

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 247

**MODUL ZA RAČUNANJE PRIANJANJA  
PROTEINA KORISTEĆI ELEKTROSTATSKI  
POTENCIJAL**

Jura Ćurić

Zagreb, lipanj 2008.

*Hvala mami i obitelji na brizi i podršci, prijateljima za dugogodišnju podršku i motivaciju, glazbi što postoji, Mili Šikiću za odličnu prijateljsku komunikaciju i neopterećenost formalnostima, te Ivanu Dokmaniću i cijelom projektnom timu na pomoći!*

## Sadržaj

1. Uvod.....	2
2. Prianjanje proteina .....	3
2.1. PDB format .....	3
2.2. Elektrostatske interakcije .....	3
2.2.1. Potencijal – uzrok postojanja naboja u prostoru.....	4
3. Reprezentacija prostorne gustoće naboja i elektrostatskog potencijala .....	6
4. Matematička razrada .....	7
4.1. Sferne polarne koordinate.....	7
4.2. Kugline funkcije .....	8
4.2.1. Realne kugline funkcije .....	8
4.3. Ortonormalnost .....	9
4.4. Rotacija .....	11
4.4.1. Wignerove rotacijske matrice .....	12
4.4.2. Realne roacijske matrice .....	13
4.5. Radijalne funkcije .....	15
4.5.1. Pogodne radijalne funkcije .....	16
4.5.2. Schrödingerova jednađba .....	16
4.5.3. Orbitale eksponencijalnog tipa .....	17
4.6. 3-D parametrizacija.....	18
5. Elektrostatska komplementarnost .....	19
5.1. Računanje elektrostatske energije vezanja kompleksa .....	19
5.2. Integrali preklapanja kao elementi traslacijskih matrica .....	20
5.3. Konačna formula za proračun elektrostatske energije vezanja kompleksa.....	23
5.4. Implementacija algoritma .....	24
5.4.1. Pretraživanje prostora interakcija .....	27
5.4.2. Pseudokod .....	27
5.5. Računanje koeficijenata.....	29
5.5.1. Računanje koeficijenata prostorne gustoće naboja.....	29
5.5.2. Računanje koeficijenata elektrostatskog potencijala .....	30
6. Rezultati i diskusija .....	32
7. Zaključak .....	33
8. Literatura .....	34
9. Naslov, sažetak i ključne riječi.....	35

# 1. Uvod

Proteini su članovi velike porodice bioloških makromolekula. Njihova strukturalna, katalitička i transportna svojstva čine ih neophodnima za život svih bića. Biokemija proteina eksperimentalno se proučava već godinama, a zahvaljujući aktualnim tehnološkim poboljšanjima (npr. u području magnetske rezonance) moguće je predvidjeti i utvrditi odnos između proteinskih molekularnih struktura i funkcija. Zadnjih godina bilježi se velik porast u komercijalnoj primjeni biomolekula gdje su najzastupljenija monoklonska antitijela kao lijekovi, te terapijski i dijagnostički postupci. Poboljšavajući sklonost vezanja proteina u svrhu pronalaska ciljanog lijeka, sam lijek može postati efektivniji, dok se povećavajući specifičnost (namjenu) lijeka mogu izbjeći mnoge neželjene popratne pojave. Jedan način koji pomaže ostvariti dane ciljeve jest računalno modeliranje interakcija proteina. Općenito se pretpostavlja da dovoljno točan model tih interakcija omogućuje sustavan dizajn i selekciju proteina sa specifičnim fiziološkim svojstvima. Mogućnost brzog i vjerodostojnog pronalaska tj. utvrđivanja komplementarnosti određenih svojstava proteina od velikog je značaja za računalno modeliranje proteinskih interakcija. Glavno svojstvo koje se proučava u tu svrhu jest komplementarnost površina dvaju proteina, tj. u kojoj se mjeri preklapanje površine dvaju proteina. Mada glavno, ovo svojstvo nije dovoljno da bi dovoljno dobro opisalo interakciju dvaju proteina. Jedna od nadogradnja navedenog modela postiže se uzimanjem u obzir elektrostatske komplementarnosti kompleksa dvaju proteina što je ujedno i tema ovog rada. Zbog svoje jakosti, elektrostatske su interakcije veoma važne za razumijevanje makromolekularnih interakcija. Uzimajući u obzir veoma male udaljenost na kojima se interakcije odvijaju (mjere se u angstromima ( $\text{\AA}$ ),  $1\text{\AA}=10^{-10}$  m), teoretski gledano, ovisnost o kvadratu udaljenosti kojem je elektrostatska sila obrnuto proporcionalna jamči veliku ulogu tih sila na svim udaljenostima gdje su interakcije moguće.

## 2. Prianjanje proteina

Općenito, problem prianja može se opisati kao računalna metoda koja nastoji odrediti najbolje slaganje, tj. najbolji kompleks dvije molekule : *receptora* i *liganda*. Pri tome je potrebno poznavati neke temeljne karakteristike pojedinih molekula (u ovom slučaju proteina) čije se vezanje ispituje. Npr., kod razmatranja komplementarnosti površina intuitivno je jasno da je potrebno znati točke koje čine površinu proteina, dok je kod razmatranja elektrostatske komplementarnosti potrebno poznavati raspodjelu naboja u proteinu. U tu svrhu, kao ulaz programa za računanje ocjene kvalitete prianjanja koriste se PDB datoteke.

### 2.1. PDB format

*Protein Data Bank* (PDB) je arhiva eksperimentalno utvrđenih trodimenzionalnih struktura bioloških makromolekula koja služi globalnoj zajednici istraživača, nastavnika i studenata. Podatci sadržani u arhivi uključuju koordinate atoma, bibliografske citate, primarnu i sekundarnu strukturu, informacije, faktore kristalografske strukture i eksperimentalne podatke dobivene nuklearnom magnetskom rezonancom. Svaka PDB datoteka organizirana je po redcima.

### 2.2. Elektrostatske interakcije

Temeljni dijelovi proteina su aminokiseline. Pojedine grupe aminokiselina sadrže skupine koje su često ionizirane, te radi postojanja naboja stvaraju elektrostatski potencijal u prostoru. Zamislimo sada dva proteina u prostoru. Budući da se proteini sastoje od ioniziranih dijelova, u oba je prisutna određena prostorna *gustoća naboja*. Svi naboji pojedinog proteina stvaraju u prostoru elektrostatski potencijal. Ponašanje naboja koji se nalaze u takvom elektrostatskom potencijalnom polju moguće je opisati energijom. Budući da je u interesu najbolje moguće slaganje dvaju proteina, jasno je da ta energija mora biti *minimalna*, tj., potencijalno polje koje stvara jedan protein mora „privlačiti“ naboje (ionizirane dijelove) drugog proteina.

### 2.2.1. Potencijal – uzrok postojanja naboja u prostoru

Prvo se promatra homogeni sustav s dielektričnom konstantom  $\varepsilon$ , bez prisutnih naboja. Elektrostatski potencijal u prostoru  $\phi$  opisan je Laplaceovom jednađbom:

$$\bar{\nabla} \left[ \bar{\nabla} \phi(\underline{r}) \right] = 0 \quad (2-1)$$

Uz prisustvo naboja, Laplaceova jednađba prelazi u Poissonovu:

$$\varepsilon \bar{\nabla} \left[ \bar{\nabla} \phi(\underline{r}) \right] = -4\pi\rho(\underline{r}) \quad (2-2)$$

Uzimajući u obzir polarizacijske naboje koji se razvijaju na granicama dielektrika, koristi se divergencija dielektrične konstante koja postaje ovisna u prostoru:

$$\varepsilon \bar{\nabla} \left[ \varepsilon(\underline{r}) \bar{\nabla} \phi(\underline{r}) \right] = -4\pi\rho(\underline{r}) \quad (2-2)$$

U svakom kompleksnom sustavu čestica u interakciji, gustoća čestice  $\delta$  može se izraziti relativno u odnosu na gustoću te iste čestice u odsustvu interakcija s drugim česticama u sustavu. Tu gustoću označavamo s  $\delta_0$  i vrijedi:

$$\delta(\underline{r}) = g(\underline{r})\delta_0(\underline{r}) \quad (2-3)$$

Omjer stvarne i prosječne gustoće čestice naziva se distribucijska funkcija  $g(\underline{r})$ :

$$g(\underline{r}) = e^{-w(\underline{r})/kT} \quad (2-4)$$

gdje je  $w(\underline{r})$  potencijal srednje sile. Naziv „potencijal srednje sile“ proizlazi iz činjenice da njegov gradijent daje srednju silu koja djeluje na česticu. Pretpostavimo li da je ionski potencijal srednje sile jednak prosječnom elektrostatskom potencijalu pomnoženom s nabojem iona, Poissonova jednađba prelazi u Poisson-Boltzmanovu:

$$\varepsilon \bar{\nabla} \left[ \varepsilon(\underline{r}) \bar{\nabla} \phi(\underline{r}) \right] = -4\pi\rho^f(\underline{r}) + 4\pi \sum_i c_i^\infty z_i q \cdot \exp\left(\frac{-z_i q \phi(\underline{r})}{kT}\right) \lambda(\underline{r}) \quad (2-5)$$

gdje je  $\rho^f(\underline{r})$  prostorna raspodjela naboja molekule,  $c_i^\infty$  je koncentracija iona  $i$  na beskonačnoj udaljenosti od molekule,  $z_i$  njegova valencija,  $q$  naboj protona,  $k$  Boltzmannova konstanta,  $T$  temperatura, a  $\lambda(\underline{r})$  opisuje pristupačnost iona.

Uz pretpostavku da je potencijal malen, jednačba (2-5) se može linearizirati te se dobiva konačan izraz za elektrostatski potencijal u prostoru:

$$\varepsilon \bar{\nabla} \left[ \varepsilon(\underline{r}) \bar{\nabla} \phi(\underline{r}) \right] = -4\pi \rho^f(\underline{r}) + 4\pi \frac{\sum_i c_i^\infty z_i^2 q^2 \phi(\underline{r})}{kT} \lambda(\underline{r}) \quad (2-6)$$

### 3. Reprezentacija prostorne gustoće naboja i elektrostatskog potencijala

Koje god svojstvo proučavali pri vezanju proteina, kao i svim drugim računalnim problemima, poželjno je da sam postupak bude što *brži i točniji*. Nažalost, želimo li postići točnije rezultate, program će biti sporiji. S druge strane, želimo li ubrzati program, često ćemo morati žrtvovati preciznost dobivenih rezultata.

Pokazalo se da postoji metoda u kojoj možemo problem dokiranja rastaviti na poddijelove, te time osigurati prilično veliku točnost i brzinu. Umjesto da funkcije gustoće naboja i elektrostatskog potencijala prikazujemo eksplicitno, u ovom će radu one biti prikazane razvojem realnih ortogonalnih radijalnih i kuglinih funkcija u Fourierov red. Glavni razlog za odabir kuglinih funkcija (sfernih harmonika) leži u činjenici da se rotacijom transformiraju među sobom na predvidljiv način, dok radijalne funkcije dobro opisuju opadanje potencijala u prostoru.

Na taj način moguće je opisati prijanjanje dva proteina u šest-dimenzionalnom prostoru. Svaki protein promatra se kao kruto tijelo sa šest stupnjeva slobode. Naravno, navedena pretpostavka nije sasvim točna jer prilikom interakcija nastaju razne konformacije proteina koje bi također trebalo uzeti u obzir prilikom interakcije dvaju proteina. Ipak, ovo svojstvo mnogo povećava kompleksnost sustava, stoga se za sad još neće razmatrati.

Kako bi se ova glavna svojstva mogla iskoristiti u 6D postupku prijanjanja, potrebno je šest stupnjeva slobode podijeliti na pet Eulerovih kutova rotacije i međumolekularnu udaljenost. Dakle, neovisno o međumolekularnoj udaljenosti, svaka od dvije molekule promatra se učvršćena u prostoru, a nad njom se vrši rotacija oko središta mase molekule. Ovakav pristup se pokazao u mnogim stvarima učinkovitiji nego kod drugih metoda prijanjanja. Npr. za razliku od također popularne FFT metode, biti će pokazano da je koeficijente razvoja funkcija određenog svojstva u Fourierov red dovoljno odrediti samo jednom, i najvažnije, proces možemo podijeliti u nekoliko nezavisnih dijelova, što omogućuje paralelizaciju procesa, te se samim time proces može mnogo ubrzati.



## 4. Matematička razrada

### 4.1. Sferne polarne koordinate

Kako je prije navedeno, elektrostatske karakteristike će biti opisane razvojem u red ortogonalnih funkcija u sfernim polarnim koordinatama. Ako  $A(\underline{r})$  predstavlja proizvoljnu funkciju u trodimenzionalnom prostoru, onda se ona u sfernim polarnim koordinatama može prikazati pomoću izraza:

$$A(\underline{r}) = \sum_{nlm}^N a_{nlm} R_{nl}(r) y_{lm}(\theta, \phi); \quad N \geq n > l \geq |m| \geq 0, \quad (4-1)$$

gdje je

- $\sum_{nlm}^N$  skraćeni zapis trostruke sume, tj.  $\sum_{nlm}^N = \sum_{n=1}^N \sum_{l=0}^{n-1} \sum_{m=-l}^l$
- $a_{nlm}$  koeficijenti razvoja koje treba naknadno odrediti
- $y_{lm}(\theta, \phi)$  realne kugline funkcije (realni sferni harmonici)
- $R_{nl}(r)$  radijalne funkcije

$N$  predstavlja veličinu reda razvoja. Za točan prikaz funkcije  $A(\underline{r})$  bilo bi potrebno pustiti  $N$  da teži u beskonačnost. Ipak, za potrebe prijanjanja biti će dovoljno uzeti  $N \leq 25$ . Iz zapisa sume je lako izračunati ukupan broj elemenata sumacije. Broj elemenata unutrašnje sumacije  $\sum_{m=-l}^l$  je očito  $2l+1$ . Razrješavanjem druge sume dobivamo:

$$\sum_{l=0}^{n-1} (2l+1) = 2 \sum_{l=0}^{n-1} l + \sum_{l=0}^{n-1} 1 = 2(0+1+2+\dots+n-1) + n = 2 \frac{(n-1)n}{2} + n = n^2 \quad (4-2)$$

te naposljetku razrješavanjem posljednje (vanjske) sume dobivamo broj koeficijenata u razvoju funkcije do određenog reda  $N$ :

$$\sum_{n=1}^N n^2 = N(N+1)(2N+1)/6 \quad (4-3)$$

## 4.2. Kugline funkcije

Laplaceova diferencijalna jednačba u sfernim polarnim koordinatama glasi

$$\nabla^2 f(\underline{r}) = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f(\underline{r})}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f(\underline{r})}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f(\underline{r})}{\partial \phi^2} = 0. \quad (4-4)$$

Rješenje kutnog djela jednačbe (dobiveno jednostavnom separacijom varijabli) su kompleksne kugline funkcije oblika

$$Y_{lm}(\theta, \phi) = \mathcal{Y}_{lm}(\mu) Z_m(\phi); \quad \mu = \cos \theta \quad (4-5)$$

gdje je

$$\mathcal{Y}_{lm}(\mu) = \left[ \frac{(2l+1)(l-m)!}{2(l+m)!} \right]^{1/2} P_l^m(\mu) \quad (4-6)$$

te

$$Z_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{jm\phi} \quad (4-7)$$

$P_l^m(\mu)$  su pridruženi Legendreovi polinomi. Budući da su na internetu dostupne biblioteke koje podržavaju računanje Legendrovih polinoma, oni neće biti ovdje bitno detaljno opisivani. Kao što je pokazano, kugline funkcije (sferni harmonici) su funkcije sfernih koordinata  $(\theta, \phi)$ .

### 4.2.1. Realne kugline funkcije

Uzimajući u obzir da kod prianjanja proteina nema potrebe za računanjem kompleksnih veličina te da je računanje realnih funkcija brže od kompleksnih, u daljnjoj razradi se koriste isključivo realne kugline funkcije koje se mogu dobiti linearnim kombiniranjem kompleksnih:

$$y_{lm}(\theta, \phi) = \begin{cases} (Y_{lm}(\theta, \phi) + Y_{lm}(\theta, \phi)^*) / \sqrt{2} & m > 0 \\ Y_{l0}(\theta, \phi) & m = 0 \\ -j(Y_{lm}(\theta, \phi) - Y_{lm}(\theta, \phi)^*) / \sqrt{2} & m < 0 \end{cases} \quad (4-8)$$

Uvrštavanjem odgovarajućih izraza, dobiva se

$$y_{lm}(\theta, \phi) = \begin{cases} \mathcal{Y}_m(\mu)(\cos m\phi) / \sqrt{\pi} & m > 0 \\ \mathcal{Y}_m(\mu) / \sqrt{2\pi} & m = 0 \\ \mathcal{Y}_{\bar{m}}(\mu)(\sin \bar{m}\phi) / \sqrt{\pi} & m < 0 \dots \bar{m} > 0 \end{cases} \quad (4-9)$$

Realne kugline funkcije se dakle mogu zapisati u obliku

$$y_{lm}(\theta, \phi) = \mathcal{Y}_{|m|}(\mu)\varphi_m(\phi) \quad (4-10)$$

gdje je

$$\varphi_m(\phi) = \begin{cases} \cos m\phi / \sqrt{\pi} & m > 0 \\ 1 / \sqrt{2\pi} & m = 0 \\ \sin \bar{m}\phi / \sqrt{\pi} & m < 0 \dots \bar{m} > 0 \end{cases} \quad (4-11)$$

Za proces prijanjanja koriste se realne kugline funkcije oblika  $y_{lm}(\theta, \phi)$ , gdje indeks  $m$  može biti negativan, uz uvjet  $|m| \leq l$ .

### 4.3. Ortonormalnost

Općenito, skup kuglinih funkcija je potpuna ortonormirana baza s obzirom na polarni integral, tj. na jediničnoj sferi. To možemo zapisati

$$\int_0^{2\pi} \int_0^\pi y_{lm}(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi = \delta_{lk} \delta_{mj}. \quad (4-12)$$

Uzmimo sada bilo koju funkciju  $\mu$ , parametriziranu u polarnim koordinatama  $(\theta, \phi)$  i prikažimo kao linearnu kombinaciju realnih kuglinih funkcija

$$\mu(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} y_{lm}(\theta, \phi). \quad (4-13)$$

Pomnožimo li danu jednadžbu s  $y_{kj}(\theta, \phi)$  i integriramo preko cijelog prostornog kuta dobivamo

$$\int_0^{2\pi} \int_0^{\pi} \mu(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi = \int_0^{2\pi} \int_0^{\pi} \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} y_{lm}(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi \quad (4-14)$$

Uzimajući u obzir da je suma pod integralom konvergentna, možemo zamijeniti redosljed integracije i sumacije čime dobivamo

$$\int_0^{2\pi} \int_0^{\pi} \mu(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} \int_0^{2\pi} \int_0^{\pi} y_{lm}(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi. \quad (4-15)$$

Ako sada primijenimo jednadžbu ortonormiranosti (4-12), dobivamo oblik

$$\int_0^{2\pi} \int_0^{\pi} \mu(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} \delta_{lk} \delta_{mj}, \quad (4-16)$$

iz kojeg slijedi da desna strana jednadžbe (4-12) ne iščezava jedino za  $l = k$  i  $m = j$ . Dakle, samo je jedan član sumacije s desne strane različit od nule što daje konačan rezultat

$$a_{lm} = \int_0^{2\pi} \int_0^{\pi} \mu(\theta, \phi) y_{lm}(\theta, \phi) \sin \theta d\theta d\phi \quad (4-17)$$

Koeficijenti  $a_{lm}$  su jedinstveni te se pokazuje da je razvoj (4-13) strogo konvergentan u smislu da se greška razvoja monotono smanjuje kako  $l$  teži u beskonačnost. Budući da su glavni članovi u razvoju oni s malim indeksima  $l$  i  $m$ , prekidanjem reda nakon dovoljno velikog  $l$  možemo dovoljno vjerno prikazati funkciju  $\mu(\theta, \phi)$ .

Time je pokazan važan razlog za odabir kuglinih funkcija, tj. *svojstvo ortonormalnosti omogućuje lako određivanje koeficijenata razvoja.*

#### 4.4. Rotacija

Drugi važan razlog za odabir kuglinih funkcija je činjenica da je prostor kuglinih funkcija zatvoren s obzirom na rotacije, te da se kugline funkcije svakog reda  $l$  rotacijom transformiraju između sebe na predvidljiv način, tj.

$$\widehat{R}(\alpha, \beta, \gamma) y_{lm}(\theta, \phi) = \sum_{m'} y_{lm'}(\theta, \phi) R_{m'm}^{(l)}(\alpha, \beta, \gamma) \quad (4-18)$$

pri čemu  $\widehat{R}(\alpha, \beta, \gamma)$  predstavlja operator rotacije,  $(\alpha, \beta, \gamma)$  su eulerovi kutevi dok  $R_{m'm}^{(l)}$  predstavlja kvadratnu realnu rotacijsku matricu dimenzija  $(2l+1) \times (2l+1)$ .

Uzmimo sad novu funkciju  $\mu'(\theta, \phi)$  koja predstavlja rotiranu funkciju  $\mu(\theta, \phi)$ , tj. neka je

$$\mu'(\theta, \phi) = \widehat{R}(\alpha, \beta, \gamma) \mu(\theta, \phi) \quad (4-19)$$

Uvrštavajući odgovarajuće izraze za svaku stranu i uzimajući  $l_{\max} = L$ , formula (4-19) prelazi u oblik

$$\sum_{l=0}^L \sum_{m=-l}^l a_{lm} y_{lm}(\theta, \phi) = \sum_{l'=0}^L \sum_{m'=-l'}^{l'} a_{l'm'} \sum_q y_{l'q}(\theta, \phi) R_{qm'}^{(l')}(\alpha, \beta, \gamma) \quad (4-20)$$

Množenjem obje strane s  $y_{kj}(\theta, \phi)$  i integrirajući preko cijelog prostornog kuta dobivamo (uz zamjenu poretka integracije i sumacije) izraz

$$\begin{aligned} \sum_{l=0}^L \sum_{m=-l}^l a_{lm} \int_0^{2\pi} \int_0^{\pi} y_{lm}(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi = \\ \sum_{l'=0}^L \sum_{m'=-l'}^{l'} a_{l'm'} \sum_q R_{qm'}^{(l')}(\alpha, \beta, \gamma) \int_0^{2\pi} \int_0^{\pi} y_{l'q}(\theta, \phi) y_{kj}(\theta, \phi) \sin \theta d\theta d\phi \end{aligned} \quad (4-21)$$

Uvažavanjem ortogonalnosti izraz (4-21) prelazi u

$$\sum_{l=0}^L \sum_{m=-l}^l a_{lm} \delta_{lk} \delta_{mj} = \sum_{l'=0}^L \sum_{m'=-l'}^{l'} a_{l'm'} \sum_q R_{qm'}^{(l')}(\alpha, \beta, \gamma) \delta_{l'k} \delta_{qj} \quad (4-22)$$

te je vidljivo da iščezavaju svi članovi sumacija osim za  $l = k = l'$  i  $m = j = q$  iz čega slijedi konačan izraz

$$a'_{lm} = \sum_{m'} a_{lm'} R_{mm'}^{(l)}(\alpha, \beta, \gamma). \quad (4-23)$$

Dani rezultat pokazuje temeljno rotacijsko svojstvo kuglinih funkcija, tj. proizvoljnu funkciju  $\mu(\theta, \phi)$  možemo rotirati da bi dobili novu funkciju,  $\mu'(\theta, \phi)$  rotirajući samo koeficijente razvoja.

#### 4.4.1. Wignerove rotacijske matrice

Nakon što su pokazana neka temeljna svojstva rotacija, potrebno je prikazati analitički alat za implementaciju rotacija.

Općenito, rotacija kompleksne kugline funkcije  $Y_{lm}(\theta, \phi)$  za kutove  $(\alpha, \beta, \gamma)$  može se zapisati

$$\widehat{R}(\alpha, \beta, \gamma) Y_{lm}(\theta, \phi) = \sum_{m'} Y_{lm'}(\theta, \phi) D_{m'm}^{(l)}(\alpha, \beta, \gamma), \quad (4-24)$$

gdje su  $D_{m'm}^{(l)}(\alpha, \beta, \gamma)$  unitarne rotacijske matrice koje je pronašao Wigner. Elementi matrica dani su s

$$D_{m'm}^{(l)}(\alpha, \beta, \gamma) = e^{-im'\alpha} d_{m'm}^{(l)}(\beta) e^{-im'\gamma} \quad (4-25)$$

gdje je

$$d_{m'm}^{(l)}(\beta) = \left[ \frac{(l+m')!(l-m')!}{(l+m)!(l-m)!} \right]^{1/2} \times \sum_{k=\max(0, m-m')}^{\min(l-m', l+m)} (-1)^{k+m-m'} \binom{l+m}{k} \binom{l-m}{l-m'-k} \times (\cos \beta / 2)^{2l+m-m'-2k} (\sin \beta / 2)^{2k+m'-m} \quad (4-26)$$

Matrice  $d^l$  su višestruko simetrične:

$$d_{m'm}^{(l)}(\beta) = (-1)^{m'-m} d_{mm'}^{(l)}(\beta) = (-1)^{m'-m} d_{\bar{m}\bar{m}}^{(l)}(\beta), \quad (4-27)$$

što znatno ubrzava izračunavanje njihovih elemenata.

#### 4.4.2 Realne rotacijske matrice

Kao što je prije spomenuto, realne kugline funkcije mogu se dobiti linearnim kombiniranjem kompleksnih kuglinih funkcija. Budući da linearne kombinacije čuvaju rotacijsku simetriju, realne kugline funkcije, kao i kompleksne imaju povoljna rotacijska svojstva:

$$\widehat{R}(\alpha, \beta, \gamma)y_{lm}(\theta, \phi) = \sum_{m'} y_{lm'}(\theta, \phi)R_{m'm}^{(l)}(\alpha, \beta, \gamma), \quad (4-28)$$

gdje su  $R_{m'm}^{(l)}(\alpha, \beta, \gamma)$  realne rotacijske matrice. Ako linearne kombinacije (4-8) prikažemo kao sumu:

$$y_{lm}(\theta, \phi) = \sum_{m'} A_{mm'}^{(l)} Y_{lm'}(\theta, \phi), \quad (4-29)$$

tada se može pokazati da je za svaki red ( $l$ ), realna rotacijska matrica  $R$  dana jednadžbom

$$R = ADA^\dagger \quad (4-30)$$

gdje je  $A^\dagger$  Hermitska transpozicija matrice  $A$ . Pažljivim računom dobivaju se izrazi za elemente rotacijske matrice (Su & Coppens, 1994).

$$R_{m'm}^{(l)} = \begin{cases} d_{m'm}^{(l)}(\beta) \cos(m\gamma + m'\alpha) + (-1)^{m'} d_{\bar{m}'m}^{(l)}(\beta) \cos(m\gamma - m'\alpha) & \dots m' > 0, m > 0 \\ d_{0m}^{(l)}(\beta) \sqrt{2} \cos(m\gamma) & \dots m' = 0, m > 0 \\ (-1)^{m'+1} d_{m'm}^{(l)}(\beta) \sin(m\gamma + m'\alpha) + d_{\bar{m}'m}^{(l)}(\beta) \sin(m\gamma - m'\alpha) & \dots m' < 0, m > 0 \\ d_{m'0}^{(l)}(\beta) \sqrt{2} \cos(m'\alpha) & \dots m' > 0, m = 0 \\ d_{00}^{(l)}(\beta) & \dots m' = 0, m = 0 \\ (-1)^{m'+1} d_{m'0}^{(l)}(\beta) \sqrt{2} \sin(m'\alpha) & \dots m' < 0, m = 0 \\ (-1)^m d_{m'm}^{(l)}(\beta) \sin(m\gamma + m'\alpha) + (-1)^{m+m'} d_{\bar{m}'m}^{(l)}(\beta) \sin(m\gamma - m'\alpha) & \dots m' > 0, m < 0 \\ (-1)^m d_{0m}^{(l)}(\beta) \sqrt{2} \sin(m\gamma) & \dots m' = 0, m < 0 \\ (-1)^{m+m'} d_{m'm}^{(l)}(\beta) \cos(m\gamma + m'\alpha) + (-1)^{m+1} d_{\bar{m}'m}^{(l)}(\beta) \cos(m\gamma - m'\alpha) & \dots m' < 0, m < 0 \end{cases} \quad (4-31)$$

Ako odaberemo  $\beta = \gamma = 0$ , dobivamo samo jednu rotaciju  $\alpha$  oko z-osi, koja ima jednostavniju formu. Budući da vrijedi:

$$\widehat{R}(\alpha, 0, 0)y_{lm}(\theta, \phi) = y_{lm}(\theta, \phi - \alpha), \quad (4-32)$$

uz pomoć trigonometrijskih formula

$$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \quad (4-33)$$

$$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta \quad (4-34)$$

se dolazi do konačnog izraza za rotaciju kugline funkcije oko z-osi

$$\widehat{R}(\alpha, 0, 0)y_{lm}(\theta, \phi) = y_{lm}(\theta, \phi) \cos m\alpha + y_{l\bar{m}}(\theta, \phi) \sin m\alpha \quad (4-35)$$

što će se kasnije pokazati, pojednostavljuje i ubrzava proces dokiranja.



#### 4.5. Radijalne funkcije

Dobra svojstva koja nam pružaju kugline funkcije mogu se upotrijebiti za parametrizaciju i superpoziciju površine proteina i prostorne raspodjele naboja i potencijala  $u$ , odnosno oko proteina. Općenito, dvodimenzionalna ploha se u trodimenzionalnom prostoru može parametrizirati pomoću dvije varijable. Tada su sve točke koje leže na nekoj plohi zadane s

$$\underline{x}(\mu, \nu) = f_x(\mu, \nu)\underline{i} + f_y(\mu, \nu)\underline{j} + f_z(\mu, \nu)\underline{k}. \quad (4-36)$$

Osim kod najjednostavnijih ploha, veoma je komplicirano odrediti funkcije  $f_x$ ,  $f_y$  i  $f_z$ . Korištenjem sfernih koordinata  $(\theta, \phi)$  dobiva se oblik

$$\underline{x}(\theta, \phi) = f_x(\theta, \phi)\underline{i} + f_y(\theta, \phi)\underline{j} + f_z(\theta, \phi)\underline{k} \quad (4-37)$$

gdje se svaka od funkcija  $f_x$ ,  $f_y$  i  $f_z$  odvojeno razvija u red sfernih harmonika. Kako bi odredili koeficijente razvoja, potrebno je prvo projicirati molekularnu površinu na sferu, te iz koordinata  $(\theta_i, \phi_i)$  očitati koordinate pojedine točke  $(x_i, y_i, z_i)$ . Koristeći ovaj pristup, moguće je željene veličine (površina proteina, prostorna raspodjela naboja i potencijala) prikazati vrlo točno, no pritom je potrebno funkcije razviti u dosta visok red ( $L=50$ ).

Ipak, pokušaj uporabe ovog pristupa prilikom dokiranja proteina pokazao se veoma neuspješnim. Drugi i mnogo jednostavniji način jest kugline funkcije proširiti radijalnim funkcijama  $R_n(r)$  koje su jednodimenzionalne i ovisne o radijalnoj komponenti. Nove funkcije baze sada su

$$F_{nlm}(\underline{r}) = R_n(r)y_{lm}(\theta, \phi), \quad (4-38)$$

i one su trodimenzionalne. Neko svojstvo površine,  $\sigma$ , biti će sada prikazano kao Fourierov razvoj navedenih trodimenzionalnih funkcija baza u red :

$$\sigma(\underline{r}) = \sum_{nlm} a_{nlm} F_{nlm}(\underline{r}). \quad (4-39)$$

#### 4.5.1. Pogodne radijalne funkcije

Kod odabira radijalnih funkcija treba obratiti pozornost na neka određena svojstva. Odabrana radijalna funkcija mora dovoljno dobro opisivati promatrana svojstva proteina (oblik površine, prostorna raspodjela naboja i potencijala) i to s takvim redom razvoja koji će biti pogodan (ne prevelik) za izvođenje procesa dokiranja na računalu. Dani zahtjevi impliciraju na uporabu Laguerreovih polinoma koji su sastavni dio rješenja Schrödingerove jednadžbe za vodikov atom.

#### 4.5.2. Schrödingerova jednadžba

Schrödingerova jednadžba je valna jednadžba koja opisuje kruženje elektrona oko pozitivno nabijene vodikove jezgre (protona) i glasi

$$\nabla^2\psi(\underline{x})+V(\underline{x})\psi(\underline{x})=E\psi(\underline{x}) \quad (4-40)$$

gdje je  $V(\underline{x})$  funkcija elektrostatskog potencijala jezgre, a  $E$  je energija sustava. Rješenja jednadžbe (4-40) su valnog oblika:

$$\psi(\underline{x})=R_{nl}(r)y_{lm}(\theta,\phi); \quad n>l\geq|m|, \quad (4-41)$$

pri čemu (cijeli) kvantni brojevi  $n, l$  i  $m$  određuju dozvoljene elektronske putanje, tj. dozvoljena stanja atoma. Radijalne funkcije  $R_{nl}(r)$  su dane izrazom (Pauling & Wilson, 1935)

$$R_{nl}(r)=N_{nl}e^{-\rho/2}\rho^l L_{n-l}^{2l+1}(\rho) \quad (4-42)$$

gdje su  $L_{n-l}^{2l+1}(\rho)$  Laguerreovi polinomi, a  $\rho$  je skalirana udaljenost od središta atoma (protona). U osnovi, valna funkcija vodikovog atoma (4-42) se sastoji od eksponencijalno prigušene funkcije pomnožene Laguerreovim polinomom i sfernim harmonikom.  $N_{nl}$  je normalizacijski koeficijent odabran tako da su navedene radijalne funkcije ortogonalne s obzirom na polarni integral:

$$\int_0^\infty R_{nl}(r)R_{n'l}(r)r^2dr=\delta_{nn'} \quad (4-43)$$

Budući da udaljenost na kojoj eksponencijalno iščezavaju raste s kvadratom od  $n$ , ovakve se funkcije nisu pokazale pogodne za modeliranje površine proteina. Kod modeliranja površine proteina poželjno je da su te funkcije lokalizirane upravo oko ljuske proteina dok bi iza nje iščezavale. S druge strane, ove su se funkcije s eksponencijalnim prigušenjem pokazale dobre za modeliranje elektrostatskog naboja i potencijala u prostoru, tj. njihovo se prigušenje dobro poklapa s opadanjem potencijala u prostoru prilikom udaljavanja od izvora potencijala.

#### 4.5.3. Orbitale eksponencijalnog tipa

Kao što je prije spomenuto, za potrebe računanja elektrostatskog potencijala i prostorne razdiobe naboja koriste se neskaliране radijalne funkcije eksponencijalnog prigušenja

$$V_{nl}(r) = \left[ (2\Lambda)^3 \frac{(n-l-1)!}{(n+l+1)!} \right]^{1/2} e^{-\rho/2} \rho^l L_{n-l-1}^{(2l+2)}(\rho), \quad (4-44)$$

gdje je  $\rho = 2\Lambda r$ . Za potrebe prianjanja proteina korišten je  $\Lambda = 1/2$ . Navedene funkcije odgovaraju onima koje predstavljaju radijalan dio rješenja valne jednačbe za vodikov atom.

#### 4.6. 3-D parametrizacija

Kao što je na početku navedeno, proizvoljna se funkcija u trodimenzionalnom prostoru može prikazati u sfernim polarnim koordinatama kao

$$A(\underline{r}) = \sum_{nlm}^N a_{nlm} R_{nl}(r) y_{lm}(\theta, \phi); \quad N \geq n > l \geq |m| \geq 0, \quad (4-45)$$

te je pokazano da se zbog ortonormalnosti funkcija baze koeficijenti  $a_{nlm}$  računaju kao

$$a_{nlm} = \int \sigma(\underline{r}) F_{nlm}(\underline{r}) dV, \quad (4-46)$$

gdje  $F_{nlm}(\underline{r}) = R_n(r) y_{lm}(\theta, \phi)$  označava složenu funkciju baze koja je trodimenzionalna (sastoji se od radijalne funkcije ovisne o  $r$  i od kuglinih funkcija ovisnih o  $(\theta, \phi)$ ).

Budući da operator rotacije ne utječe na radijalne funkcije, pokazuje se da se trodimenzionalne funkcije baze  $F_{nlm}(\underline{r})$  rotiraju kao i dvodimenzionalne:

$$\begin{aligned} \widehat{R}(\alpha, \beta, \gamma) \sigma_N(\underline{r}) &= \sum_{nlm}^N \widehat{R}(\alpha, \beta, \gamma) [a_{nlm} R_{nl}(r) y_{lm}(\theta, \phi)] \\ &= \sum_{nlm}^N R_{nl}(r) \widehat{R}(\alpha, \beta, \gamma) [a_{nlm} y_{lm}(\theta, \phi)] \\ &= \sum_{nlm}^N R_{nl}(r) y_{lm}(\theta, \phi) a'_{nlm}, \end{aligned} \quad (4-47)$$

a koeficijenti se kao i prije, među sobom transformiraju prema

$$a'_{nlm} = \sum_{m'=-l}^l a_{nlm'} R_{mm'}^{(l)}(\alpha, \beta, \gamma). \quad (4-48)$$

## 5. Elektrostatska komplementarnost

Za točan proračun elektrostatske komplementarnosti kompleksa dvaju proteina, protein je potrebno promatrati kao materiju s niskom dielektričnom konstantom dok njegovu okolinu promatramo kao medij s velikom dielektričnom konstantom. Naveden bi pristup zahtijevao proračun prostornog elektrostatskog potencijala za svaku orijentaciju protein partnera te kao takav nije praktičan u računalnom pristupu rješavanja problema elektrostatske komplementarnosti. U svrhu pojednostavljenja problema, umjesto visoko dielektričnog medija kao što je voda, pretpostavit ćemo da se proteini nalaze u vakuumu.

### 5.1. Računanje elektrostatske energije vezanja kompleksa

Elektrostatska energija uzrokovana djelovanjem elektrostatskog potencijala  $\phi(\underline{r})$  na prostornu raspodjelu naboja  $\rho(\underline{r})$  dana je formulom:

$$E = \frac{1}{2} \int \rho(\underline{r})\phi(\underline{r})dV . \quad (5-1)$$

Budući da promatramo kompleks dva proteina (označimo ih s A i B), prostornu raspodjelu naboja i elektrostatski potencijal možemo zapisati kao

$$\rho(\underline{r}) = \rho_A(\underline{r}_A) + \rho_B(\underline{r}_B) , \quad (5-2)$$

odnosno

$$\phi(\underline{r}) = \phi_A(\underline{r}_A) + \phi_B(\underline{r}_B) , \quad (5-3)$$

te sad razmatramo utjecaj potencijalnog polja koje stvaraju naboji proteina A na naboje proteina B i obrnuto.

Ako prikažemo svaku funkciju razvojem u red ortogonalnih funkcija u sfernim polarnim koordinatama, tada dobivamo da je član koji opisuje doprinos energiji (5-1) zbog utjecaja elektrostatskog potencijala proteina B na naboje proteina A jednak:

$$\begin{aligned}
\int \rho_A(\underline{r})\phi_B(\underline{r})dV &= \int \left( \sum_{nlm}^N a_{nlm}^\rho V_{nl}(r_A)y_{lm}(\theta_A, \phi_A) \sum_{n'l'm'}^N b_{n'l'm'}^\phi V_{n'l'}(r_B)y_{l'm'}(\theta_B, \phi_B) \right) dV \\
&= \sum_{nlm}^N \sum_{n'l'm'}^N a_{nlm}^\rho b_{n'l'm'}^\phi \int (V_{nl}(r_A)y_{lm}(\theta_A, \phi_A)V_{n'l'}(r_B)y_{l'm'}(\theta_B, \phi_B)) dV \\
&= \sum_{nlm}^N \sum_{n'l'm'}^N a_{nlm}^\rho b_{n'l'm'}^\phi \int V_{nl}(r_A)V_{n'l'}(r_B)\mathcal{Y}_{|m|}(\mu_A)\mathcal{Y}_{|m'|}(\mu_B)\varphi_m(\phi_A)\varphi_{m'}(\phi_B)dV
\end{aligned}
\tag{5-4}$$

Integrali s desne strane slični su dvo-ishodišnim preklapajućim integralima koji se često pojavljuju u kvantnomehaničkim problemima. Oni ovise isključivo o razmaku između centara koordinatnih sustava u kojima su opisane pojedine ljuske proteina. Jedan mogući način računanja tih integrala je numerička integracija u eliptično – sferoidnom koordinatnom sustavu. Drugi način je evaluacija analitičke formule. Potonji način daje preciznije rezultate, no zahtijeva mnogo veću kompleksnost pa tako i vrijeme računanja na računalu. Ipak, budući da integrali u formuli (5-4) ne ovise o individualnim proteinima ili svojstvima ljusaka, oni se mogu izračunati samo jednom i spremiti za sve kasnije uporabe.

Zbog navedenih razloga, u ovom će radu biti upotrebljena evaluacija analitičke formule.

## 5.2. Integrali preklapanja kao elementi translacijskih matrica

Neka funkcija  $A(\underline{r})$  opisuje neko svojstvo tijela A (elektrostatski potencijal, prostorna raspodjela naboja) a funkcija  $B(\underline{r})$  svojstvo tijela B. Tijelo A neka miruje dok se tijelo B giba uzduž z-osi za udaljenost  $T = (R, 0, 0)$ . Tada odnos (preklapanje) ta dva tijela, tj. funkcija  $A(\underline{r})$  i  $B(\underline{r})$  možemo opisati kao:

$$C(\underline{T}) = \int A(\underline{r})B(\underline{T}^{-1}\underline{r})dV. \tag{5-5}$$

Supstitucijom  $A(\underline{r}) = \sum_{nlm}^N a_{nlm} R_{nl}(r)y_{lm}(\theta, \phi)$ , odnosno  $B(\underline{r}) = \sum_{n'l'm'}^N b_{n'l'm'} R_{n'l'}(r)y_{l'm'}(\theta, \phi)$  izraz

(5-5) prelazi u

$$C(\underline{T}) = \sum_{nlm} \sum_{n'l'm'} a_{nlm} b_{n'l'm'} \times \int R_{nl}(r) y_{lm}(\theta, \phi) R_{n'l'}(r') y_{l'm'}(\theta', \phi') dV \quad (5-6)$$

$$= \sum_{nlm} \sum_{n'l'm'} a_{nlm} b_{n'l'm'} \times \int R_{nl}(r) \mathcal{Y}_{l|m|}(\mu) \varphi_m(\phi) R_{n'l'}(r') \mathcal{Y}_{l'|m'|}(\mu') \varphi_{m'}(\phi') dV$$

gdje je  $\underline{r} = (r, \theta, \phi)$ , a  $\underline{r}' = (r', \theta', \phi')$ . Budući da transliramo tijelo B (funkciju  $B(\underline{r})$ ) isključivo u smjeru z-osi, kutevi  $\phi$  i  $\phi'$  ostaju jednaki tokom translacije. To pak znači da cirkularne funkcije  $\varphi_m(\phi)$  možemo integrirati izvan integrala. Zbog ortogonalnosti, njihov je prostorni integral jednak:

$$\int \varphi_m(\phi) \varphi_{m'}(\phi') dV = \delta_{mm'}, \quad (5-7)$$

pa se u izrazu (5-6) integracija vrši samo u  $(r, \theta)$  ravnini. Budući da integral u (5-6) ovisi samo o udaljenosti  $R$ , izraz (5-5) možemo zapisati kao

$$C(R) = \sum_{nlm} \sum_{n'l'm'} a_{nlm} b_{n'l'm'} T_{nl,n'l'}^{(|m|)}(R) \delta_{mm'} \quad (5-8)$$

i interpretirati svaki element  $T_{nl,n'l'}^{(|m|)}(R)$  kao element operatora translacije. Npr., iz jednadžbe (5-8) slijedi da sume

$$b_{nlm}^R = \sum_{n'l'} T_{nl,n'l'}^{(|m|)}(R) b_{n'l'm'}, \quad (5-9)$$

odnosno nakon preimenovanja indeksa,

$$a_{nlm}^R = \sum_{n'l'} T_{n'l',nl}^{(|m|)}(R) a_{n'l'm} \quad (5-10)$$

predstavljaju translaciju tijela B u smjeru pozitivne z-osi, odnosno (ekvivalentno) translaciju tijela A u smjeru negativne z-osi. Korištenjem Besselove transformacije i normalizirani radijalnih funkcija eksponencijalnog tipa, mogu se dobiti analitički izrazi za elemente translacijskih matrica (Ritchie, 2005):

$$T_{n'l',nl}^{(|m|)}(R) = \sum_{k=|l-l'|}^{l+l'} A_k^{(|m|)} \sum_{j=0}^{n-l+n'l'-2} D_j^{(nl,n'l')} \times \frac{2}{\pi} \int_0^\infty \frac{s^{2M+k}}{(s^2+1)^{J+2}} j_k(s\Delta R) s^2 ds \quad (5-11)$$

gdje je  $M = (l + l' - k) / 2$  i  $J = j + l + l' + 2$ . Za potrebe dokiranja s obzirom na elektrostatsku komplementarnost uzima se  $\Delta = 1 / 2$ .

Koeficijenti  $A_k^{(ll'm)}$  se računaju prema izrazu

$$A_k^{(ll'm)} = (-1)^{(k+l'-l)/2+m} (2k+1) [(2l+1)(2l'+1)]^{1/2} \times \begin{pmatrix} l & l' & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l' & k \\ m & \bar{m} & 0 \end{pmatrix}. \quad (5-12)$$

Posljednja dva simbola su Wignerovi 3-j simboli, često korišteni u kvantnoj mehanici.

Koeficijenti  $D_k^{(nl,n'l')}$  dani su s

$$D_k^{(nl,n'l')} = \sum_j \sum_{j'} \delta_{k,j+j'} Y_{nlj} Y_{n'l'j'}, \quad (5-13)$$

pri čemu je

$$Y_{nlj} = \left[ \frac{1}{2} \frac{(n-l-1)!}{(n+l+1)!} \right]^{1/2} \frac{(-1)^j (2n+1)(n+l+j+1)!}{j!(n-l-j-1)!(1/2)_{l+j+2}}, \quad (5-14)$$

i

$$(x)_n = x(x+1)\dots(x+n-1) = \Gamma(x+n) / \Gamma(x), \quad (5-15)$$

gdje  $\Gamma(x)$  označava gamma funkciju.

Pokazuje se da se integral u (5-11) može računati kao:

$$\frac{2}{\pi} \int_0^\infty \frac{s^{2M+k}}{(s^2+1)^{J+2}} j_k(s\Delta R) s^2 ds = \sum_{q=0}^M \binom{M}{q} \frac{(-1)^{M+q}}{2^{J+1-q} (J+1-q)!} (\Delta R)^k \hat{k}_{J-k-q+1/2}(\Delta R). \quad (5-16)$$

gdje je  $\hat{k}_\sigma(z)$  reducirana Besselova funkcija drugog reda. Mogu se računati rekurzivno (Weniger & Steinborn, 1983):

$$\hat{k}_{1/2}(z) = \exp(-z), \quad (5-17)$$

$$\hat{k}_{3/2}(z) = (1+z) \exp(-z), \quad (5-18)$$

i

$$\hat{k}_{n+3/2}(z) = (2n+1) \hat{k}_{n+1/2}(z) + z^2 \hat{k}_{n-1/2}(z). \quad (5-19)$$



### 5.3. Konačna formula za proračun elektrostatske energije vezanja kompleksa

Nakon što su definirani svi alati za proračun elektrostatske komplementarnosti, konačna formula za proračun elektrostatske komplementarnosti dvaju proteina glasi:

$$E(R, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2) = \frac{1}{2} \sum_{nlm}^N \sum_{n'l'm'}^N \left( a'_{nlm}{}^\rho b'_{n'l'm'}{}^\phi + a'_{nlm}{}^\phi b'_{n'l'm'}{}^\rho \right) J_{nn'|l'l'|m|}(R) \delta_{mm'}, \quad (5-20)$$

gdje su  $a'_{nlm}{}^\rho$  i  $a'_{nlm}{}^\phi$  rotirani koeficijenti gustoće naboja i elektrostatskog potencijala proteina A (slično vrijedi i za protein B), a  $J(R)$  je matrica preklapajućih integrala (translacijska matrica) izračunata pomoću neskaliranih, eksponencijalno prigušenih radialnih funkcija.

Iz formule (5-20) je vidljivo da se proces dokiranja vrši u šest-dimenzionalnom prostoru. Pet stupnjeva slobode čine kutevi za koje rotiramo protein A (receptor) i protein B (ligand). Oba proteina rotiramo oko x i y osi dok oko z-osi rotiramo samo ligand. Razlog tome je simetričnost rotacije oko z-osi budući da se središta masa oba proteina nalaze na z osi. Šesti stupanj slobode predstavlja međumolekularna udaljenost R, odnosno udaljenost među središta masa proteina A, odnosno proteina B.

Uzimajući beskonačan broj članova reda, jednačba (5-20) bi bila potpuno egzaktna, no kao što je prije spomenuto, razvoj je ograničen na  $N \leq 25$  kako bi se postigao kompromis između preciznosti proračuna i brzine izvođenja samog procesa.

#### 5.4. Implementacija algoritma

Formula (5-20) izgleda jednostavno, no operacije koje je potrebno izvršiti nad koeficijentima prostorne gustoće naboja i elektrostatskog potencijala su često komplicirane, a samim time i spore. Najčešće su razlog tome ugnježdene petlje. Stvar postaje kompliciranija što želimo veću preciznost reprezentacije gustoće naboja, odnosno elektrostatskog potencijala. Stoga je potrebno pažljivo odabrati redoslijed i način operacija nad koeficijentima kako bi algoritam bio što brži i efikasniji.

Koristeći svojstva rotacija sfernih harmonika, koeficijente prostorne gustoće naboja možemo rotirati pomoću izraza

$$a'_{nlm} = \sum_{m'=-l}^l a_{nlm'}^{\rho} R_{mm'}^{(l)}(0, \beta_1, \gamma_1) \quad (5-21)$$

za protein A (receptor), odnosno pomoću izraza

$$b'_{nlm} = \sum_{m'=-l}^l b_{nlm'}^{\rho} R_{mm'}^{(l)}(0, \beta_2, \gamma_2) \quad (5-22)$$

za protein B (ligand).

Prikladno je zakretni kut  $\alpha_2$  (kut oko z-osi kojim zakrećemo ligand) računati posebno, budući da rotacija oko z-osi ima jednostavnu formu:

$$B'_{nlm} = b'_{nlm} \cos m\alpha_2 + b'_{nl\bar{m}} \sin \bar{m}\alpha_2 \quad (5-23)$$

Slične relacije vrijede za rotaciju koeficijenata elektrostatskog potencijala. Pomoću formule (5-20) se može izravno računati ukupna elektrostatska energija vezanja za sve dane orijentacije. Međutim, algoritam dan jednadžbom (5-20) složenosti je  $O(N^5)$ , te kao takav nije prikladan za izravnu uporabu. Račun je prikladnije provesti u dva dijela preslaganjem sumacija. Tj., izraze u (5-20) možemo napisati kao:

$$\sum_{nlm}^N \sum_{n'l'm'}^N (a'_{nlm} B'_{n'l'm'}) J_{nn'l'l'|m|}(R) \delta_{mm'} = \sum_{n'l'm'}^N \left( \sum_{nlm}^N a'_{nlm} J_{nn'l'l'|m|}(R) \delta_{mm'} \right) B'_{n'l'm'}, \quad (5-24)$$

odnosno

$$\sum_{nlm}^N \sum_{n'l'm'}^N \left( a'_{nlm}{}^\phi B'_{n'l'm'}{}^\rho \right) J_{m'l'|m|}(R) \delta_{mm'} = \sum_{n'l'm'}^N \left( \sum_{nlm}^N a'_{nlm}{}^\phi J_{m'l'|m|}(R) \delta_{mm'} \right) B'_{n'l'm'}{}^\rho, \quad (5-25)$$

gdje unutarnje sume označavamo kao modificirane koeficijente proteina A (receptora) koji označavaju rotirane i translirane koeficijente:

$$A'_{n'l'm'}{}^\rho(R) = \sum_{nlm}^N a'_{nlm}{}^\rho J_{m'l'|m|}(R) \delta_{mm'}, \quad (5-26)$$

odnosno

$$A'_{n'l'm'}{}^\phi(R) = \sum_{nlm}^N a'_{nlm}{}^\phi J_{m'l'|m|}(R) \delta_{mm'}. \quad (5-27)$$

Koristeći ovaj međurezultat, algoritam postaje složenosti  $O(N^3)$  za svaki rotacijski inkrement proteina B:

$$E(R, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2) = \frac{1}{2} \sum_{nlm}^N \left( A'_{nlm}{}^\rho(R) B'_{nlm}{}^\phi + A'_{nlm}{}^\phi(R) B'_{nlm}{}^\rho \right). \quad (5-28)$$

Za  $R=0$ , jednačba (5-28) predstavlja elektrostatsku energiju vezanja kompleksa koji se nalaze u istom koordinatnom sustavu. Dakle, svaki modificirani koeficijent  $A'_{nlm}$  može se interpretirati kao efektivna vrijednost određenog svojstva proteina A (receptora) koje „vidi“ protein B (ligand). Unatoč tome što za svaku orijentaciju proteina A koeficijenti  $A'_{nlm}{}^\rho$  i  $A'_{nlm}{}^\phi$  moraju biti ponovno izračunati, redukcija složenosti s  $O(N^5)$  na  $O(N^3)$  znatno ubrzava proces što posebno dolazi do izražaja za velik red  $N$ .

U procesu dokiranja pretražuje se šesterodimenzionalan prostor, te je potrebno pažljivo odrediti poredak izvršavanja pojedinih operacija. Poredak je dan u sljedećim koracima (Ritchie & Kemp, 2000).

Izraz (5-28) možemo zapisati kao:

$$E(R, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2) = \frac{1}{2} \sum_{nlm}^N \left( A_{nlm}^{\prime\rho}(R) (b_{nlm}^{\prime\phi} \cos m\alpha_2 + b_{nlm}^{\prime\phi} \sin \bar{m}\alpha_2) + A_{nlm}^{\prime\phi}(R) (b_{nlm}^{\prime\rho} \cos m\alpha_2 + b_{nlm}^{\prime\rho} \sin \bar{m}\alpha_2) \right) \quad (5-29)$$

Sumirajući članove uz  $\cos m\alpha_2$  i  $\sin \bar{m}\alpha_2$  dobiva se:

$$E(R, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2) = \sum_{m=-L}^L Q_m^+ \cos m\alpha_2 + Q_m^- \sin \bar{m}\alpha_2; \quad L = N - 1 \quad (5-30)$$

gdje su

$$Q_m^+(R, \beta_1, \gamma_1, \beta_2, \gamma_2) = \sum_{nlm'}^N \left( A_{nlm'}^{\prime\rho}(R) b_{nlm'}^{\prime\phi} + A_{nlm'}^{\prime\phi}(R) b_{nlm'}^{\prime\rho} \right) \delta_{mm'} \quad (5-31)$$

$$Q_m^-(R, \beta_1, \gamma_1, \beta_2, \gamma_2) = \sum_{nlm'}^N \left( A_{nlm'}^{\prime\rho}(R) b_{nlm'}^{\prime\phi} + A_{nlm'}^{\prime\phi}(R) b_{nlm'}^{\prime\rho} \right) \delta_{mm'} \quad (5-32)$$

Konačno, korištenjem identiteta  $\cos \bar{m}\phi = \cos m\phi$  i  $\sin \bar{m}\phi = -\sin m\phi$ , jednadžba (5-30)

se pojednostavnjuje i prelazi u realan Fourierov red po  $\alpha_2$ :

$$E(R, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2) = Q_0^{++} + \sum_{m=1}^L Q_m^{++} \cos m\alpha_2 + Q_m^- \sin m\alpha_2, \quad (5-33)$$

gdje su

$$Q_m^{++} = Q_m^+ + Q_m^+ (1 - \delta_{m0}) \quad (5-34)$$

$$Q_m^- = Q_m^- - Q_m^+ \quad (5-35)$$

U usporedbi s diskretnom FFT metodom, izraz (5-33) predstavlja kontinuiranu funkciju ovisnu o kutu  $\alpha_2$ . Koeficijenti  $Q_m^{++}$  i  $Q_m^-$  ovise samo o međumolekularnoj udaljenosti,  $R$  i rotiranim koeficijentima. Nakon što se odrede svi potrebni koeficijenti ovisni o pojedinoj orijentaciji  $(R, \beta_1, \gamma_1, \beta_2, \gamma_2)$ , uzastopnim inkrementiranjem  $\alpha_2$  vrlo je brzo moguće izračunati energiju elektrostatskog vezanja kompleksa. Najzahtjevniji računski dio uključuje računanje koeficijenata gustoće naboja i elektrostatskog potencijala te nakon toga, ponovo računanje koeficijenata  $Q_m$  za pojedinu orijentaciju.

### 5.4.1. Pretraživanje prostora interakcija

Kako bi šesterodimenzionalni prostor u kojem se odvija proces dokiranja pretražili jednoliko i ekonomično, potrebno je pažljivo odrediti kuteve rotacija. Za generiranje parova rotacijskih kuteva  $(\beta_1, \gamma_1)$  i  $(\beta_2, \gamma_2)$  koristi se set funkcija koje ravnomjerno dijele rotacijski prostor tehnikom *icosahedral tessellation*. Ona polazi od ikosaedra (pravilnog poliedra sa 20 jednakostraničnih trokuta kao stranicama) i za svaki red podjele prepolavlja bridove pri čemu od svake stranice nastaju četiri nova trokuta. Takva podjela početnog ikosaedra dovodi do površine koja za sve veće redove podjele  $k$  sve bolje aproksimira sferu, pa se taj postupak koristi i za ravnomjernu raspodjelu površine sfere. Eulerovi  $\beta$  i  $\gamma$  kutevi pravaca definiranih ishodištem i centrom tako nastalih stranica, uzimaju se kao parovi rotacijskih kuteva  $(\beta_1, \gamma_1)$  i  $(\beta_2, \gamma_2)$ . Kut  $\alpha_2$  poprima vrijednosti ravnomjerne podjele punog kuta od  $2\pi$  radijana. Dakle, rotacije proteina A i B za kutove  $(\beta_1, \gamma_1)$ , odnosno  $(\beta_2, \gamma_2)$  te dodatno zakretanje proteina B oko z-osi čine 5 stupnjeva slobode. Preostali stupanj slobode čini međumolekularna udaljenost,  $R$ . Navedene rotacije, tj. „kutno pretraživanje“ ponavlja se za pojedini  $R$  čiji je inkrement najčešće  $0.5 \text{ \AA}$ .

### 5.4.2. Pseudokod

Pseudokod programa sastoji se od niza manipulacija nad vektorima koeficijenata gustoće naboja, odnosno elektrostatskog potencijala. Svaka reprezentacija tih veličina sadrži (vidi (4-2))  $\sum_{n=1}^N n^2 = N(N+1)(2N+1)/6$  koeficijenata. Npr., uzimajući red razvoja  $N = 10, 15, 20, 25$ , dobivamo vektore koji sadrže 385, 1240, 2870, odnosno 5525 elemenata. Nad vektorima tih veličina se potom odvojeno vrše uzastopne rotacije odvojene od glavne petlje gdje se izračunava energija elektrostatskog vezanja. Nakon toga se računa elektrostatska energija vezanja koja ovisi samo o kutu  $\alpha_2$  (budući da su koeficijenti koji se koriste za proračun prethodno rotirani za  $(\beta_1, \gamma_1)$ , odnosno  $(\beta_2, \gamma_2)$  s obzirom na neku određenu međumolekularnu udaljenost  $R$ ). Pritom gledamo najmanju energiju za pojedini kut  $\alpha_2$  te ju spremamo pod jedinstvenim ključem  $(R, \beta_1, \gamma_1, \beta_2, \gamma_2)$ . Pseudokod je dan u nastavku teksta (Ritchie, 1998):

## početak 6D Docking

**za svaki  $R$**

učitaj sve  $J(R)$  preklapajuće integrale

**za svaki par  $(\beta_1, \gamma_1)$**

izračunaj  $a'^{\rho}$ ,  $a'^{\phi}$  koristeći  $a'_{nlm} = \sum_{m'=-l}^l a_{nlm} R_{mm'}^{(l)}(0, \beta_1, \gamma_1)$

izračunaj  $A'^{\rho}(R)$ ,  $A'^{\phi}(R)$  koristeći  $A'_{n'l'm'}(R) = \sum_{nlm} a'_{nlm} J_{nn'|l||m|}(R) \delta_{mm'}$

**za svaki par  $(\beta_2, \gamma_2)$**

izračunaj  $b'^{\rho}$ ,  $b'^{\phi}$  koristeći  $b'_{nlm} = \sum_{m'=-l}^l b_{nlm} R_{mm'}^{(l)}(0, \beta_2, \gamma_2)$

izračunaj  $Q^+$  koristeći  $Q_m^+(R, \beta_1, \gamma_1, \beta_2, \gamma_2) = \sum_{nlm'} (A'_{nlm'}{}^{\rho}(R) b'_{nlm'}{}^{\phi} + A'_{nlm'}{}^{\phi}(R) b'_{nlm'}{}^{\rho}) \delta_{mm'}$

izračunaj  $Q^-$  koristeći  $Q_m^-(R, \beta_1, \gamma_1, \beta_2, \gamma_2) = \sum_{nlm'} (A'_{nlm'}{}^{\rho}(R) b'_{nl\bar{m}}{}^{\phi} + A'_{nlm'}{}^{\phi}(R) b'_{nl\bar{m}}{}^{\rho}) \delta_{mm'}$

izračunaj  $Q^{++}$  koristeći  $Q_m^{++} = Q_m^+ + Q_{\bar{m}}^+(1 - \delta_{m0})$

izračunaj  $Q^{--}$  koristeći  $Q_m^{--} = Q_{\bar{m}}^- - Q_m^+$

**za sve  $\alpha_2$**

izračunaj  $e[\alpha_2] = Q_0^{++} + \sum_{m=1}^{N-1} (Q_m^{++} \cos m\alpha_2 + Q_m^{--} \sin m\alpha_2)$

**kraj petlje**

pohrani  $E[Id] = \max_{\alpha_2} (e[\alpha_2])$

**kraj petlje**

**kraj petlje**

**kraj petlje**

**kraj**

## 5.5. Računanje koeficijenata

Prije nego što se pokrene program za računanje elektrostatske energije vezanja kompleksa, potrebno je izračunati ulazne koeficijente, tj. koeficijente reprezentacije prostorne gustoće naboja i elektrostatskog potencijala.

U tu svrhu koriste se jednažbe (za protein A)

$$a_{nlm}^{\rho} = \sum_i q_i V_{nl}(r_i) y_{lm}(\theta_i, \phi_i) \quad (5-36)$$

za koeficijente reprezentacije prostorne gustoće naboja, odnosno

$$a_{nlm}^{\varphi} = \sum_i \varphi_i V_{nl}(r_i) y_{lm}(\theta_i, \phi_i) \quad (5-37)$$

za koeficijente reprezentacije elektrostatskog potencijala. Slično vrijedi i za protein B.

Dakle, iz jednažbi (5-36) i (5-37) proizlazi da je za određivanje koeficijenata dovoljno poznavati iznos i sferne koordinate naboja, odnosno potencijala u prostoru.

### 5.5.1. Računanje koeficijenata prostorne gustoće naboja

PDB format nije prikladan za računanje prostorne gustoće naboja jer se u njemu ne nalaze podatci o količini naboja atoma. Zbog toga je korišten program PDB2PQR koji automatizira mnoge zadatke oko pripreme struktura za elektrostatske proračune, od kojih je najvažnija pretvorba PDB datoteka u PQR format. U PQR formatu nalaze se svi podatci potrebni za izračun koeficijenata: iznos naboja pojedinog atoma i koordinate na kojima se naboj nalazi.

### 5.5.2. Računanje koeficijenata elektrostatskog potencijala

U svrhu računanja elektrostatskog potencijala korišten je program PotCalc. Kao ulaz, PotCalc prima PDB datoteku, konfiguracijsku datoteku i datoteku s koordinatama u kojima se treba izračunati elektrostatski potencijal. Konfiguracijska datoteka sadrži niz opcija koje se mogu podešavati, te tako utjecati na konačni izlaz programa kao što je npr. oblik zapisa rezultata.

Najpoželjnije bi naravno bilo izračunati elektrostatski potencijal na čitavom prostoru koji zauzima protein partner, no često se ne zna tko je protein partner, odnosno ako se i zna, vjerojatno je nepoznata njegova orijentacija u odnosu na promatrani protein, tj. u odnosu na protein čiji se potencijal treba izračunati. U tu svrhu potrebno je konstruirati dovoljno veliku rešetku oko promatranog proteina, tj. skup trodimenzionalnih koordinata u kojima treba izračunati elektrostatski potencijal. Također, dodatna činjenica da elektrostatske sile djeluju i na većim udaljenostima iziskuje konstrukciju velike rešetke. Točke konstruirane rešetke je potrebno zapisati u datoteku i predati programu PotCalc kao jedan od ulaznih parametara.

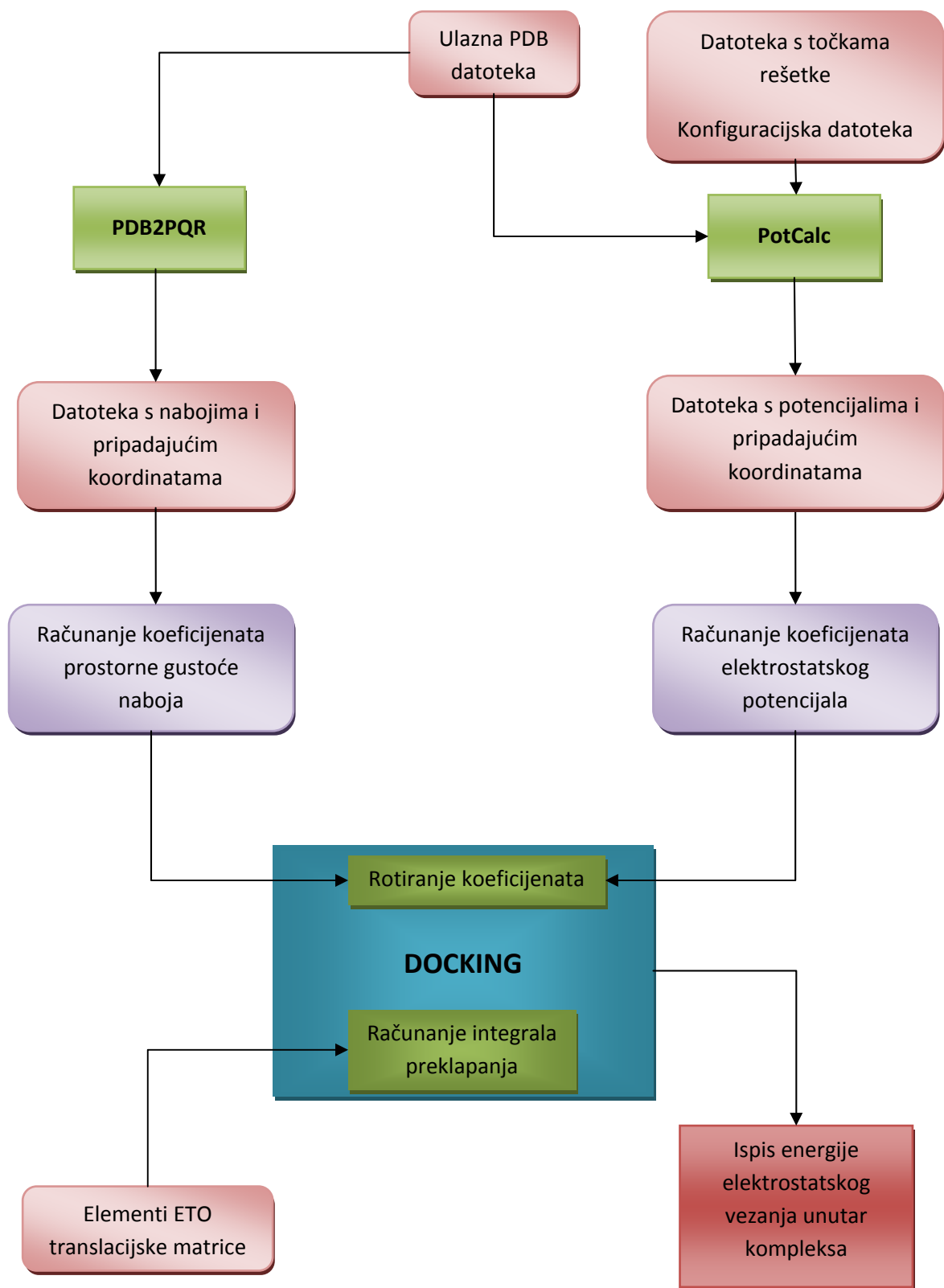
Nakon izračuna potencijala, kao izlaz iz PotCalca dobiva se datoteka u kojoj su vrijednosti odvojene točka-zarezom. Dakle, redci u datoteci su oblika

```
X; Y; Z; Potencijal
```

Koordinate koje se nalaze u PQR datoteci, odnosno u datoteci s potencijalima su kartezijske, stoga ih je prije izračuna koeficijenata potrebno pretvoriti u sferne. Također, iz izraza (5-36), (5-37) se vidi da svako računanje koeficijenata iziskuje sve ukupno  $N(N+1)(2N+1)/6$  prolazaka kroz PQR datoteku, odnosno datoteku s potencijalima. Budući da je čitanje iz datoteke sporo, prikladnije je datoteke pročitati jednom, te u posebne mape spremiti vrijednost naboja, odnosno potencijala s koordinatama kao jedinstvenim ključevima.

Nakon što su svi koeficijenti izračunati, pogodno ih je zapisati u binarnu datoteku koja predstavlja jedan od ulaza u proces prijanjanja. Na taj način postiže se odvojenost računanja koeficijenata i procesa prijanjanja što također može doprinjeti paralelizaciji cjelokupnog procesa.





**Slika 1:** Dizajn sustava za izračun energije elektrostatskog vezanja unutar kompleksa

## 6. Rezultati i diskusija

Kako bi rezultati cjelokupnog prianjanja bili vidljivi, ocjenu jačine elektrostatskog vezanja kompleksa potrebno je dodati onoj s obzirom na komplementarnost površina. Tada je ukupna energija vezanja kompleksa jednaka:

$$E_{ukupno} = \left( \frac{1391.4}{K_R} \right) E + K_H S, \quad (6.1)$$

Gdje  $K_R$  označava relativnu permitivnost, a  $K_H$  faktor hidrofobne slobodne energije. Navedeni se faktori podešavaju za pojedini kompleks. Evaluacija izraza (6.1) s visokom rezolucijom  $N$ , bila bi nadasve spora. Povoljnije je najprije evaluirati energiju komplementarnosti površina  $K_H S$  s nešto nižom rezolucijom ( $N=16$ ). Dakle, za svaku parcijalnu orijentaciju  $(R, \beta_1, \gamma_1, \beta_2, \gamma_2)$  računa se minimalna energija  $K_H S$  s obzirom na  $\alpha_2$  kao što je opisano u algoritmu. Sve orijentacije za koje je  $K_H S \geq 100$  KJ/mol automatski se odbacuju. Nakon toga, za 25% preostalih orijentacija računa se *ukupna* energija  $E_{ukupno}$ , no ovaj put s visokom rezolucijom ( $N=25$ ). Dakle, orijentacije su najprije „filtrirane“ na način da se niskom rezolucijom ispitala samo komplementarnost površina, dok se nad manjim brojem preostalih orijentacija računa ukupna energija  $E_{ukupno}$  u koju je uračunata i energija elektrostatske komplementarnosti. Budući da komplementarnost površina nije u okviru ovog rada, rezultati nisu dani eksplicitno.

U ovom je radu pretpostavljeno da se proteini nalaze u vakuumu. Preciznije bi bilo naravno, uzeti u obzir visoko dielektrični medij u kojem se proteini nalaze, no to bi zahtijevalo računanje potencijala u prostoru za svaku pojedinu orijentaciju. Navedeni pristup mnogo je veće složenosti od trenutnog i kao takav nije praktično uporabljiv. Ipak, jedna od mogućih nadogradnja bi uključivala i takav složen pristup, no samo na nekoliko najboljih kompleksa. Time bi bilo moguće vidjeti koliko se razlikuju „točan“ i „približan“ model prianjanja.

## 7. Zaključak

Budući da se kugline funkcije među sobom rotacijom transformiraju na predvidljiv način, a radijalne funkcije dobro opisuju opadanje potencijala u prostoru, pokazuje se da je prostornu gustoću naboja i elektrostatskog potencijala proteina pogodno opisati razvojem realnih ortogonalnih radijalnih i kuglinih funkcija u Fourierov red.

U okviru ovog završnog rada izrađene su računalne implementacije kuglinih funkcija i ETO radijalnih funkcija. Pomoću programa PDB2PQR proteinima su dodijeljeni naboji, a pomoću programa PotCalc u prostoru oko proteina izračunat je elektrostatski potencijal. Implementirani su moduli koji naboje i elektrostatski potencijal razvijaju u red kuglinih funkcija i ETO radijalnih funkcija. Implementirane su ETO translacijske matrice čija je ispravnost provjerena uvjetom ortogonalnosti. Time je omogućena ocjena jačine elektrostatskog vezanja kompleksa, te je pomoću nje moguće korigirati rezultate dobivene ocjenom s obzirom samo na komplementarnost površina.

## 8. Literatura

D.W. RITCHIE, G.J.L. KEMP, Protein Docking Using Spherical Polar Fourier Correlations, *PROTEINS : Structure, Function and Genetics, preprint* (1999).

D.W. RITCHIE, High-order analytic translation matrix elements for real-space six-dimensional polar Fourier correlations (2005).

D.W. RITCHIE, Parametric Protein Shape Recognition, *Doktorski rad*, University of Aberdeen, Departments of Computing Science & Molecular and Cell Biology, Aberdeen, 1998.

IVAN DOKMANIĆ, Algoritam za predviđanje strukture proteinskih kompleksa korištenjem razvoja površine u redove funkcija, *Diplomski rad*, Fakultet Elektrotehnike i Računarstva, Zagreb, 2007.

MILJENKO CRNKOVIĆ, Program za računanje elektrostatskog potencijala na površini makromolekula (2007), *Diplomski rad*. Fakultet Elektrotehnike i Računarstva, Zagreb, 2007.

ALEXANDER HEIFETZ, EPHRAIM KATCHALSKI-KATZIR and MIRIAM EISENSTEIN, Electrostatics in protein-protein docking (2002).

NOGA KOWASLMAN, MIRIAM EISENSTEIN, Inherent Limitations in Protein-Protein Docking Procedures (2006).

JUAN FERNANDEZ-RECIO, MAXIM TOTROV, RUBEN ABAGYAN, Screened charge electrostatic model in protein-protein docking simulations (2002).

## 9. Naslov, sažetak i ključne riječi

### *Modul za računanje prijanjanja proteina koristeći elektrostatski potencijal*

Mogućnost brzog i vjerodostojnog utvrđivanja komplementarnosti određenih svojstava proteina od velikog je značaja za računalno modeliranje proteinskih interakcija. Glavno svojstvo koje se proučava u tu svrhu jest komplementarnost površina dvaju proteina, tj. u kojoj se mjeri preklapaju površine dvaju proteina. Iako glavno, ovo svojstvo nije dovoljno da bi zadovoljavajuće opisalo interakciju dvaju proteina. Jedna od nadogradnji navedenog modela postiže se uzimanjem u obzir elektrostatske komplementarnosti kompleksa dvaju proteina. Zbog svoje jakosti, elektrostatske su interakcije veoma važne za razumijevanje makromolekularnih interakcija. Uzimajući u obzir veoma male udaljenosti na kojima se interakcije odvijaju, teoretski gledano, ovisnost o kvadratu udaljenosti kojem je elektrostatska sila obrnuto proporcionalna jamči veliku ulogu tih sila na svim udaljenostima na kojima su interakcije moguće.

*Ključne riječi:* protein, prijanjanje, receptor, ligand, proteinska baza podataka, prostorna gustoća naboja, elektrostatski potencijal, kugline funkcije, radijalne funkcije, ortogonalnost, orbitale eksponencijalnog tipa

### *Protein docking module using electrostatic potential*

The possibility of a fast and reliable determination of certain protein properties complementarity is of great value for the computer modeling of protein interactions. The main characteristic studied for that purpose is area complementarity of two proteins, or in what measure do areas of two proteins overlap. Although it is a major one, this characteristic is insufficient to satisfyingly describe an interaction between two proteins. One of the upgrades of the mentioned model is achieved by considering the electrostatic complementarity of a two protein complex. Because of their power, electrostatic interactions are extremely important for understanding macromolecular interactions. By taking into consideration very small distances upon which the interactions run, theoretically, the dependance of the square distance to which the electrostatic force is inversely proportionate guaranties a big role to those forces at all distances on which interactions are possible.

*Keywords:* protein, docking, receptor, ligand, protein data bank, charge density, electrostatic potential, spherical harmonic functions, radial functions, orthonormality, exponential tipe orbitals