



Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET

Kemijski odsjek

Zoe Jelić Matošević

Studentica 1. godine Poslijediplomskog sveučilišnog studija KEMIJA

Računalne metode u istraživanju alosterije

Kemijski seminar 1

prema radu

Greener, J. G. & Sternberg, M. J. *Curr. Opin. Struct. Biol.* 50, 1–8 (2018).

„Structure-based prediction of protein allostery“

Zagreb, 2020. godina

Sadržaj

§ 1. UVOD	1
§ 2. MODELI ALOSTERIČKOG MEHANIZMA	3
2.1. Klasični modeli alosterije – MWC i KNF	3
2.2. Suvremeno viđenje alosterije	4
§ 3. RAČUNALNE METODE PROUČAVANJA ALOSTERIJE	6
3.1. Analiza normalnih modova	7
3.1.1. Metode pronalaženja skrivenih alosteričkih džepova	7
3.2. Simulacije molekulske dinamike	9
3.2.1. Usmjerene simulacije molekulske dinamike.....	9
3.2.2. Proučavanje alosterije izračunom pK_a vrijednosti polarnih aminokiselina	10
3.2.3. Primjena neuralnih mreža u proučavanju alosterije	12
§ 4. ZAKLJUČAK	17
§ 5. LITERATURNI IZVORI	XVIII

§ 1. UVOD

Alosterija je svojstvo svih bioloških makromolekula te omogućava brz odgovor i prilagodbu na promjene u okolišu ili potrebe organizma i/ili stanice. Igra važnu ulogu u regulaciji biokemijskih procesa u svim biološkim sustavima jer se događa na puno manjim vremenskim skalama od promjena na razini genoma ili transkriptoma. Alosterija je u širem smislu riječi promjena do koje dolazi na jednom mjestu u biološkoj makromolekuli, a koja se odražava na aktivnost ili funkciju udaljenog mjesta na proteinu. Često se postiže reverzibilnim vezanjem alosteričkog liganda, kovalentnim modifikacijama kao što su fosforilacija, metilacija i glikozilacija, stvaranjem kompleksa s nekom većom molekulom kao što je drugi protein ili nukleinska kiselina te čak i apsorpcijom svjetlosti određene valne duljine.¹

Proučavanje alosterije zanimljivo je iz više perspektiva. Samo razumijevanje puteva kojim se alosterički signal prenosi kroz protein nužno je za kvalitetno razumijevanje biokemijskih procesa i njihove regulacija u biološkim sustavima. Iz farmakološke perspektive, alosterički modulatori imaju niz prednosti naspram ortosteričkim modulatora.¹ Ortosterički modulatori natječu se za vezanje u aktivnom mjestu proteina s nativnim ligandom te se očekuje da će imati antagonistički efekt na aktivnost proteina, dok se alosterički modulatori ne vežu u aktivnom mjestu proteina. Alosterički modulatori maraviroc² i cinacalcet³ u upotrebi za liječenje infekcije HIV-om i hiperparatireoidizma. Naposljetku, dizajn alosteričkih mjesta u postojećim proteinima kao i fuzioniranje proteina kako bi se dobio alosterički „prekidač“ predstavlja plodno područje za optimizaciju i razvoj novih sustava u području biotehnologije.

Proučavanje alosteričkih liganada izuzetno je složeno iz više razloga. Za razliku od ortosteričkih liganada, alosterički modulatori mogu imati agonistički i antagonistički učinak na aktivnost proteina. Kod razvoja novih alosteričkih modulatora u svrhu primjene u medicini i biotehnologiji predviđanje na koji način će alosterički modulator djelovati predstavlja značajan izazov.⁴ Nadalje, iako se smatra kako su gotovo svi proteini alosterički,¹ strukturni podaci o alosteričkim mjestima i modulatorima dostupni su samo za manji broj proteina. Postoji niz metoda za otkrivanje „kriptičnih“ alosteričkih mjesta za koja iz, primjerice, kristalne strukture nije očito da su dostupna ligandu.¹ Naposljetku, konformacijske promjene do kojih dolazi uslijed vezanja alosteričkog modulatora mogu se događati na vremenskim skalama

mikrosekundi ili čak milisekundi, čineći ih nedostupnima računalnim metodama kao što su simulacije molekulske dinamike (MD).⁵ Eksperimentalne metode igraju nezamjenjivu ulogu u proučavanju alosterije. Za metode strukturne računalne biofizike nužna je 3D struktura najčešće dobivena X-ray kristalografijom. Nuklearna magnetska rezonancija (NMR) se također često upotrebljava u svrhu istraživanja alosterije. Naposljetku, ukoliko se računalnim metodama identificiraju aminokiseline ključne za alosterički odgovor u nekom proteinu, važnost tih aminokiselona može se potvrditi mutacijskim eksperimentima.¹

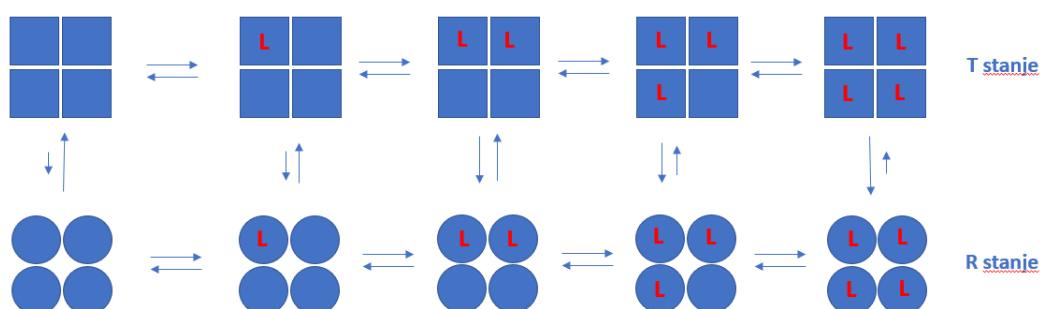
U fokusu ovog teksta biti će računalne metode istraživanja alosterije primjenjive na alosteriju proteina. Prikazati će se na koji način je moguće pronaći skrivene alosteričke džepove u proteinima te kako upotrijebiti simulacije molekulske dinamike za proučavanje alosteričkih promjena proteina unatoč ranije navedenim ograničenjima.

§ 2. MODELI ALOSTERIČKOG MEHANIZMA

2.1. Klasični modeli alosterije – MWC i KNF

Christian Bohr je 1904. uočio da afinitet hemoglobina za molekulu kisika ovisi o koncentraciji ugljikovog dioksida, fenomen koji se naziva „Bohrov efekt“.⁶ Zbog Bohrovog efekta u plućima, gdje je koncentracija CO₂ niska, hemoglobin ima visok afinitet za vezanje kisika što omogućuje efikasan prijenos kisika iz pluća u krvotok. U tkivima je zbog staničnog disanja koncentracija CO₂ visoka, što smanjuje afinitet hemoglobina za O₂ i na taj način osigurava oslobađanje kisika u tkivima, gdje je i potreban.

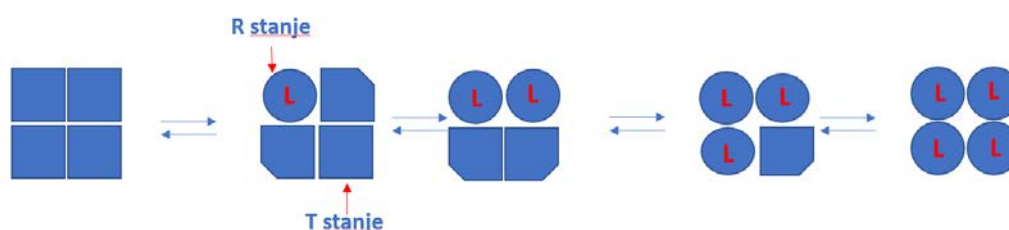
Na temelju slučaja hemoglobina razvijena su dva modela za objašnjenje alosterije – model MWC („Monod – Wyman – Changeux“) ili model KNF („Koschland – Nemethy – Filmer“). U oba modela podjedinice hemoglobina mogu postojati u dvije konformacije – R i T. Konformacija R (eng. *relaxed*) se odnosi na konformaciju hemoglobina koja veže kisik s velikim afinitetom, dok se konformacija T (eng. *tense*) odnosi na konformaciju hemoglobina koja ima nizak afinitet za vezanje kisika. Model MWC (slika 1) pretpostavlja da u multimernom proteinu (hemoglobin je tetramer) u svakom trenutku sve podjedinice moraju biti u istoj konformaciji. Pritom vezanje liganda može stabilizirati jednu ili drugu konformaciju. U slučaju hemoglobina, vezanje kisika stabilizira konformaciju R dok vezanje ugljikovog dioksida stabilizira konformaciju T.⁷



Slika 1. Model alosterije multimernog proteina MWF.

U modelu KNF (slika 2), koji se naziva i sekvencijalnim modelom, podjedinice multimera mogu se istovremeno nalaziti u različitim konformacijama. Vezanje liganda na jednu od podjedinica pritom mijenja afinitet susjednih podjedinica za ligand. Model KNF stoga

pretpostavlja mehanizam prisilnog pristajanja – vezanje alosteričkog liganda navodi protein u drugu konformaciju.⁷



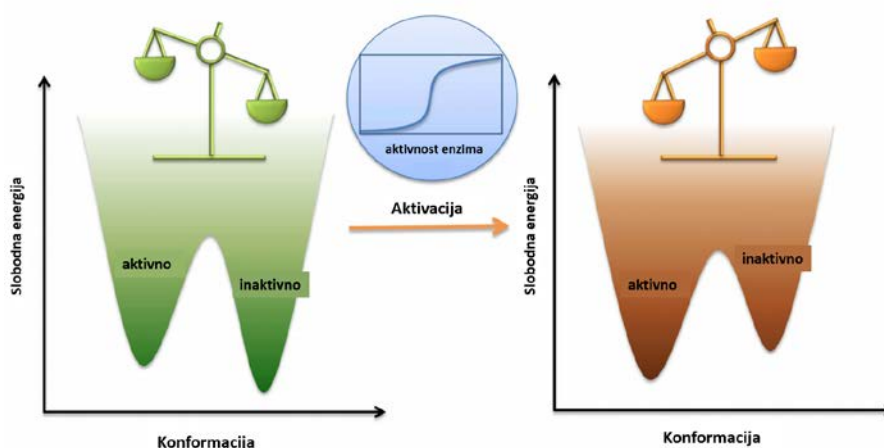
Slika 2. Model alosterije multimernog proteina KNF.

Iako su se modeli MWC i KNF pokazali kao korisni za objašnjenje fenomena alosterije u nekoliko sustava, nedostatni su za objašnjenje alosterije u svim proteinima. Ukoliko alosteriju definiramo u širem smislu kao promjenu do koje dolazi na jednom mjestu u proteinu a odražava se na nekom udaljenom mjestu u proteinu, očito je da i monomerni proteini mogu biti alosterički. Osim toga, opisani modeli se bave ravnotežom između samo dvije moguće konformacije. Viđenje alosterije prikazano u idućem poglavlju adekvatno objašnjava pojam alosterije u širem smislu te obuhvaća daleko veći broj proteina od modela MWC i KNF.

2.2. Suvremeno viđenje alosterije

Prethodno opisani modeli alosterije počivaju na dvije pretpostavke: i) svi alosterički proteini su multimeri, ii) alosteričku promjenu uvijek prati konformacijska promjena. Kao što ćemo vidjeti na nekim primjerima u idućim poglavljima, nijedna od te dvije pretpostavke ne mora uvijek biti ispunjena te, osim što i monomerni proteini mogu biti alosterički, alosterički efekt u mnogim proteinima nije popraćen većom konformacijskom promjenom.

Kako bi se navedena odstupanja od modela MWC i KNF mogla objasniti bilo je potrebno uvesti nove definicije alosterije. Dva su glavna suvremena pristupa objašnjenju alosterije. Prvi pristup izučava pomak populacije konformacijskih ensembala. Proteinima je u njihovom nativnom obliku dostupan velik broj mogućih konformacija. Pri vezanju alosteričkog liganda ne mijenja se nužno prosječna struktura proteina, već raspodjela vjerojatnosti zastupljenosti pojedine konformacije u konformacijskom prostoru, kao što je prikazano na slici 3. Proteinu je u ovom modelu i prije vezanja aktivatora bila dostupna aktivna konformacija, no vezanjem modulatora je povećana vjerojatnost zauzimanja aktivne konformacije čime se i eksperimentalno izmjerena aktivnost proteina povećava.⁶



Slika 3. Prikaz pomaka populacije konformacijskih ensemblea uslijed vezanja alosteričkog modulatora.⁸

Alosterija se može promatrati i iz perspektive alosteričkih mreža. Vezanje liganda uzrokuje perturbaciju u alosteričkom veznom mjestu koja se prenosi do aktivnog mjesta mrežom interakcija. Proučavati alosteriju iz ove perspektive znači pokušati identificirati niz interakcija i ključne aminokiseline kojima se alosterički signal provodi kroz protein.⁶

Ova dva viđenja alosterije nisu međusobno isključiva. U idućim poglavljima predstavljene su primjeri istraživanja mreže interakcija kojima se prenosi alosterički odgovor i primjeri istraživanja pomaka u konformacijskim ensembleima, no pokazat će se da je moguće i oboje istražiti istovremeno.

§ 3. RAČUNALNE METODE PROUČAVANJA ALOSTERIJE

Ukoliko je 3D struktura proteina dostupna, istraživačima se otvara mogućnost izučavanja alosterije većim brojem računalnih metoda. Ovisno o dostupnim strukturama, primjenom računalnih metoda mogu se tražiti odgovori na velik broj različitih pitanja.

U slučaju kada je alosterički modulator poznat te ako su dostupne strukture proteina u kompleksu s modulatorom i bez njega, često je moguće uočiti razlike u dvjema konformacijama već iz kristalnih struktura. No, u ovom pristupu postoje dva ograničenja: i) konformacija proteina u kristalnoj strukturi nije nužno jednaka konformaciji proteina u otopini ili u stanici zbog utjecaja slaganja u kristalnu strukturu, ii) moguće je da se u kristalnoj strukturi ne može uočiti razlika između dviju konformacija, no drugi eksperimenti pokazuju da alosterički odgovor na modulator postoji. Navedene dileme mogu se pokušati razriješiti analizom normalnih modova ili simulacijama molekulske dinamike te različitim pristupima analizi navedenih.⁵

Poznavanje razlika aktivne i inaktivne konformacije proteina ne govori ništa o samom mehanizmu aktivacije ili inhibicije funkcije proteina alosteričkim modulatorom. S obzirom da u alosteriji dolazi do prijenosa signala s alosteričkog mjesta u proteinu na neko drugo, za njezino razumijevanje potrebno je razlučiti mehanizme prijenosa alosteričkog signala. Konkretnije, potrebno je ustvrditi kojim se točno interakcijama alosterički signal provodi kroz protein, te koje su aminokiseline ključne za alosterički odgovor. U idućim poglavljima opisat će se i neki od pristupa traženju odgovora na ova pitanja.¹

Konačno, alosterička mjesta na proteinima imaju velik potencijal za primjenu u industriji kao što je farmakološka i biotehnološka. Iz tog razloga je otkrivanje dosad nepoznatih alosteričkih mjesta važan izazov u proučavanju alosterije. Ta mjesta mogu biti dostupna u konformaciji vidljivoj u kristalnoj strukturi, no moguće je i da bivaju otkrivena tek u rjeđim konformacijama proteina. Postoje mnogi pristupi za pronalaženje ovakvih mjesta, od kojih su neki rezultirali i pronalaskom novih alosteričkih modulatora.⁹

3.1. Analiza normalnih modova

Analiza normalnih modova (eng. *normal mode analysis*, NMA) je računalna metoda koja omogućava opis fluktuacija u konformaciji proteina. Pritom se protein tretira kao sustav harmoničkih oscilatora koji vibriraju oko ravnotežnog položaja. Vibracijski normalni modovi ovakvog sustava niskih frekvencija predstavljaju globalne kretnje proteina, kao što su pomaci proteinskih domena. Pokazano je kako su modovi niskih frekvencija ujedno često i biološki relevantni modovi.¹⁰ Za izučavanje alosterije važno je uočiti da i alosteričke konformacijske promjene najčešće odgovaraju modovima niske frekvencije.

Pri analizi normalnih modova proteina od najvećeg interesa su globalni modovi niske frekvencije za koje se može pretpostaviti da na njih ne utječu manja, lokalna gibanja. S obzirom na veličinu proteinskih sustava i uzevši u obzir navedenu pretpostavku, za NMA proteina često se koriste krupnozrnati modeli, odnosno modeli u kojima nije svaki atom predstavljen zasebno, već se čitav skup atoma modelira zajedno. Najčešće se pritom u obzir uzimaju samo C α atomi svake aminokiseline. U modelu elastičnih mreža (ENM, eng. *elastic network model*) svi C α atomi koji su bliži od neke zadane udaljenosti povezani su oprugama, a pokazano je da rješavanje takvog sustava harmoničkih oscilatora unatoč manjku detalja rezultira modovima niskih frekvencija koji opisuju funkcionalna kretanja proteina.¹⁰

3.1.1. Metode pronalaženja skrivenih alosteričkih džepova

Logičan kriterij za određivanje alosteričkih mjesta je da promjena dinamike proteina u mjestu za koje se pretpostavlja da je alosteričko dovodi do promjene dinamike proteina u aktivnom mjestu. Pri razvoju softvera *Allopred* ova pretpostavka je zajedno s metodama strojnog učenja korištena u svrhu razvoja modela za predviđanje alosteričkih mjesta u proteinima.

Pri razvoju softvera *Allopred* korištena je baza alosteričkih proteina ASbench¹¹. Za svaki protein iz ove baze napravljena je analiza normalnih modova pomoću modela elastičnih mreža. Pritom je za potencijalnu energiju sustava korištena funkcija

$$V = \sum_{\substack{r_{ij}^0 > R_c \\ i < j}} k (r_{ij} - r_{ij}^0)^2 \quad (1)$$

pri čemu je r_{ij} euklidska udaljenost između dva C α atoma u nekoj konformaciji a r_{ij}^0 predstavlja udaljenost C α atom u ravnotežnom položaju, R_c je maksimalna udaljenost između C α atoma za koje se modelira opruga, a k je konstanta opruge koja je ista za sve parove atoma.

Kako bi se identificirali alosterički džepovi prvo je bilo potrebno pronaći sva potencijalna vezna mjesta na proteinu. Za pronalaženje veznih mjesta korišten je program Fpocket¹². Potom je za svako od pronađenih veznih mjesta ponovljena analiza normalnih modova, tako da je konstanta opruge k za atome unutar džepova povećana 1,5 puta. Za one džepove kod kojih je povećanje konstante opruge dovelo do promjene dinamike aktivnog mjesta pretpostavlja se da su alosterički džepovi. Metrika za ovu perturbaciju je definirana izrazima

$$C_m = \sum_{i=1}^m \frac{v_i}{\omega_i} \quad (2)$$

$$v_i = \frac{1}{N_a} \sum_{j=1}^{N_a} |p_j - u_j| \quad (3)$$

C_m je mjera razlike između normalnih modova proteina u prvom slučaju (kada je konstanta k jednaka za sve parova C α atoma) i u drugom slučaju (kada je konstanta k za parove C α atoma u potencijalnom alosteričkom džepu povećana 1,5 puta). Pritom v_i predstavlja razliku odmak od ravnotežnog položaja za atome aktivnog mjesta između prve i druge NMA analize $|p_j - u_j|$ uprosječenu za sve atome aktivnog mjesta N_a . C_m se dobiva normalizacijom v_i s obzirom na frekvencije normalnih modova, ω_i . Ovom normalizacijom osigurano je da modovi najmanje frekvencije, za koje se pretpostavlja da su biološki relevantni, najviše utječu na metriku. No, istovremeno modovi viših frekvencija nisu u potpunosti zanemareni.

Za očekivati je kako će džepovi s većim brojem atoma imati i veću utjecaj na dinamiku proteina. Kako bi se izbjegla pristranost zbog veličine džepa uvedena je i druga mjera u kojoj je C_m normaliziran s obzirom na broj atoma u džepu, N_p .

$$E_m = \frac{C_m}{N_p} \quad (4)$$

Navedene mjere se potom zajedno s rezultatima programa Fpocket¹² koriste za rangiranje džepova programom SVM^{Light}¹³ u kojemu je implementirana verzija stroja potpornih vektora prilagođena za rangiranje.

Model je na skupu za testiranje na kojem nije bio treniran u 23 od 40 proteina alosterički džep rangirao kao prvi. S obzirom na to da ova metoda ne uzima u obzir unaprijed poznati ligand za predviđanje alosteričkog mjesta, moguće ju je upotrijebiti za otkrivanje novih alosteričkih džepova na proteinima.¹⁴ Program je javno dostupan na web serveru <http://www.sbg.bio.ic.ac.uk/allopred/home/> (datum pristupa 30. srpnja 2020.).

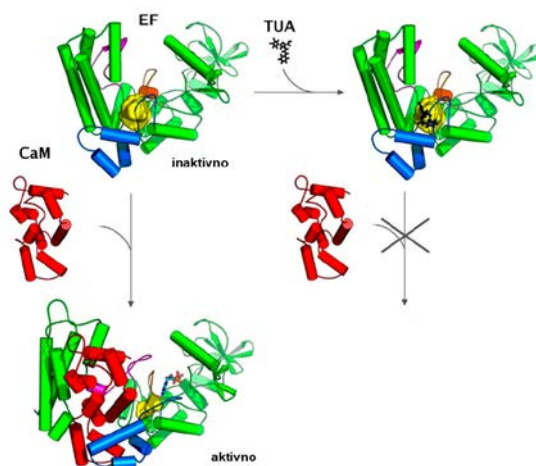
Osim predviđanja alosteričkih džepova analiza normalnih modova može se koristiti i općenitije za predviđanje konformacijskih promjena proteina. Prednost NMA je mala računaska zahtjevnost u usporedbi sa simulacijama molekulske dinamike, no ovaj pristup ima određena ograničenja. Zbog toga što se istražuju vibracije oko ravnotežnog položaja, NMA metodama nije moguće istraživati konformacijske promjene koje zahtijevaju prijelaz energetske barijere. Osim toga, pretpostavka da su modovi vibracije harmonijski nije nužno uvijek točna. Jedna od pretpostavki NMA je i da je struktura na kojoj se analiza provodi lokalni energetski minimum. Ovisno o načinu na koji je struktura dobivena, to nije uvijek točno te je potrebno obratiti posebnu pažnju i na minimizaciju strukture.^{1,7,10} NMA je izvrsna metoda za brzu analizu globalnih kretanja većeg broja proteina i neizostavan je alat u istraživanjima šireg opsega, no za detaljnu i vjerodostojniju sliku dinamike proteina najčešće se koriste simulacije molekulske dinamike.

3.2. Simulacije molekulske dinamike

3.2.1. Usmjerene simulacije molekulske dinamike

Pri istraživanju alosterije simulacijama molekulske dinamike uzorkovanje je značajan problem. Naime, alosteričke promjene se, kao što je već spomenuto, događaju u vremenskim skalama mikrosekundi ili čak milisekundi. Klasičnim simulacijama molekulske dinamike te vremenske skale su često nedostupne.

Kod usmjerenih (eng. *steered*) simulacija molekulske dinamike prijelaz početne u završnu strukturu simulira se uz djelovanje vanjske sile, što ubrzava promjenu konformacije. Laine i sur. su 2010. korištenjem usmjerenih simulacija molekulske dinamike opisali prijelaz toksina adenilil ciklaze (EF) iz *Bacillus anthracis* iz inaktivne u aktivnu, otvorenu konformaciju, proces koji je induciran vezanjem kalmodulina (CaM). Pomoću ICM PocketFinder¹⁵ programa za svaku od prijelaznih konformacija identificirani su džepovi na površini proteina. Ovim pristupom identificiran je i džep SABC (slika 4), za koji se pokazalo da se drastično smanjuje pri prijelazu iz inaktivne u aktivnu konformaciju.



Slika 4. Stvaranje kompleksa CaM-EF i inhibicija s TUA. SABC džep je prikazan žutom bojom.⁹

Veličina džepa i omjer hidrofobnih i hidrofilnih aminokiselina pokazao se kao prikladan za vezanje liganda, te je napravljeno virtualno skeniranje velikog broja liganada kako bi se odredili potencijalni alosterički regulatori. Pritom su se pozitivno vrednovali spojevi koji su se dobro vezali na SABC džep u inaktivnoj konformaciji, a negativno spojevi koji su se uspijevali vezati i na neke od prijelaznih konformacija. Identificiran je ligand (TUAdiCl) koji inhibira formiranje kompleksa CaM-EF pri mikromolarnim koncentracijama *in vitro*, a pri istim koncentracijama ne veže CaM u značajnoj mjeri. Iako se navedeni spoj još ne primjenjuje u liječenju antraksa (kojeg je *B. anthracis* uzročnik), ovaj primjer pokazuje kako računalne metode mogu biti dobar saveznik u razvoju novih alosteričkih lijekova.⁹

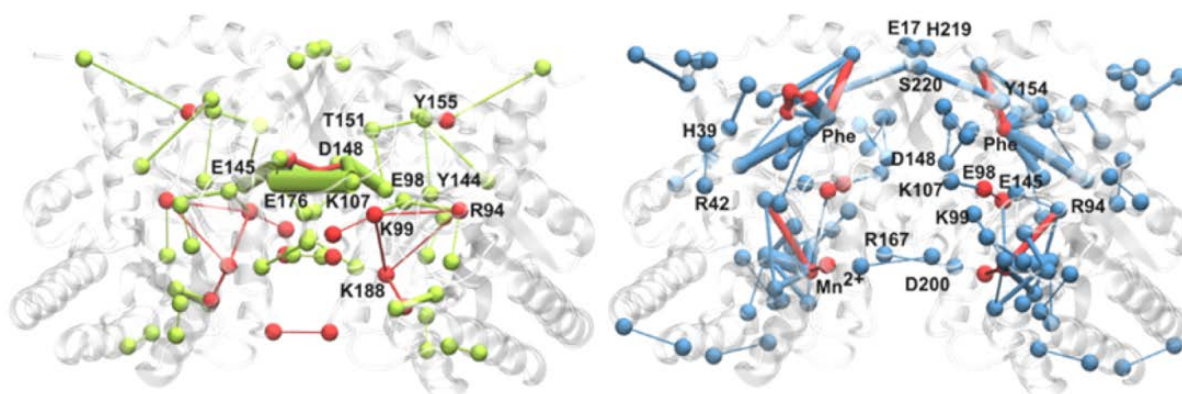
3.2.2. Proučavanje alosterije izračunom pK_a vrijednosti polarnih aminokiselina

U nekim slučajevima alosterički odgovor proteina nije praćen očitim konformacijskim promjenama, te identifikacija uzroka promjene aktivnosti proteina predstavlja pravi izazov. U takvim slučajevima potrebno je razviti osjetljivije metode za proučavanje razlika između aktivne i inaktivne konformacije.

Primjer proteina kod kojega je razlika u konformacijama aktivnog i inaktivnog stanja vrlo suptilna je i dimerni protein 3-deoksi-D-arabino-heptulosonat 7-fosfat sintaza (DAH7PS) iz bakterije *Neisseria meningitidis* koji katalizira prvi korak u biosintezi aromatskih aminokiselina, a alosterički je inhibiran vezanjem fenilalanina. Lang i sur.¹⁶ su osmislili pristup analizi simulacija molekulske dinamike u kojemu se prati promjena pK_a vrijednosti aminokiselina tijekom simulacije. Pristup se temelji na pretpostavci da aminokiselina uslijed

promjene svog okoliša najvjerojatnije mijenja i pK_a vrijednost. Ovom analizom se mogu pratiti samo polarne aminokiseline koje posjeduju skupinu koja se može protonirati ili deprotonirati jer nepolarne, kada su uklopljene u polipeptidni lanac, ne disociraju. Na navedenom proteinu ovim pristupom je pronađena mreža interakcija koja je značajno različita u aktivnoj i inaktivnoj konformaciji.

Simulacije molekulske dinamike provedene su na apo-strukturi i strukturi DAH7PS u kompleksu s fenilalaninom. Za svaku strukturu dobivenu ovim simulacijama pomoću programa PROPKA3.¹⁷ izračunate su pK_a vrijednosti za polarne aminokiseline. Korištenje velikog broja struktura dobivenih simulacijama molekulske dinamike omogućilo je izračun prosječne pK_a vrijednosti te standardne devijacije za svaku aminokiselinu, a usporedbom apo-simulacije i simulacije kompleksa protein-fenilalanin identificirano je 12 aminokiselina za koje se pK_a vrijednosti značajno razlikuje između simulacija. Identificirane su i aminokiseline čija fluktuacija pK_a vrijednosti se značajno promijenila, te aminokiseline za koje se pri vezanju fenilalanina mijenja izloženost otapalu. Izračun pK_a vrijednosti omogućio je predviđanje snage interakcija između parova aminokiselina na temelju predviđene protonacije, te je napravljena mapa interakcija za aktivnu i inaktivnu konformaciju.



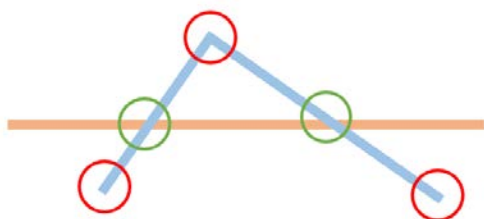
Slika 5. Prikaz mreže interakcija u aktivnoj (lijevo) i inaktivnoj (desno) konformaciji proteina DAH7PS. Odbojne interakcije su označene crvenom, a privlačne zelenom i plavom bojom.¹⁶

Identificirane su mreže interakcija aktivne i inaktivne konformacije za koje je vidljivo da se razlikuju (slika 5), razotkrivajući koje su to suptilne promjene u interakcijama aminokiselina koje dovode do promjene funkcije enzima. Rezultati prikazani na slici 5 impliciraju da E176 ima važnu ulogu u alosteričkom odgovoru na vezanje fenilalanina. U apo-strukturi pri pH 6,8 E176 bi trebao biti protoniran, dok bi u kompleksu s fenilalaninom trebao biti deprotoniran. Ukoliko je E176 uistinu važan za alosteričku regulaciju DAH7PS, pri višem pH bi zbog manjeg

udjela protoniranih E176 i efekt alosteričke inhibicije fenilalaninom trebao biti umanjen. Kako bi eksperimentalno potvrdili važnost interakcija E176 u alosteričkoj regulaciji DAH7PS, Lang i sur. su izmjerili aktivnost enzima pri pH od 6,8 do 8,3 u prisutnosti i odsutnosti fenilalanina. U skladu s izračunima pK_a vrijednosti, u odsutnosti fenilalanina pH nije imao značajan utjecaj na aktivnost, dok je u prisutnosti fenilalanina inhibicija aktivnosti enzima oslabljena pri višim pH. ITC eksperimentima provjereno je da je vezanje fenilalanina očuvano na navedenom pH rasponu.¹⁶ Na ovaj način je pokazano kako se opisanom metodom mogu identificirati suptilne promjene u mrežama interakcija kod proteina čiji alosterički odgovor ne prati značajna konformacijska promjena.

3.2.3. Primjena neuralnih mreža u proučavanju alosterije

Pri većini metoda analize trajektorija, uključujući i prethodno opisanu metodu analize pK_a vrijednosti, prvi korak je preklapanje svih dobivenih struktura. Na slici 6 prikazana je shema preklapanja dvije strukture iz koje je lako iščitati probleme s analizama temeljenima na



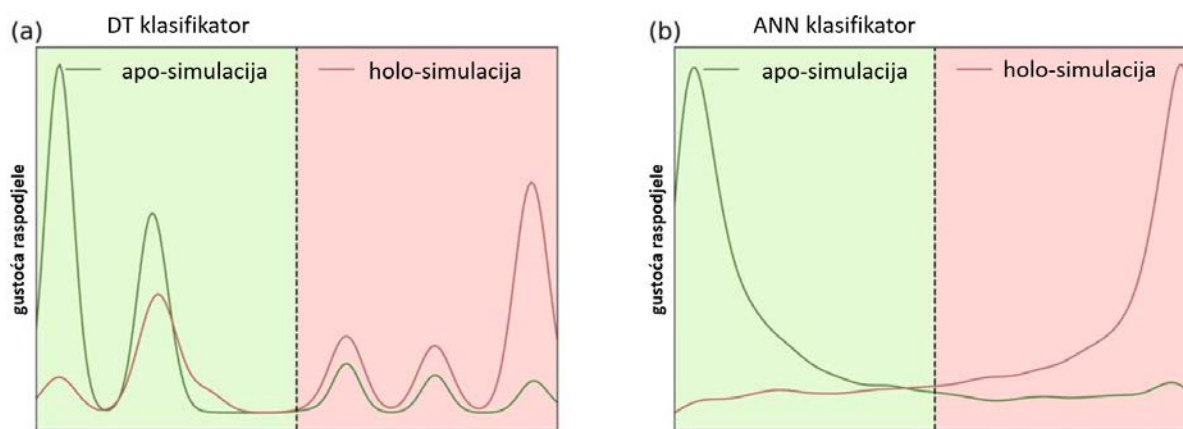
Slika 6. Shema preklapanja dviju struktura različitih konformacija.

preklapanju struktura. Jasno je da će usporedba dijelova označenih zelenim krugovima biti vjerodostojna, no usporedba dijelova označenih crvenim krugovima neće jer čak i kada je preklapanje čitave strukture optimalno, lokalna preklapanja mogu biti pogrešna. Preklapanje struktura može biti problematično u dva slučaja: i)

ukoliko, kao što je prikazano na slici, dolazi do većih pomaka u strukturi će kod preklapanja struktura samo mali broj mjesta biti ispravno preklapljen, te ii) ako je efekt od interesa dovoljno malen da i mala greška u preklapanju dijelova strukture može onemogućiti njegovo uočavanje.

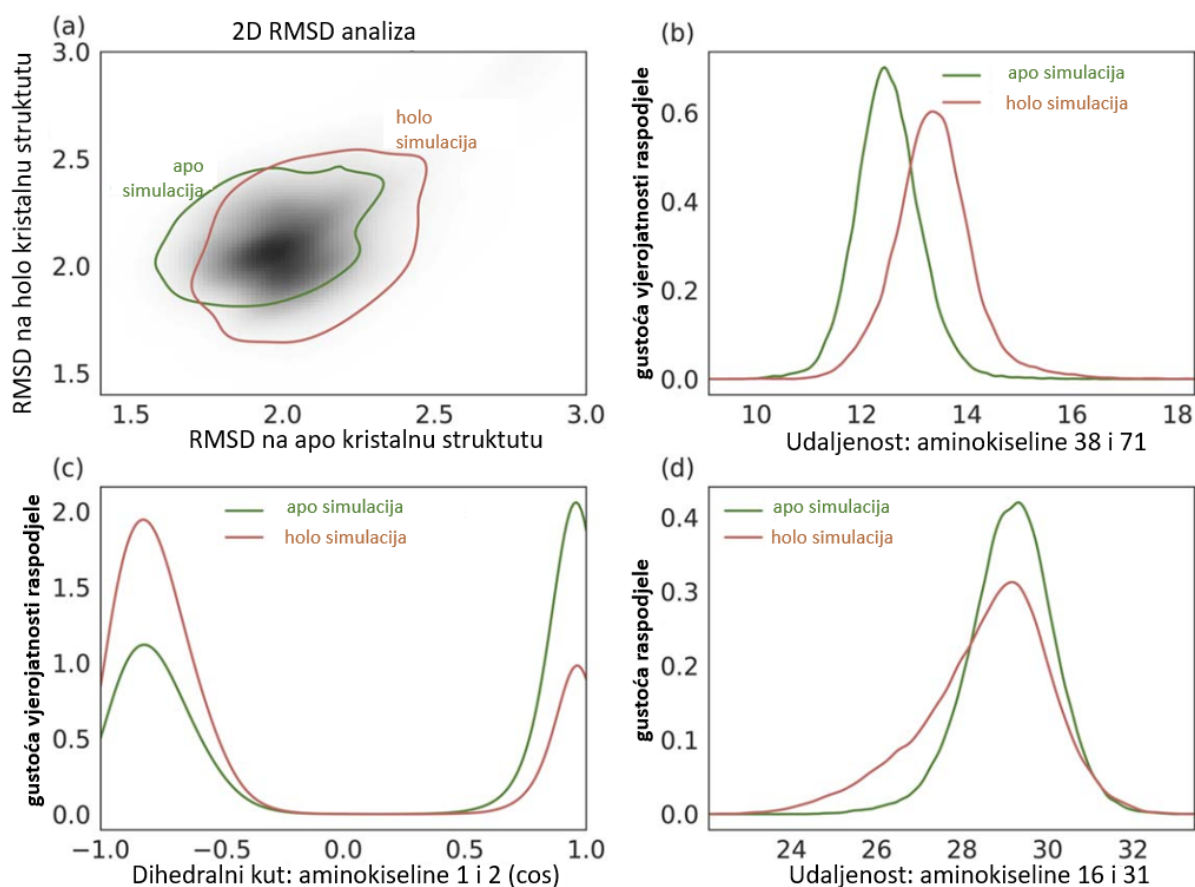
Domena PDZ2 humane tirozin-fosfataze 1E (hPTP1E) je još jedan primjer proteinske domene kod koje je alosterička promjena teško uočljiva usporedbom kristalnih struktura. Domena PDZ2 veže C-terminalne dijelove drugih proteina uslijed čega propagira alosterički signal na ostatak proteina. Kod analize simulacija domene PDZ2 kojima prethodi preklapanje struktura problematičan je upravo drugi slučaj, odnosno male greške u poravnanju struktura mogu predstavljati dovoljno snažan šum da bi se spriječilo uočavanje konformacijskih promjena. Kako bi se to izbjeglo potrebno je razviti analize kojima ne prethodi preklapanje struktura.

Alosteriju domene PDZ2 su 2018. istražili Zhou i sur. koristeći neuralne mreže (ANN, eng. *artificial neural network*) i stabla odlučivanja (DT, eng. *decision tree*).¹⁸ Nakon provedenih MD simulacija proteina u apo-obliku i s vezanim peptidom, udaljenosti Ca atoma i vrijednosti dihedralnih kuteva ψ i ϕ korištene su kao značajke za treniranje modela strojnog učenja. Analiza udaljenosti i kuteva ne zahtijeva prethodno preklapanje struktura. Pomoću stabla odlučivanja odabrane su relevantne značajke, koje su potom korištene za učenje daljnjih modela. Stabla odlučivanja i neuralne mreže su trenirani kako bi klasificirali strukture iz trajektorija prema tome dolaze li iz simulacije apo- ili holo-oblika. Osim klasifikacije, pomoću ovih modela može se izračunati i vjerojatnost pripadanja apo- ili holo-simulaciji za određenu strukturu (slika 7).



Slika 7. Gustoća raspodjele vjerojatnosti pripadanja apo- ili holo-simulaciji za apo- i holo-simulaciju domene PDZ2 dobivena DT (a) ili ANN (b) klasifikatorom.¹⁸

Analizom DTT i ANN modela identificirane su značajke ključne za klasifikaciju. Grafovi tih značajki izvrsna su ilustracija modela pomaka populacije za opis alosteričkih promjena. Već iz 2D grafa RMSD (slika 8) analize vidljivo je da protein ne zauzima dvije odvojene konformacije, već da pri vezanju liganda dolazi do blagog pomaka u zastupljenosti određenih dijelova konformacijskog prostora. Iz preklapanja struktura je stoga teško razlučiti razlike u konformaciji između dva oblika proteina. Grafovi udaljenosti dva ključna para aminokiselina te jednog dihedralnog kuta daju nešto detaljnija objašnjenja (slika 8). Kod sva tri primjera vidljivo je da ne dolazi do značajne promjene same vrijednosti kuta odnosno udaljenosti, već da se mijenja oblik krivulje gustoće raspodjele vjerojatnosti. Potvrda vjerodostojnosti modela proizlazi iz preklapanja svojstava identificiranih kao ključna s nizom drugih računalnih i eksperimentalnih istraživanja, što implicira da se ovaj pristup može primijeniti i na druge alosteričke proteine.¹⁸



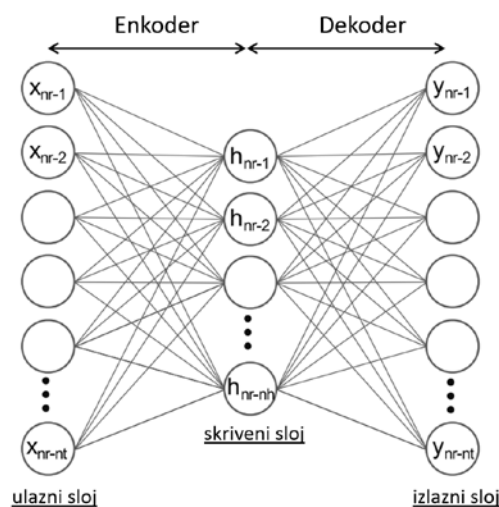
Slika 8. Analiza simulacije molekulske dinamike domene PDZ2 proteina hPTP1E u apo-obliku i u kompleksu s alosteričkim ligandom (holo). (a) 2D RMSD analiza trajektorija preklapljenij na kristalne structure apo-protein i kompleksa protein-ligand, (b) Gustoća raspodjele vjerojatnosti za udaljenost aminokiselina 38 i 71, (c) Gustoća raspodjele vjerojatnosti za kosinus dihedrallong kuta između aminokiselina 1 i 2, (d) Gustoća raspodjele za udaljenost aminokiselina 16 i 31.¹⁸

Tsuchiya i sur. također su istraživali domenu PDZ2 hPTP1E neuralnim mrežama, no primijenili su drugačiji pristup. Naime, koristili su enkoder-dekoder neuralnu mrežu kako bi razlučili koje interakcije se razlikuju između apo- i holo-oblika proteina.

Iz simulacija molekulske dinamike domene PDZ2 u apo-obliku i u kompleksu s ligandom izračunate su udaljenosti centara masa svih parova aminokiselina za svaki trenutak u simulaciji. Parovi aminokiselina su podijeljeni u skup za treniranje i skup za testiranje. Kao ulazni vektori značajki za neuralnu mrežu korišteni su vektori udaljenosti aminokiselina kroz simulaciju.

Enkoder-dekoder neuralne mreže uzimaju vektore značajki te im reduciraju dimenzionalnost smanjenjem broja neurona u idućim slojevima. U ovom primjeru prvi sloj ima 1500 neurona (što je jednako broju trenutaka u simulaciji), idući sloj ima 1000 neurona, zatim

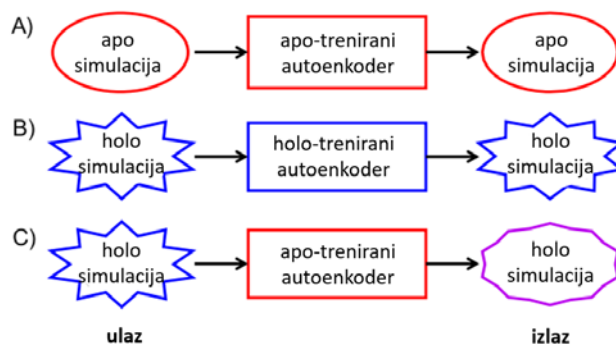
500, 200 i 100 (slika 9). Na taj način se postiže redukcija dimenzionalnosti u enkoder dijelu mreže.



Slika 9. Shema enkoder-dekoder mreže: nr predstavlja parove aminokiselina, nt trenutke u simulaciji a nh broj neurona u skrivenom sloju mreže.¹⁹

Potom dekoder dio mreže iz smanjenog prostora značajki rekonstruira izvorne vektore značajki. Mreža se trenira na način da se parametri prilagođavaju kako bi izlazni vektor bio što sličniji ulaznom.

Ideja iza ove metode temelji se na tome da će neuralna mreža ispravno rekonstruirati simulaciju apo-proteina za čije je provođenje korištena ista početna struktura kao i za simulaciju na kojoj je mreža trenirana, no simulaciju holo-proteina će izmijeniti na način da će svojstva koja se među simulacijama razlikuju tako da rekonstruirana simulacije više nalikuju simulaciji apo-strukture.



Slika 10. Shema principa određivanja ključnih alosteričkih interakcija enkoder-dekoder neuralnom mrežom.¹⁹

Usporedbom ulaza i izlaza u tom slučaju moguće je identificirati parove aminokiselina čija je dinamika različita između apo- i holo-simulacije. Radi kvantifikacije tog efekta uvedena

je mjera DIO (eng. *distance input output*) koja predstavlja razliku ulaznog i izlaznog vektora za svaki par aminokiselina. Klasteriranjem DIO vrijednosti identificirani su skupovi aminokiselina važni za alosteriju u domeni PDZ2. Aminokiseline identificirane kao ključne za alosteriju u domeni PDZ2 ovom metodom poklapaju se s aminokiselinama identificiranim u prethodnim NMR istraživanjima, što upućuje na validnost ove metode te mogućnost primjene ovakvog pristupa i na druge alosteričke proteine.¹⁹

§ 4. ZAKLJUČAK

Računalne metode su važan alat pri istraživanju alosterije u proteinima jer omogućavaju uvid u strukturu i dinamiku proteina na molekularnoj razini. Alosterija je ključan proces u biološkim sustavima te ima široke primjene u razvoju lijekova i biotehnologiji. Otkrivanje novih alosteričkih mjesta otvara mogućnost razvoja novih lijekova, a u toj potrazi su upravo računalne metode ključne. Server Allopred u kojemu je implementirana NMA predstavlja izvrstan prvi korak za takvu vrstu istraživanja.

Za detaljniju i vjerodostojniju sliku dinamike proteina valja upotrijebiti simulacije molekulske dinamike. Pri istraživanju proteina EF iz *B. anthracis* usmjerenim simulacijama molekulske dinamike identificiran je alosterički džep, što je naposljetku dovelo do identifikacije alosteričkog inhibitora aktivnog in vitro.

Kod proteina za koje se prosječna struktura ne mijenja uslijed vezanja alosteričkog modulatora potrebno su složenije metode za otkrivanje alosteričkih puteva. Promjene u pK_a vrijednostima tokom simulacije te strojno učenje neuralnim mrežama različitih arhitektura pokazali su se kao uspješni pristupa rješavanju ovog problema.

Istraživanje alosterije je aktualno područje te se za otkrivanje skrivenih alosteričkih džepova, detaljno opisivanje alosteričkih konformacijskih promjena i identifikaciju mreža interakcija kojima se prenosi alosterički signal svake godine razvija velik broj novih metoda. Iako je primjena alosteričkih modulatora u medicini i biotehnologiji tek u svojim začetcima, brzina kojom se ovo područje razvija obećava široku primjenu u budućnosti.

§ 5. LITERATURNI IZVORI

1. Greener, J. G. & Sternberg, M. J. Structure-based prediction of protein allostery. *Curr. Opin. Struct. Biol.* **50**, 1–8 (2018).
2. Selzentry (maraviroc) dosing, indications, interactions, adverse effects, and more. <https://reference.medscape.com/drug/selzentry-maraviroc-342638>.
3. Sensipar (cinacalcet) dosing, indications, interactions, adverse effects, and more. <https://reference.medscape.com/drug/sensipar-cinacalcet-342879>.
4. Amaro, R. E. Toward Understanding “the Ways” of Allosteric Drugs. *ACS Cent. Sci.* **3**, 925–926 (2017).
5. Hertig, S., Latorraca, N. R. & Dror, R. O. Revealing Atomic-Level Mechanisms of Protein Allostery with Molecular Dynamics Simulations. *PLOS Comput. Biol.* **12**, e1004746 (2016).
6. Liu, J. & Nussinov, R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLOS Comput. Biol.* **12**, e1004966 (2016).
7. Ribeiro, A. A. S. T. & Ortiz, V. A Chemical Perspective on Allostery. *Chem. Rev.* **116**, 6488–6502 (2016).
8. Tsai, C.-J. & Nussinov, R. A Unified View of “How Allostery Works”. *PLOS Comput. Biol.* **10**, e1003394 (2014).
9. Laine, E. *et al.* Use of allostery to identify inhibitors of calmodulin-induced activation of *Bacillus anthracis* edema factor. *Proc. Natl. Acad. Sci.* **107**, 11277–11282 (2010).
10. Skjaerven, L., Hollup, S. M. & Reuter, N. Normal mode analysis for proteins. *J. Mol. Struct. THEOCHEM* **898**, 42–48 (2009).

11. Huang, W. *et al.* ASBench: benchmarking sets for allosteric discovery. *Bioinforma. Oxf. Engl.* **31**, 2598–2600 (2015).
12. Le Guilloux, V., Schmidtke, P. & Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics* **10**, 168 (2009).
13. Joachims, T. Making large-Scale SVM Learning Practical. in *Advances in Kernel Methods - Support Vector Learning* (eds. Schölkopf, B., Burges, C. & Smola, A.) 169–184 (MIT Press, 1999).
14. Greener, J. G. & Sternberg, M. J. E. AlloPred: prediction of allosteric pockets on proteins using normal mode perturbation analysis. *BMC Bioinformatics* **16**, 335 (2015).
15. Abagyan, R. & Kufareva, I. The flexible pocketome engine for structural chemogenomics. *Methods Mol. Biol. Clifton NJ* **575**, 249–279 (2009).
16. Lang, E. J. M., Heyes, L. C., Jameson, G. B. & Parker, E. J. Calculated pK_a Variations Expose Dynamic Allosteric Communication Networks. *J. Am. Chem. Soc.* **138**, 2036–2045 (2016).
17. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
18. Zhou, H., Dong, Z. & Tao, P. Recognition of protein allosteric states and residues: Machine learning approaches. *J. Comput. Chem.* **39**, 1481–1490 (2018).
19. Tsuchiya, Y., Taneishi, K. & Yonezawa, Y. Autoencoder-Based Detection of Dynamic Allostery Triggered by Ligand Binding Based on Molecular Dynamics. *J. Chem. Inf. Model.* **59**, 4043–4051 (2019).