

Jihočeská univerzita v Českých Budějovicích
Přírodovědecká fakulta

Strukturní studie halogenaciddehalogenasy Tt81
z *Thermococcus thio-reducens*

Diplomová práce

Bc. Kristýna Rejžková

Školitel: Doc. Mgr. Ivana Kutá Smatanová, Ph.D.
Konzultant: Mgr. Tatyana Prudnikova, Ph.D.

České Budějovice 2019

Rejžková, K., 2019: Strukturní studie halogenaciddehalogenasy Tt81 z *Thermococcus thio-reducens*. [Structural studies of haloacid dehalogenase Tt81 from *Thermococcus thio-reducens*. Mgr. Thesis, in Czech.] – 46 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation

This master thesis is focused on the structural study of newly discovered protein Tt81 from *Thermococcus thio-reducens*. The main aim of this work is to create a model of the Tt81 protein structure based on the diffraction data obtained by X-ray diffraction of Tt81 protein crystals and confirm that protein Tt81 belongs to the halogenacid-dehalogenase superfamily.

Prohlášení

Prohlašuji, že svoji diplomovou práci jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své diplomové práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích dne 17. 4. 2019

.....

Bc. Kristýna Rejžková

Poděkování

Děkuji své školitelce Doc. Mgr. Ivaně Kuté Smatanové, Ph.D. za odborné vedení, cenné rady a připomínky. Dále bych chtěla poděkovat své konzultantce Mgr. Tatyňce Prudnikové, Ph.D. za ochotu a trpělivost při vedení praktickou částí práce a Dr. Jeroenovi R. Mestersovi za seznámení s počítačovými programy pro řešení proteinových struktur.

Poděkování patří také mé rodině a přátelům za podporu během celého studia.

OBSAH

| | |
|--|----|
| 1. Úvod..... | 1 |
| 2. Literární přehled..... | 3 |
| 2.1 Struktura krystalu..... | 3 |
| 2.2 Rentgenová difrakční analýza..... | 6 |
| 2.2.1 Historické milníky..... | 6 |
| 2.3 Princip difrakce..... | 6 |
| 2.4 Rentgenový difrakční experiment..... | 8 |
| 2.5 Zpracování difrakčních dat..... | 10 |
| 2.6 Fázový problém a jeho řešení..... | 11 |
| 2.7 Molekulové modelování a upřesnění struktury..... | 13 |
| 2.8 Validace struktury..... | 14 |
| 2.9 Měřítka kvality získané struktury..... | 15 |
| 2.9.1 R-faktor..... | 15 |
| 2.9.2 Free R-faktor..... | 16 |
| 2.9.3 Komplettnost dat..... | 16 |
| 2.9.4 Rozlišení..... | 16 |
| 2.9.5 Ramachandranův diagram..... | 16 |
| 2.9.6 B-faktory..... | 17 |
| 3. Materiál a metody..... | 18 |
| 3.1 Sběr difrakčních dat..... | 18 |
| 3.2 Zpracování difrakčních dat..... | 18 |
| 3.3 Řešení fázového problému..... | 19 |

| | |
|--|----|
| 3.3.1 Srovnání sekvencí..... | 19 |
| 3.3.2 Molekulární nahrazení..... | 19 |
| 3.4 Upřesnění struktury..... | 20 |
| 3.5 Validace struktury..... | 21 |
| 3.6 Zobrazení struktury..... | 22 |
| 3.6.1 Zobrazení sekvenčních motivů..... | 22 |
| 4. Výsledky..... | 24 |
| 4.1 Výsledky rentgenové difrakční analýzy..... | 24 |
| 4.2 Výsledky zpracování difrakčních dat..... | 24 |
| 4.3 Výsledky řešení fázového problému..... | 27 |
| 4.3.1 Výsledky srovnání sekvencí..... | 27 |
| 4.3.2 Výsledky molekulárního nahrazení..... | 27 |
| 4.4 Výsledky upřesnění struktury..... | 28 |
| 4.5 Výsledky validace struktury..... | 29 |
| 4.6 Výsledky zobrazení struktury..... | 30 |
| 4.6.1 Výsledky zobrazení sekvenčních motivů..... | 32 |
| 5. Diskuze..... | 34 |
| 6. Závěr..... | 38 |
| 7. Seznam literatury..... | 39 |

1. Úvod

Enzymy hrají v životě člověka nezastupitelnou roli. Nejenže řídí většinu biochemických procesů probíhajících v tělech všech živých organismů, ale jsou užitečné i v mnohých jiných oblastech lidského života, např. v průmyslu, nebo v environmentálních technologiích. Zajímavou nadrodinu enzymů tvoří tzv. halogenaciddehalogenasy (HAD), které se podílí na detoxikaci xenobiotik nebo vedlejších metabolických produktů. HAD využívají nukleofilní substituci pro vytvoření enzym-substrátové esterové vazby za současného štěpení vazby uhlík-halogen. Velkou pozornost přitahují tyto enzymy právě pro svůj potenciál degradovat rezistentní halogenované organické sloučeniny, které tvoří největší skupinu látek znečišťujících životní prostředí, ale také pro své využití při syntéze průmyslových chemikálií (Koonin a Tatusov, 1994; Soda a kol., 1996; Schmidberger a kol., 2007).

Abychom však mohli posoudit využití daného enzymu (proteinu), je nutné nejdříve poznat jeho třidimenzionální strukturu. Mocným nástrojem poznání atomární struktury biomakromolekul je rentgenová krystalografie. Ke dni 4. 4. 2019 obsahovala PDB databáze celkem 139 678 proteinových struktur, z nichž bylo 90 % vyřešeno právě touto metodou (RCSB PDB, 2019). Základem úspěchu rentgenové krystalografie je vypěstování dostatečně kvalitních krystalů, které poskytnou difrakční data pro řešení struktury studovaného proteinu.

Začátek teoretické části této práce je věnován stručnému uvedení do struktury proteinového krystalu, od elementární buňky a asymetrické jednotky přes krystalové mřížky až k prostorovým grupám. Následuje úvod do rentgenové difrakční analýzy, vysvětlení základních pojmů, důležité historické milníky, které vedly k řešení nejen proteinových struktur, vysvětlení principu difrakce a seznámení s průběhem rentgenového difrakčního experimentu. Zbytek teoretické části se zabývá zpracováním difrakčních dat, řešením fázového problému, molekulovým modelováním, upřesněním a validací struktury se zaměřením na počítačové programy, které k těmto účelům slouží. Podrobně jsou uvedeny počítačové programy, které byly použity při vypracování praktické části, ale jsou jmenovány i další programy s podobnými funkcemi. Na závěr teoretické části je popsáno několik měřítek kvality získané struktury.

Praktická část práce se zabývá strukturální studií putativní halogenaciddehalogenasy Tt81, která byla izolována z *Thermococcus thio-reducens*. Nejprve je popsán sběr difrakčních dat, následuje jejich zpracování a řešení fázového problému metodou molekulárního nahrazení. Dále je uvedeno molekulární modelování, upřesnění struktury, její validace a nakonec zobrazení finálního modelu proteinu Tt81.

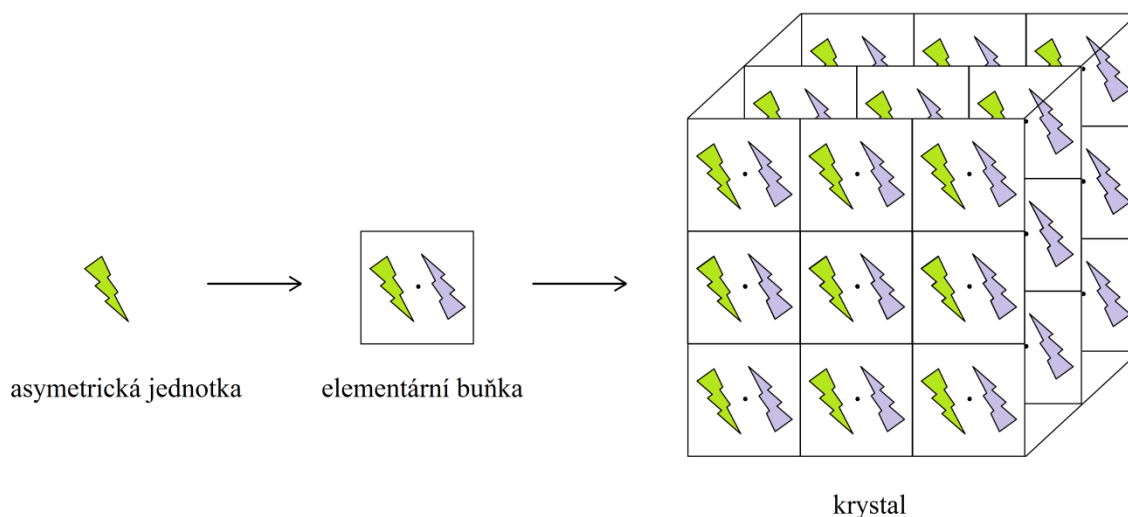
Cílem práce bylo:

- zpracovat difrakční data získaná z krystalů proteinu Tt81 z *Thermococcus thio-reducens*
- vyřešit strukturu studovaného proteinu
- upřesnit strukturu halogenaciddehalogenasy Tt81
- připravit strukturu pro deponování do PDB
- porovnat strukturu Tt81 s podobnými proteinovými strukturami

2. Literární přehled

2.1 Struktura krystalu

Krystal je definován jako pravidelné trojrozměrné uspořádání molekul, které jsou dohromady drženy slabými vazebnými interakcemi. V proteinových krystalech jsou navíc molekuly poměrně flexibilní (Rhodes, 2006). Pro snazší popis struktury a symetrie krystalů se obvykle zavádí pojem ideální krystal. Ideální krystal je nekonečný, jeho struktura je zcela pravidelná a neobsahuje žádné poruchy. Základním rysem pravidelnosti ideálního krystalu je jeho trojrozměrná translační periodičita. Charakteristické je opakování základního motivu tvořeného jedním nebo více atomy ve třech směrech napříč celým krystalem. U organických krystalů může motiv obsahovat až tisíce atomů. Reálné krystaly se ke struktuře ideálních krystalů mohou jen přibližovat (Valvoda a kol., 1992).



Obr. 1: Uspořádání asymetrické jednotky do elementární buňky aplikací inverze (černý bod = střed symetrie) a krystal složený z identických elementárních buněk (upraveno podle Rupp, 2010).

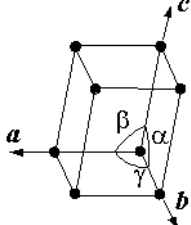
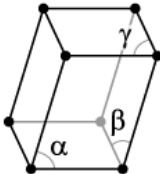
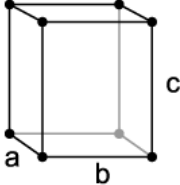
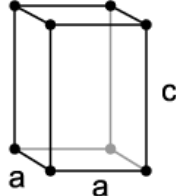
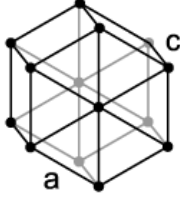
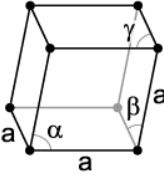
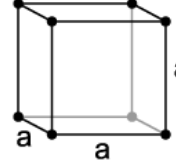
Základní stavební jednotkou krystalu je elementární buňka. Jedná se o nejmenší a nejjednodušší objemový prvek, pravidelně se opakující ve všech směrech, který reprezentuje celý krystal. Elementární buňka je definována rozměry hran (a , b , c) a úhly mezi nimi (α , β , γ). Kromě elementární buňky se také uvádí pojem asymetrická jednotka, což je nejmenší jednotka struktury, která může tvořit celý krystal, ale až po aplikaci určité operace symetrie (Rhodes, 2006), viz Obr. 1. Operací symetrie rozumíme geometrickou

transformaci, při níž jsou zachovány vzájemné vzdálenosti v tělese a po jejímž provedení nelze poznat, zda byla s tělesem nějaká transformace provedena. Mezi základní operace symetrie patří rotace, translace, inverze a zrcadlení (Vávra a Losos, 2006).

Atomy v krystalu jsou uspořádané do krystalové mřížky, složené z identických elementárních buněk. Mřížka je definována jako množina bodů se stejným a stejně orientovaným okolím. Jedná se o myšlenou mřížku, jejíž mřížové body mohou a nemusí koincidovat se středy atomů. Mřížka vystihuje translační periodicitu prostorového uspořádání atomů v krystalu a slouží k popisu struktury, například pomocí umístění atomů v jedné buňce mřížky (Valvoda a kol., 1992). Existuje sedm krystalových mřížek (soustav) – trojklonná (triklinická), jednoklonná (monoklinická), kosočtverečná (ortorombická), čtverečná (tetragonální), krychlová (kubická), šesterečná (hexagonální) a klencová (trigonální) - viz Tab. 1. Jednotlivé krystaly jsou řazeny do krystalových soustav podle společných prvků symetrie. Každá prostorová mřížka, která obsahuje střed symetrie, se označuje jako centrosymetrická a krystal s touto mřížkou jako centrosymetrický (Vávra a Losos, 2006).

K popisu struktury ideálních krystalů slouží kromě krystalové mřížky také prostorové grupy. Prostorová grupa obsahuje všechny informace o symetrii krystalu. Zahrnuje symetrii atomové struktury i makroskopickou symetrii krystalu (Valvoda a kol., 1992). Celkem je možné odvodit 230 prostorových grup, které zahrnují všechny kombinace translačních a beztranslačních prvků symetrie, přípustných v krystalografických transformacích. Prvky symetrie libovolné prostorové grupy mají v prostoru zcela určitou polohu a orientaci (Vávra a Losos, 2006). Prostorové grupy jsou označovány pomocí symbolů ze čtyř znaků. Prvním znakem je písmeno, které označuje typ mříže. Následují tři znaky označující prvky symetrie, které byly kombinovány s translacemi mříže při vytváření prostorové grupy (Valvoda a kol., 1992). Pokud známe prostorovou grupu, elementární buňku a obsah asymetrické jednotky můžeme definovat polohy všech atomů v krystalu (Rhodes, 2006).

Tab. 1: Krystalové soustavy (Upraveno podle Vávra a Losos, 2006).

| Krystalová mřížka | | Mřížkové parametry |
|-------------------|---|---|
| jednoklonná |  | $a \neq b \neq c$ $\alpha = \gamma, \beta > 90^\circ$ |
| trojklonná |  | $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90^\circ$ |
| kosočtverečná |  | $a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$ |
| čtverečná |  | $a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$ |
| šesterečná |  | $a = b \neq c$ $\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$ |
| klencová |  | $a = b = c$ $\alpha = \beta = \gamma \neq 90^\circ$ |
| krychlová |  | $a = b = c$ $\alpha = \beta = \gamma = 90^\circ$ |

2.2 Rentgenová difrakční analýza

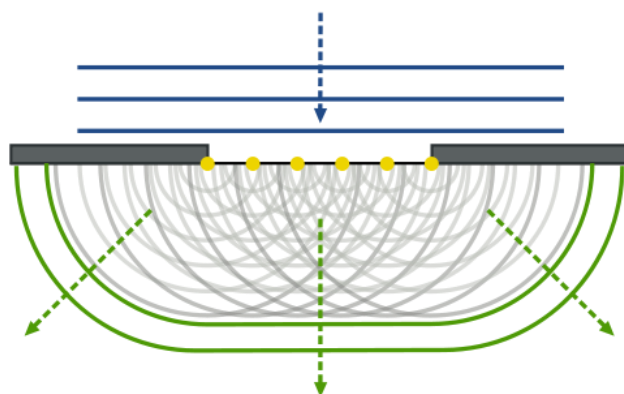
Rentgenová difrakční analýza je metoda, která slouží k určování struktury krystalických látek. Podstatou této metody je interakce elektromagnetického (rentgenového) záření s periodicky uspořádanou hmotou, respektive s elektrony atomů krystalů. Díky tomu, že rozměry krystalové mřížky jsou přibližně stejné jako vlnová délka rentgenového záření (zhruba 0,1 nm), může na krystalové mřížce docházet k difrakci neboli ohybu záření (Židek, 2015).

2.2.1 Historické milníky

Důležitým milníkem pro určování struktur látek byl rok 1895, kdy Wilhelm Conrad Röntgen objevil rentgenové záření (Rupp, 2010). Na objevu samotné rentgenové difrakce r. 1912 se podíleli Max von Laue, Walter Friedrich a Paul Knipping (Eckert, 2012). Známa skutečnost, že se rentgenové záření chová jako elektromagnetické vlnění, a že po průchodu fotonů vzorkem dochází na detektoru k zachycení difrakčního obrazce, nevedla ihned k řešení struktur. Souvislost mezi difrakčním obrazcem a vnitřní strukturou krystalů pochopili až William Henry Bragg a William Lawrence Bragg (otec a syn), kterým se v letech 1912-1922 podařilo určit strukturu řady látek. Jednalo se o struktury několika prvků (diamantu, grafitu, mědi), jednoduchých anorganických látek (např. chloridu sodného, chloridu draselného, pyritu,...) a první organické látky (naftalenu). Jejich objevy byly oceněny Nobelovou cenou za fyziku r. 1915 (Kratochvíl a kol., 2008).

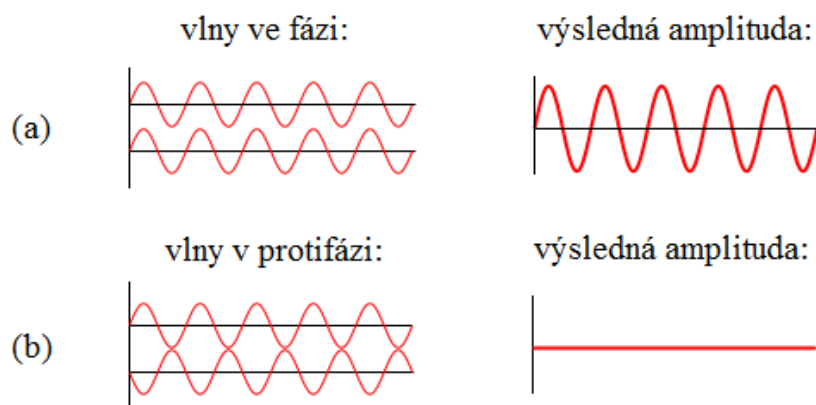
2.3 Princip difrakce

Difrakce je děj, ke kterému dochází, pokud vlnění (ať už vlnění na vodní hladině, vlnění zvukové nebo elektromagnetické – tzn. viditelné světlo, rentgenové záření či radiové vlny) narazí na překážku nebo štěrbinu. V takovém případě nastane ohyb vlnění do oblasti geometrického stínu překážky. Tento jev je popisován pomocí Huygensova principu, kdy v každém okamžiku můžeme každý bod šířící se vlny chápat jako nový zdroj vlnění, tzv. sekundárních vln (viz Obr. 2). Účinek difrakce je největší v případě, kdy se vlnová délka přibližně rovná rozměru překážky popř. štěrbinu (Rappaport, 2002).



Obr. 2: Difrakce podle Huygensova principu. (Převzato z Nordmann, 2007 [online]. Dostupné 6. 9. 2018 z WWW: <https://cs.wikipedia.org/wiki/Huygens%20AFv_princip#/media/File:Refraction_on_an_aperture_-_Huygens-Fresnel_principle.svg>.)

V momentě, kdy vlnění narazí na více překážek či štěrbin, dochází ke konstruktivnímu nebo destruktivnímu skládání sekundárních vln podle principu superpozice. Sekundární vlny mají sice stejnou vlnovou délku, ale v závislosti na poloze překážek jsou vůči sobě různě fázově posunuté. Proto v některých případech dochází k zesílení intenzity záření (konstruktivní skládání) a někdy se vlny navzájem vyruší (destruktivní skládání) (Žídek, 2015). Záleží na tom, zda se vlny nacházejí ve fázi, v protifázi, nebo někde mezi těmito dvěma stavy (viz Obr. 3).



Obr. 3: Konstruktivní (a) a destruktivní (b) skládání vln. (Upraveno podle Haade a Quibik, 2010 [online]. Dostupné 6. 9. 2018 z WWW: <https://en.wikipedia.org/wiki/Wave_interference#/media/File:Interference_of_two_waves.svg>.)

Pokud jsou odražené paprsky ve fázi, jejich amplitudy se sečtou a společně pak přispívají k tvorbě difrakčního obrazce ve formě symetricky uspořádaných skvrn (reflexí),

které jsou zachyceny na detektoru. Tento jev se popisuje pomocí Braggova zákona (Bragg a Bragg, 1913):

$$2d \sin \theta = n\lambda$$

kde d je vzdálenost mezi atomovými vrstvami krystalu, θ je úhel dopadu rentgenového paprsku, n je kladné celé číslo a λ je vlnová délka dopadajícího rentgenového paprsku (Rupp, 2010).

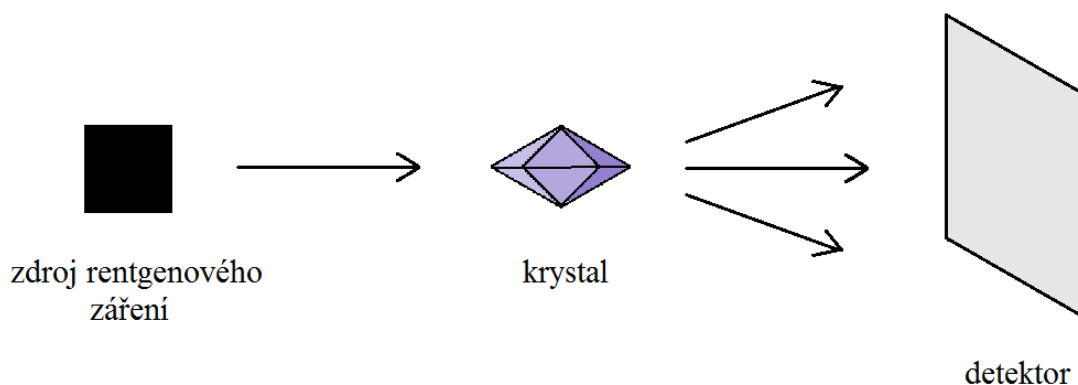
Reflexe jsou charakterizovány strukturním faktorem $F(hkl)$:

$$F(hkl) = V \int_{x=0}^1 \int_{y=0}^1 \int_{z=0}^1 \rho(xyz) e^{2\pi i(hx+ky+lz)} dx dy dz \quad (\text{Žídek, 2015})$$

kde h , k a l jsou tzv. Millerovy indexy, V je objem elementární buňky analyzovaného krystalu a ρ je elektronová hustota v bodě x , y , z v rámci elementární buňky. Millerovy indexy definují rovinu atomů v krystalu podle jejich průsečíků se třemi základními krystalografickými osami, protože se jedná o trojrozměrné těleso (Žídek, 2015). Strukturní faktor je komplexní číslo, které představuje celkový rozptyl ze všech atomů elementární buňky a zahrnuje amplitudu strukturního faktoru a relativní fázový úhel mezi odraženými vlnami (Rupp, 2010). Při určování struktury molekul z rentgenové difrakce se však nesnažíme spočítat hodnotu strukturního faktoru, ale distribuci elektronové hustoty, protože právě rozložení elektronů vypovídá o tvaru molekuly, který hledáme (Žídek, 2015).

2.4 Rentgenový difrakční experiment

Schéma rentgenového difrakčního experimentu je znázorněno na Obr. 4. Výchozím materiálem pro rentgenový difrakční experiment je monokrystal studovaného proteinu. Získání kvalitního monokrystalu je zcela limitujícím krokem celé proteinové krystalografie, protože nalezení krystalizačních podmínek může trvat i několik let a ne vždy se daný protein podaří vykrystalizovat. Všechny tři rozměry krystalu by měly být minimálně 0,01 mm a krystal by měl být bez prasklin a defektů (Kratochvíl a kol., 2008).



Obr. 4: Schéma rentgenového difrakčního experimentu.

Jako experimentální zařízení se používají různé typy difraktometrů – čtyřkruhový difraktometr, vícekruhové monokrystalové difraktometry a generátory pro sběr difrakčních dat. Některé mohou být připojeny na zdroje synchrotronového záření v rentgenové oblasti. To je pro proteinovou krystalografii výhodné, protože intenzivní a koncentrovaný synchrotronový rentgenový svazek vybudí mnohem více měřitelných difrakcí než laboratorní generátor a data jsou pak kompletnější (Kratochvíl a kol., 2008). Za zdroji synchrotronového záření musíme do zahraničí – v Evropě je jich celkem 14. Z toho čtyři jsou v Německu (BESSY v Berlíně, EMBLDESY, MPGDESY a PETRA_III v Hamburku), ostatní se nacházejí ve Španělsku, Itálii, Francii, UK, Švýcarsku, Švédsku a Rusku (BioSync, 2018).

Během experimentu je krystal vystaven paprsku rentgenového záření a jednotlivé difrakce jsou zachyceny na detektoru jako difrakční obrazec. K posbírání všech nezbytných dat pro vyřešení struktury musí být krystal otáčen o 180° , u krystalů s vyšší symetrií může být použit úhlový rozsah 90° nebo 45° (McPherson, 2002). Počet difrakcí nutných pro vyřešení struktury také souvisí se složitostí krystalu – několik desítek difrakcí u krystalů kovů až několik milionů difrakcí u krystalů virů. Doba sběru dat se pohybuje od několika hodin po několik dní (Kratochvíl a kol., 2008).

Rentgenové záření je vysoce energetické elektromagnetické vlnění, které způsobuje radiační poškození organické hmoty a tedy i proteinových krystalů (Rupp, 2010). Z toho důvodu se měření běžně provádí za nízké teploty ($100\text{--}150\text{ K}$) v tekutém dusíku. Krystal je chráněn pomocí kryoprotektantu. Nízká teplota slouží jako ochrana před radiačním poškozením a navíc potlačuje teplotně-vibrační pohyby atomů (Kratochvíl a kol., 2008).

Výsledkem experimentu jsou difrakční data v podobě jednotlivých difrakčních obrazců. Difrakční obrazec neodráží přímo vnitřní strukturu krystalu, ale je možné z něj zjistit intenzity a polohy jednotlivých difrakcí. Ze známých intenzit je možné odvodit pozice atomů a jejich teplotně vibrační parametry, z difrakčních úhlů pak rozměry elementární buňky. Některé difrakce navíc systematicky vyhasínají (mají nulovou intenzitu), z čehož lze odvodit prostorovou grupu (Kratochvíl a kol., 2008).

2.5 Zpracování difrakčních dat

Po naměření difrakčních dat následuje jejich zpracování. V současné době je možné zpracovávat data více či méně automaticky, s malým uživatelským vstupem a současně generovat určitou vizuální zpětnou vazbu o průběhu zpracování a výsledné kvalitě dat pro posouzení uživatelem. K tomuto účelu slouží např. XDSAPP ~ XDS Automation and Plotting Protocols (Sparta a kol., 2016), který kombinuje softwarové balíčky XDS (Kabsch, 1993), POINTLESS z CCP4 (Winn a kol., 2011), XDSSTAT (Diederichs, 2006), SFCHECK (Vagin a kol., 1999) a PHENIX.XTRIAGE (Zwart a kol., 2005). XDS je sofistikovaný software, který byl vyvinut pro zpracování monochromatických difrakčních dat zaznamenaných metodou rotace za použití jednoduchého krystalu (Kabsch, 1993). Dalšími dostupnými softwary pro integraci dat kromě XDS jsou iMOSFLM (Battye a kol., 2011), HKL-2000 (Otwinowski a Minor, 1997) a D^{*}TREK (Pflugrath, 1999). Tyto softwary se liší metodami autoindexace, metodami integrace a úrovní automatizace (Krug a kol., 2012).

XDSAPP (Sparta a kol., 2016) umožňuje dva režimy provozu. Prvním režimem je verze GUI vhodná pro zpracování jediné datové sady, která současně umožňuje různou míru kontroly. Druhou možností je automatické zpracování všech datových sad obsažených v adresáři bez interakcí uživatele po spuštění zpracování (Sparta a kol., 2016).

Zpracování difrakčních dat softwarem XDSAPP (Sparta a kol., 2016) zahrnuje tři kroky. Prvním krokem je indexace, která poskytuje orientaci krystalu, počáteční odhad rozměrů elementární buňky, rozlišení signálu a pozadí, určení intenzity skvrn oproti pozadí a určení počtu a tvaru užitečných skvrn (Powell, 2017). Druhým krokem je integrace, během které je určeno maximální rozlišení a POINTLESS (Winn a kol., 2011) vyhodnotí nejpravděpodobnější prostorovou grupu. Po skončení indexace i integrace je zobrazeno krátké shrnutí pro kontrolu uživatelem. Posledním krokem je analýza, kdy je analyzována kvalita dat, SFCHECK (Vagin a kol., 1999) a PHENIX.XTRIAGE (Zwart a kol., 2005)

provádí testy na případné krystalografické problémy (např. dvojčatné krystaly, pseudosymetrie) a dochází ke konverzi difrakčních dat do souborů intenzit v různých formátech (mtz, hkl, cns) vhodných pro CCP4 (Winn a kol., 2011), SHELXL (Sheldrick, 2015) a CNS (Brünger a kol., 1997). Kromě toho software sám identifikuje snímky s horší kvalitou dat a poskytuje informace o radiačním poškození krystalu během procesu sběru dat (Diederichs, 2006; Krug a kol., 2012).

Nevýhodou XDSAPP je poměrně těžkopádné použití v manuálním režimu. Významnou překážku pro méně zkušené uživatele představuje zejména manipulace s dlouhými vstupními a výstupními textovými soubory, což může být potencialem zdrojem případných chyb (Sparta a kol., 2016).

Po zpracování difrakčních dat následuje získání mapy elektronové hustoty, tedy distribuce atomů v prostoru, vypočtením funkce elektronové hustoty. Vztah mezi funkcí elektronové hustoty a strukturním faktorem bývá vyjádřen prostřednictvím Fourierovy transformace, kdy se strukturní faktory používají jako Fourierovy koeficienty pro generování mapy elektronové hustoty proteinu (Rupp, 2010):

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l F(hkl) e^{-2\pi i(hx+ky+lz)+i\alpha(hkl)} \quad (\text{Židek, 2015})$$

kde α značí fázi strukturního faktoru. Aby bylo možné vypočítat mapu distribuce elektronové hustoty, musíme znát kromě intenzit a úhlů difrakcí i údaje o fázových úhlech (fázích) difraktovaných paprsků. Tyto informace jsou však během rentgenového difrakčního experimentu ztraceny. Neznalost fází se v krystalografii označuje jako fázový problém (Kratohvíl a kol., 2008).

2.6 Fázový problém a jeho řešení

Vyřešení fázového problému je nutným krokem na cestě za vyřešením struktury krystalické látky. Allegra popisuje existenci fázového problému v rentgenové krystalografii jako fascinující dobrodružství, ve kterém se strukturní krystalograf neobejde bez chemické a fyzikální intuice, představivosti a matematických znalostí (1980).

V dnešní době existuje celá řada metod sloužících k řešení fázového problému. Tyto metody jsou založeny na skutečnosti, že neexistuje žádný formální vztah mezi amplitudami

a fázemi, jen mezi molekulární strukturou nebo hustotou elektronů a fázemi (Taylor, 2003). Taylor tyto metody rozděluje do tří skupin: přímé metody, molekulární nahrazení a izomorfní záměna (2003).

Přímé metody jsou matematické techniky, které se snaží vyřešit fázový problém pomocí matematických vztahů bez použití strukturně-chemických informací (Allegra, 1980). Jsou založené na pozitivitě a atomicitě elektronové hustoty, což vede k fázovým vztahům mezi normalizovanými strukturními faktory. Pokud jsou známy fáze některých difrakcí, je možné postupně odvodit další a další fáze až do vyřešení všech fází. Tyto metody jsou limitovány rozlišením difrakčních dat, které musí být nižší než 1,2 Å, a používají se u proteinů obsahujících do 1000 atomů (Taylor, 2003).

Pokud je dostupný homologní model s alespoň 25% sekvenční identitou, je možné použít metodu molekulárního nahrazení (Taylor, 2003). Vzhledem k rostoucímu počtu strukturních modelů v PDB databázi je tato metoda využívána nejčastěji. V dnešní době je tímto způsobem vyřešeno přibližně 80 % všech proteinových struktur. Termín nahrazení je třeba chápat jako přemístění, nikoliv substituci. Známý strukturní model se rotuje a následně translataje do elementární buňky nebo asymetrické jednotky, dokud není získáno řešení s nejlepším vztahem mezi vypočtenými difrakčními daty známého modelu a zjištěnými daty z neznámé struktury (Rossmann, 1972; Rupp, 2010). Jinými slovy, na základě dat dochází k testování různých hypotéz o orientaci, pozici a kvalitě vyhledávaného modelu (Read, 2003). Molekulární nahrazení se skládá ze tří hlavních kroků: rotace, translace a hodnocení kvality řešení. První dva kroky jsou prováděny na základě výpočtu Pattersonovy funkce. Pozice modelu se mění, dokud se vypočtená Pattersonova rotační a translační funkce neshoduje s pozorovanou Pattersonovou funkcí, odpovídající neznámé struktuře (Navaza, 1994). Homologní model, správně umístěný v neznámé struktuře, pak slouží jako počáteční model (Rupp, 2010).

Další možností, jak získat fáze difraktovaných paprsků, je metoda izomorfní záměny. Tato metoda vyžaduje přípravu derivátů studované chemické sloučeniny, ve kterých je několik málo atomů ve známých polohách nahrazeno jinými atomy. Většinou dochází k nahrazení lehkých atomů za těžší atomy kovů, tak aby nebyla narušena krystalizace. Takové deriváty označujeme za izomorfní, protože jsou strukturně stejné a liší se pouze druhem atomů v několika málo polohách (Valvoda a kol., 1992). Použití jediného derivátu u necentrosymetrických krystalů vede k dvojznačnému výsledku fázového úhlu (Bokhoven

a kol., 1951). Z toho důvodu se používá série sloučenin s izomorfní záměnou v různých polohách (Blow a Rossmann, 1961). Po zavedení atomů kovů do struktury se měří změny v intenzitách. Z těchto změn je možné vydedukovat pozice těžkých atomů a následně odpovídající fáze (Taylor, 2003). Příprava vhodných izomorfních proteinových derivátů je však velice obtížná a časově náročná, a proto u jednodušších struktur ustupuje do pozadí (Blow a Rossmann, 1961). Naopak při řešení složitých struktur proteinů je tato metoda velmi úspěšná (Valvoda a kol., 1992).

Kratochvíl a kol. navíc uvádějí metodu anomálního rozptylu (2008), která bývá často kombinována s metodou izomorfní záměny. U metody anomálního rozptylu stačí jeden krystal, ale měření se provádí při několika různých vlnových délkách. Používají se vlnové délky blízké i vzdálené od absorpční hrany atomu, který je v asymetrické jednotce krystalu zastoupen jen v malém počtu – zpravidla se jedná o těžký atom (Terwilliger, 1994). Velikost a fáze příspěvku tohoto atomu ke strukturnímu faktoru je vlivem anomálního rozptylu změněna tak, jako by se na daném místě nacházel jiný atom. Jde tedy o dva experimenty, ale vzhledem k používání jediného krystalu je tato metoda snazší na provedení než metoda izomorfní záměny. Jsou zde však kladeny vyšší nároky na přesnost, protože dochází k měření nepatrných rozdílů intenzit. Pro necentrosymetrický krystal je metodou anomálního rozptylu získáno dvojznačné řešení. Tato nejednoznačnost se odstraňuje měřením intenzit reflexí Bijvoetových párů, tedy dvojice souborů rovin (Bijvoet a kol., 1951). Na významu nabývá tato metoda především při použití zdroje synchrotronového záření, kde je téměř neomezená volba vlnové délky záření, čímž lze vyvolat anomální rozptyl ve výrazné míře na libovolném atomu. Důležitý je fakt, že při různých vlnových délkách rentgenových paprsků mění strukturní faktor, odpovídající anomálně rozptylujícím atomům v elementární buňce, svou velikost, ale ne fázi. To znamená, že fázový rozdíl mezi ne-anomálně a anomálně rozptylujícími atomy v elementární buňce není závislý na vlnové délce rentgenového záření. Výsledkem měření jsou amplitudy strukturních faktorů pro Bijvoetovy páry při několika vlnových délkách. Konečným cílem analýzy je získat průměrnou amplitudu strukturního faktoru a její fázi pro každou reflexi při určité vlnové délce (Valvoda a kol., 1992).

2.7 Molekulové modelování a upřesnění struktury

Po vyřešení fázového problému a vygenerování mapy elektronové hustoty je možné začít s molekulovým modelováním. Za tímto účelem se používá např. počítačový program

Coot (Emsley a Cowtan, 2004), který je součástí softwaru CCP4 (Winn a kol., 2011) navrženého pro makromolekulární krystalografii. Hlavním úkolem tohoto programu je vizualizace makromolekulárních struktur a dat, molekulární modelování do elektronové hustoty a validace stávajících modelů. Coot poskytuje nástroje pro zobrazení trojrozměrných dat ve formě atomových modelů, map elektronové hustoty a obecných grafických objektů. Mapu elektronové hustoty generuje na základě souboru (ve formátu .mtz), který obsahuje strukturní faktory obsahující informace o fázi (Emsley a Cowtan, 2004; Emsley a kol., 2010).

Počáteční modelování se obvykle provádí automatizovanými metodami ARP/WARP (Perrakis a kol., 2001), RESOLVE (Terwilliger, 1999) a Buccaneer (Cowtan, 2006), které většinou vyžadují rozlišení lepší než 2,50 Å. Hlavním zaměřením programu je však dokončení počátečních modelů generovaných metodou molekulárního nahrazení nebo automatickým modelováním. Coot (Emsley a Cowtan, 2004) poskytuje celou řadu funkcí sloužících k molekulovému modelování jako např. idealizaci modelu, manuální rotaci a translaci, vyhledávání ligandů, solvataci, mutaci, přidání koncových reziduí, odstranění postranních řetězců atd. Další funkcí je propojení s programem REFMAC5 (Vagin a kol., 2004), který je součástí CCP4 a slouží k upřesnění struktury. Je vhodné opakovaně používat REFMAC5 po určité době interaktivního modelování pro zdokonalení aktuálních souřadnic (Emsley a Cowtan, 2004; Emsley a kol., 2010).

Cílem programu Coot (Emsley a Cowtan, 2004) v poslední době je především poskytnutí přístupu k co největšímu počtu nástrojů potřebných pro upřesnění a validaci makromolekulární struktury. Validace je zajištěna např. pomocí Ramachandranových diagramů, kontrolou molekul vody a analýzou geometrie, torzních úhlů, teplotních faktorů a rotamerů (Emsley a kol., 2010).

2.8 Validace struktury

Vzhledem k tomu, že molekulové modelování a upřesnění struktur založené na krystalografických datech je subjektivní, jsou pro posuzování platnosti modelů vyžadovány techniky kontroly kvality. Chyby v modelu jsou téměř nevyhnutelné a účelem upřesnění struktury a přestavby modelu je detekovat a odstranit chyby, aby byl získán nejlepší možný konečný model. Lokální chyby, které mohou ovlivnit biologickou interpretaci, jsou rozšířené u struktur s nízkým rozlišením a dokonce i struktury s vysokým

rozlišením většinou obsahují několik lokálních chyb. Od takových chyb může pomoci např. webový server MolProbity (Chen a kol., 2010), který nabízí validaci kvality trojrozměrné struktury proteinů, nukleových kyselin a komplexů (Kleywegt, 2000; Davis a kol., 2007; Chen a kol., 2010).

MolProbity (Chen a kol., 2010) poskytuje podrobnou analýzu kontaktů všech atomů v modelu, upozorňuje na sterické problémy uvnitř molekul a dokáže vypočítat a zobrazit vodíkové můstky a van der Waalovy interakce. To všechno provádí na základě přidání atomů vodíků do modelu pomocí programu REDUCE (Word a kol., 1999). Mezi další funkce MolProbity patří kontrola konformačních odchylek, analýza rotamerů a Ramachandranových diagramů a testování obou orientací aminokyselin asparaginu, glutaminu a histidinu, jejichž elektronová hustota je symetrická. Všechny provedené diagnostiky jsou pak shrnuty v grafech a grafických formulářích pro manuální přestavbu dalšími úpravami např. v programu Coot (Emsley a Cowtan, 2004) nebo KiNG (Davis a kol., 2007). Po vyřešení všech problémů je struktura znovu předložena k validaci a tyto kroky se opakují, dokud nejsou odstraněny všechny problémy. Při rozlišení 2,50 Å a vyšším lze takto opravit většinu nedostatků modelu (Davis a kol., 2007; Chen a kol., 2010).

2.9 Měřítko kvality získané struktury

2.9.1 R-faktor

R-faktor nebo také faktor spolehlivosti vyjadřuje míru shody mezi krystalografickým modelem a experimentálně naměřenými rentgenovými difrakčními daty. Jinými slovy, udává míru toho, jak upřesněná struktura reflektuje zjištěná data. R-faktor je definován následujícím vztahem:

$$R = \frac{\sum |F_{obs}| - |F_{calc}|}{\sum |F_{obs}|}$$

kde F značí strukturní faktor, který je pozorovaný (F_{obs}) a vypočítaný (F_{calc}). Minimální možná hodnota R-faktoru je nula a to v případě dokonalé shody mezi experimentálními pozorováními a strukturními faktory předpovězenými z modelu. Během upřesňování strukturního modelu by se měl R-faktor postupně snižovat. Za znak správnosti modelu se u dat s vysokým rozlišením považuje hodnota R-faktoru 0,25 a nižší (Jones a kol., 1991; Kleywegt a Jones, 1995; Kleywegt a Brünger, 1996; Kleywegt, 2000; Žídek, 2015).

2.9.2 Free R-faktor

Free R-faktor nebo také R_{free} je definován stejným vztahem jako R-faktor s tím rozdílem, že se počítá jen pro malou podmnožinu odrazů (reflexí), která se nepoužívá při upřesnění modelu. R_{free} tedy měří míru, do jaké model předpovídá difrakční data testovací sady, která byla vyloučena z procesu modelování a upřesnění struktury (křížová validace). Povaha krystalografických difrakčních dat je taková, že každý odraz obsahuje informace o celé struktuře. Změny modelu, které nezlepšují jeho schopnost popsat difrakční data, by neměly zlepšit shodu modelu s daty a hodnota R_{free} by měla zůstat konstantní nebo se zvýšit. V praxi to znamená, že modely s vážnými chybami mohou být identifikovány hodnotou R_{free} nad 0,40 (Brünger, 1992; Brünger, 1993; Kleywegt a Brünger, 1996; Brünger, 1997). Rozdíl mezi hodnotami R-faktoru a R_{free} by měl být malý (Kleywegt a Jones, 1995; Kleywegt a Brünger, 1997).

2.9.3 Kompletnost dat

Kompletnost dat vyjadřuje, jaká část reflexí v rozsahu Braggových vzdáleností byla skutečně naměřena ze všech teoreticky pozorovatelných. Lze ji zjistit výpočtem a udává se v procentech. Pokud nějaké odrazy chybí, vede to ke zhoršení modelových parametrů (Hirshfeld a Rabinovich, 1973; Kleywegt, 2000; Weiss, 2001).

2.9.4 Rozlišení

V krystalografii termín rozlišení odkazuje na množství dat, které bylo nakonec použito při určování struktury. Přesnost atomových pozic závisí především na kvalitě dat, kterou odráží např. parametry jako R-faktor, ale také na limitech rozlišení dat. Při rozlišení 0,80 Å lze rozlišit jednotlivé atomy. Při rozlišení 1,50 Å lze jen stěží rozlišit sousední atomy a rozlišení 2,50 až 3,00 Å poskytuje interpretovatelnou mapu elektronové hustoty (Rhodes, 2006).

2.9.5 Ramachandranův diagram

K posouzení stereochemické kvality řešených krystalových struktur se používá Ramachandranův diagram, který je součástí validačních nástrojů - např. ProCHECK (Laskowski a kol., 1993), MOLEMAN2 (Kleywegt a Jones, 1997) a MolProbity (Chen a kol., 2010). Princip tohoto diagramu je založen na faktu, že každá aminokyselina v polypeptidovém řetězci je spjata se dvěma torzními úhly Φ a Ψ . Většina kombinací těchto

úhlů není dovolena díky sterickým zábranám mezi postranními řetězci a díky tomu lze snadno vyčíslit ty kombinace, které dovolené jsou. Ramachandranův diagram je dvoudimenzionální graf, ve kterém jsou proti sobě vyneseny kombinace úhlů Φ (osa x) a Ψ (osa y). Tento graf představuje konformačně dovolené stavy hlavního řetězce polypeptidu, které jsou projevem potenciální energie dané aminokyselinou. V grafu je možné pozorovat oblasti, které jsou silně preferovány různými aminokyselinami a odpovídají určité sekundární struktuře, a oblasti, ve kterých se aminokyseliny nacházejí jen výjimečně (Hollingsworth a Karplus, 2010; Vondrášek a Vymětal, 2016).

2.9.6 B-faktory

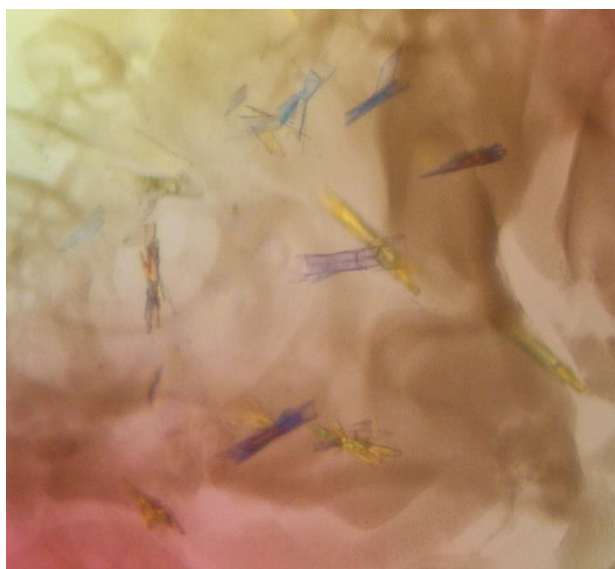
Atomové pohyby v krystalech jsou možné i při nízkých teplotách, ačkoliv jsou drasticky zredukované. Oscilační amplitudy atomů kolem jejich rovnovážných poloh jsou v krystalických strukturách monitorovány tzv. B-faktory (teplotními faktory). Pokud se snižuje rozptylová síla atomu, B-faktor se zvyšuje. B-faktory nezávisí jen na amplitudách atomových oscilací, ale také na statických nebo dynamických konformačních poruchách. Velmi vysoké B-faktory jsou obvykle spojovány s „neviditelnými“ proteinovými skupinami, jejichž struktura mizí v mapě elektronové hustoty. Molekuly s trvale vysokými B-faktory tak mohou být důsledkem problémů v modelu (Kleywegt a kol., 1996; Carugo, 2018).

Krystalické struktury s vysokým rozlišením (méně než 1,50 Å) by měly vykazovat maximální hodnotu B-faktoru 25 Å², naopak struktury s nízkým rozlišením (více než 3,30 Å) dosahují maximální hodnoty B-faktoru až 80 Å². Je zřejmé, že maximální hodnota B-faktoru se zvyšuje při horším rozlišení. K posouzení toho, jak rozumné jsou teplotní faktory modelu, existuje poměrně málo dostupných metod. Jednou z nich je např. porovnání, jak dobře se průměrný teplotní faktor modelu shoduje s hodnotou vypočítanou z dat pomocí Wilsonova diagramu. Je však dobré mít na paměti, že nízký průměr B-faktoru sám o sobě není nutně známkou vysoké kvality modelu (Kleywegt, 2000; Carugo, 2018).

3. Materiál a metody

3.1 Sběr difrakčních dat

Pro rentgenovou difrakční analýzu byly vybrány krystaly proteinu Tt81 připravené metodou sedící kapky, o velikosti $\sim 100 \times 40 \times 30 \mu\text{m}^3$, zobrazené na Obr. 5. Rentgenová difrakce byla provedena na vědeckém pracovišti Helmholtz Zentrum Berlin (HZB) v německém Berlíně se zdrojem synchrotronového záření BESY-II. Před zahájením experimentu byly krystaly zmrazeny pomocí tekutého dusíku, aby bylo zamezeno teplotně-vibračním pohybům. Rentgenové záření o vlnové délce $0,918 \text{ \AA}$ procházelo krystaly, které byly rotovány s oscilačním úhlem $0,1^\circ$. Při každém pootočení krystalu byl na detektoru typu PILATUS zaznamenán difrakční snímek. Vzdálenost krystalu od detektoru byla $297,091 \text{ mm}$. Měření bylo ukončeno po otočení krystalu o 180° .



Obr. 5: Krystaly proteinu Tt81 vybrané pro rentgenovou difrakční analýzu.

3.2 Zpracování difrakčních dat

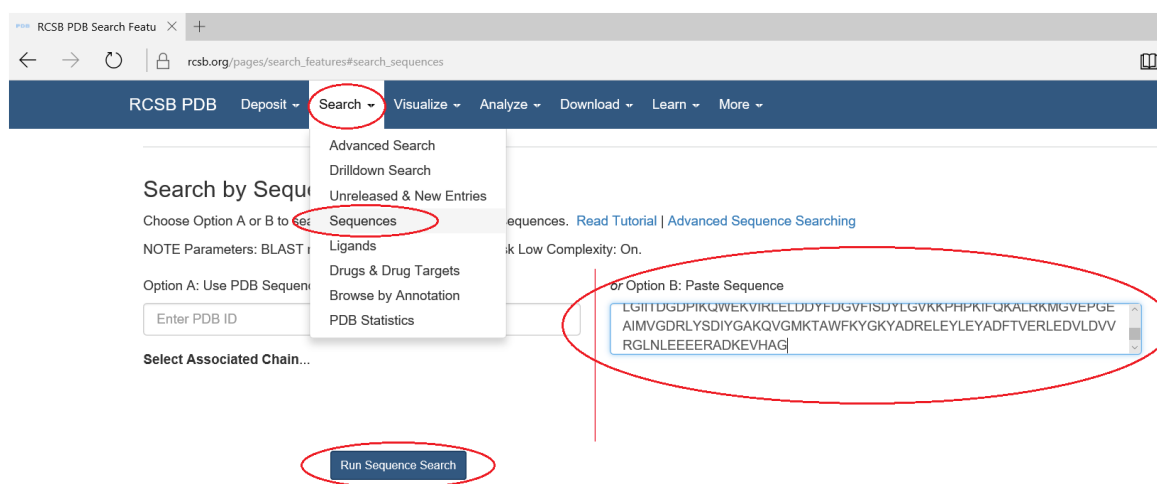
Získaná difrakční data byla zpracována pomocí programu XDSAPP (Sparta a kol., 2016) v režimu GUI během tří kroků – indexace, integrace a analýzy. Po zpracování difrakčních dat bylo zobrazeno stručné shrnutí celého procesu a byl získán výstupní soubor ve formátu .mtz obsahující informace o intenzitách jednotlivých odrazů. Tento soubor byl načten programem VIEWHKL (Krissinel a Evans, 2012), který je součástí CCP4 (Winn a kol., 2011). VIEWHKL umožňuje grafické zobrazení difrakčních dat, kdy jsou jednotlivé

odrazy zobrazeny jako kulaté skvrny, jejichž velikost a intenzita udává amplitudu nebo intenzitu odrazu (Krissinel a Evans, 2012). Soubor .mtz byl dále nahrán do programu Matthews_coef (Kantardjieff a Rupp, 2003), který je schopný vyhodnotit počet molekul v asymetrické jednotce, což je jeden z prvních kroků při stanovení makromolekulární struktury.

3.3 Řešení fázového problému

3.3.1 Srovnání sekvencí

Pro řešení fázového problému byla zvolena metoda molekulárního nahrazení, která vyžaduje již známý model podobného proteinu. Na základě znalosti aminokyselinové sekvence proteinu Tt81 bylo v RCSB PDB databázi spuštěno srovnání sekvencí pro nalezení vhodného homologu. Postup práce s RCSB PDB databází je naznačen na Obr. 6.

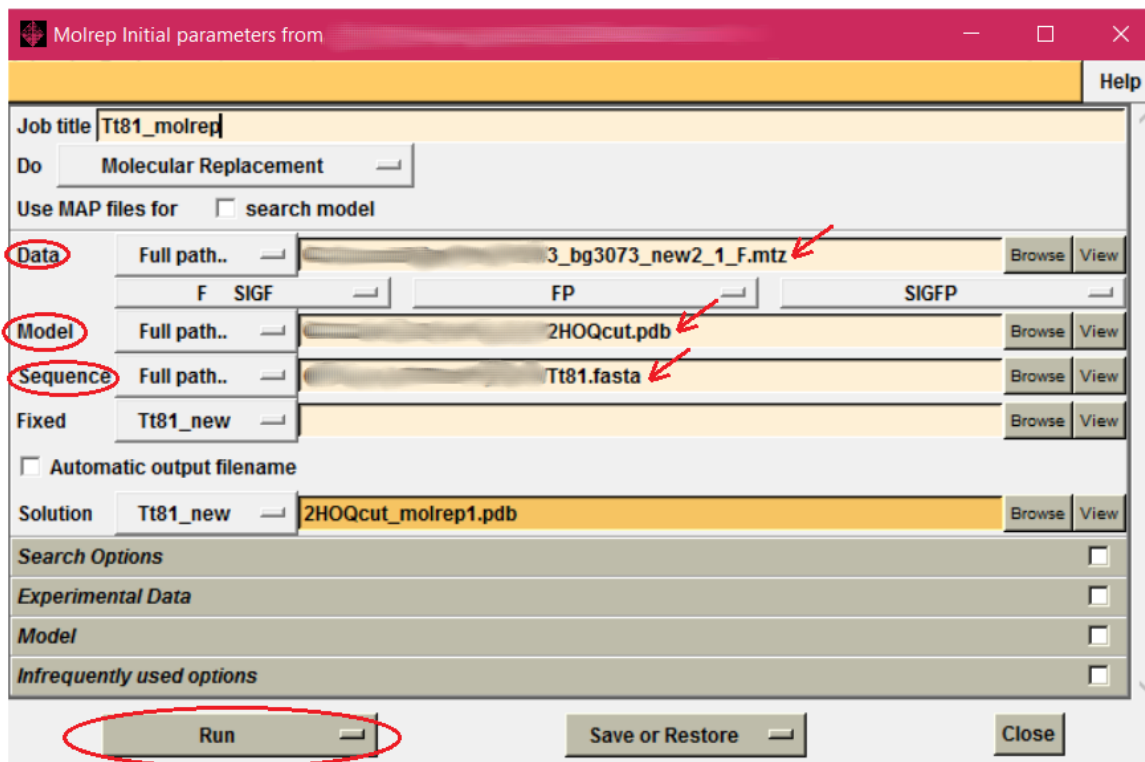


Obr. 6: Srovnání sekvencí proteinů v RCSB PDB databázi.

3.3.2 Molekulární nahrazení

Pdb soubor vyhledaného homologu (ID PDB: 2HOQ) byl stažen a následně použit pro vytvoření počátečního modelu programem MOLREP (Vagin a Teplyakov, 2000). MOLREP je součástí CCP4 (Winn a kol., 2011) a jedná se o automatický program pro molekulární nahrazení, který na základě informací z homologního proteinu, parametrů elementární buňky a symetrie krystalu dokáže vygenerovat počáteční model studovaného proteinu (Lebedev a kol., 2008). Jako výchozí data pro MOLREP byl použit pdb soubor proteinu 2HOQ, mtz soubor proteinu Tt81 a soubor ve formátu .fasta, který byl vytvořen

v textovém editoru, obsahující sekvenci proteinu Tt81. Jednotlivé kroky práce s programem MOLREP jsou vyznačeny na Obr. 7.



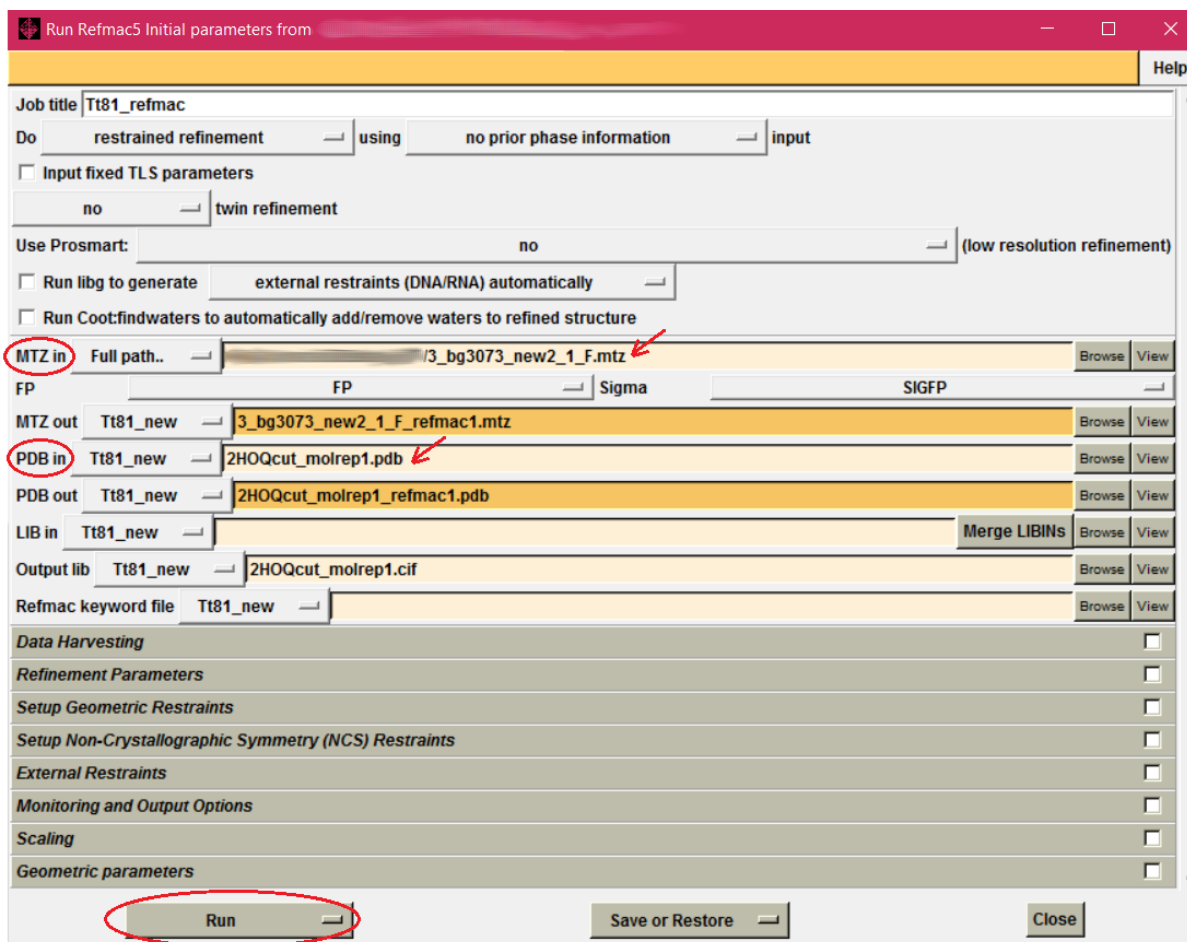
Obr. 7: Příprava počátečního modelu v programu MOLREP (Vagin a Teplyakov, 2000).

3.4 Upřesnění struktury

Počáteční model vygenerovaný programem MOLREP (Vagin a Teplyakov, 2000) byl následně upraven systémem BALBES (Long a kol., 2008), který se používá pro řešení proteinových struktur rentgenovou krystalografií metodou molekulární náhrady. Součástí BALBES je také software ARP/wARP (Perrakis a kol., 2001) sloužící pro automatizované budování modelu a jeho refinaci.

Další úpravy modelu pak byly provedeny v programu Coot (Emsley a Cowtan, 2004). Nejdříve byly zkontrolovány jednotlivé aminokyseliny řetězce. Několik chybně zařazených aminokyselin bylo zmutováno tak, aby odpovídaly aminokyselinové sekvenci proteinu Tt81. Následovalo fitování řetězců do mapy elektronové hustoty zejména použitím příkazů „Real Space Refine Zone“ a „Regularize Zone“, které zlepšují geometrii modelu. V oblastech postrádajících mapu elektronové hustoty byly aminokyseliny vymazány.

Vždy po několika úpravách v programu Coot (Emsley a Cowtan, 2004) byl spuštěn REFMAC5 (Vagin a kol., 2004) pro zdokonalení souřadnic modelu. Jako výchozí data pro REFMAC5 byl použit pdb soubor, vytvořený programem MOLREP (Vagin a Teplyakov, 2000) při tvorbě modelu, a mtz soubor proteinu Tt81 z XDSAPP (Sparta a kol., 2016). Jednotlivé kroky práce s programem REFMAC5 jsou vyznačeny na Obr. 8. Při každém dalším spuštění programu REFMAC5 byl použit pdb soubor vygenerovaný v předchozím cyklu, s uloženými změnami vytvořenými v programu Coot.



Obr. 8: Upřesnění modelu v programu REFMAC5 (Vagin a kol., 2004).

3.5 Validace struktury

Validace struktury byla provedena na internetových serverech MolProbity (Chen a kol., 2010) a ProCHECK (Laskowski a kol., 1993), které jsou dostupné online na webových stránkách *molprobity.biochem.edu* a *servicesn.mbi.ucla.edu*. ProCHECK hodnotí celkovou kvalitu struktury na základě porovnání s dobře rafinovanými strukturami

o podobném rozlišení a zdůrazňuje regiony, které vyžadují další přezkoumání, a to prostřednictvím celé řady grafů (Laskowski a kol., 1993).

Jako výchozí data pro validaci byl v obou případech použit pdb soubor proteinu Tt81 obsahující finální verzi modelu po upřesnění struktury pomocí programů Coot (Emsley a Cowtan, 2004) a REFMAC5 (Vagin a kol., 2004). Po načtení dat do MolProbity (Chen a kol., 2010) byla zvolena funkce „add hydrogens“ pro přidání atomů vodíků do struktury a poté byla spuštěna analýza kontaktů všech atomů v modelu pomocí funkce „analyze all-atom contacts and geometry“. Vyhodnocení této analýzy a grafy vygenerované serverem ProCHECK (Laskowski a kol., 1993) jsou shrnuty v kapitole 4.5.

3.6 Zobrazení struktury

K zobrazení finální struktury proteinu Tt81 byl použit software PyMOL (DeLano, 2002), který slouží ke grafickému zobrazování molekulárních dat získaných pomocí NMR spektroskopie, nebo rentgenové krystalografie. Pro zobrazení struktury stačí vložit soubor ve formátu .pdb. V pravém sloupci, vedle zobrazené struktury, se nachází ovládací panel, který obsahuje pět záložek – A, S, H, L a C. Záložka A ~ *action* nabízí odstranění a přejmenování objektů a možnosti „zoom“ a „center“, které lze použít k orientaci ve struktuře. Záložka S ~ *show* umožňuje zobrazení struktury v různých grafických podobách – tyčinkový model, mašličkový model, atd. Mezi další možnosti softwaru PyMOL patří skrytí struktury (záložka H ~ *hide*), označení vybraných částí struktury – např. označení jednotlivých aminokyselin, řetězců, nebo jednotlivých atomů (záložka L ~ *label*), a obarvení různých částí struktury (záložka C ~ *color*). Je možné obarvit např. všechny sekundární struktury stejnou barvou, nebo obarvit celý protein jako barevné spektrum, případně obarvit pouze vybranou část proteinu. Pro rychlou práci s tímto softwarem je také možné použít příkazový řádek nad zobrazenou strukturou. Jednotlivé příkazy lze dohledat v různých manuálech, např. na webových stránkách pymol.sourceforge.net.

3.6.1 Zobrazení sekvenčních motivů

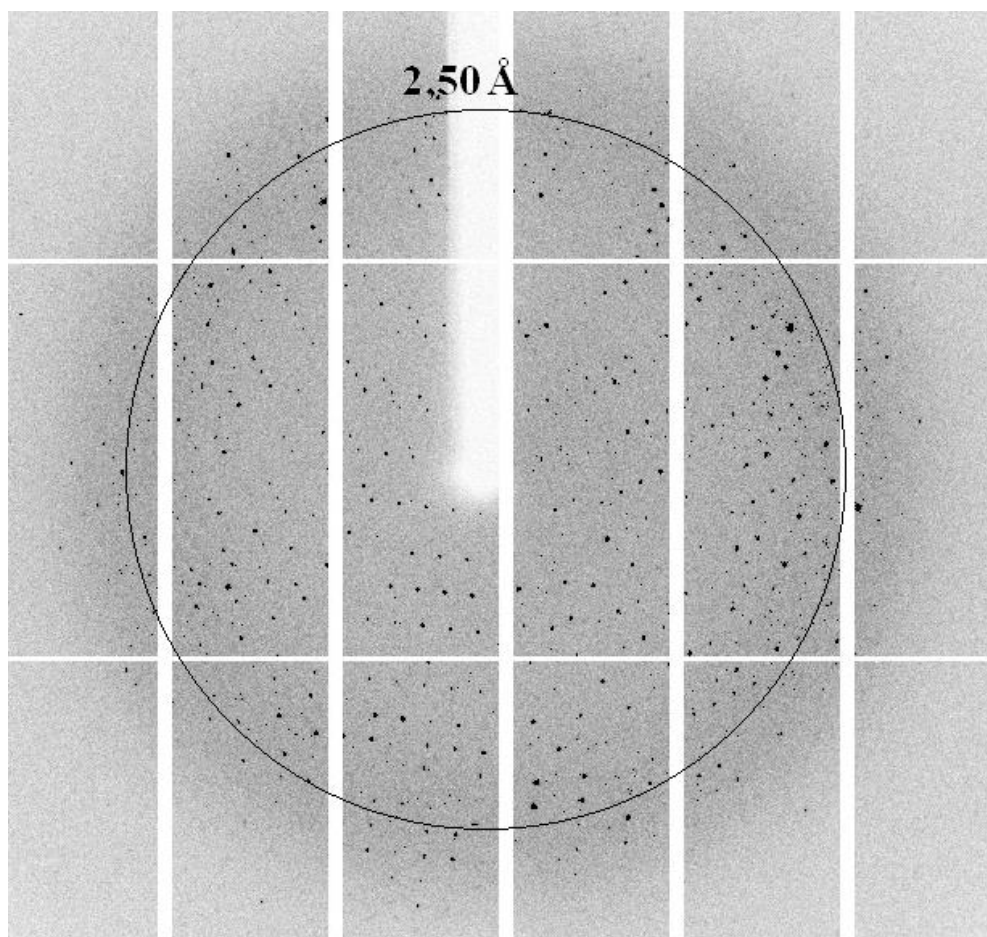
Pro nadrodinu halogenaciddehalogenas je typická přítomnost čtyř sekvenčních motivů (Aravind a kol., 1998). Pro nalezení těchto motivů byla nejdříve porovnána trojrozměrná struktura proteinu Tt81 z *Thermococcus thioreducens* s ostatními strukturami v PDB databázi pomocí DALI serveru (Holm a Laakso, 2016). DALI server je dostupný online na stránkách ekhidna2.biocenter.helsinki.fi a jako výchozí data využívá pdb soubor,

který je možné nahrát po přepnutí na záložku „PDB search“. Dalším krokem bylo porovnání sekvencí proteinů Tt81, ID PDB: 2HOQ a dvou proteinů, ID PDB: 4FFD a ID PDB: 4YGR, které byly vygenerovány DALI serverem. Za tímto účelem byl použit server ClustalW (Chojnacki a kol., 2017) dostupný online na stránkách *ebi.ac.uk/Tools/msa/clustalo*. Jako výchozí data pro ClustalW byl použit textový soubor obsahující aminokyselinové sekvence všech čtyř vybraných proteinů a získaný ALN soubor byl následně nahrán serverem ESPrpt (Robert a Gouet, 2014) na webových stránkách *esprpt.ibcp.fr*. ESPrpt ~ „Easy Sequencing in PostScript“ je program, který poskytuje sekvenční podobnosti a informace o sekundární struktuře proteinů na základě porovnaných sekvencí (Robert a Gouet, 2014). Sekvenční podobnosti proteinů Tt81, 2HOQ, 4FFD a 4YGR jsou uvedené v kapitole 4.6.1.

4. Výsledky

4.1 Výsledky rentgenové difrakční analýzy

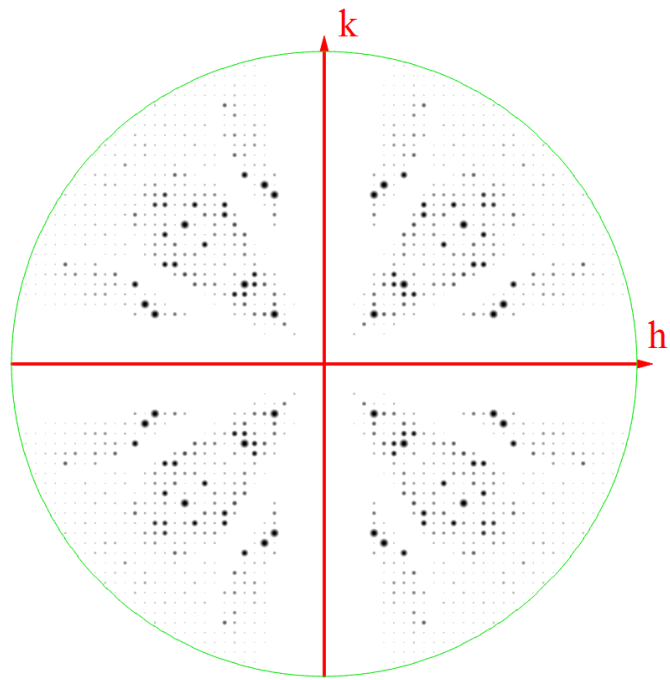
Během rentgenové difrakční analýzy bylo na detektoru zaznamenáno 1800 difrakčních snímků proteinu Tt81 s rozlišením 2,50 Å (viz Obr. 9).



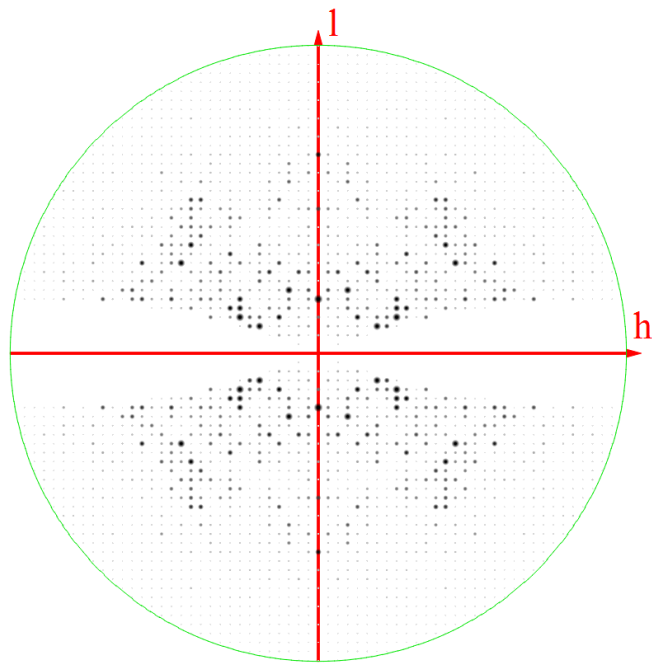
Obr. 9: Jeden z difrakčních snímků proteinu Tt81.

4.2 Výsledky zpracování difrakčních dat

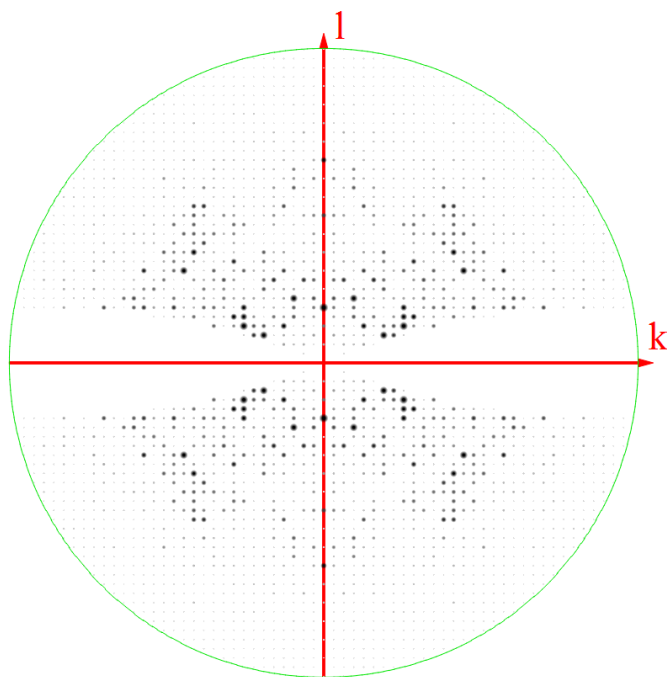
Program XDSAPP (Sparta a kol., 2016) zpracoval celkem 2146 skvrn (79 %) z difrakčních snímků a výsledkem bylo stručné shrnutí parametrů, které jsou uvedeny v Tab. 2, a mtz soubor, který po načtení programem VIEWHKL (Krissinel a Evans, 2012) poskytl grafické zobrazení difrakčních dat (viz Obr. 10). Programem Matthews_coef (Kantardjieff a Rupp, 2003) bylo zjištěno, že asymetrická jednotka obsahuje jednu molekulu proteinu Tt81 a 47 % rozpouštědla. Hodnota Matthewsovo koeficientu byla 2,32.



hk0
 (H=K=0, L=1)



h0l
 (H=L=0, K=1)



0kl

(K=L=0, H=1)

Obr. 10: Grafické zobrazení difrakčních dat pomocí programu VIEWHKL (Krissinel a Evans, 2012), kde H, K a L jsou základní krystalografické roviny.

Tab. 2: Shrnutí zpracování difrakčních dat, zobrazené programem XDSAPP (Sparta a kol., 2016).

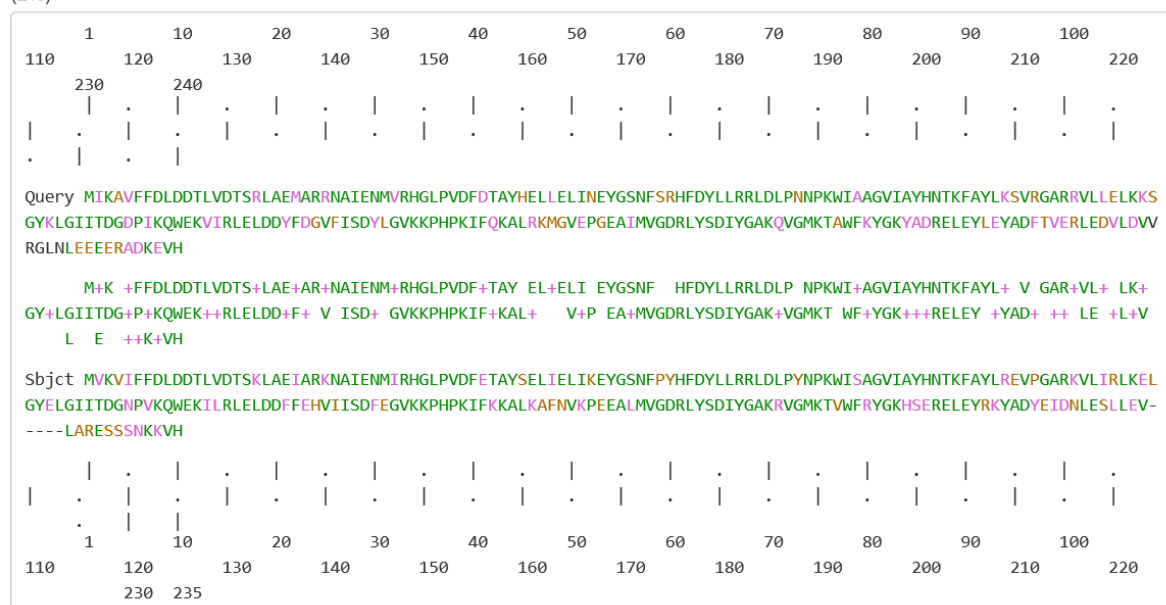
| | |
|-------------------------------|----------------------------------|
| Prostorová grupa | P4(2)2(1)2 (94) |
| Rozměry elementární buňky [Å] | 78,08 78,08 84,47 90,0 90,0 90,0 |
| Limit rozlišení [Å] | 46,22 – 2,47 (2,62 – 2,47) |
| Multiplicita | 12,61 |
| I/sigI | 19,84 (2,39) |
| R_meas [%] | 11,5 (114,5) |
| Kompletnost dat [%] | 99,7 (98,5) |
| B (Wilson) [Å ²] | 56,01 |
| Mozaicita [°] | 0,153 |
| ISa | 38,18 |

4.3 Výsledky řešení fázového problému

4.3.1 Výsledky srovnání sekvencí

Nejvyšší sekvenční podobnost s proteinem Tt81 (71 %) vykazoval po srovnání sekvencí z RCSB PDB protein s označením 2HOQ (viz Obr. 11). Soubor proteinu 2HOQ ve formátu .pdb byl z databáze stažen a následně použit pro molekulární nahrazení. Protein 2HOQ, který byl izolován z hypertermofilní bakterie *Pyrococcus horikoshii*, a jehož krystalová struktura o rozlišení 2,00 Å byla vyřešena metodou anomálního rozptylu, je považován za putativní halogenaciddehalogenasu (Arai a kol., 2006).

Length: 240 E-value: 2.11833E-100 Score: 363.229bits (931) Identities: 170/240 (71%) Positives: 209/240 (87%) Gaps: 5/240 (2%)

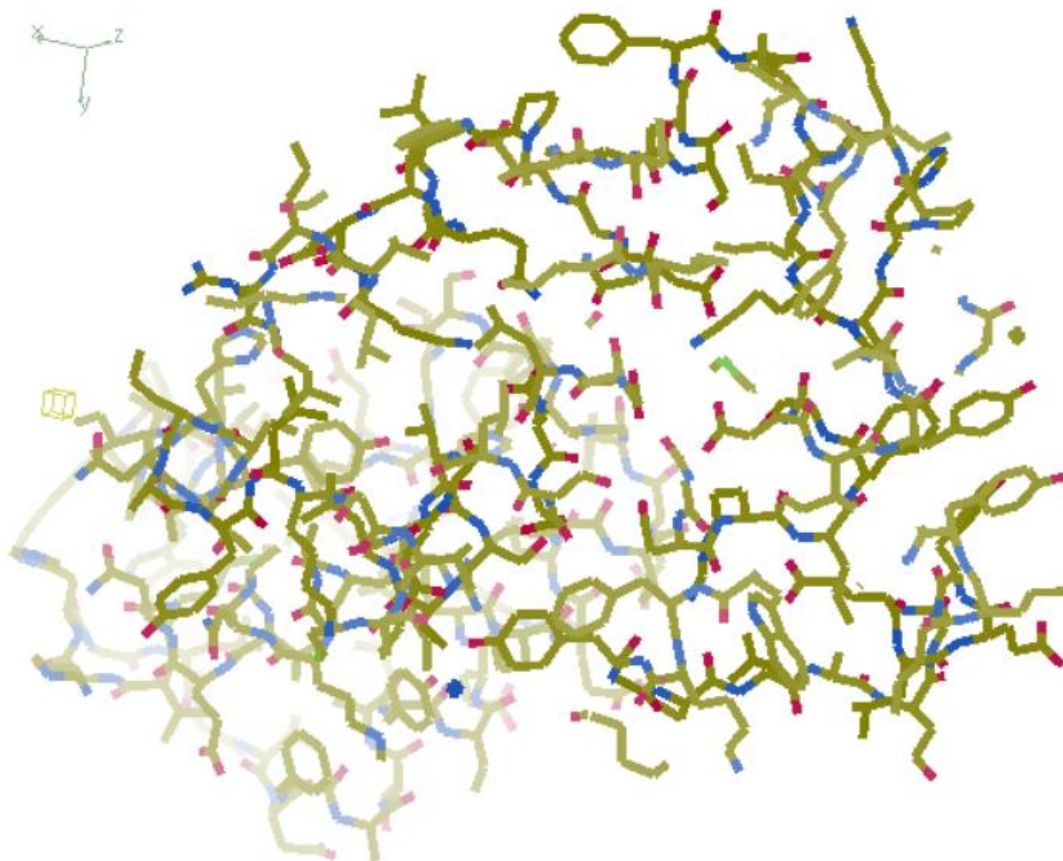


Obr. 11: Sekvenční podobnost proteinu Tt81 (horní část obrázku) a proteinu ID PDB: 2HOQ (dolní část obrázku). Zeleně jsou vyznačeny aminokyseliny, které se v obou proteinech shodují. (Vygenerováno v RCSB PDB: Protein Data Bank [online]. Dostupné 20. 9. 2017 z WWW: <<https://www.rcsb.org/>>.)

4.3.2 Výsledky molekulárního nahrazení

Fázový problém byl vyřešen metodou molekulárního nahrazení, kdy jako homolog byl použit protein 2HOQ, který vykazoval 71% sekvenční shodu s proteinem Tt81. Po zpracování výchozích dat (pdb soubor proteinu 2HOQ, mtz soubor proteinu Tt81

a soubor fasta, obsahující sekvenci proteinu Tt81) programem MOLREP (Vagin a Teplyakov, 2000) byl vygenerován počáteční model, který byl následně otevřen v programu Coot (Emsley a Cowtan, 2004), viz Obr. 12.



Obr. 12: Počáteční model vygenerovaný programem Coot (Emsley a Cowtan, 2004) - žlutě je znázorněn uhlíkový řetězec, červeně kyslík, modře dusík a zeleně síra.

4.4 Výsledky upřesnění struktury

Výsledkem počátečního modelování systémem BALBES (Long a kol., 2008) byl model upravený během 50 refinací a bylo vyhodnoceno několik parametrů, které jsou shrnuty v Tab. 3. BALBES určil jako nejvíce pravděpodobnou prostorovou grupu č. 94, stejně jako XDSAPP (Sparta a kol., 2016) při zpracování difrakčních dat.

Po úpravách struktury v programu Coot (Emsley a Cowtan, 2004) obsahoval model 228 aminokyselin z celkových 242. 14 aminokyselin z C-konce proteinu bylo z modelu vymazáno kvůli chybějící mapě elektronové hustoty.

Tab. 3: Shrnutí zpracování počátečního modelu, vygenerované systémem BALBES (Long a kol., 2008).

| | |
|------------------------------|-----------------|
| Prostorová grupa | P4(2)2(1)2 (94) |
| Limit rozlišení [Å] | 46,24 – 2,69 |
| R-faktor počátečního modelu | 0,2745 |
| R-faktor konečného modelu | 0,2465 |
| Kompletnost dat [%] | 99,61 |
| B (Wilson) [Å ²] | 49,34 |

4.5 Výsledky validace struktury

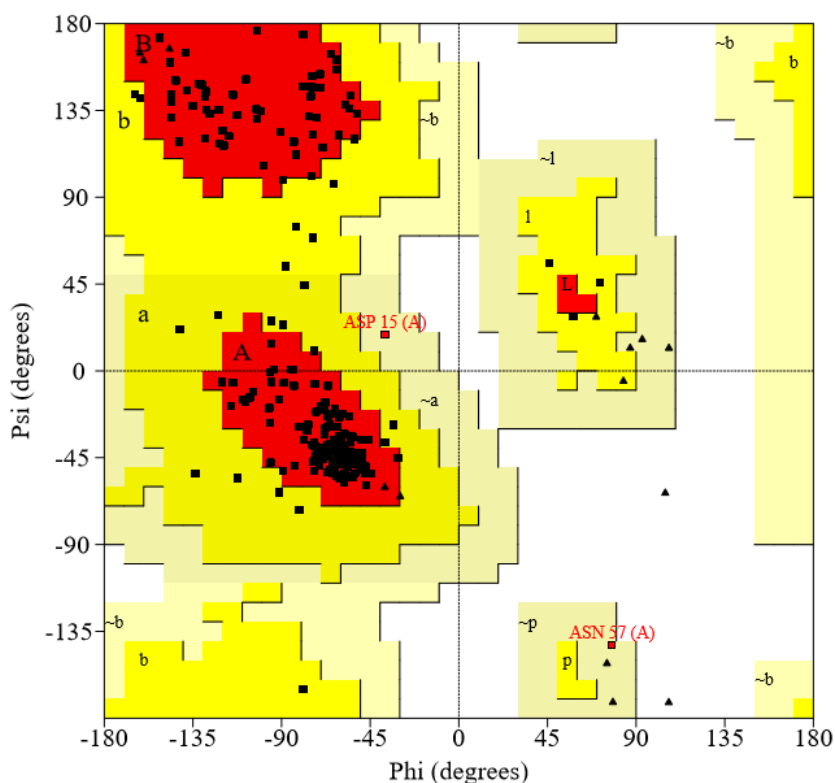
Internetový server MolProbity (Chen a kol., 2010) identifikoval jeden unikátní řetězec zahrnující hlavní řetězec a postranní řetězce, celkem obsahující 228 reziduí. Během validace nedošlo k překlopení žádné skupiny, což indikuje, že všechny asparaginy, glutaminy a histidiny jsou orientovány správně. Byla určena hodnota R-faktoru 0,2468 a hodnota free R-faktoru 0,2948. Vyhodnocení analýzy kontaktů všech atomů v modelu je shrnuto v Tab. 4.

Tab. 4: Shrnutí validace struktury internetovým serverem MolProbity (Chen a kol., 2010).

| Kontakty atomů | Clashscore, all atoms | 0,91 | 100 th percentile |
|--------------------|-----------------------|--------|------------------------------|
| Geometrie proteinu | Poor rotamers | 12 | 7,89 % |
| | Favored rotamers | 131 | 86,18 % |
| | Ramachandran outliers | 4 | 1,77 % |
| | Ramachandran favored | 208 | 92,04 % |
| | MolProbity score | 1,94 | 96 th percentile |
| | Bad bonds | 0/1748 | 0,00 % |
| | Bad angles | 2/2360 | 0,08 % |

ProCHECK (Laskowski a kol., 1993) prověřil celkem 228 reziduí a zobrazil všechny kombinace torzních úhlů Φ a Ψ Ramachandrovým diagramem, který je zobrazen na Obr. 13. 88,7 % reziduí leží v přednostně zaujímané oblasti, 10,3 % v povolené oblasti, 1,0 % ve výjimečně povolené oblasti a žádná rezidua neleží v zakázané oblasti, což potvrzuje

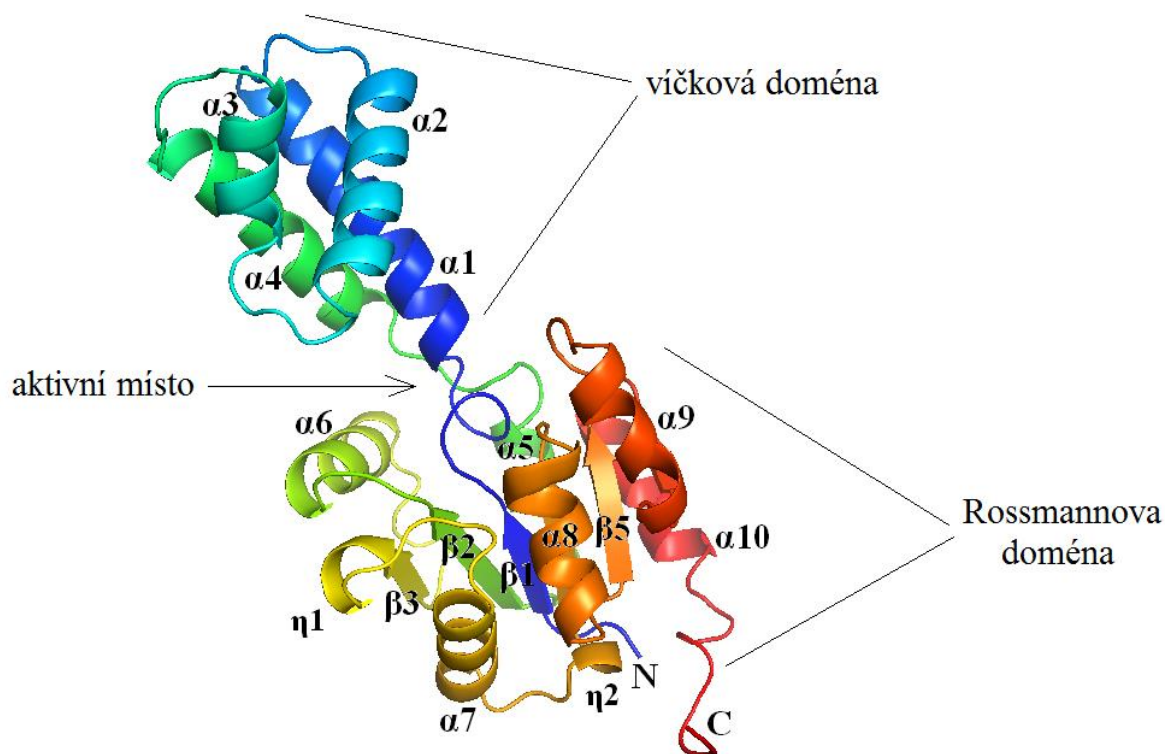
věrohodnost modelu. Serverem byla objevena zkreslená geometrie ve třech místech modelu. Jednalo se o délku vazby v hlavním řetězci u aminokyselin Tyr199 a Asp219 a o úhel vazby v hlavním řetězci u aminokyseliny Asp15.



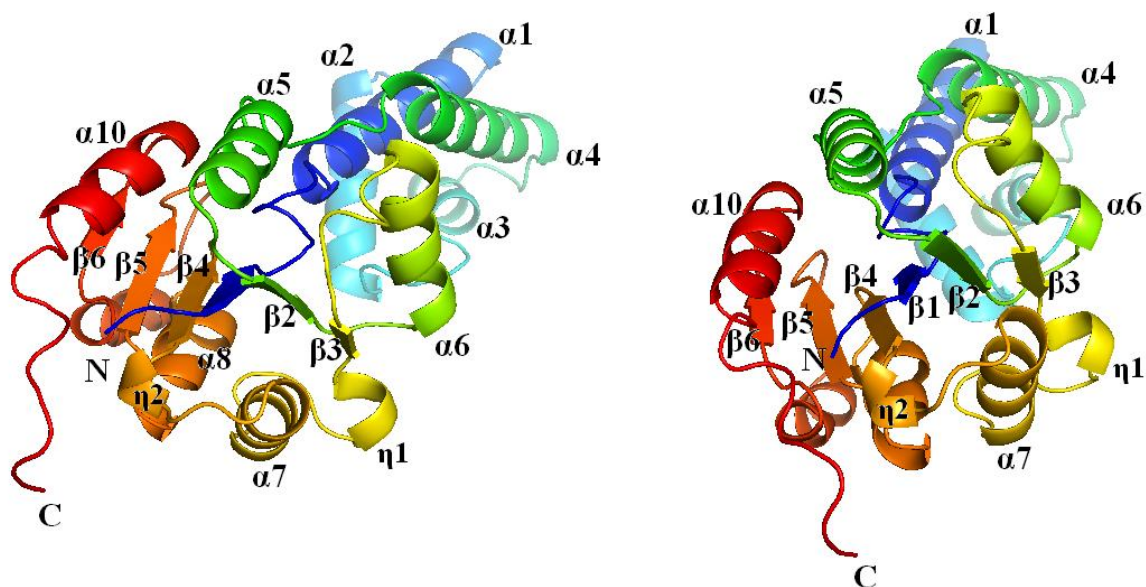
Obr. 13: Ramachandranův diagram pro strukturu proteinu Tt81 vygenerovaný serverem ProCHECK (Laskowski a kol., 1993) - červeně je znázorněna přednostně zaujímaná oblast, tmavě žlutě povolená oblast, světle žlutě výjimečně povolená oblast a bíle zakázaná oblast.

4.6 Výsledky zobrazení struktury

Finální struktura proteinu Tt81 zobrazená programem PyMOL (DeLano, 2002) je na Obr. 14 a 15. Jedná se o monomer tvořený dvěma doménami. Jádrem proteinu je tzv. Rossmannova doména, která je typická pro halogenaciddehalogenasy. Tato doména se skládá z šesti paralelně uspořádaných β -listů, v pořadí $\beta_6 - \beta_5 - \beta_4 - \beta_1 - \beta_2 - \beta_3$, které jsou lemované šesti α -helixy. Druhou doménou je víčková doména, složená ze čtyř α -helixů. Mezi těmito dvěma doménami se nachází aktivní místo proteinu. Podstatou víčkové domény je regulace přístupu substrátu k aktivnímu místu a určuje substrátovou specifitu proteinu. V rámci nadrodiny halogenaciddehalogenas se můžeme setkat s velkou strukturální a funkční diverzitou, která je způsobená právě rozmanitostí víčkových domén a jejich umístěním (Allen, 2004; Allen, 2009).



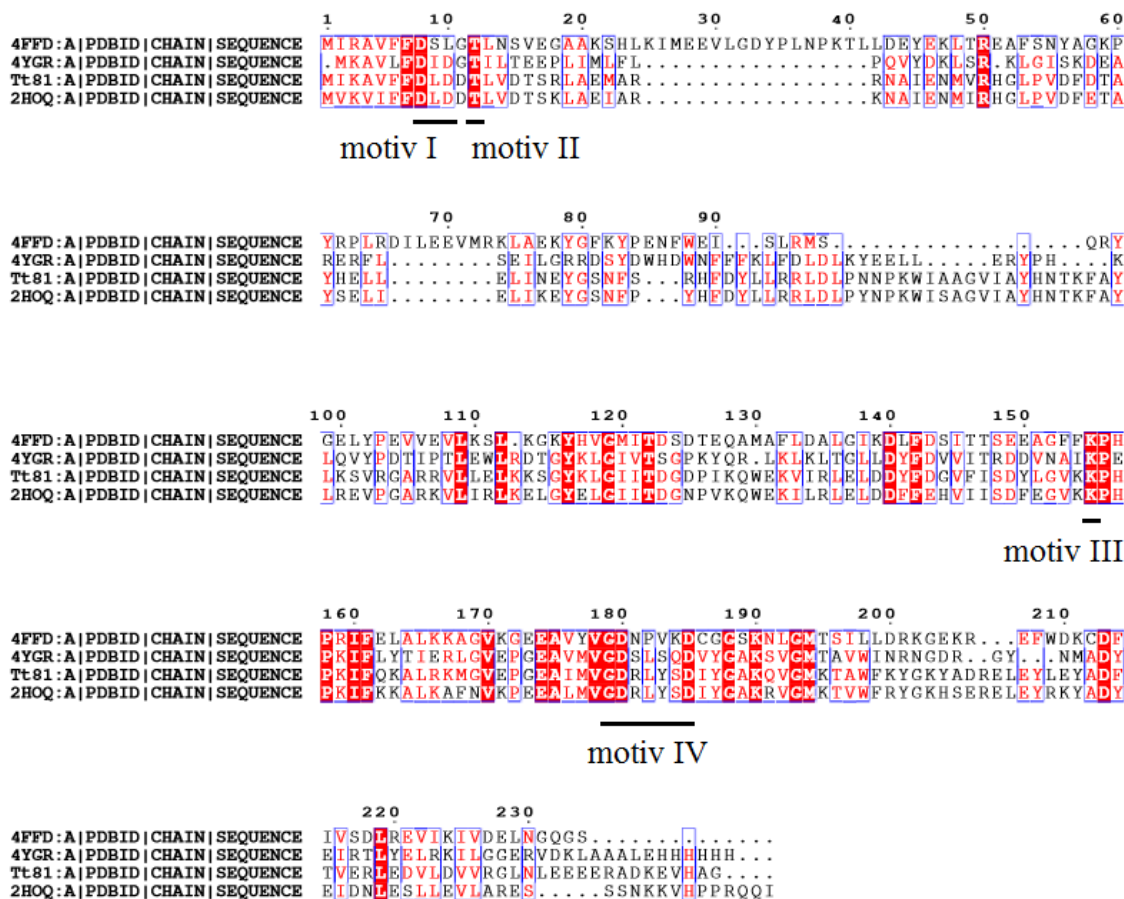
Obr. 14: Finální struktura proteinu Tt81, zobrazená programem PyMOL (DeLano, 2002), s vyznačenými doménami a aktivním místem (α značí α -helix, β značí β -list, η značí 3_{10} -helix, číslování od N-konce k C-konci proteinu).



Obr. 15: Finální struktura proteinu Tt81, zobrazená programem PyMOL (DeLano, 2002), s Rossmannovo doménou v popředí (α značí α -helix, β značí β -list, η značí 3_{10} -helix, číslování od N-konce k C-konci proteinu).

4.6.1 Výsledky zobrazení sekvenčních motivů

Na základě sekvenčních podobností proteinů Tt81, ID PDB: 2HOQ, 4FFD a 4YGR, které byly vyhodnoceny serverem ESPript (Robert a Gouet, 2014), viz Obr. 16, byly nalezeny čtyři sekvenční motivy, typické pro halogenaciddehalogenasy.



Obr. 16: Sekvenční podobnost proteinů 4FFD, 4YGR, Tt81 a 2HOQ vygenerovaná serverem ESPript (Robert a Gouet, 2014) s vyznačenými motivy I, II, III a IV.

Prvním motivem je DxD (kde x značí libovolnou aminokyselinu), který se nachází po β 1. U proteinu Tt81 se jedná o Asp8 a Asp10, které slouží ke koordinaci kofaktoru Mg^{2+} (Seal a Rose, 1987; Qian a kol., 1994; Collet a kol., 1997; Baker a kol., 1998; Morais a kol., 2000). Druhým motivem je vysoce konzervovaný threonin (Thr12), ležící mezi β 1 a α 1. Třetí motiv, vysoce konzervovaný lysin (Lys150), se nachází poblíž N-konce α 7 předcházejícím β 4. Druhý a třetí motiv společně přispívají ke stabilitě reakčních meziproduktů hydrolytické reakce (Burroughs a kol., 2006). Čtvrtým motivem jsou

konzervovaná kyselá rezidua poblíž C-konce $\beta 4$ v podobě DD, GDxxxD, nebo GDxxxxD (Burroughs a kol., 2006). V případě proteinu Tt81 se jedná o motiv GDxxxxD (Gly174, Asp175 a Asp180). Tato kyselá rezidua společně s rezidui prvního motivu jsou potřebná pro koordinaci iontu Mg^{2+} v aktivním místě (Hisano a kol., 1996; Morais a kol., 2000; Toyoshima a kol., 2000; Wang a kol., 2001; Lahiri a kol., 2002; Rinaldo-Matthis a kol., 2002; Lahiri a kol., 2003; Peisach a kol., 2004;).

5. Diskuze

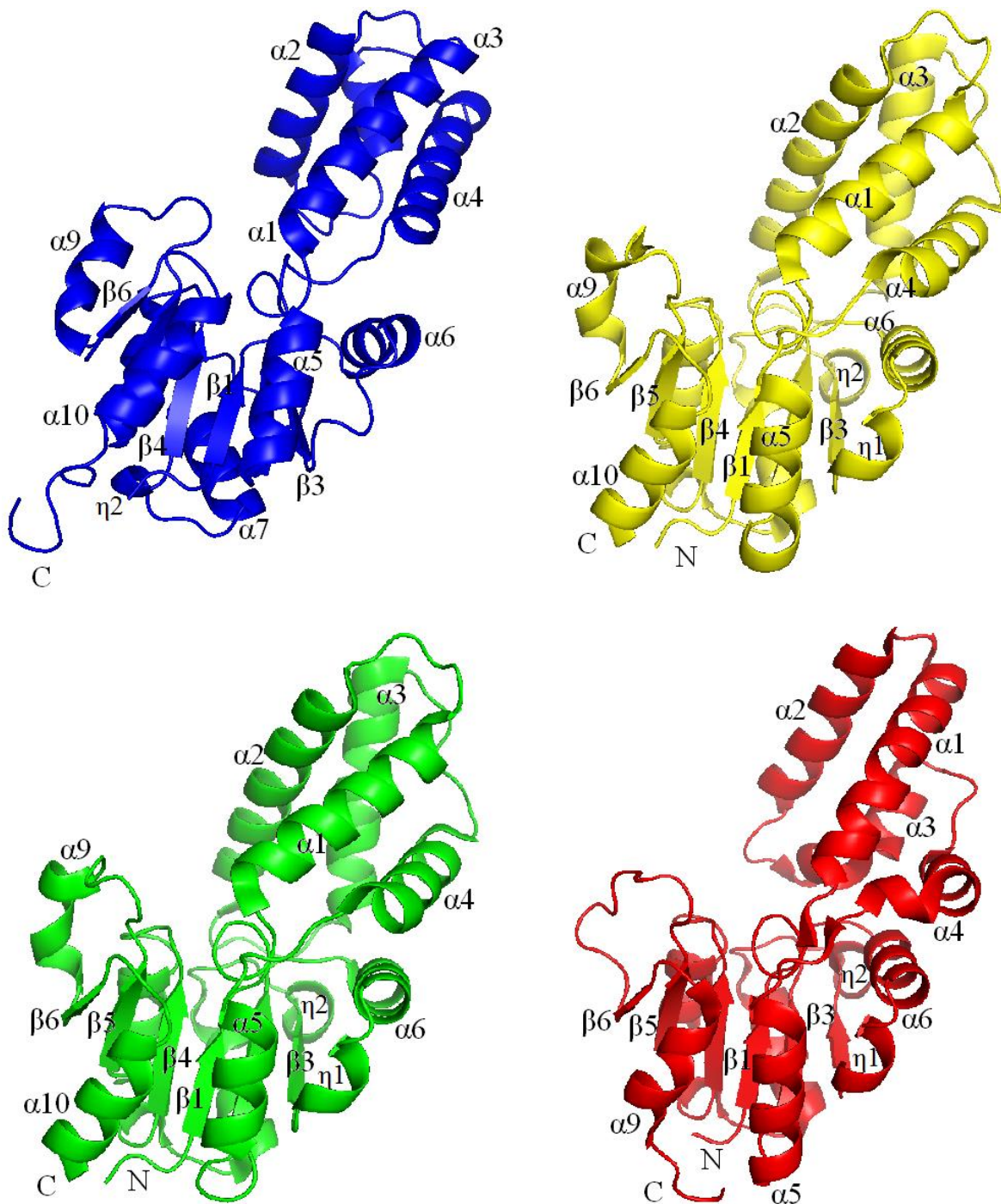
Krystaly proteinu Tt81 z *Thermococcus thio-reducens* o velikosti 100x40x30 μm^3 byly podrobeny rentgenové difrakci na vědeckém pracovišti Helmholtz Zentrum Berlin (HZB) v Berlíně na zdroji synchrotronového záření BESY-II. Vzdálenost krystalu od detektoru byla 297,091 mm a vlnová délka procházejícího záření byla 0,918 Å. Krystalem bylo otáčeno s oscilačním úhlem 0,1° dokud nebylo dosaženo otočení o 180°. Získaná difrakční data o rozlišení 2,50 Å byla zpracována pomocí XDSAPP (Sparta a kol., 2016). Fázový problém byl vyřešen metodou molekulárního nahrazení (Rossmann, 1972) za použití homologu ID PDB: 2HOQ, který vykazoval sekvenční identitu 71 % (RCSB PDB, 2016). Počáteční model byl vytvořen programem MOLREP (Vagin a Teplyakov, 2000) a struktura byla dále upřesňována pomocí programů Coot (Emsley a Cowtan, 2004) a REFMAC5 (Vagin a kol., 2004). Validace byla provedena s využitím internetových serverů MolProbity (Chen a kol., 2010) a ProCHECK (Laskowski a kol., 1993), který upozornil na několik lokálních chyb v modelu. Jednalo se o délku vazby v hlavním řetězci u aminokyselin Tyr199 a Asp219 a o úhel vazby v hlavním řetězci u aminokyseliny Asp15. Tyto lokální chyby a fakt, že 88,7 % reziduí leží v přednostně zaujímané oblasti Ramachandranova diagramu, i když ideálně by jich mělo být nad 90 % (Laskowski a kol., 1993), pravděpodobně souvisí s horším rozlišením difrakčních dat (2,50 Å). Znakem správnosti finálního modelu je hodnota R-faktoru 0,2468, hodnota free R-faktoru 0,2948 a kompletnost dat 99,61. Finální model byl zobrazen pomocí programu PyMOL (DeLano, 2002).

Porovnání trojrozměrné struktury proteinu Tt81 z *Thermococcus thio-reducens* se strukturami uloženými v PDB databázi pomocí DALI serveru (Holm a Laakso, 2016) ukázalo podobnost s halogenaciddehalogenasou ID PDB: 1X42 z *Pyrococcus horikoshii* (Arai a kol., 2006) ($Z = 23,3$; RMSD = 2,7), s HAD fosfatasou ID PDB: 4YGR z *Thermococcus onnurineus* (Ngo a kol., 2015) ($Z = 23,9$; RMSD = 3,3) a s proteinem s neznámou funkcí ID PDB: 4FFD z *Pyrococcus horikoshii* (Seetharaman a kol., v přípravě) ($Z = 25,2$; RMSD = 2,7). Struktury 1X42 a Tt81 jsou vůči sobě mírně posunuté, avšak víčková doména je v obou případech tvořena svazkem čtyř α -helixů a hlavní doména odpovídá Rossmannovu uspořádání s šesti β -listy uprostřed, které jsou obklopené šesti α -helixy. Nepatrný rozdíl mezi strukturami 1X42 a Tt81 je dán odlišným umístěním prvního 3_{10} -helixu, který se u 1X42 nachází mezi α_6 a β_3 a u Tt81 leží mezi β_3 a α_7 . Značně posunuté jsou vůči sobě struktury 4YGR a Tt81. Struktura 4YGR obsahuje v Rossmannově

doméně pouze pět α -helixů a mezi hlavní a víčkovou doménou navíc obsahuje dvě β -vlásky (před a po víčkové doméně). Téměř zcela se překrývají struktury 4FFD a Tt81, především v oblasti hlavní domény, zatímco víčkové domény jsou vůči sobě nepatrně posunuté (viz Obr. 17). Pořadí α -helixů, 3_{10} -helixů a β -listů je identické. Struktury proteinů Tt81, 2HOQ, 4FFD a 4YGR jsou zobrazeny na Obr. 18.



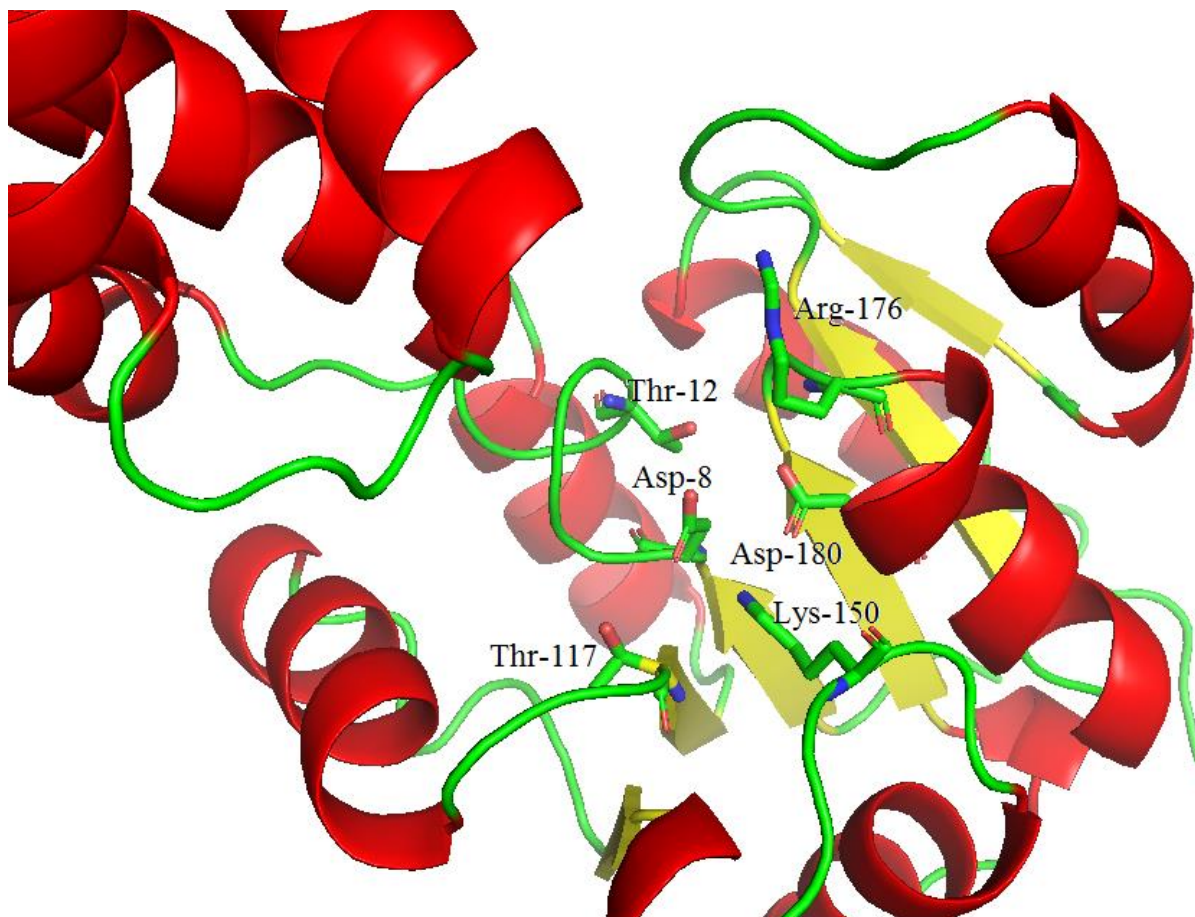
Obr. 17: Překrývající se struktury proteinů Tt81 (modře) a 4FFD (žlutě) zobrazené programem PyMOL (DeLano, 2002).



Obr. 18: Struktury proteinů Tt81 (modře), 4FFD (žlutě), 1X42 (zeleně) a 4YGR (červeně) zobrazené programem PyMOL (DeLano, 2002). α značí α -helix, β značí β -list, η značí 3_{10} -helix, číslování od N-konce k C-konci proteinu.

Ačkoliv největší shodu v trojrozměrném uspořádání se strukturou proteinu Tt81 vykazoval protein 4FFD, pro posouzení aktivního místa byl použit protein 1X42, který byl již popsán (Arai a kol., 2006). 1X42 obsahuje konzervovaná rezidua aktivního místa: Asp8, Thr12, Thr122, Lys155, Asn181 a Asp185 (Arai a kol., 2006), čemuž u proteinu Tt81

příslušně odpovídají následující rezidua: Asp8, Thr12, Thr117, Lys150, Arg176 a Asp180. Asp8 působí jako nukleofil a během katalýzy tvoří aspartylový meziprodukt (Seal a kol., 1987; Qian a kol., 1994; Collet a kol., 1997; Baker a kol., 1998; Morais a kol., 2000). Aktivní místo proteinu Tt81 bylo zobrazeno programem PyMOL (DeLano, 2002), viz Obr. 19.



Obr. 19: Aktivní místo proteinu Tt81 zobrazené programem PyMOL (DeLano, 2002) v podobě tyčinkového modelu. Zbytek modelu je zobrazen jako mašličkový diagram (červeně jsou zobrazeny α -helixy, žlutě β -listy, zeleně smyčky).

6. Závěr

Teoretická část práce se zabývá strukturou proteinových krystalů, rentgenovou difrakční analýzou a řešením proteinových struktur od zpracování difrakčních dat až po jejich validaci.

Praktická část práce je věnována strukturní studii putativní halogenaciddehalogenasy Tt81, která byla izolována z *Thermococcus thio-reducens*. V práci jsou detailně popsány jednotlivé kroky řešení struktury: sběr difrakčních dat, jejich zpracování, řešení fázového problému metodou molekulárního nahrazení, tvorba počátečního modelu, upřesnění struktury, její validace a nakonec zobrazení finálního modelu proteinu Tt81.

Krystaly proteinu Tt81 krystalizovaly v tetragonální krystalové mřížce odpovídající prostorové grupě číslo 94 ($P4_22_12$) s rozměry elementární buňky: $a = 78,08 \text{ \AA}$, $b = 78,08 \text{ \AA}$, $c = 84,47 \text{ \AA}$, $\alpha = 90,00^\circ$, $\beta = 90,00^\circ$ a $\gamma = 90,00^\circ$. Bylo zjištěno, že asymetrická jednotka obsahuje jednu molekulu proteinu Tt81. Struktura byla vyřešena metodou molekulárního nahrazení s rozlišením $2,50 \text{ \AA}$. Finální model obsahuje 228 reziduí, přičemž 14 reziduí z C-konce proteinu bylo vymazáno z důvodu chybějící mapy elektronové hustoty. Model se skládá z hlavní a víčkové domény, mezi nimiž se nachází aktivní místo proteinu. Hlavní doména vykazuje Rossmannovo uspořádání s α/β strukturou tvořenou šesti-vláknovým paralelním β -listem, který je obklopen šesti α -helixy. Strukturu víčkové domény, která je vložena do hlavní domény, tvoří svazek čtyř α -helixů. Aktivní místo je tvořeno rezidui Asp8, Thr12, Thr117, Lys150, Arg176 a Asp180. Ve struktuře byly nalezeny čtyři sekvenční motivy, které potvrzují, že protein Tt81 patří do nadrodiny halogenaciddehalogenas.

7. Seznam literatury

Allegra, G. (1980). Principles of Direct Methods of Phase Determination in Crystal Structure Analysis. In: Ladd, M. F. C., & Palmer, R. A. (1980). *Theory and Practice of Direct Methods in Crystallography*. New York, United States: Springer US.

Allen, K. N., & Dunaway-Mariano, D. (2004). Phosphoryl group transfer: evolution of a catalytic scaffold. *Trends in Biochemical Sciences*, 29(9). 495-503.

Allen, K. N., & Dunaway-Mariano, D. (2009). Markers of fitness in a successful enzyme superfamily. *Current Opinion in Structural Biology*, 19(6). 658-665.

Arai, R., Kumimoto-Niino, M., Kuroishi, C., Bessho, Y., Shirouzu, M., & Yokoyama, S. (2006). Crystal structure of the probable haloacid dehalogenase PH0459 from *Pyrococcus horikoshii* OT3. *Protein Science*, 15(2). 373-377.

Aravind, L., Galperin, M. Y., & Koonin, E. V. (1998). The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold. *Trends in Biochemical Sciences*, 23(4). 127-129.

Baker, A. S., Ciocci, M. J., Metcalf, W. W., Kim, J., Babbitt, P. C., Wanner, B. L., Martin, B. M., & Dunaway-Mariano, D. (1998). Insight into the mechanism of catalysis by the P-C bond-cleaving enzyme phosphonoacetaldehyde hydrolase derived from gene sequence analysis and mutagenesis. *Biochemistry*, 37(26). 9305-9315.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., & Leslie, A. G. W. (2011). *iMOSFLM*: a new graphical interface for diffraction-image processing with *MOSFLM*. *Acta Crystallographica*, D67. 271-281.

Bijvoet, J. M., Peerdeman, A. F., & van Bommel, A. J. (1951). Determination of the Absolute Configuration of Optically Active Compounds by Means of X-Rays. *Nature*, 168(4268). 271-272.

BioSync. A structural Biologist's Guide to High Energy Data Collection Facilities [online]. 2018. [cit. 2018-09-11]. Dostupné z WWW: <<http://biosync.sbkb.org/index.jsp>>.

Blow, D. M., & Rossmann, M. G. (1961). The Single Isomorphous Replacement Method. *Acta Crystallographica*, 14(11). 1195-1202.

- Bokhoven, C., Schoone, J. C., & Bijvoet, J. M. (1951). The Fourier synthesis of the crystal structure of strychnine sulphate pentahydrate. *Acta Crystallographica*, 4(3). 275-280.
- Bragg, W. H., & Bragg, W. L. (1913). The reflection of X-rays by crystals. *Proceedings of the Royal Society of London*, 88(605). 428-438.
- Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 355(6359). 472-475.
- Brünger, A. T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallographica*, 49(1). 24-36.
- Brünger, A. T. (1997). The free R-value: a more objective statistic for crystallography. *Methods in Enzymology*, 277. 366-396.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Krahn, J. M., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Schroeder, G. F., Simonson, T., & Warren, G. L. Crystallography & NMR System (CNS). 1997-2010 Yale University.
- Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., & Aravind, L. (2006). Evolutionary Genomics of the HAD Superfamily: Understanding the Structural Adaptations and Catalytic Diversity in a Superfamily of Phosphoesterases and Allied Enzymes. *Journal of Molecular Biology*, 361(5). 1003-1034.
- Carugo, O. (2018). How large B-factors can be in protein crystal structures. *BMC Bioinformatics*, 19(1). 61.
- Chen, V. B., Arendall III, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., & Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica*, 66(1). 12-21.
- Chojnacki, S., Cowley, A., Lee, J., Foix, A., & Lopez, R. (2017). Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. *Nucleic Acids Research*, 45(W1). W550-W553.

- Collet, J. F., Gerin, I., Rider, M. H., Veiga-da-Cunha, M., & Van Schaftingen, E. (1997). Human L-3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate. *FEBS Letters*, 408(3). 281-284.
- Cowtan, K. (2006). The Buccaneer software for automated model building. *Acta Crystallographica*, 62(Pt 9). 1002-1011.
- Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall III, W. B., Snoeyink, J., Richardson, J. S., & Richardson, D. C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, 35. 375-383.
- DeLano, W. L. (2002). *PyMOL*. DeLano Scientific, San Carlos, CA, 700.
- Diederichs, K. (2006). Some aspects of quantitative analysis and correction of radiation damage. *Acta Crystallographica*, D62. 96-101.
- Eckert, M. (2012). Max von Laue and the discovery of X-ray diffraction in 1912. *Annalen der Physik*, 524(5). A83-A85.
- Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallographica*, 60(12). 2126-2132.
- Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica*, 66(4). 486-501.
- Hisano, T., Hata, Y., Fujii, T., Liu, J. Q., Kurihara, T., Esaki, N., & Soda, K. (1996). Crystal structure of L-2-haloacid dehalogenase from *Pseudomonas* sp. YL. An alpha/beta hydrolase structure that is different from the alpha/beta hydrolase fold. *Journal of Biological Chemistry*, 271(34). 20322-20330.
- Hollingsworth, S. A., & Karplus, P. A. (2010). A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomolecular Concepts*, 1(3-4). 271-283.
- Holm, L., & Laakso, L. M. (2016). Dali server update. *Nucleic acids research* 44(W1), W351-W355.

- Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallographica*, A47. 110-119.
- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *Journal of Applied Crystallography*, 26. 795-800.
- Kantardjieff, K. A., & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science*, 12(9). 1865-1871.
- Kleywegt, G. J., & Jones, T. A. (1995). Where freedom is given, liberties are taken. *Structure*, 3(6). 535-540.
- Kleywegt, G. J., & Brünger, A. T. (1996). Checking your imagination: applications of the free R value. *Structure*, 4(8). 897-904.
- Kleywegt, G. J., & Jones, T. A. (1997). Model-building and refinement practise. *Methods in Enzymology*, 277. 208-230.
- Kleywegt, G. J. (2000). Validation of protein crystal structures. *Acta Crystallographica*, 56(3), 249-265.
- Koonin, E. V., & Tatusov, R. L. (1994). Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search. *Journal of Molecular Biology*, 244(1). 125-132.
- Kratochvíl, B., Hušák, M., Brynda, J., & Sedláček, J. (2008). Co nabízí současná RTG strukturní analýza? *Chemické listy*, 102. 889-901.
- Krissinel, E., & Evans, P. Interactive Graphical Viewer and Browser for Reflection Data [online]. 2012. [cit. 2018-10-10]. Dostupné z WWW: <<http://www.ccp4.ac.uk/newsletters/newsletter48/articles/ViewHKL/viewhkl.html>>.
- Krug, M., Weiss, M. S., Heinemann, U., & Mueller, U. (2012). XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. *Journal of Applied Crystallography*, 45(3). 568-572.

- Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., & Allen, K. N. (2002). Caught in the act: the structure of phosphorylated beta-phosphoglucomutase from *Lactococcus lactis*. *Biochemistry*, 41(26). 8351-8359.
- Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., & Allen, K. N. (2003). The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. *Science*, 299(5615), 2067-2071.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK – a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2). 283-291.
- Lebedev, A. A., Vagin, A. A., & Murshudov, G. N. (2008). Model preparation in MOLREP and examples of model improvement using X-ray data. *Acta Crystallographica*, D64. 33-39.
- Long, F., Vagin A. A., Young, P., & Murshudov, G. N. (2008). BALBES: a molecular-replacement pipeline. *Acta Crystallographica Section D*, 64(Pt 1). 125-132.
- McPherson, A. (2002). *Introduction to macromolecular crystallography*. Hoboken, New Jersey: John Wiley & Sons.
- Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., & Allen, K. N. (2000). The crystal structure of *Bacillus cereus* phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. *Biochemistry*, 39(34). 10385-10396.
- Navaza, J. (1994). AMoRe: an Automated Package for Molecular Replacement. *Acta Crystallographica*, A50. 157-163.
- Otwinowski, Z., & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, A276. 307-326.
- Peisach, E., Selengut, J. D., Dunaway-Mariano, D., & Allen, K. N. (2004). X-ray crystal structure of the hypothetical phosphotyrosine phosphatase MDP-1 of the haloacid dehalogenase superfamily. *Biochemistry*, 43(40), 12770-12779.
- Perrakis, A., Harkiolaki, M., Wilson, K. S., & Lamzin, V. S. (2001). ARP/wARP and molecular replacement. *Acta Crystallographica*, 57(Pt 10). 1445-1450.

- Pflugrath, J. W. (1999). The finer things in X-ray diffraction data collection. *Biological Crystallography*, 55(10). 1718-1725.
- Powell, H. R. (2017). X-ray data processing. *Bioscience Reports*, 37(5).
- Qian, N., Stanley, G. A., Hahn-Hagerdal, B., & Radstrom, P. (1994). Purification and characterization of two phosphoglucomutases from *Lactococcus lactis* subsp. *Lactis* and their regulation in maltose- and glucose-utilizing cells. *Journal of Bacteriology*, 176(17). 5304-5311.
- Rappaport, T. S. (2002). *Wireless communications: Principle and practice*. Upper Saddle River, New Jersey: Prentice Hall.
- RCSB PDB: Protein Data Bank [online]. Dostupné z WWW: <<http://www.rcsb.org/>>.
- Read, R. J. (2003). Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallographica*, D57. 1373-1382.
- Rhodes, G. (2006). *Crystallography Made Crystal Clear*. San Diego, United States: Elsevier Science Publishing Co Inc.
- Rinaldo-Matthis, A., Rampazzo, C., Reichard, P., Bianchi, V., & Nordlund, P. (2002). Crystal structure of a human mitochondrial deoxyribonucleotidase. *Nature Structural & Molecular Biology*, 9(10). 779-787.
- Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(W1). W320-W324.
- Rossmann, M. G. (1972). *The Molecular Replacement Method*. New York, United States: Gordon & Breach.
- Rupp, B. (2010). *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. New York, United States: Garland Science.
- Schmidberger, J. W., Wilce, J. A., Tsang, J. S. H., & Wilce, M. C. J. (2007). Crystal Structures of the Substrate Free-enzyme, and Reaction Intermediate of the HAD Superfamily Member, Haloacid Dehalogenase DehIVa from *Burkholderia cepacia* MBA4. *Journal of Molecular Biology*, 368(3). 706-717.

- Seal, S. N., & Rose, Z. B. (1987). Characterization of a phosphoenzyme intermediate in the reaction of phosphoglycolate phosphatase. *Journal of Biological Chemistry*, 262(28). 13496-13500.
- Sheldrick, G. M. (2015). Crystal structure refinement with *SHELXL*. *Structural Chemistry*, 71(1). 3-8.
- Soda, K., Kurihara, T., Liu, J.-Q., Nardi-Dei, V., Park, C., Miyagi, M., Tsunasawa, S., & Esaki, N. (1996). Bacterial 2-haloacid dehalogenases: Structures and catalytic properties. *Pure and Applied Chemistry*, 68(11). 2097-2103.
- Sparta, K. M., Krug, M., Heinemann, U., Mueller, U., & Weiss, M. S. (2016). XDSAPP2.0. *Journal of Applied Crystallography*, 49. 1085-1092.
- Taylor, G. (2003). The phase problem. *Acta Crystallographica*, 59(11). 1881-1890.
- Terwilliger, T. C. (1994). MAD Phasing: Treatment of Dispersive Differences as Isomorphous Replacement Information. *Acta Crystallographica*, D50. 17-23.
- Terwilliger, T. C. (1999). Reciprocal-space solvent flattening. *Acta Crystallographica*, D55. 1863-1871.
- Toyoshima, C., Nakasako, M., Nomura, H., & Ogawa, H. (2000). Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature*, 405(6787). 647-655.
- Vagin, A. A., Richelle, J., & Wodak, S. J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with atomic model. *Acta Crystallographica*, D55. 191-205.
- Vagin, A. A., & Teplyakov, A. (2000). An approach to multi-copy search in molecular replacement. *Acta Crystallographica*, 56(Pt12). 1622-1624.
- Vagin, A. A., Steiner, R. S., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., & Murshudov, G. N. (2004). REFMAC5 dictionary: organisation of prior chemical knowledge and guidelines for its use. *Acta Crystallographica*, D60. 2284-2295.
- Valvoda, V., Polcarová, M., & Lukáč, P. (1992). *Základy strukturní analýzy*. Praha, Česká republika: Karolinum Praha.

Vávra, V., & Losos, Z. Multimediální studijní texty z mineralogie pro bakalářské studium [online]. 2006. [cit. 2018-10-08]. Dostupné z WWW: <<http://mineralogie.sci.muni.cz/>>.

Vondrášek, J., & Vymětal, J. (2016). Konformační chování aminokyselin v peptidech a proteinech z pohledu molekulárního modelování a výpočetních metod. *Chemické listy*, 110. 385-393.

Wang, W., Kim, R., Jancarík, J., Yokota, H., & Kim, S. H. (2001). Crystal structure of phosphoserine phosphatase from *Methanococcus jannaschii*, a hyperthermophile, at 1.8 Å resolution. *Structure*, 9(1). 65-71.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica*, 67(Pt4). 235-242.

Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., Richardson, J. S., & Richardson, D. C. (1999). Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. *Journal of Molecular Biology*, 285(4). 1711-1733.

Zwart, P. H., Grosse-Kunstleve, R. W., & Adams, P. D. (2005). Xtriage and Fest: automatic assessment of X-ray data and substructure structure factor estimation. *CCP4 Newsletter on Protein Crystallography*, 43(7).

Židek, L. (2015). *Strukturní biochemie*. Skripta. Brno: Masarykova univerzita.