

Univerzita Karlova

Přírodovědecká fakulta

Studijní program: Biochemie

Studijní obor: Biochemie



Dominik Nelic

Přiřazení rezonančních frekvencí atomům proteinové páteře  
myšího proteinu Nkrp1c

Backbone Resonance Assignment of the mouse protein Nkrp1c

Bakalářská práce

Vedoucí práce: Mgr. Josef Chmelík, Ph.D.

Praha 2019

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením školitele Mgr. Josefa Chmelíka, Ph.D. a všechny použité prameny jsem řádně citoval.

V Praze dne:

Dominik Nelic

## **Abstrakt**

Protein Nkrp1c je receptor na povrchu myších přirozeně zabíječských buněk (natural killer cells) imunitního systému, který patří do rodiny C-lektin receptorů. Tyto buňky vytvářejí vrozenou imunitu proti nádorům či patogenům před vznikem protilátek.

Určení 3D struktury proteinů je často klíčem pro pochopení funkce proteinu na molekulární úrovni. Jednou z možností, jak určit strukturu proteinů na atomární úrovni, je nukleární magnetická rezonance. Náplní této bakalářské práce bylo vyhodnocení několika již naměřených spekter a přiřazení rezonančních frekvencí atomů peptidové páteře potřebných pro získání dat na predikci sekundární struktury pro protein Nkrp1c, který byl připravený rekombinantní expresí.

K vyhodnocení naměřených spekter byl využit program Sparky. Samotná predikce sekundární struktury proteinu Nkrp1c byla provedena programy Talos+ a PSIPRED. Získané výsledky byly porovnány s již publikovaným homologním modelem receptoru proteinu Nkrp1c.

**Klíčová slova:** NMR proteinu, Nkrp1c

## **Abstract**

The protein Nkrp1c is a receptor on the surface of murine natural killer cells belonging to the C-lectin receptor family. These cells create congenital immunity against tumors or pathogens before the formation of antibodies.

Determining the 3D protein structure is often the key to understanding the function of the protein at the molecular level. One way to determine the structure of proteins at the atomic level is nuclear magnetic resonance. The aim of this bachelor thesis was to evaluate several already measured spectra and to assign the resonance frequencies of the peptide backbone atoms needed to obtain data for the secondary structure prediction for Nkrp1c protein, which was prepared by recombinant expression.

The Sparky program was used to evaluate the measured spectra. The prediction of the secondary structure of the Nkrp1c protein itself was performed by programs Talos+ and PSIPRED. The obtained results were compared with the already published homologous model of the Nkrp1c protein receptor.

**Keywords:** protein NMR, Nkrp1c

# Obsah

Seznam zkratk.....	7
1 ÚVOD.....	8
1.1 Struktura proteinů.....	8
1.1.1 Ramachandranův diagram.....	11
1.2 Příprava izotopicky značeného proteinu.....	12
1.3 Určování 3D struktury proteinů.....	13
1.3.1 Rentgenová strukturní analýza.....	13
1.3.2 Kryo-elektronová mikroskopie.....	13
1.3.3 Nukleární magnetická rezonance.....	14
1.3.4 Obecný postup při řešení struktury pomocí NMR.....	15
1.4 Naměření NMR spekter – série spekter.....	16
1.4.1 Přiřazení páteře.....	16
1.5 Princip přiřazení aminokyselin.....	19
1.6 Nkrp1c.....	19
2 Cíle práce.....	20
3 Metody.....	21
3.1 Naměření NMR spekter.....	21
3.2 Vyhodnocení NMR spekter.....	21
3.3 Predikce sekundární struktury.....	23
3.3.1 Talos+.....	23

3.3.2 PSIPRED.....	24
4 Výsledky.....	25
4.1 HNCACB+CBCA(CO)NH.....	27
4.2 HNCO a HN(CA)CO.....	28
4.3 Predikce sekundární struktury.....	29
5 Diskuze.....	30
6 Závěr.....	31
7 Zdroje.....	32
8 Přílohy.....	35
8.1 Tabulka výsledků přiřazení signálů aminokyselin.....	35

## Seznam zkratek

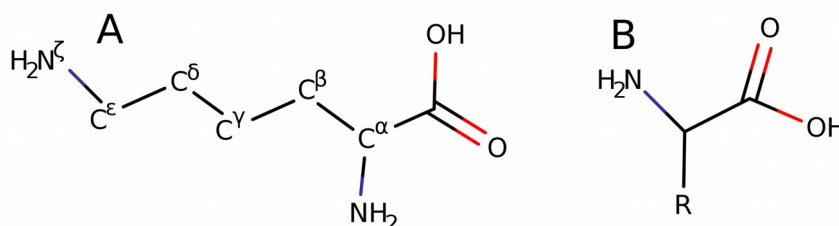
C'	karbonylový uhlík proteinové páteře
C <sup>α</sup>	α uhlík proteinové páteře
C <sup>β</sup>	β uhlík aminokyselinového zbytku
Cryo-EM	kryoelektronová mikroskopie
DNA	deoxyribonukleová kyselina (z angl. Deoxyribonucleic Acid)
HNCA	3D NMR spektrum pro přiřazení signálů C <sup>α</sup> uhlíkům proteinu
HNCACB	3D NMR spektrum pro přiřazení signálů C <sup>β</sup> a C <sup>α</sup> uhlíkům proteinu
HN(CA)CO	3D NMR spektrum pro přiřazení signálů C' uhlíkům proteinu
HNCO	3D NMR spektrum pro přiřazení signálů C' uhlíkům proteinu
HN(CO)CA	3D NMR spektrum pro přiřazení signálů C <sup>α</sup> uhlíkům proteinu
HN-HSQC	2D NMR spektrum pro přiřazení signálů dusíkům a vodíkům (z angl. Heteronuclear Single-Quantum Coherence)
MHC	hlavní histokompatibilní komplex (z angl. Major Histocompatibility Complex)
NK (buňky)	přírodně zabíječské buňky (z angl. Natural Killer cells)
NMR	nukleární magnetická rezonance
PDB	proteinová databanka (z angl. Protein Data Bank)
ppm	výraz pro jednu miliontinu (z angl. parts per milion) – měřítko, ve kterém se udávají hodnoty bezrozměrného chemického posunu
PSIPRED	program pro predikci sekundární struktury (z angl. PSI-blast based secondary structure PREDiction)
X-ray	rentgenostrukturní

# 1 ÚVOD

## 1.1 Struktura proteinů

Proteiny jsou jedny z nejběžnějších biologických makromolekul s velkou různorodostí od nejmenších proteinů až po obrovské biopolymery s molekulovou hmotností v řádech milionů. Proteiny mají řadu biologických funkcí. Ať už se jedná o bakterie nebo člověka, veškeré proteiny jsou tvořeny převážně 20 aminokyselinami kovalentně vázanými do lineární sekvence. Zajímavou vlastností buněk je, že mohou tvořit proteiny s různými vlastnostmi a aktivitou pouze spojením těchto aminokyselin do specifických sekvencí.

Všechny aminokyseliny sdílí jednu společnou vlastnost a to základní strukturu (obrázek 1B). Všechny obsahují karboxylovou skupinu  $\text{COO}^-$  a aminovou skupinu  $\text{NH}_3^+$  připojenou ke stejnému atomu uhlíku, který je označován jako  $\text{C}^\alpha$  uhlík a liší se změnou postranního řetězce (na obrázku 1B R). Při označování uhlíků v postranních řetězcích se dále používá řecká abeceda (obrázek 1A). Toto označení je použito i při přiřazování atomů v NMR (nukleární magnetická rezonance) spektrech.

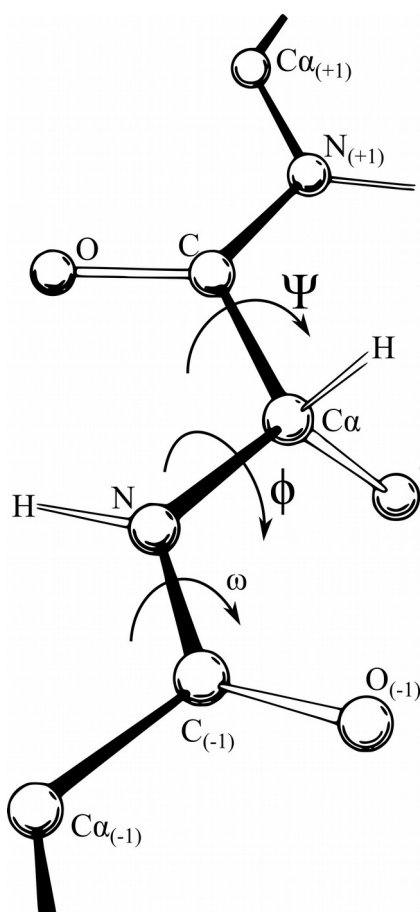


Obrázek 1: A) označení atomů v postranním řetězci, B) společná struktura pro všechny aminokyseliny (převzato z [1], upraveno v programu Marvin Sketch [2]).

Jednotlivé aminokyseliny jsou kovalentně spojovány peptidickou vazbou. Tento způsob spojení je vytvářen odštěpením vody z  $\alpha$ -karboxylové skupiny z jedné aminokyseliny a  $\alpha$ -amino skupiny z druhé [1].



Pro stanovení sekundární struktury pomocí NMR je nutno určit torzní (dihedrální) úhly  $\phi$  a  $\psi$  (fí a psí). Jedná se o úhly, které definují konformaci polypeptidového řetězce a tedy jeho sekundární strukturu [3]. Pokud máme systém obsahující 4 atomy A-B-C-D spojené jednoduchými vazbami, pak lze torzními úhly popsat konformaci mezi atomy B a C pomocí úhlu, který svírají roviny ABC a BCD. Velikost úhlů je od  $-180^\circ$  do  $+180^\circ$ . Vazby v hlavním řetězci (polypeptidická páteř) jsou označovány symboly dvou atomů, které ji tvoří, například  $N_i-C_i^\alpha$ ,  $C_i^\alpha-C_i'$ , atd. Torzní úhel popisující rotaci u vazby  $N-C^\alpha$  je označován jako  $\phi$  (fí), u vazby  $C^\alpha-C$  je označován jako  $\psi$  (psí) a u vazby  $C-N$  je označován jako  $\omega$  (omega) (viz obrázek 2) [4].



Obrázek 2: Znázornění torzních úhlů v peptidovém řetězci (převzato z [5]).

Většina proteinů má stabilní 3D strukturu, která zároveň určuje jeho funkci. Struktura proteinu je stabilizována několika druhy slabých interakcí. Vodíkové můstky,

disulfidické můstky a iontové interakce jsou nejčastěji ve specifických strukturách, které jsou termodynamicky stabilní (viz níže).

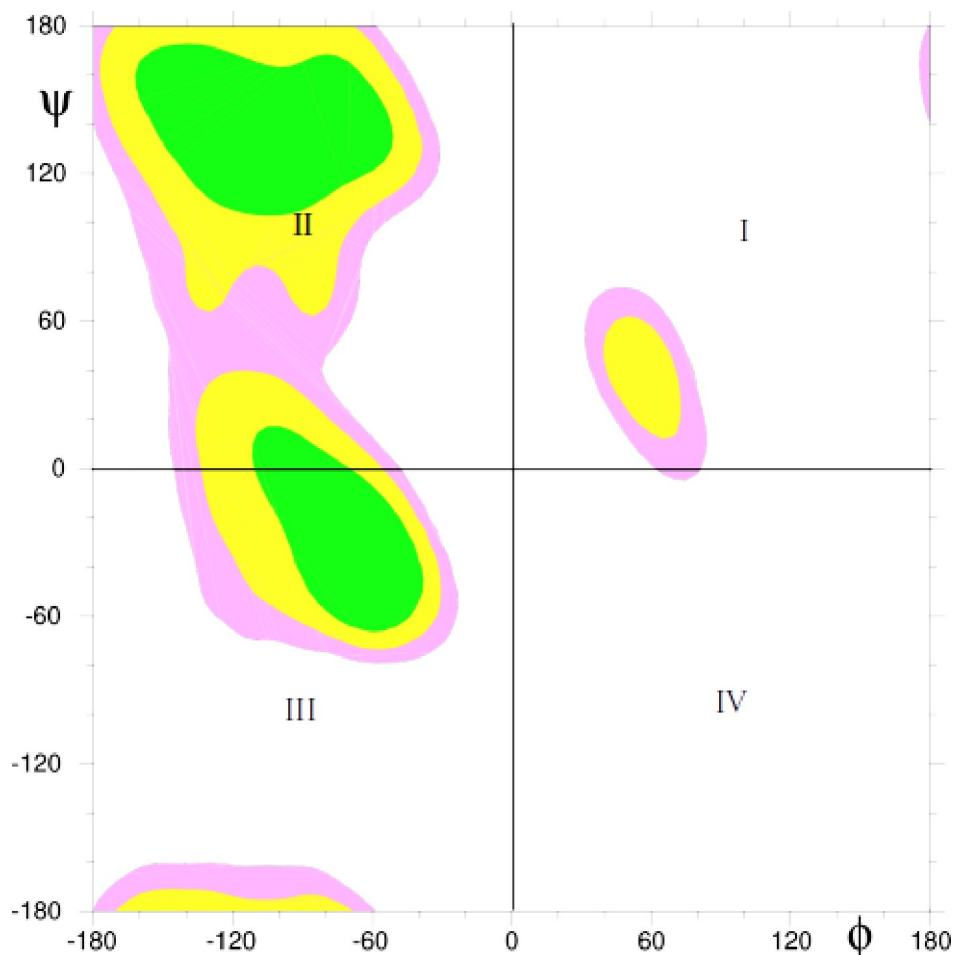
Nejjednodušší forma struktury proteinu je primární – sekvence aminokyselin v polypeptidickém řetězci. Sekvence proteinu je dána sekvencí DNA (deoxyribonukleová kyselina, z angl. deoxyribonucleic acid), která kóduje tvorbu jednotlivých proteinů. Malá změna v DNA genu může vést ke změně v aminokyselinové sekvenci proteinu a i jedna změna aminokyseliny může vyvolat změnu ve struktuře a funkci proteinu [1]. Nejznámějším příkladem je nemoc zvaná srpkovitá anémie (z angl. Sickle Cell Anemia, SCA), kdy dochází k mutaci bází kódující šestou aminokyselinu v  $\beta$ -řetězci hemoglobinu. Změna GTG páru za GAG způsobí tvorbu Valinu místo Glutamové kyseliny v syntetizovaném řetězci proteinu. To má za následek ovlivnění tvaru červených krvinek na tvar srpku. Červená krvinka s tímto tvarem není uzpůsobena pro průchod krevními cévami a dochází k jejich ucpaní, což může mít za následek například mrtvici [6].

Pojem sekundární struktura odpovídá poskládání polypeptidické páteře do určitých vzorů. Několik druhů sekundárních struktur vyskytujících se v proteinech je obzvláště stabilních. Jedno z možných uspořádání polypeptidického řetězce je šroubovicová struktura, kterou Pauling a Corey nazvali  $\alpha$  helix [7]. V této struktuře je polypeptidová páteř navinuta na imaginární osu a vedlejší skupiny aminokyselin směřují vně. Jedna opakující se jednotka odpovídá jedné otočce šroubovice, která prodlužuje vlákno o 5,4 Å a odpovídá zhruba 3,6 aminokyselinám. Ne všechny polypeptidy mají možnost zformovat se do stabilní šroubovice, protože interakce mezi jednotlivými aminokyselinovými zbytky může ovlivnit stabilizaci a destabilizaci struktury. Například, pokud je v řetězci více Glu za sebou, tak se tato část nemůže zformovat do  $\alpha$  helixu při pH 7,0 kvůli negativně nabitým částem karboxylové skupiny. Pauling a Corey dále předpověděli druhý typ opakující se struktury,  $\beta$  skládaný list. V  $\beta$  konformaci je polypeptidický řetězec prodloužen do klikaté struktury. Mezi jednotlivými řetězci dojde k rovnoběžnému spojení pomocí vodíkových můstků. Ideální  $\beta$  skládaný list je také šroubovicové uspořádání s velikou otočkou. Vedlejší polypeptidové řetězce v  $\beta$  skládaném listu mohou být buď paralelní nebo antiparalelní (stejná nebo opačná amino-karbonyl orientace).

Obecné 3D uspořádání jednotlivých atomů v proteinu je označováno jako terciární struktura. Důvodem uspořádání polypeptidového řetězce do terciárních struktur jsou interakce mezi vedlejšími řetězci aminokyselin jako jsou například vodíkové můstky, iontové interakce nebo disulfidické můstky. Dále existují proteiny, které jsou tvořeny několika polypeptidickými řetězci, též podjednotkami. Pokud jsou tyto podjednotky spojeny, je struktura proteinu označována jako kvartérní. Příkladem proteinu s kvartérní strukturou je například hemoglobin, který má 2  $\alpha$  a 2  $\beta$  podjednotky [1].

### 1.1.1 Ramachandranův diagram

Ramachandranův diagram udává pravděpodobnost (energii) jako závislost páteřních úhlů  $\psi$  ( $\psi$ ) a úhlů  $\phi$  ( $\phi$ ) každého aminokyselinového zbytku. Z tohoto diagramu lze vyčíst povolené (tzn. pravděpodobné) hodnoty torzních úhlů. Původně byl publikován v roce 1963 G.N.Ramachandranem, C.Ramakrishnanem a V.Sasisekharanem [8]. Ramachandranův diagram má 4 oblasti (rozděleno čtverci, viz obrázek 3, strana 12). První kvadrant označuje povolenou oblast pro malou část konformací. V této oblasti lze nalézt například levotočivý alfa-helix. Druhý kvadrant je oblast s nejpříznivější konformací atomů a zároveň se jedná o oblast s největší četností výskytu. Lze zde nalézt  $\beta$  skládaný list. Třetí kvadrant znázorňuje další obsáhlou oblast v diagramu. Jedná se o oblast s  $\alpha$  helixem. Poslední kvadrant je nejméně zastoupený, protože při těchto torzních úhlech dochází k velkému stericnému pnutí. Výjimku tvoří aminokyseliny prolin a glycin. Protože glycin neobsahuje uhlík  $C^\beta$  má velkou volnost v uspořádání. Naopak prolin kvůli své cyklické struktuře mnoho struktur nepovoluje [9].



Obrázek 3: Ramachandranův diagram ukazující četnost zastoupení torzních úhlů  $\phi$  ( $\phi$ ) a  $\psi$  ( $\psi$ ) v aminokyselinách - vytvořen pomocí programu MolMol (převzato a upraveno z [10])

## 1.2 Příprava izotopicky značeného proteinu

Velkým ulehčením přípravy proteinů bylo zavedení rekombinantních expresí. Dříve bylo nutné odebrat kilogramy biologického materiálu (tkáně, tělních tekutin, atd.) k získání velmi malého množství proteinu. Z teoretického hlediska je postup získání rekombinantního proteinu jednoduchý. Stačí použít gen, který kóduje hledaný protein, vektor, který máme k dispozici, a transformovat jej do hostitele, kterého si vybereme. Po indukci je protein připraven k purifikaci a přesnému určení. V praxi ale existuje několik faktorů, které mohou expresi ovlivnit. Příkladem může být špatný růst hostitelských buněk,

tvorba inkluzních tělísek (z angl. Inclusion body, IB), špatné poskládání proteinu, neaktivita proteinu nebo téměř úplně nulový výtěžek [11]. Proteiny pro NMR experimenty se běžně připravují rekombinantní expresí v *Escherichia Coli*. Nejčastěji se používá  $^{13}\text{C}$  obohacená glukóza a  $^{15}\text{N}$  obohacený chlorid amonný jako zdroj uhlíku  $^{13}\text{C}$  a dusíku  $^{15}\text{N}$ , které jsou nutné pro NMR experiment [12].

### 1.3 Určování 3D struktury proteinů

V současné době jsou pro určování 3D struktury proteinů na atomární úrovni používány čtyři metody: rentgenová strukturní analýza, nukleární magnetická rezonance, kryo-elektronová mikroskopie a výpočetní metody (in silico) [13]. Podle statistik serveru PDB (z angl. Protein Data Bank) o určených 3D strukturách lze říci, že pomocí rentgenové strukturní analýzy bylo vyřešeno zatím nejvíce struktur - 133 391 struktur (asi 90 %). Dále bylo vyřešeno pomocí NMR 12 534 struktur (asi 8 %) a pomocí kryo-elektronové mikroskopie 2841 struktur (asi 2 %) [14].

#### 1.3.1 Rentgenová strukturní analýza

Ačkoliv je rentgenová strukturní analýza (také rentgenová krystalografie, X-ray krystalografie, angl. X-ray crystallography) metoda stará více než století, je stále často používaná k řešení 3D struktury jak malých molekul, tak velkých makromolekul jako jsou proteiny [15]. První anorganická látka, u které byla určena v roce 1913 struktura pomocí X-ray krystalografie, byl diamant [16], první organická látka v roce 1923 hexamethyl tetraamin [17] a první protein v roce 1958 myoglobin [18]. V případě biomakromolekul připravit vhodně difraktující krystal je obvykle těžký úkol.

Při této metodě je krystal ozařován paprsky rentgenového záření. Krystal tyto paprsky rozloží do specifických směrů. Podle intenzity a úhlu rozložených paprsků lze poté pomocí počítačové techniky určit tří dimenzionální obraz zkoumaného vzorku [19].

#### 1.3.2 Kryo-elektronová mikroskopie

Kryo-elektronová mikroskopie (z anglického Cryo-electron microscopy, též cryo-EM) je metoda, která umožňuje určit 3D struktury molekul bez nutnosti tvorby krystalu. Vývoj této metody začal již v 70. letech minulého století, ale kvůli ne příliš citlivým detektorům se tato metoda nepoužívala tolik jako X-ray krystalografie. Tato metoda

v posledních 6 letech zaznamenala velký pokrok kvůli vývoji nových detektorů a stala se konkurenční metodou pro rentgenovou strukturní analýzu [20].

Základem je paprsek elektronů a systém elektronových čoček. Nejpoužívanější technika v dnešní době zahrnuje změření tisíců různých jednotlivých částic (z angl. single particles) uchovaných v tenké vrstvě nekystalického ledu. Protože každá molekula má v ledu různou orientaci, není nutné se vzorkem otáčet pro změření 3D struktury, jako je tomu u rentgenostrukturní analýzy. Pomocí výsledků tohoto měření z mnoha různých orientací získáme data o molekule, která lze vyhodnotit výpočetní technikou a zrekonstruovat tak 3D strukturu. Výhodou oproti rentgenové krystalografii či nukleární magnetické rezonanci je, že pro změření struktury pomocí kryo-elektronové mikroskopie je zapotřebí menší množství vzorku [21].

### 1.3.3 Nukleární magnetická rezonance

Zavedením nukleární magnetické rezonance (NMR), jako metody pro měření struktury proteinů, vedlo již v minulém století k objevení mnoha nových struktur. První NMR experimenty byly provedeny již v 60. letech minulého století [22].

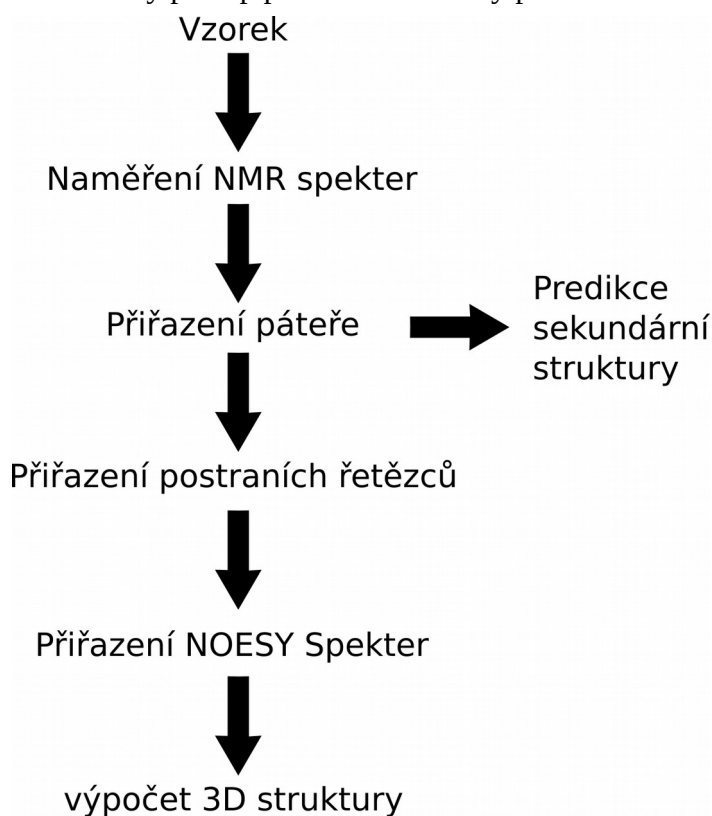
Fyzikální mechanismus techniky NMR může být shrnut pomocí názvu: nukleární (jev zahrnující jádra atomů), magnetická (pro experiment je nutné magnetické pole), rezonance (absorpce energie v dobře definované energii, tj. frekvenci). Většina prvků v periodické tabulce má alespoň 1 nukleární izotop (má nenulový jaderný spin), který interaguje s magnetickým polem a který lze použít pro NMR experimenty. Vlastnost jádra, která dovoluje tuto interakci, se nazývá spin a je vyjádřena pomocí spinového čísla. Pro NMR experimenty biomakromolekul je nutné použít jádra atomů se spinem 1/2.

Jádro vodíku (proton,  $^1\text{H}$ ), které má izotopické zastoupení téměř 100% volně v přírodě, a proto je jedno z nejdůležitějších jader ve studiích proteinů, má spin 1/2. Naopak jádro uhlíku  $^{12}\text{C}$  má spinové číslo 0 a jádro dusíku  $^{14}\text{N}$  má spinové číslo 1, proto je pro měření spekter biomakromolekul nutné izotopické obohacení. Izotopy  $^{13}\text{C}$  a  $^{15}\text{N}$  mají spin 1/2, ale jejich přirozený výskyt je nižší, a proto je nutné proteiny obohatit o tyto izotopy. Pro NMR experimenty lze v principu využít i izotopy  $^{17}\text{O}$  (se spinovým číslem 5/2) pro kyslík a  $^{33}\text{S}$  (se spinovým číslem 3/2) pro síru, ale kvůli své nízké citlivosti jsou

používány méně [23]. Jádra se sudým nukleonovým číslem a sudým protonovým číslem (například uhlík  $^{12}\text{C}$ ) mají vždy spin nulový, jádra se sudým nukleonovým číslem a lichým protonovým číslem (například dusík  $^{14}\text{N}$ ) mají vždy spin jednotkový a jádra s lichým nukleonovým číslem mají polo-číselný spin:  $1/2$ ,  $3/2$ ,  $5/2$ . Obecně platí pravidlo, že pro NMR experimenty biomakromolekul se jádra s jiným spinem než  $1/2$  téměř nepoužívají.

Principem měření NMR spekter je použití elektromagnetického pole, kdy jádro atomu vytvoří elektromagnetickou vlnu nebo oscilující změny v síle elektromagnetického pole. Rezonanční frekvence závisí primárně na typu jádra atomu a uspořádání elektronů v okolí atomu. Čím větší je magnetické pole, k tím rychlejšímu kmitání dochází [24].

#### 1.3.4 Obecný postup při řešení struktury pomocí NMR



Obrázek 4: Obecný postup při řešení struktury pomocí NMR (vytvořeno pomocí programu Inkscape)

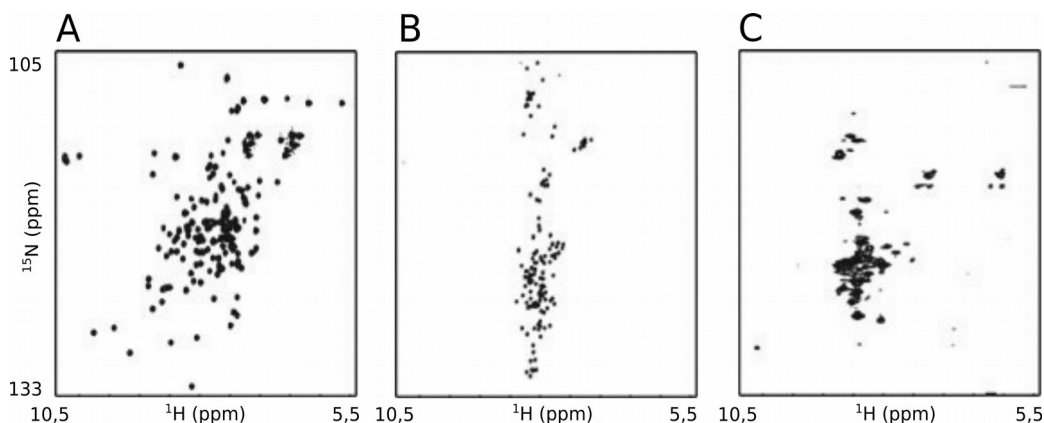
Z obrázku 4 je vidět, že samotné přiřazení páteře a postranních řetězců je nedílnou součástí NMR experimentů. Často se jedná o časově náročný proces, pokud je celý proces

prováděn manuálně. Po přiřazení páteře je již možné na základě naměřených rezonančních frekvencí provést predikci sekundární struktury.

## 1.4 Naměření NMR spekter – série spekter

### 1.4.1 Přiřazení páteře

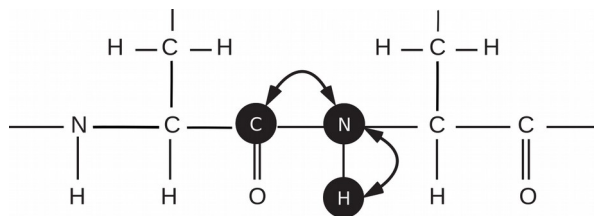
Jako první je nutné vyhodnotit  $^1\text{H}^{15}\text{N}$ -HSQC (z angl. Heteronuclear Single-Quantum Coherence) korelační spektrum. Toto spektrum nám pomůže určit rezonanční frekvence atomů vodíku a dusíku, které jsou spojené kovalentní vazbou. Celkový počet signálů poté zhruba odpovídá celkovému počtu aminokyselin v řetězci. V tomto spektru se nachází i další signály, které odpovídají amidovým skupinám postranních řetězců. Naopak nevidíme ve spektru signál prolinu, protože nemá vodík v amidové skupině.  $^1\text{H}^{15}\text{N}$ -HSQC spektrum je zároveň vhodné pro kontrolu správného sbalení proteinu. Na obrázku 5A je znázorněn protein, který má stabilní 3D strukturu. Pokud jsou ostré (s malou pološířkou) signály vysoce uskupeny při chemickém posunu vodíku od 7,5 ppm do 8,5 ppm, protein je neuspořádaný (obrázek 5B). Pokud protein agreguje do více struktur, tak vzniká velký shluk signálů s velkou pološířkou (obrázek 5C) [25].



Obrázek 5: A) Spektrum proteinu se stabilní strukturou. B) Spektrum proteinu s neuspořádanou strukturou. C) Spektrum proteinu s více strukturami (převzato z [24]).

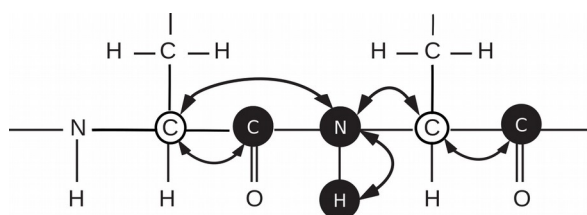
Pro přiřazení páteře proteinu je dále nutné použít 3D spektra. První experiment HNCO znázorňuje korelaci mezi chemickým posunem amidové skupiny  $^1\text{H}$  a  $^{15}\text{N}$  s chemickým posunem uhlíku  $^{13}\text{C}$  v karbonylu předchozí aminokyseliny (Obrázek 6, strana 17) [26].





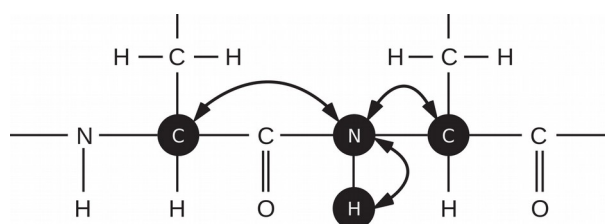
Obrázek 6: Graficky znázorněná korelace ve spektru HNCO (převzato z [26]).

Druhý experiment HN(CA)CO znázorňuje korelaci mezi chemickým posunem vlastní amidové skupiny  $^1\text{H}$  a  $^{15}\text{N}$  s chemickým posunem uhlíku  $^{13}\text{C}$  v karbonylu vlastní aminokyseliny. Pomocí těchto dvou spekter již lze vidět pořadí aminokyselin v řetězci proteinu (Obrázek 7) [27].



Obrázek 7: Graficky znázorněná korelace ve spektru HN(CA)CO (převzato z [27]).

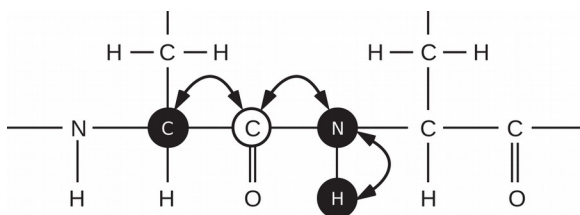
Třetí experiment HNCA měří korelaci mezi chemickým posunem amidové skupiny  $^1\text{H}$  a  $^{15}\text{N}$  s chemickým posunem  $\alpha$  uhlíku  $^{13}\text{C}$ . V tomto spektru lze dále nalézt i signál chemického posunu  $\alpha$  uhlíku předchozí aminokyseliny (Obrázek 8).



Obrázek 8: Graficky znázorněná korelace ve spektru HNCA (převzato z [28]).

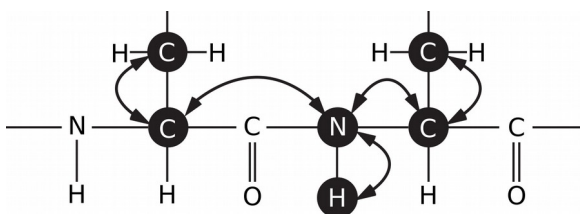
Čtvrtý experiment HN(CO)CA zobrazuje korelaci mezi chemickým posunem vlastní amidové skupiny  $^1\text{H}$  a  $^{15}\text{N}$  s chemickým posunem  $\alpha$  předchozího uhlíku  $^{13}\text{C}$ . Pro kontrolu je vhodné srovnat spektra HN(CO)CA s HNCA. Pokud je signál chemického posunu předchozího  $\alpha$  uhlíku  $^{13}\text{C}$  špatně rozpoznatelný ve spektru HNCA, je vhodné

přřadit signály CA nejprve ve spektru HN(CO)CA a výsledek použít ve spektru HNCA (Obrázek 9) [28].



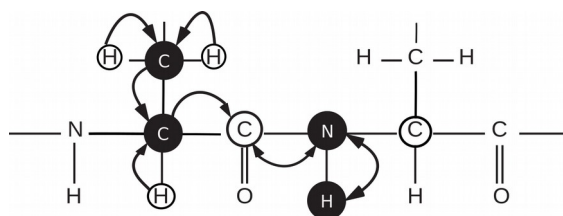
Obrázek 9: Graficky znázorněná korelace ve spektru HN(CO)CA (převzato z [28]).

Pátý experiment HNCACB pomáhá určit korelaci mezi chemickým posunem vlastní amidové skupiny  $^1\text{H}$  a  $^{15}\text{N}$  s chemickým posunem vlastních  $\alpha$  a  $\beta$  uhlíků. Dále lze v tomto spektru nalézt i signály chemických posunů předchozích  $\alpha$  a  $\beta$  uhlíků  $^{13}\text{C}$  (Obrázek 10) [29].



Obrázek 10: Graficky znázorněná korelace ve spektru HNCACB (převzato z [29]).

Šestý experiment CBCA(CO)NH určuje korelaci mezi chemickým posunem vlastní amidové skupiny  $^1\text{H}$  a  $^{15}\text{N}$  s chemickým posunem předchozích  $\alpha$  a  $\beta$  uhlíků  $^{13}\text{C}$ . Opět je vhodné toto spektrum porovnávat s předchozím spektrem HNCACB, ale zároveň je vhodné zkontrolovat chemické posuny  $\alpha$   $^{13}\text{C}$  uhlíků předchozích aminokyselin, které jsou viditelné ve spektrech HNCA a HN(CO)CA (Obrázek 11) [30].



Obrázek 11: Graficky znázorněná korelace ve spektru CBCA(CO)NH (převzato z [30]).

## 1.5 Princip přiřazení aminokyselin

Pro přiřazení jednotlivých aminokyselin je nutné určit rezonanční frekvence každého atomu v molekule proteinu. Hodnoty rezonančních frekvencí, které odpovídají signálům, jsou hledány ve 2 spektrech. Po nalezení stejné frekvence atomu v předchozí a vlastní aminokyselině, lze určit pořadí jednotlivých aminokyselin v sekvenci. Celý postup je podrobněji popsán v oddílu Metody.

## 1.6 Nkrp1c

Přírodně zabíječské buňky (z angl. Natural Killer cells, NK cells) představují velkou podtřídu lymfocytů, které hrají klíčovou roli při eliminaci virových infekcí a nádorových onemocnění. Podstatou jejich cytolytické aktivity je receptorová interakce [31].

Na povrchu NK buněk lze nalézt 2 typy receptorů – pozitivní (stimulační) a negativní (inhibiční). Stimulační receptory jsou schopné poznat látky nacházející se na povrchu různých typů buněk (např. některé druhy MHC proteinů – hlavní histokompatibilní komplex – z angl. Major Histocompatibility Complex). Tyto receptory poté vyvolají signál pro NK buňky pro cytotoxické mechanismy. Inhibiční receptory rozpoznávají proteinové komplexy MHC I skupiny na povrchu buněk. Na ně se váží cizorodé peptidy, které pocházejí například z některých virových proteinů. Jedná se o negativní signály, tedy inhibici cytotoxických mechanismů. Receptory můžeme rozdělit podle struktury na 2 druhy – NK receptory podobné imunoglobulinu a NK receptory podobné C-lektinům, kam patří právě Nkrp1c. Výsledný efekt funkce NK buněk při kontaktu s jinou buňkou je závislý na tom, převažuje-li stimulační nebo inhibiční signál. Pokud je na povrchu buňky obvyklé množství komplexů MHC skupiny I, které nesou normální peptidy, NK buňka dostane negativní signály a k aktivaci nedojde. Pokud ale na povrchu buněk není dostatečné množství komplexů MHC skupiny I nebo pokud jsou tyto molekuly upraveny tak, že nejsou rozeznány inhibičními receptory, dojde k vysílání stimulačních signálů a zabití buňky pomocí NK buňky cytotoxicky [32].

## 2 Cíle práce

- Přiřazení signálu peptidové páteře v již naměřených NMR spektrech myšního proteinu Nkrp1c.
- Předpověď sekundární struktury na základě přiřazených chemických posunů atomů peptidové páteře.
- Srovnání předpovězených motivů sekundárních struktur s dostupným 3D modelem.

## 3 Metody

### 3.1 Naměření NMR spekter

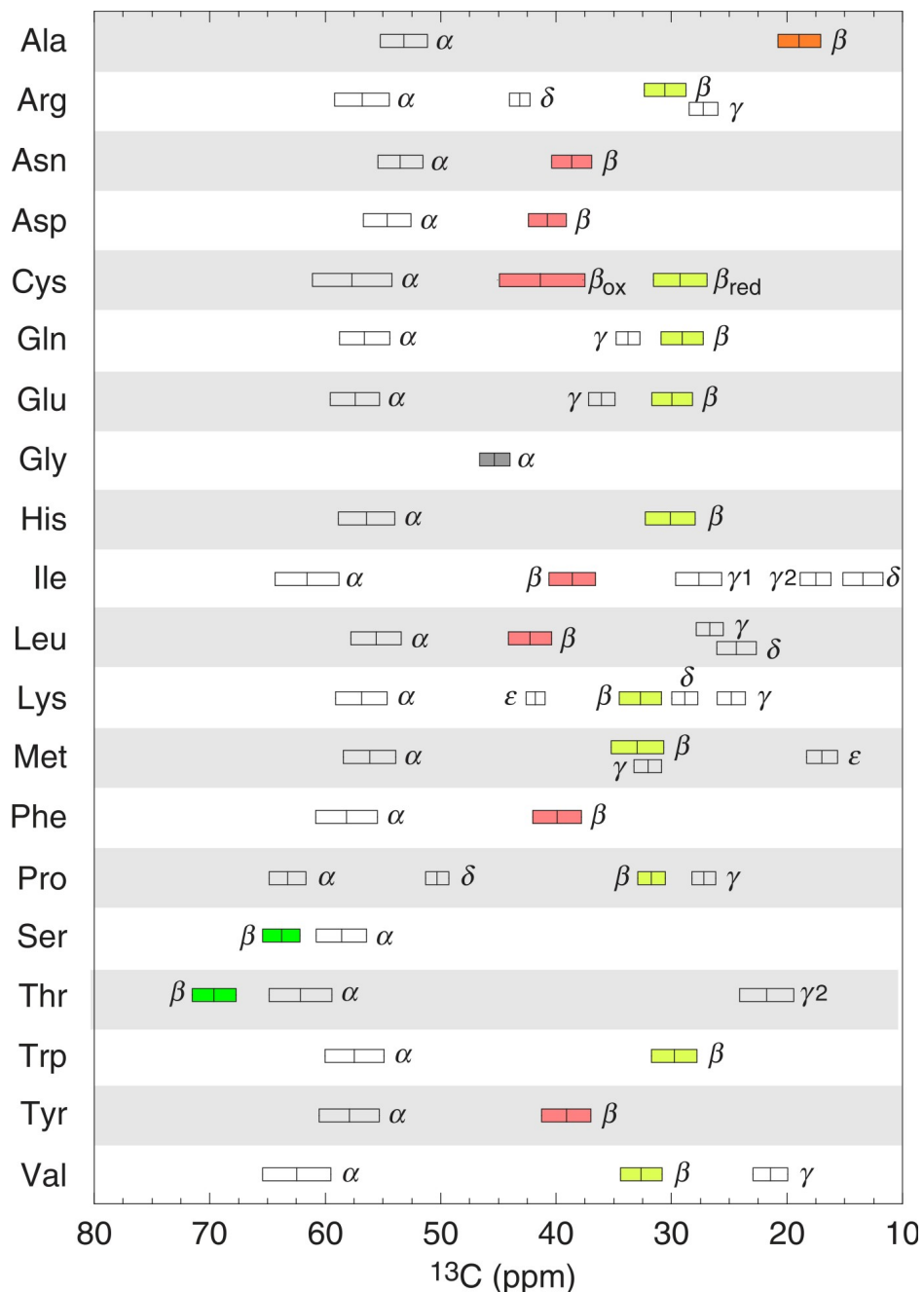
Spektra byla naměřena na spektrometru Bruker Avance III 600 MHz vybaveném citlivou TCI 1H-13C-15N kryo-sondou. Pro měření bylo použito 400  $\mu$ l 1,7 mM vzorku spolu se 100  $\mu$ l pufru (složení: 15 mM PIPES, 50 mM NaCl, 1 mM NaN<sub>3</sub>, pH = 6,8) a 10% D<sub>2</sub>O. Měření bylo prováděno při teplotě 303,1 K.

### 3.2 Vyhodnocení NMR spekter

Spektra byla vyhodnocena pomocí programu Sparky (verze 3.115) [33]. Přiřazení bylo prováděno manuálně, aby nedošlo k chybnému označení signálu a následnému špatnému přiřazení. Pokud dojde k překryvu signálu, automatický výběr může tento signál vyhodnotit jako jeden. Naopak slabý signál, který už neodpovídá signálu aminokyselin, ale jedná se pouze o šum, může vyhodnotit jako signál. Nejprve byly ve spektru <sup>1</sup>H<sup>15</sup>N-HSQC všechny signály pouze označeny, protože nelze pouze z tohoto spektra říci, o jakou aminokyselinu se jedná.

Signály ve 3D spektrech byly označeny v proton-uhlíkové korelaci. Každá aminokyselina má charakteristický <sup>13</sup>C chemický posun, pomocí kterého byly jednotlivé aminokyseliny přiřazovány.

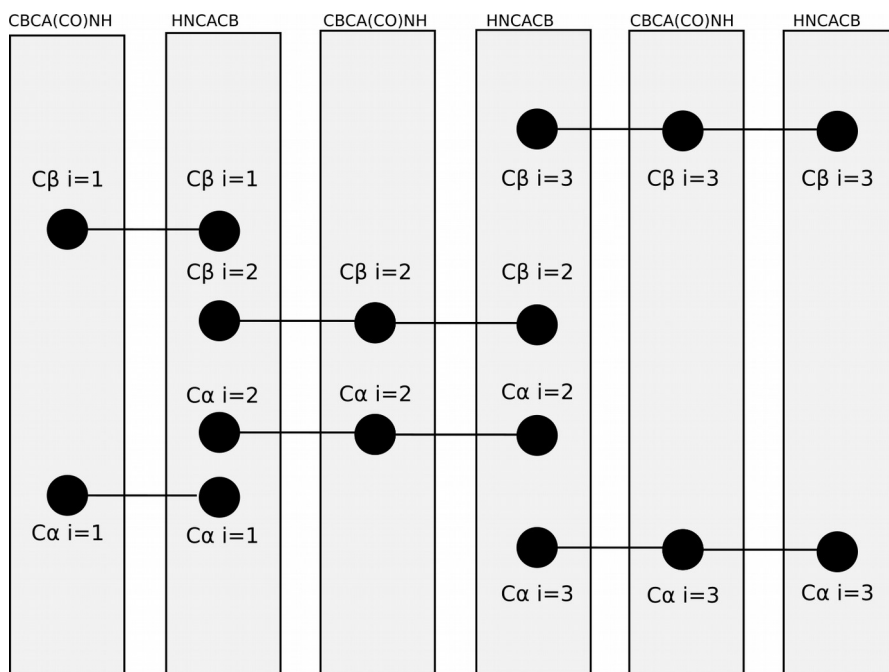
Pro správné identifikace aminokyselin je vhodné využít chemických posunů uhlíku C<sup>β</sup>, protože uhlíky C<sup>α</sup> nemají tak specifické rozdíly v chemických posunech. Chemické posuny uhlíků C<sup>β</sup> lze rozdělit do 5 skupin (obrázek 12, str. 22) - Gly nemá signál pro uhlík C<sup>β</sup>, jelikož jej nemá, Ala má charakteristický chemický posun C<sup>β</sup> při 20 ppm (na obrázku označen oranžově). Arg, redukovaný Cys, Gln, Glu, His, Lys, Met, Trp, Val patří do skupiny s chemickým posunem C<sup>β</sup> zhruba od 30 do 35 ppm. Asn, Asp, oxidovaný Cystein, Ile, Leu, Phe, Tyr patří do skupiny s chemickým posunem C<sup>β</sup> v rozmezí od 38 do 45 ppm. Do poslední skupiny patří Ser a Thr jejichž chemický posun C<sup>β</sup> je pro Ser od 62 ppm do 66 ppm a Thr od 68 do 73 ppm. Vzhledem k vysoké odlišnosti jsou aminokyseliny Ala, Ser, Thr a Gly nejsnáze přiřaditelné.



Obrázek 12: Typické chemické posuny uhlíků v aminokyselinách. Uhlíky  $\text{C}^\beta$  jsou rozděleny do několika skupin, které jsou barevně rozlišeny, pro lepší orientaci při přiřazování signálů k aminokyselinové sekvenci (převzato z [25]).

Pomocí funkce Strip Plot v programu Sparky [33] lze poté hledat jednotlivé frekvence  $\text{C}^\beta$  uhlíků aminokyselin ve spektrech HNCACB a CBCA(CO)NH. Pro kontrolu

lze využít také uhlíky  $C^\alpha$ . Pro ilustraci postupu hledání pořadí aminokyselin byl vytvořen obrázek 13. Principem této funkce je porovnání chemických posunů  $C^\beta$  a případně  $C^\alpha$ . Výsledkem je poté několik signálů s určitou pravděpodobností shody, které je nutno manuálně zkontrolovat. Pro tyto účely je vhodné také porovnat signály ve spektrech HNCO a HN(CA)CO, aby byla ověřena správnost přiřazení.



Obrázek 13: Schéma přiřazení aminokyselin ve spektrech CBCA(CO)NH a HNCACB pomocí funkce Strip plot pro 3 aminokyseliny.

### 3.3 Predikce sekundární struktury

Pro predikci sekundární struktury proteinu byly použity 2 přístupy – predikce podle chemických posunů (program Talos+) a poté predikce pouze na základě primární sekvence (program PSIPRED).

#### 3.3.1 Talos+

Talos+ je program pro predikci  $\phi$  (phi) a  $\psi$  (psi) torzních úhlů na základě změřených chemických posunů. Funguje tak, že porovná kombinaci šesti druhů chemických posunů – HN,  $H^\alpha$ ,  $C^\alpha$ ,  $C^\beta$ ,  $C'$ , N daného proteinu. Využívá metody, kdy použije triplikáty aminokyselin a porovná změřené chemické posuny s chemickými posuny

v databázi o více než 200 proteinech se známou strukturou, s vysokým rozlišením z rentgenové krystalografie a se známými chemickými posuny a porovná jejich podobnost[34].

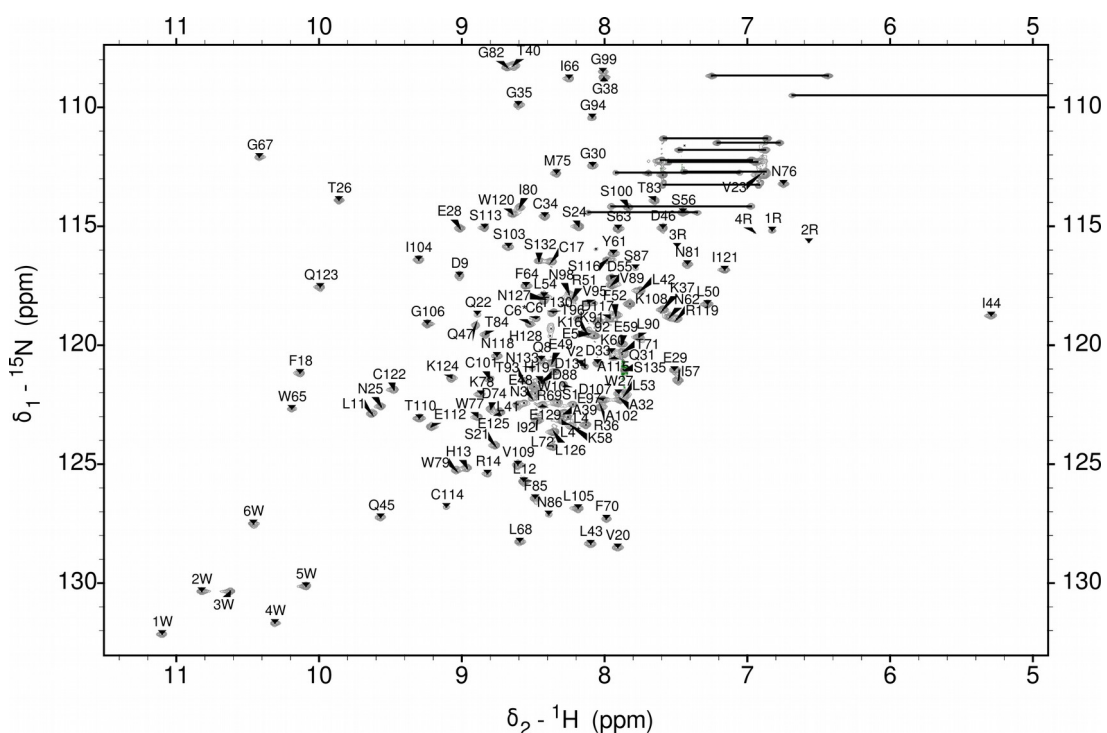
### 3.3.2 PSIPRED

PSIPRED (z angl. PSI-blast based secondary structure PREDiction) vyhodnotí na základě primární sekvence sekundární strukturu proteinu pomocí metody porovnávání sekvence a struktury jiných podobných proteinů. Principem je využití dat evolučně příbuzných proteinů [35].

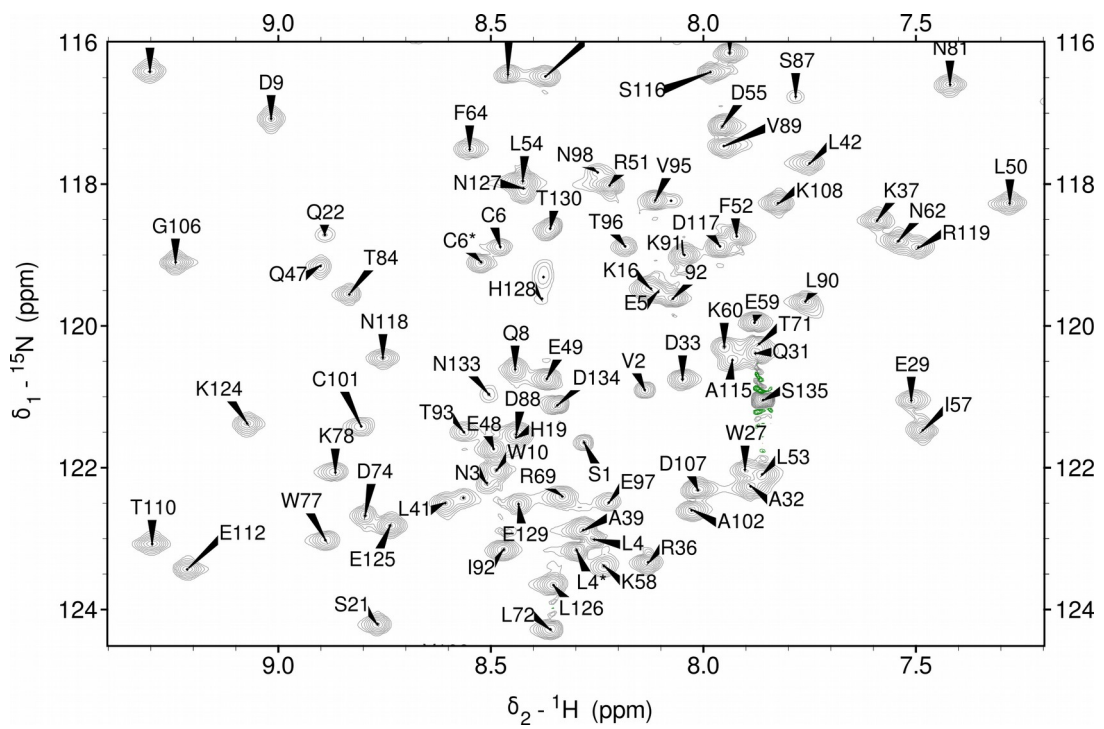


## 4 Výsledky

V této práci byly nejprve přiřazeny jednotlivé signály ve spektru  $^1\text{H}^{15}\text{N}$ -HSQC. Po zjištění přesného pořadí signálů, byly jednotlivé vrcholy signálů označeny příslušnými aminokyselinami podle sekvence. Pro větší přehlednost je uveden výřez středové části (Obrázek 15, strana 26). Na celkovém obrázku (Obrázek 14) je vidět několik spojených signálů odpovídající amidovým skupinám nacházejících se v postranním řetězci. V levé dolní části jsou dále viditelné signály NH skupin v indolovém postranním řetězci aminokyseliny tryptofanu. Pro kontrolu byl počet signálů porovnán s počtem tryptofanů v sekvenci – 6 tryptofanů odpovídá 6 signálům.



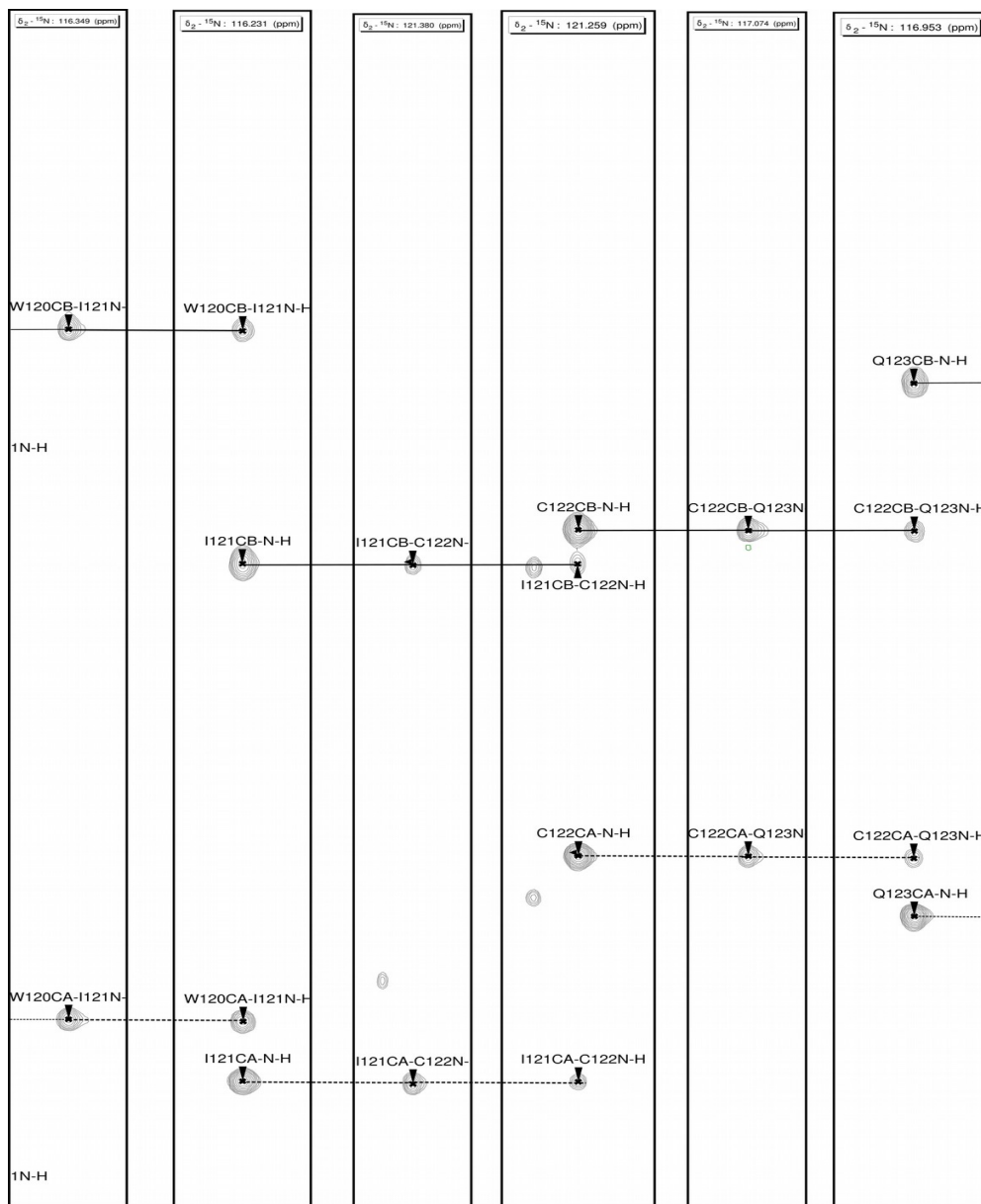
Obrázek 14: Přiřazené signály jednotlivých aminokyselin ve spektru  $^1\text{H}^{15}\text{N}$ -HSQC.



Obrázek 15: Přiřazené signály jednotlivých aminokyselin ve spektru  $^1\text{H}^{15}\text{N}$ -HSQC - detailnější zobrazení středové části.

#### 4.1 HNCACB+CBCA(CO)NH

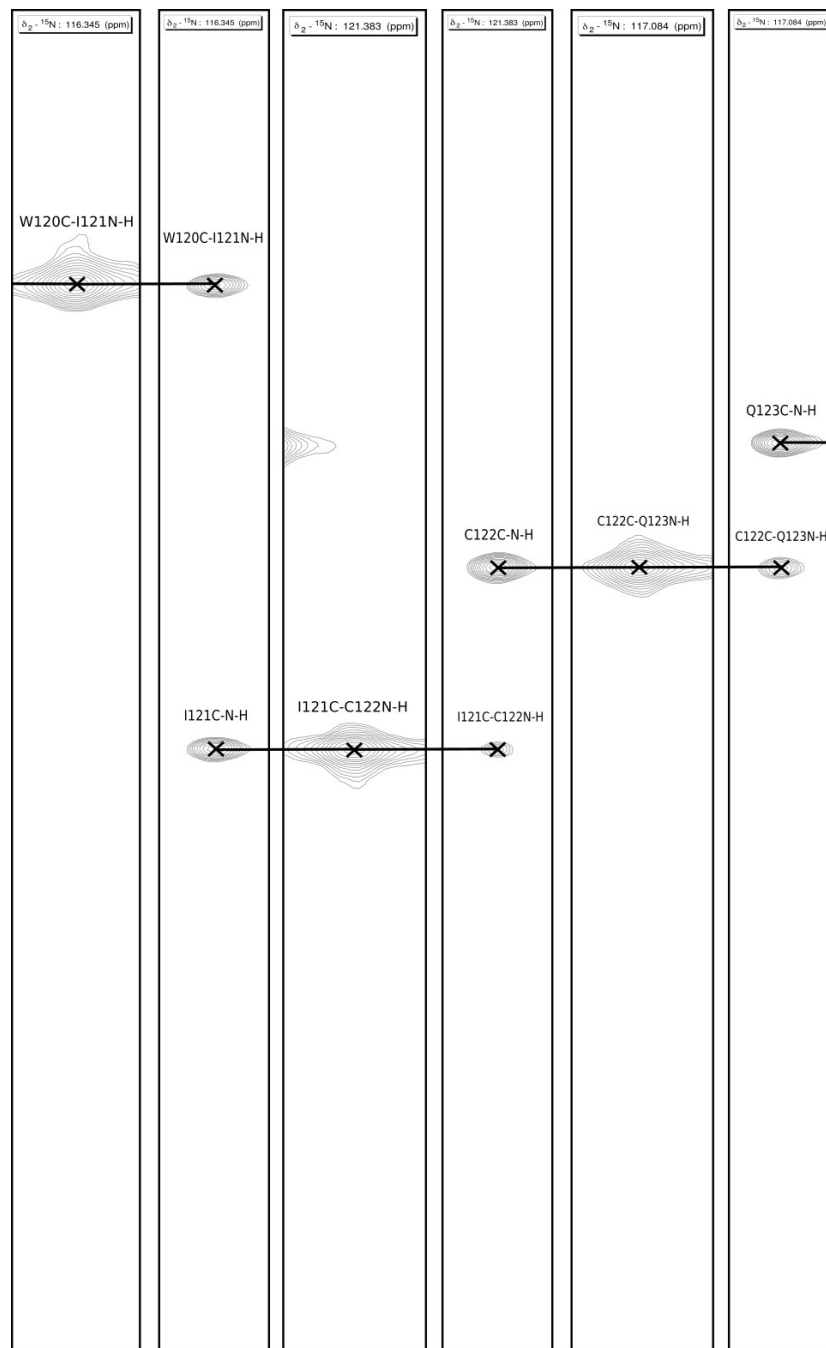
Pomocí funkce Strip plot programu Sparky byly přiřazeny jednotlivé signály aminokyselin k jejich pořadí v primární sekvenci (Obrázek 16).



Obrázek 16: Přiřazené aminokyseliny (od W120 po Q123) k signálům ve spektrech HNCACB a CBCA(CO)NH. Přerušovanou čarou jsou spojeny atomy  $\text{C}^\alpha$  a plnou čarou jsou spojeny atomy  $\text{C}^\beta$ .

## 4.2 HNCO a HN(CA)CO

Pro kontrolu byly ověřeny i signály ve spektrech HNCO a HN(CA)CO. Podle obrázku 17 je vidět správné určení pořadí signálů odpovídající aminokyselinové sekvenci.



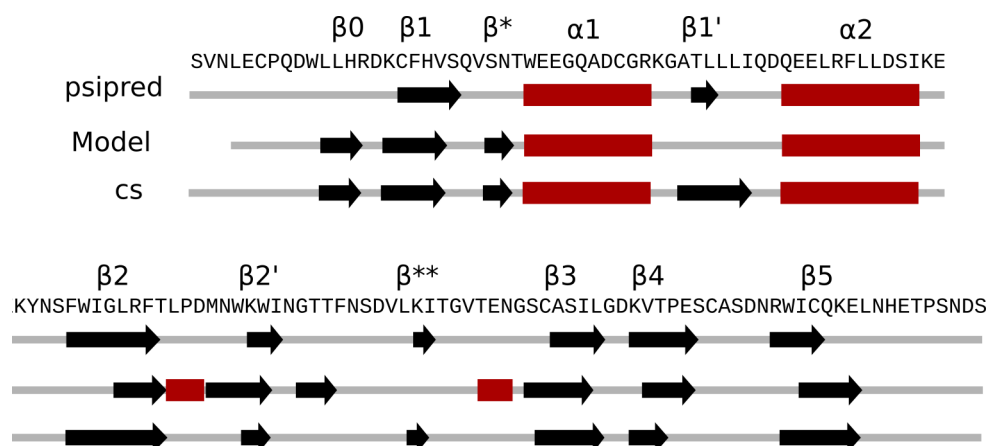
Obrázek 17: Přiřazené aminokyseliny (od W120 po Q123) k signálům ve spektrech HNCO a HN(CA)CO.

Z programu Sparky [33] byly exportovány data chemických posunů a jednotlivých přiřazení. Data byla vložena do tabulky a jsou uvedena v příloze. V tabulce jsou zaznamenány aminokyselinové zbytky, příslušné atomy a jejich chemické posuny. C značí karbonyl, CA označuje uhlík C<sup>α</sup>, CB uhlík C<sup>β</sup>, MA označuje acetylovou skupinu. Pomocí symbolu \* jsou poté označeny zdvojené signály, které odpovídají N-acetylované formě proteinu Nkrp1c.

#### 4.3 Predikce sekundární struktury

Hlavním výsledkem této bakalářské práce je predikce sekundární struktury na základě přiřazených chemických posunů a srovnání předpovězené sekundární struktury s publikovaným modelem. Po vyhodnocení spekter v programu Sparky [33] byly použity 2 programy pro predikci sekundární struktury – Talos+ a PSIPRED. Výsledky z této predikce byly poté porovnány s již publikovaným homologním modelem receptoru Nkrp1c [36]. Získaná data byla poté graficky zpracována a jsou uvedena v obrázku 18.

V obrázku 18 je uvedeno číslování jednotlivých sekundárních struktur podle univerzálního číslování rodiny C-lektinů [37]. Dva β skládané listy, které jsou v této struktuře navíc, byly označeny jako β\* a β\*\*. Srovnáním jednotlivých struktur na obrázku 18 (strana 30) lze pozorovat velkou podobnost predikovaných sekundárních struktur programy Talos+ a PSIPRED (cs a psipred, obrázek 18, strana 30) se strukturou homologního modelu receptoru Nkrp1c (Model, obrázek 18, strana 30). Predikce α helixů (červené obdelníky, obrázek 18, strana 30) je totožná s homologním modelem a predikce β skládaných listů je téměř totožná (černé šipky, obrázek 18, strana 30).



Obrázek 18: Grafické porovnání výsledků predikce sekundární struktury získané pomocí programu Talos+ (na obrázku označen jako cs = z angl. chemical shifts) [34] a PSIPRED [35] a následné porovnání s homologním modelem receptoru Nkrp1c [36]. Černé šipky znázorňují  $\beta$  skládaný list a červené obdélníky  $\alpha$  helix. Je použito obecné označení jednotlivých  $\alpha$  helixů a  $\beta$  skládaných listů charakteristické pro C-lektiny [37].

## 5 Diskuze

Myší protein Nkrp1c, který byl připraven rekombinantní expresí v *E. coli* a obohacen izotopy  $^{13}\text{C}$  a  $^{15}\text{N}$ , byl změřen pomocí metody nukleární magnetické rezonance. Pro predikci sekundární struktury bylo nutné vyhodnotit naměřená spektra.

Podle spektra HN-HSQC bylo možné ověřit, že protein je správně složený. Signály jednotlivých aminokyselin se příliš nepřekrývaly, protein tedy netvořil neuspořádané struktury nebo naopak vzorek neobsahoval více struktur a přiřazení tedy bylo možné. Pro predikci sekundárních struktur bylo použito šesti 3D spekter: HNC(O) a HN(CA)CO pro přiřazení karbonylů, HNCA a HN(CO)CA na identifikaci uhlíků  $\text{C}^\alpha$  a HNCACB a CBCA(CO)NH na přiřazení signálů  $\text{C}^\beta$ . V posledních dvou spektrech je možné ověřit správnost přiřazení uhlíků  $\text{C}^\alpha$ , protože se jejich signály v těchto spektrech také nacházejí.

Pomocí metody Strip plot programu Sparky byly přiřazeny signály k jednotlivým aminokyselinám. Nepovedlo se přiřadit amidickou skupinu u aminokyseliny D15, kde pravděpodobně došlo k rozšíření signálu zvýšenou dynamikou aminokyselinového zbytku (na milisekundové až mikrosekundové časové škále) nebo vyšší výměnou protonu v NH skupině za proton vody, a proto tento signál nemohl být přiřazen. Všechny ostatní signály byly přiřazeny k jednotlivým aminokyselinovým zbytkům.

Jedním z výsledků této práce je tabulka chemických posunů jednotlivých atomů v proteinu, která je uvedena v přílohách. Data z této tabulky byla využita pro predikci sekundární struktury v programu Talos+. Dále je možné v tabulce nalézt u některých aminokyselinových zbytků označení \*, které značí neacetylovanou formu proteinu Nkrp1c. Ve spektrech bylo viditelné zdvojení signálů aminokyselinových zbytků u N-konce, které odpovídalo právě acetylované a neacetylované formě.

Po predikci sekundární struktury programy Talos+ a PSIPRED byla získána data, která byla graficky znázorněna (viz Obrázek 18, strana 29). Do obrázku bylo přidáno znázornění sekundární struktury homologního modelu proteinu Nkrp1c. Po porovnání lze říci, že umístění alfa helixů a beta skládaných listů ve struktuře je velmi podobné.

## 6 Závěr

Shrnutí výsledků v bodech:

- Pomocí programu Sparky, použitého k vyhodnocení signálů z měření nukleární magnetické rezonance myšního proteinu Nkrp1c, byly přiřazeny jednotlivé signály peptidové páteře proteinu
- Na základě získaných chemických posunů atomů byla provedena predikce sekundární struktury pomocí programu Talos+
- Pomocí primární sekvence byla provedena predikce sekundární struktury programem PSIPRED
- Obě předpovězené sekundární struktury byly srovnány s homologním modelem proteinu



## 7 Zdroje

- 1 Nelson, D. L., Cox, M. M. a Lehninger, A. L. (2005) *Lehninger principles of biochemistry*. 4th ed.
- 2 <https://chemaxon.com/products/marvin> (přístup 1. 2. 2019)
- 3 Kodíček, M., Valentová, O. a Hynek, R. (2015) *Biochemie: chemický pohled na biologický svět*, Vysoká škola chemicko-technologická v Praze, Praha.
- 4 (1970) Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains. Tentative Rules (1969)\*. *Eur. J. Biochem.*, John Wiley & Sons, Ltd (10.1111) **17**, 193–201.
- 5 [https://commons.wikimedia.org/wiki/File:Protein\\_backbone\\_PhiPsiOmega\\_drawing.svg](https://commons.wikimedia.org/wiki/File:Protein_backbone_PhiPsiOmega_drawing.svg) (přístup 6. 4. 2019)
- 6 Fernandes, Q. (2017) Therapeutic strategies in Sickle Cell Anemia: The past present and future. *Life Sci.*, Pergamon **178**, 100–108.
- 7 Pauling, L., Corey, R. B. A Branson, H. R. (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.*, National Academy of Sciences **37**, 205–11.
- 8 Ramachandran, G. N., Ramakrishnan, C. a Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, Academic Press **7**, 95–99.
- 9 Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S. a Richardson, D. C. (2003) Structure validation by C $\alpha$  geometry:  $\phi, \psi$  and C $\beta$  deviation. *Proteins Struct. Funct. Bioinforma.* **50**, 437–450.
- 10 Koradi, R., Billeter, M. a Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.*, Elsevier **14**, 51–55.
- 11 Rosano, G. L., Ceccarelli, E. A., Neubauer, P., Bruno-Barcena, J. M. a Schweder, T. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges.
- 12 Rozbeský, D., Kavan, D., Chmelík, J., Novák, P., Vaněk, O. a Bezouška, K. (2011) High-level expression of soluble form of mouse natural killer cell receptor NKR-P1C(B6) in *Escherichia coli*. *Protein Expr. Purif.* **77**, 178–184.

- 13 Rosenbusch, J. P., Lustig, A., Grabo, M., Zulauf, M. a Regenass, M. Approaches to determining membrane protein structures to high resolution: do selections of subpopulations occur?
- 14 [http://www.rcsb.org/pdb/static.do?p=general\\_information/pdb\\_statistics/index.html](http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html)& (přístup 24. 2. 2019)
- 15 Powell, H. R. (2017) X-ray data processing. *Biosci. Rep.* **37**, BSR20170227.
- 16 Bragg, W. L. (1913) The Structure of Some Crystals as Indicated by Their Diffraction of X-rays. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **89**, 248–277.
- 17 Dickinson, R. G. a Raymond, A. L. (1923) The Crystal Structure Of Hexamethylene-Tetramine. *J. Am. Chem. Soc., American Chemical Society* **45**, 22–29.
- 18 Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. A Phillips, D. C. (1958) A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature*, Nature Publishing Group **181**, 662–666.
- 19 Helliwell, J. R. (2017) New developments in crystallography: exploring its technology, methods and scope in the molecular biosciences. *Biosci. Rep.*
- 20 Cheng, Y., Grigorieff, N., Penczek, P. A. a Walz, T. (2015) A Primer to Single-Particle Cryo-Electron Microscopy. *Cell*, Cell Press **161**, 438–449.
- 21 Murata, K. a Wolf, M. (2018) Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim. Biophys. Acta - Gen. Subj.*, Elsevier **1862**, 324–334.
- 22 Wüthrich, K. (1990) Protein structure determination in solution by NMR spectroscopy. *J. Biol. Chem.* **265**, 22059–62.
- 23 Edwards, A. J. a Reid, D. (2000) Introduction to NMR of Proteins. In *Current Protocols in Protein Science*, s 17.5.1-17.5.39, John Wiley & Sons, Inc., Hoboken, NJ, USA.
- 24 Doucleff, M., Hatcher-Skeers, M. a Crane, N. J. (2011) Protein Dynamics. In *Pocket Guide to Biomolecular NMR*, s 131–156, Springer Berlin Heidelberg, Berlin, Heidelberg.
- 25 Cavanagh, J. (2007) *Protein NMR spectroscopy : principles and practice*, Academic Press.
- 26 Clubb, R. T., Thanabal, V. a Wagner, G. (1992) A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue <sup>1</sup>HN, <sup>15</sup>N, and <sup>13</sup>C'

- chemical shifts in  $^{15}\text{N}$   $^{13}\text{C}$ -labelled proteins. *J. Magn. Reson.*, Academic Press **97**, 213–217.
- 27 Kay, L. E., Xu, G. Y. a Yamazaki, T. (1994) Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal  $\text{H}_2\text{O}$  Saturation. *J. Magn. Reson. Ser. A*, Academic Press **109**, 129–133.
- 28 Grzesiek, S. a Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.*, Academic Press **96**, 432–440.
- 29 Wittekind, M. a Mueller, L. (1993) HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *J. Magn. Reson. Ser. B*, Academic Press **101**, 201–205.
- 30 Grzesiek, S. a Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly  $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J. Biomol. NMR* **3**, 185–204.
- 31 Vivier, E. (2006) What is natural in natural killer cells? *Immunol. Lett.* **107**, 1–7.
- 32 Hořejší, V., Bartůňková, J., Brdička, T. a Spíšek, R. (2017) *Základy imunologie*, Triton.
- 33 Lee, W., Tonelli, M. a Markley, J. L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics*, Oxford University Press **31**, 1325–7.
- 34 Shen, Y., Delaglio, F., Cornilescu, G. a Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*, Springer Netherlands **44**, 213–223.
- 35 Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices 1 Edited by G. Von Heijne. *J. Mol. Biol.* **292**, 195–202.
- 36 Rozbesky, D., Sovova, Z., Marcoux, J., Man, P., Etrich, R., Robinson, C. V. a Novak, P. (2013) Structural Model of Lymphocyte Receptor NKR-P1C Revealed by Mass Spectrometry and Molecular Modeling. *Anal. Chem.* **85**, 1597–1604.
- 37 Zelensky, A. N. a Gready, J. E. (2005) The C-type lectin-like domain superfamily. *FEBS J.* **272**, 6179–6217.

## 8 Přílohy

### 8.1 Tabulka výsledků přiřazení signálů aminokyselin

\* označuje neacetylovanou část proteinu

Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun
ace0	C	177,226	E5*	CA	55,821	L12	CA	55,137
ace0	CA	24,595	E5*	CB	32,067	L12	CB	44,561
ace0	MA	2,11	C6	C	172,083	L12	H	8,567
S1	C	174,843	C6	CA	52,487	L12	HA	5,213
S1	CA	58,409	C6	CB	38,947	L12	N	125,724
S1	CB	64,087	C6	H	8,478	H13	C	173,606
S1	H	8,28	C6	HA	4,146	H13	CA	57,453
S1	HA	4,494	C6	N	118,947	H13	CB	31,532
S1	N	121,633	C6*	C	172,111	H13	H	8,965
V2	C	175,582	C6*	CA	52,549	H13	HA	4,289
V2	CA	62,151	C6*	CB	38,91	H13	N	125,146
V2	CB	32,935	C6*	H	8,522	R14	C	173,868
V2	H	8,138	C6*	N	119,095	R14	CA	58,652
V2	HA	4,193	P7	C	175,829	R14	CB	27,8
V2	N	120,913	P7	CA	62,466	R14	H	8,821
N3	C	174,675	P7	CB	31,328	R14	N	125,455
N3	CA	53,307	P7	HA	4,357	K16	C	173,392
N3	CB	39,182	Q8	C	176,718	K16	CA	55,657
N3	H	8,51	Q8	CA	58,186	K16	CB	38,045
N3	HA	4,752	Q8	CB	28,791	K16	H	8,126
N3	N	122,219	Q8	H	8,443	K16	HA	4,876
N3*	C	174,651	Q8	HA	4,134	K16	N	119,456
N3*	CA	53,292	Q8	N	120,606	C17	C	175,233
N3*	CB	39,175	D9	C	176,329	C17	CA	52,911
L4	C	176,35	D9	CA	56,579	C17	CB	44,143
L4	CA	55,041	D9	CB	39,398	C17	H	8,371
L4	CB	42,562	D9	H	9,018	C17	HA	5,907
L4	H	8,256	D9	HA	4,548	C17	N	116,477
L4	HA	4,384	D9	N	117,083	F18	C	174,565
L4	N	123,036	W10	C	175,746	F18	CA	57,53
L4*	C	176,213	W10	CA	57,557	F18	CB	43,431
L4*	CA	55,05	W10	CB	29,542	F18	H	10,127
L4*	CB	42,569	W10	H	8,488	F18	HA	6,076
L4*	H	8,297	W10	HA	5,155	F18	N	121,178
L4*	N	123,175	W10	N	122,043	H19	C	172,671
E5	C	175,868	L11	C	175,792	H19	CA	54,871
E5	CA	55,969	L11	CA	54,441	H19	CB	30,214
E5	CB	31,985	L11	CB	44,945	H19	H	8,44
E5	H	8,104	L11	H	9,631	H19	HA	4,609
E5	HA	4,398	L11	HA	4,818	H19	N	121,547
E5	N	119,536	L11	N	122,887	V20	C	174,821
E5*	C	175,812	L12	C	176,62	V20	CA	61,018

Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun
V20	CB	32,202	W27	CB	30,602	C34	N	114,563
V20	H	7,909	W27	H	7,9	G35	C	178,008
V20	HA	4,075	W27	N	122,024	G35	CA	47,786
V20	N	128,498	E28	C	180,108	G35	H	8,604
S21	C	173,86	E28	CA	59,998	G35	HA	4,028
S21	CA	59,4	E28	CB	29,408	G35	HA2	4,143
S21	CB	62,877	E28	H	9,014	G35	HA3	4,037
S21	H	8,764	E28	HA	3,773	G35	N	109,889
S21	HA	3,902	E28	N	115,061	R36	C	178,067
S21	N	124,221	E29	C	179,117	R36	CA	59,596
Q22	C	175,808	E29	CA	58,425	R36	CB	30,637
Q22	CA	55,687	E29	CB	28,121	R36	H	8,13
Q22	CB	29,984	E29	H	7,51	R36	HA	4,272
Q22	H	8,885	E29	HA	4,041	R36	N	123,348
Q22	HA	4,276	E29	N	121,046	K37	C	176,1
Q22	N	118,695	G30	C	173,665	K37	CA	55,914
V23	C	172,779	G30	CA	45,826	K37	CB	33,413
V23	CA	59,652	G30	H	8,084	K37	H	7,591
V23	CB	34,291	G30	HA2	2,059	K37	HA	4,47
V23	H	6,918	G30	HA3	1,248	K37	N	118,497
V23	HA	4,36	G30	N	112,436	G38	C	174,014
V23	N	112,868	Q31	C	178,907	G38	CA	46,779
S24	C	173,889	Q31	CA	59,018	G38	H	8,006
S24	CA	55,985	Q31	CB	27,427	G38	N	108,743
S24	CB	65,409	Q31	H	7,876	G38	QA	3,945
S24	H	8,18	Q31	N	120,41	A39	C	175,426
S24	HA	4,864	A32	C	180,312	A39	CA	50,308
S24	N	114,985	A32	CA	54,775	A39	CB	24,418
N25	C	175,23	A32	CB	17,945	A39	H	8,281
N25	CA	53,095	A32	H	7,888	A39	HA	5,127
N25	CB	41,588	A32	HA	4,034	A39	N	122,886
N25	H	9,567	A32	N	122,292	T40	C	174,723
N25	HA	5,101	D33	C	178,504	T40	CA	60,053
N25	N	122,586	D33	CA	58,398	T40	CB	73,356
T26	C	174,604	D33	CB	43,343	T40	H	8,637
T26	CA	61,837	D33	H	8,048	T40	HA	4,93
T26	CB	70,868	D33	N	120,752	T40	N	108,218
T26	H	9,861	C34	C	178,461	L41	C	177,601
T26	HA	5,558	C34	CA	55,951	L41	CA	56,732
T26	N	113,891	C34	CB	33,883	L41	CB	44,597
W27	C	177,91	C34	H	8,418	L41	H	8,608
W27	CA	62,699	C34	HA	4,524	L41	HA	5,014

Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun
L41	N	122,529	E48	N	121,754	D55	N	117,171
L42	C	174,495	E49	C	177,772	S56	C	174,371
L42	CA	56,842	E49	CA	59,734	S56	CA	60,726
L42	CB	44,43	E49	CB	30,597	S56	CB	63,686
L42	H	7,748	E49	H	8,366	S56	H	7,451
L42	HA	4,004	E49	HA	3,832	S56	HA	4,139
L42	N	117,694	E49	N	120,755	S56	N	114,406
L43	C	174,891	L50	C	177,494	I57	C	175,933
L43	CA	53,741	L50	CA	58,038	I57	CA	62,097
L43	CB	42,892	L50	CB	41,916	I57	CB	38,818
L43	H	8,098	L50	H	7,278	I57	H	7,485
L43	HA	4,476	L50	HA	3,986	I57	HA	3,876
L43	N	128,37	L50	N	118,262	I57	N	121,48
I44	C	175,744	R51	C	178,076	K58	C	177,296
I44	CA	60,659	R51	CA	59,471	K58	CA	57,822
I44	CB	38,509	R51	CB	29,752	K58	CB	31,794
I44	H	5,292	R51	H	8,221	K58	H	8,232
I44	HA	4,229	R51	HA	3,819	K58	HA	4,064
I44	N	118,783	R51	N	118,021	K58	N	123,355
Q45	C	175,477	F52	C	178,349	E59	C	176,109
Q45	CA	56,954	F52	CA	61,207	E59	CA	56,042
Q45	CB	31,087	F52	CB	38,388	E59	CB	28,803
Q45	H	9,568	F52	H	7,923	E59	H	7,88
Q45	HA	4,225	F52	HA	4,125	E59	N	119,903
Q45	N	127,243	F52	N	118,744	K60	C	176,254
D46	C	175,256	L53	C	178,646	K60	CA	56,99
D46	CA	53,455	L53	CA	58,004	K60	CB	32,522
D46	CB	42,825	L53	CB	41,614	K60	H	7,952
D46	H	7,589	L53	H	7,86	K60	HA	4,125
D46	HA	4,695	L53	HA	4,245	K60	N	120,295
D46	N	115,046	L53	N	122,113	Y61	C	175,16
Q47	C	177,708	L54	C	179,871	Y61	CA	57,052
Q47	CA	58,907	L54	CA	57,748	Y61	CB	37,76
Q47	CB	28,729	L54	CB	40,903	Y61	H	7,938
Q47	H	8,899	L54	H	8,425	Y61	HA	4,591
Q47	HA	4,014	L54	HA	4,647	Y61	N	116,132
Q47	N	119,161	L54	N	117,977	N62	C	172,312
E48	C	179,107	D55	C	177,892	N62	CA	52,391
E48	CA	59,938	D55	CA	56,317	N62	CB	41,126
E48	CB	28,791	D55	CB	40,826	N62	H	7,541
E48	H	8,49	D55	H	7,955	N62	HA	4,737
E48	HA	4,082	D55	HA	4,387	N62	N	118,792

Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun
S63	C	173,868	F70	CA	57,8	W77	HA	4,686
S63	CA	57,383	F70	CB	38,729	W77	N	123,005
S63	CB	64,304	F70	H	7,987	K78	C	177,328
S63	H	7,904	F70	HA	3,901	K78	CA	55,413
S63	HA	5,169	F70	N	127,298	K78	CB	36,746
S63	N	115,086	T71	C	172,342	K78	H	8,866
F64	C	176,448	T71	CA	61,107	K78	HA	5,013
F64	CA	55,705	T71	CB	71,26	K78	N	122,075
F64	CB	42,812	T71	H	7,869	W79	C	181,088
F64	H	8,548	T71	N	120,34	W79	CA	56,978
F64	HA	4,854	L72	C	175,843	W79	CB	31,779
F64	N	117,496	L72	CA	52,775	W79	H	9,04
W65	C	175,821	L72	CB	42,275	W79	HA	5,681
W65	CA	57,935	L72	H	8,359	W79	N	125,292
W65	CB	32,616	L72	HA	4,269	I80	C	175,817
W65	H	10,188	L72	N	124,235	I80	CA	63,79
W65	N	122,69	P73	C	176,389	I80	CB	38,213
I66	C	177,191	P73	CA	63,553	I80	H	8,594
I66	CA	59,272	P73	CB	34,296	I80	HA	4,376
I66	CB	41,948	P73	HA	4,405	I80	N	114,229
I66	H	8,248	D74	C	175,475	N81	C	176,258
I66	HA	5,329	D74	CA	56,441	N81	CA	52,568
I66	N	108,79	D74	CB	39,471	N81	CB	37,65
G67	C	172,947	D74	H	8,794	N81	H	7,421
G67	CA	49,663	D74	HA	4,149	N81	HA	4,938
G67	H	10,421	D74	N	122,699	N81	N	116,601
G67	HA2	4,658	M75	C	174,369	G82	C	174,558
G67	HA3	3,648	M75	CA	55,511	G82	CA	45,479
G67	N	112,059	M75	CB	28,437	G82	H	8,681
L68	C	173,488	M75	H	8,335	G82	HA	4,476
L68	CA	54,723	M75	HA	3,272	G82	N	108,313
L68	CB	46,475	M75	N	112,779	T83	C	174,28
L68	H	8,593	N76	C	174,104	T83	CA	62,17
L68	HA	4,873	N76	CA	51,104	T83	CB	70,669
L68	N	128,285	N76	CB	41,668	T83	H	7,647
R69	C	173,419	N76	H	6,747	T83	HA	4,58
R69	CA	54,735	N76	HA	4,843	T83	N	113,931
R69	CB	34,307	N76	N	113,2	T84	C	174,274
R69	H	8,331	W77	C	175,578	T84	CA	63,863
R69	HA	4,934	W77	CA	57,488	T84	CB	69,771
R69	N	122,424	W77	CB	29,649	T84	H	8,831
F70	C	175,47	W77	H	8,886	T84	HA	4,451

Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun
T84	N	119,556	K91	N	119,019	G99	C	173,742
F85	C	174,31	I92	C	176,633	G99	CA	45,896
F85	CA	58,632	I92	CA	60,893	G99	H	8,012
F85	CB	40,875	I92	CB	38,43	G99	N	108,497
F85	H	8,486	I92	H	8,47	G99	QA	3,939
F85	HA	4,65	I92	N	123,172	S100	C	173,185
F85	N	126,472	T93	C	173,894	S100	CA	59,312
N86	C	174,193	T93	CA	61,559	S100	CB	65,119
N86	CA	52,293	T93	CB	70,095	S100	H	7,832
N86	CB	39,62	T93	H	8,562	S100	HA	4,695
N86	H	8,389	T93	HA	4,469	S100	N	114,206
N86	HA	4,782	T93	N	121,481	C101	C	172,035
N86	N	127,109	G94	C	173,264	C101	CA	59,557
S87	C	174,727	G94	CA	44,765	C101	CB	45,825
S87	CA	57,194	G94	H	8,088	C101	H	8,802
S87	CB	63,879	G94	H2	4,204	C101	N	121,435
S87	H	7,781	G94	H3	3,966	A102	C	175,058
S87	HA	4,029	G94	N	110,419	A102	CA	52,049
S87	N	116,772	V95	C	176,199	A102	CB	21,876
D88	C	176,965	V95	CA	61,969	A102	H	8,025
D88	CA	55,552	V95	CB	33,249	A102	HA	5,187
D88	CB	40,711	V95	H	8,112	A102	N	122,583
D88	H	8,44	V95	HA	4,248	S103	C	173,552
D88	HA	4,597	V95	N	118,208	S103	CA	55,293
D88	N	121,6	T96	C	174,213	S103	CB	67,147
V89	C	175,319	T96	CA	62,49	S103	H	8,672
V89	CA	63,905	T96	CB	69,726	S103	HA	5,679
V89	CB	33,196	T96	H	8,183	S103	N	115,836
V89	H	7,95	T96	HA	4,188	I104	C	174,076
V89	HA	4,056	T96	N	118,868	I104	CA	59,287
V89	N	117,479	E97	C	176,012	I104	CB	41,644
L90	C	174,767	E97	CA	56,703	I104	H	9,299
L90	CA	53,742	E97	CB	30,52	I104	HA	5,263
L90	CB	44,121	E97	H	8,222	I104	N	116,355
L90	H	7,756	E97	HA	4,216	L105	C	177,917
L90	HA	4,631	E97	N	122,496	L105	CA	54,668
L90	N	119,662	N98	C	175,467	L105	CB	42,952
K91	C	175,915	N98	CA	53,933	L105	H	8,185
K91	CA	54,973	N98	CB	39,055	L105	HA	4,659
K91	CB	34,142	N98	H	8,248	L105	N	126,882
K91	H	8,046	N98	HA	4,571	G106	C	172,028
K91	HA	4,635	N98	N	117,838	G106	CA	47,69



Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun
G106	H	9,242	S113	HA	3,987	W120	HA	5,708
G106	HA2	3,987	S113	N	115,051	W120	N	114,473
G106	HA3	3,547	C114	C	174,201	I121	C	176,039
G106	N	119,099	C114	CA	58,792	I121	CA	59,605
D107	C	175,088	C114	CB	44,512	I121	CB	41,43
D107	CA	52,598	C114	H	9,109	I121	H	7,156
D107	CB	40,525	C114	HA	4,274	I121	HA	5,308
D107	H	8,016	C114	N	126,816	I121	N	116,829
D107	HA	4,592	A115	C	177,386	C122	C	174,727
D107	N	122,345	A115	CA	52,117	C122	CA	51,662
K108	C	175,766	A115	CB	18,826	C122	CB	40,245
K108	CA	54,958	A115	H	7,931	C122	H	9,471
K108	CB	34,651	A115	HA	4,678	C122	HA	5,982
K108	H	7,822	A115	N	120,485	C122	N	121,86
K108	HA	4,736	S116	C	170,353	Q123	C	173,824
K108	N	118,255	S116	CA	59,549	Q123	CA	53,736
V109	C	175,347	S116	CB	63,777	Q123	CB	35,048
V109	CA	61,379	S116	H	7,983	Q123	H	9,987
V109	CB	32,133	S116	HA	4,466	Q123	HA	6,139
V109	H	8,606	S116	N	116,411	Q123	N	117,554
V109	HA	4,803	D117	C	175,29	K124	C	174,832
V109	N	125,057	D117	CA	53,243	K124	CA	55,359
T110	C	172,323	D117	CB	42,031	K124	CB	37,266
T110	CA	58,523	D117	H	7,962	K124	H	9,074
T110	CB	71,429	D117	HA	4,912	K124	HA	4,851
T110	H	9,294	D117	N	118,856	K124	N	121,394
T110	HA	4,934	N118	C	176,331	E125	C	176,336
T110	N	123,068	N118	CA	52,457	E125	CA	56,467
P111	C	176,123	N118	CB	42,175	E125	CB	30,761
P111	CA	61,903	N118	H	8,752	E125	H	8,735
P111	CB	32,221	N118	HA	4,903	E125	HA	4,506
P111	HA	4,897	N118	N	120,454	E125	N	122,845
E112	C	175,692	R119	C	174,272	L126	C	176,643
E112	CA	55,594	R119	CA	56,769	L126	CA	55,332
E112	CB	35,347	R119	CB	32,601	L126	CB	42,394
E112	H	9,213	R119	H	7,496	L126	H	8,353
E112	HA	4,796	R119	HA	4,542	L126	HA	4,261
E112	N	123,413	R119	N	118,894	L126	N	123,657
S113	C	178,303	W120	C	172,682	N127	C	174,662
S113	CA	58,436	W120	CA	57,472	N127	CA	53,226
S113	CB	63,49	W120	CB	33,266	N127	CB	39,147
S113	H	8,837	W120	H	8,642	N127	H	8,424

Aminokyselinový zbytek	Atom	Chemický posun	Aminokyselinový zbytek	Atom	Chemický posun
N127	N	118,056	S135	CA	60,171
H128	C	174,563	S135	CB	65,049
H128	CA	55,513	S135	H	7,859
H128	CB	29,934	S135	N	121,067
H128	H	8,388			
H128	HA	4,707			
H128	N	119,641			
E129	C	176,284			
E129	CA	56,49			
E129	CB	30,498			
E129	H	8,436			
E129	HA	4,351			
E129	N	122,511			
T130	C	172,829			
T130	CA	59,88			
T130	CB	69,913			
T130	H	8,36			
T130	HA	4,607			
T130	N	118,602			
P131	C	176,94			
P131	CA	63,264			
P131	CB	32,225			
P131	HA	4,484			
S132	C	174,458			
S132	CA	58,261			
S132	CB	64,034			
S132	H	8,459			
S132	HA	4,467			
S132	N	116,438			
N133	C	174,853			
N133	CA	53,435			
N133	CB	39,207			
N133	H	8,504			
N133	HA	4,81			
N133	N	120,964			
D134	C	175,375			
D134	CA	54,553			
D134	CB	41,301			
D134	H	8,342			
D134	HA	4,696			
D134	N	121,128			
S135	C	178,652			

Svoluji k zapůjčení této práce pro studijní účely a prosím, aby byla řádně vedena evidence vypůjčovatelů.

Jméno a příjmení S adresou	Číslo OP	Datum vypůjčení	Poznámka