

Univerzita Hradec Králové
Přírodovědecká fakulta
Katedra chemie

Příprava a základní charakterizace fúzních
proteinů složených z PDZ3_ZO-1 a GA88

Diplomová práce

Autor práce: Bc. Jan Hurdálek

Studijní program: B1407 Chemie

Studijní obor: Chemie - spec. toxikologie a analýza škodlivin

Vedoucí práce: doc. RNDr. Lucie Zemanová, Ph.D.

Školitelka specialista: RNDr. Kristýna Vydra Boušová, Ph.D.

Hradec Králové

červenec 2023



Zadání diplomové práce

Autor: Jan Hurdálek

Studium: S21CH014NP

Studijní program: N0531A130003 Chemie

Studijní obor: Toxikologie a analýza škodlivin

Název diplomové práce: Příprava a základní charakterizace fúzních proteinů složených z PDZ_ZO-1 a GA88

Název diplomové práce AJ: The preparation and basic characterization of fusion proteins composed of PDZ3_ZO-1 and GA88

Cíl, metody, literatura, předpoklady:

Experimentální typ práce -exprese a purifikace rekombinantních fúzních proteinů a jejich základní charakterizace pomocí vybraných metod (např. SDS-PAGE, DSF, BLI)

Zpracování provedených experimentů, získaných výsledků spolu s teoretickými podklady do diplomové práce o rozsahu cca 70 stran

Yu K. et al. (2015) Synthetic fusion protein design and application. *Biotechnology advances*, 33, 155-164.

Bousova K. et al. (2021) The order of PDZ3 and TrpCage in fusion chimeres determines their properties - a biophysical characterization, *ProteinSci*, 30(8), 1653-1668.

He Y. et al. (2008) NMR structures of two designed proteins with high sequence identity but different fold and function, *PNAS*, 105(38), 14412-14417.

Zadávací pracoviště: Katedra chemie,
Přírodovědecká fakulta

Vedoucí práce: doc. RNDr. Lucie Zemanová, Ph.D.

Oponent: PharmDr. Rudolf Andráš, Ph.D.

Datum zadání závěrečné práce: 20.7.2021

Prohlášení:

Prohlašuji, že jsem diplomovou práci vypracoval samostatně a že jsem v seznamu použité literatury uvedl všechny prameny, z kterých jsem vycházel.

V Hradci Králové dne

Jan Hurdálek

Poděkování

V první řadě mnohokrát děkuji doc. RNDr. Lucii Zemanové, Ph.D. za vedení této práce, za čas, který mi věnovala, za každou zpětnou vazbu, kterou mi poskytla a ze které jsem tak dostal příležitost se přiučit a poučit. Rovněž bych jí rád poděkoval i za to, že mi pro experimentální část této práce zajistila obohacující zkušenost s laboratorní prací v ÚOCHB AV ČR.

Velký dík patří také všem členům laboratoře týmu Bioinformatika ÚOCHB AV ČR pod vedením prof. RNDr. Jiřího Vondráška, Csc., jmenovitě Ing. Veronice Vetyškové, Ph.D. a Mgr. Monice Zouharové za obětavou pomoc s osvojením použitého protokolu a metod a za konzultaci výsledků, dále Mgr. Kláře Poštulkové a Mgr. Kateřině Jiráskové, Ph.D., které byly vždy nablízku, kdykoliv jsem si nevěděl rady, a opomenuta nesmí být ani vedoucí laboratoře, RNDr. Kristýna Vydra Boušová, Ph.D., která celou mou práci v ÚOCHB AV ČR zastřešovala a poskytla mi řadu cenných rad ohledně této práce.

Dále děkuji také Mgr. Josefu Šulcovi za pomoc s vypracováním molekulárních modelů, RNDr. Lucii Bednárové, CSc. za provedení spektroskopie cirkulárního dichroismu a týmu doc. RNDr. Josefa Cvačky, Ph.D. za hmotnostně spektrometrické analýzy.

V neposlední řadě bych chtěl poděkovat také mé rodině a partnerce, kteří mi byli po celou dobu studia nepostradatelnou oporou.

Anotace

HURDÁLEK, J. *Příprava a základní charakterizace fúzních proteinů složených z PDZ3_ZO-1 a G_A88*. Hradec Králové, 2023. Diplomová práce. Univerzita Hradec Králové. Přírodovědecká fakulta. Vedoucí diplomové práce doc. RNDr. Lucie ZEMANOVÁ, Ph.D.

Fúzní proteiny jsou arteficiální proteiny připravované spojením dvou a více proteinových domén, zpravidla pomocí genového inženýrství. Jsou dnes široce uplatňovány v medicíně, vědě i chemickém průmyslu, nicméně nové poznatky o mezidoménovém ovlivňování mají potenciál odvětví designu a přípravy funkčních fúzních proteinů nebývale akcelarovat. Fúzní proteiny představené v této práci, FD5A a FD6A, jsou složeny z PDZ3 domény proteinu ZO-1 a arteficiální chameleonské domény G_A88 a molekulárním modelováním *in silico* u nich byla predikována významná mezidoménová ovlivnění mechanismem dynamické alosterie. FD5A a FD6A s histidinovými kotvami byly v rámci této práce exprimovány v *E. coli* a purifikovány pomocí afinitní chromatografie s imobilizovanými ionty kovů, dále byly kotvy z proteinů odstraněny pomocí štěpení TEV proteasou a subtraktivní afinitní chromatografie a proteiny byly přečištěny pomocí vylučovací chromatografie. Následovala charakterizace několika biofyzikálními metodami: analytické vylučovací chromatografie, dynamického rozptylu světla, analytické ultracentrifugace, cirkulárního dichroismu a diferencíální skenovací fluorimetrie. Bylo zjištěno, že ačkoliv se proteiny FD5A a FD6A prakticky neliší terciární ani sekundární strukturou, jejich tepelná stabilita je výrazně odlišná, což poukazuje na vzájemné vzdálené ovlivňování domén v těchto fúzích.

Klíčová slova

Fúzní proteiny, chimérické proteiny, PDZ domény, PDZ3 ZO-1, chameleonské proteiny, G_A88, dynamická alosterie

Annotation

HURDÁLEK, J. *The preparation and basic characterisation of fusion protein composed of PDZ3_ZO-1 and G_A88*. Hradec Králové, 2023. Diploma thesis. University of Hradec Králové. Faculty of Science. Diploma thesis supervisor doc. RNDr. Lucie ZEMANOVÁ, Ph.D.

Fusion proteins are artificial proteins prepared by connecting of two or more protein domains, generally with usage of genetic engineering methodology. Today, they are widely used in medicine, science and chemical industries, but new research in interdomain communication has the potential to accelerate the field of fusion protein design and preparation even further. For fusion proteins introduced in this work, FD5A and FD6A, comprised of the PDZ3 domain of the protein ZO-1 and artificial chameleon domain G_A88, significant interdomain influencing via the mechanism of dynamic allostery was predicted by molecular modelling *in silico*. In this thesis, FD5A and FD6A with histidine tags were expressed in *E. coli* and purified using immobilised metal affinity chromatography; the tags were cleaved with TEV protease and proteins were separated with subtractive affinity chromatography. Furthermore, characterisation using several biophysical methods such as size-exclusion chromatography, dynamic light scattering, analytical ultracentrifugation, circular dichroism, and thermal shift assay was successfully attempted. It was determined that even though FD5A and FD6A do not vary in secondary nor tertiary structure, their thermal stability is significantly different, which suggests mutual, long-distance influence of the two domains in these fusions.

Key words

Fusion proteins, chimeric proteins, PDZ domains, PDZ3 ZO-1, chameleon proteins, G_A88, dynamic allostery

Seznam zkratek

A	= <u>a</u> lanin (v proteinových sekvencích), <u>a</u> denin (v sekvencích DNA)
ABL1	= homolog virového onkogenu Abelsonovy myší leukemie 1 (z angl. <u>A</u> belson murine leukemia viral oncogene homolog <u>1</u>)
ABR	= aktin vázající oblast (z angl. <u>a</u> ctin <u>b</u> inding <u>r</u> egion)
AF-6	= <u>a</u> fadin
AMP	= <u>a</u> mpicilin
AMPA	= kyselina α -amino-3-hydroxy-5-methyl-4-isoxazolpropionová (z angl. α -amino-3-hydroxy-5- <u>m</u> ethyl-4-isoxazolepropionic <u>a</u> cid)
AP	= alkalická fosfatasa (z angl. <u>a</u> lkaline <u>p</u> hosphatase)
aPKC	= atypická protein kinasa C (z angl. <u>a</u> typical <u>p</u> rotein <u>k</u> inase <u>C</u>)
ASEC	= analytická vylučovací chromatografie (z angl. <u>a</u> nalytical <u>s</u> ize- <u>e</u> xclusion <u>c</u> hromatography)
AtoB	= acetoacetyl-CoA thiolasa
AUC	= <u>a</u> lytická <u>u</u> ltra <u>c</u> entrifugace
BCR	= protein klastrového regionu zlomového bodu (z angl. <u>b</u> reakpoint <u>c</u> luster <u>r</u> egion protein)
BPB	= bromfenolová modř (z angl. <u>b</u> romophenol <u>b</u> lue)
BFP	= modrý fluorescenční protein (z angl. <u>b</u> lue <u>f</u> luorescent <u>p</u> rotein)
BRET	= bioluminiscenční rezonanční přenos energie (z angl. <u>b</u> ioluminescence <u>r</u> esonance <u>e</u> nergy <u>t</u> ransfer)
C	= <u>c</u> ystein (v proteinových sekvencích), <u>c</u> ytosin (v sekvencích DNA)
CASK	= vápník-dependentní serinová protein kinasa (homolog Lin-2) (z angl. <u>c</u> alcium/calmodulin-dependent <u>s</u> erine protein <u>k</u> inase)
CASP	= soutěž kritického hodnocení predikce struktury (z angl. <u>C</u> ritical <u>A</u> ssessment of <u>S</u> tructure <u>P</u> rediction)
CAT	= <u>c</u> hloramfenikol <u>a</u> cetyl <u>t</u> ransferasa
CBB	= barvivo <u>C</u> oomassie <u>B</u> rilliant <u>B</u> lue
CBP	= kalmodulin vázající peptid (z angl. <u>c</u> almodulin <u>b</u> inding <u>p</u> eptide)
CD	= <u>c</u> irkulární <u>d</u> ichroismus
CFP	= <u>c</u> yanový fluorescenční protein
CLIC1	= intracelulární chloridový kanál 1 (z angl. <u>c</u> hloride intracellular <u>c</u> hannel <u>1</u>)
CTB	= doména B exotoxinu <i>Vibrio cholerae</i> (z angl. <u>c</u> holera <u>t</u> oxin domain <u>B</u>)
CuAAC	= Huisgenova <u>Cu</u> -katalyzovaná <u>a</u> zid- <u>a</u> lkyn [3+2] cykloadice
Cx	= konexiny (z angl. <u>c</u> onnexin; Cx43, Cx46, Cx50, Cx45, Cx31.9 a Cx47)
D	= aspartát
Dlg	= <u>D</u> iscs <u>l</u> arge homology (Dlg1-4, homology Discs large proteinů <i>Drosophila</i>)
DLS	= dynamická rozptyl světla (z angl. <u>d</u> ynamic <u>l</u> ight <u>s</u> cattering)
DMSO	= <u>d</u> imethylsulfoxid
dNTP	= deoxynukleotid trifosfát (z angl. <u>d</u> eoxynucleotide <u>t</u> riphosphate)
DSF	= diferenční skenovací fluorimetrie, také TSA = <u>t</u> hermal <u>s</u> hift <u>a</u> ssay
DTT	= <u>d</u> ithio <u>t</u> reitol
Dvl	= proteiny rodiny <u>D</u> ishevelled (homology stejnojmenného Dsh z <i>Drosophila</i>)

E	= glutamát (v proteinových sekvencích), eluční frakce (v při popisu purifikace)
<i>E. coli</i>	= <i>Escherichia coli</i>
F	= feny l alanin (v proteinových sekvencích), frakce (při popisu purifikace)
Fc	= region protilátek (z angl. fragment crystallizable)
FcRn	= Fc receptor, neonatální
FPLC	= rychlá proteinová kapalinová chromatografie (z angl. fast protein liquid chromatography)
FRET	= Försterův/fluorescenční rezonanční přenos energie (z angl. fluorescence resonance energy transfer)
FT	= proše d ší frakce (z angl. flow-through)
Fz	= proteiny rodiny Frizzled
G	= glycin (v proteinových sekvencích), guanin (v sekvencích DNA)
GBD	= GTPasová protein vázající doména (z angl. GTPase protein-binding domain)
GFP	= zelený fluorescenční protein (z angl. green fluorescence protein; GFP, GFP2)
GKAP-1	= protein kotvící G kinasu 1 (z angl. G kinase-anchoring protein 1)
GluN	= rodina podjednotek NMDA glutamátového receptoru
GluR	= rodina podjednotek AMPA glutamátového receptoru
Gm	= guanidin
GRIP-1	= protein interagující s glutamátovým receptorem 1 (z angl. glutamate receptor-interacting protein 1)
GST	= glutathion-S-transferasa
GTP	= guanosintrifosfát (z angl. guanosin triphosphate)
GuK	= doména homologní guanylát kinase
H	= histidin
HIC	= chromatografie hydrofobních interakcí (z angl. hydrophobic interactions chromatography)
HIV	= virus lidské imunitní nedostatečnosti (z angl. human immunodeficiency virus)
HMGR	= hydroxy-methylglutaryl-CoA reduktasa
HMGS	= hydroxy-methylglutaryl-CoA synthasa
HPV	= lidský papillomavirus (z angl. human papillomavirus)
HSA	= lidský sérový albumin (z angl. human serum albumin)
HSPG	= heparansulfátový proteoglykan
HTLV	= lidský virus leukemie T-buněk (z angl. human T-lymphoma virus)
I	= isoleucin
IDA	= iminodiocetová kyselina (z angl. iminodiacetic acid)
IEC	= iontově výměnná chromatografie (z angl. ion exchange chromatography)
Ig	= rodina imunoglobulinů (IgG, huIgG, IgA, IgM, IgE)
IMAC	= afinitní chromatografie s imobilizovanými ionty kovů (z angl. immobilized metal affinity chromatography)
IPA	= isopropanol (isopropylalkohol)
IPTG	= isopropyl-β-D-1-thiogalaktopyranosid

JAM	= molekuly adheze spojů (z angl. <u>j</u> unctional <u>a</u> dhesion <u>m</u> olecule; JAM-A, JAM-B a JAM-C)
K	= lysin
Kex2	= <u>k</u> exin
KIF	= kinesiny a kinesinu podobné proteiny (KIF5, KIF1A, KIF13B, KIF1B α a KIF17)
L	= <u>l</u> eucin
LARG	= xylosyl- a glukuronyltransferasa <u>L</u> ARGE2
LLGL	= homolog 1 proteinu Lgl (z angl. <u>l</u> ethal(<u>2</u>) <u>g</u> iant <u>l</u> arvae protein homolog 1)
LPA	= lysofosfatidová kyselina (z angl. <u>l</u> ysophosphatidic <u>a</u> cid)
LTB	= doména B tepelně labilního enterotoxinu <i>E. coli</i> (z angl. heat <u>l</u> abile enterotoxin domain <u>B</u>)
LTD	= dlouhodobá deprese (z angl. <u>l</u> ong <u>t</u> erm <u>d</u> epression)
Ltn	= lymfotaktin (formy Ltn10 a Ltn40)
LTP	= dlouhodobá potenciace (z angl. <u>l</u> ong <u>t</u> erm <u>p</u> otentiation)
M	= <u>m</u> ethionin
MAGUK	= homolog guanylát kinasy asociované s membránou (z angl. <u>m</u> embrane-associated <u>g</u> uanylate <u>k</u> inases homolog)
MALDI	= <u>m</u> atrici <u>a</u> sistovaná <u>l</u> aserová <u>d</u> esorpce/ <u>i</u> onizace
MARVEL	= doména proteinu MAL a příbuzných proteinů vezikulárního transportu a membránového spojení (z angl. <u>M</u> AL <u>a</u> nd <u>r</u> elated proteins for <u>v</u> esicle trafficking and membrane <u>l</u> ink)
MBP	= maltosu vázající protein (z angl. <u>m</u> altose- <u>b</u> inding protein)
mGluR	= <u>m</u> etabotropní <u>g</u> lutamátové <u>r</u> eceptory
MS	= hmotnostní spektrometrie (z angl. <u>m</u> ass <u>s</u> pectrometry)
MUPP1	= protein s několika PDZ doménami 1 (z angl. <u>m</u> ultiple <u>P</u> DZ domain protein <u>1</u>)
MWCO	= mezní hodnota molekulové hmotnosti (z angl. <u>m</u> olecular <u>w</u> eight <u>c</u> utoff)
N	= asparagin
NHERF1	= regulační kofaktor Na ⁺ /H ⁺ výměny (z angl. <u>N</u> a ⁺ / <u>H</u> ⁺ <u>e</u> xchange <u>r</u> egulatory <u>c</u> ofactor)
NLGN	= <u>n</u> euro <u>l</u> igin
NMDA	= N-metyl-D-aspartát (z angl. <u>N</u> - <u>m</u> ethyl- <u>D</u> - <u>a</u> spartic acid)
NMR	= <u>n</u> ukleární <u>m</u> agnetická <u>r</u> ezonance
nNOS	= <u>n</u> euronální <u>N</u> O-synthasa
NRXN	= <u>n</u> eure <u>x</u> in
NTA	= nitrilotrioctová kyselina (z angl. <u>n</u> itrilo <u>t</u> ri <u>a</u> cetic acid)
O/N	= přes noc (z angl. <u>o</u> ver <u>n</u> ight)
OCLDBR	= okludin vázající oblast (z angl. <u>o</u> cc <u>l</u> u <u>d</u> in- <u>b</u> inding <u>r</u> egion)
P	= prolin (v proteinových sekvencích), <u>p</u> eleta (při popisu purifikace)
PAGE	= <u>p</u> oly <u>a</u> krylamidová <u>g</u> elová <u>e</u> lektroforéza
Par	= proteiny a komplex defektivního oddělení (z angl. <u>p</u> artition/partioning defective; PAR-6, PAR-3)
PCR	= polymerázová řetězová reakce (z angl. <u>p</u> olymerase <u>c</u> hain <u>r</u> eaction)
PDZ	= doména proteinů Postsynaptic density 95, <u>D</u> iscs-large a <u>Z</u> onula occludens 1

PICK-1	= protein interagující s protein kinasou C- α 1
PIP	= Fosfatidylinositolfosfát (z angl. <u>phosphatidylinositolphosphate</u> ; PI(4,5)P ₂)
PMF	= hmotnostní otisk prstu peptidu (z angl. <u>peptide mass fingerprinting</u>)
PSD-95	= protein postsynaptické hustoty 95 (z angl. <u>Postsynaptic density 95</u>)
PTPN4	= nереceptorová tyrosin-protein fosfatasa typu 4 (z angl. <u>tyrosine-protein phosphatase non-receptor type 4</u>)
Q	= glutamin
R	= arginin (v proteinových sekvencích), separace (v FPLC separacích; z angl. <u>run</u>)
RP-HPLC	= vysokoúčinná kapalinová chromatografie s obrácenými fázemi (z angl. <u>reverse-phase high-performance liquid chromatography</u>)
S	= serin
SARS	= akutní respirační syndrom (z angl. <u>severe acute respiratory syndrome</u>)
SDS	= dodecylsírán sodný (z angl. <u>sodium dodecylsulphate</u>)
SEC	= vylučovací chromatografie, také gelová filtrace (angl. <u>size exclusion chromatography</u>)
SH3	= doména 3 homologní se Src (z angl. <u>Src homology 3</u>)
Shank1	= protein s doménami SH3 a několika ankyrinovými repeticemi (z angl. <u>SH3 and multiple ankyrin repeat domains protein 1</u>)
SN	= supernatant
T	= threonin (v proteinových sekvencích), thymine (v sekvencích DNA)
Taq	= <i>Thermus aquaticus</i>
TEMED	= tetramethylethyldiamin
TEV	= virus leptané mozaiky tabáku (z angl. <u>Tobacco etch virus</u>)
TIAM-1	= protein indukující invazi a metastázi lymfomu T-buněk 1 (z angl. <u>T-lymphoma invasion and metastasis-inducing protein 1</u>)
TOF	= analyzátor doby letu (z angl. <u>time of flight</u>)
Tris	= tris(hydroxymethyl)aminomethan
TRX	= thioredoxin
V	= valin
VLP	= virům podobná částice (z angl. <u>virus-like particle</u>)
W	= tryptofan (v proteinových sekvencích), promytí (při popisu purifikace; z angl. <u>wash</u>)
Wnt	= protein <u>Wingless/Integrin-1</u> a s ním spojená signalizační kaskáda
Y	= Tyrosin
YFP	= žlutý fluorescenční protein (z angl. <u>yellow fluorescence protein</u>)
ZO	= Zonula occludens proteiny (ZO-1 až ZO-3)
ZU5	= doména proteinů Zonula occludens 1 a <u>unc5</u> podobnému netrinovému receptoru

Obsah

Úvod	13
1 Teoretická část	14
1.1 PDZ domény	14
1.1.1 Obecná charakteristika PDZ domén	15
1.1.1.1 Struktura PDZ domén	15
1.1.1.2 Vazební partneři PDZ domén.....	16
1.1.2 Fyziologické funkce PDZ.....	18
1.1.2.1 Rozdělení proteinů obsahujících PDZ domény	18
1.1.2.2 Synaptický přenos	18
1.1.2.3 Buněčné spoje	20
1.1.2.4 Signalizace	22
1.1.2.5 Onemocnění spojená s PDZ-PBM interakcemi	23
1.1.3 PDZ3 ZO-1.....	25
1.1.3.1 Dynamická alosterie PDZ3 ZO-1.....	27
1.1.3.2 Charakteristika použité PDZ3.....	27
1.2 G_{A88} , G_{B88}	29
1.2.1 Chameleonské proteiny	29
1.2.2 Charakteristika G_{A88}	32
1.3 Fúzní proteiny	34
1.3.1 Design a příprava rekombinantních fúzních proteinů.....	35
1.3.1.1 Výběr domén a linkerů	35
1.3.1.2 Metody přípravy proteinových fúzí	36
1.3.2 Praktická využití rekombinantních fúzních proteinů	38
1.3.2.1 Terapeutika	38
1.3.2.2 Biokatalyzátory.....	38
1.3.2.3 Analytické metody.....	39
2 Cíle práce	40
3 Experimentální část	41
3.1 Vybavení.....	41
3.1.1 Přístrojové vybavení a materiál	41
3.1.2 Software.....	42
3.2 Chemický a biologický materiál	43

3.2.1	Chemikálie.....	43
3.2.2	Biologický materiál.....	43
3.2.3	Komerční směsi, kity a standardy.....	44
3.2.4	Pufry, roztoky a média.....	44
3.2.5	Příprava polyakrylamidových gelů.....	45
3.3	Metodika.....	46
3.3.1	Predikce struktur proteinů FD5A a FD6A <i>in silico</i>	46
3.3.2	Příprava bakteriálních kultur a exprese studovaných proteinů.....	46
3.3.2.1	Transformace <i>E. coli</i>	46
3.3.2.2	Mutagenese pET15b_FD6A pomocí PCR.....	46
3.3.2.3	Expese proteinů v <i>E. coli</i>	47
3.3.3	Purifikace exprimovaných proteinů.....	48
3.3.3.1	Příprava lyzátu <i>E. coli</i>	48
3.3.3.2	Afinitní chromatografie s imobilizovanými ionty kovů (IMAC).....	48
3.3.3.3	Vizualizace průběhu IMAC pomocí SDS-PAGE.....	48
3.3.3.4	Denaturace, odštěpení histidinové kotvy TEV proteasou, renaturace.....	48
3.3.3.5	Subtraktivní IMAC.....	49
3.3.3.6	Vylučovací chromatografie (SEC).....	49
3.3.3.7	SDS-PAGE a MS frakcí z SEC.....	50
3.3.4	Biofyzikální charakterizace.....	51
3.3.4.1	Analytická SEC.....	51
3.3.4.2	Dynamický rozptyl světla (DLS).....	51
3.3.4.3	Analytická ultracentrifugace (AUC).....	51
3.3.4.4	Cirkulární dichroismus (CD).....	52
3.3.4.5	Diferenční skenovací fluorimetrie (DSF).....	52
3.4	Výsledky a diskuse.....	53
3.4.1	Design a struktury proteinů FD5A a FD6A.....	53
3.4.2	Transformace <i>E. coli</i>	56
3.4.3	Příprava plazmidu pro expresi G _A 88.....	56
3.4.4	Purifikační proces FD5A, FD6A, G _A 88.....	59
3.4.4.1	Purifikace rekombinantního FD5A.....	61
3.4.4.2	Purifikace rekombinantního FD6A.....	65
3.4.4.3	Purifikace rekombinantního G _A 88.....	69

3.4.5	Biofyzikální charakterizace FD5A, FD6A, G _A 88 a PDZ3	74
3.4.5.1	Analytická SEC	74
3.4.5.2	Dynamický rozptyl světla (DLS)	78
3.4.5.3	Analytická ultracentrifugace (AUC)	81
3.4.5.4	Cirkulární dichroismus (CD)	88
3.4.5.5	Diferenční skenovací fluorimetrie (DSF)	90
	Závěr	93
	Bibliografie	94
	Příloha I	112
	Příloha II	113
	Příloha III	116
	Příloha IV	117
	Příloha V	118

Úvod

Fúzní proteiny se řadí mezi perspektivní nástroje moderní molekulární biologie, bioinženýrství a medicíny, nicméně vztah mezi jejich složením a funkcí je komplexnější, než by se mohlo zdát. Vzájemné ovlivňování proteinových domén, např. mechanismem dynamické alosterie, může signifikantně přispívat k regulaci aktivity i proteinové stabilitě, přesto jde o málo prozkoumaný děj jak ve fúzních, tak i přírodních proteinech.

Jednou z proteinových rodin s dobře popsanou schopností dynamické alosterie jsou PDZ domény. Tyto jsou významně zastoupeny v multidoménných proteinech prokaryotních i eukaryotních organismů, člověka nevyjímaje. Ujímají se velkého množství rolí včetně buněčné regulace, tvorby buněčných spojů a facilitace synaptického přenosu, a to díky své schopnosti vázat k sobě jiné proteiny a vytvářet klíčové proteinové komplexy.

V této práci byly připraveny fúzní proteiny, pro něž byly *in silico* predikovány možné mezidoménné interakce, obsahující třetí PDZ doménu proteinu ZO-1 spojenou s arteficiálním chameleonským proteinem GA88. Následně byly tyto proteiny charakterizovány širokým spektrem biofyzikálních metod – analytickou vylučovací chromatografií, dynamickým rozptylem světla, analytickou ultracentrifugací, cirkulárním dichroismem a diferenciální skenovací fluorimetrií – za účelem stanovení jejich hydrodynamických vlastností, složení ve smyslu sekundární struktury a teplotní stability.

1 Teoretická část

1.1 PDZ domény

Na pojem proteinová doména lze nahlížet z několika perspektiv. Z hlediska proteinové struktury byl tento pojem zprvu vyčleněn pro kompaktní části proteinu vymezené od sebe v prostoru, z pohledu biochemie jde o funkční podjednotky proteinů, zatímco z hlediska evolučního označuje tento termín části proteinů, jejichž sekvence jsou silně zachovány mezi příbuznými organismy. Ačkoliv to tyto definice v principu nevyžadují, ve skutečnosti často všechny tři perspektivy vymezují v proteinu coby domény velmi podobné oblasti, neboť prostorová struktura determinuje funkci, která má následně tendenci být evolučně konzervována (Ponting a Russell, 2002).

Zatímco některé proteinové domény jsou evolučně relativně nové, vyskytují se např. pouze v živočišné říši, jiné rodiny, např. domény proteinů centrálního metabolismu a některé domény, jejichž ústřední funkcí je zajišťování protein-protein interakcí, byly identifikovány jako extrémně staré, přítomné napříč celým popsáním evolučním vývojem. Jednou z nich jsou právě PDZ domény (Ponting a Russell, 2002; Sakarya et al., 2010; Kennedy, 1995). Ty jsou široce zastoupeny od prokaryotických proteinů po ty lidské, přičemž jejich konzervativní struktura i mezi vzdáleně příbuznými druhy vypovídá o esenciálních rolích, kterých se mohou v buňkách ujímat (Ponting, 1997; Liu a Fuentes, 2019).

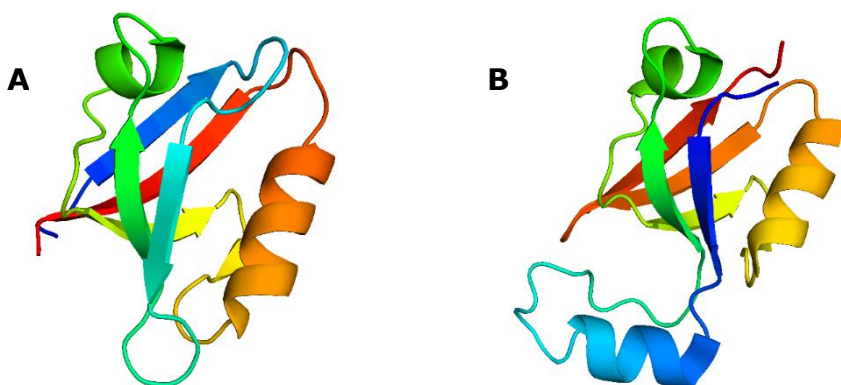
Jejich jméno představuje akronym odvozený od prvních proteinů, v nichž byly v první polovině 90. let 20. století identifikovány – *Postsynaptic density 95 (PSD-95) protein*, důležitý protein savčího synaptického přenosu (viz 2.1.2.2), *Discs-large*, protein přepážkového spoje (z angl. septate junction) epiteliálních buněk v *Drosophila*, a *Zonula occludens 1 (ZO-1)*, klíčový protein těsných spojů (z angl. tight junction) u savců (viz 2.1.2.3) (Kennedy, 1995). PDZ domény nemají katalytickou aktivitu, jsou však schopny vázat krátké peptidové sekvence; na základě těchto interakcí pak vykonávají většinu svých známých fyziologických funkcí (Lee a Zheng, 2010).

1.1.1 Obecná charakteristika PDZ domén

1.1.1.1 Struktura PDZ domén

Kanonické PDZ jsou malé globulární domény čítající zpravidla kolem 80-100 aminokyselin složené z šesti β -skládaných listů a dvou α -helixů (nazývaných někdy krátký a dlouhý) v pořadí $\beta 1$ - $\beta 2$ - $\beta 3$ - $\alpha 1$ - $\beta 4$ - $\beta 5$ - $\alpha 2$ - $\beta 6$, přičemž jejich C-konec a N-konec se nachází v prostoru velmi blízko sobě. Jejich vazebné místo tvoří kavita mezi $\beta 2$ a $\alpha 2$ společně se smyčkou (z angl. loop) mezi $\beta 1$ a $\beta 2$ (tzv. karboxylát vázající smyčka) (Lee a Zheng, 2010; Liu a Fuentes, 2019; Harris a Lim, 2001). Rovněž smyčky mezi $\beta 2$ a $\beta 3$ a mezi $\alpha 2$ a $\beta 6$ se mohou podílet na vazbě některých ligandů s některými PDZ, ovlivňovat svou délkou a složením dynamiku domény a vazebnou specifitu (Tyler et al., 2010).

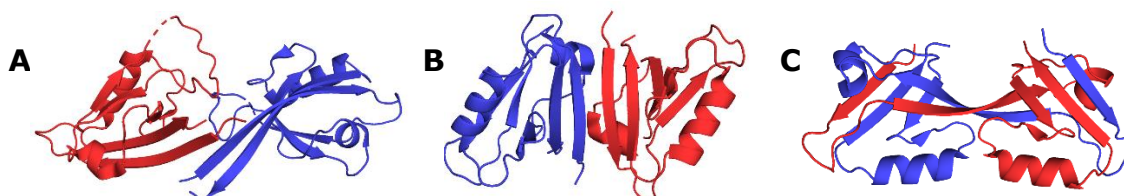
Za nekanonické PDZ (či PDZ-like) jsou označovány podobně prostorově uspořádané domény složené z pěti až šesti β -skládaných listů a dvou až tří α -helixů, kde jsou však sekundární struktury řazeny v jiném pořadí (srov. *Obr. 1A* a *Obr. 1B*) (Lee a Zheng, 2010; Runyon et al., 2007). Zatímco sekvenční identita PDZ domén může tedy být velmi nízká (okolo 20 %), je to právě charakteristická prostorová struktura, která je zachována u prakticky všech PDZ domén (van Ham, 2003).



Obr. 1: Porovnání PDZ domén. A: kanonická struktura třetí PDZ domény lidského ZO-1 (PDB ID: 3TSV). B: nekanonická PDZ doména lidské serinové proteasy Htra2 (PDB ID: 1LCY), upraveno pomocí PyMOL Molecular Graphics System, Verze 2.4.1, Schrödinger, LLC.

Některé PDZ domény jsou náchylné k dimerizacím či oligomerizacím. Homodimerizace mohou probíhat vzájemnými interakcemi $\beta 1$ listů (viz *Obr. 2A, 2B*), vazebné místo pro ligand buďto není touto interakcí změněno (např. PDZ6 GRIP-1 homodimer), nebo jsou jeho rozpoznávací vlastnosti vůči ligandu dimerizací posíleny (např. Shank-1 homodimer) (Im et al., 2003a; Im et al., 2003b). Speciálním typem je homodimer PDZ2 ZO-1 či ZO-2, v němž dochází k vnitřnímu propletení dvou PDZ domén výměnou $\beta 2$ listů (tzv. β -strand-swapped či domain-swapped dimer) (viz *Obr. 2C*), přičemž tento dimer si uchovává dvě vazebná místa. Obdobným mechanismem se rovněž tvoří heterodimery ZO-1/ZO-2 a ZO-1/ZO-3 (nicméně ZO-2 a ZO-3 heterodimer netvoří) (Wittchen et al., 1999;

Fanning et al., 2007). PDZ-PDZ interakce (dimerizace, oligomerizace) napomáhají scaffoldové funkci (tj. funkci opory při sestavování proteinových komplexů) proteinů obsahujících PDZ domény a podílí se na stavbě proteinových komplexů, kde mohou facilitovat držení jednotlivých funkčních komponent v optimální blízkosti. PDZ-PDZ interakcí je patrně schopno velké množství PDZ domén, kolem 30 % savčích, nicméně konkrétní role řady z nich zatím nebyly stanoveny (Chang et al., 2011; Liu a Fuentes, 2019).



Obr. 2: Dimerizace PDZ domén. A: Shank1 homodimer (PDB ID: 1Q3P). B: PDZ6 GRIP-1 homodimer (PDB ID: IN7F). C: PDZ2 ZO-1 homodimer (PDB ID: 3CYY), upraveno pomocí softwaru PyMOL Molecular Graphics System, Verze 2.4.1, Schrödinger, LLC.

1.1.1.2 Vazební partneři PDZ domén

Jak již bylo zmíněno v úvodu této kapitoly, PDZ domény mají schopnost kotvit k sobě specifické sekvence, tzv. PDZ vázající motivy (z angl. PDZ-binding motif, PBM), obsažené v některých proteinech, ale také další ligandy zcela jiného charakteru. Vzhledem k této promiskuitě se ligandy PDZ domén kategorizují na kanonické a nekanonické. Mezi kanonické patří sekvence specifické C-konce peptidů a proteinů (C-terminální PBM), což koreluje se scaffoldovou funkcí této domény. Sekvence aminokyselinových reziduí ligandu se popisují od C-terminálního (P_0) směrem k N-konci (P_{-1} , P_{-2} atd.), přičemž obvykle je do vazby zapojeno 4-7 reziduí (P_0 - P_{-3} resp. P_{-6}) (Liu a Fuentes, 2019). Rozdělení těchto ligandů podle obecné struktury bylo postupně expandováno ze dvou na šestnáct skupin, nejčastěji jsou však používány tři až čtyři hlavní třídy (ačkoliv autoři se jejich vymezení ne vždy shodují) (viz *Tabulka 1*) (Stricker et al., 1997; Tonikian et al., 2008; van Ham, 2003; Liu a Fuentes, 2019; Lee a Zheng, 2010). Výzkumy na úrovni proteomu však poukazují na možnost, že dělení ligandů je čistě arbitrární, mnohé PDZ domény mohou interagovat s ligandy více tříd a s různou afinitou (Stiffler et al., 2007).

Nejen C-terminální, ale i interní aminokyselinové sekvence mohou sloužit jako ligandy pro PDZ domény, a to i se zapojením P_{+1} až P_{+3} reziduí (Penkert et al., 2004). Ačkoliv jsou tyto interakce považovány za nekanonické, ukazuje se, že PDZ domény schopné interakcí s interními peptidovými sekvencemi patrně mohou tvořit na úrovni genomu významný podíl, kolem 50 % (Lenfant et al., 2010).

Tabulka 1: Obecné struktury ligandů PDZ domén. rozdělené do tří hlavních tříd (I-III) podle P₂ a šestnácti podtříd (1a-4a) podle celého řetězce. X označuje libovolnou aminokyselinu, Ω představuje aromatické aminokyseliny (F, W, Y), Ψ vyjadřuje alifatické aminokyseliny (V, I, L, M), zatímco Φ mohou být všechny hydrofobní aminokyseliny (Ω+Ψ) (Tonikian et al., 2008).

Třída	P ₆	P ₅	P ₄	P ₃	P ₂	P ₁	P ₀ (C-konec)
I	X	X	X	X	T/S	X	Φ
II	X	X	X	X	Φ	X	Φ
III	X	X	X	X	D/E	X	Φ
1a	X	Φ	K/R	X	S	D	V
1b	X	Ω	K/R	E	T	S/T/R/K	Φ
1c	X	Φ	Φ	E	T	X	L
1d	X	X	X	E	T	X	V
1e	X	X	X	X	T	W	Ψ
1f	X	X	Ω	Ω	T	W	Ψ
1g	Φ	Φ	Φ	T/S	T/S	Ω	Ψ
1h	X	Φ	Φ	D/E	T/S	W	Ψ
2a	X	X	F	D	Ω	Ω	C
2b	X	W	X	Ω	F	D	V
2c	X	X	W	Ω	Φ	D	Ψ
2d	Φ	Φ	X	D/E	Φ	Φ	Φ
2e	X	X	X	Φ	Φ	Φ	Φ
2f	X	X	X	D/E	Φ	Ω	Φ
3a	X	W	Ω	S/T	D	W	Ψ
4a	X	X	Ω	Φ	G	W	F

Další skupinu nekanonických ligandů tvoří lipidy. Fosfatidylinositolfosfáty (PIPs) fungují v buněčných membránách jakožto verzatilní signální molekuly a jsou tak rozpoznávány řadou proteinových domén, včetně právě PDZ (Liu a Fuentes, 2019). Nejlépe prozkoumány jsou v současnosti vztahy s PI(4,5)P₂, který může interagovat s širokou škálou PDZ domén (např. PDZ1 a PDZ2 syntenin1, PDZ2 ZO-1 a ZO-2) kompetitivně s peptidovým ligandem (Meerschaert et al., 2009; Liu a Fuentes, 2019), avšak jeho afinita k jiným PDZ doménám (např. CASK) může být naopak kanonickým ligandem posilována (Ivarsson et al., 2013). Odhaduje se, že asi 20-30 % savčích PDZ domén je schopno interakce s PIPs (Chen et al., 2012; Ivarsson et al., 2013). Řada PDZ domén (např. Dlg1, Dlg4, NHERF1 PDZ1) rovněž váže cholesterol, který může dokonce sloužit jako modulátor jejich kánonické funkce (např. Dvl2) (Sheng et al., 2012; Sheng et al., 2014).

1.1.2 Fyziologické funkce PDZ

1.1.2.1 Rozdělení proteinů obsahujících PDZ domény

Proteiny, které obsahují PDZ doménu, často v několika kopiích, jsou zpravidla vícedoménové (viz *Obr. 3*) a bylo jich charakterizováno široké spektrum – jen v lidském genomu přes 650 (dle databáze SMART) (Schultz et al., 1998) –, přičemž standardně se dle struktury a funkce tyto proteiny rozdělují do čtyř podrodin:

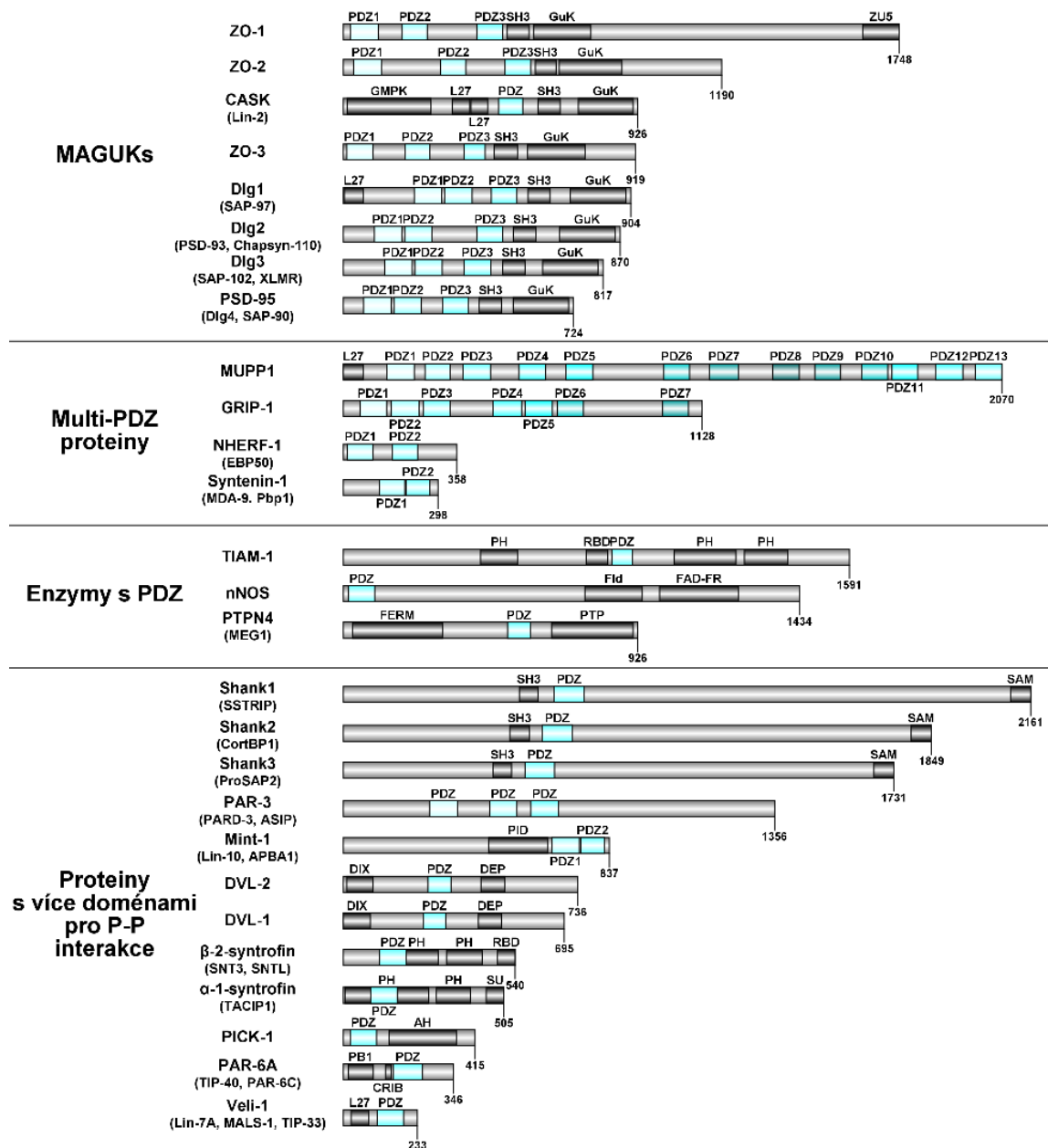
- homology guanylát-kinasy asociované s membránou (MAGUKs) obsahující charakteristickou sérii domén 1-3 PDZ-SH3-GuK jsou spojované s transportem, signalizací, ale hlavně scaffoldovou funkcí;
- multi-PDZ proteiny obsahující pouze PDZ domény (2-13×) vykazují zpravidla čistě scaffoldovou funkci;
- enzymy obsahující PDZ, které se mohou být pomocí své PDZ domény stávat součástí enzymových komplexů např. buněčné signalizace;
- proteiny s několika různými doménami pro protein-protein interakce jsou rozmanitou skupinou s regulačními, signalizačními, strukturními a scaffoldovými funkcemi (Liu a Fuentes, 2019; Harris a Lim, 2001).

1.1.2.2 Synaptický přenos

V lidském organismu zastávají proteiny obsahující PDZ domény řadu rolí, jednou z těch nejdůležitějších je koordinace glutamatergní synaptické transmise. Ta je klíčová v procesech synaptické plasticity excitačních drah – dlouhodobé potenciace (LTP) a dlouhodobé deprese (LTD) –, esenciálních nástrojích učení a paměti. (Maffei, 2018; Kim a Sheng, 2004)

Glutamátový receptor kyseliny α -amino-3-hydroxy-5-methyl-4-isoxazolpropionové (AMPA receptor) je ionotropní kanál tvořený heterotetramerní strukturou podjednotek GluR2 (dimer) a GluR1/GluR3/GluR4 (homodimer). PDZ-PBM interakce jsou důležitou součástí jeho intracelulárního transportu a regulace, a to prostřednictvím rodiny Discs-large homologů (Dlg1-4), proteinu interagujícího s glutamátovým receptorem 1 (GRIP-1) a proteinu interagujícího s protein kinasou C- α 1 (PICK-1) (Kim a Sheng, 2004; Liu a Fuentes, 2019).

Podjednotky GluR2 a GluR3 obsahují C-terminální PBM, který je schopen interakce s PDZ5 GRIP-1; tato interakce jednak patrně pomáhá stabilizovat postsynaptické i intracelulární AMPA receptory, také však zajišťuje transport AMPA receptorů po mikrotubulech z Golgiho aparátu k synaptické štěrbině (tzv. anterográdní transport) prostřednictvím proteinových motorů, kinesinu 1 (KIF5) a kinesinu podobnému proteinu KIF1A. (Kim a Sheng, 2004; Schnapp, 2003). PBM GluR2 a GluR3 rovněž interaguje s PDZ doménou PICK-1; jakou roli má jeho interakce s AMPA receptory, však není objasněno, a ačkoliv jeho genový knockout (v překladu „vyřazení“ genu) minimalizuje LTD, není na rozdíl od knockoutu GRIP-1 letální (Kim a Sheng, 2004; Takamiya et al., 2004).



Obr. 3: Schémata vybraných lidských proteinů s PDZ doménami (modře) rozdělená do příslušných podrodin. Ostatní domény vyznačeny černě. Zpracováno z databáze Uniprot (uniprot.org) pomocí softwaru DOG 2.0 (Ren et al., 2009).

Z rodiny Discs-large homologů interaguje s AMPA receptory přímo pouze Dlg1, a to pomocí své PDZ2, která váže C-terminální PBM GluR1 (Zhou et al., 2008; Leonard et al., 1998; Cai et al., 2002). Dlg1 dále interaguje s mikrotubul-dependentními proteinovými motory pro anterográdní transport (KIF13B, KIF1B α a KIF17) a nebo s aktin-dependentními proteinovými motory pro retrográdní transport (myosin VI) a napomáhá tak exocytóze i endocytóze AMPA receptorů (Walch, 2013; Kim a Sheng, 2004). Knockout Dlg1 je letální, nicméně nikoliv v souvislosti s jeho synaptickou funkcí, ale z důvodu narušení spojů mezi

buňkami (Caruana a Bernstein, 2001) – v postsynaptické oblasti zůstává hladina AMPA receptorů i při knockoutu Dlg1 stejná (Klöcker et al., 2002).

Dlg4 (tj. PSD-95) tvoří komplex s AMPA receptory nepřímo, přes stargazin (CACNG2) (Ma et al., 2014). Stargazin interaguje s GluR1, GluR2 a GluR3, zatímco Dlg4 s KIF1B α (anterográdní transport) (Chen et al., 2000; Kim a Sheng, 2004). Zvýšení množství (over-exprese) Dlg4 potencuje glutamátergní přenos, zatímco jeho knockout přenos pomocí AMPA receptorů utlumuje (a zamezuje LTD); tento efekt však může být kompenzován zvýšenou expresí Dlg1, který jeho funkci následně zastane (Schlüter et al., 2006).

Glutamátový N-metyl-D-aspartátový (NMDA) receptor je ionotropní kanál tvořený heterotetramerní strukturou podjednotek GluN1 (dimer) a GluN2/GluN3 (homodimer nebo heterodimer). I tento protein interaguje s rodinou Dlg PMB-PDZ interakcemi, přičemž Dlg1, Dlg3 a Dlg4 figurují významně v jeho intracelulární distribuci (Liu a Fuentes, 2019; Sans et al., 2000; Kim a Sheng, 2004; Elias et al., 2008). Zejména pak Dlg4 reprezentuje signifikantní scaffoldový protein, díky němuž mohou být NMDA receptory napojeny k neuronální NO-synthase (nNOS). Jejich komplex se ukotvuje k plazmatické membráně a spouští řadu pro učení a paměť (na synaptické úrovni) vitálních signalizačních kaskád (Tochio et al., 2000), nicméně zvýšenou aktivací glutamátových drah může také nadprodukovat NO a způsobovat excitotoxické patologické stavy (Sattler et al., 1999).

Také metabotropní glutamátové receptory (mGluRs) interagují pomocí C-terminálních PBM, a to s PDZ PICK-1, a tato interakce se významně podílí na stabilizaci mGluR na membráně a na presynaptické negativní zpětné regulaci vyplavování glutamátu (Perroy et al., 2002). Obdobné PDZ-PBM interakce (s jinými proteiny) byly identifikovány i u dalších receptorů – serotoninových, α -adrenergních i β -adrenergních, κ -opioidních, adenosinových, lutropinových a řady dalších (Romero et al., 2011).

Role výše zmíněných proteinů v nervové soustavě je samozřejmě komplexnější, GRIP-1 byl nalezen také v presynaptických neuronech, kde se podílí na synaptickém zrání, a také v GABAergním systému, kde je jeho funkce nejasná (Li et al., 2005; Naylor et al., 2005; Wyszynski et al., 2002); CASK se rovněž vyskytuje v presynaptickém neuronech, kde kotví neurexin 1 β (NRXN), který transsynapticky váže komplex neuroligin 1 (NLGN)/PDZ3 Dlg4 a zajišťuje tak adhezi presynaptického neuronu s postsynaptickým, stabilizuje a moderuje synaptický přenos (Missler et al., 1998; Südhof, 2008); Dlg4 váže svou SH3 doménou také kainátové glutamátové receptory (Garner et al., 2000) atd. Zajímavým faktem rovněž zůstává, že hladiny jednotlivých zástupců Dlg rodiny se během synaptického zrání mění – zatímco množství Dlg3 v buňkách s časem klesá, je tedy patrně zapojen do synaptogeneze, množství Dlg2 a Dlg4 se zvyšuje a figurují tak zřejmě v zrání samotném (Sans et al., 2000; Elias et al., 2008).

1.1.2.3 Buněčné spoje

V konstituci buněčných spojů, volných (gap junction, *nexus*), adhesních (adherens junction, *zonula adherens*), těsných (tight junction, *zonula occludens*) atd. hrají v principu esenciální roli protein-protein interakce, které tvoří jejich funkční, strukturní i regulační

fundament (Adil et al., 2021), přičemž u všech zmíněných typů spojů se v těchto interakcích hojně uplatňují právě PDZ domény (Liu a Fuentes, 2019).

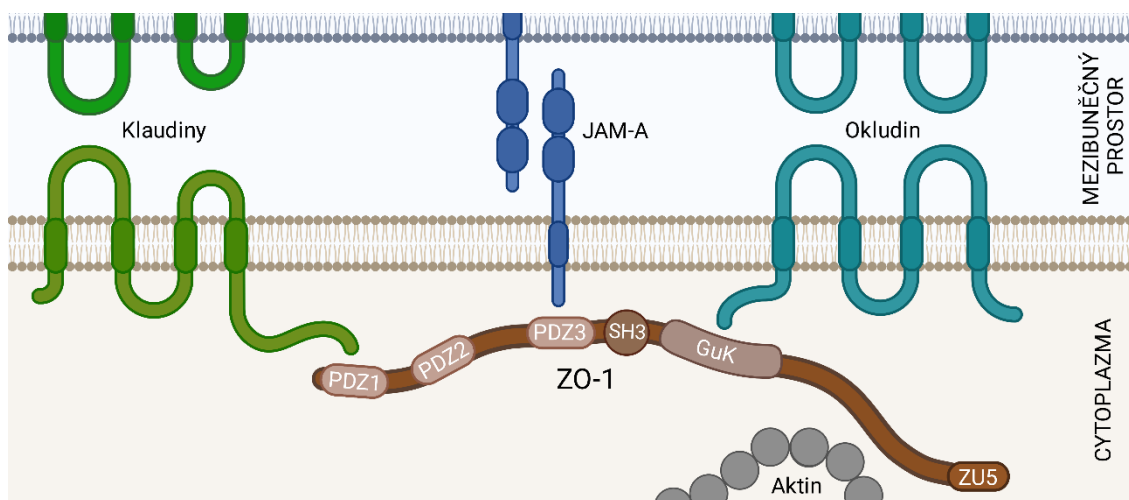
Buňky spojené volnými spoji vykazují schopnost silné metabolické a energetické intercelulární komunikace zajišťované konexony, které vytváří transmembránové kanály pro univerzální látkovou výměnu malých molekul (průměr pod 15 Å, tj. cca 1 kDa). Obě buňky přispívají za účelem formace tohoto kanálu jedním konexonem (*hemi*-kanálem), hexamerem složeným z šesti proteinů konexinu (lidských konexinů je rozlišováno 21 ve třídách α - δ , jejich exprese je tkáňově specifická) (Wei et al., 2004; Kumar a Gilula, 1996). Celý *nexus* je následně tvořen shlukem těchto kanálů, který je ohraničen množstvím *hemi*-kanálu v oblasti tzv. *perinexu*. (Rhett a Gourdie, 2012).

Ačkoliv jsou proteiny rodiny Zonula occludens (ZO-1 až ZO-3), jak jméno napovídá, spojovány primárně s těsnými spoji, prokazují také schopnost ovlivňovat významně spoje volné (Wei et al., 2004). ZO-1 je pomocí své PDZ2 (i v homodimerní formě, viz 2.1.1.1) (Fanning et al., 2007) schopen interagovat s C-terminálními PBM celé řady konexinů třídy α , zejména s Cx43 (tj. α 1), Cx46 (α 3), Cx50 (α 8), Cx31.9 (α 11) a Cx47 (α 12) (Nielsen et al., 2003; Li et al., 2004a). Prokazatelná je rovněž analogická interakce ZO-2 s Cx43 (Singh et al., 2005). Interakce ZO-1/Cx43 je regulována fosforylací Cx43 pomocí proto-onkogen tyrosin kinasy sarkomu (c-Scr) (Toyofuku et al., 2001) a má regulační funkci – aktivně snižuje velikost volných spojů (Hunter et al., 2005), a to pravděpodobně stabilizací konexonových *hemi*-kanálů v oblasti *perinexu*, kde se ZO-1 akumuluje, a blokáde expanze *nexu* (Rhett a Gourdie, 2012).

Těsné spoje jsou přítomny mezi epiteliálními buňkami, které musí tvořit pevnou bariéru bránící difuzi mezi mnohdy značně odlišnými prostředními apikální a bazální strany. Permeabilita těsných spojů se může lišit dle požadavků dané tkáně v selektivitě i rozsahu a je definována jejich proteinovým složením a hustotou sítě pojících vláken (z angl. sealing strands) (Alberts et al., 2002).

Mezi složky těsných spojů se řadí transmembránové proteiny jako kladiny, okcludin (resp. proteiny obsahující tzv. MARVEL doménu), JAM-A, intracelulární proteiny jako ZO (viz Obr. 4) nebo Par-komplex a také membránový cholesterol (Shigetomi et al., 2018; Liu a Fuentes, 2019). Kladiny (lidských bylo identifikováno přes 20) obsahují konzervativní C-terminální PBM, díky kterému jsou schopny vazby na PDZ1 doménu ZO-1, ZO-2 i ZO-3 (Itoh et al., 1999), a následně čtyřikrát prostupují membránou. Extracelulární části kladinů se pak účastní homodimerizací (Van Itallie et al., 2011) nebo heterodimerizací (Gong et al., 2015) za vzniku póru selektivnímu k náboji o bazálně-apikální permeabilitě pod 4 Å (Otani a Furuse, 2020).

Skupina imunoglobulinu podobných proteinů JAM (z angl. junctional adhesion molecules) čítá členy JAM-A, JAM-B a JAM-C. Zatímco JAM-B a JAM-C tvoří extracelulární heterodimerní struktury, které se účastní buněčné adheze (zejména lymfocytů k endotelu) (Lamagna et al., 2005), JAM-A je jednou z klíčových komponent těsných spojů. Intracelulárním C-terminálním PBM se ukotvuje k PDZ3 ZO-1 a extracelulární N-terminální část homodimerizuje (až oligomerizuje) se sousedícím JAM-A, čímž vzniká větší pór o permeabilitě pod 60 Å (Otani a Furuse, 2020; Itoh et al., 2001).



Obr. 4: Schéma interakcí mezi proteiny těsného spoje. Zpracováno pomocí online nástroje BioRender (biorender.com).

Jak je evidentní z výše uvedených informací, jsou to proteiny rodiny ZO, které lze v rámci těsných spojů považovat za ubikvitní scaffold. ZO-1 a ZO-2 se navíc přímo upoutává k vláknům F-aktinu (a nepřímo i k dalším komponentám cytoskeletu) a zajišťuje tak celému proteinovému komplexu oporu o cytoskelet (Fanning et al., 1998; Van Itallie et al., 2017; Wittchen et al., 1999; Paris et al., 2008). Při knockoutu ZO-1 i ZO-2 dochází prenatálním úmrtím spojenými s buněčnými spoji, nicméně knockout ZO-3 letální není, ani se neprojevuje ve fenotypu (Xu et al., 2008; Katsuno et al., 2008).

Regulace těsných spojů je rovněž ovlivňována proteinovými komplexy tzv. systému buněčné polarizace (z angl. cell polarity system), mezi nimiž se za ústředního aktéra označuje Par (z angl. partition-defective) komplex. Ten je tvořen scaffoldovým proteinem s PDZ doménou PAR-6 (partitioning defective 6 homolog, A/C, B, D/G) (Paris et al., 2008; Assémat et al., 2008), atypickou protein kinasou C (aPKC ζ a λ/ι) (Hirano et al., 2005) a PAR-3 (v epitelu nejčastěji PAR-3A) (Assémat et al., 2008), přičemž tento komplex podporuje tvorbu těsných spojů (antagonizováno podobným komplexem PAR-6/aPKC/LLGL) (Lin et al., 2000; Paris et al., 2008; Schneeberger a Lynch, 2004). PAR-3 má dále schopnost se asociovat pomocí své PDZ3 s JAM-A, což celému komplexu umožňuje lokalizaci u těsných spojů (Itoh et al., 2001), do nichž může aPKC následně zasahovat fosforylací ZO-1, okludinu a klaudinů (antagonizováno proteinovou fosfatázou 2A) (Nunbhakdi-Craig et al., 2002).

1.1.2.4 Signalizace

S buněčnými spoji souvisí také fenomén buněčné adheze a migrace, která je regulována mimo jiné transmembránovými proteoglykanovými receptory, např. syndekany (členové třídy heparansulfátových proteoglykanů – HSPGs). Tyto signalizují vnější podněty (cytokiny, chemokiny, kolagen) dovnitř buňky napojením intracelulárního C-konce na protein kinasu C- α a její následnou aktivaci (Keum et al., 2004), nebo interakcemi jejich

C-terminálního PBM s PDZ doménami, např. s proteinem indukujícím invazi a metastázi lymfomu T-buněk 1 (TIAM-1) (Shepherd et al., 2010). Obě tyto možnosti mohou interferovat se signalizačními dráhami Rho GTPas (RhoA, Rac1), esenciálními komponentami celé řady buněčných procesů – buněčného růstu, transkripce, organizace cytoskeletu a rovněž buněčné adheze (Van Aelst a D'Souza-Schorey, 1997; Mackay a Hall, 1998).

Do kaskád Rho GTPas však zasahují proteiny s PDZ doménami i jinak, a ani těmito cestami nejsou jejich signalizační úlohy vyčerpány. Např. plazmatické růstové faktory Wingless/Int-1 (Wnt), významné pro embryonální vývoj, ale i buněčnou homeostázu dospělých tkání (Clevers, 2006; Malbon a Wang, 2005), spouštějí po vazbě na svoje receptory rodiny Frizzled (Fz) (Bhanot et al., 1996; Wang et al., 1996) tzv. Wnt signalizační kaskádu. Je známa řada forem, cest Wnt signalizace (kanonická, tj. β -katenin-dependentní, a několik nekanonických), ve valné většině však prominentně figuruje právě interakce Wnt/Fz s proteinem s PDZ doménou Dishevelled (Wawrzak et al., 2009; Semenov et al., 2007). Zatímco kanonická Wnt stimuluje buněčnou migraci a potlačuje adhezi (Nusse, 2012), nekanonické Wnt signalizace produkují (mimo jiné ovlivněním dráhy Rho GTPas) širší spektrum komplexních kauzalit interferujících s vlastnostmi cytoskeletu, systémem buněčné polarita i genovou transkripcí (Semenov et al., 2007; Liu a Fuentes, 2019; Clark et al., 2012).

1.1.2.5 Onemocnění spojená s PDZ-PBM interakcemi

Ve spojitosti s buněčnými spoji a adhezí tvoří rovněž proteiny obsahující PDZ domény výhodné cíle pro patogenní viry. Řada virových proteinů obsahuje PBM, často C-terminálních, které interakcemi s buněčnými PDZ mohou narušovat buněčné pochody a adhezi, a to až do bodu onkogeneze (např. adenoviry, lidský virus leukemie T-buněk – HTLV, či lidský papillomavirus – HPV). S PDZ doménami však interagují i proteiny neonkogenních virů, např. viru vztekliny, SARS (SARS-CoV i SARS-CoV-2) a chřipkového viru A (Liu a Fuentes, 2019; Shepley-McTaggart et al., 2021).

Další onemocnění mohou být spojena s nesprávnou funkcí fyziologických proteinů s PDZ doménami, zpravidla excesivními interakcemi. Nejčastěji jde o onemocnění onkologického nebo neurologického charakteru (viz *Tabulka 2*), jak predestinují jejich známé funkce popsané výše (viz 1.1.2.2, 1.1.2.3, 1.1.2.4). PDZ domény tak tvoří perspektivní farmakologické cíle, ať už ve smyslu specifických antivirotik nebo potenciálních léčiv celé řady nádorových (např. inhibitory Wnt signalizace skrze Dvl) a neurologických onemocnění (např. inhibitory NMDAR/Dlg4/nNOS komplexu); v tomto směru je však zatím vývoj ve velmi ranném stádiu (Liu a Fuentes, 2019).

Tabulka 2: Popsané interakce PDZ domén, jež se podílí na vzniku onemocnění. Převzato a upraveno z Liu a Fuentes, 2019.

PDZ-PBM interakce	Funkce	Onemocnění
PDZ1 Dlg4/NMDAR PDZ2 Dlg4/nNOS	Regulace glutamátergního přenosu	Mrtvice, ischemie mozku
PDZ1 Dlg4/5-HT _{2A}	Regulace GPCR signalizace a neuronální excitace	Neuropatické bolesti
PDZ PICK-1/ErbB2 PDZ PICK-1/GluR2 PDZ PICK-1/DAT	Regulace exocytózy a endocytózy receptorů	Deprese, schizofrenie, neuropatické bolesti, rakoviny
PDZ Shank-3/GKAP	Regulace synaptické aktivity	Autismus, retardace
PDZ CASK/NRXN PDZ3 Dlg4/NLGN PDZ syntropin/NLGN3	Regulace transsynaptického spojení	Poruchy autistického spektra
PDZ1 syntenin 1/c- <i>Src</i>	Posttranslační signalizace	<i>Glioblastoma multiforme</i> , melanom
PDZ AF-6/BCR PDZ LARG/CD44 PDZ TIAM-1/sydekan 1	Regulace buněčné migrace a dynamiky cytoskeletu	Akutní lymfoblastická leukemie, rakovina prsu a prostaty
PDZ Par-6/Par-3 PDZ Par-3/PTEN PDZ1 Dlg1/β-katenin	Regulace tvorby těsných spojů	Epitelo-mezenchymální tranzice buněk tumoru
PDZ NRXNR2/LPA	Regulace CTFR signalizace	Cystická fibróza

1.1.3 PDZ3 ZO-1

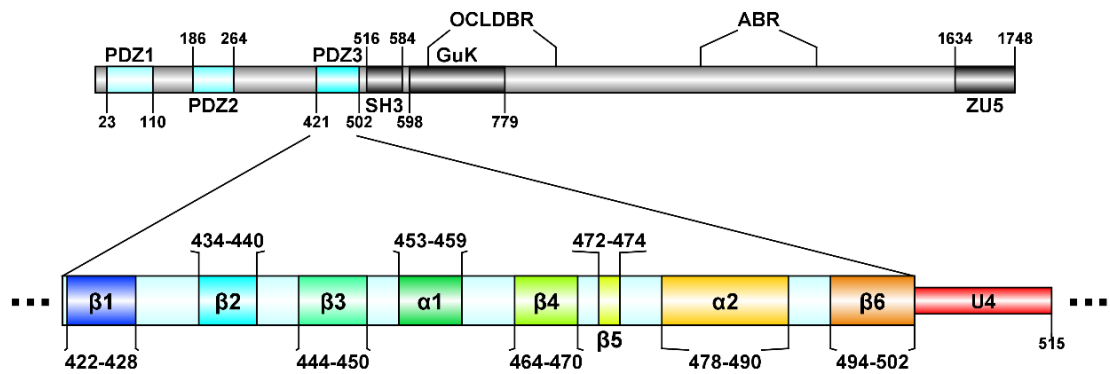
Role proteinu ZO-1 byla již široce diskutována v předešlých kapitolách (viz 1.1.2.3). Jedná se o vícedoménový protein ze skupiny MAGUKS o délce 1748 aminokyselin a velikosti 195,5 kDa (uniprot.org), jehož hlavní doposud popsané role souvisí s volnými a těsnými spoji. V obou případech zde působí jako scaffoldové centrum, které pomocí svých domén pro protein-protein interakce umožňuje mnoha komponentám správnou lokalizaci, funkci a stabilizaci o cytoskelet (Fanning et al., 1998; Nielsen et al., 2003; Itoh et al., 1999). Na jeho N-konci a v meziprostoru jeho šesti domén (PDZ1-PDZ3, SH3, GuK, ZU5, viz Obr. 5), se nachází šest, tzv. unikátních oblastí (U1-U6), silně variabilních mezi proteiny rodiny ZO. Vzhledem k faktu, že společně tvoří téměř dvě třetiny celého ZO-1, mají tyto oblasti patrně značný vliv na jeho prostorové uspořádání a funkci, ačkoliv u většiny z nich je predikována neuspořádaná struktura (výjimku tvoří U4) (Bewley et al., 2013; Bauer et al., 2010).

PDZ1 doména ZO-1 je zapojována především při konstrukci těsných spojů, kde ukotvuje kladiny (C-konec -NYV/-DYV/-AYV ad.) (Itoh et al., 1999), nicméně minoritně se podílí i na stavbě spojů volných, a to zejména v mozkové tkáni, nadledvinkách a pankreatu, neboť interaguje také s Cx36 (-SAYV) (Li et al., 2004b; Li et al., 2004c). S proteiny této rodiny se však častěji váže spíše PDZ2 doména, a to např. s Cx31.9 (-DLAI), Cx43 (-DLDV), Cx46 (-DLAI), Cx47 (-TVWI) a Cx50 (-DLTV) (Nielsen et al., 2003; Li et al., 2004a), jak již bylo předesláno výše. Tato doména je rovněž schopna homodimerizace výměnou $\beta 2$ listů, která zvyšuje její specifitu vůči Cx43 (Fanning et al., 2007; Chen et al., 2008b), a heterodimerizace s PDZ2 ZO-2 a ZO-3 (Wittchen et al., 1999; Fanning et al., 2007).

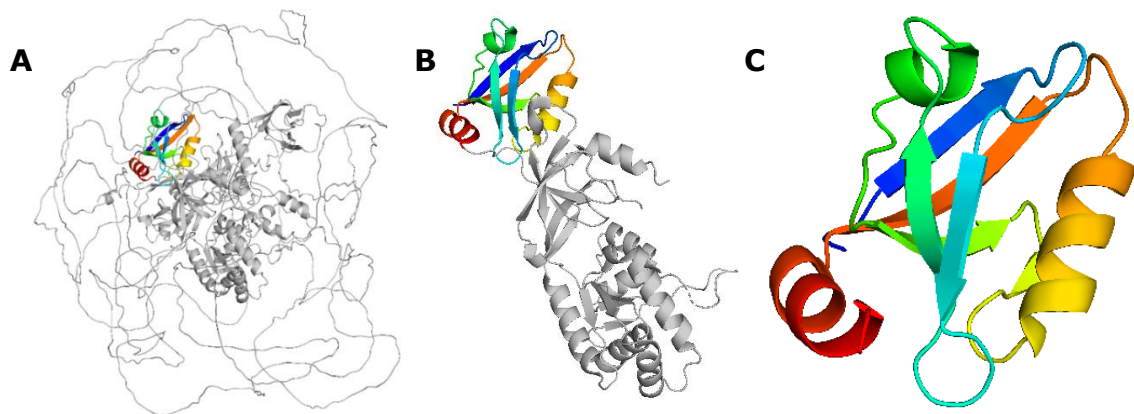
PDZ3 doména ZO-1, která je součástí fúzních proteinů studovaných v této práci, je třetí z jeho domén (viz Obr. 5 a Obr. 6A). Je tvořena aminokyselinami na pozicích 421-502 (viz Tabulka 3) v terciární struktuře kanonické PDZ domény (viz Obr. 6C), přičemž na jejím C-konci ji následuje α -helikální unikátní oblast U4 (označovaná též jako nativní linker), napojující PDZ3 na SH3 doménu. Tento linker hraje významnou roli ve schopnosti PDZ3 vázat své ligandy, neboť do těchto interakcí se zapojuje rovněž doména SH3, jejíž orientaci vůči PDZ3 udržuje právě nativní linker (Pan et al., 2011; Nomme et al., 2011).

Tabulka 3: Sekvence PDZ3 domény ZO-1 (černě) a jejího nativního linkeru (šedivě) včetně vymezení v rámci proteinu ZO-1 (číslování nad sekvencí). Data dostupná z databáze Uniprot, ID Q07157 (uniprot.org).

425	430	435	440	445	450	455	460
SMKLV	KFRKG	DSVGL	RLAGG	NDVGI	FVAGV	LEDSP	AAKEG
465	470	475	480	485	490	495	500
LEEGD	QILRV	NNVDF	TNIIR	EEAVL	FLLDL	PKGEE	VTILA
505	510	515					
QKKKD	VYRRI	VESDV					



Obr. 5: Schématické zobrazení proteinu ZO-1 a jeho PDZ3 domény. Výše schéma ZO-1, vyznačeny jsou jeho domény (PDZ1-3 modře, SH3, GuK, ZU5 černě), vazebná místa (okludin vázající oblast – OCLDBR – a aktin vázající oblast – ABR) a unikátní oblasti U1-U6 (šedě). Níže jeho PDZ3 doména s nativním linkerem (U4), vyznačeny jsou sekundární struktury, jejich pozice a pořadí. Data dostupná z databázi Uniprot, ID Q07157 (uniprot.org) a RSCB PDB, ID 4Q2Q, 3TSZ a 3TSV (rcsb.org). Zpracováno pomocí softwaru DOG 2.0 (Ren et al., 2009).



Obr. 6: Prostorové uspořádání PDZ3 domény proteinu ZO-1. PDZ3 vyznačena barevně, barvy korespondují s Obr. 5. A: ZO-1 (in silico predikce), převzato z databáze AlphaFold, ID AF-Q07157-F1 (alphafold.ebi.ac.uk). B: Tandem PDZ3-SH3-GuK (PDB ID: 3TSZ). C: PDZ3 doména s nativním linkerem (PDB ID: 3TSV) (obě stanoveno rentgenovou krystalografií). Upraveno pomocí softwaru PyMOL Molecular Graphics System, Verze 2.4.1, Schrödinger, LLC.

Mezi známé ligandy PDZ3 ZO-1, jejichž interakce byly detailně popsány, patří JAM-A (-SFLV) a Cx45 (-SVWI). I₀ PBM Cx45 je patrně vázáno vodíkovými můstky s využitím V433, G434 a L435, W₋₁ interaguje hydrofobně s L435 a L451, zatímco V₋₂ je ukotveno hydrofobní kapsou tvořenou V484, L485, L549 a L554 (poslední dva jsou součástí domény SH3) a S₋₃ může být vázáno pomocí R436 (Pan et al., 2011). JAM-A je zprostředkována interakcí V₀ hydrofilně s V433, G434, a L435 a hydrofobně s V484 a L488, F₁ s hydrofobní kapsou V484, L549 a L554, S₃ hydrofilně s L437 (Nomme et al., 2011). Mezi těmito vazbami je zřejmá značná shoda, obě zapojují především $\beta 2$ list, $\alpha 2$ helix PDZ3 a také několik hydrofobních reziduí z vedlejší SH3 domény.

1.1.3.1 Dynamická alosterie PDZ3 ZO-1

Jak vyplývá z podstaty interakce ligandů PDZ3 ZO-1 s jejími partnery (viz 1.1.3), na regulaci funkce PDZ3 ZO-1 mají značný vliv strukturní komponenty, které doménu obklopují, zejména pak nativní linker (Pan et al., 2011; Boušová et al., 2021). Nejedná se však o přímou alosterickou regulaci, spíše o vzdálené vzájemné působení, relativně nedávno popsáný jev nazývaný dynamická alosterie (Grutsch et al., 2016).

Mechanismus dynamické alosterie není zatím dobře známý, mohou se na ní podílet sterické, termodynamické a elektrostatické efekty či specifické vlastnosti linkerů, hlavním rozdílem od standardní alosterické regulace je však ovlivnění funkce beze změny konformace domény (Kumawat a Chakrabarty, 2017; Ma et al., 2011). Nejen PDZ3 ZO-1, ale i PDZ domény jiných proteinů (PDZ3 PSD-95, PDZ PAR-6 ad.) vykazují podobně chování, přičemž hlavní centra dynamické alosterie se nachází v $\alpha 1$ helixu, nativním linkeru a smyčce po $\alpha 2$ helixu (Stevens a He, 2022; Petit et al., 2009).

Dynamická alosterie byla pozorována i u zcela jiných proteinových rodin neobsahujících PDZ domény, např. cyklofilinů (Holliday et al., 2017).

1.1.3.2 Charakteristika použité PDZ3

V rámci této práce byla použita mírně modifikovaná varianta PDZ3 ZO-1 obsahující histidinovou kotvu, štěpné místo pro TEV proteasu, krátkou vektorovou sekvenci (přirozeně se vyskytující ve vektoru pET15b) a tzv. signální sekvenci, tj. chybu designu, dle nukleární magnetické rezonance nestrukturovanou a bez interakce s PDZ3 (Boušová, 2023, soukromá konverzace) (viz *Tabulka 4*). Charakterizace biofyzikálními metodami byly prováděny s proteinem s histidinovou kotvou i bez ní (po štěpení TEV proteasou).

Tabulka 4: Sekvence a charakteristika PDZ3 domény použité v experimentální části této práce. Molekulová hmotnost byla predikována online nástrojem ProtParam (Gasteiger et al., 2005). PDZ3 (tmavě zeleně), histidinová kotva (tmavě modře), signální sekvence (světle modře), vektorová sekvence (šedivě), štěpná sekvence pro TEV proteasu (oranžově).

Protein		Počet reziduí				M_w [Da]		
6His-PDZ3		114				12 539,27		
5	10	15	20	25	30	35	40	
MGSSH	HHHHH	SSGEN	LYFQG	HMLEG	SHMIL	RPSMK	LVKFR	
45	50	55	60	65	70	75	80	
KGDSV	GLRLA	GGNDV	GIFVA	GVLED	SPAAG	EGLEE	GDQIL	
85	90	95	100	105	110	115		
RVNNV	DFTNI	IREEA	VLFLI	DLPKG	EEVTI	LAQK		
PDZ3		95				10 327,95		
5	10	15	20	25	30	35	40	
GHMLE	GSHMI	LRPSM	KLVKF	RKGDS	VGLRL	AGGND	VGIFV	
45	50	55	60	65	70	75	80	
AGVLE	DSPAA	KEGLE	EGDQI	LRVNN	VDFTN	IIREE	AVLFL	
85	90	95						
LDLPK	GEEVT	ILAQK						

1.2 G_A88, G_B88

1.2.1 Chameleonské proteiny

V roce 1994 vyhlásili Rose a Creamer v rámci své práce o predikování vztahu mezi aminokyselinovou sekvencí a způsobem, jakým se protein skládá, tzv. Paracelsovu výzvu, jejíž podmínkou splnění byla transformace jednoho globulárního proteinu na jiný s odlišnou terciární strukturou při zachování alespoň poloviny jeho primární struktury (Rose a Creamer, 1994). Dva z pokusů podmínku naplnit, *de novo* designovaný paracelsin-43 (Jones et al., 1996) a hybridní Crotein-G (Yuan a Clarke, 1998), skončily neúspěchem. Nebyly schopny očekávané terciární struktury *in vitro* ani částečně dosáhnout bez použití radikálních podmínek (extrémně nízké teploty, koncentrace), jevily se neuspořádané nebo silně náchylné k agregaci (Dalal et al., 1997a).

Protein Janus, navržený *de novo* se shodou 50 % s B1 doménou proteinu G bakterií *Streptococcus* (s $\alpha/4\beta$ konformací) a očekávanou 4α -helikální konformací bakteriálního proteinu Rop (Dalal et al., 1997b), byl prvním z kandidátů, který naplnil Paracelsovu výzvu (Rose, 1997) a stal se tak jedním z nejdůležitějších milníků na poli proteinového designu a arteficiálních proteinů. Po prvním *de novo* designu proteinu (Regan a DeGrado, 1988) a vyhlášení proslulé soutěže CASP (z angl. Critical Assessment of Structure Prediction) (Moult et al., 1995; Moult, 2005) a od ní odvozených soutěží (Janin, 2002), byla Paracelsova výzva dalším krokem v chápání centralizované, lokální (tj. determinované jen určitým počtem aminokyselin v sekvenci, nikoliv rovnoměrně distribuované) kontroly terciární struktury proteinů (Lattman a Rose, 1993; Dalal et al., 1997b; Rose a Creamer, 1994).

V návaznosti na tyto práce vznikla řada dalších *de novo* designovaných i hybridních proteinů, někdy označovaných jako chameleonské proteiny (z angl. chameleon proteins), neboť obsahují řadu sekvencí, které jsou schopné zaujímat vícero sekundárních struktur (tj. chameleonských sekvencí). Jedním z oblíbených templátů zůstal protein G bakterií *Streptococcus* (Dalal a Regan, 2000; Blanco et al., 1999; Roy et al., 2014), který je schopen interakcí se sérovými proteiny – s albuminem (HSA) pomocí domén G_A (A1-2/3) a s imunoglobulinem G (IgG) pomocí domén G_B (B1-2/3) (Nezlin, 1998). Právě tyto domény se staly výchozí komponentou další skupiny konstruktů.

G_A nabírá konformaci tří α -helixů, zatímco G_B je složeno z jednoho α -helixu a čtyř β -listů, přičemž tyto dvě domény sdílí pouhých 16 % (9 reziduí) sekvence. Substitucemi v obou sekvencích byla zkonstruována série proteinů se zvyšující se sekvenční identitou, ovšem se zachováním původních konformací G_A a G_B, a to proteiny G_A30/G_B30, G_A77/G_B77 a G_A88/G_B88 (vzájemná procentuální identita udána číslem v názvu) (Alexander et al., 2007). Později přibyly G_A91/G_B91, G_A95/G_B95 a dokonce G_A98/G_B98, které se strukturně odlišují pouze v jediné aminokyselině, přesto se mají jinou konformaci (3α , resp. $\alpha/4\beta$) (viz Tabulka 5) (Alexander et al., 2009).

Tabulka 5: Sekvence domén proteinu G a od nich odvozených chameleonských proteinů. Shodné aminokyseliny tučně, zahrnuty obě varianty G_B88. Stability stanoveny v 0,1 M draselném pufru (pH 7.2). Převzato z Alexander et al. 2007 a Alexander et al. 2009.

Protein	Sekvence	Stabilita [kcal/mol]
G _A	MEAVDANSLA QAKEEAIKEL KQYGIGDYI KLINNAKTVE GVESLKNEIL KALPTE	6
G _B	MTYKLILNKG TLKGETTTEA VDAATAEKVF KQYANDNGVD GEWTYDDATK TFTVTE	7
G _A 30	MEAVDANSLA QAKEEAIKEL KQYGIGEKYI KLINNAKTVE GVWSLKNEIL KALPTE	6
G _B 30	MTYKLILNKG TLKGETTTEA VDAATAEKYF KLYANDKTVE GEWTYDDATK TFTVTE	4,5
G _A 77	TTYKLILNLK QAKEEAIKEL VDAGIAEKYI KLIANAKTVE GVWTLKDEIL TFTVTE	5
G _B 77	TTYKLILNKG QLKEEAITEA VDAATAEKYF KLYANAKTVE GVWTKDETK TFTVTE	4
G _A 88	TTYKLILNLK QAKEEAIKEL VDAGIAEKYI KLIANAKTVE GVWTLKDEIL TFTVTE	4
G _B 88	TTYKLILNLK QAKEEAIKEL VDAATAEKYF KLYANAKTVE GVWTKDETK TFTVTE	2
G _B 88b	TTYKLILNLK QAKEEAITEA VDAGIAEKYF KLYANAKTVE GVWTKDEIK TFTVTE	~4
G _A 88	TTYKLILNLK QAKEEAIKEL VDAGIAEKYI KLIANAKTVE GVWTLKDEIL TFTVTE	4
G _A 91	TTYKLILNLK QAKEEAIKEL VDAGIAEKYI KLIANAKTVE GVWTLKDEIL TFTVTE	~3
G _B 91	TTYKLILNLK QAKEEAIKEA VDAGIAEKYF KLIANAKTVE GVWTKDEIK TFTVTE	~3
G _A 95	TTYKLILNLK QAKEEAIKEL VDAGIAEKYI KLIANAKTVE GVWTLKDEIK TFTVTE	~3
G _B 95	TTYKLILNLK QAKEEAIKEA VDAGIAEKYF KLIANAKTVE GVWTKDEIK TFTVTE	~3
G _A 98	TTYKLILNLK QAKEEAIKEL VDAGIAEKYF KLIANAKTVE GVWTLKDEIK TFTVTE	Nízká
G _B 98	TTYKLILNLK QAKEEAIKEL VDAGIAEKYF KLIANAKTVE GVWTKDEIK TFTVTE	Nízká

Vysvětlení tohoto jevu není triviální. Mechanismus překladu primární struktury na terciární patří mezi nejvýznamnější problémy moderní molekulární biologie, a to i přes rozmach využití strojového učení pro predikce a simulace (Noé et al., 2020; Jumper et al., 2021; Mirdita et al., 2022; Baek et al., 2021; AlQuraishi, 2019). Důležitým faktorem pro zachování konformace předlohy domény může být dle některých autorů zachování stejného hydrofobního jádra proteinu (Banach et al., 2016; Fabian et al., 2020), či zachování preference sekundární struktury C-terminálních nebo N-terminálních aminokyselin (Alexander et al., 2009; Saravanan et al., 2010) nebo určitých velmi stabilních strukturních komponent, např. u G_B skupiny druhé β-vlášeny (z angl. hairpin) a dlouhého α-helixu, které zajišťují ustanovení specifických stabilních intermediátů (Sikosek et al., 2016; Morrone et al., 2011).

Dynamika přeměny jedné konformace v jinou není však pouze problémem arteficiálních domén navržených *de novo* kvůli vědecké výzvě; existuje řada příkladů přírodních (tzv. metamorfických) proteinů, kde byly podobné tranzice pozorovány, některé jsou dokonce strukturně zcela identické (tj. jeden protein zaujímá dvě nativní konformace bez interakce s ligandem) (Bryan a Orban, 2010; Murzin, 2008). Např. chemokin třídy C lymfotaktin (Ltn) se ve fyziologických podmínkách (*in vitro* i *in vivo*) vyskytuje v rovnováze dvou stavů označovaných jako Ltn10 a Ltn40 v poměru cca 1:1, které mají odlišnou konformaci (α/3β monomer resp. 4β dimer) a funkčně se doplňují (Tuinstra et al., 2008; Volkman et al., 2009). Do této skupiny lze řadit také priony, infekční bílkovinné částice, které podléhají konformační přeměně z majoritně α-helikální struktury na konformaci

s větším zastoupením β -skládaných listů, která aktivuje jejich patogenní potenciál (Harrison et al., 1999).

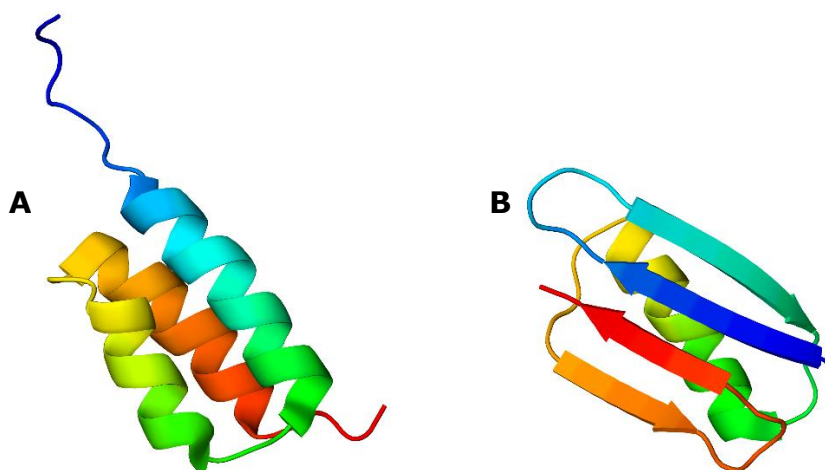
Špatné skládání proteinů (a následná agregace) je rovněž faktorem celé řady neurodegenerativních proteinopatií, jakými jsou např. Alzheimerova a Parkinsonova (zapojené proteiny jsou někdy označovány jako prionoidy), a také některých onemocněních zapříčiněných dědičnými polymorfismy, které mohou vyústit ve ztrátu funkce postiženého proteinu (např. Gaucherova choroba), nebo jeho zisku role patologické (např. Huntingtonova choroba) (Sahni et al., 2015; Hartl, 2017; Cohen a Kelly, 2003; Scheckel a Aguzzi, 2018). Porozumění mechanismu rozdílného skládání strukturně velmi blízkých proteinů by tak samozřejmě představovalo výhodný výchozí bod pro vývoj nových terapeutik, např. na bázi farmakologických chaperonů (Tripathi et al., 2021; Bernier et al., 2004; Gámez et al., 2018).

1.2.2 Charakteristika G_A88

G_A88 (Obr. 7A) je původně 56 aminokyselin dlouhá chameleonská sekvence odvozená od G_A domény proteinu G s konformací 3 α . Má 88% sekvenční identitu s G_B88 (Obr. 7B) s konformací α /4 β odvozené od domény G_B (viz Tabulka 6), jak již bylo zmíněno. Kromě konformace přejímají obě tyto sekvence i vazebnou specifitu domén, od kterých jsou odvozeny – G_A88 tak interaguje s HSA, u G_B88 je zachována interakce IgG (Alexander et al., 2009).

Tabulka 6: Porovnání sekvencí G_A88 a G_B88 (obě publikované varianty). Rezidua shodná s G_A vyznačena zeleně, s G_B červeně, s oběma tmavě žlutě, rezidua zcela zaměněná od výchozích struktur černě. Nad sekvencemi naznačena sekundární struktura. Převzato z Alexander et al. 2007 a 2009.

		~~~~~ $\alpha$ 1~~~~~	~~~~~ $\alpha$ 2~~~~~	~~~~~ $\alpha$ 3~~~~~	
G _A 88	<b>T</b>	<b>TYKLILNLKQAK</b>	<b>EEAIKELVDAGIAEKYIKLI</b>	<b>ANAKTVEGVWTLKDEIL</b>	<b>TFTVTE</b>
		_____ $\beta$ 1_____	_____ $\beta$ 2_____	~~~~~ $\alpha$ 1~~~~~	_____ $\beta$ 3_____
G _B 88	<b>T</b>	<b>TYKLILNLKQAK</b>	<b>EEAIKELVDAATAEKYFKLY</b>	<b>ANAKTVEGVW</b>	<b>TYKDETKTFTVTE</b>
G _B 88b	<b>T</b>	<b>TYKLILNLKQAK</b>	<b>EEAITEAVDAGTAEKYFKLY</b>	<b>ANAKTVEGVW</b>	<b>TYKDEIKTFTVTE</b>



Obr. 7: Struktury chameleonských proteinů G_A88 a G_B88 dle He et al., 2008. A: G_A88 (PDB ID: 2JWS), B: G_B88 (PDB ID: 2JWU).

V rámci této práce byla použita modifikovaná varianta G_A88 domény obsahující histidinovou kotvu, štěpné místo pro TEV proteasu a krátkou vektorovou sekvencí (viz Tabulka 7). Charakterizace biofyzikálními metodami byly prováděny po odstranění histidinové kotvy.

Tabulka 7: Sekvence a charakteristika  $G_{A88}$  použité v experimentální části této práce. Molekulová hmotnost byla predikována online nástrojem ProtParam (Gasteiger et al., 2005).  $G_{A88}$  (rudě), histidinová kotva (tmavě modře), vektorová sekvence (šedivě), štěpná sekvence pro TEV proteasu (oranžově).

Protein		Počet reziduí				$M_w$ [Da]		
<b>6His-$G_{A88}$</b>		<b>80</b>				<b>9 088,37</b>		
5	10	15	20	25	30	35	40	
<b>MGSSH</b>	<b>HHHHH</b>	<b>SSGEN</b>	<b>LYFQG</b>	<b>HMLET</b>	<b>TYKLI</b>	<b>LNLKQ</b>	<b>AKEEA</b>	
45	50	55	60	65	70	75	80	
<b>IKELV</b>	<b>DAGIA</b>	<b>EKYIK</b>	<b>LIANA</b>	<b>KTVEG</b>	<b>VWTLK</b>	<b>DEILT</b>	<b>FTVTE</b>	
<b>$G_{A88}$</b>		<b>61</b>				<b>6 877,05</b>		
5	10	15	20	25	30	35	40	
<b>GHMLE</b>	<b>TTYKL</b>	<b>ILNLK</b>	<b>QAKEE</b>	<b>AIKEL</b>	<b>VDAGI</b>	<b>AEKYI</b>	<b>KLIAN</b>	
45	50	55	60	65				
<b>AKTVE</b>	<b>GVWTL</b>	<b>KDEIL</b>	<b>TFTVT</b>	<b>E</b>				

### 1.3 Fúzní proteiny

Jako fúzní (rovněž chimérické) proteiny jsou označovány proteiny obsahující domény pocházející z několika různých proteinů (Uhlén et al., 1992; Yu et al., 2015). Rovněž se tímto pojmem označuje zcela nepříbuzná skupina virových proteinů zajišťujících vstup virů do buněk (fúzi s buněčnou membránou) (Poehlmann a Simmons, 2013), nicméně tyto zde dále diskutovány nebudou a v rámci této práce bude termín fúzní protein vyhrazen pouze pro první z definic.

Přestože o fúzních proteinech se dnes hovoří především v kontextu syntetických proteinů, existují rovněž přírodní proteinové fúze, tj. např. enzymatické komplexy, které se v nižších organismech (např. prokaryotech) vyskytují v podobě několika proteinů, zatímco ve vyšších organismech (např. eukaryotech) jako jeden protein obsahující všechny odpovídající domény. Příkladem budiž pyrrolin-5-karboxylát synthasa s glutamát kinasovou a glutamyl-fosfát reduktasovou aktivitou (Pérez-Arellano et al., 2010), či acetyl-CoA karboxylasa s biotin karboxylasovou a karboxyl transferasovou aktivitou (Tong, 2005). V obou těchto příkladech vzniká fúze se stejnou funkcí jako jednotlivé komponenty, avšak s vyšší aktivitou, jde tedy o evoluční optimalizaci katalyzovaných procesů. To naznačují také experimenty s tryptofan synthetasou hub (rovněž fúzní protein), jejíž dvě podjednotky jsou řazeny v pořadí zajišťující nejlepší aktivitu i přesto, že v genomu bakterií se vyskytují (jako jednotlivé proteiny) v pořadí opačném (Burns et al., 1990).

Ve vzácnějších případech mohou však fúzí vznikat proteiny se zcela novou funkcí jako je tomu např. u fúze KUA-UEV (Long, 2000), nebo může být funkce ztracena jako u onkogenních proteinů akutní (např. MLL-ENL) a chronické (např. BCR-ABL1) myeloidní leukemie (Rapin a Porse, 2014; Strohl, 2017), popřípadě není funkce pozměněna, ale fúze nevykonává sérii reakcí, která by poskytovala na první pohled zřejmou evoluční výhodu, např. u glutamyl-tRNA/prolyl-tRNA synthetasy, která je jedinou bifunkční aminoacyl-tRNA synthetasou (Berthonneau a Mirande, 2000).

V přírodě jsou fúze proteinů realizovány přirozenými evolučními procesy, a to na úrovni genů. V nižších organismech mohou vznikat fúze např. inserty či bodovými mutacemi v terminačním kodonu, zatímco ve vyšších organismech může být mechanismus jejich vzniku komplexnější, např. chromozomovými přestavbami (z angl. chromosomal rearrangements, např. vzájemnou translokací), či tzv. read-through transkripcí s alternativním sestřihem (také nazýváno intergenový sestřih), kdy jsou z primárního transkriptu vystřiženy části nesoucí terminační kodon prvního z proteinů a iniciační kodon druhého z proteinů – sekvence je tak podrobena translaci jako jeden celistvý protein (Long, 2000; Berthonneau a Mirande, 2000; Strohl, 2017).

### 1.3.1 Design a příprava rekombinantních fúzních proteinů

Obtížným úkolem přípravy syntetických funkčních fúzních proteinů je jejich design paralelizující fungující přírodní fúze, avšak bez času nutného na evoluční optimalizaci. Ukazuje se, že pro zachování řady vlastností (např. rozpustnosti, skládání a funkce) je esenciální jak pořadí, tak i způsob, jakým jsou fúzované domény spojeny, čemuž musí být uzpůsoben proces jejich designu (Sachdev a Chirgwin, 1998; Boušová et al., 2021; Yu et al., 2015).

#### 1.3.1.1 Výběr domén a linkerů

S počtem popsaných přírodních domén dosahujícím téměř dvaceti tisíc (dle databází InterPro a CDD) (Paysan-Lafosse et al., 2022; Wang et al., 2023) je množství kombinací do proteinových fúzí v praxi téměř neomezené – jen pro dvoudoménové fúze se pohybuje ve stovkách milionů. Design fúzních proteinů však zpravidla není v principu náhodný – na základě požadované funkce nebo vlastnosti jsou vybírány vhodné domény, které jsou s to ji zajistit, např. v případě potenciálních terapeutik se často vychází z popsaných biochemických aktérů patologického procesu, na něž je toto vyvíjeno, či jejich interakčních partnerů, a další domény snižující degradabilitu *in vivo* (viz 1.3.2.1).

Ačkoliv v přírodních proteinech může docházet ke změnám v pořadí částí sekvence proteinu bez významného narušení terciální struktury v rámci jevu nazývaného cirkulární permutace (např. nezávislou genetickou fúzí dvou jednodoménových proteinů či sérií duplikací a delecí dvoudoménového proteinu) (Lutz a Bornscheuer, 2009), který je rovněž důležitým nástrojem evoluce proteinů (např. výše zmíněný příklad tryptofan synthetasy hub), u syntetických fúzních proteinů může změna pořadí domén požadované vlastnosti velmi nepříznivě ovlivňovat. Obvykle se tak zpočátku připravují a testují obě varianty fúze (často označované forward a reverse), z nichž je následně vybrána ta preferovanější (Yu et al., 2015; Sachdev a Chirgwin, 1998; Kirubakaran et al., 2016).

Vhodné spojení obou domén je taktéž klíčové. Používají se zpravidla krátké aminokyselinové sekvence (tzv. linkery) se specifickými vlastnostmi; mohou zajišťovat fúzi flexibilitu či naopak rigiditu (podporovat či omezovat volný pohyb domén vůči sobě v prostoru), nebo obsahovat štěpné místo, v němž může být fúze rozdělena. Flexibilní linkery jsou používány pro fúze, kde je třeba uchovat přímou interakci mezi fúzovanými doménami, nebo mezi nimi jednoduše vytvořit fyzickou mezeru. Jsou složeny z malých a hydrofilních aminokyselin, nejčastěji G, S, ale také E, K a dalších (viz *Tabulka 8*). Rigidní linkery lze použít, pokud je cílem zafixování vzdálenosti mezi fúzovanými doménami a zabránění nežádoucím interakcím, které mohou zhoršovat aktivitu či výtěžek fúze. Jedná se buďto o  $\alpha$ -helikální struktury, nejčastěji obsahující sekvence EAAAK, nebo sekvence s vysokým zastoupením P (viz *Tabulka 8*) (Chen et al., 2013; Patel et al., 2022).

V případě, že cílem je vytvoření štěpitelného fúzního proteinu, který se *in vivo* rozpadá na jednotlivé domény, lze použít linkery se specifickými štěpnými sekvencemi proteas (např. CRRRRR*EAEAC štěpitelná Kex2 nebo Ste13 proteasami, či PLG*LWA štěpitelná MMP-1 proteasou), často ve formě cyklopeptidu se snadno redukovatelným disulfidovým můstkem mezi koncovými C linkeru (Chen et al., 2013; Strohl, 2017).

Tabulka 8: Sekvence vybraných často využívaných flexibilních a rigidních linkerů. Příklady z Chen et al., 2013, Kirubakaran et al., 2016 a Strohl, 2017.

	Sekvence	n
Flexibilní	(G) _n	5-8
	GGSG	
	(GGGS) _n	1-4
	KESGSVSSEQLAQFRSLD	
	EGKSSGSGSESKST	
	GSAGSAAGSGEF	
Rigidní	(EAAAK) _n	1-3
	A(EAAAK) _n A	2
	A(EAAAK) _n ALEA(EAAAK) _n A	4
	PAPAP	
	(XP) _n , X = A, E, K ad.	5-17

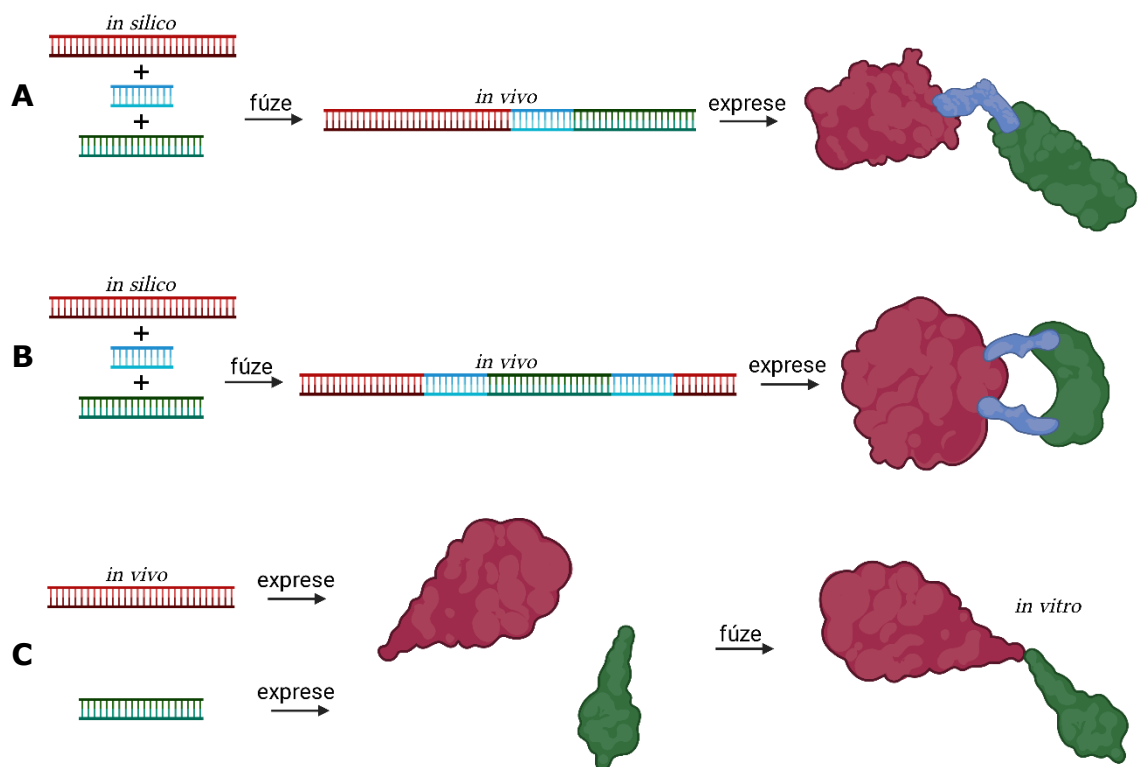
### 1.3.1.2 Metody přípravy proteinových fúzí

Spojení jednotlivých domén do jedné fúze může být provedeno několika způsoby, buďto na úrovni genu nebo na úrovni proteinu. Nejsnadnějším a také nejužívanějším způsobem je tandemová fúze (také end-to-end fúze) (viz Obr. 8A), kdy jsou sekvence DNA kódující obě fúzované domény umístěny v expresním vektoru ihned za sebe (bez terminačního kodonu mezi nimi, se zachováním čtecího rámce) a jsou tak exprimovány jako jedna sekvence. Při použití tohoto postupu se velmi často přistupuje k zavedení linkeru mezi tyto dvě sekvence, zejména pakliže je funkce jedné z domén koncentrována na konci, na kterém je spojena s doménou druhou, a může tak docházet k jejímu negativnímu ovlivnění narušením dynamiky, či pokud není konec dostatečně flexibilní a mohou se objevovat nežádoucí sterické efekty bránící správnému skládání. Zmíněné negativní jevy mohou způsobovat snížení stability exprimovaných fúzí, tvorbu inkluzních tělísek a zhoršení až ztrátu funkce fúzovaných domén. V optimálním případě jsou však funkce i struktura obou domén zachovány (Yu et al., 2015; Rizk et al., 2012; Strohl, 2017).

Složitější alternativou je doménová inserce (viz Obr. 8B), kdy je sekvence jedné z domén (vkládané) vmezeřena do sekvence druhé domény (hostitelské). Analogické děje jsou u přírodních multidoménových proteinů relativně časté, funkce domén tedy nemusí nutně vyžadovat zachování kontinuity jejich sekvence, ale spíše prostorové konformace. Vkládané domény jsou proto zpravidla menší než hostitelské a mají C-konec a N-konec v těsné blízkosti tak, aby jejich zavedení do hostitelské domény významně neovlivnilo její skládání. Výběr místa, v němž bude doména vložena, je rovněž esenciální. Zpravidla se inserce provádí na vnějších otočkách, které se nezapojují do funkce hostitelské domény, a to s použitím dvou linkerů. Touto cestou vzniká rigidnější, stabilnější fúze odolnější vůči proteolýze *in vivo* (Doi a Yanagawa, 1999; Yu et al., 2015).

V případě, že exprese kompletních fúzních proteinů (z fúzí na bázi genu) nemá dostatečný výtěžek nebo nevykazuje kýženou strukturu či funkci z důvodu špatného složení, je nasnadě zvážit třetí, fundamentálně odlišný přístup, a to posttranslační fúzi (viz Obr. 8C). Místo kompletního fúzního proteinu jsou v tomto případě exprimovány fúzované domény

samostatně, příslušné konce jsou však modifikovány tak, aby mohly být následně spojeny (Yu et al., 2015). Fúze je připravena ze syntetizovaných domén *in vitro* enzymatickou ligací (řídčeji méně specifickou chemickou ligací, např. Huisgenovou azid-alkyn [3+2] cykloadicí, tzv. CuAAC) (Hatzakis et al., 2006; Xiao a Tolbert, 2009). Za tímto účelem lze použít specifické aminokyselinové sekvence na koncích obou domén, buďto za účelem vytvoření tradičních spojů peptidovou vazbou, např. C-terminální LPETGX_n motiv a N-terminální GGG kotva při ligaci sortasou A (Levary et al., 2011), nebo méně tradičních spojení jinými typy vazeb, např. dvě GGGY kotvy při párování (z angl. coupling) tyrosinů křenovou peroxidasou za vzniku dityrosinového můstku (Minamihata et al., 2011). Tento přístup je extrémně užitečný zejména pro multidoménové fúze (Yu et al., 2015).



Obr. 8: Metody přípravy fúzních proteinů. A: Tandemová fúze. B: Doménová inserce. C: Posttranslační fúze. Domény a jejich příslušné kódující sekvence červeně a zeleně, linker modře. Zpracováno pomocí online nástroje BioRender ([biorender.com](http://biorender.com)).

### 1.3.2 Praktická využití rekombinantních fúzních proteinů

Prvním využitím toho, co dnes bývá označováno jako proteinové fúze, byly afinitní kotvy. Spojením kýženeho proteinu s jinou doménou (kotvou) lze zlepšit jeho expresi, skládání a rozpustnost a zvýšit jeho rezistenci vůči proteolýze, přičemž objemné kotvy (např. maltosu vázající protein – MBP, thioredoxin – TRX, či glutathion-S-transferasa – GST) v těchto ohledech zpravidla fungují lépe než malé. Zároveň, jak název napovídá, pomocí nich lze kýžený protein oddělit od jiných pomocí afinitní chromatografie. Postupně však využití fúzních proteinů přerostla z nástroje pro purifikaci k širokým uplatněním v analytické biochemii, proteinové terapii i chemickém průmyslu (Smyth et al., 2003; Yu et al., 2015).

#### 1.3.2.1 Terapeutika

Terapeutik na bázi fúzních proteinů existuje již dnes na trhu celá řada a mnohé další jsou stále předmětem vývoje. Obecně je lze rozdělit do tří hlavních skupin:

- Fúze farmakologicky aktivního peptidu/proteinu s proteinem s vysokým biologickým poločasem – Nejčastěji s Fc fragmentem lidského imunoglobulinu G (huIgG1-Fc), či lidským sérovým albuminem (HSA) výsledkem je prodloužení doby působení (na 4 dny až 4 týdny) díky recyklačnímu mechanismu zprostředkovaného neonatálním Fc receptorem (FcRn) a obecné zlepšení farmakokinetických vlastností. Např. etanercept (Enbrel), Eftrenonacog alfa (Alprolix), rFIX-FP (Idelvion);
- Fúze farmakologicky či toxikologicky aktivního peptidu/proteinu s doménou cílící do specifických tkání/buněk – Využívané pro cílení (z angl. targeting) terapeutika např. cytokinů (často IL-2) či buněčných jedů do postižené tkáně, často protinádorová léčiva. Např. denileukin diktilox (ONTAK);
- Fúze několika domén s vlastní farmakologickou aktivitou – Tvoří se multispecifický protein, nejčastěji fúze monoklonálních protilátek. Např. blinatumomab (Blinicyto) (Strohl, 2017; Schmidt, 2013; Silver et al., 2021; Kontermann, 2012).

Speciálním případem jsou pak vakcíny na bázi proteinových fúzí. Pracuje se s fúzemi řady antigenů (odvozených od HIV-1, *Bacillus anthracis*, *Vibrio cholerae*, enterotoxigenní *E. coli* ad.) s Fc fragmenty různých imunoglobulinů (např. IgA, IgE, IgG1, IgG3, IgM), netoxickými imunogenními doménami bakteriálních exotoxinů (např. CTB, LTB), či virům podobnými částicemi (VLPs), které působí jako adjuvans. Tato nevšední technologie, která je zatím však relativně v plenkách, má potenciál zajistit vakcíny proti nemocem, pro něž zatím prevence touto cestou nebyla možná (Czajkowsky et al., 2012; Soria-Guerra et al., 2011; Moyle, 2017).

#### 1.3.2.2 Biokatalyzátory

Zatímco anorganické katalyzátory jsou v chemické průmyslové produkci naprostým standardem, moderní řešení k provádění některých složitých syntéz bylo nalezeno v použití rekombinantních enzymů. Tento přístup se již dnes s výhodou využívá při industriální výrobě řady látek, např. aspartamu, vanilinu, penicilinu G, islatraviru, nootkatonu či akrylamidu, přičemž mezi nejvyužívanější enzymové rodiny patří alkohol dehydrogenasy,

ketoreduktasy, transaminasy, aldolasy, acylasy, halohydrin dehalogenasy, Baeyer–Villiger monooxygenasy ad. Mezi klíčové výhody biokatalytické syntézy se řadí rychlost, nízká cena, vysoká selektivita, bezpečnost pro člověka i životní prostředí a v neposlední řadě možnost optimalizace cílenou evolucí a dalšími metodami proteinového inženýrství (Bell et al., 2021; Bornscheuer, 2018).

Jednou z nich je i fúze enzymů. Při tandemovém spojení domén katalyzujících po sobě jdoucí reakce lze totiž v některých případech dosáhnout i několikanásobné zvýšení výtěžku díky přiblížení aktivních míst. Např. tvorbou bifunkční aldolasy/kinasy prodlužující řetězec dihydroxyacetonu vytvořením C-C vazby se substrátovým aldehydem se zvýší rychlost reakce 20× (Iturrate et al., 2010).

Ještě lepších výsledků bylo dosaženo s použitím syntetických fúzních scaffoldových proteinů, nikoliv přímého spojení katalytických proteinů (Vanderstraeten a Briers, 2020). Tímto způsobem byl vytvořen např. proteinový komplex katalyzující sérii tří reakcí složený z acetoacetyl-CoA thiolasy (AtoB), hydroxy-methylglutaryl-CoA synthasy (HMGS) a hydroxy-methylglutaryl-CoA reduktasy (HMGR), spojený fúzním scaffoldovým proteinem s GTPasovou protein-vázací doménou (GBD), SH3 doménou a PDZ doménou, který vykazoval 77× vyšší výtěžek než samostatné enzymy v roztoku (Dueber et al., 2009).

### 1.3.2.3 Analytické metody

Proteinová fúze rovněž umožňuje úpravu studovaného proteinu tak, aby mohl být charakterizován. Za tímto účelem se používají zpravidla fluorescenční proteiny jako zelený fluorescenční protein (GFP, GFP2 ad.) a jeho varianty (cyanový – CFP, žlutý – YFP, modrý – BFP) a luciferasy (v přítomnosti svých substrátů – luciferinů). Pomocí nich lze kýžený protein sledovat *in vivo*, což je možnost jinak velmi obtížně dostupná. S využitím takových fúzí lze snadno detekovat expresi a lokalizaci studovaného proteinu v buňkách pomocí fluorescenční mikroskopie, či studovat protein-protein a protein-ligand interakce metodou Försterova (fluorescenčního) rezonančního přenosu energie (FRET) (nejčastěji s využitím fúzí CFP-protein 1 a YFP-protein 2, méně BFP-protein 1 a GFP-protein 2) a bioluminescenčního rezonančního přenosu energie (BRET) (fúze luciferasa-protein 1 a YFP/GFP2-protein 2) (Pollok, 1999; Snapp, 2005; Bacart et al., 2008; Shimosono a Miyawaki, 2008).

Pro evaluaci exprese lze využít i jiné fúzní partnery, které poskytují specifické reakce, jako chloramfenikol acetyltransferasa (CAT), alkalická fosfatasa (AP)  $\beta$ -galaktosidasa nebo  $\beta$ -laktamasa, pro něž existují publikované protokoly a komerčně dostupné testovací sady (Strohl, 2017).



## 2 Cíle práce

V obecné rovině lze za cíl předložené diplomové práce považovat přípravu a biofyzikální charakterizaci dvou fúzních proteinů, FD5A a FD6A, složených z PDZ3 a G_A88 domén, ve smyslu hydrodynamických vlastností a teplotní stability.

Konkrétní cíle práce byly:

- Příprava bakteriálních kultur pro expresi proteinů FD5A, FD6A a samostatné G_A88
- Expres FD5A, FD6A a G_A88 s histidinovou kotvou a jejich purifikace
- Odstranění histidinové kotvy pomocí proteasy a subtraktivní afinitní chromatografie
- Přečištění FD5A, FD6A a G_A88 vylučovací chromatografií
- Charakterizace FD5A, FD6A a G_A88 pomocí analytické vylučovací chromatografie
- Charakterizace FD5A, FD6A a G_A88 pomocí dynamického rozptylu světla
- Charakterizace FD5A, FD6A a G_A88 pomocí analytické ultracentrifugace
- Charakterizace FD5A, FD6A a G_A88 pomocí cirkulárního dichroismu
- Charakterizace FD5A, FD6A a G_A88 pomocí diferenciální skenovací fluorimetrie

## 3 Experimentální část

### 3.1 Vybavení

#### 3.1.1 Přístrojové vybavení a materiál

Analytická ultracentrifuga Optima AUC (Beckman Coulter, Brea, USA)  
rotor AN50Ti

Analytické váhy L303i (Bel Engineering, Monza, Itálie)

Centrifuga Avanti JXN-30 (Beckman Coulter, Brea, USA)  
rotor JLA-8.1000

Centrifuga Multifuge X1R (Thermo Fisher Scientific, Waltham, USA)  
rotor TX-200  
rotor Fiberlite F15-8x50cy

Denzitometr Biospectrometer Kinetic (Eppendorf, Hamburg, Německo)

Dialyzační membrána Spectra/Por 1 Standard RC Tubing  
(Thermo Fisher Scientific, Waltham, USA)

DLS analyzátor Zetasizer Nano ZS (Malvern Panalytical, Malvern, UK)

Elektroforetický systém Mini-PROTEAN Tetra System (BioRad, Hercules, USA)

Elektroforetický zdroj PowerPac Basic (Bio-Rad, Hercules, USA)

Filtry Whatman ME25 (Cytiva, Chicago, USA)

Filtry Whatman 13mm SPARTAN (Cytiva, Chicago, USA)

Flow-box Secuflow (Waldner, Wangen, Německo)

FPLC ÄKTA Pure Protein Purification System (Cytiva, Chicago, USA)  
kolona Superdex 75 Increase 10/300 GL  
Fraction Collector F9-R

Hlubokomrazící box VIP Plus MDF U76V (Panasonic, Osaka, Japonsko)

Hmotnostní spektrometr Finnigan MALDI/TOF  
(Thermo Fisher Scientific, Waltham, USA)

Inkubátor třepací Incubator Shaker ES-60 (BTLab, St. Louis, USA)

Inkubátor třepací Innova S44i (Eppendorf, Hamburg, Německo)

Koncentrátor Amicon Ultra – 15 Ultracel – 10K (Merck, Rahway, USA)

Magnetická míchačka ARE Hot Plate Stirrer (VELP Scientifica, Usmate, Itálie)

Magnetická míchačka ESP Ultraflat Magnetic Stirrer  
(VELP Scientifica, Usmate, Itálie)

Mikrocentrifuga Micro Star 17 (VWR, Radnor, USA)

Millipore Milli-Q Synthesis (Merck, Rahway, USA)  
cartridge Q-Gard 1

NanoPhotometer NP80 (Implen, Mnichov, Německo)

PCR cyklér T100 Thermal Cycler (Bio-Rad, Hercules, USA)

Peristaltická pumpa Ecoline Vc-MS/CA8-6 (Ismatec, Glattbrugg, Švýcarsko)

pH metr Edge Blu (Hanna Instruments, Smithfield, USA)

qPCR cyklér qTOWER3/G (Analytik Jena, Jena, Německo)

Skener CanoScan LiDE 220 (Canon, Ota, Japonsko)

Sonikátor UP200St (Hielscher, Teltow, Německo)

Spektrometr cirkulárního dichroismu J-815 CD (JASCO, Tokyo, Japonsko)

Termoblok TDB-120 (Biosan, Riga, Litva)

Vodní lázeň SWB6D Digital Water Bath (Cole-Parmer, Stone, UK)

Vortex Genie 2T (Scientific Industries, Bohemia, USA)

### 3.1.2 Software

CDPro (Sreerama a Woody, 2000)

CollabFold (Mirdita et al., 2022)

DSFworld (Wu et al., 2020)

GUSSI, verze 1.4.2 (Brautigam, 2015)

HullRad, verze 9 (online) (Fleming a Fleming, 2018)

HYDROPRO, verze 10 (kompilace 54) (García De La Torre et al., 2000; Ortega et al., 2011)

NEB Tm calculator ([tmcalculator.neb.com](http://tmcalculator.neb.com))

PDBMD2CD (online) (Drew a Janes, 2020)

PyMOL Molecular Graphics System, Verze 2.4.1, Schrödinger, LLC

Sedfit, verze 15.01b (Schuck, 2000; Brown a Schuck, 2008)

Sednterp, verze 20130813 BETA (Laue et al., 1992)

UltraScan III, verze 4.0 (6731) (Demeler, 2005)

## 3.2 Chemický a biologický materiál

### 3.2.1 Chemikálie

2-merkptoethanol (Carl Roth, Karlsruhe, Německo)  
Ampicilin (AMP) (Carl Roth, Karlsruhe, Německo)  
Bromfenolová modř (BPB) (Thermo Fisher Scientific)  
Coomassie Brilliant Blue (CBB) (Carl Roth, Karlsruhe, Německo)  
Dihydrogenfosforečnan draselný ( $\text{KH}_2\text{PO}_4$ ) (PENTA, Praha, ČR)  
Dimethylsulfoxid (DMSO) (Sigma-Aldrich, St. Louis, USA)  
Dithiotreitol (DTT) (Sigma-Aldrich, St. Louis, USA)  
Dodecylsírán sodný (SDS) (Carl Roth, Karlsruhe, Německo)  
fenylmethylsulfonyl fluorid (PMSF) (Carl Roth, Karlsruhe, Německo)  
Glycerol (Sigma-Aldrich, St. Louis, USA)  
Glycin (Lach-Ner, Neratovice, ČR)  
Guanidin hydrochlorid (Gm-HCl) (Carl Roth, Karlsruhe, Německo)  
Hydrogenfosforečnan sodný dihydrát ( $\text{Na}_2\text{HPO}_4 \cdot \text{H}_2\text{O}$ ) (PENTA, Praha, ČR)  
Hydroxid sodný (NaOH) (PENTA, Praha, ČR)  
Chlorid draselný (KCl) (Lach-Ner, Neratovice, ČR)  
Chlorid sodný (NaCl) (Lach-Ner, Neratovice, ČR)  
Imidazol (Carl Roth, Karlsruhe, Německo)  
Isopropanol (IPA) (Lach-Ner, Neratovice, ČR)  
Isopropyl  $\beta$ -D-1-thiogalaktopyranosid (IPTG) (BioWorld, Irving, USA)  
Kyselina boritá (Carl Roth, Karlsruhe, Německo)  
kyselina ethylendiamintetraoctová (EDTA) (PENTA, Praha, ČR)  
Kyselina chlorovodíková (HCl) (PENTA, Praha, ČR)  
Močovina (Lach-Ner, Neratovice, ČR)  
NP-40 (Merck, Rahway, USA)  
Peroxodisírán amonný (APS) (Sigma-Aldrich, St. Louis, USA)  
SYPRO Orange (Sigma-Aldrich, St. Louis, USA)  
Tetramethylethylendiamin (TEMED) (Bio-Rad, Hercules, USA)  
tris(hydroxymethyl)aminomethan (Tris) (Carl Roth, Karlsruhe, Německo)  
tris(hydroxymethyl)aminomethan hydrochlorid (Tris-HCl)  
(Carl Roth, Karlsruhe, Německo)  
Tween 20 (Carl Roth, Karlsruhe, Německo)

### 3.2.2 Biologický materiál

LB médium dle Millera (dehydrované) (Sigma-Aldrich, St. Louis, USA)  
Lysozym (Thermo Fisher Scientific, Waltham, USA)  
PBM JAM-A (PSPSPSTSSFLV) (GenScript, Piscataway, USA)  
Petriho misky s agarem s AMP  
(připraveno na ÚOCHB AV ČR v rámci servisních laboratoří)  
Plazmidy pET15b pro expresi FD5A (pET15b_FD5A), resp. FD6A (pET15b_FD6A)  
s histidinovými kotvami (GenScript, Piscataway, USA)  
Fw: GTTACCGAGTAATAACCGGCGCCGGCGCCGGGTAGCCACATGATTCTG  
Rv: GCCGGCGCCGGTTATTACTCGGTAACGGTAAAGGTCAGGATTTTCGTCCTTCAGGG

Primery pro PCR (Eurofins, Lucemburk, Lucembursko)  
Protein PDZ3 (připraveno na ÚOCHB AV ČR skupinou Bioinformatika)  
Q5 Hot Start High-Fidelity DNA polymerasa (New England Biolabs, Ipswich, USA)  
TEV proteasa s His-tagem (připraveno na ÚOCHB AV ČR skupinou Bioinformatika)  
Zásobní kultury kompetentní *Escherichia coli* BL21(DE3) (OD ≈ 0,3)

### 3.2.3 Komerční směsi, kity a standardy

dNTPs (New England Biolabs, Ipswich, USA)  
GC enhancer (5×) (New England Biolabs, Ipswich, USA)  
Chelatační sepharosa Chelating Sepharose Fast Flow (Cytiva, Chicago, USA)  
Kit pro purifikaci plazmidové DNA High-Speed Plasmid Mini Kit  
(Geneaid, Nová Tchaj-pej, Taiwan)  
Marker Precision Plus Protein Dual Xtra Standard (Bio-Rad, Hercules, USA)  
Marker proteinových velikostí Precision Plus Protein Standard  
(Bio-Rad, Hercules, USA)  
Q5 pufr (5×) (New England Biolabs, Ipswich, USA)  
Směs akrylamidu a bisakrylamidu (29:1) 30 % ROTIPHORESE NF  
(Carl Roth, Karlsruhe, Německo)

### 3.2.4 Pufry, roztoky a média

Barvicí roztok – CBB (0,007 %), HCl (0,1 %)  
Denaturační pufr – Tris-HCl (50 mM), NaCl (500 mM), Gm-HCl (6 M)  
Elektroforetický pufr (10×) – Tris (250 mM), glycin (1,92 M), SDS (1 %)  
Eluční pufr – Tris-HCl (50 mM), NaCl (0,5 M), imidazol (600 mM), NP-40 (0,01 %),  
pH 7,4  
Lyzační pufr – PBS (1×), NaCl (500 mM), imidazol (20 mM), NP-40 (0,05 %), pH 7,5  
PB Pufr – NaCl (137 mM), KCl (2,7 mM), Na₂HPO₄ · H₂O (10 mM), KH₂PO₄ (1,4 mM),  
pH 7,5  
PBS pufr 10× – NaCl (1,37 M), KCl (27 mM), Na₂HPO₄ (100 mM), KH₂PO₄ (18 mM),  
pH 7,4  
Promývací pufr – Pufr E + Pufr F (9:1), NP-40 (0,01 %), pH 7,5  
Pufr E – PBS (1×), NaCl (500 mM), imidazol (1 mM), pH 7,4  
Pufr F – PBS (1×), NaCl (500 mM), imidazol (600 mM), pH 7,5  
Pufr pro separační gel (4×) – Tris-HCl (1,5 M), SDS (0,4 %), pH 8,8  
Pufr pro zaostřovací gel (4×) – Tris-HCl (500 mM), SDS (0,4 %), pH 6,8  
TEV pufr 1 – Tris-HCl (50 mM), NaCl (500 mM), glycin (100 mM), Tween 20  
(0,005 %), DTT (1 mM), EDTA (0,5 mM), PMSF (0,1 mM), pH 7,5  
TEV pufr 2 – Tris-HCl (50 mM), NaCl (500 mM), glycin (100 mM), Tween 20  
(0,005 %), pH 7,5  
Vyzazovací roztok – NaCl (500 mM), EDTA (100 mM)  
Vzorkový pufr pro SDS-PAGE (5×) – Tris-HCl (60 mM), glycerol (25 %), SDS (2 %),  
2-merkapt ethanol (5 %), BPB (0,1 %), pH 6,8

### 3.2.5 Příprava polyakrylamidových gelů

Bylo smíseno příslušné množství komponent (viz *Tabulka 9*), které tvořily polymerační směs separačního gelu. Ta byla rychle vpravena mezi skla umístěná v sestavené aparatuře pro plnění gelů, kde polymerovala cca 30 min pod vrstvou IPA. Ten byl po skončení reakce odstraněn, mezi skla byla přidána polymerační směs zaostřovacího gelu (viz *Tabulka 9*) a do ní byl zasunut hřebínek tvořící jamky gelu. Gel byl znovu ponechán k zatuhnutí a následně skladován zabalen v buničině napuštěné elektroforetickým pufrům (4 °C).

*Tabulka 9: Objemy jednotlivých komponent pro přípravu gelů pro jeden 1mm SDS-PAGE. Míseno v tomto pořadí.*

	Separací gel (15%)	Zaostřovací gel
Akrylamid/bis-akrylamid (30%)	2,45 ml	0,35 ml
Separací pufr (4×)	1,3 ml	-
Zaostřovací pufr (4×)	-	0,65 ml
Destilovaná voda	1,25 ml	1,45 ml
10% APS	49 µl	29 µl
TEMED	5,5 µl	3,25 µl

### 3.3 Metodika

#### 3.3.1 Predikce struktur proteinů FD5A a FD6A *in silico*

Sekvence FD5A a FD6A byly vloženy do online nástroje ColabFold (Mirdita et al., 2022). Predikce byla provedena metodou MSA párování s nastavením templátového módu pdb70 s číslem relaxace 5 pod vedením Mgr. Josefa Šulce ze skupiny Bioinformatika ÚOCHB AV ČR. Struktury byly vizualizovány v softwaru PyMOL Molecular Graphics System, Verze 2.4.1, Schrödinger, LLC.

#### 3.3.2 Příprava bakteriálních kultur a exprese studovaných proteinů

##### 3.3.2.1 Transformace *E. coli*

Kompetentní buňky *E. coli* BL21 a *E. coli* TOP10 (100  $\mu$ l, OD  $\approx$  0,3) byly rozmrazeny (4  $^{\circ}$ C) a k nim byly přidány zakoupené plazmidy pET15b_FD5A, resp. pET15b_FD6A (GenScript, Piscataway, USA). Tato směs byla inkubována (25 min, 4  $^{\circ}$ C) a následně vystavena tepelnému šoku (60 s, 42  $^{\circ}$ C) a znovu zchlazena (2 min, 4  $^{\circ}$ C). Bylo přidáno LB médium (900  $\mu$ l), kultura byla inkubována na třepačce (1 h, 37  $^{\circ}$ C, 160 RPM) a následně přenesena (200  $\mu$ l) na Petriho misku s agarem s AMP, rozetřena a dále inkubována (37  $^{\circ}$ C). Transformace byly provedeny ve spolupráci s týmem bioinformatické laboratoře ÚOCHB.

##### 3.3.2.2 Mutageneze pET15b_FD6A pomocí PCR

Kolonií buněk *E. coli* TOP10 s plazmidem pET15b_FD6A bylo z Petriho misky inokulováno LB médium (5 ml) s AMP (5  $\mu$ l 30 mg/ml roztoku) a kultura byla přes noc (O/N) ponechána na třepačce (190 RPM, 37  $^{\circ}$ C). Z těchto buněk byla zpětně izolována plazmidová DNA (pET15b_FD6A) pomocí komerčního kitu High-Speed Plasmid Mini Kit (Geneaid, Nová Tchai-pej, Taiwan) dle postupu předepsaného výrobcem (Geneaid Biotech Ltd., 2017, viz *Příloha I*) včetně volitelného kroku. Koncentrace získaného plazmidu byla stanovena spektrometricky z  $A_{260}$  na zařízení NanoPhotometer (Implen, Mnichov, Německo) v módu pro analýzu DNA. Dále byly připraveny shodné směsi (á 50  $\mu$ l, složení viz *Tabulka 10*) pro polymerázovou řetězovou reakci (PCR), s nimiž byla metoda provedena s různými teplotami hybridizace pro různé směsi (nastavení metody viz *Tabulka 11*) pomocí PCR cykléru T100 Thermal Cycler (Bio-Rad, Hercules, USA).

*Tabulka 10: Složení PCR směsí použitých pro mutagenezi pET15b_FD6A.*

pET15b_FD6A	10 ng
Fw primer	0,5 $\mu$ M
Rv primer	0,5 $\mu$ M
Q5 polymerasa	1 U
Q5 pufr	1 $\times$
GC Enhancer	1 $\times$

Tabulka 11: Nastavení metody PCR. Popsány jsou jednotlivé fáze PCR a korespondující teploty a časy. Hybridizace byla provedena v gradientu teplot, pro každou směs při jiné teplotě.

Fáze	Teplota	Čas
Počáteční denaturace	98 °C	30 s
Denaturace	98 °C	10 s
Hybridizace	62-72 °C	20 s
Extenze	72 °C	240 s
Finální extenze	72 °C	300 s
Po skončení	4 °C	∞

Po proběhnutí PCR byla k reakčním směsím přidána *DpnI* restriktasa (á 10 U) (Thermo Fisher Scientific, Waltham, USA) a směsi byly inkubovány (120 min, 37 °C). Výslednými směsmi byly transformovány a inkubovány kompetentní buňky *E. coli* BL21 analogicky k 3.3.2.1. Plazmidy (pET15b_GA88) byly znovu, z obdobně připravených kultur v LB a za pomoci kitu High-Speed Plasmid Mini Kit (Geneaid, Nová Tchaj-pej, Taiwan), izolovány a jejich sekvence byla ověřena společností Eurofins (Eurofins, Lucemburk, Lucembursko).

### 3.3.2.3 Expres proteinů v *E. coli*

Kolonií buněk *E. coli* BL21 s plazmidem (pET15b_FD5A, pET15b_FD6A, resp. pET15b_GA88) pro daný studovaný protein s histidinovou kotvou (6His-FD5A, 6His-FD6A, resp. 6His-GA88) z Petriho misky bylo inokulováno LB médium (8×5 ml) s AMP (8×5 µl 30 mg/ml roztoku) a kultura byla inkubována na třepačce (190 RPM, 37 °C, O/N). Kultury byly sloučeny, homogenizovány pomocí vortexu a následně po 4,5 ml distribuovány do 8×1 l LB média s AMP (8×1 ml 30 mg/ml roztoku) v Erlenmeyerových baňkách. Tyto kultury byly umístěny do třepačky, inkubovány (150 RPM, 37 °C) a průběžně podrobovány denzitometrickým stanovením. Po dosažení  $OD_{600} \approx 0,6$  byly podmínky zmírněny (22 °C, 77 RPM). Po dosažení  $OD_{600} \approx 1$  bylo do média přidáno IPTG (8×1 ml 0,5 mM roztoku) a směs byla ponechána na třepačce (150 RPM, 22 °C, O/N).

Kultury byly centrifugovány (3500 RPM/1150×g, 4 °C, 20 min), sloučeny a koncentrovány, až z nich byly získány 2 pelety buněk (každá ze 4 l kultury). Pelety byly resuspendovány v lyzačním pufru (320 ml), sloučeny, homogenizovány a rozděleny do dvou zamrazovacích lahví (2×160 ml). Nakonec byl přidán lysozym (cca 2×1,5 mg) a buňky byly uskladněny (-20 °C).



### 3.3.3 Purifikace exprimovaných proteinů

#### 3.3.3.1 Příprava lyzátu *E. coli*

Roztok buněk v lyzačním pufru (160 ml) z kroku 3.3.2.3 byl rozmrazen (4 °C, O/N), následně k němu bylo přidáno PMSF (160 µl 100 µM roztoku). Celá směs byla v ledové lázni sonikována pomocí sonikátoru UP200St (Hielscher, Teltow, Německo) do dosažení viskozity podobné vodě (15-20 min v cyklech 5 s pulz, 10 s prodleva). Z lyzátu byl odebrán vzorek pro elektroforézu, který byl samostatně centrifugován (příprava vzorků peleta – P, supernatant – SN) a zbytek lyzátu byl následně rovněž centrifugován (13 500 RPM/21 200×g, 4 °C, 1 h) a supernatant obsahující protein byl oddělen od pelety.

#### 3.3.3.2 Afinitní chromatografie s imobilizovanými ionty kovů (IMAC)

Kolona pro IMAC byla naplněna chelatační sepharosou Chelating Sepharose Fast Flow (5 ml) (Cytiva, Chicago, USA), která byla promyta vodou (20 ml), vyvazovacím roztokem (30 ml) a znovu vodou (20 ml). Následným promytím síranem nikelnatým (60 ml 0,1 M roztoku) byly navázány nikelnaté ionty (vznik matrice Ni/IDA); průtok byl facilitován pumpou. Kolona byla promyta vodou (20 ml), pak kondicionována puftrem E (20 ml + 10 µl 20% NP-40), a následně byl postupně nanášen supernatant z kroku 3.3.3.1. Nezachycený, prošedší roztok (FT) byl sbírán. Po nanesení veškerého supernatantu byla kolona promývána promývacím puftrem (400 ml), přičemž prvních 10 ml bylo sbíráno (W). Pumpa byla odpojena a byla zahájena gravitační eluce zachyceného proteinu z afinitní matrice pomocí elučního pufru (25 ml). Bylo sesbíráno 20 frakcí á 1 ml (E1-E20), ty byly dočasně uchovávány na ledu.

#### 3.3.3.3 Vizualizace průběhu IMAC pomocí SDS-PAGE

Byly připraveny 1mm 15% SDS-polyakrylamidové gely (složení viz 3.2.5). 20 µl každého ze vzorků P, SN, FT, W, E1-E6 a E8 bylo smíšeno s 5 µl vzorkového pufru a tepelně denaturováno (90 °C, 10 min). Gel byl vložen do lázně s elektroforetickým puftrem v elektroforetické vaně, na gel byl do jamek nanesen marker molekulových hmotností Precision Plus Protein Standard nebo Precision Plus Protein Dual Xtra Standard (5 µl) (Bio-Rad, Hercules, USA) a postupně také všechny vzorky (á 18 µl) a elektroforéza byla spuštěna (200 V, cca 60 min). Gel byl následně povařen ve vodě, obarven pomocí barvicího roztoku a odbarven v horké vodě.

#### 3.3.3.4 Denaturace, odštěpení histidinové kotvy TEV proteasou, renaturace

Eluční frakce byly sloučeny, rozředěny denaturačním puftrem (5×, v případě FD6A část 10×), přeneseny do dialyzační membrány Spectra/Por 1 (MWCO 6-8 kDa) (Thermo Fisher Scientific, Waltham, USA) a denaturovány (4 °C, O/N) v denaturačním pufru (1 l) za pomalého míchání. Následně byla membrána přenesena do TEV pufru 1 (2 l) a dialyzována (4 °C, 4 h) za rychlého míchání. Po každé půlhodině byla membrána kontrolována na vznik agregátů a promíchána.

Po přenesení proteinu do TEV pufru 1 byl odebrán vzorek na gel (preTEV), přidána TEV proteasa s histidinovou kotvou (6His-TEV) (250 U na 1 mg proteinu) (vlastní výroba laboratoře) a směs byla inkubována (30 °C, 90 min). Z této směsi byl odebrán vzorek na gel (postTEV) a poté byla dialyzována (4 °C, O/N) proti TEV pufru 2.

### 3.3.3.5 Subtraktivní IMAC

Kolona byla připravena analogicky k 3.3.3.2 a kondicionována TEV pufrům 2 z dialýzy. Roztok proteinů byl po odebrání vzorku na gel (preSAC) postupně nadávkován a gravitačním průtokem skrz kolonu byla získána celá nezachycená frakce (postSAC). Kolona byla promyta za facilitace pumpou TEV pufrům 2 z dialýzy (50 ml), sbíráno bylo zvláště prvních 10 ml (W1) a zbylých 40 ml (W2).

Frakce postSAC obsahující kýžený protein byla smísena s W1 a předběžně koncentrována pomocí koncentrátoru Amicon Ultra 15 (MWCO 10 kDa) (Merck, Rahway, USA) centrifugací (2500-3500 RPM/850-1650×g, 1-2 min intervaly, pak kontrola agregace a homogenizace roztoku) na koncentraci proteinu cca 1 mg/ml (měřeno spektrometricky z  $A_{280}$  na zařízení NanoPhotometer) (Implen, Mnichov, Německo) v módu pro proteinová stanovení pomocí molárních extenčních koeficientů, viz *Tabulka 12*) a ponechána do dalšího dne (4 °C).

*Tabulka 12: Molární extenční koeficienty studovaných proteinů při 280 nm. Predikce online nástroje ProtParam (Gasteiger et al., 2005).*

Protein	$\epsilon$ [ $M^{-1} \cdot cm^{-1}$ ]
6His-PDZ3	1490
FD5A	8480
FD6A	8480
G _A 88	8480

Celý proces odštěpení histidinové kotvy byl rovněž vizualizován pomocí SDS-PAGE. Tento proces probíhal analogicky k 3.3.3.3, pouze se vzorky nasbíranými během procesu denaturace, inkubace s 6His-TEV proteasou, renaturace a subtraktivní IMAC.

### 3.3.3.6 Vylučovací chromatografie (SEC)

Připravený protein (cca 1 mg/ml, viz 3.3.2.5) byl dále koncentrován do množství 4-10 ml (cca 3-7 mg/ml dle náchylnosti k agregaci) pomocí koncentrátoru Amicon Ultra 15 (MWCO 10 kDa, 2 000-3 500 RPM/850-1650×g, 30-120 s intervaly, pak kontrola agregace a homogenizace roztoku).

Takto získaný protein byl filtrován filtrem Whatman (velikost pórů 0,2  $\mu m$ ) (Cytiva, Chicago, USA) a podroben automatické SEC separaci na koloně Superdex 75

(Cytiva, Chicago, USA) (nástřík cca 2 ml, průtok 0,7 ml/min) na instrumentu pro rychlou proteinovou kapalinovou chromatografii (FPLC) ÄKTA Pure Protein Purification System (Cytiva, Chicago, USA) s UV detektorem (280 nm) a záchytem frakcí á 1 ml, při níž byl zároveň převeden do PB pufru. V jednotlivých frakcích byla spektrofotometricky pomocí NanoPhotometeru (Implen, Mnichov, Německo) stanovena koncentrace proteinu a vhodné frakce byly alikvotovány a zamrazeny (-80 °C).

### 3.3.3.7 SDS-PAGE a MS frakcí z SEC

Z chromatogramu (závislost  $A_{280}$  na prošlém objemu) ze SEC (3.3.2.6) byly vybrány frakce vykazující nejvyšší koncentraci proteinu a ty byly podrobeny SDS-PAGE analogicky k 3.3.3.3. Vybraný proužek odpovídající velikostně studovanému proteinu byl podroben hmotnostně spektrometrické (MS) analýze (in gel digestion), a tak byla ověřena správnost jeho sekvence. MS analýza byla provedena skupinou Hmotnostní spektrometrie ÚOCHB AV ČR pod vedením doc. RNDr. Josefa Cvačky, Ph.D. pomocí MALDI-TOF-MS analýzy (Thermo Fisher Scientific, Waltham, USA).

### 3.3.4 Biofyzikální charakterizace

#### 3.3.4.1 Analytická SEC

FD5A, FD6A byly naředěny na koncentraci 0,5 mg/ml (30  $\mu$ M), G_A88 na koncentraci 0,2 mg/ml (30  $\mu$ M) a následně analyzovány metodou SEC na koloně Superdex 75 (Cytiva, Chicago, USA) (nástržik 0,5 ml, průtok 0,7 ml/min) na FPLC ÅKTA Pure Protein Purification System (Cytiva, Chicago, USA). Pro každý z proteinů byly provedeny 3 analýzy.

#### 3.3.4.2 Dynamický rozptyl světla (DLS)

FD5A, FD6A byly naředěny na koncentrace 1 mg/ml (60  $\mu$ M) a 0,5 mg/ml (30  $\mu$ M), G_A88 bylo naředěno na koncentraci 0,5 mg/ml (70  $\mu$ M). Koncentrace byly určeny spektrometricky na NanoPhotometeru (Implen, Mnichov, Německo). PDZ3 o koncentraci 0,5 mg/ml bylo získáno zředěním poskytnutého zásobního roztoku (0,7 mg/ml) v poměru 7:3 (zásobní roztok : PB pufru). Vzorky byly filtrovány filtry Whatman (velikost pórů 0,2  $\mu$ m) (Cytiva, Chicago, USA) a centrifugovány (13 500 RPM/17 000 $\times$ g, 4  $^{\circ}$ C, 30 min).

Horní vrstva vzorku byla odebrána do křemenné kyvety a změřena pomocí instrumentu Zetasizer Nano (Malvern Panalytical, Malvern, UK) (20 měření á 10 s, 20  $^{\circ}$ C). Použité vzorky byly denaturovány v PB pufru s 8M močovinou (4  $^{\circ}$ C, 72 h), renaturovány v PB pufru (4  $^{\circ}$ C, 4 h), upraveny na koncentraci cca 0,5 mg/ml, filtrovány a centrifugovány a znovu podrobeny stejnému měření.

#### 3.3.4.3 Analytická ultracentrifugace (AUC)

Byly připraveny vzorky FD5A a FD6A o  $A_{280} = 0,8$  při délce kyvety 1,2 cm (cca 1,2 mg/ml, 75  $\mu$ M). Vzorek G_A88 byl připraven o  $A_{280} = 1$  při délce kyvety 1,2 cm (cca 0,6 mg/ml, 80  $\mu$ M). Vzorek 6His-PDZ3 byl připraven  $A_{280} = 0,65$  při délce kyvety 1,2 cm (cca 4,5 mg/ml, 375  $\mu$ M). Tyto vzorky byly před měřením přefiltrovány filtry Whatman (velikost pórů 0,2  $\mu$ m) (Cytiva, Chicago, USA). Skenování pro zjištění absorpčního maxima bylo provedeno s FD6A na přístroji Optima AUC proti PB pufru (3 000 RPM, 180-780 nm). Následně byly všechny vzorky podrobeny analýze sedimentační rychlosti na ultracentrifuze Optima AUC (Beckman Coulter, Brea, USA) proti PB pufru (48 000, 30 000 a 15 000 RPM, 281 nm, 500 skenů á 60 s, 20  $^{\circ}$ C nebo 4  $^{\circ}$ C). Hustota (1,007 g/ml) a viskozita (1,59313 cP) pufru a parciální specifické objemy proteinů (viz *Tabulka 13*) byly predikovány pomocí softwaru Sednterp, verze 20130813 BETA (Laue et al., 1992), data byla následně analyzována pomocí softwaru Sedfit, verze 15.01b (Schuck, 2000; Brown a Schuck, 2008) s použitím modelu kontinuální distribuce sedimentačního koeficientu  $c(S)$  a intervalem spolehlivosti 68 % ( $1\sigma$ ). Grafické zpracování bylo provedeno softwarem GUSSI, verze 1.4.2 (Brautigam, 2015).

Tabulka 13: Parciální specifické objemy (při 4 °C) zkoumaných proteinů vypočtené pomocí softwaru Sednterp, verze 20130813 BETA (Laue et al., 1992).

Protein	$\bar{v}_4$ [ml/g]
6His-PDZ3	0,73081
FD5A	0,74985
FD6A	0,74980
G _A 88	0,75953

#### 3.3.4.4 Cirkulární dichroismus (CD)

FD5A, FD6A a G_A88 byly naředěny na koncentraci 0,5 mg/ml a podrobeny CD analýze (PB pufr, 20 °C, spektrální rozsah 195-280 nm, frekvence skenů 10 nm/min, doba odezvy 8 s, šířka spektrálního pásu 1 nm, kyveta křemenná o tloušťce 0,05 cm, vyhodnocení v programu CDPro po korekci výchozích hodnot) (Sreerama a Woody, 2000). Analýzu provedla RNDr. Lucie Bednářová, CSc., ÚOCHB AV ČR na přístroji J-815 CD (JASCO, Tokyo, Japonsko).

#### 3.3.4.5 Diferenční skenovací fluorimetrie (DSF)

FD5A, FD6A a G_A88 byly naředěny na zásobní roztoky o koncentraci 320 µg/ml, 160 µg/ml a 80 µg/ml. Tyto byly smíseny s PB pufrém s barvivem SYPRO Orange do konečné koncentrace barviva 1:500 či 1:1000 a koncentrace proteinu 40 µg/ml, 20 µg/ml, 10 µg/ml v triplicátu. Měření bylo provedeno pomocí přístroje qPCR cykléru qTOWER3/G (Analytik Jena, Jena, Německo) v teplotním rozmezí 25 °C až 90 °C s inkrementem 1 °C za 60 s a měřením fluorescence Ex/Em 492/610 nm. Další měření byla prováděna s výslednou koncentrací barviva 1:500 a proteinů FD5A, FD6A, G_A88 20 µg/ml v teplotních rozmezích od 10 °C do 65-80 °C s inkrementem 1 °C za 60 s nebo 0,2 °C za 45 s (oktaplikát). Ve stejném nastavení byla provedena měření v přítomnosti 1 % DMSO a také v přítomnosti PBM JAM-A ve finálních koncentracích 10 µM a 20 µM a 1% DMSO (triplikát). Teploty tání pro všechna stanovení byly určeny pomocí online nástroje DSFworld (Wu et al., 2020).

## 3.4 Výsledky a diskuse

### 3.4.1 Design a struktury proteinů FD5A a FD6A

V rámci této práce byly charakterizovány dvě fúze domén PDZ3 ZO-1 a G_A88 propojené rigidním linkerem PAPAP, FD5A (forward, PDZ3-linker-G_A88) (viz *Tabulka 14*) a FD6A (reverse, G_A88-linker-PDZ3) (viz *Tabulka 15*), pro které byla *in silico* predikována dynamická alosterie (Kirubakaran et al., 2016). U podobných fúzí, FD3A a FD4A, stejného původu (Kirubakaran et al., 2016) složených z PDZ3 a malé  $\alpha$ -helikální domény TrpCage spojených flexibilním linkerem bylo pozorováno značné mezidoménové ovlivnění (Boušová et al., 2021).

*Tabulka 14: Sekvence a charakteristika FD5A použité v experimentální části této práce. Molekulová hmotnost byla predikována online nástrojem ProtParam (Gasteiger et al., 2005). PDZ3 (tmavě zeleně), G_A88 (rudě), linker (černě), histidinová kotva (tmavě modře), vektorová sekvence (šedivě), signální sekvence (světle modře), štěpná sekvence pro TEV proteasu (oranžově).*

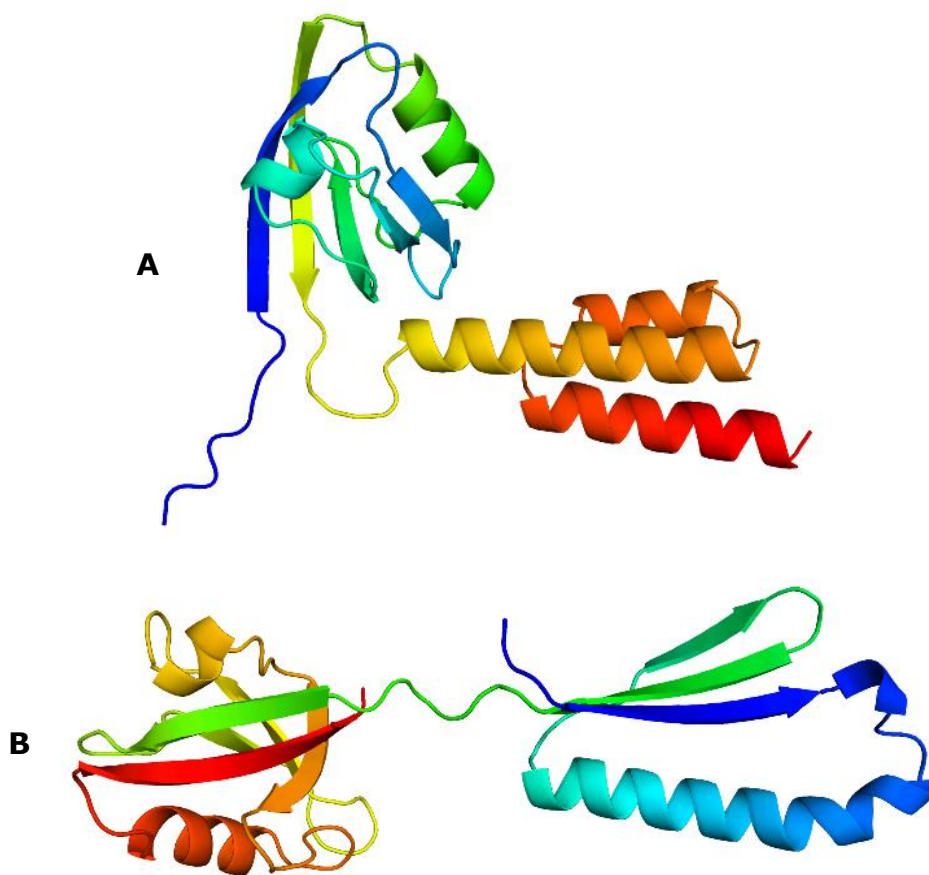
Protein		Počet reziduí				$M_w$ [Da]		
<b>6His-FD5A</b>		175				19 264,15		
5	10	15	20	25	30	35	40	
<b>MGSSH</b>	<b>HHHHH</b>	<b>SSGGH</b>	<b>MLEEN</b>	<b>LYFQG</b>	<b>SHMIL</b>	<b>RPSMK</b>	<b>LVKFR</b>	
45	50	55	60	65	70	75	80	
<b>KGDSV</b>	<b>GLRLA</b>	<b>GGNDV</b>	<b>GIFVA</b>	<b>GVLED</b>	<b>SPAAK</b>	<b>EGLEE</b>	<b>GDQIL</b>	
85	90	95	100	105	110	115	120	
<b>RVNNV</b>	<b>DFTNI</b>	<b>IREEA</b>	<b>VLFLI</b>	<b>DLPKG</b>	<b>EEVTI</b>	<b>LAQKP</b>	<b>APAPT</b>	
125	130	135	140	145	150	155	160	
<b>TYKLI</b>	<b>LNLKQ</b>	<b>AKEEA</b>	<b>IKELV</b>	<b>DAGIA</b>	<b>EKYIK</b>	<b>LIANA</b>	<b>KTVEG</b>	
165	170	175						
<b>VWTLK</b>	<b>DEILT</b>	<b>FTVTE</b>						
<b>FD5A</b>		151				16 485,17		
5	10	15	20	25	30	35	40	
<b>GSHMI</b>	<b>LRPSM</b>	<b>KLVKF</b>	<b>RKGDS</b>	<b>VGLRL</b>	<b>AGGND</b>	<b>VGIFV</b>	<b>AGVLE</b>	
45	50	55	60	65	70	75	80	
<b>DSPAA</b>	<b>KEGLE</b>	<b>EGDQI</b>	<b>LRVNN</b>	<b>VDFTN</b>	<b>IIREE</b>	<b>AVLFL</b>	<b>LDLPK</b>	
85	90	95	100	105	110	115	120	
<b>GEEVT</b>	<b>ILAQK</b>	<b>PAPAP</b>	<b>TTYKL</b>	<b>IILNK</b>	<b>QAKEE</b>	<b>AIKEL</b>	<b>VDAGI</b>	
125	130	135	140	145	150	155		
<b>AEKYI</b>	<b>KLIAN</b>	<b>AKTVE</b>	<b>GVWTL</b>	<b>KDEIL</b>	<b>TFTVT</b>	<b>E</b>		

Tabulka 15: Sekvence a charakteristika FD6A použité v experimentální části této práce. Molekulová hmotnost byla predikována online nástrojem ProtParam (Gasteiger et al., 2005). PDZ3 (tmavě zeleně), GA88 (rudě), linker (černě), histidinová kotva (tmavě modře), vektorová sekvence (šedivě), štěpná sekvence pro TEV proteasu (oranžově).

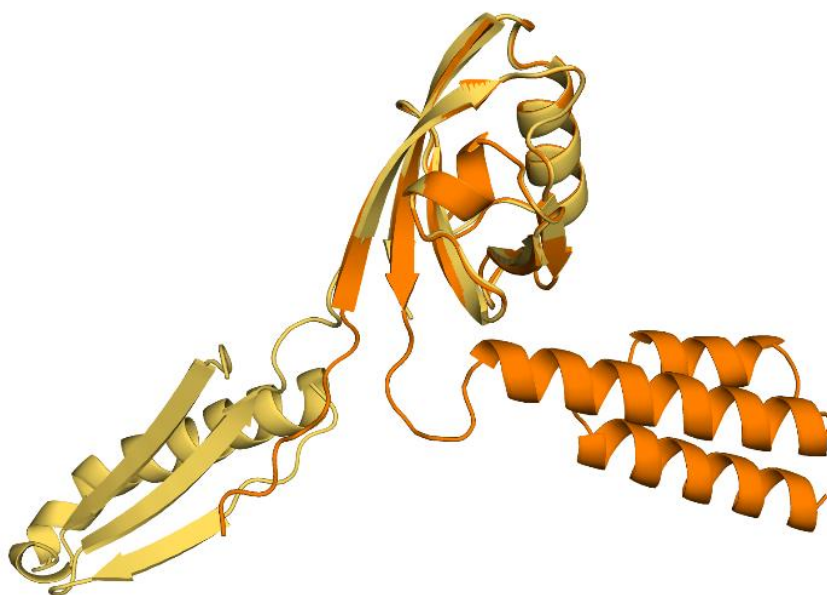
Protein		Počet reziduí				$M_w$ [Da]		
<b>6His-FD6A</b>		174				19 008,81		
5	10	15	20	25	30	35	40	
<b>MGSSH</b>	<b>HHHHH</b>	<b>SSGLV</b>	<b>PRGSH</b>	<b>MLEEN</b>	<b>LYFQG</b>	<b>GGTTY</b>	<b>KLILN</b>	
45	50	55	60	65	70	75	80	
<b>LKQAK</b>	<b>EEAIK</b>	<b>ELVDA</b>	<b>GIAEK</b>	<b>YIKLI</b>	<b>ANAKT</b>	<b>VEGVW</b>	<b>TLKDE</b>	
85	90	95	100	105	110	115	120	
<b>ILTFT</b>	<b>VTEPA</b>	<b>PAPMK</b>	<b>LVKFR</b>	<b>KGDSV</b>	<b>GLRLA</b>	<b>GGNDV</b>	<b>GIFVA</b>	
125	130	135	140	145	150	155	160	
<b>GVLED</b>	<b>SPAAK</b>	<b>EGLEE</b>	<b>GDQIL</b>	<b>RVNNV</b>	<b>DFTNI</b>	<b>IREEA</b>	<b>VLFLN</b>	
165	170	175						
<b>DLPKG</b>	<b>EEVTI</b>	<b>LAQK</b>						
<b>FD6A</b>		145				15 677,16		
5	10	15	20	25	30	35	40	
<b>GGGTT</b>	<b>YKLIL</b>	<b>NLKQA</b>	<b>KEEAI</b>	<b>KELVD</b>	<b>AGIAE</b>	<b>KYIKL</b>	<b>IANAK</b>	
45	50	55	60	65	70	75	80	
<b>TVEGV</b>	<b>WTLKD</b>	<b>EILTF</b>	<b>TVTEP</b>	<b>APAPM</b>	<b>KLVKF</b>	<b>RKGDS</b>	<b>VGLRL</b>	
85	90	95	100	105	110	115	120	
<b>AGGND</b>	<b>VGIFV</b>	<b>AGVLE</b>	<b>DSPAA</b>	<b>KEGLE</b>	<b>EGDQI</b>	<b>LRVNN</b>	<b>VDFTN</b>	
125	130	135	140	145				
<b>IIREE</b>	<b>AVLFL</b>	<b>LDLPK</b>	<b>GEEVT</b>	<b>ILAQK</b>				

Z těchto primárních struktur proteinů FD5A a FD6A byla predikována jejich sekundární a terciární struktura pomocí online nástroje ColabFold (Mirdita et al., 2022) pracujícího na bázi algoritmu AlphaFold2 organizace Google DeepMind. Tento nástroj se stal velmi oblíbeným po rezolutních výhrách v CASP13 a 14 zejména díky své přesnosti. Pracuje na principu umělé neuronové sítě s hlubokým učením a koevoluční analýzy (AlQuraishi, 2019; Baek et al., 2021).

Z predikovaných struktur (viz Obr. 9) se zdá, že oba proteiny nabírají v prostoru relativně podobné konformace (viz Obr. 10), FD5A je však kompaktnější. Ani jeden z modelů nepoukazuje na významné mezidoménové interakce mezi PDZ3 a GA88, což je pravděpodobně dáno užitím rigidního linkeru. Zajímavý fakt je, že GA88 v modelu FD6A, nikoliv však FD5A, získává strukturu GB88. Zda se jedná o efekt tandemové proteinové fúze na konformaci, či pouze chybu predikčního algoritmu, musí být determinováno experimentálně.



Obr. 9: Predikované struktury studovaných fúzí. A: FD5A. B: FD6A. Vytvořeno pomocí softwaru PyMOL Molecular Graphics System, Verze 2.4.1, Schrödinger, LLC.



Obr. 10: Proložení modelů FD5A (oranžově) a FD6A (žlutě) (sekvenční alignment PDZ3). Vytvořeno pomocí softwaru PyMOL Molecular Graphics System, Verze 2.4.1, Schrödinger, LLC.

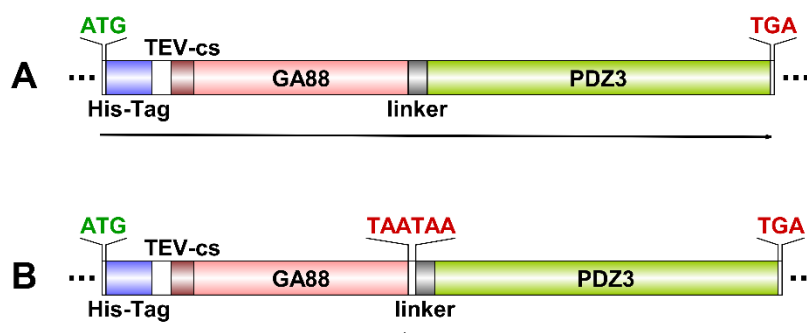


### 3.4.2 Transformace *E. coli*

Zakoupenými plazmidy pET15b_FD5A a pET15b_FD6A byly transformovány buňky *E. coli* BL21(DE3) a TOP10 dle standardního laboratorního postupu. Je záhodno vysvětlit, že zatímco *E. coli* BL21(DE3) je kmen primárně určený k proteinové expresi (často právě s pET plazmidy s T7 promotorem) díky snížené proteolytické schopnosti (deficit proteas *Lon* a *OmpT*) (Gopal a Kumar, 2013), *E. coli* TOP10 byly vyvinuty přímo ke klonování plazmidů (dle informací poskytovaných výrobcem) a za těmito účely byly využity také v této práci. Existuje samozřejmě celá řada alternativ ke kmenům BL21(DE3) (např. RIL) i TOP10 (např. XL-1 Blue), výběr by se povětšinou měl řídit zvolenou metodou transformace a specifickými nároky proteinu (Chan et al., 2013; Gopal a Kumar, 2013).

### 3.4.3 Příprava plazmidu pro expresi G_A88

Plazmid pro expresi G_A88 nebyl, na rozdíl od plazmidů pET15b_FD5A a pET15b_FD6A, zakoupen, neboť jej lze připravit vložením terminačních kodonů do pET15b_FD6A, a to pomocí metody PCR. Tato metoda je jednou z nejrozšířenějších technik molekulární biologie, neboť nabízí možnost exponenciální *in vitro* amplifikace templátové DNA, nutná je však předchozí znalost její sekvence. S touto informací mohou být navrženy vhodné primery, vymezující amplifikovaný úsek a sloužící jako počátek polymerace (Rahman et al., 2013). Pakliže se tyto primery jen mírně odlišují od sekvence komplementární k templátu, jsou tyto odlišnosti do amplifikovaných molekul zapracovány (viz Obr. 11), v čemž spočívá princip mutagenese pomocí PCR (Hemsley et al., 1989).



Obr. 11: Znáznornění procesu mutagenese pET15b_FD6A pomocí PCR. A: Úsek genu pro FD6A před PCR (pET15b_FD6A). B: Stejný úsek po PCR (pET15b_GA88). Start kodon (ATG), úsek kódující histidinovou kotvu (His-Tag) modře, místo štěpení pro TEV proteasu (TEV-cs) hnědě, G_A88 doménu červeně, linker šedivě a PDZ3 doménu zeleně; terminační kodony (TGA, TAA). Šipky pod diagramy ukazují rozsah translace takového genu. Zpracováno pomocí softwaru DOG 2.0 (Ren et al., 2009).

Pro tento případ byly proto designovány primery Fw (z angl. forward) a Rv (z angl. reverse) (viz *Tabulka 16*) dosedající v místě napojení G_{A88} k linkeru, v nichž byly přítomny dva terminační kodony (TAATAA, resp. TTATTA).

Terminační kodon TAA patří vedle TGA k nejoblíbenějším (oba použité v této práci, viz *Obr. 11*), hojně zastoupeným v genomu *E. coli*. Volba terminačních kodonů závisí však také na bázi tento kodón následující. Zatímco kombinace TAA-T je extrémně efektivní v terminaci translace (cca 80 %), TAA-C, která by vznikla vnesením pouze jediného TAA, patří do skupiny signifikantně méně účinných (cca 30 %) (Poole et al., 1995). Zdvojení terminačních kodonů má v tomto případě tedy dvojí benefit – v případě chyby primární terminace existuje pojistka ve formě terminace sekundární, ale i zavedení značně efektivní TAA-T sekvence, která chybovost výrazně snižuje.

Vzhledem k vysokému podílu guaninových a cytosinových bází v obou primerech se i teplota tání primerů pohybovala velmi vysoko (cca 85 °C). Využití těchto teplot při hybridizaci samozřejmě není možné (maximum doporučené výrobcem Q5 je 72 °C) (New England BioLabs Inc., 2023), proto bylo připraveno několik vzorků a každý byl hybridizován při jiné teplotě (gradient od 62 do 72 °C).

*Tabulka 16: Strukturní a biofyzikální popis primerů použitých pro mutagenezi (Fw, Rv) v porovnání s výchozí sekvencí (FD6A). Úseky kódující strukturní komponenty jsou vyznačeny barevně: červeně G_{A88}, šedivě PAPAP linker, zeleně PDZ3. Terminační kodony tučně. Teplota tání a zastoupení guaninových a cytosinových bází (GC) byly spočteny pomocí online nástroje NEB Tm calculator ([tmcalculator.neb.com](http://tmcalculator.neb.com)).*

FD6A	... CCCTGAAGGACGAAATCCTGACCTTTACCGTTACCGAG- -CCGGCGCCGGCGCCGGGTAGCCACATGATTCTG ...	T _m [°C]	GC [%]
<b>Fw</b>	<b>GTTACCGAGTAATAA</b> CCGGCGCCGGCGCCGGGTAGCCACATGATTCTG	85	60
<b>Rv</b>	GCCGGCGCCGG <b>TTATTA</b> CTCGGTAACGGTAAAGGTCAGGATTT <b>CGTCCTTCAGGG</b>	84	56

V PCR probíhá amplifikace DNA enzymaticky, pomocí DNA polymeras (obdobně jako *in vivo*), přičemž těchto enzymů je široce využíváno hned několik. Mezi nejlevnější, ale také nejméně spolehlivé varianty se řadí DNA polymerasa z *Thermus aquaticus* (*Taq* polymerasa), hojně využívaná v např. rutinních testech; její chybovost se odhaduje na 1-20×10⁻⁵ chyb na bázi za duplikaci. Dnes však existuje řada alternativ, např. polymerasy Deep Vent, Phusion, *Pfu* či KOD, nabízející řádově vyšší spolehlivost. Q5 polymerasa, použitá v této práci, se však v současnosti řadí do samého čela s chybovostí více než 250× nižší než *Taq* polymerasa, proto je i přes vyšší cenu ideální pro delikátnější amplifikace, kde existují vzhledem k dalším záměrům se vzorkem vysoké nároky na sekvenční správnost (Potapov a Ong, 2017; McInerney et al., 2014).

Po proběhnutí PCR reakce bylo nutné odstranit výchozí plazmidovou DNA, která terminační kodony neobsahuje. Za tímto účelem byla použita *DpnI* restriktasa, protein *Diplococcus pneumoniae*, který štěpí methylovanou DNA v místech G-A^{Me}-T-C v obou

řetězcích (Lacks a Greenberg, 1977). Tento enzym bývá pro tyto účely využíván, neboť je s to štěpit výlučně DNA, která se v minulosti nacházela v buňkách (označení methylocí na adeninu) (Shenoy a Visweswariah, 2003). Získanou směsí mohly být následně transformovány *E. coli* bez rizika, že dojde ke vzniku směsi produkující FD6A i G_A88.

Plazmid byl z transformovaných buněk znovu izolován, aby mohla být ověřena jeho sekvence, tj. úspěšnost PCR mutageneze a transformace. Toto stanovení provedla společnost Eurofins se dvěma vzorky (pro sekvenování v obou směrech) plazmidu z každé hybridizační teploty. Ačkoliv se výpočetně stanovená optimální hybridizační teplota pohybovala výrazně výše než použité teploty (viz výše), sekvenční správnost (přítomnost dvou terminačních kodonů) vykazovaly vzorky ze všech teplot gradientu 62-72 °C. Pro další práci byly použity buňky s pET15b_G_A88 připraveným při teplotě hybridizace 64 °C.

#### 3.4.4 Purifikační proces FD5A, FD6A, G_A88

Purifikace proteinů představuje kritický krok charakterizačního procesu – je zásadní protein získat v optimální kvalitě i kvantitě pro další využití biofyzikálními metodami. Pro purifikaci FD5A, FD6A a G_A88 v této práci byl použit protokol publikovaný v Kirubakaran et al., 2016 (viz 3.3.2 a 3.3.3), neboť se v minulosti prokázal jako vhodný pro jiné fúzní proteiny s PDZ doménou (Kirubakaran et al., 2016; Boušová et al., 2021).

Uvolnění proteinů z buněk probíhalo sonikační lyzací buněčné kultury za přítomnosti lysozymu a detergentu a následnou centrifugací lyzátu. Lyzace může být prováděna mnoha mechanickými, chemickými a biologickými přístupy, nebo jejich kombinacemi, přičemž každý přístup má své výhody a nevýhody, může produkovat nižší výtěžky proteinu (např. enzymatická lyze, lyze detergenty), být pomalejší (alkalická lyze) nebo poškozovat citlivější proteiny (tepelná lyze, lyze detergenty, tlakovými homogenizátory). Sonikace (také nazývaná ultrasonická kavitace) je značně univerzální metoda umožňující ve velkém měřítku lyzovat prakticky jakékoliv kultury. Ačkoliv je zpravidla prováděna za intenzivního chlazení, existuje zde nebezpečí poškození velmi citlivých proteinů a v případě setrvání exprimovaného proteinu v buněčných zbytcích (po centrifugaci tvořící peletu) může být obtížné tyto proteiny dále purifikovat (Shehadul Islam et al., 2017).

Získaný roztok proteinů byl podroben IMAC, a to s využitím N-terminální histidinové kotvy (tzv. His-Tag, sekvence GSSHHHHHSSG), v moderní molekulární biologii velmi favorizované peptidové kotvy s afinitou k bivalentním iontům. Tyto ionty (Cu²⁺, Zn²⁺, Co²⁺, nejčastěji Ni²⁺) jsou upoutány chelatačním ligandem (iminodioctová kyselina – IDA, častěji nitrilotrioctová kyselina – NTA) k matici (zpravidla agarosa) v separační koloně (Block et al., 2009), kde jsou následně odděleny promýváním od nečistot a eluovány imidazolem. Ačkoliv existuje celá řada dalších peptidových značek (MBP, GST, kalmodulin vázající peptid – CBP, epitop FLAG ad.) (Young et al., 2012) a příslušných komerčních matic, vysoká kapacita a nízké pořizovací náklady (i za cenu horší čistoty) u IMAC s použitím histidinové kotvy vysvětluje její rozšířenost (Lichy et al., 2005).

Malá velikost histidinové kotvy je rovněž benefitem, neboť díky ní strukturu většiny proteinů ovlivňuje jen zanedbatelně a nemusí být pro charakterizaci odštěpována (Carson et al., 2007). Některé proteiny označené touto kotvou se však mohou od nativních lišit, vykazovat pozměněnou strukturu, sníženou rozpustnost a enzymatickou aktivitu, a to i v závislosti na umístění kotvy (Woestenenk et al., 2004; Zhao a Huang, 2016; Yeon et al., 2014; Meng et al., 2015) (což souvisí s jevy diskutovanými v 1.3.1), proto se někdy přistupuje k jejímu odstranění. Po odštěpení histidinové kotvy lze navíc protein znovu podstoupit IMAC (subtraktivní), kdy dojde k dalšímu přečištění.

Odštěpení kotev lze provádět chemicky, nicméně enzymatický přístup je zde jednoznačně preferován z důvodu vyšší specifity a minimalizace risku denaturace a modifikace reziduí. Používané enzymy (proteasy, peptidasy) bývají v praxi často opatřeny stejnou kotvou, kterou štěpí, a mohou tak být efektivně eliminovány ze vzorku stejnou subtraktivní IMAC jako odštěpený fragment kotvy (Arnau et al., 2006; Nilsson et al., 1997). Vhodných proteas existuje široké spektrum (faktor Xa, enterokinasa, trombin, sortasa A, DAPasa ad.), každá nabízí unikátní místo štěpení (z angl. cleavage site), také však unikátní výhody a nevýhody. Ve spojitosti s histidinovou kotvou se však nejčastěji skloňuje proteasa viru leptané

mozaiky tabáku (z angl. Tobacco etch virus – TEV) se štěpným místem ENLYFQ*G/S použita rovněž v této práci, která je vysoce efektivní i selektivní a co do produkce nenáročná (Arnau et al., 2006; Raran-Kurussi et al., 2017; Kapust et al., 2002).

Afinitní chromatografie (imunoafinitní chromatografie nebo IMAC) dnes jednoznačně mezi metodami pro purifikaci rekombinantních proteinů dominuje; odlišné přístupy (iontové výměnná chromatografie – IEC, chromatografie hydrofobních interakcí – HIC – ad.) se používají zřídka a zpravidla v kombinaci s IMAC. Tak tomu bylo i v této práci, kde po IMAC následovala vylučovací chromatografie (z angl. size exclusion chromatography – SEC, také nazývaná gelová filtrace) na FPLC. Tento typ kapalinového chromatografu se vyznačuje instrumentální optimalizací pro separaci biomolekul (oproti vysokoúčinné kapalinové chromatografii – HPLC – používá nižší tlaky, dávkování většího množství vzorku a méně organických rozpouštědel), přičemž principiálně může pracovat v několika módech, včetně právě HIC, IEC a SEC. Poslední zmíněná, použitá v této práci, představuje častý finální krok purifikací, umožňuje odstranění zbytkových kontaminantů separací dle hydrodynamického poloměru, oddělení agregovaných či oligomerních stavů a výměnu pufru. Na koloně však dochází k ředění naneseného vzorku (3-5×) a její kapacita samozřejmě rovněž není neomezená, proto se zpravidla před FPLC protein koncentruje (Madadlou et al., 2011; Wingfield, 2015; Janson, 2011).

Celý purifikační proces byl vizualizován pomocí elektroforézy v polyakrylamidovém gelu v přítomnosti dodecylsírany sodného (SDS-PAGE). Pomocí ní lze orientačně usuzovat na čistotu připraveného proteinu, ale také jeho přibližnou velikost, neboť SDS zde funguje jako denaturační agens, který propůjčuje proteinům uniformní náboj, takže jsou prostupem skrz polyakrylamidový gel v elektrickém poli děleny výhradně na základě své molekulové hmotnosti. Vyznačuje se instrumentální i manuální nenáročností, rychlostí provedení analýzy a nízkou cenou, což z ní činí jednu z nejrozšířenějších a nejpoužívanějších metod biochemie. Alternativy pro ověřování čistoty proteinů, jako kapalinové chromatografie (vysokoúčinná s obrácenými fázemi – RP-HPLC – nebo FPLC), DLS a MS se používají zřídka, ve specifických případech a často ve spojitosti se současnou charakterizací. Detekce absorbance ( $A_{280}$ ) po SEC, rovněž použitá v této práci, poskytuje paralelní data o čistotě a velikosti proteinu v konečné fázi purifikace (Rhodes a Laue, 2009).

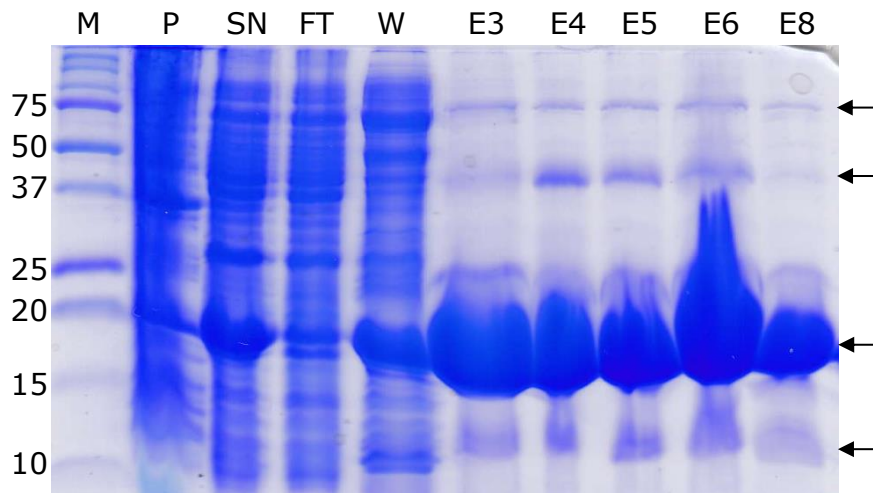
Ověření sekvence proteinů bylo provedeno pomocí MS. Pro ověření identity proteinů se často používá Western blot využívající interakci antigen-protilátka; tato metoda však nepřístupňuje možnost ověření celé, ani většiny, primární struktury izolovaného proteinu. Zatímco v dřívějších dobách se pro tyto účely užívalo sekvenace Edmanovým odbouráváním následujícím Western blot (Gevaert a Vandekerckhove, 2000), dnes je tato technika překonána metodami hmotnostního otisku prstu peptidu (z angl. peptide mass fingerprinting – PMF –, částečné pokrytí, hledání specifických sekvencí) a sekvenace pomocí tandemové hmotnostní spektrometrie (MS/MS, úplné pokrytí sekvence). Prvním krokem těchto přístupů je izolace již denaturovaného proteinu z elektroforetického (SDS-PAGE) gelu, dále je pokračováno jeho enzymatickou fragmentací (nejčastěji trypsinem), příslušnou MS analýzou a *in silico* evaluací (Cottrell, 2011; Hjernø, 2006). V této práci byla sekvence ověřena metodou PMF, a to pomocí hmotnostního spektrometru s maticí asistovanou laserovou desorpcí/ionizací (MALDI) a analyzátozem doby letu (TOF), které se velmi často využívají právě pro analýzu biopolymerů, neboť MALDI dokáže obstojně pracovat s roztoky

pufří a solí a poskytuje velmi měkkou ionizaci a jakožto technika založená na pulzní bázi se přirozeně páruje s rovněž pulzním analyzátozem, tj. TOF (Gevaert a Vandekerckhove, 2000).

#### 3.4.4.1 Purifikace rekombinantního FD5A

Dle purifikačního protokolu pro fúzní proteiny (Kirubakaran et al., 2016) byl nejprve FD5A s histidinovou kotvou (6His-FD5A) exprimován v *E. coli*. Tyto kultury byly po inkubaci lyzovány a po odstranění buněčných zbytků byl získaný roztok proteinů podroben IMAC na Ni/IDA za zisku 20 elučních frakcí á 1 ml. V rámci celého procesu byly sbírány vzorky pro SDS-PAGE: vzorek pelety se zbytky buněčného lyzátu (P), roztoku proteinů z buněčného lyzátu (supernatant – SN), frakce tohoto roztoku prošedší skrz afinitní matici (flow-through – FT), frakce proteinů uvolněných při promývání této matrice (wash – W) a některé eluční frakce (E3-E6, E8).

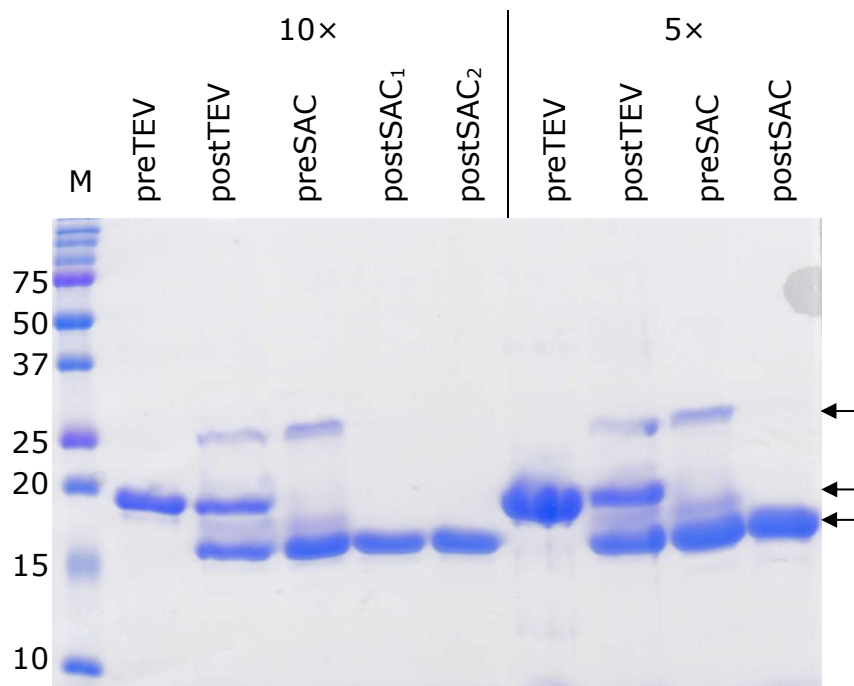
Z elektroforetického gelu (viz *Obr. 12*) se zdá, že exprimovaný rekombinantní 6His-FD5A po lyzaci buněk dobře přešel do roztoku, neboť protein o velikosti 6His-FD5A (15-20 kDa) se vyskytoval ve vysoké koncentraci v SN, nikoliv však v P. Vazba na afinitní matici byla rovněž dostatečná, jak napovídá FT frakce, kde se protein zmíněné velikosti vyskytuje minimálně, a ačkoliv část domnělého 6His-FD5A byla promýváním odstraněna (viz W), všechny eluční frakce vykazovaly vysoký výtěžek (v případě E4-6 dokonce takový, že došlo při tepelné denaturaci k agregaci a zatuhnutí vzorků, takže nemohly být ani celé přeneseny na gel). Čistota proteinu v elučních frakcích nebyla dokonalá (kontaminanty cca 10, 37 a 75 kDa velké ad., jejich signály jsou však amplifikovány nanesením velmi koncentrovaných roztoků), nicméně majorita proteinů vlastních *E. coli* byla pomocí IMAC od kýženého proteinu s histidinovou kotvou oddělena. Pro zvýšení čistoty by bylo nasnadě optimalizovat potenci (koncentraci imidazolu) a množství promývacího pufří, nicméně vzhledem k přítomnosti dalších purifikačních kroků, kterými mohou být nečistoty odstraněny, lze IMAC považovat za úspěšnou.



Obr. 12: SDS-PAGE frakcí purifikačního procesu FD5A. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), peleta lyzátu buněčné kultury (P), supernatant lyzátu (SN), frakce na IMAC nezachycených proteinů (FT), proteiny vymyté promývacím pufrém (W), 3. až 6. a 8. eluční frakce (E3-E6, E8).

Vzhledem k neznámým vlastnostem zkoumaného fúzního proteinu byly všechny eluční frakce sloučeny a část naředěna 5× a část 10× denaturačním pufrém (u méně koncentrovaného vzorku lze v případě nestability očekávat menší náchylnost k agregaci v klidovém stavu) (Treuheit et al., 2002). Další kroky purifikace byly následně prováděny pro oba tyto vzorky paralelně. Po denuraci byly vzorky inkubovány s 6His-TEV proteasou pro odštěpení histidinové kotvy, renaturovány a 6His-TEV proteasa, neštěpený protein a fragmenty histidinové kotvy byly odstraněny subtraktivní IMAC. Tento proces byl rovněž vizualizován na SDS-PAGE pomocí průběžně odebíraných vzorků: před přidáním 6His-TEV proteasy (preTEV 5×, preTEV 10×), po inkubaci s 6His-TEV proteasou 90 min při 30 °C (postTEV 5×, 10×), po inkubaci s 6His-TEV O/N při 4 °C (preSAC 5×, 10×), frakce proědí afinitní maticí při subtraktivní IMAC (postSAC 5×, postSAC 10× část 1 a část 2).

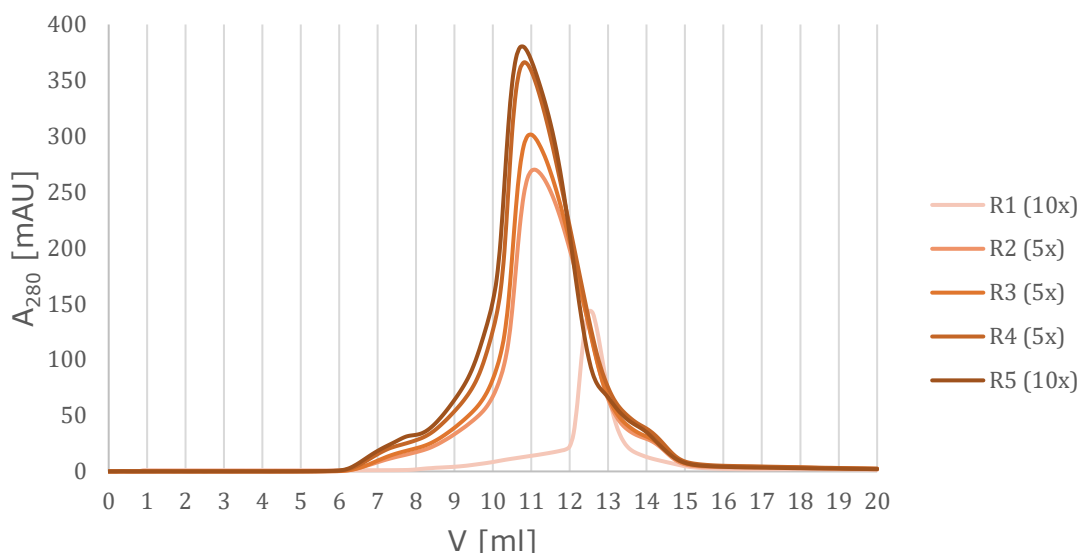
Na elektroforetickém gelu (viz Obr. 13) je patrný rozdíl v kompozici jednotlivých vzorků. Před přidáním 6His-TEV proteasy lze v obou vzorcích (preTEV 5×, 10×) pozorovat prakticky čistý (kontaminanty viditelné na Obr. 12 nelze již rozeznat vzhledem k ředění) výchozí protein o velikosti blízké 20 kDa (domnělý 6His-FD5A), po inkubaci s 6His-TEV proteasou lze pozorovat část tohoto proteinu zkrácenou na cca 15 kDa (odštěpená histidinová kotva již opustila gel), zbytek výchozího proteinu (cca 20 kDa) a samozřejmě také 6His-TEV proteasu (cca 25 kDa). Po dialýze přes noc se již výchozí protein ve směsi v signifikantním množství nevyskytoval. Po provedení subtraktivní IMAC se již ve vzorcích (postSAC 5×, 10×) nacházel ve detekovatelném množství pouze protein o velikosti cca 15 kDa, což odpovídá velikosti FD5A (16,485 kDa). Tento purifikační krok lze tedy rovněž považovat za úspěšný.



Obr. 13: SDS-PAGE procesu odštěpení histidinové kotvy FD5A pomocí TEV proteasy. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), dále vzorky 10x naředěné části: před přidáním proteasy (preTEV), po inkubaci (postTEV), před subtraktivní IMAC (preSAC) a po této IMAC, část 1 (postSAC₁) a 2 (postSAC₂); následně vzorky 5x naředěné části: před přidáním proteasy (preTEV), po inkubaci (postTEV), před subtraktivní IMAC (preSAC) a po této IMAC (postSAC).

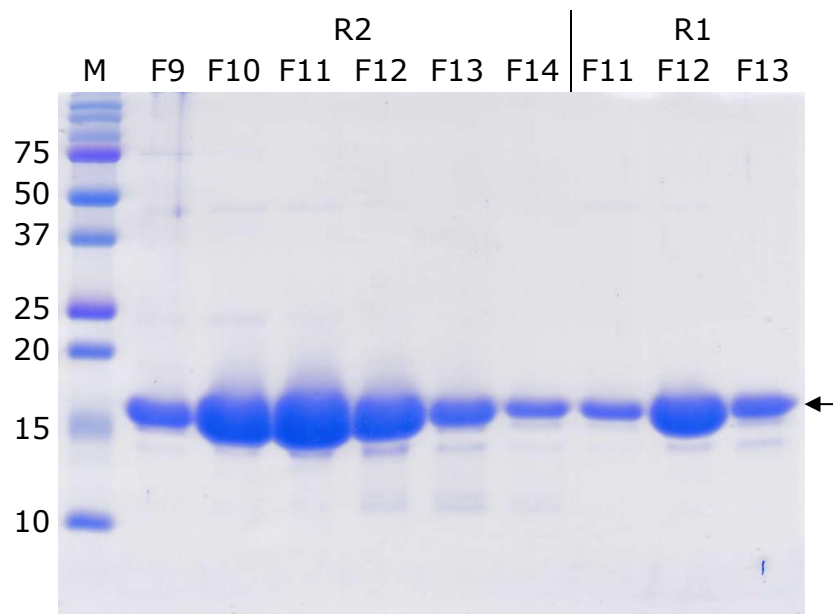
Posledním purifikačním krokem bylo provedení vylučovací chromatografie pro přečištění a frakcionaci vzorků podle molekulové velikosti. Při koncentrování v koncentrátorech Amicon byly roztoky pečlivě monitorovány, v oblasti filtrační membrány docházelo k akumulaci proteinu, což se projevovalo zvýšením viskozity. Agregaci bylo předcházeno resuspendací. Takto bylo připraveno 5×2 ml (3 z 10×, 2 z 5× ředěné části) roztoku o koncentraci proteinu 4-6 mg/ml. Tyto vzorky byly postupně nadávkovány na FPLC (5 separací v módu SEC) kde byly děleny a převedeny do fosfátového (PB) pufru. Dle získaných chromatogramů z detekce absorbance při 280 nm (viz Obr. 14, kde byla data vertikálně transformována tak, aby minimum jejich absorbance odpovídalo nule a vyjádřeny tedy byly relativní absorbance) byly v každé separaci selektovány frakce (sbírány po 1 ml, číslování odpovídá retenčnímu objemu, kde byl sběr započat) s vysokou koncentrací proteinu. První separace (R1) byla poznamenána chybou při kontrole instrumentu – při nanášení vzorku na kolonu došlo ke ztrátě části produktu únikem malého množství kapaliny ze systému z důvodu nedostatečně těsnícího spoje, což je rovněž příčinou posunu chromatografického píku o cca 2 ml i snížení intenzity (viz Obr. 14).





Obr. 14: Chromatogramy SEC separací FD5A, tj. závislosti relativní absorbance při 280 nm na retenčním objemu pro pět separací (R1-R5) provedených z původních koncentrátů vzorků ředěných 5× a 10×.

Pro první dvě analýzy byla čistota vybraných získaných frakcí dále ověřena pomocí SDS-PAGE. První separace byla provedena ze směsného koncentrátu 10×, druhá pak z koncentrátu 5× ředěné části. Vzhledem k tomu, že i ostatní separace byly provedeny ze stejných koncentrátů, a to v krátkém časovém intervalu, lze u nich předpokládat obdobnou kompozici odpovídajících frakcí. Ve všech frakcích na elektroforetickém gelu (viz Obr. 15) byl přítomen kýžený protein velikostně odpovídající FD5A (cca 15 kDa) v dobré čistotě. Rané frakce (9.-11. druhé separace, 11. a 12. první separace) prokazovaly minimum kontaminantů a velmi vysoký podíl domnělého FD5A, pozdější frakce (12.-14. frakce druhé separace) již obsahovaly větší než zanedbatelné množství menších proteinů v poměru ke kýženému analytu. Toto odpovídá rovněž datům z chromatogramu, kde lze před hlavním vrcholem i za ním pozorovat průběh neodpovídající rozložení dle Gaussovy funkce. Tuto deformaci lze přisoudit právě kontaminacím, potenciálně také vysoké koncentraci nadávkovaného vzorku (Hagel, 2011).



Obr. 15: SDS-PAGE frakcí prvních analýz FD5A pomocí SEC. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), následují frakce druhé (R2) separace: 9. až 14. frakce (F9-R14), pak frakce z první (R1) separace: 11. až 13. frakce (F11-F13).

Celkově bylo ze všech pěti separací získáno 18 ml kvalitních frakcí s koncentrací v rozmezí od 10  $\mu$ M po 180  $\mu$ M se sumárním výtěžkem proteinu cca 27 mg. MS analýza potvrdila identitu proteinu při zjištění 98% shody se sekvencí FD5A (viz Příloha II). Získané frakce byly rozděleny na alikvoty, zamrazeny a sloužily dále jako zásobní vzorky pro biofyzikální charakterizace FD5A.

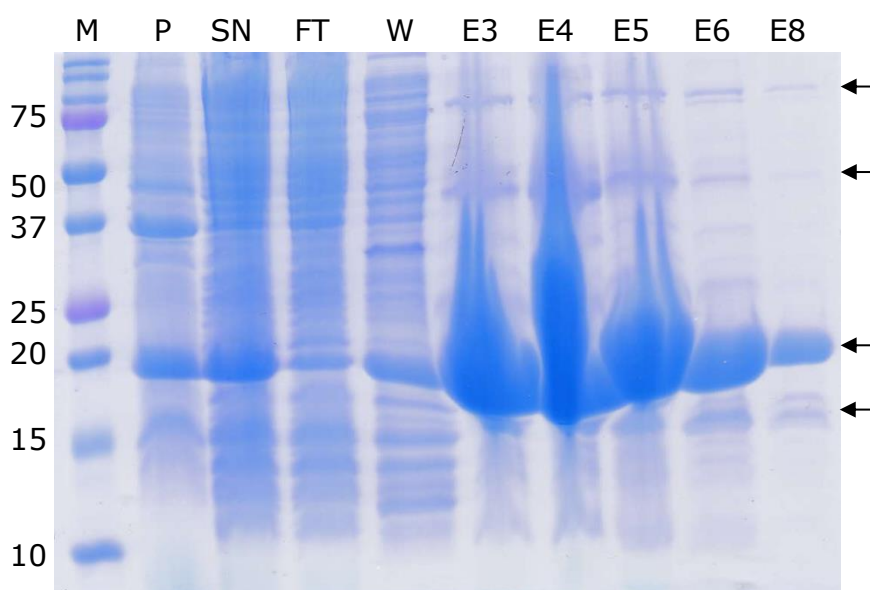
#### 3.4.4.2 Purifikace rekombinantního FD6A

Expres 6His-FD A v *E. coli*, izolace i IMAC probíhaly shodně s postupem uvedeným v 3.3.3 a 3.4.4.1, rovněž SDS-PAGE byla provedena se shodně získanými vzorky z průběhu purifikace.

Z elektroforetického gelu (viz Obr. 16) se zdá, že ačkoliv majorita zkoumaného proteinu (velikost cca 20 kDa) přešla ze zbytků buněk v lyzátu do roztoku (SN), lze jej pozorovat také v peletě. Tato skutečnost by mohla naznačovat horší rozpustnost 6His-FD6A oproti 6His-FD5A, tvorbu inkluzních tělísek nedodržením optimálních podmínek při expresi, nebo neúplnou lyzaci buněčné kultury (Wingfield, 2015; Kane a Hartley, 1988; Shehadul Islam et al., 2017). Ačkoliv první z hypotéz se zdá být zpochybněna pozorováním při koncentrování FD6A a FD5A (viz níže, 3.4.4.1), kdy FD6A vykazoval dobrou rozpustnost, neměl by být opomenut potenciální negativní vliv histidinové kotvy na rozpustnost popsany u některých proteinů (Woestenenk et al., 2004). Příčinu tohoto jevu by však s jistotou determinovala pouze další experimentální práce. Kdyby bylo primární motivací maximalizovat výtěžek purifikace, mohly by být paralelně použity alternativní protokoly

pro purifikaci proteinů z buněčné pelety (Wingfield, 2015), popřípadě zopakování exprese s jinou kotvou (Waugh, 2005), získaný výtěžek byl však zhodnocen jako dostačující.

Část proteinu rovněž prošla afinitní maticí bez interakce (FT), nebo byla vymyta promývacím pufrém (W). Nicméně jsou to eluční frakce, kde lze pozorovat značné množství kýženého proteinu (cca 20 kDa). Přítomné jsou, podobně jako v purifikaci FD5A, kontaminanty (cca 75, 37, 25 a 15 kDa a menší), nicméně vzorek je viditelně velmi koncentrovaný a relativní poměr kontaminantů vůči proteinu o velikosti 6His-FD6A je marginální.

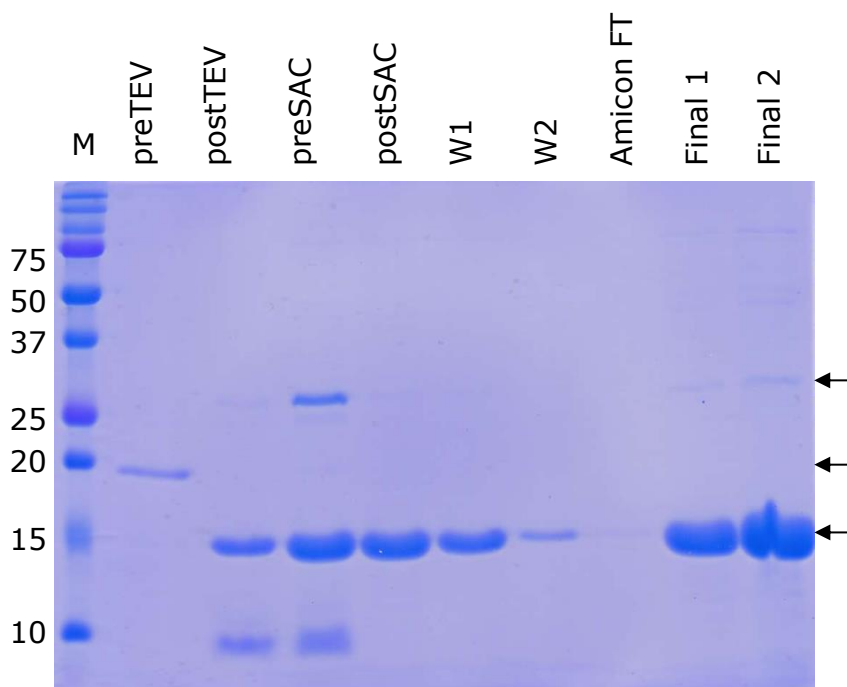


Obr. 16: SDS-PAGE frakcí purifikačního procesu FD6A. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), peleta lyzátu buněčné kultury (P), supernatant lyzátu (SN), frakce na IMAC nezachycených proteinů (FT), proteiny vymyté promývacím pufrém (W), 3. až 6. a 8. eluční frakce (E3-E6, E8).

Domnělý 6His-FD6A byl dále 5× naředěn a denaturován. Proces odštěpení histidinové kotvy byl vizualizován na SDS-PAGE pomocí průběžně odebíraných vzorků: denaturovaného roztoku před přidáním 6His-TEV proteasy (preTEV), po inkubaci s 6His-TEV proteasou 90 min při 30 °C (postTEV), po inkubaci s 6His-TEV O/N při 4 °C (preSAC), frakce nezadržená na afinitní maticí při subtraktivní IMAC (postSAC) a frakce vymyté TEV pufrém 2 (W1, W2). Elektroforetické analýze byl podroben také filtrát koncentrátoru Amicon (Amicon FT) a koncentrované vzorky připraveného proteinu (Final 1 a 2).

Na elektroforetickém gelu (viz Obr. 17) je, podobně jako u FD5A (viz 3.4.4.1, Obr. 13), zřejmá predikovaná přeměna obsaženého proteinu. Domnělý 6His-FD6A (cca 20 kDa) v preTEV se po přidání 6His-TEV proteasy (postTEV a preSAC) štěpí na fragment histidinové kotvy (menší než 10 kDa) a protein o velikosti cca 15 kDa, což odpovídá FD6A (15,485 kDa). Tento protein byl, na rozdíl od malého fragmentu a 6His-TEV proteasy, v roztoku zachován při průchodu afinitní maticí subtraktivní IMAC (postSAC, W1, W2). Zatímco první

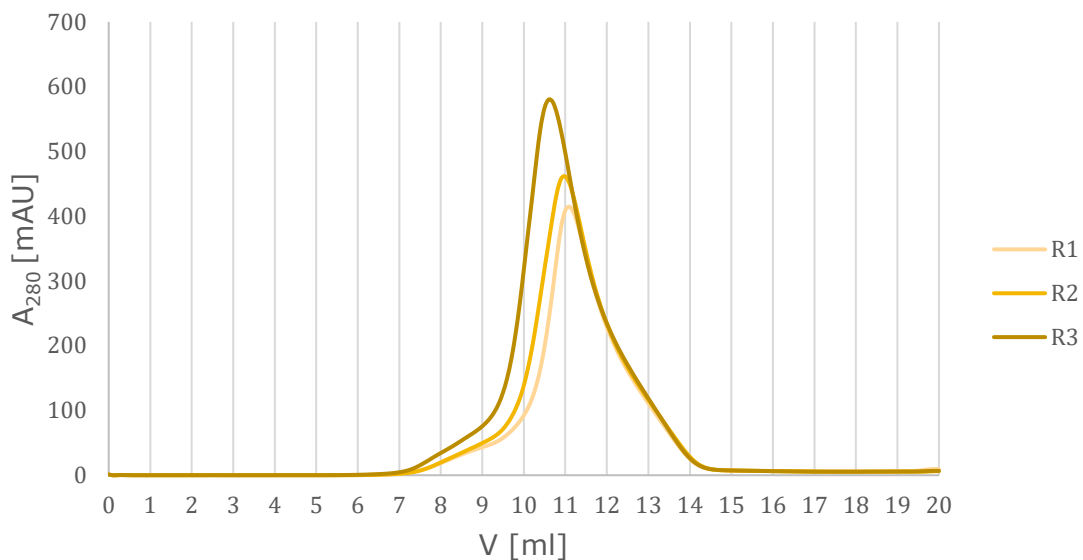
promývání (W1) bylo s to z afinitní matrice uvolnit významné množství proteinu, ve vzorku z druhého promytí (W2) byl protein už silně zředěn. Z časových důvodů nebyla tato frakce dále koncentrována, ačkoliv pro maximalizaci výtěžku by to samozřejmě představovalo optimální postup. Z gelu (Obr. 17) je rovněž patrné, že při koncentrování nedocházelo k průchodu kýženého proteinu membránou do filtrátu a že finální koncentrát obsahuje minimum kontaminací, potenciálně však marginální zbytek 6His-TEV proteasy (velikost cca 25 kDa); může však jít i o docela jiný protein podobné velikosti (srov. Obr. 16 a Obr. 17).



Obr. 17: SDS-PAGE procesu odštěpení histidinové kotvy FD6A pomocí TEV proteasy a následného koncentrování. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), následuje vzorek před přidáním TEV proteasy (preTEV), po inkubaci s tímto enzymem (postTEV), před provedením subtraktivní IMAC (preSAC), frakce nezadržená na afinitní matrici (postSAC), frakce vymytá z afinitní matrici 10 ml TEV pufru 2 (W1), frakce vymytá z afinitní matrici dalšími 40 ml TEV pufru 2 (W2), dále filtrát z koncentrátoru (Amicon FT) a koncentrované roztoky: 1 mg/ml (Final 1) a 2,5 mg/ml (Final 2).

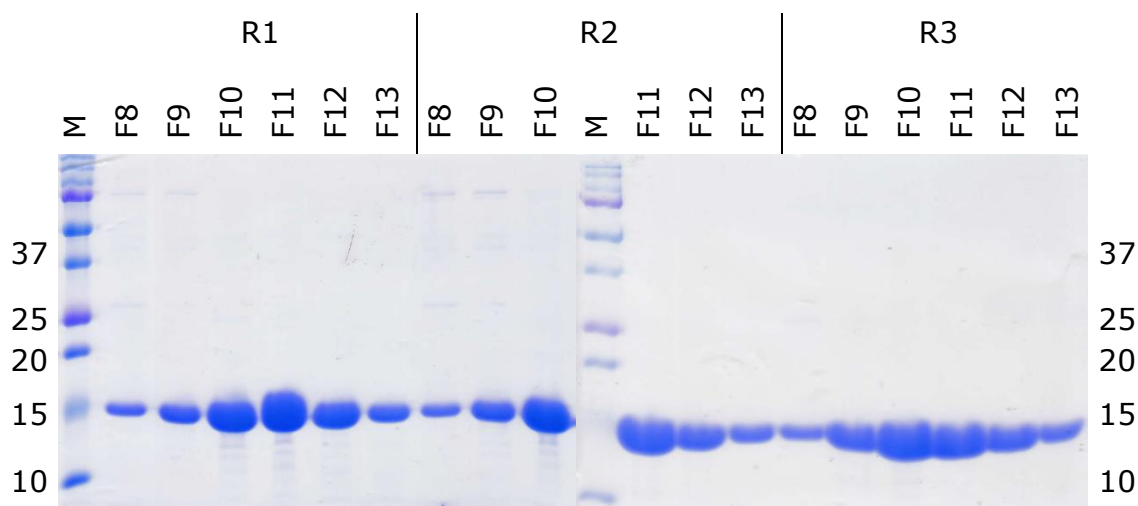
Při koncentrování v koncentrátorech Amicon byly roztoky pečlivě monitorovány a resuspendovány, podobně jako u FD5A, nicméně se zdálo, že FD6A efektu zvyšování viskozity akumulací proteinu v blízkosti membrány podléhá méně. Tuto autorovu observaci však nelze považovat za pravidlo, neboť existuje celá řada faktorů, které agregaci ovlivňují (jako přítomnost exogenních kontaminantů, teplota, úspěšnost samovolné renaturace atd.) (Wang a Roberts, 2018; Mahler et al., 2009; Kim et al., 1993), přičemž mezi vzorky FD5A a FD6A mohla existovat nezohledněná diference.

Protein byl koncentrován na 3×2 ml roztoku o koncentraci 4-7 mg/ml. Tyto vzorky byly postupně nadávkovány na FPLC (3 separace v módu SEC) kde byly děleny a převedeny do PB pufru. Dle získaných chromatogramů z detekce absorbance při 280 nm (viz *Obr. 18*, vyobrazeny relativní absorbance) byly v každé separaci selektovány frakce s vysokou koncentrací proteinu.



*Obr. 18: Chromatogramy SEC separací FD6A, tj. závislosti relativní absorbance při 280 nm na retenčním objemu pro tři separace (R1-R3).*

Tyto frakce byly podrobeny SDS-PAGE analýzám (viz *Obr. 19*), které ukázaly ve všech z nich protein o velikosti FD6A o velmi vysoké čistotě, srovnatelné s FD5A. I v tomto případě lze tvar neodpovídající Gaussovu rozložení vysvětlit nedostatečným rozlišením, vliv může mít rovněž vysoké množství silně koncentrovaného vzorku (Hagel, 2011).



Obr. 19: SDS-PAGE frakcí prvních separací FD6A pomocí SEC. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo a vpravo), následují frakce první (R1) separace: 8. až 13. frakce (F8-R13); pak frakce z druhé (R2) separace: 8. až 10. frakce (F8-F10), marker molekulových hmotností, frakce druhé separace (pokračování z předchozího gelu): 11. až 13. frakce (F11-R13); pak frakce z třetí (R3) separace: 8. až 13. frakce (F8-F13).

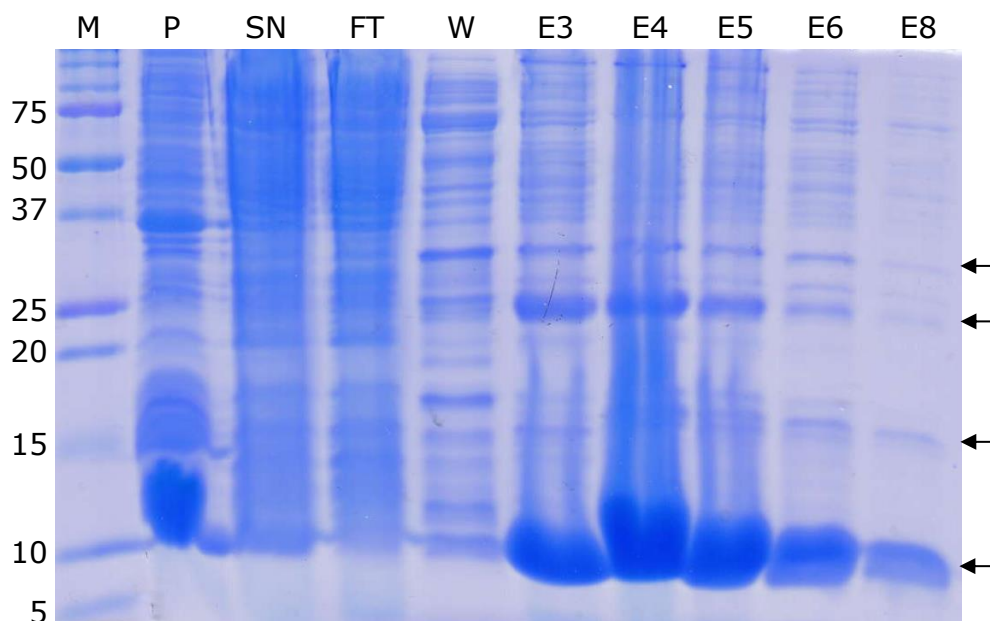
Celkově bylo ze všech tří separací získáno 8 ml kvalitních frakcí s koncentracemi v rozmezí od 90  $\mu\text{M}$  po 280  $\mu\text{M}$  se sumárním výtěžkem proteinu cca 22 mg. MS analýza potvrdila identitu proteinu při zjištění 100% shody se sekvencí FD6A (viz Příloha II). Získané frakce byly rozděleny na alikvoty, zamrazeny a sloužily dále jako zásobní vzorky pro biofyzikální charakterizace FD6A.

#### 3.4.4.3 Purifikace rekombinantního $G_{A88}$

Expres 6His- $G_{A88}$  v *E. coli*, izolace i IMAC probíhaly shodně s postupem uvedeným 3.3.3 a 3.4.4.1, rovněž SDS-PAGE byla provedena se shodně získanými vzorky z průběhu purifikace.

Na rozdíl od předešlých analytů, je u samostatné 6His- $G_{A88}$  z gelu (viz Obr. 20) patrné, že přechází z buněčného lyzátu do roztoku neochotně. V peletě (P) lze rozeznat velmi významný podíl proteinu odpovídajícímu velikostně 6His- $G_{A88}$  (10 kDa), zatímco v supernatantu je stejný protein zastoupen velmi slabě. Je nutné podotknout, že použitý purifikační protokol (Kirubakaran et al., 2016) byl vyvinut a optimalizován (Boušová et al., 2021) pro purifikaci fúzních proteinů s PDZ doménou a do této množiny  $G_{A88}$  nespadá. Použité inkubační podmínky, afinitní kotva, lyzační technika i složení pufrů tedy, jak je patrné, nejsou pro  $G_{A88}$  ideální volbou. Vzhledem k cílům této práce však musí být tato nedokonalost opomenuta.

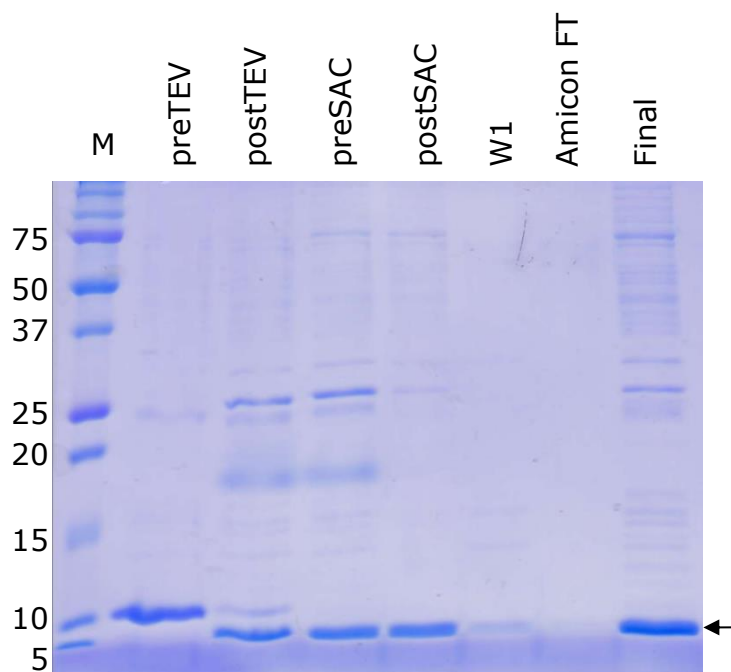
Ačkoliv se zdá, že supernatant (SN) obsahuje kýženého proteinu pouze nepatrné množství, v elučních frakcích jej lze pozorovat ve významné koncentraci. Patrné jsou však ale i kontaminanty (cca 15, 25, 30 kDa ad.).



Obr. 20: SDS-PAGE frakcí purifikačního procesu  $G_{A88}$ . Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), peleta lyzátu buněčné kultury (P), supernatant lyzátu (SN), frakce na IMAC nezachycených proteinů (FT), proteiny vymyté promývacím pufrem (W), 3. až 6. a 8. eluční frakce (E3-E6, E8).

Eluce byly sloučeny, 5× naředěny a denaturovány. Proces odštěpení histidinové kotvy byl vizualizován na SDS-PAGE pomocí průběžně odebíraných vzorků, shodně se štěpením FD6A (viz 3.4.4.2), vyjma W2. Elektroforetické analýze byl podroben také filtrát koncentrátoru Amicon (Amicon FT) a koncentrát získaného proteinu (Final).

Elektroforetický gel (viz Obr. 21) potvrzuje štěpení 6His- $G_{A88}$  TEV proteasou. Mezi vzorky před a po přidání TEV proteasy je patrné snížení velikosti z cca 10 kDa na cca 5 kDa, což odpovídá  $G_{A88}$  (6 877 kDa). Zatímco proužek 6His-TEV proteasy (cca 25 kDa) se ztrácí po subtraktivní IMAC, domnělý  $G_{A88}$  afinitní matricí zadržen nebyl (srov. preSAC, postSAC). Ačkoliv spadá  $G_{A88}$  velikostně pod MWCO použitého koncentrátoru (10 000 kDa), byl při daných otáčkách (2000-2500 RPM) filtrem zadržován (Amicon FT). Toto lze vysvětlit faktem, že reálné hodnoty MWCO se mohou pohybovat i výrazně pod hodnotami poskytovanými výrobcem (Kim et al., 1994; Hilal et al., 2007),  $G_{A88}$  by se však rovněž mohl vyskytovat v oligomerních formách. Koncentrát (Final) vykazoval přítomnost mnoha kontaminantů (cca 25, 75 kDa ad.) větších než  $G_{A88}$  v relativně vysokém poměru ke kýženému proteinu. Tento stav, který nebyl u FD5A ani FD6A pozorován, je přímým důsledkem nízkého výtěžku  $G_{A88}$  během izolace.

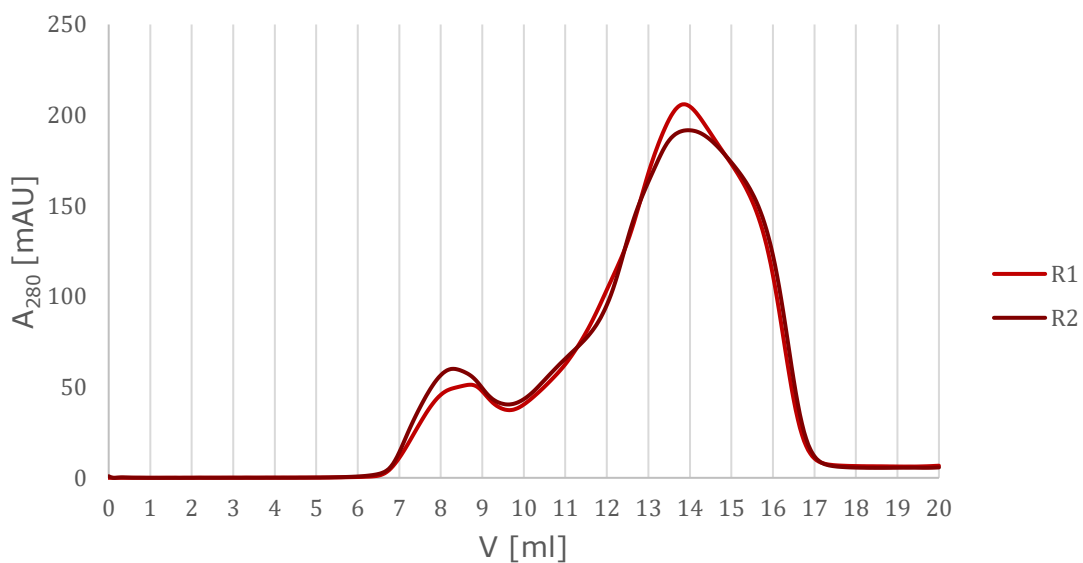


Obr. 21: SDS-PAGE procesu odštěpení histidinové kotvy FD6A pomocí TEV proteasy a následného koncentrování. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), následuje vzorek před přidáním TEV proteasy (preTEV), po inkubaci s tímto enzymem (postTEV), před provedením subtraktivní IMAC (preSAC), frakce nezadržená na afinitní matrici (postSAC), frakce vymytá z afinitní matrice 10 ml TEV pufru 2 (W1), dále filtrát z koncentrátoru (Amicon FT) a koncentrovaný roztoky o 1 mg/ml (Final).

Během koncentrování v koncentrátorech Amicon byly roztoky pečlivě monitorovány a resuspendovány, shodně jako u předešlých vzorků.  $G_{A88}$  v tomto ohledu vykazovalo výrazně horší vlastnosti než fúzní proteiny; zatímco FD5A i FD6A ochotně překonávaly hranici 3 mg/ml při stáčení na 2500 RPM s promícháním po 60 s,  $G_{A88}$  jen stěží při 2000 RPM s resuspendací á 30 s. V případě  $G_{A88}$  jsou *in vivo/in vitro* vlastnosti kompromisem se striktními kritérii, dle kterých byla této doména designována (viz 1.2), snížená rozpustnost a náchylnost k agregaci může proto být častým nežádoucím efektem podobných *in silico* designů (He et al., 2008).

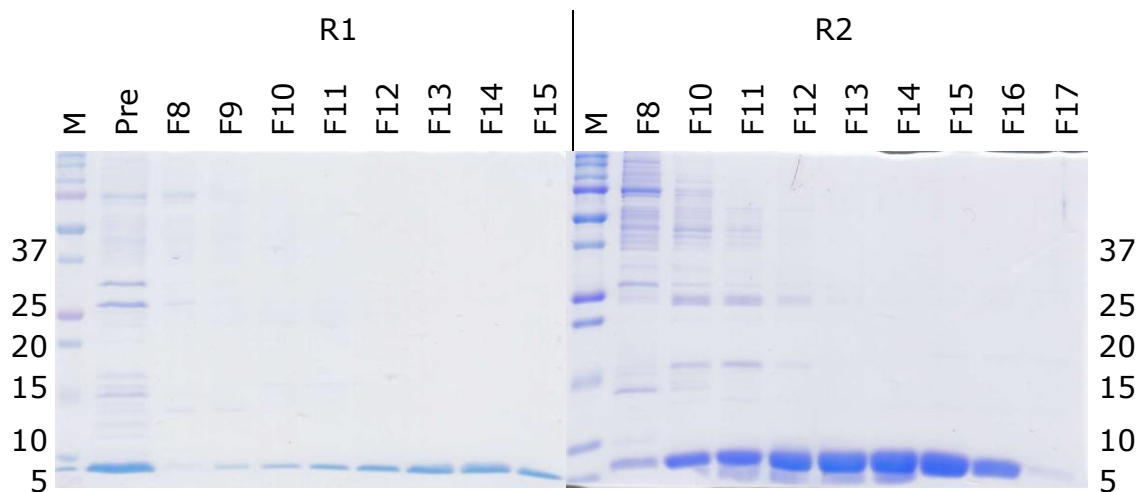
Protein byl koncentrován na 2×2 ml roztoku o koncentraci cca 3 mg/ml. Tyto vzorky byly postupně nadávkovány na FPLC (2 separace v módu SEC), kde byly děleny a převedeny do PB pufru. Dle získaných chromatogramů z detekce absorbance při 280 nm (viz Obr. 22, vyobrazeny relativní absorbance) byly v každé separaci selektovány frakce s vysokou koncentrací proteinu a ty podrobeny analýze pomocí SDS-PAGE.





Obr. 22: Chromatogramy SEC separací  $G_A88$ , tj. závislosti relativní absorbance při 280 nm na retenčním objemu pro obě separace (R1, R2).

Na elektroforetických gelech obou separací (viz Obr. 23) je viditelné v prvních frakcích (F8-F11) silné zastoupení kontaminantů, což koreluje s chromatogramy, kde lze pozorovat dobře rozlišený pík v této oblasti. Pozdější frakce (F13-F16) se zdají být prakticky čistý protein odpovídající velikostně  $G_A88$ , důležité je však podotknout, že pakliže by produkt obsahoval kontaminanty menší 5 kDa (např. fragmenty), pro jejich vizualizaci by bylo třeba užít gelu s vyšší koncentrací polyakrylamidu.



Obr. 23: SDS-PAGE frakcí prvních separací  $G_{A88}$  pomocí SEC. Zleva marker molekulových hmotností (M, jednotlivé hmotnosti v kDa jsou přiřazeny vlevo), následuje vzorek před SEC o koncentraci 1 mg/ml (Pre) a frakce první (R1) separace: 8. až 15. frakce (F8-R15), marker molekulových hmotností, následují frakce druhé (R2) separace: 8. a 10. až 17. frakce (F8, F10-R17).

Celkově bylo z obou separací získáno 10 ml kvalitních frakcí s koncentracemi v rozmezí od 25  $\mu$ M po 110  $\mu$ M se sumárním výtěžkem proteinu cca 5,5 mg. MS analýza potvrdila identitu proteinu při zjištění 100% shody se sekvencí  $G_{A88}$  (viz Příloha II). Získané frakce byly alikvotovány, zamrazeny a sloužily dále jako zásobní vzorky pro biofyzikální charakterizace  $G_{A88}$ .

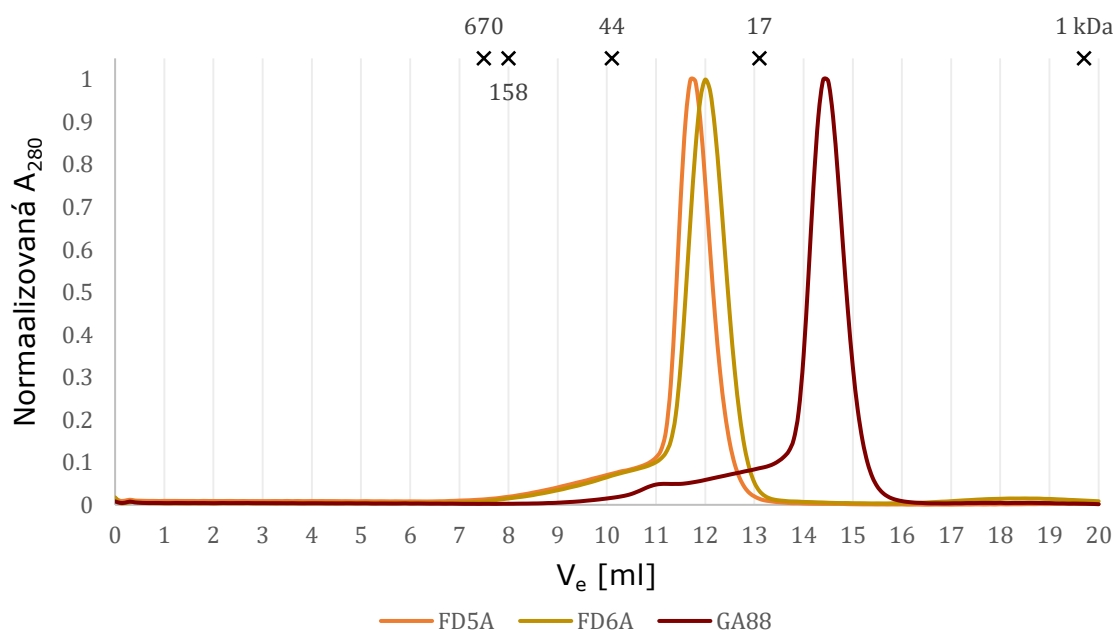
### 3.4.5 Biofyzikální charakterizace FD5A, FD6A, G_A88 a PDZ3

#### 3.4.5.1 Analytická SEC

Jak již bylo předesláno dříve (viz 3.4.4) metoda SEC se provádí pomocí FPLC instrumentace. Řadí se mezi kapalinové chromatografie, přičemž stacionární fázi tvoří často dextran zesíťovaný epichlorohydrinem (Separation Pharmacia Dextran – Sephadex) a mobilní fázi zpravidla vodné roztoky na bázi pufrů (při použití organických fází při zachování principu separace, např. pro separaci polymerů, se hovoří o gelové permeační chromatografii) (Hagel, 2011; Ó'Fágáin et al., 2011). Pro sledování průběhu separace může být použita celá řada detekčních metod. Mezi nejčastější patří UV spektrometrie (rovněž v této práci, zpravidla měření absorbance při 280 nm), IR spektrometrie, dynamický rozptyl světla, viskometrie a hmotnostní spektrometrie (Barth et al., 1998; Hagel, 2011).

Kromě použití SEC pro přečišťování (viz 3.4.4) a asistování skládání (refoldu) proteinů (Freydell et al., 2010) ji lze adoptovat i pro analytické účely, např. pro stanovování orientační molekulové hmotnosti (Andrews, 1964) a hydrodynamického poloměru (La Verde et al., 2017; Kunji et al., 2008) studovaného proteinu, ale také např. pro studium protein-protein interakcí (resp. druhého viriálního koeficientu  $B_2$  roztoku) (Bloustine et al., 2003) a proteinové agregace (Fekete et al., 2014).

Proteiny FD5A, FD6A a G_A88 byly podrobeny vylučovací chromatografii v nízkých koncentracích (30  $\mu$ M) za účelem porovnání s markerem molekulových hmotností (Gel Filtration Standard, Biorad). Data ze tří měření pro každý protein byla zprůměrována, normalizována a zanesena do grafu závislosti absorbance (280 nm) na elučním objemu (viz Obr. 24).



Obr. 24: Chromatogramy ASEC proteinů FD5A, FD6A a G_A88 (průměr ze tří měření pro každý z proteinů, normalizováno), včetně elučních objemů jednotlivých proteinů markeru molekulových hmotností (nad grafem, v kDa).

Přesné eluční objemy ( $V_e$ ), tj. maxima, pro jednotlivé proteiny byly stanoveny metodou první derivace (viz *Tabulka 17*).

*Tabulka 17: Vlastnosti studovaných proteinů zjištěné ASEC. Eluční objemy ( $V_e$ ), distribuční koeficienty ( $K_d$ ), rozdělovací koeficienty ( $K_{av}$ ), zdánlivé Stokesovy poloměry ( $R_S$ ) a molekulové hmotnosti globulárního ekvivalentu ( $M_{w,g}$ ) ( $R_S$  a  $M_{w,g}$  vypočtené z  $K_d$  a  $K_{av}$  byly shodné na uvedené dvě desetinná místa).*

Protein	$V_e$ ( $\sigma$ ) [ml]	$K_d$	$K_{av}$	$R_S$ [nm]	$M_{w,g}$ [kDa]
FD5A	11,747 (0,034)	0,348	0,257	2,59	29,37
FD6A	12,002 (0,128)	0,369	0,273	2,49	26,56
G _A 88	14,449 (0,025)	0,570	0,421	1,74	10,14

Za předpokladu, že největší z proteinů standardu (cca 670 kDa) byl eluován v mrtvém objemu kolony (v souladu s vlastnostmi kolony poskytovanými výrobcem) (Cytiva, 2023), zatímco nejmenší ze složek standardu (cca 1 kDa) představuje malou molekulu dobře zadržovanou ve stacionární fázi, lze určit distribuční koeficient ( $K_d$ ) jednotlivých proteinů (standard, vzorek) dle vzorce (La Verde et al., 2017):

$$K_d = \frac{V_e - V_0}{V_i}$$

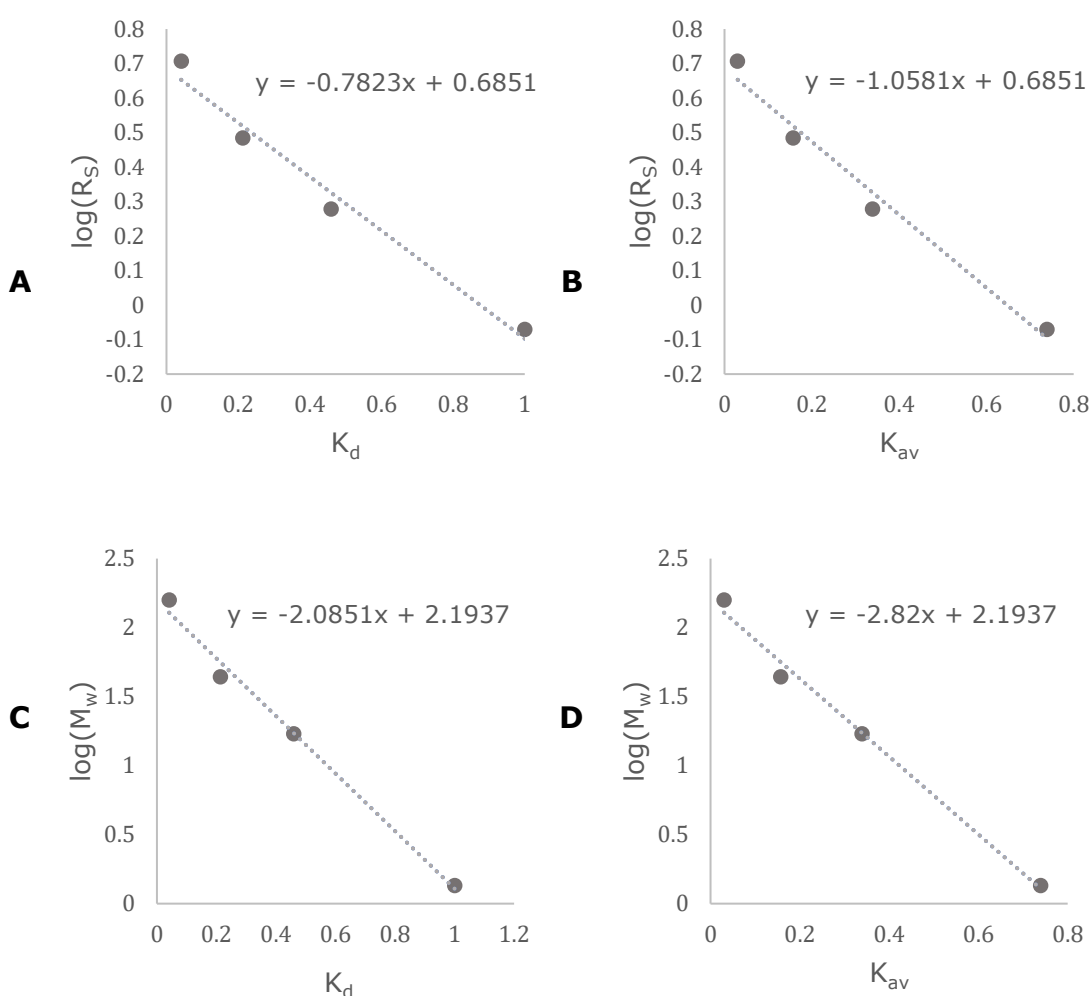
kde  $V_e$  představuje eluční objem daného proteinu,  $V_0$  mrtvý objem kolony (tj. eluční objem složky o velikosti 670 kDa, 7,5 ml) a  $V_i$  objem stacionární fáze (tj. eluční objem složky o velikosti 1 kDa, 19,7 ml, zmenšený o mrtvý objem). Paralelně byly rovněž vypočítány pro všechny proteiny rozdělovací koeficienty ( $K_{av}$ ) dle vzorce (La Verde et al., 2017):

$$K_{av} = \frac{V_e - V_0}{V_t - V_0}$$

kde  $V_t$  je celkový objem kolony daný výrobcem (24 ml) (Cytiva, 2023). Na základě známých Stokesových hydrodynamických poloměrů ( $R_S$ ) a molekulových hmotností jednotlivých proteinů standardu (La Verde et al., 2017; Talmard et al., 2007; Voisin et al., 1991; Erickson, 2009) (viz *Tabulka 18*) byly sestaveny kalibrační přímky (viz *Obr. 25*), tj. závislosti dekadického logaritmu  $R_S$  či  $M_w$  na  $K_d$ , resp.  $K_{av}$ . Dosazením do jejich lineárních regresí byly získány hodnoty zdánlivých Stokesových poloměrů a molekulových hmotností globulárních ekvivalentů (tj. globulárních proteinů o shodném zdánlivém Stokesově poloměru) FD5A, FD6A a G_A88 (viz *Tabulka 17*)

Tabulka 18: Vlastnosti standardu použitého pro výpočet Stokesových poloměrů studovaných proteinů. Molekulové hmotnosti složek ( $M_w$ ) a příslušné eluční objemy ( $V_e$ ), distribuční a rozdělovací koeficienty ( $K_d$ , resp.  $K_{av}$ ). Stokesovy poloměry převzaté z La Verde et al., 2017; Voisin et al., 1991, Erickson, 2009 a Talmard et al., 2007.

$M_w$ [kDa]	$V_e$ [ml]	$K_d$	$K_{av}$	$R_S$ [nm]
669	7,5	0	0	8,6
158	8,0	0,041	0,030	5,1
43	10,1	0,213	0,158	3,05
17	13,1	0,459	0,339	1,9
1,36	19,7	1	0,739	0,85



Obr. 25: Kalibrační přímky pro stanovení zdánlivého Stokesova poloměru a molekulové hmotnosti globulárního ekvivalentu FD5A, FD6A a  $G_{A88}$ . A: Závislost dekadického logaritmu Stokesova poloměru složek standardu na jejich distribučním koeficientu. B: Závislost dekadického logaritmu Stokesova poloměru složek standardu na jejich rozdělovacím koeficientu. C: Závislost dekadického logaritmu molekulové hmotnosti složek standardu na jejich distribučním koeficientu. D: Závislost dekadického logaritmu molekulové hmotnosti složek standardu na jejich rozdělovacím koeficientu.

Z analýz pomocí ASEC vyplývá, že fúzní proteiny FD5A a FD6A mají podobný zdánlivý Stokesův poloměr, cca 2,55 nm, což odpovídá globulárnímu proteinu o velikosti cca 27,5 kDa. Predikované velikosti FD5A a FD6A jsou však nižší (cca 16 kDa), lze tak usuzovat, že terciární struktura těchto fúzí je asymetrická, neboť v prostoru zaujmají z hydrodynamického hlediska stejný objem jako globulární proteiny o větší velikosti. Toto koresponduje s molekulárními modely představenými výše (viz *Obr. 9*), kde vykazují oba fúzní proteiny cylindrický tvar. Jiné vysvětlení nekonzistence mezi predikovanou velikostí proteinů a velikostí globulárních ekvivalentu by mohlo být, že se obě fúze vyskytují v dimerní podobě, více globulární než proteiny standardu. Tato možnost je však nepravděpodobná, neboť ani u  $G_A88$ , ani u PDZ3 ZO-1 nebyl pozorován proces dimerizace (Alexander et al., 2007; He et al., 2008; Pan et al., 2011; Itoh et al., 1999) a vznik takto kompaktního dimeru u proteinů s predikovaným asymetrickým tvarem se zpravidla neočekává.

Stejnou logikou lze interpretovat i výsledky pro  $G_A88$ . Na základě zdánlivého Stokesova poloměru (1,7 nm) a velikosti globulárního ekvivalentu (10 kDa, predikováno 6,9 kDa) lze usuzovat, že se jedná o menší, více globulární protein než studované fúze. Toto odpovídá struktuře  $G_A88$  (viz *Obr. 7*), která se jeví globulární. I v tomto případě jde zcela jistě o monomerní formu proteinu, v níž se  $G_A88$  vyskytuje podle literatury výhradně (Alexander et al., 2007; Alexander et al., 2009; He et al., 2008).

### 3.4.5.2 Dynamický rozptyl světla (DLS)

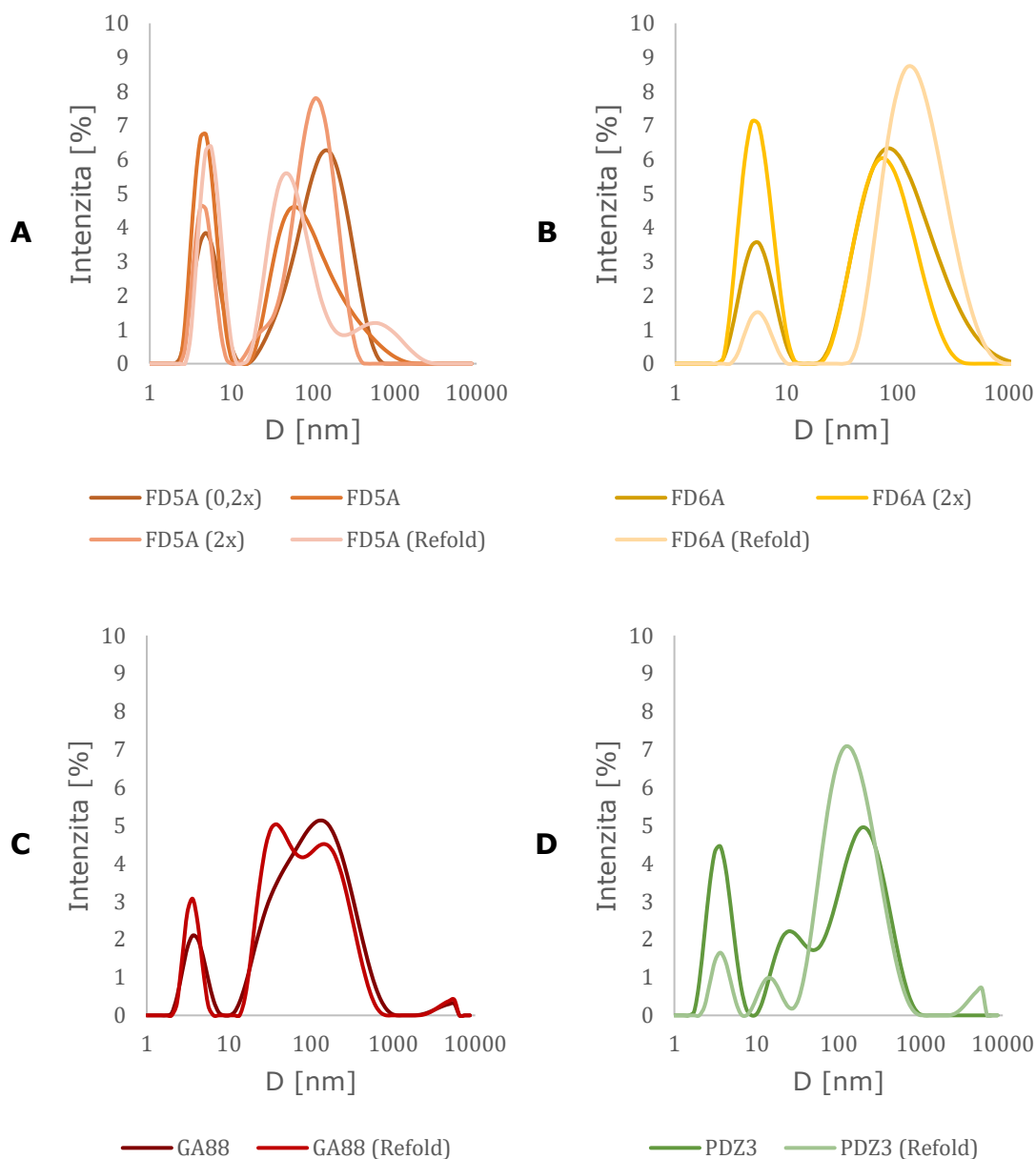
Dynamický rozptyl světla je metoda používaná pro charakterizaci proteinů a nanočástic založená na analýze rozptylu laserového paprsku. Tento rozptyl je zapříčiněn Brownovým pohybem částic v roztoku, závisí tak především na teplotě, viskozitě rozpouštědla a velikosti částice. Intenzitní fluktuace daná rozptylem je v čase zaznamenávána (tzv. autokorelační funkce) a následnými matematickými metodami převáděna na čtení ve vztahu k velikosti stanovované částice s využitím translačního difuzního koeficientu ( $D_t$ ) dle Stokesovy-Einsteinovy rovnice:

$$R_S = \frac{k_B T}{6\pi\eta_{20,w} D_t}$$

kde  $k_B = 1,380649 \cdot 10^{-23}$  J/K je Boltzmannova konstanta,  $\eta_{20,w} = 1,002$  cP viskozita vody při 20 °C a  $T = 293,5$  K je termodynamická teplota odpovídající 20 °C (Stetefeld et al., 2016).

Analýze metodou DLS byly podrobeny vzorky FD5A, FD6A, GA88 a PDZ3 v různých koncentracích. Automatickou transformací primárních dat, tj. autokorelačních funkcí (viz *Příloha III*), pomocí metody analýzy kumulantů dle Koppela (Koppel, 1972) byly získány intenzitní distribuce pro jednotlivá měření (viz *Obr. 26*). Byl proveden i pokus o zvýšení čistoty měřených vzorků pomocí denaturace (8M urea) a následné opětovné renaturace (tzv. refold), který se neprokázal jako úspěšný; dva dobře rozlišené píky, nativní formy a agregátu, byly přítomny i v takto upravených vzorcích.

Z intenzitních distribucí (tj. závislosti procentuální intenzity  $I$  na hydrodynamické velikosti  $D$ ) všech proteinů byly spočteny hydrodynamické (Stokesovy) poloměry metodou první derivace pro první z píků (viz *Tabulka 19*). Je nutno podotknout, že tyto poloměry jsou vzhledem k snížené přesnosti metody analýzy kumulantů pro polydisperzní systémy (tj. analýzy s neuniformní populací částic) spíše orientační a pro přesnější stanovení by musely být připraveny proteiny v lepší čistotě (Stetefeld et al., 2016).



Obr. 26: Intenzitní distribuce (analýza kumulantů autokorelační funkce) jednotlivých proteinů. A: FD5A, koncentrace 0,1 mg/ml (0,2×), 0,5 mg/ml, 1 mg/ml (2×) a refoldovaný 0,5 mg/ml při 20 °C. B: FD6A, koncentrace 0,5 mg/ml a refoldovaný 0,5 mg/ml při 20 °C a 1 mg/ml (2×) při 4 °C. C: GA88, koncentrace 0,5 mg/ml a refoldovaný 0,5 mg/ml při 20 °C. D: PDZ3, koncentrace 0,5 mg/ml a refoldovaný 0,5 mg/ml při 20 °C.

Primárním cílem tohoto stanovení bylo však zjištění čistoty proteinových vzorků. Z intenzitních distribucí je patrné, že ve všech vzorcích jsou přítomny alespoň dvě populace, menší z píků zaujímající zpravidla rozmezí 1-10 nm a následující druhý, široký pík, povětšinou pokrývající celý zbytek škály (10-1000 nm). Matematické vyjádření odchýlení od ideálního monodisperzního systému (s jedinou populací částice) popisuje tzv. relativní polydisperzita (%Pd), která se v monodisperzních systémech pohybuje v hodnotách do cca 20 % (Lorber et al., 2012); v předložených měřeních však zpravidla



přesahovala 90 %, což dokazuje polydisperzní charakter vzorků. Pro dvě ze složek takového systému platí Rayleighho aproximace:

$$I_a = \frac{100a^6 \cdot N_a}{a^6 \cdot N_a + b^6 \cdot N_b}$$

kde  $I_a$  je intenzitní distribuce složky A o hydrodynamickém průměru  $a$  přítomné v počtu  $N_a$  molekul a  $N_b$  je počet molekul složky B formy o velikosti  $b$  (Stetefeld et al., 2016). Pro stanovení čistoty (poměru nativní formy – A – vůči agregátu – B) lze rovnicí přepsat takto:

$$\frac{N_a}{N_b} = \frac{I_a \cdot b^6}{a^6 \cdot (100 - I_a)}$$

Aby bylo zajištěno, že vypočtená hodnota tohoto poměru bude nižší, než reálná hodnota, byly jako velikosti agregátu použity nejnižší hodnoty hydrodynamického průměru vykazující jakýkoliv intenzitní signál, které následovaly po píku nativní formy (viz *Tabulka 19*).

*Tabulka 19: Výsledky analýzy pomocí dynamického rozptylu světla. Uvedeny hydrodynamické poloměry (polovina naměřeného hydrodynamického průměru  $D$ , průměr) a minimální poměr nativní formy vůči agregátu vypočtený dle Rayleighho aproximace (pouze z původních, nikoliv z refoldovaných vzorků).*

	$R_S$ [nm]	Min. $N_a : N_b$
FD5A	2,39	~300 : 1 (99,71 %)
FD6A	2,87	~600 : 1 (99,84 %)
G _A 88	1,82	~180 : 1 (99,22 %)
PDZ3	1,77	~120 : 1 (99,48 %)

Pro další analýzu na AUC jsou vhodné pouze vzorky, které mají čistotu studované formy (tj. nativní) nad 95 % (Schuck a Braswell, 2000; Cole et al., 2008), což v tomto případě splňují FD5A, FD6A i G_A88 (PDZ3 rovněž, z praktických důvodů byl však použit jiný vzorek s histidinovou kotvou).

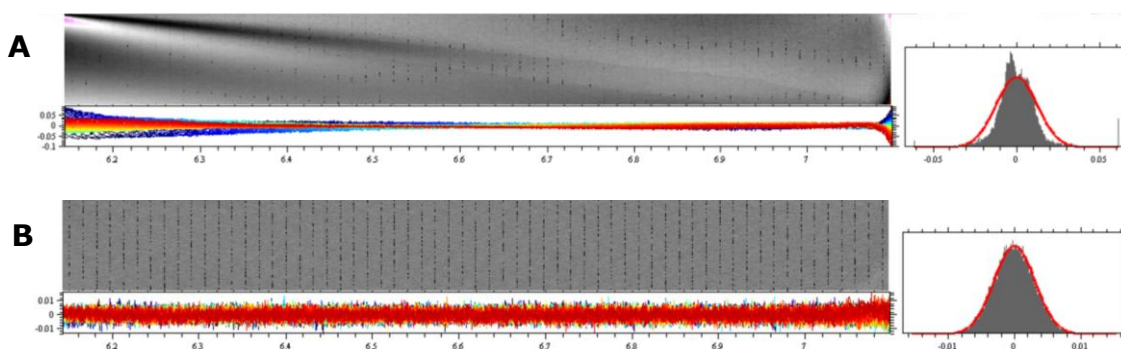
Ačkoliv v této práci bylo použito primárně pro ověření čistoty vzorků, potenciál DLS pro kvalitativní analýzu a charakterizace proteinů je významně vyšší (např. pro stanovení molekulové hmotnosti, pro studium termální stability, či studium skládání proteinů), muselo by však být adoptováno ve větším měřítku (DLS screeningy). Nedostatky této práce ve smyslu neúplného využití tohoto potenciálu by však mohly být podnětem pro navazující výzkum.

### 3.4.5.3 Analytická ultracentrifugace (AUC)

Analytická ultracentrifugace, příbuzná častěji využívanější separační centrifugaci, využívá prostup makromolekul roztokem zapříčiněný působením centrifugální síly při vysokých otáčkách k charakterizaci těchto makromolekul. K detekci prostupu jsou následně využívány spektroskopické metody (absorbance, fluorescence). AUC lze provádět dvěma metodami – analýzou sedimentační rychlosti (hydrodynamický přístup) a analýzou sedimentační rovnováhy (termodynamický přístup). Zatímco sedimentační rychlost je určována zpravidla při velmi vysokých otáčkách (15 000 až 60 000 RPM) v rámci hodin, sedimentační rovnováhu (bod, kdy je difuze perfektně vyvážena sedimentací) při nižších otáčkách (pod 15 000 RPM) v rámci několika dní. Obě lze dobře aplikovat pro stanovování molekulové hmotnosti analytů, analýza rovnováhy je však v principu v tomto ohledu přesnější, neboť čtení nezávisí na tvaru makromolekuly. Z analýzy sedimentační rychlosti je však možné získat řadu dalších podstatných hydrodynamických parametrů, především sedimentační koeficient, frikční poměr a Stokesův poloměr (Laue a Stafford III, 1999; Howlett et al., 2006; Lebowitz et al., 2009), proto byla použita v této práci.

Analýza sedimentační rychlosti pomocí AUC byla provedena pro FD5A, FD6A, G_A88 i PDZ3. Poslední zmíněný protein musel být použit s histidinovou značkou (6His-PDZ3), neboť sám o sobě nevykazuje absorpenci při 281 nm, která byla stanovována, z důvodu absence aromatických aminokyselin (viz *Tabulka 4*). Vzhledem k nutnosti dodržení doporučeného absorpčního rozsahu ( $A_{280} = 0,65$  v 1,2 cm kyvetě), musely být G_A88 i PDZ3 nanášeny v jiných koncentracích a absorbancích než FD5A a FD6A.

Byly provedeny analýzy při 48 000, 30 000 RPM při 20 °C a 30 000 a 15 000 RPM při 4 °C, přičemž nejlepší distribuce pro FD5A a FD6A byly pozorovány při 15 000 RPM a 30 000 RPM, pro G_A88 a 6His-PDZ3 při 30 000 RPM, všechny při 4 °C (posuzováno z tvaru absorpční křivky a reziduálního histogramu a bitmapy, viz *Obr. 27*).



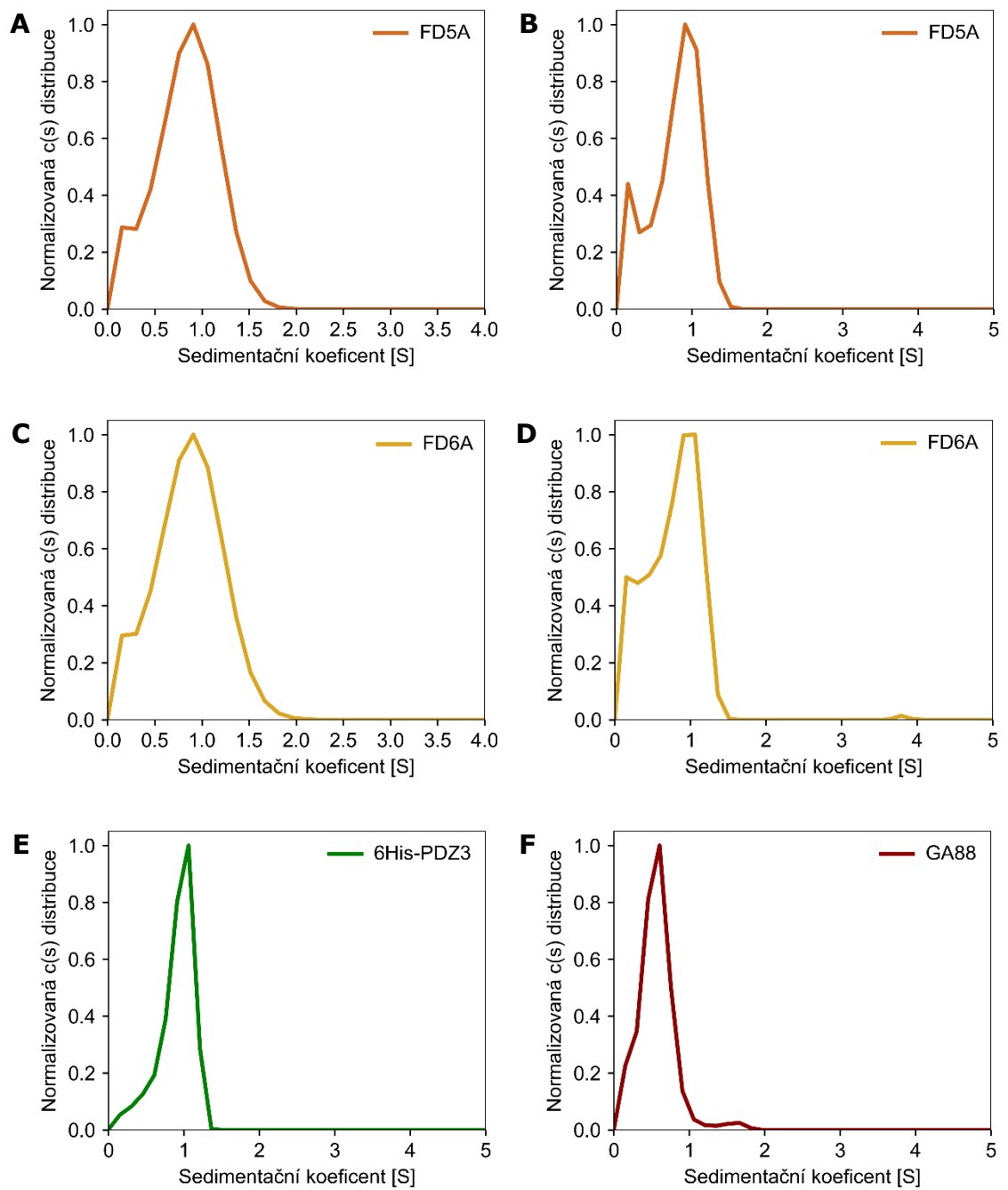
*Obr. 27: Příklady reziduálních bitmap a histogramů. A: Příklad vyřazeného měření (FD5A, 48 000 RPM, 20 °C). Histogram (vpravo) neodpovídá rozložení podle Gausse. Bitmapa (vlevo) vykazuje nerovnoměrné rozložení. B: Příklad použitého měření (FD5A, 15 000 RPM, 4 °C). Histogram (vpravo) odpovídá Gaussovu rozložení. Bitmapa (vlevo) ukazuje rozložení bez disturbancí. Zpracováno pomocí softwaru Sedfit, verze 15.01b (Schuck, 2000; Brown a Schuck, 2008).*

Pomocí softwaru Sedfit, verze 15.01b (Schuck, 2000; Brown a Schuck, 2008) byla data použita v modelu kontinuální  $c(s)$  distribuce, čímž byly stanoveny molekulové hmotnosti a frikční poměry ( $f/f_0$ ) (viz *Tabulka 20*) všech proteinů a rovněž vytvořeny grafy závislosti  $c(s)$  distribuce na sedimentačním koeficientu pro všechny proteiny (viz *Obr. 28* a *Obr. 29*).

*Tabulka 20: Sumarizace vlastností stanovených analýzou sedimentační rychlosti (4 °C, 15 000 nebo 30 000 RPM) pro jednotlivé proteiny. Experimentální molekulová hmotnost ( $M_w$ ), standardizovaný (voda, 20 °C) sedimentační koeficient ( $s_{20,w}$ ) a frikční poměr ( $f/f_0$ ) byly určeny pomocí softwaru Sedfit, verze 15.01b (Schuck, 2000; Brown a Schuck, 2008), Stokesův hydrodynamický poloměr ( $R_s$ ) pomocí softwaru Sednterp, verze 20130813 BETA (Laue et al., 1992).*

Protein	$\nu$ [RPM]	$M_w$ [kDa]	$f/f_0$	$s_{20,w}$ [S]	$R_s$ [nm]
FD5A	15 000	16,3	1,41	1,45	2,48
FD6A	15 000	15,3	1,40	1,42	2,41
FD5A	30 000	16,6	1,45	1,43	2,52
FD6A	30 000	15,3	1,42	1,41	2,46
G _A 88	30 000	7,5	1,37	0,96	1,53
6His-PDZ3	30 000	12,6	1,27	1,47	1,95

Pro FD5A byla v měřeních při 15 000 a 30 000 RPM stanovena molekulová hmotnost 16,3, resp. 16,6 kDa, což odpovídá vypočtené hodnotě (16,5 kDa). Pro FD6A byla stanovena molekulová hmotnost na 15,3 kDa v obou analýzách, což rovněž odpovídá predikované hodnotě (15,7 kDa). Pro G_A88 byla stanovena molekulová hmotnost 7,5 kDa, která rovněž relativně dobře odpovídá predikované velikosti monomerní formy (6,9 kDa) a pozorováním z analytické SEC (viz 3.4.5.1). Velikost 6His-PDZ3 byla stanovena na 12,6 kDa, což téměř dokonale odpovídá vypočtené hodnotě (12,5 kDa).



Obr. 28: Závislost normalizované  $c(s)$  distribuce proteinů na sedimentačním koeficientu. A: FD5A při 15 000 RPM, 4 °C. B: FD5A při 30 000 RPM, 4 °C. C: FD6A při 15 000 RPM, 4 °C. D: FD6A při 30 000 RPM, 4 °C. E: 6His-PDZ3 při 30 000 RPM, 4 °C. F: GA88 při 30 000 RPM, 4 °C. Zpracováno pomocí softwaru GUSI 1.4.2 (Brautigam, 2015).

Frikční poměry představují poměry frikčního koeficientu proteinu  $f$  a frikčního koeficientu  $f_0$  hypotetické, dokonale kulové, pevné, inertní částice při pohybu rozpouštědlem, který je daný vztahem odvozeným od Stokesova zákona:

$$f_0 = 6\pi\eta r, f = 6\pi\eta R_S$$

kde  $r$  je poloměr této hypotetické částice,  $\eta$  viskozita rozpouštědla a  $R_S$  Stokesův hydrodynamický poloměr proteinu (oficiální stránky softwaru Sedfit, [sedfitsedphat.github.io](http://sedfitsedphat.github.io), Cole et al., 2008). Vyjadřují tak odchyl od takové hypotetické částice (mající  $f/f_0 = 1$ ) způsobené asymetrií, vrásčitostí, deformací a hydratací proteinu, indikují tedy jeho tvar. Pro globulární proteiny se pohybují hodnoty frikčních poměrů kolem 1,05 až 1,3, pro fibrilární a nestrukturované proteiny přesahují 2 (Unzai, 2018; Garbett et al., 2010). Jediný ze studovaných proteinů spadajících do zmíněného rozmezí pro globulární proteiny je 6His-PDZ3, pro který byl stanoven frikční poměr na 1,27. Toto odpovídá krystalograficky determinované struktuře (Nomme et al., 2011), kde 6His-PDZ3 zaujímá značně kulovitou nativní konformaci. V literatuře nejsou frikční poměry PDZ domén extenzivně popsány, je však známo, že frikční poměr třetí PDZ domény whirlinu, samotné i v komplexu s PBM myosinu 15a, je 1,4 (Zhu et al., 2020). Je však nutné podotknout, že tato hodnota byla zjištěna v jiném pufru (Tris) a za jiné teploty (20 °C) i jiných otáček (42 000 RPM), proto je konkluzivní srovnání s 6His-PDZ3 velmi obtížné.

Frikční poměry FD5A a FD6A se od sebe liší minimálně, s  $f/f_0 \approx 1,43$  pro FD5A a  $f/f_0 \approx 1,41$  pro FD6A (průměry z měření při 15 000 a 30 000). Tyto hodnoty naznačují, že oba proteiny nabývají asymetrické prostorové konformace. Toto odpovídá predikovaným strukturám (viz Obr. 9), kde se oba fúzní proteiny jeví jako válcovité, a rovněž koresponduje s analytickou SEC, kde jsou FD5A i FD6A eluovány před globulárním standardem odpovídající velikosti (viz Obr. 24). Obdobný frikční poměr (1,37) vykazuje rovněž G_A88, rovněž molekula mírně podlouhlého tvaru dle NMR struktur (viz Obr. 7).

Integrací křivky  $c(s)$  distribuce v softwaru Sedfit, verze 15.01b (Schuck, 2000; Brown a Schuck, 2008) byly stanoveny sedimentační koeficienty standardizované pro vodu při 20 °C (viz Tabulka 20). Hodnoty sedimentačního koeficientu na rozdíl frikčních poměrů zohledňují také nerovnoměrnou hydrataci molekuly, dle Svedbergova vztahu:

$$s = \frac{M_w(1 - \bar{v}_t\rho_{t,b})}{N_A f}$$

kde  $s$  je sedimentační koeficient,  $M_w$  molekulová hmotnost,  $\bar{v}_t$  parciální specifický objem proteinu při teplotě měření,  $\rho$  hustota pufru při teplotě měření,  $N_A = 6,02214076 \cdot 10^{23} \text{ mol}^{-1}$  Avogadrova konstanta a  $f$  frikční koeficient proteinu. Dalším přepočtem lze sedimentační koeficient standardizovat pro vodné prostředí při 20 °C:

$$s_{20,w} = s \cdot \frac{1 - \bar{v}_t\rho_{t,b}}{1 - \bar{v}_{20}\rho_{20,w}} \cdot \frac{\eta_{20,w}}{\eta_{t,b}}$$

kde  $s_{20,w}$  je standardizovaný sedimentační koeficient,  $\eta_{t,b}$  viskozita pufru při teplotě měření a  $\rho_{20,w} = 0,998234 \text{ g/ml}$  hustota vody při 20 °C a  $\bar{v}_{20}$  parciální specifický objem proteinu při 20 °C (oficiální stránky softwaru Sednterp, [bitc.sr.unh.edu](http://bitc.sr.unh.edu)).

Pro 6His-PDZ3 byla vypočtena hodnota  $s_{20,w} \approx 1,47$  S. Toto odpovídá analogickým měřením v literatuře, kde se tyto koeficienty PDZ domén pohybují v obdobných hodnotách (viz *Tabulka 21*).

*Tabulka 21: Sedimentační koeficienty různých PDZ domén, porovnání. Původní data z některých zdrojů (*) přepočítána z  $s$  na  $s_{20,w}$  pomocí software Sednterp, verze 20130813 BETA (Laue et al., 1992) dle uvedených podmínek.*

Protein	Doména	$s_{20,w}$ [S]	Zdroj	Poznámka
ZO-1	PDZ3	1,47	Vlastní práce	N-6His kotva
PTPN3	PDZ	1,5	(Genera et al., 2019)	-
MLLT4	PDZ	1,5*	(Zhu et al., 2022)	-
MAST2	PDZ	1,59	(Delhommel et al., 2015)	V komplexu
Whirlin	PDZ3	1,3*	(Zhu et al., 2020)	Upravený C-konec
MAST2	PDZ	2,07	(Delhommel et al., 2015)	Dimer
ZO-1	PDZ2	2,17*	(Zhu et al., 2022)	Dimer

Standardizované sedimentační koeficienty FD5A a FD6A se od sebe významně neliší (1,45 resp. 1,42 S v analýze při 15 000 RPM a 1,43 resp. 1,47 S v analýze při 30 000 RPM), což v kontextu podobných frikčních poměrů naznačí analogické vlastnosti obou proteinů ve smyslu velikosti, tvaru i hydratace. Tuto observaci lze vysvětlit velmi podobným aminokyselinovým složením těchto proteinů, blízkostí C-konce a N-konce PDZ3 domény a rigiditou použitého linkeru, který může stabilizovat tvar obou proteinů. G_A88 má výrazně nižší sedimentační koeficient než předešlé proteiny (0,96 S), charakteristický pro proteiny o velikosti nižší než 10 kDa (Fredericq, 1957; Khan et al., 2013; Liu et al., 2006).

Finálním krokem bylo vypočtení Stokesova hydrodynamického poloměru ( $R_S$ ) pomocí softwaru Sednterp, verze 20130813 BETA (Laue et al., 1992). Hydrodynamický poloměr představuje poloměr hypotetické kulové částice o stejném frikčním koeficientu, jaký vykazuje protein. Lze jej získat ze sedimentačního koeficientu či standardizovaného sedimentačního koeficientu:

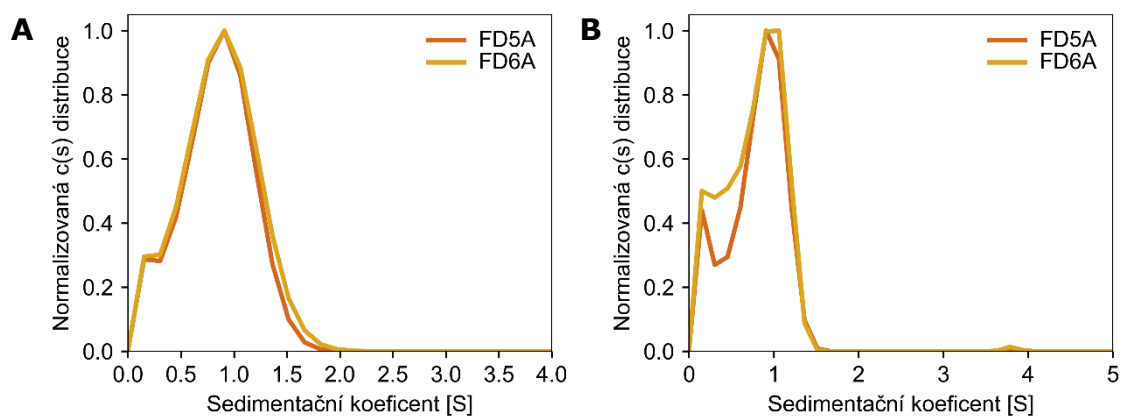
$$R_S = \frac{M_w(1 - \bar{v}_t \rho_{t,b})}{\eta_{t,b} s N_A} = \frac{M_w(1 - \bar{v}_{20} \rho_{20,w}) R_0}{N_A f_0 s_{20,w}}$$

kde  $R_0$  je tzv. ekvivalentní poloměr (poloměr bezvodé koule o objemu proteinu) daný  $\bar{v}_{20}$  a  $M_w$  (oficiální stránky softwaru Sednterp, [bitc.sr.unh.edu](http://bitc.sr.unh.edu)).

Stokesův poloměr 6His-PDZ3 byl stanoven na 1,95 nm. Toto rámcově odpovídá např. datům PDZ domény proteinu Shank (2,1 nm) (Im et al., 2010) či třetí PDZ Dlg4 (1,8 nm) (Salinas-García et al., 2023; Murciano-Calles et al., 2014).

Stokesovy poloměry FD5A (2,48 resp. 2,52 nm) a FD6A (2,41 resp. 2,46 nm) nebyly, podobně jako u frikčních poměrů i sedimentačních koeficientů významně rozdílné, konzistentně s observacemi z ASEC (viz *Tabulka 4*). Lze tedy vyvozovat, že tyto proteiny

jsou i z hlediska hydrodynamiky velmi podobné. Toto zjištění je v souladu s předchozími zjištěními o charakteru analogicky připravených fúzních proteinů s PDZ3 a TrpCage, FD3A a FD4A, které rovněž vykazovaly významnou schodu v těchto charakteristikách, lišily se však mírně flexibilitou proteinu (šířkou píku) a schopností dimerizace (Boušová et al., 2021). Podobné rozdíly nebyly u FD5A a FD6A pozorovány (viz Obr. 29), je však záhodno podotknout, že v této práci byly podrobeny analýze sedimentační rychlosti fúzní proteiny pouze o jediné koncentraci, 75  $\mu\text{M}$  (u FD4A byl pozorován významný podíl dimeru až při 60  $\mu\text{M}$ ) a že v FD5A a FD6A jsou domény propojeny rigidním linkerem (v FD3A a FD4A flexibilním) (Kirubakaran et al., 2016; Boušová et al., 2021).



Obr. 29: Proložení grafů analýz sedimentačních rychlostí FD5A a FD6A. A: Při 15 000 RPM, 4 °C. B: Při 30 000 RPM, 4 °C. Kromě artefaktů způsobených chybou fitu na začátku škály (B) nejsou patrné významné rozdíly.

Na základě experimentálních dat získaných sedimentační analýzou byly rovněž evaluovány modely FD5A a FD6A představené v 3.4.1 (viz Obr. 9), a to predikcí hydrodynamických veličin z těchto modelů *in silico* několika algoritmy – HYDROPRO, verze 10 (kompilace 54) (García De La Torre et al., 2000; Ortega et al., 2011), UltraScan III, verze 4.0 (6731) (Demeler, 2005) a HullRad, verze 9 (Fleming a Fleming, 2018) (princip algoritmů viz Příloha IV). Všechny stanovené hodnoty byly vloženy do souhrnné tabulky (viz Tabulka 22). Z tohoto porovnání se zdá, že ačkoliv FD6A patrně odpovídá predikovanému modelu (z hlediska hydrodynamických charakteristik), FD5A se od něj patrně slabě liší. Zdá se, že terciární struktura FD5A může být méně kompaktní, než naznačoval model, s protáhlým tvarem bližším FD6A.

Tabulka 22: Srovnání experimentálních hydrodynamických dat (Exp.) pro FD5A a FD6A z AUC, DLS a ASEC a hodnot predikovaných různými výpočetními nástroji za využití různých modelů a výpočetních metod (viz text, Příloha IV). Zpracováno na základě dat ze softwaru HYDROPRO, verze 10 (kompilace 54) (García De La Torre et al., 2000; Ortega et al., 2011), softwaru UltraScan III, verze 4.0 (6731) (Demeler, 2005) a online nástroje HullRad, verze 9 (Fleming a Fleming, 2018).

		$M_w$ [Da]	$f/f_0$	$s$ [S]	$s_{20,w}$ [S]	$R_s$ [nm]
FD5A	Exp. (AUC)	<b>16 300</b>	<b>1,41</b>	<b>0,89</b>	<b>1,45</b>	<b>2,48</b>
	(AUC)	<b>16 600</b>	<b>1,45</b>	<b>0,89</b>	<b>1,43</b>	<b>2,52</b>
	(DLS)	-	-	-	-	<b>2,39</b>
	(ASEC)	-	-	-	-	<b>2,59</b>
	UltraScan (ZENO)	16 484	1,31	-	1,59	2,23
	(GRPY)	16 484	1,34	-	1,55	2,29
	(SMI)	16 484	1,31	-	1,59	2,23
	(O-ZENO)	16 484	1,35	-	1,54	2,30
	(O-GRPY)	16 484	1,38	-	1,51	2,34
	HullRad	16 488	1,34	-	1,55	2,28
	HYDROPRO (1)	-	-	0,98	1,46	2,41
	(2)	-	-	1,02	1,50	2,35
	(4)	-	-	0,98	1,47	2,41
	FD6A	Exp. (AUC)	<b>15 300</b>	<b>1,40</b>	<b>0,86</b>	<b>1,42</b>
(AUC)		<b>15 300</b>	<b>1,42</b>	<b>0,86</b>	<b>1,41</b>	<b>2,46</b>
(DLS)		-	-	-	-	<b>2,87</b>
(ASEC)		-	-	-	-	<b>2,49</b>
UltraScan (ZENO)		15 674	1,42	-	1,41	2,39
(GRPY)		15 674	1,45	-	1,39	2,43
(SMI)		15 674	1,42	-	1,42	2,37
(O-ZENO)		15 674	1,45	-	1,38	2,43
(O-GRPY)		15 674	1,48	-	1,36	2,48
HullRad		15 679	1,37	-	1,47	2,29
HYDROPRO (1)		-	-	0,89	1,32	2,53
(2)		-	-	0,94	1,40	2,39
(4)		-	-	0,93	1,39	2,40

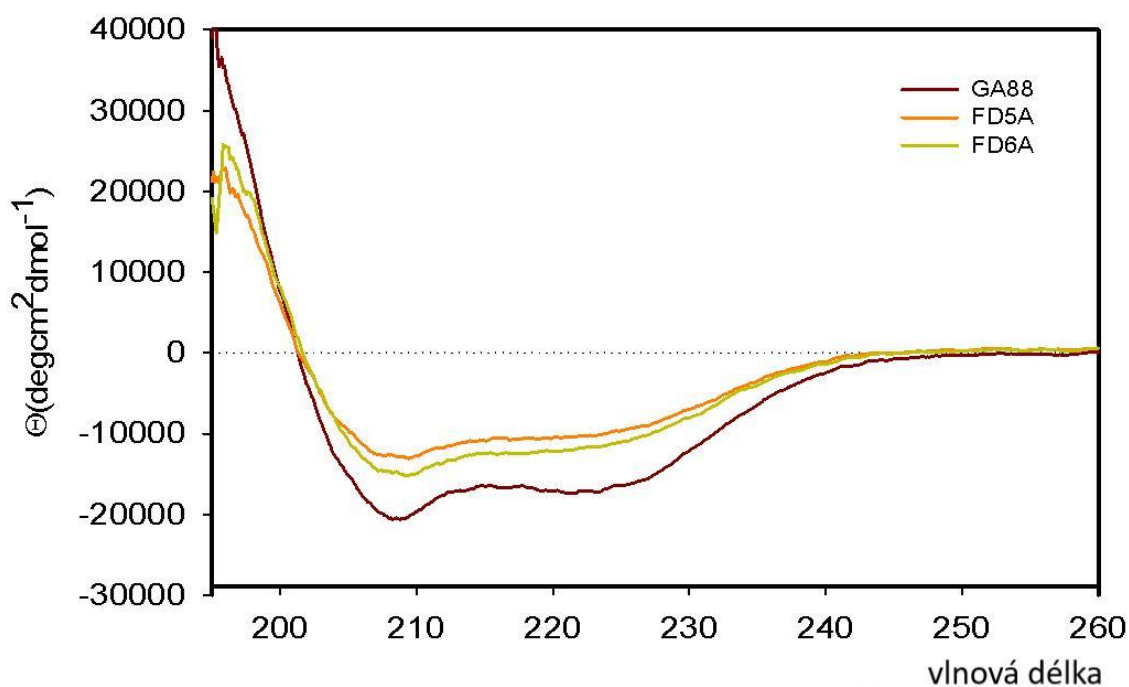


### 3.4.5.4 Cirkulární dichroismus (CD)

Spektrometrie cirkulárního dichroismu je spektrální metoda založená na rozdílné absorpci levotočivého a pravotočivého cirkulárně polarizovaného světla vzorkem (tzv. elipticita). Je využívána pro studium struktury proteinů, neboť každý typ sekundární struktury poskytuje charakteristický průběh CD spektra v oblasti dalekého UV (180-240 nm) díky absorbanci peptidových vazeb, zatímco terciární strukturu lze stanovovat pomocí elipticit v oblasti blízkého UV (260-320 nm) díky aromatickým aminokyselinám (Kelly a Price, 2000; Kelly et al., 2005).

Dalším využitím mohou být denaturační studie, kdy je sledována změna sekundární či terciární struktury se stoupající teplotou, nebo při postupném dávkování denaturačního činidla (např. močoviny). Tuto metodu lze použít např. za účelem optimalizace složení pufru tak, aby v něm měl protein vyšší stabilitu či pro stabilitní charakterizace a určení teploty tání (obdobně jako diferenciální skenovací fluorimetrie, viz 3.4.5.5) (Ranjbar a Gill, 2009).

Provedené měření (viz *Obr. 30*) ukazuje, že fúzní proteiny FD5A a FD6A vykazují CD spektra s negativním maximem při 209 nm, průběhem velmi podobná, což značí podobnost ve složení z hlediska sekundárních struktur. Odpovídá analogickým studiím fúzních proteinů FD3A a FD4A metodou CD v Boušová et al., 2021, kde se tyto dva rovněž mezi sebou odlišovaly pouze marginálně. Stanovené spektrum FD5A rovněž tvarem relativně dobře odpovídá předchozí publikované charakterizaci tohoto proteinu metodou CD, porovnání je však ztíženo faktem, že tato byla provedena v jiném pufru a s nižší koncentrací proteinu (Kirubakaran et al., 2016).



*Obr. 30: CD spektra (závislosti reziduální elipticity na vlnové délce) v daleké UV oblasti studovaných proteinů FD5A, FD6A a GA88.*

V kontextu dat z AUC, kde rovněž nebyl patrný významný rozdíl mezi FD5A a FD6A se tedy zdá, že oba proteiny jsou z hlediska tvaru a struktury extrémně podobné, bez předchozí znalosti prakticky nerozlišitelné.

Pro další evaluaci predikcí struktury FD5A a FD6A provedených pomocí nástroje ColabFold (Mirdita et al., 2022) byla z těchto struktur pomocí online nástroje PDBMD2CD (Drew a Janes, 2020) předpovídána očekávaná CD spektra (princip algoritmu viz Příloha IV) a následně byly kompozice sekundární struktury FD5A a FD6A predikované tímto nástrojem a experimentální data porovnána (viz Tabulka 23).

*Tabulka 23: Složení sekundární struktury studovaných proteinů zjištěné z CD spekter pomocí programu CPro (Exp.) a predikované pomocí online nástroje PDBMD2CD (Drew a Janes, 2020) na základě modelů struktur (Pred.).*

		$\alpha$ -helix [%]	$\beta$ -skládaný list [%]	otočka [%]	neuspořádané [%]
FD5A	Exp.	43,3	11,3	16,1	29,3
	Pred.	41,7	22,5	8,0	27,8
FD6A	Exp.	50,6	6,9	13,8	28,7
	Pred.	26,2	35,2	13,8	24,8
G _A 88	Exp.	65,7	4,6	8,5	21,2
	Pred.	64,3	0	7,1	28,6

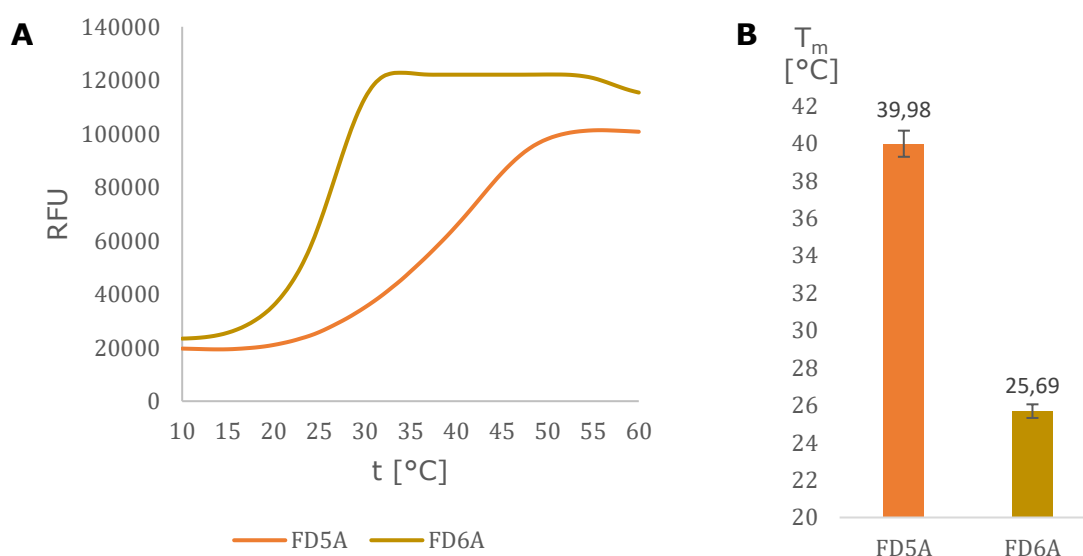
Před zabřednutím do extenzivních interpretací je nutné postulovat, že v tomto případě jsou predikovaná i experimentální data spíše orientačními hodnotami a jsou nejspíše zatíženy značnými chybami. V případě predikovaných hodnot je chybovost dána především nedokonalostí empirických algoritmů, které nástroj PDBMD2CD využívá, u experimentálních dat může chybu přinášet částečná denaturace a agregace (měření při 20 °C), interference složek pufru (např. soli) a potenciálních kontaminantů. Srovnáním hodnot pro G_A88 se však zdá, že se relativně dobře shodují, a to mezi sebou i s daty popsány v literatuře (Morrone et al., 2011).

Ani mezi experimentálními a predikovanými hodnotami FD5A není významný rozdíl, z hlediska sekundární struktury lze tedy model FD5A považovat za relativně věrný. FD6A se však od predikované struktury liší více, zejména v poměru  $\alpha$ -helixů a  $\beta$ -skládaných listů. Zdá se tak, že predikce softwaru ColabFold (Mirdita et al., 2022), že G_A88 v této fúzi nabírá konformaci G_B88, je s největší pravděpodobností mylná, snad důsledek nebývalé sekvenční podobnosti těchto dvou domén, kvůli níž je software není s to dobře rozlišit.

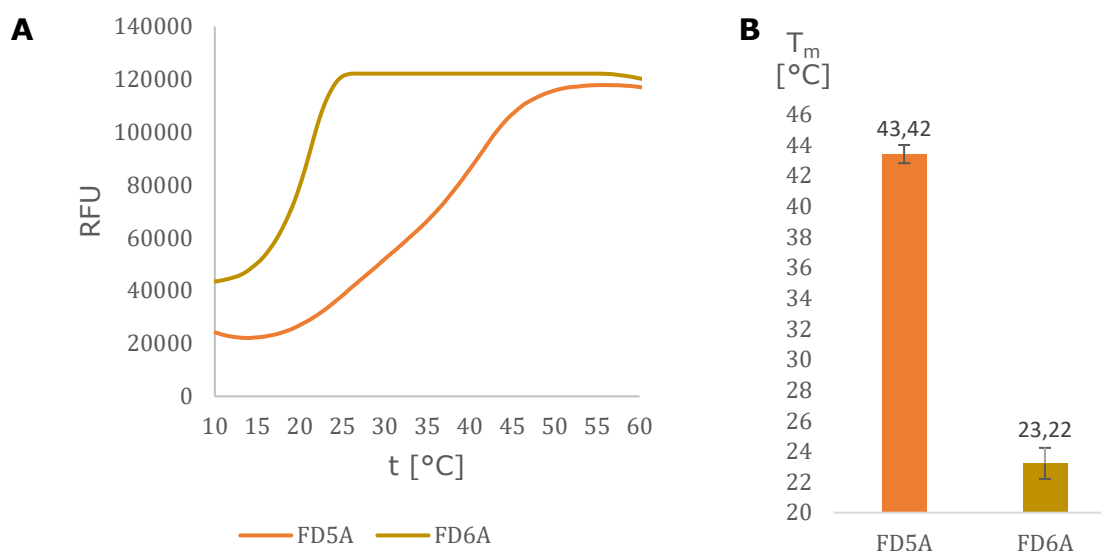
### 3.4.5.5 Diferenční skenovací fluorimetrie (DSF)

Diferenční skenovací fluorimetrie (také TSA – thermal shift assay) je metoda využívající qPCR instrumentaci ke studiu termální stability (stanovení teploty tání –  $T_m$ ) a detekci interakčních partnerů proteinů. Její výhodou je kromě nenáročné instrumentace také fakt, že na rozdíl od např. CD či izotermální titrační kalorimetrie (metody využívané pro podobné účely) je zde možnost screeningů mnoha vzorků v relativně krátkém časovém intervalu, ač na úkor přesnosti. Při DSF je sledována fluorescence lipofilního barviva (často SYPRO Orange s excitací při vlnové délce 488 nm a emisí v rozmezí 500-610 nm), která je ve vodě zhášena, ale stoupá v lipofilním prostředí reziduí denaturovaného proteinového hydrofobního jádra (Niesen et al., 2007; Gao et al., 2020; Santiveri et al., 2017).

Pro FD5A, FD6A v PB pufru byly optimální koncentrace proteinu a barviva SYPRO Orange pro další měření stanoveny na finální ředění SYPRO Orange 1:500 a koncentrace proteinu 20  $\mu\text{g/ml}$  optimalizačním měřením v několika kombinacích složení směsi (viz Příloha V). S tímto složením byly provedeny další analýzy (teplotní inkrement 0,2  $^{\circ}\text{C}$  za 45 s a 1  $^{\circ}\text{C}$  za 60 s), přičemž každý protein byl analyzován v oktaplikátu (viz Obr. 31 a Obr. 32). Teploty tání ( $T_m$ ) byly určeny pomocí online nástroje DSFworld (Wu et al., 2020), který využívá techniky prokládání experimentálních dat do empiricky sestavených modelů.



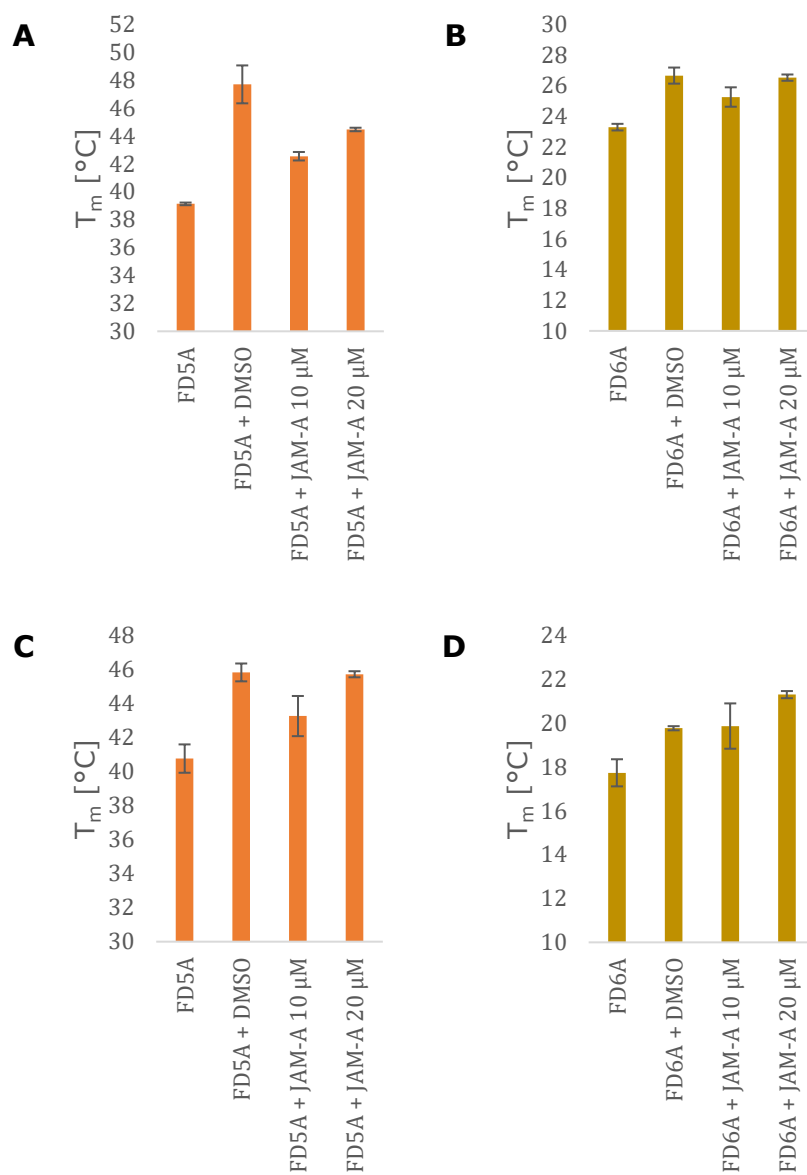
Obr. 31: Výsledky analýzy pomocí DSF (teplotní inkrement 1  $^{\circ}\text{C}$  za 60 s). A: Průměrné křivky tání FD5A a FD6A (20  $\mu\text{g/ml}$ ) v PB pufru. B: Porovnání teplot tání FD5A a FD6A s chybovými úsečkami.



Obr. 32: Výsledky analýzy pomocí DSF (teplotní inkrement 0,2 °C za 45 s.). A: Průměrné křivky tání FD5A a FD6A (20 µg/ml) v PB pufru. B: Porovnání průměrných teplot tání FD5A a FD6A s chybovými úsečkami.

Z výsledků analýzy vyplývá, že FD5A ( $T_m \approx 40\text{ °C}$ ) a FD6A ( $T_m \approx 25,7\text{ °C}$ ) mají velmi odlišnou teplotní stabilitu (v PB pufru), a to i přes velmi blízké složení ve smyslu terciární a sekundární struktury (viz 3.4.5.3 a 3.4.5.4). Toto pozorování koresponduje se stanoveními teplotní stability fúzních proteinů FD3A a FD4A (pomocí CD), které se od sebe rovněž z hlediska stability odlišovaly ( $T_m = 46$ , resp.  $59\text{ °C}$ ) (Boušová et al., 2021). Zdá se tak, že existuje určité mezidoménové ovlivnění ve studovaných proteinech FD5A a FD6A, ačkoli jeho podstata je zatím samozřejmě nejistá.

Zda tato mezidoménová interakce může ovlivnit funkci PDZ3 domény ve fúzích FD5A a FD6A zůstává rovněž otázkou, neboť následná stanovení metodou DSF (teplotní inkrement 0,2 °C za 45 s a 1 °C za 60 s) za přítomnosti PBM proteinu JAM-A (sekvence PSPSPSTSSFLV) (viz Obr. 33) neprokázala konzistentně významnou změnu teploty tání proteinů v přítomnosti tohoto potenciálního ligandu (viz Obr. 33). Zdá se však, že přítomnost 1 % DMSO v roztoku FD5A i FD6A silně stabilizuje. Záhodno by samozřejmě bylo provést rozsáhlejší sérii testů, ideálně v několika pufrch a za použití různých potenciálních ligandů (např. PBM Cx45, který vykazuje silnější interakci s PDZ3 než JAM-A) (Pan et al., 2011), popřípadě citlivějších metod pro studium protein-ligand interakcí jako izotermální titrační kalorimetrie (Santiveri et al., 2017).



Obr. 33: Průměrné teploty tání FD5A a FD6A (20  $\mu\text{g}/\text{ml}$ ) v PB pufru, s 1 % DMSO, s 10  $\mu\text{M}$  JAM-A a 1 % DMSO a s 20  $\mu\text{M}$  JAM-A a 1 % DMSO s chybovými úsečkami. A: FD5A, teplotní inkrement 1  $^{\circ}\text{C}$  za 60 s. B: FD6A, teplotní inkrement 1  $^{\circ}\text{C}$  za 60 s. C: FD5A, teplotní inkrement 0,2  $^{\circ}\text{C}$  za 45 s. D: FD6A, teplotní inkrement 0,2  $^{\circ}\text{C}$  za 45 s.

## Závěr

V rámci této práce byly připraveny geneticky modifikované bakterie *E. coli*, v nichž byly exprimovány proteiny FD5A, FD6A a G_A88. Tyto proteiny byly následně purifikovány pomocí afinitní chromatografie, jejich histidinové značky byly odstraněny pomocí štěpení TEV proteasou a subtraktivní afinitní chromatografie. Finální přečištění bylo provedeno pomocí SEC. Celý proces purifikace byl vizualizován pomocí SDS-PAGE a identita proteinů byla potvrzena pomocí MS. Fúze FD5A a FD6A jsou nové proteiny, přičemž FD6A nikdy předtím nebyl připraven, přesto jejich exprese a purifikace proběhly s vysokým výtěžkem bez nutnosti optimalizace protokolu.

Získané proteiny FD5A a FD6A byly následně poprvé charakterizovány pomocí několika metod s cílem jejich vzájemného srovnání. Zatímco z analýz metodami stanovujícími sekundární (CD) a terciární strukturu (ASEC, DLS, AUC) se zdálo, že oba proteiny jsou prakticky shodné, pomocí DSF bylo zjištěno, že z hlediska termální stability se FD5A a FD6A velmi významně liší. Toto naznačuje, že mezi doménami existuje ovlivnění, a to i přes použití rigidního linkeru, který by měl zamezit přímým interakcím mezi doménami.

Vhodné by tak bylo v rámci navazujícího výzkumu porovnat tyto výsledky s analogicky připravenými proteinovými fúzemi s např. flexibilním linkerem, G_B88 doménou, jinou PDZ doménou, či mutacemi v oblastech, které se pravděpodobně podílí na dynamické alosterii, stejně tak jako prozkoumat do detailů vzájemné ovlivnění vazebných vlastností obou domén pomocí dalších testů s JAM-A, popř. Cx45 na DSF či ITC a nalezení dalších, unikátních vazebných partnerů každé fúze, např. metodou pull-down. Paralelně s těmito testy by mělo nadále pokračovat i molekulární modelování těchto fúzí a hledání korelací mezi *in vitro* a *in silico* výsledky, které mohou přinést lepší porozumění dynamické alosterii a napomoci usnadnění úspěšného funkčního designu dalších fúzních proteinů pro praktická využití.

## Bibliografie

1. ABDUL-GADER, A. et al. A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics*. 2011, **27**(12), 1630–1636. DOI: 10.1093/bioinformatics/btr234.
2. ADIL, M. S. et al. Cell-cell junctions: structure and regulation in physiology and pathology. *Tissue Barriers*. 2021, **9**(1), 1848212. DOI: 10.1080/21688370.2020.1848212.
3. ALBERTS, B. et al. *Molecular biology of the cell*. 4th ed. New York : Garland Science, 2002. 1548 p. ISBN: 978-0-8153-3218-3.
4. ALEXANDER, P. A. et al. A minimal sequence code for switching protein structure and function. *Proceedings of the National Academy of Sciences*. 2009, **106**(50), 21149–21154. DOI: 10.1073/pnas.0906408106.
5. ALEXANDER, P. A. et al. The design and characterization of two proteins with 88% sequence identity but different structure and function. *Proceedings of the National Academy of Sciences*. 2007, **104**(29), 11963–11968. DOI: 10.1073/pnas.0700922104.
6. ALQURAISHI, M. AlphaFold at CASP13. VALENCIA, A. (ed.). *Bioinformatics*. 2019, **35**(22), 4862–4865. DOI: 10.1093/bioinformatics/btz422.
7. ANDREWS, P. Estimation of the molecular weights of proteins by Sephadex gel-filtration. *Biochemical Journal*. 1964, **91**(2), 222–233. DOI: 10.1042/bj0910222.
8. ARNAU, J. et al. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification*. 2006, **48**(1), 1–13. DOI: 10.1016/j.pep.2005.12.002.
9. ASSÉMAT, E. et al. Polarity complex proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2008, **1778**(3), 614–630. DOI: 10.1016/j.bbamem.2007.08.029.
10. BACART, J. et al. The BRET technology and its application to screening assays. *Biotechnology Journal*. 2008, **3**(3), 311–324. DOI: 10.1002/biot.200700222.
11. BAEK, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021, **373**(6557), 871–876. DOI: 10.1126/science.abj8754.
12. BANACH, M. et al. Structural Role of Hydrophobic Core in Proteins-Selected Examples. *Journal of Proteomics & Bioinformatics*. 2016, **9**(11). DOI: 10.4172/jpb.1000416.
13. BARTH, H. G. et al. Size Exclusion Chromatography and Related Separation Techniques. *Analytical Chemistry*. 1998, **70**(12), 251–278. DOI: 10.1021/a1980015t.
14. BAUER, H. et al. The Dual Role of Zonula Occludens (ZO) Proteins. *Journal of Biomedicine and Biotechnology*. 2010, **2010**, 1–11. DOI: 10.1155/2010/402593.
15. BELL, E. L. et al. Biocatalysis. *Nature Reviews Methods Primers*. 2021, **1**(1), 46. DOI: 10.1038/s43586-021-00044-z.
16. BERNIER, V. et al. Pharmacological chaperones: potential treatment for conformational diseases. *Trends in Endocrinology & Metabolism*. 2004, **15**(5), 222–228. DOI: 10.1016/j.tem.2004.05.003.
17. BERTHONNEAU, E. a MIRANDE, M. A gene fusion event in the evolution of aminoacyl-tRNA synthetases. *FEBS Letters*. 2000, **470**(3), 300–304. DOI: 10.1016/S0014-5793(00)01343-0.
18. BEWLEY, M. C. et al. A complex affair: Attraction and repulsion make occludin and ZO-1 function! *Tissue Barriers*. 2013, **1**(1), e23496. DOI: 10.4161/tisb.23496.

19. BHANOT, P. et al. A new member of the frizzled family from *Drosophila* functions as a Wingless receptor. *Nature*. 1996, **382**(6588), 225–230. DOI: 10.1038/382225a0.
20. BLANCO, F. J. et al. Exploring the conformational properties of the sequence space between two proteins with different folds: an experimental study 1 Edited by A. R. Fersht. *Journal of Molecular Biology*. 1999, **285**(2), 741–753. DOI: 10.1006/jmbi.1998.2333.
21. BLOCK, H. et al. Immobilized-Metal Affinity Chromatography (IMAC). In *Methods in Enzymology - Guide to Protein Purification*. 2nd. Elsevier, 2009 [cit. 2023-03-19], 439–473. DOI: 10.1016/S0076-6879(09)63027-5. ISBN: 978-0-12-374536-1.
22. BLOUSTINE, J. et al. Measurements of Protein-Protein Interactions by Size Exclusion Chromatography. *Biophysical Journal*. 2003, **85**(4), 2619–2623. DOI: 10.1016/S0006-3495(03)74684-0.
23. BORNSCHEUER, U. T. The fourth wave of biocatalysis is approaching. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2018, **376**(2110), 20170063. DOI: 10.1098/rsta.2017.0063.
24. BOUŠOVÁ, K. et al. The order of PDZ3 and TrpCage in fusion chimeras determines their properties - a biophysical characterization. *Protein Science*. 2021, **30**(8), 1653–1666. DOI: 10.1002/pro.4107.
25. BRAUTIGAM, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. In *Methods in Enzymology*. Elsevier, 2015 [cit. 2023-03-31], 109–133. DOI: 10.1016/bs.mie.2015.05.001. ISBN: 978-0-12-802908-4.
26. BROWN, P. H. a SCHUCK, P. A new adaptive grid-size algorithm for the simulation of sedimentation velocity profiles in analytical ultracentrifugation. *Computer Physics Communications*. 2008, **178**(2), 105–120. DOI: 10.1016/j.cpc.2007.08.012.
27. BRYAN, P. N. a ORBAN, J. Proteins that switch folds. *Current Opinion in Structural Biology*. 2010, **20**(4), 482–488. DOI: 10.1016/j.sbi.2010.06.002.
28. BURNS, D. M. et al. Evolution of the tryptophan synthetase of fungi. Analysis of experimentally fused *Escherichia coli* tryptophan synthetase alpha and beta chains. *Journal of Biological Chemistry*. 1990, **265**(4), 2060–2069. DOI: 10.1016/S0021-9258(19)39940-5.
29. CAI, C. et al. Selective Binding of Synapse-associated Protein 97 to GluR-A  $\alpha$ -Amino-5-hydroxy-3-methyl-4-isoxazole Propionate Receptor Subunit Is Determined by a Novel Sequence Motif. *Journal of Biological Chemistry*. 2002, **277**(35), 31484–31490. DOI: 10.1074/jbc.M204354200.
30. CARSON, M. et al. His-tag impact on structure. *Acta Crystallographica Section D Biological Crystallography*. 2007, **63**(3), 295–301. DOI: 10.1107/S0907444906052024.
31. CARUANA, G. a BERNSTEIN, A. Craniofacial Dysmorphogenesis Including Cleft Palate in Mice with an Insertional Mutation in the *discs large* Gene. *Molecular and Cellular Biology*. 2001, **21**(5), 1475–1483. DOI: 10.1128/MCB.21.5.1475-1483.2001.
32. CLARK, C. E. J. et al. The Tangled Web of Non-Canonical Wnt Signalling in Neural Migration. *Neurosignals*. 2012, **20**(3), 202–220. DOI: 10.1159/000332153.
33. CLEVERS, H. Wnt/ $\beta$ -Catenin Signaling in Development and Disease. *Cell*. 2006, **127**(3), 469–480. DOI: 10.1016/j.cell.2006.10.018.
34. COHEN, F. E. a KELLY, J. W. Therapeutic approaches to protein-misfolding diseases. *Nature*. 2003, **426**(6968), 905–909. DOI: 10.1038/nature02265.



35. COLE, J. L. et al. Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium. In *Methods in Cell Biology*. Elsevier, 2008 [cit. 2023-05-15], , 143–179. DOI: 10.1016/S0091-679X(07)84006-4. ISBN: 978-0-12-372520-2.
36. COTTRELL, J. S. Protein identification using MS/MS data. *Journal of Proteomics*. 2011, **74**(10), 1842–1851. DOI: 10.1016/j.jprot.2011.05.014.
37. CZAJKOWSKY, D. M. et al. Fc-fusion proteins: new developments and future perspectives. *EMBO Molecular Medicine*. 2012, **4**(10), 1015–1028. DOI: 10.1002/emmm.201201379.
38. DALAL, S. et al. Protein alchemy: Changing  $\beta$ -sheet into  $\alpha$ -helix. *Nature Structural Biology*. 1997b, **4**(7), 548–552. DOI: 10.1038/nsb0797-548.
39. DALAL, S. et al. Transmuting  $\alpha$  helices and  $\beta$  sheets. *Folding and Design*. 1997a, **2**(5), R71–R79. DOI: 10.1016/S1359-0278(97)00036-9.
40. DALAL, S. a REGAN, L. Understanding the sequence determinants of conformational switching using protein design. *Protein Science*. 2000, **9**(9), 1651–1659. DOI: 10.1110/ps.9.9.1651.
41. DELHOMMEL, F. et al. Deciphering the unconventional peptide binding to the PDZ domain of MAST2. *Biochemical Journal*. 2015, **469**(1), 159–168. DOI: 10.1042/BJ20141198.
42. DEMELER, B. UltraScan - A Comprehensive Data Analysis Software Package for Analytical Ultracentrifugation Experiments. In SCOTT, D. et al. (ed.). *Analytical Ultracentrifugation*. The Royal Society of Chemistry, 2005 [cit. 2023-05-12], 210–230. DOI: 10.1039/9781847552617-00210. ISBN: 978-0-85404-547-1.
43. DOI, N. a YANAGAWA, H. Insertional gene fusion technology. *FEBS Letters*. 1999, **457**(1), 1–4. DOI: 10.1016/S0014-5793(99)00991-6.
44. DREW, E. D. a JANES, R. W. PDBMD2CD: providing predicted protein circular dichroism spectra from multiple molecular dynamics-generated protein structures. *Nucleic Acids Research*. 2020, **48**(W1), W17–W24. DOI: 10.1093/nar/gkaa296.
45. DUEBER, J. E. et al. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology*. 2009, **27**(8), 753–759. DOI: 10.1038/nbt.1557.
46. ELIAS, G. M. et al. Differential trafficking of AMPA and NMDA receptors by SAP102 and PSD-95 underlies synapse development. *Proceedings of the National Academy of Sciences*. 2008, **105**(52), 20953–20958. DOI: 10.1073/pnas.0811025106.
47. ERICKSON, H. P. Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online*. 2009, **11**(1), 32–51. DOI: 10.1007/s12575-009-9008-x.
48. FABIAN, P. et al. Alternative Hydrophobic Core in Proteins—The Effect of Specific Synergy. *Symmetry*. 2020, **12**(2), 273. DOI: 10.3390/sym12020273.
49. FANNING, A. S. et al. Domain Swapping within PDZ2 Is Responsible for Dimerization of ZO Proteins. *Journal of Biological Chemistry*. 2007, **282**(52), 37710–37716. DOI: 10.1074/jbc.M707255200.
50. FANNING, A. S. et al. The Tight Junction Protein ZO-1 Establishes a Link between the Transmembrane Protein Occludin and the Actin Cytoskeleton. *Journal of Biological Chemistry*. 1998, **273**(45), 29745–29753. DOI: 10.1074/jbc.273.45.29745.
51. FEKETE, S. et al. Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *Journal of Pharmaceutical and Biomedical Analysis*. 2014, **101**, 161–173. DOI: 10.1016/j.jpba.2014.04.011.

52. FLEMING, P. J. a FLEMING, K. G. HullRad: Fast Calculations of Folded and Disordered Protein and Nucleic Acid Hydrodynamic Properties. *Biophysical Journal*. 2018, **114**(4), 856–869. DOI: 10.1016/j.bpj.2018.01.002.
53. FREDERICQ, E. The Molecular Weight of Insulin in Dioxane-Water Solutions. *Journal of the American Chemical Society*. 1957, **79**(3), 599–601. DOI: 10.1021/ja01560a026.
54. FREYDELL, E. J. et al. Size-exclusion chromatographic protein refolding: Fundamentals, modeling and operation. *Journal of Chromatography A*. 2010, **1217**(49), 7723–7737. DOI: 10.1016/j.chroma.2010.10.038.
55. GÁMEZ, A. et al. Protein misfolding diseases: Prospects of pharmacological treatment: GÁMEZ et al. *Clinical Genetics*. 2018, **93**(3), 450–458. DOI: 10.1111/cge.13088.
56. GAO, K. et al. Theory and applications of differential scanning fluorimetry in early-stage drug discovery. *Biophysical Reviews*. 2020, **12**(1), 85–104. DOI: 10.1007/s12551-020-00619-2.
57. GARBETT, N. C. et al. Sedimentation Velocity Ultracentrifugation Analysis for Hydrodynamic Characterization of G-Quadruplex Structures. In BAUMANN, P. (ed.). *G-Quadruplex DNA, Methods in Molecular Biology*. Totowa, NJ : Humana Press, 2010 [cit. 2023-05-02], 97–120. *Methods in Molecular Biology*. DOI: 10.1007/978-1-59745-363-9_7. ISBN: 978-1-58829-950-5.
58. GARCÍA DE LA TORRE, J. et al. Calculation of Hydrodynamic Properties of Globular Proteins from Their Atomic-Level Structure. *Biophysical Journal*. 2000, **78**(2), 719–730. DOI: 10.1016/S0006-3495(00)76630-6.
59. GARNER, C. C. et al. PDZ domains in synapse assembly and signalling. *Trends in Cell Biology*. 2000, **10**(7), 274–280. DOI: 10.1016/S0962-8924(00)01783-9.
60. GASTEIGER, E. et al. Protein Identification and Analysis Tools on the ExpASY Server. In WALKER, J. M. (ed.). *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press, 2005 [cit. 2023-06-03], 571–607. DOI: 10.1385/1-59259-890-0:571. DOI: 10.1385/1-59259-890-0:571. ISBN: 978-1-58829-343-5.
61. GENERA, M. et al. Structural and functional characterization of the PDZ domain of the human phosphatase PTPN3 and its interaction with the human papillomavirus E6 oncoprotein. *Scientific Reports*. 2019, **9**(1), 7438. DOI: 10.1038/s41598-019-43932-x.
62. GEVAERT, K. a VANDEKERCKHOVE, J. Protein identification methods in proteomics. *Electrophoresis*. 2000, **21**(6), 1145–1154. DOI: 10.1002/(SICI)1522-2683(20000401)21:6<1145::AID-ELPS1145>3.0.CO;2-Z.
63. GONG, Y. et al. Biochemical and biophysical analyses of tight junction permeability made of claudin-16 and claudin-19 dimerization. NUSRAT, A. (ed.). *Molecular Biology of the Cell*. 2015, **26**(24), 4333–4346. DOI: 10.1091/mbc.E15-06-0422.
64. GOPAL, G. J. a KUMAR, A. Strategies for the Production of Recombinant Protein in Escherichia coli. *The Protein Journal*. 2013, **32**(6), 419–425. DOI: 10.1007/s10930-013-9502-5.
65. GRUTSCH, S. et al. NMR Methods to Study Dynamic Allostery. LIU, J. (ed.). *PLOS Computational Biology*. 2016, **12**(3), e1004620. DOI: 10.1371/journal.pcbi.1004620.
66. HAGEL, L. Gel Filtration: Size Exclusion Chromatography. In JANSON, J.-C. (ed.). *Methods of Biochemical Analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2011 [cit. 2023-04-01],, 51–91. DOI: 10.1002/9780470939932.ch3. ISBN: 978-0-470-93993-2.
67. VAN HAM, M. PDZ domains-glye and guide. *Molecular Biology Reports*. 2003, **30**(2), 69–82. DOI: 10.1023/A:1023941703493.

68. HARRIS, B. Z. a LIM, W. A. Mechanism and role of PDZ domains in signaling complex assembly. *Journal of Cell Science*. 2001, **114**(18), 3219–3231. DOI: 10.1242/jcs.114.18.3219.
69. HARRISON, P. M. et al. Thermodynamics of model prions and its implications for the problem of prion protein folding 1 1Edited by A. R. Fersht. *Journal of Molecular Biology*. 1999, **286**(2), 593–606. DOI: 10.1006/jmbi.1998.2497.
70. HARTL, F. U. Protein Misfolding Diseases. *Annual Review of Biochemistry*. 2017, **86**(1), 21–26. DOI: 10.1146/annurev-biochem-061516-044518.
71. HATZAKIS, N. S. et al. Synthesis and single enzyme activity of a clicked lipase–BSA hetero-dimer. *Chem. Commun*. 2006(19), 2012–2014. DOI: 10.1039/B516551B.
72. HE, Y. et al. NMR structures of two designed proteins with high sequence identity but different fold and function. *Proceedings of the National Academy of Sciences*. 2008, **105**(38), 14412–14417. DOI: 10.1073/pnas.0805857105.
73. HEMSLEY, A. et al. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Research*. 1989, **17**(16), 6545–6551. DOI: 10.1093/nar/17.16.6545.
74. HILAL, N. et al. Characterization and retention of UF membranes using PEG, HS and polyelectrolytes. *Desalination*. 2007, **206**(1–3), 568–578. DOI: 10.1016/j.desal.2006.02.077.
75. HIRANO, Y. et al. Structure of a Cell Polarity Regulator, a Complex between Atypical PKC and Par6 PB1 Domains. *Journal of Biological Chemistry*. 2005, **280**(10), 9653–9661. DOI: 10.1074/jbc.M409823200.
76. HJERNØ, K. Protein Identification by Peptide Mass Fingerprinting. In RUNE, M. *Mass Spectrometry Data Analysis in Proteomics*. New Jersey : Humana Press, 2006 [cit. 2023-03-24], , 61–76. DOI: 10.1385/1-59745-275-0:61. DOI: 10.1385/1-59745-275-0:61. ISBN: 978-1-59745-275-5.
77. HOLLIDAY, M. J. et al. Networks of Dynamic Allostery Regulate Enzyme Function. *Structure*. 2017, **25**(2), 276–286. DOI: 10.1016/j.str.2016.12.003.
78. HOWLETT, G. et al. Analytical ultracentrifugation for the study of protein association and assembly. *Current Opinion in Chemical Biology*. 2006, **10**(5), 430–436. DOI: 10.1016/j.cbpa.2006.08.017.
79. HUNTER, A. W. et al. Zonula Occludens-1 Alters Connexin43 Gap Junction Size and Organization by Influencing Channel Accretion. *Molecular Biology of the Cell*. 2005, **16**(12), 5686–5698. DOI: 10.1091/mbc.e05-08-0737.
80. CHAN, W.-T. et al. A comparison and optimization of methods and factors affecting the transformation of *Escherichia coli*. *Bioscience Reports*. 2013, **33**(6), e00086. DOI: 10.1042/BSR20130098.
81. CHANG, B. H. et al. A Systematic Family-wide Investigation Reveals that ~30% of Mammalian PDZ Domains Engage in PDZ-PDZ Interactions. *Chemistry & Biology*. 2011, **18**(9), 1143–1152. DOI: 10.1016/j.chembiol.2011.06.013.
82. CHEN, J. et al. Domain-swapped dimerization of ZO-1 PDZ2 generates specific and regulatory connexin43-binding sites. *The EMBO Journal*. 2008b, **27**(15), 2113–2123. DOI: 10.1038/emboj.2008.138.
83. CHEN, L. et al. Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms. *Nature*. 2000, **408**(6815), 936–943. DOI: 10.1038/35050030.
84. CHEN, X. et al. Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews*. 2013, **65**(10), 1357–1369. DOI: 10.1016/j.addr.2012.09.039.

85. CHEN, Y. et al. Genome-wide Functional Annotation of Dual-Specificity Protein- and Lipid-Binding Modules that Regulate Protein Interactions. *Molecular Cell*. 2012, **46**(2), 226–237. DOI: 10.1016/j.molcel.2012.02.012.
86. IM, Y. J. et al. Crystal Structure of GRIP1 PDZ6-Peptide Complex Reveals the Structural Basis for Class II PDZ Target Recognition and PDZ Domain-mediated Multimerization. *Journal of Biological Chemistry*. 2003b, **278**(10), 8501–8507. DOI: 10.1074/jbc.M212263200.
87. IM, Y. J. et al. Crystal Structure of the Shank PDZ-Ligand Complex Reveals a Class I PDZ Interaction and a Novel PDZ-PDZ Dimerization. *Journal of Biological Chemistry*. 2003a, **278**(48), 48099–48104. DOI: 10.1074/jbc.M306919200.
88. IM, Y. J. et al. Structural Basis for Asymmetric Association of the  $\beta$ PIX Coiled Coil and Shank PDZ. *Journal of Molecular Biology*. 2010, **397**(2), 457–466. DOI: 10.1016/j.jmb.2010.01.048.
89. ITOH, M. et al. Direct Binding of Three Tight Junction-Associated MAGUKs, ZO-1, ZO-2, and ZO-3, with the COOH Termini of Claudins. *Journal of Cell Biology*. 1999, **147**(6), 1351–1363. DOI: 10.1083/jcb.147.6.1351.
90. ITOH, M. et al. Junctional adhesion molecule (JAM) binds to PAR-3. *Journal of Cell Biology*. 2001, **154**(3), 491–498. DOI: 10.1083/jcb.200103047.
91. ITURRATE, L. et al. Preparation and Characterization of a Bifunctional Aldolase/Kinase Enzyme: A More Efficient Biocatalyst for C–C Bond Formation. *Chemistry - A European Journal*. 2010, **16**(13), 4018–4030. DOI: 10.1002/chem.200903096.
92. IVARSSON, Y. et al. Prevalence, Specificity and Determinants of Lipid-Interacting PDZ Domains from an In-Cell Screen and In Vitro Binding Experiments. ARKOWITZ, R. A. (ed.). *PLoS ONE*. 2013, **8**(2), e54581. DOI: 10.1371/journal.pone.0054581.
93. JANIN, J. Welcome to CAPRI: A Critical Assessment of PRedicted Interactions. *Proteins: Structure, Function, and Genetics*. 2002, **47**(3), 257–257. DOI: 10.1002/prot.10111.
94. JONES, D. T. et al. Towards meeting the Paracelsus challenge: The design, synthesis, and characterization of paracelsin-43, an  $\alpha$ -helical protein with over 50% sequence identity to an all- $\beta$  protein. *Proteins: Structure, Function, and Genetics*. 1996, **24**(4), 502–513. DOI: 10.1002/(SICI)1097-0134(199604)24:4<502::AID-PROT9>3.0.CO;2-F.
95. JUMPER, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021, **596**(7873), 583–589. DOI: 10.1038/s41586-021-03819-2.
96. KANE, J. F. a HARTLEY, D. L. Formation of recombinant protein inclusion bodies in *Escherichia coli*. *Trends in Biotechnology*. 1988, **6**(5), 95–101. DOI: 10.1016/0167-7799(88)90065-0.
97. KAPUST, R. B. et al. The P1' specificity of tobacco etch virus protease. *Biochemical and Biophysical Research Communications*. 2002, **294**(5), 949–955. DOI: 10.1016/S0006-291X(02)00574-0.
98. KATSUNO, T. et al. Deficiency of Zonula Occludens-1 Causes Embryonic Lethal Phenotype Associated with Defected Yolk Sac Angiogenesis and Apoptosis of Embryonic Cells. MOSTOV, K. (ed.). *Molecular Biology of the Cell*. 2008, **19**(6), 2465–2475. DOI: 10.1091/mbc.e07-12-1215.
99. KELLY, S. M. et al. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. 2005, **1751**(2), 119–139. DOI: 10.1016/j.bbapap.2005.06.005.

100. KELLY, S. a PRICE, N. The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Current Protein & Peptide Science*. 2000, **1**(4), 349–384. DOI: 10.2174/1389203003381315.
101. KENNEDY, M. Origin of PDZ (DHR, GLGF) domains. *Trends in Biochemical Sciences*. 1995, **20**(9), 350. DOI: 10.1016/S0968-0004(00)89074-X.
102. KEUM, E. et al. Syndecan-4 regulates localization, activity and stability of protein kinase C-alpha. *Biochemical Journal*. 2004, **378**(3), 1007–1014. DOI: 10.1042/bj20031734.
103. KHAN, S. et al. The Solution Structure of Heparan Sulfate Differs from That of Heparin. *Journal of Biological Chemistry*. 2013, **288**(39), 27737–27751. DOI: 10.1074/jbc.M113.492223.
104. KIM, E. a SHENG, M. PDZ domain proteins of synapses. *Nature Reviews Neuroscience*. 2004, **5**(10), 771–781. DOI: 10.1038/nrn1517.
105. KIM, K. J. et al. A comparative study of techniques used for porous membrane characterization: pore characterization. *Journal of Membrane Science*. 1994, **87**(1–2), 35–46. DOI: 10.1016/0376-7388(93)E0044-E.
106. KIM, K. J. et al. Some factors determining protein aggregation during ultrafiltration. *Biotechnology and Bioengineering*. 1993, **42**(2), 260–265. DOI: 10.1002/bit.260420216.
107. KIRUBAKARAN, P. et al. Artificial proteins as allosteric modulators of PDZ3 and SH3 in two-domain constructs: A computational characterization of novel chimeric proteins: Chimeric Two-Domain Protein Design. *Proteins: Structure, Function, and Bioinformatics*. 2016, **84**(10), 1358–1374. DOI: 10.1002/prot.25082.
108. KLÖCKER, N. et al. Synaptic glutamate receptor clustering in mice lacking the SH3 and GK domains of SAP97: Synaptic glutamate receptors in SAP-97 mutant mice. *European Journal of Neuroscience*. 2002, **16**(8), 1517–1522. DOI: 10.1046/j.1460-9568.2002.02228.x.
109. KONTERMANN, R. E. Antibody–cytokine fusion proteins. *Archives of Biochemistry and Biophysics*. 2012, **526**(2), 194–205. DOI: 10.1016/j.abb.2012.03.001.
110. KOPPEL, D. E. Analysis of Macromolecular Polydispersity in Intensity Correlation Spectroscopy: The Method of Cumulants. *The Journal of Chemical Physics*. 1972, **57**(11), 4814–4820. DOI: 10.1063/1.1678153.
111. KUMAR, N. M. a GILULA, N. B. The Gap Junction Communication Channel. *Cell*. 1996, **84**(3), 381–388. DOI: 10.1016/S0092-8674(00)81282-9.
112. KUMAWAT, A. a CHAKRABARTY, S. Hidden electrostatic basis of dynamic allostery in a PDZ domain. *Proceedings of the National Academy of Sciences*. 2017, **114**(29). DOI: 10.1073/pnas.1705311114.
113. KUNJI, E. R. S. et al. Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography. *Methods*. 2008, **46**(2), 62–72. DOI: 10.1016/j.ymeth.2008.10.020.
114. LA VERDE, V. et al. Determination of Hydrodynamic Radius of Proteins by Size Exclusion Chromatography. *BIO-PROTOCOL*. 2017, **7**(8). DOI: 10.21769/BioProtoc.2230.
115. LACKS, S. a GREENBERG, B. Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation. *Journal of Molecular Biology*. 1977, **114**(1), 153–168. DOI: 10.1016/0022-2836(77)90289-3.

116. LAMAGNA, C. et al. Dual Interaction of JAM-C with JAM-B and  $\alpha_M \beta_2$  Integrin: Function in Junctional Complexes and Leukocyte Adhesion. *Molecular Biology of the Cell*. 2005, **16**(10), 4992–5003. DOI: 10.1091/mbc.e05-04-0310.
117. LATTMAN, E. E. a ROSE, G. D. Protein folding--what's the question? *Proceedings of the National Academy of Sciences*. 1993, **90**(2), 439–441. DOI: 10.1073/pnas.90.2.439.
118. LAUE, T. M. et al. Computer-aided Interpretation of Sedimentation Data for Proteins. In HARDING, S. E. et al. (ed.). *Analytical ultracentrifugation in biochemistry and polymer science*. Cambridge: Royal Society of Chemistry, 1992, , 90–125. ISBN: 0-85186-345-0.
119. LAUE, T. M. a STAFFORD III, W. F. MODERN APPLICATIONS OF ANALYTICAL ULTRACENTRIFUGATION. *Annual Review of Biophysics and Biomolecular Structure*. 1999, **28**(1), 75–100. DOI: 10.1146/annurev.biophys.28.1.75.
120. LEBOWITZ, J. et al. Modern analytical ultracentrifugation in protein science: A tutorial review. *Protein Science*. 2009, **11**(9), 2067–2079. DOI: 10.1110/ps.0207702.
121. LEE, H.-J. a ZHENG, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Communication and Signaling*. 2010, **8**(1), 8. DOI: 10.1186/1478-811X-8-8.
122. LENFANT, N. et al. A genome-wide study of PDZ-domain interactions in *C. elegans* reveals a high frequency of non-canonical binding. *BMC Genomics*. 2010, **11**(1), 671. DOI: 10.1186/1471-2164-11-671.
123. LEONARD, A. S. et al. SAP97 Is Associated with the  $\alpha$ -Amino-3-hydroxy-5-methylisoxazole-4-propionic Acid Receptor GluR1 Subunit. *Journal of Biological Chemistry*. 1998, **273**(31), 19518–19524. DOI: 10.1074/jbc.273.31.19518.
124. LEVARY, D. A. et al. Protein-Protein Fusion Catalyzed by Sortase A. NAJBAUER, J. (ed.). *PLoS ONE*. 2011, **6**(4), e18342. DOI: 10.1371/journal.pone.0018342.
125. LI, R.-W. et al. GRIP1 in GABAergic synapses. *The Journal of Comparative Neurology*. 2005, **488**(1), 11–27. DOI: 10.1002/cne.20566.
126. LI, X. et al. Association of connexin36 with zonula occludens-1 in HeLa cells, betaTC-3 cells, pancreas, and adrenal gland. *Histochemistry and Cell Biology*. 2004b, **122**(5), 485–498. DOI: 10.1007/s00418-004-0718-5.
127. LI, X. et al. Connexin47, connexin29 and connexin32 co-expression in oligodendrocytes and cx47 association with zonula occludens-1 (zo-1) in mouse brain. *Neuroscience*. 2004a, **126**(3), 611–630. DOI: 10.1016/j.neuroscience.2004.03.063.
128. LI, X. et al. Neuronal connexin36 association with zonula occludens-1 protein (ZO-1) in mouse brain and interaction with the first PDZ domain of ZO-1. *European Journal of Neuroscience*. 2004c, **19**(8), 2132–2146. DOI: 10.1111/j.0953-816X.2004.03283.x.
129. LICHTY, J. J. et al. Comparison of affinity tags for protein purification. *Protein Expression and Purification*. 2005, **41**(1), 98–105. DOI: 10.1016/j.pep.2005.01.019.
130. LIN, D. et al. A mammalian PAR-3–PAR-6 complex implicated in Cdc42/Rac1 and aPKC signalling and cell polarity. *Nature Cell Biology*. 2000, **2**(8), 540–547. DOI: 10.1038/35019582.
131. LIU, X. a FUENTES, E. J. Emerging Themes in PDZ Domain Signaling. In *International Review of Cell and Molecular Biology*. Elsevier, 2019 [cit. 2023-02-22], , 129–218. DOI: 10.1016/bs.ircmb.2018.05.013. ISBN: 978-0-12-816703-8.

132. LIU, Y. et al. The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity. *Cancer Cell*. 2006, **9**(4), 249–260. DOI: 10.1016/j.ccr.2006.03.012.
133. LONG, M. A New Function Evolved from Gene Fusion. *Genome Research*. 2000, **10**(11), 1655–1657. DOI: 10.1101/gr.165700.
134. LORBER, B. et al. Protein analysis by dynamic light scattering: Methods and techniques for students. *Biochemistry and Molecular Biology Education*. 2012, **40**(6), 372–382. DOI: 10.1002/bmb.20644.
135. MA, B. et al. Dynamic Allostery: Linkers Are Not Merely Flexible. *Structure*. 2011, **19**(7), 907–917. DOI: 10.1016/j.str.2011.06.002.
136. MA, T. M. et al. Serine Racemase Regulated by Binding to Stargazin and PSD-95. *Journal of Biological Chemistry*. 2014, **289**(43), 29631–29641. DOI: 10.1074/jbc.M114.571604.
137. MACKAY, D. J. G. a HALL, A. Rho GTPases. *Journal of Biological Chemistry*. 1998, **273**(33), 20685–20688. DOI: 10.1074/jbc.273.33.20685.
138. MADADLOU, A. et al. Fast Protein Liquid Chromatography. In WALLS, D.; a LOUGHRAN, S. T. (ed.). *Protein Chromatography, Methods in Molecular Biology*. Totowa, NJ : Humana Press, 2011 [cit. 2023-03-24], , 439–447. Methods in Molecular Biology. DOI: 10.1007/978-1-60761-913-0_25. ISBN: 978-1-60761-912-3.
139. MAFFEI, A. Long-Term Potentiation and Long-Term Depression. In MAFFEI, A. *Oxford Research Encyclopedia of Neuroscience*. Oxford University Press, 2018 [cit. 2023-03-04]. DOI: 10.1093/acrefore/9780190264086.013.148. ISBN: 978-0-19-026408-6.
140. MAHLER, H.-C. et al. Protein aggregation: Pathways, induction factors and analysis. *Journal of Pharmaceutical Sciences*. 2009, **98**(9), 2909–2934. DOI: 10.1002/jps.21566.
141. MALBON, C. C. a WANG, H. Dishevelled: A Mobile Scaffold Catalyzing Development. In *Current Topics in Developmental Biology*. Elsevier, 2005 [cit. 2023-03-27], , 153–166. DOI: 10.1016/S0070-2153(05)72002-0. ISBN: 978-0-12-153172-0.
142. MCINERNEY, P. et al. Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase. *Molecular Biology International*. 2014, **2014**, 1–8. DOI: 10.1155/2014/287430.
143. MEERSCHAERT, K. et al. The PDZ2 domain of zonula occludens-1 and -2 is a phosphoinositide binding domain. *Cellular and Molecular Life Sciences*. 2009, **66**(24), 3951–3966. DOI: 10.1007/s00018-009-0156-6.
144. MENG, D. Q. et al. Expression and role of cyclophilin B in stomach cancer. *Genetics and Molecular Research*. 2015, **14**(2), 5346–5354. DOI: 10.4238/2015.May.22.5.
145. MINAMIHATA, K. et al. Protein Heteroconjugation by the Peroxidase-Catalyzed Tyrosine Coupling Reaction. *Bioconjugate Chemistry*. 2011, **22**(11), 2332–2338. DOI: 10.1021/bc200420v.
146. MIRDITA, M. et al. ColabFold: making protein folding accessible to all. *Nature Methods*. 2022, **19**(6), 679–682. DOI: 10.1038/s41592-022-01488-1.
147. MISSLER, M. et al. The Making of Neurexins. *Journal of Neurochemistry*. 1998, **71**(4), 1339–1347. DOI: 10.1046/j.1471-4159.1998.71041339.x.
148. MORRONE, A. et al. The Denatured State Dictates the Topology of Two Proteins with Almost Identical Sequence but Different Native Structure and Function. *Journal of Biological Chemistry*. 2011, **286**(5), 3863–3872. DOI: 10.1074/jbc.M110.155911.

149. MOULT, J. A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction. *Current Opinion in Structural Biology*. 2005, **15**(3), 285–289. DOI: 10.1016/j.sbi.2005.05.011.
150. MOULT, J. et al. A large-scale experiment to assess protein structure prediction methods. *Proteins: Structure, Function, and Genetics*. 1995, **23**(3), ii–iv. DOI: 10.1002/prot.340230303.
151. MOYLE, P. M. Biotechnology approaches to produce potent, self-adjuvanting antigen-adjuvant fusion protein subunit vaccines. *Biotechnology Advances*. 2017, **35**(3), 375–389. DOI: 10.1016/j.biotechadv.2017.03.005.
152. MURCIANO-CALLES, J. et al. A thermodynamic study of the third PDZ domain of MAGUK neuronal protein PSD-95 reveals a complex three-state folding behavior. *Biophysical Chemistry*. 2014, **185**, 1–7. DOI: 10.1016/j.bpc.2013.10.005.
153. MURZIN, A. G. Metamorphic Proteins. *Science*. 2008, **320**(5884), 1725–1726. DOI: 10.1126/science.1158868.
154. NAYLOR, D. E. et al. Trafficking of GABA_A Receptors, Loss of Inhibition, and a Mechanism for Pharmacoresistance in Status Epilepticus. *The Journal of Neuroscience*. 2005, **25**(34), 7724–7733. DOI: 10.1523/JNEUROSCI.4944-04.2005.
155. NEZLIN, R. Interactions Outside the Antigen-Combining Site. In *The Immunoglobulins*. Elsevier, 1998 [cit. 2023-04-10], 219–cp1. DOI: 10.1016/B978-012517970-6/50006-0. ISBN: 978-0-12-517970-6.
156. NIELSEN, P. A. et al. Lens Connexins  $\alpha$ 3Cx46 and  $\alpha$ 8Cx50 Interact with Zonula Occludens Protein-1 (ZO-1). *Molecular Biology of the Cell*. 2003, **14**(6), 2470–2481. DOI: 10.1091/mbc.e02-10-0637.
157. NIESEN, F. H. et al. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*. 2007, **2**(9), 2212–2221. DOI: 10.1038/nprot.2007.321.
158. NILSSON, J. et al. Affinity Fusion Strategies for Detection, Purification, and Immobilization of Recombinant Proteins. *Protein Expression and Purification*. 1997, **11**(1), 1–16. DOI: 10.1006/prev.1997.0767.
159. NOÉ, F. et al. Machine learning for protein folding and dynamics. *Current Opinion in Structural Biology*. 2020, **60**, 77–84. DOI: 10.1016/j.sbi.2019.12.005.
160. NOMME, J. et al. The Src Homology 3 Domain Is Required for Junctional Adhesion Molecule Binding to the Third PDZ Domain of the Scaffolding Protein ZO-1. *Journal of Biological Chemistry*. 2011, **286**(50), 43352–43360. DOI: 10.1074/jbc.M111.304089.
161. NUNBHAKDI-CRAIG, V. et al. Protein phosphatase 2A associates with and regulates atypical PKC and the epithelial tight junction complex. *Journal of Cell Biology*. 2002, **158**(5), 967–978. DOI: 10.1083/jcb.200206114.
162. NUSSE, R. Wnt Signaling. *Cold Spring Harbor Perspectives in Biology*. 2012, **4**(5), a011163–a011163. DOI: 10.1101/cshperspect.a011163.
163. Ó'FÁGÁIN, C. et al. Gel-Filtration Chromatography. In WALLS, D.; a LOUGHRAN, S. T. (ed.). *Protein Chromatography, Methods in Molecular Biology*. Totowa, NJ : Humana Press, 2011 [cit. 2023-07-04], , 25–33. Methods in Molecular Biology. DOI: 10.1007/978-1-60761-913-0_2. ISBN: 978-1-60761-912-3.
164. ORTEGA, A. et al. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophysical Journal*. 2011, **101**(4), 892–898. DOI: 10.1016/j.bpj.2011.06.046.



165. OTANI, T. a FURUSE, M. Tight Junction Structure and Function Revisited. *Trends in Cell Biology*. 2020, **30**(10), 805–817. DOI: 10.1016/j.tcb.2020.08.004.
166. PAN, L. et al. The Structure of the PDZ3-SH3-GuK Tandem of ZO-1 Protein Suggests a Supramodular Organization of the Membrane-associated Guanylate Kinase (MAGUK) Family Scaffold Protein Core. *Journal of Biological Chemistry*. 2011, **286**(46), 40069–40074. DOI: 10.1074/jbc.C111.293084.
167. PARIS, L. et al. Structural organization of the tight junctions. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2008, **1778**(3), 646–659. DOI: 10.1016/j.bbamem.2007.08.004.
168. PATEL, D. K. et al. Linkers: A synergistic way for the synthesis of chimeric proteins. *Protein Expression and Purification*. 2022, **191**, 106012. DOI: 10.1016/j.pep.2021.106012.
169. PAYSAN-LAFOSSE, T. et al. InterPro in 2022. *Nucleic Acids Research*. 2022, **51**(D1), D418–D427. DOI: 10.1093/nar/gkac993.
170. PENKERT, R. R. et al. Internal recognition through PDZ domain plasticity in the Par-6–Pals1 complex. *Nature Structural & Molecular Biology*. 2004, **11**(11), 1122–1127. DOI: 10.1038/nsmb839.
171. PÉREZ-ARELLANO, I. et al. Pyrroline-5-carboxylate synthase and proline biosynthesis: From osmotolerance to rare metabolic disease: P5CS and Proline Biosynthesis. *Protein Science*. 2010, **19**(3), 372–382. DOI: 10.1002/pro.340.
172. PERROY, J. et al. PICK1 is required for the control of synaptic transmission by the metabotropic glutamate receptor 7. *The EMBO Journal*. 2002, **21**(12), 2990–2999. DOI: 10.1093/emboj/cdf313.
173. PETIT, C. M. et al. Hidden dynamic allostery in a PDZ domain. *Proceedings of the National Academy of Sciences*. 2009, **106**(43), 18249–18254. DOI: 10.1073/pnas.0904492106.
174. POLLOK, B. Using GFP in FRET-based applications. *Trends in Cell Biology*. 1999, **9**(2), 57–60. DOI: 10.1016/S0962-8924(98)01434-2.
175. PONTING, C. P. Evidence for PDZ domains in bacteria, yeast, and plants: Novel PDZ domains. *Protein Science*. 1997, **6**(2), 464–468. DOI: 10.1002/pro.5560060225.
176. PONTING, C. P. a RUSSELL, R. R. The Natural History of Protein Domains. *Annual Review of Biophysics and Biomolecular Structure*. 2002, **31**(1), 45–71. DOI: 10.1146/annurev.biophys.31.082901.134314.
177. POOLE, E. S. et al. The identity of the base following the stop codon determines the efficiency of in vivo translational termination in Escherichia coli. *The EMBO Journal*. 1995, **14**(1), 151–158. DOI: 10.1002/j.1460-2075.1995.tb06985.x.
178. POTAPOV, V. a ONG, J. L. Examining Sources of Error in PCR by Single-Molecule Sequencing. KALENDAR, R. (ed.). *PLOS ONE*. 2017, **12**(1), e0169774. DOI: 10.1371/journal.pone.0169774.
179. RAHMAN, M. T. et al. Polymerase Chain Reaction (PCR): A Short Review. *Anwer Khan Modern Medical College Journal*. 2013, **4**(1), 30–36. DOI: 10.3329/akmmcj.v4i1.13682.
180. RANJBAR, B. a GILL, P. Circular Dichroism Techniques: Biomolecular and Nanostructural Analyses- A Review. *Chemical Biology & Drug Design*. 2009, **74**(2), 101–120. DOI: 10.1111/j.1747-0285.2009.00847.x.

181. RAPIN, N. a PORSE, B. T. Oncogenic fusion proteins expressed in immature hematopoietic cells fail to recapitulate the transcriptional changes observed in human AML. *Oncogenesis*. 2014, **3**(6), e106–e106. DOI: 10.1038/oncsis.2014.22.
182. RARAN-KURUSSI, S. et al. Removal of Affinity Tags with TEV Protease. In BURGESS-BROWN, N. A. (ed.). *Heterologous Gene Expression in E. coli, Methods in Molecular Biology*. New York, NY: Springer New York, 2017 [cit. 2023-03-22], , 221–230. *Methods in Molecular Biology*. DOI: 10.1007/978-1-4939-6887-9_14. ISBN: 978-1-4939-6885-5.
183. REGAN, L. a DEGRADO, W. F. Characterization of a Helical Protein Designed from First Principles. *Science*. 1988, **241**(4868), 976–978. DOI: 10.1126/science.3043666.
184. REN, J. et al. DOG 1.0: illustrator of protein domain structures. *Cell Research*. 2009, **19**(2), 271–273. DOI: 10.1038/cr.2009.6.
185. RHETT, J. M. a GOURDIE, R. G. The perinexus: A new feature of Cx43 gap junction organization. *Heart Rhythm*. 2012, **9**(4), 619–623. DOI: 10.1016/j.hrthm.2011.10.003.
186. RHODES, D. G. a LAUE, T. M. Chapter 38 Determination of Protein Purity. In *Methods in Enzymology*. Elsevier, 2009 [cit. 2023-03-24], 677–689. DOI: 10.1016/S0076-6879(09)63038-X. ISBN: 978-0-12-374536-1.
187. RIZK, M. et al. End-to-end gene fusions and their impact on the production of multifunctional biomass degrading enzymes. *Biochemical and Biophysical Research Communications*. 2012, **428**(1), 1–5. DOI: 10.1016/j.bbrc.2012.09.142.
188. ROCCO, M. a BYRON, O. Computing translational diffusion and sedimentation coefficients: an evaluation of experimental data and programs. *European Biophysics Journal*. 2015, **44**(6), 417–431. DOI: 10.1007/s00249-015-1042-9.
189. ROMERO, G. et al. Role of PDZ Proteins in Regulating Trafficking, Signaling, and Function of GPCRs: Means, Motif, and Opportunity. In *Advances in Pharmacology*. Elsevier, 2011 [cit. 2023-03-11], , 279–314. DOI: 10.1016/B978-0-12-385952-5.00003-8. ISBN: 978-0-12-385952-5.
190. ROSE, G. D. Protein folding and the Paracelsus challenge. *Nature Structural Biology*. 1997, **4**(7), 512–514. DOI: 10.1038/nsb0797-512.
191. ROSE, G. D. a CREAMER, T. P. Protein folding: Predicting predicting. *Proteins: Structure, Function, and Genetics*. 1994, **19**(1), 1–3. DOI: 10.1002/prot.340190102.
192. ROY, A. et al. Computing the Relative Stabilities and the Per-Residue Components in Protein Conformational Changes. *Structure*. 2014, **22**(1), 168–175. DOI: 10.1016/j.str.2013.10.015.
193. RUNYON, S. T. et al. Structural and functional analysis of the PDZ domains of human HtrA1 and HtrA3. *Protein Science*. 2007, **16**(11), 2454–2471. DOI: 10.1110/ps.073049407.
194. SAHNI, N. et al. Widespread Macromolecular Interaction Perturbations in Human Genetic Disorders. *Cell*. 2015, **161**(3), 647–660. DOI: 10.1016/j.cell.2015.04.013.
195. SACHDEV, D. a CHIRGWIN, J. M. Order of Fusions between Bacterial and Mammalian Proteins Can Determine Solubility in Escherichia coli. *Biochemical and Biophysical Research Communications*. 1998, **244**(3), 933–937. DOI: 10.1006/bbrc.1998.8365.
196. SAKARYA, O. et al. Evolutionary Expansion and Specialization of the PDZ Domains. *Molecular Biology and Evolution*. 2010, **27**(5), 1058–1069. DOI: 10.1093/molbev/msp311.

197. SALINAS-GARCÍA, M. C. et al. pH-Driven Polymorphic Behaviour of the Third PDZ Domain of PSD95: The Role of Electrostatic Interactions. *Crystals*. 2023, **13**(2), 218. DOI: 10.3390/cryst13020218.
198. SANS, N. et al. A Developmental Change in NMDA Receptor-Associated Proteins at Hippocampal Synapses. *The Journal of Neuroscience*. 2000, **20**(3), 1260–1271. DOI: 10.1523/JNEUROSCI.20-03-01260.2000.
199. SANTIVERI, C. M. et al. A Biophysical Toolkit for Molecular Interactions. In JOHN WILEY & SONS LTD (ed.). *eLS*. Chichester, UK : John Wiley & Sons, Ltd, 2017 [cit. 2020-12-28], , 1–16. DOI: 10.1002/9780470015902.a0027015. ISBN: 978-0-470-01590-2.
200. SARAVANAN, K. M. et al. Sequence and structural analysis of two designed proteins with 88% identity adopting different folds. *Protein Engineering Design and Selection*. 2010, **23**(12), 911–918. DOI: 10.1093/protein/gzq070.
201. SATTLER, R. et al. Specific Coupling of NMDA Receptor Activation to Nitric Oxide Neurotoxicity by PSD-95 Protein. *Science*. 1999, **284**(5421), 1845–1848. DOI: 10.1126/science.284.5421.1845.
202. SEMENOV, M. V. et al. SnapShot: Noncanonical Wnt Signaling Pathways. *Cell*. 2007, **131**(7), 1378.e1-1378.e2. DOI: 10.1016/j.cell.2007.12.011.
203. SHEHADUL ISLAM, M. et al. A Review on Macroscale and Microscale Cell Lysis Methods. *Micromachines*. 2017, **8**(3), 83. DOI: 10.3390/mi8030083.
204. SHENG, R. et al. Cholesterol modulates cell signaling and protein networking by specifically interacting with PDZ domain-containing scaffold proteins. *Nature Communications*. 2012, **3**(1), 1249. DOI: 10.1038/ncomms2221.
205. SHENG, R. et al. Cholesterol selectively activates canonical Wnt signalling over non-canonical Wnt signalling. *Nature Communications*. 2014, **5**(1), 4393. DOI: 10.1038/ncomms5393.
206. SHENOY, A. R. a VISWESWARIAH, S. S. Site-directed mutagenesis using a single mutagenic oligonucleotide and DpnI digestion of template DNA. *Analytical Biochemistry*. 2003, **319**(2), 335–336. DOI: 10.1016/S0003-2697(03)00286-0.
207. SHEPHERD, T. R. et al. The Tiam1 PDZ Domain Couples to Syndecan1 and Promotes Cell–Matrix Adhesion. *Journal of Molecular Biology*. 2010, **398**(5), 730–746. DOI: 10.1016/j.jmb.2010.03.047.
208. SHEPLEY-MCTAGGART, A. et al. SARS-CoV-2 Envelope (E) protein interacts with PDZ-domain-2 of host tight junction protein ZO1. KOVAL, M. (ed.). *PLOS ONE*. 2021, **16**(6), e0251955. DOI: 10.1371/journal.pone.0251955.
209. SHIGETOMI, K. et al. Adherens junctions influence tight junction formation via changes in membrane lipid composition. *Journal of Cell Biology*. 2018, **217**(7), 2373–2381. DOI: 10.1083/jcb.201711042.
210. SHIMOZONO, S. a MIYAWAKI, A. Engineering FRET Constructs Using CFP and YFP. In *Methods in Cell Biology*. Elsevier, 2008 [cit. 2023-07-10], , 381–393. DOI: 10.1016/S0091-679X(08)85016-9. ISBN: 978-0-12-372558-5.
211. SCHECKEL, C. a AGUZZI, A. Prions, prionoids and protein misfolding disorders. *Nature Reviews Genetics*. 2018, **19**(7), 405–418. DOI: 10.1038/s41576-018-0011-4.
212. SCHLÜTER, O. M. et al. Alternative N-Terminal Domains of PSD-95 and SAP97 Govern Activity-Dependent Regulation of Synaptic AMPA Receptor Function. *Neuron*. 2006, **51**(1), 99–111. DOI: 10.1016/j.neuron.2006.05.016.
213. SCHNAPP, B. J. Trafficking of signaling modules by kinesin motors. *Journal of Cell Science*. 2003, **116**(11), 2125–2135. DOI: 10.1242/jcs.00488.

214. SCHNEEBERGER, E. E. a LYNCH, R. D. The tight junction: a multifunctional complex. *American Journal of Physiology-Cell Physiology*. 2004, **286**(6), C1213–C1228. DOI: 10.1152/ajpcell.00558.2003.
215. SCHUCK, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophysical Journal*. 2000, **78**(3), 1606–1619. DOI: 10.1016/S0006-3495(00)76713-0.
216. SCHUCK, P. a BRASWELL, E. H. Measuring Protein-Protein Interactions by Equilibrium Sedimentation. *Current Protocols in Immunology*. 2000, **40**(1). DOI: 10.1002/0471142735.im1808s40.
217. SCHULTZ, J. et al. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences*. 1998, **95**(11), 5857–5864. DOI: 10.1073/pnas.95.11.5857.
218. SIKOSEK, T. et al. Theoretical Insights into the Biophysics of Protein Bi-stability and Evolutionary Switches. JERNIGAN, R. L. (ed.). *PLOS Computational Biology*. 2016, **12**(6), e1004960. DOI: 10.1371/journal.pcbi.1004960.
219. SILVER, A. B. et al. Engineered antibody fusion proteins for targeted disease therapy. *Trends in Pharmacological Sciences*. 2021, **42**(12), 1064–1081. DOI: 10.1016/j.tips.2021.09.009.
220. SINGH, D. et al. Connexin 43 Interacts with Zona Occludens-1 and -2 Proteins in a Cell Cycle Stage-specific Manner. *Journal of Biological Chemistry*. 2005, **280**(34), 30416–30421. DOI: 10.1074/jbc.M506799200.
221. SMYTH, D. R. et al. Crystal structures of fusion proteins with large-affinity tags. *Protein Science*. 2003, **12**(7), 1313–1322. DOI: 10.1110/ps.0243403.
222. SNAPP, E. Design and Use of Fluorescent Fusion Proteins in Cell Biology. *Current Protocols in Cell Biology*. 2005, **27**(1). DOI: 10.1002/0471143030.cb2104s27.
223. SORIA-GUERRA, R. E. et al. Two decades of plant-based candidate vaccines: a review of the chimeric protein approaches. *Plant Cell Reports*. 2011, **30**(8), 1367–1382. DOI: 10.1007/s00299-011-1065-3.
224. SREERAMA, N. a WOODY, R. W. Estimation of Protein Secondary Structure from Circular Dichroism Spectra: Comparison of CONTIN, SELCON, and CDSSTR Methods with an Expanded Reference Set. *Analytical Biochemistry*. 2000, **287**(2), 252–260. DOI: 10.1006/abio.2000.4880.
225. STETEFELD, J. et al. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical Reviews*. 2016, **8**(4), 409–427. DOI: 10.1007/s12551-016-0218-6.
226. STEVENS, A. O. a HE, Y. Allosterism in the PDZ Family. *International Journal of Molecular Sciences*. 2022, **23**(3), 1454. DOI: 10.3390/ijms23031454.
227. STIFFLER, M. A. et al. PDZ Domain Binding Selectivity Is Optimized Across the Mouse Proteome. *Science*. 2007, **317**(5836), 364–369. DOI: 10.1126/science.1144592.
228. STRICKER, N. L. et al. PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences. *Nature Biotechnology*. 1997, **15**(4), 336–342. DOI: 10.1038/nbt0497-336.
229. STROHL, W. Chimeric Genes, Proteins. In *Reference Module in Life Sciences*. Elsevier, 2017 [cit. 2023-06-11], B9780128096338061000. DOI: 10.1016/B978-0-12-809633-8.06199-9. ISBN: 978-0-12-809633-8.
230. SÜDHOF, T. C. Neuroligins and neuroligins link synaptic function to cognitive disease. *Nature*. 2008, **455**(7215), 903–911. DOI: 10.1038/nature07456.

231. TAKAMIYA, K. et al. A direct functional link between the multi-PDZ domain protein GRIP1 and the Fraser syndrome protein Fras1. *Nature Genetics*. 2004, **36**(2), 172–177. DOI: 10.1038/ng1292.
232. TALMARD, C. et al. Amyloid-Beta Peptide Forms Monomeric Complexes With Cull and ZnII Prior to Aggregation. *ChemBioChem*. 2007, **8**(2), 163–165. DOI: 10.1002/cbic.200600319.
233. TOCHIO, H. et al. Formation of nNOS/PSD-95 PDZ dimer requires a preformed  $\beta$ -finger structure from the nNOS PDZ domain. *Journal of Molecular Biology*. 2000, **303**(3), 359–370. DOI: 10.1006/jmbi.2000.4148.
234. TONG, L. Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. *Cellular and Molecular Life Sciences*. 2005, **62**(16), 1784–1803. DOI: 10.1007/s00018-005-5121-4.
235. TONIKIAN, R. et al. A Specificity Map for the PDZ Domain Family. EDDY, S. R. (ed.). *PLoS Biology*. 2008, **6**(9), e239. DOI: 10.1371/journal.pbio.0060239.
236. TOYOFUKU, T. et al. c-Src Regulates the Interaction between Connexin-43 and ZO-1 in Cardiac Myocytes. *Journal of Biological Chemistry*. 2001, **276**(3), 1780–1788. DOI: 10.1074/jbc.M005826200.
237. TREUHEIT, M. J. et al. Inverse Relationship of Protein Concentration and Aggregation. *Pharmaceutical Research*. 2002, **19**(4), 511–516. DOI: 10.1023/A:1015108115452.
238. TRIPATHI, S. et al. Nanoparticles as Artificial Chaperons Suppressing Protein Aggregation: Remedy in Neurodegenerative Diseases. In SARMA, H. et al. (ed.). *Biobased Nanotechnology for Green Applications, Nanotechnology in the Life Sciences*. Cham: Springer International Publishing, 2021 [cit. 2023-04-11], , 311–338. *Nanotechnology in the Life Sciences*. DOI: 10.1007/978-3-030-61985-5_12. ISBN: 978-3-030-61984-8.
239. TUINSTRAL, R. L. et al. Interconversion between two unrelated protein folds in the lymphotactin native state. *Proceedings of the National Academy of Sciences*. 2008, **105**(13), 5057–5062. DOI: 10.1073/pnas.0709518105.
240. TYLER, R. C. et al. Distal Interactions within the par3–VE-Cadherin Complex. *Biochemistry*. 2010, **49**(5), 951–957. DOI: 10.1021/bi9017335.
241. UHLÉN, M. et al. Fusion proteins in biotechnology. *Current Opinion in Biotechnology*. 1992, **3**(4), 363–369. DOI: 10.1016/0958-1669(92)90164-E.
242. UNZAI, S. Analytical ultracentrifugation in structural biology. *Biophysical Reviews*. 2018, **10**(2), 229–233. DOI: 10.1007/s12551-017-0340-0.
243. VAN AELST, L. a D’SOUZA-SCHOREY, C. Rho GTPases and signaling networks. *Genes & Development*. 1997, **11**(18), 2295–2322. DOI: 10.1101/gad.11.18.2295.
244. VAN ITALLIE, C. M. et al. Claudin-2 Forms Homodimers and Is a Component of a High Molecular Weight Protein Complex. *Journal of Biological Chemistry*. 2011, **286**(5), 3442–3450. DOI: 10.1074/jbc.M110.195578.
245. VAN ITALLIE, C. M. et al. Visualizing the dynamic coupling of claudin strands to the actin cytoskeleton through ZO-1. NUSRAT, A. (ed.). *Molecular Biology of the Cell*. 2017, **28**(4), 524–534. DOI: 10.1091/mbc.e16-10-0698.
246. VANDERSTRAETEN, J. a BRIERS, Y. Synthetic protein scaffolds for the colocalisation of co-acting enzymes. *Biotechnology Advances*. 2020, **44**, 107627. DOI: 10.1016/j.biotechadv.2020.107627.
247. VOISIN, T. et al. Solubilization and hydrodynamic properties of active peptide YY receptor from rat jejunal crypts. Characterization as a Mr 44,000 glycoprotein. *Journal*

- of Biological Chemistry*. 1991, **266**(17), 10762–10767. DOI: 10.1016/S0021-9258(18)99083-6.
248. VOLKMAN, B. F. et al. Chapter 3 Lymphotactin Structural Dynamics. In *Methods in Enzymology*. Elsevier, 2009 [cit. 2023-04-12], 51–70. DOI: 10.1016/S0076-6879(09)05403-2. ISBN: 978-0-12-374907-9.
  249. WALCH, L. Emerging Role of the Scaffolding Protein Dlg1 in Vesicle Trafficking: Dlg1 in the Vesicle Trafficking. *Traffic*. 2013, **14**(9), 964–973. DOI: 10.1111/tra.12089.
  250. WANG, J. et al. The conserved domain database in 2023. *Nucleic Acids Research*. 2023, **51**(D1), D384–D388. DOI: 10.1093/nar/gkac1096.
  251. WANG, W. a ROBERTS, C. J. Protein aggregation – Mechanisms, detection, and control. *International Journal of Pharmaceutics*. 2018, **550**(1–2), 251–268. DOI: 10.1016/j.ijpharm.2018.08.043.
  252. WANG, Y. et al. A Large Family of Putative Transmembrane Receptors Homologous to the Product of the Drosophila Tissue Polarity Gene frizzled. *Journal of Biological Chemistry*. 1996, **271**(8), 4468–4476. DOI: 10.1074/jbc.271.8.4468.
  253. WAUGH, D. S. Making the most of affinity tags. *Trends in Biotechnology*. 2005, **23**(6), 316–320. DOI: 10.1016/j.tibtech.2005.03.012.
  254. WAWRZAK, D. et al. Frizzled–PDZ scaffold interactions in the control of Wnt signaling. *Advances in Enzyme Regulation*. 2009, **49**(1), 98–106. DOI: 10.1016/j.advenzreg.2009.01.002.
  255. WEI, C.-J. et al. CONNEXINS AND CELL SIGNALING IN DEVELOPMENT AND DISEASE. *Annual Review of Cell and Developmental Biology*. 2004, **20**(1), 811–838. DOI: 10.1146/annurev.cellbio.19.111301.144309.
  256. WINGFIELD, P. T. Overview of the Purification of Recombinant Proteins. *Current Protocols in Protein Science*. 2015, **80**(1). DOI: 10.1002/0471140864.ps0601s80.
  257. WITTCHEN, E. S. et al. Protein Interactions at the Tight Junction. *Journal of Biological Chemistry*. 1999, **274**(49), 35179–35185. DOI: 10.1074/jbc.274.49.35179.
  258. WOESTENENK, E. A. et al. His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors. *Journal of Structural and Functional Genomics*. 2004, **5**(3), 217–229. DOI: 10.1023/B:jsfg.0000031965.37625.0e.
  259. WU, T. et al. *Three Essential Resources to Improve Differential Scanning Fluorimetry (DSF) Experiments*. *Biochemistry*, 2020 [cit. 2021-05-02]. DOI: 10.1101/2020.03.22.002543.
  260. WYSZYNSKI, M. et al. Interaction between GRIP and Liprin- $\alpha$ /SYD2 Is Required for AMPA Receptor Targeting. *Neuron*. 2002, **34**(1), 39–52. DOI: 10.1016/S0896-6273(02)00640-2.
  261. XIAO, J. a TOLBERT, T. J. Synthesis of N-Terminally Linked Protein Dimers and Trimers by a Combined Native Chemical Ligation-CuAAC Click Chemistry Strategy. *Organic Letters*. 2009, **11**(18), 4144–4147. DOI: 10.1021/ol9016468.
  262. XU, J. et al. Early Embryonic Lethality of Mice Lacking ZO-2, but Not ZO-3, Reveals Critical and Nonredundant Roles for Individual Zonula Occludens Proteins in Mammalian Development. *Molecular and Cellular Biology*. 2008, **28**(5), 1669–1678. DOI: 10.1128/MCB.00891-07.
  263. YEON, Y. J. et al. Effect of His-tag location on the catalytic activity of 3-hydroxybutyrate dehydrogenase. *Biotechnology and Bioprocess Engineering*. 2014, **19**(5), 798–802. DOI: 10.1007/s12257-014-0089-2.

264. YOUNG, C. L. et al. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. *Biotechnology Journal*. 2012, **7**(5), 620–634. DOI: 10.1002/biot.201100155.
265. YU, K. et al. Synthetic fusion protein design and applications. *Biotechnology Advances*. 2015, **33**(1), 155–164. DOI: 10.1016/j.biotechadv.2014.11.005.
266. YUAN, S.-M. a CLARKE, N. D. A hybrid sequence approach to the paracelsus challenge. *Proteins: Structure, Function, and Genetics*. 1998, **30**(2), 136–143. DOI: 10.1002/(SICI)1097-0134(19980201)30:2<136::AID-PROT3>3.0.CO;2-L.
267. ZHAO, D. a HUANG, Z. Effect of His-Tag on Expression, Purification, and Structure of Zinc Finger Protein, ZNF191(243-368). *Bioinorganic Chemistry and Applications*. 2016, **2016**, 1–6. DOI: 10.1155/2016/8206854.
268. ZHOU, W. et al. GluR1 Controls Dendrite Growth through Its Binding Partner, SAP97. *The Journal of Neuroscience*. 2008, **28**(41), 10220–10233. DOI: 10.1523/JNEUROSCI.3434-08.2008.
269. ZHU, Y. et al. Deciphering the Unexpected Binding Capacity of the Third PDZ Domain of Whirlin to Various Cochlear Hair Cell Partners. *Journal of Molecular Biology*. 2020, **432**(22), 5920–5937. DOI: 10.1016/j.jmb.2020.09.012.
270. ZHU, Y. et al. Interactions of Severe Acute Respiratory Syndrome Coronavirus 2 Protein E With Cell Junctions and Polarity PSD-95/Dlg/ZO-1-Containing Proteins. *Frontiers in Microbiology*. 2022, **13**, 829094. DOI: 10.3389/fmicb.2022.829094.
271. ZUK, P. J. et al. GRPY: An Accurate Bead Method for Calculation of Hydrodynamic Properties of Rigid Biomacromolecules. *Biophysical Journal*. 2018, **115**(5), 782–800. DOI: 10.1016/j.bpj.2018.07.015.
272. *Fusion protein technologies for biopharmaceuticals: applications and challenges*. S. R. Schmidt (ed.). Hoboken, NJ: Wiley, 2013. 624 p. ISBN: 978-0-470-64627-4.
273. *Protein engineering handbook*. S. Lutz, a U. T. Bornscheuer (ed.). Weinheim : Wiley-VCH, 2009. 3 p. ISBN: 978-3-527-31850-6.
274. *Protein purification: principles, high resolution methods, and applications*. J.-C. Janson (ed.). 3rd ed. Hoboken, NJ: John Wiley & Sons, 2011. 517 p. Wiley series in methods of biochemical analysis. ISBN: 978-0-471-74661-4.
275. *Viral entry into host cells*. S. Poehlmann, a G. Simmons (ed.). New York, NY. Austin, Tex : Springer Science+Business Media Landes Bioscience, 2013. Advances in experimental medicine and biology. ISBN: 978-1-4614-7651-1.

## Internetové zdroje

1. Cytiva, 2023. *Superdex™ 75 Increase* [online]. Cytiva [cit. 12. 4. 2023]. Dostupné z: <https://www.cytivalifesciences.com/en/us/shop/chromatography/prepacked-columns/size-exclusion/superdex-75-increase-p-06188#tech-spec-table>
2. Geneaid, 2017. High-Speed Plasmid Mini Kit [online]. Geneaid [cit. 9. 7. 2023]. Dostupné z: <https://www.geneaid.com/data/files/1605261423275877369.pdf>
3. Invitrogen. *One Shot® TOP10 Electrocomp™ E. coli* [online]. Thermo Fisher Scientific [cit. 20. 6. 2023]. Dostupné z: [https://assets.fishersci.com/TFS-Assets/LSG/manuals/oneshottop10_ecomp_man.pdf](https://assets.fishersci.com/TFS-Assets/LSG/manuals/oneshottop10_ecomp_man.pdf)
4. New England Biolabs, 2013. *PCR Using Q5® High-Fidelity DNA Polymerase (M0491)* [online]. New England Biolabs [cit. 14. 6. 2023]. Dostupné z: <https://international.neb.com/protocols/2013/12/13/pcr-using-q5-high-fidelity-dna-polymerase-m0491>
5. New England Biolabs, 2023. *Tm Calculator* [online]. New England Biolabs [cit. 25. 5. 2023]. Dostupné z: <https://tmcalculator.neb.com/#!/main>
6. *RSCB Protein Data Bank*, 2023 [online]. RCSB PDB Core Operations [20. 6. 2023]. Dostupné z: <https://www.rcsb.org>
7. *Sednterp*, 2012 [online]. University of New Hampshire [cit. 3. 7. 2023]. Dostupné z: [https://bitc.sr.unh.edu/index.php?title=Main_Page](https://bitc.sr.unh.edu/index.php?title=Main_Page)
8. *UniProt*, 2023 [online]. UniProt consortium [cit. 2. 7. 2023]. Dostupné z: <https://www.uniprot.org>



## Příloha I

Protokol pro purifikaci plazmidové DNA pomocí komerčního kitu High-Speed Plasmid Mini Kit (Geneaid) poskytnutý výrobcem kitu. Dostupné z: <https://www.geneaid.com/data/files/1605261423275877369.pdf>

Harvesting	<ul style="list-style-type: none"> <li>Transfer 1.5 ml of cultured bacterial cells to a microcentrifuge tube.</li> <li>Centrifuge at 14-16,000 x g for 1 minute then discard the supernatant.</li> <li>If using more than 1.5 ml of cultured bacterial cells, repeat the Harvesting Step.</li> </ul>
Step 1 Re-suspension	<ul style="list-style-type: none"> <li>Add <b>200 µl of PD1 Buffer</b> (make sure RNase A was added) to the tube.</li> <li>Re-suspend the cell pellet by vortex or pipetting.</li> </ul>
Step 2 Lysis	<ul style="list-style-type: none"> <li>Add <b>200 µl of PD2 Buffer</b> (make sure any precipitates are dissolved).</li> <li>Mix gently by inverting the tube 10 times.</li> </ul> <p>NOTE: Do not vortex to avoid shearing the genomic DNA.</p> <ul style="list-style-type: none"> <li>Let stand at room temperature for at least 2 minutes (do not exceed 5 minutes).</li> </ul> <p>NOTE: This will ensure the lysate is homogeneous.</p>
Step 3 Neutralization	<ul style="list-style-type: none"> <li>Add <b>300 µl of PD3 Buffer</b> and mix immediately by inverting the tube 10 times.</li> </ul> <p>NOTE: Do not vortex to avoid shearing the genomic DNA.</p> <ul style="list-style-type: none"> <li>Centrifuge at 14-16,000 x g for 3 minutes.</li> </ul>
Step 4 DNA Binding	<ul style="list-style-type: none"> <li>Place a <b>PD Column</b> in a <b>2 ml Collection Tube</b>.</li> <li>Add the supernatant from Step 3.</li> <li>Centrifuge at 14-16,000 x g for 30 seconds then discard the flow-through.</li> <li>Place the <b>PD Column</b> back in the <b>2 ml Collection Tube</b>.</li> </ul>
Step 5 Wash	<p><b>Optional: For Sequencing</b></p> <ol style="list-style-type: none"> <li>Add <b>400 µl of W1 Buffer</b> into the <b>PD Column</b>.</li> <li>Centrifuge at 14-16,000 x g for 30 seconds then discard the flow-through.</li> <li>Place the <b>PD Column</b> back in the <b>2 ml Collection Tube</b>.</li> <li>Proceed with <b>Wash Buffer</b> addition.</li> </ol> <ul style="list-style-type: none"> <li>Add <b>600 µl of Wash Buffer</b> (make sure ethanol was added) into the <b>PD Column</b> then centrifuge at 14-16,000 x g for 30 seconds.</li> <li>Discard the flow through.</li> <li>Place the <b>PD Column</b> back in the <b>2 ml Collection Tube</b>.</li> <li>Centrifuge at 14-16,000 x g for 3 minutes to dry the column matrix.</li> <li>Place the <b>PD Column</b> in a new 1.5 ml microcentrifuge tube.</li> </ul>
Step 6 DNA Elution	<p>If a higher DNA concentration is required, use 30 µl of Elution Buffer. If maximum DNA yield is required, use 100 µl of Elution Buffer (DNA concentration will be diluted).</p> <ul style="list-style-type: none"> <li>Add <b>50 µl of Elution Buffer</b> or TE into the CENTER of the <b>PD Column</b> matrix.</li> <li>Let stand for at least 2 minutes to allow the <b>Elution Buffer</b> or TE to be completely absorbed.</li> <li>Centrifuge at 14-16,000 x g for 2 minutes to elute the purified DNA.</li> </ul> <p>NOTE: Performing a second elution step will increase yield between 10-30%.</p> <ul style="list-style-type: none"> <li>Transfer the flow-through back into the CENTER of the <b>PD Column</b> matrix.</li> <li>Centrifuge at 14-16,000 x g for 2 minutes to elute the purified DNA.</li> </ul>

## Příloha II

Stručné výsledky MS analýz FD5A, FD6A a G_A88. Součástí jsou testované sekvence proteinů včetně nedetekovaných míst (červeně), specifické sekvence, které byly stanoveny a jejich modifikace, počet detekcí (tzv. #PSMs) a pozice ve výchozí sekvenci, kterým odpovídají.

### FD5A

GHMLEGSHMI LRPSMKLVKF RKGDSVGLRL AGGNDVGIFV AGVLEDSPA KEGLEEGDQI LRVNVDFTN  
IIREEAVLFL LDLPKGEEVT ILAQKPAPAP TTYKLILNLK QAKEEAIKEL VDAGIAEKYI KLIANAKTVE  
GVWTLKDEIL TFTVTE

Detekovaná sekvence	Modifikace	Počet	Pozice
[K] .DEILTFTVTE . [-]	-	6	147-156
[K] .EEAIKELVDAGIAEK . [Y]	-	34	114-128
[K] .EEAIKELVDAGIAEKYIK . [L]	-	3	114-131
[R] .EEAVLFLLDLPK . [G]	-	11	74-85
[R] .EEAVLFLLDLPKGEEVTILAQKPAPAPTTYK . [L]	-	2	74-104
[K] .EGLEEGDQILR . [V]	-	20	52-62
[K] .EGLEEGDQILRVNVDFTNIIR . [E]	1xDeamid. [N13]	1	52-73
[K] .EGLEEGDQILRVNVDFTNIIREEAVLFLLDLPK . [G]	1xDeamid. [N/Q]	4	52-85
[K] .ELVDAGIAEK . [Y]	-	18	119-128
[K] .ELVDAGIAEKYIK . [L]	-	4	119-131
[K] .ELVDAGIAEKYIKLIANAK . [T]	-	4	119-137
[K] .FRKDSVGLR . [L]	-	24	20-29
[K] .GEEVTILAQKPAPAPTTYK . [L]	-	32	86-104
[K] .GEEVTILAQKPAPAPTTYKLILNLK . [Q]	-	6	86-110
[-] .GHMLEGSHMILRPSMK . [L]	-	11	1-16
[-] .GHMLEGSHMILRPSMK . [L]	1xOxidace [M]	28	1-16
[-] .GHMLEGSHMILRPSMK . [L]	2xOxidace [M]	21	1-16
[-] .GHMLEGSHMILRPSMK . [L]	3xOxidace [M3; M9; M15]	27	1-16
[R] .KGDVGLR . [L]	-	17	22-29
[R] .LAGGNDVGIFVAGVLEDSPA . [E]	-	13	30-51
[R] .LAGGNDVGIFVAGVLEDSPA . [E]	1xDeamid. [N5]	2	30-51
[R] .LAGGNDVGIFVAGVLEDSPAKEGLEEGDQILR . [V]	-	15	30-62
[R] .LAGGNDVGIFVAGVLEDSPAKEGLEEGDQILR . [V]	1xDeamid. [N5]	1	30-62
[K] .LILNLK . [Q]	-	12	105-110
[K] .LILNLKQAK . [E]	-	1	105-113
[K] .LILNLKQAKEEAIK . [E]	-	24	105-118
[K] .LILNLKQAKEEAIK . [E]	1xDeamid. [N4]	4	105-118
[K] .QAKEEAIKELVDAGIAEK . [Y]	-	50	111-128
[K] .QAKEEAIKELVDAGIAEK . [Y]	1xDeamid. [Q1]	6	111-128
[K] .TVEGVWTLKDEILTFTVTE . [-]	-	37	138-156
[R] .VNVVDFTNIIR . [E]	-	13	63-73
[R] .VNVVDFTNIIREEAVLFLLDLPK . [G]	-	8	63-85
[R] .VNVVDFTNIIREEAVLFLLDLPK . [G]	1xDeamid. [N3]	2	63-85
[R] .VNVVDFTNIIREEAVLFLLDLPKGEEVTILAQKPAPAP TTYK . [L]	-	11	63-104

[R] . VNNVDFTNI IREEAVLFLLDLPKGEEVTILAQKPAPAP TTYK. [L]	2xDeamid. [N8; Q32]	1	63-104
[K] . YIKLIANAK. [T]	-	2	129-137
[K] . YIKLIANAK. [T]	1xDeamid. [N7]	2	129-137

## FD6A

GGGTTYKLIL NLKQAKEEAI KELVDAGIAE KYIKLIANAK TVEGVWTLKD EILTFTVTEP APAPMKLVK  
RKGDSVGLRL AGGNDVGIFV AGVLEDSPAK KEGLEEGDQI LRVNNVDFTN IREEAVLFL LDLPKGEEVT  
ILAQK

Detekovaná sekvence	Modifikace	Počet	Pozice
[K] . DEILTFTVTEPAPAPMK. [L]	1xOxidace [M16]	4	50-66
[K] . DEILTFTVTEPAPAPMKLVK. [F]	1xOxidace [M16]	1	50-69
[K] . EEAIKELVDAGIAEK. [Y]	-	10	17-31
[K] . EEAIKELVDAGIAEKYIK. [L]	-	14	17-34
[R] . EEAVLFLLDLPK. [G]	-	6	124-135
[R] . EEAVLFLLDLPKGEEVTILAQK. [K-]	-	2	124-145
[K] . EGLEEGDQILR. [V]	-	36	102-112
[K] . EGLEEGDQILRVNNVDFTNIIR. [E]	-	3	102-123
[K] . EGLEEGDQILRVNNVDFTNIIR. [E]	1xDeamid.[Q8]	1	102-123
[K] . ELVDAGIAEK. [Y]	-	3	22-31
[K] . ELVDAGIAEKYIK. [L]	-	4	22-34
[K] . ELVDAGIAEKYIKLIANAK. [T]	-	16	22-40
[K] . ELVDAGIAEKYIKLIANAK. [T]	1xDeamid. [N17]	5	22-40
[K] . FRKGSVGLR. [L]	-	46	70-79
[K] . GDSVGLRLAGGNDVGIFVAGVLEDSPAK. [E]	-	3	73-101
[K] . GEEVTILAQK. [K-]	-	8	136-145
[-] . GGGTTYKLILNLK. [Q]	-	34	1-13
[-] . GGGTTYKLILNLK. [Q]	1xDeamid. [N11]	4	1-13
[-] . GGGTTYKLILNLKQAK. [E]	-	16	1-16
[-] . GGGTTYKLILNLKQAK. [E]	1xDeamid.[N/Q]	2	1-16
[KR] . KGSVGLR. [L]	-	2	72-79
[R] . KGSVGLRLAGGNDVGIFVAGVLEDSPAK. [E]	-	6	72-101
[R] . KGSVGLRLAGGNDVGIFVAGVLEDSPAK. [E]	1xDeamid. [N13]	5	72-101
[R] . LAGGNDVGIFVAGVLEDSPAK. [E]	-	41	80-101
[R] . LAGGNDVGIFVAGVLEDSPAK. [E]	1xDeamid. [N5]	4	80-101
[R] . LAGGNDVGIFVAGVLEDSPAAKEGLEEGDQILR. [V]	-	21	80-112
[R] . LAGGNDVGIFVAGVLEDSPAAKEGLEEGDQILR. [V]	1xDeamid. [N5]	1	80-112
[R] . LAGGNDVGIFVAGVLEDSPAAKEGLEEGDQILRVN NVDFTNIR. [E]	-	3	80-123
[R] . LAGGNDVGIFVAGVLEDSPAAKEGLEEGDQILRVN NVDFTNIR. [E]	1xDeamid. [N35]	1	80-123
[K] . LIANAKTVEGVWTLKDEILTFTVTEPAPAPMK. [L]	-	2	35-66
[K] . LIANAKTVEGVWTLKDEILTFTVTEPAPAPMK. [L]	1xDeamid. [N4]	3	35-66
[K] . LIANAKTVEGVWTLKDEILTFTVTEPAPAPMK. [L]	1xOxidace [M31]	3	35-66
[K] . LIANAKTVEGVWTLKDEILTFTVTEPAPAPMK. [L]	1xDeamid. [N4]; 1xOxidace [M31]	3	35-66
[K] . LILNLKQAK. [E]	-	2	8-16

[K] . LILNLKQAKEEAIK . [E]	-	12	8-21
[K] . LILNLKQAKEEAIK . [E]	2xDeamid. [N4; Q7]	3	8-21
[K] . QAKEEAIKELVDAGIAEK . [Y]	-	15	14-31
[K] . QAKEEAIKELVDAGIAEK . [Y]	1xDeamid. [Q1]	7	14-31
[K] . TVEGVWTLK . [D]	-	1	41-49
[K] . TVEGVWTLKDEILTFTVTEPAPAPMK . [L]	-	11	41-66
[K] . TVEGVWTLKDEILTFTVTEPAPAPMK . [L]	1xOxidace [M25]	15	41-66
[K] . TVEGVWTLKDEILTFTVTEPAPAPMKLVK . [F]	-	6	41-69
[K] . TVEGVWTLKDEILTFTVTEPAPAPMKLVK . [F]	1xOxidace [M25]	27	41-69
[R] . VNNVDFTNIIR . [E]	-	6	113-123
[R] . VNNVDFTNIIR . [E]	1xDeamid. [N3]	4	113-123
[R] . VNNVDFTNIIREEAVLFLLDLPK . [G]	-	9	113-135
[R] . VNNVDFTNIIREEAVLFLLDLPK . [G]	1xDeamid. [N8]	2	113-135
[R] . VNNVDFTNIIREEAVLFLLDLPKGEEVTILAQK . [K-]	-	11	113-145
[K] . YIKLIANAK . [T]	-	3	32-40
[K] . YIKLIANAK . [T]	1xDeamid. [N7]	3	32-40

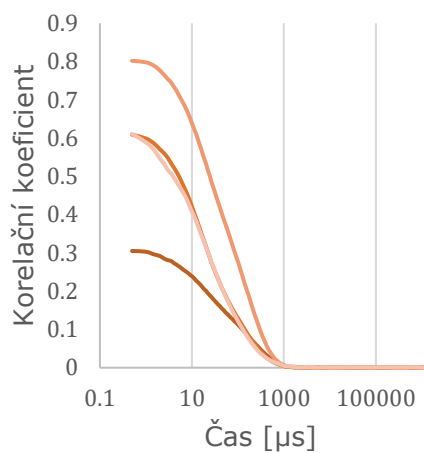
## GA88

TTYKLILNLK QAKEEAIKEL VDAGIAEKYI KLIANAKTVE GVWTLKDEIL TFTVTE

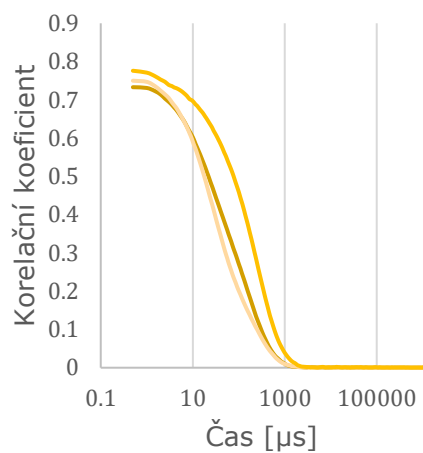
Detekovaná sekvence	Modifikace	Počet	Pozice
[K] . EEAIKELVDAGIAEK . [Y]	-	8	14-28
[K] . EEAIKELVDAGIAEKYIK . [L]	-	14	14-31
[K] . ELVDAGIAEK . [Y]	-	3	19-28
[K] . ELVDAGIAEKYIK . [L]	-	12	19-31
[K] . ELVDAGIAEKYIKLIANAK . [T]	-	8	19-37
[K] . LIANAKTVEGVWTLKDEILTFTVTE . [-]	-	17	32-56
[K] . LIANAKTVEGVWTLKDEILTFTVTE . [-]	1xDeamid.[N4]	3	32-56
[K] . LILNLKQAK . [E]	-	1	5-13
[K] . LILNLKQAKEEAIK . [E]	-	6	5-18
[K] . QAKEEAIKELVDAGIAEK . [Y]	-	11	11-28
[K] . QAKEEAIKELVDAGIAEK . [Y]	1xDeamid. [Q1]	2	11-28
[-] . TTYKLILNLK . [Q]	-	6	1-10
[-] . TTYKLILNLKQAK . [E]	-	1	1-13
[K] . TVEGVWTLKDEILTFTVTE . [-]	-	47	38-56
[K] . YIKLIANAK . [T]	-	4	29-37
[K] . YIKLIANAK . [T]	1xDeamid. [N7]	6	29-37

## Příloha III

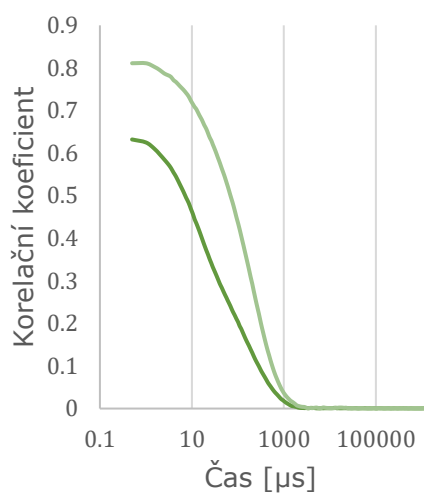
Výchozí data získaná měřeními DLS v podobě autokorelačních funkcí.



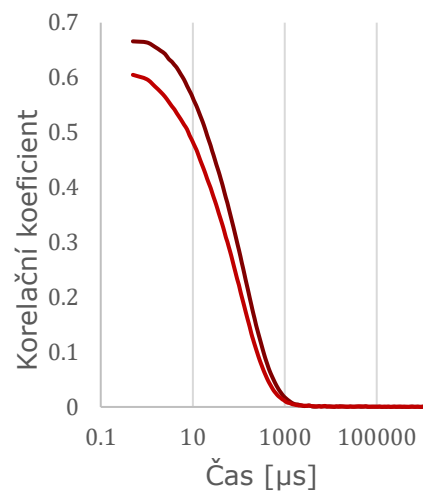
FD5A (0,2x)    FD5A  
FD5A (2x)    FD5A (Refold)



FD6A    FD6A (2x, 4 °C)  
FD6A (Refold)



PDZ3    PDZ3 (Refold)



GA88    GA88 (Refold)

## Příloha IV

Použité nástroje pro predikci vlastností z modelovaných terciárních struktur (ColabFold) (Mirdita et al., 2022) včetně hrubého popisu jejich algoritmu a využití v této práci:

Pomocí softwaru HYDROPRO, verze 10 (kompilace 54) (García De La Torre et al., 2000; Ortega et al., 2011) byly z molekulové hmotnosti a parciální specifického objemu (Sednterp, verze 20130813 BETA) (Laue et al., 1992) a předpovídané terciární struktury proteinů (ColabFold) (Mirdita et al., 2022) predikovány hodnoty sedimentačních koeficientů ( $s$ ,  $s_{20,w}$ ) a translačního difuzního koeficientu ( $D_t$ , Stokesův poloměr přepočítán dle Stokesovy-Einsteinovy rovnice, viz výše) pomocí tří modelů: atomového modelu s obalovými výpočty (model 1), reziduálního modelu s obalovými výpočty (model 2), a reziduálního modelu s korálovými výpočty (model 4) s výpočtem hydrodynamických charakteristik metodou inverze supermatice (SMI).

Pomocí softwaru UltraScan III, verze 4.0 (6731) (Demeler, 2005) byly z terciární struktury proteinů (ColabFold) (Mirdita et al., 2022) predikovány molekulová hmotnost, frikční koeficient, standardizovaný sedimentační koeficient a Stokesův poloměr, a to pomocí modelu korálek SoMo a překryvu korálek SoMo (O) a následným výpočtem hydrodynamických charakteristik metodou Zeno algoritmu (ZENO), generalizované aproximace dle Rotne-Prager-Yamakawa (GRPY), nebo SMI (Rocco a Byron, 2015; Zuk et al., 2018).

Pomocí online nástroje HullRad, verze 9 (Fleming a Fleming, 2018) byly z terciární struktury proteinů (ColabFold) (Mirdita et al., 2022) predikovány molekulová hmotnost, frikční koeficient, standardizovaný sedimentační koeficient, Stokesův poloměr, a to modelem konvexního trupu proteinu.

Pomocí online nástroj PDBMD2CD (Drew a Janes, 2020) byla z terciární struktury proteinů (ColabFold) (Mirdita et al., 2022) predikována CD spektra a jejich kompozice sekundárních struktur. Tento software využívá sadu 83 referenčních proteinů (SMP180) se známými spektry (Abdul-Gader et al., 2011) a dva nezávislé algoritmy založené na empirické statistické analýze: a) na základě strukturní podobnosti dodaných struktur s referenčními proteiny je proveden fit do jejich spekter s následnou optimalizací jednotlivých příspěvků; b) z dodané struktury je zjištěn podíl jednotlivých komponent sekundární struktury (7 kategorií) a na základě příspěvků těchto kategorií ve spektrech referenčních proteinů je sestaveno spektrum dodané struktury. Výsledné predikované CD spektrum zahrnuje výsledky obou algoritmů (Drew a Janes, 2020).

## Příloha V

DSF, optimalizační měření s různými koncentracemi proteinů (40, 20 a 10  $\mu\text{g/ml}$ ) a barviva SYPRO Orange (1:1000 a 1:500), průměry z triplikátů. Pro další analýzy byly vybrány byly koncentrace poskytující dobré křivky pro všechny tři proteiny.

