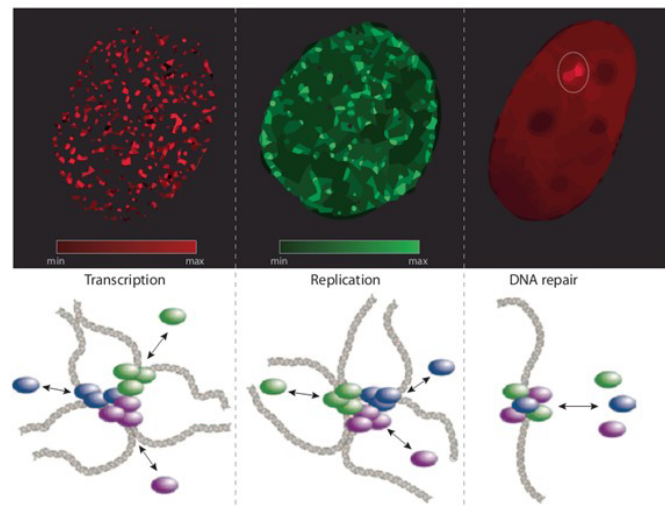
Slika 2.43 prikazuje organizaciju hromozoma kokoške u ćelijskom jedru kroz poređenje metafaze i interfaze. Metafaza je jedna od faza deobe ćelije u kojoj se hromozomi poređaju u sredini ćelije. U metafazi su hromozomi jasno razdvojeni, što omogućava njihovu identifikaciju pomoću specifičnog bojenja. U interfazi, isti hromozomi su dekoncentrovani i raspoređeni unutar jedra, pri čemu svaki zauzima sopstvenu prostornu oblast, poznatu kao hromozomska teritorija. Napomena: ćelije kokoške su diploidne, sa po dve kopije svakog hromozoma. Hromozom Z je polni hromozom.

Slika 2.44 prikazuje prostorna organizacija osnovnih nuklearnih procesa, transkripcije, replikacije i popravke DNA, u jedru eukariotske ćelije. Ovi procesi se ne odvijaju ravnomerno u celom jedru, već su organizovani u specifične regione gde je koncentracija aktivnosti visoka. Transkripcija i replikacija odvijaju se u žarištima, dok se popravka DNA lokalizuje na mestima oštećenja. Na molekularnom nivou, proteinski kompleksi se dinamički okupljaju na DNA i omogućavaju izvršavanje ovih procesa, što ukazuje na visoku prostornu i funkcionalnu organizaciju genoma u jedru.

Slika 2.45 prikazuje hijerarhijsku organizaciju DNA u eukariotskoj ćeliji, počev od dvostruke spirale prečnika oko 2 nm, koja se obmotavanjem oko histonskih proteina organizuje u nukleosome



Slika 2.43: Hromozomske teritorije u ćelijskom jedru.



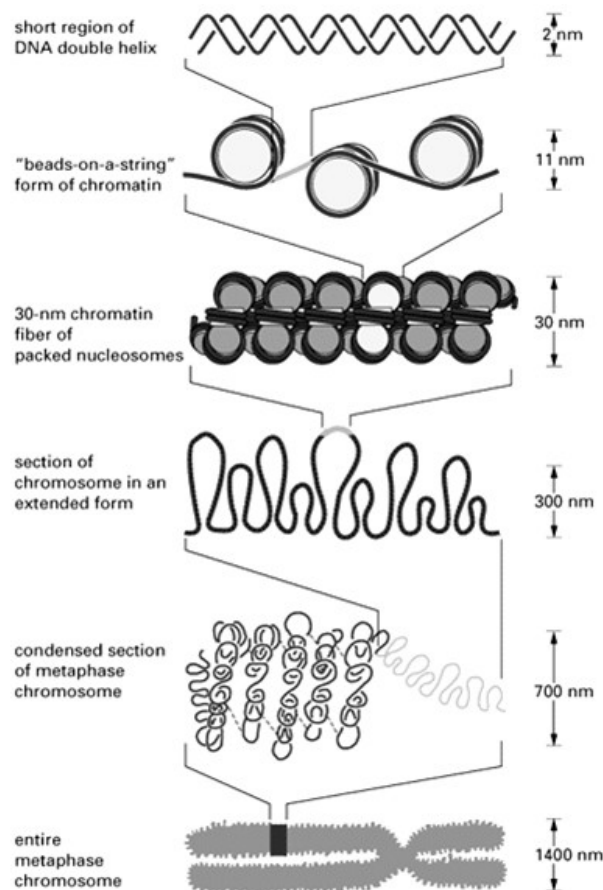
Slika 2.44: Prikaz ključnih procesa u jedru: transkripcije, replikacije i popravke DNA

i formira „perle na koncu“ strukturu debljine oko 11 nm. Daljim sabijanjem nastaje 30 nm vlakno, koje se zatim organizuje u veće petlje vezane za proteinski skelet. Ove petlje se dodatno kondenzuju u deblje strukture koje na kraju formiraju potpuno kondenzovan metafazni hromozom, prečnika oko 1400 nm. Ovakva višestepena organizacija omogućava efikasno pakovanje DNA i regulaciju genske ekspresije.

Slika 2.46 prikazuje metafazni hromozom čoveka snimljen skenirajućim elektronskim mikroskopom. Hromozom je u maksimalno kondenzovanom stanju i sastoji se od dve sestrinske hromatide povezane u oblasti centromere (braon zona). Krajnji crveno obojeni delovi hromozoma, telomere, imaju zaštitnu ulogu i sprečavaju degradaciju genetskog materijala. Površina hromozoma (zelenkasta struktura) odražava visoko sabijen hromatin, što omogućava efikasno pakovanje DNA i njenu pravilnu raspodelu tokom ćelijske deobe.

Slika 2.47 prikazuje mitotički hromozom sa jasno izraženom centromerom, koja se uočava kao suženje između dve sestrinske hromatide. Centromera sadrži visoko ponavljajuće sekvence DNA i predstavlja mesto formiranja *kinetohora*, proteinske strukture koja omogućava vezivanje mikrotubula deobnog vretena. Ova interakcija je ključna za pravilno razdvajanje hromozoma tokom mitoze i mejoze.

Kariotip predstavlja sistematski prikaz hromozoma jedne ćelije koji se klasifikuju prema dužini, morfologiji i obrascu traka. Hromozomi se najpre poređaju po veličini, od najdužih ka

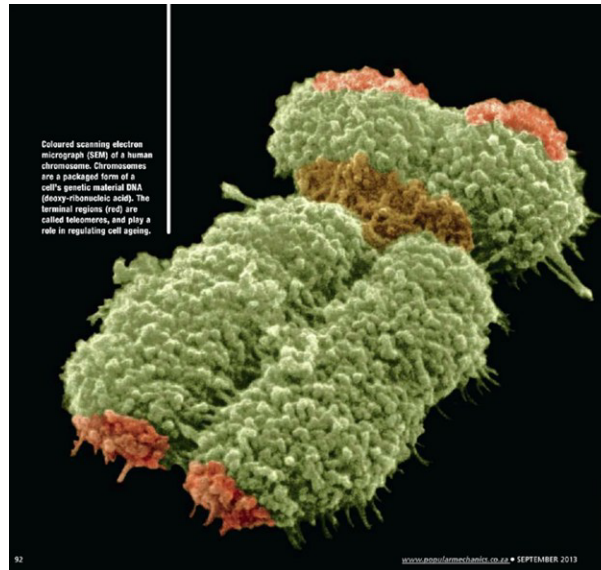


Slika 2.45: Hijerarhijsku organizacija DNA u eukariotskoj ćeliji

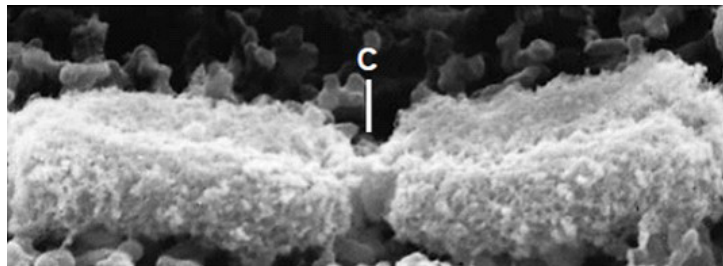
najkraćim, a zatim se razlikuju prema položaju centromere na metacentrične (centromera je u sredini, kraci su približno jednake dužine), submetacentrične (centromera je pomaknuta od sredine) i akrocentrične (centromera je blizu kraja). Dodatna identifikacija postiže se analizom karakterističnih obrazaca traka dobijenih različitim metodama bojenja, kao što su G, Q, R i C trake. Ove trake predstavljaju različite metode bojenja hromozoma koje omogućavaju vizualizaciju specifičnih obrazaca duž hromozoma i koriste se za njihovu identifikaciju i analizu strukturnih promena.

Slika 2.48 prikazuje kariotip čoveka, odnosno organizovan prikaz svih hromozoma raspoređenih u 23 para. Prvih 22 para čine autosomi, dok poslednji par predstavljaju polni hromozomi, koji u ovom slučaju odgovaraju muškom polu (XY). Hromozomi su poredani prema veličini i morfologiji, pri čemu svaki par čine homologni hromozomi poreklom od oba roditelja. Analiza kariotipa omogućava da se utvrdi da li je broj hromozoma normalan i da li je njihova građa očuvana. Kod čoveka je uobičajen kariotip 46 hromozoma, odnosno 23 para. Kada se pregledom kariotipa ustanovi odstupanje u broju, govori se o *numeričkim aberacijama*, a kada je promenjena struktura nekog hromozoma, govori se o *strukturnim aberacijama*.

Numeričke aberacije nastaju kada postoji višak ili manjak celih hromozoma ili čitavih setova hromozoma. Najpoznatiji primer je trisomija 21, kod koje osoba ima tri kopije hromozoma 21 umesto dve, što dovodi do Daunovog sindroma. Mogu se javiti i monosomije, kada nedostaje jedan hromozom iz para, na primer monosomija X kod Turnerovog sindroma. Dakle, analizom



Slika 2.46: Hromozom čoveka snimljen skenirajućim elektronskim mikroskopom



Slika 2.47: Hromozom - centromera

kariotipa može se direktno videti da li neki hromozom postoji u dve, tri ili samo jednoj kopiji.

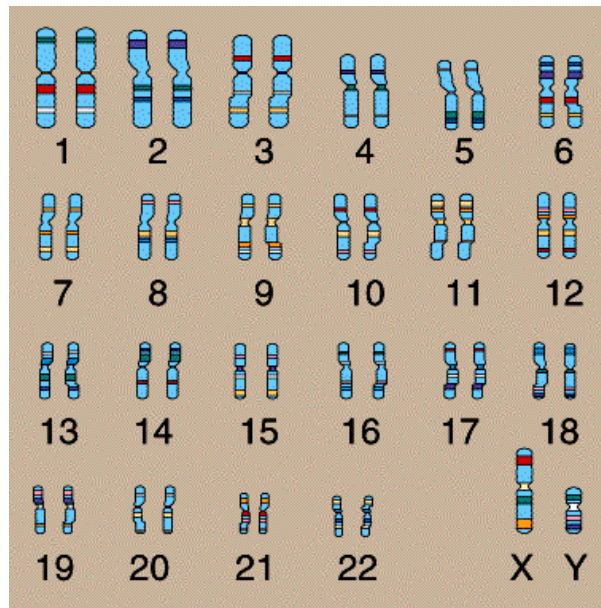
Strukturne aberacije znače da broj hromozoma može biti normalan, ali je promenjena unutrašnja građa nekog od njih. Na kariotipu se takve promene uočavaju poređenjem dužine hromozoma, položaja centromere i obrasca pruga nakon bojenja. Ako neki hromozom izgleda kraće, ima neobičan oblik, ili deo koji bi trebalo da pripada jednom hromozomu zapravo pripada drugom, to ukazuje na strukturnu aberaciju.

Zato je kariotip važan u dijagnostici:

- urođenih genetičkih poremećaja,
- razvojnih anomalija,
- nekih malignih bolesti, naročito leukemija, gde su česte translokacije hromozoma.

Slika 2.49 prikazuje humani kariotip sa G-trakama, gde su hromozomi raspoređeni u parove prema dužini i morfologiji. Svaki hromozom poseduje karakterističan obrazac tamnih i svetlih traka koji omogućava njegovu jedinstvenu identifikaciju. Tamne trake odgovaraju regionima sa većim stepenom kondenzacije i nižom genskom aktivnošću, dok svetle trake predstavljaju transkripciono aktivnije delove genoma. Analiza ovih obrazaca omogućava detekciju strukturnih promena kao što su delecije, duplikacije i translokacije.

Slika 2.50 prikazuje odnos između broja gena i veličine hromozoma u humanom genomu. Iako veći hromozomi sadrže više DNA, broj gena nije proporcionalan njihovoj veličini. Na primer, neki



Slika 2.48: Kariotip

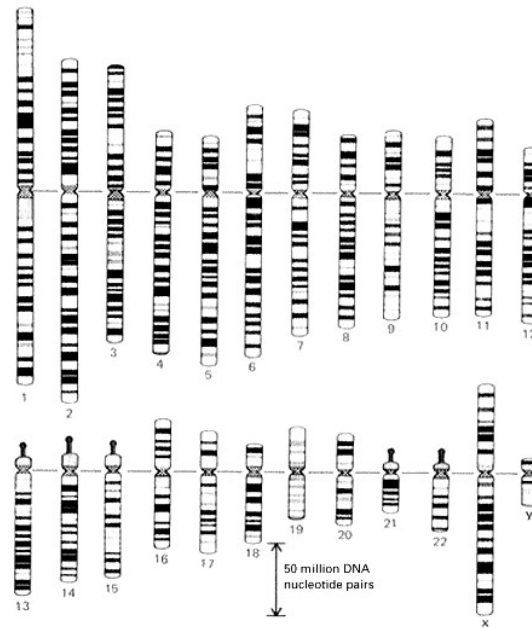
manji hromozomi, poput hromozoma 19, imaju visoku gustinu gena, dok veći hromozomi mogu sadržati relativno manje gena. Takođe se uočava da polni hromozom Y sadrži znatno manje gena u poređenju sa ostalim hromozomima.

2.8.3 Ploidija

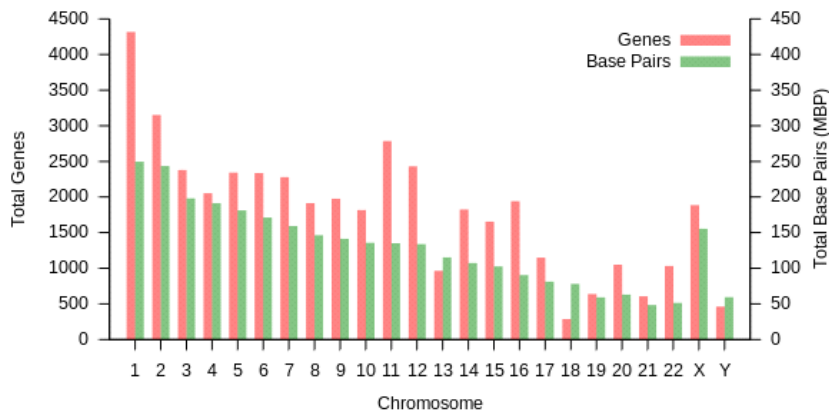
Ploidija je broj hromozomskih kompleta u ćeliji. Euploidija (slika 2.51) je pojava kada ćelija ili organizam ima umnožak haploidnog broja hromozoma (npr. $2n$ – diploidan, $3n$ tetraploidan, itd. broj hromozoma). Aneuploidija je pojava kada ćelija ili organizam ima nedostatak ili povećan broj nekog, pojedinačnog hromozoma (npr. u slučaju Daunovog sindroma trizomija hromozoma).

Slika 2.52 prikazuje primere poliploidnih ćelija kod sisara, koje sadrže više od dva seta hromozoma. Takve ćelije se javljaju u različitim tkivima i imaju specifične funkcije, kao što su proizvodnja trombocita u megakariocitima, detoksikacija u hepatocitima ili kontrakcija u kardiomiocitima. Broj kopija genoma može varirati od nekoliko do nekoliko stotina, u zavisnosti od tipa ćelije. Mikroskopski prikaz ilustruje organizaciju jedra i raspored genetskog materijala u ovim ćelijama.

Slika 2.53 prikazuje kariotip poliploidne ćelije jetre, u kojoj su hromozomi raspoređeni prema veličini, ali se uočava prisustvo četiri kopije svakog hromozoma, što ukazuje na tetraploidno stanje ($4n$). Za razliku od tipičnih diploidnih ćelija koje sadrže dve kopije svakog hromozoma, poliploidne ćelije imaju povećan broj genoma, što im omogućava veću metaboličku aktivnost i specijalizovane funkcije. Ovakva pojava predstavlja normalnu adaptaciju ćelije na funkcionalne zahteve organizma.



Slika 2.49: Humani kariotip

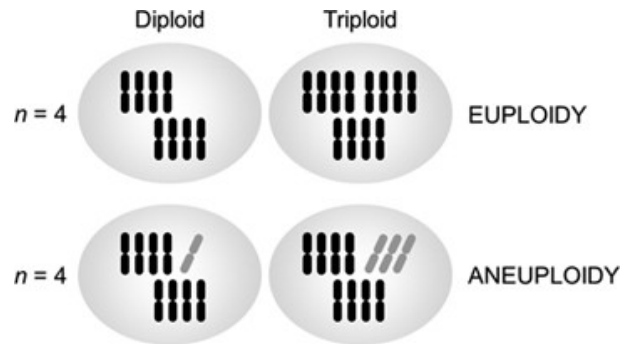


Slika 2.50: Odnos hromozoma i broja gena

2.9 Virusi

Virusi su infektivni, acelularni (nemaju ćelijsku strukturu) biološki entiteti koji se sastoje od nukleinske kiseline (DNA ili RNA) obavijene proteinskim omotačem (kapsidom), a ponekad i lipidnom membranom, i mogu se razmnožavati isključivo unutar živih ćelija domaćina koristeći njihov metabolički aparat.

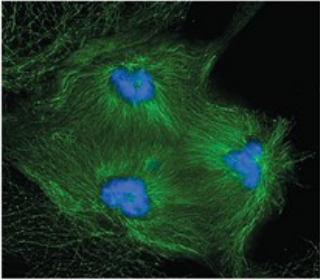
Slika 2.54 prikazuje klasifikaciju genetičkih elemenata prema njihovoj mobilnosti i stepenu „sebičnosti“, odnosno sposobnosti da se šire nezavisno od domaćina. Na dnu se nalaze hromozomi i organelna DNA, koji su stabilni i prenose se uglavnom vertikalno. Iznad njih su mobilni genetički elementi, poput plazmida i transpozona, koji mogu menjati svoju poziciju unutar genoma ili se prenositi između ćelija. Konjugativni elementi i virusi imaju još veću sposobnost horizontalnog prenosa, dok litički virusi predstavljaju najmobilniju i najnezavisniju grupu, jer se brzo razmnožavaju i često dovode do lize (razaranja) ćelije domaćina. Ova klasifikacija ilustruje



Slika 2.51: Euploidija i aneuploidaija

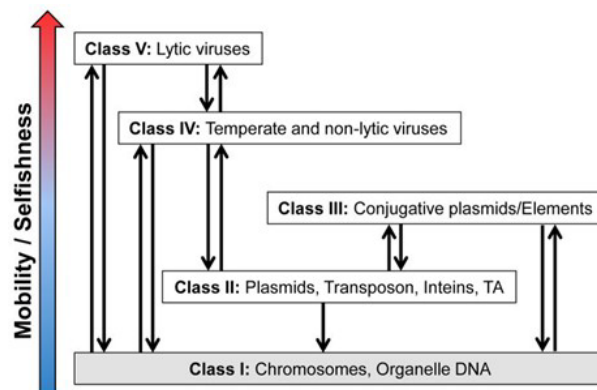
POLYPLOID CELL TYPES IN MAMMALS			
CELL TYPE	LOCATION	FUNCTION	NUMBER OF GENOME COPIES
Megakaryocyte	Bone marrow	Producing blood clotting platelets	Up to 128
Hepatocyte	Liver	Detoxification, metabolism	Typically 4 to 16
Trophoblast giant cell	Embryo	Promote implantation	Up to 1000
Cardiomyocyte	Heart	Contraction	Typically 4

Bonus DNA. The polyploid cells in mammalian bodies differ in their location, function, and number of chromosome sets (table). In a liver cell (right), the multiple chromosome copies (blue) have sorted into three clusters in preparation for cell division.



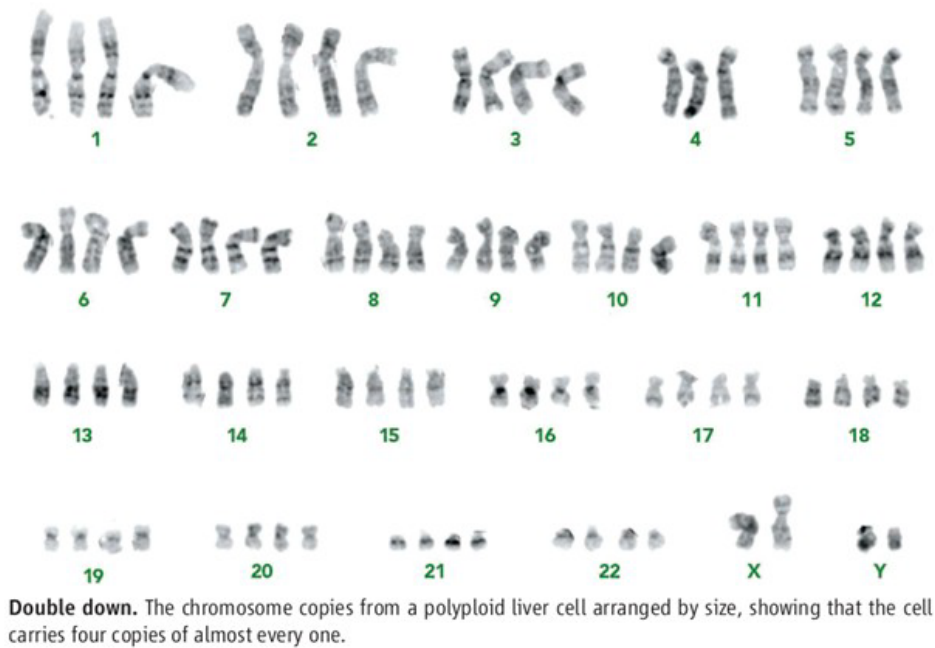
Slika 2.52: Ploidija kod sisara

kontinuum između stabilnog genoma i visoko mobilnih genetičkih elemenata.



Slika 2.54: Klasifikacija genetičkih elemenata prema njihovoj mobilnosti i stepenu „sebičnosti“

Slika 2.55 prikazuje Baltimore klasifikaciju virusa, koja ih deli u sedam grupa prema tipu genetičkog materijala i načinu sinteze mRNA. Virusi mogu imati DNA ili RNA, jednolančanu ili dvolančanu, a neki koriste i reverznu transkripciju. Bez obzira na razlike u genomu, svi virusi moraju proizvesti mRNA kako bi se sintetisali virusni proteini. Ova klasifikacija naglašava centralnu ulogu mRNA u životnom ciklusu virusa i omogućava razumevanje različitih mehanizama replikacije.



Slika 2.53: Kariotip poliploidne ćelije jetre

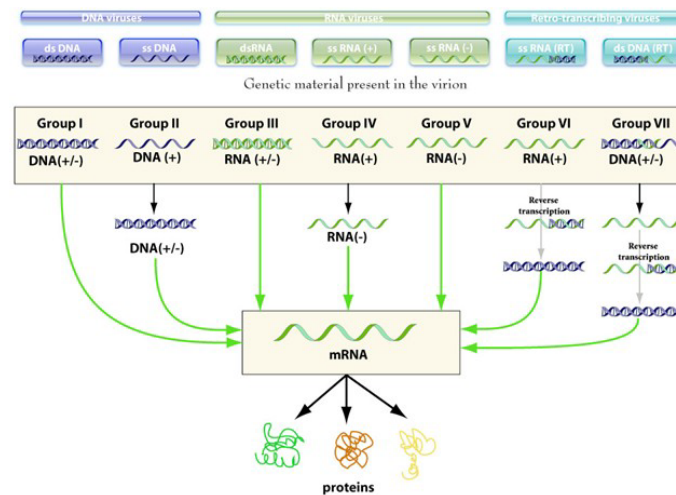
2.9.1 Korona virusi i SARS-2 CoV

Coronaviridae su familija RNA membranama ograničenih virusa koji inficiraju vodozemce, ptice i sisare. Viralni genom se sastoji od jednolančane (ss) +RNA (mRNA, kodirjuća RNA, RNA može odmah da se koristi kao mRNA), dužine između 26 do 32 kb. Ribonukleoproteinski kompleks (RNA je povezana sa proteinima) je zaštićen E-proteinima (Envelope protein, mali protein, (~20 nm), učestvuje u formiranju i stabilnosti virusa), u obliku latice i deteline i dvostrukom membranom u koju su integrisani S-šiljak (Spike protein formira „šiljke“ i omogućava ulazak u ćeliju) i M-membranski proteini (M protein daje oblik virusu) i koji daju virusu izgled krune pod elektronskim mikroskopom.

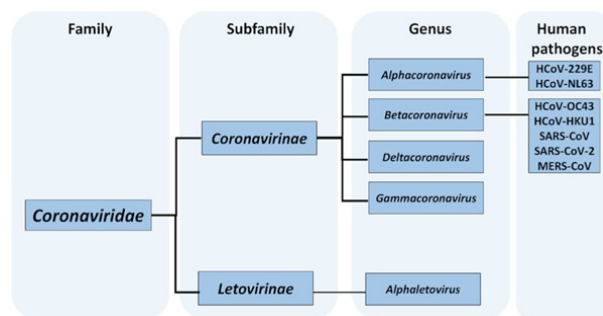
Slika 2.56 prikazuje taksonomsku klasifikaciju familije Coronaviridae, koja se deli na dve podfamilije: Orthocoronavirinae i Letovirinae. U okviru podfamilije Orthocoronavirinae razlikuju se četiri roda: Alphacoronavirus, Betacoronavirus, Deltacoronavirus i Gammacoronavirus. Humani patogeni uglavnom pripadaju rodovima Alpha i Beta, pri čemu su najznačajniji predstavnici SARS-CoV, MERS-CoV i SARS-CoV-2 iz roda Betacoronavirus.

Slika 2.57 prikazuje virusne čestice koronavirusa pod elektronskim mikroskopom. Virioni imaju sferičan oblik i sastoje se od ribonukleoproteinskog kompleksa koji sadrži jednolančanu RNA, okruženog lipidnim omotačem. Na površini omotača nalaze se S-proteini u obliku šiljaka, koji virusu daju karakterističan izgled krune i omogućavaju vezivanje za ćelije domaćina. Ova struktura je ključna za infektivnost i ulazak virusa u ćeliju.

Slika 2.58 prikazuje organizaciju genoma koronavirusa i proteine koje on kodira. Genom se sastoji od jednolančane RNA pozitivnog smera dužine oko 30 kb i sadrži više otvorenih čitalačkih okvira (ORF). Najveći deo genoma čine ORF1a i ORF1b, koji kodiraju nestrukturane proteine uključene u replikaciju i transkripciju virusne RNA. U završnom delu genoma nalaze se geni za strukturne proteine S, M, E i N, koji učestvuju u formiranju virusne čestice i infekciji ćelije



Slika 2.55: Baltimore klasifikacija virusa

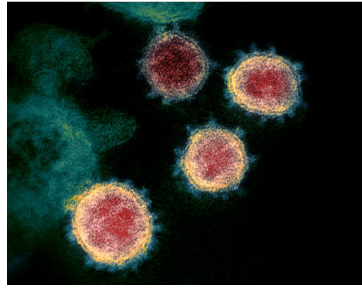


Slika 2.56: Taksonomska klasifikacija familije Coronaviridae

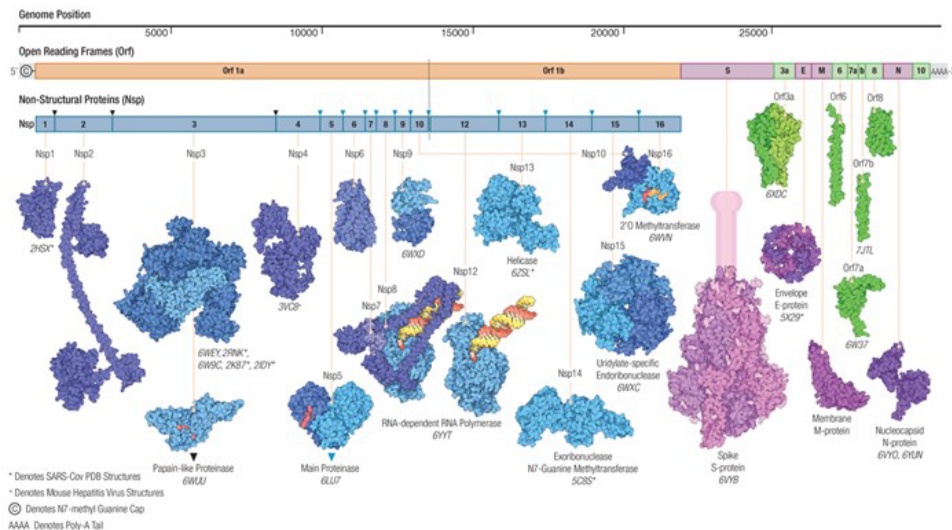
domaćina. Pored njih, prisutni su i pomoćni proteini koji imaju ulogu u regulaciji interakcije virusa sa domaćinom.

Slika 2.59 prikazuje organizaciju gena SARS-CoV-2 i njegov odnos sa srodnim koronavirusima. Gornji deo ilustruje raspored strukturnih, pomoćnih i preklapajućih gena duž genoma, naglašavajući visoku gustinu genetičkih informacija. Donji deo prikazuje filogenetske odnose između SARS-CoV-2 i drugih koronavirusa, posebno onih koji potiču od šišmiša i pangolina. Uočava se da su strukturni geni visoko očuvani, dok pomoćni geni pokazuju veću varijabilnost, što ukazuje na njihovu ulogu u adaptaciji virusa na različite domaćine. Na slici su prikazani samo geni nizvodno od ORF1ab, počev od Spike gena S.

Slika 2.60 prikazuje životni ciklus SARS-CoV-2 virusa u ćeliji domaćina. Virus se vezuje za ACE2 receptor i ulazi u ćeliju, nakon čega oslobađa svoju jednolančanu RNA pozitivnog smera. RNA se direktno koristi kao mRNA za sintezu virusnih proteina i formiranje replikacionog kompleksa, koji omogućava umnožavanje genoma i sintezu subgenomskih RNA. Strukturni proteini se sintetišu u endoplazmatičnom retikulumu i Golgijevom aparatu, gde se vrši sastavljanje novih virusnih čestica. Formirani virioni se zatim transportuju u vezikulama i oslobađaju iz ćelije.



Slika 2.57: Koronavirusne čestice (virione) pod elektronskim mikroskopom



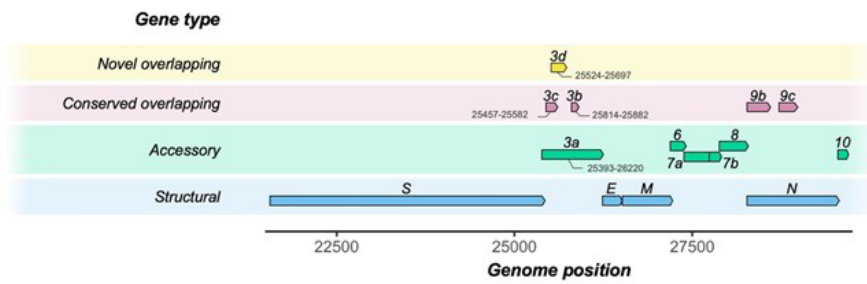
Slika 2.58: Koronavirusne čestice (virione) pod elektronskim mikroskopom

2.10 Određivanje primarne strukture (sekvencionisanje) DNA

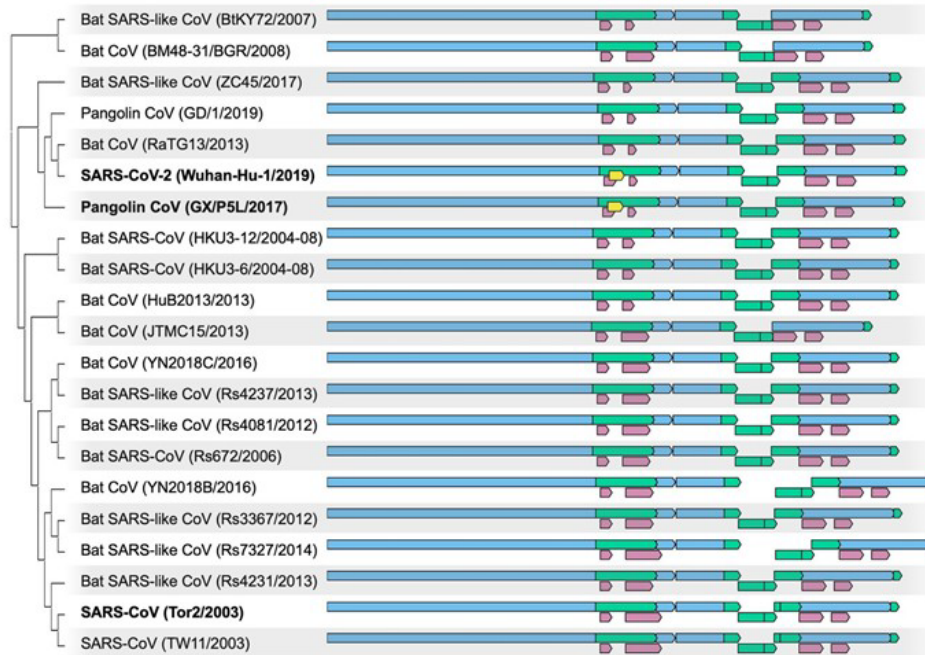
Frederick Sanger je britanski biohemičar i dvostruki dobitnik Nobelove nagrade, poznat po razvoju metode sekvenciranja DNA koja je postavila temelje moderne molekularne biologije i bioinformatike.

Sanger sekvenciranje (slika 2.61) je metoda određivanja redosleda nukleotida u DNA zasnovana na kontrolisanom prekidu sinteze lanca. Proces započinje pripremom matrične DNA i vezivanjem prajmera, nakon čega se u reakciju dodaju standardni nukleotidi (dNTP) i mala količina dideoksinukleotida (ddNTP). Tokom sinteze DNA, povremeno se ugrađuje ddNTP koji nema 3' OH grupu, što dovodi do prekida daljeg produžavanja lanca. Kao rezultat nastaje skup fragmenata različitih dužina, pri čemu svaki završava određenim nukleotidom (A, C, G ili T). Fragmenti se zatim razdvajaju gel elektroferezom u odvojenim trakama za svaki nukleotid, a očitavanjem rasporeda traka od najkraćih ka najdužim fragmentima određuje se tačan redosled baza u analiziranoj DNA sekvenci.

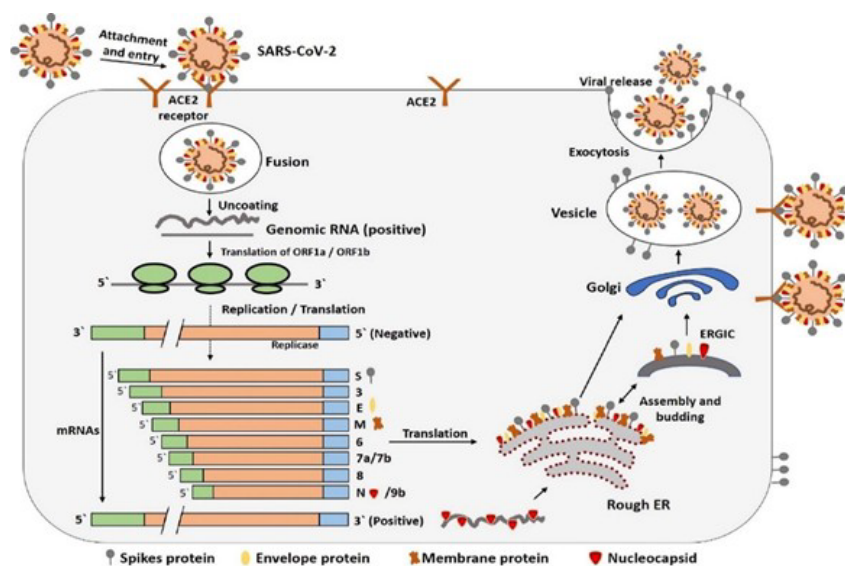
A SARS-CoV-2 (Wuhan-Hu-1)



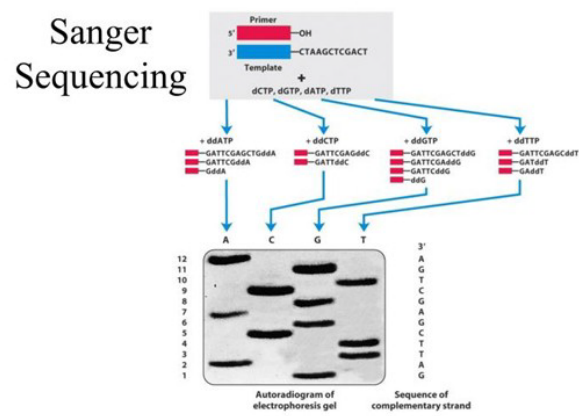
B Severe acute respiratory syndrome-related coronavirus species members



Slika 2.59: Genetski repertoar i evolutivni odnosi pripadnika vrste koronavirusa povezanih sa teškim akutnim respiratornim sindromom (SARS).



Slika 2.60: Životni ciklus SARS-CoV-2 virusa u ćeliji domaćina



Slika 2.61: Sanger sekvenciranje

Kratki uvod u hemiju i biologiju

belančevina

3.1 Amino kiseline

3.1.1 Amino kiseline koje ulaze u sastav proteina

Amino kiseline (AK) su organska jedinjenja koja u svom sastavu imaju amino (baznu, $-NH_2$) i karboksilnu (kiselinsku, $-COOH$) grupu. Do sada je iz prirodnih proizvoda izolovano preko 500 različitih AK. Amino kiseline koje na istom ugljenikovom atomu (alfa C-atom, C_α) imaju amino i karboksilnu grupu su označene kao α amino kiseline (ili, ređe, 2-amino kiseline), ili proteinogene AK, što znači da ulaze u sastav proteina. Za C_α atom je vezan i atom H. AK se međusobno razlikuju po radikalima (R, eng. residue), tj. bočnim grupama (nizovima). U upotrebi su trivijalni (biohemijski) nazivi AK, tako npr. glicin (Gly, gr. glikos – sladak) je trivijalni naziv za 2-aminoacetic acid (aminoethanoic acid).

Uobičajeni nazivi amino kiselina su:

- Alanin – verovatno od aldehid + „an“ (radi praktičnosti) + amin (1849)
- Arginin – kristališe kao srebrna so, od lat. argentum (srebro) (1886)
- Asparagin – prvi put izolovan iz špargle (1813)
- Aspartat – sličan asparaginu (1836)
- Glutamat – prvi put identifikovan u biljnom proteinu glutenu (1866)
- Glutamin – sličan glutamatu (1866)
- Glicin – od grč. glykys (sladak), ima sladak ukus (1848)
- Cistein – od grč. kystis (bešika), otkriven u kamenu u bešici (1882)
- Histidin – prvi put izolovan iz sperme jesetre, naziv od grč. histos (tkivo) (1896)
- Isoleucin – izomer leucina
- Leucin – od grč. leukos (beo), formira bele kristale (1820)
- Lizin – produkt hidrolize proteina, od grč. lysis (razlaganje) (1891)

- Metionin – bočni lanac sadrži atom sumpora (grč. theion) sa metil grupom (1928)
- Fenilalanin – alanin sa fenil grupom (1883)
- Prolin – izvedeno iz „pirolidin“, jer formira pirolidinski prsten (1904)
- Serin – od lat. sericum (svila), čest u svili (1865)
- Treonin – sličan šećeru sa četiri ugljenika – treози (1936)
- Triptofan – izolovan iz tripsinskog digestata proteina + grč. phainein (pojaviti se) (1890)
- Tirozin – pronađen u siru, od grč. tyros (sir) (1890)
- Valin – derivat valerijanske kiseline iz biljke roda Valeriana (1906)

Tabela 3.1 prikazuje dvosmislene i grupne oznake amino kiselina, koje se koriste u bioinformatički i analizi sekvenci kada identitet neke amino kiseline nije jednoznačno određen ili kada se posmatraju grupe sa sličnim osobinama. Na primer, oznake poput B (Asx) i Z (Glx) predstavljaju alternativu između dve srodne amino kiseline, dok oznaka X označava bilo koju ili nepoznatu amino kiselinu. Pored toga, uvedene su i oznake za funkcionalne grupe, kao što su hidrofobne, hidrofilne, aromatične, alifatične, male i naelektrisane amino kiseline, koje omogućavaju pojednostavljeno posmatranje biohemijskih svojstava proteina.

Osobine amino kiselina određene su njihovim bočnim grupama (R), koje utiču na strukturu i funkciju proteina. Hidrofobne amino kiseline teže da se nalaze u unutrašnjosti proteina, dok su hidrofilne najčešće izložene na površini i učestvuju u interakcijama sa vodom i drugim molekulima. Naelektrisane amino kiseline (pozitivne i negativne) igraju važnu ulogu u formiranju jonskih veza i stabilizaciji strukture. Aromatične amino kiseline doprinose stabilnosti kroz specifične interakcije, dok male amino kiseline omogućavaju fleksibilnost proteinskog lanca. Ove osobine su ključne za razumevanje savijanja proteina, njihove funkcije i interakcija u ćeliji.

3.1.2 Fizličko hemijske osobine amino kiselina

Amino kiseline su *hiralne* (osim glicina) jer se centralni $C\alpha$ atom vezuje za četiri različite grupe: amino, karboksilnu, vodonik i bočni lanac. Za neki molekul kažemo da je hiralan ako nema unutrašnju simetriju i ako se ne može poklopiti sa svojim likom u ogledalu. Zbog toga postoje dve konfiguracije, D i L, koje predstavljaju međusobne ogledalske slike i ne mogu se preklopiti. Određivanje konfiguracije može se izvršiti pomoću CORN pravila, pri čemu se posmatra redosled grupa oko $C\alpha$ atoma kada je vodonik orijentisan unazad.

Pravilo za određivanje L-konfiguracije amino kiselina i (tzv. CORN pravilo) je: ako se vodonik nalazi iza ravni posmatranja, raspored grupa $CO \rightarrow R \rightarrow N$ ide u smeru kazaljke na satu, što definiše L-oblik. U proteinima su uglavnom zastupljene L-amino kiseline. (slika 3.1).

Pravilo za određivanje D-konfiguracije amino kiselina je analogno L-konfiguraciji, ali sa suprotnim rasporedom grupa. Posmatra se $C\alpha$ atom tako da je vodonik (H) okrenut unazad (iza ravni posmatranja). Gleda se redosled grupa: $CO \rightarrow R \rightarrow N$. Ako taj redosled ide suprotno smeru kazaljke na satu, amino kiselina je D-konfiguracije.

Hidropatija je odnos bočnog lanca AK prema vodi (hidrofilan ili hidrofoban). Indeks hidropatije je broj koji predstavlja hidrofilne (hfl) ili hidrofobne (hfb) karakteristike bočnog

Oznaka	3-slovo	1-slovo	Uključene AK	Kodoni
Bilo koja / nepoznata	Xaa	X	Sve	NNN
Asparagin ili aspartat	Asx	B	D, N	RAY
Glutamin ili glutamat	Glx	Z	E, Q	SAR
Leucin ili izoleucin	Xle	J	I, L	YTR, ATH, CTY
Hidrofobne	Φ		V, I, L, F, W, Y, M	NTN, TAY, TGG
Aromatične	Ω		F, W, Y, H	YWY, TTY, TGG
Alifatične (ne-aromatične)	Ψ		V, I, L, M	VTN, TTR
Male	π		P, G, A, S	BCN, RGY, GGR
Hidrofilne	ζ		S, T, H, N, Q, E, D, K, R	VAN, WCN, CGN, AGY
Pozitivno naelektrisane	+		K, R, H	ARR, CRY, CGR
Negativno naelektrisane	-		D, E	GAN

Napomena: N = bilo koja baza, R = purini (A, G), Y = pirimidini (C, T).

Tabela 3.1: Dvosmislene i grupne oznake amino kiselina

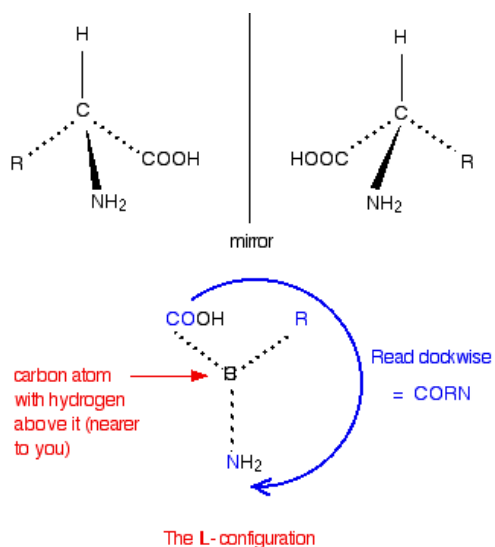
lanca. Postoji više skala (eksperimentalne, teorijske), od kojih je najpoznatija Kajt-Dulitlova (Jack Kyte and Russell F. Doolittle, 1982). Pozitivnim vrednostima su označene hidrofobne AK (npr.: 4,5, 4,2), a negativnim hidrofobne (npr.: -4,5, -3,9). Hidropatijski indeks amino kiselina je prikazan u tabeli 3.2.

Smatra se da prilikom formiranja proteinske 3D strukture hidropatija igra najznačajniju ulogu. Hidrofobne amino kiseline se kod globularnih proteina nalaze unutar proteinske globule, a hidrofobne spolja, na površini proteina (slika 3.2).

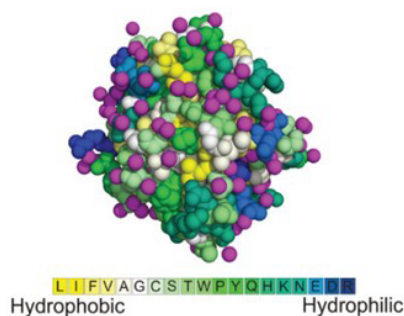
Postoje različiti hidropatski koeficijenti za amino kiseline koji su dobijeni različitim metodama (eksperimentalnim i teorijskim proračunima). Neki od njih su prikazani u tabeli 3.3.

Prema bočnim lancima AK se najčešće klasifikuju u grupe:

1. nepolarne alifatične (G, A, V, L, I, P) – hidrofobne AK koje se najčešće nalaze u unutrašnjosti proteina i doprinose stabilnosti strukture putem hidrofobnih interakcija;
2. polarne (koje mogu da ostvaruju interakcije sa vodom) nešaržirane (S, T, C, M, N, Q) – mogu formirati vodonične veze i često se nalaze na površini proteina ili u aktivnim mestima;
3. negativno šaržirane (kisele) (D, E) – nose negativan naboj i učestvuju u elektrostatičkim interakcijama i katalizi;
4. pozitivno šaržirane (bazne) (K, R, H) – nose pozitivan naboj i često učestvuju u vezivanju za negativno naelektrisane molekule (npr. DNK);
5. aromatične i heterociklične (F, Y, W) – sadrže aromatične prstenove, učestvuju u π - π interakcijama i često imaju ulogu u stabilizaciji proteinske strukture.



Slika 3.1: L-AK i D-AK



Slika 3.2: Uticaj hidropatije. Izvor: KESSEL, NIR BEN-TAL Introduction to Proteins Structure, Function, and Motion, 2e

Osim toga, amino kiseline se mogu klasifikovati i prema veličini i molekularnoj masi („small“ – male i „tiny“ – sićušne), zapremini (\AA^3), površini (\AA^2) itd. Podela amino kiselina prema fizičko-hemijskim osobinama može se videti na slici 3.3.

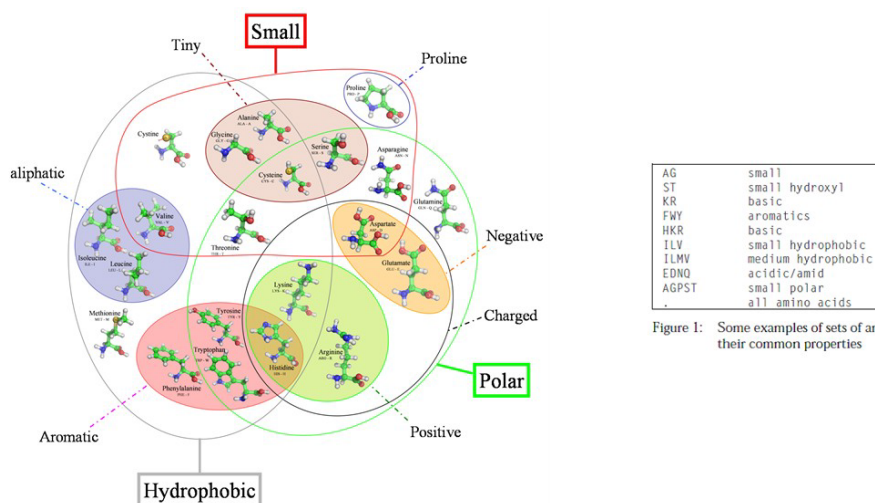


Figure 1: Some examples of sets of amino acids and their common properties

Slika 3.3: Fizičko-hemijske karakteristike AK

Amino kiselina	Hidropatijski indeks
Izoleucin	4.5
Valin	4.2
Leucin	3.8
Fenilalanin	2.8
Cistein/Cistin	2.5
Metionin	1.9
Alanin	1.8
Glicin	-0.4
Treonin	-0.7
Triptofan	-0.9
Serin	-0.8
Tirozin	-1.3
Prolin	-1.6
Histidin	-3.2
Glutaminska kiselina	-3.5
Aspartinska kiselina	-3.5
Asparagin	-3.5
Glutamin	-3.5
Lizin	-3.9
Arginin	-4.5

Tabela 3.2: Hidropatijski indeks amino kiselina

Tabela 3.4 prikazuje osnovne osobine amino kiselina uključujući strukturu bočnog lanca, hidrofobnost, polarnost i kiselinsko-bazne karakteristike. Ove osobine imaju ključnu ulogu u formiranju 3D strukture proteina i njihovoj funkciji.

Amino kiselina	Kratko	Oznaka	Bočni lanac	Hidrofobna	pKa	Polarna	pH osobina	Mala	Sićušna	Tip
Alanin	A	Ala	-CH ₃	✓	-	✗	-	✓	✓	alifatična
Cistein	C	Cys	-CH ₂ SH	✓	8.55	✓	kisela	✓	✓	-
Aspartat	D	Asp	-CH ₂ COOH	✗	3.67	✓	kisela	✓	✗	-
Glutamat	E	Glu	-CH ₂ CH ₂ COOH	✗	4.25	✓	kisela	✗	✗	-
Fenilalanin	F	Phe	-CH ₂ C ₆ H ₅	✓	-	✗	-	✗	✗	aromatična
Glicin	G	Gly	H	✓	-	✗	-	✓	✓	-
Histidin	H	His	-CH ₂ C ₃ H ₃ N ₂	✗	6.54	✓	slabo bazna	✗	✗	aromatična
Izoleucin	I	Ile	-CH(CH ₃)CH ₂ CH ₃	✓	-	✗	-	✗	✗	alifatična
Lizin	K	Lys	-(CH ₂) ₄ NH ₂	✗	10.40	✓	bazna	✗	✗	-
Leucin	L	Leu	-CH ₂ CH(CH ₃) ₂	✓	-	✗	-	✗	✗	alifatična
Metionin	M	Met	-CH ₂ CH ₂ SCH ₃	✓	-	✗	-	✗	✗	alifatična
Asparagin	N	Asn	-CH ₂ CONH ₂	✗	-	✓	-	✓	✗	-
Prolin	P	Pro	ciklična	✓	-	✗	-	✓	✓	-
Glutamin	Q	Gln	-CH ₂ CH ₂ CONH ₂	✗	-	✓	-	✗	✗	-
Arginin	R	Arg	-(CH ₂) ₃ NH-C(NH)NH ₂	✗	12.3	✓	jako bazna	✗	✗	-
Serin	S	Ser	-CH ₂ OH	✗	-	✓	-	✓	✓	-
Treonin	T	Thr	-CH(OH)CH ₃	✗	-	✓	-	✓	✗	-
Selenocistein	U	Sec	-CH ₂ SeH	✗	5.43	✓	kisela	✓	✓	-
Valin	V	Val	-CH(CH ₃) ₂	✓	-	✗	-	✗	✗	alifatična
Triptofan	W	Trp	-CH ₂ C ₈ H ₆ N	✓	-	✗	-	✗	✗	aromatična
Tirozin	Y	Tyr	-CH ₂ C ₆ H ₄ OH	✗	9.84	✓	slabo kisela	✗	✗	aromatična

*pKa je mera jačine kiseline, predstavlja pH vrednost pri kojoj je koncentracija protonovanog i deprotonovanog oblika funkcionalne grupe jednaka (50% : 50%).

Tabela 3.4: Osobine amino kiselina

AK	ZIMM	FAUC	BULL	COTHIA	JANIN	GUY	HEIJ	HW	KD	MEEK	MEEKKR	GUO
Ala	0.83	0.31	-200	0.38	1.70	0.10	-12.04	-0.5	1.8	0.5	1.1	2.2
Arg	0.83	-1.01	-120	0.01	0.00	-1.90	-39.23	3.0	-4.5	0.8	-0.4	0.9
Asn	0.09	-0.60	-80	0.12	0.40	-0.48	4.25	0.2	-3.5	0.8	-4.2	-0.8
Asp	0.64	-0.77	-200	0.15	0.40	0.78	23.22	3.0	-3.5	-8.2	-1.6	-2.6
Cys	1.48	1.54	-450	0.50	4.60	-1.42	3.95	-1.0	2.5	-6.8	7.1	2.6
Gln	0.00	-0.22	160	0.07	0.30	-0.95	2.16	0.2	-3.5	-4.8	-2.9	-0.2
Glu	0.65	-0.64	-300	0.18	0.30	0.83	16.81	3.0	-3.5	-16.9	0.7	-0.2
Gly	0.10	0.00	0	0.36	1.80	0.33	-7.85	0.0	-0.4	-0.7	0.2	0.0
His	1.10	0.13	-120	0.17	0.80	-0.50	6.28	-0.5	-3.2	-3.5	-0.7	2.2
Ile	2.52	1.80	2260	0.60	3.10	-1.13	-18.32	-1.8	4.5	13.9	8.5	8.3
Leu	3.07	1.70	2460	0.45	2.40	1.18	-17.79	-1.8	3.8	8.8	11.0	9.0
Lys	1.60	-0.99	-350	0.03	0.05	-1.40	9.71	3.0	-3.9	0.1	-1.9	0.0
Met	1.40	1.23	1470	0.40	1.90	-1.59	-8.86	-1.3	1.9	4.8	5.4	6.0
Phe	2.75	1.79	2330	0.50	2.20	-2.12	-21.98	-2.5	2.8	13.2	13.0	9.0
Pro	2.70	0.72	-980	0.18	0.60	2.73	5.82	0.0	-1.6	6.1	4.4	2.2
Ser	0.14	-0.04	-300	0.22	0.80	0.52	-1.54	0.3	-0.8	-0.8	-3.2	2.0
Thr	0.53	0.26	-920	0.23	0.70	0.07	-4.15	-0.4	-0.7	2.7	-1.7	0.3
Trp	0.37	2.25	2010	0.27	1.60	-0.51	-16.19	-3.4	-0.9	14.9	17.0	9.5
Tyr	2.97	0.96	2240	0.15	1.50	-0.21	-1.51	-2.3	-1.3	6.1	7.4	4.6
Val	1.79	1.22	1560	0.54	2.90	-1.27	-16.22	-1.5	4.2	2.7	5.9	5.7

Tabela 3.3: Vrednosti hidrofobnosti amino kiselina prema 12 različitih skala

Selenocistein i pirolizin (tabela 3.5) predstavljaju tzv. 21. i 22. proteinogene amino kiseline. Ugrađuju se u proteine tokom translacije pomoću specifičnih signalnih sekvenci i specijalizovanih mehanizama.

Amino kiselina	3-slovo oznaka	1-slovo oznaka	Molekulska masa
Selenocistein	Sec	U	168.064
Pirolizin	Pyl	O	255.313

Tabela 3.5: 21. i 22. amino kiselina

Aminokiselinski indeks (AAindex) (<https://www.genome.jp/aaindex/>) je baza podataka numeričkih indeksa koji predstavljaju različite fizičko-hemijske i biohemijske karakteristike amino kiselina.

Slika 3.4 prikazuje organizaciju AAindex baze kroz minimalno razapinjuća stabla, gde svaki čvor predstavlja jedan indeks, a veze između čvorova označavaju sličnost između fizičko-hemijskih i bioloških osobina amino kiselina, grupisanih po tipu osobina.

Tabela 3.6 prikazuje najvažnije strukturne i funkcionalne osobine amino kiselina. Boje označavaju osnovne grupe amino kiselina prema prirodi bočnog lanca: hidrofobne, polarne, kisele, bazne, aromatične i posebne/dvosmislene oznake. U tabeli 3.7 je prikazana legenda boja za klasifikaciju amino kiselina.

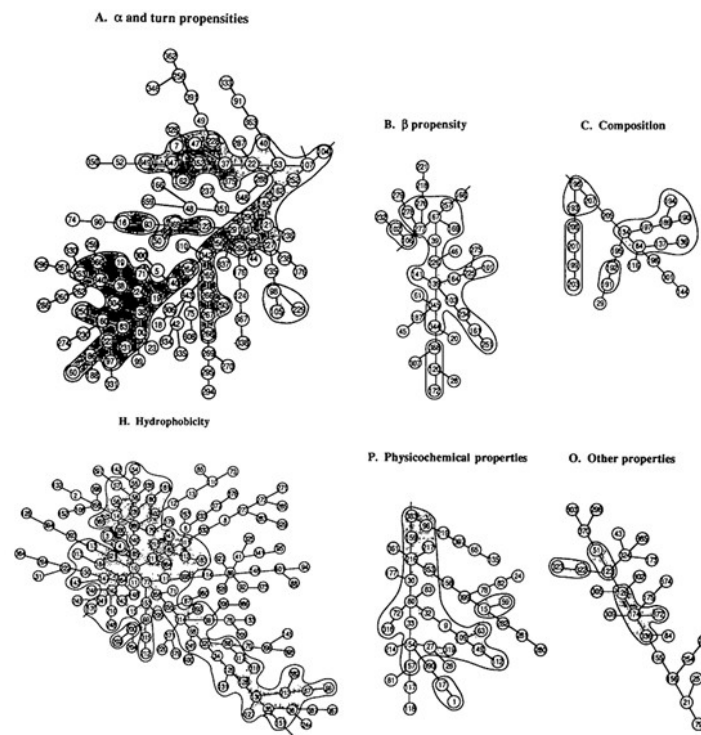


Fig. 5. Enlarged drawing of the minimum spanning tree of amino acid indices. Each amino acid index is identified in the text by the single-letter classification code, A, B, C, H, P or O, followed by the number listed in the AAindex, available on the Internet.

Slika 3.4: Organizacija AAindex baze

Tabela 3.6: Osobine i napomene o amino kiselinama

Amino kiselina	1-slovo	3-slovo	Opis
Alanin	A	Ala	Veoma zastupljena, mala i svestrana amino kiselina. Ponaša se skoro neutralno i može se nalaziti i u hidrofilnim regionima na površini proteina i u hidrofobnim regionima u unutrašnjosti.
Asparagin ili aspartat	B	Asx	Oznaka koja se koristi kada na datoj poziciji može stajati ili asparagin ili aspartat.
Cistein	C	Cys	Atom sumpora lako gradi veze sa teškim metalima. U oksidacionim uslovima dve cisteinske ostatke mogu formirati disulfidnu vezu i dati cistin. Disulfidne veze stabilizuju terciarnu strukturu proteina i povećavaju otpornost na denaturaciju.
Aspartat	D	Asp	Kisela amino kiselina sa hidrofilnom bočnom grupom i jakim negativnim nabojem. Obično se nalazi na površini proteina, doprinosi rastvorljivosti i često učestvuje u vezivanju pozitivno naelektrisanih molekula i jona.

Amino kiselina	1-slovo	3-slovo	Opis
Glutamat	E	Glu	Po osobinama sličan aspartatu, ali ima duži i nešto fleksibilniji bočni lanac. Negativno je naelektrisan i često učestvuje u jonskim interakcijama.
Fenilalanin	F	Phe	Esencijalna amino kiselina. Sadrži veliki hidrofobni aromatični prsten i zbog toga je često usmerena ka unutrašnjosti savijenog proteinskog molekula.
Glicin	G	Gly	Najmanja amino kiselina. Zbog velike fleksibilnosti manje je česta kao krut strukturni element.
Histidin	H	His	U blago kiselim uslovima azot u imidazolnom prstenu može promeniti protonaciju, pa histidin menja i konformaciju polipeptida. Zbog toga je važan u katalizi enzima i regulaciji strukture.
Izoleucin	I	Ile	Esencijalna, alifatična i izrazito hidrofobna amino kiselina. Zajedno sa leucinom i valinom često se nalazi u unutrašnjosti proteina i doprinosi pravilnom savijanju.
Leucin ili izoleucin	J	Xle	Oznaka koja se koristi kada na datoj poziciji može stajati leucin ili izoleucin.
Lizin	K	Lys	Esencijalna bazna amino kiselina sa dugim fleksibilnim bočnim lancem i pozitivno naelektrisanim krajem. Važna je za vezivanje negativno naelektrisanih molekula, npr. DNK.
Leucin	L	Leu	Esencijalna amino kiselina, slična izoleucinu i valinu. Hidrofobna je i doprinosi stabilizaciji unutrašnjosti proteina.
Metionin	M	Met	Esencijalna amino kiselina sa sumporom. Često je prva amino kiselina ugrađena u novonastali protein. Njena metil grupa može biti aktivirana i koristiti se u različitim reakcijama prenosa grupa.
Asparagin	N	Asn	Polarna, nešaržirana amino kiselina. Slična aspartatu, ali umesto karboksilne ima amidnu grupu. Često učestvuje u vodoničnim vezama.
Pirolizin	O	Pyl	Retka amino kiselina, slična lizinu, sa dodatim pirolinskim prstenom. Prisutan je u nekim mikroorganizmima.
Prolin	P	Pro	Posebna amino kiselina sa cikličnom strukturom koja ograničava pokretljivost lanca. Često uvodi okrete i prekida regularne sekundarne strukture poput α -heliksa i β -listova.

Amino kiselina	1-slovo	3-slovo	Opis
Glutamin	Q	Gln	Sličan glutaminskoj kiselini, ali sadrži amidnu grupu umesto karboksilne. U proteinima učestvuje u vodoničnim vezama i služi kao rezervoar amonijaka.
Arginin	R	Arg	Funkcionalno sličan lizinu. Sadrži snažno baznu guanidinsku grupu i često učestvuje u elektrostatičkim interakcijama i vezivanju negativno naelektrisanih molekula.
Serin	S	Ser	Sadrži kratki bočni lanac sa hidroksilnom grupom. Veoma je hidrofilan, često učestvuje u vodoničnim vezama i katalitičkim mestima enzima.
Treonin	T	Thr	Esencijalna amino kiselina sa hidroksilnom grupom. Po osobinama slična serinu i često učestvuje u vodoničnim vezama i posttranslacionim modifikacijama.
Selenocistein	U	Sec	Naziva se 21. amino kiselinom. Strukturno je sličan cisteinu, ali umesto sumpora sadrži selen i ugrađuje se specifičnim mehanizmom tokom translacije.
Valin	V	Val	Esencijalna, hidrofobna i alifatična amino kiselina. Stabilizuje unutrašnjost proteina i često se nalazi u hidrofobnom jezgru.
Triptofan	W	Trp	Esencijalna amino kiselina sa velikim aromatičnim prstenom. Fluorescentna je i služi kao prekursor serotonina. Često doprinosi stabilnosti proteinske strukture.
Nepoznata	X	Xaa	Oznaka koja se koristi kada amino kiselina nije poznata ili nije značajna za datu analizu.
Tirozin	Y	Tyr	Aromatična amino kiselina, slična fenilalaninu i triptofanu. Prekursor je melanina, epinefrina i tiroidnih hormona. Može biti fluorescentna, mada je fluorescencija često oslabljena prenosom energije na triptofan.
Glutamat ili glutamin	Z	Glx	Oznaka koja se koristi kada na datoj poziciji može stajati glutamat ili glutamin.

Neke amino kiseline se mogu posttranslaciono modifikovati. Najčešće kovalentne modifikacije bočnih lanaca uključuju: fosforilacije (dodaje se fosfatna grupa, npr. fosfo-serin, fosfo-tirozin), glikozilacije (dodaju se šećeri (glikani)), lipidacije (dodaju se lipidi (masti)), zatim formiranje „cross-link“ disulfidnih (-S-S-) veza između dva Cys, itd. Posttranslacione modifikacije igraju važnu ulogu u regulaciji proteinske aktivnosti, međucelijskoj komunikaciji itd. Tabela 3.8 prikazuje neke posttranslacione modifikacije amino kiselina.

Boja	Opis
	Hidrofobne (nepolarne) amino kiseline – nalaze se u unutrašnjosti proteina
	Polarne nešaržirane amino kiseline – učestvuju u vodoničnim vezama
	Negativno šaržirane (kisele) amino kiseline – nose negativan naboj
	Pozitivno šaržirane (bazne) amino kiseline – nose pozitivan naboj
	Aromatične amino kiseline – sadrže aromatične prstenove
	Posebne / dvosmislene oznake – specijalne ili neodređene amino kiseline

Tabela 3.7: Legenda boja za klasifikaciju amino kiselina

Tabela 3.8: Posttranslacione modifikacije amino kiselina

Amino kiselina	Oznaka	Modifikacije
Alanin	Ala	N-acetilacija (N-kraj)
Arginin	Arg	deiminacija u citrulin, metilacija
Asparagin	Asn	deamidacija u Asp/izoAsp, N-vezana glikozilacija
Aspartat	Asp	izomerizacija u izoaspartat
Cistein	Cys	disulfidne veze, oksidacija (sulfenilna, sulfinska, sulfonska), palmitoilacija, S-nitrozilacija
Glutamin	Gln	ciklizacija u piroglutamat, deamidacija
Glutamat	Glu	ciklizacija u piroglutamat (N-kraj), gama-karboksilacija
Glicin	Gly	N-miristoilacija (N-kraj), N-acetilacija (N-kraj)
Histidin	His	fosforilacija
Izoleucin	Ile	–
Leucin	Leu	–
Lizin	Lys	acetilacija, ubikvitinacija, SUMOilacija, metilacija, hidroksilacija
Metionin	Met	N-acetilacija (N-kraj), oksidacija
Fenilalanin	Phe	–
Prolin	Pro	hidroksilacija
Serin	Ser	fosforilacija, O-vezana glikozilacija, N-acetilacija (N-kraj)
Treonin	Thr	fosforilacija, O-vezana glikozilacija, N-acetilacija (N-kraj)
Triptofan	Trp	oksidacija, formiranje kinurenina
Tirozin	Tyr	sulfatacija, fosforilacija
Valin	Val	N-acetilacija (N-kraj)

Proteom je skup svih proteina koji se ekspresuju u određenoj ćeliji, tkivu ili organizmu u datim uslovima i vremenskom trenutku. Tabela 3.9 prikazuje da je zastupljenost amino kiselina u

proteomima arheja, bakterija i eukariota je veoma slična, što ukazuje na evolutivnu konzerviranost proteinskog sastava, pri čemu dominiraju male i hidrofobne amino kiseline, dok su aromatične i reaktivne amino kiseline ređe. Male razlike postoje: eukarioti imaju više Ser, Pro, Gln, a bakterije malo više Ala i Leu. Esencijalne amino kiseline su one koje organizam ne može sam da sintetizuje i koje se moraju unositi putem ishrane.

Tabela 3.9: Kodoni, učestalost i esencijalnost amino kiselina

Amino kiselina	1-slovo	3-slovo	Kodoni	Arheje (%)	Bakterije (%)	Eukarioti (%)	Čovek (%)	Esencijalna
Alanin	A	Ala	GCU, GCC, GCA, GCG	8.2	10.06	7.63	7.01	Ne
Cistein	C	Cys	UGU, UGC	0.98	0.94	1.76	2.3	Uslovno
Aspartat	D	Asp	GAU, GAC	6.21	5.59	5.4	4.73	Ne
Glutamat	E	Glu	GAA, GAG	7.69	6.15	6.42	7.09	Uslovno
Fenilalanin	F	Phe	UUU, UUC	3.86	3.89	3.87	3.65	Da
Glicin	G	Gly	GGU, GGC, GGA, GGG	7.58	7.76	6.33	6.58	Uslovno
Histidin	H	His	CAU, CAC	1.77	2.06	2.44	2.63	Da
Izoleucin	I	Ile	AUU, AUC, AUA	7.03	5.89	5.1	4.33	Da
Lizin	K	Lys	AAA, AAG	5.27	4.68	5.64	5.72	Da
Leucin	L	Leu	UUA, UUG, CUU, CUC, CUA, CUG	9.31	10.09	9.29	9.97	Da
Metionin	M	Met	AUG	2.35	2.38	2.25	2.13	Da
Asparagin	N	Asn	AAU, AAC	3.68	3.58	4.28	3.58	Ne
Pirolizin	O	Pyl	UAG*	0	0	0	0	Ne
Prolin	P	Pro	CCU, CCC, CCA, CCG	4.26	4.61	5.41	6.31	Ne
Glutamin	Q	Gln	CAA, CAG	2.38	3.58	4.21	4.77	Ne
Arginin	R	Arg	CGU, CGC, CGA, CGG, AGA, AGG	5.51	5.88	5.71	5.64	Uslovno
Serin	S	Ser	UCU, UCC, UCA, UCG, AGU, AGC	6.17	5.85	8.34	8.33	Ne
Treonin	T	Thr	ACU, ACC, ACA, ACG	5.44	5.52	5.56	5.36	Da
Selenocistein	U	Sec	UGA**	0	0	0	> 0	Ne
Valin	V	Val	GUU, GUC, GUA, GUG	7.8	7.27	6.2	5.96	Da
Triptofan	W	Trp	UGG	1.03	1.27	1.24	1.22	Da
Tirozin	Y	Tyr	UAU, UAC	3.35	2.94	2.87	2.66	Uslovno
Stop kodon	–	Term	UAA, UAG, UGA	–	–	–	–	–

Slika 3.5 prikazuje prosečnu zastupljenost amino kiselina u proteinima izraženu u procentima, na osnovu analize velikog broja proteinskih sekvenci. Najzastupljenije su alifatične i male amino kiseline, poput leucina, alanina, glicina i serina, koje doprinose stabilnosti proteinske strukture. Sa druge strane, aromatične i sumporne amino kiseline, kao što su triptofan i cistein, znatno su ređe zbog složenije biosinteze i specifičnih hemijskih osobina. Ova raspodela odražava evolutivnu optimizaciju između stabilnosti, funkcionalnosti i energetske troškova sinteze proteina. Stubić amino kiseline je obojen prema njenoj osobini: alifatične - svetlosiva, kisele — roze, male hidrosilne (Ser i Thr) — zelena, bazne — svetloplava, aromatične (Phe, Trp, Tyr) — crna, amidne — bela, sumporne — žuta.

Novije evolucionarne studije ukazuju da amino kiseline : Gly, Ala, Asp, Val, Ser, Pro, Glu, Leu, Thr (tj. G, A, D, V, S, P, E, L, T) pripadaju grupi amino kiselina *ranog* genetskog koda, dok AK: Cys, Met, Tyr, Trp, His, Phe (tj. C, M, Y, W, H, F) predstavljaju *kasnije* proširenje genetskog koda. Rane amino kiseline (Gly, Ala, Asp, Val, Ser, Pro, Glu, Leu, Thr) smatraju se „starijim” jer su hemijski jednostavnije, tj. imaju jednostavne bočne lance (često male ili alifatične) i lako se sintetizuju bez složenih metaboličkih puteva. Nasuprot tome, „kasnije” amino kiseline (Cys, Met, Tyr, Trp, His, Phe) imaju složenije strukture (npr. aromatične prstenove ili sumporne grupe) i njihova biosinteza zahteva razvijenije enzimske sisteme, zbog toga se smatra da su se pojavile kasnije tokom evolucije genetskog koda.

Vennovi dijagrami 3.6 ilustruju povezanost između hemijskih osobina amino kiselina, njihove pozicije u genetskom kodu i funkcionalne uloge u proteinima. U gornjem dijagramu raspored amino kiselina zavisi od centralnog nukleotida kodona i njihovih fizičko-hemijskih karakteristika, pri čemu se jasno vidi preklapanje različitih osobina kao što su hidrofobnost, polaritet i veličina. Donji dijagram prikazuje učestalost pojavljivanja amino kiselina u katalitičkim mestima proteina, gde su najčešće zastupljene one sa reaktivnim ili naelektrisanim bočnim lancima (npr. His,

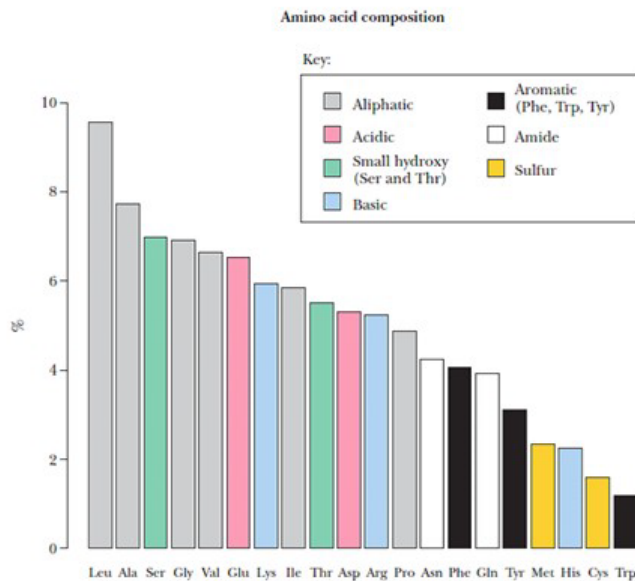


FIGURE 5.16 Amino acid composition: frequencies of the various amino acids in proteins for all the proteins in the SWISS-PROT protein knowledgebase. These data are derived from the amino acid composition of more than 100,000 different proteins (representing more than 40,000,000 amino acid residues). The range is from leucine at 9.55% to tryptophan at 1.18% of all residues.

Slika 3.5: Prosećnu zastupljenost amino kiselina u proteinima

Asp, Glu, Lys, Arg). Katalitička mesta proteina (ili aktivna mesta enzima) su specifični delovi proteinske strukture u kojima se odvija hemijska reakcija. Ovi rezultati ukazuju na vezu između strukture genetskog koda i funkcionalnih svojstava proteina, što podržava hipoteze o evolutivnom razvoju genetskog koda i izboru amino kiselina u biološkim sistemima.

3.1.3 Polimerizacija amino kiselina - peptidna veza

Peptidna veza (amidna veza) (slika 3.7a) je kovalentna hemijska veza koja se formira između dve amino kiseline kada karboksilna grupa jedne amino kiseline reaguje sa amino grupom druge amino kiseline, uz otpuštanje molekula vode (H_2O). Ovaj proces je dehidrataciona sinteza (takođe je poznata kao reakcija kondenzacije). Rezultujuća funkcionalna grupa sa četiri atoma, $C(=O)NH-$, naziva se peptidna veza, dok je rezultujući molekul amid.

Peptidna veza je rezonantno hibridizovana (slika 3.7b), tj. ima parcijalne karakteristike dvostruke veze ($C=N$ (jednostruka veza) i $C=N$ (dvostruka veza)), pa oko nje nije moguća slobodna rotacija (ω diedralni ugao je normalno 180°). Na ovaj način se formira proteinski polipeptidni lanac (primarna struktura), čija dužina može da varira od reda veličine 10 do desetina hiljada AK (humani protein titin se sastoji od preko 35.000 AK).

Peptidna veza apsorbuje UV svetlost u opsegu 190–230 nm zbog elektronskih tranzicija u karbonilnoj grupi ($C=O$), što je čini osetljivom na UV zračenje i omogućava njenu primenu u spektroskopskim analizama proteina.

Slika 3.8 prikazuje planarni i rigidni karakter peptidne grupe $CO-NH$. U delu (a) sivi paralelogram označava ravan peptidne grupe u polipeptidu. Standardne dužine veza su izražene u Å. Obratiti pažnju da je $C-N$ veza u peptidnoj grupi kraća od $C-N$ veza u α -ugljeniku, što ukazuje na parcijalni karakter dvostruke veze peptidne veze. U delu (b) prikazana je elektronska

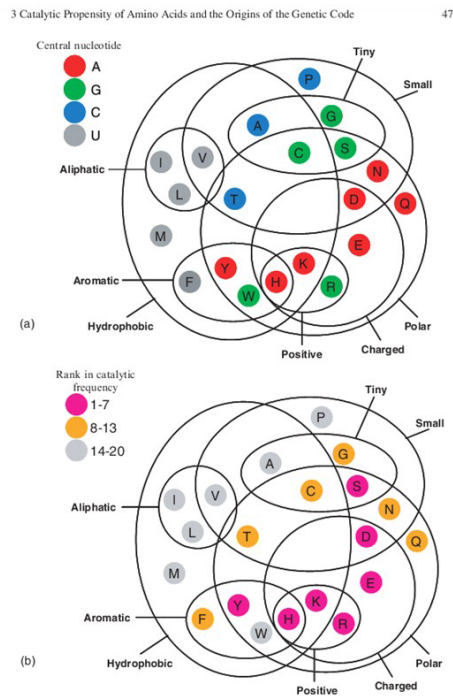


Fig. 7 Venn diagrams of amino acids (chemical sets from Taylor, 1986). (a) Distribution of amino acids based on the middle letter of the genetic code. (b) distribution of amino acids according to catalytic frequency ranks and chemical properties

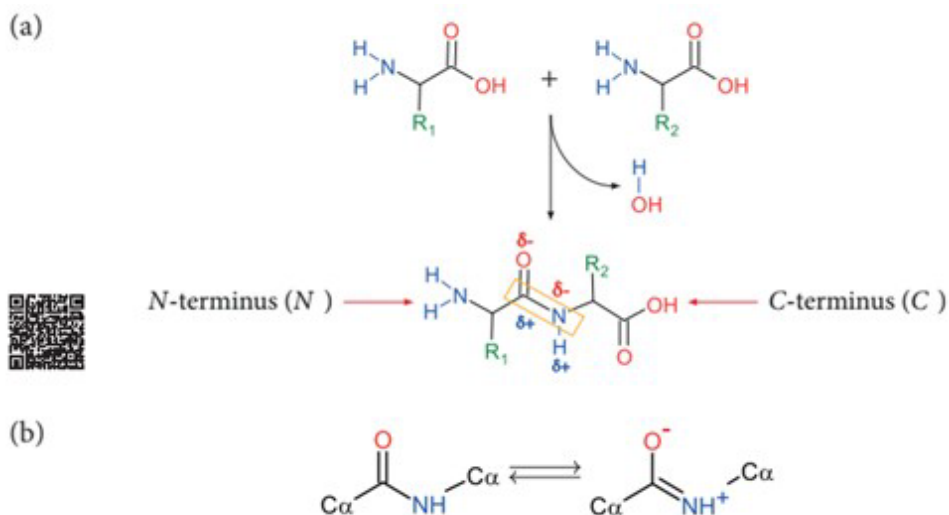
Slika 3.6: Venovi dijagrami amino kiselina (hemijski skupovi prema Tayloru, 1986). (a) Raspodela amino kiselina na osnovu srednjeg slova genetskog koda, (b) raspodela amino kiselina prema rangovima katalitičke učestalosti i hemijskim osobinama.

rezonanca peptidne veze, tj. da peptidna veza nema jednu fiksnu strukturu, već je kombinacija više mogućih elektronskih rasporeda (rezonantnih struktura). Elektronska rezonanca peptidne veze nastaje delokalizacijom elektrona između C=O i C–N veza, čime C–N dobija parcijalni karakter dvostruke veze, što dovodi do planarnosti i rigidnosti peptidne grupe. Slika 3.9 prikazuje detaljnu geometriju (rastojanja i uglove pojedinih atoma) peptidne veze između dve amino kiselina.

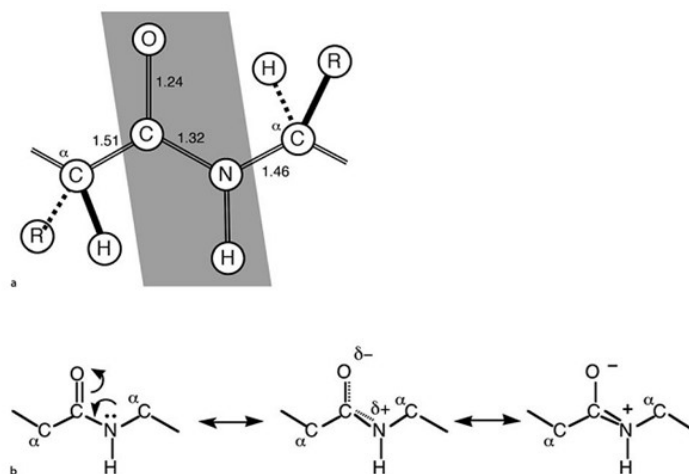
3.1.4 Diedralni uglovi

Diedralni uglovi u proteinima definišu se pomoću četiri uzastopna atoma i uključuju φ (C–N–C α –C), ψ (N–C α –C–N), ω (C α –C–N–C α), kao i χ uglove bočnih lanaca, koji određuju trodimenzionalnu konformaciju proteina.

Peptidna veza može postojati u dve konformacije, cis i trans, koje se razlikuju po relativnom položaju bočnih lanaca susednih amino kiselina (slika 3.11). U trans konfiguraciji bočni lanci su raspoređeni na suprotnim stranama peptidne veze, čime se minimizuju steričke smetnje i postiže veća stabilnost, zbog čega je ovaj oblik dominantan u proteinima. Nasuprot tome, u cis konfiguraciji bočni lanci se nalaze na istoj strani, što dovodi do steričkih sudara i znatno manje stabilnosti, pa se ovaj oblik retko javlja. Sterički sudari su situacije kada se atomi ili grupe atoma u molekulu previše približe jedni drugima, tako da dolazi do odbijanja između njihovih elektronskih oblaka. Izuzetak predstavlja aminokiselina prolin, kod koje je cis konformacija češća zbog specifične ciklične strukture njenog bočnog lanca.



Slika 3.7: Formiranje peptidne veze



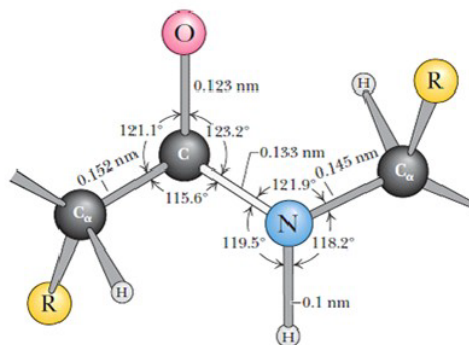
Slika 3.8: Planarni i rigidni karakter peptidne grupe CO–NH

Ramačandranov dijagram (slika 3.12) prikazuje dozvoljene kombinacije diedralnih uglova φ i ψ u proteinskom lancu, pri čemu su dozvoljene konformacije određene prostornim ograničenjima i izbegavanjem steričkih sudara između atoma. Ove dozvoljene oblasti odgovaraju stabilnim sekundarnim strukturama, kao što su α -heliksi i β -trake, dok su ostale kombinacije uglova energetski nepovoljne i zato se ne javljaju u proteinima.

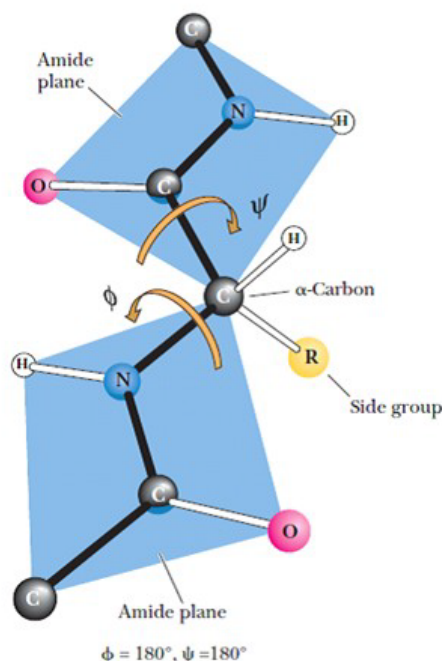
3.2 Nivoi organizacije polipeptidnih lanaca

3.2.1 Primarna struktura proteina

Primarna struktura proteina definisana je redosledom, brojem i sadržajem amino kiselina u peptidnom lancu. Proteini koji se sastoje od manje od 50 amino kiselina definišu se kao *peptidi* (oligopeptidi, npr. neuropeptidi, hormoni, antibakterijski peptidi), dok se polipeptidi sa većim brojem amino kiselina definišu kao proteini. Prema konvenciji, proteinski lanac se predstavlja sa α -amino grupom prve, sa leve strane, amino kiseline („N-terminus“) i α -karboksilnom grupom desno



Slika 3.9: Geometrija peptidne veze

Slika 3.10: Diedralni uglovi φ i ψ

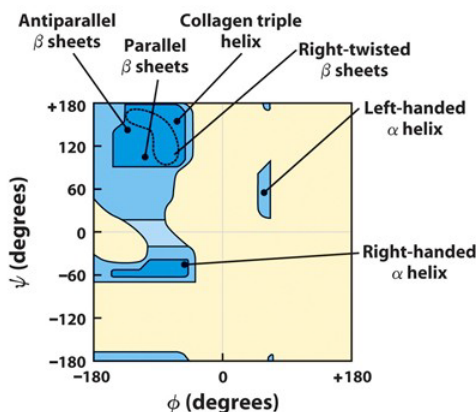
(„C-terminus“). Redosled amino kiselina u proteinskom lancu sadrži informaciju o proteinskoj 3D strukturi. Informacija sadržana u redosledu amino kiselina proteina realizuje se kroz nekovalentne interakcije između različitih bočnih lanaca amino kiselina (R), koje omogućavaju formiranje pravilne 3D strukture proteina.

Kovalentne interakcije su jake hemijske veze zasnovane na deljenju elektrona između atoma (npr. peptidna veza (C–N), disulfidna veza (–S–S– između Cys)), dok su nekovalentne interakcije slabije međumolekulske sile koje ne uključuju deljenje elektrona, ali su ključne za stabilizaciju trodimenzionalne strukture proteina (npr. vodonične veze, hidrofobne interakcije).

Višestrukim poravnanjem se mogu uočiti *konsenzus sekvence* ili linearni motivi (konzervativne ili ultrakonzervativne), koje sadrže identične AK ili AK sa istim fizičko-hemijskim karakteristikama (npr. hidrofobne AK). *Sekvencioni motivi* su kratke konsenzus sekvence koje su zajedničke za veliki broj proteina, nezavisno od stepena njihove homologije. *Kratki linearni motivi*, ili minimotivi („Short Linear Motifs – SLiMs“, kod eukariota „Eukaryotic Linear Motif – ELM“), učestvuju



Slika 3.11: Trans i cis konformacije peptidne veze



Slika 3.12: Ramačandranov dijagram

u protein–protein interakcijama (npr. adresiranje („targeting“) – signal za lokalizaciju u jedru, različitim modifikacijama; motiv za N-glikozilaciju je NP[ST]P, vezivanje liganada, kao i mestima koja prepoznaju endopeptidaze itd).

Višestrukim poravnanjem (tri ili više) proteinskih sekvenci (primarnih struktura) možemo odrediti stepen sličnosti pojedinih proteina (slika 3.13). Sekvence se označavaju kao *homologne* ako potiču od zajedničkog pretka. Postoji nekoliko vrsta homolognih sekvenci:

- *ortologne*: proteinske sekvence u različitim vrstama koje dele zajedničkog pretka i zadržavaju istu funkciju;
- *paralogne*: proteinske sekvence u istoj vrsti, nastale putem duplikacije i mogu imati različite funkcije;
- *ksenologne*: proteinske sekvence u različitim vrstama, nastale kao proizvod horizontalnog genskog transfera, najčešće imaju istu funkciju.

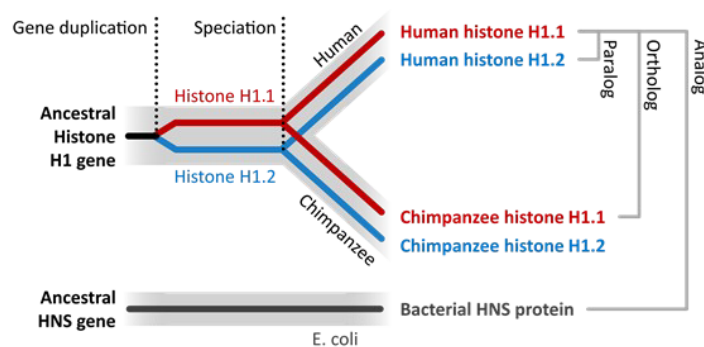
Analogne proteinske sekvence u različitim vrstama imaju različito evolutivno porijeklo, razlikuju se, ali imaju sličnu funkciju.

Slika 3.14 ilustruje evolutivne odnose između proteinskih sekvenci kroz procese duplikacije gena i specijacije. Duplikacijom gena nastaju paralogni proteini unutar iste vrste, koji mogu divergovati u funkciji, dok specijacijom nastaju ortologni proteini u različitim vrstama koji najčešće zadržavaju istu funkciju. Nasuprot tome, analogni proteini nemaju zajedničko evolutivno porijeklo, ali obavljaju slične funkcije kao rezultat konvergentne evolucije.

AK ponovci (eng. protein repeats) je pojava kada se jedan sekvencioni motiv pojavljuje u

Q5E940_BOVIN	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_HUMAN	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_MOUSE	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_RAT	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_CHICK	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_BANSY	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
Q7ZUG3_BRASEY	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_ICTPU	-----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_DROME	-----MVRENKAANKAQYFKVVFLEDFPKCFIVGADNVGSKOMQIRMSLRGK-AVVLGKNTMMRKAIRGHLENN--PALE	76
RLA0_DICDI	-----MSGAG-SKRKKLFIEKATKLFITDKMIVAEADYVGSLOLQIRKSLRGI-CAVLMGKNTMMRKAIRGHLENN--PALE	75
Q54LPO_DICDI	-----MSGAG-SKRKNVFEKATKLFITDKMIVAEADYVGSLOLQIRKSLRGI-CAVLMGKNTMMRKAIRGHLENN--PALE	75
RLA0_PLAF8	-----MAKLSKQKKQMYIEKLSLTIQQYSKILIVHYDVGSDNOMASVYKSLRGL-AILMGKNTMMRKAIRGHLENN--PALE	76
RLA0_SULAC	-----MIGLAVTTTKKIAKWKVDEVAELTKLTKHTIIIIANIEGFPADKLHEIRKSLRGL-ADIKVYKNTLFIKAAKNAAG----DLYK	79
RLA0_SULTO	-----MRIMAVITQERKIAKWKIEEYKLELQKREYHTIIIIANIEGFPADKLHEIRKSLRGL-ADIKVYKNTLFIKAAKNAAG----DLYS	80
RLA0_SULSO	-----MKRLALALKQKRVASWKLIEYKELTELKKSNTLILGIEGFPADKLHEIRKSLRGL-ADIKVYKNTLFIKAAKNAAG----DLYI	80
RLA0_AERPE	MSVVSIVGQMYKREKPIPEWKTLMLELEELFSKRRVVLADITGDFVYVYRVRKKLWKKYDMMVAKKRIILRAMKAAGLE----LDDN	86
RLA0_PYRAE	MMLAIGKRRYVTRQIPARKVKIVSEATELLQKYPYVFLFDLHGLSRIILHEVRYRLRRY-GVTKIIPFLFKIAFTKVVYGG----IPAE	85
RLA0_METAC	MAEERHHTHEIPQWKDEIENIKELIQSHKVFQGVVTEGILATKMKIRRDLDKDV-AVLKVRNTLIERALNQLG----ETIP	78
RLA0_METMA	MAEERHHTHEIPQWKDEIENIKELIQSHKVFQGVVTEGILATKMKIRRDLDKDV-AVLKVRNTLIERALNQLG----ESIP	78
RLA0_ARCFU	MAAVRGS-----PPEYKVRAVEEIKRMISSKVVVAIVSFRNVPAQOMOKIRREFRGK-AEIKVYKNTLIERALDALG----GDYL	75
RLA0_METKA	MAVKAKGQPPSSGYPKVAEKKRREVEKELKELMDEYENGLVLDLGGIPAPLOEIRAKLRERDITIRMSRNTLMRIKALEKLEDER--PELE	88
RLA0_METH	MAHVAEKKKKEVQELHDLKGVVYVGIANLADIPAROLOKMRQTLRDS-ALIRMSKNTLISLALAKAGREL--ENVD	74
RLA0_METTL	MITAESEHKIAPWKIEEVNKLKELKNGQIVALVDHMEVPAPOLOEIRDKIR-QTMLKMSRNTLIERAIKEVAEETGNPEFA	82
RLA0_METVA	MIDAKSEHKIAPWKIEEVNKLKELKNSANVIALVDHMEVPAPOLOEIRDKIR-QTMLKMSRNTLIERAIKEVAEETGNPEFA	82
RLA0_METJA	METKVAHVAPWKIEEVNKLKELKNSANVIALVDHMEVPAPOLOEIRDKIR-DKVKLRMSRNTLIERAIKEVAEELNPNKLA	81
RLA0_PYRAB	MAHVAEKKKKEVEELANLKSVPVIALVDVSSMPAYPLSQMRRLIRENGLLRVSNTLIELAIKKAAKELGKPELE	77
RLA0_PYRHO	MAHVAEKKKKEVEELAKLKSVPVIALVDVSSMPAYPLSQMRRLIRENGLLRVSNTLIELAIKKAAKELGKPELE	77
RLA0_PYRFU	MAHVAEKKKKEVEELANLKSVPVIALVDVSSMPAYPLSQMRRLIRENGLLRVSNTLIELAIKKAAKELGKPELE	77
RLA0_PYRKO	MAHVAEKKKKEVEELANLKSVPVIALVDVAGVPAYPLSKMRDKLR-GKALLRVSNTLIELAIKKAAKELGQPELE	76
RLA0_HALMA	MSEESERKTEIENKQEEVDVAVIEMIESYSEVGVVNIAGIPRDLQDMRRDLHGT-AELRVSNTLIERALDDVD----DGLF	79
RLA0_HALVO	MSESEVRQTEVPEWKKRQEVDELVDLIESYSEVGVVNIAGIPRDLQDMRRDLHGT-AAYRMSRNTLIERALDEVN----DGLF	79
RLA0_HALSA	MSEEQRTTEVPEWKKRQEVDELVDLIESYSEVGVVNIAGIPRDLQDMRRDLHGT-AALRMSRNTLIERALDEAG----DGLD	79
RLA0_THEAC	MKEVSQKKELVNEITDRIKASRSVAIVDAGIRRDLQDMRRDLHGT-INLKVYKNTLIERALDGLD----EKLS	72
RLA0_THEVO	MRKINPKKKEIVSELAADITKSKAVAVDIKGVRELDQDMRRDLHGT-VKIKVYKNTLIERALDGLD----EKLT	72
RLA0_PICTO	MTEPAQWKIDFVKNLENEINRSKVAIVSIKGLRNNFQKIRNSIRDK-ARIKVSARLLRLAIENYK----NNIV	72
tuIuler	1.....10.....20.....30.....40.....50.....60.....70.....80.....90	

Slika 3.13: Višestruko poravnanje proteinskih sekvenci



Slika 3.14: Homologe sekvenci

proteinskoj sekvenci dva ili više puta. Sekvencioni motiv – ponovak može biti monoaminokiselinski (npr.: (Gln) n , (Asn) n , (Ser) n ; (Gln) n ponovci su povezani sa nekim patološkim stanjima). Sekvencioni motivi se često reflektuju i na nivou sekundarne i 3D strukture. Ponovci se mogu pojaviti u tandemu (jedan do drugog, oko 14% proteina u proteomu), ili mogu biti rasuti po proteinskoj sekvenci; takođe mogu biti direktni ili inverzni. Oko polovine tandem ponovaka (posebno oni do 10 AK) imaju neuređenu strukturu, dok duži ponovci (30–40 AK) obično formiraju neku sekundarnu strukturu.

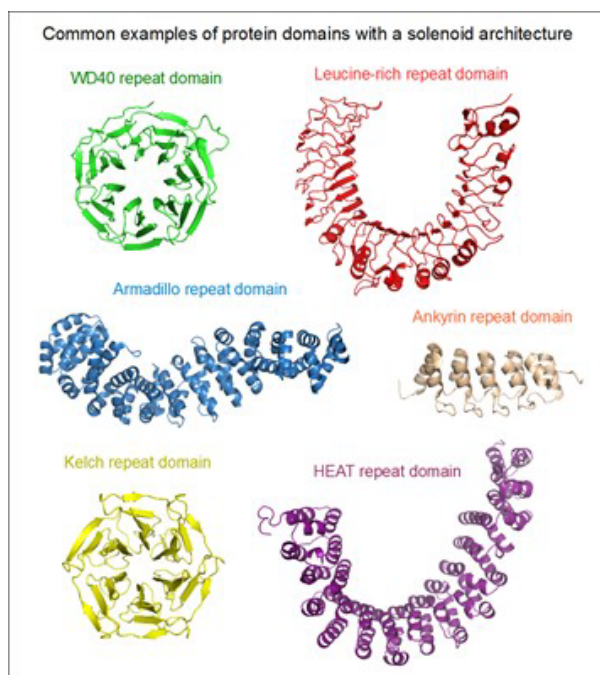
Na slici 3.15 je prikazano višestruko poravnanje pentapeptida sa nepotpunim A-L ponovkom. Ovakav ponovak može rezultirati u tandem ponovcima istih sekundarnih struktura (slika 3.16).

Proteinske familije

Proteinske familije su grupe proteina koje dele zajedničko evoluciono poreklo, odnosno homologne su, i imaju značajnu sličnost u primarnim (konsenzus sekvence) i 3D strukturama. Trenutno je definisano preko 60.000 proteinskih familija. Na osnovu višestrukog poravnanja primarnih

HglK	78	117	CDIRFANLEGADLTDANLAAASHKSNLRRANLCRATLNR
HglK	118	157	CNISEADLTESDANEALFCQAVFTEVEAHGLRLYRAKVSQ
HglK	158	197	AQLMGAHLHQAYAPEADFSAVAIAVDLRWANLRKTNFRG
HglK	203	242	GNFRGANLTQADFTGANLKGANLRGANLVGTNLQRADLSD
spkB	454	493	QNLVGVILAKAFVPGINCYQANLTNANFEQAELTRADFGK
spkB	504	543	ANLSDAYFGYADLRGADLRGANLNGVNFKYANLQGANFSG
slr1819	20	59	LKLPGINLEAADLIGIVLNEADLRGANLFCYLNRRANLGGQ
slr1819	60	99	ANLVAANLSGASLNQADLAGADLRSANFHGAMLQGAILRD
slr1819	100	139	SDMTLATLQDTNLIQADLRGADLSGATLTGACLRGANMRQ
slr1819	151	190	AILGRADLQGANMKGVDLSRADLSYANLKEANLRDVDLRK
slr1819	191	230	ADLSYANLKGALLTDANLSGAKLNGADLQANLMRAKISE
slr1819	231	270	AEMTAVNCQGAIMTHVNLNRTNLTGSNLSFTRMNSADLSR
slr1819	271	310	ANITKANLQEAELIEAFFARANLTEANFINANLVRADLMS
consensus			A L A L A L A L A L A L A L A L

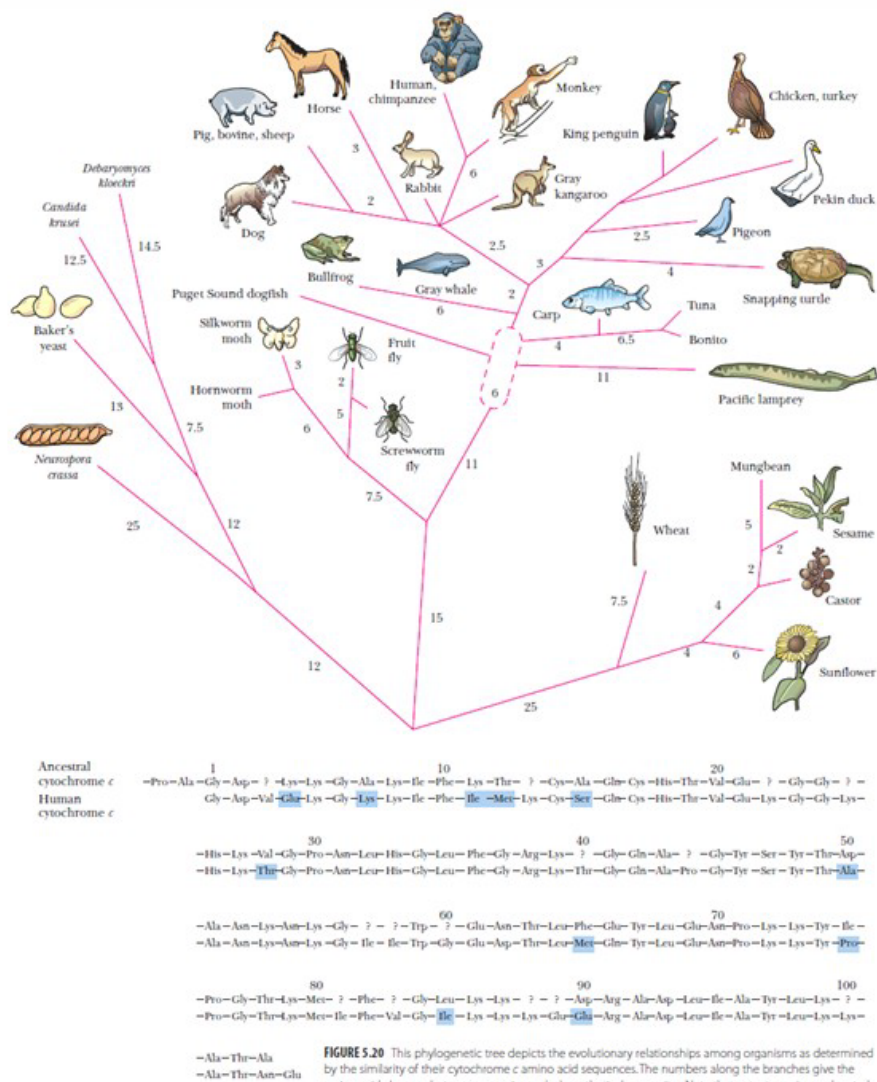
Slika 3.15: Višestruko poravnanje



Slika 3.16: Tandem ponovci istih sekundarnih struktura (domena).

struktura formirana je Pfam baza podataka proteinskih familija (<http://pfam.xfam.org/>).

Poređenjem primarnih struktura sličnih proteina iz različitih organizama moguće je odrediti njihove filogenetske relacije (na primeru citohroma c). Citohrom c je često korišćen protein za poređenje primarnih struktura između različitih organizama, jer se nalazi kod gotovo svih eukariota i ima sličnu funkciju kod svih vrsta, te njegova sekvencija omogućava procenu evolutivne srodnosti i rekonstrukciju filogenetskih odnosa. Filogenetsko stablo na slici 3.17 prikazuje evolutivne odnose među organizmima, određene sličnošću njihovih sekvenci citohroma c. Brojevi duž grana predstavljaju broj razlika u aminokiselinama između vrsta i hipotetičkog pretka. Postojeće vrste prikazane su samo na krajevima grana. Ispod je prikazana sekvencija humanog citohroma c, koja je upoređena sa inferisanom sekvencijom predstavljenom na osnovi stabla korelacija. (Adaptirano prema Creighton, T.E., 1983. *Proteins: Structure and Molecular Properties*. San Francisco: W.H. Freeman.)



Slika 3.17: Filogenetsko stablo napravljeno na osnovu sličnosti sekvenci citohroma *c* organizama

3.2.2 Sekundarna struktura proteina

Sekundarna struktura proteina je definisana φ i ψ diedralnim uglovima polipeptidnog lanca. To je lokalna, regularna, podstruktura, koja nastaje kao rezultat formiranja vodoničnih veza između CO i NH grupa polipeptidnog lanca.

Dva osnovna tipa sekundarnih struktura su zavojnica (helikoida) i trakasta struktura. Zavisno od broja zakokreta razlikujemo tri osnovna tipa helikoida (najčešće desnih): α -zavojnica (α -helix), 3_{10} zavojnica (3_{10} -helix), π zavojnica (π -helix) i PPII zavojnica (PPII-helix, levi). Postoje dve glavne trakaste strukture: paralelna i antiparalelna β -struktura (β -strand).

Tabela 3.10 prikazuje strukturne parametre sekundarnih struktura proteina zavojnica i trakastih struktura. Za svaku strukturu dati su diedralni uglovi (φ , ψ), broj aminokiselinskih ostataka po jednom zavoju, translacija po ostatku, kao i karakteristike vodoničnih veza. Helikoidne strukture (α , 3_{10} , π i PPII) razlikuju se po geometriji i načinu formiranja vodoničnih veza između ostataka (npr. α -helix ima stabilnu vezu tipa $i \rightarrow i + 4$). Većina heliksa je desnoruka, osim PPII

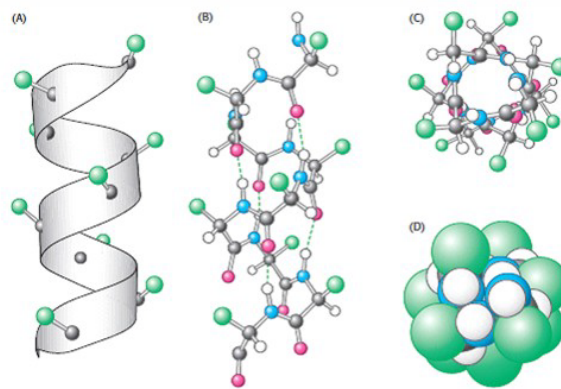
heliksa koji je levoruk. β -strukture (paralelne i antiparalelne) nemaju definisan broj ostataka po zavoju, već formiraju proširene lance povezane vodoničnim vezama između različitih lanaca (eng. inter-strand). Razlike u diedralnim uglovima određuju njihovu prostornu organizaciju i stabilnost.

Tip elementa	Element	Hiralnost	φ, ψ	Broj AK po zavoju	Pomeraj po AK	Poluprečnik	H-veze
Zavojnica	α	Desna	$-57^\circ, -47^\circ$	3.6	1.5 Å	2.3 Å	$i \rightarrow i + 4$
	3_{10}	Desna	$-49^\circ, -26^\circ$	3.0	2.0 Å	1.9 Å	$i \rightarrow i + 3$
	π	Desna	$-57^\circ, -70^\circ$	4.4	1.1 Å	2.8 Å	$i \rightarrow i + 5$
	PPII	Leva	$-75^\circ, 145^\circ$	3.0	3.1 Å	1.4 Å	
β -lanac	Paralelan	–	$-119^\circ, +113^\circ$	–	3.4 Å		Medulančane
	Antiparalelan	–	$-139^\circ, +135^\circ$	–	3.2 Å		Medulančane

Tabela 3.10: Karakteristike najčešćih sekundarnih struktura proteina. Izvor: AMIT KESSEL, NIR BEN-TAL Introduction to Proteins Structure, Function, and Motion, 2e

Slika 3.18 prikazuje različite načine reprezentacije α -zavojnice (α -helix):

- (A) Prikazana je tzv. *ribbon* (trakasta) reprezentacija, gde se jasno vidi spiralna (helikoidna) struktura polipeptidnog lanca. Bočne grupe (R) su prikazane kao sfere koje izlaze spolja iz zavojnice.
- (B) Model *kuglica–štapić* (eng. ball-and-stick) prikazuje detaljnu hemijsku strukturu, uključujući atome i veze. Isprekidane linije predstavljaju vodonične veze između NH i CO grupa koje stabilizuju α -zavojnicu.
- (C) Pogled duž ose zavojnice (eng. end view) pokazuje da je struktura cilindrična, sa centralnim kičmenim delom (eng. backbone) u unutrašnjosti, dok su bočni lanci orijentisani ka spolja.
- (D) *Space-filling model* (model ispunjen prostorom) ilustruje gusto pakovanje atoma u unutrašnjosti zavojnice, naglašavajući kompaktnost strukture i način na koji atomi zauzimaju prostor.

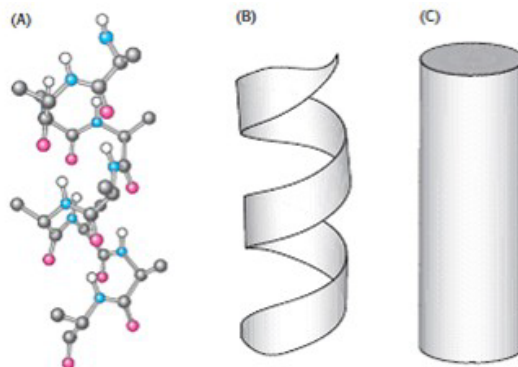


Slika 3.18: Struktura α -zavojnice

I slika 3.19 prikazuje različita načina reprezentacije α -zavojnice:

- (A) Model *kuglica–štapić*;
- (B) trakasta reprezentacija;

(C) Idealizovani geometrijski model u obliku cilindra predstavlja uprošćen prikaz α -zavojnice, gde je naglašena njena pravilna, cilindrična struktura.



Slika 3.19: Struktura zavojnice

Tabela 3.11 prikazuje osnovne geometrijske karakteristike tri glavna tipa zavojnice. Za svaki tip dati su:

- broj ostataka po jednom zavoju - broj aminokiselina potrebnih da se polipeptidni lanac obrne za jedan pun krug (360°) oko ose heliksa;
- translacija po ostatku - pomeraj duž ose heliksa koji nastaje kada se doda jedna aminokiselina (jedan ostatak);
- poluprečnik zavojnice - rastojanje od ose heliksa do atoma (ili kičme polipeptidnog lanca);
- korak zavojnice (eng. pitch) - rastojanje duž ose heliksa koje odgovara jednom punom zavoju spiralne strukture.

$$\text{korak zavojnice} = (\text{broj ostataka po zavoju}) \times (\text{translacija po ostatku})$$

α -zavojnica predstavlja najčešći i najstabilniji oblik, dok 3_{10} - i π -zavojnice imaju specifične geometrijske osobine i ređe su zastupljene u proteinima.

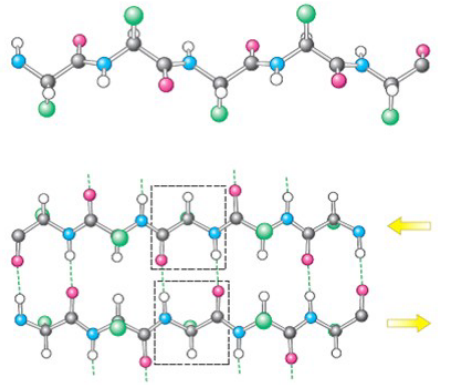
Geometrijski atribut	α -zavojnica	3_{10} -zavojnica	π -zavojnica
Broj ostataka po zavoju	3.6	3.0	4.4
Translacija po ostatku	1.5 Å (0.15 nm)	2.0 Å (0.20 nm)	1.1 Å (0.11 nm)
Poluprečnik zavojnice	2.3 Å (0.23 nm)	1.9 Å (0.19 nm)	2.8 Å (0.28 nm)
Korak zavojnice	5.4 Å (0.54 nm)	6.0 Å (0.60 nm)	4.8 Å (0.48 nm)

Tabela 3.11: Strukturne karakteristike tri glavna oblika proteinskih heliksa

Slika 3.20 prikazuje trakaste sekundarne strukture proteina, odnosno β -lanac (eng. β -strand) i β -ploču (eng. β -sheet):

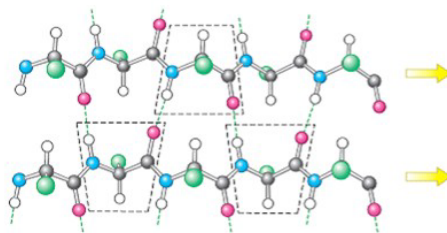
(A) Prikazan je pojedinačni β -lanac, gde se vidi izdužena (razvučena) konformacija polipeptidnog lanca. Bočni lanci (označeni zelenom bojom) naizmenično su usmereni iznad i ispod ravnog lanca, što je karakteristično za β -strukture.

- (B) Prikazana je β -ploča formirana povezivanjem više β -lanaca. U ovom slučaju radi se o antiparalelnoj β -ploči, gde susedni lanci imaju suprotan smer (N- i C-krajevi su orijentisani u različitim pravcima). Vodonične veze (isprekidane linije) formiraju se između NH i CO grupa susednih lanaca, pri čemu svaka aminokiselina iz jednog lanca može ostvariti interakciju sa odgovarajućom aminokiselinom iz susednog lanca, čime se stabilizuje struktura.



Slika 3.20: Trakaste strukture

Slika 3.21 prikazuje dva β -lanca koji su orijentisani u istom smeru (paralelno), što znači da su njihovi N- i C-krajevi usmereni u istom pravcu. Polipeptidni lanci imaju izduženu konformaciju, a bočni lanci (radikali, R) su naizmenično raspoređeni iznad i ispod ravni svakog β -lanca. Vodonične veze (isprekidane linije) formiraju se između NH i CO grupa susednih lanaca, stabilizujući strukturu. Za razliku od antiparalelne β -ploče, gde se aminokiseline uparuju jedna prema jednoj, u paralelnoj β -ploči jedna aminokiselina iz jednog lanca ostvaruje vodonične veze sa dve različite aminokiseline iz susednog lanca. Ovakav način povezivanja dovodi do nešto manje optimalne geometrije vodoničnih veza u odnosu na antiparalelnu β -ploču, ali i dalje omogućava stabilnu strukturu.



Slika 3.21: Paralelna β -ploča

Okreti (eng. turns) i petlje (eng. loops) predstavljaju neregularne elemente sekundarne strukture proteina (eng. coil) koji omogućavaju promenu pravca polipeptidnog lanca i povezivanje različitih sekundarnih strukturnih elemenata.

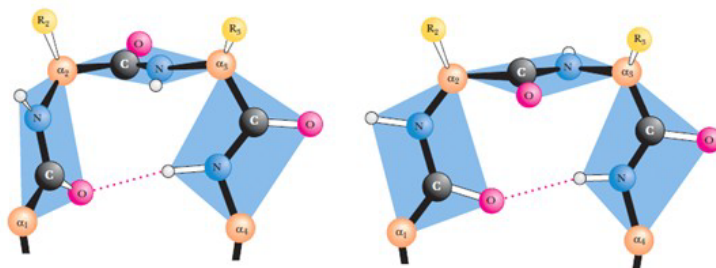
Okreti su kratki segmenti polipeptidnog lanca koji dovode do naglog savijanja strukture. Najčešći su β -okreti (eng. β -turns), koji obuhvataju četiri uzastopna aminokiselinska ostatka ($i, i + 1, i + 2, i + 3$). Karakteristična osobina β -okreta je postojanje vodonične veze između

CO grupe ostatka i i NH grupe ostatka $i + 3$, što stabilizuje ovu konformaciju. U zavisnosti od vrednosti diedralnih uglova razlikuju se različiti tipovi β -okreta.

Petlje su duži i fleksibilniji segmenti polipeptidnog lanca koji nemaju regularnu, ponavljajuću strukturu. Za razliku od okreta, petlje mogu obuhvatati različit broj aminokiselina i često nemaju stabilizujuće vodonične veze karakteristične za α -zavojnice ili β -ploče.

Okreti i petlje su ključni za pravilno savijanje proteina i formiranje njegove trodimenzionalne strukture.

Slika 3.22 prikazuje strukture dva tipa β -okreta (eng. β -turn), koji se nazivaju i čvrsti okreti ili β -savijanja (eng. β -bend). Prikazana su dva osnovna tipa β -okreta (Tip I i Tip II), koji omogućavaju naglu promenu pravca polipeptidnog lanca. β -okret obuhvata četiri uzastopna aminokiselinska ostatka ($i, i + 1, i + 2, i + 3$), pri čemu dolazi do savijanja lanca tako da se početni i krajnji deo približavaju. Stabilizacija strukture ostvaruje se vodoničnom vezom (isprekidana linija) između CO grupe ostatka i i NH grupe ostatka $i + 3$. Razlika između Tipa I i Tipa II potiče od različitih vrednosti diedralnih uglova centralnih ostataka ($i + 1$ i $i + 2$), što dovodi do različite prostorne konfiguracije. β -okreti su česti u proteinima i često se nalaze na površini molekula, gde omogućavaju kompaktno pakovanje i povezivanje sekundarnih strukturnih elemenata.



Slika 3.22: Strukture dva tipa β -okreta

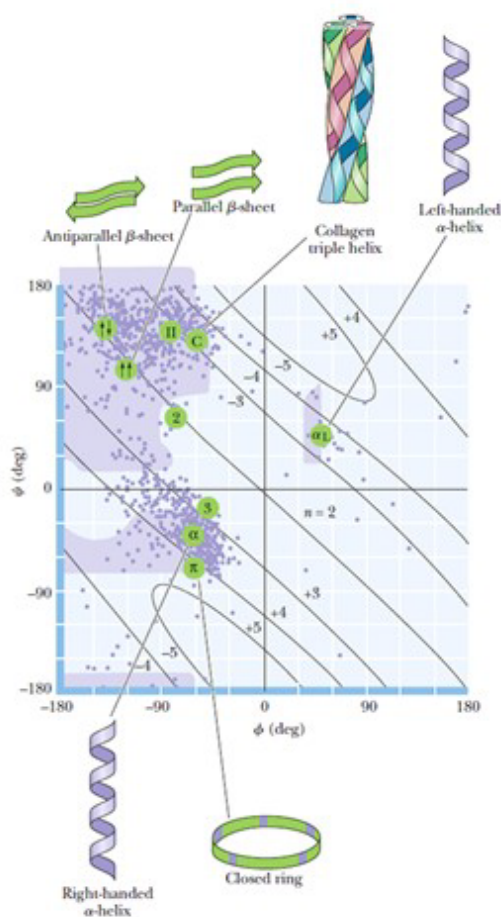
Ramačandranov dijagram na slici 3.23 prikazuje kombinacije uglova φ (osa x) i ψ (osa y), koje određuju konformaciju polipeptidnog lanca. Zasenčene oblasti označavaju dozvoljene (energetski povoljne) regione konformacionog prostora, dok su nezasenčene oblasti sterički nepovoljne zbog sudara atoma. Različite regije na dijagramu odgovaraju tipičnim sekundarnim strukturama:

- β -ploče (paralelne i antiparalelne) nalaze se u gornjem levom delu dijagrama,
- desnoruka α -zavojnica nalazi se u donjem levom delu,
- levoruka α -zavojnica nalazi se u gornjem desnom delu,
- kolagenska trostruka zavojnica zauzima specifičan region dijagrama.

Tačke na dijagramu predstavljaju eksperimentalno određene vrednosti uglova za različite aminokiseline u proteinima. Ramačandranov dijagram je ključan alat za analizu i validaciju strukture proteina, jer omogućava identifikaciju stabilnih i nestabilnih konformacija.

3.2.3 Aminokiseline i sekundarne strukture

Aminokiseline imaju različite sklonosti ka formiranju određenih sekundarnih struktura, što je posledica njihovih fizičko-hemijskih osobina.



Slika 3.23: Okreti i petlje

Relativne učestalosti aminokiselina u sekundarnim strukturama su prikazane u tabeli 3.12. Boje označavaju dominantnu sklonost aminokiselina ka određenoj sekundarnoj strukturi. Vidljivo je da hidrofobne aminokiseline (npr. Val, Ile, Phe) favorizuju β -ploče, dok fleksibilne i male aminokiseline (npr. Gly, Pro) preferiraju okrete. α -zavojnice su stabilizovane aminokiselinama koje omogućavaju formiranje povoljnih vodoničnih veza bez steričkih prepreka.

Slika 3.24 prikazuje raspodelu dužina segmenata sekundarnih struktura proteina za α -zavojnice, β -ploče i neuređene regione. Na x-osi prikazana je dužina segmenata sekundarne strukture (broj aminokiselinskih ostataka), dok y-osa predstavlja frekvenciju pojavljivanja segmenata određene dužine. Uočava se da su coil i β -segmenti najčešće kratki, sa najvećom frekvencijom pri malim dužinama (oko 2–6 ostataka), nakon čega frekvencija naglo opada. α -zavojnice imaju širu raspodelu i često dostižu veće dužine, što ukazuje na njihovu veću stabilnost i sposobnost formiranja dužih kontinuiranih struktura. Za sve tipove struktura frekvencija opada sa porastom dužine segmenata, što znači da su veoma dugi segmenti ređi u proteinskim strukturama.

Sekundarne strukture se mogu iz podataka sa PDB-a anotirati pomoću *Dictionary of Protein Secondary Structure* (DSSP) koda (Wolfgang Kabsch, Christian Sander). Tipovi sekundarnih struktura su:

Aminokiselina	α -zavojnica	β -ploča	Okret
Ala	1.29	0.90	0.78
Cys	1.11	0.74	0.80
Leu	1.30	1.02	0.59
Met	1.47	0.97	0.39
Glu	1.44	0.75	1.00
Gln	1.27	0.80	0.97
His	1.22	1.08	0.69
Lys	1.23	0.77	0.96
Val	0.91	1.49	0.47
Ile	0.97	1.45	0.51
Phe	1.07	1.32	0.58
Tyr	0.72	1.25	1.05
Trp	0.99	1.14	0.75
Thr	0.82	1.21	1.03
Gly	0.56	0.92	1.64
Ser	0.82	0.95	1.33
Asn	0.94	0.72	1.41
Asp	0.90	0.76	1.28
Pro	0.52	0.64	1.91
Arg	0.96	0.99	0.88

Tabela 3.12: Relativne učestalosti aminokiselina u sekundarnim strukturama (istaknute dominantne sklonosti)

- G = 3-turn helix (3_{10} helix). Min. dužina 3 ostatka.
- H = 4-turn helix (α -helix). Min. dužina 4 ostatka.
- I = 5-turn helix (π -helix). Min. dužina 5 ostataka.
- T = hydrogen bonded turn (3, 4 ili 5 okreta).
- E = extended strand u paralelnoj i/ili antiparalelnoj β -ploči. Min. dužina 2 ostatka.
- B = ostatak u izolovanom β -mostu (jedan par vodoničnih veza u β -strukturi).
- S = savijanje (*bend*) (jedina oznaka koja nije zasnovana na vodoničnim vezama).

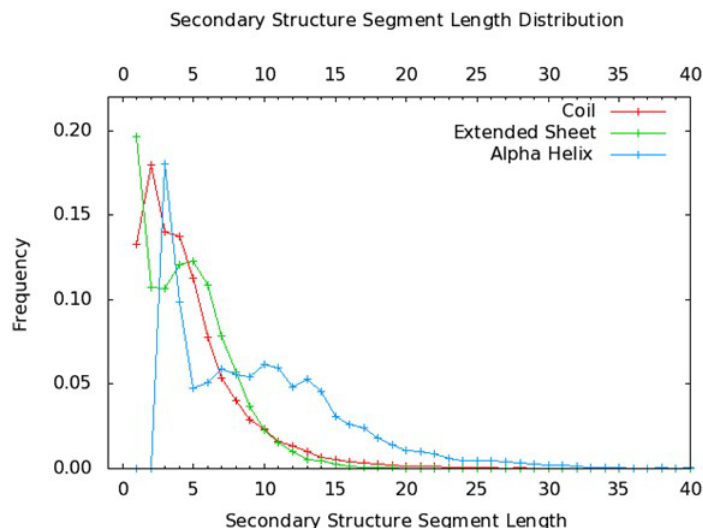
3.2.4 Lokalna proteinska struktura – strukturalni alfabet

Strukturalni alfabet je pokušaj da se 3D struktura proteina svede na kombinaciju ograničenog broja malih lokalnih struktura nazvanih prototipovi i može se upotrebiti za pojednostavljenje i analizu proteinskih struktura, kao i za predikciju strukture polipeptidnog lanca.

Protein blok (PB) (slika 3.25) je pokušaj da se 3D struktura proteina prikaže pomoću 16 prototipova dužine 5 aminokiselina, koji najbolje aproksimiraju pojedine proteinske 3D strukture.

3.2.5 Supersekundarna struktura, motiv, domen

Jednostavni strukturalni elementi proteina se postepeno organizuju u sve složenije nivoe strukture, što je ključno za funkciju proteina. Slika 3.26 prikazuje organizaciju proteinske



Slika 3.24: Dužine sekundarnih struktura

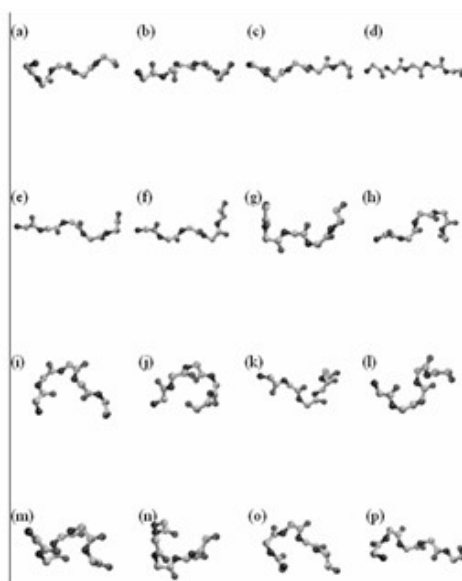
strukture, od jednostavnijih ka složenijim nivoima:

- **Sekundarna struktura** predstavlja osnovne lokalne elemente.
- **Motiv** predstavlja kombinaciju nekoliko sekundarnih struktura koje formiraju karakterističan i često ponavljajući obrazac u proteinima.
- **Domen** je veća, stabilna i funkcionalna celina proteina koja se sastoji od više motiva i može se savijati nezavisno od ostatka molekula.
- **Tercijarna struktura** predstavlja kompletnu trodimenzionalnu strukturu jednog polipeptidnog lanca, nastalu međusobnim interakcijama svih njegovih delova.
- **Kvaternerna struktura** opisuje način na koji se više polipeptidnih lanaca (subjedinica) organizuje i međusobno povezuje u funkcionalni proteinski kompleks.

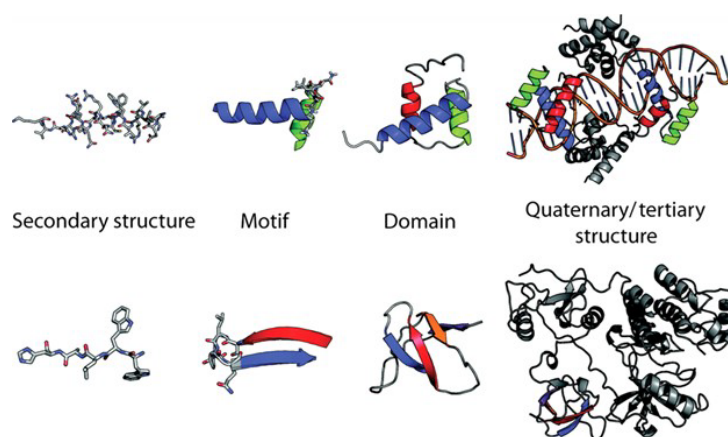
Supersekundarna struktura je kompaktna 3D kombinacija nekoliko susednih sekundarnih struktura ili motiva (npr.: helikoida–okret–helikoida), koja je manja od proteinskog domena ili subjedinice. Deluje kao centar nukleacije (mesto odakle počinje savijanje proteina) prilikom formiranja 3D strukture proteina (eng. protein folding). Helikoidne supersekundarne strukture su: heliks–ukosnica (dve α -zavojnice povezane kratkim okretom (ukosnicom)), heliks–ugao (dve α -zavojnice povezane okretom pod određenim uglom) itd. Trakaste superstrukture su: β -ukosnica (dva β -lanca povezana kratkim okretom), β -ugao (povezivanje β -lanaca pod određenim uglom) itd. Mešovite: β - α - β motiv, itd.

Strukturalni motiv je kratak segment proteinske strukture (jedna ili više sekundarnih struktura), koji je zajednički za veći broj proteina ili se pojavljuje više puta u datom proteinu. Motivi su često povezani sa određenom funkcijom ili evolutivnim poreklom. Na primer, motivi kao što su β -meandar ili „grčki ključ“ opisuju način povezivanja β -lanaca, dok neki motivi, poput cink-prsta (Zn prst), imaju i specifičnu ulogu, kao što je vezivanje DNK.

Slika 3.27 prikazuje hijerarhijsku organizaciju proteinske strukture i prelaz od jednostavnih ka složenijim strukturnim nivoima. Na nivou sekundarne strukture prikazani su osnovni elementi:



Slika 3.25: 16 protein blokova obeleženi malim latiničnim slovima a–p.



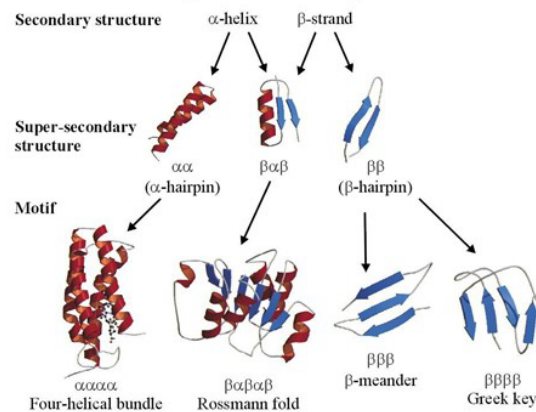
Slika 3.26: Organizacija proteinske strukture po nivoima

α -zavojnica i β -lanac, koji predstavljaju lokalne konformacije polipeptidnog lanca. Na nivou supersekundarne strukture prikazane su jednostavne kombinacije ovih elemenata, kao što su $\alpha\alpha$ (heliks–ukosnica), $\beta\alpha\beta$ i $\beta\beta$ (β -ukosnica). Ove kombinacije dalje formiraju strukturalne motive, koji su stabilni i često ponavljajući obrasci u proteinima.

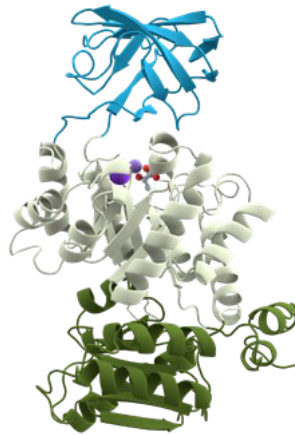
Strukturni domeni (moduli) su veći, relativno stabilni i kompaktni segmenti proteinske 3D strukture, koji se nezavisno formiraju od ostatka proteina i sastoje se od 50 do 250 aminokiselina. Neki domen može da se pojavi u više različitih proteina (i u više kopija), kao posledica splajsovanja (alternativnog spajanja) mRNA. Spajanje mRNA je proces uklanjanja introna i povezivanja eksona u zrelu mRNA, dok alternativno spajanje (isti gen može dati više različitih mRNA, jer neki eksoni mogu biti uključeni ili izostavljeni) omogućava nastanak različitih proteina iz istog gena. Proteini eukarija se obično sastoje od više strukturno-funkcionalnih domena: receptorski domen (prepoznaje i vezuje signalnu molekulu), transmembranski domen (omogućava proteinu da bude usidren u ćelijskoj membrani), katalitički domen (obavlja hemijsku reakciju).

Enzim piruvat kinaza (PDB: 1PKN) prikazan na slici 3.28 se sastoji od 3 domena koji su različito obojeni.

Hierarchical organization of protein structure



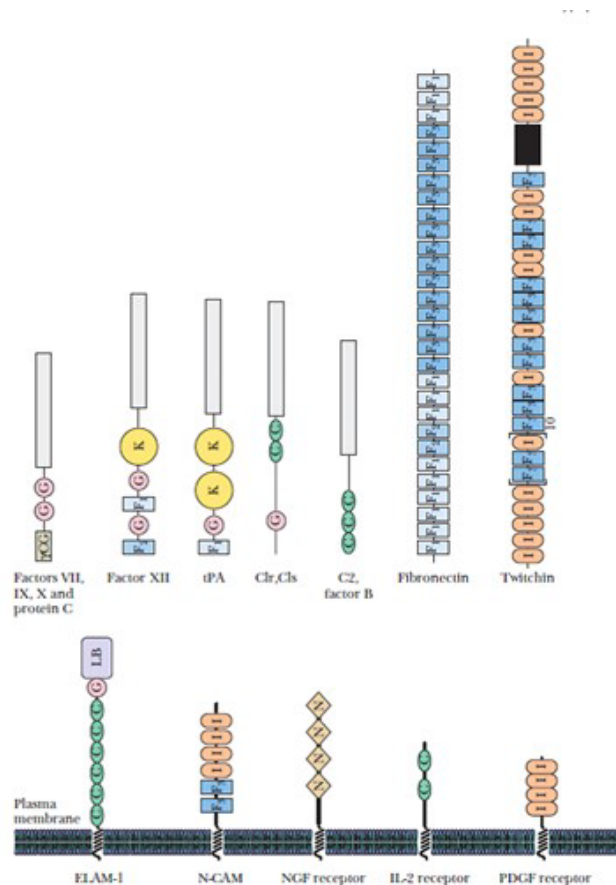
Slika 3.27: Hijerarhijska organizacija proteinske strukture



Slika 3.28: Enzim piruvat kinaza (PDB: 1PKN)

Slika 3.29 prikazuje modularnu organizaciju proteina i način na koji se različiti strukturni domeni kombinuju u složene proteinske molekule. U gornjem delu prikazani su različiti proteini (npr. faktori koagulacije, fibronektin, titin), sastavljeni od ponavljajućih modula (domena), koji su označeni različitim oblicima i bojama. Svaki modul predstavlja specifičan strukturni domen sa određenom funkcijom (npr. vezivanje jona kalcijuma, interakcija sa drugim proteinima ili molekulima), a isti tip domena može se pojavljivati više puta unutar jednog proteina. Posebno je istaknuta modularnost velikih proteina, kao što su fibronektin i titin, koji sadrže veliki broj ponavljajućih domena organizovanih u linearne nizove. U donjem delu slike prikazani su membranski proteini (npr. ICAM-1, NCAM, NGF receptor, IL-2 receptor, PDGF receptor), koji se sastoje od ekstracelularnih domena, transmembranskog segmenta i intracelularnog dela.

Slika 3.30 prikazuje strukturu trostruke helikoide, karakteristične za fibrilarne proteine (proteini koji imaju izdužen, vlaknast oblik i prvenstveno imaju strukturnu/mehaničku ulogu u organizmu) kao što je kolagen. Prikazane su tri polipeptidne niti koje su međusobno isprepletene i formiraju stabilnu strukturu. Trostruka helikoide omogućava visoku mehaničku čvrstoću i otpornost na istezanje, što je ključno za vezivna tkiva.



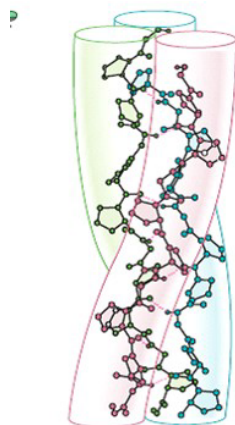
Slika 3.29: Modularnu organizacija proteina

3.2.6 Tercijalna i kvaternarna struktura proteina

Tercijalna proteinska (3D) struktura je posledica međusobnih interakcija bočnih aminokiselinskih lanaca (i prostetičkih grupa – neproteinski deo proteina koji je čvrsto vezan za njega i neophodan za njegovu biološku funkciju, npr. Fe u hemoglobinu) pod fiziološkim uslovima. Tercijalna struktura je definisana položajem pojedinih atoma i atomskih grupa u prostoru.

Slika 3.31 prikazuje različite tipove interakcija koje učestvuju u formiranju tercijalne strukture proteina:

- **Jonske interakcije (ion–ion)** nastaju između suprotno naelektrisanih grupa (npr. NH_3^+ i COO^-) i predstavljaju jedne od najjačih nekovalentnih interakcija, sa energijama u opsegu od 40 do 100 kJ/mol.
- **Jon–dipol interakcije** uključuju interakciju između naelektrisanih grupa i polarnih molekula ili grupa, sa energijama od približno 4 do 40 kJ/mol.
- **Vodonične veze** formiraju se između donora vodonika (npr. NH) i akceptora (npr. C=O) i imaju ključnu ulogu u stabilizaciji sekundarnih i tercijalnih struktura, sa energijama do oko 40 kJ/mol.
- **Hidrofobne interakcije** nastaju između nepolarnih grupa (npr. alkilnih lanaca) koje se grupišu kako bi izbegle kontakt sa vodom, sa energijama u opsegu od 4 do 40 kJ/mol.



Slika 3.30: Fibrilarni proteini – trostruka helikoida (kolagen)

- **Kovalentne veze**, kao što su disulfidni mostovi ($-S-S-$), predstavljaju najjače veze u ovom kontekstu i dodatno stabilizuju strukturu proteina, sa energijama od 80 do 600 kJ/mol.

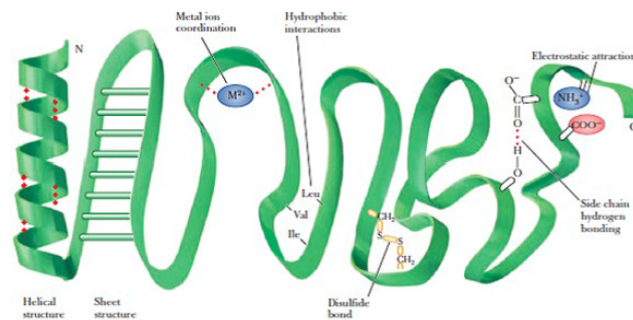
Slika ilustruje da stabilnost proteinske strukture zavisi od kombinacije različitih slabijih i jačih interakcija koje zajedno omogućavaju pravilno savijanje makromolekula.

Ion-ion	$-\text{NH}_3^+ \cdots \text{O}^- \text{C}=\text{O}$	40-100 kJ
Ion-dipole	$\text{C}_6\text{H}_5\text{-OH} \cdots \text{O}^- \text{C}=\text{O}$	4-40 kJ
Hydrogen bonds	$-\text{NH} \cdots \text{O}=\text{C}$	0-40 kJ
Hydrophobic interactions	$-\text{H}_2\text{C}-\text{CH}_3 \cdots \text{CH}_3-\text{CH}_2-$	4-40 kJ
Covalent bonds	$-\text{S}-\text{S}-$	80-600 kJ

Slika 3.31: Neke od interakcija koje omogućavaju korektno formiranje strukture proteina, kao i interakcije između proteina i drugih molekula.

Slika 3.32 prikazuje sile koje stabilizuju tercijalnu strukturu proteina. Na levoj strani prikazane su sekundarne strukture, α -zavojnica i β -ploča, koje su stabilizovane vodoničnim vezama duž kičme polipeptidnog lanca. Tercijalna struktura nastaje daljim savijanjem lanca, pri čemu ključnu ulogu imaju interakcije između bočnih lanaca aminokiselina. Koordinacija metalnog jona (npr. M^{2+}) omogućava stabilizaciju strukture vezivanjem za specifične aminokiseline. Hidrofobne interakcije dovode do grupisanja nepolarnih aminokiselina (npr. Leu, Val, Ile) u unutrašnjosti proteina, daleko od vodene sredine. Disulfidne veze ($-S-S-$) predstavljaju kovalentne veze između cisteinskih ostataka i značajno doprinose stabilnosti strukture. Elektrostatičke interakcije (jonske veze) nastaju između suprotno naelektrisanih grupa, kao što su NH_3^+ i COO^- . Vodonične

veze između bočnih lanaca dodatno stabilizuju prostornu organizaciju proteina.



Slika 3.32: Sile koje stabilizuju tercijalnu strukturu proteina

3.2.7 Protein Data Bank (PDB)

Protein Data Bank (PDB), <http://www.rcsb.org/>, je javno dostupna baza podataka koja sadrži eksperimentalno određene trodimenzionalne strukture proteina i njihovih kompleksa. Svaka struktura ima jedinstveni identifikacioni kod (PDB ID) i može se koristiti za analizu strukture, funkcije i interakcija biomolekula.

	1	2	3	4	5	6	7	8	Tens Place		
	1	2	3	4	5	6	7	8	Ones Place		
ATOM	171	N	ALA	A	22	-23.530	15.854	23.632	1.00	8.12	N
ATOM	172	CA	ALA	A	22	-22.101	16.127	23.631	1.00	10.49	C
ATOM	173	C	ALA	A	22	-21.714	17.207	24.627	1.00	11.94	C
ATOM	174	O	ALA	A	22	-22.515	18.047	25.045	1.00	13.86	O
ATOM	175	CB	ALA	A	22	-21.623	16.582	22.260	1.00	5.03	C
ATOM	176	N	LEU	A	23	-20.471	17.079	25.047	1.00	12.22	N
.....											

Atom Coordinates

Slika 3.33: Primer koordinata atoma u bazi PDB

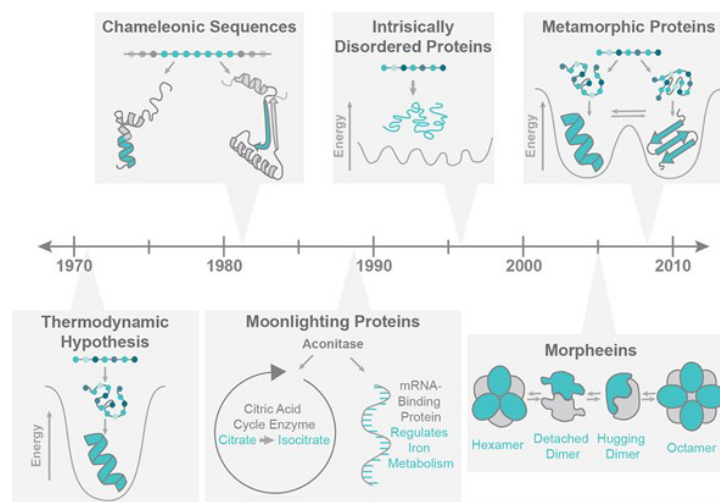
3.2.8 Metamorfni i neuređeni proteini

Nedavno je otkriveno nekoliko grupa proteina koji ne poštuju Anfinsenovo pravilo: „one sequence, one fold“ (jedna sekvenca, jedna struktura). Otkriveni su proteini (proteinski regioni) koji nemaju fiksnu strukturu koji se nazivaju **neuređeni proteini** (*Intrinsically disordered proteins – IDP*), kao i kratke *kameleonske sekvence* koje mogu imati dve (ili više) različite sekundarne strukture (energetski povoljna stanja) i nazivaju se **metamorfni proteini**. Poseban tip metamorfni proteina su morfini – proteini koji mogu formirati dva ili više različitih homo (istih) oligomera.

Slika 3.34 prikazuje razvoj savremenih koncepata u razumevanju proteinske strukture i funkcije, posebno odstupanja od klasičnog principa „jedna sekvenca – jedna struktura“:

- **Kameleonske sekvence** predstavljaju kratke delove proteina koji mogu zauzimati različite sekundarne strukture u zavisnosti od okruženja.
- **Intrinzično neuređeni proteini (IDP)** nemaju stabilnu, jedinstvenu 3D strukturu, već postoje kao skup fleksibilnih konformacija, što im omogućava funkcionalnu prilagodljivost.
- **Metamorfni proteini** mogu prelaziti između više stabilnih strukturnih stanja (različitih „fold“-ova), pri čemu svako stanje može imati različitu funkciju.
- **Termodinamička hipoteza** opisuje savijanje proteina kao kretanje ka energetsom minimumu, pri čemu više lokalnih minimuma može odgovarati različitim stabilnim konformacijama.
- **Moonlighting proteini** su proteini koji obavljaju više različitih funkcija bez promene svoje primarne strukture (npr. enzim koji ima i regulatornu ulogu).
- **Morfni (morphesins)** predstavljaju posebnu klasu metamorfnih proteina koji mogu formirati različite oligomerne strukture (npr. dimer, heksamer, oktamer) ponovnim udruživanjem svojih podjedinica, što utiče na njihovu funkciju.

Donja vremenska linija ilustruje istorijski razvoj ovih koncepata, od 1970-ih do savremenih istraživanja, ukazujući na evoluciju razumevanja proteinske dinamike i funkcionalnosti.



Slika 3.34: Bitni koncepti za razumevanje proteinske strukture i funkcije

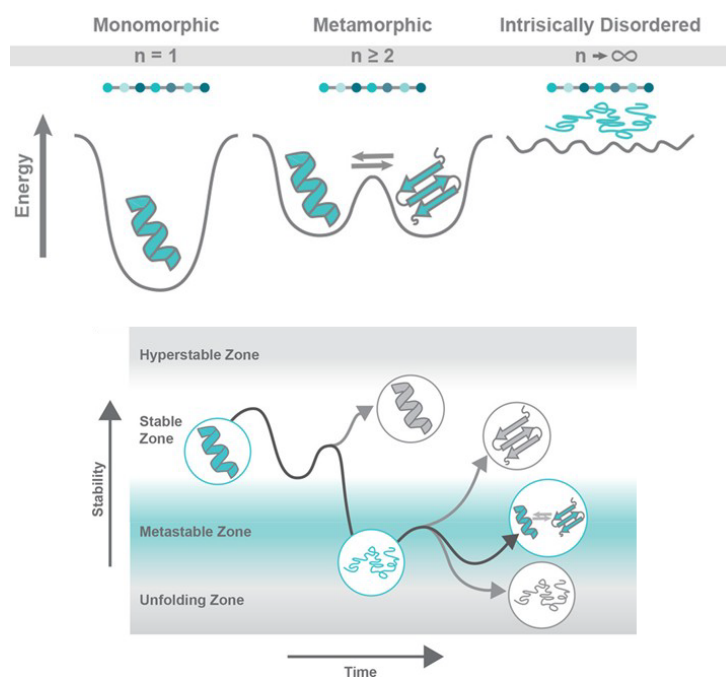
Slika 3.35 prikazuje različite tipove proteinskih konformacija i njihov odnos sa energetsom pejzažom (način prikaza kako se energija proteina menja u zavisnosti od njegove strukture (konformacije)) i stabilnošću. U gornjem delu prikazane su tri klase proteina:

- **Monomorfni proteini** ($n = 1$) imaju jednu stabilnu konformaciju koja odgovara jednom globalnom energetsom minimumu.
- **Metamorfni proteini** ($n \geq 2$) mogu postojati u više stabilnih konformacija, koje odgovaraju različitim lokalnim minimumima na energetsom pejzažu, između kojih protein može prelaziti.
- **Intrinzično neuređeni proteini** ($n \rightarrow \infty$) nemaju jedinstvenu stabilnu strukturu, već postoje kao skup mnogih fleksibilnih konformacija bez jasno definisanog minimuma.

Energetski pejzaž prikazan je kao niz dolina (minimuma), gde svaka dolina predstavlja stabilno ili metastabilno stanje proteina. U donjem delu slike prikazan je odnos između stabilnosti i vremena:

- **Stabilna zona** odgovara dubokim energetskim minimumima, gde je struktura dugotrajno očuvana.
- **Metastabilna zona** predstavlja pliće minimume, gde protein može prelaziti između više konformacija.
- **Zona razvijanja (unfolding)** odgovara nestabilnim stanjima u kojima protein gubi svoju strukturu.
- **Hiperstabilna zona** predstavlja izuzetno stabilne konformacije sa veoma malom verovatnoćom promene.

Slika ilustruje da proteinske strukture nisu statične, već dinamične, i da njihova funkcija često zavisi od sposobnosti prelaska između različitih konformacionih stanja.



Slika 3.35: Različiti tipovi proteinskih konformacija

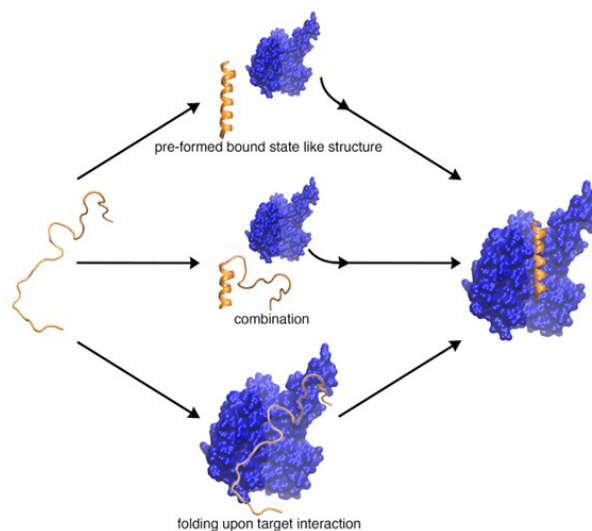
Neuređeni proteini (NP) i neuređeni proteinski regioni (NPR) nemaju fiksnu (uređenu) strukturu, tj. nemaju stabilne vrednosti φ i ψ uglova polipeptidnog lanca. U zavisnosti od stepena neuređenosti, mogu postojati u više konformacionih oblika: nasumični lanac, preistopljena globula, istopljena globula, kao i potpuno uređeni (fiksni) polipeptidni lanac. Neuređeni delovi proteina često učestvuju u interakcije sa drugim molekulima. Prilikom vezivanja za ligand (npr. hem i globin) mogu preći u uređeno, fiksno stanje. Funkcionalno su povezani sa regulatornim ćelijskim procesima, za razliku od enzima koji najčešće imaju stabilnu i uređenu strukturu. Poremećaj u prisustvu uređenih–neuređenih regiona može dovesti do patoloških stanja (npr. neuroloških bolesti i maligniteta).

Slika 3.36 prikazuje mehanizme vezivanja IDP za ciljne molekule. IDP može ostvariti interak-

ciju sa partnerom na više načina.

- Gornji mehanizam prikazuje situaciju u kojoj IDP najpre poprime strukturu sličnu vezanom stanju, a zatim se vezuje za ciljni molekul.
- Donji mehanizam prikazuje vezivanje pri kome IDP prvo dolazi u kontakt sa partnerom, a zatim se savija (formira uređenu strukturu) tokom ili nakon interakcije (eng. folding upon binding).
- Srednji mehanizam predstavlja kombinaciju prethodna dva pristupa, gde se delimično strukturiranje dešava pre i tokom vezivanja.

Slika ilustruje da IDP proteini mogu menjati svoju konformaciju u zavisnosti od interakcije sa drugim molekulima, što im omogućava veliku funkcionalnu fleksibilnost.

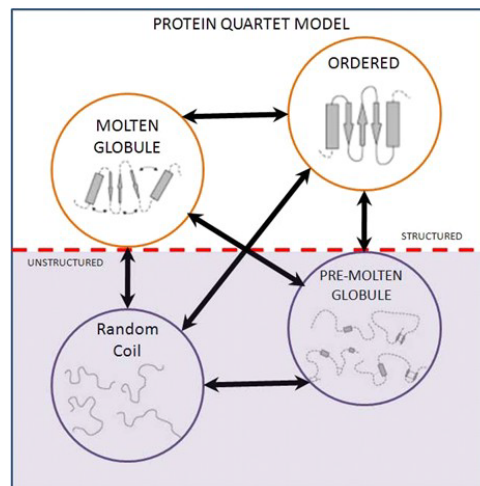


Slika 3.36: Mehanizmi vezivanja intrinzično neuređenih proteina (IDP) za ciljne molekule

Slika 3.37 prikazuje *protein kvartet model* Uverskoga, koji opisuje različite konformacione forme proteina u zavisnosti od stepena uređenosti i stabilnosti. Model obuhvata četiri osnovna stanja proteina:

- **Uređeno stanje (ordered)** – stabilna, dobro definisana trodimenzionalna struktura.
- **Istaljena globula (molten globule)** – delimično uređeno stanje sa očuvanom sekundarnom strukturom, ali bez potpuno definisane tercijalne strukture.
- **Pre-istaljena globula (pre-molten globule)** – još manje uređeno stanje sa delimičnom strukturom.
- **Nasumični lanac (random coil)** – potpuno neuređeno stanje bez stabilne strukture.

Isprekidana linija razdvaja uređene i neuređene regione. Strelice između stanja ukazuju na mogućnost prelaska proteina iz jednog konformacionog stanja u drugo, u zavisnosti od uslova sredine i interakcija sa drugim molekulima. Model ilustruje da proteini nisu uvek u jednoj fiksnoj strukturi, već mogu postojati u ravnoteži između više stanja različitog stepena uređenosti.



Slika 3.37: Protein kvartet model Uverskoga

3.2.9 Globularni proteini

Globularni proteini su proteini kompaktne, približno sferne strukture koji su uglavnom rastvorljivi u vodi i imaju različite funkcionalne uloge u organizmu. Slika 3.38 prikazuje veličinu i organizaciju globularnih proteina, kao i njihovu strukturnu raznovrsnost. U tabelarnom delu prikazani su različiti proteini sa njihovom molekulskom masom (M_r), brojem aminokiselinskih ostataka po lancu i organizacijom subjedinica. Proteini se značajno razlikuju po veličini, od malih molekula kao što su insulin i citohrom c, do velikih kompleksa kao što su glutamin sintetaza i miozin. Neki proteini su sastavljeni od jednog polipeptidnog lanca (monomeri), dok drugi imaju više podjedinica. Donji deo slike prikazuje prostorne modele odabranih proteina nacrtane u istoj razmeri, što omogućava direktno poređenje njihovih dimenzija i oblika. Uočava se da globularni proteini imaju kompaktan, približno sferni oblik, ali mogu značajno varirati u veličini i kompleksnosti.

3.3 Interakcije između proteina

3.3.1 Međumolekulske interakcije

Interaktom je skup svih međumolekulskih interakcija u datoj ćeliji. Najčešće se pod interaktomom podrazumeva protein–protein interaktom (PPI) ili mreža (PIN). Interaktomi se mogu odrediti eksperimentalno (yeast two-hybrid system (Y2H)) ili kompjuterski (npr. prenos funkcije na osnovu homologije (predviđanje funkcije ili interakcija proteina na osnovu sličnosti sa drugim, već okarakterisanim proteinima), ML). Na osnovu interaktoma mogu se predvideti proteinske funkcije, pratiti patološka stanja itd.

Interaktom se obično predstavlja u obliku grafa. Na slici 3.39 je prikazan ljudski proteinski interaktom. Kružići predstavljaju proteine, a linije međusobne interakcije.

Slika 3.40 prikazuje poremećeni (patološki) proteinski interaktom povezan sa šizofrenijom. Uočava se kompleksna mreža sa velikim brojem međusobno povezanih proteina, gde pojedini proteini imaju centralnu ulogu (tzv. „hub“ proteini) sa velikim brojem interakcija. Crveni čvorovi označavaju proteine koji su povezani sa patološkim stanjem (u ovom slučaju šizofrenijom),

Protein	M_r	Number of Residues per Chain	Subunit Organization
Insulin (bovine)	5,733	21 (A) 30 (B)	$\alpha\beta$
Cytochrome <i>c</i> (equine)	12,500	104	α_1
Rihonuclease A (bovine pancreas)	12,640	124	α_1
Lysozyme (egg white)	13,930	129	α_1
Myoglobin (horse)	16,980	153	α_1
Chymotrypsin (bovine pancreas)	22,600	13 (α) 132 (β) 97 (γ)	$\alpha\beta\gamma$
Hemoglobin (human)	64,500	141 (α) 146 (β)	$\alpha_2\beta_2$
Serum albumin (human)	68,500	550	α_1
Hexokinase (yeast)	96,000	200	α_1
γ -Globulin (horse)	149,900	214 (α) 446 (β)	$\alpha_2\beta_2$
Glutamate dehydrogenase (liver)	332,694	500	α_6
Myosin (rabbit)	470,000	2,000 (heavy, β) 190 (α) 149 (α') 160 (β)	$h_2\alpha_1\alpha'\beta_2$
Ribulose biphosphate carboxylase (spinach)	560,000	475 (α) 123 (β)	$\alpha_6\beta_8$
Glutamine synthetase (<i>E. coli</i>)	600,000	468	α_{12}

*Illustrations of selected proteins listed in Table 4.2 are drawn to constant scale.
Adapted from Goodsell, D.S., and Olson, A.J., 1993. Soluble proteins: Size, shape and function. Trends in Biochemical Sciences 18:65-68.

Slika 3.38: Opis nekih globularnih proteina

dok ostali čvorovi predstavljaju proteine koji učestvuju u širem interaktomu. Poremećaji u interakcijama između proteina mogu dovesti do disfunkcije ćelijskih procesa, što se odražava na nastanak i razvoj bolesti. Slika ilustruje da šizofrenija nije posledica poremećaja jednog proteina, već kompleksnih promena u mreži proteinskih interakcija.

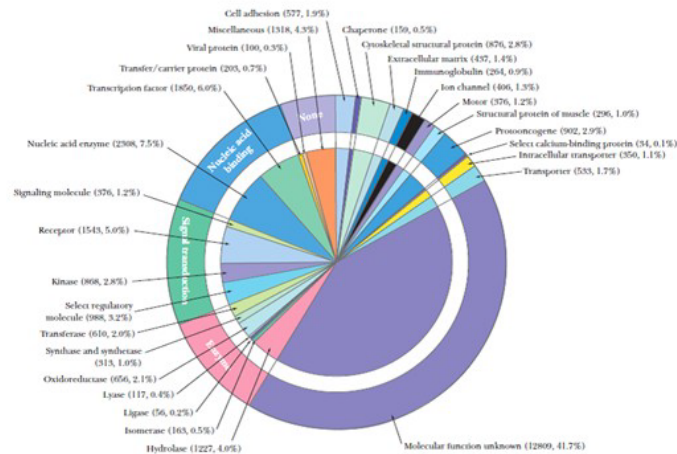
Postoji više javno dostupnih baza interaktoma (npr. STRING, BioGRID, IntAct) koje sadrže podatke o protein–protein interakcijama i omogućavaju analizu mreža interakcija u ćeliji.

3.3.2 Funkcija proteina

Proteini su glavni izvršni molekuli u ćeliji. Proteinske funkcije se mogu razmatrati na nivou *molekula* (molekulske interakcije, npr. enzimi – kataliza) ili na nivou ćelije (biološka funkcija, npr. imuni odgovor – antitela).

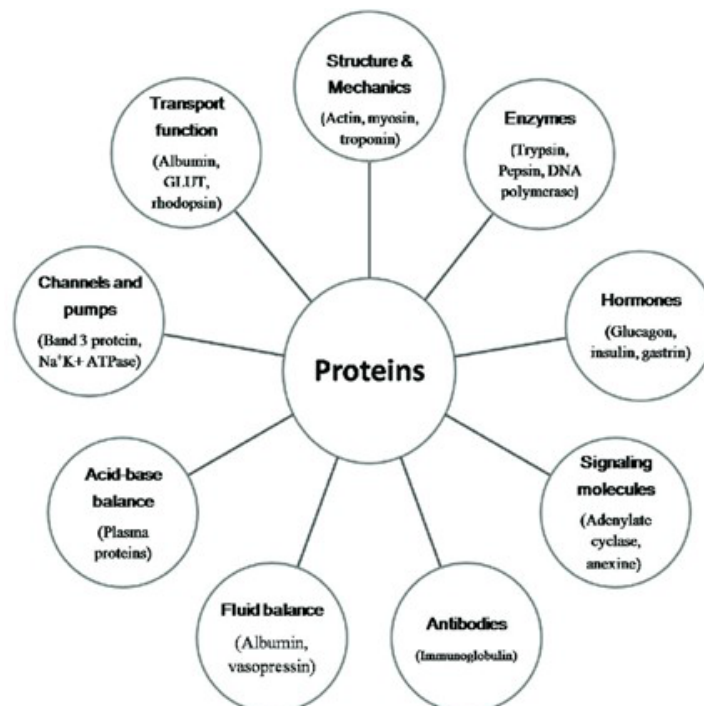
Slika 3.41 prikazuje raspodelu funkcija proteina u ljudskom genomu, organizovanu prema njihovoj molekularnoj i biološkoj ulozi. Proteini su grupisani prema molekularnoj funkciji (unutrašnji prsten) i biološkoj ulozi (spoljašnji prsten). Najveći deo proteina pripada kategoriji nepoznate molekulske funkcije, što ukazuje da značajan broj proteina još uvek nije u potpunosti funkcionalno okarakterisan. Među poznatim funkcijama dominiraju:

- enzimi (npr. transferaze, kinaze, hidrolaze),
- receptori i signalni proteini,



Slika 3.41: Funkcije proteina na molekulskom nivou

- **Antitela** (imunoglobulini) imaju ključnu ulogu u imunom odgovoru.
- **Regulacija kiselinsko-bazne ravnoteže** (plazma proteini) i **ravnoteže tečnosti** (npr. albumin, vazopresin) održavaju homeostazu organizma.
- **Kanali i pumpe** (npr. Na^+/K^+ -ATPaza) regulišu transport jona kroz membranu.
- **Transportni proteini** (npr. albumin, GLUT, rodopsin) omogućavaju prenos molekula kroz organizam i ćelijske membrane.
- **Strukturni i mehanički proteini** (npr. aktin, miozin, troponin) obezbeđuju oblik ćelije i omogućavaju kretanje.

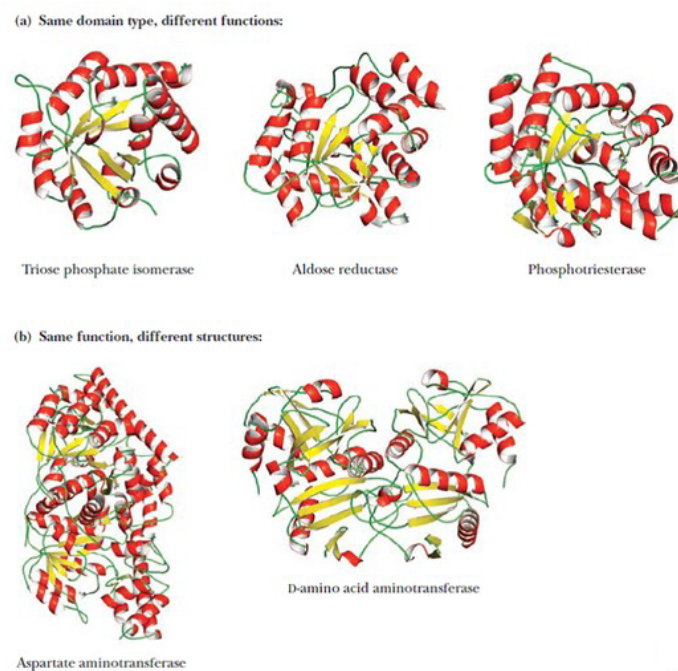


Slika 3.42: Funkcije proteina na biološkom nivou

Slika 3.43 ilustruje odnos između strukture i funkcije proteina kroz dva osnovna slučaja:

- (a) **Isti tip domena – različite funkcije.** Prikazani su proteini koji imaju sličnu trodimenzionalnu strukturu (isti ili sličan domen), ali obavljaju različite funkcije. To pokazuje da slična struktura ne mora nužno značiti istu funkciju, jer male razlike u aktivnom mestu ili interakcijama mogu dovesti do različitih bioloških uloga.
- (b) **Ista funkcija – različite strukture.** Prikazani su proteini koji obavljaju istu ili sličnu funkciju (npr. katalizu iste vrste reakcije), ali imaju različitu ukupnu strukturu. Ovo ukazuje na to da različite proteinske arhitekture mogu evolutivno razviti sposobnost obavljanja iste funkcije.

Slika naglašava da veza između strukture i funkcije proteina nije jednoznačna: ista struktura može imati više funkcija, a ista funkcija može biti ostvarena različitim strukturnim rešenjima.



Slika 3.43: Odnos između strukture i funkcije proteina

Proteini koji mogu obavljati više različitih funkcija iako imaju istu aminokiselinsku sekvencu i strukturu se nazivaju *moonlighting* proteini (ili „gene sharing“). U tabeli 3.13 su dati primeri moonlighting proteina.

3.4 Klasifikacija proteina

Proteini se mogu klasifikovati prema raznim kriterijumima: *hemijskim* (rastvorljivosti u polarnim ili nepolarnim rastvaračima, prisustvu različitih prostetičkih grupa, molekularnoj masi itd.), *strukturnim* (homologije na nivou primarne, sekundarne ili 3D strukture), prema *morfologiji* (globularni i fibrilarni), *interakcijama* sa drugim molekulima (enzimi, transportni proteini, antitela itd.), *ćelijskoj lokalizaciji* (membranski, citoplazmatski) itd.

Grupa	Protein	Organizam	Primarna funkcija	Dodatna (moonlighting)
Životinje	Aconitaza	H. sapiens	Enzim TCA ciklusa	Homeostaza gvožđa
	ATF2	H. sapiens	Transkripcioni faktor	Odgovor na oštećenje DNK
	Kristalini	Razni	Strukturni protein sočiva	Različite enzimske funkcije
	Citohrom c	Razni	Energetski metabolizam	Apoptoza
	ERK2	H. sapiens	MAP kinaza	Transkripcioni represor
Biljke	Heksokinaza	A. thaliana	Metabolizam glukoze	Signalizacija glukoze
	Presenilin	P. patens	γ -sekretaza	Funkcija citoskeleta
Gljive	Aconitaza	S. cerevisiae	Enzim TCA ciklusa	Stabilnost mtDNK
	Aldolaza	S. cerevisiae	Glikolitički enzim	Sklapanje V-ATPaze
	Enolaza	S. cerevisiae	Glikolitički enzim	Fuzija vakuola / transport tRNK
	Galaktokinaza	K. lactis	Metabolizam galaktoze	Indukcija gena
	HSP60	S. cerevisiae	Mitohondrijski šaperon	Stabilizacija DNK
Prokarioti	Aconitaza	M. tuberculosis	Enzim TCA ciklusa	Regulacija gvožđa
	Enolaza	S. pneumoniae	Glikolitički enzim	Vezivanje plazminogena
	GroEL	E. aerogenes	Šaperon	Toksin za insekte
	MutD	E. coli	Biosinteza ćelijskog zida	Inhibicija giraze
	Tioredoksin	E. coli	Antioksidans	Podjedinica DNK polimeraze
Protisti	Aldolaza	P. vivax	Glikolitički enzim	Invazija ćelija domaćina

Tabela 3.13: Primeri „moonlighting“ proteina – proteini sa više funkcija

3.4.1 Strukturna klasifikacija proteina

Strukturna klasifikacija proteina zasniva se na njihovoj trodimenzionalnoj organizaciji, odnosno na načinu rasporeda sekundarnih struktura u prostoru (eng. protein fold). Proteini se grupišu prema sličnostima u strukturi i evolutivnom poreklu, koristeći klasifikacione sisteme kao što su SCOP, CATH i FSSP. Ovi sistemi omogućavaju bolje razumevanje odnosa između strukture i funkcije proteina, kao i identifikaciju strukturno sličnih proteina. Postoji nekoliko klasifikacionih sistema:

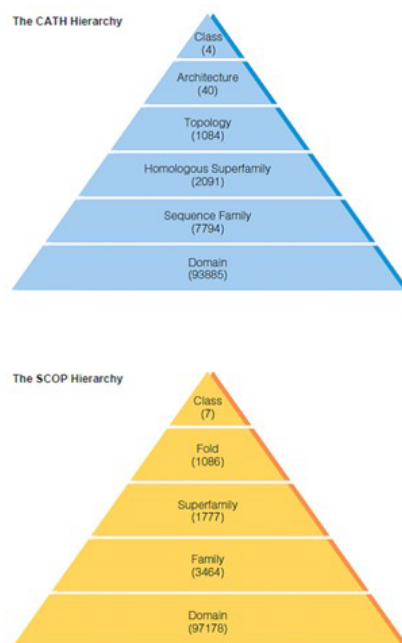
- SCOP (Structural Classification of Proteins) <http://scop.mrc-lmb.cam.ac.uk/>
- CATH (Class, Architecture, Topology, Homology) <https://www.hsls.pitt.edu/obrc/index.php?page=URL1100794593>
- FSSP (Families of Structurally Similar Proteins)

Slika 3.44 prikazuje hijerarhijsku organizaciju strukturne klasifikacije proteina kroz dva glavna sistema, CATH i SCOP. Proteini su u SCOP-u klasifikovani ručno, dok CATH kombinuje manuelne i automatske procedure. Brojevi označavaju broj elemenata u svakoj kategoriji. CATH klasifikacija organizovana je u sledeće nivoe:

- klasa (eng. Class) – zasnovana na tipu sekundarne strukture,
- arhitektura (eng. Architecture) – opšti raspored sekundarnih struktura,
- topologija (eng. Topology) – način povezivanja sekundarnih elemenata,
- homologna superfamilija (eng. Homologous superfamily) – proteini sa zajedničkim evolutivnim poreklom,
- familija sekvence (eng. Sequence family),
- domen (eng. Domain).

SCOP klasifikacija ima sličnu hijerarhiju:

- klasa (eng. Class),
- fold – opšta 3D struktura,
- superfamilija (eng. Superfamily),
- familija (eng. Family),
- domen (eng. Domain).



Slika 3.44: Hijerarhije u SCOP i CATH klasifikacijama

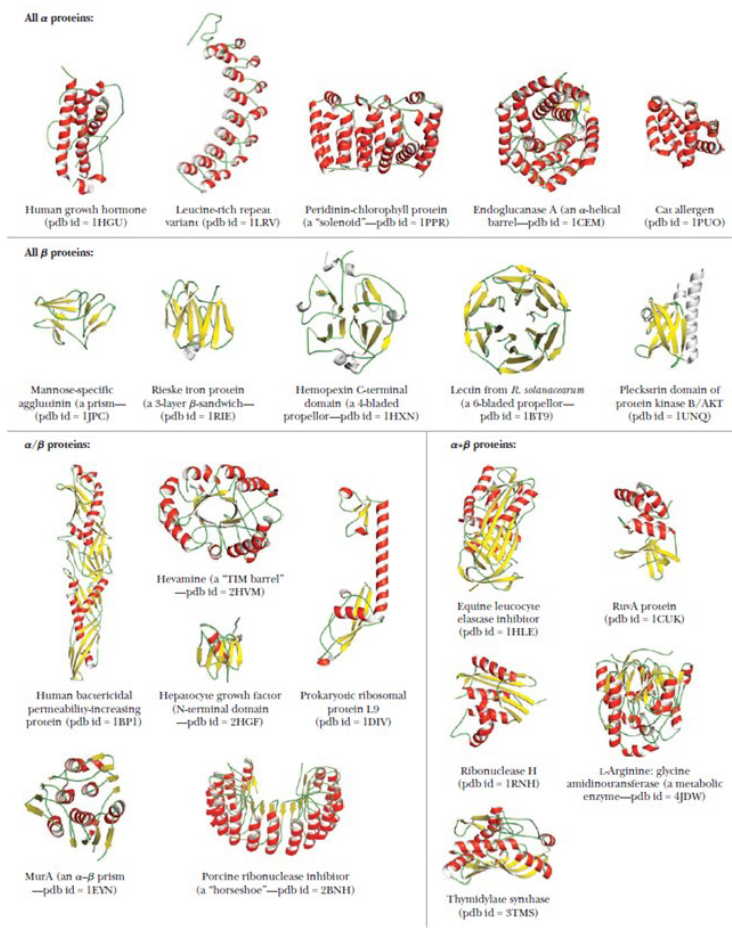
Slika 3.45 prikazuje četiri glavne klase proteinskih struktura prema SCOP klasifikaciji, zasnovane na zastupljenosti i rasporedu sekundarnih struktura:

- **All α proteini** – strukture u kojima dominiraju α -zavojnice, koje formiraju kompaktne helikoidne motive.
- **All β proteini** – proteini kod kojih dominiraju β -ploče.
- **α/β proteini** – proteini kod kojih su α -zavojnice i β -ploče međusobno izmešane unutar istog domena.
- **$\alpha + \beta$ proteini** – proteini kod kojih su α i β strukture prisutne, ali su prostorno odvojene u različitim delovima domena.

Svaka klasa ilustrovana je primerima proteina sa poznatim PDB identifikatorima, čime se prikazuje raznovrsnost trodimenzionalnih struktura.

3.5 Eksperimentalne metode za određivanje strukture proteina

Eksperimentalne metode za određivanje strukture proteina omogućavaju uvid u njihovu trodimenzionalnu organizaciju na atomskom nivou. Najznačajnija metoda je difrakcija X-zraka,



Slika 3.45: Klase u SCOP klasifikaciji

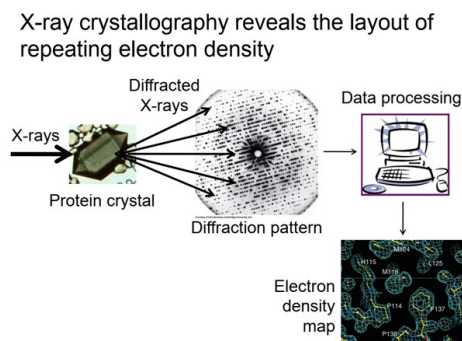
koja se zasniva na rasipanju X-zraka na kristalu proteina i omogućava određivanje rasporeda atoma. Pored nje, koriste se i NMR spektroskopija, pogodna za analizu proteina u rastvoru, kao i kriogena elektronska mikroskopija (cryo-EM), koja omogućava proučavanje velikih proteinskih kompleksa bez potrebe za kristalizacijom.

3.5.1 Metoda zasnovana na difrakciji X-zraka

Slika 3.46 prikazuje princip određivanja strukture proteina metodom difrakcije X-zraka:

- X-zraci se usmeravaju na kristal proteina, gde dolazi do njihovog rasipanja (difrakcije) usled interakcije sa elektronima u atomima proteina.
- Difraktovani X-zraci formiraju karakterističan difrakcioni obrazac, koji sadrži informacije o prostornom rasporedu atoma u kristalu.
- Dobijeni podaci se zatim obrađuju računski kako bi se rekonstruisala mapa elektronske gustine.
- Na osnovu mape elektronske gustine moguće je odrediti položaj atoma i izgraditi trodimenzionalni model proteinske strukture.

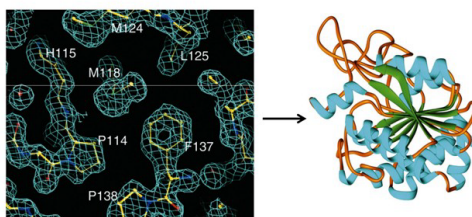
Slika 3.47 prikazuje kako se na osnovu mape elektronske gustine određuje trodimenzionalna struktura proteina. Levo je prikazana mapa elektronske gustine, gde mrežasta struktura predsta-



Slika 3.46: Metoda zasnovana na difrakciji X-zraka

vlja raspodelu elektrona u prostoru, a označene tačke odgovaraju položajima atoma u proteinu. Na osnovu ovih podataka moguće je precizno odrediti položaj pojedinačnih aminokiselinskih ostataka i njihovih atoma. Desno je prikazan model proteina, rekonstruisan iz mape elektronske gustine, gde su jasno vidljive sekundarne strukture poput α -zavojnica i β -ploča.

Electron density map allows for positioning of protein atoms, revealing structure



Slika 3.47: Određivanje 3D strukture proteina na osnovu mape elektronske gustine

3.5.2 2D Nuklearno-magnetna rezonantna spektroskopija

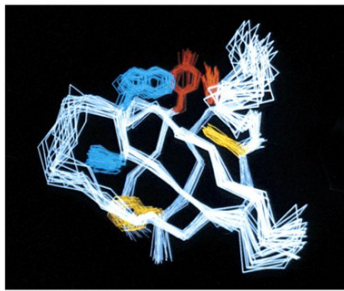
2D nuklearno-magnetna rezonantna spektroskopija (2D NMR) je eksperimentalna metoda koja omogućava određivanje strukture proteina u rastvoru, na osnovu interakcija između atomskih jezgara. Ova tehnika daje informacije o udaljenostima i prostornim odnosima između atoma, što omogućava rekonstrukciju trodimenzionalne strukture proteina. Posebno je pogodna za proučavanje manjih i srednje velikih proteina u uslovima bliskim fiziološkim.

Slika 3.48 prikazuje kako se podaci dobijeni NMR spektroskopijom koriste za određivanje strukture proteina:

- NMR eksperimenti obezbeđuju ograničenja rastojanja između atoma (eng. distance constraints), zasnovana na interakcijama između protona.
- Na osnovu ovih ograničenja računski se generiše više mogućih konformacija proteina koje zadovoljavaju eksperimentalne podatke.
- Prikazani skup linija predstavlja više modela strukture, gde se podudaranje delova strukture ukazuje na dobro definisane regione, dok rasipanje linija označava fleksibilnije ili manje precizno određene delove.

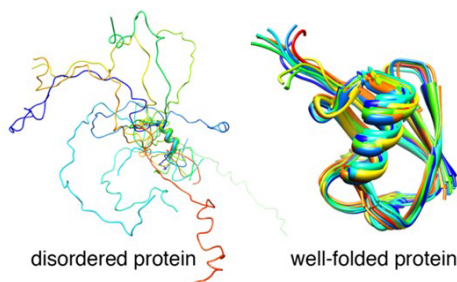
Primetiti da *NMR* ne daje jednu jedinstvenu strukturu kao difrakcija *X*-zrakom, već skup mogućih konformacija koje su u skladu sa eksperimentalnim podacima.

NMR-derived distance constraints are used to calculate likely protein conformations



Slika 3.48: 2D NMR

Slika 3.49 prikazuje poređenje intrinzično neuređenog proteina i dobro savijenog (uređenog) proteina. Levo je prikazan neuređeni protein, čije različite konformacije (prikazane različitim bojama) pokazuju veliku varijabilnost i ne preklapaju se dobro, što ukazuje na odsustvo stabilne 3D strukture. Desno je prikazan dobro savijen protein, gde se različite konformacije gotovo potpuno preklapaju, što ukazuje na stabilnu i dobro definisanu strukturu. Ovakvo poređenje ilustruje razliku između fleksibilnih, dinamičkih proteina (IDP) i stabilnih globularnih proteina sa jasno određenom strukturom.



Slika 3.49: Poređenje IDP i urednog proteina

Bioinformatics Note

Koda:

*Čovek je nesumnjivo najsloženiji sistem za obradu informacija koji postoji na Zemlji. Ukupan broj bitova koji se svakodnevno obrađuju u svim informacionim procesima u ljudskom telu iznosi $3 * 10^{24}$. Ovo uključuje sve svesne, kao i sve nesvesne aktivnosti, pri čemu prve obuhvataju upotrebu jezika i informacije potrebne za kontrolu voljnih pokreta, dok druge uključuju kontrolu unutrašnjih organa i hormonskih sistema. Broj bitova koji se svakodnevno obrađuju u ljudskom telu veći je za više od milion puta od ukupne količine ljudskog znanja pohranjenog u svim bibliotekama sveta, što iznosi oko 10^{18} bitova.*

* Tekst predstavlja ilustrativan primer iz bioinformatike koji naglašava ogromnu količinu informacija koje ljudski organizam obrađuje, često korišćen kao motivacioni zaključak u literaturi. Broj bitova koje ljudski organizam obrađuje procenjuje se primenom teorije informacija na neuronske, biohemijske i fiziološke procese, ali predstavlja samo grubu aproksimaciju složenosti sistema.

Bibliografija

- [1] Bruce Alberts, Dennis Bray, Karen Hopkin, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Essential Cell Biology*. Garland Science, New York, NY, 3rd edition, 2010.
- [2] Bruce Alberts, Dennis Bray, Karen Hopkin, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter. *Essential Cell Biology*. Garland Science, New York, NY, 4th edition, 2014.
- [3] Marcello Barbieri. *The Organic Codes: An Introduction to Semantic Biology*. Cambridge University Press, Cambridge, UK, 2003. 312 pages.
- [4] Marcello Barbieri. *Code Biology: A New Science of Life*. Springer International Publishing, Cham, Switzerland, 2015.
- [5] Helen Pearson. Genetics: what is a gene? *Nature*, 441(7092):398–401, 2006.
- [6] Elizabeth Pennisi. Dna study forces rethink of what it means to be a gene. *Science*, 316(5831):1556–1557, 2007.
- [7] Petter Portin and Adam S. Wilkins. The Evolving Definition of the Term “Gene”. *Genetics*, 205(4):1353–1364, 2017.
- [8] Marko Vitas and Andrej Dobovišek. Towards a general definition of life. *Origins of Life and Evolution of Biospheres*, 49(1–2):77–88, 2019.