

Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Biochemie

Studijní obor: Biochemie



Michal Hanigovský

Strukturní a funkční charakterizace inhibice koronavirové methyltransferázy

Structural and functional characterization of inhibition of a coronaviral
methyltransferase

Bakalářská práce

Vedoucí práce:

Mgr. et Mgr. Evžen Bouřa, Ph.D.

Praha 2022

Prohlášení:

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením školitele Mgr. et. Mgr. Evžena Bouři, Ph.D. a všechny použité zdroje jsem řádně citoval.

V Praze dne 23.5.2022

Podpis:

Poděkování

Velice rád bych poděkoval vedoucímu bakalářské práce Mgr. et Mgr. Evženu Bouřovi, Ph.D. a konzultantce RNDr. Petře Kraččíkové, Ph.D. za vedení práce, odborné konzultace, rady a pomoc, a také za jejich ochotu a čas během celého působení na ÚOCHB AV ČR.

Současně bych rád poděkoval RNDr. Martinu Klímovi, Ph.D. a RNDr. Janu Šilhánovi Ph.D. za pomoc při řešení proteinové struktury, a v neposlední řadě všem členům skupiny Evžena Bouři za cenné rady.

Abstrakt

Vysoce patogenní koronavirus SARS-CoV-2 (Severe Accute Respiratory Syndrome Coronavirus 2) způsobil světovou pandemií onemocnění COVID-19 (coronavirus disease 2019) a představuje hrozbu celému lidstvu v mnoha směrech. Do dnešního dne je datováno více než 510 miliónu potvrzených případů a přes 6 miliónu úmrtí podle Světové zdravotnické organizace (WHO, <https://covid19.who.int/> [29.4.2022]). Funkcí virových methyltransferáz je maskování a ochrana genomu před imunitním systémem hostitelské buňky. Důležitým aspektem pro stabilizaci virové RNA je tvorba RNA čepičky, kterou zprostředkovává 2'-O-RNA methyltransferáza (MTase, z angl. methyltransferase) a 7-methylguaninový nukleotid připojený k 5'-trifosfátovému konci (m⁷GpppNm). 2'-O-MTase aktivita je asociována s proteinem nsp16 (z angl. non-structural protein 16), který pro svou správnou funkci vyžaduje přítomnost proteinu nsp10 jako kofaktoru.

Abychom zjistili strukturu proteinového komplexu nsp10-nsp16, byly v této práci rekombinantně připraveny proteiny nsp10 a nsp16 expresí v bakteriích *Escherichia coli*. Pro obdržení krystalové struktury byl proteinový komplex inhibován látkou TO383 v aktivním místě. Zjištěná data mohou posloužit pro vývoj antivirotik založených na inhibici methyltransferázy a destabilizaci virové RNA.

Klíčová slova:

SARS-CoV-2, methyltransferáza, nsp16, nsp10, struktura, inhibice

Abstract

Highly pathogenic coronavirus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused global pandemic of coronavirus disease 2019 (COVID-19). Up to date, more than 510 million cases of COVID-19 and over 6.2 million deaths were confirmed according to World Health Organization (WHO, <https://covid19.who.int/> [29.4.2022]). The function of viral methyltransferases is crucial in genome protection against the immune system of the host cell. Fundamental aspect for stabilization of viral RNA is the formation of an RNA cap, which is mediated by 2'-O-RNA methyltransferase (MTase), which attached 7-methylguanosine nucleotide to 5'triphosphates end (m7GpppNm). Activity of 2'-O-MTase is associated with non-structural protein 16 (nsp16), which requires the presence of nsp10 as a cofactor.

For structure determination of the nsp10-nsp16 protein complex, nsp10 and nsp16 proteins were prepared recombinantly by expression in bacteria *Escherichia coli*. To obtain the crystal structure the protein complex was inhibited with a small molecule inhibitor TO383. The acquired structural data can be helpful in the development of antiviral drugs based on methyltransferase inhibition and viral RNA destabilization.

Key words:

SARS-CoV-2, methyltransferase, nsp16, nsp10, structure, inhibition

Seznam zkratk a symbolů

(v/v)	Objemová procenta
(w/v)	Hmotnostní procenta v objemu
Å	Angström, jednotka délky $1 \text{ Å} = 10^{-1} \text{ nm}$
ADAM17	Metalopeptidázové domény 17 skupiny ADAM
ACE2	Angiotensin konvertující enzym 2, z angl. angiotensin-converting enzyme 2
<i>E.coli</i>	bakterie <i>Escherichia coli</i>
ELFO	elektroforéza
DNA	Deoxyribonukleová kyselina, z angl. Deoxyribonucleic acid
DMSO	Dimethylsulfoxid
DPP4	Dipeptidylpeptidáza 4
GTP	Guanosintrifosfát
kb	Jednotka délky nukleových kyselin, z angl. kilobase
M	koncentrace v mol/l (molarita)
$\text{Mg}(\text{COO})_2 \cdot 4\text{H}_2\text{O}$	Tetrahydrát octanu hořečnatého
MERS-CoV	Middle East respiratory syndrome coronavirus
MOPS	3-(N-morpholino)propansulfonová kyselina
MTasa	Methyltransferasa
NiCo21 (DE3)	Bakteriální kmen <i>E.coli</i>
NaH_2PO_4	Dihydrogenfosforečnan sodný
Nsp	Nestrukturní protein, z angl. non-structural protein
KH_2PO_4	Dihydrogenfosforečnan draselný
hACE2	Lidský ACE2
hCoVs	Lidské koronaviry, z angl. Human coronaviruses
HEPES	4-(2-hydroxyethyl)-1-piperazinethansulfonová kyselina
His-tag	Polyhistidinová kotva
OD_{600}	Optická hustota při vlnové délce 600nm
ORF	Otevřený čtecí rámec, z angl. Open Reading Frame
PEG	Polyethylen glykol
PDB	Proteinová databáze, z angl. Protein Data Bank

pH	Vodíkový exponent
pp	Polyprotein
ppp	Trifosfát
RBD	Doména vázající se na receptor, z angl. Receptor-Binding Domain
rpm	Otáčky za minutu, z angl. revolution per minute
RNA	Ribonukleová kyselina, z angl. Ribonucleic acid
SAH	S-adenosyl homocystein
SAM	S-adenosyl methionin
SARS-CoV	Severe acute respiratory syndrome coronavirus
SDS	Dodecylsírán sodný
ssRNA	jednovláknová RNA, z angl. single-stranded RNA
SUMO	protein SUMO, z angl. small ubiquitin-like modifier
TO383	Označení organické molekuly syntetizované na ÚOCHB AV ČR v.v.i.
TMPRSS-2	Transmembránová serin proteáza-2 TMPRSS-2
WHO	Světová zdravotnická organizace, z angl. World health organisation

Obsah

1	Literární přehled	9
1.1	Koronaviry.....	9
1.1.1	Infikování hostitele.....	9
1.1.2	SARS-CoV.....	10
1.1.3	MERS-CoV.....	11
1.1.4	SARS-CoV-2.....	12
1.2	Methyltransferasy.....	15
1.2.1	SARS-CoV-2 methyltransferasy.....	16
2	Cíl práce	20
3	Materiál a metody	21
3.1	Laboratorní technika a pomůcky.....	21
3.2	Použité chemikálie.....	22
3.2.1	Roztoky, pufrы a média.....	23
3.2.2	Složení gelů.....	24
3.2.3	Enzymy, plazmidy a bakteriální kmeny.....	24
3.3	Metody.....	25
3.3.1	Expresе a purifikace proteinů nsp10 a nsp16 SARS-CoV-2.....	25
3.3.2	Krystalizace proteinového komplexu.....	27
4	Výsledky	29
4.1	Expresе a purifikace nsp10 a nsp16.....	29
4.2	Krystalizace proteinového komplexu nsp10/nsp16 s ligandem TO383.....	31
4.3	Krystalová struktura proteinového komplexu nsp10/nsp16 s ligandem TO383.....	33
5	Diskuse	37
6	Závěr	39
7	Seznam použité literatury	40

1 Literární přehled

1.1 Koronaviry

Koronaviry patří do rodiny *Coronaviridae* a podrodiny *Coronavirinae* se zabalenou RNA, které jsou široce distribuované mezi ptáky a savci [1,2]. Primárně způsobují respirační a střevní choroby. V ojedinělých případech mohou vyvolat neurologické onemocnění nebo hepatitidu. Šíření je zprostředkováno na mezidruhové úrovni a infekce může být akutní, ale i chronická. Infekce je přenášena především respirační a fekálně – orální cestou [1]. Koronaviry nebyly považovány za vysoce patogenní až do vypuknutí epidemie SARS (severe acute respiratory syndrome) v roce 2002 a 2003 v čínské provincii Guangdong. Do té doby koronaviry cirkulovaly v lidské populaci a způsobovaly víceméně mírné infekce. O 10 let později se v krajinách středního východu objevil vysoce patogenní koronavirus Middle East respiratory syndrome coronavirus (MERS-CoV) [2].

Značná odlišnost této rodiny virů je jejich genomová velikost. Koronaviry jsou nositelem největšího genomu v rámci všech RNA virů, včetně RNA virů se segmentovaným genomem. Tato obsáhlá kódovací kapacita poskytuje mnoho strategií pro genomovou expresi, ze kterých ne všechny jsou úplně známé [1]. Koronaviry vytvářejí zabalené a sférické částice o průměru 100-160 nm. Jejich genom tvořený jedno-vláknovou RNA (ssRNA, z angl. single-stranded RNA) s pozitivní orientací dosahuje velikosti 27–32 kb. Dvě třetiny RNA z 5'-konce kódují polyprotein pp1ab, který je štěpen na 16 nestrukturních proteinů s transkripční a replikační funkcí. 3'-konec RNA obsahuje genetickou informaci pro čtyři strukturní proteiny tvořící obal virové RNA – proteiny S („*spike*“), M („*membrane*“), E („*envelope*“) a N („*nucleocapsid*“). Kromě genů kódujících strukturní proteiny jsou zde i doplňkové, druhově specifické geny neesenciální pro replikaci viru [2].

1.1.1 Infikování hostitele

Většina koronavirů způsobuje onemocnění jen u specifického druhu. Přesto může nakažení jiného druhu vyústit v omezenou virovou replikaci, obvykle bez onemocnění. Koronaviry způsobují převážně respirační a střevní potíže, nicméně mohou také způsobit hepatitidu, infekční peritonitidu, nefritidu, myokarditidu, případně neurologické, reprodukční nebo imunologické poruchy [3]. Schopnost koronavirů adaptovat se a nakazit jiné druhy může

z části pramenit z širokého spektra vstupu viru do buňky a její infikování. *In vivo*, koronaviry se vážou na apikální straně respiračních a střevních epiteliálních buněk a jsou uvolňovány z apikální strany, bazolaterální membrány nebo obou, v závislosti na druhu viru. [4]. Samotný vstup viru do buňky je kontrolován povrchovým spike glykoproteinem, jakožto nositelem schopnosti vázat se na cílový receptor a procházet membránou. Po navázání je virus přenesen receptorově kontrolované endocytózy využitím klatrin nebo caveolin dependentní dráhy. Byly však zaznamenány i jiné vstupné dráhy. Endocytóza za využití S proteinu je proteolyticky závislý proces. K fúzi viru s buňkou je také nutná konformační změna komplexu spike proteinu s receptorem a možná acidifikace endosomálních vezikul [5]. Koronavirový S glykoprotein patří do I. třídy virových fúzních proteinů. Aktivace S proteinu je často doprovázena proteolytickými procesy hostitelskými proteázami, která je typická pro virové fúzní proteiny I. třídy. Mnoho proteáz se prokazatelně potvrdilo v roli koronavirové aktivace [4]. Vskutku, inhibice intracelulárních proteáz způsobila zablokování vstupu viru a fúzi [5]. Pro úspěšnou infekci koronaviry jakožto obalenými viry je vedle navázání na hostitelský receptor dalším důležitým krokem splynutí virového obalu s buněčnou membránou. Fúze virového obalu s membránou vyvolá uvolnění virálního genomu do cytoplazmy hostitelské buňky a umožní replikaci viru [3].

1.1.2 SARS-CoV

Lidské koronaviry HCoV (z angl., Human coronaviruses) podléhají rychlé evoluci díky délce svého genomu a časté rekombinaci. V posledních 20 letech se zvýšil výskyt HCoV v různých částech světa a jsou spojeny se vznikem fatálních pneumonií. První záznam o vypuknutí onemocnění severe acute respiratory syndrome (SARS) se datuje do listopadu 2002 v čínském městě Foshan. To přerostlo v globální epidemii v roce 2003 s celosvětovou 10% smrtností [6]. SARS-CoV patří do druhu *Betacoronavirů* z B linie [7].

Genom SARS-CoV obsahuje základní segment, který sdílí s ostatními koronaviry, a také dodatečných 8 otevřených čtecích rámců (ORFs, z angl. open reading frames). Těchto 8 čtecích rámců kóduje proteiny, které jsou specifické pro SARS-CoV a nevykazují homologii s doplňkovými proteiny jiných koronavirů. Zpočátku byly doplňkové proteiny identifikovány jako pomocné ve virové replikaci *in vitro*. Ukázalo se však, že některé proteiny se účastní procesu interakce viru a hostitele během infekce *in vivo*. Některé studie naznačily,

že této proteiny poskytují viru biologické výhody při infekci hostitelské buňky, což může způsobovat patogenitu SARS-CoV [8].



Obr.1.1.2: Schématický diagram genomové organizace SARS-CoV. Od 5' konce ORFs zodpovědné za transkripci/replikaci (ORF1a, ORF1b), strukturní proteiny S („spike“), E („envelope“), M („membrane“), N („nucleocapsid“) a osm dodatkových proteinů (3a, 3b, 6, 7a, 7b, 8a, 8b, 9) [8].

Receptorem viru SARS-CoV je protein ACE2 (z angl., angiotensin-converting enzyme 2), Tento receptor je široce rozšířen v dýchacích cestách, a to na epitelálních buňkách průdušnice, průdušek, bronchiálních serózních žláz a alveolů, stejně jako na alveolárních monocytech a makrofázích. SARS-CoV atakuje tyto buňky pomocí S1 podjednotky spike proteinu v níž se nachází receptor-vázající doména RBD (z angl. Receptor Binding Domain), která se váže na tento receptor [6].

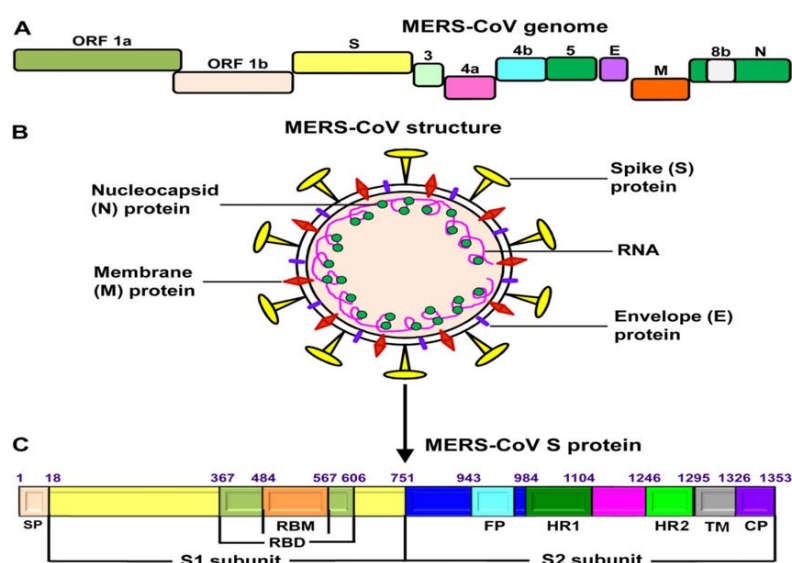
1.1.3 MERS-CoV

V červenci 2012, deset let po první zmínce SARS-CoV, byl izolován nový patogenní koronavirus z hlenu 60 ročního pacienta v Jeddah, Saudské Arábií. Tento pacient po jedenácti dnech od přijetí do nemocnice zemřel na pneumonii a selhání ledvin. U tohoto koronaviru byli zaznamenány také nosokomiální infekce. Volný pohyb poté přispěl k rozšíření viru do jiných krajín arabského poloostrova. Z toho důvodu byl později pojmenován Middle East Respiratory Syndrome coronavirus (MERS-CoV). Celkově bylo v období od července 2012 do 16. října 2018 zaznamenáno 2254 potvrzených laboratorních případů s alespoň 800 úmrtí ve více než 27 krajínách Středního Východu, Severní Afriky, Evropy a Asie. To představuje mortalitu vyšší než 35 % [9, 10].

MERS-CoV patří do linie C *Betacoronavirů* a obsahuje genom o velikosti přibližně 30 kb. Podobně jako u jiných koronavirů, první dvě třetiny z 5' konce genomu kódují proteiny pro replikaci a transkripci, zbylá třetina kóduje strukturní proteiny [10]. Na rozdíl od jiných betakoronavirů, MERS-CoV genome neobsahuje hemagglutinin-enterasu. MERS-CoV

obsahuje 5 doplňkových proteinů, které podobně jako u SARS-CoV mohou zlepšit infekčnost viru. Těto rozdíly mohli zapříčinit větší citlivost pro indukci a signalizace interferonů I. typu v imunitní odpovědi MERS-CoV oproti SARS-CoV [9].

Podobně jako při SARS-CoV, spike protein MERS-CoV se štěpí na receptor-binding S1 doménu a membrane-fusion S2 doménu. Pro navázání viru na hostitelskou buňku je zde také přítomná RBD na S1 doméně. MERS-CoV se váže na jiný cílový receptor. S1 doména interaguje s enzymem DPP4, který se nachází na povrchu většiny buněk. Sice RBD SARS-CoV a MERS-CoV sdílejí vysoký stupeň strukturní podobnosti v jádru subdomén, avšak liší se svou funkcí, což vysvětluje rozdílné cílové receptory zmíněné výše.



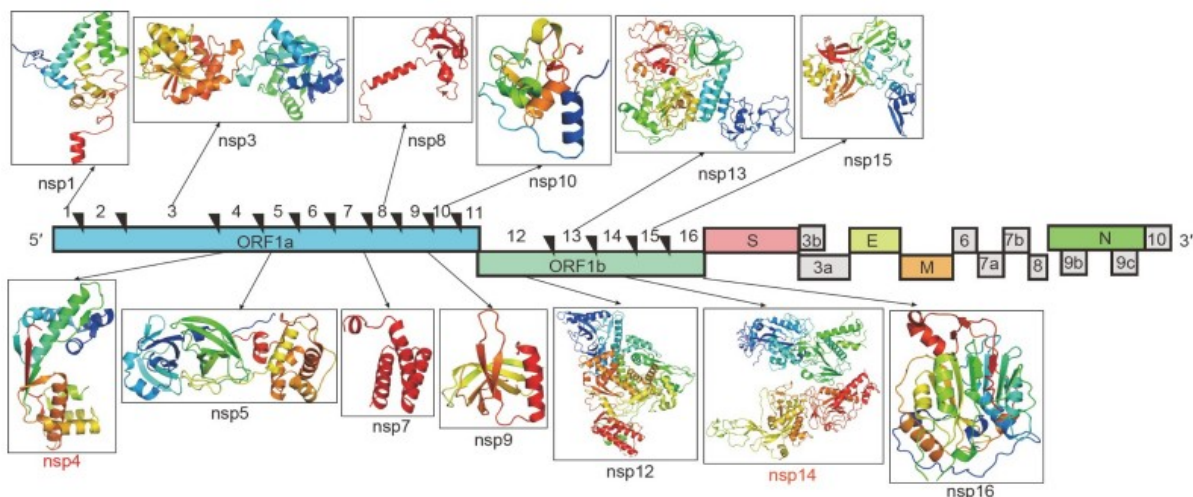
Obr.1.1.3: Schématický diagram genomu MERS-CoV (A). Od 5'konce (zleva) virální funkční proteiny a jiné proteiny s neznámou funkcí (ORF 1a a 1b), strukturní spike protein (S), obal (E), membrána (M), nukleokapsid (N) a pět dodatkových proteinů (3, 4a, 4b, 5, 8b). **Schématická struktura hlavních virálních proteinů (B).** **Schématická struktura S proteinu (C),** rozděleného na S1 a S2 podjednotku [11].

1.1.4 SARS-CoV-2

První záznam o severe acute respiratory syndrome koronaviru 2 (SARS-CoV-2) byl evidován 31. prosince 2019 v čínském městě Wuhan, a to Čínským Ministerstvem Zdravotnictví a Světovou Zdravotnickou Organizací (WHO) [12]. 30. ledna 2020 byl zapsán jako „ohrožení veřejného zdraví mezinárodního významu“ a 11. března 2020 byl změněn WHO

na „Pandemii“. Neobvyklá virální pneumonie nazvaná „Coronavirus disease 2019“ (COVID-19) začala být tedy hrozbou pro celosvětovou populaci. K 29. dubnu 2022 zaznamenala WHO více než 510,2 miliónů potvrzených nakažených a 6,2 miliónů úmrtí. SARS-CoV-2 vícenásobně překonal MERS-CoV i SARS-CoV v počtu nakažených i rozsahu zasažených oblastí [12,13,14].

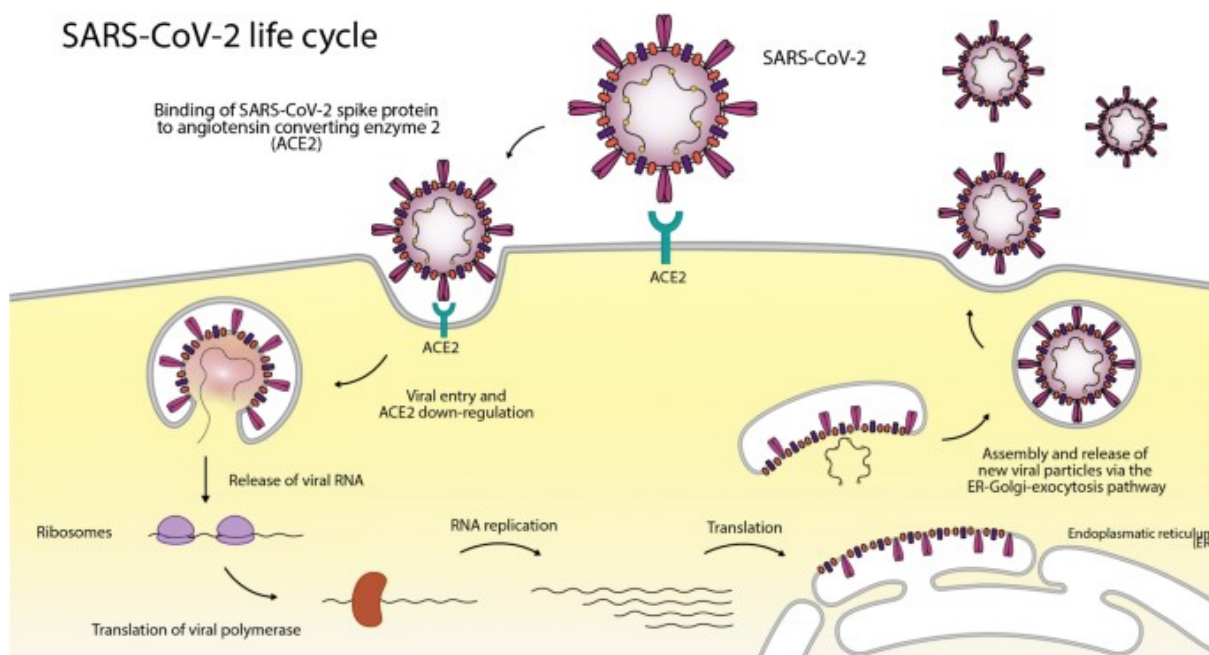
SARS-CoV-2 ze skupiny *Betakoronavirů* z B linie sdílí 79% genomu se SARS-CoV, 50% genomu s MERS-CoV a přibližně 70-80% genomu s různými netopýřními koronaviry. Obsahuje ssRNA genom s pozitivní orientací o velikost přibližně 30 kb. Většina proteinů kódovaných v genomu SARS-CoV-2 má podobnou délku jako příslušné proteiny v SARS-CoV. Aminokyselinová podobnost ve čtyřech strukturních proteinech je více než 90 % se SARS-CoV, liší se pouze v genu S. V prvních dvou třetinách genomu od 5'-konce jsou zakódovaný polyprotein pp1a a pp1ab. Tyto polyproteiny jsou proteolyticky štěpeny na 16 nestrukturních proteinů (nsps z angl., non-structural proteins). Nestrukturní proteiny jsou regulátory replikace a transkripce RNA, a také se podílejí na samotné genomové replikaci a transkripci jakožto podjednotky v transkripčních/replikačních komplexech. Se SARS-CoV má více než 85% podobnost aminokyselinové sekvence [13,15]. Genom pozůstává z 14 funkčních ORFs, včetně strukturních proteinů, dvou nekódujících oblastí na obou koncích, nsps a přídatných proteinů. ORF1a a ORF1b kóduje 16 nsps (nsp1-nsp16). Dále tu jsou čtyři strukturní proteiny, spike protein (S), membránový protein (M), envelope protein (E) a nucleocapsid protein (N), důležité pro složení viru. Devět doplňkových proteinů mají selektivní výhodu v infikovaném hostiteli [13].



Obr.1.1.4a: Schématický diagram genomu SARS-CoV-2 se strukturou 16 nestruturních proteinů. Od 5' konce (zleva) ORF1a kódující 11 NSPs (*nsp1-nsp11*), ORF1b kódující 4 *nsp*s (*nsp12-nsp16*), čtyři strukturní proteiny (*S*, *E*, *M* a *N*) a šedě vyznačených devět dodatkových proteinů [13].

S protein, zodpovědný za navázání viru a membránovou fúzi, se skládá ze dvou podjednotek S1 a S2. Svou S1 podjednotkou, která obsahuje RBD se váže na lidský ACE2 (hACE2, z angl. human angiotensin-converting enzyme) receptor na membráně buňky. RBD SARS-CoV-2 má přibližně o 10 až 20-krát vyšší afinitu k hACE2 než SARS-CoV, a také se silněji váže k rozpustnému hACE2. To může mít za následek vyšší infekčnost.

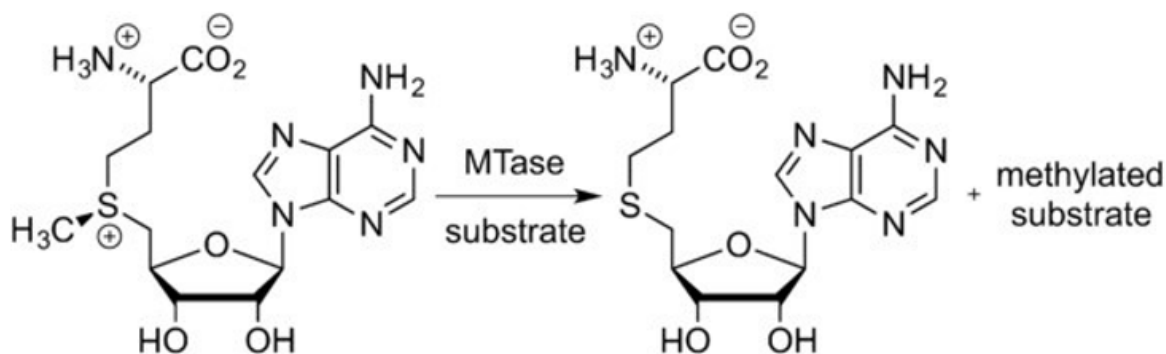
K virální fúzi s hostitelskou membránou pomocí S2 podjednotky je potřebná přítomnost transmembránové proteasy serin proteasa-2 (TMPRSS-2) a ADAM metalopeptidázové domény 17 (ADAM17). Tudíž, vir vstupuje do buňky pomocí endocytózy a vypouští svou RNA pro replikaci a translaci, za využití hostitelského mechanismu s následnou produkcí nových virových částic, které jsou externalizované exocytózou [16].



Obr. 1.1.4b: Životní cyklus SARS-CoV-2. SARS-CoV-2 se váže na receptor ACE2 nacházejícího se na membráně buňky. Následná fúze s membránou vyústí k uvolnění virové ssRNA a replikaci genomu. Nově replikovaná RNA je translatována a vznikají proteiny potřebné pro vytvoření nových virových částic [16].

1.2 Methyltransferasy

Biologické systémy jsou složité komplexy různých komponent, které mezi sebou interagují a dávají tak možnost vzniku nových vlastností organismů. Nicméně, i jediná změna v molekule může mít velký dopad na celý organismus. Malá modifikace DNA, RNA nebo interagujících proteinů v epigenetické regulaci genové aktivity je provázána změnou v biologické aktivitě. Jednou z epigenetických regulací je enzymatická aktivita S-adenosyl-L-methionine (SAM) dependentních MTase. Hlavní úlohou, těchto enzymů je metylace od malých molekul, až po proteiny a oligonukleotidy. Při této reakci dochází ku katalytickému kovalentnímu přenosu methylové skupiny z donoru SAM na cílovou látku. Aby k tomu došlo, methyltransferázy tvoří komplex s cílovou molekulou a SAM. Komplex tří molekul zahrnuje mnoho intermolekulárních interakcí, kterých principy jsou podobné pro všechny systémy. Přenosná metylová skupina SAM je vázaná na sulfonovém centru spájejícím methioninový zbytek se sacharidovým. Molekula je náchylná na nukleofilní atak a metylová skupina lehko podléhá substituční reakci [17].



Obr.1.2: Přenos metylové skupiny pomocí katalytické aktivity methyltransferasy. V levé části reakce se nachází SAM s navázanou $-CH_3$ skupinou na síře. Napravo je vzniklý SAH s methylovaným substrátem [17].

1.2.1 SARS-CoV-2 methyltransferasy

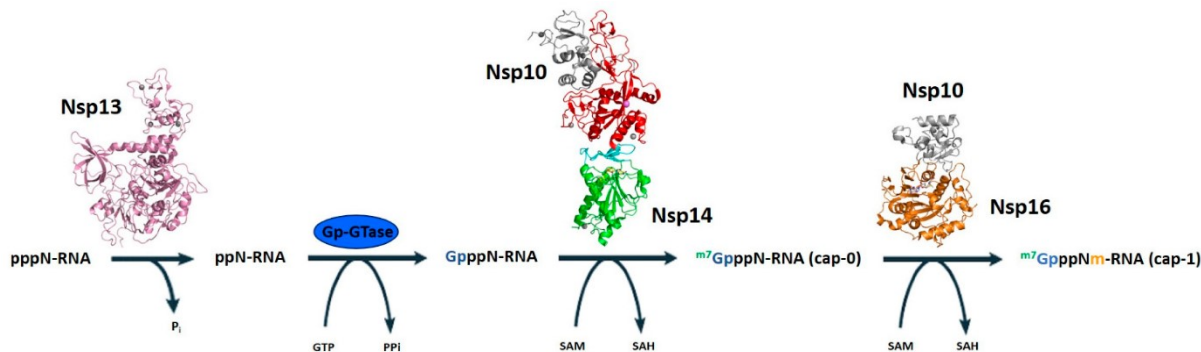
Koronaviry patří do skupiny virů s jedno-vláknovou RNA s pozitivní orientací, které replikují v hostitelské cytoplazmě. K replikaci dochází přes RNA replikační/transkripční systém a 16 nestrukturních proteinů. Mezi tyto proteiny patří také nsp16/nsp10 proteinový komplex, který má za úlohu metylaci předposledního nukleotidu virální RNA čepičky na pozici 2'-O ribózy [18].

1.2.1.1 Mechanismus tvorby RNA čepičky a její metylace

Koronaviry jsou chráněny čepičkou na svém 5'-konci RNA. Táto čepička je tvořená z N7-methylované GTP molekuly připojené k prvnímu transkribovanému nukleotidu pomocí 5'-5' trifosfátové vazby. K syntéze RNA čepičky slouží C-konec proteinového komplexu nsp14 a nsp10. Ta vykazuje (N7-guanine)-methyltransferasovou aktivitu. RNA čepička-0 (m^7GpppN -RNA) je složená z 7-methylguanozinu (m^7G) připojeného přes tři fosfát (ppp) k 5'-nukleozidu (N) RNA.

Proces tvorby čepičky začíná hydrolýzou 5' γ -fosfátu vznikající RNA (pppN-RNA) za účasti nsp13 helikasy, která slouží jako RNA 5' trifostasa. Následně Guanintransferasa (GTase) přesune GMP molekulu na 5' difosfátový konec RNA (ppN-RNA), který vede k tvorbě GpppN-RNA. Methylace na N7 pozici guanosinu probíhá prostřednictvím proteinu nsp14 a svou N7-methyltransferasovou domény formuje čepičku-0 (m^7GpppN -RNA). Poslední část je metylace 2'-O ribózy na prvním transkribovaném nukleotidu skrze nsp16 za vzniku finální

čepičky-1 (m^7GpppN_m -RNA). RNA čepička je důležitá z hlediska stability RNA a brání před rozpoznáním imunitním systémem hostitelské buňky. Mechanismus tvorby RNA čepičky je proto zajímavým cílem pro konstrukci anti virálních léků [19].

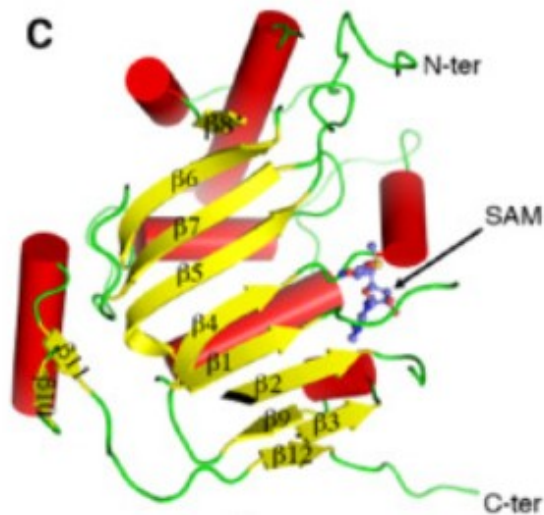


Obr.1.2.1.1: Schéma syntézy RNA čepičky v SARS-CoV-2. Syntéza je zajištěna pomocí čtyř enzymů: *nsp13* (růžová), *GTase* (modrá), *nsp14* (zelená) a *nsp16* (oranžová). *nsp10* přítomný v *nsp14* a *nsp16* slouží jako kofaktor [19].

1.2.1.2 Struktura *nsp10/nsp16* methyltransferasy a vazebného místa pro SAM

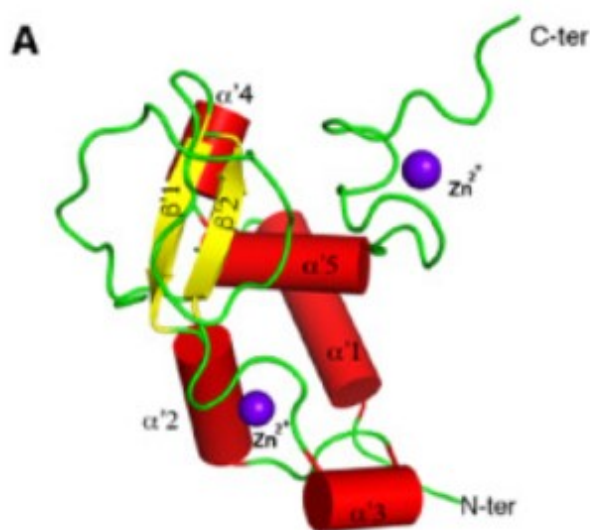
Proteiny *nsp10* a *nsp16* jsou zakódovány v ORF1ab části genu. Proteasa *nsp5* slouží k odštěpení proteinů z polyproteinů *pp1a* a *pp1ab*. *Nsp10* je štěpena z *pp1a* a *pp1ab*, zatímco *nsp16* se uvolňuje jen z *pp1ab*. Samostatný protein *nsp16* je nestabilní, proto s *nsp10* tvoří heterodimer. *Nsp10* zde působí jako kofaktor.

Nsp16 pozůstává z 298 aminokyselinových zbytků (ORF1ab 6799-7096) a molekulová hmotnost je rovná 33 323 Da. Katalytická 2'-O MTasa část se nachází na proteinu *nsp16* a je složená dvanáct β -listů a sedm α -helixů [20, 21]. V centrální části se nachází Rossman-like- β -sheet záhyb pozůstávající z jedenácti α -helixů a sedmých β -listů. Tyto β -listy mají kónické rozložení ve formátu 3-2-1-4-5-7-6, ve kterém jediným antiparalelním vláknem je $\beta 7$. β -listy jsou obklopeny α -helixy a smyčkami [20].



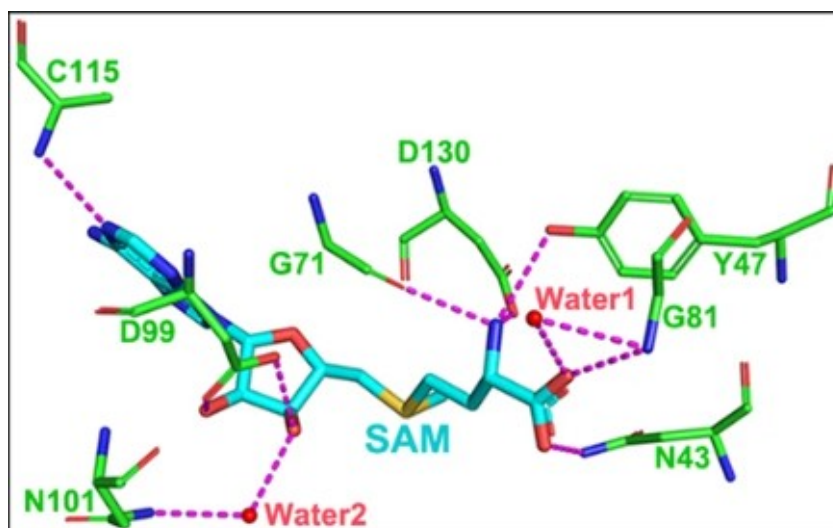
Obr. 1.2.1.2a: Sekundární struktura nsp16 s Rossman-like- β -sheet záhybem. Žlutou barvou jsou vyznačeny β -listy, červenou barvou α -helixy a zelenou barvou zbytek řetězce [20].

Protein nsp10 je složen z 133 aminokyselinových zbytků a jeho molární hmotnost je 14 789 Da. Povrch tohoto proteinu je hydrofobní a pozitivně nabitý, a tedy dokáže interagovat s negativně nabitým hydrofobním povrchem nsp16. Táto interakce vede k stabilizaci SAM vazebného místa. Centrální část je tvořena dvojicí antiparalelních β -listů obklopených α -helixy z jedné strany. Dvě ionty zinku se nacházejí v helikální doméně. První je koordinován přes Cys-74, Cys-77, His-83 a Cys-90. Druhý zinek je koordinován přes Cys-117, Cys-120, Cys-128 a Cys-130 [20].



Obr. 1.2.1.2b: Sekundární struktura *nsp10*. α -helixy jsou vyznačeny červenou barvou a β -listy barvou žlutou. Zeleně je vyznačen řetězec [20].

SAM vazebné místo nacházející se na *nsp16* je tvořeno aminokyselinami Asn-43, Tyr-47, Gly-71, Gly-81, Asp-99, Asn-101, Cys-115 a Asn-130. Této rezidua spolu se sousedními molekulami vody tvoří vodíkové vazby a zachytávají SAM. Rezidua zapojující se do vazby SAM jsou konzervativní pro SARS-CoV i MERS-CoV. Z tohoto důvodu se naskytuje možnost vytvořit antivirotikum na bázi inhibice SAM vazebního místa [21].



Obr. 1.2.1.2c: Interakce SAM s aminokyselinami na proteinu *nsp16*. Modrou barvou je zvýrazněn SAM, zelenou barvou interagující aminokyseliny. Červené tečky vyznačují molekuly vody, červené části na molekulách představují kyslíky a modře jsou označeny dusíky [21].

2 Cíl práce

V posledních desítkách let začaly být koronaviry hrozbou pro lidské zdraví v globálním měřítku. Doposud nebyla nalezena vhodná antivirotika pro léčbu. SAM-dependentní methyltransferasy jsou důležité pro replikaci a chrání svou RNA před imunitním systémem hostitele. Pro svou vysoce konzervativní strukturu v aktivním místě jsou vhodným kandidátem pro hledání možných inhibitorů.

V této práci byly připraveny rekombinantní proteiny nsp10 a nsp16 z koronaviru SARS-CoV-2 bakteriální expresí a jejich následná purifikace. Důležitou částí bylo vytvoření krystalů komplexu proteinu s malými molekulami, které jsme mohli analyzovat difrakčním experimentem a vytvořit strukturní model.

3 Materiál a metody

3.1 Laboratorní technika a pomůcky

Centrifuga 5418 R, rotor FA-45-18-11	Eppendorf (DE)
Centrifuga Allegra X-15 R, rotor SX4750A	Beckman Coulter (USA)
Centrifuga Avanti J-26S XPI, rotor JLA-9.1000	Beckman Coulter (USA)
Centrifuga Avanti J-30I, rotor JA-30.50 Ti	Beckman Coulter (USA)
Dounceho homogenizátor	P-Lab (ČR)
Chromatografický systém ÄKTA pure	Cytiva (USA)
CrystalGryphon	Art Robbins Instruments (USA)
Kolona HiLoad 16/600 Superdex 200 pg	GE Healthcare (USA)
Koncentrátor Vivaspin Turbo 4 (MWCO 10 000)	Sartorius AG (DE)
Koncentrátor Vivaspin Turbo 4 (MWCO 3 000)	Sartorius AG (DE)
Krystalizační destičky Intelli-Plate 96-2 Shallow Well	ARI (USA)
Krystalizační robot Mosquito Crystal	TPP Labtech (GB)
Magnetická míchačka KMO-2 Basic	IKA (USA)
Mikroskop SZX10	Olympus (JP)
Mikroskop UV fluorescenční Uvex-M	JANSi (USA)
Míchačka válcová TubeRoller R3005	Benchmark Scientific (USA)
Předvážky EMB 2000-2	KERN & SOHN (DE)
Rentgenová difrakční stanice	
MicroMax TM -007 HF	Rigaku (USA)
Goniometr AFC11 Partial- χ 4-axis Goniometer	Rigaku (USA)
Optika Confocal X-RAY Optical assembly VariMax	Rigaku (USA)
Detektor PILATUS 300K	Dectris (USA)
Chlazení Cryostream 800	Oxford Cryosystems (UK)
Sběrač frakcí Fraction Collector F9-C	Cytiva (USA)
Sonikátor Q700	QSonica (USA)
Sonikační sonda CL-334	QSonica (USA)
Spektrofotometr BioPhotometer plus	Eppendorf (Německo)
Spektrofotometr Nanodrop ND-1000	Thermo Scientific (USA)
Třepačka Incubating Mini Shaker	VWR (ČR)

Třepačka s inkubátorem Innova 44	New Brunswick Scient. (USA)
Vodní lázeň JB Aqua 2 Plus	Grant Instruments (UK)
Vortex Genie 2	Scientific Industries (USA)
Zdroj pro SDS elektroforézu PowerPac HV Basic	Bio-rad (USA)

3.2 Použité chemikálie

β – merkaptoenatol (β -ME)	Carl Roth (DE)
Akrylamid / Bisakrylamid	Carl Roth (DE)
Sodná sůl ampicilinu (AMP)	Carl Roth (DE)
Coomassie Brilliant Blue G 250	Sigma-Aldrich (USA)
Chlorid sodný (NaCl)	Penta (ČR)
Dialyzační membrána	
Membra-Cel MD34 (MWCO 14000)	Carl Roth (DE)
Dimethylsulfoxid (DMSO)	Sigma-Aldrich (USA)
Ethylendiamintetraoctová kyselina (EDTA)	Carl Roth (DE)
Glycerol	Lach-Ner (ČR)
Octová kyselina (CH ₃ COOH)	Penta (ČR)
Dodecylsírán sodný (SDS)	Carl Roth (Německo)
Sírán hořečnatý (MgSO ₄)	Penta (ČR)
Standard SDS elektroforézy	
PageRuler Plus (10-250 kDa)	Thermo Scientific (USA)
Talon Ni-NTA Resin	Thermo Scientific (USA)
Peroxodisírán amonný (NH ₄) ₂ S ₂ O ₈	Carl Roth (DE)
TEMED	Carl Roth (DE)
Tris(2-carboxyethyl) fosfin hydrochlorid (T-CEP)	Sigma-Aldrich (USA)
Trizma base (Tris-Cl)	Sigma-Aldrich (USA)

3.2.1 Roztoky, pufrů a média

Transformace plazmidové DNA a bakteriální exprese proteinů

- 1000x koncentrovaný roztok ampicilinu AMP: 100 mg/ml AMP (vodný roztok), 50% (v/v) ethanol
- LB médium dle Lennox: 0,5% (w/v) kvasničný extrakt, 1 % (w/v) trypton, 85 mM NaCl
- ZY médium: 1% (w/v) trypton, 0,5% (w/v) kvasničný extrakt, 2,8 mM Glukóza, 6 mM laktóza, 0,5% (v/v) glycerol, 1 mM MgCl₂, 25 mM (NH₄)₂SO₄, 50 mM KH₂PO₄, 50 mM Na₂HPO₄
- 20x koncentrovaný fosfátový pufr: 1 M Na₂HPO₄ · 12 H₂O, 1M KH₂PO₄, 0,5 M (NH₄)₂SO₄
- 50x koncentrovaný 5052 pufr (+100 mg MgCl₂): 25% glycerol, 2,5% glukóza, 10% laktóza

Izolace proteinů

- Lyzační pufr: 50 mM Tris-HCl pH 8,0; 500 mM NaCl; 3 mM β – ME; 10% glycerol (v/v), 20 mM imidazol, 5 mM MgSO₄

Afinitní chromatografie na agarózovém nosiči Ni-NTA

- Ekvilibrační pufr: 50 mM Tris-HCl pH 8,0; 500 mM NaCl; 3 mM β – ME; 10% glycerol (v/v); 20 mM imidazol; 5 mM MgSO₄
- Promývací pufr: 50 mM Tris-HCl pH 8,0; 1 M NaCl; 3 mM β – ME; 10% glycerol (v/v); 20 mM imidazol; 5 mM MgSO₄
- Eluční pufr: 50 mM Tris-HCl pH 8,0; 500 mM NaCl; 3 mM β – ME; 10% glycerol (v/v); 300 mM imidazol; 5 mM MgSO₄
- Dialyzační pufr: 50 mM Tris-HCl pH 8,0; 200 mM NaCl; 3 mM β – ME; 10% glycerol (v/v); 20 mM imidazol; 5 mM MgSO₄

Gelová permeační chromatografie

- Ekvilibrační pufr B: 10 mM Tris-HCl pH 8,0; 150 mM NaCl; 5% glycerol (v/v); 1 mM T-CEP

Elektroforéza v polyakrylamidovém gelu

- Tris-glycinový pufr: 25 mM Tris NaCl; 0,25 M glycin; 0,1% (w/w) SDS pH 8,8
- Vzorkový pufr: 125 mM Tris pH 6,8; 20% (v/v) glycerol; 0,2% (v/v) bromfenolová modř; 4% (w/v) SDS; 10% (v/v) β – ME
- Barvicí roztok pro gely SDS-ELFO: 3 mM Coomassie Brilliant Blue, 1,5 M octová kyselina, 50% (v/v) methanol

Komerční sady krystalizačních roztoků

- JCSG I Qiagen (DE)
- JCSG II Qiagen (DE)
- JCSG III Qiagen (DE)
- JCSG IV Qiagen (DE)
- ProPlex Molecular Dimension (USA)

3.2.2 Složení gelů

- Zaostrňovací gel SDS elektroforézy: 125 mM Tris-Cl pH 6,8; 4% (v/v) Akrylamid/Bisakrylamid (37,5:1); 0,1% (w/v) SDS; 0,005% (v/v) TEMED; 0,05% (w/v) $(\text{NH}_4)_2\text{S}_2\text{O}_8$
- Separáčnící gel SDS elektroforézy: 375 mM Tris-Cl pH 8,8; 15% (w/v) Akrylamid / Bisakrylamid (37,5:1); 0,1% (w/v) SDS; 0,001% (v/v) TEMED; 0,05% (w/v) $(\text{NH}_4)_2\text{S}_2\text{O}_8$

3.2.3 Enzymy, plazmidy a bakteriální kmeny

SUMO proteáza s histidinovou kotvou	Připraveno v laboratoři
Plasmid pro expresi koronavirového nsp10 s His _{8x} -SUMO kotvou	Připraveno v laboratoři
Plasmid pro expresi koronavirového nsp16 s His _{8x} -SUMO kotvou	Připraveno v laboratoři
<i>Escherichia coli</i> NiCo21 (DE3)	Připraveno v laboratoři

3.3 Metody

3.3.1 Exprese a purifikace proteinů Nsp10 a Nsp16 SARS-CoV-2

3.3.1.1 Transformace kompetentních bakterií *E.coli* NiCo21(DE3)

Kompetentní bakterie *Escherichia coli*, uskladněné v 1,5 ml mikrozkušavkách při -80°C byly rozmrazeny na ledu. Transformace bakterií proběhla metodou tepelného šoku. K rozmražené 35 μ l suspenzi bylo odděleně přidáno 1,5 μ l plazmidu nesoucího gen pro nsp10 a nsp16. Mikrozkušavky byly uloženy do vodní lázně s teplotou 42 °C po dobu 40 sekund. Poté byly obě zkumavky s bakteriemi ponechány na ledě přibližně 5 minut. K suspenzím bylo přidáno 0,5 ml LB média a zkumavky byly uloženy do inkubační třepačky o teplotě 37°C s otáčkami 220 rpm na 30 minut. Do 13 ml zkumavek bylo mezitím rozpipetováno 5ml LB média a 5 μ l zásobního roztoku AMP abychom dostali výslednou koncentraci 100 μ g/ml. Bakteriální suspenze byly jednotlivě přeneseny do roztoků LB média a inkubovány přes noc v třepačkách za stejných podmínek.

3.3.1.2 Exprese rekombinantních proteinů v buňkách *E.coli* NiCo21(DE3)

Do tří dvoulitrových Erlenmayerových baněk obsahujících 1l ZY média obohaceného o 1x koncertovaný fosfátový pufr a 1x koncertovaný 5052 pufr, bylo přidáno po 1,65 ml inokul nesoucího gen pro Nsp10. K ZY médiu byl přidán 1 ml 1000x koncentrovaného AMP. S buněčnou suspenzí nesoucí gen pro Nsp16 bylo obdobně postupováno. Šest Erlenmayerových baněk bylo uloženo do orbitálních třepaček s teplotou 37 °C a otáčkami 220 rpm. Po přibližně 5 hodinách dosáhly hodnoty optických densit suspenzí při vlnové délce 600 nm (OD_{600}) rozmezí 0,7 – 0,8. Následně byla teplota v třepačce snížena na 18 °C a buňky inkubovány přes noc.

3.3.1.3 Izolace proteinů z buněk *E.coli* NiCo21(DE3)

Bakteriální suspenze byla přelita do 1l kyvety a centrifugována (centrifuga Avanti J-26S XPI, rotor JLA-9.1000) při otáčkách 6000 \times g po dobu 6 minut a při 4 °C. Vzniklý supernatant byl vylit a do kyvety přidána další bakteriální suspenze. Centrifugace probíhala za stejných podmínek. Takto bylo pracováno se všemi bakteriálními suspenzemi, až nám zůstali jednotlivě pelety pro proteiny nsp10 a nsp16. Při práci s peletami i dalším zacházením s proteiny byly vzorky udržovány na ledu nebo v chladu.

Bakteriální pelety byly resuspendovány v asi 30 ml lyzačního pufru za pomoci Vortexu. K homogenizaci pelety byl použit Dounceho homogenizátor. Bakterie byly lyzovány pomocí sonikace (amplituda 70 Hz, doba pulsu 1 sekunda, doba přestávky pulsu 3 sekundy, celkový čas 2 minuty). Lyzované buňky byly převedeny do 50ml zkumavek a centrifugovány (Avanti J-30I, rotor JA-30,50 Ti) po dobu 30 minut při 4 °C. Supernatanty byly jednotlivě přelity do 50ml zkumavek. Po celou dobu jsme zvláště pracovali s proteiny nsp10 a nsp16 až do gelové permeační chromatografie.

3.3.1.4 Afinitní chelatační chromatografie

K purifikaci proteinů nsp10 a nsp16 bylo využito agarozového nosiče Ni-NTA, který se svými Ni²⁺ ionty váže na imidazolový kruh histidinů přítomných na His-tag značce. 6 ml nosiče Ni-NTA bylo individuálně pipetováno a propláchnuto v plastových kolonách destilovanou vodou a ekvilibračním pufrem. Poté byl nosič Ni-NTA resuspendován v 5ml ekvilibračního pufru a souměrně rozdělen do zkumavek se supernatantem. Zkumavky s proteiny byly přesunuty na válcovou míchačku, kde byly inkubovány 30 minut při 4 °C. Zkumavky byly poté centrifugovány (centrifuga Allegra X-15R, rotor SX7450A) po dobu 3 minut při otáčkách 1620 rpm a 4 °C. Supernatanty byly opatrně vylity a do každé zkumavky bylo přidáno zhruba 40 ml promývacího pufru na resuspendaci. Všechny zkumavky byly znovu centrifugovány za stejných podmínek. Po vylití supernatantu bylo přidáno 40 ml ekvilibračního pufru, peleta opět resuspendována a centrifugována za obdobných podmínek. Tento krok byl zopakován 3×. Po poslední centrifugaci byl nosič Ni-NTA resuspendován v 10 ml ekvilibračního pufru a vložen do plastové kolony. Ekvilibrační pufr prokapal přes kolonu a bylo přidáno 4 ml elučního pufru. Eluát byl zachycen do 15ml plastové zkumavky.

3.3.1.5 Dialýza a štěpení poly-histidinové kotvy SUMO proteasou

4 ml eluátu byly převedeny do celulózy dialyzační membrány (MWC 14 000) a byl přidán 100μl alikvot SUMO proteasy. Štěpení probíhalo přes noc při 4 °C uložením membrán do 2l dialyzačního pufru.

3.3.1.6 Reversní afinitní chelatační chromatografie

Důležitým krokem bylo oddělení SUMO proteázy s histidinovou kotvou a odštěpených His_{8x}-tag značek od proteinů. K tomu byla využita reversní afinitní chromatografie na nosiči Ni-NTA. Aktivace nosiče Ni-NTA proběhla pomocí 100 mg/ml koncentrovaného NiCl₂

roztoku a promytím ekvilibračním pufrem. 5 ml ekvilibračního pufru bylo použito k resuspendaci nosiče Ni-NTA a přidáno k proteinu ve zkumavce. Inkubace trvala 30 minut při 4 °C na válcové míchačce. Po vyjmutí z míchačky byl vzorek vložen do plastové kolony a rovnou zachytáván eluát do 15ml zkumavky.

3.3.1.7 Gelová permeační chromatografie

Proteiny nsp10 a nsp16 byly spojeny v koncentrátoru (MWCO 10 000) a koncentrovány na objem pod 5 ml třínásobnou centrifugací (Allegra X-15R, rotor SX4750A) při otáčkách 4000 rpm a 4 °C po dobu 15 minut. Kolona HiLoad 16/600 superdex 200 pg byla připojena na systém ÄKTA a promyta ekvilibračním pufrem B. Před nanesením byl vzorek centrifugován 10 minut při 4500 rpm a 4 °C, kvůli sedimentaci možných nečistot. Následně byly proteiny nastříknuty do 5ml smyčky a jímány 1,5ml frakce.

3.3.1.8 Diskontinuální elektroforéza v prostředí SDS (SDS-PAGE)

Elektroforéza v 15% separačním a 4% zaostřovacím polyakrylamidovém gelu sloužila k analýze jímáných frakcí. Po 2 µl 6x koncentrovaného vzorkového pufru a 10 µl každé vybrané frakce bylo naneseo do jamek gelu. Do první jamky bylo naneseo 5 µl proteinového standardu. SDS-PAGE probíhala 45 minut při konstantním napětí 220 V a laboratorní teplotě.

3.3.2 Krystalizace proteinového komplexu

3.3.2.1 Příprava proteinového komplexu s ligandem

Potřebné koncentrace proteinu přibližně 5 mg/ml bylo dosaženo přenesením vzorku do koncentrátoru (MWCO 10 000) a opakovanou centrifugací při 4500 rpm, 4 °C v 15 minutových intervalech. Absorbance vzorku při 280nm oproti ekvilibračnímu pufru B z gelové permeační chromatografie byla měřena na přístroji Nanodrop ND-1000. Protein byl smíchán s 10mM roztokem ligandu TO383 tak, aby výsledná koncentrace ligandu činila 1 mM.

3.3.2.2 Krystalizace metodou sedící kapky

Přístroj CrystalGryphon byl využit pro pipetování 80 µl komerčních krystalizačních roztoků (JCSG I-III, Proplex) do rezervoárů v krystalizačních destičkách Intelli-Plate 96-2 Shallow Well. Protein s ligandem TO383 byl rozpipetován po 300 nl do krystalizačních jamek

v krystalizační destičce, bylo k němu přidáno 150 nl krystalizačního roztoku z rezervoáru a celá destička byla přelepena páskou. Poměr proteinu s ligandem a krystalizačního roztoku v jamce tedy činil 2:1. Tímto způsobem byly připraveny finální 4 krystalizační destičky.

Krystalizační destičky byly ponechány při 18 °C a každé 2 dny kontrolovány pod optickým mikroskopem SZX10. K rozlišení proteinových krystalů od krystalů soli sloužil UV fluorescenční mikroskop Uvex-M.

3.3.2.3 Měření rentgenové difrakce

Proteinové krystaly byly zachyceny smyčkou a ponořeny do plastových vialek uložených v kazetě (tzv. „*pucku*“) v tekutém dusíku. Zmražený krystal byl přenesen na goniometr v difrakční stanici a analyzován pomocí rotační anody. Krystal byl kontrastně chlazen přívodem par kapalného dusíku.

4 Výsledky

4.1 Exprese a purifikace nsp10 a nsp16

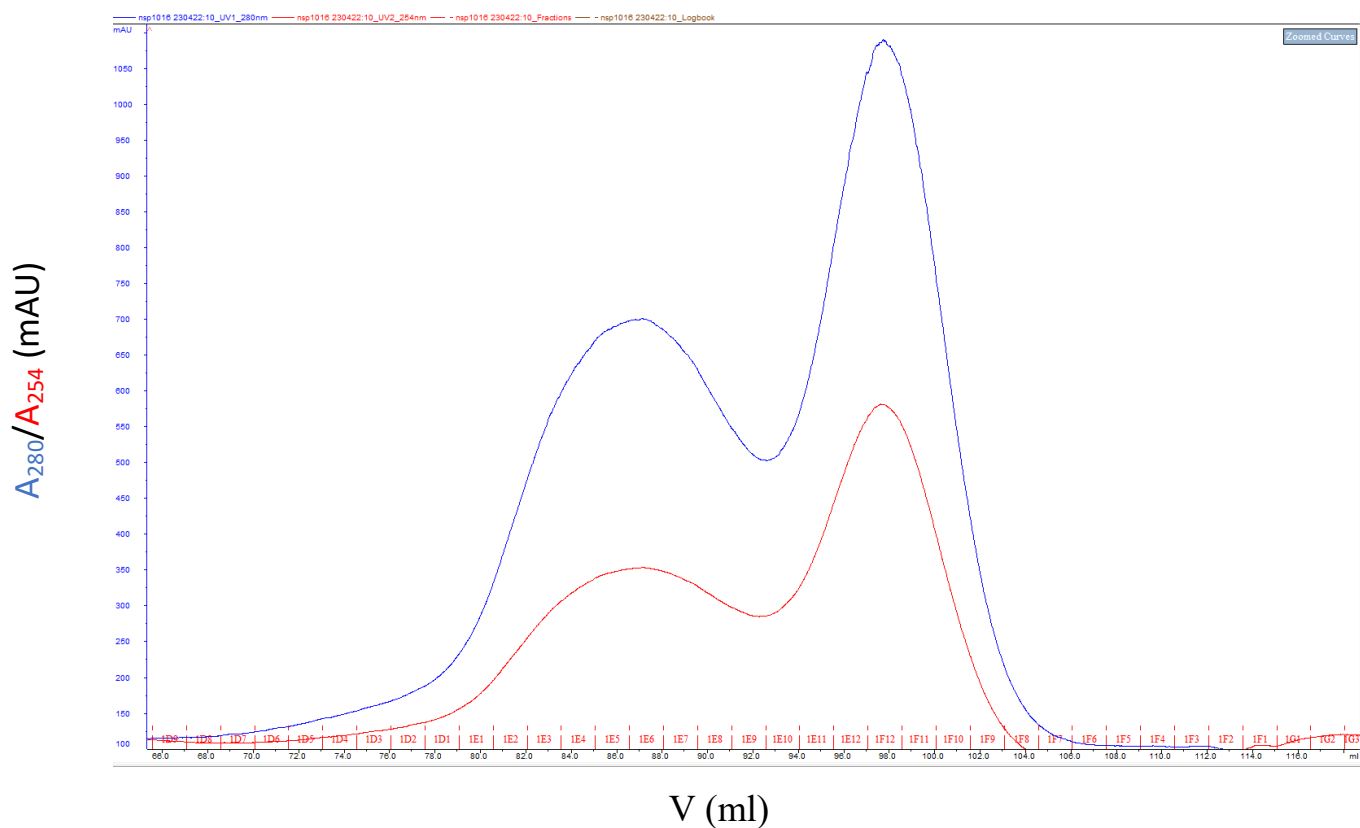
Kompetentní buňky *E.coli* NiCo21(DE3) byly transformovány plazmidem pSUMO nesoucím gen pro nsp10 a nsp16 s poly-histidinovou kotvou His_{8x}-SUMO. Nsp16 je tvořen z 298 aminokyselin s molekulovou hmotností 33 323 Da. Nsp10 je menší protein, skládá ze 133 aminokyselinových zbytků a jeho molekulová hmotnost je 14 789 Da. Exprese těchto proteinů probíhala v autoindukčním médiu ZY.

Izolace proteinů začala lyzí bakteriální buněk a sonifikací. Proteiny s His_{8x}-tag kotvou byly zachyceny na nosiči Ni-NTA a mohla být využita afinitní chelatační chromatografie. Z výsledných eluovaných proteinů byl odebrán vzorek pro SDS-PAGE analýzu.

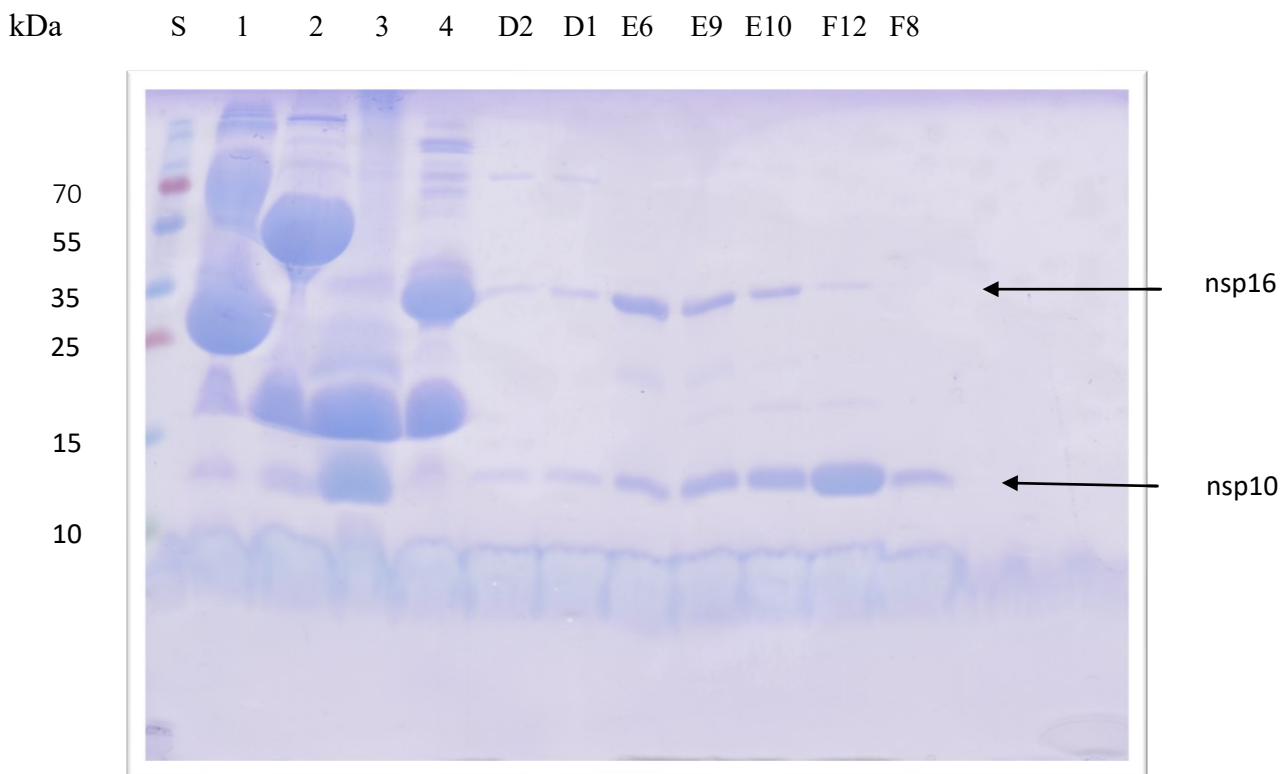
Eluované proteiny byly přeneseny do dialyzačních membrán spolu se SUMO proteázou štěpící poly-histidinovou kotvu. Dialýza probíhala přes noc při 4 °C proti dialyzačnímu pufru. Po ukončení byly odebrány vzorky pro SDS-PAGE analýzu.

Dále byla provedena reversní afinitní chelatační chromatografie, kde byly na nosiči Ni-NTA zachyceny odštěpené His_{8x}-tag značky z rekombinantních proteinů a také SUMO proteáza, která měla rovněž navázanou poly-histidinovou kotvu.

Před gelovou permeační chromatografií byly vzorky nsp10 a nsp16 spojeny a zakoncentrovány pod 5 ml. Vzorek byl nanesen do smyčky a rozdělen do 1,5ml frakcí na základě různé velikosti molekul (Obr. 4.1a, str. 30)



Obr. 4.1a: Chromatogram po gelové permeační chromatografii. Kolona HiLoad 16/600 Superdex 200 pg, 4°C, 1 ml/min. Modrá křivka znázorňuje absorbanci proteinů při vlnové délce 280 nm, červená křivka představuje absorbanci při 254 nm. Ve spodní části jsou červeně vyznačeny frakce a černě je označen eluovaný objem.



Obr. 4.1b: Výsledek analýzy frakcí proteinů *nsp10* a *nsp16* metodou SDS-PAGE v 15% polyakrylamidovém gelu. Dráha S představuje proteinový standard, jehož molekulové hmotnosti jsou uvedeny v jednotkách kDa nalevo. V dráze 1 a 2 jsou proteiny *nsp10* a *nsp16* před dialýzou s navázanou His_{8x}-tag značkou. *Nsp10* + His_{8x}-tag se nachází těsně nad hodnotou 25 kDa, zatímco *nsp16* + His_{8x}-tag je v úrovni 55 kDa. V drahách 3 a 4 se nacházejí *nsp10* a *nsp16* po dialýze. Spodní proužek v dráze 3 představuje *nsp10*, proužek nad ním v rozmezí 15-25 kDa reprezentuje SUMO proteázu. Proužek v 4 dráze pod 35 kDa patří *nsp16* a rovněž v rozmezí 15-25 kDa se nachází SUMO proteáza. D2 – F8 jsou jednotlivé frakce z gelové permeační chromatografie.

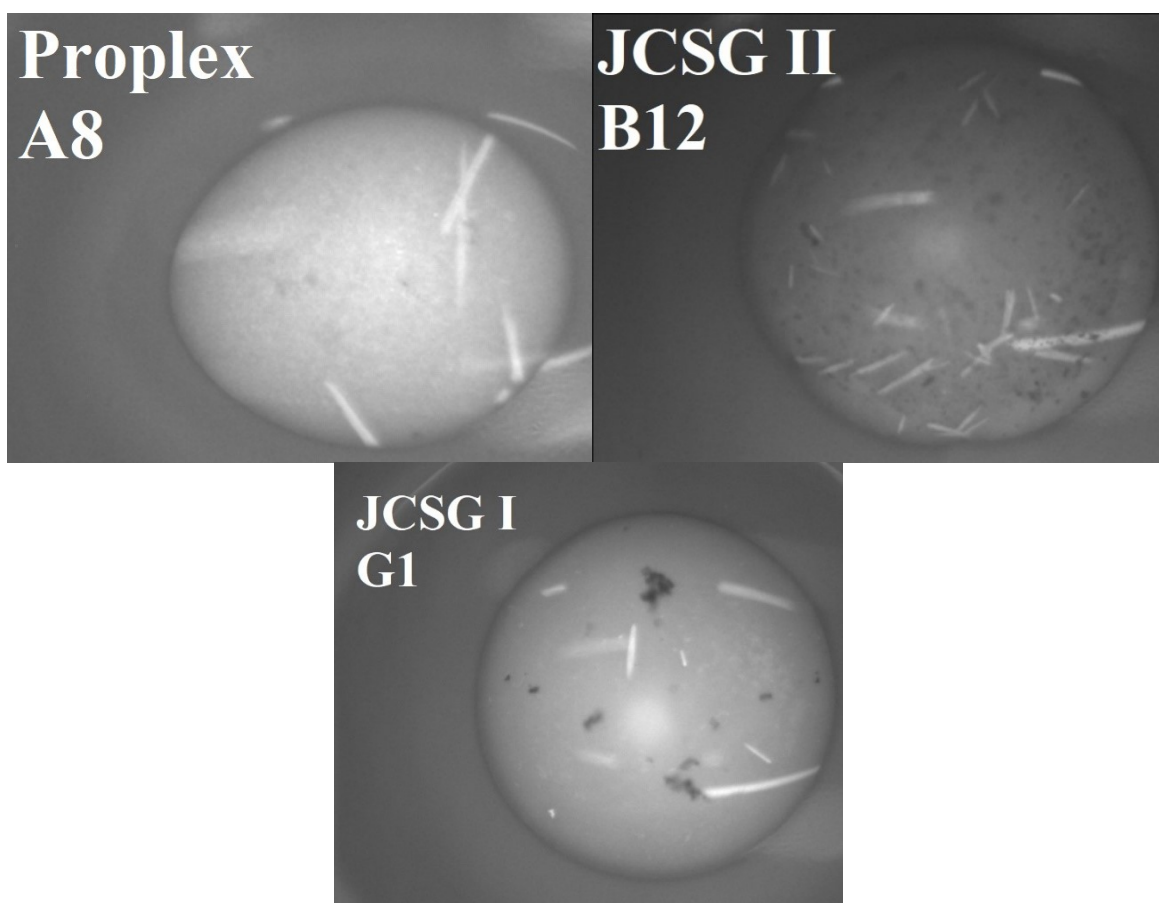
4.2 Krystalizace proteinového komplexu *nsp10/nsp16* s ligandem TO383

Před krystalizací byly spojeny frakce D2 až E10 (Obr. 4.1b str.31), které obsahovaly protein *nsp10* a *nsp16*. Tyto frakce byly přeneseny do koncentrátoru a opakovanou centrifugací (dle kapitoly 3.3.2.1, str. 27) byla dosažena výsledná koncentrace proteinu 5.27 mg/ml. Příprava krystalů probíhala dle kapitoly 3.3.2.2, str. 27.

Po dvou dnech bylo vidět rostoucí jehlice krystalů téměř v každé podmínce. Ve více než polovině všech jamek byly patrné proteinové precipitáty. Ke kontrole krystalů byl využit mikroskop.

Čtvrtý den byly vyloveny tenké jehlice z podmínek (Obr. 4.2a, str. 32):

- Proplex pozice A8 [0,1 M $\text{Mg}(\text{CH}_3\text{COO})_2 \cdot 4\text{H}_2\text{O}$; 0,1 M MOPS; pH 7,5; 12% w/v PEG 8000]
- JCSG II pozice B12 [0,6 M NaH_2PO_4 ; 0,6 M KH_2PO_4 ; 0,075 M Na HEPES pH 7,5; 25% (v/v) glycerol]
- JCSGI I pozice G1 [0,18 M tri-Ammonium citrát; 20% (w/v) PEG 3350]



Obr 4.2a: Proteinové krystaly *nsp10/nsp16* s ligandem *TO383*. Snímky byly pořízené pod UV mikroskopem při 10x zvětšení.

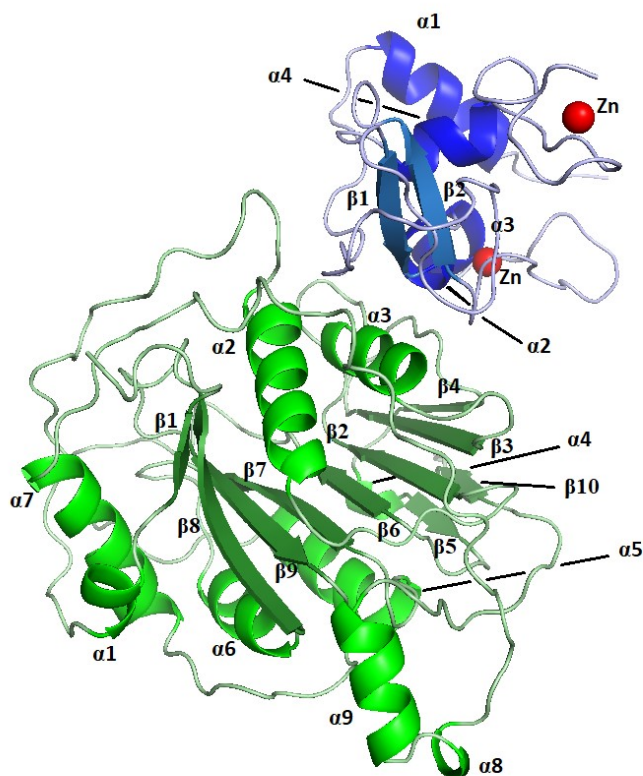
4.3 Krystalová struktura proteinového komplexu nsp10/nsp16 s ligandem TO383

Krystaly z krystalizačních roztoků byly zamraženy v kapalném dusíku podle kapitoly (3.3.2.3, str. 28). K měření difrakce byl použit krystal z podmínky Proplex A8 (Obr.4.2a, str. 32) při vlnové délce 1,54187 Å. Rozlišení z tohoto difrakčního experimentu dosahovalo hodnoty 2,7 Å.

Výsledná difrakční data byla integrována využitím programu XDS. Molekulárním nahrazením vyřešenou strukturou nsp10/nsp16 MTasy se sinefunginem (PDB ID: 6YZ1) [22] byl vyřešen fázový problém v programu Phaser. Úprava modelu struktury byla provedena v programu Coot. Oba programy Phaser i Coot byly součástí komplexního softwarového balíčku Phenix. Vybrané statistiky z měření dat a řešení struktury jsou uvedeny v tabulce 4.3, str. 33. Výsledná molekulová struktura byla upravena v programu PyMol.

Parametr	Hodnota
Rozlišení [Å]	45,78 – 2,70
Vlnová délka záření [Å]	1,54187
Počet reflexí	17104
Prostorová grupa	P1
R-work	0,2251
R-free	0,2844
RMS (bond)	0,0086
RMS (angles)	1,008
Clashscore	13,2
Ramachandran favoured (%)	90,15
Ramachandran outliers (%)	0,74
Rotamer outliers	8,28

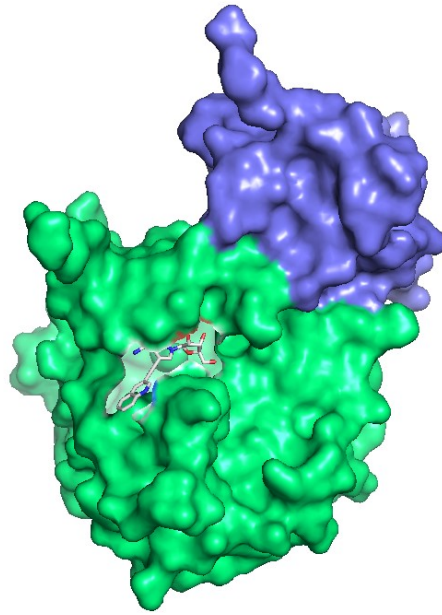
Tab. 4.3: Vybrané statistiky z měření dat a řešení krystalové struktury pro proteinový komplex nsp10/nsp16 s ligandem TO383.



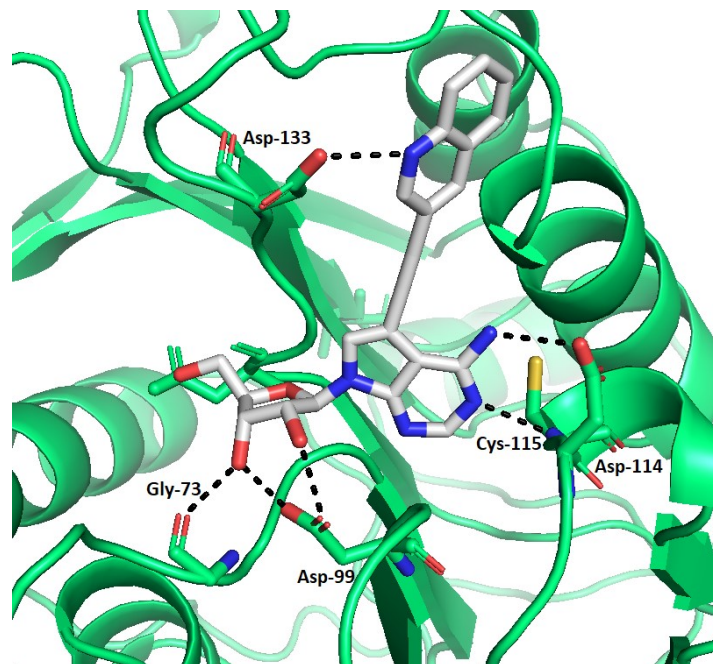
Obr. 4.3a: Struktura proteinového komplexu nsp10/nsp16 MTasy. Nsp16 v zelené barvě s vyznačenými α -helixy (světle zeleně) a β -listy (tmavě zeleně). Nsp10 vykreslen modrou barvou se zvýrazněnými α -helixy (tmavě modrá) a β -listy (světle modrá). Červenou barvou jsou zvýrazněny zinečnaté ionty.

Struktura proteinu nsp10 se skládá ze čtyř α -helixů a dvou antiparalelních β -listů. Zároveň se nej vážou dva ionty zinku. První zinek se nachází mezi $\alpha 2$ a $\alpha 3$ helixem a je vázán přes tři cysteinové zbytky (Cys-74, Cys-77, Cys-90) a jeden histidinový (His-83). Druhý iont zinku se nachází při C-konci a je vázán přes čtyři cysteinové zbytky (Cys-117, Cys-120, Cys-128, Cys-130). Ionty zinku slouží k stabilizaci proteinu nsp10.

Nsp16 se skládá z deseti β -listů a devíti α -helixů. β -listy $\beta 1$ - $\beta 9$ jsou kromě $\beta 5$ uspořádány paralelně a antiparalelně v centru proteinu. α -helixy se nachází na vnější straně a obklopují centrální část. Uspořádání α -helixů stabilizuje nsp16.

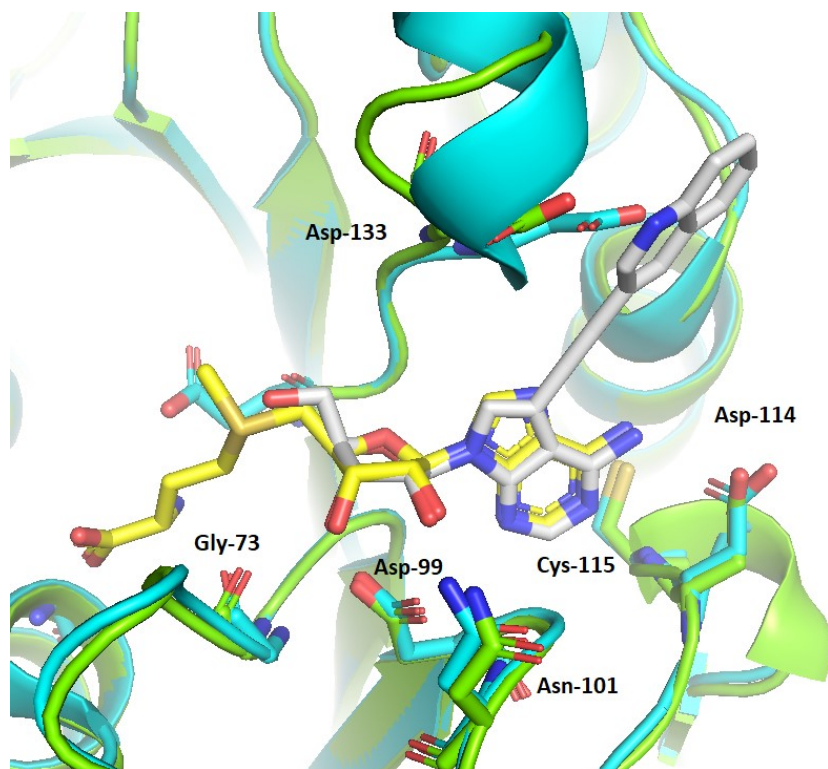


Obr. 4.3b: Vyobrazení povrchu komplexu nsp10/nsp16 s TO383 navázaným do vazebného místa. Zelenou barvou je vyznačen protein nsp16, modrou barvou nsp10. Ve vazebném místě pro SAM na povrchu nsp16 je vyobrazen navázaný ligand TO383.



Obr. 4.3c: Ligand TO383 ve vazebném místě nsp16. TO383 je vyobrazen světle šedou barvou, modrou barvou dusíky a červenou barvou kyslíky. Černou přerušovanou čárou jsou znázorněny interakce s aminokyselinami.

Nukleosidová část ligandu TO383 se převážně váže ke stejným aminokyselinám jako SAM. K této konzervativní části patří vodíková vazba k Asp-99, Cys-115 a Asp-114. Absence methioninového zbytku způsobila možnost vázat se ke Gly-73. Syntetizovaná část na umělém nukleosidu vytvořila novou vodíkovou vazbu na Asp-133.



Obr. 4.3d: Superpozice ligandu TO383 a SAM ve vazebném místě nsp16. Zelenou barvou je vyznačen komplex nsp16 s TO383 (šedá barva) a tyrkysovou barvou je vyznačen komplex nsp16 se SAM (žlutá barva). Superpozice proteinového komplexu nsp10/nsp16 s TO383 byla zhotovena v programu PyMol se strukturou nsp10/nsp16 v komplexu se SAM (PDB ID: 7JYY) [23].

5 Diskuse

V posledních dvou dekadách jsou patogenní koronaviry na vzestupu, proto je kladen důraz na lepší poznání této skupiny virů, abychom v případě nového smrtelného koronaviru byli schopni efektivněji reagovat. Jedním z atraktivních cílů terapeutického zásahu jsou methyltransferázy, které se vyznačují vysokou konzervovaností. Zároveň jsou důležité pro ochranu virového genomu před imunitním systémem. Na základě těchto skutečností by objev inhibitorů vázajících se do aktivního místa methyltransferáz mohl zastavit virovou infekci.

V této bakalářské práci byly připraveny rekombinantní proteiny nsp10 a nsp16 viru SARS-CoV-2. Prvním krokem bylo transformování kompetentních bakterií *E.coli* NiCo21(DE3) expresními plazmidy pSUMO. Proteiny byly jednotlivě exprimovány v buňkách *E.coli* a vypurifikovány (kapitola 3.3.1, str. 25). Z důvodu precipitace proteinu nsp16 byla v dialyzačním pufru snížena koncentrace NaCl (na 200 mM), avšak i po snížení iontové síly pufru byly patrné občasné precipitace. Možným vysvětlením může být vysoká koncentrace imidazolu (300mM) v elučním pufru po afinitní chromatografii. Nsp10 je stabilnější a méně náchylný na změnu podmínek, jelikož nevykazoval známky precipitace. Po odštěpení poly-histidinové kotvy a reversní afinitní chromatografií byly proteiny spojeny. Finálním krokem purifikace byla gelová permeační chromatografie. Jednotlivé purifikační kroky byly zkontrolovány metodou SDS-PAGE (Obr. 4.1b, str. 31). Zde je vidět rozdíl mezi proteiny nsp10 a nsp16 s navázanou poly-histidinovou kotvou a čistými proteiny nsp10 a nsp16. Po gelové permeační chromatografii byl vyhodnocen chromatogram (Obr. 4.1a, str. 30) a příslušné frakce byly rovněž zkontrolovány pomocí SDS-PAGE (Obr. 4.1b, str. 31). Frakce obsahující komplex nsp10/nsp16 byly využity pro další kroky. Frakce odpovídající menší molekulové hmotnosti oproti komplexu nsp10/nsp16 obsahovaly pouze protein nsp10, což ukazuje produkci většího množství tohoto proteinu při bakteriální expresi.

Během koncentrování proteinového komplexu pro krystalizační experiment (na cca. 5 mg/ml) nebyla pozorována žádná samovolná precipitace.

Ligand s označením TO383 byl rozpuštěn v DMSO a smíchán s proteinovým komplexem ve finální koncentraci 1 mM. Připravený komplex nsp10/nsp16 s ligandem TO383 byl smíchán s roztoky z komerčních krystalických sad v poměru protein:krystalizační roztok 2:1 a krystalizován při 18 °C. Po dvou dnech byly pozorovány precipitáty ve více než polovině jamek, ale také rostoucí proteinové krystaly jehlicového tvaru. Vyšší koncentrace proteinu

způsobovala precipitaci, bez růstu krystalu, téměř v každé jamce. Vybrané krystaly (Obr. 4.2a, str. 32) byly zamrazeny v tekutém dusíku.

Difrakční experiment probíhal na rotační anodě při vlnové délce 1,54187 Å. Nasbíraná difrakční data byla nejdříve integrována v programu XDS. V programu Phaser byl vyřešen fázový problém molekulovým nahrazením. Dodatečné úpravy struktury komplexu byly řešeny v programu Coot. Rozlišení dosáhlo hodnoty 2,7 Å.

Proteinová struktura nsp16 (Obr. 4.3a, str. 34, zeleně) obsahuje deset β -listů a devět α -helixů. Oproti publikacím krystalové struktury nsp16 s navázaným SAM je zde neshoda. Ta totiž obsahuje dvanáct β -listů a sedm α -helixů (Obr. 1.2.1.2a, str. 18). K povšimnutí je rozdíl právě dvou sekundárních struktur. U nsp10 je nesoulad mezi α -helixy, získaná struktura nsp10 se skládá ze čtyř α -helixů, zatímco publikované struktury se SAM jich mají pět. β -listy v nsp10 jsou identické. Může se jednat o změnu struktury, zapříčiněnou navázáním látky TO383 do místa pro SAM. Identické jsou rovněž aminokyseliny koordinující zinkové ionty.

Látka TO383 neobsahuje methionin, a tedy se neváže se k reziduíům Asn-43, Tyr-47, Gly-71, Gly-81 a Asp-130. Absence methioninu způsobila vznik nové interakce mezi 2'OH skupinou sacharidového zbytku a Gly-73. Vazba sacharidu s Asp-99 je stejná, nicméně posunutí molekuly dalo možnost vzniku nové vazby mezi 3'OH skupinou sacharidu a druhým kyslíkem Asp-99. Zúžení prostoru v tomto místě zapříčinilo nepřítomnost molekuly vody a vazbu s Asn-101. Nicméně ve výsledku se sacharidový zbytek váže třemi vazbami oproti dvěma vazbám se SAM. Upravený adenin tvoří stejnou vazbu s reziduem Cys-115 jako v případě SAM, navíc vznikla nová vazba s Asp-114. Svou artificiální částí tvoří interakci s aminokyselinou Asp-133.

Ze superpozice proteinových struktur (Obr. 4.3d, str. 36) je ve vazebném místě patrný posun celkové struktury TO383 vůči SAM. Důsledkem tohoto posunu je změna konformace reziduí Gly-73, Asn-101, Asp-114 a Asp-133. Asp-99 a Cys-115 jsou téměř nezměněné.

Inhibitor TO383 tvoří vodíkové vazby se šesti aminokyselinami, zatímco SAM jich tvoří deset.

Strukturní analýza vazebního místa s inhibitorem TO383 je dobrý základ pro návrh inhibitorů s potenciálem stát se léčivými.

6 Závěr

Cílem této bakalářské práce bylo připravit rekombinantní proteiny nsp10 a nsp16 z viru SARS-CoV-2 a detailně popsat inhibici methyltransferasové aktivity koronavirové 2'-O-RNA methyltransferasy, tj. komplexu nsp10/nsp16 pomocí blokace vazebného místa SAM ligandem TO383. Plazmidy nesoucí gen jednotlivě pro nsp10 a nsp16 byly transformovány do buněk *E.coli*. Expresí a purifikací proteinů byl obdržen proteinový komplex nsp10/nsp16. Krystalizací v různých krystalizačních roztocích metodou sedící kapky byl nakonec získán dobře difraktující krystal z roztoku Proplex A8.

Krystal byl analyzován rentgenovou difrakcí a byla získaná data pro řešení proteinové struktury. Podařilo se vyřešit strukturu komplexu a popsat rozdíly v interakci proteinového komplexu nsp10/nsp16 s ligandem TO383 oproti substrátu SAM. Získaná strukturní data mohou posloužit k pro vývoj ještě lepších inhibitorů koronavirové 2'-O-RNA methyltransferasy a eventuálně i účinných koronavirových antivirových léčiv.

7 Seznam použité literatury

- [1] Masters, Paul S. The Molecular Biology of Coronaviruses. *Advances in Virus Research*, 2006, **66**, 193-292. ISSN: 0065-3527. DOI: 10.1016/S0065-3527(06)66005-3
- [2] Cui, J.; Li, F.; Shi, Z.-L. Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology*, 2019, **17**(3), 181-192. ISSN: 1740-1534. DOI: 10.1038/s41579-018-0118-9
- [3] Millet, J. K.; Whittaker, G. R. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research*, 2015, **202**, 120-134. ISSN: 0168-1702. DOI: 10.1016/j.virusres.2014.11.021
- [4] Holmes, K. V. CORONAVIRUSES (CORONAVIRIDAE). *Encyclopedia of Virology (Second Edition)*, 1999, 291-298. ISBN: 978-0-12-227030-7. DOI: 10.1006/rwvi.1999.0055
- [5] Hulswit, R. J. G; de Haan, C. A. M; Bosch B.-J. Chapter Two - Coronavirus Spike Protein and Tropism Changes. *Advances in Virus Research*, 2016, **96**, 29-57. DOI: DOI: 10.1016/bs.aivir.2016.08.004
- [6] Kirtipal, N.; Bharadwaj, S.; Kang, S. G. From SARS to SARS-CoV-2, insights on structure, pathogenicity and immunity aspects of pandemic human coronaviruses. *Infection, Genetics and Evolution*, 2020, **85**, 104502. ISSN: 1567-1348. DOI: 10.1016/j.meegid.2020.104502
- [7] Lu, G.; Wang, Q.; Wang, G. F. Bat-to-human: spike features determining 'host jump' of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends in Microbiology*, 2015, **23**(8), 468-478. ISSN: 0966-842X, 1878-4380. DOI: 10.1016/j.tim.2015.06.003
- [8] Liu, D. X.; Fung, T. S.; Chong, K. K.-L.; Shukla, A.; Hilgenfeld, R. Accessory proteins of SARS-CoV and other coronaviruses. *Antiviral Research*, 2014, **109**, 97-109. ISS: 0166-3542. DOI: 10.1016/j.antiviral.2014.06.013
- [9] Song, Z.; Xu, Y.; Bao, L.; Zhang, L.; Yu, P. Qu, Y.; Zhu, H.; Zhao, W.; Han, Y.; Qin, C. From SARS to MERS, Thrusting Coronaviruses into the Spotlight. *Viruses*, 2019, **11**(1), 59. ISSN: 1999-4915. DOI: 10.3390/v11010059

- [10] Chafekar, A.; Fielding, B. C. MERS-CoV: Understanding the Latest Human Coronavirus Threat. *Viruses*, 2018, **10**(2), 93. ISSN: 1999-4915. DOI: 10.3390/v10020093
- [11] Du, L.; Tai, W.; Zhou, Y.; Jiang, S. Vaccines for the prevention against the threat of MERS-CoV. *Expert Review of Vaccines*, 2016, **15**(9), 1123-1134. ISSN: 1476-0584. DOI: 10.1586/14760584.2016.1167603
- [12] Ting, M.; Suzuki, J. B. SARS-CoV-2: Overview and Its Impact on Oral Health. *Biomedicines*, 2021, **9**(11), 1690. ISSN: 2227-9059. DOI: 10.3390/biomedicines9111690
- [13] Bai, C.; Zhong, Q.; Gao, G. F. Overview of SARS-CoV-2 genome-encoded proteins | SpringerLink. *Science China Life Sciences*, 2022, **65**, 280-294. DOI: 10.1007/s11427-021-1964-4
- [14] WHO Coronavirus (COVID-19) Dashboard. [vid. 29.4.2022]. <https://covid19.who.int/>
- [15] Hu, B.; Guo, H.; Zhou, P.; Shi, Z.-L. Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, 2021, **19**(3), 141-154. ISSN: 1740-1534. DOI: 10.1038/s41579-020-00459-7
- [16] Beyerstedt, S.; Casaro, E. B.; Rangel, É. B. COVID-19: angiotensin-converting enzyme 2 (ACE2) expression and tissue susceptibility to SARS-CoV-2 infection. *European Journal of Clinical Microbiology & Infectious Diseases*, 2021, **40**(5), 905-919. ISSN: 1435-4373. DOI: 10.1007/s10096-020-04138-6
- [17] Deen, J.; Vranken, C.; Leen, V.; Neely R. K.; Janssen, K. P. F.; Hofkens, J. Methyltransferase-Directed Labeling of Biomolecules and its Applications. *Angewandte Chemie International Edition*, 2017, **56**(19), 5182-5200. ISSN: 1521-3773. DOI: 10.1002/anie.201608625
- [18] Lin, S.; Chen, H.; Ye, F.; Chen, Z.; Yang, F.; Zheng, Y.; Cao, Y.; Qiao, J.; Yang, S.; Lu, G. Crystal structure of SARS-CoV-2 nsp10/nsp16 2'-O-methylase and its implication on antiviral drug design. *Signal Transduction and Targeted Therapy*, 2020, **5**(1), 1-4. ISSN: 2059-3635. DOI: 10.1038/s41392-020-00241-4
- [19] Romano, M.; Ruggiero, A.; Squeglia, F.; Maga, G.; Berisio, R. A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *Cells*, 2020, **9**(5), 1267. ISSN: 2073-4409. DOI: 10.3390/cells9051267

- [20] Rosas-Lemus, M.; Minasov, G.; Shuvalova, L.; Inniss, N. L.; Kiryukhina, O.; Wiersum, G.; Kim, Y.; Jedrzejczak, R.; Maltseva, N. I.; Endres, M.; Jaroszewski, L.; Godzik, A.; Joachimiak, A.; Satchell, K. J. F. The crystal structure of nsp10-nsp16 heterodimer from SARS-CoV-2 in complex with S-adenosylmethionine. *bioRxiv*, 2020, 2020.04.17.047498. DOI: 10.1101/2020.04.17.047498
- [21] Bank, RCSB Protein Data. RCSB PDB – 6YZ1. The crystal structure of SARS-CoV-2 nsp10-nsp16 methyltransferase complex with Sinefungin. <https://www.rcsb.org/structure/6yz1>
- [22] Bank, RCSB Protein Data. RCSB PDB – 7JYY. Crystal Structure of SARS-CoV-2 Nsp16/10 Heterodimer in Complex with (m⁷GpppA)pUpUpApApA (Cap-0) and S-Adenosylmethionine (SAM). <https://www.rcsb.org/structure/7JYY>