# *AB INITIO* VÝPOČTY CHEMICKÝCH POSUNŮ V BIOMOLEKULÁCH

**JIŘÍ CZERNEK A VLADIMÍR SKLENÁŘ***

*Laboratoř struktury a dynamiky biomolekul, Přírodovědecká fakulta, Masarykova univerzita, Kotlářská 2, 611 37 Brno, e-mail: sklenar@chemi.muni.cz, czernek@chemi.muni.cz*

Došlo dne 9.II.1999

Klíčová slova: výpočty *ab initio*, chemické posuny, biomolekuly

## Obsah

1. Úvod
2. Teoretická část
   2.1. Chemický posun
   2.2. Přehled výpočetních metod
3. Studované problémy
   3.1. Výpočty pro peptidy a báze nukleových kyselin
   3.2. Konformační závislost chemických posunů
   3.3. Vliv vodíkových vazeb
4. Výhled do budoucna

## 1. Úvod

Spektroskopie nukleární magnetické rezonance (NMR) patří již řadu let k hlavním experimentálním nástrojům studia struktury biomolekul a dynamiky procesů v nich probíhajících. Jednou z nejdůležitějších informací, kterou NMR spektrum obsahuje, jsou chemické posuny. Tuto informaci je však možné pro strukturní studie molekul využít jen ve speciálních případech. Uvedená skutečnost je důsledkem toho, že chemické posuny, jak uvidíme v následující kapitole, přímo souvisí s elektronovou strukturou, a jsou tedy ovlivněny ději, jež způsobují změnu distribuce elektronů v molekule. Mezi nejdůležitější z nich patří interní flexibilita (experimentálně zjištěný posun je průměrem přes hodnoty příslušné geometrickým uspořádáním nabytým během trvání experimentu NMR) a intermolekulární interakce (blíže v kapitolách 3.2. a 3.3.). V obou případech není závislost posunu na těchto procesech obecně popsána.

Atraktivní možností, jak výše uvedené vlivy studovat, je provedení výpočtu *ab initio*, tj. neempirického výpočtu, který využívá pouze kvantovou mechaniku a základní fyzikální konstanty. V 90. letech došlo v oblasti výpočtu chemických posunů k prudkému nárůstu počtu publikací jak metodického (přehledný článek[1]), tak aplikačního rázu (viz např. ročenky[2]). Díky rozvoji kvantově chemické teorie jsou současné přístupy *ab initio* schopny poskytnout údaje o chemickém stínění v malých molekulách s poměrně velkou přesností. Rozvoj moderních počítačových technologií v 90. letech zpřístupnil širokému okruhu zájemců pracovní stanice a víceprocesorové superpočítače s výkonem řádu Gflop. Tato technologie se tak stala běžnou součástí vybavení akademických i průmyslových laboratoří. Tyto příznivé okolnosti vedou k rostoucímu zájmu o studium chemických posunů, případně faktorů, jež je ovlivňují, nejen u malých organických molekul, ale i systémů o větší molekulové hmotnosti a fragmentů biomakromolekul jako jsou proteiny a nukleové kyseliny. Jako příklad lze uvést vyčíslení posunů $^{13}C$ v molekule taxolu ($C_{47}H_{51}NO_{14}$, použito 1037 bázových funkcí) se střední kvadratickou odchylkou[3] 6,4 ppm a popis vlivu sekundární a terciární struktury na chemické posuny izotopů $^{1}H$, $^{13}C$, $^{15}N$ a $^{19}F$ v bílkovinách[4,5].

Tento referát shrnuje základní metody výpočtů *ab initio* chemických posunů. Na vybraných příkladech jsou demonstrovány možnosti aplikací těchto výpočtů a diskutovány některé problémy spojené s jejich realizací. Problematika zvoleného tématu je mezioborová. Schopnost nejen přesného vyčíslení absolutního stínění malých molekul, ale reprodukce rozdílů posunů v biologicky zajímavých, komplikovaných systémech je velkou výzvou teoretickým chemikům. Detailní znalost vztahu mezi strukturou a posunem či vlivem mezimolekulárních interakcí na chemické stínění může v mnoha případech usnadnit interpretaci experimentálních NMR dat.

## 2. Teoretická část

### 2.1. Chemický posun

Jádra v molekule jsou obklopeny elektrony, jejichž tok ve vnějším magnetickém poli $\vec{B}$ indukuje vznik lokálního pole $\vec{B}^{lok}$. Toto pole způsobuje, že dvě NMR aktivní jádra s odlišným chemickým okolím rezonují při různých frekvencích. Indukci na daném jádře lze vyjádřit vztahem

$$\vec{B}_{\alpha}^{lok} = (1 - \sigma_{\alpha\beta})\, \vec{B}_{\beta} \qquad (1)$$

kde $\vec{B}_{\alpha}^{lok}$ je složka magnetického pole ve směru $\alpha$ indukovaná tokem elektronů, který je důsledkem působení vnějšího magnetického pole ve směru $\beta$. Chemické stínění je vzhledem k asymetrii rozložení elektronů obecně tenzorovou veličinou charakterizovanou tenzorem druhého řádu $\sigma_{\alpha\beta}$ s devíti elementy. Vhodnou volbou souřadnicové soustavy je možné tento tenzor převést do tvaru, v němž diagonální členy $\sigma_{xx}$, $\sigma_{yy}$, $\sigma_{zz}$ jsou jeho hlavními složkami. Jejich průměr

$$\sigma^{izo} = 1/3\, (\sigma_{xx} + \sigma_{yy} + \sigma_{zz}) \qquad (2)$$

potom reprezentuje izotropní stínění daného jádra. Zatímco v pevné fázi je v případě studia jednotlivých krystalů možné anizotropní tenzor chemického stínění plně charakterizovat,

* Korespondující autor

v plynné a kapalné fázi dochází následkem rychlého a neuspořádaného pohybu k zprůměrňování jeho jednotlivých složek. Za předpokladu, že všechny orientace molekuly vzhledem ke směru vnějšího magnetického pole jsou stejně pravděpodobné, je měřitelná pouze veličina $\sigma^{izo}$. Izotropní stínění se charakterizuje nejběžněji parametrem $\delta$, nazývaným chemický posun, který je vyjádřen rozdílem rezonanční frekvence měřeného jádra $\nu = \gamma.(1 - \sigma^{izo})$. $B_0$ a frekvence zvoleného standardu $\nu_{st} = \gamma.(1 - \sigma_{st}^{izo})$. $B_0$ normalizovaným vzhledem k indukci vnějšího magnetického pole

$$\delta = (\nu - \nu_{st}) / \nu_{st} = (\sigma_{st}^{izo} - \sigma^{izo}) / (1 - \sigma_{st}^{izo}) \qquad (3)$$

Při teoretickém výpočtu chemického posunu postupujeme tak, že nejprve některou z metod popsaných dále určíme $\sigma^{izo}$ vybraného jádra. Poté tuto hodnotu odečteme od izotropního stínění standardu získaného buď teoreticky nebo experimentálně; např. stínící konstanta uhlíku $^{13}C$ v jednom z nejběžnějších standardů, tetramethylsilanu, (TMS) $\sigma^{izo}$ ($^{13}C$) je 186,4 ppm (cit.[6]).Takto získaný chemický posun můžeme pak porovnávat s naměřenou hodnotou.

Základy kvantově mechanické teorie chemického stínění pochází ze začátku 50. let (cit.[7,8]). Stručně a srozumitelně jsou popsány ve sborníku[9]. Jelikož podstata stínění je elektronická, vychází se při výpočtu pro atomové jádro s magnetickým momentem $\vec{\mu}$ ve vnějším magnetickém poli $\vec{B}$ z celkového hamiltoniánu $\hat{H}(\vec{\mu}, \vec{B})$. Energie jádra $E(\vec{\mu}, \vec{B})$ závisí na indukci vnějšího magnetického pole $\vec{B}$ a velikosti $\vec{\mu}$. S použitím Rayleighovy-Schrödingerovy poruchové teorie lze obdržet rozvoj energie $E(\vec{\mu}, \vec{B})$ podle $\vec{\mu}$ a $\vec{B}$ a posléze seřazením členů úměrných současně $\vec{\mu}$ i $\vec{B}$ separovat dva příspěvky ke stínění, a to diamagnetickou a paramagnetickou složku:

$$\sigma_{\alpha\beta} = \left(\sigma_{\alpha\beta}^{dia} + \sigma_{\alpha\beta}^{para}\right) \qquad (4)$$

Diamagnetická část tenzoru chemického stínění popisuje schopnost vnějšího magnetického pole způsobovat proudění náboje v základním elektronickém stavu systému. Její směr je obecně opačný než směr vnějšího magnetického pole, proto zvyšuje stínění daného jádra a tím způsobuje posun jeho rezonančního signálu k menším hodnotám $\delta$. Hodnotu diamagnetické složky lze vypočítat na základě znalosti vlnové funkce základního stavu molekuly $|0\rangle$:

$$\sigma^{dia} \propto \left\langle 0 \left| \frac{\boldsymbol{r}_j \boldsymbol{r}_p}{\boldsymbol{r}_j^3} \right| 0 \right\rangle \qquad (5)$$

kde $\boldsymbol{r}_j$ udává polohu elektronů vzhledem ke studovanému jádru a $\boldsymbol{r}_p$ vzhledem k počátku vektorového potenciálu vnějšího magnetického pole (viz dále). Jelikož výsledek je málo citlivý ke kvalitě použité báze výpočtu *ab initio* a rychle konverguje při jejím rozšiřování, je jeho získání pro kvantovou chemii relativně snadným úkolem. Vážnější problém představuje výpočet paramagnetické složky. Paramagnetické stínění lze interpretovat jako schopnost vnějšího magnetického pole vyvolávat tok elektronů molekulou, a to využitím orbitalů, jež nejsou obsazeny v základním stavu. $\sigma^{para}$ působí ve směru blízkém k orientaci $\vec{B}$ a je zodpovědné za zvyšování chemického posunu. Hodnota $\sigma^{para}$ je dána součtem

$$\sigma^{para} \propto -\sum_q \frac{1}{E_q - E_0} \left( \langle \boldsymbol{q} | \frac{\boldsymbol{l}_j}{\boldsymbol{r}_j^3} | 0 \rangle \langle 0 | \boldsymbol{l}_p | \boldsymbol{q} \rangle \right) \qquad (6)$$

přes excitované stavy q systému; $E_q$ a $|\boldsymbol{q}\rangle$ jsou vlastní energie resp. vlnové funkce jednotlivých stavů a $\boldsymbol{l}$ je operátor úhlového momentu. Při výpočtu se obvykle postupuje tak, že excitované stavy jsou popsány virtuálními orbitaly přibližné vlnové funkce. Tato povětšinou drastická aproximace (přesnou hodnotu $\sigma^{para}$ je možné obdržet pouze použitím kompletní báze[10]) vede k podstatným chybám. Je užitečné si všimnout, že velikost paramagnetického členu je nepřímo úměrná rozdílu energií základního a příslušného excitovaného stavu. V orbitalové formulaci je tento jev vyjádřen energetickými rozdíly mezi obsazenými a neobsazenými orbitaly. Obecně, pokud molekula má nízko ležící excitovaný stav (resp. malý rozdíl nejvyššího obsazeného a nejnižšího neobsazeného orbitalu), paramagnetické stínění příslušných atomů bude velké.

Vnější magnetické pole $\vec{B}$ vstupuje do výpočtu *ab initio* chemického stínění prostřednictvím vektorového potenciálu $\vec{A}$ tohoto pole. Ve vztazích (*5*) a (*6*) se objevují členy závislé na počátku vektorového potenciálu. Polohu tohoto počátku lze umístit do libovolného bodu. Ačkoliv v případě přesného výpočtu by provedená volba vliv na výsledek mít neměla, hodnotu přibližného kvantově chemického řešení výrazně ovlivňuje. Při posunu by součet $\sigma^{dia}$ a $\sigma^{para}$ měl zůstat konstantní. V praxi však je chyba vzniklá při výpočtu $\sigma^{para}$ větší než v případě $\sigma^{dia}$ a pro různě zvolené počátky $\vec{A}$ tak dostáváme rozdílné hodnoty chemického stínění vybraného atomu studované molekuly. Tento artefakt se anglicky označuje jako gauge [geidž] problem. Podle způsobu, jak se k tomuto problému přistupuje, lze kvantově chemické metodiky výpočtu chemického stínění rozdělit do několika skupin.

## 2.2. Přehled výpočetních metod

První přístupy[11,12] vycházející z celkového hamiltoniánu $\hat{H}(\vec{\mu}, \vec{B})$ a označované jako Common Origin Approach (COA) neřešily gauge problem. Složky tenzoru studovaného jádra se počítaly vzhledem k počátku na tomto jádře. Tento přístup vyžaduje pro dosažení shody s experimentem použití obrovské báze. Je možné dokázat, že s kompletní bází poskytuje COA přesné výsledky[13,14]. Prakticky je tato metoda však použitelná pouze pro dvouatomové molekuly. Pro studium větších systému bylo vypracováno několik metod, které nejsou závislé na volbě počátku souřadnic vektorového potenciálu. V souvislosti s výpočty biomolekul je třeba zmínit GIAO[15,16] (Gauge Including Atomic Orbitals) a IGLO[17,18] (Independent Gauge for Molecular Orbitals) (na nasazení v této oblasti zatím čekají LORG[19,20] a CSGT[21,22]). Jejich podstatou je explicitní zahrnutí vektorového potenciálu do atomových (v případě GIAO) resp. lokalizovaných molekulových (IGLO) orbitalů. Z praktického hlediska je důležité, že metoda GIAO, jež byla pro výpočty molekul s biologickou relevancí použita ve většině případů, při rozšiřování báze konverguje k přesnému řešení rychleji. Získané výsledky jsou tedy s využitím srovnatelné báze většinou lepší než při aplikaci přístupu IGLO.

Metodiky pro výpočet chemického stínění lze rovněž rozdělit podle způsobu, kterým počítají jednotlivé složky tenzoru $\sigma_{\alpha\beta}$. Většinou jsou založeny na konečné poruchové teorii. Ta při popisu chemického stínění vychází z definičního vztahu

$$\sigma_{\alpha\beta} = \left. \frac{\partial^2 E\left(\vec{\mu}, \vec{B}\right)}{\partial \vec{\mu}_a \partial \vec{B}_\beta} \right|_{\vec{\mu}=0, \vec{B}=0} \qquad (8)$$

Prakticky lze $\sigma_{\alpha\beta}$ vypočítat ze vztahu

$$\sigma_{\alpha\beta} = \boldsymbol{h}^{11}_{\alpha\beta} P^0 + \boldsymbol{h}^{01}_{\alpha\beta} P^1 \qquad (9)$$

Tvary operátorů $\boldsymbol{h}^{11}_{\alpha\beta}$ a $\boldsymbol{h}^{01}_{\alpha}$ jsou uvedeny v již citovaném sborníku[9]. $\boldsymbol{P}^0$ je matice hustoty neporušeného systému, $\boldsymbol{P}^1$ matice hustoty prvního řádu popisující míchání obsazených a neobsazených orbitalů v přítomnosti vnějšího magnetického pole. Výpočet $\boldsymbol{P}^0$ je rutinní záležitostí, $\boldsymbol{P}^1$ se získává řešením spřažených Hartreeho-Fockových (HF) rovnic s magnetickým polem jako poruchou (Coupled Perturbed Hartree-Fock Equations). Základním kvantově chemickým přístupem pro nalezení vlnové funkce systému a jeho energetických hladin je Hartreeho-Fockova metoda selfkonzistentního pole (SCF), tj. použití jednodeterminantového přiblížení a hledání efektivního pole pro odpuzování elektronů. Konečná poruchová teorie chemického stínění ve spojení s takto nalezenou vlnovou funkcí a některou z metod nezávislých na volbě počátku vektorového potenciálu (např. GIAO) tvoří příslušný přístup CHF (např. CHF-GIAO). Jím získané hodnoty jsou v literatuře často označovány jen jako data HF resp. SCF.

Elektronová korelace, tj. rozdíl repulse elektronů vyjádřené pomocí efektivního pole a vypočítané přesně, má v některých případech zásadní vliv na popis stínění. Pro některé molekuly (včetně malých, např. dusík) ani použití rozsáhlé báze v rámci žádného z přístupů CHF neposkytne posuny blízké experimentálním hodnotám. Pro výpočty chemického stínění fragmentů biomolekul byly použity dvě metodiky zahrnující elektronovou korelaci (korelované). K přesnému popisu menších systémů je možno využít Møllerovu-Plessetovu (MP) poruchovou teorii. Její podstatou je, že rozdíl mezi výpočtem se zahrnutím přesné elektronové repulse a s využitím efektivního pole v rovnicích HF se vyjadřuje formou poruchy různých řádů. Nejznámější je poruchová teorie druhého řádu (MP2), která je důležitá např. při výpočtu přesných molekulárních geometrií. Výpočet stínění metodou MP používá matic hustoty a postupu naznačeného rovnicí (*9*). Výpočetní vyjádření odpovídajících matic hustoty bylo však nalezeno teprve v 90. letech[23,24].

Další korelované metody jsou založeny na teorii funkcionálu hustoty[25] (Density Functional Theory, DFT). Tato teorie popisuje přímo elektronovou hustotou, jež automaticky určuje počet elektronů a ostatní elektronické charakteristiky – tedy i chemické stínění. Formálně je podobná teorii SCF. Orbitaly a jím odpovídající energie hledá pomocí rovnic Kohnových-Shamových (KS), jež jsou obdobou rovnic HF; pro systém s poruchou používá rovnice CPKS (Coupled Perturbed Kohn Sham Equations). Z několika implementací DFT pro výpočet stínění se jeví jako nejatraktivnější SOS-DFPT[26,27] (Sum-Over-States Density Functional Perturbation Theory) ve spojení s metodikou IGLO. Po nalezení orbitalů KS neřeší rovnice s poruchou jako alternativní přiblížení UDFT[3] (Uncoupled DFT, nespřažená teorie funckionálu hustoty pro popis chemického stínění). Tím je metodika SOS-DFPT-IGLO tzv. proces $N^3$. Toto označení zjednodušeně znamená, že výpočetní čas roste s třetí mocninou počtu bázových funkcí. UDFT a CHF mají závislost $N^4$ a poruchová teorie MP2 je dokonce proces $N^5$. Tato skutečnost ve spojení s přesností danou zahrnutím elektronové korelace činí použití SOS-DFPT-IGLO pro biomolekuly velmi přitažlivým (viz např.[28,29]).

## 3. Studované problémy

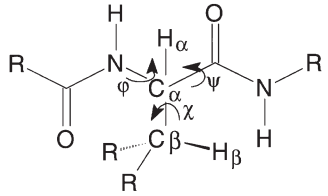
### 3.1. Výpočty pro peptidy a báze nukleových kyselin

Následující část je věnována využití výše uvedené teorie při studiu vybraných stavebních jednotek biopolymerů, a to bází nukleových kyselin, aminokyselin a modelů peptidů a bílkovin. Místo detailního citování konkrétních výsledků (podrobnosti viz uvedená literatura) hledá text spíše odpovědi na následující otázky: Jaké jsou velikosti změn chemických posunů různých atomů v závislosti na změnách struktury, geometrie a okolí studovaných molekul; které metody je třeba použít pro postihnutí těchto změn (jsou přístupy SCF dostatečné nebo zahrnutí korelační energie poskytne jiný obraz?) a konečně, jakou geometrii je vhodné použít pro daný výpočet *ab initio*?

Systematické studium stínění atomů bází nukleových kyselin bylo cílem pouze jediné studie[30]. Jedná se o výpočty přístupem IGLO bez zahrnutí korelační energie a s použitím relativně malých (kvality přibližně triple zeta) bází. Obě tyto okolnosti přispěly k odchylkám od experimentu velikosti až 13 ppm pro uhlíkové a 30 ppm pro dusíkové posuny. V současné době lze metodami DFT ve větších bázích dosáhnout podstatně přesnějších výsledků[31]. Detailně byly prostudovány dusíky uracilu[32]; o této práci se zmíníme při diskusi vlivu vodíkových vazeb na posun v další kapitole.

Odlišná situace panuje v oblasti složek bílkovin – aminokyselin resp. peptidů. Dvě významné empirické studie se zabývaly statistickými korelacemi mezi posuny uhlíků páteře bílkoviny (obr. 1) a její sekundární strukturou[33,34]. Je zřejmé, že pochopení a schopnost předpovědi těchto závislostí by usnadnila řešení struktur bílkovin. Proto se také několik nezávisle pracujících skupin autorů snažilo o rozřešení otázek uvedených na začátku této kapitoly[26,35-47].

Základním údajem, jenž do výpočtu *ab initio* vstupuje, jsou vnitřní souřadnice dané molekuly. Experimentální údaje, pro větší molekuly získávané zpravidla z dat rentgenové strukturní nebo NMR analýzy, jsou ovlivněny konkrétním uspořádáním v krystalu nebo interakcemi v roztoku (více v kapitole 3.3.). Jejich přesnost není mnohdy dostatečná a pro modelové systémy jsou tato data povětšinou nedostupná. Proto se obvyk-




Obr. 1. **Schématické znázornění vazebné struktury peptidického řetězce**. Dihedrální úhly páteře φ, resp. ψ jsou definovány jako $C^O$–N–$C_\alpha$–$C^O$ resp. N–$C_\alpha$–$C^O$–N, χ označuje dihedrální úhel bočního řetězce $H_\beta$–$C_\beta$–$C_\alpha$–$H_\alpha$

le provádí optimalizace geometrie. Tento výpočet představuje samostatný problém se specifickými požadavky na volbu metody a použití vhodné báze. Kvalita geometrie je velmi důležitá pro získání správných hodnot absolutního stínění atomů aminokyselin. Přístupem CHF-GIAO s kvalitní bází byly vypočteny[35] závislosti stínění na vazebných délkách $r$ a vazebných úhlech $\vartheta$. Nalezené hodnoty jsou řádu $10^1$ ppm.$A^{-1}$ pro $d\sigma/dr$ a $10^{-1}$ ppm.$(1^o)^{-1}$ pro $d\sigma/d\vartheta$. Je známo, že směrnice závislosti $ds/dr$ je záporná, protože s rostoucí vazebnou délkou klesají příslušné excitační energie. Výsledkem je zvětšení paramagnetického členu, který je záporný. Z tohoto důvodu způsobuje podhodnocení vazebné délky nadhodnocení stínění. Porovnáme-li výsledky získané metodami HF a MP2 s využitím střední báze pro optimalizaci geometrie, jsou vazebné délky z HF výpočtů kratší než při použití metody MP2. Ukazuje se však, že nadhodnocení stínění v důsledku použití méně správné HF geometrie je systematické, tzn. že rozdíly posunů mezi jednotlivými atomy zůstávají zachovány[38,43]. Při sledování vlivu dihedrálního úhlu na stínění lze příspěvek změn vazebných délek a úhlů jako výsledek geometrické optimalizace zanedbat, poněvadž pro aminokyseliny a modely peptidů jsou změny stínění z více než 90 % určovány příslušným dihedrálním úhlem[38].

Obecně platí, že HF metody nadhodnocují energetické rozdíly mezi obsazenými a virtuálními orbitaly, což vede (vztah (*6*)) k podhodnocení paramagnetického členu a ve svém důsledku ke zvýšení chemického stínění ve srovnání s výsledky korelovaných metod. Orbitaly $\pi$ peptidické skupiny mají tendenci ležet vysoko v souboru obsazených orbitalů. Současně jak nevazebné orbitaly typu n volných elektronových párů karbonylového kyslíku a amidického dusíku, tak orbitaly $\pi^*$ patří mezi nejníže ležící neobsazené orbitaly. Proto se použití přístupu SCF na absolutním stínění tohoto fragmentu dramaticky projeví (např. CHF-GIAO nadhodnocuje stínění karbonylového uhlíku o cca 20 ppm, ale $C_\alpha$ pouze do 1 ppm, ve srovnání s MP2-GIAO za použití stejné báze pro tutéž geometrii modelového peptidu[43]). Pro výzkum biomolekul je zásadní, zda teoretické metody jsou s to reprodukovat trendy v chemických posunech při strukturních změnách. V tomto ohledu je situace příznivá – rozdíly stínění mezi jednotlivými konformacemi lze adekvátně popsat již na úrovni SCF; pro správnější určení absolutních hodnot posunů při různých hodnotách dihedrálního úhlu byl navržen zajímavý postup[43]: Pro jeden konformer se vypočítá stínění přístupem MP2 a rozdíl $\delta$(MP2) – $\delta$(HF) pro tento úhel je přičten k hodnotám získaným pro všechny studované dihedrální úhly metodou HF.

### 3.2. Konformační závislost chemických posunů

Experimentálně stanovený chemický posun je průměrem přes všechna prostorová uspořádání molekul během doby snímání NMR signálu. Z teorie vyplývá, že čas, po který je molekula v určité konfiguraci, lze určit na základě znalosti energetické hyperplochy. Hodnota posunu v každé nabyté konfiguraci je dána bodem na hyperploše stínění, která vyjadřuje závislost chemického posunu na uvažovaných vnitřních souřadnicích daného systému. Jak již bylo uvedeno, v případě bílkovin jsou známy empirické korelace mezi posuny a sekundární resp. terciární strukturou. Mezi nejdůležitější patří vyšší posuny $^{13}C$ na $C_\alpha$ a $C_\beta$ v $\alpha$-šroubovici než v $\beta$-hřebenu a naopak zvýšení posunů $^1H$ $\alpha$-vodíků při přechodu z $\alpha$-šroubovice do $\beta$-hřebenu. Je proto na místě pokusit se o sestrojení hyperploch stínění bílkovin na základě výpočtů *ab initio*. Ukazuje se, že pro reprodukci hlavních experimentálních trendů stačí uvažovat závislost posunu na několika málo dihedrálních úhlech modelových sloučenin. Konkrétně jde o úhly $\phi$ a $\psi$ při popisu páteře biomolekuly a úhel $\chi$ pro určování příspěvků bočních řetězců. Jako modely byly použity systémy od glycinu až po heptapeptid HCO – $(Gly)_3$ – Ala – $(Gly)_3$ – $NH_2$ (cit.[38]). Postup byl ověřován tak, že pro vybranou bílkovinu (v klíčové práci[36] byla zkoumána nukleasa *Staphyloccoca*), jejíž krystalová struktura (a tedy i úhly $\phi$, $\psi$, $\chi$ jednotlivých reziduí) i NMR spektrum jsou známy, se odečítaly $\delta$ z vypočtené hyperplochy stínění a porovnávaly s experimentálními údaji pro odpovídající hodnoty torzí. Shoda posunů $^{13}C$ vypočtených s naměřenými je dobrá (např. pro $C_\alpha$ dvanácti alaninových reziduí byla směrnice závislosti experimentálních hodnot na vypočtených 0,85 s regresním koeficientem 0,94). Shodu teorie s experimentem je možné ještě zlepšit zahrnutím elektrostatických interakcí a vodíkových můstků (viz dále). Diskusi dosud publikovaných dat pro jednotlivé atomy několika aminokyselin lze nalézt v článku[35].

Pro správný popis posunů těžších atomů první periody hyperplocha stínění modelového peptidu nestačí. $^{15}N$, $^{17}O$ a $^{19}F$ jsou (kromě ovlivnění mezimolekulárními interakcemi popsanému v kapitole 3.3.) mnohem citlivější k silám dalekého dosahu. V případě bílkovin je proto stínění těchto jader silně ovlivněno přítomností polárních skupin. V principu samozřejmě lze tento efekt popsat *ab initio*, prakticky však by výpočet hyperplochy stínění velkého fragmentu biomolekuly byl ne-li neproveditelný, tak přinejmenším velmi nákladný. Proto se hledaly strategie, jež by umožnily zahrnutí vlivu interakcí dalekého dosahu bez zvyšování počtu atomů s bázovými funkcemi. Při nejrozšířenějším přibližném popisu[36] se celkové stínění $\sigma$ daného atomu předpokládá ve tvaru součtu

$$\sigma = \sigma_e + \sigma_p + \sigma_o \qquad (10)$$

kde $\sigma_e$ je stínění vypočtené některou z dosud uvedených metod (elektronický příspěvek, nazývaný také příspěvek sil krátkého dosahu, protože se určuje pouze pro malou okolní část kolem daného atomu biomolekuly), $\sigma_p$ vyjadřuje elektrostatickou polarizaci v okolí jádra (příspěvek sil dalekého dosahu), a $\sigma_o$ zahrnuje ostatní vlivy na stínění (magnetické – susceptibilitu vzorku, vliv kruhových proudů a pod.). Člen $\sigma_o$ je pro atomy jiné než vodík malý; mívá hodnoty řádu chyb určení ostatních členů z (*10*) a proto se zanedbává.

Pro výpočet $\sigma_p$ se ujaly dvě metody podrobně popsané v práci[48]. První z nich atomy zkoumaného fragmentu, které nepopisuje člen $\sigma_e$, aproximuje bodovými náboji. Tyto jsou formou multipólů nižších řádů zahrnuty do hamiltoniánu pro výpočet složek tenzoru stínění. Výsledkem je hodnota, která je součtem $\sigma_e$ a $\sigma_p$. Tento postup je označován jako metoda Charge Field Perturbation (CFP). Druhý přístup, tzv. Multipole Shielding Polarizability (MSP), využívá polarizabilit stínění, tj. derivací stínění podle složek vnějšího elektrostatického pole. Při předpokládaném tvaru závislosti stínění na elektrickém poli lze $\sigma_p$ stanovit na základě klasicky určené hodnoty elektrostatického pole (a jeho gradientů) v místě studovaného jádra. V práci[36] byly studovány obě zmíněné metodiky. Přístup MSP úspěšně reprodukoval $^{15}N$ stínění amidických dusíků

valinových reziduí a metoda CFP správně určila $^{19}F$ posuny 5-F-tryptofanů zkoumaných proteinů.

Výpočet trojrozměrné struktury biomakromolekul z NMR dat je možné schematicky popsat takto: Nejprve jsou z vícedimenzionálních NMR spekter přiřazeny jednotlivé signály konkrétním atomům ve studované molekule. S využitím nukleárního Overhauserova efektu (NOE) jsou určeny vzdálenosti vodíkových atomů menší než 5A. Ze změřených spin-spinových interakčních konstant jsou stanoveny hodnoty některých dihedrálních úhlů pomocí Karplusových rovnic. Tyto údaje společně se základními molekulárními parametry slouží následně jako vstup pro geometrickou optimalizaci pomocí metod distanční geometrie a molekulové dynamiky. Uvedené metody s využitím poměrně jednoduchých silových polí hledají výsledné struktury nejlépe splňující vstupní vazebné podmínky. V optimálním případě lze naznačeným postupem nalézt globální energetické minimum studovaného systému.

Poněvadž chemický posun je jednoznačně diktován strukturním uspořádáním, nabízí se myšlenka využít hyperploch stínění k vymezení dalších vazebných podmínek pro optimalizaci geometrie. Byly navrženy dva možné postupy[40]:

První z nich využívá naměřených posunů jader pro získání hodnot dihedrálních úhlů. Např. z dat $^{13}C$ a $^{1}H$ je možné v případě aminokyselin zjistit úhly φ, ψ a χ. Pro zahrnutí takto získaných torzí do molekulové dynamiky byla vyvinuta metodika tzv. plochy *Z* (Z-surface)[41]. V obecném případě plocha Z udává pravděpodobnost, s jakou soubor spektroskopických parametrů $\{P_i\}$ (např. $^{13}C$ chemických posunů) závisí na konkrétním prostorovém uspořádání atomů (např. dihedrálních úhlech φ, ψ bílkoviny). Známe-li funkční závislost $P = P(f(\alpha, \beta, \ldots))$, je pravděpodobnost, že experimentální hodnota $P_{exp}$ odpovídá geometrickým parametrům α, β, …, definována plochou *Z*

$$Z = \exp\left(-\frac{P_{exp} - f(\alpha,\beta,\ldots)^2}{W}\right) \qquad (11)$$

W je „šířka prohledávání“ vyjadřující výpočetní resp. experimentální nepřesnosti. Při řešení struktury s využitím plochy *Z* jsou kromě experimentálních dat do výpočtu zahrnuty i dihedrální úhly φ, ψ, χ získané z naměřených chemických posunů. Tyto údaje slouží jako vstupní data pro výpočet molekulové dynamiky programem[49] X-PLOR, který hledá energetická minima konzistentní s uvedenými daty. Výsledkem tedy jsou mj. nové dihedrální úhly φ', ψ', χ', pro něž je možno zpětně vypočítat chemické posuny metodami *ab initio* za účelem ověření nalezené geometrie.

Druhou možností je vypočítat pro experimentální úhly hodnoty posunů *ab initio* a tyto posuny zahrnout jako vazebné podmínky minimalizace. Pro tyto účely byla vyvinuta modifikovaná verze X-PLORu označovaná jako DPLOR (cit.[40]). Výslednou strukturu lze opět podrobit kontrole určením teoretických posunů.

### 3.3. Vliv vodíkových vazeb

Všechny běžné sekundární struktury nalezené v bílkovinách leží v povolených oblastech konformační mapy torzních úhlů φ, ψ (Ramachandranův diagram). Pokud konformace spadající do některé z těchto oblastí má být sice stericky povoleným, ale nestálým uspořádáním, musí být stabilizovaná. Na této stabilizaci se významně podílejí vodíkové můstky. Nyní se zmíníme o nejdůležitějších pracích[32,37,46], které simulovaly ovlivnění chemického stínění intermolekulárními vodíkovými vazbami mezi strukturními prvky molekul s biologickou relevancí. Autoři nereprodukovali pouze hodnotu izotropního stínění studovaných atomů, ale snažili se získat kompletní informaci, kterou lze analýzou tenzoru chemického stínění obdržet, tj. tři hlavní složky $\sigma_{xx}$, $\sigma_{yy}$, $\sigma_{zz}$ a jejich orientaci určenou třemi směrovými kosiny vzhledem ke zvolené soustavě souřadnic molekuly. Experimentálně lze uvedená data získat z NMR měření v pevné fázi. Z teoretického hlediska je při výpočtu podstatné, že k faktorům jako neúplnost báze, chyba metody (např. nezahrnutí korelační energie), kvalita geometrie a vliv mezimolekulárních interakcí je těchto šest hodnot obecně mnohem citlivějších než izotropní stínění $\sigma^{izo}$.

Pro studium vlivu mezimolekulárních interakcí existují dva základní přístupy. První z nich do výpočtu *ab initio* explicitně zahrnuje okolní molekuly. V již zmíněné studii[32] byly přístupem CHF-GIAO v malé bázi (D95**) reprodukovány hlavní složky tenzoru $^{15}N$ stínění dusíků uracilu se střední kvadratickou odchylkou 25,1 ppm a směrnicí lineární závislosti naměřených hodnot proti vypočteným –0,94. Krystal uracilu je uspořádán tak, že dusíkové atomy sousedí s kyslíky okolních molekul mřížky. Vliv tohoto uspořádání na stínění byl simulován přidáním formaldehydu k oběma dusíkům, a to do vzdálenosti stejné jako v krystalové struktuře uracilu. Shoda teorie s experimentem je v případě tohoto postupu ještě lepší (pro uvedenou bázi klesla střední kvadratická odchylka na 16,9 ppm a směrnice byla 1,01).

$^{13}C$ posuny karbonylových uhlíků podobně jako posuny atomů $C_\alpha$ a $C_\beta$ silně korelují se sekundární strukturou bílkovin – např. alaninová rezidua β-hřebenů jsou v průměru o 4,6 ppm více stíněna oproti α-šroubovicím[50]. Cílem práce[37] bylo reprodukovat tento experimentální trend. Jelikož předchozí výpočty *ab initio*, které vliv vodíkových vazeb neuvažovaly, neuspěly[42,43], autoři vytvořili jednoduchý model – N-methylacetamid s vodíkovou vazbou na formamid – a pro jeho různé geometrie zkoumali metodou CHF-GIAO s dostatečně velkou (6-311++G(2d,2p)) bází stínění karbonylového uhlíku. Experimentální tendence byla reprodukována. Výborné shody s výše uvedeným empirickým faktem bylo dosaženo pro N-formylpentaalaninamid jako model alaninových reziduí bílkovin. Při použití optimalizované geometrie jeho α-šroubovice a β-hřebenu CHF-GIAO přístup s bází 6-31G předpověděl rozdíl 4,9 ppm.

Aminokyseliny jsou v krystalu přítomny ve formě obojetného iontu, tj. s $-CO_2^-$ a $-NH_3^+$ skupinou, u něhož se dá předpokládat silné ovlivnění stínění karbonylových uhlíků mezimolekulárními interakcemi. Práce[48] zabývající se α-glycinem v pevné fázi toto potvrdila: Výpočet SOS-DFPT-IGLO s adekvátní bází pro izolovaný obojetný ion poskytl $^{13}C$ posun karbonylového uhlíku zatížený značnou chybou ($\delta_{teor} = -29{,}1$ ppm, $\delta_{exp} = -48{,}6$ ppm) a dokonce zaměnil pořadí hlavních složek tenzoru chemického stínění. Zahrnutí dvou $NH_4^+$ molekul v sousedství $-CO_2^-$ skupiny do výpočtu zlepšilo výsledky jen částečně: $\delta_{teor} = -35{,}6$ ppm při správném pořadí hlavních složek. Teprve rozšíření tohoto klastru o dva $HCO_2^-$ fragmenty v okolí $-NH_3^+$ skupiny glycinu poskytlo pro toto obtížné jádro výsledky v dobré shodě s experimentem ($\delta_{teor} = -43{,}8$ ppm, $\delta_{exp} = -48{,}6$ ppm).

Do výpočtu *ab initio* je taktéž možno explicitně zahrnout

molekuly rozpouštědla a studovat tak vliv solvatace, která může mít podstatný vliv na chemické stínění některých jader. (Je dobré si uvědomit, že např. hyperplochy stínění zmíněné v kapitole 3.2. byly získány výpočty *ab initio* izolovaných molekul. Použitelnost hyperploch pro popis dvourozměrné struktury je umožněna příznivým faktem, že změny posunů uhlíků $C_\alpha$ a $C_\beta$ závisí prakticky pouze na lokální struktuře, tj. na hodnotách $\varphi$ a $\psi$ při zanedbatelném vlivu současných změn ostatních vnitřních souřadnic, a nejsou ovlivněny prostředím, v němž se konkrétní reziduum v biomolekule nachází.) Pro peptidy byl vliv solvatace výpočetně studován na úrovni CHF-GIAO (cit.[45]). K modelovému dipeptidu byly postupně přidány dvě molekuly vody, jedna do okolí karbonylového kyslíku, druhá k amidickému protonu, a zkoumány změny tenzoru chemického stínění atomů peptidické vazby. Byla pozorována aditivita vlivu hydratace na složky tenzoru a jeho orientaci (*1*) (příslušný údaj pro strukturu s dvěma molekulami vody byl přibližně dvojnásobkem hodnoty v monohydrátu) a výrazná změna orientace tenzoru amidického protonu při hydrataci (*2*). Pokud by uvedené závislost byly prokázány i pro větší modelové peptidy a vyšší rozsah solvatace, mohly by najít použití při určování stupně solvatace z měření změn posunů (*1*) či určování geometrie solvatovaného fragmentu na základě znalosti orientace tenzorů chemického stínění (*2*).

Druhou principiální možností postihnutí vlivu mezimolekulárních interakcí na chemické stínění je nasazení výše popsaného přístupu CFP (cit.[48]), což tomto případě znamená zahrnutí bodových nábojů atomů tvořících vodíkové vazby se studovaným fragmentem. V publikaci[39] jsou výsledky výpočtů pro threonin a tyrosin v pevné fázi, kde, jak již bylo uvedeno, se aminokyseliny nacházejí ve formě obojetných iontů a jsou proto silně ovlivněny intermolekulárními vodíkovými můstky. Pro složky tenzorů stínění $^{13}C$ izolovaných obojetných iontů autoři obdrželi střední kvadratickou odchylku mezi teoreticky vypočtenými a experimentálními hodnotami 12,3 ppm. Pro zlepšení výsledku výpočtu byly intermolekulární interakce simulovány tak, že základní buňka obsahující čtyři aminokyseliny byla obklopena dalšími 28 molekulami v geometrické konfiguraci odpovídající uspořádání v krystalové mříži. Na atomy jedné z aminokyselin základní buňky poté umístili funkce báze 6-31G** a všechny ostatní aminokyseliny reprezentovali bodovými náboji. Střední kvadratická odchylka klesla o polovinu (na 6,4 ppm).

Již dříve zmíněná práce[36] využila pro modelování vodíkové vazby také přístupu CFP. Ke kyslíku peptidické skupiny modelového dipeptidu byly umístěny bodové náboje reprezentující amoniak. Byla zkoumána závislost rozdílu mezi $^{15}N$ stíněním peptidického dusíku v tomto uspořádání a v izolovaném dipeptidu na délce vodíkové vazby. Pro typickou délku vodíkových můstků v bílkovinách 3,0 A byla hodnota tohoto rozdílu přibližně –1 ppm. Kontrolní výpočet, jenž celý klastr počítal *ab initio*, poskytl hodnotu blízkou –2 ppm. Tento model vlivu vodíkových vazeb na dusíkové stínění je maximálně zjednodušený a reprodukuje pouze znaménko změn (experimentální rozdíly jsou větší než –30 ppm).

## 4. Výhled do budoucna

Z výše prezentovaného přehledu vyplývá, že v současnosti existuje již několik dobře zvládnutých metod výpočtu chemického stínění použitelných pro rozměrné systémy. Diskutované metodiky dokáží nejenom správně předpovědět hodnotu chemického posunu, ale i teoreticky zkoumat konstituční, geometrické, rozpouštědlové a jiné vlivy ovlivňující chemický posun. Tím umožňují porozumění jeho podstatě a faktorům, na nichž tato důležitá experimentální veličina závisí.

Jelikož se moduly pro výpočet chemického stínění stávají součástí běžně dostupných kvantově chemických programů, není daleko doba, kdy pro fragmenty biomolekul bude možné počítat chemické posuny přímo na pracovních stanicích NMR laboratoří. Velký nárůst aplikací lze očekávat hlavně v případě nukleových kyselin. Jistě se nezastaví ani vývoj teorie. Příznivé důsledky pro výpočty *ab initio* posunů biomolekul bude mít další rozvoj teorie funkcionálu hustoty. V této souvislosti je očekávána hlavně výpočetně efektivní metodika s výměnně korelačním funkcionálem přímo závisejícím na magnetickém poli[51]. V budoucnu bude snad možné i pro větší molekuly přesněji předpovídat a lépe pochopit hodnoty chemických posunů, podobně jak lze již dnes určovat jejich geometrie a energie.

LITERATURA


1. Chesnut D. B.: Annu. Rep. NMR Spectrosc. *29*, 71 (1994).
2. Jameson C. J., v knize: *Nuclear Magnetic Resonance, Specialist Periodical Reports* (Webb G.A., ed.), sv. 8–25. The Royal Society of Chemistry, London 1980–1996.
3. Cheeseman J. R., Trucks G. W., Keith T. A., Frish M. J.: J. Chem. Phys. *104*, 5497 (1996).
4. de Dios A. C., Pearson J. G., Oldfield E.: Science *260*, 1491 (1993).
5. de Dios A.C.: J. Prog. Nuc. Magn. Spectrosc. *29*, 229 (1996).
6. Jameson A. K., Jameson C. J.: Chem. Phys. Lett. *134*, 461 (1987).
7. Ramsey N. F.: Phys. Rev. *78*, 699 (1950).
8. Ramsey N. F.: Phys. Rev. *86*, 243 (1952).
9. Beveridge D. L., v knize: *Semiempirical Methods of Electronic Structure Calculation* (Segal G. A., ed.), str.163. Plenum Press, New York 1977.
10. Hehre W. J., Radom L., Schleyer P. v. R., Pople J. A.: *Ab Initio Molecular Orbital Theory*. Wiley, New York 1986.
11. Stevens R. M., Pitzer R. M., Lipscomb W. N.: J. Chem. Phys. *38*, 550 (1963).
12. Ditchfield R., Miller D. P., Pople J. A.: J. Chem. Phys. *49*, 2960 (1968).
13. Epstein S. T.: J. Chem. Phys. *42*, 2897 (1965).
14. Epstein S. T.: J. Chem. Phys. *58*, 1592 (1973).
15. Ditchfield R.: Mol. Phys. *27*, 789 (1974).
16. Wolinski K., Hinton J. F., Pulay P.: J. Am. Chem. Soc. *112*, 8251 (1990).
17. Kutzelnigg W.: Israel J. Chem. *19*, 193 (1980).
18. Schindler M., Kutzelnigg W.: J. Chem. Phys. *76*, 1919 (1982).
19. Hansen Aa. E., Bouman T. D.: J. Chem. Phys. *82*, 5035 (1985).
20. Bouman T. D., Hansen Aa. E.: Chem. Phys. Lett. *82*, 510 (1988).
21. Keith T. A., Bader R. F. W.: Chem. Phys. Lett. *210*, 223 (1993).

22. Keith T. A., Bader R. F. W.: J. Chem. Phys. *99*, 3669 (1993).
23. Gauss J.: Chem. Phys. Lett. *191*, 614 (1992).
24. Gauss J.: Chem. Phys. Lett. *229*, 198 (1994).
25. Parr R. G., Yang W.: *Density – Functional Theory of Atoms and Molecules*. Clarendon Press, Oxford 1989.
26. Malkin V. G., Malkina O. L., Casida M. E., Salahub D. R.: J. Am. Chem. Soc. *116*, 5898 (1994).
27. Malkin V. G., Malkina O. L., Eriksson L. A., Salahub D. R., v knize: *Theoretical and Computational Chemistry* (Seminario J. M., Politzer P., ed.), sv. II, str. 273. Elsevier, Amsterdam 1995.
28. Hricovíni M., Malkina O. L., Bízik F., Turi Nagy L., Malkin V. G.: J. Phys. Chem. A *101*, 9756 (1997).
29. Case D. A.: J. Biomol. NMR *6*, 341 (1995).
30. Schindler M.: J. Am. Chem. Soc. *110*, 6623 (1988).
31. Czernek J., Sklenář V.: J. Phys. Chem. *A 103*, 4089 (1999).
32. Anderson-Altmann K. L., Phung C. G., Mavromoustakos S., Zheng Z., Facelli J. C., Poulter C. D., Grant D. M.: J. Phys. Chem. *99*, 10454 (1995).
33. Spera S., Bax A.: J. Am. Chem. Soc. *113*, 5490 (1991).
34. Wishart D. S., Sykes B. D., Richards F. M.: J. Mol. Biol. *222*, 311 (1991).
35. de Dios A. C., Pearson J. G., Oldfield E.: J. Am. Chem. Soc. *115*, 9768 (1993).
36. de Dios A. C., Pearson J. G., Oldfield E.: Science *260*, 1491 (1993).
37. de Dios A. C., Oldfield E.: J. Am. Chem. Soc. *116*, 11485 (1994).
38. de Dios A. C., Oldfield E.: J. Am. Chem. Soc. *116*, 5307 (1994).
39. de Dios A. C., Laws D. D., Oldfield E.: J. Am. Chem. Soc. *116*, 7784 (1994).
40. Le H., Pearson J. G., de Dios A. C., Oldfield E.: J. Am. Chem. Soc. *117*, 3800 (1995).
41. Pearson J. G., Wang J., Markley J. L., Le H., Oldfield E.: J. Am. Chem. Soc. *117*, 8823 (1995).
42. Jiao D., Barfield M., Hruby V. J.: J. Am. Chem. Soc. *115*, 10883 (1993).
43. Sulzbach H. M., Schleyer P. v. R., Schaefer III H. F.: J. Am. Chem. Soc. *116*, 3967 (1994).
44. Chesnut D. B., Phung C. G.: Chem. Phys. Lett. *183*, 505 (1991).
45. Chesnut D. B., Phung C. G., v knize: *Nuclear Magnetic Shieldings and Molecular Structure* (Tosell J. A., ed.), str. 221. Kluwer, Amsterdam 1993.
46. Malkin V. G., Malkina O. L., Salahub D. R.: J. Am. Chem. Soc. *117*, 3294 (1995).
47. Woolf T. B., Malkin V. G., Malkina O. L., Salahub D. R., Roux B.: Chem. Phys. Lett. *239*, 18 (1995).
48. de Dios A. C., Oldfield E.: Chem. Phys. Lett. *205*, 108 (1993).
49. Brunger A. T.: *X-PLOR*. Yale University Press, New Haven 1992.
50. Wishart D. S., Sykes B. D.: J. Biomol. NMR *4*, 171 (1994).
51. Harris R. A., Salsbury F. R., Jr.: J. Chem. Phys. *109*, 2609 (1998).

**J. Czernek and V. Sklenář** *(Laboratory of Structure and Dynamics of Biomolecules, Faculty of Science, Masaryk University, Brno):* ***Ab initio*** **Calculations of NMR Chemical Shifts in Biomolecular Systems**

A survey of the theory of *ab initio* calculations of chemical shifts of $^{1}$H, $^{13}$C, $^{15}$N and $^{19}$F in NMR spectra and its most important applications in the study of biologically relevant molecules are given. The review is focused on theoretical investigation of constitutional, geometric, and hydration-interaction effects on the chemical shift.