

# Izračun slobodne energije pomoću simulacija molekulske dinamike

---

**Cvjetinović, Robert**

**Undergraduate thesis / Završni rad**

**2023**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:217:171449>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-11-29**



*Repository / Repozitorij:*

[Repository of the Faculty of Science - University of Zagreb](#)





Sveučilište u Zagrebu  
PRIRODOSLOVNO-MATEMATIČKI FAKULTET  
Kemijски odsjek

Robert Cvjetinović

Student 3. godine Preddiplomskog sveučilišnog studija KEMIJA

# IZRAČUN SLOBODNE ENERGIJE POMOĆU SIMULACIJA MOLEKULSKE DINAMIKE

**Završni rad**

Rad je izrađen u Zavodu za fizikalnu kemiju

Mentor rada: izv. prof. dr. sc. Gordan Horvat

Zagreb, 2023.



Datum predaje prve verzije Završnog rada:

5. srpnja 2023.

Datum ocjenjivanja Završnog rada i polaganja Završnog ispita:

8. rujna 2023.

Mentor rada: izv. prof. dr. sc. Gordan Horvat

Potpis:



## Sadržaj

|  |      |
|--|------|
| § SAŽETAK.....   | VII  |
| § 1. UVOD.....   | 1    |
| 1.1. Slobodna energija.....                            | 1    |
| 1.2. Simulacije molekulske dinamike .....              | 2    |
| § 2. PRIKAZ ODABRANE TEME .....                        | 4    |
| 2.1. Perturbacija slobodne energije .....              | 4    |
| 2.2. <i>Umbrella sampling</i> .....                    | 11   |
| 2.3. Metadinamika.....                                 | 15   |
| 2.4. Primjena metoda za izračun slobodne energije..... | 23   |
| § 3. LITERATURNI IZVORI.....                           | XXIX |



## § Sažetak

Slobodna energija je termodinamička veličina bitna za procjenu stabilnosti nekog sustava i spontanosti procesa čija se promjena, ovisno o uvjetima pri kojima se sustav nalazi, izjednačava s Gibbsovom ili Helmholtzovom energijom. Promjena te veličine za neki proces može izračunati simulacijama molekulske dinamike rješavanjem Newtonovih jednadžbi gibanja kroz vrijeme, čime se generiraju različite konfiguracije na temelju kojih se izračunava ukupna promjena. No zbog ograničenih računalnih resursa teško je postići dobro uzorkovanje ukoliko je proces spor (npr. difuzija nekog iona u otopini do aktivnog mjesta proteina) ili ukoliko su prisutne visoke energetske barijere. Zbog toga se koriste metode kojima se poboljšava uzorkovanje faznog prostora. Perturbacija slobodne energije temelji se na zapisu Hamiltonijana za perturbirano stanje kao zbroj Hamiltonijana referentnog sustava i dodatnog, perturbacijskog člana. Ukoliko je poznat perturbacijski član, dovoljno je uzorkovati neperturbirani sustav i za svaku generiranu točku u faznom prostoru izračunati razliku Hamiltonijana i iz toga je jednostavno izračunati razliku slobodne energije. Iako je izvod egzaktna za bilo kakvu perturbaciju, potrebno je dobro preklapanje važnih dijelova faznog prostora perturbiranog i referentnog sustava jer je uzorkovanje referentnog sustava ograničeno računalnim resursima. *Umbrella sampling* (US) je metoda u kojoj se potencijalnoj energiji sustava dodaje dodatni član koji djeluje samo na određene stupnjeve slobode (reakcijske koordinate) i time se poboljšava uzorkovanje konfiguracijskog prostora. Adaptivna US metoda poboljšava uzorkovanje iterativnim izgrađivanjem vanjskog potencijala u oblik postane suprotan plohi slobodne energije duž reakcijskih koordinata čime raspodjela reakcijskih koordinata postaje konstantna, odnosno sustav se ponaša kao da ne postoje energetske barijere duž reakcijskih koordinata. To se provodi nizom simulacija, tako da se podaci iz prethodnih koriste za poboljšavanje vanjskog potencijala za sljedeću simulaciju. Drugi način je da se koriste harmonijski potencijali u nizu prozora koji su centrirani na različitim mjestima duž reakcijskih koordinata. Time se poboljšava uzorkovanje oko tih vrijednosti, a kombiniranjem podataka iz svih prozora može se izgraditi globalni profil slobodne energije. Metadinamika, slično adaptivnoj US metodi iterativno izgrađuje vanjski potencijal dok ne postane jednak negativnoj vrijednosti slobodne energije duž stupnjeva slobode na koje se primjenjuje. No to radi tako da se potencijalnoj energiji dodaju Gaussijani u određenim vremenskim intervalima. U konačnici, kada Gaussijani ispune



minimumu na plohi potencijalne energije, sustav će se slobodno kretati duž kolektivnih varijabli, a vanjski će potencijal biti jednak negativnoj vrijednosti slobodne energije (uz određenu konstantu).

## § 1. UVOD

### 1.1. Slobodna energija

Ukupno povećanje entropije sustava i okoline kriterij je spontanosti nekog procesa. Helmholtzova i Gibbsova energija su funkcije stanja koje se odnose samo na sustav i čija promjena pri određenim uvjetima vrijedi kao kriterij spontanosti procesa. Helmholtzova energija definirana je kao:

$$A = U - TS \quad (1.1)$$

pri čemu je  $U$  unutrašnja energija,  $T$  temperatura, a  $S$  entropija sustava. Promjena Helmholtzove energije zatvorenog sustava pri konstantnom volumenu i temperaturi, ukoliko se vrši samo  $pV$  rad, kriterij je spontanosti procesa. Definicija Gibbsove energije je:

$$G = H - TS \quad (1.2)$$

pri čemu je  $H$  entalpija. Kao i kod Helmholtzove energije, promjena Gibbsove energije zatvorenog sustava pri konstantnom tlaku i temperaturi, ukoliko se vrši samo  $pV$  rad, kriterij je spontanosti procesa. Particijska funkcija je funkcija koja iz molekulskih svojstava daje termodinamička.<sup>1</sup> Mikrokanonska particijska funkcija (particijska funkcija pri uvjetima konstantnog volumena, broja čestica i unutrašnje energije) u klasičnoj aproksimaciji za  $N$  identičnih čestica je:

$$Q_{NVE} = \frac{1}{N! h^{3N}} \int \delta[H(\mathbf{q}, \mathbf{p}) - E] d\mathbf{q}d\mathbf{p} \quad (1.3)$$

pri čemu je  $H(\mathbf{q}, \mathbf{p})$  Hamiltonijan za konfiguraciju, tj. točka u faznom prostoru  $(\mathbf{q}, \mathbf{p})$ , a  $\delta[H(\mathbf{q}, \mathbf{p}) - E]$  je Diracova delta koja je jednaka nuli osim ako su  $H(\mathbf{q}, \mathbf{p})$  i  $E$  jednaki, odnosno prilikom integriranja samo se pribrajaju one konfiguracije s energijom  $E$ . Faktor  $N!$  proizlazi iz identičnosti čestica, pri čemu se izraz (1.3) može modificirati tako da su čestice različite,<sup>2</sup> a  $h^{3N}$  služi kako bi rezultat bio konzistentan s kvantnomehaničkim rezultatom na visokoj temperaturi.<sup>3</sup> Entropija je definirana kao:

$$S = k_B \ln(Q_{NVE}) \quad (1.4)$$

U  $NVT$  ansamblu je izraz za particijsku funkciju glasi:

$$Q_{NVT} = \frac{1}{N! h^{3N}} \int e^{-\beta H(\mathbf{q}, \mathbf{p})} d\mathbf{q}d\mathbf{p} \quad (1.5)$$

dok se Helmholtzova se energija može definirati i kao:

$$A = -\beta^{-1} \ln(Q_{NVT}) \quad (1.6)$$

Particijska funkcija se može definirati za  $NpT$  ansambl, a onda se u jednadžbi analognoj (1.6) izračunava Gibbsova energija umjesto Helmholtzove. Hamiltonijan sadrži članove vezane uz kinetičku i potencijalnu energiju. No kinetički se dio može analitički riješiti, tj. prointegrirati i particijska se funkcija može zapisati kao produkt kinetičkog dijela (koji je particijska funkcija za idealni plin) i konfiguracijskog dijela. Onda se može definirati konfiguracijski integral:

$$Z_{NVT} = \int e^{-\beta U(\mathbf{q})} d\mathbf{q} \quad (1.5)$$

pri čemu je  $U(\mathbf{q})$  potencijalna energija sustava u ovisnosti o konfiguraciji  $\mathbf{q}$ . Konfiguracijski integral pomnožen s  $V^{-N}$  daje konfiguracijski dio particijske funkcije  $Q_{NVT}$ . S obzirom da je kinetički dio analitički rješiv on nas ne zanima tijekom simulacije, zbog čega i Monte–Carlo metode, koje za razliku od molekulske dinamike ne izračunavaju brzine čestica, mogu na temelju potencijalne energije konfiguracija predvidjeti svojstva sustava.<sup>b)</sup>

## 1.2. Simulacije molekulske dinamike

Termodinamička svojstva za kemijski zanimljive sustave ne mogu se dobiti analitički, ali ih je moguće dobiti simulacijama molekulske dinamike. Molekulska dinamika rješavanjem Newtonovih jednadžbi gibanja generira trajektoriju koja opisuje promjenu pozicija i količina gibanja atoma u sustavu pri čemu se jednadžbe gibanja rješavaju numerički. Sile se mogu izračunati kvantno-mehanički ili iz polja sila. Polja sila sadržavaju funkcionalne oblike potencijala interakcije između atoma te parametre kojima se potencijalna energija izračunava iz konfiguracije sustava. Sustav se nalazi u kutiji te se primjenjuju periodički uvjeti na granice kutije kako bi se otklonili površinski efekti, što znači da kada čestica izađe iz kutije, sa suprotne strane će u kutiju ući njoj identična čestica. No uz periodičke granične uvjete broj interakcija koje se trebaju izračunati bio bi beskonačan. Stoga se koristi konvencija minimalne slike, odnosno zanemaruju se interakcije između čestica i svih slika ostalih čestica izvan određene udaljenosti.<sup>2</sup>

Prema ergodskoj hipotezi prosjek ansambla neke observable jednak je vremenskom prosjeku te observable, koji se može dobiti simulacijama molekulske dinamike. Fazni prostor za sustav od  $N$  atoma je  $6N$  dimenzionalni prostor čije koordinate čine položaji i količine gibanja svih atoma u sva tri smjera prostora.<sup>2</sup> Iako je teško dokazati da je ergodičnost sustava zadovoljena, većina fizikalno zanimljivih sustava ponaša se u skladu s ergodskom hipotezom. Pravi problem je kvazi-neergodičnost koja najčešće proizlazi zbog visokih energetske barijere

koje sustav u ograničenom vremenu simulacije ne prelazi, uslijed čega ne dolazi do uzorkovanja svih dijelova faznog prostora, iako su konfiguracije u skladu sa zakonima očuvanja. Čak i da postoji put u kojima su energetske barijere između dijelova faznog prostora dovoljno niske, ukoliko je taj put jako uzak, tijekom simulacije će se sustav ponašati neergodski. Istražen dio faznog prostora će ovisiti o početnoj konfiguraciji. Moguće je i da je difuzija sustava vrlo spora, pa zbog toga nije zadovoljen uvjet ergodičnosti, ali to je rjeđi slučaj. Dobro uzorkovanje faznog prostora važno je za izračun termodinamičkih veličina poput slobodne energije. Metodama poboljšanog uzorkovanja moguće je bolje istražiti fazni prostor olakšavanjem prijelaza preko visokih energetske barijere ili ubrzavanjem rijetkih događaja poput vezanja liganda u vezno mjesto proteina.<sup>3</sup>

Cilj ovog rada je predstaviti tri metode kojima se poboljšava uzorkovanje faznog prostora: *Umbrella sampling*, metadinamiku i perturbaciju slobodne energije. Za svaku će biti predstavljena teorijska pozadina te strategije za provođenje simulacija i statističku analizu podataka. Na kraju će na primjerima biti ilustrirana primjenjivost metoda ovisno o promatranom procesu.

## § 2. PRIKAZ ODABRANE TEME

### 2.1. Perturbacija slobodne energije

Perturbacija slobodne energije metoda je izračuna razlike slobodne energije između dva sustava, 0 i 1, koja se temelji na uzorkovanju jednog, referentnog sustava (0) i perturbiranju referentnog sustava u željeni sustav (1).<sup>4</sup> Ukoliko je sustav 0, koji se sastoji od  $N$  čestica, opisan Hamiltonijanom  $H_0(\mathbf{x}, \mathbf{p}_x)$ , koji je funkcija  $3N$  prostornih koordinata i  $3N$  količina gibanja čestica, onda se Hamiltonijan sustava 1 može zapisati kao:

$$H_1(\mathbf{x}, \mathbf{p}_x) = H_0(\mathbf{x}, \mathbf{p}_x) + \Delta H(\mathbf{x}, \mathbf{p}_x) \quad (2.1)$$

U kanonskom  $NVT$  ansamblu je razlika Helmholtzove energije dana jednadžbom:

$$\Delta A = -\beta^{-1} \ln \frac{Q_1}{Q_0} = -\beta^{-1} \ln \frac{\iint e^{-\beta H_1(\mathbf{x}, \mathbf{p}_x)} d\mathbf{x} d\mathbf{p}_x}{\iint e^{-\beta H_0(\mathbf{x}, \mathbf{p}_x)} d\mathbf{x} d\mathbf{p}_x} \quad (2.2)$$

gdje su  $Q_1$  i  $Q_0$  kanonske particijske funkcije, a  $\beta = (k_B T)^{-1}$ ,  $k_B$  Boltzmannova konstanta, a  $\iint d\mathbf{x} d\mathbf{p}_x$  označava  $6N$ -struki integral po svim koordinatama i količinama gibanja čestica. Uvrštavanjem (2.1) u (2.2) dobiva se:

$$\Delta A = -\beta^{-1} \ln \iint e^{-\beta \Delta H(\mathbf{x}, \mathbf{p}_x)} P_0(\mathbf{x}, \mathbf{p}_x) d\mathbf{x} d\mathbf{p}_x \quad (2.3)$$

pri čemu je  $P_0(\mathbf{x}, \mathbf{p}_x)$  funkcija gustoće vjerojatnosti da se sustav 0 nađe u mikrostanju određenom koordinatama čestica  $(\mathbf{x}, \mathbf{p}_x)$ :

$$P_0(\mathbf{x}, \mathbf{p}_x) = \frac{e^{-\beta H_0(\mathbf{x}, \mathbf{p}_x)}}{\iint e^{-\beta H_0(\mathbf{x}, \mathbf{p}_x)} d\mathbf{x} d\mathbf{p}_x} \quad (2.4)$$

Jednadžba (2.3) može se zapisati i kao:

$$\Delta A = -\beta^{-1} \ln \langle e^{-\beta \Delta H(\mathbf{x}, \mathbf{p}_x)} \rangle_0 \quad (2.5)$$

gdje  $\langle \dots \rangle_0$  označava prosjek po ansamblu sustava 0, koji je za ergodske sustave jednak vremenskom prosjeku.<sup>3</sup> Jednadžbe (2.3) i (2.5) govore kako je samo potrebno znati vjerojatnost (odnosno gustoću vjerojatnosti) da se sustav nađe u određenom mikrostanju, što se može dobiti provođenjem simulacije molekulske dinamike za sustav 0, te razliku u energiji konfiguracija sustava 0 i 1. Kako se kinetički dio particijske funkcije može analitički riješiti, najčešće se gleda razlika potencijalne energije konfiguracija dvaju sustava,  $\Delta U(\mathbf{x})$ , kako bi se izračunala razlika u slobodnoj energiji:

$$\Delta A = -\beta^{-1} \ln \langle e^{-\beta U(\mathbf{x})} \rangle_0 \quad (2.6)$$

pri čemu sada definirana funkcija gustoće vjerojatnosti da se sustav nađe u konfiguraciji određenoj koordinatama ( $\mathbf{x}$ ) glasi:

$$P_0(\mathbf{x}) = \frac{e^{-\beta U_0(\mathbf{x})}}{\int e^{-\beta U_0(\mathbf{x})} d\mathbf{x}} \quad (2.7)$$

Izračunom razlike potencijalne energiji između konfiguracija referentnog i perturbiranog sustava može se dobiti raspodjela razlika potencijalne energije  $P_0(\Delta U)$  i onda se razlika u slobodnoj energiji može zapisati kao jednostruki integral koji ovisi samo o razlici potencijalne energije:<sup>3</sup>

$$\Delta A = -\beta^{-1} \ln \int e^{-\beta \Delta U} P_0(\Delta U) d\Delta U \quad (2.8)$$

Perturbacijom slobodne energije se u teoriji može izračunati razlika u slobodnoj energiji ukoliko dva sustava imaju isti fazni prostor, odnosno ukoliko se Hamiltonijani kojima su opisani razlikuju za  $\Delta H(\mathbf{x}, \mathbf{p}_\mathbf{x})$ , tj.  $\Delta U(\mathbf{x})$ , kakva god ta perturbacija bila. Također, barem u teoriji nije bitno koji je smjer perturbacije, odnosno pretvara li se sustav 0 u sustav 1 ili sustav 1 u sustav 0. Ukoliko je  $\Delta A$  razlika u slobodnoj energiji kada se sustav 0 pretvori u sustav 1, a  $\Delta U(\mathbf{x})$  razlika potencijalne energije između sustava 1 i 0,  $\Delta U(\mathbf{x}) = U_1(\mathbf{x}) - U_0(\mathbf{x})$ , onda je<sup>4</sup>:

$$e^{-\beta \Delta A} = \langle e^{-\beta \Delta U} \rangle_0 \quad (2.9)$$

$$e^{+\beta \Delta A} = \langle e^{+\beta \Delta U} \rangle_1 \quad (2.10)$$

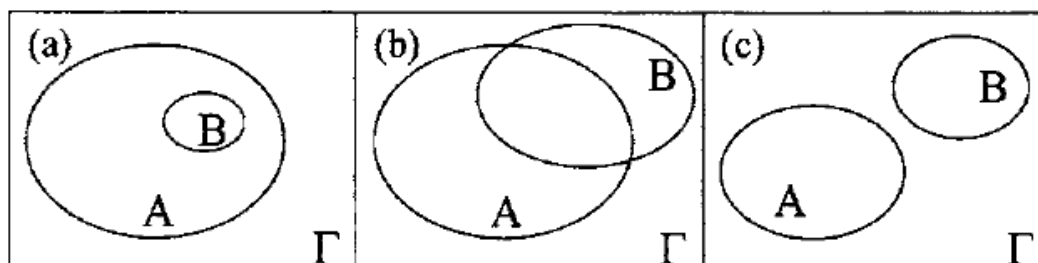
No u praksi dolazi do nepotpune pretrage faznog prostora zbog ograničenosti trajanja simulacije, stoga je potrebno provoditi male perturbacije<sup>3</sup> i pažljivo odabrati smjer perturbacije, odnosno referentni sustav koji će se uzorkovati.<sup>5</sup>

Konačno dugom simulacijom sustav se neće naći u svim konfiguracijama u faznom prostoru  $\Gamma$ . Jednadžba (2.7) govori da u slučaju kanonskog ansambla funkcija gustoće vjerojatnosti da se sustav nađe u određenoj konfiguraciji ovisi samo o energiji te konfiguracije, stoga energija određuje važnost konfiguracije.<sup>5</sup> Za sustav se može definirati „važan“ potprostor faznog prostora,  $\Gamma'$  tako da vrijedi:

$$\frac{\int_{\Gamma'} e^{-\beta U(\mathbf{x})} d\mathbf{x}}{\int_{\Gamma} e^{-\beta U(\mathbf{x})} d\mathbf{x}} = f \quad (2.11)$$

pri čemu je  $f$  neka vrijednost blizu 1. To znači da su tim potprostorom obuhvaćene sve konfiguracije niske energije koje će sustav vjerojatno posjetiti tijekom simulacije i koje će imati veliki doprinos slobodnoj energiji.<sup>5</sup> Ukoliko imamo dva sustava, A i B, koji se mogu

međusobno prevesti jedan u drugi perturbacijom, jedna od stvari koja će odrediti primjerenost korištenja perturbacije slobodne energije za izračun razlike slobodne energije je preklapanje značajnih potprostora faznog prostora. Slika 1. prikazuje tri odnosa značajnih faznih potprostora.



Slika 1. a) Značajni dio faznog prostora sustava B je podskup značajnog dijela faznog prostora sustava A. b) Značajni dijelovi faznog prostora dvaju sustava se djelomično preklapaju. c) Značajni dijelovi faznog prostora sustava A i B se ne preklapaju. (ref. 5)

U jednadžbi (2.3) se vidi da postoje dva bitna doprinosa razlici slobodne energije, funkcija gustoće vjerojatnosti da se neperturbirani sustav nađe u određenoj konfiguraciji te  $e^{-\beta\Delta H(\mathbf{x},\mathbf{p}_\mathbf{x})}$ , tj.  $e^{-\beta\Delta U(\mathbf{x})}$ . Da bi se dobro izračunala razlika slobodne energije, potrebno je da neperturbirani sustav bude dobro uzorkovan u regiji gdje je  $e^{-\beta\Delta U(\mathbf{x})}$  velik, odnosno tamo gdje perturbirani sustav nije puno veće energije od neperturbiranog. Kod slučaja a) na slici 1. postoje dva moguća odabira za smjer perturbacije. Simulacijom sustava A će se perturbacijom dobiti sve važne konfiguracije za sustav B, stoga će rezultati biti točni. Iznimka tomu je kada je značajni fazni potprostor sustava B jako mali dio značajnog potprostora sustava A jer je tada mala vjerojatnost da će se referentni sustav A naći u bitnoj konfiguraciji.<sup>5</sup> U obrnutom smjeru će biti dobiveni precizni, ali netočni rezultati jer uzorkovanjem sustava B neće biti dobivene bitne konfiguracije sustava A. Ukoliko se za slučaj b) na slici 1. perturbacija provodi u jednom koraku, postojat će nepouzdanost u razlici slobodne energije. Smjer perturbacije će biti odabran tako da je perturbirani sustav onaj kod kojeg je manji dio značajnog faznog potprostora leži izvan regije preklapanja. Perturbacija slobodne energije za slučaj c) na slici 1. nije primjenjiva, jer koji god se sustav simulirao, neće biti dovoljno posjećene konfiguracije koje najviše doprinose slobodnoj energiji drugog sustava, a upravo je za te konfiguracije ključno izračunati  $\Delta U(\mathbf{x})$ . Razlika slobodne energije će biti bliska nuli jer će  $e^{-\beta\Delta U(\mathbf{x})}$  biti blizak nuli s obzirom da će se zbrajati doprinosi samo onih konfiguracija kod kojih je potencijalna energija perturbiranog

sustava puno veća, odnosno za koje vrijedi  $\Delta U(\mathbf{x}) \gg 0$ . No moguće je stratifikacijskim metodama konstruirati međustanja u simulacijama, i na taj način poboljšati rezultate za slučajeve b) i c)<sup>2,5</sup>.

Raspodjele razlike potencijalne energije dobivene uzorkovanjem sustava 0,  $P_0(\Delta U)$  i sustava 1,  $P_1(\Delta U)$  nisu neovisne, ali općenito nisu jednake:

$$P_0(\Delta U)e^{\beta\Delta A} = P_1(\Delta U)e^{\beta\Delta U} \quad (2.12)$$

$\Delta U$  i  $\Delta A$  u jednadžbi (2.12) označavaju promjenu pri prelasku iz sustava 0 u sustav 1. Za neki uzorak se mogu odrediti granične razlike potencijalne energije,  $\Delta U_1$  i  $\Delta U_2$ , takve da tijekom uzorkovanja nikad ne dođe do  $\Delta U$  koji je manji od  $\Delta U_1$  ili veći od  $\Delta U_2$ . To odgovara repovima funkcija raspodjele razlike potencijalne energije. Jednadžbe (2.9) i (2.10) se mogu zapisati kao<sup>1</sup>:

$$e^{-\beta\Delta A} = \int_{-\infty}^{\Delta U_1} e^{-\beta\Delta U} P_0(\Delta U) d\Delta U + \int_{\Delta U_1}^{\Delta U_2} e^{-\beta\Delta U} P_0(\Delta U) d\Delta U \quad (2.13)$$

$$+ \int_{\Delta U_2}^{+\infty} e^{-\beta\Delta U} P_0(\Delta U) d\Delta U$$

$$e^{+\beta\Delta A} = \int_{-\infty}^{\Delta U_1} e^{+\beta\Delta U} P_1(\Delta U) d\Delta U + \int_{\Delta U_1}^{\Delta U_2} e^{+\beta\Delta U} P_1(\Delta U) d\Delta U \quad (2.14)$$

$$+ \int_{\Delta U_2}^{+\infty} e^{+\beta\Delta U} P_1(\Delta U) d\Delta U$$

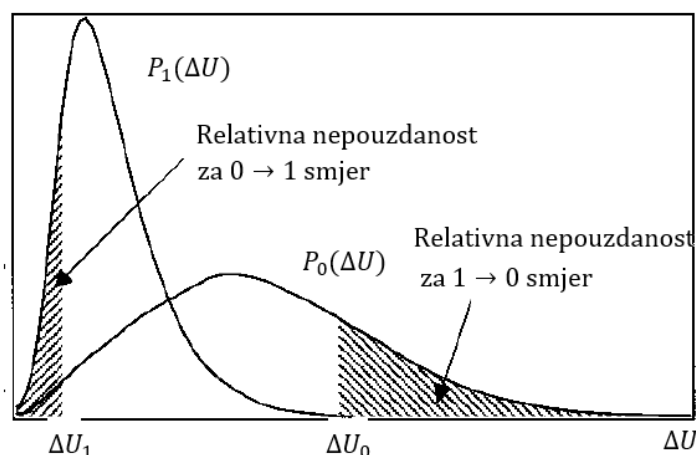
Tijekom simulacije konačne duljine neće doći do konfiguracija takvih da razlika potencijalne energije ne bude unutar intervala  $[\Delta U_1, \Delta U_2]$ , no u stvarnosti će raspodjela razlike potencijalne energije biti mala, ali različita od 0. Iz ovih jednadžbi se vidi da će za transformaciju  $0 \rightarrow 1$  puno bitniji doprinos nepouzdanosti imati interval  $\langle -\infty, \Delta U_1 \rangle$  jer iako je  $P_0(\Delta U)$  blizak nuli u tom intervalu,  $e^{-\beta\Delta U}$  je puno veći nego kada je u intervalu  $\langle \Delta U_2, +\infty \rangle$ . Slično, za transformaciju  $0 \rightarrow 1$  puno bitniji doprinos nepouzdanosti imati interval  $\langle \Delta U_2, +\infty \rangle$ . Na temelju toga su izvedeni izrazi za relativnu nepouzdanost:

$$\delta e_{0 \rightarrow 1} = \frac{e^{-\beta\Delta A^{sim}} - e^{-\beta\Delta A^{točna}}}{e^{-\beta\Delta A^{točna}}} = - \int_{-\infty}^{\Delta U_1} P_1(\Delta U) d\Delta U \quad (2.15)$$



$$\delta e_{1 \rightarrow 0} = \frac{e^{+\beta \Delta A^{sim}} - e^{+\beta \Delta A^{točna}}}{e^{+\beta \Delta A^{točna}}} = - \int_{\Delta U_2}^{+\infty} P_0(\Delta U) d\Delta U \quad (2.16)$$

gdje je  $\Delta A^{sim} = A_1 - A_0$  promjena slobodne energije dobivena simuliranjem sustava 0 i perturbiranjem u sustav 1 za (2.14), odnosno simuliranjem sustava 1 i perturbiranjem u sustav 0 za (2.15).<sup>4</sup> Na slici 2. je grafički prikazano značenje jednadžbi (2.14) i (2.15).



Slika 2. Relativna nepouzdanost  $e^{-\beta \Delta A}$  za  $0 \rightarrow 1$  smjer je jednaka negativnoj površini ispod raspodjele dobivene perturbacijom sustava 1 u sustav 0,  $P_1(\Delta U)$ , s manjom energijom od  $\Delta U_1$ . Relativna nepouzdanost  $e^{+\beta \Delta A}$  za  $1 \rightarrow 0$  smjer je jednaka negativnoj površini ispod  $P_0(\Delta U)$  s većom energijom od  $\Delta U_2$ . (ref. 4)

Da bi se odredila nepouzdanost promjene slobodne energije ukoliko nam nije poznata stvarna vrijednost, potrebno je poznavati  $P_0(\Delta U)$  ili  $P_1(\Delta U)$  te  $\Delta U_1$  ili  $\Delta U_2$ , ovisno o smjeru. Potrebno je procijeniti granične razlike potencijalne energije zato jer će za svaku neovisnu simulaciju jednake duljine biti dobivena različita vrijednost te će često biti precijenjena točnost. Traženjem maksimuma raspodjele  $P(\Delta U_1)$  tj.  $P(\Delta U_2)$  dobivnih velikim brojem ponavljanja neovisnih simulacija jednakih duljina dobivene su formule:

$$\left. \frac{\partial \ln (P_0(\Delta U))}{\partial \Delta U} \right|_{\Delta U = \Delta U_1^*} = M P_0(\Delta U_1^*) \quad (2.17)$$

$$\left. \frac{\partial \ln (P_1(\Delta U))}{\partial \Delta U} \right|_{\Delta U = \Delta U_2^*} = -M P_1(\Delta U_1^*) \quad (2.18)$$

gdje  $\Delta U_1^*$  i  $\Delta U_2^*$  označavaju najvjerojatnije granične razlike potencijalne energije, a M broj neovisnih uzorkovanja raspodjele slobodne energije u simulaciji. Što je veći broj uzoraka, to je  $\Delta U_1^*$  manji, a  $\Delta U_2^*$  veći, odnosno to je veća točnost promjene slobodne energije.<sup>3,4</sup>

Parametar reda,  $\lambda$ , je varijabla koja govori u kojem se stanju sustav nalazi. Može biti geometrijska koordinata, ali i označavati koje su interakcije prisutne, npr. koliki je naboj čestice. Odabran je tako da je za neko međustanje:

$$H(\lambda_i) = \lambda_i H_1 + (1 - \lambda_i) H_0 \quad (2.19)$$

tj.  $\lambda_i = 1$  opisuje sustav 1, a  $\lambda_i = 0$  opisuje sustav 0. Kako je slobodna energija funkcija stanja, njena promjena ovisi samo o početnom i konačnom stanju, a ne kojim je putem sustav došao do konačnog stanja. Promjena slobodne energije se računa tako da se simulira referentno stanje i perturbira u neko međustanje, izračuna se promjena slobodne energije za tu perturbaciju, a potom se simulira to međustanje i perturbira u sljedeće međustanje nakon čega se dobivena promjena slobodne energije pribroji prethodnoj. To se ponavlja sve dok se ne dođe do konačnog, željenog stanja. Za  $i$ -tu perturbaciju (u  $i + 1$  međustanje) je promjena Hamiltonijana:

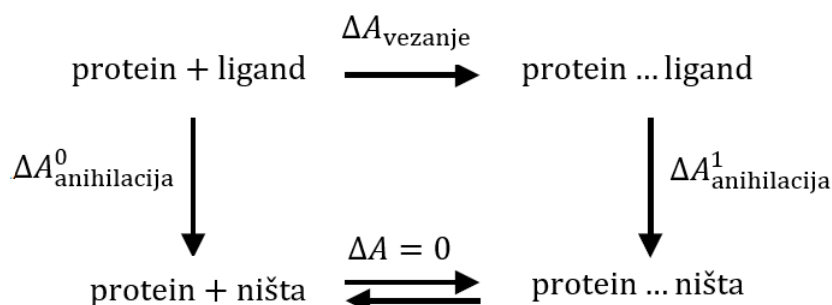
$$\Delta H_i = (\lambda_{i+1} - \lambda_i)(H_1 - H_0) = \Delta \lambda_i \Delta H \quad (2.20)$$

Nije nužno da se izračuna slobodna energija tako da se sustav 0 postupno perturbira u sustav 1. Moguće je i konstruirati neko međustanje M i potom simulirati i sustav 0 i sustav 1, potom ih perturbirati u to međustanje. Ukupna promjena slobodne energije onda se može dobiti iz relacije:

$$\Delta A_{0 \rightarrow 1} = \Delta A_{0 \rightarrow M} - \Delta A_{1 \rightarrow M} \quad (2.21)$$

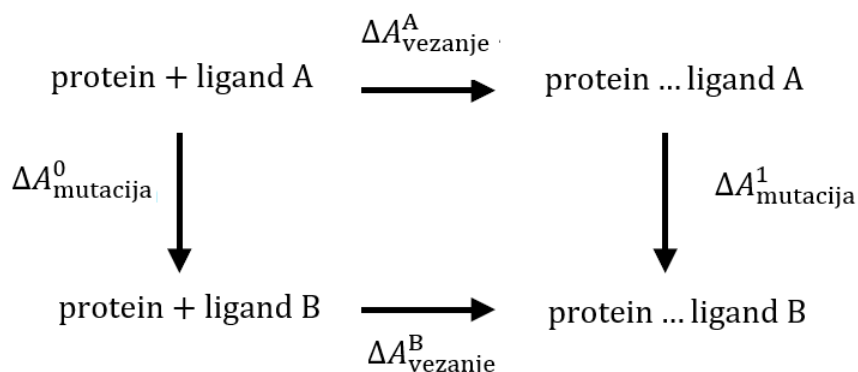
gdje je  $\Delta A_{0 \rightarrow M}$  promjena slobodne energije za perturbaciju sustava 0 u međustanje M, a  $\Delta A_{1 \rightarrow M}$  promjena slobodne energije za perturbaciju sustava 1 u međustanje M.

Termodinamički ciklus je skup reverzibilnih transformacija sustava takvih da se završetkom ciklusa sustav vraća u početno stanje. Koristan je zato jer često postoji niz perturbacija koje su jednostavnije za provesti od željene, a onda izračunom promjene slobodne energije za te korake dobiti slobodnu energiju za željeni korak. Na slici 3. je prikazan primjer termodinamičkog ciklusa.



Slika 3. Termodinamički ciklus za vezanje liganda u vezno mjesto proteina. (ref. 3)

Teško bi bilo definirati parametar  $\lambda$  koji će opisivati prijelazak liganda iz otopine u vezno mjesto, pogotovo s obzirom da su često proteini alosteričke molekule kod kojih dolazi do konformacijskih promjena uslijed vezanja. Alternativni način dobivanja veličine  $\Delta A_{vezanje}$  je simulacijom u kojoj se otopina proteina i liganda perturbira u otopinu u kojoj je prisutan samo protein anihiliranjem liganda. Anihilacija znači da se ligand perturbiranjem postupno pretvara u česticu plina koja dijeli prostor s otopinom, ali više ne interagira s okolinom, to jest s otapalom i proteinom. Zatim se simulira sustav u kojem se ligand nalazi u veznom mjestu otopljenog proteina te ponovno dolazi do anihilacije liganda. No kako bi proces bio reverzibilan, potrebno je ograničiti gibanje liganda u veznom mjestu čak i kad se ponaša kao čestica plina. Za male ligande dovoljno je ograničiti translaciju, a za veće poput benzena potrebno je ograničiti i rotaciju kako bi ključne interakcije s veznim mjestom bile ostvarene u slučaju kad se provodi proces stvaranja liganda, koji je suprotan anihilaciji, u kojem ligand ponovno interagira s okolinom. Donja strelica na slici 3. označava promjenu slobodne energije povezanu sa smanjenjem entropije uslijed ograničenja gibanja liganda u neinteragirajućem stanju.<sup>6</sup> Česti je problem izračun promjene slobodne energije za vezanje različitih liganada u vezno mjesto proteina radi određivanja kandidata za koje je vezanje najjače, primjerice ukoliko se želi dizajnirati neki inhibitor proteina. Na slici 4. je prikazan termodinamički ciklus za takav problem.



Slika 4. Termodinamički ciklus za određivanje promjene slobodne energije za vezanje različitih liganada. (ref. 3)

Razlika slobodnih energija vezanja je  $\Delta A_{vezanje}^B - \Delta A_{vezanje}^A = \Delta A_{mutacija}^1 - \Delta A_{mutacija}^0$ .

## 2.2. Umbrella sampling

*Umbrella sampling* (US) je metoda kojom se poboljšava uzorkovanje faznog prostora. Ukoliko je energetska barijera između dva stanja puno veća od  $k_B T$ , mala je vjerojatnost da će u ograničenom vremenu simulacije sustav dobro istražiti oba stanja. Ukoliko s  $\xi$  označimo reakcijsku koordinatu, tj. parametar reda, funkcija gustoće vjerojatnosti da se sustav nađe u stanju opisanom reakcijskom koordinatom  $\xi$  dana je izrazom:

$$P_n(\xi) = \frac{\iint \delta[\xi(\mathbf{q}) - \xi] e^{-\beta U(\mathbf{q})} d\mathbf{q}}{\iint e^{-\beta U(\mathbf{q})} d\mathbf{q}} \quad (2.22)$$

pri čemu je  $\delta[\xi(\mathbf{q}) - \xi]$  Diracova delta funkcija koja iznosi 0 ukoliko je  $\xi(\mathbf{q}) - \xi$  različito od 0, tj. pribrajaju se samo Boltzmannovi faktori onih konfiguracija  $\mathbf{q}$  koje su opisane parametrom  $\xi$ .<sup>3,7</sup> Ukoliko se potencijalnoj energiji doda dodatni član koji ovisi samo o koordinati  $\xi(\mathbf{q})$ ,  $\omega_i(\xi(\mathbf{q}))$ , dobit ćemo pristranu funkciju gustoće vjerojatnosti,  $P'_i(\xi)$ , da se sustav nađe u stanju opisanom koordinatom  $\xi$ :

$$P'_i(\xi) = \frac{\iint \delta[\xi(\mathbf{q}) - \xi] e^{-\beta[U(\mathbf{q}) + \omega_i(\xi(\mathbf{q}))]} d\mathbf{q}}{\iint e^{-\beta[U(\mathbf{q}) + \omega_i(\xi(\mathbf{q}))]} d\mathbf{q}} \quad (2.23)$$

S obzirom da potencijal ovisi o  $\xi(\mathbf{q})$ , a u brojniku se pribrajaju samo konfiguracije kod kojih je  $\xi(\mathbf{q}) = \xi$ , to se može zapisati kao:

$$P'_i(\xi) = e^{-\beta \omega_i(\xi)} \frac{\iint \delta[\xi(\mathbf{q}) - \xi] e^{-\beta U(\mathbf{q})} d\mathbf{q}}{\iint e^{-\beta[U(\mathbf{q}) + \omega_i(\xi(\mathbf{q}))]} d\mathbf{q}} \quad (2.24)$$

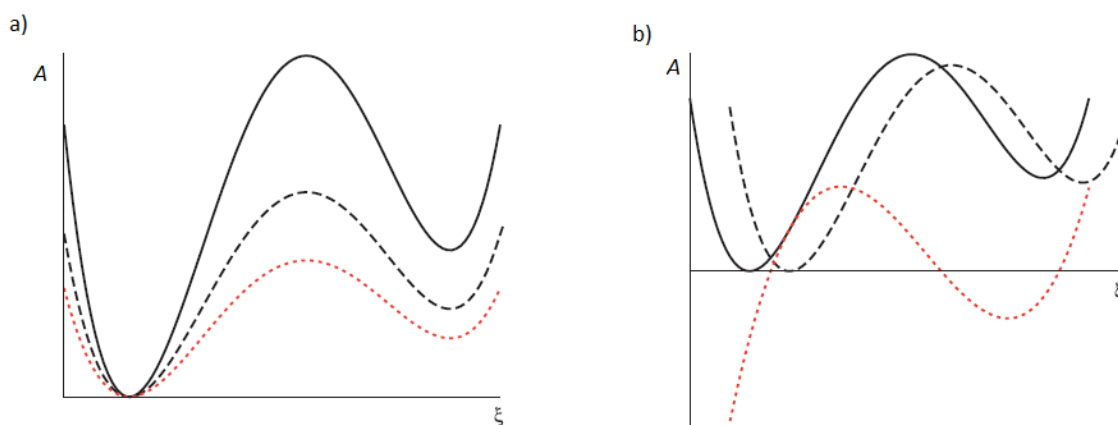
Iz jednadžbi (2.22) i (2.24) slijedi da se prava funkcija gustoće vjerojatnosti može zapisati kao:

$$\begin{aligned} P_n(\xi) &= P'_i(\xi) e^{\beta \omega_i(\xi)} \frac{\iint e^{-\beta \omega_i(\xi(\mathbf{q}))} \cdot e^{-\beta U(\mathbf{q})} d\mathbf{q}}{\iint e^{-\beta U(\mathbf{q})} d\mathbf{q}} = \\ &= P'_i(\xi) e^{\beta \omega_i(\xi)} \langle e^{-\beta \omega_i(\xi)} \rangle_n \end{aligned} \quad (2.25)$$

Pri čemu je  $\langle e^{-\beta \omega_i(\xi)} \rangle_n$  prosječna vrijednost  $e^{-\beta \omega_i(\xi)}$  po ansamblu za nepristrani sustav. Izraz za Helmolzovu energiju u ovisnosti o reakcijskoj koordinati glasi:

$$\begin{aligned} A_i(\xi) &= -\beta^{-1} \ln(P'_i(\xi)) - \omega_i(\xi) - \beta^{-1} \ln(\langle e^{-\beta \omega_i(\xi)} \rangle_n) = \\ &= -\beta^{-1} \ln(P'_i(\xi)) - \omega_i(\xi) + F_i \end{aligned} \quad (2.26)$$

Budući da je  $F_i$  konstanta, ukoliko se odabere takav potencijal  $\omega_i(\xi)$  da je suprotan vrijednosti slobodne energije za nepristrani sustav, onda će  $P'_i(\xi)$  biti konstanta, stoga će sve konfiguracije koje odgovaraju različitim vrijednostima reakcijske koordinate biti jednako posjećivane. No to bi zahtijevalo prethodno poznavanje plohe slobodne energije, što nije uvijek realistično.



Slika 5. Prava slobodna energija (ispunjena linija), vanjski potencijal  $\omega$  (iscrtkana linija) i slobodna energija nakon djelovanja vanjskim potencijalom (točkasta linija) u ovisnosti o reakcijskoj koordinati. a) Vanjski potencijal nije jednak krivulji slobodne energije, ali njegovim djelovanjem se snižavaju energijske barijere, odnosno krivulja slobodne energije nakon djelovanja vanjskim potencijalom postaje ravnija b) Iako potencijal ima isti oblik kao slobodna energija, budući da je malo pomaknut duž reakcijske koordinate u odnosu na pravu slobodnu energiju, ne postiže se poravnavanje profila slobodne energije nakon djelovanja vanjskim potencijalom, pa se ne očekuje poboljšanje rezultata u odnosu na simulaciju sustava gdje se ne dodaje vanjski potencijal (ref. 3)

Jedan od načina da se poboljša uzorkovanje je korištenje potencijala čiji se oblik iterativno poboljšava tako da se dobije što je moguće uniformnija funkcija gustoće vjerojatnosti da će u pristranoj simulaciji biti posjećena konfiguracija s koordinatom  $\xi$ , što je glavna ideja adaptivne *Umbrella sampling* metode. Drugi pristup je korištenje harmonijskih potencijala u više prozora  $i$ , tj. simulacija u kojima je sustav podvrgnut potencijalu oblika  $\omega_i(\xi) = \frac{K}{2}(\xi - \xi_i)^2$ . Potom se rezultati tih simulacija kombiniraju da bi se dobila ukupna razlika slobodne energije.<sup>7</sup>

Da bi se kombinirali rezultati prikupljeni u različitim prozorima, potrebno je izračunati  $F_i$ , no za to je potrebno poznavati  $\langle e^{-\beta\omega_i(\xi)} \rangle_n$ , prosječnu vrijednost  $e^{-\beta\omega_i(\xi)}$  po ansamblu za nepristrani sustav, a US se koristi zato što se običnim simulacijama ne može postići dovoljno dobro uzorkovanje. Odabir metode kojom će se kombinirati slobodne energije u ovisnosti o reakcijskim koordinatama je proizvoljan. Analiza nepouzdanosti, ali i način stratifikacije razlikuju se ovisno o metodi koja se koristi za kombiniranje rezultata iz različitih prozora.

Budući da je slobodna energija u ovisnosti o reakcijskoj koordinati kontinuirana funkcija, moguće je preklapanjem nepristranih profila slobodne energije dobivenih u različitim prozorima rekonstruirati profil slobodne energije duž reakcijske koordinate uz dodanu proizvoljnu koordinatu, koja je posljedica nemogućnosti izravnog određivanja  $F_l$  u bilo kojem od prozora.<sup>7,8</sup> No problem je što je potrebno značajno preklapanje prozora kako bi se smanjile nepouzdanosti, odnosno veliki dio prikupljenih podataka služi samo tome da se smanji nepouzdanost.<sup>7</sup> Zato su razvijene metode kod kojih je potrebno manje preklapanje, poput metode analize vaganim histogramima (*Weighted histogram analysis method*, WHAM) dok za *Umbrella* integraciju preklapanje nije potrebno, ali je poželjno budući da se time smanjuje nepouzdanost.<sup>7,10</sup> WHAM metoda kombinira rezultate tako da se ukupna, nepristrana funkcija raspodjele reakcijske koordinate  $\xi$  zapiše kao linearna kombinacija nepristranih funkcija raspodjele dobivenih u svim prozorima, a potom se minimizira varijanca ukupne raspodjele s obzirom na koeficijente u linearnoj kombinaciji.<sup>7</sup> Vjerojatnost da se  $\xi$  nađe u  $l$ -toj košari histograma koji ima  $M$  košara kada nije primijenjen vanjski potencijal je dana jednadžbom:

$$P_l = \frac{\sum_{i=1}^S n_{il}}{\sum_{i=1}^S N_i e^{-\beta \omega_i(\xi_l)} f_i} \quad (2.27)$$

pri čemu  $\sum_{i=1}^S n_{il}$  sumira broj puta koliko se sustav našao u stanju opisanom s  $\xi_l$  u svim prozorima od  $i = 1$  do  $S$ ,  $N_i$  predstavlja ukupan broj konfiguracija u  $i$ -tom prozoru, a  $f_i$  je dan jednadžbom:

$$f_i = \frac{1}{\sum_{l=1}^M e^{-\beta \omega_i(\xi_l)} P_l} \quad (2.28)$$

što je samo inverz diskretnog zapisa  $\langle e^{-\beta \omega_i(\xi)} \rangle_n$  u jednadžbi (2.25). Jednadžbe (2.27) i (2.28) mogu se iterativno rješavati tako da se pretpostavi skup  $\{f_i\}$  u svim prozorima, koji služe za izračun skupa nepristranih vjerojatnosti da će reakcijska koordinata biti u određenoj košari histograma,  $\{P_l\}$ , a zatim se dobiveni skup  $\{P_l\}$  koristi za novi skup  $\{f_i\}$  i taj se postupak ponavlja dok se ne postigne konvergencija. Problem takvog rješavanja je što razlika slobodne energije između iteracija nije dobar pokazatelj odstupanja od točnog rješenja, zbog čega su osmišljeni točniji i brži načini rješavanja sustava jednadžbi.<sup>11</sup>

Statistička nepouzdanost dobivene raspodjele iz jednadžbe (2.27) pretežito proizlazi iz nepoznavanja  $f_i$ . Samo histogrami susjednih prozora imaju značajno preklapanje, stoga je za udaljene prozore nepouzdanost kumulativna. Nepouzdanost se može procijeniti u slučaju da je

širina košara histograma  $\Delta\xi$  konstantna te ako se koriste harmonijski potencijali  $\frac{K}{2}(\xi - \xi_i)$  centrirani na mjestu  $\xi_i = \xi_0 + i\Delta\xi$  pomoću jednadžbe:

$$\text{var}(A(\xi)) \approx (K\Delta\xi)^2 \cdot \sum_{i=1}^{\frac{\xi-\xi_0}{\Delta\xi}} \text{var}(\bar{\xi}_i) \quad (2.29)$$

pri čemu je  $\text{var}(\bar{\xi}_i)$  kvadrat nepouzdanosti srednje vrijednosti reakcijske koordinate  $\xi$  u  $i$ -tom prozoru.<sup>11</sup> Budući da vanjski potencijal djeluje samo na reakcijsku koordinatu  $\xi$ , ostali će procesi biti jednako spori kao i u normalnoj simulaciji.<sup>8</sup> Član  $\text{var}(\bar{\xi}_i)$  će biti znatno podcijenjen ukoliko vrijeme simulacije nije dovoljno dugo da se odvijaju ti spori procesi. Visoke energetske barijere mogu postojati i duž koordinata ortogonalnih na  $\xi$ . Tada će u ekstremnim slučajevima sustav ostati u samo jednom metastabilnom stanju opisanom ortogonalnim koordinatama, pa se iz simulacije neće moći vidjeti da uzorkovanje nije dovoljno dobro, ali će nepouzdanost slobodne energije biti velika. Da je unesena sustavna nepouzdanost neadekvatnim uzorkovanjem u ortogonalnim koordinatama može se vidjeti tako da se usporede rezultati simulacija u kojima je vanjski potencijal primijenjen samo na  $\xi$  te kada vanjski potencijal djeluje i na neku od ortogonalnih koordinata.<sup>11</sup> Još jedan uzrok nepouzdanosti je sama diskretizacija, odnosno korištenje histograma. Pretpostavka je da se slobodna energija u nekoj točki može aproksimirati srednjom slobodnom energijom u intervalu oko te točke. Uz dovoljno mali interval, tj. dovoljno usku košaru, slobodna će energija unutar intervala biti linearna, no i u tom će se slučaju funkcija gustoće vjerojatnosti da se sustav nađe u konfiguraciji opisanom s  $\xi$  mijenjati eksponencijalno. Zbog toga će srednja vrijednost raspodjele vjerojatnosti u centru intervala biti veća nego stvarna vrijednost, pa će slobodna energija biti sustavno podcijenjena. Smanjivanjem intervala smanjuje se sustavna nepouzdanost, no povećava se statistička nepouzdanost. Procjenom tih nepouzdanosti može se odabrati optimalna širina košara u histogramu.<sup>12</sup>

U adaptivnoj US metodi provodi se  $i$  simulacija s  $N_i$  koraka, u prvoj od kojih je vanjski potencijal  $\omega_1(\xi_1, \dots, \xi_L) = 0$  nakon čega se radi histogram reakcijskih koordinata. Vjerojatnosti u trenutnom prozoru  $P_{k,\dots,l}$  da se sustav nađe u konfiguraciji u kojoj je  $\xi_1$  opisan s  $\xi_k, \dots, \xi_L$  s  $\xi_l$  dobiju se kombinacijom svih prikupljenih podataka WHAM metodom (jednadžbe (2.27) i (2.28), uz modifikaciju da se  $\sum_{l=1}^M e^{-\beta\omega_l(\xi_l)} P_l$  u  $f_i$  zamijeni s izrazom za slučaj kada vanjski potencijal djeluje na više reakcijskih koordinata,  $\sum_{l=1}^{M_l} \dots \sum_{k=1}^{M_k} e^{-\beta\omega_l(\xi_k, \dots, \xi_l)} P_{k,\dots,l}$ ). Iz skupa

vjerojatnosti  $\{P_{k,\dots,l}\}$  se uz ekstrapolaciju dobije diskretna verzija vanjskog potencijala,  $\{\omega_{i+1}(\xi_k, \dots, \xi_l)\}$  koji će se koristiti za sljedeću simulaciju. Ekstrapolacija se radi kako bi se u sljedećem prozoru bolje uzorkovale neistražene regije prostora reakcijskih koordinata. Važno je i da potencijal bude dovoljno glatka funkcija kako ne bi nastali lokalni minimumi kada se iz diskretne verzije napravi kontinuirani vanjski potencijal. Kontinuirani vanjski potencijal se može zapisati kao linearna kombinacija produkata polinoma ili trigonometrijskih funkcija koje čine L baza vektorskih prostora ukoliko vanjski potencijal djeluje na L reakcijskih koordinata:

$$\omega_{i+1}(\xi_1, \dots, \xi_L) = \sum_{\alpha_1=1}^{D_1} \dots \sum_{\alpha_L=1}^{D_L} a_{\alpha_1 \dots \alpha_L} f_{\alpha_1}^{(1)}(\xi_1) \cdot \dots \cdot f_{\alpha_L}^{(L)}(\xi_L) \quad (2.30)$$

pri čemu je za svaki  $\xi_N, N = 1, 2, \dots, L$  definiran linearno nezavisan skup od  $D_N$  polinoma ili trigonometrijskih funkcija  $\{f_1^{(N)}(\xi_N), \dots, f_{D_N}^{(N)}(\xi_N)\}$ . Koeficijenti u linearnoj kombinaciji se određuju metodom najmanjih kvadrata tako da se što vjernije reproducira diskretni vanjski potencijal,  $\{\omega_{i+1}(\xi_k, \dots, \xi_l)\}$ . Kada se dobije  $\omega_{i+1}(\xi_1, \dots, \xi_L)$ , njime se djeluje na reakcijske koordinate u idućem prozoru, nakon čega se ponavlja postupak obrade podataka sve dok se ne postigne dobro uzorkovanje željenog dijela konformacijskog prostora.<sup>13</sup>

### 2.3. Metadinamika

Metadinamika je metoda u kojoj se, slično kao kod US metode, Hamiltonianu dodaje dodatni član, tj. potencijal koji djeluje samo na određene stupnjeve slobode koji se nazivaju kolektivne varijable.<sup>14</sup> Ova metoda može se smatrati posebnim slučajem adaptivne US metode kada je vrijeme simulacija između ažuriranja potencijalne energije toliko kratko da se kolektivna varijabla ne mijenja.<sup>15</sup> Na kolektivne varijable  $S = (s_1, \dots, s_{N_{CV}})$  djeluje vanjski potencijal koji ovisi o prethodnim vrijednostima i može se zapisati kao suma Gaussijana:

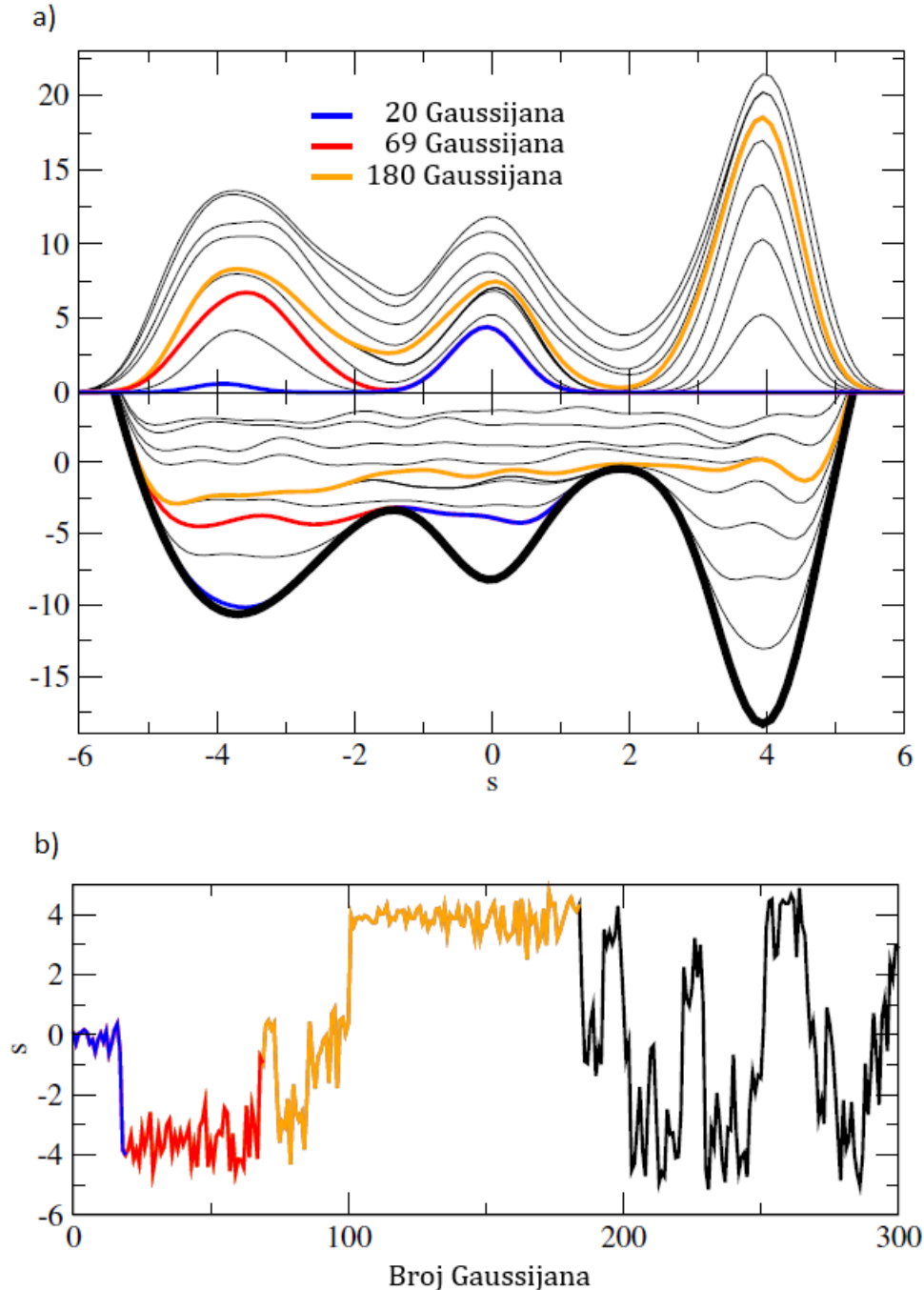
$$V_G(S, t) = \sum_{i=1}^{t/\tau_G} w_G e^{-\sum_{\alpha=1}^{N_{CV}} \frac{(s_\alpha - s_\alpha(i\tau_G))^2}{2\sigma_\alpha^2}} \quad (2.31)$$

$w_G$  je visina Gaussijana koji se dodaje potencijalnoj energiji, a  $\sigma_\alpha$  je širina Gaussijana duž kolektivne varijable  $s_\alpha$ . Gaussijani su centrirani u kolektivnoj varijabli u vremenu  $t = i\tau_G$ ,  $s_\alpha(i\tau_G)$  i na taj način se otežava ponovno posjećivanje konfiguracije koja je opisana istom kolektivnom varijablom.<sup>16,17</sup> Ukoliko je  $\tau_G$  dovoljno mali, umjesto (2.31) pravilo ažuriranja vanjskog potencijala može se zapisati preko vremenske derivacije tog izraza:



$$\dot{V}_G(S, t) = \frac{W_G}{\tau_G} e^{-\sum_{\alpha=1}^{N_{CV}} \frac{(s_\alpha - s_\alpha(i\tau_G))^2}{2\sigma_\alpha^2}} \quad (2.32)$$

Na slici 6. je prikazan učinak dodavanja vanjskog potencijala na kolektivnu varijablu.



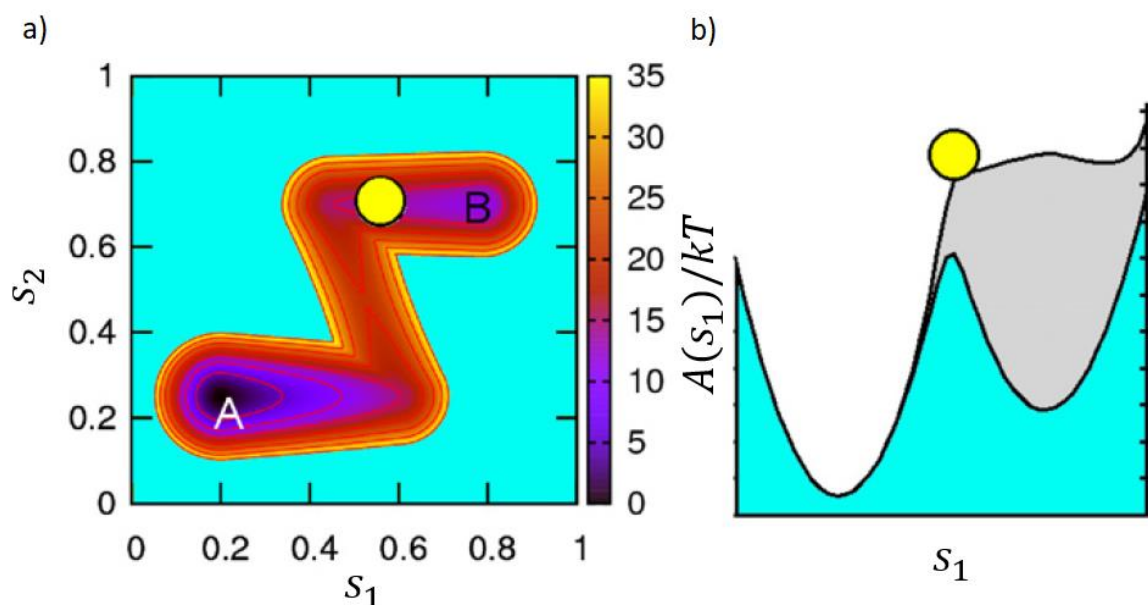
Slika 6. a) Gore: Izolinije prikazuju vanjski (dodani) potencijal u ovisnosti o kolektivnoj varijabli  $s$  pri različitim vremenima, tj. nakon određenog broja dodanih Gaussijana. Dolje: Izolinije prikazuju ukupni potencijal, tj. zbroj vanjskog potencijala i originalnog potencijala

(podebljana crna linija). b) Trajektorija sustava pod utjecajem ukupnog potencijala prikazanog pod a) (ref. 17)

Nakon što se doda 20 Gaussijana (plava linija), popuni se prva udubina duž kolektivne varijable u plohi potencijalne energije, te se počnu istraživati konfiguracije sljedećeg minimuma, što se očituje padom vrijednosti kolektivne varijable na vrijednost od oko  $-4$  (slika 6. b)). Nakon što se doda 69 Gaussijana (crvena linija) popuni se i drugi minimum i podjednako se posjećuju konfiguracije u rasponu vrijednosti kolektivnih varijabli od  $-5$  do  $1$ , sve dok se ne doda nešto više od 100 Gaussijana uslijed čega sustav prijeđe u treći minimum. Taj se minimum popunjava sve dok se ne doda 180 Gaussijana, a nakon toga se sve konfiguracije podjednako posjećuju. Kada vanjski potencijal popuni sve minimume slobodne energije, najbolja aproksimacija plohe slobodne energije u ovisnosti o kolektivnim varijablama je negativni vremenski prosjek vanjskog potencijala:

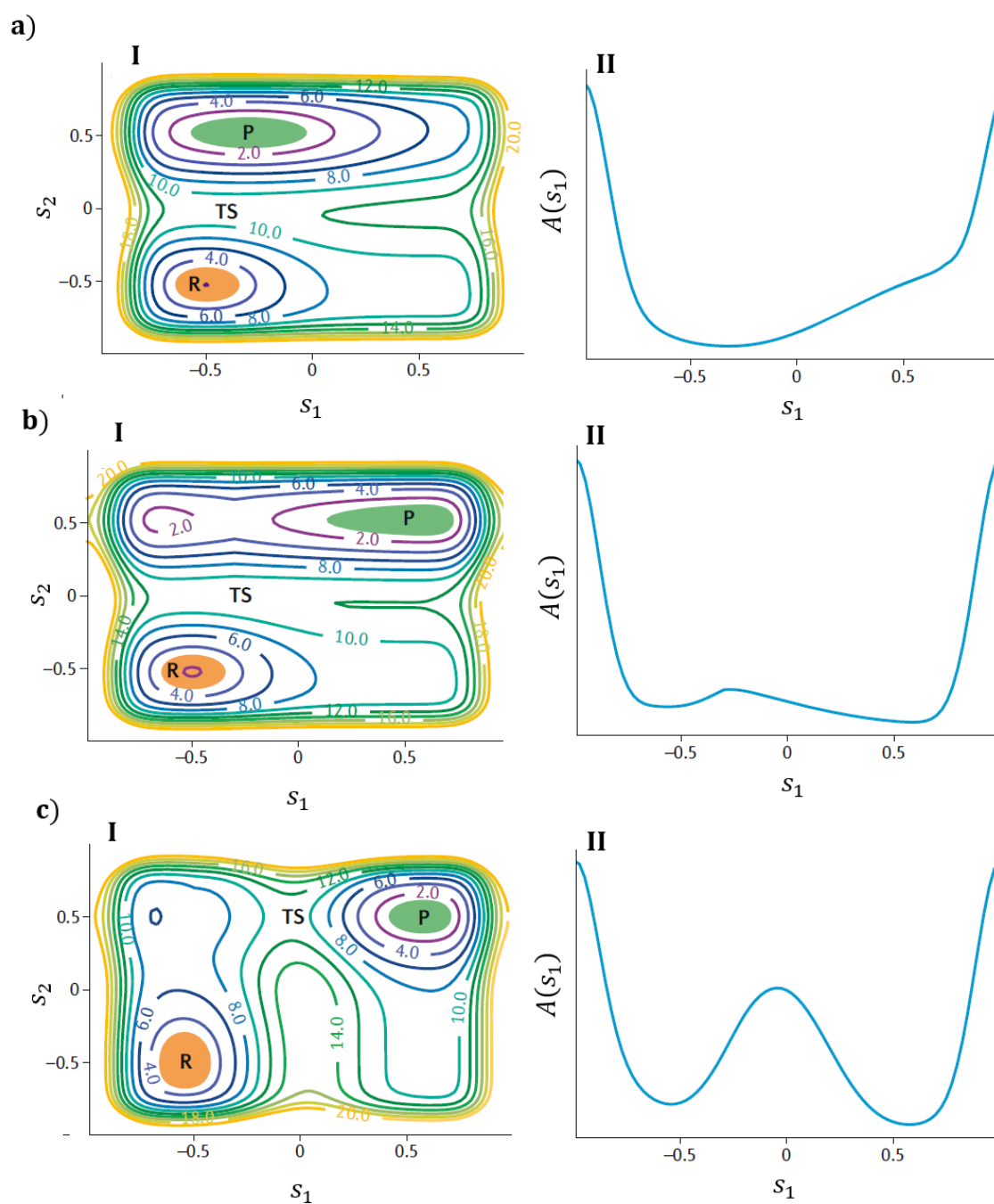
$$A(S) \approx -\frac{1}{t - t_p} \sum_{t_p}^t V_G(S, t) \quad (2.33)$$

pri čemu je  $t_p$  vrijeme potrebno da se popune svi minimumi, nakon čega će dinamika postati difuzivna. U limesu kada  $t$  teži u beskonačnost slobodna energija je jednaka negativnom vremenskom prosjeku vanjskog potencijala ukoliko je dinamika u prostoru kolektivnih varijabli adijabatski raspregnuta od dinamike u ostalim varijablama (ali u praksi taj uvjet nije uvijek nužan).<sup>15</sup> Ukoliko se ne primjenjuje potencijal na sve spore varijable, nakon punjenja oba minimuma dinamika neće postati difuzivna.<sup>15,17</sup> Takav slučaj je prikazan na slici 7.



Slika 7. Nivo-krivulje potencijalne energije u ovisnosti o kolektivnim varijablama  $s_1$  i  $s_2$   
 b) Profil slobodne energije duž varijable  $s_1$  te generirani vanjski potencijal (sivo) (ref. 17)

Na slici 7. a) se vidi da između stanja A i B postoji energetska barijera i duž kolektivnih varijable  $s_1$  i duž varijable  $s_2$ . Ukoliko se samo primjenjuje vanjski potencijal duž varijable  $s_1$ , čak i nakon što se popuni minimum B, neće biti vjerojatan prelazak u stanje A budući da postoji i barijera duž varijable  $s_2$ , što se vidi na slici 7. b). Ukoliko sustav i prijeđe u stanje A, nakon što se i taj minimum popuni neće biti vjerojatan prelazak u stanje B, stoga će se naizmjenično prepunjavati minimumi, a vanjski potencijal neće konvergirati u negativnu vrijednost slobodne energije. No vrijeme potrebno da se popuni minimum duž kolektivne varijable  $s_\alpha$  je proporcionalno  $(\sigma_\alpha)^{-N_{CV}}$  pri čemu je  $\sigma_\alpha$  širina Gaussijana, a  $N_{CV}$  broj kolektivnih varijabli.<sup>17</sup> Razlog tomu je to što je mala vjerojatnost da će biti posjećena konfiguracija opisana istim skupom kolektivnih varijabli ukoliko je njihov broj velik.<sup>16,17</sup> To znači da je za puno kolektivnih varijabli potrebno koristiti široke Gaussijane, čiji je nedostatak da ne mogu dobro opisati detalje plohe slobodne energije na skalama manjim od njihove širine. Stoga je za dobivanje kvalitetne slobodne energije potrebno vanjskim potencijalom djelovati na što manji broj kolektivnih varijabli.<sup>17</sup> Kolektivne varijable bi trebale i jasno razlikovati sva metastabilna, ali i prijelazna stanja,<sup>15,17</sup> što je prikazano na slici 9.



Slika 8. **I** Nivo-krivulje potencijalne energije u ovisnosti o kolektivnim varijablama  $s_1$  i  $s_2$  i **II** profili slobodne energije duž varijable  $s_1$  kada vanjski potencijal djeluje samo na  $s_1$  kada **a)**  $s_1$  ne razlikuje reaktante R, produkte P ni prijelazno stanje TS **b)**  $s_1$  ne razlikuje reaktante R, i prijelazno stanje TS, ali od njih razlikuje produkte P **c)**  $s_1$  jasno razlikuje reaktante R, produkte P i prijelazno stanje TS (ref. 15)

Na slici 8. a) (I) je prikazan slučaj kad reaktant (R), prijelazno stanje (TS) i produkt (P) imaju sličnu vrijednost kolektivne varijable  $s_1$ . Tada će pripadajući profil slobodne energije (slika 8. a) (II)) imati samo jedan minimum. Ukoliko reaktant i produkt imaju različite vrijednosti za kolektivnu varijablu  $s_1$ , ali reaktant posjeduje i prijelazno stanje slične vrijednosti (slika 8. b) (I)) profil slobodne energije će imati dva minimuma, ali prijelazno stanje neće biti dobro opisano (slika 9. b) (II)). U tom slučaju će prijelazi sustava iz reaktanta u produkt biti rijetki. Ako reaktant, prijelazno stanje i produkt imaju dovoljno različite vrijednosti za kolektivnu varijablu  $s_1$  (slika 8. c) (I)) profil slobodne energije će imati dva minimuma te sedlastu točku za prijelazno stanje, tj. dobro će opisivati reakciju (slika 8. c) (II)).

No čak i kada su dobro odabrane kolektivne varijable, u klasičnoj metadinamici postoje problemi s konvergencijom. U jednoj simulaciji slobodna energija neće potpuno konvergirati, nego će njena vrijednost fluktuirati oko točne vrijednosti. Da bi se poboljšao rezultat, trebao bi se smanjiti omjer visine Gaussijana i vremenskog koraka između dodavanja Gaussijana potencijalu,  $w_G/\tau_G$ . Bez obzira povećá li se interval između dodavanja Gaussijana ili smanji li se visina Gaussijana, produljit će se popunjavanje minimuma na plohi potencijalne energije. Također, ukoliko predugo traje simulacija, može se dogoditi da se sustav pogurne u minimume koji nisu fizikalno značajni.<sup>18</sup> *Well-Tempered* metadinamika te je probleme riješila tako da se potencijal ažurira prema jednadžbi:

$$\dot{V}_G(S, t) = \frac{w_G}{\tau_G} e^{-\frac{V(S(t), t)}{k_B \Delta T}} e^{-\sum_{\alpha=1}^{N_{CV}} \frac{(s_\alpha - s_\alpha(t \tau_G))^2}{2\sigma_\alpha^2}} \quad (2.34)$$

gdje je  $\dot{V}_G(S, t)$  vremenska derivacija vanjskog potencijala, a  $\Delta T$  je parametar s mjernom jedinicom temperature. Svaki Gaussijan koji se u trenutku  $t$  stavi na mjesto  $S(t)$  skaliran je s  $e^{-\frac{V(S(t), t)}{k_B \Delta T}}$  i na taj način je visina Gaussijana niža na mjestima gdje je već dodan veliki vanjski potencijal.<sup>16,18</sup> U limesu kada  $t$  teži u beskonačnost, vanjski je potencijal jednak udjelu slobodne energije (uz određenu konstantu):

$$V_G(S, t) = \frac{-\Delta T}{T + \Delta T} A(S) + C \quad (2.35)$$

odnosno, ta funkcija više neće konvergirati u negativnu vrijednost slobodne energije. Tada je sustav uravnotežen, budući da se  $V_G$  ne mijenja, te je raspodjela kolektivnih varijabli:

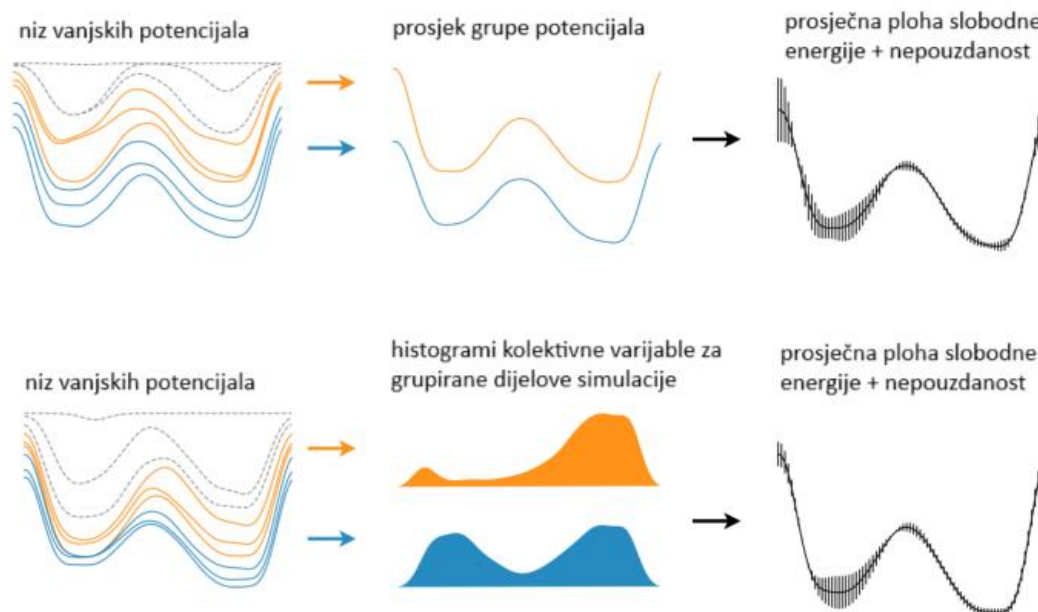
$$P(S) \propto e^{-\frac{A(S) + V(S, t \rightarrow \infty)}{kT}} \propto e^{-\frac{A(S)}{k(T + \Delta T)}} \quad (2.36)$$

što je izraz za kanonsku raspodjelu kolektivnih varijabli pri temperaturi  $T + \Delta T$ . Ukoliko je  $\Delta T = 0$  onda je  $V_G(S, t) = 0$ , a raspodjela kolektivnih varijabli je jednaka raspodjeli koja se dobije kada se ne koristi vanjski potencijal. Ukoliko  $\Delta T$  teži u beskonačnost,  $V_G(S, t) = -A(S)$ , odnosno rezultat je ekvivalentan onome koji se dobije kada se promatra prosječna vrijednost potencijalne energije iz ansambla simulacija obične metadinamike.<sup>16,18</sup> Ostale vrijednosti  $\Delta T$  određuju efektivnu temperaturu pri kojoj će se istraživati prostor kolektivnih varijabli, što znači da  $\Delta T$  određuje koji će dio plohe slobodne energije biti istražen. To znači da se može smanjiti statistička nepouzdanost povećanjem vremena simulacije bez da se sustav pogurne u minimume koji nisu fizikalno značajni.<sup>18</sup> Primjer takvog problema je potpuno narušavanje sekundarne strukture proteina ukoliko se promatra njegova konformacijska promjena. No u standardnoj metadinamici moguće je nariniti dodatni potencijal koji će držati slobodne varijable u određenom rasponu. Na granicama prostora kolektivnih varijabli (koje se mogu i prirodno pojaviti, primjerice koordinacijski broj ne može biti negativan) pojavljuju se sustavne nepouzdanosti zbog diskontinuiteta plohe slobodne energije. Nepouzdanosti se povećavaju s trajanjem simulacije. Stoga se može postaviti uvjet da se izvan raspona kolektivne varijable,  $s'$ , potencijalna energija ne mijenja, tj.:

$$\left. \frac{dV_G(s, t)}{ds} \right|_{s \notin s'} = 0 \quad (2.37)$$

Ovo funkcionira samo ako se radi o jednoj kolektivnoj varijabli  $s$ , umjesto o skupu njih,  $S$ . Također, i dalje se dodaju Gaussijani iako je kolektivna varijabla izvan intervala  $s'$  budući da njihovi „repovi“ doprinose vrijednosti  $V_G(s, t)$  unutar intervala.<sup>19</sup>

Analiza nepouzdanosti se razlikuje za standardnu i *Well-Tempered* metadinamiku. Procedura je prikazana na slici 9.



Slika 9. a) Procedura za izračun nepouzdanosti u standardnoj metadinamici. Plavom i narančastom bojom su obilježeni potencijali koji su grupirani, a sivom iscrtkanom linijom potencijali na početku simulacije koji se ne koriste za izračun prosjeka ni nepouzdanosti. b) Procedura za izračun nepouzdanosti u *Well-Tempered* metadinamici. Plavom i narančastom bojom su obilježeni potencijali iz dijelova simulacija koji su grupirani, te histogrami kolektivne varijable dobiveni u tim grupama, a sivom iscrtkanom linijom potencijali na početku simulacije koji se ne koriste za izračun prosjeka ni nepouzdanosti. (ref. 15)

U standardnoj metadinamici se odbace vanjski potencijali na početku simulacije (prikazani sivom iscrtkanom linijom na slici 9. a)), a zatim se vremenski korelirani potencijali grupiraju i izračuna se srednja vrijednost grupa. Za rekonstrukciju plohe slobodne energije i izračun nepouzdanosti koriste se prosjeci grupa. U *Well-Tempered* metadinamici smanjuje se visina Gaussijana, što se na slici 9. b) vidi i iz toga što vanjski potencijali postaju sve grupiraniji i sličniji s vremenom simulacije. Stoga se za izračun nepouzdanosti ne koriste sami potencijali, nego histogrami kolektivnih varijabli na koje djeluje vanjski potencijal. Histogrami se grupiraju i iz grupa se izračuna nepouzdanost histograma, koja se onda može preračunati u nepouzdanost plohe slobodne energije.<sup>15</sup> Nepouzdanost se može procijeniti i prije simulacije na temelju parametara potencijala. Izveden je izraz za procjenu nepouzdanosti na temelju pretpostavke da potencijal evoluira pod utjecajem Langevinove dinamike:

$$\bar{\epsilon} = C(d) \sqrt{\frac{S \cdot \delta s \cdot w_G}{D \cdot \beta \cdot \tau_G}} \quad (2.38)$$

pri čemu je  $\delta s$  širina, a  $w_G$  visina Gaussijana,  $\tau_G$  vrijeme između dodavanja Gaussijana,  $D$  difuzijski koeficijent,  $\beta = (k_B T)^{-1}$ , a  $S$  je korijen veličine prostora kolektivnih varijabli.  $C(d)$  ovisi o dimenzionalnosti prostora kolektivnih varijabli.  $D$ ,  $\beta$  i  $S$  su svojstva sustava, a prilikom postavljanja simulacije biraju se širina i visina te vrijeme između dodavanja Gaussijana.  $\tau_G$  i  $w_G$  nisu nezavisni te će nepouzdanost ovisiti o njihovom omjeru  $\omega = w_G/\tau_G$ . No ukoliko su  $\tau_G$  i  $w_G$  veliki, to će stvarati diskontinuitete u slobodnoj energiji u ovisnosti o vremenu, stoga će biti nešto bolje koristiti manji  $\tau_G$ . Odabir širine Gaussijana ovisi o veličini sustava  $S$ . Zapravo je nepouzdanost manja što je omjer  $S/\delta s$  manji, što se vidi iz slučaja kada se nepouzdanost zapiše kao funkcija koja eksplicitno ovisi o vremenu simulacije. No to bi značilo da je nepouzdanost najmanja za jako veliku širinu Gaussijana  $\delta s$ , a kao što je već spomenuto, nije moguće dobro opisati detalje plohe slobodne energije na skalama manjim od širine Gaussijana. Pokazano je da kada je  $\delta s > 0,1S$ , tada se nepouzdanost ne može procijeniti ovom jednačbom.<sup>20</sup>

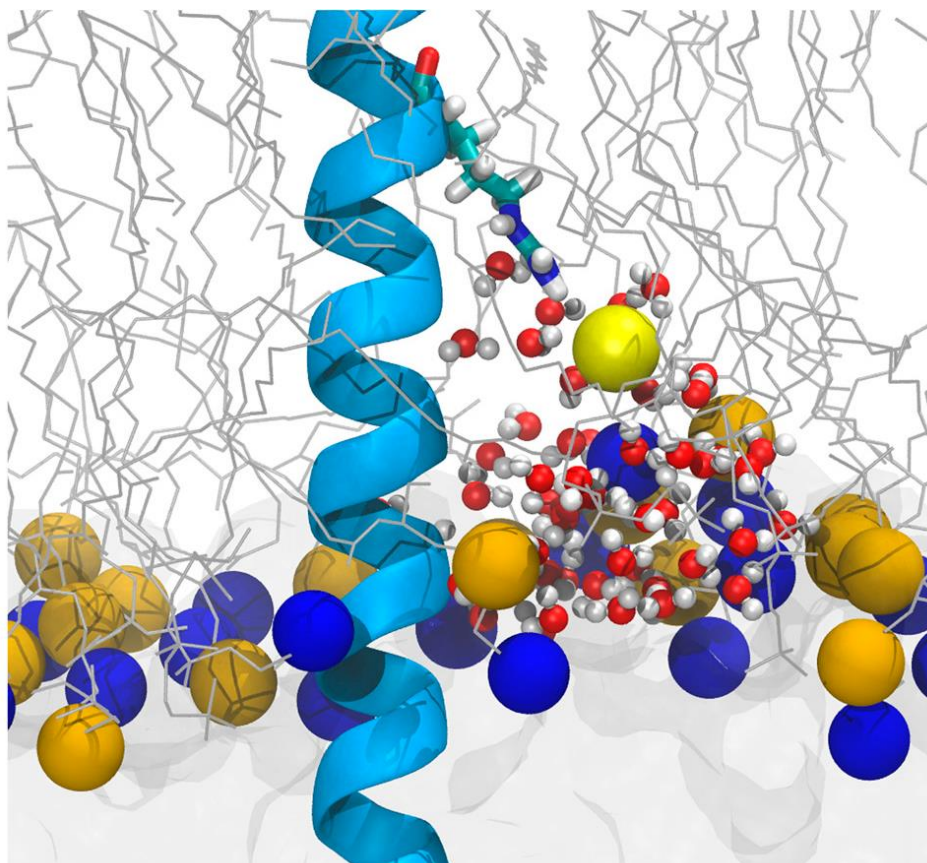
## 2.4. Primjena metoda za izračun slobodne energije

Sec translokon je kanal za proteine kroz koji rastući polipeptidni lanac može proći lipidni dvosloj prolazeći kroz transverzalna vrata ili se može u njega ugraditi prolazeći kroz lateralna vrata ako se radi o transmembranskom proteinu. Eksperimentalno je napravljena skala slobodne energije za svih 20 proteinogenih aminokiselina koja predviđa hoće li se neki aminokiselinski slijed pomoću Sec translokona ugraditi u membranu ili proći kroz lipidni dvosloj. Postoje pokušaji da se simulacijama molekulske dinamike objasni što je izvor lakoće umetanja nabijenih aminokiselina u lipidni dvosloj. No simulacije molekulske dinamike precjenjuju slobodnu energiju umetanja nabijenih aminokiselina kada se promatra direktno umetanje polipeptida u lipidni dvosloj.

Kako bi se izračunala slobodna energija za umetanje arginina u lipidni dvosloj, konstruirana je efektivno beskonačna  $\alpha$  – zavojnica od leucina, u kojoj je jedan leucin zamijenjen argininom.<sup>21</sup> Alkemijskom transformacijom leucina u geometrijskoj sredini lipidnog dvosloja u arginin te istovremenom transformacijom arginina daleko od dvosloja dobiva se slobodna energija za prijenos arginina u polileucinskoj zavojnici iz vodene otopine u lipidni dvosloj. To se radi mijenjanjem parametra reda postepeno, u 50 koraka, kako bi se osiguralo dobro



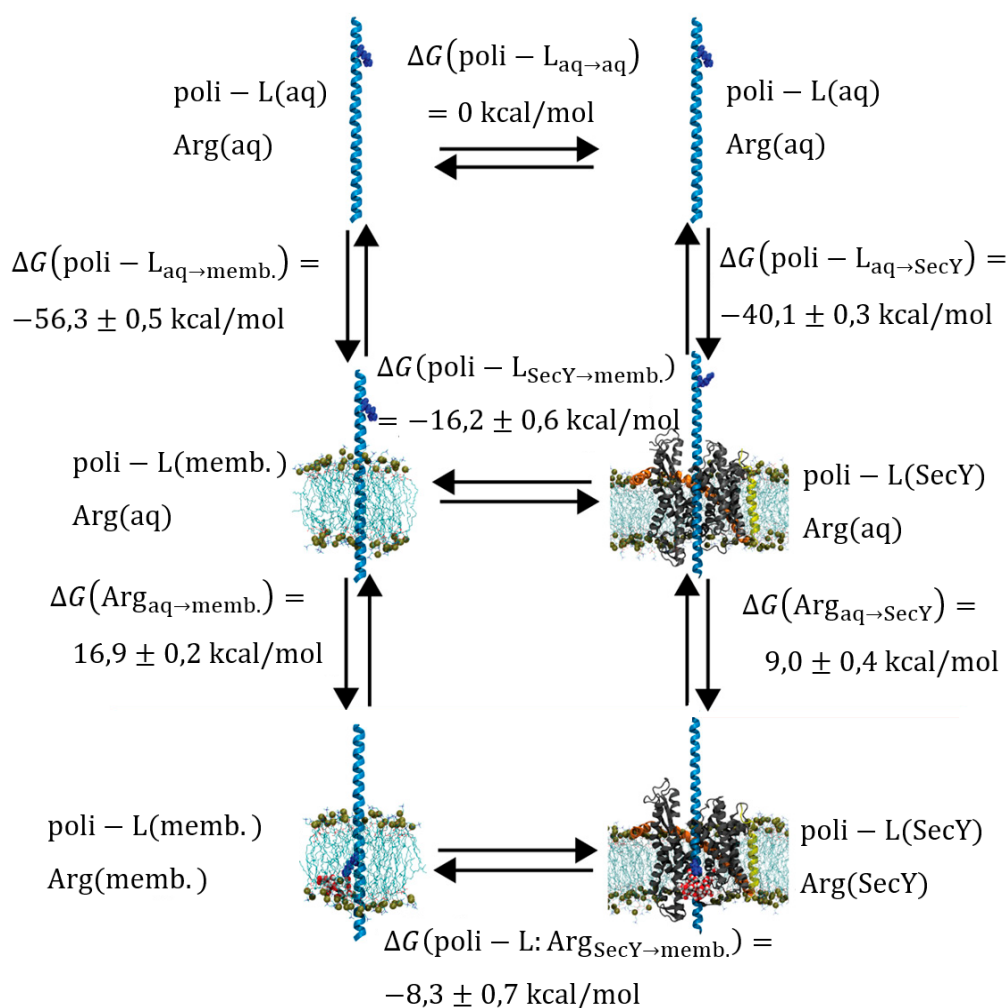
preklapanje faznog prostora. U međustanjima se smanjuje interakcija arginina s okolinom na mjestu udaljenom od lipidnog dvosloja, a na tom istom mjestu se istovremeno pojačavaju interakcije leucina s okolinom, no bočni ogranci tih aminokiselina međusobno ne interagiraju. Zbog naboja arginin ulazi u lipidni dvosloj koordiniran molekulama vode. Dolazi do lokalne deformacije dvosloja, ali ostaje strukturno očuvan,<sup>21</sup> što je prikazano na slici 10.



Slika 10. Prikaz leucinske zavojnice s jednim argininom u središtu lipidnog dvosloja. Arginin je koordiniran s molekulama vode te fosforom iz glave fosfolipida (žuta kuglica). (ref. 21)

No arginin u zavojnici leucina može se unijeti u lipidni dvosloj korištenjem vanjskog potencijala koji djeluje na geometrijsku koordinatu. U tu svrhu je korišten US metoda s harmonijskim potencijalima oblika  $\omega_i(z) = \frac{k_i}{2}(z - z_i)^2$  nakon čega su rezultati obrađeni korištenjem WHAM metode. Tim je pristupom dobivena slobodna energija za prijenos arginina u leucinskoj zavojnici iz vode u lipidni dvosloj  $\Delta G(\text{Arg}_{\text{aq} \rightarrow \text{memb.}}) = 17 \text{ kcal/mol}$ .<sup>22</sup> To se odlično slaže s vrijednošću dobivenom alkemijskim transformacijama arginina i leucina uz korištenje *Bennett acceptance ratio* (BAR) metode koja kombinira transformacije u oba smjera kako bi se smanjila statistička nepouzdanost,  $\Delta G(\text{Arg}_{\text{aq} \rightarrow \text{memb.}}) = 16,9 \pm 0,2 \text{ kcal/mol}$ .<sup>21,22</sup>

No prilikom ulaska arginina u membranu događale su se ogromne deformacije dvosloja. U svakom prozoru potrebno je relaksirati lipidni dvosloj, što u nekim prozorima zahtijeva dulju ekvilibraciju.<sup>22</sup> Iako je moguće dobiti profil slobodne energije korištenjem vanjskog potencijala koji djeluje na geometrijsku koordinatu, računalno je to vrlo zahtjevno. Budući da radi usporedbe s eksperimentalnim podacima potrebno samo znati ukupno promjenu slobodne energije između početnog i konačnog stanja, puno je isplativije koristiti niz nefizikalnih međustanja kako bi se u konačnici dobila promjena slobodne energije za unos arginina u lipidni dvosloj.<sup>21,22</sup> No dobivene vrijednosti slobodne energije za ove procese ne slažu se s onima dobivenim na temelju biološke skale hidrofobnosti aminokiselina, zato jer se ne uzima u obzir stabilizacijska uloga translokona SecY. Tu se vidi dodatna prednost korištenja alkemijskih transformacija budući da je vrlo jednostavno konstruirati termodinamički ciklus u kojem je konačni korak prijenos zavojnice s argininom iz translokona u lipidni dvosloj. Termodinamički ciklus je prikazan na slici 11.



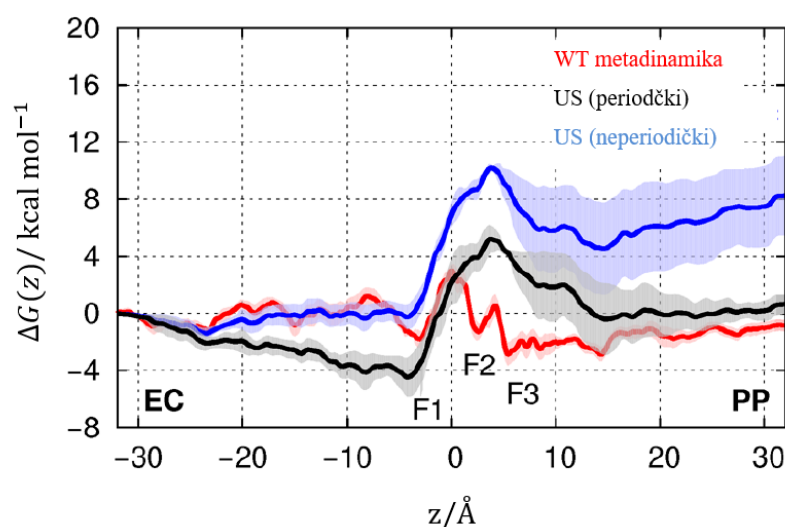
Slika 11. Termodinamički ciklus za izračun promjene Gibbsove slobodne energije za umetanje leucinske zavojnice s argininom u lipidni dvosloj. (ref. 21)

Unos arginina iz vodene otopine u SecY je napravljen na isti način kao i unos u lipidni dvosloj, tako da se nizom alkemijskih transformacija leucin u središtu SecY pretvori u arginin, a arginin u vodenoj otopini u leucin, čime se dobije  $\Delta G(\text{Arg}_{\text{aq}} \rightarrow \text{SecY}) = 9,0 \pm 0,4 \text{ kcal/mol}$ . Arginin je i u SecY koordiniran molekulama vode no transmembranske su zavojnice SecY djelomično hidratizirane, što uz manje deformacije lipidnog dvosloja čine SecY puno povoljnijim okruženjem za arginin nego što je to lipidni dvosloj, zbog čega je  $\Delta G(\text{Arg}_{\text{aq}} \rightarrow \text{SecY})$  skoro dva puta manji od  $\Delta G(\text{Arg}_{\text{aq}} \rightarrow \text{memb.})$ . Da bi se dobila vrijednost slobodne energije za prijenos arginina u leucinskoj zavojnici iz SecY u lipidni dvosloj,  $\Delta G(\text{poli-L: Arg}_{\text{SecY}} \rightarrow \text{memb.})$ , potrebno je izračunati slobodnu energiju za prijenos leucinske zavojnice iz SecY u lipidni

dvosloj,  $\Delta G(\text{poli} - L_{\text{SecY} \rightarrow \text{memb.}})$ . To se radi tako da se računa slobodna energija za prijenos potpuno otopljene leucinske zavojnice u lipidni dvosloj,  $\Delta G(\text{poli} - L_{\text{aq} \rightarrow \text{memb.}})$  odnosno SecY,  $\Delta G(\text{poli} - L_{\text{aq} \rightarrow \text{SecY}})$ . No ni ta se perturbacija ne odvija izravno, već se leucinska zavojnica iz vakuuma prebacuje u otopinu, tj. u lipidni dvosloj/SecY i zatvaranjem termodinamičkih ciklusa (koji nisu prikazani na slici 11.) se izračunava promjena slobodne energije za prijenos leucinske zavojnice iz otopine u dvosloj, tj. u SecY. Prilikom tih alkemijskih transformacija arginin nije bio prisutan. Postoji velika razlika između slobodnih energija ovisno o tome prenosi li se zavojnica iz vakuuma u dvosloj/SecY ili obrnuto, tj. ovisno radi li se anihilacija ili kreacija zavojnice u membrani ili u translokonu. Budući da se arginin tijekom simulacije nalazi na leucinskoj zavojnici, ne može se izravno usporediti dobivena vrijednost promjene Gibbsove energije s vrijednošću dobivenom na temelju biološke skale hidrofobnosti aminokiselina, budući da ona kvantificira doprinose pojedinačnih aminokiselina stabilizaciji transmembranskog segmenta. No izračunom razlika između promjena slobodnih energija za umetanje leucina i arginina na leucinskoj zavojnici u lipidni dvosloj, uz određene korekcije je izvedena  $\Delta G$  za umetanje arginina koja iznosi 6,5 kcal/mol. Iako metoda i dalje precjenjuje težinu umetanja arginina (koja iznosi 2,6 kcal/mol), rezultat je točniji nego u slučaju direktnog prijenosa iz vodene otopine u lipidni dvosloj, zanemarujući ulogu SecY.<sup>21</sup>

U drugom primjeru je pomoću metadinamike i US simulacija okarakteriziran prolazak kloridnog, dihidrogenfosfatnog iona te antibiotika fosmidomicina kroz kanal bakterijskog trimernog proteina OprO.<sup>23</sup> Korištena je *Well-Tempered* inačica metadinamike pri čemu je vanjski potencijal djelovao na z-komponentu vektora razlike centara masa između  $\alpha$ -ugljikovih atoma monomera OprO i otopljenih tvari koje prolaze kroz kanal tog monomera. Istovremeno se promatrao prolazak triju istih otopljenih tvari kroz tri monomera OprO. Kod US simulacija korišten je harmonijski potencijal u više prozora koji je djelovao na istu kolektivnu varijablu kao i kod simulacija metadinamike, te je opet promatran istovremeni prolazak triju otopljenih tvari kroz monomere. Rezultati su kombinirani WHAM metodom. Plohe slobodnih energija za prolazak kloridnih iona kroz različite monomere OprO-a dobivene korištenjem *Well-Tempered* metadinamike i US metodom se odlično poklapaju. Plohe slobodne energije dobivene kao prosjek ploha dobivenih za monomere OprO-a se za US i *Well-Tempered* metadinamiku također poklapaju u položajima, ali nešto manje u relativnim visinama maksimuma i dubinama minimuma. Lošiji su rezultati dobiveni za fosmidomicin te dihidrogenfosfatne ione. Razlog tomu je što te molekule posjeduju rotacijske stupnjeve slobode koji neće biti dobro uzorkovani

u uskoj regiji kanala. U tom slučaju do izražaja dolazi i razlika između ova dva pristupa: kada fosmidomicin uđe u vrlo usku regiju kanala, unutar njega ne može rotirati, ali može izaći iz te regije i ući natrag reorijentiran. To se redovito događa kada se popuni minimum na plohi potencijalne energije u metadinamici, jer se tada događa slobodna difuzija. No kada se harmonijskim potencijalima fiksira reakcijska koordinata unutar nekog intervala, fosmidomicin se ne može reorijentirati tako da izađe iz te uske regije. Zbog toga su se morale provesti neovisne simulacije u kojima je na početku fosmidomicin bio u dvije suprotne orijentacije. Iako su rezultati upitne točnosti za fosmidomicin, u simulacijama metadinamike je barem dobiven simetrični profil slobodne energije. Također, manja je standardna devijacija slobodne energije za prolazak fosmidomicina kroz monomere. Rezultati US simulacija su poboljšani tako da je WHAM metoda korištena uz periodičku korekciju, no i dalje su rezultati *Well-Tempered* metadinamike bili pouzdaniji. Dobiveni profili slobodne energije prikazani su na slici 12.<sup>23</sup>



Slika 12. Prosječni profili slobodnih energija dobivenih prolaskom fosmidomicina kroz različite monomere, dobiveni pomoću *Well-Tempered* metadinamike te US metodom, uz periodičku korekciju prilikom korištenja metode analize otežanim histogramima (US periodički) i bez periodičke korekcije (US neperiodički). (ref. 23)

### § 3. LITERATURNI IZVORI

1. I. N. Levine *Physical Chemistry*, McGraw-Hill, Inc., New York, 2009., str. 37., 109–115.
2. M. P. Allen, D. J. Tildesley *Computer Simulation of Liquids*, Oxford University Press, London, 2017., str. 1–58., 95–99.
3. C. Chipot, A. Pohorille, *Free Energy Calculations. Theory and applications in chemistry and biology*, Springer Verlag Berlin Heidelberg, 2007., str. 1–91., 199–220.
4. N. Lu, D.A. Kofke, *J. Chem. Phys.* **114** (2001) 7303–7311.
5. D. Wu, D.A. Kofke, *J. Chem. Phys.* **123** (2005) 054103
6. J. Hermans, L. Wang, *J. Am. Chem. Soc.* **119** (1997) 2707–2714.
7. J. Kästner, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1** (2011) 932–942.
8. B. Roux, *Comput. Phys. Commun.* **91** (1995) 275–282.
9. T. C. Beutler, W. F. van Gunsteren, *J. Chem. Phys.* **100** (1994) 1492–1497.
10. J. Kästner, W. Thiel, *J. Chem. Phys.* **124** (2006) 234106
11. F. Zhu, G. Hummer, *J. Comput. Chem.* **33**(4) (2012) 453–465.
12. M. N. Kobrak, *J. Comput. Chem.* **24**(12) (2003) 1437–1446
13. C. Bartels, M. Karplus, *J. Comput. Chem.* **18** (1997) 1450–1462.
14. A. Barducci, M. Bonomi, M. Parrinello, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1** (2011) 826–843.
15. G. Bussi, A. Laio, *Nat. Rev. Phys.* **2** (2020) 200–212.
16. G. Bussi, D. Branduardi, *Rev. Comput. Chem.* **28** (2015) 1–49.
17. A. Laio, F. L. Gervasio, *Rep. Prog. Phys.* **71** (2008) 126601
18. A. Barducci, G. Bussi, M. Parrinello, *Phys. Rev. Lett.* **100** (2008) 020603
19. F. Baftizadeh, P. Cossio, F. Pietrucci, A. Laio, *Current Physical Chemistry* **2**(1) (2012) 79–91.
20. A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, M. Parrinello, *J. Phys. Chem. B* **109**(14) (2005) 6714–6721.
21. J. Gumbart, C. Chipot, K. Schulten, *Proc. Natl. Acad. Sci. U.S.A.* **108**(9) (2011) 3596–3601.
22. S. Dorairaj, T. Allen, *Proc. Natl. Acad. Sci. U.S.A.* **104**(12) (2007) 4943–4948.

23. V. K. Golla, J. D. Prajapati, M. Joshi, U. Kleinekathöfer, *J. Chem. Theory Comput.* **16**(4) (2020) 2751–2765.