

KONFORMAČNÍ CHOVÁNÍ AMINOKYSELIN V PEPTIDECH A PROTEINECH Z POHLEDU MOLEKULÁRNÍHO MODELOVÁNÍ A VÝPOČETNÍCH METOD

Věnováno Zdeňku Havlasovi k jeho 65. narozeninám.

JIŘÍ VONDRÁŠEK^a a JIŘÍ VYMĚTAL^{a,b}

^a Ústav organické chemie a biochemie AV ČR, v.v.i., Flemingovo náměstí 2, 166 10 Praha 6, ^b Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
jiri.vondrasek@uochb.cas.cz

Došlo 27.1.16, přijato 15.2.16.

Klíčová slova: molekulové modelování, konformační stavy aminokyselin, vzorkování konformačního prostoru, molekulová dynamika, Ramachandranova mapa

Obsah

1. Úvod
2. Chemická struktura aminokyselin a jejich vlastnosti
3. Od aminokyselin k peptidům
4. Energetické stavy peptidu jako funkce torzních parametrů hlavního řetězce
5. Metody výpočetní chemie pro mapování konformačního prostoru aminokyselin a peptidů
 - 5.1. Korekce parametrů empirické funkce pomocí přesných *ab initio* metod
 - 5.2. Přesnost popisu konformačních stavů empirickými metodami
6. Vliv sekvenčního okolí na konformační chování aminokyselin v peptidech
7. Závěr a shrnutí

1. Úvod

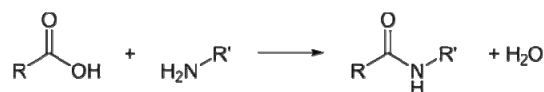
Aminokyseliny jsou základní stavební prvky proteinů a jejich fyzikálně chemické vlastnosti určují chování těchto lineárních polymerů do té míry, že je v některých případech možné z proteinové sekvence předpovídat i jejich strukturu¹. Tradiční dělicí hranice mezi proteiny a peptidy byla založená na přesvědčení, že protein je oproti peptidu charakterizovatelný právě přítomností stabilního prostorového uspořádání zvaného „struktura proteinu“, která je zároveň termodynamickým minimem Gibbsovy volné energie příslušné lineární kombinace aminokyselin. Objev a charakterizace tzv. nestrukturovaných proteinů², anglicky „Intrinsically Disordered Proteins“ ale nastolil otázku, do jaké míry je délka proteinového řetězce a vlastnosti aminokyselin, ze kterých se protein skládá, určující pro jeho prostorové uspořádání a funkci a v čem tedy spočívá

skutečný rozdíl mezi proteiny a peptidy a jejich chováním.

Právě případ nestrukturovaných proteinů, které nemají stabilní prostorové uspořádání³ a tedy jediné termodynamické minimum, které determinuje jejich funkci, poukázal na skutečnost, že protein může existovat v celém strukturálním spektru stavů od plně nestrukturovaných forem, přes částečně strukturované až k rigidním prostorovým útvarům. Nestrukturované proteiny tvoří tedy jeden ze základních typů společně s membránovými, fibrilárními a globulárními proteiny. Experimentální charakterizace dalších nestrukturovaných proteinů vyvolaly diskusi o tradičním paradigmatu, že funkce proteinu je jednoznačně závislá na jeho prostorovém uspořádání. Objevy posledních let v různých odvětvích strukturní biologie ukazují, že to, co hraje hlavní roli ve funkci proteinů, je jejich dynamické chování. Tato dynamika může být jak reflexí komplexního uspořádání proteinu v jeho stabilní prostorové formě, tak odrazem chování relativně nezávislých lineárních úseků proteinové sekvence, jejichž dynamika je blízká chování izolovaných peptidů.

2. Chemická struktura aminokyselin a jejich vlastnosti

Živé systémy v našem světě podléhají tzv. centrálnímu dogmatu molekulární biologie, kdy lineární informace obsažená v DNA je pomocí transkripčního a translačního mechanismu postupně přepisována do molekul RNA a poté do proteinů. Proteinová abeceda je vytvořena 20 odlišnými typy aminokyselin kódovanými genomovou DNA. Všechny tyto aminokyseliny mají společný centrální uhlíkový atom (C_α), ke kterému je vázán vodíkový atom, aminoskupina (NH_2) a karboxyskupina ($COOH$). To, co rozlišuje jednu aminokyselinu od druhé, je tzv. postranní řetězec kovalentně navázaný na C_α atom na poslední valenci jeho sp^3 uspořádání. Aminokyseliny jsou spojovány za sebou do lineárního řetězce v procesu proteinové syntézy utvořením peptidové vazby, kdy karboxylová skupina jedné aminokyseliny reaguje s aminoskupinou té následující a uvolňuje tak jednu molekulu vody. Formování po sobě jdoucích peptidových vazeb vytváří tzv. hlavní řetězec nebo také páteř proteinu, ze které ční jednotlivé postranní řetězce aminokyselin, ze kterých je protein složen (obr. 1).



Obr. 1. Schéma vzniku peptidové vazby spojením dvou aminokyselin

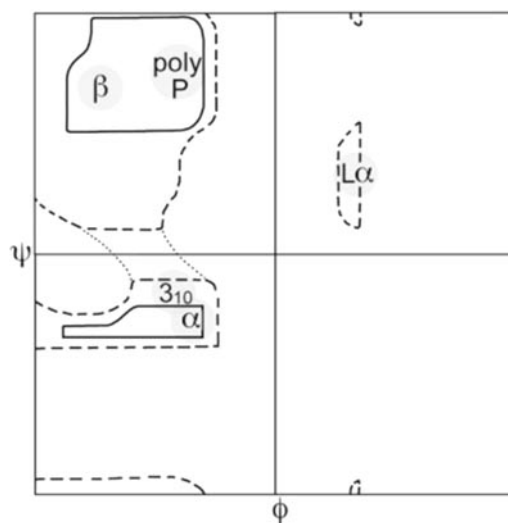
Aminokyseliny je možné rozdělit podle charakteru postranního řetězce do 3 různých kategorií charakterizovaných chemickou povahou jeho funkční skupiny. V první kategorii jsou zařazeny aminokyseliny se striktně hydrofobním postranním řetězcem (Ala, Val, Leu, Ile, Phe, Pro a Met), do druhé kategorie spadají aminokyseliny s nabitým postranním řetězcem (Asp, Glu, Arg, Lys) a konečně třetí kategorie obsahuje aminokyseliny s polárním postranním řetězcem (Ser, Thr, Cys, Asn, Gln, His, Tyr a Trp). Jen pro úplnost je třeba dodat, že všechny proteiny vznikající proteosyntézou se skládají z aminokyselin v jediné chirální formě – L.

3. Od aminokyselin k peptidům

Polypeptidový řetězec se nechá charakterizovat konformačními vlastnostmi, kde základní a klíčovou úlohu hraje peptidová vazba mezi sousedními aminokyselinami. Systém tří vazeb mezi dvěma C_α atomy se nazývá peptidová jednotka a její rigidní charakter omezuje konformační variabilitu peptidového řetězce, takže jediné stupně volnosti, které se váží k páteři polypeptidu, jsou vazby C_α atomu, označované jako ϕ a ψ . Každá peptidová jednotka může tedy rotovat kolem těchto dvou vazeb – vazba $N-C_\alpha$ se nazývá ϕ a vazba $C_\alpha-C'$ z téhož C_α atomu se nazývá ψ (obr. 2).

Tímto způsobem je každá aminokyselina svázána se dvěma torzními úhly ϕ a ψ a protože jsou to jediné stupně volnosti na peptidové páteři, je tato kompletně určena, jestliže jsou udány tyto úhly. Velká většina kombinací úhlů ϕ a ψ není dovolena díky sterickým zábranám mezi postranními řetězci a je proto velmi jednoduché vyčíslit takové kombinace, které dovolené jsou. Kombinace úhlů ϕ a ψ vynesené ve dvoudimenzionálním grafu proti sobě se nazývají Ramachandranův diagram⁴, který je pojmenován po indickém biofyzikovi G. N. Ramachandranovi, který první určil stericky povolené oblasti na tomto grafu (obr. 3).

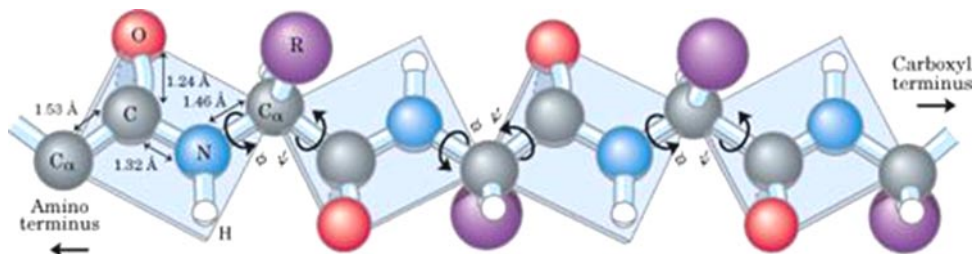
Na schematickém obrázku Ramachandranova diagramu jsou jasně vidět oblasti, které jsou průnikem všech 20 charakteristik základních proteinových aminokyselin. Je nicméně jasné, že právě charakter a struktura postranních řetězců jednotlivých aminokyselin, která je odlišná, se



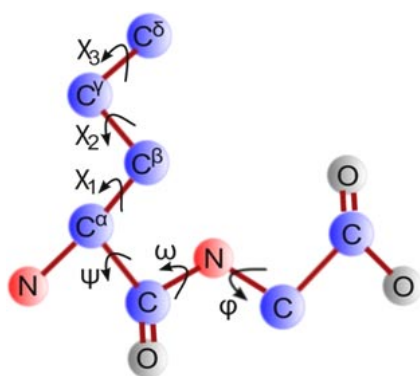
Obr. 3. Ramachandranův diagram

také projeví na intervalu dovolených hodnot úhlů ϕ a ψ . Každý postranní řetězec delší než alanin může v principu rotovat kolem vazeb mezi atomy postranního řetězce. Je obecné pravidlo v chemii, že energeticky nejvýhodnější uspořádání pro dva tetraedrálně koordinované uhlíkové atomy je tzv. „staggered“ konformace. Protože postranní řetězce aminokyselin mohou obsahovat více takových uhlíkových atomů, existuje $n \times 3$ kombinací takových torzních úhlů, které může postranní řetězec aminokyseliny zaujímat. Těmto stavům se říká rotamerní stavy a společně s konformačními stavy hlavního řetězce tvoří úplný geometrický popis polypeptidového řetězce (obr. 4).

Ramachandranův diagram představuje konformačně dovolené stavy hlavního řetězce polypeptidu, které jsou projevem potenciální energie dané aminokyselinou určené sérií vnitřních souřadnic – v našem případě torzních úhlů ϕ a ψ . Protože na energetické hyperploše určené těmito souřadnicemi nevystupují jako další parametry rotamerní torzní úhly, dá se na tuto hyperplochu potenciální energie pohlížet jako na degenerovanou, kde degenerace závisí na počtu stupňů volnosti pro daný postranní řetězec. Dokona-



Obr. 2. Geometrické parametry peptidového řetězce (převzato z wordpress.com/tag/peptide)

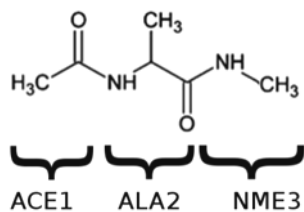


Obr. 4. Torzní úhly definované v aminokyselině

lé zmapování povrchu potenciální energie pro danou aminokyselinu představuje obrovskou výzvu pro metody výpočetní chemie, neboť ne každá metoda je z hlediska zmapování takového povrchu vhodná a metody se liší i přesností, s jakou dané lokální minimum lokalizují.

4. Energetické stavy peptidu jako funkce torzních parametrů hlavního řetězce

V této souvislosti je třeba připomenout základní parametry mapování energetické funkce aminokyselin v závislosti na kontextu, ve kterém se nacházejí. Triviálním modelovým systémem je aminokyselina alanin, která má jak na C-konci, tak na N-konci chránící skupiny, C-konec je chráněn jako *N*-methylamid a N-konec je acetylován. Odpovídající název této molekuly obsahující dvě peptidové vazby je alanindi-peptid (obr. 5).



Obr. 5. Alanindi-peptid jako základní modelová jednotka pro studium konformačního chování

5. Metody výpočetní chemie pro mapování konformačního prostoru aminokyselin a peptidů

Prozkoumání konformačního prostoru takové molekuly je závislé na charakteru a rozvětvenosti postranního řetězce a právě jeho vlastnosti mohou významnou měrou komplikovat mapování konformačního prostoru.

V případě, že postranní řetězec není tvořen jedinou methylovou skupinou, jako je tomu u alaninu, jde o netriviální problém mapování *n*-rozměrné energetické funkce, kde *n* je počet torzních úhlů aminokyseliny. Dalším významným parametrem je metoda, která se pro mapování použije. Tyto metody se dají rozdělit do 3 hlavních skupin.

První skupinu tvoří empirické potenciály – tedy funkce, jejichž parametry jsou získány empiricky z modelových systémů. Energie je v nich vyjádřena jako funkce určitého počtu členů zahrnujících parametry chemické vazby – délky vazeb, úhly mezi nimi a torze a nekovalentní interakce. Tyto metody jsou pro prohledávání konformačního prostoru nejpoužívanější, avšak jejich použití naráží na jeden zásadní problém. Empirické parametry nejsou jednak přesné (nedá se např. zahrnout anharmonicitu) a jednak nejsou obecně přenositelné, což znamená, že pro každou konkrétní úlohu by se měla parametrizace provést znovu. To je samozřejmě nemožné, a proto se interpretace výsledků získaných pomocí empirických potenciálů musí pečlivě analyzovat a srovnat nejlépe s experimentem.

Druhá velká skupina metod jsou metody semiempirické. V rámci řešení stacionární Schrödingerovy rovnice se zavádějí určité aproximace, jako je třeba zahrnutí pouze valenčních elektronů, a poté se řeší zjednodušené rovnice. V těchto rovnicích se kromě toho vyskytují empirické parametry (stejně jako v empirických metodách získané z experimentu), které řešení dále zjednodušují. Pro hledání energetického minima nebo prohledávání konformačního prostoru jsou tyto metody vhodné pouze v případech, kdy je verifikována jejich přesnost.

Přesnost zmíněná v předchozím odstavci je dosažitelná pouze na úrovni tzv. *ab initio* metod. I mezi těmito metodami jsou značné rozdíly, takže se dá obecně říci, že pro spolehlivé mapování konformačního povrchu aminokyselin či peptidů je nutné použít pouze takové metody, které zahrnují elektronovou korelaci a používají dostatečně velké báze funkce (Møller-Pleset – MP, coupled clusters – CCSD(T) či configuration interaction – CI metody). Problém použití těchto metod spočívá v tom, že jsou sice velmi přesné, ale časově nesmírně náročné, takže přesné zmapování energetického povrchu všech aminokyselin je proveditelné pouze za předpokladu dostatečného množství strojového času a není možné pomýšlet na přesné mapování delších peptidů.

Zmíněná omezení diskutovaných metod se nicméně nechají řešit. Poměrně nejintenzivnější je úsilí vylepšit přesnost empirických potenciálů, které se poté používají pro realistické vzorkování konformačního prostoru peptidů a proteinů metodami molekulové dynamiky^{5–7}. Hlavní problém je spojen s nevyvážeností tendencí některých oblastí na Ramachandranově diagramu pro jednotlivé aminokyseliny, což činí spolehlivý popis termodynamických vlastností komplikovaným, pokud už ne rovnou nemožným. Jediný způsob, který umožňuje ukotvení výsledků vzorkování konformačního prostoru, je srovnání s experimentálními daty. Výsledky některých studií publikované v poslední době poskytují dostatečné množství

materiálu pro testování teoretických metod a přesnosti aproximací na dostatečné úrovni spolehlivosti^{8,9}. Konceptně jde o velmi zajímavý problém, protože plocha potenciální energie získaná pomocí pokročilých *ab initio* metod nijak negarantuje správný profil volné energie, již je mnohem složitější vypočítat, a to především díky efektu okolního prostředí. Je třeba také zmínit, že většina metod, jako např. metody založené na funkci elektronové hustoty (density function theory – DFT) a metody semiempirické, neposkytují korektní charakteristiky termodynamických vlastností zkoumaných molekul¹⁰. Na druhé straně jsou to pouze pokročilé metody kvantové chemie, jež garantují, že plocha potenciální energie (potential energy surface – PES) aminokyselin a malých peptidů odráží vliv elektronové korelace, jež má daleko větší vliv na nekovalentní stabilizaci molekul než se původně předpokládalo¹¹.

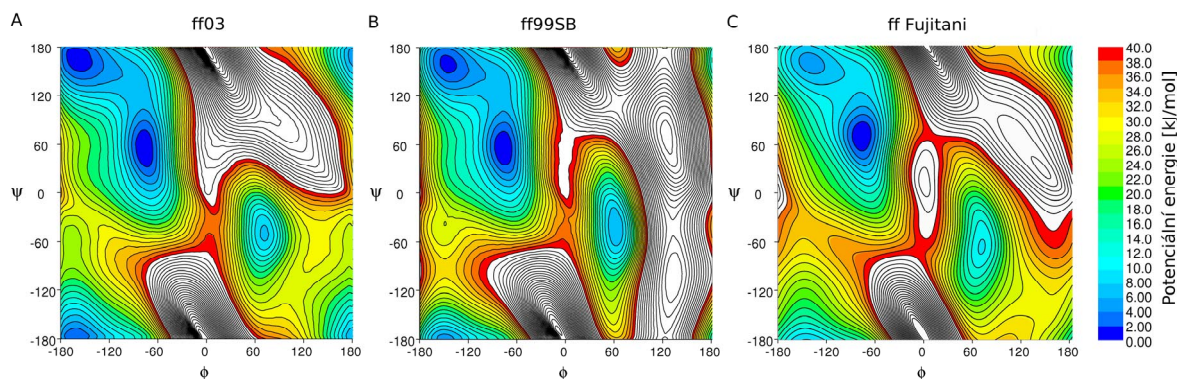
5.1. Korekce parametrů empirické funkce pomocí přesných *ab initio* metod

V poslední době vzniklo množství prací na téma korekce empirických funkcí, které si kladly za cíl především zlepšit parametry pro dihedrální úhly u peptidů a proteinů,

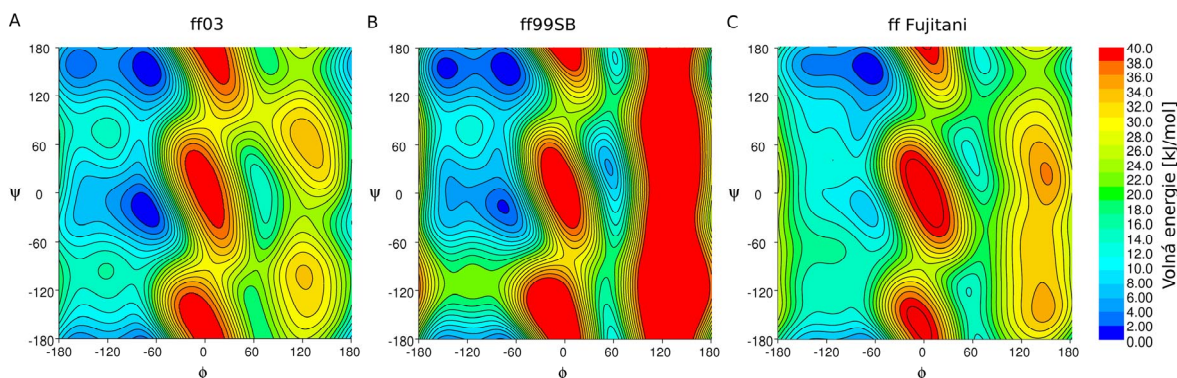
včetně publikování dalších kompletních verzí empirických potenciálů typu amber^{12,13}. První prací publikovanou na nejvyšší úrovni přesnosti byla práce Fujitaniho a spol.¹⁴ v roce 2009. Ve studiích, které jsme prováděli¹⁵ na základě navrženého korekčního schématu, jsme určovali plochu potenciální energie pro molekulu alanindipeptidu jako prvního kroku pro přesný popis termodynamického chování peptidu.

Pro rekonstrukci povrchu Gibbsovy volné energie lze použít metadynamiku, tedy metodu umožňující realistické vzorkování konformačního prostoru. Tato metoda navržená v roce 2002 autory Laio a Parinello¹⁶ byla úspěšně aplikována v různých oblastech. Cílem tedy je ověřit hypotézu, že přesný popis plochy potenciální energie automaticky zaručuje realistické termodynamické chování dané molekuly.

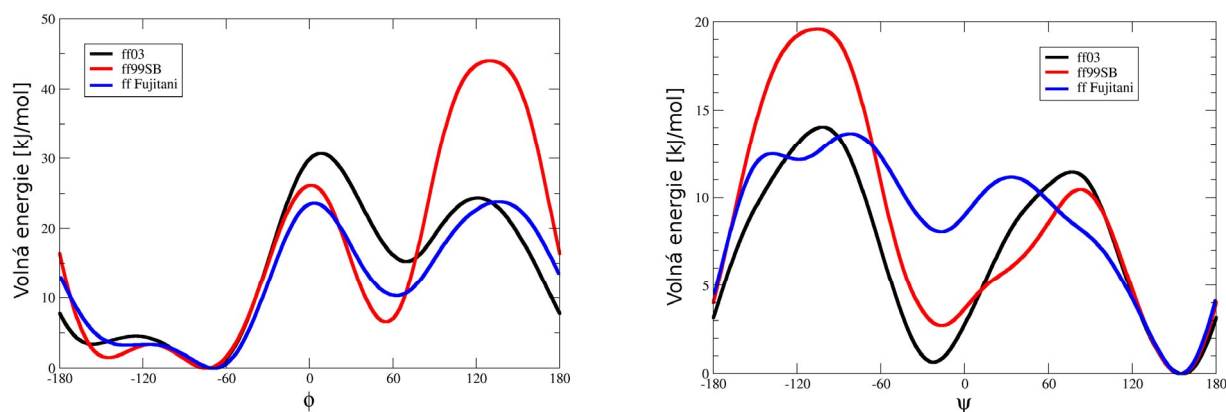
Na uvedeném obrázku (obr. 6) je ukázán rozdíl v popisu plochy potenciální energie alanindipeptidu ve vakuu pomocí dvou empirických potenciálů typu AMBER (ff03 a ff99SB) a potenciálu, ve kterém je již zabudována korekce navržená Fujitanim. Je patrné, že jak poloha, tak i preference pro jednotlivá minima na ploše potenciální energie jsou velmi rozdílné a tyto rozdíly by se měly pro-



Obr. 6. Plochy potenciální energie alanindipeptidu získané pomocí tří různých empirických potenciálů. (Převzato z cit.¹⁵)



Obr. 7. Plocha volné energie alanindipeptidu. (Převzato z cit.¹⁵)



Obr. 8. Srovnání profilů volné energie podél torzních koordinát v závislosti na použitém empirickém potenciálu. (Převzato z cit. ¹⁵)

jevit na profilu volné energie alanindipeptidu ve vodě získané metadynamikou. Jak ukazuje obr. 7, jde skutečně o efekt, který je netriviální a na první pohled patrný z uvedených obrázků. Vyplývá z něj, že kvalita takto získaného povrchu je zásadním způsobem určena přesností popisu PES ve vakuu.

Pokud se pozorně podíváme na energetický profil podél jednotlivých torzních koordinát (obr. 8), je možné konstatovat, že všechny zkoumané empirické potenciály poskytly kvalitativně podobné profily podél parametru ϕ . Na druhé straně je patrné, že ff03 a ff99SB jsou si sice podobné i podél parametru ψ , avšak kvalitativně se liší ve srovnání s potenciálem obsahujícím *ab initio* korekci.

Uvedené výsledky jdou důležité ze dvou hledisek. Prvním je zjištění, že torzní parametr v empirickém potenciálu musí být korigován systematicky a podle výsledků na úrovni pokročilých *ab initio* metod. Druhým důležitým zjištěním je, že takto upravenou plochu potenciální energie lze použít pro realistické vzorkování povrchu volné energie, která je v dobré shodě s experimentálně zjištěnými charakteristikami. V konkrétním příkladu alanindipeptidu je pozoruhodné zjištění, že jako termodynamicky nejstabilnější je tzv. polyprolin II konformace, jež zjevně odráží jasnou závislost na parametru ψ .

Tyto výsledky ukazují, že cesta k realistickému popisu konformačního chování peptidů a proteinů vede přes přesný popis chování vytvářejících aminokyselin v dipeptidech a stanovení korekční funkce pro torze v empirickém potenciálu. Měli bychom být ovšem připraveni i na fakt, že bude třeba prozkoumat chování delších peptidů – např. tripeptidů – jako modelu mapujícího také vliv bezprostředně sousedících aminokyselin. Tato strategie je podle všech dalších zjištění nejen důležitá pro popis chování globulárních proteinů, ale lze ji úspěšně použít i při studiu nestrukturovaných proteinů či krátkých peptidů a jejich dynamického chování.

5.2. Přesnost popisu konformačních stavů empirickými metodami

Z velkého počtu studií konformačního chování peptidů a proteinů vyplývá, že klíčovým prvkem použitelnosti daného popisu je kvalita empirického potenciálu, lépe řečeno kvalita parametrů empirické funkce¹⁷. Právě vyváženost mezi trendy pro helikální či extendované struktury je esenciální vlastností každého empirického potenciálu používaného pro popis proteinových struktur a procesu jejich vytváření nebo denaturace¹⁸. Jak ukázaly nedávne studie na příkladu krátkých peptidů obsahujících alanin, jsou parametry používaných empirických potenciálů výrazně prohelikální⁵. Dalším důkazem nedokonalosti používaných potenciálů bylo selhání při predikci kinetických a termodynamických vlastností vznikajícího helixu v peptidech. Mezi nejpoužívanější empirické potenciály v současnosti patří potenciály AMBER (ff03 a ff99SB) reparametrizované poměrně nedávno na vlastnosti torzí peptidového hlavního řetězce (ff03* a ff99SB*), dále pak sada potenciálů CHARMM, která zmíněné korekce torzí ve svých posledních verzích obsahuje také¹⁹. Přestože korekce torzního členu pro peptidový hlavní řetězec přinesla jisté vylepšení, je zde další problém, který je spojen s parametrizací rotamerních stavů jednotlivých postranních řetězců. Ty jsou ve většině převzaty z modelů malých molekul obsahujících stejné funkční skupiny a nemohou tak zohlednit vliv prostředí peptidu nebo proteinu. Jedinou výjimkou je empirický potenciál typu OPLS-AA/L²⁰, který zmíněné nedostatky překonává použitím reparametrizovaných torzních úhlů získaných z *ab initio* výpočtů již uvedeným způsobem. Je tedy nasnadě, že reparametrizaci jakéhokoli empirického potenciálu by měla předcházet studie zaměřená na jeho kvalitu, a poté reparametrizovat pouze takový potenciál, u kterého výpočet korekčního členu bude nejméně pracný a namáhavý.

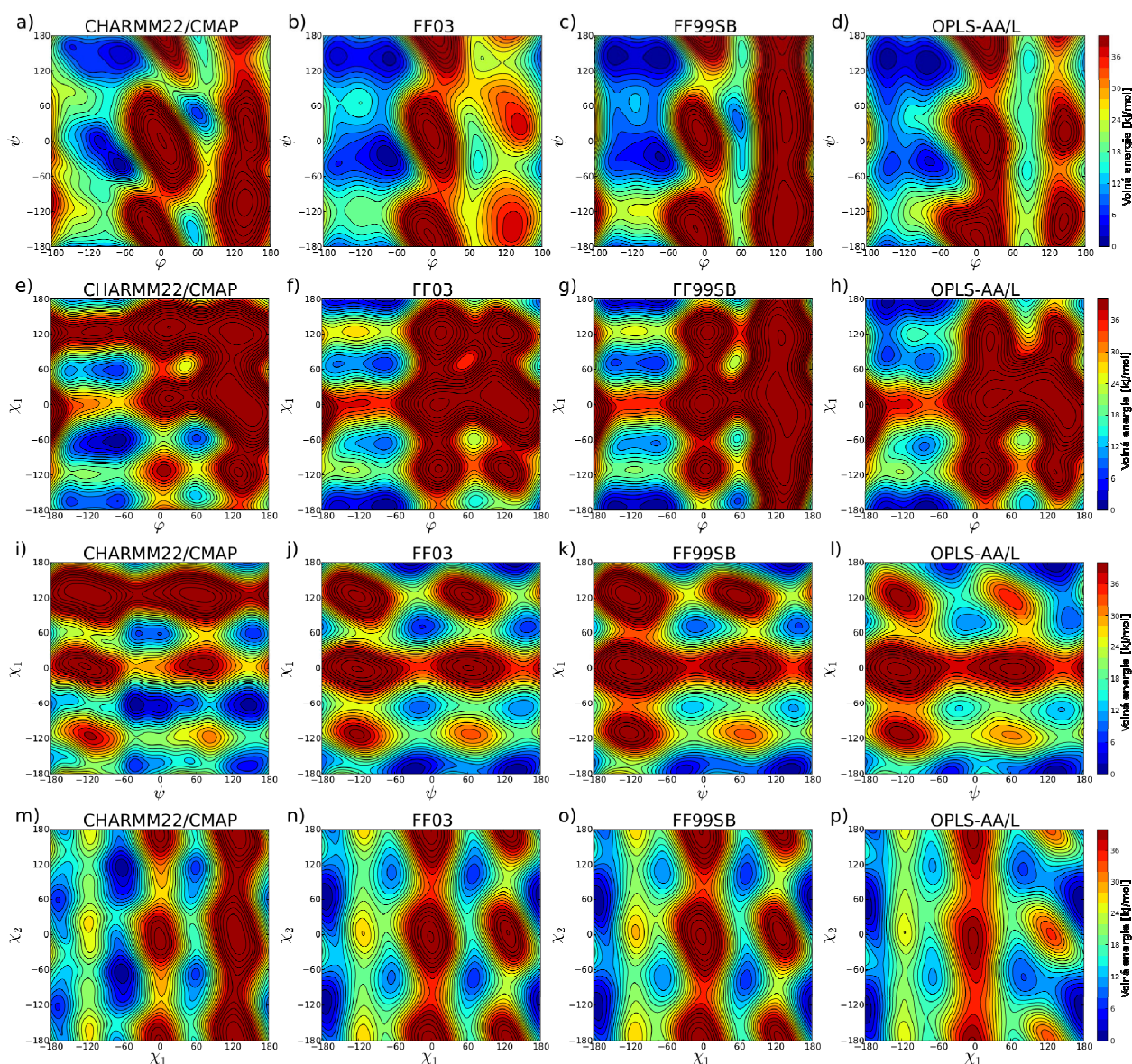
Z výše uvedeného vyplývá, že kritické srovnání empirických potenciálů je nezbytným požadavkem pro realistické mapování konformačního stavu peptidů, jejich struk-

turních preference, jakož i polohy maxim a minim a bariér mezi nimi. Výsledný povrch volné energie je poté nejen spolehlivým ukazatelem strukturních preferencí, ale především poskytuje informace o vzájemném propojení parametrů v torzním prostoru.

V nedávno publikované studii²¹ jsme se detailně zabývali srovnáním 4 vybraných empirických potenciálů a jejich vlastností, které mají vliv na konformační charakteristiky 20 proteinoidních aminokyselin, včetně protonačních stavů aminokyselin nabitých. Aminokyseliny jsme modelovali podobně jako v případě alanindipeptidu – aminokyselinu opatřenou chránicí skupinou na N- a C-konci acetylem a *N*-methylamidem. Už první srovnání preferencí

v torzním prostoru přineslo překvapivé výsledky. Jak vyplývá z uvedených diagramů, vykazují všechny testované empirické potenciály jen velmi málo společných rysů. Týká se to jak preference v torzním prostoru úhlů ϕ a ψ , tak i preference pro jednotlivé torzní úhly charakterizující rotamerní stavy, v konkrétním případě je znázorněna aminokyselina Asp (obr. 9).

Podrobným srovnáním zkoumaných charakteristik všech studovaných aminokyselin jsme dospěli k několika důležitým závěrům. Lze říci, že studované empirické funkce se výrazně liší v predikci energetických minim a jejich umístění jak v torzním prostoru hlavního řetězce, tak v torzním prostoru postranních řetězců. Není jasné, zda



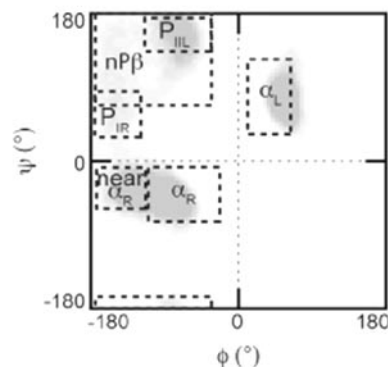
Obr. 9. Profily volné energie pro kyselinu asparátovou. Dvoudimenzionální mapy volné energie byly získány metadynamikou za použití empirických potenciálů – CHARMM22/CMAP, ff03, ff99SB a OPLS-AA/L. (Převzato z cit. ²¹)

jsou tyto rozdíly způsobeny chybnou parametrizací nebo nedostatky výpočetního modelu. Podařilo se nám ale ukázat, že rozdíl mezi empirickými potenciály typu AMBER je způsoben především rozdílným modelem pro určení parciálních nábojů. Domníváme se, že zjištěná nekonzistence studovaných empirických potenciálů může zásadním způsobem ovlivňovat interpretace chování biomolekul, konkrétně procesu proteinového sbalování a denaturace a vlivu prostředí na jejich konformační vlastnosti. Stále zde však zůstává prostor pro optimalizaci, ačkoli lze dopředu počítat s tím, že budeme muset rezignovat na univerzální parametrizace a jejich přenositelnost. Spíše lze očekávat speciální parametrizace pro konkrétní systémy, které budou splňovat podmínky kladené na spolehlivý popis dynamiky peptidů a proteinů.

6. Vliv sekvenčního okolí na konformační chování aminokyselin v peptidech

Jednou z možných cest mapování konformačního chování aminokyselin v peptidech a proteinech je postup, při kterém se bere do úvahy bezprostřední sekvenční a strukturní kontext dané aminokyseliny např. v tri- nebo pentapeptidu. Právě pentapeptid je nejmenší peptidovou jednotkou, ve které může vzniknout stabilní helikální konformace tvořená všemi vytvářejícími aminokyselinami. Jde tedy o mapování aminokyselinových preferencí v kontextu dalších aminokyselin. Je samozřejmě nemožné určit takové chování pro všechny kombinace všech 20 aminokyselin v pentapeptidu, existuje nicméně způsob jak zjistit, zdali tento vliv okolí ovlivňuje i chování a preference dané aminokyseliny. Za tímto účelem se konstruuji tzv. „guest-host“ peptidy, kdy vliv okolí (a především preference pro torzní úhly ϕ a ψ) je modelován glycinem nebo alaninem²². Daná aminokyselina si pak uchovává částečně vlastnosti, které demonstrovala v uspořádání dipeptidu, a částečně je přizpůsobuje danému peptidovému kontextu reprezentovaným především délkou peptidu. Takové studie jsou velmi důležité pro popsání vlivu nekovalentních interakcí v peptidu na jeho konformační vlastnosti a dynamické chování.

V poslední právě publikované studii²³ jsme použili metody molekulové dynamiky na „guest-host“ alaninových pentapeptidech (AAXAA), které sloužily jako model pro nesbalený či denaturovaný stav proteinu. To nám pomohlo při určení skutečných preferencí jednotlivých aminokyselin v tomto modelovém systému. Vytvořili jsme reprezentativní statistické soubory, které byly v dobré shodě s experimentálně určenými hodnotami chemických posunů v NMR, a ukázali, že jejich chování je daleko od předpokládaného náhodného chování. Preference pro konkrétní konformační stav jednotlivých aminokyselin byly nejen vzájemně rozdílné, ale podléhaly také vlivu prostředí (močovina) nebo podmínkám tepelné denaturace proteinů. Simulace molekulovou dynamikou dále umožnily detailně prozkoumat reprezentativní konformační uspořádání daných souborů a vysvětlit tak hodnoty souborově-

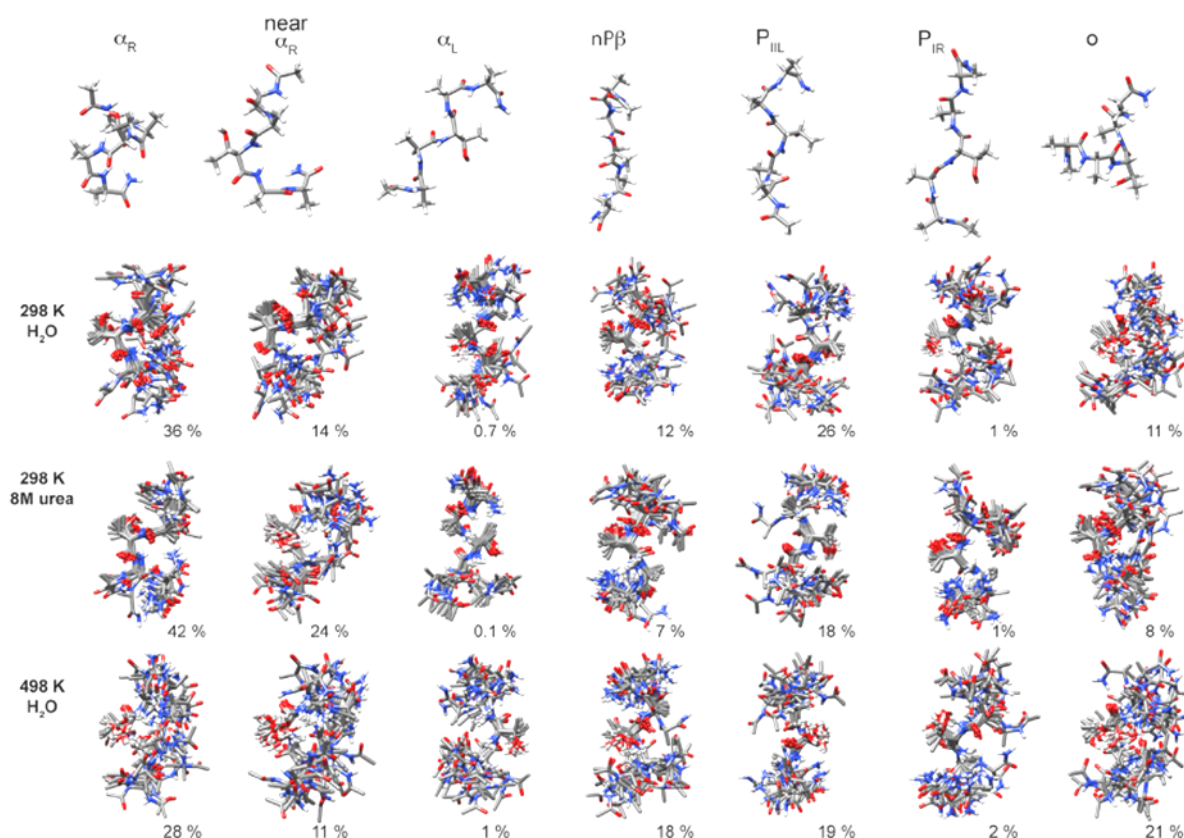


Obr. 10. Oblasti na Ramachandranově diagramu s vyznačenými specifickými konformacemi hlavního řetězce: pravotočivý α -helix (α_R), levotočivý α -helix (α_L), polyproline II levotočivý helix (PIIL), polyproline I pravotočivý helix (PIR) a zbývající β extendovaná oblast (nP β)

zprůměrovaného chemického posunu. Uveřejnění takových dat umožní lepší a realističtější vytváření strukturních modelů a poskytuje základ pro vysvětlení a interpretaci experimentů. Výsledky studie a její závěry lze dobře dokumentovat na příkladu konformačního chování threoninu za normálních podmínek ve vodném prostředí, za zvýšené teploty, která modeluje tepelnou denaturaci (498 K), a konečně v 8M směsi močoviny a vody. Pro interpretaci je nutné vysvětlit, jaké konformační oblasti jsou pro daný aminokyselinový řetězec definovány a jakým způsobem kvantitativně určit pravděpodobnost výskytu v tomto stavu. K tomuto účelu slouží obr. 10, na kterém jsou v jednotlivých kvadrantech Ramachandranova diagramu znázorněny rozdílné konformační oblasti polypeptidového řetězce, které charakterizují typické sekundární strukturní prvky proteinů a peptidů, přechodů mezi nimi a unikátní konformační stavy polypeptidového řetězce.

Konformační chování threoninu v pentapeptidu AATAA za různých podmínek se nechá velmi dobře ilustrovat pomocí následujícího obr. 11.

Na obr. 11 jsou znázorněny struktury pentapeptidu charakterizující rozdílné konformační stavy definované na Ramachandranově diagramu. Příslušné populace za různých podmínek ukazují na velikost konformační variability pro daný stav a dobře ilustrují reprezentativní soubor, který danou konformaci charakterizuje. Výsledky těchto studií potvrdily, že jednotlivé aminokyseliny mají jasné a vzájemně rozdílné preference a tyto preference mohou být modulovány vlastnostmi rozpouštědla nebo teplotou. Ačkoli konformační preference v nativních podmínkách jsou většinou zachovány i v roztoku močoviny, neplatí totéž pro chování a preference aminokyselin při tepelné denaturaci. Nicméně je třeba zdůraznit, že pro většinu reziduí je jejich konformační chování invariantní vzhledem ke změně podmínek, a jde tedy o chování malého počtu aminokyselin, které diktují strukturní chování celého



Obr. 11. Jednotlivé typy konformací pro X aminokyselinu pentapeptidu AAXAA ve vodě, 8M močovíně a ve vodě při teplotě 498 K a jejich příslušné reprezentativní soubory. (Převzato z cit.²³)

peptidu jako reakci na změnu podmínek. To, co je asi nejdůležitějším výsledkem této studie, je zjištění, že výsledky simulací dobře reprodukuje chování aminokyselin v globulárních proteinech, což indikuje, že strukturní preference v proteinech jsou důsledkem strukturních preferencí jednotlivých aminokyselin.

7. Závěr a shrnutí

Konformační chování aminokyselin v proteinech je určujícím parametrem procesu označovaného jako sbalování proteinů. Dlouhou dobu se předpokládalo, že toto chování je důsledkem především vzájemné interakce postranních řetězců a jejich konformační chování diktované hlavním řetězcem je více či méně uniformní (s výjimkou prolinu a glycinu). Jak se však ukazuje v poslední době, je chování aminokyselin silně závislé jak na sekvenčním kontextu peptidu či proteinu, tak i na jeho délce. Chování izolovaných postranních řetězců nelze tedy promítat do jejich chování v proteinech či peptidech. Navíc se ukazuje, že přesný energetický popis konformačního chování aminokyselin je základním požadavkem pro parametrizaci

empirických potenciálů a tento popis není obecně transferabilní, pokud je bezprostřední strukturní a chemické okolí aminokyseliny značně odlišné. Úspěšné pokusy reparametrizaci nejvíce používaných potenciálů pro studium strukturovaných proteinů tak přispěly k přesnější interpretaci nejen jejich dynamického chování a stavů, ale umožňují popis konformačních souborů proteinů, které za normálních podmínek pevnou strukturu nemají.

LITERATURA

1. Dill K. A., Ozkan S. B., Weikl T. R., Chodera J. D., Voelz, V. A.: *Curr. Opin. Struct. Biol.* 17, 342 (2007).
2. Tompa P., Han K. H.: *Physics Today* 65, 64 (2012).
3. Dyson H. J., Wright P. E.: *Nat. Rev. Mol. Cell Biol.* 6, 197 (2005).
4. Ramachandran G. N., Ramakrishnan C., Sasisekharan V.: *J. Mol. Biol.* 7, 95 (1963).
5. Best R. B., Buchete N. V., Hummer G.: *Biophys. J.* 95, L7 (2008).
6. Daggett V.: *Molecular-Dynamics Simulations of Peptides*. Abstr. Pap. Am. Chem. Soc. 209, 90-COMP (1995).

7. Hornak V., Abel R., Okur A., Strockbine B., Roitberg A., Simmerling C.: *Proteins: Struct., Funct., Bioinf.* **65**, 712 (2006).
8. Alonso J. L., Echenique P.: *J. Comput. Chem.* **27**, 238 (2006).
9. Nakano T., Kaminuma T., Sato T., Akiyama Y., Uebayasi M., Kitaura K.: *Chem. Phys. Lett.* **318**, 614 (2000).
10. Valdes H., Pluhackova K., Pitonak M., Rezac J., Hobza P.: *Phys. Chem. Chem. Phys.* **10**, 2747 (2008).
11. Valdes H., Spiwok V., Rezac J., Reha D., Abo-Riziq A. G., de Vries M. S., Hobza P.: *Chem. - Eur. J.* **14**, 4886 (2008).
12. MacKerell A. D.: *J. Comput. Chem.* **25**, 1584 (2004).
13. Perczel A. S., Farkas O., Jákl I., Topol I. A.: *J. Comput. Chem.* **24**, 1026 (2003).
14. Fujitani H., Matsuura A., Sakai S., Sato H., Tanida Y.: *J. Chem. Theor. Comput.* **5**, 1155 (2009).
15. Vymetal J., Vondrasek J.: *Chem. Phys. Lett.* **503**, 301 (2011).
16. Laio A., Parrinello M.: *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12562 (2002).
17. Piana S., Lindorff-Larsen K., Shaw D. E.: *Biophys. J.* **100**, L47 (2011).
18. Lindorff-Larsen K., Maragakis P., Piana S., Eastwood M. P., Dror R. O., Shaw D. E.: *PLoS One*, **7**, e32131 (2012), doi: 10.1371/journal.pone.0032131.
19. Best R. B., Hummer G.: *J. Phys. Chem. B* **113**, 9004 (2009).
20. Kaminski G. A., Friesner R. A., Tirado-Rives J., Jorgensen W. L.: *J. Phys. Chem. B* **105**, 6474 (2001).
21. Vymetal J., Vondrasek J.: *J. Chem. Theory Comput.* **9**, 441 (2013).
22. Beck D. A. C., Alonso D. O. V., Inoyama D., Daggett V.: *Proc. Natl. Acad. Sci. U.S.A.* **105**, 12259 (2008).
23. Towse C. L., Vymetal J., Vondrasek J., Daggett V.: *Biophys. J.* **110**, 348 (2016).

J. Vondrášek^a and J. Vymětal^{a,b} (^a *Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Prague*, ^b *Department of Biochemistry, University of Zurich, Zurich, Switzerland*): **Conformational States of Amino Acids in Proteins and Peptides by Means of Computational Chemistry Methods**

The application of the physical principles governing biomolecules in order to adopt their unique 3D structure implies a need for reasonably accurate methods able to map conformational hyperspace at the Gibbs free-energy level. In the case of proteins, the question is rather complex and includes the issue of at which level the local conformational preferences of amino acids in short sequential context are propagated into the final protein structure. The role of amino acid side chains is not quite clear in the structural context because the backbone preferences define the number of possibilities available for the side chains to interact. This could crucially constrain the spatial possibility of finding the near-energy optimum geometry for each interacting side chain.

It is well-known that the precise balance between propensities for helical and extended structures is an essential property of any force field intended for simultaneous description of folding α -helical, β -sheet, or mixed protein architectures, as well as their unfolded states. As was documented in the not so distant past, the simulations of short alanine-based peptides have revealed overrated helical tendencies of current force fields.

In this paper we discussed the issue from three different perspectives. First, how to improve empirical force field parameters by corrections obtained from high level *ab initio* calculations. Second, how these corrections can be implemented into a force field in a relatively simple way, and third, how the sampling of conformational space by metadynamics reflects the corrections towards a better agreement with experimental results.