Strukturní biochemie

Skripta k přednášce C9530 Přírodovědecké fakulty Masarykovy univerzity

Lukáš Žídek
Na přípravě se podíleli Jiří Damborský, Eva Fadrná, Jaromír Marek, Markéta Munzarová, Pavel Plevka

6. prosince 2019
# Obsah

1 Procházka ZOO struktur 1

1 Struktura v chemii a biochemii 3
   1.1 Chemická struktura a geometrie 3
      1.1.1 Konfigurace a konformace 3
      1.1.2 Geometrický popis molekul 4
      1.1.3 Definice torzního úhlu 7
   1.2 Konformace a energie 8
      1.2.1 Energie a entropie 8
      1.2.2 Síly působící v biomakromolekulách 9
      1.2.3 Celková energie a výsledná konformace 10
      1.2.4 Termodynamika a kinetika 11

2 Struktura proteinů 13
   2.1 Proteiny jako příklad biomakromolekul 13
   2.2 Stavební jednotky 13
   2.3 Torzní úhly 16
   2.4 Primární struktura 18
   2.5 Sekundární struktura 19
      2.5.1 Struktura alfa 19
      2.5.2 Struktura beta 21
   2.6 Terciární struktura 23
      2.6.1 Spojovací prvky 23
      2.6.2 Alfa-proteiny 24
      2.6.3 Beta-proteiny 29
      2.6.4 Alfa+beta-proteiny 34
      2.6.5 Alfa/beta-proteiny 34
      2.6.6 Proteiny bez sekundární struktury 34
   2.7 Kvartérní struktura a vyšší struktury 34

3 Struktura nukleových kyselin 37
   3.1 Stavební jednotky 37
   3.2 Torzní úhly 39
   3.3 Primární struktura 40
   3.4 Sekundární struktura 40
   3.5 Terciární struktura 43
   3.6 Kvartérní struktura a vyšší struktury 48
# OBSAH

## 4 Struktura oligo- a polysacharidů

4.1 Stavební jednotky ........................................ 49
4.2 Torzní úhly .................................................. 49
4.3 Primární struktura .......................................... 51
4.4 Sekundární struktura ...................................... 51
4.5 Terciární struktura .......................................... 51

## 5 Struktura biologických membrán

5.1 Membrány a biomakromolekuly ........................................ 55
5.2 Stavební jednotky ............................................. 55
5.3 Geometrie lipidových útvarů ...................................... 56
5.4 Tekutost membrán ............................................. 56
5.5 Konformace lipidů a struktura membrány

5.5.1 Vliv polární skupiny ...................................... 56
5.5.2 Vliv hydrofobního řetězce ................................ 57
5.5.3 Vliv složení ................................................ 58
5.6 Vyšší struktury ................................................ 58

## II Lov struktur

## 6 Výpočetní metody

6.1 Molekulové modelování ....................................... 61
6.1.1 Model molekuly .............................................. 61
6.1.2 Minimalizace energie ...................................... 62
6.2 Kvantové metody .............................................. 63
6.2.1 Kvantová mechanika atomů ................................ 63
6.2.2 Molekulové orbitály ........................................ 64
6.2.3 Metody \textit{ab initio} ......................................... 66
6.2.4 Semiempirické metody ...................................... 69
6.3 Molekulová mechanika .......................................... 69
6.3.1 Hyperplocha potenciální energie ......................... 70
6.3.2 Silové pole .................................................. 72
6.3.3 Implicitní a explicitní solvent ............................. 73
6.3.4 Metoda PME ................................................. 73
6.3.5 Molekulová dynamika ...................................... 75

## 7 Příprava vzorku biomakromolekul

7.1 Izolace z přírodního materiálu ................................ 79
7.2 Metody molekulární biologie ................................... 80
7.2.1 Principy molekulární genetiky ............................ 80
7.2.2 Expressní vektory .......................................... 81
7.2.3 Molekulové klonování ..................................... 82
7.2.4 Místně řízená mutageneze .................................. 83
7.2.5 Bakteriální exprese ......................................... 83
7.2.6 Purifikace exprimovaných proteinů ..................... 84
7.3 Chemická syntéza ............................................. 84
7.3.1 Syntéza peptidů .............................................. 84
7.3.2 Syntéza oligonukleotidů .................................. 86
7.3.3 Syntéza oligosacharidů ................................... 86
<table>
<thead>
<tr>
<th>OBSAH</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7.4 Enzymová syntéza</td>
<td>86</td>
</tr>
<tr>
<td>7.4.1 Enzymová syntéza nukleových kyselin</td>
<td>86</td>
</tr>
<tr>
<td>7.4.2 Enzymová syntéza proteinů</td>
<td>87</td>
</tr>
<tr>
<td>7.4.3 Enzymová syntéza oligosacharidů</td>
<td>87</td>
</tr>
<tr>
<td>8 Metody určování sekvence</td>
<td>89</td>
</tr>
<tr>
<td>8.1 Sekvenace proteinů</td>
<td>89</td>
</tr>
<tr>
<td>8.1.1 Edmanovo odbourávání</td>
<td>89</td>
</tr>
<tr>
<td>8.1.2 Sekvenace hmotnostní spektrometrie</td>
<td>90</td>
</tr>
<tr>
<td>8.2 Sekvenace nukleových kyselin</td>
<td>92</td>
</tr>
<tr>
<td>8.2.1 Chemická metoda Maxama a Gilberta</td>
<td>93</td>
</tr>
<tr>
<td>8.2.2 Enzymová neboli dideoxymetoda (Sangerova)</td>
<td>93</td>
</tr>
<tr>
<td>8.3 Sekvenace oligosacharidů</td>
<td>93</td>
</tr>
<tr>
<td>9 Úvod do experimentálních metod strukturní biochemie</td>
<td>97</td>
</tr>
<tr>
<td>10 Optické metody</td>
<td>99</td>
</tr>
<tr>
<td>10.1 Elektromagnetické záření</td>
<td>99</td>
</tr>
<tr>
<td>10.2 Světlo mezi molekulami</td>
<td>100</td>
</tr>
<tr>
<td>10.2.1 Světlo mezi malými netečnými molekulami</td>
<td>101</td>
</tr>
<tr>
<td>10.2.2 Světlo si vyměňuje energii s molekulami</td>
<td>101</td>
</tr>
<tr>
<td>10.2.3 Vlnová rovnice mezi molekulami</td>
<td>101</td>
</tr>
<tr>
<td>10.3 Intenzita světla a spektrum</td>
<td>102</td>
</tr>
<tr>
<td>10.4 Světlo a sekundární struktura</td>
<td>102</td>
</tr>
<tr>
<td>10.4.1 Polarizované světlo a sekundární struktura</td>
<td>103</td>
</tr>
<tr>
<td>10.5 A co jádra?</td>
<td>106</td>
</tr>
<tr>
<td>10.5.1 Vibrací spektroskopie</td>
<td>106</td>
</tr>
<tr>
<td>10.5.2 Vibrace jader v peptidové vazbě</td>
<td>108</td>
</tr>
<tr>
<td>10.6 Rozptyl světla</td>
<td>108</td>
</tr>
<tr>
<td>10.6.1 Ramanova spektra</td>
<td>109</td>
</tr>
<tr>
<td>11 Rentgenová krystalografie</td>
<td>111</td>
</tr>
<tr>
<td>11.1 Ohyb záření (difrakce)</td>
<td>111</td>
</tr>
<tr>
<td>11.1.1 Virtuální nahrada mikroskopu</td>
<td>111</td>
</tr>
<tr>
<td>11.1.2 Braggovy rovnice</td>
<td>113</td>
</tr>
<tr>
<td>11.1.3 Difrakce na krystále</td>
<td>114</td>
</tr>
<tr>
<td>11.1.4 Strukturní faktor</td>
<td>116</td>
</tr>
<tr>
<td>11.2 Krystalizace biomakromolekul</td>
<td>117</td>
</tr>
<tr>
<td>11.3 Měření difrakce</td>
<td>119</td>
</tr>
<tr>
<td>11.3.1 Zdroj záření</td>
<td>119</td>
</tr>
<tr>
<td>11.3.2 Goniostat</td>
<td>120</td>
</tr>
<tr>
<td>11.3.3 Detektor</td>
<td>120</td>
</tr>
<tr>
<td>11.3.5 Vyhodnocení dat</td>
<td>120</td>
</tr>
<tr>
<td>11.4 Řešení fázového problému</td>
<td>121</td>
</tr>
<tr>
<td>11.4.1 Pattersonova mapa</td>
<td>121</td>
</tr>
<tr>
<td>11.4.2 Metoda molekulárního přemísťení</td>
<td>122</td>
</tr>
<tr>
<td>11.4.3 Průměřní metody</td>
<td>123</td>
</tr>
<tr>
<td>11.4.4 Metody izomorfního nahrazení</td>
<td>123</td>
</tr>
<tr>
<td>11.4.5 Metody využívající anomální rozptyl</td>
<td>125</td>
</tr>
</tbody>
</table>
# Nukleární magnetická rezonance

12.1 Magnetické chování jader ........................................ 131  
12.2 Ovlivnění jader vnějšími poli .................................... 132  
12.3 Interakce jader s nejbližším okolím .......................... 135  
  12.3.1 Interakce s magnetickými poli párových elektronů  
  12.3.2 Průmá interakce s magnetickými dipolí sousedních jader  
  12.3.3 Interakce mezi magnetickými momenty jader zprostředkovaná elektrony vazeb ............ 138  
12.4 Spektrometr a experiment NMR ............................... 140  
12.5 Vícevrstvá spektroskopie a korelace frekvencí .......... 143  
12.6 Vztah mezi spektre NMR a strukturou ...................... 144  
12.7 Heteronukleární korelace a sekvenci přiřazení ......... 145  
12.8 Homonukleární korelace a postranní řetězec ............. 146  
12.9 NMR spektra a struktura molekule ......................... 147  
12.10 Sekundární struktura ........................................ 148  
12.11 Vypočet konformace proteinu ............................... 149  
12.12 Struktura nukleových kyselin pomocí NMR ......... 150  
  12.12.1 Nukleární Overhauserov jev a vzdálenosti protonů ............ 151  
  12.12.2 Trívazebné interakce a torzní úhly ................. 151  
  12.12.3 Vypočet konformace nukleových kyselin ............... 151  
12.13 Struktura oligosacharidů pomocí NMR .................. 151  
12.14 Polychy molekule v NMR .................................. 152  
12.15 Další vývoj metody ........................................ 152  

# Bioinformatika

13.1 Co to je bioinformatika ....................................... 159  
13.2 Genomové projekty ........................................ 159  
13.3 Biologické databáze ....................................... 160  
13.4 Databáze sekvencí proteinů ................................ 160  
  13.4.1 Primární databáze sekvencí proteinů .......... 160  
  13.4.2 Složené databáze sekvencí proteinů .... 163  
  13.4.3 Sekundární databáze sekvencí proteinů .... 163  
13.5 Databáze sekvencí DNA ..................................... 164  
13.6 Databáze proteinových struktur ......................... 165  
13.7 Vyhledávací systémy ....................................... 165  
13.8 Párové přiřazení (pairwise alignment) sekvencí .... 166  
13.9 Žítačské přiřazení (multiple alignment) .......... 167  
13.10 Předpovídání struktury ze sekvence ................. 169  
  13.10.1 Předpověď sekundární struktury ............. 169  
13.11 Předpověď foldu .......................................... 169  
  13.11.1 Homologní modelování ......................... 169  

III Dodatky ...................................................... 173  
A Energii různých interakcí .................................... 175  
B Peptidová vazba jako příklad chromoforu ............... 177  
  B.1 Atomové orbitály ....................................... 177  
  B.2 Hybridní orbitály ....................................... 177
### OBSAH

<table>
<thead>
<tr>
<th>Kapitola</th>
<th>Název</th>
<th>Strana</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.3</td>
<td>Molekulové orbitaly</td>
<td>180</td>
</tr>
<tr>
<td>B.4</td>
<td>Energie stavů a vlnové délky přechodu</td>
<td>182</td>
</tr>
<tr>
<td>B.5</td>
<td>Tranzitní dipóly a pravděpodobnosti přechodu</td>
<td>183</td>
</tr>
<tr>
<td>C</td>
<td>Metody chemické syntézy na pevném nosiči</td>
<td>185</td>
</tr>
<tr>
<td>C.1</td>
<td>Karbodiimidová metoda syntézy peptidů</td>
<td>185</td>
</tr>
<tr>
<td>C.2</td>
<td>Fosforamiditová metoda syntézy nukleotidů</td>
<td>186</td>
</tr>
<tr>
<td>D</td>
<td>Anomální rozptyl</td>
<td>186</td>
</tr>
<tr>
<td>E</td>
<td>Přiřazení spekter NMR nukleových kyselin s využitím skalárních interakcí</td>
<td>188</td>
</tr>
</tbody>
</table>
Část I

Procházka ZOO struktur
Kapitola 1

Struktura v chemii a biochemii

1.1 Chemická struktura a geometrie

1.1.1 Konfigurace a konformace

Chceme-li studovat strukturu biologicky zajímavých látek, musíme se nejdříve dohodnout na tom, co vlastně pojmem *struktura*myslimme. Chemické látky se skládají z atomů různých prvků. První informaci, kterou o neznámé látce můžeme získat, je tedy údaj, kolik atomů jednotlivých prvků molekula látky obsahuje. Například můžeme zjistit, že molekula naší látky obsahuje čtyři atomy uhlíku, dva atomy kyslíku a osm atomů vodíku. Budeme znát *složení* látky

1

ale stále ještě nebudeme znát strukturu látky.

Teprve když zjistíme, které atomy jsou navzájem spojeny kovalentní vazbou

2

můžeme říci, zda zkoumaná látkou je kyselina máśelná, 2-buten-1,4-diol, nebo třeba 3-hydroxybutanon.

Ani teď však nemůžeme být jednoznačně dány, o jaké chemické látku jde. Předpokládejme, že řekneme, že zkoumaná látku je 2-buten-1,4-diol. Tímto názvem lze ale označit dvě různé chemikálie s různými vlastnostmi. Tyto dvě látky se budou lišit uspořádáním atomů kolem dvojně vazby. Budou-li obě koncové vazby C–C na jedné straně dvojně vazby, půjde o izomer *cis* (chemicky úplný název látky bude *(Z)*-2-buten-1,4-diol), budou-li koncové C–C vazby směřovat na opačné strany, půjde o izomer *trans* *(E)*-2-buten-1,4-diol).

Pokud by zkoumanou látkou byl 3-hydroxybutanon, měli bychom podobný problém. Museli bychom si vybrat ze dvou molekul, z nichž jedna je zrcadlovým obrazem druhé. Úplný chemický popis by tedy byl *(R)*-3-hydroxybutanon nebo *(S)*-3-hydroxybutanon.

Případ vzájemného uspořádání atomů propojených kovalentními vazbami, tak jak jsme jej probírali v předchozích odstavcích, se nazývá *konfigurace*. Zatímco kyselina mášelá se vyskytuje v jediné konfiguraci, 2-buten-1,4-diol a 3-hydroxybutanon mají po dvou různých konfiguracích (obrázek 1.1).

Se znalostí konfigurace již umíme popsat molekulu například chemickým vzorečkem (*s* vyznačením chirality). Pokud budeme chtít z obrazce a kuliček sestavit model zkoumané látky (jako na obrázku 1.1), třeba kyseliny mášelé, potřebujeme znát víc, než jen konfiguraci. Musíme se například rozhodnout, jestli uhlíkový řetězec ohneme tak, aby byla karboxylová kyselina nejbližší koncovému methylu, nebo jestli bude uhlíkový řetězec připomínat klikový hřídel a karboxylová skupina bude od methylové co nejdále. Konkrétní rozmístění atomů v prostoru se nazývá *konformace*. Mezi konfigurací a konformací

1 Chemici často mluví o *kompozici* látky.

2 Informace udávající pouze to, které atomy jsou spojeny, se nazývá *konektivita*. Chemika ovšem zajímá i to zda jde o vazbu jednoduchou nebo dvojou. Chemický popis atomů a vazeb v molekule se nazývá *konstituce*.

3 Účinně řečeno, jde o *antipody*, tedy *enantiostry* s opačnou *chirality*. Takové látky se liší směrem, kterým stáčí rovině polarizované světlo.
je důležitý rozdíl. Konformací našich molekul můžeme ménit pouhým otáčením kolem kovalentních vazeb, zatímco ke změně konfigurace je třeba kovalentní vazbu přerušit a vytvořit novou.

Konfigurace a konformace hrají důležitou roli v rozdílech chápání pojmu určené struktury v kontextu malých a velkých molekul. Pokud izolujeme z přírodního zdroje nějakou nízkomolekulární látku (tedy látku, jejíž molekuly se skládají z desítek atomů), například antibiotikum, musíme především určit konfiguraci této látky. Určování konformace je v takovém případě okrajová záležitost. U malých molekul, jako je kyselina maselná, jsou rotace kolem vazeb tak snadné a rychlé, že nemá většinou smysl o jednotlivých konformerech mluvit. Složitější molekuly se sice vyskytují v jedné či více z několika jasně definovaných konformací, určení těchto konformací ale stále představuje jednodušší část určování struktury. Určováním struktury malých molekul se tedy můžou především určováním konfigurace.


Určovániím struktury biomakromolekul se proto můžou především určování konformace. A to je úkol obtížný.

1.1.2 Geometrický popis molekul

Z matematického pohledu k určení propojení atomů stačí znát topologii molekulky, zatímco určení konformace vyžaduje znalost přesné geometrie. Pokud jsme si v předchozím odstavci řekli, že určování struktury biomakromolekul je vlastně popis jejich konformace, musíme se nejprve seznámit s geometrickým popisem molekul.

Geometrie nám umožňuje vytvořit jiznodušší obraz molekulky, tedy jejích model (zde slovo „model“ znamená opravdu to, co si pod ním zpravidla představíme). Pro potřeby strukturní biochemie většinou stačí považovat atomy za body v prostoru a chemické vazby za úsečky, které body spojují. Uvedomme si, že k vytvoření modelu nestačí znát pouze polohy všech atomů, ale musíme také vědět, mezi kterými atomy je kovalentní vazba (musíme určit topologii molekulky).

Polohu atomů v prostoru můžeme určit několika způsoby. Prvním způsobem, který nás asi napadne,
1.1. CHEMICKÁ STRUKTURA A GEOMETRIE

je udat souřadnice \( x \), \( y \), \( z \) každého atomu v obvyklé kartézské souřadné soustavě (obrázek 1.2). Tento způsob používá například databáze struktur PDB (Protein Data Bank), jak ukazuje příklad na obrázku 1.3. Pro popis modelu molekuly, která je tvůrčena \( N \) atomy, potřebujeme \( 3N \) souřadnic.

Obrázek 1.2: Kartézská souřadná soustava. Osa \( x \) má k nám, osa \( y \) doprava, osa \( z \) nahoru.

Výhoda kartézských souřadnic je v tom, že je můžeme přímo a jednoduše použít k zobrazení modelu molekuly. Kartézské souřadnice mají ale také své nevýhody. Představme si, že molekulu definovanou na obrázku 1.3 otočíme o 90° kolem osy \( x \). Polohy atomů pak budou udány souřadnicemi uvedenými na obrázku 1.4. Jak je vidět, hodnoty dvou souřadnic každého atomu se podstatně změnily – přitom jde stále o tutéž molekulu vody, která má naprosto stejný tvar. Proto se často polohy atomů udávají jiným způsobem. Pro molekulu skládající se z \( N \) atomů opět potřebujeme \( 3N \) údajů. Z těchto 3\( N \) hodnot budou tři udávat polohu těžiště molekuly, další tři natočení molekuly jako celku vzhledem k souřadné soustavě a zbývající údaje (kterých je pochopitelně \( 3N - 6 \)) vzájemnou polohou atomů uvnitř molekuly.

Prvních šest údajů nás zpravidla nezajímá, neboť se mění, jak se molekula pohybuje Brownovým pochodem. Zajímá nás pouze posledních 3\( N - 6 \) údajů, které jsou pro danou molekulu stejné, ať už je její poloha v souřadné soustavě jakákoliv. Tyto údaje se proto nazývají vnitřní souřadnice molekuly. Jako vnitřní souřadnice je zvykem udávat vzdálenosti mezi atomy, které jsou vázány kovalentní vazbou (vazebné vzdálenosti) a úhly, které kovalentní vazby svárají (vazebné úhly). Model molekuly vody, který jsme si uvedli jako příklad, je určen dvěma vazebními vzdálenostmi |\( O-H1 \)| a |\( O-H2 \)|, které jsou obě rovny 95,7 pm a jedním vazebným úhlem \( \angle H1O-H2 \), který je roven 104,5°.

---

6 Souřadná souřadnice musí být definována tak, jak ukazuje obrázek 1.2. Kdybychom otočili směr jedné osy, popisovali bychom opačný enantiomer.

7 Uvedené počty platí pouze pro molekulu, která není lineární. Lineární biomakromolekuly se ovšem nevyvíjejí, takže se jim nemusíme zabývat.

Stavbu modelu molekuly si tedy můžeme popsat následujícím způsobem:

- Polohu prvního atomu (označme si jej A) udávají tři souřadnice, které tímto definují polohu celé molekuly v souřadné soustavě. Jsou to vnitřní souřadnice vnější, které nás většinou nezajímají, v případě jediného atomu nemá smysl mluvit o vnitřních souřadnicích molekuly.
- Polohu druhého atomu (B) udává jedna vnitřní souřadnice, vazebná délka |A–B|, a dvě vnější souřadnice, které definují natočení první vazby vůči souřadné soustavě.
- Polohu třetího atomu (C) udávají dvě vnitřní souřadnice, vazebná délka |B–C| a vazebný úhel ∠A–B–C, a jedna vnější souřadnice, která udává natočení druhé vazby vůči souřadné soustavě (teprve tím je definováno natočení celé molekuly v soustavě souřadnic).

Jiná možnost, jak tento úhel popsat, je určit úhel, který svází atomy, ve kterých leží atomy H1,01,02 a O1,02,H2. Jako vnitřní souřadnice bychom mohli stejně dobře použít dihedrální úhel, tedy doslova úhel mezi dvěma stěnami.
1.1. CHEMICKÁ STRUKTURA A GEOMETRIE

• Polohu čtvrtého atomu (D) udávají tři vnitřní souřadnice. Pokud jej připojíme k atomu C, bude to vazebná délka |C–D|, vazebný úhel $\angle \text{B–C–D}$ a tozní úhel popisující rotaci kolem vazby B–C.

• Polohu všech dalších atomů udávají vždy tři vnitřní souřadnice, v případě nevětveného řetězce (lineárního nebo cyklického) délka vazby s předchozím atomem, úhel mezi dvěma předchozími vazbami a tozní úhel popisující rotaci kolem předposlední vazby.

1.1.3 Definice tozního úhlu

Zatímco definice vazebná délky a vazebnáho úhlu jsou asi každému zřejmé, definice tozního úhlu si zaslouží pečlivou pozornost. Pro každou kovalentní vazbu, která není koncová, definujeme tozní úhel (popisující rotaci kolem této vazby) následujícím způsobem:

1. Nalezneme řetězec atomů A–B–C–D, který definuje rotaci kolem vazby B–C. Pokud jsou na atom B nebo C (nebo oba) naváženy více než dva atomy, určíme atomy A a D podle následujících pravidel:

   • Je-li vazba B–C součástí hlavního řetězce (páteré) makromolekuly, zvolíme atomy A a D tak, aby byly také součástí hlavního řetězce (páteré).
   • Je-li vazba B–C součástí postranního řetězce makromolekuly, zvolíme atomy A a D tak, aby měly nejvyšší prioritu podle pravidel používaných v organické chemii (zavedených Cahnem, Ingoldem a Prelogem).

2. Umístíme molekulu v prostoru tak, aby vazba B–C byla kolmá na rovinu papíru. Podíváme se na molekulu ve směru vazby B–C (atom B je blíže k nám a atom C je v zákrytu za ním) a promítneme vazby A–B a C–D do roviny papíru.

3. Změříme úhel mezi průměty vazeb A–B a C–D. Hodnoty úhlu udáváme v rozsahu od $-180^\circ$ do $+180^\circ$, kladné hodnoty nabývá tozní úhel v případě, že vazba C–D je pootočena oproti vazbě A–B po směru hodinových ručiček (obrázek [L6]).

Kromě udání číselných hodnot lze přibližnou hodnotu tozního úhlu popsat slovně. Používá se k tomu různá názvosloví, která jsou také ukažána na obrázku [L6]. První z nich v závade označení pro nejvyhodnější konformaci. Tak jako v případě dvojně vazby C=C jsou výrazně nejvýhodnější hodnoty tozního úhlu 0 a $+180^\circ$ (izomery cis a trans), v případě jednoduché vazby C–C mezi dvěma sp³ uhliky jsou nejvyhodnější tři tozní úhly, které se označují gauche$(+)$ ($+60^\circ$), gauche$(-)$ ($-60^\circ$) a trans ($180^\circ$).

Druhé názvosloví vychází z rozdělení konformací na dvě skupiny. První skupina, označovaná $syn$, zahrnuje konformace, ve kterých jsou průměty vazeb A–B a C–D vzájemně pootočeny o méně než 90° jedním či druhým směrem (léží tedy ve stejně polovině kruhu). Naopak druhá skupina, označovaná $anti$, zahrnuje konformace, ve kterých jsou průměty vazeb A–B a C–D vzájemně pootočeny o více než 90° (léží tedy ve opačně polovině kruhu). Tyto dvě skupiny lze kombinovat o údaj popisující, zda čtverce atomů leží spíše v jedné rovině (periplanární konformace) nebo ve dvou různých rovinách (klinální konformace). Podle tohoto názvosloví tedy hovoříme například o kladném synperiplanárním úhlu (0 až $+30^\circ$, zkratka $sp$), nebo o záporném antiklinálním úhlu ($-150^\circ$ až $-90^\circ$, zkratka $-ac$).

---

KAPITOLA 1. STRUKTURA V CHEMII A BIOCHEMII

1.2 Konformace a energie

1.2.1 Energie a entropie

Naší hlavním cílem je popsat strukturu (tedy konformaci) biomakromolekul. Výsledný statický geometrický model je důležitým základem pro pochopení dynamiky a funkce biologicky zajímavých molekul na atomární úrovni. V celé této knížce se tedy budeme zaměřovat na popisnou stránku biochemie.

Hned v úvodu bychom si měli zdůraznit, že snaha o pochopění základních principů stavby molekul nás vede k velmi zjednodušenému pohledu na molekuly. A nejen nás, stejně jak chemik, který kreslí vzorce a montuje z kuliček a trubek modely molekul. Prvním zjednodušením je již to, že ve vzorcích spojení mezi atomy působí síly, které jsou popsány stejnými fyzikálními zákony. Silové působení mezi atomy, které nejsou spojeny kovalentní vazbou, může být dokonce silnější než některé kovalentní vazby.

V sekci 1.1.1 jsme na kovalentní vazbě založili rozlišování mezi konfigurací a konformací. V přírodě ale mezi konfigurací a konformací žádná jasná hranice není – otáčení kulem některé dvojně vazby může být snazší, než otáčení kulem určité jednoduché vazby. Znamená to, že můžeme strukturní vzorce založené na konfiguraci a modely založené na konformaci odhodit jako zbytečné a matoucí? Určitě ne, vzorce a modely jsou velmi užitečné pomůcky, které pomáhají chápát základní pravidla a souvislosti výstavby molekul. Musíme si ale uvědomit, že pomocí nich můžeme základní principy ilustrovat, ne však fyzikálně vysvětlit. Proto bychom se měli v úvodu zamyslet nad fyzikálními zákony, které stavby molekul určují. Tedy nad tím, proč se zkoumaná molekula vyskytuje v určité konformaci.

1.2. KONFORMACE A ENERGIE

můžeme pouze porovnávat energie různých konformací mezi sebou. Molekuly se ani v buňkách ani ve zkoumavce nevyskytují v jediné, energeticky nejvýhodnější konformaci. Energetickým popisem souboru molekul v různých konformacích se zabývá termodynamika. Termodynamika obvykle popisuje soubor molekul ve stavu, kdy jsou jednotlivé konformace v rovnováze. Zabývá se celkovými počty molekul v jednotlivých konformacích, aniž by se starala o to, že každá z molekul stále přechází z jedné konformace do druhé. Z pohledu termodynamiky je nás model struktury obrázek jedné, reprezentativní, nejlépe nejvýhodnější, konformace z mnoha, ve kterých můžeme molekulu nalézt.

Zastoupení konformací v souboru molekul není dáno jen energetickou výhodností jednotlivých konformací. Termodynamicky je také výhodné mít určitou konformaci pestrost. To popisuje termodynamika veličinou entropie. Většina, která zahrnuje oba zmíněné příspěvky, energii jednotlivých konformací i jejich pestrost, se nazývá volná energia.

1.2.2 Síly působící v biomakromolekulách

Energie související se strukturou molekul je téměř výhradně elektrického původu. Z praktických důvodů se často vyjadřují jednak jednotlivé příspěvky k celkové energii. Z těchto příspěvků zmíníme alespoň nejdůležitější.

- **Kovalentní vazby** určují konfiguraci molekuly. Tyto vazby popisují nejvýhodnější prostorové rozdělení elektronů z pohledu jejich vzájemných interakcí a jejich interakcí s jádramy atomů. Popis kovalentních vazeb je poměrně složitý, vyžadující kvantový pohled.

- **Vodíkové vazby** se tvoří mezi vodíkem vázaným velmi polární vazbou na atom zvaný donor (nejčastěji na kyslík, dusík, síru) a atomem zvaným akceptor, který může poskytnout elektrony pro vytvoření vazby (opět nejčastěji kyslík, dusík, síra). Vodíkové vazby se podobají kovalentním vazbám a jejich popis je velmi složitý (vyžadující kvantový přístup). Vodíkové vazby jsou ale slabší než kovalentní a proto se nepovažují za vazby určující konfiguraci. Velmi výrazně však zvyšují určitou konformaci.

- **Sterické odpuzování (repulze)** elektronů se projeví, když se snášíme přibližit atomy tak, že se překrývají oblasti nejpravděpodobnějšího výskytu elektronů. Energie odpuzování klesá s dvánáctou mocninou vzdálenosti.

- **Iontové interakce** představují jednoduché elektrostatické odpuzování stejně nabízených iontů a přitahování opačně nabízených iontů. Energie těchto interakcí klesá s první mocninou vzdálenosti iontů. Síla iontových interakcí je neprůměrná elektrické permitivitě (nebo i dielektrické konstantě prostředí). Ve vodě jsou elektrostatické interakce 80-krát slabší než ve vakuumu a v přítomnosti dalších rozpuštěných iontů jsou stále vzhledem k vodě.

- **Dipolové interakce** jsou síly působící mezi elektrickými dipoly, které vznikají v molekule vytvořením nadbytku kladného náboje v jedné části molekuly a nadbytku záporného náboje v jiné části molekuly. Molekula jako celek je přitom navenek neutrální. Elektrický dipol má každá polární vazba.

---

10 Na rozdíl od entropie je mezně určit entropii v absolutních číslech. Například entropie dokonale uspořádaných molekul je rovna nule.

1.2.3 Celková energie a výsledná konformace

Výhodnost určité konformace je dána nejenom vzájemným energetickým působením mezi jednotlivými částmi makromolekuly, ale také interakcí s okolními molekulami (především s molekulami rozpuštědla) a entropií celé soustavy makromolekula-rozpuštědla. Uvedme si jaké obecné chování můžeme očekávat u biomakromolekuly rozpuštěné ve vodném roztoku iontů solí (tedy v prostředí odpovídajícím cytoplazmě).

1. Atomy, které mohou tvořit vodíkové vazby, se budou snažit rozmístit tak, aby mohlo vzniknout co nejvíce vodíkových vazeb. Ale pozor, vodíkové vazby mohou vzniknout i mezi atomy makromolekuly a rozpuštědla a mezi různými molekulami rozpuštědla navzájem. Dva postranní řetězce biomakromolekuly mohou mířit

- buď ven z makromolekuly, do rozpuštědla, se kterým vytvoří vodíkové vazby,
- nebo dovnitř makromolekuly, kde vytvoří vodíkové vazby mezi sebou navzájem. Molekuly rozpuštědla, které z prvního případu tvoří vodíkové vazby s postranními řetězcem, mohou v tomto případě vytvořit více vodíkových vazeb mezi sebou.

Energie spojená se vznikem vodíkové vazby je poměrně velká (−12 až −38 kJ mol⁻¹), v obou případech proto získáme velký příspěvek k energii jednotlivých konformací. Protože ale jediné, co můžeme o energii říci, je rozdíl mezi jednotlivými konformacemi, a ten může být velmi malý, je většinou těžké odhadnout, jestli budou postranní řetězce tvořit vodíkové vazby mezi sebou, nebo jestli budou otočeny ven do rozpuštědla. Z pohledu entropie se na první pohled zdá, že orientace řetězců ven z makromolekuly je výhodnější, protože takový řetězec má víc volnosti (může se vyskytovat ve více konformacích, než když je součástí pevně strukturovaného jádra biomakromolekuly). Neschle ale zapomenout na to, že toto orientace je naopak nevýhodná z pohledu molekulu rozpuštědla, které musí vytvořit takzvaný solvatační obal – konstrukci kolem postranního řetězce, která je dosti přesně vymezena jeho tvarem. Molekuly vody, které se zapojují do tvorby solvatačního obalu budou mít menší volnosti než molekuly tvořící součást volného rozpuštědla.

2. Ionizované postranní řetězce, které nesou volný elektrický náboj, se budou snažit rozmisťit tak, aby se náboje opačného znaménka co nejvíce přitahovaly a náboje stejného znaménka co nejméně odpuštaly. Dva postranní řetězce, které nesou opačný náboj, mohou mířit

- buď ven z makromolekuly a výhodně interagovat s ionty rozpuštěných solí,
• nebo dovnitř makromolekuly a vytvořit iontovou vazbu mezi sebou navzájem.

Energie iontových interakcí je srovnatelná s energié vodíkových vazeb a stejně jako v případě vodíkových vazeb je těžké odhadnout, který z uvedených případů je výhodnější.

3. Polární postranní řetězce se budou snažit rozmístit tak, aby interakce jejich trvalých dipólů s ionty a okolními dipóly byly co nejvýhodnější. Polární řetězec mohou mřížit:
   • buď ven z makromolekuly a výhodně interagovat s ionty rozpuštěnými solí a s dipóly rozpuštědla,
   • nebo dovnitř makromolekuly a výhodně interagovat s ionizovanými postranními řetězci, s ionty pevně vázanými na makromolekulu, nebo s dipóly ostatních částí makromolekuly.

Energie dipólových interakcí je obecně nižší než energie iontových interakcí, ale elektrických dipólů je v biomakromolekule více než zcela ionizovaných skupin. Stejně jako v předchozích případech je obtížné odhadnout, zda je výhodnější konformace s polárními řetězci mřížit makromolekuly nebo ven do rozpouštědla.

4. Nepolární postranní řetězce biomakromolekuly obklopené vodným roztokem se budou snažit směřovat dovnitř makromolekuly a vytvořit těsný kontakt s ostatními nepolárními řetězci, stejně jako se olej ve vodě shlukuje do kapek. Tento takzvaný hydrofobní efekt nelze uspokojivě vysvětlit z pohledu energetických příspěvků. Výhodně van der Waalsovy interakce by totiž byly silnější, kdyby dipóly indukované v nepolárních postranních řetězcích mohly interagovat s trvalými dipóly rozpouštědla a s rozpouštěnými ionty, ne pouze s indukovanými dipóly ostatních postranních řetězců. Pro hydrofobní efekt je rozhodující entropie rozpouštědla. Vytvoření solvatačních obalů kolem nepolárních řetězců totiž značně omezuje volnost molekul vody.

Na rozdíl od předchozích případů, kdy ionizované, polární a vodíkové vazby tvořící postranní řetězce mohly mřížit dovnitř makromolekuly nebo do rozpouštědla, jsou nepolární řetězce výrazně hydrofobní – jednoznačně se snaží neinteragovat s okolní vodou. Situace by byla opačná, kdyby biomakromolekula nebyla obklopena vodným roztokem, ale nějakou nepolární látkou. Takovému případu se blíží například části proteinů procházejících přes membránu. V oblasti membrány projeví nepolární postranní řetězce svůj lipofilní charakter a ochotně interagují s okolními lipidy. Naopak polární a ionizované postranní řetězce se budou vždy snažit mřížit dovnitř makromolekuly, kde se jim nabízejí energeticky výhodné interakce.

1.2.4 Termodynamika a kinetika

Dosud jsme se zabývali pouze energetickým popisem souboru molekul v rovnováze. Pro studium struktury není ale důležité jen energetické srovnání konformací, ale také znalost rychlosti, se kterou může jedna konformace přecházet v druhou. Popisem rychlostí takových změn konformace se zabývá kinetika. Rychlost přeměny konformací je opět dána energií, ne však energií nejvýhodnějších konformací, ale naopak energií nejméně výhodné konformace, kterou musí molekula projít na cestě od jedné výhodné konformace k jiné.

Srovnání kinetického a termodynamického popisu je ukázáno na obrázku [14.3]. Pro zjednodušení zde popisujeme konformační změny spojenou pouze se změnou jednoho tořenho úhlu. Během rotace kolem vazby l se mění energie molekuly v roztoku. Tenká čára ukazuje průběh vnitřní energie jedné konformace, zatímco tlustá čára popisuje volnou energii souboru molekul v různých konformacích ostatních vazeb. Tlustá čára tedy zahrnuje i příspěvek entropie. Rozdíl mezi údolním na křivce energie určuje pravděpodobnost, že se bude molekula vyskytovat v první konformaci, s postranním řetězecem mřížit do nitra makromolekuly, nebo v druhé konformaci, s postranním řetězcem mřížit do vodného roztoku.
KAPITOLA 1. STRUKTURA V CHEMII A BIOCHEMII

Obrázek 1.7: Srovnání termodynamického a kinetického popisu konformační změny. Vlevo je schematicky znázorněna rotace kolem vazby 1. Z možných konformací způsobených rotací kolem vazby 2 jsou znázorněné tři. Vpravo je energetický popis děje. Tenkou čarou je znázorněna závislost vnitřní energie na rotaci kolem vazby 1, tlustou čarou je znázorněna závislost volné energie na rotaci kolem vazby 1. Energetický rozdíl znázorněný uprostřed grafu energie (černě) udává, jaké bude zastoupení jednotlivých konformací v rovnovážné směsi. Energetický rozdíl vyznačený vlevo (zeleně) udává rychlost přeměny levé konformace na pravou, energetický rozdíl vyznačený vpravo (červeně) udává rychlost zpětné konformační změny.

Naproti tomu energetický rozdíl mezi údolím a vrcholem na křivce, takzvaná aktivační energie, určuje rychlost, se kterou bude přecházet konformace odpovídající danému údolí na konformaci druhou. Čím je tato energetická hradba vyšší, tím bude změna pomalejší. Zatímco u malých molekul jsou aktivační energie zanedbatelné a konformační změny probíhají velmi rychle, u biomakromolekul se setkáváme s energetickými hradbami tak vysokými, že změny konformace jsou téměř nemožné. Tak se může stát, že molekula zůstane po celý svůj život uvězněna v jednom údolí, ačkoli jiné údolí by bylo energeticky ještě výhodnější.

\[12\] Přesněji řečeno, tím bude menší pravděpodobnost, že ke změně dojde.
Kapitola 2

Struktura proteinů

2.1 Proteiny jako příklad biomakromolekul

Existují dva důvody, proč je vhodné probírat struktury proteinů (česky bílkovin) jako první ze struktur biomakromolekul. Prvním důvodem je velký význam proteinů a bezpočet funkcí, které v buňce hrají. Strukturou proteinů se proto zabývá nejvíce biochemiků. Druhý důvod je pedagogický. Na strukturách proteinů dobře vyniknou obecná pravidla, se kterými se budeme později setkávat i u jiných biomakromolekul.

Jedním z těchto pravidel je stavebnicový charakter struktury. Z několika málo součástek si budeme moci sestavit základní strukturní motivy, jejichž další kombinaci získáme i ty nejsložitější struktury. Jiným obecným rysem je hierarchie struktur. To znamená, že strukturní motivy, které můžeme sestavit ze základních dílů biochemické stavebnice, slouží jako stavební díly složitějších strukturních motivů. A tyto složitější struktury jsou zase stavebními kameny ještě složitějších strukturních útvarů a tak dále. Tato hierarchie je u proteinů nejlépe definovaná – biochemici rozlišují primární, sekundární, tertiární strukturu i struktury vyšší.

V této chvíli asi není jasné, co je tak pozoruhodné na zmíněných rysech biomakromolekulárních struktur – stavebnicovém a hierarchickém uspořádání. Následující stránky by měly ukázat, že tyto základní principy jsou klíčem k pochopení stavby biologicky zajímavých molekul. Složitost makromolekuly není dána složitostí základních stavebních jednotek, ale neuvěřitelným množstvím možností, jak můžeme poměrně jednoduché stavební jednotky komбинovat. Živé tvorby nejsou vyjmační tím, z jakých atomů se skládají, ale vynikají způsobem, jak jsou uspořádány atomy v jejich molekulách, molekuly v buňkách, buňky v orgánech. O kráse katedrál se také dozvíme více z papírového modelu, který nese informaci o architektuře, aniž by říkal cokoli o materiálním složení, než z příslušných hromad kamene, písku a vápna, které naopak nenesou informaci o myšlence architekta.

Obecná pravidla stavby biomakromolekul bychom mohli dlouho teoreticky probírat. Mohli bychom se snažit uvěznit tato pravidla do co nejpřesnějších definic, což by se nám stejně nepodařilo. Pojďme se raději podívat na konkrétní příklad biomakromolekulární architektury – na molekuly proteinů.

2.2 Stavební jednotky

Základními stavebními jednotkami proteinů jsou α-l-aminokyseliny. Ačkoli je skutečná syntéza proteinů v buňce velmi složitá, formálně můžeme molekulu proteinu získat spojováním karboxylových a aminových skupin kondenzační reakcí, při které se odstěpuje molekula vody. Vzniklý produkt se nazývá peptid. V peptidech jsou zbytky aminokyselin (aminokyselinová rezidua) spojeny amidovou vazbou, která se často označuje jako vazba peptidová. Při kondenzaci většího počtu aminokyselin vzniká lineární
molekula, jejíž pátí je tvořena atomy zkondenzovaných karboxyllových a aminových skupin a zbytek molekuly aminokyselin tvůr postranní řetězec. Označení peptid a protein jsou ponecháno nejednoznačná a do jisté míry synonymní. V literatuře se můžeme setkat s několika způsobů rozlišování mezi peptidem a proteinem:

- **Peptidy a proteiny se rozlišují podle počtu aminokyselinových zbytků.** Molekuly skládající se ze dvou až deseti aminokyselinových zbytků se označují jako oligopeptidy, molekuly složené z 11 až 100 aminokyselin se označují jako polypeptidy a molekuly obsahující více aminokyselinových zbytků se označují jako proteiny. Přestože je toto dělení oblibené autory učebnic, je zcela umělé a o struktuře nic nevyhodnotí.

- **Jako peptidy se označují kratší řetězce, které snadno přechází z jedné z mnoha možných konformací do jiné a netvoří stabilní trojrozměrné struktury.** Naopak jako proteiny se označují delší řetězce, které mají zřetelnou tendenci vyskytovat se v jedné nejvyhodnější konformaci (nebo v závislosti na vnějších podmínkách měnit konformaci v rámci malého počtu možných konformací) a mají tedy dobře definovanou protonovou strukturu. Je zřejmé, že mezi takto definovanými peptidy a proteiny není jasně vymezená hranice, ale spíše plynulý přechod. Navíc může být část i značné velké molekuly proteinu málo uspořádaná, připomínající flexibilní peptid.

- **Při studiu složitějších molekul, které kromě aminokyselinového řetězce obsahují chemicky odlišnou část, se jako polypeptid označuje samotný peptidový řetězec a jako protein celá složená molekulka.** Jiné názvosloví nazývá peptidovou část apoprotein a celou složenou molekulu holoprotein.

- **Při pohledu na vzorce 20 proteinogenních aminokyselin si jevíme chemickou různost postranních řetězců.** Tato pestrost má zásadní význam pro strukturu i funkci proteinů a vysvětluje, proč jeden typ makromolekuly může mít tak odlišné chemické a fyzikální vlastnosti a hrát tak různé role v životě buněk. Z pohledu interakcí, které určují energetickou výhodnost konformace proteinu, si můžeme aminokyseliny rozdělit na několik typů.

- **Aminokyseliny, které jsou schopny vytvářet iontovou vazbu, zahrnují arginin, lysin, histidin, tyrosin, cystein, kyselinu asparagovou a kyselinu glutamovou. Podmínky, za kterých jsou tyto aminokyseliny ionizovány, jsou dány především kyselostí (hodnotou pH) a elektrickou permitivitou jejich nejbližšího okolí.** Arginin a lysin tvoří kationty v kyselém a neutrálním prostředí, histidin tvoří kationty v kyselém prostředí. Kyselina asparagová a glutamová tvoří anionty v neutrálním a zásaditém prostředí, cystein v zásaditém a tyrosin jen v silné zásaditém prostředí. Konkrétní hodnoty disociační konstanty závisí na elektrické permitivitě okolí. V hydrofobním jádře proteinu budou postranní řetězce hůře ionizovány, takže kationty se budou tvorit až při nižším pH a anionty až při vyšším pH.

---

1 V molekulách proteinů izolovaných z buněk nacházíme více než 20 různých aminokyselin. „Nestandardní“ aminokyselinové zbytky ale v naprosté většině případů vznikají chemickými úpravami již vzniklého polypeptidového řetězce. Tyto úpravy se nazývají posttranslační modifikace a zahrnují zavedení jednoduchých funkčních skupin jako hydroxyl nebo fosfát i navázání velkých molekul jako mastné kyseliny, porfyriny (hem, chlorofyl), oligosacharidy a další.

2 Dekadický logaritmus disociační konstanty s opačným znaménkem je roven pH, při kterém je právě polovina molekul disociována.
2.2. STAVEBNÍ JEDNOTKY

Obrázek 2.1: Vzorce aminokyselin s atomy označením atomů podle doporučení IUPAC-IUB. Purpurově jsou zvýrazněny atomy, které se podílejí na tvorbě vodíkových vazeb. Barvy trojpísmenných a jednopísmenných zkratek aminokyselin rozlišují krátké postranní řetězce (černá), nepolární postranní řetězce (zelená), polární postranní řetězce (purpurová), kyselé (záporně nabité) postranní řetězce (červená) a zásadité (kladně nabité) postranní řetězce (modrá).
KAPITOLA 2. STRUKTURA PROTEINŮ

- Postranní řetězce schopné dipolárních interakcí jsou zároveň schopné tvořit vodíkové vazby. Patří mezi ně všechny aminokyseliny schopné tvořit ionty a navíc serin, threonin, asparagin, glutamin a tryptofan.
- Silný hydrofobní efekt vykazují alifatické postranní řetězce prolinu, leucinu, isoleucinu, valinu a do jisté míry alaninu. Míra hydrofobních a lipofilních vlastností je dána délku alifatického řetězce. Dále můžeme mezi hydrofobní postranní řetězce zařadit aromatické aminokyselinové zbytky tryptofanu, tyrosinu a fenylalaninu a sirné zbytky cysteinu a methioninu.
- Glycin a do jisté míry i alanin, serin a threonin můžeme považovat za zvláštní skupinu ,,malých“ postranních řetězců. Vzhledem k jejich velikosti je zřejmo když ovšem hydrofilní ani hydrofobní charakter. Tím, že zabírají malý prostor poskytují zvýšenou volnost okolním postranním řetězcům.
- Prolin má netypickou konfiguraci (cyklický postranní řetězec), která významně omezuje možné konformační proteínové pátere.
- Cystein je vyjmačný v tom, že může tvořit kovalentní disulfidovou vazbu s postranním řetězcem cysteinu v jiné části proteinu. To je velký zásah do topologie proteinové pátěře, který má zásadní vliv na konformační proteinu.

Stejně jako je důležitá rozdílnost postranních řetězců, tak je významná neměnnost uspořádání na prvních dvou ulících. Všechny aminokyseliny můžeme považovat za deriváty alaninu, všechny zachovávají stejnou konfiguraci

\[ \alpha \]-uhlíku (přestože zrcadlová konfigurace

\[ \alpha \]-uhlíku by byla z chemického pohledu stejně dobrá možná). Tato jednotnost základních stavebních dílů umožňuje stavebnicové uspořádání popsané v sekci 2.1. Stejná konfigurace

\[ \alpha \]-uhlíku otvírá dveře téměř neomezenému kombinování aminokyselin při zachování pravidelných struktur. Glycin, který je achirální, a prolin, který se liší konfigurací na dusíku, jsou výjimky, které potvrzují pravidlo a dávají ještě více vyniknout logice stavebnicového uspořádání.

2.3 Torzní úhly

Jak jsme diskutovali v sekci 1.1.2 konformaci biomakromolekuly jednoznačně určují torzní úhly. Podívejme se proto na torzní úhly, se kterými se setkáváme v proteinech (obrázek 2.2).

Konformace peptidové pátěře aminokyseliny je určena třemi torzními úhly, \( \phi \), \( \psi \) a \( \omega \). Čtvrtecky atomů, které torzní úhel definují (A–B–C–D na stránce 7) jsou pro uvedené úhly následující. V reziduu číslo \( i \) je úhel \( \phi \) je definován atomy \( C_\alpha^{i-1}, N_i, C_\alpha^i, C_i' \), úhel \( \psi \) atomy \( N_i, C_\alpha^i, C_i', N_{i+1} \) a úhel \( \omega \) atomy \( C_\alpha^i, C_i', N_{i+1}, C_{i+1} \).

Pro toto peptidová vazba se chová částečně jako dvojná vazba, úhel \( \phi \) nabývá pro většinu aminokyselin hodnot blízkých 180° (konformace trans). Výjimkou je prolin, který se vyskytuje i v cis konformaci (0°). Konformace peptidové pátěře je tak definována především hodnotami úhlů \( \phi \) a \( \psi \). Peptidové konformery je proto možno graficky znázornit jako body v grafu, jehož osy tvoří hodnoty úhlů \( \phi \) a \( \psi \). Takový graf se nazývá Ramachandranův diagram. Běžné konformace peptidové pátěře budou

\[ \text{Vodíkové vazby je schopna tvořit každá peptidová vazba, zde je řeč Pouze o postranných řetězcích.} \]
\[ \text{Aromatické aminokyseliny, zařazené do skupinu hydrofobních postranních řetězců, mohou krom van der Waalsovských sil interagovat prostřednictvím elektronů v delokalizovaných π-orbitalích. Totěž platí o d-orbitalích síry. Je-li interakce svým popisem připomínající kovalentní vazbu, ale mnohonásobné slabší. V našem zjednodušeném výčtu sil působících v biomakromolekulách jsme si je neuvědomí.} \]
\[ \text{Na otázku proč se v proteinech vyskytují právě L-aminokyseliny dosud neznáme uspokojivou odpověď.} \]
\[ \text{Pokud stejnou strukturu můžeme postavit z většího počtu možných kombinací základních stavebních prvků, je více pravděpodobné, že se s takovou strukturou v buře setkáme.} \]
\[ \text{Výjádřeno pojmy molekulových orbitálů, o kterých bude řeč v kapitole 6 má Amidový dusík peptidové vazby má asi 40% podíl hybridizace sp^2.} \]
Obrázek 2.2: Definice torzních úhlů hlavního řetězce (nahoře) a torzního úhlu \( \chi^1 \) postranních řetězců (dole) proteinů. Vlevo dole je definován \( \chi^1 \) nevětvených řetězců, uprostřed threoninu a isoleucinu, vpravo valinu.
Obrázek 2.3: Ramachandranův diagram s vyzačeným názvoslovím peptidových konformací. Modré jsou označeny konformace výhodnější pro t-aminokyseliny. Červené jsou označeny konformace výhodnější pro d-aminokyseliny, které se běžně v geneticky kódovaných proteinech nevyskytují.

diskutovány později. Na obrázku je uvedena obecná konvence názvosloví peptidových konformerů. Pro chirální aminokyseliny (všechny běžné kromě glyciny) existují většinou dvojice (zrcadlové obrazy) konformerů, z nichž jeden (označený modré) je výhodný pro t-aminokyseliny (tu běžně vyskytuje v proteinách), druhý (označený červeně) by byl výhodnější pro d-aminokyseliny:

Konformace postranních řetězců je dána torzními úhly, které se značí písmenem $\chi$, a jejichž počet je pochopitelně různý pro různé aminokyseliny. Úhel $\chi$, definovaný atomy N, C$\alpha$, C$\beta$, X se nachází nejčastěji v oblasti gauche($-$) nebo trans. Konformace gauche($+$) je nevýhodná stericky kvůli přiblížení peptidové vazby, častěji se vyskytuje u seriny a threoninu, kde se stejných důvodů umožňuje tvorbu vodíkové vazby mezi kyslíkem postranního řetězce a peptidovou vazbou. Valin a isoleucin, které mají dva $\gamma$ uhliky, se vyskytují převážně v konformaci, ve které není žádný $\gamma$ uhlík orientován gauche($+$) vzhledem k hlavnímu řetězci (pozor na záložnost doporučeného názvosloví organické chemie: u isoleucinu se taková konformace označuje gauche($-$), ale u valinu trans).

### 2.4 Primární struktura

Poté, co jsme prošli základní definice, se již můžeme podívat na vlastní struktury. Jak jsme se již zmínili v sekci 2.3, zavedli biochemici hierarchickou posloupnost struktur. Nejúrovní úrovni této hierarchie je primární struktura.

Primární struktura je pořadí (sekvence) aminokyselin v polypeptidu. Protože konfigurace jednotlivých aminokyselin známe, poskytuje nám primární struktura většinou informaci o konfiguraci proteinu (s jednou důležitou výjimkou – z primární struktury stále nevímeme, které cysteiny jsou spojeny disulfidovou vazbou). Naopak se z primární struktury nedovíme nic o konformaci proteinu.

Podle konvence začíná sekvence N-termínální aminokyselinou (obsahuje volnou aminoskupinu) a končí C-termínální aminokyselinou (obsahuje volnou karboxyskupinu). Tímto směrem také probíhá syntéza proteinů v buňce. Pořadí aminokyselin je dáno genetickým kódem, který je zapsán v molekule deoxyribonukleové kyseliny (DNA). Každým třetím monomerním jednotkám DNA (nukleotidům), které
2.5. SEKUNDÁRNÍ STRUKTURA

Zatímco definice primární struktury byla celkem jednoznačná, u vyšších struktur se již definice různých autů často liší. My budeme mít pod pojmem sekundární struktura na mysli pravdelnou, libovolnou opakovatelnou konformaci motive polypeptidové pátěří. U proteinů se setkáváme se dvěma základními typy sekundárních struktur.

2.5.1 Struktura alfa

Struktura α je pravotočivá šroubovice neboli helix (obrázek 2.4, pokud chcete skripta v elektronické formě a jste připojeni na internet, můžete si model α šroubovice prohlédnout na svém počítači). Víšneme si, že poskládáním aminokyselin do vyšší struktury vznikl nový tvar, který je sám o sobě chirální. Tato chiralita nijak nesouvisí s chiralitou aminokyselin. Kdybychom postavili z glycinu α-chiralitu nijak nesouvisí s chiralitou aminokyselin. Kdybychom postavili si, že poskládáním aminokyselin do vyšší struktury vznikl nový tvar, který je sám o sobě chirální. Tato struktura α je stabilizována vodíkovými vazbami, které jsou orientovány tématř rovnoběžně s osou helixu a tedy i s jeho výsledným dipolovým momentem (nejvýhodnější.

Rozdíl mezi levotočivou a pravotočivou šroubovicí je v konformaci stavebních dílů, nikoli v jejich konfiguraci. Na úrovni sekundárních struktur můžeme takřikává, abychom mohli postavit z glycinu pravotočivou šroubovicí, která by byla zrcadlovým obrazem α šroubovice. Ale pozor, pokud sestavíme levotočivou šroubovicí z jiných aminokyselin, zrcadlový obraz α šroubovice nezískáme. Vzájemnými zrcadlovými obrazy jsou totiž pravotočivé šroubovice postavené z l aminokyselin (α-šroubovice) a levotočivé šroubovice postavené z D aminokyselin.

Úhly φ a ψ odpovídají v α-struktuřích konformaci gauche (ideální α šroubovice je definována hodnotami φ = −57° a ψ = −47°). α-Šroubovice je stabilizována vodíkovými vazbami, které jsou orientovány tématř rovnoběžně s osou helixu a tedy i s jeho výsledným dipolovým momentem (nejvýhodnější.

8 Tato univerzální definice genetického kódu má své výjimky. Ponekud jednodušší rozporoznávání nukleotidů například v mitochondriálních větě k drobným odchylkám kodovacích pravidel. Jiným příkladem výjimky je využití některých kodonů pro zabudování selenu do molekuly cysteinu (místo síry).
9 V současné době je již pro několik organismů (včetně člověka) známa téměř veškerá genetická informace. U těchto organismů známe tedy sekvence všech nukleotidů genové DNA. Stáčí nám tedy určit, kde jsou geny jednotlivých proteinů začínají a končí (a také, kde jsou přerušeny někódujícími úseky) k tomu, abychom znali sekvence všech proteinů, alespoň v okamžiku, kdy jsou v buněčné syntetizovány (v mnoha případech se část proteinu odštěpí dříve, než začne protein plnit svoji úlohu).
10 Sekundární struktury lze často (alespoň formálně matematicky) popsat jako šroubovice.
11 Pravidla IUPAC-IUB definují nejen konfiguraci cítrštehų, používanou na úrovni atomů, ale i konfiguraci šroubovcových útvarů. Pravotočivé útvary se označují symbolem P, levotočivé symbolem M.
12 Stále samozřejmě mluvíme o konformaci peptidové pátěře, nikoli postranních řetězcích.
KAPITOLA 2. STRUKTURA PROTEINŮ

Obrázek 2.4: Model α-šroubovice.

Obrázek 2.5: Model šroubovice 3_{10}.

orientace). Protože vodíkové vazby se tvoří mezi amidovými vodíky a karbonylovými kyslíky peptidových vazeb, jsou aminokyseliny v sousedních závitech šroubovice navzájem posunuty zhruba o jednu třetinu. Na jeden závit tak připadá 3,6 (přibližně 3\(\frac{2}{3}\)) aminokyselin.

V proteinové chemii se používá speciální názvosloví šroubovic. Každá šroubovice je popsaná dvěma čísly. První číslo udává počet aminokyselin na jeden závit, u α-šroubovice tedy 3,6. Druhé číslo je definované složitěji. Je založeno na tom, že sousední vodíkové vazby uzavírají ve šroubovicích pravidelné smyčky atomů spojených kovalentními a vodíkovými vazbami. V α-helixech je každá taková smyčka tvořena 13 atomy. Podle názvosloví odvozeného od počtu reziduů na závit a od počtu atomů ve smyčkách by se tedy α-šroubovice označovala jako helix 3\(\frac{2}{3}\).

Struktura α-šroubovice je energeticky nejvýhodnější stočená forma peptidového řetězce. Těsnější vinutí šroubovice (α \(\frac{2}{3}\) reziduň) tvoří strukturu helixu 3_{10} (3 aminokyseliny na závit, 10 atomů ve smyčkách, viz obrázek 2.5.1 a model). Tato šroubovice je méně výhodná z pohledu dipolových interakcí (vodíkové vazby jsou orientovány šikmo k ose helixu) a sterického bránění (postranní řetězce jsou v zákrytu). S touto strukturou se setkáváme zejména na C-koncích α-helixů, kde již karbonyly nemají k dispozici amidové vodíky v dalším závitu, se kterými by mohly tvořit stabilizující vodíkové vazby. Je tedy výhodnější šroubovicí „utáhnout“ do helixu 3_{10}, který umožňuje vytvořit více vodíkových vazeb při stejném počtu aminokyselin.

Pokud je šroubovice vinuta naopak o \(\frac{2}{3}\) reziduň volněji než α-helix, tvoří tzv. π-šroubovicí (helix 4.4_{16}, viz obrázek 2.6 a model). Vodíkové vazby jsou v π-helixu sice orientovány rovnoběžně s osou, energeticky je ale π-helix nevýhodný a ve strukturách proteinů velmi vzácný. Důvodem je tvorba dutiny uvnitř π-šroubovice (tím přicházíme o výhodné van der Waalsovy interakce) a sterické pnutí vedoucí k rozevření úhlu N–Cα–C′.

Všechny uvedené šroubovice se nacházejí na Ramachandranově diagramu v oblasti α konformací.
2.5. SEKUNDÁRNÍ STRUKTURA

Obrázek 2.6: Model π-šroubovice.

(gauche(−), gauche(−)). Na obrázku 2.5[1] je tato oblast α povolena pro L-aminokyselinou označena černě. Vyjímečnou aminokyselinou je prolín[13], který se uprostřed pravotočivé α-šroubovice nevyskytuje ze sterických důvodů (kolize s karbonylem předchozí peptidové vazby).

Zrcadlový obraz pravotočivé α-šroubovice odpovídá levotočivé šroubovicí s konformací gauche(+), gauche(+), gauche(+), gauche(−) (červená oblast). V této konformaci se často nalézají achirální glyciny, občas v ní nalezneme asparagin, pro který je nevyhodnost této konformace kompenzována možností tvorby vodíkových vazeb mezi páteří a postranním řetězcem.

2.5.2 Struktura beta


Zcela natažená trans,trans konformace peptidové páteře by odpovídala strukturní skládanému listu, ve kterém by následující aminokyseliny byly navzájem otočeny o 180° (formálně by se dala tato struktura považovat za šroubovicí, ve které na jeden závit případnou dvě rezidua). V proteinech se zcela natažená konformace nevyskytuje, hodnoty úhlu φ a ψ jsou menší než 180° (φ = −139°, ψ = +135° v antiparalelních a φ = −119°, ψ = +113° v paralelních β-listech). Peptidová páteř v β-strukturách je tedy mírně pravotočivá (i β-struktury jsou vlastně pravotočivé šroubovice) a v důsledku toho jsou β-listy levotočivé zkrucené. Skutečnost, že pravotočivost v rámci jednotlivého peptidového řetězce způsobuje levotočivost vzájemného uspořádání různých řetězců, jistě působí na první dojem záhadné. Všelijaké zřejmější, když si dobře prohlédneme schema na obrázku 2.9.

Oblast povolených β-konformací v Ramachandranově diagramu je poměrně rozsáhlá (obrázek 2.10).

Obrázek 2.7: Ramachandranův diagram s černě vyznačenou oblastí konformací aminokyselin v šroubovicích proteinů (strukturách α). Červeně je vyznačena oblast levotočivých šroubovic, které by byly výhodnější pro D-aminokyseliny.

Obrázek 2.8: Model dvou vláken antiparalelního (nahoře) a paralelního (dole) β-listu.
2.6. TERCÍÁRNÍ STRUKTURA

2.6.1 Spojovací prvky

Molekuly některých proteinů jsou tvořeny jedinou monotónně se opakující sekundární strukturou (například keratin). Daleko častěji se ale prvky sekundárních struktur stavebnicově skládají do přibližně globulárních molekul. Takové skládání polypeptidových řetězců se nazývá terciární strukturou. Aby byla molekulová stavebnice proteinů použitelná, musí krom pravidelných sekundárních struktur (α a β) obsahovat i spojovací díly.

Stejně jako α-helixy nejsou β-strukturny výhodné pro prolin, kde cyklická struktura postranního řetězce vyžaduje úhel $\phi \approx -60^\circ$.

Obrázek 2.9: Schematické znázornění pravotočivé orientace vodíkových vazeb (vyznačeny černě a čárkovaně) a levotočivého zkruhu β-listů (peptidová páteř je vyznačena červeně). Vlevo pohled shora, vpravo pohled z boku.

Obrázek 2.10: Ramachandranův diagram s vyznačenou oblastí konformací aminokyselin v β-listech proteinů.

2.6 Terciární struktura

---

14 Sekundární struktury se ve vláknitých (fibrilárních) proteinech nevyskytují izolovaně, ale splétají se do vyšších struktur, srovnatelných s kvartérními strukturami globulárních proteinů (viz sekc. 2.7). Vytvoření takových vyšších struktur je možné pouze tehdy, když do sebe sekundární struktury přesně zapadají. Takové přesné uspořádání většinou vyžaduje určitou sekvenci aminokyselin, která se ve fibrilárních proteinech mnohokrát opakuje. Vidíme, že zde struktura není nezvislá na sekvenci, příčinu ale nenajdeme na úrovni struktur sekundárních, ale vyšších.
Spojovacích prvků se ve strukturách proteinů vyskytuje celá řada. Nejjednodušší jsou takzvané 
*ohyby* (*turns*). Jediná aminokyselina v $\gamma$ konformaci tvoří *inverzní $\gamma$-ohyb* (*obyčejný* $\gamma$-ohyb je pro $\alpha$-aminokyselinu stericky nevýhodný).

Tři takzvané *reverzní ohyby* (značené I, II, III) se skládají ze dvou aminokyselin a jsou stabilizované vodíkovým můstkem mezi vodíkem peptidové vazby s následující aminokyselinou a karbonylovým kyslíkem peptidové vazby s předcházející aminokyselinou. Ohyb III je vlastně výsek z výše popisované $3_{10}$-šroubovice. Reverzní ohyby jsou poměrně stabilní a často se vyskytují na povrchu proteinů.

Zrcadlové obrazy reverzních ohybů jsou označovány I', II' a III'. Ramachandranovy diagramy zmíněných jednoduchých ohybů jsou uvedeny na obrázku 2.11.

Základní hierarchie struktur (primární, sekundární, terciární) je pro detailní popis struktur proteinů příliš hrubá. Proto se v posledních letech objevily v literatuře jemnější stupně v rámci terciární struktury. Názvosloví takového jemnějšího dělení není ale dosud ustálené a často vychází z terminologie různých počítačových programů, které se snaží struktury proteinů analyzovat. Zkusme se tedy popsat vznikající složitost prostorového uspořádání atomů pomocí pojmů běžných ve strukturní biochemii, s tím, že se soustředíme spíš na podstatu věci, než na názvoslovná úskalí.

Ohyby (příklady některých byly zmíněny výše) spolu s úseků sekundárních struktur $\alpha$ a $\beta$ tvoří vyšší strukturní prvky, takzvané *supersekundární motivy*. Jejich kombinace vznikají *foldy*, které dále tvoří *domény* (struktury s funkcí autonomní) a z domén je složena trojrozměrná struktura celého polypeptidového řetězce. Následující přehled *tříd proteinů* ilustruje tuto hierarchii proteinové architektury na strukturách skládaných z různých stavebnicových prvků.

### 2.6.2 Alfa-proteiny

Proteiny tvořené $\alpha$-šroubovicemi označujeme jako *$\alpha$-proteiny*. Nejjednodušším supersekundárním motivem $\alpha$-proteinů je dvojice antiparalelních $\alpha$-šroubovic spojených krátkým ohybem (v anglické literatuře označovaná *helix-turn-helix*). Typické jsou ohyby skládající se ze dvou, tří, nebo čtyř aminokyselin. Ramachandranovy diagramy typických $\alpha$-ohybů jsou uvedeny na obrázku 2.12.

Obrázek 2.11: Ramachandranův diagram s vyobrazenými konformacemi reverzních ohybů. Šipka naznačuje směr od první ke druhé aminokyselině ohybu, v ohybu III a III' mají obě aminokyseliny stejnou konformaci.

![Ramachandranův diagram](image-url)
2.6. TERCÍÁRNÍ STRUKTURA

Obrázek 2.12: Ramachandranův diagram s vyznačenými konformacemi α-ohybů tvořených dvěma (černě v levém diagramu), třemi (červeně v levém diagramu), nebo čtyřmi (dvě možnosti jsou uvedeny různými barvami v pravém diagramu) aminokyselinami. Sípky ukazují od první ke druhé a od druhé ke třetí aminokyselině.

Jiným příkladem supersekundárního motivu je „zlomený helix“ (α-corner), ve kterém je pravidelná α-šroubovice přerušena trojicí aminokyselin ve neobykle konformaci (obrázek [2.13]).


Obrázek 2.13: Model zlomeného helixu (α-corner).

Obrázek 2.14: Supersekundární motiv helix-turn-helix a fold svazku tvořeného čtyřmi šroubovicemi. Vpravo je propojení helixů ohyby znázorněno schematicky.

Obrázek 2.15: Možné topologie svazků čtyř helixů znázorněné schematicky podle obrázku 2.14. Šipky ukazují směr polypeptidového řetězce, tlustou čárou jsou kresleny kroužky natočené k nám N-koncem.
2.6. TERCÍÁRNÍ STRUKTURA

polární dutinu, která spojuje vodná prostředí na obou stranách membrány (například cytoplasmu s mezibuněčnou tekutinou).

Ve většině proteinů nejsou helixy uspořádány rovnoběžně do svazků. Jednotlivé šroubovice navzájem svrají určitý úhel. Logiku takového uspořádání můžeme opět hledat v rozložení postranních řetězců podél heliky. V případě šroubovice 310 tvoří každý čtvrť (například první, pátý, devátý, atd.) postranní řetězec jakýsi hřebínek na boku šroubovice. V případě α-helixu není každá čtvrť α-aminokyselina nad sebou, ale jsou pootočeny vždy o 33,7°. V důsledku toho není spojnictví postranních řetězců lišících se o tři aminokyseliny rovnoběžná s osou heliky (jak tomu bylo u šroubovice 310), ale svrás s ní úhel asi 26°. Jednodušeji řečeno, náš hřebínek postranních řetězců se vine kolem šroubovice. Dvě šroubovice spolu často interagují tak, že do sebe hřebínce postranních řetězců zapadají jako dvě poloviny zdrhovadla (zipu). Protože se hřebíinky vinou kolem šroubovice, musí být osy heliků vzájemně otočeny o −52° (dvojnásobek změněných 26°), aby do sebe hřebíinky ideálně zapadly (pravá část obrázku 2.16). Proto velice často nacházíme ve strukturách proteinů šroubovice svrající spolu úhel kolem mínus padesáti stupňů. Kromě hřebínu tvoreného každou čtvrtou aminokyselinou (ťichému mu i +4-hřebínek) můžeme nalézt i pro aminokyseliny lišící se o tři rezidua (i +3-hřebínek) nebo pro sousední aminokyseliny (i +1-hřebínek). Pokud má i +4-hřebínek zapadat do i +3-hřebínu nebo i +1-hřebínu, musí šroubovice svříhat úhel +23° respektive −105°.


Je zajímavé, že foldy složené z více než šesti šroubovic už do pravidelného mnohostěnů uspořádány nejsou. V těchto foldech jsou jeden nebo dva heliky umístěny uvnitř mnohostěnu nebo svazku šroubovic.
Obrázek 2.17: Ideální uspořádání šroubovic v hranách pomyslných mnohostěnů. Umístění šroubovic jsou vyznačena barevně.
2.6. TERCIÁRNÍ STRUKTURA

Takové centrální helixy obsahují především nepolární postranní řetězce.

Půjdeme-li v hierarchii struktur ještě výše, dospějeme na úroveň domény. Tímto termínem se označují uspořádané úseky polypeptidového řetězce, které mají nejen pravidelnou strukturu, ale často i svou dobře definovanou funkci (například schopnost specificky vázat jinou molekulu, nebo katalyzovat enzymovou reakci). Od podjednotek oligomerních proteinů, o kterých se zmíníme v sekci 2.7, se domény vlastně liší jen tím, že jsou součástí jediného peptidového řetězce. Příkladem α-domény mohou být dva čtyřhelixové svazky.

2.6.3 Beta-proteiny

Proteiny tvořené skládanými listy bývají nazývány β-proteiny. Obdobou supersekundárního motivu helix-turn-helix, který jsme si uvedli pro α-proteiny, je u β-proteinů takzvaná β-vlášenka (β-hairpin), která je součástí struktury zobrazené na obrázku 2.18 (vlášenka je v pravé horní části struktury).

Název je inspirován tvarovým motivem. V β-hairpinch jsou dvě antiparalelní β-vláčky spojena krátkým ohybem (takzvaný β-turn), který může být tvořen dvěma, třemi, nebo čtyřmi aminokyselinami. V případě dvouaminokyselinových β-ohybů jde nejčastěji o výše zmíněné typy I' a II'. Ramachandranovy diagramy tří- a čtyřaminokyselinových β-ohybů jsou uvedeny v levé části obrázku 2.19.

Spodní část struktury na obrázku 2.18 je β-obdobou ,,zlomeného helixu“ a nazývá se β-corner. Pravá část obrázku 2.19 znázorňuje konformace glycinu na vnitřním vláknu β-corneru, kde dochází ke zlomu a aminokyselině, kterou je přerušeno párováním β-listu na vnější straně zlomu.

Kombinací supersekundárních struktur vznikají β-foldy, stejně jak jsme si to popsali u α-proteiny. Skládáním většího počtu antiparalelních β-vláken vznikají β-meandry (horní část obrázku 2.20). Možných propojení β-vláken je samozřejmě více. Jako příklad β-folds s odlišnou topologií můžeme uvést motivy řeckého klíče (Greek key) a „rolády“ (jellyroll), ukázané ve spodní části obrázku 2.20.

Na rozdíl od jasně definovaných šroubovic jsou β-listy mnohem volnější, jak ukazuje mnohem větší konformační volnost v Ramachandranově diagramu. Struktura β-proteinů je jen vyjímečně tvořena
Kapitola 2. Struktura proteinů

Obrázek 2.19: Ramachandranův diagram s vyznačením konformací tří nebo čtyř aminokyselin v ohýbu β-vlášenky (vevo) a konformaci glyciny, na kterém dochází ke zlomu v β-corneru (vpravo). Význam šipek je podobný jako na obrázku 2.12.


Jiným typickým foldem β-proteinů je takzvaný β-soudek (β-barrel), který umožňuje β-proteinům vytvořit příbližně kulový útvar, s nepolárními postranními řetězci orientovanými dovnitř a polárními ven (obrázek 2.23). Na rozdíl od β-sandwiche svárají protilehlá vlákna v β-soudcích téměř pravý úhel. Vláčka tvořící soudek neleží rovnoběžně s osou soudku, ale svárají s ní určitý úhel. Tento úhel je u různých struktur různý, stejně jako počet vláken tvořících soudek (od čtyř po dvacet). Jen některé kombinace ale umožňují vypnout soudek postranními řetězcemi a vytvořit tak nepolární dutinu.

V bohatství foldech β-proteinů můžeme narazit i na další uspořádání β-listů. Jednotlivé listy mohou být uspořádány jako lopatky vrtule či turbiny, mluvíme pak o foldu β-propeller. Listy mohou tvořit stěny trojbokého hranolu (β-prism), s vláknky orientovanými kolmo k podstavce nebo rovnoběžně s podstavou. Struktury β-sandwiche se podobá takzvaný β-helix, ve kterém jsou protilehlá vlákna téměř rovnoběžná a velmi málo zkroucená. Celé listy pak jsou měrně štrobucivě zkrouceny (pravotěsné nebo levotěsné).

Foldy β-proteinů nám ukazují, jak lze s rozdílnou architekturou dosáhnout stejněho cíle, který jsem se zabývali v případě α-proteinů. Tímto cílem je vytvoření stabilní struktury s možností oddělení vodního a nevodního prostředí. I u β-proteinů jsou z foldů stavěny domény, které dále tvoří celkovou

Obrázek 2.20: Nahoře ukázka tvorby β-meandrů kombinací β-vlásenek, dole motivy řeckého klíče (vlevo) a „rolády“.
KAPITOLA 2. STRUKTURA PROTEINŮ

Obrázek 2.21: Vlevo pravotočivý β-β-β motiv (se schematickým znázorněním topologie), vpravo uspořádání postranních řetězců mezi listy β-sandwiche (ukázána jsou jen dvě vlákna v každém listu).

Obrázek 2.23: Rozvinutý list tvořící $\beta$-soudek (nahoře) a schematický nákres soudku (vlevo dole). Pohled dovnitř soudku (vpravo dole) ukazuje uspořádání nepolárních postranních řetězců uvnitř soudku. Postranní řetězce směřující před rovínou a za rovínou nákresu (po svinutí dovnitř soudku a ven) jsou odlíšeny plnou respektive tečkovanou čaron obrysu kroužku. Vodíkové vazby jsou vyznačeny čárkovaně.
terciární strukturu.

2.6.4 Alfa+beta-proteiny

Mnohé proteiny obsahují α- i β-foldy, které jsou spojeny až na úrovni vyšších strukturních celků. Takové proteiny, na jejichž stavbě se podílejí oddělené α-motivy a β-motivy označujeme zkratkou α + β.

2.6.5 Alfa/beta-proteiny

Kromě α+β-proteínů s oddělenými částmi tvořenými buď α nebo β strukturami se setkáváme s proteiny, v nichž se kombinují α a β struktury již na úrovni supersekundárních motivů. Pro tyto proteiny se používá označení α/β. Jako příklad supersekundárního motivu tvořeného oběma typy sekundárních struktur můžeme uvést β-α-β motiv (obrázek 2.24 nahoře). Jde o dvě paralelní β-vlákná, mezi která je vložen opačně orientovaný α-helix. Stejně jako v případě β-β-β motivu se v proteinech vyskytuje výhodnější pravotočivé uspořádání.

Pokračujeme-li v kombinování prvků sekundárních struktur α a β, získáme β-α-β-α-β strukturu, která se jinak nazývá Rossmanův fold a tvoří vazební místa pro nukleotidy (obrázek 2.24 nahoře). Popsané β-α-β motivy mohou být uspořádány do vyšších struktur podobně jako β-listy. Na obrázku 2.24 dole je příklad domény tvořené dvěma Rossmanovými foldy.


2.6.6 Proteiny bez sekundární struktury

Ve velmi vzácnych případech neobsahuje architektura proteínu žádné prvky pravidelné sekundární struktury. Jde o malé proteiny, jejichž struktury jsou stabilizovány disulfidovými vazbami (agglutinin) nebo koordinačními vazbami s navázaným kovovým iontem (například ferredoxin).

2.7 Kvarterní struktura a vyšší struktury

Polypeptidy tvořící terciární struktury mohou dále agregovat a tvořit oligomerní molekuly. Architektura takových komplexů se označuje jako kvarterní struktura. Oligomeru se mohou skládat buď z několika stejných podjednotek (homooligomery) nebo z podjednotek odlišných (heterooligomery). Na rozdíl od domén tvoří podjednotky samostatné polypeptidové řetězce, které mohou být vzájemně vázány kovalentními disulfidovými vazbami mezi cysteiny jednotlivých podjednotek, nebo různými nekovalentními interakcemi, které jsme diskutovali v sekci 1.2.2. Je také časté, že funkce proteinů souvisí s opakovanou disociací a zpětnou asociací podjednotek.

Velká část proteínu obsahuje jako součást nepeptidovou molekulu. Může jít o jednoduché kovové ionty nebo o složité organické molekuly, jako například vitaminy, chlorofyl, hem, nukleotidy, oligosacharidy, mastné kyseliny. Nepeptidová část může být k peptidovému řetězci vázána kovalentně (v enzymologii se takto vázaná molekula nazývá prostetická skupina) nebo nekovalentně.
Obrázek 2.24: Vlevo nahoře pravotočivý $\beta$-$\alpha$-$\beta$ motiv (se schematickým znázorněním topologie), vpravo nahoře schematické znázornění topologie Rossmanova foldu, dole schematické znázornění topologie domény tvořené dvěma Rossmanovými foldy.

Obrázek 2.25: Schematické znázornění topologie TIM-barrelu s vyznačenou polohou postranních řetězců vyplňujících soudek (vlevo) a rozvinutý list tvořící TIM-barrel (vpravo).
Proteiny se mohou dále shlukovat do nadmolekulárních útvarů (například ribozomů), ve kterých často interaguji s molekulami větším a složitějším, než jsou samy – jde zejména o oligosacharidy a nukleové kyseliny. Konečně se mohou proteiny podílet se na stavbě buněčných membrán, organel a jiných biologických útvarů, které svými rozměry přesahují hranice mikrosvěta, jak jsme je zvyklí chápat v chemii i biochemii.
Kapitola 3

Struktura nukleových kyselin

3.1 Stavební jednotky

V sekci 2.2 jsme viděli, že „abeceda proteinů“ má dvacet písmen, kterými je dvacet geneticky kódovaných proteinogenních aminokyselin. „Abeceda“ nukleových kyselin je mnohem jednoduší, tvoří ji pouze čtyři písmena genetického kódu, čtyři nukleotidy. Chemická struktura stavebních jednotek je ale složitější, než v případě proteinů. Každý nukleotid se skládá z pětiuhlíkatého monosacharidu, na který je navážána esterovou vazbou kyselina fosforečná a N-glykosidickou vazbou aromatický heterocyklus, zvaný báze (obrázek 3.1).

Atomy nukleotidů se číslí podle pokynů IUPAC-IUB, které se liší od standardního chemického názvosloví. Sacharidová a aromatická část nukleotidu se číslí zvláště, proto se pro odlišení čísla atomů sacharidu označují čárkou.

Struktury stavebních jednotek ribonukleové kyseliny (RNA) a deoxyribonukleové kyseliny (DNA) se liší ve dvou detailech. Jak již název napovídá, první distinkce se týká sacharidu. V případě RNA je jím ribosa, zatímco v molekule DNA se vyskytuje deoxyribose, která postrádá hydroxylovou skupinu na uhlíku C2'. Tato zdánlivá drobnost je příčinou velkých rozdílů ve struktuře a funkci RNA a DNA.

Zbytek kyseliny fosforečné je navážán v poloze 5' ribosy či deoxyribosy. Tak jako peptidový řetězec vznikal kondenzací karboxylové a aminové skupiny za vzniku amidové vazby, řetězec nukleové kyseliny vzniká kondenzací fosfátové skupiny vázané na uhlík C5' s hydroxyuskupinou vázanou na C3' následujícího nukleotidu. Kyselina fosforečná je tedy v polymukleotidovém řetězci vázána dvěma esterovými vazbami. Ribosy či deoxyribozy spojené přes fosfát tvoří hlavní řetězec nukleových kyselin (nazývaný cukr-fosfatová páteř), který je obdobou peptidové páteře proteinů.


Ve struktuře jedné z bází, pyrimidinu hydroxylovaného v polohách 2 a 4, je druhý chemický rozdíl mezi RNA a DNA. Zatímco v RNA se vyskytuje přímo 2,4-dihydroxypyrimidin (nazývá se uracil), v DNA nacházejího se jednotlivé methylou skupinou na uhlíku C5 (označovaný jako thymín). Prvním písmenem genetické abecedy je tedy U (=uracil, v RNA) nebo T(=thymín, v DNA). Druhou pyrimidinovou bází je cytosín (C). Všimněme si, jak se liší struktury těchto bází vzhledem k možnosti tvorby vodíkových vazeb. Uracil a thymín mají donor v poloze 3 a akceptory v polohách 2 a 4. U cytosinu jsou v důsledku náhrady hydroxylovy uhlíku a methyloskupinu role donoru a akceptoru v polohách 3 a 4

1 Uvědomme si, že v deoxyriboze po utvoření této vazby nevhýbá žádný volný hydroxyl. Proto je DNA mnohem méně reaktivní než RNA, ve které je na ribose ještě hydroxyl v poloze C2'.
3.2. TORZNÍ ÚHLY

Zatímco konformace peptidové páteře aminokyselin je určena třemi torzními úhly, k popisu cukrofosfátové páteře nukleových kyselin je třeba šest úhlů. Zde jsou uvedeny 11 variant toktickej konformace.

Jako aromatické heterocykly jsou báze nukleových kyselin ploché. V molekule nukleové kyseliny mohou báze vzájemně interagovat buď plochami nebo hranami. Pokud jsou báze uspořádány nad sebou tak, že se k sobě blíží celou plochou (která je vymezena elektrony v delokalizovaných orbitálech, jež dodávají bázi aromatický charakter), dochází ke složité interakci zvané stacking. Pokud neapaku báze k sobě blíží hranami, jsou velmi důležité vodíkové vazby mezi atomy, které tvoří příslušné hrany.

Jeden z dusíků bází (N1 pyrimidinových a N9 purinových) je vázan glykosidickou vazbou ke uhlíku C1′ ribosy nebo deoxyribosy. Ribosa s navázanou bází (bez fosfátu) se nazývá nukleosid. Nukleosidy tvořené uracilem, thymidin, cytidin, adenosin a guanosin se nazývají (v stejném pořadí) uridín, thymín, cytidín, adenin a guanin.

3.2 Torzní úhly

4 vyměněné.

Jakyhosi zrcadlový obraz vztahu, jaké jsme popsalí u pyrimidinových bazí, platí i pro další dvě písmena genetického kódu. Jsou to báze odvozené od většího heterocyklu purinu. Adenin (A) má v poloze 6 aminoskupinu (donor) a dusík N1 slouží jako akceptor. Guanin (G) má v poloze 6 hydroxyl a v poloze 2 aminoskupinu, takže ve výsledné tautomerní formě jsou donory v polohách 1 a 2 a akceptor v poloze 6.

Jako aromatické heterocykly jsou báze nukleových kyselin ploché. V molekule nukleové kyseliny mohou báze vzájemně interagovat buď plochami nebo hranami. Pokud jsou báze uspořádány nad sebou tak, že se k sobě blíží celou plochou (která je vymezena elektrony v delokalizovaných orbitálech, jež dodávají bázi aromatický charakter), dochází ke složité interakci zvané stacking. Pokud se naopak báze k sobě blíží hranami, jsou velmi důležité vodíkové vazby mezi atomy, které tvoří příslušné hrany.

Jeden z dusíků bází (N1 pyrimidinových a N9 purinových) je vázan glykosidickou vazbou ke uhlíku C1′ ribosy nebo deoxyribosy. Ribosa s navázanou bází (bez fosfátu) se nazývá nukleosid. Nukleosidy tvořené uracilem, thymidin, cytidin, adenosin a guanosin se nazývají (ve stejném pořadí) uridín, thymín, cytidín, adenin a guanin.

2 Pozor na odlišný význam koncovky -on v cytosinu a purinových nukleosidu.
3 Tyto torzní úhly, definované atomy, které přímo tvoří cyklický řetězec, se nazývají endocyklické.
4 V anglické literatuře se pro tento druh konformace používá označení pucker.
\[ \nu_i = \Psi_m \cos (P + 144(i - 2)) \] (3.1)

Definice týkající se konformací a pseudorotace jsou opět dány pravidly IUPAC-IUB.

3.3 Primární struktura

Význam pojmu primární struktura je u nukleových kyselin stejný jako u proteinů. Pořadí (sezvětku) nukleotidů se podle konvenční čtení čte od 5' konce do 3' konce vlákna nukleové kyseliny. Zejména v případě sekvence dvoušroubových sekvencí je vhodné 5' a 3' konce vlákén označit. Jak jsme zmínilí v sekci 2.4, je určení sekvence nukleových kyselin poměrně snadné a primární struktura (konfigurace) je vždy známá dříve, než je rešena trojrozměrná struktura (konformace).

3.4 Sekundární struktura

Na základě analogie s proteýny sekundární strukturou označujeme pravidelný, libovolně opakovatelný konformacní motiv poly(nukleotidové) řetězce. Základními sekundárními strukturami nukleových kyselin jsou dvouválcové šroubovice stabilizované vodíkovými vazbami mezi bázemi. Všimněme si důležitého rozdílu mezi proteiny a nukleovými kyselinami: u sekundárních struktur proteínů jsme vždy hovořili o vodíkových vazbách mezi atomy páteře, zatímco dvojšroubovice nukleových kyselin jsou spojené sítí vodíkových vazeb mezi bázemi, které tvoří postranní řetězce polynukleotidu. Sekundární struktura není proto určena pouze konformací páteře, jako v případě proteínů. Je proto třeba definovat všech osm tozmicích úhlů určujících konformaci nucleotidu, aby byl strukturální motiv jednoznačný a popsán.

Stejně jako proteiny, tvoří nukleové kyseliny dva základní typy sekundárních struktur, A a B. Na rozdíl od proteinových α a β struktur se zmíněné sekundární struktury nukleových kyselin liší poměrně méně. V obou případech jde o pravotočivou dvojšroubovicí, ve které jsou antiparalelní vlákna spojena vodíkovými vazbami.

Způsob párování bází je klíčem k pochopení struktury dvojšroubovicí nukleových kyselin. Vzpomeňme si na sekci 3.1, kde jsme si popisovali uspořádání donorů a akceptorů vodíkových vazeb v bázích. Zkusme si nakreslit jednu purinovou a jednu pyrimidinovou bázi. Začněme šestičlennými kruhy, natočené „napříč“ tak aby dusík N1 purinu a dusík N3 pyrimidinu ležely proti sobě. Nakresláme vzorce bázi tak, aby jedna glykosidická vazba sněžila vpravo dolů a druhá vlevo dolů (uspořádání cis). V uspořádání, které jsme si právě nakreslili (obrázek 3.3), se báze stáčejí hranami, které obsahují nejvíce donorů a akceptorů vodíkových vazeb. Aby se vodíkové vazby mohly vytvořit, musí ležet akceptor proti donoru. To je možné pouze tehdy, když guanin tvoří pár s cytosinem a adenin s uracilem nebo thyminem.

Všimněme si, že všechny páry popsány v předchozím odstavci mají téměř stejný tvar. Páruje se vždy menší pyrimidinová báze s větší purinovou a vcelku nezáleží na tom, která je vpravo a která vlevo. Tento způsob párování v se se názvá podle Watsona a Cricka a zaručuje věrné předávání genetické informace. Mám-li jedno vlákno, které kóduje dědičnou informaci, může se s každou jeho bází párovat jen jedna určitá báze vlákna druhého. Pokud bude druhé vlákno dosyntetizováno tak, aby se všechny báze párovaly, ponešlo stejnou informaci. Alfa pozor, toto druhé vlákno nebude kopí protního. Vzhled mezi nimi bude stejný jako vztah mezi formou a odlitkem, tepře další syntézu vlákna, které by se párovalo s vláknom novým, bychom získali stejně pořadí bází.

Stejný tvar všech párů bází podle Watsona a Cricka umožňuje, aby byly součástí pravidelného útvaru – dvojšroubovice (viz obrázek 3.7). V ní musí mít správnou hodnotu také osm tozmicích úhlů, které popisují konformaci nucleotidu. V obou typech dvojšroubovice je konformace taková, že v paru nakresleném tak, jak jsme výše popsalí, vystupuje hlavní řetězec báze nakreslené vlevo z roviny papíru a druhý vlevo dolů pod rovinu papíru na nám a hlavní řetězec báze nakreslené vpravo mří dolů pod rovinu papíru. Tabulka 3.1 udává typické
Obrázek 3.2: Pseudorotační kruh (deoxy)ribosy. Uvnitř kruhu je uvedeno symbolické značení konformací, vně kruhu pseudorodační fáze ve stupních a slovní označení konformací. Na obvodu diagramu jsou schematicky znázorněny struktury jednotlivých konformerů. Atomy jsou symbolizovány kroužky (nejdále od středu je vždy kyslík O4'). Symbol v kroužku uvádí, leží-li daný atom v rovině nákresu (○), nad rovinou (+), nebo pod rovinou nákresu (−). Znaménka u vazeb udávají, zda je daný torzní úhel kladný nebo záporný.
KAPITOLA 3. STRUKTURA NUKLEOVÝCH KYSELIN

Tabulka 3.1: Hodnoty torzních úhlů ideálních dvojšroubovic.

<table>
<thead>
<tr>
<th></th>
<th>α</th>
<th>β</th>
<th>γ</th>
<th>δ</th>
<th>ε</th>
<th>ζ</th>
<th>χ</th>
<th>P</th>
<th>Ψ_m</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-DNA</td>
<td>−41</td>
<td>136</td>
<td>38</td>
<td>139</td>
<td>−133</td>
<td>−57</td>
<td>78</td>
<td>155</td>
<td>40</td>
</tr>
<tr>
<td>A-DNA</td>
<td>−90</td>
<td>−149</td>
<td>47</td>
<td>83</td>
<td>175</td>
<td>45</td>
<td>27</td>
<td>13</td>
<td>40</td>
</tr>
<tr>
<td>A-RNA</td>
<td>−80</td>
<td>175</td>
<td>49</td>
<td>83</td>
<td>−147</td>
<td>−78</td>
<td>13</td>
<td>13</td>
<td>40</td>
</tr>
</tbody>
</table>

Jak ukazuje obrázek [3.3], cukří fosfátová pátéř ovíjí dvojšroubovice zvenčí, takže zbytky kyseliny fosforečné (za běžných podmínek v disociovaném stavu, tedy záporně nabité) jsou velmi dobře přístupné. Báze je naopak uvnitř dvojšroubovice, ne však zcela skrytá. Mezi páteri jsou dostatečně široké žlábky, které umožňují jiným molekulám (nebo jiným částem téhož polynukleotidové retízce) interagovat s bázemi. Každou bázi si můžeme zjednodušeně představit jako trojúhelník (obrázek 3.4). Tyto trojúhelníky jsou párovány jednou hranou (Watson-Crickovou), druhou hranou je přístupná z žlábku po jedné straně páter a třetí hranou je přístupná z žlábku po druhé straně páter.

Z vnějšího pohledu se liší vzhled žlábků mezi páteri jednotlivých vláken na povrchu dvojšroubovice. Dvojšroubovice B má jeden žlábek širší a hlubší (označuje se hluboký nebo velký), druhý užší a mělký (označuje se mělký nebo malý), zatímco dvojšroubovice A má jeden žlábek velmi úzký a hluboký (označuje se hluboký) a druhý široký a velmi mělký (označuje se mělký). V obou případech do hlubokého žlábku směřuje hranu, která se u purínů nazývá Hoogsteenova a u pyrimidínů „C-H“ hranu. Do mělkého žlábku naopak směřuje hranu v blízkosti glykosidické vazby, pro kterou byl proto navržen název „sugar-edge“ (do části můžeme přeložit „ sacharidová“).

Při pohledu podél osy dvojšroubovice je patrná dutina ve středu dvojšroubovice A. Dvojšroubovice A má také ponekud pozvolnější stoupání (11 residuí na závit) ve srovnání s dvojšroubovice B (10 residuí na závit).  

5 Ve snaze o jednotnost názvosloví se často používá název „Hoogsteenova“ i pro „C-H“ hranu pyrimidínů.
Obrázek 3.4: Zjednodušený popis bází nukleových kyselin pomocí trojúhelníků (s označením jednotlivých hran).

Pro popis vzájemné polohy bází ve dvojšroubovicích byly definovány takzvané helikální parametry (obrázky 3.5 a 3.6). Na první pohled je patrný rozdíl v inklinci. Zatímco báze v dvojšroubovicí B jsou těsně kolmé k ose, normálně báze v dvojšroubovicí A svírají s osou úhel asi 20°. Dvojšroubovice B a A jsou ukázány na obrázku 3.7 a na modelech A a B.

Podobně jako u proteinů se konformace DNA neomezují na popsané typické sekundární struktury. V závislosti na sekvenci a fyzikálních podmínkách se setkáváme s mnoha odlišnými strukturami (viz například zprohýbaná levotočivá dvojšroubovice Z u sekvencí opakující se dvojice CG).

3.5 Terciární struktura

DNA se jako nosič genetické informace nejčastěji vyskytuje ve formě dvojšroubovice B (i když se za určitých fyziologických podmínek a na určitých místech můžeme setkat s neobvyklými čtyřšroubovivými, větvěnými a jinými strukturami).

Prostorové uspořádání RNA je naproti tomu pestřejší a připomíná daleko více terciární struktury proteinů. Molekuly RNA jsou tvořeny úseky dvojšroubovic A (nazývaných stems, stonky) spojenými různými spojovacími prvky.

Stejně jako v proteinových ohybech se u spojovacích prvků nukleových kyselin setkáváme s odlišnými hodnotami torzních úhlu a odlišnou tvorbou vodíkových vazeb. Za zmínku stojí, že se párují i báze vy- skytující se mimo pravidelné dvojšroubovice. Na rozdíl od kanonických páru, popsaných v sekci 3.4, je zde mnoho možností, jak sň vodíkových vazeb vytvořit. Mimo dvojšroubovice nejsou tvary páru limi- továny pravidelnou strukturou, takže guanin se může dobře párovat i s uracilem či adeninem. Vodíkové vazby se tvoří nejen mezi dvěma Watson-Crickovými hranami, ale možné jsou všechny kombinace tří hran (Watson-Crickova, Hoogsteenova, sacharidová). Glykosidické vazby navíc nemusí být v poloze cis, ale možné je i poloha trans, kdy glykosidické vazby mří opačným směrem. Celkově je tak k dispozici 12 typů párování (obrázek 3.8), příklady některých nekanonických páru jsou uvedeny na obrázku 3.9.

V každém páru je využita jen jedna hrana jednotlivých bází, takže báze párována jedním způsobem (například podle Watsona a Cricka) může stále interagovat další hrany (například Hoogsteenovou). Takovým způsobem vznikají triplety bází, které mohou spojovat dvojšroubovice s dalším vlákem, aniž by se výrazně měnila pravidelná dvoušroubovicová struktura.

Pravidelné dvojšroubovice a struktury zmíněné v předchozím odstavci mohou tvořit supersekundární motivy podobné proteinovým (obrázek 3.10), jako vlásenkové smyčky (hairpin loops) a zlomy v dvojšroubovici.
Obrázek 3.5: Translační helikální parametry. V levém horním rohu je znázorněna souřadná soustava, ve které se helikální parametry udávají. Parametry v horní řadě udávají posunutí páru bází jako celku (souběžná posunutí), parametry ve střední řadě udávají vzájemně posunutí bází v jednotlivých párech vůči sobě (protisměrná posunutí), parametry v dolní řadě udávají vzájemná posunutí sousedních pář bází. Posunutí ve směru osy x, y, z jsou seřazena zprava doleva.
3.5. TERCÍARNÍ STRUKTURA

Obrázek 3.6: Rotační helikální parametry. V levém horním rohu je znázorněna souřadná soustava, ve které se helikální parametry udávají. Parametry v horní řadě udávají natočení párů bází jako celků (souběžné rotace), parametry ve střední řadě udávají vzájemná natočení bází v jednotlivých párech vůči sobě (protisměrné rotace), parametry v dolní řadě udávají vzájemná natočení sousedních párů bází. Rotace kolem osy $x$, $y$, $z$ jsou seřazeny zprava doleva.
Obrázek 3.7: Modely DNA šroubovic B a A.

Obrázek 3.8: Možné způsoby párování bází. Báze jsou znázorněny jako trojúhelníky s udáním směru hlavního řetězce (⊙ směřuje dopředu z roviny nákresu, □ směřuje dozadu za rovinu nákresu). Hraný jsou rozlišeny dílkou stran trójúhelníku – nejlépe Watson-Crickova, prostřední sacharidová, nejdelší Hoogsteenova nebo „C-H“ Nadmoří jsou znázorněna párování s cis uspořádáním glykosidických vazeb, dole s trans uspořádáním glykosidických vazeb.
Obrázek 3.9: Příklady nekanonických párů bází: vlevo nahoře cis párování uracilu (Watson-Crickovou hranou) s adeninem (Hoogsteenovou hranou), vpravo nahoře trans párování dvou adeninů Hoogsteenovými hranami, vlevo dole cis párování dvou cytosinů sacharidovou a ,,C-H“ (,,Hoogsteenovou“) hranou, vpravo dole trans párování cytosinu a guaninu Watson-Crickovými hranami.
KAPITOLA 3. STRUKTURA NUKLEOVÝCH KYSELIN

Obrázek 3.10: Příklady strukturních motivů nukleových kyselin. Zleva vlákenková smyčka (*hairpin loops*), zlom s včleněným nepárovaným nukleotidem *bulge, internal loop, four-way junction, pseudoknot, kissing hairpin*.

Supersekundární struktury typické pro nukleové kyseliny jsou dále různé vnitřní smyčky (*internal loops*) a větvení (*junctions*).

Z vyšších strukturních prvků můžeme jmenovat trojšroubovicová uspořádání, kvadruplexy (útvary tvořené čtyřmi vláky), pseudouzly (*pseudoknots*), neobvykle vzájemně párované smyčky romanticky označované *kissing hairpins* (obrázek 3.10).

Domény RNA jsou definovány stejně jako v případě proteinů. Například velká podjednotka bakteriální ribosomové RNA je tvořena šesti doménami, z nichž každá má specifickou funkci a samostatnou strukturu.

### 3.6 Kvartérní struktura a vyšší struktury

Pojem kvartérní struktury můžeme převzít tak, jak jsme si jej definovali pro proteiny. Nukleové kyseliny často tvoří funkční oligomerní komplexy s proteiny. Podobně jako proteiny se ribonukleové kyseliny podílí na stavbě nadmolekulárních útvarů (opět můžeme jmenovat ribosom).


---

*Jednotlivá vlákna musí být kruhová nebo pevně uchycená například k proteinům, jinak by se nadšroubovice samovolně rozplela.*
Kapitola 4

Struktura oligo- a polysacharidů

4.1 Stavební jednotky

Nejnápadnějším rysem, který bude naší diskusi struktur oligosacharidů a polysacharidů provázet, je neobyčejná pestrost. Mnohem větší strukturní různorodost sacharidů ve srovnání s proteiny a nukleovými kyselinami se u oligo- a polysacharidů objevuje již na úrovni monomerních jednotek. Základní stavební jednotky se liší počtem uhlíků (nejčastěji jsou hexosy a pentosy), funkční skupinou (aldosy a ketosy), geometrickou isomerií hydroxylů (2ⁿ optických isomerů, kde n je počet chirálních uhlíků), možnou substitucí (sulfátem, karboxylem, aminem, N-acetylaminem), počtem uhlíků v kruhu (furanosy a pyranosy), konformací cukerného kruhu, polohou a počtem hydroxylů, které tvoří glykosidickou vazbu se sousedním reziduem. Nejběžnější monosacharidové jednotky jsou uvedeny na obrázku 4.1. Názvosloví sacharidů se řídí pokyny IUPAC-IUB.

4.2 Torzní úhly

Monosacharidové jednotky se vyskytují nejčastěji ve formě pětičlenných a šestičlenných kruhů. Sacharidy s pětičlenným cyklem se nazývají furanosy, zatímco sacharidy s šestičlenným cyklem se označují jako pyranosy.


Pokud oba protilehlé atomy leží nad rovinou kruhu, mluvíme o vaničkové konformaci. Takových vaniček můžeme sestrojit šest: 1₄B, 2₅B, 3₀B, 1₄B, 2₅B, 3₀B (horní indexy opět značí atomy nad
KAPITOLA 4. STRUKTURA OLIGO- A POLYSACHARIDŮ

Obrázek 4.1: Vzorce nejčastějších monosacharidových jednotek oligo- a polysacharidů. Ze dvou možných anomérů α a β je vždy uveden jen jeden.
rovinou a dolní indexy atomy pod rovinou). Popis jednotlivých konformací se řídí pravidly IUPAC-IUB.

Stabilita pyranosových konformerů je tak velká, že je lze považovat za rigidní strukturní jednotky. Konformaci polysacharidového řetězce lze proto zpravidla popsat dvěma \((\phi, \psi)\) nebo třemi \((\phi, \psi, \omega)\) torzními úhly. Úhel \(\phi\) definuje torzi kolem glykosidické vazby. Úhel \(\omega\) je nezbytný pokud glykosidickou vazbu tvorící hydroxyl není vázaný přímo na uhlík pyranového nebo furanosového kruhu, ale na methylenový uhlík (například u \(1 \rightarrow 6\) vázaných aldopyranos nebo \(1 \rightarrow 5\) vázaných ketofuranos).

4.3 Primární struktura

Vzhledem k možnosti větvení a k zmíněné pestrosti monomerních jednotek je definice primární struktury oligo- a polysacharidů složitější než v případě proteinů a nukleových kyselin. Pravidla pro zápis primární struktury oligosacharidů lze nalézt v doporučených IUPAC-IUB.

Další důležitou odlišností od proteinů a nukleových kyselin je skutečnost, že primární strukturu oligo- a polysacharidů nelze přímo vyčíst z genetického kódu. Určování primární struktury je proto v případě sacharidů mnohem obtížnější, než v případě proteinů a nukleových kyselin.

4.4 Sekundární struktura

Pestrost sekundárních struktur polysacharidů ostře kontrastuje s dvojicí typických strukturních motivů proteinů a nukleových kyselin. Analogue s popisem sekundárních struktur proteinů jsou ale zřejmé. Stejně jako u proteinů, můžeme sekundární struktury polysacharidů formálně považovat za šroubovice. K jejich popisu bývá používána dvě čísla: počet monosacharidových jednotek na jeden závit \(n\) a stoupání šroubovice na jednu monosacharidovou jednotku \(h\) (nebo stoupání šroubovice na jeden závit \(P\)). Sekundární struktury jsou opět stabilizovány vodíkovými městky mezi závity těže šroubovice (jako u \(\alpha\)-helixu proteinů) nebo spojující dvě polysacharidová vlákna (jako v \(\beta\)-listech proteinů). Sekundární struktury polysacharidů můžeme delít na šroubovice, ve které na jeden závit případně dvě rezidua, například cellulose (vzor model, analogie s \(\alpha\)-sroubovickými proteinovými řetězci), a natažené (vzor model, analogie s \(\beta\)-listy proteinů). Některé struktury polysacharidů nelze zařadit do této dvou kategorií, jako například velmi flexibilní dextrany vázané \(1 \rightarrow 6\) glykosidickou vazbou nebo naopak těsně skládané bakteriální (1 \(\rightarrow 2\)) glukany.

4.5 Terciární struktura

Sacharidy mohou tvořit velmi dlouhé řetězce, často s širokou hmotnostní distribucí. Díky větvení mohou vznikat polysacharidy přibližně kulového tvaru. Naopak přesně definované oligosacharidové struktury plní důležitou úlohu v molekulovém rozpoznávání. Na rozdíl od proteinů jsou terciární struktury polysacharidů méně prozkoumány.

Příkladem architektury polysacharidů může být celulosa, základní složka buněčných stén rostlin. Sekundárním motivem celulosy je natažený řetězec \(\beta\)-d-glukopyranos-\(4\)C\(_1\) spojených \((1 \rightarrow 4)\) glykosidickou vazbou. Jde tedy formálně o šroubovici s dvěma \(\beta\)-d-glukopyranosovými jednotkami na jeden „závit“ jinými slovy, \(4\)C\(_1\)-ždílčky \(\beta\)-d-glukopyranos se střídají tak, že kyslík pyranového kruhu leží

---

1 Kromě popsaných konformací, ve kterých jsou atomy ležící mimo rovinu odděleny dvěma atomy kruhu, vyskytují se vzácně i konformace, kde leží mimo rovinu atomy oddělené jen jedním atomem (šikmé ždílčky, \(S\)) nebo sousedními atomy (polozdílčky, \(H\)). Podobně jako v případě furanové pyranové konformace popsat pomocí pseudorotace. Vzhledem k tomu, že šestíčlenné cykly vyžadují tři nezávislé parametry, nevystačíme s pseudorotačním kruhem, ale musíme definovat pseudorotační koúl.
KAPITOLA 4. STRUKTURA OLIGO- A POLYSACHARIDŮ

Obrázek 4.2: Uspořádání řetězců v celulose Iα (vlevo) a Iβ (vpravo). Střídající se β-d-glukopyranosové kruhy s kyslíkem nad a pod rovinou řetězce jsou znázorněny jako plné a prázdné obdélníky.

Jednou nad rovinou a podruhé pod rovinou řetězce. V tomto uspořádání leží všechny hydroxyly v ekvatoriální poloze (v rovině kruhu). Vytváří se tak polární prostředí po bocích celulosového řetězce a nepolární prostředí nad a pod rovinou řetězce. Natažené řetězce se k sobě skládají nepolárními plochami a tvoří krystalovou strukturu označovanou I. Kromě hydrofobních interakcí je tato struktura stabilizována sítí vodíkových vazeb v rámci jednoho řetězce i mezi sousedními řetězci. Je zajímavé, že glykosidická vazba celulosy I je v konformaci trans-gauche a řetězce směřují stejným směrem (paralelní uspořádání), ačkoli energeticky výhodnější je gauche-trans konformace glykosidické vazby a střídavá (antiparalelní) orientace řetězů. Přirozená forma celulosy je tedy metastabilní[1] na energeticky nejvýhodnější formu, označovanou II, přechází až umělým zpracováním. Celulosa I se vyskytuje ve dvou formách Iα a Iβ, které se liší vzájemným uspořádáním řetězců (obrázek 4.2). Formy Iα a Iβ se mohou střídat i v rámci jednoho celulosového vlákna, jehož délka dosahuje do 10 000 β-d-glukopyranosových jednotek.

Jiným příkladem polysacharidu je zásobná látka rostlin, škrob. Základní stavební jednoty škroubu jsou α-d-glukopyranosy-[4]C1, spojené (1 → 4) glykosidickou vazbou. Řetězce α-d-glukopyranos mohou tvořit různé šroubovicové struktury, z nichž nejdůležitější je paralelní dvojsroubovice s výškou závitu 0,5 nm a šesti až sedmi jednotkami na jeden závit. Stejně jako v případě celulosy je důležité uspořádání hydroxylů, které směřují na povrch šroubovic, takže vnitřek šroubovic je nepolární. Škrob se skládá z nevětvených řetězců amylosy a občích molekul amylopektinu, které jsou tvořeny přibližně milionem α-d-glukopyranosových jednotek, a které obsahují i větveni prostřednictvím (1 → 6) glykosidické vazby. Větvení níjak nevadí dvojsroubovicové struktuře, naopak ji spíše stabilizuje. Dvojsroubovicové úseky mezi větveninami jsou základem vrstevnaté struktury škrobových zrn (obrázek 4.3). V těchto vrstvách jsou dvojsroubovice pravidelně uspořádány jako v krystalech. Řetězce amylosy jsou nejspíše náhodné rozptýleny jako jednoduché šroubovice mezi amylopektinovými řetězci (zvláště na povrchu škrobových zrn), částečně se mohou také podílet na tvorbě dvojsroubovic.

Stejně jako jiné biomakromolekuly tvoří oligo- a polysacharidy funkční komplexy s proteiny a podílejí se na tvorbě výsších struktur buněčné architektury. Je také důležité si uvědomit, že velkou skupinou biomakromolekul jsou glykoproteiny, jejichž oligosacharidová část je co do velikosti, důležitosti a stability konformace srovnatelná s částí proteinovou.

[1] Přítomnost metastabilní celulosy I je vysvětlována tím, že celulosová vlákna se tvoří krátce po biosyntéze, během kterérostou celulosové řetězce jedním směrem, a na přeskládání do výhodnějšího antiparalelního uspořádání není dost času.
Obrázek 4.3: Schematické znázornění větvení amylopektiny. Rovnoběžné čáry představují paralelní dvojsroubovice.
Kapitola 5

Struktura biologických membrán

5.1 Membrány a biomakromolekuly

Biologické membrány se od biomakromolekul (proteinů, nukleových kyselin, polysacharidů) zásadně liší. Zatímco biomakromolekuly tvorí dlouhý polymerní řetězec, v němž jsou stavební jednotky pevně spojeny kovalentní vazbou, membrány se skládají z molekul, mezi kterými působí především slabé van der Waalsovy síly. Přesto můžeme nalézt důležitý rys společný membránám i biomakromolekulám. V obou případech je dobře známá konfigurace jednotlivých stavebních jednotek a popisem struktury se myslí určení geometrického uspořádání těchto stavebních jednotek.

5.2 Stavební jednotky

Biologické membrány jsou tvořeny látkami, které souhrnně označujeme \textit{lipidy}. Nejdě se o název odrážející chemické složení, ale spíše fyzikální vlastnosti. Molekuly lipidů obsahují relativně malou polární, často ionizovanou část, která má výrazné hydrofilní chování. Zbytek molekuly je nepolární, výrazně hydrofobní. Molekula lipidu je tedy „vnitřně rozpolcená“ – jedna část vyhledává polární (vodná) prostředí, druhá část nepolární prostředí. Chemická struktura lipidů je velmi rozmanitá, k nejběžnějším patří následující tři typy:

- \textit{Glycerolfosfolipidy} (\textit{fosfatidy}) – k molekule glycerolu jsou vázány esterovou vazbou dvě mastné kyseliny a jeden zbytek kyseliny fosforečné. Z mastných kyselin jsou nejběžnější kyselina palmitová (šestnáctiuhlíková nasyacená kyselina), osmnáctiuhlíkové kyseliny (nasyacená stearová, s jednou dvojnou vazbou v polohě 9 oleová, s dvěma dvojnými vazbami v polohách 9 a 12 linoleová) a kyselina arachidonová (dvacetiuhlíkatá se čtyřmi dvojnými vazbami v polohách 5, 8, 11, 14).
  Na kyselinu fosforečnou je často další esterovou vazbou vázan kyslík ethanolaminu, serinu, cholinu, nebo inositolu (cukerný alkohol, na který bývá vázan další jeden nebo dva zbytky kyselin fosforečné).

- \textit{Sfingolipidy} – k molekule dvacetniuhlíkatého aminodiolu sfingosinu je vázána amidovou vazbou jedna mastná kyselina a jeden zbytek kyseliny fosforečné, který může být dále esterifikován (nejčastější cholinem).

- \textit{Steroly} – v živočišných buňkách cholesterol
5.3 Geometrie lipidových útvarů

Pokud přidáme lipidy ke směsi nemísitelného polárního a nepolárního rozpouštědla (voda a olej), budou ochotné tvořit rozhraní těchto dvou rozpouštědel. Hydrofobní část lipidu bude mířit do oleje a hydrofilní do vody. Pokud přidáme lipidy pouze do vodného roztoku, budou se jejich molekuly orientovat tak, aby nepolární části byly co nejméně v kontaktu s vodou (hydrofobní efekt). Stejně jako v případě nepolárních postranních řetězců proteinů (viz sekcí 1.2.3) je příčinou převětšené entropie solvatačních obalů, které musí rozpoutět dvojvrstvu vytvořit kolem nepolární části lipidu.

Tvar útvaru, který v důsledku hydrofobního efektu vznikne, závisí na tvaru molekuly lipidu. Pokud je nepolární část lipidu malá, podobá se jeho molekula kuličce, jehož podstavu tvoří polární část. Takovéto kužely můžeme poskládat do zhruba kulového útvaru, který bude uvnitř nepolární a zvenčí polární.

Takový útvar se nazývá **micela**.

Lipidy, které jsme si vyjmenovali v sekcii 5.2 mají hydrofobní a hydrofilní část přibližně stejnou, takže jejich molekuly se podobají válci. Na rozdíl od kuželů můžeme válce skládat do plochých útvarů, které se budou orientovat stejně, jako když k sobě přiložíme dva krajnice namazaného chleba. Vznikne **dvojvrstva** tvořená lipidy, které budou orientovány hydrofobní částí proti sobě. Taková dvojvrstva je základem biologických membrán.

5.4 Tekutost membrán

Voda a jiné látky se mohou vyskytovat v různých skupenstvích, která se liší uspořádáním a volností molekul v trojrozměrném prostoru. V případě biomembrán se setkáváme s dvourozměrnou analogií, mluvime o různých fázích lipidové dvojvrstvy. Při nízkých teplotách a velkém natěšení lipidových molekul na malé ploše se tvoří dvojrozměrné krystaly. Mají-li molekuly trochu více volnosti, dochází k vytvoření dvojrozměrného gelu, tato fáze se označuje $L_\beta$ (obrázek 5.1). Zvýšením teploty vzniká **fáze kapalného krystalu** $L_\alpha$, ve které mají molekuly více volnosti, než v plochém gelu. Teprve dalšími fázovými přechody vzniká fáze kapalného krystalu $L_\alpha$.

V biomembránách se vyskytují lipidy převážně v fázi kapalných krystalů, velmi blízko fázovému přechodu na gelovou strukturu. Mluví se často o **tekulé mozaice**, navenek tvořící jasně definovanou membránu, ve které se však mohou molekuly volně pohybovat.

5.5 Konformace lipidů a struktura membrány

5.5.1 Vliv polární skupiny

Polární část lipidu se v membránách vyskytuje v poměrně přesné definované konformaci. Tato konformace je dána požadavkem, aby molekula jako celek byla zhromažďována a mohla tak neomezeně tvořit ploché struktury. Například v glycerolofosfolipidech obsahujících ethanolamin jsou v gelové fázi

---

1 Vedle micel a dvojvrstev existuje celá řada možných uspořádání, ve kterých se budou nacházet lipidy, jejichž tvar bude nejčastěji mezí válcem a kuželem. V buněčných membránách se setkáváme s podobnou geometrií například v místě styku dvou membrán.
5.5. KONFORMACE LIPIDŮ A STRUKTURA MEMBRÁNÝ

obrázek 5.1: Fázové přechody lipidových dvojvrstev.


5.5.2 Vliv hydrofobního řetězce

V sekci 5.4 jsme zmínili, že tekutost membrán je dána volností pohybu jednotlivých molekul. Není proto divu, že konformace hydrofobních řetězců úzce souvisí s tekutostí biomembrány. Pokud je hydrofobní řetězec mastné kyseliny nasycený a pokud jsou všechny vazby v tomto řetězci v konformaci $trans$, zabírá tento řetězec v rovině membrány plošku o průměru 0,42nm. Takové řetězce tedy mohou být velmi těsně poskládány a vyskytují se v tuhých gelových fázích. Rotace kolem jediné vazby do konformace gauche(+) vede k ohnutí řetězce, který naráz zabírá několikanádobně více místa. Takováto konformační změna tedy znamená obrovské rozvolnění struktury v daném místě. Změněný dramatický efekt však může být v přístím okamžiku kompenzován další konformační změnou, při které předje vazba vzdálená o dva uhlíky od místa zlomu do konformace gauche(−). Dva opačné ohyby se navzájem téměř vyrusí a dvakrát zlomený řetězec bude zabírat jen asi o 10% více místa než původní.

V čem se bude lišit chování hydrofobního řetězce v případě, že jde o nenasycenou mastnou kyselinu? Pokud budou všechny vazby v konformaci (trans), bude zabírat mnohem více místa než nasycená. Mastné kyseliny totiž obsahují dvojná vazby v konfiguraci cis, které vedou k podobnému ohybu řetězce,
jako konformační změny popsané v předchozím odstavci. Pokud ale bude dva uhlíky od dvojně vazby změněna konformace z trans na gauche, bude ohyb řetězce do značné míry vykompenzován.

Můžeme tedy shrnout, že změna konformace z uspořádání trans do polohy gauche může mít v různé situaci opačný účinek. V případě nasycených mastných kyselin vede ke zvýšení tekutosti, zatímco v případě nenasycených nebo již jednou konformačně deformovaných mastných kyselin vede ke zvýšení tuhosti membrány.

5.5.3 Vliv složení
Podobný účinek jako konformační změny postranního řetěze mají i změny složení membrány. Například přidávek cholesterolu do těsně uspořádané, relativně tuhé membrány tvořené nasycenými mastnými kyselinami naruší strukturu membrány a zvýší její tekutost. Naopak přidávek cholesterolu k relativně tekuté membráně obsahující značné množství nenasycených mastných kyselin vede ke zvýšení tuhosti, protože molekula cholesterolu může vyplňovat dutiny, které vznikly v důsledku přítomnosti dvojných vazeb.

5.6 Vyšší struktury
Tak jako se biomakromolekuly navzájem spojují a tvoří vyšší komplexy, tak i biomembrány často interagují s různými biomakromolekuly. Lipidy tak tvoří řadu polovinu hmoty buněčných membrán. Velká množství proteinů nabývá správné konformace pouze v přítomnosti membrán. Membránóvé proteiny mohou lipidovou dvojvrstvou procházet, a to i několikrát, nebo k ní mohou být připojeny z jedné či druhé strany. Interakce mezi proteiny a membránou mohou být dány hydrofobním jevem, ale proteiny se mohou k lipidům vázat i kovalentními vazbami. Podobně mohou být k lipidům dvojvrstev vázány i polysacharidy. Funkce membrán není omezena jen na udržování tvaru buněk. Rozhodující část tvorby energie a regulace mnoha buněčných procesů jsou přímo spojeny s rozdělením prostoru biologickými membránami.
Část II

Lov struktur
Výpočetní metody

6.1 Molekulové modelování

6.1.1 Model molekuly


Abychom odlišili náš popis struktury molekuly od skutečné molekuly, budeme vypočítanou strukturu nazývat **modelem molekuly**. Slovo „model“ budeme používat v trochu obecnějším významu, než je konstrukce z trubiček a kuliček, kterou jsme nazvali modelem v sekci 1.1.1. Model bude pro nás popisem všech vlastností struktury, které nám výpočet poskytne. Samotný výpočet, kterým model získáme, můžeme označit jako molekulové modelování.


Hledání modelu molekuly je ovšem mnohem složitější, než se na první pohled může zdát. I kdybychom uměli energii přesně spočítat, nic nám nezaručí, že naše molekula opravdu odpovídá nejnižší energii, globálnímu minimu. Již sama existence izomerů naznačuje, že vedle sebe mohou stát molekuly s různou energií, aniž by se hned přeměnily v tu, která je z nich energeticky nejvýhodnější. Genetický kód je také založen na tom, že může existovat obrovské množství konfigurací stejné sady atomů. Tato různost konfigurací umožňuje kódování dědičné informace. Nestačí nám tedy znát termodynamiku, musíme znát i kinetiku, která nám říká, jak dlouho mohou atomy setrvat v určitém stavu, i když existuje stav pro ně výhodnější. A to mohou být z pohledu lidského života věky.

Spokojme se tedy s tím, že nemá smysl počítať, která konfigurace atomů má nejnižší energii. Řekněme, že víme jaká je konfigurace naší molekuly. Způsobem, jakým tuto informaci získat, věnujeme kapitolu 8. Není možné vypočítat alespoň nejvýhodnější konformaci, která se experimentálně určuje.
mnohem obtížněji, než konfigurace?

Pro malé molekuly je analýza konformací poměrně snadná. Výpočty energií nejsou sice tak jednoduché, jak by se mohlo na první pohled zdát, ale dobré počítače s ními celkem slušně poradí. Přechody mezi jednotlivými konformacemi jsou rychlé, takže můžeme předpokládat, že ve zkumavce máme takové počty jednotlivých konformérů, jaké předpovídají jejich energie. Vše se ale komplikuje, když přecházíme ke složitéjším molekulám. Výpočty energií jsou složitější a složitější, takže musíme používat metody cím dál tím více zjednodušené a méně přesné. Přestože i u velkých molekul zpravidla hledáme konformaci s nejnižší energií, nemáme jistotu, že se v ní molekula opravdu nachází. Přechody mezi stavby s různou konformací jsou totiž pro velké molekuly obtížné a může se stát, že molekula po celý svůj život setrvá v konformaci s vyšší energií. Následující stránky by měly ukázat, kam a z ve výpočtech konformací biomakromolekul umíme dnes dojít a jakým způsobem se potýkáme s obtížemi, které ze složitosti biologicky zajímavých molekul plynou.

6.1.2 Minimalizace energie

Základní myšlenkou molekulového modelování je hledání stavu molekuly s nejnižší energií. Energie molekuly závisí na množství proměnných, které popisují stav atomů v molekulě. Budeme proto s energií zacházet jako s funkci mnoha proměnných. Kdyby energie závisela jen na jedné proměnné, násli bychom minimum snadno podle následujících pouček:

- Jestliže je první derivace funkce pro určitou hodnotu proměnné rovna nule, pak má funkce v daném bodě minimum, maximum, nebo inflexní bod.

- Jestliže je druhá derivace funkce pro stejnou hodnotu proměnné větší než nula, má funkce v daném bodě minimum (záporná druhá derivace odpovídá maximu a nulová druhá derivace odpovídá inflexnímu bodu).

Pokud jde o funkci více proměnných, nestačí k hledání minima jediné číslo. Musíme počítat derivace funkce podle všech proměnných. Průběh funkce v určitém bodě popisují vektory, které obsahují první a druhé derivace v daném bodě podle všech souřadnic. Matematicky se vektor prvních derivací označuje jako gradient a vektor druhých derivací jako hessián.


Metod, kterých můžeme k hledání modelu molekuly použít, je celá řada. Dají se rozdělit na metody založené na kvantové mechanice, které se snaží nejvíce přiblížit přesnému popisu přírodních zákonů, a na metody takzvané molekulové mechaniky, které se snaží o co největším podpis molekuly poskytující uspokojivý odhad energie.

---

1Vedle metod založených na derivacích, jako je gradientová metoda, je možné použít i nederivační metody, jako je metoda simplexová (na počátku náhodně zvolíme proměnné, ty potom měníme a sledujeme, které změny vedou ke snížení energie). Volba nejvyhodnější metodě závisí na počtu proměnných, na tom jak jsme daleko od minima a na tom, jak citlivé závisí energie na jednotlivých proměnných. V průběhu minimalizace bývá užitečné metody stříhat.
6.2 Kvantové metody

6.2.1 Kvantová mechanika atomů

I ty největší biomakromolekuly se skládají z atomů. Atomy jsou tvořeny jádrem a elektrony. Jak postupujeme k menším a menším částicím, přestávají platit fyzikální zákony tak, jak jsme na ně zvyklí. Pro popis malých částic musíme mít naši klasické mechaniky použít obecnější metodu – kvantovou mechaniku. V případě elektronů je kvantové chování již tak výrazné, že nám neumožňuje říci přesně, kde se elektron přes podobně začne amorfickým řást v daný okamžik. To ale neznamená, že mítlo výskytu elektronů je zcela náhodné. Pro popis elektronů se používá komplexní matematická funkce (jako proměnné v ní vystupují souřadnice a čas), které fyzici říkají vlnová funkce. Místo toho, abychom sledovali, jak se působením vnějších sil mění polohy a rychlosti elektronů, sledujeme jak se mění vlnová funkce \( \psi \). Kvantová mechanika nám pak poskytuje předpisy, jak z vlnové funkce získat informace o fyzikálních veličinách elektronů. Jednou takovou užitečnou informací je pravděpodobnost, že se v daný okamžik elektron vyskytuje v určitém místě. Tuto pravděpodobnost můžeme spočítat jako druhou mocninu (absolutní hodnoty) vlnové funkce \( \psi \), tedy jako součin vlnové funkce a funkce s ní komplexně sdružené (tedy s opačným znaménkem imaginární části, označuje se hvězdičkou): \( \psi \psi^\ast \).

V atomech se elektrony vyskytují jen v určitých stavech. Každý stav má určitou energii (energie se tedy nemění spojitě, ale poskakuje od stavu ke stavu) a je popsan určitou vlnovou funkcí. Vlnová funkce elektronu v atomu nebo molekule se nazývá orbital. Nemůžeme tedy říci, že se v určitém stavu nachází elektron v určitém místě, ale můžeme říci, že každý stav má určitou pravděpodobnost výskytu elektronů a že tuto pravděpodobnost můžeme spočítat pomocí uvedeného předpisu \( \psi \psi^\ast \).


Ostatní atomy mají více než jeden elektron a přesné řešení kvantových rovnic, které poskytují jejich vlnových funkcí, není možné. Často se proto používá zjednodušené řešení, nazývané Hartreeho-Fockova metoda. Zjednodušením je zde předpoklad, že jednotlivý elektron nevidí své kolegy jako jednotlivé

---

2 Tyto změny vlnové funkce popisuje obecný předpis známý pod jménem Schrödingerova rovnice.
3 Toto platí nejen pro kvantovou mechaniku, ale i pro mechaniku klasickou. Problém tří těles je v astronomii stejně nerešitelný jako problém tří částic v kvantové fyzice.
částice, ale jako jakýsi rozmazaný průměrný záporný náboj\footnote{Slůvko „nevidí“ zde označuje elektrostatické odpuzování mezi stejně nabitými částicemi. Z fyzikálního pohledu potom řešíme pohyb jediného elektronu v elektrickém poli tvořeném kladně nabitým jádrem a rozptýleným záporným nábojem.} Slůvko „nevidí“ zde označuje elektrostatické odpuzování mezi stejně nabitými částicemi. Z fyzikálního pohledu potom řešíme pohyb jediného elektronu v elektrickém poli tvořeném kladně nabitým jádrem a rozptýleným záporným nábojem.\footnote{Pokud předpokládáme, že obláček ostatních elektronů kolem jádra má pěkně kulový tvar, dostaneme stejně tvary atomových orbitálů jako v případě atomu vodíku. Chemici se proto neostýchají používat tyto orbitály, které jsou odvozeny z vodíkových, i pro ostatní atomy.}

Stejně jako jakýsi rozmazaný průměrný záporný náboj. Slůvko „nevidí“ zde označuje elektrostatické odpuzování mezi stejně nabitými částicemi. Z fyzikálního pohledu potom řešíme pohyb jediného elektronu v elektrickém poli tvořeném kladně nabitým jádrem a rozptýleným záporným nábojem. \footnote{Pokud předpokládáme, že obláček ostatních elektronů kolem jádra má pěkně kulový tvar, dostaneme stejně tvary atomových orbitálů jako v případě atomu vodíku. Chemici se proto neostýchají používat tyto orbitály, které jsou odvozeny z vodíkových, i pro ostatní atomy.}

Když sestavíme z atomů molekulu, změníme prostředí, ve kterém se elektrony pohybují. Nepřekvapí proto, že musíme hledat nové stavy elektronů. Tyto nové stavy budou popsané novými vlnovými funkcemi (novými orbitaly). Je tedy nutno znovu rozložovat \textit{atomové orbitaly} elektronů v izolovaných atomech a \textit{molekulové orbitaly} elektronů v molekulách, jinak bychom nemohli popsat vznik chemických vazeb.

Při hledání molekulových orbitalů pochopitelně narážíme na stejnou překážku, se kterou jsme se setkali u atomů s více elektronů: kvantové rovnice nelze přesně řešit. Můžeme opět použít starý trik a hledat řešení pro jediný elektron, zatímco ostatní elektrony nahradíme jejich elektrickým polem (Hartreeho-Fockova metoda). Molekuly jsou ale složitější než atomy tím, že neobsahují jen více elektronů, ale i více jader. Naštěstí jsou jádra mnohem těžší než elektrony a pohybují se daleko pomaleji. Tyto pomalé pohyby lze bez větších chyb zanedbat, když popisujeme rychlé hemžení elektronů. Abychom tento postup použít, musíme samozřejmě vědět, kde se nacházejí jádra v molekule. Jakmile známe polohy jader, můžeme vypočítat tvar elektrického pole, ve kterém se nacházejí elektrony a počítat jejich orbitály.

Pokud chceme spočítat neznámou strukturu, polohy jader atomů pochopitelně neznáme. V tom případě postupujeme tak, že provádíme kvantový výpočet opakovaně s různým rozmístěním jader atomů a hledáme, které rozmístění atomů je nejvýhodnější. Tento postup se nazývá \textit{optimizace geometrie}, jde vlastně o minimalizaci energie molekuly. O takovýchto minimalizácích budeme hovořit v sekci 6.2.3, takže řešení problému polohy jader zatím odložíme. V následujících řádcích budeme pro jednoduchost předpokládat, že polohy jader známe.

Vraťme se teď k hledání molekulových orbitalů (jak jsme si řekli, předpokládáme že polohy jader známe). Pokud se nám nechce hledat vhodný tvar pole ostatních elektronů pro Hartreeho-Fockovu metodu, můžeme využít toho, co už známe. Kvantová mechanika nás učí, že z určitého počtu atomových orbitálů můžeme získat stejný počet molekulových orbitalů \textit{lineární kombinací}. Srozumitelněji řečeno, jednotlivé vlnové funkce (atomové orbitály) vynásobíme vhodným koeficientem a pak je všechny sečteme. Abychom získali tolik molekulových orbitálů, kolik bylo atomových, musíme koeficienty srovnat s různými koeficienty tolikrát, z kolika atomových orbitálů vycházíme. Tato metoda, sdíle obyčejného chemika blízká, bývá nazývána \textit{metodou molekulových orbitálů jako lineárních kombinací atomových orbitalů (metodou MO-LCAO)}. Důvod její obliby je jednoduchý, můžeme použít dobré známé atomové orbitály a bez velkého počítání kvalitativně popsat chemické vlastnosti molekuly.

Chceme-li použít metodu MO-LCAO, musíme vědět, jakým způsobem orbitály sčítat a odčítat. Jako vodítko slouží tato obecná pravila:

- **Výsledný molekulový orbital musí odpovídat symetrii (tvaru) molekuly.**
- **Zpravidla stačí použít \textit{valenční} atomové orbitály (atomové orbitály popisující elektrony, které mají v daném atomu nejvyšší energii, jsou tedy největší vážány k jádru a nejochotněji mění svůj**

\footnote{Zjednodušení víceelektronového problému na jednoelektronový znamená, že se elektrony budou chovat, jako by o sobě navzájem „nevěděli“. Ve skutečnosti ovšem elektrony nejsou nezávislé, jejich výskyt v atomu je korelovan. Chyba, které se dopustíme tím, že tuto korelací zanedbáme, se projeví zvýšením energie o příspěvek známý korelační energie.}

\footnote{V literatuře se můžete setkat také s názvem \textit{metoda konzistentného pole} nebo \textit{metoda SCF} (z anglického self-consistent field). Toto označení vychází ze způsobu, jakým se hledá popis zprůměrovaného působení ostatních elektronů.}

\footnote{Toto zjednodušení se nazývá Bornova-Oppenheimerova aproximace.}

stav za tvorby chemické vazby).

- Atomové orbitaly, které sčítáme, musí mít vhodný tvar a vzájemnou orientaci, musí se co nejvíce překrývat.

- Sčítáme anebo odečítáme orbitaly, které popisují elektrony s přibližně stejnou energií.

Jako příklad si můžeme uvést hledání molekulových orbitalů vody (obrázek 6.2). Chceme získat orbitaly, které věrně zobrazí tvar molekuly vody. Molekula vody je symetrická vzhledem k otočení o 180° kolem osy, která přísluší úhlu H–O–H. Navíc má dvě zrcadlové roviny symetrie které se protínají ve zmíněné osce v jedné leží všechny tři atomy a druhá je na ní kolmá. Valenční atomové orbitaly, které máme k dispozici, jsou dva orbitaly 1s atomu vodíku, jeden orbital 2s atomu kyslíku a tři orbitaly 2p atomu kyslíku. Když tyto orbitaly přeložíme přes sebe tak, aby odpovídaly tvaru molekuly vody a aby se co nejvíce překryly, získáme šest molekulových orbitalů.

stejnou symetrii jako čtyřstěn. Pro látky jako karboxylové kyseliny nebo alkeny, ve kterých je uhlík obklopen jen třemi atomy v rozích pomyslného trojúhelníka, si připraví jinou sadu tří orbitálů, které mají požadovanou symetrii. Podrobnější popis postupu tvorby molekulových orbitálů pro biologicky zajímavý příklad peptidové vazby můžete najít v dodatku B.

### 6.2.3 Metody ab initio

Po kvalitativním popisu molekulových orbitálů přistupme k vlastnímu kvantitativnímu výpočtu orbitálů, které popisují biologicky zajímavé molekuly. Nejseriážejší způsob, jakým lze tyto výpočty provádět, je odmítnout všechny nápovědy, které nám nabízí experiment a zkušenosti popsané v literatuře. Takový výpočet se hrdě opírá jen o obecné zákony kvantové mechaniky a základní fyzikální konstanty a označuje se jako přístup ab initio.

Jak jsme si již řekli, přesný výpočet je možný jen pro nejjednodušší případy. Proto metody ab initio používají nejrůznějších zjednodušení a přibližně reší hledají pomocí výkonných počítačů. Metody ab initio můžeme použít k výpočtu energie molekuly v určité konformaci (takzvaný single-point výpočet), nebo k hledání energeticky nejvýhodnější konformace, případně k výpočtu jiných vlastností molekuly. Hledání stavu molekuly s nejnižší energií (minimalizace energie), které nás zajímá především, probíhá podle zásad uvedených v sekci 6.1.2.

1. Vypočteme energii v určité konformaci.
2. Vypočteme síly, které působí na atomy v dané konformaci. Sílu lze spočítat jako derivaci energie v daném bodě v různých směrech, tedy gradient.
3. Necháme působit vypočtené síly na atomy a vypočteme, jak tyto síly změní konformaci.
4. Znovu vypočteme energii a síly, teď již pro pozměněnou konformaci.

Během výpočtu ab initio je nutné počítat vlnové funkce (orbitaly), které popisují chování elektronů v molekule. V sekci 6.2.2 jsme si ukázali, že molekulové orbitály lze získat lineární kombinací atomových orbitalů. Nemusíme ovšem využívat z atomů atomové vlny, molekulové orbitaly můžeme vypočítat lineární kombinaci nejrůznějších sad orbitalů. Taková sada funkcí, jejíž kombinací lze získat požadovaný orbital, se matematicky nazývá báze. Volba vhodné báze funkcí je pro výpočty velmi důležitá.

Nejpřirozenější bázi funkcí pro výpočty ab initio jsou staré dobré atomové orbitaly. Jde o vlnové funkce, které exponentiálně klesají se vzdáleností od jádra. Funkce vypočítané z rovnic kvantové mechaniky jde o exponenciální závislost na vzdálenosti r

\[ r = a_0 + a_1 r + a_2 r^2 e^{-br} \]

kde \( a_0, a_1, a_2, b \) jsou konstanty. Pro vytvoření báze se mnohočleny mnohočleny vznášejí vlnovou funkcí jakýkoliv exponentiální křivku. Samozřejmě tak získáme vlnové funkce na hory vzdálené atomovým orbitalům a nemůžeme očekávat, že tyto funkce budou dobře popisovat naši molekulu. Když ale sečteme několik Gaussových křivek, získáme křivku, která má tvar velmi podobný exponentiále. Takto dospějeme k popisu molekuly, který je vhodný pro výpočty, za cenu, že báze obsahuje větší počet funkcí zvaných

7Od začátku.
8Metody ab initio ovšem nejsou omezeny na konformační změny, můžeme pomocí nich popsat i chemické reakce.
9Přesněji řečeno, u funkcí vypočítačních z rovnic kvantové mechaniky jde o exponentiální závislost na vzdálenosti r.

Převážně výpočet ab initio se provádí pomocí vznášenou mnohočlen, který obsahuje celočíselné mocniny vzdálenosti (například \( (a_0 + a_1 r + a_2 r^2 e^{-br}) \)). Nám bude stačit, když se zaměříme na exponentiální část zjednodušených funkcí.
6.2. KVANTOVÉ METODY

Hustoty, protože energii vyjadřujeme jako funkci elektronové hustoty, tajemné elektrické pole, které teorie DFT zavádí. Pokud chceme získat i vlnovou funkci, musíme ji tronu v prostoru (takovou funkci elektronové hustoty), která odpovídá nejnízší energii elektronu v tom molekuly a jiné užitečné veličiny. Metoda DFT postupuje opačně. Hledáme proto takové rozložení elektronů jako polarizací funkce uhlíku, zatímco báze 6-31G∗∗ přidává navíc orbitaly p jako polarizační funkce vodíku.


Na první pohled se zdá, že jsme si moc nepomohli. Stejně jako v Hartreeho-Fockově metodě musíme složité dodatečně zahrnovat to, že elektrony o sobě „vědí“. Je tu a zásadní rozdíl. Ono tajemné elektrické pole metody DFT je jen jedno pro celou molekulu, stejně pro všechny elektrony. Opravy tedy nemusíme provádět způsobem „každý s každým“, jako v Hartreeho-Fockově metodě, ale jenom jednou pro každý elektron. Pomocí metod DFT lze proto dnes provádět ab-initio kvantové výpočty i s několika desítkami atominů.

Dalším důležitým rozdílem mezi Hartreeho-Fockovou metodou a metodou DFT je způsob, jakým získáme kžený výsledek – energeticky vzhodné rozložení elektronů v prostoru, které udává tvar molekuly. Hartreeho-Fockova metoda je založena na tom, že z vlnových funkcí můžeme spočítat energii i rozložení elektronů v prostoru. Hledáme proto takovou vlnovou funkci, která odpovídá nejnižší energii. Z této vlnové funkce pak můžeme vypočítat rozložení elektronů (elektronovou hustotu) v základním stavu molekuly a jiné užitečné veličiny. Metoda DFT postupuje opačně. Hledáme přímě takové rozložení elektronů v prostoru (takovou funkci elektronové hustoty), která odpovídá nejnižší energii elektronů v tom tajemném elektrickém poli, které teorie DFT zavádí. Pokud chceme získat i vlnovou funkci, můžeme ji vypočítat zpětně z elektronové hustoty. Tím se dostáváme i k názvu metody. Mluvíme o funkcionálu hustoty, protože energii vyjadřujieme jako funkci elektronové hustoty.

---

10Orbitaly, které odpovídají polarizačním funkcím, mají vyšší veďejší kvantové číslo l, než odpovídá zaplněným orbitalům daného atomu.

11Funkce je matematický předpis, jak přiřadit hodnotu nějaké veličiny (v našem případě energie) určitým hodnotám proměnných (v našem případě souřadnic). Funkcionál je matematický předpis, jak přiřadit hodnotu nějaké veličiny (v našem případě energie) určité funkci jako celku (v našem případě funkce elektronové hustoty).
Vzdálenost od jádra $r$

$g_1 = c_1 e^{-k_1 r^2}$
$g_2 = c_2 e^{-k_2 r^2}$
$g_3 = c_3 e^{-k_3 r^2}$

$g_1 + g_2 + g_3$

6.3. MOLEKULOVÁ MECHANIKA

6.2.4 Semiempirické metody

Náročnost metod ab initio vedla k mnoha snahám o zjednodušení výpočtu. Při použití těchto metod se vzdáváme snahy o fyzikálně přesný popis molekuly a raději hledáme cestičky, jak některou část výpočtu nahradit dosazením nějakého experimentálního údaje, nebo jak se určitěmu výpočtu zcela vyhnout.

Prvním zjednodušením semiempirických metod je, že výpočet omezují jen na elektronu s nejvyšší energií, které se podílejí na tvorbě chemických vazeb (valenční elektrony). Druhým zjednodušením je, že se omezíme na elektrony s nejvyšší energií, které se podílejí na tvorbě chemických vazeb (valenční elektrony).

Další zjednodušením se týká překryvů atomových orbitalů, ze kterých vyrábíme orbitaly molekulové. V sekci 6.2.2 jsme si řádili, že pro tvorbu molekulových orbitalů (pro tvorbu chemických vazeb) je nutný co nejlepší překryv atomových orbitalů. V semiempirických metodách se některé méně významné překryvy nedělají. Podle toho, co všechno zanedbáme, mluvíme o metodách NDDO, INDO, CNDO. Nejpatrnější je metoda NDDO, naopak nejdrastičtější je metoda CNDO.

Zanedbáním překryvů se dopouštíme chyby v výpočtu. Tuto chybu můžeme trochu napravit tím, že nahradíme číselné hodnoty překryvů nějakým opravným parametrem. Tento opravný parametr můžeme buď odhadnout z tvaru atomových orbitalů, nebo vypočítat z výsledků různých experimentů.

Zvláštní roduin metod tvoří takzvané Hückelovy metody, které zavádějí opravné parametry ne místo vypočítaných překryvů, ale přímo místo hodnot potenciální energie v rovnicích kvantové mechaniky. Parametry v Hückelových metodách vlastně odpovídají ionizačním energiím atomů (energiím, které jsou potřeba k odtržení elektronu z atomu). Proto je hledána parametře v tomto případě velmi snadné a takzvaná rozšířená Hückelova metoda (EHM, Extended Hückel Method) poskytuje parametry pro všechny prvky periodické tabulky. EHM jsou nejlepší pro výpočty vlastností molekul obsahujících atomy přechodných prvků, naopak se nepříliš hodí pro optimalizaci tvaru molekuly.

Semiempirické metody umožňují kvantové výpočty pro molekuly skládající se ze stovek atomů. Je možné vypočítat molekulové orbitaly, energie jednotlivých stavů, rozložení náboje v molekule, řady vazeb a spektrální vlastnosti molekul, které zmíněme v kapitole 10. Semiempirické metody mohou za-
hrnovat i vliv rozpouštědla na chování molekuly. Další z jeho výhod je, že je možné vypočítat i vliv rozpouštědla na chování molekuly. Další z jeho výhod je, že je možné vypočítat i vliv rozpouštědla na chování molekuly.

Semiempirické metody umožňují kvantové výpočty pro molekuly skládající se ze stovek atomů. Je možné vypočítat molekulové orbitaly, energie jednotlivých stavů, rozložení náboje v molekule, řady vazeb a spektrální vlastnosti molekul, které zmíněme v kapitole 10. Semiempirické metody mohou za-
hrnovat i vliv rozpouštědla na chování molekuly. Další z jeho výhod je, že je možné vypočítat i vliv rozpouštědla na chování molekuly. Další z jeho výhod je, že je možné vypočítat i vliv rozpouštědla na chování molekuly. Další z jeho výhod je, že je možné vypočítat i vliv rozpouštědla na chování molekuly. Další z jeho výhod je, že je možné vypočítat i vliv rozpouštědla na chování molekuly. Další z jeho výhod je, že je možné vypočítat i vliv rozpouštědla na chování molekuly.

### 6.3.1 Hyperplocha potenciální energie

Vzájemné působení atomů (ve skutečnosti zprostředkováno elektrony) napodobujeme tím, že atomy různě spojíme ideálními pružinkami (obrázek 6.4). Jak si jistě vzpomenete ze střední školy, energii dvou kuliček spojených ideální pružinou můžeme popsat rovnicí \( E = K_r(r - r_0)^2 \), \( \text{(6.1)} \)

kde \( r \) je okamžitá vzdálenost kuliček, \( r_0 \) je rovnovážná vzdálenost kuliček a \( K_r \) je tuhost pružiny.

V molekulové mechanice stačí dosadit za \( r_0 \) optimální vazebnou vzdálenost a máme energetický popis dvojice atomů spojených kovalentní vazbou.

Energii spojenou s deformací vazebných úhlů můžeme popsat podobně (obrázek 6.5)

\[ E = K_\theta (\theta - \theta_0)^2, \]

\( \text{(6.2)} \)

kde \( \theta \) je okamžitá hodnota vazebného úhlu a \( \theta_0 \) je nejvýhodnější hodnota vazebného úhlu.

---

\(^{14}\)Presněji dvou hmotných bodů, tedy kuliček nekonečně malých.

\(^{15}\)Z pohledu matematiky se taková pružina chová jako harmonický oscilátor.
6.3. MOLEKULOVÁ MECHANIKA

Obrázek 6.6: Příklad zjednodušeného popisu energie spojené se změnou torzního úhlu.

Na torzní úhel už s pružinkou nevystačíme. Musíme se ptát, jak se asi bude měnit energie, když budeme točit kolem chemické vazby. Na základě jednoduchých úvah dospějeme k rovnici

\[
E = \frac{V_0}{2} (1 + \cos n(\phi - \phi_0)),
\]

kde \(V_0\) je energie nejméně výhodné konformace, \(n\) požadovaný počet minim energie, \(\phi\) okamžitá hodnota torzního úhlu a \(\phi_0\) optimální hodnota torzního úhlu (obrázek 6.6).


Umělé vztahy představené rovnicemi 6.1 až 6.3 molekulová mechanika kombinuje s elektrickými interakcemi mezi atomy, které jsme popsalí v sekci 1.2.2. Elektrické síly můžeme popsat rovnicemi shrnutými v dodatku A. Nejčastěji se používá elektrostatická energie dvou iontů, která klesá s jejich vzdáleností (rovnicí 13.1) a energie van der Waalsových sil, které klesají s řadou mocnin vzdálenosti atomů (rovnicí 13.5 a 13.6). Mezi atomy navíc působí velmi silné odpudivé síly, které brání jejich kontaktu. Tyto síly klesají s dvanáctou mocninou vzdálenosti atomů. Silová působení popsané v tomto odstavci můžeme shnout do rovnice 6.4.

\[
E = \frac{A}{r^{12}} - \frac{B}{r^6} + \frac{C Q_1 Q_2}{r},
\]

kde \(Q_1\) a \(Q_2\) jsou náboje atomů, \(r\) jejich vzdálenost a \(A, B, C\) konstanty úmernosti.

Všimněme si, že energie vyjádřená v rovnicích 6.1 až 6.4 vždy závisí buď na vzdálenosti atomů, nebo na vazebném či torzním úhlu – prostě na nějakých souřadnicích atomů. Kdybychom však jen dvě takové

16Kdybychom měli například dvě velké skupiny spojené disulfidovou vazbou, budeme očekávat, že nejvýhodnější bude torzní úhel \(\phi = \pm 180^\circ\) (trans) a nejméně výhodný torzní úhel \(\phi = 0^\circ\) (cis). Takový průběh energie by dobře popsal funkcí \(\frac{1}{2}(1 + \cos \phi)\). V případě ethanu budou nejvýhodnější torzní úhly \(\phi = \pm 180^\circ\) (trans) a \(\phi = \pm 180^\circ\) (gauche) a nejméně výhodný torzní úhly \(\phi = \pm 180^\circ\) a \(\phi = 0^\circ\). Tomuto průběhu by odpovídala nejlépe funkce \(\frac{1}{2}(1 + \cos 3\phi)\).

17Constraint je donucení, přesné vymezení.

18Restraint je omezení, nátlak.
souřadnice, třeba úhly $\phi$ a $\psi$ peptidové vazby, mohli bychom si přímo graf energie jako funkce souřadnic zobrazit. Na trojrozměrném grafu, jehož základnou by byl Ramanchandranův diagram, bychom viděli zprohýbanou plochu připomínající krajinu. Údolí by odpovízala výhodným konformacím s minimální energií a pohoří by představovala maximu energie (obrázek 6.7). Tato plocha nad Ramanchandranovým grafem se nazývá hyperplocha potenciální energie (zkratka $PES$ z „potential energy surface“). V makromolekule máme souřadnic mnohem více, takže plochu bychom museli kreslit ne nad dvourozměrným diagramem, ale nad mnoharozměrným prostorem (takzvaný $konformační$ $prostor$). To sice neumíme, ale matematickému určení takové nepředstavitelné hyperplochy a hledání jejích minín nic nebrání.

6.3.2 Silové pole


Soubor rovnic (6.1) a (6.4) a číselné hodnoty parametrů $K$, $V$, $A$, $B$, $C$ pro seskupení atomů, která nacházíme v biomakromolekulách, se nazývají silové pole a představují srdce každé metody moleku-
6.3. MOLEKULOVÁ MECHANIKA


6.3.3 Implicitní a explicitní solvent

Zjednodušený popis energie umožňuje molekulové mechanice výpočty nejen pro celé biomakromolekuly, ale i pro nejblíží okolí těchto molekul. Táh je možno studovat biologicky zajímavé molekuly v jejich přirozeném prostředí, ve vodném roztoku. Tento přitažlivý úkol se s kooperačně nese otázku, jakým způsobem vodu do výpočtu zahrnout. Vliv rozpouštědla se popisuje jednou ze dvou metod. První z nich, zvaná metoda implicitního solventu, je výpočtěm velmi snadná. Využívá toho, že síla elektrických interakcí je nepříměm úměrná elektrické permitivitě prostředí. Nejjednoduššího přibližení vodním prostředí je možné dosáhnout dosazením správné hodnoty elektrické permitivity. Používá se buď konstantní, vůči stěně konstantní, nebo permitivita závislá na vzdálenosti atomů.


6.3.4 Metoda PME


Rozumný kompromis mezi přesností a rychlostí výpočtu nabízí metoda PME (Particle Mesh Ewald). Tato metoda je vhodná zejména pro výpočty nukleových kyselin, v nichž každý nukleotid obsahuje záporné nabité fosfát a elektrostatické síly se proto projevují velmi výrazně.

Metoda PME zjednodušuje výpočet tím, že elektrostatické interakce počítá pro jednotlivé molekuly vody jen do určité vzdálenosti od biomakromolekuly. Ve větších vzdálenostech elektrostatické interakce nezanedbává, ale jejich energii počítá nikoli jako funkci vzdálenosti, ale jako funkci převážné hodnoty vzdálenosti. Údaj tomu, že naše molekuly jsou umístěny v periodicky se opakujících krabiciích, můžeme pro výpočty veličin závislých na převážné hodnotě vzdálenosti použít velmi rychlý výpočetní metodologii pro Fourierovu transformaci.
Obrázek 6.8: Periodické okrajové podmínky pro výpočet struktury proteinu s explicitním solventem (znázorněn červenými tečkami).
6.3. MOLEKULOVÁ MECHANIKA

6.3.5 Molekulová dynamika

Minimalizace energie má jedno důležité omezení – „umí běhat jen z kopce“. To znamená, že metoda dospěje vždy na dno tohoto údolí hyperplochy potenciální energie, ve kterém se už od začátku nachází. Nás ovšem zajímají všechna údolí, zejména to nejhlubší. Proto se minimalizace často kombinuje s metodami, které nás umí přenést i přes pohoří potenciální energie. Tyto metody připomínají postup detektiva, který hledá ve městě podezřelé. Může k tomu přistupovat různým způsobem:

1. Detektiv může postupovat systematicky, dům od domu. Tomu odpovídá metoda systematického prohledávání (grid search). Jako příklad si můžeme vzít opět Ramanchandranův diagram. Budeme systematicky zvyšovat úhly $\phi$ a $\psi$ po pěti stupních a pro každou dvojici úhlů vypočítáme energii. Abychom prohledali celý diagram, musíme takových výpočtů provést 72 krát 72, to je 5 184.

2. Detektiv může zcela náhodně brát za kliky různých domů. Podobně my můžeme volit hodnoty úhlů $\phi$ a $\psi$ náhodně, hodit si vždycky kostkou nebo zatačit ruletou. Celkem trefně se tyto metody označují Monte Carlo.


4. Detektiv může sledovat pohyb ve městě a čekat, až pachatel sám dorazí na místo činu. Takovým způsobem funguje molekulová dynamika, kterou se v této části budeme zabývat především.

Molekulová dynamika je metoda napodobování pohybů v molekule. Pomocí této metody můžeme nejen prohledávat konformační prostor, ale studovat i samotné molekulové pohyby.

Věrna heslu molekulové mechaniky, molekulová dynamika nevychází z pohybové rovnice kvantové chemie (Schrödingerova rovnice), ale používá jednodušího popisu klasickými Newtonovými pohybovými rovnicemi. Postup výpočtu probíhá podle následujícího scénaře.

1. Metoda začíná s modelem molekuly, který je popsán nějakými počátečními souřadnicemi atomů v molekule.

2. Silové pole nám řekne, jaká je energie celé molekuly v místě určeném danými souřadnicemi atomů.

3. Vypočítáme síly, které působí na jednotlivé atomy (síla je derivace energie podle souřadnice).


5. Nové souřadnice použijeme místo počátečních v kroku 1 a celý postup opakujeme. Síly, zrychlení, rychlosti a souřadnice se mění po skokcích (celý postup se označuje jako leap-frog algorithm).

---

19V praxi asi použijeme funkci počítače, která generuje pseudonáhodná čísla.
20Takto postupuje program CICADA.
21Problémem tohoto přístupu je, že detektiv místo činu dopředu nezná. Může sedět za komínem a rozhlížet se, kde se co děje, část výhledu mu ale budou zakrývat nedaleké mrakodrapy. Pokud bude pachatel koupit v jejich skrytu, zůstane neodhalen.

V kroku jsme zmínili, že výpočet zahajujeme s nějakými počátečními souřadnicemi. Tyto počáteční souřadnice, které definují startovní konformaci, musíme odhod vit vzát. Za počáteční konformaci můžeme zvolit buď konformaci, kterou už známe z experimentu (třeba nepřesně), nebo náhodně sbalenou molekulu.


Konformaci můžeme sledovat prostřednictvím očí, když si co chvíli (zpravidla po několika stovkách časových kroků) souřadnice atomů zapíšeme a necháme počítač následně model struktury. Každý obrázek modelu představuje náhodnou makromolekulu v určitém čase, spojeném těchto snímků vznikne film pohybující se makromolekulou. Jiným způsobem popisu trajektorie je graf, do kterého vynášeme, jak se mění některý údaj o molekule v čase. Tímto údajem může být energie nebo střední kvadratická odchylka určitých souřadnic.

Pohyblivost molekuly jsme si v úvodu této sekcí zavedli proto abychom se mohli pohybovat v celém konformaciálním prostoru, nejen směrem k nejléžšímu údolí. Dynamika molekul, tak jak jsme si ji v předchozích odstavcích popsali, za rozumných teplot nestáčí k překonání vysokých energetických pohrůží. Proto se často provádí takzvané simulované žhání molekul, znázorňené na obrázku. To znamena, že se teplota zvýší na obrovskou hodnotu, často tisíce Kelvinů. Atomy se při této teplotě pohybují tak rychle, že jejich kinetická energie je vysoko než jakékoli pohrůží na hyperploše potenciálové energie. S takovou energií můžeme nad hyperplochou volně poletovat. Potom molekulu postupně ochlazujeme, kinetickou energii klesá, až „přistane“ v nějakém údolí na hyperploše, často daleko od startovní konformace. Za biologicky přijatelných teplot se pak molekula pohybuje v rámci nového údolí, jehož dno můžeme nalézt minimalizací energie. Když takový výpočet opakujeme s místně odlíněnými startovními podmínkami, 

\[ \text{rychlost} \sim \sqrt{\frac{2k_B T}{m}} \]
Obrázek 6.9: Princip metody simulovaného žíhání. Při zahřátí na vysokou teplotu (červená čára) získá molekula větší kinetickou energii než nejvyšší hradby potenciální energie a může dospět do konformace odpovídající globálnímu minimu energie. Naopak při laboratorní teplotě (azurová čára) může molekula dospět pouze do konformace odpovídající lokálnímu minimu.

je velká naděje, že skončíme v různých údolích. Takto můžeme prohledat velkou část konformačního prostoru a snad najít i konformaci s nejnižší energií.
Kapitola 7

Příprava vzorku biomakromolekul

7.1 Izolace z přírodního materiálu

Zpracování přírodního materiálu (rozbitých buněk, tělních tekutin a nejrůznějších produktů, které organismus vylučuje) je nejstarším způsobem získávání biomakromolekul. Výhodou tohoto postupu je snadná dostupnost větších množství takového materiálu. Nevýhodou je složitost směsi, ze které chceme naši biomakromolekulu získat.

Přírodní vzorek, z kterého chceme získat jednu čistou látku, zpravidla obsahuje tisíce jiných látek. Získání jediné požadované je nutno navrhnout vhodný postup postupně čištění (purifikační schéma). Takový postup zahrnuje vhodnou kombinaci různých separačních technik (metod dělení látek).

Zmíníme tedy pouze obecné přístupy k purifikaci různých biologicky zajímavých molekul.

Pro čištění složitých směsí biologických látek byla vyvinuta řada separačních metod. Tyto metody dělí látky na základě různé interakce molekul směsí s vnějším polem nebo s molekulami přítomnými v separačním zařízení. Volba metody závisí na množství látky, které chceme získat. K určení sekvence postačují nanogramová i menší látky, zatímco určení prostorové struktury vyžaduje několik miligramů čisté látky. Obecně lze říci, že v prvních krocích purifikačního postupu se používají metody, které silně nedělí látky příliš účinně, ale mohou zpracovat velké množství vzorku. V závěru napak přichází ke slovu techniky s omezenou kapacitou (množstvím vzorku, které lze zpracovat), ale vysokou účinností dělení. Navíc je třeba kombinovat metody, které dělí látky podle různých vlastností (například podle velikosti molekul, sily iónových interakcí a velikosti hydrofobního efektu) – to se týká zejména proteinů.

Jako příklad purifikaciho schematu proteinu můžeme uvést postup zahrnující srážení sranem amonným, iontoměničovou, hydrofobní a vylučovací chromatografií (viz kap. 7.1). Malá množství vysoce čistého proteinu pro určení sekvence můžeme získať elektroforézou nebo chromatografií na reverzní fázi v kapilárním provedení. Po určení částečné sekvence je pak možné vyhledat gen příslušného proteinu a větší množství vzorku připravit metodami molekulární biologie (viz sekce 7.2).

Nukleové kyseliny se obvykle získávají iontoměničovou chromatografií nebo extrakcí do směsi chloroformu a fenolu a následným alkoholovým srážením. K dělení DNA podle velikosti molekul se používá elektroforéza v gelu. Díky párování bází lze izolovat nukleovou kyselinu, která obsahuje pouze určitou sekvenci nukleotidů. Takto izolovanou DNA obvykle vnášíme metodami molekulární genetiky (viz kap. 7.2) do jiného organismu, se kterým se snáze pracuje. Jinou možností je určit sekvenci z malého množství nukleové kyseliny a připravit větší množství chemicky (viz sekce 7.3).

Srážení a extrakce slouží i k hrubému přečištění lipidů a polysacharidů. K jemnějšímu dělení slouží opět chromatografie na reverzní nebo normální fázi. Polysacharidy se často dělí pomocí iontoměničové chromatografie nebo podle velikosti vylučovací chromatografie. Protože stavba sacharidů a biomembrán není přímo geneticky kodována, jsou na studium přírodního materiálu odkázáni mnohem více, než v případě proteinů a oligonukleotidů.

Lipidové dvojvrstvy můžeme připravit poměrně jednoduše smícháním příslušných lipidů. V takovém případě ale nejde o biomembrány v pravém slova smyslu, ale spíše o modelové útvary, které postrádají membránové proteiny a jiné složky. Studium procesů, které se na membránách odehrávají, proto vyžaduje připravu dvojvrstvy obsahující potřebné membránové proteiny. Ty jsou obvykle izolovány z přírodního materiálu s použitím vhodného detergentu, který umožní jejich rozpuštění (solubilizaci).

7.2 Metody molekulární biologie

7.2.1 Principy molekulární genetiky

V počátcích molekulární biologie se zájem badatelů soustředil na pochopení toho, jak buňka předává a zpracovává dědičnou informaci (na molekulární genetiku). Tento cíl vyžadoval nejprve zvládnutí experimentální práce s molekulou, která genetickou informací nese, s DNA. Rozvoj metod práce s nukleovými kyselinami umožnil, že molekulárně biologické techniky jsou dnes využívány nejen k připravě vzorků nukleových kyselin, ale i vzorků proteinů. Metody genového inženýrství umožňují připravu velkého množství proteinu v cizím organismu (se kterým se snadněji pracuje), zavedené afinity známky (která dramaticky zjednodušuje purifikaci) a určitou manipulaci s proteinem (jako zámenu isotopů pro potřeby nukleární magnetické rezonance nebo zámenu látek za selen využívaný v rentgenové strukturní analýze).

Metody molekulární biologie vycházejí ze způsobu, jakým se předává genetická informace v buňce. Tento soubor složitých dějů si můžeme shrnout do několika bodů. Zaměříme se přitom na bakterie, se kterými se pracuje nejméně.

- Nositělem genetické informace je molekula DNA. Většina dědičné informace je v buňce přítomna v jedné kopii ve formě jedné nebo více dlouhých dvojšírovacích, zvaných chromozomy. Chromozom bakterie Escherichia coli, se kterou molekulární biologové pracují nejčastěji, je kruhová molekula obsahující čtyři miliony nukleotidů. V bakteriálních buňkách je navíc malá část genetické informace přítomna ve formě krátkých kruhových molekul DNA, zvaných plasmidy. Obvyklá délka plasmid je několik tisíc nukleotidů.

- Molekula DNA je z chemického hlediska souvislý řetězec nukleotidů. Z biologického hlediska si ale tento řetězec můžeme rozdělit na kratší úseky, z nichž každý nese informaci o jedné vlastnosti.

1K určení sekvence není třeba zachovat přirozenou konformaci.
7.2. METODY MOLEKULÁRNÍ BIOLOGIE

Buňky. Tyto úseky se nazývají geny. Geny jsou části informace, nikoli struktury, sekvence nukleotidů kódující jeden gen mohou být několikrát přerušeny sekvencemi, které nesouvisí s informací daného genu (takové sekvence se nazývají introny), nebo může určitý úsek DNA kódovat dva různé překládající se geny.

- Během buněčného dělení se vytvoří kopie (replika) chromozomu, takže obě dcerinné buňky obsahují po jedné molekule chromozomové DNA. Tento proces se nazývá replikace.

- Buňka si svou dědičnou informaci důsledně uchovává a předává následujícím generacím. Za určitých okolností se ale mohou buňky genetickou informací vyměňovat. Je-li například život bakterie ohrožen nějakým antibiotikem, začne bakteriální buňka ochotně přijímat plasmidy v naději, že některý z plasmidů bude obsahovat gen, který ji před antibiotikem ochrání (gen rezistence). Popsaná změna genetické informace se nazývá transformace buňky.

- Proces, kterým se určitá dědičná vlastnost projeví, se nazývá exprese genu. Nejběžnějším případem je syntéza proteinu kódovaného určitým genem (exprese proteinu).

- Má-li dojít k expresi, musí být nejdříve příslušný gen „přepsán“ do molekuly RNA. Tento proces interpretace genetického kódu se nazývá translace. K vlastnímu dekódování dochází během takzvané aktivace aminokyseliny, kdy je každý aminokyselinový zbytek navázán na příslušnou kódující sekvenci nukleotidů (introny) procesem zvaným sestřih.

- Během exprese proteinů je genetická informace přeložena podle kóduvící tabulky ze sekvence nukleotidů do sekvence aminokyselin. Tento proces interpretace genetického kódu se nazývá translace. K vlastnímu dekódování dochází během takzvané aktivace aminokyseliny, kdy je každý aminokyselinový zbytek navázán na příslušnou kódující sekvenci nukleotidů (introny) procesem zvaným sestřih.

- Může dojít k expresi, musí být nejdříve příslušný gen „přepsán“ do molekuly RNA. Tento proces interpretace genetického kódu se nazývá translace. K vlastnímu dekódování dochází během takzvané aktivace aminokyseliny, kdy je každý aminokyselinový zbytek navázán na příslušnou kódující sekvenci nukleotidů (introny) procesem zvaným sestřih.

- Během exprese proteinů je genetická informace přeložena podle kóduvící tabulky ze sekvence nukleotidů do sekvence aminokyselin. Tento proces interpretace genetického kódu se nazývá translace. K vlastnímu dekódování dochází během takzvané aktivace aminokyseliny, kdy je každý aminokyselinový zbytek navázán na příslušnou kódující sekvenci nukleotidů (introny) procesem zvaným sestřih.

- Má-li dojít k expresi, musí být nejdříve příslušný gen „přepsán“ do molekuly RNA. Tento proces interpretace genetického kódu se nazývá translace. K vlastnímu dekódování dochází během takzvané aktivace aminokyseliny, kdy je každý aminokyselinový zbytek navázán na příslušnou kódující sekvenci nukleotidů (introny) procesem zvaným sestřih.

- Během exprese proteinů je genetická informace přeložena podle kóduvící tabulky ze sekvence nukleotidů do sekvence aminokyselin. Tento proces interpretace genetického kódu se nazývá translace. K vlastnímu dekódování dochází během takzvané aktivace aminokyseliny, kdy je každý aminokyselinový zbytek navážen na příslušnou kódující sekvenci nukleotidů (introny) procesem zvaným sestřih.

7.2.2 Expresní vektory

Studium eukaryotních proteinů často probíhá podle naznačeného schematu „od genu k proteinu“. Jak bude zmíněno v sekci , často se i u proteinu nejprve izolovaného z přirozeného zdroje postupuje „od proteinu přes gen k proteinu“. Eukaryotní gen se (ve formě odpovídající již sestřihané mRNA) zavede do molekuly DNA schopné nezávislé replikace a exprese (takzvaného vektoru, nejčastěji bývá vektorom plasmid). Tím se vytvoří takzvaný expresní konstrukt, kterým se transformuje vhodný kmen Escherichia coli. Dnes jsou k dispozici kmeny, se kterými se pohodlně a bezpečně pracuje a které jsou ochotné podle dodané genetické informace vyrábět velká množství proteinů pro E. coli zcela cizích.

K přesné definovánímu vyjmoucí úsek DNA se používají zvláštní enzymy zvané restrikční endonukleasy, které specificky rozpoznávají krátké sekvence (nejčastěji šesti nukleotidů). Vektory, do kterých úsek DNA vkládáme jsou uměle připraveny tak, aby obsahovaly:

1. větší počet míst rozpoznávaných jednotlivými restrikčními endonukleasami,
2. sekvence nezbytné pro replikaci a expresi vektoru v bakterii (sekvence určující počátek replikace, bakteriální promotor, což je sekvence rozpoznávána jako začátek transkripcie, místo vazby na ribozomy),
3. sekvenci kódující takzvanou afinity na značku (několik aminokyselinových zbytků, obvykle histidinových, které se pevně váží k určitému chromatografickému nosiči),
4. selektivní značku, které umožní rozoznat buňky, do kterých opravdu pronikl použitý vektor. Nejčastěji se používá gen rezistence k vybranému antibiotiku. Bakterie se pak pěstují v prostředí obsahujícím toto antibiotikum, takže přežijí pouze ty buňky, které přijaly expresní vektor.

7.2.3 Molekulové klonování

Pro úspěšnou expresi proteinu je nezbytné získat velké množství potomků jediné buňky, jejíž genetickou informaci jsme upravili zavedením genu pro nás protein. V takovém společenství, které se nazývá klon, nesou všechny buňky požadovanou genetickou informaci (gen našeho proteinu).

Základní postup molekulového klonování zahrnuje tři kroky:

1. Příprava genu, který chceme do bakterie vložit

Předpokládejme, že máme k dispozici (například v jiném plasmidu) genu proteinu, odpovídající již sestřižené mRNA. Tento gen musíme rozmnožit způsobem, který napodobuje replikaci v živé buňce. Dnes se nejčastěji používá takzvaná polymerasová řetězová reakce (PCR z anglického polymerase chain reaction), kde celý proces provádíme ve zkumavce a řídíme jej pouze změnami teploty reakční směsi.

Využíváme přitom zvláštní vlastnosti enzymu, který replikaci provádí. Tento enzym neumí totiž syntézu nového vlákna zahájit, potřebuje již kousek dvojšroubovice, kterou dále umí prodlužovat. Proto musí být ve směsi kromě dlouhého vlákna DNA (obsahující gen) i takzvané primery – krátké úseky DNA, jejichž báze se párují s konci dlouhého vlákna a vytvářejí tak potřebné zárodky dvojšroubovice.

V praxi tedy postupujeme tak, že chemicky nasynetizujeme (nebo si necháme na objednávku vyrobit) dva krátké oligonukleotidy, které nám budou sloužit jako primery. Nesmí být příliš krátké, to by se nevytvorily stabilní zárodky dvojšroubovice, naopak syntéza dlouhých primer˚ u by byla dražší. Vhodná délka je asi 18 nukleotid˚ u. Primery nám umožňují nejen vybrat si, který úsek DNA chceme replikovat, ale také zavést několik koncových nukleotid˚ u podle naší potřeby. Toho využíváme k zavedení sekvence, které rozpoznají námí zvolené restrikční endonukleasy.

Připravíme reakční směs obsahující DNA s genem, primery, volnými nukleotidy (ve form˚ e trifosf˚ at˚ u, které vyžaduje replikační enzym), enzymem a dalšími látkami potřebnými pro dobrý průběh reakce.

Postupně méníme teplotu reakční směsi. Při ochlazení se spojí primery s dlouhým vláknem DNA, při optimální teplotě dochází k syntéze nového vlákna DNA podle starého (obě vlákná vytvoří dvojšroubovice), při zahřátí se dvojšroubovice rozpadne na dvě vlákná. V dalším kroku tedy už začínáme se dvěma vlákná a po stejném teplotním cyklu dostaneme vlákná čtyři, v příštím kroku osm, pak šestnáct, a tak dále geometrickou řadou.

Nakonec provedeme štěpení DNA, kterou chceme do vektoru vložit, restrikčními endonukleasami. Po štěpení restrikčními endonukleasami vznikají krátké přesahující jednovláknové sekvence, kterými se později dvě molekuly DNA spojí na základě párování bází.

2. Vložení genu do vektoru


Po smícháním vektoru s genem dojde k párování konců obou molekul DNA. Molekuly ovšem dosud nejsou kovalentně spojeny. Dalším krokem je proto vytvoření fosfodiesterových vazeb mezi genem a vek- torem pomocí specifických enzymů. Tím vznikne expresní konstruk, který můžeme vložit do bakterie.
7.2. METODY MOLEKULÁRNÍ BIOLOGIE

Transformace


Transformované buňky necháme množit na živném půdě (mědu) s antibiotikem, proti kterému jsou rezistentní buňky obsahující použitý vektor. Bakteriální kulturu potom rozetrem na agarovou plotnu tak, aby vyrostly oddělené kolonie bakterií. Každá taková kolonie je tvořena potomky jedné buňky, jde tedy o jeden klon. Analýzou DNA jednotlivých kolonií zjistíme, které opravdu obsahují vektor se začleněným genem.

Kolonii pak pipetou nebo páratkem přeneseme do živné půdy a necháme rozmnožit. Následnou selekcí se vyberou klony buněk, které opravdu obsahují vektor se začleněným genem.

7.2.4 Místně řízená mutageneze

V některých případech nechceme zkoumat pouze protein přítomný ve studovaném organismu, ale chceme jej nějakým způsobem pozměnit. Takových změn genetické informace, nebo-li mutací, lze dosáhnout pomocí metod řízené mutageneze.

Genetickou informaci naklonovanou do vhodného vektoru lze pozměnit zavedením mutací na přesně daném místě. Běžné strategie místně řízené mutageneze lze rozdělit na dva druhy.

1. První postup je založen na vyštěpující dvojřetězecového fragmentu DNA v místě, které má být mutováno (přesněji řečeno mezi nejbližšími místy rozpoznávanými restrikčními nukleasami) a na nahrazení tohoto fragmentu uměle připraveným fragmentem obsahujícím požadovanou mutací.

2. Druhý postup využívá toho, že dostatečně dlouhý (asi 18 párů bází) úsek dvojšroubovice může být účinné párován i v případě, že jeden pár bází není komplementární (netvoří pár dle Watsona a Cricka). Vektor s naklonovaným genem se rozdělí na jednotlivá vlákna a jedno z nich se nechá párovat s uměle připraveným oligonukleotidem obsahujícím jeden vyměněný (nepárující) nukleotid, takzvanou bodovou mutací. Tento oligonukleotid pak slouží jako primer pro dosyntetizování druhého vlákna vektoru. Mutace je tak zavedena do jednoho vlákna výsledného vektoru, které je třeba izolovat. Tento druhý postup je sice složitější, ale je levnější, neboť nemusí být syntetizován celý dvojřetězecový úsek molekuly DNA.

7.2.5 Bakteriální exprese

Před expresi se transformované buňky namnoží v potřebném objemu kultivačního média a vlastní expresi se zahájí přidaním látky, která je nezbytná pro transkripci genu v plasmidě3 – takzvanou induci. Nejčastěji používaným organismem je E. coli, která představuje dobře definovaný organismus s geneticky upravenými kmeny umožňujícími vysokou expresi cizích proteinů. Pro E. coli je také komerčně dostupný bohatý výběr vektorů.

Bakteriální exprese má ale i své nevýhody. Bakterie neumožňují sestřih a posttranslační modifikace eukaryotních proteinů. Proteiny mohou být využívány ne jako rozpustné makromolekuly ve správné konformaci, ale v nerozpustných shlucích zvaných inkluzní tělíska a disulfidové můstky nemusí být v

3Používané plasmidy jsou navrženy tak, že místo slouzící jako počátek transkripce, takzvaný promotor, je rozpoznávána příslušným enzymem pouze tehdy, je-li přítomna určité látka. Nejčastěji se používá promotor, který bakterie používají k zahájení transkripce genů pro produkcii léků. Tento promotor je běžně nefunkční, protože se v jeho blízkosti váže protein, který jej blokuje. Když ale přidáme látku, která je schopna vázat se na tento protein (používá se isopropylthiogalaktosid), protein se váže k přidané látkě místo k DNA a uvolní tak promotor. Volný promotor je rozpoznán příslušným enzymem a expresi může začít.
správně vytvořeny. Posledně zmíněné překážky se dají částečně obejít vhodnou volbou podmínek exprese (například snížení teploty).

7.2.6 Purifikace exprimovaných proteinů

Po sklizení buněk je třeba rozbít buněčnou stěnu a izolovat exprimovaný protein. Pro usnadnění poslední fáze se často připravuje protein ne ve své přirozené formě, ale spojený s jiným peptidem. Druhý peptid může být buď krátký (několik aminokyselin), nebo může jít o celý protein.


7.3 Chemická syntéza

7.3.1 Syntéza peptidů

Syntéza biomakromolekul s přesně daným pořadím monomerních jednotek je poměrně obtížný chemický úkol. Nejčastěji se používá syntéza na pevném nosiči, na který je navázána poslední jednotka rostoucí řetězce (obrázek 7.1). Vezmeme si náš oblíbený příklad proteinu a zamysleme se, jak bychom jej mohli připravit chemicky.

Každá aminokyselina obsahuje několik funkcíních skupin, které spolu mohou reagovat. Chceme-li z aminokyselin syntetizovat peptid, musíme zajišťovat, aby

- spolu reagovaly skupiny, které chceme spojit a
- nereagovaly skupiny, které chceme zachovat nedotčeny.


Funkční skupiny, které spolu reagovat nemají, potřebujeme naopak chránit, neboť *blokovat*. Po skončení syntézy musíme samorozčlenit peptid *odblokovat* (šetřením odstranit chránící skupinu), abychom získali molecule identickou s přirozeným peptidem. Chránící skupiny jsou na obrázku 7.1 představeny kroužkem, plný a prázdný kroužek mají ukázat, že je často dobré mít různé blokující skupiny, které lze odstranit nezávisle.

Obecně se postup chemické syntézy můžeme rozdělit do několika kroků, které budeme čidlovat podle obrázku 7.1. Nejprve musíme připravit jednotlivé stavební jednotky s vhodnými aktivovanými a blokovanými funkčními skupinami (1). Potom navážeme první (nebo poslední) monomerní jednotku na pevný nosič (2). Z navázaných jednotek selektivně odstěpíme chránící skupinu, která až dosud blokovala

---

4Nejčastěji se používají sekvence šesti histidinů, které komplekzuje dvojmocně katióny nikelnaté, kobalnaté, zinečnaté, měďnaté. Protein se izoluje metalochelatační afinitní chromatografií. Tato metoda využívá kolon s chelatovanými katióny. Při příchodu kolonu komplekzuje kation histidiny. Síla vazby je mnohonásobně vyšší pro umělou afinitní značky obsahující několik histidinů za sebou. Afinitní vázání proteinu je možno vyměnit snížením pH (protonace histidinů), imidazolem (kompetice s histidinem), kyselinou ethylendiamintetraoctovou (komplekzuje katióny lépe než kolona), nebo odstěpením afinitní značky od proteinu (pokud je afinitní značka oddělena sekvencí rozpoznávanou specifickou endopeptidasou).
Obrázek 7.1: Obecný postup chemické syntézy biopolymeru. Jednotlivé kroky jsou popsány v textu.
konec rostoucí řetězce (3). Přidáme aktivovanou a blokovanou druhou (nebo předposlední) monomerní jednotku (4) a prodloužíme rostoucí řetězec (5). Celý postup opakujeme tolikrát, kolik jednotek má syntetizovaný polymer obsahovat (6). Nakonec hotový biopolymer uvolníme z nosiče a odstraníme všechny chráničí skupiny (7).

K chemické syntéze peptidů se obvykle používá takzvaná *karbodiimidová metoda*, která je popsaná v Dodatku C.1. Tento metodou je možno syntetizovat i delší peptidové řetězce. Tento postup je ale nákladný, proto se syntetizují většinou krátké peptidy, zatímco větší proteiny se připravují metodami molekulární biologie (sekce 7.2).

### 7.3.2 Syntéza oligonukleotidů

Postup chemické syntézy nukleových kyselin vychází ze stejných principů jako peptidová syntéza. Klíčem k úspěchu je zaručit, aby se vytvořil fosfodiester mezi 5′ a 3′ hydroxyly a poslední OH-skupina kyseliny fosforečné zůstala volná. Nejčastěji se používá takzvaná *fosforamiditová metoda*, která je popsána v Dodatku C.2. V této metodě je málo reaktivní ester kyseliny fosforečné nahrazen reaktivním diisopro- pylumidem kyseliny fosforitové, jejíž OH-skupina je chráněna β-kynoethylovou skupinou.

Fosforamiditová syntéza se běžně používá k přípravě krátkých a středně dlouhých oligonukleotidů pro potřeby molekulového klonování. Pomocí metod molekulární genetiky můžeme oligonukleotidy spojovat do delších úseků DNA. Takto je možné uměle připravovat celé geny, nutná je pouze znalost sekvence.

### 7.3.3 Syntéza oligosacharidů


### 7.4 Enzymová syntéza

V sekci 7.2 jsme popsal, jakým způsobem lze získat protein (a nukleové kyseliny) biosyntézou v živé buňce. V sekci 7.3 jsme zmínil o metodách přípravy biomakromolekul chemickou syntézou. Enzymová syntéza představuje jakýsi přechod mezi biologickou a chemickou přípravou. Stejně jako v bakteriální expresi využívá složitých biologických katalyzátorů – enzymů, syntéza však neprobíhá v živé buňce, ale „ve zkumavce“ s použitím izolovaných enzymů.

#### 7.4.1 Enzymová syntéza nukleových kyselin

Nejčastějším případem enzymové syntézy je *polymerážová řetězová reakce (PCR)*, již zmíněná v sekci 7.2. Tato metoda je poměrně jednoduchá, protože k syntéze DNA stačí jeden enzym (DNA-polymerasa). Dalšími složkami reakční směsi jsou *templát* (molekula DNA nesoucí genetickou informaci, podle které syntéza probíhá), *primary* (krátké oligonukleotidy vymezující tu část templátu, podle které syntéza opravdu proběhne), *stavební jednotky* (čtyři nukleosidtrifosfáty, které slouží jako aktivované nukleotidy) a několik malých pomocných látek. V praxi se používá enzym z bakterií žijících v horkých pražnících, takže chod reakce lze řídit pohybem změnou teploty. Účinnost metody je mimořádná, z jediné molekuly

---

5Enzym se není ani zahráží téměř k bodu varu, takže jej není třeba během reakce přidávat.
templatu lze připravit dostatečné množství DNA, aby se s ní dalo pohodlně pracovat. Pomocí enzymové syntézy je možné připravit i molekuly RNA, je ale nutno použít jiný enzym (RNA-polymerasu).

### 7.4.2 Enzymová syntéza proteinů

V poslední době byla metoda podobná PCR zavedena i pro syntézu proteinů. Jde o takzvanou *bez-buněčnou expresi proteinů* (*cell-free protein expression*). Ve srovnání z PCR je bezbuněčná proteinová exprese mnohem náročnější, protože syntéza proteinů v buňkách je mnohem složitější, než replikace DNA. Exprese proteinů zahrnuje transkripci (syntézu mRNA podle templatu DNA) a translaci, která probíhá na složitých nadmolekulárních útvarech – ribozomech. Navíc nejsou aminokyseliny aktivované jednoduchým navázaním zbytků kyseliny fosforečné jako nukleotidy, nýbrž vazbou na relativně složité molekuly tRNA, které zaručují správné přeložení genetické informace do sekvence aminokyselin. Proto musí reakční směs pro bezbuněčnou expresi proteinu obsahovat množství enzymů včetně celých ribozomů, izolovaných z bakteriálních buňek. Templatem je opět DNA (ve formě vhodného vektoru), takže enzymy musí zajistit i transkripci do mRNA. Jako stavební bloky slouží obvykle volné aminokyseliny, které jsou aktivovány až v reakční směsi (ta musí obsahovat potřebné enzymy a tRNA izolované z buňek). Je ale možné použít aminokyseliny již vázané na tRNA. Výhodou této možnosti je, že nejsme omezeni na 20 geneticky kódovaných aminokyselin (synteticky můžeme připravit a do molekuly proteinu zařadit i aminokyseliny, které se v přírodě nevyskytují).

V současné době nachází bezbuněčná syntéza uplatnění zejména ve dvou případech. První případ zahrnuje situace, kdy není možné protein v buňce připravit (protein je pro buňku toxický, nebo jej enzymy buňky štěpí). Druhým případem jsou situace, kdy chceme připravit protein, ve kterém jsou aminokyseliny jasně daným způsobem upraveny (začlenění nepřirozených nebo izotopově značených aminokyselin na určité místo proteinu).

### 7.4.3 Enzymová syntéza oligosacharidů

KAPITOLA 7. PŘÍPRAVA VZORKU BIOMAKROMOLEKUL
Kapitola 8

Metody určování sekvence

8.1 Sekvenace proteinů


1. Pomocí specifických enzymů můžeme připravit kratší peptidy, které je možno sekvenovat úplně. Máme-li alespoň dvě sady peptidů s vzájemně se překrývajícími sekvencemi, můžeme kombinaci částečných sekvencí určit celkové pořadí aminokyselin proteinu.

2. V posledních letech je běžnější druhý postup, kdy určíme pouze částečnou sekvenci proteinu (například přibližně 15 aminokyselin N-konce proteinu). Z této informace je možné odvodit nukleotidovou sekvenci kódující příslušný protein (přesněji řečeno sekvenci s několika nejednoznačnými nukleotidy v důsledku degenerace genetického kódu). Pomocí této sekvence identifikujeme gen studovaného proteinu a sekvenaci provádíme na úrovni nukleových kyselin (metody sekvenace nukleových kyselin jsou rychlejší a méně pracně než metody sekvenace proteinů).


8.1.1 Edmanovo odbourávání

Podstatou Edmanovy metody je adice fenylisothiokyanátu na koncovou aminoskupinu proteinu. Takto derivatizovanou aminokyselinu je možno selektivně hydrolyzovat a pomocí kapalinové chromatografie

---

1 Jako příklad můžeme uvést trypsin štěpící na karboxylovém konci lysinu a argininu, nebo bakteriální endopeptidasy jako peptidasa V8 (z bakterie *Staphylococcus aureus*) štěpící na karboxylovém konci kyseliny glutamové.
KAPITOLA 8. METODY UŘÍCOVÁNÍ SEKVENCE

určit, které aminokyselině odstěpený fenylothiohydantoin odpovídal. Celý postup se provádí v automatických sekvenátech. Úspěšnost Edmanova odbourávání vyžaduje, aby se v každém kroku odstěpila úplně právě jedna aminokyselina. Chyby způsobené nedokonalostí se hromaď a omezují délku peptidu, který je možné sekvenovat. Kdyby stěpení probíhalo z 90 %, což je na první pohled slušná účinnost, druhý krok by proběhl pouze z 81 % (0,9²) a při určování sedmé aminokyseliny už bychom správného produktu získali méně než polovinu (0,9⁷ = 0,478). V praxi se většinou stanovuje sekvence zhruba patnácti aminokyselin.

8.1.2 Sekvenace hmotnostní spektrometrie

Analýza derivatizovaných aminokyselin odstěpovaných v jednotlivých krocích Edmanova odbourávání vyžaduje spolehlivou chromatografickou metodu. Existuje však jiná vlastnost, kterou se aminokyseliny liší a která je mnohem jednoznačněji definována, než chromatografické chování. Tento vlastnost je hmotnost. Přesné měření tak malé hmotnosti, jako je hmotnost jednotlivých molekul, není samozřejmě jednoduché. Metody, které nám to umožňují, se souhrnně označují jako hmotnostní spektrometrie (mass spectrometry, zkratka MS) a byly vyvinuty už před mnoha lety.


Největším problémem hmotnostní spektrometrie proteinů (a jiných biomakromolekul) nebylo po dlouhá léta vlastní měření hmotnosti, ale vytvoření iontu. Biomakromolekuly jsou sice v roztoku přítomny jako ionty, ale je velmi těžké tyto ionty z roztoku uvolnit a poslat do hmotnostního detektoru. Dnes známe dvě velmi účinná řešení tohoto problému. Prvním řešením je takzvaný elektrospray. Jde o zajímavý úkaz známý již dlouhá léta. Přivádíme-li do oblasti silného elektrického pole roztok iontů, kapky roztoku se odparují (ztrácí molekuly vody), takže se zmenšuje jejich povrch a roste poměr náboje k povrchu, povrchové napětí nestačí udržet kapku pohromadě, takže se tříší na menší a menší kapky, které se ještě rychleji odparují a vše pokračuje tak dlouho, až se odparí všechny molekuly vody a prostorem letí „nahý“ iont makromolekuly.

Jiným způsobem, jak získat ionty biomakromolekul, je odparit roztok a vzniklý prášek ostřelovat záblesky laseru. V nejčastějším provedení jsou k biomakromolekulé přidávány organické kyseliny (takzvaná matrix), které absorbují energii laseru, chrání makromolekulu a dodávají ji vhodný náboj. Tato metoda se nazývá Matrix-Assisted Laser Desorption Ionization (MALDI).

Pomocí popsaných metod hmotnostní spektrometrie můžeme určit molekuly ve vodních roztocích nebo vykrystalizované s vhodnou matricí, je možno studovat fragmenty peptidů. Tyto fragmenty se buď připraví před analýzou nebo, častoji, vznikají během vlastního měření rozpadem molekulu excitovaných laserovou ionizací (takzvaný post-source decay) nebo srázkami s molekulami vhodného plynu (collision-induced dissociation).

Takto získaná hmotnostní spektra mají vzhled řeščíků a z rozdílu hmotnosti jednotlivých fragmentů lze odvodit sekvenci původního peptidu. Rozlišení jedné hmotnostní jednotky stačí k určení většího

¹Přesněji řečeno, hmotnostní spektrometrie měří poměr hmotnosti a náboje nabité molekuly.
²K tomu lze například použít enzymy, které nespecifickým odstěpují aminokyseliny z jednoho konce, jako karboxypeptidasa Y.

Hmotnostní spektrometry detekují pouze nabité částice (ionty). Dojde-li ke stěpení peptidového řetězce s nábojem +1, může kladný náboj zustat buď na N-koncové části peptidu (takový stěp se označuje písmenem b), nebo na C-koncové části (stěp y). Analýza spektre je navíc komplikována tím, že ke stěpení hlavního řetězce může docházet nejen v místě peptidové vazby, ale i mezi uhlíky (vznikají stěpy a x) nebo mezi dusíkem a α-uhlíkem (vznikají stěpy c a z). Určování sekvence proteinů je proto často založeno na počátečním vyhodnocení spektre, které může zahrnovat všechny uvedené komplikace. Příklad sekvence krátkého peptidu (Ala-Thr-Gly-Cys) je ukázán na obrázku 8.1 Skutečné fragmentační spektrum, získané metodou MALDI-TOF-MS, ukazuje obrázek 8.2.

Sekvenace proteinů pomocí hmotnostní spektrometrie je užitná i v případech, kdy je známá sekvence DNA (gen proteinu). Pokud totiž dochází k posttranslačním modifikacím proteinu (viz sekce 2.2), lze tyto modifikace z rozdílu hmotnosti odhalit.

8.2 Sekvenace nukleových kyselin


V principu lze pro sekvencii nukleových kyselin použít hmotnostní spektrometrii jako u proteinů. V současné době je ale toto použití spíše výjimkou (něž popsaná enzymová sekvence umožňuje určit sekvenci délky několika stavek nukleotidů, zatímco hmotnostní spektrometrie je limitována několika desítkami). Pokud potřebujeme zkomplikovat delší úseky nukleových kyselin, přípravíme kratší překrývající se délky pomocí specifických endonukleas, podobně jako u proteinů.

Obrázek 8.2: Fragmentační spektrum peptidu získané pomocí MALDI-TOF hmotnostní spektrometrie.
8.3. SEKVENACE OLGOSACHARIDŮ

8.2.1 Chemická metoda Maxama a Gilberta

Při sekvenaci nukleových kyselin chemickou (odbourávací) metodou je nejprve třeba označit jeden konec vlákna, například 5'-konec radioaktivním fosforem \(^{32}\)P. Poté potřebujeme čtyři chemické reakce, které přeruší vlákno v místech jednotlivých bází. Pro přerušení vlákna na pyrimidínech se používá štěpení pyrimidinového kruhu hydrazinem. Reaktivita bází klesá v pořadí uracil > cytosin > thymín, takže v případě sekvování DNA lze zvolit podmínky, za kterých je štěpen cytosin i thymín nebo pouze cytosin (reakce thymínu je potlačena v přítomnosti 2M NaCl). V přítomnosti piperidinu pak dochází k selektivní hydrolyzě fosfoesterové vazby nukleotidu s odbouraným pyrimidinovým kruhem. Při štěpení na purínech se nejprve báze methylují dimethylsulfátem. Guanin se methyluje na dusíku 7, zatímco adenin se methyluje na dusíku 3. Hydrolyzou v zásaditém prostředí se přednostně štěpí N-glykosidická vazba s N7-methylguaninem, zatímco v kyselém prostředí se odštěpuje guanin i adenin. Fosfoesterová vazba se pak hydrolyzuje alkaličně při 90 °C. Popsané reakce se provádějí za podmínek zvolených tak, aby se každá analyzovaná vlákno štěpilo v průměru jednou (pochopitelně na různém místě). Vznikají tak čtyři sady radioaktivně značených oligonukleotidů, které při elektroforezě vytvoří žebříčky (obrazek 8.3), jejichž porovnáním lze snadno odečíst sekvenci studovaného vlákna DNA.

Popsaná metoda je vzhledem ke své pracnosti používána méně, lze ji však přizpůsobit například vyskytu nestandardních bází, kde selhává níže popsaná dideoxymetoda.

8.2.2 Enzymová neboli dideoxymetoda (Sangerova)

V současnosti je nejběžnější metoda sekvenace DNA založená na enzymové syntéze oligonukleotidů, ve které sekvováním oligonukleotid slouží jako templát. Přidáním vhodného množství dideoxyanalogueu trifosfátů, ze kterých je oligonukleotid sekvován (ddATP, ddGTP, ddTTP, ddCTP) je zajistěno náhodné předčasně ukončení syntézy oligonukleotidů (tedy tvorba žebříčku v elektrofořogramu, jak je ukázáno na obrázku [8.4]). Sekvence, kterou při dideoxymetodě čteme, není sekvenci vlákna DNA použité jako templát, ale sekvenci komplementárního vlákna. Čtená sekvence tedy odpovídá již sekvenci mRNA, podle které se syntetizuje protein.


8.3 Sekvenace oligosacharidů

Sekvence oligosacharidů se většinou provádí pomocí hmotnostní spektrometrie. Často se kombinuje použití enzymů štěpících glykosidickou vazbu (glykosidas) specifických pro různé sacharidy a typy glykosidické vazby s fragmentaci během měření. Analýza spektre je obdobná jako u proteinů, vzhledem ke složitosti primárních struktur sacharidů je ale sekvenace oligosacharidů mnohem obtížnější.

---

\(^{5}\)Dideoxyanalogue postrádá hydroxyl v poloze 3'. Proto po jeho zabudování nemůže dojít k dalšímu prodlužování řetězce.
Obrázek 8.3: Princip chemické metody Maxama a Gilberta. Ve spodní části obrázku je schematicky znázorněna poloha radioaktivních proužků na gelu, ve kterém byly elektroforézou rozděleny fragmenty vzniklé v jednotlivých reakcích. Šipka vlevo udává v jakém směru oligonukleotidy putovaly během elektroforézy.
Obrázek 8.4: Princip dideoxymetody. Ve spodní části je schematicky znázorněn obarvený gel po elektroforéze produktů jednotlivých reakcí. Šipka vlevo udává v jakém směru oligonukleotidy putovaly během elektroforézy. Všimněte si, že čtená sekvence (uprostřed gelu), neodpovídá vláknu DNA, které jsme použili jako templát pro sekvencii, ale vláknu komplementárnímu.
Obrázek 8.5: Výstup automatického sekvenátoru využívajícího fluorescenční značky. Obrázek ukazuje záznam čtyř fluorescenčních detektorů, které snímají fluoreskující oligonukleotidy vycházející z elektroforetické kapiláry. Červeně je značen signál fluorofo ru navázaného na ddTTP, modré na ddCTP, zeleně na ddATP, černé na ddGTP (obrázek poskytl Laboratoř molekulární fyziology restlin Přírodovědecké fakulty MU v Brně).
Kapitola 9

Úvod do experimentálních metod strukturní biochemie


Metody výzkumu biologicky zajímavých molekul. Optické metody, se kterými se seznámit se dozvedeme v kapitole 10, se podobají čichu a chuti, popisují složení molekuly, ne jejich tvar. Tak jako ochutnáváme stylovou cokolázu, můžeme odhadnout obsah cukru či alkoholu, orientačně dichroismus nám řekne, kolik procent skladaného listu či α-sřevbovice obsahuje, ale neposkytuje nám trojrozměrný model molekuly. Pokud nás zajímá prostorové uspořádání, musíme použit jiné přístupy.


Tak, jak jsme si popsali metody založené na difrakci záření jako obdobu zraku, můžeme najít i metody představující analogii hmatu. Jde o skupinu metod označovanou jako atomová silová mikroskopie (atomic force microscopy, AFM). Při těchto metodách se používá sonda s velmi ostrým hrotem, která „osahává“ zkoumaný vzorek podobně jako jehla gramofonu „ohmatává“ povrch klasické vinylové desky.

Kapitola 10

Optické metody

Cílem tohoto krátkého shrnutí principů optické spektroskopie je poskytnout základní přehled tomu, kdo zatím neměl příležitost do tajů světa molekul a elektromagnetických polí proniknout. Hned na začátku si musíme uvědomit, že vstupujeme do světa, kde se naplnější kvantové jevy, na které nejsme z běžného života známy. Odpovídající fyzikální popis (kvantová elektrodynamika) by si vyžádal mnohem delší povídky. Zde se pokusíme získat alespoň hrubou přehled o tom, jak je možné dostupat ke stavbě biomakromolekul z pozorování jejich interakcí s elektromagnetickým zářením.

10.1 Elektromagnetické záření

Pokusme se nejdříve trochu popsat základní pojmy. Tedy, co je to světlo? Jistě vědíte, že světlo je elektromagnetické záření. Elektromagnetické pole můžeme popsat čtyřmi veličinami: elektrickou intenzitou $\vec{E}$ a indukcí $\vec{D}$ a magnetickou intenzitou $\vec{H}$ a indukcí $\vec{B}$. Tyto veličiny musí splňovat takzvané Maxwellovy rovnice. Jedním z řešení Maxwellových rovnic je rovnice rovinné vlny

$$\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(\omega t - \vec{k} \cdot \vec{r})} = \vec{E}_0 (\cos(\omega t - \vec{k} \cdot \vec{r}) + i \sin(\omega t - \vec{k} \cdot \vec{r}))$$ (10.1)

Řešení je zde zapsáno v komplexním tvaru. Reálná část této rovnice nám říká, jaký bude vektor elektrické intenzity v čase $t$ a v místě určeném polohovým vektorem $\vec{r}$. Součin $\omega t$ popisuje kmitání v čase, skalární součin $\vec{k} \cdot \vec{r}$ popisuje vlnění v prostoru (vlnový vektor $\vec{k}$ míří směrem, kterým se vlna šíří), $\vec{E}_0$ je amplituda. Tak, jak je rovnice 10.1 napsána, popisuje průběh pole, pro které je vektor $\vec{E}$ v čase nula a v počátku souřadných soustav $\vec{E}(0,0) = \vec{E}_0$. Jinak bychom k argumentu (tedy k výrazu $\omega t - \vec{k} \cdot \vec{r}$) museli přidat ještě opravný člen, fázi. Jinými slovy, závislost by již nezačínala v maximu jako kosinusovka, ale v jiném bodě (při fázi rovné $90^\circ$ by začínala v nule, jako sinusovka). Často se také z argumentu vytkne $2\pi$ a vlnový vektor $\vec{k}$ se nahradí vektorem $\vec{s} = \vec{k}/2\pi$ a úhlová rychlost $\omega$ frekvencí $\nu = \omega/2\pi$. Velikost vektoru $\vec{s}$ je rovna převázané hodnotě vlnové dély $\lambda$ stejně jako frekvence $\nu$ je rovna převázané hodnotě periody kmitání $T$. Graficky je rovnice 10.1 znázorněna na obrázku 10.1 Protože nás zajímá průběh funkce $\vec{E}$ v prostoru i čase, je obrázek 10.1 rozdělen do dvou částí. V horní části je zachyceno kmitání v čase (v určitém bodě v prostoru, například v počátku souřadných soustav), dolní část ilustruje tvar vlny v prostoru (v určitém okamžiku, například v čase nula). Vektor $\vec{E}$ opisuje

---

1 Připomínáme si, že komplexní číslo $C$ se skládá ze dvou částí. Můžeme je buď zapsat pomocí amplitudy (maximální hodnoty) $C_0$ a fáze $\alpha$ jako $C = C_0 e^{i\alpha} = C_0 (\cos \alpha + i \sin \alpha)$ nebo je vyjádřit jako reálnou i imaginární složku $C = C_{Re} + i C_{Im}$, kde $C_{Re} = C_0 \cos \alpha$ a $C_{Im} = C_0 \sin \alpha$. Dále si připomínáme, že druhou mocninu komplexního čísla definujeme jako součin tohoto čísla s číselným nás komplexně sdruženým (komplexně sdružená čísla se liší jen ve znaménku imaginární části):$|C|^2 = (C_{Re} - i C_{Im})(C_{Re} + i C_{Im}) = C_0^2$. 

99

kosinovou křivku (v obrázku jsou šipkami znázorněny vektory $\vec{E}$ ve dvou časech a ve dvou bodech v prostoru).

Všimněte si, že v rovnici \[10.1\] i na obrázku \[10.1\] vystupuje lineárně polarizované světlo, tedy světlo, jehož vektor $\vec{E}$ kmitá pouze v jednom směru. Na obrázku \[10.1\] jsme si zvolili souřadnou soustavu tak, že vektor $\vec{E}$ kmitá ve směru osy $x$ a vlna se šíří ve směru osy $z$ (to znamená, že vlnový vektor $\vec{k}$, nebo samozřejmě $\vec{s}$, leží ve směru $z$). Kdybychom si chtěli znázornit vektor magnetické indukce $\vec{B}$, získali bychom stejný obrázek pouze by $\vec{B}$ kmital ve směru osy $y$ (Maxwellovy rovnice vyžadují, aby $\vec{E}$ a $\vec{B}$ byly navzájem kolmá). Světelné zdroje, které známe z běžného života, poskytují ovšem záření, jehož vektory $\vec{E}$ a $\vec{B}$ kmitají všemi směry (nepolarizované světlo). V popisu nepolarizované vlny má cenu sledovat pouze velikost (absolutní hodnotu) elektrické intenzity, $E$ (a magnetické indukce $B$). Rovnici \[10.1\] budeme tedy psát ve tvaru

$$
E(\vec{r}, t) = E_0 e^{i(\omega t - \vec{k} \cdot \vec{r})} = (\cos(\omega t - \vec{k} \cdot \vec{r}) + i \sin(\vec{k} \cdot \vec{r} - \omega t)).
$$

(10.2)

10.2 Světlo mezi molekulami

V kapitolce \[10.1\] jsme si popsal elektromagnetickou vlnu v prostoru bez molekul, ve vakuu. Co se s touto vlnou stane, když vstoupí do prostředí tvořeného nějakými molekulami? A co se stane s elektronem molekul ve vhodném elektromagnetickém moři? Elektrony jako elektricky nabité částice pocítí méněcí se pole a rozkmitají se s přibližně stejnou frekvencí. Tyto kmitání vlnového pole můžeme popsat jako elektrické dipolové momenty. Dipoly jsou ovšem sány také zdroji elektrického pole. Chceme-li tedy popsat celkové pole, musíme k oscilujícímu poli přidat oscilující polé vybuzených dipolových momentů. Když bude nakonec elektromagnetická vlna prostředí našich molekul opouštět, bude jím pozměněna. Pokud budeme rozumět tomu, jak méně molekuly prohazující vlnu, budeme umět z porovnání světla, které do vzorku s molekulami vstupuje, se světlem, které ze vzorku s molekulami vychází, získat informaci o našich molekulách.

---

2 V našem popisu jsme pominuli veličiny $\vec{D}$ a $\vec{H}$, které jsou ve vakuu přímo úměrné $\vec{E}$ a $\vec{B}$ ($\vec{D} = \varepsilon_0 \vec{E}$, $\vec{B} = \mu_0 \vec{H}$). V prostředí molekul musí navíc $\vec{D}$ popisovat elektrickou polarizaci a $\vec{B}$ magnetizaci látky. Vliv molekul na záření bude podrobněji probrán později.
10.2. SVĚTLO MEZI MOLEKULAMI

10.2.1 Světlo mezi malými netechními molekulami

Začněme nejjednodušším případem, kdy světlo prochází mezi molekulami zanedbatelných rozměrů, a kdy nedochází k přenosu energie na molekuly (nebo naopak). V takovém případě můžeme vzorek s molekulou považovat za homogenní prostředí. Světlo postupuje jako roviná vlna vzorkem a v každém okamžiku rozechvává novou rovinu elektrických dipolových momentů elektronů v molekulách. Ty knížají se stejnou frekvencí, ale s posunutou fází. Takto vybuzené vlny se šíří z dipolů všemi směry a ve všech směrech se vzájemně téměř vyrusí, kromě směrů rovnoběžných s původní (budící) vlnou. Výsledkem jsou tedy opět roviné vlny, které se skládají s původní (budící vlnou). Výsledná vlna se za běžných okolností trochu opoždí ve srovnání s vlnou budící.

3 Toto nejjednodušší působení molekul na procházející světlo popisujeme veličinou index lomu \( n \), která je rovna poměru rychlosti světla v prostředí molekul a ve vakuu.

10.2.2 Světlo si vyměňuje energii s molekulami


10.2.3 Vlnová rovnice mezi molekulami

Jakým způsobem musíme upravit rovnici (10.1), aby nám popisovala světelnou vlnu nejen ve vakuu, ale v prostředí molekul? Úprava se týká vektoru \( \vec{k} \). Zatímco ve vakuu je velikost \( k \) dán jednoduše vlnovou délkou \( k = 2\pi/\lambda \), v prostředí molekul musíme tuto hodnotu vynásobit ještě takzvaným komplexním indexem lomu. Jak název napovídá, jde o komplexní číslo tvaru \( n - i\kappa \), jehož reálná část je nás známý index lomu \( n \) a imaginární část \( \kappa \) je takzvaný index absorpce. Po dosazení do rovnice (10.1)

\[
\vec{E}(\vec{r}, t) = \vec{E}_0 e^{i(\omega t - n\vec{k}\vec{r})} = \vec{E}_0 e^{-i\kappa\vec{k}\vec{r}} e^{i(\omega t - n\vec{k}\vec{r})} = \vec{E}_0 e^{-\kappa\vec{k}\vec{r}} (\cos(\omega t - n\vec{k}\vec{r}) + i\sin(\omega t - n\vec{k}\vec{r}))
\]

vidíme, že jsme dosáhli toho, co potřebujeme. Zavedení indexu lomu \( n < 1 \) nám popisuje zpomalěný postup vlny a zavedení indexu absorpce \( \kappa \) nám popisuje tlumení v důsledku absorpce molekulami (imaginární číslo \( \kappa \) dalo po vynásobení imaginární jednotkou reálné číslo \( -\kappa \)).

3 Opoždování vln znamená, že elektrony jsou zdrojem vln s fází posunutou o \( 90^\circ \). V některých případech může být fáze posunuta až \( 270^\circ \) a předhubát fázi budící vlnu. Je proto nutné rozlišovat fázovou rychlost (ktará může být vysší než rychlost světla ve vakuu \( c \), protože fáze nenese žádnou informaci) od rychlosti šíření signálu (ktará nemůže překročit rychlost světla ve vakuu \( c \), protože důsledek nemůže být rychlejší než příčina).

4 Označení pochází z popisu situace, kdy světlo přechází z jednoho prostředí do druhého. Pokud dopadá roviná vlna na rozhraní mezi prostředními šikmo, musí se šířit v jiném směru (lámě se), aby zůstala spojitá. Úhel, pod kterým se vlna láme, můžeme spočítat právě z indexů lomu jednotlivých prostředí (Snellův zákon lomu).
10.3 Intenzita světla a spektrum

Intenzita světla \( I \) je dána druhou mocninou velikosti elektrické intenzity \( I(\vec{r}) = \vec{E}(\vec{r})\vec{E}^*(\vec{r}) = E_0^2\varepsilon^{-2\kappa l}\).

Ve spektrometrii často měříme absorbanci \( A = \log(I_0/I) \) při průchodu vzorkem tloušťky \( l \).

Z uvedených vztahů vyplývá, že běžně užívaný absorpci vrchní indexu absorbce: \( \varepsilon = 4\pi\kappa/(C\Lambda I) \), kde \( C \) je koncentrace molekul ve vzorku.

Na čem závisí schopnost molekul absorbovat energii světelného záření? Pravděpodobnost absorbce je v tím větším, čím je frekvence záření bližším energii přechodu. Splnění této rezonanční podmínky nám výkonu, tedy změny energie za jednotku času.

Oligo- a polysacharidy ve většině neabsorbuje v vhodné oblasti. Pokud sacharidy neobsahují například acetylovanou aminoskupinu (tedy systém podobný peptidové vazbě), neblížší použitelná vibrací delší vlnové délka 175 nm (přechod na π-orbitalu glykosidické vazby). Tak krátké vlnové délky vyžadují speciální zařízení (je například nutné odstranit kyslík rozpuštěný ve vodě) a přípravu se málo.

Další důležitou skupinou, jejíž absorbce můžeme měřit, je peptidová vazba. Nejvýraznější přechody mezi elektronovými stavby peptidové vazby jsou popsané v Dodatku [B]. Absorpcí při přechodu \( n \to \pi^* \) je intenzivnější než absorbce peptidová, její vlnová délka (kolem 230 nm) je na hranici použitelnosti běžných spektrofotometrů. Absorpcí při přechodu \( \pi^* \to \pi^* \) se nevylučuje přesto již šířeněji. Důvodem je kratká vlnová délka (190 nm).

Oligo- a polysacharidy většinou neabsorbuje ve vhodné oblasti. Pokud sacharidy neobsahují například acetylovanou aminoskupinu, tedy systém podobný peptidové vazbě, neblížší použitelná vlnová délka je 175 nm (přechod \( n \to \sigma^* \) elektronu glykosidické vazby). Tak krátké vlnové délky vyžadují speciální zařízení (je například nutné odstranit kyslík rozpuštěný ve vodě) a přípravu se málo.


10.4 Světlo a sekundární struktura

Jakým způsobem odráží schopnost absorbovat světlo sekundární strukturu (pravidelně uspořádaní monomerních jednotek biopolyméru)? Představme si monomerní jednotku, jejíž elektrony mohou přecházet z jednoho stavu do druhého. Elektrony v těchto dvou stavech budou popisovat dvě vlnové funkce a
10.4. SVĚTLO A SEKUNDÁRNÍ STRUKTURA

přechod mezi těmito stavby bude popsán jedním tranzitním elektrickým dipolovým momentem. Utvořme z takových monomerních jednotek dimer. Každá jednotka bude moci přecházet z jednoho stavu do druhého, celkem máme čtyři stavy a dva tranzitní dipóly. Kdyby byly tyto dvě monomerní jednotky od sebe izolovány, měly by dva stejné energie (stav s první jednotkou v nižším stavu a druhou ve výšším) a stav s druhou jednotkou v nižším stavu a první ve výšším). V dimeru jsou ale tranzitní dipóly blízko sebe a navzájem interagují (první leží v elektrickém poli druhého a naopak). Proto nezískáme výše zmíněné stavy se stejnou energií, ale stavy jiné. Stav, odpovídající energeticky výhodnějšímu uspořádání dipolů (s energií o něco sníženou) a stav odpovídající energeticky nevýhodnějšemu uspořádání dipolů (s energií o něco zvýšenou). Proto bude dimer absorbovat světlo při dvou vlnových délích.

Pokryné stavy budou samozřejmě odpovídat i nové vlnové délce (absorpční koeficienty). Absorpce dimeru a výsledné tranzitní dipolové momenty budou udávat pravděpodobnost absorpcí při rozdělení jednotlivých tranzitních dipolových momentů. Rozdíly energií nám budou udávat vlnové délky absorpcí dimeru a výsledné tranzitní dipolové momenty budou udávat pravděpodobnost absorpcí při jednotlivých vlnových délích (absorpční koeficienty).

Pokud budeme ve výstavbě makromolekul pokračovat, můžeme z monomerních jednotek vybudovat například pravděpodobnou šroubovicí. Získáme tedy sekundární strukturu. Z tranzitních dipolů (s energií o něco zvýšenou) a stav odpovídající energeticky nevýhodnějšímu uspořádání dipolů (s energií o něco sníženou). Proto bude dimer absorbovat světlo při dvou vlnových délích. Popsaným stavům budou samozřejmě odpovídat i nové vlnové délce a z těchto vlnových funkcí budeme moci spočítat výsledné tranzitní elektronové dipolové momenty dimeru (budou úměrné součtu a rozdílu jednotlivých tranzitních dipolových momentů). Rozdíly energií nám budou udávat vlnové délky absorpcí dimeru a výsledné tranzitní dipolové momenty budou udávat pravděpodobnost absorpcí při jednotlivých vlnových délích (absorpční koeficienty).

10.4.1 Polarizované světlo a sekundární struktura

Více se o sekundární struktuře můžeme dozvědět, když použijeme světlo, jehož vektor \( \vec{E} \) (a \( \vec{B} \)) kmitá v jednom směru. Takové lineárně polarizované světlo můžeme rozložit na dvě kruhové polarizované složky, jejichž směr kmitání se otáčí (u jedné doprava a u druhé doleva). Jak světelná vlna postupuje, amplituda \( \vec{E} \) opisuje šroubovicí, pro jednu složku pravotočivou a pro druhou levotočivou (obrázek 10.2 dole). Jednotlivé složky lineárně polarizovaného světla mají tedy odlišnou chiralitu (helicitu). Pokud takové světlo vstoupí do prostředí chiralních molekulek, bude každá složka ovlivněna jinak. Rychlosti postupu levotočivé a pravotočivé vlny (udávané jejich indexy lomu \( n \)) budou různé. Pokud molekuly nebudou absorbovat energii světla, projeví se rozdílnost indexů lomů \( n \) pro pravotočivou a levotočivou složku tak, že jejich opětovným složením získáme lineárně polarizované světlo kmitající v jiném směru (obrázek 10.2 uprostřed). Tento jev se nazývá optická otáčivost a jeho závislost na vlnové délce optická rotační disperze.

Častěji než změny indexu lomu měříme změny indexu absorpce \( \kappa \). Pokud chiralní molekuly světlo absorbuje, je pravděpodobnost absorpcí jednotlivých složek různá. Sečteme-li složky zpět, získáme místo knižnic po úsece elips (obrázek 10.2 dole). Poměr dlouhé a krátké poloviny elipsy (elipticita) je úměrný rozdílu absorpcních koeficientů (a tedy rozdílu \( \kappa \)) pro jednotlivé složky. Tento jev se nazývá cirkulární dichroismus (CD). Protože CD závisí přímo na chiralitě, můžeme se z CD spektro dozvědět užitečné in-

---

6 Obvykle sledujeme pokles absorbance.
7 Matematicky lze pravotočivou a levotočivou složku popsat rovnicemi

\[
E_R(r, t) = \frac{1}{\sqrt{2}} (E_{0,x} + iE_{0,y})e^{i(\omega t - (\kappa n_L - \kappa n_R)\vec{k}\cdot\vec{r})} \\
E_L(r, t) = \frac{1}{\sqrt{2}} (E_{0,x} - iE_{0,y})e^{i(\omega t - (\kappa n_L - \kappa n_R)\vec{k}\cdot\vec{r})},
\]

kde \( E_{0,x} \) a \( E_{0,y} \) jsou vektory délky \( E_0 \) ve směru \( x \) respektive \( y \).
Obrázek 10.2: Rozklad lineárně polarizovaného spektra na kruhově polarizované složky (nahoře) a ilustrace principu optické otáčivosti (uprostřed) a cirkulárního dichroismu (dole). Oproti obrázkům 10.1 je souřadná soustava otočena o 90°, takže osa x míří vzhůru a osa z dozadu kolmo k rovině nákresu. Poloha vektoru \( \vec{E} \) (a jeho jednotlivých složek) v určitém čase je znázorněna červenou šipkou. Černé je vyznačen průmět roviny, ve které kmitá vektor \( \vec{E} \) lineárně polarizovaného světla a spirál, po kterých se pohybují vektory \( \vec{E} \) kruhově polarizovaných složek.
formace o pravidelných chirálních strukturách biomakromolekul. Například u neuspořádaných proteinů pozorujeme v CD spektrech záporné maximum při 200 nm. Struktury skládaného listu vylučují pozitivní maximum při 198 nm a méně intenzivní negativní maximum při 215 nm. Konéčné α-šroubovice mají pozitivní maximum při 190 nm a dvě méně intenzivní negativní maxima při 208 nm a 222 nm (obrázek 10.3).

Pokud protein obsahuje různé druhy sekundární struktury, výsledné spektrum si můžeme představit jako součet tří základních spekter (náhodná konformace, skládaný list, α-šroubovice) v poměru odpovídajícím zastoupení těchto tří konformací v proteinu. Zpětně můžeme ze spektre odhadnout, jak velká část proteinu je v jednotlivých konformacích. Mohli bychom to udělat ručně, kdybychom odečítali elipticitu při třech různých vlnových dĺžkách a řešili pak soustavu tří rovnic o třech neznámých. V praxi většinou takovou analýzu svěříme počítači, který přesněji rozloží naměřené spektrum na jednotlivé složky.

Podobné rozdíly v CD spektrech můžeme pozorovat i u nukleových kyselin. Analýza spektre nuk-
KAPITOLA 10. OPTICKÉ METODY

Obrázek 10.4: CD spektra dvojšroubovic DNA tvořených opakovanou sekvencí CG: dvojšroubovice B (plná čára), dvojšroubovice A (čárkované) a levotočivá dvojšroubovice Z (tečkované).

levých kyselin je ale složitější, protože každý pár bází má jiné spektrum (spektra proteinů závisí na druhu aminokyseliny velmi málo). Jako příklad si můžeme ukázat spektra dvojšroubovic DNA tvořených opakující se sekvencí CG (obrázek 10.4). O této sekvenci je známo, že kromě běžné pravotočivé dvojšroubovice B a pravotočivé dvojšroubovice A může tvořit levotočivou dvojšroubovicí Z se zlomy v průběhu cukršťatové řady. Na obrázku je vidět, že pravotočivé dvojšroubovice mají podobná (ale rozlišitelná) spektra, zatímco spektrum levotočivé šroubovice Z má přibližně opačný průběh.

10.5 A co jádra?

10.5.1 Vibrační spektroskopie

Zatím jsme při popisu průchodu světla prostředím molekul věnovali pozornost jen elektronům. Ani jádra atomů však neleží nehybné v určité poloze. Jádra se vůči sobě pohybují knižovným pohybem. Polohu

Zákony kvantové mechaniky vyžadují, aby jádra kmitala jen s určitými frekvencemi. Stejně jako elektrony, i jádra se v molekule vyskytují v určitých stavových polohách. Kmitání se různě vztahují k určitém funkcím. Intenzitou přechodů do vyšších stavů pro jednotlivé vibrace je dána opět frekvence povrchové spektrum (na obrázku červené). Ty nám vlastně udávají směr polarizovaného světla rezonujícího s jednotlivými vibracemi. Energetické rozdíly jsou pro vibrací jader nízké o jeden až dva řády ve srovnání s elektronem, takže k rezonanci dochází s infračerveným světlem (frekvence \(10^{13} \text{ a } 10^{14} \text{ Hz})

Kvantová mechanika také definuje podmínky, za kterých můžeme absorpce v infračervených spektrech pozorovat. Důležitá podmínka pro to, aby byl tranzitní dipolový moment nenulový a vibrací přechod povolený se týká stálého dipolového momentu našich jednotlivých skupin (viz Dodatek 5 rovnice 13.34). Stálý dipolový moment se musí během vibrací měnit (vibrace nesmí být symetrická). Přechod, při kterých se stálý dipolový moment nemění a jsou v absorpčních spektrech zaznamenána, mohou být ale povolene v takzvaných Ramanových spektrech, o kterých se zmíníme v sekci 10.6.1.


---

8 Uvědomme si ale, že se tak dopouštěme určité nepřesnosti – kmitání našich tří jader bude spojeno s vibracemi zbyteku makromolekuly.

9 Ve vibrační spektroskopii je zvykem místo frekvence \(\nu\) udávat vlnočet \(\bar{\nu} = 1/\lambda\), zpravidla v jednotkách cm\(^{-1}\).
Ve vibračních spektrech můžeme pozorovat i cirkulární dichroismus, pokud použijeme polarizovaného infračerveného záření (citlivost takového měření je ale velmi nízká a tepře v posledních letech se objevily na trhu přístroje, které taková měření umožňují).

10.5.2 Vibrace jader v peptidové vazbě

Na závěr povídání o vibracích jader atomů se podíváme na konkrétní příklad využití vibračních spektre, na určování sekundární struktury atomů pomocí vibrací peptidové vazby. Opět použijeme zjednodušený popis, který vychází z předpokladu, že vibrace jednotlivých vazeb jsou do značné míry nezávislé na po- hybech okolních atomů. Ze zkušenosti víme, že vibrace kolem 3400 cm\(^{-1}\) je možné připsat vazbě N–H, vibrace kolem 1700 cm\(^{-1}\) je možné připsat vazbě C=O, vibrace kolem 1650 cm\(^{-1}\) je možné připsat vazbě C=N. Tento přístup je ale v mnoha případech příliš hrubý. Lepší výsledky poskytne zjednodušení opatrnější, které nepracuje s izolovanými dvojicemi atomů, ale rozumně zvolenými funkčními skupinami. V sekci 10.5 jsme si popisali šest možných vibrací skupiny složené ze tří atomů. Pokud se vrátíme k peptidové vazbě, situace spražených vibrací bude ještě složitější. Můžeme popsat devět spřažených pohybů. V praxi se nejčastěji využívá vibrace označovaná AMID I a AMID II (v infračervených spektrech; obrázek 10.6) nebo AMID I a AMID III (v Ramanových spektrech; obrázek 10.6), jejichž frekvence závisí na vodě, kterou vodíkové vazby ve vodě a na druhu sekundární struktury.

Postup při analýze sekundární struktury si můžeme ukázat na příkladu vibrace AMID I. Kdyby vodíkové vazby netvořil žádné vodíkové vazby, pozorovali bychom vibraci kolem 1670 cm\(^{-1}\). Vodě prochází do srovnání s molekulami vody, které zpomalují vibraci AMID I. V neuspořádaných molekulech můžeme pozorovat zpomalenou vibraci AMID I peptidových vazeb vystavených vodě v oblasti 1640–1650 cm\(^{-1}\). Co se stane při vytvoření sekundární struktury? Vodíkové vazby v α-sroubovici jsou slabší v srovnání s vodní vodíkové vazbě a tedy na druhu sekundární struktury.

Pokud na průchozí spektru vibrace AMID I můžeme připravit některé technické překážky. První z nich je přítomnost vibrací vody okolo 1650 cm\(^{-1}\), která překrývá nejzajímavější část spektra. Moderní přístroje dokáží vibrace vody od spektra spolehlivě odlišit, díve se zpravidla měření prováděla v težké vodě, která vibruje pomaleji. Další obtíži vyplývají z toho, že vibrace jednotlivých druhů sekundárních struktur se liší jen velmi málo a ve spektrech spojují také tento problém pomáhá rešit matematické zpracování spektra. Dnes je proto možné na uzavření zastoupení jednotlivých druhů sekundárních struktur. Při analyze spektrů je možné připravit se s rozměrům zastoupení sklopněho listu, tuto techniku může být rozlišení α-sroubovice od neuspořádané části proteinu (jak jsme si řekli, vibrace neuspořádané struktury značně závisí na tom, jestli tvoří vodíkové vazby s vodou).

10.6 Rozptyl světla

V sekci 10.2.1 jsme se zabývali světlem, které prochází prostředím molekul, aniž by si vyměňovalo s molekulami energii (jeho energie neodpovídá energii žádného přechodu mezi stavy molekuly). Předpokládali

10.6.1 Ramanova spektra

Zatím jsme mluvili o rozptylu světla, při kterém nedochází k výměně energie s molekulou (elastický rozptyl). Dalo by se očekávat, že pokud ozáříme molekulu světlem o vlnové délce, kterou molekulu neabsorbuje, bude mít rozptylé světlo stejnou vlnovou délkou. Neplatí to tak docela. Když vzorek molekul ozáříme velmi intenzivním světlem určité vlnové délky (obykle jde o laserový paprsek), pozorujeme v rozptyleném světle nejen světlo stejně vlnové délky, ale i světlo, které vyzářily molekuly, u nichž nejprve došlo k přechodu do nížšího (nebo vyššího) vibračního stavu. Frekvence tohoto světla je snížena (nebo zvýšena) o příspěvek, který odpovídá rozdílu energií vibračních stavů, mezi kterými došlo k přechodu. Závislost intenzity rozptyleného světla na popsané změně frekvence se nazývá Ramanovo spektrum.


Ramanova spektra nám umožňují studovat vibrace stejně jako infračervená vibrační spektra. Podmínky pozorovatelnosti určité vibrace jsou ale různé (v případě molekul se středem symetrie jsou dokonce

---

10.6. ROZPTYL SVĚTLA

Obrázek 10.6: Vybrané vibrace peptidové vazby.
vibrační přechody povolené v infračervených spektrech zakázané v Ramanových spektrech a naopak). Symetrické vibrace jsou dobře viditelné v Ramanových spektrech, zatímco nesymetrické vibrace jsou viditelné v infračervených spektrech. Pomocí Ramanových spekter tedy můžeme proto zkoumat sekundární strukturu stejně jako pomocí infračervených spekter, pouze se musíme zaměřit na jiné vibrace.
Kapitola 11

Rentgenová krystalografie

11.1 Ohyb záření (difrakce)

11.1.1 Virtuální náhrada mikroskopu

V kapitole 9 jsme narazili na nepříjemnou skutečnost. Nemůžeme sestrojit mikroskop s zvětšením potřebným pro pozorování jednotlivých molekul, protože neexistují čočky schopné zaostřovat „světlo“ dostatečně krátké vlnové dély (rentgenové paprsky o vlnové délce kolem 0.1 nm). Čočky mikroskopu modifikují dráhu světla jednoduchým způsobem, který můžeme nahradit výpočetní operací. Podívejme se na obrázek velmi jednoduchoho mikroskopu [11.1A].

Co se vlastně v takovém mikroskopu děje? Na sledovaný objekt dopadá světlo, které je objektem rozptylováno do všech směrů. Čočka mikroskopu lámá rozptýlené paprsky a zaostřuje je do obrazu, který může být zvětšený a převrácený, jako v našem případě.


Pro pochopení zesilujícího efektu krystalu a hlavně pro vyhodnocení pozorovaného difraktovaného záření je třeba porozumět tomu, co se s koherentním zářením při rozptylu na molekulářích děje. Jak jsme si již řekli v kapitole 10, elektromagnetická vlna rozkmitává elektrony molekul. V případě elastického rozptylu je každý kmitající elektron zdrojem nového záření (o stejné vlnové délce jako budící záření), které se šíří do všech směrů (obrázek [11.2]). Vlny vyzářené různými elektrony mají tedy stejnou vlnovou délku, ale v závislosti na poloze elektronu v krystalu mají různou délkou optické dráhy a jsou vůči sobě různě fázově posunuté (obrázek [11.2]). Právě na fázovém posunu záleží, jestli se paprsky budou skládat konstruktivně a záření se zesílí (to když mají jednotlivé vlny podobnou fází), nebo jestli se amplitudy odečtou a vlny vzájemně vyruší (to když mají fází opačnou).

Základní otázka tedy zní: Kdy mají rozptýlené vlny stejnou fází? Odpověď na tuto otázku vysvětlí, proč krystal funguje jako zesilovač rozptýleného záření a proč pozorovaný difrakční obrazec krystalu obsahuje jen omezený počet skvrn.
Obrázek 11.1: Vlevo princip nejjednoduššího mikroskopu, vpravo příklad difrakčních dat.

Obrázek 11.2: Rozptyl záření na jednom atomu (A) a difrakce na více atomech (B). Záření do zvýrazněného směru rozptýlené na modré označených atomech se skládá konstruktivně (vlna šířicí se daným směrem má stejnou fázi), zatímco záření rozptýlené na červeně označeném atomu je fázově posunuto.
11.1. OHYB ZÁŘENÍ (DIFRAKCE)

11.1.2 Braggovy roviny

Kdy dojde k rozptylu koherentních světelných paprsků tak, aby měly všechny stejnou fázi? Existují dvě možnosti:

1. Paprsky mají stejnou délku optické dráhy

Rozptylné záření se šíří ze zdroje všemi směry. Představme si, že objektom proložíme rovinu (může mít libovolnou orientaci a můžete si ji představit také jako zrcadlo, jak ukazuje obrázek 11.3A). Jeden z možných směrů šíření rozptylného záření je takový, že úhel, pod kterým záření vychází (úhel odrazu od původního rovinu), se rovná úhlu dopadu zdrojového paprsku na tuto rovinu. Všechny rozptylné paprsky, které se od původního světla „odraží“ pod stejným úhlem, pod jakým na rovinu dopadl zdrojový paprsek, mají stejnou fázi, bez ohledu na to, v kterém místě na rovinu (zrcadlo) dopadly. Obrázek 11.3A vysvětluje proč. Dva koherentní paprsky mají ve chvíli, kdy protnou úsečku HK stejnou fázi. První z nich narazí na zrcadlo pod úhlem \( \theta \) v místě označeném K. Z tohoto bodu se šíří rozptylné záření stejnou směry, nás ale teď zajímá pouze směr s úhlem odrazu \( \theta \). Čím je tento směr vyjádřen? Tím, že dráha, kterou musí urazit druhý paprsek před dopadem \( (\Delta') \) navíc proti prvnímu paprsku, je stejná jako dráha, kterou zase zrovna urazí první paprsek po odrazu \( (\Delta) \). A to bez ohledu na to, jak jsou body K a L daleko od sebe. Ve chvíli, kdy paprsky protnou úsečku LM, mají opět stejnou fázi. Povšimněte si také, že směr paprsku se po odrazu změní o dvojnásobek úhlu dopadu, tedy o 2\( \theta \). Pokud budeme chápat díra a difrakčních obrazců určit úhel odrazu od původního rovinu, musíme odchylovku difrakovaného paprsku od původního směru vždy vydělit dvěma. Kvůli analogii mezi odrazem světla od nějakého zrcadla (= reflektoru), která odráží paprsky do různých směrů, tímto původným zrcadlům říkají krystalografové Braggovy roviny. Zákon odrazu říká, že úhel dopadu světla od zrcadla se rovná úhlu dopadu. Tato vlastnost zrcadla zajistí, že odrážené paprsky si zachovávají stejnou relativní fázi, jako měly před dopadem na zrcadlo. Díky tomu můžeme v zrcadlech pozorovat obrazy předmetů.

2. Rozdíl délky optických drah paprsků je roven celočíselnému násobku jejich vlnové délky

Zatím jsem se zabývali tím, co se stane, když se dvě paprsky odrážejí od stejné Braggovy roviny. Co když se od každého odrazi od jiné roviny[1] (obrázek 11.3B)? Délky optických drah těchto paprsků budou v takovém případě různé. Odrážený paprsek proto nebudou mít stejnou fázi vždycky, ale jen tehdy, když bude rozdíl deltek optických drah paprsků roven celočíselnému násobku vlnové délky \( \lambda \). Jaká musí být vzdálenost Braggových rovin, aby byla tato podmínka splněna? Vysvětlení je na obrázku 11.3B. Rozdíl v délce optických drah paprsků řádí v vzdálenosti původních rovin a na úhlu dopadu zdrojového paprsku (úhel odrazu musí být stejný, jako úhel dopadu). Rozdíl v délce optických drah paprsků odrážených od různých rovin je roven dvojnásobku vzdálenosti \( \Delta \). Obrázek 11.3B ukazuje, že úhel HKL je stejný jako úhel \( \theta \). Součet úhlu v trojúhelníku je 180° a součet úhlů KLH a \( \theta \) je 90°. Pak můžeme spočítat, že \( \Delta = d \sin \theta \). Aby měly paprsky stejnou fázi po rozptylu na různých rovinách, je třeba, aby rozdíl v délce jejich optických drah (dvojnásobek vzdálenosti \( \Delta \)) byl roven celohodnotnému násobku vlnové délky zařízení \( 2\Delta = N\lambda \) (kde \( N \) značí přirozené číslo). Závislost mezi vlnovou délkou záření a vzdáleností mezi Braggovými rovinami je tedy \( N\lambda = 2d \sin \theta \).

Analyza obrázku 11.3B nám poskytla důležitý vztah:

\[
d = \frac{N\lambda}{2 \sin \theta},
\]

Objekty, které leží na Braggových rovinách rozptylují světlo tak, že jednotlivé paprsky budou mít shodnou fázi a budou se tedy skládat konstruktivně (jejich amplitudy se budou sčítat). Naproti tomu

KAPITOLA 11. RENTGENOVÁ KRystalografie

Objekty, které mají mezi sebou jinou vzdálenost než je vzdálenost Bragových rovin budou rozptylovat světlo s odlisnou fází.

Rovnice \[ \theta = \frac{n \lambda}{d} \] známá jako Braggův zákon, nám také říká, že větší úhly ohybu pozorované v difrakčních obrazcích odpovídají menším vzdálenost mezi Braggovými rovinami. Tato neprímo úměra je základem konceptu takzvaného \( \text{reciprokého prostoru} \), se kterým krystalografové často pracují.

Graficky lze vztah \[ \theta = \frac{n \lambda}{d} \] vyjádřit obrázkem \( C \). Zelené, modré a červené jsou vyznačeny Braggovy roviny různé vzdálené od roviny první (černé). Sípky odpovídajících barev ukazují směry, ve kterých má světlo rozptylováno na jednotlivých rovinách stejnou fází (celkový dráhový rozdíl \( 2\Delta \) se rovná vlhové déle \( \lambda \)).

Pro pochopení difrakce je dobré zamyslet se nad tím, jaký nejménisí a největší difrakční úhel můžeme pozorovat. Rozptyl světla pod nejménisím (nulovým) úhlem nastane, když se paprsek neodchyluje od původního směru \( \theta = 0^\circ \) a dráhový rozdíl mezi paprsky rozptylenými jednotlivými elektrony je nulový. To odpovídá nekončené vzdálenosti mezi Braggovými rovinami. Paprsek putující v původním směru proto nenese žádnou informaci o vzdálenostech, nese pouze informaci o celkovém počtu elektronů (,,hmotě'' objektu) . Naopak největší difrakční úhel odpovídá kolmému odrazi (\( \theta = 90^\circ \)), kdy se paprsek odráží zpět do směru, ze kterého přiletěl (směr paprsku se mění o 180\(^\circ\) ). V takovém případě se dráhový rozdíl rovná polovině vlhové délky. To vysvětluje, proč nemůžeme za normálních okolností pomocí elektromagnetických (nebo jiných) vln zkoumat objekty menší, než je polovina vlhové délky záření (takzvaný difrakční limit).

11.1.3 Difrakce na krystalu

Pro vysvětlení difrakce jsme použili popis zkoumaného objektu jako soustavy rovin (zrcadel), na kterých jsou rozmístěny elektrony. Zatím jsme se ale nezabývali tím, jaká je orientace Braggových rovin v krystalu. Orientace těchto rovin určuje směry, do kterých krystaly vysílají intenzivní rozptylené záření. Nyní se bliží podíváme na to, kde se ony roviny (zrcadlá) v krystalu nalézají. Ideální krystal je tvořen
velkým počtem uspořádaných molekul. Nejmenší jednotka krystalu, která se pravidelně opakuje ve všech směrech, se nazývá elementární buňka. Jak jsme si již říkali, krystal zasiluje rozptýlené záření v určitých směrech. Intenzivní rozptýlené záření pozorujeme ve směrech, do kterých rozptýluje různé elementární buňky záření se stejnou fází. Aby měly všechny paprsky v určitém směru stejnou fází, musí Braggovy roviny procházet stejnými body ve všech elementárních buňkách. Pak jsou splněny obě podmínky, kdy dochází k rozptylu rentgenových paprsků tak, aby v určitých směrech měly všechny rozptýlené paprsky stejnou fází. Tyto podmínky jsou splněny, když bude vzdálenost mezi rovinami (zrcadly) rovna délce hrany jedné elementární buňky v krystalu, nebo když na délku jedné hrany připadne celočíselný násobek rovin. Na obrázku 11.4A vidíme dvojrozměrný krystal s elementárními buňkami vyznačenými černě. Jednotlivé hrany buňek mají velikosti \( a \) a \( b \). Červené roviny protínají hranu v polovině. Skáčeš-li z jedné červené roviny na druhou, posuneme se proto o délku hrany \( a \) přesně po dvou skočích. Hrany \( b \) tyto roviny vůbec neprotínají. Červené roviny můžeme proto označit indexem (2 0). Modré roviny jsou vzdálené o délku hrany \( a \) a hrany \( b \) neprotínají. Označují se (1 0), protože do sousední buňky nám stačí jeden skok. Zelené roviny dělí hranu \( b \) na tři části, hranu ve směru \( a \) neprotínají. Označení zelených rovin je (0 3).

Trochu složitější příkazy Braggových rovin jsou ukázány na obrázku 11.4B. Když se podíváte na červené roviny, zjistíte, že skok na sousední rovinu nás posune o jednu elementární buňku jak ve směru \( \vec{a} \), tak ve směru \( \vec{b} \). Jde tedy o rovinu (1 1). Modré roviny jsou uspořádány podobně, ale je tu jeden rozdíl. Pokud se skokem na sousední rovinu posuneme o jednu délky hrany \( dál \) ve směru \( \vec{a} \), zjistíme, že ve směru \( \vec{b} \) se ocitneme na jednu buňku \( zpět \). Takovou rovinu označujeme (1 −1). U zelených rovin k posunutí o jednu buňku ve směru \( \vec{a} \) potřebujeme dva skočky, zatímco ve směru \( \vec{b} \) se o buňku dál dostaneme jedním skokem. Jde tedy o rovinu (2 1). Krytá používané pro krystalografickou analýzu mají tři rozměry. K popisu rovin, které je protínají v různých místech tak potřebujeme tři čísla. V krystalografii se tato tři čísla označují písmenky \( h, k, l \) a nazývají Millerovy indexy.

Jak víme z kapitoly 10 k rozptylu elektromagnetického záření dochází především na elektronech.

\[
F(h k l) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(x \, y \, z) e^{2\pi i (hx + ky + lz)} \, dx \, dy \, dz,
\]  

(11.2)

kde \( F(h k l) \) je komplexní číslo popisující reálnou a imaginární složku rozptyleného záření, \( h, k, l \) jsou Millerovy indexy a \( V \) je objem elementární buňky analyzovalého krystalu. Integraci provádíme přes celý objem elementární buňky ve všech třech rozměrech vyjádřených frakčními souřadnicemi \( x, y, z \). Tyto souřadnice jsou definovány jako zlomky hran elementární buňky, například \( x = 0,25 \) znamená, že elektron se nachází v jedné čtvrtině elementární buňky ve směru hrany a. Hodnota \( \rho(x \, y \, z) \) je tedy elektronová hustota v bodě \( x, y, z \) v rámci elementární buňky.

Při určování struktury molekul z rentgenové difrakce se ale nesnažíme spočítat hodnotu strukturálního faktoru. Naopak chceme spočítat distribuci elektronové hustoty, která nám ukáže tvar molekuly. Roz-
11.2 Krystalizace biomakromolekul

Pustili jsme se do rentgenové krystalografie biologicky zajímavých makromolekul, aniž bychom si řekli, jak krystaly tak velkých molekul můžeme získat. Přítom úspěch v určování struktury biomakromolekul závisí hlavně na tom, jestli se nám podaří připravit dostatečně veliký a kvalitní krystal, abychom jej mohli použít jako mřížku pro difrakci rentgenového záření. Pojďme proto tuto chybu rychle napravit.

Pro pochopení procesu krystalizace je nutné vědět něco o chování roztoků pevných látek. Nazveme-li nějakou chemikálii „pevnou látkou“, znamená to, že za běžných podmínek jsou její molekuly natěsnány na sebe bez větší volnosti pohybu. Pokud jsou molekuly poskládány pravidelně, mluvíme o krystalu. Co se stane, když takovou pevnou látku vložíme do vody (nebo jiného rozpouštědla)? Bude-li takové

\[ \rho(x y z) = \frac{1}{V} \sum_h \sum_k \sum_l F(h k l) e^{-2\pi i(hx + ky + lz)}. \]  

(11.3)

Jednotlivé členy rovnice jsou stejně, jako v rovnici 11.2. Integrace byla nahrazena sumací, protože k difrakci dochází pouze ve směrech definovaných parametry krystalu (podle orientace "arcadel" určené Millerovými indexy \((h k l)\)).

Strukturní faktor je komplexní číslo, které se skládá z reálné a imaginární části, stejně jako světelná ona má amplitudu a fázi. Rovnicí 11.3 můžeme tedy přepsat jako

\[ \rho(x y z) = \frac{1}{V} \sum_h \sum_k \sum_l |F(h k l)| e^{-2\pi i(hx + ky + lz)} + \alpha(h k l). \]  

(11.4)

Oproti rovnici 11.3 jsme provedli nahrazení \(F(h k l) = |F(h k l)| e^{i\alpha(h k l)}\), kde \(|F(h k l)|\) značí amplitudu a \(\alpha(h k l)\) fázi strukturního faktoru. Elektronová hustota je vždy reálným číslem.

11.2 Krystalizace biomakromolekul

Pustili jsme se do rentgenové krystalografie biologicky zajímavých makromolekul, aniž bychom si řekli, jak krystaly tak velkých molekul můžeme získat. Přítom úspěch v určování struktury biomakromolekul závisí hlavně na tom, jestli se nám podaří připravit dostatečně veliký a kvalitní krystal, abychom jej mohli použít jako mřížku pro difrakci rentgenového záření. Pojďme proto tuto chybu rychle napravit.

Pro pochopení procesu krystalizace je nutné vědět něco o chování roztoků pevných látek. Nazveme-li nějakou chemikálii „pevnou látkou“, znamená to, že za běžných podmínek jsou její molekuly natěsnány na sebe bez větší volnosti pohybu. Pokud jsou molekuly poskládány pravidelně, mluvíme o krystalu. Co se stane, když takovou pevnou látku vložíme do vody (nebo jiného rozpouštědla)?
KAPITOLA 11. RENTGENOVÁ KRYSTALOGRAFIE


Při rychlém odpařování dochází k tomu, že koncentrace roste prudce a brzy se začne tvořit velké množství zárodků krystalů. Ty se postupně rychle, ale brzy vyčerpají rozpouštěnou látku, takže výsledkem bude obrovské množství mikroskopických krystalů, které nejsou pro přesné určování struktury použitelné. Při pomalém odpařování naopak vzniká nejprve přesycený roztok. Pokud vložíme do přesyceného roztoku jeden malý krystal, budou se k němu molekuly rozpouštěné látky připojovat a můžeme tak vypěstovat velký pravidelný krystal. Proto je přesycený roztok, jako stav, ve kterém krystal může růst, ale samovolně nevzniká, pro pěstování velkých krystalů velmi důležitý.


Obličejem krystalizačním postupem je metoda visicí kapky, ilustrovaná vlevo na obrázku 11.6. Připravíme si asi mililitr roztoku látky, které krystalografové říkávají srážedlo. Může to být organická nebo anorganická sůl, organické rozpouštědlo, nebo polyethylenglykol. Z roztoku srážedla odebereme několik mikrolitrů, smícháme se stejným objemem roztoku zkoumané biomakromolekuly a kápneme na vhodně upravené sklikářské. Koncentrace biomakromolekuly v kapce by měla být taková, aby za dané koncentrace srážedla byl její roztok téměř nasyčený. Potom umístíme roztok srážedla do malé nádobky a sklikem nádobku přikryjeme tak, aby kapka visela nad hladinou srážedla. Nádobka se sklíčkem musí být dobře utěsněna, aby páry vody nemohly unikat ke kapce. Co se v nádobce bude dít? Koncentrace srážedla v kapce je po smíchání s roztokem biomakromolekuly poloviční ve srovnání s roztokem srážedla v nádobce. Tento rozdíl koncentrace se bude během času vyrovnávat. Pokud je srážedlo netěžké, bude k vyrovnávání docházet tak, že molekuly vody se budou vypražovat a koncentrováni roztok srážedla je bude postupně absorbovat. Protože objem kapky je mnohem menší, než objem roztoku srážedla (obykle zhruba tisíckrát), objem visicí kapky se zhruba v důsledku změněná původní poloviční na poloviční. Během zmenšování kapky poroste koncentrace biomakromolekuly, vznikne přesycený roztok a u určitém okamžiku může dojít k vytvoření zárodku krystalu (takzvaného nukleázního centra). Tento krystal se začne růst a během růstu se spotřebuje část biomakromolekuly, takže jejich koncentrace v roztoku poklesne na úroveň přesyceného roztoku, v níž jsou krystaly stabilní, ale nová nukleázní centra nevznikají. V ideálním případě pokles koncentrace dříve, než se zační vytvořit více zárodků, a v kapce nám poroste jediný krystal. Takové ideální (nebo alespoň téměř ideální) podmínky není snadné nalézt, obvykle je nutné vyzkoušet různé koncentrace různé kapky do různé látky

Metoda visicí kapky má různé alternativy. Pokud nám nestačí objem kapky, kterou kapilární síly

118
11.3. MĚŘENÍ DIFRAKCE

Obrázek 11.6: Krystalizační metody visící kapky (vlevo), sedící kapky (uprostřed) a mikrodávek (vpravo).

udrží viset na víčku, můžeme použít *metodu sedící kapky* (obrázek 11.6 uprostřed). Moderní laboratoře si svěřují zdlouhavé hledání vhodných podmínek automatickým robotům, pro které je výhodná *metoda mikrodávek* (*microbatch method*). Při ní je do parafinového oleje umístěno velké množství kapíček s roztokem obsahujícím srážedlo a biomakromolekulu (obrázek 11.6 vpravo). V oleji nedochází k odparování vody, takže se musíme střetnout s podmínkami, kdy je již na začátku koncentrace biomakromolekulu taková, aby vzniklo jen malé množství zárodků krystalů. Tyto větší nároky na počáteční složení jsou vyváženy tím, že automat může vyzkoušet mnohem více počátečních podmínek.

Jak vidíme, pěstování krystalů biomakromolekul je spíše umění založené na pokusu a omylu, než vědecká disciplína. Jsme-li úspěšní, na konci našeho snažení nás čeká krásný pravidelný krystal o rozměrech desetin milimetrů. Zajímávající je, že vzhled je ale uspořádaný molekulami v takovém krystalu.

Možná vás překvapí, jak výrazně se krystaly biomakromolekul liší od toho, co nám při slově *krystal* většinou vytané na mysli a co zpravidla odráží naší znalosti o obsahu slánky či cukřenky. Složité makromolekuly mají dosti nepravidelný tvar a je pro ně proto obtížné nalézt pravidelné krystalové uspořádání. Krystal proto obsahuje různé dutinky mezi jednotlivými makromolekulami, které jsou vyplněny krystalizačním roztokem. Objem roztoku je často srovnatelný s objemem zkoumané molekuly (kolísá mezi 30% až 70% celkového objemu krystalu makromolekul). Dutinky mezi makromolekulami jsou dost velké na to, aby se v nich mohla voda poměrně volně pohybovat. Přítomnost dutinek má pro krystalografii příjemné i méně příjemné důsledky. Značná volnost molekuly vody a některých částí makromolekulu způsobuje, že se makromolekuly mohou v krystalu částečně pohybovat a přesnost určování struktury je proto horší, než v případě krystalů malých molekul, které bývají těsně uspořádané. Na druhou stranu voda v dutinkách zaručuje, že ani v krystalu není biomakromolekula příliš náhodná ve „suchém stavu“. Prostředí, v kterém se molekula nachází, se spíše podobá velmi koncentrovanému roztoku a struktury biomakromolekul v krystaloch se většinou velmi podobají strukturám v roztoku, o kterých bude řeč v kapitole 12.

11.3 Měření difrakce

Od přípravy krystalu nyní postoupíme k vlastnímu měření. Přístroj pro měření difrakce (*difraktometr*) má tři hlavní části: zdroj rentgenového záření, zařízení pro uchycení krystalu v různých polohách (*goniostat*) a detektor.

11.3.1 Zdroj záření

Rentgenové záření je elektromagnetická vlnně o velmi krátké vlnové délce, tedy vysoké energii. Obecně jsou zdroji elektromagnetických vln nabité částice, které mění směr nebo rychlost pohybu. V klasické Roentgenové lampě dochází k vyzarování, když elektryny urychlené v silném elektrickém poli prudce brzdí při dopadu na kladnou elektrodu. Pro měření difrakce biomakromolekul se ale dnes již většinou
používají jako zdroje záření *synchrotrony*. V tomto případě už lze těžko hovořit o přístroji, protože může jít o zařízení velikosti malého městečka. V synchrotronech jsou elektrony urychleny v lineárních a kruhových urychlovačích na rychlost blízkou rychlosti světla a přivedeny do takzvaného akumulačního prstence. V něm se pohybují po přibližně kruhové dráze, která se skládá z rovných úseků spojených ohybovými magnety. V každé zatáčce vyzařují elektrony úzký paprsek velmi intenzivního polarizovaného elektromagnetického záření. Protože pro většinu měření potřebujeme záření o jedné vlnové délce, je důležitou součástí zdroje *monochromátor*, zařízení, které nám umožní vybrat záření s přesně definovanou vlnovou délku.

### 11.3.2 Goniostat

Během měření je krystal uchycen v takzvaném *goniostatu*, který umožňuje krystalu během měření otáčet. V některých experimentech se používá goniostat, který umožňuje krystal natočit podle více os a tak jej orientovat libovolným způsobem. Většinou ale vystačíme s jednoosým goniostatem, který otáčí krystalom tam a zpět v rozsahu několika stupňů.

### 11.3.3 Detektor


### 11.3.4 Sběr dat

Abychom mohli popsat trojrozměrnou strukturu objektu musíme se na něj podívat ze všech možných směrů. Podobně, pro určení struktury makromolekuly z difrákčních dat musíme naměřit difrákční data ve všech možných orientacích krystalu. Při sběru difrákčních dat proto používáme goniostat, který krystalom postupně otáčí.

### 11.3.5 Vyhodnocení dat

Zpracování naměřených dat můžeme rozdělit do tří kroků. Nejprve musíme určit mřížkové parametry a orientaci krystalu (vektory \(a\), \(b\), \(c\)). Známe-li mřížkové parametry a orientaci krystalu, známe ,,adresy“ (indexy \(h\), \(k\), \(l\)) míst, kde můžeme reflexe pozorovat. Víme tedy, v kterém místě na detektoru máme měřit intenzitu \(I\).

V dalším kroku musíme určit celkové intenzity jednotlivých reflexů, odečíst pozadí a stanovit chybu měření. Většinu reflexů můžeme určit pomocí různých metod, odečít pozadí a stanovit měřenou hodnotu.

Ve třetím kroku redukujeme naměřené intenzity na amplitudy strukturních faktorů. Intenzity naměřené detektorem jsou úměrné druhé mocnině absolutní hodnoty strukturního faktoru. Konstanty úměrnosti ale závisí na čidě dalších faktorů: na objemu krystalu, na rychlosti, se kterou krystalem otáčíme, na absorpci záření krystalem, na polarizaci záření, na postupném poškozování krystalu dopadajícím rentgenovým zářením. Tyto vlivy se mohou lišit pro jednotlivé reflexy, proto musí být provedeny různé opravy naměřených intenzit, abychom získali porovnatelné hodnoty. Tato část zpracování výsledků difrákčního
11.4 ŘEŠENÍ FÁZOVÉHO PROBLÉMU

Ukáznali jsme si, že rovnici [11.3] můžeme přepsat jako

\[ \rho(xyz) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} |F(hkl)| e^{-2\pi i (hx+ky+lz)+i\alpha(hkl)} \]  

(rovnice [11.5]), kde |F(hkl)| značí amplitudu a α(hkl) fázi strukturního faktoru.

Kdybychom uměli měřit amplitudu i fázi strukturního faktoru, stačilo by prostě nechat počítadlo zpracovat naměřené data podle rovnice [11.5]. Získali bychom ihned rozložení pravděpodobnosti výskytu elektronů v elementární buňce (takzvanou mapu elektronové hustoty), což je vlastně tvar molekuly.

Ve skutečnosti ale během experimentu pouze naměříme, kolik fotonů v daném směru dopadne na detektor, ne jakou mají fázi. Počet fotonů je uměrný druhé mocnině amplitudy vlny, informaci o fázi ztrácíme. Teto komplikaci říkáme fázový problém a velká část krystalografie je věnována jeho řešení.

### 11.4.1 Pattersonova mapa


Pattersonova mapa je mapou meziatomových vektorů. Jinými slovy, pokud bychom v mapě elektronové hustoty nalezli vysokou hustotu elektronů v místě popsaném vektorem \( \vec{r}_1 \) a v místě popsaném vektorem \( \vec{r}_2 \), v Pattersonové mapě bychom viděli maxima v místech pospaných vektory \( \vec{r}_1 - \vec{r}_2 \) a \( \vec{r}_2 - \vec{r}_1 \). Výšky maxim by přitom odpovídaly součinu výšek maxim v místech \( \vec{r}_1 \) a \( \vec{r}_2 \) mapy elektronové hustoty. Na obrázku [11.8] je ukázka Pattersonovy mapy jednoduché molekuly. Pattersonova mapa je vlastně kombinace několika posunutých obrazů molekuly, s tím, že každý z obrazů je posunut tak, aby jeden atom ležel v počátku buňky. Protože mapa zobrazuje vektory typu \( \vec{r}_1 - \vec{r}_2 \) i \( \vec{r}_2 - \vec{r}_1 \), najdeme v ní i převrácené obrazy molekuly.

Pokud není atomů moc, můžeme z Pattersonovy mapy určit původní polohy atomů v elementární buňce. Pattersonova mapa nám tedy umožní získat tvar molekuly i bez znalosti fáze! S rostoucím počtem atomů takový postup ale brzy selhává. Z poloh \( N \) atomů v elementární buňce můžeme sestavit \( N^2 - N \) nenulových rozdílových vektorů. Pro desetiatomovou molekulu tak získáme mapu s devadesáti maximy, které je možné rozumně analyzovat. Ale pro malý protein skládající se z tisíc atomů bude Pattersonova mapa obsahovat 999 000 maxim, která splynou v nerozlišitelnou směs. V kombinaci s dalšími triky je ale Pattersonovů přístup základem metod řešení fázového problému v krystalografii velkých molekul.

### 11.4.2 Metoda molekulárního přemístění

Metoda molekulárního přemístění (molecular replacement) umožňuje vyřešit fázový problém v případě, že máme dobrý model molekuly, její strukturu určujeme. V praxi to znamená, že již dříve někdo vyřešil strukturu dostatečně podobné molekuly, což je dnes již běžný případ. Podobnost terciárních struktur dvou proteinů většinou odráží podobnost jejich primárních struktur, takže svornávání sekvencí získáme dobrý odhad toho, zda bude molekulární přemístění úspěšné. Pokud je známá struktura dostatečně kompletní a sekvence obsahuje alespoň 40% identických aminokyselin, je velká naděje, že metoda molekulárního přemístění nám umožní vypočítat fázové úhly strukturních faktorů.

Molekulové přemístění vyžaduje, abychom modelovou strukturu vložili ve správné orientaci a do správného místa v rámci elementární buňky našeho krystalu. K tomu, abychom molekulu natočili, musíme určit tři rotační úhly, a k tomu, abychom ji umístili v elementární buňce, musíme zadat tři polohové parametry, popisující posunutí v prostoru. Obecně je molekulové přemístění šestirozmerný
11.4. ŘEŠENÍ FÁZOVÉHO PROBLÉMU


probém. Většinou lze ale úkol řešit ve dvou krocích, nejdříve spočítat orientaci a pak správně natočenou molekulu posunout do správného místa.

K pochopení rotačních a translačních funkcí nám pomůže Pattersonova mapa. Ačkoli pro molekulu velikosti proteinu nezískáme rozlišené rozdílové vektory, způsob jejich akumulace je specifický pro strukturu a orientaci makromolekuly. Vektory v Pattersonově mapě můžeme rozdělit na dva druhy. Intramolekulární vektory (mezi atomy v téže molekule) závisí pouze na orientaci, ne na pozici molekuly v buňce. Intermolekulární vektory (mezi atomy různých molekul) závisí na natočení i na pozici, takže po určení správné orientace nám pomohou najít potřebné posunutí. Protože vzdálenosti v rámci molekuly jsou v průměru kratší, intramolekulární vektory jsou většinou kratší. Proto můžeme rotaci spočítat, když použijeme pouze část Pattersonovy mapy v okolí počátku.

11.4.3 Přímé metody

Pokud předpokládáme, že se molekuly skládají z atomů podobných tvarů s kladnou elektronovou hustotou, můžeme nalézt statistickou závislost mezi sadami strukturních faktorů. Tyto vztahy nám mohou pomoci najít hodnoty fází. Zmíněné statistické závislosti jsou dostatečně výrazné pouze pro krystaly s velmi vysokým rozlišením a nepříliš velkým počtem atomů v elementární buňce (v řádu stovek). Proto jsou přímé metody použitelné k určení struktur proteinů jen vyjimečně. V rámci experimentálních přístupů k řešení fázového problému jsou ale přímé metody užitečné.

11.4.4 Metody izomorfního nahrazení

Metody izomorfního nahrazení jsou založeny na předpokladu, že dokážeme krystal modifikovat tak, abychom ovlivnili strukturní faktor, ale ne strukturu a uspořádání makromolekul v krystalové mířece.
OBRÁZEK 11.9: Ovlivnění celkového strukturního faktoru přítomností těžkého atomu. Vlevo je znázorněna poloha tří lehkých atomů (nakreslených červeně, zeleně a modře) a jednoho těžkého atomu (nakresleného azurovou barvou) mezi Braggovými rovinami, vpravo Argandův diagram ukazující sčítání strukturních faktorů jednotlivých atomů (strukturní faktory a atomy jsou nakresleny stejnými barvami). Černá šipka představuje výsledný strukturní faktor pro molekulu bez těžkého atomu ($F_P$) a s těžkým atomem ($F_{PD}$).


metody vycházejí z předpokladu, že těžké atomy neovlivní konformaci biomakromolekuly ani tvar krys-
talové míryžky – odtud i název izomorfní nahrazení, tedy nahrazení biomakromolekuly jejím derivátem,
który má stejný tvar. Aby izomorfní nahrazení bylo úspěšné mířkové parametry krystalů makromole-
kuly s těžkým kovem a bez těžkého kovu se nemí lišit o více než 1 %.

Pokud se nám podaří připravit dva krystaly se stejnými mířkovými parametry, jeden obsahující
pouze náviní protein a druhý protein s navázánými těžkými atomy, můžeme naměřit difrakční data
obou. Rozdíly v pozorovaných intenzitách rozptýleného záření odpovídajících reflexů z těchto dvou ex-
perimientů budou přispívat těžkým atomům. Z těchto rozdílů můžeme vypočítat Pattersonovu
mapu, která bude zachycovat pouze rozložení těžkých atomů v elementární buňce. Pokud elementární
buňka obsahuje malý počet těžkých atomů, je možné takovou Pattersonovu mapu analyzovat a určit, kde
se těžké atomy v buňce nacházejí. Ze známého rozmístění těžkých atomů dokážeme Fourierovou trans-
formaci vypočítat jejich strukturní faktor včetně fáze. Jak nám to pomůže v hledání fáze strukturního
faktoru proteinu samotného?

Podstatnou řešení je skutečnost, že strukturní faktor krystalu s těžkým atomem můžeme vyjádřit
jako součet dvou komplexních čísel. Prvním z těchto komplexních čísel je strukturní faktor těžkých
atomů získaný například z Pattersonovy mapy (pro něj známe amplitudu i fázi), druhým je strukturní
faktor makromolekuly, který se rovná strukturnímu faktoru krystalu tvořeného biomakromolekulou
bez těžkého atomu (pro ten známe amplitudu, ale ne fázi). Sčítání komplexních čísel je výhodné zobra-
zat v dvojrozměrném grafu, kde na vodorovné osu vynášíme reálnou část a na svislou imaginární
část (obrázek [11.10A]). Nás případ je takto graficky znázorněn na obrázku [11.10B]. Ve skutečnosti ovšem
známe směr jen pro šipku $\vec{F}_D$, u ostatních šipek známe jen délkou. Řešení tohoto problému je ukázáno na
obrázku [11.10C]. Nejdříve zakreslíme vektor $\vec{F}_D$, pro který známe velikost i směr. Do obrázku zakreslíme
jak vektor $\vec{F}_P$, který popisuje strukturní faktor makromolekuly spolu s těžkým kovem. Pro tento vek-
tor známe pouze amplitudu, ale ne fázi. Informaci o amplitudě nakreslíme jako kružnicí středovanou ve
počátku souřadnic a poloměrem odpovídajícím amplitudě $|F_P|$. Nakonec zakreslíme do obrázku to,
co víme o vektoru $\vec{F}_P$, který odpovídá strukturnímu faktoru nativní makromolekuly. Opět známe pouze
amplitudu, ale ne fázi. Takže $|F_P|$ zakreslíme jako kružnicí. Součet $\vec{F}_P$ a $\vec{F}_D$ musí odpovídat $\vec{F}_D$. Střed
kružnice znázorňující amplitudu $|F_P|$ proto umístíme na konec vektoru $\vec{F}_D$. Kružnice znázorňující am-
plitudu $|F_P|$ a $\vec{F}_D$ se obvykle protnou ve dvou bodech. Jednoduché izomorfní nahrazení (SIR, Single
Isomorphous Replacement) proto neumožňuje jednoznačné určení fázi strukturních faktorů makromole-
kuly. Je třeba připravit alespoň dva krystaly s různými těžkými atomy. Graficky lze řešení znázornit
opět v Argandově diagramu, do kterého v tomto případě zakreslíme dva známé vektory (strukturní
faktory těžkých atomů ve dvou derivátech) a při kružnice s poloměrem odpovídajícím amplitudám struk-
turních faktorů nativní makromolekuly a jejich dvou derivátech. Jednoznačné řešení fázného problému je
pak dáno průsečníkem všech tří kružnic. Toto je podstata metody vícenásobného izomorfního nahrazení
(MIR, Multiple Isomorphous Replacement). Metoda MIR je základní metodou řešení fázného problému
krystalů makromolekul neznámé struktury.

11.4.5 Metody využívající anomální rozptyl

Až dosud jsme předpokládali, že při průchodu rentgenovým záření krystalem se paprsek chová k elek-
trónům zcela netežně, že nedochází k výměně energie mezi molekulou a zářením. V takovém případě
interagují všechny elektrony se zářením stejně. Pro difrakce je důležitý jen rozdíl fázi jednotlivých
rozptýlených paprsků určený relativní polohou elektronů mezi Braggovými rovinami, ne absolutní hod-
notou fáze. Proto můžeme fázi záření rozptýleného na elektronu v počátku sousední soustavy krystalu
definovat jako nulovou. Ostatní elektrony budou rozptýlovat s odlišnou fází. Elektrony, které leží na
Braggových rovinách, budou rozptýlovat se stejnou fází ve směru definovaném Braggovou rovinou, fáze
záření rozptýleného na elektronech mezi Braggovými rovinami je posunuta úměrně kolmé vzdálenosti
elektronu od Braggovy rovin (plná šipky na obrázku [11.11]). Stejná sada Braggových rovin popisuje ale

11.4. ŘEŠENÍ FÁZOVÉHO PROBLÉMU
KAPITOLA 11. RENTGENOVÁ KRystalografie

Imaginární osa

Reálná osa

A

B

C

D

Obrázek 11.10: Strukturní faktor v Argandově diagramu (A) a řešení fázového problému izomorfním nahrazením (B, C a D). Na obrázku B jsou nakresleny strukturní faktory nederivatizované biomakromolekuly ($\vec{F}_P$), biomakromolekuly derivatizované těžkým atomem ($\vec{F}_{PD}$) a těžkých atomů ($\vec{F}_D$). Fázi známe pouze pro $\vec{F}_D$, pro ostatní strukturní faktory známe jen amplitudy. Neznámé fáze můžeme určit způsobem ukázáným na obrázky C. Do Argandova diagramu zakreslíme známý vektor $\vec{F}_D$. Dále zakreslíme čárkovanou kružnicí o poloměru $|\vec{F}_{PD}|$, která reprezentuje všechny možné vektory $\vec{F}_{PD}$. Nakonec zakreslíme i plnou kružnicí o poloměru $|\vec{F}_P|$, znázorňující možné vektory $\vec{F}_P$. Proto víme, že součet $\vec{F}_D$ a $\vec{F}_P$ je rovný $\vec{F}_{PD}$, víme také, že počátek vektoru $\vec{F}_{PD}$ (a tedy střed čárkované kružnice) leží v počátku vektoru $\vec{F}_D$, zatímco počátek vektoru $\vec{F}_P$ musíme umístit do konce vektoru $\vec{F}_D$ (aby na sebe šipky představující vektory navazovaly a znázorňovaly tak součet $\vec{F}_D + \vec{F}_P = \vec{F}_{PD}$, jak ukazuje obrázek B). Proto musíme střed plné kružnice položit do konce vektoru $\vec{F}_D$. Řešení ukazuje obrázek D. Místo, kam směřují vektory $\vec{F}_P$ i $\vec{F}_{PD}$, musí ležet na obou kružnicích, přísečník kružnic nám tedy určuje konec obou hledaných vektorů. Obecně se ale plná a čárkovaná kružnice protinou ve dvou bodech, úkohá má proto dvě možná řešení. Správné řešení je nakresleno černě, nesprávné červeně.
11.4. ŘEŠENÍ FÁZOVÉHO PROBLÉMU

Obrázek 11.11: Friedelův pár v nepřítomnosti těžkého atomu. Vlevo je znázorněna poloha lehkých atomů (nakreslených červeně, zeleně a modře) mezi Braggovými rovinami, vpravo Argandův diagram ukazující sčítání strukturálních faktorů jednotlivých atomů (strukturální faktory a atomy jsou nakresleny stejnými barvami). Částečné jsou nakresleny strukturální faktory pro odraz od spodní strany Braggových rovin. Černé šipky představují výsledné strukturální faktory \( \vec{F}_{(hkl)} \) a \( \vec{F}_{(h-k-l)} \). Tyto vektory mají stejnou reálnou složku a opačnou imaginární složku, jde tedy o komplexně sdružené čísla.

Reálná osa

Imaginární osa

\[ \vec{F}_{(hkl)} = \vec{F}_{(h-k-l)} \]

jaký odraz na opačné straně roviny, tedy pod přesně opačným úhlem (částečné šipky na obrázku 11.11). Jestliže odraz paprsku znázorněného plnou čarou na obrázku 11.11 odpovídá rovině s Millerovými indexy \( (hkl) \), druhou stranou stejné roviny (od které se odráží částečný paprsek) popisují Millerovy indexy \( (-h-k-l) \). Dvojice reflexů s indexy \( (hkl) \) a \( (-h-k-l) \) se označuje jako Friedelův pár. Pokud nedochází k interakci záření s elektronami, mají strukturální faktory Friedelova páru fázi stejně velikosti, ale opačného znamění. Matematicky řečeno, jde o komplexně sdružená čísla \( F(hkl) = F^*(-h-k-l) \). Jinými slovy, částečné šipky na obrázku 11.11 jsou zrcadlovým obrazem plných šipk.

Pokud ale dochází k interakci mezi zářením a elektronem, jinými slovy, pokud se frekvence záření blíží rezonanční frekvenci \( E \), situace se změní. Pohlcený foton vybuduje přechod elektrona do jiného stavu a vyzářený foton bude mít jinou amplitudu i fázi. Tento rozdíl se nazývá anomální rozptyl. Proto strukturální faktory Friedelova páru nejsou v přítomnosti anomálního rozptylu komplexně sdružené.

Rezonanční frekvence elektronů v atomech vodíku, uhličí, dusíku, kyslíku a fosforu jsou příliš daleko vlnových délek rentgenového záření používaného pro analýzu makromolekul. Proto i metody využívající anomální rozptyl vyžadují přítomnost těžších atomů. Pokud určíme rozmístění atomů, pro které pozorujeme anomální rozptyl, v elementární buňce, můžeme vypočítat amplitudu i fázi strukturálních faktorů \( F_{PD}(hkl) \) a \( F_{PD}(-h-k-l) \) a vytvořit tak fázový problém. Pokud jsme schopni rozlišit intenzity reflexů s opačnými hodnotami indexů \( h, k, l \), můžeme tedy anomální rozptyl využít pro řešení fázového problému podobně jako v případě isomorfního nahrazení. Více informací o tomto přístupu můžete najít v Dodatku D.

K anomálnímu rozptylu dochází, když se frekvence rentgenových paprsků blíží rezonanční frekvenci \( E \).

---

4Rezonanční frekvence je daná rozdílem energií dvou stavů elektronů v ozařovaných atomech, vyloučeným Planckovou konstantou, \( \Delta E/h \). V případě absorpcie rengenových paprsků mluvíme často o absorpční hraně, protože od určité frekvence je k dispozici velké množství přechodů.

11.5 Určení struktury z difrakčních dat


Zpětovníní struktury můžeme provádět nejen ručně, ale i s pomocí programů, které jsme si popsali v kapitole 6. Používá se při tom metody molekulové mechaniky (sekce 6.3), které jsou dostatečně rychlé i pro velké molekuly. Je zde ale důležitý rozdíl. Během výpočtu se kromě hledání konformace s nejmenším energií snažíme získat konformaci nejlépe odpovídající experimentálně naměřeným amplitudám strukturních faktorů.


Obrázek 11.12 ukazuje část modelu molekuly proteina zabudovanou do získané mapy elektronové hustoty.

Abychom vyhověli požadavkům na dostatečné počet měřených reflexů vzhledem k počtu proměnných, je výhodné snížit počet proměnných. Můžeme předpokládat, že struktura aromatických kruhů v aminokyselinách nebo bazích nukleových kyselin není ovlivněna celkovou konformací makromolekuly. Je proto zbytečné počítat se současní sestavu atomů aromatických skupin, jako by byly nezávislé. Počet stupňů volnosti je proto možné snížit na jediný torzní úhel, který popisuje orientaci aromatického kruhu jako celku. Takové dosazení pevně daných souřadnic některých atomu je příkladem použití terčů vazebných

128 KAPITOLA 11. RENTGENOVÁ KRYSALOGRAFIE
11.6 Kontrola správnosti

Během určování struktury jsme mnohokrát použili různých odhadů, oprav, ručních zásahů a přibližných výpočtů. Je tu tedy nebezpečí, že se nám do modelu vloudily nejrůznější chyby. Proto musíme pečlivě zkontrolovat, jestli získaná struktura nevypadá podezřele.

Nejpoužívanějším měřítkem presnosti struktury je $R$-faktor, definovaný jako

$$R = \frac{\sum_{i=1}^{N} |I_i - I_{calc}|}{\sum_{i=1}^{N} I_i}$$

$t$-tato odchylka může být různá v různých směrech, tato různost se ale u makromolekul obvykle zanedbává.
$R = \frac{\sum |F_i^{\exp}| - |F_i^{\text{mod}}|}{\sum |F_i^{\exp}|}$, \hspace{1cm} (11.6)

kde $|F_i^{\exp}|$ jsou amplitudy strukturních faktorů určených z experimentálních dat a $|F_i^{\text{mod}}|$ jsou amplitudy strukturních faktorů vypočítaných z atomárního modelu. Cílem je struktura přesnější, tím je $R$ nižší. Během zpřesňování strukturního modelu by se tedy měl $R$ snížovat. Nízkého $R$ lze ovšem dosáhnout i uměle, když například doplníme molekuly vody tam, kde by měla být část proteinu. Abychom odhalili takováto zdánlivá „vylepšení“, rozdělíme si intenzity na dvě části. Z větší části (90 %) vypočítáme strukturu, z menší části (10 %) počítáme $R$. Od okamžiku, kdy se takto vypočtený $R$ již nesnížuje (i když $R$ určený ze všech intenzit stále klesá), víme, že se nám nedaří model zpřesňovat.

Kromě testování přesnosti je dobré se podívat, zda hodnoty vazebních délek, vazebních úhlů a torzních úhlů odpovídají hodnotám běžným pro organické molekuly. Jakékoli neobvyklé hodnoty jsou podezřelé – je větší pravděpodobnost, že jsme udělali chyby, než že jsme objevili něco dosud nevýřešeného.
Kapitola 12

Nukleární magnetická rezonance

12.1 Magnetické chování jader

V úvodu kapitoly o využití magnetických vlastností jader atomů k určování struktury molekul je na místě otázka, odkud se magnetismus jader bere. Zdrojem magnetického pole je každý elektrický náboj, který se pohybuje. Pokud je zdrojem smyčka, kterou protéká elektrický proud, popisuje sílu takového magnetu magnetický moment μ. Magnetický moment je úmerný momentu hybnosti, který popisuje zakřivený pohyb náboje (elektronů) ve smyčce. Konstanta úmernosti je nazývá 

\[ \text{magnetogyrický poměr} \]

a obvykle se označuje písmenem \( \gamma \).


Vyměníme teď v našem příkladě Zemi za elektron a Slunce za jádro atomu. Elektron se pohybuje kolem našeho jádra podobným způsobem, jako kdyby putoval kruhovou smyčkou z drátu. V elektrotechnice se velikost magnetického momentu takové smyčky definuje jako součin proudu a plochy smyčky. Směr vektoru tohoto orbitálního magnetického momentu je stejný jako směr orbitálního momentu hybnosti. Pokud bychom předpokládali, že elektron obíhá po kruhové dráze, nedalo by nám moc práce spočítat, že magnetogyrický poměr je roven jedné polovině plochy náboje k hmotnosti. Podobně bychom mohli popsat vlastní magnetický moment, spojený s otáčením elektronu kolem své vlastní osy (se spinem).

Kdybychom zkoumali pohyb elektronu kolem jádra podle zákonů kvantové mechaniky a teorie relativity, došli bychom k trochu paradoxnímu závěru, že pohyb elektronů nemůžeme popsat jako pohyb Země (otáčení kolem vlastní osy), ale že elektrony přesto mají jak orbitální moment hybnosti, tak vlastní moment hybnosti (často nazývaný jednoduše spin). Proto mají elektrony také orbitální i vlastní magnetický moment. Hodnota magnetogyrického povahu vlastního momentu hybnosti se ale liší od výsledku, který jsme si uvedli v předchozím odstavci. Správnou hodnotu velmi dobře předpovídat kvantová mechanika v kombinaci s teorií relativity. Kvantová mechanika nám také říká, že ze všech stavů, ve kterých se může magnetický moment osamoceného elektronu v magnetickém poli nacházet, jsou jenom dva stacionární (pokud je elektron ve stacionárním stavu, setrvá v něm tak dlouho, dokud jej vnější příčina nedonutí stav změnit). V chemii se tyto stacionární stavy popisují hodnotou spinového kvantového čísla +\( \frac{1}{2} \) a −\( \frac{1}{2} \), nebo symbolicky šipkami ↑ a ↓. Kdyby se elektron utrhl z atomu a letěl volně prostorem, ztratil by orbitální moment hybnosti, ale vlastní moment hybnosti by mu zůstal. Vlastní

\[^1\] Matematicky definujeme moment hybnosti \( \vec{L} \) jako vektorový součin \( \vec{L} = \vec{r} \times \vec{p} = m(\vec{r} \times \vec{v}) \), kde \( \vec{v} \) je rychlost a \( \vec{r} \) vektor popisující polohu Země vzhledem k Slunci.

131
KAPITOLA 12. NUKLEÁRNÍ MAGNETICKÁ REZONANCE

Moment hybnosti mají i částice, ze kterých se skládá jádro. Vlastní moment hybnosti a magnetický moment jádra se skládají dosti složitým způsobem z příslušných momentů částic, které jádro tvorí a drží pohromadě. Proto se magnetogyrické poměry jader počítají velmi obtížně, dájí se ale velmi přesně změřit.

Pro zkoumání struktur biomakromolekul se téměř výhradně používají jader, která se podobají elektronu tím, že jejich magnetické momenty mají dva stacionární stavy. Z jader, která potkávávají makromolekulách, jde o izotopy \(^{31}\)P a \(^{1}\)H (což je proton), která tvoří většinu jádra fosforu a naprostou většinu jáder vodíku v přírodě. Navíc k tomuto druhu jader patří \(^{13}\)C a \(^{15}\)N, ty však tvoří pouze 1,1 % jader uhlíku a 0,4 % jader dusíku, takže často musíme biomakromolekuly těmito jadermi obohatit.

Pro naši metodu je zásadní, že v přítomnosti vnějšího magnetického pole koná vektor magnetického momentu \(\vec{p}_{\text{precess}}\) kolem vektoru magnetické indukce tohoto pole. Měřenou veličinou ve spektroskopii nukleární magnetické rezonance je frekvence \(\nu_0\) tohoto precesního pohybu. Frekvenci i směr precesního pohybu můžeme elegantně popsat vektorem \(\vec{\omega}_0\), jehož velikost je rovna \(2\pi\nu_0\) a jeho směr určuje osu precese. Směr precesního pohybu nám udává pravidlo pravé ruky. Úhlová rychlost závisí pouze na magnetogyrickém poměru \(\gamma\) a na magnetické indukci pole \(B\):

\[
\vec{\omega}_0 = -\gamma\vec{B}.
\]

Protože \(\gamma\) je průměrná konstanta, měřením úhlové rychlosti \(\vec{\omega}_0\) nebo frekvence \(\nu_0\) vlastně měříme sílu (indukci) magnetického pole v místě jádra.

12.2 Ovlivnění jader vnějšími poli

Kvantová mechanika je nezbytná pro správný popis magnetických momentů jednotlivých jader. Při měření spektre nukleární magnetické rezonance (NMR) ale pracujeme s makroskopickými vzorky (většinou s necelým mililitrem roztoku molekul, které zkoumáme), které obsahují miliardy miliard jader (velmi zjednodušená ilustrace je ukázána na obrázku 12.1). Proto celkové magnetické vlastnosti takových vzorků popisuje dobře klasická, nekvantová fyzika.


Velněk celkového magnetického momentu dělenu objemem vzorku se nazývá magnetizační moment, kterým magnetické momenty jader jsou velmi slabé. Velikost magnetizačního momentu je energie, kterou jader vznikne při svrchní energie, která je nejvýhodnější energeticky pro jader.

Po umístění vzorku do magnetického pole pomalu naroste počet magnetických momentů, které míří stejným směrem, jako \(\vec{B}_0\) (tento směr je energeticky výhodnější). Nadbytek magnetických momentů orientovaných podél \(\vec{B}_0\) se navenek projeví jako nenulový celkový magnetický moment (obrázek 12.2). Velikost celkového magnetického momentu dělenu objemem vzorku se označuje magnetickou indukcí a zpravidla označuje písmenem \(M\). V tepelné rovnováze je tento nadbytek malý i v nejsilnějších magnetech, protože magnetické momenty jader jsou velmi slabé a jen nepatrně přispívají k energii molekul.

\[\text{Jinými slovy, vektor magnetického momentu se otáčí pod určitým úhlem kolem vektoru magnetické indukce, tedy kolem vektoru, který popisuje směr a sílu magnetického pole. Obdobně koná precesní pohyb rotočený setračník v gravitačním poli, díky tomu se například udržíme v sedle jízdního kola.}\]
12.2. OVLIVNĚNÍ JADER VNĚJŠÍMI POLI


Obrázek 12.2: Vlevo je ukázána distribuce magnetických momentů jader ve vnějším magnetickém poli o indukci \( \mathbf{B}_0 \) (fialová šipka). Celkový magnetický moment je znázorněn jako azurová šipka. Uprostřed je ukázána distribuce magnetických momentů odpovídající magnetizaci (azurowá šipka) vychýlené mírně doprava dalším magnetickým polem o indukci \( \mathbf{B}_1 \) (tenká fialová šipka), které otáčí magnetické momenty po směru hodinových ručiček. Vpravo je ukázána tato distribuce otočená o polovinu kruhu v římském precese kolem \( \mathbf{B}_0 \). Slabší pole o indukci \( \mathbf{B}_1 \), otácející magnetické momenty po směru hodinových ručiček, v tento okamžik vrací vektor magnetizace (azurowou šipku) zpět do svislého směru.
KAPITOLA 12. NUKLEÁRNÍ MAGNETICKÁ REZONANCE

Přítomnost vnějšího magnetického pole síce učiní vzorek navenek magnetickým, neumožní nám ale měřit frekvenci precese magnetických momentů (alespoň ne běžným způsobem). Ačkoli nás vzorek obsahuje miliardy miliard magnetických momentů, které konají precesní pohyb s (téměř) stejnou frekvencí, detektor tuto frekvenci neregistruje, protože celkový magnetický moment je orientován podél \( \vec{B}_0 \) (otáčením kolem svislého směru se azurová šipka na obrázku [12.2] nijak nemění). Abychom mohli frekvenci měřit, musíme zajistit, aby azurová šipka (celkový magnetický moment) mířila jinam, než vektor \( \vec{B}_0 \), nejlépe kolmo k němu. Již víme, že vektor magnetického momentu se otáčí kolem vektoru magnetické indukce. To nám napovídá, že k otočení azurové šipky můžeme použít další magnetické pole, s indukcí \( \vec{B}_1 \) kolmou ke směru \( \vec{B}_0 \). Pokud bude vektor indukce \( \vec{B}_1 \) mířit směrem k nám, bude otáčet azurovou šipku na obrázku [12.2] po směru hodinových ručiček. Ale pozor! Vektor \( \vec{B}_1 \) nemůže jednoduše mířit jedním směrem kolmým na \( \vec{B}_0 \). Jakmile bude azurová šipka nepatrně vychýlena po směru hodinových ručiček ze směru \( \vec{B}_0 \), tedy na našem obrázku doprava, začne kolem \( \vec{B}_0 \) konat precese. Potřebujeme vektor \( \vec{B}_1 \), který osciluje mezi směrem k nám a od nás se stejnou frekvencí, s jakou se vektor magnetizace (azurová šipka) otáčí kolem \( \vec{B}_0 \). V tom případě bude \( \vec{B}_1 \) stále více vychylovat azurovou šipku ze svislého směru. V okamžiku, kdy je azurová šipka vychýlena doleva, míří \( \vec{B}_1 \) ke směru \( \vec{B}_0 \) a otáčí magnetizaci po směru hodinových ručiček, tedy ještě více vlevo. Vektor magnetizace se tak více a více vzájemně otočí a vznikne tak malé magnetické pole otáčející se s frekvencí precese magnetických momentů pozorovaných jader. Pokud bude v blízkosti cívka nebo smyčka z vodičového drátu, bude otáčející se pole v drátu indukovat střídavé elektromotorické napětí, oscilující s frekvencí precese. Toto napětí představuje signál, který v experimentu NMR měříme.

Obrázek 12.3: Vlevo znázornění pohybu vektoru magnetizace (jako azurová dráha na povrchu koule o poloměru \( |\vec{M}| \)) v přítomnosti oscilujícího magnetického pole radiové vlny. Vpravo distribuce magnetických momentů jader po sklopení vektoru magnetizace do směru kolmého na \( \vec{B}_0 \).

\[^{3}\text{V našich úvahách i ve skutečném experimentu bude velikost } \vec{B}_1 \text{ vždy mnohem menší, než velikost } \vec{B}_0.\]
12.3 Interakce jader s nejbližším okolím

Kromě uměle vyvolaných polí interagují jádra s magnetickými polí v jejich bezprostředním (molekulárním) okolí. Tyto interakce jsou pro nás velmi zajímavé, protože nesou informaci o nejbližších atomech v okolí jádra. Jenom díky těmto interakcím můžeme použít NMR ke zkoumání vlastností molekul. Ve všech biomakromolekulách se setkáváme se třemi typy interakcí, o kterých se teď zmíníme trochu podrobněji.

12.3.1 Interakce s magnetickými polí párovaných elektronů


Magnetické momenty jsou absolutně netečné k mechanickému otočení molekuly. Ať molekulu natočíme jakkoli, magnetické momenty konají precizní pohyb kolem směru vnějšího magnetického pole. Magnetické momenty jsou cítivé změny magnetických polí. Pokud v určitém okamžiku srážka s rozpuštědlem roztocí pozorovanou molekulu náhodou právě takovou rychlostí, jako je rychlost precizního pohybu, dojde k tomu, co jsme si popsal pro radiové vlny. Rozkmitání magnetického pole (v našem případě v důsledku nárazu rozpuštědla) ve směru kolmému na vnější pole s frekvencí rezonující s frekvencí procese určitého magnetického momentu začne otáčet magnetický moment po spirále ukázané na obrázku [12.3].z vlevo. Po spirále se můžeme pohybovat nahoru nebo dolů, v průměru budou častečky ty, které budou magnetický moment tlačit do energeticky výhodnějšího směru (tedy do směru rovnoběžného s vnějším polem). Na rozdíl od radiového vln, které mají stálé stejnou frekvenci, otáčení molekuly zasažené rozpuštědlem bude mít tu správnou rychlost a směr jen na okamžik, než přijde další náraz, který vzácnou sourou pokazí. Magnetický moment se proto posune po azurové spirále jen malíčko a pak dlouho čeká, až zase přijde ta správná srážka, aby se posunul po spirále o konek dál. Mezičasová dojde ke srážce, která bude magnetický moment zpátky. Jak jsme si ale rečli, jde o náhodný jev, který nakonec v souboru všech magnetických momentů povede k takovému rozložení magnetických momentů, které bude odpovídat termodynamické rovnováze při dané teplotě.
Tak jako kolísání magnetického pole kolmo na vnější pole vede k ustavení rovnovážného rozložení magnetických momentů a vzniku celkové magnetizace po vložení vzorku do magnetu, tak samozřejmě navrací magnetizaci zpět do rovnováhy poté, co jsme ji radiovými vlnami vychytili do kolmého směru. Díky kolmým výchylkám tedy signál získáváme (kdoby nevznikla v rovnováze celková magnetizace, nebylo by co kládět radiovými vlnami do měřitelného směru), ale i ztrácíme. Na ztrátu signálu má ale další vliv ale i kolísání rovnoběžné s vnějším konstantním magnetickým polem \( \vec{B}_0 \), tedy v něm uspořádání svídl sedmíčky. Kolem svídlého směru (směru \( \vec{B}_0 \)) totiž magnetické momenty konají precesní pohyb s frekvencí, kterou měříme. Pokud se všechy magnetické momenty pohybují se stejnou frekvencí, tedy \( \text{koherentně} \), celková magnetizace bezměřeného rotuje se stejnou, přesně danou, frekvencí kolem \( \vec{B}_0 \). V důsledku pohybů molekul se ale každé jádro nachází v malé jiném a stále se měním magnetickém poli. Protože frekvence precese je přímo úměrná magnetickému poli, tak se v každém okamžiku otáčí každý magnetický moment s nepatrné odloučenou frekvencí. Po delší době toto kolísání frekvencí vede k „rozostření“ celkové magnetizace a tedy úbytku měřeného signálu. Tento efekt je zvlášť silný u velkých, pomalu se pohybujících molekul. To je jeden z důvodů, proč je velmi obtížné pomoci NMR zkoumat proteiny skládající se z více než 200 aminokyselin.

Diskuse kolísání frekvencí přirozeně vede k otázce, jakou frekvenci v experimentu NMR vlastně měříme. Asi nepřekvapí, že jde o frekvenci průměrnou, tedy o průměr frekvencí magnetického momentu pozorovaného jádra ve všech různě natočených molekulách v roztoku. Rozdílné rozložení elektronové hustoty způsobuje, že různá jádra téhož izotopu (například proton v methylové skupině a proton v aromatickém systému) mají různou průměrnou precesní frekvenci. Ve spektrech pozorujeme takzvaný chemický posun, většinou označovaný \( \delta \). Protože precesní frekvence umíře velmi přesně, dokážeme dobře rozlišit chemický posun jader v různých skupinách. Chemický posun se obvykle vyjadřuje v milionech precesní frekvence (jednotka ppm) a pro malé molekuly je nejčastěji využíván informací o struktuře. Hodnoty chemického posunu protonů v různých funkčních skupinách jsou uvedeny na obrázku[12.5].

Rozložení elektronové hustoty samozřejmě obsahuje cennou informaci o struktuře (vizpomeňte na rentgenovou difrakci). V současné době ale nedokážeme vliv interakcí s elektronky na frekvence jader ve velkých molekulách dostatečně přesně předpovědět a hodnoty chemického posunu se využívají k určování struktur jen omezené.

### 12.3.2 Přímá interakce s magnetickými dipóly sousedních jader

Kromě elektronových oblasti se v sousedství pozorovaných magnetických momentů nacházejí také magnetické momenty okolních jader. Tyto magnetické momenty jsou zdroji malých lokálních polí, jejichž velikost a směr závisí na natočení celé molekuly (podobně jako velikost a směr polí způsobených orbitálními magnetickými momenty elektronů). Magnetický dipólový moment sousedního jádra může zvyšovat nebo snižovat magnetické pole v místě sledovaného jádra, v závislosti na tom, jaké polohy se sousední jádro nachází vzhledem ke sledovanému jádru. Pokud je spojnice jader kolmá na směr indukce vnějšího pole, jako vlovo na obrázku[12.6], magnetické pole sousedního jádra má v místě pozorovaného jádra opačný směr a vnější pole proto snižuje. Pokud ale otocíme molekulu o 60° (obrázek[12.6] vpravo), míří pole sousedního jádra v místě pozorovaného jádra (šikmo) vzhůru a vnější pole tak zesiluje. Pokud je magnetická indukce sousedního jádra přesně kolmá na indukci vnějšího pole (obrázek[12.6] uprostřed), nemá magnetický moment sousedního jádra na pole v místě pozorovaného jádra žádný vliv. Úhel mezi spojnicí jader a indukcí vnějšího pole, pro který toto nastane, se nazývá magický úhel.

---

5Fyzikálně nejprírozenější by bylo srovnání s volným protonem. Volné protony ale nejsou chemikálie, kterou bychom si mohli objednat z katalogů chemických firem. Proto se používají jiné, stabilní látky: v organické chemii tetramethylsilan (TMS) a v biochemii 3-(trimethylsilyl)-propan-1-sulfonová kyselina, která je dobře rozpustná ve vodě.

---
12.3. INTERAKCE JADER S NEJBLÍZŠÍM OKOLÍM


Dalším důsledkem interakce s magnetickým dipolovým momentem sousedního jádra je takzvaný nukleární Overhauserův jev (NOE, Nuclear Overhauser Effect); pokud se selectivně ozaříme radiovými vlnami jen sledované jádro a ponecháme sousední jádro v rovnováží,[6] interakce mezi magnetickými momenty jader způsobí nejen to, že momenty obou jader nakonec dospějí do rovnováhy, ale také to, že rovnovážné rozložení magnetických momentů sousedního jádra (netknutými radiovými vlnami) bude dočasně narušeno. Ozářením jednoho jádra tak způsobíme nerovnováhu magnetických momentů jader v blízkém okolí, což můžeme sledovat jako změnu intenzity měřeného signálu. Protože účinnost tohoto jevu klesá s šestou mocninou vzdálenosti mezi jádry, používá se NOE velmi často k měření vzdálenosti mezi jádry.

Srovnejme si teď vlivy magnetických momentů elektronů a vlastních magnetických momentů jader. Dipol-dipolové interakce mezi jádry a nerovnoměrné rozložení indukovaných polí elektronů, nazývané anizotropie chemického posunu, závisí na orientaci molekuly a v důsledku pohybů molekul vedou ke kolísání lokálních polí. Toto kolísání způsobuje relaxaci, tedy návrat rozložení magnetických momentů do termodynamické rovnováhy. Průměrná hodnota kolísajících polí v izotropním prostředí je v praxi se, zvláště pro velké molekuly, používá trochu jiné uspořádání. Průběžně se z rovnováhy vychýlí obě jádra, ale každé s jinou frekvencí, danou jeho frekvencí precese. Signál se pak zpracovává jako dvourozměrné spektrum.
KAPITOLA 12. NUKLEÁRNÍ MAGNETICKÁ REZONANCE

průpadě dipól-dipólůvé interakce nulová, v průpadě interakcí s orbitalními momenty elektronů nemulová, závislá na rozložení elektronové hustoty v okolí pozorovaného jádra a popsaná relativním číslem zvaným chemický posun. Jedním z projevů relaxace je nukleární Overhauserův jev, způsobený dipól-dipólůvou interakcí, který slouží k měření vzdáleností mezi jádry.

12.3.3 Interakce mezi magnetickými momenty jader zprostředkovaná elektrony vazeb

Na první pohled se zdá, že v diskusi interakcí s orbitalním magnetickým momentem elektronů a vlastními magnetickými momenty okolních jader jsme vyčerpali všechna možná pole v okolí pozorovaných magnetických momentů. Přesto jsme se nezmínili o jeho zajímavém a důležitém druhu interakcí, nazývaném J-interakce (J-coupling). Zatímco jsme dosud dávali vždy přednost klasickému popisu velkých souborů magnetických momentů, v případě J-interakce vyjde z mikroskopického popisu dvojice interagujících jader. Vlastně začneme ještě jednodušším případem, interakcí magnetického momentu jádra s vlastním magnetickým momentem elektronu (obrázek 12.7). Ponecháme přítom stranou, že ve skutečnosti není poloha elektronu přesně lokalizována, že náš popis je klasický, ne kvantový. Jak lze čekat, tato interakce se velmi podobá dipól-dipólůvé interakci s magnetickým momentem jiného jádra (obrázek 12.7), která se v izotropním prostředí zprůměruje k nule. S jednou důležitou výjimkou.

V jednom případě magnetické pole v místě pozorovaného jádra nezávisí na vzájemné orientaci jádra a elektronu (a nemá tedy smysl počítat průměr přes všechny orientace). Jde o případ, kdy se elektron vyskytuje přesně v místě jádra (obrázek 12.8). V našem klasickém obrázku to samozřejmě vypadá jako nesmysl, ale jádro a elektron nejsou kulečníkové koule. Kvantová mechanika mlví pouze o pravděpodobnosti výskytu elektronu v určitém místě a tuto pravděpodobnost popisují orbitaly. V nejednodušším případě atomu vodíku se elektron v základním stavu nachází v orbitalu s, a ten má v místě jádra nenulovou hodnotu (obrázek 12.8). Elektron se tedy očividně může s určitou (nenulovou) pravděpodobností nacházet v místě jádra. V tomto případě má magnetické pole elektronu v místě jádra stejný směr jako magnetický moment elektronu (obrázek 12.8). Energetický příspěvek této interakce je proto přímo úměrný $-\vec{\mu}_\text{e} \cdot \vec{\mu}_\text{ jádro} = -|\mu_\text{e}| |\mu_\text{ jádro}| \cos \alpha$, kde $\alpha$ je úhel mezi směry

---

\[
\delta = \nu - \nu(\text{Si(CH}_3)_4) \\
\nu = (1 + \delta) \nu(\text{Si(CH}_3)_4)
\]

Obrázek 12.5: Typické hodnoty chemického posunu protonu v různých funkčních skupinách.

\[ \nu = \frac{\nu - \nu(\text{Si(CH}_3)_4)}{\text{posun o } \delta} \]

---

\footnote{Přesněji řečeno, druhé mocniny amplitud vlnových funkcí, které jednotlivé orbitaly představují.}
12.3. INTERAKCE JADER S NEJBLÍŽŠÍM OKOLÍM

Obrázek 12.6: Interakce pozorovaného magnetického momentu s magnetickým dipolovým momentem sousedního jádra. Pozorované jádro je znázorněno azurovou barvou, sousední jádro zeleně. Vnější homogenní magnetické pole je znázorněno tlustou fialovou šipkou v místě azurového jádra, magnetické pole sousedního jádra je znázorněno zelenými magnetickými indukčními čarami (jejich hustota a tečna v daném místě udávají velikost a směr vektoru magnetické indukce, malé šipky vyznačují směr). Čtyři obrázky zachycují čtyři orientace molekuly pohybující se Brownovým pohybem v roztoku, θ je úhel mezi černou spojnici jader (nemusí mezi nimi být kovalentní vazba) a fialovým směrem indukce vnějšího magnetického pole. Azurové jádro se během rotace postupně dostává z oblasti, ve které pole zeleného jádra snižuje vnější magnetické pole (zelené šipky mříří dolů), do oblasti, kde pole zeleného jádra vnější pole zvyšuje (zelená šipka mříří nahoru).

magnetických momentů jádra (na obrázku 12.8A nevyznačený) a elektronu. Tato interakce je nazývána Fermiho kontaktní a tvoří většinou nejdloužitejší část J-interakce mezi dvěma jádry.


---

8 J-interakce je pozorovatelná mezi jádry spojenými dvěma, třemi, vyjímečně až pěti kovalentními vazbami.
Obrázek 12.7: Interakce pozorovaného magnetického momentu s magnetickým dipolovým momentem elektronu. Pozorované jádro je znázorněno azurovou barvou, elektron červeně. Vnější homogenní magnetické pole je znázorněno tlustou fialovou šipkou v místě azurového jádra, magnetické pole vlastního magnetického momentu elektronu je znázorněno červenými magnetickými indukčními čarami (jejich hustota a tečna v daném místě udávají velikost a směr vektoru magnetické indukce, malé šipky vyznačují směr). Čtyři obrázky zachycují čtyři vzájemné polohy elektronu vůči jádru, \( \theta \) je úhel mezi černou spojnicí jádra a elektronu a fialovým směrem indukce vnějšího magnetického pole. Azurové jádro se během rotace postupně dostává z oblasti, ve které pole elektronu snižuje vnější magnetické pole (červené šipky mířící dolů), do oblasti, kde pole elektronu vnější pole zvyšuje (červená šipka mířící nahoru).

úhlu se nazývá Karplusova křivka (obrázek 12.9).


Nyní bychom měli samozřejmě postoupit od popisu jedné dvojice jader ke statistickému popisu miliard miliard takových dvojic ve skutečném vzorku. Takový popis je ale velmi složitý. Proto se spokojme s tím, že si řekneme, že při měření dvou jader citlivých vzájemnou \( J \)-interakcí pozorujeme pro každé jádro dvě precesní frekvence lišící se o hodnotu interakční konstanty označované \( J \). Tyto frekvence se při pohybu molekule nemění, pokud nedochází ke změnám konformace v místě jader (vazebných delek, úhlů a zejména torzních úhlů). Rozdíl precesních frekvencí způsobený \( J \)-interakcí také nezávisí na vnějším magnetickém poli, hodnota rozdílu vyjádřená v Hz je stejná ve všech spektrometrech. Pokud je jádro ovlivněno \( J \)-interakcí s více okolními jádry, vlivy těchto interakcí se kombinují.

### 12.4 Spektrometr a experiment NMR

Jak vypadá spektrometr, tedy zařízení, které nám umožní magnetické momenty ovlivňovat tak, jak jsme si popsal, a měřit jejich frekvenci? Nezbytnou součástí spektrometu je silný magnet, vytvářející velmi homogenní magnetické pole o indukci \( B_0 \). Jako zdroj slabšího oscilujícího magnetického pole \( B_1 \) využíváme magnetickou složku elektromagnetických vln. Stejně jako v případě optické spektroskopie tedy vystavujeme jádra elektromagnetickému vlnění. Frekvence tohoto vlnění musí rezonovat s frekvencí precese magnetického momentu jádra, které pozorujeme. Proto hovoříme o nukleární magnetické rezonanci. Energetické rozdíly mezi orientacemi magnetických momentů jader jsou velmi malé, proto je i frekvence elektromagnetického pole v NMR mnohem nižší (v oblasti velmi krátkých radiových vln, řádově \( 10^8 \) Hz), než frekvence záření v optické spektroskopii. Vidíme, že další důležitou částí spektro-
12.4. SPEKTROMETR A EXPERIMENT NMR


metru NMR je radiový vysílač.

Nej jednodušší experiment NMR vypadá tak, že vzorek ozařujeme radiovými vlnami po dobu, během které se sklopi vektor magnetizace \( \vec{M} \) z původního směru rovnoběžného s \( \vec{B}_0 \) do roviny kolmé na \( \vec{B}_0 \). Po vypnutí elektromagnetického záření bude \( \vec{M} \) v této rovině rotovat kolem \( \vec{B}_0 \). Co to znamená z praktického pohledu měření? Zatímco na začátku, v rovnovážném stavu, směřoval vektor nehybné jedného směrem, po zarážení po určité dobu \( \vec{M} \) rotuje. Jádra sama se nám tak stala zdroji proměnného elektromagnetického pole, tedy vysílači elektromagnetického záření. Přesněji řečeno, jádra byla vždycky zdroji proměnných elektromagnetických polí, ale tato pole se navzájem rušila. Ozářením radiovou vlnou jsme ale soubor jader ovlivnili takovým způsobem, že se jejich proměnná pole sčítají a tvoří měřitelné elektromagnetické záření.

Toto záření má frekvenci shodnou s precessní frekvencí magnetických momentů jader a můžeme je zachytit detektorem v principu podobným rozhlasovému přijímači. Precesní frekvence stejných jader v různých místech molekuly se přitom mírně liší v důsledku interakce s elektrony a ostatními jádry, které jsme si podrobně popsali v části 12.3. Signál zachycený radiovým přijímačem tedy nemá jednu frekvenci, ale skládá se z množství vln s mírně odlišnými frekvencemi. Úplně stejně vypadá signál ve frekvenčně modulovaném rozhlasovému vysílání na velmi krátkých vlnách. V rozhlasovém studiu se zachytí akustické vlny vlasu končící na redaktora mikrofonem, převedenou na elektrické oscilace se stejnou frekvencí (v rozsahu hertzů až kilohertzů) a smíší s vysokofrekvenční radiovou vlnou (o frekvenci kolem 100 MHz). Vzniklá vysokofrekvenční radiová vlna je tak modulovaná nízkofrekvenčním audio-signálem. Tato vlna je zesílena, vysílácí anténou šířena do prostoru a posléze zachycena anténnou vašeho rozhlasového přijímače. V něm se z modulované radiové vlny odstraněna vysokofrekvenční složka a zesílena nízkofrekvenční

---

9V běžných spektrometrech je tato doba dlouhá približně 10 µs, závisí na konkrétním spektrometru, nastavení výkonu vysílače i složení vzorku. Při studiu proteínů a nucleových kyselin se snažíme získat co největšího signálu, tedy sklopit magnetizaci přesně do kolmého směru. Proto vždy před vlastním experimentem měříme, jak dlouhá doba potřebná ke sklopení magnetizace o 90°, ze svislého do vodorovného směru. Tomuto pomocném měření říkáme kalibrace 90° pulzu radiových vln.

10Přirovnání k radiu je vskutku na místě. Ve spektrometrech s magnetem o indukci 9 až 12 tesla vysílají jádra uhliku na stejné frekvenci jako vaše oblíbená FM stanice.

Na tuto vysokofrekvenční složku je přijímač naladen.
KAPITOLA 12. NUKLEÁRNÍ MAGNETICKÁ REZONANCE

Obrázek 12.9: Závislosti třívazebných interakčních konstant na hodnotě torzního úhlu \( \phi \) proteinu (Karplusovy křivky).

audio-složka rozhlasového signálu je poslána do reproduktoru, který ji převede na zvukové vlnění věrně kopírující akustické vlnění zachycené ve studiu.

Přijímač spektrometru funguje úplně stejně. Pro jednoduchost si můžeme představit, že jsme jádra ozářili radiovou vlnou se stejnou frekvencí, jako je precesní frekvence srovnávací látky, pomocí které definujeme nulu na škálě chemického posunu. Chemický posun pak hraje stejnou roli jako frekvence hlasu přepínávák v rozhlasovém vysílání. Po vypnutí vysílači anténa přijímač zachytí signál generovaný magnetickými momenty jader, který je modulován chemickým posunem. Ze zesíleného signálu jáder přijímač spektrometru odstraní vysokofrekvenční složku, složku o stejné frekvenci, se kterou jsme jádra ozařovali. Ze signálu zbytek směsí vln s frekvencemi odpovídající chemickým posunům jader v různých místech molekuly. Takový výsledek měření signálu protonu je ukázán na obrázku 12.10A pro velmi jednoduchou molekulu, pro acetaldehyd (CH\(_3\)—CH=O) se čtyřmi atomy vodíku ve dvou chemických skupinách, třemi v methylové skupině a jedním v aldehydové.

Obrázek 12.10A pěkně ukazuje, že odečtení frekvencí přímo ze záznamu měření signálu je i pro nejjednodušší molekuly prakticky nemožné. Proto záznam převádíme do digitální formy a dále zpracováváme. Nejčastěji je převod používám následně Fourierovou transformaci. Na rozdíl od rentgenové krystalografie pracujieme se závislosti měřeného signálu na čase. Výsledkem Fourierovy transformace je v případě NMR spektrum intenzity signálu pro různé frekvence. Tato intenzita je blízká nule pro frekvence, které v signálu nevyskytovály, ale vysoká pro hodnoty blízké frekvencím precesního pohybu magnetických momentů v měřeném vzorku. Spektrum acetaldéhydu je ukázáno na obrázku 12.10B. Při pohledu na spektrum okamžitě vidíme, že trojice protonů methylové skupiny koná presči pohyb s frekvencí asi o 1 000 Hz vyšší, než je frekvence radiových vln vysílače, zatímco frekvence aldehydového protonu je téměř o 5 000 vyšší než vysílač frekvence.

Obrázky 12.10C,D nám připomínají, že jsme při povídání o modulaci chemickým posunem zanedbali, že magnetické momenty protonů interagují také mezi sebou prostřednictvím elektronů tří kovalentních
12.5 Vícerozměrná spektroskopie a korelace frekvencí

V kapitole 12.2 jsme si popsali princip moderního spektrometru, který používá krátké ozařovací pulzy elektromagnetického záření. Při studiu malých molekul nám často stačí jednoduchý experiment s jedním ozařovacím pulzem, který sklopl magnetizaci pozorovaných jader do vodorovné roviny. Po nějakou dobu (přibližně sekundu) magnetizace každého jádra v této rovině rotuje s precesní frekvencí magnetických momentů, které ji tvoří. Během této doby můžeme měřit signál indukovaný v přijímači anténu, než náhodné interakce s okolím přivedou magnetické momenty zpět do rovnováhy. V případě biomakromolekul s takovým experimentem nevystačíme, vzniká spektrum obsahující příliš mnoho frekvencí, které se často překrývají, a spektrum není možné vyhodnotit. Naštěstí ozařovací pulzy umožňují nejen počáteční sklopení magnetizace, ale i další manipulace s magnetickými momenty jader v průběhu experimentu. Tyto manipulace mohou být poměrně složité. Jejich cílem je vytvořit v molekule takové stavy, při kterých


Vazeb, které leží mezi methylovými vodíky a aldehydovým vodíkem. Ve skutečnosti pozorujeme pro methylové protony signály o dvou frekvencích liších se v důsledku interakce s aldehydovým vodíkem. Podobně signál aldehydového protonu je rozštěpen na čtyři signály v důsledku interakce s protony methylové skupiny.

12.5 Vícerozměrná spektroskopie a korelace frekvencí

V kapitole 12.2 jsme si popsali princip moderního spektrometru, který používá krátké ozařovací pulzy elektromagnetického záření. Při studiu malých molekul nám často stačí jednoduchý experiment s jedním ozařovacím pulzem, který sklopl magnetizaci pozorovaných jader do vodorovné roviny. Po nějakou dobu (přibližně sekundu) magnetizace každého jádra v této rovině rotuje s precesní frekvencí magnetických momentů, které ji tvoří. Během této doby můžeme měřit signál indukovaný v přijímači anténu, než náhodné interakce s okolím přivedou magnetické momenty zpět do rovnováhy. V případě biomakromolekul s takovým experimentem nevystačíme, vzniká spektrum obsahující příliš mnoho frekvencí, které se často překrývají, a spektrum není možné vyhodnotit. Naštěstí ozařovací pulzy umožňují nejen počáteční sklopení magnetizace, ale i další manipulace s magnetickými momenty jader v průběhu experimentu. Tyto manipulace mohou být poměrně složité. Jejich cílem je vytvořit v molekule takové stavy, při kterých
se jaderné magnetické momenty nepohybují nezávisle jeden na druhém, ale jsou určitým způsobem provázané. Toto provázáni mimo jiné umožňuje, že můžeme v jednom experimentu měřit více frekvencí. Mluvíme pak o vícerozměrných spektrech.

Vícerozměrná spektra neobsahují jen informaci o rezonančních frekvencích jednotlivých jader, ale nesou v sobě i údaj o tom, která jádra mají navzájem provázané pohyby magnetických momentů. Říkáme, že mezi takovými jádry existuje korelace. Frekvence, které ve vícerozměrných spektrech korelujeme, mohou být frekvence magnetických momentů stejného izotopu (nejčastěji \(^{1}\text{H}\)) v různých místech molekuly. Pak mluvíme o homonukleární korelací. Může jít ale o frekvence jader různých překů, například \(^{13}\text{C}\), \(^{15}\text{N}\), nebo \(^{31}\text{P}\). V takovém případě, nazývaném heteronukleární korelace, zpravidla potřebujeme, aby tyto izotopy byly přítomny v téměř všech jáderech daného typu (například, aby \(^{13}\text{C}\) netvořil asi jedno procento atomů uhlíku, jak v přírodě, ale aby alespoň 95% uhlíku mělo jádro \(^{13}\text{C}\)). Můžeme proto říci, že studium proteinů pomocí NMR bylo umožněno rozvojem dvou metod. První z nich je expresní metodou heteronukleárního korelace, která nám umožňuje heteronukleární korelace dosáhnout.

Příklad dvourozměrného spektra s heteronukleární korelací je ukázán na obrázku 12.11. V tomto spektru jsou provázané (korelovány) frekvence jádra \(^{1}\text{H}\) a \(^{15}\text{N}\). Použitá korelace je založena na interakcích mezi magnetickými momenty \(^{1}\text{H}\) a \(^{15}\text{N}\), a to interakcí zprostředkovaných elektrony tvořící kovalentní vazbu mezi vodíkem a dusíkem. Ve srovnání s měřením jednorozměrných spektrech frekvencí \(^{1}\text{H}\) a \(^{15}\text{N}\) má tento dvourozměrný experiment několik výhod:

1. Frekvence \(^{15}\text{N}\) měříme s mnohem větší citlivostí, než v jednorozměrném experimentu, který začíná vychýlením magnetizace \(^{15}\text{N}\). Je tomu tak proto, že v našem dvourozměrném experimentu začínáme sklopením magnetizace \(^{1}\text{H}\) a signál \(^{1}\text{H}\) také měříme. Precesní frekvenci \(^{15}\text{N}\) měříme pouze neprímou, experiment je navržen tak, že signál \(^{1}\text{H}\) je modulován frekvencí magnetizace \(^{15}\text{N}\), která se uprostřed experimentu ocitá ve vodorovné rovině. Tak je signál měřen s citlivostí odpovídající velikosti celkové magnetizace \(^{1}\text{H}\), která je v mnohem větší, než magnetizace \(^{15}\text{N}\).

2. Ve spektru nevidíme frekvence všech jáder vodíků (těch je v molekule zkoumaného proteínu několik tisíc), ale pouze frekvence magnetických momentů protonů vázaných přímo na dusík (těch je jenom přibližně dvě tisíce). Proto se frekvence jedrů v různých aminokyselinách a různých chemických skupinách aminokyselin mnohem méně překrývají.

3. Tím, že jsou signály rozprostřeny ve dvou dimenzích, jsou mnohem lépe rozlišeny. Pokud jsou shodou okolností téměř stejné frekvence dvou protonů, většinou jsou dostatečně odlišné frekvence dusíků \(^{15}\text{N}\), na které jsou tyto protony navázány. Při počivé prohlídce spektra jsme tak schopni rozlišit dvojice frekvencí téměř všech NH skupin v proteinu.


12.6 Vztah mezi spektry NMR a strukturou

Zatím jsme se snahili pochopit co to NMR je a jak je možné získat spektra NMR. Změříme teď úhel pohledu a podívejme se na to, co nám NMR spektroskopie může říci o molekulách, které studujeme.

Interakce jader se svým okolím mohou být využity k získání informací o struktuře molekul. Obecná strategie určování struktury pomocí NMR se dá rozdělit na dva druhy experimentů.
První typ experimentů slouží k přířazení naměřených frekvencí odpovídajícím atomům. Znamená to, že vyjde ze známé konfigurace biomakromolekuly a, obrazeně řečeno, na co nejvíce atomů nalépíme štítek, na kterém bude napsána hodnota frekvencí. V těchto experimentech zpravidla měříme dvourozměrná a třírozměrná spektra, která korelují frekvence jednotlivých jader a umožňují tak jejich přířazení. Je-li například v trojrozměrném spektru každý signál definován trojicí frekvencí \(^1\)H, \(^{13}\)C, \(^{15}\)N, koreluje takové spektrum trojice atomů H, C, N, které k sobě nějakým způsobem „patří“. Podle druhu experimentu může „patřit“ znamenat například to, že se dané atomy vyskytují v téže skupině H—C—N. Výsledkem prvního typu experimentů je tedy tabulka udávající rezonanční frekvence jednotlivých jader v molekule. Každé jádro v tabulce je tak označeno jeho rezonanční frekvenci.

Druhý typ experimentů opět využívá korelací, ale tentokrát již vztahujících se ne k topologii, ale k trojrozměrné struktuře. Například dvoul protonů v experimentu sledujícím nukleární Ohradovský jev mezi jádry známé, že dané protony jsou si blízko v prostoru. Kvantitativní měření této korelace (NOE) umožní přibližně určit vzdálenost daných jader. Výsledkem druhého typu experimentů je opět tabulka, která tentokrát shrnuje poznatky o trojrozměrné struktuře, získané z NMR experimentů. Bez využití tabulky první bychom ale více veděli pouze, že například protony o frekvenci 8,73 ppm a 9,64 ppm jsou vzdáleny 0,3 až 0,4 nm. Kombinace výsledků experimentů prvního a druhého typu nám však umožní říci, že takto blízko jsou ve struktuře například vodíky peptidových vazeb osmě a dvacáté aminokyseliny studovaného proteinu.

12.7 Heteronukleární korelace a sekvenční přířazení

Přířazení frekvencí atomům se obvykle provádí ve dvou krocích. Nejprve se přířadí frekvence jader peptidové páté (sekvenční přířazení) a v druhém kroku se provede přířazení postranních řetězců.

Jak jsme si již řekli, interpretace jednorozměrných protonových spektrech proteinů je téměř nemožná ze dvou důvodů. V důsledku obrovského množství signálů je spektrum špatně rozlišené a není známé, který signál náleží kterému atomu. Přiřadíme-li proteinu známý \(^{15}\)N a měříme-li dvourozměrné spektrum s \(^1\)H-\(^{15}\)N korelací, rozlišení se dramaticky zlepší rozvinutím spektra do druhé dimenze (frekvence \(^{15}\)N) a snížením počtu signálů (spektrum obsahuje pouze \(^1\)H-\(^{15}\)N korelace, tedy jen protony vázané na dusík). Jdem-li dále a označíme-li proton \(^{15}\)N a \(^{13}\)C, můžeme měřit spektra trojrozměrná, která nejen že budou ještě lépe rozlišená, ale umožní i určit, které aminokyseliny daný signál náleží. Jinými slovy, heteronukleární korelací (korelace mezi jádry různého typu, která jsou spojena kovalentní vazbou) můžeme použít pro sekvenční přířazení.

Postup sekvenčního přířazení si můžeme ukázat na příkladu krátkého peptidu, složeného z pěti aminokyselin. Obrázek [12.12] ukazuje jedno dvourozměrné spektrum takového peptidu, které poskytuje korelaci \(^1\)H-\(^{15}\)N a, dvou trojrozměrné spektra, označovaná HNCA a HN(CO)CA. Pokud pentapeptid neobsahuje NH skupiny v postranních řetězcích, uvidíme v dvourozměrném spektru čtyři signály (měřitelné budou pouze signály NH skupiny v peptidové vazbě, koncová aminokyselina signály neposkytne, protože její protony snadno disociují a během měření se rychle vyměňují s protony vody). Trojrozměrná spektra rozšiřují korelací na uhlíky α. Ve spektrech HNCA (na obrázku [12.12] modré) jsou korelovány protony a dusíky NH skupiny (v peptidové vazbě) aminokyseliny číslo k s α-uhlíkem téže k-té aminokyseliny. Naopak ve spektrech HN(CO)CA (na obrázku [12.12] červené) jsou korelovány protony a dusíky NH skupiny aminokyseliny číslo k opět s α-uhlíkem, ale tentokrát aminokyseliny číslo k = 1 (tedy prostřednictvím karboxylového uhlíku). Všechny si, že promítnutí spektrum HNCA nebo HN(CO)CA získáme stejnou korelací, jako vidíme v dvourozměrném spektrech. Jinými slovy, signály v trojrozměrných spektrech leží významně v „komínku“ stojícím nad místem signálu v dvourozměrném spektre („komínky“ jsou také znázorněny na obrázku [12.12].

\(^{12}\)NMR experimenty používané pro studium biomakromolekul se často označují podle jader, která jsou v nich korelována. Například zkratka HM(CO)CA nám říká, že jsou korelovány vodík, dusík a uhlík α (CA). Uhlík α je přítom korelován prostřednictvím karboxylového uhlíku (CO), frekvence karboxylového uhlíku ale není měřena, proto je CO v závorce.
KAPITOLA 12. NUKLEÁRNÍ MAGNETICKÁ REZONANCE


Srovnáním HNCA a HN(CO)CA spektre tedy porovnáváme korelace těže NH skupiny s α-uhlíkem dvou sousedních aminokyselin. Zprostředkováváme přes frekvence α-uhlíku se tak dozvíme, které dvě NH skupiny následují v sekvenci po sobě. Můžeme tedy určit pořadová čísla všech NH skupin. K sekvenčnímu přiřazení se častěji využívá dvojice spektre, ve které jsou korelovány nejen α, ale i β uhlíky (spektre HNCACB a CBCA(CO)NH). Zahrnutí β uhlíků snižuje možnost náhodného výskytu shodné frekvence v uhlíkové dimenzi a proto tato varianta experimentů bývá používána nejčastěji.


12.8 Homonukleární korelace a postranní řetězce

Zatím jsme hovořili o korelaci mezi jádry různého druhu (heteronukleární korelace). Velmi užitečné jsou také korelace mezi jádry stejného druhu, tedy mezi jádry, která mají přibližně stejné rezonanční frekvence (homonukleární korelace). Množinu takových jader, která jsou vzájemně propojena interakcí zprostředkovanými elektrony vaze, nazýváme spinový systém. Alifatické aminokyseliny tvoří izolované spinové systémy jader 1H a 13C, zatímco aromatické aminokyseliny se skládají ze dvou izolovaných

13Jiná zhruba tohoto spektra odráží to, že korelace jsou získány trochu jiným způsobem. Experiment HN(CO)CACB je také možné použít (a někdy nutné použít, konkrétně když jsou vodíky aminokyselin s výjimkou NH skupin označené deuteriem), ale jeho citlivost je nižší.
spinových systémů (frekvence \(^{13}\)C alifatické a aromatické části se liší natolik, že je nemůžeme považovat za stejný typ jader).

Metody NMR umožňují měřit spektra, ve kterých jsou korelována všechna jádra spinového systému (jde o takzvaný TOCSY experiment). Protože se sekvenčního příručení známé již frekvence \(\alpha\) a \(\beta\) uhlíků a \(\alpha\) a amidoxy protonů, homonukleární korelace v rámci aminokyselin umožní určit, který atom postranného řetězce patří které aminokyselině. V praxi se takováto homonukleární korelace kombinuje s výše popsanou korelací heteronukleární, která umožňuje zvýšit rozlišení. Příkladem experimentu pro určování frekvencí postranných řetězců je experiment HCCH-TOCSY, ve kterém korelujeme všechny uhlíky spinového systému navázující pomocí TOCSY a každý z uhlíků navíc heteronukleární s protony, které jsou na něj vzdálené. Spojení alifatických a aromatických spinových systémů aromatických aminokyselin je založeno na heteronukleárních korelacích využívaných několika jednovazebných interakcí zprostředkovaných elektrony vazeb. Například experiment (HB)CB(CGCD)HD koreluje alifatický \(\beta\) uhlík s aromatickým \(\alpha\) protonem. Speciální experimenty jsou používány pro některé skupiny postranných řetězců (guanidinovou v argininu, indolovou v tryptofanu, imidazolovou v histidinu).

12.9 NMR spektra a struktura molekul

Seznámili jsme se již s první fázi studia biomarkomolekul pomocí NMR, s fázi, ve které přizpůsobuji frekvence jednotlivým atomům. Pojdeme se teď podívat, jak můžeme pomocí NMR určovat struktury proteinů.

V části 12.3.2 jsme zmínili, že přímé interakce mezi blízkými magnetickými dipolovými momenty jsou příčinou nukleárního Overhauserova efektu (NOE). Intenzitu tohoto jevu můžeme měřit pomocí experimentu zvaného NOESY (zkratka Nuclear Overhauser Effect SpectroscopY). Tento experiment využívá homonukleární korelace, která není založena na interakci magnetických momentů zprostředkované elektrony vazeb, ale na přímé interakci mezi magnetickými momenty, které jsou si blízko v prostoru. Proto intenzitu signalů v NOESY spektru (příklad ukázaný na obrázku 12.13) přímo odpovídá intenzitě NOE. Strnný pokles intenzity NOE s rostoucí vzdáleností jader se často využívá v určování trojrozměrné struktury molekul a nukleových kyselin. S využitím známých vzdáleností protonů (například v CH spinových systémů) můžeme NOESY spektra kalibratovat. Víme například, že vcinální protony v aromatickém kruhu (tedy protony v poloze ortho) jsou vzdálené 0.25 nm. Z teorie dale vímé, že intenzita NOE klesá s šestou mocninou vzdálenosti. Pokud tedy ve spektru například vidíme, že signál odpovídající korelací dvou vicinálních protonů tyrosinu číslo 47 (tedy protony \(Y47\) H\(1\) a \(Y47\) H\(2\)) má intenzitu 64 jednotek a signál mezi protony \(V6\) a \(V\) (valinu 6 a alaninu 124 m8 intenzitu jedné jednotky, víme, že vzdálenost mezi protony \(V6\) H\(6\) a \(A124\) H\(15\) je poloviční, než vzdálenost protonů \(Y47\) H\(2\) a \(Y47\) H\(4\). Tímto přimočarým způsobem můžeme z intenzit signálů ve spektro určit přímo meziatomové vzdálenosti (obrázek 12.14 vpravo). Na základě meziatomových vzdáleností byly určeny první struktury proteinů počítané z experimentálních dat poskytnutých NMR. Mezi různými údaji o struktuře, které NMR spektra poskytují, hrájí meziatomové vzdálenosti dodnes nejdůležitější roli.

Ačkoli je NOE velmi užitečný, není jediným zdrojem informace o konformaci molekul, který NMR poskytuje. Důležitou, byť spíše doplňkovou, úlohu hrají třetí údaje. Jedněm z nich je \(J\)-interakce mezi magnetickými momenty jader oddělených třemi kovalentními vazbami.

Proteže torzní úhly definojí konformaci, je jejich určování z trivazebných \(J\)-interakcí velmi užitečné.14

14Zkratka anglického TOTally Correlated SpectroscopY. Tento experiment je založen na tom, že během měření stále „opravujeme“ mírně se lišící frekvence jader v rámci spinového systému. Proto se jádra nechávají jako jádra s přibližné stejnou frekvencí, ale jako jádra s přibližné stejnou frekvencí.
15Přijínosná vzdálenost znamená 26 = 64 krát větší intenzitu.
16Interakční konstanty, které intenzitu interakcí zprostředkovaných elektrony popisují, se určují buď z intenzit nebo z rozštěpení signálů ve spektech. První postup (takzvané Q3 metody) je založen na porovnání dvou spekter (nebo dvou signálů v téže spektru), z nichž v jednom se interakce uplatní a v druhém je potlačena. Druhý postup využívá nejčastěji takzvaného E.COSY principu, který umožňuje měření i velmi malých interakčních konstant. E.COSY spektra obsahují...
K jednoznačnému určení torzního úhlu ovšem většinou nestačí změřit jednu interakční konstantu, protože Karplusova krivka může nabývat stejné hodnoty pro čtyři různé torzní úhly. Je proto nutno měřit interakce mezi různými atomy definujícími torzi kolem téže vazby. Jak si ale ukážeme v části [12.10], i jediný typ interakční konstanty nám může hodně prozradit o sekundární struktuře.

Konformaci molekuly neodráží jen interakce mezi magnetickými momenty jader, ale i interakce s orbitálními momenty elektronů. Údaj, kterým tyto interakce v roztoku popisujeme, chemický posun, je dokonce nejdůležitější informací při strukturní analýze malých molekul založené na NMR. Chemický posun závisí na elektronové hustotě v okolí měřeného jádra. V rentgenové krystalografii či elektronové mikroskopii je rozložení elektronové hustoty právě to, co se snažíme určit. Vypočítání rozložení elektronové hustoty biomakromolekul z chemických posunů jader je ale příliš složitý úkol. Proto při strukturní analýze proteínů a nukleových kyselin využíváme chemický posun jen jako doplňkový údaj, poskytující informaci o lokální konformaci, nejbližším okolí pozorovaného jádra. Nejčastěji využíváme empirické (odpozorované) korelace mezi hodnotami torzních úhlů a chemických posunů k tomu, abychom odhadli hodnoty torzních úhlů $\phi$ a $\psi$ v přátří proteinu.

Poslední údaj o konformaci, který zmíníme, vyžaduje zvláštní přípravu vzorku proteinu či nukleové kyseliny. Tímto údajem je opět přímá interakce mezi magnetickými momenty blízkých jader. Tentkrát nejde o vliv okamžitých interakcí na návrat magnetických momentů do rovnováhy, jak tomu bylo případě NOE, ale opravdu o měření průměrné hodnoty této interakce. A tato průměrná hodnota je v běžných roztocích nulová. Jak ji tedy můžeme získat?

To, co jsme si o dipol-dipolových interakcích dosud řekli, vycházelo z předpokladu, že se mohou molekuly v roztoku otácer ve všech směrech. Pokud ale některý směr zvýhodněme (umístíme molekulu do anizotropného prostředí, například do kapalného krystalu, který se uspořádá v silném magnetickém poli spektrometru), nebude již průměr pole sousedních jader nulový. V tomto případě pozorujeme ve spektrech štěpení signálů, které závisí na částečně orientací molekuly ve zkoumavci. Měření zbytkových dipolárních interakcí nám umožní určit polohu vektoru mezi interagujícími jádry všech vnějších souřadnic soustavě. Tato vnější souřadná soustava je určena osami, které popisují volnost molekuly rotovat v různých směrech. Tyto směry jsou zase dány směrem $\vec{B}_0$ a druhem kapalného krystalu.

Pokud změříme zbytkovou dipolární interakci pro jednu dvojici jader (geometricky popsanou výše zmíněným vektorem směřujícím od jednoho jádra k druhému), neumožní nám to ještě jednoznačně určit orientaci této dvojice v naší vnější souřadné soustavě. Pokud ale naměříme zbytkových interakcí více v rámci části proteinu, jejichž lokální konformace známe, například v dobře určených $\alpha$-sroubovících, můžeme opravdu spočítat, jak jsou tyto sroubovice v proteinu orientovány (obrázek [12.14] vpravo). Totéž platí například pro vzájemnou orientaci domén většího proteinu.

Zbytkové dipolární interakce jsou velmi užitečné proto, že vůbec nezávisí na vzdálenosti části molekuly, jejíchž vzájemnou orientaci určujeme. Je tomu tak proto, že všechny orientace vlastně určujeme všichni jedno společné souřadnice soustavě. V tom se liší zbytkové dipolární interakce od NOE, které nám poskytnou údaje pouze o kontaktu části molekuly, jejichž protony se nacházejí v prostoru blízko sebe. Z toho vyplývá, že zbytkové dipolární interakce hrají nezastupitelnou úlohu při studiu malých kompaktních molekul, pro které potřebujeme určit vzájemnou orientaci částí, které jsou od sebe v prostoru dáleko.

### 12.10 Sekundární struktura

Detailní popis struktury na základě NMR měření vyžaduje poměrně složité zpracování získaných dat, které si popíšeme v sekci [12.11]. Teď se ale zkusme podívat na to, co se o struktuře můžeme dozvědět rychle a bez složitých analýz.

**Signály rozštěpené ve dvou dimenziích.** Ze čtyřice signálů vzniklých štěpením se dva odcitkou a zbyvající dva sečtou. Pokud využijeme silné interakce v jedné dimenzi, můžeme odcitk a velmi malé štěpení (slabou interakci) v druhé dimenzi.
12.11. VÝPOČET KONFORMACE PROTEINU

Na základě několika základních experimentů umíme popsat prvky sekundární struktury proteinu. Využíváme k tomu především údaje, které jsou shrnuty v následujícím výčtu.


2. Vyhodnocením NOE mezi α-protony a amidovými protony peptidové vazby můžeme určit sekundární strukturu podle následujících pravidel. Ve skládaných listech je vzdálenost mezi α-protony a amidovými protony sousední aminokyseliny menší než vzdálenost mezi dvěma amidovými protony sousedních aminokyselin, zatímco ve šroubovících je tomu naopak. Navíc v α-šroubovících je malá vzdálenost mezi α-protony a β-protony aminokyseliny vzdálené o tři rezidua. Dobre pozorovatelné jsou i NOE mezi α-protony a amidovými protony aminokyselin vzdálených o tři a čtyři rezidua. Ve šroubovicí 310 je na rozdíl od α-šroubovice vzdálenost mezi α-protony a amidovými protony aminokyselin vzdálených o tři rezidua kratší než vzdálenost mezi α-protony a β-protony těchto aminokyselin.

3. V téměř rovinném uspořádání skládaného listu jsou silnější J-interakce mezi α-protony a amidovými protony peptidové vazby než ve šroubovících. Hodnoty těchto snadno měřitelných trávazebných interakčních konst - ant bývají ve skládaných listech vyšší než 8 Hz a ve šroubovících nižší než 5 Hz. V praxi se proto obvykle skládané struktury rozlišují pouze na základě této interakce, aniž by se měřily všechny interakční konstanty potřebné k matematicky jednoznačnému výpočtu torzního úhlu φ. Torzní úhel ψ se zpravidla z J-interakcí nepočítá vůbec, protože příslušné interakční konstanty jsou špatně měřitelné.

12.11 Výpočet konformace proteinu

Z výsledků NMR experimentů získáme velké množství geometrických údajů, zejména vzdálenosti mezi atomy vodíku, některé torzní úhly, informaci o sekundární struktuře a podobně. Na rozdíl od elektronové hustoty, kterou získáme z rentgenové difrakce, z výsledků NMR měření není na první pohled patrný tvar molekuly. Jsme v podobné situaci, jako když chceme nakreslit plán domu a k dispozici máme jen tabulku číselných vzdáleností. Musíme proto výsledky měření zpracovat do grafické formy. Výpočetní metody, které umožňují ze vzdáleností vytvořit trojrozměrný model, jsou označovány jako distanční geometrie.


Abychom získali chemicky rozumný model proteinu, používáme pro výpočet konformace metod molekulového modelování, které „vědí“, jak mají aminokyseliny vypadat. Nejčastěji se používá molekulová dynamika, ve které jako startovní konformaci můžeme použít třeba výsledek distanční geometrie. Mezi výpočtem struktury, jak jsme jej popisovali v kapitole 5 a využitím molekulové dynamiky v rámci 17

17Výjimkou je chirality. Ze vzdálenosti nejsme schopni určit, který z vzájemných zrcadlových obrazů je ten pravý. Naštěstí víme, že proteinky jsou stavěny z l-aminokyselin, takže nás tato nedokonalost nijak neomezuje.
určování struktury z výsledků NMR je ale důležitý rozdíl. Pokud vycházíme z výsledků NMR experimentů, používáme naměřené hodnoty (vzdálenosti atomů, torzní úhly, zbytkové dipolové interakce) jako dodatečné vazné podmínky. Molekulová dynamika se tedy nesnaží najít celkově nejvýhodnější konformaci, ale vybírat chemicky nejprávější strukturu v rámci možností, kterým odpovídají experimentální data. Jinými slovy, NMR vymeze na hyperplášť potenciální energie menší území a cílem molekulové dynamiky je najít v tomto území nejhlubší údolí.


12.12 Struktura nukleových kyselin pomocí NMR

Ačkoli se při určování struktur nukleových kyselin využívá stejné principy jako v případě proteinů, postup práce je v mnohem odlišný. Důvodem je nákladnost izotopového značení nukleových kyselin a odlišnost struktury, která činí NMR nukleových kyselin experimentálně obtížnější. Se srovnáním s proteiny se NOESY experimenty často používají nejen v druhé fázi řešení struktury, ale i k přiřazení frekvencí. Techniky frekvenčního přiřazení můžeme u nukleových kyselin dělit na experimenty využívající interakce zprostředkované elektrony kovalentních vazeb a na experimenty využívající NOE (dipolární interakce prostorem). Oba přístupy mají své výhody a nevýhody a v praxi se často kombinují. Zde si uvedeme jen příklady postupů založených na využití NOE.

Metody využívající NOE k frekvenčnímu přiřazení jsou založeny na korelacích protonů nacházejících se blízko sebe v prostoru. V principu tedy nevyžadují izotopové značení vzorku. V jednodušších případech lze tedy tímto způsobem studovat snadno dostupné vzorky nukleových kyselin v přirozeném zastoupení. Pokud je již k dispozici značený vzorek, kombinují se tyto metody často s heteronukleární korelací s uhliky a protony, protože takto lze zvýšit rozlišení rozšířeného spektra do další dimenze. Nevýhodou tohoto přístupu je, že vzdálenosti mezi protony a tedy i pozorované korelace závisí na konformaci dané molekuly nukleové kyseliny.

**Korelace mezi imino (a amino) protony**


**Korelace báze-cukr**

Pro přiřazení frekvencí jsou cenné korelace H6/H8 protonů s protony H1', H2' (silnější u dvojšroubovice A), H2" (silnější u dvojšroubovice B) pentosového krku. Protože lze pozorovat korelace s vlastní (deoxy)ribosou a s cukrem předchozího nukleotidu, poskytují zmíněné korelace údaje pro sekvenční přiřazení (obrázek [12.16]). Zároveň získáváme informaci o přiřazení cukru a báze v rámci jednoho rezídua.
12.13. STRUKTURA OLIGOSACHARIDŮ POMOCÍ NMR

Korelace protonů cukru

Korelace mezi protony v rámci (deoxy)ribosového kruhu lze využít k přiřazení frekvencí protonů spinovým systémům jednotlivých bází. Navíc lze pozorovat korelace mezi protony cukrů sousedících nukleotidů. Nejsilnější bývají korelace H5''(i)-H2''(i-1) v dvojšroubovicích B a H5''(i)-H1'(i-1) v dvojšroubovících A.

12.12.1 Nukleární Overhauserův jev a vzdálenosti protonů

Využití NOE k určování vzdáleností mezi protony je u nukleových kyselin stejné jako u proteinů. Vzhledem k nižšímu počtu protonů v nukleových kyselinách bývá ale počet získaných meziatomových vzdáleností u nukleových kyselin nižší ve srovnání s proteiny. Při určování struktury závisí proto výsledná konformace více na závěrečném výpočtu molekulovou dynamikou, než tomu je v případě proteinů.

12.12.2 Třívazebné interakce a torzní úhly

Třívazebné úhly v molekulárních nukleových kyselinách se také obvykle určují z hodnot třívazebných interakčních konstant. Relativně velké množství třívazebných interakcí je dostupné pro určování konformací cukrů. Při interpretaci těchto dat ale můžeme narazit na komplikaci způsobené přítomností více konformací ve vzorku. Torzí kolem glykosidické vazby (χ) určujeme z třívazebných interakcí mezi atomy H1'-C2 a H1'-C6 u pyrimidinů a H1'-C4 a H1'-C8 u purinů. Z torzních úhlů patří lze poměrně přesně určit β (z třívazebných interakcí mezi atomy H5''-P5', H5'-P5', C4'-P5'), ε (z třívazebných interakcí mezi atomy H3'-P3', C2'-P5', C4'-P3') a γ (z třívazebných interakcí mezi atomy H5''-H4', H5'-H4', H5''-C3', H5'-C3').

12.12.3 Výpočet konformace nukleových kyselin.


12.13 Struktura oligosacharidů pomocí NMR

Určování struktury (nederivatizovaných) sacharidů pomocí NMR je komplikováno nedostupností izotopově značených vzorků a špatným rozlišením protonových spektér (ze struktury vyplývá, že většina protonů má velmi podobné okolí). Sekvenční přiřazení se provádí pomocí NOE, přiřazení v rámci sacharidových jednotek je založeno na NOESY a TOCSY experimentech. Počet NOE korelací pro určování struktury mezi různými jednotkami sacharidů je většinou malý. K určování torzních úhlů v rámci jednoho cukerného kruhu lze využít třívazebné interakce mezi protony. K určování torzí kolem glykosidických vazeb je nutno měřit třívazebné interakce mezi protony a uhlíky (v přirozeném zastoupení
KAPITOLA 12. NUKLEÁRNÍ MAGNETICKÁ REZONANCE

12.14 Pohyby molekul v NMR

Během měření NMR sledujeme nejen amplitudu a frekvenci záření jednotlivých jader, ale také postupné slábnutí tohoto záření − relaxaci. Signál během měření slábne protože se postupně ztrácí což není nic jiného než nám dobře známý nukleární Overhauserův efekt). Z těchto příznaků měříme relaxaci jednotlivých jader, například rychlost, s kterou se v roztoku otáčí celá molekula. Program Relaxace umožňuje měřit více různých podmínek.

Při sledování dynamiky většinou mieríme tři relakční údaje (rychlost ztráty koherencí, rychlost návratu do rovnováhy a účinnost přenosu energie na sousední jádra).

Informace o pohyblivosti může být důležitá pro pochopení funkce studované molekuly.

12.15 Další vývoj metody

Použitelnost NMR pro studium biomakromolekul má v současnosti různé omezení. Některá z nich vyplývá cizí činnost studovaných molekul a rozdíl mezi energiemi základního a excitovaného stavu.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s relaxací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.

Další omezení souvisí s koncentrací. Obecně lze říci, že čím je větší molekula, tím pomaleji se pohybuje v roztoku, tím více času mají jádra interagovat a tím rychlejší je relaxace. Například u proteinů obsahujících více než 200 aminokyselin je relaxace tak rychlá, že se nesmí mít dovolit pohyb na několik sekund po excitaci, ale měření můžeme opakovat dejíře po jedné sekundě.
v důsledku interakcí se sousedním jádrem. Při určitém magnetickém poli tak téměř nedochází ke ztrátě koherence. Příspěvky k relaxaci v důsledku interakcí se vzdálenějšími protony lze potlačit přípravou vzorku proteinu značeného deuteriem ($^2$H), v případě proteinů expresí v mediu připraveném z těžké vody ($^2$H$_2$O). Tímto způsobem lze studovat i velké biomakromolekuly, například proteiny složené z tisíců aminokyselin.

Dosud jsme mluvili jen o rozpustných proteinech. Velké úsilí je také vynáškáno na studium membránových proteinů. Membránové proteiny můžeme zkoumat buď v roztoku, kdy je zabudováno do malých lipidových micel, nebo přímo v membránách v pevné fázi. Měření v pevné fázi jsou technicky náročnější než měření v roztoku a je mnohem obtížnější určit přesnou konformaci větších molekul. Struktury několika proteinů však již byly určeny pomocí NMR v pevné fázi.
Obrázek 12.11: Příklad dvourozměrného NMR spektra proteinu, poskytujícího korelace $^1$H a $^{15}$N. Spektrum je ukázáno jako vrstevnicová mapa, na svislé a vodorovné ose jsou vyneseny precesní frekvence $^{15}$N a $^1$H, vrstevnice popisují intenzitu signálu.
Obrázek 12.12: Příklad sekvenčního přiřazení pomocí experimentů HNCA a HN(CO)CA. Nahoře je schema korelace v jednotlivých experimentech (černé korelace $^1$H a $^{15}$N, modré korelace v HNCA, červené korelace v HN(CO)CA) a naměřená spektra. Dole je postup sekvenčního přiřazení.

Obrázek 12.14: Schematické znázornění strukturních informací, které můžeme z NMR spekter získat pro aminokyseliny, které leží od sebe daleko v sekvenci proteinu. Vlevo je zeleně znázorněno měření vzdálenosti mezi protony methylových skupin v různých α-šroubovicích. Tuto vzdálenost lze určit z hodnoty NOE mezi měřenými protony. Vpravo je zelenými šipkami ukázána vzájemná orientace dvou α-šroubovic, kterou lze spočítat z hodnot zbytkových interakcí měřených v jednotlivých šroubovicích.

Obrázek 12.15: Příklady NOESY spekter oblasti imino-protonů bází. Červeně je znázorněna korelace mezi protony v sekvencích (tečkování je označena nepozorovatelná korelace mezi protony s identickou rezonanční frekvencí) a sekvenci informace ve spektrech. Vzhledem k tomu, že jde o diagonálně symetrická spektra, sekvenci informaci lze sledovat dvěma způsoby (plná a přerušovaná červená čára).
Obrázek 12.16: Vlevo schema korelací mezi protony H1' (v pentose) a H6 nebo H8 (v bázi), vpravo výsek oblasti NOESY spektra, která obsahuje tyto korelace. Červeně je znázorněna sekvenční informace. Protože je na obrázku ukázána pouze část spektra, sekvenční informaci můžeme sledovat jen jedním způsobem. Všimněte si, že proton první báze je korelován pouze s H1' vlastní pentosy a proton H1' poslední báze je korelován pouze s H6 nebo H8 protonem vlastní báze (v těchto případech chybí korelace se sousedem) – na základě toho je možné určit směr sekvence.
KAPITOLA 12. NUKLEÁRNÍ MAGNETICKÁ REZONANCE
Kapitola 13

Bioinformatika

13.1 Co to je bioinformatika

V předchozích kapitolách jsme si popsal možnosti, pomocí kterých můžeme zkoumat strukturu velkých biologických molekul. Za poměrně krátkou dobu používání těchto metod se biochemici a biologové dosud nikdy o struktuře biomakromolekul velmi mnoho. Dokonce lze říci, že je jich poměrně mnoho, neboť již v průběhu času se v zářivě zavěšených struktur zváženou. Metody určování struktur biomakromolekul stále nejsou dokonalé a na jejich zlepšení se usiluje nejvíce. Můžeme proto očekávat, že se nové informace o strukturách bude dodávávat na vědecký fond, ale také na hlavní biochemika padá, bude rostoucí, ale přesněji. Protože nové poznatky nelze zvládnout lidskou silou, posédliváním v knihovně s tužkou v ruce, vznikl nový vědní obor, který se snaží vznést do záplavy informací řád s pomocí počítačů. Pro tento obor se ují název bioinformatika.

Bioinformatika se narodila v osmdesátých letech dvacátého století. Matkou její byla se vývojem počítačové technologie, otcem zejména sekvenování co nejvíce DNA. Přečtení úseky DNA bylo nutno vzájemně srovnávat a analyzovat. Úroveň výpočetní techniky byla schopná mnoho dat zvládnout. O několik let později začalo rychle růst počet biomakromolekul, u kterých se podařilo určit prostorovou strukturu. Také strukturní informace bylo třeba třídit. Dnes je bioinformatika nejdůležitější aplikací počítačů v biologických vědách. Chceme-li tento nový obor definovat, můžeme říci, že bioinformatika je informační technologie sloužící k uchování, zpracování a analýze biologických dat.

13.2 Genomové projekty

V sekci 8.2 jsme si ukázali, že určitý sekvenci úseků DNA je poměrně snadné. To připoměnil molekulární biolog na myšlenku přečíst celou genetickou informaci některých organismů. Již v roce 1977 byl sekvenován genom vírového X174, obsahujícího 5 000 nukleotidů a 11 genů. V roce 1995 byla určena první úplná sekvence gen Photoshop Influenciae, s 1,8 miliony nukleotidů a 1 700 genů. Cílem genového projektu bylo získat úplné sekvence DNA modelových organismů: střevní bakterie Escherichia coli, kvasinky Saccharomyces cerevisiae, cerva Cenorhabditis elegans, huseníčku rolního (Arabidopsis thaliana), mušky Drosophila melanogaster, myšky Mus musculus, psa (Canis familiaris) a autora celého projektu (Homo sapiens).

KAPITOLA 13. BIOINFORMATIKA

Zmínili jsme se již v sekci 7.2, že vyjádření (exprese) informace uložené v DNA je složitý proces. Je proto dobré si uvědomit jednotlivé úrovni genetické informace:

1. **Genom** je informace zapsaná v DNA tvořící chromosomy, je společná všem buňkám organismu.
2. **Transkripom** je část genomu, která je exprimována v buňce v dané fázi jejího vývoje.
3. **Proteom** je sada vzájemně interagujících molekul geneticky kódovaných proteinů, která dává buňkám její individuální charakter.


### 13.3 Biologické databáze

Biologické databáze jsou vlastně elektronické kartotéky biologicky zajímavých údajů. Díky své elektronické formě jsou mnohem pružnější než to, co si obvykle pod pojmem „kartotéka“ představíme (plechové skříně plné kartíček z tvrdého papíru). Údaje v elektronických databázích můžeme mnohem snadněji třídit, analyzovat, porovnávat s jinými daty.


### 13.4 Databáze sekvencí proteinů

#### 13.4.1 Primární databáze sekvencí proteinů

Známé sekvence proteinů jsou uchovávány v různých primárních databázích. Zkusíme si teď udělat přehled alespoň těch nejdůležitějších a nejúžitečnějších.
13.4. DATABÁZE SEKVENCÍ PROTEINŮ

13.4.1. Příklad jednoduchého textového souboru.
Obrázek 13.2: Databáze odkazů na odborné články jako příklad relační databáze.


- **PIR** (Protein International Resource) je americká databáze založená na myšlence co nejrychlejšího zveřejňování nových sekvencí. Tato databáze obsahuje největší počet položek s poměrně nízkou úrovní anotací. Evropským centrem, které shromažďuje sekvence ve spolupráci s PIR je Martinsried Institute for Protein Sequences (MIPS).

- **TrEMBL** je doplněk databáze SWISS-PROT. Obsahuje všechny sekvence z databáze EMBL, které kódují nějaký protein. Anotace jsou tvořeny automaticky počítačem, což umožňuje rychlý, ale méně kvalitní popis sekvence. Zkontrolované položky jsou potom ručně anotovány a přesouvány do databáze SWISS-PROT.

- **NRL-3D** je databáze sekvencí všech proteinů, jejichž prostorová struktura (konformace) byla určena a uložena v databázi PDB (sekcí [13.6]).


---

1Databáze sekvencí DNA, která bude zmíněna v sekci [13.3]
13.4. Databáze sekvencí proteinů

13.4.2 Složené databáze sekvencí proteinů

Nepohodlí vyplývající z toho, že existují více databází proteinových sekvencí, se pokouší odstranit složené databáze, které kombinují údaje z primárních databází a snaží se, aby se v nich každá sekvence ocítila jen jednou. Možná vás překvapí, že i složených databází je několik. Všechy se snaží sjednotit údaje z několika primárních databází, každá z nich ale přístupuje k tomuto úkolu jinak. Složené databáze se liší tím, zda kladou více důraz na kvalitu nebo aktuálnost a také tím, kterou primární databázi používají jako základní zdroj informací.

NRDB je databáze poskytující nejaktuálnější přehled sekvencí. OWL, zaměřený především na SWISS-PROT, je pomalejší, ale lépe anotovaný. MIPSX je naopak založen na PIR. Za složenou databázi můžeme považovat i kombinaci SWISS-PROT+TrEMBL, která spojuje přednosti obou svých součástí.

13.4.3 Sekundární databáze sekvencí proteinů

Zatím jsme hovořili pouze o databázích, které obsahují jednotlivé sekvence proteinů s příslušným popisem. Podívejme se nyní na databáze sekundární, které obsahují již výsledky analýzy sekvencí. Nejdříve si ale musíme vysvětlit, co takovým výsledkem analýzy sekvence vlastně je.


Nejjednodušším příkladem abstrakce je sekvenční motiv. Sekvence proteinů obsahuje obvykle úseky, které jsou velmi podobné částem sekvencí jiných proteinů a úseky, které se naopak ostatním sekvencím podobají málo. Úseky sekvencí, které se s drobnými obměnami objevují v různých proteinech, se nazývají motivy. Podle zacházení s motivy můžeme sekundární databáze dělit na následující tři skupiny:


2. Databáze obsahující soubor všech motivů v sekvenci (takzvaný fingerprint, česky otisk prstu sekvence). Fingerprints mohou být definovány na základě frekvenci matic (tabulek užívajících, jak často se k některé aminokyselině nachází na daném místě; příkladem je databáze PRINTS) nebo vážených matic (tabulek užívajících nejen kolikrát je na daném místě například glycin a ne alanin, ale zahrnující také pravděpodobnost, že dojde k náhodné záměně glycinu za alaninu; příkladem je databáze BLOCKS).
3. Databáze obsahující nejen motivy, ale výsledky analýz celých sekvencí. Používanou abstrakcí v těchto databázích jsou *profily* (tabulky udávající, jak často se na daném místě nachází určitá aminokyselina nebo jak často na tomto místě dochází k inzerci nebo deleci*) příkladem je databáze *PROFILES* nebo *skryté Markovovy modely* (matematické statistické algoritmy, které popisují výstavbu sekvence jako sled inzercí, delecí, nebo shod v aminokyselině na každém místě sekvence; příkladem je databáze *Pfam*).

4. Databáze INTERPRO, která se snaží sloučit všechny výše zmíněné databáze.


### 13.5 Databáze sekvencí DNA

V sekci 13.4 jsme si popsali databáze, které uchovávají sekvence proteinů. V databázích jsou proteínové sekvence uloženy jako pořadí aminokyselin pomocí nám dobře známé abecedy, která se skládá z dvaceti písmen. Ve skutečnosti je ovšem většina sekvencí proteinů určena na základě sekvencí DNA a z abecedy DNA do abecedy proteinů přeložena pomocí genetické kóduvací tabulky (takzvaná koncepční translace). Známých sekvencí úseků DNA je ale více než těch, které tvorí geny určitého proteinu. Během desetiletí byla v laboratořích přečtena spousta úseků DNA, které nekódují žádný protein, nebo kratkých úseků, ze kterých nemůžeme určit úplnou sekvenci proteinu. Aby bylo vneseno jakýsi řad i do těchto dat, bylo založeno několik databází, které ukládají sekvence úseků DNA. Nejdůležitější jsou následující tři databáze:

- **EMBL** (*European Molecular Biology Laboratory*). Sekvence jsou do této databáze vkládány buď přímo těmi, kteří sekvenci provádějí, nebo jsou převzaty z výsledků genomových projektů, z odborné literatury, nebo z patentů. Počet položek v databázi se každý rok přibližně zdvojnásobí. Pro vyhledávání sekvencí v EMBL slouží vyhledávací SRS (viz sekce 13.7).

- **GenBank** je obdobná americká databáze vyvinutá v National Center for Biotechnology Information (NCBI) při National Institute of Health (NIH) v Bethesda. Data jsou v ní rozdělena do 17 divizí a k jejich vyhledávání slouží systém Entrez (viz sekci 13.7).

- **DNA Data Bank of Japan** (DDBJ) je třetí důležitou databází, vyvinutou japonským Národním ústavem genetiky. K vyhledávání v ní slouží systém DBGet (viz sekci 13.7).

Všechny tři uvedené databáze úce spolupracují, každý den si vyměňují nové položky. Obsahem by se proto neměly lišit, takže rozhodnutí, kterou databázi využít, vyplývá hlavně z toho, který vyhledávací systém nám více vyhovuje.

Vedle zmíněné trojice databází, která uchovává všechny DNA sekvence s anotací, existují i databáze specializované. Jako příklad můžeme uvést databázi *GSDB*, která uchovává sekvence určené v rámci genomových projektů, nebo databázi *dbEST*, která obsahuje takzvané *exprimované sekvenci značky (EST)*.

---

*)Delece je vypuštění aminokyselin ze sekvence, inzerci je zařazení další aminokyseliny.

*EST* (*Expressed Sequence Tag*) je kratký následně zvolený úsek sekvence cDNA (viz sekce 7.2.3), pomocí kterého můžeme sledovat, jestli dochází k expresi daného genu v určité tkáně.
13.6 Databáze proteinových struktur

Presná prostorová struktura byla určena jen pro malou část proteinů, jejichž sekvence známé. I tak se známé struktury počítají na tisíce. Nepřekvapí proto, že byly vytvořeny databáze, ve kterých jsou struktury proteinů (a nejen proteinů) uloženy.

Nejdůležitější databáze struktur je PDB (Protein Data Bank), která byla vyvinuta v Brookhaven National Laboratory a je udržována v rámci Research Collaboratory for Structural Bioinformatics (RCSB). PDB uchovává struktury všech proteinů (ale i jiných biomakromolekul), které byly určeny rentgenovou difrací, nukleární magnetickou rezonancí, kryoskopíí, nebo výpočtem. Struktury jsou uloženy v textových souborech, které obsahují informace o proteinu a souřadnice jeho atomů. Každá struktura je označena jedinečným identifikačním kódem, složeným ze čtyř znaků (číslic a písmen), například „1DUF“. Struktury je možné hledat také podle klíčových slov. Souřadnice atomů můžete uložit ve svém počítači, nebo si můžete trojrozměrný model molekuly prohlédnout na internetových stránkách PDB.

Kromě PDB existují i databáze obsahující analýzy struktur, podobně jako sekundární databáze obsahující analýzy sekvencí. Databáze PDBSum obsahuje shrnutí informací z PDB a přehled analýzy struktury. Další databáze analyzující struktury slouží klasifikaci proteinů. Nejstaršími takovými databázemi jsou SCOP a CATH.

Databáze SCOP (Structural Classification of Proteins) byla vyvinuta v MRC Laboratory of Molecular Biology ve Velké Británii. Struktury jsou porovnávány kombinací automatických a ručních metod a tříděny na třech úrovnících. Nejvýše stojí v hierarchii fold. Do stejného foldu patří proteiny se stejnými elementy sekundární struktury, ale bez známky příbuznosti. Další úrovni je superfamília (superfamily), sdružující proteiny s podobnou strukturou i funkcí (což svědčí pro jejich společný původ), ale s velmi rozdílnými sekvencemi. Konečně do stejné rodiny (family) patří proteiny, které se podobají strukturou i sekvencí (víc než 30% aminokyselin shodných).

Odlišnou strukturní hierarchií, obsazenou přímo v názvu, používá databáze CATH (Class-Architecture-Topology-Homology), vyvinutá na University College London. Nejvýše stojí v hierarchii třída (class). Každý protein je zařazen do jedné ze čtyř tříd: α-proteiny, β-proteiny, α-β-proteiny (obsahují proteiny α/β i α+β) a proteiny s nízkým obsahem sekundární struktury. Další úrovni je architektura (architecture), která popisuje vzájemné prostorové uspořádání prvků sekundární struktury, aniž by se zabývala tím, jak jsou tyto prvky propojeny. Na třetí úrovni, topology, jsou pak proteiny dále tříděny podle topologie spojení prvků sekundární struktury. Čtvrtou úrovni je homology. Do skupiny se stejnou homologii patří proteiny, které mají více než 35% sekvence shodné (předpokládá se, že mají společného předchůdce). Na rozdíl od databáze SCOP je CATH tvořena automaticky.

Vzhledem k nejednotné třídění proteinů nás asi nepřekvapí, že různé databáze poskytují různé výsledky klasifikace. Ani terminologie zde není ustálená. Dá se říci, že názvosloví klasifikace proteinů má často původ právě v programech, které se o automatické třídění proteinů pokoušejí.

13.7 Vyhledávací systémy


Systém SRS (Sequence Retrieval System) je udržován EBI (podobně jako databáze EMBL). Jde o síťový prohledávač zvolených molekulárně-biologických databází, které hypertextově propojuje. SRS umožňuje vzájemné indexování (propojení pomocí odkazů) databází obsahujících textové soubory, ať už jde o sekvence, struktury, nebo citace literatury.

Prohledávání databází často používáme k tomu, abychom nalezli členy proteinových rodin v různých biologických druzích. Databáze *Entrez* prohledává podle textových údajů, ale tento přístup často sehnává kvůli chybějícím nebo neúplným anotacím sekvencí. Proto je spolehlivější porovnávat přímo sekvence a snadno se odhalit evoluční příbuznosti (*homologie*) na základě podobnosti sekvencí. Na tento úkol se zaměříme v následujícím podkapitole.

### 13.8 Párové příložení (pairwise alignment) sekvencí

Chceme-li porovnat dvě sekvence, hledáme způsob, jak je příložit vedle sebe tak, aby proti sobě bylo co nejvíce stejných aminokyselin (obrázek [13.3]). Takový způsob se nazývá anglicky *pairwise alignment*, což můžeme do čestiny přeložit jako *párové příložení*. Nás úkol je komplikován tím, že u jedné sekvence mohlo dojít k vyřazení (*deleti*) některé aminokyseliny, nebo naopak k zařazení (*inseri*) aminokyseliny nové. Proto lze lepšího příložení dosáhnout tehdy, když máme možnost do sekvence vkládat i mezery (anglicky *gap*). Tohoto prostředku musí být ale využíváno s měrou. Neomezovaným vkládáním mezer můžeme samozřejmě dosáhnout toho, že i úplně odlišné sekvence přímo způsobí, aby proti sobě výzdvížely lépe stejných aminokyseliny. Již z tohoto jednoduchého příkladu je zřejmé, že párové příložení je poměrně složitý úkol. Řešíme jej pomocí matematických modelů, jejíchž vlastnosti méníme nastavením různých parametrů.
13.9. VÍČENÁSOBNÉ PŘILOŽENÍ (MULTIPLE ALIGNMENT)

Nalézt párové přiložení celé dlouhé sekvence (globální přiložení) je výpočetně velmi náročné. Algoritmy párového přiložení se proto často opírají o hledání podobnosti v rámci krátkých úseků sekvence (lokální přiložení). Tak pracují nejběžnější algoritmy *FastA* a *BLAST*. Vylepšená verze algoritmu *BLAST*, která umožňuje vložení mezer, je dnes nejpoužívanější metodou prohledávání databází. Algoritmus nejprve vyhledává přiložení krátkých úseků sekvence bez mezer, pak se snaží rozšířit přiložení tak, aby proti sobě leželo co nejvíce shodných aminokyselin. Pro měření shody se počítá určité skóre, takže hledáme páry s vysokým skóre (*high-scoring pairs*). Nakonec se přiložení vylepší vložením mezer. Abychom mohli odhadnout, nakolik přiložení odhaluje skutečnou příbuznost sekvencí, výsledky algoritmu obsahují i údaj o pravděpodobnosti, že přiložení je dílem náhody (*hodnota p*) v algoritmu *BLAST*, nebo očekávaný počet úseků dané délky, který jsou náhodou shodné v prohledávane databázi (*hodnota E*). Máme-li přiložení věřit, musí být hodnoty *p* i *E* co nejméně.

13.9 Vícenásobné přiložení (multiple alignment)

Často chceme vědět nejen to, jak se sobě navzájem podobají dvě sekvence. Chceme zkoumat vzájemnou podobnost v rámci celé skupiny sekvencí, která například patří do stejného proteinové rodiny. Snažíme se proto přiložit k sobě více sekvencí tak, abychom vytvořili tabulku, ve které bude na každém řádku jedna sekvence a ve sloupcích pod sebou budou co nejpodobnější aminokyseliny. Tato metoda, zvaná *multiple alignment* (vícenásobné přiložení), je neoposredně podobná analýze genových rodin a tvorbě sekundárních databází.

Nejlepší způsob přiložení je poskládat sekvence tak, aby každý sloupec tabulky tvořily aminokyseliny, které jsou na stejném místě ve struktuře proteinu. Protože prostorovou strukturu známe jen pro omezený počet proteinů, musíme často zakládat tabulky vícenásobného přiložení jen na základě sekvencí, podle podobnosti aminokyselin. Nápiklad kyselina asparagová by se měla nejlépe očitnout ve sloupecu tvořeném kyselinou asparagovou. Pokud to nelze, tak by měla být ve sloupci tvořené zejména kyselinou glutamovou (tě se nejvíce podobá náboji) nebo asparaginem (tomu se podobá velikostí a tvarem postranního řetězce). Jíž méně je kyselina asparagová podobná krátkým polárním aminokyselinám a velmi se liší od aminokyselin s dlouhými nepolárními postranními řetězci.

Je možné hledat najednou nejlépe přiložení celé skupiny sekvencí. Takovým způsobem pracují *simultánní metody*, které jsou však výpočetně náročné. Častěji se postupně budovává tabulky přiložení pomocí *progressivních (postupných) metod*. Oběhovým programem založeným na postupném přístupu je *CLUSTAL*. Tento program k sobě postupně přikládá dvojice sekvencí, trďí je podle vzájemné podobnosti a vyváží výsledný vztah přibuznosti, jaký znáte ze slechtických rodokmenů. Nejpodobnější sekvence odpovídají sousedním větvím stromu, zatímco málo podobné sekvence představují vzdálené větve. Rozšiřenou verzí tohoto programu je *CLUSTAL W*, který se na základě mezer již vložených do sekvencí snaží vkládat mezery další a tak hledat způsoby přiložení velmi málo podobných sekvencí (obrázek 13.4).

Jiným užívaným programem je *PSI-BLAST* (*Position-Specific Iterated BLAST*), který kombinuje metody párového a vícenásobného přiložení. V první fázi hledá nejlépe párové příbuzenství. V rámci nejlepších výsledků tohoto kroku provede vícenásobné přiložení a vytvoří profil, o kterém jsme mluvili při diskusi sekundárních databází v sekci [13.4.3] na základě profilu znovu prohledává databáze a snaží se najít vzdáleně příbuzné sekvence, které v minulém kroku přehlédli. Takovýmto opakovaným prohledáváním umí *PSI-BLAST* odhalit i velmi slabou podobnost. *PSI-BLAST* je proto jednou z nejcitlivějších metod detekce vzdáleně příbuzných. Tato výhoda ale může být i nevýhodou. Do profilu se nám mohou vložit sekvence tvůrčí krátkými stále se opakujícími úseky (takzvané oblasti s nízkou komplexitou). Příkladem může být kolagen, v němž se pravidelně opakuje trojice aminokyselin Gly-Pro-Pro. V takovém případě prohledávání selže a snaží se nám vymít zpátky podobnost ze zcela odlišných proteinů.
**Obrázek 13.4:** Příklad vícenásobného přiložení (výstup programu **CLUSTAL W**).
13.10 Předpovídání struktury ze sekvence


13.10.1 Předpověď sekundární struktury

Pokusme se nejprve předpovědět sekundární strukturu na základě sekvence (primární struktury). To, co o sekundárních strukturách víme z části 2.5, zchladí nás optimismus hned na začátku. Jíž podle definice jsou sekundární struktury pravidelná prostorová uspořádání, která by neměla záviset na postranních řetězcích aminokyselin (a tedy ani na sekvenci). Ve skutečnosti tato nezávislost není úplná. Statistická analýza ukazuje, že určitá sekundární struktura jsou pro určité aminokyseliny o něco výhodnější a pro jiné aminokyseliny naopak méně výhodné. Byly vyvinuty algoritmy, které předpovídají, s jakou pravděpodobností bude aminokyselina v určitém místě sekvence součástí α-šroubovice, β-listu, pravidelného ohybu, nebo nepravidelné smyčky. Tyto algoritmy mohou být založeny na statistickém vyhodnocení, analýze tvaru postranních řetězců, nebo na analýze příbuznosti (homologie). Všechny metody ale vycházejí z pravidel, která byla formulována na základě sekvencí těch proteinů, jejichž prostorová struktura byla určena experimentálně. Moderní metody jsou založeny na vícenásobném přeložení takových sekvencí a na hledání shody (konsensu) několika různých algoritmů. Úspěšnost metod předpovídání sekundárních strutur je dnes lepší než 70%. Příklad předpovědi sekundární struktury je ukázán na obrázku 13.5.

13.11 Předpověď foldu


13.11.1 Homologní modelovaní

Příbuznost sekvencí (homologii) můžeme využít k předpovědi přesné prostorové struktury proteinu, pokud byla pro příbuznou sekvenci určena struktura experimentálně. Postup si můžeme rozdělit do čtyř kroků:
Obrázek 13.5: Příklad předpovědi sekundární struktury (výstup programu *PSIPRED*).
1. Provedeme příložení naší sekvence se sekvencemi homologních proteinů, pro něž je známá prosto-rová struktura. Takové proteiny budou pro naše modelování sloužit jako výchozí formy (TEMPLÁTY).

2. Ze struktury templátu vezmeme polohy uhlíků peptidové pátě a postranních řetězců. Pravidelné strukturní motivy použijeme jako základ modelu naší molekuly.

3. Snášíme se namodelovat méně pravidelné části struktury – ohyby a smyčky.

4. Navržený model dále upřesňujeme a ověřujeme jeho správnost.


Na závěr povídání o bioinformatice a využití databáze jsme se opět dostali k jádru našeho zájmu – k prostorové struktuře biomakromolekul. Vidíme, že analýza údajů o již známých biomakromolekulách nám může říci mnoho o struktuře biomakromolekuly dosud neznámé. Z tohoto pohledu můžeme i bioinformatiku považovat za metodu určování struktury biologicky zajímavých molekul.
Část III

Dodatky
A. ENERGIE RŮZNÝCH INTERAKCÍ

V jednoduchých případech lze elektrostatické působení mezi částmi molekul popsat nepříliš složitými rovnicemi. Uvádíme zde několik důležitých příkladů.

Energie 

E

interakce mezi náboji 

Q

1 a 

Q

2, které se nacházejí ve vakuu ve vzájemné vzdálenosti 

r

je dáná

E = \frac{1}{4\pi \varepsilon_0} \frac{Q_1 Q_2}{r},

(13.1)

kde \varepsilon_0 je elektrická permitivita vakuua (8,854,10^{−12} \text{ F m}^{−1}).

Energie 

E

interakce mezi nábojem 

Q

1 a trvalým elektrickým dipolovým momentem o velikosti 

p

2 = q_2 l_2, který odpovídá dvojici nábojů +q_2 a −q_2 ve vzájemné vzdálenosti 

l

je dáná

E = \frac{1}{4\pi \varepsilon_0} \frac{Q_1 p_2}{r^2} \cos \theta,

(13.2)

kde \theta a r jsou úhel a vzdálenost definované na obrázku [13.6A].

Energie 

E

interakce mezi trvalými elektrickými dipolovými momenty o velikostech 

p

2 = q_1 l_1 a 

p

2 = q_2 l_2 je dána
\[ E = \frac{1}{4\pi\varepsilon_0} \frac{p_1 p_2}{r^3} (\sin \theta_1 \sin \theta_2 \cos (\phi_1 - \phi_2) - 2 \cos \theta_1 \cos \theta_2), \quad (13.3) \]

kde úhly \( \theta_1, \theta_2, \phi_1, \phi_2 \) a vzdálenost \( r \) jsou definovány na obrázku [13.6].

Energie \( E \) interakce mezi nábojem \( Q_1 \) a okamžitém elektrickým dipólem indukovaným ve skupině atomů o elektrické polarizovatelnosti \( \alpha_2 \) je dána

\[ E = -\frac{1}{32\pi^2 \varepsilon_0} \frac{Q_1^2 \alpha_2}{r^4}, \quad (13.4) \]

kde \( r \) je vzdálenost mezi nábojem \( Q_1 \) a skupinou atomů.

Energie \( E \) interakce mezi trvalým elektrickým dipólovým momentem o velikosti \( p_1 = q_1 l_1 \) a okamžitým elektrickým dipólem indukovaným ve skupině atomů o elektrické polarizovatelnosti \( \alpha_2 \) je dána

\[ E = -\frac{1}{32\pi^2 \varepsilon_0} \frac{p_1^2 \alpha_2}{r^6} (3 \cos^2 \theta + 1), \quad (13.5) \]

kde \( \theta \) a \( r \) jsou úhel a vzdálenost definované na obrázku [13.6].

Energie \( E \) interakce mezi okamžitými dipólovými momenty indukovanými mezi skupinami atomů o elektrických polarizovatelnostech \( \alpha_1 \) a \( \alpha_2 \), které jsou považovány za harmonické oscilátory kmitající s frekvencemi \( f_1 \) a \( f_2 \), je dána

\[ E = -\frac{3}{4} \frac{h f_1 f_2}{f_1 + f_2} \frac{\alpha_1 \alpha_2}{r^6}, \quad (13.6) \]

kde \( h \) je Planckova konstanta (6,626.10\(^{-34}\) J s).
B. PEPTIDOVÁ VAZBA JAKO PŘÍKLAD CHROMOFORU

Pokusme se podívat podrobněji na to, jak se interakce se světlem projeví na malé části molekuly proteinu, na peptidové vazbě. Začneme tím, že si popíšeme jednotlivé stavy elektronů této skupiny. Víme, že tyto stavy popisujeme pomocí vlnových funkcí. Abychom takovému popisu porozuměli, zkusíme si (dostí jenom odvozit vlnové funkce peptidové vazby.

B.1 Atomové orbitaly

Ze střední školy jistě dobře znáte schematické znázornění atomových orbitalů v prostoru (Obrázek 13.7).

Tyto obrázky jsou vlastně prostorovou obdobou izobar na meteorologických mapách. Vyznačují místa v prostoru, kde má vlnová funkce stejnou hodnotu. Obvykle se pro každý orbital znázorňují jen dvě „izobary“, pro jednu kladnou (zde znázorněna modře) a jednu zápornou (zde znázorněna červeně) hodnotu vlnové funkce. V peptidové vazbě máme jeden vodík, pro který nám bude stačit nejnižší orbital 1s. Dále máme v peptidové vazbě po jednom atomu uhlíku, kyslíku a dusíku. Když budeme předpokládat, že s dvojicí elektronů v nejnižším stavu 1s se nic nestane, stačí nám pro každý z těchto atomů čtyři orbitaly. To je dohromady třináct orbitalů.

B.2 Hybridní orbitaly

Z popsaných 15 atomových orbitalů chceme sestavit molekulové orbitaly, kterých musí být také patnáct. Nové orbitaly můžeme spočítat jako lineární kombinaci atomových orbitalů. To znamená, že každý atomový orbital vynásobíme určitým koeficientem a tyto příspěvky sečteme. Je výhodné tento výpočet dělat nadvakrát, nejprve sestrojit takzvané hybridní orbitaly, ze kterých se nám výsledné molekulové orbitaly už budou dobře počítat.

Začneme s karbonylovým uhlikem. Budeme předpokládat, že peptidová vazba leží v rovině xy. V této rovině tvoří uhlik tři σ-vazby s okolními atomy. Pro jejich popis si připravíme tři hybridní orbitaly, které

\[
1s = \sqrt[4]{\pi} \left( x^2 + y^2 + z^2 \right)^{3/2} e^{-Zr}, \quad 2s = \sqrt[4]{2\pi} \left( x^2 + y^2 + z^2 \right)^{3/2} e^{-Zr}, \quad 2p_x = \frac{1}{\sqrt{2\pi}} \left( x^2 + y^2 + z^2 \right)^{5/2} e^{-Zr},
\]

\[
2p_y = \frac{1}{4\sqrt{2\pi}} \left( x^2 + y^2 + z^2 \right)^{5/2} e^{-Zr}, \quad 2p_z = \frac{1}{4\sqrt{2\pi}} \left( x^2 + y^2 + z^2 \right)^{5/2} e^{-Zr},
\]

kde \( r^2 = x^2 + y^2 + z^2 \), Z je počet protonů v jádře a \( a \) je poloměr dráhy elektronu nejnižšího jádra v klasickém Bohrově popisu.
budeme označovat sp\(^2\). Vypočítáme je ze tří atomových orbitalů 2s, 2p\(_x\), 2p\(_y\):

\[
\begin{align*}
sp_1^2 &= \sqrt{\frac{1}{3}} 2s + \sqrt{\frac{2}{3}} 2p_x \\
sp_2^2 &= \sqrt{\frac{1}{3}} 2s - \sqrt{\frac{1}{6}} 2p_x + \sqrt{\frac{1}{2}} 2p_y \\
sp_3^2 &= \sqrt{\frac{1}{3}} 2s - \sqrt{\frac{1}{6}} 2p_x - \sqrt{\frac{1}{2}} 2p_y
\end{align*}
\] (13.7, 13.8, 13.9)

Graficky si výpočet orbitalů sp\(^2\) můžeme znázornit tak, jak ukazuje obrázek 13.8.

V horní části obrázku jsou znázorněny orbitalu při pohledu shora na rovinu xy, v dolní části při stejném úhlu pohledu jako na předchozím obrázku. Podobným způsobem můžeme vytvořit i tři sp\(^2\) orbitály dusíku. Trochu jiné orbitaly vypočítejme pro kyslík. Spolu s orbitálem 2s vezmeme pouze jeden orbital 2p (2p\(_x\)), ze kterých sestrojíme dva orbitály: sp\(^2\) a s\(^2\)p.

6 Pokud vám není jasné, proč jsme za koeficienty zvolili právě ty podivné odmocniny, vzpomeňte, jak jsme definovali pravděpodobnost elektronu v určitém místě v prostoru: \(\psi\psi^*\) (pro naše orbitály zvolené jako reálné funkce můžeme psát jednoduše \(\psi^2\)). Pokud bude elektron ve stavu popsaném orbitálem \(\psi\), musí platit, že součet pravděpodobností výskytu ve všech místech prostoru dá jedničku (někde elektron být musí). Matematicky řečeno, \(\int_V \psi^2 dV = 1\) (V značí, že integrujeme přes celý objem). Zkusme si to třeba pro orbital sp\(^2\)

\[
1 = \int_V (sp_1^2)^2 dV = \int_V \left(\sqrt{\frac{1}{3}} 2s + \sqrt{\frac{2}{3}} 2p_x\right)^2 dV = \frac{1}{3} \int_V (2s)^2 dV + \frac{2}{3} \int_V (2p_x)^2 dV + \frac{2\sqrt{2}}{3} \int_V 2s2p_x dV = \frac{1}{3} + \frac{2}{3} + 0
\] (13.10)

Protože poslední integrál je nulový (2s2p\(_x\) je lichá funkce, to je dáno symetrií orbitalů), můžeme říci, že pravděpodobnost výskytu v sp\(^2\) je díana z jedné třetiny pravděpodobností výskytu v 2s a ze dvou třetin pravděpodobností výskytu v 2p\(_x\).
B. PEPTIDOVÁ VAZBA JAKO PŘÍKLAD CHROMOFORU

Obrázek 13.9: Orbitaly s²p

\[ sp^2 = \sqrt{\frac{1}{3}}2s + \sqrt{\frac{2}{3}}2p_x \]

\[ s^2p = \sqrt{\frac{2}{3}}2s - \sqrt{\frac{1}{3}}2p_x \]

Grafické znázornění je obdobné předchozímu (obrázek 13.9).

A nakonec budeme potřebovat jeden sp³ orbital od každého α-uhlíku, který vypočteme z orbitalu 2s a ze tří orbitalů 2p (takto získáme čtyři orbitaly sp³, zbývající tři bychom mohli použít na popis dalších tří σ-vazeb α-uhlíku)

\[ sp^3_1 = \frac{1}{2}2s + \frac{1}{2}2p_x + \frac{1}{2}2p_y + \frac{1}{2}2p_z \]

\[ sp^3_2 = \frac{1}{2}2s + \frac{1}{2}2p_x - \frac{1}{2}2p_y - \frac{1}{2}2p_z \]

\[ sp^3_3 = \frac{1}{2}2s - \frac{1}{2}2p_x - \frac{1}{2}2p_y + \frac{1}{2}2p_z \]
\[ sp^3 = \frac{1}{2}2s - \frac{1}{2}2p_x + \frac{1}{2}2p_y - \frac{1}{2}2p_z \]  

(13.17)

**B.3 Molekulové orbitaly**

Konečně jsme připraveni vypočítat molekulové orbitaly peptidové vazby. Podívejme se nejprve na orbitaly \( \sigma \). V peptidové skupině máme 5 \( \sigma \)-vazeb, budeme tedy počítat 10 \( \sigma \) orbitalů (5 vazebných a 5 antivazebných). K výpočtu použijeme jeden 1s orbital vodíku, sedm sp\(^2\) orbitalů (po třech z uhlíku a dusíku a jeden kyslíkový) a dva sp\(^3\) orbitaly sousedních \( \alpha \)-uhlíků.

\[
\sigma(C_\alpha^1 C) = \sqrt{\frac{1}{2}}sp^3(C_\alpha^1) + \sqrt{\frac{1}{2}}sp^3(C) \\
\sigma^*(C_\alpha^1 C) = \sqrt{\frac{1}{2}}sp^3(C_\alpha^1) - \sqrt{\frac{1}{2}}sp^3(C) \\
\sigma(CO) = \sqrt{\frac{1}{2}}sp^2(C) + \sqrt{\frac{1}{2}}sp^2(O) \\
\sigma^*(CO) = \sqrt{\frac{1}{2}}sp^2(C) - \sqrt{\frac{1}{2}}sp^2(O) \\
\sigma(CN) = \sqrt{\frac{1}{2}}sp^2(C) + \sqrt{\frac{1}{2}}sp^2(N) \\
\sigma^*(CN) = \sqrt{\frac{1}{2}}sp^2(C) - \sqrt{\frac{1}{2}}sp^2(N) \\
\sigma(NH) = \sqrt{\frac{1}{2}}sp^1(N) + \sqrt{\frac{1}{2}}1s(H) \\
\sigma^*(NH) = \sqrt{\frac{1}{2}}sp^1(N) - \sqrt{\frac{1}{2}}1s(H) \\
\sigma(NC_\alpha^2) = \sqrt{\frac{1}{2}}sp^3(N) + \sqrt{\frac{1}{2}}sp^3(C_\alpha^2) \\
\sigma^*(NC_\alpha^2) = \sqrt{\frac{1}{2}}sp^3(N) - \sqrt{\frac{1}{2}}sp^3(C_\alpha^2) 
\]

(13.18) - (13.27)

Graficky je tvorba orbitalů \( \sigma \) a \( \sigma^* \) ze dvou sp\(^2\) orbitalů znázorněna na obrázku [13.10].

Všimněte si, že při tvorbě hybridních atomových orbitalů jsme kombinovali orbitaly, které vycházely z jednoho atomu, tedy z jednoho bodu v prostoru. Teď, když tvoríme molekulové orbitaly, kombinujeme atomové orbitaly různých atomů (na obrázku označených X a Y), tedy vycházející ze dvou bodů v prostoru.

Když už máme popsané vazby \( \sigma \), podívejme se teď na vazby \( \pi \). Mezi původními patnácti atomovými orbitaly byly tři orbitaly 2p\(_z\), které jsme nepoužili k hybridizaci. Tyto orbitaly leží nad a pod rovinou \( xy \) a můžeme je použít k tvorbě tří orbitalů \( \pi \) (obrázek [13.11]).

\[
\pi_1 = \sqrt{\frac{1}{3}}2p_z(O) + \sqrt{\frac{1}{3}}2p_z(C) + \sqrt{\frac{1}{3}}2p_z(N) \\
\pi_2 = \sqrt{\frac{1}{3}}2p_z(O) + \sqrt{\frac{1}{3}}2p_z(N) \\
\pi_3^* = \sqrt{\frac{1}{3}}2p_z(O) - \sqrt{\frac{1}{3}}2p_z(C) + \sqrt{\frac{1}{3}}2p_z(N) 
\]

(13.28) - (13.30)
Obrázek 13.10: Tvorba orbitalů $\sigma$.

Obrázek 13.11: Tvorba orbitalů $\pi$. 
Po vytvoření $\sigma$ a $\pi$ orbitálů nám zbyly ještě dva nepoužité orbitály kyslíku ($2p_y$ a $s^2p$). Tyto orbitály tvoří nevazebné molekulové orbitály $n$ a $n'$.

\begin{align*}
    n &= 2p_y(O) \\
    n' &= s^2p(O)
\end{align*}

**B.4 Energie stavů a vlnové délky přechodů**

Víme, že rezonanční frekvence přechodů elektronů ze základního stavu do stavu vyššího je dána rozdílem energií těchto stavů. Podívejme se tedy na energie stavů popsány jednotlivými orbitály. Schematicky je energetické rozložení orbitálů ukázáno na obrázku [13.12].

Zleva doprava jsou nakresleny orbitály tak, jak jsme je postupně počítali. Rámeček vymezuje to, co nás zajímá (tedy skupinu CONH s vazbami na sousední uhlíky). Různé typy molekulových orbitálů jsou znázorněny barevně, pomocí barev je naznačeno, kombinací kterých atomových orbitálů jsme jednotlivé typy molekulových orbitálů získali. Dále je na obrázku zachyceno obsazení orbitálů elektronů v základním stavu. Konečně jsou na obrázku znázorněny dva přechody do vyšších stavů. Jde o přechody, které ve spektrech obvykle vidíme v ultrafialové oblasti. Světlo o vlnové délce kolem 190 nm je pohlceno při přechodu z nevazebného $\pi_2$ orbitalu do antivazebného $\pi^*_1$ orbitalu. Kolem 220 nm pozorujeme absorpci při přechodu z nevazebného n orbitalu do antivazebného $\pi^*_3$ orbitalu.
B. PEPTIDOVÁ VAZBA JAKO PRÍKLAD CHROMOFORU

B.5 Tranzitní dipóly a pravděpodobnosti přechodů

Pokud chceme znát i pravděpodobnost přechodu (tedy absorpční koeficient), musíme umět spočítat také tranzitní elektrické dipólové momenty. Co to vlastně je elektrický dipólový moment? Je-li záporný náboj elektronu rozdělen symetricky kolem kladného náboje jádra (jak by napovídaly tvary orbitálů), je dipólový moment nulový. Pokud ale dojde k přemístění elektronů takovým způsobem, že výsledný záporný náboj elektronů bude vzdálen od jádra o vzdálenost r ve směru \( \vec{r} \), můžeme takovou polarizaci popsat elektrickým dipólovým momentem

\[
\vec{\mu}_e = q\vec{r}
\]

(13.33)

kde \( q \) je výsledný náboj. Protože kvantová teorie uměje popsat pravděpodobnost výskytu elektronu (a tak i rozložení náboje) pomocí vhovné funkce, pomůže nám i nalézt dipólový moment:

\[
\vec{\mu}_e = q \int_V \psi \psi^* dV
\]

(13.34)

místo náboje jsme vlastně do rovnice [13.33] dosadili vyjádření pravděpodobnosti výskytu elektronu. Tak, jak jsme si v rovnici [13.34] vypočítali dipólový moment molekuly s elektrony ve stavu \( \psi \), můžeme si spočítat i tranzitní elektronový dipólový moment přechodu ze stavu \( \psi_1 \) do stavu \( \psi_2 \):

\[
\vec{\mu}_{e,tt} = q \int_V \psi_2 \psi_1^* dV
\]

(13.35)

Integrály v rovnicích [13.34] a [13.35] znamenají, že musíme sečíst všechny lokální příspěvky \( d\vec{\mu}_{e,tt} \) k celkovému tranzitnímu momentu \( \vec{\mu}_{e,tt} \). Vezmeme si například přechod z orbitálu \( 2s \) do orbitálu \( 2p_\pi \). Zkusme nejprve sčítat příspěvky k \( \vec{\mu}_{e,tt} \) podél osy \( z \). V horní polovině (tedy pro kladné \( z \)) mají obě vhovné funkce kladné znaménko, takže součin \( \psi_1 \psi_2 \) bude kladný a příspěvek \( \psi_1 \psi_2 \) bude také kladný.

Naopak v dolní polovině (pro záporné \( z \)) mají vhovné funkce opačné znaménko, takže součin \( \psi_1 \psi_2 \) bude záporný a příspěvek \( \psi_1 \psi_2 \) opět kladný (z je teď záporné), Tedy příspěvky k tranzitnímu momentu \( \vec{\mu}_{e,tt} \) budou vždy podél osy \( z \). Co když budeme podobně sčítat příspěvky podél osy \( x \)? V horní části obrázku jsou obě vhovné funkce kladné pro kladné i záporné \( x \). Proto se nám příspěvky v místě s kladným \( x \) vyrůstí s příspěvky v místě se záporným \( x \) (obě vhovné funkce jsou symetrické, sudé vzhledem k \( x \), takže \( \psi(x) = \psi(-x) \)). Stejně tak to dopadne i v dolní polovině obrázku, kde je zase součet \( \psi_1 \psi_2 \) záporný pro kladné i záporné \( x \). Totéž bychom zjistili i podél osy \( y \). To znamená, že celkový příspěvek ve směru \( x \) a \( y \) bude nulový, vektor \( \vec{\mu}_{e,tt} \) bude směřovat ve směru osy \( z \).

Podobnou analýzu znamená orbitála \( \pi_2 \) a \( \pi_3^* \) bychom zjistili, že tranzitní elektrický dipólový moment přechodu \( \pi_2 \rightarrow \pi_3^* \) je kolmý k ose \( z \) a směřuje k dusíku ke kyslíku. Naopak analýza orbitálů \( n \) a \( \pi_3^* \) ukazuje, že všechny složky tranzitního elektrického dipólového momentu jsou nulové. Přechod \( \pi_2 \rightarrow \pi_3^* \) je tedy elektricky povolený, zatímco přechod \( n \rightarrow \pi_3^* \) je elektricky zakázaný.

Tranzitní elektrický moment popisuje přesně pravděpodobnost výskytu elektronu (a tedy i přesun náboje) určitým směrem. Jde tedy o popis jakéhokoli "translační" změny během přechodu. Kromě toho dochází při přechodu i k rotačním změnám pravděpodobnosti výskytu elektronu a tedy ke kruhovému toku náboje. Jak víte z fyziky, kruhový tok náboje vyvolá vznik magnetického pole. Kromě tranzitního elektrického dipólového momentu je tedy přechod mezi stavý popsaný i tranzitním magnetickým dipólovým momentem. Analýza znamená a tvaru orbitálů by nám například ukázala, že přechod \( n \rightarrow \pi_3^* \) má nenulový tranzitní magnetický dipólový moment, je tedy magneticky povolený.

Vzájemné interakce elektrických a dipólových momentů způsobují rozdílnost v absorpcii levotwicí a pravotwicí polarizovaného světla (cirkulární dichroismus). Zatímco zrcadlově symetrické molekuly mají elektrické a magnetické momenty navzájem kolmé a tedy neinteragující, v chirálních molekulách k interakcí dochází. Peptidová vazba, která nás především zajímá, je ovšem roviný útvar a sama o sobě je tedy zrcadlově symetrická. V proteinech však dochází k interakci mezi tranzitními momenty různých
važe. Rozdíly mezi absorpcí kolem 220 nm souvisí především s interakcí elektrického tranzitního momentu jedné vazby s magnetickým tranzitním momentem jiné vazby. Naopak rozdíly absorpcí kolem 195 nm jsou dány vzájemným působením elektrických dipólů různých vazeb. To, že církuální dichroismus závisí na interakcích dipólů jednotlivých vazeb vysvětluje, proč je tato technika tak užitečná ke zkoumání sekundárních struktur, které se liší právě vzájemnou orientací peptidových vazeb.
C. Metody chemické syntézy na pevném nosiči

C.1 Karbodiimidová metoda syntézy peptidů

1. Zablokujeme všechny α-aminoskupiny a všechny reaktivní skupiny postranních řetězců. Jako blokující skupina α-aminoskupiny se používá 9-fluorenylmethyloxykarbonyl (Fmoc) nebo terti-butyloxykarbonyl (tBoc).

2. C-koncovou aminokyselinu karboxyskupinou naváže me na pevný nosič. Výhoda použití pevného nosiče je v tom, že rostoucí peptid zůstává připojený k nosiči, takže po každém kroku můžeme snadno vyměnit reakční roztok.

3. Odblokujeme α-aminoskupinu C-koncovou aminokyseliny vázané na nosič.


5. Nakonec odblokujeme všechny funkční skupiny, uvolněme peptid z nosiče a vyčistíme jej.

C.2 Fosforamiditová metoda syntézy nukleotidů

1. Zablokujeme všechny aminoskupiny a hydroxyly. Po celou dobu syntézy je třeba chránit aminoskupiny bází a v případě RNA 2′-hydroxyl. Jako blokující skupina 5′-hydroxylu se používá dimethoxytrityl.

2. 3′-koncový nukleotid naváže me přes 3′-hydroxyl na pevný nosič.

3. Kyselinou trichloroctovou odblokujeme 5′-hydroxyl koncového nukleotidu vázaného na nosič.


5. Jodem zoxidujeme triester kyseliny fosforité na triester kyseliny fosforečné.

6. Předchozí tři kroky opakujeme tolikrát s příslušným aktivovaným nukleotidem, kolik nukleotidů má obsahovat vzniklý produkt.

7. Nakonec odblokujeme 5′-hydroxyl prvního nukleotidu, amoniakem odštěpíme poslední nukleotid z nosiče a kyanoethyl z fosfátů, zahrátím uvolníme ostatní chráněné funkční skupiny a oligonukleotid přečistíme.

---

7 Použití různých blokujících skupin umožňuje selektivní odblokování jedné funkční skupiny. Například v amoniaku dochází k odštěpení Fmoc, ale ne tBoc, který se naopak odštěpuje v kyselině trifluoroctové.
8 nebo N-hydroxysukcinimidem.
D Anomální rozptyl

V kapitole 11.4.5 jsme si představili anomální rozptyl jako jev, který nám může pomoci vyřešit fázový problém. Následující text nabízí konkrétnější popis.


Pokud se ale frekvence záření blíží rezonanční frekvenci, pohlcený foton vybuduje přechod elektronu do jiného stavu a vyzářený foton bude mít jinou amplitudu i fázi\(^9\). Tento rozdíl je znázorněn krátkou čárovou šipkou na obrázku 11.11B a nazývá se anomální rozptyl. Všimněte si, že tato šipka pootočí vektory \(\vec{F}_D(hkl)\) i \(\vec{F}_D(hkl)\) ve stejném směru (změní fázi obou strukturálních faktorů o stejnou hodnotu). Po přetížení příspěvku těžkého atomu (včetně anomálního rozptylu) k \(\vec{F}_P(hk l)\) a \(\vec{F}_P(-h-k-l)\) získáme výsledné strukturální faktory \(\vec{F}_{PD}(hk l)\) a \(\vec{F}_{PD}(-h-k-l)\), které mají různou amplitudu – plná a čárována kružnice na obrázku 13.13A mají různý poloměr. Strukturní faktory Friedlova páru nejsou v přítomnosti anomálního rozptylu komplexně sdružené \(F(hk l) \neq \bar{F}^{*}(-h-k-l)\). Tento rozdíl vynikne lépe, když si místo černé čárovány šipky \(\vec{F}_{PD}(-h-k-l)\) nakreslíme její zrcadlový obraz (což není níc jiného, než komplexně sdružený vektor \(\vec{F}_{PD}^{*}(-h-k-l)\)). Obrázek 13.13C ukazuje, že vektor \(\vec{F}_{PD}^{*}(-h-k-l)\) se nepřekryvá s plnou šipkou \(\vec{F}_{PD}(hk l)\) a má i jinou amplitudu (různé poloměry kružnice). Rozdílná amplituda se projeví v difrakčním obrazci jako různá intenzita reflexů s indexy \((hk l)\) a \((-h-k-l)\).

Pokud jsme schopni rozlišit intenzity reflexů s opačnými hodnotami indexů \(h, k, l\), můžeme anomální rozptyl využít pro řešení fázového problému podobně jako v případě isomorfního nahrazení. Graficky je takový postup ukázán na obrázku 13.14.

\(^9\)Pokud budou všechny rozptylovat stejně, nebude ovlivněna amplituda, ale jen fáze
Obrázek 13.13: Friedelův pár a anomální rozptyl. Obrázek A zachycuje situaci v nepřítomnosti těžkého atomu, obrázky B a C v případě, že je přítomen těžký atom vykazující anomální rozptyl. Vlevo je znázorněna poloha lehkých atomů (nakreslených červeně, zeleně a modře) a jednoho těžkého atomu (nakresleného azurovou barvou) mezi Braggovými rovinami, vpravo Argandův diagram ukazující sčítání strukturních faktorů jednotlivých atomů (strukturní faktory a atomy jsou nakresleny stejnými barvami). Čárkované jsou nakresleny strukturní faktory pro odraz od spodní strany Braggových rovin. Černé šipky představují výsledné strukturní faktory pro molekulu bez těžkého atomu ($\vec{F}_P$) a s těžkým atomem ($\vec{F}_{PD}$). Obrázek C zachycuje srovnání strukturního faktoru $\vec{F}_{PD}(hkl)$ s hodnotou komplexně sdruženou se strukturálním faktorem $\vec{F}_{PD}(-h-k-l)$ (vektor $\vec{F}_{PD}(-h-k-l)$). Tučné azurové šipky znázorňují fázový rozdíl v důsledku anomálního rozptylu.
E  Přiřazení spektre NMR nukleových kyselin s využitím skalárních interakcí

Jak bylo zmíněno výše, frekvenční přiřazení proteinů je založeno na jednovazebných skalárních interakcích. Výhodou tohoto přístupu je nezávislost na konkrétní konformaci. Frekvenční přiřazení nukleových kyselin pomocí skalárních interakcí nařizí na to, že jednovazebné interakce často nejsou k dispozici a je nutno využívat dvou- i třívazebných interakcí. V důsledku toho je tento přístup u nukleových kyselin komplikován nízkou citlivostí a závislostí na konformaci u některých experimentů (jak dobře víme, třívazebné interakce závisí na torzních úhlech).

Přiřazení páteře (sekvenční přiřazení)

Fosfátorové skupiny páteře nukleových kyselin obsahují kyslík, který nemá vhodné izotopy pro NMR. Sekvenční přiřazení je proto založeno především na třívazebných interakčních konstantách $^3J(PC)$ a $^3J(PH)$. První interakce využívá například HCP experiment, který poskytuje korelace $H'(i)-C'(i)-P(i)$ a $H''(i-1)-C''(i-1)-P(i)$ (obdoba HNCA a HN(CO)CA u proteínů). Druhou interakci využívá HPHCH experiment, který poskytuje korelace $H''(i)-C''(i)-P(i)$, $H''(i)-C''(i)-P(i)$ a $H'(i-1)-C'(i-1)-P(i)$. Třetí možností je PCCH-TOCSY experiment (varianta HCCH-TOCSY experimentu popsaného pro proteiny), který využívá $^3J(PC)$ interakce ke korelací fosforu s pentosovým kruhem, následuje homonukleární TOCSY korelace cukerných uhlíků a heteronukleární korelace s navázaným protonem. Tak je možné korelovat fosfor s různými atomy (deoxy)ribosy.

Přiřazení v cukrech

K přiřazení protonů u uhlíků v (deoxy)ribose se používá HCCH-TOCSY experiment, který byl popsán pro proteiny.

Korelace báze-cukr

Ke korelací H6 (pyrimidinů) a H8 (purinů) s H1′ cukru se používá dvojice HCN experimentů, z nichž jeden excituje uhlíky alifatické a druhý aromatické. Získáme tak korelace H6-C6-N1(pyrimidinů), H8-C8-N9 (purinů) a H1′-C1′-N1 (pyrimidinů) a H1′-C1′-N9 (purinů). Opět jde o obdobu HNCA a HN(CO)CA u proteínů. Zmíněných korelací lze dosáhnout i použitím jednoho experimentu (HCNCH), jehož citlivost je ale nižší.

Přiřazení v bázích

Základním úkolem přiřazení v bázích nukleových kyselin je přiřazení jednotlivých protonů. Ke korelací protonů v polohách 5 a 6 pyrimidinů lze využít jednovazebné interakce $^1J(CH)$ a $^1J(CC)$. Korelace vzdálenějších protonů využívá často vícevazebných skalárních interakcí, které bývají v aromatických heterocyklických spinových systémech bází poměrně intenzivní (napak některé jednovazebné skalární interakce jsou velmi slabé). Často se používá princip TOCSY. Jako příklad můžeme uvést opět HCCH-TOCSY experiment, používaný ke korelací H2 a H8 u adeninu, nebo HNC-TOCSY-CH experiment (na rozdíl od předchozího koreluje v prvním kroku imino-protony s uhlíky přes imino-dusík, tedy obdobně jako v HNCA experimentu), používaný ke korelací H1 a H8 u guaninu.
Obrázek 13.14: Využití anomálního rozptylu k řešení fázového problému. Na obrázku A jsou zakresleny strukturální faktory. \( \vec{F}_P \) značí strukturální faktor nativní biomakromolekuly (stejné pro \((hkl)\) a \((-h-k-l)\)), který není ovlivněn anomálním rozptylem. Vektory \( \vec{F}_D \) (různé pro \((hkl)\) a \((-h-k-l)\)) jsou strukturální faktory těžkých atomů, které vykazují anomální rozptyl. Vektory \( \vec{F}_{PD} \) jsou výsledné strukturální faktory derivátu s těžkými atomy. Ve skutečnosti známe pouze vektory \( \vec{F}_D(h,k,l) \) a \( \vec{F}_D(-h,-k,-l) \) a amplitudy \( |F_P|, |F_{PD}(h,k,l)| \) a \( |F_{PD}(-h-k-l)| \). Obrázky B až C ukazují princip řešení. Nejprve si nakreslíme plnou kružnici o poloměru \( |F_P| \), která zachycuje veškerou informaci o strukturálním faktoru nativní molekuly (obrázek B). Potom zaneseme do diagramu údaje o derivátech. Vektory \( \vec{F}_{PD}(h,k,l) \) a \( \vec{F}_{PD}(-h-k-l) \) můžeme zakreslit přímo. Jak jsme si vysvětlili v popisu obrázku 11.10, musí střed plné kružnice, znázorňující amplitudu \( |F_P| \), ležet v konci vektoru \( \vec{F}_P \). To platí pro \((h,k,l)\) i \((-h-k-l)\), proto zakreslíme vektory \( \vec{F}_D(h,k,l) \) a \( \vec{F}_D(-h-k-l) \) tak, aby směřovaly do středu plné kružnice, jak je ukázáno na obrázku C. Nakonec zakreslíme tečkanou a čárkovou kružnicí odpovídající strukturálním faktorům \( \vec{F}_{PD} \), které jsou ovlivněny anomálním rozptylem a liší se znaménkem indexů \( h, k, l \). Tyto kružnice mají proto různé poloměry. Jejich posunutí vzhledem k plné kružnici, odpovídající strukturálnímu faktoru nativní biomakromolekuly neovlivněné anomálním rozptylem, je dáno směry \( \vec{F}_D \) (středy kružnic leží v počátcích \( \vec{F}_D(h,k,l) \) a \( \vec{F}_D(-h-k-l) \)). Podobně jako při použití isomorfního nahrazení nám řešení udávají průsečíky kružnic. Tam, kde se všechny kružnice protínají, leží konce hledaných vektorů (obrázek D).