Identification of novel macrocyclic inhibitors of PSD-95 using tailored mRNA encoded libraries

Ellen Walse

DIVISION OF PURE AND APPLIED BIOCHEMISTRY | DEPARTMENT OF CHEMISTRY | FACULTY OF ENGINEERING LTH | LUND UNIVERSITY

IN COLLABORATION WITH

DEPARTMENT OF DRUG DESIGN AND PHARMACOLOGY | UNIVERSITY OF COPENHAGEN

MASTER THESIS 2024

Identification of macrocyclic inhibitors of PSD-95 using tailored mRNA encoded libraries

Subject:
KBKM05 Degree Project in Applied Biochemistry for Engineers

Author:
Ellen Walse

Supervisor at University of Copenhagen:
Kristian Strømgaard, Professor, Department of Drug Design and Pharmacology, University of Copenhagen

Supervisor at LTH:
Johan Svensson Bonde, Senior Lecturer, Division of Pure and Applied biochemistry, Department of Chemistry, Faculty of Engineering, Lund University

Examiner:
Lei Ye, Professor, Division of Pure and Applied biochemistry, Department of Chemistry, Faculty of Engineering, Lund University

Acknowledgements

The research presented in this thesis was conducted as part of my master's thesis project in applied biochemistry. All laboratory work was carried out between January and June 2024, in the Strømagard lab and in collaboration with the Rogers lab at Centre for Biopharmaceuticals, Department of Drug Design and Pharmacology, University of Copenhagen. The project was supervised by Professor Kristian Strømgaard and Assistant Professor Christian Bech-Bartling at University of Copenhagen, and by Senior Lecturer Johan Svensson Bonde at the Faculty of Engineering, Lund University.

To begin with, I want to thank Professor Kristian Strømgaard for accepting my application of a master's thesis project in his research group, and for allowing me to pursue such an interesting project. What I will bring with me from this experience is not only appreciation for peptide and protein research, but also professional and personal growth.

I want to direct special gratitude to Assistant Professor Christian Bech-Bartling for being my daily supervisor during this project, supporting me in the lab, with writing as well as with daily challenges or questions I had.

I also want to thank Associate Professor Jospeh M. Rogers, for his guidance in teaching me all aspects of the RaPID system, as well as providing the necessary materials, equipment, and lab space. A big thanks to PhD student Jing Zhao and Postdoc Fabian Hink in the Rogers group for answering my daily questions. I have learned a great deal from this experience and have become a better scientist because of it.

A big thank you to everyone in the Strømgaard lab for being such an easygoing, joyful, and welcoming group.

Lastly, I want to thank my parents for their support and for cheering me on during this period.

Development of a peptide library to find protein-binders for treatment of stroke-related brain damage

Stroke is a health problem all over the world. It happens when blood vessels in the brain are blocked, causing devastating brain-damage or even death. Current drugs for stroke must be given quickly and be of the right kind to work. Therefore, new drugs treating stroke-related brain-damage are needed. Peptides are short chains of amino acids and can bind to protein surfaces that small molecules - the traditional type of drug – cannot. A method to identify cyclic peptide drug-candidates is to reprogram the genetic code from which peptides are produced, to create a large library with up to a trillion peptides. By displaying the library towards a disease-related protein, a peptide that binds to the protein can be found. In this project, this method was used against a stroke-related protein in the brain. The project discovered a peptide similar to previously identified peptides using the same method. Key amino acids were identified important for binding, agreeing with the literature.

Stroke is the second leading cause of death in the world. This health issue is caused by blockage of blood vessels in the brain, which causes parts of the brain to die, leaving the patient with brain-damage that can affect speech, the ability to walk and move, or in the worst-case scenario, death. There are only a few existing drugs today that treat brain-damage caused by stroke. However, they need to be given to the patient in a very narrow time-period from disease start to be effective.

The history of drug development towards all kinds of diseases have long been dominated by small molecules or larger drugs like antibodies (proteins that protect you when unwanted substances enter the body). Today's drug development methods are exploring the functions of peptides. Peptides are short chains of amino acids linked by peptide bonds. The human body naturally contain peptides, where maybe the most commonly known peptide is insulin, that controls the bodies energy supply. Peptides advantage as drugs are their ability to bind to protein surfaces that small molecules are unable to, expanding the ways of how to treat diseases. Their disadvantage is that they are broken down fast by the body's own enzymes. They also have difficulties entering cells and passing the barrier surrounding the brain. Cyclic peptides are peptide chains with a ring structure. Development of cyclic peptides is a possible solution for slower breakdown by the body's enzyme. However, to develop successful cyclic peptides as drugs they need to have structural characteristics different from natural amino acids. There are different methods to do this, but one way is to reprogram the genetic code from which peptides are produced so that they contain non-natural amino acids. Creation of such peptides is done by using a cell-free translation machinery (the process in which peptides are produced

using RNA molecules as templates) and man-made ribosome-related enzymes. Imagine being able to use this method to create a library containing more a trillion unique cyclic peptides with non-natural amino acids, and then displaying this library against a protein target with the aim to find a binder. This is a relatively new and very powerful method that expands the way for how to discover cyclic peptides as drugs.

In this thesis, performed at University of Copenhagen, this powerful method was used to create a library of a trillion cyclic peptides containing nine to thirteen amino acids. The library was displayed against a protein present in the brain and that is known to be related to the disease-mechanism of stroke. The aim of the study was to identify a cyclic peptide with structural characteristics so that it would not only bind to the stroke-related protein, but also be small enough to pass the barrier surrounding the brain compared with previously discovered peptides. The overall purpose of the study was to get further down the road towards developing a new, more effective drug for treatment of brain-damage caused by stroke.

Surprisingly, a longer peptide than in the initial library was discovered using the described method. However, this peptide was almost identical to previously identified peptides, using the same method. This points to the method being powerful for drug discovery. Key amino acid patterns of the peptide were identified as important for binding to the stroke-related protein. These key amino acids agreed with earlier studies. Lastly, this study provides an additional base for future attempts in finding a cyclic peptide drug-candidate against stroke, a devastating disease.

Abstract

Macrocyclic peptides are promising therapeutic modalities due to their ability to target protein-protein interactions (PPIs) with high specificity and affinity. In contrast to their linear analogues, macrocycles are conformationally constrained and the lack of N- and C-terminal make them more resistant to metabolic degradation. Advances in *de novo* peptide drug discovery, such as the Random non-standard peptides integrated discovery (RaPID) system, provides a powerful mRNA display technology and facilitates generation of large (10^{12}) genetically diverse macrocyclic peptide libraries for development of therapeutic entities. The post-synaptic density protein 95 (PSD-95) is an abundant scaffolding protein present within the post synaptic density (PSD) of excitatory synapses. Inhibiting the neuronal nitric oxide synthase (nNOS)/PSD-95 interaction is a promising therapeutic strategy for preventing brain damage from acute ischemic stroke, a disease responsible as the second leading cause of death in the world. This thesis focused on the screening of a randomized cyclic peptide library, specifically comprising nine to thirteen amino acids (AAs), expressed using the RaPID platform. Employing next generation sequencing led to the discovery of a potential 17-residue macrocyclic peptide binder towards the PDZ2 domain of PSD-95 similar to previously discovered second-generation peptides. The 17 AAs containing peptide exhibited certain internal binding motifs previously described crucial for binding in the literature. Lastly, structure-activity relationships predicted with machine learning tools yielded a foundation for future optimization-efforts towards a lead-candidate.

List of abbreviations

ABU	α-Aminobutyric acid
AA	L-α-amino acids
aa-tRNA	Aminoacyl-tRNA
ABT	Amino-derivatized benzyl thioester
AcOH	Acetic acid
aFx	amino flexizyme
AIS	Acute ischemic stroke
AMPARs	α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors
ATP	Adenosine triphosphate
BBB	Blood-brain barrier
Bmt	(4R)-4-[(E)-2-butenyl]-4,N-dimethyl-I-threonine
BSA	Bovine serum albumin
bp	Base pair
ClAc	Chloro-acetyl
cDNA	Complementary DNA
CME	Cyanomethyl ester
CMP	Cell membrane permeability
CPP	Cell-penetrating peptide
Da	Dalton
DBE	Dinitrobenzyl ester
dFx	dinitro flexizyme
DLG-1	Disks large homolog 1
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DTT	Dithiothreitol
E. coli	*Escherichia coli*
eFx	enhanced flexizyme
EtOH	Ethanol
FIT	Flexible *in vitro* translation
FP	Fluorescence polarization
FPLC	Fast protein liquid chromatography
GK	Guanylate kinase

GKAP	Guanylate kinase-associated protein
HBDs	Hydrogen bond donors
HEPES	4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid
HIV-1	Human immunodeficiency virus type 1
$^{ini}tRNA_{CAU}$	Initiator tRNA
IC_{50}	Half-maximal inhibitory concentration
iPrOH	Isopropanol
kb	Kilobase
K_d	Dissociation constant
K_i	Inhibitory constant
MAGUK	Membrane-associated guanylate kinase
Me	Methyl
mRNA	Messenger RNA
ncAA	Non-canonical amino acid
NGS	Next Generation Sequencing
NMDARs	*N*-methyl-D-aspartate receptors
nNOS	Neuronal nitric oxide synthase
NO	Nitric oxide
*N*PEG	Amino-polyethylene glycol
NTPs	Nucleoside triphosphates
OH	Hydroxy
PAE	Predicted aligned error
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PCSK9	Proprotein subtilisin/kexin type 9
PDB	Protein Data Bank
PDBePISA	Protein Data Bank in Europe Proteins, Interfaces, Structures and Assemblies
PDZ	Postsynaptic density-95, Discs-large, Zona occludens 1
PEG	Polyethylene glycol
plDDT	Predicted local distance difference test
pMCAO	Permanent middle cerebral artery occlusion
PPI	Protein-protein interaction
PSD	Postsynaptic density
PSD-95	Postsynaptic density protein 95

PURE	Protein synthesis using recombinant elements
qPCR	Quantitative PCR
RaPID	Random non-standard peptides integrated discovery
RMSD	Root mean square deviation
RNA	Ribonucleic acid
RNase	Ribonuclease
RT	Room temperature
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SH3	Src homology 3
Shank	SH3 and ankyrin repeat-containing protein
ssDNA	Single stranded DNA
TAT	Trans-activating transcriptional activator
tPA	Tissue plasminogen activator
tRNA	Transfer RNA
UV	Ultraviolet
ZO-1	Zona occludens-1

AAs are referred to as their three-letter code in the text.

DNA and RNA sequences are written with capital letters.

Figures were made either with Biorender.com or with ChemDraw 23.1.1.

Table of contents

1. Introduction

1.1 Postsynaptic density (PSD)

Cell-to-cell communication in the nervous system is achieved via an electrochemical signal transmission at the synapse. Neurotransmitters are released from the presynaptic terminal located at the end of an axon in response to an electrical signal and diffuses across the synaptic cleft to bind specific receptors at the postsynaptic terminal of the synapse.[1] The majority of excitatory synapses in the mammalian nervous system are glutamatergic, releasing glutamate to mediate *N*-methyl-D-aspartate receptors (NMDARs) and α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid receptors (AMPARs). Upon binding, the receptors open to allow influx of Na^+ ions and in the case of NMDARs also Ca^{2+} ions to further propagate the electrical signal. In the excitatory postsynaptic membrane these receptors are embedded in a complex and dense network of scaffolding and adaptor proteins, signaling enzymes, adhesives and cytoskeletal components known as the postsynaptic density (PSD).[2] The discovery of the PSD initially occurred through electron microscopy in 1956[3], where electron-dense thickenings were observed beneath postsynaptic membranes (Figure 1.1 a). Subsequently, it was acknowledged that PSDs are composed of densely packed proteins, forming disc-shaped mega-assemblies[3] that are 30-50 nm thick and 200-800 nm wide.[2]

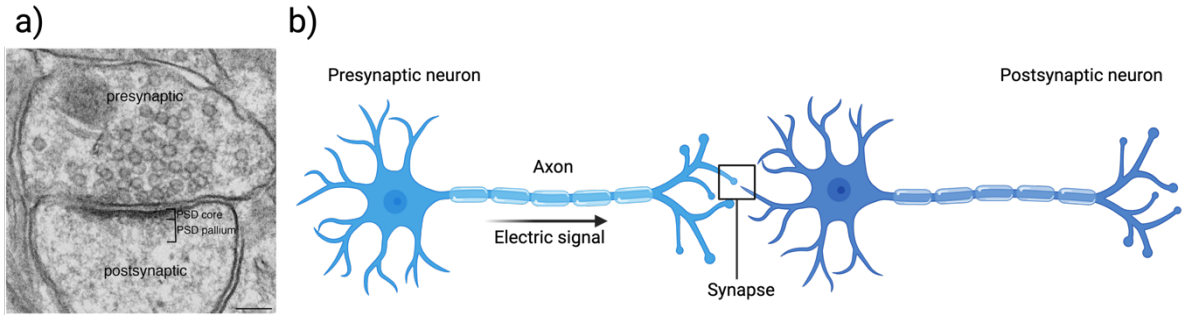

Figure 1.1: The post synaptic density (PSD). a) Electron micrograph of the PSD in an excitatory glutamatergic synapse, with highlighted core and pallium in 0.1 µm scale.[4] **b)** Neurons process and transmits cellular signals, where the axon carries the electric signal, and neurotransmitters are released at the synapse to communicate with the postsynaptic neuron.

1.2 PDZ domains

The PDZ domains are named after the proteins from which the PDZ sequence motif were originally identified, namely postsynaptic density protein 95 (PSD-95), discs large (DLG1) and zona occludens 1 (ZO-1). PDZ domains are specialized peptide-binding domains commonly present in multi-domain scaffolding proteins.[5] These proteins play crucial roles in trafficking, anchoring, and clustering glutamate receptors involving assembly and disassembly of signaling

protein complexes, facilitating synapse transmission in neurons.[2] Consequently, PDZ-domain containing proteins are associated with a range of diseases and disorders.[6]

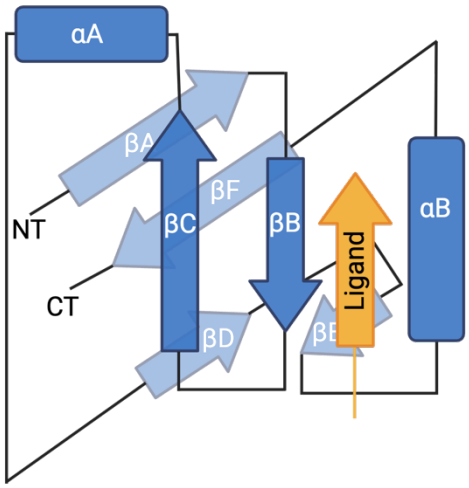

Figure 1.2: The PDZ domain's canonical fold. Schematic representation of a canonically folded PDZ domain (blue) binding to a ligand (orange), modified from Olsen *et al*.[6]

Moreover, PDZ domains are small modular protein entities constituted of 80-110 residues and were first discovered in the 1990s.[6] More precisely, 268 PDZ domains have so far been identified in 151 different human proteins.[6] All PDZ domains generally share a compact globular fold comprising six anti-parallel beta-sheets and two alpha-helices arranged in a βA-βB-βC-αA-βD-βE-αB-βF configuration. Other structural characteristics of PDZ domains are their closely positioned N- and C-terminal regions, which is believed to have contributed to the insertion of PDZ domains into larger proteins during evolution. PDZ domains typically recognize peptide motifs at the extreme C-terminus of their protein interaction partner. In particular, the first amino acid of the C-terminus (P_0) and the third amino acid (P_{-2}) of the peptide ligand dictate the binding to the PDZ binding pocket positioned between βB and αB secondary structures (Figure 1.2). Moreover, the importance of positions P_0 and P_{-2} has led to a classification of PDZ domains into three classes, based on different PDZ domain-binding motifs of their ligands (Table 1.2).[6]

Table 1.2: Classification of PDZ domains based on canonical binding modes. Three classes of PDZ domains have been determined based on the C-terminal PDZ binding motif, where X is any amino acid and Ψ a hydrophobic residue.

Class	Binding motif
I	T/S-X-Ψ-COOH
II	Ψ-X-Ψ-COOH
III	D/E-X-Ψ-COOH

1.3 PSD-95

One of the most abundant and thoroughly characterized scaffolding proteins within the PSD is the postsynaptic density protein 95 (PSD-95).[5] PSD-95 (UniProt: P78352) comprises 725 amino acids, belongs to the Discs large (DLG) family and is classified as a member of the membrane-associated guanylate kinase (MAGUKs) class of proteins present at excitatory synapses. The MAGUK family specifically interacts with NMDA receptors through the PDZ domains and play a crucial role in cell-cell adhesion and regulation of receptor function and recruitment.[7] Structurally, PSD-95 consist of three class I PDZ domains (PDZ1-3), followed by a Src homology 3 (SH3) domain and a guanylate kinase-like (GK) domain organized into the PDZ1-2 tandem and the PSG supramodule (Figure 1.3.1).[7]

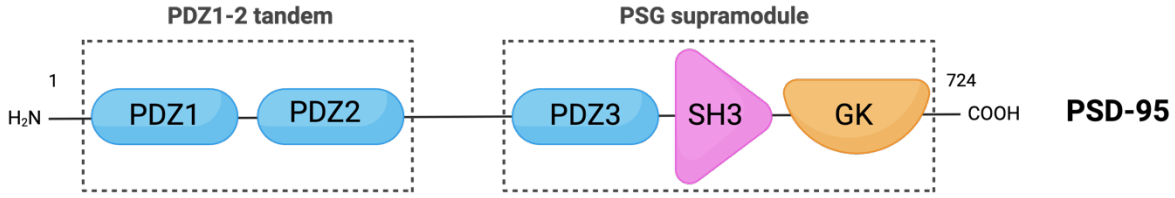

Figure 1.3.1: PSD-95. Schematic representation of PSD-95 consisting of three PDZ domains, a SH3 and a GK domain organized into the PDZ1-2 tandem and the PSG supramodule.

The biological function of PSD-95 includes the simultaneous binding to NMDA receptors and the enzyme neuronal nitric oxide synthase (nNOS) through its PDZ1 and PDZ2 domains (Figure 1.3.2).[8] More specifically, PSD-95 interacts with the C-terminus tail of the GluN2B subunit of NMDA receptors with either PDZ1 or PDZ2. Interaction with nNOS is established through the non-canonical binding of the nNOS β-hairpin motif based on 30 residues[9] into the PDZ2 domain's binding pocket.[7] The formation of the ternary NMDAR/PSD-95/nNOS complex couples the influx of Ca^{2+} ions caused by glutamate release and activation of the NMDA receptor with nNOS activation and the production of nitric oxide (NO).[7] PSD-95 not only interacts with various signaling enzymes, it is also linked with other scaffolding proteins within the PSD through protein-protein interactions (PPIs). Including guanylate kinase-associated protein (GKAP), SH3 and ankyrin repeat-containing protein (Shank), and Homer.[5] Subsequently, these proteins bind to additional signaling and cytoskeletal molecules forming a large network in the PSD.

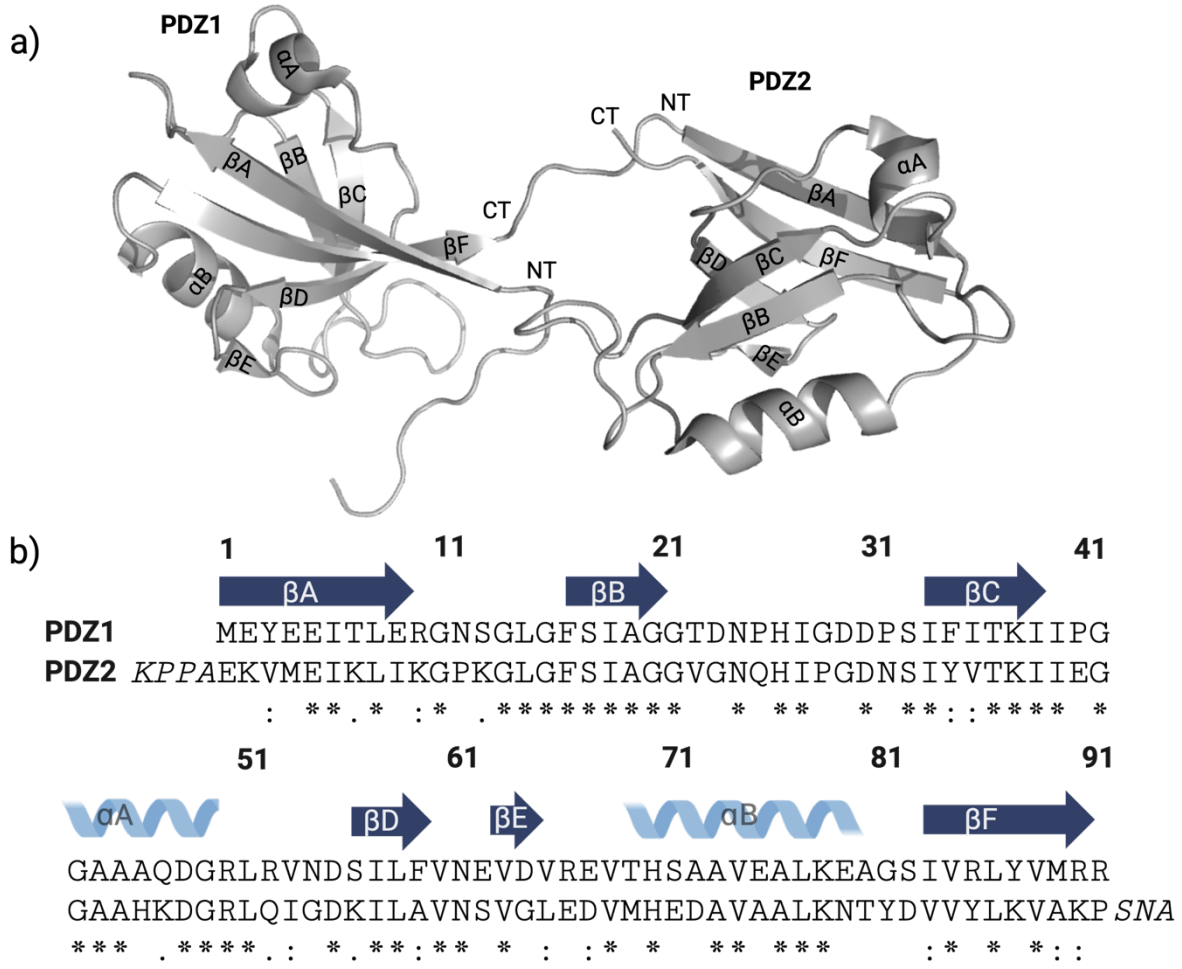

Figure 1.3.2: The PDZ1-2 tandem of PSD-95. a) Cartoon representation of the X-ray crystal structure of the human PDZ1-2 tandem of PSD-95 (UniProt: P78352, PDB: 3ZRT). **b)** Sequence alignment of the PDZ1-2 tandem with highlighted residues involved in their respective secondary structure. '*' is an identical residue, ':' a conservative mutation, '.' a semi-conservative mutation and blank a non-conservative mutation.

1.4 Targeting PSD-95 for treatment of acute ischemic stroke

Stroke is the second leading cause of death in the world, and a leading cause of long-term neurological disability as well as dementia.[10] During cerebral ischemia (stroke), blockage of blood vessels reduces perfusion to the affected brain tissue, which restricts oxygen and nutrient supply to the surrounding brain cells. In the infarct core, which suffers from severe or complete blockage, cells undergo irreversible damage or cell death (infarction). The penumbra, referring to the brain tissue surrounding the infarct core, suffers moderate blood flow reduction, resulting in functional impairment of its cells. If the blood flow is restored the brain-function can be recovered. However, after the immediate damage, the brain tissue suffers a secondary neuronal death, caused by excitotoxicity induced by overstimulation of NMDA receptors.[7] More specifically, the excess release of glutamate into the synaptic cleft leads to NMDA receptor over-activation, which results in abnormal influx of Ca^{2+} ions. Elevated levels of Ca^{2+} triggers

nNOS-mediated generation of cytotoxic NO in excess (Figure 1.4.1). NO is a highly reactive molecule and when generated in excess it facilitates toxicity by reacting with superoxide free radicals forming the oxidant peroxynitrite, which lead to the oxidation of proteins, DNA-damage and finally apoptotic neuronal cell death.[7] Monomeric and dimeric ligands binding to PDZ1-2 of PSD-95 can inhibit the ternary NMDAR/PSD-95/nNOS complex formation, uncoupling NMDAR Ca^{2+} influx activity and nNOS NO generation, and thereby circumventing excitotoxicity.[8] Since Ca^{2+} influx is unaffected by PSD-95 inhibition, and it is rather the downstream events of NMDA receptor activation that are affected, the PDZ1-2 tandem of PSD-95 is considered a particularly promising target for the treatment of acute ischemic stroke (AIS).[8] PSD-95 is actively pursued in the clinic as a drug target in the context of AIS[8, 11] (Bach *et al.*, Aarts *et al.*).

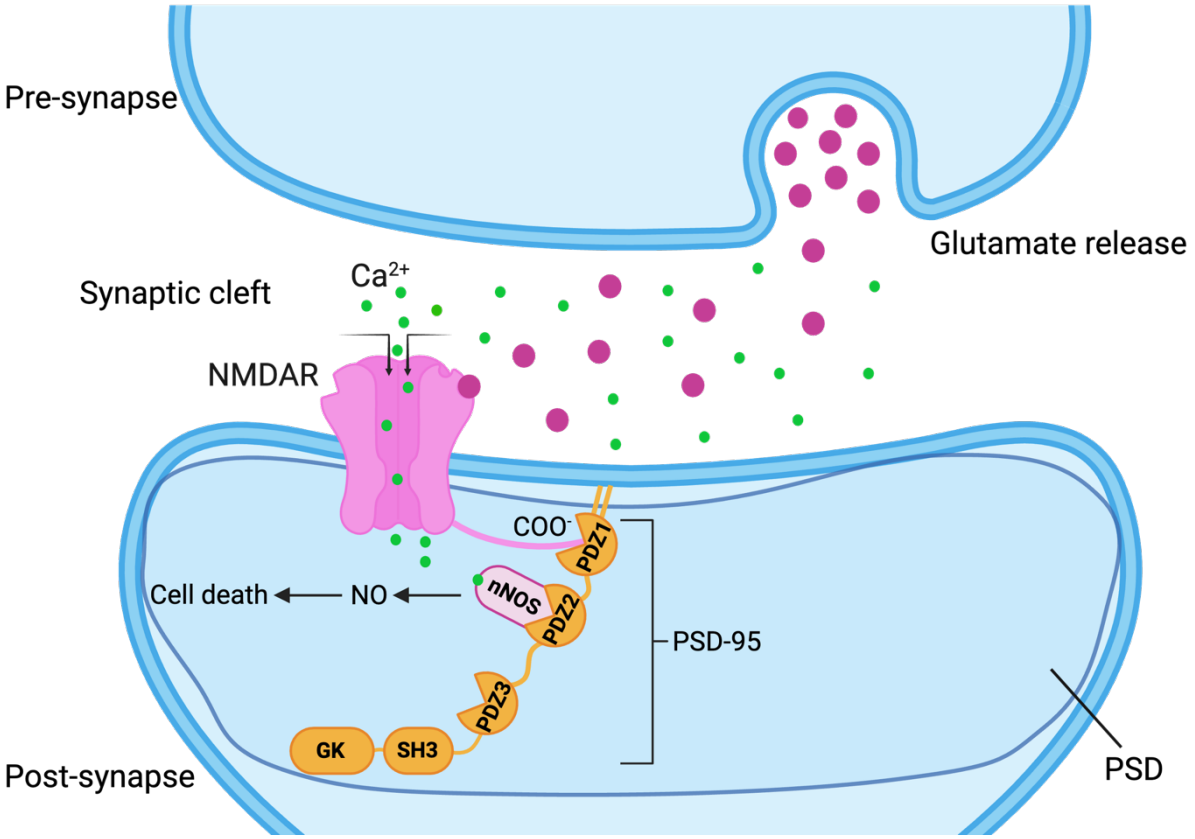

Figure 1.4.1: The role of PSD-95 in acute ischemic stroke. Overactivation of NMDAR due to excessive release of glutamate during stroke causes formation of the ternary NMDAR/PSD-95/nNOS complex, leading to nNOS activation and excitotoxicity. Figure modified from Bach *et al.*[8]

NERINETIDE also known as Tat-NR2B9c or NA-1, represents the first-in class compound targeting PSD-95 for AIS treatment.[11] The peptide developed in Toronto, Canada by NoNO Inc. was first reported in 2002. NERINETIDE is a 20-mer peptide composed of nine amino acids corresponding to the C-terminal of the GluN2B subunit of the NMDA receptor, fused to

the cell-penetrating peptide (CPP) moiety HIV-1 Tat peptide (Figure 1.4.2 a).[11] It binds to the PDZ-2 domains of PSD-95, thereby dissociating the NMDAR/PSD-95/nNOS complex, uncoupling NMDAR and nNOS activity.[11] In the Phase III clinical trial ESCAPE-NA1, NERINETIDE's neuroprotective effect was evaluated in addition to the thrombolytic drug alteplase (tissue plasminogen activator, tPA), which is the standard of care for patients with ischemic stroke.[12] NERINETIDE as not effective in patients also receiving alteplase. However, the group of patients only treated with NERINETIDE showed significant neuroprotective effects and showed reduced mortality.[12] In 2021 the lack of effect of treatment with NERINETIDE co-administered with tPA could be rationalized.[13] It was found that NERINETIDE is subject to proteolytic cleavage if administered after tPA, due to tPA's generation of the serine protease plasmin. However, besides being subjected to proteolytic cleavage[13], NERINETIDE suffers from low affinity (K_i = 4.6 μM) towards PDZ1-2 of PSD-95[8], limiting its bioactivity.

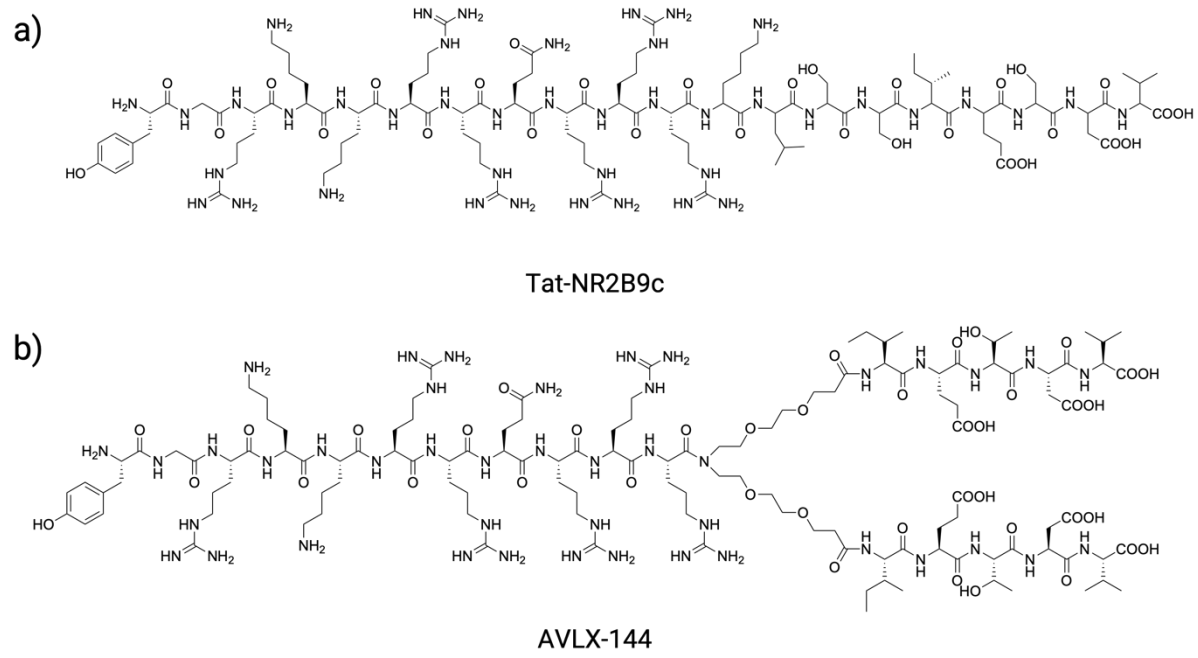

Figure 1.4.2: Monomeric Tat-NR2B9c and dimeric AVLX-144. Chemical structures of the PSD-95 inhibitors **a)** Tat-NR2B9c (NERINETIDE) developed by NoNO Inc. in 2002[11] and **b)** AVLX-144 developed by the Strømgaard group at University of Copenhagen in 2012.[8]

The dimeric peptide inhibitor AVLX-144 target the PDZ1-2 domains of PSD-95. AVLX-144 and improved second generation compounds have successfully been developed by the Strømgaard group at University of Copenhagen.[8] AVLX-144 is a Tat-*N*-dimer consisting of two 5-mer peptides (IETDV) linked to a monodisperse amino-polyethylene glycol (*N*PEG) linker, and the Tat (YGRKKRRQRRR) CPP moiety (Figure 1.4.2 b). The five amino acids correspond to the C-terminal of the GluN2B subunit of NMDA receptors. The dimeric peptide

binds to the PDZ1-2 domains simultaneously in a bivalent fashion. Thus, explaining the high affinity with a K_i value of 4.6 nM, determined in an *in vitro* fluorescence polarization (FP) assay with increasing concentrations of PDZ1-2. Interestingly, NERINETIDE exhibited an affinity of 4.6 μM (K_i value), confirming AVLX-144 having 1000-fold improved affinity compared to monomeric NERINETIDE. AVLX-144 has demonstrated enhanced stability in blood plasma and penetration of the blood-brain-barrier (BBB). The dimeric peptide was further examined for neuroprotective properties in an *in vivo* mice model of AIS. More precisely, a single intravenous injection (3.0 nmol/g) of AVLX-144 was injected in mice subjected to permanent middle cerebral artery occlusion (pMCAO). Injection occurred 30 min post-ischemic insult, and treated animals showed 40% reduction in ischemic tissue damage after 6 h and 37% after 48 h in the postsurgical survival period compared with saline injected mice. Post-ischemic motor functions like grip-strength were also significantly improved in the rodents. In contrast, NERINETIDE showed no significant reduction in infarct volume when tested in the same study using the same dosing regimen.[8] The neuroprotectant properties, enhanced plasma stability and affinity of AVLX-144 prompted further clinical assessment and a Phase I clinical trial has successfully been completed (NCT02930018).[14]

Despite partial clinical success of AVLX-144, the PDZ1-2 domain's limited binding specificity to this dimeric ligand remain a challenge.[15] In addition, its relatively poor BBB permeability and proteolytic stability are still problems to overcome.[8] A potential solution to increase binding affinity, stability profile and permeability involves the development of macrocyclic peptides.[16]

1.5 Macrocyclic peptides

Macrocyclic peptides are polypeptide structures bearing a ring structure, formed through various macrocyclization reactions linking head-to-tail or side-chain to side-chain through chemically stable bonds such as amide, thioether, or disulfide bonds, amongst others.[17] Macrocyclic peptides, ranging in size from 700 to 2000 Da[16], serve as a versatile drug modality, filling the gap between small molecules (up to 500 Da) and larger biologics such as antibodies (150 000 Da)[18] prevalent in the current pharmaceutical industry.[19] In contrast to traditional, well-established drug entities, peptides can target intracellular protein-protein interactions (PPIs). PPIs are protein surfaces considered “undruggable” for small molecules due to the lack of well-defined hydrophobic binding pockets and are inaccessible for antibodies. Macrocyclic

peptides, however, are capable of binding to these featureless protein surfaces with high affinity and specificity.[16] Similar to small molecules, macrocyclic peptides are synthetically accessible, allowing for lead optimization efforts through conventional medicinal chemistry methods.[16] Furthermore, linear peptides become unstructured in solution and only adopt its defined 3D-structure upon binding, leading to an entropic penalty. Their cyclic peptide counterparts are conformationally constrained, and their rigidity lowers entropic states, contributing to stronger and more selective interaction with target proteins by minimizing the entropic penalty upon binding.[20] Additionally, cyclic peptides are more resistant to hydrolysis of enzymes in the blood plasma compared with linear analogues due to the lack of N- and C-terminals.[17] However, compared with small molecules, cyclic peptide's proteolytic stability is still considered poor. Thus, introducing non-canonical amino acids into the structure will render it more resistant to proteolysis.[19] A major challenge for all peptide structures are their poor cell membrane permeability (CMP) and bioavailability, as most bioactive peptides generally violate Lipinski's rule of five, a criterion for CMP and bioavailability.[16] A rare example of an orally available cyclic peptide drug is the known structure cyclosporine A (Figure 1.5). Which demonstrates that the combination of peptide backbone modifications such as *N*-methylations with few OH-groups can enhance CMP and bioavailability by reducing the number of hydrogen bond donors (HBDs).[21]

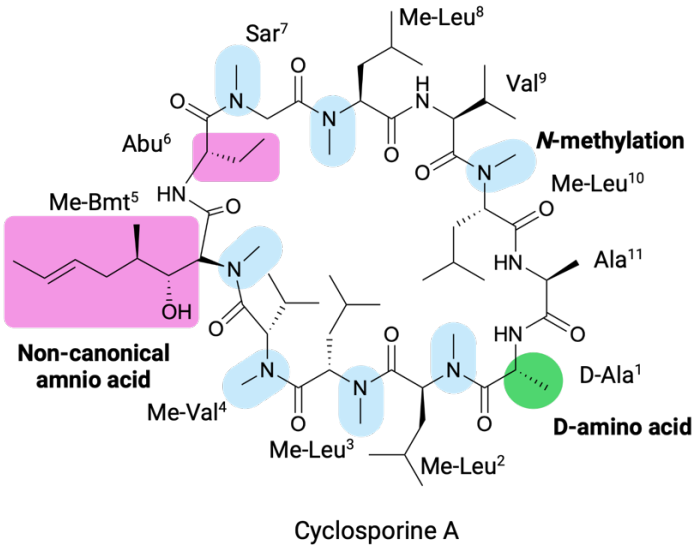

Figure 1.5: Structure of Cyclosporine A. Chemical structure of the orally available cyclic peptide drug Cyclosporine A with highlighted backbone *N*-methylations (blue), D-amino acids (green) and other no-canonical amino acids (pink).

Recent research by Tanada *et al*., highlights a highly *N*-alkylated and lipophilic 11-mer cyclic peptide with improved CMP, oral bioavailability, and with sufficient bioactivity targeting an intracellular protein. Thereby, introducing a cyclic peptide with drug-likeness and promising

therapeutic effects.[21] Additionally, Tanada *et al.* has established a list of criterions for the generation of drug-like hit peptides. Cyclic peptides should comprise nine to eleven amino acids, incorporate six or more *N*-alkylated residues in the backbone, exhibit a *C*LogP value of 12.9 or higher, and have no charged functional groups.[21] Research recently provided by the Strømgaard group in the context of targeting the PDZ1-2 tandem of PSD-95 for treatment of AIS, demonstrates a different approach in contrast to the linear GluN2B C-terminal derived peptides. The design of a 23-mer cyclic peptide mimicking the nNOS β-hairpin with incorporated non-canonical amino acids provides the most potent macrocyclic peptide so far towards a PDZ domain, and with high affinity compared to linear analogues (K_i = 30 nM towards PDZ2, determined by FP).[9] Further attempts to miniaturize this macrocyclic peptide have subsequently been made by the Strømgaard group. The RaPID mRNA display technology[22] was employed to develop a cyclic peptide comprising 17 amino acids (fIETTFAGNRIETSFTC) which exhibited an affinity of 500 nM (K_i value determined via FP) towards PDZ2 of PSD-95 (*unpublished results*, patent application pending).[a] Moreover, macrocyclic peptides as therapeutic modalities are attracting a lot of attention, and this new class of drugs will broaden the scope of therapeutic targets.[21]

1.6 *De novo* peptide drug discovery

Although macrocyclic peptides with non-canonical structures have proven to be promising therapeutics, the source from which they can be retrieved have historically been limited to methods such as mimicking hormonal peptides or extraction from natural sources. For more effective design of cyclic peptide drugs, high-throughput screening of peptide libraires is necessary. Advances of technology for discovery of *de novo* macrocyclic peptides have led to the construction and screening of genetically encoded macrocyclic peptide libraries of up to a trillion unique sequences.[22]

The development of the Random non-standard peptides integrated discovery (RaPID) system by the Suga laboratory at The University of Tokyo[22] allows for the discovery of more than 10^{12} cyclic peptides expressed *in vitro* trough an affinity-based selection towards a protein target. The RaPID system employs genetic code reprogramming using the flexible *in vitro* translation

[a] PhD thesis, Flora Alexopoulou, Postdoc, Department of Drug Design and Pharmacology, University of Copenhagen – Targeting protein-protein interactions using peptide technologies, November 2022 (was provided digitally)

(FIT) system. The FIT system includes the use of flexizymes which are flexible artificial ribozymes responsible for aminoacylating tRNA, together with mRNA display technology. Flexizyme have the capability to attach any amino acid to the 3'-end of the tRNA molecule, forming aminoacyl-tRNA. This process enables the incorporation of non-canonical amino acids, thereby reprogramming the genetic code and expanding the structural diversity of peptides produced trough translation.[22] One of the most renowned applications of the RaPID system involves preparation of libraries with synthesized peptides cyclized trough a thioether linkage reaction. In this process, the start codon of the peptide sequence encodes the non-canonical *N*-chloroacetylated (*N*-ClAc) amino acid instead of formyl methionine (fMet). Additionally, downstream amino acids in the sequence are randomized, with the terminal amino acid designated as Cys. Puromycin is subsequently linked to the 3'-end of each mRNA molecule, thereby terminating the elongation of the peptide chain, and resulting in peptides labeled with their respective mRNA template. Finally, the *N*-ClAc amino acid undergoes spontaneous posttranslational ring closure with the Cys residue to form the thioether linkage. In contrast to naturally derived macrocyclic peptides cyclized trough a disulfide bond, the thioether linkage in RaPID derived peptides is stable under physiological reducing conditions.[22] Moreover, the commercialization of the RaPID platform through the company PeptiDream Inc. established in 2006 in Tokyo, Japan has paved the way for macrocyclic peptide drug discovery.[23]

2. Objectives

The objectives for this project include the application of the RaPID technology with the *aim* to develop novel macrocyclic PSD-95 inhibitors comprising fewer than 17 amino acids. The specific objectives are:

I. A randomized mRNA library of cyclic peptides containing nine to thirteen amino acids will be prepared.

II. The library will be panned against the PDZ2 domain of PSD-95 with the aim to identify high affinity hits.

3. Methods

3.1 RaPID

3.1.1 Flexizyme technology and genetic code reprogramming

Flexizyme (Fx) is a 45-nucelotide[24] artificial ribozyme that catalyzes tRNA acylation, enabling preparation of acyl-tRNAs charged with non-proteogenic amino acids[25] such as D-amino acids, *N*-methyl amino acids or *α*-amino acids.[26] The acylation occurs at the 3'-hydroxyl group by Fx's ability of recognizing the 3' end ACC codon of the tRNA molecule.[25] In addition, Fx recognizes the aromatic side chain of amino acid derivatives activated by a cyano-methyl ester (CME), and nonaromatic amino acid derivatives activated by either dinitrobenzyl ester (DBE)[25] or amino-derivatized benzyl thioester (ABT).[27] Flexizyme structures are specific to their respective activated amino acid substrate and has hence received the names enhanced flexizyme (eFx), dinitro flexizyme (dFx)[25] and amino flexizyme (aFx)[27] respectively (Figure 3.1.1.1 a). After generation of the charged acyl-tRNA molecule, Fx can be separated from its attached tRNA by RNase (ribonuclease)[24], an RNA degrading enzyme.[28]

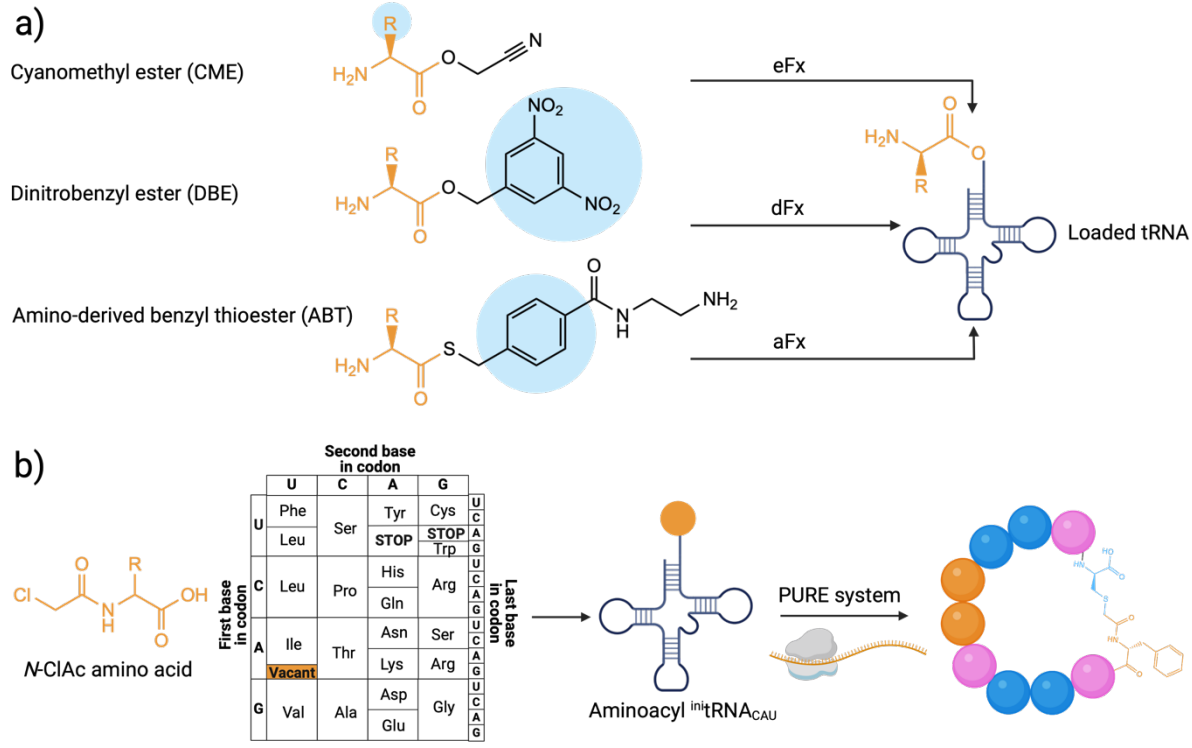

Figure 3.1.1.1: Flexizyme tRNA aminoacylation and RaPID synthesis of cyclic peptides. a) Esterified amino acids with activation groups CME, DBE and ABT are recognized by flexizyme, which catalyzes tRNA aminoacylation. **b)** Genetic code reprogramming of the initiation codon to *N*-ClAc modified amino acids and translation using the PURE system generates cyclic peptides cyclized posttranslationally with a downstream cysteine.

In *Escherichia coli (E. coli)* translation is initiated trough acylation of $^{ini}tRNA_{CAU}$ with formyl-methionine (fMet) with the help of ribozyme. In the RaPID system, the translation initiation event is genetic code reprogrammed to deplete fMet and reassigning the initiation codon to the desired non-proteinogenic amino acid, a technique first reported by Goto *et al.*[29] Combining genetic code reprogramming, flexizyme technology and the cell-free translation machinery, namely the protein synthesis using recombinant elements (PURE) system, provides efficient expression of peptides with non-canonical amino acids *in vitro* (Figure 3.1.1.1 b).[30] The PURE system, derived from *E. coli*, includes reconstituted purified translation factors such as initiation, elongation, and release factors, ribosomes, T7 RNA polymerase NTPS, tRNAs, aminoacyl-tRNA synthetases and more required for translation.[30] The PURE system, first reported in 2001 by Shimizu *et al.*[30], provides superior control over individual components, allowing for easy manipulation to achieve desired translational outcomes. In addition, the PURE system represents a key breakthrough in cell-free synthetic biology.[30] In the thioether cyclized peptide library using the RaPID system mentioned in section 1.6, the AUG initiation translation codon is reassigned to the *N*-ClAc modified amino acid, which is readily charged to $^{ini}tRNA_{CAU}$ by flexizyme (eFx), and the peptide chain is further translated using the PURE system. Subsequently, the synthetized peptide undergoes spontaneous posttranslational ring closure with a downstream Cys residue to form the nonreducible thioether linkage (Figure 3.1.1.2).[22]

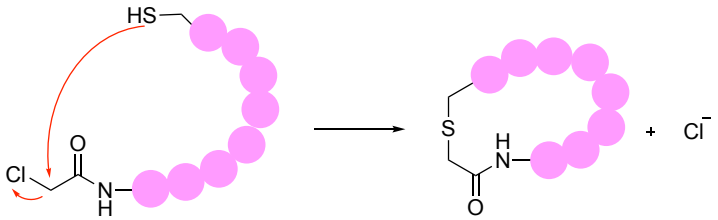

Figure 3.1.1.2: Peptide macrocyclization via head to side-chain cyclization in the RaPID system. Schematic representation of the head to side-chain cyclization strategy in the RaPID system with thioether ring-closure.

3.1.2 mRNA display

The RaPID system employs standard mRNA display technique for high-throughput affinity screening of the expressed peptide library. First, a complementary DNA (cDNA) library is transcribed to obtain the working mRNA (Figure 3.1.2.3, step 1).[31] The mRNA encodes the start codon, downstream randomized amino acids and a Cys residue for posttranslational ring closure.[22] Secondly, the mRNA is attached to a DNA spacer and a puromycin molecule

(Figure 3.1.2.1 b) linked together via a PEG linker at the mRNA's 3' end through enzymatic ligation (Figure 3.1.2.1 a and Figure 3.1.2.3, step 2).[31]

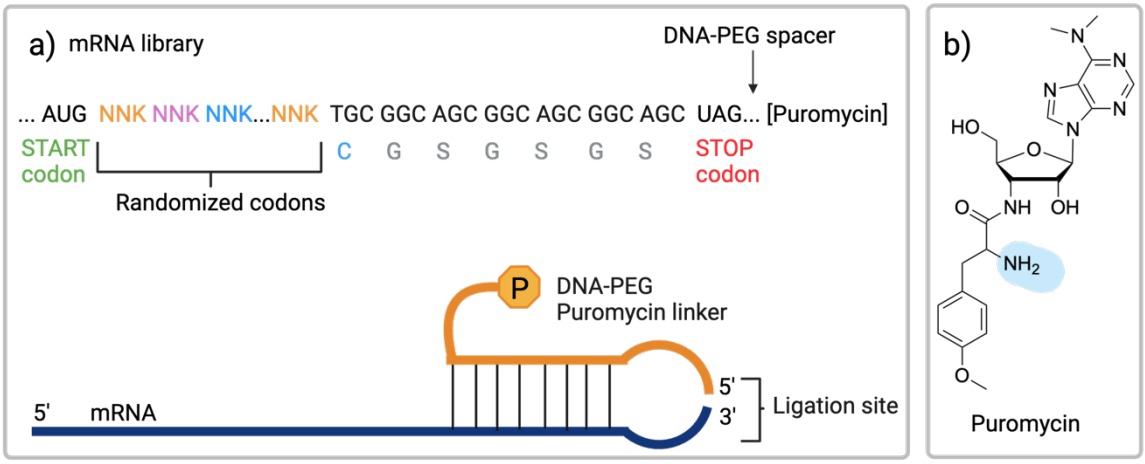

Figure 3.1.2.1: The mRNA library is linked with puromycin. a) Sequence design and schematic representation of the mRNA library in the RaPID system, showing how mRNA is linked with a DNA-spacer and a puromycin molecule via a PEG-linker. **b)** Chemical structure of a puromycin molecule with highlighted amine group (blue) that reacts with the RaPID-synthetized peptide.

Next, the ribosome in the PURE translation mixture reads mRNA from 5' to 3' and binds aminoacyl-tRNA carrying the encoded amino acid at its A-site. The ribosome's P-site binds peptidyl-tRNA which houses the translated peptide chain. Translation is halted at the stop codon (UAG), after which the ligated puromycin enters the A-site and becomes covalently attached to the translated peptide chain terminating the elongation (Figure 3.1.2.2).[31,22] The puromycin linkage to the peptide chain enables specific mRNA sequences to be enriched based on the properties of their encoded peptide.[31] The mRNA also acts as a barcode for the identification of expressed peptides in subsequent sequencing.[22]

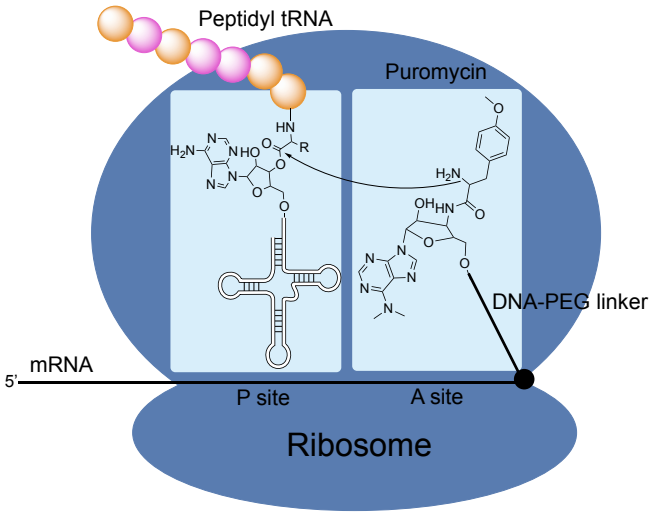

Figure 3.1.2.2: RaPID synthetized peptides are linked with their encoding mRNA sequence. The ribosome's P-site binds peptidyl-tRNA housing the peptide and entering of puromycin in the A-site mediates covalent linkage of cyclic peptides with their encoding mRNA sequence. Figure modified from Horiya *et al.*[32]

After translation and puromycin ligation (Figure 3.1.2.3, step 3) the mRNA-peptide fusion is reverse transcribed creating a library of cDNA-mRNA duplexes attached to their encoding peptide (Figure 3.1.2.3, step 4). Reverse transcription is followed by a negative affinity selection towards solely magnetic beads, to eliminate peptide species that binds to the beads (Figure 3.1.2.3, step 5). The negative selection is followed by a positive affinity selection towards the target protein immobilized to the magnetic beads (Figure 3.1.2.3, step 6).[31] Typically, streptavidin magnetic beads, specifically streptavidin Dynabeads are used. The beads are 2.8 μm in size, monodisperse, magnetic and hydrophobic, covalently linked to streptavidin.[33] The target protein of interest is readily biotinylated, which involves the attachment of biotin at the protein's C-terminus via an AviTag (a 15-mer peptide, GLNDIFEAQKIEWHE).[34] The streptavidin beads bind to biotin, immobilizing the protein[33] enabling pulldown of binding peptides.[33] Following affinity selection, cDNA-mRNA duplexes bound to their encoding peptide are eluted from the protein-bead mixture through heat treatment. Next, the cDNA sequences are recovered and amplified using PCR, completing one round of mRNA display selection (Figure 3.1.2.3, step 7). Repeated rounds of mRNA display selection are performed until a significant enrichment of sequences and thereby enrichment of active peptide species is observed.[31] Upon reaching a satisfactory number of selection rounds, the cDNA templates encoding the active peptide species are sequenced using Next Generation Sequencing (NGS) to identify potential RaPID-derived macrocyclic peptide ligands (Figure 3.1.2.3, step 8).[22]

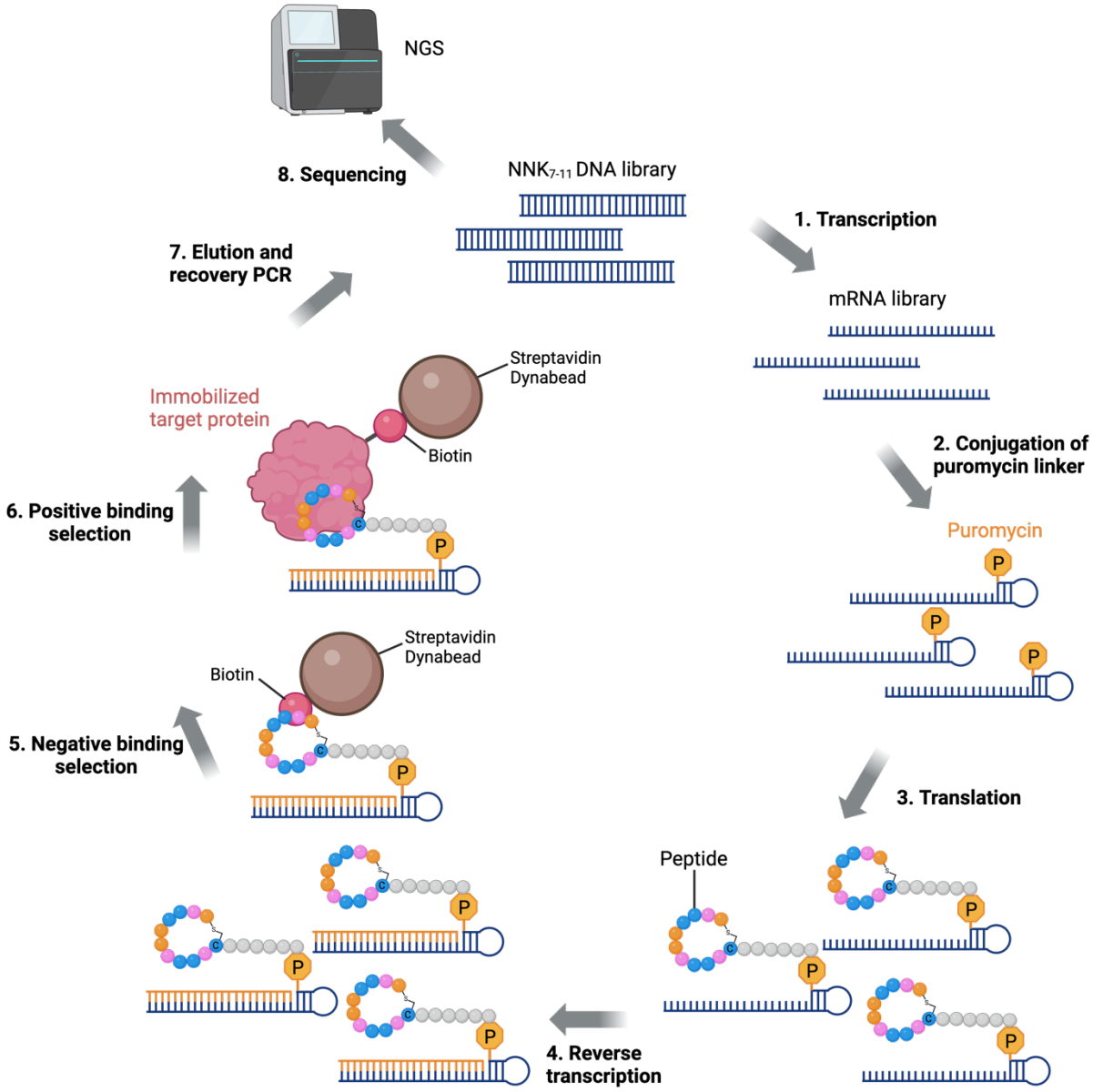

Figure 3.1.2.3: mRNA display cycle in the RaPID system. Schematic overview of mRNA display used in the RaPID system. A cDNA library is transcribed into mRNA, which is linked to puromycin, and attached to the depicted peptide upon translation. mRNA sequences are reverse transcribed, panned against magnetic beads and next panned against the target protein which is immobilized. After elution of hit peptides cDNA are recovered, amplified and sequenced with NGS identifying the hit peptide sequences.

3.2 Next generation sequencing (NGS)

Next generation sequencing is a technology used to determine DNA or RNA sequences, which was introduced for commercial use in 2005.[35] In contrast to traditional Sanger sequencing, where DNA strands are sequenced one at a time, NGS allows for the parallel sequencing of thousands of genes or whole genomes in a short period of time.[36] Most commonly used NGS system is the Illumina sequencing platform. This method operates on a sequencing-by-synthesis principle, enabling real-time detection of incorporated bases. Initially,

a library of DNA or cDNA (for RNA sequencing) fragments of specific size is prepared through physically, enzymatically, or chemical fragmentation. Adaptors are then ligated to both ends of each fragment through annealing. Subsequently, fragments are immobilized on a flow cell surface via base pairing between the adaptors and oligonucleotides on its surface. Following amplification with PCR, clusters of identical sequences are formed. The cyclic sequencing of these clusters is achieved with the sequencing-by-synthesis methodology, where fluorescently labeled nucleotides are attached to the fragments. As each base is incorporated, it emits a unique fluorescent signal which is captured by a detector, enabling determination of the sequence of each cluster.[37]

3.2 General methods

3.2.1 SDS-PAGE

SDS-PAGE (sodium dodecyl-sulfate polyacrylamide gel electrophoresis), first reported in the early 1970s[38], is a commonly employed method for effectively separating protein mixtures with high resolution. Proteins are initially denatured and coated with an anionic detergent that imparts a negative charge to them, proportional to their molecular mass. Subsequently, these negatively charged proteins are separated in gel electrophoresis based on their respective molecular masses as they migrate through a porous acrylamide gel matrix.[38]

3.2.2 Agarose gel electrophoresis

Agarose gel electrophoresis is a reliable method for separating DNA fragments of different sizes, typically ranging from 100 bp to 25 kbp. DNA samples are loaded into pre-cast wells within a gel and an electric current is applied. The phosphate backbone of DNA is negatively charged, causing DNA fragments to migrate towards the positively charged cathode under the influence of the electric field. Due to DNA's uniform mass-to-charge ratio, fragments are separated by size in a pattern such that the distance traveled by each fragment is inversely proportional to the logarithm of its molecular weight.[39]

3.2.3 FPLC

Fast protein liquid chromatography (FPLC) is a high-performance and resolution liquid chromatography method developed in 1982[40] by the Swedish company Pharmacia. Originally designed for proteins, FPLC is also applicable to other biological samples including oligonucleotides such as DNA and RNA, and plasmids. Similarly to other chromatography

methods, FPLC separates components based on their affinity towards the mobile phase (fluid) and stationary phase (porous solid). In FPLC, the mobile phase consists of an aqueous buffer, regulated by a pump with varied composition sourced from two reservoirs: buffer A (running buffer) and buffer B (elution buffer). The stationary phase contains resin beads, typically made of cross-linked agarose, packed into a column. Applying a gradient (by increasing the elution buffer concentration) causes bound proteins to elute from the stationary phase and pass through a UV-detector. The UV-detector measures protein concentration by detecting absorption of UV light at 280 nm. Each protein appears as a distinct peak on the chromatogram, allowing for collection.[40]

4. Results and discussion

4.1 Discovery of first- and second-generation cyclic PSD-95-PDZ2 inhibitors

As shown by Balboa *et al*.[9] developing macrocyclic peptides mimicking the nNOS β-hairpin by incorporating non-canonical amino acids (ncAAs) is a promising avenue of drug development in the context of inhibiting the nNOS/PSD-95 interaction. The purpose of inhibiting this PPI is to protect against brain damage caused by acute ischemic stroke (AIS), a disease responsible as the second leading cause of death in the world.[10] A first-generation cyclic PSD-95 inhibitor have thereby been developed by the Strømgaard group. Where a 23-mer nNOS β-hairpin mimetic peptide (Table 4.1) identified the ^{107}LETTF^{111} motif as potentially important for the PDZ2 binding interface. This compound exhibited an affinity towards PDZ2 of PSD-95 of 30 nM (determined by FP).[9] In an attempt to identify second-generation cyclic peptide PSD-95 inhibitors, the RaPID platform was applied in a collaboration between the Strømgaard and the Rogers groups. Employing this powerful mRNA display technology, PhD student Flora Axelopoulou identified a potent cyclic PSD-95-PDZ2 inhibitor (Figure 4.1 A). Evaluation of binding properties towards PDZ2 of the RaPID-derived peptides was established by FP assay. The obtained peptide, JFL03, a 17-mer macrocyclic peptide (Table 4.1) exhibited relatively high binding affinity towards PDZ2 of PSD-95 (K_i = 500 nM) (Figure 4.1 B, C, D and E) (*unpublished results*, patent application pending)[b]. A RaPID deep mutational scan of JFL03 was additionally performed in efforts to identify critical residues for

[b] PhD thesis, Flora Alexopoulou, Postdoc, Department of Drug Design and Pharmacology, University of Copenhagen – Targeting protein-protein interactions using peptide technologies, November 2022 (was provided digitally)

binding. It was found that substitutions in most positions, including the "TTF" motif, was deleterious for binding. However, substitutions in positions P_7, P_8, P_9 and P_{16} were tolerated to some degree (Table 4.1). These studies were used as a starting point for the identification of smaller macrocyclic peptides targeting PDZ2 of PSD-95. Where the aim was to identify macrocycles exhibiting similar affinities towards PSD-95-PDZ2 as previously discovered peptides, while being more permeable.

Table 4.1: First- and second-generation cyclic PSD-95-PDZ2 inhibitors. Sequence alignment of the first-generation nNOS β-hairpin peptide and the second-generation JFL03 peptide inhibitors of PSD-95-PDZ2. Residues in bold are identified as crucial for binding.

Peptide	Length (AA)	K_i	Sequence
nNOS	23	30 nM	^{105}TH**LETTF**T**G**DGT**P**K**T**IRVTQ^{124} – pGQ$_\alpha$
JFL03	17	500 nM	f**IETTF**AGN**RIETSF**TC

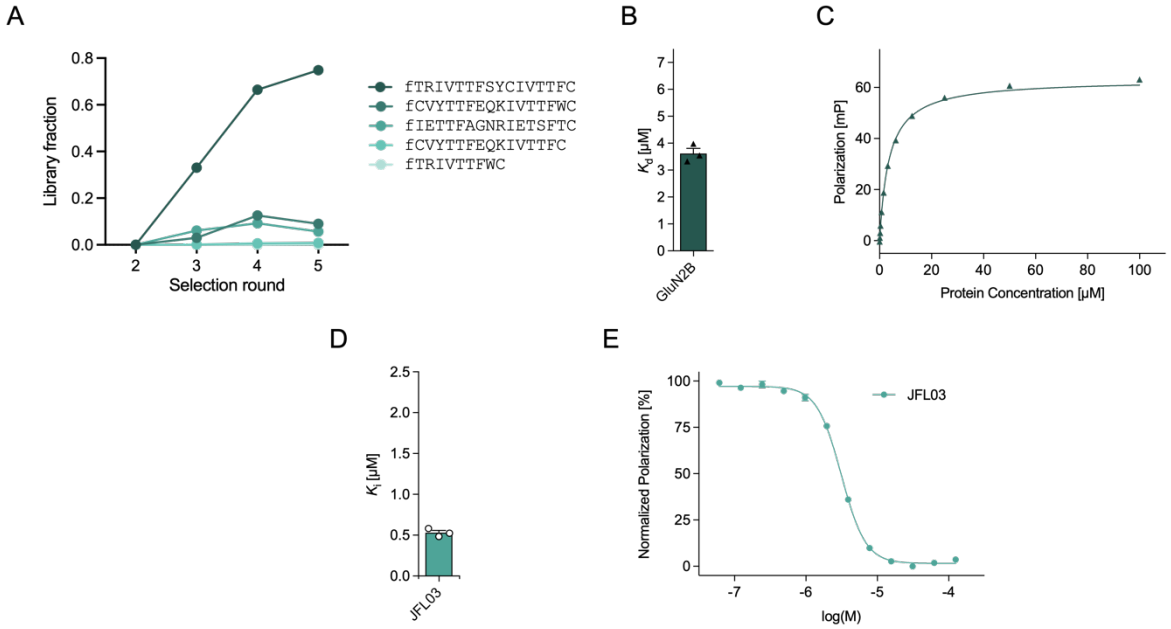

Figure 4.1: RaPID selection against PSD-95-PDZ2 generated a 17-mer second-generation cyclic inhibitor exhibiting potent binding determined by FP assay. A) NGS data analysis of the positive selection for rounds two to five, comparing the fraction of DNA reads for each individual peptide sequence. **B) C)** A FP assay was performed for determination of the binding affinity of the TAMRA-GluN2B C-terminal peptide (C_{probe} = 5 µM) to PSD-95-PDZ2, exhibiting K_d = 3.6 µM with data presented as mean ± SEM, n = 3. **D)** K_i value of the JFL03 peptide towards PSD-95-PDZ2. Data is presented as mean ± SEM, n = 3. **E)** Concentration-response curve for the RaPID-derived JFL03 peptide measured in an FP assay for affinity against PDZ1-2 (C_{PDZ1-2} = 5 µM) using the TAMRA-GluN2B probe (C_{probe} = 50 nM).

Research recently conducted by Othi *et al.*[41] and Tanada *et al.*[21] focused on cyclic peptides with the aim to determine the structural features granting peptides drug-likeness. The research

demonstrated that 11-residue peptides show better possibility to meet their drug-like criteria. Validation of their drug discovery concept was made by mRNA display towards the intracellular protein KRAS, identifying an 11-mer with affinity for KRAS. Additionally, Merck & Co.[42] has developed other mRNA display-based peptides with potential therapeutic effects. They reported an orally available peptide-based inhibitor of proprotein subtilisin/kexin type 9 (PCSK9), which represents one of the first clinical candidates derived from an mRNA display technique. Specifically, Merck[42] screened 7-12 mer peptide libraries against PCSK9.

In this thesis the RaPID platform was thereby applied for high-throughput screening against PSD-95-PDZ2 with a peptide library designed to contain nine to thirteen AAs. Having the optimal length of eleven amino acids centered, thought to have improved permeability, bioavailability and bioactivity. In contrast to the previously identified JFL03 peptide inhibitor of PSD-95-PDZ2 (Table 4.1) the initiation codon in the screened library were genetic code reprogrammed to encode L-Phe instead of D-Phe. The overall aim of the thesis was to identify smaller potentially therapeutic peptide-based modalities for AIS treatment.

4.2 Determination of PSD-95-PDZ2 bead-binding properties

Macrocyclic peptide discovery using the RaPID system requires the target protein to be purified and labeled with an affinity tag to efficiently sort out peptide binders and non-binders. Biotin is commonly used and are attached to the proteins C-terminal via an AviTag, due to biotin's high affinity to ($K_D = 10^{-15}$ M)[43] and slow dissociation from streptavidin functionalized beads.[43] Therefore, prior to RaPID screening, recombinant protein expression, purification as well as biotin labeling of the PSD-95-PDZ2 protein construct was performed.[c] Additionally, the protein's degree of biotinylation as well as loading capacity to the streptavidin functionalized beads was determined. Both parameters are important for performing the RaPID selection and hence a bead binding assay, in which a fixed concentration of protein was incubated with increasing amount of bead slurry was performed. Bound and unbound protein samples were run in an SDS-PAGE gel together with individual protein and beads samples. The volume of beads required to reach protein saturation was determined. From the gel, the biotinylation percentage of the PDZ2 domain was estimated to be 80%, and the protein loading capacity calculated to 1 pmol/µL (Figure 4.2).

[c] Protein expression, purification and biotin labeling of PSD-95-PDZ2 was performed by Fabian Hink, Postdoc, Department of Drug Design and Pharmacology, University of Copenhagen

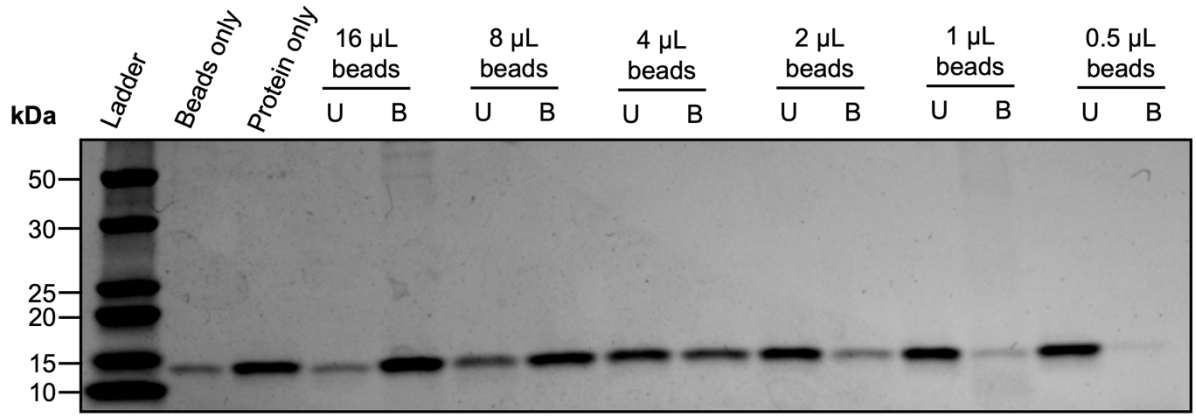

Figure 4.2: SDS-PAGE analysis determined PDZ2-Avi-BTN's bead binding properties. The bead binding assay showed that PSD-95-PDZ2 is 80% biotinylated (C_{PDZ2} = 2 µM) and that the beads are saturated in the 16 µL bead sample (DynaBeads™ M-280 Streptavidin) determined with SDS-PAGE analysis.

More precisely, the degree of biotinylation for PSD-95-PDZ2 could be estimated by comparing the 'protein-only' sample with the 'unbound-proteins' from the 16 µL bead slurry, resulting in the rough estimation of 80%. Saturation of beads occurred already at 16 µL bead slurry (Figure 4.2). Since 20 pmol 80% biotinylated protein was loaded onto the beads, the loading capacity was calculated to 1 pmol protein per µL bead slurry. Resulting values could further be used for determination of bead slurry volume and amount of protein needed in the RaPID selection. Values for the positive selection in round one was calculated to 106 µL bead slurry and 397.5 pmol protein. For subsequent rounds 3.6 µL bead slurry and 13.5 pmol protein would be needed.

4.3 RaPID PSD-95-PDZ2 selection

In preparation for the RaPID selection, a DNA library encoding for seven to eleven randomized amino acids (NNK7-11) was prepared. Randomized codons N = A/C/G/T and K = G/T allow for a variability of $(4 * 4 * 2)^{NNK\ size}$ of each NNK library. Smaller NNK library sizes comprise a higher number of molecules when present in the same concentration as the larger libraries. Therefore, in the final library NNK7 was included with a ratio of 0.1, while all possible DNA sequences of NNK8-11 was present. Further on, genetic code reprogramming of the initiator amino acid enables the expression of macrocyclic peptides as described in section 3.1.1. To do so, both flexizyme (eFx) and $^{ini}tRNA_{CAU}$ were synthesized using chemical biology methods, purified by FPLC and RNA extraction methods to obtain final concentrations of 300 µM. Subsequently, *N*-terminally ClAc-L-Phe-CME was loaded onto $^{ini}tRNA_{CAU}$ by eFx during a 2 h reaction. The readily aminoacetylated ClAc-L-Phe-$^{ini}tRNA_{CAU}$ molecules were included in the *in vitro* PURE translation system, containing 19 amino acids and excluding Met

and the formyl donor. After which the *in vitro* translation machinery incorporated ClAc-L-Phe in position P_0 in the translated peptide sequences. The terminal amino acid was encoded to Cys and peptides of various length consequently cyclized post-translationally via a thioether-bond as described in section 3.1.1.

The application of an mRNA display technology enabled attachment of a DNA-PEG-puromycin tag to the 3' end of the mRNA library sequences. This was followed by puromycin-mediated linkage of the cyclic peptides to their respective encoding mRNA sequence during translation using the PURE system.[30] The mRNA-peptide fusions were reverse transcribed yielding mRNA-cDNA duplexes. The library was incubated with immobilized biotin-tagged PSD-95-PDZ2 for a positive affinity screening. Unbound peptides were washed away and bound peptides heat-eluted, which was followed by qPCR analysis to quantify the number of recovered DNA molecules. The DNA library was recovered and amplified using PCR, then transcribed to mRNA for another RaPID selection round. Six subsequent selection rounds were conducted, where prior to the positive affinity screening a negative affinity selection was included. In the negative selection the library was incubated five times with solely streptavidin-functionalized magnetic beads to eliminate non-binders. qPCR analysis of the input, positive, and negative selection libraries was employed to monitor the progress of peptide enrichment by calculating the library recovery rate.

qPCR analysis of each round of the RaPID selection showed a notable enrichment of peptide encoding cDNA sequences by round five, as indicated by the increased amount of recovered DNA library in the positive selection compared to the negative selection (Figure 4.3.1 A). Implying that peptide species present are more likely to bind to PSD-95-PDZ2 than the streptavidin functionalized beads. The highest level of enrichment was observed in round six, with nearly 8% of the peptide library being recovered (Figure 4.3.1 A). These findings indicate the presence of potential PSD-95-PDZ2 binders within the library. In round seven, a significant decrease in the percentage of recovered library in the negative selection was observed, visualized more clearly in Figure 4.3.1 B, further supporting the presence of potential binders of PSD-95-PDZ2 within the library.

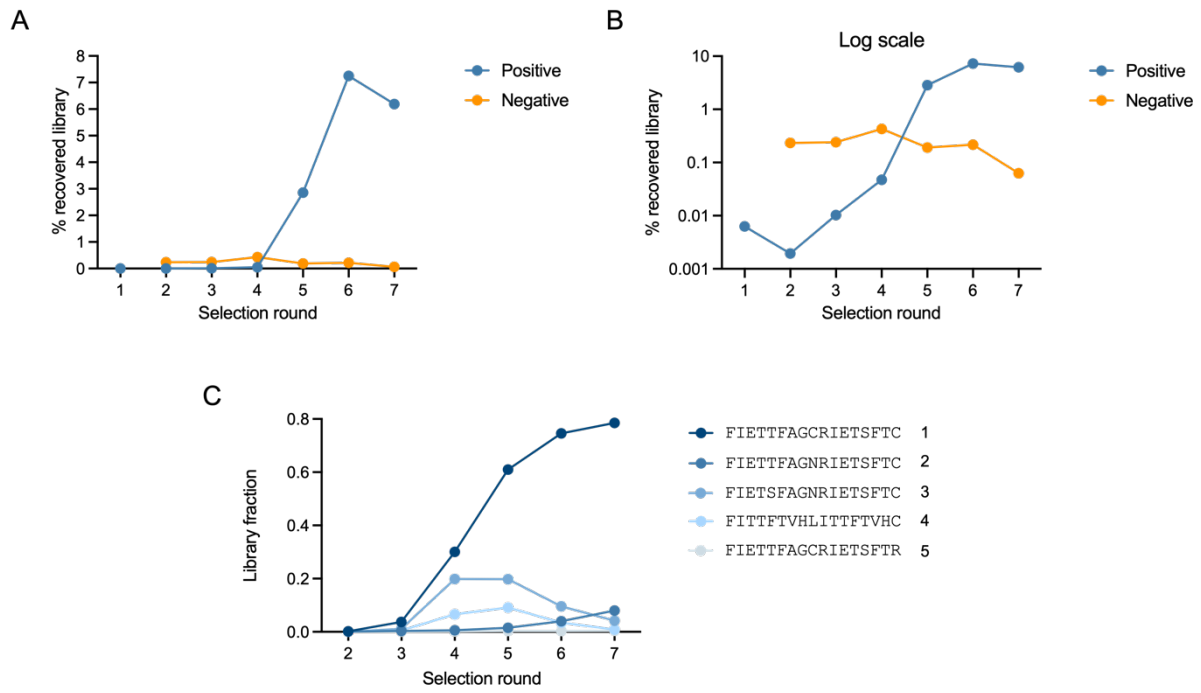

Figure 4.3.1: RaPID selection against PSD-95-PDZ2 enriched cyclic 17-mers similar to second-generation compounds. **A)** Comparison of cDNA recovery rates between positive and negative selection determined with qPCR for all seven RaPID rounds against PSD-95-PDZ2. **B)** Comparison of logarithmic cDNA recovery rates between positive and negative selection determined with qPCR for all seven RaPID rounds against PSD-95-PDZ2. **C)** NGS data analysis of the positive selection of rounds two to seven, comparing the fraction of DNA reads for each unique peptide sequence.

After completing seven rounds the RaPID selection was stopped and positive and negative library samples from rounds two to seven (POS) and three to seven (NEG) were prepared and sent for Illumina NGS sequencing at Azenta Life Sciences, where 10.000 reads were performed to identify the enriched peptide sequences present in the library. The fastq files provided by Azenta were processed and analyzed using a python script (Miniconda Python 3.8 MacOSX pkg)[d]. Surprisingly, decoding of the cDNA sequences revealed the enrichment and identification of peptide species comprising 17 amino acids (Figure 4.3.1 C). The initial library was designed to contain 9 to 13 amino acids (NKK7-11), suggesting that the library have been contaminated with also longer templates (e.g NKK15). The presence of longer peptide species could also be attributed to the accumulation of PCR product artefacts during amplification of the library after each selection round. Multiple PCR cycles were performed (up to 20), which can increase the PCR bias and can lead to accumulation of *Taq* DNA polymerase errors.[44] Interestingly, the NGS results revealed that the library is dominated by one peptide species present in a library fraction as high as 0.8 (Figure 4.3.1 C), indicating that this peptide is a

[d] Python scripts were provided by Associate Professor Joseph M. Rogers, Department of Drug Design and Pharmacology, University of Copenhagen

potential PSD-95-PDZ2 binder. Additionally, sequencing results revealed that the top-hit peptide comprise identical residues as the previously discovered second-generation JFL03 peptide, except for the initiator amino acid designated as L-Phe instead of D-Phe and an internal cysteine. It thus appears that the N-terminal ClAc group can cyclize with two different downstream Cys residues, forming either a cyclic or a lariat peptide structure (Figure 4.3.2). Lariat peptide structures are known to be more prone to degradation as its C-terminal is exposed and not protected by the constrained cyclic structure.[17] However, there are existing lariat structures exhibiting therapeutic effects.[45] Which does not completely rule out the top-hit peptide as a potential therapeutic compound.

A study by Iwasaki *et al.*[46] demonstrated that for peptides containing the N-terminal ClAc group and two competing Cys residues expressed *in vitro* under the reprogrammed genetic code, the nearest upstream Cys generally dictates the spontaneous thioether bond formation. Thus, leaving the downstream Cys unreacted. One exception, however, was that if the Cys is embedded in position two, then the downstream Cys predominantly reacts with the ClAc group to form the macrocycle. These findings suggests that the top-hit peptide is a lariat structure (Figure 4.3.2 A) and will likely not form the macrocycle. However, Iwakasi *et al.* performed experiments with a very flexible peptide (FTYSHFGPCLTWVKPQC). This contradicts the previous statement regarding a preferred lariat structure, since the RaPID-derived top-hit peptide is likely not as flexible. Therefore, to be able to draw any profound conclusions the structure-preference need to be addressed with further experiments.

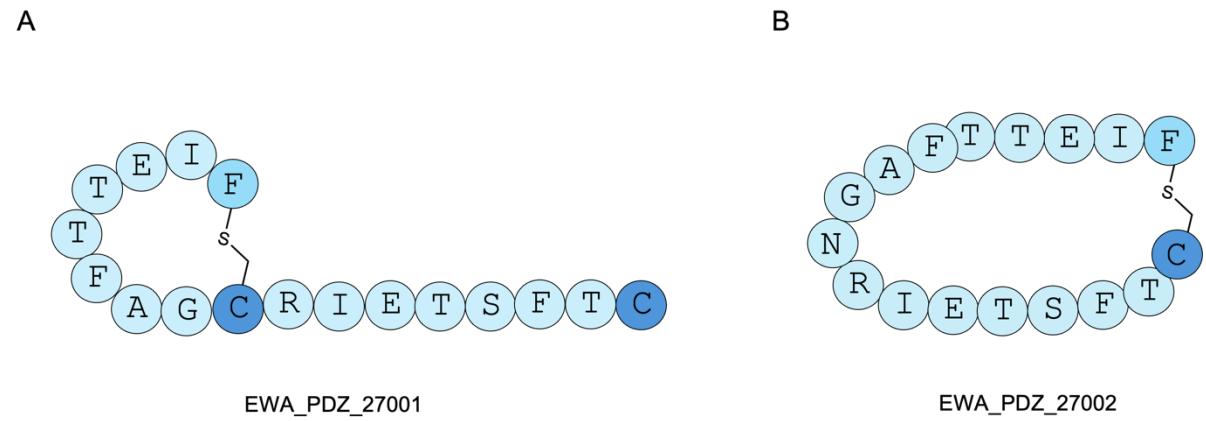

Figure 4.3.2: Structures of the two most enriched peptide sequences in the RaPID selection against PSD-95-PDZ2. A) Structure of the top-enriched peptide in the RaPID selection cyclized with its internal cysteine. **B)** Structure of the second most enriched peptide in the RaPID selection, identical to the second-generation cyclic scaffold[e] except for L-Phe in position P_0 due to different genetic programming approaches.

Notably, the second most enriched peptide from the RaPID selection (Figure 4.3.1 C) comprise completely identical amino acid residues as the previous RaPID-derived 17-mer (JFL03). Except for the AA in position P_0 designated as L-Phe (Figure 4.3.2 B) instead of D-Phe, due to different genetic code reprogramming approaches.

Further on, a heat map was created from the NGS data to visualize the enrichment of the top ten hit peptide structures from the RaPID selection across each individual round (Figure 4.3.3 A). The color coding corresponds to the number of reads of each peptide sequence, where a darker color represents a higher number of reads. The NGS data reveal dynamic changes in peptide enrichment throughout the selection. However, all top-hit peptide seqeunces are 17-mers and no shorter peptides were present amongst the top ten. Although, NGS analysis revealed smaller peptides were present in the libraires amongst the first RaPID rounds.

[e] PhD thesis, Flora Alexopoulou, Postdoc, Department of Drug Design and Pharmacology, University of Copenhagen – Targeting protein-protein interactions using peptide technologies, November 2022 (was provided digitally)

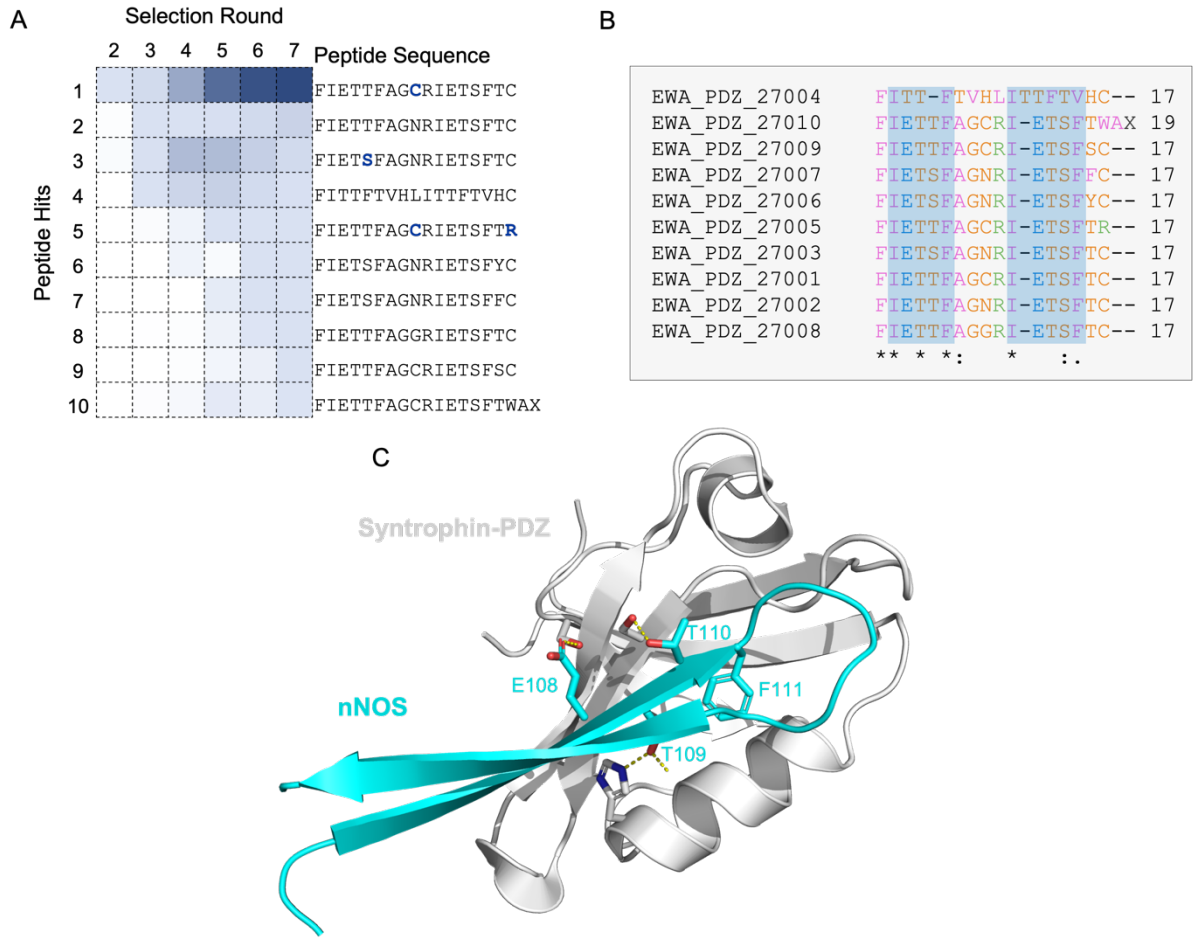

Figure 4.3.3: RaPID selection against PSD-95-PDZ2 enriched peptide sequences related through the literature-described "IETTF" motif[9] and an "IETSF" motif. A) Heat map of the 10 top-hit peptides in the RaPID selection against PSD-95-PDZ2 show dynamic changes in peptide enrichment. **B)** Sequence comparison performed with the Clustal Omega tool[47] between the 10 top-enriched peptide in the RaPID PSD-95-PDZ2 selection show two frequent binding motifs "IETTF" and "IETFS". '*' is an identical residue, ':' a conservative mutation, '.' a semi-conservative mutation and blank a non-conservative mutation. **C)** Cartoon representation of the X-ray crystal structure of rat syntrophin-PDZ (grey) binding to the nNOS peptide (cyan), where residues found to be important for binding are presented as sticks (PDB: 1QAV).[48]

Noteworthy, the majority of expressed hit peptide sequences are related through a "IETTF" motif, which previously has been shown important for binding of the nNOS β-hairpin to PDZ2 of PSD-95 by the Strømgaard group.[9,f] Multiple sequence alignment using the publicly available Clustal Omega tool[47] was applied to further investigate motif-relations. Confirming that most expressed sequences are related through a "IETTF" motif (Figure 4.3.3 B), similar to the nNOS β-hairpin mimetic peptide's ^{107}LETTF^{111} internal binding motif.[9] Sequence alignment also reveals a relation via an "IETSF" motif (Figure 4.3.3 B), which have similar properties as the 5-mer motif in the dimeric AVLX-144 compound (IETDV).

[f] PhD thesis, Flora Alexopoulou, Postdoc, Department of Drug Design and Pharmacology, University of Copenhagen – Targeting protein-protein interactions using peptide technologies, November 2022 (was provided digitally)

4.4 Peptide library diversity

A bubble diagram was created to investigate how the diversity of the library related to peptide length changes throughout the rounds of the RaPID selection. The top 100 enriched peptide sequences (starting with Phe and excluding stop codons) in each round were analyzed based on length and NGS frequency (Figure 4.4).

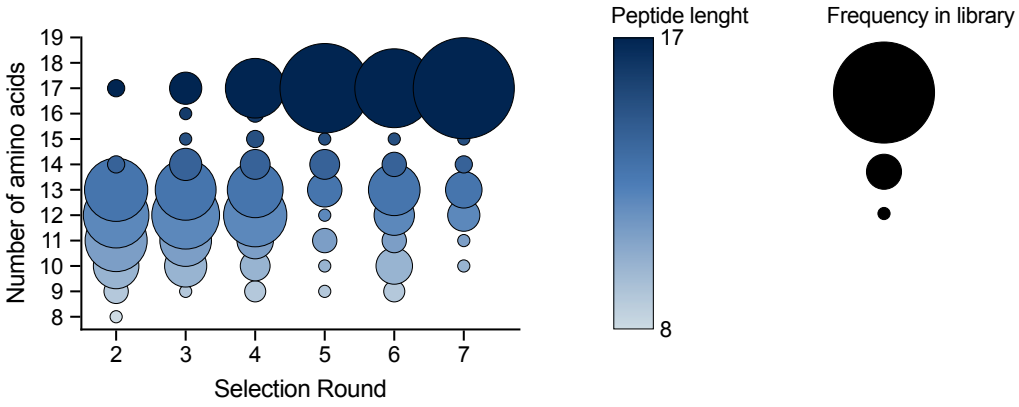

Figure 4.4: The RaPID selection against PSD-95-PDZ2 became dominated by 17-mer sequences. A bubble plot of peptide length frequency in each round in the RaPID PSD-95-PDZ2 selection show that peptide sequences containing 17 AAs became increasingly enriched in each round.

In round two, the library comprised a variety of shorter peptide sequences of significant diversity. However, as the selection progressed, a noticeable shift towards longer sequences, particularly 17-mer species, was observed. Longer peptide seqeunces became increasingly enriched after round four, eventually dominating the selection by round seven. Notably, longer peptide species were already present by round two, suggesting early contamination of the library with RNA templates encoding longer peptides. The reduction in the amount of shorter peptide species after each round suggests a possible binding preference of the PDZ2 domain towards longer sequences. However, it is premature to draw any definite conclusions at this stage. Further research is required to confirm this potential binding preference.

4.5 Potential binding properties of novel PSD-95-PDZ2 inhibitors

Protein structure prediction is commonly applied in structure-based drug discovery to better understand protein binding, affinity and selectivity. For this purpose, protein-complex structure prediction was performed using the publicly available ColabFold tool.[49] The sequences for PSD-95-PDZ2 (UniProt: P78352) and the previously discovered second-generation peptide JFL03 (fIETTFAGNRIETSFTC) were connected using a ":" separator and put into ColabFold's batch processing. The protein-peptide complex was predicted in template mode

"mmseqs2", which enabled ColabFold to search for existing PDB structures, thereby improving prediction accuracy (Figure 4.5.1 A). Similarly, PSD-95-PDZ2 was predicted in complex with the top-hit peptide EWA_PDZ_27001, for which the lariat structure was put into ColabFold's batch processing (Figure 4.5.1 B). Important to mention is that ColabFold cannot predict fully cyclic structures, hence the gap(s) in the predicted peptide structures (Figure 4.5.1 A and B).

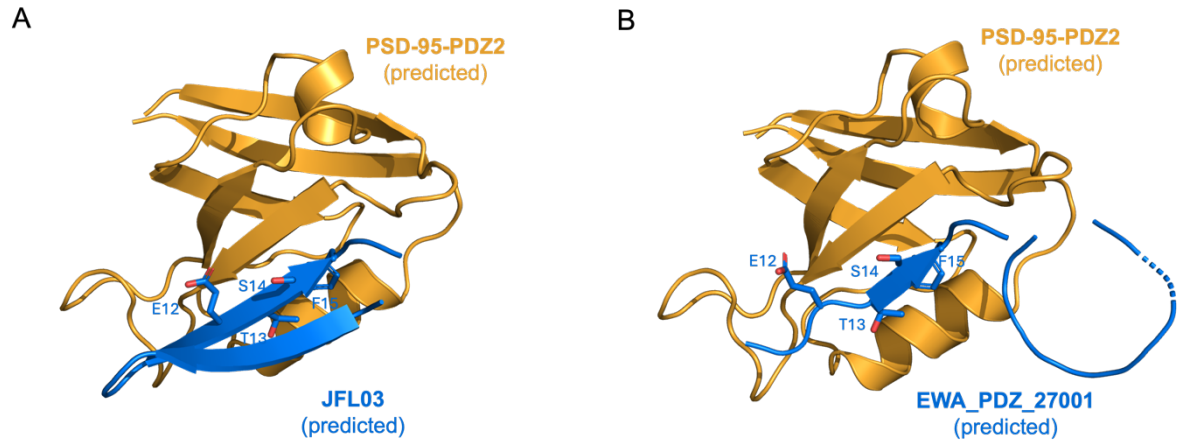

Figure 4.5.1: ColabFold predicted secondary structures of PSD-95-PDZ2 in complex with RaPID-derived peptides. Cartoon representation of predicted secondary structure of PSD-95-PDZ2 (blue) in complex with **A)** second-generation JFL03 peptide (orange), where the ^{12}ETSF^{15} motif is presented as sticks show similar binding properties as the nNOS β-hairpin mimetic peptide. **B)** RaPID-derived EWA_PDZ_27001 peptide (orange), where the ^{12}ETSF^{15} motif is presented as sticks show binding similarities with conventional linear peptides.

To further elucidate the structure-activity relationship of the RaPID-derived peptides with the target protein, the PDBePISA tool was employed. PDBePISA is an interactive web server used to analyze macromolecular assemblies from X-ray crystallography data.[50] This tool was applied to the predicted protein-peptide complexes to identify interactions and pinpoint residues crucial for the pharmacophore of the peptide drug. Interaction analysis of the predicted PDZ2/JFL03 complex showed several key hydrogen-bond interactions, including the side chains of PDZ2 residues S18, T37 and H70 and peptide residues S14, E12 and T13, respectively, comprising bond lengths of 2.75 Å, 2.94 Å and 3.93 Å. The analysis also implied that the PDZ2 domain engages in multiple hydrogen-bond interactions with the peptide's backbone. Interaction analysis of PDZ2/EWA_PDZ_27001 showed similar key residue interactions, as well as peptide backbone interactions. Findings suggests that the "ETS" motif is potentially important for the PDZ2 binding interface and hence the peptide's pharmacophore.

Encouragingly, this “ETS” motif was present in most of the top enriched RaPID-derived peptides (Figure 4.3.3 B).

Comparison with the known complex syntophin-PDZ/nNOS (Figure 4.3.3 C) implies that the “ETS” motif obtain similar binding properties towards PDZ2 as the previously identified “ETT” motif in the nNOS β-hairpin (Figure 4.5 B). Interestingly, since the second-generation compound JFL03 also contains a “ETT” motif in one of its β-sheets, it seems like it has flipped 180° horizontally to bind to PDZ2 with the “ETS” motif in the second β-sheet instead compared to the “ETT” motif in the nNOS β-hairpin. In addition, residue F13 in both predicted complexes (Figure 4.5.1 A and B), seems to locate itself into a hydrophobic pocket in the PDZ2 binding interface, similarly as the nNOS β-hairpin mimetic peptide’s F111 residue (Figure 4.3.3 C). Moreover, the prediction of the lariat peptide structure EWA_PDZ_27001 in complex with PDZ2 (Figure 4.5.1 B) suggests that the lariat tail binds similar as conventional linear PDZ2 binding peptides. The prediction also implies that the cyclic part of the lariat peptide seems to be of less importance for binding to PDZ2. Which raises speculations concerning a potential binding preference of PDZ2 towards linear peptides. However, it is important to emphasize, that these structures are only predictions and that ColabFold uses existing PDB structures as reference, thereby generating models similar to existing structures which are based on true X-ray crystallography data. Hence, the suggested structure-activity relationship of the RaPID-derived PSD-95 binders in the predicted models are very speculative and need to be further assessed and confirmed.

Further on, to assess the accuracy of ColabFold’s predictions, the known crystal structure the syntrophin-PDZ domain in complex with nNOS was aligned with the predicted models using PyMOL. The protein model in PDZ2/JFL03 (Figure 4.5.2 A) exhibited a root mean square deviation (RMSD) value of 0.44 Å when aligned with the syntrophin-PDZ domain. Since success of such alignments is observed when the RMSD is < 3 Å[51], this indicates a high level of accuracy in the prediction. Similarly, the protein model in PDZ2/EWA_PDZ_27001 (Figure 4.5.2 B) exhibited an RMSD value of 0.46 Å when aligned with the PDZ domain. Additionally, to evaluate the structural similarity between the PDZ domain of syntrophin and PDZ2 of PSD-95, sequence alignment was performed using PyMOL with a known crystal structure of the PSD-95-PDZ2 domain (PDB: 3ZRT).[8] The alignment yielded an RMSD value of 0.57 Å, confirming the structural similarity between the two PDZ domains (Figure 4.5.2 C).

The structural similarity supports the accuracy of the predicted PDZ2 domain in the two models (Figure 4.5.2 A and Figure 4.5.2 B).

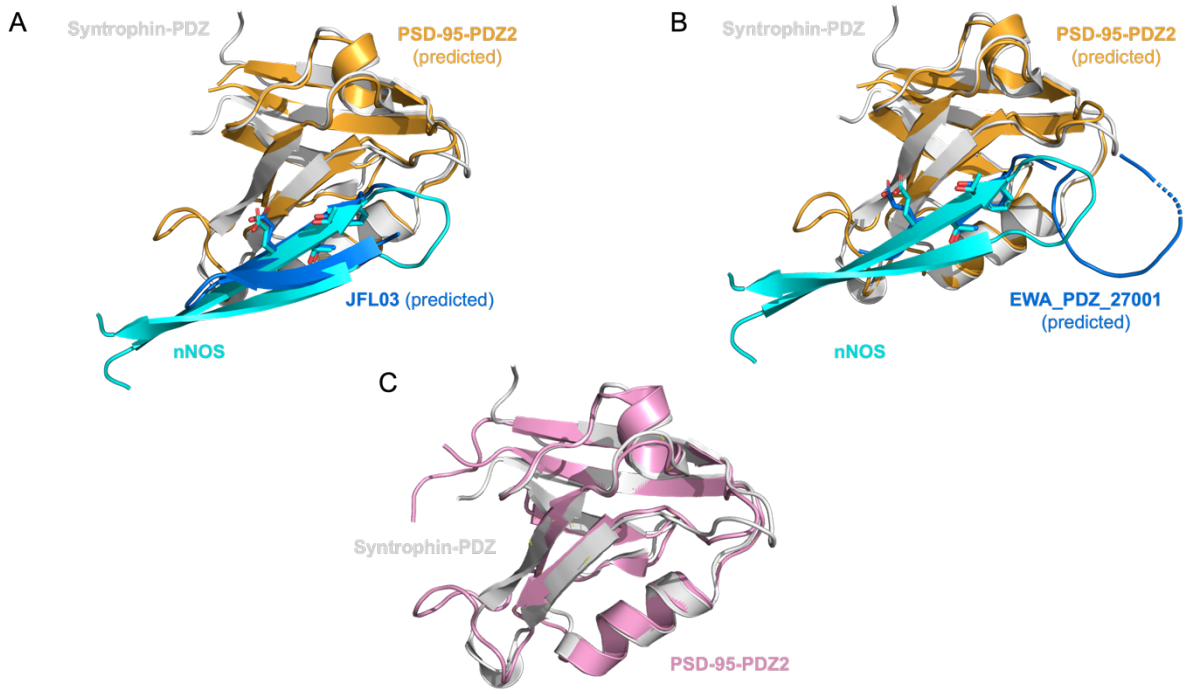

Figure 4.5.2: ColabFold predicted secondary structures of PSD-95-PDZ2 in complex with RaPID-derived peptides aligned with a known PDZ domain structure binding to the nNOS peptide. A) Alignment of predicted secondary structure of PSD-95-PDZ2 (blue) in complex with JFL03 peptide (orange) with syntrophin-PDZ/nNOS (PDB: 1QAV)[48] implies similar pharmacophores. **B)** Alignment of predicted secondary structure of PSD-95-PDZ2 (blue) in complex with EWA_PDZ_27001 peptide (orange) with syntrophin-PDZ/nNOS (PDB: 1QAV)[48] suggests similar residues crucial for binding. **C)** PDS-95-PDZ2 (PDB: 3ZRT)[8] aligned with syntrophin-PDZ/nNOS (PDB: 1QAV).[48]

Moreover, it appears that the cyclic JFL03 scaffold adopts a secondary structure similar to the first-generation nNOS β-hairpin mimetic peptide, which contains a loop conformation (^{112}TGDGT^{116}) connecting the two antiparallel β-sheets. The loop in the nNOS β-hairpin includes a glycine residue, whose side chain is small and commonly found in protein loops.[52] Both RaPID-derived JFL03 and EWA_PDZ_27001 peptides contain a glycine residue between their β-sheets, suggesting that, if adopting a full macrocycle, they would also form a loop connecting the β-sheets. As previously mentioned, the ColaFold prediction of PDZ2/JFL03 implies that the peptide has flipped 180° horizontally, positioning the thioether towards the PDZ2 binding interface rather than its glycine residue and the potential loop conformation. In future work, it would be interesting to further investigate the impact the positioning might have on the peptides' pharmacophore.

5. Conclusion and future perspectives

The development and structural optimization efforts of macrocyclic peptides inhibitors of the nNOS/PSD-95 PPI holds significant potential in mitigating brain damage caused by AIS. This thesis explored the evolution of cyclic peptide inhibitors using the RaPID platform to generate a large, genetically encoded library of cyclic peptides targeting PSD-95-PDZ2. Experiments focused on a randomized peptide library of smaller, and potentially more permeable cyclic peptides than previously employed libraries. The focus was specifically on peptide sequences comprising nine to thirteen AAs, hitting the optimal length of eleven AAs thought to have improved membrane permeability and bioactivity. The bead-binding assay ensured efficiency of the RaPID selection, confirming an 80% biotinylation degree of PSD-95-PDZ2 and a loading capacity of 1 pmol/µL bead slurry. The RaPID selection, spanning seven rounds, successfully enriched specific peptide species, obtaining a noticeable increase by round five, with a peak in round six. Decoding of NGS results revealed an unexpected dominance of 17 AAs containing peptide sequences, suggesting contamination or formation of PCR product artefacts yet providing a potential high-affinity binder. The discovered binder was almost completely identical to the previous RaPID-derived second-generation cyclic scaffold, implying a potential binding preference of PSD-95-PDZ2 for longer sequences. Simultaneously, it shows that the RaPID platform is a strong method for high-throughput screening of peptides toward a protein target.

The structural predictions using ColabFold and interaction analysis via PDBePISA underscored binding potential of the identified peptides. The key motifs “IETSF” and “IETTF” were identified as potentially critical for the pharmacophore, aligning with previous findings related to the nNOS β-hairpin. The predicted models, to some degree validated by RMSD values, suggested a binding preference of PDZ2 of PSD-95 towards linear peptides, although further experiential validation is needed. In conclusion, the top enriched RaPID-derived peptide species EWA_PDZ_27001, resembling the second-generation cyclic peptide, exhibit promising characteristics as a PSD-95-PDZ2 inhibitor. The structure-activity relationship gained from this study provide a foundation for future optimization and validation, potentially leading to effective therapeutic modalities for stroke-related brain damage. Future steps involve detailed structural studies and functional assays to confirm the binding preferences and efficacy of these novel PSD-95 inhibitors.

6. Experimental procedures

6.1 General methods

FPLC. Purification of eFx and tRNA were obtained with an ÄKTA FPLC system using ion exchange chromatography technology. RNA samples were run in an DEAE FF HiTrap IEX 1 mL column (Cytvia™) with a gradient of increasing buffer B from 0 to 100% with 2%/mL in buffer A. RNA were eluted into 1 mL fractions in 1.5 mL Eppendorf tubes in the fraction tray.

NanoDrop. RNA and DNA concentrations were determined with a NanoDrop 2000 spectrophotometer (Thermo Scientific™). 1.6 µL sample was pipetted into the NanoDrop and A260 in ng/µL was measured. Concentrations were calculated with the formula: [RNA conc. in µM] = A260 in ng/µL / (0.34*length of RNA in nt).

SDS-PAGE. Protein samples were mixed with 3.3 µL SDS-PAGE loading buffer consisting of 4x Laemmli sample buffer (BioRad™) and *b*-mercaptoethanol in a 9:1 ratio. Samples were heated at 95°c for 5 min. 10 µL of protein samples and 3 µL Precision Plus Protein Dual Color Standard (BioRad™) was loaded in a 4-20%, 15-well comb, 15 µL Mini-PROTEAN TGX precast gel (BioRad™) The gel was run in x1 Tris-glycine SDS (TGS) buffer at 180 V for 34 min. The gel was stained with Quick Coomassie Stain for 1 h and destined overnight in MilliQ water.

Agarose gel electrophoresis. 4% agarose gels were made through mixing 4 g Agarose 3:1 HRB™ with 100 mL 1x TAE buffer. The mixture was boiled in a microwave until the agarose was dissolved. 5 µL SDS was added, and gels with wells were cast. 3 µL of the DNA samples were mixed with equal amount x2 orange DNA loading buffer before loading into the gel next to a 60, 100, 130 bp DNA ladder premixed with x2 orange dye. Gels were run in x1 TAE buffer at 150 V for 15 min.

6.2 Chemical biology

6.2.1 Bead binding assay

32 µL slurry of DynaBeads™ M-280 Streptavidin (Invitrogen™) were pelleted using a magnetic rack and storage solution removed. The beads were washed with 30 µL wash buffer and then resuspended in 32 µL of the same buffer. Aliquots of beads with 16, 8, 4, 2, 1, and 0.5 µL slurry were transferred into new 8-strip PCR tubes. The beads were pelleted, supernatant discarded and 20 µL of 2 µM protein solution was added to each aliquot to allow the protein to bind to the beads. Samples were incubated for 15 min at 4°C with rotation. Biotin was added to a final concentration of 12.5 mM and samples were incubated while rotating for another 15 min at 4°C to allow biotin to bind to any free sites. After incubation samples were spun down and supernatant collected into new tubes, which represented the “unbound” samples for

later SDS-PAGE analysis. The beads were washed with wash buffer x3 and resuspended in 10 µL buffer, which represented the "bound" samples for the SDS-PAGE.

6.2.2 Flexizyme (eFx) synthesis

6.2.2.1 Klenow primer extension

Initially, primer extension was carried out using the Klenow enzyme to synthetize a long dsDNA from two ssDNA oligos with a >18 bp overlap. Buffer2 (included with the Klenow enzyme purchased from NEB) was mixed with a forward primer (Fx_PE_v2_F) and a reverse primer (eFx_PE_v2_R) in RNAse free water to a total volume of 400 µL. The mixture was then divided into 100 µL aliquots in 8-strip PCR tubes. Primers were separated and overlapped using the T100™ Thermal cycler (BioRad). The lid was heated to 105°C and cycles were run at 94°C for 2 min, then ramped down to 50°C at 0.2°C/s, followed by constant 50°C (5 min) and lasty ramped down to 37°C at 0.2°C/s. 2.5 µL 10 mM dNTPS and 1 µL Klenow enzyme was next added into each of the four PCR tubes and put back into the Thermal cycler set at 37°C for 30 min to allow primer extension.

6.2.2.2 Phenol chloroform extraction and ethanol precipitation

The Klenow PCR product was transferred to two 0.6 mL tubes and purified via phenol-chloroform extraction in the fume hood. First 1/10 volume of 3 M NaCl was added, then 200 µL PCI (phenol-chloroform-isoamyl alcohol at 25:24:1). Tubes were centrifuged at 13,000 rpm for 5 min at 25°C. The top aqueous phase was then transferred to new tubes to which 200 µL CI (chloroform-isoamyl alcohol at 24:1) was added and centrifuged again at 13,000 rpm, 5 min at 25°C. The top aqueous phase was transferred to new tubes to which 2x the volume 200 proof EtOH was added and mixed thoroughly. Tubes were centrifuged at 13,000 rmp, 15 min at 25°C and then the supernatant was discarded. The pellets were washed by adding 70% EtOH (100 µL), followed by centrifugation at 13,000 rpm, 3 min at 25°C and the supernatant discarded. Pellets were dried using KimWipe placed over the tubes, and then dissolved in 50 µL RNase-free H_2O.

6.2.2.3 Transcription DNase and EDTA treatment

The purified Klenow PCR DNA product was transcribed into RNA. Transcription was carried out in a 5 mL centrifuge tube in a mixture containing 10x T7 buffer, 1 M DTT, 1 $MgCl_2$, 25 mM each dNTPs, DNA template (the Klenow product), RNasin RNase inhibitor (Promega™) T7 RNA polymerase (homebrew) and Yeast iPPase (homebrew) and incubated for 3 h at 37°C. After transcription, 275 µL x10 DNase buffer and 36 µL DNase was added and samples were incubated for another 60 min at 37°C. Finally, 0.5 M EDTA pH 8.0 (438 µL) was added, and eFx were purified with FPLC.

6.2.2.4 Acidic PCI and ethanol precipitation

After purification with FPLC where the most concentrated fractions were pooled into 5 mL tubes, equal volume of acidic PCI was added in a fume hood. Samples were centrifuged at 13,000 rmp for 10 min at 25°C. The top aqueous phase was collected and mixed with equal volume CI and centrifuged fat 13,000 rpm, 10 min at 25°C. Aqueous phase was collected and

x2 volume of 200 proof EtOH was added and spun at 13,000 rpm, 15 min at 25°C. The supernatant was discarded, and the pellet washed with 70% EtOH through centrifugation for 3 min at 25°C. Pellet was dried before resuspension with RNase-free H_2O. After concentration determination with NanoDrop, 0.3 µL RNasin RNase inhibitor was added and eFx (final conc. of 250 µM) were stored in -80°C.

6.2.3 $^{ini}tRNA_{CAU}$ synthesis

Synthesis of tRNA molecules were performed with the same method as for eFx, but with primers ini_tRNA_PE_v2_F57 (forward), ini_tRNA_PE_v2_R38 (reverse) and initial sample volume of 100 µL. Additionally, after Klenow primer extension and before phenol chloroform extraction and ethanol precipitation, PCR in the T100™ Thermal Cycler using the enzyme Pfu-sso7d (Phusion homebrew) was carried out. The PCR Phusion mixture contained RNase-free ddH_2O, x5 Phusion HF, 10 mM dNTPs, 100 µM forward primer (ini_tRNA_PCR_v2_F40), reverse primer (ini_3_R20_OMe), DNA template (i.e the Klenow product) and the enzyme Pfu-sso7d. Samples were split into eight 8-strip PCR tubes and 15 PCR cycles were run with the program in Table 6.2.

Table 6.2.1: Temperature and time specifications for the program run in the PCR T100™ Thermal Cycler for $^{ini}tRNA_{CAU}$ synthesis.

°C	Time (s)
98	60 (*intial denaturation)*
98	10 (*cycle denaturation)*
55	30 (*cycle annealing*)
72	40 (*cycle extension*)
72	60 (*final extension*)
20	∞ (*hold temp.*)

6.2.4 Generation of the NKK7-11 mRNA library

In this thesis the library was designed with NNK sequences comprising 7-11 randomized amino acids, giving a variability of $\# = (4 * 4 * 2)^{NNK\ size}$ for each NNK size. Since the smaller NNK sizes have a higher number of molecules when in the same concentration as the larger ones, NNK7 was added with a ratio of 0.1, while NNK8-11 was added with a 1:1 ratio. The appropriate mRNA NNK template stocks (NNK*n* CGS, 20 µM) were thawed and 6.1 µL of each NNK library (NNK7 was prior diluted x10) were mixed in a 1.5 mL Eppendorf tube. RNase-free H_2O was added to receive a final volume of 100 µL and 5 µM concentration.

6.2.5 qPCR DNA standards

A reverse transcription reaction (RvTr) was prepared in an 8-strip PCR tube on ice. Consisting of 8 µL 5xSSIV buffer (Thermo Fisher™), 0.5 mM dNTPs, 2 µM reverse primer CGS3an_RvTr_R26, 0.2 µM NNK9 mRNA library, 0.2 µL RNase inhibitor and 1 µL MutD9 reverse transcriptase (homebrew, 5 µM) in RNase-free H_2O to a total volume of 40 µL. The

RvTr was incubated in a PCR machine for 1h at 45°C. Under the assumption that there were 100% reverse transcription, a 60.2x dilution of the RvTr reaction with 0.1% Triton X-100 was made to yield a 2e9 DNA molecules/µL solution. A serial dilution was then made taking 10 µL from the previous sample and diluting with 90 µL 0.1% Triton X-100 yielding solutions containing 2e8, 2e7, 2e6, 2e5 and 2e4 molecules/µL stored in -20°C for later use in qPCR analysis.

6.2.6 Flexizyme aminoacylation of tRNA

Synthetized flexizyme (eFx) was used for aminoacylation and charging of the initiator amino acid *N*-ClAc-L-Phe onto the synthetized $^{ini}tRNA_{CAU}$ molecules. HEPES-KOH pH 7.6 (83 mM), eFx (41.7 µM), $^{ini}tRNA_{CAU}$ (41.7 µM) and RNase-free H_2O was mixed in a 0.6 mL tube, heated at 95°C for 2 min and left in RT for 5 min to allow the RNA to fold. 3 M $MgCl_2$ was added with a final conc. of 0.75 M, and mixture left for another 5 min in RT. After cooling of the reaction mixture on ice, the amino acid substrate *N*-ClAc-L-Phe was added (5 mM) and the reaction incubated on ice for 2 h. After end of reaction acidic buffer was added (0.24 M AcOH pH 5.2) and mixture split into the desirable number of Eppendorf tubes in equal amounts (one tube for each RaPID round). Aminoacetylated tRNA (aa-tRNA) was precipitated by addition of 2x amount of EtOH, followed by centrifugation at 13 000 rpm, 25°C for 15 min. aa-tRNA pellets were washed with 70% EtOH, 0.1 M AcOH and centrifuged for 10 min at 13 000 rpm, 25°C two times, and finally washed with 70% EtOH and centrifuged for 3 min at 13 00 rpm, 25°C. The supernatant was discarded, pellets air dried and stored in -80°C for later use in the RaPID selection. Stability of aa-tRNA varies on the amino acid substrate but should not be stored for more than four weeks.

Table 6.2.2: Primer sequences used for preparation of qPCR standards and synthesis of flexizyme and $^{ini}tRNA_{CAU}$.

Primer	Sequence
Fx_PE_v2_F	GAATTTAATACGACTCACTATAGGATCGAAAGATTTCCGC
eFx_PE_v2_R	ACCTAACGCTAATCCCCTTTCGGGGCCGCGGAAATCTTTCGATCC
ini_tRNA_PE_v2_F57	GGCGGGGTGGAGCAGCCTGGTAGCTCGTCGGGCTCATAACCCGAAGATCGTCGGTTC
ini_tRNA_PE_v2_R38	TGGTTGCGGGGGCCGGATTTGAACCGACGATCTTCGGG
ini_tRNA_PCR_v2_F40	GAATTTAATACGACTCACTATAGGCGGGGTGGAGCAGCCT
ini_3_R20_OMe	TGGTTGCGGGGGCCGGATTT
CGS3an_RvTr_R26	TCCTAGCTGCCGCTGCCGCTGCCGCA

6.3 RaPID

6.3.1 Puromycin ligation

Ligation of puromycin to the mRNA library was performed in a mixture composed of Pur-linker (1.5 µM), mRNA library (1 µM), ATP (1 mM), 10x T4 RNA ligase buffer (NEB) and T4 RNA ligase (Thermo Fisher™) at RT for 30 min. Following addition of EDTA pH 8.0 (final conc. of 9.8 µM), the mRNA-pur complex was purified using Mag-Bind® TotalPure NGS beads (Omega Bio-tek, Inc). 24.6 µL bead slurry was mixed with 21 µL iPrOH (isopropanol) and the supernatant removed. The pellet was washed x2 with 90% EtOH and then air dried. Beads were resuspended in 4 µL RNase-free ddH_2O and after 5 min the supernatant containing the mRNA-pur complex library with assumed concentration of 5 µM was collected.

6.3.2 PURE translation

PURE translation solution was prepared by mixing SolB (22.5%) and SolA (15.5%) solutions with a 1 mM mixture composed of all canonical amino acids except for Met (19 aa). The rounds of RaPID selection were performed using 25 pmol of mRNA-puromycin ligated library and 250 pmol of ClAc-L-Phe-$^{ini}tRNA_{CAU}$. The first RaPID round was performed in a 150 µL translation mixture, while the subsequent rounds were performed in a 5 µL translation mixture. Followed by mixing, the translation solution was incubated at 37°C for 40 min. After the first incubation 30 µL (R1) or 1 µL (R2+) of 100 mM EDTA pH 8.0 was added and the mixture incubated in 37°C for another 30 min to dissociate the ribosome from the mRNA-puromycin-peptide products and promote thioether-cyclization of the peptides.

6.3.3 Reverse transcription

The reverse transcription of the translated mRNA-peptide fusion was performed in a mixture composed of the translation product, RvTr-premix, 17 mM $Mg(OAc)_2$ and 0.17 µM MutD9 reverse transcriptase (homebrew). The RvTr mixture was incubated in the PCR machine at 45°C for 30 min to generate the mRNA-cDNA duplex fused with puromycin and its encoding peptide. After reverse transcription 265 µL (R1) or 8.8 µL (R2+) of x2 blocking solution was added and sample put on ice. From the blocked translation solution, 0.5 µL was diluted into 200 µL 0.1% Triton for later qPCR as the INPUT sample.

6.3.4 Binding selection

6.3.4.1 Negative selection

Negative selection rounds with the purpose of eliminating peptides binding to the streptavidin magnetic beads were performed for RaPID selection R2+. Five negative selection samples were prepared by transferring 54 µL bead slurry (3x positive selection amount) to an 8-strip PCR tube. Storage solution was removed and 100 µL x1 selection buffer added before splitting it into 2x50 µL. BTN was subsequently added to one of the samples with a final concentration of 25 µM and incubated at 4°C for 15 min while rotating. Both samples were washed again with 30 µL selection buffer. 50 µL selection buffer was added and beads combined. 20 µL bead slurry was added into five new 8-strip PCR tubes on ice. The supernatant was removed from the first negative bead sample and the blocked translation solution containing the translated

peptides added and incubated at 4°C rotating for 30 min. The sample was spun down, pellet with magnet and the supernatant with translated peptides transferred to the second negative beads sample for incubation. This procedure was repeated for all five negative bead samples. After the last incubation, the recovered peptide library was transferred to the positive beads for the subsequent positive selection.

6.3.4.2 Positive selection

Protein target was immobilized on Dynabeads™ M-280 Streptavidin magnetic beads (Invitrogen™) for the positive selection. 107 µL (R1) or 3.6 µL bead slurry (R2+) was pelleted, washed with selection buffer (x2) and transferred to a new tube. Protein solution was subsequently added to beads (107 pmol in R1 and 3.6 pmol in R2+) and rotated for 15 min at 4°C. 25 µM final concentration of BTN was next added followed by rotation for 15 min at 4°C. Beads were washed with selection buffer (x2), transferred to new tubes before adding the recovered peptide library solution from the fifth negative beads sample. The positive sample was incubated for 30 min at 4°C.

6.3.5 Heat elution of hit peptides

The negative beads were recombined and washed (x3) with selection buffer. Negative beads were resuspended in 20 µL 0.1% Triton and 1 µL 500 µM BTN before heat treatment at 95°C for 5 min to elute peptides bound to the beads. The supernatant was collected representing the NEG sample for later qPCR analysis. The positive beads were after incubation washed with selection buffer (x3) and transferred to a new tube after each wash. 50 µL of elution PCR mix was added before heat treatment at 95°C for 5 min. The supernatant containing the recovered peptide library was collected representing the POS sample for later qPCR.

6.3.6 qPCR analysis

Quantitative PCR (qPCR) was used to quantify the amount of DNA molecules in the input and output samples from the RaPID selection. A qPCR master mix (MM) was prepared on ice. 19 µL of the qPCR MM was mixed with 1 µL of the sample to be analyzed in qPCR tubes on ice. The qPCR DNA standards, the INPUT, NEG and POS samples from the RaPID selection together with a no-DNA control (0.1% Triton) were run in a MyGO Pro qPCR thermocycler for 35 cycles with the program in Table 6.3.1.

Table 6.3.1: Temperature and time specifications for the program run on the MyGO Pro qPCR thermocycler.

°C	Time (s)
95	30 (*intial denaturation)*
95	20 (*cycle denaturation)*
61	20 (*cycle anneal*)
72	30 (*cycle extension*)

6.3.7 Recovery PCR

The cDNA libraries were mixed with 0.5 µL Pfu-sso7d enzyme, recovered and amplified with PCR run in a T100™ Thermal Cycler with the program in Table 6.3.2. The number of PCR cycles were based on qPCR results. cDNA samples were subsequently analyzed in an agarose gel (section 6.1), confirming enough PCR cycles were run. The cDNA libraries were then isolated using x1.23 volume of Mag-Bind® TotalPure NGS beads (Omega Bio-tek, Inc) together with x1.05 volume of iPrOH. Washing of pelleted beads using 85% EtOH was repeated x2 followed by air drying on ice (5 min). After resuspension in RNase-free ddH_2O the beads were pelleted and supernatant containing the isolated DNA libraries collected.

Table 6.3.2: Temperature and time specifications for the recovery PCR program run with the T100™ Thermal Cycler.

°C	Time (s)
98	30 (*intial denaturation)*
98	10 (*cycle denaturation)*
61	30 (*cycle annealing*)
72	40 (*cycle extension*)
72	60 (*final extension*)
20	∞ (*hold temp.*)

6.3.8 Transcription

The isolated DNA libraries were transcribed in 37°C for 3 h in a mixture of 10x T7 buffer, 25 mM NTPs, 1 M $MgCl_2$, 100 mM DTT, T7 RNA polymerase (Thermo Fisher™) and RNase-free water. After incubation 0.1x volume of 500 mM EDTA was added, and mRNA purified using Mag-Bind® TotalPure NGS beads (Omega Bio-tek, Inc) together with iPrOH. Washing of pelleted beads using 85% EtOH was repeated x2 followed by air drying on ice (5 min). After resuspension in RNase-free ddH_2O the beads were pelleted and supernatant containing the mRNA library collected. The mRNA concentration was measured with Nanodrop and diluted to a final mRNA library concentration of 5 µM to be used in the subsequent RaPID selection round.

Table 6.3.3: Primer seqeunces used in the RaPID selection and preparation of the RaPID libraries.

RaPID primers	Sequence
lib_NNK7_R83	TTTCCGCCCCCCGTCCTAGCTGCCGCTGCCGCTGCCGCAMNNMNNMNNMNNMNNMNNMNNCATATGTATATCTCCTTCTTAAA
lib_NNK8_R86	TTTCCGCCCCCCGTCCTAGCTGCCGCTGCCGCTGCCGCAMNNMNNMNNMNNMNNMNNMNNMNNCATATGTATATCTCCTTCTTAAA

lib_NNK9_R89	TTTCCGCCCCCCGTCCTAGCTGCCGCTGCCGCTGCCGC AMNNMNNMNNMNNMNNMNNMNNMNNMNNCATAT GTATATCTCCTTCTTAAA
lib_NNK10_R92	TTTCCGCCCCCCGTCCTAGCTGCCGCTGCCGCTGCCGC AMNNMNNMNNMNNMNNMNNMNNMNNMNNMNNCA TATGTATATCTCCTTCTTAAA
lib_NNK11_R95	TTTCCGCCCCCCGTCCTAGCTGCCGCTGCCGCTGCCGC AMNNMNNMNNMNNMNNMNNMNNMNNMNNMNNM NNCATATGTATATCTCCTTCTTAAA
CGS3an_RvTr_R23	TAGCTGCCGCTGCCGCTGCCGCA
T7g10M_CEL_F51	GAATTTAATACGACTCACTATAGGGTTAACTTTAAGA AGGAGATATACATA
GS3an_R36	TTTCCGCCCCCCGTCCTAGCTGCCGCTGCCGCTGCC

Table 6.3.4: Buffers and reagent mixtures compositions used in the RaPID selection.

Buffer/reagent mixture	Composition
x2 Blocking buffer	100 mM HEPES-KOH pH 7.6, 200 mM NaCl, 0.1% Tween20, 0.2% acetylated BSA (20 mg/mL)
x1 Selection buffer	50 mM HEPES-KOH pH 7.6, 100 mM NaCl, 0.05% Tween20
RvTr-premix	3 mM dNTPs, 10 µM CGS3an_RvTr_R23, 150 mM Tris pH 8.3
Elution PCR mix	20% 5x Phusion HF buffer, 200 µM dNTPs, 0.6 µM T7g10M_CEL_F51 primer, 0.6 µM GS3an_R36 primer, 25 µM BTN
qPCR master mix	1% 10x Key Buffer, 200 µM dNTPs, 0.5 µM T7g10M_CEL_F51 primer, 0.5 µM GS3an_R36 primer, 0.5 µM SYBR Green, 0.4% Taq DNA Polymerase (VWR™)

6.4 NGS sample preparation

cDNA library samples from rounds two to seven from the positive selection and from rounds three to seven from the negative selection were prepared for NGS. A unique pair of forward and reverse primers were chosen for each cDNA library sample (Table 6.4). Each library sample was mixed with a PCR mix containing 10 µL 5x Phusion HF buffer, 200 µM dNTPs, 0.6 µM forward oligo, 0.6 µM reverse oligo, 0.5 µL template library cDNA and 0.5 µL Pfu-sso7d enzyme. Followed by amplification with PCR run with the Phusion PCR program in Table 6.2 to receive a final cDNA concentration of 0.2 µM. The number of PCR cycles run were calculated from qPCR analysis results. To confirm enough cycles were run, each cDNA sample was subsequently analyzed with agarose gel electrophoresis (section 6.1). If a band was

visible in the gel, enough cycles were run, and if not, 3 more cycles were run, and sample analyzed again in the agarose gel. All cDNA samples from the positive selection were pooled in an Eppendorf tube, and negative samples pooled in another. The POS and NEG mixtures were purified with x1.8 volume of Mag-Bind® TotalPure NGS beads (Omega Bio-tek, Inc), and washed x2 with 80% EtOH. Mixtures were diluted with RNase-free H_2O to a final cDNA concentration of 20 ng/µL measured with NanoDrop, labeled and shipped to Azenta Life Sciences for NGS.

Table 6.4: The unique barcodes used for identification of each cDNA library sample after NGS.

Primer	Barcode
an13Rd2_GS3_r1	GTTGGAA
an13Rd2_GS3_r2	AAATCGC
an13Rd2_GS3_r3	AGCGTGT
an13Rd2_GS3_r4	ATGGGCT
an13Rd2_GS3_r5	CTCAGTG
Rd1T7g10M_f6	AGCTCAT
Rd1T7g10M_f7	AATGCGA
Rd1T7g10M_f8	CCATTTG

7. References

(1) Caire, M. J.; Reddy, V.; Varacallo, M. *Physiology, Synapse*; StatPearls Publishing, 2018.

(2) Wei, F.; Mingjie, Z. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nature Reviews Neuroscience* **2009**, *10*, 87-99. DOI: 10.1038/nrn2540.

(3) Menglong, Z.; Xudong, C.; Dongshi, G.; Haowei, W.; Penger, T.; Mingjie, Z. Reconstituted Postsynaptic Density as a Molecular Platform for Understanding Synapse Formation and Plasticity. *Cell* **2018**, *174* (5), 1172-1187. DOI: 10.1016/j.cell.2018.06.047.

(4) Dosemeci1, A.; Weinberg, R. J.; Reese, T. S.; Tao-Cheng, J.-H. The Postsynaptic Density: There Is More than Meets the Eye. *Frontiers in Synaptic Neuroscience* **2016**, *8* (23). DOI: 10.3389/fnsyn.2016.00023.

(5) Eunjoon, K.; Morgan, S. PDZ domain proteins of synapses. *Nature Reviews Neuroscience* **2004**, *5*, 771-781. DOI: 10.1038/nrn1517.

(6) Christensen, N. R.; Čalyševa, J.; Fernandes, E. F. A.; Lüchow, S.; Clemmensen, L. S.; Haugaard-Kedström, L. M.; Strømgaard, K. PDZ Domains as Drug Targets. *Advanced Therapeutics* **2019**, *2* (7). DOI: 10.1002/adtp.201800143.

(7) Ugalde-Triviño, L.; Díaz-Guerra, M. PSD-95: An Effective Target for Stroke Therapy Using Neuroprotective Peptides. *Int. J. Mol. Sci.* **2021**, *22* (22), 12585. DOI: 10.3390/ijms222212585.

(8) Bach, A.; Clausen, B. H.; Møller, M.; Strømgaard, K. A high-affinity, dimeric inhibitor of PSD-95 bivalently interacts with PDZ1-2 and protects against ischemic brain damage. *PNAS* **2012**, *109* (9), 3317-3322. DOI: 10.1073/pnas.1113761109.

(9) Balboa, J. R.; Essig, D. J.; Ma, S.; Karer, N.; Clemmensen, L. S.; Pedersen, S. W.; Joerger, A. C.; Knapp, S.; Østergaard, S.; Strømgaard, K. Development of a Potent Cyclic Peptide Inhibitor of the nNOS/PSD-95 Interaction. *Journal of Medicinal Chemistry* **2023**, *66* (1), 976-990. DOI: 10.1021/acs.jmedchem.2c01803.

(10) Collaborators, G. S. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Neurology* **2021**, *20* (10), 795-820. DOI: 10.1016/S1474-4422(21)00252-0.

(11) Aarts, M.; Liu, Y.; Liu, L.; Besshoh, S.; Arundine, M.; Gurd, J. W.; Wang, Y.; W.Salter, M.; Tymianski, M. Treatment of Ischemic Brain Damage by Perturbing NMDA Receptor-PSD-95 Protein Interactions. *Science* **2002**, *298* (5594), 846-850. DOI: 10.1126/science.1072873.

(12) Hill, M. D.; Goyal, M.; Menon, B. K.; Nogueira, R. G.; McTaggart, R. A.; Demchuk, A. M.; Poppe, A. Y.; Brian H Buck; Field, T. S.; Dowlatshahi, D.; et al. Efficacy and safety of

nerinetide for the treatment of acute ischaemic stroke (ESCAPE-NA1): a multicentre, double-blind, randomised controlled trial. *The Lancet* **2020**, *395* (10227), 878-887. DOI: 10.1016/S0140-6736(20)30258-0.

(13) Mayor-Nunez, D.; Ji, Z.; Sun, X.; Teves, L.; Garman, J. D.; Tymianski, M. Plasmin-resistant PSD-95 inhibitors resolve
effect-modifying drug-drug interactions between
alteplase and nerinetide in acute stroke. *Science Translational Medicine* **2021**, *13* (588). DOI: 10.1126/scitranslmed.abb1498.

(14) Successfully completes phase 1 clinical trial. Avilex Pharma: https://www.avilexpharma.com/2021/05/26/completes-phase-1-clinical-trial/, 2021.

(15) Lim, I. A.; Hall, D. D.; Hell, J. W. Selectivity and Promiscuity of the First and Second PDZ Domains of PSD-95 and Synapse-associated Protein 102*. *The Journal of Biological Chemistry* **2002**, *227* (June 14), 21697–21711. DOI: 10.1074/jbc.M112339200.

(16) Buyanova, M.; Pei, D. Targeting intracellular protein–protein interactions with macrocyclic peptides. *Trends in Pharmacological Sciences* **2022**, *43* (3), 234-248. DOI: 10.1016/j.tips.2021.11.008

(17) Joo, S. H. Cyclic Peptides as Therapeutic Agents and Biochemical Tools. *Biomolecules & Therapeutics, The Korean Society of Applied Pharmacology* **2012**, *20* (1), 19-26. DOI: 10.4062/biomolther.2012.20.1.019.

(18) Blanco, M.-J.; Gardinier, K. M.; Namchuk, M. N. Advancing New Chemical Modalities into Clinical Studies. *ACS Medical Chemistry Letters* **2022**, *13* (11), 1691-1698. DOI: 10.1021/acsmedchemlett.2c00375.

(19) Vinogradov, A. A.; Yin, Y.; Suga*, H. Macrocyclic Peptides as Drug Candidates: Recent Progress andRemaining Challenges. *Journal of the American Chemical Society* **2019**, *141* (10), 4167-4181. DOI: 10.1021/jacs.8b13178.

(20) Glas, A.; Bier, D.; Hahne, G.; Rademacher, C.; Ottmann, C.; Grossmann, T. N. Constrained Peptides with Target-Adapted Cross-Links as Inhibitors of a Pathogenic Protein–Protein Interaction. *Angewandte Chemie International Edition* **2014**, *53* (9), 2253-2507. DOI: 10.1002/anie.201310082.

(21) Tanada, M.; Tamiya, M.; Matsuo, A.; Chiyoda, A.; Takano, K.; Ito, T.; Irie, M.; Kotake, T.; Takeyama, R.; Kawada, H.; et al. Development of Orally Bioavailable Peptides Targeting an Intracellular Protein: From a Hit to a Clinical KRAS Inhibitor. *Journal of the American Chemical Society* **2023**, *145* (30), 16287-16948. DOI: 10.1021/jacs.3c03886.

(22) Goto, Y.; Suga, H. The RaPID Platform for the Discovery of Pseudo-Natural Macrocyclic Peptides. *Accounts of Chemical Research* **2021**, *54* (18), 3505-3618. DOI: 10.1021/acs.accounts.1c00391.

(23) *Company Info / Access*. PeptiDream Inc, https://www.peptidream.com/en/company/profile/ (accessed 2024-04-25).

(24) Xiao, H.; Murakami, H.; Suga, H.; Ferré-D'Amaré, A. R. Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. *nature* **2008**, *454*, 358-361. DOI: 10.1038/nature07033.

(25) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *nature methods* **2006**, *3* (5), 357–359. DOI: 10.1038/nmeth877.

(26) Kuschert, S.; Stroet, M.; Chin, Y. K.-Y.; Conibear, A. C.; Jia, X.; Lee, T.; Bartling, C. R. O.; Strømgaard, K.; Güntert, P.; Rosengren, K. J.; et al. Facilitating the structural characterisation of non-canonical amino acids in biomolecular NMR. *Magnetic Resonance* **2023**, *4* (1), 57-72. DOI: 10.5194/mr-4-57-2023.

(27) Niwa, N.; Yamagishi, Y.; Murakami, H.; Suga, H. A flexizyme that selectively charges amino acids activated by a water-friendly leaving group. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19* (14), 3892-3894. DOI: 10.1016/j.bmcl.2009.03.114.

(28) Rosenberg, H. F. RNase A ribonucleases and host defense: an evolving story. *Journal of Leukocyte Biology* **2008**, *83*, 1079–1087. DOI: 10.1189/jlb.1107725.

(29) Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. Reprogramming the Translation Initiation for the Synthesis of Physiologically Stable Cyclic Peptides. *ACS Chemical Biology* **2008**, *3* (2), 120–129. DOI: 10.1021/cb700233t.

(30) Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. Cell-free translation reconstituted with purified components. *Nature Biotechnology* **2001**, *19*, 751–755. DOI: 10.1038/90802.

(31) Huang, Y.; Wiedmann, M. M.; Suga*, H. RNA Display Methods for the Discovery of Bioactive Macrocycles. *Chemical Reviews* **2019**, *119* (17), 9723-10392. DOI: 10.1021/acs.chemrev.8b00430.

(32) Horiya, S.; Bailey, J. K.; Krauss, I. J. Chapter Four - Directed Evolution of Glycopeptides Using mRNA Display. *Methods in Enzymology,* **2017**, *597*, 83-141. DOI: 10.1016/bs.mie.2017.06.029.

(33) *Dynabeads™ M-280 Streptavidin*. https://www.thermofisher.com/order/catalog/product/11205D (accessed 2024-04-29).

(34) Fairhead, M.; Howarth, M. *Site-Specific Biotinylation of Purified Proteins Using BirA*; Humana Press, 2014. DOI: 10.1007/978-1-4939-2272-7_12.

(35) Mandlik, J. S.; Patil, A. S.; Singh, S. Next-Generation Sequencing (NGS): Platforms and Applications. *Journal of Pharmacy and Bioallied Sciences* **2024**, *16* (1), 41-45. DOI: 10.4103/jpbs.jpbs_838_23.

(36) Qin, D. Next-generation sequencing and its clinical application. *Cancer Biology & Medicine* **2019**, *16* (1), 4-10. DOI: 10.20892/j.issn.2095-3941.2018.0055.

(37) Lu, Y.; Shen, Y.; Warren, W.; Walter, R. *Next Generation Sequencing in Aquatic Models*; Kulski, J. K., Ed.; 2016. DOI: 10.5772/61657.

(38) Nowakowski, A. B.; Wobig, W. J.; Petering, D. H. Native SDS-PAGE: high resolution electrophoretic separation of proteins with retention of native properties including bound metal ions *Metallomics* **2014**, *6* (5), 1068–1078. DOI: 10.1039/c4mt00033a.

(39) Lee, P. Y.; Costumbrado, J.; Hsu, C.-Y.; Kim, Y. H. Agarose Gel Electrophoresis for the Separation of DNA Fragments. *Journal of Visualized Experiments* **2012**, *20* (62), 3923. DOI: 10.3791/3923.

(40) Moreno-Arribas, M. V.; Polo, M. C. CHROMATOGRAPHY | High-performance Liquid Chromatography. In *Ecyclopedia of Food Sciences and Nutrition (second Edition*, Caballero, B. Ed.; Academic Press, 2003; pp 1274-1280. DOI: 10.1016/B0-12-227055-X/00232-7.

(41) Ohta, A.; Tanada, M.; Shinohara, S.; Morita, Y.; Nakano, K.; Yamagishi, Y.; Takano, R.; Kariyuki, S.; Iida, T.; Matsuo, A.; et al. Validation of a New Methodology to Create Oral Drugs beyond the Rule of 5 for Intracellular Tough Targets. *Journal of the American Chemical Society* **2023**, *145* (44), 23891-24436. DOI: 10.1021/jacs.3c07145.

(42) Iskandar, S. E.; Bowers, A. A. mRNA Display Reaches for the Clinic with New PCSK9 Inhibitor. *ACS Medicinal Chemistry Letters* **2022**, *13* (9), 1378-1523. DOI: 10.1021/acsmedchemlett.2c00319.

(43) WEBER, P. C.; OHLENDORF, D. H.; WENDOLOSKI, J. J.; SALEMME, F. R. Structural Origins of High-Affinity Biotin Binding to Streptavidin. **1989**, *243* (4887), 85-88. DOI: 10.1126/science.2911722.

(44) Acinas, S. G.; Sarma-Rupavtarm, R.; Klepac-Ceraj, V.; Polz, M. F. PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. *Applied and Environmental Microbiology* **2005**, *71* (12), 8966 - 8969. DOI: 10.1128/aem.71.12.8966-8969.2005.

(45) Patel, K.; Walport, L. J.; Walshe, J. L.; Solomon, P. D.; Low, J. K. K.; Tran, D. H.; Mouradian, K. S.; Silva, A. P. G.; Wilkinson-White, L.; Norman, A.; et al. Cyclic peptides can engage a single binding pocket through highly divergent modes. *PNAS* **2020**, *117* (43), 26728-26738. DOI: 10.1073/pnas.2003086117.

(46) Iwasaki, K.; Goto, Y.; Katoh, T.; Suga, H. Selective thioether macrocyclization of peptides having the N-terminal 2-chloroacetyl group and competing two or three cysteine

residues in translation. *Organic & Biomolecular Chemistry* **2012**, *30* (13). DOI: 10.1039/c2ob25306b.

(47) Madeira, F. b.; Pearce, M.; Tivey, A. R.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research* **2022**, *50*, 276-279. DOI: 10.1093/nar/gkac240.

(48) Hillier, B. J.; Christopherson, K.; Prehoda, K. E.; Bredt, D. S.; Lim, W. A. Unexpected Modes of PDZ Domain Scaffolding Revealed by Structure of nNOS-Syntrophin Comple. **1999**, *284* (5415), 812-815. DOI: 10.1126/science.284.5415.812.

(49) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: making protein folding accessible to all. *Nature Methods* **2022**, *19*, 679-682. DOI: 10.1038/s41592-022-01488-1.

(50) *PDBePISA (Proteins, Interfaces, Structures and Assemblies)*. 2024. (accessed 2024-05-20).

(51) Reva, B. A.; Finkelstein, A. V.; Skolnick, J. What is the probability of a chance prediction of a protein structure with an rmsd of 6 å? *Folding and Design* **1998**, *3* (2), 141-147.

(52) Krieger, F.; Möglich, A.; Kiefhaber, T. Effect of Proline and Glycine Residues on Dynamics and Barriers of Loop Formation in Polypeptide Chains. *Journal of the American Chemical Society* **2005**, *127* (10), 3346-3352. DOI: 10.1021/ja042798i.

(53) Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **2021**, *50* (D1), D439-D444. DOI: 10.1093/nar/gkab1061.

(54) Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **2013**, *29* (21), 2722-2728. DOI: 10.1093/bioinformatics/btt473.

Appendix A

RaPID mRNA library composition

mRNA NNK7-11 5’-3’:

TAATACGACTCACTATAGGGTTAACTTTAAGAAGGAGATATACATATG(NNK)$_{7-11}$TGCGGCAGCGGCAGCGGCAGCTAGGACGGGGGGCGGAAA

Color coding:

TATA = TATA box –Translation will start after the TATA box.

AGGA = RBS – ribosome binding site

ATG = Translation initiation site

(NNK)$_{7-11}$ = The randomized library sequence (N=A/T/C/C, K=G/T)

TGC = Precoded downstream Cys residue for posttranslational thioether ring-closure

GGC = Gly residue

AGC = Ser residue

TAG = Translation termination site

GACGGGGGGCGG = mRNA spacer

Figure S1: The DNA-PEG-puromycin linker consists of a 5’-phosphorylated DNA oligo (complementary to the mRNA’s 3’ end), a PEG spacer (5x iSp18), two CC DNA bases and a puromycin molecule, which mimics the CCA codon of tyrosyl-tRNA.

Table S1: Fractions of individual NNK libraries prepared for the RaPID selection against PSD-95-PDZ2.

Library	Variability	Library fraction	DNA molecules in 100 µL library mixture
NNK7	3.44E+10	0.1	7.33E+12
NNK8	1.10E+12	1	7.33E+13
NNK9	3.52E+13	1	7.33E+13
NNK10	1.13E+15	1	7.33E+13
NNK11	3.60E+15	1	7.33E+13

Appendix B

qPCR data and DNA library recovery rate calculations

Table S2: qPCR data and calculated recovery PCR cycles for POS samples in all RaPID selection rounds against PDS-95-PDZ2.

Selection round	qPCR C_{DNA} POS (molecules/µL)	moles/µL	µmoles/L (µM)	Recovery PCR cycles
R1	8.27E+07	1.37E-16	1.37E-04	9.51
R2	4.25E+05	7.06E-19	7.06E-07	17.1
R3	2.64E+06	4.38E-18	4.38E-06	14.5
R4	5.71E+06	9.48E-18	9.48E-06	13.4
R5	7.95E+08	1.32E-15	1.32E-03	6.24
R6	1.41E+09	2.34E-15	2.34E-03	5.42
R7	1.82E+09	3.03E-15	3.03E-03	5.05

Table S3: qPCR data and calculated recovery PCR cycles for NEG samples in RaPID selection rounds two to seven against PDS-95-PDZ2.

Selection round	qPCR C_{DNA} NEG (molecules/µL)	moles/µL	µmoles/L (µM)	Recovery PCR cycles
R2	1.22E+08	2.02E-16	2.02E-04	7.95
R3	1.47E+08	2.44E-16	2.44E-04	7.68
R4	1.23E+08	2.05E-16	2.05E-04	7.93
R5	1.26E+08	2.10E-16	2.10E-04	7.90
R6	9.99E+07	1.66E-16	1.66E-04	8.24
R7	4.40E+07	7.30E-17	7.30E-05	9.42

Table S4: qPCR data for INPUT samples in all RaPID selection rounds against PDS-95-PDZ2. Recovery rates were calculated from qPCR data from INPUT, POS and NEG samples.

Selection round	qPCR C_{DNA} INPUT (molecules/µL)	INPUT molecules in translation solution	POS molecules in elute solution	NEG molecules in elute solution	Recovery rate POS (%)	Recovery rate NEG (%)
R1	3.09E+08	6.56E+13	4.14E+09		6.31E-03	
R2	1.55E+08	1.09E+12	2.13E+07	2.56E+09	1.94E-03	0.234
R3	1.82E+08	1.28E+12	1.32E+08	3.09E+09	1.03E-02	0.240
R4	8.57E+07	6.05E+11	2.85+E08	2.59E+09	4.72E-02	0.428
R5	1.97E+08	1.39E+12	3.97E+10	2.65E+09	2.86	0.191
R6	1.38E+08	9.71E+11	7.03E+10	2.10E+09	7.25	0.216
R7	2.09E+08	1.47E+12	9.12E+10	9.23E+08	6.19	6.27E-02

Appendix C

List of top peptide sequence enrichments for POS and NEG libraires

Table S5: Peptide sequence library fractions from next generation sequencing for POS samples in RaPID selection rounds two to seven against PDS-95-PDZ2. '*' corresponds to stop codons.

No	R2	R3	R4	R5	R6	R7	Peptide seq	Peptide name (POS)
1	1.91E-03	3.67E-02	0.299	0.608	0.741	0.780	FIETTFAGCRIETSFTC	EWA_PDZ_27001
2	1.47E-04	2.88E-03	5.46E-03	1.49E-02	3.99E-02	7.95E-02	FIETTFAGNRIETSFTC	EWA_PDZ_27002
3	2.93E-04	1.03E-02	0.195	0.192	8.42E-02	3.19E-02	FIETSFAGNRIETSFTC	EWA_PDZ_27003
4	0	3.85E-04	3.28E-03	5.37E-03	1.15E-02	1.01E-02	FIETSFAGNRIETSFTCGC	EWA_PDZ_27004
5	0	4.71E-03	6.60E-02	9.12E-02	3.41E-02	8.10E-03	FITTFTVHLITTFTVHC	EWA_PDZ_27005
6	0	0	1.09E-03	1.84E-03	4.61E-03	4.67E-03	FITTFTVHLITTFTVHC	EWA_PDZ_27006
7	0	2.88E-04	7.80E-04	2.76E-03	3.13E-03	3.50E-03	FIETTFAGCRIETSFTR	EWA_PDZ_27007
8	0	0	7.80E-04	3.53E-04	2.97E-03	3.50E-03	FIETSFAGNRIETSFYC	EWA_PDZ_27008
9	0	0	0	1.27E-03	2.31E-03	3.50E-03	FIETSFAGNRIETSFFC	EWA_PDZ_27009
10	0	9.61E-05	1.58E-04	8.48E-04	1.65E-03	2.98E-03	FIETTFAGGRIETSFTC	EWA_PDZ_27010
11	0	0	1.58E-04	6.36E-04	9.88E-04	2.66E-03	FIETTFAGCRIETSFSC	EWA_PDZ_27011
12	0	1.92E-04	4.68E-04	1.34E-03	1.15E-03	1.75E-03	FIETTFAGCRIETSFTWAX	EWA_PDZ_27012
13	0	9.61E-05	0	1.41E-04	1.15E-03	1.68E-04	FIETTFAGYRIETSFTC	EWA_PDZ_27013
14	0	9.61E-05	3.12E-04	6.36E-04	9.88E-04	1.68E-04	FIETSFAGNRIETTFTC	EWA_PDZ_27014
15	0	0	9.36E-04	1.48E-03	1.48E-03	1.62E-03	FIETTFAGCRIETSFTFGS	EWA_PDZ_27015
16	0	0	0	2.12E-04	4.94E-04	1.23E-03	FIETSFAGNRIETSFTLC	EWA_PDZ_27016
17	0	1.92E-04	1.72E-03	2.47E-03	1.48E-03	1.23E-03	FIETSFAGCRIETSFTC	EWA_PDZ_27017
18	0	0	0	2.12E-04	4.94E-04	1.04E-03	FIETTFAGWRIETSFTC	EWA_PDZ_27018
19	0	0	0	7.78E-04	6.59E-04	9.72E-04	YTYD*DYFCWLSY*DFVYLR	EWA_PDZ_27019
20	0	0	3.12E-04	2.12E-04	4.94E-04	9.72E-04	FIETTFSGCRIETSFTC	EWA_PDZ_27020
21	0	0	0	5.65E-04	4.94E-04	9.72E-04	FIETTFAGCRIETSFTSAX	EWA_PDZ_27021
22	0	0	1.56E-04	7.07E-04	3.29E-04	9.72E-04	FIETTFAGCRIETSFTCAX	EWA_PDZ_27022
23	0	9.61E-05	0	4.24E-04	3.29E-04	9.07E-04	YT*LRLLLLVVVLRLRLLA	EWA_PDZ_27023
24	0	0	1.56E-04	9.90E-04	9.88E-04	9.07E-04	FIETSFVGNRIETSFTC	EWA_PDZ_27024
25	0	0	0	2.12E-04	3.29E-04	7.13E-04	FIETTFTGCRIETSFTC	EWA_PDZ_27025
26	0	0	4.68E-04	4.95E-004	9.88E-04	6.48E-04	YTHD*DYFCWLSY*DFVYLR	EWA_PDZ_27026

27	0	0	3.12E-04	1.41E-04	3.29E-04	6.48E-04	FIETTFAGCRIETSFTCGQX	EWA_PDZ_27027
28	0	9.61E-05	0	0	8.24E-04	6.48E-04	FIETSFAGNRIETSFTIC	EWA_PDZ_27028
29	0	0	3.12E-04	7.07E-05	6.59E-04	5.83E-04	FIETTFAGNRIETSFTCGC	EWA_PDZ_27029
30	0	0	0	1.41E-04	4.94E-04	5.83E-04	FIETTFAGFRIETSFTC	EWA_PDZ_27030
31	0	9.61E-05	7.80E-04	6-36E-04	8.24E-04	5.83E-04	FIETTFAGCRIETSFIC	EWA_PDZ_27031
32	0	9.61E-05	1.56E-04	5.65E.04	1.15E-03	5.18E-04	FIETTFVGCRIETSFTC	EWA_PDZ_27032
33	0	0	0	0	8.24E-04	5.18E-04	FIETTFAGCRIETSFFC	EWA_PDZ_27033
34	0	9.61E-05	0	7.07E-05	8.24E-04	4.51E-04	FIETTFAGRRIETSFTC	EWA_PDZ_27034
35	0	0	0	1.41E-04	3.29E-04	4.51E-04	FIETTFAGCRIETSFYC	EWA_PDZ_27035
36	0	0	3.12E-04	3.53E-04	1.15E-03	4.51E-04	FIETTFAGCRIETLFTC	EWA_PDZ_27036
37	0	0	0	3.53E-04	3.29E-04	4.51E-04	FIETSFAGNRIETSFTGC	EWA_PDZ_27037
38	0	0	0	4.95E-04	4.94E-04	3.89E-04	FIETTLLVVVLRLRLLA	EWA_PDZ_27038
39	0	0	0	2.12E.04	1.65E-04	3.89E-04	FIETTFAGCRIETSLTC	EWA_PDZ_27039
40	0	0	0	2.83E-04	1.65E-04	3.89E-04	FIETTFAGCRIETSFTCSS	EWA_PDZ_27040
41	0	0	1.56E-04	5.65E-04	3.29E-04	3.89E-04	FIETSFAGNRIETSFSC	EWA_PDZ_27041
42	0	0	0	1.41E-04	0	3.89E-04	FIETSFAGNRIETSFLGC	EWA_PDZ_27042
43	0	0	0	0	3.29E-04	3.24E-04	FIETTFAGNRIETSFTR	EWA_PDZ_27043
44	0	0	1.56E-04	2.12E-04	1.65E-04	3.24E-04	FIETTFAGCRIETSFAC	EWA_PDZ_27044
45	0	0	0	7.07E-05	1.65E-04	3.24E-04	FIETSFAGNRIETTFYC	EWA_PDZ_27045
46	0	0	0	1.41E-04	3.29E-04	3.24E-04	FIETSFAGNRIETSFTVC	EWA_PDZ_27046
47	0	0	0	7.07E-05	4.94E-04	3.24E-04	FIETSFAGNRIETSFTCC	EWA_PDZ_27047
48	0	0	0	0	1.65E-04	2.59E-04	FIETTFAYNRIETSFTC	EWA_PDZ_27048
49	0	0	0	0	0	2.59E-04	FIETTFAGNRIETSFTFGS	EWA_PDZ_27049
50	0	0	0	7.07E-05	0	2.59E-04	FIETTFAGCRIETSFTGGS	EWA_PDZ_27050
51	0	0	0	0	0	2.59E-04	FIETTFAGCRIETSFTCGTAAAATQ*DRKSTRX	EWA_PDZ_27051
52	0	0	1.56E-04	3.53E-04	6.59E-04	2.59E-04	FIETSFFAGNRIETSFTC	EWA_PDZ_27052
53	0	0	3.12E-04	7.78E-04	3.29E-04	2.59E-04	FIETSFAGNRIETSFIC	EWA_PDZ_27053
54	0	0	0	2.12E-04	3.29E-04	2.59E-04	FIETSFAGNRIETSFGIC	EWA_PDZ_27054
55	0	0	0	7.07E-05	0	1.94E-04	YI*LRLLLLVVVLRLRLLA	EWA_PDZ_27055

56	0	0	0	7.07E-05	0	1.94E-04	FIVTTFAGCRIETSFTC	EWA_PDZ_27056
57	0	0	4.68E.04	9.19E-04	3.29E-04	1.94E-04	FITTFTVHLITTFTVHFGS	EWA_PDZ_27057
58	0	0	0	1.41E-04	3.29E-04	1.94E-04	FIETTLAGCRIETSFTC	EWA_PDZ_27058
59	0	0	0	2.12E.04	0	1.94E-04	FIETTFASCRIETSFTC	EWA_PDZ_27059
60	0	0	0	1.41E-04	1.65E-04	1.94E-04	FIETTFAGCRTETSFTC	EWA_PDZ_27060
61	0	0	0	1.41E-04	0	1.94E-04	FIETTFAGCRIVTSFTC	EWA_PDZ_27061
62	0	0	0	7.07E-05	0	1.94E-04	FIETTFAGCRIGTSFTC	EWA_PDZ_27062
63	0	0	7.80E-04	3.53E-04	9.88E-04	1.94E-04	FIETTFAGCRIETSFTRGS	EWA_PDZ_27063
64	0	0	0	0	0	1.94E-04	FIETTFAGCRIETSFTLC	EWA_PDZ_27064
65	0	0	0	0	0	1.94E-04	FIETTFAGCRIETSFTCGKRQRQLLSEIGRAHX	EWA_PDZ_27065
66	0	0	0	0	0	1.94E-04	FIETTFAGCRIETSFTCGAAAAATQ*DRKSTRX	EWA_PDZ_27066
67	0	0	0	2.12E-04	1.65E-04	1.94E-04	FIETTFAGCRIETSFTAGS	EWA_PDZ_27067
68	0	9.61E-05	1.56E-04	0	0	1.94E-04	FIETTFAGCRIET*FTC	EWA_PDZ_27068
69	0	0	0	7.07E-05	0	1.94E-04	FIETTFAGCRIESSFTC	EWA_PDZ_27069
70	0	0	0	7.07E-05	0	1.94E-04	FIETTFA	EWA_PDZ_27070
71	0	0	0	0	0	1.94E-04	FIETSFAGNRIETSFTTC	EWA_PDZ_27071
72	0	0	0	1.41E-04	1.65E-04	1.94E-04	FIESTFAGCRIETSFTC	EWA_PDZ_27072
73	0	0	0	7.07E-05	3.29E-04	1.94E-04	FIEATFAGCRIETSFTC	EWA_PDZ_27073
74	0	0	0	0	0	1.94E-04	FFETTFAGCRIETSFTC	EWA_PDZ_27074
75	0	0	0	7.07E-05	0	1.30E-04	YTYN*DYFCWLSY*DFVYLR	EWA_PDZ_27075
76	0	0	0	0	0	1.30E-04	YTHD*DYFCW*SY*DFVYLR	EWA_PDZ_27076
77	0	0	0	0	0	1.30E-04	YTFIETTFAGCRIETSFTC	EWA_PDZ_27077
78	0	0	0	7.07E-05	0	1.30E-04	YT*LRLLLLVVVLRLRLPA	EWA_PDZ_27078
79	0	0	0	2.38E-04	3.29E-04	1.30E-04	YT*LRLLLLVIVLRLRLLA	EWA_PDZ_27079
80	0	0	0	0	1.65E-04	1.30E-04	YT*LRLLLLVIVLRLRFFA	EWA_PDZ_27080
81	0	0	0	7.07E-05	1.65E-04	1.30E-04	YPYD*DYFCWLSY*DFVYLR	EWA_PDZ_27081
82	0	0	0	7.07E-05	0	1.30E-04	SPSSFSSFCWLSY*DFVYLR	EWA_PDZ_27082
83	0	0	0	7.07E-05	1.65E-04	1.30E-04	SPSSCDYFCWLSY*DFVYLR	EWA_PDZ_27083
84	0	4.81E-04	1.15E-02	4.17E-03	3.29E-04	1.30E-04	FQDWYPLWYCLHYC	EWA_PDZ_27084

85	0	0	1.56E-04	0	1.65E-04	1.30E-04	FNETTFAGCRIETSFTC	EWA_PDZ_27085
86	0	4.81E-04	2.80E-03	1.13E-03	4.94E-04	1.30E-04	FITTFTVHLITTFA	EWA_PDZ_27086
87	0	0	0	7.07E-05	0	1.30E-04	FIRLLLLVVVLRLRLLA	EWA_PDZ_27087
88	0	0	0	2.12E-04	1.65E-04	1.30E-04	FIETTSAGCRIETSFTC	EWA_PDZ_27088
89	0	0	0	0	1.65E-04	1.30E-04	FIETTFAGSRIETSFTC	EWA_PDZ_27089
90	0	0	0	0	0	1.30E-04	FIETTFAGNRIETTFTC	EWA_PDZ_27090
91	0	0	0	0	0	1.30E-04	FIETTFAGNRIETSFTTC	EWA_PDZ_27091
92	0	0	0	0	0	1.30E-04	FIETTFAGNRIETSFTCVS	EWA_PDZ_27092
93	0	0	0	0	0	1.30E-04	FIETTFAGNRIETSFTCGRRQRQLLSEIGRAHX	EWA_PDZ_27093
94	0	0	0	0	1.65E-04	1.30E-04	FIETTFAGNRIETLFTC	EWA_PDZ_27094
95	0	0	0	7.07E-05	3.29E-04	1.30E-04	FIETTFAGCRIETTFTC	EWA_PDZ_27095
96	0	0	0	7.07E-05	0	1.30E-04	FIETTFAGCRIETSLLA	EWA_PDZ_27096
97	0	0	0	2.12E-04	3.29E-04	1.30E-04	FIETTFAGCRIETSFYLR	EWA_PDZ_27097
98	0	0	1.56E-04	7.07E-05	4.94E-04	1.30E-04	FIETTFAGCRIETSFTVGS	EWA_PDZ_27098
99	0	0	0	7.07E-05	0	1.30E-04	FIETTFAGCRIETSFTSGS	EWA_PDZ_27099
100	0	0	0	7.07E-05	0	1.30E-04	FIETTFAGCRIETSFTIC	EWA_PDZ_27100

Table S6: Peptide sequence library fractions from next generation sequencing for NEG samples in RaPID selection rounds three to seven against PDS-95-PDZ2. '*' corresponds to stop codons.

No	R3	R4	R5	R6	R7	Peptide seq	Peptide name (NEG)
1	4.43E-03	4.04E-03	0.107	0.321	0.590	FIETTFAGCRIETSFTC	EWA_PDZ_27001
2	3.94E-03	4.77E-03	0.157	0.396	0.226	FITTFTVHLITTFTVHC	EWA_PDZ_27002
3	4.43E-03	1.62E-03	4.41E-02	7.87E-02	3.45E-02	FIETSFAGNRIETSFTC	EWA_PDZ_27003
4	1.64E-03	3.24E-04	1.02E-03	6.45E-03	2.57E-02	FIETTFAGNRIETSFTC	EWA_PDZ_27004
5	8.21E-05	4.85E-04	5.20E-03	1.32E-02	6.26E-03	FITTFTVHLITTFA	EWA_PDZ_27005
6	0	0	7.14E-04	2.78E-03	4.95E-03	FIETSFAGNRIETSFTCGC	EWA_PDZ_27006
7	8.21E-05	0	6.12E-04	2.33E-03	3.94E-03	FIETTFAGCRIETSFTCGC	EWA_PDZ_27007

8	0	0	8.16E-04	4.39E-03	2.24E-03	FITTFTVHLITTFTVHCGC	EWA_PDZ_27008
9	0	0	8.16E-04	2.60E-03	2.09E-03	FIETTLLVVVLRLRLLA	EWA_PDZ_27009
10	0	0	0	1.79E-04	1.70E-03	FIETTFAGYRIETSFTC	EWA_PDZ_27010
11	0	0	3.06E-04	1.43E-03	1.70E-03	FIETTFAGCRIETSFTA	EWA_PDZ_27011
12	0	0	0	6.27E-04	1.55E-03	FIETSFAGNRIETSFYC	EWA_PDZ_27012
13	0	0	5.10E-04	1.34E-03	1.55E-03	FIETSFAGCRIETSFTC	EWA_PDZ_27013
14	8.21E-05	1.62E-04	1.02E-02	6.72E-03	1.31E-03	FQDWYPLWYCLHYC	EWA_PDZ_27014
15	0	0	1.02E-03	1.52E-03	1.24E-03	FITTFTVHLITTFTVHFGS	EWA_PDZ_27015
16	0	0	1.02E-04	5.37E-04	1.16E-03	FIETSFAGNRIETSFFC	EWA_PDZ_27016
17	0	0	2.04E-03	4.48E-04	1.08E-03	FIETTFVGCRIETSFTC	EWA_PDZ_27017
18	0	0	0	4.48E-04	1.01E-03	FIETTFAGWRIETSFTC	EWA_PDZ_27018
19	0	0	0	1.79E-04	1.01E-03	FIETTFAGFRIETSFTC	EWA_PDZ_27019
20	0	8.09E-05	4.08E-04	6.27E-04	1.01E-03	FIETTFAGCRIETSFTRv	EWA_PDZ_27020
21	8.21E-05	0	3.06E-04	1.52E-03	9.28E-04	FITTFTVHLITTFTVLC	EWA_PDZ_27021
22	0	8.09E-05	0	3.58E-04	8.50E-04	YTHD*DYFCWLSY*DFVYLR	EWA_PDZ_27022
23	0	0	1.02E-04	1.79E-04	8.50E-04	FITTFTVRLITTFTVHC	EWA_PDZ_27023
24	0	0	1.02E-04	1.79E-04	8.50E-04	FIETTFAGCRIETSFSC	EWA_PDZ_27024
25	0	0	4.08E-04	5.37E-04	7.73E-04	FIETTFAGCRIETSFTWAX	EWA_PDZ_27025
26	0	0	0	3.58E-04	7.73E-04	FIETSFVGNRIETSFTC	EWA_PDZ_27026
27	0	0	0	0	6.96E-04	YTYY*DYFCWLSY*DFVYLR	EWA_PDZ_27027
28	0	0	2.04E-04	5.37E-04	6.96E-04	FITTFTVHLITTFTVLR	EWA_PDZ_27028
29	0	0	1.02E-04	8.59E-04	6.96E-04	FITTFTVHLITTFTVHWAX	EWA_PDZ_27029
30	0	8.09E-05	1.02E-04	4.48E-04	6.96E-04	FIETTFCWLSY*DFVYLR	EWA_PDZ_27030
31	0	0	0	0	6.19E-04	FITTFTVHLITTFAAAAATQ*DRKSTRLNSSHX	EWA_PDZ_27031

32	0	0	0	0	5.41E-04	FIETTFAGCRIETSFTCGA AAAATQ*DRKSTRX	EWA_PDZ_27032
33	0	0	2.04E-04	1.79E-04	4.64E-04	YT*LRLLLLVVVLRLRLLA	EWA_PDZ_27033
34	8.21E-05	0	2.14E-03	1.07E-03	4.64E-04	FIIYKNRILGYGC	EWA_PDZ_27034
35	0	0	0	8.95E-05	4.64E-04	FIETTFSGCRIETSFTC	EWA_PDZ_27035
36	0	0	5.10E-04	4.48E-04	4.64E-04	FIETTFAGCRIETSFTFGS	EWA_PDZ_27036
37	0	8.09E-05	5.31E-03	2.24E-03	4.64E-04	FFVITWIDSKAFIFPC	EWA_PDZ_27037
38	0	8.09E-05	4.08E-04	3.58E-04	3.87E-04	YTYD*DYFCWLSY*DFVYL R	EWA_PDZ_27038
39	0	0	5.10E-04	1.79E-04	3.87E-04	FITTFTVHLITTFTVRC	EWA_PDZ_27039
40	0	0	0	4.48E-04	3.87E-04	FITTFTVHC	EWA_PDZ_27040
41	0	0	0	0	3.87E-04	FIETTLLVVVLRLRLLAAA AATQ*DRKSTRLNX	EWA_PDZ_27041
42	0	8.09E-05	2.04E-04	1.79E-04	3.87E-04	FIETTFAGRRIETSFTC	EWA_PDZ_27042
43	0	0	0	0	3.87E-04	FIETTFAGNRIETSFTCGC	EWA_PDZ_27043
44	0	0	0	1.79E-04	3.87E-04	FIETTFAGCRIETSLTC	EWA_PDZ_27044
45	0	0	0	0	3.87E-04	FIETTFAGCRIETSFYC	EWA_PDZ_27045
46	0	0	1.02E-04	3.58E-04	3.87E-04	FIETTFAGCRIETSFTCAX	EWA_PDZ_27046
47	8.21E-05	0	1.02E-04	2.69E-04	3.87E-04	FIETTFAGCRIETSFIC	EWA_PDZ_27047
48	0	2.43E-04	3.57E-03	1.97E-03	3.09E-04	FNILVWRRGIPSLWNVC	EWA_PDZ_27048
49	0	0	1.02E-04	0	3.09E-04	FITTTFAGCRIETSFTC	EWA_PDZ_27049
50	0	0	2.04E-04	1.79E-04	3.09E-04	FITTFTVHLITTFTVQC	EWA_PDZ_27050
51	0	0	0	0	3.09E-04	FITTFTVHLITTFTVHCCS	EWA_PDZ_27051
52	0	0	4.08E-04	9.85E-04	3.09E-04	FITTFTVHLITTFTVHA	EWA_PDZ_27052
53	0	1.62E-04	2.45E-03	2.24E-03	3.09E-04	FIRFYKGQWYISC	EWA_PDZ_27053
54	0	0	0	0	3.09E-04	FIETTFAGNRIETSFTA	EWA_PDZ_27054
55	0	0	0	2.69E-04	3.09E-04	FIETTFAGGRIETSFTC	EWA_PDZ_27055

56	0	0	0	3.58E-04	3.09E-04	FIETTFAGCRIETSFTSAX	EWA_PDZ_27056
57	0	0	0	0	3.09E-04	FIETTFAGCRIETSFTCGG GSGSYSVRSEEHTX	EWA_PDZ_27057
58	0	0	1.02E-04	8.95E-05	3.09E-04	FIETTFAGCRI*TSFTC	EWA_PDZ_27058
59	8.21E-05	0	3.06E-04	0	3.09E-04	FIETTCWLSY*DFVYLR	EWA_PDZ_27059
60	0	0	0	8.95E-05	3.09E-04	FIETSFAGNRIETSFTLC	EWA_PDZ_27060
61	0	0	0	1.79E-04	3.09E-04	FIETSFAGNRIETSFTIC	EWA_PDZ_27061
62	0	0	0	0	3.09E-04	FIETSFAGNRIETSFTCC	EWA_PDZ_27062
63	0	0	1.02E-04	8.95E-05	2.32E-04	YTYIETTFAGCRIETSFTC	EWA_PDZ_27063
64	0	1.62E-04	1.33E-03	8.95E-05	2.32E-04	FTRLGRFTTHGTFLYIC	EWA_PDZ_27064
65	0	1.62E-04	3.57E-03	2.33E-03	2.32E-04	FLIVWTVPRRWLC	EWA_PDZ_27065
66	0	0	1.02E-04	4.48E-04	2.32E-04	FITTLRFI*LRLLRFIA	EWA_PDZ_27066
67	0	0	0	9.85E-04	2.32E-04	FITTFTVHLITTFTVYC	EWA_PDZ_27067
68	0	1.62E-04	4.08E-04	4.49E-05	2.32E-04	FITTFTVHLITTFTVHLR	EWA_PDZ_27068
69	0	0	0	8.95E-05	2.32E-04	FITTFTFHLITTFTVHC	EWA_PDZ_27069
70	0	0	0	0	2.32E-04	FIGTTFAGCRIETSFTC	EWA_PDZ_27070
71	0	0	1.02E-04	1.79E-04	2.32E-04	FIETTLLVVVLRLRLPA	EWA_PDZ_27071
72	0	0	2.04E-04	2.69E-04	2.32E-04	FIETTLAGCRIETSFTC	EWA_PDZ_27072
73	0	0	1.02E-04	8.95E-05	2.32E-04	FIETTFAVVVLRLRLLA	EWA_PDZ_27073
74	0	0	1.02E-04	1.79E-04	2.32E-04	FIETTFASCRIETSFTC	EWA_PDZ_27074
75	0	0	0	0	2.32E-04	FIETTFAGCRLETSFTC	EWA_PDZ_27075
76	0	0	3.06E-04	4.48E-04	2.32E-04	FIETTFAGCRIETSFTRGS	EWA_PDZ_27076
77	0	0	0	1.79E-04	2.32E-04	FIETTFAGCRIETSFTGGS	EWA_PDZ_27077
78	0	0	0	0	2.32E-04	FIETTFAGCRIETSFTCGR GSGSYSVRSEEHTX	EWA_PDZ_27078
79	0	0	0	8.95E-05	2.32E-04	FIETTFADCRIETSFTC	EWA_PDZ_27079

80	0	0	0	8.95E-05	2.32E-04	FIETSFAGNRIETSFTGC	EWA_PDZ_27080
81	0	0	2.04E-04	5.37E-04	2.32E-04	FIETSFAGNRIETSFTA	EWA_PDZ_27081
82	0	0	3.06E-04	2.69E-04	2.32E-04	FIETSFAGNRIETSFSC	EWA_PDZ_27082
83	0	0	0	1.79E-04	2.32E-04	FIETSFAGNRIETSFFA	EWA_PDZ_27083
84	0	0	2.04E-04	3.58E-04	2.32E-04	FIETFTVHLITTFTVHC	EWA_PDZ_27084
85	0	0	0	0	1.65E-04	FIATTFAGCRIETSFTC	EWA_PDZ_27085
86	0	0	1.02E-04	8.95E-05	2.32E-04	FIATFTVHLITTFTVHC	EWA_PDZ_27086
87	0	0	0	0	2.32E-04	FI*TTFAGCRIETSFTC	EWA_PDZ_27087
88	0	0	0	0	2.32E-04	FFETTFAGCRIETSFTC	EWA_PDZ_27088
89	0	0	0	0	1.55E-04	YTYN*DYFCWLSY*DFVYL RQRQLLSEIGRAHV*T	EWA_PDZ_27089
90	0	0	0	0	1.55E-04	YTSD*DYFCWLSY*DFVYL R	EWA_PDZ_27090
91	0	0	0	0	1.55E-04	YTHDYDFYGSFDYDFYGSL RQRQLLSEIGRAHV*T	EWA_PDZ_27091
92	0	0	0	0	1.55E-04	YT*LRLLLLVVVLRLRLLA AAAATQ*DRKSTRLNS	EWA_PDZ_27092
93	0	0	0	1.79E-04	1.55E-04	YT*D*DYFCWLSY*DFVYL R	EWA_PDZ_27093
94	0	0	0	2.69E-04	1.55E-04	YIYD*DYFCWLSY*DFVYL R	EWA_PDZ_27094
95	0	0	0	8.95E-05	1.55E-04	YD*DYFCWLSY*DFVYLR	EWA_PDZ_27095
96	0	0	0	0	1.55E-04	SPSSCDSFCWLSY*DFVYL R	EWA_PDZ_27096
97	0	0	1.02E-04	8.95E-05	1.55E-04	FIVTTFAGCRIETSFTC	EWA_PDZ_27097
98	0	0	2.04E-04	8.95E-05	1.55E-04	FITTLTVHLITTFTVHC	EWA_PDZ_27098
99	0	0	0	8.95E-05	1.55E-04	FITTFTVYLITTFTVHC	EWA_PDZ_27099
100	0	0	0	0	0	FITTFTVHLITTYGSLR	EWA_PDZ_27100

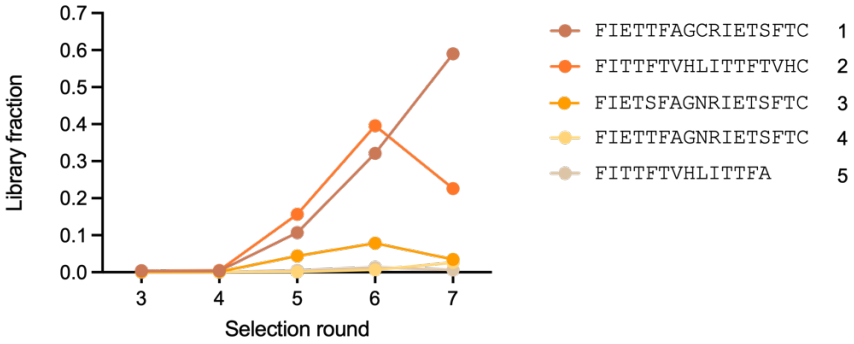

Figure S2: RaPID negative enrichments. NGS data analysis of the negative selection of RaPID rounds three to seven, comparing the fraction of DNA reads for each unique peptide sequence.

Appendix D

ColabFold supplementary information

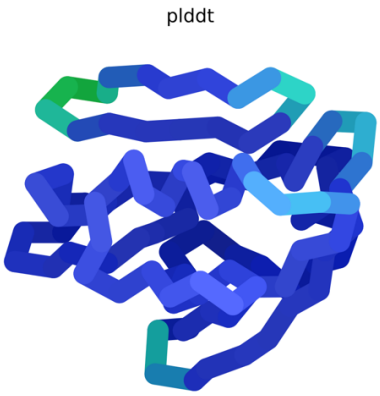

Figure S3: Model confidence of PDZ2/JFL03. Different colors indicate ColabFolds's level of confidence in its prediction. Very high predicted local distance difference test (plDDT > 90) corresponds to blue, plDDT < 50 to red.[53]

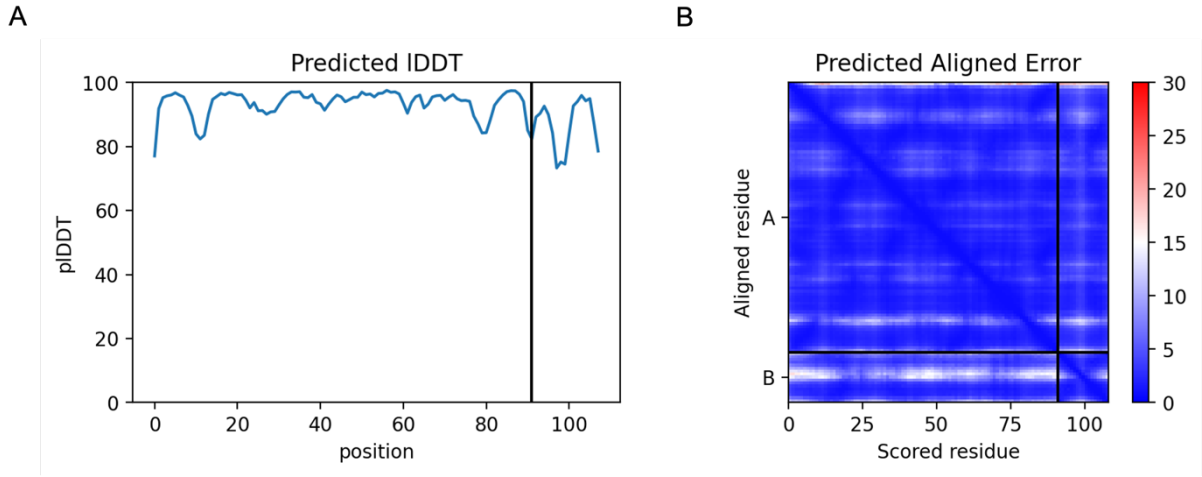

Figure S4: ColabFold confidence measures. A) Predicted Local Distance Difference Test (plDDT) is a superposition-free score that evaluates local distance differences of all atoms in the model.[54] **B)** Predicted aligned error (PAE) is a measure of how confident ColbFold is in the relative position of two residues within the predicted structure. PAE is defined as the expected positional error at residue X, measured in Å, if the predicted and actual residue were aligned on residue Y.[53]

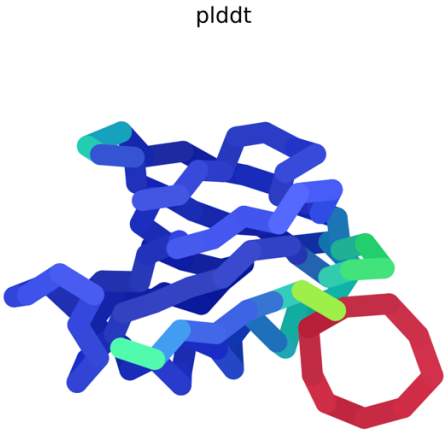

Figure S5: Model confidence of PDZ2/EWA_PDZ_27001. Different colors indicate ColabFolds's level of confidence in its prediction. Very high predicted local distance difference test (plDDT > 90) corresponds to blue, plDDT < 50 to red.[53]

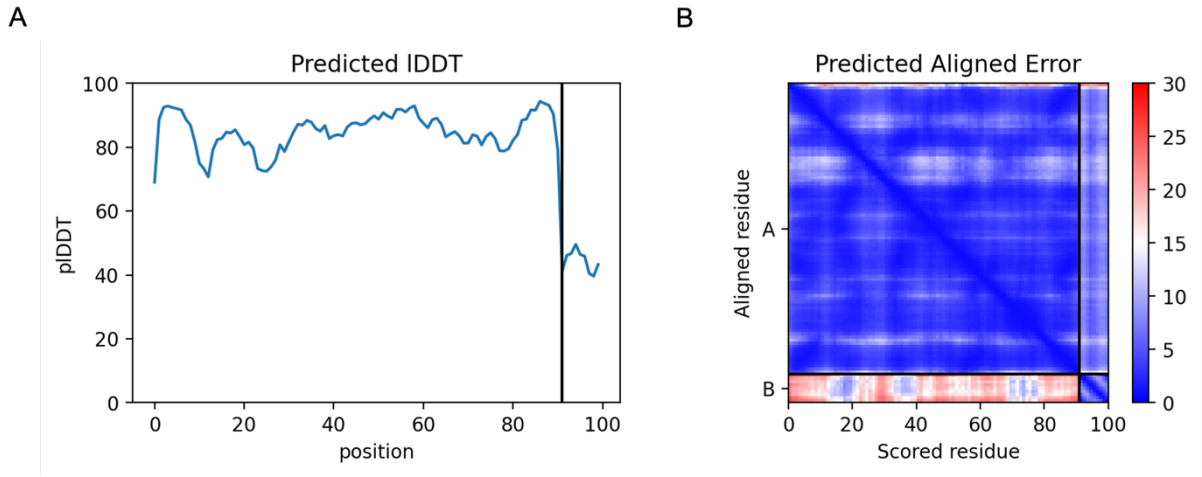

Figure S6: ColabFold confidence measures. A) Predicted Local Distance Difference Test (plDDT) is a superposition-free score that evaluates local distance differences of all atoms in the model.[54] **B)** Predicted aligned error (PAE) is a measure of how confident ColbFold is in the relative position of two residues within the predicted structure. PAE is defined as the expected positional error at residue X, measured in Å, if the predicted and actual residue were aligned on residue Y.[53]

Appendix E

PSD-95 protein sequence

Sequence of the human post synaptic density protein 95 (PSD-95)

|UniProt: P78352|

MDCLCIVTTKKYRYQDEDTPPLEHSPAHLPNQANSPPVIVNTDTLEAPGYELQVNGTEGE
MEYEEITLERGNSGLGFSIAGGTDNPHIGDDPSIFITKIIPGGAAAQDGRLRVNDSILFV
NEVDVREVTHSAAVEALKEAGSIVRLYVMRRKPPAEKVMEIKLIKGPKGLGFSIAGGVGN
QHIPGDNSIYVTKIIEGGAAHKDGRLQIGDKILAVNSVGLEDVMHEDAVAALKNTYDVVY
LKVAKPSNAYLSDSYAPPDITTSYSQHLDNEISHSSYLGTDYPTAMTPTSPRRYSPVAKD
LLGEEDIPREPRRIVIHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGDQIL
SVNGVDLRNASHEQAAIALKNAGQTVTIIAQYKPEEYSRFEAKIHDLREQLMNSSLGSGT
ASLRSNPKRGFYIRALFDYDKTKDCGFLSQALSFRFGDVLHVIDASDEEWWQARRVHSDS
ETDDIGFIPSKRRVERREWSRLKAKDWGSSSGSQGREDSVLSYETVTQMEVHYARPIIIL
GPTKDRANDDLLSEFPDKFGSCVPHTTRPKREYEIDGRDYHFVSSREKMEKDIQAHKFIE
AGQYNSHLYGTSVQSVREVAEQGKHCILDVSANAVRRLQAAHLHPIAIFIRPRSLENVLE
INKRITEEQARKAFDRATKLEQEFTECFSAIVEGDSFEEIYHKVKRVIEDLSGPYIWVPA
RERL