Sveučilište u Zagrebu
PRIRODOSLOVNO-MATEMATIČKI FAKULTET

Boris Gomaz

MOLEKULSKA DINAMIKA I KONFORMACIJSKE PROMJENE FOSFORILAZE PURINSKIH NUKLEOZIDA

DOKTORSKI RAD

Mentor:
dr. sc. Zoran Štefanić

Zagreb, 2025. godina

University of Zagreb
FACULTY OF SCIENCE

Boris Gomaz

MOLECULAR DYNAMICS AND CONFORMATIONAL CHANGES OF PURINE NUCLEOSIDE PHOSPHORYLASE

DOCTORAL DISSERTATION

Supervisor:
Dr. Zoran Štefanić

Zagreb, 2025 year

Ova disertacija izrađena je na Zavodu za fizičku kemiju Instituta Ruđer Bošković u Laboratoriju za kemijsku i biološku kristalografiju pod vodstvom dr. sc. Zorana Štefanića u okviru projekta Hrvatske zaklade za znanost pod nazivom „Alosterički komunikacijski putevi u oligomernim enzimima“ (IP-2019-04-6764).

Sve što je ovdje napisano ne bi postojalo bez ljudi koji su bili uz mene, direktno ili indirektno.
Znaju oni tko su – i znaju koliko mi to znači.

Sadržaj

SADRŽAJ **9**
SAŽETAK ..... **XIII**
ABSTRACT ..... **XV**
§ 1. UVOD **1**
§ 2. LITERATURNI PREGLED **5**
2.1. Bakterija *Helicobacter pylori* **5**
2.1.1. Epidemiologija H. pylori *6*
2.1.2. Patofiziologija H. pylori *8*
2.1.3. Bolesti povezane uz H. pylori *10*
2.1.4. Antibiotska rezistencija *11*
2.2. Fosforilaza purinskih nukleozida (PNP) **12**
2.2.1. Metabolizam purina kod H. pylori *15*
2.2.2. Struktura PNP-a iz H. pylori *17*
2.2.3. Mehanizam katalize *20*
2.2.4. Inhibitori fosforilaze purinskih nukleozida *23*
2.3. (Molekulska) dinamika proteina **24**
2.3.1. Alosterija *26*
2.3.2. Ramachandranovi kutovi *28*
2.3.3. Metode MD simulacija *30*
2.3.4. Analiza podataka *32*
§ 3. EKSPERIMENTALNI DIO **37**
3.1. Kristalne strukture **37**
3.2. Simulacije molekulske dinamike **42**
3.3. Analiza trajektorija **44**
3.4. Programi i računala **48**
§ 4. REZULTATI I RASPRAVA **49**
4.1. Razvoj metodologije MDavocado dijagrama **50**
4.2. Analiza H8 *α*-zavojnice i aktivnog mjesta **56**
4.2.1. Analiza glavnih komponenata *58*
4.2.2. Analiza struktura bez supstrata *60*
4.2.3. Analiza struktura sa supstratima *64*

4.3. Korelacije u gibanjima aminokiselina 75
4.4. Prostorni raspored korelacija 81
4.5. Vremenska analiza i točke promjene 86
§ 5. ZAKLJUČAK 90
§ 6. POPIS OZNAKA, KRATICA I SIMBOLA 93
§ 7. LITERATURNI IZVORI 97
§ 8. ŽIVOTOPIS CV

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Kemijski odsjek

Doktorska disertacija

SAŽETAK

MOLEKULSKA DINAMIKA I KONFORMACIJSKE PROMJENE FOSFORILAZE PURINSKIH NUKLEOZIDA

Boris Gomaz
Institut Ruđer Bošković, Bijenička cesta 54, 10000 Zagreb, Hrvatska

Helicobacter pylori, bakterija odgovorna za brojne gastrointestinalne bolesti, predstavlja značajan istraživački izazov zbog svoje iznimno visoke prevalencije i sve veće otpornosti na antibiotike. Jedan od ključnih enzima u njenom metabolizmu je fosforilaza purinskih nukleozida, jer ova bakterija nema sposobnost *de novo* sinteze nukleotida. Eksperimentalno utvrđene strukture daju detaljan uvid u građu ovog enzima, ali razumijevanje njegovih lokalnih i globalnih konformacijskih promjena zahtijeva proučavanje njegove dinamike. Razvijena je metodologija koja omogućuje vizualizaciju lokalnih dinamičkih promjena u enzimu, i praćenje korelacija koje su omogućile uvid u sinkronizirane konformacijske prijelaze. Simulacije molekularne dinamike ponudile su detaljan prikaz dinamičkih promjena u ovom enzimu, identificirajući ključne regije kao što su središnje heksamerno područje, sučelja dimera i aktivna mjesta enzima i putove komunikacije među njima. Istraživanje također ukazuje na vezu navedenih rezultata i kretanja fosfatnih iona te otvaranja i zatvaranja aktivnog mjesta enzima, nudeći smjernice za buduća eksperimentalna istraživanja i potencijalne terapijske pristupe.

(123 stranica, 47 slika, 1 tablica, 135 literaturnih navoda, jezik izvornika: hrvatski)

Rad je pohranjen u Središnjoj kemijskoj knjižnici, Horvatovac 102a, Zagreb i Nacionalnoj i sveučilišnoj knjižnici, Hrvatske bratske zajednice 4, Zagreb.

Ključne riječi: alosterija, fosforilaza purinskih nukleozida, *Helicobacter pylori*, konformacijske promjene, MDavocado dijagram, molekulska dinamika

Mentor: dr. sc. Zoran Štefanić, viši znanstveni suradnik

Rad prihvaćen: 05. veljače 2025.

Ocjenitelji:
1. prof. dr. sc. Biserka Prugovečki
2. prof. dr. sc. Tomica Hrenar
3. dr. sc. Borislav Kovačević, v. zn. sur.

University of Zagreb
Faculty of Science
Department of Chemistry

Doctoral Thesis

ABSTRACT

MOLECULAR DYNAMICS AND CONFORMATIONAL CHANGES OF PURINE NUCLEOSIDE PHOSPHORYLASE

Boris Gomaz
Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Hrvatska

Helicobacter pylori, a bacterium responsible for numerous gastrointestinal diseases, represents a significant research challenge due to its extremely high prevalence and increasing antibiotic resistance. One of the key enzymes in its metabolism is purine nucleoside phosphorylase, as this bacterium lacks the ability to synthesize new nucleotides *de novo*. Experimentally determined structures provide detailed insight into the structure of this enzyme, but understanding its local and global conformational changes requires studying its dynamics. A methodology has been developed that allows visualization of local dynamic changes in the enzyme, and monitoring of correlations that have provided insight into synchronized conformational transitions. Molecular dynamics simulations have provided a detailed view of the dynamic changes in this enzyme, identifying key regions such as the central hexameric region, dimer interfaces, and active sites of the enzyme and the communication pathways between them. The research also suggests a connection between the above results and the movement of phosphate molecules and the opening and closing of the active site, offering directions for future experimental studies and potential therapeutic approaches.

(123 pages, 47 figures, 1 table, 135 references, original in Croatian)

Thesis deposited in Central Chemical Library, Horvatovac 102A, Zagreb, Croatia and National and University Library, Hrvatske bratske zajednice 4, Zagreb, Croatia.

Keywords: allostery, conformational changes, *Helicobacter pylori*, MDavocado diagrams, molecular dynamics, purine nucleoside phosphorylase

Supervisor: Dr. Zoran Štefanić, Senior Research Associate

Thesis accepted: 5th of February 2025.

Reviewers:
1. Dr. Biserka Prugovečki, Professor
2. Dr. Tomica Hrenar, Professor
3. Dr. Borislav Kovačević, Senior Research Associate

§ 1. UVOD

Helicobacter pylori (*H. pylori*) je gram-negativna bakterija koja kolonizira sluznicu želuca i dvanaesnika, uzrokujući različite poremećaje gornjeg gastrointestinalnog trakta, uključujući gastritis, čir na želucu i, u nekim slučajevima, adenokarcinom želuca.[1] Zbog svoje sposobnosti prilagodbe i otpornosti na veoma nepogodne uvjete u želucu, *H. pylori* postala je izazov mnogih istraživanja. Ovom bakterijom zaraženo je više od 43 % svjetske populacije,[2] a tom broju svakako doprinosi laka transmisija patogena i njegova sve veća otpornost na antibiotike.[3] Kod proučavanja posebnu pažnju privlače enzimi ključni za metaboličke procese bakterije, jer oni predstavljaju potencijalne mete za razvoj novih terapijskih strategija.

Jedan od takvih enzima je fosforilaza purinskih nukleozida (engl. *purine nucleoside phosphorylase*, PNP). PNP katalizira fosforolitičku razgradnju purinskih nukleozida u purinske baze i ribozu-1-fosfat, čime sudjeluje u jednom od dva moguća procesa sinteze nukleotida.[4,5] Ovaj enzim je od velike važnosti za *H. pylori* budući da ona nema mogućnost sinteze nukleotida *de novo*, već se u potpunosti oslanja na purinski reciklirajući put, a u njemu je jedan od ključnih enzima upravo PNP.[6] Zbog svoje esencijalne funkcije, ovaj enzim je sve više u fokusu istraživanja ne samo kako bi se bolje razumjeli biološki mehanizmi ovog patogena, već i kako bi se razvile nove terapije koje ciljaju ključne enzime.

Enzimi PNP su strukturalno i kinetički okarakterizirani kod mnogo organizama, te se pokazalo da su po mnogo čemu zanimljivi proteini. Naime nalazimo ih kod nižih organizama kao heksamerne proteine, dok su kod viših organizama to trimeri. Trodimenzijska struktura pojedinog monomera iz heksamera i trimera je vrlo slična, usprkos relativno niskoj podudarnosti primarnih sekvenci.[7] Uz ovu strukturnu sličnost PNP-ovi pokazuju i zanimljivost u samoj enzimskoj kinetici koja ukazuje na postojanje alosteričke regulacije, a koja opet upućuje na to da pojedini monomeri u proteinu izmjenjuju informacije između aktivnih mjesta. No mehanizam kako to rade je u velikoj mjeri nerazjašnjen.[8]

Za rasvjetljavanje na koji način dolazi do ovakvih mehanizama kod PNP enzima potrebno je posegnuti za metodama koje uključuju dinamičku sliku proteina. Molekularno-dinamičke (MD) simulacije predstavljaju moćan alat za proučavanje dinamike bioloških makromolekula na atomskom nivou te one omogućuju detaljnu analizu gibanja proteina u vremenu, pružajući uvid u procese koji nisu dostupni eksperimentalnim metodama poput

rendgenske strukturne analize. Takva analiza PNP-a može dati dublji uvid u alosteričku regulaciju i dinamiku enzima. MD simulacije omogućuju praćenje konformacijskih promjena, identificiranje ključnih regija odgovornih za funkciju enzima te razumijevanje alosteričkih prijenosa unutar proteinske strukture. No, te simulacije neizostavno sadrže previše podataka da bi se sagledala cjelokupna slika i nužno je napraviti radikalnu redukciju podatkovnog prostora, a da se pritom ne izgubi informacija koja je potencijalno od interesa za alosteriju. Postavlja se pitanje kako iz mnoštva podataka dostupnih u MD simulacijama izdvojiti one koje nose potpis ovakve alosteričke komunikacije?

Jedan od mogućih načina takve redukcije je analiza konformacija proteinske okosnice koja je dobro opisana torzijskim Ramachandranovim kutovima φ i ψ. Mogući su i drugi odabiri parametara, ali je pretpostavka da će veće konformacijske promjene proteina zasigurno biti vidljive u ovim kutnim varijablama, a one pritom neće biti previše ovisne o lokalnim promjenama kao što su npr. rotacije aminokiselinskih bočnih lanaca.

Cilj ovog istraživanja bio je razviti i primijeniti novi pristup analize proteinske dinamike koji omogućava praćenje vremenski ovisnih promjena u strukturi enzima te konformacijskih promjena koje se dešavaju u njegovom aktivnom mjestu. Jedan od naglasaka u rezultatima ovog istraživanja je bila alosterička interakcija te korelirane promjene i komunikacijski putovi unutar same proteinske strukture. Svrha disertacije je poboljšati razumijevanje konformacijskih prijelaza u enzimu PNP te njihove uzroke i posljedice koje se dešavaju na atomskoj razini u samoj trodimenzijskoj strukturi. Ovakva dinamička ponašanja u enzimima moguće je proučavati pomoću dugotrajnih simulacija molekulske dinamike. Upravo MD simulacije služe kao izvor velikog seta podataka koji se analizirao na razne načine poput izračuna torzijskih kutova i položaja kako cijelog proteina tako i njegovih manjih dijelova. Također su trajektorije poslužile za izračun koreliranih kretanja aminokiselina te komunikacijskih putanji unutar strukture enzima. Polazna hipoteza disertacije glasi da su konformacijske promjene vremenski nepravilno distribuirane, ali da do njih dolazi zbog funkcionalno važnih događanja u strukturi proteina te da se signal za regulaciju konformacijskih stanja prenosi okosnicom enzima, odnosno promjenama njegovih torzijskim kutova.

S tim ciljem razvijena je nova metodologija i program za analizu lokalnih promjena okosnice enzima, nazvan MDavocado.[9] Dok su tradicionalno ovi kutovi korišteni za prikaz statičke konformacije proteina, ovaj novi način analize stavlja Ramachandranove kutove (međuovisne φ i ψ kutove) u vremenski ovisan kontekst. MDavocado dijagrami, kao krajnji

produkt ove analize, prikazuju vremenski razvoj ovisnosti φ–ψ kutova svake aminokiseline posebno i to za cijelo vrijeme trajanja MD simulacije neovisno o njezinoj duljini. Ovom analizom jasno se mogu utvrditi regije koje odgovaraju α-zavojnici ili β-ploči budući da je njihova međuovisnost φ i ψ kutova jasno definirana u Ramachandranovom prostoru. Ovim vremenski ovisnosnim dijagramima postaje očito kada dolazi do bilo kakve promjena u sekundarnoj strukturi i to sve skupa je vidljivo za cijeli protein ili za neku njegovu regiju na jednom dijagramu. Ovakav pristup na neki način zadržava sve prednosti klasičnog Ramachandranovog dijagrama, ali mu daje dodatnu dinamičku komponentu kojom možemo precizno pratiti dinamiku proteina.

Budući da se radi o istovremenom prikazu velikog broja podataka na ovakvim dijagramima korištene su suvremene metode napredne vizualizacije podataka. Ispostavlja se da aminokiseline u proteinu tijekom simulacija u većem dijelu održavaju stalnu konformaciju, no u jednom manjem dijelu skaču između dva, ili u rjeđim slučajevima, tri ili četiri stacionarna stanja dobro odijeljena u MDavocado dijagramima. Aminokiseline kod kojih se ovo događa, pokazuju znatnu međusobnu korelaciju u ovim gibanjima, što upućuje da povezanost ovakvih događaja nije slučajna. Prostornim mapiranjem koreliranih aminokiselina dobivene su regije proteina koje sudjeluju u interakciji tijekom simulacija, te su se najsnažnijima pokazale upravo regije centralnog područja heksamera PNP-a, sučelja dimera ovog enzima te područja aktivnih mjesta. Jedan od glavnih ciljeva ove analize je povezivanje ovih područja s konformacijskom promjenom koja se dešava prilikom enzimske reakcije, a to je već od prije poznato zatvaranje aktivnog mjesta do kojeg dolazi segmentacijom terminalne α-zavojnice. Usporednim strukturnim i konformacijskim analizama utvrđeno je da kod PNP-a dolazi do promjene konformacije iz otvorene u zatvorenu i obratno što se jasno može detektirati iz pojedinih analiza. Jedno od mogućih uzročno-posljedičnih događaja koji su povezani s konformacijskom promjenom je ulazak, odnosno izlazak fosfata iz aktivnog mjesta. Također je napravljena analiza prostorno i vremenski bliskih promjena konformacije proteinske okosnice i tako su konstruirani putovi kojima se te promjene prostiru. Ovi putovi dobro se podudaraju s područjima u kojima prevladavaju znatno korelirane aminokiseline, što dodatno potvrđuje ispravnost analize i moguću vezu ovih putova s alosteričkom komunikacijom između monomera u enzimu.

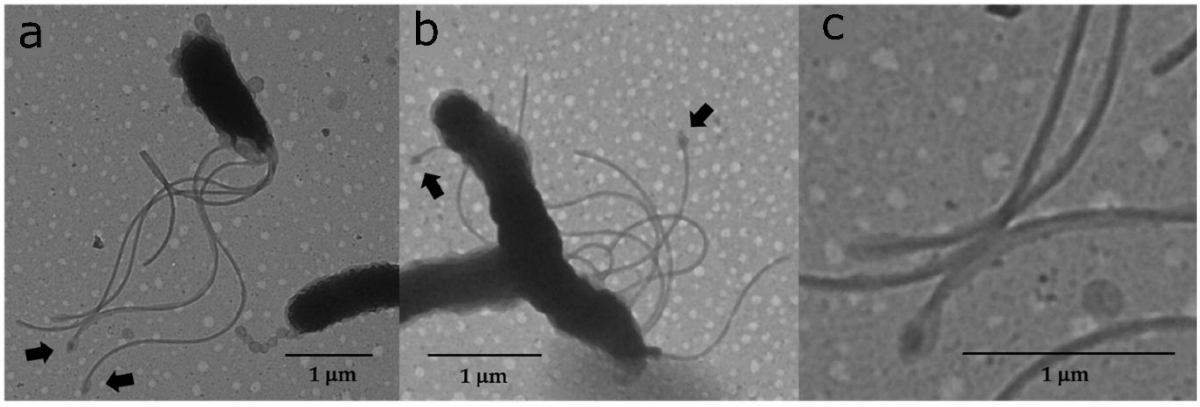
Ukupno gledano, provedene MD simulacije PNP-a iz *H. pylori* značajno doprinose razumijevanju dinamičkih svojstava ovog enzima. Rezultati ukazuju na važnost pojedinih regija

u funkcionalnosti enzima te potvrđuju da ovakve simulacije pružaju uvide u molekularne mehanizme koji se odvijaju na vremenskim i prostornim skalama nedostupnim klasičnim eksperimentalnim metodama.

§ 2. LITERATURNI PREGLED

2.1. Bakterija *Helicobacter pylori*

Gastroenterologija je grana medicine koja se bavi proučavanjem probavnog trakta što zapravo uključuje nekoliko organa istovremeno poput želuca, crijeva, jetre, itd. Važna godina u razvoju patofiziologije probavnog sustava je svakako 1982. kada su Marshall i Warren prvi put opisali bakteriju *Helicobacter pylori* (*H. pylori*).[1] Značaj njihovog otkrića je da su uspjeli uzgojiti mikrobnu kulturu (iz tkiva dobivenih biopsijom) te da su promatrani mikrobi bili gram negativni i da su imali bičeve, a za razvoj su im bili potrebni mikroaerofilni uvjeti. Ove novootkrivene bakterije mogućeg *Campylobacter* roda bile su prisutne kod gotovo svih pacijenata koji su imali aktivan kronični gastritis, peptički ulkus na stijenki želuca ili dvanaesnika. Zaključili su kako ova bakterija potencijalno ima ulogu u etiologiji navedenih bolesti, mišljenje koje i danas prevladava.[2,10] Prikaz bakterije te njene morfologije može se vidjeti na slici 1 koja prikazuje izduženi oblik same bakterije te mnogobrojne bičeve koji su karakteristični za *H. pylori*.



Slika 1. Prikaz morfologije (a i b) te bičeva (c) bakterije *H. pylori* pomoću transmisijskog elektronskog mikroskopa (TEM) (preuzeto i prilagođeno iz reference [11]).

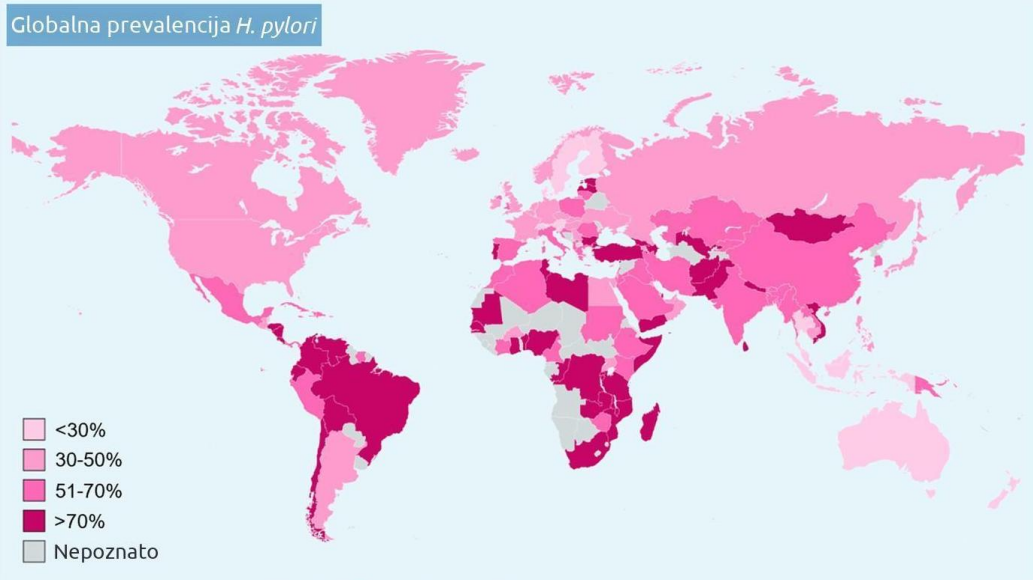
H. pylori veoma je rasprostranjen patogen te se smatra da je njime zaraženo više od 43 % svjetske populacije. Ovaj postotak se tijekom posljednjih godina smanjuje što je zapravo rezultat smanjenja širenja bakterije posljedično boljem načinu života i većem standardu

življenja.[2] Prevalencija *H. pylori* nešto je manja od tog postotka u razvijenim zemljama, dok je u zemljama u razvoju ona nešto veća.[12] Dva su razloga zašto ova bakterija tako uspješno preživljava, a to su jako dobra adaptacija na kisele uvjete želuca, njenog staništa te sve veća otpornost na antibiotike koji se koriste kod njene eradikacije. Prilagođavanje na stanište želuca je kod ove bakterije jedinstveno. Naime, bakterija za povećanje pH vrijednosti citoplazme i periplazme, koristi poseban mehanizam proizvodnje ureaze kako bi se kiselost smanjila. Također kako bi *H. pylori* izbjegla takve nepovoljne uvjete uvlači se u mukozno tkivo želuca. U tu svrhu bakterija je helikalnog oblika posjeduje brojne bičeve koji joj olakšavaju kretanje.[13] Skoro dva desetljeća od otkrića ove bakterije terapija koja se koristi u eradikaciji je bila poprilično jednostavna u odnosu na današnju terapiju. Za liječenje se najprije koristio jedan antibiotik, a nakon toga se uvela dvojna antibiotska terapija koja je bila uspješna u liječenju.[14,15] Zbog progresivnog razvoja otpornosti bakterije *H. pylori* na antibiotike, liječenje s godinama postaje sve izazovnije. Zbog toga se ovaj patogen danas svrstava u klasu I ljudskih kancerogena.[3] Također, jedan od važnijih faktora razvoja rezistencije može biti i sve veće ili nepravilno korištenje ili pretjerano korištenje antimikrobnih sredstava kod ljudi, životinja ili biljaka.[16] Posljedično, trenutna terapija predložena od strane svjetske gastroenterološke organizacije je kombinacija više vrsta antibiotika. Dvije najčešće terapije koje se koriste u eradikaciji *H. pylori* su: trojna terapija antibioticima PPI-AC (inhibitori protonskih pumpi, amoksicilin i klaritromicin) i četverovalentna terapija s bizmutom PPI-BTM (inhibitori protonskih pumpi, bizmut, tetraciklini i metronidazol).[17]

2.1.1. Epidemiologija H. pylori

H. pylori vjerojatno je jedna od najrasprostranjenijih bakterija u svijetu što je čini predmetom mnogih epidemioloških istraživanja današnjice.[18–21] Globalna studija iz 2017. godine pokazala je da je bakterijom *H. pylori* zaraženo više od polovice svjetskog stanovništva[22] no isti je časopis objavio novu studiju 2024. godine iz koje se može zaključiti kako se taj postotak smanjio te iznosi nešto više od 43 %.[2] U oba ova pregledna članka može se vidjeti da se zaraženost bakterijom sve više smanjuje u razvijenim dijelovima svijeta dok ona ostaje konstantno visoka kod stanovništva zemalja u razvoju. Također sveukupno manja zaraženost ovim patogenom mogla bi se objasniti sve boljim načinom života općenito te poboljšanim sanitarnim uvjetima u većini zemalja posljednjih desetljeća. Prevalencija se uvelike razlikuje kod odraslih i kod djece i adolescenata, gdje djeca i adolescenti imaju manji postotak

zaraženosti od odraslih, ali se trend smanjenja u zadnjem desetljeću pokazao samo kod odraslih.[2] Globalno je *H. pylori* najviše raširen u zemljama u razvoju u Južnoj Americi, Africi ili na Bliskom Istoku, gdje sam postotak varira od oko 24 % u Oceaniji do 70 % u Africi.[23] Na slici 2 može se vidjeti globalna prevalencija bakterije *H. pylori*.



Slika 2. Globalna prevalencija bakterije *H. pylori* (preuzeto i prilagođeni iz reference [24]).

U Hrvatskoj se također postotak prevalencije smanjio tijekom godina. Istraživanje iz 1997. pokazalo je da je zaraženost kod osoba između 20 i 70 godina iznosila oko 60 % gdje se taj postotak razlikovao prema dobnoj skupini ispitanika. Mlađi su imali manji postotak zaraženosti, a stariji su imali veći. Također je istraživanje napravljeno prema geografskoj rasprostranjenosti što je ukazivalo na veću zaraženost na južnom dijelu Hrvatske, a manju u centralnom i sjevernom dijelu.[25] Između 2016. i 2018. godine provedeno je istraživanje na pacijentima hospitaliziranim na traumatološkim odjelima u pet hrvatskih bolnica (jedna bolnica iz Osijeka, Rijeke i Splita te dvije bolnice iz Zagreba) koji nisu imali znakove peptične ili druge kronične bolesti te se kod njih ustanovila prevalencija od 46 %.[26] Veća studija, koja bi obuhvatila svo stanovništvo, uključujući djecu i adolescente, nije napravljena do pisanja ove disertacije.

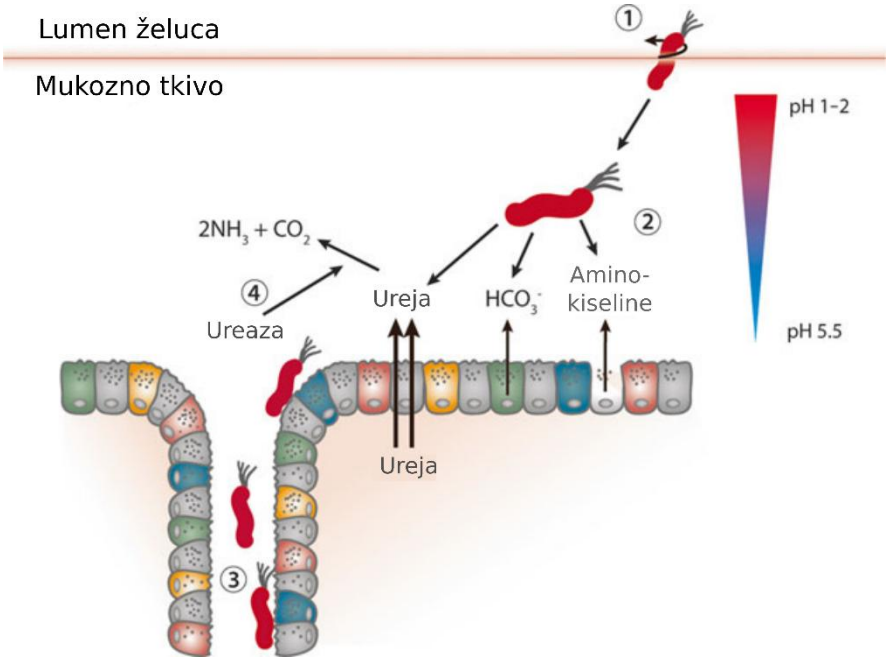
Rasprostranjenosti bakterije *H. pylori* doprinosi vrlo vjerojatno i lakoća i mnogostrukost puteva njezine transmisije. Najčešći način prijenosa ove bakterije je čovjek-čovjek prijenos, no ostaje nejasno zašto kod nekih ljudi ne dođe do kolonizacije ove bakterije, dok su drugi trajno zaraženi. Identificirane su tri glavne rute prijenosa *H. pylori*, a to su oralna transmisija, feko-

oralna transmisija i gastro-oralna transmisija.[27] Prijašnjih godina, najviše u zemljama u razvoju, jedna od glavnih ruta prijenosa bila je feko-oralna transmisija čemu je ponajviše mogao biti razlog konzumacija zaražene vode ili hrane.[28] No, sve češće, ova se ruta smatra sporednom zbog opažanja prijenosa oralnom transmisijom, kako u razvijenim zemljama, tako i u zemljama u razvoju. Također, prijenos između članova obitelji pokazao se kao glavni faktor u transmisiji, budući da se bakterija prenosi već u dječjoj dobi, najčešće s majke na dijete.[29,30]

2.1.2. Patofiziologija H. pylori

Bakterija *H. pylori* kolonizira želudac bolesnika što zahtjeva patološko istraživanje, no s obzirom da ovaj patogen ne utječe samo na želudac već i na ostatak organizma, problem je potrebno sagledati sa patofiziološkog aspekta. Patofiziologija je znanstvena disciplina medicine koja spaja znanje fiziološkog odnosno zdravog stanja organizma i patološkog odnosno nezdravog stanja organizma. Drugim riječima, pokušava objasniti funkcionalne promjene koje se zbivaju kod pojedinaca posljedično nekoj zarazi ili patološkom stanju.[31] Ljudski želudac, pri fiziološkom odnosno zdravom stanju organizma ima pH lumena od oko 1,4 te se on povećava prilikom unosa hrana na pH od oko 4.[32] Bakterija *H. pylori* pripada skupini neutralofila što znači da može preživjeti između pH 4 i pH 8 te se može razmnožavati između pH 6 i pH 8.[33] Kako bi izbjegla ovako kisele uvjete svog staništa, bakterija je razvila nekoliko načina obrane. Za povećanje pH vrijednosti citoplazme i periplazme, signalizira se proizvodnja ureaze kako bi se ta kiselost smanjila, koju pokušava izbjeći, kao što je već gore navedeno, uvlačenjem u mukozno tkivo želuca.[13] U sljedećim odlomcima detaljnije su opisani svaki od načina preživljavanja i aklimatizacije *H. pylori*.

Mobilnost *H. pylori* prva je i jedna od najvažnijih adaptacija ove bakterije. Sam rod i naziv ove bakterije, *Helicobacter*, govori o njenoj morfologiji, odnosno o njenom helikalnom obliku. Kako je gore navedeno, najčešći način prijenosa ove bakterije je oralnom transmisijom, što znači da bakterija iz želučanog sadržaja mora doći do mukoznog tkiva u želucu. Njen helikalni oblik i brojni bičevi omogućuju joj prolazak kroz mukoznu barijeru koja prekriva površinu želučanog epitela. Kod kolonizacije želučanog epitela, bakterija se suočava sa sadržajem koji izlučuju stanice epitela, te se zbog svog kemotaktičnog odgovora mogu kretati ka ili suprotno od njih.[34,35] Jednom kad se smjeste u pogodno okruženje, uključuju ostale načine obrane od stanišnih kiselih uvjeta. Cijeli proces prikazan je na slici 3.



Slika 3. Model kolonizacije bakterije *H. pylori* u mukozno tkivo želuca. (1) Helikalni oblik i bičevi omogućavaju ulazak bakterije u mukozno tkivo želuca. (2) Bakterija kemotaktički detektira ureju, hidrogenkarbonat i razne aminokiseline što joj omogućava kretanje prema površini želučanog epitela. (3) *H. pylori* se veže za receptore adhezina na površini stanica. (4) Pomoću mehanizma aklimatizacije bakterija uspijeva preživjeti u takvim iznimno kiselim uvjetima. (preuzeto i prilagođeno iz reference [10])

Jedan od glavnih mehanizama u obrani od kiselih uvjeta je enzim ureaza te je samim time i neophodan za bakterijsko preživljavanje. Proizvodnja ovog proteina je veoma konzervirana kod *H. pylori* te je zanimljivost da čak 10 % svih proteina ove bakterije čini upravo ureaza.[36] Aktivnost ovog proteina kod intaktne bakterije minimalna je kod neutralnog pH te se povećava kako se pH smanjuje, gdje doseže svoju maksimalnu aktivnost između pH 6,5 i pH 2,5. Iz navedenog se može zaključiti da je mehanizam djelovanja ove bakterije veoma koordiniran te je unos ureje u citoplazmu dozvoljen pri niskom pH, a kod višeg pH dolazi do zatvaranja odnosno blokiranja unosa ureje u stanicu. Također ovakav pristup je veoma opravdan budući da ureaza katalizira reakciju raspada ureje na ugljičnu kiselinu, a potom na ugljikov dioksid i amonijak koji u konačnici neutraliziraju okruženje bakterije, gdje se ona lakše kreće i veže na epitelne stanice mukoznog tkiva.[33] Uz ureju koju unosi u svoje stanice, bakterija *H. pylori* također ju i proizvodi. Jedna je od rijetkih bakterija za koje se zna da imaju, jedinstveno konfiguriran, ciklus ureje.[37] Posljednji korak ciklusa ureje je reakcija arginaze, odnosno kataliza

reakcije arginina u ornitin gdje kao nusprodukt nastaje ureja. Ornitin ponovo ulazi u ciklus, a ureu zatim razgrađuje ureaza na isti način kako je gore opisano. Ureaza je dakle zaslužna za inicijalnu kolonizaciju bakterije te također za održavanje kronične infekcije.[38] Zbog konstantnog prisustva, bakterija također izlučuje i razne citotoksične proteine koji posljedično uzrokuju nastanak raznih bolesti.

2.1.3. Bolesti povezane uz H. pylori

Sama kolonizacija bakterije *H. pylori* ne smatra se bolešću kao takvom, već infekcijom odnosno stanjem koje utječe na razvoj raznih kliničkih bolesti gornjeg gastrointestinalnog trakta ili drugih gastroduodenalnih poremećaja.[39] Prva bolest koja se javlja prilikom ulaska bakterije u želudac i njene kolonizacije u mukozno tkivo je akutni gastritis. Prvi znakovi gastritisa su pojava dispepsije u prvom tjednu inkubacije gdje simptomi postaju najintenzivniji između 9-og i 12-og dana.[40] Potom se sami od sebe utišavaju i većina ih nestaje nakon dva tjedna od početka infekcije.[41] Ipak, kad kolonizacija bakterije potraje duže infekcija se počinje smatrati kroničnom. Kod kroničnog gastritisa javlja se jasna korelacija promjene količine izlučene želučane kiseline i mjesta razvijanja upale. Tako postoje dva slučaja gdje se smanjuje ili povećava proizvodnja želučane kiseline te se posljedično javljaju pangastritis (zahvaća antrum i tijelo želuca) ili antralni gastritis (zahvaća samo antrum želuca).[39] Oba ova stanja mogu se dalje razviti u još teže poremećaje probavnog sustava. Kod pojave simptoma dispepsije te ako nije došlo do razvoja ulkusa, dijagnosticira se funkcionalna dispepsija. Ona je specifična prema tome što su prisutni svi simptomi tegoba gornjeg gastrointestinalnog trakta, ali nema nikakvih strukturnih promjena tkiva u tom istom području. Također se postavlja ovakva dijagnoza za sve neistražene ili neliječene simptome dispepsije.[42] Simptomi dispepsije prisutni su kod oko 20 % svjetske populacije (bez drugih gastrointestinalnih tegoba).[43] Ako se primijeti promjena u strukturi mukoznog tkiva koje prodire barem 0,5 cm do mišićne sluznice, riječ je o čiru želuca ili dvanaesnika (ovisno gdje se nalazi promatrano mukozno tkivo). Želučani ulkus nastaje kao posljedica pangastritisa te se promjene tkiva uočavaju najčešće na granici antruma i tijela želuca na manjoj krivini želuca.[44] Duodenalni ulkus najčešće se javlja na početnom dijelu dvanaesnika koji je najviše izložen kiselini želuca. Postotak oboljelih od peptičkog ulkusa u posljednje vrijeme je u opadanju.[45] Jadan od najgorih ishoda koji se događa kod pacijenata zaraženih bakterijom *H. pylori* je razvitak raka želuca te je upravo *H. pylori* potvrđena kao uzrok kod 60 % dijagnoza.[46] Klasificirana su dva tipa raka želuca: intestinalni i difuzni tip. Želučani

adenokarcinom intestinalnog tipa nastaje u nekoliko koraka gdje najprije nastaje kronični gastritis koji stupnjevito napreduje u atrofiju, crijevnu metaplaziju i displaziju te na kraju u rak crijevnog tipa.[47] Difuzni tip raka želuca se smatra da najčešće nastaje uslijed genetskih poremećaja te je manje povezan sa okolišnim čimbenicima kao što je to slučaj kod intestinalnog raka želuca.[48] Uz gore navedene bolesti infekcija *H. pylori* povezana je uz još mnoštvo dijagnoza poput želučanog MALT limfoma, gastroezofagealnog refluksa te drugih van crijevnih bolesti poput astme, rozaceje, anemije uzrokovane nedostatkom željeza, idiopatske trombocitopenične purpure, itd.[39]

2.1.4. Antibiotska rezistencija

Otkriće bakterije *H. pylori* bilo jedno od važnijih otkrića u polju gastroenterologije. Bolestima koje su do tada bile neizlječive, poput peptičkog ulkusa, znao se uzrok koji se odjednom mogao tretirati antibioticima. No, korištenje jednog antibiotika 1990-ih godina više nije bilo dovoljno što je zahtijevalo uvođenje novog dvojnog tretmana.[14] Takva terapija koristila se do otprilike prije 10 godina kada je bilo potrebno uvesti novi način eradikacije bakterije zbog njenog progresivnog razvoja rezistencije na sve više antibiotika.[15] Postoje razni mehanizmi kako bakterija može razviti rezistenciju na antibiotike, a neki od njih su:[49]

1. Unošenje mobilnog genetskog materijala (koji sadrži gene za rezistenciju) u stanicu. Bakterije mogu razmijeniti genetski materijal na više načina. Konjugacija je izmjena plazmidne DNA te je taj mehanizam kodiran u samom plazmidu. Izmjena plazmida može se desiti samo horizontalnim prijenosom te bakterije akceptori postaju rezistentne u kratkom vremenskom periodu. Uz plazmide postoje i transpozoni i integroni koji su također genetski fragmenti koji obično nose antibiotsku rezistenciju. Za njihov međugenomski prijenos potrebni su bakteriofazi te se sam proces zove transdukcija. U bakterijskom okruženju također postoje razni dijelovi genetskog materijala od mrtvih bakterija koji se, u specifičnom trenutku bakterijskog životnog ciklusa, mogu integrirati u samu bakteriju. Ovakvim procesom transformacije bakterije također mogu steći antibiotsku rezistenciju.
2. Točkastim mutacijama koje modificiraju strukturu antibiotske mete. Takve mutacije obično su dosta česte u bakterijskom DNA te se događaju zbog mistranslacijskih pogrešaka. Također mnoge od točkastih mutacije nemaju utjecaja niti na funkciju niti na strukturu proteina, dok rjeđe, one mogu promijeniti samu fiziologiju bakterija pa tako

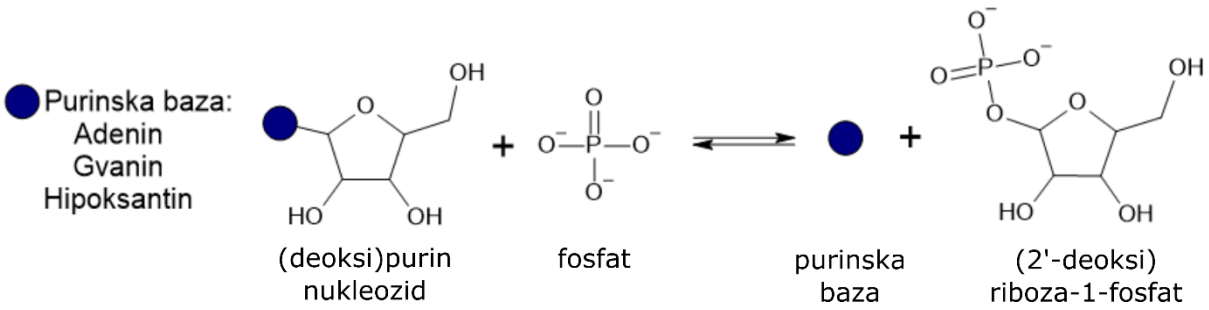
i njenu antibiotsku rezistenciju. Kod *H. pylori* točkaste mutacije su esencijalan mehanizam stjecanja rezistencije te se tako stečena rezistencija prenosi vertikalnim prijenosom na novonastale bakterije.

Jedan od rjeđe primijećenih mehanizama obrane od antibiotika kod bakterije *H. pylori* je prekomjerna ekspresija refluksnih pumpi. Kako bi se postigla stanična homeostaza i izbjegla visoka koncentracija inhibitora, refluksne pumpe služe za izbacivanje antibiotika van stanica, čime se izbjegava njihovo djelovanje.[50]

Poznato je šest vrsta antibiotika koji se mogu koristiti kod liječenja infekcije *H. pylori*. Problem rezistencije najveći je kod dvije skupine: makrolidi i kinolonski antibiotici. Kod obje ove skupine antibiotika dolazi do točkaste mutacije njihovih meta u domaćinu čime se stječe rezistencija na korišteni antibiotik.[51,52] Kod iduće skupine antibiotika koju predstavljaju amoksicilin, tetraciklin i rifabutin rijetko su pronađeni sojevi *H. pylori* koji su rezistentni na njih što indicira da se sva ova tri antibiotika mogu koristiti kod liječenja infekcije. Šesta skupina antibiotika je 5-nitroimidazola koji se inače koriste kod tretiranja anaerobnih bakterija. Neki spojevi, pripadnici ove skupine, pokazali su se dobrima kod tretiranja mikroaerofilnih bakterija poput *H. pylori*. Trenutno preporučeno liječenje infekcije bakterijom *H. pylori* ima dvije glavne smjernice, a to su trojna terapija i četverovalentna terapija. Trojna terapija antibioticima PPI-AC sastoji se od inhibitora protonskih pumpi, amoksicilina i klaritromicina, a četverovalentna terapija s bizmutom PPI-BTM se sastoji od inhibitora protonskih pumpi, bizmuta, tetraciklina i metronidazola.[17]

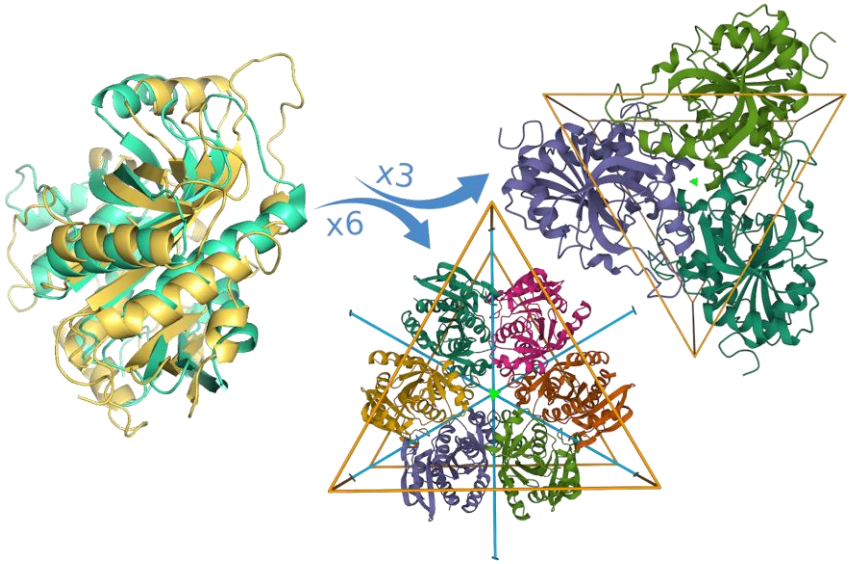
2.2. Fosforilaza purinskih nukleozida (PNP)

Purin-nukleozid:fosfat riboziltransferaza, odnosno fosforilaza purinskih nukleozida, odnosno PNP (E.C. 2.4.2.1) enzim je koji katalizira reverzibilnu reakciju fosforolitičkog cijepanja glikozidne veze kod purin (deoksi)ribonukleozida (slika 4). Dakle, PNP pripada podskupini enzima glikoziltransferaza zato što se ribozilna skupina prenosi s (deoksi)purin nukleozida na fosfat.[4,5]



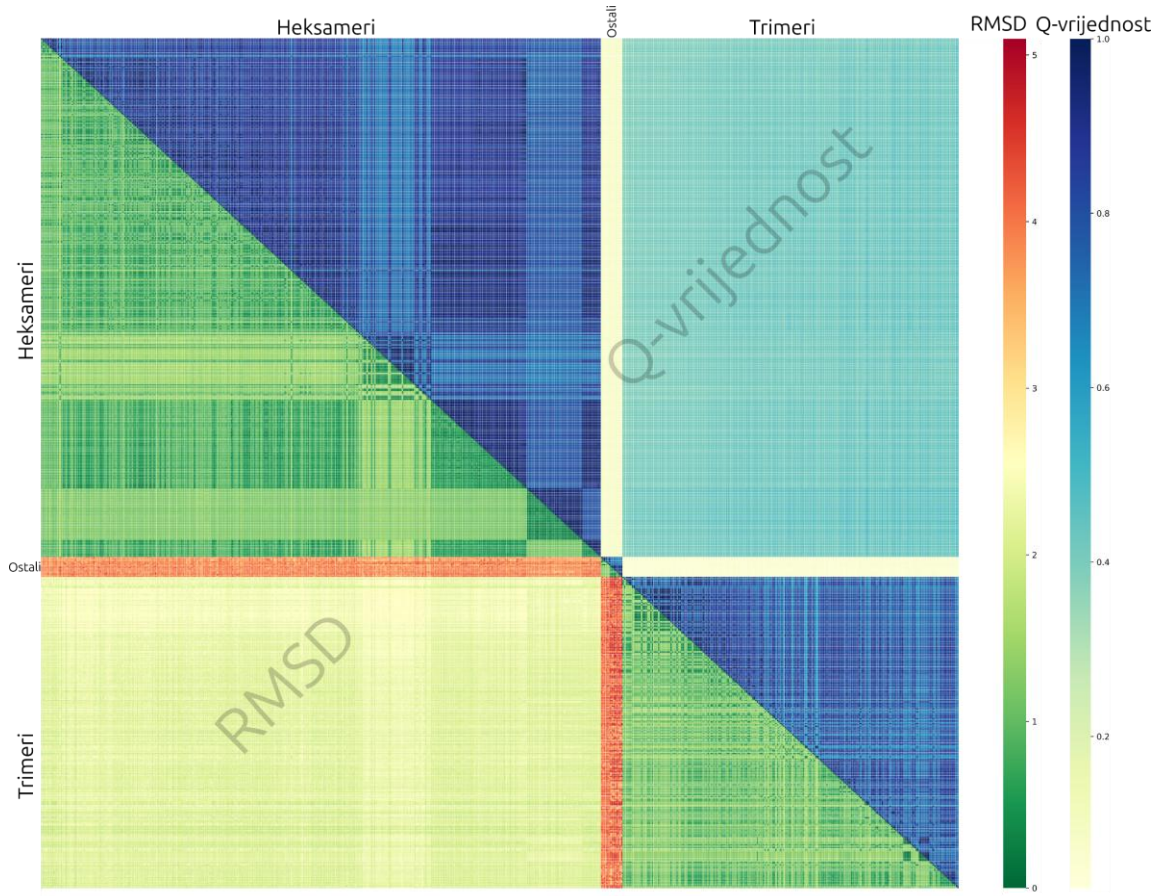
Slika 4. Reakcijska shema enzima PNP.

Poznata su dva tipa PNP-a identificirana u raznim organizmima. Prva skupina enzima su PNP male molekulske mase odnosno homotrimeri koji pretežito kataliziraju 6-oksopurine i pripadajuće nukleozide. Većinom su izolirani iz eukariotskih organizama, dok se rijeđe mogu pronaći i kod prokariotskih organizama. Druga skupina enzima su PNP velike molekulske mase odnosno homoheksameri te oni imaju širi spektar katalitičke aktivnosti budući da prihvaćaju i 6-oksopurine i 6-aminopurine te pripadajuće nukleozide kao supstrate. Najčešće ih se može pronaći kod mikroorganizama. Obje vrste PNP-a prihvaćaju gvanin-, hipoksantin- i adenin-(2'−deoksi)ribonukleozid kao supstrat.[53] Aminokiselinske sekvence trimera i heksamera imaju sličnost od samo 20-30 %. Stoga je vrlo zanimljiva velika sličnost trodimenzijskih struktura njihovih monomernih podjedinica (slika 5). No unatoč velikoj sličnosti trodimenzijskih struktura pojedinih monomera, potpuno je različita njihova organizacija u kvaternu strukturu. Poznato je da je oligomerizacija neophodna za njihovu funkciju, te je poznato da ni trimerni ni heksamerni PNP ne mogu vršiti svoju katalitičku funkciju kao monomeri. Odnosno, i jedna i druga forma, neophodno moraju biti u svom oligomernom stanju.[54,55]



Slika 5. Strukturno sravnjene monomernih podjedinica trimernog PNP-a (žuto obojena struktura) i heksamernog PNP-a (zeleno obojena struktura). Vidljivo je odlično slaganje monomera dok je udruživanje u kvaternu strukturu potpuno različito. Naime kod eukariota je PNP trimer, dok je kod prokariota heksamer (preuzeto i prilagođeno iz reference [7]).

Evolucijski gledano, heksamerni PNP iz nižih organizama preteče su trimernim PNP-ovima iz viših organizama. Zanimljiva je činjenica da su ovi enzimi sačuvali svoju originalnu funkciju i važnost unatoč velikoj promjeni u organizaciji kako kvaterne strukture, tako i primarne sekvence. Pretpostavlja se da je njihova funkcija sačuvana baš zbog sličnosti trodimenzijskih struktura njihovih monomera (40-60 %) koja je dvostruko veća od sekvencijalne sličnosti monomera (20-30 %).[7] Također, analiza strukturnog sravnjivanja monomernih lanaca svih određenih struktura u makromolekulskoj bazi podataka PDB[56] je to pokazala. Radi usporedbe svih struktura heksamera i trimera koje su poznate u PDB-u napravljena su sveobuhvatna međusobna sravnjivanja monomernih podjedinica pomoću programa GESAMT.[57] Kao jedna od mjera slaganja trodimenzijskih struktura monomera koristi se Q-vrijednost (engl. *Q-score*).[58] Izračunate su velike Q-vrijednosti (između 0,8 i 1,0) i male vrijednosti RMSD-a (između 0,0 i 1,5) kako za monomere heksamernih PNP-ova, tako i za monomere trimernih PNP-ova. Rezultati za strukturno sravnjene monomera heksamera i trimera pokazuju da Q-rezultat iznosi između 0,4 i 0,6, dok RMSD vrijednost iznosi između 2,0 i 3,0 (slika 6) što zapravo može poslužiti kao potvrda da su funkcija i mehanizam djelovanja ovog enzima očuvani upravo zbog očuvane trodimenzijske strukture monomera.

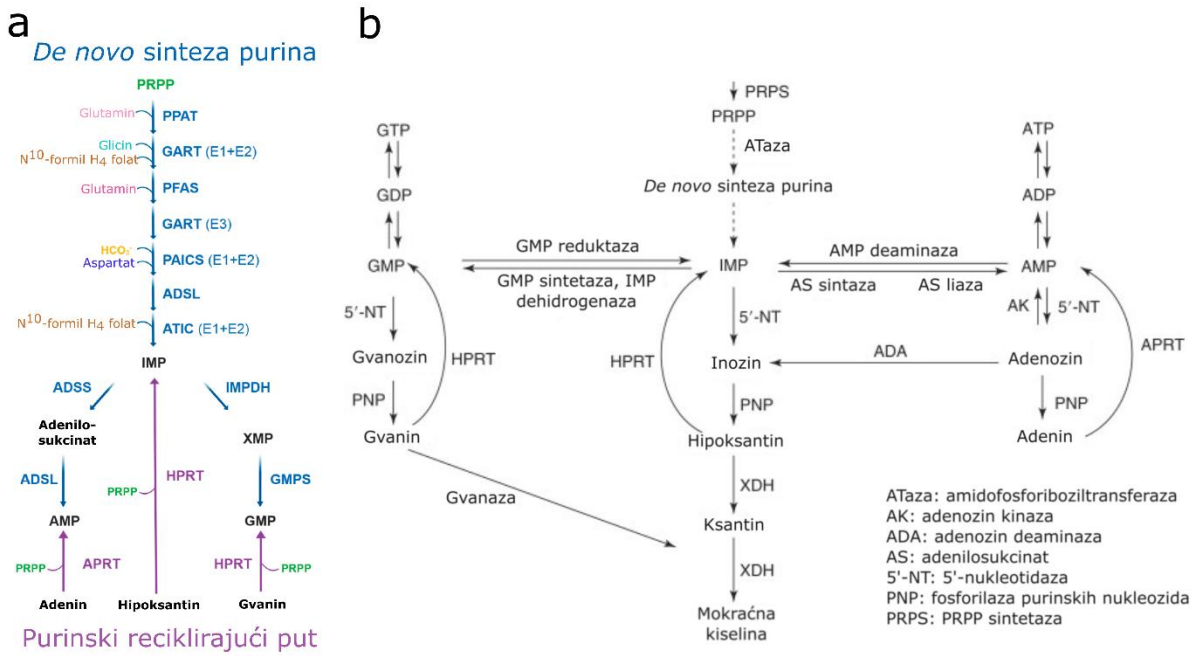


Slika 6. Distribucija rezultata Q-vrijednosti (gornji desni trokut) i RMSD-a (donji lijevi trokut) za strukturo sravnjenje svih lanaca monomera eksperimentalno riješenih struktura enzima PNP gdje se jasno odvajaju heksameri od trimera. (preuzeto i prilagođeno iz reference [7])

2.2.1. Metabolizam purina kod H. pylori

Purini, uz pirimidine, jedni su od esencijalnijih građevnih jedinica u živom svijetu. Organizmi za opstanak trebaju purine kao gradivne jedinice njihovog genetskog materijala. Zbog velike važnosti u velikoj većini slučajeva imaju dobro razvijene sustave njihove biosinteze kojima ih mogu sami sintetizirati. Uz to što su sastavni dio DNK i RNK, sudjeluju u mnogim signalnim putevima, kofaktori su kod brojnih enzimskih reakcija, služe kao molekule za pohranu kemijske energije, itd.[59] Sinteza purina *de novo*, odnosno njihova biosinteza odvija se na način da stanice koriste aminokiseline (glutamin, glicin, asparaginsku kiselinu) i druge male molekule/ione (fosforibozil pirofosfat, hidrogenkarbonat, N^{10}-formil H_4 folat) za sintezu purinskog prstena. Nakon dugog reakcijskog puta od 10 koraka sintetiziran je inozin monofosfat koji se zatim dalje prevodi u gvanozin monofosfat i adenozin monofosfat te njihove derivate (slika 7a).[60] Drugi

način sinteze, ako kod nekog organizma djelomično ili u potpunosti nedostaju geni za proteine koji sudjeluju u *de novo* sintezi, je purinski reciklirajući put (engl. *purine salvage pathway*). Kod ovog načina sinteze dolazi do recikliranja već sintetiziranih baza te njihova ponovna upotreba. Također stanice mogu unositi dijelove razgrađenih ili mrtvih dijelova drugih organizama što se zatim isto reciklira i ponovo koristi. Fosforilaza purinskih nukleozida katalizira reakcije pretvorbe gvanozina u gvanin, inozina u hipoksantin te adenozina u adenin. Ovaj također složeni reakcijski put prikazan je na slici 7b.[6]



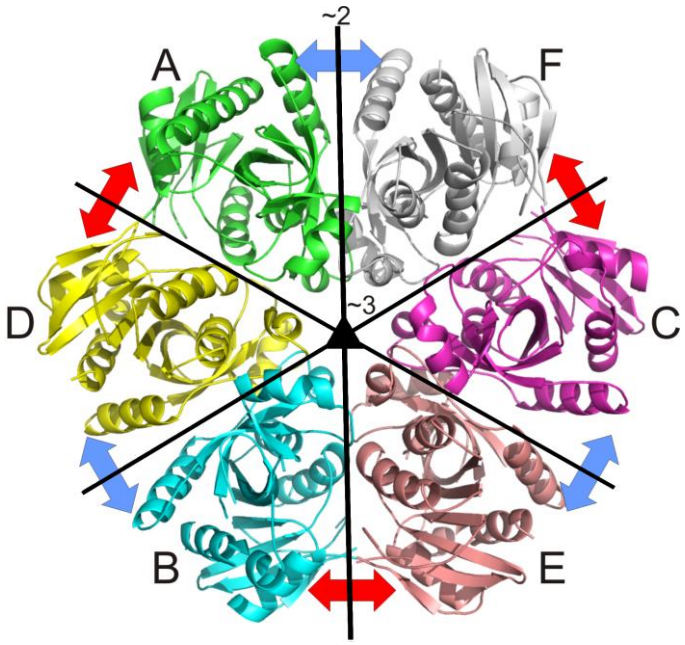
Slika 7. Dva načina sinteze purina. (a) Shematski prikaz *de novo* biosinteze purina te enzimi koji sudjeluju. (preuzeto i prilagođeno iz reference [60] (b) Purinski reciklirajući put te enzimi koju sudjeluju u katalitičkim reakcijama. (preuzeto i prilagođeno iz reference [6])

Utvrđeno je da bakteriji *H. pylori* nedostaju geni za sintezu purina *de novo*. Bioinformatičkom analizom je potvrđeno da kod nje nedostaju geni za skupinu enzima koji su zaslužni za sintezu inozin monofosfata (IMP) koji je preteča srodnih i važnijih gvanozin monofosfata (GMP) i adenozin monofosfata (AMP). To znači da se *H. pylori* oslanja isključivo na purinski reciklirajući put i unos već gotovih purina ili njihovih derivata u stanice. Nadalje, delecijom gena *deoD* koji je zaslužan za ekspresiju enzima fosforilaze purinskih nukleozida utvrđena je nemogućnost rasta bakterije na mediju bogatom nukleozidima ili purinskom bazom adenin.[61]

Daljnje istraživanje purinskog reciklirajućeg puta i enzima PNP od velike je važnosti zbog potencijalnih novih terapija protiv ovog sve opasnijeg patogena. Budući da se PNP nalazi u svim reakcijskim putevima kod kojih dolazi do sinteze purinskih baza (slika 5B), može se reći da je on jedan od važnijih enzima kod ove bakterije što ga čini veoma potentnom metom novo dizajniranih lijekova. Ovakvim ciljanim djelovanjem lijekova PNP postaje limitirajući faktor u samom preživljenju bakterije *H. pylori* budući da ona nema drugih načina i izvora purina koji su neophodni za njen život.[62]

2.2.2. Struktura PNP-a iz H. pylori

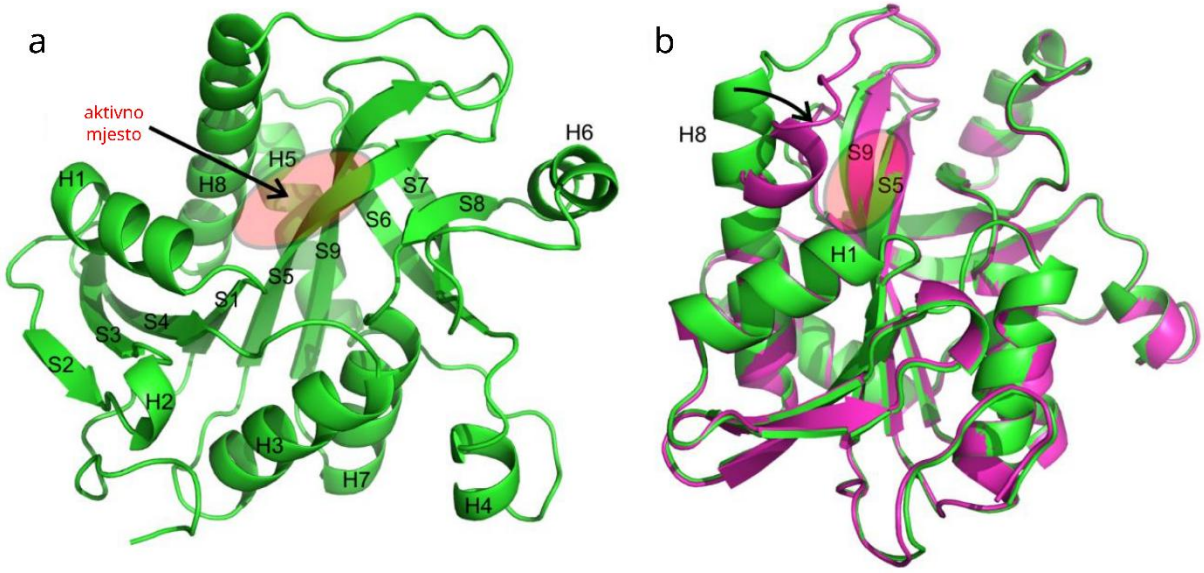
Kao što je već ranije opisano, PNP kod prokariotskih organizama gradi heksamer, što znači da je njegova kvaterna struktura sačinjena od šest jednakih monomernih podjedinica. Kod PNP iz *H. pylori* svaki monomer se sastoji od 233 aminokiseline. Konačna struktura ovog enzima nastaje tako da se monomeri prvo udružuju u dimere i to tako da jedan monomer upotpunjava aktivno mjesto onog drugog donirajući mu dvije aminokiseline: histidin His4 i arginin Arg43.



Slika 8. Cjelokupna struktura jednog homoheksamera enzima PNP koja prikazuje položaje osi simetrije (jedna trigira označena je trokutom, a tri digire označene su punom crtom) na samoj strukturi. Crvenim strelicama je označena intradimerno sučelje, dok je plavim strelicama označena interdimerno sučelje. Slovima su označeni pojedinačni monomerni lanci. (preuzeto i prilagođeni iz reference [63])

Tako formiran dimer ima približnu simetriju rotacije za 180° odnosno os drugog reda (digiru). Tri tako nastala dimera spajaju se u heksamer preko osi trećeg reda (trigire) što daje približnu simetriju rotacije od oko 120° koja je okomita na sve tri osi drugog reda.[63] Na slici 8 može se vidjeti raspodjela monomera u prostoru gdje su približno označene simetrijske osi (~2 označava položaj digira, ~3 označava položaj trigire) koje stvaraju kvaternu strukturu heksamera. Također su na slici slovima označeni monomeri te oni grade dimerne strukture na sljedeći način: AD dimer, BE dimer i CF dimer.

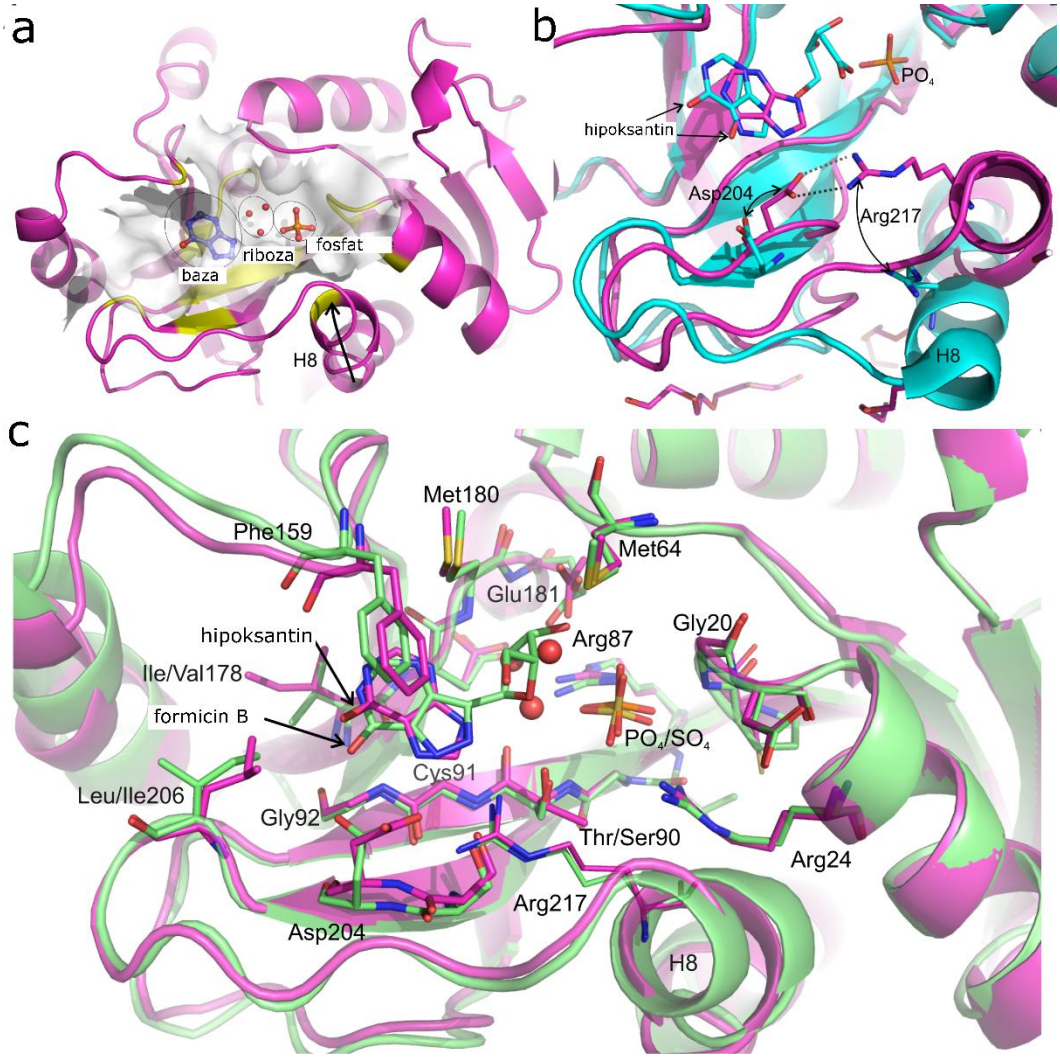
Kod svakog organizma sekvenca nekog enzima se manje ili više razlikuje od njemu srodnih enzima te se to isto odnosi i na PNP iz bakterije *H. pylori*.[53] Trodimenzijska struktura jednog monomera ovog enzima sastoji se od devet *β*-ploča i osam α-zavojnica. Aktivno mjesto proteina omeđeno je *α*-zavojnicama H1 i H8 te *β*-pločama S5 i S9 (slika 9) te je sastavljeno od tri glavna vezna mjesta, regije za vezanje fosfata, riboze i purinske baze.



Slika 9. Prikaz praznog aktivnog mjesta enzima PNP iz bakterije *H. pylori* te sekundarnih struktura koje ga omeđuju u (a) otvorenoj konformaciji i (b) zatvorenoj konformaciji (preuzeto i prilagođeno iz reference [63]).

Vezna šupljina u kojoj se smješta fosfat najdublje je smještena u samoj strukturi te ju čine aminokiseline Gly20, Arg24, Arg87, Thr90 te Arg43 iz susjednog monomera. Zatim slijedi vezna šupljina za ribozu koju čine aminokiseline Met64, Arg87, Thr90, Met180 i Glu181.

Najdalje mjesto i ono najviše izloženo otopini je vezna šupljina za purinsku bazu koju čine aminokiseline Cys91, Gly92, Phe159, Ile178 i Ile206 (slika 10).



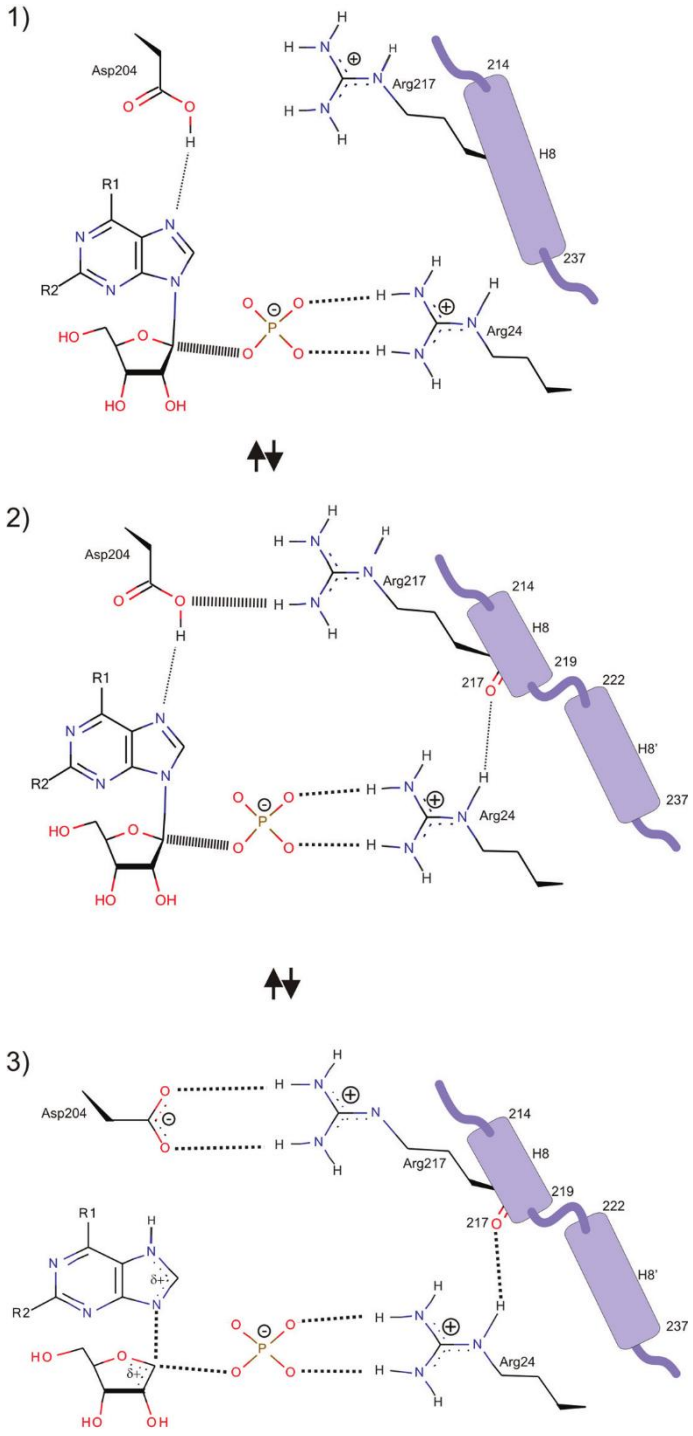
Slika 10. Struktura aktivnog mjesta enzima PNP iz *H. pylori*. (a) Prikaz tri regije, odnosno veznih mjesta za fosfat, ribozu i bazu. (b) Konformacijska promjena segmentacije α-zavojnice H8 te razlika otvorenog (plava struktura iz bakterije *B. subtilis*) i zatvorenog (roza struktura iz bakterije *H. pylori*) aktivnog mjesta. (c) Sravnjene strukture zatvorenih aktivnih mjesta kod *E. coli* (zeleno) i *H. pylori* (rozo) te prikaz aminokiselina koje se nalaze u blizini liganada. (preuzeto i prilagođeno iz reference [63])

Također u aktivnom mjestu sudjeluje i aminokiselina His4 iz susjednog lanca. Ona se nalazi na početku lanca u petlji koja je vrlo fleksibilna te stoga može, ali i ne mora koordinirati supstrate enzima. Prilikom odvijanja reakcije, aktivno mjesto PNP-a doživljava strukturnu promjenu,

odnosno zatvaranje aktivnog mjesta pa se tako razlikuju dvije konformacije svakog monomera: otvorena i zatvorena. Strukturna promjena dešava se kod H8 α-zavojnice koja se segmentira u dvije manje α-zavojnice odvojene kratkom petljom između aminokiselina 219 i 222 što čine glutamat Glu219, serin Ser220, fenilalanin Phe221 i aspartat Asp222. Ovakva konformacijska promjena približava još dvije aminokiseline (Asp204 i Arg217) koordinacijskoj sferi liganada koji zatim sudjeluju i u katalitičkoj reakciji.[63,64]

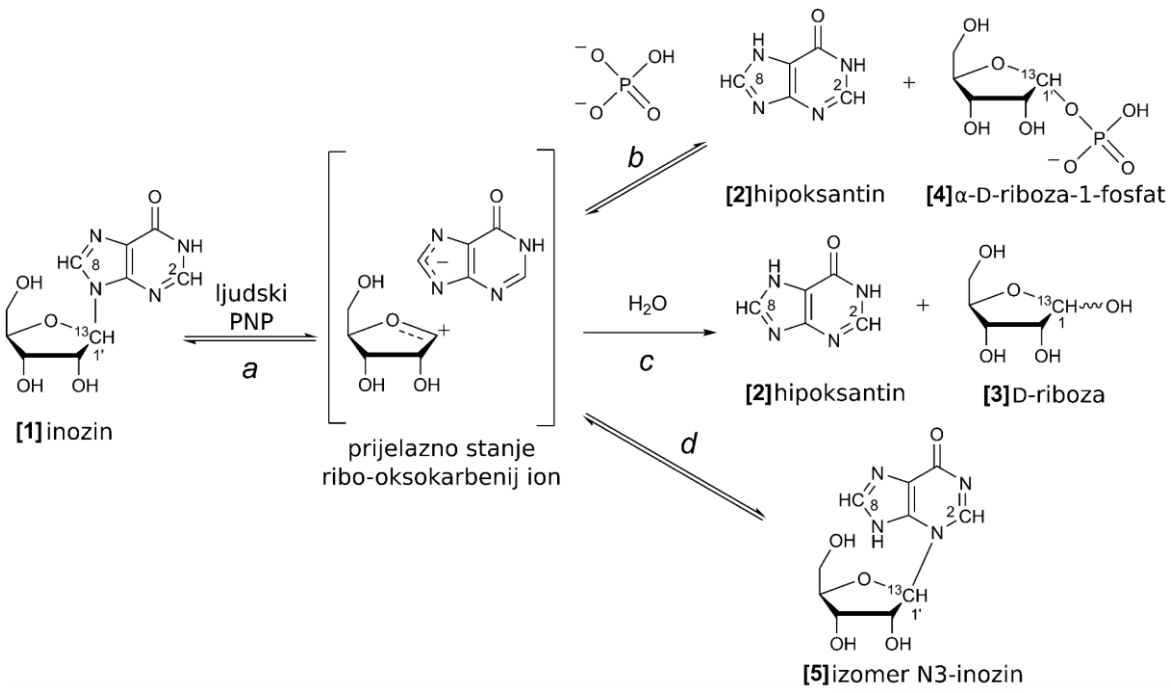
2.2.3. Mehanizam katalize

Prema reakciji sa slike 4 koju katalizira enzim PNP, jasno je da se u aktivno mjesto enzima moraju smjestiti dva supstrata kako bi se reakcija uspješno odvijala, neovisno o smjeru u kojem se reakcija odvija. Sam mehanizam navedene reakcije nije do kraja rasvijetljen, iako ih je nekoliko predloženo. Kod ljudskog (trimernog) enzima PNP predložena su tri različita mehanizma, koja su u suštini ili nukleofilna supstitucija ili nukleofilna adicija. Kao najvažnije aminokiseline u sva tri mehanizma spominju se His86 i Asn243. Histidin djeluje kao kiselo/bazni katalizator te se također nalazi u katalitičkoj trijadi kod jednog od predloženih mehanizama dok asparagin djeluje kao stabilizator nastalog međuprodukta. [65–67] Kod prokariota predložena su trenutno dva veoma slična mehanizma, a osnovna razlika je broj aminokiselina koje sudjeluju u reakciji. Različit broj aminokiselina posljedica je različitih struktura PNP-ova prema kojima je mehanizam predložen budući da su određene trodimenzijske strukture enzima izoliranih iz različitih organizama (*P. falciparum* i *E. coli*). Oba mehanizma predlažu da se reakcija odvija samo ako dođe do konformacijske promjene iz otvorenog aktivnog mjesta u zatvoreno što supstratima približava aspartat koji se nalazi na S9 β-ploči netom prije H8 α-zavojnice (ekvivalent tom aspartatu kod *H. pylori* je Asp204). Taj aspartat zatim može protonirati N7 poziciju purinske baze koja se nalazi u aktivnom mjestu čime se cijepa glikozidna veza između riboze i baze. U najdubljoj regiji gdje je smješten fosfat nalazi se nekoliko arginina koji stabiliziraju negativan naboj na fosfatu. Nastali kation na ribozi i fosfatni anion reagiraju pri čemu dolazi do regeneracije početnog aspartata i nastanka produkata.[64,68] Na reakcijskoj shemi slike 11 mogu se vidjeti koraci navedene reakcije.



Slika 11. Predloženi mehanizam reakcije katalizirane enzimom PNP iz bakterije *E. coli*. 1) Asp204 ostvaruje vodikovu vezu s purinskom bazom. 2) Događa se konformacijska promjena segmentacije H8 α-zavojnice te se Asp204 i Arg217 približavaju supstratima. 3) Protonacija purinske baze na N7 poziciji čime nastaje pozitivno nabijeno prijelazno stanje koji omogućava nastanak krajnjih produkata. (preuzeto i prilagođeno iz reference [4])

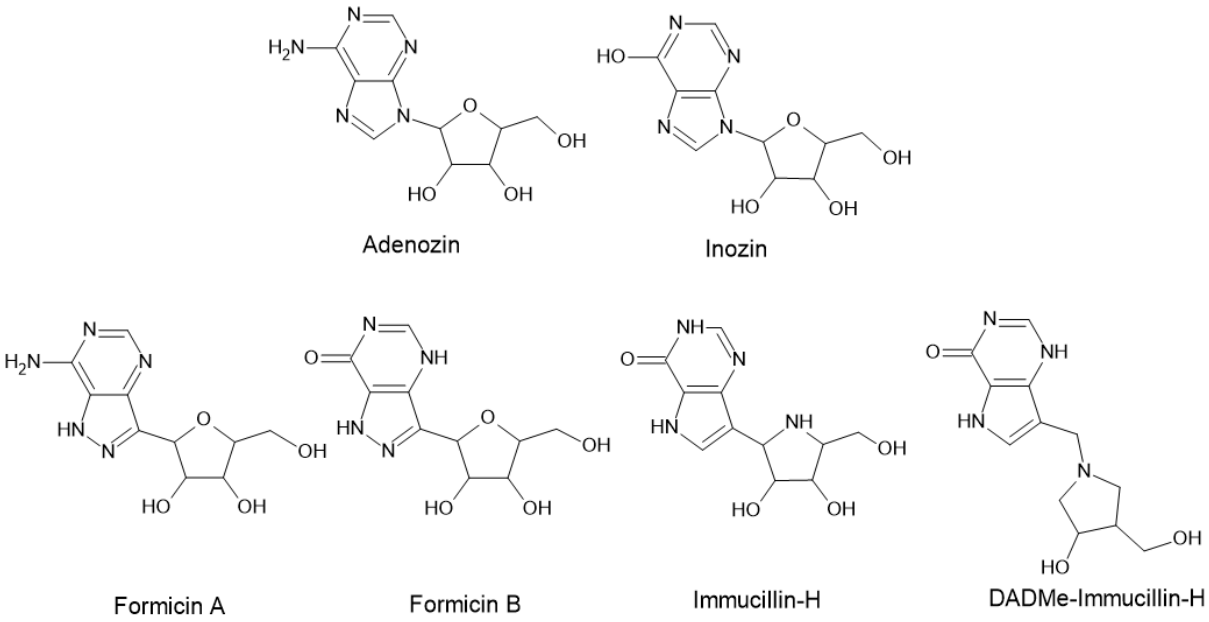
Kod ljudskog enzima PNP također je istraživana i njegova supstratna nespecifičnost, odnosno mogućnost katalitičke aktivnosti s drugim supstratima. Ova reakcija započinje kiselom katalizom nukleozida koja rezultira pucanjem veze između riboze i baze čime nastaje prijelazno stanje ribo-oksokarbenij ion. Tako nastalo prijelazno stanje ima tri različita reakcijska puta i samim time različite produkte. U prisustvu hidrogen fosfata dolazi do gore već opisane reakcije i nastajanja riboze-1-fosfata i hipokstantina. Drugi potencijalni put je ireverzibilna reakcija s vodom, koja konkurira prethodnoj reakciji, a posebice kod manjih koncentracija fosfata u vodenom mediju. Treća, ne toliko izražena, ali moguća reakcija je izomerizacija, odnosno ponovno spajanje riboze i baze na njenom drugom dušikovom atomu, odnosno nastajanje izomera N3-inozina. Također, moguća je i povratna reakcija čime ponovo nastaje nukleozid.[69] Svi ovi reakcijski putovi prikazani su na reakcijskoj shemi slike 12. Opisan reakcijski profil također je istražen i kod goveđeg enzima PNP (brzina reakcije $7{,}5 \cdot 10^{-4}$ s^{-1}) kod kojeg je hidrolitička reakcija nešto brža od one kod reakcije ljudskog enzima PNP (brzina reakcije $5 \cdot 10^{-6}$ s^{-1}).[70,71] Sve ovo ukazuje da su potrebna dodatna istraživanja te također biokemijska analiza aktivnosti prokariotskih enzima, koji bi zbog strukturne sličnosti mogli pokazati slične rezultate.



Slika 12. Reakcije katalizirane enzimom PNP te prijelazno stanje koje ima više potencijalnih smjerova za daljnji razvitak same reakcije. (preuzeto i prilagođeno iz reference [69])

2.2.4. Inhibitori fosforilaze purinskih nukleozida

Inhibitori se najčešće konstruiraju tako da oponašaju prijelazno stanje koje nastaje tijekom enzimske reakcije te se takvi nazivaju analozima prijelaznog stanja.[72–74] Kod enzima PNP najčešće se priređuju upravo analozi prijelaznog stanja budući da oni pokazuju najveću inhibiciju. Inhibitori nekih eukariotskih i prokariotskih organizama, dobro su poznati i istraženi. Jedni od prvih i najviše istraženih inhibitora PNP-a su forodesin odnosno Immucillin-H te DADMe-Immucillin-H kod kojih su izmjerene pikomolarne vrijednosti konstante disocijacije. Ova istraživanja potvrđena su kod tri različita PNP-a: goveđeg (*Bos taurus*), koji je trimerni PNP, ljudskog (*Homo sapiens*) koji je trimerni PNP te malarijskog (*Plasmodium falciparum*) koji je heksamerni PNP.[74] Kod heksamernih PNP-ova također su poznati inhibitori koji se nazivaju formicini, odnosno formicin A i formicin B. Ovi inhibitori razlikuju se od prethodnih zbog svoje specifičnosti prema heksamernim PNP-ovima budući da strukturno odgovaraju njihovom aktivnom mjestu, odnosno prema njemu su dizajnirani. Tom specifičnošću se isključuje mogućnost toksičnosti prema npr. ljudskom enzimu ako bi se ovim inhibitorima tretirali neki patogeni čiji je domaćin čovjek.[53,75] Svi ovi inhibitori su analozi nekog od supstrata PNP-a poput inozina ili adenozina te se sličnost prirodnih supstrata i inhibitora može vidjeti na slici 13. Ono po čemu se ti inhibitori razlikuju su razne modifikacije koje se ugrađuju na molekule supstrata (inozina ili adenozina) kako bi se postigla razna poboljšana svojstva poput bolje inhibicije, smanjene citotoksičnosti ili u krajnjem slučaju veoma selektivna inhibicija enzima točno određenog organizma. Aktivna mjesta u PNP-u raznih organizama imaju sličnu infrastrukturu no ipak postoje razlike koje omogućavaju optimizaciju priređivanja inhibitora. Za organizam *H. pylori* nema dovoljno istraživanja koja bi upućivala na konkretne molekule kao inhibitore enzima PNP. Nekoliko eksperimenta pokazali su se kao dobar način istraživanja potencijalnih inhibitora za PNP, a samim time i *H. pylori*. Pokazano je da su analozi purina kod kojih se supstituiraju pozicije 2 i/ili 6 dobri inhibitori sa mikromolarnom vrijednosti konstante disocijacije. Uz inhibiciju samog PNP enzima, također su prikazani i obećavajući rezultati inhibicije rasta *H. pylori* bakterijske kulture.[8,62,76]

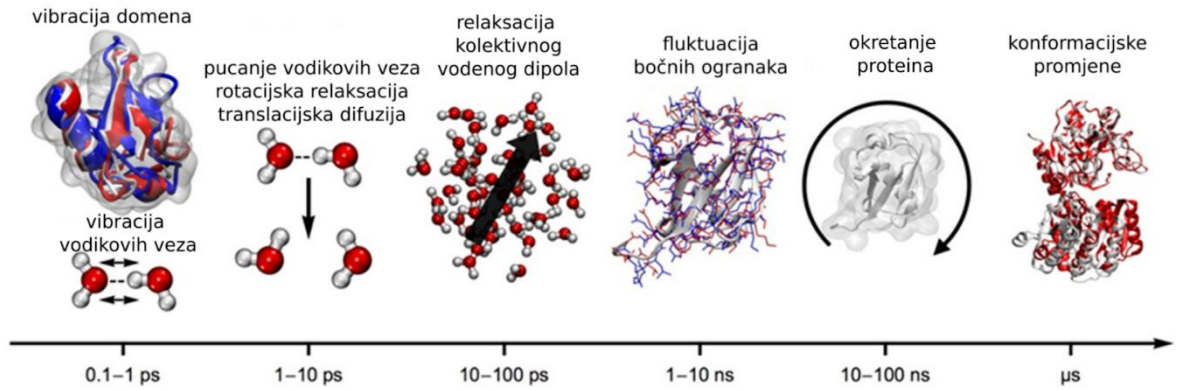


Slika 13. Strukture molekula adenozina i inozina te njihovih analoga (formicin A, formicin B, Immucillin-H i DADMe-Immucillin-H) koji predstavljaju inhibitore enzima HpPNP.

2.3. (Molekulska) dinamika proteina

Biološke makromolekule predmet su mnogih istraživanja zbog svoje izuzetne važnosti u živom svijetu. Proteini, uz nukleinske kiseline svakako spadaju u najvažnije gradivne blokove živog svijeta. Pritom je uloga proteina mnogostruka: oni mogu biti gradivni elementi, transportni, regulatorni, ili pak potpomagati kemijske reakcije. Takva raznolikost uloga proteina omogućena je nezamislivo velikim brojem načina na koji se dvadeset gradivnih aminokiselina mogu međusobno kombinirati. Enzimima nazivamo proteine koji imaju katalitičku aktivnost, tj. ubrzavaju odvijanje kemijskih reakcija, često i za mnogo redova veličina. Mnoge reakcije u živom svijetu ne bi uopće bile moguće da nisu katalizirane enzimima. Ono što čini enzime jedinstvenima je njihova trodimenzijska struktura i ona uvelike određuje njihovu specifičnost prema njihovim supstratima. Što je uži krug supstrata koji mogu sudjelovati u reakciji koje katalizira određeni enzim, to je njegova kemijska specifičnost veća.[77] Razumijevanjem detaljne strukture enzima na atomskoj razini bavi se strukturna biologija, a otkriće da se proteini mogu kristalizirati te da se posljedično može odrediti njihova trodimenzijska struktura pomoću rendgenske kristalografije, uvelike je promijenilo svijet biologije i kemije.[78,79] Rendgenska kristalografija je nesumnjivo temeljna metoda za proučavanje molekula na atomskoj razini. No

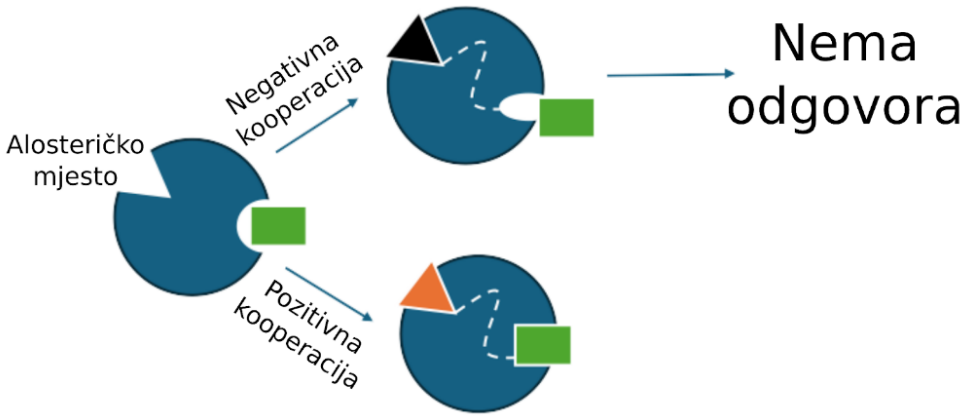
ono što tom metodom dobivamo je vremenski uprosječena slika položaja atoma molekule koju proučavamo, tj. kristalografija kao metoda ne daje gotovo nikakav uvid u dinamiku samih molekula. Jedni način na koji kristalografski vidimo da su neki dijelovi molekule dinamični, je taj da su ti dijelovi previše fleksibilni i nije moguće odrediti položaje tih dijelova molekule. To dodatno ovisi o veličini same molekule i maksimalnoj rezoluciji prikupljenih difrakcijskih podataka, ali jednostavno znači da je kristalografija nepovoljna metoda za praćenje dinamike molekula.[80] Kako bi se doskočilo ovom problemu, da dinamika proteina nije vidljiva u rendgenskom eksperimentu, koriste se računalne simulacije molekulske dinamike. Pored znanja o strukturi nekog proteina, iz koje se proučavaju strukturni elementi, aktivno mjesto te samim time i funkcija tog proteina, potrebno je proučavati i njegovu dinamiku. Za njegovu funkciju nisu bitni samo položaji i vrste aminokiselina, već i način na koji se one kreću i interagiraju kako jedne s drugima, tako i sa supstratima koji ulaze u aktivno mjesto. Kristalografske strukture i dalje igraju vrlo važnu ulogu budući da se upravo one koriste kao polazište za pripremu molekulskih simulacija. Molekulsko dinamičkim (MD) simulacijama možemo simulirati gibanje proteina na širokom rasponu vremenskih skala. Vrijeme trajanja molekulskih simulacija potrebno je prilagoditi ovisno o rezultatu koji se želi dobiti i naravno, dostupnoj računalnoj snazi.[81] Npr. za praćenje nastanka vodikovih (nekovalentnih) veza potrebno je simulirati na vremenskoj skali od 10 ps, za promjenu rotamera aminokiselina potrebno je simulirati oko 10 ns, dok je za veće konformacijske promjene potrebno simulirati oko i više od 1 μs (slika 14), dok je za simulaciju makroskopskih gibanja pojedinih domena u proteinu potrebno gledati čak i duža vremena .



Slika 14. Hijerarhija vremenske skale događaja u proteinima kao orijentir za određivanje trajanja MD simulacija. (preuzeto i prilagođeno iz reference [81])

2.3.1. Alosterija

Uz samo praćenje dinamike proteina, jedan od fenomena koji također o njoj ovisi je alosterija. Alosterička regulacija opisuje proces u kojem se alosterički modulator veže na neko alosteričko mjesto udaljeno od aktivnog mjesta pri čemu dolazi do konformacijske promjene koja zatim utječe na enzimsku aktivnost. Ako je riječ o povećanju aktivnosti ili konformacijskoj promjeni koja pozitivno utječe na ishod reakcije, radi se o alosteričkim aktivatorima. S druge strane, ako je pak riječ o smanjenju funkcije ili čak inhibiciji samog enzima radi se o alosteričkim inhibitorima.[82] Primjer negativne i pozitivne kooperativnosti alosteričkog modulatora prikazan je na slici 15. Ako se na alosteričko mjesto veže alosterički modulator može doći do konformacijske promjene koja onemogućava vezanje supstrata u aktivno mjesto ili pak može doći do pozitivne kooperacije kod koje se događa konformacijska promjena koja pomaže u vezanju supstrata u aktivno mjesto enzima. Alosterija i s njom povezane konformacijske promjene važni su dinamički procesi koji se mogu pratiti MD simulacijama čime se dobiva uvid u njihov vremenski tijek i mehanizam.



Slika 15. Primjer negativne i pozitivne kooperacije alosteričkog modulatora kod nekog proteina. Plavo je označen protein, zeleni kvadrat označava supstrat, crni trokut označava alosterički inhibitor, a narančasti trokut alosterički aktivator.

Dva načina alosteričke modulacije koji mogu opisati većinu alostičkih efekata su Monod-Wyman-Changeux (MWC) model i Koshland-Némethy-Filmer (KNF) model. MWC model temelji se na pretpostavci da protein ima dva (ili više) konformacijska stanja te da postoji ravnoteža između tih relaksiranih (R) i napetih (T) stanja. Vezanje alosteričkog modulatora na

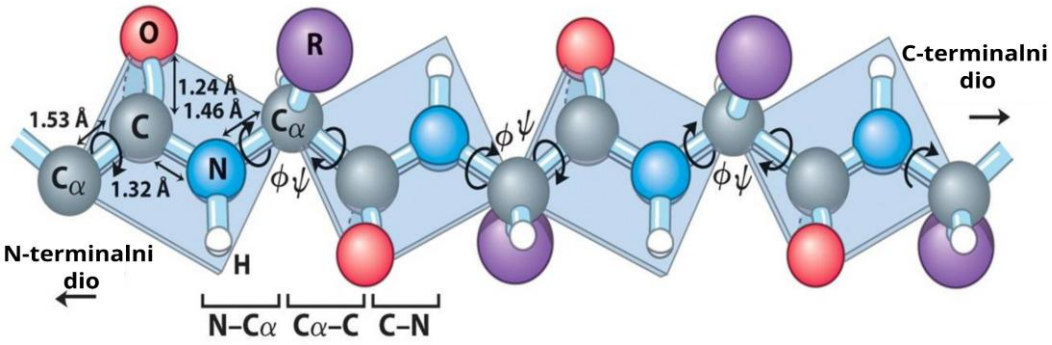
protein utječe tako da se pomiče ravnoteža prema jednom od navedenih stanja što dovodi do promjene aktivnosti enzima.[83] Jedan od glavnih primjera ovog modela je hemoglobin kod kojeg dolazi do vezanja molekule kisika u jedno aktivno mjesto što povećava afinitet vezanja u preostalim podjedinicama proteina.[84] Kod KNF modela vezanjem alosteričkog modulatora dolazi do lokalnih konformacijskih promjena koje se zatim prenose do drugih mjesta u proteinu. No, ne smije se miješati sa MWC modelom kod kojeg dolazi do simultane konformacijske promjene u ostalim podjedinicama enzima. Kod KNF modela se sekvencijalno događaju konformacijske promjene u podjedinicama proteina, a npr. vezanje supstrata u jednoj od njih samo neznatno utječe na vezna mjesta susjednih podjedinica te one postaju osjetljivije za vezanje supstrata.[85] Kinetičke studije HpPNP enzima pokazale su zanimljive rezultate. Budući da ovaj enzima ima dva supstrata (fosfat i purinska baza) u aktivnom mjestu, napravljena su mjerenja gdje je koncentracija jednog supstrata konstantna, a druga se varira. Kod reakcija kod kojih se varira koncentracija fosfata (koncentracija purinske baze je konstantna), uočava se kinetika modela alosteričkog mjesta s negativnom suradnjom između podjedinica dok se u suprotnom slučaju, kada se varira koncentracija purinske baze (koncentracija fosfata je konstantna), uočava klasična Michaelis-Menten kinetika.[86] Navedeno istraživanje i potreba za daljnjim analizama poticaj su za proučavanje kako PNP enzima tako i alosterije koja je potencijalno jedna od ključnih mehanizama u kinetici ovog enzima.[8]

Alosterija je izuzetno važna kod istraživanja novih terapijskih lijekova budući da se radi o novom inovativnom pristupu kod kojeg se vezanjem lijekova na alosteričko mjesto može direktno utjecati na povećanje ili smanjene aktivnosti enzima. Važno je napomenuti da alosterička mjesta nemaju veliku aminokiselinsku konzerviranost (kao što je to slučaj za aktivna mjesta) u svojim homolognim enzimima što daje prednost alosteričkim lijekovima u pogledu specifičnog vezanja na određeni enzim. Razvijanjem ovakvih lijekova omogućava se dvojno djelovanje na ciljni enzim gdje se u terapiji mogu koristiti i dosad već poznati lijekovi koji se smještaju u aktivno mjesto te također lijekovi koji se smještaju u alosteričko mjesto.[87]

Konformacijske promjene i komunikacijski putovi između aminokiselina i/ili podjedinica koje se događaju uslijed alosteričke regulacije su od ključne važnosti za njeno razumijevanje. Klasičnim strukturnim analizama ove dinamičke procese nije moguće okarakterizirati zbog čega je potrebno razviti nove napredne tehnike u raznim granama znanosti (poput molekulske dinamike) kako bi se što bolje istražili alosterički mehanizmi i sve što je njihov uzrok ili posljedica.

2.3.2. Ramachandranovi kutovi

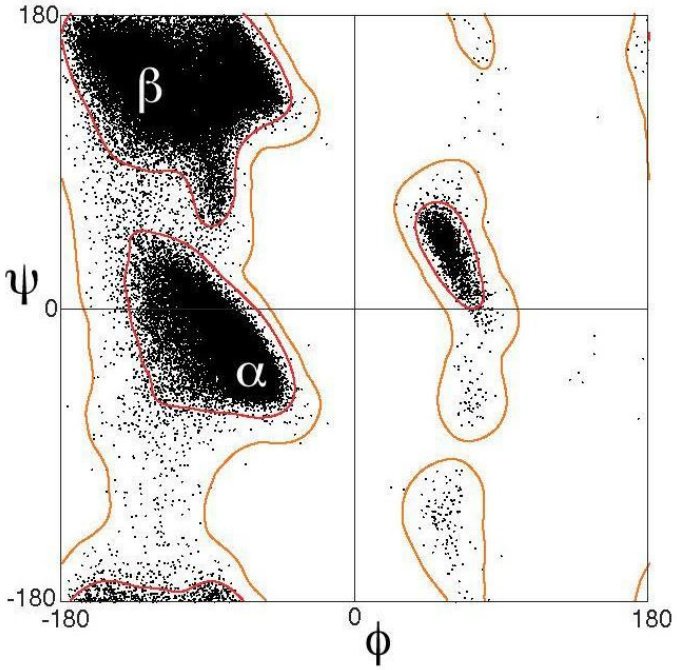
Vremenska dinamika proteina u MD simulacijama rezultira nizom položaja (i ponekad brzina) svih atoma u simulaciji u diskretnim vremenskim intervalima. Takav zapis nazivamo trajektorijom, i ona sadrži informaciju za rekonstrukciju cijelog vremenskog tijeka simulacije. Iz trajektorije po potrebi možemo uzimati samo određeni podskup koordinata radi praćenja dijelova proteina koji su od interesa (npr. koordinate samo C_α atoma pojedinih aminokiselina, ili centre mase pojedinih domena i sl.). Osobito povoljan izbor za praćenje dinamike su torzijski kutevi duž proteinske okosnice i bočnih ogranaka. Kutovi imaju tu prednost što su translacijski invarijantni i neovisni o ukupnom pomaku simuliranog proteina u vremenu, radi kojeg je potrebno izvršiti "centriranje" trajektorije. Pritom se ima sloboda izbora koje torzijske kutove uzeti u obzir. Povijesno gledano značajni su Ramachandranovi kutovi φ i ψ [88] koji dobro opisuju strukturu proteinske okosnice i peptidne veze i koji su dobro definirani za sve osim prve i zadnje aminokiseline u pojedinom proteinskom lancu. To su zapravo torzijski kutovi (definirani pomoću 4 atoma) aminokiselinske okosnice (slika 16). Kut φ definiran je s karbonilnim ugljikovim atomom prethodne aminokiseline (C_{n-1}), dušikovim atomom (N_n), ugljikom na koji je vezan bočni ogranak ($C_{\alpha,n}$) i karbonilnim ugljikovim atomom (C_n), dok je kut ψ definiran dušikovim atomom (N_n), ugljikom na koji je vezan bočni ogranak ($C_{\alpha,n}$), karbonilnim ugljikovim atomom (C_n) i dušikovim atomom iz iduće aminokiseline (N_{n+1}).



Slika 16. Prikaz Ramachandranovih kutova φ i ψ na shematskom prikazu polipeptidne okosnice. R je oznaka aminokielinskog ostatka. (preuzeto i prilagođeno iz reference [77])

Ramachandranov dijagram tradicionalno se koristi za prikaz ukupne raspodjele konformacija okosnice proteina u jednoj strukturi, gdje se jedna aminokiselina prikazuje kao jedna točka u φ–ψ dijagramu (slika 17). Ovakav prikaz omogućuje da se odmah uoče konformacije koje

padaju izvan dopuštenih područja, što čini ove dijagrame neizostavnim alatom kod validacije strukture proteina. Zbog svoje prilično složene prirode i velikog broja točaka, Ramachandranov dijagram rijetko se koristio kao alat za praćenje konformacija okosnice proteina kroz vrijeme, budući da bi to iziskivalo vizualizaciju velikog broja takvih dijagrama na jednom grafu. No s napretkom vizualizacijskih metoda, u posljednje vrijeme javlja se interes za vizualizaciju dinamičkih podataka korištenjem dijagrama baziranih na Ramachandranovim kutovima.[89–91]

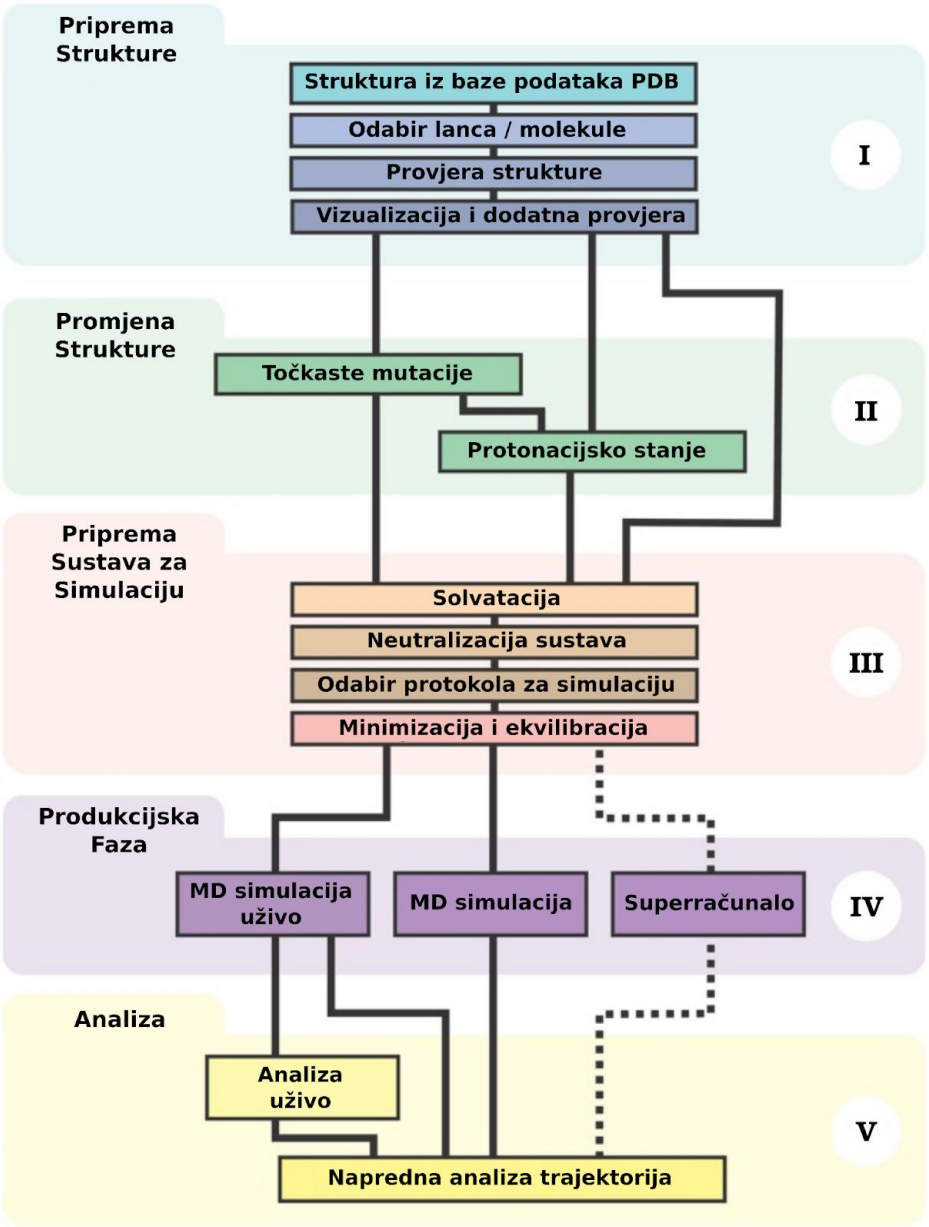


Slika 17. Statistička distribucija ovisnosti φ i ψ kutova kod više od 500 struktura visoke rezolucije iz PDB baze podataka.[56] Područje unutar crvene linije označava vrijednosti φ i ψ kutova koje je dominantno kod aminokiselina. Područje između crvene i narančaste linije označava vrijednosti φ i ψ kutova koja su dopuštene kod aminokiselina. Sve izvan narančaste linije označava područja φ i ψ kutova koji nisu dopušteni. Vrijednosti dopuštenih kutova su u najvećoj mjeri određene steričkim smetnjama bočnih ogranaka aminokiselina te tako vidimo koja su preferirana područja za α-zavojnicu i β-ploču. Treće najdominantnije područje koje nije označeno na slici, a nalazi se u desnom gornjem kvadrantu odgovara aminokiselinama koje se najčešće nalaze u lijevo zakrenutoj zavojnici, u okretima ili zavojima. (preuzeto i prilagođeno iz reference [92])

2.3.3. Metode MD simulacija

Simulacije molekulske dinamike opisuju gibanje molekulskog sustava koristeći zakone klasične mehanike, a za opis veza između atoma najčešće se upotrebljava aproksimacija harmonijskog oscilatora. U iterativnom postupku, i za svaki korak simulacije se za sustav intereagirajućih čestica numerički rješavaju Newtonove jednadžbe gibanja čime se dobivaju trajektorije atoma, dok se potencijali najčešće računaju molekulsko mehaničkim poljima sila. Većina polja sila su empirijska i definirana su kao suma svih veznih (istezanje veza, promjena veznih ili torzijskih kutova) i neveznih doprinosa (elektrostatske ili van der Waalsove interakcije). Jedan od najzahtjevnijih koraka prilikom izračuna trajektorija je upravo izračunavanje neveznih doprinosa ukupnoj potencijalnoj energiji, budući da se moraju izračunati milijuni parova interakcija u svakom vremenskom koraku.[93] Dakle, MD simulacijama mogu se oponašati i proučavati gore navedene promjene poput istezanja kovalentnih veza, promjena veznih kutova ili pak neveznih interakcija. Polje sila sadrži formule za izračun potencijalne energije sustava i sila iz koordinata svih čestica sustava te također većinu parametara koji se koriste u tim formulama. Aproksimacije koje se koriste u ovim izračunima veoma su dobre, ali imaju svoja ograničenja te ne uključuju izračun kvantnih doprinosa tako da u MD simulacijama ne može doći do formiranja ili pucanja kemijskih veza.[94] Jedno od najčešće korištenih polja sila je AMBER polje sila te je od 2018. godine ff14SB[95] postalo gotovo standardno za proteinske simulacije. Proces simuliranja proteinskih gibanja započinje odabirom početne strukture koja je najčešće rezultat neke od metoda za dobivanje precizne trodimenzijske strukture na atomskoj razini poput rendgenske kristalografije, elektronske mikroskopije ili otopinske nuklearne magnetske rezonancije (NMR). Sve makromolekulske strukture pohranjene su u bazu podataka Protein Data Bank (PDB)[56] koja u trenutku pisanja ove disertacije broji ukupno 223,532 eksperimentalno riješenih struktura. Zbog veoma dinamičnih dijelova proteina, kao što je opisano gore, moguće da su ti dijelovi izostavljeni prilikom deponiranja struktura u PDB. Također, većina struktura nema u svojim koordinatama vodikove atome zbog nemogućnosti njihove detekcije prilikom eksperimenta, osim za mjerenja jako visoke rezolucije. Nakon odabira strukture potrebno je ispraviti detalje poput dopunjavanja nedostajećih dijelova strukture, dodavanja vodika, provjere i dodavanja naboja aminokiselina te se potom struktura priprema za daljnje korake. Struktura proteina se zatim smješta u simulacijsku kutiju u vodenu otopinu (kako bi se simulirali fiziološki uvjeti) te se sustav neutralizira dodavanjem natrijevih ili kloridnih iona. Definiraju se periodični rubni

uvjeti što znači da je simulacijska kutija sa svih strana okružena svojim replikama. Potom slijedi minimizacija sustava, odnosno energetska minimizacija što je zapravo proces pronalaska optimalnog razmještaja atoma u prostoru gdje je teoretski ukupna međuatomska sila na svakom atomu približno jednaka nuli, u realnom slučaju gdje je ta sila približno jednaka zadanoj minomalnoj sili. Dvije najčešće metode koje se koriste su robusna metoda najstrmijeg spusta (engl. *method of steepest descent*) i finija metoda konjugiranog gradijenta (engl. *conjugate gradient method*). Obje ove metode koriste se činjenicom da se na temelju smjera i iznosa gradijenta funkcije polja sila, može procijeniti smjer i udaljenost minimuma na plohi potencijalne energije od same polazišne točke. Nakon minimizacije slijedi uravnoteženje sustava, što znači povećanje temperature i tlaka sustava na željene, obično također fiziološke uvjete. Nakon ekvilibracije, slijedi produkcijska, odnosno MD, faza koja simulira gibanje atoma u sustavu čime se pokušava oponašati sustav na molekulskoj razini kroz zadani vremenski period. Zbog temperature koja je uvedena u sustav atomi imaju određenu kinetičku energiju te se oni konstantno gibaju, a na početku simulacije dodijeljene su im početne brzine prema Maxwell-Boltzmanovoj raspodjeli za zadanu temperaturu. Rezultat MD simulacija su vremenski zapisi gibanja proteina (trajektorije) koje predstavljaju niz trodimenzijskih struktura u prostoru u vremenu (slika 18). Dobivenu trajektoriju potrebno je analizirati kako bi se što bolje razumio promatrani protein.[94]



Slika 18. Hodogram za pripremu, izvođenje, analizu i vizualizaciju MD simulacija podijeljeno u okvirnih pet dijelova koji sumiraju cjelokupni proces. (preuzeto i prilagođeno iz reference [96])

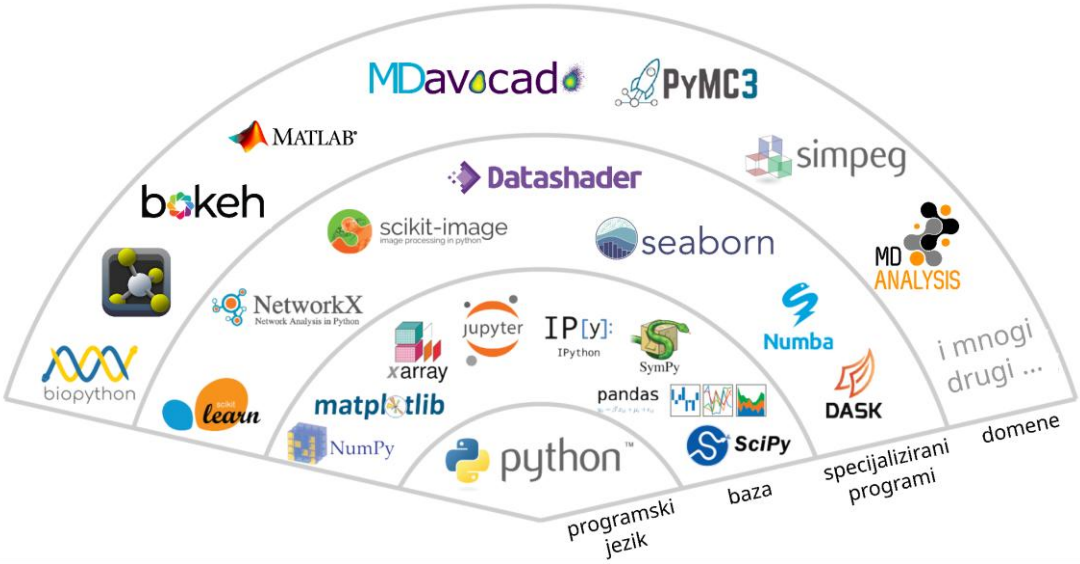
2.3.4. Analiza podataka

Kao što se može i pretpostaviti, MD simulacije generiraju izuzetno velike količine podataka, posebice ako se radi o dugim simulacijama velikih sustava koji sadrže mnogo atoma. Stoga izazov predstavlja kako iz mnoštva podataka izvući smislene informacije, tj. kako u moru nasumičnih gibanja proteina pronaći ona koja su doista uzročno-posljedično povezana. S druge strane, velika količina podataka omogućava da se tom problemu pristupi upotrebom raznih

metoda upotrebe znanosti o podacima i strojnog učenja (engl. *big data* and *machine learning,* ML). Strojno učenje je znanstvena disciplina u okviru umjetne inteligencije koja proučava izradu i upotrebu algoritama za obradu podataka te pronalaska novog znanja u tim podacima. Princip ML je da se odabrani model prilagođava i traži povezane uzorke u podacima koji se obrađuju. Jednom kada je model naučen, naučeno znanje primjenjuje se na podatke koji nisu sudjelovali u početnom učenju, odnosno ti algoritmi naposljetku mogu vršiti zadatke bez direktnih uputa.[97] U posljednje vrijeme strojno učenje se sve više koristi i pronalazi se primjena algoritama strojnog učenja u mnogim poljima kao što su prepoznavanje govora, procesiranje teksta, medicine pa i kemije. Jedan od važnijih primjera iz polja kemije je svakako AlphaFold,[98] program koji predviđa trodimenzijske strukture proteina i nukleinskih kiselina te njihovih potencijalnih kompleksa te također interakcija do kojih dolazi prilikom njegovog stvaranja. Ovaj program izvrstan je primjer algoritma strojnog učenja, budući da je model za predviđanje struktura treniran upravo na podacima iz baze podataka PDB, te se program zatim primijenio na još neobjavljene strukture.[99]

Jedan poseban dio strojnog učenja je podatkovno rudarenje (engl. *data mining*) koje je usko vezano uz različite grane znanosti: strojno učenje, statistika i baze podataka, a može se opisati kao proces kojim se iz velikih skupova podataka ekstrahiraju i otkrivaju uzorci.[100] Svrha ovakve obrade podataka je zapravo smanjenje njihove količine i dimenzionalnosti čime se dobivaju novi kompaktniji skupovi podataka prikladni za daljnju obradu. Novi skupovi podataka mogu zatim pokazati neke nepoznate i zanimljive uzorke u originalnom setu podataka, podaci se mogu grupirati što rezultira klasterima podataka, mogu se detektirati anomalije u podacima što omogućava izolaciju nekog seta podataka i njegovu daljnju obradu. Otkriveni uzorci su rezultat rudarenja podataka i predstavljaju sažetak ulaznog seta podataka te se mogu koristiti u daljnjim analizama strojnog učenja poput prediktivne analize.[101] U novije vrijeme sveprisutna je upotreba metoda strojnog učenja u analizi MD simulacija gdje se algoritmi upotrebljavaju upravo zbog prevelike količine podataka i nemogućnosti njihove obrade na klasičan način.[102–104] Svemu tome potrebno je pridonijeti i skupom alata koji omogućavaju brzu obradu i razvoj novih algoritama. Jedan od krovnih alata u tom smislu je Python kao programski jezik vodeći u implementaciji programa i algoritama za procesiranje i analiziranje podataka,[105] a to se odlično nadovezuje na široku bazu programa opće namjene koji su već razvijeni u tom programskom jeziku. Kod znanstvene primjene Python se uz pomoć tih dodatnih programa transformira u visoko kvalitetan programski jezik koji postaje izuzetno brz i koristan alat za

obradu podataka, a ujedno je i dovoljno fleksibilan da se ubrza uz dodatne ekstenzije.[106] Na slici 19 može se vidjeti dio velike palete programa implementiranih u Pythonu koji se koriste prilikom obrade znanstvenih podataka, manipulacije velikih setova podataka, vizualizacije ili kod specijalizirane obrade određenih setova podataka.

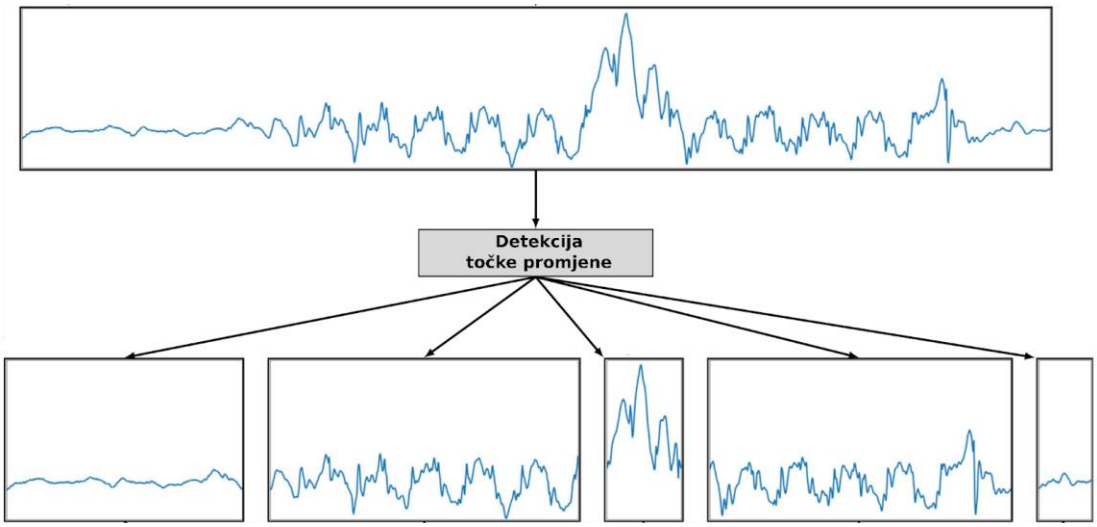


Slika 19. Dio širokog ekosistema programskog jezika Python koji predstavlja jednostavnost i širinu ovog programskog jezika kroz sve alate koji su dostupni za implementaciju kod analize podataka. (preuzeto i prilagođeno iz reference [107])

Primjer jednog takvog programa koji su napisani u programskom jeziku Python je MDAnalysis koji je specijalizirani program za manipulaciju MD trajektorija.[108,109] MDAnalysis omogućava obradu trajektorija i struktura proteina, nukleinskih kiselina, lipida i drugih biomolekula. Također se mogu izračunati različite fizikalne i strukturne karakteristike, kao što su međuatomske udaljenosti, kutovi ili torzijski kutovi. Također on je građen kao skup alata od kojih se mogu napraviti prilagođeni alati potrebni za specifične analize. Ova biblioteka se također integrira s drugim Python alatima, što je čini moćnim alatom za istraživanje i analizu MD simulacija. Neke od specifičnih analiza koje se mogu napraviti uz pomoć programa MDAnalysis su izračun standardne devijacije promjene položaja (RMSD), analize redukcije dimenzionalnosti, npr. analiza glavnih komponenata (engl. *principal component analysis*) ili grupiranja (engl. *clustering*).

Kako trajektorije možemo promatrati kao vremenske sljedove podataka (engl. *time series*) od posebnog interesa su programi kojima možemo analizirati takve vremenske serije.

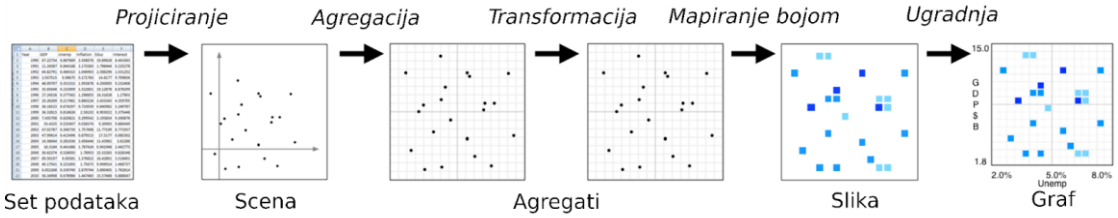
Posebice nas zanima kako u ogromnom mnoštvu takovih vremenskih slijedova automatski odrediti određene događaje, bilo da se radi o skokovima u vrijednosti ili promjenama npr. standardne devijacije i sl. Ovdje također postoji cijeli niz programskih alata za analizu promjena u vremenskim serijama, ali posebno se izdvaja program Ruptures.[110] Program za analizu promjena u vremenskim serijama Ruptures predstavlja program koji se lako može primijeniti na analizi dinamičkih podataka poput trajektorija. Trajektorije naime sadrže razne podatke u obliku vremenskih ovisnosti: vrijednosti koordinata u vremenu, raznih RMSD vrijednosti ili promjene kutova kroz vrijeme. Ono u čemu pomaže Ruptures je automatsko procesiranje i detekcija točaka promjene (engl. *changepoints*, slika 20). To omogućava paralelno procesiranje velikog broja pojedinih karakteristika trajektorije, kao što je npr. automatsko pridjeljivanje promjena u Ramachandranovim kutevima. Ovim programom moguće je detektirati promjene u frekvenciji samog signala kao i promjenu u amplitudi signala.



Slika 20. Veliki broj znanstvenih varijabli dolazi u obliku vremenske serije podataka. Zajedničko im je da predstavljaju promjenu neke veličine u vremenu. Na njih možemo primijeniti razne algoritme za detekciju promjena, npr. promjene frekvencije ili amplitude signala. (preuzeto i prilagođeno iz reference [110])

Uz problem automatskog procesiranja vremenskih serija (koje mogu u slučaju trajektorija sadržavati i milijune točaka) javlja se problem prikaza tako velikih količina podataka na jednoj slici. Ukoliko npr. želimo prikazati međusobnu ovisnost jednog promatranog kuta o drugom to ne možemo učiniti prikazom točkica bilo kako male veličine, jer će zasigurno doći do njihova prekrivanja. Mora se pribjeći drugim načinima prikaza. S tom svrhom može se koristiti još

jedan Python program pod imenom Datashader.[111] Njegova primarna namjena je vizualizacija velikih skupova podataka, odnosno golemog broja točaka. Dizajniran je za efikasno prikazivanje podataka s milijunima ili čak milijardama točaka, gdje standardne metode vizualizacije često postaju neupotrebljive zbog prevelike količine informacija na samo jednom grafu. Zapravo tehnika "agregacije" podataka koja se tu koristi čini ga neovisnim o samom broju podataka. Umjesto da prikazuje sve podatke, Datashader agregira i prilagođava podatke tako da se zadrže ključne informacije te da se pritom vide i analiziraju obrasci i trendovi u originalnom setu podataka (slika 21). Veliki set podataka se najprije projicira, zatim se ti podaci agregiraju i transformiraju, nakon čega dolazi do mapiranja podataka pomoću boje i njihova ugradnja u graf odnosno vizualizacija.



Slika 21. Prikaz hodograma vizualizacije velikih setova podataka pomoću programa Datashader. (preuzeto i prilagođeno iz reference [111])

§ 3. EKSPERIMENTALNI DIO

Prije početka eksperimentalnih istraživanja i analize podataka bilo je potrebno pronaći odgovarajuće početne strukture u bazi makromolekulskih struktura PDB. Cilj pretraživanja bio je pronalazak modelnih struktura enzima izoliranih iz bakterije *Helicobacter pylori* koje će se koristiti kao početne točke za MD simulacije. Nakon odabira nekoliko odgovarajućih struktura, generirane su MD trajektorije koje prikazuju gibanje proteina u prostoru (simulacijskoj kutiji) kroz vrijeme (trajanje simulacije). Kako bi se detaljnije proučile konformacijske promjene enzima potrebno je simulirati taj enzima u što dužem vremenskom intervalu. Dobivene trajektorije analizirane su pomoću raznih alata s ciljem otkrivanja uzročno-posljedičnih veza koje stoje iza otvaranja i zatvaranja aktivnih mjesta monomernih podjedinica PNP-a. Osim ovih globalnih konformacijskih promjena, proučavane su i lokalne promjene, uključujući analizu Ramachandranovih kutova svih aminokiselina. Za potrebe analize razvijene su vlastite programske rutine i metodologije u programskom jeziku Python. Jedan od korištenih alata, nazvan MDavocado, pokazao se izuzetno korisnim u interpretaciji i analizi MD trajektorija.[9]

3.1. Kristalne strukture

Kao polazišna točka svih simulacija korištene su kristalografski određene strukture PNP-a iz bakterije *H. pylori* (HpPNP) te su one izdvojene u tablici 1. Pretraga PDB baze podataka rezultirala je s 13 HpPNP struktura kod kojih se uočava različita distribucija otvorenih, odnosno zatvorenih aktivnih mjesta. Također se kod nekih struktura u njihovim aktivnim mjestima nalaze različiti ligandi koji imitiraju supstrate tog enzima. Kao što je navedeno u literaturnom pregledu, ove strukture koriste se kao polazišne točke MD simulacija kojima se pokušava što bolje uzorkovati konformacijski prostor. Izabrano je 8 struktura za nastavak istraživanja čije se strukture međusobno razlikuju na osnovi konformacijskog stanja aktivnih mjesta te prisustvu liganada u njima. Tablica 1 je podijeljena na dvije manje tablice gdje gornja predstavlja pretežito otvorene strukture a donja pretežito zatvorene strukture. Dodatno su naglašeni monomerni lanci (A – F) koji se nalaze u otvorenoj (O) konformaciji zelenom bojom te lanci koji se nalaze u zatvorenoj (Z) konformaciji crvenom bojom. Svaka tablica prikazuje četiri

strukture gdje se razlikuju *apo* (strukture bez liganada) i *holo* (strukture s ligandima) strukture. Stupac O/Z označava gore opisanu raspodjelu otvorenih i zatvorenih aktivnih mjesta, dok stupci PO4 (fosfat), HPA (hipoksantin) i FMC (formicin A) prikazuju prisutnost navedenih liganada u odgovarajućem monomernom lancu, odnosno govore o tome nalazi li se navedena u struktura u svojoj *apo* ili *holo* formi. Prazno polje indicira da u tom monomeru nema navedenog liganda, dok + indicira prisustvo tog liganda u aktivnom mjestu enzima.

Tablica 1. Raspodjela otvorenih (gornja tablica) i zatvorenih (donja tablica) struktura te supstrata u njihovim aktivnim mjestima. S lijeve strane nalaze se nazivi monomernih lanaca (A, B, C, D, E i F) te je kod struktura zelenom bojom označeno otvoreno (O) aktivno mjesto i crvenom bojom zatvoreno (Z) aktivno mjesto. Strukture bez liganada su označene oznakom *apo*, fosfat je označen kao *PO4*, a dušikova baza kao *FMC* (formicin A) i *HPA* (hipoksantin). Oznaka + potvrđuje da se navedeni ligand nalazi u aktivnom mjestu pripadajućeg monomernog lanca.

	Otvorene strukture											
	5mx4_apo			6f52_apo			6g7x_PO4			5mx6_PO4_HPA		
	O/Z	PO4	HPA	O/Z	PO4	FMC	O/Z	PO4	FMC	O/Z	PO4	HPA
A	O			O			O	+		O	+	+
B	O			O			O	+		O	+	+
C	Z			O			O	+		Z	+	+
D	O			O			O	+		O	+	+
E	O			O			O	+		O	+	+
F	O			O			O	+		O	+	+

	Zatvorene strukture											
	5lu0_apo			6f4x_apo			5lu0_PO4			6f4x_PO4_FMC		
	O/Z	PO4	FMC	O/Z	PO4	FMC	O/Z	PO4	FMC	O/Z	PO4	FMC
A	Z			Z			Z	+		Z	+	+
B	Z			O			Z	+		O		+
C	Z			Z			Z	+		Z	+	+
D	O			Z			O			Z		+
E	Z			Z			Z	+		Z	+	+
F	O			O			O			O		+

Struktura HpPNP enzima sastoji se od šest podjednakih lanaca duljine 233 aminokiseline, označenih slovima A-F (slika 22). Iz baze podataka odabrano je šest struktura koje imaju svoje međusobne razlike koje su opisane u sljedećem popisu:

1. Struktura 5lu0 enzima HpPNP

   Struktura ima dva otvorena aktivna mjesta (lanci D i F) te četiri zatvorena aktivna mjesta (lanci A, B, C i E). U otvorenim aktivnim mjestima nema liganada dok se u zatvorenim aktivnim mjestima nalaze fosfati, dakle svako zatvoreno aktivno mjesto sadrži jedan fosfatni ion. Kao takva, struktura 5lu0, koristila se kao početna struktura za MD simulaciju prema oznaci *5lu0_po4*. Također, kako bi se što bolje ispitao konformacijski prostor sustava i utjecaj samih liganada, iz ove strukture obrisani su ioni fosfata te je takva struktura poslužila kao još jedna početna struktura za MD simulaciju prema oznaci *5lu0_apo*. Ova početna struktura svrstana je u zatvorenu strukturu bez liganada.

2. Struktura 5mx4 enzima HpPNP

   Struktura ima pet otvorenih aktivnih mjesta (lanci A, B, D, E i F) te jedno zatvoreno aktivno mjesto (lanac C). U otvorenim aktivnim mjestima ove strukture nisu pronađeni nikakvi ligandi, dok je u aktivnom mjestu lanca C pronađen fosfatni ion i hipoksantin koji su prije početka simulacije obrisani. Kao takva, struktura 5mx4, koristila se kao početna struktura za MD simulaciju prema oznaci *5mx4_apo*.

3. Struktura 5mx6 enzima HpPNP

   Struktura ima pet otvorenih aktivnih mjesta (lanci A, B, D, E i F) te jedno zatvoreno aktivno mjesto (lanac C). U svim aktivnim mjestima također se nalaze fosfat i dušikova baza hipoksantin. Kao takva, struktura 5mx6, koristila se kao početna struktura za MD simulaciju prema oznaci *5mx6_po4_hpa*.

4. Struktura 6f4x enzima HpPNP

   Struktura ima dva otvorena aktivna mjesta (lanci B i F) te četiri zatvorena aktivna mjesta (lanci A, C, D i E). U aktivnim mjestima lanaca A, C i E nalaze se fosfati, dakle svaki od njih sadrži jedan fosfatni ion te je također u aktivnim mjestima svih lanaca nađena dušikova baza formicin A. Kao takva, struktura 6f4x, koristila se kao početna struktura za MD simulaciju prema oznaci *6f4x_po4_fmc*. Također, kako bi se što bolje ispitao konformacijski prostor sustava i utjecaj samih liganada, iz ove strukture obrisani su ioni fosfata i molekule formicina A te je takva struktura poslužila kao još jedna početna

struktura za MD simulaciju prema oznaci *6f4x_apo*. Ova početna struktura svrstana je u zatvorenu strukturu bez liganada.

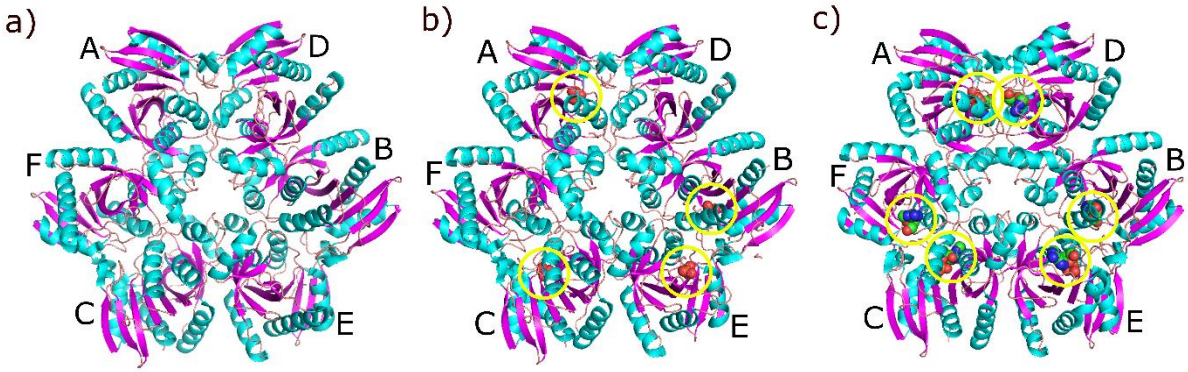
5. Struktura 6f52 enzima HpPNP

   Struktura ima šest otvorenih aktivnih mjesta (lanci A, B, C, D, E i F). U svim aktivnim mjestima ove strukture nisu pronađeni nikakvi ligandi. Kao takva, struktura 6f52, koristila se kao početna struktura za MD simulaciju prema oznaci *6f52_apo*.

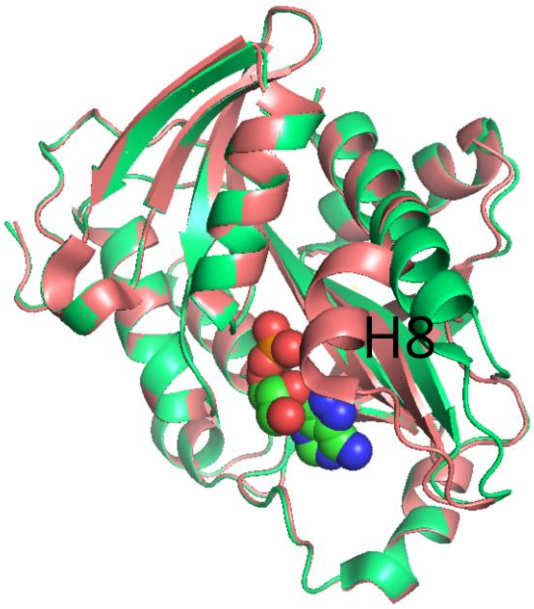
6. Struktura 6g7x enzima HpPNP

   Struktura ima šest otvorenih aktivnih mjesta (lanci A, B, C, D, E i F). U svim aktivnim mjestima također se nalaze fosfati, dakle svaki od njih sadrži jedan fosfatni ion. Kao takva, struktura 6g7x, koristila se kao početna struktura za MD simulaciju prema oznaci *6g7x_po4*.

Primjer otvorene strukture HpPNP enzima bez liganada (a), zatvorene strukture s fosfatom (b) i zatvorene strukture s oba liganda (c) nalazi se na slici 22. Također su izdvojena dva aktivna mjesta koja prikazuju H8 α-zavojnicu u otvorenoj konformaciji (zelena boja) i u zatvorenoj konformaciji (crvena boja) (slika 23). Sve slike proteina napravljene su pomoću programa PyMOL.[112]



Slika 22. Primjer (a) otvorene strukture HpPNP enzima bez liganada, (b) zatvorene strukture s fosfatom i (c) zatvorene strukture s oba liganda (fosfat i formicin A). Ligandi su na slikama zaokruženi žutim kružnicama. Konkretno, slike predstavljaju strukture *6f52_apo, 5lu0_po4* i *6f4x_po4_fmc*. (preuzeto i prilagođeno iz reference [8])



Slika 23. Dva aktivna mjesta HpPNP enzima koja prikazuju H8 α-zavojnicu u otvorenoj konformaciji (zelena boja) i u zatvorenoj konformaciji (crvena boja). Ligandi (fosfat i formicin A) iz zatvorene konformacije označeni su kuglicama. (preuzeto i prilagođeno iz reference [8])

3.2. Simulacije molekulske dinamike

Navedenih osam struktura HpPNP enzima poslužile su kao početne strukture za MD simulacije. Osim same strukture HpPNP enzima i eventualno liganada (prema opisanoj pripremi struktura), sve ostalo iz početnih datoteka je obrisano, npr. kristalne vode, molekule ili ioni zaostali iz kristalizacijske otopine. Strukture je potom bilo potrebno provjeriti zbog njihove kompletnosti, odnosno ima li prekida u njihovoj trodimenzijskoj strukturi te popraviti, pomoću programa Coot,[113] ako se takvi prekidi pronađu. Također, ako su u kristalnim strukturama pronađene neuređene regije, uzimala se u obzir samo jedna od predloženih konformacija. Za sve daljnje korake i samu MD simulaciju bile su potrebne datoteke topologije i koordinata za čije generiranje se koristio program Amber20.[114] Prilikom stvaranja ovih datoteka bilo je potrebno dodati atome vodika cijeloj strukturi HpPNP enzima budući da oni najčešće nedostaju kod eksperimentalno riješenih struktura. Parametrizacija HpPNP enzima napravljena je pomoću polja sila AMBER ff14SB[95] budući da su u njemu sadržani paramteri za aminokiseline, dok su ligandi parametrizirani pomoću polja sila GAFF[115] koristeći program Antechamber[116] dostupan u programskom paketu Amber20. Sustav se zatim dodatno pripremio prije izvođenja simulacija

minimizacije i ekvilibracije. Parametriziran sustav smješta se u simulacijsku kutiju oblika krnjeg oktaedra te se ona ispunjava molekulama vode tipa TIP3P.[117] Ako sustav nije neutralan, potrebno je dodati natrijeve (pozitivne) ili kloridne (negativne) ione kako bi se ukupan naboj uravnotežio.[118] Na simulacijsku kutiju također su primjenjeni periodični bočni uvjeti, koji osiguravaju ukupnost sadržaja kutije kroz cijelo simulacijsko vrijeme. Ovako pripremljen sustav spreman je za simulaciju minimizacije. Minimizacija sustava izvršila se u tri koraka koristeći dvije različite metode za izračun: metoda najstrmijeg spusta i metoda konjugiranog gradijenta. Tijekom inicijalnog koraka optimizacije energije primijenjeno je ograničenje kretanja atoma okosnice proteina primjenom konstatne sile od 2000 kcal mol^{-1} $Å^{-2}$ tijekom cijelog trajanja simulacije od ukupno 50.000 koraka. U drugom koraku također se koristila metoda najstrmijeg spusta, no nisu se koristila nikakva ograničenja prilikom simulacije te je ukupno trajanje bilo 25.000 koraka. U trećem koraku koristila se metoda konjugiranog gradijenata, također bez ograničenja i u ukupnom trajanju od 25.000 koraka. Ekvilibracija odnosno temperaturno uravnoteženje vršilo se postepenim povećanjem temperature sustava od 0 K do 300 K u šest koraka. Tijekom tih koraka ograničenje na kretanje atoma okosnice proteina je postepeno smanjeno od početne korištene konstantne sile od 2000 kcal mol^{-1} $Å^{-2}$ do nekorištenja nikakvog ograničenja kretanja. Za zagrijavanje sustava korišten je NVT ansambl (kanonski ansambl, ansambl stanja koja imaju jednake vrijednosti broja čestica *N*, volumena *V* i temperature *T*) te je svaki od ovih koraka trajao 40 ps što je ukupno vrijeme temperaturne ekvilibracije 240 ps. Tlačno uravnoteženje, do tlaka jednakog atmosferskom tlaku, je finalni korak prije produkcijske faze. Kod druge ekvilibracije korišten je NPT ansambl (izotermno-izobarni ansambl, statističko mehanički ansambl mikrostanja s konstantnim brojem čestica *N*, tlakom *P* i temperaturom *T*) te je njeno ukupno trajanje 500 ps. Simulacija molekulske dinamike provodila se ukupno 1 μs te je temperatura održavana konstantnom pomoću Langevinova termostata[119] s frekvencijom kolizije of 1 ps^{-1} dok je tlak održavan konstantnim pomoću Berendsenova barostata[120] s izotropnim skaliranjem položaja. Ova produkcijska faza simulirana u NPT ansamblu računata je uz pomoć koda ubrzanog na GPU jedinici[121,122] dostupnog u programskom paketu Amber20. Sve navedene simulacije izvedene su ili na lokalnom superračunalu ili na računalnom klasteru Srce Sveučilišta u Zagrebu.

3.3. Analiza trajektorija

Analiza trajektorija započeta je provjerom sustava tijekom cijelog vremena simulacije. Ako sustav nije bio stabilan i uravnotežen, simulacije je bilo potrebno ponoviti. Najčešće korištena metoda za provjeru stabilnosti je izračun RMSD vrijednosti cijelog proteina. Ova analiza napravljena je na način da se kao referentna struktura koristi struktura u prvom koraku simulacije te se računa srednje kvadratno odstupanje struktura svih drugih koraka simulacije u odnosu na nju. Na trajektorije koje su stabilne tijekom cijelog trajanja simulacije ukazuje postupna stabilizacija RMSD vrijednosti na njenom početku i njezina konstantnost do kraja simulacije. U ovakvom slučaju može se nastaviti analiza pomoću složenijih alata. Za analizu trajektorija koristili su se alati u programskom jeziku Python. Kao jedan od glavnih programa za analizu MD trajektorija koristio se program MDAnalysis. U njemu su napravljene osnovne analize trajektorija poput RMSD analiza (pojedinačnih dijelova proteina), usporedbe i sravnjivanja trajektorija, izračuna Ramachandranovih kutova aminokiselina itd. Dobiveni podaci poslužili su za daljnje složenije analize kao što je vremenska korelacija raznih promjena kod aminokiselina. Za te daljnje analize koristio su se široki spektar Python modula posebno namijenjenih analizi podataka: NumPy,[123] SciPy,[124] pandas,[125] scikit-learn,[126] Biopython,[127] itd. Kako bi se utvrdilo dolazi li do kakve globalne konformacijske promjene, napravljena je analiza glavnih komponenata (PCA analiza) iz koje su dobivene dominantne komponente kretanja proteina. Analiza glavnih komponenata izračunata je na cjelokupnom setu podataka, odnosno na svim trajektorijama istovremeno, budući da se radi o uzorkovanju istog sustava. Spajanjem trajektorija dobivaju se još veći setovi podataka što može uvesti popriličan šum kod njihovog analiziranja, pogotovo ako konformacijska promjena nije previše izražena. Navedenom analizom promatra se protein kao cjelina, što može biti jedan od uzroka ne uočavanja ili težeg uočavanja (manjih) globalnih konformacijskih promjena.

Analiza trajektorija nastavljena je u smjeru proučavanja manjih dijelova proteina, kao što su analiza samo H8 α-zavojnice ili analiza aminokiselina i liganada u aktivnom mjestu. Prilikom analize aktivnog mjesta koristila se metoda određivanja sekundarne strukture aminokiselina H8 α-zavojnice, DSSP analiza (engl. *Define Secondary Structure of Proteins*). DSSP analiza temelji se na izračunu vodikovih veza u okosnici proteina čime su zapravo i definirane sekundarne strukture.[128] Ovom metodom pridjeljuju se elementi sekundarne strukture svim aminokiselinama na temelju pozicije atoma iz konformacije njihove okosnice. Analizom podataka dobivenih ovom metodom može se također utvrditi dolazi li do konformacijskih

promjena koje utječu na sekundarne strukture u proteinu. Kao dio analize aktivnih mjesta, proučavali su se kretanje fosfata (ulazak ili izlazak u aktivno mjesto) kod nekih simularnih struktura. Izračunati su njihovi kontakti i udaljenosti koji su uspoređeni s unaprijed poznatim kristalografskim podacima. Za izračun sekundarnih struktura, udaljenosti liganada od određenih aminokiselina i kontakti u aktivnom mjestu korištene su metode dostupne u Amberovom alatu cpptraj.[129,130]

Uz navedene globalne promjene također su se proučavale i lokalne promjene. Jedna mogućnost promatranja lokalnih promjena je praćenje promjena kutova u glavnom lancu proteina. Kutovi okosnice odražavaju velike konformacijske promjene (otvaranje i zatvaranje aktivnog mjesta odnosno prijelom α-zavojnice) u strukturi proteina. Uz ove veće konformacijske promjene, ova analiza može primijetiti i puno suptilnije promjene. Pomoću MDAnalysis programa pratili su se torzijski kutovi svih aminokiselina u HpPNP-u kako bi se dobili dijagrami ovisnosti φ i ψ kuta (svake pojedinačne aminokiseline) u vremenskom intervalu trajanja simulacije, kao svojevrsni vremenski ovisni Ramachandranovi dijagrami. Tako dobiveni dijagrami sadržavaju milijune točaka pa nije moguće koristiti uobičajene programe za vizualizaciju. U tu svrhu koristio se program Datashader kao dio šireg Python vizualizacijskog okružja HoloViz, koji je specijaliziran za analizu i vizualizaciju velikih količina podataka. Uz sam izračun i vizualizaciju vremenski ovisnih Ramachandranovih dijagrama, napravljen je program, MDavocado koji se može koristiti kao automatski alat za izračun ovakvih dijagrama te je univerzalno primjenjiv na bilo koju trajektoriju. Program je napisan u programskom jeziku Python te upute se za uporabu nalaze se na web stranici platforme GitHub: https://github.com/zoranstefanic/MDavocado (datum pristupa 12. rujna 2024.). Također su izračunate "sve-sa-svima" korelacije u promjenama navedenih kutova kod svih aminokiselina koje nose informaciju o istovremenosti pojedinih događaja u HpPNP enzimu. Ovako se dobivaju vremenske ovisnosti kutova u vidu vremenskih serija. Korelacija se zapravo izračunava tako da se uspoređuju dvije takve serije te se može izračunati njihov linearni korelacijski koeficijent definiran sljedećim izrazom:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \qquad (1)$$

gdje x_i i y_i predstavljaju vrijednosti bilo koja dva kuta, a indeks i ide po svim vremenskim točkama (i = 1 – 400.000). No budući da kutovi predstavljaju kružne (cikličke ili cirkularne) varijable čije vrijednosti se protežu od 0 do 360°, ispravnije je koristiti takozvanu cirkularnu statistiku. Tako je umjesto gornjeg linearnog korelacijskog koeficijenta korišten tzv. cirkularni korelacijski koeficijent definiran kao:

$$r_c = \frac{\sum_{i=1}^{n} sin(\phi_i - \bar{\phi}) sin(\psi_i - \bar{\psi})}{\sqrt{\sum_{i=1}^{n} sin^2(\phi_i - \bar{\phi})} \sqrt{\sum_{i=1}^{n} sin^2(\psi_i - \bar{\psi})}} \qquad (2)$$

Tako definiran cirkularni korelacijski koeficijent r_c može se onda izračunati za sve parove kutova φ i ψ za pojedinu trajektoriju. Računalno je to vrlo zahtjevna operacija budući da je potrebno skalarno pomnožiti sve parove stupaca u tablici kutova (spremljene u programu Pandas) sa preko milijardu elemenata. No uz neke matematičke operacije taj se izračun može svesti na množenje matrice dimenzija (2792, 400000) sa matricom dimenzija (400000, 2792). Uz korištenje vrlo brzih vektoriziranih operacija u Python programu za numeričko računanje Numpy i gotovih funkcija za cirkularne korelacije u programu AstroPy[131] taj izračun postaje moguć u relativno kratkom vremenu. Takav izračun daje simetričnu korelacijsku matricu dimenzija (2792, 2792) koja sadrži 3.896.236 nezavisnih korelacija.

Također na tim istim vremenskim serijama kutova primijenile su se metode pronalaska točaka promjene (engl. *change points*) upotrebom alata Ruptures. Budući da svaka aminokiselina ima φ ili ψ kut (osim prve i zadnje), vremenske serije su se računale istovremeno na oba od njih te bi se pomoću ovog programa zapravo tražile promjene u vrijednosti nekog od ova dva kuta. Usporedno prateći oba moguće je odrediti u kojem trenutku dolazi do promjene kod neke određene aminokiseline. Bitno je napomenuti da ovaj program automatski detektira promjene pa je potrebno obratiti pozornost dolazi li do promjene koja ima značaja (dovoljno velika amplituda) ili do neke koja nema. Ovaj programski paket omogućio je praćenje vremenske udaljenosti detektiranih događaja te također to povezanost s već opisanim cirkularnim korelacijama. Samo vizualna provjera može potvrditi da se kod nekih dviju visoko koreliranih aminokiselina mogu pronaći vremenski veoma bliske točke promjene nekog ili oba φ ili ψ kuta. Naposljetku je bilo potrebno kombinirati sve navedene pristupe u jedinstvenu sliku,

uz upotrebu naprednih vizualizacijskih tehnika pomoću različitih Python modula (NetworkX,[132] Bokeh,[133] Seaborn,[134] Matplotlib[135]).

Na kraju je bilo potrebno iz svih pronađenih točaka promjena na neki način automatizirati pronalazak putova komunikacije. Tu se kao pogodnom pokazala reprezentacija cijelog HpPNP enzima kao posebno konstruiranog grafa gdje aminokiseline predstavljaju čvorove (engl. *nodes*). Za svaku aminokiselinu definira se njezina okolina kao sve aminokiseline koje su udaljene manje od određene udaljenosti *d* (npr. $d(C_\alpha\text{-}C_\alpha) < 10$ Å). Onda se definira da između aminokiselina postoji jedna veza (engl. *edge*) za svaki par njihovih točka promjena koje su bliske, tj. gotovo istovremene. Tehnički je još bolje ne gledati točnu vremensku udaljenost već povući veze za sve moguće parove točaka promjene ali tako da se vezama pridjele težine (engl. *weights*) koje su znatne samo ako su točke vremenski blizu. Ukupan rezultat takvog postupka je da se cijela trajektorija prikaže kao neusmjereni graf s težinama (engl. *undirected weighted graph*). Naravno, mnogi će čvorovi u tom grafu imati jako puno veza prema drugim čvorovima, ali poanta cijelog postupka je da se problem prebaci u oblik za koji već postoji rješenje u vidu pronalaska svih mogućih putova između svih čvorova. To možemo napraviti upotrebom već dobro razrađenih algoritama za procesiranje grafova, a ovdje je korišten Python program Networkx. U njemu je moguće čak i za ovako kompleksne grafove s jako puno veza pronaći najkraće putove između bilo koja dva čvora (tj. aminokiseline). Tijekom pronalaženja putova program Networkx izračunava i zbroj svih težina duž pojedinog puta i od svih mogućih putova (kojih između svakog para čvorova u ovakvom kompleksnom grafu ima jako puno) pronalazi onaj s najmanjom ukupnom težinom. Zapravo traže se samo oni putovi za koje je ukupna težina jako mala a to znači da su pojedine promjene jako blizu u vremenu, što je oznaka njihove vjerojatne povezanosti.

Vizualizacija je zapravo jedna od najzahtjevnijih komponenata kod ovakve vrste analiza budući da se radi o izuzetno velikim količinama podataka. Za neke od analiza napravljena je redukcija podataka odnosno smanjena je količina uzorka koji se vizualizira. Zbog složenosti analize i velikog skupa podataka, upotrebom svih navedenih alata i pristupa, razvile su se vlastite programske rutine pomoću kojih su se pronašli različiti vremenski uzorci ponašanja aminokiselina, njihove međusobne interakcije te komunikacija između oligomernih podjedinica.

3.4. Programi i računala

Tijekom istraživanja korišten je širok spektar računalnih programa i resursa prilagođenih simulacijama molekulske dinamike, analizi velikih skupova podataka i vizualizaciji proteinskih struktura. Molekulske simulacije primarno su se provodile pomoću programskog paketa AMBER, dok su se neke od analiza, poput DSSP analize, provodile pomoću programskog paketa AmberTools. Uz AMBER, korišten je i programski paket GROMACS u svrhu ispitivanja funkcionalnosti i brzine te planiranja daljnjih simulacija.

Dobiveni podaci iz simulacija analizirani su pomoću programskog jezika Python, koristeći biblioteke NumPy, SciPy, AstroPy, scikit-learn, Pandas, Biopython i MDAnalysis za numeričke proračune, statističku analizu i obradu trajektorija molekularne dinamike. Program MDAnalysis poslužio je za većinu obrada podataka i izračun određenih geometrijskih parametara kao što su vrijednosti RMSD, udaljenosti između atoma, izračun torzijskih kutova, itd. Kako bi se dublje analizirale komunikacijske putanje unutar enzima HpPNP, koristio se NetworkX za graf teorijsku analizu, dok se za identifikaciju ključnih prijelaznih stanja primjenjivao program Ruptures, algoritam za detekciju promjena u vremenskim serijama. Razvijen je također i vlastiti alat MDavocado, koji omogućava sažetu vizualizaciju i analizu kutnih korelacija unutar proteinskih struktura, čime se značajno poboljšava interpretacija dinamičkih podataka. Za vizualizaciju velikih količina podataka na jednom grafu koristio se program Datashader, koji je specijaliziran upravo za to.

Vizualizaciju struktura i dinamičkih procesa provodila se korištenjem programa PyMOL koji je omogućio jasan uvid u promjene konformacija i međumolekularne interakcije. Za vizualizaciju obrađenih podataka i izradu grafova također su se koristili programi dostupni u programskom jeziku Python poput Seaborn, Matplotlib i Bokeh.

Za izvođenje simulacija koristila se lokalna radna stanica Supermicro opremljena višejezgrenim procesorima i grafičkom karticom koja je poslužila za cjelokupnu pripremu sustava i samu produkcijsku fazu MD simulacija. Također je za sve obrade podataka u svrhu procesiranja i vizualizacije korištena navedena lokalna radna stanica. Uz Supermicro, MD simulacije su se izvodile na superračunalnim klasterima Isabella (Sveučilišni računalni centar Sveučilišta u Zagrebu) i EPCC (Sveučilište u Edinburghu).

§ 4. REZULTATI I RASPRAVA

Osnovna motivacija za ovo istraživanje je pronalazak načina i putova komunikacije između pojedinih dijelova proteina s nekim udaljenim dijelovima proteina, pojava koja se naziva alosterija. To je u osnovi fenomen dinamičke prirode te ako ga želimo rasvijetliti moramo posegnuti za metodama koje uzimaju u obzir dinamiku samog proteina. Kao polazna točka ovog istraživanja služi skup kristalografski određenih struktura enzima purinske nukleozidne fosforilaze iz bakterije *H. pylori* (HpPNP). Iako te strukture daju trodimenzijsku strukturu tih enzima na atomskoj rezoluciji, iz njih samih ne možemo doznati gotovo ništa o samoj dinamici enzima jer su te strukture po definiciji statične. Stoga su istraživanja bazirana na MD simulacijama dinamike tih enzima od izuzetne pomoći kod ovakve tematike. Točnije radi se o šest kristalografski određenih struktura HpPNP enzima (PDB kodovi: 5lu0, 5mx4, 5mx6, 6f4x, 6f52, 6g7x) i dvije iz njih izvedene uklanjanjem određenih liganada te su one poslužile kako bi se što bolje ispitao konformacijski prostor gibanja ovog enzima. Korištenje većeg početnog (različitog) skupa podataka i izvođenje simulacija na što dužoj vremenskoj skali rezultiralo je raznim promjenama u HpPNP enzimu, odnosno promjenama konformacije kako na maloj, tako i na velikoj skali. Osnovna pretpostavka alosteričke komunikacije u HpPNP enzimu je da se ona odvija preko konformacijskih promjena na razini pojedinih aminokiselina, odnosno promjenom u okosnici samih aminokiselina ili u promjeni njihovih rotamera. Uz lokalne promjene, u HpPNP enzimima se mogu desiti i globalne konformacijske promjene koje mogu dovesti do npr. snažnijeg vezanja supstrata ili promjene u enzimskoj reakciji. U sklopu ovog istraživanja razvijena je metoda i programski alat za praćenje lokalnih promjena u vremenu čime se posljedično mogu pratiti i globalne konformacijske promjene. Također su pronađene poveznice između vezanja fosfata u aktivno mjesto i konformacijskih promjena (njegova otvaranja, odnosno zatvaranja). Pomoću izračunatih cirkularnih korelacija vremenskih serija promjena kutova okosnice φ i ψ te njihova povezivanja s konformacijskim promjenama pronađeni su centri koji imaju veliku važnost kako u komunikaciji između monomernih podjedinica ovog enzima tako i u komunikaciji koja ili uzrokuje globalnu konformacijsku promjenu ili joj je posljedica. Cjelokupna analiza i rezultati dio su interaktivne mrežne stranice ALOKOMP (https://alokomp.irb.hr/, datum pristupa 13. ožujka 2025.) koja se navodi u daljnjim odjeljcima kako bi se bolje opisala dinamička priroda nekih od rezultata. Statičke slike

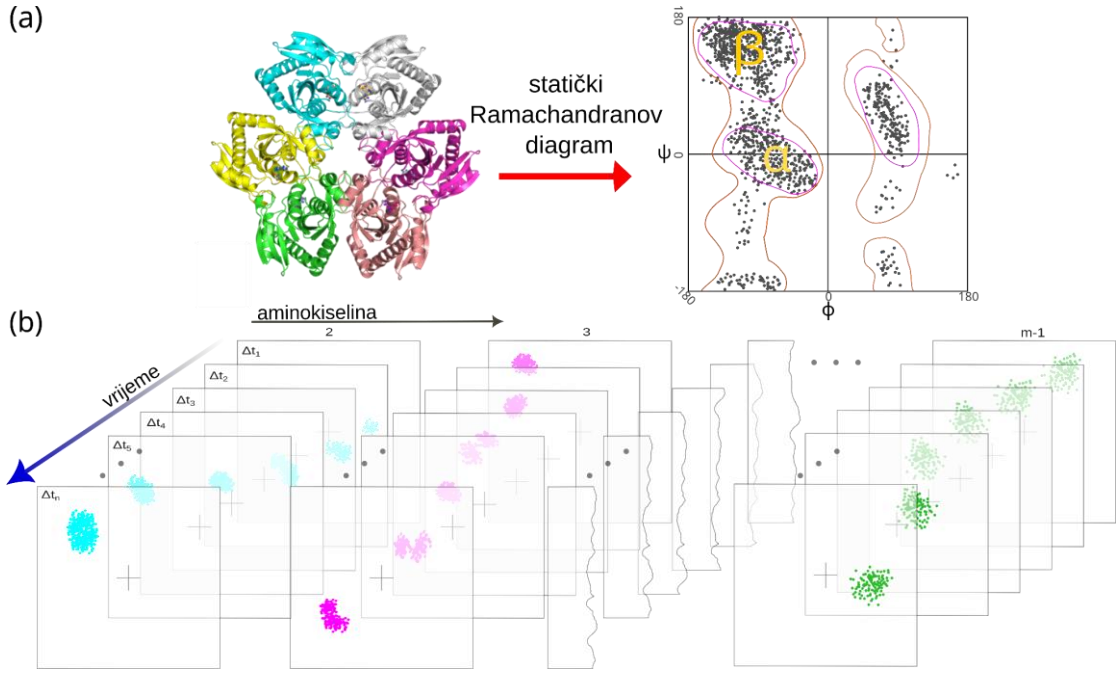
u disertaciji vode na mrežne stranice ALOKOMP gdje se mogu vidjeti njihovi dinamički ekvivalenti. Ovakvim računalnim istraživanjem dobiva se snažna osnova za daljnje eksperimente kojima bi se još jednom potvrdila važnost PNP enzima kod eradikacije bakterije *H. pylori*.

4.1. Razvoj metodologije MDavocado dijagrama

Praćenje konformacijskih promjena kroz MD simulacije može se izvesti promatranjem položaja atoma u aminokiselinama ili npr. torzijskih kutova. Kutovi imaju određenu prednost jer su neovisni o translaciji atoma unutar same trajektorije. Kao odabir torzijskih kutova koji dobro odražavaju konformaciju proteina, a pogotovu njegove okosnice, odabrani su Ramachandranovi kutovi φ i ψ. Ramachandranov dijagram (slika 13) danas je standardan način validacije strukture proteina i u tu svrhu koristi se već dugi niz godina. Ramachandranovi kutovi opisuju torziju okosnice svake aminokiseline u proteinu, osim prve i zadnje u svakom lancu, jer za prvu aminokiselinu nije dobro definiran torzijski kut φ, a za zadnju torzijski kut ψ. U klasičnom Ramachandranovom dijagramu prikazuje se međusobna ovisnost φ i ψ kutova za svaku pojedinu aminokiselinu, tj. svaka aminokiselina je jedna točka na tom dijagramu u φ-ψ ravnini čije se vrijednosti najčešće prostiru od -180° do 180°. U ovom se dijagramu ističu pojedine zone koje su određene dopuštenim vrijednostima ovih kutova a koje su određene steričkim smetnjama aminokiselinskih bočnih ogranaka. Nadalje, ova područja su karakteristična za sekundarnu strukturu pa se na dijagramu jasno izdvajaju područja α-zavojnice i β-nabrane ploče. Ramachandranov dijagram je dakle pogodan za brzi pregled na statičku sliku proteinske konformacije u danom trenutku, i jedna struktura odgovara jednom Ramachandranovom dijagramu. On sadrži potencijalno veliki broj točaka koji je ovisan samo o veličini promatranog proteina, ali najčešće se radi o stotinama ili tisućama. Ako želimo vidjeti dinamiku proteina, statički Ramachandranovi dijagrami ne mogu dati nikakav uvid u nju. U svom nemodificiranom obliku vremenska ovisnost gibanja (promjena kutova) rezultirala bi slaganjem niza Ramachandranovih dijagrama jedan na drugi za sve vremenske točke u trajektoriji, a takvih je najčešće na stotine tisuća. U sklopu ove disertacije tom se problemu pristupilo na novi način. Umjesto da se cijeli Ramachandranovi dijagrami slažu jedan na drugi, Ramachandranov dijagram je razložen na pojedine aminokiseline te se takvi posebni manji dijagrami slažu u vremenu. Takvi zasebni posebno prilagođeni dijagrami nazvani su

MDavocado dijagrami (što dolazi od engl. akronima **MD** ***A****nalysis and* ***V****isualization* ***Of*** ***C****orrelated* ***A****ngular* ***D****iagrams*).

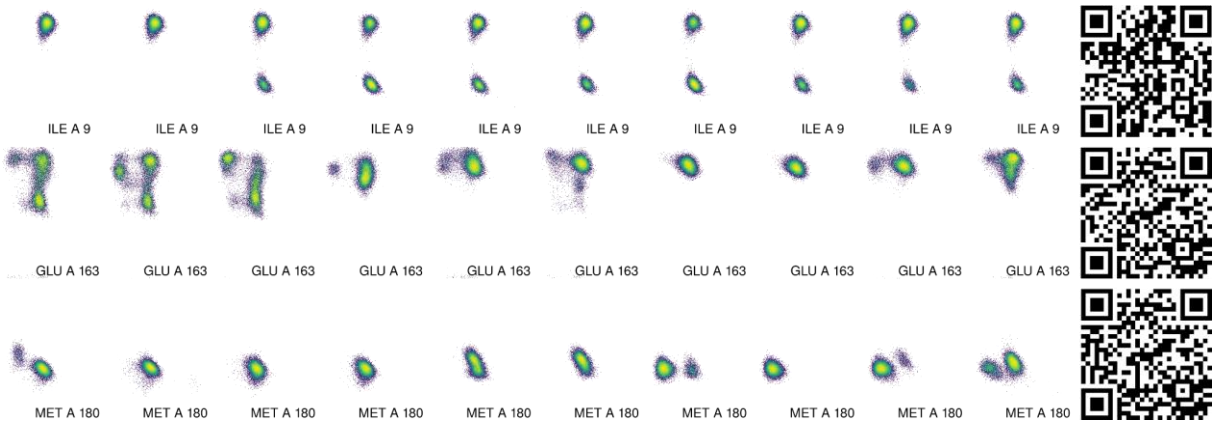
U tu svrhu napravljen je program MDavocado, za analizu i vizualizaciju kretanja proteina pomoću vremenski ovisnih kutnih dijagrama. Ovim programom trajektorije se obrađuju tako što se izračunavaju Ramachandranovi kutovi za svaku aminokiselinu tijekom cijele trajektorije. Dobiveni podaci se prikazuju u setovima individualnih dijagrama za svaku aminokiselinu zasebno, što se razlikuje od klasičnih prikaza, tako da se φ i ψ ovisnost vizualizira u vremenu. Najveća razlika između klasičnog Ramachandranovog i MDavocado dijagrama je količina točaka koje se moraju prikazati na svakom od njih. U klasičnom Ramachandranovom dijagramu broj točaka odgovara broju aminokiselina i ima onoliko dijagrama koliko koraka ima simulacija, dok kod MDavocado dijagrama broj točaka odgovara broju koraka simulacije i dijagrama ima koliko protein ima aminokiselina. Kod razvoja programa MDavocado korištena je tehnika agregacije točaka, dostupna u programu Datashader, koja omogućava vizualizaciju velikog broja točaka na jednom dijagramu čime se zapravo dobiva m-2 (gdje m predstavlja broj aminokiselina) dijagrama (slika 24). Naime, prva i zadnja aminokiselina nemaju definirana oba, φ i ψ, kuta te su one izuzete iz same analize. Cijela MD simulacija ima N koraka sa ukupnim vremenom simulacije T. Ukupno vrijeme simulacije se kod ove analize podijeli na n jednakih dijelova čime se dobivaju vremenski intervali $\Delta t = T/n$. Budući da je broj koraka simulacije N obično puno veći od nazivnika n, svaki pojedini interval će svejedno imati veliki broj točaka (u ovom slučaju oko 40.000). Konstrukcijom ovakvih dijagrama može se precizno detektirati kretanje svake pojedinačne aminokiseline u određenom vremenskom intervalu trajektorije.



Slika 24. Procedura izgradnje MDavocado dijagrama. (a) Konformacija proteina u određenom trenutku u vremenu može se prikazati statičkim Ramachandranovim dijagramom gdje svaka točka predstavlja jednu aminokiselinu. Dijagram ilustrira cjelinu, odnosno cijeli protein odjednom te se na njemu vide sve kombinacije φ i ψ kutova koje odgovaraju konformaciji aminokiselina kod formiranja sekundarnih struktura. (b) MDavocado dijagrami prikazuju dinamiku proteina u vremenu. Ukupno vrijeme trajektorije podijeljeno je u *n* jednakih vremenskih intervala Δt za svaku aminokiselinu numeriranu od *2* do *m-1*. Na takav način konstruirani su φ i ψ dijagrami gdje svaki kvadrat odgovara agregaciji točaka u vremenskom intervalu. (preuzeto i prilagođeno iz reference [9])

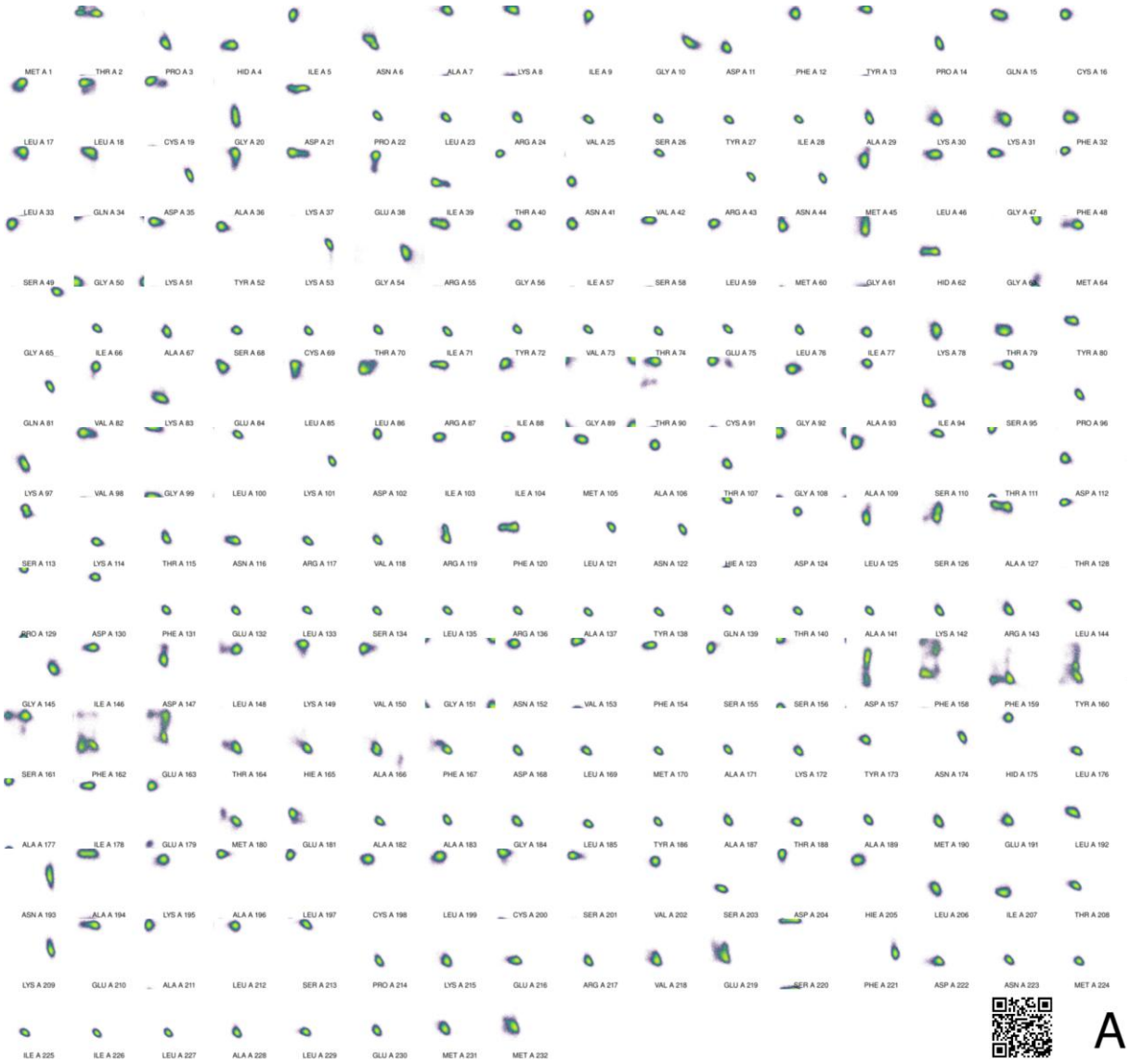
Tehnika agregacije dostupna u programu Datashader omogućava redukciju podataka i njihovu vizualizaciju na način da se prikaže veliki broj točaka na dvodimenzijskoj ravnini bez da se točke preklapaju i bez da se izgube esencijalne promjene u podacima koje bi se možda nekim drugim tehnikama zagubile zbog velike količine (šume) podataka. U suštini ova metoda projicira sve točke na dvodimenzijsku ravninu, agregira te podatke i reprezentira gustoću točaka u svakom pikselu pomoću prikladnog raspona boja. Također, ova tehnika automatski skalira proces agregacije što omogućava već spomenuti pregledni raspored svih točaka i izbjegava gubitak nedominantnih promjena u setu podataka. Redukcijom podataka na određen broj dijagrama jednakih vremenskih intervala mogu se utvrditi konformacijske promjene svake aminokiseline pojedinačno kroz cijelo trajanje trajektorije. Na slici 25 može se vidjeti 10

pojedinačnih slika koje odgovaraju intervalima MDavocado dijagrama za tri različite aminokiseline iz simulacije *6f4x_po4_fmc* gdje je prisutna vidljiva karakteristična promjena φ ili ψ kuta čija se progresija jasno može iščitati iz ovakvog prikaza.



Slika 25. Prikaz promjene φ ili ψ kuta triju različitih aminokiselina (izoleucin Ile9, glutamat Glu163 i metionin Met180) pomoću 10 vremenskih intervala MDavocado dijagrama. S lijeva na desno vidi se vremenski slijed simulacije. Na krajnjem desnom rubu slike nalazi se QR-kod koji vodi na ALOKOMP mrežne stranice gdje se za svaku od ovih aminokiselina može vidjeti animirana slika (GIF) promjene kutova u MDavocado dijagramu. Koordinatne osi i vrijednosti uklonjene su zbog lakšeg prikaza vremenske progresije.

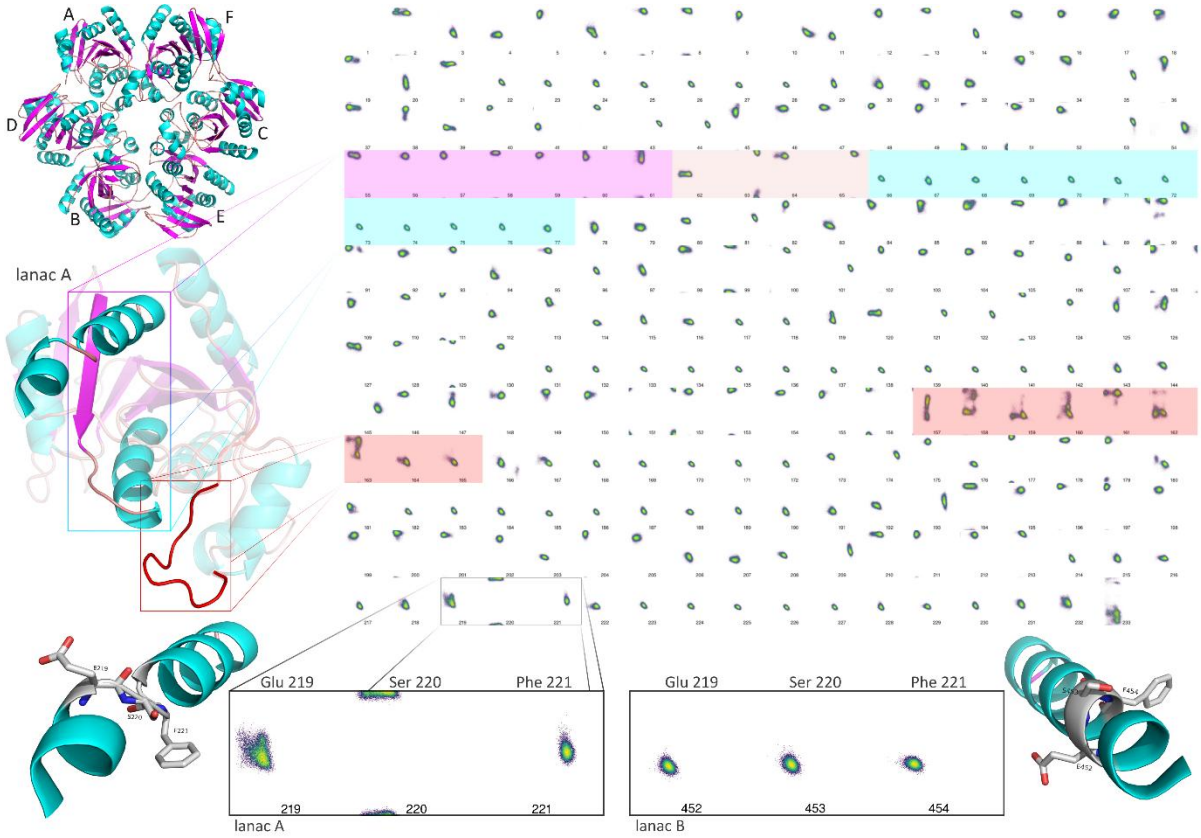
Kombinacijom dobivenih slika u vremenu, odnosno redanjem tih slika jedne nakon druge dobiva se glavni rezultat MDavocado programa kao vremenski ovisna slika (GIF slika) promjene Ramachandranovih kutova. Dodatnom kombinacijom gifova svih aminokiselina, dobiva se animacija kretanja pojedinih dijelova proteina: lanaca, domena ili proteina u cijelosti. Primjerice takav prikaz kretanja cijelog lanca iz simulacije *6f4x_po4_fmc* koji uključuje aminokiseline sa slike 25 može se vidjeti na slici 26. Pojedine slike u ovim radu su nužno samo statični prikazi, i u takvim slučajevima je dodan QR kod koji vodi na mrežne stranice ALOKOMP gdje se može vidjeti njihova animirana verzija. Ovakav pristup omogućava sveobuhvatan pregled kretanja odnosno dinamike dijela proteina od interesa u jedinstvenoj animaciji. Samo promatranjem ovakvih GIF animiranih slika odmah se mogu uočiti lokalni dijelovi proteina kod kojih dolazi ili ne dolazi do promjene kutova što ukazuje na dijelove proteina koji su fleksibilni, odnosno rigidni.



Slika 26. Slika jednog vremenskog intervala kretanja svih aminokiselina lanca A iz simulacije strukture *6f4x_po4_fmc* pomoću MDavocado dijagrama.

MDavocado dijagrami pružaju jako dobar uvid u gibanje aminokiselina tijekom MD simulacija, otkrivajući ključne značajke koje su često skrivene u sirovim podacima trajektorija. Rezultati ovih analiza jasno pokazuju da je gibanje aminokiselina u velikoj mjeri uvjetovano njihovom sekundarnom strukturom. Naime, aminokiseline koje se nalaze u stabilnim sekundarnim strukturama, poput α-zavojnice i β-ploče, uglavnom su stacionarne i pokazuju vrlo ograničenu mobilnost. S druge strane, najdinamičniji dijelovi proteina su petlje i fleksibilne regije koje povezuju ove sekundarne strukture. Ovakvi podaci jasno su vidljivi u dijagramima koje su rezultat MDavocado analize (slika 27). Dijagrami također omogućuju globalni pregled

dinamike cijelog proteina, pružajući prikaz koji jasno ističe pojave sinkroniziranih promjena između različitih područja proteina. Takve sinkronizirane prijelaze često je teško uočiti klasičnim tehnikama analize MD simulacija, no uz MDavocado pristup postaju vrlo uočljive. Na globalnoj razini, ovi dijagrami otkrivaju da velika većina aminokiselina skoro cijelo vrijeme trajanja simulacije provodi u jednom dominantnom, stacionarnom stanju. Međutim, manji broj aminokiselina pokazuje sposobnost izmjene između dva ili, u nekim slučajevima, nekoliko jasno definiranih stanja u φ–ψ prostoru. Prijelazi između ovih različitih stanja potencijalno su ključni za funkcionalne procese poput otvaranja i zatvaranja aktivnih mjesta ili interakcije s ligandima. Bez ovakvog pojednostavljenja podataka iz MD trajektorija i njihove vizualne projekcije, takve dinamičke promjene u proteinima vrlo je teško uočiti. MDavocado dijagrami, zahvaljujući svojoj sposobnosti integracije i kompresije velike količine podataka, predstavljaju ključan alat u analizi molekularnih simulacija i otkrivanju skrivenih obrazaca u dinamici bioloških makromolekula.



Slika 27. Skupni prikaz MDavocado dijagrama jednog lanca strukture *6f4x_po4_fmc* gdje su jasno definirane regije poput *α*-zavojnice (tirkizno) i *β*-ploče (rozo) koje pokazuju ograničenu mobilnost. Uz ove stacionarne regije proteina se ističu i veoma mobilne regije poput petlji (crveno). (preuzeto i prilagođeni iz reference [9])

4.2. Analiza H8 *α*-zavojnice i aktivnog mjesta

Konformacijska promjena koja se uočava kod HpPNP enzima je segmentacija H8 *α*-zavojnice na dvije manje *α*-zavojnice. H8 je oznaka osme po redu *α*-zavojnice u svakom monomernom lancu enzima HpPNP te je ona sastavljena od 20 aminokiselina koje imaju pozicije od 213 do 233 u sekvenci samog enzima. Segmentacija navedene *α*-zavojnice događa se između aminokiselina serin Ser220 i fenilalanin Phe221 čime se dobivaju dvije manje *α*-zavojnice odijeljene veoma kratkom petljom. Na slici 9b u Literaturnom pregledu može se vidjeti navedena konformacijska promjena. Također, kada se govori o nesegmentiranoj *α*-zavojnici, odnosno o njenom izduženom obliku, riječ je o otvorenoj konformaciji aktivnog mjesta. U

suprotnom slučaju, kada se govori o segmentiranoj α-zavojnici, riječ je o zatvorenoj konformaciji aktivnog mjesta.

Za analizu prethodno opisane konformacijske promjene koristi se nekoliko različitih metoda kako bi se što bolje opisala segmentacija H8 α-zavojnice. Analiza sekundarnih struktura, odnosno DSSP analiza (engl. *Define Secondary Structure of Proteins*) dodjeljuje simbole aminokiselinama čime se te aminokiseline kategoriziraju u neku od poznatih sekundarnih struktura. Ova analiza napravljena je pomoću programa *secstruct* dostupnog u Amber programskom paketu. Suština ovakve metode je izračun pripadajuće sekundarne strukture svake aminokiseline na temelju pozicija amidnog dušika i karbonilnog kisika iz okosnice. Metoda prepoznaje različite sekundarne strukture poput paralelnih i antiparalelnih β-ploča ili α-zavojnice i 3-10 α-zavojnice. Rezultat DSSP analize prikazan je bojama gdje ljubičasta boja označava aminokiseline koje nisu kategorizirane u nijednu sekundarnu strukturu. Zelenom bojom su označene aminokiseline svrstane u kategoriju sekundarne strukture 3-10 α-zavojnice, α-zavojnice i π (3-14) α-zavojnice. Žutom bojom su označene aminokiseline svrstane u kategoriju sekundarne strukture okreta. Oznake 3-10 i 3-14 kod α-zavojnica označavaju način formiranja aminokiselina u sekundarnoj strukturi, odnosno koliko aminokiselina čine jedan ciklički motiv vodikovih veza u α-zavojnici. 3-10 α-zavojnica ima 10 članova što ju čini zbijenijom dok 3-14 α-zavojnica ima 14 članova što ju čini raširenijom.

Također se za analizu konformacijske promjene koriste MDavocado dijagrami kod kojih su jasno definirane regije sekundarnih struktura α-zavojnice ili β-ploče te se odstupanje iz navedenih regija može smatrati promjenom sekundarne strukture za promatranu aminokiselinu.

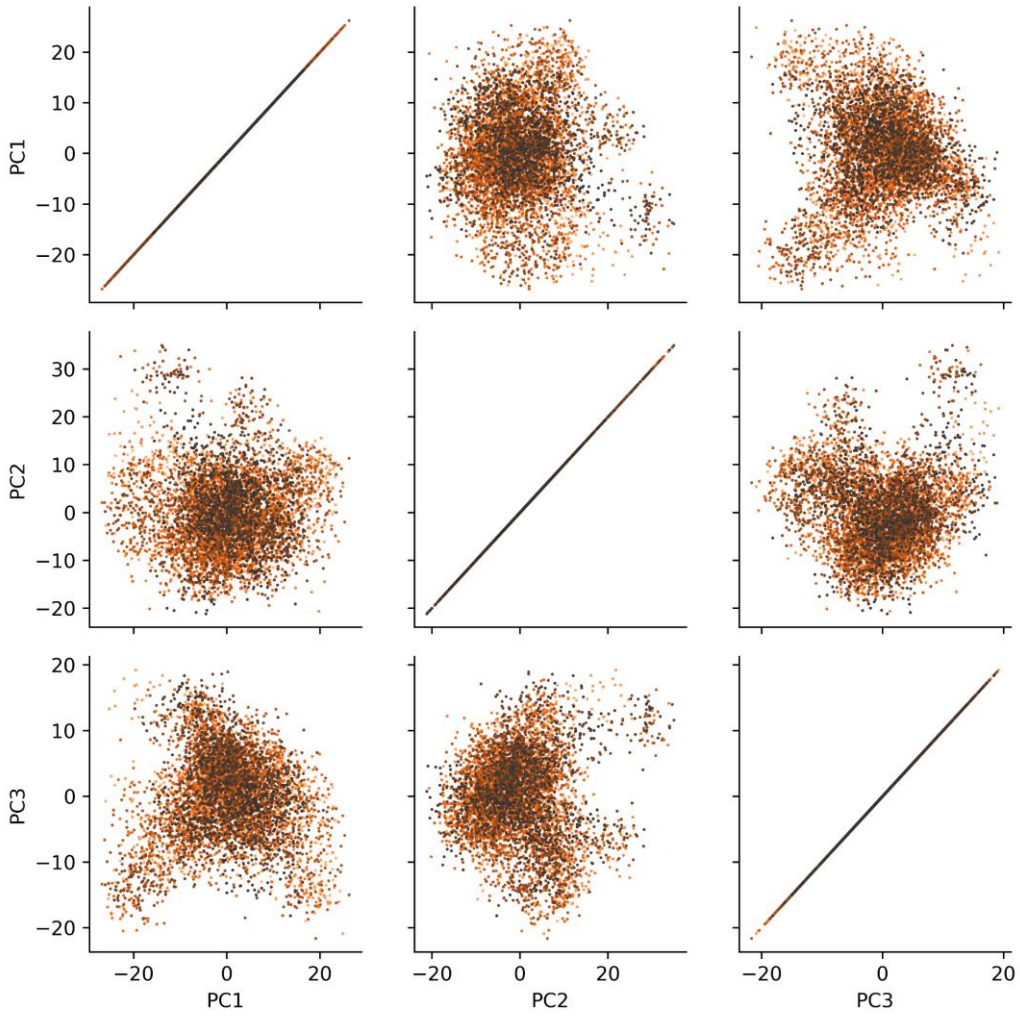
Još jedna od metoda proučavanja konformacijske promjene nekog dijela proteina je izračun RMSD vrijednosti. Pomoću ove metode uspoređuju se promatrana struktura i referentna struktura te se računa srednja udaljenost atoma u promatranoj strukturi svakog koraka simulacije naspram atoma u referentnoj strukturi. Kod analize H8 α-zavojnice računala se RMSD vrijednost samo atoma okosnice sadržanih u zavojnici naspram istih tih atoma iz referentne strukture. Kao referentna struktura koristila se kristalna struktura PDB koda 6f52 budući da se kod nje aktivno mjesto svih 6 monomernih podjedinica nalazi u otvorenoj konformaciji, odnosno H8 α-zavojnica je u nesegmentiranom obliku. Iz RMSD analize može se ugrubo tvrditi dolazi li do promjene konformacije aktivnog mjesta ili ne, budući da se ovom analizom računa srednja udaljenost atoma promatrane strukture od atoma referentne strukture što bi značilo da bilo kakav pomak doprinosi povećanju RMSD vrijednosti. Rezultati ove

analize mogu se tumačiti na način da ako su RMSD vrijednosti niske, promatrana α-zavojnica se nalazi u nesegmentiranoj formi, a ako dolazi do povećanja RMSD vrijednosti, kod promatrane α-zavojnice dolazi do pomaka, odnosno segmentacije.

4.2.1. Analiza glavnih komponenata

Jedna od najučestalijih metoda za praćenje globalnih konformacijskih promjena je analiza glavnih komponenata (PCA) enzima budući da se njome dobivaju glavne kretnje enzima u MD simulaciji. U slučaju HpPNP radi se o promatranju zatvaranja, odnosno otvaranja šest aktivnih mjesta što je dosta izazovno budući da se radi o malom prostornom pomaku u odnosu na cijeli protein. Rezultati te analize to i potvrđuju. Analiza je napravljena spajanjem svih izračunatih trajektorija budući da se radi o istovjetnom sustavu te se iz nje nažalost nisu mogle iščitati nikakve jasno odvojene konformacijske promjene. Pretpostavka je da se u ovom slučaju radi o prevelikom setu podataka te također o prevelikom šumu koji nastaje tijekom analize. Stoga se konformacijske promjene od našeg interesa nisu izdvojile te ih posljedično nije bilo moguće ni detektirati. Grupiranje je korišteno kako bi se identificirali klusteri u samom setu podataka i time automatski izdvojila različita konformacijska stanja ovog enzima. Zbog široke i guste raspodjele podataka koristila se hijerarhijska metoda grupiranja koja najviše odgovara ovakvom tipu podataka. Upotrebom Ruptures programa na RMSD podacima promjene iz otvorenog aktivnog mjesta u zatvoreno i obratno, pronađene su točke promjene što daje točno određeni vremenski trenutak. Izolirani su svi takvi događaji te su se iz trajektorija izolirali dijelovi na način da se uzelo 5000 koraka ispred i iza svakog od njih i kreirala se "nova" trajektorija koja sadrži samo događaje od interesa, odnosno samo otvaranje/zatvaranje aktivnog mjesta. Spajanjem takvih dijelova pokušalo se izbaciti što više podataka koji uvode šum u rezultate. Analizom glavnih komponenata na novo dobivenoj trajektoriji također nisu pronađene promjene koje bi se izdvojile i time potvrdile globalne konformacijske promjene. Još jedan način smanjenja uzorkovanja je strukturno skaliranje podataka gdje se uzimaju u obzir samo aminokiseline koje su dio H8 α-zavojnice, te su ovom analizom uočeni manji pomaci u vrijednostima prvih triju glavnih komponenata što može ukazivati na manje lokalne promjene u samoj zavojnici budući da ona ima fleksibilan početak i kraj (slika 28). Vizualizacija trajektorije, odnosno projekcija glavnih komponenata na koordinate proteina (vizualizacija konformacijskih promjena) nije uzeta u obzir zbog kumulativne varijance glavnih komponenata

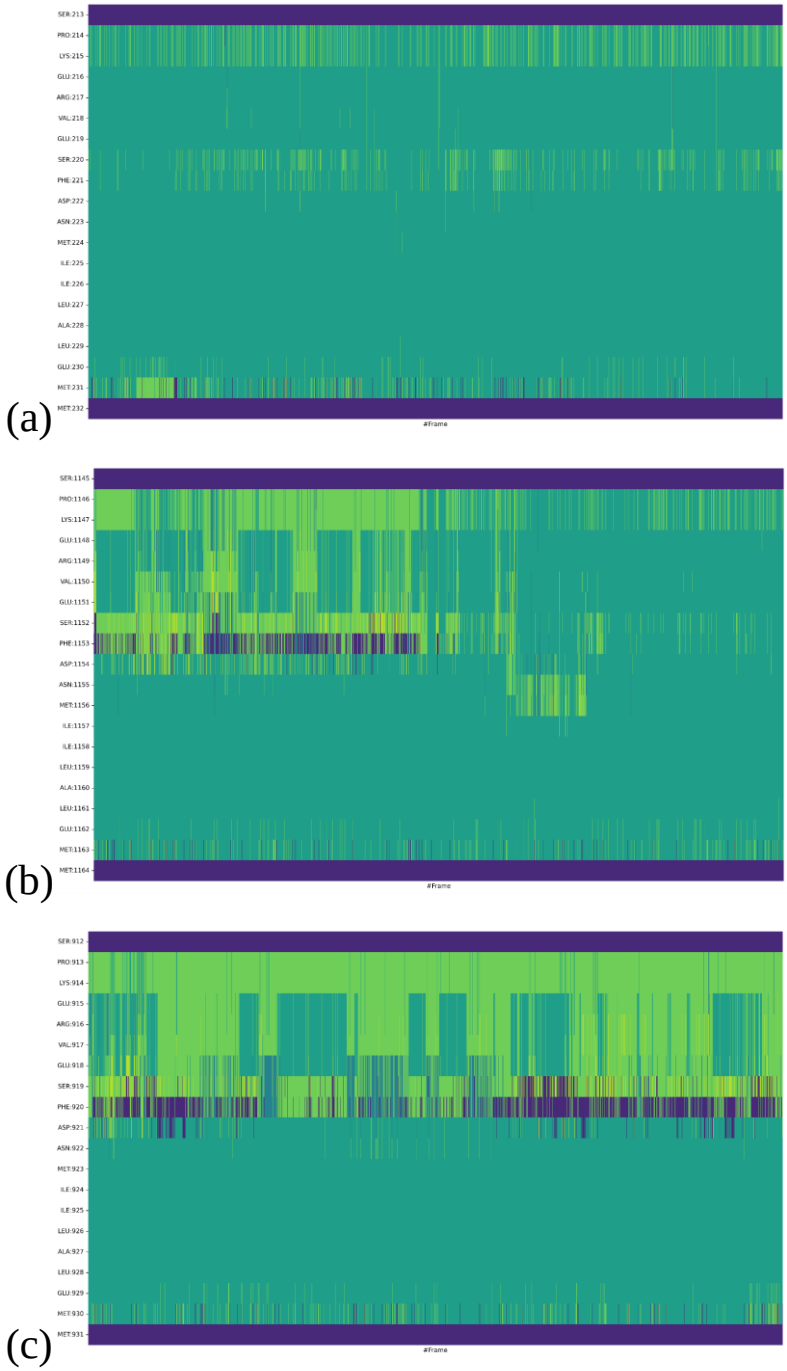
koje iznose 17,0 %, 14,6 % i 9,6 %. Zbog svega navedenog, PCA se nažalost nije koristila u daljnjoj analizi te je bilo potrebno pronaći druge načine i metode kako klasificirati promjene iz otvorenog aktivnog mjesta u zatvoreno i obratno. Kao zamjena za analizu glavnih komponenata koristile su se slične analize, poput analize sekundarnih struktura (DSSP analiza) ili MDavocado analize, kod kojih se proučava promjena konformacija na lokalnoj razini te je to opisano u sljedećim odjeljcima. Također se iskoristila metoda smanjivanja uzorkovanja baš zbog problema koji su se pojavili u ovoj analizi.



Slika 28. Međusobna ovisnost prvih triju glavnih komponenti HpPNP enzima, odnosno kretanja svih 6 H8 α-zavojnica na smanjenom setu podataka nastalom smanjenjem uzorkovanja na vremenskoj i prostornoj razini.

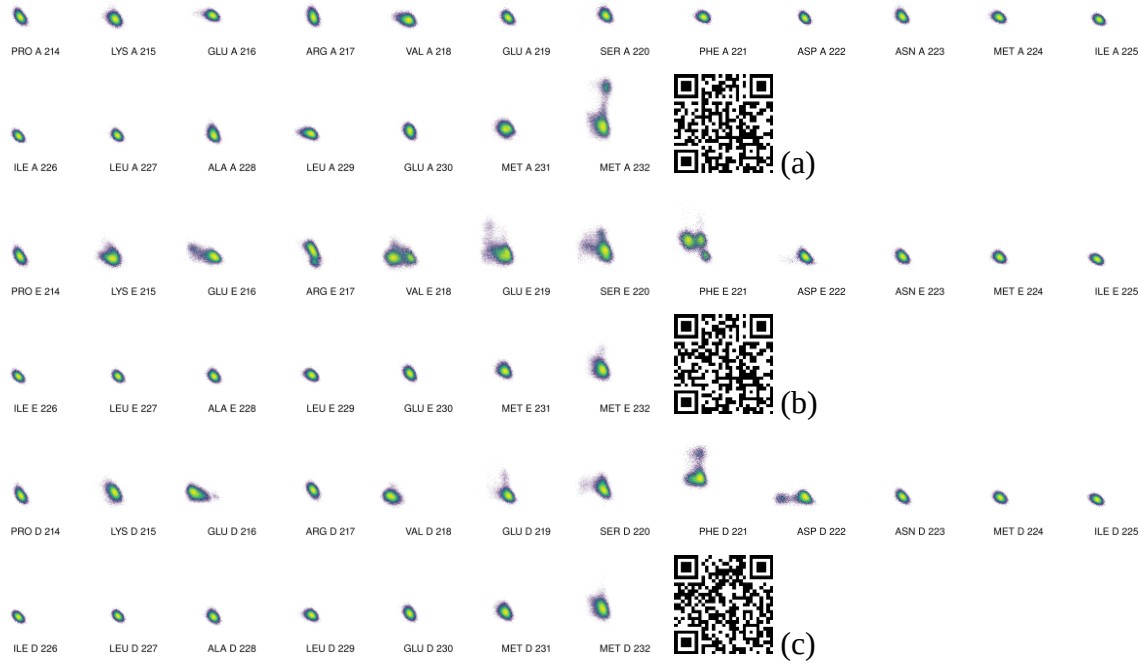
4.2.2. Analiza struktura bez supstrata

Kod struktura koje nemaju supstrata u aktivnom mjestu (*5lu0_apo, 5mx4_apo, 6f4x_apo* i *6f52_apo*) dolazi do ravnotežne promjene konformacijskog stanja aktivnog mjesta iz otvorenog u zatvoreno, no uočavaju se također i slučajevi kod kojih se konformacije zadržavaju tijekom cijelog trajanja simulacije. Iz DSSP analize može se vidjeti kako kod simulacija kod kojih je α-zavojnica nesegmentirana na početku, ona ostaje takva kroz cijelo vrijeme trajanja simulacije. Također se mogu pronaći i one simulacije kod kojih dolazi do jasnijeg razdvajanja konformacija odnosno konkretnije se vidi otvorenost i zatvorenost aktivnog mjesta, te se također uočava konformacija koja pokazuje da je α-zavojnica segmentirana kroz većinu vremena trajanja simulacije. Na slici 29 mogu se vidjeti primjeri za svaki od ova tri navedena slučaja, (a) prikazuje stalno otvorenu konformaciju aktivnog mjesta trajektorije *6f52_apo* replike 1, (b) prikazuje promjenu iz zatvorene u otvorenu konformaciju aktivnog mjesta trajektorije *6f4x_apo* replike 1, a (c) prikazuje zatvorenu konformaciju aktivnog mjesta trajektorije *6f4x_apo* replike 3 kroz cijelo vrijeme trajanja simulacije.

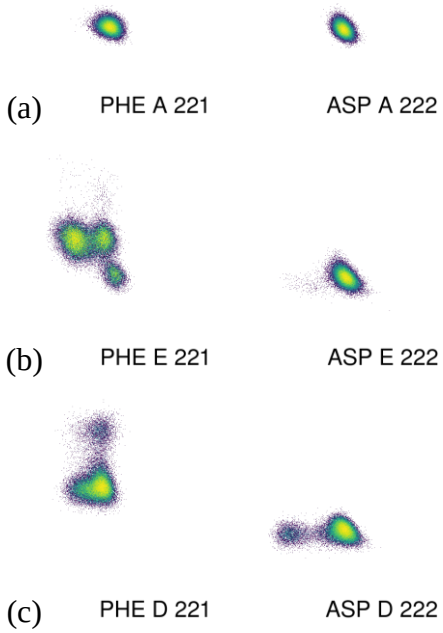


Slika 29. DSSP analiza simuliranih struktura (a) *6f52_apo*, (b) *6f4x_apo* i (c) *6f4x_apo* bez supstrata gdje se uočavaju (a) u potpunosti otvorena konformacija aktivnog mjesta, (b) promjena konformacije iz zatvorene u otvorenu konformaciju aktivnog mjesta te (c) zatvorena konformacija aktivnog mjesta u većem dijelu trajanja simulacije. Na svim slikama boje indiciraju različite sekundarne strukture: ljubičasta ne označava nikakvu sekundarnu strukturu, zelena boja predstavlja 3-10 α-zavojnice, α-zavojnice i π (3-14) α-zavojnice, dok žuta boja predstavlja okret.

Također na slici 30 vide se MDavocado dijagrami svake od pripadajućih simulacija, odnosno svaki prikaz DSSP analize odgovara jednom MDavocado dijagramu. Kako se kod MDavocado dijagrama zapravo radi o dinamičkim slikama, odnosno GIF animiranim slikama koje su sastavljeni od 10 slika, svaka od tih slika odgovara desetini trajektorije pa se one mogu nazvati vremenskim intervalima što bi značilo da npr. prvi vremenski interval odgovara prvoj desetini trajanja MD trajektorije. Na slici 30a prikazan je prvi (od deset) vremenski interval simulacije te se kod njega može vidjeti distribucija kutova koji odgovaraju α-zavojnici što upućuje na otvorenu konformaciju. Na slici 30b prikazan je peti (od deset) vremenskih intervala što bi odgovaralo polovici simulacije, a prema DSSP analizi (slika 29b) prijelazu iz zatvorene (ljubičasta i žuta boja kod aminokiselina Ser220 i Phe221 ukazuju na segmentiranu α-zavojnicu) u otvorenu konformaciju aktivnog mjesta (zelena boja kod aminokiselina Ser220 i Phe221 ukazuje na nesegmentiranu α-zavojnicu). Na slici 30b se može vidjeti poremećaj vrijednosti kuta aminokiseline Phe221 i njoj prethodnih aminokiselina (Lys215, Glu216, Arg217, Val218, Glu219 i Ser220) što indicira da se događa djelomična segmentacija α-zavojnice. Na slici 30c prikazan je drugi (od deset) vremenski interval trajanja simulacije budući da je kod njega najviše izražena promjena koja ukazuje na zatvorenu konformaciju aktivnog mjesta. Pokazatelj navedene promjene je odstupanje međuovisnosti torzijskih kutova aminokiselina Phe221 i Asp222 iz regije koja indicira na sekundarnu strukturu α-zavojnice. Na slici 31 mogu se vidjeti uvećani MDavocado dijagrami aminokiselina Phe221 i Asp222 te se na njima jasnije mogu vidjeti promjene u kutovima MDavocado dijagrama.



Slika 30. Prikaz MDavocado dijagrama simuliranih struktura (a) *6f52_apo*, (b) *6f4x_apo* i (c) *6f4x_apo* bez supstrata gdje su prikazane (a) otvorena konformacija aktivnog mjesta, (b) aktivno mjesto kod kojeg se mijenja konformacija iz zatvorene u otvorenu konformaciju (prikazan je peti od deset vremenskih intervala) te (c) zatvorena konformacija aktivnog mjesta kod koje se može vidjeti povećanje ψ kuta kod aminokiseline fenilalanin Phe221 i smanjenje φ kuta kod aminokiseline aspartat Asp222. Također se uz svaku sliku nalaze pripadni QR kodovi koji vode na dinamički prikaz promjena.



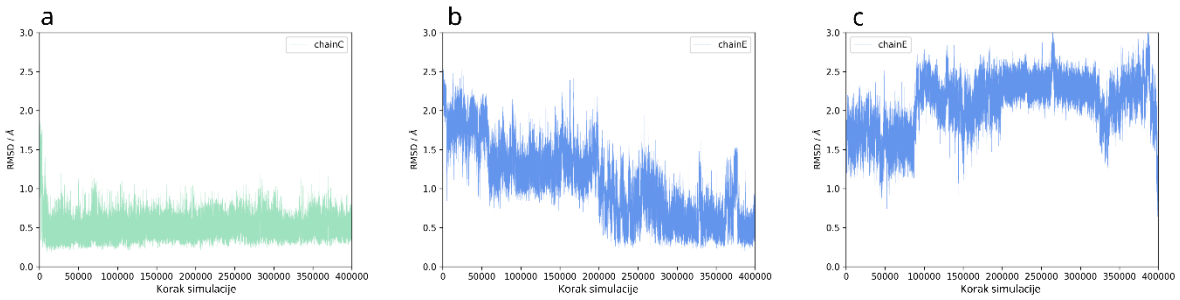
Slika 31. Uvećan prikaz MDavocado dijagrama aminokiselina fenilalanin Phe221 i aspartat Asp222 simuliranih struktura (a) *6f52_apo*, (b) *6f4x_apo* i (c) *6f4x_apo* kod kojih se uočava konformacijska promjena zatvaranja aktivnog mjesta. (a) odgovara otvorenoj konformaciji aktivnog mjesta, (b) odgovara djelomično zatvorenoj konformaciji aktivnog mjesta, a (c) odgovara zatvorenoj konformaciji aktivnog mjesta.

4.2.3. Analiza struktura sa supstratima

Kako bi se vidjelo utječu li supstrati na promjene konformacije aktivnih mjesta napravljene su iste analize kao i u prethodnom odjeljku, no u ovom slučaju koristile su se preostale trajektorije. Uočeno je da ligandi formicin A i hipoksantin u potpunosti izlaze iz strukture aktivnog mjesta enzima kod većeg broja simulacija pa se stoga veći fokus stavio na fosfat. Uočeno je da se ioni fosfata također pomiču iz svojih inicijalnih pozicija u kristalnim strukturama, ali uvijek ostaju u džepu aktivnog mjesta.

Prva analiza koja je napravljena kod struktura sa supstratom je analiza RMSD vrijednosti te su na slici 32 izdvojena dva krajnja slučaja, kod kojih pretežito ne dolazi do promjene u vrijednosti RMSD te jedan kod kojeg se ona mijenja. Ako navedena vrijednost varira oko vrijednosti od 0,5 Å znači da ne dolazi do promjene u odnosu na referentnu strukturu te se zaključuje da je H8 α-zavojnica u navedenom lancu nesegmentirana (slika 32a). Ako ta

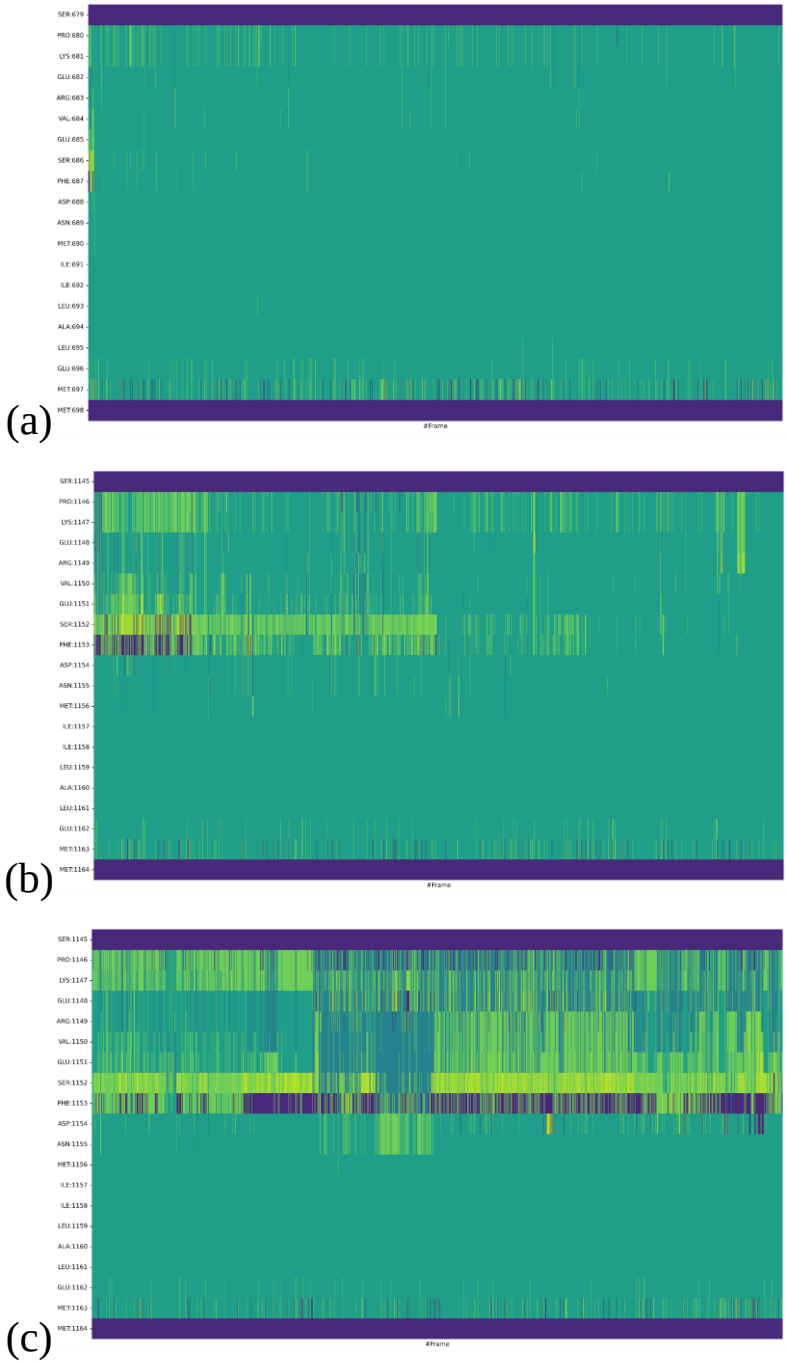
vrijednost varira između vrijednosti 1,5 Å i 2 Å znači da dolazi do promjene u odnosu na referentnu strukturu što može ukazivati da se H8 α-zavojnica segmentirala (slika 32c). Na slici 32b nalazi se primjer gdje H8 α-zavojnica mijenja svoju konformaciju tijekom vremena trajanja simulacije. Ono što se dodatno može utvrditi iz ovakvih grafova je, ako dolazi do promjene u konformaciji, točan trenutak kada se ta konformacijska promjena dešava. Iz b grafa ovisnosti na slici 32 vidi se da dolazi do postupnog otvaranja aktivnog mjesta (prelazak iz segmentirane u nesegmentiranu formu H8 α-zavojnice) gdje se uočava prvi pad RMSD vrijednosti kod oko 60.000 koraka (150 ns) te drugi pad RMSD vrijednosti nešto prije 300.000 koraka (750 ns). RMSD vrijednost, kao što se može vidjeti, je dosta skokovite naravi što uvodi sumnju u samu analizu i iziskuje dodatne analize koje će potvrditi i bolje objasniti konformacijske promjene koje se dešavaju kod HpPNP enzima.



Slika 32. Primjer ovisnosti RMSD vrijednosti o broju koraka trajanja simulacije koje prikazuju primjer (a) otvorene konformacije aktivnog mjesta (nesegmentirana H8 α-zavojnica u C lancu simulirane strukture *5lu0_po4* u odnosu na referentnu strukturu *6f52_apo*), (b) promjene konformacije aktivnog mjesta (segmentacija H8 α-zavojnica u E lancu simulirane strukture *6f4x_po4_fmc* u odnosu na referentnu strukturu *6f52_apo*) i (c) zatvorene konformacije aktivnog mjesta (segmentirana H8 α-zavojnica u E lancu simulirane strukture *6f4x_po4_fmc* u odnosu na referentnu strukturu *6f52_apo*).

Uz izračun promjene RMSD vrijednosti H8 α-zavojnice napravljena je i DSSP analiza. Na slici 33 mogu se vidjeti primjeri različitih konformacija aktivnog mjesta kao što je to slučaj i kod izračuna RMSD vrijednosti (slika 32), odnosno analiza je prikazana za iste lance iz istih trajektorija. Istim redoslijedom mogu se vidjeti primjeri otvorene konformacije aktivnog mjesta trajektorije *5lu0_po4* iz replike 3 (slika 33a), promjene iz zatvorene u otvorenu konformaciju aktivnog mjesta iz trajektorije *6f4x_po4_fmc* iz replike 1 (slika 33b) i zatvorene konformacije

aktivnog mjesta iz trajektorije *6f4x_po4_fmc* iz replike 3 (slika 33c). Pomoću ove metode se također može približno odrediti vremenski interval u kojem se događa potencijalna konformacijska promjena, odnosno u kojem trenutka se događa promjena u sekundarnoj strukturi H8 α-zavojnice. Također se uz vremensku komponentu može odrediti kod kojih se aminokiselina dešava ovakva promjena. Sa slike 33 može se zaključiti kako do promjene najčešće dolazi kod aminokiselina serin Ser220 i fenilalanin Phe221.

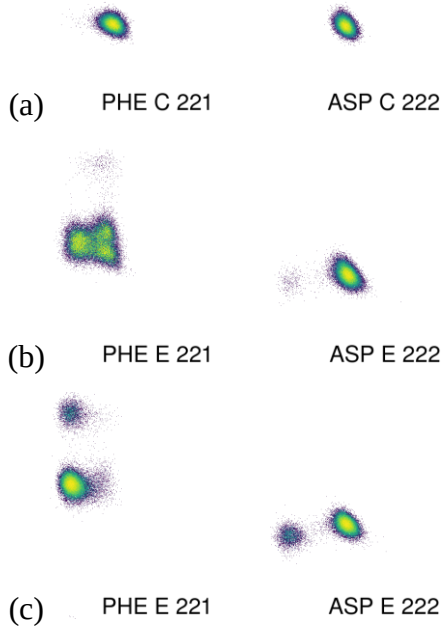


Slika 33. Prikaz rezultata DSSP analize okosnice atoma H8 α-zavojnice simuliranih struktura (a) *5lu0_po4*, (b) *6f4x_po4_fmc* i (c) *6f4x_po4_fmc*. Slika (a) predstavlja primjer otvorene konformacije aktivnog mjesta, slika (b) predstavlja primjer kod kojeg se dešava konformacijska promjena aktivnog mjesta te slika (c) predstavlja primjer zatvorene konformacije aktivnog mjesta. Na svim slikama boje indiciraju različite sekundarne strukture: ljubičasta ne označava nikakvu sekundarnu strukturu, zelena boja predstavlja 3-10 α-zavojnice, α-zavojnice i π (3-14) α-zavojnice, dok žuta boja predstavlja okret.

Metodom MDavocado dijagrama mogu se vidjeti razlike u lokalnom ponašanju kutova okosnice za gornja tri slučaja tijekom konformacijskih promjena. MDavocado dijagrami za aminokiseline koje odgovaraju H8 α-zavojnici za gornja tri slučaja (simulirane strukture (a) *5lu0_po4*, (b) *6f4x_po4_fmc* i (c) *6f4x_po4_fmc*) prikazani su na slici 34. Na toj slici mogu se vidjeti tri primjera: otvorena konformacija aktivnog mjesta (slika 34a), djelomično zatvorena konformacija aktivnog mjesta (slika 34b) i zatvorena konformacija aktivnog mjesta (slika 34c). Kod otvorene konformacije može se vidjeti kako su vrijednosti na MDavocado dijagramu smještene u područje koje odgovara vrijednostima α-zavojnice. Kod zatvorene konformacije aktivnog mjesta može se vidjeti da dolazi do promjene vrijednosti međuovisnosti φ i ψ kuta kod aminokiselina fenilalanin Phe221 i aspartat Asp222 što indicira proces segmentacije α-zavojnice, no ni u jednom slučaju ne dolazi do potpune segmentacije i stabilizacije φ i ψ kutova ovih aminokiselina.

Slika 34. MDavocado dijagrami simuliranih struktura (a) *5lu0_po4*, (b) *6f4x_po4_fmc* i (c) *6f4x_po4_fmc* područja H8 α-zavojnice kod (a) otvorene konformacije aktivnog mjesta, (b) djelomično zatvorene konformacije aktivnog mjesta i (c) zatvorene konformacije aktivnog mjesta. Na (b) i (c) dijelovima vidi se segmentacija zavojnice koja se događa kod aminokiselina fenilalanin Phe221 i aspartat Asp222 uz također njima prethodne aminokiseline Val218 – Ser220.

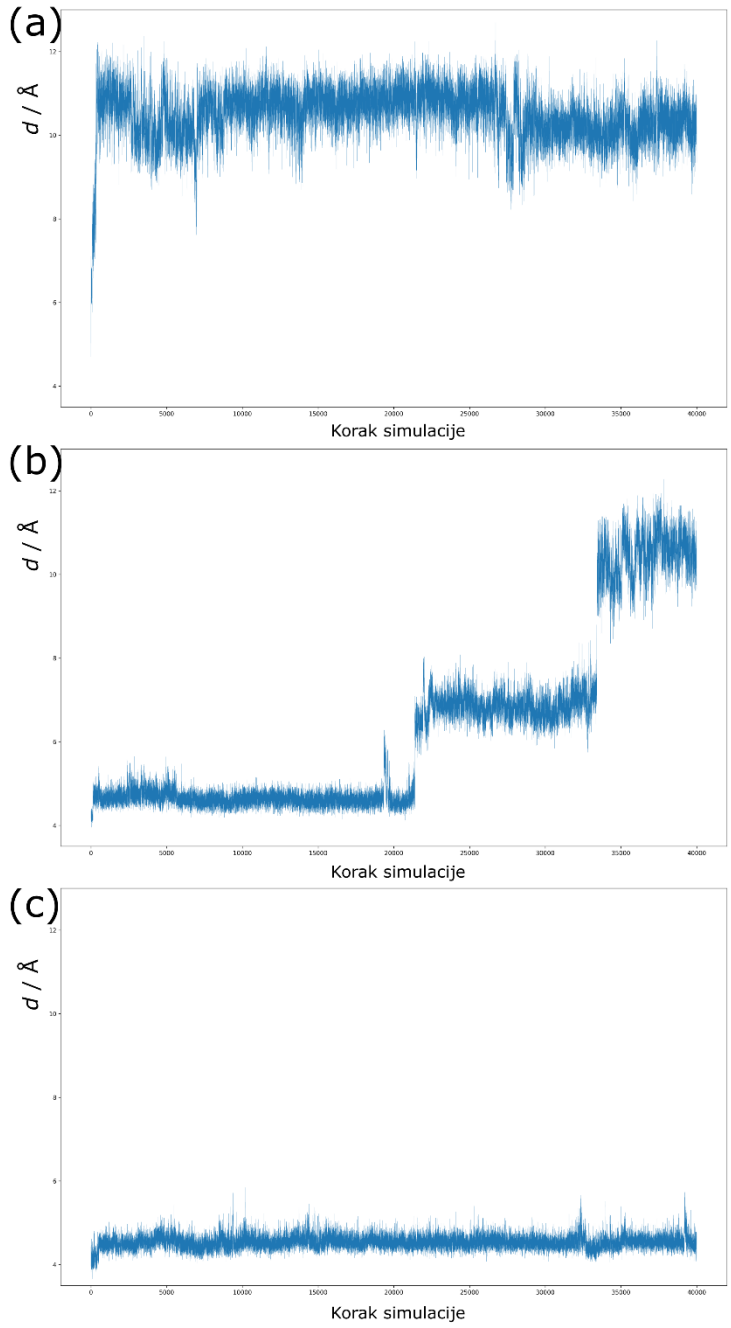
Na slici 34 prikazan samo je jedan (od deset) vremenskih intervala trajanja MD simulacije. Prikazani rezultati analize trajektorije na slici 34 odgovaraju onima na slici 33, dakle (a) na obje slike odgovara *5lu0_po4* trajektoriji replike 3, (b) na obje slike odgovara *6f4x_po4_fmc* trajektoriji replike 1, a (c) na obje slike odgovara *6f4x_po4_fmc* trajektoriji replike 3. Za prikaz ove analize kod otvorene konformacije aktivnog mjesta odabran je drugi (od deset) vremenski interval (na slici 33a vidi se da je α-zavojnica u drugom vremenskom intervalu simulacije u potpunosti u svojoj nesegmentiranoj formi). Kod MDavocado dijagrama djelomično zatvorene konformacije aktivnog mjesta odabran je prvi (od deset) vremenski interval budući da se kod njega uočava djelomična segmentacija α-zavojnice u samo nekoliko vremenskih točaka što odgovara porastu ψ kuta kod aminokiseline Phe221 i smanjenje φ kuta kod aminokiseline Asp222. Također se uočavaju jača gibanja na dijagramima kod susjednih aminokiselina, točnije kod aminokiselina valin Val218, glutamat Glu219 i serin Ser220. Kod zatvorene konformacije aktivnog mjesta odabran je sedmi (od deset) vremenski interval trajanja simulacije zbog uočavanja većeg broja kutova koji odgovaraju potpunoj segmentaciji α-zavojnice što se vidi na aminokiselinama Phe221 i Asp222 (slika 34c). Također se uočavaju promjene na tri njima prethodne aminokiseline, Val218, Glu219 i Ser220, što odgovara promjenama i kod djelomično zatvorene konformacije aktivnog mjesta (slika 34b). U ostalim vremenskim intervala simulacije ne dolazi često do ovakve segmentacije α-zavojnice, odnosno takva promjena nije stalna kroz cijelo trajanje MD simulacije. Na slici 35 uvećani su MDavocado dijagrami za aminokiseline fenilalanin Phe221 i aspartat Asp222 kako bi se bolje vidjela navedena promjena: (a) odgovara otvorenoj konformaciji aktivnog mjesta, (b) odgovara djelomičnom zatvaranju aktivnog mjesta, a (c) odgovara zatvorenoj konformaciji aktivnog mjesta.



Slika 35. Uvećani prikaz MDavocado dijagrama simuliranih struktura (a) *5lu0_po4,* (b) *6f4x_po4_fmc* i (c) *6f4x_po4_fmc* za aminokiseline fenilalanin Phe221 i aspartat Asp222 kod kojih se uočava konformacijska promjena zatvaranja aktivnog mjesta. (a) odgovara otvorenoj konformaciji aktivnog mjesta, (b) odgovara djelomično zatvorenoj konformaciji aktivnog mjesta, a (c) odgovara zatvorenoj konformaciji aktivnog mjesta.

DSSP analiza kao analiza sekundarne strukture aminokiselina H8 α-zavojnice pokazala se odličnim odabirom za analiziranje ovakvog tipa podataka i promjena koje se događaju kod HpPNP enzima. Iz analize su proizašli rezultati koji potvrđuju promjenu konformacijskog stanja aktivnog mjesta što se lako iščitava pomoću DSSP grafova. Također se se pomoću ove analize uspješno utvrdile aminokiseline kod kojih se događa navedena konformacijska promjena. Također, kao nastavak ove analize, napravljena je i MDavocado analiza proteina koja se kao i DSSP analiza odnosi samo na H8 α-zavojnicu. Ova analiza je potvrdila konformacijske promjene aktivnog mjesta. Dodatno su ovom analizom utvrđene aminokiseline felinalanin Phe221 i aspartat Asp222 kod kojih dolazi do segmentacije zavojnice i zatvaranja odnosno otvaranja aktivnog mjesta. Navedenim analizama se također može utvrditi i u kom trenutku u simulaciji dolazi do promjene konformacijske promjene, no ostaje neobjašnjeno zašto se kod simulacija bez supstrata javljaju slični, ako ne i isti slučajevi promjena konformacija.

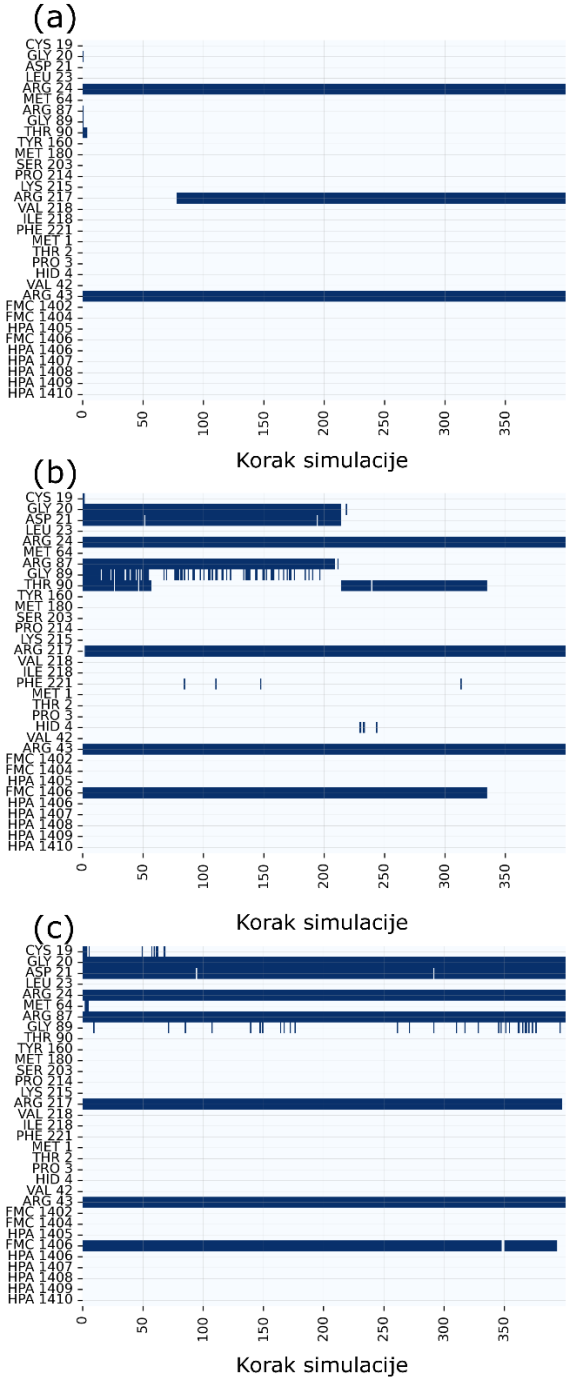
Usporedno s analizom segmentacije H8 α-zavojnice napravljena je i analiza proučavanja iona fosfata te njegove koordinacije. Ovdje izdvojene analize na određenim trajektorijama odgovaraju istim trajektorijama na kojima je napravljena analiza DSSP (slika 33) i MDavocado (slika 34) analiza. Trodimenzijski raspored aktivnog mjesta HpPNP-a poznat je iz kristalnih struktura te se iz njega mogu iščitati aminokiseline koje koordiniraju fosfat, a koje dušikovu bazu. Jedna od ključnih aminokiselina kod same koordinacije, ali i kod reakcijskog mehanizma je glicin Gly20. Zbog lokacije ovog glicina koja se nalazi duboko u džepu aktivnog mjesta, napravljen izračun udaljenosti fosfata od glicina. Točnije, promatrala se udaljenost centra mase glicina i centra mase fosfata. Jednostavnije, ako se u nekom aktivnom mjestu nalazi fosfat, izračunata je udaljenost fosfata od glicina koji se nalazi u tom istom aktivnom mjestu. Na slici 36 vidi se ovisnost navedene udaljenosti o tijeku trajanja simulacije za (a) podjedinicu s otvorenom konformacijom aktivnog mjesta, (b) podjedinicu s djelomično otvorenom konformacijom aktivnog mjesta i (c) podjedinicu kod koje je konformacija aktivnog mjesta većinu vremena trajanja simulacije zatvorena. Na navedenoj slici može se vidjeti da se kod otvorene konformacije aktivnog mjesta (a) fosfat u samom početku nalazi blizu glicina (5 - 6 Å) te on odmah izlazi iz svoje primarne lokacije u aktivnom mjestu na lokaciju koja je udaljenja 10 - 11 Å od glicina. Kod simulacije kod koje dolazi do promjene konformacije aktivnog mjesta (b), fosfat se smješta u tri zasebne lokacije u samom aktivnom mjestu koje su na različitim udaljenostima od glicina (prva lokacija na ~4,5 Å, druga lokacija na ~7 Å i treća lokacija na 10 - 11 Å). Kod zatvorene konformacije aktivnog mjesta (c) fosfat se zadržava cijelo vrijeme na jednoj lokaciji koja se nalazi na udaljenosti od ~4,5 Å od glicina što znači da on ostaje u svojoj primarnoj lokaciji u aktivnom mjestu.



Slika 36. Ovisnost udaljenosti centara masa glicina Gly20 i pripadajućeg im fosfata o tijeku trajanja simulacije simuliranih struktura (a) *5lu0_po4*, (b) *6f4x_po4_fmc* i (c) *6f4x_po4_fmc* za (a) podjedinicu u otvorenoj konformaciji aktivnog mjesta, (b) podjedinicu s djelomično otvorenom konformacijom aktivnog mjesta i (c) podjedinicu kod koje aktivno mjesto većinu trajanja simulacije u zatvorenoj konformaciji.

Iduća analiza koja uključuje ligand je analiza interakcija fosfata s aminokiselinama tijekom boravka u aktivnom mjestu, odnosno izračunati su svi kontakti koje fosfat ostvaruje s

aminokiselinama. Navedena analiza također je napravljena na istim trajektorijama kod kojih su već napravljene i gore navedene analize. Fosfat u početku simulacije kod koje je aktivno mjesto u otvorenoj konformaciji (slika 37a) ima kontakt s glicinom Gly20, argininom Arg87, glicinom Gly89 i treoninom Thr90 te ih odmah gubi. Zadržavaju se kontakti s argininom Arg24 i argininom Arg43 iz susjednog monomera enzima kroz cijelo vrijeme trajanja trajektorije dok se kontakt s argininom Arg217 ostvaruje nešto prije četvrtine simulacije. Kod simulacije kod koje dolazi do promjene konformacije aktivnog mjesta iz zatvorenog u otvoreno (slika 37b) mogu se vidjeti različiti kontakti s obzirom na poziciju fosfata u aktivnom mjestu. Kad se fosfat nalazi na ~4,5 Å stalne i konstantne kontakte ostvaruje s glicinom Gly20, aspartatom Asp21 i argininom Arg87 te povremene kontakte s glicinom Gly89 i treoninom Thr90. Nakon malo više od polovice trajanja simulacije, fosfat se pomakne na drugu svoju poziciju udaljenu od glicina Gly20 za ~7 Å te na toj poziciji gubi sve dosad ostvarene kontakte osim onog s treoninom Thr90. Pri kraju trajanja simulacije, fosfat se pomiče na svoju treću lokaciju prilikom čega gubi i kontakt s treoninom Thr90, ali i s formicinom A iz tog aktivnog mjesta. Tijekom cijelog trajanja simulacije, fosfat ostvaruje kontakt s argininom Arg43, argininom Arg217 i argininom iz susjedne podjedinice Arg43. U trećem slučaju, kada se u podjedinici javlja zatvorena konformacija aktivnog mjesta kroz cijelo vrijeme trajanja simulacije (slika 37c), također se pojavljuju i kontakti koji su ostvareni u cijeloj trajektoriji. Ti kontakti su kontakti s glicinom Gly20, aspartatom Asp21, argininom Arg24, argininom Arg87, argininom Arg217, argininom iz susjedne podjedinice Arg43 i s formicinom A iz tog aktivnog mjesta. Povremeno se pojavljuju kontakti s glicinom Gly89 te na početku s cisteinom Cys19 i metioninom Met64. Navedena analiza potvrđuje aminokiseline koje ostvaruju interakciju s fosfatom na različitim lokacijama u koje se on smješta u aktivnom mjestu. Ono što je najvažnije, je lokacija koja se nalazi najbliže glicinu Gly20, budući da je to mjesto ono iz kojeg fosfat može djelovati fosforolitički odnosno ono u koje se fosfat smješta tijekom odvijanja reakcije samog enzima (reakcijska shema prikazana na slici 4 u Literaturnom pregledu). Aminokiseline koje su važne za njegovo smještanje u aktivno mjesto su glicin Gly20, aspartat Asp21 i arginin Arg87. Također u tome sudjeluje i glicin Gly89 te arginin Arg24, arginin Arg217 i arginin iz susjednog monomera enzima Arg43. Navedena tri arginina koordiniraju fosfat kada se on nalazi na sva tri svoje lokacije pa se može zaključiti da oni zajedno s fosfatom mijenjaju svoju lokaciju u aktivnom mjestu tijekom trajanja simulacija.



Slika 37. Prikaz aminokiselina koje ostvaruju interakciju s fosfatom kroz vrijeme trajanja simulacije za simulirane strukture (a) *5lu0_po4*, (b) *6f4x_po4_fmc* i (c) *6f4x_po4_fmc*. Plavo su prikazana razdoblja u kojima pojedina aminokiselina ima kontakt s fosfatnim ionom. Slika (a) prikazuje podjedinicu enzima koja ima otvorenu konformaciju aktivnog mjesta, (b) prikazuje podjedinicu s promjenom konformacije iz zatvorenu u otvorenu te (c) prikazuje podjedinicu kod koje je aktivno mjesto u zatvorenoj konformaciji kroz cijelo vrijeme trajanja simulacije.

Analiza aktivnog mjesta pokazala se izuzetno zanimljivom zbog svog poklapanja s dosadašnjim analizama. Fosfat u aktivnom mjestu pokazao se od iznimne važnosti zbog svog utjecaja na njegovo konformacijsko stanje što se potvrdilo izračunom udaljenosti fosfata od glicina Gly20. Povezivanjem analize izračuna udaljenosti fosfata, DSSP analiza i analiza MDavocado dijagrama pronađene su uzročno-posljedične veze u pomicanju fosfata u aktivnom mjestu i promjena njegovog konformacijskog stanja. Ako se fosfat nalazi u svom glavnom (fosforilitičkom) mjestu koje se nalazi najdublje u enzimu (najbliže Gly20) aktivno mjesto se nalazi u zatvorenoj konformaciji. U slučaju kada se fosfat barem malo udalji od Gly20, dolazi do nestabilnosti u konformaciji zatvorenog aktivnog mjesta, odnosno dolazi do njegova otvaranja. Dodatno su potvrđena različita mjesta interakcije u aktivnom mjestu u koja se smješta fosfat te su potvrđene aminokiseline s kojima ostvaruje kontakt tijekom tih boravaka.

4.3. Korelacije u gibanjima aminokiselina

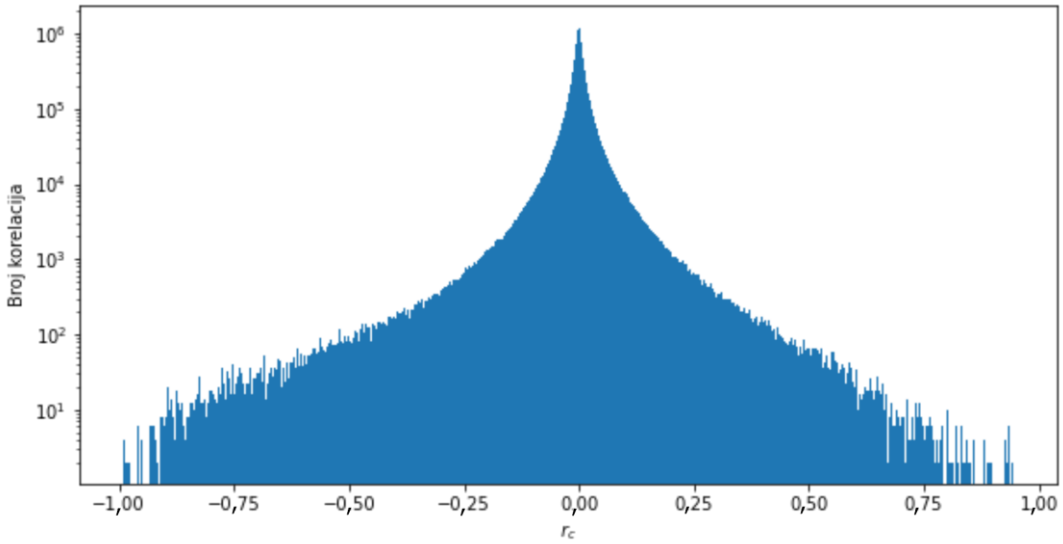
Kako je već rečeno, glavni interes istraživanja u ovom radu je pronalaženje načina kako promjene položaja ili konformacije pojedine aminokiseline utječe na ostale aminokiseline u proteinu tijekom MD simulacija. Pretpostavka je da postoji određeni utjecaj promjena konformacije neke aminokiseline na one u njenoj neposrednoj blizini te također djelovanje koje se može prenositi i na neke udaljene aminokiseline kroz protein alosteričkim putovima. Za praćenje međusobnog utjecaja gibanja pojedinih aminokiselina prirodno je pratiti kako su njihova gibanja korelirana u vremenu. Rečeno je da je pojedini torzijski kut u trajektoriji predstavljen jednostavno kao vremenski slijed vrijednosti. U svim simulacijama to je obično vremenski slijed od oko 400.000 točaka, jer su trajektorije sadržavale toliko vremenskih koraka za svaki kut. Ukupno gledano to se može promatrati kao tablica vrijednosti gdje stupci predstavljaju kutove (parni brojevi stupaca predstavljaju kut φ, neparni brojevi kut ψ), dok retci predstavljaju broj koraka u MD simulaciji (slika 38). Svaki lanac HpPNP-a ima 233 aminokiseline i njih ima šest, dobivenom broju se oduzima dva (jer prvi i zadnji kut nisu definirani) te se množi s dva i dobiva se ukupan broj stupaca od 2792.

	0	1	2	3	4	5	6	7	8	9	...	2782	2783
0	-88.504910	149.427615	-45.448280	-31.754288	-93.755635	6.939630	-144.293941	122.032357	-77.375456	-18.884050	...	-61.030922	-49.517219
1	-56.667459	163.372281	-64.406816	-5.534525	-124.664968	-1.064152	-139.534503	120.128182	-75.818870	-34.602927	...	-55.261020	-45.849335
2	-80.454549	169.681512	-59.991390	-19.662685	-101.239606	0.410726	-155.687192	122.568511	-75.152084	-34.698919	...	-64.331584	-54.352792
3	-75.437152	166.609182	-60.252445	-22.192870	-103.760326	-0.816714	-138.602037	117.429539	-68.933347	-35.352185	...	-73.987084	-55.011225
4	-60.443230	153.477652	-77.768531	5.263440	-99.742512	-14.350381	-141.110443	121.027219	-70.736054	-15.217790	...	-61.269148	-31.070922
...	...	...	...	...	...	...	...	...	...	...	...	...	...
399995	-62.336561	149.362298	-66.296173	-26.481905	-87.349151	-26.429451	-139.544403	149.130030	-113.769277	-0.208671	...	-60.668850	-44.166496
399996	-59.890229	139.829326	-52.467241	-27.680054	-83.452686	-30.926797	-134.552276	134.820972	-106.195231	5.455429	...	-49.330857	-44.897809
399997	-46.569288	148.625733	-70.502616	-20.439895	-78.508644	-20.415724	-149.767665	133.532573	-126.580411	20.805950	...	-70.771867	-42.251431
399998	-57.786678	143.095584	-60.305619	-38.832738	-60.387826	-35.922851	-145.418067	134.028227	-107.749809	4.789811	...	-45.773107	-46.726752
399999	-54.010861	147.834658	-60.697371	-20.412023	-88.565128	-29.639222	-149.998348	134.945733	-109.503752	16.832216	...	-69.543274	-37.858975

400000 rows × 2792 columns

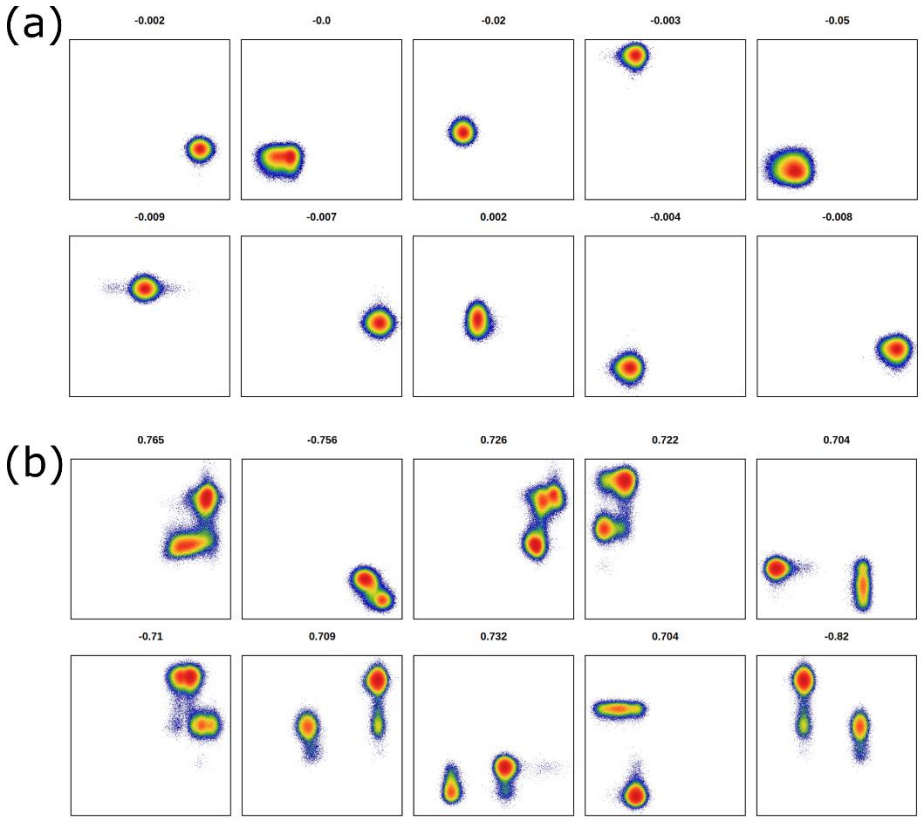
Slika 38. Svaka MD simulacija rezultira matricom koja je spremljena u obliku Pandas tablice koja sadrži po dva kuta za svaku aminokiselinu. Sveukupno ovakva matrica sadrži približno $1{,}12 \cdot 10^9$ elemenata i zauzima oko 8,8 GB prostora na računalu. Vrijednosti u tablici predstavljaju torzijske Ramachandranove kutove (parni brojevi stupaca predstavljaju kut φ, neparni brojevi kut ψ) izražene u stupnjevima, a zbog velike količine podataka i izvorne analize brojevi nisu prikazani standardnim hrvatskim jezikom.

Nakon izračuna korelacija za svaku trajektoriju može se uočiti nekoliko zanimljivih značajki. Prvo je da je velika većina korelacija grupirana oko nule dok je samo manji dio raspoređen ispod $-0{,}5$ i iznad 0,5 te taj udio korelacija obično iznosi oko 0,05 % (slika 39). To je očekivan rezultat jer zapravo korelacije izračunavamo između jako dugih vremenskih sljedova koji imaju razmjerno visoku razinu šuma što zapravo znači da vrijednost Ramachandranovih kutova nikad nije uvijek ista već da ona varira oko neke srednje vrijednosti. Potrebno je naglasiti da je ordinata na logaritamskoj skali što daje dodatnu težinu.



Slika 39. Tipičan izgled distribucija korelacija na primjeru jedne od simuliranih trajektorija. Na ordinati ovog histograma je dat broj korelacija a na apscisi je cirkularni korelacijski koeficijent r_c.

Ovakva distribucija korelacija dodatno potkrepljuje činjenicu da velika većina aminokiselina provodi cijelo vrijeme trajektorije stacionarno oko samo jedne točke uz veće ili manje raspršenje oko te točke (slika 26). Da bismo imali značajnu korelaciju (recimo veću od 0,5 po apsolutnoj vrijednosti) potrebno je da i jedna i druga aminokiselina kod kojih se računa korelacijski koeficijent imaju barem dva odijeljena stanja između kojih dolazi do promjene barem jednog od Ramachandranovih kutova, φ ili ψ (slika 40). Na svakoj od manjih slika prikazano je 400.000 točaka programom Datashader te su na slici a) prikazane slabo korelirane aminokiseline gdje se može uočiti da su one približno sferične i postoji samo jedna domena. Kao što je već istaknuto prije, takve korelacije su u izrazitoj većini. Na drugom dijeli slike b) prikazane su visoko korelirane aminokiseline (kojih je zapravo jako malo) te se uočava da postoje najmanje dvije odvojene domene i to u oba kuta, koje ukazuju na korelirane skokove između njih.

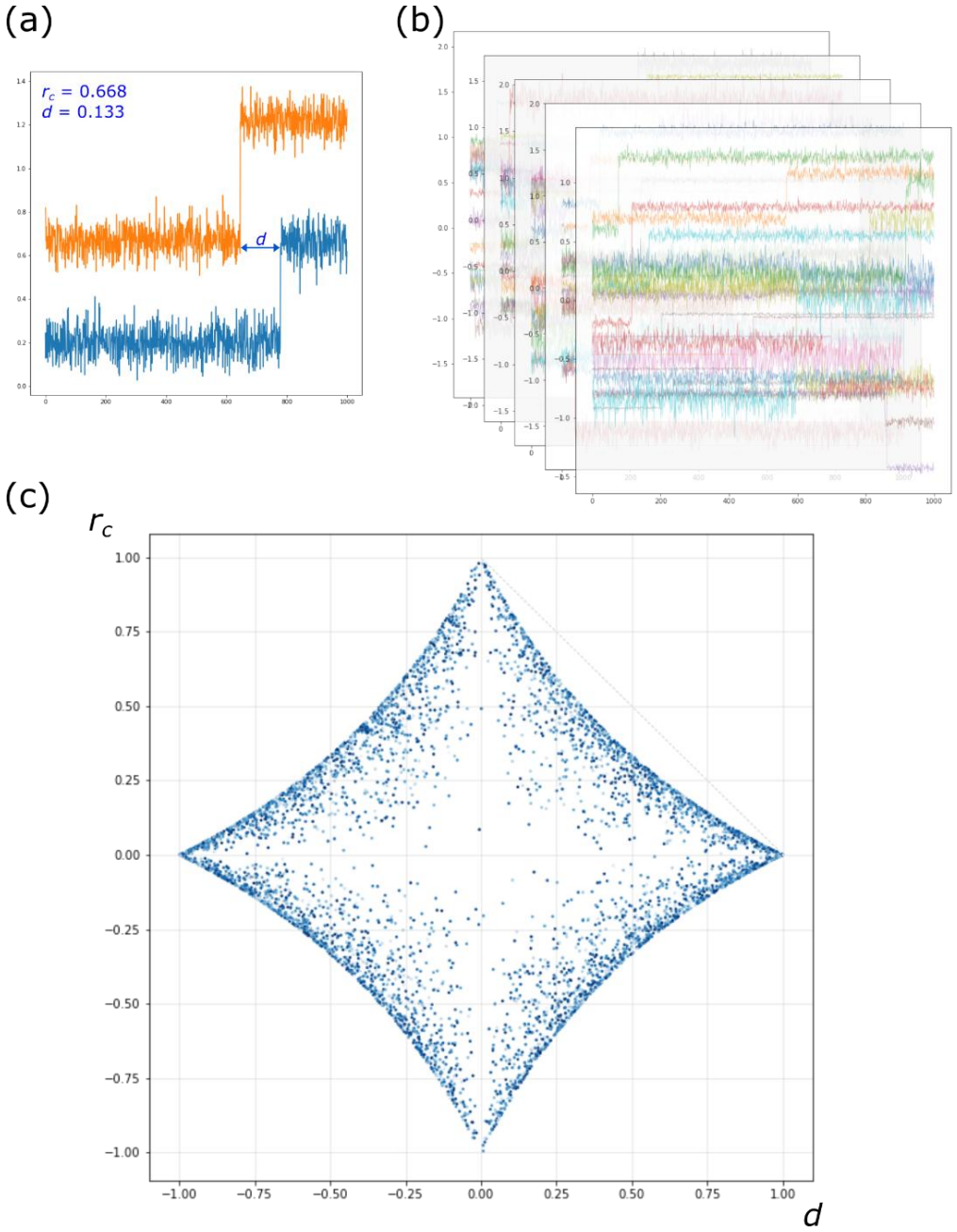


Slika 40. Prikaz međusobnih ovisnosti nasumično izabranih kutova φ ili ψ za jednu od trajektorija. Iznad svake male slike nalazi se izračunata vrijednost cirkularnog korelacijskog koeficijenta. Na svakoj manjoj slici prikazano je 400.000 točaka programom Datashader. (a) Slabo korelirane aminokiseline. (b) Visoko korelirane aminokiseline.

Ukupno gledano može se zaključiti da znatne korelacije postoje samo između kutova koji pokazuju nekoliko (najmanje dvije, a to je i daleko najčešće) odvojenih domena u φ-ψ ravnini. To znači da će one aminokiseline koje pokazuju skokove biti značajno korelirane samo sa drugim aminokiselinama jednakog ponašanja, što čini korelacije u ovom kontekstu vrlo pogodnim načinom praćenja međusobnog utjecaja aminokiselina jedne na drugu. Ostaje samo još pitanje kako korelacije među jako koreliranim aminokiselinama ovise o vremenskoj udaljenosti njihovih točaka promjene, dakle onih u kojima dođe do skoka. Za pretpostaviti je da će korelacija biti to veća što su točke promjene međusobno bliže, te da će najveća korelacija biti između onih kutova čije se promjene događaju praktički istovremeno. Kao dodatnu potvrdu gore izrečenog napravljeni su sintetički podaci koji oponašaju kutove koji imaju samo jedan skok u svoj vrijednosti. Načinjen je skup od 1000 signala koji su dobiveni dodavanjem šuma na običan stepenasti signal (engl. *step-function*) sa slučajnim položajem skoka i iznosom šuma

te je među njima napravljena korelacija (slika 41). U ovom skupu podataka definirane su dvije vrijednosti, a to su vrijednosti cirkularnog korelacijskog koeficijenta r_c i vremenske udaljenosti događaja skoka u vrijednosti kuta između dva seta podataka d što se može vidjeti na slici 41a. Ovako definirana vrijednost d je normalizirana te iznosi 0 ako su skokovi u dva seta podataka blizu i 1 ako su oni jako udaljeni. Kako bi se dobio što veći set podataka, generirani su različiti primjeri u kojima se variraju dva navedena parametra (r_c i d) što se može vidjeti na slici 41b. Na slici 41c nalazi se prikaz ovisnosti korelacije o udaljenosti između točaka promjene čime je pokazano da je korelacija najveća onda kada je udaljenost najmanja. Zanimljivo da su točke korelacija ograničene unutar zakrivljenog područja koje odstupa od linearnog pada korelacija s udaljenošću (crtkana linija u gornjem desnom kvadrantu). Vidljiva je i aglomeracija točaka prema toj krivulji koja je praktički neovisna o nivou šuma što je prikazano kao obojanost točaka u nijansama plave boje.

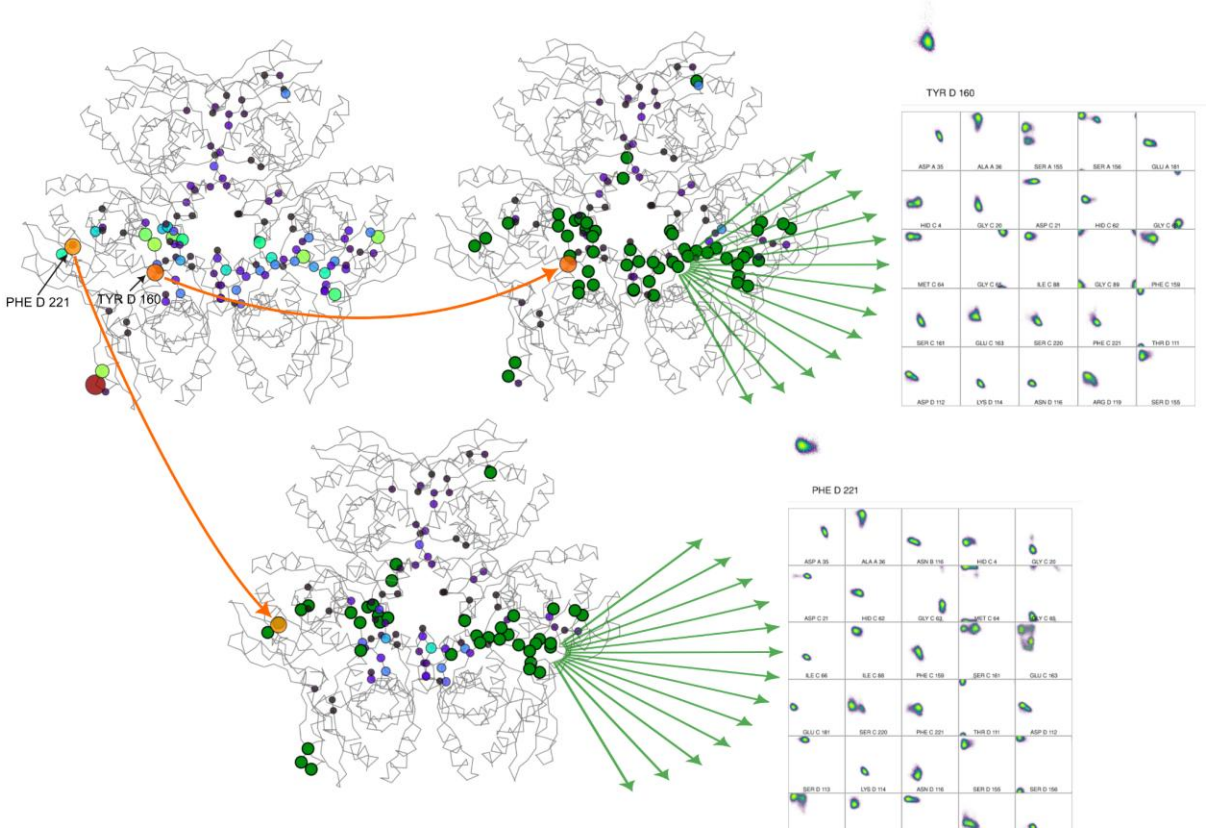
Provedene simulacije zato pružaju dodatne dokaze koji potvrđuju tezu o povezanosti značajnih korelacija između nekih aminokiselina koje imaju skokovite promjene njihovih φ-ψ kutova. Osim što jasno pokazuju uzročno-posljedične veze između ovih promjena, rezultati ističu da jačina tih korelacija ovisi o vremenskoj udaljenosti između pojedinih skokova. Drugim riječima, promjene koje se događaju u kraćim vremenskim intervalima pokazuju izraženiju međusobnu povezanost iz čega se mogu razlučiti ne samo važni aminokiselinski parovi već i dijelovi proteina koji mogu utjecati na alosteričke putove ili konformacijske promjene.



Slika 41. Izračun korelacija na sintetičkom setu od 1000 simuliranih signala "kutova" s jednom točkom promjene (koja je jedan skok s nasumičnim iznosom i smjerom). (a) Primjer jednog para signala na kojem je izražena vrijednost korelacije i udaljenosti skokova u dva seta podataka. Umjetno generirani podaci predstavljaju vrijeme simulacije na ordinati i vrijednost kutova na apscisi. (b) Mnoštva nasumično generiranih signala. Umjetno generirani podaci predstavljaju vrijeme simulacije na ordinati i vrijednost kutova na apscisi. (c) Prikaz ovisnosti cirkularne korelacije (r_c) o udaljenosti (d) između točaka promjene.

4.4. Prostorni raspored korelacija

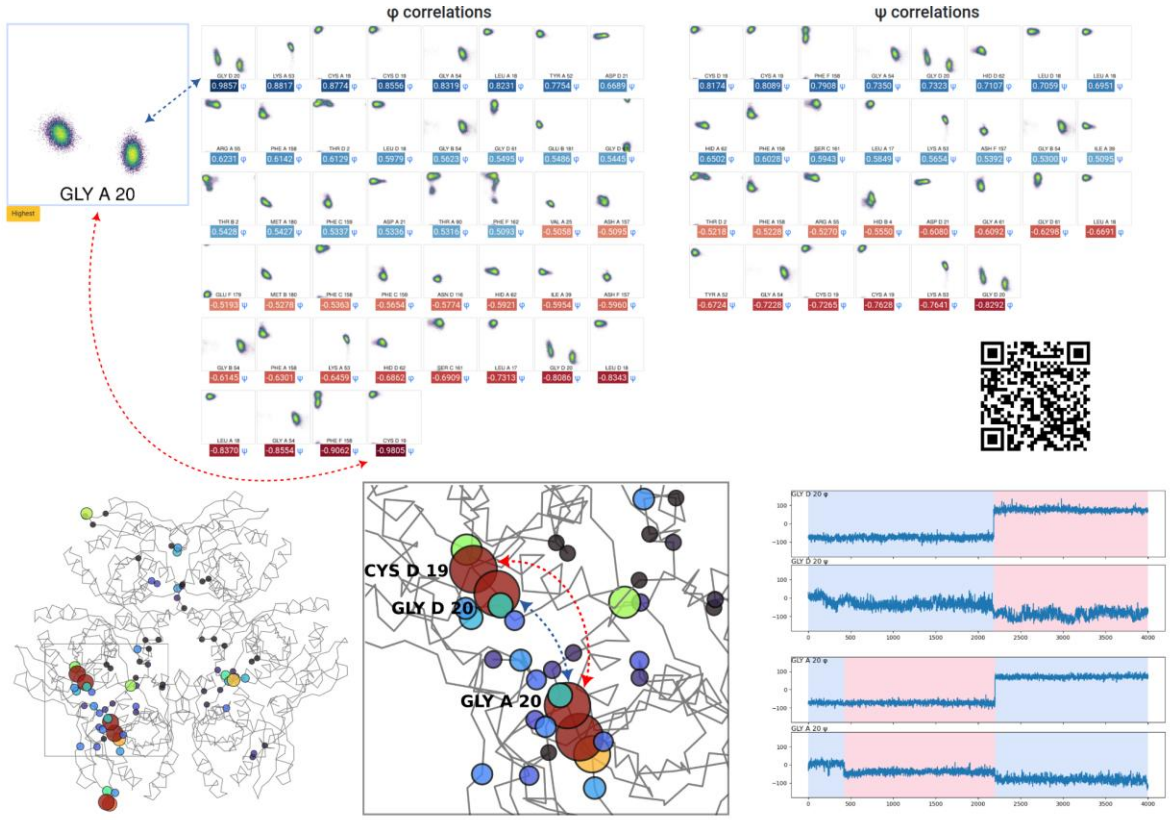
Dosad navedene činjenice nisu ni na koji način bile povezane s prostornim raspodjelom aminokiselina. Jedna od pretpostavki je da korelacije imaju određenu ovisnost o blizini drugih aminokiselina zbog lokalnog prijenosa promjena s jedne na drugu, čisto mehaničkim prijenosom informacija. Za prostorni prikaz aminokiselina i njihovih korelacija izrađena je posebna interaktivna aplikacija bazirana na mrežnim tehnologijama i JavaScript biblioteci D3.js. Koordinate svih aminokiselina tj. njihovih C_α atoma projicirane su na 2D ravninu koja prolazi središtem enzima HpPNP što daje minimalno prekrivanje i optimalnu projekciju. Aminokiseline su onda povezane linijama sa susjednima u primarnoj sekvenci te se dobro uočavaju obrisi sekundarnih struktura u svih šest lanaca. Na takav prikaz dodane su sve korelirane aminokiseline koje imaju vrijednost korelacije iznad apsolutne vrijednosti 0,5 s barem nekom drugom aminokiselinom. U interaktivnoj inačici u internetskom pregledniku se prelaskom preko bilo koje od koreliranih aminokiselina prikažu MDavocado dijagrami za sve korelirane aminokiseline. Pritom su veličinom i bojom istaknute one aminokiseline koje su korelirane s velikim brojem drugih aminokiselina, tj. njihova veličina (što veći krug, to veća više korelacija) i boja (spektar koji ide od crne (malo korelacija) do crvene (puno korelacija)) prikazuju njihovu važnost u korelacijskom smislu (https://alokomp.irb.hr/md/corrd3_all/1458, slika 42).



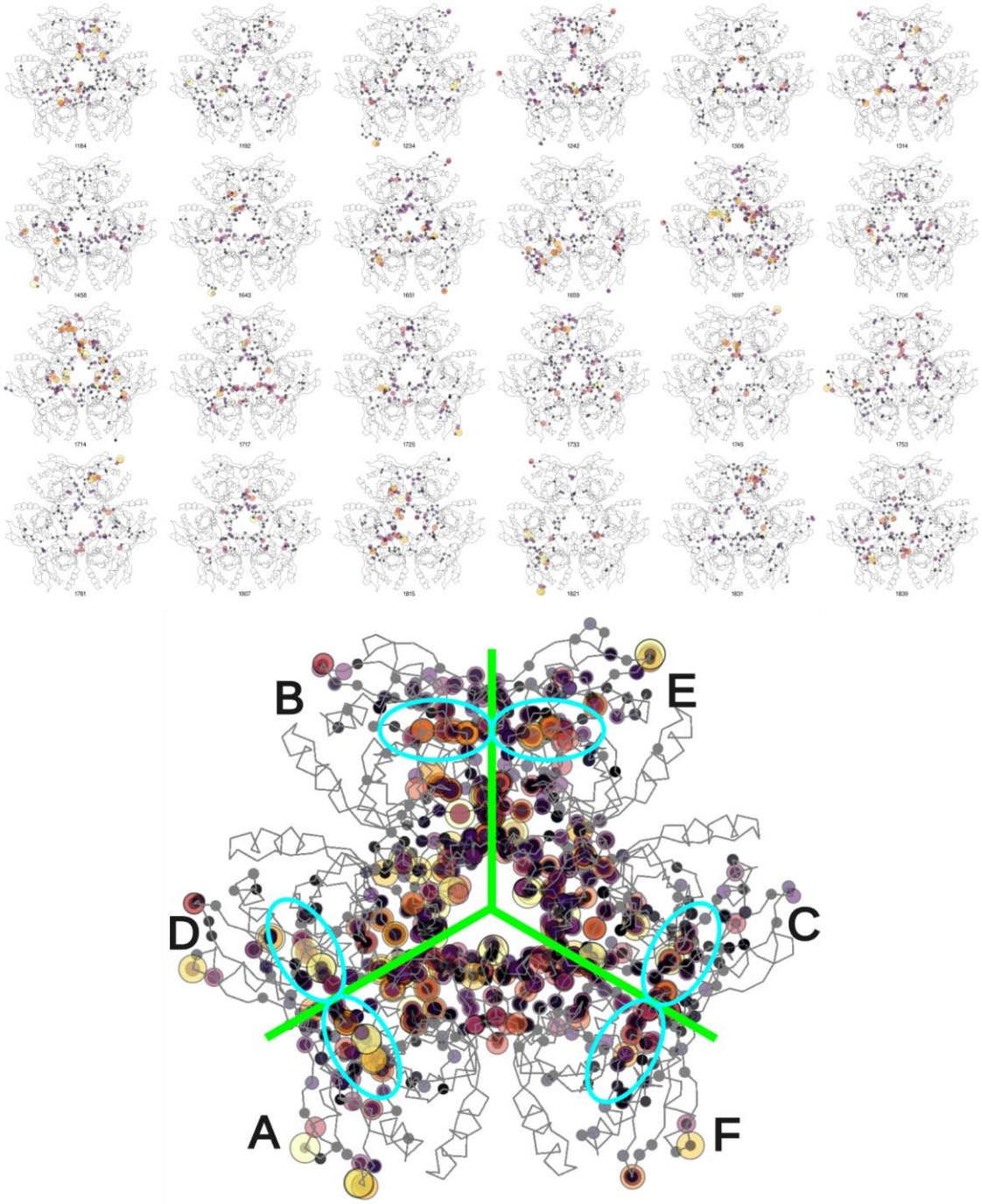
Slika 42. Shematski prikaz funkcionalnosti dostupne na mrežnoj stranici projekta (https://alokomp.irb.hr/md/corrd3_all/1458) za prikaz prostornih ovisnosti korelacija. Korelirane aminokiseline su prikazane većim ili manjim krugovima, ovisno o broju drugih aminokiselina koje su s njome korelirane. Prelaskom preko položaja te aminokiseline (npr. Phe221 iz lanca D ili Tyr168 iz lanca D u gornjim primjerima) postanu vidljive sve takve aminokiseline (zeleni krugovi), a na desnoj strani se istovremeno pojave njihovi MDavocado dijagrami (simbolizirano zelenim strelicama). Ovakav prikaz omogućava simultano promatranje prostornog rasporeda i dinamike aminokiselina. U slučaju jako koreliranih aminokiselina vidljive su njihove sinkronizirane promjene.

Iz ovakvih prostornih prikaza upadljivo je nekoliko stvari. U većini proučavanih trajektorija postoji nekoliko aminokiselina koje su dominantne, tj. korelirane s većinom ostalih. One se na neki način ponašaju kao "okidači" promjene i pritom zadaju ritam ostalima, što se osobito dobro vidi na njihovim MDavocado dijagramima (slika 43). Druga važna značajka je da korelirane aminokiseline nisu ravnomjerno raspoređene unutar proteinske strukture. Iako je prisutna određena razina stohastičnosti i varijabilnosti u raspodjeli, analizirane trajektorije pokazuju da su korelirane aminokiseline uglavnom koncentrirane u središnjem dijelu proteina, na sučeljima

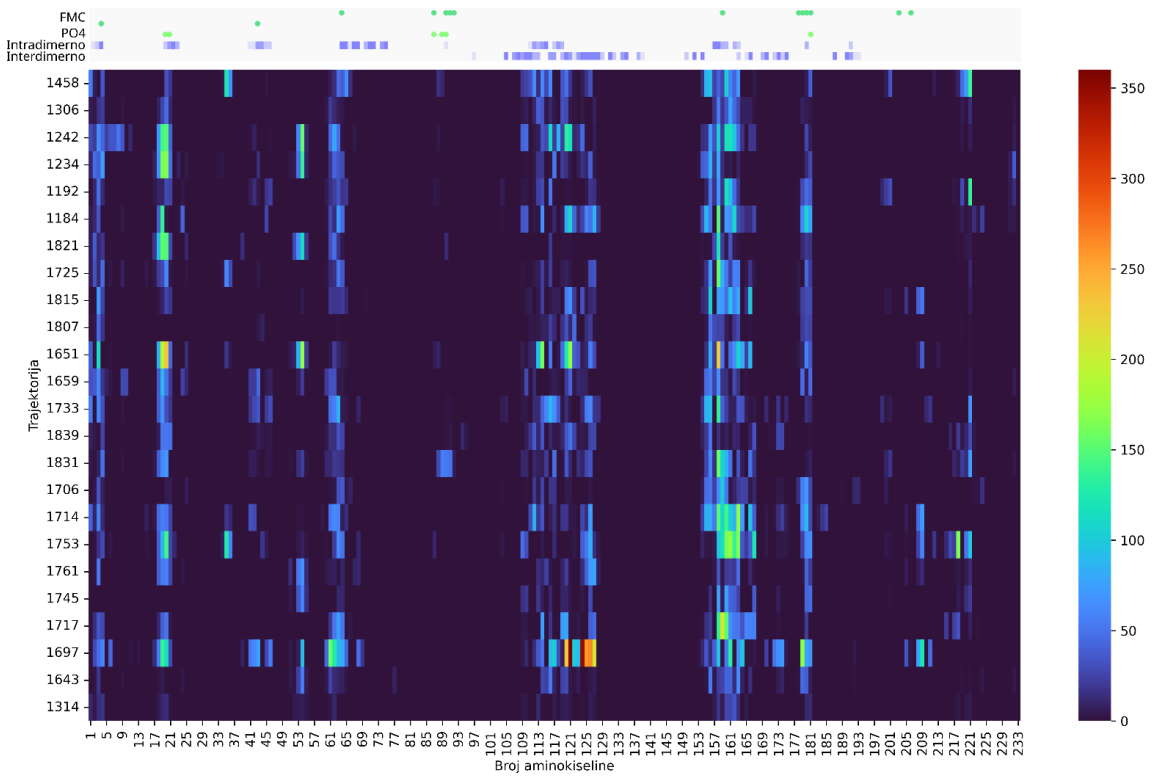
između dimernih jedinica te u udubinama koje vode prema aktivnim mjestima enzima (slika 44). Kao još jedan način prikaza napravljena je toplinska karta (engl. *heatmap*) prosječnog broja korelacija kod svake aminokiseline u svim simuliranim trajektorijama gdje boja prikazuje količinu korelacija. Crveno je označen veliki broj prosječnih korelacija, a plavo mali broj korelacija. Kako bi bilo jasnije koja regija u sekvencijalnom prikazu aminokiselina odgovara kojoj strukturnoj regiji, na gornjoj strani dijagrama dodana je anotacije ostvarenih kontakata iz kristalografske strukture PDB koda 6f4x. Zelenim kuglica označene su aminokiseline koje koordiniraju ligande: prva dva retka odgovaraju koordinaciji formicina A, a treći redak odgovara koordinaciji fosfata. Ljubičastim linijama označene su regije koje ostvaruju kontaktne sa susjednim monomernim lancima pa tako prvi redak označava intradimerno sučelje, a drugi označava interdimerno sučelje (slika 45). Iz slike se jasno vidi da ova vizualizacija odgovara prethodnoj te su istaknute regije kod kojih se javlja najviše koreliranih aminokiselina.



Slika 43. Primjer interaktivnog prikaza za jednu jako koreliranu aminokiselinu (Gly20 iz lanca A koja je dio aktivnog mjesta) a koji je dostupan za sve aminokiseline. Gore lijevo je MDavocado dijagram aminokiseline u pitanju a desno su takvi dijagrami njoj koreliranih aminokiselina s prikazom iznosa korelacije. Klikom na bilo koju od njih dobiva se pripadni prikaz za tu aminokiselinu. Dolje je uvećano područje A-D dimera i njihovih aktivnih mjesta. Crtkano su prikazane dvije najjače korelacije a dolje desno je vremenski prikaz razvoja kutova za Gly20 iz A i D lanaca gdje se jasno vidi razlog tako visoke korelacije. Primjetno je jako slično ponašanje tih dvaju aminokiselina iako su udaljene u proteinu, tj. nisu u neposrednom kontaktu.



Slika 44. Prostorni raspored koreliranih aminokiselina u svim trajektorijama, s brojem simulacije naznačenim ispod svake projekcije. Vidljiva je koncentracija takvih aminokiselina pretežito oko središnje šupljine heksamera i na dodirnim površinama pojedinih monomera, pogotovu onih između dvaju monomera koji čine dimere (AD, BE i CF, zeleno) te prema aktivnim mjestima (tirkizno). Desni dio slike dobiven je preklapanjem svih 24 slika s lijeve strane, odnosno predstavlja prosjek svih MD trajektorija.



Slika 45. Toplinska karta prosječnog broja korelacija svake aminokiseline u svim simuliranim trajektorijama (crveno je označen veliki broj prosječnih korelacija, a plavo mali broj korelacija). Na gornjoj strani dijagrama dodana je anotacije ostvarenih kontakata iz kristalografske strukture PDB koda 6f4x. Zelenim kuglicama označene su aminokiseline koje ostvaruju kontakt sa supstratima: prva dva retka odgovaraju interakcijama s formicinom A, a treći redak odgovara interakcijama s fosfatom. Ljubičastim linijama označene su regije koje ostvaruju kontaktne sa susjednim monomernim lancima pa tako prvi redak označava intradimerno sučelje, a drugi označava interdimerno sučelje.

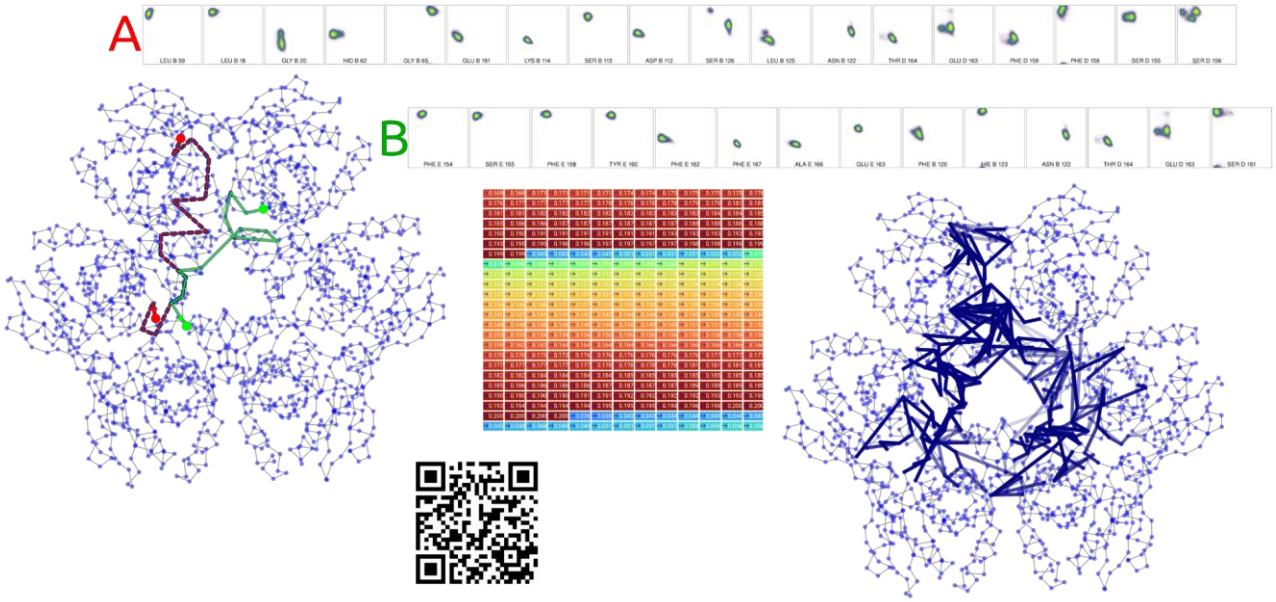
4.5. Vremenska analiza i točke promjene

Iz prethodnih odjeljaka je vidljivo da korelacije među aminokiselinama nisu jednoliko raspodijeljene po proteinu, te da određena područja u prostoru pokazuju znatno veću koreliranost u odnosu na druga (slika 44). To najvjerojatnije znači da su područja koja su znatnije korelirana odgovorna za svojevrstan prijenos informacije s jednog dijela proteina na drugi. No sama koreliranost aminokiselina predstavlja mjeru sinkroniziranosti promjena u Ramachandranovim kutovima gledanu kroz cijelo trajanje simulacije i sama po sebi ne daje više informacija o vremensko-prostornom tijeku tih promjena. Detaljnija analiza vremenskog

tijeka promjena i njihova veza s prostornom blizinom aminokiselina na kojima se događaju logičan je korak ka pronalaženju putova alosteričke komunikacije u proteinu.

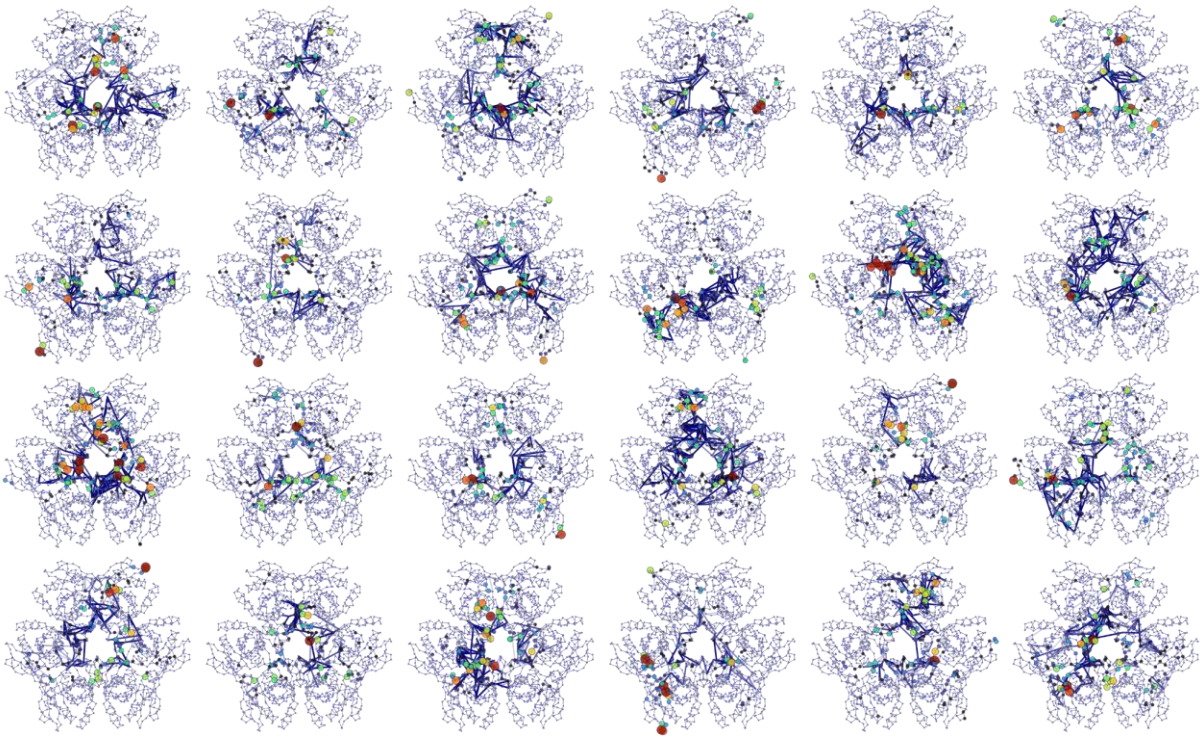
Kao model za takve alosteričke putove možemo uzeti promatranje promjena koje su bliske u prostoru i vremenu, jer je razumno pretpostaviti da promjene kutova na jednoj aminokiselini najprije utječu na svoju neposrednu okolinu, a isto tako da su ograničene na neko kraće vremensko razdoblje neposredno nakon same promjene. Nadalje, pretpostavljajući model uzročno posljedične povezanosti pretpostavlja se da su promjene vremenski orijentirane, te da promjene kutova u nekom trenutku mogu utjecati samo na promjene kasnije u vremenu i nikako obrnuto. Dakle, potreban je robustan i automatiziran način pretrage promjena u Ramchandranovim kutovima tijekom trajanja simulacije, a već je rečeno da se to svodi na traženje naglih skokova u iznosima tih kutova tj. vremenskim serijama, te da takvih skokova u velikom dijelu simulacija ili nema ili ima samo nekoliko (do desetak). U tu svrhu korišten je program specifično dizajniran za traženje promjena u vremenskim serijama Ruptures. Kao rezultat te analize za svaku aminokiselinu u svakoj od simulacija dobije se niz od nekoliko točaka u vremenu zajedno s apsolutnim iznosima tih promjena (promjenama Ramachandranovih kutova). Možemo pretpostaviti da se na neki način promjene prenose s jedne aminokiseline na aminokiseline u njezinoj neposrednoj okolici, što može predstavljati mehanizam prijenosa informacije (alosterije) u proteinu. Kao podloga tog mehanizma mora postojati kako prostorna tako i vremenska lokalizacija tih promjena. Dakle, problem se svodi na to da se pronađu putovi prijenosa prostorno-vremenski lokaliziranih promjena. Na prvi pogled, imajući u vidu kompleksnost i veliki broj ovakvih promjena i njihov složen prostorno-vremenski raspored, čini se da je pronalazak svih mogućih putova ovog tipa u proteinu zadatak za koji je gotovo nemoguće naći efikasno rješenje. Za pronalazak ovakvih putova komunikacije upotrebljena je reprezentacija teorije grafova, gdje je protein predstavljen kao graf u kojem su aminokiseline čvorovi koji su povezani vezama koje su izvedene iz njihovih vremenskih promjena.

Konačni rezultat ovakve analize na takvim grafovima je opet jedan veliki skup mogućih putova, iako su parametri koji su zadani relativno strogi (udaljenost d je uzeta kao 10 Å, a vremenski razmak između događaja je ~1/100 ukupnog trajanja simulacije). Njihov broj varira od oko 300 pa sve do gotovo 3000 u pojedinoj trajektoriji, dok je dužina putova i do 18 aminokiselina u pojedinim slučajevima. Primjer takvih putova prikazan je na slici 46.



Slika 46. Prikaz putova komunikacije između pojedinih dijelova proteina na primjeru trajektorije *5mx6_po4_hpa* koja ukupno ima 2746 ovakvih putova. Lijevo su označena dva puta A i B koji se na jednom dijelu i preklapaju, a gore su prikazani MDavocado dijagrami za oba puta. U sredini je dio koji prikazuje interaktivnu tablicu u kojoj su svu putovi složeni prema dužini i ukupnoj težini a desno je preklop svih mogućih putova. Vidljivo je da iako ih ima jako puno ti putovi ipak zauzimaju samo određeno područje bliže središtu proteina.

Na slici 47 prikazan je usporedni prikaz svih ovako pronađenih putova za svih 24 simulacija. Istovremeno je na tim dijagramima dan i prikaz korelacija sa slike 44. Uočljivo je da su unatoč velikom broju putova oni za pojedinu trajektoriju smješteni u pojedinim područjima koje nikada ne prelazi 20-30 % proteina čak i u slučajevima s najviše putova. U svim simulacijama vidljivo je vrlo dobro slaganje područja kojim prolaze putovi s područjima visoko koreliranih aminokiselina. Također je vidljivo da su i putovi kao i korelacije situirani uglavnom na područjima između dimera i uglavnom oko središnje šupljine unutar heksamera, što još dodatno ukazuje na važnost ovih područja kao komunikacijskih kanala. Treba naglasiti da iako su korelacije i putovi donekle povezani, jer se vidi da je za koreliranost potrebno bar nekoliko točaka promjene, koje su pak podloga izračuna putova, nije isprva posve razumljivo zašto bi se ovi putovi morali prostirati istim područjima. Korelacije predstavljaju vremenski prosjek kroz cijelu trajektoriju, a putovi su usko vezani za točke promjene u jednom malom dijelu vremena. Ovo saznanje predstavlja dodatnu nezavisnu potvrdu važnosti ovih područja u ovom proteinu.



Slika 47. Prikaz svih putova u svim trajektorijama i njihova usporedba s korelacijama. Putovi su označeni plavim linijama, a korelirane aminokiseline su označene obojanim kružnicama. Putovi se grupiraju, kao i korelirane aminokiselina, na sučeljima dimera i na centralnoj šupljini heksamera.

§ 5. ZAKLJUČAK

Istraživanja u ovom radu baziraju se na kristalografski određenim strukturama enzima fosforilaze purinskih nukleozida (PNP) iz bakterije *Helicobacter pylori* (HpPNP), koje unatoč tome što daju uvid u strukturu na atomskoj razini, ne mogu dati odgovor na dinamičko ponašanje ovog enzima. Izabrano je šest kristalografskih struktura HpPNP enzima PDB kodova: 5lu0, 5mx4, 5mx6, 6f4x, 6f52 i 6g7x. Danas se smatra da ovaj enzim kao i velika većina ostalih proteina pokazuje svojstva alosterije, međutim mehanizam kako se ovaj fenomen odvija je daleko od toga da bude sasvim razjašnjen. Za njegovo objašnjenje ne možemo se osloniti samo na statičku sliku koja se dobiva iz kristalografske strukture, već se mora uključiti i dinamika proteina koju generiramo *in silico*. Rezultati takvih simulacija su često ogromni skupovi podataka u vidu koordinata atoma zapisanih u vremenu, trajektorije.

U ovom radu razvijen je inovativan pristup koji se bazira na već dugo poznatim Ramachandranovim kutovima φ i ψ, koji su tradicionalno korišteni za opis konformacije proteinske okosnice, no samo u statičkom kontekstu. Novorazvijen program nazvan MDavocado promatra te kutove, ali tako da svaku aminokiselinu gleda zasebno i prati vremenski slijed razvoja φ i ψ kutova u tzv. MDavocado dijagramu. Ovakav pristup odmah daje informaciju o dijelovima proteina koji su dinamički aktivni a koji su statični. Konkretno kod HpPNP enzima dobivaju se dijagrami iz kojih su naizgled očiti dijelovi proteina koji su dinamični te oni koji su u istoj konformaciji tijekom cijelog trajanja simulacije. Oni daju i vrlo karakteristične razlike za dijelove HpPNP enzima koji prolaze kroz značajne konformacijske promjene (kao što je segmentacija H8 α-zavojnice). Iz MDavocado dijagrama na istraživanim strukturama HpPNP enzima vidi se da većina aminokiselina zapravo miruje, a manji broj njih radi određeni broj skokova između dobro definiranih stanja. Malobrojni skokovi koji se uočavaju međusobno su uzročno-posljedično povezani i predstavljaju neki vid prijenosa signala kroz protein što može biti jedan od segmenata alosteričke komunikacije.

Dio MDavocado programa je također i njegova nadogradnja, odnosno izračun svih cirkularnih korelacija između svih aminokiselina. Pokazalo se da tek jedan izuzetno mali postotak (oko 0,05 % u svim simulacijama) od ukupnog broja korelacija prelazi apsolutnu vrijednost korelacijskog koeficijenta od 0,5. Taj rezultat je donekle razumljiv jer je i velika većina kutova zapravo statična uz određen šum, odnosno gibaju se oko neke ravnotežne

vrijednosti. Korelacija između takva slična dva signala uvijek je blizu nule. Postalo je jasno da je za znatniju korelaciju između dva signala potrebno da oba imaju neku skokovitu promjenu, te da je korelacija veća što su te promjene vremenski bliže. Prikazom svih koreliranih aminokiselina ustanovljeno je da se one ne nalaze slučajno raspodijeljene po proteinu već da se predominantno nalaze na područjima na dodiru dvaju monomera koji čine funkcionalni dimer u heksameru HpPNP-a (A-D, B-E ili C-F sučeljima), zatim oko aktivnog mjesta te oko središnje šupljine u heksameru.

Sljedeći korak nadogradnje MDavocado dijagrama bio je izračun svih mogućih točaka promjene u vrijednosti Ramachandranovih kutova za svaku aminokiselinu kako bi se pratio vremenski tijek komunikacije koju vidimo iz korelacija. Točke promjene (promjena ili φ ili ψ kuta) su automatski pronađene i njihov broj se kreće od nule do desetak po aminokiselini. Tu je napravljen još jedan iskorak te je taj problem prebačen u domenu teorije grafova, te su relativno brzo izračunati svi putovi kojima se pojedine promjene šire po proteinu. Zanimljivost koja je proizašla iz ovih rezultat je da su ti putovi uvijek ograničeni na određeno manje područje (okvirno od 20 do 30 %) samog proteina. Preklapanje ovim komunikacijskih putova s koreliranim aminokiselinama predstavlja dobar putokaz za traženje alosteričke komunikacije u HpPNP enzimu.

Povezivanjem promjena konformacije aktivnog mjesta, smještanje i koordinacija fosfata kao jednog od supstrata te komunikacijskih putova u određenom vremenskom intervalu pokazuju potencijalna uzročno-posljedična događanja u ovom proteinu koja dosad nikad nisu opisana. Analiza aktivnog mjesta i H8 α-zavojnice pokazala je usku povezanost segmentacije α-zavojnice i pozicije fosfata u samom aktivnom mjestu, odnosno stabilizacije zatvorene konformacije aktivnog mjesta ako se fosfat nalazi u svom katalitičkom džepu. Također je opisan i sam proces segmentacije α-zavojnice gdje su pomoću nekoliko metoda uočene jasne promjene u sekundarnoj strukturi aminokiselina fenilalanina Phe221 i aspartata Asp222 te također njima prethodnih aminokiselina valina Val218, glutamata Glu219 i serina Ser220.

Sveukupni rezultati ovog istraživanja značajno doprinose razumijevanju dinamičkih svojstava HpPNP enzima, jednog od ključnih enzima u metabolizmu bakterije *H. pylori*. Korištenjem MD simulacija, omogućeno je praćenje lokalnih i globalnih konformacijskih promjena enzima na vremenskim i prostornim skalama koje su nedostupne klasičnim eksperimentalnim metodama. Posebna pažnja posvećena je identifikaciji regija koje igraju

ključnu ulogu u alosteričkoj komunikaciji, kao i analizi sinkroniziranih konformacijskih prijelaza koji utječu na funkciju enzima.

Ovim radom dobiva se posve novi pogled na molekulske simulacije enzima kao metodu praćenja njihove dinamike. Dobiveni rezultati pružaju novi uvid u alosteričku regulaciju i funkcionalnu dinamiku HpPNP enzima, pri čemu identificirani komunikacijski putovi i konformacijske promjene upućuju na moguće mehanizme alosteričke modulacije. Međutim, njihova kvaliteta i interpretacija ovise o primijenjenim metodama propagacije dinamike, posebice u kontekstu ograničenja molekulskih simulacija, kao što su duljina simulacija, korištena polja sila te mogućnost preciznog uzorkovanja konformacijskih stanja. Identifikacija ključnih aminokiselina odgovornih za alosteričku komunikaciju otvara mogućnost razvoja novih terapijskih strategija usmjerenih na inhibiciju ovog enzima, što može doprinijeti borbi protiv infekcija uzrokovanih bakterijom *H. pylori*.

§ 6. POPIS OZNAKA, KRATICA I SIMBOLA

5'-NT	5'-nukleotidaza
5lu0_apo	struktura PDB koda 5lu0 koja nema liganada u aktivnom mjestu
5lu0_po4	struktura PDB koda 5lu0 s fosfatom u aktivnom mjestu
5mx4_apo	struktura PDB koda 5mx4 koja nema liganada u aktivnom mjestu
5mx6_po4_hpa	struktura PDB koda 5mx6 s fosfatom i hipoksantinom u aktivnom mjestu
6f4x_apo	struktura PDB koda 6f4x koja nema liganada u aktivnom mjestu
6f4x_po4_fmc	struktura PDB koda 6f4x s fosfatom i formicinom A u aktivnom mjestu
6f52_apo	struktura PDB koda 6f52 koja nema liganada u aktivnom mjestu
6g7x_po4	struktura PDB koda 6g7x s fosfatom u aktivnom mjestu
ADA	adenozin deaminaza
ADSL	adenilosukcinat liaza
ADSS	adenilosukcinat sintaza
AK	adenozin kinaza
Ala	alanin
ALOKOMP	naziv interaktivne mrežne stranice cjelokupnih rezultata
AMBER	programski paket za izvođenje MD simulacija
AMP	adenozin monofosfat
APRT	adenin fosforibozil transferaza
Arg	arginin
AS	adenilosukcinat
Asn	asparagin
Asp	aspartat
ATaza	amidofosforibozil transferaza
ATIC	5-aminoimidazol-4-karboksamid ribonukleotid formil transferaza
B. subtilis	bakterija *Bacilus subtilis*
$C_{\alpha}/C_{\alpha,n}$	ugljik na koji je vezan bočni ogranak
C_n	karbonilni ugljikov atom
C_{n-1}	karbonilni ugljikov atom prethodne aminokiseline
Cys	cistein
d	udaljenost od analizirane aminokiseline

DNA/DNK	deoksiribonukleinska kiselina
deoD	oznaka gena zasložnog za ekspresiju PNP enzima
DSSP	metoda određivanja sekundarne strukture aminkiselina
E. coli	bakterija *Escherichia coli*
ff14SB	najčešće korišteno polje sila za izračun MD simulacija
FMC	formicin A
GAFF	generalno AMBER polje sila
GART	glicinamid ribonukleozid transformilaza
GESAMT	program za strukturno sravnjivanje makromolekulskih struktura
Gln	glutamin
Glu	glutamat
Gly	glicin
GMP	gvanozin monosfat
GMPS	gvanin monofosfat sintaza
GPU	grafička procesorska jedinica
H1-H8	oznaka svih 8 α-zavojnica PNP enzima
His/Hid	histidin
HPA	hipoksantin
HPRT	hipoksantin fosforibozil transferaza
HpPNP	enzim PNP iz bakterije *H. pylori*
H. pylori	bakterija *Helicobacter pylori*
Ile	izoleucin
IMP	inozin monofosfat
IMPDH	inozin monofosfat dehidrogenaza
KNF	Koshland-Némethy-Filmer model alosterije
Leu	leucin
Lys	lizin
m	ukupan broj aminokiselina
MALT	limfoidno tkivo povezano sa sluznicom
MD	molekulska dinamika
Met	metionin
ML	strojno učenje
MWC	Monod-Wyman-Changeux model alosterije
n	broj jednakih dijelova trajektorije
N	broj koraka simulacije

N	broj čestica u simulaciji
NMR	nuklearna magnetska rezonancija
N_n	dušikov atom
N_{n+1}	dušikov atom sljedeće aminokiseline
NPT	ansambl stanja koja ima imaju jednake vrijednosti čestica, tlak i temperaturu
NVT	ansambl stanja koja ima imaju jednake vrijednosti čestica, volumen i temperaturu
P	tlak
PAICS	fosforibozilaminoimidazol karboksilaza i fosforibozil-imidazolsukcinokarboksamid sintaza
PCA	analiza glavnih komponenti
PDB	makromolekulska baza podataka
P. falciparum	bakterija *Plasmodium falciparum*
PFAS	fosforibozil formilglicinamid sintaza
pH	oznaka kiselosti medija
Phe	fenilalanin
PNP	fosforilaza purinskih nukleozida
PO4	fosfatni ion
PPAT	fosforibozil pirofosfat amidotransferaza
PPI-AC	trojna terapija antibiotika (inhibitori protonskih pumpi, amoksicilin i klaritromicin)
PPI-BTM	četverovalentna terapija antibiotika (inhibitori protonskih pumpi, bizmut, tetraciklini i metronidazol)
Pro	prolin
PRPP	fosforibozil pirofosfat
PRPS	PRPP sintetaza
QR-kod	tip je matričnog dvodimenzionalnog koda koji vodi na ALOKOMP stranice
Q-vrijednost	mjera slaganja trodimenzijskih struktura proteina
R	aminokiselinski bočni ogranak
r_c	cirkularni korelacijski koeficijent
RMSD	korijen srednje kvadratne devijacije položaja atoma
r_{xy}	linearni korelacijski koeficijent
S1-S9	oznaka svih 9 β-ploča PNP enzima
Ser	serin
Δt	vremenski interval trajektorije
T	ukupno vrijeme trajanja simulacije
T	temperatura

TEM	transmisijski elektronski mikroskop
Thr	treonin
Trp	triptofan
Tyr	tirozin
V	volumen
Val	valin
x_i	vrijednost kuta x
y_i	vrijednost kuta y

§ 7. LITERATURNI IZVORI

1. B. Marshall, J. R. Warren, *The Lancet* **323** (1984) 1311–1315.
2. Y.-C. Chen, P. Malfertheiner, H.-T. Yu, C.-L. Kuo, Y.-Y. Chang, F.-T. Meng, Y.-X. Wu, J.-L. Hsiao, M.-J. Chen, K.-P. Lin, C.-Y. Wu, J.-T. Lin, C. O'Morain, F. Megraud, W.-C. Lee, E. M. El-Omar, M.-S. Wu, J.-M. Liou, *Gastroenterology* **166** (2024) 605–619.
3. M. Miftahussurur, Y. Yamaoka, D. Y. Graham, *Expert. Rev. Mol. Med.* **19** (2017) e4.
4. Z. Štefanić, G. Mikleušević, M. Narczyk, B. Wielgus-Kutrowska, A. Bzowska, M. Luić, *Croat. Chem. Acta* **86** (2013) 117–127.
5. G. Mikleušević, Z. Štefanić, M. Narczyk, B. Wielgus-Kutrowska, A. Bzowska, M. Luić, *Biochimie* **93** (2011) 1610–1622.
6. K. Ichida, M. Hosoyamada, T. Hosoya, H. Endou, u: *Genetic Diseases of the Kidney*, Elsevier, 2009, str. 651–660.
7. B. Gomaz, Z. Štefanić, *Symmetry (Basel)* **16** (2024) 124.
8. M. Narczyk, B. Bertoša, L. Papa, V. Vuković, I. Leščić Ašler, B. Wielgus-Kutrowska, A. Bzowska, M. Luić, Z. Štefanić, *FEBS J.* **285** (2018) 1305–1325.
9. B. Gomaz, A. Pandini, A. Maršavelski, Z. Štefanić, *J. Chem. Inf. Model*. 64 (2024) 5742–5748.
10. N. Kim, u: *Helicobacter Pylori*, Springer Nature Singapore, Singapore, 2023, pp. 3–21.
11. K. Rosinke, V. J. Starai, T.R. Hoover, *Cells* **13** (2024) 1438.
12. W. Delport, S. W. van der Merwe, *Best Pract. Res. Clin. Gastroenterol.* **21** (2007) 215–236.
13. S. Ansari, Y. Yamaoka, *Helicobacter* **22** (2017).
14. T. Lind, S. V. van Zanten, P. Unge, R. Spiller, E. Bayerdörffer, C. O'Morain, K. D. Bardhan, M. Bradette, N. Chiba, M. Wrangstadh, C. Cederberg, J. Idström, *Helicobacter* **1** (1996) 138–144.
15. P. Malfertheiner, F. Megraud, C. A. O'Morain, J. Atherton, A. T. R. Axon, F. Bazzoli, G. F. Gensini, J. P. Gisbert, D. Y. Graham, T. Rokkas, E. M. El-Omar, E. J. Kuipers, *Gut* **61** (2012) 646–664.
16. C. J. L. Murray, K. S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, S. C. Johnson, A. J. Browne, M. G. Chipeta, F. Fell, S. Hackett, G. Haines-Woodhouse, B. H. Kashef Hamadani, E. A. P. Kumaran, B. McManigal, S. Achalapong, R. Agarwal, S. Akech, S. Albertson, J. Amuasi, J. Andrews, A. Aravkin, E. Ashley, F.-X. Babin, F. Bailey, S. Baker, B. Basnyat, A. Bekker, R. Bender, J. A. Berkley, A. Bethou, J. Bielicki, S. Boonkasidecha, J. Bukosia, C. Carvalheiro, C. Castañeda-Orjuela, V. Chansamouth, S. Chaurasia, S. Chiurchiù, F. Chowdhury, R. Clotaire Donatien, A. J. Cook, B. Cooper, T. R. Cressey, E. Criollo-Mora, M. Cunningham, S. Darboe, N. P. J. Day, M. De Luca, K. Dokova, A. Dramowski,

S. J. Dunachie, T. Duong Bich, T. Eckmanns, D. Eibach, A. Emami, N. Feasey, N. Fisher-Pearson, K. Forrest, C. Garcia, D. Garrett, P. Gastmeier, A. Z. Giref, R. C. Greer, V. Gupta, S. Haller, A. Haselbeck, S. I. Hay, M. Holm, S. Hopkins, Y. Hsia, K. C. Iregbu, J. Jacobs, D. Jarovsky, F. Javanmardi, A. W. J. Jenney, M. Khorana, S. Khusuwan, N. Kissoon, E. Kobeissi, T. Kostyanev, F. Krapp, R. Krumkamp, A. Kumar, H. H. Kyu, C. Lim, K. Lim, D. Limmathurotsakul, M. J. Loftus, M. Lunn, J. Ma, A. Manoharan, F. Marks, J. May, M. Mayxay, N. Mturi, T. Munera-Huertas, P. Musicha, L. A. Musila, M. M. Mussi-Pinhata, R. N. Naidu, T. Nakamura, R. Nanavati, S. Nangia, P. Newton, C. Ngoun, A. Novotney, D. Nwakanma, C. W. Obiero, T. J. Ochoa, A. Olivas-Martinez, P. Olliaro, E. Ooko, E. Ortiz-Brizuela, P. Ounchanum, G. D. Pak, J. L. Paredes, A. Y. Peleg, C. Perrone, T. Phe, K. Phommasone, N. Plakkal, A. Ponce-de-Leon, M. Raad, T. Ramdin, S. Rattanavong, A. Riddell, T. Roberts, J. V. Robotham, A. Roca, V. D. Rosenthal, K. E. Rudd, N. Russell, H. S. Sader, W. Saengchan, J. Schnall, J. A. G. Scott, S. Seekaew, M. Sharland, M. Shivamallappa, J. Sifuentes-Osornio, A. J. Simpson, N. Steenkeste, A. J. Stewardson, T. Stoeva, N. Tasak, A. Thaiprakong, G. Thwaites, C. Tigoi, C. Turner, P. Turner, H. R. van Doorn, S. Velaphi, A. Vongpradith, M. Vongsouvath, H. Vu, T. Walsh, J. L. Walson, S. Waner, T. Wangrangsimakul, P. Wannapinij, T. Wozniak, T. E. M. W. Young Sharma, K. C. Yu, P. Zheng, B. Sartorius, A. D. Lopez, A. Stergachis, C. Moore, C. Dolecek, M. Naghavi, *The Lancet* **399** (2022) 629–655.

17. P. Katelaris, R. Hunt, F. Bazzoli, H. Cohen, K. M. Fock, M. Gemilyan, P. Malfertheiner, F. Mégraud, A. Piscoya, D. Quach, N. Vakil, L. G. Vaz Coelho, A. LeMair, J. Melberg, *J. Clin. Gastroenterol.* **57** (2023) 111–126.
18. R. Deng, H. Zheng, H. Cai, M. Li, Y. Shi, S. Ding, *Front. Immunol.* **13** (2022).
19. Y. Gong, K. Zhai, L. Sun, L. He, H. Wang, Y. Guo, J. Zhang, *Microbiol. Spectr.* **11** (2023).
20. X. Yao, S. Xiao, L. Zhou, *Helicobacter* **26** (2021).
21. H. Zheng, P. Xia, W. Fu, S. Ding, *Helicobacter* **29** (2024).
22. J. K. Y. Hooi, W. Y. Lai, W. K. Ng, M. M. Y. Suen, F. E. Underwood, D. Tanyingoh, P. Malfertheiner, D. Y. Graham, V. W. S. Wong, J. C. Y. Wu, F. K. L. Chan, J. J. Y. Sung, G. G. Kaplan, S. C. Ng, *Gastroenterology* **153** (2017) 420–429.
23. Y. Li, H. Choi, K. Leung, F. Jiang, D. Y. Graham, W. K. Leung, *Lancet Gastroenterol. Hepatol.* **8** (2023) 553–564.
24. N. Aumpan, V. Mahachai, R. Vilaichone, *JGH Open* **7** (2023) 3–15.
25. V. Babuš, V. Presečki, M. Katičić, M. Balija, I. Zorić, L. Kronja, A. Sabo, J. Vrličak, S. Čuković-Čavka, *Liječnički Vjesnik : Glasilo Hrvatskoga Liječničkog Zbora* **119** (1997) 139–142.
26. J. Strinić, Antimikrobna Osjetljivost Sojeva *Helicobacter Pylori*, Diplomski rad, Sveučilište u Zagrebu, 2022.
27. S. Suerbaum, P. Michetti, *N. Eng. J. Med.* **347** (2002) 1175–1186.
28. M. Rahat, M. Saqib, M. Ahmed, M. Suleman, S. M. Ismail, H. Mumtaz, H. Khalid, *Ann. Med. Surg.* **85** (2023) 2756–2760.

29. B. Drumm, G. I. Perez-Perez, M. J. Blaser, P. M. Sherman, *N. Eng. J. Med.* **322** (1990) 359–363.

30. M. Kivi, Y. Tindberg, M. Sörberg, T. H. Casswall, R. Befrits, P. M. Hellström, C. Bengtsson, L. Engstrand, M. Granström, *J. Clin. Microbiol.* **41** (2003) 5604–5608.

31. M. Schmid, M. Knoblauch, u: *Pathophysiology*, Springer New York, New York, NY, 1979, str. 261–315.

32. S. Teyssen, S. T. Chari, J. Scheid, M. V. Singer, *Dig. Dis. Sci.* **40** (1995) 247–255.

33. D. R. Scott, E. A. Marcus, D. L. Weeks, A. Lee, K. Melchers, G. Sachs, *Infect. Immun.* **68** (2000) 470–477.

34. S. Foynes, N. Dorrell, S. J. Ward, R. A. Stabler, A. A. McColm, A. N. Rycroft, B. W. Wren, *Infect. Immun.* **68** (2000) 2016–2023.

35. M. L. Worku, Q. N. Karim, J. Spencer, R. L. Sidebotham, *J. Med. Microbiol.* **53** (2004) 807–811.

36. L. T. Hu, H. L. Mobley, *Infect. Immun.* **58** (1990) 992–998.

37. V. M. Hernández, A. Arteaga, M. F. Dunn, *FEMS Microbiol. Rev.* **45** (2021).

38. A. W. Debowski, S. M. Walton, E.-G. Chua, A. C.-Y. Tay, T. Liao, B. Lamichhane, R. Himbeck, K. A. Stubbs, B. J. Marshall, A. Fulurija, M. Benghezal, *PLoS Pathog.* **13** (2017) e1006464.

39. J. G. Kusters, A. H. M. van Vliet, E. J. Kuipers, *Clin. Microbiol. Rev.* **19** (2006) 449–490.

40. D. Y. Graham, *Gut* **53** (2004) 1235–1243.

41. A. J. Morris, M. R. Ali, G. I. Nicholson, G. I. Perez-Perez, M. J. Blaser, *Ann. Intern. Med.* **114** (1991) 662–663.

42. J. Tack, N. J. Talley, M. Camilleri, G. Holtmann, P. Hu, J.-R. Malagelada, V. Stanghellini, *Gastroenterology* **130** (2006) 1466–1479.

43. A. C. Ford, S. Mahadeva, M. F. Carbone, B. E. Lacy, N. J. Talley, *The Lancet* **396** (2020) 1689–1702.

44. S. J. O. Veldhuyzen van Zanten, M. F. Dixon, A. Lee, *Gastroenterology* **116** (1999) 1217–1229.

45. J. J. Y. Sung, E. J. Kuipers, H. B. El-Serag, *Aliment. Pharmacol. Ther.* **29** (2009) 938–946.

46. A. H. Chang, J. Parsonnet, *Clin. Microbiol. Rev.* **23** (2010) 837–857.

47. J. Ma, H. Shen, L. Kapesa, S. Zeng, *Oncol. Lett.* **11** (2016) 2959–2964.

48. G. Nardone, *Aliment. Pharmacol. Ther.* **17** (2003) 75–81.

49. P. Courvalin, R. Leclercq, L. B. Rice, *Antibiogram*, ESKA Publishing, Paris, France, 2011.

50. K. van Amsterdam, A. Bart, A. van der Ende, *Antimicrob. Agents Chemother.* **49** (2005) 1477–1482.

51. J. Versalovic, D. Shortridge, K. Kibler, M. V. Griffy, J. Beyer, R. K. Flamm, S. K. Tanaka, D. Y. Graham, M. F. Go, *Antimicrob. Agents Chemother.* **40** (1996) 477–480.
52. R. A. Moore, B. Beckthold, S. Wong, A. Kureishi, L. E. Bryan, *Antimicrob. Agents Chemother.* **39** (1995) 107–111.
53. A. Bzowska, E. Kulikowska, D. Shugar, *Pharmacol. Ther.* **88** (2000) 349–425.
54. A. Dyzma, B. Wielgus-Kutrowska, A. Girstun, Z. J. Matošević, K. Staroń, B. Bertoša, J. Trylska, A. Bzowska, *Int. J. Mol. Sci.* **24** (2023) 2157.
55. B. Bertoša, G. Mikleušević, B. Wielgus-Kutrowska, M. Narczyk, M. Hajnić, I. Leščić Ašler, S. Tomić, M. Luić, A. Bzowska, *FEBS J.* **281** (2014) 1860–1871.
56. H. M. Berman, *Nucleic Acids Res.* **28** (2000) 235–242.
57. E. Krissinel, *J. Mol. Biochem.* **1** (2012) 76–85.
58. E. Krissinel, K. Henrick, *Acta Crystallogr. D Biol. Crystallogr.* **60** (2004) 2256–2268.
59. L. Hedstrom, *Chem. Rev.* **109** (2009) 2903–2928.
60. E. Villa, E. Ali, U. Sahu, I. Ben-Sahra, *Cancers (Basel)* **11** (2019) 688.
61. G. Liechti, J. B. Goldberg, *J. Bacteriol.* **194** (2012) 839–854.
62. M. Narczyk, M. I. Wojtyś, I. Leščić Ašler, B. Žinić, M. Luić, E. K. Jagusztyn-Krynicka, Z. Štefanić, A. Bzowska, *J. Enzyme Inhib. Med. Chem.* **37** (2022) 1083–1097.
63. Z. Štefanić, G. Mikleušević, M. Luić, A. Bzowska, I. Leščić Ašler, *Int. J. Biol. Macromol.* **101** (2017) 518–526.
64. G. Koellner, A. Bzowska, B. Wielgus-Kutrowska, M. Luić, T. Steiner, W. Saenger, J. Stępiński, *J. Mol. Biol.* **315** (2002) 351–371.
65. G. V. Isaksen, K. H. Hopmann, J. Åqvist, B. O. Brandsdal, *Biochemistry* **55** (2016) 2153–2162.
66. F. Canduri, V. Fadel, L. A. Basso, M. S. Palma, D. S. Santos, W. F. de Azevedo, *Biochem. Biophys. Res. Commun.* **327** (2005) 646–649.
67. M. D. Erion, J. D. Stoeckler, W. C. Guida, R. L. Walter, Steven E. Ealick, *Biochemistry* **36** (1997) 11735–11748.
68. A. Chaikuad, R. L. Brady, *BMC Struct. Biol.* **9** (2009) 42.
69. M. Ghanem, A. S. Murkin, V. L. Schramm, *Chem. Biol.* **16** (2009) 971–979.
70. P. C. Kline, V. L. Schramm, *Biochemistry* **31** (1992) 5964–5973.
71. P. C. Kline, V. L. Schramm, *Biochemistry* **34** (1995) 1153–1162.
72. V. L. Schramm, *ACS Chem. Biol.* **8** (2013) 71–81.
73. V. L. Schramm, *Chem. Rev.* **118** (2018) 11194–11258.
74. E. Ringia, V. Schramm, *Curr. Top. Med. Chem.* **5** (2005) 1237–1258.
75. A. Bzowska, E. Kulikowska, D. Shugar, *Biochim. Biophys. Acta, Protein Struc. Mol. Enzymol.* **1120** (1992) 239–247.
76. A. Bzowska, L. Magnowska, Z. Kazimierczuk, *Z. Naturforsch. C* **54** (1999) 1055–1067.

77. D. L. Nelson, M. M. Cox, *Lehninger Principles of Biochemistry*, 6th ed., W. H. Freeman, New York, NY, 2012.

78. C. C. F. Blake, D. F. Koenig, G. A. Mair, A. C. T. North, D. C. Phillips, V. R. Sarma, *Nature* **206** (1965) 757–761.

79. L. N. Johnson, G. A. Petsko, *Trends. Biochem. Sci.* **24** (1999) 287–289.

80. A. McPherson, *Introduction to Macromolecular Crystallography*, Wiley, 2009.

81. Y. Xu, M. Havenith, *J. Chem. Phys.* **143** (2015).

82. J. Liu, R. Nussinov, *PLoS Comput. Biol.* **12** (2016) e1004966.

83. J. Monod, J. Wyman, J.-P. Changeux, *J. Mol. Biol.* **12** (1965) 88–118.

84. M. H. Ahmed, M. S. Ghatge, M. K. Safo, in: 2020, pp. 345–382.

85. D. E. Koshland, G. Némethy, D. Filmer, *Biochemistry* **5** (1966) 365–385.

86. L. Michaelis, M. M. L. Menten, *FEBS Lett.* **587** (2013) 2712–2720.

87. T. P. Kenakin, in: *Pharmacology in Drug Discovery and Development*, Elsevier, 2017, pp. 101–129.

88. G. N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, *J. Mol. Biol.* **7** (1963) 95–99.

89. S. W. Park, B. H. Lee, S. H. Song, M. K. Kim, *J. Struct. Biol.* **215** (2023) 107939.

90. A. A. Rosenberg, N. Yehishalom, A. Marx, A. M. Bronstein, *PNAS* **120** (2023).

91. B. Tam, Z. Qin, B. Zhao, S. Sinha, C. L. Lei, S. M. Wang, *Int. J. Mol. Sci.* **25** (2024) 850.

92. S. C. Lovell, I. W. Davis, W. B. Arendall, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson, *Proteins: Struc., Funct. Bioinf.* **50** (2003) 437–450.

93. B. Rizzuti, *Biochim. Biophys. Acta, Proteins Proteomics* **1870** (2022) 140757.

94. E. Lindahl, in: A. Kukol (Ed.), *Molecular Modeling of Proteins*, 2015, pp. 3–26.

95. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory. Comput.* **11** (2015) 3696–3713.

96. J. V. Ribeiro, R. C. Bernardi, T. Rudack, J. E. Stone, J. C. Phillips, P. L. Freddolino, K. Schulten, *Sci. Rep.* **6** (2016) 26536.

97. M. Kubat, *An Introduction to Machine Learning*, Springer International Publishing, Cham, 2015.

98. J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis, J. M. Jumper, *Nature* **630** (2024) 493–500.

99. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A.

Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **596** (2021) 583–589.

100. M. Jordan, J. Kleinberg, B. Schölkopf, *Pattern Recognition and Machine Learning*, n.d.

101. J. Han, M. Kamber, J. Melton, S. Buxton, T. J. Teorey, S. S. Lightstone, T. P. Nadeau, J. Celko, I. Witten, E. Frank, G. C. Simsion, G. C. Witt, J. Schiller, A. Voisard, T. Halpin, K. Evans, P. Hallock, B. Maclean, S. Ceri, P. Fraternali, A. Bongio, M. Brambilla, S. Comai, M. Matera, A. R. Simon, U. Fayyad, G. G. Grinstein, A. Wierse, P. Rigaux, M. Scholl, A. Voisard, *Designing Data-Intensive Web Applications*, n.d.

102. A. Pandini, A. Fornili, F. Fraternali, J. Kleinjung, *Bioinformatics* **29** (2013) 2053–2055.

103. F. Maschietto, B. Allen, G. W. Kyro, V. S. Batista, *J. Chem. Phys.* **158** (2023).

104. D. K. Brown, D. L. Penkler, O. Sheik Amamuddy, C. Ross, A. R. Atilgan, C. Atilgan, Ö. Tastan Bishop, *Bioinformatics* **33** (2017) 2768–2771.

105. www.python.org/ (preuzeto 09. kolovoza 2024.).

106. T. E. Oliphant, *Comput. Sci. Eng.* **9** (2007) 10–20.

107. https://jupytearth.org/ (preuzeto 09. kolovoza 2024.).

108. R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, O. Beckstein, in: 2016, pp. 98–105.

109. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, *J. Comput. Chem.* **32** (2011) 2319–2327.

110. C. Truong, L. Oudre, N. Vayatis, *Signal Processing* **167** (2020) 107299.

111. J. A. Bednar, J. Crail, J. Crist-Harif, P. Rudiger, G. Brener, C. B, J. Mease, J. Signell, J.-L. Stevens, M. Liquet, B. Collins, I. Thomas, A. Thorve, S. Bird, thuydotm, esc, kbowen, N. Abdennur, O. Smirnov, S. H. Hansen, maihde, A. Oriekhov, A. Ahmadia, B. A. B. Jr, C. H. Brandt, C. Tolboom, E. G., J. Bourbeau, J. J. Schmidt, M. G. Delicious., *Zonodo* (2022).

112. L. Schrödinger, W. DeLano, https://www.pymol.org/ (preuzeto 03. listopada 2024).

113. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **66** (2010) 486–501.

114. R. Salomon-Ferrer, D. A. Case, R. C. Walker, *WIREs Comput. Mol. Sci.* **3** (2013) 198–210.

115. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **25** (2004) 1157–1174.

116. J. Wang, W. Wang, P. A. Kollman, D. A. Case, *J. Mol. Graph. Model.* **25** (2006) 247–260.

117. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **79** (1983) 926–935.

118. I. S. Joung, T. E. Cheatham, *J. Phys. Chem. B* **112** (2008) 9020–9041.

119. R. J. Loncharich, B. R. Brooks, R. W. Pastor, *Biopolymers* **32** (1992) 523–535.

120. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **81** (1984) 3684–3690.

121. A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **8** (2012) 1542–1555.

122. R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, R. C. Walker, *J. Chem. Theory Comput.* **9** (2013) 3878–3888.

123. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T. E. Oliphant, *Nature* **585** (2020) 357–362.

124. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. Pietro Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C. N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D. A. Nicholson, D. R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G. A. Price, G.-L. Ingold, G. E. Allen, G. R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J. T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J. V. de Miranda Cardoso, J. Reimer, J. Harrington, J. L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville, M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N. J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P. A. Brodtkorb, P. Lee, R. T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T. J. Pingel, T. P. Robitaille, T. Spura, T. R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y. O. Halchenko, Y. Vázquez-Baeza, *Nat. Methods* **17** (2020) 261–272.

125. W. McKinney, *Data Structures for Statistical Computing in Python* in: 2010, pp. 56–61.

126. F. Pedregosa, V. Michel, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, J. Vanderplas, D. Cournapeau, F. Pedregosa, G. Varoquaux, A. Gramfort, B. Thirion, O. Grisel, V. Dubourg, A. Passos, M. Brucher, M. Perrot, É. Duchesnay, *Scikit-Learn: Machine Learning in Python* Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos. Matthieu Perrot, 2011.

127. P. J. A. Cock, T. Antao, J. T. Chang, B. A. Chapman, C. J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, M. J. L. de Hoon, *Bioinformatics* **25** (2009) 1422–1423.

128. W. Kabsch, C. Sander, *Biopolymers* **22** (1983) 2577–2637.

129. D. R. Roe, T. E. Cheatham, *J. Chem. Theory Comput.* **9** (2013) 3084–3095.

130. D. R. Roe, T. E. Cheatham, *J. Comput. Chem.* **39** (2018) 2110–2117.

131. The Astropy Collaboration, A. M. Price-Whelan, P. L. Lim, N. Earl, N. Starkman, L. Bradley, D. L. Shupe, A. A. Patil, L. Corrales, C. E. Brasseur, M. Nöthe, A. Donath, E. Tollerud, B. M. Morris, A. Ginsburg, E. Vaher, B. A. Weaver, J. Tocknell, W. Jamieson, M. H. van Kerkwijk, T. P. Robitaille, B. Merry, M. Bachetti, H. M. Günther, T. L. Aldcroft, J. A. Alvarado-Montes, A. M. Archibald, A. Bódi, S. Bapat, G. Barentsen, J. Bazán, M. Biswas, M. Boquien, D. J. Burke, D. Cara, M. Cara, K. E. Conroy, S. Conseil, M. W. Craig, R. M. Cross, K. L. Cruz, F. D'Eugenio, N. Dencheva, H. A. R. Devillepoix, J. P. Dietrich, A. D. Eigenbrot, T. Erben, L. Ferreira, D. Foreman-Mackey, R. Fox, N. Freij, S. Garg, R. Geda, L. Glattly, Y. Gondhalekar, K. D. Gordon, D. Grant, P. Greenfield, A. M. Groener, S. Guest, S. Gurovich, R. Handberg, A. Hart, Z. Hatfield-Dodds, D. Homeier, G. Hosseinzadeh, T. Jenness, C. K. Jones, P. Joseph, J. B. Kalmbach, E. Karamehmetoglu, M. Kałuszyński, M. S. P. Kelley, N. Kern, W. E. Kerzendorf, E. W. Koch, S. Kulumani, A. Lee, C. Ly, Z. Ma, C. MacBride, J. M. Maljaars, D. Muna, N. A. Murphy, H. Norman, R. O'Steen, K. A. Oman, C. Pacifici, S. Pascual, J. Pascual-Granado, R. R. Patil, G. I. Perren, T. E. Pickering, T. Rastogi, B. R. Roulston, D. F. Ryan, E. S. Rykoff, J. Sabater, P. Sakurikar, J. Salgado, A. Sanghi, N. Saunders, V. Savchenko, L. Schwardt, M. Seifert-Eckert, A. Y. Shih, A. S. Jain, G. Shukla, J. Sick, C. Simpson, S. Singanamalla, L. P. Singer, J. Singhal, M. Sinha, B. M. Sipőcz, L. R. Spitler, D. Stansby, O. Streicher, J. Šumak, J. D. Swinbank, D. S. Taranu, N. Tewary, G. R. Tremblay, M. de Val-Borro, S. J. Van Kooten, Z. Vasović, S. Verma, J. V. de Miranda Cardoso, P. K. G. Williams, T. J. Wilson, B. Winkel, W. M. Wood-Vasey, R. Xue, P. Yoachim, C. Zhang, A. Zonca, *Astrophys. J.* **935** (2022) 167.

132. A. A. Hagberg, D. A. Schult, P. J. Swart, *Exploring Network Structure, Dynamics, and Function using NetworkX* 2008, pp. 11–15.

133. https://bokeh.pydata.org (preuzeto 02. rujna 2024).

134. M. Waskom, *J. Open Source Softw.* **6** (2021) 3021.

135. J. D. Hunter, *Comput. Sci. Eng.* **9** (2007) 90–95.

§ 8. ŽIVOTOPIS

Osobne informacije

Boris Gomaz
Institut Ruđer Bošković (IRB), Bijenička cesta 54, 10 000 Zagreb
bgomaz@irb.hr
Datum rođenja 14/12/1995

Radno iskustvo

Veljača 2021. – danas	Asistent, doktorand, Institut Ruđer Bošković, Bijenička cesta 54, 10 000 Zagreb, Hrvatska, Zavod za fizičku kemiju, Laboratorij za kemijsku i biološku kristalografiju

Obrazovanje

Listopad 2021. – danas	Poslijediplomski doktorski studij Kemije Prirodoslovno-matematički fakultet, Sveučilište u Zagrebu, Horvatovac 102A, 10000 Zagreb, Hrvatska Smjer : Anorganska i strukturna kemija Voditelj rada: dr. sc. Zoran Štefanić
Listopad 2018. – listopad 2020.	Magistar kemije (24. rujna 2020.) Prirodoslovno-matematički fakultet, Sveučilište u Zagrebu, Horvatovac 102A, 10000 Zagreb, Hrvatska Smjer: Anorganska kemija i biokemija Diplomski rad: *Utjecaj halogenidnih iona na kristalizaciju te molekulsku i kristalnu strukturu bakrovog(II), niklovog(II) i kobaltovog(II) derivata ljudskog inzulina* Voditeljice rada: prof. dr. sc. Biserka Prugovečki, dr. sc. Ivana Kekez
Listopad 2014. – listopad 2018.	Prvostupnik kemije (07. rujna 2018.) Prirodoslovno-matematički fakultet, Sveučilište u Zagrebu, Horvatovac 102A, 10000 Zagreb, Hrvatska Završni rad: *San o plavoj ruži* Voditeljica rada: prof. dr. sc. Višnja Vrdoljak

Rujan 2010. – lipanj 2014.	Prirodoslovno-matematička gimnazija, Druga gimnazija Varaždin, 42 000 Varaždin, Hrvatska

Publikacije

1. **B. Gomaz**, A. Pandini, A. Maršavelski, Z. Štefanić, *J. Chem. Inf. Model.* **64(15)** (2024) 5742–5748.
2. **B. Gomaz**, Z. Štefanić, *Symmetry* **16(1)** (2024) 124.

Priopćenja na znanstvenim skupovima i školama

Ukupno: 9 (tijekom diplomskog studija 1, u sklopu izrade disertacije 8)
Usmena priopćenja: 7
Posterska priopćenja: 1

Izdvojena priopćenja tijekom doktorata:
1. 8th Faculty of Science PhD Student Symposium, Zagreb, Hrvatska, 2024.; usmeno priopćenje
B. Gomaz, A. Pandini, A. Maršavelski, Z. Štefanić "Deciphering dynamics of purine nucleoside phosphorylase"

2. PyCon DE & PyData Berlin 2024, Berlin, Njemačka, 2024.; usmeno priopćenje
Z. Štefanić. **B. Gomaz** "How Python helped us uncover secrets of protein motion"

3. Hot Topics in Contemporary Crystallography 6 (HTCC 6), Dubrovnik, Hrvatska, 2024.; usmeno priopćenje
B. Gomaz, A. Pandini, A. Maršavelski, Z. Štefanić "Computational analysis of purine nucleoside phosphorylase dynamics"

4. EMBO Lecture Course Structural biophysics of biomolecular complexes, Istanbul, Turska, 2023.; usmeno priopćenje
B. Gomaz, Z. Štefanić "Selected Talks from Abstracts: Allostery analysis of oligomeric enzymes by combining X-ray structures and MD simulations with advanced visualization"

5. The Twenty-ninth Slovenian-Croatian Crystallographic Meeting, Topolšica, Slovenija, 2023.; usmeno priopćenje
B. Gomaz, Z. Štefanić "Understanding allostery in oligomeric enzymes through big data visualization"

6. Hot Topics in Contemporary Crystallography 5 (HTCC5), Dubrovnik, Hrvatska, 2023.; usmeno priopćenje
Z. Štefanić, **B. Gomaz** "Exploring the structural and dynamic data"

7. Solutions in Chemistry, Sveti Martin na Muri, Hrvatska, 2022.; postersko priopćenje
B. Gomaz, Z. Štefanić "Structure comparison and structural alignments of oligomeric purine nucleoside phosphorylases"

8. The Twenty-eighth Croatian-Slovenian Crystallographic Meeting, Poreč, Hrvatska, 2022.; usmeno priopćenje
B. Gomaz, Z. Štefanić "Oligomeric symmetry of purine nucleoside phosphorylases"

Sudjelovanje u radionicama i školama

1. Hot Topics in Contemporary Crystallography 6, Dubrovnik, Hrvatska, travanj 2024.
2. Hot Topics in Contemporary Crystallography 5, Dubrovnik, Hrvatska, travanj 2023.

Organizacija

1. Volonter u organizaciji i provedbi Hot Topics in Contemporary Crystallography 6 (HTCC6), Dubrovnik, 2024.
2. Volonter u organizaciji i provedbi Hot Topics in Contemporary Crystallography 5 (HTCC5), Dubrovnik, 2023.
3. Volonter u organizaciji i provedbi European Powder Diffraction Conference (EPDIC), Šibenik, 2022.

Nagrade

1. Nagrada za najbolje usmeno priopćenje na Hot Topics in Contemporary Crystallography 6 (HTCC6) dodijeljeno od strane Selvite naslova: "Computational analysis of purine nucleoside phosphorylase dynamics"

Projekti

1. Projekt kratkotrajnog boravka dodijeljen od strane Hrvatske zaklade za znanost: Program mobilnosti – odlazna mobilnost asistenata
20. listopada 2023. – 18. prosinca 2023. Brunel University London

2. Projekt kratkotrajnog boravka dodijeljen od strane HPC-Europa3 Transnational Access programme: EC-funded research visit
20. rujna 2021. – 2. prosinca 2021. Brunel University London

Članstva

Hrvatsko kemijsko društvo