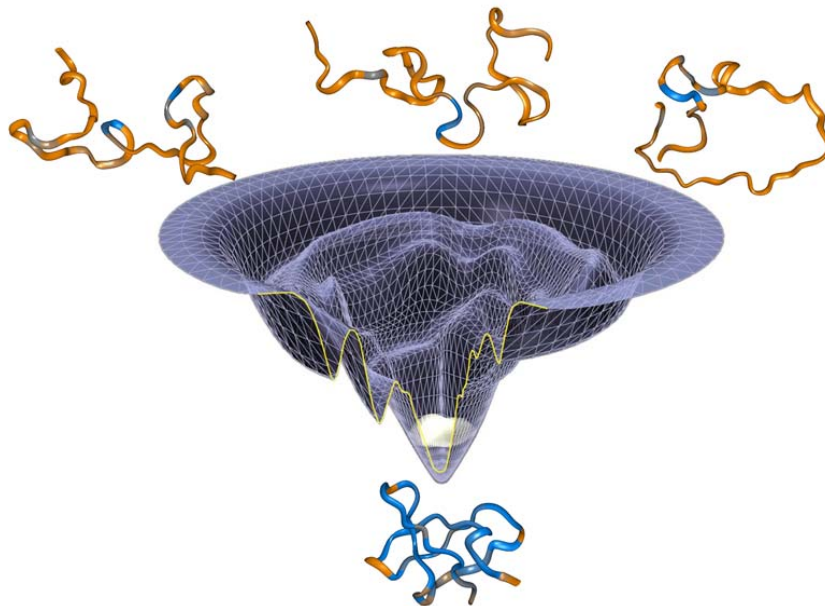


Univerza v Ljubljani
Fakulteta za matematiko in fiziko
Oddelek za fiziko

Seminar pri predmetu Fizika mehkih snovi

Statistična mehanika zvijanja proteinov



Aljaž Godec
Kemijski Inštitut, Ljubljana

Ljubljana, 2008

Povzetek

V seminarju bomo na kratko opredelili zapleteno strukturo proteinov in predstavili fizikalni pristop, s katerim jo lahko napovedujemo iz AK zaporedja. Prikazali bomo pojmovna razlika med proteini in naključnimi polipeptidi, ki so prav tako sestavljenih iz zaporedij AK, in razloženo dejstvo, da so proteini evolucijsko izbrane polipeptidne sekvence. Predstavili bomo osnove statistične mehanike proteinskih makromolekul in gonilne sile zvijanja proteinov. Elektrostatsko komponento proste energije solvatacije bomo opisali z analitičnim posplošenim Bornovim modelom. Predstavili bomo tudi teorijo posplošenega molekularnega polja (GMFA), najprej splošno, nato še aplicirano na opis hidrofobnega efekta, tako na molekularni kot tudi na mezo in makroskopski skali. V zadnjem odseku strnemo nekaj ugotovitev računalniških simulacij o prispevkih posameznih tipov interakcij k prosti energiji zvijanja in stabilnosti nativnega stanja. V dodatku izpeljemo Yvon-Born-Green formalizem tekočin, ki je osnova za GMFA teorijo.

Kazalo

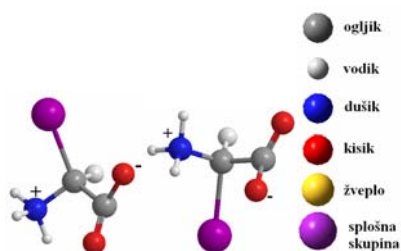
1. Uvod	4
1. 1 Zgradba, struktura in nativno stanje	4
2. Fizika proteinov	6
2. 1 Molekularno-mehanski model proteina	7
2. 2 Prosta energija zvijanja proteina in termodinamska stabilnost nativnega stanja	10
2. 3 Intramolekularne interakcije	11
2. 4 Eksplicitno toplilo	13
2. 5 Implicitno toplilo	13
2. 5. 1 Analitični posplošeni Bornov model	13
2. 5. 2 Teorija hidrofobnega efekta	18
2. 5. 2. 1 Generalized mean field (GMF) teorija solvatacije sferičnega delca v LJ tekočini	19
2. 5. 2. 2 Hidratacija hidrofobnega sferičnega delca, hidrofobne interakcije v vodi in implikacije na zvijanje proteinov	26
3. Pripisevek posameznih intramolekularnih interakcij k prosti energiji zvijanja in stabilnosti nativne konformacije	29
4. Zaključek	30
Dodatek: Yvon-Born-Green formalizem tekočin	31
Literatura	34

1. Uvod

Proteini so biološke makromolekule, ki jih najdemo v vseh živih bitjih. Od vseh skupin bioloških molekul (ogljikovih hidratov, lipidov nukleinskih kislin) so proteini najbolj raznolika skupina. Kljub temu da so sestavljeni iz skromnega nabora osnovnih gradnikov (aminokislin), izkazujejo osupljivo sposobnost tvorbe različnih struktur z visoko specifičnimi katalitičnimi mesti na površini. Posledično proteini na različnih fizioloških nivojih izvajajo vse zahtevne in zapletene procese v živih bitjih.

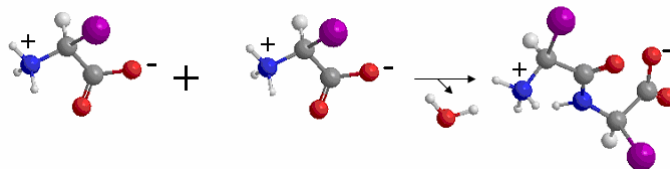
1.1 Zgradba, struktura in nativno stanje

Kljub strukturalni raznolikosti so proteini s kemijskega stališča homogena skupina. So linearni heteropolimeri, sestavljeni iz aminokislin (AK). Aminokislina je sestavljena iz centralnega α -C atoma, na katerega so vezane kislina karboksilna skupina (-COOH), bazična amino skupina (-NH₂) in kislina, bazična ali nevtralna stranska veriga (-R_s). V fizioloških pogojih sta tako amino kot tudi karboksilna skupina ionizirani (slika 1).



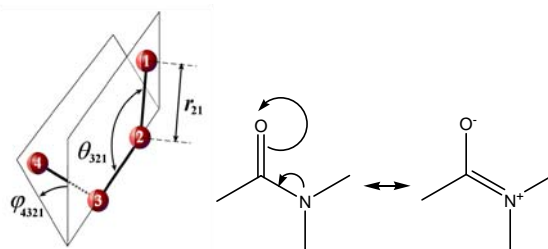
Slika 1: Splošna struktura L- (na sliki levo) in D- (na sliki desno) aminokislin in legenda posameznih vrst atomov.

Če stranska veriga -R_s ni enaka kateri od ostalih treh (kot pri glicinu, kjer je -R_s=H), pravimo, da je α -C atom asimetričen in sta možni dve konfiguraciji, L in D (glej sliko 1). Z izjemo določenih eksotičnih organizmov so proteini v živih bitjih sestavljeni le iz L-aminokislin. Vzrok za to do danes ni znan. Genetski kod živih bitij kodira 20 različnih aminokislin. Aminokislina se preko peptidnih vezi povezujejo v polipeptidne verige. Nastanek dipeptida je prikazan na sliki 2.



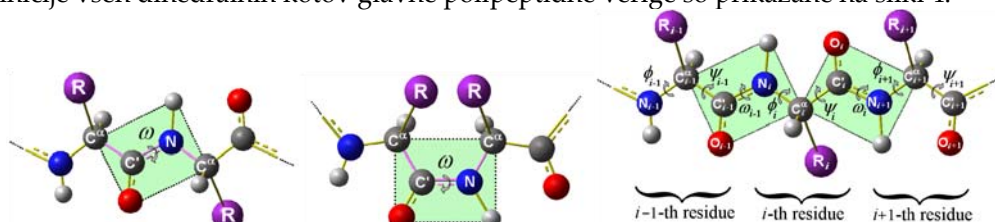
Slika 2: Nastanek dipeptida.

Kot je razvidno iz naslova, nas zanima predvsem zvijanje proteinov, čemur se bomo podrobneje posvetili v nadaljevanju, kjer bomo s pojmom protein označevali polipeptidno verigo iz L-aminokislin brez posttranslacijskih modifikacij, ki bo locirana v čisti vodi. Kovalentna struktura (AK zaporedje) in stereoizomerija (vse AK so L- in konfiguracije na asimetričnih centrih v stranskih verigah določene) sta določeni, kar je pri manjših rigidnih molekulah že dovolj za določitev tridimenzionalne strukture. Polimeri, še posebej pa proteini, imajo prostostne stopnje, ki jim pravimo, da so mehke (soft-mode), saj majhne spremembe energije povzročijo drastične spremembe v pozicijah atomov in funkcionalnih skupin. Taka spremenljivka so dihedralni koti (slika 3- levo), ki opisujejo rotacije okoli enojnih vezi. V proteinih so nekateri mehki dihedralni koti locirani na stranskih verigah. Sprememba teh kotov spremeni strukturo le lokalno, medtem ko lahko sprememba dihedralnih kotov v polipeptidni verigi drastično spremeni relativne pozicije velikega števila parov atomov, zato jim moramo posvetiti posebno pozornost.



Slika 3: (levo): Tipična definicija notranjih koordinat: r_{21} je dolžina vezi med atomoma 2 in 1 (odvisna od vrste atomov 2 in 1 in njihovih bližjih sosedov), θ_{321} je kot med vezema (2,1) in (3,2) in lahko zavzema vrednosti od 0 do 180° , ϕ_{4321} je dihedralni kot, ki opisuje rotacijo okoli vezi (3,2); definiran je kot kot med ravnino, ki vsebuje atome 1,2 in 3, ter ravnino ki vsebuje atome 2,3 in 4 in lahko zavzema vrednosti med -180° in 180° . (desno): Delokalizacija elektronskega para v peptidni vezi.

Enojne vezi med atomi so načeloma (razen če so prisotne sterične ovire) prosto vrtljive. Peptidna vez pa ima zaradi resonance delni karakter dvojne vezi, saj je elektronski par delokaliziran med C'-O in C'-N vezema (glej sliko 3- desno ter sliko 4). Zaradi resonance ima karbonylni kisikov atom parcialno negativni naboj in amidni dušik parcialno pozitivni naboj, kar povzroči majhen električni dipol. Delni karakter dvojne vezi je vzrok za to, da atomi v peptidni vezi težijo h koplanarnosti (glej sliko 4) in tvorijo tako imenovano peptidno ravnino (zelena barva na sliki 6). Koplanarnost omogoča le dve konformaciji, cis in trans (slika 4). Konformacija cis ustreza dihedralnemu kotu $\omega \approx 0^\circ$; v tej konformaciji sta α -C atoma na enaki strani C-N vezi. Konformacija trans ustreza dihedralnemu kotu $\omega \approx \pm 180^\circ$; v tej konformaciji sta α -C atoma na različnih straneh C-N vezi. Večina peptidnih vezi je v trans konformaciji, z izjemo vezi s prolinom, ki zaradi večje togosti zelo poveča verjetnost cis vezi (razmerje verjetnosti trans/cis je pribl. 3000:1, če ne sodeluje prolin in 15:1 v primeru vezi s prolinom [2]). Definicije vseh dihedralnih kotov glavne polipeptidne verige so prikazane na sliki 4.



Slika 4: trans peptidna vez (levo), cis peptidna vez (sredina) in definicije vseh dihedralnih kotov, ki opisujejo rotacije okoli vezi med atomi v polipeptidni verigi (desno).

Zaradi delnega karakterja dvojne vezi peptidna vez torej ni prosto vrtljiva. Energetska bariera za rotacijo med konformacijama je približno 20 kcal/mol [3], kar je okoli 40 krat večje od termične energije v fizioloških pogojih, zaradi česar je spontana izomerizacija trans-cis izredno počasna (v realnem času se statistično skoraj ne zgodi). V ribosomih se sintetizirajo vse vezi v trans obliki [4]. V nekaterih primerih pa mora biti določena peptidna vez v cis konformaciji, da se lahko protein ustrezno zvije in opravlja svojo funkcijo. V teh primerih narava uporabi trik. Zaradi visoke rotacijske bariere je izomerizacija v cis encimsko katalizirana z encimi peptidil-prolin izomerazami (PPI). Slednje in ostale posttranslacijske modifikacije je potrebno upoštevati pri izdelavi algoritmov za napovedovanje tridimenzionalne strukture.

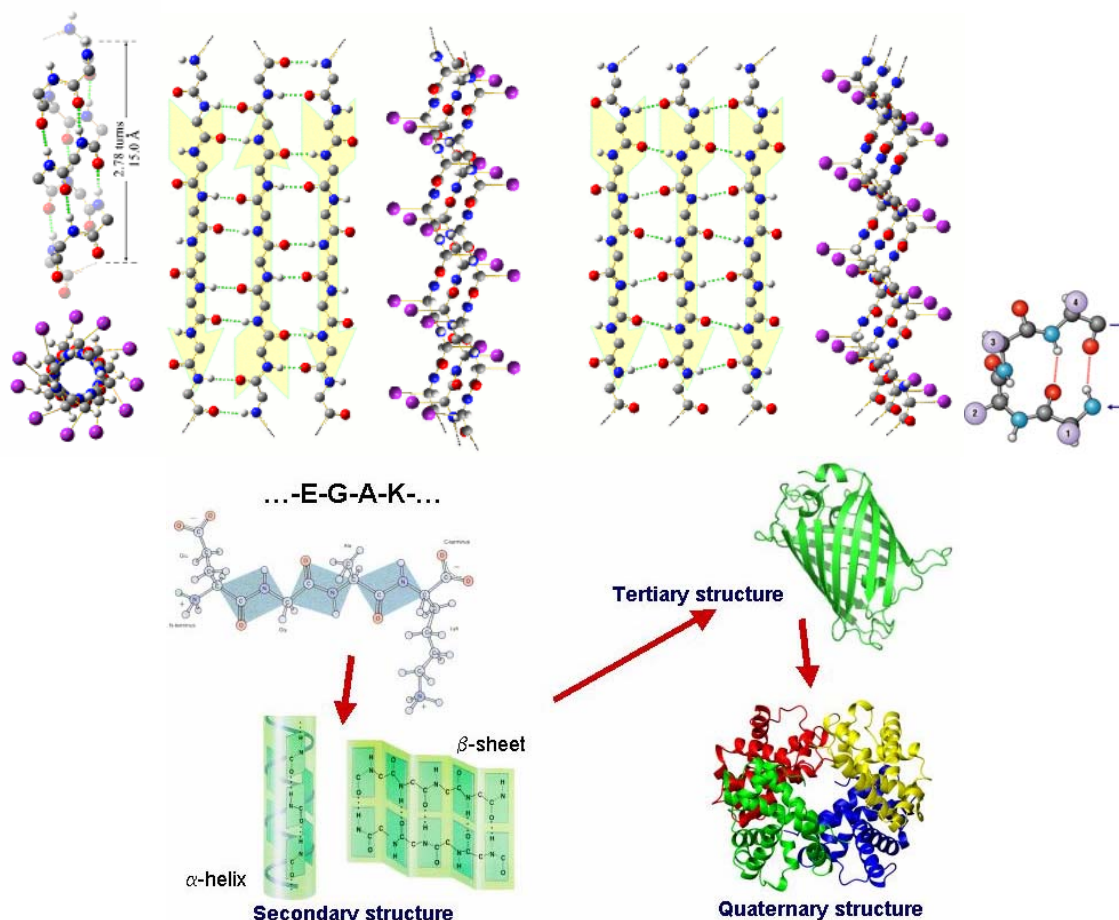
V nadaljnji obravnavi bomo predpostavili, da so vse AK (tudi tiste s prolinom) v trans konfiguraciji, tako bo konformacija proteina določena s dihedralnima kotoma ϕ in ψ , ki opisujeta rotacijo okoli obeh vezi z α -C atomom v polipeptidni verigi. Za elementi primarne strukture (AK zaporedje) so elementi sekundarne strukture naslednji nivo organizacije proteina. Elemente sekundarne strukture opredelimo kot določeno število zaporednih AK preostankov, ki imajo enake ϕ in ψ vrednosti. Elemente sekundarne strukture lahko okarakteriziramo tudi z vzorci vodikovih vezi, ki so v resnici vzrok njihove večje stabilnosti. Delimo jih na vijačnice (najbolj znana in pogosta je α -vijačnica), β -ploskve in β -zavoje (slika 5).

Naslednji nivo organizacije proteinov predstavlja tridimenzionalna ureditev elementov sekundarne strukture v prostoru skupaj z neregularnimi elementi strukture (deli med regularnimi elementi sekundarne strukture). Temu nivoju organizacije pravimo terciarna struktura (slika 5). Določitev

terciarne strukture proteina iz primarne (AK zaporedja) je predmet teorije zvijanja proteinov, ki jo bomo podrobneje obdelali v nadaljevanju. Če je protein sestavljen iz več podenot, potem strukturi asociata teh podenot pravimo kvarterna struktura proteina (slika 5), ki se ji v tem seminarju ne bomo posvetili. V nadaljevanju bomo na osnovi primarne strukture poskušali razložiti in napovedati elemente sekundarne in terciarne strukture.

Nativno stanje proteina je z biokemijskega stališča definirano kot stanje proteina v fizioloških pogojih (fiziološki pH, temperatura, tlak, ionska moč,...), v katerih je protein biološko aktiven. Nativna struktura fiziološko aktivnega proteina je zelo kompleksna. Ker se polipeptidne verige sintetizirajo linearno na ribosomu (ne sintetizira se že zvita veriga), bi lahko sklepali, da je za zvijanje v nativno stanje odgovoren kateri od celičnih "strojev", kar bi praktično onemogočilo napovedovanje konformacije iz AK zaporedja. V petdesetih letih prejšnjega stoletja pa je Christian B. Anfinsen eksperimentalno pokazal v seriji poskusov, da je vsa informacija o nativni strukturi zakodirana v AK zaporedju [5]. Za delo je prejel Nobelovo nagrado.

V seminarju se bomo posvetili statistični termodinamiki zvijanja in preko le-te poskušali fizikalno opredeliti nativno stanje in razložiti gonilno silo zvijanja.



Slika 5: α -vijačnica (zgoraj levo) antiparalelna β -ploskev, (zgoraj druga z leve) in paralelna β -ploskev (zgoraj tretja z leve) in β -zavoj (zgoraj desno; vodikove vezi v zavoju so označene z rdečimi črtami), povzetek celotne strukture (spodaj). V pogledu od zgoraj so zaradi večje preglednosti α -vodiki izpuščeni in smeri verig označene z rumenimi puščicami. Vodikove vezi so prikazane z zelenimi črtkanimi črtami. V navpični postavitvi α -vijačnice so stranske verige zaradi večje preglednosti izpuščene. Vodikove vezi so označene z zeleno.

2. Fizika proteinov

V tem odseku bomo predstavili atomistični molekularno mehanski model proteina, ki temelji na Born-Oppenheimerjevi aproksimaciji. Atomi predstavljajo klasične delce z določenim parcialnim nabojem, ki je posledica polarizacije vezi med atomi in disociacije titrabilnih funkcionalnih skupin.

2. 1 Molekulsko-mehanski model proteina

Zamislimo si protein, sestavljen iz M atomov z Evklidskimi koordinatami $(\mathbf{R}_i)=(X_i, Y_i, Z_i)$; $i=1,\dots,M$ in notranjimi koordinatami $\mathbf{q}=(q_i)$, $i=1,\dots,3M-6$ ($N-1$ vezi + $N-2$ kotov med vezmi + $N-3$ dihedralnih kotov). Molekula je potopljena v rezervoar z N togimi molekulami vode, katerega koordinate bomo označili z $\mathbf{r}=(x_i, y_i, z_i, \omega_i, \phi_i, \chi_i)$; $i=1,\dots,N$, kjer so x_i, y_i, z_i koordinate središča mase molekule vode in ω_i, ϕ_i, χ_i Eulerjevi koti, ki določajo orientacijo molekule topila. Niz vseh možnih mikroskopskih stanj sistema predstavlja fazni prostor, ki ga bomo označili z $\Gamma_M \times \Gamma_N$. Predpostavili bomo konstantno temperaturo in volumen (konst. št. molekul vode). Hamiltonian za tak sistem ima obliko

$$H(\mathbf{r}^N, \mathbf{R}^M, \boldsymbol{\Pi}_N, \boldsymbol{\pi}_M) = \sum_N \frac{\boldsymbol{\Pi}_N^2}{2m_N} + \sum_M \frac{\boldsymbol{\pi}_M^2}{2m_M} + U(\mathbf{r}^N, \mathbf{R}^M) \quad (1)$$

kjer sta $\boldsymbol{\Pi}_N$ in $\boldsymbol{\pi}_M$ gibalni količini molekul topila in atomov proteina, m_N in m_M sta masi molekule topila in atoma proteina in $U(\mathbf{r}^N, \mathbf{R}^M)$ potencialna energija. Če predpostavimo, da lahko sistem opišemo z kanoničnim ansamblom, je porazdelitvena funkcija dana z

$$Q = \frac{1}{h^{M+N} N!} \int_{\Gamma_M \times \Gamma_N} \exp[-\beta H(\mathbf{r}^N, \mathbf{R}^M, \boldsymbol{\Pi}_N, \boldsymbol{\pi}_M)] d\mathbf{r}^N d\mathbf{R}^M d\boldsymbol{\Pi}_N d\boldsymbol{\pi}_M \quad (2)$$

kjer je h Planckova konstanta, $\beta=1/(kT)$, kjer je k Boltzmannova konstanta in T temperatura, člen $N!$ pa nastopa zaradi kvantno-mehanske nerazločljivosti molekul vode. Ker nas prvenstveno zanimajo konformacijske lastnosti proteina, lahko gibalne količine vode in atomov proteina izpovprečimo [6]. Porazdelitvena funkcija tako dobi bolj preprosto obliko

$$Q = \frac{Z}{N! \Lambda^{3N} \prod_{i=1}^M \Lambda^3} \quad (3),$$

kjer je Z klasični konfiguracijski integral

$$Z = \int \exp[-\beta U(\mathbf{r}^N, \mathbf{R}^M)] d\mathbf{r}^N d\mathbf{R}^M \quad (4)$$

in Λ termična de Broglieva valovna dolžina. Integracija po momentih atomov proteina (izpovprečenje) je striktno upravičena le v primeru konformacije, ki je blizu native konformacije, saj sicer transformacija spremenljivk ni invariantna [7]. Težavi se lahko izognemo, tako da razdelimo konformacijski fazni prostor na končne podprostore [8]. Z upoštevanjem enačbe 3 je Helmholtzova prosta energija dana z

$$F = -kT \ln Q = -kT \ln Z + kT \ln \left(N! \Lambda^{3N} \prod_{i=1}^M \Lambda^3 \right) \quad (5).$$

Formalno lahko izvedemo tudi integracijo po koordinatah topila, če vpeljemo potencial povprečne sile (potential of mean force), W ,

$$\exp(-\beta W) = Z_{ww}^{-1} \int \exp[-\beta U(\mathbf{r}^N, \mathbf{R}^M)] d\mathbf{r}^N \quad (5),$$

kjer je Z_{ww} konfiguracijski integral čistega topila,

$$Z_{ww} = \int \exp(-\beta U_{ww}) d\mathbf{r}^N \quad (6).$$

Tako lahko zapišemo celotni konfiguracijski integral v obliki

$$Z = Z_{ww} \int \exp(-\beta W) d\mathbf{R}^M \quad (7)$$

kjer je odvisen le od koordinat atomov proteina. Če je Hamiltonian aditiven, kar velja za večino molekulske-mehanskih polj sil [9], lahko naredimo še naslednje poenostavitve. Hamiltonian lahko razčlenimo na posamezne interakcijske prispevke: $U=U_{mm}+U_{mw}+U_{ww}$, kjer prvi člen predstavlja intramolekularne interakcije, drugi interakcije proteina z vodo in tretji interakcije med molekulami topila. Tako dobimo

$$\exp(-\beta W) = \exp(-\beta U_{mm}) Z_{ww}^{-1} \int \exp(-\beta U_{mw} - \beta U_{ww}) d\mathbf{r}^N \quad (8).$$

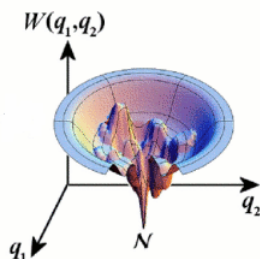
Potencial povprečne sile lahko zapišemo kot $W=H_{mm}+X$, kjer je

$$\exp(-\beta X) = Z_{ww}^{-1} \int \exp(-\beta U_{mw} - \beta U_{ww}) d\mathbf{r}^N \quad (9),$$

kar lahko preoblikujemo v

$$X = -kT \ln \langle \exp(-\beta U_{mw}) \rangle_0 \equiv \Delta G^{solv} \quad (10),$$

kjer povprečimo po ansamblu čistega topila (o). Enačba (10) je znana kot presežni kemijski potencial ali prosta energija solvatacije [10]. Potencial povprečne sile tako sestavljata dva prispevka: intramolekularna energija in solvacijska prosta energija. Namesto izraza potencial povprečne sile lahko uporabimo tudi efektivna energija. W definira hiperploskev v konformacijskem faznem prostoru proteina v prisotnosti ekvilibriranega topila (po spremembi konformacije proteina je potreben končni čas za vzpostavitev novega ravnotežja molekul topila, torej predpostavimo, da se to zgodi hipoma oz. prej kot se ponovno spremeni konformacija, zato ekvilibrirano topilo). To hiperploskev imenujemo tudi energetski relief (slika 6) in določa termodinamiko in kinetiko makromolekulskih konformacijskih prehodov.



Slika 6: Hipotetični energetski relief. Zaradi preglednosti sta ohranjeni le dve prostostni stopnji q_1 in q_2 . N predstavlja nativno stanje in predpostavljamo, da gre za globalni minimum.

Razčlenitev efektivne energije lahko formalno izvedemo tudi v primeru, ko potencial ni-parsko aditiven. Če gre za za interakcije med tremi telesi, lahko zapišemo

$$U = (U_{mm} + U_{mmm}) + (U_{mw} + U_{ww} + U_{www} + U_{wmm} + U_{mmw}) \quad (12)$$

efektivna energija pa je potem dana z

$$W = (U_{mm} + U_{mmm}) - kT \ln \langle \exp(-\beta(U_{mw} + U_{mmw} + U_{mw})) \rangle_0 \quad (13),$$

vendar je obdelava W v tem primeru bistveno bolj zahtevna.

Za formalni opis makromolekule je pogosto ugodnejša uporaba notranjih koordinat, \mathbf{q} (slika 3). Transformacijska matrika je odvisna le od dolžin vezi in kotov med vezmi, zato je približno konstantna za vse konformacije makromolekule in jo lahko vzamemo izpod integrala v enačbi (7). Mi jo bomo vključili v oznako $d\mathbf{q}$. Z integracijo po šestih zunanjih koordinatah (opisujejo postavitev celotne molekule v prostoru: x, y, z in trije Eulerjevi koti), ki jo lahko opravimo, saj je sistem homogen, dobimo $V8\pi^2$, tako da enačba (7) dobi obliko

$$Z = Z_{ww} V 8\pi^2 \int \exp(-\beta W) d\mathbf{q} \quad (14).$$

Verjetnost, da makromolekulo v stanju \mathbf{q} , $p(\mathbf{q})$, je dana z

$$p(\mathbf{q}) = \frac{\exp(-\beta W(\mathbf{q}))}{\int \exp(-\beta W(\mathbf{q})) d\mathbf{q}} \quad (15).$$

Iz enačbe (5) je razvidno, da konformacijsko distribucijo proteina v raztopini določa efektivna energija. Če upoštevamo enačbi (14) in (15) lahko zapišemo

$$\begin{aligned} \int p(\mathbf{q}) \ln p(\mathbf{q}) d\mathbf{q} &= \int p(\mathbf{q}) \{-\beta W(\mathbf{q}) - \ln Z + \ln Z_{ww} + \ln 8\pi^2 V\} d\mathbf{q} \\ &= -\ln Z + \ln Z_{ww} + \ln 8\pi^2 V - \beta \int p(\mathbf{q}) W(\mathbf{q}) d\mathbf{q} \end{aligned} \quad (16)$$

in enačba (5) dobi obliko

$$\begin{aligned} F &= F^0 + kT \ln \left(\frac{\prod_{i=1}^M \Lambda^3}{8\pi^2 V} \right) + \int p(\mathbf{q}) \{H_{mm}(\mathbf{q}) + \Delta G^{solv}(\mathbf{q})\} d\mathbf{q} + kT \int p(\mathbf{q}) \ln p(\mathbf{q}) d\mathbf{q} \\ &= F^0 + kT \ln \left(\frac{\prod_{i=1}^M \Lambda^3}{8\pi^2 V} \right) + \langle W \rangle - TS^{konfig} \end{aligned} \quad (17),$$

kjer je F^0 prosta energija čistega topila in drugi člen idealni prispevek translacije in rotacije proteina. Tretji člen predstavlja povprečno efektivno energijo (povprečna intramolekularna energija + povprečna prosta energija solvatacije) in zadnji člen prispevek konfiguracyjske entropije proteina. Entropija topila je vključena v ΔG^{solv} in $p(\mathbf{q})$. Gibbsovo prosto energijo dobimo tako, da enačbi (17) prištejemo člen pV , ki je pri fizioloških pogojih zanemarljiv in tako lahko Helmholtzovo in Gibbsovo prosto energijo uporabljamo izmenljivo.

Za energijo dobimo ob opoštevanju ekviparticijskega izreka za kinetične energije izraz

$$E = kT^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V} = \frac{3}{2} kT(N+M) + \int p(\mathbf{q}, \mathbf{r}^N) U(\mathbf{q}, \mathbf{r}^N) d\mathbf{q} d\mathbf{r}^N \quad (18)$$

kjer prvi člen predstavlja kinetično energijo, drugi pa potencialno energijo. Verjetnost, da najdemo sistem v stanju $(\mathbf{q}, \mathbf{r}^N)$ je dana z izrazom

$$p(\mathbf{q}, \mathbf{r}^N) = p(\mathbf{q}) p(\mathbf{r}^N | \mathbf{q}) \quad (19)$$

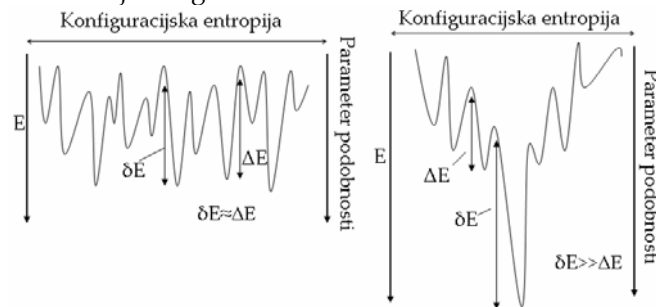
kjer je $p(\mathbf{r}^N | \mathbf{q})$ pogojna verjetnostna distribucija, da najdemo topilo v stanju \mathbf{r}^N , pri pogoju da je protein v konformaciji \mathbf{q} . Ob upoštevanju enačb (7) in (19) lahko zapišemo

$$E = E_{kin} + \int p(\mathbf{q}) U_{mm}(\mathbf{q}) d\mathbf{q} + \int p(\mathbf{q}) \int p(\mathbf{r}^N | \mathbf{q}) \{U_{mw}(\mathbf{q}, \mathbf{r}^N) + U_{ww}(\mathbf{r}^N)\} d\mathbf{q} d\mathbf{r}^N \quad (20)$$

kjer drugi člen predstavlja povprečno intramolekularno potencialno energijo in tretji povprečno energijo topljenec-topilo in topilo-topilo. Entropijo dobimo z relacijo $(E-F)/T$.

Izpeljane enačbe nam omogočajo, da formalno opredelimo nativno stanje proteina, in sicer je to distribucija konfiguracij proteina, $p(\mathbf{q})$, ki minimizira funkcional proste energije v enačbi (17) pri fizioloških pogojih (variacijska minimizacija enačbe (17) da enačbo (15)). Definicija vključuje fleksibilnost in nered v nativnem stanju. Povprečna efektivna energija teži k temu, da je protein lokaliziran v najglobljih dolinah hiperploskve efektivne energije $[y]$, konfiguracijska entropija pa poskuša narediti $p(\mathbf{q})$ čim bolj enakomerno. Posledično nativno stanje, ki vsebuje konformacije z najnižjo prosto energijo, ni nujno konformacija z najnižjo efektivno energijo, saj so lahko določene najgloblje doline W tako 'ozke', da je vibracijska entropija proteina v teh dolinah zelo majhna.

Proteini pa niso naključni heteropolimeri 20-ih aminokislin (naključno zaporedje AK), temveč evolucijsko izbrane sekvence z edinstvenim zvitjem [11]. Naključni heteropolimeri nimajo dobro definirane energetskega minimuma- obstaja veliko število konformacij z nizko energijo. Pri teh pride do hidrofobnega kolapsa (tako da so hidrofobne AK senčene pred vodo s hidrofilnimi AK), pri katerem protein preide v energetsko bolj ugodno stanje, vendar je razlika energije primerljiva z velikostnim redom tipičnih fluktuacij energije. Kolabiranih stanj s primerljivo energijo je veliko (slika 7- levo). Pri proteinskih sekvencah (slika 7- desno) je eno oz. zelo ozka distribucija stanj izrazito bolj stabilna in razlika energije bistveno večja od tipičnih fluktuacij energije. Velika energetska razlika med kolabiranimi in nativnim stanjem pri proteinski sekvencah predstavlja gonilno silo, ki pelje sistem do nativnega stanja. Energetske bariere so reda velikosti fluktuacij energije, a ko se protein enkrat znajde v nativnem stanju, pri danih pogojih ne more več pobegniti iz globalnega minimuma (fluktuacije energije niso dovolj). Izrazit minimum je posledica evolucijsko izbranega AK zaporedja, ki omogoča maksimalno število energetske ugodnih topoloških kontaktov. Pri naključnih heteropolimerih po hidrofobnem kolapsu ne nastanejo le ugodni kontakti.



Slika 7: Shematski prikaz energetskega reliefa naključnega heteropolimera (levo) in evolucijsko izbrane proteinske sekvence (desno). δE predstavlja stabilnost konformacije v globalnem minimumu in ΔE tipične fluktuacije energije. Širina minimuma (lokalnih in globalnega) predstavlja konfiguracijsko entropijo.

2.2 Prosta energija zvijanja in termodinamska stabilnost nativnega stanja

S statistično mehanskim pristopom lahko izpeljemo prosto energijo zvijanja proteina s stališča interakcij in distribucije mikroskopskih konformacijskih stanj. Najprej razdelimo konfiguracijski prostor v podprostore A , ki vsebujejo različne konfiguracije proteina, kot to analogno storimo pri preučevanju ravnotežja med različnimi konformacijskimi izomeri manjših organskih molekul. Prosta energija konformacijskega niza A je dana z

$$F_A = -kT \ln Z_A + kT \ln \left(N! \Lambda^{3N} \prod_{i=1}^M \Lambda^3 \right) \quad (21)$$

kjer je

$$Z_A = Z_{ww} 8\pi^2 V \int_A \exp(-\beta W) d\mathbf{q} \quad (22)$$

Integral v enačbi (22) je po vseh konformacijah, ki jih vsebuje niz A . Po definiciji je $\sum_A Z_A = Z$. Prosto energijo konformacijskega niza A lahko zapišemo tudi analogno kot v enačbi (17)

$$\begin{aligned}
F_A &= F^0 + kT \ln \left(\frac{\prod_{i=1}^M \Lambda^3}{8\pi^2 V} \right) + \int_A p_A(\mathbf{q}) \{U_{mm}(\mathbf{q}) + \Delta G^{solv}(\mathbf{q})\} d\mathbf{q} + kT \int_A p_A(\mathbf{q}) \ln p_A(\mathbf{q}) d\mathbf{q} \\
&= F^0 + kT \ln \left(\frac{\prod_{i=1}^M \Lambda^3}{8\pi^2 V} \right) + \langle W \rangle_A - TS_A^{konfig}
\end{aligned} \tag{23}$$

kjer je $p_A(\mathbf{q})$ verjetnostna distribucija normalizirana znotraj niza A

$$p_A(\mathbf{q}) = \frac{\exp(-\beta W(\mathbf{q}))}{\int_A \exp(-\beta W(\mathbf{q})) d\mathbf{q}} \tag{24}$$

Prva člena v enačbi (23) sta enaka za vse konformacije, tretji člen predstavlja povprečno efektivno energijo stanja A in zadnji člen konfiguraciono entropijo stanja A. Razlika proste energije med dvema stanjema A in B je potem dana z

$$\begin{aligned}
\Delta F &= F_B - F_A = F[p_B(\mathbf{q})] - F[p_A(\mathbf{q})] = \langle W \rangle_B - \langle W \rangle_A - T[S_B^{konfig} - S_A^{konfig}] \\
&= \Delta \langle W \rangle - T\Delta S^{konfig} = \Delta \langle H_{mm} \rangle + \Delta \langle \Delta G^{solv} \rangle - T\Delta S^{konfig}
\end{aligned} \tag{25}$$

kjer oznaka $F[p(\mathbf{q})]$ pomeni, da je prosta energija funkcional verjetnostne distribucijske funkcije. Če A predstavlja denaturirano stanje (konformacije znotraj niza A) in B nativno stanje (konformacije znotraj niza B) predstavlja enačba (25) prosto energijo zvijanja. Termodinamska stabilnost nativnega stanja je rezultat ravnovesja med efektivno energijo, ki favorizira nativno stanje, in konfiguracijo ki favorizira denaturirano stanje. Sprememba efektivne energije je povezana z globino doline energetske hiperploskve, v kateri je locirano nativno stanje. Med stanjema je lahko seveda prisotna energetska bariera, katere višina določa kinetično stabilnost nativnega stanja. Entropijska izguba pri lokalizaciji proteina v najgloblji dolini hiperploskve je ocenjena na nekaj sto kcal/mol [13]. Celotna sprememba proste energije pri zvijanju je ponavadi med 5 in 15 kcal/mol [14], globina doline nativnega stanja na hiperploskvi efektivne energije je prav tako reda velikosti nekaj sto kcal/mol [15].

2. 3 Intramolekularne interakcije

Kot smo pokazali v prejšnjem odseku, je potencial povprečne sile sestavljen iz prispevka intramolekularnih interakcij in solvatacijske proste energije. Za to potrebujemo potencialno energijo kot funkcijo koordinat atomov in model solvatacije. S potencialno energijo lahko direktno določimo relativne stabilnosti posameznih konformacij. Sile, ki delujejo na atome, dobimo z odvodom potenciala po atomskih koordinatah. S temi silami lahko določimo dinamične in termodinamske lastnosti sistema, npr. z reševanjem Newtonovih enačb gibanja lahko preučujemo spreminjanje atomskih pozicij s časom ali pa iz pozicij izračunamo povprečne lastnosti, npr. entalpijo [12,13]. Z drugimi odvodi potencialne energije po odmikih iz ravnovesnih leg lahko dobimo konstante sil, s pomočjo katerih lahko določimo normalne načine nihanja. Le ti predstavljajo alternativni pristop k dinamiki v harmonski limiti in so uporabni za vključitev kvantno-mehanskih popravkov vibracijskih prispevkov k termodinamskim količinam. Da lahko izračunamo hiperploskve potencialne energije za proteine z ustrežno točnostjo in hitrostjo, moramo uporabiti enostavne modele, ki so kalibrirani s fitanjem eksperimentalnih podatkov ali kvantno-mehanskih izračunov. Ko imamo opravka s proteini z nekaj sto do nekaj tisoč atomi, moramo za zanesljiv izračun interakcijskih energij uporabiti empirične potencialne funkcije. Le te niso splošno uporabne, kot npr. kvantno-mehanski izračuni ampak le za sisteme, na katerih so kalibrirane. Potentialno energijo razdelimo na vezne prispevke

(dolžine vezi, koti med vezmi, dihedralni koti) in nevezne interakcije (van der Waalsove in elektrostatske interakcije): $U_{mm} = U_{vezne} + U_{nevezne}$. Najprej si pogledjmo vezne prispevke. Kovalentno vez ponavadi modeliramo s harmonskim potencialom. Enačba, ki opisuje potencialno energijo harmonskega modela (slika 8- skrajno levo), ima obliko

$$U_B = k_{ij}^B (r_{ij} - r_0)^2 \quad (26),$$

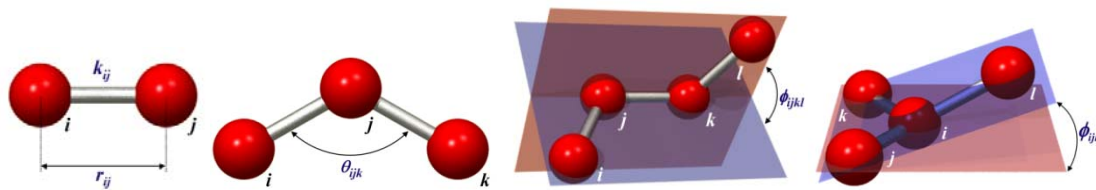
kjer je k_{ij}^B konstanta sile vezi, r_{ij} je razdalja med atomoma i in j (dolžina vezi) in r_0 je razdalja med atomoma, kjer ima U_B minimum. Harmonski potencial je Taylorjeva aproksimacija (najnižjega reda) bolj zapletenih in bolj realističnih modelov v bližini referenčne razdalje (energetskega minimuma). Harmonski kotni potencial (slika 8- druga z leve) s konstanto sile k_θ je podan z enačbo

$$U_\theta = k_\theta (\theta_{ijk} - \theta_0)^2 \quad (27)$$

H kotnemu potencialu ponavadi dodamo še Urey-Bradley-ev potencial z namenom omejitve gibanja vezi, ki opisujeta kot. Ta potencial v bistvu predstavlja fiktivno vez med atomoma i in k in ga opišemo z enačbo

$$U_{UB} = k_{ik}^{UB} (r_{ik} - r_0^{UB})^2 \quad (28),$$

kjer je k_{ik}^{UB} konstanta sile UB potenciala, r_{ik} razdalja med atomoma i in k in r_0^{UB} razdalja med i in k , kjer ima U_{UB} minimum.



Slika 8: Od leve proti desni si sledijo: Potencial vezi, kotni potencial, potencial dihedralnega kota in potencial neustrezne torzije.

Poznamo dva tipa torzijskih potencialov, dihedralni potencial (oz. standardni torzijski) in potencial nestandardne torzije (improper torsion potential). Oba se nanašata na štiri atome, ki so tako ali drugače povezani (slika 8- druga z desne in skrajno desna). Potencial dihedralnega kota je odvisen od štirih zaporednih atomov (slika 8- druga z desne), potencial nestandardne torzije pa je odvisen od treh atomov, ki so postavljeni okoli četrtega (slika 8- skrajno desno).

Dihedralni potencial upoštevamo z namenom omejitve rotacije okoli vezi, nestandardno torzijo pa z namenom ohranjanja kiralnosti tetraedričnega atoma ali z namenom ohranjanja koplanarnosti atomov. Glavna razlika med omenjenima potencialoma je definicija torzijskega kota in oblika potencialne funkcije. Razliko v definiciji torzijskega kota lahko odpravimo, če atome pametno označimo (slika 8- druga z desne in skrajno desna). Standardni torzijski kot je kot med ravnino, ki vsebuje atome i , j in k , ter ravnino, ki vsebuje atome j , k in l .

Bariero za rotacijo okoli vezi najbolje opiše kosinusna oblika torzijskega potenciala

$$U_D = k_{ijkl}^C (1 + \cos(n_{ijkl} \phi_{ijkl} - \phi_0)) \quad (29)$$

kjer je k_{ijkl}^C konstanta sile kosinusnega torzijskega potenciala, ϕ_0 je kot, pri katerem ima U_D minimum in n_{ijkl} multipliciteta, ki pove število minimumov, če vez zavrtimo za polni krog (360°). Multipliciteta je neničelno pozitivno celo število.

Nestandardni torzijski potencial je prav tako odvisen od kota med ravninama, ki vsebujeta atome i , j in k in atome j , k in l (saj smo 'iznajdljivo' poimenovali atome-glej sliko 8- skrajno desno). Zato lahko potencial opišemo z enakimi enačbami kot standardni torzijski potencial. Ker pa je nestandardni

torzijski potencial v glavnem vpeljan za ohranitev planarnosti v strukturi in ima zato le en minimum, ga lahko opišemo z enostavnim harmonskim potencialom

$$U_{imp} = k_{imp} (\phi_{ijkl} - \phi_0)^2 \quad (30).$$

Med nevezne potenciale spadata van der Waalsov in elektrostatski potencial (slika 9). Van der Waalsove interakcije najpogosteje opisujemo z Lennard-Jones-ovim potencialom

$$U_{vdW} = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (31)$$

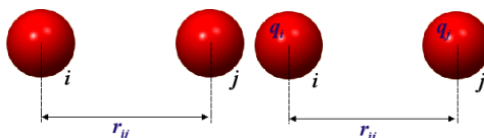
kjer je ϵ_{ij} vrednost potenciala v minimumu, σ_{ij} je razdalja med atomoma, kjer je $U_{vdW}=0$. Ponavadi so parametri ϵ in σ dani za posamezne atome, zato moramo za izračun parskih parametrov uporabiti Lorentz-Berthelot-ova pravila mešanja

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad \epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \quad (32)$$

Ker bolj elektronegativni atomi močnejše privlačijo elektrone, je porazdelitev naboja v molekuli neenakomerna in jo najpogosteje predstavljamo s točkastimi naboji, lokaliziranimi v molekuli. Tudi ione opišemo kot točkaste naboje. Energija elektrostatske interakcije med dvema točkastima nabojema je dana s Coulomb-ovim zakonom

$$U_E = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}} \quad (33),$$

kjer sta q_i in q_j naboja atomov i in j , ϵ_0 je permisivnost vakuumu in ϵ_r dielektrična konstanta medija.



Slika 9: VdW (levo) in elektrostatski (desno) potencial.

2. 4 Eksplicitno toplilo

Za popolno solvatacijo proteina je potrebnih na tisoče molekul vode, ki jih lahko v modelu predstavimo eksplicitno. Eksplicitno toplilo predstavlja veliko oviro pri računalniških simulacijah, saj se večino časa porabi za računanje interakcij med molekulami vode, nas pa zanima protein. Da bi se izognili temu problemu, so bili vpeljeni različni implicitni modeli topila (npr. posplošen Bornov model, Poisson-Boltzmann, model nehomogene tekočine)

2. 5 Implicitno toplilo

Za opis solvatacije proteinov in makromolekul na splošno obstaja več različnih modelov, kot so npr. kontinuumski elektrostatski, model nehomogene tekočine, različni empirični solvatacijski modeli ipd. Za solvatacijo sta najpomembnejši elektrostatika in hidrofobni efekt.

2. 5. 1 Posplošeni Bornov model solvatacije

Za opis elektrostatske komponente proste energije solvatacije se najpogosteje uporabljajo modeli, ki temeljijo na elektrostatici na nivoju kontinuuma in temeljijo na razdelitvi celotnega prostora na dve

regiji, med katerima obstaja dielektrična meja. Topilo v tem primeru predstavlja visoko dielektrični kontinuum, ki interagira z atomi proteina, ki imajo manjšo dielektrično konstanto. Elektrostatski potencial lahko dobimo neposredno, tako da numerično rešimo Poissonovo enačbo

$$\nabla \cdot \varepsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \quad (34)$$

kjer je $\varepsilon(\mathbf{r})$ prostorsko odvisna dielektrična konstanta, $\phi(\mathbf{r})$ predstavlja elektrostatski potencial in $\rho(\mathbf{r})$ gostoto naboja. Z nekaj dodatnimi aproksimacijami pa lahko izpeljemo tudi analitični izraz za opis proste energije solvatacije, kot npr. v primeru posplošenih Bornovih modelov (Generalized Born - GB). Mi bomo predstavili različico GB, ki je najprimernejša za opis solvatacije makromolekul, kot so proteini [18]. V modelu vsak atom makromolekule opišemo kot sfero z van der Waalsovimi radijem σ_i s parcialnim nabojem q_i , ki se nahaja v središču sfere. Parcialne naboje atomov dobimo z rentgensko kristalografijo, NMR v trdnem ali kvantno-mehanskimi izračuni in so del empiričnih potencialnih funkcij (CHARMM, AMBER, GROMOS). Notranjost atoma je enakomerno zapolnjena z dielektrikom z nizko dielektrično konstanto ε_i (Ponavadi upoštevamo, da je dielektričnost znotraj atoma enaka dielektričnosti vakuumu, torej $\varepsilon_i=1$). Molekulo obdaja topilo z visoko dielektrično konstanto ε_s .

Bornovo solvatacijsko energijo za topljenec z N naboji q_1, \dots, q_N v kontinuumu topila, izpeljemo tako, da najprej zapišemo elektrostatsko energijo kot integral gostote energije električnega polja $u(\mathbf{x}) = \mathbf{D}^2(\mathbf{x}) / (8\pi\varepsilon(\mathbf{x}))$ po celotnem prostoru

$$A^{el} = \int_{R^3} u(\mathbf{x}) d^3x = \frac{1}{8\pi\varepsilon_s} \int_{topilo} \mathbf{D}^2(\mathbf{x}) d^3x + \frac{1}{8\pi\varepsilon_i} \int_{topljenec} \mathbf{D}^2(\mathbf{x}) d^3x \quad (35)$$

kjer je $\mathbf{D}(\mathbf{r})$ gostota električnega polja in $\varepsilon(\mathbf{r})$ prostorsko odvisna dielektrična konstanta. $\mathbf{D}(\mathbf{r})$ je dan s superpozicijo posameznih nabojev, $\mathbf{D}(\mathbf{r}) = \sum_i \mathbf{D}_i(\mathbf{r})$ kjer $i=1, \dots, N$. Simbol R^3 označuje, da integriramo po neskončnem 3 dimenzionalnem kartezičnem prostoru. Enačbo (35) lahko zapišemo drugače, če ji prištejemo in odštejemo integral $\mathbf{D}^2(\mathbf{x}) / (8\pi\varepsilon_s)$ po volumnu topljenca

$$A^{el} = \frac{1}{8\pi\varepsilon_s} \int_{R^3} \mathbf{D}^2(\mathbf{x}) d^3x + \frac{1}{8\pi} \left(\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon_s} \right) \int_{topljenec} \mathbf{D}^2(\mathbf{x}) d^3x \quad (36)$$

$\mathbf{D}(\mathbf{r})$ v obeh integralih enačbe (36) mora izpolnjevati robne pogoje električnega polja na meji med topljencem in topilom, in sicer se tangencialni komponenti električnega polja $\mathbf{E} = \mathbf{D}/\varepsilon$ in normalna komponenta gostote električnega polja pri prehodu meje med topilom in topljencem ne spremenita. Gostoto električnega polja v prvem integralu lahko aproksimiramo s Coulombskim poljem. Slednje povzroči relativno napako v elektrostatski energiji le za nekaj % [19]. Aproksimacijo lahko upravičimo tudi s kvalitativnim argumentom: Če je topljenec majhen, je večina nabojev močno izpostavljena topilu, tako da je odstopanje polja od Coulombskega majhno. Če pa je topljenec velik, prevladujoči del elektrostatske energije predstavlja prispevek drugega integrala v enačbi (36), tako da z aproksimacijo ne naredimo velike napake.

Za točkaste naboje je gostota električnega polja dana z $\mathbf{D} = \sum_i q_i \varepsilon(\mathbf{x}) (\mathbf{x} - \mathbf{x}_i) / |\mathbf{x} - \mathbf{x}_i|^3$. Po integraciji \mathbf{D}^2 po celotnem prostoru dajo izvendiagonalni členi $2\mathbf{D}_i \cdot \mathbf{D}_j$ Coulombsko interakcijsko energijo, $(1/4\pi\varepsilon_s) \int_{R^3} \mathbf{D}_i \cdot \mathbf{D}_j d^3x = q_i q_j / (\varepsilon_s r_{ij})$, medtem ko diagonalni elementi, ki predstavljajo prispevke lastne energije \mathbf{D}_i^2 , divergirajo. Singularnost lahko obidemo, če predpostavimo končno velikost nabojev, tako da si zamislimo, da je naboj q_i enakomerno porazdeljen po površini sfere z radijem R_i (Bornov pristop [20]). Potem je lastna energija naboja i v homogenem dielektriku dana z $(1/8\pi\varepsilon_s) \int_{R^3} \mathbf{D}_i^2 d^3x = q_i^2 / (2\varepsilon_s R_i)$. Prvi integral v enačbi (36) je torej dan s Coulombsko in Bornovo energijo v topilu z dielektrično konstanto ε_s . Potrebno je še izvrednotiti drugi integral v enačbi (36) po volumnu topljenca. Celotno elektrostatsko energijo topljenca v dielektričnem mediju lahko razdelimo na prispevke lastnih energij in interakcijske prispevke:

$$A^{el} = \sum_i A_i^{lastna} + \sum_{i<j} A_{ij}^{int} \quad (37)$$

$$A_i^{lastna} = \frac{q_i^2}{2\varepsilon_s R_i} + \frac{1}{8\pi} \left(\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon_s} \right) \int_{topljenec} \mathbf{D}_i^2(\mathbf{x}) d^3x \quad (38)$$

$$A_{ij}^{int} = \frac{q_i q_j}{\varepsilon_s r_{ij}} + \frac{1}{8\pi} \left(\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon_s} \right) \int_{topljenec} \mathbf{D}_i(\mathbf{x}) \cdot \mathbf{D}_j(\mathbf{x}) d^3x \quad (39)$$

Celotno elektrostatsko energijo topljenca v topilu pa lahko zapišemo še drugače, in sicer kot vsoto celotne elektrostatske energije v vakuumu in proste energije solvatacije, ΔA , ali drugače

$$\Delta A = A_{razl}^{el} - A_{vak}^{el} \quad (40)$$

Za takšen opis moramo upoštevati tri dielektrične konstante, dielektrično konstanto znotraj atoma topljenca, ε_i (ponavadi $\varepsilon_i=1$), dielektrično konstanto topila, ε_s in vakuuma ($\varepsilon=1$). Celotno energijo topljenca v vakuumu zapišemo preprosto kot vsoto Bornove energije in Coulombske interakcije, saj ima sistem homogen dielektrični odziv

$$A_{vak}^{el} = \sum_i \frac{q_i^2}{2\varepsilon_i R_i} + \sum_{i<j} \frac{q_i q_j}{\varepsilon_i r_{ij}} \quad (41)$$

ΔA predstavlja v tem pogledu delo, ki je potrebno za prenos topljenca iz vakuuma v dielektrik in je enako ΔG^{solv} . S kombinacijo enačb (37-41) lahko zapišemo

$$\Delta A = \sum_i \Delta A_i^{lastna} + \sum_{i<j} \Delta A_{ij}^{int} \quad (42)$$

$$\Delta A_i^{lastna} = A_i^{lastna} - \frac{q_i^2}{2\varepsilon_i R_i} \quad (43)$$

$$\Delta A_{ij}^{int} = A_{ij}^{int} - \frac{q_i q_j}{\varepsilon_i r_{ij}} \quad (44)$$

Pri izpeljavi GB enačbe za prosto energijo solvatacije najprej celotno elektrostatsko energijo proteina v topljencu zapišemo kot vsoto energije proteina v vakuumu in solvatacijske proste energije, ki jo v literaturi imenujejo tudi polarizacijska energija. Če upoštevamo enačbo (41), lahko zapišemo GB izraz za celotno elektrostatsko energijo makromolekule v topilu v obliki

$$A_{razl}^{el} = \sum_i \frac{q_i^2}{2\varepsilon_i R_i} + \sum_{i<j} \frac{q_i q_j}{\varepsilon_i r_{ij}} + \sum_{i,j} \Delta G_{ij}^{solv} \quad (45)$$

kjer smo uporabili $\Delta A = \Delta G^{solv}$. Diagonalni elementi 'matrike solvatacijske proste energije', ΔG_{ii}^{solv} , predstavljajo atomske solvatacijske proste energije, izvendagonalni elementi ΔG_{ij}^{solv} pa vključujejo senčene Coulombske intrakcije med nabojema i in j . Oglejmo si analitična izraza parske solvatacijske energije, $\Delta G_2^{solv} = \Delta G_{ii}^{solv} + \Delta G_{jj}^{solv} + \Delta G_{ij}^{solv} + \Delta G_{ji}^{solv}$ za delca i in j v limitah $r_{ij} \rightarrow \infty$ in $r_{ij} \rightarrow 0$. V primeru ko sta delca daleč narazen, se interakcijska energija med nabojema i in j asimptotično približuje Coulombski interakciji in je prosta solvatacijska energija dana z

$$\lim_{r_{ij} \rightarrow \infty} \Delta G_2^{solv} = -\frac{1}{2} \left(\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon_s} \right) \left(\frac{q_i^2}{\sigma_i} + \frac{q_j^2}{\sigma_j} \right) - \left(\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon_s} \right) \frac{q_i q_j}{r_{ij}} \quad (46)$$

V drugem primeru ko $r_{ij} \rightarrow 0$, se sferi zlijeta v eno in če predpostavimo še $\sigma_i \approx \sigma_j$, lahko tak sistem aproksimiramo s sfero z radijem $(\sigma_i + \sigma_j)^{1/2}$ in celokupnim nabojem $(q_i + q_j)$. Prosta energija solvatacije je v tem primeru enostavno razlika Bornovih energij take sfere v vakuumu in topilu

$$\lim_{r_{ij} \rightarrow 0} \Delta G_2^{solv} = -\frac{1}{2} \left(\frac{1}{\epsilon_i} - \frac{1}{\epsilon_s} \right) \frac{(q_i + q_j)^2}{\sqrt{\sigma_i \sigma_j}} \quad (47).$$

Podobnost enačb (46) in (47) namiguje *Ansatz* za posplošeno Bornovo solvatacijsko prosto energijo

$$\Delta G^{solv} = -\frac{1}{2} \sum_{i,j} \left(\frac{1}{\epsilon_i} - \frac{1}{\epsilon_s} \right) \frac{q_i q_j}{f_{ij}^{GB}(r_{ij})} \quad (48)$$

kjer je $f^{GB_{ij}}(r_{ij})$ določena gladka funkcija, ki je odvisna le od atomskih radijev σ_i in medatomskih razdalj r_{ij} . Zahtevamo, da $f^{GB_{ij}}$ interpolira med limitnima primeroma in da mora biti takšna, da izpolnjuje $f^{GB_{ii}} = \sigma_i$, in da gre $f^{GB_{ij}} \rightarrow r_{ij}$, ko gre $r_{ij} \rightarrow \infty$ ter $f^{GB_{ij}} \rightarrow (\sigma_i \sigma_j)^{1/2}$, ko gre $r_{ij} \rightarrow 0$. Najpogosteje uporabljena oblika $f^{GB_{ij}}$, ki ustreza tem zahtevam, je dana z enačbo

$$f_{ij}^{GB}(r_{ij}) = \left[r_{ij}^2 + R_i R_j \exp(-r_{ij}^2 / 4R_i R_j) \right]^{\frac{1}{2}} \quad (49)$$

kjer je $R_i \geq \sigma_i$ efektivni Bornov radij in vključuje efekt, da sta lahko atoma obdana s sosednjimi atomi, ki izpodrivajo topilo in tako znižujejo polarizacijsko energijo. Večji efektivni radiji povzročijo manjšo ΔG^{solv} . Vrednosti teh radijev opisujejo stopnjo 'zakopanosti' atoma v proteinu in so odvisne le od medsebojnih pozicij, radijev in parcialnih nabojev sosednih atomov (atomov, ki jih obdajajo). Tako je celotna elektrostatska energija makromolekule v topilu v GB približku dana z

$$A_{\text{razt}}^{el} = \sum_i \frac{q_i^2}{2\epsilon_i R_i} + \sum_{i < j} \frac{q_i q_j}{\epsilon_i r_{ij}} - \frac{1}{2} \sum_{i,j} \left(\frac{1}{\epsilon_i} - \frac{1}{\epsilon_s} \right) \frac{q_i q_j}{\left[r_{ij}^2 + R_i R_j \exp(-r_{ij}^2 / 4R_i R_j) \right]^{\frac{1}{2}}} \quad (50),$$

GB prosta energija solvatacije pa z enačbo

$$\Delta G_{GB}^{solv} = -\frac{1}{2} \sum_{i,j} \left(\frac{1}{\epsilon_i} - \frac{1}{\epsilon_s} \right) \frac{q_i q_j}{\left[r_{ij}^2 + R_i R_j \exp(-r_{ij}^2 / 4R_i R_j) \right]^{\frac{1}{2}}} \quad (51)$$

Sedaj potrebujemo še izraz za R_i . Volumen topljenca definiramo kot volumen, ki je nedosegljiv za molekule vode. Schaefer in Froemel sta na podlagi primerjalne analize aproksimacije zrcalnih nabojev z analitično Kirkwoodovo rešitvijo za sferični protein [21] pokazala, da lahko pri računanju lastne energije naboja uporabimo kar gostoto električnega polja naboja v vakuumu [19] in elektrostatsko energijo atoma k zapišemo kot lastni prispevek in prispevke medsebojno neodvisnih parskih interakcij atoma i z ostalimi k atomi, $k \neq i$. Sprememba lastne energije izoliranega atoma po solvataciji je potem dana z

$$\Delta G_i^{solv} = \frac{\tau q_i^2}{8\pi} \int_{\text{topilo}} \frac{1}{r^4} d^3 r \quad (52)$$

kjer je $\tau = (1/\epsilon_i - 1/\epsilon_s)$ in smo atom postavili v središče koordinatnega sistema. Enačba mora po definiciji biti ekvivalentna spremembi proste energije sfere z efektivnim Bornovim radijem r_i po solvataciji:

$$\frac{\tau q_i^2}{8\pi} \int_{\text{topilo}} \frac{1}{r^4} d^3 r = -\frac{\tau q_i^2}{2R_i} \quad (53)$$

Efektivni radij je potem definiran kot

$$R_i^{-1} = \frac{1}{4\pi} \int_{\text{topilo}} \frac{1}{r^4} d^3 r \quad (54)$$

Sedaj si predstavljajmo dve sferi K_i in K_k z izhodiščema na razdalji r_{ik} in končnima radijema R_i in R_k . Sfera i se nahaja v izhodišču koordinatnega sistema, sfera k pa na pozitivni z-osi. Vpeljemo funkcijo $f(r)$:

$$f(r) = \begin{cases} \frac{1}{r^4}, & \text{če } r < R_i \\ 0, & \text{sicer} \end{cases} \quad (55)$$

Integral v enačbi (52) v primeru atoma i , ki ima naboj enakomerno porazdeljen po površini sfere z radijem R_i , v prisotnosti atoma k je potem dan z enačbo [19]

$$H(r_{ik}, R_i, R_k) = \int_{K_k} f(r) d^3r = 4\pi \int_{K_k/K_i} \frac{dr}{r^2} \quad (56)$$

kjer $K_k/K_i = K_k - (K_i \cap K_k)$ označuje sfero k minus področje preseka sfer i in k . Integral na desni strani enačbe (56) lahko zapišemo tudi preko deleža površine sfere z radijem r s središčem v atomu i , ki je izpostavljen topilu v prisotnosti sfere k z radijem σ_k , C_i :

$$\int_{K_k/K_i} \frac{dr}{r^2} = \int_{\sigma_i}^{\infty} \frac{dr}{r^2} C_i \quad (57)$$

Velja tudi $C_i = 1 - H_i$, kjer je H_i delež senčene površine sfere z radijem r zaradi atoma k z radijem σ_k na razdalji r_{ik} . Sedaj lahko zapišemo

$$\int_{\sigma_i}^{\infty} \frac{dr}{r^2} C_i = \sigma_i^{-1} - \int_{\sigma_i}^{\infty} \frac{dr}{r^2} H_i(r_{ik}, \sigma_k) \quad (58)$$

Z upoštevanjem enačb (54-57) je efektivni Bornov radij dan z enačbo

$$R_i^{-1} = \sigma_i^{-1} - \sum_{k \neq i} \int_{\sigma_i}^{\infty} \frac{dr}{r^2} H_i(r_{ik}, \sigma_k) \quad (59).$$

Enačba (59) je eksaktna, če ni presekov med atomi. V tem primeru lahko integral v enačbi (59) izračunamo analitično [22]. Efektivni Bornov radij lahko potem analitično zapišemo kot

$$R_i^{-1} = \sigma_i^{-1} - \frac{1}{2} \sum_{k \neq i} \left[\frac{1}{L_{ik}} - \frac{1}{U_{ik}} + \frac{r_{ik}}{4} \left(\frac{1}{U_{ik}^2} - \frac{1}{L_{ik}^2} \right) + \frac{1}{2r_{ik}} \ln \frac{L_{ik}}{U_{ik}} + \frac{\sigma_k^2}{4r_{ik}} \left(\frac{1}{L_{ik}^2} - \frac{1}{U_{ik}^2} \right) \right] \quad (60)$$

kjer je

$$\begin{aligned} L_{ik} &= 1 && \text{če } r_{ik} + \sigma_k \leq \sigma_i \\ L_{ik} &= \sigma_i && \text{če } r_{ik} - \sigma_k \leq \sigma_i < r_{ik} + \sigma_k \\ L_{ik} &= r_{ik} - \sigma_k && \text{če } \sigma_i \leq r_{ik} - \sigma_k \end{aligned} \quad (61)$$

in

$$\begin{aligned} U_{ik} &= 1 && \text{če } r_{ik} + \sigma_k \leq \sigma_i \\ U_{ik} &= r_{ik} - \sigma_k && \text{če } \sigma_i < r_{ik} + \sigma_k \end{aligned} \quad (62)$$

Ker protein ni raztopljen v čisti vodi, temveč so prisotni še raztopljeni ioni soli (predvsem Na^+ in Cl^-), moramo v GB model k elektrostatskemu senčenju nabojev atoma proteina z drugimi atomi proteina še senčenje zaradi protionov. Ponovno si oglejmo spremembo interakcijske energije zaradi solvatacije za dva sferična nabita delca, ki sta zelo daleč narazen (enačba (46) brez prispevkov atomskih solvacijskih energij: $\Delta G_2^{\text{solv}} - \Delta G_i^{\text{solv}} - \Delta G_j^{\text{solv}}$) v Debye-Hückel-ovi limiti:

$$\lim_{r_{ij} \rightarrow \infty} \Delta G_{ij; i \neq j}^{\text{solv}} = - \left(\frac{1}{\epsilon_i} - \frac{e^{-\kappa r_{ij}}}{\epsilon_s} \right) \frac{q_i q_j}{r_{ij}} \quad (63)$$

kjer je κ Debye-Hückel-ova dolžina ($\kappa^2 = \sum_i [q_i^2 n_i / (\epsilon_0 \epsilon_s k T)]$), kjer je n_i številska gostota iona i). Solvacijska prosta energija za izoliran sferični ion z radijem a je v Debye-Hückel-ovi limiti dana z enačbo [21]

$$\Delta G_a^{\text{solv}} = - \frac{1}{2} \left(\frac{1}{\epsilon_i} - \frac{1}{\epsilon_s} \right) \frac{q_a^2}{a} - \frac{q_a^2 \kappa}{2\epsilon_s (1 + \kappa a)} \quad (64)$$

kjer je b radialna razdalja, znotraj katere so izključeni ioni soli ($b-a$ je radij iona soli). Limitna primera, ko gre $r_{ij} \rightarrow \infty$ in $r_{ij} \rightarrow 0$ lahko do dobrega približka dobimo ze enostavno substitucijo

$$\left(\frac{1}{\varepsilon_i} - \frac{1}{\varepsilon_s} \right) \rightarrow \left(\frac{1}{\varepsilon_i} - \frac{e^{-\kappa f_{ij}^{GB}}}{\varepsilon_s} \right) \quad (65)$$

v enačbo (48). Ko gre $r_{ij} \rightarrow \infty$, dobimo enačbo (63) direktno. Ko gre $r_{ij} \rightarrow 0$, postanejo (po razvoju do drugega reda v κ) od soli odvisni členi

$$-\frac{q_a^2 \kappa}{2\varepsilon_s \left(1 + \frac{1}{2} \kappa R_i \right)} \quad (66)$$

Po razvoju eksponenta v enačbi (65) po κ se enačbi (65) in (64) ujemata do linearnih členov, v kvadratnih členih v κ , pa dobimo namesto b v imenovalcu $R_i/2$. Ta razlika posledično napove efekte soli, ki so nekoliko večji od tistih, napovedanih z numeričnim reševanjem Poisson-Boltzmannove enačbe. Slednje je verjetno posledica tega, da v GB aproksimaciji ne upoštevamo končne velikosti ionov.

Če vstavimo enačbo (60) v enačbo (50) in uporabimo (65), dobimo analitični izraz, ki opisuje elektrostatsko komponento proste energije solvatacije proteina kot funkcijo atomskih koordinat proteina. Vendar pa je tak pristop k solvataciji pomankljiv, saj vključuje le elektrostatski prispevek solvatacijske proste energije, ne vključuje pa hidrofobnih efektov, ki so zelo pomembni za zvijanje proteina, saj le ti vodijo zvijanje do stanja, v katerem so hidrofobni deli proteina ščiteni pred vodo s hidrofobnimi. Zato si moramo ogledati še teorijo hidrofobnega efekta.

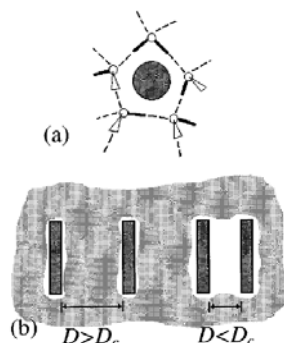
2. 5. 2 Teorija hidrofobnega efekta

Znano je, da efektivne interakcije med nepolarnimi funkcionalnimi skupinami v vodi igrajo pomembno vlogo pri stabilnosti in strukturnih lastnostih bioloških struktur na mezoskopski skali. Vendar pa zaenkrat kvantitativno razumevanje teh efektov še izostaja. Vzrok za to je dualna narava hidrofobnih efektov v vodi glede na velikostno skalo. Hidrofobne interakcije med majhnimi hidrofobnimi topljenci (npr. nižjimi alkani) v razredčenih vodnih raztopinah se bistveno razlikujejo od tistih med večjimi strukturami ali v primeru večjih koncentracij nepolarnih skupin.

Hidrofobne skupine ne tvorijo vodikovih vezi z vodo in tvorijo področja, iz katerih je topilo izključeno. Če so te skupine majhne, se voda lahko reorganizira okoli njih brez izgube vodikovih vezi med lastnimi molekulami (slika 10 a). Entropijska izguba zaradi reorientacije molekul ima za posledico majhno topnost majhnih nepolarnih skupin v vodi. Mehanizem in prosta energija tega efekta je poznana tudi kvantitativno [23]. Vendar pa ta efekt ne implicira asociacije majhnih hidrofobnih skupin v vodi, temveč bi po tej teoriji voda morala take skupine razmikati [23].

V bližini večjega hidrofobnega objekta (npr. večjega asociata majhnih hidrofobnih skupin) je reorganizacija molekul vode brez prekinitve mreže vodikovih vezi geometrijsko nemogoča. Izguba vodikovih vezi lahko povzroči 'sušenje' (tik ob površini hidrofobnega objekta voda preide v plinsko stanje), tako da tik ob površini objekta nastane medfaza plin/tekočina. Sušenje lahko povzroči močne privlačne interakcije med takimi hidrofobnimi objekti, kar je lepo razvidno, če si zamislimo dve večji hidrofobni planarni površini v vodi (slika 10 b levo). Izguba vodikovih vezi povzroči, da se voda odmakne od površine, tako da nastane ob površini tanka plast vodne pare. Fluktuacije v tako nastalih medfazah lahko pri ustrezni razdalji medsebojni razdalji površin destabilizirajo in izključijo preostalo tekočino med površinama. Tako nastalo neravnovesje tlakov povzroči privlak med površinama (slika 10 b desno). Tudi ta pojav je poznan in kvantitativno razložen [24]. Ker pa je sušenje nelinearen pojav, je makroskopska manifestacija pogojena z lokalno strukturo na majhni skali, povezava med njima pa z omenjenima pristopoma ni dosegljiva. V primeru zvijanja proteinov v vodi je, pri popolnoma raztegnjeni konformaciji, hidrofobnost posledica efektov na majhnih skalah, v zvitih proteinih pa

efektov sušenja na velikih skalah. Če hočemo podrobneje razumeti proces zvijanja proteinov (prehod iz enega režima v drugega) moramo torej razumeti hidrofobno solvatacijo in hidrofobne efekte tudi na vmesnih skalah.



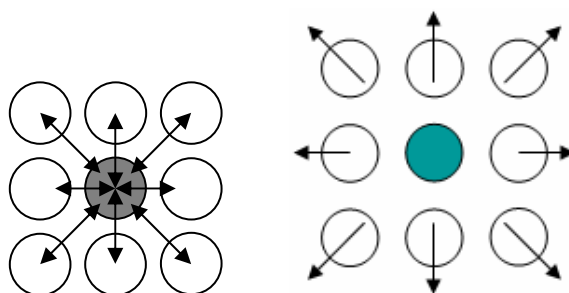
Slika 10: (a) Shematski prikaz lokalne strukture vode okoli manjšega hidrofobnega topljenca (s črtkanimi črtami so označene vezi). (b) Shematski prikaz lokalne strukture vode okoli večjih paralelnih hidrofobnih plošč (šrafirano področje predstavlja območje, kjer je gostota vode enaka bulk (volumski fazi) gostoti vode v tekočem stanju, belo področje pa regije, kjer je gostota vode enaka bulk (volumski fazi) gostoti vodne pare).

Za opis hidrofobnih efektov pri solvataciji proteina zaenkrat še ni dobrega analitičnega modela. Zato se bomo podrobneje osredotočili na opis solvatacije sferičnega delca v Lennard-Jones tekočini s formalizmom posplošenega molekularnega polja (GMF), ki temelji na YBG formalizmu in modifikaciji opisa molekularnega polja. Z nekoliko poenostavljeno različico te teorije lahko kvantitativno opredelimo hidrofobno hidratacijo (solvatacijo v vodi) ter hidrofobne interakcije med cilindri v odvisnosti od njihove velikosti. Razumevanje teh pojavov pa nam bo dalo vpogled na njihov pomen pri zvijanje proteinov.

2. 5. 2. 1 Generalized mean field (GMF) teorija solvatacije sferičnega delca v LJ tekočini [25-28]

Zamislimo si sferični topljenec s trdim odbojem v Lennard-Jones (LJ) topilu. Slednje je dober model za hidrofobno solvatacijo, saj tak topljenec s topilom ne tvori privlačnih interakcij (kot hidrofobni objekt v vodi ne tvori vodikovih vezi z vodo), hkrati pa povzroči efekte izključenega volumna.

V večini konfiguracij LJ tekočine se privlačne sile na izbran delec izključijo med nasprotnimi pari delcev (slika 11 levo).



Slika 11: Shematski prikaz parskega izničenja privlačnih sil na izbran delec topila (sive barve) v homogeni LJ tekočini (levo) in neizničenih interakcij med delci topila okoli topljenca (modro), ki s topilom interagirira s strdim odbojem (desno).

Tako ostanejo le korelacije zaradi izključitve volumna, ki so posledica kratkosežnih močnih odbojnih sil. Ta kratkosežni odboj lahko dodobra aproksimiramo z enakomernim referenčnim sistemom trdih sfer. Argument izničenja privlačnih sil mora (sicer manj točno) veljati tudi v primeru prisotnosti majhnega sferičnega topljenca s trdim odbojem, ki je enako velik kot LJ delci. Odziv gostote na tak (mirujoči) delec bi moral biti podoben oscilirajoči radialni distribucijski funkciji v referenčnem sistemu trdih sfer, z maksimumom gostote pri stiku. Ker pa tak topljenec ne tvori privlačnih interakcij z LJ delci topila, bodo LJ delci pri naraščajočem radiju topljenca čutili naraščajočo 'neizničeno' ali

'neuravnoteženo' neto privlačno silo, ki kaže stran od topljenca, zaradi delcev, ki so bolj oddaljeni od topljenca (slika 11 desno).

Ta neizničena privlačna interakcija povzroči še drugo, počasneje se spreminjajočo, komponento v odzivu gostote na topljenec, ki tekmuje z oscilirajočimi korelacijami izključenega volumna in povzroči manjšo gostoto LJ delcev na stiku s topljencem. Končni profil gostote je potem vsota obeh komponent odziva gostote na topljenec in je odvisen od tega, kako blizu je termodinamsko stanje začetnega homogenega LJ sistema faznemu ravnotežju plin/tekočina in od velikosti neuravnotežene privlačne sile, ki je povezana z velikostjo topljenca (sfere).

Neizničene privlačne interakcije bomo aproksimirali s splošnim molekularnim poljem, v katerem so lokalno povprečni efekti privlačnih interakcij zamenjani z efektivnim enodelčnim potencialom. Strukturo nehomogenega LJ sistema (LJ sistem s topljencem) bomo opisali s strukturo referenčnega sistema trdih sfer v sutreznem efektivnem referenčnem polju (*effective reference field*-ERF), ki je sestavljeno iz dveh delov: golega zunanega polja trdega topljenca (trdi odboj) in počasi se spreminjajočega dela, ki opisuje neizničene privlačne interakcije.

Izhajamo iz velekanoničnega ansambla s fiksiranim kemijskim potencialom μ^B , ki določa ρ^B , gostoto homogene tekočine daleč stran od delca kjer je zunanje polje $\phi(\mathbf{r})=0$. Delci tekočine interagirajo z zunanjim poljem trdega delca, $\phi(\mathbf{r})$, in med seboj z LJ potencialom, $w(r)\equiv u_0(r)+u_1(r)$, ki je razdeljen na hitro padajoči odbojni del u_0 in počasi se spreminjajoči del zaradi privlačnih interakcij, u_1 . Tako je $u_0(r)=w(r)+\varepsilon$ za $r\leq r_0$ (kjer je r_0 razdalja pri kateri LJ sila menja predznak (razdalja pri minimumu potenciala) in ε običajni LJ energetski parameter) in $u_1(r)=-\varepsilon$ za $r\leq r_0$ in $u_1(r)=w(r)$ sicer. Strukturo nehomogenega LJ sistema opredelimo s strukturo bolj preprostega nehomogenega referenčnega sistema (ki vsebuje le odbojne interakcije, torej LJ odboj) in kemijskim potencialom μ^B in pripadajočo gostoto v 'bulk' fazi, ρ^B , ki se nahaja v zunanjem referenčnem polju (ERF), $\phi_R(\mathbf{r})$. Kako izbrati ERF?

Glede na to, da je referenčni sistem čim boljši približek pravega sistema, je smiselno, da je oblika ERF takšna, da sta rezultirajoči lokalni singletni gostoti obeh sistemov enaki v vsaki točki sistema:

$$\rho_0(\mathbf{r};[\phi_R]) = \rho(\mathbf{r};[\phi]) \quad (67)$$

kjer "0" označuje, da so parske interakcije referenčnega sistema izključno odbojne, u_0 , $[\phi_R]$ pa označuje, da sta distribucijski funkciji funkciji ustreznega zunanega polja. ERF lahko dobimo tako, da odštejemo prvi enačbi YGB hierarhije (glej dodatek, enačba (D23)) intrinzičnega in referenčnega sistema ob pogoju, da za ERF velja enačba (67). Prva člena enačbe (D23) za oba sistema sta zaradi (67) enaka, tako po preurejanju členov dobimo

$$-\nabla_1[\phi_R(\mathbf{r}_1) - \phi(\mathbf{r}_1)] = -\int d\mathbf{r}_2 \rho_0(\mathbf{r}_2|\mathbf{r}_1;[\phi_R]) \nabla_1 u_1(r_{12}) - \int d\mathbf{r}_2 \{ \rho(\mathbf{r}_2|\mathbf{r}_1;[\phi]) - \rho_0(\mathbf{r}_2|\mathbf{r}_1;[\phi_R]) \} \nabla_1 w_1(r_{12}) \quad (68)$$

kjer je $\rho_0(\mathbf{r}_2|\mathbf{r}_1;[\phi_R]) \equiv \rho_0(\mathbf{r}_1, \mathbf{r}_2;[\phi_R]) / \rho_0(\mathbf{r}_1;[\phi_R])$ pogojna singletna gostota, torej gostota pri \mathbf{r}_2 pri pogoju, da je en delec fiksiran pri \mathbf{r}_1 . Če enačba (67) zagotovi podobno lokalno okolico odbojnih sredic v obeh tekočinah (v področjih, kjer prevladujejo (identične) odbojne sile), ki potem bistveno (odločilno) vplivajo na korelacije v gostoti preko efekta izključitve volumna, potem bosta tudi pogojni singletni gostoti podobni. Slednje pa pomeni, da bo zadnji člen v enačbi (68) majhen, vsaj pri velikih gostotah, ko prevladujejo efekti izključenega volumna. Ta izjava glede strukture omogoča vpeljavo posplošene teorije molekularnega (ali povprečnega) polja (generalized MF theory-GMF), v kateri ima referenčna tekočina še vedno netrivialne parske korelacije in tudi korelacije višjega reda, ki so posledica odbojnih sil.

Zadnji člen (68) pa očitno izgine tudi pri zelo nizkih gostotah, kjer velja $\phi_R = \phi$. Če na osnovi omenjenih argumentov popolnoma zanemarimo zadnji člen enačbe (68), dobimo približno enačbo za ERF

$$-\nabla_1 [\phi_R(\mathbf{r}_1) - \phi(\mathbf{r}_1)] = -\int d\mathbf{r}_2 \rho_0(\mathbf{r}_2; [\phi_R]) \nabla_1 u_1(r_{12}) \quad (69)$$

Povezava s teorijo povprečnega (ali molekularnega) polja postane še bolj očitna, če pogojno singletno gostoto zamenjamo kar s singletno gostoto $\rho_0(\mathbf{r}_2; [\phi_R])$. Ta aproksimacija je bistveno boljša, kot izgleda na prvi pogled, saj je razlika med tema funkcijama največja, ko je \mathbf{r}_2 blizu \mathbf{r}_1 , vendar pa je pri majhnih meddelčnih razdaljah tudi $\nabla_1 u_1 = 0$. Gradient lahko potem pišemo pred integralom v enačbi (69) in enačbo (69) integriramo, tako da izeberemo konstanto integracije pri pogoju, da je gostota daleč stran od topljenca, kjer je $\phi_R = \phi$, enaka ρ^B . Dobimo enačbo

$$\phi_R(\mathbf{r}_1) - \phi(\mathbf{r}_1) = \int d\mathbf{r}_2 \rho_0(\mathbf{r}_2; [\phi_R]) u_1(r_{12}) + 2\rho^B a \quad (70)$$

kjer $a \equiv -\frac{1}{2} \int d\mathbf{r}_2 u_1(r_{12})$ ustreza interakcijskemu parametru v van der Waalovi teoriji homogenih tekočin. ϕ_R je torej splošno sestavljen iz originalnega polja ϕ in počasi se spreminjajočega dela ϕ_S

$$\phi_S(\mathbf{r}_1) \equiv \int d\mathbf{r}_2 \rho_0(\mathbf{r}_2) u_1(r_{12}) + 2\rho^B a \quad (71)$$

ki opisuje prostorsko spreminjanje privlačnih interakcij (neizničeni privlačni sili) v nehomogeni tekočini. Sedaj si za ilustracijo pogledimo termodinamske implikacije enačbe (70) v primeru konstantnega polja. V velikokanoničnem ansamblu je to ekvivalentno spremembi kemijskega potenciala, ki povzroči spremembo v homogeni (enakomerni) gostoti. V tem primeru lahko uporabimo enačbo (70) za povezavo med kemijskim potencialom v referenčnem in LJ sistemu. Če sta $\mu(\rho)$ in $\mu_0(\rho)$ kemijska potenciala kot funkciji gostote v LJ in referenčnem sistemu (ki sta oba homogena), potem je $\phi^{\rho} = \mu^B - \mu(\rho)$ vrednost konstantnega polja, ki spremeni gostoto v LJ sistemu iz ρ^B na ρ , $\phi_R^{\rho} = \mu_0^B - \mu_0(\rho)$ analogno polje, ki v referenčnem sistemu povzroči enako spremembo gostote. Z uporabo enačb (67) in (70) dobimo MF približek, ki povezuje polji oz. kemijska potenciala v homogenem referenčnem in homogenem LJ sistemu: $\mu(\rho) = \mu_0(\rho) - 2a\rho$. V limiti homogenega sistema ta enačba opisuje vse efekte privlačnih interakcij preko konstantnega parametra a .

Sedaj si oglejmo odziv na splošno polje ϕ_R . Ker kemijski potencial v porazdelitveni funkciji velikokanoničnega ansambla deluje kot konstantno polje, je pripadajoča gostota $\rho_0(\mathbf{r}; [\phi_R], \mu_0^B) \equiv \rho_0(\mathbf{r})$ funkcional ϕ_R in funkcija μ^B in je odvisna le od razlike med njima. Zato lahko za vsako pozicijo \mathbf{r}_1 definiramo premaknjen kemijski potencial $\mu_0^{\mathbf{r}_1} \equiv \mu_0^B - \phi_R(\mathbf{r}_1)$ in premaknjeno polje $\phi_0^{\mathbf{r}_1}(\mathbf{r}) \equiv \phi_R(\mathbf{r}) - \phi_R(\mathbf{r}_1)$, za vsak \mathbf{r} pa mora veljati $\rho_0(\mathbf{r}; [\phi_R], \mu_0^B) = \rho_0(\mathbf{r}; [\phi_0^{\mathbf{r}_1}], \mu_0^{\mathbf{r}_1})$. Po definiciji premaknjeno polje izgine pri $\mathbf{r} = \mathbf{r}_1$. Če se ϕ_R spreminja počasi, potem je polje zelo majhno pri \mathbf{r} blizu \mathbf{r}_1 . V tem primeru lahko gostoto pri \mathbf{r}_1 aproksimiramo z $\rho_0(\mathbf{r}_1; [\phi_R], \mu_0^B) = \rho_0(\mathbf{r}_1; [0], \mu_0^B) \equiv \rho_0^{\mathbf{r}_1}$, torej z gostoto homogene tekočine pri premaknjenem kemijskem potencialu $\mu_0^{\mathbf{r}_1}$. $\rho_0^{\mathbf{r}_1}$ definira hidrostatski odziv na ϕ_R . $\rho_0^{\mathbf{r}_1}$ je odvisna od lokalne vrednosti ϕ_R preko odvisnosti od $\mu_0^{\mathbf{r}_1}$. Če se polje spreminja po prostoru dovolj počasi, je ta hidrostatski približek zelo dober [26]. $\rho_0^{\mathbf{r}_1}$ lahko definiramo tudi s pomočjo kemijskega potenciala homogenega referenčnega sistema: $\mu_0(\rho_0^{\mathbf{r}_1}) \equiv \mu_0^B - \phi_R(\mathbf{r}_1)$. Če ima ϕ_R jedro s trdim odbojem ($\phi_R = \infty$), bo v tem jedru gostota enaka 0 (kar ustreza $\mu_0^{\mathbf{r}_1} = -\infty$). Vendar pa hidrostatski približek zaradi striktno lokalnega odziva odpove pri opisu nelokalnih korelacij, ki so posledica trdega odboja. Sedaj si oglejmo, kako rešiti to pomankljivost.

Najprej zapišimo lokalni hidrostatski odziv na ϕ_R

$$\mu_0(\rho_0^{\text{fi}}) = \mu_0^{\text{B}} - \phi(\mathbf{r}_1) - \int d\mathbf{r}_2 \rho_0(\mathbf{r}_2) u_1(r_{12}) - 2\rho^{\text{B}} a \quad (72)$$

Enačba (72) definira hidrostatsko gostoto ρ_0^{fi} preko lokalne vrednosti ϕ_R pri \mathbf{r}_1 , ki pa vključuje integral preko celotne gostote po vseh ostalih \mathbf{r}_2 točkah prostora. Da lahko eksplicitno določimo ρ_0^{fi} , vpeljemo izraz $\Delta\rho_0(\mathbf{r}_2) = \rho_0(\mathbf{r}_2) - \rho_0^{\text{fi}}$, ga vstavimo v enačbo (72), na obeh straneh odštejemo $2a\rho_0^{\text{fi}}$ in dobimo

$$\mu_0(\rho_0^{\text{fi}}) - 2a\rho_0^{\text{fi}} = \mu_0^{\text{B}} - \phi(\mathbf{r}_1) - 2\rho^{\text{B}} a - \int d\mathbf{r}_2 [\rho_0^{\text{fi}} - \rho_0^{\text{fi}}] u_1(r_{12}) - \int d\mathbf{r}_2 [\Delta\rho_0(\mathbf{r}_2)] u_1(r_{12}) \quad (73)$$

V klasični MF teoriji zadnji člen zanemarimo. Vendar pa lahko $\Delta\rho_0(\mathbf{r}_2)$ enostavno izračunamo z uporabo teorije linearnega odziva, ki natančno povezuje majhne spremembe v polju in gostoti. Splošna enačba linearnega odziva v splošnem polju ϕ_R s kemijskim potencialom μ^{B} pri dani temperaturi se glasi [29]

$$-\beta\delta\phi_R(\mathbf{r}_1) = \int d\mathbf{r}_2 \chi_0^{-1}(\mathbf{r}_1, \mathbf{r}_2; [\rho_0]) \delta\rho_0(\mathbf{r}_2) \quad (74)$$

in povezuje majhne perturbacije polja in gostote preko inverzne odzivne funkcije

$$\chi_0^{-1}(\mathbf{r}_1, \mathbf{r}_2; [\rho_0]) \equiv \frac{\delta(\mathbf{r}_1 - \mathbf{r}_2)}{\rho_0(\mathbf{r}_1)} - c_0(\mathbf{r}_1, \mathbf{r}_2; [\rho_0]) \quad (75)$$

kjer je c_0 direktna (Ornstein-Zernike) korelacijska funkcija sistema z gostoto $\rho_0(\mathbf{r})$. Ker nas zanimajo perturbacije okoli homogenega sistema, dobi c_0 obliko $c_0(r_{12}; \rho)$. Ali lahko s takim pristopom določimo tudi odziv na veliko perturbacijo polja? Linearna relacija mora pri zelo veliki perturbaciji polja (za ilustracijo npr. neskončnem polju) zagotovo odpovedati. Nasprotno enačba (74) najbolj velja za tiste \mathbf{r}_1 , za katere je polje majhno (oz. izgine). Preko integracije po vseh \mathbf{r}_2 pa poveže spremembe gostote v regiji, kjer je polje zelo majhno (oz. izgine) s spremembami v regijah, v katerih je polje veliko. Ta pogoj za optimalno veljavnost teorije linearnega odziva je zadoščen vedno, kadar uporabimo enačbo (74) za določanje spremembe gostote pri vsakem \mathbf{r}_1 glede na gostoto homogene tekočine, ρ_0^{fi} , kot posledico premaknjenega polja. V tem primeru postavimo $\chi_0^{-1}(\mathbf{r}_1, \mathbf{r}_2; [\rho_0]) = \chi_0^{-1}(r_{12}; \rho_0^{\text{fi}})$ in $\delta\phi_R = \phi_R^{\text{fi}}$ ter $\delta\rho_0(\mathbf{r}_2) = \rho_0(\mathbf{r}_2) - \rho_0^{\text{fi}}$. Ker je $\phi_R^{\text{fi}}(\mathbf{r})$ pri \mathbf{r}_1 po definiciji enako 0, dobimo

$$0 = \int d\mathbf{r}_2 \chi_0^{-1}(r_{12}; \rho_0^{\text{fi}}) [\rho_0(\mathbf{r}_2) - \rho_0^{\text{fi}}] \quad (76)$$

kar ob opoštevanju enačbe (75) lahko zapišemo kot

$$0 = \int d\mathbf{r}_2 \frac{\delta(\mathbf{r}_1 - \mathbf{r}_2)}{\rho_0^{\text{fi}}} [\rho_0(\mathbf{r}_2) - \rho_0^{\text{fi}}] - \int d\mathbf{r}_2 c_0(r_{12}; \rho_0^{\text{fi}}) [\rho_0(\mathbf{r}_2) - \rho_0^{\text{fi}}] \quad (77)$$

Po integraciji levega integrala v enačbi (77) dobimo enačbo

$$0 = \frac{\rho_0(\mathbf{r}_1)}{\rho_0^{\text{fi}}} - 1 - \int d\mathbf{r}_2 c_0(r_{12}; \rho_0^{\text{fi}}) [\rho_0(\mathbf{r}_2) - \rho_0^{\text{fi}}] \quad (78)$$

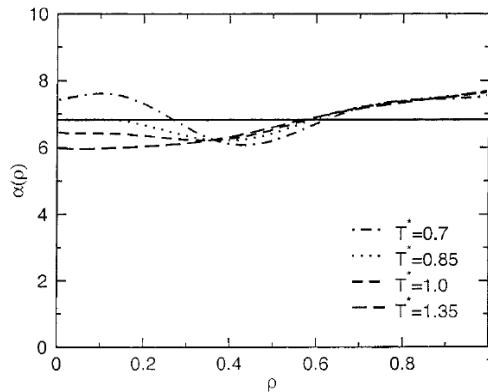
ki po preurejanju členov dobi končno obliko

$$\rho_0(\mathbf{r}_1) = \rho_0^{\text{fi}} + \rho_0^{\text{fi}} \int d\mathbf{r}_2 c_0(r_{12}; \rho_0^{\text{fi}}) [\rho_0(\mathbf{r}_2) - \rho_0^{\text{fi}}] \quad (79)$$

Enačbo (79) imenujemo enačba hidrostatičnega linearnega odziva (*hydrostatic linear response-HLR*). Polje v enačbi nastopa le implicitno, in sicer preko lokalnega efekta na $\rho_0^{\mathbf{r}_1}$. Osnova za uporabo HLR enačbe je poznavanje $c_0(r_{12}; \rho)$ in $\mu_0(\rho)$ homogenega referenčnega sistema. Ustrezni približki teh funkcij so znani iz GMSA teorije [30] ali bolj enostavne Percus-Yevick aproksimacije [29]. Tako lahko z rešitvijo linearne integralske HLR enačbe določimo $\rho_0(\mathbf{r}_1)$ za vsak \mathbf{r}_1 .

Sedaj lahko na osnovi premisleka hidrostatične limite popravimo še MF enačbo (70). V limiti homogenega sistema enačba ustrezno opiše vse efekte privlačnih interakcij preko parametra a . Po drugi strani pa bo napovedala premaknjene (MF) vrednosti gostote plinske in tekoče faze, če bomo z njo želeli opisati počasi se spreminjajočo medfazo tekočina-plin [31].

Če želimo dobiti kvantitativno ustrezne vrednosti 'bulk' termodinamskih lastnosti, lahko zamenjamo konstanto a s funkcijo α , ki je odvisna (upamo, da čim manj) od temperature in gostote. α lahko enostavno empirično določimo, če poznamo točno analitično enačbo stanja homogene tekočine (LJ in referenčnega sistema). Namesto da uporabimo MF enačbo $\mu(\rho) = \mu_0(\rho) - 2a\rho$, predpostavimo, da je $\mu(\rho)$ znana iz točne enačbe stanja (npr. iz enačbe stanja LJ tekočine s 33-parametri, ki so jo predlagali Johnson in sodelavci [32]). Imamo tudi točen izraz za $\mu_0(\rho)$ (tekočina iz trdih sfer). Ti funkciji lahko analogno MF povežemo z enačbo $\mu(\rho) = \mu_0(\rho) - 2\alpha(\rho)\rho$, kjer je funkcija α izbrana tako, da je enačba veljavna. Ker je tudi najenostavnejša MF aproksimacija kvalitativno ustrezna, pričakujemo, da bo razmerje $\alpha(\rho)/a$ reda velikosti 1. Kot je razvidno s slike 12, je temu dejansko tako.



Slika 12: Funkcija $\alpha(\rho)/a$, dobljena iz modificirane MF enačbe ($\mu(\rho) = \mu_0(\rho) - 2\alpha(\rho)\rho$) na osnovi natančne enačbe stanja LJ tekočine in tekočine iz trdih sfer, v odvisnosti od temperature (v LJ enotah) [28].

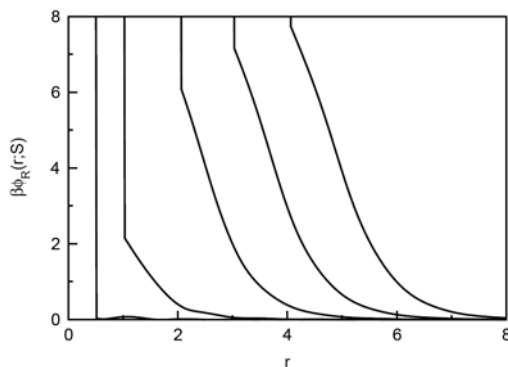
Te rezultate za konstantno polje lahko uporabimo za določitev v hidrostatski limiti počasi se spreminjajočega polja. Razmeroma enostaven način, s katerim lahko to naredimo, je naslednji. MF enačbo (70) zapišemo v modificirani obliki

$$\phi_R(\mathbf{r}_1) - \phi(\mathbf{r}_1) = \frac{\alpha(\rho_0^{\mathbf{r}_1})}{a} \int d\mathbf{r}_2 \rho_0(\mathbf{r}_2; [\phi_R], \mu_0^B) u_1(r_{12}) + 2\alpha(\rho^B) \rho^B \quad (80)$$

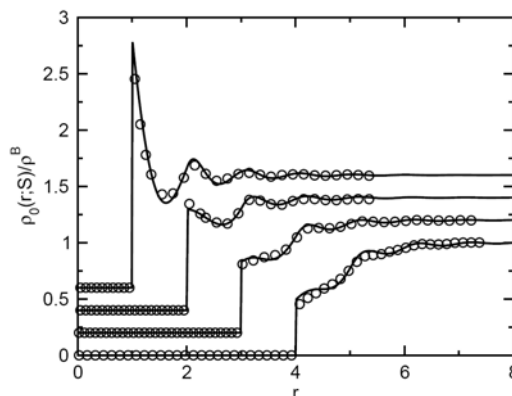
Enačba (80) je modificirana MF enačba (MMF) in predstavlja enačbo (70), v kateri je integral pomnožen s faktorjem $\alpha(\rho_0^{\mathbf{r}_1})/a$, ki je dovisen od \mathbf{r}_1 preko odvisnosti $\rho_0^{\mathbf{r}_1}$ od lokalne vrednosti polja $\phi_R(\mathbf{r}_1)$, konstanta integracije pa je zamenjana z ustrezno limitno vrednostjo tako modificiranega integrala.

MMFA teorija (oz. MVDW teorija) zahteva sočasno rešitev dveh enačb, enačbe (73) oz. (80) in enačbe (79). Enačba (80) določa $\rho_0^{\mathbf{r}_1}$, torej lokalni odziv na ERF (z enačbo $\mu_0(\rho_0^{\mathbf{r}_1}) \equiv \mu_0^B - \phi_R(\mathbf{r}_1)$), enačba (79) pa določa celotni nelokalni odziv $\rho_0(\mathbf{r}_1)$. Celotni proces reševanja enačb lahko torej opišemo z dvostopenjskim iteracijskim postopkom. Za dani približek ERF, se v prvem koraku določi hidrostatska gostota $\rho_0^{\mathbf{r}_1}$ iz enačbe (73) oz. (80), če jo zapišemo analogno kot (73). V drugem koraku se upoštevajo še nelokalni oscilirajoči prispevki k $\rho_0^{\mathbf{r}_1}$ (ki jih splošno povzročijo hitre spremembe

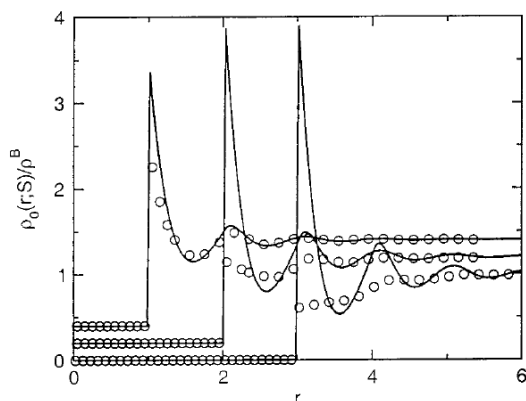
zunanjšega polja ϕ) z uporabo lokalno optimirane uporabe teorije linearnega odziva- enačbe (79). Dobljeno novo gostoto $\rho_0(\mathbf{r}_i)$ nato uporabimo za ponovni izračun približka ERF in koraka ponavljamo, dokler algoritem ne skonvergira. Rezultat tega postopka je ERF in profil gostote. Rezultat v primeru sferičnega delca, ki s toplilom interagira s trdim odbojnim potencialom, je prikazan na slikah 13 in 14. Na sliki 15 so prikazani še profili gostote v referenčni LJ tekočini v primeru golega zunanjšega potenciala (trdega odboja) v odvisnosti od velikosti topljenca.



Slika 13: Zunanje molekularno referenčno polje sferičnega topljenca s trdim odbojem v LJ tekočini pri radiju topljenca 0.5, 1.0, 2.0, 3.0 in 4.0 v LJ enotah



Slika 14: Profil gostote LJ tekočine v prisotnosti sferičnega topljenca s trdim odbojem v LJ tekočini pri radiju topljenca 1.0, 2.0, 3.0 in 4.0 v LJ enotah. Zaradi preglednosti so profili za radij topljenca 1.0, 2.0, 3.0 premaknjeni navzgor za 0.6, 0.4 in 0.2 enote. S krogi so označeni rezultati Monte-Carlo simulacij.



Slika 15: Profil gostote referenčne LJ tekočine v prisotnosti golega zunanjšega polja ϕ pri radiju topljenca 1.0, 2.0 in 3.0 v LJ enotah. Zaradi preglednosti so profili za radij topljenca 1.0, 2.0 premaknjeni navzgor za 0.4 in 0.2 enote. S krogi so označeni rezultati Monte-Carlo simulacij v 'polni' LJ tekočini.

V primeru majhnega topljenca ($S < 0,7$) privlačne interakcije med delci LJ topila ne povzročijo spremembe zunanjšega polja (trdega odboja s topljencom), kar je dobro razvidno na sliki 13 za $S=0,5$. Efekti privlačnih interakcij postanejo pomembni pri $S=1,0$ (ko je velikost topljenca enaka velikosti delca topila), kar je razvidno iz slike 13, saj nad tem radijem topljenca odboj postaja mehkejši.

Kontaktna vrednost gostote pri $S=1,0$ je še vedno večja od 'bulk' vrednosti, vendar pa manjša kot v primeru gostote ob prisotnosti golega zunanega potenciala (trdega odboja- slika 15). Pri $S>1,0$ pride do deplecije (izčrpanja) topila ob topljencu- kontaktna vrednost gostote je manjša od bulk vrednosti (slika 14). Večji kot je topljenec, večja je deplecija delcev topila na površini.

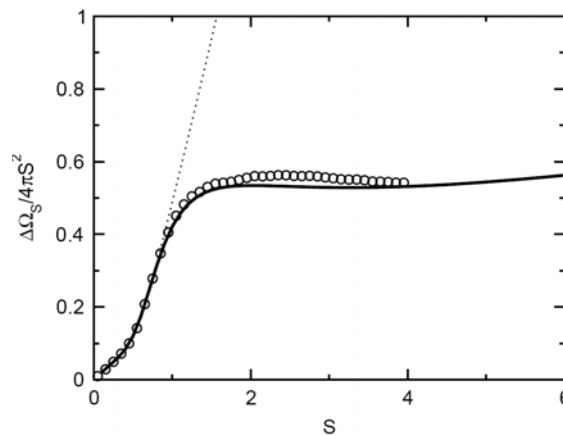
Sedaj si oglejmo še prosto energijo solvatacije sferičnega delca v LJ tekočini. Kemijski potencial sferičnega topljenca A v LJ topilu, $\mu_A(LJ)$, formalno zapišemo kot

$$\mu_A(LJ) = kT \ln[\rho_A(LJ)\Lambda_A^3] + \Delta\Omega_S \quad (81)$$

kjer je $\rho_A(LJ)$ številska gostota topljenca, Λ_A je termična valovna dolžina topljenca in $\Delta\Omega_S$ presežni kemijski potencial ali prosta energija solvatacije - delo ki ga je potrebno opraviti, ko 'vklopimo' interakcije med topilom in topljencem. Po teoriji skaliranja delcev (scaled particle theory) lahko $\Delta\Omega_S$ opišemo z enačbo

$$\Delta\Omega_S = 4\pi S^3 kT \int_0^1 \lambda^2 \rho_\lambda(\lambda S) d\lambda \quad (82)$$

kjer je S radij delca topljenca, $\rho_\lambda(\lambda S)$ je gostota topila na kontaktu z delcem z radijem S in λ fiktivna integracijska spremenljivka, ki teče od 0 do 1 ('vklopimo' interakcije med topljencem in topilom). Integracijo po λ lahko zamenjamo z vsoto in profile gostote izračunamo za posamezne vrednosti λ pri konstantnem μ^B . S takim pristopom lahko izračunamo prosto energijo solvatacije v odvisnosti od velikosti sferičnega topljenca (slika 16)



Slika 16: Prosta energija solvatacije na enoto površine topljenca v odvisnosti od radija topljenca. Krogi označujejo rezultate Monte-Carlo simulacij, črtna črta pa predstavlja $\Delta\Omega_S / 4\pi S^2$ ob prisotnosti golega zunanega polja (trdega odboja).

Za majhne topljence privlačne sile med delci topila ne igrajo pomembne vloge in solvatacijska prosta energija se dobro ujema z energijo v prisotnosti golega zunanega polja (trdega odboja s topljencem). Obnašanje $\Delta\Omega_S / 4\pi S^2$ se drastično spremeni pri $S \approx 0,75$. Nad $S \approx 1,25$ doseže $\Delta\Omega_S / 4\pi S^2$ plato, ki ga lahko razumemo, če prosto energijo razdelimo na volumski prispevek ($V_S = 4\pi S^3/3$) in površinski prispevek ($A_S = 4\pi S^2$)

$$\Delta\Omega_S \approx V_S p^B + A_S \gamma_S \quad (83)$$

Tu prvi člen predstavlja delo potrebno za odstranitev delcev tekočine iz volumna ki ga zaseda topljenec (p^B je tlak v 'bulk' fazi tekočine), drugi del pa delo za nastanek medfaze (γ_S je medfazna

napetost topljenec/topilo). Prvi člen je zanemarljiv za tukaj preučevane velikosti topljenca, drugi del pa je neodvisen od velikosti topljenca za velike topljence.

2. 5. 2. 2 Hidratacija hidrofobnega sferičnega delca, hidrofobne interakcije v vodi in implikacije na zvijanje proteinov

Prikazane principe lahko razširimo na hidratacijo hidrofobnih površin in hidrofobne interakcije v vodi [33]. Zamislimo si trdo sfero v vodi, postavljeno v izhodišču koordinatnega sistema. Tak sistem nam lahko služi kot model hidratacije hidrofobnega objekta, saj sfera ne tvori vodikovih vezi z vodo in izpodrine vodo iz volumna, ki ga zaseda objekt. Privlačne van der Waalsove interakcije med vodo in topljencem ne spremenijo kvalitativne slike fenomena in jih lahko zlahka vključimo v teorijo z nekaj dodatnimi modifikacijami [34]. Zaradi večje preglednosti jih bomo zanemarili.

Majhne fluktuacije gostote v vodi lahko opišemo z Gaussovo distribucijo. Z računalniškimi simulacijami je bilo pokazano [35], da tudi relativno velike fluktuacije, ki povzročijo nastanek lukenj z velikostjo molekul, lahko opišemo z enako Gaussovo distribucijo. Pratt in Chandler sta prva na osnovi eksperimentalno določene odzivne funkcije za vodo razvila kvantitativno teorijo za opis solvatacije majhnih nepolarnih delcev v vodi, ki vključuje spremembe strukture in proste energije, ki jih povzroči izključitev vode iz volumna, ki ga zaseda topljenec [36]. Teorija je bila kasneje še razširjena in teoretično močnejše podkrepljena [35]. Na osnovi zaključkov teh del lahko sklepamo, da lahko z uporabo ustreznih odzivnih funkcij posplošeno teorijo molekularnega polja, predstavljeno v prejšnjem odseku na primeru LJ topila, razširimo na vodne sisteme. Poglavitna lastnost, ki je ne vključuje uporaba Gaussove distribucije za opis verjetnosti fluktuacij gostote, je nastanek medfaze v primeru večjih topljencev. Nastanek medfaze je splošen pojav, zato razlika v lokalni strukturi vode in LJ tekočine nima velikega vpliva na kvalitativno sliko pojava.

Za opis solvatacije sferičnega topljenca, ki z vodo reagira s trdim odbojnim potencialom, je uporabno, da prvi korak reševanja MMFA enačb interpretiramo še drugače. Kot smo pokazali, lahko ERF razdelimo na ϕ (originalno polje-trdi odboj) in ϕ_s (počasi se spreminjajoči del ERF, ki vključuje neizničene privlačne interakcije med molekulami topila). Če želimo upoštevati še privlačne interakcije med topilom in topljencem, jih vključimo v ϕ_s . Zamislimo si najprej lokalni hidrostatični odziv na ϕ_s , ρ_s^i . Ker je odziv na ERF strogo lokalni, se ρ_s^i razlikuje od ρ_0^i (odziv na celotno ERF) le v področju znotraj delca topljenca, kjer je ρ_0^i enaka 0, ρ_s^i pa ostaja zvezna in gladka. Slednje nam omogoča, da lahko pri določanju ρ_0^i v prvem koraku reševanja MMFA enačb, predzadnji člen na desni strani enačbe (73) (oz. analogno zapisane enačbe (80)) aproksimiramo z gradientom, kot bomo pokazali kasneje. Ker ϕ_s opisuje privlačne interakcije, povezane z nastankom medfaze med topljencem in topilom, lahko ρ_s^i interpretiramo kot medfazno komponento polnega odziva gostote na ERF, $\rho_0(\mathbf{r}_i)$. V naslednjem koraku upoštevamo še odziv na ϕ (trdi odboj) in vse nelokalne efekte z enačbo (79). V tem primeru $\rho_0(\mathbf{r}_i)$ izgine znotraj topljenca in, odvisno od tega, v kakšni meri je znižana gostota ρ_s^i blizu topljenca, povzroči nelokalne oscilirajoče popravke ρ_s^i zunaj topljenca, ki jih izračunamo s teorijo linearne odziva. Drugače povedano, v drugem koraku upoštevamo nelokalne Gaussove fluktuacije okoli gladke medfzne komponente profila gostote, ki so posledica trdega odboja. Ne-Gaussovo komponento, ki je povezana z nastankom medfaze in je posledica neizničenih privlačnih interakcij med delci topila, upoštevamo v prvem koraku. Končni profil je tako razmerje med komponentama.

Glede na to interpretacijo teorije posplošenega molekularnega polja lahko sklepamo, da lahko medfazno komponento v vodnem sistemu vsaj kvalitativno opišemo z MMFA pristopom (podobno kot v primeru LJ topila). V LJ sistemu lahko neizničene privlačne interakcije med delci topila (v enačbi (71)) opišemo še drugače, in sicer v smislu grobozrnjene ('coarse-grained') gostote, $\bar{\rho}(\mathbf{r}_i)$, z normalizirano obtežitveno funkcijo, ki je proporcionalna u_i :

$$-2a\bar{\rho}(\mathbf{r}_1) \equiv \int d\mathbf{r}_2 \rho(\mathbf{r}_2) u_1(r_{12}) \quad (84)$$

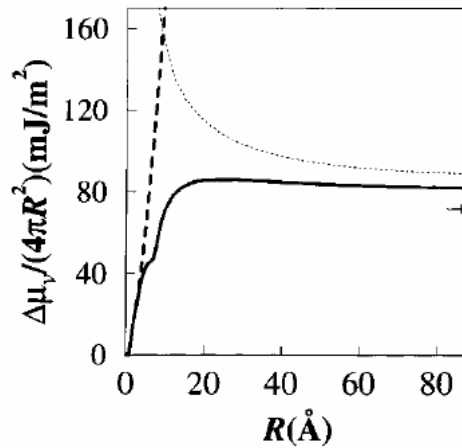
Lum in sodelavci [33] so pokazali, da neizničene privlačne interakcije v vodi lahko opišemo s podobno grobozrnjeno gostoto vode, kjer grobozrnimo znotraj ustreznega dometa privlačnih interakcij v vodi, τ .

Sedaj se vrnemo na (v prejšnjem odstavku na novo interpretiran) prvi korak, v katerem upoštevamo le odziv na ϕ_s , in ρ_s^r . Predzadnji člen na desni strani enačbe (73) razvijemo okoli ρ_s^r do drugega reda. Zaradi simetrije u_1 je člen prvga reda po integraciji po \mathbf{r}_2 enak 0, člen drugega reda pa lahko vzamemo izpod integrala, saj ga obravnavamo pri \mathbf{r}_1 (in zato ni odvisen od \mathbf{r}_2). Ob upoštevanju tega in zveze (84) dobi enačba (73) obliko

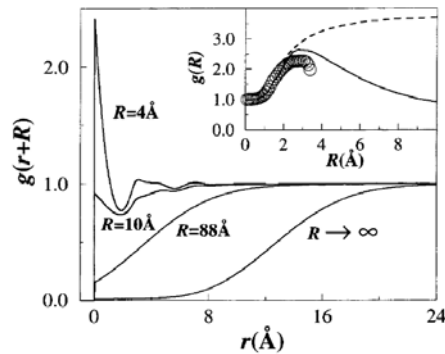
$$\mu_0(\rho_s^r) = \mu^B + m\nabla^2 \rho_s^r + 2a \left[\bar{\rho}(\mathbf{r}_1) - \bar{\rho}_s^r \right] \quad (85)$$

kjer je $m=1/2(\int(r_{12})^2 u(r_{12}) dr_2)$. Enačba (85) je enaka enačbi, ki so jo Lum *in sod.* izpeljali iz gostotnega funkcionala proste energije. Ker bi počasi se spreminjajoča medfazna komponenta ρ_s^r morala biti neodvisna od lokalne strukture, so Lum *in sod.* uporabili enostavno van der Waalsovo (vdW) obliko za $\mu(\rho)$ [33], ki direktno interpolira v področju dveh faz. Parametra a in b so izbrali tako, da sta reproducirala gostoto in stisljivost vode pri temperaturi 298 K. Privzeli so, da je parameter m enak površinski napetosti vode, iz relacije med a in m pa so določili skalo povprečenja (τ) za $\bar{\rho}(\mathbf{r})$, za katero so uporabili kar Gausovo utež. Kriteriji izbire teh parametrov so sicer precej poljubni in vdW relacija ni ravno ustrezna za opis relacije med gostoto energije in površinsko napetostjo vode, vendar so Lum *in sod.* pokazali, da majhne spremembe teh parametrov niso bistveno vplivale na kvalitativne lastnosti rezultatov, ki so jih pri tem dobili.

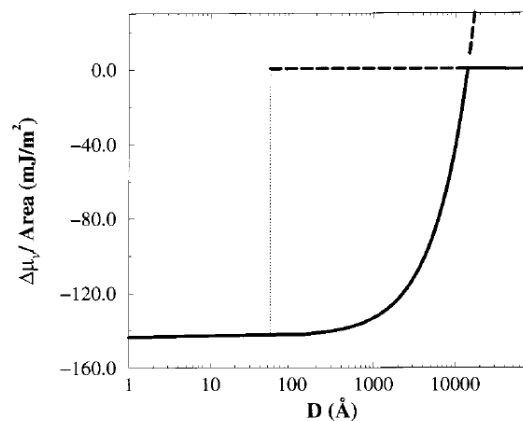
Lum *in sod.* so uporabili standardno obliko odzivne funkcije v njihovem analogu enačbe (76), v kateri so uporabili eksperimentalno določeno radialno distribucijsko funkcijo za tekočo vodo



Slika 17: Prosta energija solvatacije v odvisnosti od radija sferičnega topljenca. ($\Delta\Omega_s / 4\pi S^2 = \Delta\mu_v / 4\pi R^2$)

Slika 18: Radialna distribucijska funkcija $g(R+r) = \rho(R+r) / \rho^B$.

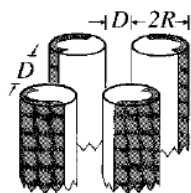
Rezultati, ki so jih dobili, so prikazani na slikah 17 in 18 in se kvalitativno ujemajo z rezultati topljenja v LJ topilu. Na biološko relevantni nanometrski skali opazimo prehod iz 'močenja' (kontaktna gostota večja od bulk vrednosti) na 'sušenje' (kontaktna gostota manjša od bulk vrednosti). Pri izračunih solvatacijskih prostih energij so uporabili nekoliko manj natančen pristop, in sicer so uporabili izraz $-kT \ln P(R)$, kjer je $P(R)$ verjetnost, da najdemo v vodi luknjo velikosti R . Predpostavili so, da je ta verjetnostna porazdelitev Gaussova. Rezultati se kvalitativno ujemajo s solvatacijskimi prostimi energijami v LJ topilu. Izračunali so še solvatacijske proste energije dveh neskončnih ravnih plošč (ki z vodo interagirajo s trdim odbojem) v odvisnosti od razdalje med njima (D) z enačbo (85). Za določeno območje D ima enačba (85) dve rešitvi, torej v tem območju obstajata dve solvatacijski energiji (slika 19).



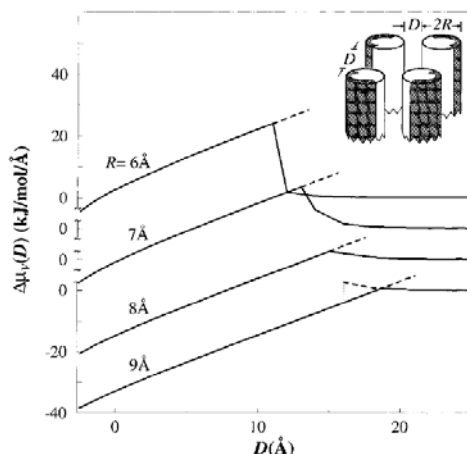
Slika 19. Solvatacijska prosta energija na enoto površine v odvisnosti od medsebojne razdalje plošč. Polna črta predstavlja stabilno območje, črtkana metastabilno in pikčasta spodnjo mejo metastabilnosti tekoče faze.

Iz slike 19 je razvidno, da je tudi pri relativno velikih razdaljah plinska faza termodinamsko bolj stabilna od tekoče. Tekoča faza je v širokem intervalu D metastabilna. V metastabilnem območju je uparevanje posredovano s fluktuacijami medfaz vode, ki morajo biti tako velike, da med ploščama nastanejo mostički plina. Slednje se zgodi zelo redko, zato voda ostane v tekoči obliki do medsebojne razdalje plošč, pri kateri postane mehansko nestabilna (spinodalna meja označena na sliki 19 s točkasto črto). Ta mejna razdalja je pribl. 5 nm.

Podobno so ugotovili še za hidrofobne interakcije med štirimi paralelnimi cilindri. Da bi se približali problemu zvijanja proteinov, so izračunali še hidrofobne interakcije med cilindri, katerih zunanji deli so imeli hidrofilne lastnosti, deli, ki pa so bili obrnjeni drug proti drugemu, pa hidrofobne (slika 20). Slednje je model za interakcije med sekundarnimi vijačnimi strukturnimi elementi proteinov, ki imajo obliko valja in podobne hidrofilne/hidrofobne lastnosti.



Slika 20: interakcije med hidrofoniimi cilindri s hidrofilno zunanostjo (senčeno področje).



Slika 21: Solvacijske proste energije na enoto dolžine štirih neskončnih paralelnih cilindrov (s slike 20) v odvisnosti od razdalje med površinami. Polne črte predstavljajo stabilna področja, črtkane metastabilna, točkasta črta pa spodnjo mejo metastabilnosti tekoče faze (faze z veliko gostoto).

V primeru cilindrov z radijem $R \geq 0,8$ nm, opazimo območje razdalj D , kjer obstaja metastabilno tekoče (stanje z visoko gostoto) - označeno na sliki 21 s črtkano črto. V primeru manjših cilindrov področja metastabilnega tekočega stanja ne opazimo. Plinska faza (faza z nizko gostoto) je v vseh primerih stabilna vsaj do $D \leq 1,1$ nm. Metastabilna plinska stanja opazimo pri razdaljah od pribl. 1,3 nm (pri $R=0,6$ nm) do pribl. 2,1 nm (pri $R=0,9$ nm). Stabilnost oz. metastabilnost te faze predstavlja gonilno silo asociacije cilindrov. Pri $R=0,9$ nm energetska bariera (dvig solvacijske proste energije pri približevanju) za asociacijo izgine celo v področju metastabilne plinske faze. V tem področju torej lahko pričakujemo hitro in močno asociacijo cilindrov za $R \geq 0,9$ nm. Dejstvo, da izračuni v tem področju radijev cilindrov in medsebojnih razdalj med njimi niso pokazali energetske bariere, ne pomeni, da te zagotovo ni, saj teorija LCW ne upošteva fluktuacij medfaz. Le-te lahko namreč v primeru, ko se med cilindri nahaja metastabilna plinska faza, povzročijo nastanek tekočih mostičkov med cilindri, kar zmanjša neravnovesje tlakov zunaj cilindrov in med njimi (okolice cilindrov je tlak večji kot v področju med njimi), ki povzroča privlak med njimi, slednje pa predstavlja bariero asociacije.

3. Pripisevek posameznih intramolekularnih interakcij k prosti energiji zvijanja in stabilnosti native konformacije

Pri računalniških metodah za računanje proste energije zvijanja proteinov igra pomembno vlogo koncept vodi dostopne površine (Accessible Surface Area). V ASA modelu se solvacijska prosta energija atoma opiše kot produkt solvacijskega parametra (solvacijska energija na enoto površine) in površine atoma, ki je dostopna molekulam topila. Topilu dostopna površina (ASA) je površina, ki obdaja področje izključenega volumna in se izračuna iz molekulskih vdW radijev. Solvacijski parameter je dobljen empirično s korelacijsko analizo energij prenosa manjših organskih molekul iz nepolarne organske faze (npr. oktanol) v vodo, ki jih izračunamo iz meritev porazdelitvenih koeficientov, dobljenih npr. iz tekočinske kromatografije. Porazdelitveni koeficienti predstavljajo razmerje energij prenosa iz vakuumu v vodo in iz vakuumu v olje. S pomočjo molekulske mehanike

se nato izračuna energija molekule v vakuumu, ki posledično omogoči izračun energije prenosa iz vakuuma v vodo.

Iz atomističnih računalniških simulacij molekulske dinamike (MD) na osnovi CHARMM polja sil in ASA ter Poisson-Boltzmann (PB; opisuje elektrostatsko komponento) modelov solvatacije je znano [37,38], da je prispevek veznih intramolekularnih interakcij (odsek 2.3) k spremembi potencialne energije pri zvijanju proteina v primerjavi z neveznimi zelo majhen. Po rezultatih teh simulacij 60-70 % spremembe potencialne energije prispevajo vdW interakcije in 25-30 % elektrostatske interakcije. Vodikove vezi so v CHARMM potencialni funkciji opisane s Coulombskim potencialom. Prispevke interakcij k prosti energiji zvijanja pa so razdelili na prispevke polarnih in nepolarnih funkcionalnih skupin. Med prispevke polarnih so uvrstili vdW interakcije med polarnimi skupinami, intramolekularne elektrostatske interakcij, elektrostatske (PB) prispevke proste energije solvatacije in $\frac{1}{2}$ prispevkov interakcij med polarnimi in nepolarnimi skupinami. Med prispevke nepolarnih skupin so uvrstili vdW interakcije med nepolarnimi skupinami, $\frac{1}{2}$ prispevkov vdW interakcij med polarnim in nepolarnimi skupinami in hidrofobne prispevke (ASA) proste energije solvatacije. Ugotovili so, da prevladujoči prispevek k prosti energiji zvijanja proteinov predstavljajo prispevki nepolarnih skupin, kar se ujema s tradicionalno idejo, da je hidrofobni kolaps glavna gonilna sila zvijanja. Vloga polarnih skupin je zagotavljanje specifičnosti in edinstvenosti konformacij z najnižjo energijo, kar pomeni, da so proste energije ostalih nenativnih konformacij večje od proste energije nativne konformacije zaradi neugodnih polarnih interakcij.

Tak model za opis solvatacijske energije ima dve večji pomanjkljivosti. Prva je ta, da je dobljena solvatacijska energija polarnih skupin prevelika [38], druga pa, kot smo pokazali v odseku 2. 5. 2, da je solvatacijski parameter neodvisen od velikosti topljenca (v tem primeru atoma oz. funkcionalne skupine) samo v primeru velikih topljencev (kar atomi in izolirane funkcionalne skupine vsekakor niso). Posledično ASA model v najboljšem primeru relativno ustrezno opiše le hidrofobno komponento solvatacijske proste energije zvite (nativne) konformacije. V primeru iztegnjene konformacije je z ASA modelom napovedana solvatacijska prosta energije precenjena, torej je prispevek hidrofobnih efektov k zvitju proteina še toliko večji.

4. Zaključek

V seminarju smo na kratko opredelili zapleteno strukturo proteinov in predstavili fizikalni pristop, s katerim jo lahko napovedujemo iz AK zaporedja. Prikazali smo pojmovno razliko med proteini in naključnimi polipeptidi, ki so prav tako sestavljenih iz 20 AK, in pokazali, da so proteini evolucijsko izbrane polipeptidne sekvence. Predstavili smo osnove statistične mehanike proteinskih makromolekul in gonilne sile zvijanja proteinov. Elektrostatsko komponento proste energije solvatacije smo opisali z analitičnim posplošenim Bornovim modelom. Predstavili smo tudi GMFA teorijo, najprej splošno, nato še aplicirano na opis hidrofobnega efekta tako na molekularni kot tudi mezo in makroskopski skali. V zadnjem odseku pa smo omenili še nekaj zaključkov računalniških simulacij glede prispevkov posameznih tipov interakcij k prosti energiji zvijanja in stabilnosti nativnega stanja.

Računalniške simulacije so sicer najmočnejše teoretično orodje biofizike in nudijo veliko informacij o bioloških procesih na molekularni (oz. atomistični) ravni, vendar pa zaradi omejenosti na nanosekundne časovne skale ne omogočajo vpogleda v proces zvijanja proteinov, ki se dogaja na mikrosekundni skali. Zato vsak teoretično razdelan korak, ki prispeva k bolj poglobljenemu razumevanju posameznih efektov, predstavlja osnovo za nadaljnji razvoj in izpopolnjevanje simulacijskih orodij, kar je bila tudi motivacija tega seminarja.

Če združimo zaključke iz računalniških simulacij in teorije iz odseka 2, lahko izrišemo naslednjo fizikalno sliko zvijanja proteinov. Poglavitna gonilna sila za nastanek sekundarne strukture proteinov je zmanjšanje izpostavljenosti hidrofobnih funkcionalnih skupin molekulam vode, pri čemer je specifična konformacija elementov sekundarne strukture stabilnejša od nespecifične kolabirane konformacije (v kateri so hidrofobne kisline pravtako senčene pred vodo s hidrofilnimi, vendar pa

interakcije med aminokislinami niso optimalne) zaradi specifičnih interakcij med polarnimi skupinami (vodikovih vezi in elektrostatskih ter dipolnih interakcij). Hidrofobni deli elementov sekundarne strukture asociirajo naprej zaradi hidrofobnega efekta na mezoskopski skali, kjer nastanek specifičnih interakcij (vodikovih vezi in elektrostatskih ter dipolnih interakcij) med polarnimi deli elementov sekundarne strukture zagotovi večjo stabilnost specifične terciarne strukture nativnega stanja glede na nespecifične asociate elementov sekundarne strukture. Ker je osnova hidrofobnega efekta entropijska sprememba vode, lahko torej zaključimo z izjavo, da je zvijanje proteinov vodeno predvsem s preurejanjem molekul vode do strukture, v kateri je ob optimalnih intramolekularnih interakcijah v najmanjši možni meri zmotena mreža vodikovih vezi med molekulami vode.

Dodatek: Yvon-Born-Green (YBG) formalizem tekočin [29,39]

Zamislimo si sistem točkastih delcev (le zaradi večje preglednosti, sicer lahko, kot bomo pokazali, upoštevamo končno velikost) pri temperaturi T in kemijskim potencialom μ v posodi z volumnom V . Velikokanonični in kanonični potencial tega sistema sta povezana z relacijo

$$\exp[-\beta\Omega(T, \mu, V)] = \sum_N \exp(\beta\mu_N) \exp[-\beta F(T, N, V)] \quad (D1)$$

kjer je $\Omega(T, \mu, V)$ velikokanonična prosta energija in $A(T, N, V)$ kanonična (Helmholtzova) prosta energija sistema z N delci. Glede na nadaljni postopek izbira ansambla ni pomembna, vendar je termodinamska povprečja potrebno definirati v izbranem ansamblu. Mi bomo Yvon -Born-Green formalizem izpeljali v kanoničnem ansamblu, ekvivalentne rezultate pa dobimo tudi v velikokanoničnem ansamblu. Za sistem delcev s parsko aditivnimi silami lahko dobimo F iz konfiguracijskega integrala

$$\exp[-\beta F(T, N, V)] = \frac{1}{N! \Lambda^{3N}} \int_V d\mathbf{r}_1 \dots d\mathbf{r}_N \exp\left(-\beta \sum_{i>j} V_2(|\mathbf{r}_i - \mathbf{r}_j|)\right) \quad (D2)$$

kjer je V_2 parski potencial. Tukaj naredimo dve modifikaciji enačbe (D2). Najprej odstranimo omejitve integrala po konfiguracijskem prostoru, tako da vpeljemo enodelčni potencial $\phi_1(\mathbf{r})$, ki strmo naraste, ko se \mathbf{r} približa steni posode z volumnom V . Nato odstranimo še omejitve na parsko aditivne sile, tako da potencialno energijo N delcev, U_N , zapišemo bolj splošno kot

$$\begin{aligned} U_N(\mathbf{r}_1, \dots, \mathbf{r}_N) &= \sum_{j=1}^N \phi_1(\mathbf{r}_j) + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \phi_2(\mathbf{r}_i, \mathbf{r}_j) + \frac{1}{3!} \sum_{\substack{i,j,k=1 \\ i \neq j \neq k}}^N \phi_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \dots \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\substack{j_1=1 \\ j_1 \neq j_2 \neq \dots \neq j_n}}^N \dots \sum_{j_n=1}^N \phi_n(\mathbf{r}_{j_1}, \dots, \mathbf{r}_{j_n}) \end{aligned} \quad (D3)$$

kjer so potencialna polja $\{\phi_n\}$ ($n=1,2,\dots$) simetrične funkcije argumentov ($\phi_2(\mathbf{r}, \mathbf{r}') = \phi_2(\mathbf{r}', \mathbf{r})$), ki opisujejo n -delčno interakcijo. Glede na sistem lahko določeni členi ϕ_n izginejo. U_N je torej funkcija konfiguracije delcev $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$ in funkcional ϕ_N -polj, kar označimo z $U_N = U_N(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; [\phi])$, kjer je $[\phi] = [\phi_1, \phi_2, \dots]$. Podobno je tudi F zdaj funkcija T in N in funkcional ϕ_N -polj, $F = F(T, N; [\phi])$. Posledično F več ni eksplicitna funkcija V , saj je ta določen s funkcionalno odvisnostjo od ϕ_1 (bolj eksplicitno od $\phi_1(\mathbf{r}; V)$). Enačbo (D2) lahko sedaj zapišemo v obliki

$$\exp[-\beta F(T, N; [\phi])] = \frac{1}{N! \Lambda^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N \exp\{-\beta U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\phi])\} \quad (D4)$$

Konfiguracijski intergral ni več omejen, saj ima v prisotnosti $\phi_1(\mathbf{r}; V)$ vsaka konfiguracija, ki ne pripada V , statistično utež 0 (saj je takrat $U_N = \infty$). Desna stran enačbe (D4) vsebuje integral po vseh možnih konfiguracijah $\{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N\}$, zato mora biti invariantna na vsako gladko obrnljivo transformacijo

koordinat, $\mathbf{r}_j \rightarrow \mathbf{r}'_j$ ($j=1, \dots, N$). Če zapišemo to transformacijo s pomočjo lokalnega umeritvenega polja, $\mathbf{u}(\mathbf{r})$, kot $\mathbf{r}'_j = \mathbf{r}_j + \mathbf{u}(\mathbf{r}_j)$, potem lahko novo konfiguracijo $\{\mathbf{r}'_1, \dots, \mathbf{r}'_N\}$ zapišemo kot funkcijo začetne $\{\mathbf{r}_1, \dots, \mathbf{r}_N\}$ in funkcional $\mathbf{u}(\mathbf{r})$, ki smo ga vpeljali za umeritev različnih transformacij $\{\mathbf{r}_j\} \rightarrow \{\mathbf{r}'_j\}$:

$$U_N(\mathbf{r}'_1, \dots, \mathbf{r}'_N; [\phi]) \equiv U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\mathbf{u}], [\phi]) \quad (\text{D5})$$

$\mathbf{u}(\mathbf{r})=0$ ustreza začetni konfiguraciji. V tem primeru tudi prosta energija postane funkcional umeritvenega polja in enačbo (D4) lahko zapišemo kot

$$\begin{aligned} & \exp[-\beta F(T, N; [\mathbf{u}], [\phi])] = \\ & = \frac{1}{N! \Lambda^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N \text{Jac}(\mathbf{r}_1, \dots, \mathbf{r}_N; [\mathbf{u}]) \exp\{-\beta U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\mathbf{u}], [\phi])\} \\ & = \frac{1}{N! \Lambda^{3N}} \int d\mathbf{r}'_1 \dots d\mathbf{r}'_N \exp\{-\beta U_N(\mathbf{r}'_1, \dots, \mathbf{r}'_N; [\phi])\} \\ & = \frac{1}{N! \Lambda^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N \exp\{-\beta U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\phi])\} \\ & = \exp[-\beta F(T, N; [\phi])] \end{aligned} \quad (\text{D6})$$

kjer smo zaradi večje preglednosti uporabili $U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\mathbf{u}=0], [\phi]) = U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\phi])$ in $F(T, N; [\mathbf{u}=0], [\phi]) = F(T, N; [\phi])$. Jac označuje Jacobi-jevo transformacijsko matriko, ki je prav tako funkcional $\mathbf{u}(\mathbf{r})$:

$$\text{Jac}(\mathbf{r}_1, \dots, \mathbf{r}_N; [\mathbf{u}]) \equiv \frac{\partial(\mathbf{r}'_1, \dots, \mathbf{r}'_N)}{\partial(\mathbf{r}_1, \dots, \mathbf{r}_N)} \quad (\text{D7})$$

Ker sta $\{\mathbf{r}_j\}$ in $\{\mathbf{r}'_j\}$ neomejeni spremenljivki, je prosta energija v enačbi (D6) neodvisna od izbire umeritvenega polja:

$$F(T, N; [\mathbf{u}], [\phi]) = F(T, N; [\mathbf{u}=0], [\phi]) \quad (\text{D8})$$

Enačba (D8) nam pove, da je prosta energija invariantna na transformacijo koordinat v konfiguracijskem prostoru.

Sedaj pa si, namesto končne spremembe konfiguracije sistema, predstavljajmo infinitizimalno spremembo konfiguracije $\delta\mathbf{u}(\mathbf{r})$. Enačbo (D8) lahko potem zapišemo kot

$$\begin{aligned} \delta F(T, N; [\mathbf{u}], [\phi]) & = F(T, N; [\mathbf{u}], [\phi]) - F(T, N; [0], [\phi]) \\ & = \int d\mathbf{r} \left. \frac{\delta F(T, N; [\mathbf{u}], [\phi])}{\delta \mathbf{u}(\mathbf{r})} \right|_{T, N, \phi, \mathbf{u}=0} \cdot \delta \mathbf{u}(\mathbf{r}) + \mathcal{O}(\delta \mathbf{u}^2) \\ & = 0 \end{aligned} \quad (\text{D9})$$

in ker lahko $\delta\mathbf{u}$ poljubno izberemo (enačba D8), mora veljati, da

$$\left. \frac{\delta F(T, N; [\mathbf{u}], [\phi])}{\delta \mathbf{u}(\mathbf{r})} \right|_{T, N, \phi, \mathbf{u}=0} = 0 \quad (\text{D10})$$

kar bomo zaradi večje preglednosti zapisali kot

$$\left. \frac{\delta F[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right|_0 = 0 \quad (\text{D11})$$

kar pa pomeni, da T, N, ϕ in vsa polja ϕ_n ($n=1,2,\dots$) držimo konstantne (sledi, da je tudi V fiksiran), medtem ko odvod izvednotimo pri $\mathbf{u}=0$. V tem poenostavljenem zapisu imamo torej

$$\delta F[\mathbf{u}] = \int d\mathbf{r} \left. \frac{\delta F[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right|_0 \cdot \delta \mathbf{u}(\mathbf{r}) + O(\delta \mathbf{u}^2) = 0 \quad (\text{D12})$$

Iz enačbe (D12) lahko razberemo, da pri (namišljeni) spremembi konfiguracije za $\delta \mathbf{u}$ ni opravljeno nobeno delo oz. da je sila, ki je konjugirana (sklopljena) s tem (namišljenim) premikom v konfiguracijskem faznem prostoru enaka 0. Izračunajmo sedaj to silo.

V poenostavljenem zapisu imamo

$$-\beta F[\mathbf{u}] = \ln \frac{1}{N! \Lambda^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N \text{Jac}(\mathbf{r}_1, \dots, \mathbf{r}_N; [\mathbf{u}]) \exp\{-\beta U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\mathbf{u}], [\phi])\} \quad (\text{D13})$$

$$\begin{aligned} -\beta \left. \frac{\delta F[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right|_0 &= \ln \frac{\exp(F[\mathbf{u}])}{N! \Lambda^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N \left(\frac{\delta \text{Jac}}{\delta \mathbf{u}(\mathbf{r})} \exp(-\beta U_N) - \beta \frac{\delta U_N}{\delta \mathbf{u}(\mathbf{r})} \text{Jac} \exp(-\beta U_N) \right) \\ &= -\beta \left\langle \frac{\delta U_N[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right\rangle_{\mathbf{u}} + \left\langle \frac{\delta \text{Jac}[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right\rangle_{\mathbf{u}} \end{aligned} \quad (\text{D14})$$

Ob upoštevanju $\text{Jac}[\mathbf{u}=0]=1$ (enačba (D7)) in enačbe (D11) pa lahko zapišemo še

$$\left. \frac{\delta F[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right|_0 = \left\langle \frac{\delta U_N[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right|_0 + kT \left\langle \frac{\delta \text{Jac}[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \right|_0 = 0 \quad (\text{D15})$$

Povprečja $\langle \dots \rangle$ so vzeta po kanoničnem ansamblu. Sedaj moramo vpeljati le še distribucijske funkcije, in sicer tako da enačbo (D3) zapišemo kot

$$U_N(\mathbf{r}_1, \dots, \mathbf{r}_N; [\phi]) = \sum_{n=1}^{\infty} \frac{1}{n!} \int d\mathbf{x}_1, \dots, d\mathbf{x}_n \hat{\rho}_n(\mathbf{x}_1, \dots, \mathbf{x}_n; \{\mathbf{r}_j\}) \phi_n(\mathbf{x}_1, \dots, \mathbf{x}_n) \quad (\text{D16})$$

kjer so $\hat{\rho}_n$ mikroskopske gostote:

$$\begin{aligned} \hat{\rho}_1(\mathbf{x}_1; \{\mathbf{r}_j\}) &= \sum_{j=1}^N \delta(\mathbf{x}_1 - \mathbf{r}_j) \\ \hat{\rho}_2(\mathbf{x}_1, \mathbf{x}_2; \{\mathbf{r}_j\}) &= \sum_{i=1}^N \sum_{j=1}^N \delta(\mathbf{x}_1 - \mathbf{r}_i) \delta(\mathbf{x}_2 - \mathbf{r}_j) \quad , \text{ itd.} \end{aligned} \quad (\text{D17})$$

Povprečja mikroskopskih gostot po ustreznem ansamblu (mi bomo nadaljevali v kanoničnem) definirajo običajne distribucijske funkcije (ali lokalne gostote):

$$\rho_1(\mathbf{x}_1) = \langle \hat{\rho}_1(\mathbf{x}_1) \rangle, \quad \rho_2(\mathbf{x}_1, \mathbf{x}_2) = \langle \hat{\rho}_2(\mathbf{x}_1, \mathbf{x}_2) \rangle, \quad \text{itd.} \quad (\text{D18})$$

Iz enačbe (D16) sledi, da so ρ_n konjugirane s ϕ_n , tako da je

$$\frac{\delta[-\beta F(T, N; [\phi])]}{\delta[-\beta \phi_n(\mathbf{x}_1, \dots, \mathbf{x}_n)]} = \frac{1}{n!} \rho_n(\mathbf{x}_1, \dots, \mathbf{x}_n) \quad (\text{D19})$$

Ob upoštevanju (D14-D18) ter simetričnosti ϕ_n lahko zapišemo

$$\begin{aligned}
& \left\langle \frac{\delta U_N[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \Big|_0 \right\rangle = \\
& = \frac{\delta}{\delta \mathbf{u}(\mathbf{r})} \left(\sum_{n=1}^{\infty} \frac{1}{n!} \int d\mathbf{x}_1, \dots, d\mathbf{x}_n \rho_n(\mathbf{x}_1, \dots, \mathbf{x}_n) [\mathbf{u}(\mathbf{x}_1) \cdot \nabla_1 + \dots + \mathbf{u}(\mathbf{x}_n) \cdot \nabla_n] \phi_n(\mathbf{x}_1, \dots, \mathbf{x}_n) \right) \Big|_0 \quad (\text{D20}) \\
& = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int d\mathbf{x}_1, \dots, d\mathbf{x}_{n-1} \rho_n(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{r}) \nabla \phi_n(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{r})
\end{aligned}$$

in podobno še

$$\left\langle \frac{\delta \text{Jac}[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{r})} \Big|_0 \right\rangle = \frac{\delta}{\delta \mathbf{u}(\mathbf{r})} \left(\int d\mathbf{x}_1 \rho_1(\mathbf{x}_1) \nabla_1 \cdot \delta \mathbf{u}(\mathbf{x}_1) \right) \Big|_0 = -\nabla \rho_1(\mathbf{r}) \quad (\text{D21})$$

Če vstavimo (D20) in (D21) v (D15) dobimo

$$kT \nabla \rho_1(\mathbf{r}) + \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \int d\mathbf{x}_1, \dots, d\mathbf{x}_{n-1} \rho_n(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{r}) \nabla \phi_n(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{r}) = 0 \quad (\text{D22})$$

kar lahko eksplicitno zapišemo kot

$$kT \nabla \rho_1(\mathbf{r}) + \rho_1(\mathbf{r}) \nabla \phi_1(\mathbf{r}) + \int d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}') \nabla \phi_2(\mathbf{r}, \mathbf{r}') + \dots = 0 \quad (\text{D23})$$

Enačba je prvi člen YBG hierarhije v prisotnosti enodelčnih, dvodelčnih, ... sil. Z vpeljavo Ursellovih funkcij (imenovanih tudi posplošeni strukturni faktorji) dobimo še ostale člene hierarhije, npr. drugi člen

$$kT \nabla \rho_2(\mathbf{r}, \mathbf{r}') + \rho_2(\mathbf{r}, \mathbf{r}') \nabla \phi_1(\mathbf{r}) + \rho_2(\mathbf{r}, \mathbf{r}') \nabla \phi_2(\mathbf{r}, \mathbf{r}') + \int d\mathbf{r}'' \rho_3(\mathbf{r}, \mathbf{r}', \mathbf{r}'') \nabla \phi_3(\mathbf{r}, \mathbf{r}', \mathbf{r}'') + \dots = 0 \quad (\text{D24})$$

itn. Splošno dobimo n-ti člen hierarhije, če n-1 krat operiramo na enačbo (D15)

$$\frac{\delta^{n-1}}{\delta \phi_1(\mathbf{x}_1) \dots \delta \phi_1(\mathbf{x}_{n-1})} \left(\frac{\delta F[\mathbf{u}]}{\delta \mathbf{u}(\mathbf{x}_n)} \right) \Big|_0 = 0 \quad (\text{D25})$$

Literatura

- [1] A. M. Lesk, *Introduction to Protein Architecture*, Oxford University Press, Oxford, 2001.
- [2] M. S. Weiss, A. Jabs, and R. Hilgenfeld, *Peptide bonds revisited*, *Nat. Struct. Biol.* 5 (1998) 676.
- [3] J. F. Swain and L. M. Gierash, *A new twist for an Hsp70 chaperone*, *Nat. Struct. Biol.* 9 (2002) 406.
- [4] V. I. Lim and A. S. Spirin, *Stereochemical analysis of ribosomal transpeptidation conformation of nascent peptide*, *J. Mol. Biol.* 188 (1986) 565.
- [5] C. B. Anfinsen, *Principles that govern the folding of protein chains*, *Science* 181 (1973) 223.
- [6] D. A. McQuarrie, *Statistical Mechanics*, Harper & Row, New York, 1976.
- [7] P. Echenique, I. Calvo, and J. L. Alonso, *Quantum mechanical calculation of the effects of stiff and rigid constraints in the conformational equilibrium of the Alanine dipeptide*, *J. Comp. Chem.* 27 (2006) 1748. ; P. Echenique, *Introduction to protein folding for physicists*, *Contemporary Physics* 48 (2007) 81.
- [8] T. Lazaridis and M. Karplus, *Effective energy function for proteins in solution*, *PROTEINS: Struct. Funct. Gen.* 35 (1999) 133.
- [9] A. D. MacKerell, D. Bashford, M. Bellott, *et al.*, *All atom empirical potential for molecular modeling and dynamics studies of proteins*, *J. Phys. Chem. B* 102 (1998) 3586.
- [10] A. Ben-Naim, *Standard thermodynamics of transfer. Uses and misuses*, *J. Phys. Chem.* 82 (1978) 792.
- [11] R. Elber, M. Karplus, *Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin*, *Science* 235 (1987) 318.
- [12] J.-E. Chea, M. R. Friedel, A. Bumketner, *Simulation of protein folding*, *Rev. Comp. Chem.* 22 (2006) 169.
- [13] W. Kauzmann, *Some factors in the interpretation of protein denaturation*, *Adv. Prot. Chem.* 14 (1959) 1.
- [14] G. I. Makhatadze, P. L. Privalov, *Energetics of protein structure*, *Adv. Prot. Chem.* 47 (1995) 307.
- [15] M. Schaefer, M. Karplus, *A Comprehensive Treatment of Continuum Electrostatics*, *J. Phys. Chem.* 100 (1996) 1578.
- [16] J. A. McCammon, S. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, 1987.
- [17] C. L. Brooks III, M. Karplus, B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure, & Thermodynamics*, *Adv. Chem. Phys.* LXXI, John Wiley & Sons, 1988.
- [18] A. Onufirev, D. Bashford, D. A. Case, *Modification of the Generalized Born Model Suitable for Macromolecules*, *J. Phys. Chem.* 104 (2000) 3712.
- [19] M. Schaefer, C. J. Froemmel, *A Precise Analytical Method for Calculating the Electrostatic Energy of Macromolecules in Aqueous Solution*, *J. Mol. Biol.* 216 (1990) 1045.
- [20] M. Born, *Volumen und Hydratationswärme der Ionen*, *Z. Phys. Chem.* 1 (1920) 45.
- [21] J. G. Krikwood, *Theory of Solutions of Molecules Containing Widely Separated Charges with Special Application to Zwitterions*, *J. Chem. Phys.* 2 (1934) 351.
- [22] G. D. Hawkins, C. J. Cramer, D. G. Truhlar, *Pairwise descreening of solute charges from dielectric medium*, *Chem. Phys. Lett.* 246 (1995) 122.
- [23] L. R. Pratt, D. Chandler, *Theory of hydrophobic effect*, *J. Chem. Phys.* 67 (1977) 3683.
- [24] G. Gompper, M. Hauser, A. A. Kornyshev, *"Confined" water and hydrophobic attraction as a result of metastable coordination, stabilized by hydrophobic surfaces*, *J. Chem. Phys.* 101 (1994) 3378.
- [25] J. D. Weeks, K. Katsov, K. Vollmyr, *Roles of repulsive and attractive forces in determining the structure of nonuniform liquids: Generalized mean field theory*, *Phys. Rev. Lett.* 81 (1998) 4400.
- [26] K. Katsov, J. D. Weeks, *Density fluctuations and the structure of a nonuniform hard sphere fluid*, *Phys. Rev. Lett.* 86 (2001) 440.
- [27] J. D. Weeks, *CONNECTING LOCAL STRUCTURE TO INTERFACE FORMATION: A Molecular Scale van der Waals Theory of Nonuniform Liquids*, *Ann. Rev. Phys. Chem.* 53 (2002) 533.
- [28] K. Katsov, J. D. Weeks, *On the mean field treatment of attractive interactions in nonuniform simple fluids*, *J. Chem. Phys.* 105 (2001) 6738.
- [29] J.P. Hansen, I. R. McDonald, *Theory of simple liquids*, Academic Press, London, 1986.

- [30] E. Weissman, J. L. Lebowitz, D. Henderson, *Solution of the mean spherical approximation for the density profile of a hard-sphere fluid near a wall*, Mol. Phys. 32 (1976) 1373.
- [31] F. van Swol, J. R. Henderson, *Wetting and drying transitions at a fluid-wall interface. Density-functional theory versus computer simulation. II*, Phys. Rev. A 43 (1991) 2932.
- [32] J. K. Johnson, J. A. Zollweg, K. E. Gubbins, *The Lennard-Jones equation of state revisited*, Mol. Phys. 78 (1993) 591.
- [33] K. Lum, D. Chandler, J. D. Weeks, *Hydrophobicity at Small and Large Length Scales*, J. Phys. Chem. B 103 (1999) 4570.
- [34] D. M. Huang, D. Chandler, *The Hydrophobic Effect and the Influence of Solute-Solvent Attractions*, J. Phys. Chem. B 106 (2002) 2047.
- [35] G. Hummer, S. Garde, A. E. García, A. Pohorille, and L. R. Pratt, *An information theory model of hydrophobic interactions*, PNAS 93 (1996) 8951.
- [36] L. R. Pratt, D. Chandler, *The theory of the hydrophobic effect*, J. Chem. Phys. 67 (1977) 3683.
- [37] T. Lazaridis, G. Archontis, M. Karplus, *Enthalpic contribution to protein stability: atom-based calculations and statistical mechanics*, Adv. Prot. Chem. 47 (1995) 231.
- [38] T. Lazaridis, M. Karplus, *Thermodynamics of protein folding: a microscopic view*, Biophys. Chem. 100 (2003) 367.
- [39] M. Baus, R. Lovett, *A direct derivation of the profile equations of Buff-Lovett-Mou-Wertheim from the Born-Green-Yvon equations for a non-uniform equilibrium fluid*, Physica A 181 (1992) 329.