



Targeting PRMT5 Protein-Protein Interactions with Peptidic Modulators

Dissertation

for the achievement of the academic degree of the
Doctor in Natural Sciences
(Dr. rer. nat.)

Submitted to
the Faculty of Chemistry and Chemical Biology
Technical University Dortmund

by
Adrian Krzyzanowski, M.Sc.

born on 4th August 1994
in Inowrocław, Poland

Dortmund, 2022

The work presented in this thesis was performed during the time period from October 2018 to November 2022 under the supervision of Prof. Dr. Dr. h.c. Herbert Waldmann and the guidance of Dr. Peter 't Hart at the Faculty of Chemistry and Chemical Biology of the Technical University Dortmund and the Max Planck Institute of Molecular Physiology, Dortmund.

1st Examiner: Prof. Dr. Dr. h.c. H. Waldmann

2nd Examiner: Dr. A. Brunschweiler

*Dla Mamy i Taty
oraz Babci i Dziadka*

Acknowledgements

The past few years have been definitely a fascinating journey, full of daring challenges and rewarding experiences, packed with equal amounts of disappointments as well as pleasant surprises and successes. I would like to believe that the events of last four years let me grow in equal measure both as a scientist and a person, producing in the result a more mature and experienced researcher, that can tackle even the most taxing problems. However, during this journey I was never on my own and I do not think I would reach the current point without the help and support of a number of dear to me mentors, friends and family.

There are many people I need to thank but I will start with the captain of the ship. This whole experience would not have been possible without Prof. Waldmann giving me a chance and accepting me to his amazing group, here in Dortmund. Prof. Waldmann is one of the most impressive scientists that I had a pleasure to interact with, with an amazing pack of experience and knowledge, and I am absolutely grateful that I was able to learn from him. Prof. Waldmann was always full of support and understanding, especially when things were getting tough. I am also very grateful to Prof. Waldmann for creating a warm and welcoming atmosphere in the group, that always felt like second home to me.

I thank Dr. Peter 't Hart for the amazing supervision over the years. Peter played a pivotal role in my scientific growth, sharing his vast research expertise with me, while always staying positive and upbeat. Peter always prioritised a personal well-being of students over the ruthless progress in projects, always providing the necessary support and being the solid rock and compass, which kept me sane even during the most challenging times. I cannot express how grateful I am for having the privilege of working on the scientific projects with him. Peter also taught me a lot about how an exemplary leader should act and how he should treat his inferiors, a lesson that surely will be invaluable for me in the future.

Dr. Hélène Adihou and Dr. Romain Tessier both deserve a massive hug for the countless discussions about scientific and not so scientific topics, always having fresh perspectives on the things in the lab and the things in life, teaching me so much about the scientific reality. I am thankful for all the support and encouragement that I was given. I could not imagine working on my doctoral degree without them always having my back.

I also thank Dr. Raphael Gasper-Schönenbrücher for his tireless assistance. Raphael was always eager to brainstorm and provide the best support he was able to, always giving 110%. I very much enjoyed the hours that we spent together in the cold room fishing crystals, freezing our bottoms off.

Dr. Jana Flegel, Aylin Binici and Lara Dötsch are absolute superstars that provided me with the basics necessary for performing biological experiments on my own, answering five million often less than smart questions, while always being positive, supportive and encouraging. I also thank them for their friendship and invaluable mental support.

Sasikala Thavam and Jens Warmers and Beate Schölermann were always there for me when I needed help with machines or laboratory matters. I really appreciate their help over the years, as I could always count on them. I also would like to thank Thomas Manteufel, a trainee at the time, for help with a preparation of a number of chemical samples. Dr. Mahyar Akbarzadeh is thanked for his help with biological experiments and exploratory analysis of my peptides.

I absolutely need to thank Dr. Georg Niggemeyer and Dr. Caitlin Davies for constant support on the personal front. They were always there for me when I needed to talk to somebody, providing a very healthy and balanced perspective on everything happening around. Similar acknowledgement goes to Dr. Stefan Zimmermann.

My thanks also go to Dr. Michael Grigalunas and Dr. Axel Pahl for all the support, discussions and collaborations during the final phase of my doctoral degree and my time at the MPI. I very much appreciate their quite often ingenious ideas and all the time that they devoted for meetings with me. I learned a lot from them.

I cannot express enough my gratitude to the members of 't Hart group, especially Jessica Nowacki, Joseph Openy, Stefan Schmeing, Gulshan Amrahova, Jen-Yao Chang and Dr. Sunit Pal. I could always count on their help whether in a professional or more personal context. Jessica Nowacki and Joseph Openy deserve here a particular recognition. I feel exceptionally lucky for the opportunity to work with them.

Yannik Mantel is thanked for his unwavering enthusiasm and endless supply of encouragement over the years. I would also like to thank this family for the support.

Frau Rose provided invaluable support on the non-scientific front, helping me with a flood of German documents and paper work. I also would like to thank her for the absolutely perfect organisation of various matters in the group and support with miscellaneous concerns during my time at the MPI.

Protein Chemical Facility, or as it was formerly known, Dortmund Protein Facility is kindly acknowledged for the contribution to my doctoral projects. I especially would like to thank Dr. Jan-Erik Hoffmann, Dr. Eyad Fansa, Astrid Sander, Melina Terbeck, Deike Reinke, as well as the newer members of the facility for all the support and hard work. The facility expressed and purified all the

proteins used in the described here studies. I am not sure if my research could have taken place without the assistance that they provided.

Dr. Christoph Peter and Lea Esser from the Heinrich Heine University Düsseldorf are thanked for the help with the biological analysis of my peptides and their scientific contribution to the presented here work. Christiane Heitbrink and Eva Wieczorek are acknowledged for the collection of HRMS data. I would like to also thank Edelris, and especially Dr. Renaud Prudent and Dr. Anthony Willaume for their contributions to the PRMT5 projects. I also thank Refeyn, as well as FIDA Biosystems for giving me access to their instruments and their kind assistance with the sample analyses. I acknowledge the Paul Scherrer Institut, Villigen, Switzerland for provision of synchrotron radiation beamtime at the SLS and would like to thank Dr. Vincent Olieric and Dr. John Beale for their assistance.

Prof. Christian Hedberg, Prof. Michael Elofsson, Dr. Lindon Moodie and Dr. Michael Saleeb are also acknowledged here for providing me with an excellent scientific training, giving me a strong enough foundation to commence a doctoral degree in chemistry. The lessons they had taught me proved to be invaluable during the doctoral degree.

I thank Aventis Foundation and Stiftung Stipendien-Fonds of the Verbandes der Chemischen Industrie for providing a very generous doctoral scholarship.

I would also like to thank Dr. Brunschweiler for kindly assuming the role of the second examiner and the evaluation of my doctoral work.

Last, but definitely not least, I would like to thank the most important people in my life, my mum, my dad and my grandparents. Without them I would not have been able to reach this stage of education and finish my thesis work. I am forever grateful for their immeasurable love and sacrifices.

Table of Contents

Abstract	IX
Kurzzusammenfassung	XI
Introduction	1
1.1. Protein-Protein Interactions	1
1.1.1. General Characteristics of Protein-Protein Interactions.....	1
1.1.2. Hot Spots.....	5
1.1.3. PPI Pockets.....	6
1.1.4. Targeting PPIs and the Role of Small Molecules as PPI Modulators.....	7
1.2. Peptides and Peptidomimetics as PPI Inhibitors	12
1.2.1. Classification of Peptidomimetics.....	12
1.2.2. Structural Protein Elements.....	14
1.2.3. Peptide Limitations, Selected Modifications and Macrocyclisation Strategies	16
1.2.5. Fundamental Principles of Fmoc-based Solid Phase Peptide Synthesis	26
1.3. Methyltransferase PRMT5	29
1.3.1. Introduction to Protein Arginine Methyltransferases	29
1.3.2. PRMT5 in Disease.....	30
1.3.3. PRMT5 Structure and Selected Interactions.....	32
Aim and Objectives of the Thesis	36
PART A: Modulating Interactions Between PRMT5 and MEP50	37
3.1. Brief Introduction to PART A.....	37
3.2. Results and Discussion	39
3.2.1. Exploring the PRMT5-MEP50 PPI Interface	39
3.2.2. MEP50 Insertion Finger Mimetics.....	43
3.2.3. TIM Barrel Loop Mimetics.....	46
3.2.4. Protein Expression	49
3.2.5. Compound Evaluation.....	51
3.2.6. Histone Tail Peptides.....	55
3.3. Summary and Conclusions.....	57
PART B: Targeting Potential Interactions Between SUZ12 and MEP50	59
4.1. Brief Introduction to PART B.....	59
4.2. Results and Discussion	60
4.2.1. Identification of the MEP50 binding sequence in SUZ12.....	60
4.3. Summary and Conclusions.....	66

PART C: Investigating a Novel PPI Interface Between PRMT5 and its Adaptor Proteins	67
5.1. Brief Introduction to PART C.....	67
5.2. Results and Discussion	69
5.2.1. Identification and Biophysical Analysis of the Consensus Sequence.....	69
5.2.2. Initial Attempted RioK1 Peptide Cyclisation	74
5.2.3. Crystallographic Elucidation of the Interaction Between PRMT5 and RioK1 Peptide.....	77
5.2.4. Analysis of the Interactions Between PRMT5 and Full-length pICln and RioK1	80
5.3. Summary and Conclusions.....	83
PART D: Targeting Interactions of PRMT5 with its Adaptor Proteins	84
6.1. Brief Introduction to PART D.....	84
6.2. Results and Discussion	86
6.2.1. Development of the Macrocyclic PAPIIs.....	86
6.2.2. Evaluation of the Second Generation Macrocyclic PAPIIs	91
6.2.3. Optimisation of PAPIIs	97
6.2.4. Evaluation of the Third Generation Macrocyclic PAPIIs	100
6.3. Summary and Conclusions.....	106
Experimental Section	107
7.1. General Methods	107
7.2. Linear Peptide Synthesis	108
7.3. Cyclic Peptide Synthesis.....	109
7.3.1. General Approach	109
7.3.2. Amide Bond Mediated Cyclisation in Solution.....	110
7.3.2. Amide Bond Mediated Cyclisation on Solid Support	110
7.3.3. RCM Mediated Cyclisation on Solid Support	112
7.4. Protein Expression and Purification.....	112
7.4.1. Protein Expression	112
7.4.2. Protein Purification	113
7.5. Mass Photometry.....	113
7.6. Protein Labelling with Alexa 488.....	113
7.7. Fluorescence Polarisation	113
7.7.1. Direct Binding Assay.....	113
7.7.2. Competitive Binding Assay.....	114
7.8. Flow Induced Dispersion Analysis.....	114
7.9. Stability Assay	114
7.10. MTase-Glo TM Activity Assay	115
7.11. Nanoscale Differential Scanning Fluorimetry	115

7.12. Thermal Shift Assay	115
7.13. Protein Crystallisation, Data Collection and Analysis	116
7.14. Cell Culture	116
7.15. SDMA Pattern Analysis	117
7.15.1. Cell Treatment with PAPIIs	117
7.15.2. Cell Lysis and Immunoblotting	117
7.15. Pulldown	117
7.16. GFP Immunopurification and Immunoblotting.....	118
7.16.1. Generation of S100 Extract	118
7.16.2. Inhibitor Treatment and Immunopurification	118
7.16.3. Immunoblotting	118
7.17. Computational Modelling	119
Bibliography	120
Appendix	141
8.1. Supplementary Experimental Data	141
8.1.1. Supplementary Tables	141
8.1.2. Supplementary Figures	151
8.1.3. Peptide Purity HPLC Chromatograms	166
8.2. Publication List	242
8.3. Eidesstattliche Versicherung (Affidavit).....	243

Abstract

Protein-protein interactions (PPIs) are crucial for mediating nearly every cellular process in living cells, allowing to form complex molecular machines and pathways governing the cells. Modulation of PPIs is regarded as a promising approach for influencing cellular behaviour with a great potential in drug discovery. Protein arginine methyltransferase 5 (PRMT5) is an important enzyme, methylating arginine residues of its substrates. PRMT5 is involved in a very large number of various cellular processes, and has strong ties to a plethora of disorders including cancer, neurodegenerative and cardiovascular disease, as well as diabetes. Due to the crucial role of PRMT5 and its still not fully understood cellular functions and regulatory mechanisms, the enzyme has been a subject of numerous biochemical and biological investigations, as well as a target for drug discovery. PRMT5 is also known to mediate a number of PPIs with diverse adaptor proteins, such as the obligate binding partner MEP50, and adaptor proteins RioK1, pICln and COPR5. The adaptor proteins can regulate the enzymatic activity and substrate selectivity by directing PRMT5 to the appropriate cellular components. Modulation of PRMT5 PPIs thus appears as an attractive alternative to the inhibition of the active methylation site, giving an opportunity for significantly more precise targeting of the PRMT5 functionality in cells.

This thesis describes pathways for development of peptidomimetics intended for the inhibition of various PPIs formed by PRMT5 or its adaptor protein MEP50. The focus was put first on the PPI between PRMT5 and MEP50 (PART A). A virtual alanine scan was performed for the interface residues of PRMT5 and MEP50, and together with the available crystal structures of the complex the results indicated possible sites for targeting with PPI inhibitors. Two native protein loops were identified as promising templates for construction of inhibitory peptide analogues, resulting in a set of linear and macrocyclic compounds. The peptide analogues were tested for binding to the PRMT5-MEP50 complex, as well as for the ability to inhibit the enzymatic activity of the methyltransferase. None of the synthesised compounds showed any activity in the employed assays.

In the second part, the thesis describes the analysis of a potential PPI between SUZ12 and MEP50 (PART B). SUZ12 is a component of important protein assemblies called polycomb repressive complexes responsible for the regulation of gene expression. A number of SUZ12 fragments spanning different protein regions were synthesised and tested for the direct interaction with the PRMT5-MEP50 complex. The initial analysis indicated that the short SUZ12 sequence CPWCTL, found within the SUZ12 zinc finger, is responsible for the interaction with the methyltransferase complex. Further analysis employing truncated PRMT5 protein constructs revealed that the identified motif can strongly interact with the TIM barrel domain of PRMT5 in the absence of MEP50, which was in disagreement

with previously published findings. Experiment testing for the interaction between sequence CPWCTL and control proteins showed very strong non-specific binding. In light of the obtained results any further peptidomimetic development based on the SUZ12 sequence was terminated.

The final two parts of the thesis (PART C and D) focus on the PPIs between PRMT5 and adaptor proteins RioK1, pICln and COPR5. A consensus sequence GQF[D/E]DA[D/E], identified in the terminal regions of the adaptor proteins, was determined to be responsible for mediating the interactions with PRMT5. Using different PRMT5 protein truncations it was concluded that the consensus sequence binds to the non-catalytic TIM barrel domain of PRMT5. Alanine scans performed on a RioK1-derived sequence allowed to determine the contribution of the peptide side chains to the protein binding. A co-crystal structure of the TIM barrel bound to a RioK1 peptide was successfully obtained, revealing the structural characteristics of the identified interface. The biophysical data were then used for the design of highly stable macrocyclic PRMT5 adaptor protein interaction inhibitors (PAPIIs). Various macrocycle sizes, cyclisation linkages and amino acid sequence modifications were tested, resulting in a potent cyclic PRMT5 binder ($K_D = 89$ nM). The developed molecule could inhibit the PPI between PRMT5 and full length pICln when tested *in vitro* ($IC_{50} = 654$ nM). Examination of the compound activity in cellular lysates revealed a surprising selectivity towards the inhibition of the PRMT5 interactions with RioK1 over the interactions with pICln. The developed PAPIIs did not suppress the enzymatic activity of the methyltransferase and did not inhibit the interaction with MEP50.

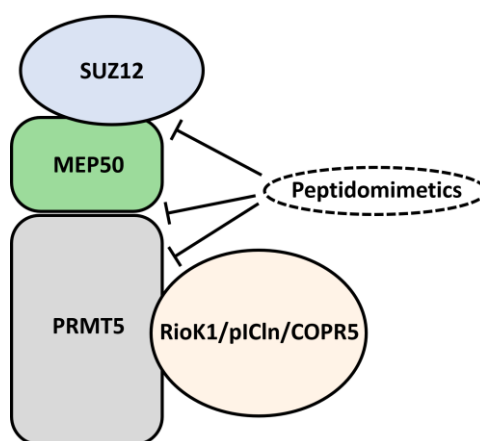


Figure 1. Schematic representation of the investigated PRMT5 PPIs. The aim of the thesis was to develop peptidomimetics capable of inhibiting the depicted interactions.

Kurzzusammenfassung

Protein-Protein-Wechselwirkungen (PPI) sind entscheidend für die Vermittlung fast aller zellulärer Prozesse in lebenden Zellen und ermöglichen die Bildung komplexer molekularer Maschinen und Wege, die die Zellen steuern. Die Modulation von PPIs wird als vielversprechender Ansatz zur Beeinflussung des Zellverhaltens angesehen, der ein großes Potenzial für die Arzneimittelforschung birgt. Die Protein-Arginin-Methyltransferase 5 (PRMT5) ist ein wichtiges Enzym, das die Argininreste seiner Substrate methyliert. PRMT5 ist an sehr vielen verschiedenen zellulären Prozessen beteiligt und steht in engem Zusammenhang mit einer Vielzahl von Krankheiten wie Krebs, neurodegenerativen und kardiovaskulären Erkrankungen sowie Diabetes. Da PRMT5 eine entscheidende Rolle spielt, seine zellulären Funktionen und Regulierungsmechanismen aber noch nicht vollständig geklärt sind, ist das Enzym Gegenstand zahlreicher biochemischer und biologischer Untersuchungen und ein Ziel für die Arzneimittelentwicklung. Es ist auch bekannt, dass PRMT5 eine Reihe von PPIs mit verschiedenen Adaptorproteinen bildet, wie dem obligaten Bindungspartner MEP50 und den Adaptorproteinen RioK1, pICln und COPR5. Die Adaptorproteine können die Enzymaktivität und die Substratspezifität regulieren, indem sie PRMT5 zu den entsprechenden zellulären Komponenten leiten. Die Modulation der PRMT5-PPIs scheint somit eine attraktive Alternative zur Hemmung der aktiven Methylierung zu sein und bietet die Möglichkeit, die PRMT5-Funktionalität in den Zellen wesentlich präziser zu steuern.

Die vorliegende Arbeit beschreibt Wege für eine Entwicklung von Peptidomimetika, für die Hemmung verschiedener PPIs, von PRMT5 oder seinem Adaptorprotein MEP50. Der Fokus wurde zunächst auf die PPI zwischen PRMT5 und MEP50 gelegt (TEIL A). Ein virtueller Alanin-Scan wurde an den Schnittstellenresten von PRMT5 und MEP50 durchgeführt, und zusammen mit den verfügbaren Kristallstrukturen des Komplexes zeigten die Ergebnisse mögliche Angriffspunkte für PPI-Inhibitoren auf. Zwei native Proteinschleifen wurden als vielversprechende Vorlagen für die Konstruktion von hemmenden Peptidomimetika identifiziert, was zu einer Reihe von linearen und makrozyklischen Verbindungen führte. Die Peptidomimetika wurden auf ihre Bindung an den PRMT5-MEP50-Komplex sowie auf ihre Fähigkeit zur Hemmung der enzymatischen Aktivität der Methyltransferase getestet. Keine der synthetisierten Verbindungen zeigte in den verwendeten Assays Aktivität.

Im zweiten Teil der Arbeit (TEIL B) wurde eine mögliche PPI zwischen SUZ12 und MEP50 beschrieben. SUZ12 ist eine Komponente an wichtigen polycomb-repressiven Komplexen, die für die Regulierung der Genexpression verantwortlich sind. Eine Reihe von SUZ12-Fragmenten, die verschiedene Proteinregionen abdecken, wurde synthetisiert und auf eine direkte Interaktion mit dem PRMT5-MEP50-Komplex getestet. Die erste Analyse ergab, dass die kurze SUZ12-Sequenz CPWCTL, ein Motiv innerhalb des SUZ12-Zinkfingers, für die Interaktion mit dem Methyltransferase-Komplex verantwortlich ist. Weitere Analysen mit verkürzten PRMT5-

Proteinkonstrukten ergaben, dass das identifizierte Motiv in Abwesenheit von MEP50 stark mit der TIM-Barrel-Domäne von PRMT5 interagieren kann, was im Widerspruch zu früher veröffentlichten Ergebnissen stand. Experimente zur Untersuchung der Interaktion zwischen der Sequenz CPWCTL und Kontrollproteinen zeigten eine sehr starke unspezifische Bindung. Angesichts der erzielten Ergebnisse wurde eine weitere peptidomimetische Entwicklung auf der Grundlage der SUZ12-Sequenz eingestellt.

TEIL C und D der Arbeit konzentrieren sich auf die PPIs zwischen PRMT5 und den Adaptorproteinen RioK1, pICln und COPR5. Eine Konsensussequenz GQF[D/E]DA[D/E], in den terminalen Regionen der Adaptorproteine, wurde als verantwortlich für die Vermittlung der Interaktionen mit PRMT5 ermittelt. Unter Verwendung verschiedener PRMT5-Proteinsequenzen wurde festgestellt, dass die Konsensussequenz an die nichtkatalytische TIM-Fassdomäne von PRMT5 bindet. Anhand von Alanin-Scans, die an einer von RioK1 abgeleiteten Sequenz durchgeführt wurden, konnte der Beitrag der Peptidseitenketten zur Proteinbindung bestimmt werden. Es wurde erfolgreich eine Co-Kristallstruktur des an ein RioK1-Peptid gebundenen TIM-Fasses erhalten, die die strukturellen Merkmale der identifizierten Schnittstelle aufzeigt. Die biophysikalischen Daten wurden dann für das Design von hochstabilen makrozyklischen PRMT5-Adaptorprotein-Interaktionsinhibitoren (PAPIIs) verwendet. Es wurden verschiedene Makrozyklusgrößen, Zyklisierungsbindungen und Aminosäuresequenzmodifikationen getestet, was zu einem potenten zyklischen PRMT5-Binder ($K_D = 89 \text{ nM}$) führte. Das entwickelte Molekül konnte im in-vitro-Test die PPI zwischen PRMT5 und pICln in voller Länge hemmen ($IC_{50} = 654 \text{ nM}$). Die Untersuchung der Aktivität der Verbindung in Zelllysaten ergab eine überraschende Selektivität in Bezug auf die Hemmung der PRMT5-Interaktionen mit RioK1 gegenüber der Interaktion mit pICln. Die entwickelten PAPIIs unterdrückten weder die enzymatische Aktivität der Methyltransferase noch hemmten sie die Interaktion mit MEP50.

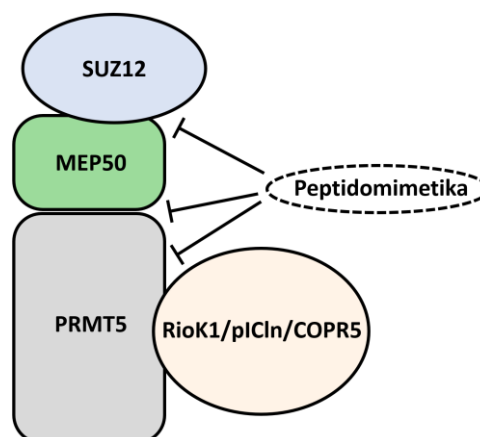


Abbildung 1. Schematische Darstellung der in der Arbeit untersuchten PRMT5-PPIs. Ziel der Arbeiten war es, Peptidomimetika zu entwickeln, die in der Lage sind, die dargestellten Wechselwirkungen zu hemmen.

Introduction

1.1. Protein-Protein Interactions

1.1.1. General Characteristics of Protein-Protein Interactions

Proteins rarely act alone and interactions between proteins, or so called protein-protein interactions (PPIs), are fundamental and crucial for the correct functioning of cells as they are involved in nearly every cellular process. PPIs allow for the formation of complex protein machineries, occurrence of intricate cellular pathways and their regulation, and thus, they facilitate the existence of a very advanced cellular interactome.^[1–6] PPIs are responsible for the assembly of such notable multiprotein constructs as the nuclear pore complex,^[7–10] spliceosome machinery,^[11,12] cellular motors including ATP synthase,^[13–16] chaperones,^[17,18] cytoskeletal structures,^[19,20] or chromatin remodelling complexes,^[21,22] just to name a very few examples. Aberrant PPIs are a well-established cause of many diseases, due to either an over-stabilisation of interactions or a weakening of the crucial associations between proteins.^[23] Modulation of PPIs therefore offers remarkable opportunity for influencing cellular development and behaviour, and from a chemical biology perspective it seems to be pivotal to have a comprehensive understanding of these interactions.

Due to the prevalence of PPIs between a wide range of diverse proteins, many different types of PPIs can be distinguished with regards to the composition of the involved units, orientation, affinity and the stability of the interactions (Figure 2). The first distinction can be made based on the identity of the interacting proteins. The interactions can occur between identical protein chains, affording homooligomers, or can form between different chains giving heterooligomers.^[24–26] The four most common multimeric stoichiometries of biological assemblies as determined based on the analysis of protein data bank (PDB) entries are homodimers (23.6% of all entries), homotetramers (6.3%), heterodimers (6.3%), and homotrimers (2.8%), with a total of 14.2% of all PDB entries being heteromeric.^[27] These statistics may not, however, necessarily represent the commonality of such interactions in living cells, but rather reflect the arrangements of structures conducive to crystallisation.

Interfaces between proteins are further classified as isologous for homodimeric and symmetric interactions where the same surface on both proteins is involved, or as heterologous for asymmetric interactions for both homodimers and heterodimers. Alternatively, interfaces can be categorised as pseudoisologous for cases where they are approximately symmetrical in heterodimers with homologous domains.^[25,27]

Another classification of interactions is made based on the 'obligation' of proteins to form interactions with other units in complexes. If a subunit of a complex is not stable on its own *in vivo*, the PPI is categorised as obligate. The obligate complexes are normally formed right at the time of the protein biogenesis, and thus are also referred to as folding complexes. On the other hand, if the protomers of a complex can exist independently of each other, the PPI between them is considered to be non-obligate, for example in case of antibody and antigen interactions or signalling protein complexes. The non-obligate complexes are also known as recognition complexes.^[24–26,28]

PPIs can also be classified as permanent or transient depending on the life-time and stability of the interactions. Permanent interactions are normally irreversible and occur in very stable complexes. Transient interactions are found in complexes which protomers can freely and transiently associate and dissociate, allowing a particular protein to change its binding partner and be involved in many different complexes with different cellular functions and purposes.^[25,26,29] The transient interaction can be further broadly categorised as weak, with the dissociation constant (K_D) typically in the micromolar range, or as strong, with the K_D values generally in the nanomolar range and below.^[26,30] The transient interactions are also normally associated with the signalling and the regulation of various cellular pathways, often related to diseases, and are thus of particular medicinal interest. Due to the nature of the obligate interactions they are typically permanent. Non-obligate complexes are characterised predominantly by the transient interactions but can also be permanent.^[25,26,29]

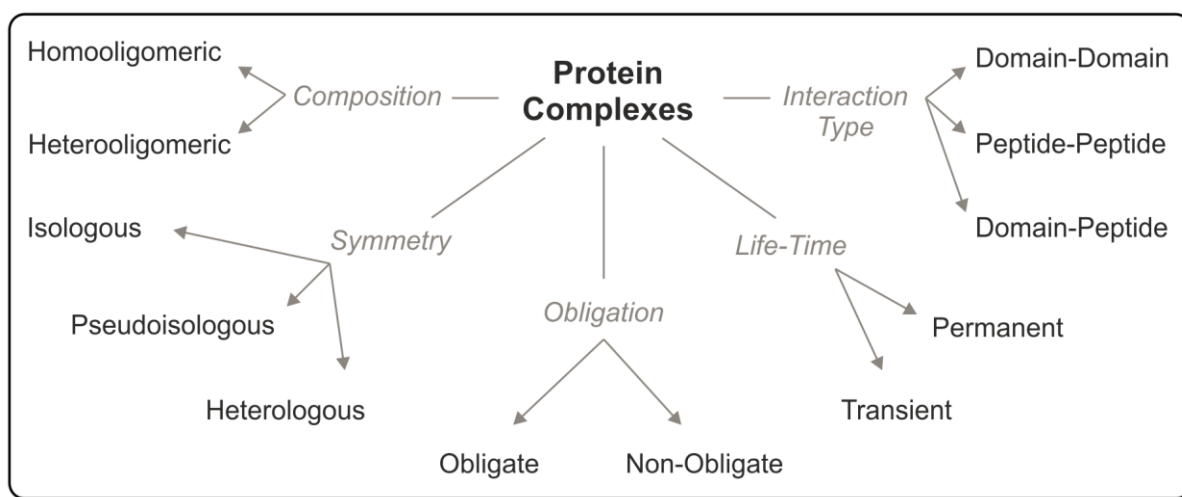


Figure 2. Schematic representation of different PPI classifications applied to protein complexes.

It is worth noting that one PPI can fall into different categories depending on the current environment, surrounding physiological conditions and the particular state and modifications of the involved proteins. Interactions can be obligate and permanent under certain set of conditions but can change to non-obligate and transient in different milieu. PPIs can often be controlled through changing local concentrations of complex subunits and by influencing the affinity of protomers to the other units in

a complex, for example through modifications to the physicochemical environment of the proteins or through post translational modifications (PTMs) and allosteric changes to the conformations.^[25,29,31–35]

Structural studies of complexes found additionally a few general characteristics for different types of PPIs that are worth mentioning here. The binding interfaces found among transient non-obligate complexes have a considerable number of polar and charged residues, similarly to the exterior of proteins, whereas the interfaces in the permanent and obligate complexes are more hydrophobic resembling the interior of a globular protein.^[25,26,28,36–38] The salt bridges and hydrogen bonds connecting protomers together are thus more common for the transient complexes.^[36] The surface area involved in PPIs is typically below 1500 Å² for the transient interactions and in the range of 1500–10000 Å² for the permanent complexes.^[25,26,28,39,40] The explanation for these polarity and size trends seems to be the necessity for the transient and non-obligate protomers to be water soluble without aggregation and the ability to fold independently when existing in isolation without the binding partners.^[25,26,37,41] Considering the structural interface differences, non-obligate complexes appear to have more irregular secondary structures that impart flexibility and can facilitate protomer dissociation, whereas typically only the obligate complexes have β-turns present at their interfaces.^[26,38] It was also observed that proteins with PPI interfaces characterised by large surface areas (> ca. 1000 Å²) often have a tendency for going through conformational changes when forming a complex.^[25,36,39] Moreover, the binding free energy ΔG between most of the complex protomers is uncorrelated to the polarity or size of the formed PPI interfaces, however, a weak correlation was shown for cases of complexes with co-expressed and co-localised protein subunits.^[25,39,42]

PPI occurrences can be mediated through domain-domain, peptide-peptide (e.g. coiled-coil) or domain-peptide interactions (Figure 3).^[26,43,44] The domain-peptide interactions are conservatively estimated to account for up to 40% of all PPIs in cells.^[45] The domain-peptide interactions are also called transient peptide-mediated interactions as they are mostly of transient nature, occurring often between a globular domain and a relatively short linear sequence. The interaction domains typically recognise a very specific binding motif of the peptide (consensus motif usually of three to ten residues), normally found at one of the termini of the partner protein or within one of its loops or disordered regions, and while bound to the domains the peptidic motifs generally assume a well-defined conformation.^[26,46,47] Conventionally, the peptidic consensus motif is energetically favourable and accounts for the most of the binding free energy in the domain-peptide PPIs. The consensus motif can be surrounded by a varying number of contextual residue interactions, which allows to subtly regulate the strength and specificity of the PPI. In some cases, however, these contextual interactions are unfavourable and can weaken the interactions with a binding protein, leading to more transient

states.^[46,48] The surface area of domain-peptide interfaces is typically between 200 and 500 Å², which as expected is considerably smaller than the average area of domain-domain interfaces.^[26,49] It also appears that majority of peptides binding the interaction domains do so in the coiled conformation, with only a small fraction interacting through either an α -helix or a β -sheet. Peptides have a significantly higher density of H-bonds at the interface with the binding protein partners than in the case of the domain-domain interactions. The H-bonds between peptides and the binding domains also involve considerably more main chain atoms. The number of H-bonds formed is dependent on the peptide conformation, with the α -helices making on average fewer H-bonds than β -sheets. The interacting α -helical peptides are normally of amphiphilic nature, binding the protein domains with the hydrophobic surface. Another notable feature of the domain-peptide interactions is that the protein domains interacting with the peptidic sequences typically do not undergo any meaningful conformational changes. Peptides, on the other hand, can be in a highly flexible state when unbound, and thus need to pay a considerable entropic penalty upon the interaction with the protein domains and folding into a specific conformation.^[49] This sizable entropic penalty appears to explain the need for the high observed H-bond density at the domain-peptide interfaces and the lack of further conformational changes to the interacting protein domains that could presumably otherwise induce further energetic costs. In some cases, the binding peptide motif in a protein can be already folded prior to the contact with the other protomer, and this is caused by the stabilising effect of the remainder of the protein structure acting as an internal scaffold.^[44]

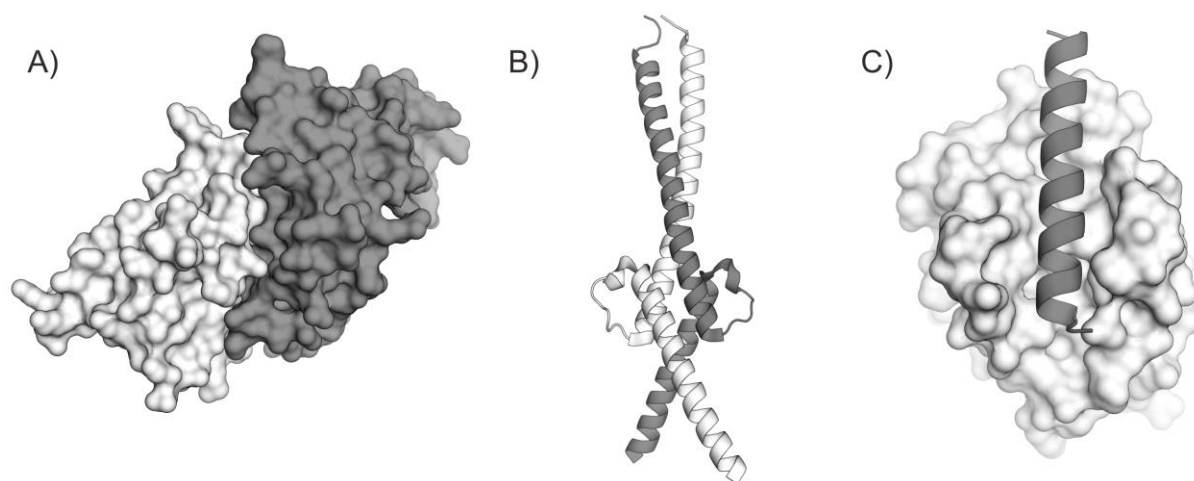


Figure 3. Exemplary crystal structures representing PPIs mediated through different interaction types: **A)** Domain-domain interaction (PBD ID: 1Z92).^[50] **B)** Coiled-coil peptide-peptide interaction (PBD ID: 1NKP).^[51] **C)** Domain-peptide interaction (PBD ID: 2XA0).^[52]

Specific parameters of PPIs appear to be finely tuned for the corresponding biological function that given protein complexes have, as well as to the underlying regulatory mechanisms governing the cellular behaviours. It is believed that PPIs evolved optimising the efficacy of the relevant cellular

functions, resulting in the most appropriate protomer affinities and complex stabilities for the required tasks. The evolutionary optimisation of the PPIs is not tantamount to the strength increase in the protomer affinities, as both strong and weak associations of complexes have their specific purposes in cellular processes. The evolution and existence of obligate complexes could be caused by the instability of subunits in isolation as described before or may be in some cases the absolute necessity for the cooperation between proteins, that otherwise would have no other functions, bestowing on them a particular activity. It is also possible that the existence of certain complexes is accidental and meaningless to the function of their individual subunits.^[25] PPI interface residues in proteins are generally more conserved than in the remainder of the protein surface.^[53-55] More granularly, the interface residues in the permanent complexes are also more conserved than the residues involved in the transient PPIs.^[53,56] On the interaction surfaces there are typically only a handful of heavily conserved residues with a significant energetical contribution to the protein binding. Such key residues are referred to as hot residues or hot spots.^[53,57,58]

1.1.2. Hot Spots

Hot spots are the residues that upon mutation to Ala cause a significant decrease in the binding free energy ($\Delta\Delta G$), when measured in interactions with a binding partner. Typically, the required $\Delta\Delta G$ needed to classify a residue as a hot spot is assumed to be ≥ 2.0 kcal/mol, whereas the requirement for a classification as a red-hot residue is $\Delta\Delta G$ of 4.0 kcal/mol or higher. Red-hot residues are the strongest contributors to the interactions, however they are very unusual and rarely found.^[59,60]

The hot spot composition is not random, and the most commonly found amino acids are Trp (21%), Arg (13%) and Tyr (12%). Trp is a large amino acid and its mutation to Ala can form a considerable cavity on the protein surface that results in a complex destabilisation. Additionally, it can strongly interact with other residues through π -interactions and H-bonds, and due to the large hydrophobic surface it can shield weak H-bonds from water. Arg can form a strong salt bridge with negatively charged residues, and contribute up to five H-bonds through the guanidinium containing side chain. Tyr is three times more likely to be found as a hot spot than Phe, presumably due to its H-bond forming ability, in addition to the capacity of establishing π -interactions. The amino acids that are the least likely to be hot spots are Leu, Ser, Thr and Val.^[59,61]

Hot residues are not distributed in a homologous fashion over the interface area, but instead, they tend to form densely packed clusters, also known as hot regions.^[58-60] According to an already well-established and accepted O-ring theory, a hot region is normally surrounded by other, unimportant for the energy of binding residues (O-ring), shielding the hot region from detrimental interactions with

water molecules that could disrupt the H-bonds and the charged interactions.^[60–62] The O-ring normally has hydrophobic properties. Thus, the residues of hot regions typically have minimal or no solvent accessible surface area (SASA).^[61]

The number of hot spots present at the interface of protein-protein or protein-peptide interaction is often proportional to the size of the interface.^[49,60] In case of peptides, a hot spot normally is needed for every third residue, in order to assure appropriate binding energy of the interaction necessary to counterbalance the considerable entropic costs sustained during the binding event with a protein.^[49]

As the experimental, mutational determination of hot spots is a challenging and very time-consuming task, many computational methods, based on both traditional programming and machine learning solutions, have been developed for the detection of interface hot residues in bound and unbound proteins. The *in silico* approach can be complementary to the classical mutational studies and can help to very rapidly and with satisfactory accuracy elucidate many functional and mechanistic considerations regarding the PPIs.^[63,64]

1.1.3. PPI Pockets

Pockets are important geometrical features populating the interfaces of PPIs (Figure 4). Unfilled pockets can be created after an interaction between proteins is established. In such case the surface of one or both binding partners is concave and left vacant. The unfilled pockets constitute approximately 5 to 20% of the interface area, where the area of the unfilled pockets increases with the increase of the PPI interface surface. It is assumed that these pockets may be necessary for biological functions by allowing dynamic movements of the housed residues, as well as for the required binding interface flexibility. The unfilled pockets also appear to lower the PPI stability.^[60] Complemented pockets are another motif found at the PPI interfaces, although they occur less frequently than the unfilled pockets. Complemented pockets are concave regions of non-trivial shape on a protein that are tightly filled by protruding chains of the binding partners, and their size is poorly correlated with the PPI surface area. The complemented pockets and the corresponding protrusions are rich in hot residues, where approximately 60% of all hot residues and over 90% of the red-hot residues are located in these surface features. The complementary pockets are thus instrumental for the formation of strong PPIs, as they have the most significant contribution to the $\Delta\Delta G$ of binding, and therefore, they are potentially a very interesting target for PPI modulation.^[60] Curiously, the process of the protomer association into a complex rarely imparts conformational changes that lead to the considerable rearrangement of the pre-existing pocket organisation.^[60]

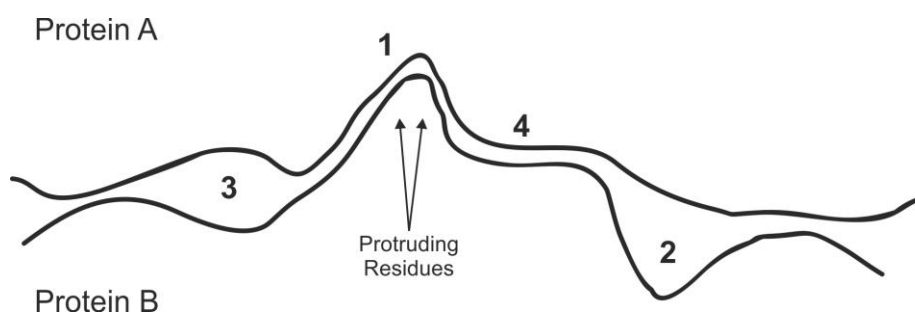


Figure 4. Schematic representation of the PPI interface pockets: complementary pocket (1), unfilled pocket created by a concave region on protein B (2), unfilled pocket created by shallow depression of the both interacting proteins (3) and matched, flat interface region without non-trivial geometry (4). The figure is based on the image from Li *et al.*, 2004.^[60]

1.1.4. Targeting PPIs and the Role of Small Molecules as PPI Modulators

Dysregulation and aberrant dynamics of PPI networks in cells is characteristic for many diseases, and thus, modulation of selected interactions between specific proteins appears to be of particular interest. Modulation of PPIs can lead to the development of novel therapeutics as an alternative to targeting active enzyme sites, receptors and channels.^[44,65–70] Compounds that can inhibit or alternatively stabilise PPIs have great pharmaceutical and research potential as prospective drugs and molecular probes. Such modulators could induce system changes through affecting a broader web of protein interactions. Moreover, PPI modulators have a potential to cause modifications to signalling or regulatory networks without completely shutting them down.^[66,71] And although inhibition of PPIs typically provides better target control than a complete enzyme or receptor shutdown, a protein can use the same binding interface to interact with a number of different binding partners, meaning that there is also a potential for a PPI inhibitor to have unspecific and inadvertent consequences caused through the disruption of multiple interactions, leading to potentially undesired biological effects.^[30,72,73]

Until recently, PPIs were considered undruggable due to the ineffectiveness of the classical strategies normally used for the development of small molecules targeting enzymes and other proteins with deep and distinctive cavities. Following advancements in structural and computational biology, as well as the development of a better understanding of PPIs and adjustments to the screening methods, PPIs became more tractable, however, they still remain a very challenging and daunting target.^[44,66–70]

One of the major reasons why PPIs are so difficult to target is that PPI interfaces are typically much larger and flatter than the classical protein-ligand interaction interfaces, and are also characterised by diverse types of topologies.^[44,74] Additionally, PPI interfaces, contrary to many enzymes, receptors and channels, did not evolve ligand cavities to host small molecules, rendering a discovery of PPI

modulators more challenging. For enzymes and receptors, the structural information concerning the known ligand cavity can be used as a starting point for development of a novel binder. Identification of PPI cavities appropriate for targeting with small molecules is more difficult. Complexes suspended in solution are typically not completely rigid and can undergo conformational movements, often inducing an appearance of transient openings that can be targeted with ligands. The opening of such cavities can be possibly detected through molecular dynamics simulations.^[67,69,75]

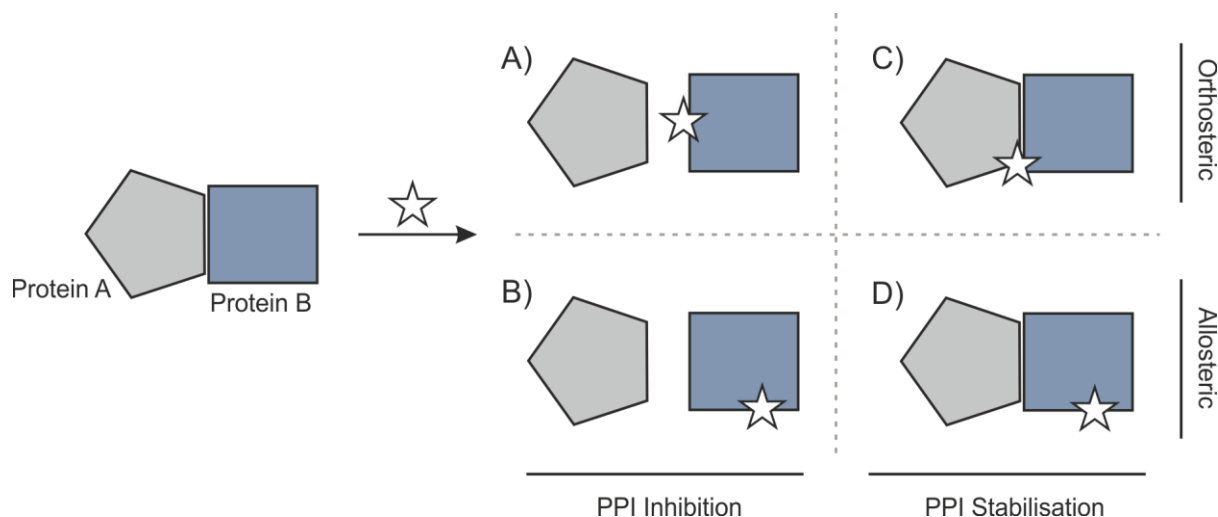


Figure 5. Schematic representation of different binding modes of PPI modulators: **A)** Orthosteric PPI inhibitor binding protein B. **B)** Allosteric PPI inhibitor binding protein B. **C)** Orthosteric PPI stabiliser binding to protein A and B. **D)** Allosteric PPI stabiliser binding to protein B.

The figure is based on the image from Fischer *et al.*, 2015.^[76]

PPIs can be inhibited, preventing the protein complex partners from binding to each other, or they can be stabilised (Figure 5). Both strategies can achieve different results on a biochemical and functional level, and change the cell behaviour accordingly.^[69,76,77] It appears, however, that the development of PPI stabilisers involves all the difficulties associated with the development of PPI inhibitors, but the challenge is exacerbated by a very limited understanding of the principles governing the PPI stabilisation, as well as a lack of the chemical consensus caused by the wide variation of the molecular structures of the identified stabilisers. Most PPI stabilisers have been found serendipitously, with only a handful being a result of a rational design.^[68,69,77]

From the perspective of PPI inhibition, the druggable PPI interfaces typically have a relatively small buried surface area, are more hydrophobic and do not undergo significant conformational changes.^[70] The interfaces can be divided into four classes: narrow and wide, as well as, tight and loose. The narrow PPI interfaces with a surface area below 2500 Å² and counterintuitively tight interactions with the K_D values below 200 nM appear to be more receptive to inhibition. A potential explanation of this observation could be that the tight and narrow PPI interfaces are the most similar to the enzymatic

active sites targeted by the classical inhibitors, and in consequence, the regular techniques of finding binders are the most relevant for this particular class of PPIs.^[30] In addition to the orthosteric mode of action the PPI inhibitors can also disrupt a complex formation through allosteric changes caused by binding to the site on one of the protomers away from the PPI interface (Figure 5).^[76] Among different PPI types targeted by small molecules, domain-peptide interfaces where the inhibitor binds to the globular domain have been identified to be the most tractable, presumably due to the interactions being dominated by a relatively compact, dominant and continuous binding epitope, fulfilling the requirements of the tight and narrow interface classification.^[49,70,78] Small molecule inhibitors of the domain-peptide interactions typically mimic the native interactions of the peptidic motif in the hot-spot region of the binding pocket. Domain-domain interactions on the other hand are less amenable for targeting with small molecules due to the relatively large and flat interfaces. Peptide-peptide interactions are also very challenging from the perspective of the binder design and discovery as they normally lack any well-defined binding sites and there is a tendency for the interacting peptides to be in a disordered state when the complex subunits are unbound. The known inhibitors of peptide-peptide interactions have weak potency and the mechanism of their interaction with the targets is generally unknown.^[44]

Small molecules capable of targeting PPI interfaces normally do not have the classical drug-like properties. This observed discrepancy could be justified by considering that the classical drug-like characteristics are a result of small molecules often having to compete with endogenous small molecules for binding to their protein targets with a relatively small pocket surface area, typically below 1000 Å².^[74,79] Developing an effective PPI inhibitor may require to target a number of hotspots and hot regions located on an often flat and extended surface, which makes it difficult for the active molecule to be confined within the conventional drug discovery parameters.^[44,68,80] Thus, identified small PPI inhibitory molecules are normally larger and more lipophilic than classical drug-like structures, and they are also more rigid and form fewer H-bond interactions. PPI inhibitors typically also have lower ligand efficiency scores.^[79]

The classical structure-based design and high throughput screening (HTS) of small molecules are somewhat effective for the tight and narrow interfaces, however, in general HTS screenings against PPIs have been shown to give low hit rates and typically weakly potent binders.^[30,44] The low hit rate could be caused by a general difficulty to bind the targeted protein surface, but could also be a result of a library bias towards molecules aimed at targeting conventional ligand pockets.^[44] Fragment-based drug discovery (FBDD) has shown relatively satisfactory effectiveness against PPIs, as well as a positive correlation between the number of fragment binding sites and hot spot regions on a target protein. An advantage of the FBDD method over the conventional screening approach is that molecular

fragments are unlikely to have a bias towards any particular class of targets and interactions.^[44,81] On the other hand, fragment hits are typically of low affinity and may not be able to disrupt the target PPI, and thus alternative and relatively low throughput biophysical techniques are used for screening, such as the thermal shift assay (TSA), surface plasmon resonance (SPR), isothermal titration calorimetry (ITC), as well as ligand- and protein-based nuclear magnetic resonance (NMR) to name a few.^[44] To account for the low affinity a modification of the FBDD method can be used, where the fragments are covalently attached to the protein of interest using a reversible covalent connection or so called tether. The tethers are normally formed between the fragments and either Cys or Lys of the protein chain. If needed the required amino acid can be engineered onto the surface of the targeted protein. As the tether can be easily broken, the fragments with no intrinsic affinity to the target are not retained on the surface of the protein, but the weakly binding fragments are stabilised with the equilibrium favouring the protein modification that is detectable through an assessment with mass spectroscopy. The tethered fragments are also amenable to a further analysis with X-ray crystallography, allowing to elucidate the exact interactions with the protein and design an optimised binder. The tethering method has been successfully applied to the development of both PPI inhibitors and stabilisers.^[82–86] Despite some successes with the FBDD approach targeting PPIs, linking the identified binding fragments together into one molecule and optimising the potency is typically a very challenging task.^[87] Covalent small molecule PPI inhibitors are also known, where the covalent linkage results in high affinity compounds with a prolonged time of target modulation. As with the tethered FBDD strategy, one significant limitation of the covalent binders is that they require a presence of an appropriate reactive amino acid in the vicinity of the binding site on the protein surface.^[88–90] Application of covalent PPI inhibitors may also in some cases lead to undesired effects, such as high toxicity, caused by the irreversible binding and the produced active metabolites.^[90,91] The lead optimisation of any PPI inhibitor hits from screening campaigns can be particularly challenging especially when dealing with flat and featureless surfaces.^[44]

Studies in structural biology are of considerable importance for a development of PPI inhibitors, as such development is often supported by an analysis of protein interfaces with X-ray crystallography or protein NMR studies. The obtained structural information about the protein surface can then be used for *in silico* virtual screening against digital molecular libraries and molecular fragments.^[92–94] During screening for PPI modulators, however, significant challenges can arise in terms of the accuracy of the analysis, especially in cases when considerable conformational changes of protein surfaces occur upon forming a complex.^[44]

Despite the advances in the understanding of the PPIs and the associated limitations concerning the screening and development process of small molecule PPI modulators, the discovery of novel

modulators with the traditional small molecule drug discovery techniques still remains a truly significant challenge. An alternative and more intuitive approach to the development of the PPI modulators is a structure-based design strategy, where compounds are created based on the structures of the pre-existing binding motifs present at the targeted interface.^[95] The structure-based design and the resulting peptidomimetic modulators of PPIs are elaborated in the following section 1.2.

1.2. Peptides and Peptidomimetics as PPI Inhibitors

1.2.1. Classification of Peptidomimetics

As the traditional screening of physical and virtual small molecule libraries in search for novel PPI modulators remains a very difficult and daunting task, designing inhibitors based on an already existing structural protein motifs appears as an appealing probe and drug discovery approach. The compounds based on the existing protein motifs have typically higher molecular weight than the classical small ligands, and can normally establish a broad network of non-covalent interactions with the target, which can be beneficial for binding to relatively flat and featureless PPI surfaces.^[44,68,95–97] The presence of well-defined secondary structures at the interfaces such as distinct α -helices, β -sheets and loops is very important from the perspective of the inhibitor design, as it allows to synthesise compounds that mimic and compete with the native structures at the PPI interfaces.^[44,68,95] One limitation of utilising the native binding PPI motifs for inhibition, however, is that when extracted from the full protein sequence, short peptides tend to lose their secondary structures. This poses as a challenge, as conformational freedom is a cause for the peptide proclivity towards proteases, contributes to the low bioavailability and is also a reason for typically low binding affinity. The low affinity of flexible peptides to their targets can be explained by the entropic penalty paid upon binding to the partner protein, when the peptide is required to assume a specific conformation.^[49,95]

To counterbalance the negative effects of excising peptides from the full proteins, several methods have been devised which allow to modify peptides through cyclisation or chain modifications, or through constructing artificial structures that mimic the interactions with protein based on the native relationship between the peptide and the protein. Such modified structures are referred to as peptidomimetics. Conventionally, the peptidomimetics are classified into three types. Type I peptidomimetics are short peptides that assume the native secondary structure through an artificial stabilisation of the conformation aided by incorporating minor modifications into the peptide sequence. Type I mimetics tend to match the original peptidic backbone atom-to-atom and conserve the important interactions with the binding site. Type II are functional mimetics which in some cases can be direct structural analogues of the peptides but this is not a required criterium for the classification, as these mimetics can also bind to subsites that are different from the interaction sites used by the native peptides. Type II mimetics are small nonpeptide molecules. Type III are topographical mimetics that are based on novel nonpeptidic templates unrelated to the native peptide.^[98]

As the traditional peptidomimetic classification does not encompass the modern advances in the field, a new and improved classification was introduced in 2016 by Grossmann and colleagues, dividing the

peptidomimetics into four classes A through D.^[95] Classes A and B are allocated for molecules with strongly peptidic character, whereas classes C and D include small molecules (Figure 6). Class A mimetics are peptides that are based on the native peptide sequence with only minor modifications introduced to some of the amino acids in order to stabilise the active conformation. In class A mimetics, the positions and orientations of the main chain and the side chains are in close alignment with the structures of the original binding sequence. Class B mimetics are class A mimetics with further structural modifications, including incorporation of non-proteinogenic amino acids and small molecule fragments, as well as changes to the backbone. Class B also includes foldamers and peptoids, where the topology of the side chains resembles the orientation of the side chains in the native peptide. Class C mimetics are compounds with strong small molecule character, where the backbone of the original peptide is completely replaced by a new scaffold. The C class mimetics orient their substituents in a manner that allows for a formation of the key interactions originally existing between the precursor peptide and the target protein. Lastly, class D of mimetics are small molecules that have no direct structural or functional connection to the side chains of the original peptide, but they are able to mimic the mode of action of the native peptide. Such mimetics may be developed as a result of an optimisation of the class C mimetics, or they can be found in small molecule screening campaigns.^[95]

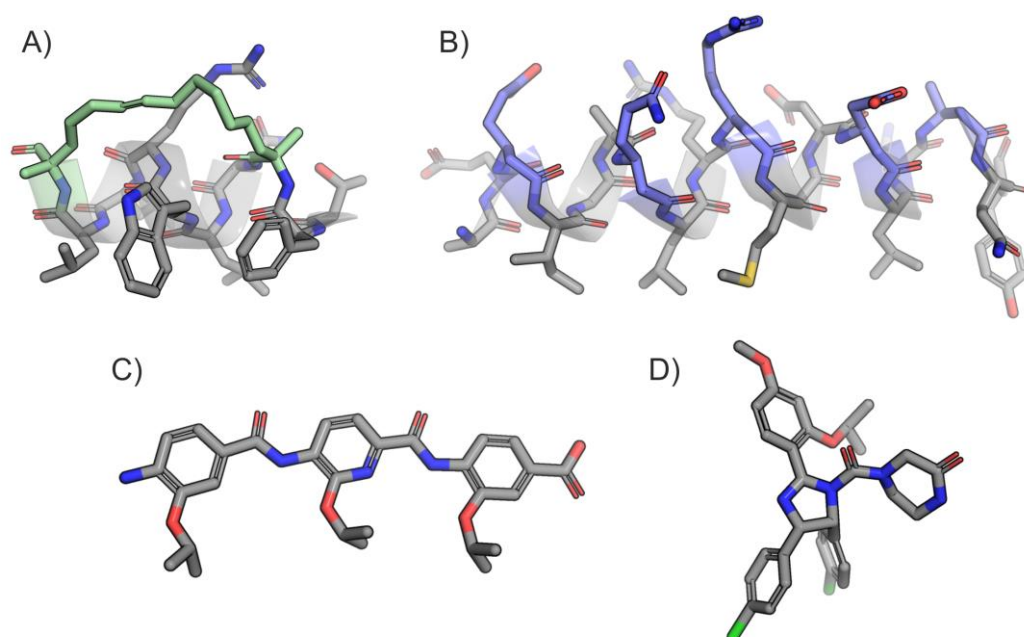


Figure 6. Exemplary peptidomimetics of classes A-D: **A)** Representant of the class A peptidomimetics. Stabilised α -helical peptide binding to MDM2 protein. The artificial α -helix stabilising motif is coloured as mint green (PDB ID: 3V3B).^[99] **B)** Representant of the class B peptidomimetics. Helical binder of MCL-1 protein, where some of the α -amino acids are replaced by β -amino acids containing an addition backbone methylene group. The β -amino acids are coloured as purple (PDB ID: 4BPI).^[100] **C)** Representant of the class C peptidomimetics. Structural mimetic of a helical MCL-1 binder.^[101] **D)** Representant of the class D peptidomimetics. Small molecule binding MDM2, inhibiting the p53-MDM2 PPI (PDB ID: 4HG7).^[102]

1.2.2. Structural Protein Elements

Structural protein elements such as turns, helices and sheets, are the key building blocks used for the construction of peptidomimetic PPI inhibitors, and thus, a short introduction of those elements is in place.

Turns are irregular secondary protein structure elements with up to six amino acid residues in length, that allow the peptidic chain to fold back on itself, and therefore, they are responsible for imparting the globular character onto proteins.^[103,104] The turns can be classified into different families based on their specific intraturn hydrogen bond pattern between the backbone carbonyl group and the backbone hydrogen of the amide group (Figure 7A). If the carbonyl group is assigned position i in the main chain and the H-bond connected amide group position $i+n$ in the same chain, where n is a number of residues, tight turn families γ , β , α and π will have n equal to 2, 3, 4 and 5 respectively (CO_i to NH_{i+n}).^[105] Turn families with a reverse conformation can also be distinguished, where the H-bond is formed between the amide group of the first backbone residue, i , and the carbonyl group of the last backbone residue, $n+i$ (NH_i to CO_{i+n}). These are called δ - and ϵ -turns, with n equal to 1 and 2 respectively.^[103,106] There are also β -, α - and π -turns that can be classified as “open” or “disordered”, where there is no intraturn hydrogen bond, and the distinction is made based on a specific distance between the first and last $\text{C}\alpha$ in the turn.^[103,105] The turns do not need to be isolated from each other, and they can often be found in sequence as a multiple turn type e.g. in combinations $\gamma\beta$, $\beta\gamma$, $\gamma\gamma$ or $\beta\beta$.^[107]

Another structural element in proteins is a β -strand. A β -strand is a simple and common structural motif where amide bonds of the backbone are in nearly coplanar arrangement and the side chains pointing below and above the plane formed by the backbone. The amino acids constituting a specific β -strand do not interact with each other through intramolecular H-bonds. β -strands are typically not found in isolation and H-bonds tend to be formed between different β -strands, creating β -sheet structures, where the alignment of the strands in respect to one another can be parallel or antiparallel (Figure 7B). The parallel β -sheets have their component β -strands oriented in the same direction (both strands run from the N-terminus to the C-terminus), and the antiparallel β -sheets have their strands oriented in the opposite directions (one strand runs from the N-terminus to the C-terminus, whereas the other runs from the C-terminus to the N-terminus). β -sheets are regular structural elements that have primarily a scaffolding role, but they can also contribute to the crucial recognition structures in proteins.^[108] They allow for “neutralisation” of the backbone polarity in the context of the hydrophobic protein interior due to the extensive intra-structural H-bonding across the sheets.^[109,110] Two consecutive antiparallel β -strands that are connected to each other through H-bonds and linked by a previously described turn element comprise a common structure referred to as a β -hairpin.^[111]

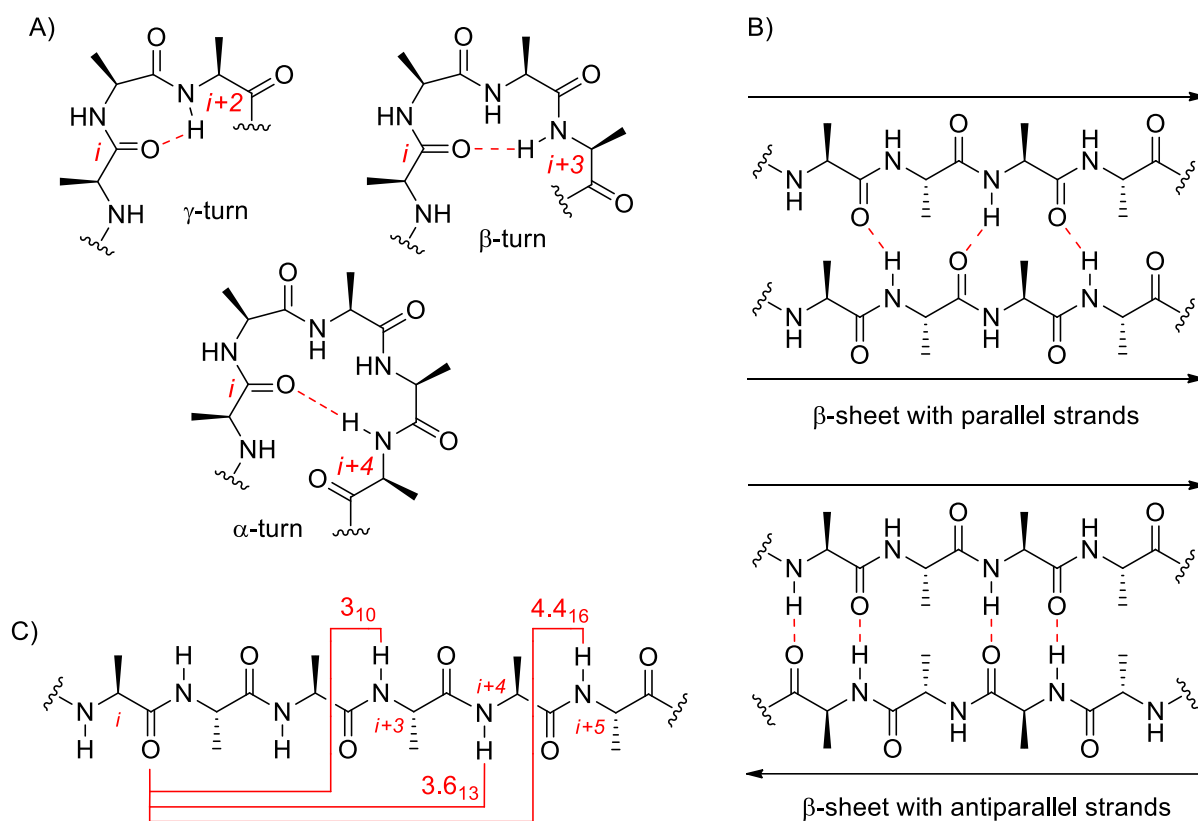


Figure 7. Schematic representation of the most important structural protein elements: **A)** γ -, β - and α -Turns. **B)** β -Sheets with parallel and antiparallel strands. **C)** 3_{10} , 3.6_{13} and 4.4_{16} Helices.

The red lines represent H-bonds. The figure is based on the images from Pelay-Gimeno *et al.*, 2015.^[95]

Another common structural elements in proteins are helices. Helices can be divided into three different types, depending on the H-bond patterns that they incorporate, where each type is designated using n_x notation with n representing the number of residues per full turn and the subscript x signifying the number of atoms in the turn closed by an H-bond.^[112,113] Helices are thus classified as 3_{10} , 3.6_{13} (also called α -helices) or 4.4_{16} (also known as π -helices; Figure 7C).^[113–116] Helices form intramolecular H-bonds between the amide bonds of the backbone, connecting CO_i to NH_{i+n} , where i is a sequential number of the residue in a protein, with i equal to 3, 4 and 5 for 3_{10} , α - and π -helices, respectively.^[113] Similarly to the β -sheets, the internally formed H-bonding network allows to neutralise the backbone polarity of the helical structures.^[109,114] α -Helices are the most common class, occurring in over 30% of secondary protein structures, followed by π -helices which are estimated to be present in ca. 15% of all proteins and 3_{10} helices are constituted by approximately 4% of the protein residues.^[117–121] The explanation for the domination of α -helices in the protein structures could be the more optimal geometry in comparison to the alternative helical arrangements.^[113,115,116,121,122] Despite the energetic handicap of the π -helices, their popularity within proteins can be explained by their involvement and enhancement of various protein functions.^[121] Approximately 60% of the PPI interfaces features α -helices, and among those, ca. 60% utilises one face of the helix containing hot

spot residues with the remaining interfaces using two or three faces populated with hot spots, emphasising the important role of these structural elements in the context of protein complexes and the modulation of their interactions.^[123]

Mimicking peptidic turns, sheets and helices can be a viable strategy to inhibit PPIs, and thus, class A and B of peptidomimetics can be afforded through stabilisation of these key structural elements using such modifications and enhancements as turn inducing, β -sheet-enforcing and other nonproteinogenic amino acids, various macrocyclisation strategies, as well as incorporation of N-methylation into the sequence to name a few.^[95] Examples and benefits of the stabilising peptide modifications are presented in the following section 1.2.3.

1.2.3. Peptide Limitations, Selected Modifications and Macrocyclisation Strategies

In a context of biological applications, unmodified peptides face a number of challenges regarding cell permeability, oral bioavailability, metabolic stability and very quick renal clearance from the body.^[97] The metabolic stability is one of the first hurdles that need to be overcome in order to develop an effective peptidic PPI inhibitor. Proteases are ubiquitous components occurring in every living cell, and are found in several hundred known variants. Disordered linear peptides based on native sequences are thus very exposed and vulnerable to the rapid *in vivo* degradation, resulting in short half-lives ($t_{1/2}$) often in a range of only a few minutes.^[97,124,125] Various strategies improving the resistance to proteases through structural and sequence modifications of peptides are often very effective and relatively amicable to implement. Macrocyclisation can be a very effective modification improving the peptide resistance to proteases. One example of a successful peptide stabilisation is cyclopeptide termed cyclopeptide targeting the NK-1 receptor, that has over seven times higher $t_{1/2}$ (30 min) than the linear analogue (4 min) when tested in liver slices.^[126] Another example of a proteolytical stability increase achieved through macrocyclisation is a bicyclic peptide inhibiting a PPI between MTA1 and RbAp48, with a $t_{1/2}$ of over 90 min in cell lysate, significantly longer than the $t_{1/2}$ of 16 min for the linear counterpart derived from MTA1.^[127] Mutations of the peptidic sequence through introduction of non-proteinogenic amino acids, as well as N-methylation of the backbone amide bonds can also afford structures with remarkable stabilities as exemplified by the natural product cyclosporin A, a cyclic structure with non-canonical amino acids and seven N-methylated peptide bonds, and the prostate cancer drug Degarelix with six unnatural residues in a ten amino acid long, linear sequence.^[128–131] As proteases cleave peptidic amide bonds, the bond replacement with an enzymatically more stable bioisostere can be an attractive strategy as shown for example by an analogue of glucagon-like peptide-1 containing a thioamide at a critical position in the sequence, and thus, increasing the

stability to dipeptidyl peptidase 4 up to 750-fold.^[132] Modifications of linear peptide termini by N-acetylation and C-amidation are also known to considerably improve the enzymatic stability, and can be easily introduced during peptide synthesis.^[133–135]

Arguably, the second most important characteristic of peptides judged from the perspective of biological experiments, after a reasonable enzymatic stability, is the ability to enter cells. Regular peptide sequences under normal circumstances are not expected to be transported into the cells, however, there are known sequences capable of penetrating membranes, so called cell penetrating peptides (CPPs), and in some cases they can even enable to transport attached cargo into the cells.^[136] The problem of assuring that peptides are capable of penetration into the cytosol is non-trivial, chiefly due to the poorly understood uptake mechanisms for diverse structures and sequences and the difficult to predict interactions between the cells, the CPPs and the cargos, under a particular set of conditions. The level and mode of uptake is also strongly depended on the cell line tested, the associated cargo, type of the interaction between the cargo and the CPPs, as well as concentrations of the CPPs.^[136,137] Cells very quickly interiorise and recycle large portions of their surface area and volume, thus, creating potentially conducive circumstances for the peptide internalisation through endocytosis.^[138] In fact, a significant number of CPPs enters cells through macropinocytosis and receptor mediated endocytosis that can be assisted by clathrins or caveolins.^[139–143] The peptide transport into the cells can also occur through lipid raft-mediated endocytosis.^[144] A challenging but crucial, yet still not fully understood and limiting step for a successful cellular uptake of CPPs is the efficient escape from endosomes, where the peptides can be trapped and degraded over time.^[139,145–147] Alternative peptide pathways into the cells, that have been less extensively explored, involve transporters such as organic anion-transporting polypeptide (OATP) and proton-coupled oligopeptide transporters (SCL15).^[97,148,149] Peptides with the right physical and chemical characteristics can undergo direct penetration or membrane transduction into cells in an energy independent manner. There are currently a few models of direct peptidic penetration mechanisms, including the barrel-stave pore formation model, the carpet model, as well as the inverted micelle model.^[150–157] The barrel-stave pore formation model assumes that amphipathic peptides form α -helices within the membranes with the hydrophilic side chains directed to the inner side of the transient pore, allowing for the hydrophobic segment of peptides to cross the membrane.^[150–152,157] According to the carpet model, the positively charged peptide residues interact with the negatively charged membrane, increasing the local fluidity of the membrane, and thus, a diffusion into the cytosol.^[153,154,157] The inverted micelle model is based on a concept of inverted micelles encapsulating the peptides. The transport into the cytosol occurs inside a micelle formed at the cell membrane, which can then release the peptides into the inside of the cell.^[155–157] The understanding of peptide penetration is further

complicated by the fact that the uptake mechanisms are not mutually exclusive and the entrance into the cells can occur simultaneously through different pathways.^[97,158,159]

There are a number of modifications that allow to impart or improve the cell penetrating ability of peptides. Natural and artificial CPPs can be conjugated to the cargo intended for the cellular transport. Common CPPs used are TAT, penetratin and poly-Arg peptides such as R8 or R9 with eight and nine Arg residues, respectively.^[136,160] The TAT sequence (RKKRRQRRR) is derived from the Tat protein encoded by the HIV-1 virus, required for virus replication.^[161–164] An exemplary application of this sequence was performed for the aforementioned bicyclic inhibitor of the PPI between MTA1 and RbAp48, allowing the peptide to cross the membrane and exert the intended biological effect.^[127] Penetratin is a somewhat longer sequence (QIKIWFQNRRMKWKK) found in the Antennapedia protein of *Drosophila melanogaster*.^[155,164,165] A conjugate between peptide KLA and penetratin is another example of how a CPP can transform ineffective agent with very limited cell permeability (KLA) into efficient species that shows potent cytotoxic activity against cancer cell lines.^[166] Highly positively charged poly-Arg sequences, typically with eight Arg residues or more, are also able to induce cell penetration, where the increase in the number of the Arg residues is proportional to the increase in cellular uptake.^[167,168] Exemplary application of the poly-Arg CPPs is enhancement of the insulin permeability through epithelial cell membranes.^[169]

The peptides ability to penetrate cells can also be facilitated by relatively small non-peptidic tags such as asparagusic acid or a fluorour tag.^[159,170] The asparagusic acid tag utilises endogenous thiol groups located at the cellular surface and due to disulphide exchange reactions between the tag and the exofacial thiols, the cargo covalently associates with the cells resulting in enhanced cellular uptake.^[171–173] An alternative approach is fluoroalkylation of the cargo, where the fluorour-tagged peptides assemble into highly proteolytically stable nanostructures. The tags are attached via a disulphide bond to the cargo. The nanostructures are internalised through various uptake pathways, and the cargo is released in a result of the disulphide bond reduction by intracellular glutathione (GSH). The fluorour tag has been recently shown to work very effectively in conjugation with the KLA peptide, penetrating into cells and excreting a cytotoxic effect against cancer cells, as well as reducing a tumour size when tested in mice.^[159] Covalent attachment of CPPs and tags improving the cellular uptake is not always necessary, as non-covalent shuttle peptides can also be a potential solution used for delivering desired cargo into cells. A shuttle called dTAT is a dimeric TAT peptide connected via an N-terminal disulfide bond, where each TAT sequence is additionally labelled with the fluorophore tetramethylrhodamine (TAMRA). The conjugation to TAMRA proved to enhance the endosomal leakage following macropinocytosis. When a desired cargo is mixed with dTAT and administered onto cells, the presence of dTAT allows for a cellular co-internalisation of the cargo material.^[147,174,175] A similar

principle was utilised for a rational design of amphiphilic peptide shuttles containing in their sequences a hydrophobic cluster and hydrophilic/cationic elements. The amphiphilic shuttle peptides have been shown to deliver cargos into a number of cell lines.^[176,177] Very recently, a novel shuttle cellular delivery system was developed allowing to transport a variety of macromolecules into cells, utilising peptidic coacervate microdroplet structures. Appropriately designed peptide sequences are conjugated through their Lys residues to a disulphide bond-containing protective moieties, forming a phase separating microdroplets that can encapsulate within them large cargo. The coacervates have been shown to be taken up by the cells through a cholesterol-dependent lipid rafting. Once inside the cytosol, the disulphide bond in the protective group of a Lys side chain is reduced by GSH, leading to the release of the Lys side chain and in effect causing the coacervate dissociation and cargo release.^[178]

Once the problem of peptide stability and cellular uptake is solved, the obtained agents could constitute very effective molecular probes for biological and biochemical investigations. However, to reach the full therapeutic utility and effectiveness further concerns often need to be addressed, especially in respect to the oral bioavailability and renal clearance of peptides. In addition to the low pH of the stomach and a number of proteases produced by the gastrointestinal (GI) tract, the orally administered peptides need to also survive the first-pass metabolism and clearance. In order for the peptide to be absorbed into the bloodstream from the GI tract, it needs to penetrate typically in a passive manner either through tight junctions of the GI cells or pass transcellularly through enterocytes, a process significantly more challenging than traversing just one membrane during a treatment of isolated cell lines.^[97,179] Oral absorption of peptides can be increased through formulation using enzyme inhibitors, pH modifiers and absorption enhancers such as chelating agents, detergents, surfactants, mucoadhesive polymers and formulation vehicles, for example in form of emulsions, nanoparticles or microspheres.^[179–183] Following the entrance into the bloodstream, the peptides can be rapidly filtered out by kidneys, even when the structures of the peptides remain unmodified. Most smaller peptides below 5 kDa can pass easily through glomeruli (the filtering structures in kidneys) and due to the typically poor reabsorption in the proximal tubule they are rapidly removed from the system. One of the solutions to the renal clearance is the increase in the hydrodynamic radius of molecules which prevents the filtration. This can be achieved by the peptide association with serum albumin, a protein that transports lipophilic compounds. If a peptide is sufficiently hydrophobic it can naturally bind to albumin, otherwise a lipophilic motif such as a fatty acid can be attached to the peptide in order to ensure its association with albumin, thus extending the plasma retention time.^[133,179] An example of a peptide-based molecule tagged with a fatty acid with a long plasma retention can be a clinically approved insulin analogue called insulin detemir.^[184–186] An alternative

approach to extending the plasma retention is the attachment of long polyethylene glycol (PEG) chains to the peptide termini.^[133]

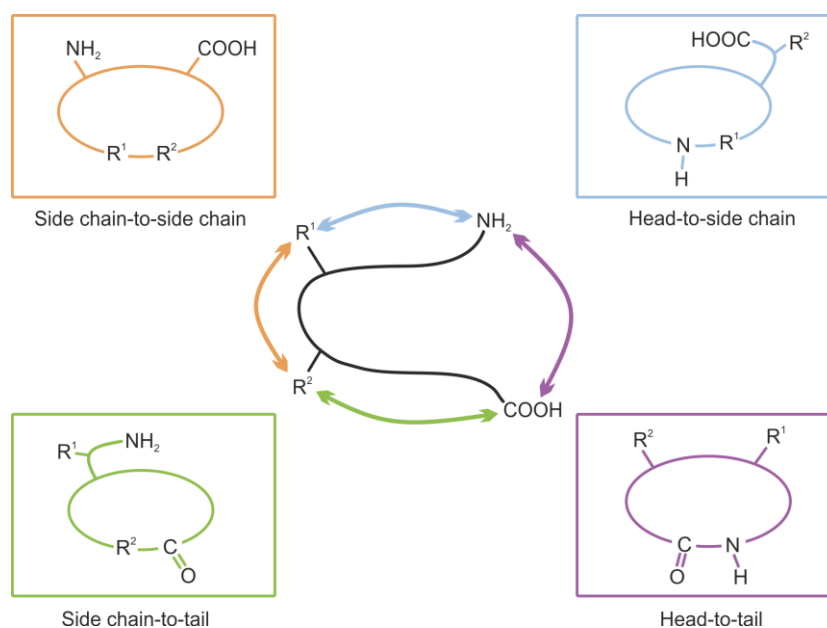


Figure 8. The majority of peptide cyclisation techniques utilise one of the four connection types: side chain-to-side chain, head-to-side chain, side chain-to-tail or head-to-tail.

The figure is based on the image from White and Yudin, 2011.^[187]

Peptide macrocyclisation is not only used for increasing the stability of peptides, but adequate cyclisation also introduces a level of conformational restraint and structural preorganisation that allows to minimise the entropic penalty paid upon binding to the protein target, and thus, significantly increases the compound affinity in comparison to its flexible linear analogue.^[49,97,188] Peptide cyclisation often requires special synthetic and structural considerations, as the geometry of the peptide bonds is normally not conducive to cyclisation into small and medium sized macrocycles. This problem is alleviated when larger macrocycles are obtained, however, the large ring sizes still require caution at the formulation stage in order to prevent intermolecular linking.^[187] Certain peptides with small ring sizes can be exceptionally difficult to synthesise, as demonstrated for cancerostatic marine cyclotetrapeptides by Schmidt and Langer.^[189] The very short peptidic sequences have a substantial tendency to dimerise or to epimerise at the C-terminal residue involved in the cyclisation.^[189] The macrocyclisation typically needs to be performed under conditions of high dilution or pseudo-dilution, either in solution or with the peptides attached to a solid support. High dilution allows to minimise the chances of oligomerisation which is typically crucial while trying to obtain cyclic peptides with smaller ring sizes.^[187,190–192] The success of cyclisation attempts can also strongly depend on the specific sequence of the peptide.^[189,193]

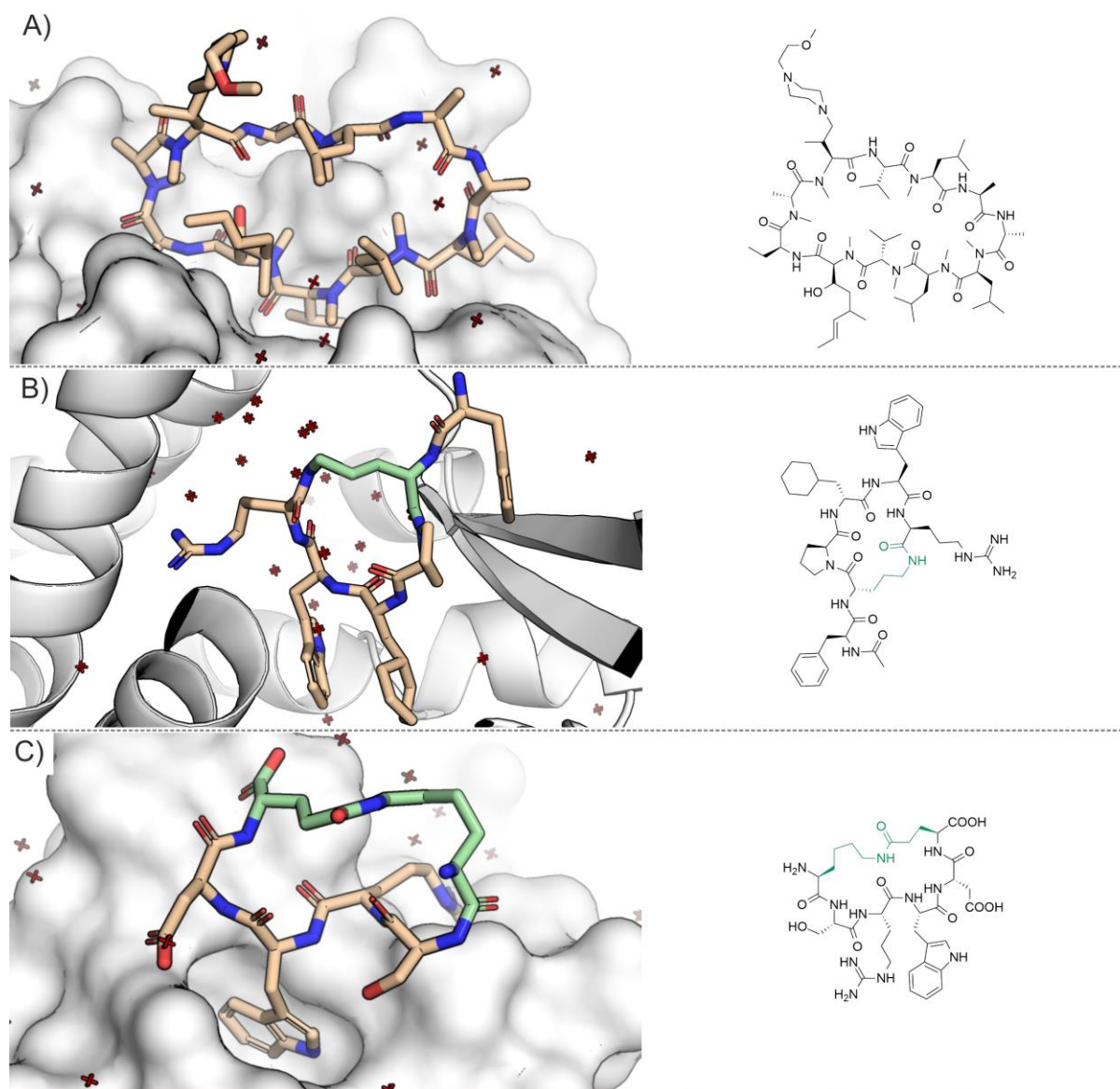


Figure 9. Exemplary peptidomimetic PPI inhibitors: **A)** Head-to-tail cyclised peptide NIM258 bound to protein cyclophilin D (PDB ID: 4TOT).^[194] **B)** Side chain-to-tail cyclised peptide PMX53 interacting with CtaR receptor (PDB ID: 6C1R).^[195] **C)** Amide-based side chain-to-side chain macrocyclic binder of UHM domain (PDB ID: 5LSO).^[196]

There is a plethora of existing cyclisation techniques available, but the majority of the peptide macrocyclisation strategies involve a connection of side chain-to-side chain, head-to-side chain, side chain-to-tail and head-to-tail (Figure 8). The ring is frequently closed through lactamisation, lactonization or disulphide bond formation.^[96,179,187,197–199] Macrocycle NIM258 is an example of a head-to-tail cyclised peptide, which shows inhibitory activity against cyclophilin and hepatitis C virus. In NIM258 the amine of the N-terminus is connected to the C-terminal carboxylic acid forming a new amide bond (Figure 9A).^[194] Another cyclopeptide example called PMX53, an antagonist of CtaR receptor found on immune cells, is a side chain-to-tail macrocycle that was designed to mimic the C-terminal turn of the C5a protein interacting with the CtaR receptor. PMX53 incorporates a connection between the side chain of Lys with the C-terminal carboxylic acid (Figure 9B).^[195,200] The amide-based

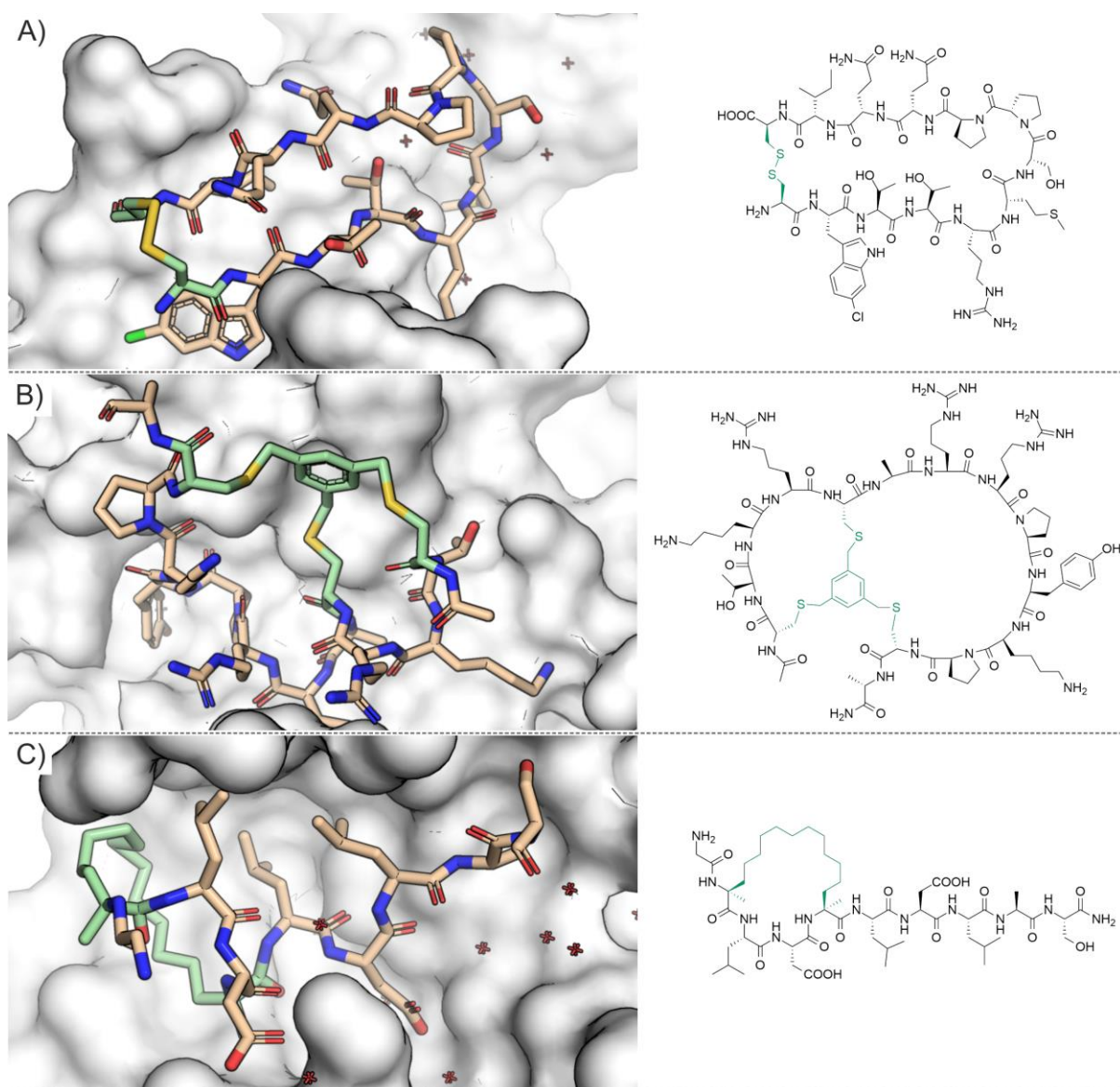


Figure 10. Exemplary peptidomimetic PPI inhibitors: **A)** AMA-1 binder, cyclised via a disulphide bond (PDB ID: 4Z0F).^[201] **B)** RbAp48 binder cyclised using the CLIPS technology (PDB ID: 6ZRD).^[127] **C)** RCM-based peptide bound to protein 14-3-3 ζ (PDB ID: 4N84).^[202]

side chain-to-side chain connection where Lys is connected to Glu is exemplified by a binder of a U2AF homology motif (UHM) involved in the regulation of the alternative pre-mRNA splicing processes (Figure 9C).^[196] Macrocycles are also created using disulphide bonds formed between Cys residues or Cys analogues e.g. homocysteine which has a longer side chain.^[201,203,204] The disulphide-based linkers are relatively unstable and thus unreliable as they are prone to reduction in a cellular environment. An improvement in the chemical stability of the linker can be achieved by replacing the disulphide connections with thioesters.^[205] The utility of the disulphide bond for macrocyclisations is demonstrated on the example of a peptidic antimalarial agent. PPI inhibitor of the protein antigen 1 (AMA1) interactions with rhopty neck protein 2 (RON2), prevents the host infection by the malarial

parasite. The inhibitor was designed based on the C-terminal loop of RON2, where the peptidic termini are connected via a disulphide bond stabilising a β -sheet structure of the peptide (Figure 10A).^[201]

In addition to the classical linking strategies peptides can be also cyclised using more elaborate methods.^[95,187,206] Chemical linkage of peptides onto scaffolds (CLIPS) technology is a convenient and quick synthetic technique where thiols located on a peptide react with arylmethylene bromides resulting in mono-, bi- and even tricyclic peptides (Figure 11A).^[207–209] As an example, the already aforementioned PPI inhibitor of the MTA1 and RbAp48 interactions utilises the CLIPS technology to form the bicyclic frame of the peptide (Figure 10B).^[127] A similar approach developed by Pentelute and co-workers uses a perfluoroaromatic molecule such as hexafluorobenzene for linking peptidic thiols generating bis-arylated macrocycles (Figure 11B).^[210] The Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC), known as the “click” reaction, commonly applied as a bio-compatible technique for ligations, is also used for peptide cyclisations where the azide and alkyne groups are transformed into a triazole linker (Figure 11C).^[211–213] A double click reaction between a peptide and a compound containing two alkynes, allowing to introduce an extensive additional functionalisation to the native peptide, is also possible.^[214–216] When the appropriate linker is used the peptidic conformations of macrocycles can be controlled dynamically. The incorporation of photoswitchable azobenzene linkers into peptides allows to flip between *trans* and *cis* isomers after irradiation with light allowing to activate or deactivate the cyclised peptide towards the target protein (Figure 11D).^[217–220] A common macrocyclisation approach is formation of a hydrocarbon linker through ring closing metathesis (RCM) that uses ruthenium-based catalysts such as the Grubbs catalyst or the related Hoveyda-Grubbs catalyst (Figure 11E).^[202,221–227] RCM produces an unsaturated linker that can be reduced to a fully saturated form.^[202,223,227] An example of a bioactive peptide incorporating an RCM-based hydrocarbon linker is a PPI inhibitor of the interactions between virulence factor exoenzyme S of the pathogenic bacterium *Pseudomonas aeruginosa* and the human protein 14-3-3, where the linker stabilises an irregular conformation of the peptide (Figure 10C).^[202]

Obtaining a population of compounds with only a single conformation is challenging and cyclisation can often be insufficient to attain this goal, thus, further strategies that allow to introduce conformationally restraining elements to the compound structures can be applied.^[95] The amino acid population occurring at β -turns overrepresents certain amino acids, such as Pro, Gly, Asn and Asp.^[103,228] Pro is the only proteinogenic amino acid with its sidechain connected to the backbone amine, which results in a significantly limited conformational freedom and the absence of the amide proton in the peptidic sequence, preventing H-bond interactions with the backbone nitrogen. Thus, Pro has a crucial structural function in proteins, shifting the conformational ratio towards the *cis* isomer of the amide bond.^[229–233] The D isomer of Pro can also act as an β -turn inducer and stabiliser,

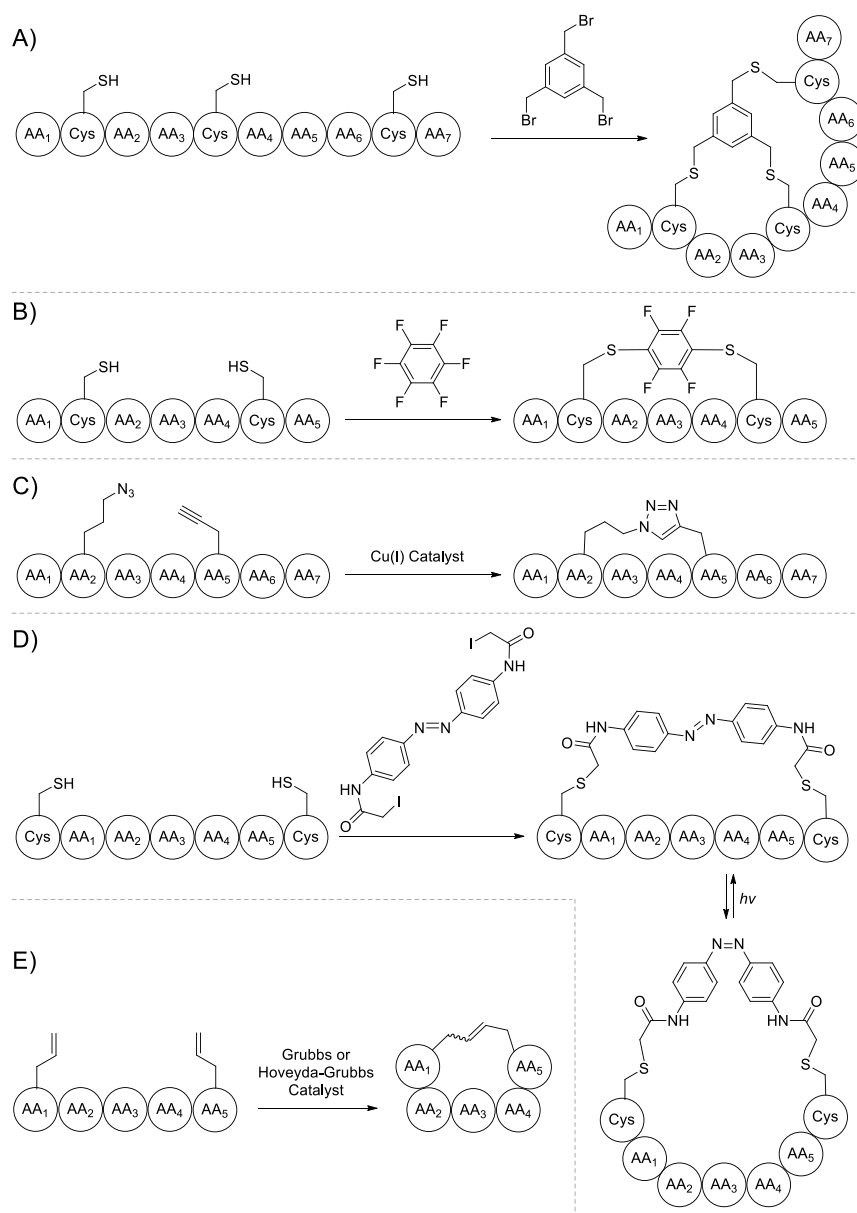


Figure 11. Schematic representation of exemplary peptide cyclisation approaches: **A)** CLIPS strategy resulting in a bicyclic peptide. **B)** Peptide cyclisation using hexafluorobenzene. **C)** Cyclisation based on a CuAAC click reaction. **D)** Incorporation of a photoswitchable azobenzene linker, which can flip between *trans* and *cis* isomers. **E)** Peptide cyclisation through RCM.

and motifs such as D-Pro-L-Pro and D-Pro-Gly have been commonly and reliably used for inducing β -turns and antiparallel β -hairpins in peptidic structures.^[234–242] An introduction of the D-Pro-DADME (1,2-diamino-1,1-dimethylethane) linker to a peptide works in an analogous manner but results in parallel β -sheet structures.^[243,244] The D-Pro-L-Pro motif has been successfully used to design cyclic β -hairpin mimetics inhibiting PPI between p53 and HDM2 proteins.^[245,246] α -Aminoisobutyric acid or Aib is a nonproteinogenic amino acid and is another example of a β -turn inducing element through conformational restriction of peptidic flexibility that can be used in combination with Gly as an Aib-Gly motif.^[247] As described before, peptidic amide bonds can be N-methylated, modifying the

conformational preferences and possible interactions of affected structures.^[130] N-methylation changes the number of H-bond interactions that a peptide can form, as it removes an H-bond donor from the structure, and it also changes the equilibrium of the *cis-trans* conformer equilibrium in a similar manner to the impact of Pro, promoting the *cis* conformation.^[130,229,248,249] The N-methylation of the peptidic backbone can also be a useful method for preventing aggregation and fibril formation.^[250–253] The stabilisation of β -hairpin arrangements can also be achieved through an approach based on non-covalent interactions. Peptidic chains can be equipped with N-terminal acetylated Trp as well as a C-terminal Trp-Thr-Gly motif, where the stabilisation is achieved through interactions between the indole groups of the tryptophans, and the H-bonds between the terminal regions.^[254] A similar strategy involves Trp zippers or trpzip, where reversible interactions between Trp-Trp in cross-strand pairs are utilised for β -hairpin stabilisation.^[255,256] β -hairpin formation can also be stabilised through introduction of amino acids that can connect through π -cation interactions, typically Trp and a positively charged, side chain-methylated amino acid.^[257–259] The β -sheet arrangements can be induced and stabilised through incorporating β -strand-enforcing amino acids. Amino acids such as 1,2-dihydro-3(6H)-pyridinone (Ach) and 5-hydrazino-2-methoxybenzoic acid (Hao), are conformationally constrained and can mimic the H-bond pattern of β -sheets, thus promoting this type of structure in a peptide.^[260,261] Small molecule fragments can be incorporated into a peptide replacing and mimicking the backbone of the turn, that can potentially result in molecules with improved pharmacokinetic properties, increasing an overall resemblance to small molecules.^[95,262–265] Normally, in order for a small cyclic peptide to form a β -sheet arrangement, the peptide needs to have appropriate amino acid composition, the number of its ring amino acids must satisfy rule $2(2n + 1)$ where n is a positive integer, and have only two equally separated type I' or type II' β -turns necessary for the correct residue alignment.^[266]

Macrocyclisation techniques are also very often used for stabilising, or stapling, peptides in their α -helical conformations, where the molecule is covalently cross-linked at positions i and $i+3$ or i and $i+4$ spanning one turn or alternatively positions i and $i+7$ spanning over two helical turns.^[95,267–269] A double stapling with two separate linkers is also viable approach, as well as using two cross-links attached to the peptide through a central spiro connection, thus, affording a “stitched” helix.^[270,271] The length of the linker is a very important parameter impacting the helicity and the target recognition.^[95,204] The α -helix stapling is known to often improve the cell penetrating ability of the modified peptide.^[136,158,272] The presence of an amino acid at the N-terminus of an α -helix that can form an H-bond between its side chain and the backbone of another residue can nucleate a helix.^[273,274] An artificial capping can be introduced to the N-terminus of an α -helix, a motif that is also called an N cap, to mimic such nucleating behaviour and covalently stabilise the turns.^[273,275] As helix

nucleation is a very energetically demanding process the incorporation of an N cap into a structure can be very effective for generating successful peptidomimetic PPI inhibitors.^[275–278] This approach allows to stabilise the structure without compromising the side chain interactions with the protein target.^[279]

Another possible peptide modification, resulting in the class B peptidomimetics, involves incorporating β -amino acids, building blocks that contain an additional methylene group attached to the backbone of a conventional α -amino acid. The side chain of the β -amino acids can be located at the carbon C_2 or C_3 , or the amino acids can have two side chains simultaneously on C_2 and C_3 .^[95,280,281] A structure composed of only β -amino acids is referred to as a β -peptide, but the β -amino acids and α -amino acids can be also combined together affording a heterogenous oligomer. In such hybrid structures, the α -helical mimicry, as well as the protease stability can be improved, where the α -amino acids can be used for the target recognition and the β -amino acids contribute to the helicity of the structure.^[282–286] Alternatively, the side chain of amino acids in peptides can be moved from the α -carbon to the amide nitrogen, affording peptoids or poly-N-substituted glycines. Peptoids are typically characterised by improved stability to proteases, as well as improved cell permeability especially when cyclised.^[287–293] The peptoid monomers can be combined together with regular α -amino acids into one structure, affording effective PPI inhibitors.^[294,295]

There is a wide range of possible peptide modifications available, but the impact of the peptide alterations on the interactions with the target protein and the selectivity in the context of modulating PPIs is highly depended on a specific sequence and structure of the peptide. Thus, the modifications need to be carefully tailored based on the in-depth analysis of each particular interaction case.

1.2.5. Fundamental Principles of Fmoc-based Solid Phase Peptide Synthesis

In 1963, synthesis on solid support was introduced by Merrifield eliminating the need to tediously synthesise peptides in solution.^[296] The solution-based method required peptides to be purified after each step, whereas, solid phase peptide synthesis (SPPS) allows to grow peptidic chains on a polymeric resin, where solvents, reagents and unattached by-products can be simply washed out of the resin. SPPS also allows to use a large excess of reagents, helping to drive the coupling reactions to completion, and all steps can be performed in one vessel with resin washes in-between the steps, eliminating the need for the transfer of the peptidic material. Thus, SPPS in combination with robotics enables automating the highly repetitive procedure of the peptide synthesis.^[134]

In SPPS, the synthesis occurs from the C-terminal end of the peptide to its N-terminus. Normally in the first step, the N-terminally protected amino acid is attached through its C-terminus to the hydroxyl or amino resin, which results after cleavage in a C-terminal acid or amide, respectively. The peptide is now assembled in a linear fashion, repeating the cycles of N-terminal deprotections and coupling steps. In order to prevent branching and side reactivity, the functionalised side chains of the amino acids used for coupling are protected with groups that are orthogonally cleaved to the temporary protective group of the α -amine, and are stable under the conditions of the chain synthesis.^[134,297] Due to the repetitive nature of the method, the N-terminal group protecting the α -amine needs to be removed under relatively mild conditions that do not interfere with the chain integrity. Once the desired peptidic sequence is synthesised, the side-chain protective groups are removed and the peptide is cleaved from the resin (Figure 12). Due to the nature of the method, normally only one purification stage is needed in order to isolate the desired product after cleaving from the resin.^[134]

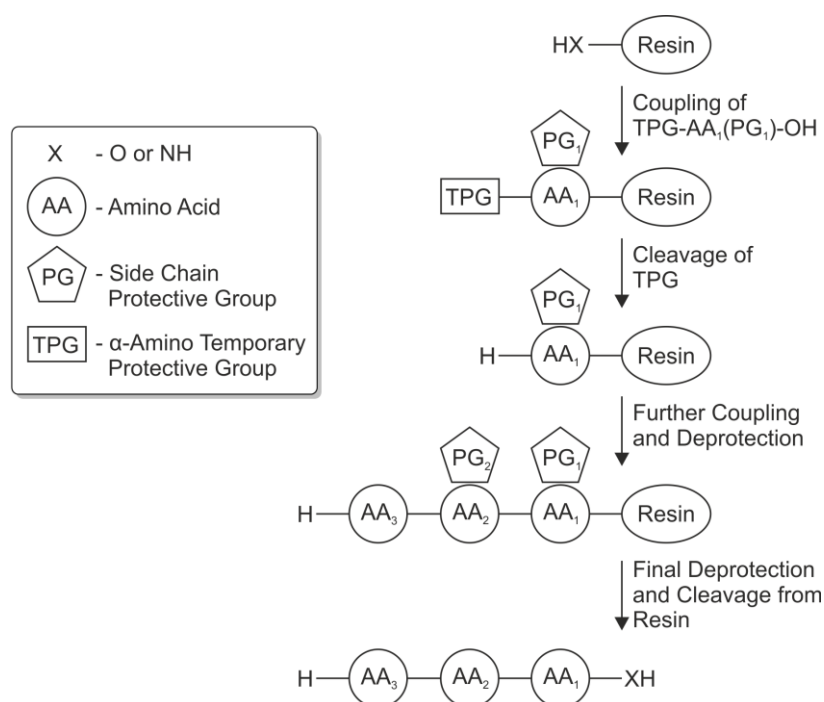


Figure 12. Schematic representation of the main steps in SPPS.

Today, the most common SPPS methodology is based on the Fmoc approach, where the α -amino groups are temporarily protected by the Fmoc group, a successor of Boc-based SPPS. For the Boc-based SPPS highly dangerous and toxic acids are used at the deprotection stages, and special apparatus is needed for carrying out the side-chain removal and detachment from the resin. The Fmoc SPPS is safer and does not require highly specialised equipment.^[134,298] Fmoc is removed under basic conditions using primary or secondary amines, and the most commonly used base for this task is piperidine. The standard groups protecting the side chains are cleaved during the detachment from

the resin, which occurs under acidic conditions. The acid normally used for the final cleavage is TFA.^[134,297] The amino acids need to be activated in order to perform the coupling reaction and extend the peptidic chain.^[134] Some of the most commonly used activating/coupling reagents are PyBOP (phosphonium-based reagent), DIC (carbodiimide), HCTU, HBTU, HATU and COMU (uranium-based reagents). The coupling reactions are performed under basic conditions, where the tertiary amine, DIPEA, is used as base that does not interfere with the Fmoc protective group.^[134,299] When the Fmoc SPPS is used with the appropriate, selectively removable protective groups, it allows for creating macrocycles on resin applying a number of the available cyclisation strategies.^[187,206,300–304]

1.3. Methyltransferase PRMT5

1.3.1. Introduction to Protein Arginine Methyltransferases

Protein methyltransferases (PMTs) are enzymes that use S-adenosyl-L-methionine (SAM or AdoMet) as a cofactor (the source of the methyl group) in order to methylate proteins, producing S-adenosylhomocysteine (SAH) as a by-product. PMTs typically methylate basic nitrogens, and this includes the N-terminal α -amino groups, as well as the side chains of Arg and Lys residues. There are two main classes of PMTs: protein arginine methyltransferases (PRMTs) and protein lysine methyltransferases (PKMTs). As the names suggest, PRMTs methylate selected Arg residues on the targets and PKMTs methylate Lys residues.^[305–308] In humans up to 11 PRMTs and approximately 50 PKMTs have been identified.^[308,309] The known human PRMTs can be divided into three distinct types, depending on the methylation pattern that can be afforded by the enzyme. PRMTs can monomethylate or symmetrically and asymmetrically dimethylate nitrogens of the guanidine moiety in the target Arg residues (Figure 13). Type I enzymes (PRMT1-4, 6, 8) produce ω -N^G-monomethylarginine (MMA) and asymmetric ω -N^G,N^G-dimethylarginine (ADMA). PRMT5 and PRMT9, as well as potentially, PRMT10 and PRMT11 belong to the type II PRMTs, and they generate ω -N^G-monomethyl and symmetric ω -N^G,N^G-dimethylarginine (SDMA) residues.^[305,308,310] It needs to be noted for the type II enzymes, that PRMT9 (FBXO11) and PRMT11 (FBXO10) are not considered as true PRMTs by some researchers, due to the contested proof of the methyltransferase activity for these proteins.^[307,311] PRMT7 is a type III PRMT capable of producing only ω -N^G-monomethylarginine products (Figure 13).^[305,308,310,312,313]

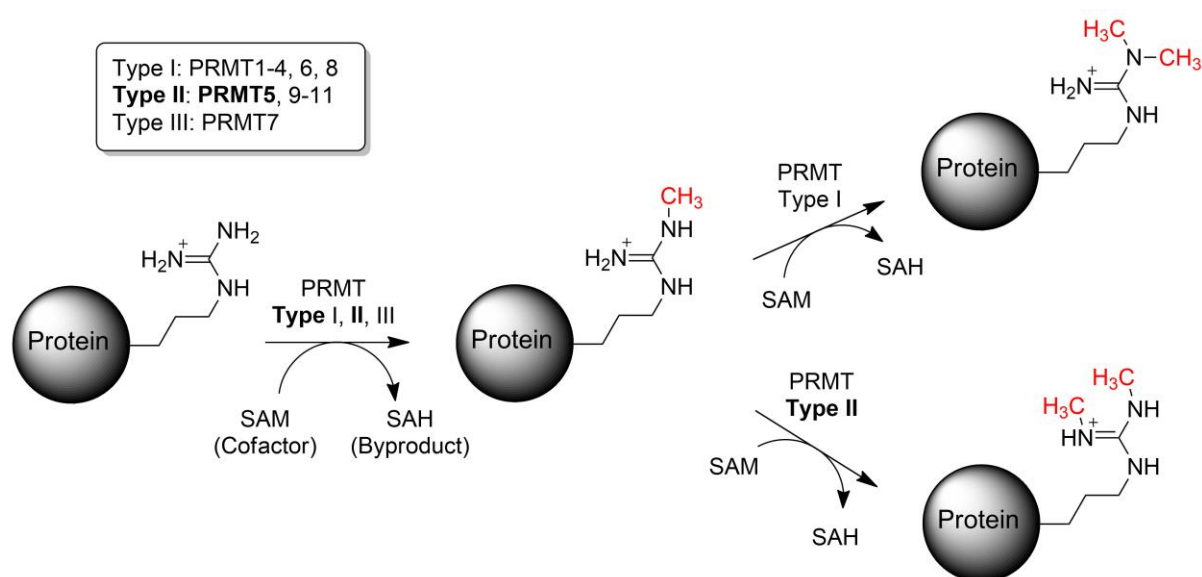


Figure 13. A schematic representation of the methylation activity of the different PRMT types found in humans. PRMT5 is a type II PRMT affording monomethylated and symmetrically dimethylated Arg residues.

The methylation of the guanidinium group of Arg does not alter its charge, but results in a higher hydrophobicity of the side chain, an increased bulk and a decreased potential to form H-bonds.^[307] As Arg can form strong interaction with carboxyl, phosphoryl and guanyl groups, the methylation caused by PRMTs can have a very considerable effect on complex formation and cell signalling, for example, it can weaken the interactions with nucleic acids.^[307,314–316] A significant portion of the Arg-methylated proteins is associated with different steps in RNA metabolism such as RNA biogenesis or splicing. The proteins are also linked to gene expression, translation, chromatin reorganisation or even protein targeting to membrane, microtubule-based processes and responses to the DNA damage.^[306,317] Due to the high importance and a vast number of key functions in cells and a promising potential for inhibition with medicinal relevance, PRMT5 (also known as Jak-binding protein 1, JBP1) has been of great interest among the scientific community.^[306,318–324] The functions of PRMT5 include, among many, promoting the assembly of the structures necessary for the correct functioning of the spliceosome, influencing the central nervous system where PRMT5 is involved in the myelination process and oligodendrocyte differentiation, affecting osteoclast differentiation and bone resorption, as well as affecting the functions of the cardiac tissues.^[11,325,334–337,326–333] PRMT5 also appears to be able to act as either transcriptional repressor or activator, depending on the complexation status and type of the adaptor protein that it interacts with.^[320]

1.3.2. PRMT5 in Disease

PRMT5 is a very influential cellular component that can methylate and change behaviour of such key transcription factors like p53, NF- κ B, E2F-1 or MYCN, as well as interact with and methylate histone tails and spliceosomal proteins, thus, PRMT5 is significantly involved in the regulation of the cellular growth, differentiation and proliferation.^[325–329,338–342] Expectedly, the aberrant expression of PRMT5 in cells has been associated with a wide variety of different cancer cell types, as well as with neurodegenerative and cardiovascular diseases and diabetes.^[306,320,321]

The PRMT5 enzyme is typically strongly overexpressed in cancer and its overexpression often results in poor prognosis for the affected patients.^[343–350] The elevated levels of PRMT5 can lead to the reduced expression of tumour suppressor genes, and thus, this can result in the development of cancer. PRMT5 is strongly linked to breast cancer and it has been shown that PRMT5 can significantly alter the function of programmed cell death 4 (PDCD4) tumour suppressor, accelerating the tumour growth.^[343] PRMT5 can affect the regulation of the alternative splicing, as well as methylation of cell-fate regulators, leading to the attenuation of the cell cycle arrest, thus, contributing to the breast cancer development.^[344,351] The enzyme was also found to contribute to the maintenance and

propagation of breast cancer stem cells, significant in patient relapse, and to play an important role in the resistance to chemotherapeutics.^[352,353] Additionally, PRMT5 involvement is well-established for hepatocellular carcinoma (HCC), where the PRMT5 knockdowns result in inhibited cancer cell proliferation and tumorigenesis.^[354–356] PRMT5 activity can lead to acceleration of the growth of cancer cells and promotion of HCC metastasis.^[355,356] Increased levels of PRMT5 in glioblastoma cell lines, as well as patient-derived primary tumours correlate with tumour growth through silencing specific suppressor genes. Inhibition of PRMT5 activity, however, results in cell-cycle arrest, apoptosis and loss of cell migratory activity, and thus, increased survival in aggressive glioblastoma models.^[347] PRMT5 plays also a significant role in other types of cancer, for example, overexpression of PRMT5 found in lymphoid cancer leads to repression of suppressor genes, resulting in increased growth of the cancer cells.^[349] PRMT5 is a crucial factor promoting growth of prostate cancer cells through epigenetic activation of transcription of the androgen receptor, where inhibition or knock-down of PRMT5 results in the suppression of the cancer cell lines.^[350] Antiproliferative and proapoptotic genes are selectively suppressed through PRMT5 methylation of p53 in lymphomas, and the inhibition of PRMT5 decreases the colony-forming activity of the lymphoma cells.^[357]

The involvement and the link of PRMT5 to human diseases are not only limited to cancer. PRMT5 has also been associated with neurodegenerative disorders: Alzheimer's disease and Huntington's disease, where the methyltransferase appears to be downregulated and the enhanced expression of PRMT5 improves the survival of neuronal cells.^[358,359] The impact of PRMT5 onto the cardiovascular system seems to be context dependent where both the under and overexpression of this methyltransferase can have negative implications.^[321,337,360–363] Finally, PRMT5 is linked to metabolic pathways involved in diabetes, suggesting this methyltransferase as a potential therapeutic target for this disease.^[321,364–367] There is still no comprehensive understanding of the role of PRMT5 in non-cancer diseases and further investigations are required in order to fully elucidate the involvement of PRMT5 in these human disorders.^[321] One rather clear and intuitive conclusion seems to emerge when the analysis of the involvement of PRMT5 into various diseases is made. Due to the extensive regulatory impact of PRMT5 onto the cell functioning, an imbalance and deviation from the natural endogenous PRMT5 levels appears to result in disorders and can be considered as a potential risk factor.

Due to the very high expected biochemical and clinical usefulness of modulating the PRMT5 activity, a high number of active site PRMT5 inhibitors have been developed to date, that act through the competition with the substrate, competition with the cofactor SAM, dual SAM/substrate inhibition, covalent SAM inhibition or through allosteric effects (Figure 14).^[306,313,322–324,368] There are currently a number of PRMT5 inhibitors tested in clinical trials: GSK3326595 (Epizyme and GSK, trials:

NCT04676516 and NCT02783300), JNJ-64619178 (Johnson & Johnson, trial: NCT03573310), AMG 193 (Amgen, trial: NCT05094336), TNG908 (Tango Therapeutics, trial: NCT05275478), PRT811 (Prelude Therapeutics, trial: NCT04089449), PRT543 (Prelude Therapeutics, trial: NCT03886831), and MRTX1719 (Mirati Therapeutics, trial: NCT05245500).^[324,369] The high clinical interest with this methyltransferase and the active inhibitor development observed across the pharmaceutical industry and academic centres further support the important role and involvement of PRMT5 in cellular regulation processes and diseases.

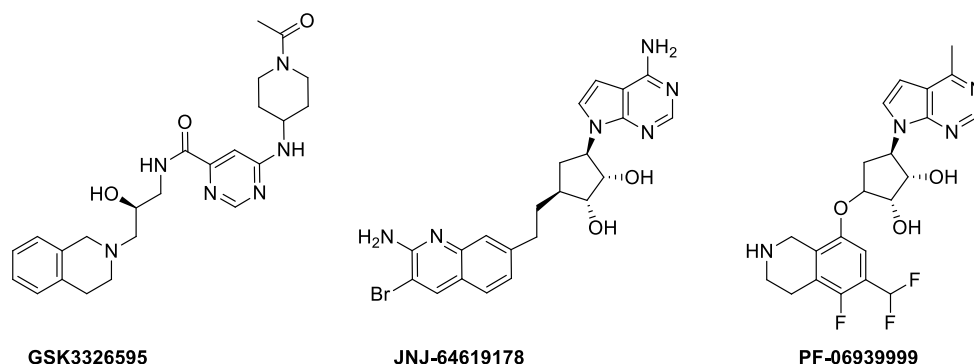


Figure 14. Exemplary active site inhibitors of PRMT5. Molecules GSK3326595 and JNJ-64619178 are currently tested in clinical trials, but the trial of PF-06939999 was recently terminated.^[369]

1.3.3. PRMT5 Structure and Selected Interactions

The structure of the human PRMT5 methyltransferase complex with methylome protein 50 (MEP50, also known as WDR77) was solved for the first time by Antonyamy and colleagues, using X-ray crystallography, with the results published in 2012.^[370] Shortly after, the structure of the PRMT5-MEP50 complex of *Xenopus laevis* (African clawed frog) was also released, with high sequence identity (84%) to human PRMT5.^[371] More recently, the structural conclusion drawn from the crystallography-based experiments was corroborated by observing the human PRMT5-MEP50 complex with the means of cryo-electron microscopy.^[372]

PRMT5 is a 73 kDa protein composed of three domains (Figure 15A): the N-terminal domain triose-phosphate isomerase (TIM) barrel also known as $(\beta\alpha)_8$ -barrel (residues 1-292), Rossmann fold (residues 293-466) and the C-terminal β -barrel domain (467-637). The Rossmann fold and β -barrel closely bind and interact with each other creating a methyltransferase domain. The Rossmann fold contains the SAM binding motif, whereas the β -barrel domain interacts with the substrate.^[370–372] TIM barrel domains are the most common enzyme folds, normally composed of eight units, where each unit is composed of a β -strand and α -helix which are connected together with a $\beta\alpha$ loop. The units of

the barrel are linked through $\alpha\beta$ loops. The eight parallel β -strands form a central, circular β -sheet surrounded by the α -helices. The fold of the TIM barrels is not sequence dependent but is rather caused by the distribution of charged, polar and nonpolar residues. The majority of the TIM barrels are single domain proteins, however, many form homooligomeric assemblies or are a part of multienzyme complexes.^[373] TIM barrels have a potential to catalyse a large number of reactions, where most of these transformations are involved in energy, macromolecule or small molecule metabolism.^[374] There is, however, no prove or indication for the catalytic activity of the TIM barrel of PRMT5. The interface between the TIM barrel domain and the catalytic domains in PRMT5 is stabilised by extensive charged and hydrophobic interactions and the surface area of this interface is equal to 821 \AA^2 .

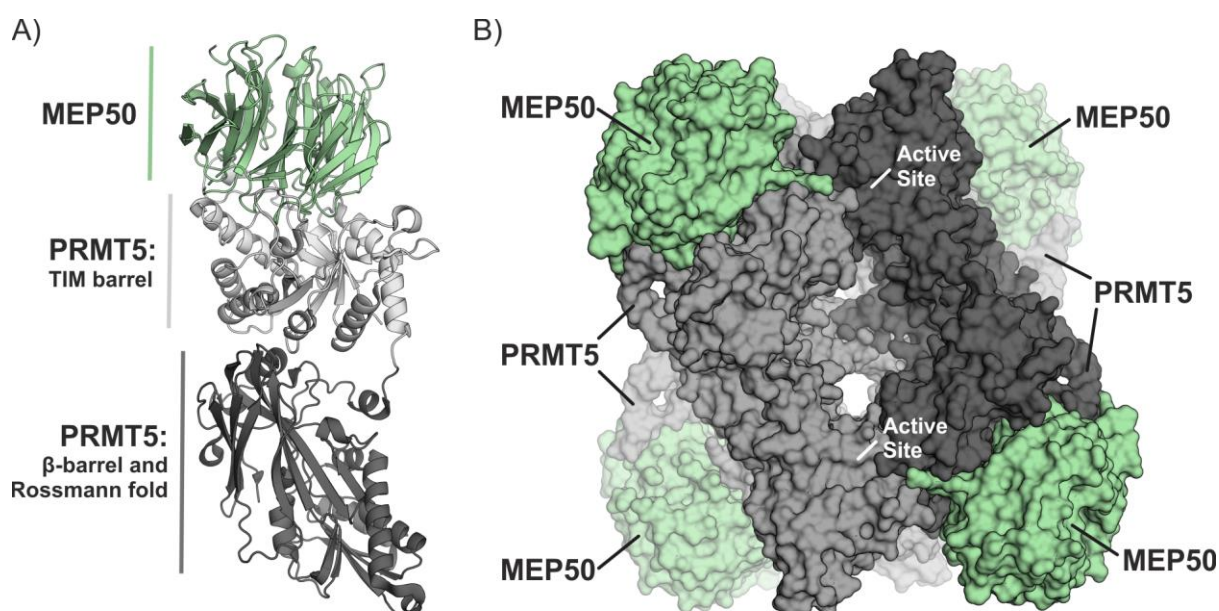


Figure 15. Structure of the PRMT5-MEP50 complex: **A)** Heterodimer of PRMT5-MEP50 (PDB ID: 4GQB).^[370] **B)** Heterooctameric PRMT5-MEP50 (PDB ID: 4GQB). The heterooctameric form is found as a native assembly.^[370]

MEP50 is a 37 kDa WD40 repeat protein found to associate with PRMT5.^[327] WD40 repeat proteins are one of the most abundant domains found in eukaryotes. The WD40 repeats are characterised by the β -propeller architecture with the presence of tandem repeats composed of ca. 40 to 60 amino acids, often with Trp and Asp residues located at the C-termini of the repeats, thus, justifying the name given to those domains. WD repeats normally have four to eight antiparallel sheets arranged as blades composing the β -propeller fold where each sheet has typically four β -strands. The blades of the protein assume an arrangement around a central cavity. So far no WD40 domain has been shown to have any enzymatic activity, but the domains are known to act as adaptors for protein-protein and protein-DNA interactions.^[375–377] MEP50 forms a very stable and tightly bound complex with PRMT5,

and appears to be an obligate *in vivo* binding partner of PRMT5.^[370,378] Knockdown of PRMT5 does not result in free MEP50 and the knockdown of MEP50 does not afford free PRMT5 or its subcomplexes.^[378] It is also known that the knockdown of PRMT5 or MEP50 causes the loss of expression of the other complex partner.^[379] Although, it is possible to obtain isolated human PRMT5 in the absence of MEP50, such afforded protein has a propensity to aggerate and does not crystallise.^[370] MEP50 interacts with PRMT5 solely through the TIM barrel domain via a PPI interface with a large surface area of over 2000 Å², and extensive charged and Van der Waals interactions. PRMT5-bound MEP50 has seven β-propeller blades, where the C-terminal blade has only three β-strands and the remaining six blades are composed of four strands.^[370–372]

The PRMT5 and MEP50 proteins associate into a large 438 kDa heterooctameric complex (PRMT5₄-MEP50₄) with four core units of PRMT5 and four TIM-barrel-bound MEP50 macromolecules (Figure 15B). The complex can be also described as a tetramer of PRMT5-MEP50 heterodimers. The heterodimers in the tetramer are arranged in an alternating head-to-tail manner, and the interactions between TIM barrels and the catalytic domains of different PRMT5 monomers allow the oligomerisation and formation of the heterooctamer.^[370–372]

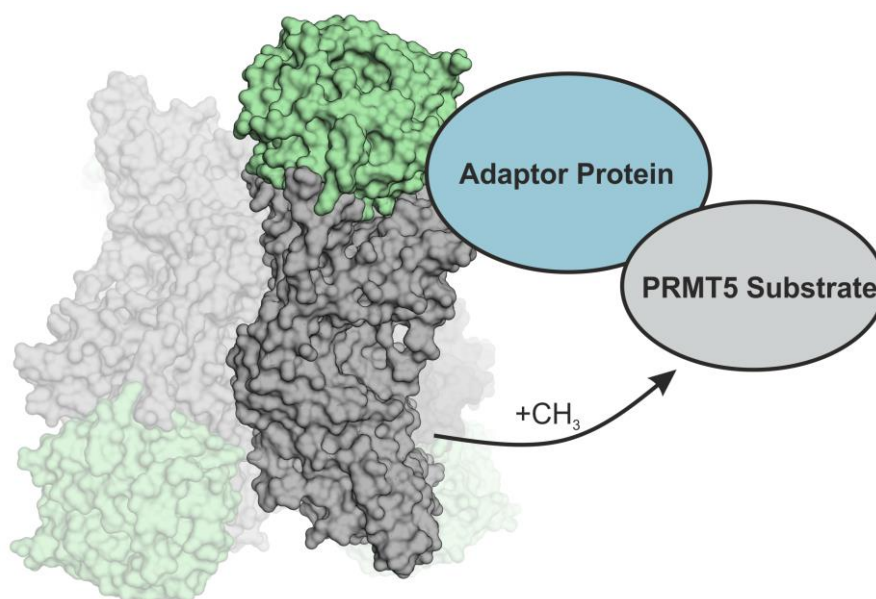


Figure 16. A schematic representation of the methylation activity of PRMT5 (PDB ID: 4GQB). Many substrates of the enzyme are known to be methylated through an assistance by an adaptor protein.

The activity and substrate specificity of PRMT5 is regulated in part by interactions with various adaptor proteins which present the appropriate cellular components for methylation (Figure 16). As noted above, WD40 repeat proteins tend to mediate interactions with other macromolecules, and MEP50 is known to work as a substrate recruiter for PRMT5, interacting with histones and possibly with other

macromolecules such as suppressor of zeste 12 (SUZ12) protein.^[371,380] PRMT5-MEP50 also interacts with such proteins as methylome subunit pICln which recruits Sm proteins for methylation or Rio kinase 1 (RioK1) adaptor protein that allows methylation of nucleolin and nuclear factor 90 (NF90) by the enzyme.^[325,326,381–383] Another interaction partner of the methyltransferase is cooperator of PRMT5 (COPR5), resulting in preferential methylation of histone H4 and recruitment onto chromatin at the target gene cyclin E1 (CCNE1).^[384] A noteworthy PRMT5 adaptor is methyl CpG binding domain 2 (MBD2) connecting the enzyme to specific chromatin regions and the Mi-2/NuRD chromatin remodelling complex.^[385] It is evident that the exact role and regulation mechanisms of PRMT5 in cells are not fully understood, and further investigations regarding the biochemistry of the enzyme and the interactions of PRMT5 with its substrates and the adaptor proteins are needed.

Aim and Objectives of the Thesis

PRMT5 plays an important role in regulation of cellular biochemistry, has high significance in a number of diseases and is associated with a number of principal proteins. However, the mechanisms of its activity and selectivity regulation are still poorly understood. Therefore, the aim of this thesis is to develop a novel PPI inhibitor targeting PRMT5 or one of its adaptor proteins. A successful PRMT5 PPI inhibitor is expected to allow more in-depth investigation of the enzyme and its complexes in biologically relevant environment and expand the current understanding of this intricate methyltransferase.

PART A will describe a development of inhibitors of the PPI between PRMT5 and its most prominent binding partner MEP50. Available crystal structure of the PRMT5-MEP50 complex will be used for identification of the most important regions contributing to the targeted interaction. Peptidomimetics will be designed and synthesised based on the identified binding protein motifs, and biophysical and enzymatic assays will be used for assessing the potential inhibitory activity of the developed compounds. If successful, the most effective structures will be further optimised and analysed under biologically relevant conditions.

PART B will describe an investigation regarding the PPI between MEP50 and SUZ12 proteins. A number of various peptidic SUZ12 fragments will be synthesised and tested for binding to PRMT5-MEP50, in order to find the motif responsible for the interaction with MEP50. If the binding motif is identified successfully, the afforded sequence will be used for a development of the first-in-class MEP50-SUZ12 PPI inhibitor, which will be then further evaluated in biophysical and biological assays.

PART C will concern PRMT5 interactions with three adaptor proteins: RioK1, pICln and COPR5. A common adaptor protein binding motif will be identified based on the protein sequence information and biophysical investigations utilising different PRMT5 protein constructs. The found motif will be characterised structurally in order to provide a foundation for future peptidomimetic PPI inhibitor development.

PART D will build on the information gathered in PART C, and the main objective will be development and optimisation of an effective PRMT5/adaptor protein PPI inhibitor. Various macrocyclic structures will be constructed and tested. The afforded compounds will be assessed using biophysical assays, and if a working inhibitor is successfully obtained, the compound will be further evaluated in biologically relevant experiments.

A novel PRMT5 PPI modulator will be potentially an invaluable tool for biochemistry and biology, applicable to a broad range of investigations concerning the PRMT5-MEP50 complex.

PART A: Modulating Interactions Between PRMT5 and MEP50

3.1. Brief Introduction to PART A

MEP50 is the most crucial binding partner of human PRMT5, significantly influencing key PRMT5 properties. Antonyamy *et al.*, 2012, reported that the full heterooctameric PRMT5-MEP50 complex has a significantly higher methylation activity towards histone peptides than the enzyme devoid of MEP50. The most pronounced effect of the partnering at MEP50 on the activity is observed for the ability of the enzyme to dimethylate histone peptides. Substrate dimethylation by isolated PRMT5 is particularly compromised in comparison to the performance of the full complex. The absence of MEP50 also appears to impact the capacity of PRMT5 to form the core homotetrameric assembly, as the isolated enzyme seems to only homodimerise. This indicates an important allosteric influence of MEP50 on the behaviour of PRMT5. As also noted in section 1.3.3. of the main introduction, the absence of MEP50 negatively impacts the stability of the isolated methyltransferase, promoting enzyme aggregation.^[370] Burgos *et al.*, reached similar conclusions^[370] based on their experimental results, which indicated the direct interaction of histones with MEP50 and active site remodelling effect in the complex caused by the WD40 repeat protein, affecting the PRMT5 methylation activity. Interestingly, a dose-dependent effect on the enzymatic methylation of H4 peptide was observed when PRMT5 was tested with different stoichiometric ratios of MEP50. The highest activity was observed for the fully assembled complex with the 4:4 PRMT5/MEP50 ratio.^[386]

An oncogenic mutation of Janus kinase 2 (JAK2) enhances its interaction with PRMT5, leading to the phosphorylation of Tyr residues in the Rossmann fold at positions 297, 304 and 307. The wild type of JAK2 has no effect on the enzyme. The kinase activity of oncogenic mutant JAK2 appears to disrupt the PRMT5-MEP50 complex *in vivo*. The treatment with JAK Inhibitor I was shown to block the phosphorylation of PRMT5 and rescue the interaction between the methyltransferase and MEP50, indicating that it is the kinase activity but not the PPI with JAK2 that influences complex formation.^[387] The phosphorylation by mutated JAK2 was also shown to impair the histone methyltransferase activity of PRMT5, which is consistent with the previous observations made for PRMT5 in the absence of MEP50 by Antonyamy *et al.*, 2012 and Burgos *et al.*, 2015.^[370,386,387]

The presence of MEP50 in the methyltransferase complex is crucial for its full activity, at least in the context of the histone methylation. MEP50 may be a necessary enzyme partner due to the potential effect on the conformation of the PRMT5 active site, stabilisation of PRMT5 or a role as a substrate presenter to the catalytic site. The understanding of the exact influence of MEP50 on the PRMT5

structure and biochemistry still remains largely incomplete, and considering the key importance of histone methylation for the epigenetic regulation in cells, it appears that further attempts to investigate the PPI between MEP50 and PRMT5 are needed.

The aim of the study in PART A is to develop a molecular probe capable of inhibiting the PPI between MEP50 and PRMT5. Taking into consideration all the previously discussed difficulties associated with a discovery of small molecule PPI inhibitors, and all the benefits associated with macrocyclic peptides in the context of challenging PPI modulation, an attempt is made to synthesise a set of peptides and peptidomimetics with a potential to influence the PRMT5-MEP50 interaction. The compounds will be designed based on the already existing structural protein elements of the interface. If successful, the afforded inhibitor could be used as an invaluable tool for a closer examination of the *in vivo* activity and the biochemical regulatory mechanisms of PRMT5 based on the association with MEP50. The potential peptidomimetic PPI inhibitor could also be the first step in a development towards the first clinically applicable PRMT5-MEP50 PPI modulator.

3.2. Results and Discussion

3.2.1. Exploring the PRMT5-MEP50 PPI Interface

In order to determine which regions of the interface between PRMT5 and MEP50 are the most crucial for the PPI, a virtual Ala scan was performed on the interacting residues. The scan was conducted using the DrugScore^{PPI} webserver service which can estimate changes in the binding free energy between proteins upon Ala mutations, utilising a knowledge-based function. DrugScore^{PPI} is intended

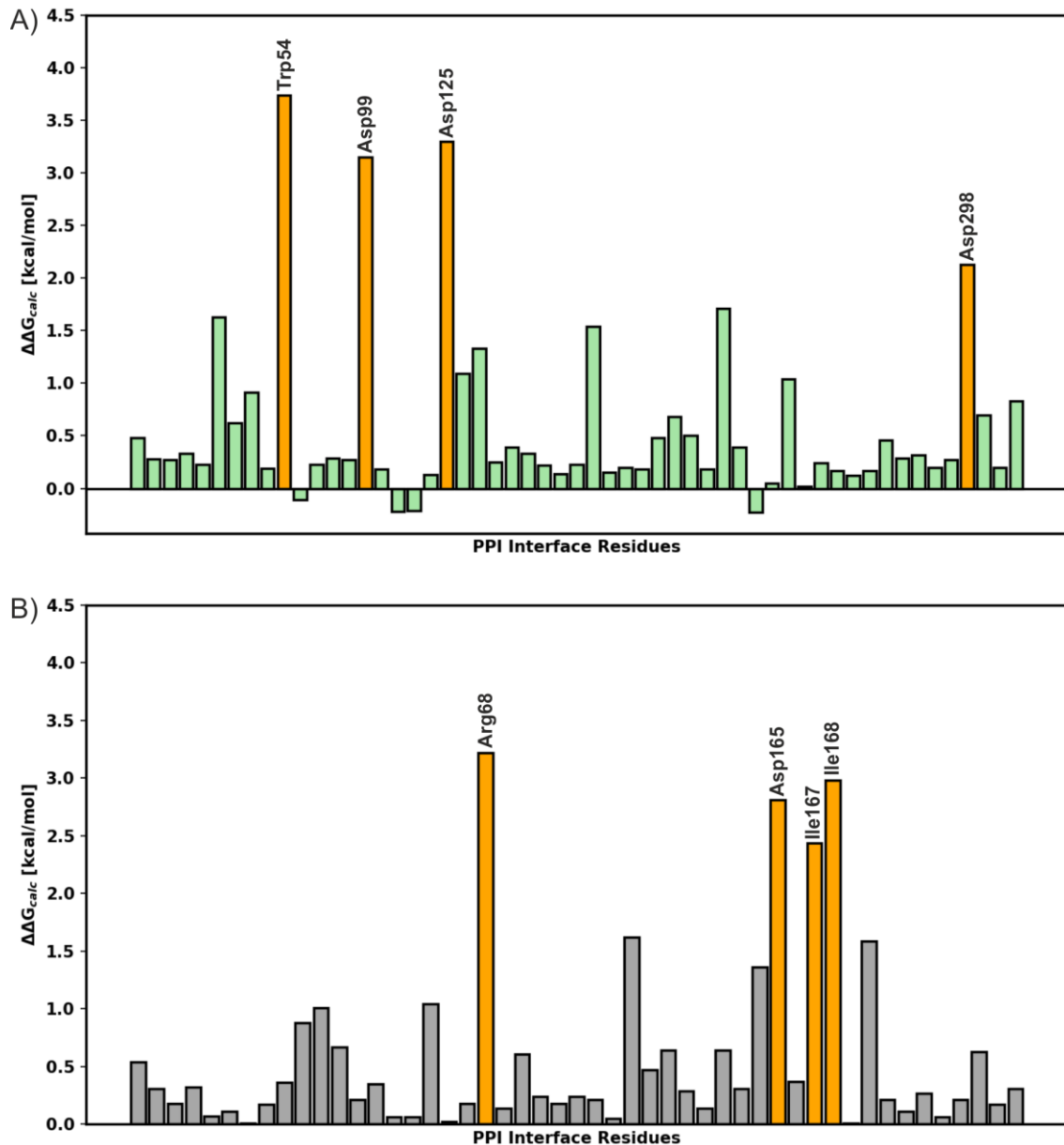


Figure 17. Overview of the virtual Ala scan performed with DrugScore^{PPI} at the PPI interface between the PRMT5 TIM barrel domain and MEP50: **A)** Ala scan of the MEP50 interface residues. Hot residues ($\Delta\Delta G_{calc} \geq 2.0$ kcal/mol) are marked as orange. **B)** Ala scan of the TIM barrel interface residues. Hot residues ($\Delta\Delta G_{calc} \geq 2.0$ kcal/mol) are marked as orange.

for finding hot-spot residues.^[388] The crystal structure provided for the calculations was the human PRMT5-MEP50 structure obtained by Antonyamy *et al.* (PDB ID: 4GQB). The virtual mutations were performed for the surface residues of both PRMT5 and MEP50. The scan returned four hot-spots ($\Delta\Delta G_{\text{calc}} \geq 2.0$ kcal/mol) at the surface of MEP50: Trp54, Asp99, Asp125 and Asp298, and four hot-spots in the TIM barrel of PRMT5: Arg68, Asp165, Ile167 and Ile168 (Figure 17, Table S1 and S2).

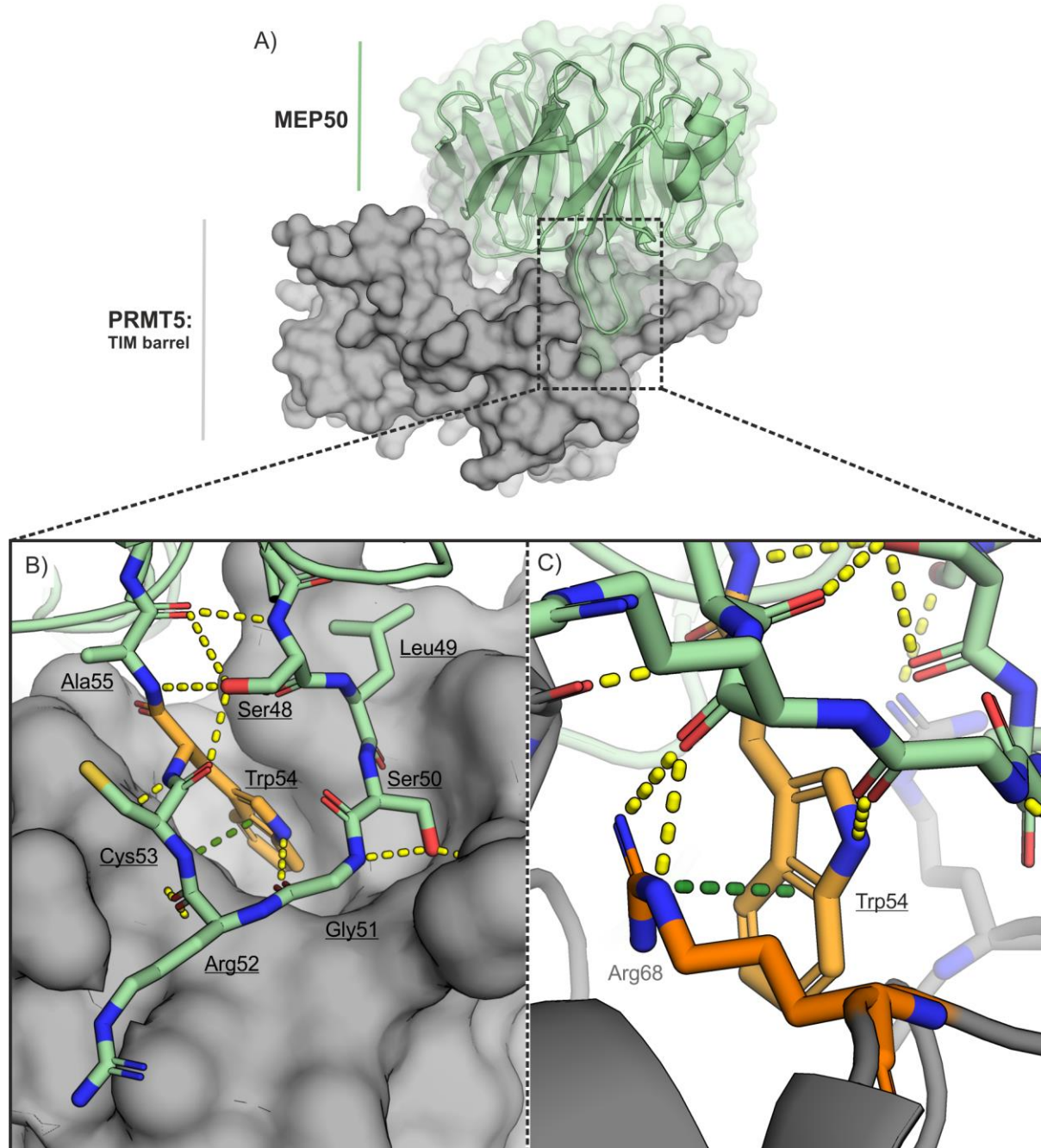


Figure 18. Structure and interactions of the MEP50 insertion finger (PDB ID: 4GQB): **A)** Overview of the MEP50-TIM barrel interaction. **B)** Expanded view of the insertion finger. **C)** Close-up of the interactions between Trp54 (MEP50) and Arg68 (PRMT5). Hot-spot residues ($\Delta\Delta G_{\text{calc}} \geq 2.0$ kcal/mol) as determined with DrugScore^{PP1}, are marked as orange. H-bonds and π -cation interactions are visualised with yellow and green dashed lines, respectively.

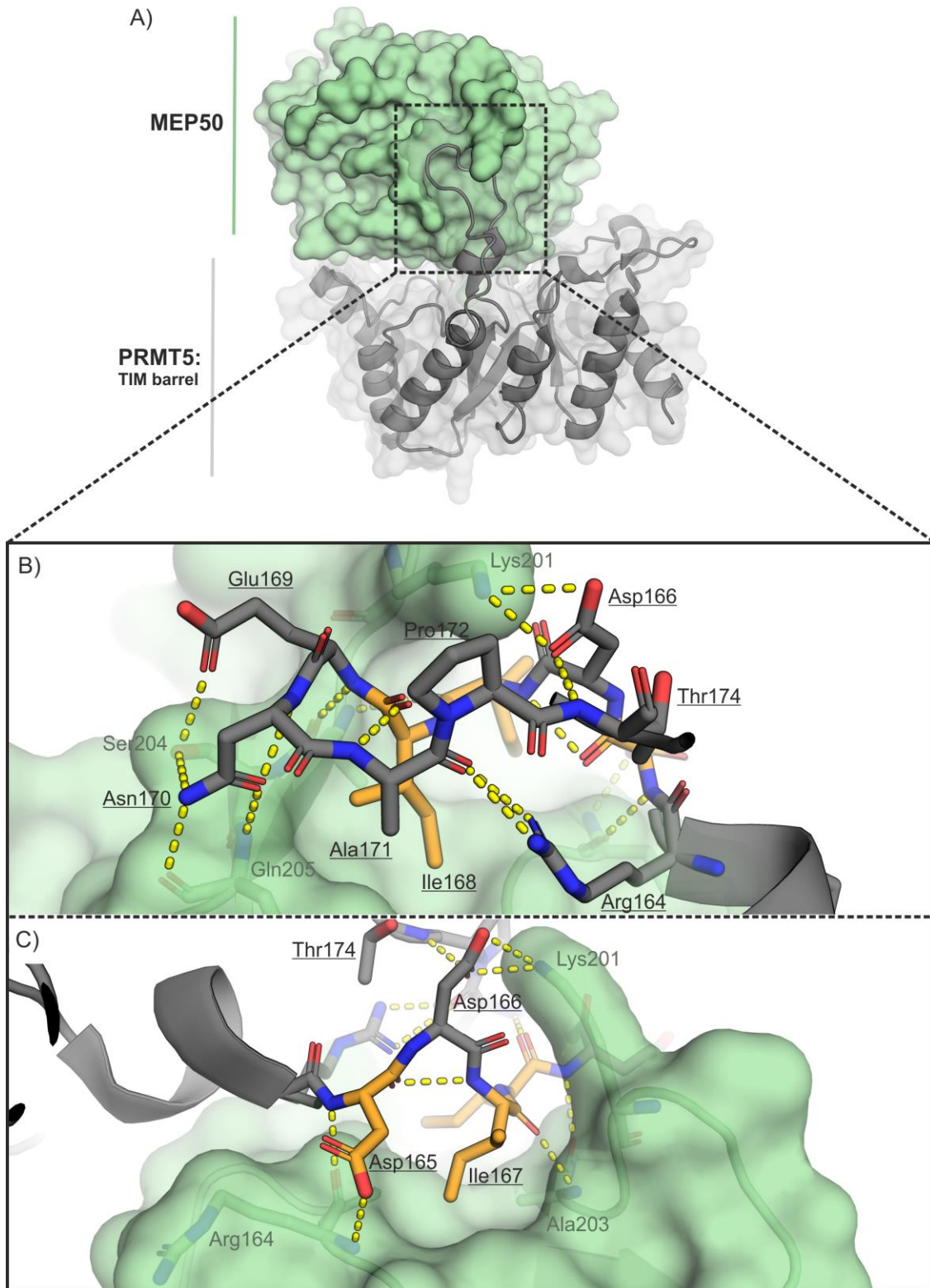


Figure 19. Structure and interactions of the TIM barrel protruding loop (PDB ID: 4GQB): **A)** Overview of the MEP50-TIM barrel interaction. **B)** and **C)** Expanded views of the TIM barrel loop shown from different perspectives. Hot-spot residues ($\Delta\Delta G_{\text{calc}} \geq 2.0$ kcal/mol) as determined with DrugScore^{PPI}, are marked as orange. H-bonds interactions are visualised with yellow and green dashed lines, respectively. H-bond interactions are visualised with yellow dashed lines.

A closer analysis of the location of the hot-spot MEP50 residues revealed that Asp99, Asp125 and Asp298 are located in relatively flat and featureless regions of the interface. Residue Trp54 on the other hand is integrated into a distinctive protein loop with a relatively large interaction area with PRMT5 (Figure 18). This protruding and largely solvent exposed loop is referred to as the MEP50 insertion finger. The insertion finger is known to be an important element of the complex, as mutations within its sequence can lead to the impairment of kinetic parameters of binding histones and SAM, indicating its significant role in the methyltransferase activity and suggesting the involvement of this loop in substrate binding.^[386] The insertion finger is connected to PRMT5 through a π -cation interaction between the hot-residue Trp54 (MEP50) and Arg68 (PRMT5), as well as through H-bond interactions between Ser50, Arg52 and Cys53 in the MEP50 loop and the residues of PRMT5. The conformation of the insertion finger appears to be further stabilised by the extensive intra-chain H-bond interactions formed between Ser48 and backbones of Ala55 and Cys53. There is also an H-bond interaction present between the indole nitrogen of Trp54 and the carbonyl oxygen of Gly51 (Figure 18B and C). Considering the interactions, the conformation, the location and the important role of the insertion finger, this structure seems to be potentially favourable for targeting and mimicking with peptide-based probes.

Three out of four predicted hot-spot residues (Asp165, Ile167 and Ile168) of the TIM barrel domain are located in a distinctive and also considerably solvent exposed loop placed in a shallow groove of MEP50 (Figure 19). The loop is characterised by a very extensive H-bond network of both intra- and inter-chain connections. Residues Asp165, Asp166, Ile167, Glu169 and Asn170 of the PRMT5 loop are involved in H-bond interactions with the surface residues of MEP50. Additionally, the loop conformation is shaped by a number of the intra-chain H-bonds: Arg164-Ala171, Asp165-Ile167, Asp166-Thr173 and Ile168-Ala171 (Figure 19B and C). Taking into consideration the high density of the hot-spot residues and the inter-chain H-bonds, as well as the location and the conformation of the TIM barrel loop, this structure also appears to be viable for mimicking with peptidomimetics, presenting a potential attractive region for targeting in the context of the PRMT5-MEP50 PPI inhibition.

Both the insertion finger of MEP50 and the protruding loop of the TIM barrel domain appear to have conformations viable for macrocyclisation, and could be potentially mimicked by appropriate cyclic peptides.

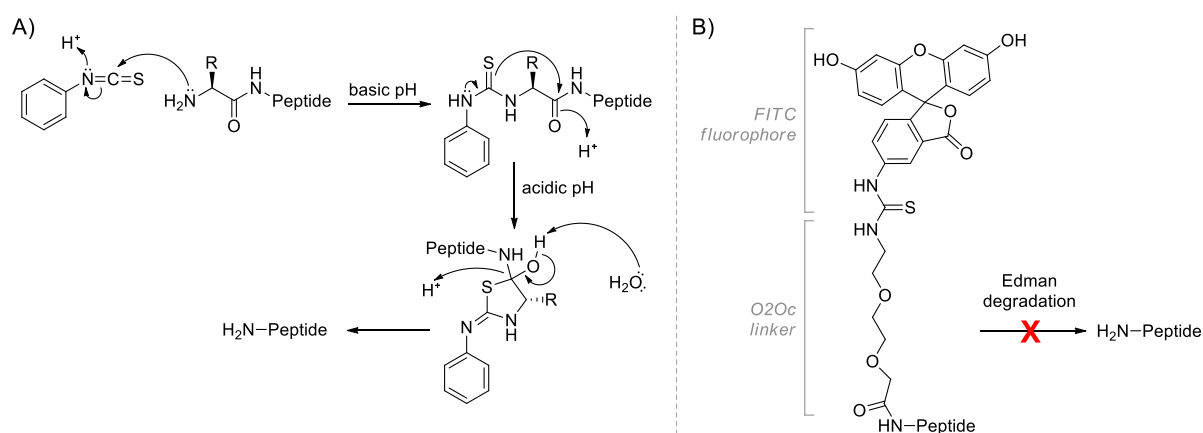
3.2.2. MEP50 Insertion Finger Mimetics

The synthesis of the MEP50 insertion finger mimetics was started by obtaining linear peptidic sequences of different lengths (7 to 12 residues), where the number of the terminal amino acids was varied around the common core sequence of the loop encompassing residues Ser48 through Trp54: SLSGRCW (Table 1). The linear peptides were obtained using the standard, automated and microwave assisted Fmoc-based synthesis on rink amide resin, resulting in C-terminal amides. The peptides were synthesised in two versions, either as N-terminally acetylated sequences (**1-5**) intended for use in activity-based readouts with the PRMT5-MEP50 complex, or as fluorescently labelled sequences (**6-10**) intended to be tested for the ability to directly interact with PRMT5. The labelled peptides are tagged with fluorescein isothiocyanate (FITC) which is separated from the peptidic chain with a PEG-based 8-amino-3,6-dioxaoctanoic linker called O2Oc. The presence of the linker is intended to prevent a direct steric interference of FITC with the target protein and to avoid Edman degradation that would result in the elimination of the fluorophore together with the N-terminal amino acid from the sequence during the final deprotection stage of the SPPS with TFA (Scheme 1).^[389]

Table 1. Sequences of linear acetylated and fluorescently labelled peptides, derived from the MEP50 insertion finger.

Peptide	Sequence
1	Ac-SLSGRCW-NH ₂
2	Ac-SLSGRCWA-NH ₂
3	Ac-SLSGRCWAG-NH ₂
4	Ac-SLSGRCWAGS-NH ₂
5	Ac-ASLSGRCWAGS-NH ₂
6	FITC-O2Oc-SLSGRCW-NH ₂
7	FITC-O2Oc-SLSGRCWA-NH ₂
8	FITC-O2Oc-SLSGRCWAG-NH ₂
9	FITC-O2Oc-SLSGRCWAGS-NH ₂
10	FITC-O2Oc-ASLSGRCWAGS-NH ₂

FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker.



Scheme 1. Edman degradation: **A)** Example of Edman degradation resulting in a loss of the N-terminal amino acid. **B)** The presence of the O2Oc linker prevents the loss of the FITC fluorophore via Edman degradation.

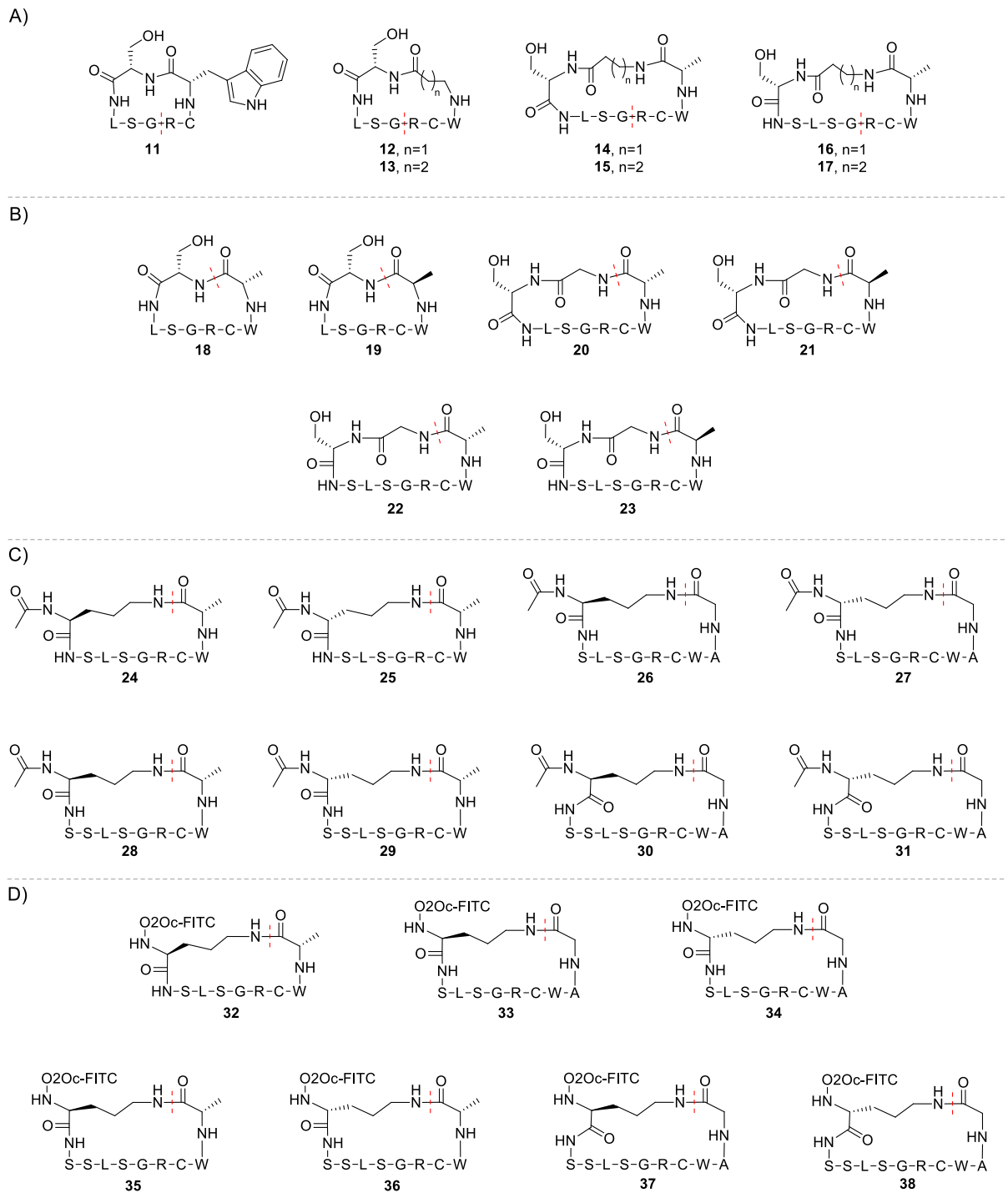
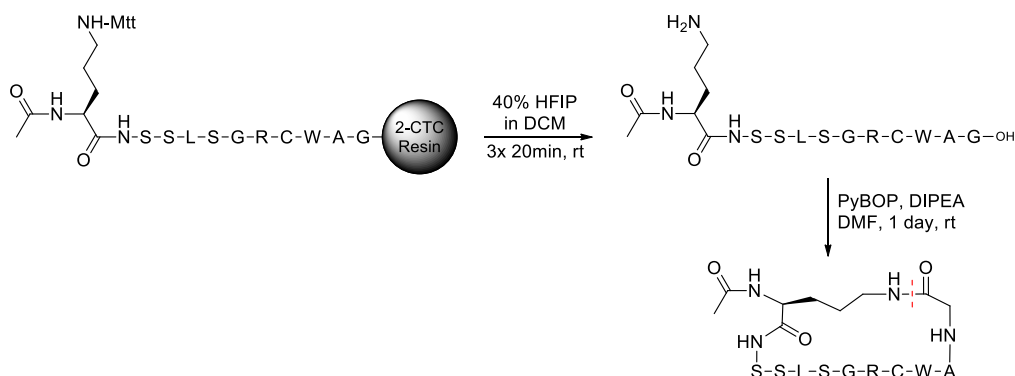


Figure 20. Structures of cyclic peptides intended to mimic the insertion finger of MEP50: **A)** Head-to-tail cyclised peptides, with cyclisation between the C-terminal carboxylic acid of Gly and N-terminal amine of Arg. **B)** Head-to-tail cyclised peptides, with cyclisation between the C-terminal carboxylic acid of L/D-Ala and N-terminal amine of Ser or Gly. **C)** Side chain-to-tail cyclised peptides with cyclisation between the side chain of Orn and carboxylic acid of Ala or Gly. **D)** Peptides labelled with a fluorescent tag, cyclised in a side chain-to-tail manner between the side chain of Orn and the peptidic C-terminus. The red dashed line between amino acids indicates the cyclisation site. FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker.

The successful synthesis of the linear sequences was followed by a pursuit of macrocyclic peptides intended to mimic the insertion loop of MEP50, based on the assumption that cyclopeptides could afford better target affinities and higher stability in envisaged biophysical and biological assays.^[49,97,126,127,189] Peptides with several different ring sizes were designed and synthesised ranging from 21 atoms in the ring structure for the smallest cyclic compound **11** to the maximum of 36 ring atoms for compounds **30**, **31**, **37** and **38** (Figure 20). Peptides were connected either in a head-to-tail fashion (**11-23**) or through side chain-to-tail connections (**24-38**). The side chain-to-tail peptides were also obtained with fluorescent tags (**32-38**), where FITC is connected to the head of the peptide through the O2Oc linker. The smallest cyclopeptide **11** is linked through an amide bond between Ser48 and Trp54, and the compounds with larger rings incorporate either a higher number of amino acids from the MEP50 insertion loop sequence or utilise non-proteinogenic α -, β - and γ -amino acids, as well as amino acids in D-configuration. Such a range of linkers was explored with the intention that some of the synthesised compounds would be able to assume a similar conformation to the native MEP50 insertion finger in aqueous solution, without paying a significant entropic penalty.



Scheme 2. Synthetic cyclisation strategy exemplified for peptide **30**. The strategy yields side chain-to-tail cyclised peptides. 2-CTC = 2-chlorotrityl chloride.

All linear peptidic sequences used for cyclisation were synthesised on 2-chlorotrityl chloride (2-CTC) resin, allowing for cleavage of the sequences under mild acidic conditions without the removal of the standard side chain protective groups, thus preventing side-reactions during the following cyclisation step. The desired linear peptides were cleaved from 2-CTC resin using solutions of HFIP (20-40%) in DCM, resulting in sequences with C-terminal carboxylic acids. In case of the side chain-to-tail cyclopeptides, the linking side chain of ornithine (Orn) was initially protected with acid labile Mtt group that was removed at the stage of the resin cleavage with the HFIP solution. The cyclisation connecting a free amine and carboxylic acid was then performed using PyBOP and DIPEA at the maximum peptide concentration of 1 mM in DMF (Scheme 2). The relatively high peptide dilution

allows to minimise the undesired formation of dimeric and oligomeric forms of the macrocycles. The ring size of seven or more amino acids is generally conducive to macrocyclisation and according to expectations, peptides **11-38** were successfully cyclised without any further complications.^[187]

3.2.3. TIM Barrel Loop Mimetics

The synthesis of the TIM barrel loop mimetics commenced by obtaining N-terminally acetylated linear peptides with a common sequence of seven amino acids spanning from Arg164 through Asn170: RDDIEN, and a varied number of flanking residues at the N- and C-terminus (**39-45**; Table 2). All the linear peptides contain in their sequence three amino acids identified previously as hot-spot residues: Asp165, Ile167 and Ile168.

Table 2. Sequences of linear acetylated peptides, derived from the TIM barrel loop interacting with MEP50.

Peptide	Sequence
39	Ac-RDDIEN-NH ₂
40	Ac-RDDIENA-NH ₂
41	Ac-IRDDIENA-NH ₂
42	Ac-RDDIENAP-NH ₂
43	Ac-IRDDIENAP-NH ₂
44	Ac-RDDIENAPT-NH ₂
45	Ac-IRDDIENAPT-NH ₂

Due to the intricate conformation of the loop and the arrangements of the side chains, presenting an extensive intra- and intermolecular H-bond network, four different macrocyclisation approaches were devised resulting in a wide range of ring sizes spanning between 30 and 43 atoms (Figure 21). The first design attempts to connect the side chain of Arg164 with the C-terminus of Asn170. To facilitate such connection the N-terminal Arg is replaced by either Lys or Orn, whose side chain is then linked to Asn170 through hydrocarbon chains of varied length using amide bonds for attachment. This strategy afforded compounds **46-51** with the ring sizes between 30 and 33 atoms (Figure 21).

The linear sequences of **46-51** were obtained on 2-CTC resin through the means of the standard Fmoc-based SPPS, where the desired hydrocarbon connector was directly attached to the resin linker and the N-terminal Lys or Orn was side chain-protected with the Mtt group. The steps towards the cyclic structures were analogous to the synthesis of **32-38**. The removal of the peptide from the resin with a concomitant cleavage of the Mtt group was achieved with 40% HFIP in DCM. The following cyclisation was performed in solution under high dilution (maximum peptide concentration of 1 mM) using PyBOP and DIPEA in DMF (Scheme 3).

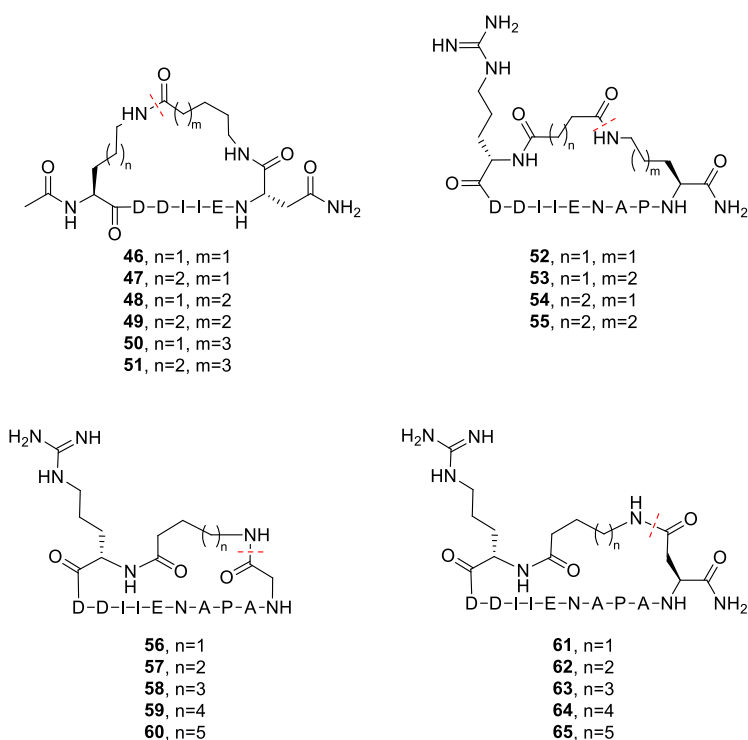
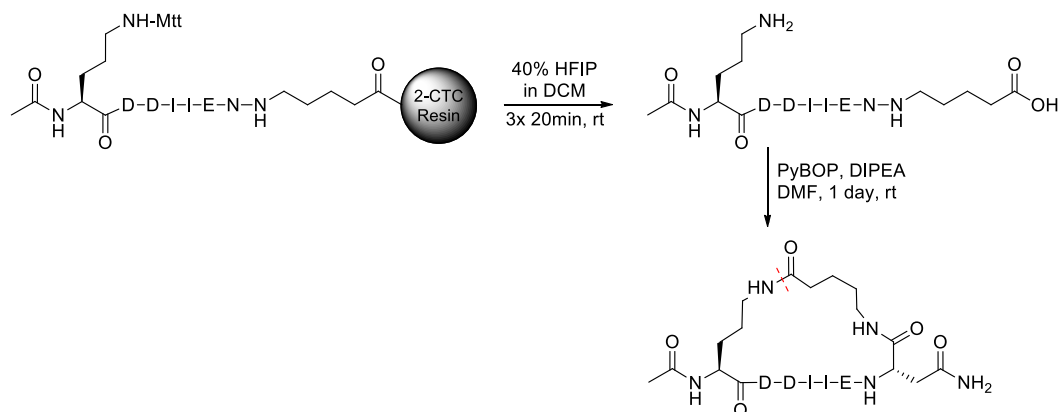


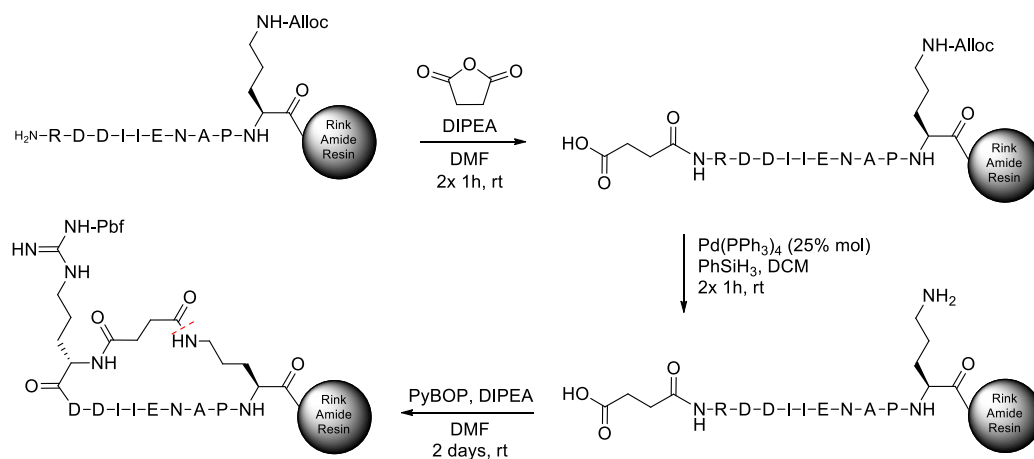
Figure 21. Structures of cyclic peptides intended to mimic the protruding loop of the TIM barrel domain interacting with MEP50. The red dashed line indicates the cyclisation site.



Scheme 3. Synthetic cyclisation strategy exemplified for peptide **46**. The strategy yields side chain-to-tail cyclised peptides. 2-CTC = 2-chlorotryl chloride.

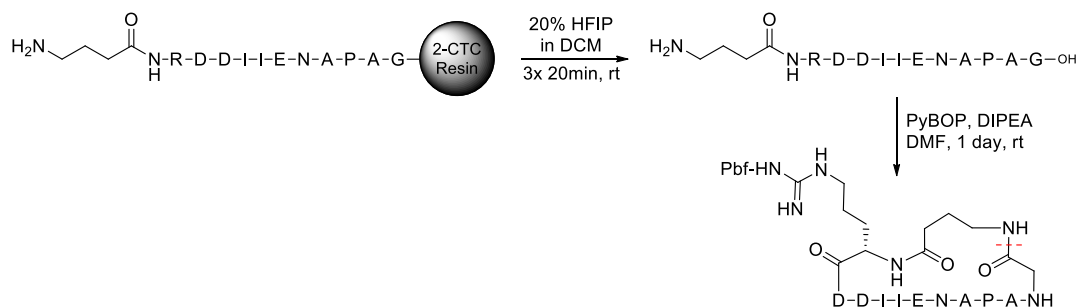
The second cyclisation design attempts to connect the termini of Arg164 and Pro172 through a linker. In this approach Lys or Orn is coupled to the C-terminal Pro172 and its side chain is tethered to the N-terminus of Arg164 via a hydrocarbon linker of varied length. The compounds that resulted from this strategy are **52-55** with a ring size between 37 and 39 atoms (Figure 21). The compounds were prepared on low substitution rink amide resin with Alloc protected C-terminal Lys or Orn. Following the SPPS of the linear sequences, the N-terminal amine of Arg164 underwent a reaction with a cyclic

anhydride of the desired length, affording a linker with a free carboxylic acid intended for a further amide coupling. The Alloc group was removed from Lys or Orn using $\text{Pd}(\text{PPh}_3)_4$ in the presence of PhSiH_3 in anhydrous DCM. The resulting free amine was then coupled intramolecularly with the free carboxylic acid, affording resin-bound and side chain protected cyclic peptides **52-55** (Scheme 4). The use of the low substitution resin allowed to prevent the undesired cross-coupling reactions between separate peptide chains.



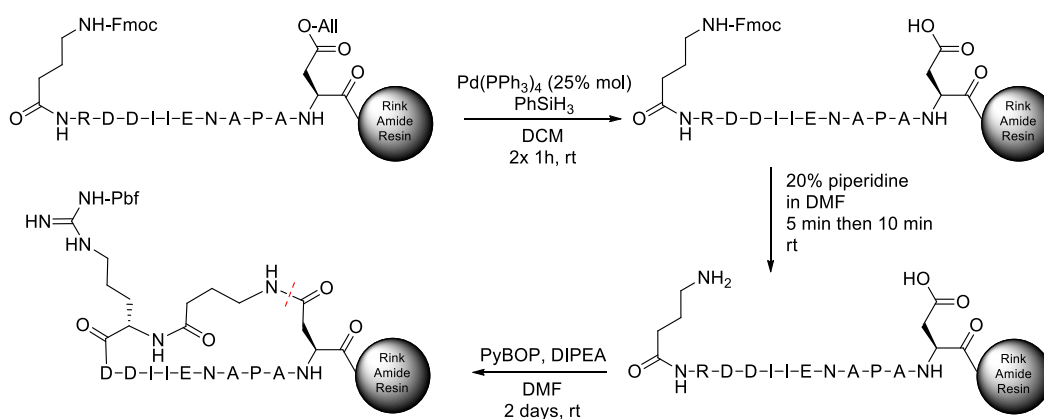
Scheme 4. Synthetic cyclisation strategy exemplified for peptide **52**. The strategy affords head-to-side chain cyclised peptides.

The third approach envisaged larger macrocycles where Thr174 is replaced by Ala to prevent unnecessary, non-specific interactions of the Thr side chain. The connection is made between the N-terminus of Arg164 and the C-terminal carboxylic acid of the residue at position 174, using a hydrocarbon chain and Gly as a linker. This design afforded cyclopeptides **56-60** with the ring sizes between 38 and 42 atoms (Figure 21). The linear precursors were synthesised on 2-CTC resin with Gly at the C-terminus and the desired linker on the N-terminal end of the peptide. The peptides were removed from the resin through the treatment with 20% HFIP in DCM, which was followed by cyclisation in solution (Scheme 5).



Scheme 5. Synthetic cyclisation strategy exemplified for peptide **56**. The strategy affords head-to-tail cyclised peptides. 2-CTC = 2-chlorotrityl chloride.

The final macrocyclisation design is analogous to the structures of **56-60**, however, Gly is replaced by Asp, whose sidechain is then used for completing the ring of 39 to 43 atoms. Compounds **61-65** are based on this design (Figure 21). The linear sequences of **61-65** were obtained through the SPPS synthesis on low substitution rink amide resin with C-terminal Asp protected as allyl ester and N-terminally attached linker with its amine protected by the Fmoc group. The allyl ester on the Asp side chain was then removed through the treatment with $\text{Pd}(\text{PPh}_3)_4$, which was followed by the removal of the Fmoc group, exposing a free amine. The subsequent macrocyclisation via amide bond formation was performed on resin yielding side chain protected cyclopeptides (Scheme 6).



Scheme 6. Synthetic cyclisation strategy exemplified for peptide **61**. The strategy affords head-to-side chain cyclised peptides.

3.2.4. Protein Expression

In order to test the ability of the synthesised peptidomimetics to interact with PRMT5 and MEP50 in biophysical and enzymatic assays, isolated proteins were required.

All proteins used for the experiments described in this work were expressed and purified by the team at the Protein Chemistry Facility at the Max Planck Institute of Molecular Physiology in Dortmund. Attempts were made to express the human heterooctameric PRMT5-MEP50 complex or its isolated components in *Escherichia coli*, however, none of the expression strategies tested by the facility afforded the desired constructs. The successful co-expression of PRMT5-MEP50 was only achieved using HighFive insect cells, where PRMT5 was labelled with a cleavable N-terminal 6His-MBP tag and MEP50 was tagged with removable N-terminal 6His-TRX. The tags were removed from the proteins at the purification stage, using HRV-3C PreScission protease. The human complex was obtained with only short N-terminal Gly-Pro overhangs attached to the otherwise native protein sequences. Attempts to

express and isolate either PRMT5 or MEP50 alone were unsuccessful, testing all standard tags and expression conditions offered by the protein facility (Table S3).

The full length PRMT5 protein can oligomerise, with a significant role in this process performed by the catalytic domain. The catalytic domain contains the interchain salt-bridge formed by Arg488 and Asp491 contributing to the complex dimerisation (PRMT5₂), as well as Ser321, Arg368, Thr400, Asp531 and Trp603 interacting with the residues of the TIM barrel domain involved in the assembly of the homotetramer (PRMT5₄). Based on the distribution of the interchain interactions across the PRMT5 domains and the spacial arrangements of the domains from different PRMT5 units comprising the full complex, it was inferred that protein constructs with the truncated catalytic domains (the β -barrel and Rossmann fold) would lose the ability to form PRMT5 homodimers and homotetramers.^[370–372] Such constructs were deemed desirable as they could potentially offer the ability to inhibit the interaction between the full-length PRMT5 and MEP50 proteins, or allow to determine in more unambiguous manner the specific interaction domain of the synthesised peptides and peptidomimetics intended to bind to the PRMT5-MEP50 complex.

A truncated version of the human PRMT5 protein corresponding to the TIM barrel domain (1-292 aa) was expressed in *E. coli* and successfully isolated. The protein was expressed with a SUMO-tag, which was cleaved with SUMO protease at the purification stage for a portion of the expressed protein. The isolated domain proved to be buffer soluble in the absence of MEP50, either in the presence or absence of the SUMO tag. The protein was also co-expressed together with MEP50 (both proteins initially labelled with the N-terminal 6His-TRX tag), giving an untagged TIM-MEP50 heterodimer after the protein purification. The heterodimer could only be expressed in the HighFive insect cells.

The molecular weight of the expressed proteins was evaluated using mass photometry. Mass photometry detects particle binding/unbinding events at the interface between the solution and a solid glass surface, based on the associated light scattering effect characterised by the appropriate contrast that can be detected and counted by the instrument. As different particle sizes are expected to give different contrasts in mass photometry, the distribution of the particle sizes in solution can be estimated without a need to interfere with the sample and to modify or label the particles.^[390,391] The analysis of the expressed full length PRMT5-MEP50 construct returned a mass estimation of 407 ± 22 kDa which approximately corresponds to the heterooctameric assembly (PRMT5₄-MEP50₄, calculated mass of 438 kDa) or alternatively to the complex without one MEP50 unit (PRMT5₄-MEP50₃, calculated mass of 401 kDa; Figure 22A). The exact character of the complex is difficult to determine due to the nature of the technique. The isolated TIM-MEP50 (calculated mass of 70 kDa) afforded the mass estimation of 68 ± 15 kDa, corroborating the initial hypothesis, that the truncation of the catalytic domains would result in the inability to form the oligomers (Figure 22B).

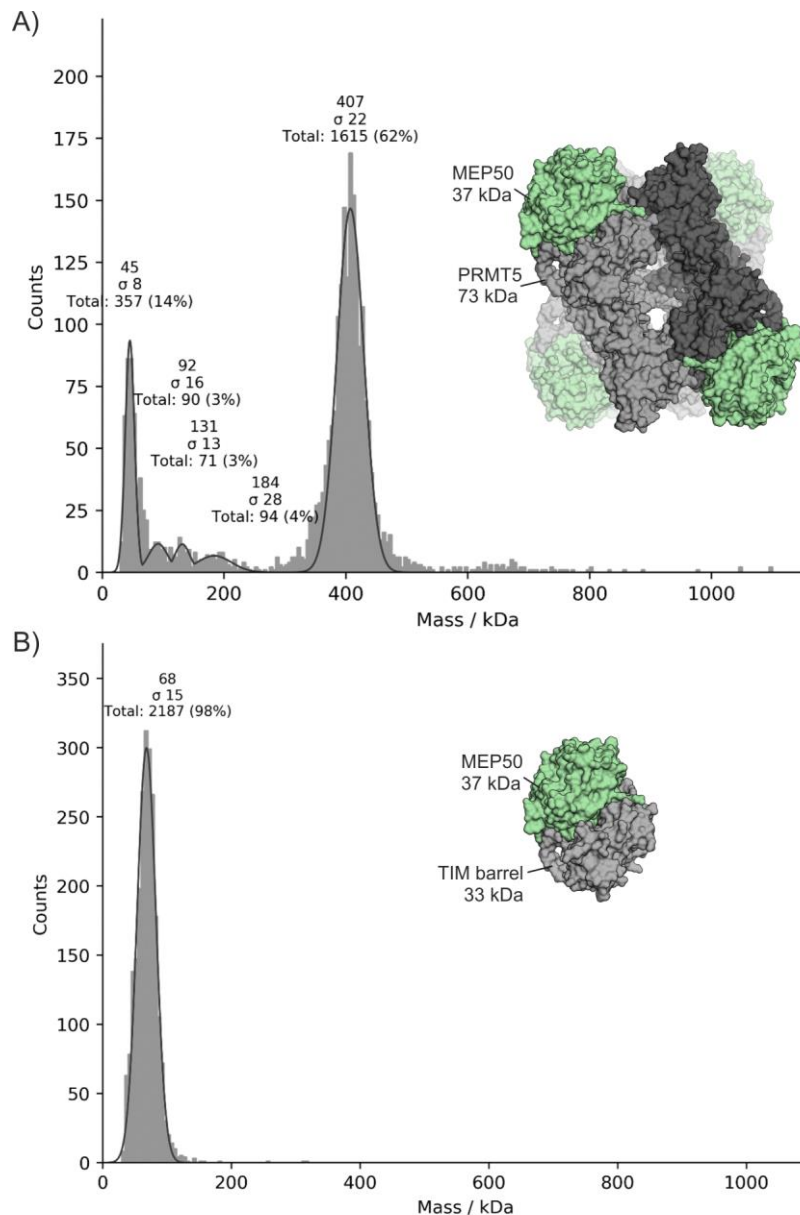


Figure 22. Results of the mass photometry measurements of the expressed proteins: **A)** Results for full-length human PRMT5-MEP50 complex. **B)** Results for the truncated complex, TIM-MEP50. The protein complex structures are presented based on the crystal structure 4GQB.

3.2.5. Compound Evaluation

As the presence of MEP50 in the methyltransferase complex is known to be important for the methylation activity of the PRMT5 enzyme, the synthesised peptidomimetics intended to disrupt the interaction between MEP50 and PRMT5 were tested for their ability to impact the methylation rate of an H4 histone tail peptide (Ac-SGRGKGGKGLGKGGAKRHRKV-NH₂).^[370,386,387] The assay used to this end was MTase-Glo™, which detects the SAH by-product produced as the result of the substrate methylation (Figure 13, section 1.3.1.). The presence of SAH is translated into ADP by the assay

reagent, which is subsequently converted into luminescent light using the MTase-Glo™ detection solution containing luciferase and luciferin. The amount of the produced light is proportional to the SAH concentration generated by PRMT5.^[392] The inhibition of the PRMT5 enzymatic activity, thus, can be directly detected with the assay. A small compound which is an active site inhibitor of PRMT5 called EPZ015666 was used as a positive control in the experiment (Figure 23A).^[393] All peptidomimetics of the MEP50 insertion finger (**1-5** and **11-31**) were soluble in the assay buffer, however, ten of the synthesised PRMT5 loop mimetics (**41, 48, 49, 54, 58-60, 62, 64** and **65**) formed gel-like aggregates following a preparation of peptide solutions in the required buffer, and thus, they were excluded from

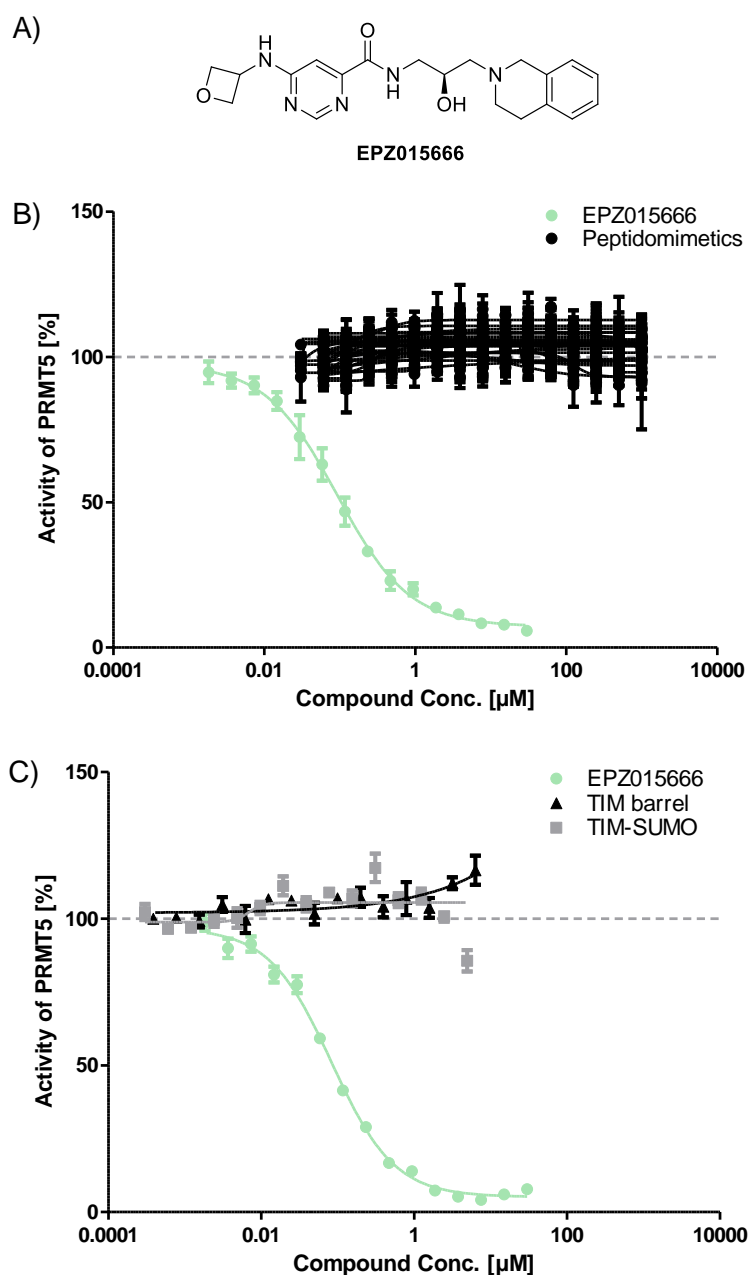


Figure 23. Methyltransferase activity analysis: **A)** Structure of the commercial PRMT5 active site inhibitor. **B)** MTase-Glo™ assay with EPZ015666 and the buffer soluble, unlabelled peptidomimetics. **C)** MTase-Glo™ assay with EPZ015666, TIM barrel domain and TIM-SUMO protein construct.

further analysis. None of the tested mimetics of either the insertion finger of MEP50 or the PRMT5 loop had a detectable effect on the methyltransferase activity of the full-length human PRMT5-MEP50 complex, despite observing a strong enzyme inhibition by EPZ015666 (IC_{50} of 93 ± 40 nM; Figure 23B). As a control experiment, the compounds were also tested under the assay conditions without the enzyme present, in order to ensure that they do not induce luminescence. The enzymatic activity assay proved that the expressed and isolated PRMT5-MEP50 complex is catalytically active.

The assessment of the peptidomimetics in the MTase-GloTM assay was followed by testing the influence of the isolated TIM barrel domain and the TIM-SUMO protein construct on the activity of the PRMT5-MEP50 complex. In a similar manner to the peptidomimetics, the exposure of the methyltransferase complex to the TIM barrel proteins had no detectable effect on the enzyme activity (Figure 23C).

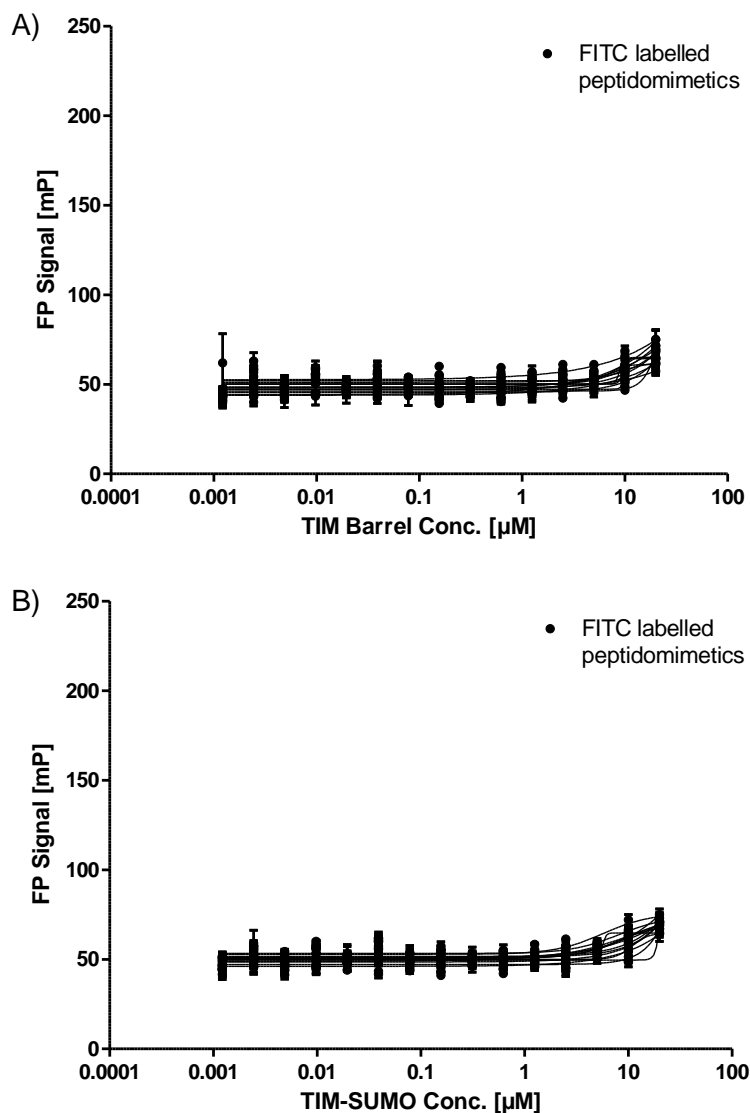


Figure 24. FP analysis of the fluorescently labelled peptidomimetics (**6-10** and **32-38**): **A)** FP analysis for peptide binding to the unlabelled TIM barrel domain. **B)** FP analysis for binding to the SUMO tagged TIM barrel domain.

The observed results suggest that the synthesised linear and cyclised peptides, as well as the TIM barrel constructs are unable to disrupt the PRMT5-MEP50 complex. This could be caused by inappropriate cyclisation of the peptidic sequences and misfolded interaction face of the expressed TIM barrel proteins, or by the very strong interactions between the protomers in the presumably obligate PRMT5-MEP50 complex. Alternatively, the PPI between MEP50 and PRMT5 are partially inhibited but such action has insufficient effect on the enzyme activity. Burgos *et al.* showed that even a loss of two out of four MEP50 proteins from the heterooctameric construct (PRMT5₄-MEP50₄) still allows to conserve over 75% of the original methyltransferase activity.^[386] The tested compounds could thus suppress the interaction of the proteins to a limited extent, but unless the inhibition is complete the impact on the activity might be negligible.

The MTase-GloTM-based experiments were followed by a binding assessment with a fluorescence polarisation (FP) assay. The fundamental principle of the FP assay is based on the detection of the changes in the light polarisation emitted by a fluorophore attached to the tested molecule, where the degree of the polarisation of the emitted light is inversely related to the rotation of the fluorophore. When a fluorophore is attached to a small molecule, the light is largely depolarised, however, when the fluorophore is bound to a larger particle, its speed of rotation is reduced and the degree of the polarisation of the emitted light increases.^[394–396] The fluorescently labelled linear and cyclic mimetics of the MEP50 insertion finger (**6-10** and **32-38**) were tested for the ability to bind directly to the isolated, unlabelled TIM barrel domain (Figure 24A) or to the TIM barrel domain labelled with the SUMO tag (Figure 24B). None of the tested compounds displayed detectable binding activity to the used proteins. The inability of the peptides to bind to the surface of the TIM barrel constructs could have been caused by a misguided peptide design, although this explanation is unlikely due to the inclusion of the linear peptides of various lengths among the tested compounds. Alternatively, the peptides might have inherently very low affinity towards the target, based on the potentially weak native interaction between the peptide chain and the surface of the protein. The surface area on the isolated TIM barrel domains could have also been shaped differently than the PPI surface of PRMT5 in the presence of MEP50, preventing the peptides from the appropriate, native interactions, that are observed in the crystal structure of the full PRMT5-MEP50 complex.

The combined negative results obtained from the MTase-GloTM and FP assays highlight the difficulty of designing PPI inhibitors and protein-binding probes, and suggest that targeting of the interactions between PRMT5 and MEP50 might not be feasible with the selected strategy applied here.

During writing of this doctoral thesis, Asberry *et al.* reported a discovery of the first small molecule inhibitor of the PRMT5-MEP50 interaction.^[397] The compound was developed following the initial hit in a screening campaign involving bimolecular fluorescence complementation-based assay in cells and

virtual screen. The PPI inhibitor is believed to bind to the hydrophobic pocket of the TIM barrel domain that houses the MEP50 insertion finger, displacing the hot-spot residue Trp54, thus, supporting the outlined here initial decision and reasoning for targeting this particular site on PRMT5. The analysis performed by Asberry *et al.* revealed that their inhibitor could impact a number of cellular pathways associated with PRMT5, crucial for the survival and the proliferation of prostate and lung cancer cells. Additionally, the compound appeared to be inhibiting the transforming growth factor- β (TGF- β), protein which can act as either tumour suppressor or promoter, extensively involved in cancer progression.^[397,398] However, biophysical evidence of the PRMT5-MEP50 PPI inhibition is lacking from their report.

3.2.6. Histone Tail Peptides

In order to further explore the potential PPIs of the PRMT5-MEP50 complex, the interaction of the TIM barrel and MEP50 with the human histone tail peptides H2A, H4 and H3 were tested. It was hypothesised that the histone tails could interact with MEP50 in a similar manner to the binding between H3 and WDR5 (a protein with a similar WD40 fold).^[399] MEP50 could therefore potentially act as a PRMT5 substrate recruiter through a direct interaction with the histone peptides. Six fluorescently labelled histone tail peptides were synthesised, placing the FITC label either on the N-terminal (**66-68**) or the C-terminal end of the sequences (**69-71**; Table 3). The variations in the FITC placement were applied in order to minimise the chance of label interference with the protein. The synthesised peptides were tested for binding to the TIM-MEP50 complex, a construct without the catalytic methyltransferase domain. None of the analysed histone tail peptides showed binding activity in the performed FP assay, not indicating interaction between the human histone tails and MEP50 or the TIM barrel (Figure 25). These observations seem to support the interaction and the substrate recognition model between PRMT5-MEP50 and histones proposed by Shechter *et al.*, where MEP50 binds to the substrate region distal to the methylated Arg residue on the histone tails, orienting the tails towards the active site of PRMT5.^[371,386]

Table 3 Sequences of fluorescently labelled H2A, H4 and H3 histone tail peptides.

Peptide	Sequence
66 (H2A)	Fitc-O2Oc-SGRGKQGGKARAKAKTRSSRA-NH ₂
67 (H4)	Fitc-O2Oc-SGRGKGGKGLGKGGAKRHRKV-NH ₂
68 (H3)	Fitc-O2Oc-ARTKQTARKSTGGKAPRKQLATKAARKSA-NH ₂
69 (H2A)	Ac-SGRGKQGGKARAKAKTRSSRA-O2Oc-K (FITC) -NH ₂
70 (H4)	Ac-SGRGKGGKGLGKGGAKRHRKV-O2Oc-K (FITC) -NH ₂
71 (H3)	Ac-ARTKQTARKSTGGKAPRKQLATKAARKSA-O2Oc-K (FITC) -NH ₂

FITC = fluorescein isothiocyanate.

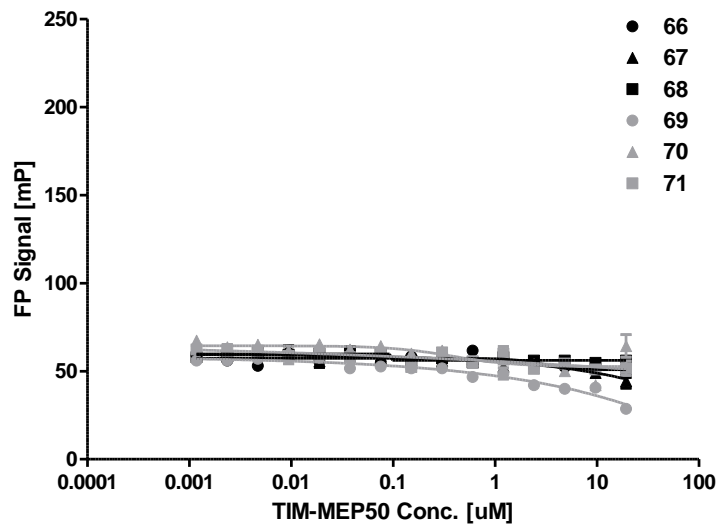


Figure 25. FP analysis of the direct histone tail binding to the TIM-MEP50 protein complex.

3.3. Summary and Conclusions

MEP50 plays an important role in histone methylation by PRMT5, and development of inhibitors of the interaction between these two proteins could afford an invaluable tool for elucidating the enzyme biochemistry and regulatory mechanisms. A working PRMT5-MEP50 PPI inhibitor could also contribute to the development of novel clinically applicable compounds targeting the PRMT5-MEP50 complex.

Analysis of the amino acid contribution to the binding between PRMT5 and MEP50 via a virtual Ala scan suggested two distinctive protein loops as viable structures for mimicking with various linear and cyclic peptides. The identified regions were the MEP50 insertion finger and a protruding loop of PRMT5, both with seemingly extensive contributions to the PPI of interest. A number of linear and cyclic analogues of the loops were synthesised, where the macrocycles were obtained through different in-solution and on-resin cyclisation strategies resulting in compounds with various ring sizes. The afforded peptidomimetics were tested *in vitro* using different protein constructs. The MTase-Glo™ assay measuring the enzymatic activity of PRMT5-MEP50 complex, showed no effect by the peptidomimetics on the methylation rate of the methyltransferase complex. Similar observations were made when the TIM barrel domain and the TIM-SUMO construct were used with the intention to inhibit the PRMT5-MEP50 activity. Fluorescently labelled peptidomimetics based on the sequence of the MEP50 insertion finger were tested in an FP assay for direct binding to the SUMO-tagged and untagged TIM barrel domain. The performed FP assay indicated no binding activity. The observed results have a number of potential explanations. The selected protein sequences may have an intrinsically weak binding affinity insufficient for the PPI inhibition or detection of binding to the tested TIM barrel domains under the used conditions. Additionally, the interaction between PRMT5 and MEP50 might be too strong in comparison to the tested compounds. The negligible effect on the enzyme activity could be caused by the incomplete inhibition of the PPI between PRMT5 and MEP50. It is also possible that the shape of the interfacing surface of the isolated TIM barrel domains significantly different from the interface topology observed in the available crystal structures of the PRMT5-MEP50 complex, preventing the formation of the desired interactions between the tested peptides and proteins. These negative results pointed the attention of the author towards alternative PPIs involving the PRMT5-MEP50 complex.

The removal of the methyltransferase domain from PRMT5 prevents the resulting constructs from forming oligomeric complexes, as proved through mass photometry. The truncated TIM-MEP50 protein construct was used to determine whether MEP50 can interact with human histone tail sequences. No detectable binding was found between the histone tails and TIM-MEP50, supporting

the model of the arrangement between PRMT5-MEP50 and histones, where the body of the histone interacts with MEP50, orienting the tail towards the active site of PRMT5.

PART B: Targeting Potential Interactions Between SUZ12 and MEP50

4.1. Brief Introduction to PART B

In 2006, while performing a set of pull-down experiments, Furuno *et al.* observed interactions between MEP50 and the SUZ12 protein.^[380] SUZ12 is a crucial 83 kDa component of the polycomb repressive complexes 2/3/4 (PRC2/3/4) and required for their activity.^[400–403] PRCs are heterogeneous protein assemblies composed of a number of various combinations of subunits. PRC2/3/4 play an important role in the regulation of the gene expression, inducing gene silencing through methylation of Lys residues located in histone tails.^[400] The potential association between SUZ12 and MEP50 also implicates the involvement of PRMT5, suggesting that PRMT5-MEP50 could be conceivably an additional and transient component of the SUZ12-containing PRC assemblies, assisting with the methylation of Arg residues. Such supposition appears to be likely considering the well-known histone methylation activity of the PRMT5-MEP50 complex.^[342,386,404,405]

Potential identification of the protein binding motif responsible for mediating the observed SUZ12-MEP50 interaction could lead to a rapid development of peptidomimetic PPI inhibitors and allow more in-depth analysis of the nature of this interaction. Peptidomimetics disrupting the SUZ12-MEP50 PPI could also be used for elucidating the plausible involvement and association of PRMT5-MEP50 with the PRC assemblies, shedding more light onto the biochemical mechanisms and activity of these complex gene expression regulators.

Considering the above motivation, the aim of the study in PART B is to develop a peptidomimetic capable of inhibiting the potential interaction between SUZ12 and MEP50. The objectives are to identify the exact sequence mediating the interaction based on the initial observations of Furuno *et al.*, optimise and cyclise the determined binding sequence, and if possible perform biophysical and biochemical analyses of the investigated PPI.

4.2. Results and Discussion

4.2.1. Identification of the MEP50 binding sequence in SUZ12

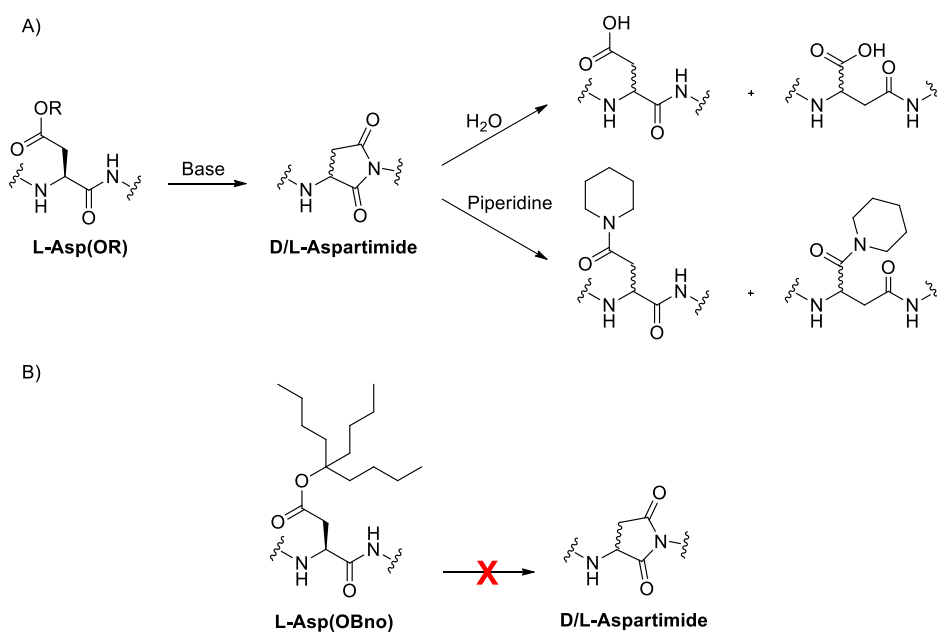
Furuno *et al.* determined the SUZ12 sequence of residues 403-536 as the protein fragment responsible for interacting with MEP50.^[380] Based on this information, it was planned to synthesise a number of peptides of different length spanning over the identified, binding SUZ12 region (**72-78**, Table 4), aiming to potentially find a short amino acid motif responsible for the interactions with MEP50.

Table 4. Sequences of fluorescently labelled SUZ12-derived peptides.

Fitc-O2Oc-KESLTTDLQTRKEKDTFNPENRQKLRIFYQFLYNNNTRQQT-NH ₂	72 (403-442)
73 (443-478) Fitc-O2Oc-EARDDLHCPWCTLNCRKLYSLKHLKLCNCRFIFNY-NH ₂	
74 Fitc-O2Oc-YQFLYNNNTRQQTEARDDLHCPWCTLNCR-NH ₂	
75 Fitc-O2Oc-EARDDLHCPWCTLNCR-NH ₂	
76 Fitc-O2Oc-KLYSLKHLKLCNCR-NH ₂	
77 Fitc-O2Oc-IFYQFLYNNNTRQQTEA-NH ₂	
78 Fitc-O2Oc-QFLYNNNTRQQ-NH ₂	
79 Fitc-O2Oc-DLHCPWCTLNCR-NH ₂	
80 Fitc-O2Oc-LHCPWCTLNCR-NH ₂	
81 Fitc-O2Oc-HCPWCTLNCR-NH ₂	
82 Fitc-O2Oc-CPWCTLNCR-NH ₂	
83 Fitc-O2Oc-NTRQQTEARD-NH ₂	

FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxoactanoic linker.

The linear sequences were synthesised using automated SPPS with microwave-assisted heating of the coupling and the Fmoc removal steps. As the peptides were intended for use with FP-based assays, the SUZ12-derived sequences were N-terminally labelled with FITC connected through a O2Oc linker. Due to the substantial length, sequences **72-74** were synthesised on the fully polyethylene glycol-based ChemMatrix[®] resin, which is more appropriate and effective for use with challenging, hydrophobic, highly structured, large or poly-Arg peptides than the typical polystyrene-based solid support.^[406] During the synthesis of **74**, a considerable amount of by-product, with a molecular mass corresponding to the piperidine adducts resulting from the reaction with aspartimide, could be detected.^[407,408] The formation of aspartimides during the SPPS of **74** was expected due to the central location of the two consecutive Asp residues in the sequence: Asp446 and Asp447. In order to prevent the aspartimide formation, resulting in the racemisation and backbone alteration (Scheme 7A), the standard t-butyl side chain protective group of Asp was replaced with an acid removable trialkylcarbinol-based protective group, i.e. 5-n-butyl-5-nonyl (Bno), shown to be very effective at reducing the aspartimide formation (Scheme 7B).^[409] To further minimise the rate of the aspartimide formation, the piperidine solution used for the Fmoc removal was supplemented with 0.5 M oxyma, an approach shown to significantly reduce the base-driven side reaction through an acid-mediated pH modulation.^[410,411] The changes to the synthetic protocol of **74** were successful and resulted in no detectable presence of aspartimide or its piperidine adducts during the peptide synthesis and purification.



Scheme 7. Aspartimide formation: **A)** Formation of aspartimide from protected Asp residue, followed by the ring opening through hydrolysis or addition of piperidine. **B)** Bno side chain protective group prevents the aspartimide formation.

The obtained peptides were tested in an FP assay against the human PRMT5-MEP50 complex. During the sample preparation **73** proved to be insoluble in the assay buffer, and thus, it was excluded from the analysis. Peptide **76** appeared to be causing aggregation resulting in aberrant readout in the assay. It seemed that **74** and **75**, compounds sharing the common sequence (residues 443-458) containing a Cys₂His₂ zinc finger motif, showed a detectable, although weak binding to the protein complex (Figure 26A). Zinc fingers from the classic Cys₂His₂ fold group are Zn binding protein elements containing two conserved Cys and His residues interacting with the Zn²⁺ ion (Figure 27).^[412] The coordination to the Zn ion provides additional stability to the zinc finger fold.^[413,414] Most of the human Cys₂His₂ zinc finger proteins appear to bind DNA, however, the DNA binding ability requires two or more tandem Cys₂His₂ zinc finger motifs, which is not a case for SUZ12.^[415,416] The Cys₂His₂ zinc finger motifs can also mediate binding to proteins and RNA, and some can interact with both DNA and proteins.^[417,418] The zinc finger of SUZ12 appears to be crucial for mediating the interactions with Jarid2, one of the PRC2 protein subunits, thus, assuring the ability of the complex to interact with chromatin.^[401] It is possible, however, that the SUZ12 zinc finger can also interact with other proteins like MEP50. Furuno *et al.* considered the involvement of the SUZ12 zinc finger in the binding of the protein to MEP50, but they dismissed this hypothesis after observing the PPI between MEP50 and the SUZ12 fragment with the zinc finger Cys residues mutated to Gly.^[380] The conclusions rejecting the zinc finger as a binding motif may, however be premature, as the mutations of the Cys residues involved in the interaction with Zn ion can destabilise the fold, but they are presumably unlikely to impact the potential protein interaction motif itself, thus, resulting in the preserved PPI.

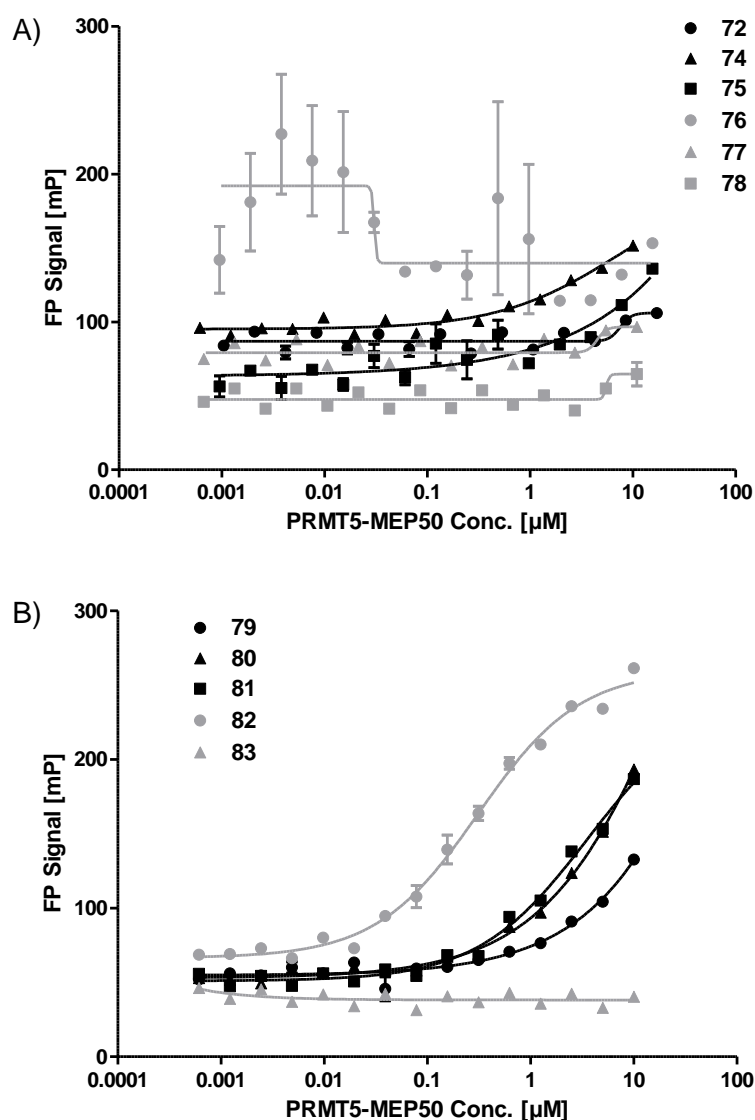


Figure 26. FP analysis of the direct interactions between SUZ12-derived peptides and the full length PRMT5-MEP50 complex: **A)** FP results for the interaction between peptides **72, 74-78** and PRMT5-MEP50. **B)** FP results for the interactions between peptides **79-83** and PRMT5-MEP50.

Considering the promising binding results observed for **74** and **75**, their binding sequences were divided into smaller peptides **79-82** of different lengths containing the amino acids of the zinc finger loop and peptide **83** with the more N-terminally oriented sequence of **74**, to explore the likely fragment combinations (Table 4). The peptides were tested in an FP assay for direct binding to PRMT5-MEP50. Compound **83** showed no interaction with the complex, but the peptides based on the zinc finger loop (**79-82**) could bind to PRMT5-MEP50 (Figure 26B). The longest loop-derived sequence **79** was binding with the lowest affinity, and the shortest peptide **82** with six amino acids (Figure 27) was binding the strongest with a K_D of 309 ± 74 nM. It was thus tentatively assumed that the motif responsible for the interactions between SUZ12 and MEP50 was contained in the short sequence of the zinc finger loop: CPWCTL. This finding seemed reasonable considering the aforementioned ability

of the zinc finger motifs to mediate PPIs.^[417,418] Encouraged by the afforded results, it was decided to synthesise the first fluorescently labelled cyclic peptides **84** and **85** (Figure 28), based on the sequence of **79**, where the macrocyclic design was guided by the crystal structure of SUZ12.^[401] The cyclopeptides were macrocyclised in solution through a connection between the side chain of Orn or Lys and the tail of the compound.

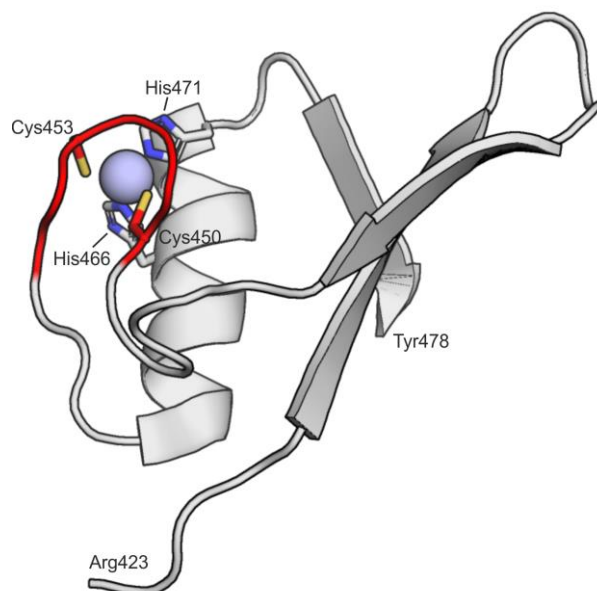


Figure 27. Crystal structure of the human SUZ12 fragment (residues 423-478 are shown) encompassing the Cys₂His₂ zinc finger motif (PDB ID: 5WAI). The red colouration designates the amino acid sequence corresponding to peptide **82**.

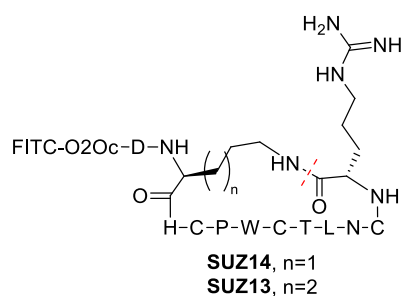


Figure 28. Structures of cyclic peptides **84** and **85** based on the sequence of **79**. The peptides are cyclised using the side chain-to-tail approach. FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker.

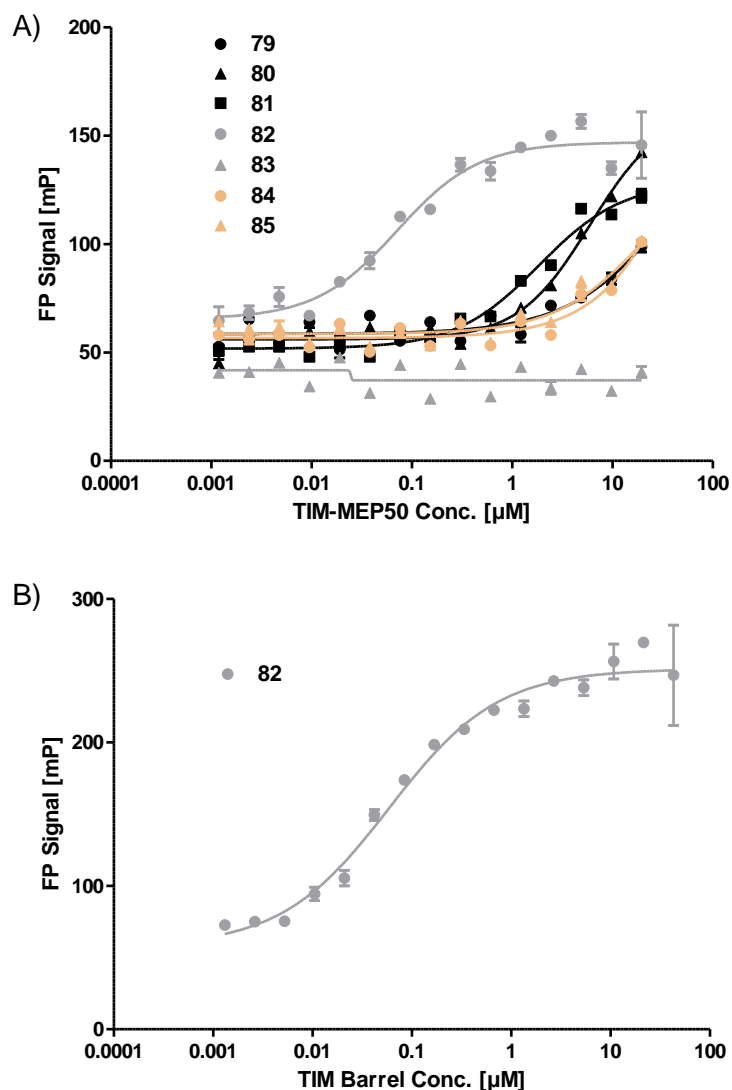


Figure 29. FP analysis of the direct interactions between SUZ12-derived peptides and truncated methyltransferase protein constructs: **A)** FP results for the interaction between peptides **79-85** and TIM-MEP50. **B)** FP results for the interactions between peptide **82** and the isolated TIM barrel domain of PRMT5.

Peptides **79-85** were tested in an FP assay against the TIM-MEP50 protein complex in order to eliminate the Rossmann and β -barrel domains of PRMT5 as the binding sites for the SUZ12 zinc finger sequences (Figure 29A). The results were similar to the analysis with the full methyltransferase complex, where **82** was binding the strongest (K_D of 71 ± 17 nM) and **83** did not show any activity. The cyclised peptides **84** and **85** did not show any significant affinity improvement to the target over the linear sequence **79**. As a control experiment confirming the identity of the interacting domain, another FP assay was conducted where **82** was tested with the isolated TIM barrel domain and no MEP50 present (Figure 29B). Surprisingly, the compound could bind to the TIM domain with a K_D value of 59 ± 22 nM, similar to the affinity observed when tested with TIM-MEP50. These unexpected results imply that either the conclusions reported by Furuno *et al.* indicating MEP50 as the interaction domain

of SUZ12 were inaccurate, and the PPI is actually mediated through the TIM barrel domain of PRMT5, or alternatively, the observed interactions are highly non-specific.^[380] In order to determine whether **82** binds non-specifically to proteins, the compound was tested in an FP assay against the GST (Figure 30A) and BSA (Figure 30B) proteins. The sequence could bind very strongly to both control proteins, resulting in a K_D of $1.1 \pm 0.03 \mu\text{M}$ with GST and $36 \pm 15 \text{ nM}$ with BSA, proving the non-specific character of the interactions.

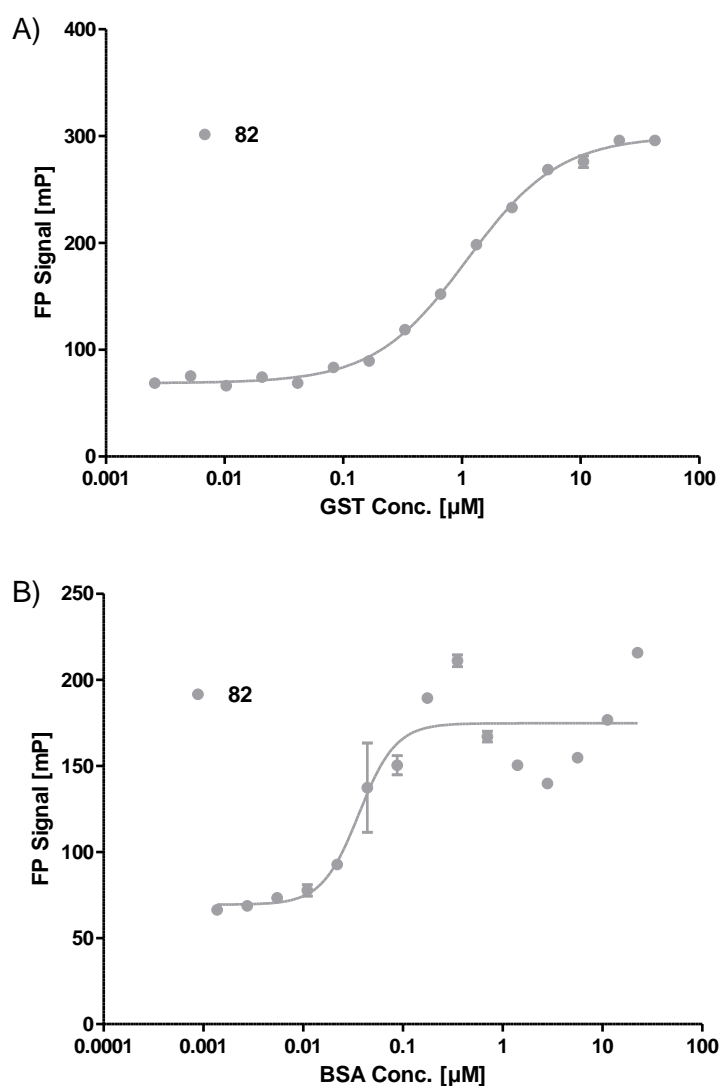


Figure 30. FP analysis of the direct interactions between **82** and control proteins: **A)** FP results for the interaction between the peptide and GST. **B)** FP results for the interaction between the peptide and BSA.

In light of the non-specific interactions of the SUZ12 zinc finger-derived peptide, it was decided to terminate any further development of peptidomimetics based on this sequence. It is possible that the interaction between MEP50 and SUZ12 identified by Furuno *et al.* is a result of a multitude of weak interactions distributed across larger surface area between the proteins and could require a specific structural fold involving longer amino acid sequence, not a short peptidic motif.

4.3. Summary and Conclusions

The PPI between MEP50 and SUZ12 was investigated, based on a previous report of the interactions between these two proteins.^[380] The SUZ12 fragment of residues 403-536 was divided into shorter peptide sequences intended for biophysical testing with different PRMT5-based protein constructs. A short sequence of compound **82**, CPWCTL, was initially identified as a tentative binding motif when tested with the full length PRMT5-MEP50 complex and the truncated TIM-MEP50 construct, affording K_D values of 309 ± 74 nM and 71 ± 17 nM, respectively. The binding sequence was corresponding to the loop of the zinc finger motif in SUZ12 and was also tested with the isolated TIM barrel domain, where a strong binding interaction was observed (K_D of 59 ± 22 nM). The interaction with the TIM barrel revealed a discrepancy between the previously reported findings regarding the SUZ12-MEP50 PPI and the here observed interaction with the TIM domain of PRMT5. The results suggested potential non-specific interactions between the tested proteins and the peptide. Sequence **82** was further assessed for binding to GST and BSA, giving K_D values of 1.1 ± 0.03 μ M and 36 ± 15 nM, respectively. The binding of **82** to the control proteins corroborated the conjecture of the non-specific interactions with the methyltransferase-based proteins. Considering the strong non-specific peptide binding, any further synthesis of the SUZ12 peptidomimetics based on the zinc finger sequence was not pursued. It is possible that the reported SUZ12-MEP50 association is mediated through a number of weaker interactions distributed over longer protein sequence and could additionally require a specific protein fold for the effective complexation.

PART C: Investigating a Novel PPI Interface Between PRMT5 and its Adaptor Proteins

5.1. Brief Introduction to PART C

PRMT5 interacts with a number of various adaptor proteins which can regulate the substrate specificity and the enzyme activity by presenting suitable cellular targets for methylation. One of the most important PRMT5 adaptor proteins is pICln, which presents three Sm proteins SmD1, SmD3 and SmB to the enzyme.^[325,326] The symmetric dimethylation of the Sm proteins allows to increase their affinity to the survival of motor neurons (SMN) complex, allowing then to assemble the Sm cores onto proper small nuclear RNAs (snRNAs), and thus, afford small nuclear ribonucleoproteins (snRNPs) essential for spliceosomal activity in cells.^[11,325,326] More recently, the PRMT5-pICln assembly was also shown to act as a master epigenetic activator of DNA double-strand break repair genes, important for cell survival and genome integrity.^[419]

PRMT5 also interacts with RioK1, an adaptor protein which recruits nucleolin for methylation by the enzyme.^[382] Nucleolin is involved in DNA metabolism affecting such processes as DNA repair, recombination and replication. Nucleolin also impacts RNA regulatory mechanisms including ribosome assembly, transcription and microRNA processing.^[420] Nucleolin is one of the most highly methylated proteins with approximately a third of its Arg residues methylated.^[421] The methylation of nucleolin has a strong effect on its interaction with DNA and RNA.^[422,423] Very recently, NF90 has been identified as a new PRMT5 substrate, and shown to also interact with the PRMT5-MEP50 complex through the association with RioK1.^[383] NF90 is a ubiquitous double-stranded RNA-binding protein associated with RNA metabolism.^[424] Interestingly, RioK1 and pICln bind to PRMT5 in a mutually exclusive manner.^[382]

It has been proposed that COPR5 works as a chromatin adaptor for PRMT5, recruiting the enzyme to function on a selected set of its target genes. The PRMT5-COPR5 complex interacts with the transcription start site of the cyclin E1 (CCNE1) gene, causing gene repression through promoter regulation. COPR5 can also modulate the histone target specificity of PRMT5, resulting in the preferential methylation of histone H4.^[384]

A better understanding of the nature of the interactions between PRMT5 and its adaptor proteins could contribute to elucidating the regulatory mechanisms governing the activity and specificity of the methyltransferase, and could allow for the development of new PPI inhibitors modulating the cellular behaviour of PRMT5. This chapter outlines the path taken to identify the consensus binding sequence of the pICln, RioK1 and COPR5 adaptor proteins used for the substrate association with PRMT5, and describes the novel PPI interface found on the surface of the TIM barrel domain. The work relies on

synthetic peptides used for investigating PRMT5-adaptor protein interactions by means of biophysical assays and protein crystallography. The available methyltransferase protein truncations were utilised to determine the binding domain and measure the binding affinity. An alanine scan of the residues in the binding sequence was performed, identifying and quantifying the contributions of individual amino acids to the PPI. A co-crystal structure was obtained for the TIM barrel domain of PRMT5 bound to a RioK1-derived peptide, confirming the biochemical assay results. The results and conclusions are in agreement with a recent and independent report by Mulvaney *et al.*, concerning the PRMT5-adaptor protein interface and the associated interactions.^[425]

5.2. Results and Discussion

5.2.1. Identification and Biophysical Analysis of the Consensus Sequence

Guderian *et al.* reported competitive binding of pICln and RioK1 to PRMT5 which suggested that either those two proteins share potentially overlapping PPI regions, or they bind to the same site on the surface of PRMT5, possibly indicating a common binding motif used for the interaction with the methylosome.^[382] The then available structural information about pICln and RioK1 was very limited, and thus, the initial analysis focused on the amino acid sequences of the proteins, in hope of finding meaningful common patterns between them. The sequences were compared using the protein BLAST alignment tool.^[426,427] The search resulted in the identification of compelling similarities between the C-terminal region of pICln (residues 223-234) and the N-terminal sequence of RioK1 (residues 9-20). A further similarity search to the identified sequences within the *Homo sapiens* protein data set afforded two additional protein fragments: the C-terminal region of COPR5 (residues 177-183) and the central region of the disabled1 (DAB1) protein (residues 388-394; Figure 31A). The results seemed to be encouraging as pICln, RioK1 and COPR5 were known PRMT5 adaptor proteins, with their potential short consensus sequences located in the terminal regions, which is consistent with the characteristics of the typical domain-peptide interactions.^[26,46,47] The identified sequences also either overlapped with or were fully incorporated into the protein regions previously believed to be important for the interactions with PRMT5. The AD3 acid domain region of pICln (residues 230-237), as well as its PH domain (residues 1-134) were previously proposed to mediate the interactions with the methyltransferase.^[381,428] Additionally, the N-terminal sequence of RioK1 (residues 1-120) and the C-terminal protein fragment of COPR5 (residues 141-184) are potentially involved in binding to PRMT5 (Figure 31B).^[382,384] DAB1, which is an adapter protein participating in signal transduction processes with a significant involvement in neuron development, has not been previously associated with either PRMT5 or MEP50.^[429,430]

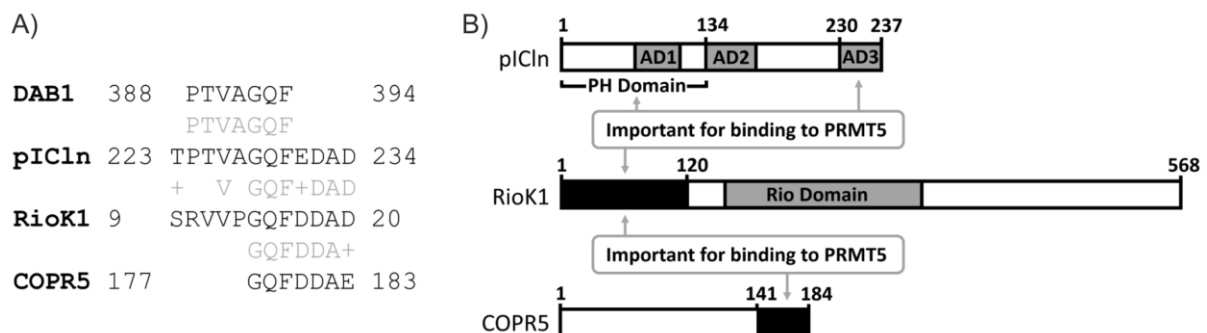


Figure 31. Protein sequences with identified alignment: **A)** BLAST alignments found for proteins DAB1, pICln, RioK1 and COPR5. **B)** Schematic representation of pICln, RioK1 and COPR5 proteins, with the proposed PRMT5 interaction domains indicated.

Fluorescently labelled linear peptides **86**, **87**, **88** and **89** (Table 5), derived from the identified alignment sequences of RioK1, pICln, COPR5 and DAB1, respectively, were prepared using Fmoc-based SPPS for FP analysis with the available PRMT5-based protein constructs. Due to the presence of up to three Asp residues in the C-terminal region of **86-88**, the peptides were regarded as prone to aspartimide formation, and thus, Fmoc removal during the synthesis was performed using 20% piperidine in DMF with an addition of 0.5 M oxyma in order to prevent the degeneration of the sequences and appearance of undesired side products.^[410,411] In result, no aspartimide or piperidine adducts were detected during the synthesis or purification of the peptides.

Table 5. Sequences and K_D values for the initial RioK1-, pICln-, COPR5- and DAB1-derived peptides **86-89**.

Peptide	Sequence	K_D Value (μM) ^[a]		
		Full length PRMT5-MEP50	TIM-MEP50	TIM barrel
86	FITC-O2Oc-SRVVPGQFDDAD-NH ₂	0.52 ± 0.04	0.84 ± 0.15	2.8 ± 0.07
87	FITC-O2Oc-TPTVAGQFEDAD-NH ₂	1.14 ± 0.04	2.1 ± 0.15	5.8 ± 0.11
88	FITC-O2Oc-MVFETGQFDDAED-NH ₂	0.55 ± 0.12	0.46 ± 0.00	2.4 ± 0.32
89	FITC-O2Oc-PTVAGQF-NH ₂	n.b.	n.b.	n.d.

^[a]As determined with FP. n.b. = no binding, n.d. = not determined, FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker.

When tested with the full length PRMT5-MEP50 complex, peptides **86**, **87** and **88** bound to the methylosome with K_D values of 0.52 μM , 1.14 μM and 0.55 μM (Table 5, Figure 32A). Sequence **89** showed no detectable binding under the tested conditions. In order to determine the interacting domain for peptides **86-88**, the compounds were also tested against TIM-MEP50 and the TIM barrel (Table 5, Figure 32B and C). The peptides displayed similar affinities for all tested protein constructs, although the determined interactions were weaker for the isolated TIM barrel, presumably due to the absence of the stabilising effect by the MEP50 unit. The obtained results suggested the TIM barrel as the domain responsible for the interactions with the peptides derived from the adaptor proteins. The inability of **89** to bind to PRMT5 indicated a crucial role of the negatively charged C-terminal peptidic region of the analysed compounds. The binding peptides **86-88** share a common sequence of seven amino acids: GQF[D/E]DA[D/E]. The identified consensus sequence was thus hypothesised to act as a binding motif mediating the interactions with PRMT5. Curiously, the pICln-derived sequence **87** gave consistently comparatively higher K_D values than **86** and **88** with every tested protein, thus also tentatively suggesting the involvement of the N-terminal region in the interaction with the protein. Compounds **86-88** did not show non-specific binding when tested with the control proteins GST and BSA, further indicating the genuine character of the observed peptide-PRMT5 interactions (Figure S1 and S2).

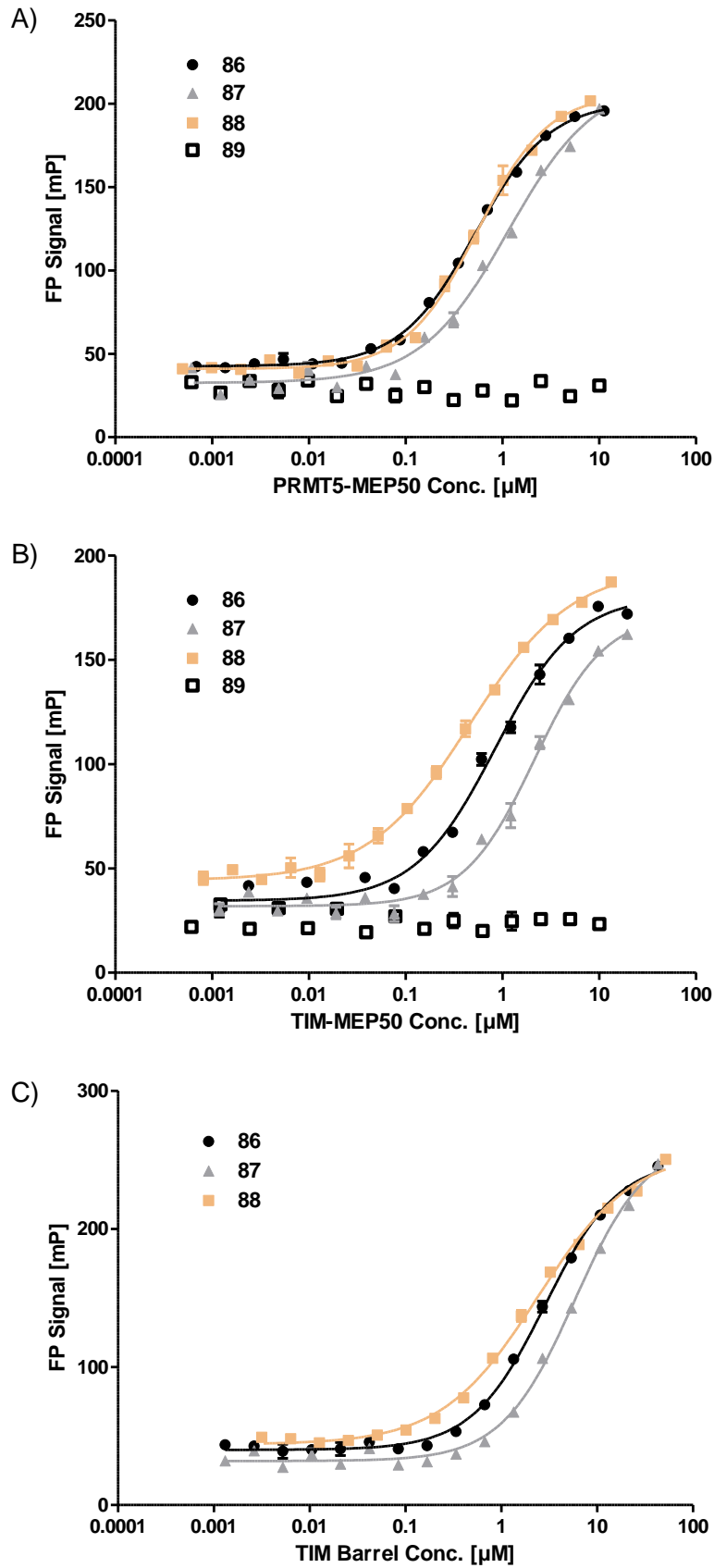


Figure 32. Results for the FP analysis of the direct interactions between the fluorescently labelled adaptor protein-derived peptides and different PRMT5-based protein constructs: **A)** Analysis of the peptide binding to full-length PRMT5-MEP50. **B)** Analysis of the interactions between peptides and TIM-MEP50. **C)** Analysis of the peptide binding to the isolated TIM barrel domain of PRMT5.

Mulvaney *et al.* also identified the sequence GQF[D/E]DA[D/E] of pICln, RioK1 and COPR5 as the PRMT5 binding motif.^[425] The identification of the motif was achieved through sequence alignment and biophysical analysis. They observed that pICln-derived sequences interacted with PRMT5 with lower affinity than RioK1 and COPR5 peptides, which is consistent with the results presented herein (Table 5, Figure 32).

The approach based on the identification of sequences shared by unrelated proteins mediating interactions with a common protein has been successfully used before to identify modulators of PPIs.^[431–433] A prototypical example utilising such a method for finding a PPI inhibitor was reported by Mochly-Rosen *et al.*, who recognised a short sequence corresponding to a homology region of unrelated protein kinase C (PKC) binders, resulting in a synthesis of an effective peptidic PKC PPI inhibitor.^[432] The previous successful applications of such an approach based on a similarity between protein sequences seem to further strengthen the validity of the reasoning outlined herein.

Table 6. Sequences and K_D values of pICln-, RioK1- and COPR5-derived, fluorescently labelled linear peptides.

Peptide	Sequence	K_D Value (nM) ^[a]
86	FITC-O2Oc-SRVVPGQFDDAD-NH ₂	522 ± 45
87	FITC-O2Oc-TPTVAGQFEDAD-NH ₂	1145 ± 38
88	FITC-O2Oc-MVFETGQFDDAED-NH ₂	549 ± 119
89	FITC-O2Oc-PTVAGQF-NH ₂	n.b.
90	FITC-O2Oc-LLMSRVVPGQFDDADSSD-NH ₂	279 ± 41
91	Ac-LLMSRVVPGQFDDADSSDK (O2Oc-FITC) -NH ₂	374 ± 16
92	FITC-O2Oc-TPTVAGQFEDADVDH-NH ₂	520 ± 40
93	FITC-O2Oc-SRVVPGQFDDADSSD-NH ₂	295 ± 11
94	Ac-SRVVPGQFDDADSSDK (O2Oc-FITC) -NH ₂	1108 ± 83
95	Fitc-O2Oc-VDTTPTVAGQFEDAD-NH ₂	2378 ± 817
96	Fitc-O2Oc-LLMSRVVPGQFDDAD-NH ₂	192 ± 26
97	Ac-LLMSRVVPGQFDDADK (O2Oc-FITC) -NH ₂	290 ± 25
98	FITC-O2Oc-SRVVPGQFDD-NH ₂	>5000
99	FITC-O2Oc-SRVVPGQF-NH ₂	n.b.
100	FITC-O2Oc-TPTVAGQF-NH ₂	n.b.
101	FITC-O2Oc-VPGQFDDAD-NH ₂	905 ± 104
102	FITC-O2Oc-GQFDDAD-NH ₂	>5000
103	FITC-O2Oc-FDDAD-NH ₂	n.b.
104	Ac-SRVVPGQFDDADK (O2Oc-FITC) -NH ₂	716 ± 36
105	Ac-VPGQFDDADK (O2Oc-FITC) -NH ₂	208 ± 74
106	Ac-GQFDDADK (O2Oc-FITC) -NH ₂	>5000
107	Ac-FDDADK (O2Oc-FITC) -NH ₂	n.b.

^[a]As determined with FP using the native PRMT5-MEP50 complex. n.b.: no binding, FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker.

In order to further investigate the interactions of the adaptor protein peptides with PRMT5, a number of various sequences of different lengths with extended or truncated termini were synthesised (Table

6, Figure S3-S5). Nearly every tested sequence was obtained in two variants, where the fluorescent label was attached either to the N-terminus through a PEG-based linker or the C-terminal end via the side chain of a Lys residue and O²OC, to assure that the labelling does not interfere with the peptide binding to the protein. The placement of the FITC label indeed had a significant effect on the interactions for some peptides (compare peptide **93** to **94**, or **101** to **105**, Table 6). The sequence of **86** could be trimmed by removing the first three N-terminal residues without significant impact on the affinity. Compounds **101** ($K_D = 905$ nM) and **105** ($K_D = 208$ nM) contained the shortest but still effective binding sequence among the tested combinations. Compounds **101** and **105** presented the nine amino acid sequence VPGQFDDAD, where the C-terminal labelling in **105** seemed more appropriate than the N-terminal FITC location in **101**, affording considerably better affinity for PRMT5. Any further removal of the N-terminal amino acids resulted in a substantial loss of the binding ability (peptides **102**, **103**, **106** and **107**). Similarly, any truncations of the C-terminal end was particularly detrimental for the K_D values (peptides **98-100**). The tested N- and C-terminal extensions of the **86** and **87** sequences in general showed improvement in the peptide affinities to the protein target (**90-93**, **96** and **97**), but not in comparison to sequence **105**. All binding sequences contained the consensus motif GQF[D/E]DA[D/E].

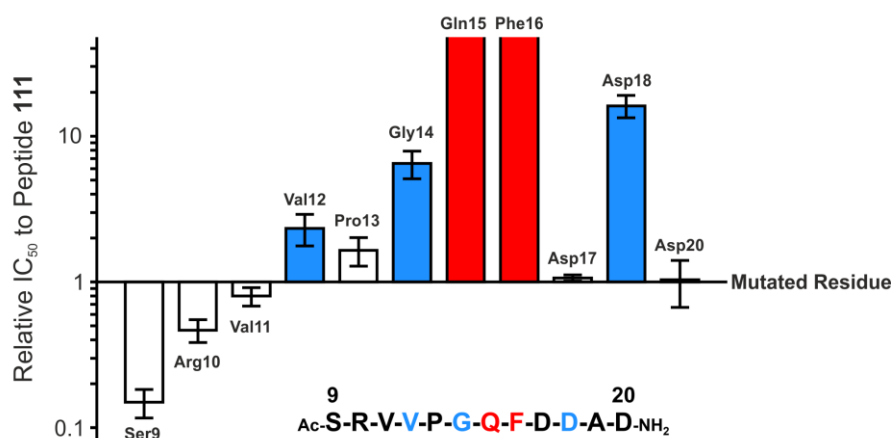
The assays examining the ability of the fluorescently labelled peptides to directly interact with PRMT5 were followed by competitive FP experiments, where compound **93** was used as a tracer rivalling for binding with unlabelled peptides (Table 7). The initial competitive analysis of compound **108**, the unlabelled equivalent of tracer **93**, confirmed that the FITC tag does not contribute to the binding through an introduction of strong non-specific interactions (Table 7, Figure S6). Additionally, the evident competition between the tracer and peptides **109-111** proved that RioK1, pICln and COPR5 derived peptides bind to the same interaction site on the surface of the TIM barrel domain (Table 7, Figure S6).

An alanine scan was performed for the sequence of the RioK1-based peptide **111** in order to determine the individual contributions of the amino acid side chains to the investigated interaction (Table 7, Figure S7). The side chains of Gln15 and Phe16 were the most significant for binding, and the mutations to Ala caused an over 50-fold decrease in the affinities (peptides **115** and **116**). Thus, Gln15 and Phe16 were considered to be hot-spot residues. The experiment also revealed a positive contribution of Val12 (**119**), and Asp18 (**113**) side chains to the peptide-protein interaction. A significant affinity drop caused by replacing Gly14 with Ala suggested that Gly could be crucial for the correct peptide conformation upon binding to the protein (compound **117**). The mutation of the residues in the N-terminal SRV motif (peptides **120-122**) had a positive effect on the peptide affinity, and is in agreement with the observations made during the analysis of the peptide truncations

(compare compound **104** to **105**, Table 6). The Ala scan results are consistent with the amino acid contributions for the interaction between a pICln-derived peptide and PRMT5, reported by Mulvaney *et al.*^[425]

Table 7. Sequences and IC₅₀ values of pICln-, RioK1- and COPR5-derived linear peptides and the results of the alanine scan for peptide **111**.

Peptide	Sequence	IC ₅₀ (μM) ^[a]
108	Ac-SRVVPGQFDDADSSD-NH ₂	1.5 ± 0.4
109	Ac-TPTVAGQFEDAD-NH ₂	14.6 ± 0.9
110	Ac-MVFETGQFDDAED-NH ₂	3.4 ± 0.2
111	Ac-SRVVPGQFDDAD-NH ₂	6.0 ± 0.7
112	Ac-SRVVPGQFDDAA-NH ₂	6.2 ± 2.2
113	Ac-SRVVPGQFDAAD-NH ₂	97 ± 17
114	Ac-SRVVPGQFADAD-NH ₂	6.4 ± 0.3
115	Ac-SRVVPGQADDAD-NH ₂	>300
116	Ac-SRVVPGAFDDAD-NH ₂	>300
117	Ac-SRVVPAQFDDAD-NH ₂	39 ± 8.5
118	Ac-SRVVAGQFDDAD-NH ₂	9.9 ± 2.2
119	Ac-SRVAPGQFDDAD-NH ₂	14 ± 3.4
120	Ac-SRAVPGQFDDAD-NH ₂	4.8 ± 0.7
121	Ac-SAVVPGQFDDAD-NH ₂	2.8 ± 0.5
122	Ac-ARVVPQFDDAD-NH ₂	0.9 ± 0.2



^[a]As determined with FP using the native PRMT5-MEP50 complex and compound **93** as a fluorescent tracer.

5.2.2. Initial Attempted RioK1 Peptide Cyclisation

Considering the identified importance of Gly in the binding sequences which potentially suggests a specific conformational dependence on this amino acid, as well as the presence of the PG motif, that is often found in β -turns, indicating a plausible turn in the peptides, it was hypothesised that the RioK1 binding fragments could form a β -hairpin upon the interaction with PRMT5.^[103,228] Following this assumption, a number of fluorescently labelled cyclic RioK1-derived peptides was synthesised, capable of adopting and stabilising a β -sheet arrangement. The compounds were designed according

to the $2(2n + 1)$ rule described in section 1.2.3., thus, incorporating 14 amino acids, as well as two turning motifs placed at equal distances from each other within the structures (peptides **123-128**, Figure 33).^[266] Combinations of various amino acid pairs intended to stabilise the turns were incorporated into the peptides including the highly effective D-Pro-L-Pro motif.^[234–242] The macrocycles were synthesised starting from a Glu residue with allyl protection of the C-terminus (Scheme 8). The amino acid was attached to the rink amide resin through its free carboxylic acid located in the side chain. The linear sequences were obtained through standard SPPS, incorporating a Lys(Mtt) residue, intended for use as a handle that could enable coupling to a fluorescent tag or alternatively a cell penetrating motif. The allyl group protecting the C-terminus was then removed using $\text{Pd}(\text{PPh}_3)_4$ and the free termini were connected together via amide coupling on solid support. The macrocyclisation was followed by the removal of protective Mtt group from Lys, and coupling of O2Oc linker and FITC. The labelled cyclic peptides were then cleaved from the resin using TFA with concomitant side chain deprotection of the remaining residues (Scheme 8). As the rink amide resin had been used for the peptide synthesis, the cleavage from solid support affords an amide group, transforming the initially used Glu amino acid to the desired Gln residue. A linear peptide **129** incorporating the D-Pro-L-Pro motif was synthesised as a control to the cyclised versions of the RioK1-inspired sequences (Figure 30).

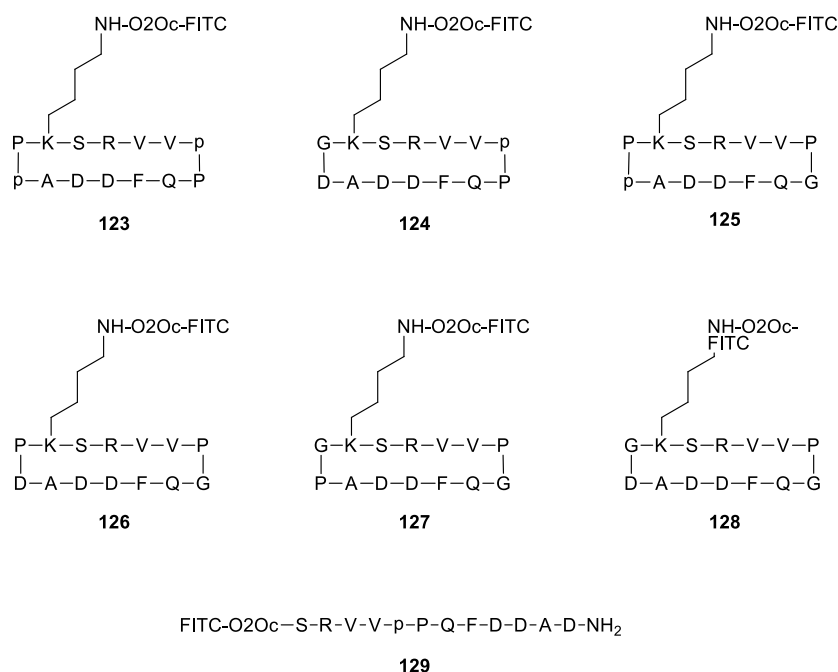
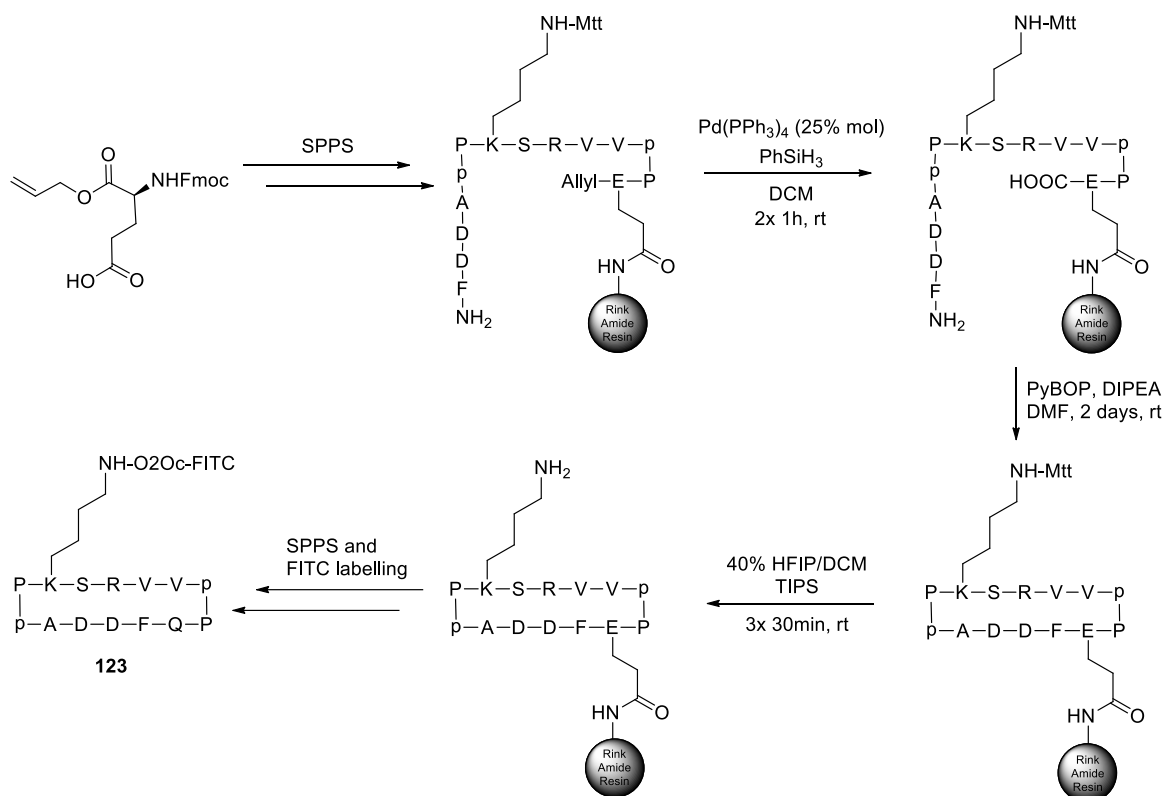


Figure 33. Structures of the RioK1-derived compounds intended to mimic a presumed β -hairpin conformation of the PRMT5-bound peptides. p = D-Pro, FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker.



Scheme 8. Exemplary synthesis of cyclic peptide **123** intended to mimic a β -hairpin conformation. The same synthetic steps were applied in order to obtain compounds **124-128**. FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker

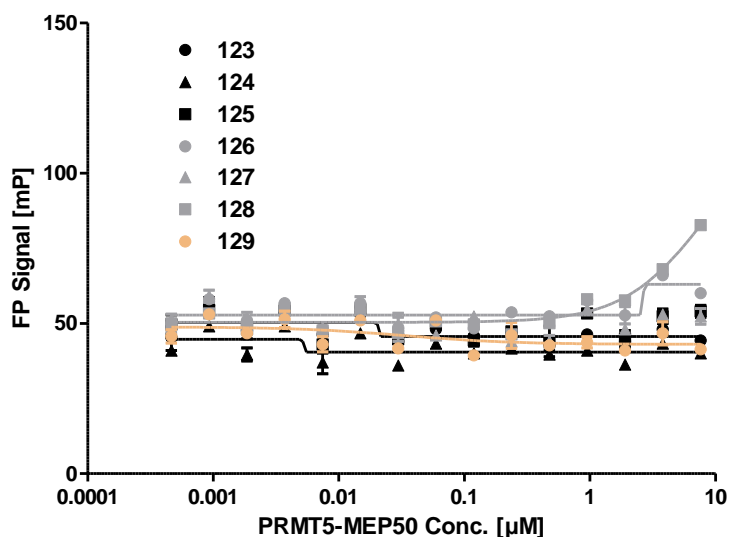


Figure 34. Results of FP-based analysis of the interactions between fluorescently labelled peptides **123-129** and the full-length PRMT5-MEP50 complex.

Compounds **123-129** were tested for direct binding to the PRMT5-MEP50 complex in an FP assay. Cyclic peptide **128** showed only very weak affinity towards the target, with the remaining compounds

presenting no detectable interactions with PRMT5. The obtained results strongly suggested that the RioK1 peptide does not assume the envisaged β -hairpin arrangement upon binding to PRMT5, and more detailed structural elucidation was needed in order to understand the PPI interface and allow to design more effective peptidomimetics.

5.2.3. Crystallographic Elucidation of the Interaction Between PRMT5 and RioK1 Peptide

More detailed insight into the structural relationship between the adaptor proteins and PRMT5 at the novel PPI interface was obtained by X-ray crystallography. With the assistance of Dr. Raphael Gasper of the Max Planck Institute in Dortmund, as well as Dr. Vincent Olieric and Dr. John Beale at the Swiss Light Source of the Paul Scherrer Institute, a co-crystal structure of the unlabelled human PRMT5 TIM barrel bound to the RioK1-derived peptide **108** was obtained. The structure was resolved using data up to a resolution of 2.55 Å (PDB ID: 7BOC). Attempts of producing a co-crystal structure of the adaptor peptide bound to either the available full length PRMT5-MEP50 methyltransferase or the TIM-MEP50 complex were unsuccessful.

Typical TIM domains contain a circular β -sheet core composed of eight parallel β -strands, and this pattern has also been seen across the representative PRMT5-MEP50 structures.^[370–373,425] In the co-crystal structure of the TIM barrel, the fold shows an unusual seven-stranded β -sheet where one of the β -strands is placed in an antiparallel orientation (Figure 35A and S8). The atypical conformation of the domain core was corroborated by the additionally obtained anomalous density maps based on the sulphur anomalous signal of the Cys and Met residues, as well as the anomalous signal of a Pt-derivative resulting from the soaking of the crystals in a solution of potassium tetrachloroplatinate (Figure S9). A plausible explanation for the observed unusual core structure of the crystallised TIM barrel domain is the absence of the stabilising MEP50 protein and the truncation of the catalytic domains of PRMT5. The regions of the TIM barrel between residues 1-40 and 52-76 that normally encompass and interact with the insertion finger of MEP50 were disordered. Unexpectedly, the protein fragment constituting the PRMT5 loop interacting with MEP50 (residues 165-178; the loop was discussed in PART A) was also unstructured. The disordered regions of the TIM domain were compared with the results obtained using DrugScore^{PPI}, which estimated the energy contributions to the PPI by the interfacing residues of PRMT5 and MEP50, based on a crystal structure of the full-length methylosome complex (Table S1 and S2).^[388] A number of the TIM barrel residues with the calculated $\Delta\Delta G$ values above 1 kcal/mol were located in the disordered protein regions. The residues with the considerable predicted contribution to the PPI found in the unstructured fragments are: Arg62, Arg68,

Arg164, Asp165, Ile167, Ile168 and Asn170. The results obtained with the DrugScore^{PPI} service, thus further corroborate the expected stabilisation effect of the MEP50 unit onto PRMT5.

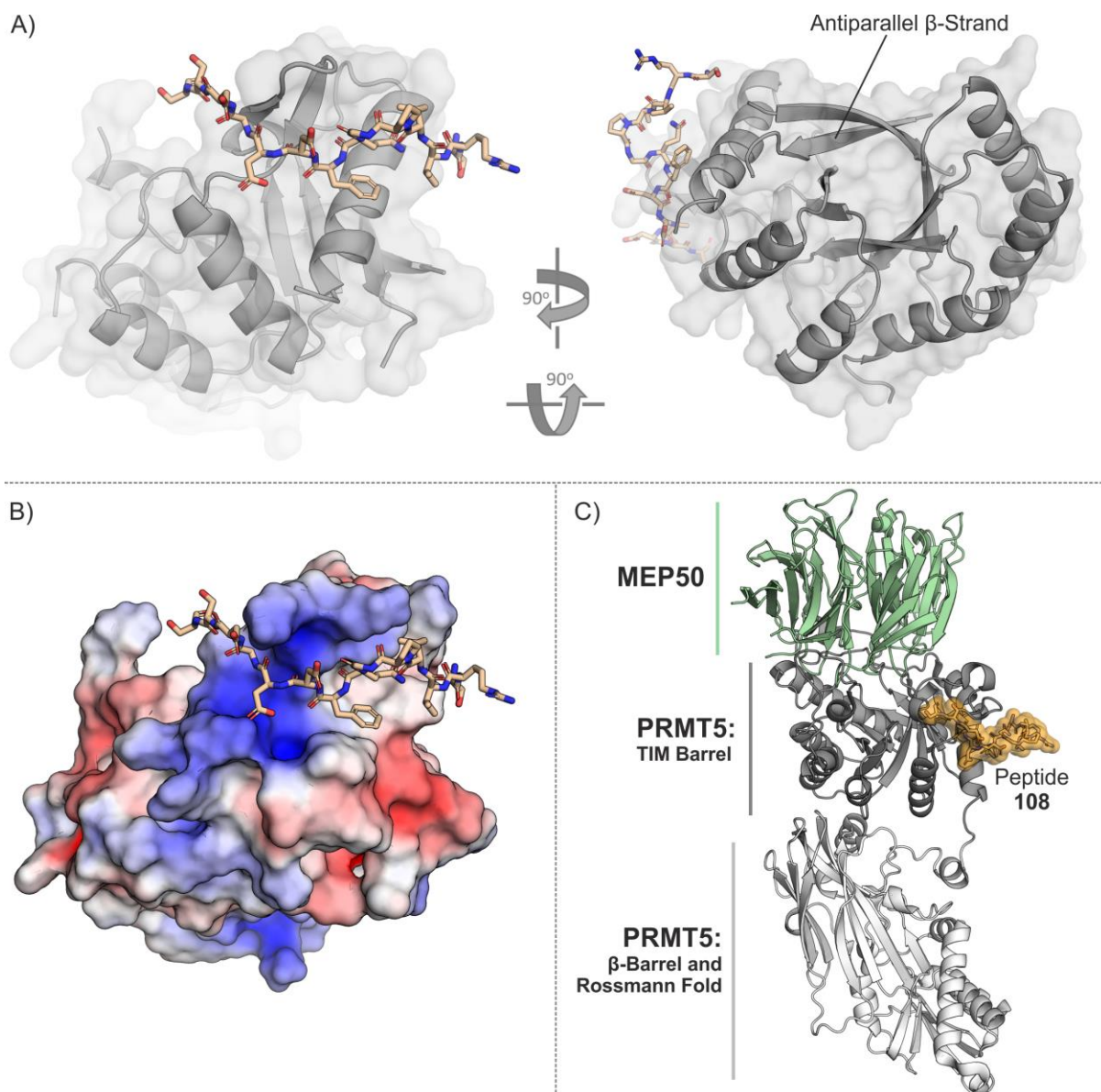


Figure 35. Visualisation of the structural data for the interaction between the TIM barrel of PRMT5 and RioK1-derived peptide **108**: **A)** Obtained co-crystal structure of the TIM barrel domain and RioK1 peptide **108** (PDB ID: 7BOC). **B)** Co-crystal structure of the TIM barrel and **108** (PDB ID: 7BOC). The electrostatic potential of the TIM barrel domain is visualised, where blue, white and red surface colouration designates positive, neutral and negative electrostatic potential, respectively. **C)** Visualisation of the peptide location in the context of the heterodimer of the PRMT5-MEP50 complex. The representation was afforded through a superposition of the obtained here TIM barrel-bound RioK1 peptide and the PDB structure 4GQB.

The RioK1-derived peptide is bound to a shallow groove on the surface of the TIM barrel, and is located away from the unstructured protein regions which normally interface with MEP50 (Figure 35, 36 and S10). The groove spans between residues Lys240, Lys241, Lys248 and Tyr286. A portion of the binding site has a positive electrostatic potential and is housing the negatively charged C-terminal region of

peptide **108**, whereas, the remaining area of the groove interfacing with the compound is electrostatically neutral (Figure 35B). Residues Ser21 and Ser22 on the C-terminal end are fixed in place due to the location of the crystal symmetry mate (Figure S11). The N-terminal residues SRV appear to be strongly solvent exposed. The central peptide fragment corresponding to the identified consensus sequence GQF[D/E]DA[D/E] forms a number of interactions with the protein domain. The side chains of peptidic Asp20 and protein Lys240 demonstrate a potential for interacting with each other, although due to the large and unrestricted freedom of movement of Lys240 side chain, and thus, a very high expected entropic penalty upon forming a salt bridge with Asp20, the energetic contribution to the PPI was assumed to be negligible. This conjecture appears to be in line with the results of the Ala scan afforded for compound **112**, where the mutation of Asp20 to Ala had no effect on the measured binding affinity to PRMT5 (Table 7). A salt bridge is formed between Asp18 and Lys248, and the considerable strength of this interaction is also reflected in the biophysical analysis with compound **113** (Table 7). The backbone carbonyl groups of Asp17 and Gln15 interact through H-bonds with the side chain of Gln239. The peptide assumes an S-shaped double β -turn involving the hot-spot residues Gln15 and Phe16, allowing to orient appropriately their side chains in relation to the surface of TIM barrel. The participation of Val12 and Pro13 in the double turn explains the observed significant affinity loss caused by the truncation of these amino acids in the examined sequences (peptides **102** and **106**, Table 6). The initial inference about the involvement of the Pro-Gly motif into a turn motif proved to be accurate, despite the fact that the bound peptide does not assume a β -beta hairpin arrangement. Considering the location of the binding interface on the TIM barrel, it appears to be possible for four adaptor proteins to bind simultaneously to the heterooctameric PRMT5-MEP50 complex (Figure S12).

The crystallographic results presented here, as well as the drawn structural conclusions regarding the discussed domain-peptide interface are in agreement with the data generated by Mulvaney *et al.*, who co-crystallised adaptor peptides with the full-length PRMT5-MEP50 complex. The calculated root-mean-square deviation (RMSD) between the RioK1 peptides in the complex presented here and the structure afforded by Mulvaney *et al.* is very low and equal to 0.41 Å, thus, indicating a very strong similarity between the two peptide orientations (Figure 36B).^[425]

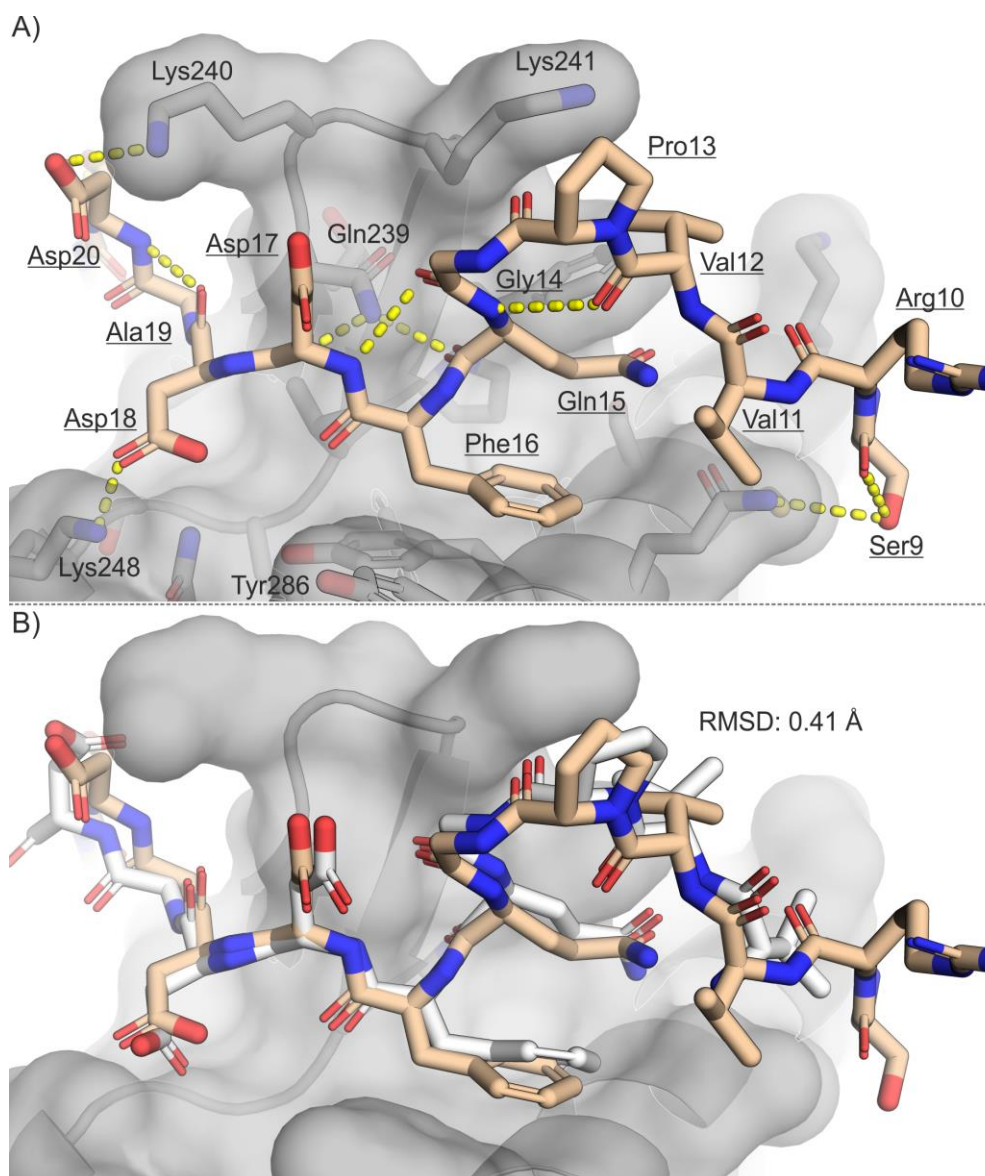


Figure 36. Visualisation of the RioK1-derived peptide **108** in the binding groove of the TIM barrel domain: **A)** Depiction of the peptide position and interactions in the binding groove (PDB ID: 7BOC). **B)** Comparison of the obtained here peptide structure (wheat, PDB ID: 7BOC) with the structure afforded by Mulvaney *et al.* (white, PDB ID: 6V0N).

5.2.4. Analysis of the Interactions Between PRMT5 and Full-length pICln and RioK1

It was deemed potentially informative to compare the binding results obtained for isolated peptides based on the identified consensus motif of the pICln/RioK1/COPR5 adaptor proteins with the performance of full-length proteins. To this end, the expression and isolation of pICln, RioK1 and COPR5 were attempted at the Protein Chemistry Facility of the Max Planck Institute in Dortmund. Full-length human pICln N-terminally tagged with 6His-TRX was successfully expressed in *E. coli* and purified. Full-length RioK1 tagged with 6His needed to be expressed in HighFive insect cells. The tags

were cleaved from the proteins during the purification process. Despite a number of attempts COPR5 could not be successfully isolated.

In order to determine the binding affinity of pICln and RioK1 to PRMT5-MEP50 using FP assay, the proteins needed to be labelled with an appropriate fluorophore. pICln and RioK1 were thus treated with Alexa488 dye containing a reactive succinimidyl ester, labelling primary amines in the proteins. The labelled adaptor proteins were tested for binding against the full-length PRMT5-MEP50 complex, as well as the truncated TIM-MEP50 assembly. Protein pICln gave K_D values of 9.1 ± 1.5 nM and 4.1 ± 0.6 nM for the full and truncated complex, respectively (Figure 37A). RioK1 tested with PRMT5-MEP50 returned a K_D of 34.1 ± 9.9 nM, and when analysed with TIM-MEP50 it afforded a K_D of 6.3 ± 0.2 nM (Figure 37B).

The FP analysis of the PPIs was complemented by flow induced dispersion analysis (FIDA). FIDA utilises the properties of particle diffusion, where large compounds diffuse slowly, and small compounds can diffuse at a faster rate. The compound of interest, referred to as an indicator, is fluorescently labelled to allow detection, and the interactions with its binding partner can be investigated based on the changes in the observed diffusion rate which is dependent on the ratio between the free and complexed indicator in the analysed samples.^[434,435] This technique can be thus used for determining the K_D values between two interacting components. The analysis of the Alexa488 labelled pICln binding to PRMT5-MEP50 and TIM-MEP50 gave K_D values of 39.3 and 64.1 nM. The testing of fluorescently tagged RioK1 afforded a K_D of 1.5 nM with PRMT5-MEP50 and 20.6 nM with TIM-MEP50 (Table S4 and Figure S13). The results obtained with FP and FIDA are comparable (Table 8), indicating a significantly higher affinity of the full-length adaptor proteins to the methylosome than the observed affinities of the short peptides containing the consensus sequence. It is likely that the full-length proteins utilise additional sites mediating the investigated PPIs. Pesiridis *et al.* previously indicated the PH domain of pICln as an additional PRMT5 interaction site alongside the acidic AD3 region which contains the binding motif GQF[D/E]DA[D/E].^[381] Considering the similar affinity between pICln and RioK1 to the tested methyltransferase protein complexes, it seems plausible for RioK1 to also have additional interaction region coexisting with the consensus sequence. Alternatively, the much higher binding strength of the full-length proteins could be a result of a structural preorganisation of the binding motif caused by the supporting interactions within the adaptor proteins, significantly lowering the entropic penalty upon binding to PRMT5.

Cox *et al.* performed a similar experiment using microscale thermophoresis and Alexa488 labelled GST-PRMT5, affording a comparable K_D value of 100 ± 29 nM with RioK1, supporting the results presented here.^[383]

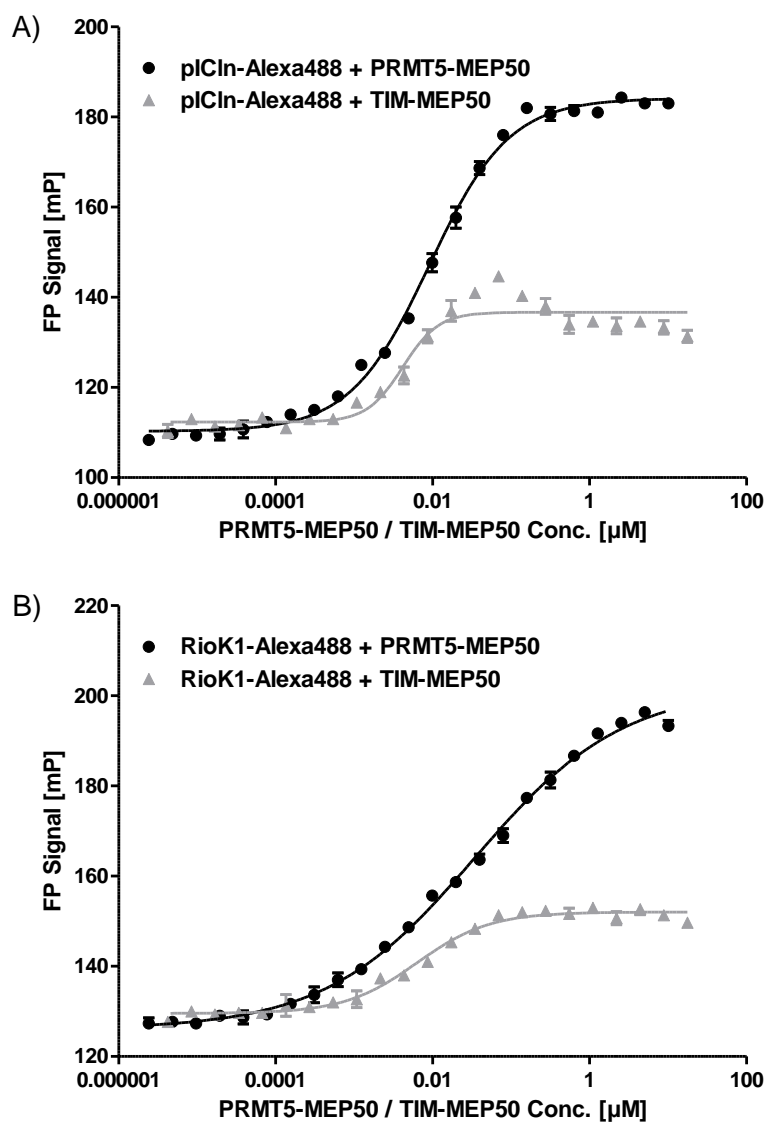


Figure 37. FP assay results for the investigation of the direct interactions between the fluorescently labelled full-length adaptor proteins and PRMT5-MEP50 or TIM-MEP50: **A)** Results for the analysis with pICln. **B)** Results of the analysis with RioK1.

Table 8. Overview and comparison of the K_D values obtained for the binding of the adaptor proteins pICln and RioK1 to PRMT5-MEP50 or TIM-MEP50. The results were obtained through analysis with FP and FIDA.

Method	Protein Complex	K_D [nM]	
		pICln-Alexa488	RioK1-Alexa488
FP	PRMT5-MEP50	9.1	34.1
FIDA	PRMT5-MEP50	39.3	1.5
FP	TIM-MEP50	4.1	6.3
FIDA	TIM-MEP50	64.1	20.6

5.3. Summary and Conclusions

PRMT5 adaptor proteins are important for regulating PRMT5 substrate specificity. PPIs between PRMT5 and the adaptor proteins pICln, RioK1 and COPR5 are crucial for the correct cell functioning by directing PRMT5 to the appropriate targets. A consensus binding motif GQF[D/E]DA[D/E] was identified in pICln/RioK1/COPR5 through a sequence alignment search. FP assays were used to determine that the adaptor peptides bind to the TIM barrel domain of PRMT5. The shortest RioK1 peptide capable of interacting with the methyltransferase is the nine amino acid long sequence VPGQFDDAD, where Gln and Phe were identified as hot-spot residues. The results obtained in the biophysical assays were confirmed by a co-crystal structure of the TIM barrel domain bound to a RioK1-derived peptide. The peptide is wedged into a shallow groove on the surface of the domain, where the negatively charged C-terminus of the peptide resides in an electrostatically positive region of the groove. The peptide also forms a number of H-bonds with the protein, and presents a double S-shaped β -turn involving residues VPGQFD. The interactions between PRMT5-MEP50/TIM-MEP50 and the full-length proteins pICln and RioK1 were assessed through FP and FIDA-based analyses. The experiments revealed that the full-length proteins bind the methyltransferase with significantly higher affinity than the short peptides containing the identified consensus binding sequence. The presented findings are in agreement with the independent report released by Mulvaney *et al.*^[425]

PART D: Targeting Interactions of PRMT5 with its Adaptor Proteins

6.1. Brief Introduction to PART D

The discovery of the novel adaptor protein interaction motif and the subsequent characterisation of the PPI interface between the pICln/RioK1/COPR5 adaptor proteins and PRMT5 opened avenues for development of new PRMT5 PPI inhibitors.^[425,436] Considering the important contribution of pICln, RioK1 and COPR5 to cellular functions, and still poorly understood involvement of PRMT5 in the regulation of various cellular processes and mechanisms, compounds which can modulate the enzyme interactions with its adaptor proteins, and thus, impact the substrate selectivity, seem to be highly desirable. Analysis performed by Mulvaney *et al.* indicated that symmetric demethylation of at least 25 PRMT5 substrates is mediated through the newly discovered PPI interface on the surface of the TIM barrel.^[425] Targeting the adaptor protein PPI interface without disrupting the global enzymatic activity of the methylosome could allow for more precise investigation of the enzyme biochemistry than what is currently possible with the active site inhibitors, minimising or fully eliminating the undesired impact on the PRMT5 substrates that are of no interest to a given study. Thus, a focused inhibition of the adaptor protein PPI in PRMT5 should cause only more relevant biological responses to the investigated cellular mechanisms, increasing the confidence in the observations and conclusion drawn based on the enzyme modulation.

Very recently, McKinney *et al.* utilised the findings concerning the PRMT5 adaptor protein PPI interface and developed the first-in-class covalent groove binder able to inhibit the interactions between the enzyme and the adaptor proteins.^[437] The afforded inhibitor showed only moderate activity in the μM range and was reportedly prone to side-reactivity with GSH. The compound was obtained after extensive screening campaigns with over 900.000 compounds from diverse libraries, combined with an NMR-based fragment screen and a virtual pharmacophore screen. Despite the considerable efforts, the performed screening campaigns resulted in only one weakly binding compound cluster, highlighting the immense difficulty of the PPI inhibitor development.^[437] Considering the results obtained by McKinney *et al.*, it was deemed appropriate to apply an alternative approach to finding potent inhibitors targeting the interface between PRMT5 and its adaptor proteins pICln/RioK1/COPR5.

In this chapter, steps taken towards a development of a potent macrocyclic PRMT5 adaptor protein interaction inhibitor (PAPII) are described, based on the existing binding epitope of the PPI. The strategy to obtain potent PAPIIs involved structure-guided design of macrocycles, extensively tested

through FP assays against the PRMT5-MEP50 complex. Macrocyclisation significantly increased the stability of the peptide towards proteases. Various side chain modifications, as well as backbone N-methylations were explored, considerably improving the affinity of the cyclopeptide to the target. The resulting compound was tested in a biologically relevant environment, disrupting the PRMT5-adaptor protein PPI. The afforded peptidomimetic showed unexpected selectivity towards inhibition of the PRMT5-RioK1 PPI over the interaction involving pICln.

6.2. Results and Discussion

6.2.1. Development of the Macrocyclic PAPIIs

The designs towards a macrocyclic PAPII were based on the previously obtained co-crystal structure of the PRMT5 TIM barrel domain bound to the RioK1-derived peptide, as well as the biophysical data afforded through the FP analyses described in PART C. It appeared feasible structurally to connect the side chain of Asp17 and the α -carbon of Gly14 with a covalent linker (Figure 38A). Such an approach would presumably allow to stabilise the double β -turn in the peptide, as well as encompass the hot-spot residues Gln15 and Phe16, and thus, the cyclisation was expected to positively contribute to the stability against proteases and improve binding affinity. The choice of the envisaged connection points was further supported by the Ala scan results where the substitution of Asp17 with Ala had no significant effect on the binding affinity to PRMT5 (Table 7, section 5.2.1.). Mutation of Gly14 to Ala resulted in a significant loss in the ability to bind to the target. *In silico* manipulations of the peptidic structures, however, suggested that Gly14 could be replaced by a D-amino acid in order to facilitate a feasible ring connection, as well as, that Asp17 could be supplanted by either L- or D-residue. Computer models also indicated a need for either a four or five atom long linker between the α -carbons in the peptidic backbone (Figure 38B).

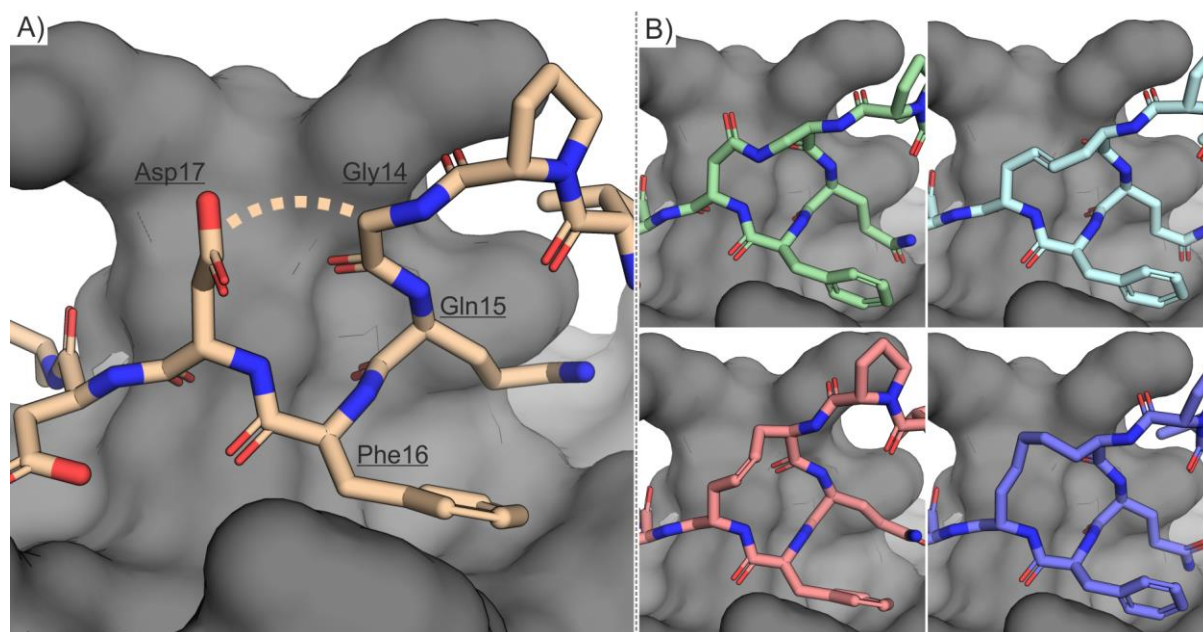


Figure 38. Designing macrocyclic RioK1 peptide: **A)** A fragment of RioK1-derived peptide bound to the groove of the PRMT5 TIM barrel domain. The dashed line indicates a proposed connection for macrocyclisation. **B)** Computer models of cyclic RioK1 peptides, with a four atom amide-based linker (green), four atom unsaturated RCM linker (*trans* - blue, *cis* - salmon) and five atom long saturated RCM linker (purple).

To validate whether the presence of D-amino acids at position 14 and 17 was tolerated by the RioK1 peptide, compounds **130-132** were synthesised and tested in an FP assay for binding to PRMT5-MEP50 in the competition with fluorescently labelled **93** (Table 9 and Figure S14). Replacement of either Gly14 or Asp17 with D-Ala (peptides **130** and **131**) resulted in only a moderate decrease in the affinity when compared to the native sequence of **111**, but the simultaneous incorporation of two D-Ala residues in compound **132** had a very negative impact on the ability to interact with the enzyme.

Table 9. Sequences and IC₅₀ values for RioK1-derived peptides.

Peptide	Sequence ^[a]	IC ₅₀ (μM) ^[b]
111	Ac-SRVVPGQFDDAD-NH ₂	6.0 ± 0.7
130	Ac-SRVVPGQFaDAD-NH ₂	14 ± 2.7
131	Ac-SRVVPaQFDDAD-NH ₂	16 ± 5.9
132	Ac-SRVVPaQFaDAD-NH ₂	198 ± 99

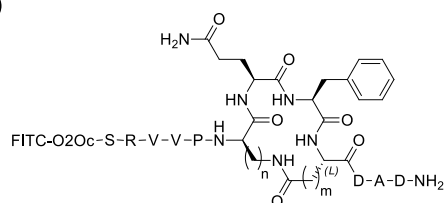
^[a]D-Ala is designated by the symbol "a".

^[b]As determined with FP using the native PRMT5-MEP50 complex and compound **93** as a fluorescent tracer.

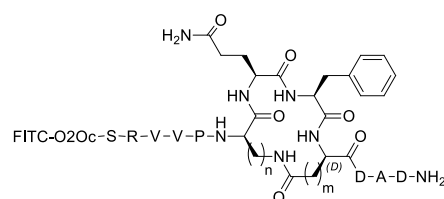
The first generation of macrocycles was based either on ring connection through an amide bond (peptides **133-140**) or a hydrocarbon linker resulting from RCM (**141-144**), incorporating from 14 to 16 atoms (Figure 39A). The compounds were synthesised on a low substitution rink amide resin, which allowed to prevent cross-linking between peptide chains during the cyclisation step. The piperidine solution used for the Fmoc removal was supplemented with 0.5 M oxyma, in order to suppress the formation of aspartimides. The linear peptide sequences intended for the cyclisation via an amide bond formation (**133-140**) incorporated an Asp/Glu residue with the side chain protected by the allyl group, and diaminobutyric acid (Dab)/diaminopropionic acid (Dap) protected by Mtt. The allyl group was first removed by Pd(PPh₃)₄ and then Mtt was cleaved with HFIP. The macrocyclisation was performed on resin, coupling the afforded free amine and carboxylic acid with PyBOP (Scheme 9A). In cases where the allyl protected Asp was used, a very high amount of aspartimide related by-products was detected during the synthesis of the peptidic sequences, and thus, it was necessary to protect the preceding backbone nitrogen with the 2,4-dimethoxybenzyl (Dmb) group. Dmb was installed through the reductive amination between a free primary amine on the N-terminus of the peptide fragment and 2,4-dimethoxybenzaldehyde. As Dmb causes a considerable steric hinderance for a subsequent coupling reaction, the installation of the Fmoc-Asp(OAll) residue was repeated three times in the presence of COMU and oxyma (Scheme 9B). The cyclopeptides containing the hydrocarbon linkers (**141-144**) were obtained through an RCM reaction under microwave irradiation using the 2nd generation Hoveyda-Grubbs catalyst, and where it was required the afforded unsaturated linker was reduced with 2,4,6-triisopropylbenzenesulfonyl hydrazide (TPSH) at 50 °C (Scheme 10A). Where it was feasible, the *cis* and *trans* isomers were separated during the final purification step using preparative

HPLC. In order to facilitate a straightforward analysis of the direct peptide interactions with PRMT5 in an FP assay, allowing for the determination of the K_D values, the compounds were fluorescently labelled with FITC on the N-termini.

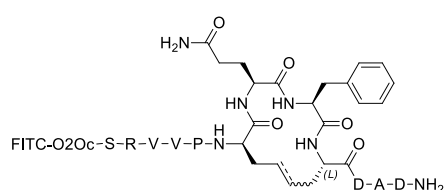
A)



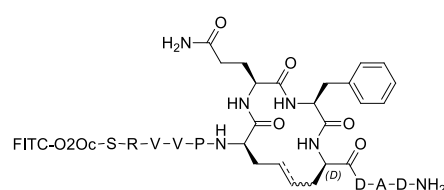
133: $n=1, m=1$ ($K_D > 2 \mu\text{M}$)
134: $n=1, m=2$ ($K_D > 2 \mu\text{M}$)
135: $n=2, m=1$ ($K_D > 2 \mu\text{M}$)
136: $n=2, m=2$ ($K_D > 2 \mu\text{M}$)



137: $n=1, m=1$ ($K_D > 2 \mu\text{M}$)
138: $n=1, m=2$ ($K_D > 2 \mu\text{M}$)
139: $n=2, m=1$ ($K_D > 2 \mu\text{M}$)
140: $n=2, m=2$ ($K_D > 2 \mu\text{M}$)

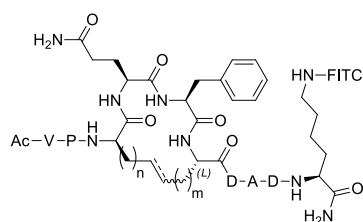


141: unsaturated ($K_D = 719 \pm 41 \text{ nM}$)
142: saturated ($K_D = 1.29 \pm 0.11 \mu\text{M}$)

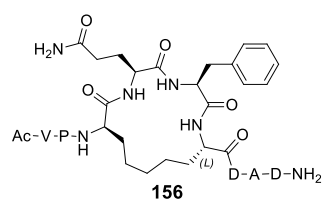


143: unsaturated
(Isomer 1: $K_D > 2 \mu\text{M}$;
Isomer 2: $K_D > 2 \mu\text{M}$)
144: saturated ($K_D > 2 \mu\text{M}$)

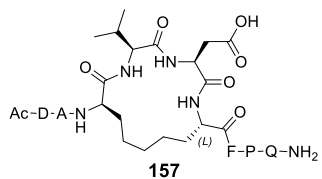
B)



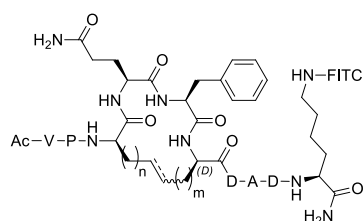
145: $n=1, m=1$, unsaturated ($K_D = 452 \pm 1 \text{ nM}$)
146: $n=1, m=2$, unsaturated
(Isomer 1: $K_D = 352 \pm 49 \text{ nM}$
Isomer 2: $K_D = 1.68 \pm 0.92 \mu\text{M}$)
147: $n=2, m=1$, unsaturated ($K_D = 562 \pm 92 \text{ nM}$)
148: $n=2, m=2$, unsaturated ($K_D > 2 \mu\text{M}$)
149: $n=2, m=1$, saturated ($K_D = 430 \pm 96 \text{ nM}$)
150: $n=2, m=2$, saturated ($K_D > 2 \mu\text{M}$)



156



157

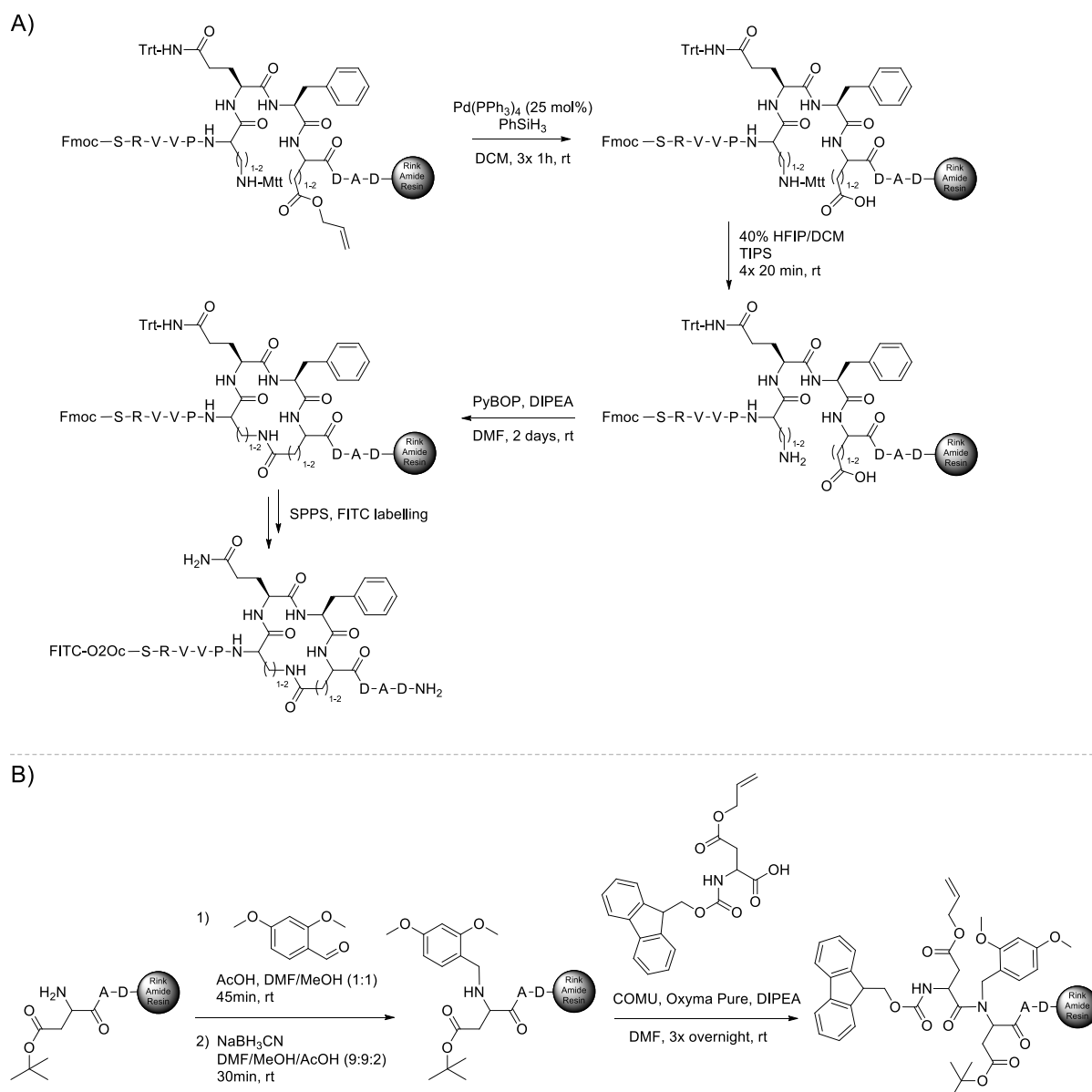


151: $n=1, m=2$, unsaturated ($K_D > 2 \mu\text{M}$)
152: $n=2, m=1$, unsaturated ($K_D > 2 \mu\text{M}$)
153: $n=2, m=2$, unsaturated ($K_D > 2 \mu\text{M}$)
154: $n=2, m=1$, saturated ($K_D > 2 \mu\text{M}$)
155: $n=2, m=2$, saturated ($K_D > 2 \mu\text{M}$)

Ac-V-P-G-Q-F-D-D-A-D-K(FITC)-NH₂
158: ($K_D = 670 \pm 187 \text{ nM}$)

Ac-V-P-G-Q-F-D-D-A-D-NH₂
159

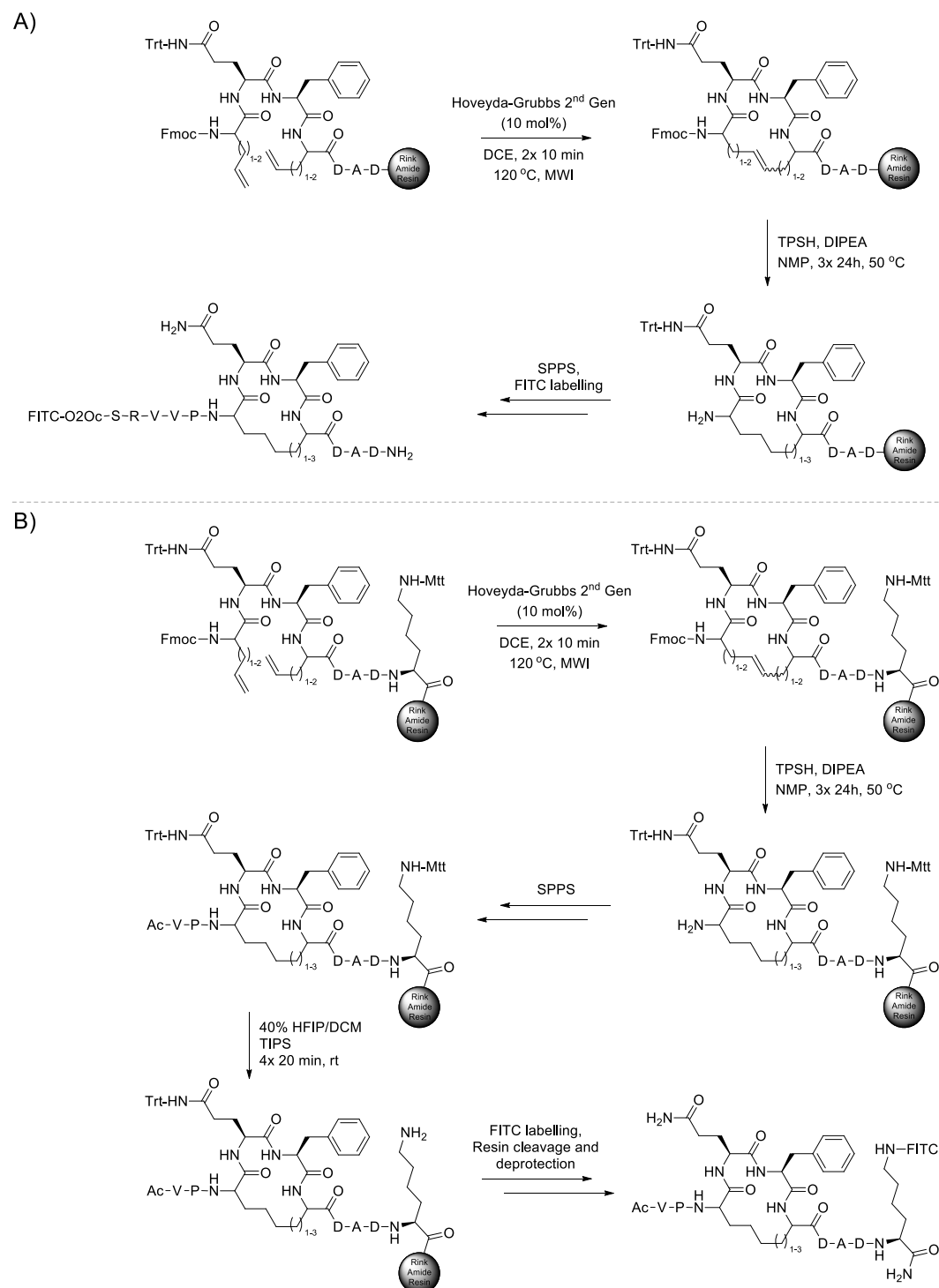
Figure 39. Structures of macrocyclic PAPIIs and results of direct binding FP with the PRMT5-MEP50 complex: **A)** First generation of cyclic PAPIIs. **B)** Second generation of cyclic PAPIIs and linear peptide analogues. FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker



Scheme 9. Synthesis of first generation PAPIIs incorporating amide-based linkers: **A)** Cyclisation strategy for amide-based macrocycles. **B)** On-resin protection with Dmb and subsequent coupling to Fmoc-Asp(OAll)-OH. Backbone protection with Dmb was necessary in order to prevent the aspartimide formation by Asp(OAll). FITC = fluorescein isothiocyanate, O2Ooc = 8-amino-3,6-dioxaoctanoic linker

The first generation of macrocycles (**133-144**) was tested for binding to PRMT5-MEP50. The best affinity was displayed by compounds obtained through the RCM cyclisation incorporating an L-amino acid in the position of Asp17: **141** (K_D of 719 ± 41 nM) containing a double bond in the ring, and **142** (K_D of 1.29 ± 0.11 μ M) with the saturated hydrocarbon linker. The analogous compounds **143** and **144** with a D-amino acid in place of Asp17 showed no binding (Figure S15), confirming the earlier experimental outcomes observed for sequence **132** (Table 9). The amide-based cyclopeptides performed poorly, showing only weak to negligible interactions with PRMT5 under the applied assay conditions. A trend, however, was observed, where the five atom long amide-based linkers performed

better than the four and six atom long connections (Figure S15). These results suggested that further inhibitor development should focus on the better performing RCM-based linkers, and that the linker length is a crucial factor which needs to be optimised.



Scheme 10. Synthesis of PAPIIs containing RCM-based linkers: **A)** Synthesis of the first generation PAPIIs. The double bond reduction with TPSH was omitted for compounds with unsaturated linkers. **B)** Synthesis of the second generation PAPIIs. The double bond reduction with TPSH was omitted for compounds with unsaturated linkers. FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker.

The design of the second generation macrocyclic PAPIIs took also into consideration the previously observed affinity improvement caused by the truncation of the N-terminal SRV motif and the transfer of the FITC tag to the C-terminal part of the peptide (compare peptides **86**, **101** and **105**, Table 6 in section 5.2.1.). The new set of fluorescently labelled macrocycles **145-155** (Figure 39B) was synthesised in an analogous manner to the first generation RCM compounds **141-144**, but a Mtt protected Lys residue was placed at the C-terminal end of the peptide, acting as a connector and allowing for the labelling with FITC before the final peptide cleavage from resin and global side chain deprotection (Scheme 10B). Cyclopeptides **145-155** contained between four and six carbon atoms in the linker, prepared either as unsaturated (**145-148** and **151-153**) or saturated hydrocarbon chain (**149**, **150**, **154** and **155**), with varying stereochemistry of the ring. In case of the unsaturated linkers, when possible, the position of the double bond was placed at various distances from the peptide backbone (compare **146** to **147** and **151** to **152**), in order to explore varying ring conformations. The macrocycles with D-amino acid at the original position of Asp17 (**151-155**) showed only weak binding to PRMT5-MEP50 (Figure S16). The best results were obtained for peptides with five atoms in the linker and an L-amino acid in place of Asp17, giving similar K_D values of 352 ± 49 nM and 430 ± 96 nM for **146** (one of the isolated isomers) and **149**, respectively. Both compounds achieved improved potency over the linear analogue **158** (K_D of 670 ± 187 nM), which was tested under the same conditions and using the same protein batch for the consistency of the results (Figure 39B). Considering that only one of the very challenging to isolate double bond isomers of compound **146** could bind with high potency to the target protein, it was decided to continue further experimentation requiring larger amounts of the compound with the scaffold of peptide **149**. It was noted that the experimental binding results, where only compounds with four or five atoms in the linker showed potent affinity towards PRMT5, are in agreement with the initial conclusions drawn based on the computational cyclopeptide models.

6.2.2. Evaluation of the Second Generation Macrocyclic PAPIIs

The cyclisation of the linear RioK1-derived sequences was expected to improve the peptide stability to proteolytic digestion. Resistance to proteases over an appropriate time period is an important feature necessary for the application of a compound in a number of experiments performed under biologically relevant conditions. To evaluate the compound stability, compound **156** - an unlabelled analogue of **149** (Figure 39B), was incubated in U2OS cell lysate up to seven days at 37 °C, and the linear peptide **159** was tested under the same conditions for comparison (Figure 40). The macrocycle showed remarkable stability with the estimated half-life of 12.5 days, whereas, the linear peptide was

characterised by the half-life of only 4.4 hours. The cyclisation resulted in a 70-fold stability increase, justifying the assumed macrocyclic design.

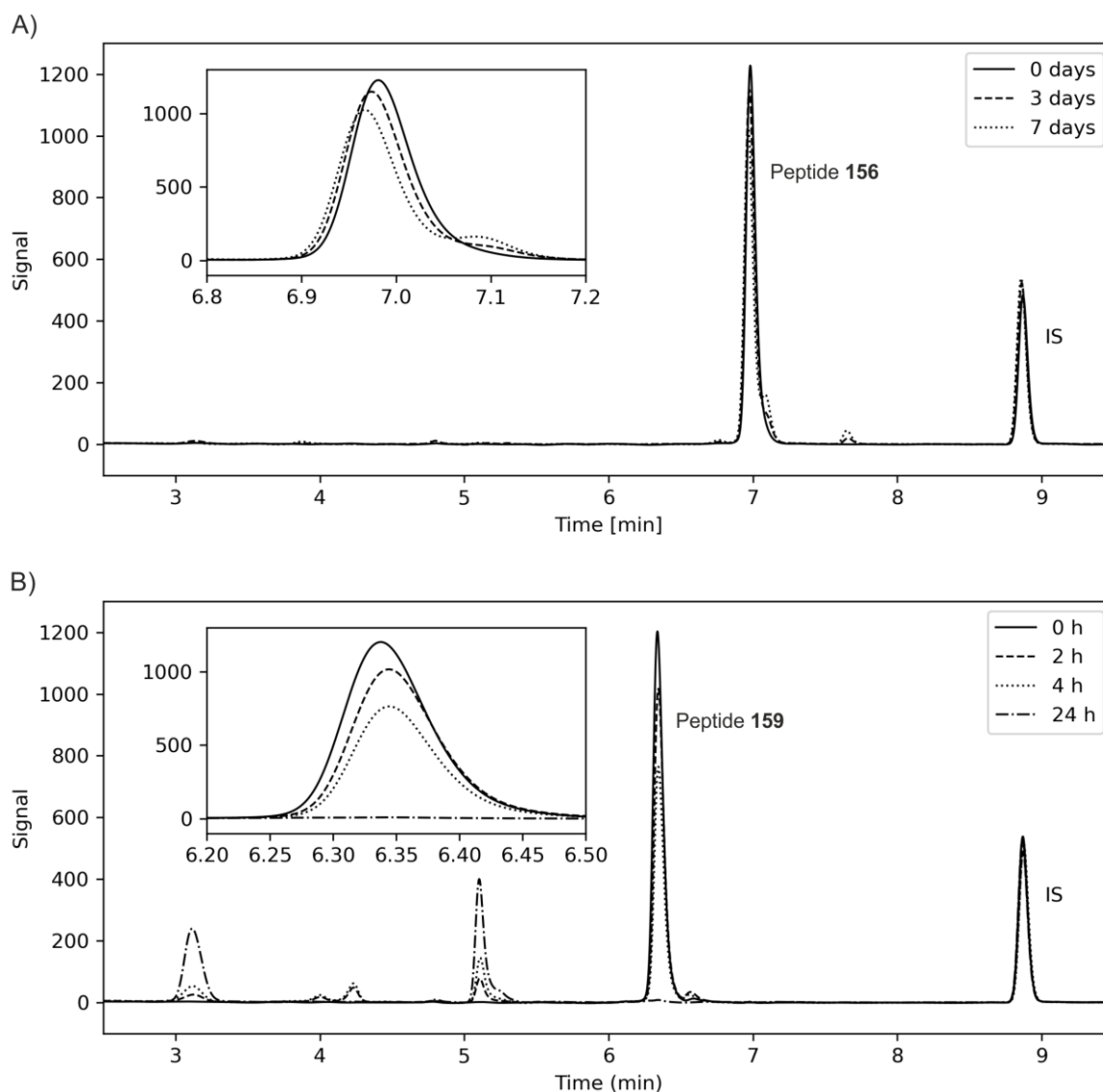


Figure 40. HPLC analysis of peptide samples obtained from the stability assay performed in U2OS cell lysate: **A)** Overlay of selected chromatograms for cyclopeptide **156** samples collected between 0-7 days. A significant portion of the peptide remains intact even after 7 days of incubation in the lysate. **B)** Overlay of selected chromatograms for linear peptide **159** samples collected between 0-24h. The peptide is completely degraded after 24h. IS: internal standard.

The macrocycle evaluation was followed by confirming that **156** does not interfere with the active site of the methyltransferase, allowing for the uninhibited direct methylation of the enzyme substrates. The cyclopeptide was analysed using the luminescence-based MTase-Glo™ assay and commercial PRMT5 active site inhibitor EPZ015666 as a control. While EPZ015666 resulted in a potent suppression of H4 histone tail methylation by the enzyme, compound **156** had no observable effect on the methyltransferase activity (Figure 41A). This expected, yet very crucial observation allowed for the continuation of the experimental analysis with **156**.

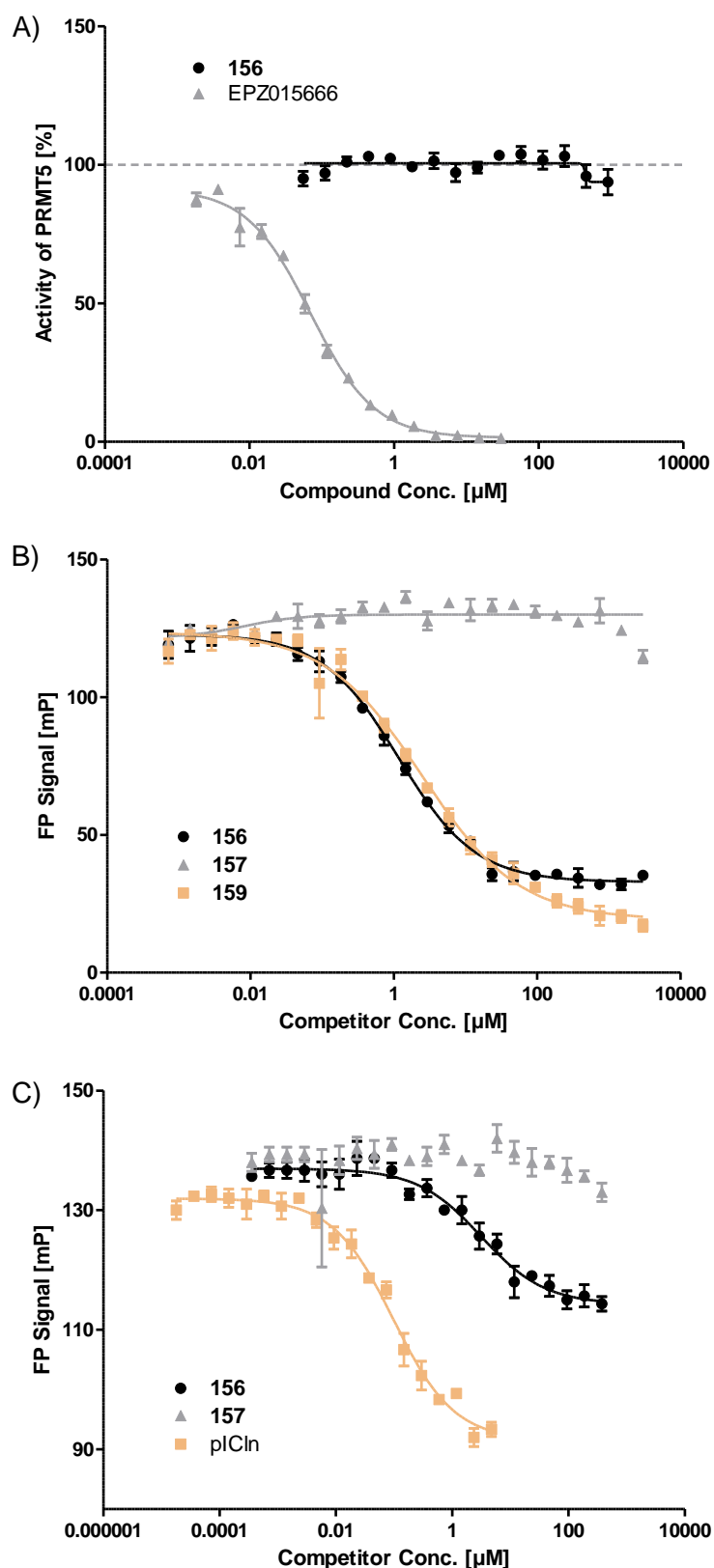


Figure 41. Results of enzymatic and biophysical assays for the second generation PAPII: **A)** Results for MTase-GloTM methyltransferase assay performed with the full length PRMT5-MEP50 complex and H4 histone tail peptide as a substrate. Active site inhibitor EPZ015666 suppresses the enzymatic activity of PRMT5, whereas cyclopeptide **156** has no observable effect on the methylation rate. **B)** Results for competitive FP assay with the full length PRMT5-MEP50 complex and compound **149** used as a tracer. **C)** Results for competitive FP assay with the full length PRMT5-MEP50 complex and Alexa488-labelled pICln protein used as a tracer.

Macrocycle **156** and linear peptide **159** were also tested in an FP experiment competing for binding to the PRMT5-MEP50 complex against fluorescently labelled cyclic **149**. Compound **156** afforded an IC_{50} of $1.17 \pm 0.29 \mu\text{M}$ and **159** gave a higher IC_{50} of $2.49 \pm 0.22 \mu\text{M}$. Scrambled control compound **157** could not compete with the fluorescent probe (Figure 41B). The experiment proved that the cyclic and linear peptides bind to the same site on PRMT5, as well as, that the FITC labelling does not intervene with the enzyme interaction. The cyclopeptide was examined for the ability to compete with the full-length pICln protein conjugated to Alexa488 dye for binding to PRMT5-MEP50. Compound **156** was capable of suppressing the very strong interaction between pICln and the PRMT5-MEP50 complex with an IC_{50} of $3.2 \pm 2.3 \mu\text{M}$, but the scrambled control **157** did not show activity against the targeted PPI (Figure 41C). Despite many attempts, execution of a similar experiment utilising fluorescently labelled RioK1 failed, most likely due to the excessive aggregation of the proteins.

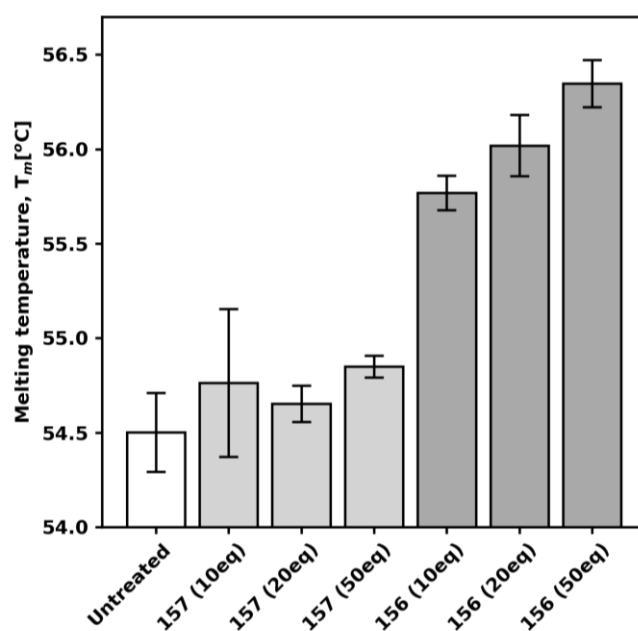


Figure 42. Results of nanoDSF analysis for the second generation cyclic PAPII.

The interaction between macrocycle **156** and the full-length PRMT5-MEP50 complex was examined using nanoscale differential scanning fluorimetry (nanoDSF). The analysis was performed by Dr. H el ene Adihou. These measurements revealed a clear, dose dependent stabilisation effect of **156**, but not for scrambled cyclic analogue **157**. In comparison to the untreated PRMT5-MEP50 complex, 10 equiv. of **156** induced a ΔT_m of $1.27 \pm 0.09 \text{ }^\circ\text{C}$ and in the presence of 50 equiv. ΔT_m reached $1.85 \pm 0.12 \text{ }^\circ\text{C}$ (Figure 42).

Attempts were also made to prepare a co-crystal structure of the macrocyclic PAPIIs with PRMT5, however, no crystals were obtained.

In order to evaluate the effect of macrocycle **156** on protein methylation in cells, the macrocycle, due to its structural characteristics, needed to be equipped with a cell penetrating peptide (CPP) sequence to facilitate permeability through membranes. The CPP of choice was the popular TAT motif which was attached to the peptide through the Lys residue placed on the C-terminus of the cyclopeptide. The CPP was attached either directly to Lys (**160**), or a O2Oc linker was inserted between Lys and the TAT sequence (**161**, Figure 43). The compounds were synthesised completely on resin using the same strategy as outlined in Scheme 10B. The resin used, however, was not the standard polystyrene-based rink amide, but fully polyethylene glycol-based resin ChemMatrix®.^[406] The application of ChemMarix® was needed due to the presence of the difficult to synthesise, poly-Arg TAT sequence. Additionally, the amino acid coupling steps during the TAT sequence construction were performed using the strong activator COMU, together with oxyma and DIPEA in DMF containing 0.4 M LiCl solution acting as a chaotropic agent. The LiCl solution was used to further improve the efficiency of the synthesis by breaking potential peptidyl-aggregates on the solid phase support.^[438–441]

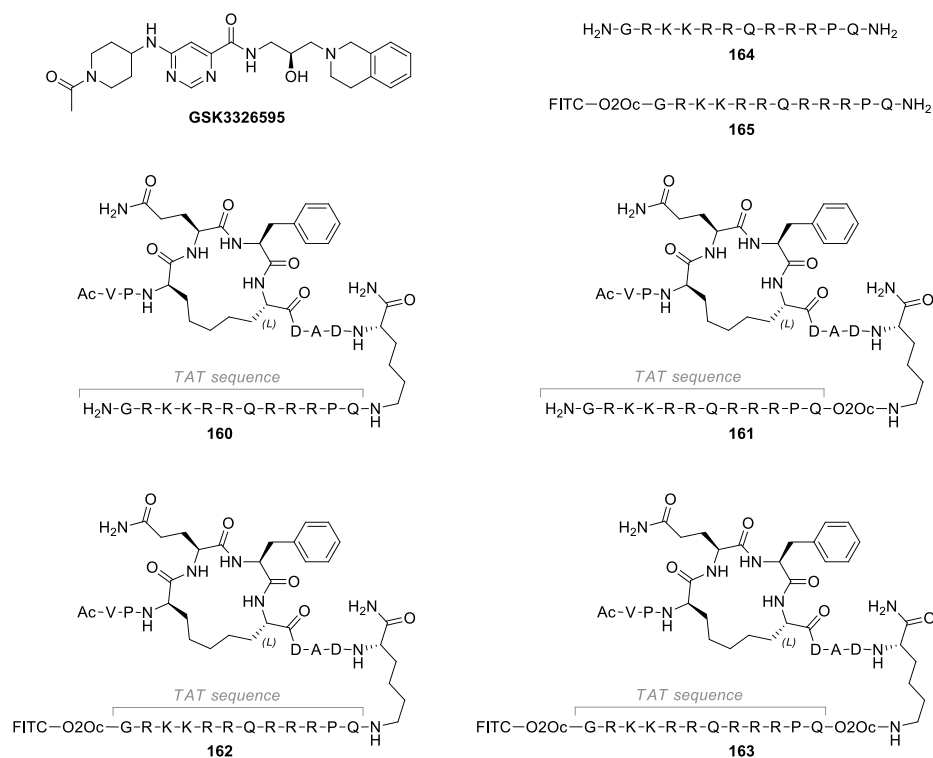


Figure 43. Structures of compounds intended for biological analysis. FITC = fluorescein isothiocyanate, O2Oc = 8-amino-3,6-dioxaoctanoic linker

To confirm that the TAT-conjugated cyclopeptides can still effectively bind to PRMT5, fluorescently labelled compounds **162** and **163** were made (Figure 43) and tested in a direct binding FP experiment with the PRMT5-MEP50 complex *in vitro*. The compounds could bind to the enzyme with the K_D value of 249 ± 27 nM for **162** and 389 ± 13 nM for **163**, proving that the CPP does not negatively interfere

with the investigated cyclopeptide-protein interaction. The fluorescently labelled control TAT sequence **165** showed no meaningful binding to the complex (Figure 44).

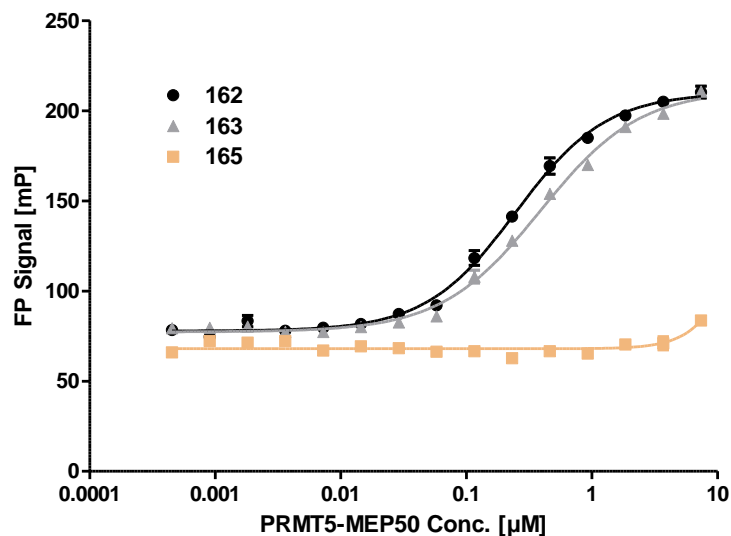


Figure 44. Results of the direct binding FP assay with the PRMT5-MEP50 complex.

In order to determine if a treatment of intact cells with cyclopeptide **160** or **161** has an effect on the SDMA protein methylation pattern, initial experiments were performed using HECK293T and MCF7 cells with the concentration of the macrocycles in the cell medium up to 50 µM. Whereas the commercial PRMT5 active site inhibitor GSK3326595 could nearly completely suppress the antibody-detected SDMA methylation of proteins, **160** and **161** showed no impact on the SDMA pattern in comparison to the control TAT sequence **164** (Figure S17). Additionally, Dr. Christoph Peter and Lea Esser of Heinrich Heine University Düsseldorf performed indicative immunoprecipitation (IP) experiments intended to analyse the effectiveness of the synthesised cyclopeptides to impact the PPI between PRMT5 and the adaptor proteins. Intact cells overexpressing either GFP-RioK1 or GFP-PRMT5 were treated with **156**, **160** or **161** at concentrations up to 100 µM, followed by the IP with anti-GFP antibody and detection of the PRMT5, MEP50, RioK1 and pICln proteins. Based on the initially afforded results, it was concluded that the tested compounds exhibited no observable *in cellulo* effect on the targeted PPI. The IP results obtained by the collaborators were in line with the in-house observations for the SDMA protein methylation, suggesting that either the cell-penetration ability of the peptides was insufficient to exert an observable *in cellulo* effect or that the affinity of the cyclopeptides to PRMT5 was not high enough to effectively compete with the strongly binding adaptor proteins in the native intracellular environment. Thus, the collaborators further tested **156** in lysates of cells overexpressing GFP-RioK1, GFP-pICln, GFP-MEP50 or GFP-PRMT5, eliminating the need for the cyclopeptide to penetrate through a membrane, yet retaining the highly biologically relevant context of native cellular composition and PPIs. Based on the initial IP results obtained from the cell lysates

only a very weak and disputable inhibitory effect was observed onto the PRMT5-RioK1 interaction in the lysate with overexpressed GFP-PRMT5, which further suggested that an effective PPI inhibition in cells would require higher affinity to PRMT5 than what was achieved by the existing scaffold of **156**.

6.2.3. Optimisation of PAPIIs

Due to the very weak and arguable effects exerted on the PRMT5-adaptor protein PPIs in biologically relevant environment by the second generation PAPIIs, and no observable effect on intact cells, it became apparent that a sequence optimisation was needed in order to improve the cyclopeptide affinity to the target. To this end, 43 different modifications to the side chains of amino acids were envisaged (Figure 45), where the choice of the modifications was roughly directed by the previously obtained co-crystal structure of the peptide-bound TIM barrel. Most of the modifications focused on the Phe residue due to the relatively exposed location of its side chain in the PRMT5 groove providing opportunity for extending the interactions with the surface of the protein and generally good availability of the commercial Phe derivatives. The proteogenic and non-proteogenic modifications were incorporated into the scaffold of linear peptide **158**, and the resulting peptides were tested in a direct binding FP assay for interactions with PRMT5-MEP50. Ten of the tested compounds gave K_D values better than the K_D for the native sequence of **158** (670 nM), where in each case the beneficial modification was either to Phe16 or Asp20 (Table 10 and S5, Figure S18-S21). The highest affinities were achieved by the Asp20 modification to Gla (**174**, K_D of 245 nM) or by replacing Phe16 by either Phe(4-Cl) (**173**, K_D of 250 nM) or Phe(4-NO₂) (**175**, K_D of 232 nM), in each case increasing the affinity three-fold. The beneficial effect of Gla incorporation onto the peptide affinity can be easily explained by the increased negative charge enhancing the interactions with the electrostatically positive region of PRMT5 and presumably with Lys240 (Figure 35B and 36A, section 5.2.3.). The exact reason for the strongly positive impact of attaching the chloro and nitro group to the Phe residue is not clear, despite attempts to computationally model the modified peptide with PRMT5. However, it can be assumed that at least part of the reason is the decrease in the electron density of the aromatic ring, considering all the alternative, beneficial variations to Phe (Table 10). It is also evident that Phe ring substitution at position 4 was strongly preferred over the other locations. The most potent modifications were combined together in compounds **176** and **177**, affording K_D values of 144 ± 46 nM and 78 ± 13 nM, respectively (Table 10). Thus, compound **177** incorporating both Phe(4-NO₂) and Gla showed nearly nine-fold affinity increase when compared to the original sequence of **158**.

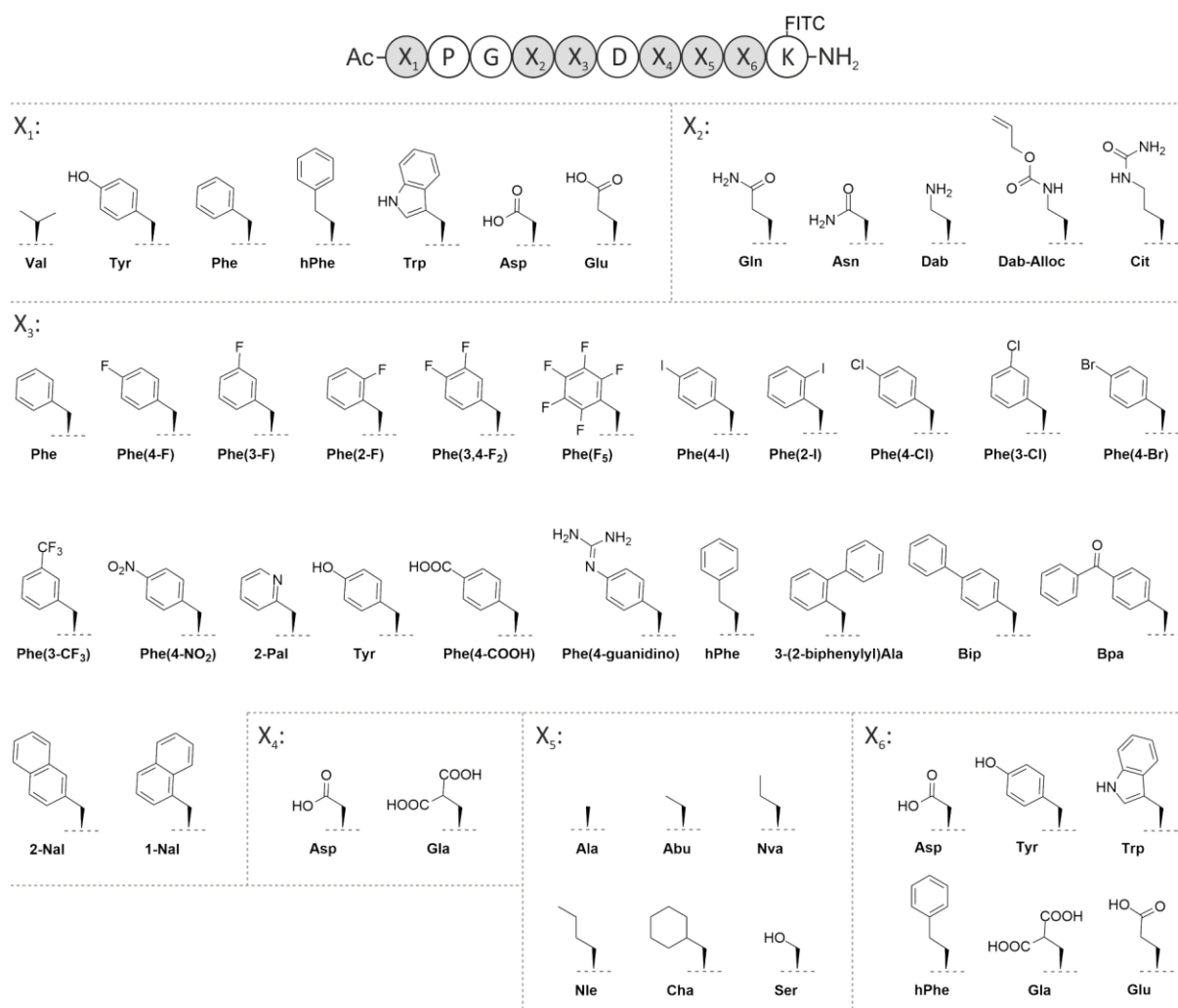


Figure 45. Overview of the explored amino acid modifications to the original RioK1-derived sequence of compound **158**. FITC = fluorescein isothiocyanate.

Table 10. Overview of the amino acid modification improving peptide affinity towards PRMT5-MEP50 over the original sequence of **158**.

Peptide	Modified Amino Acid	Sequence	K _D (nM) ^[a]
158	–	Ac-VPGQFDDADK (FITC) -NH ₂	670 ± 187
166	X ₃	Ac-VPGQ- Phe (3-F) -DDADK (FITC) -NH ₂	654 ± 23
167	X ₆	Ac-VPGQFDDA- Glu -K (FITC) -NH ₂	531 ± 23
168	X ₃	Ac-VPGQ- Bip -DDADK (FITC) -NH ₂	447 ± 41
169	X ₃	Ac-VPGQ- Phe (3, 4-F₂) -DDADK (FITC) -NH ₂	412 ± 31
170	X ₃	Ac-VPGQ- Phe (4-F) -DDADK (FITC) -NH ₂	391 ± 35
171	X ₃	Ac-VPGQ- Phe (4-Br) -DDADK (FITC) -NH ₂	317 ± 17
172	X ₃	Ac-VPGQ- Phe (4-I) -DDADK (FITC) -NH ₂	279 ± 22
173	X ₃	Ac-VPGQ- Phe (4-Cl) -DDADK (FITC) -NH ₂	250 ± 6
174	X ₆	Ac-VPGQFDDA- Gla -K (FITC) -NH ₂	245 ± 45
175	X ₃	Ac-VPGQ- Phe (4-NO₂) -DDADK (FITC) -NH ₂	232 ± 46
176	X ₃ and X ₆	Ac-VPGQ- Phe (4-Cl) -DDA- Gla -K (FITC) -NH ₂	144 ± 46
177	X ₃ and X ₆	Ac-VPGQ- Phe (4-NO₂) -DDA- Gla -K (FITC) -NH ₂	78 ± 13

^[a]As determined with FP using the native PRMT5-MEP50 complex. FITC = fluorescein isothiocyanate.

As backbone N-methylation can positively contribute to the favourable binding conformation of a peptide, and thus, increase the affinity to target, several amide bond N-methylations were explored.^[442,443] The methyl group was placed only on the amide bonds which were not involved in the direct peptide-protein intermolecular H-bonds (Table 10). Compounds **178-183** were obtained through standard SPPS using commercially available Fmoc protected and N-methylated amino acids. The N-terminal methylation of the peptide sequence in the synthesis of **183** proved challenging as during the acid cleavage and deprotection, the N-acetylated and methylated Val residue was rapidly cleaved resulting in a deletion sequence and no desired product. Similar behaviour for N-terminally acetylated and methylated peptides in the presence of TFA was previously reported by Kim *et al.*, 2014, who circumvented the undesired transformation by simply acetylating the unprotected peptide after the TFA cleavage using Ac₂O and DIPEA in DMF.^[444] Compound **183** was successfully afforded following the method by Kim *et al.* The synthesised N-methylated peptides were tested in a direct binding FP assay against the methyltransferase complex. Compound **178** showed similar affinity to PRMT5-MEP50 as unmodified **158**, and compounds **179-183** were less potent (Table 10 and Figure S22).

Table 10. Explored backbone N-methylations incorporated into the sequence of **158**, and the corresponding K_D values.

Peptide	Sequence	K _D (μM) ^[a]
178	Ac-VPGQFDDA-NMe D -K (FITC) -NH ₂	0.6 ± 0.1
179	Ac-VPGQFDD-NMe A -DK (FITC) -NH ₂	>2
180	Ac-VPGQFD-NMe D -ADK (FITC) -NH ₂	>2
181	Ac-VPGQ-NMe F -DDADK (FITC) -NH ₂	>2
182	Ac-VP-NMe G -QFDDADK (FITC) -NH ₂	>2
183	Ac-NMe V -PGQFDDADK (FITC) -NH ₂	>2

^[a]As determined with FP using the native PRMT5-MEP50 complex. FITC = fluorescein isothiocyanate.

The sequence modifications positively contributing to PRMT5 affinity were incorporated into the previously developed macrocyclic scaffold of **149**, resulting in the third generation PAPII compound **184**, which binds to the methyltransferase complex with the K_D value of 89 ± 11 nM. The scrambled control macrocycle **185** did not show meaningful affinity to the target (Figure 46 and 47). The final optimisation idea concerned the length of the N-terminal overhang in the cyclopeptide. It was hypothesised that the N-terminal Val-Pro motif was superfluous, due to the conformation of the compound being sufficiently stabilised by the ring structure. To this end, Val and Pro were removed from **184**, giving truncated cyclic **186** (Figure 46). The N-terminal motif proved however to be crucial for the effective interaction with PRMT5, as **186** showed significantly impaired affinity to the protein (Figure 47).

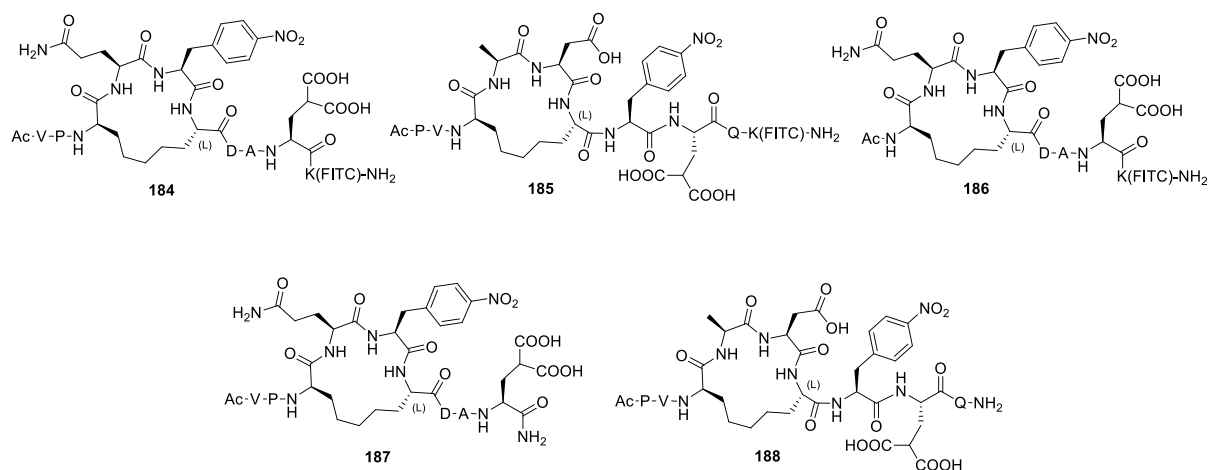


Figure 46. Structures of the third generation cyclic PAPIIs. FITC = fluorescein isothiocyanate.

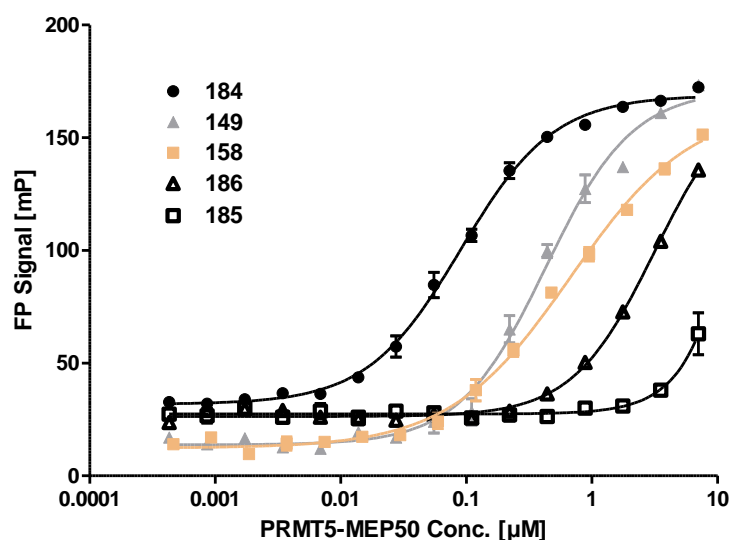


Figure 47. Assessment of the second and third generation PAPIIs in the direct binding FP with the PRMT5-MEP50 complex.

6.2.4. Evaluation of the Third Generation Macrocylic PAPIIs

The unlabelled analogue of the most optimal cyclopeptide **184**, compound **187**, was tested in a competitive FP assay (using **184** as tracer) and compared to the second generation PAPII **156** and linear sequence **159**. The assay afforded an IC₅₀ of 356 ± 59 nM for **187**, 1.06 ± 0.24 μM for **156** and 5.3 ± 3.7 μM for linear compound **159**, proving that **187** interacts with the same binding site as the previously synthesised constructs, and does not show any indications of nonspecific interactions with the protein (Figure 48A).

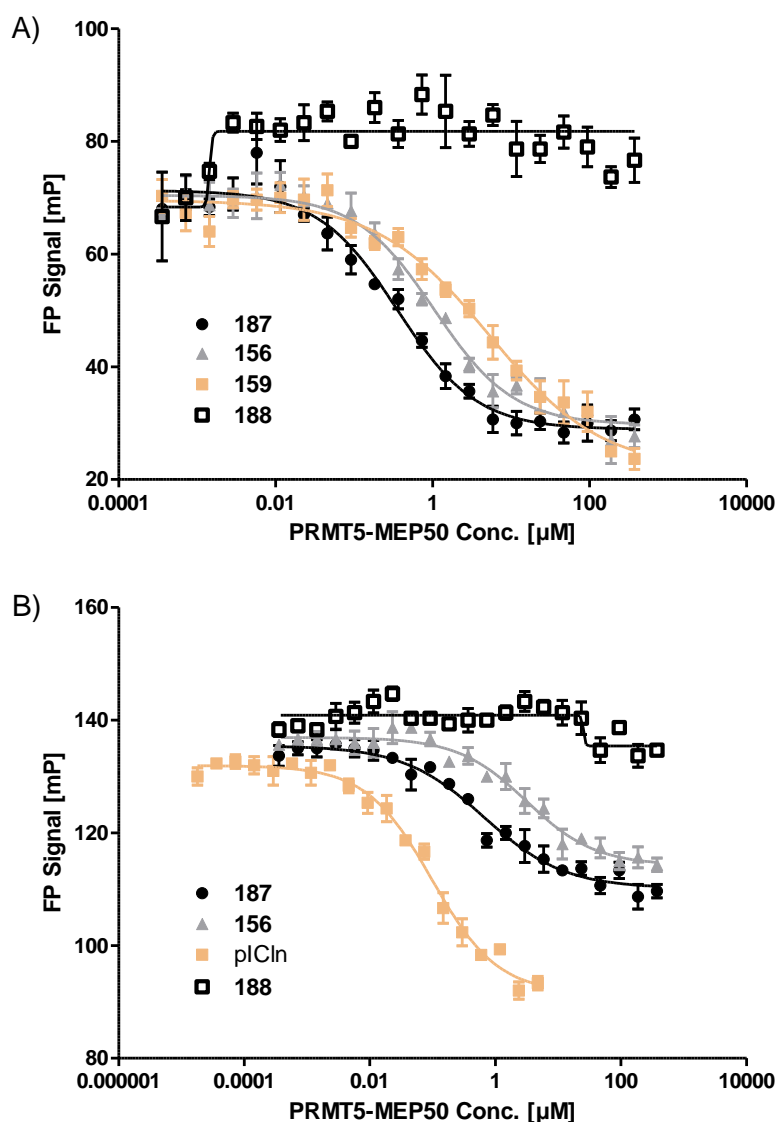


Figure 48. Results of the competitive binding FP assays with the full length PRMT5-MEP50 complex: **A)** Competitive FP assay using the third generation fluorescent PAPII **184** as a tracer. **B)** Competitive FP assay using the Alexa488-labelled pICln protein as a tracer.

The unlabelled compounds were also tested *in vitro* for the ability to disrupt the PPI between PRMT5-MEP50 and the fluorescently labelled full-length pICln protein. Cyclopeptide **187** proved to be a significantly more effective PPI inhibitor (IC_{50} value of 654 ± 476 nM) than the second generation PAPII **156** (IC_{50} value of 3.2 ± 2.36 μM). Scrambled control **188** did not disrupt the PPI (Figure 48B). The compounds were also tested using the thermal shift assay (TSA) with the truncated TIM-MEP50 protein construct by Dr. Anthony Willaume and Dr. Renaud Prudent from Edelis. Treatment with compound **187** revealed a clear stabilisation effect on the protein complex resulting in an average melting temperature increased by 2.27 $^{\circ}\text{C}$. No stabilisation was observed when the control **188** was tested (Figure 49A).

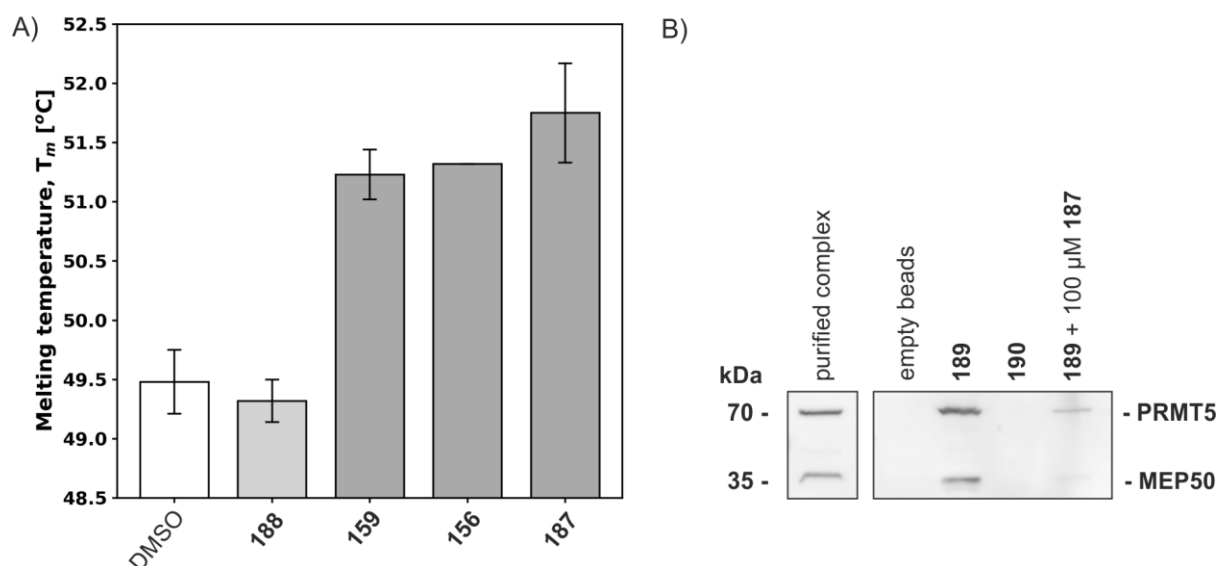
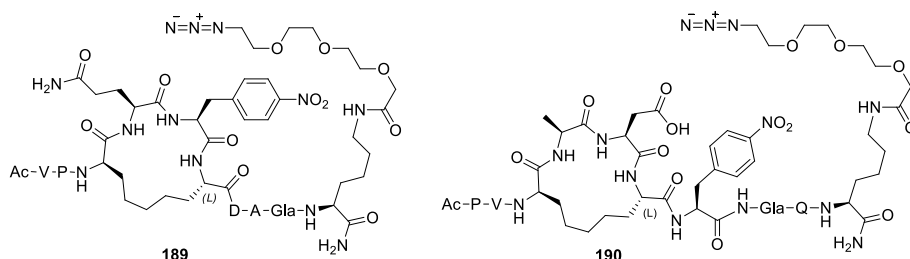


Figure 49. Evaluation of PRMT5-PAPII interactions: **A)** Results of the TSA analysis of linear and cyclic peptides with the TIM-MEP50 complex. **B)** Pull-down assay using MCF7 cell lysate. Active probe **189** and scrambled control macrocycle **190** were immobilised on DBCO beads. Immobilised **189** was also tested in the competitive conditions of lysate containing 100 μ M of non-immobilised active PRMT5 binder **187**.

In order to prove that compound **187** binds to PRMT5-MEP50 under biologically relevant conditions, active compound **189** and scrambled cyclopeptide **190**, equipped with a C-terminally placed azide group were employed to synthesise affinity isolation probes (Figure 50A). The compounds were coupled to dibenzocyclooctyne (DBCO) agarose beads via a copper-free 1,3-dipolar cycloaddition reaction. The beads containing immobilised **189**, **190**, as well as the empty DBCO beads were incubated with MCF7 cell lysate. The active probe **189** was able to successfully enrich the target PRMT5 and MEP50, whereas the empty beads and scrambled control **190** did not show interactions with the protein complex (Figure 49B). As expected, there was no observed suppression of the interaction within the complex between PRMT5 and MEP50 caused by **189**. The experiment was also performed using competitive conditions where immobilised **189** was incubated in cellular lysate containing 100 μ M of non-immobilised **187**. Compound **187** could compete with the immobilised active probe, however, even at 100 μ M concentration residual interactions with PRMT5-MEP50 could still be observed (Figure 49B). This result can be explained by the avidity effect, where the apparent interaction strength results from multiple interactions between all the binding partners.^[445] As the PRMT5-MEP50 heterooctamer has a capacity to interact simultaneously with four active **189** probes, in order for the complex to be released from the beads, **187** needs to displace all attached **189** anchors. As it is known that avidity can significantly increase the biological effectiveness of employed agents, it is expected that as indicated in the pull-down experiment, multivalent PRMT5-MEP50 binders would be significantly more effective at inhibiting PRMT5/adaptor protein interactions than the current generation of presented here monovalent PAPIIs.^[445,446] A proposition of dimeric PRMT5

binders (**191** and **192**) based on the structure of **187** and **189** is depicted in Figure 50B. The macrocycles could be connected to each other via a simple PEG linker attached to either the N- or C-termini through click reactions. An expected downside of such structures would be a very poor and difficult to optimise cell permeability, where most likely the compounds would need to be administered in combination with cellular shuttles facilitating delivery to the cytosol.^[147,174–178] The development and biological assessment of multivalent PRMT5 binders, however, is left for future explorations.

A)



B)

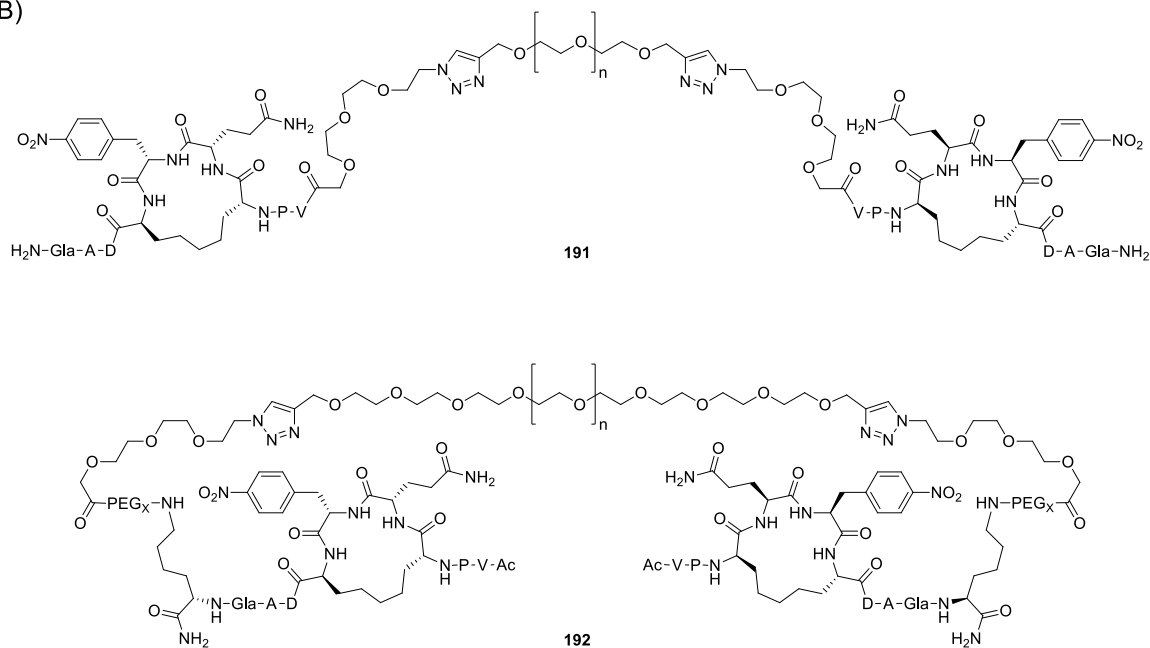


Figure 50. Structures of PAPII probes: **A)** Azide functionalised third generation PAPIIs. Compound **189** is an active PRMT5 binder, whereas **190** is a scrambled control. **B)** Proposed structures of multivalent, dimeric PRMT5/adaptor protein inhibitors, based on the third generation PAPIIs.

Compound **187** and its scrambled equivalent **188** were tested by the collaborators from the Heinrich Heine University Düsseldorf for the inhibition of the PPI between PRMT5-MEP50 and the adaptor proteins RioK1 and pICln in cellular lysates. Conjugate GFP-PRMT5 was overexpressed in Flp-In T-REx 293 cells, which was followed by compound incubation in lysate, and then IP with an anti-GFP antibody. Compounds **187**, **188** and DMSO control showed enrichment of PRMT5 and MEP50. The treatment with control cyclopeptide **188** and DMSO lead also to the enrichment of RioK1 and pICln,

but the application of active compound **187** successfully suppressed the PPI between PRMT5 and RioK1. Surprisingly, **187** had little effect on the enrichment of pICln, thus, suggesting adaptor protein inhibition selectivity in biologically relevant media (Figure 51A). In order to validate these results, the collaborators also repeated the experiments overexpressing GFP-pICln and GFP-RioK1 and performing IP to enrich these adaptor proteins and their binding partners. Active PAPII **187** was again able to fully inhibit the PPI between RioK1 and PRMT5-MEP50 but not the interactions concerning pICln (Figure 51B). A possible explanation of this behaviour could be the existence of an additional binding site on pICln which is not targeted by **187**, as it was previously suggested in the published literature by Pesiridis *et al.*^[381] Compound **187** is however effective at inhibiting the interactions between PRMT5-MEP50 and pICln *in vitro*, as shown in the presented here FP assay (Figure 48B), which could suggest that additional proteins could contribute to and strengthen the interaction between PRMT5-MEP50 and pICln under more native and biologically relevant conditions.

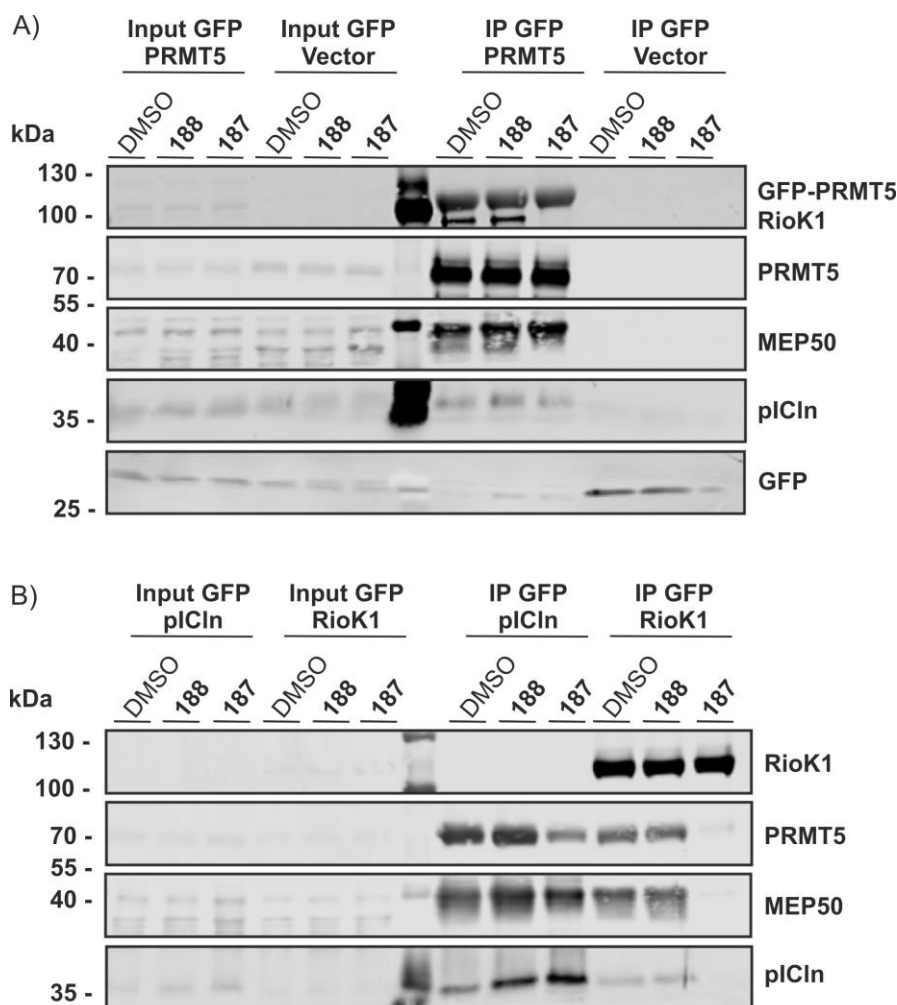


Figure 51. IP assays using cell lysates: **A)** GFP-IP from lysate of Flp-In T-Rex 293-GFP and Flp-In T-Rex 293-GFP-PRMT5 overexpressing cells, evaluating compound **187** and scrambled **188** at 50 μ M. **B)** GFP-IP from Flp-In T-Rex 293-GFP-pICln and Flp-In T-Rex 293-GFP-RioK1 cytoplasmic extract, testing **187** and **188** at 50 μ M.

The inadvertent selectivity towards RioK1 over pICln by the developed here PAPII can be regarded as an advantage, considering that the only other known PRMT5/adaptor protein PPI inhibitor BRD0639, binding the interaction site covalently, does not appear to show such discriminatory ability.^[437] This compound developed here can constructively contribute to future investigations concerning the biochemistry of PRMT5 and its adaptor proteins.

More work, however, will need to be done in order to render cyclopeptide **187** cell permeable, and thus, allowing to utilise this compound in investigations involving intact cells. It is proposed that in future biological studies **187** is used in conjunction with shuttle peptides.^[147,174–178] Although this approach would require precise optimisation of the applied assay conditions, it would allow to use **187** without additional and presumably very arduous structural modifications.

6.3. Summary and Conclusions

The previously obtained co-crystal structure of the RioK1 peptide and the TIM barrel domain, as well as the biophysical data collected for linear adaptor peptides (PART C) in conjunction with computer-based models were used for designing a series of macrocyclic PAPIIs. A number of different ring sizes and linkers with varying stereochemistry were explored, resulting in an optimal scaffold of compound **149**, containing five-atom long saturated hydrocarbon linker. Cyclopeptide **149** has a K_D value of 430 nM when measured with the target complex PRMT5-MEP50, exhibiting also a stabilising effect onto the protein complex. The optimised scaffold showed remarkable proteolytic stability in cell lysate with the estimated half-life 70-fold higher compared to the linear analogue. Unlabelled compound **156** does not interfere with the direct methyltransferase activity of PRMT5, but it can inhibit the PPI between purified PRMT5 and pICln proteins with an IC_{50} value of 3.2 μ M when tested *in vitro*. Testing the compound under more biologically relevant conditions revealed a need to improve the cyclopeptide affinity. A number of amino acid modifications were explored, resulting in optimised compound **184** binding to PRMT5-MEP50 with a significantly improved K_D of 89 nM. The unlabelled analogue **187** was also capable of disrupting the PRMT5-pICln interaction with an IC_{50} value of 654 nM, when tested in an FP assay. Compound engagement with the PRMT5-MEP50 complex was further confirmed by pull-down experiments from cell lysates. A strong avidity effect was observed suggesting that multivalent PRMT5 binders could be potentially very effective agents for inhibiting PRMT5/adaptor protein interactions, indicating a pathway for future developments. IP experiments with compound **187** showed an unexpected selectivity for inhibition of the methyltransferase interactions with RioK1 over pICln. This surprising selectivity suggests still incomplete understanding of the nature of the PRMT5-pICln interactions. Although future optimisations are necessary in order to render compound **187** cell permeable, the presented herein results demonstrate a valid approach to developing potent macrocyclic inhibitors of challenging PPIs. The here afforded PAPII can be used in future biochemical investigations, potentially allowing to gain a better understanding of the function and regulation mechanisms governing the activity of PRMT5.

Experimental Section

7.1. General Methods

All chemical reagents and solvents, unless stated otherwise, were purchased from commercial suppliers: Sigma-Aldrich, Merck, Novabiochem, Acros Organics, abcr, Activate Scientific, Fluorochem, Carbolution Chemicals, Carbosynth, Iris Biotech, Carl Roth, TCI Deutschland, Fisher Scientific, Fisher Chemicals, Biosolve, Chem-Impex, Serva Electrophoresis, Gerbu Biotechnik, Calbiochem, Thermo Scientific or VWR International. The air and moisture sensitive reactions were performed under an inert atmosphere of Ar gas. Organic solvents for moisture sensitive reactions were dried through storage over activated 3Å molecular sieves, for at least 24h.

Automated peptide synthesis was performed using either CEM Microwave Peptide Synthesiser connected to CEM Discover microwave reactor (CEM Corporation, USA) or Syro I peptide synthesiser (Multisynth GmbH, Germany). The RCM reactions were performed in the CEM Discover microwave reactor (CEM Corporation, USA). Preparative scale HPLC purification was carried out either on an Agilent Infinity or Infinity II LC-MS system equipped with a 125 mm x 21 mm, 5 µm or 125 mm x 10 mm, 5 µm Macherey-Nagel Nucleodur C18 Gravity column (Macherey-Nagel GmbH & Co. KG, Germany) and detection at 210 nm.

Purity of the final peptide products was determined at 210 nm with an Agilent Infinity HPLC system applying the elution system: 5% → 65% MeCN (0.1% TFA) in H₂O (0.1% TFA) over 14 min, using either 50 mm x 3 mm, 1.8 µm Macherey-Nagel Nucleodur C18 Gravity column and a flow rate of 0.56 ml/min (Method A) or 125 mm x 3 mm, 5 µm Macherey-Nagel Nucleodur C4 Gravity and a flow rate of 1 ml/min (Method B). Alternatively, an Agilent Infinity II HPLC system was used, equipped with 150 mm x 2.1 mm, 2.7 µm Agilent InfinityLab Poroshell 120 EC-C18, with a flow rate of 0.4 ml/min and the elution system 5% → 95% MeCN (0.1% TFA) in H₂O (0.1% TFA) over 20 min (Method C). HRMS analyses were performed on a LTQ Orbitrap mass spectrometer (Thermo Fisher Scientific, USA) using electrospray ionisation.

Concentrations of PRMT5-MEP50 were measured with the Bradford assay, using Protein Assay Dye Reagent from Bio-Rad (Cat. #500-0006; Bio-Rad Laboratories, USA) and a dilution series of BSA for calibration. The concentration of pICln and RioK1 proteins was determined using NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific, USA). Fluorescence- and luminescence-based readouts were performed on the Spark Multimode Microplate Reader (Tecan Trading AG, Switzerland). Fluorescence-based assays were carried out in black, low volume, round bottom 384-well plates (ref. 4514, Corning Incorporated, USA), and the luminescence-based assays in white, low volume, round

bottom 384-well plates (ref. 4512, Corning Incorporated). Mass photometry analyses were done on the OneMP mass photometer (Refeyn Ltd, UK). Flow induced dispersion analysis (FIDA) was performed using Fida 1 platform (FIDA Biosystems ApS, Denmark) with LED 488 nm detector and the analysis was done using FIDA software Version 2.01 (FIDA Biosystems ApS, Denmark).

The antibodies used for the pull-down and immunoprecipitation experiments were as follows: anti-PRMT5 antibody (sc-376937, Santa Cruz Biotechnology) and anti-MEP50 (2828S, Cell Signaling Technology). The detection was done using the fluorescent secondary antibodies: IRDye 680RD goat anti-rabbit and IRDye 800CW goat anti-mouse. The following primary antibodies were used for GFP immunopurification and immunoblotting: anti-PRMT5 (2252, CST), anti-RioK1 (NBP1-30103, Novus Biologicals) anti-MEP50 (2823, CST), anti-GFP (3H9, Chromotek), anti-pICln (sc-393525, Santa Cruz). The detection of proteins was carried out with the following fluorescent secondary antibodies: IRDye 680LT goat anti-rabbit, IRDye 800LT goat anti-mouse and IRDye 800CW goat anti-rat. The antibodies used for the assessment of the SDMA patterns were as follows: anti-SDMA (13222S, Cell Signaling Technology), anti-vinculin (V9131, Sigma) and anti-GAPDH (G8795, Sigma). The detection was done using the fluorescent secondary antibodies: IRDye 800CW donkey anti-rabbit and IRDye 680RD donkey anti-mouse.

7.2. Linear Peptide Synthesis

All linear peptide were obtained using standard Fmoc chemistry and solid-phase peptide synthesis (SPPS) methods.^[134] The peptides were normally synthesised on a polystyrene-based Rink Amide AM resin (substitution ca. 0.7 mmol/g, 100-200 mesh), with the exception of difficult to afford histone peptides **66-74**, **164** and **165**, which were synthesised on the polyethylene glycol-based Rink Amide ChemMatrix[®] resin (substitution ca. 0.5 mmol/g, 35-100 mesh).

Peptide synthesis was normally performed using 4 equiv of amino acid, 4 equiv of PyBOP and 8 equiv of DIPEA in DMF. The TAT-based sequences were synthesised using 4 equiv of amino acid, 3.6 equiv COMU, 4 equiv oxyma pure and 8 equiv of DIPEA in DMF with the addition of 0.4M LiCl. Fmoc removal was achieved with 20% piperidine solution in DMF, and where applicable, 0.5M oxyma pure was added to the solution, in order to suppress the aspartimide formation.

The coupling of amino acids on the CEM synthesiser was carried out for 5 min at 75 °C under microwave irradiation (MWI), with an exception for Arg residues, where a modified coupling method was applied (rt to 75 °C, 30 min under MWI, followed by a second coupling at 75 °C over 5 min under MWI) and during coupling of Cys and His where the maximum applied temperature was 50 °C. Fmoc removal was performed at 75 °C under MWI for 30s followed by a second deprotection at 75 °C for 3

min. The peptide synthesis conducted on the Syro synthesiser was done through double 50 min amino acid coupling at rt. Fmoc removal was done twice at rt for 5 min for each residue. Manual amino acid coupling was typically performed for 1h at rt with resin agitation induced by bubbling Ar gas through the reaction mixture. Fmoc removal was achieved by applying the piperidine solution over 5 min and then over 10 min. In all cases, the SPPS steps were always followed by resin washes with DMF.

Where applicable, acetylation of the resin-bound peptides was performed using Ac₂O (10 equiv) and DIPEA (12 equiv) in DMF over 30 min at rt.

In order to facilitate the C-terminal FITC labelling of the peptides, sidechains of the C-terminal lysines were protected by the acid sensitive Mtt group instead of the standard Boc group. Mtt was selectively removed by the treatment of the resin-bound peptides with 40% HFIP solution in DCM containing 2.5% TIPS, 4x 20 min at rt.

In all applicable cases, Fmoc-AEEA-OH (4 equiv) was double-coupled to the peptide using the manual SPPS with PyBOP (4 equiv) and DIPEA (4 equiv) in DMF at rt over 2h, and Fmoc was cleaved with 20% piperidine in DMF with 0.5 M oxyma pure over 5 min, and then 10 min, at rt. The PEGylated peptide sequence was reacted with FITC isomer I (2 equiv) in presence of DIPEA (4 equiv) in DMF over 1h at rt, and the labelling with FITC was repeated overnight. Global deprotection and cleavage from the resin was achieved through treatment with TFA/H₂O/DODT/TIPS (90 : 5 : 2.5 : 2.5 v/v) normally over 1-3h at rt, over 6h for the histone tail peptides and overnight in case of the TAT-derived peptides. The peptides were triturated in cold Et₂O and washed three more times with cold Et₂O, dried, dissolved in MeCN:H₂O (1:1 v/v; containing 0.1% TFA) and lyophilised. The crude peptides were purified by preparative HPLC (using MeCN/MeOH + 0.1% TFA and H₂O + 0.1% TFA as buffers) to afford the final products.

The N-terminally methylated peptide **183** was synthesised following the modified protocols by Kim *et al.*, 2014, in order to avoid the truncation of the N-terminal amino acid.^[444] The N-terminus of **183** was acetylated after the installation of the remaining peptide elements and TFA cleavage. The peptide was treated with Ac₂O (1.05 equiv) and DIPEA (2 equiv) in DMF over 5 min at rt, which as followed by titration in cold Et₂O, lyophilisation and purification with preparative HPLC.

7.3. Cyclic Peptide Synthesis

7.3.1. General Approach

Synthesis of the linear peptide intermediates for cyclisation, as well as further modifications were achieved through the standard Fmoc SPPS techniques on a polystyrene-based Rink Amide AM resin LL

(100-200 mesh) and 2-CTC resin (100-200 mesh) or polyethylene glycol-based Rink Amide ChemMatrix[®] resin (35-100 mesh). The 2-CTC resin was utilised for peptides intended to be cyclised in solution and the ChemMatrix[®] resin was used for peptides containing the TAT sequence, in all other cases the Rink Amide AM resin LL was applied. Resins used for macrocyclisation procedures on solid support had the following substitutions: Rink Amide AM resin LL – maximum substitution of 0.35 mmol/g, Rink Amide ChemMatrix[®] resin – substitution of ca. 0.5 mmol/g. Peptides were synthesised using the same general protocols as outlined for the linear peptides in section 7.2. Additionally, coupling to N₃-AEEEA (3 equiv) was achieved using COMU (2.7 equiv), oxyma pure (3 equiv) and DIPEA (6 equiv) in DMF overnight at rt.

7.3.2. Amide Bond Mediated Cyclisation in Solution

The peptides intended for cyclisation in solution (**11-38**, **46-51**, **56-60**, **84** and **85**) were cleaved from the 2-CTC resin through treatment with 20-40% HFIP in DCM, three times for 20 min at rt, yielding peptides with free C-terminal carboxylic acids. In case of compounds **32-38**, **46-51**, **84** and **85** the treatment with 40% HFIP also removed the protective Mtt group located on the amino acid side chain, exposing an amine group. The acid was evaporated in vacuo and the resulting solid was resuspended in DMF to the maximum concentration of 1 mM. The macrocyclisation through coupling of the free amine and carboxylic acid was performed by treatment with PyBOP (2 equiv) and DIPEA (4 equiv) over 1 day at rt. The solvent was completely evaporated under reduced pressure, and the remaining side chain protecting groups were removed by treatment with TFA/H₂O/DODT/TIPS (90 : 5 : 2.5 : 2.5 v/v) over 1-3h at rt, followed by purification with preparative HPLC.

7.3.2. Amide Bond Mediated Cyclisation on Solid Support

Cyclopeptides 52-55:

Linear peptides with exposed N-terminal amino acid were treated with either succinic or glutaric anhydride (10 equiv) and DIPEA (10 equiv) in DMF two times for 1h at rt, resulting in an N-terminal linker with a free carboxylic acid. The peptides had a C-terminal Orn or Lys residue protected by Alloc group, which was then deprotected through treatment of the peptides with Pd(PPh₃)₄ (25 mol %) and PhSiH₃ (30 equiv) in anhydrous DCM, two times for 1h, under a protective atmosphere of argon. Resin was then washed with DCM (4x 30s), DMF (4x 30s), 0.1 M solution of cupral in DMF (5x 5min) and finally with DMF (4x 30s). The macrocyclisation was achieved using PyBOP (2 equiv) and DIPEA (4 equiv) in DMF over 2 days at rt.

Cyclopeptides 61-65:

The allyl group on the C-terminal Asp residue was removed using Pd(PPh₃)₄ (25 mol %) and PhSiH₃ (30 equiv) in anhydrous DCM, two times for 1h at rt, under argon. Resin was washed with DCM (4x 30s), DMF (4x 30s), 0.1 M solution of cupral in DMF (5x 5min) and with DMF (4x 30s). Fmoc protecting the amino group of N-terminally located linker was removed with 20% piperidine in DMF at rt over 5 min and then 10 min. The on-resin cyclisation was then performed in presence of PyBOP (2 equiv) and DIPEA (4 equiv) in DMF over 2 days at rt.

Cyclopeptides 123-128:

Linear peptides synthesised towards macrocycles **123-128**, were attached to the resin via a side chain of the C-terminal residue, whereas the carboxyl group attached directly to the α -carbon was protected with an allyl group. The allyl group was removed using Pd(PPh₃)₄ (25 mol %) and PhSiH₃ (30 equiv) in anhydrous DCM, two times for 1h at rt, under argon. Resin was washed with DCM (4x 30s), DMF (4x 30s), 0.1 M solution of cupral in DMF (5x 5min) and with DMF (4x 30s). The macrocyclisation was achieved through combining the free N-terminal amine and the C-terminal carboxylic acid using PyBOP (2 equiv) and DIPEA (4 equiv) in DMF over 2 days at rt. The cyclisation was followed by Mtt removal from Lys side chain using 40% HFIP and 2.5 % TIPS in DCM, three times for 30 min at rt, allowing to install a fluorescent tag on an macrocyclic ring.

Cyclopeptides 133-140:

Amide-cyclised peptides **133-140** were synthesised through incorporation of allyl-protected Asp or Glu and Mtt-protected Dap or Dab into the linear peptidic chain. Allyl group was selectively removed through treatment of the resin-bound peptide with Pd(PPh₃)₄ (25 mol %) and PhSiH₃ (30 equiv) in anhydrous DCM, three times for 1h, under a protective atmosphere of argon. Resin was then washed with DCM (4x 30s), DMF (4x 30s), 0.1 M solution of cupral in DMF (5x 5min) and finally with DMF (4x 30s). Mtt was removed from Dab/Dap by treatment of the solid support attached peptide with a solution of 40% HFIP in DCM, containing 2.5% TIPS, four times for 20 min at rt. The Mtt removal was followed by washing of the resin with DCM (4x 30s) and DMF (4x 30s).

The cyclisation was achieved by coupling the side chains of Asp/Glu and Dab/Dap, where the resin-bound peptide was treated with PyBOP (2 equiv) and DIPEA (4 equiv) in DMF over 2 days at rt. Allyl protected Asp in case of peptides **133**, **135**, **137** and **139** proved to form a significant amount of aspartimide, and Dmb protection of the preceding residue was necessary: Fmoc deprotected, resin-bound DAD sequence was washed with a mixture of DMF/MeOH/AcOH (9 : 9 : 2) for 5 min, followed by a wash with DMF/MeOH (1 : 1, 4x 30s). The resin was treated with 2,4-dimethoxybenzaldehyde (10 equiv) in DMF/MeOH (1 : 1) for 45 min at rt and washed with DMF/MeOH (1 : 1, 4x 30s). NaBH₃CN (20 equiv) was added in DMF/MeOH/AcOH (9 : 9 : 2), and the resin was shaken for 30 min, at rt, then

washed with DMF/MeOH/AcOH (9 : 9 : 2), NMP, 5% DIPEA in NMP, NMP and DMF (4x), affording the Dmb protected DAD peptidic intermediate on solid support. Dmb-DAD gave a negative Kaiser test result and the product could be observed by LC-MS after a cleavage from the resin. Fmoc-Asp(OAll) (4 equiv) was coupled to this sequence using COMU (3.6 equiv), oxyma pure (4 equiv), DIPEA (8 equiv) in DMF, three times for 24h, at rt. Uncoupled Dmb-DAD sequence was capped by treatment with Ac₂O (10 equiv) and DIPEA (12 equiv) in DMF over 30 min at rt.

7.3.3. RCM Mediated Cyclisation on Solid Support

Dry resin was swelled in argon-flushed DCE for 20 min. Hoveyda-Grubbs 2nd generation catalyst (10 mol%) was added and the mixture flushed with argon, and heated under MWI for 10 min at 120 °C. The mixture was flushed with argon and heated again under MWI for 10 min at 120 °C with a new portion of the catalyst (10 mol%), washed with DCM (4x) and DMF (4x). Where applicable, the double bond, created in RCM, was reduced: resin was washed with NMP (4x), then TPSH (30 equiv) and DIPEA (46 equiv) were added to the resin suspended in NMP. The mixture was shaken for 24h at 50 °C, washed with NMP (2x), and the process was repeated two more times with fresh portions of the reagents. The progress of the double bond reduction was monitored by LC-MS. The resin was washed with NMP (2x), DMF (4x), DCM (4x), Et₂O (2x), H₂O (4x) and DMF (4x). When possible, the *cis* and *trans* isomers were separated.

7.4. Protein Expression and Purification

7.4.1. Protein Expression

For expression of the proteins, all constructs were subcloned into pOPIN vectors. Full-length human pICln, N-terminally tagged with 6His-TRX followed by the 3C cleavage site, and TIM barrel domain (1-292 aa) of human PRMT5, N-terminally tagged with 6His-SUMO, were expressed in BL21 CodonPlus (DE3) RIPL strain of Escherichia coli. Bacteria with the respective plasmids were cultured in Terrific Broth medium. Expression of TIM barrel was auto-induced, with incubation of the starter-culture (starting OD of ~0.05) at 35 °C for 2.5h, followed by an overnight incubation (20-24h) at 25 °C. Expression of pICln was induced at OD ~0.6-0.8 using 0.3 mM IPTG and E. coli was incubated at 20 °C for 20h. Bacteria were harvested by centrifugation and lysed.

Expression of full-length human RioK1 tagged with N-terminal 6His-3C and co-expression of full-length human PRMT5 (N-terminal 6His-MBP with 3C) with MEP50 (N-terminal 6His-TRX with 3C), and of TIM barrel (1-292 aa) with MEP50 (both N-terminally tagged with His-TRX with 3C cleavage site) was done

using the Flash Bac Expression System (Oxford Expression Technologies, UK) with HighFive insect cells in Gibco Sf-900 III SFM medium. The cells were harvested by centrifugation and lysed.

7.4.2. Protein Purification

Protein purification was performed on HisTrap FF crude 5 ml Ni-based column, followed by the on-column treatment with 6His-tagged HRV-3C PreScission protease, or in case of the SUMO-tagged TIM barrel, with 6His-tagged SUMO protease. The cleaved proteins were then further purified using gel filtration chromatography with the HiLoad 26/60 Superdex 75 or 200 prep grade column. PRMT5-MEP50, TIM-MEP50 and TIM barrel were stored in a buffer containing 50 mM HEPES, 250 mM NaCl, 1 mM TCEP with the pH adjusted to 8.0, whereas, pICln was stored in a buffer of 20 mM phosphate, 100 mM NaCl, 5 mM TCEP and pH of 7.5, whereas RioK1 was stored in a buffer of 50 mM Tris, 200 mM NaCl, 10% glycerol, 1 mM TCEP at pH 8.0.

7.5. Mass Photometry

Autofocus of the mass photometer was set using 18 μ l of the stock buffer of 50 mM HEPES, 250 mM NaCl, 1 mM TCEP, 8.0 pH at rt. To that was added 2 μ l of either 500 nM PRMT5-MEP50 complex or 500 nM TIM-MEP50 complex solution in the stock buffer at rt, giving the final working concentration of 50 nM, followed by the acquisition of the movies and the automated data analysis.

7.6. Protein Labelling with Alexa 488

To a solution of the full-length pICln protein (ca. 1 mg/ml) in 0.2 M bicarbonate buffer with 1 mM TCEP at pH 8.3, was added 10 mM Alexa 488 NHE-ester in DMSO (8 equiv). The solution was kept on ice, in darkness, overnight. The protein was washed eight to ten times with a buffer intended for the fluorescence polarisation assay (50 mM HEPES, 250 mM NaCl, 1 mM TCEP, 8.0 pH), using an Amicon[®] Ultra-0.5 mL spin filter (Merck, Germany). Protein concentration was determined using the NanoDrop spectrophotometer.

7.7. Fluorescence Polarisation

7.7.1. Direct Binding Assay

The assay was performed in a buffer of 50 mM HEPES, 250 mM NaCl, 1 mM TCEP, 0.01% (v/v) Tween 20, pH 8.0, in black, 384 well-plates, with a total volume of 10 μ l per well. The analysed FITC labelled

peptides were tested at a final concentration of 1 nM, and the unlabelled PRMT5-MEP50 protein complex was titrated as two-fold dilution series. The plates were incubated at room temperature for up to 2h, and analysed on a plate reader using 485 nm excitation and 535 nm emission wavelength. The assay was performed in triplicates.

7.7.2. Competitive Binding Assay

The competitive binding assay was performed in a buffer of 50 mM HEPES, 250 mM NaCl, 1 mM TCEP, 0.01% (v/v) Tween 20, 8.0 pH, in black, 384 well-plates, with a total volume of 10 μ l per well. The FITC labelled peptides or the Alexa488-labelled pICln protein, at a final concentration of 1 nM, were mixed with PRMT5-MEP50 (at the final concentration of 194 nM when used with peptide **184**, 600 nM with compound **93** and **149**, and 27 nM in the case of Alexa488-labelled pICln), and the appropriate non-labelled peptide or protein was titrated as two-fold dilution series. The plates were incubated at room temperature for up to 1h 45 min, and analysed on a plate reader using 485 nm excitation and 535 nm emission wavelength. The experiment was performed in triplicates.

7.8. Flow Induced Dispersion Analysis

FIDA experiments were performed using Fida 1 platform with LED 488 nm detector and PEG-coated capillary with inner diameter of 75 μ m. The analyses were performed in a buffer of 50 mM HEPES, 250 mM NaCl, 1 mM TCEP, 8.0 pH and 0.01% (v/v) Tween 20. The column was equilibrated with buffer, followed by the injection of the indicator sample (Alexa488 labelled pICln or RioK1), and subsequent injection of the indicator-analyte mixtures (pICln-Alexa488 or RioK1-Alexa488 mixed with unlabelled PRMT5-MEP50 or TIM-MEP50). The indicator was used at a concentration of 10 nM. The analysis was performed at 25 °C. The obtained data were evaluated using FIDA software Version 2.01 with the fraction setting fixed to 75% or 100% and fixed selection of Rh for one species option. The samples were analysed as technical triplicate.

7.9. Stability Assay

To whole cell lysate prepared from U2OS cells with the freeze-thaw method was added the appropriate peptide, resulting in a mixture of 0.5 mg/ml protein and 600 μ M peptide in PBS, incubated at 37 °C. Samples were taken at 0 min, 15 min, 30 min, 1h, 2h, 4h, 1-day, 3-day and 7-day timepoint, and mixed with equal volume of ice-cold ethylparaben solution in MeOH (0.1 mg/ml), used as the internal standard. The samples were kept on ice for 15 min and centrifuged at 16,873 \times g at 4 °C for 5 min. The obtained supernatant was then analysed by HPLC (Method B), the peptide and the internal

standard peaks at 210 nm were integrated and the ratio of the peptide peak surface area to the internal standard was calculated. The experiment was conducted in duplicates.

7.10. MTase-Glo™ Activity Assay

The methyltransferase activity of the expressed and purified PRMT5-MEP50 complex was tested in the presence of the active site inhibitor EPZ015666 and the appropriate peptidomimetics, using the MTase-Glo™ Methyltransferase Assay from Promega (Promega Corporation, USA).^[392] The assay was performed in 1X reaction buffer of 20 mM Tris, 50 mM NaCl, 1 mM EDTA, 3 mM MgCl₂, 0.1 mg/ml BSA, 1 mM DTT at pH = 8.0, in white, 384 well-plates. A dilution series of tested compounds was prepared in a mixture of 2 μM S-adenosylmethoinine, 1 μM H4 histone tail peptide and 2X MTase-Glo™ Reagent in the 1X reaction buffer (2.5 μl per well), and equal volume of 200 nM PRMT5-MEP50 protein solution was added to give a total volume of 5 μl per well. The plate was incubated at 20-21 °C for 1h. MTase-Glo™ Detection Solution (5 μl) was added, the plate was incubated at 20-21 °C for 1h, followed by a luminescence measurement at a plate reader. The experiment was performed as triplicates.

7.11. Nanoscale Differential Scanning Fluorimetry

To 5 μM of PRMT5-MEP50 complex was added 10, 20 or 50 equiv of peptides in 250 mM NaCl, 50 mM HEPES, 1 mM TCEP, pH = 8.0. The samples were incubated at room temperature for 30 minutes and then high sensitivity nanoDSF capillaries were filled with 10 μl of the sample. All the capillaries were placed in the sample holder of the Prometheus NT.48 instrument and a gradient of 1 °C·min⁻¹ from 20 to 90 °C was applied. Melting temperatures were determined after characterisation of the inflection point of the intrinsic fluorescence curve for the ratio 350/330 nm using the first derivative of the corresponding curve. All experiments were performed as quadruplicates.

7.12. Thermal Shift Assay

Thermal Shift Assays of peptides on the TIM-MEP50 complex were carried out into a 96 well PCR plate (LightCycler® 480 Multiwell Plates 96, white, 04729692001). The purified protein complex was appropriately diluted in a buffer containing 50 mM HEPES (pH = 8.0), 250 mM NaCl, 1 mM TCEP. All assay experiments used 5.36 μg protein per well and 140 nl 5000X Sypro Orange (Invitrogen) up to a total volume of 25 μl, with a resultant protein concentration of 3 μM and 5X SYPRO. Peptides were supplied at 10 mM concentration in DMSO. The PCR plates were sealed with optical seal (LightCycler®

480 Sealing Foil, 04729757001), shaken, and centrifuged after protein and compounds were added. Thermal scanning (25 to 95°C at 1°C·min⁻¹) was performed using a real-time PCR setup (LightCycler® 480 – Roche) and fluorescence intensity was measured after every temperature increment step. Analysis of the raw data was performed using internally developed software. Statistical validation of T_m shift relevance was performed using Student test (n=6).

7.13. Protein Crystallisation, Data Collection and Analysis

The protein was crystallised in a sitting drop setup, using 100 nl reservoir solution (0.1M NaAc, 0.2 M CaAc, 26-30% (v/v) PEG 400, at pH 4.18-4.73) and 100 nl of protein-peptide solution (5 mg/ml protein, 450 µM peptide 23) in a buffer of 50 mM HEPES, 150 mM NaCl, 1 mM TCEP, 10% (v/v) glycerol and pH = 8.0. The crystals were grown for two weeks at 4 °C, and were flash frozen in liquid nitrogen. Datasets of the crystals were taken at the Suisse Light Source X10SA beamline. The data were processed and scaled using Xia2/DIALS.^[447,448] The structure could be solved using a truncated model of PDB 4X60 (chain A) with Phaser.^[449] Cycles of manual and automated refinement was performed with phenix.refine and Coot.^[450,451]

In order to produce anomalous data for structure verification, crystals of TIM barrel-peptide complex were soaked in 2 mM solution of K₂PtCl₄ for 9 min. A dataset was taken at the X10SA beamline at a wavelength of 0.85Å. Similarly, another crystal of the complex was measured at X10SA at a wavelength of 2Å for sulphur anomalous data. Both datasets were solved using the previously refined TIM barrel-peptide structure as phaser search model and refined using phenix-refined. Anomalous maps were calculated from the mtz files in coot using ANOM and ANOMPH amplitudes and phases.

The Matthews coefficient of 4.6 deviates substantially from the ideal value of 2.3 indicating a solvent content of 73.25%. Attempts to place a second TIM barrel structure or parts thereof into the asymmetric unit have failed in every program tested. The MR rotation function score indicates the presence of a single TIM barrel molecule only in the asymmetric unit.

7.14. Cell Culture

MCF7, HEK293-T and U2OS cells were cultured in DMEM medium, supplemented with 10% foetal bovine serum, 1% sodium pyruvate, 1% nonessential amino acids, at 37 °C in a 5% CO₂ humidified atmosphere. MCF7 medium was further supplemented with 10 µg/ml insulin. Generation of inducible Flp-In T-Rex 293 cells system expressing GFP, GFP-PRMT5, GFP-RioK1 and GFP-pICln was carried out according to manufacturer's instructions (Invitrogen, Thermo Fisher Scientific) and has been described

previously.^[452,453] All Flp-In T-Rex 293 cells were cultured in DMEM (4.5 g/l D-glucose; Gibco, Thermo Fisher Scientific) supplemented with 10% (v/v) FCS (Biochrom, Merck), 100 U/ml Penicillin and 100 µg/ml Streptomycin (Gibco, Thermo Fisher Scientific) in a 5% CO₂ humidified atmosphere at 37 °C. For induction of GFP, GFP-PRMT5, GFP-RioK1 and GFP-pICln expression, Flp-In T-Rex 293 cell lines were stimulated with 0.1 µg/ml Doxycycline (Clontech) for 18 h.

7.15. SDMA Pattern Analysis

7.15.1. Cell Treatment with PAPIIs

Experiment was performed with HEK293-T and MCF7 cells. HEK293-T cells were treated with 1 µM GSK3326595 PRMT5 inhibitor, 10 µM or 50 µM CPP-labelled **160** and **161** macrocycles, 10 µM or 50 µM TAT sequence **164** and DMSO. HEK293-T cells were incubated at 37 °C in a 5% CO₂ humidified atmosphere for 2 days. MCF7 cells were treated with 1 µM GSK3326595 PRMT5 inhibitor, 50 µM CPP-labelled **160** and **161** macrocycles, 50 µM TAT sequence **164** and DMSO. MCF7 cells were incubated at 37 °C in a 5% CO₂ humidified atmosphere for 3 days.

7.15.2. Cell Lysis and Immunoblotting

The cells were washed with PBS and then harvested through lysis with SDS buffer (40 µM Tris, 8% v/v glycerol, 2% SDS, pH 6.8). The collected mixtures were sonicated and centrifuged at 16,873 × g at 4 °C for 5 min. The supernatant was collected, and DTE (final concentration of 80 µM) and bromophenol blue (final concentration of 0.02%) were added. The samples were heated shaking for 10 min at 95 °C and 350 rpm. The resulting solutions were analysed by western blotting using 10% acrylamide gels and PVDF membranes. The SDMA proteins of interest were detected using anti-SDMA antibody, selecting either anti-vinculin or anti-GAPDH antibody to use vinculin or GAPDH as a loading control, and were visualised using corresponding secondary IRDye antibodies.

7.15. Pulldown

DBCO agarose beads (300 µl; Jena Bioscience, Germany) were washed three times with PBS (600 µl, pH 7.4). To the beads suspended in PBS (600 µl) was added a 10 mM solution of azide labelled peptide in DMSO, or pure DMSO. The obtained suspensions were incubated overnight at 4 °C. The beads were washed with PBS three times for 10 min. MCF7 cells were harvested and lysed using triton X100 lysis buffer (NaCl 150 mM, HEPES 25 mM, TCEP 1 mM, 1% triton x100, 1% NP40 alternative, protease and phosphatase inhibitor, pH 7.4). Empty beads, beads bound to an active peptide (**189**), or to a

scrambled peptide (**190**), were added to lysate (6 mg/ml) and rotated at 4 °C overnight. Beads attached to the active peptides were also incubated in lysate containing non-immobilised active macrocycle (**187** at 100 µM). Beads were collected by centrifugation (800 x g for a 1 min) and washed three times for 10 min with PBS (500 µl). The beads were suspended in SDS-PAGE protein loading buffer (20 µl; 40 µM Tris, 8% v/v glycerol, 2% SDS, 80 µM dithioerythritol (DTE), 0.02% bromophenol blue, pH 6.8), and heated shaking for 10 min at 95 °C and 350 rpm. The obtained solutions were analysed by western blotting, using 10% acrylamide gels, PVDF membranes and the relevant proteins were fluorescently detected using antibodies against PRMT5 and MEP50 and the corresponding secondary IRDye antibodies.

7.16. GFP Immunopurification and Immunoblotting

7.16.1. Generation of S100 Extract

Harvested Flp-In T-REx 293 cells were incubated with Roeder A buffer³² in three times sample weight for 10 min at room temperature, dounced 10 times and adjusted to 150 mM NaCl. After centrifugation at 17,000 g for 30 min the supernatants (S100 extracts) were used for immunopurification and treatment with inhibitors.

7.16.2. Inhibitor Treatment and Immunopurification

For the treatment the S100 extracts of GFP, GFP-PRMT5, GFP-RioK1 and GFP-pICln were incubated with DMSO, active **187** and scrambled **188**, used at a concentration of 50 µM. The incubation with the S100 extract was carried out for 1 h at room temperature. For GFP immunopurification S100 extracts were incubated with GFP-Trap_A beads (ChromoTek) at 4 °C for 2 h with rotation. Purified proteins were washed 3 times with washing buffer (150 mM NaCl, 50 mM Tris/HCl pH 7.5, 1 mM EDTA, 1 mM EGTA and 0.01% Igepal with protease inhibitors), eluted in sample buffer [375 mM Tris pH 7.5; 25.8% (w/v) glycerol; 12.3% (w/v) SDS; 0.06% (w/v) bromophenol blue; 6% (v/v) β-mercaptoethanol; pH 6.8] and analysed by immunoblotting.

7.16.3. Immunoblotting

Immunopurification samples were separated by Tris/Tricine or Tris/Glycine SDS gel electrophoresis and transferred to PVDF membranes (Immobilon-FL, Merck Millipore).^[454] The immunoblot analysis was performed using the indicated antibodies against RioK1, PRMT5, MEP50, pICln and GFP and signals were detected with an Odyssey LI-COR Imaging System (LI-COR Biosciences, USA).

7.17. Computational Modelling

Computational modelling was performed using the Maestro environment, version 12.3.013, and the Schrödinger suite, release 2020-1 (Schrödinger Inc., USA). Protein preparation was done with the Protein Preparation Wizard (Schrödinger), with the amino acid protonation states refined using PROPKA at pH set to 8.0.^[455] The macrocyclic compounds were built upon the co-crystal structure obtained for RioK1-derived peptide sequence and TIM barrel of PRMT5. Using the Conformational Search tool (Schrödinger) with the OPLS_2005 force field and the water solvent, different conformations were generated for the cyclic core of the molecules with all the remaining atoms fixed.^[456] Ligand refinement was performed with Glide v8.6 (Schrödinger), using first the standard precision, then followed by the extra precision ligand refinement.^[457–459] The obtained models were evaluated visually.

Bibliography

- [1] B. Alberts, *Cell* **1998**, *92*, 291–294.
- [2] T. Pawson, P. Nash, *Science* **2003**, *300*, 445–452.
- [3] M. Vidal, M. E. Cusick, A. L. Barabási, *Cell* **2011**, *144*, 986–998.
- [4] J. Hardin, G. Bertoni, L. J. Kleinsmith, *Becker's World of the Cell*, Pearson, Boston, **2011**.
- [5] K. Venkatesan, J. F. Rual, A. Vazquez, U. Stelzl, I. Lemmens, T. Hirozane-Kishikawa, T. Hao, M. Zenkner, X. Xin, K. Il Goh, M. A. Yildirim, N. Simonis, K. Heinzmann, F. Gebreab, J. M. Sahalie, S. Cevik, C. Simon, A. S. de Smet, E. Dann, A. Smolyar, A. Vinayagam, H. Yu, D. Szeto, H. Borick, A. Dricot, N. Klitgord, R. R. Murray, C. Lin, M. Lalowski, J. Timm, K. Rau, C. Boone, P. Braun, M. E. Cusick, F. P. Roth, D. E. Hill, J. Tavernier, E. E. Wanker, A. L. Barabási, M. Vidal, *Nat. Methods* **2009**, *6*, 83–90.
- [6] B. Lehne, T. Schlitt, *Hum. Genomics* **2009**, *3*, 291.
- [7] A. Hoelz, E. W. Debler, G. Blobel, *Annu. Rev. Biochem.* **2011**, *80*, 613–643.
- [8] D. H. Lin, A. Hoelz, *Annu. Rev. Biochem.* **2019**, *88*, 725–783.
- [9] G. Huang, X. Zhan, C. Zeng, X. Zhu, K. Liang, Y. Zhao, P. Wang, Q. Wang, Q. Zhou, Q. Tao, M. Liu, J. Lei, C. Yan, Y. Shi, *Cell Res.* **2022**, *32*, 349–358.
- [10] S. Petrovic, D. Samanta, T. Perriches, C. J. Bley, K. Thierbach, B. Brown, S. Nie, G. W. Mobbs, T. A. Stevens, X. Liu, A. Hoelz, *bioRxiv* **2021**, *9798*, 2021.10.26.465796.
- [11] D. J. Battle, M. Kasim, J. Yong, F. Lotti, C. K. Lau, J. Mouaikel, Z. Zhang, K. Han, L. Wan, G. Dreyfuss, *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 313–320.
- [12] X. Zhang, C. Yan, J. Hang, L. I. Finci, J. Lei, Y. Shi, *Cell* **2017**, *169*, 918–929.e14.
- [13] M. J. Davey, D. Jeruzalmi, J. Kuriyan, M. O'Donnell, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 826–835.
- [14] J. T. Canty, R. Tan, E. Kusakci, J. Fernandes, A. Yildiz, *Annu. Rev. Biophys.* **2021**, *50*, 549–574.
- [15] W. Junge, N. Nelson, *Annu. Rev. Biochem.* **2015**, *84*, 631–657.
- [16] M. Yoshida, E. Muneyuki, T. Hisabori, *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 669–677.
- [17] M. P. Mayer, *Mol. Cell* **2010**, *39*, 321–331.
- [18] R. J. Ellis, *Philos. Trans. R. Soc. B Biol. Sci.* **2013**, *368*, 3–7.
- [19] T. D. Pollard, *Nature* **2003**, *422*, 741–745.
- [20] H. Y. Kueh, T. J. Mitchison, *Science* **2009**, *325*, 960–963.
- [21] C. R. Clapier, B. R. Cairns, *Annu. Rev. Biochem.* **2009**, *78*, 273–304.
- [22] C. R. Clapier, J. Iwasa, B. R. Cairns, C. L. Peterson, *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 407–422.

- [23] G. Zinzalla, D. E. Thurston, *Future Med. Chem.* **2009**, *1*, 65–93.
- [24] S. Jones, J. M. Thornton, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13–20.
- [25] I. M. A. Nooren, J. M. Thornton, *EMBO J.* **2003**, *22*, 3486–3492.
- [26] S. E. Acuner Ozbabacan, H. B. Engin, A. Gursoy, O. Keskin, *Protein Eng. Des. Sel.* **2011**, *24*, 635–648.
- [27] Q. Xu, R. L. Dunbrack, *Curr. Opin. Struct. Biol.* **2019**, *55*, 34–49.
- [28] P. Block, J. Paern, E. Hüllermeier, P. Sanschagrin, C. A. Sotriffer, G. Klebe, *Proteins Struct. Funct. Bioinforma.* **2006**, *65*, 607–622.
- [29] J. R. Perkins, I. Diboun, B. H. Dessailly, J. G. Lees, C. Orengo, *Structure* **2010**, *18*, 1233–1243.
- [30] M. C. Smith, J. E. Gestwicki, *Expert Rev. Mol. Med.* **2012**, *14*, 1–20.
- [31] M. G. Su, J. T. Y. Weng, J. B. K. Hsu, K. Y. Huang, Y. H. Chi, T. Y. Lee, *BMC Syst. Biol.* **2017**, *11*, DOI 10.1186/s12918-017-0506-1.
- [32] S. Wang, A. O. Osgood, A. Chatterjee, *Curr. Opin. Struct. Biol.* **2022**, *74*, 102352.
- [33] B. P. Cossins, A. D. G. Lawson, *Molecules* **2015**, *20*, 16435–16445.
- [34] C. Sheng, G. Dong, Z. Miao, W. Zhang, W. Wang, *Chem. Soc. Rev.* **2015**, *44*, 8238–8259.
- [35] J. P. Changeux, *Drug Discov. Today Technol.* **2013**, *10*, e223–e228.
- [36] L. Lo Conte, C. Chothia, J. Janin, *J. Mol. Biol.* **1999**, *285*, 2177–2198.
- [37] S. Ansari, V. Helms, *Proteins Struct. Funct. Genet.* **2005**, *61*, 344–355.
- [38] S. De, O. Krishnadev, N. Srinivasan, N. Rekha, *BMC Struct. Biol.* **2005**, *5*, 1–16.
- [39] I. M. A. Nooren, J. M. Thornton, *J. Mol. Biol.* **2003**, *325*, 991–1018.
- [40] H. Zhu, F. S. Domingues, I. Sommer, T. Lengauer, *BMC Bioinformatics* **2006**, *7*, 27.
- [41] J. Mintseris, Z. Weng, *Proteins Struct. Funct. Genet.* **2003**, *53*, 629–639.
- [42] N. Brooijmans, K. A. Sharp, I. D. Kuntz, *Proteins Struct. Funct. Genet.* **2002**, *48*, 645–653.
- [43] B. Apostolovic, M. Danial, H. A. Klok, *Chem. Soc. Rev.* **2010**, *39*, 3541–3575.
- [44] D. E. Scott, A. R. Bayly, C. Abell, J. Skidmore, *Nat. Rev. Drug Discov.* **2016**, *15*, 533–550.
- [45] E. Petsalaki, R. B. Russell, *Curr. Opin. Biotechnol.* **2008**, *19*, 344–350.
- [46] A. Stein, P. Aloy, *PLoS One* **2008**, *3*, 1–10.
- [47] P. Aloy, R. B. Russell, *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 188–197.
- [48] A. Stein, R. A. Pache, P. Bernadó, M. Pons, P. Aloy, *FEBS J.* **2009**, *276*, 5390–5405.
- [49] N. London, D. Movshovitz-Attias, O. Schueler-Furman, *Structure* **2010**, *18*, 188–199.
- [50] M. Rickert, X. Wang, M. J. Boulanger, N. Goriatcheva, K. C. Garcia, *Science* **2005**, *308*, 1477–1480.

- [51] S. K. Nair, S. K. Burley, *Cell* **2003**, *112*, 193–205.
- [52] B. Ku, C. Liang, J. U. Jung, B. H. Oh, *Cell Res.* **2011**, *21*, 627–641.
- [53] E. D. Levy, J. B. Pereira-Leal, *Curr. Opin. Struct. Biol.* **2008**, *18*, 349–357.
- [54] W. S. J. Valdar, J. M. Thornton, *Proteins Struct. Funct. Genet.* **2001**, *42*, 108–124.
- [55] D. R. Caffrey, *Protein Sci.* **2004**, *13*, 190–202.
- [56] J. Mintseris, Z. Weng, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10930–10935.
- [57] T. Clackson, J. A. Wells, *Science* **1995**, *267*, 383–386.
- [58] O. Keskin, B. Ma, R. Nussinov, *J. Mol. Biol.* **2005**, *345*, 1281–1294.
- [59] I. S. Moreira, P. A. Fernandes, M. J. Ramos, *Proteins Struct. Funct. Bioinforma.* **2007**, *68*, 803–812.
- [60] X. Li, O. Keskin, B. Ma, R. Nussinov, J. Liang, *J. Mol. Biol.* **2004**, *344*, 781–795.
- [61] A. A. Bogan, K. S. Thorn, *J. Mol. Biol.* **1998**, *280*, 1–9.
- [62] R. M. Ramos, L. F. Fernandes, I. S. Moreira, *Comput. Biol. Chem.* **2013**, *44*, 31–39.
- [63] E. Cukuroglu, H. B. Engin, A. Gursoy, O. Keskin, *Prog. Biophys. Mol. Biol.* **2014**, *116*, 165–173.
- [64] D. Ovek, Z. Abali, M. E. Zeylan, O. Keskin, A. Gursoy, N. Tuncbag, *Curr. Opin. Struct. Biol.* **2022**, *72*, 209–218.
- [65] D. P. Ryan, J. M. Matthews, *Curr. Opin. Struct. Biol.* **2005**, *15*, 441–446.
- [66] L. C. Cesa, A. K. Mapp, J. E. Gestwicki, *Front. Bioeng. Biotechnol.* **2015**, *3*, 1–18.
- [67] J. A. Wells, C. L. McClendon, *Nature* **2007**, *450*, 1001–1009.
- [68] H. Lu, Q. Zhou, J. He, Z. Jiang, C. Peng, R. Tong, J. Shi, *Signal Transduct. Target. Ther.* **2020**, *5*, DOI 10.1038/s41392-020-00315-3.
- [69] M. Rosell, J. Fernández-Recio, *Expert Opin. Drug Discov.* **2018**, *13*, 327–338.
- [70] N. London, B. Raveh, O. Schueler-Furman, *Curr. Opin. Chem. Biol.* **2013**, *17*, 952–959.
- [71] M. R. Arkin, A. Whitty, *Curr. Opin. Chem. Biol.* **2009**, *13*, 284–290.
- [72] O. Keskin, R. Nussinov, *Structure* **2007**, *15*, 341–354.
- [73] P. M. Kim, L. J. Lu, Y. Xia, M. B. Gerstein, *Science* **2006**, *314*, 1938–1941.
- [74] A. C. Cheng, R. G. Coleman, K. T. Smyth, Q. Cao, P. Souldard, D. R. Caffrey, A. C. Salzberg, E. S. Huang, *Nat. Biotechnol.* **2007**, *25*, 71–75.
- [75] S. Eyrisch, V. Helms, *J. Med. Chem.* **2007**, *50*, 3457–3464.
- [76] G. Fischer, M. Rossmann, M. Hyvönen, *Curr. Opin. Biotechnol.* **2015**, *35*, 78–85.
- [77] S. A. Andrei, E. Sijbesma, M. Hann, J. Davis, G. O’Mahony, M. W. D. Perry, A. Karawajczyk, J. Eickhoff, L. Brunsveld, R. G. Doveston, L. G. Milroy, C. Ottmann, *Expert Opin. Drug Discov.* **2017**, *12*, 925–940.

- [78] N. London, B. Raveh, D. Movshovitz-Attias, O. Schueler-Furman, *Proteins Struct. Funct. Bioinforma.* **2010**, *78*, 3140–3149.
- [79] A. P. Higuieruelo, A. Schreyer, G. R. J. Bickerton, W. R. Pitt, C. R. Groom, T. L. Blundell, *Chem. Biol. Drug Des.* **2009**, *74*, 457–467.
- [80] A. L. Jochim, P. S. Arora, *ACS Chem. Biol.* **2010**, *5*, 919–923.
- [81] A. P. Turnbull, S. M. Boyd, B. Walse, **2014**, 13–26.
- [82] D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon, J. A. Wells, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 9367–9372.
- [83] A. C. Braisted, J. D. Oslob, W. L. Delano, J. Hyde, R. S. McDowell, N. Waal, C. Yu, M. R. Arkin, B. C. Raimundo, *J. Am. Chem. Soc.* **2003**, *125*, 3714–3715.
- [84] J. Hyde, A. C. Braisted, M. Randal, M. R. Arkin, *Biochemistry* **2003**, *42*, 6475–6483.
- [85] E. Sijbesma, K. K. Hallenbeck, S. Leysen, P. J. De Vink, L. Skóra, W. Jahnke, L. Brunsveld, M. R. Arkin, C. Ottmann, *J. Am. Chem. Soc.* **2019**, *141*, 3524–3531.
- [86] M. Wolter, D. Valenti, P. J. Cossar, L. M. Levy, S. Hristeva, T. Genski, T. Hoffmann, L. Brunsveld, D. Tzalis, C. Ottmann, *Angew. Chemie - Int. Ed.* **2020**, *59*, 21520–21524.
- [87] O. Ichihara, J. Barker, R. J. Law, M. Whittaker, *Mol. Inform.* **2011**, *30*, 298–306.
- [88] S. E. Dalton, S. Campos, *ChemBioChem* **2020**, *21*, 1080–1100.
- [89] B. Li, D. Rong, Y. Wang, *Curr. Top. Med. Chem.* **2019**, *19*, 1872–1876.
- [90] S. S. Cheng, G. J. Yang, W. Wang, C. H. Leung, D. L. Ma, *J. Hematol. Oncol.* **2020**, *13*, 1–14.
- [91] J. D.S., W. E., C. B.F., *Future Med. Chem.* **2010**, *2*, 949–964.
- [92] H. R. Lawrence, Z. Li, M. L. Richard Yip, S. S. Sung, N. J. Lawrence, M. L. McLaughlin, G. J. McManus, M. J. Zaworotko, S. M. Sebti, J. Chen, W. C. Guida, *Bioorganic Med. Chem. Lett.* **2009**, *19*, 3756–3759.
- [93] J. Scheper, M. Guerra-Rebollo, G. Sanclimens, A. Moure, I. Masip, D. González-Ruiz, N. Rubio, B. Crosas, Ó. Meca-Cortés, N. Loukili, V. Plans, A. Morreale, J. Blanco, A. R. Ortiz, À. Messeguer, T. M. Thomson, *PLoS One* **2010**, *5*, DOI 10.1371/journal.pone.0011403.
- [94] W. Tian, X. Han, M. Yan, Y. Xu, S. Duggineni, N. Lin, G. Luo, Y. M. Li, X. Han, Z. Huang, J. An, *Biochemistry* **2012**, *51*, 724–731.
- [95] M. Pelay-Gimeno, A. Glas, O. Koch, T. N. Grossmann, *Angew. Chemie - Int. Ed.* **2015**, *54*, 8896–8927.
- [96] A. Tapeinou, M. T. Matsoukas, C. Simal, T. Tselios, *Biopolymers* **2015**, *104*, 453–461.
- [97] A. A. Vinogradov, Y. Yin, H. Suga, *J. Am. Chem. Soc.* **2019**, *141*, 4167–4181.
- [98] A. S. Ripka, D. H. Rich, *Curr. Opin. Chem. Biol.* **1998**, *2*, 441–452.
- [99] S. Baek, P. S. Kutchukian, G. L. Verdine, R. Huber, T. A. Holak, K. W. Lee, G. M. Popowicz, *J. Am. Chem. Soc.* **2012**, *134*, 103–106.

- [100] B. J. Smith, E. F. Lee, J. W. Checco, M. Evangelista, S. H. Gellman, W. D. Fairlie, *ChemBioChem* **2013**, *14*, 1564–1572.
- [101] X. Cao, J. L. Yap, M. K. Newell-Rogers, C. Peddaboina, W. Jiang, H. T. Papaconstantinou, D. Jupiter, A. Rai, K. Y. Jung, R. P. Tubin, W. Yu, K. Vanommeslaeghe, P. T. Wilder, A. D. MacKerell, S. Fletcher, R. W. Smythe, *Mol. Cancer* **2013**, *12*, DOI 10.1186/1476-4598-12-42.
- [102] B. Anil, C. Riedinger, J. A. Endicott, M. E. M. Noble, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1358–1366.
- [103] O. Koch, G. Klebe, *Proteins Struct. Funct. Bioinforma.* **2009**, *74*, 353–367.
- [104] C. M. Venkatachalam, *Biopolymers* **1968**, *6*, 1425–1436.
- [105] K. C. Chou, *Anal. Biochem.* **2000**, *286*, 1–16.
- [106] C. Toniolo, E. Benedetti, *Crit. Rev. Biochem. Mol. Biol.* **1980**, *9*, 1–44.
- [107] K. Guruprasad, M. S. Prasad, G. R. Kumar, *J. Pept. Res.* **2000**, *56*, 250–263.
- [108] W. A. Loughlin, J. D. A. Tyndall, M. P. Glenn, D. P. Fairlie, **2004**.
- [109] O. Koch, J. Cole, P. Block, G. Klebe, *J. Chem. Inf. Model.* **2009**, *49*, 2388–2402.
- [110] L. Pauling, R. B. Corey, *Proc. Natl. Acad. Sci.* **1951**, *37*, 251–256.
- [111] B. L. Sibanda, J. M. Thornton, *Nature* **1985**, *316*, 170–174.
- [112] L. Bragg, J. C. Kendrew, M. F. Perutz, *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.* **1950**, *203*, 321–357.
- [113] M. N. Fodje, S. Al-Karadaghi, *Protein Eng.* **2002**, *15*, 353–358.
- [114] L. Pauling, R. B. Corey, *Proc. Natl. Acad. Sci.* **1951**, *37*, 235–240.
- [115] B. W. Low, R. B. Baybutt, *J. Am. Chem. Soc.* **1952**, *74*, 5806–5807.
- [116] J. Donohue, *Proc. Natl. Acad. Sci.* **1953**, *39*, 470–478.
- [117] E. N. Baker, R. E. Hubbard, *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97–179.
- [118] D. J. Barlow, J. M. Thornton, *J. Mol. Biol.* **1988**, *201*, 601–619.
- [119] M. E. Karpen, P. L. De Haseth, K. E. Neet, *Protein Sci.* **1992**, *1*, 1333–1342.
- [120] L. Pal, G. Basu, *Protein Eng.* **1999**, *12*, 811–814.
- [121] R. B. Cooley, D. J. Arp, P. A. Karplus, *J. Mol. Biol.* **2010**, *404*, 232–246.
- [122] P. De Santis, E. Giglio, A. M. Liquori, A. Ripamonti, *Nature* **1965**, *206*, 456–458.
- [123] B. N. Bullock, A. L. Jochim, P. S. Arora, *J. Am. Chem. Soc.* **2011**, *133*, 14220–14223.
- [124] X. S. Puente, L. M. Sánchez, C. M. Overall, C. López-Otín, *Nat. Rev. Genet.* **2003**, *4*, 544–558.
- [125] P. K. Madala, J. D. A. Tyndall, T. Nall, D. P. Fairlie, *Chem. Rev.* **2010**, *110*, PR1–PR31.
- [126] G. Byk, D. Halle, I. Zeltser, G. Bitan, Z. Selinger, C. Gilon, *J. Med. Chem.* **1996**, *39*, 3174–

- 3178.
- [127] P. 't Hart, P. Hommen, A. Noisier, A. Krzyzanowski, D. Schüler, A. T. Porfetye, M. Akbarzadeh, I. R. Vetter, H. Adihou, H. Waldmann, *Angew. Chemie Int. Ed.* **2021**, *60*, 1813–1820.
- [128] A. Rüegger, M. Kuhn, H. Lichti, H.-R. Loosli, R. Huguenin, C. Quiquerez, A. von Wartburg, *Helv. Chim. Acta* **1976**, *59*, 1075–1092.
- [129] S. Gupta, A. Bakran, R. Johnson, M. Rowland, *Br. J. Clin. Pharmacol.* **1989**, *27*, 353–357.
- [130] J. Chatterjee, C. Gilon, A. Hoffman, H. Kessler, *Acc. Chem. Res.* **2008**, *41*, 1331–1342.
- [131] J. E. Frampton, K. A. Lyseng-Williamson, *Drugs* **2009**, *69*, 1967–1976.
- [132] X. Chen, E. G. Mietlicki-Baase, T. M. Barrett, L. E. McGrath, K. Koch-Laskowski, J. J. Ferrie, M. R. Hayes, E. J. Petersson, *J. Am. Chem. Soc.* **2017**, *139*, 16688–16695.
- [133] M. Werle, A. Bernkop-Schnürch, *Amino Acids* **2006**, *30*, 351–367.
- [134] M. Amblard, J.-A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* **2006**, *33*, 239–254.
- [135] G. Mezö, F. Hudecz, in *Pept. Synth. Appl.*, Humana Press, Totowa, NJ, **2005**, pp. 63–76.
- [136] F. Milletti, *Drug Discov. Today* **2012**, *17*, 850–860.
- [137] G. Tünnemann, R. M. Martin, S. Haupt, C. Patsch, F. Edenhofer, M. C. Cardoso, *FASEB J.* **2006**, *20*, 1775–1784.
- [138] R. M. Steinman, I. S. Mellman, W. A. Muller, Z. A. Cohn, *J. Cell Biol.* **1983**, *96*, 1–27.
- [139] F. Heitz, M. C. Morris, G. Divita, *Br. J. Pharmacol.* **2009**, *157*, 195–206.
- [140] H. T. McMahon, E. Boucrot, *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 517–533.
- [141] M. Kaksonen, A. Roux, *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 313–326.
- [142] M. C. Kerr, R. D. Teasdale, *Traffic* **2009**, *10*, 364–371.
- [143] J. P. Lim, P. A. Gleeson, *Immunol. Cell Biol.* **2011**, *89*, 836–843.
- [144] A. Fittipaldi, A. Ferrari, M. Zoppé, C. Arcangeli, V. Pellegrini, F. Beltram, M. Giacca, *J. Biol. Chem.* **2003**, *278*, 34141–34149.
- [145] A. El-Sayed, S. Futaki, H. Harashima, *AAPS J.* **2009**, *11*, 13–22.
- [146] A. Erazo-Oliveras, N. Muthukrishnan, R. Baker, T. Y. Wang, J. P. Pellois, *Pharmaceuticals* **2012**, *5*, 1177–1209.
- [147] J. K. Allen, D. J. Brock, H. M. Kondow-McConaghy, J. P. Pellois, *Biomolecules* **2018**, *8*, DOI 10.3390/biom8030050.
- [148] J. König, A. Seithel, U. Gradhand, M. F. Fromm, *Naunyn. Schmiedeberg's. Arch. Pharmacol.* **2006**, *372*, 432–443.
- [149] D. E. Smith, B. Cléménçon, M. A. Hediger, *Mol. Aspects Med.* **2013**, *34*, 323–336.
- [150] G. Baumann, P. Mueller, *J. Supramol. Cell. Biochem.* **1974**, *2*, 538–557.

- [151] M. S. P. Sansom, *Eur. Biophys. J.* **1993**, *22*, 105–124.
- [152] E. Gazit, Y. Shai, W. J. Lee, P. T. Brey, *Biochemistry* **1994**, *33*, 10681–10692.
- [153] Y. Pouny, D. Rapaport, Y. Shai, A. Mor, P. Nicolas, *Biochemistry* **1992**, *31*, 12416–12423.
- [154] S. Thennarasu, A. Tan, R. Penumatchu, C. E. Shelburne, D. L. Heyl, A. Ramamoorthy, *Biophys. J.* **2010**, *98*, 248–257.
- [155] D. Derossi, S. Calvet, A. Trembleau, A. Brunissen, G. Chassaing, A. Prochiantz, *J. Biol. Chem.* **1996**, *271*, 18188–18193.
- [156] A. Prochiantz, *Curr. Opin. Neurobiol.* **1996**, *6*, 629–634.
- [157] G. Guidotti, L. Brambilla, D. Rossi, *Trends Pharmacol. Sci.* **2017**, *38*, 406–424.
- [158] Q. Chu, R. E. Moellering, G. J. Hilinski, Y. W. Kim, T. N. Grossmann, J. T. H. Yeh, G. L. Verdine, *Medchemcomm* **2015**, *6*, 111–119.
- [159] G. Rong, C. Wang, L. Chen, Y. Yan, Y. Cheng, *Sci. Adv.* **2020**, *6*, DOI 10.1126/sciadv.aaz1774.
- [160] K. Melikov, L. V. Chernomordik, *Cell. Mol. Life Sci.* **2005**, *62*, 2739–2749.
- [161] A. D. Frankel, C. O. Pabo, *Cell* **1988**, *55*, 1189–1193.
- [162] M. Green, M. Ishino, P. M. Loewenstein, *Cell* **1989**, *58*, 215–223.
- [163] E. Vivès, P. Brodin, B. Lebleu, *J. Biol. Chem.* **1997**, *272*, 16010–16017.
- [164] F. Said Hassane, A. F. Saleh, R. Abes, M. J. Gait, B. Lebleu, *Cell. Mol. Life Sci.* **2010**, *67*, 715–726.
- [165] A. H. Joliot, A. Triller, M. Volovitch, C. Pernelle, A. Prochiantz, *New Biol.* **1991**, *3*, 1121–1134.
- [166] I. D. Alves, M. Carré, M. P. Montero, S. Castano, S. Lecomte, R. Marquant, P. Lecorché, F. Burlina, C. Schatz, S. Sagan, G. Chassaing, D. Braguer, S. Lavielle, *Biochim. Biophys. Acta - Biomembr.* **2014**, *1838*, 2087–2098.
- [167] P. A. Wender, D. J. Mitchell, K. Pattabiraman, E. T. Pelkey, L. Steinman, J. B. Rothbard, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13003–13008.
- [168] G. Tünnemann, G. Ter-Avetisyan, R. M. Martin, M. Stöckl, A. Herrmann, M. C. Cardoso, *J. Pept. Sci.* **2008**, *14*, 469–476.
- [169] M. Morishita, N. Kamei, J. Ehara, K. Isowa, K. Takayama, *J. Control. Release* **2007**, *118*, 177–184.
- [170] A. Tirla, M. Hansen, P. Rivera-Fuentes, *Synlett* **2018**, *29*, 1289–1292.
- [171] A. G. Torres, M. J. Gait, *Trends Biotechnol.* **2012**, *30*, 185–190.
- [172] G. Gasparini, E. Bang, G. Molinard, D. V Tulumello, S. Ward, S. O. Kelley, A. Roux, N. Sakai, S. Matile, **2014**.
- [173] G. Gasparini, G. Sargsyan, E.-K. Bang, N. Sakai, S. Matile, *Angew. Chemie* **2015**, *127*, 7436–7439.

- [174] A. Erazo-Oliveras, K. Najjar, L. Dayani, T. Y. Wang, G. A. Johnson, J. P. Pellois, *Nat. Methods* **2014**, *11*, 861–867.
- [175] K. Najjar, A. Erazo-Oliveras, J. P. Pellois, *J. Vis. Exp.* **2015**, *2015*, 1–9.
- [176] T. Del’Guidice, J. P. Lepetit-Stoffaes, L. J. Bordeleau, J. Roberge, V. Théberge, C. Lauvaux, X. Barbeau, J. Trottier, V. Dave, D. C. Roy, B. Gaillet, A. Garnier, D. Guay, *PLoS One* **2018**, *13*, 1–26.
- [177] S. Krishnamurthy, C. Wohlford-Lenane, S. Kandimalla, G. Sartre, D. K. Meyerholz, V. Théberge, S. Hallée, A. M. Duperré, T. Del’Guidice, J. P. Lepetit-Stoffaes, X. Barbeau, D. Guay, P. B. McCray, *Nat. Commun.* **2019**, *10*, 1–12.
- [178] Y. Sun, S. Y. Lau, Z. W. Lim, S. C. Chang, F. Ghadessy, A. Partridge, A. Miserez, *Nat. Chem.* **2022**, *14*, 274–283.
- [179] D. S. Nielsen, N. E. Shepherd, W. Xu, A. J. Lucke, M. J. Stoermer, D. P. Fairlie, *Chem. Rev.* **2017**, *117*, 8094–8128.
- [180] A. Bak, D. Leung, S. E. Barrett, S. Forster, E. C. Minnihan, A. W. Leithead, J. Cunningham, N. Toussaint, L. S. Crocker, *AAPS J.* **2015**, *17*, 144–155.
- [181] C. B. Fox, J. Kim, L. V. Le, C. L. Nemeth, H. D. Chirra, T. A. Desai, *J. Control. Release* **2015**, *219*, 431–444.
- [182] J. Renukuntla, A. D. Vadlapudi, A. Patel, S. H. S. Boddu, A. K. Mitra, *Int. J. Pharm.* **2013**, *447*, 75–93.
- [183] B. J. Aungst, *AAPS J.* **2012**, *14*, 10–18.
- [184] P. Kurtzhals, S. Havelund, I. Jonassen, B. Kiehr, U. D. Larsen, U. Ribel, J. Markussen, *Biochem. J.* **1995**, *312*, 725–731.
- [185] L. Heinemann, K. Sinha, C. Weyer, M. Loftager, S. Hirschberger, T. Heise, *Diabet. Med.* **1999**, *16*, 332–338.
- [186] M. Lechleitner, F. Hoppichler, *Wiener Medizinische Wochenschrift* **2011**, *161*, 300–304.
- [187] C. J. White, A. K. Yudin, *Nat. Chem.* **2011**, *3*, 509–524.
- [188] K. Deyle, X. D. Kong, C. Heinis, *Acc. Chem. Res.* **2017**, *50*, 1866–1874.
- [189] U. Schmidt, J. Langer, *J. Pept. Res.* **1997**, *49*, 67–73.
- [190] M. Malesevic, U. Strijowski, D. Bächle, N. Sewald, *J. Biotechnol.* **2004**, *112*, 73–77.
- [191] J. Chatterjee, B. Laufer, H. Kessler, *Nat. Protoc.* **2012**, *7*, 432–444.
- [192] B. K. W. Chung, C. J. White, A. K. Yudin, *Nat. Protoc.* **2017**, *12*, 1277–1287.
- [193] V. Cavallaro, P. Thompson, M. Hearn, *J. Pept. Sci.* **1998**, *4*, 335–343.
- [194] J. Fu, M. Tjandra, C. Becker, D. Bednarczyk, M. Capparelli, R. Elling, I. Hanna, R. Fujimoto, M. Furegati, S. Karur, T. Kasprzyk, M. Knapp, K. Leung, X. Li, P. Lu, W. Mergo, C. Miault, S. Ng, D. Parker, Y. Peng, S. Roggo, A. Rivkin, R. L. Simmons, M. Wang, B. Wiedmann, A. H. Weiss, L. Xiao, L. Xie, W. Xu, A. Yifru, S. Yang, B. Zhou, Z. K. Sweeney, *J. Med. Chem.* **2014**, *57*, 8503–8516.

- [195] H. Liu, H. R. Kim, R. N. V. K. Deepak, L. Wang, K. Y. Chung, H. Fan, Z. Wei, C. Zhang, *Nat. Struct. Mol. Biol.* **2018**, *25*, 472–481.
- [196] P. K. A. Jagtap, D. Garg, T. G. Kapp, C. L. Will, O. Demmer, R. Lührmann, H. Kessler, M. Sattler, *J. Med. Chem.* **2016**, *59*, 10190–10197.
- [197] A. Parenty, X. Moreau, J.-M. Campagne, *Chem. Rev.* **2006**, *106*, 911–939.
- [198] S. H. Joo, *Biomol. Ther.* **2012**, *20*, 19–26.
- [199] A. K. Malde, T. A. Hill, A. Iyer, D. P. Fairlie, *Chem. Rev.* **2019**, *119*, 9861–9914.
- [200] A. M. Finch, A. K. Wong, N. J. Paczkowski, S. K. Wadi, D. J. Craik, D. P. Fairlie, S. M. Taylor, *J. Med. Chem.* **1999**, *42*, 1965–1974.
- [201] G. Wang, N. Drinkwater, D. R. Drew, C. A. MacRaid, D. K. Chalmers, B. Mohanty, S. S. Lim, R. F. Anders, J. G. Beeson, P. E. Thompson, S. McGowan, J. S. Simpson, R. S. Norton, M. J. Scanlon, *J. Mol. Biol.* **2016**, *428*, 3986–3998.
- [202] A. Glas, D. Bier, G. Hahne, C. Rademacher, C. Ottmann, T. N. Grossmann, *Angew. Chemie - Int. Ed.* **2014**, *53*, 2489–2493.
- [203] A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris, A. F. Spatola, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 11273–11278.
- [204] A. K. Galande, K. S. Bramlett, J. O. Trent, T. P. Burris, J. L. Wittliff, A. F. Spatola, *ChemBioChem* **2005**, *6*, 1991–1998.
- [205] A. K. Galande, K. S. Bramlett, T. P. Burris, J. L. Wittliff, A. F. Spatola, *J. Pept. Res.* **2004**, *63*, 297–302.
- [206] H. C. Hayes, L. Y. P. Luk, Y. H. Tsai, *Org. Biomol. Chem.* **2021**, *19*, 3983–4001.
- [207] P. Timmerman, J. Beld, W. C. Puijk, R. H. Meloen, *ChemBioChem* **2005**, *6*, 821–824.
- [208] P. Timmerman, W. C. Puijk, R. H. Meloen, *J. Mol. Recognit.* **2007**, *20*, 283–299.
- [209] P. Timmerman, W. C. Puijk, R. S. Boshuizen, P. Van Dijken, J. W. Slootstra, F. J. Beurskens, P. W. H. I. Parren, A. Huber, M. F. Bachmann, R. H. Meloen, *Open Vaccine J.* **2009**, *2*, 56–67.
- [210] A. M. Spokoynny, Y. Zou, J. J. Ling, H. Yu, Y. Lin, B. L. Pentelute, *J. Am. Chem. Soc.* **2013**, *135*, 5946–5949.
- [211] J. E. Moses, A. D. Moorhouse, *Chem. Soc. Rev.* **2007**, *36*, 1249–1262.
- [212] M. Scrima, A. Le Chevalier-Isaad, P. Rovero, A. M. Papini, M. Chorev, A. M. D'Ursi, *European J. Org. Chem.* **2010**, 446–457.
- [213] S. A. Kawamoto, A. Coleska, X. Ran, H. Yi, C.-Y. Yang, S. Wang, *J. Med. Chem.* **2012**, *55*, 1137–1146.
- [214] O. Torres, D. Yüksel, M. Bernardina, K. Kumar, D. Bong, *ChemBioChem* **2008**, *9*, 1701–1705.
- [215] Y. H. Lau, P. De Andrade, S. T. Quah, M. Rossmann, L. Laraia, N. Sköld, T. J. Sum, P. J. E. Rowling, T. L. Joseph, C. Verma, M. Hyvönen, L. S. Itzhaki, A. R. Venkitaraman, C. J.

- Brown, D. P. Lane, D. R. Spring, *Chem. Sci.* **2014**, *5*, 1804–1809.
- [216] Y. H. Lau, P. De Andrade, N. Sköld, G. J. McKenzie, A. R. Venkitaraman, C. Verma, D. P. Lane, D. R. Spring, *Org. Biomol. Chem.* **2014**, *12*, 4074–4077.
- [217] J. R. Kumita, O. S. Smart, G. A. Woolley, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3803–3808.
- [218] L. Chi, O. Sadovski, G. A. Woolley, *Bioconjug. Chem.* **2006**, *17*, 670–676.
- [219] M. R. Jafari, L. Deng, P. I. Kitov, S. Ng, W. L. Matochko, K. F. Tjhung, A. Zeberoff, A. Elias, J. S. Klassen, R. Derda, *ACS Chem. Biol.* **2014**, *9*, 443–450.
- [220] S. Bellotto, S. Chen, I. Rentero Rebollo, H. A. Wegner, C. Heinis, *J. Am. Chem. Soc.* **2014**, *136*, 5880–5883.
- [221] T. D. Clark, M. R. Ghadiri, *J. Am. Chem. Soc.* **1995**, *117*, 12364–12365.
- [222] J. F. Reichwein, C. Versluis, R. M. J. Liskamp, *J. Org. Chem.* **2000**, *65*, 6187–6195.
- [223] J. L. Stymiest, B. F. Mitchell, S. Wong, J. C. Vederas, *Org. Lett.* **2003**, *5*, 47–49.
- [224] S. M. Miles, R. J. Leatherbarrow, S. P. Marsden, W. J. Coates, *Org. Biomol. Chem.* **2004**, *2*, 281–283.
- [225] B. Wels, J. A. W. Kruijtzter, K. Garner, W. A. J. Nijenhuis, W. H. Gispen, R. A. H. Adan, R. M. J. Liskamp, *Bioorganic Med. Chem.* **2005**, *13*, 4221–4227.
- [226] R. N. Chapman, P. S. Arora, *Org. Lett.* **2006**, *8*, 5825–5828.
- [227] T. Y. Yuen, C. J. Brown, Y. Xue, Y. S. Tan, F. J. Ferrer Gago, X. E. Lee, J. Y. Neo, D. Thean, H. Y. K. Kaan, A. W. Partridge, C. S. Verma, D. P. Lane, C. W. Johannes, *Chem. Sci.* **2019**, *10*, 6457–6466.
- [228] P. Y. Chou, G. D. Fasman, *J. Mol. Biol.* **1977**, *115*, 135–175.
- [229] H. Kessler, *Angew. Chemie Int. Ed. English* **1970**, *9*, 219–235.
- [230] D. A. Torchia, *Biochemistry* **1972**, *11*, 1462–1468.
- [231] W. A. Thomas, M. K. Williams, *J. Chem. Soc. Chem. Commun.* **1972**, *0*, 994.
- [232] P. Keim, R. A. Vigna, A. M. Nigen, J. S. Morrow, F. R. Gurd, *J. Biol. Chem.* **1974**, *249*, 4149–4156.
- [233] C. A. Evans, D. L. Rabenstein, **1974**, 7312–7317.
- [234] D. K. Chalmers, G. R. Marshall, **1995**, 5927–5937.
- [235] S. K. Awasthi, S. Raghothama, P. Balaram, *Biochem. Biophys. Res. Commun.* **1995**, *216*, 375–381.
- [236] T. S. Haque, J. C. Little, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 6975–6985.
- [237] M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342–345.
- [238] T. S. Haque, S. H. Gellman, *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.
- [239] S. R. Raghothama, S. K. Awasthi, P. Balaram, *J. Chem. Soc. Perkin Trans. 2* **1998**, 137–

- 143.
- [240] J. A. Robinson, *Acc. Chem. Res.* **2008**, *41*, 1278–1288.
- [241] H. E. Stanger, S. H. Gellman, *J. Am. Chem. Soc.* **1998**, *120*, 4236–4237.
- [242] I. Karle, H. N. Gopi, P. Balaram, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5160–5164.
- [243] J. D. Fisk, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **2000**, *122*, 5443–5447.
- [244] J. M. Langenhan, I. A. Guzei, S. H. Gellman, *Angew. Chemie - Int. Ed.* **2003**, *42*, 2402–2405.
- [245] R. Fasan, R. L. A. Dias, K. Moehle, O. Zerbe, J. W. Vrijbloed, D. Obrecht, J. A. Robinson, *Angew. Chemie - Int. Ed.* **2004**, *43*, 2109–2112.
- [246] R. Fasan, R. L. A. Dias, K. Moehle, O. Zerbe, D. Obrecht, P. R. E. Mittl, M. G. Grütter, J. A. Robinson, *ChemBioChem* **2006**, *7*, 515–526.
- [247] L. R. Masterson, M. A. Etienne, F. Porcelli, G. Barany, R. P. Hammer, G. Veglia, *Biopolymers* **2007**, *88*, 746–753.
- [248] F. Piriou, K. Lintner, S. Femandjian, P. Fromageot, M. C. Khosla, R. R. Smeby, F. M. Bumpus, *Proc. Natl. Acad. Sci.* **1980**, *77*, 82–86.
- [249] B. Laufer, J. Chatterjee, A. O. Frank, H. Kessler, *J. Pept. Sci.* **2009**, *15*, 141–146.
- [250] A. Kapurniotu, A. Schmauder, K. Tenidis, *J. Mol. Biol.* **2002**, *315*, 339–350.
- [251] M. Tatarek-Nossol, L. M. Yan, A. Schmauder, K. Tenidis, G. Westermark, A. Kapurniotu, *Chem. Biol.* **2005**, *12*, 797–809.
- [252] L. M. Yan, M. Tatarek-Nossol, A. Velkova, A. Kazantzis, A. Kapurniotu, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 2046–2051.
- [253] L. M. Yan, A. Velkova, M. Tatarek-Nossol, E. Andreetto, A. Kapurniotu, *Angew. Chemie - Int. Ed.* **2007**, *46*, 1246–1252.
- [254] B. L. Kier, I. Shu, L. A. Eidenschink, N. H. Andersen, *Proc. Natl. Acad. Sci.* **2010**, *107*, 10466–10471.
- [255] L. Wu, D. McElheny, R. Huang, T. A. Keiderling, *Biochemistry* **2009**, *48*, 10362–10371.
- [256] C. M. Santiveri, M. A. Jiménez, *Biopolymers* **2010**, *94*, 779–790.
- [257] R. M. Hughes, M. L. Waters, *J. Am. Chem. Soc.* **2005**, *127*, 6518–6519.
- [258] R. M. Hughes, M. L. Benschoff, M. L. Waters, *Chem. - A Eur. J.* **2007**, *13*, 5753–5764.
- [259] A. J. Riemen, M. L. Waters, *Biochemistry* **2009**, *48*, 1525–1531.
- [260] S. T. Phillips, M. Rezac, U. Abel, M. Kossenjans, P. A. Bartlett, *J. Am. Chem. Soc.* **2002**, *124*, 58–66.
- [261] J. S. Nowick, D. L. Holmes, G. Mackin, G. Noronha, A. J. Shaka, E. M. Smith, *J. Am. Chem. Soc.* **1996**, *118*, 2764–2765.
- [262] U. Nagai, K. Sato, *Tetrahedron Lett.* **1985**, *26*, 647–650.

- [263] M. G. Hinds, N. G. J. Richards, J. A. Robinson, *J. Chem. Soc., Chem. Commun.* **1988**, 1, 1447–1449.
- [264] R. Hirschmann, K. C. Nicolaou, S. Pietranico, J. Salvino, E. M. Leahy, Paul A. Saprengeler, G. Furst, A. B. Smith, C. D. Strader, M. A. Cascieri, M. R. Candelore, C. Donaldson, W. Vale, L. Maechler, *J. Am. Chem. Soc.* **1992**, 114, 9217–9218.
- [265] W. C. Ripka, G. V. De Lucca, A. C. Bach, R. S. Pottorf, J. M. Blaney, *Tetrahedron* **1993**, 49, 3593–3608.
- [266] A. C. Gibbs, L. H. Kondejewski, W. Gronwald, A. M. Nip, R. S. Hodges, B. D. Sykes, D. S. Wishart, *Nat. Struct. Biol.* **1998**, 5, 284–288.
- [267] M. Chorev, E. Roubini, R. L. McKee, S. W. Gibbons, M. E. Goldman, M. P. Caulfield, M. Rosenblatt, *Biochemistry* **1991**, 30, 5968–5974.
- [268] D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh, P. G. Schultz, *J. Am. Chem. Soc.* **1991**, 113, 9391–9392.
- [269] C. E. Schafmeister, J. Po, G. L. Verdine, *J. Am. Chem. Soc.* **2000**, 122, 5891–5892.
- [270] G. H. Bird, N. Madani, A. F. Perry, A. M. Princiotto, J. G. Supko, X. He, E. Gavathiotis, J. G. Sodroski, L. D. Walensky, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 14093–14098.
- [271] G. J. Hilinski, Y. W. Kim, J. Hong, P. S. Kutchukian, C. M. Crenshaw, S. S. Berkovitch, A. Chang, S. Ham, G. L. Verdine, *J. Am. Chem. Soc.* **2014**, 136, 12314–12322.
- [272] Y. Tian, Y. Jiang, J. Li, D. Wang, H. Zhao, Z. Li, *ChemBioChem* **2017**, 18, 2087–2093.
- [273] R. Aurora, G. D. Rose, *Protein Sci.* **1998**, 7, 21–38.
- [274] S. Penel, E. Hughes, A. J. Doig, *J. Mol. Biol.* **1999**, 287, 127–143.
- [275] A. B. Mahon, P. S. Arora, *Drug Discov. Today Technol.* **2012**, 9, e57–e62.
- [276] E. Cabezas, A. C. Satterthwait, *J. Am. Chem. Soc.* **1999**, 121, 3862–3875.
- [277] A. B. Mahon, P. S. Arora, *Chem. Commun.* **2012**, 48, 1416–1418.
- [278] S. Kushal, B. B. Lao, L. K. Henchey, R. Dubey, H. Mesallati, N. J. Traaseth, B. Z. Olenyuk, P. S. Arora, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, 110, 15602–15607.
- [279] C. H. Douse, S. J. Maas, J. C. Thomas, J. A. Garnett, Y. Sun, E. Cota, E. W. Tate, *ACS Chem. Biol.* **2014**, 9, 2204–2209.
- [280] R. P. Cheng, S. H. Gellman, W. F. DeGrado, *Chem. Rev.* **2001**, 101, 3219–3232.
- [281] D. Seebach, J. Gardiner, *Acc. Chem. Res.* **2008**, 41, 1366–1375.
- [282] W. S. Horne, S. H. Gellman, *Acc. Chem. Res.* **2008**, 41, 1399–1408.
- [283] L. K. A. Pils, O. Reiser, *Amino Acids* **2011**, 41, 709–718.
- [284] S. De Pol, C. Zorn, C. D. Klein, O. Zerbe, O. Reiser, *Angew. Chemie - Int. Ed.* **2004**, 43, 511–514.
- [285] G. V. M. Sharma, P. Nagendar, P. Jayaprakash, P. Radha Krishna, K. V. S. Ramakrishna, A. C. Kunwar, *Angew. Chemie* **2005**, 117, 6028–6032.

- [286] M. A. Schmitt, S. H. Choi, I. A. Guzei, S. H. Gellman, *J. Am. Chem. Soc.* **2006**, *128*, 4538–4539.
- [287] R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, D. C. Spellmeyer, R. Tan, A. D. Frankel, D. V. Santi, F. E. Cohen, P. A. Bartlett, *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 9367–9371.
- [288] H. Kessler, *Angew. Chemie Int. Ed. English* **1993**, *32*, 543–544.
- [289] K. Kirshenbaum, A. E. Barron, R. A. Goldsmith, P. Armand, E. K. Bradley, K. T. V. Truong, K. A. Dill, F. E. Cohen, R. N. Zuckermann, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 4303–4308.
- [290] R. N. Zuckermann, *Biopolymers* **2011**, *96*, 545–555.
- [291] Y. U. Kwon, T. Kodadek, *Chem. Biol.* **2007**, *14*, 671–677.
- [292] N. C. Tan, P. Yu, Y. U. Kwon, T. Kodadek, *Bioorganic Med. Chem.* **2008**, *16*, 5853–5861.
- [293] M. K. Shin, Y. J. Hyun, J. H. Lee, H. S. Lim, *ACS Comb. Sci.* **2018**, *20*, 237–242.
- [294] J. T. Nguyen, C. W. Turck, F. E. Cohen, R. N. Zuckermann, W. A. Lim, *Science* **1998**, *282*, 2088–2092.
- [295] O. Demmer, A. O. Frank, F. Hagn, M. Schottelius, L. Marinelli, S. Cosconati, R. Brack-Werner, S. Kremb, H. J. Wester, H. Kessler, *Angew. Chemie - Int. Ed.* **2012**, *51*, 8110–8113.
- [296] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
- [297] F. Albericio, A. Isidro-Ilobet, A. Mercedes, **2009**, 2455–2504.
- [298] P. Alewood, D. Alewood, L. Miranda, S. Love, W. Meutermans, D. Wilson, *Methods Enzymol.* **1997**, *289*, 14–29.
- [299] C. A. G. N. Montalbetti, V. Falque, *Tetrahedron* **2005**, *61*, 10827–10852.
- [300] C. F. McCusker, P. J. Kocienski, F. T. Boyle, A. G. Schätzlein, *Bioorganic Med. Chem. Lett.* **2002**, *12*, 547–549.
- [301] A. A. Aimetti, R. K. Shoemaker, C. C. Lin, K. S. Anseth, *Chem. Commun.* **2010**, *46*, 4061–4063.
- [302] W. L. Xu, A. L. Cui, X. X. Hu, X. F. You, Z. R. Li, J. S. Zheng, *Tetrahedron Lett.* **2015**, *56*, 4796–4799.
- [303] P. M. Cromm, S. Schaubach, J. Spiegel, A. Fürstner, T. N. Grossmann, H. Waldmann, *Nat. Commun.* **2016**, *7*, 4–10.
- [304] J. Openy, G. Amrahova, J. Chang, A. Noisier, P. 't Hart, *Chem. – A Eur. J.* **2022**, DOI 10.1002/chem.202201121.
- [305] P. A. Boriack-Sjodin, K. K. Swinger, *Biochemistry* **2016**, *55*, 1557–1569.
- [306] Q. Wu, M. Schapira, C. H. Arrowsmith, D. Baryshte-Lovejoy, *Nat. Rev. Drug Discov.* **2021**, *20*, 509–530.

- [307] J. Fuhrmann, K. W. Clancy, P. R. Thompson, *Chem. Rev.* **2015**, *115*, 5413–5461.
- [308] S. S. Wolf, *Cell. Mol. Life Sci.* **2009**, *66*, 2109–2121.
- [309] V. M. Richon, D. Johnston, C. J. Sneeringer, L. Jin, C. R. Majer, K. Elliston, L. F. Jerva, M. P. Scott, R. A. Copeland, *Chem. Biol. Drug Des.* **2011**, *78*, 199–210.
- [310] A. Di Lorenzo, M. T. Bedford, *FEBS Lett.* **2011**, *585*, 2024–2031.
- [311] Y. C. Wang, C. Li, *FEBS J.* **2012**, *279*, 932–945.
- [312] C. I. Zurita-Lopez, T. Sandberg, R. Kelly, S. G. Clarke, *J. Biol. Chem.* **2012**, *287*, 7859–7870.
- [313] Y. Wang, W. Hu, Y. Yuan, *J. Med. Chem.* **2018**, *61*, 9429–9441.
- [314] B. J. Calnan, B. Tidor, S. Biancalana, D. Hudson, A. D. Frankel, *Science* **1991**, *252*, 1167–1171.
- [315] N. M. Luscombe, R. A. Laskowski, J. M. Thornton, *Nucleic Acids Res.* **2001**, *29*, 2860–2874.
- [316] D. J. Mandell, I. Chorny, E. S. Groban, S. E. Wong, E. Levine, C. S. Rapp, M. P. Jacobson, *J. Am. Chem. Soc.* **2007**, *129*, 820–827.
- [317] J. Xu, S. Richard, *Mol. Cell* **2021**, *81*, 4357–4368.
- [318] H. Kim, Z. A. Ronai, *Cell Stress* **2020**, *4*, 199–215.
- [319] F. Zhu, L. Rui, *Genes Dis.* **2019**, *6*, 247–257.
- [320] H. Shailesh, Z. Z. Zakaria, R. Baiocchi, S. Sif, *Oncotarget* **2018**, *9*, 36705–36718.
- [321] A. Motolani, M. Martin, M. Sun, T. Lu, *Life* **2021**, *11*, 1–16.
- [322] H. Lin, J. I. Luengo, *Bioorganic Med. Chem. Lett.* **2019**, *29*, 1264–1269.
- [323] X. Li, C. Wang, H. Jiang, C. Luo, *Expert Opin. Ther. Pat.* **2019**, *29*, 97–114.
- [324] K. Feustel, G. S. Falchook, *J. Immunother. Precis. Oncol.* **2022**, *5*, 58–67.
- [325] G. Meister, C. Eggert, D. Bühler, H. Brahms, C. Kambach, U. Fischer, *Curr. Biol.* **2001**, *11*, 1990–1994.
- [326] W. J. Friesen, S. Paushkin, A. Wyce, S. Massenet, G. S. Pesiridis, G. Van Duyne, J. Rappsilber, M. Mann, G. Dreyfuss, *Mol. Cell. Biol.* **2001**, *21*, 8289–8300.
- [327] W. J. Friesen, A. Wyce, S. Paushkin, L. Abel, J. Rappsilber, M. Mann, G. Dreyfuss, *J. Biol. Chem.* **2002**, *277*, 8243–8247.
- [328] H. Brahms, L. Meheus, V. De Brabandere, U. Fischer, R. Lührmann, *Rna* **2001**, *7*, 1531–1542.
- [329] H. Brahms, J. Raymackers, A. Union, F. De Keyser, L. Meheus, R. Lührmann, *J. Biol. Chem.* **2000**, *275*, 17122–17129.
- [330] M. Bezzi, S. X. Teo, J. Muller, W. C. Mok, S. K. Sahu, L. A. Vardy, Z. Q. Bonday, E. Guccione, *Genes Dev.* **2013**, *27*, 1903–1916.

- [331] C. J. Braun, M. Stanciu, P. L. Boutz, J. C. Patterson, D. Calligaris, F. Higuchi, R. Neupane, S. Fenoglio, D. P. Cahill, H. Wakimoto, N. Y. R. Agar, M. B. Yaffe, P. A. Sharp, M. T. Hemann, J. A. Lees, *Cancer Cell* **2017**, *32*, 411-426.e11.
- [332] A. Radzishuskaya, P. V. Shliaha, V. Grinev, E. Lorenzini, S. Kovalchuk, D. Shlyueva, V. Gorshkov, R. C. Hendrickson, O. N. Jensen, K. Helin, *Nat. Struct. Mol. Biol.* **2019**, *26*, 999–1012.
- [333] J. Huang, G. Vogel, Z. Yu, G. Almazan, S. Richard, *J. Biol. Chem.* **2011**, *286*, 44424–44432.
- [334] Y. Dong, C. Song, Y. Wang, Z. Lei, F. Xu, H. Guan, A. Chen, F. Li, *Cell. Signal.* **2017**, *34*, 55–65.
- [335] S. K. Kota, C. Roening, N. Patel, S. B. Kota, R. Baron, *Bone* **2018**, *117*, 37–46.
- [336] P. Beltran-Alvarez, A. Espejo, R. Schmauder, C. Beltran, R. Mrowka, T. Linke, M. Batlle, F. Pérez-Villa, G. J. Pérez, F. S. Scornik, K. Benndorf, S. Pagans, T. Zimmer, R. Brugada, *FEBS Lett.* **2013**, *587*, 3159–3165.
- [337] M. Chen, B. Yi, J. Sun, *J. Biol. Chem.* **2014**, *289*, 24325–24335.
- [338] H. Wei, B. Wang, M. Miyagi, Y. She, B. Gopalan, D.-B. Huang, G. Ghosh, G. R. Stark, T. Lu, *Proc. Natl. Acad. Sci.* **2013**, *110*, 13516–13521.
- [339] E. C. Cho, S. Zheng, S. Munro, G. Liu, S. M. Carr, J. Moehlenbrink, Y. C. Lu, L. Stimson, O. Khan, R. Konietzny, J. McGouran, A. S. Coutts, B. Kessler, D. J. Kerr, N. B. L. Thangue, *EMBO J.* **2012**, *31*, 1785–1797.
- [340] J. H. Park, M. Szemes, G. C. Vieira, Z. Melegh, S. Malik, K. J. Heesom, L. Von Wallwitz-Freitas, A. Greenhough, K. W. Brown, Y. G. Zheng, D. Catchpoole, M. J. Deery, K. Malik, *Mol. Oncol.* **2015**, *9*, 617–627.
- [341] M. Jansson, S. T. Durant, E. C. Cho, S. Sheahan, M. Edelmann, B. Kessler, N. B. La Thangue, *Nat. Cell Biol.* **2008**, *10*, 1431–1439.
- [342] B. P. Pollack, S. V. Kotenko, W. He, L. S. Izotova, B. L. Barnoski, S. Pestka, *J. Biol. Chem.* **1999**, *274*, 31531–31542.
- [343] M. A. Powers, M. M. Fay, R. E. Factor, A. L. Welm, K. S. Ullman, *Cancer Res.* **2011**, *71*, 5579–5587.
- [344] M. Rengasamy, F. Zhang, A. Vashisht, W. M. Song, F. Aguilo, Y. Sun, S. De Li, W. Zhang, B. Zhang, J. A. Wohlschlegel, M. J. Walsh, *Nucleic Acids Res.* **2017**, *45*, 11106–11120.
- [345] N. Stopa, J. E. Krebs, D. Shechter, *Cell. Mol. Life Sci.* **2015**, *72*, 2041–2059.
- [346] X. Bao, S. Zhao, T. Lius, Y. Liu, Y. Liu, X. Yang, *J. Histochem. Cytochem.* **2013**, *61*, 206–217.
- [347] F. Yan, L. Alinari, M. E. Lustberg, L. K. Martin, H. M. Cordero-Nieves, Y. Banasavadi-Siddegowda, S. Virk, J. Barnholtz-Sloan, E. H. Bell, J. Wojton, N. K. Jacob, A. Chakravarti, M. O. Nowicki, X. Wu, R. Lapalombella, J. Datta, B. Yu, K. Gordon, A. Haseley, J. T. Patton, P. L. Smith, J. Ryu, X. Zhang, X. Mo, G. Marcucci, G. Nuovo, C. H. Kwon, J. C. Byrd, E. A. Chiocca, C. Li, S. Sif, S. Jacob, S. Lawler, B. Kaur, R. A. Baiocchi, *Cancer Res.* **2014**, *74*, 1752–1765.

- [348] X. Han, R. Li, W. Zhang, X. Yang, C. G. Wheeler, G. K. Friedman, P. Province, Q. Ding, Z. You, H. M. Fathallah-Shaykh, G. Y. Gillespie, X. Zhao, P. H. King, L. B. Nabors, *J. Neurooncol.* **2014**, *118*, 61–72.
- [349] S. Pal, R. A. Baiocchi, J. C. Byrd, M. R. Grever, S. T. Jacob, S. Sif, *EMBO J.* **2007**, *26*, 3558–3569.
- [350] X. Deng, G. Shao, H. T. Zhang, C. Li, D. Zhang, L. Cheng, B. D. Elzey, R. Pili, T. L. Ratliff, J. Huang, C. D. Hu, *Oncogene* **2017**, *36*, 1223–1231.
- [351] D. Hu, M. Gur, Z. Zhou, A. Gamper, M. C. Hung, N. Fujita, L. Lan, I. Bahar, Y. Wan, *Nat. Commun.* **2015**, *6*, DOI 10.1038/ncomms9419.
- [352] K. Chiang, A. E. Zielinska, A. M. Shaaban, M. P. Sanchez-Bailon, J. Jarrold, T. L. Clarke, J. Zhang, A. Francis, L. J. Jones, S. Smith, O. Barbash, E. Guccione, G. Farnie, M. J. Smalley, C. C. Davies, *Cell Rep.* **2017**, *21*, 3498–3513.
- [353] Z. Wang, J. Kong, Y. Wu, J. Zhang, T. Wang, N. Li, J. Fan, H. Wang, J. Zhang, R. Ling, *Breast Cancer Res. Treat.* **2018**, *168*, 531–542.
- [354] H. Jiang, Y. Zhu, Z. Zhou, J. Xu, S. Jin, K. Xu, H. Zhang, Q. Sun, J. Wang, J. Xu, *Cancer Med.* **2018**, *7*, 869–882.
- [355] K. Zhu, Y. Peng, J. Hu, H. Zhan, L. Yang, Q. Gao, H. Jia, R. Luo, Z. Dai, Z. Tang, J. Fan, J. Zhou, *Carcinogenesis* **2020**, *41*, 130–138.
- [356] L. Liu, X. Zhao, L. Zhao, J. Li, H. Yang, Z. Zhu, J. Liu, G. Huang, *Cancer Res.* **2016**, *76*, 1260–1272.
- [357] Y. Li, N. Chitnis, H. Nakagawa, Y. Kita, S. Natsugoe, Y. Yang, Z. Li, M. Wasik, A. J. P. Klein-Szanto, A. K. Rustgi, J. Alan Diehl, *Cancer Discov.* **2015**, *5*, 288–303.
- [358] X. Quan, W. Yue, Y. Luo, J. Cao, H. Wang, Y. Wang, Z. Lu, *J. Neurochem.* **2015**, *134*, 969–977.
- [359] T. Ratovitski, N. Arbez, J. C. Stewart, E. Chighladze, C. A. Ross, *Cell Cycle* **2015**, *14*, 1716–1729.
- [360] S. Cai, R. Liu, P. Wang, J. Li, T. Xie, M. Wang, Y. Cao, Z. Li, P. Liu, *Front. Pharmacol.* **2020**, *11*, 1–11.
- [361] B. Tan, Q. Liu, L. Yang, Y. Yang, D. Liu, L. Liu, F. Meng, *BMC Cardiovasc. Disord.* **2019**, *19*, 1–10.
- [362] D. P. Harris, S. Bandyopadhyay, T. J. Maxwell, B. Willard, P. E. DiCorleto, *J. Biol. Chem.* **2014**, *289*, 15328–15339.
- [363] D. P. Harris, U. M. Chandrasekharan, S. Bandyopadhyay, B. Willard, P. E. DiCorleto, *PLoS One* **2016**, *11*, e0148905.
- [364] Z. Jia, F. Yue, X. Chen, N. Narayanan, J. Qiu, S. A. Syed, A. N. Imbalzano, M. Deng, P. Yu, C. Hu, S. Kuang, *Adv. Sci.* **2020**, *7*, 1–18.
- [365] S. E. LeBlanc, S. Konda, Q. Wu, Y. J. Hu, C. M. Osowski, S. Sif, A. N. Imbalzano, *Mol. Endocrinol.* **2012**, *26*, 583–597.

- [366] W. W. Tsai, S. Niessen, N. Goebel, J. R. Yates, E. Guccione, M. Montminy, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8870–8875.
- [367] J. Ma, X. He, Y. Cao, K. O'Dwyer, K. M. Szigety, Y. Wu, B. Gurung, Z. Feng, B. W. Katona, X. Hua, *J. Endocrinol.* **2020**, *244*, 41–52.
- [368] R. L. Palte, S. E. Schneider, M. D. Altman, R. P. Hayes, S. Kawamura, B. M. Lacey, M. S. Mansueto, M. Reutershan, P. Siliphaivanh, C. Sondey, H. Xu, Z. Xu, Y. Ye, M. R. Machacek, *ACS Med. Chem. Lett.* **2020**, *11*, 1688–1693.
- [369] NIH, "ClinicalTrials.gov," can be found under <https://clinicaltrials.gov/>, **n.d.**
- [370] S. Antonysamy, Z. Bonday, R. M. Campbell, B. Doyle, Z. Druzina, T. Gheyi, B. Han, L. N. Jungheim, Y. Qian, C. Rauch, M. Russell, J. M. Sauder, S. R. Wasserman, K. Weichert, F. S. Willard, A. Zhang, S. Emtage, *Proc. Natl. Acad. Sci.* **2012**, *109*, 17960–17965.
- [371] M. C. Ho, C. Wilczek, J. B. Bonanno, L. Xing, J. Seznec, T. Matsui, L. G. Carter, T. Onikubo, P. R. Kumar, M. K. Chan, M. Brenowitz, R. H. Cheng, U. Reimer, S. C. Almo, D. Shechter, *PLoS One* **2013**, *8*, 57008.
- [372] D. E. Timm, V. Bowman, R. Madsen, C. Rauch, *PLoS One* **2018**, *13*, 1–13.
- [373] R. Sterner, B. Höcker, *Chem. Rev.* **2005**, *105*, 4038–4055.
- [374] N. Nagano, C. A. Orengo, J. M. Thornton, *J. Mol. Biol.* **2002**, *321*, 741–765.
- [375] B. P. Jain, S. Pandey, *Protein J.* **2018**, *37*, 391–406.
- [376] C. Xu, J. Min, *Protein Cell* **2011**, *2*, 202–214.
- [377] C. U. Stirnimann, E. Petsalaki, R. B. Russell, C. W. Müller, *Trends Biochem. Sci.* **2010**, *35*, 565–574.
- [378] H. Chen, B. Lorton, V. Gupta, D. Shechter, *Oncogene* **2017**, *36*, 373–386.
- [379] K. Saha, G. Adhikary, R. L. Eckert, *J. Invest. Dermatol.* **2016**, *136*, 214–224.
- [380] K. Furuno, T. Masatsugu, M. Sonoda, T. Sasazuki, K. Yamamoto, *Biochem. Biophys. Res. Commun.* **2006**, *345*, 1051–1058.
- [381] G. S. Pesiridis, E. Diamond, G. D. Van Duyne, *J. Biol. Chem.* **2009**, *284*, 21347–21359.
- [382] G. Guderian, C. Peter, J. Wiesner, A. Sickmann, K. Schulze-Osthoff, U. Fischer, M. Grimmler, *J. Biol. Chem.* **2011**, *286*, 1976–1986.
- [383] J. Cox, L. M. Esser, M. Jüdt, K. Schmitz, K. Reiffert, M. Grimmler, B. Stork, S. Wesselborg, C. Peter, *Biol. Chem.* **2022**, *403*, 907–915.
- [384] M. Lacroix, S. El Messaoudi, G. Rodier, A. Le Cam, C. Sardet, E. Fabbrizio, *EMBO Rep.* **2008**, *9*, 452–458.
- [385] X. Le Guezennec, M. Vermeulen, A. B. Brinkman, W. A. M. Hoeijmakers, A. Cohen, E. Lasonder, H. G. Stunnenberg, *Mol. Cell. Biol.* **2006**, *26*, 843–851.
- [386] E. S. Burgos, C. Wilczek, T. Onikubo, J. B. Bonanno, J. Jansong, U. Reime, D. Shechter, *J. Biol. Chem.* **2015**, *290*, 9674–9689.

- [387] F. Liu, X. Zhao, F. Perna, L. Wang, P. Koppikar, O. Abdel-Wahab, M. W. Harr, R. L. Levine, H. Xu, A. Tefferi, A. Deblasio, M. Hatlen, S. Menendez, S. D. Nimer, *Cancer Cell* **2011**, *19*, 283–294.
- [388] D. M. Krüger, H. Gohlke, *Nucleic Acids Res.* **2010**, *38*, 480–486.
- [389] R. A. Laursen, *J. Am. Chem. Soc.* **1966**, *88*, 5344–5346.
- [390] D. Cole, G. Young, A. Weigel, A. Sebesta, P. Kukura, *ACS Photonics* **2017**, *4*, 211–216.
- [391] G. Young, N. Hundt, D. Cole, A. Fineberg, J. Andrecka, A. Tyler, A. Olerinyova, A. Ansari, E. G. Marklund, M. P. Collier, S. A. Chandler, O. Tkachenko, J. Allen, M. Crispin, N. Billington, Y. Takagi, J. R. Sellers, C. Eichmann, P. Selenko, L. Frey, R. Riek, M. R. Galpin, W. B. Struwe, J. L. P. Benesch, P. Kukura, *Science* **2018**, *360*, 423–427.
- [392] K. Hsiao, H. Zegzouti, S. A. Goueli, *Epigenomics* **2016**, DOI 10.2217/epi.15.113.
- [393] K. W. Duncan, N. Rioux, P. A. Boriack-sjodin, M. J. Munchhof, L. A. Reiter, C. R. Majer, L. Jin, L. D. Johnston, E. Chan-penebre, K. G. Kuplast, M. P. Scott, R. M. Pollock, N. J. Waters, J. J. Smith, M. P. Moyer, R. A. Copeland, R. Chesworth, **2016**, DOI 10.1021/acsmchemlett.5b00380.
- [394] N. J. Moerke, *Curr. Protoc. Chem. Biol.* **2009**, *1*, 1–15.
- [395] W. A. Lea, A. Simeonov, *Expert Opin. Drug Discov.* **2011**, *6*, 17–32.
- [396] R. T. Raines, in *Protein-Protein Interact. Methods Appl. Second Ed.*, **2015**, pp. 323–327.
- [397] A. M. Asberry, X. Cai, X. Deng, U. Santiago, S. Liu, H. S. Sims, W. Liang, X. Xu, J. Wan, W. Jiang, C. J. Camacho, M. Dai, **2022**, DOI 10.1021/acs.jmedchem.2c01000.
- [398] Y. Drabsch, P. ten Dijke, *Cancer Metastasis Rev.* **2012**, *31*, 553–568.
- [399] A. Schuetz, A. Allali-Hassani, F. Martín, P. Loppnau, M. Vedadi, A. Bochkarev, A. N. Plotnikov, C. H. Arrowsmith, J. Min, *EMBO J.* **2006**, *25*, 4245–4252.
- [400] K. Gall Trošelj, R. Novak Kujundzic, D. Ugarkovic, *Clin. Epigenetics* **2016**, *8*, 1–10.
- [401] S. Chen, L. Jiao, M. Shubbar, X. Yang, X. Liu, *Mol. Cell* **2018**, *69*, 840-852.e5.
- [402] R. Cao, Y. Zhang, *Mol. Cell* **2004**, *15*, 57–67.
- [403] A. N. Rai, M. L. Vargas, L. Wang, E. F. Andersen, E. L. Miller, J. A. Simon, *Mol. Cell. Biol.* **2013**, *33*, 4844–4856.
- [404] K. Ancelin, U. C. Lange, P. Hajkova, R. Schneider, A. J. Bannister, T. Kouzarides, M. A. Surani, *Nat. Cell Biol.* **2006**, *8*, 623–630.
- [405] Q. Zhao, G. Rank, Y. T. Tan, H. Li, R. L. Moritz, R. J. Simpson, L. Cerruti, D. J. Curtis, D. J. Patel, C. D. Allis, J. M. Cunningham, S. M. Jane, *Nat. Struct. Mol. Biol.* **2009**, *16*, 304–311.
- [406] F. García-Martín, M. Quintanar-Audelo, Y. García-Ramos, L. J. Cruz, C. Gravel, R. Furic, S. Côté, J. Tulla-Puche, F. Albericio, *J. Comb. Chem.* **2006**, *8*, 213–220.
- [407] Y. Yang, W. V. Sweeney, K. Schneider, S. Thörnqvist, B. T. Chait, J. P. Tam, *Tetrahedron Lett.* **1994**, *35*, 9689–9692.

- [408] R. Behrendt, P. White, J. Offer, *J. Pept. Sci.* **2016**, *22*, 4–27.
- [409] R. Behrendt, S. Huber, P. White, *J. Pept. Sci.* **2016**, *22*, 92–97.
- [410] R. Subirós-Funosas, A. El-Faham, F. Albericio, *Biopolymers* **2012**, *98*, 89–97.
- [411] T. Michels, R. Dölling, U. Haberkorn, W. Mier, *Org. Lett.* **2012**, *14*, 5218–5221.
- [412] J. H. Laity, B. M. Lee, P. E. Wright, *Curr. Opin. Struct. Biol.* **2001**, *11*, 39–46.
- [413] A. D. Frankel, J. M. Berg, C. O. Pabo, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 4841–4845.
- [414] M. S. Lee, G. P. Gippert, K. V. Soman, D. A. Case, P. E. Wright, *Science* **1989**, *245*, 635–637.
- [415] S. A. Wolfe, L. Nekludova, C. O. Pabo, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 183–212.
- [416] H. S. Najafabadi, S. Mnaimneh, F. W. Schmitges, M. Garton, K. N. Lam, A. Yang, M. Albu, M. T. Weirauch, E. Radovani, P. M. Kim, J. Greenblatt, B. J. Frey, T. R. Hughes, *Nat. Biotechnol.* **2015**, *33*, 555–562.
- [417] K. J. Brayer, S. Kulshreshtha, D. J. Segal, *Cell Biochem. Biophys.* **2008**, *51*, 9–19.
- [418] K. J. Brayer, D. J. Segal, *Cell Biochem. Biophys.* **2008**, *50*, 111–131.
- [419] J. L. Owens, E. Beketova, S. Liu, S. L. Tinsley, A. M. Asberry, X. Deng, J. Huang, C. Li, J. Wan, C. D. Hu, *iScience* **2020**, *23*, 100750.
- [420] K. Abdelmohsen, M. Gorospe, *RNA Biol.* **2012**, *9*, 799–808.
- [421] M. A. Lischwe, K. D. Roberts, L. C. Yeoman, H. Busch, *J. Biol. Chem.* **1982**, *257*, 14600–14602.
- [422] B. Raman, C. Guarnaccia, K. Nadassy, S. Zakhariev, A. Pintar, F. Zanuttin, D. Frigyes, C. Acatrinei, A. Vindigni, G. Pongor, S. Pongor, *Nucleic Acids Res.* **2001**, *29*, 3377–3384.
- [423] G. J. Pellar, P. J. DiMario, *Chromosoma* **2003**, *111*, 461–469.
- [424] S. Castella, R. Bernard, M. Corno, A. Fradin, J. C. Larcher, *Wiley Interdiscip. Rev. RNA* **2015**, *6*, 243–256.
- [425] K. M. Mulvaney, C. Blomquist, N. Acharya, R. Li, M. J. Ranaghan, M. O’Keefe, D. J. Rodriguez, M. J. Young, D. Kesar, D. Pal, M. Stokes, A. J. Nelson, S. S. Jain, A. Yang, Z. Mullin-Bernstein, J. Columbus, F. K. Bozal, A. Skepner, D. Raymond, S. LaRussa, D. C. McKinney, Y. Freyzon, Y. Baidi, D. Porter, A. J. Aguirre, A. Ianari, B. McMillan, W. R. Sellers, *Mol. Cell* **2021**, *81*, 3481-3495.e7.
- [426] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J. Mol. Biol.* **1990**, *215*, 403–410.
- [427] “Protein BLAST: search protein databases using a protein query,” can be found under https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome
- [428] F. Emma, R. Sanchez-Olea, K. Strange, *Biochim. Biophys. Acta - Mol. Cell Res.* **1998**, *1404*, 321–328.

- [429] B. W. Howell, F. B. Gertler, J. A. Cooper, *EMBO J.* **1997**, *16*, 121–132.
- [430] R. Homayouni, D. S. Rice, M. Sheldon, T. Curran, *J. Neurosci.* **1999**, *19*, 7507–7515.
- [431] D. W. Carr, Z. E. Hausken, I. D. C. Fraser, R. E. Stofko-Hahn, J. D. Scott, *J. Biol. Chem.* **1992**, *267*, 13376–13382.
- [432] D. Mochly-Rosen, H. Khaner, J. Lopez, B. L. Smith, *J. Biol. Chem.* **1991**, *266*, 14866–14868.
- [433] M. C. Souroujon, D. Mochly-Rosen, *Nat. Biotechnol.* **1998**, *16*, 919–924.
- [434] H. Jensen, J. Østergaard, *J. Am. Chem. Soc.* **2010**, *132*, 4070–4071.
- [435] N. N. Poulsen, N. Z. Andersen, J. Østergaard, G. Zhuang, N. J. Petersen, H. Jensen, *Analyst* **2015**, *140*, 4365–4369.
- [436] A. Krzyzanowski, R. Gasper, H. Adihou, P. t. Hart, H. Waldmann, *ChemBioChem* **2021**, *22*, 1908–1914.
- [437] D. C. McKinney, B. J. McMillan, M. J. Ranaghan, J. A. Moroco, M. Brousseau, Z. Mullin-Bernstein, M. O’Keefe, P. McCarren, M. F. Mesleh, K. M. Mulvaney, F. Robinson, R. Singh, B. Bajrami, F. F. Wagner, R. Hilgraf, M. J. Drysdale, A. J. Campbell, A. Skepner, D. E. Timm, D. Porter, V. K. Kaushik, W. R. Sellers, A. Ianari, *J. Med. Chem.* **2021**, *64*, 11148–11168.
- [438] A. El-Faham, R. S. Funosas, R. Prohens, F. Albericio, *Chem. - A Eur. J.* **2009**, *15*, 9404–9416.
- [439] A. El-Faham, F. Albericio, *J. Pept. Sci.* **2010**, *16*, 6–9.
- [440] D. Seebach, A. Thaler, A. K. Beck, *Helv. Chim. Acta* **1989**, *72*, 857–867.
- [441] A. Thaler, D. Seebach, F. Cardinaux, *Helv. Chim. Acta* **1991**, *74*, 628–643.
- [442] S. Sagan, P. Karoyan, O. Lequin, G. Chassaing, S. Lavielle, *Curr. Med. Chem.* **2012**, *11*, 2799–2822.
- [443] S. M. Vogen, N. J. Paczkowski, L. Kirnarsky, A. Short, J. B. Whitmore, S. A. Sherman, S. M. Taylor, S. D. Sanderson, *Int. Immunopharmacol.* **2001**, *1*, 2151–2162.
- [444] S. Kim, G. Biswas, S. Park, A. Kim, H. Park, E. Park, J. Kim, Y. U. Kwon, *Org. Biomol. Chem.* **2014**, *12*, 5222–5226.
- [445] S. Erlendsson, K. Teilum, *Front. Mol. Biosci.* **2021**, *7*, 1–13.
- [446] E. Johansson, R. Caraballo, N. Mistry, G. Zocher, W. Qian, C. D. Andersson, D. L. Hurdiss, N. Chandra, R. Thompson, L. Frängsmyr, T. Stehle, N. Arnberg, M. Elofsson, *ACS Chem. Biol.* **2020**, *15*, 2683–2691.
- [447] G. Winter, *J. Appl. Crystallogr.* **2010**, *43*, 186–190.
- [448] G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, G. Evans, *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 85–97.
- [449] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read,

- J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- [450] P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 352–367.
- [451] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 486–501.
- [452] K. Schmitz, J. Cox, L. M. Esser, M. Voss, K. Sander, A. Löffler, F. Hillebrand, S. Erkelenz, H. Schaal, T. Kähne, S. Klinker, T. Zhang, L. Nagel-Steger, D. Willbold, S. Seggewiß, D. Schlütermann, B. Stork, M. Grimmler, S. Wesselborg, C. Peter, *Nucleic Acids Res.* **2021**, *49*, 6437–6455.
- [453] A. S. Löffler, S. Alers, A. M. Dieterle, H. Keppeler, M. Franz-Wachtel, M. Kundu, D. G. Campbell, S. Wesselborg, D. R. Alessi, B. Stork, *Autophagy* **2011**, *7*, 696–706.
- [454] U. K. LAEMMLI, *Nature* **1970**, *227*, 680–685.
- [455] M. H. M. Olsson, C. R. SØndergaard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
- [456] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- [457] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739–1749.
- [458] T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750–1759.
- [459] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177–6196.

Appendix

8.1. Supplementary Experimental Data

8.1.1. Supplementary Tables

Table S1. Results of the DrugScore^{PPI} analysis of the interface between the TIM barrel domain of PRMT5 and MEP50 in the PDB structure 4GQB for the MEP50 residues. Entries with $\Delta\Delta G_{\text{calc}} > 1$ kcal/mol are shown in bold.

MEP50 Residue	$\Delta\Delta G_{\text{calc}}$ [kcal/mol]	Degree of Buridness	Possible Saltbridges
ARG29	0.48	3.07	No
GLN30	0.28	4.05	No
GLU32	0.27	8.4	Yes
SER47	0.33	9.34	No
SER48	0.23	7.74	No
LEU49	1.63	3.72	No
SER50	0.62	9.25	No
ARG52	0.91	1.54	Yes
CYS53	0.19	3.93	No
TRP54	3.74	8.33	No
GLU80	-0.11	4.78	No
VAL83	0.23	6.37	No
ASP85	0.29	9.92	No
SER98	0.27	9.25	No
ASP99	3.15	6.43	Yes
SER100	0.18	6.23	No
TYR122	-0.22	6.04	No
GLU123	-0.21	2.09	No
HIS124	0.13	7.54	No
ASP125	3.3	5.07	No
ASP126	1.09	5.07	No
ILE127	1.33	6.34	No
VAL128	0.25	8.60	No
LYS145	0.39	5.00	No
ILE147	0.33	6.39	No
LYS150	0.22	7.55	No
SER162	0.14	4.72	No
TYR163	0.23	3.18	No
ARG164	1.54	5.01	No
HIS166	0.15	6.6	No
GLN169	0.2	3.02	No
GLU188	0.18	3.16	No
ASP189	0.48	10.14	No
ARG191	0.68	6.10	Yes
LEU193	0.5	7.39	No
TRP195	0.18	3.43	No
LYS201	1.71	5.00	No
SER204	0.39	9.25	No
GLN205	-0.23	4.05	No
ILE206	0.05	1.83	No
TYR213	1.04	3.22	No
GLU233	0.02	3.16	No
ASN234	0.24	2.37	No
SER255	0.17	4.72	No
GLN256	0.12	4.96	No
CYS257	0.17	5.21	No
GLU276	0.46	3.16	No
ASP277	0.29	9.99	No
CYS278	0.32	7.85	No

Table S2. Results of the DrugScore^{PPI} analysis of the interface between the TIM barrel domain of PRMT5 and MEP50 in the PDB structure 4GQB for the TIM barrel residues. Entries with $\Delta\Delta G_{\text{calc}} > 1$ kcal/mol are shown in bold.

TIM Barrel Residue	$\Delta\Delta G_{\text{calc}}$ [kcal/mol]	Degree of Buridness	Possible Saltbridges
Asn21A	0.54	5.89	No
Cys22A	0.31	5.29	No
Val23A	0.18	5.44	No
Glu25A	0.32	6.86	Yes
Ile26A	0.07	2.42	No
Met43A	0.11	3.67	No
Val45A	0.01	1.11	No
Phe46A	0.17	5.92	No
His47A	0.36	8.70	No
Arg49A	0.88	6.10	Yes
Phe50A	1.01	6.74	No
Lys51A	0.67	2.50	No
Arg52A	0.21	5.31	No
Glu53A	0.35	9.48	No
Gln56A	0.06	1.99	No
Glu57A	0.06	2.15	No
Arg62A	1.04	4.52	Yes
Gln66A	0.02	3.93	No
Thr67A	0.18	4.82	No
Arg68A	3.22	7.68	No
Ser69A	0.14	4.72	No
Leu72A	0.61	3.72	No
Lys85A	0.24	8.32	No
Trp89A	0.18	6.30	No
Arg91A	0.24	4.56	No
Ser94A	0.21	6.23	No
Lys95A	0.05	0.87	No
Val96A	1.62	6.43	No
Lys98A	0.47	1.68	No
Ile99A	0.64	4.50	No
Asn102A	0.29	7.10	No
Asn127A	0.14	4.81	No
Glu161A	0.64	5.37	Yes
Leu163A	0.31	4.50	No
Arg164A	1.36	6.80	No
Asp165A	2.81	3.78	No
Asp166A	0.37	5.07	No
Ile167A	2.44	3.67	No
Ile168A	2.98	8.28	No
Glu169A	0.01	3.10	No
Asn170A	1.59	5.89	No
Thr173A	0.21	7.33	No
His175A	0.11	3.80	No
Thr233A	0.27	8.52	No
Ser234A	0.06	1.60	No
Thr267A	0.21	7.33	No
The269A	0.63	6.08	No
Asn270A	0.17	8.32	No
His271A	0.31	7.59	No

Table S3. Overview of expression systems, vectors and tags initially tested for the expression of human PRMT5 and MEP50 proteins. None of the tested protein constructs could be successfully obtained in isolation.

Expression System	Protein	Vector	Tag
<i>E. coli</i>	PRMT5	pOPIN	N-6His C-8His N-6His-3C N-6His-MBP-3C N-6His-SUMO N-6His-GST-3C N-6His-TRX-3C
		pET19	N-6His-3C C-3C-6His
	MEP50	pOPIN	N-6His C-8His N-6His-3C N-6His-MBP-3C N-6His-SUMO N-6His-GST-3C N-6His-TRX-3C
		pET19	N-6His-3C C-3C-6His
HighFive and Sf9 Insect Cells	PRMT5	pOPIN	N-6His C-8His N-6His-3C N-6His-MBP-3C N-6His-SUMO N-6His-GST-3C N-6His-TRX-3C
	MEP50	pOPIN	N-6His C-8His N-6His-3C N-6His-MBP-3C N-6His-SUMO N-6His-GST-3C N-6His-TRX-3C

Table S4. FIDA results for pICln and RioK1 interactions with PRMT5 protein complexes.

	pICln-Alexa488		RioK1-Alexa488	
	PRMT5-MEP50	TIM-MEP50	PRMT5-MEP50	TIM-MEP50
Affinity constant	39.26 nM	64.11 nM	1.53 nM	20.62 nM
Indicator size	3.53 nm	3.66 nm	3.06 nm	2.89 nm
Complex size	7.10 nm	4.30 nm	6.37 nm	3.73 nm
R ²	0.999	0.996	0.922	0.801

Table S5. Sequences of molecules containing amino acid modifications resulting in weaker affinity to PRMT5-MEP50 than the original linear peptide **158**.

Molecule	Sequence
S1	Ac- Tyr -PGQFDDADK (FITC) -NH ₂
S2	Ac- Phe -PGQFDDADK (FITC) -NH ₂
S3	Ac- hPhe -PGQFDDADK (FITC) -NH ₂
S4	Ac- Trp -PGQFDDADK (FITC) -NH ₂
S5	Ac- Asp -PGQFDDADK (FITC) -NH ₂
S6	Ac- Glu -PGQFDDADK (FITC) -NH ₂
S7	Ac-VPG- Asp -FDDADK (FITC) -NH ₂
S8	Ac-VPG- Dab -FDDADK (FITC) -NH ₂
S9	Ac-VPG- Dab (Alloc) -FDDADK (FITC) -NH ₂
S10	Ac-VPG- Cit -FDDADK (FITC) -NH ₂
S11	Ac-VPGQ- Phe (2-F) -DDADK (FITC) -NH ₂
S12	Ac-VPGQ- Phe (F₅) -DDADK (FITC) -NH ₂
S13	Ac-VPGQ- Phe (2-I) -DDADK (FITC) -NH ₂
S14	Ac-VPGQ- Phe (3-Cl) -DDADK (FITC) -NH ₂
S15	Ac-VPGQ- Phe (3-CF₃) -DDADK (FITC) -NH ₂
S16	Ac-VPGQ- 2-Pal -DDADK (FITC) -NH ₂
S17	Ac-VPGQ- Tyr -DDADK (FITC) -NH ₂
S18	Ac-VPGQ- Phe (4-COOH) -DDADK (FITC) -NH ₂
S19	Ac-VPGQ- Phe (4-guanidino) -DDADK (FITC) -NH ₂
S20	Ac-VPGQ- hPhe -DDADK (FITC) -NH ₂
S21	Ac-VPGQ- 3- (2-biphenyl) -Ala -DDADK (FITC) -NH ₂
S22	Ac-VPGQ- Bpa -DDADK (FITC) -NH ₂
S23	Ac-VPGQ- 2-Nal -DDADK (FITC) -NH ₂
S24	Ac-VPGQ- 1-Nal -DDADK (FITC) -NH ₂
S25	Ac-VPGQFD- Gla -ADK (FITC) -NH ₂
S26	Ac-VPGQFDD- Abu -DK (FITC) -NH ₂
S27	Ac-VPGQFDD- Nva -DK (FITC) -NH ₂
S28	Ac-VPGQFDD- Nle -DK (FITC) -NH ₂
S29	Ac-VPGQFDD- Cha -DK (FITC) -NH ₂
S30	Ac-VPGQFDD- Ser -DK (FITC) -NH ₂
S31	Ac-VPGQFDDA- Tyr -K (FITC) -NH ₂
S32	Ac-VPGQFDDA- Trp -K (FITC) -NH ₂
S33	Ac-VPGQFDDA- hPhe -K (FITC) -NH ₂

Table S6. HRMS data for synthesised peptides.

Molecule	HRMS	
	m/z calculated	m/z found
1	849.40358 [M+H] ⁺	849.40315
2	920.44070 [M+H] ⁺	920.44034
3	1007.47273 [M+H] ⁺	1007.47263
4	1064.49419 [M+H] ⁺	1064.49450
5	1222.56333 [M+H] ⁺	1222.56410
6	1341.50270 [M+H] ⁺	1341.50355
7	1412.53982 [M+H] ⁺	1412.54060
8	1499.57185 [M+H] ⁺	1499.57283
9	1556.59331 [M+H] ⁺	1556.59360
10	1714.66245 [M+H] ⁺	1714.66258
11	790.36647 [M+H] ⁺	790.36617
12	861.40358 [M+H] ⁺	861.40352
13	875.41923 [M+H] ⁺	875.41930
14	932.44070 [M+H] ⁺	932.44093
15	946.45635 [M+H] ⁺	946.45654
16	1019.47273 [M+H] ⁺	1019.47318
17	1033.48838 [M+H] ⁺	1033.48895
18	861.40358 [M+H] ⁺	861.40372
19	861.40358 [M+H] ⁺	861.40404
20	918.42505 [M+H] ⁺	918.42500
21	918.42505 [M+H] ⁺	918.42576
22	1005.45708 [M+H] ⁺	1005.45745
23	1005.45708 [M+H] ⁺	1005.45817
24	1017.49346 [M+H] ⁺	1017.49355
25	1017.49346 [M+H] ⁺	1017.49386
26	1074.51492 [M+H] ⁺	1074.51554
27	1074.51492 [M+H] ⁺	1074.51558
28	1104.52549 [M+H] ⁺	1104.52601
29	1104.52549 [M+H] ⁺	1104.52613
30	1161.54695 [M+H] ⁺	1161.54760
31	1161.54695 [M+H] ⁺	1161.54722
32	755.29993 [M+2H] ²⁺	755.30066
33	783.81066 [M+2H] ²⁺	783.81124
34	783.81066 [M+2H] ²⁺	783.81118
35	798.81594 [M+2H] ²⁺	798.81582
36	798.81594 [M+2H] ²⁺	798.81653
37	827.32668 [M+2H] ²⁺	827.32638
38	827.32668 [M+2H] ²⁺	827.32708
39	915.45304 [M+H] ⁺	915.45558
40	986.49015 [M+H] ⁺	986.49186
41	1099.57421 [M+H] ⁺	1099.57552
42	1083.54291 [M+H] ⁺	1083.54636
43	1196.62698 [M+H] ⁺	1196.62949
44	1184.59059 [M+H] ⁺	1184.59286
45	1297.67466 [M+H] ⁺	1297.67818
46	955.47310 [M+H] ⁺	955.47565
47	969.48875 [M+H] ⁺	969.48967
48	969.48875 [M+H] ⁺	969.48965
49	983.50440 [M+H] ⁺	983.50526

Appendix

50	983.50440 [M+H] ⁺	983.50540
51	997.52005 [M+H] ⁺	997.52212
52	1237.61714 [M+H] ⁺	1237.61975
53	1251.63279 [M+H] ⁺	1251.63560
54	1251.63279 [M+H] ⁺	1251.63572
55	1265.64844 [M+H] ⁺	1265.65112
56	1237.61714 [M+H] ⁺	1237.62025
57	1251.63279 [M+H] ⁺	1251.63678
58	1265.64844 [M+H] ⁺	1265.65311
59	1279.66409 [M+H] ⁺	1279.66880
60	1293.67974 [M+H] ⁺	1293.68474
61	1294.63860 [M+H] ⁺	1294.64174
62	1308.65425 [M+H] ⁺	1308.65823
63	1322.66990 [M+H] ⁺	1322.67198
64	1336.68555 [M+H] ⁺	1336.68800
65	1350.70120 [M+H] ⁺	1350.70633
66	1346.18587 [M+2H] ²⁺	1346.18590
67	1312.69059 [M+2H] ²⁺	1312.69193
68	1200.98125 [M+3H] ³⁺	1200.98243
69	954.49485 [M+3H] ³⁺	954.49568
70	932.16466 [M+3H] ³⁺	932.16531
71	755.01277 [M+5H] ⁵⁺	755.01310
72	1373.92337 [M+4H] ⁴⁺	1373.93147
73	1243.09157 [M+4H] ⁴⁺	1243.09787
74	1379.25042 [M+3H] ³⁺	1379.25276
75	822.32372 [M+3H] ³⁺	822.32763
76	1108.55520 [M+2H] ²⁺	1108.56019
77	1342.08155 [M+2H] ²⁺	1342.08896
78	979.90995 [M+2H] ²⁺	979.91523
79	1993.74884 [M+H] ⁺	1993.75151
80	1722.62079 [M+H] ⁺	1722.62497
81	1506.52754 [M+H] ⁺	1506.52922
82	1255.42570 [M+H] ⁺	1255.43126
83	1751.70782 [M+H] ⁺	1751.71434
84	989.36241 [M+2H] ²⁺	989.36276
85	996.37023 [M+2H] ²⁺	996.37044
86	1838.73264 [M+H] ⁺	1838.74403
87	1783.67921 [M+H] ⁺	1783.68691
88	1018.86367 [M+2H] ²⁺	1018.86502
89	1252.49794 [M+H] ⁺	1252.50244
90	1243.01976 [M+2H] ²⁺	1243.02345
91	1328.07253 [M+2H] ²⁺	1328.07628
92	1067.92038 [M+2H] ²⁺	1067.92328
93	1064.41546 [M+2H] ²⁺	1064.41816
94	1149.46822 [M+2H] ²⁺	1149.47222
95	1049.91476 [M+2H] ²⁺	1049.91657
96	1098.47426 [M+2H] ²⁺	1098.47733
97	1183.52703 [M+2H] ²⁺	1183.53018
98	1652.66858 [M+H] ⁺	1652.66956
99	1422.61470 [M+H] ⁺	1422.61635
100	1353.54562 [M+H] ⁺	1353.54723
101	1496.53109 [M+H] ⁺	1496.53270

Appendix

102	1300.40991 [M+H] ⁺	1300.41144
103	1115.32987 [M+H] ⁺	1115.33101
104	1004.92272 [M+2H] ²⁺	1004.92569
105	1666.63658 [M+H] ⁺	1666.63968
106	1470.51544 [M+H] ⁺	1470.51881
107	1285.43540 [M+H] ⁺	1285.43804
108	1635.72449 [M+H] ⁺	1635.72651
109	1291.58009 [M+H] ⁺	1291.58152
110	1544.62096 [M+H] ⁺	1544.62257
111	1346.63352 [M+H] ⁺	1346.63623
112	1302.64369 [M+H] ⁺	1302.64621
113	1302.64369 [M+H] ⁺	1302.64633
114	1302.64369 [M+H] ⁺	1302.63783
115	1270.60222 [M+H] ⁺	1270.60403
116	1289.61206 [M+H] ⁺	1289.61450
117	1360.64917 [M+H] ⁺	1360.64443
118	1320.61787 [M+H] ⁺	1320.61953
119	1318.60222 [M+H] ⁺	1318.60393
120	1318.60222 [M+H] ⁺	1318.60389
121	1261.56952 [M+H] ⁺	1261.57166
122	1330.63860 [M+H] ⁺	1330.63818
123	1034.95911 [M+2H] ²⁺	1034.96144
124	1023.93055 [M+2H] ²⁺	1023.93316
125	1014.94346 [M+2H] ²⁺	1014.94578
126	1023.93055 [M+2H] ²⁺	1023.93302
127	994.92781 [M+2H] ²⁺	994.93060
128	1003.91490 [M+2H] ²⁺	1003.91768
129	1878.76394 [M+H] ⁺	1878.76453
130	1302.64369 [M+H] ⁺	1302.64518
131	1360.64917 [M+H] ⁺	1360.65083
132	1316.65934 [M+H] ⁺	1316.66100
133	925.37795 [M+2H] ²⁺	925.37895
134	932.38578 [M+2H] ²⁺	932.38677
135	932.38578 [M+2H] ²⁺	932.38673
136	939.39360 [M+2H] ²⁺	939.39476
137	925.37795 [M+2H] ²⁺	925.37888
138	932.38578 [M+2H] ²⁺	932.38677
139	932.38578 [M+2H] ²⁺	932.38673
140	939.39360 [M+2H] ²⁺	939.39480
141	916.88287 [M+2H] ²⁺	916.88398
142	917.89069 [M+2H] ²⁺	917.89179
143 isomer 1	916.88287 [M+2H] ²⁺	916.88396
143 isomer 2	916.88287 [M+2H] ²⁺	916.88397
144	917.89069 [M+2H] ²⁺	917.89172
145	758.29791 [M+2H] ²⁺	758.29659
146 isomer 1	765.30573 [M+2H] ²⁺	765.30446
146 isomer 2	765.30573 [M+2H] ²⁺	765.30464
147	765.30573 [M+2H] ²⁺	765.30440
148	772.31356 [M+2H] ²⁺	772.31227
149	766.31356 [M+2H] ²⁺	766.31240
150	773.32138 [M+2H] ²⁺	773.32047
151	765.30573 [M+2H] ²⁺	765.30441

Appendix

152	765.30573	[M+2H] ²⁺	765.30475
153	772.31356	[M+2H] ²⁺	772.31239
154	766.31356	[M+2H] ²⁺	766.31275
155	773.32138	[M+2H] ²⁺	773.32047
156	1014.48909	[M+H] ⁺	1014.48752
157	1014.48909	[M+H] ⁺	1014.48727
158	1521.56272	[M+H] ⁺	1521.56580
159	1004.43197	[M+H] ⁺	1004.42952
160	915.86219	[M+3H] ³⁺	915.86281
161	964.22016	[M+3H] ³⁺	964.22076
162	1093.89875	[M+3H] ³⁺	1093.89898
163	1142.25672	[M+3H] ³⁺	1142.25669
164	541.01212	[M+3H] ³⁺	541.01189
165	719.04868	[M+3H] ³⁺	719.04846
166	1539.55330	[M+H] ⁺	1539.55557
167	1535.57837	[M+H] ⁺	1535.58145
168	1597.59402	[M+H] ⁺	1597.59731
169	1557.54388	[M+H] ⁺	1557.54775
170	1539.55330	[M+H] ⁺	1539.55660
171	1599.47323	[M+H] ⁺	1599.47609
172	1647.45936	[M+H] ⁺	1647.46158
173	1555.52375	[M+H] ⁺	1555.52694
174	1579.56820	[M+H] ⁺	1579.57062
175	1566.54780	[M+H] ⁺	1566.55130
176	1613.52923	[M+H] ⁺	1613.53080
177	1624.55328	[M+H] ⁺	1624.55454
178	1535.57837	[M+H] ⁺	1535.58195
179	1535.57837	[M+H] ⁺	1535.58169
180	1535.57837	[M+H] ⁺	1535.58178
181	1535.57837	[M+H] ⁺	1535.58196
182	1535.57837	[M+H] ⁺	1535.58188
183	1535.57837	[M+H] ⁺	1535.58070
184	1634.61040	[M+H] ⁺	1634.61054
185	1634.61040	[M+H] ⁺	1634.61064
186	1438.48922	[M+H] ⁺	1438.48948
187	1117.47964	[M+H] ⁺	1117.48050
188	1117.47964	[M+H] ⁺	1117.48031
189	1460.66521	[M+H] ⁺	1460.66729
190	1460.66521	[M+H] ⁺	1460.66757
H4 tail peptide (MTase-Glo™ Substrate)	1066.64103	[M+2H] ²⁺	1066.64198
S1	1585.55764	[M+H] ⁺	1585.56081
S2	1569.56272	[M+H] ⁺	1569.56604
S3	1583.57837	[M+H] ⁺	1583.58125
S4	1608.57362	[M+H] ⁺	1608.57625
S5	1537.52125	[M+H] ⁺	1537.52441
S6	1551.53690	[M+H] ⁺	1551.54015
S7	1507.54707	[M+H] ⁺	1507.55099
S8	1493.56781	[M+H] ⁺	1493.56975
S9	1577.58894	[M+H] ⁺	1577.59177
S10	1550.58927	[M+H] ⁺	1550.59317
S11	1539.55330	[M+H] ⁺	1539.55666
S12	1611.51561	[M+H] ⁺	1611.51901

Appendix

S13	1647.45936 [M+H] ⁺	1647.46197
S14	1555.52375 [M+H] ⁺	1555.52705
S15	1589.55011 [M+H] ⁺	1589.55440
S16	1522.55797 [M+H] ⁺	1522.56122
S17	1537.55764 [M+H] ⁺	1537.56120
S18	1565.55255 [M+H] ⁺	1565.55556
S19	1578.59542 [M+H] ⁺	1578.59706
S20	1535.57837 [M+H] ⁺	1535.58161
S21	1597.59402 [M+H] ⁺	1597.59787
S22	1625.58894 [M+H] ⁺	1625.59170
S23	1571.57837 [M+H] ⁺	1571.58183
S24	1571.57837 [M+H] ⁺	1571.58167
S25	1579.56820 [M+H] ⁺	1579.57063
S26	1535.57837 [M+H] ⁺	1535.58199
S27	1549.59402 [M+H] ⁺	1549.59737
S28	1563.60967 [M+H] ⁺	1563.61310
S29	1603.64097 [M+H] ⁺	1603.64485
S30	1537.55764 [M+H] ⁺	1537.56164
S31	1569.59911 [M+H] ⁺	1569.60176
S32	1592.61509 [M+H] ⁺	1592.61751
S33	1567.61984 [M+H] ⁺	1567.62296

Table S7. Data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

Wavelength	0.99986
Resolution range	46.74 - 2.55 (2.641 - 2.55)
Space group	P 31 2 1
Unit cell	96.63 96.63 112.68 90 90 120
Total reflections	409259 (41998)
Unique reflections	20296 (1991)
Multiplicity	20.2 (21.1)
Completeness (%)	99.58 (99.50)
Mean I/sigma(I)	12.05 (1.33)
Wilson B-factor	71.65
R-merge	0.1916 (4.157)
R-meas	0.1966 (4.26)
R-pim	0.04392 (0.9258)
CC1/2	0.999 (0.829)
CC*	1 (0.952)
Reflections used in refinement	20232 (1983)
Reflections used for R-free	1012 (99)
R-work	0.2269 (0.3443)
R-free	0.2552 (0.3418)
CC(work)	0.959 (0.819)
CC(free)	0.836 (0.751)
Number of non-hydrogen atoms	1844
macromolecules	1844
Protein residues	225
RMS(bonds)	0.010
RMS(angles)	1.30
Ramachandran favored (%)	94.01
Ramachandran allowed (%)	5.53
Ramachandran outliers (%)	0.46
Average B-factor	88.41
macromolecules	88.41
Number of TLS groups	7

8.1.2. Supplementary Figures

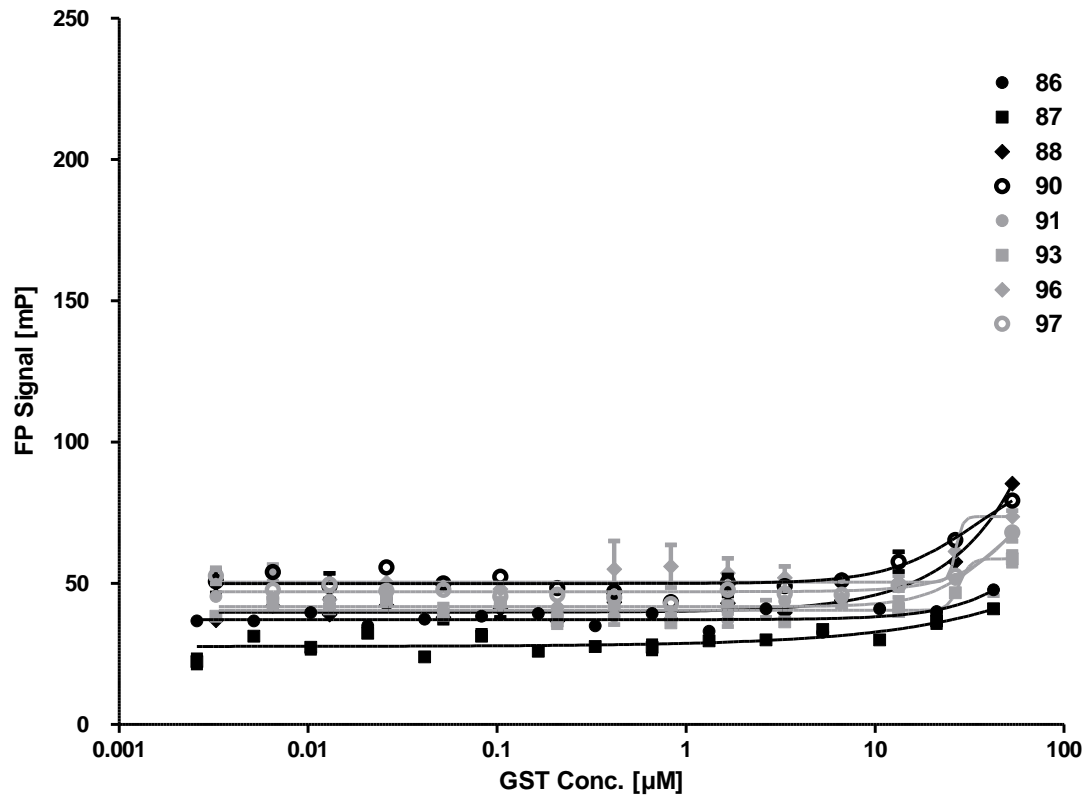


Figure S1. Direct Binding FP with GST protein.

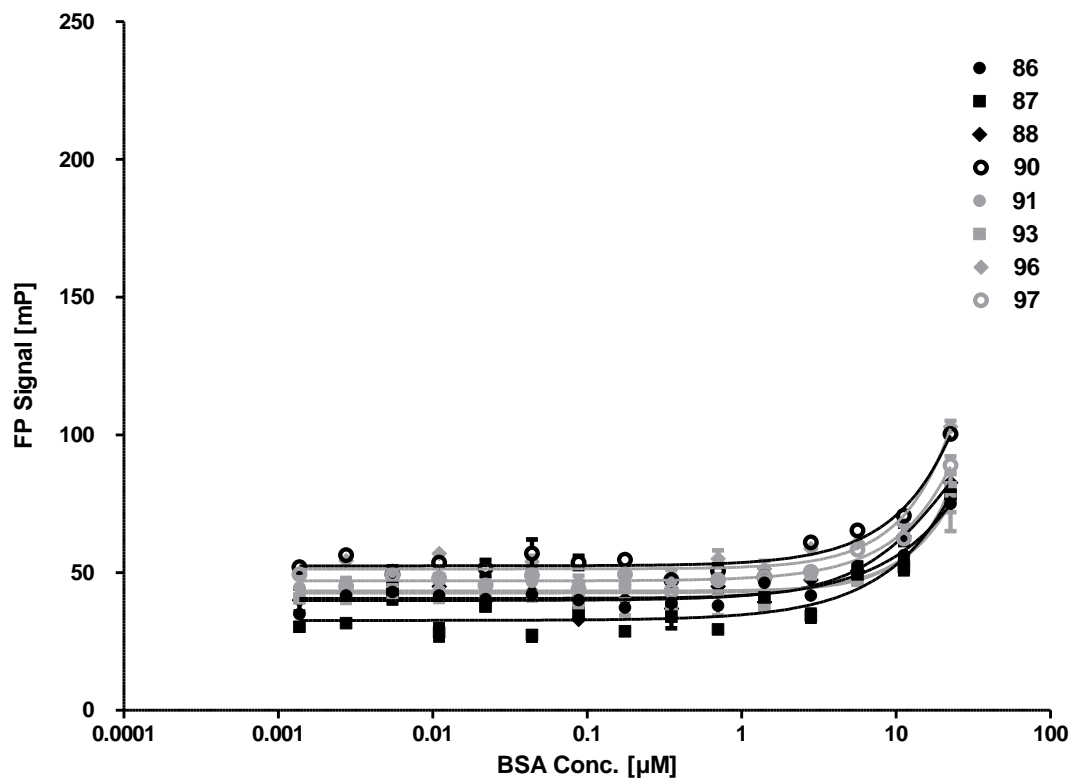


Figure S2. Direct binding FP with BSA protein.

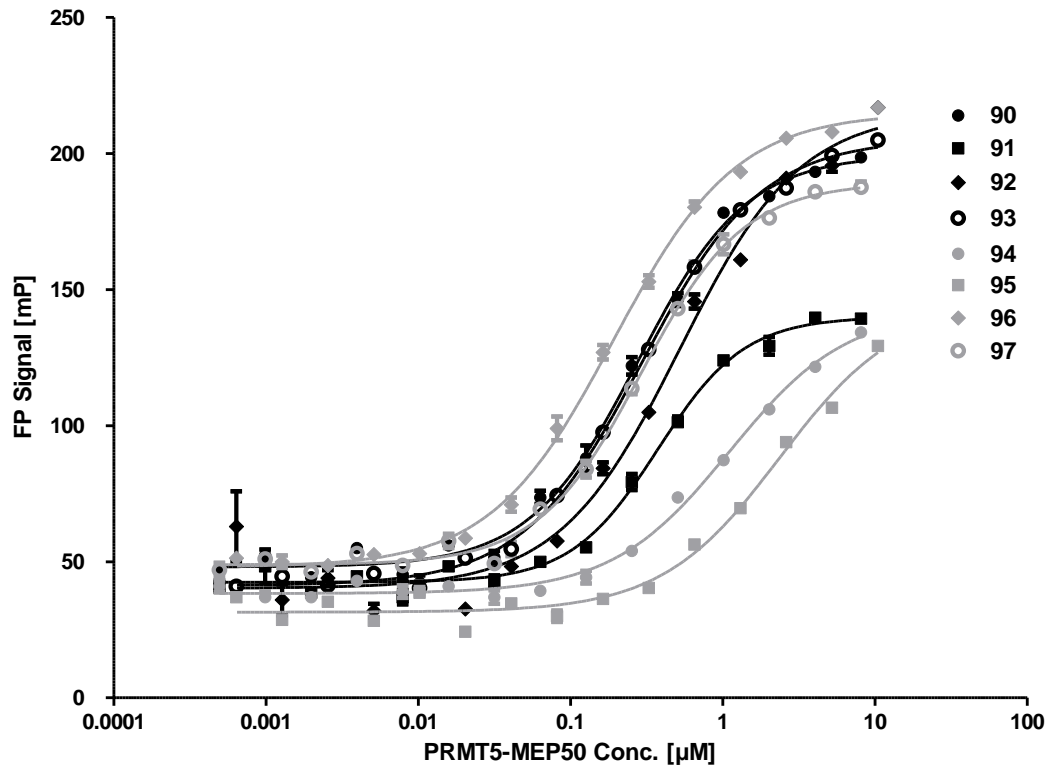


Figure S3. Direct binding FP with PRMT5-MEP50.

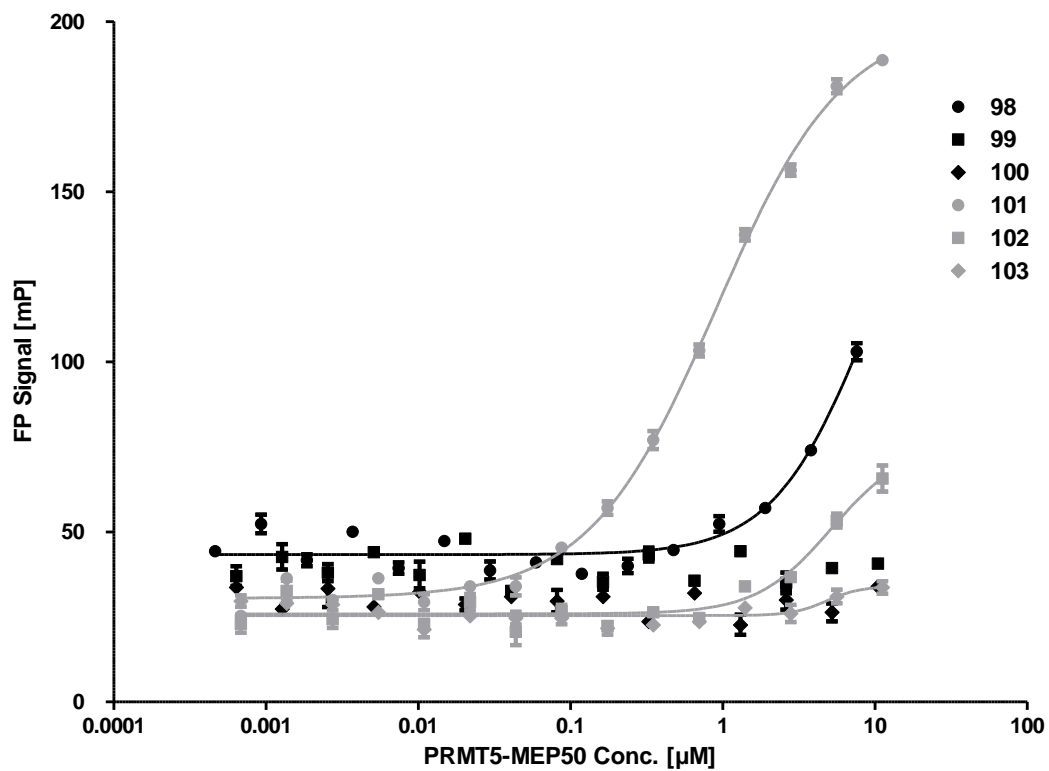


Figure S4. Direct binding FP with PRMT5-MEP50.

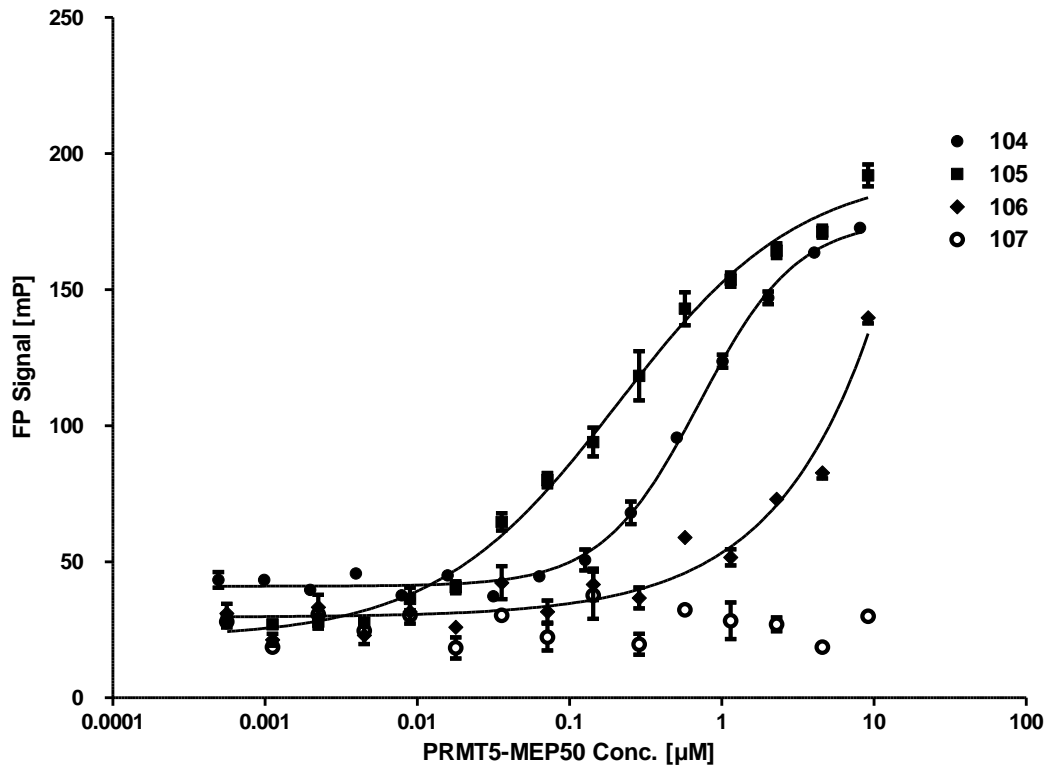


Figure S5. Direct binding FP with PRMT5-MEP50.

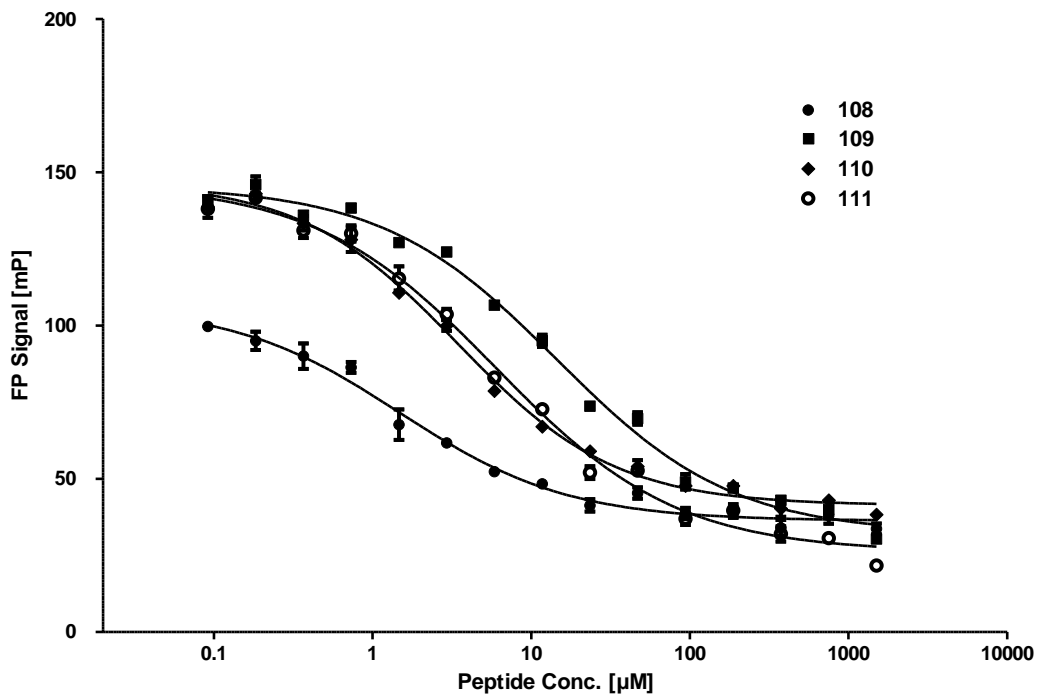


Figure S6. Direct binding FP with PRMT5-MEP50.

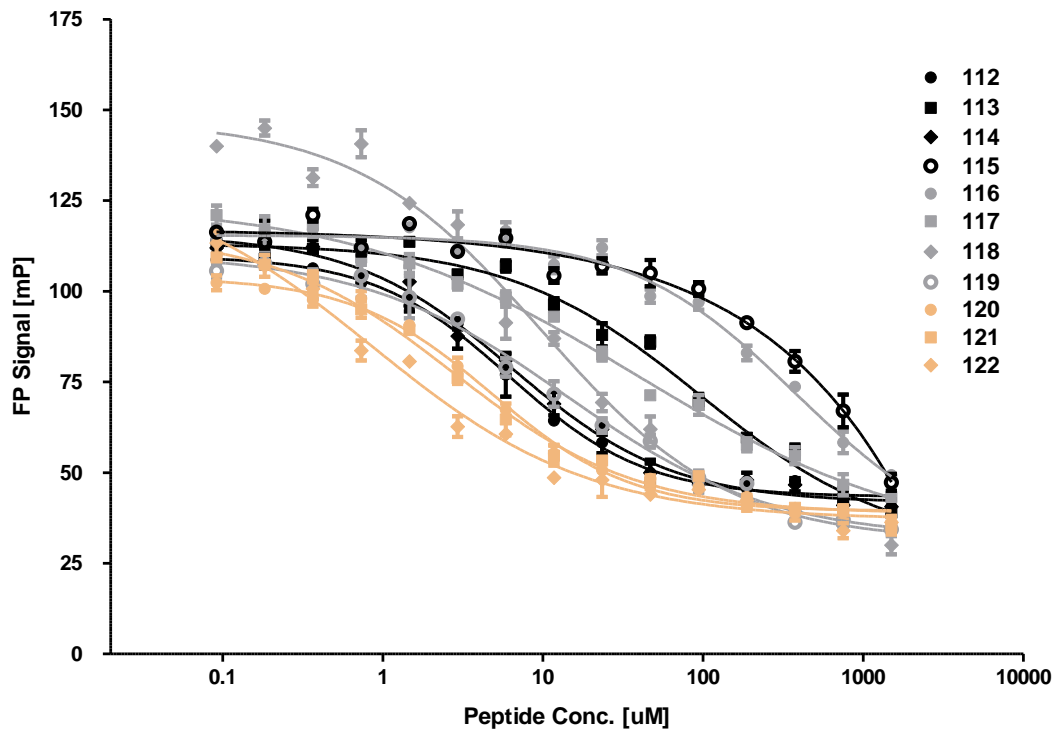


Figure S7. Competitive binding FP with PRMT5-MEP50 and fluorescently labelled compound **93** used as a tracer.

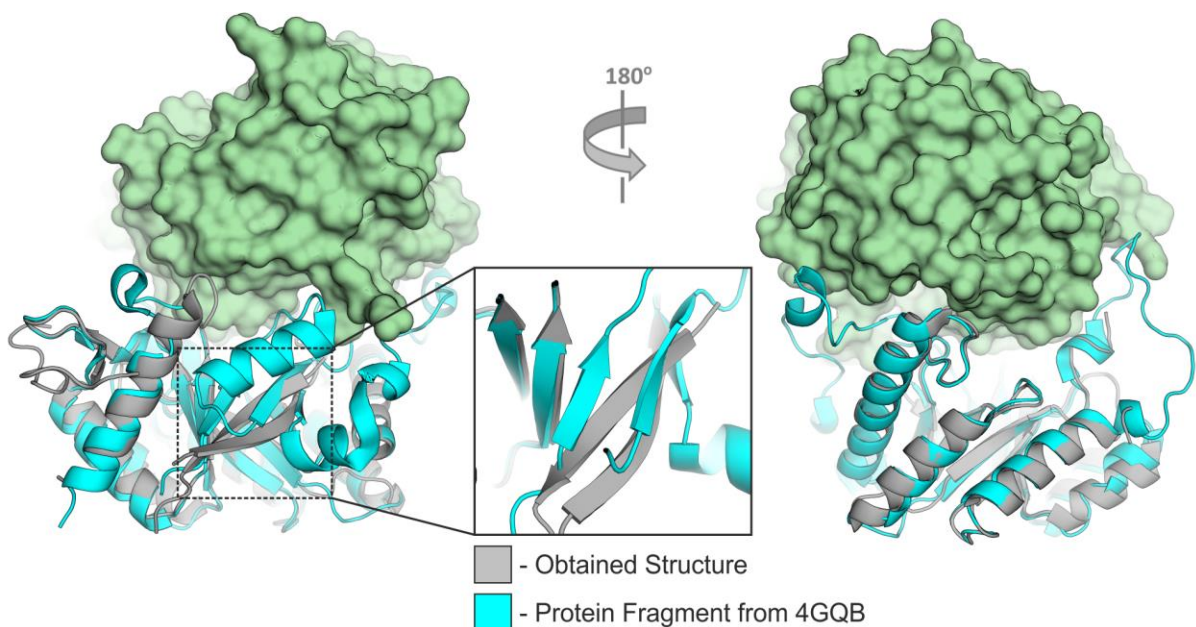


Figure S8. Superposition of the crystallised TIM barrel domain of PRMT5 (PDB ID: 7BOC, gray) and a fragment of PRMT5-MEP50 (PDB ID: 4GQB, turquoise and mint green). The expanded image focuses on the area of the modified β -strand protein core, with the antiparallel strand.

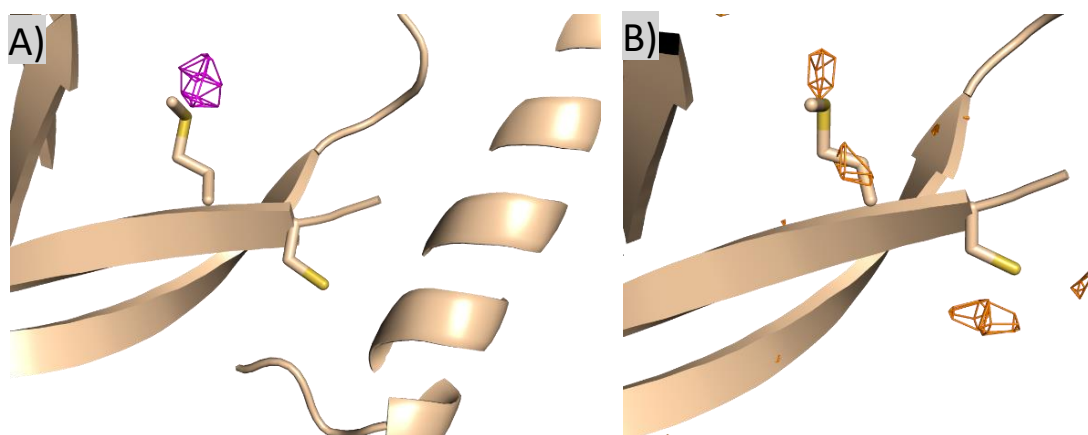


Figure S9. Depiction of the anomalous density calculations: **A)** The structure of the antiparallel strand in the TIM barrel domain with the side chains of Cys42 and Met43 residues visualised. The purple density represents the anomalous density of a Pt-derivative at 0.86Å (contour level of 3.7). **B)** The structure of the antiparallel strand in the TIM barrel domain with the side chains of Cys42 and Met43 residues visualised. The orange density shows the anomalous signal of the native protein crystal at 2.0 Å (contour level of 2.7).

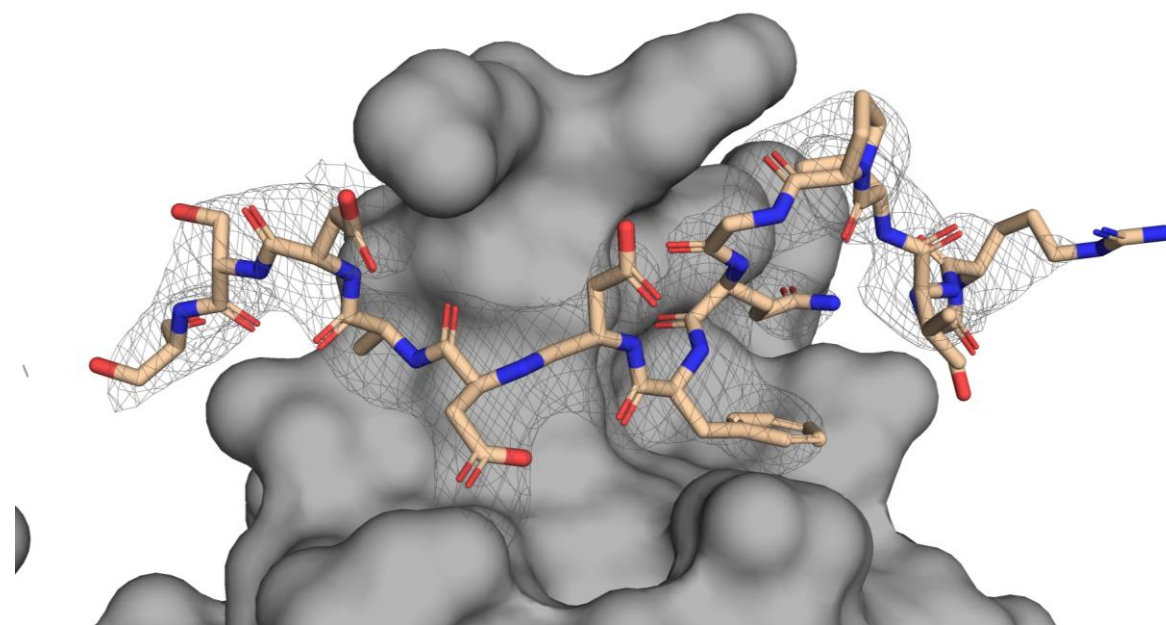


Figure S10. Visualisation of the 2Fo-Fc electron density for peptide **108** bound to the TIM barrel domain (PDB ID: 7BOC). The electron density is contoured at 2.5σ .

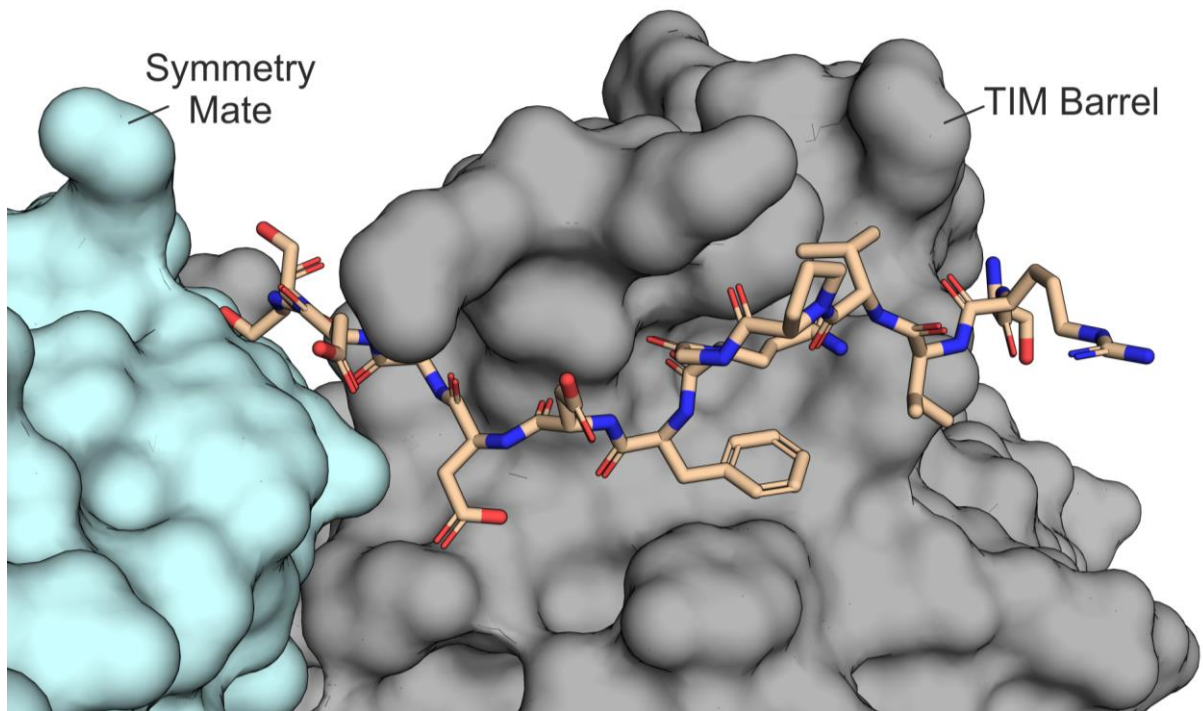


Figure S11. TIM barrel symmetry partner fixes the C-terminal fragment of the protein-bound peptide **108**.

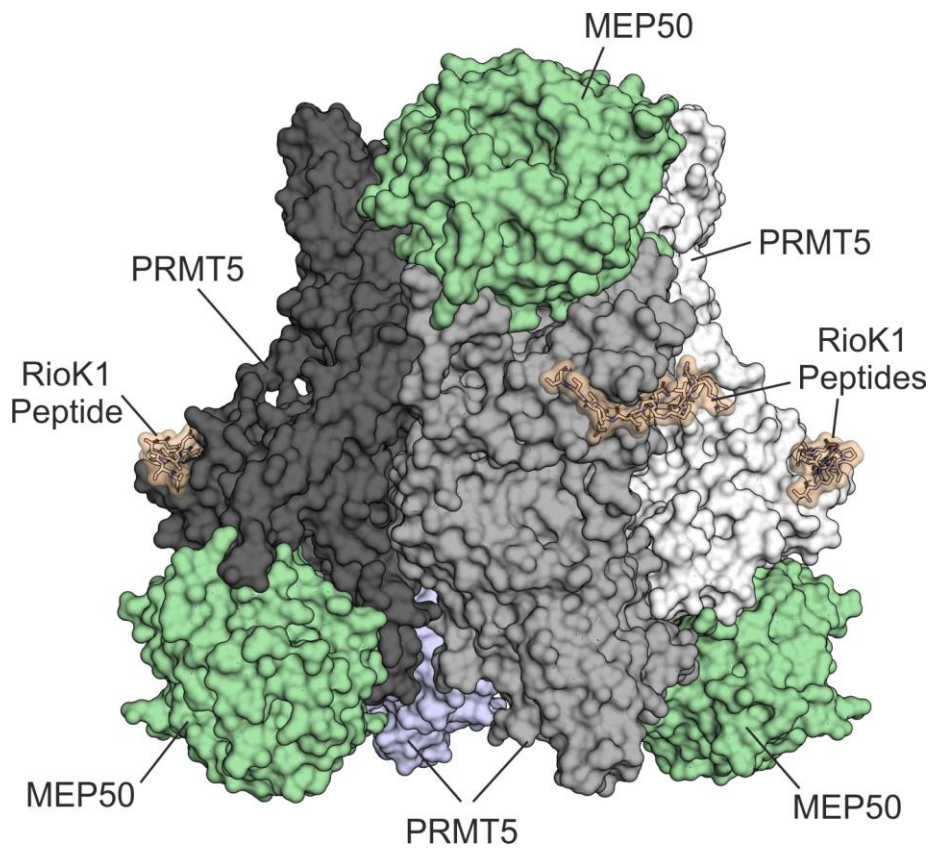


Figure S12. Depiction of the heterooctameric PRMT5-MEP50 complex (PDB ID: 4GQB) with peptide **108** structures fitted through a superposition with the obtained TIM-peptide co-crystal structure (PDB ID: 7BOC).

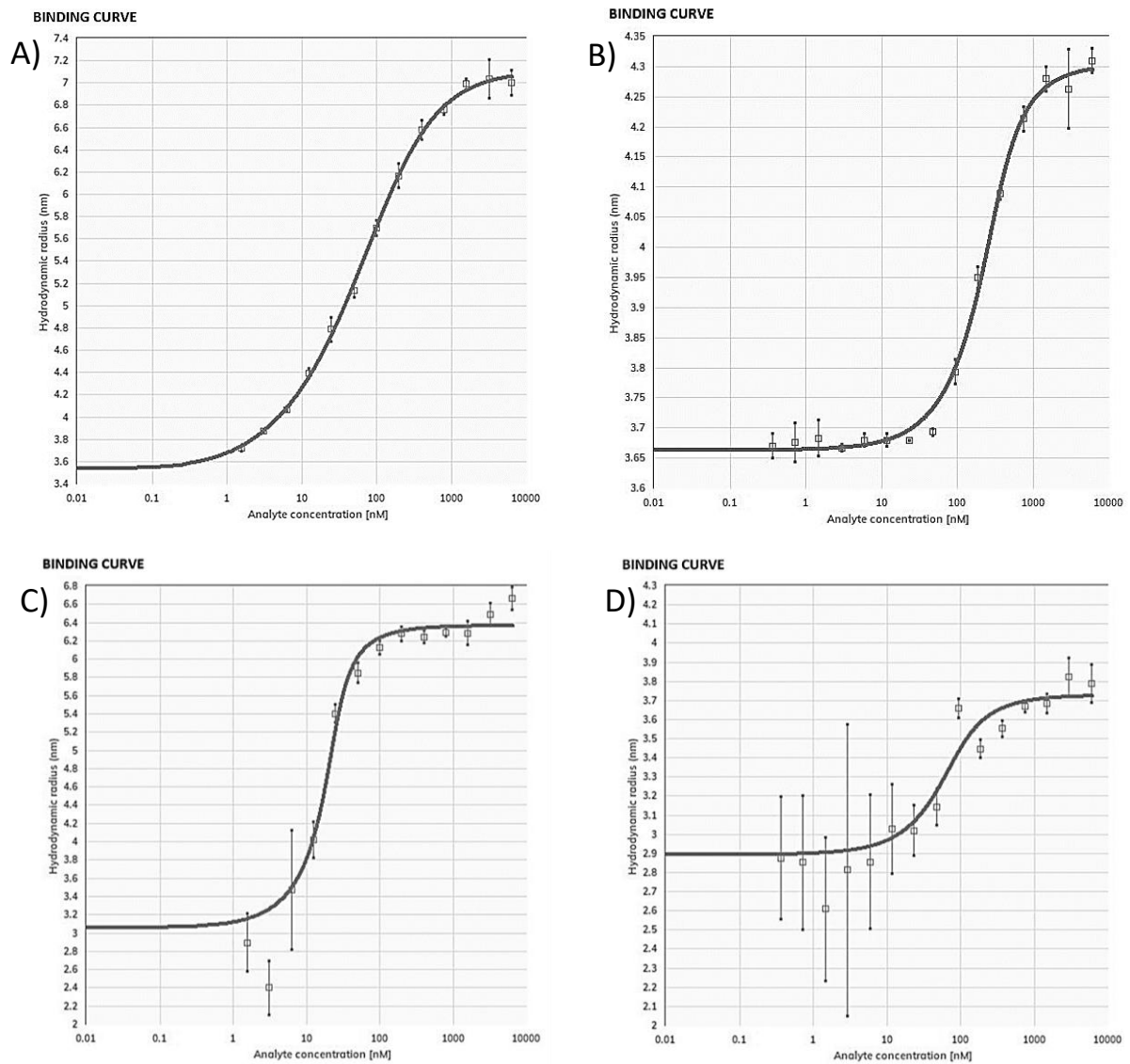


Figure S13. Binding curves obtained for the performed FIDA measurements: **A)** Results for the interactions between pICln-Alexa488 and PRMT5-MEP50. **B)** Results for the interactions between pICln-Alexa488 and TIM-MEP50. **C)** Results for the interactions between RioK1-Alexa488 and PRMT5-MEP50. **D)** Results for the interactions between RioK1-Alexa488 and TIM-MEP50.

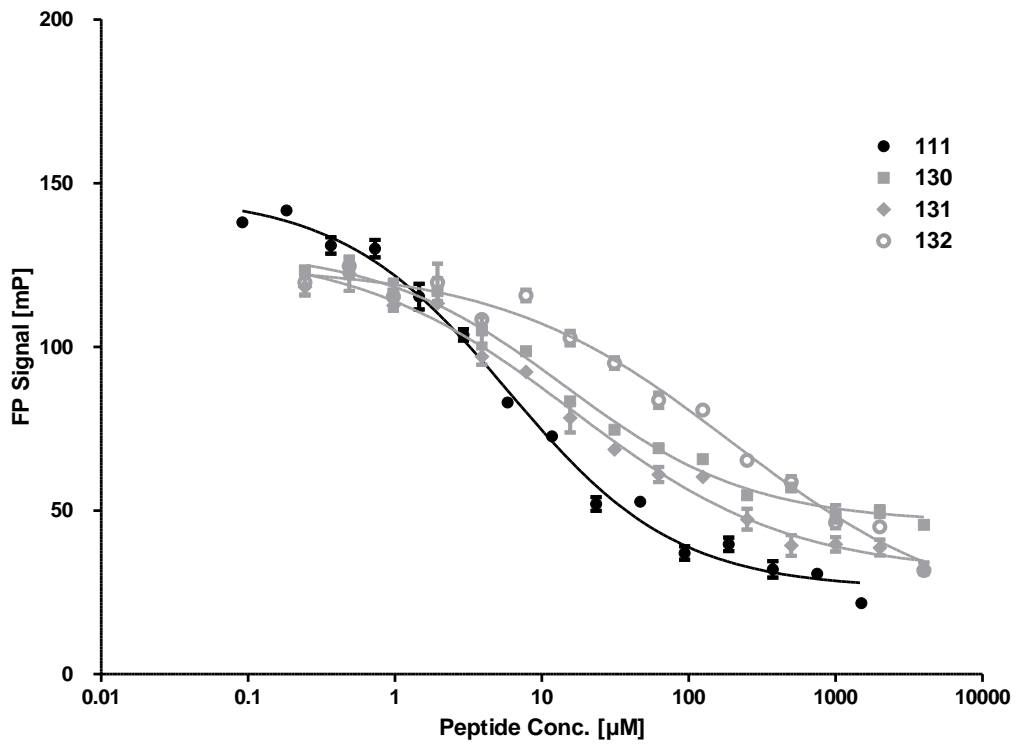


Figure S14. Competitive binding FP with PRMT5-MEP50 and fluorescently labelled compound **93** used as a tracer.

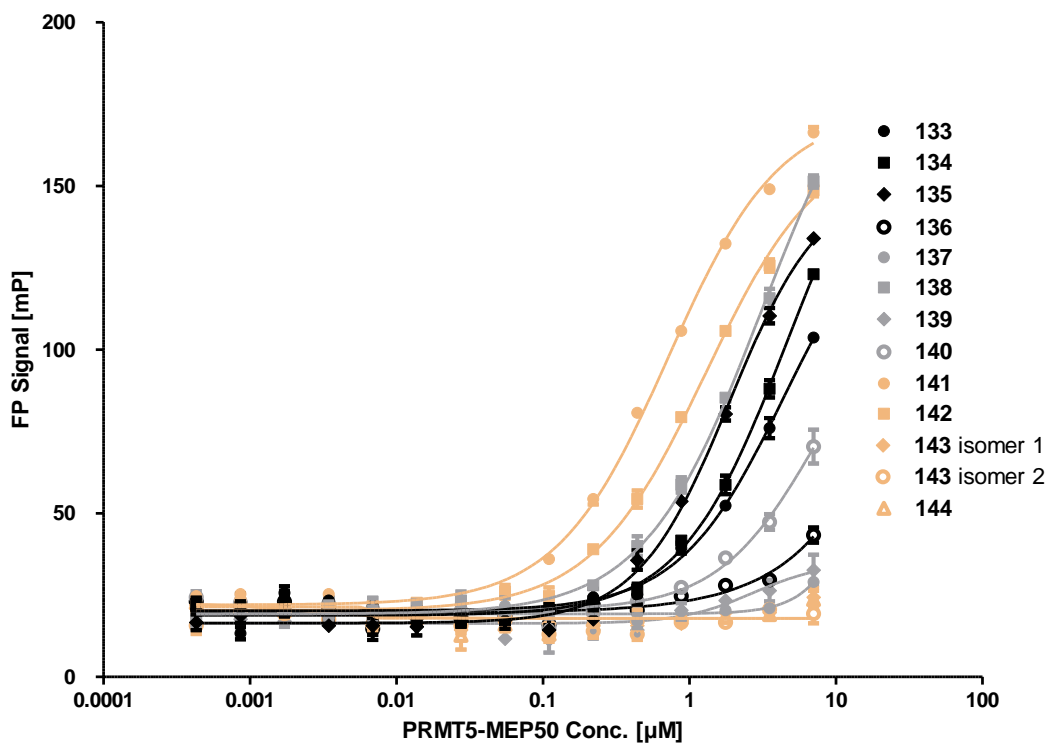


Figure S15. Direct binding FP with PRMT5-MEP50.

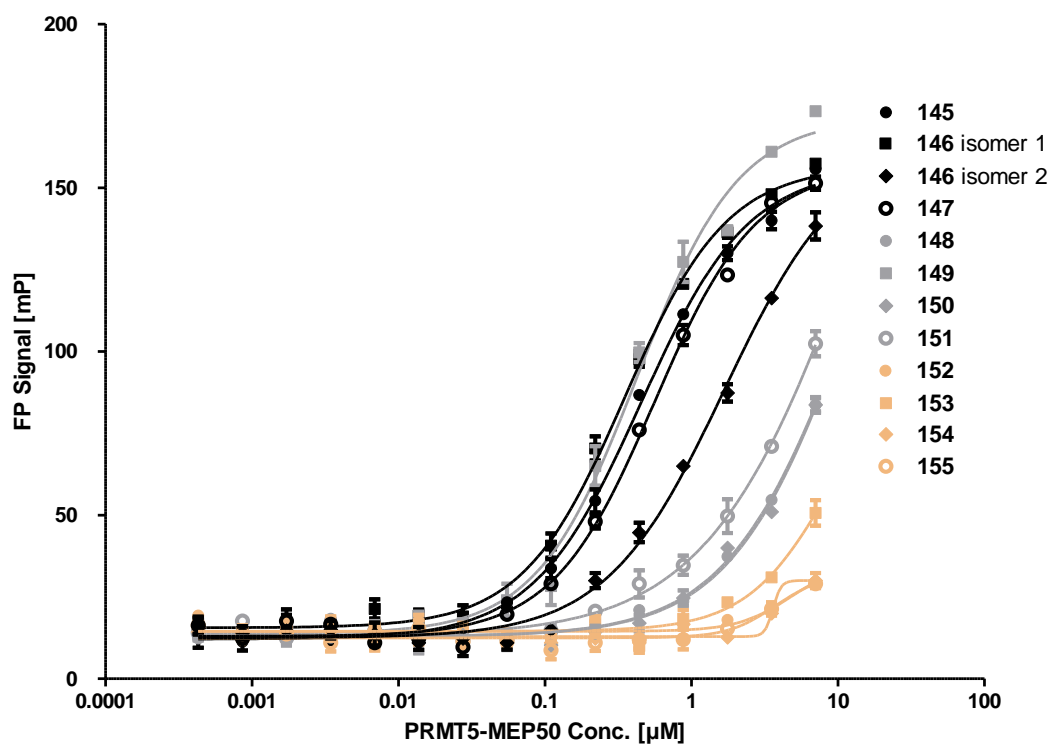


Figure S16. Direct binding FP with PRMT5-MEP50.

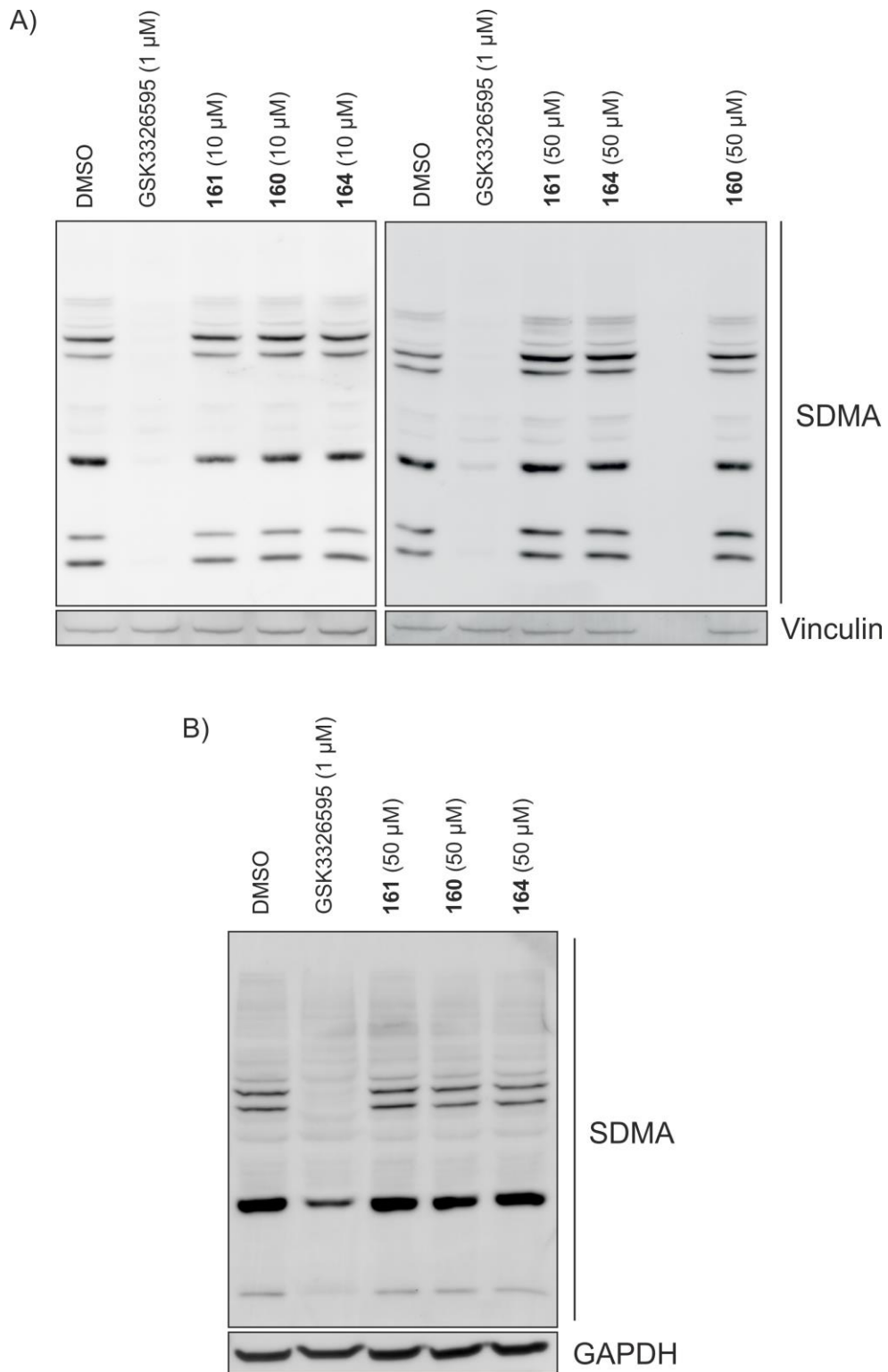


Figure17. Immunoblotting visualising SDMA proteins in treated cells: **A)** Results of treatment of HEK293-T cells. **B)** Results of treatment of MCF7 cells.

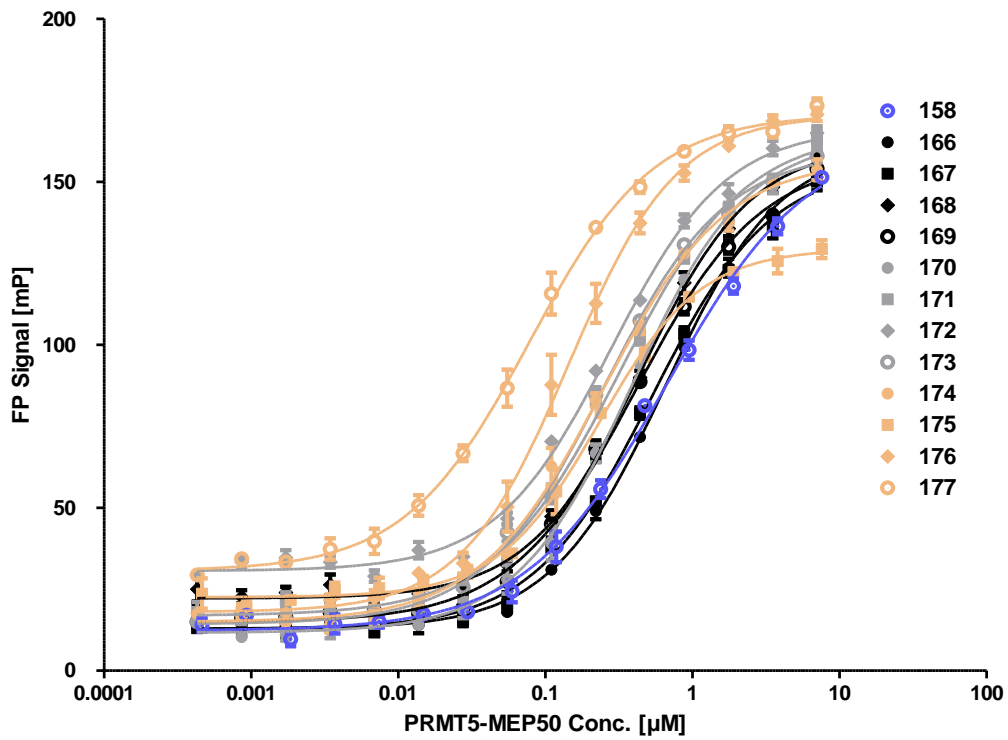


Figure S18. Direct binding FP with PRMT5-MEP50.

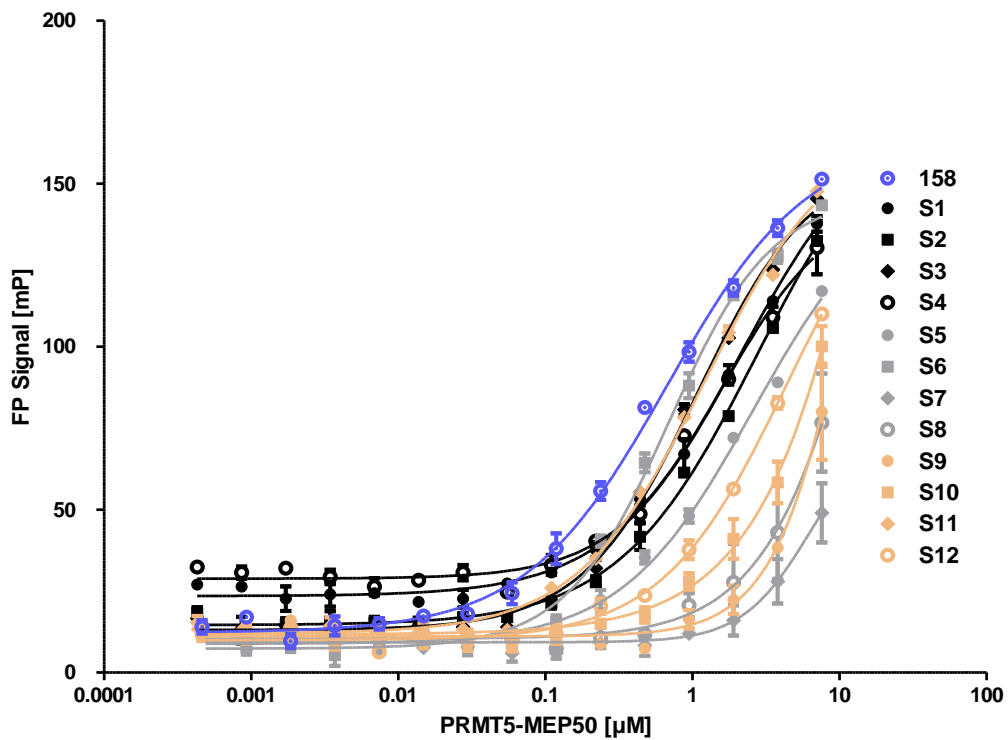


Figure S19. Direct binding FP with PRMT5-MEP50.

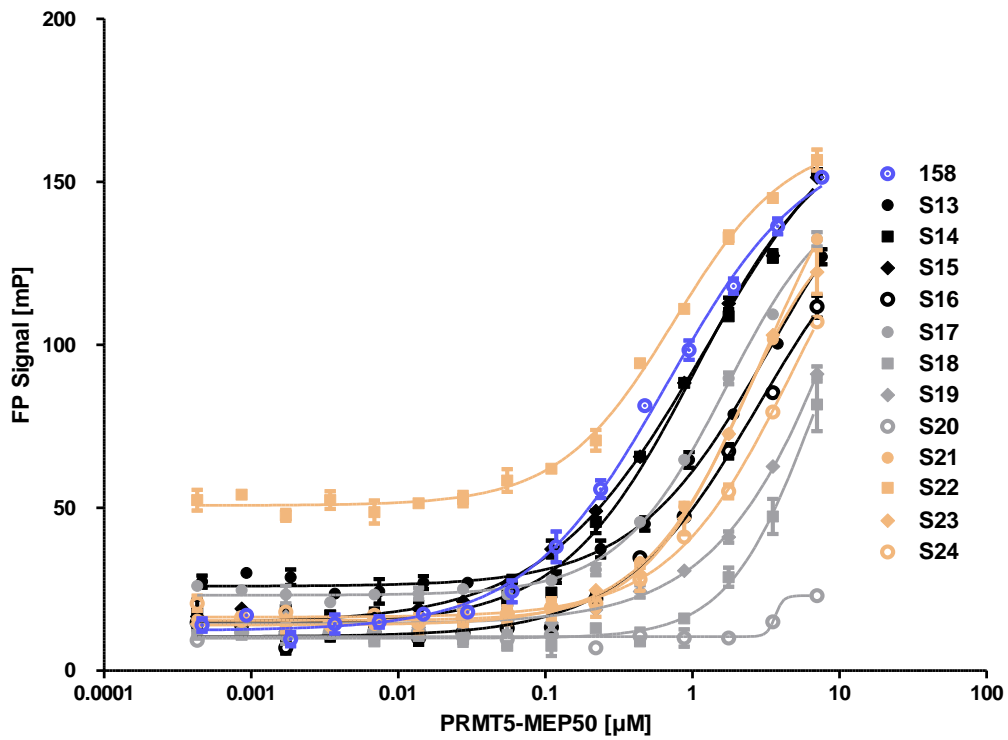


Figure S20. Direct binding FP with PRMT5-MEP50.

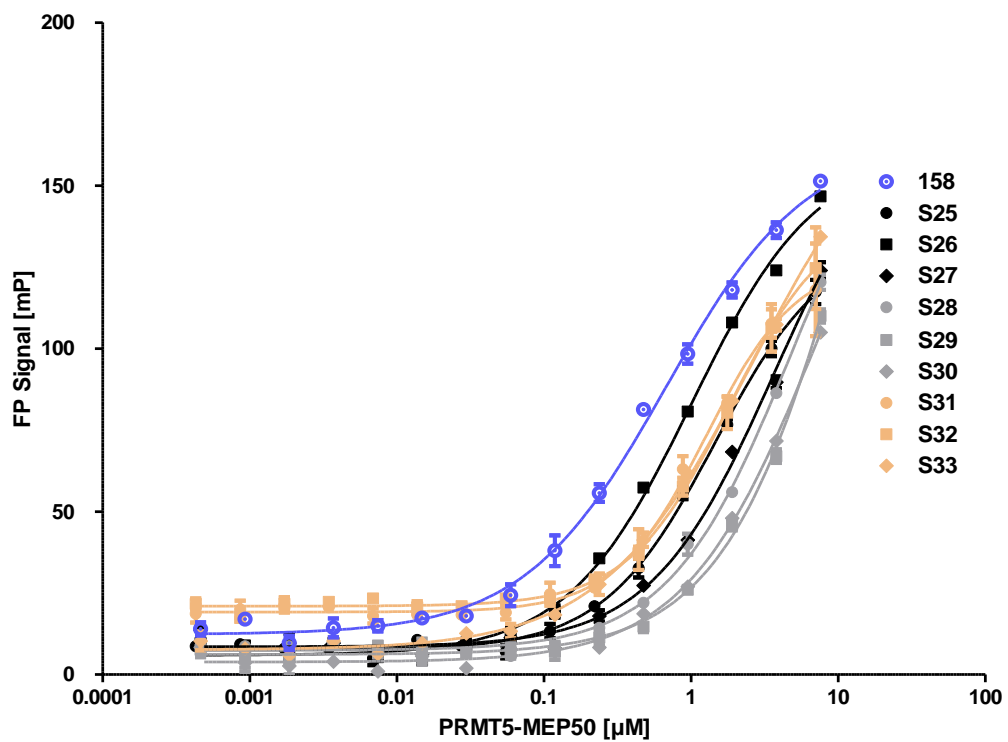


Figure S21. Direct binding FP with PRMT5-MEP50.

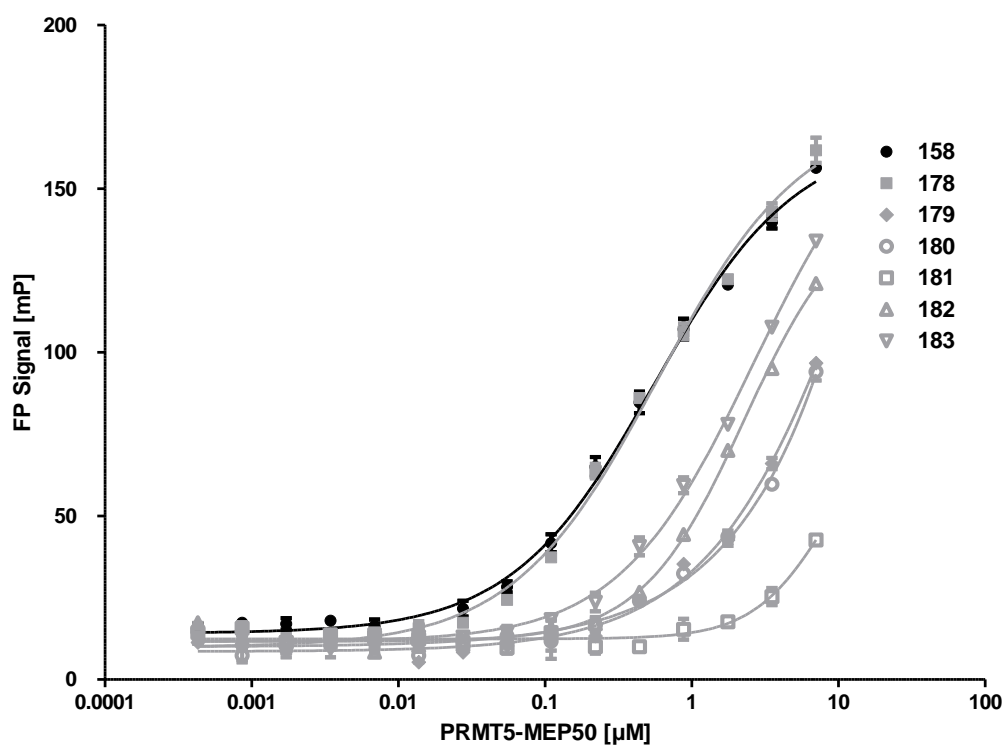


Figure S22. Direct binding FP with PRMT5-MEP50.

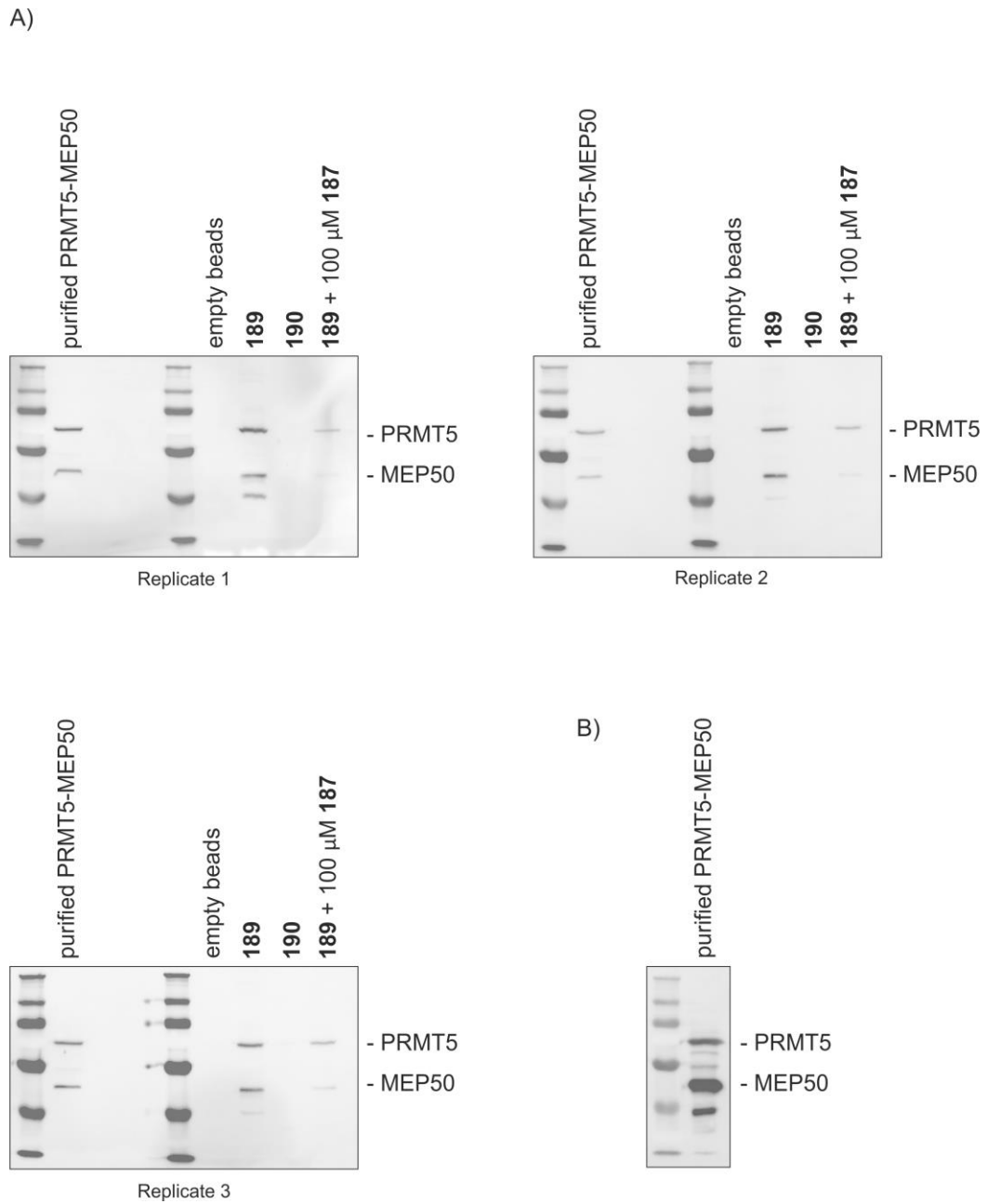


Figure 23. Results of pull-down experiments: **A)** Pull-down from MCF7 cellular lysates. **B)** Western blot of the purified PRMT5-MEP50 at a higher concentration than used in A) for comparison

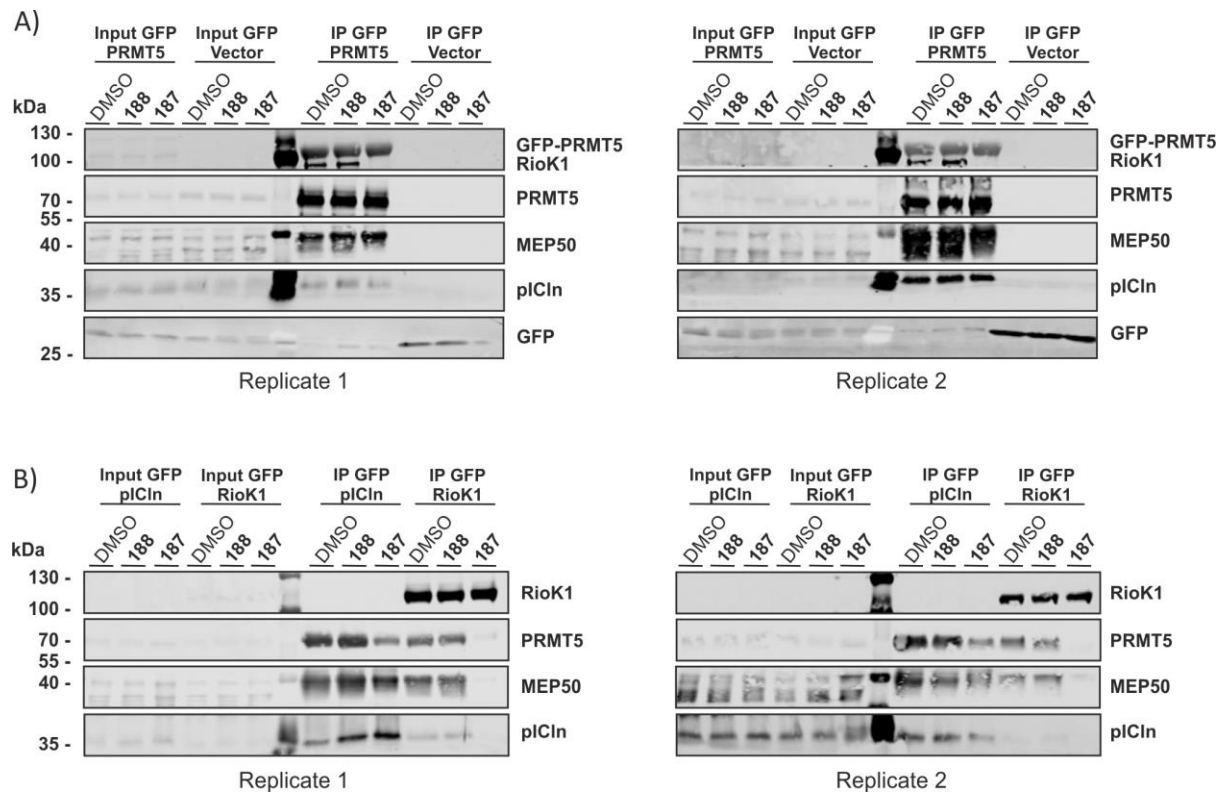
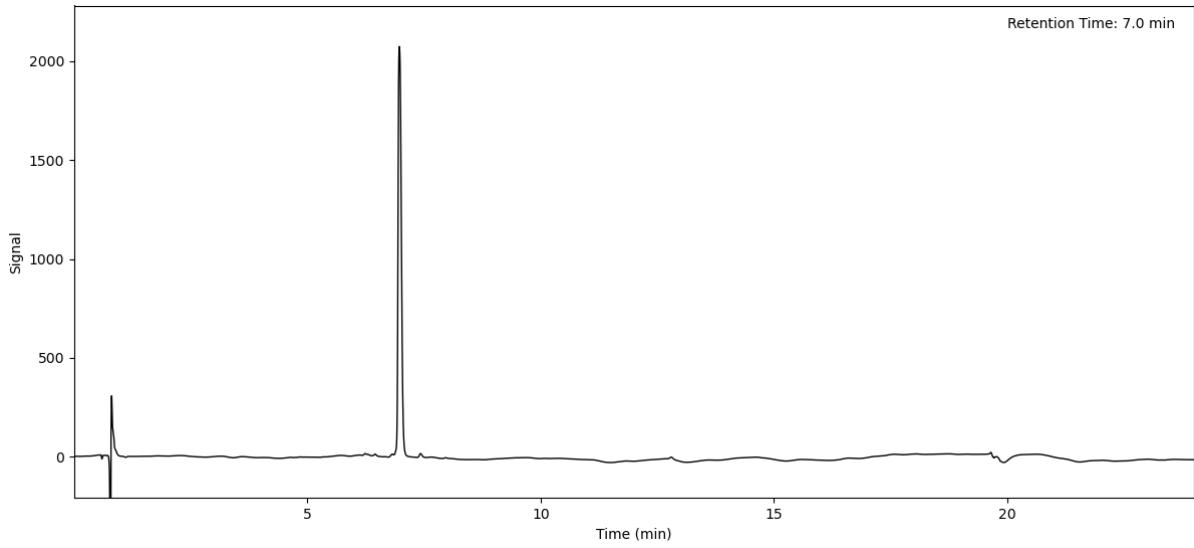


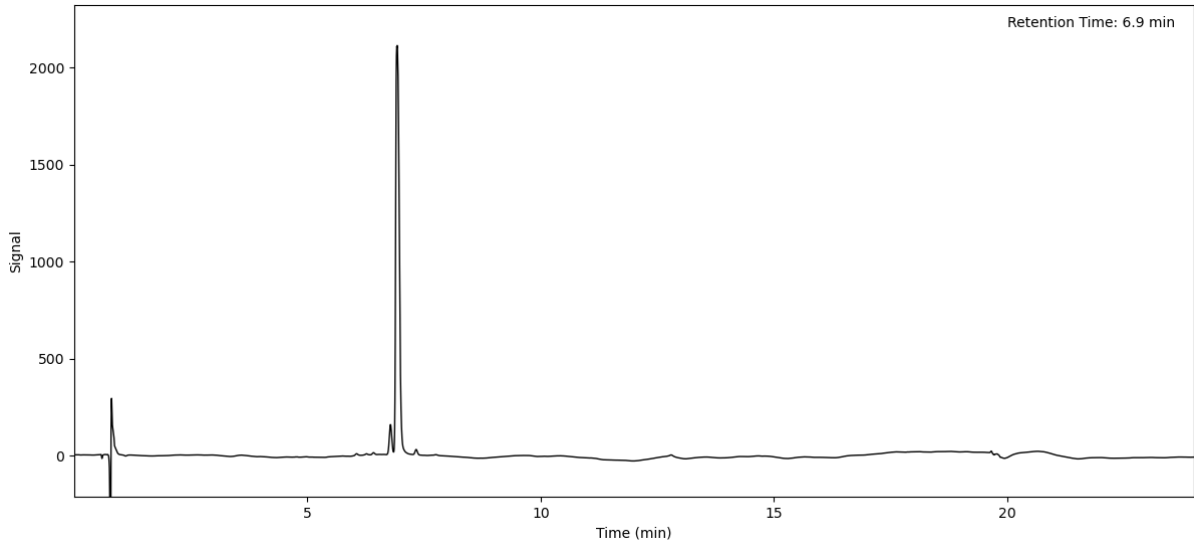
Figure S24. IP assays using cell lysates: **A)** GFP-IP from lysate of Flp-In T-Rex 293-GFP and Flp-In T-Rex 293-GFP-PRMT5 overexpressing cells, evaluating compound **187** and scrambled **188** at 50 μ M. **B)** GFP-IP from Flp-In T-Rex 293-GFP-pICln and Flp-In T-Rex 293-GFP-RioK1 cytoplasmic extract, testing **187** and **188** at 50 μ M.

8.1.3. Peptide Purity HPLC Chromatograms

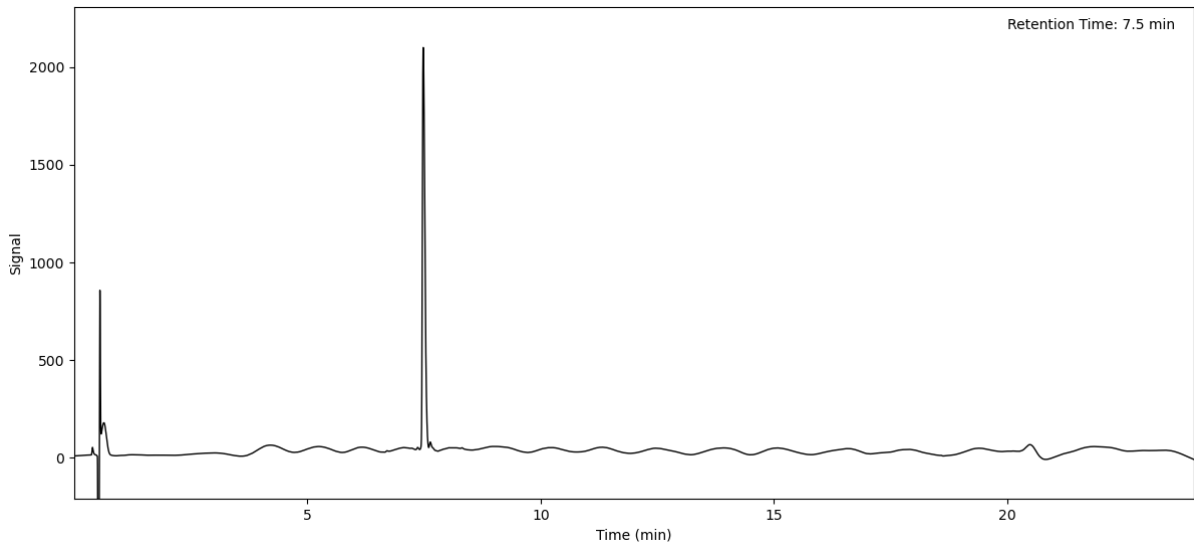
Peptide 1 (Method B)



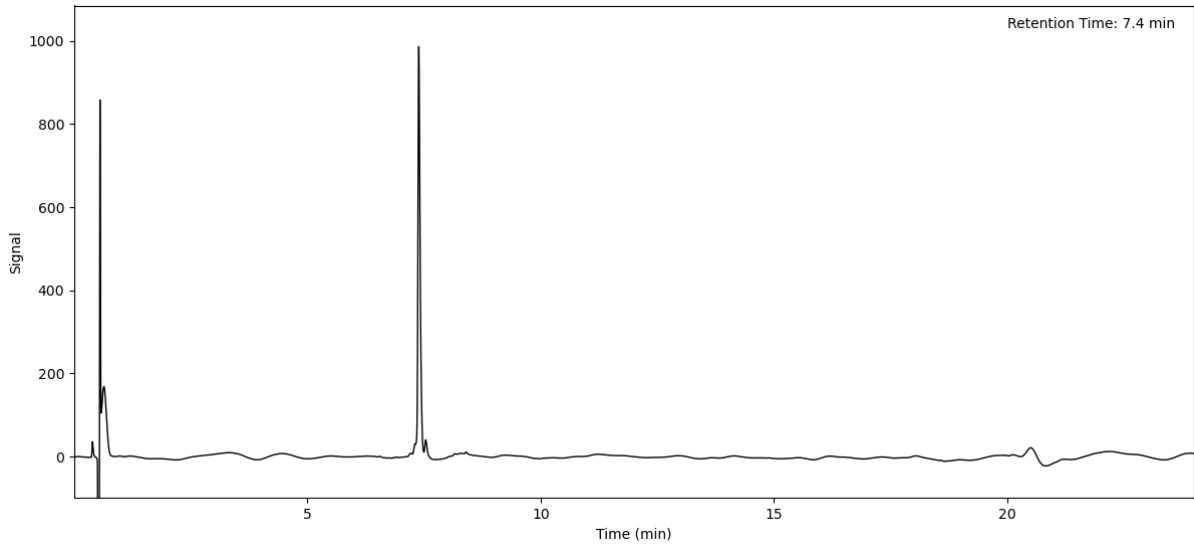
Peptide 2 (Method B)



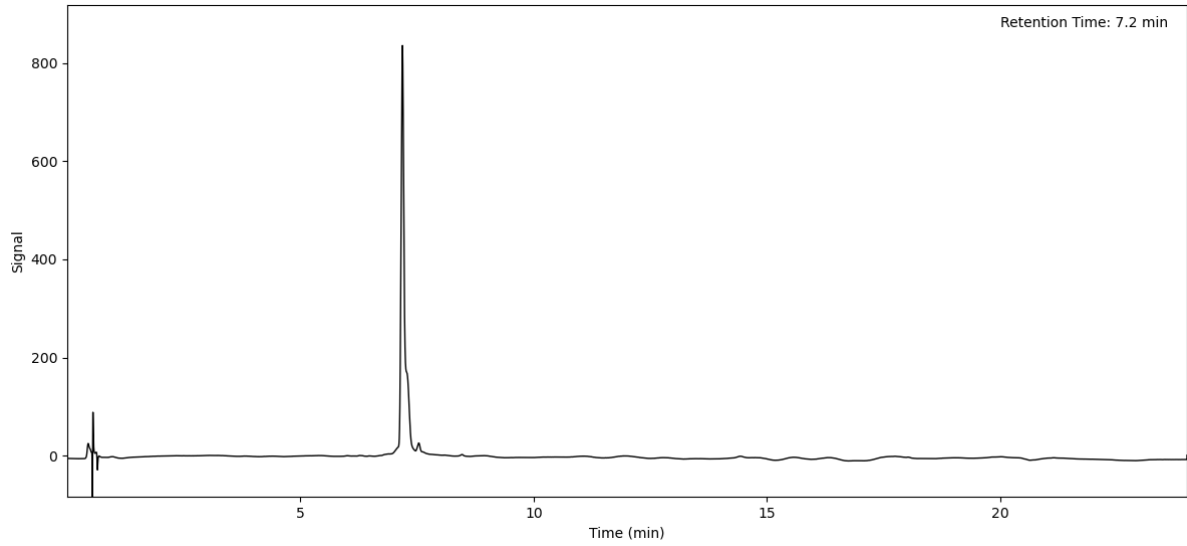
Peptide 3 (Method A)



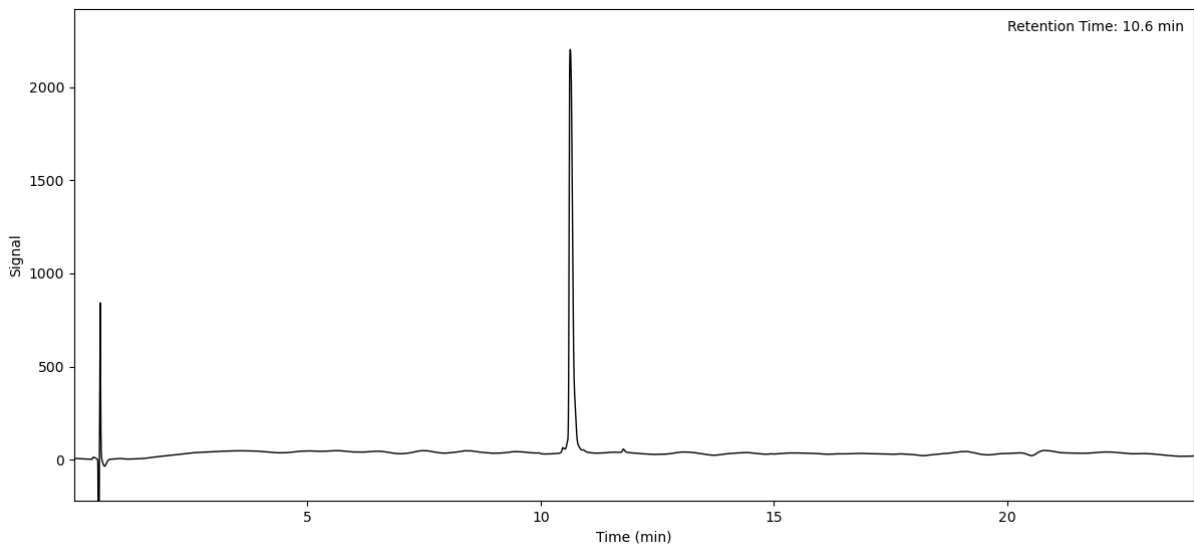
Peptide 4 (Method A)



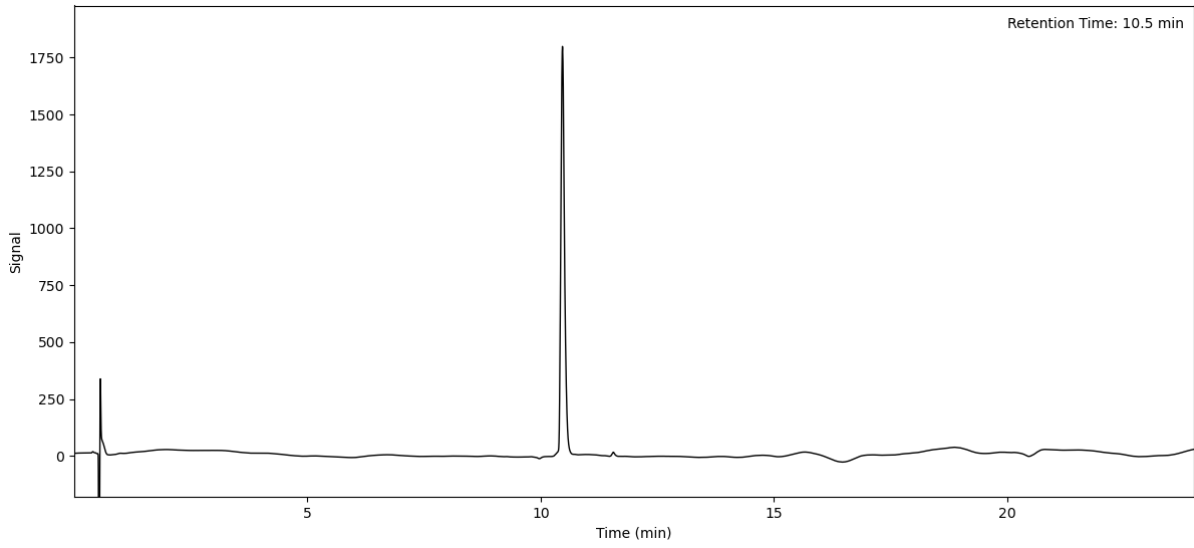
Peptide 5 (Method A)



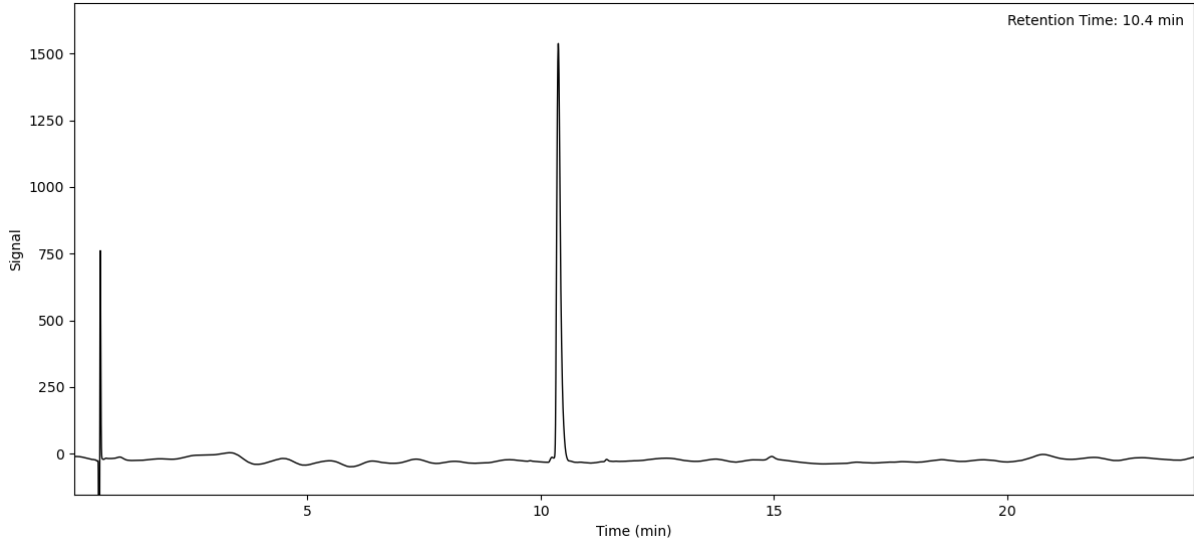
Peptide 6 (Method A)



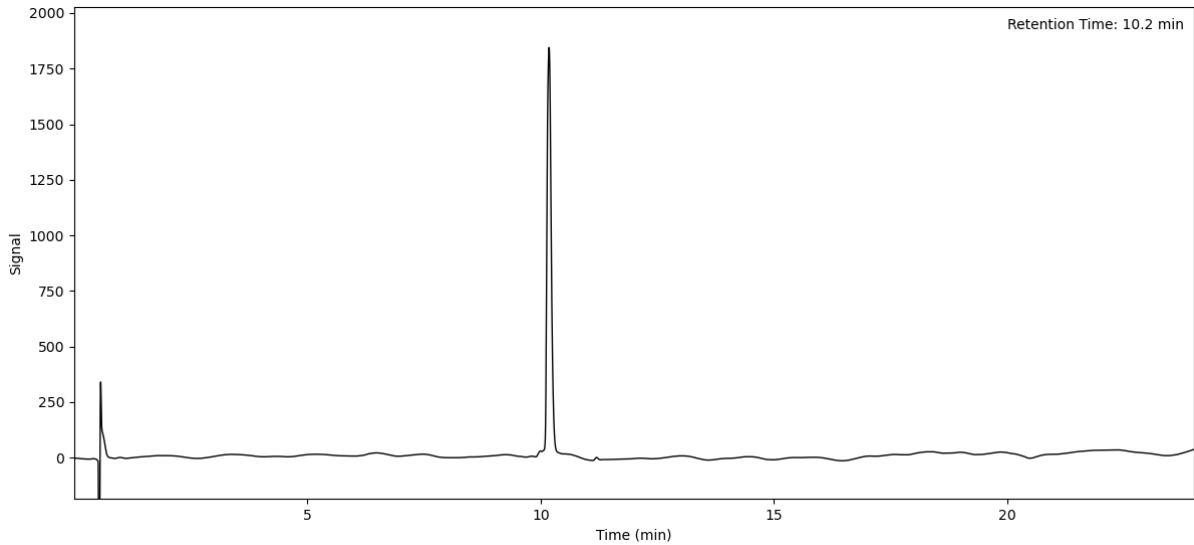
Peptide 7 (Method A)



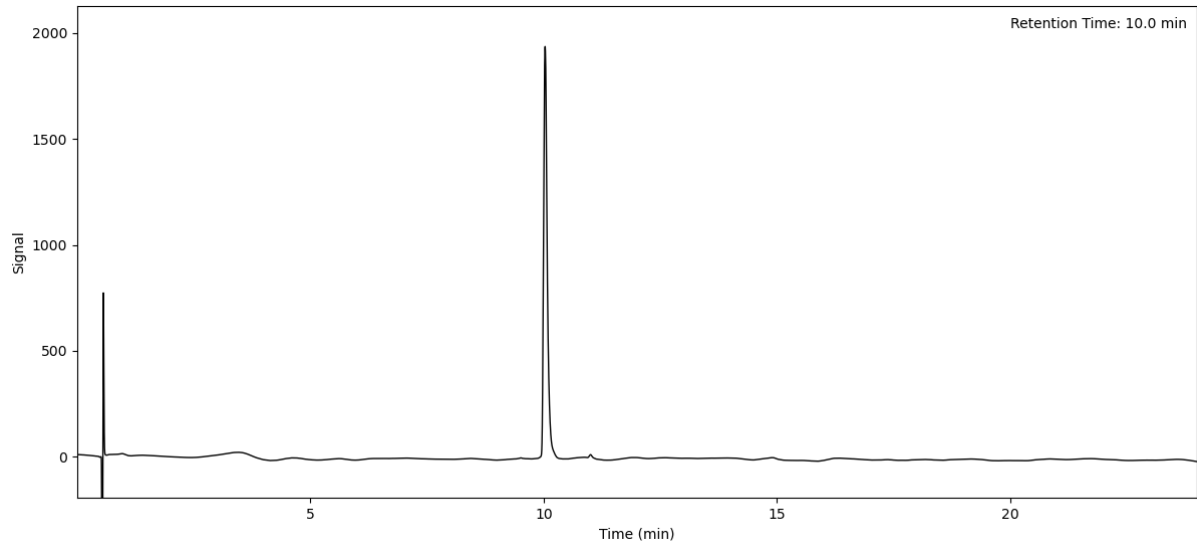
Peptide 8 (Method A)



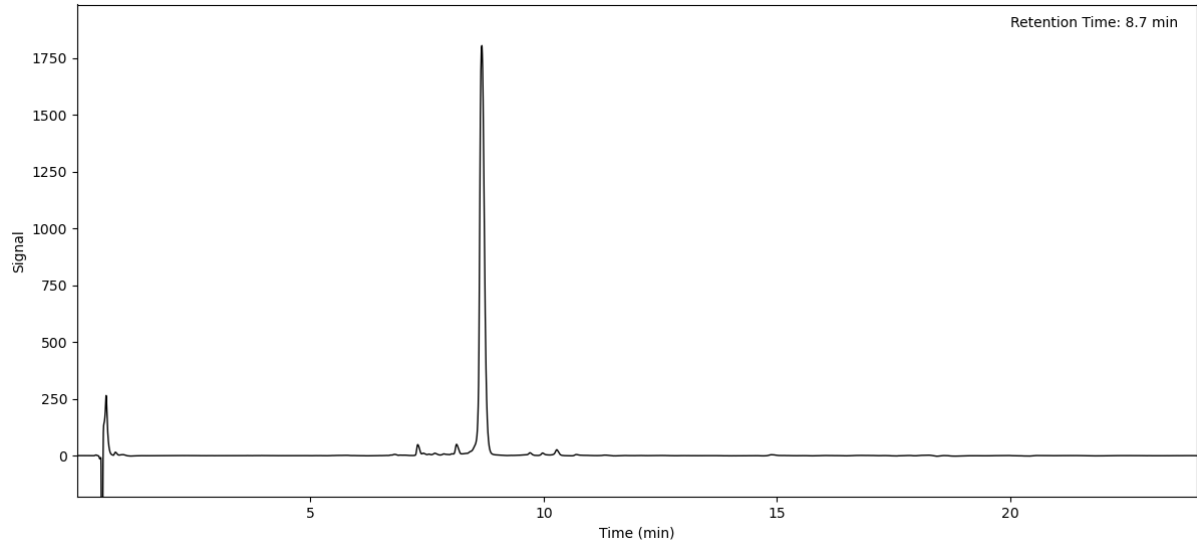
Peptide 9 (Method A)



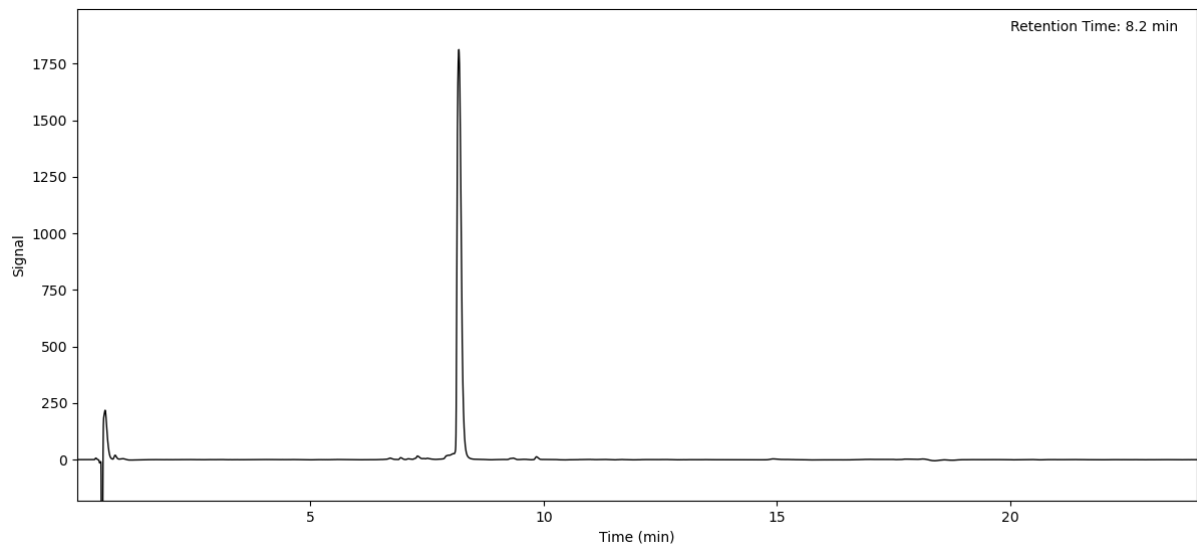
Peptide 10 (Method A)



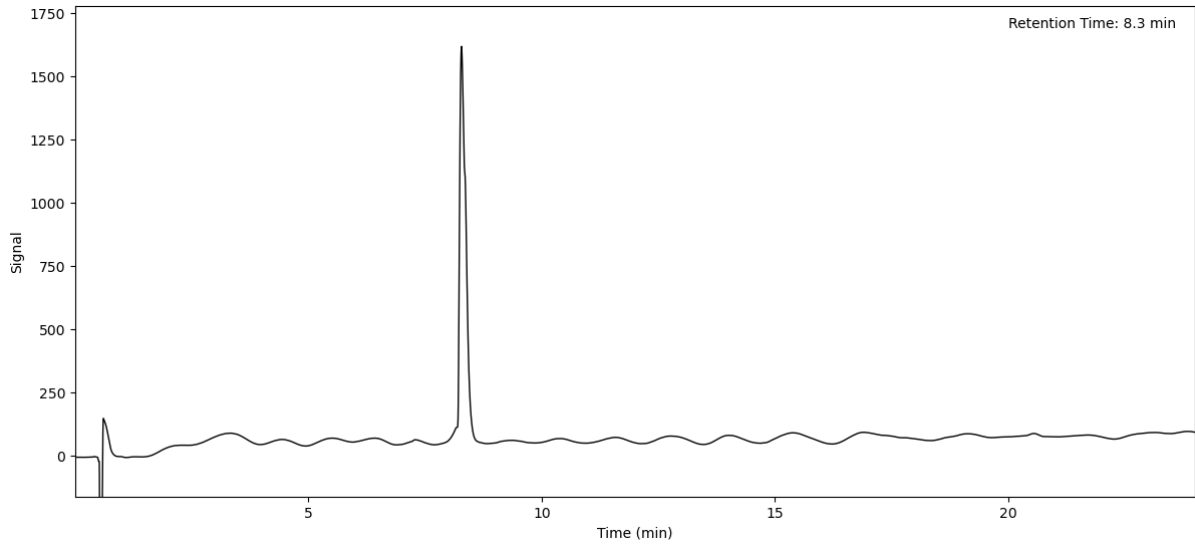
Peptide 11 (Method A)



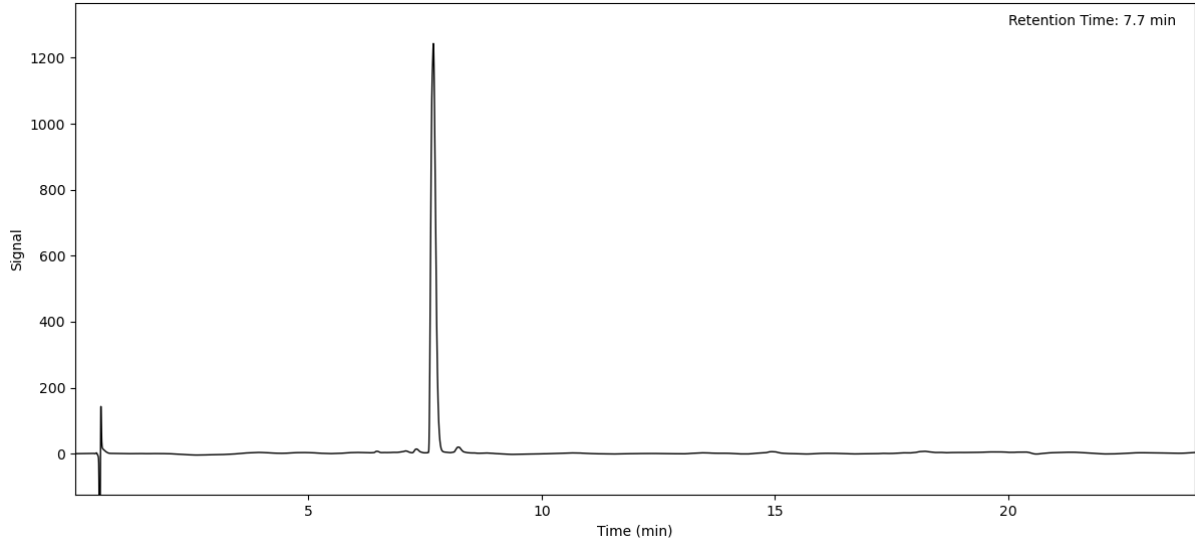
Peptide 12 (Method A)



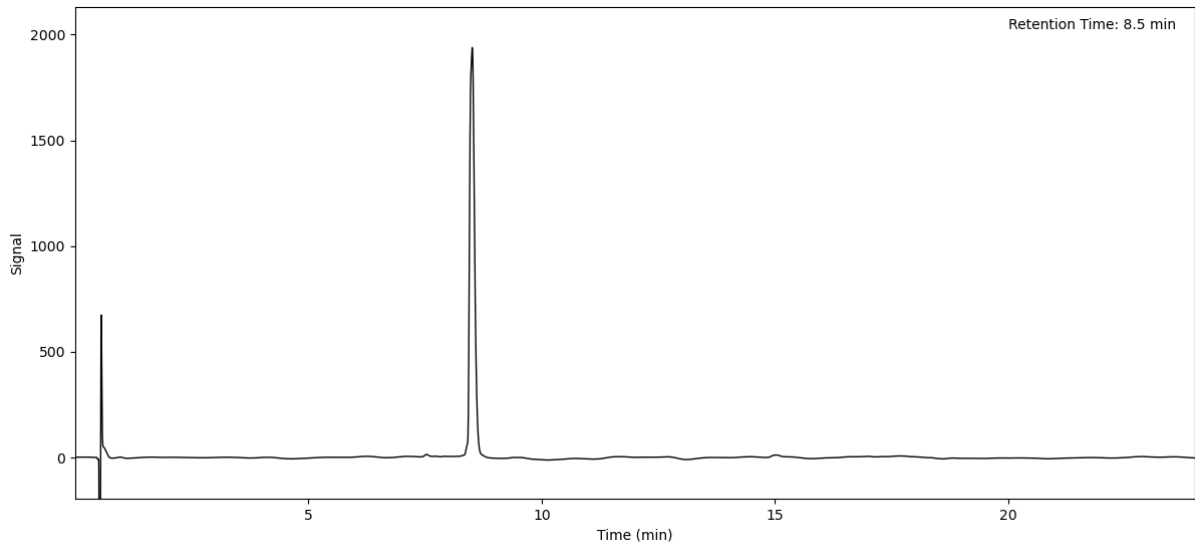
Peptide 13 (Method A)



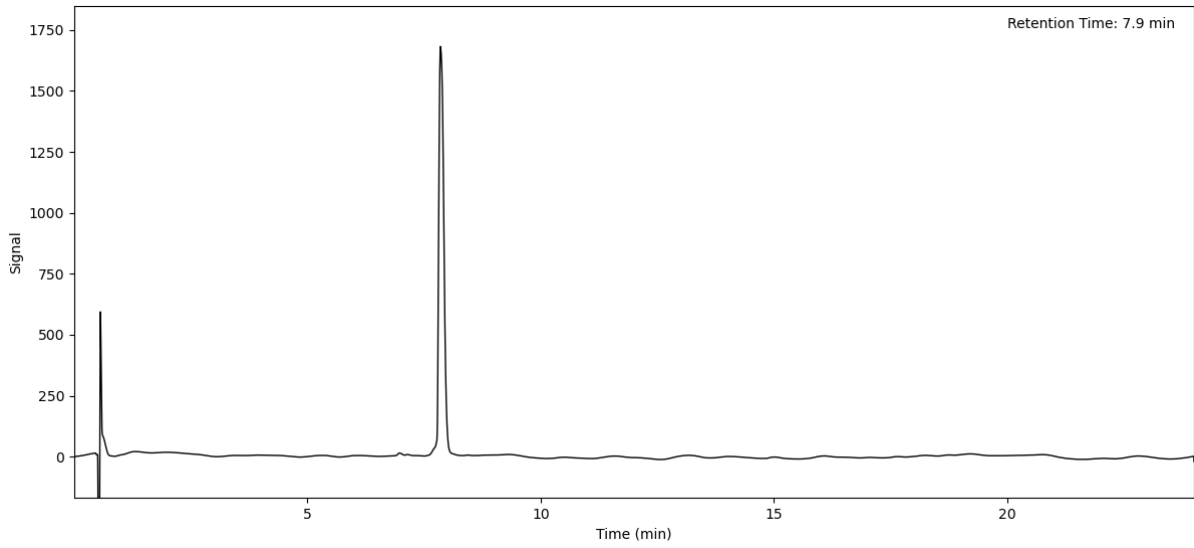
Peptide 14 (Method A)



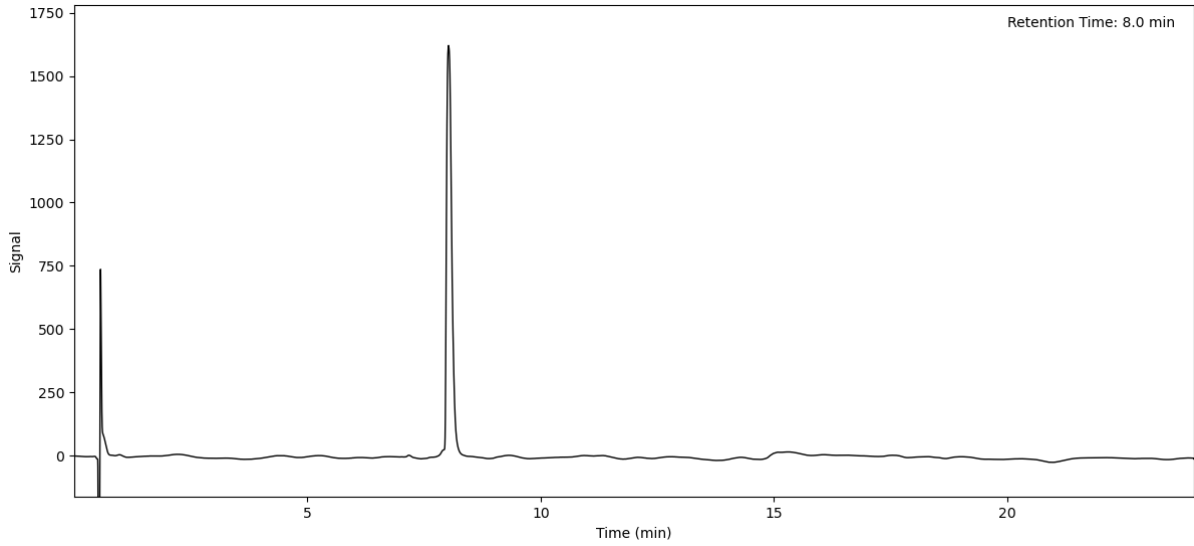
Peptide 15 (Method A)



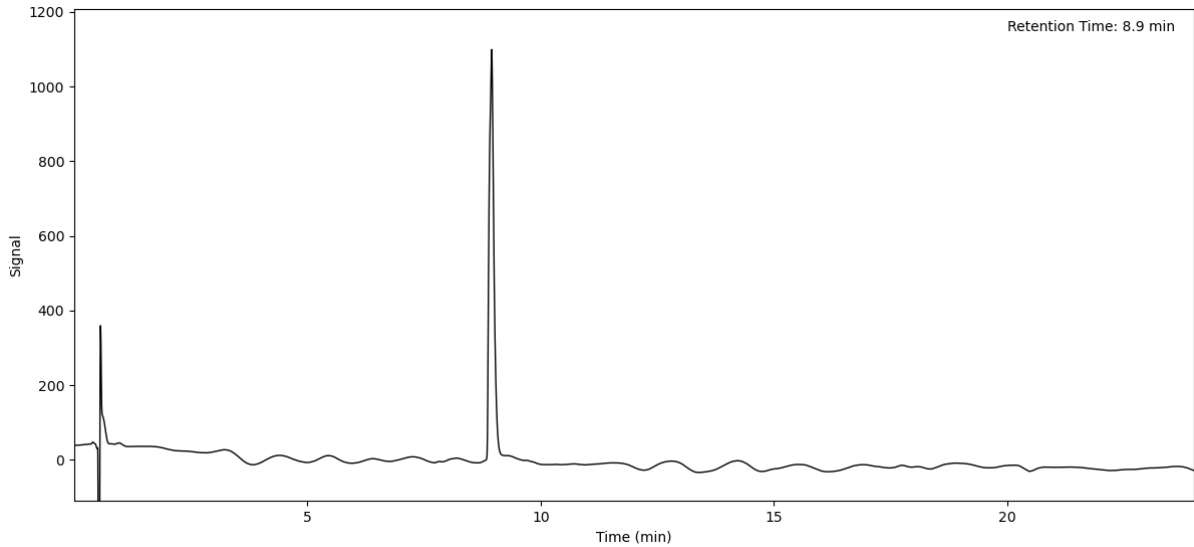
Peptide 16 (Method A)



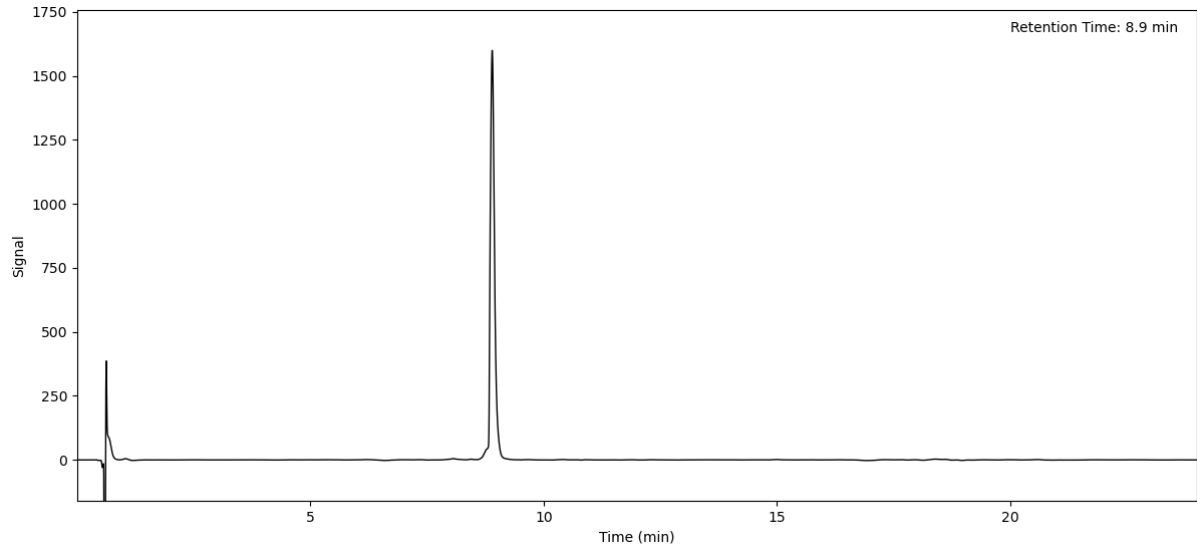
Peptide 17 (Method A)



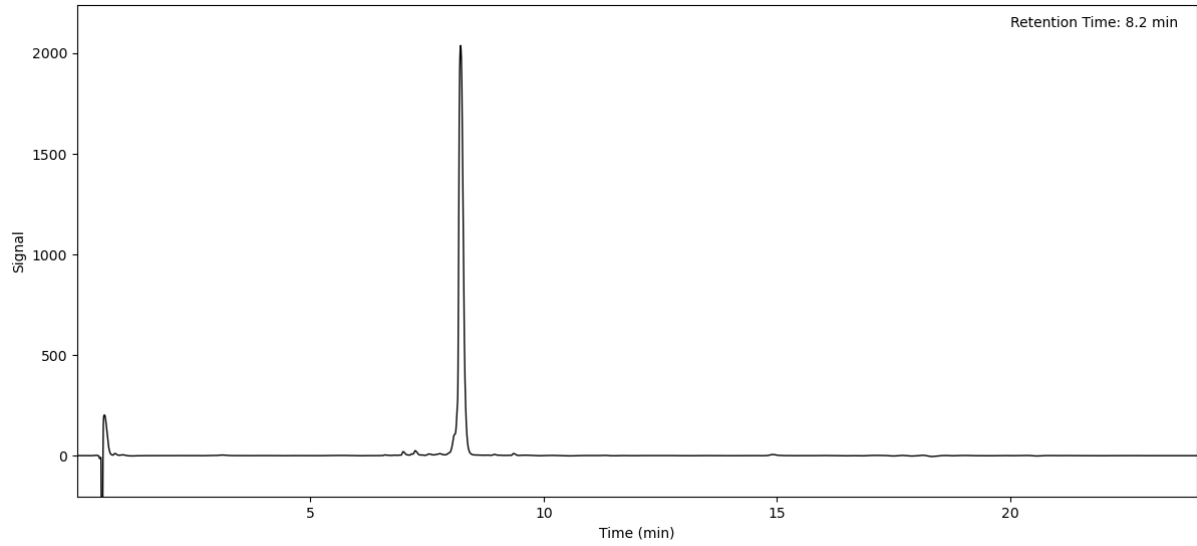
Peptide 18 (Method A)



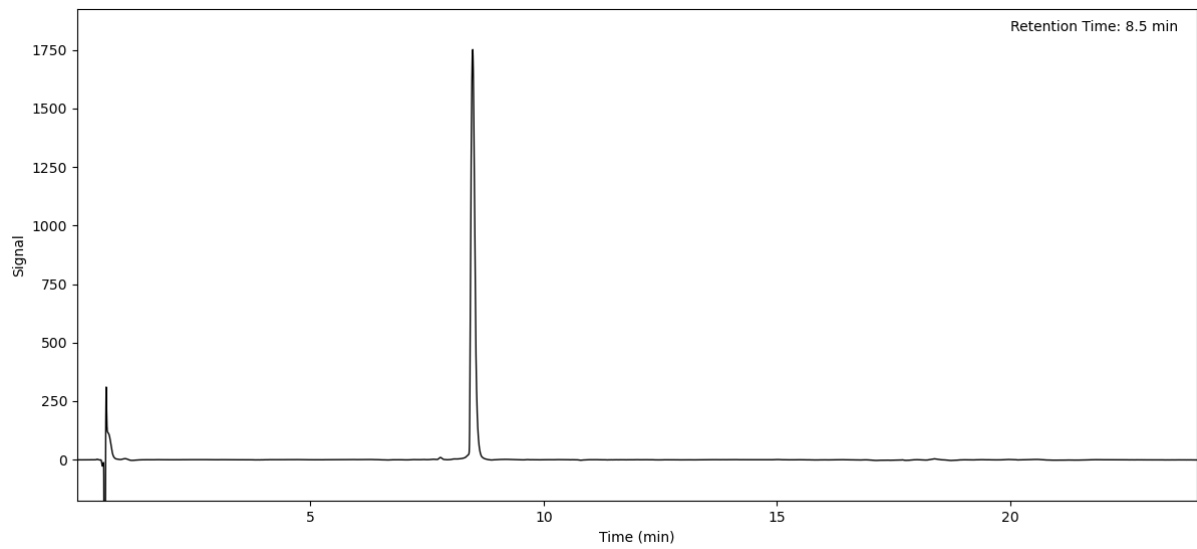
Peptide 19 (Method A)



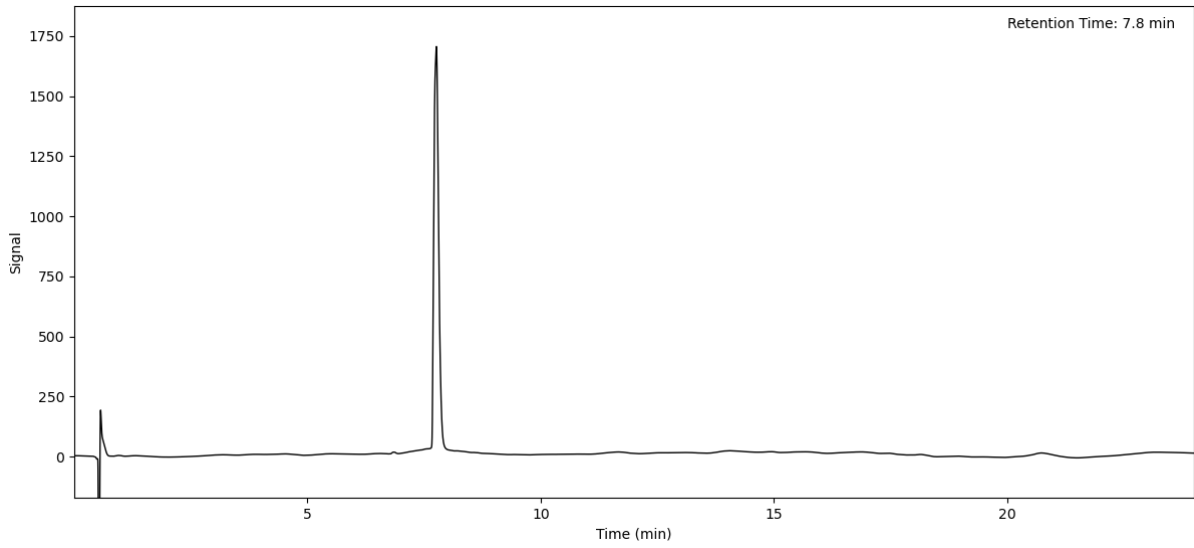
Peptide 20 (Method A)



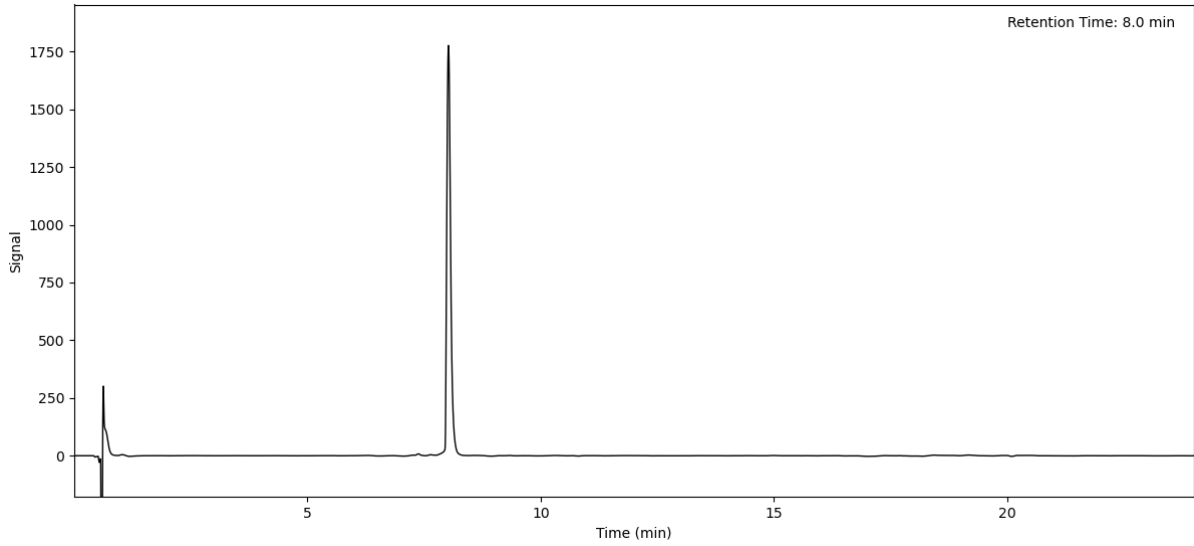
Peptide 21 (Method A)



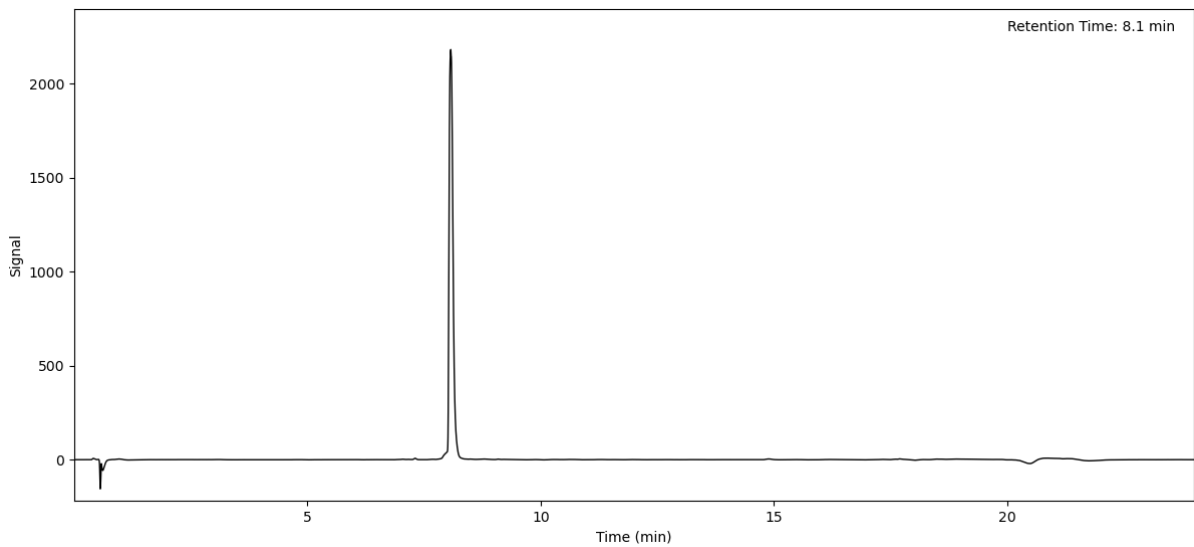
Peptide 22 (Method A)



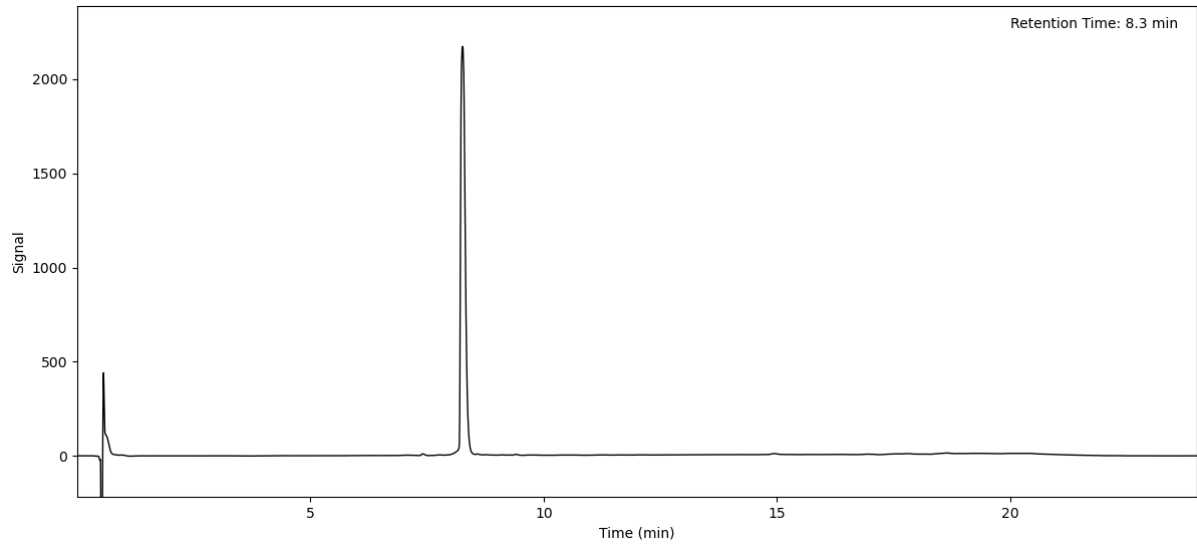
Peptide 23 (Method A)



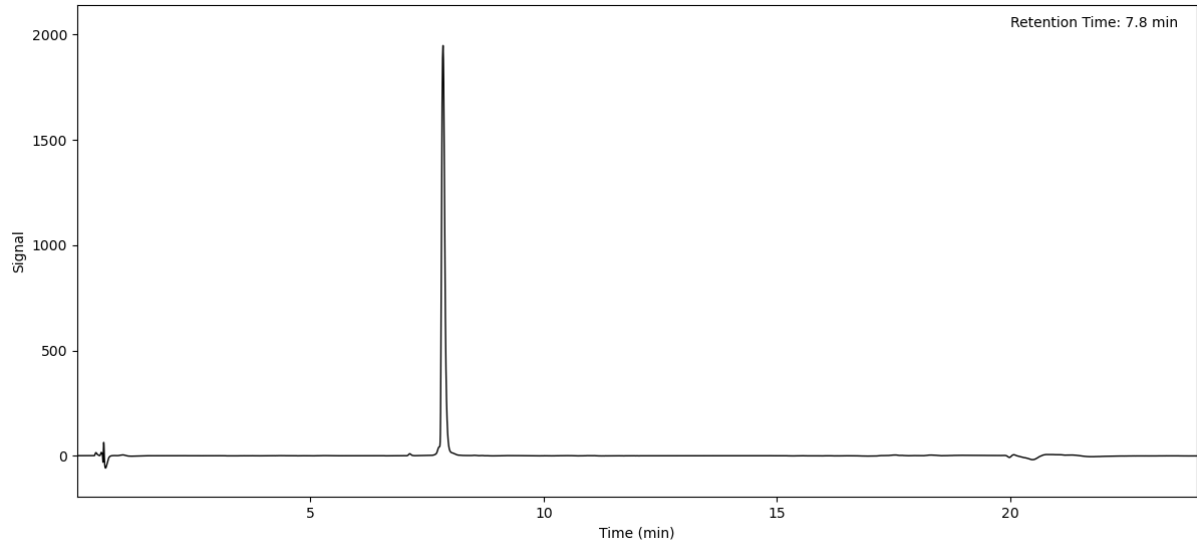
Peptide 24 (Method A)



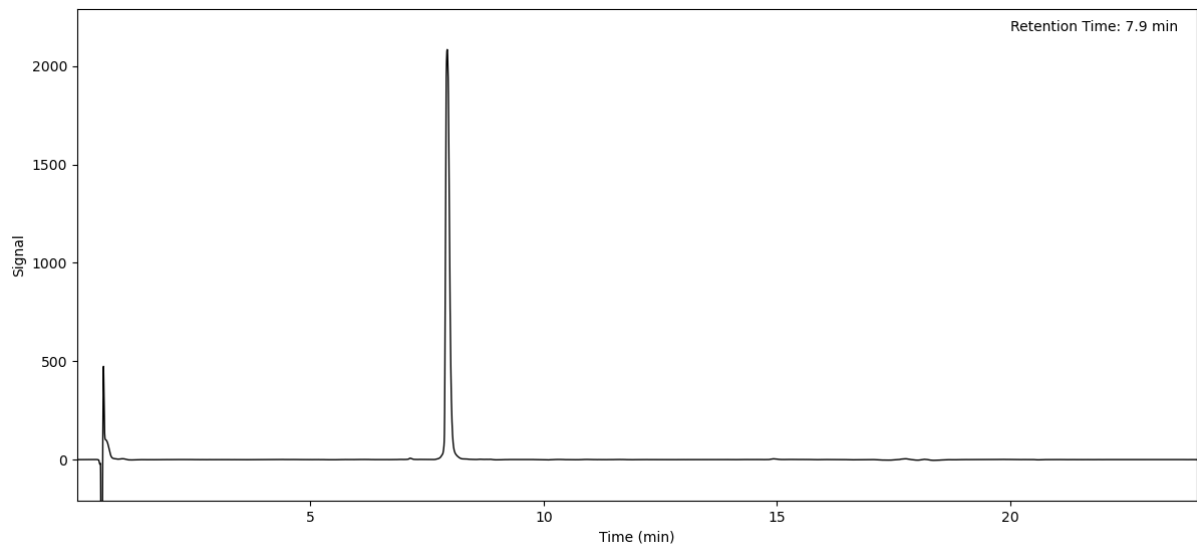
Peptide 25 (Method A)



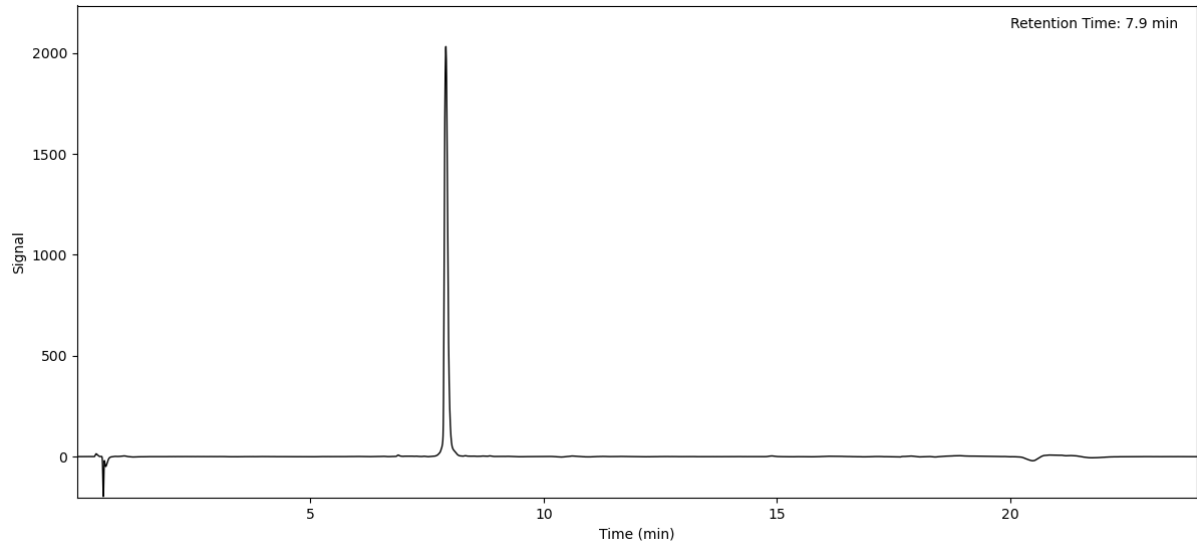
Peptide 26 (Method A)



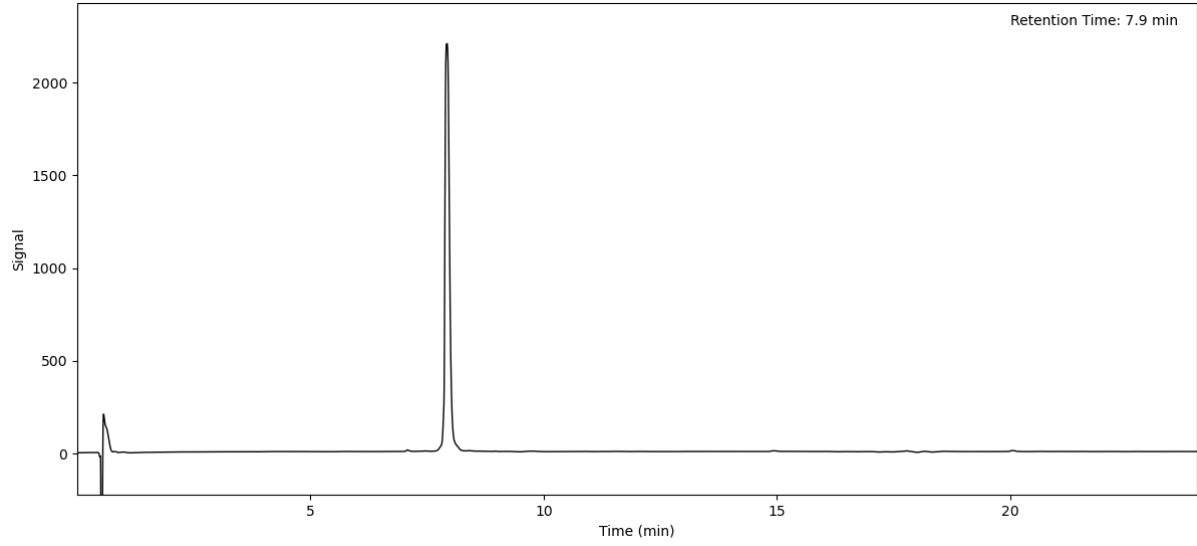
Peptide 27 (Method A)



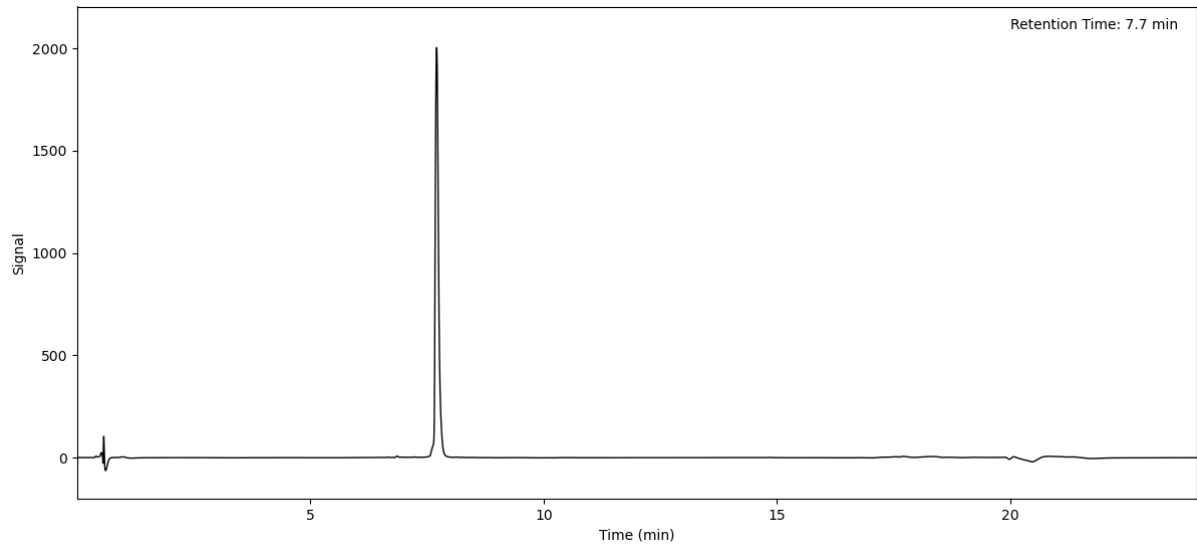
Peptide 28 (Method A)



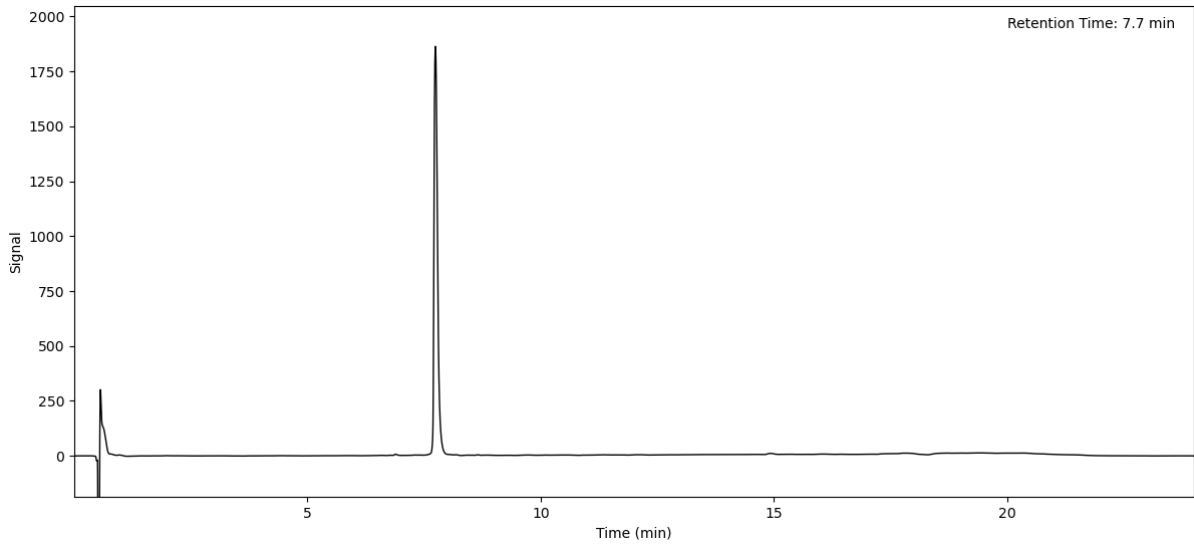
Peptide 29 (Method A)



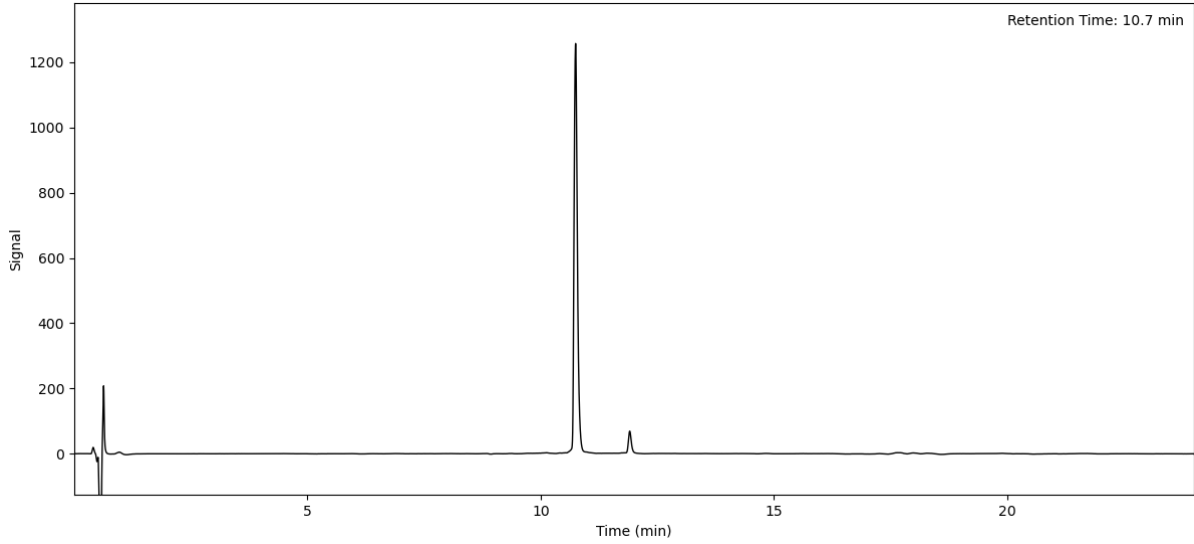
Peptide 30 (Method A)



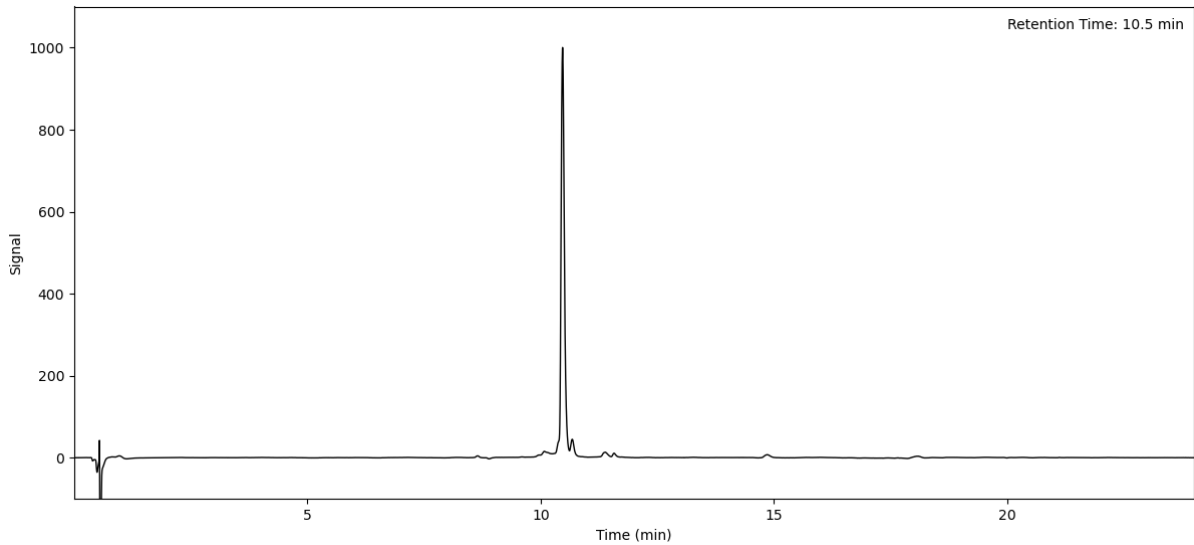
Peptide 31 (Method A)



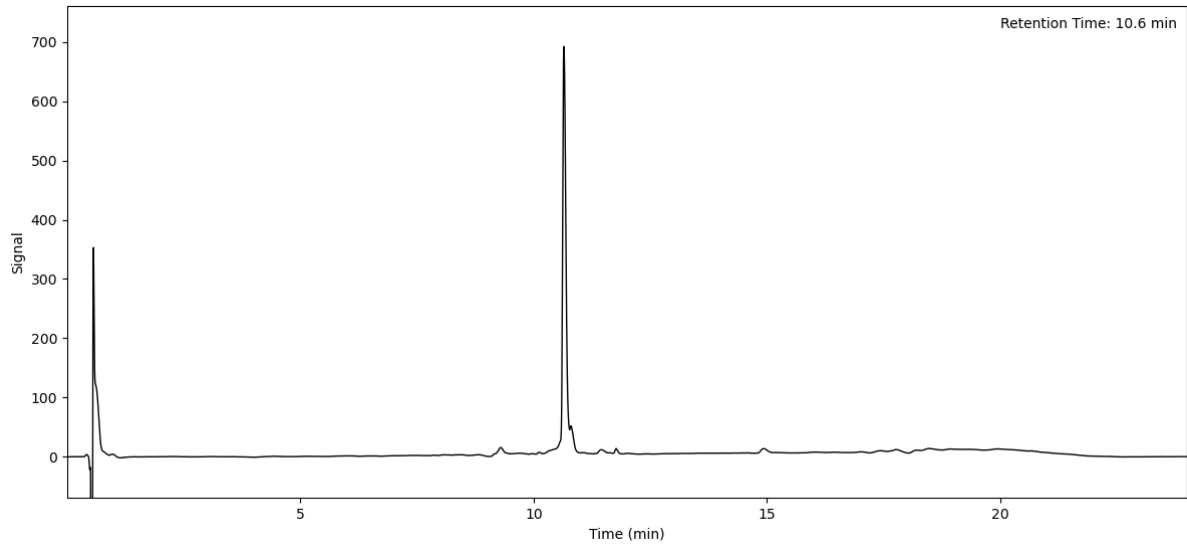
Peptide 32 (Method A)



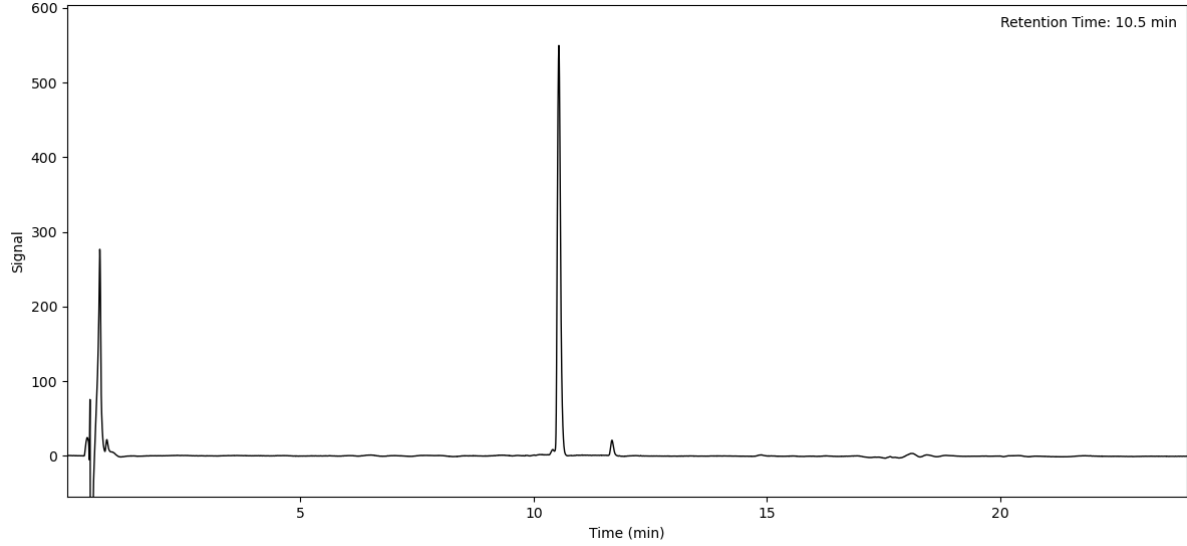
Peptide 33 (Method A)



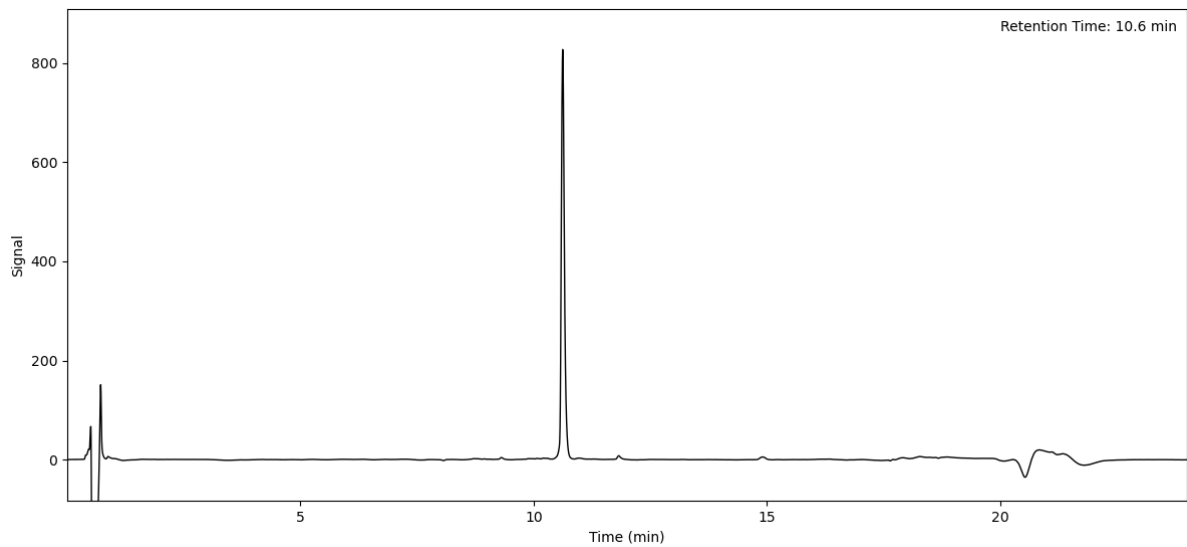
Peptide 34 (Method A)



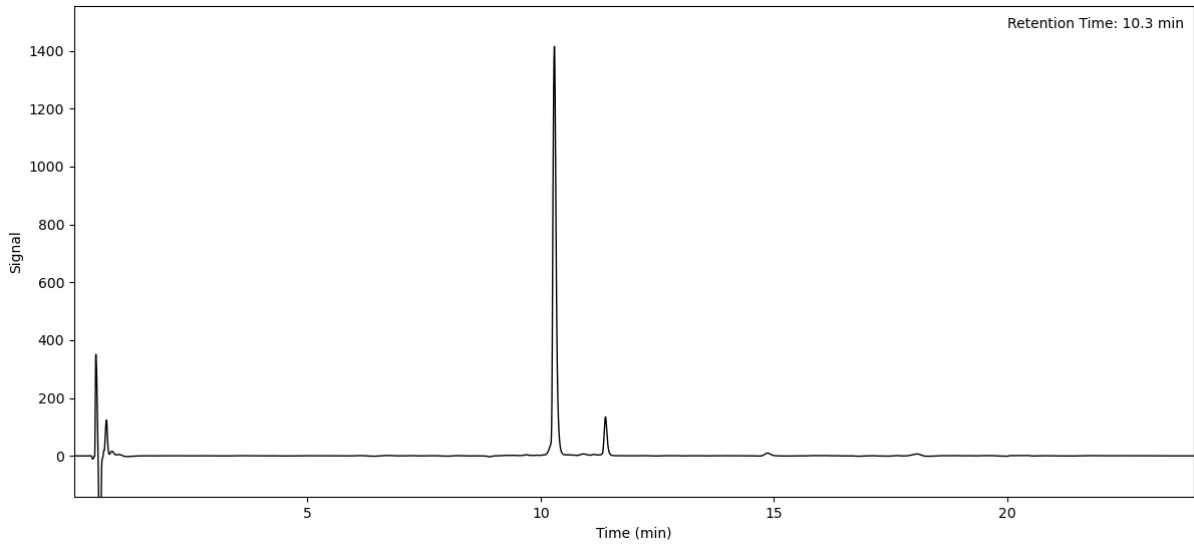
Peptide 35 (Method A)



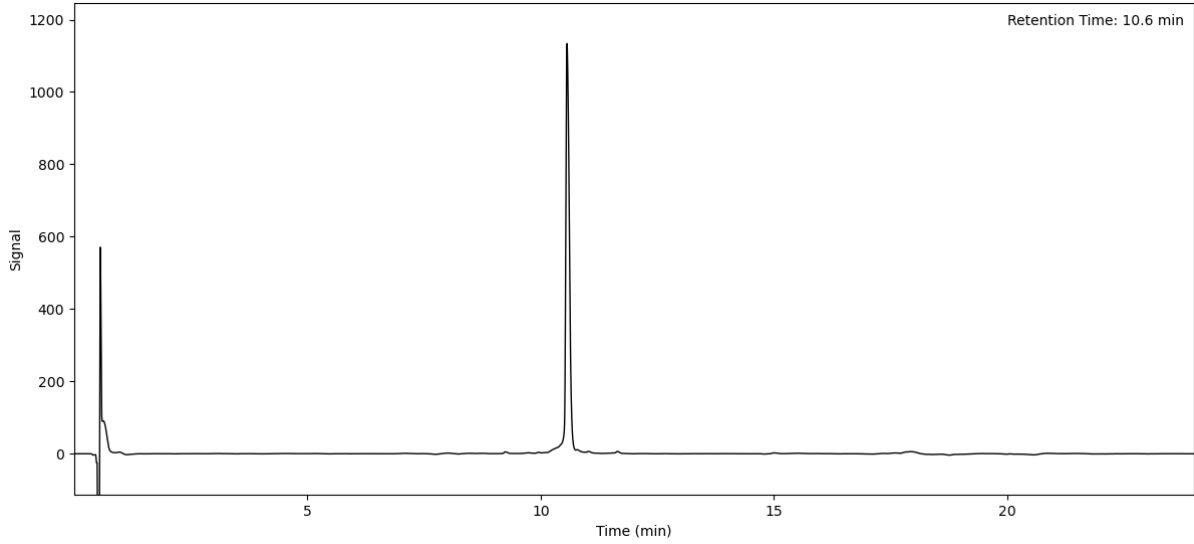
Peptide 36 (Method A)



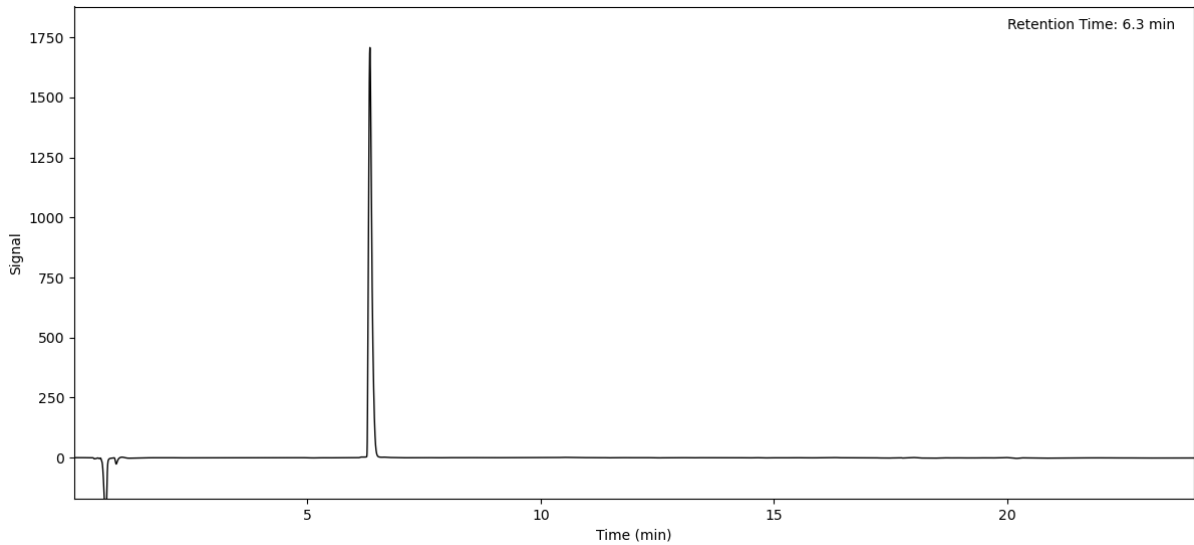
Peptide 37 (Method A)



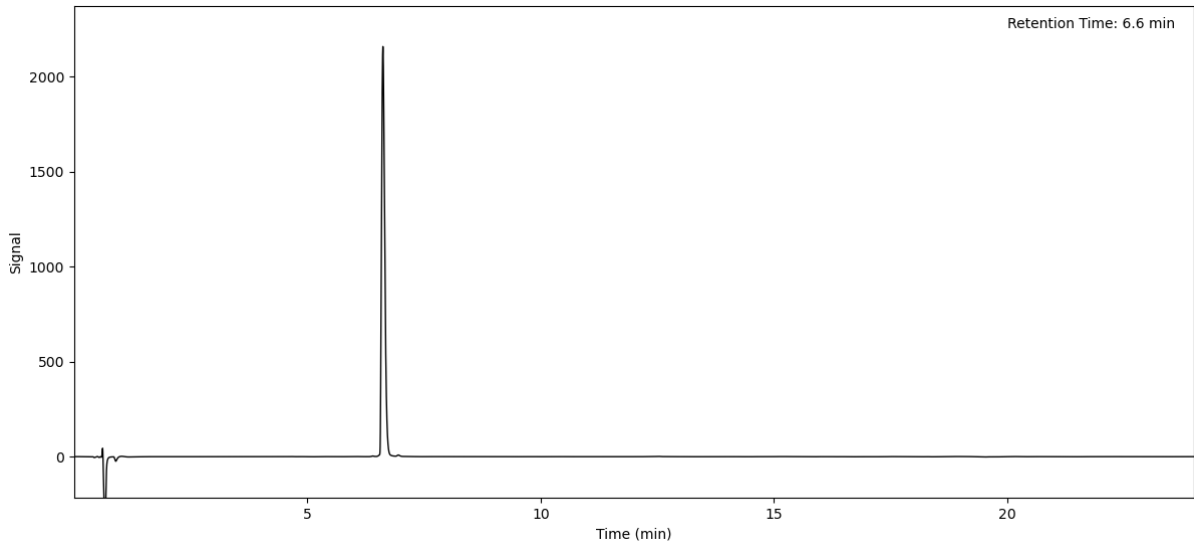
Peptide 38 (Method A)



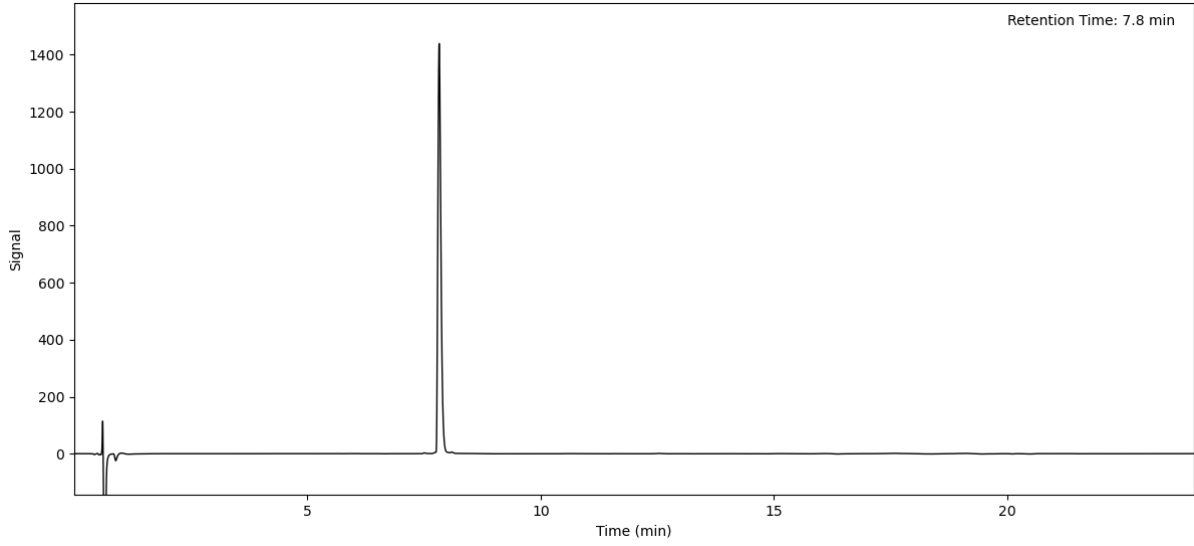
Peptide 39 (Method A)



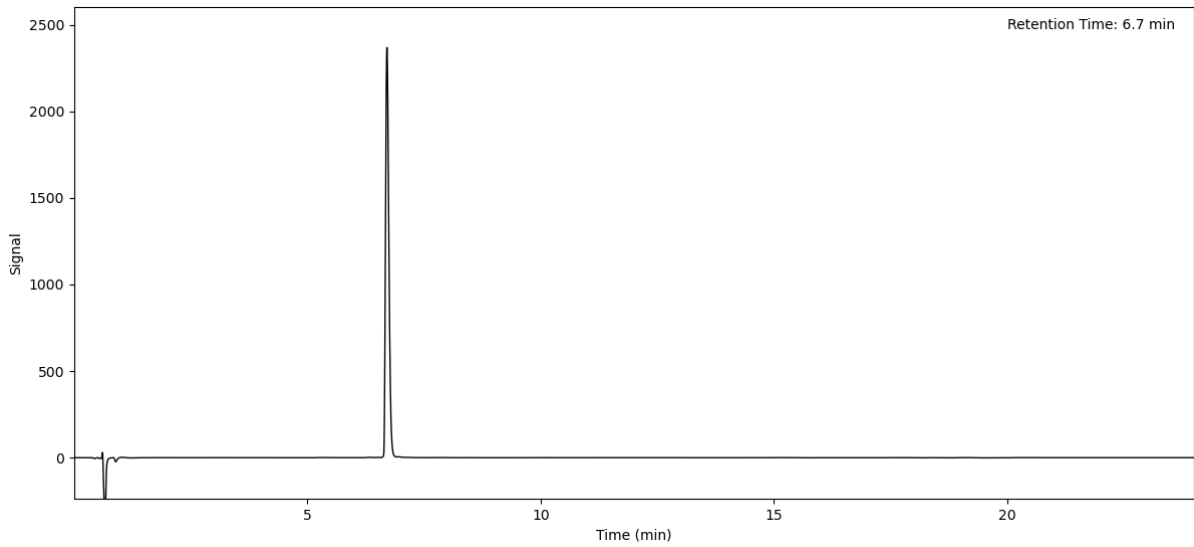
Peptide 40 (Method A)



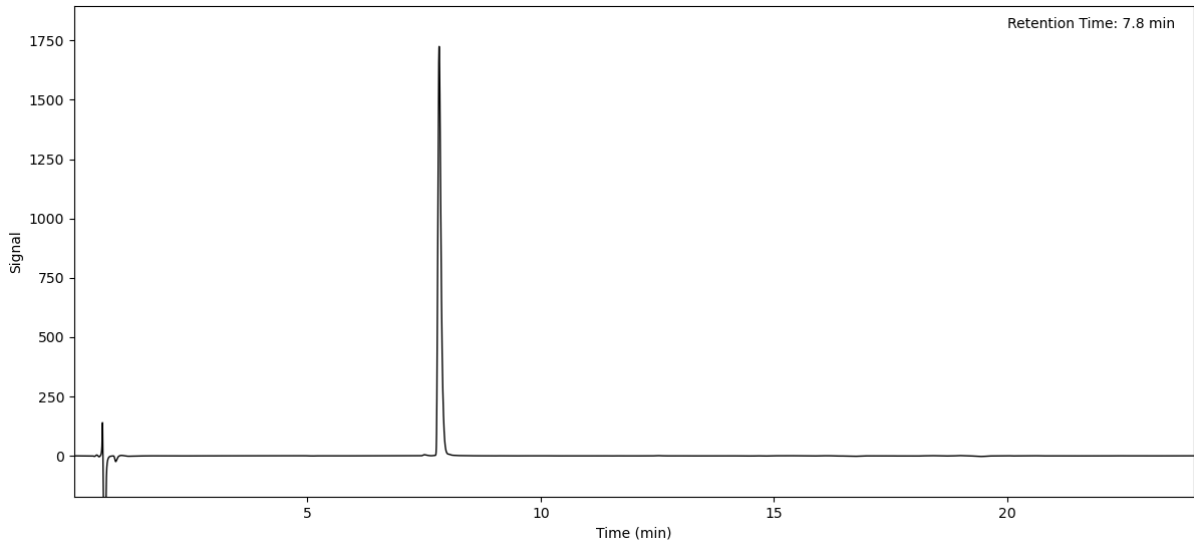
Peptide 41 (Method A)



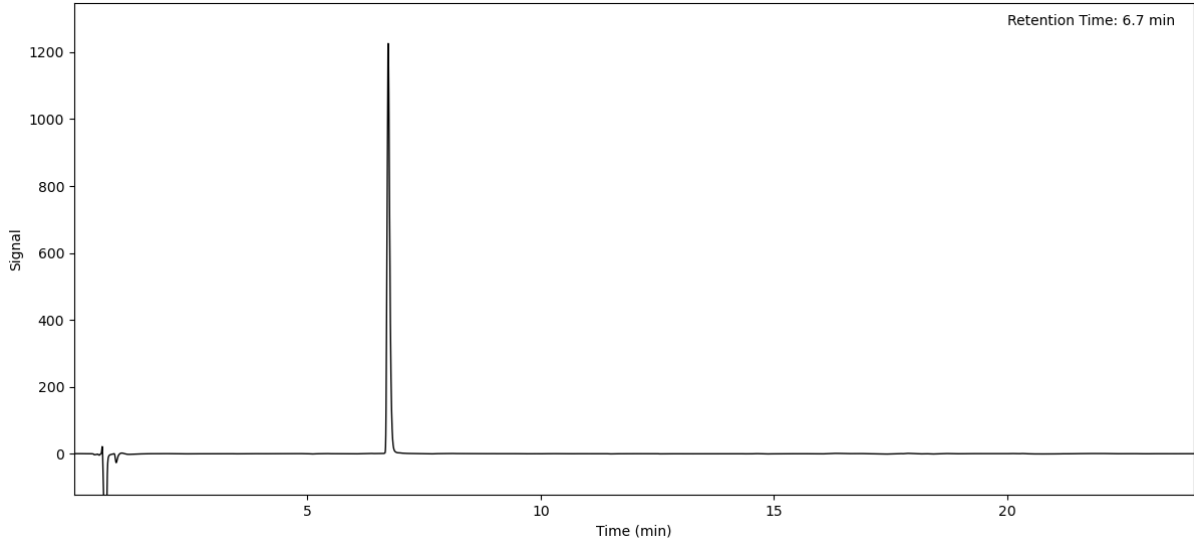
Peptide 42 (Method A)



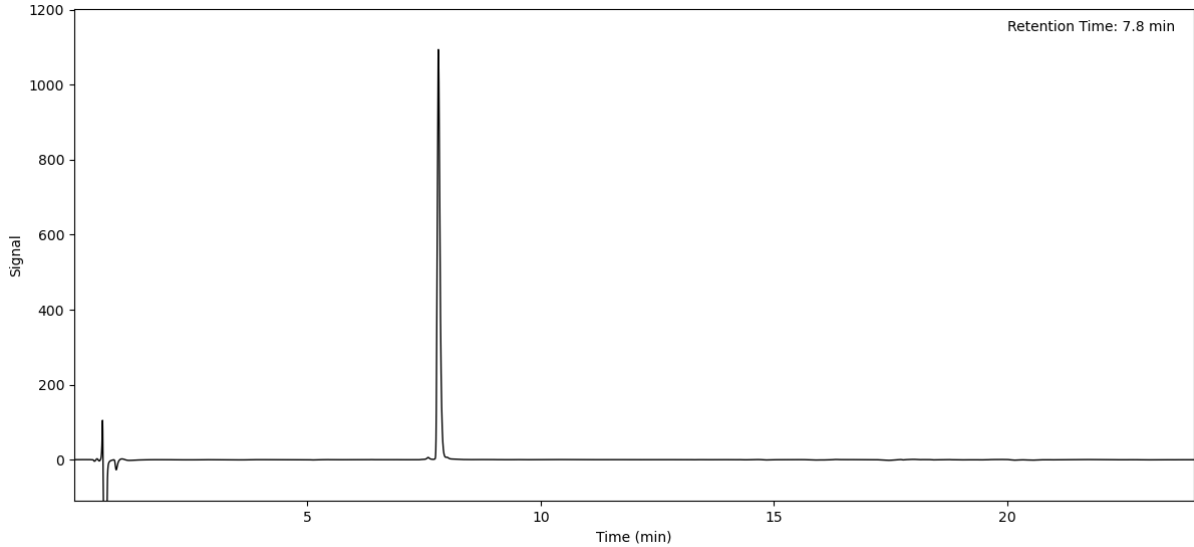
Peptide 43 (Method A)



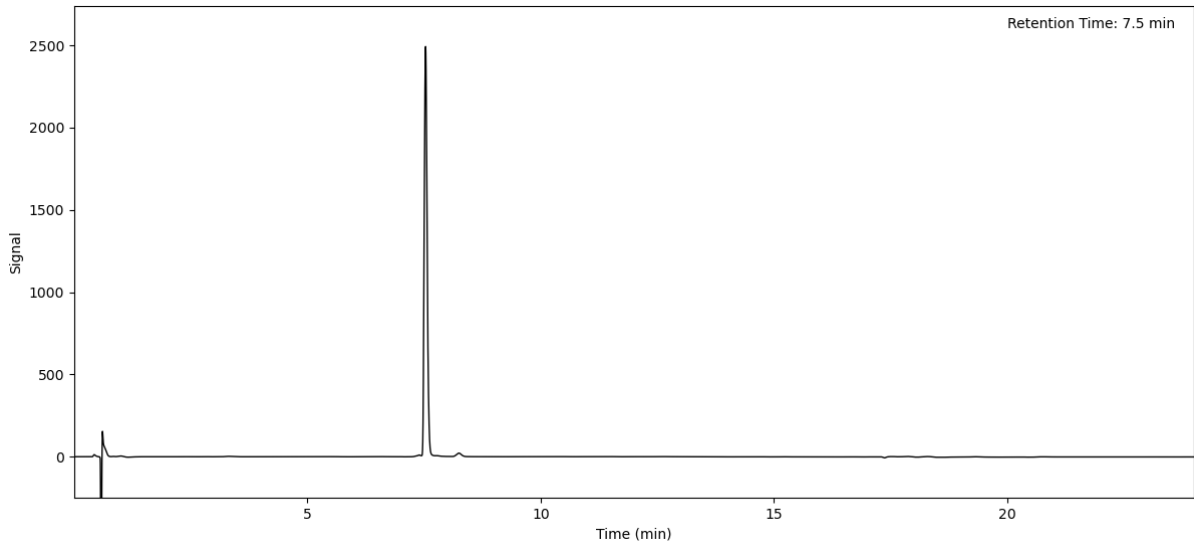
Peptide 44 (Method A)



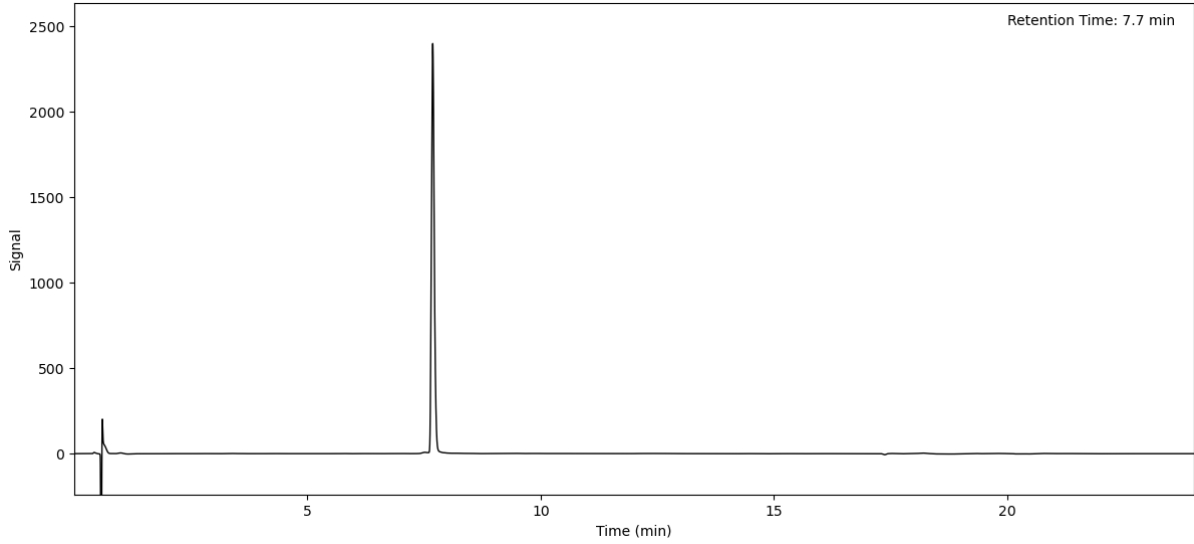
Peptide 45 (Method A)



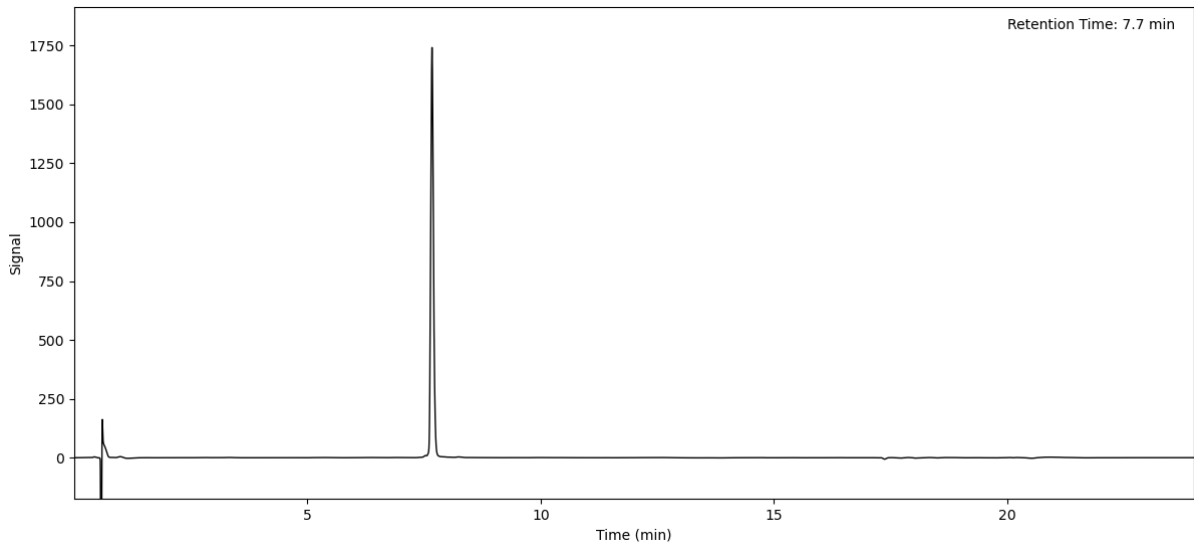
Peptide 46 (Method A)



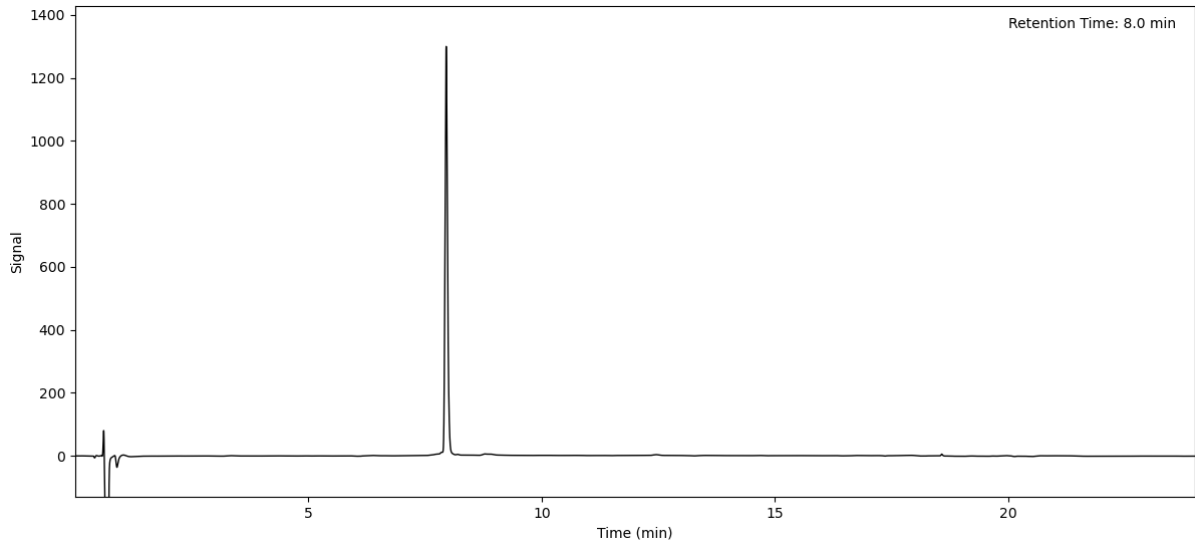
Peptide 47 (Method A)



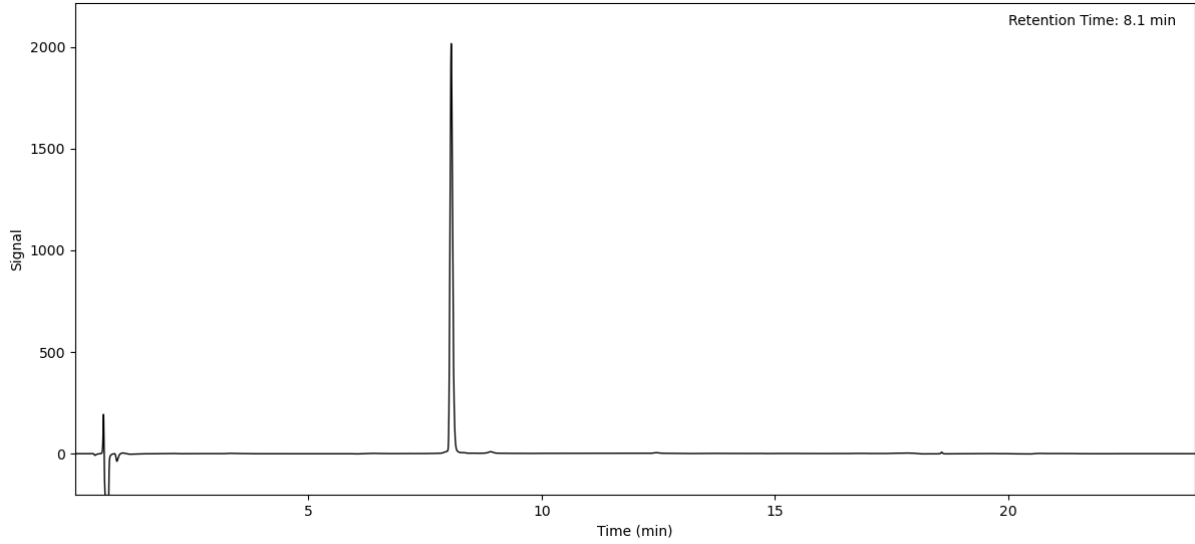
Peptide 48 (Method A)



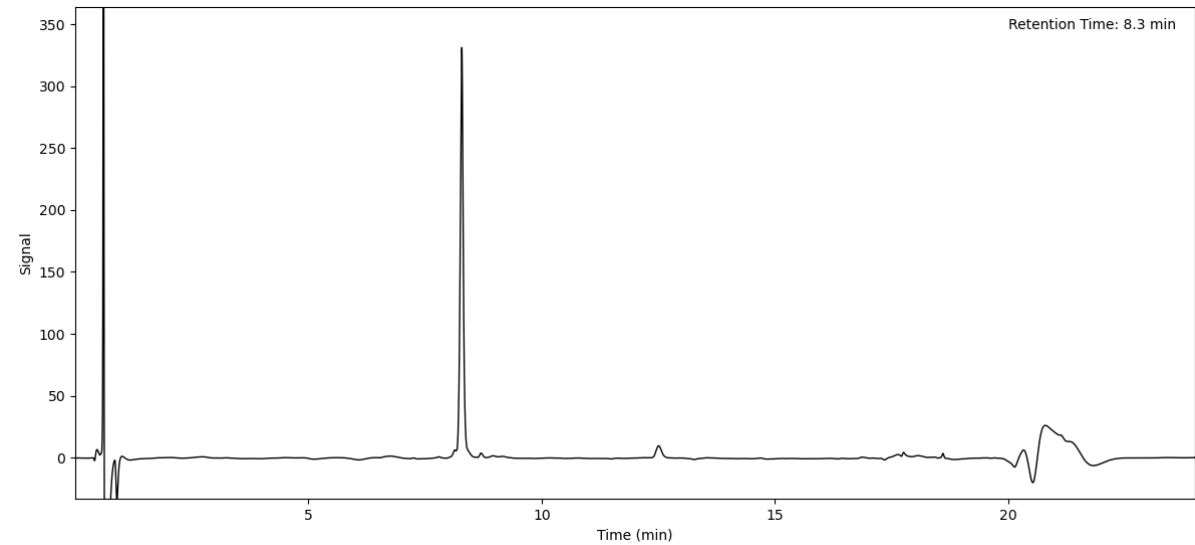
Peptide 49 (Method A)



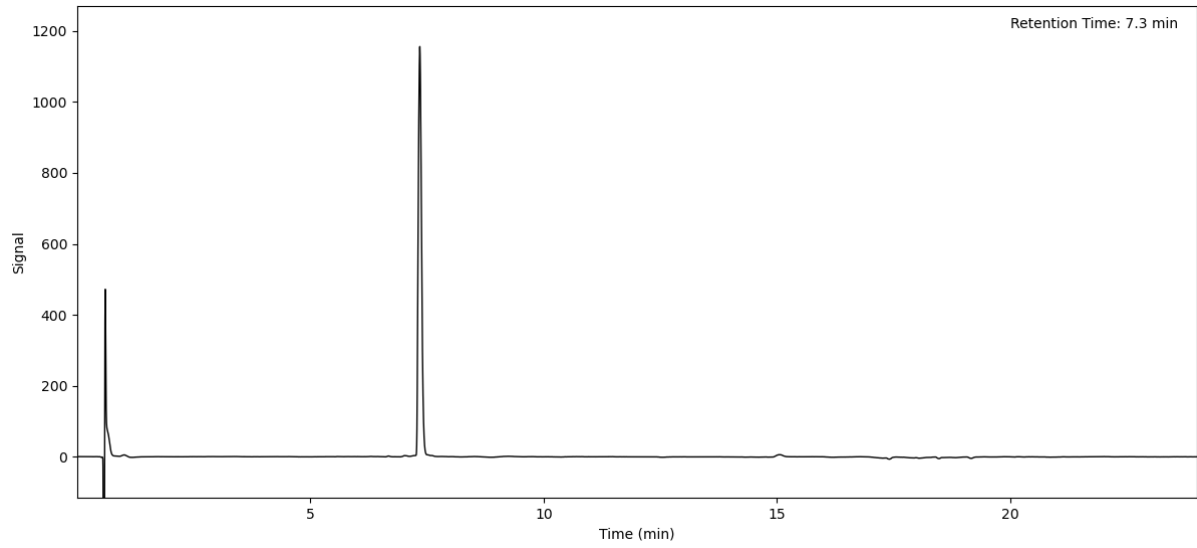
Peptide 50 (Method A)



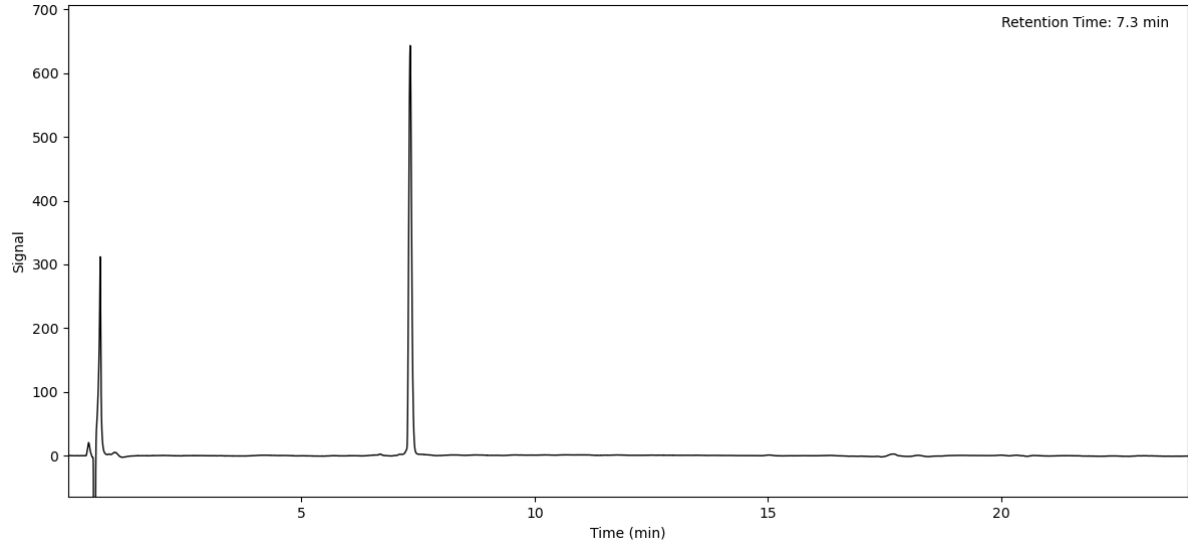
Peptide 51 (Method A)



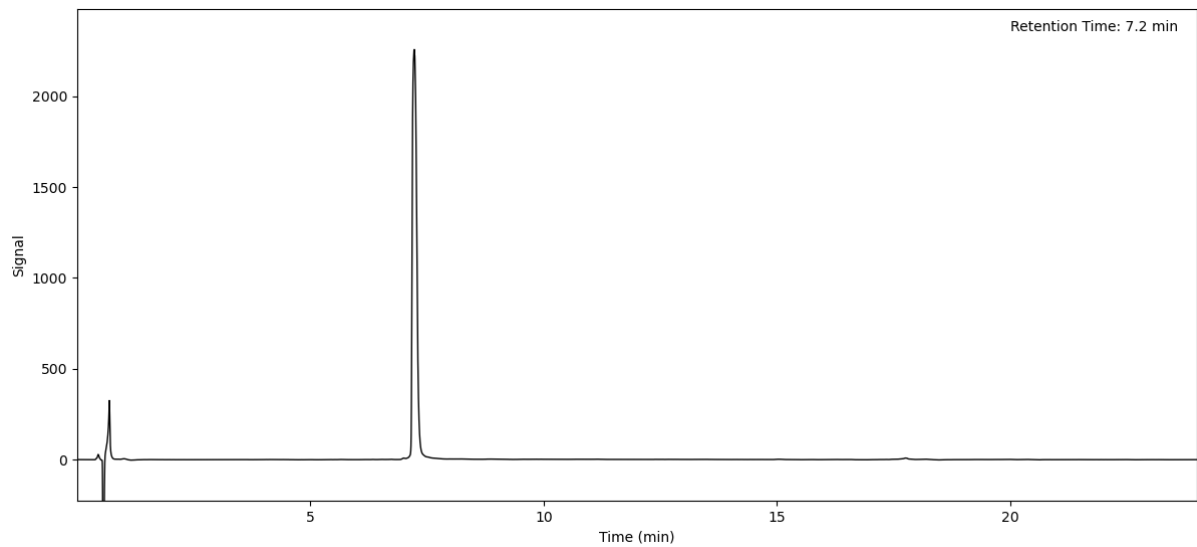
Peptide 52 (Method A)



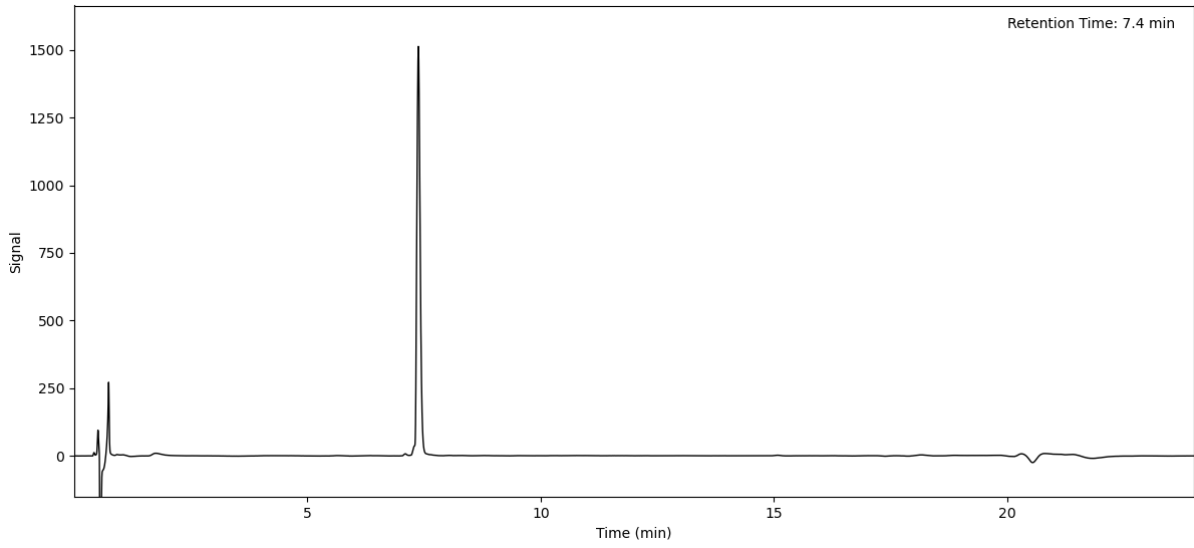
Peptide 53 (Method A)



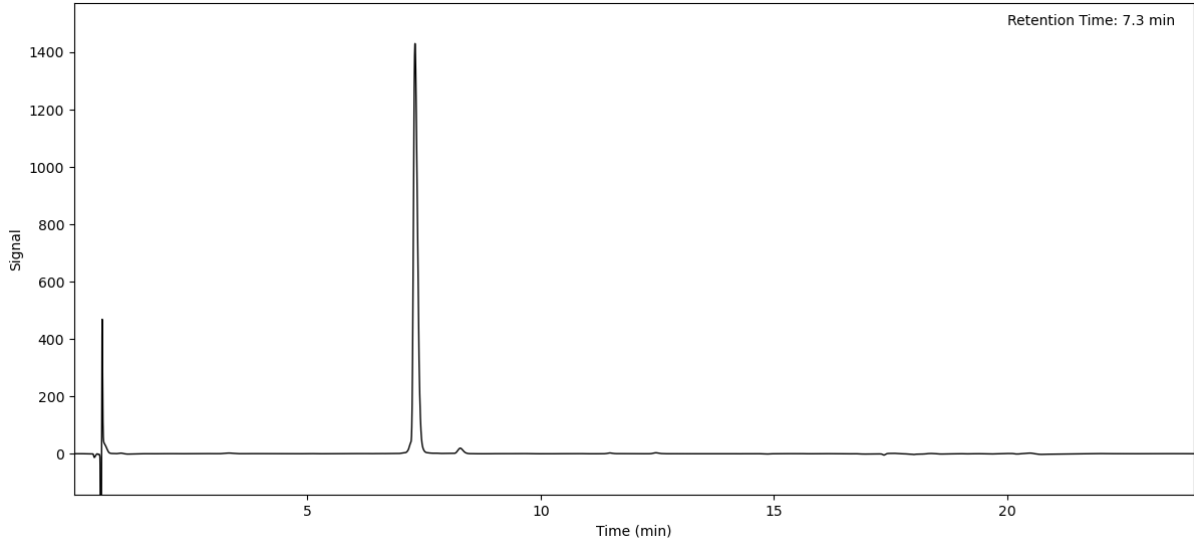
Peptide 54 (Method A)



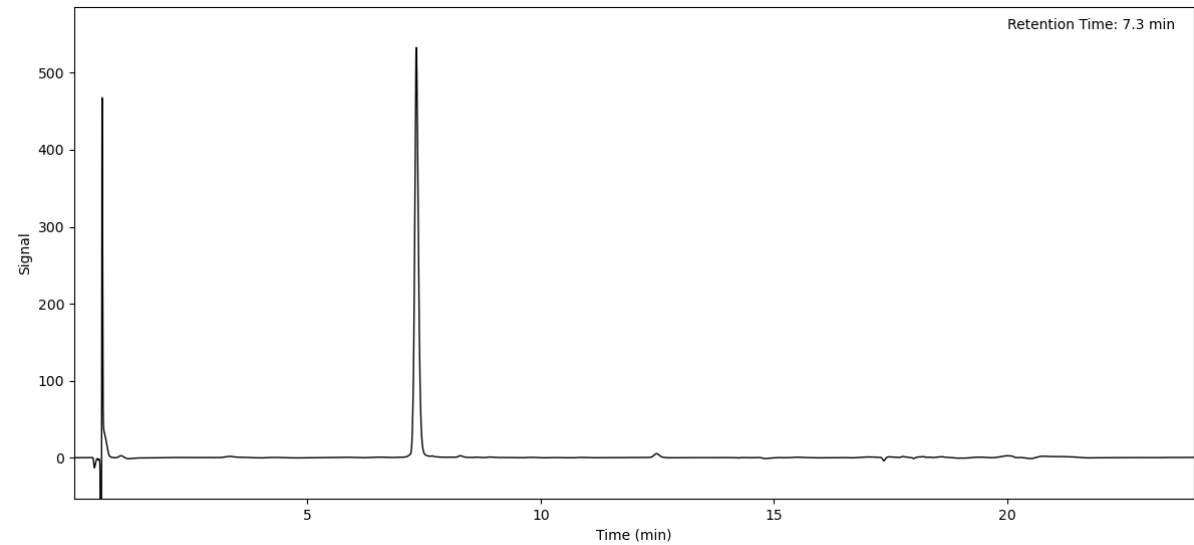
Peptide 55 (Method A)



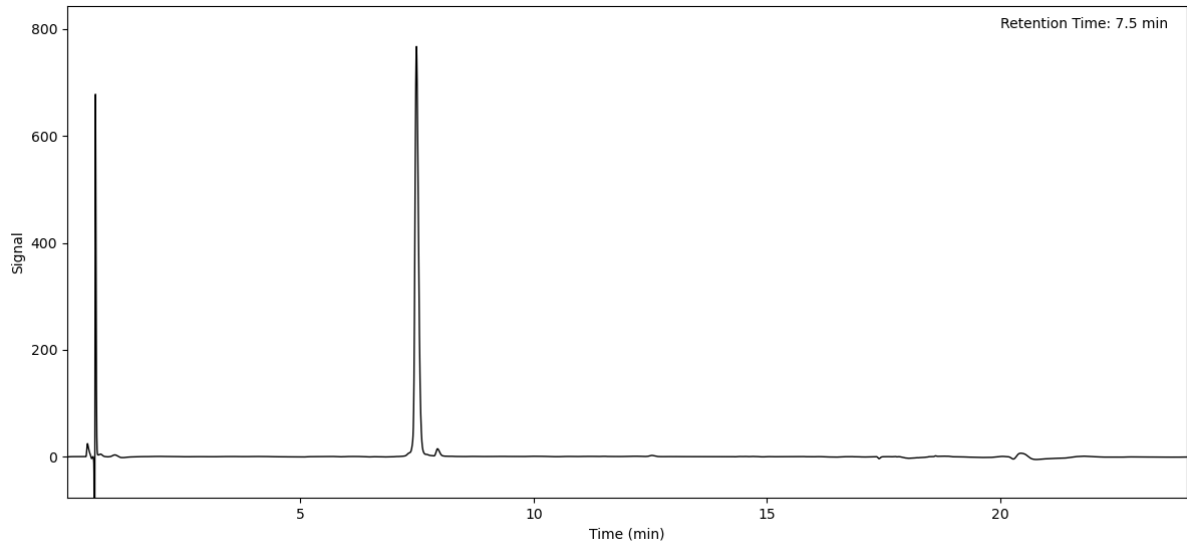
Peptide 56 (Method A)



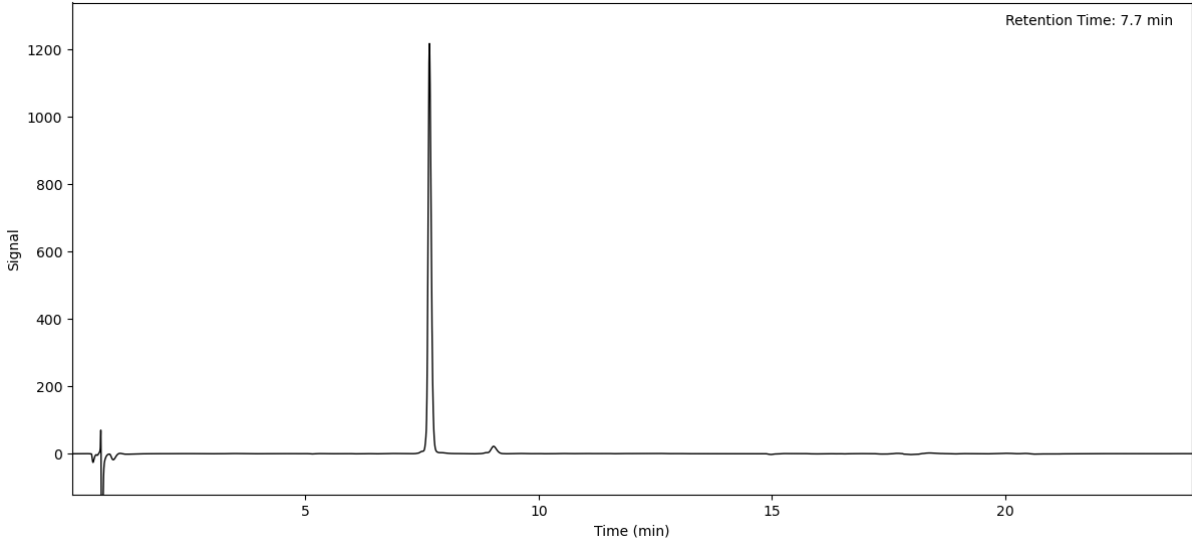
Peptide 57 (Method A)



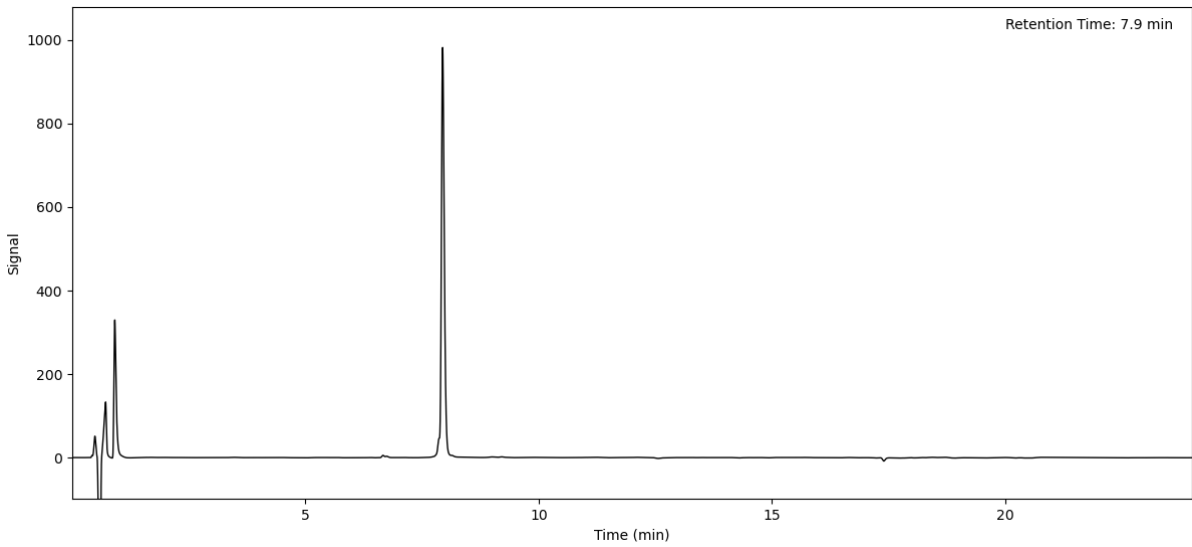
Peptide 58 (Method A)



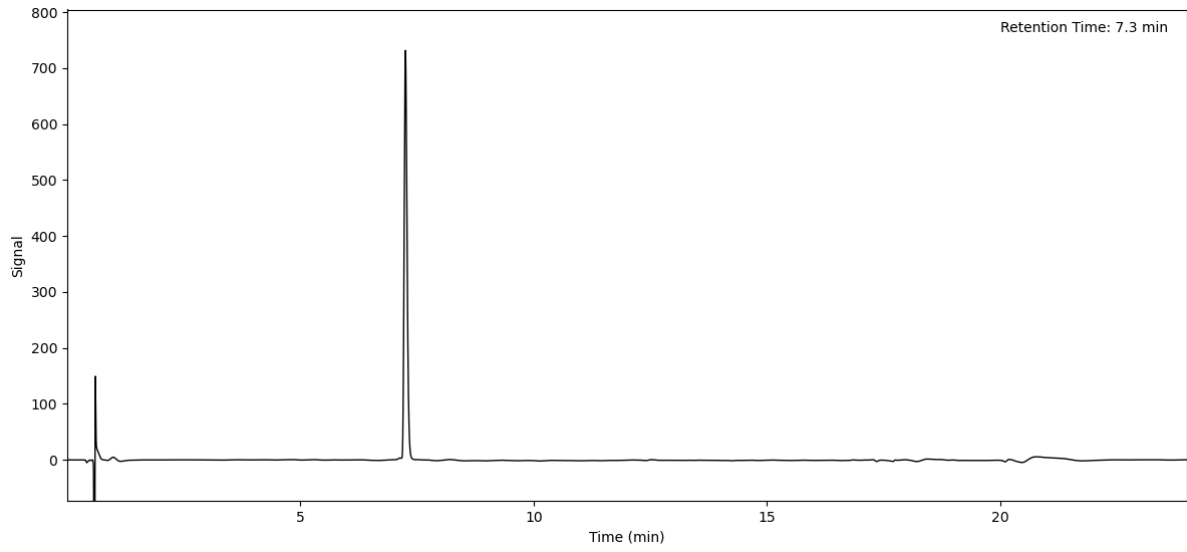
Peptide 59 (Method A)



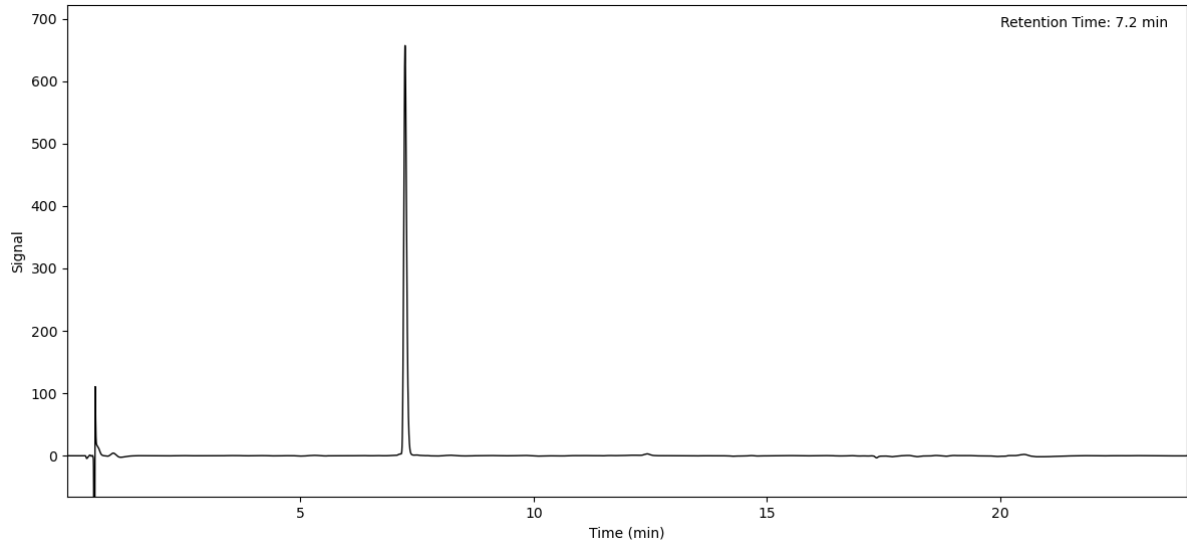
Peptide 60 (Method A)



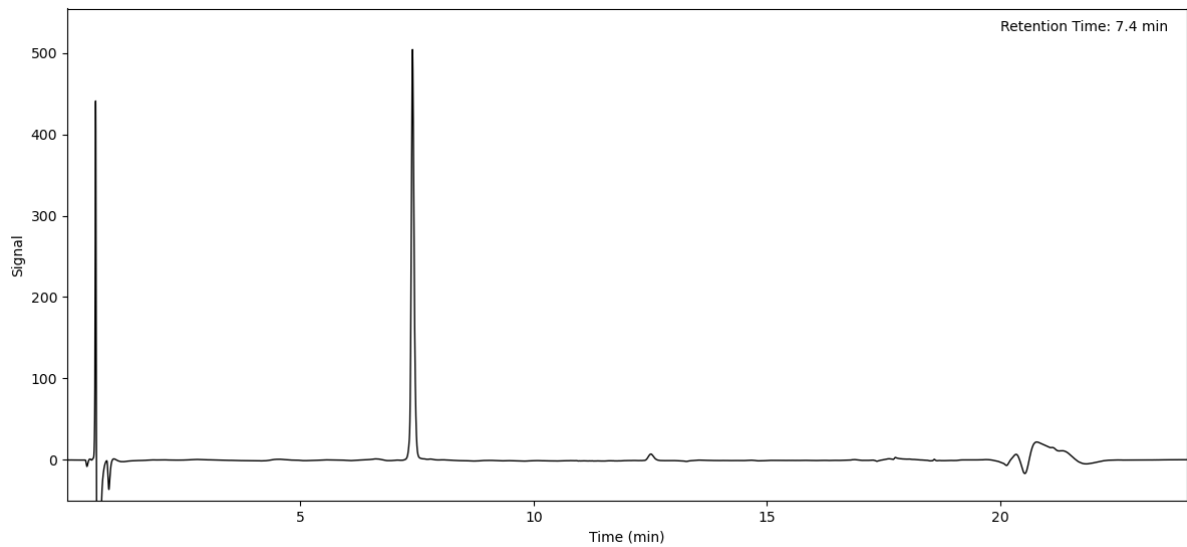
Peptide 61 (Method A)



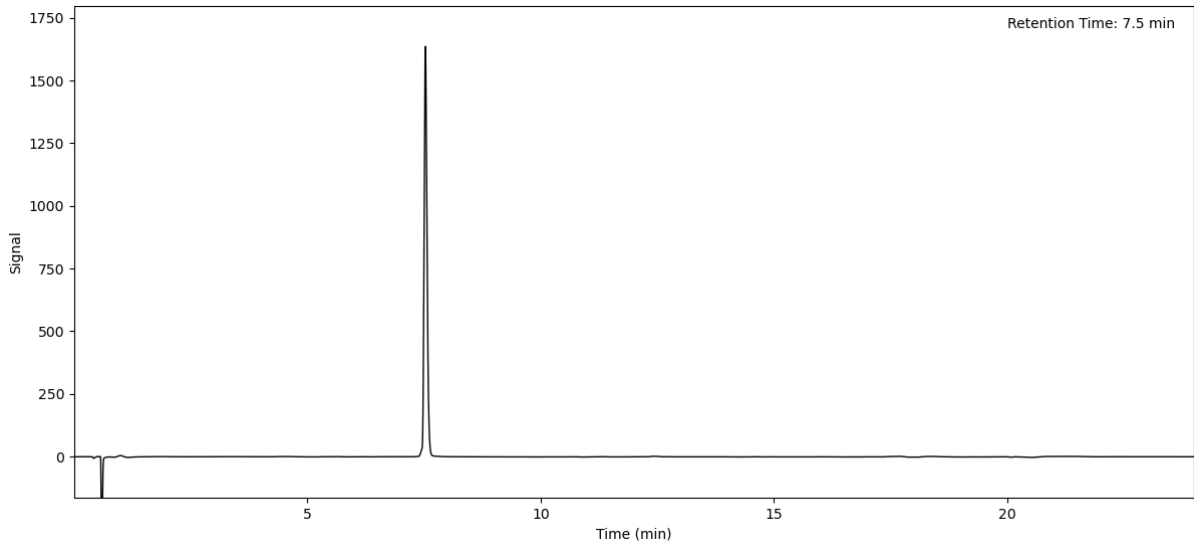
Peptide 62 (Method A)



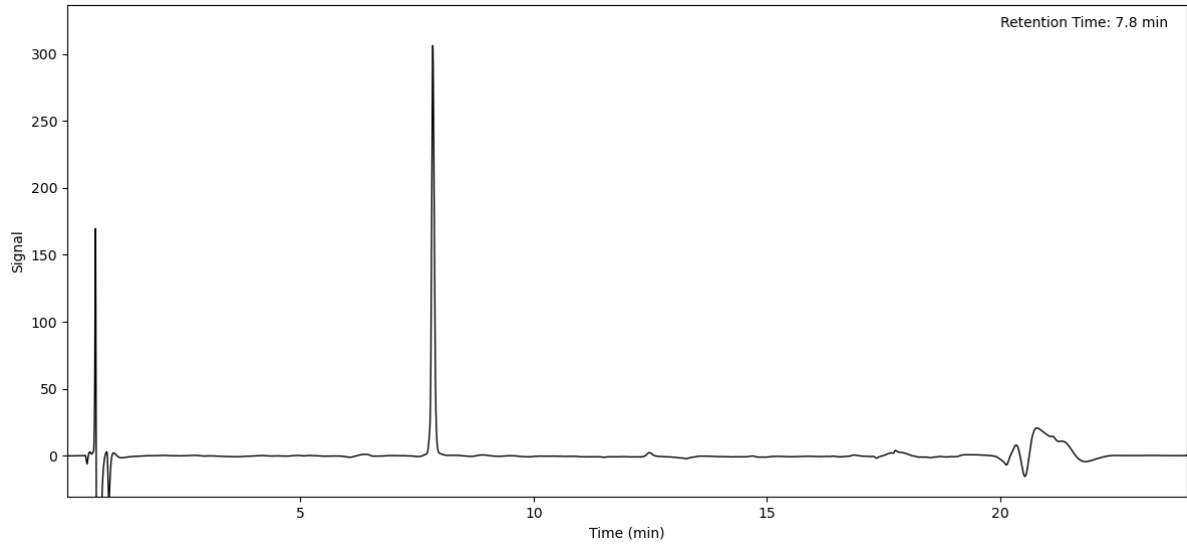
Peptide 63 (Method A)



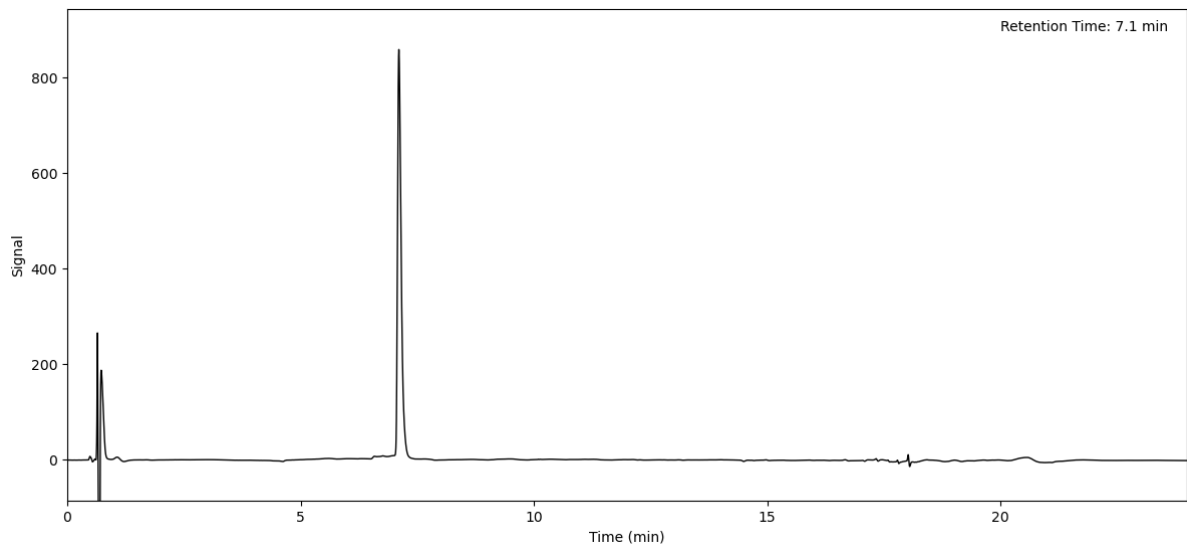
Peptide 64 (Method A)



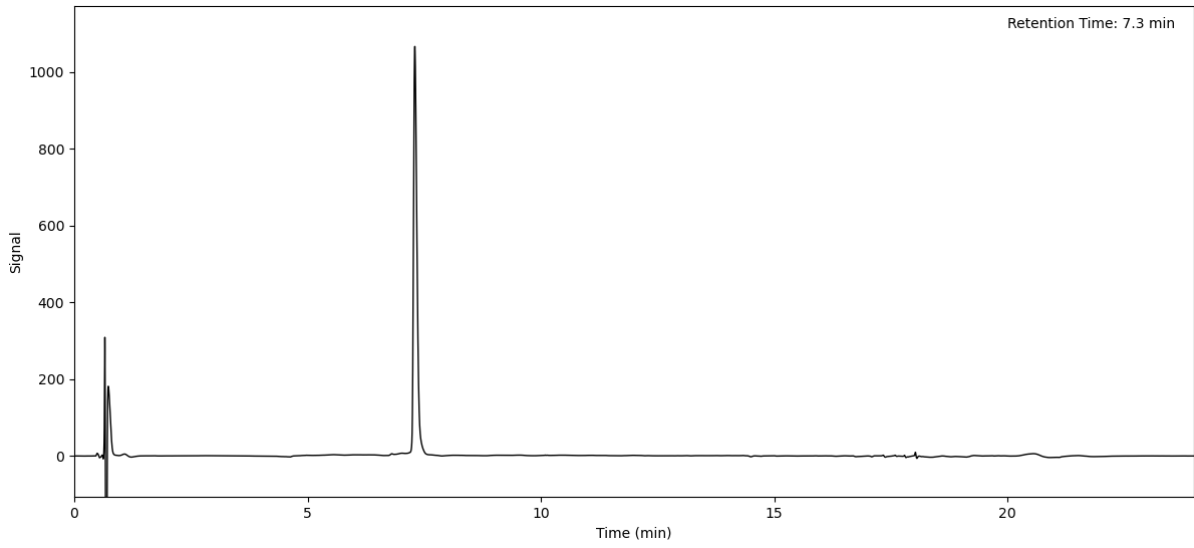
Peptide 65 (Method A)



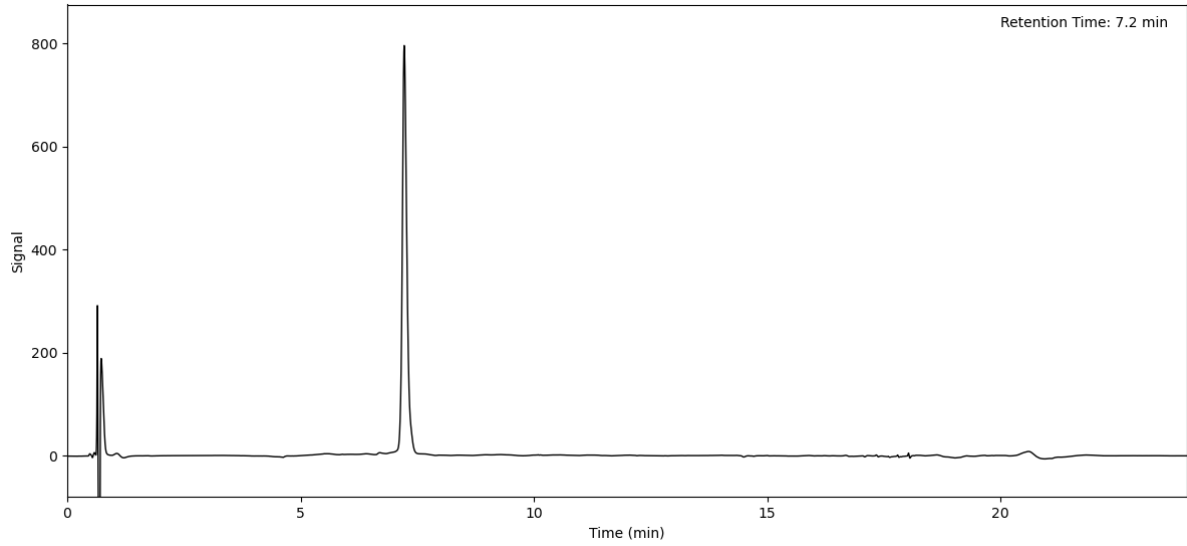
Peptide 66 (Method A)



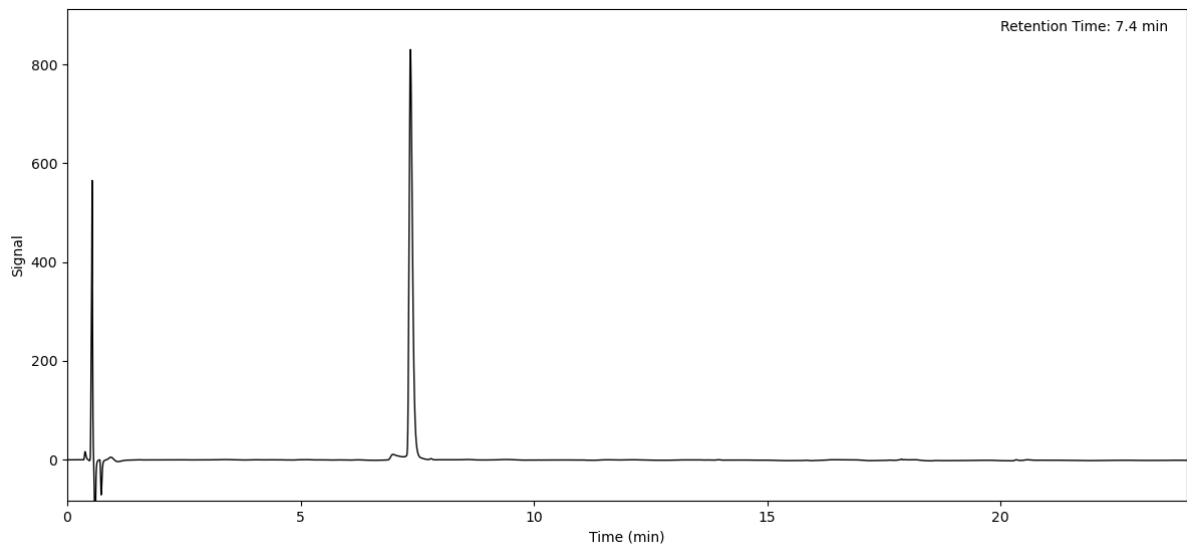
Peptide 67 (Method A)



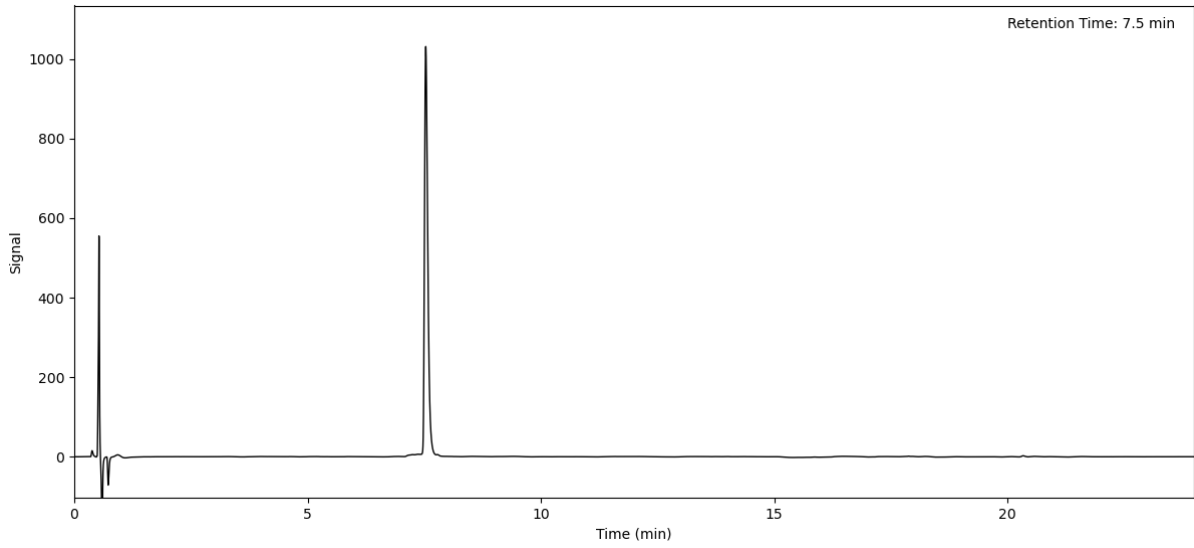
Peptide 68 (Method A)



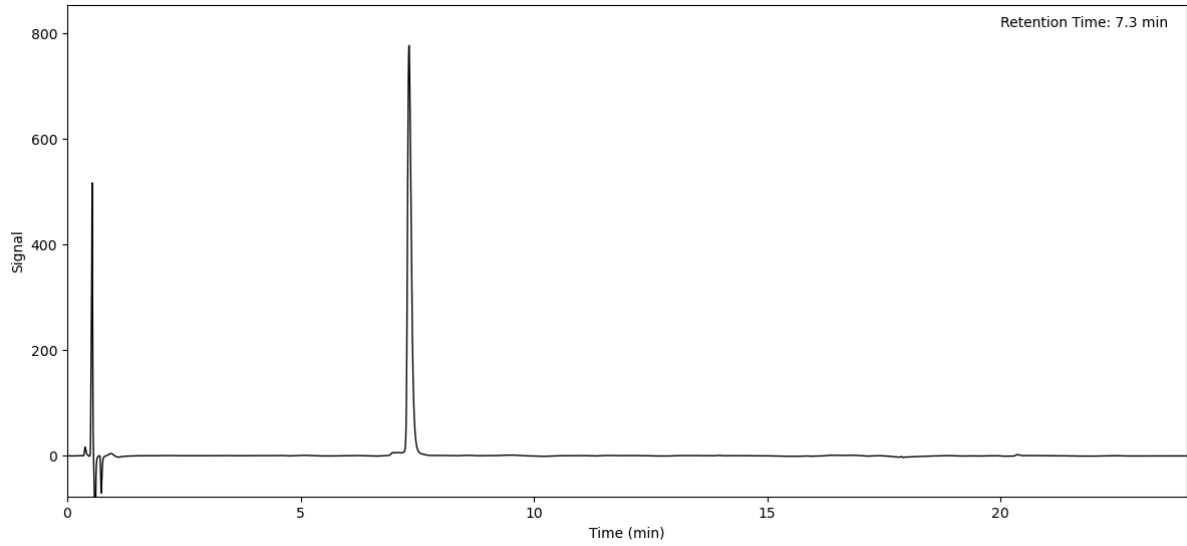
Peptide 69 (Method A)



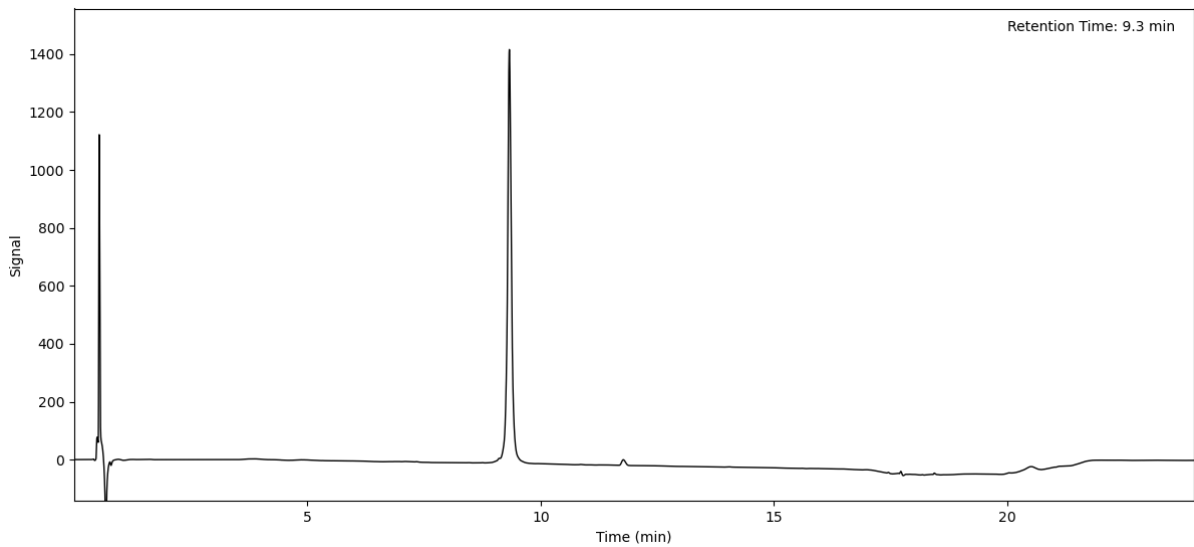
Peptide 70 (Method A)



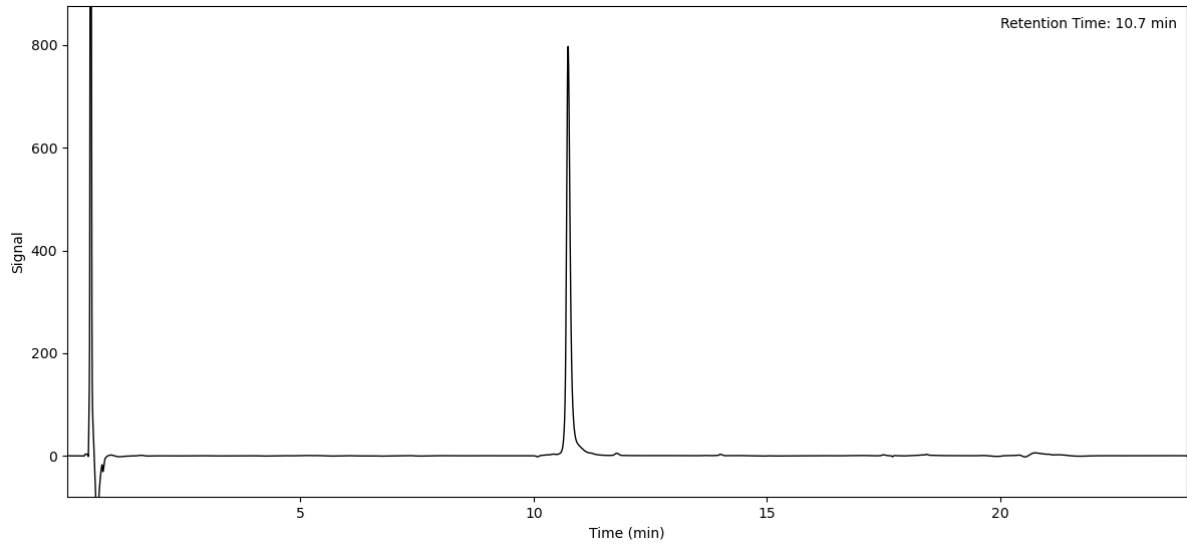
Peptide 71 (Method A)



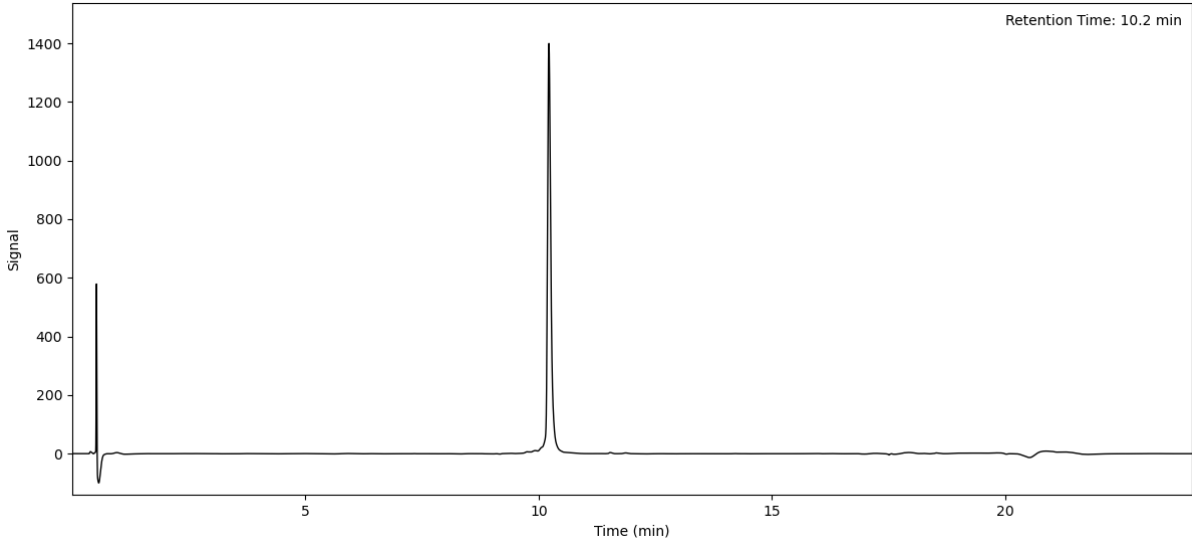
Peptide 72 (Method A)



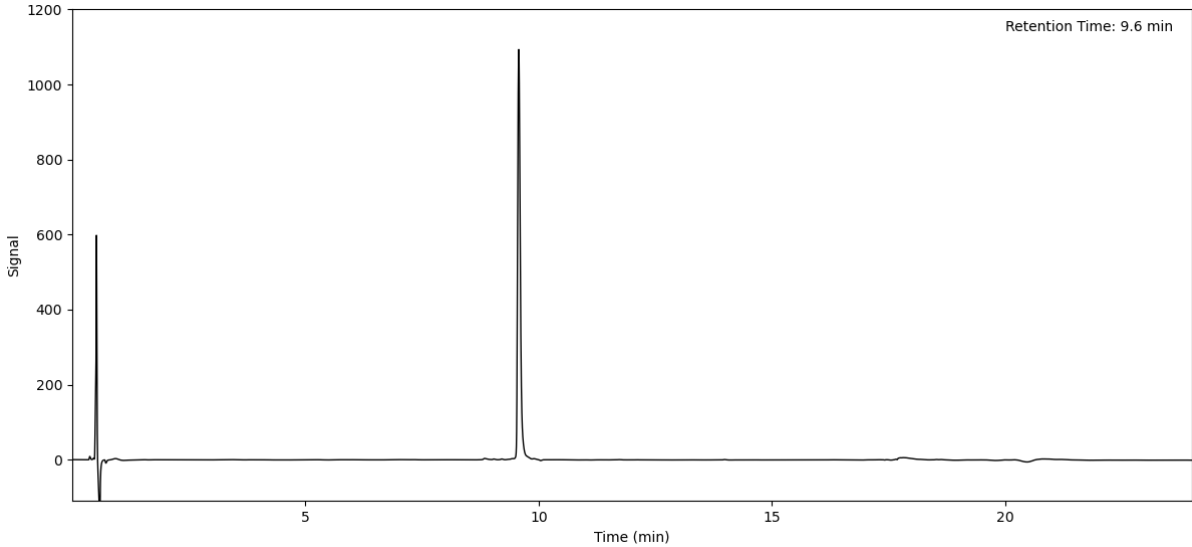
Peptide 73 (Method A)



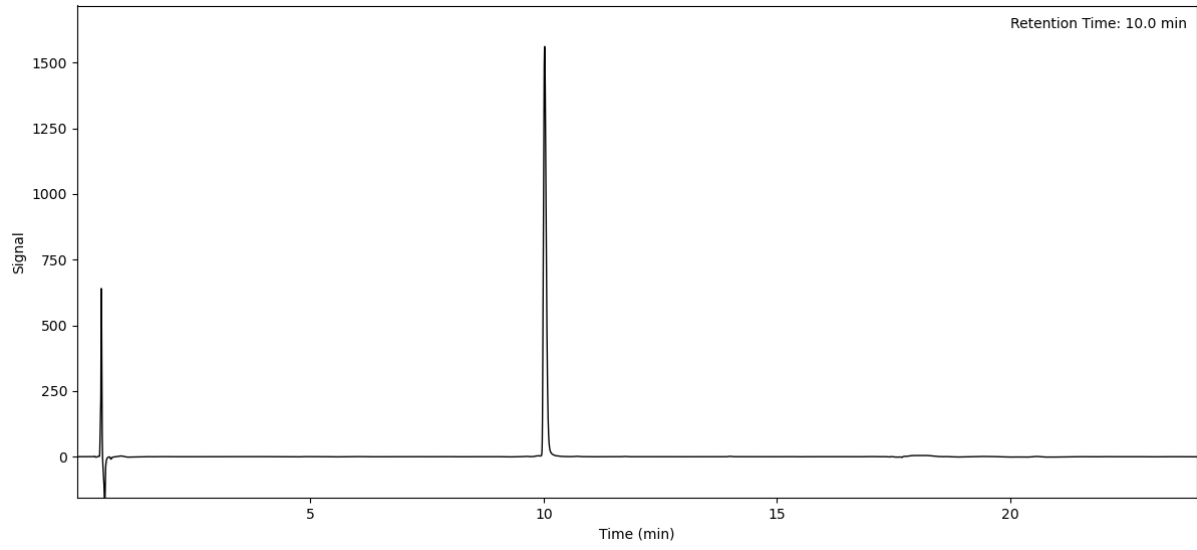
Peptide 74 (Method A)



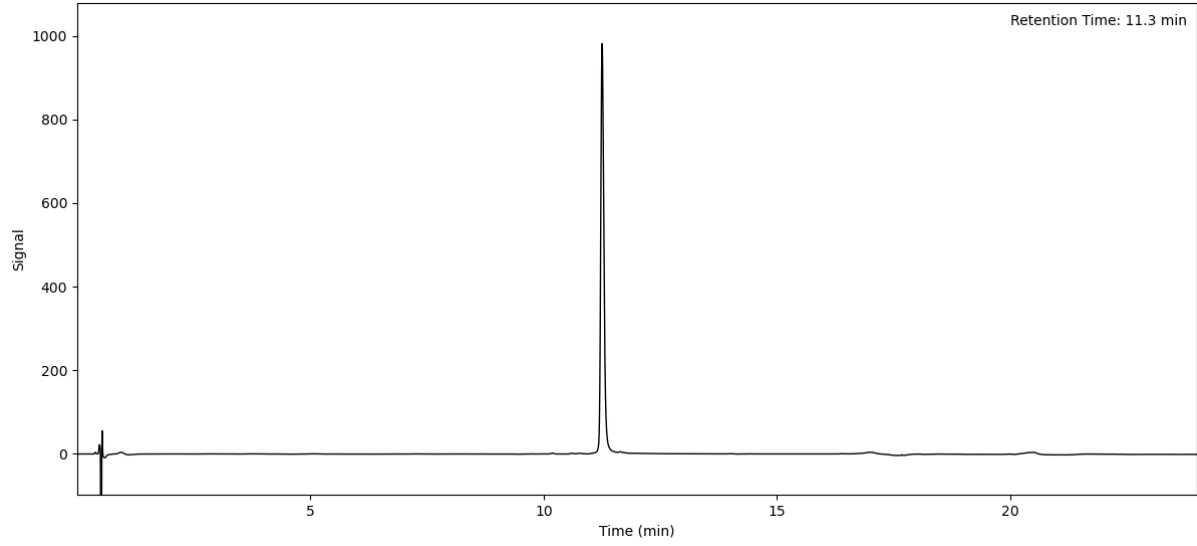
Peptide 75 (Method A)



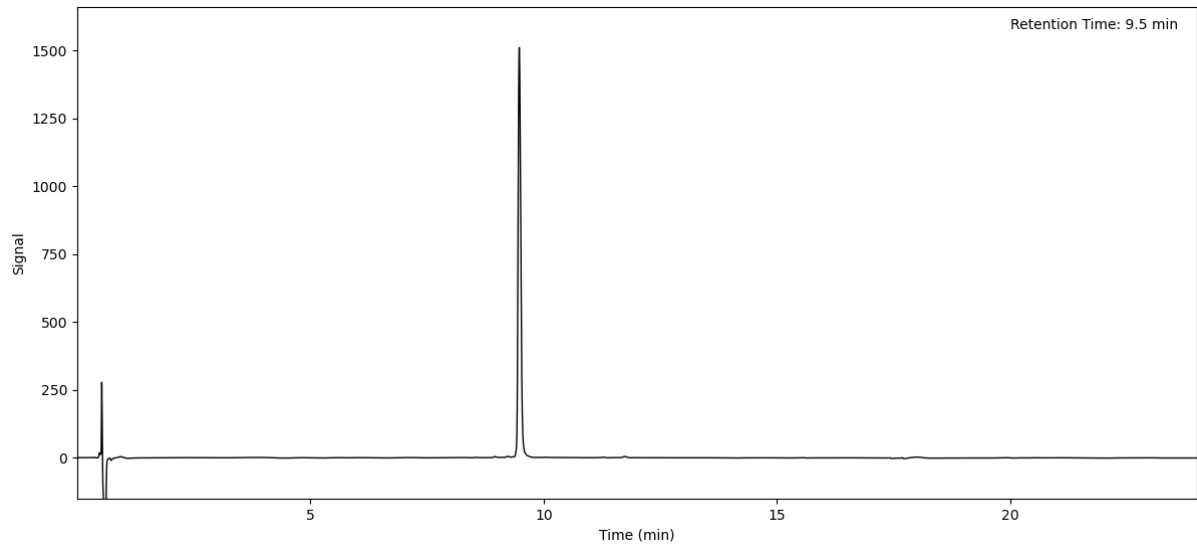
Peptide 76 (Method A)



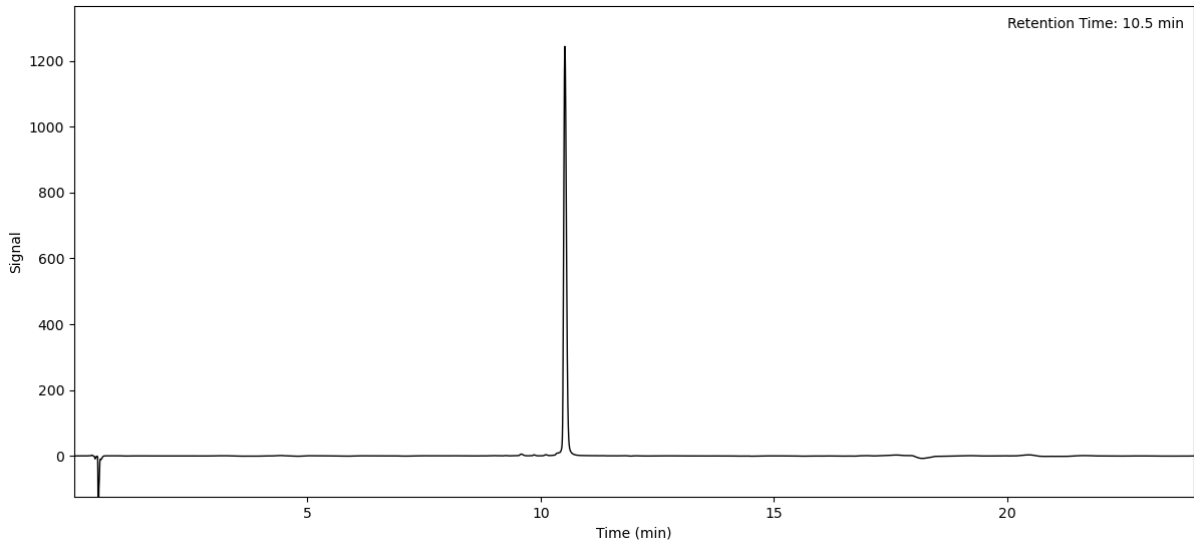
Peptide 77 (Method A)



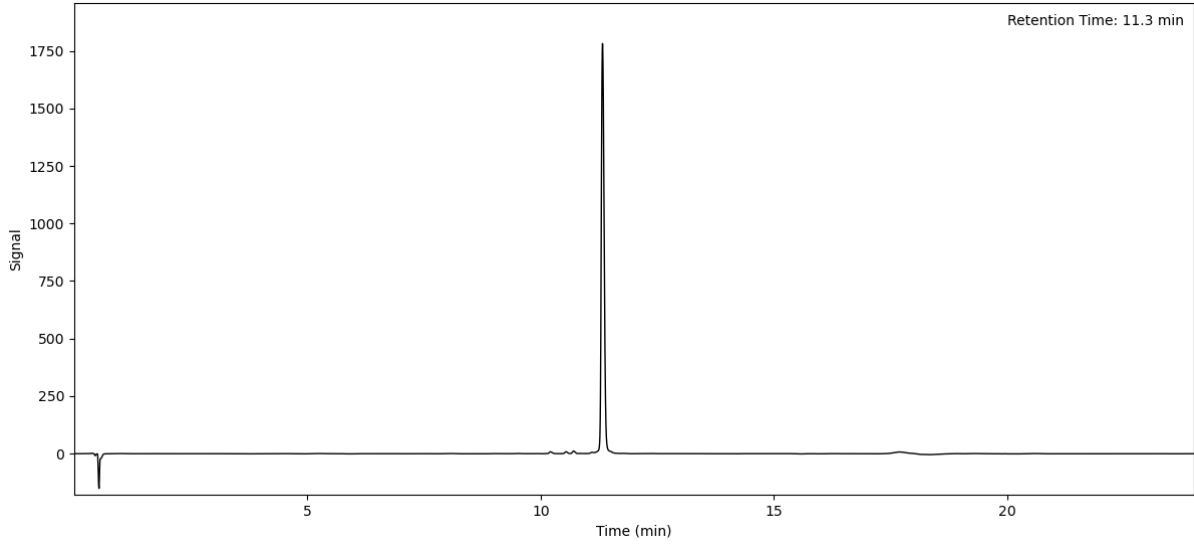
Peptide 78 (Method A)



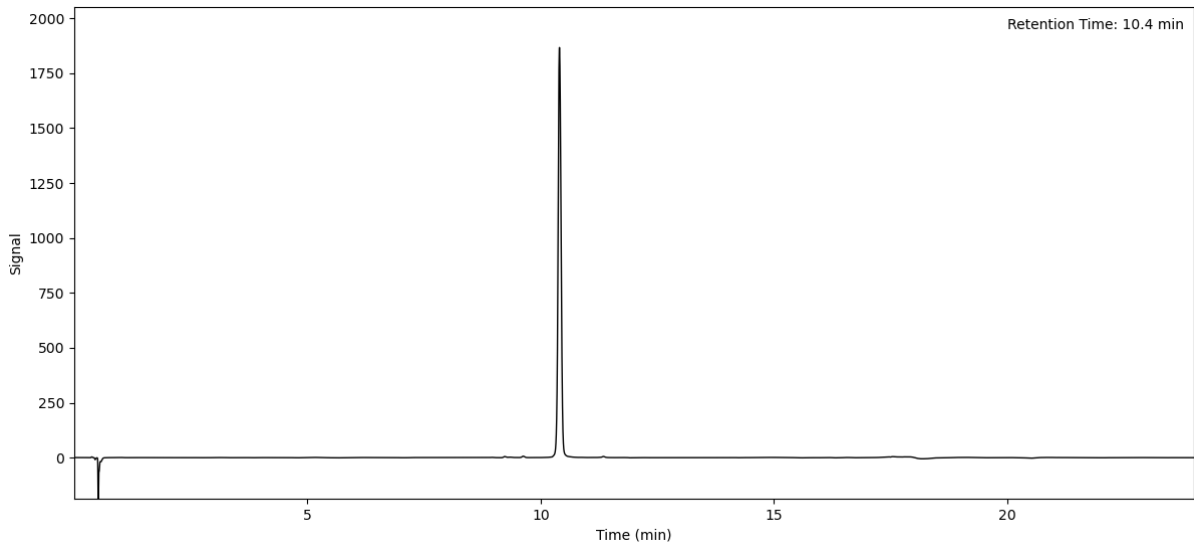
Peptide 79 (Method A)



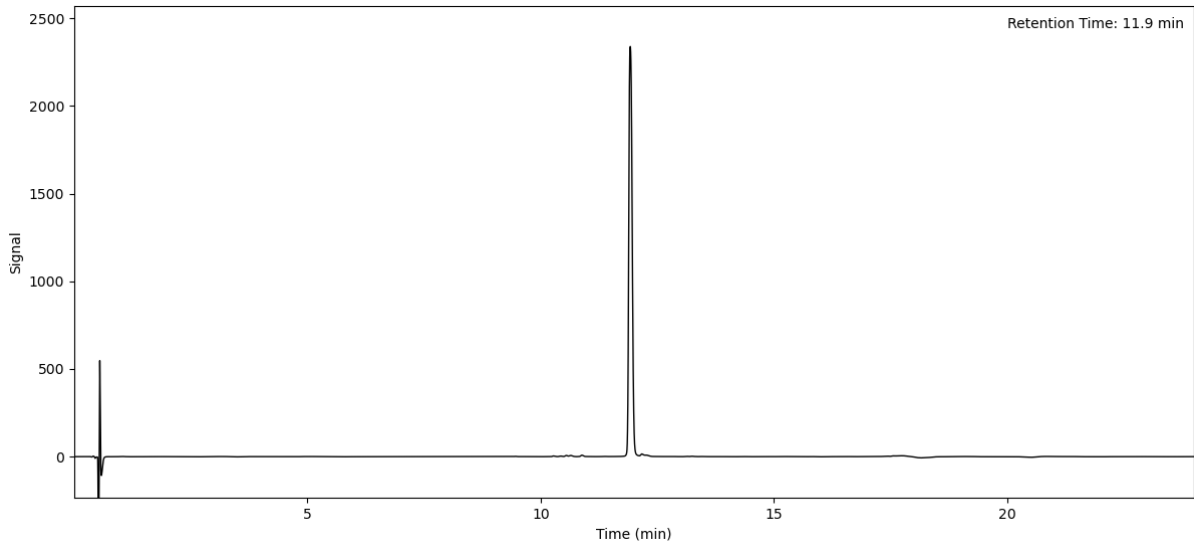
Peptide 80 (Method A)



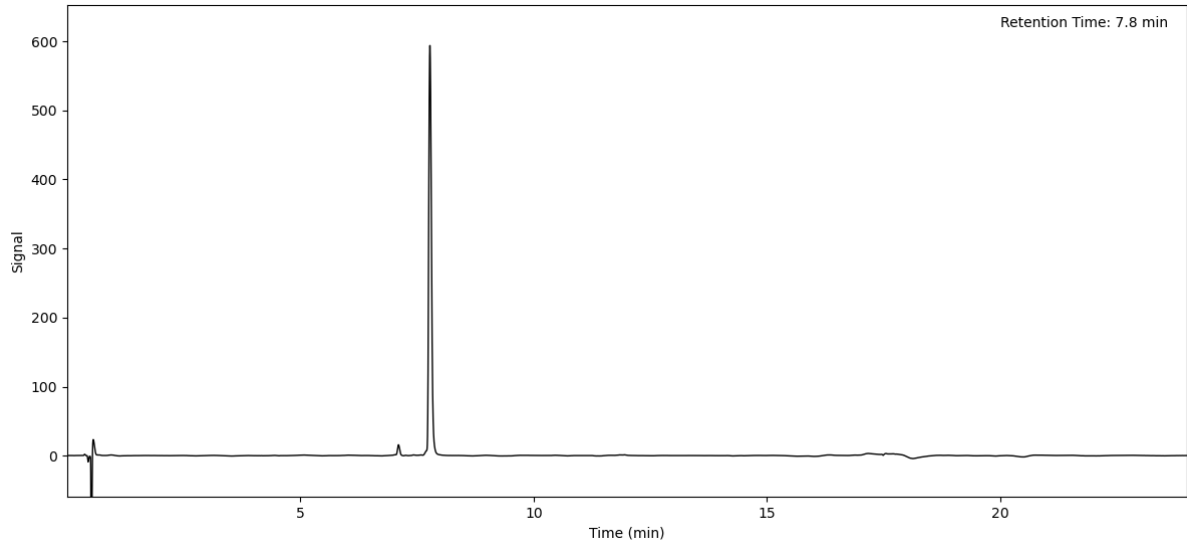
Peptide 81 (Method A)



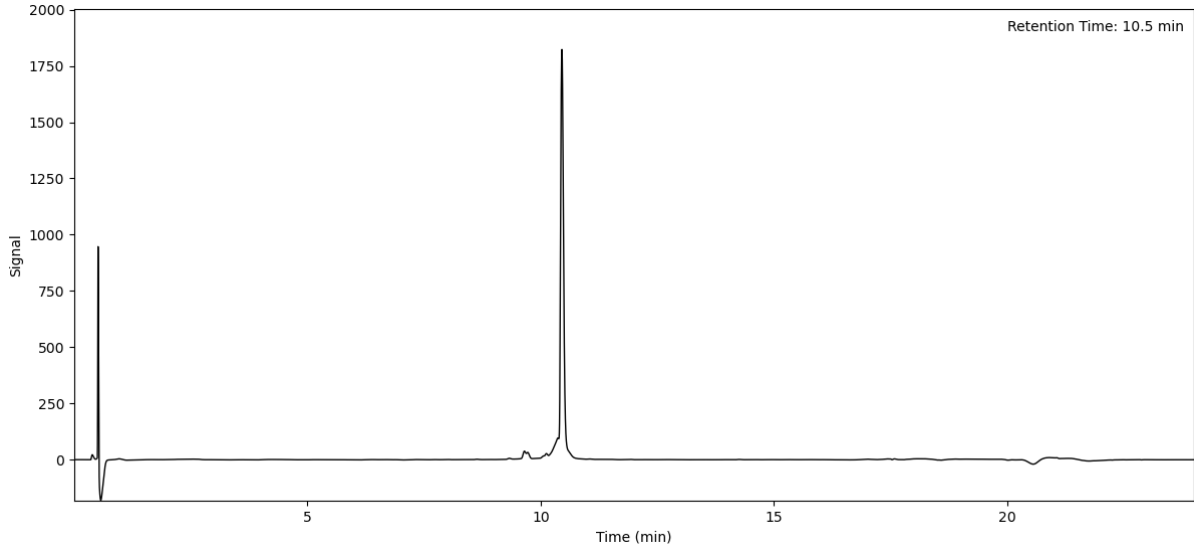
Peptide 82 (Method A)



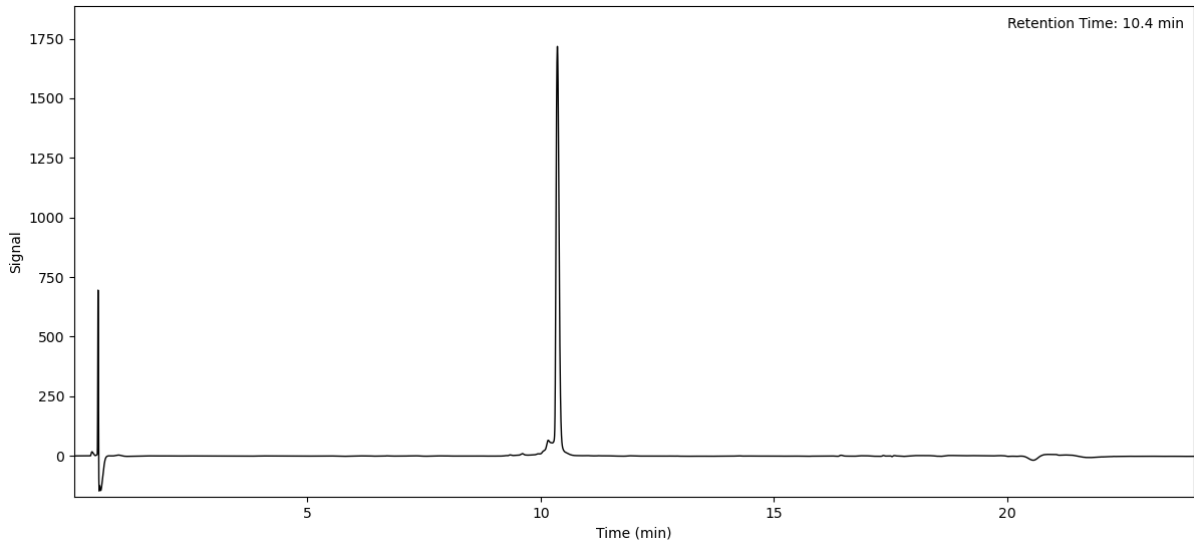
Peptide 83 (Method A)



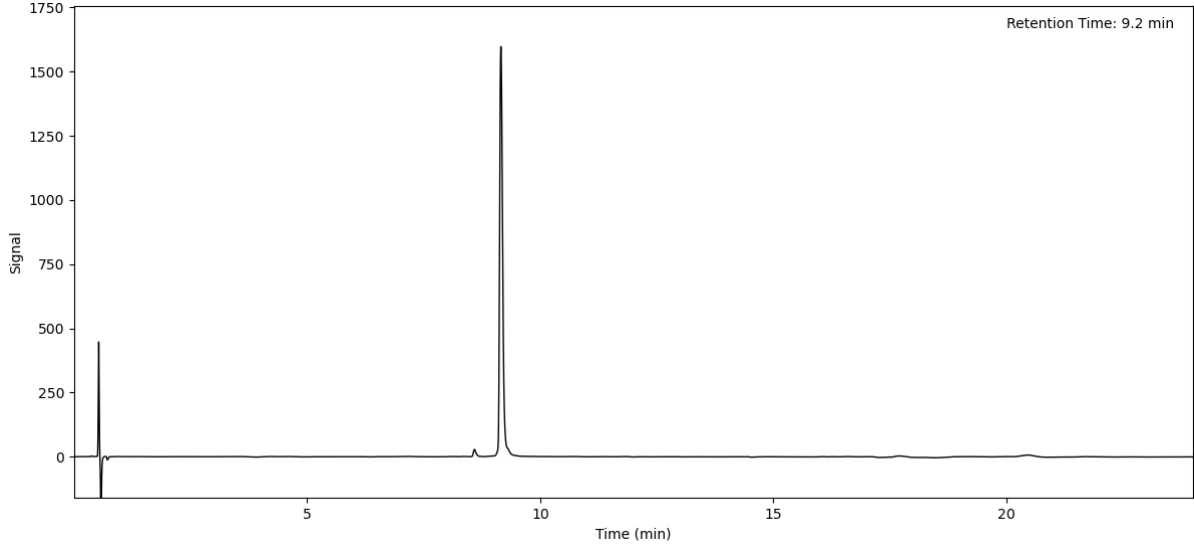
SUZ14_FITC (Method A)



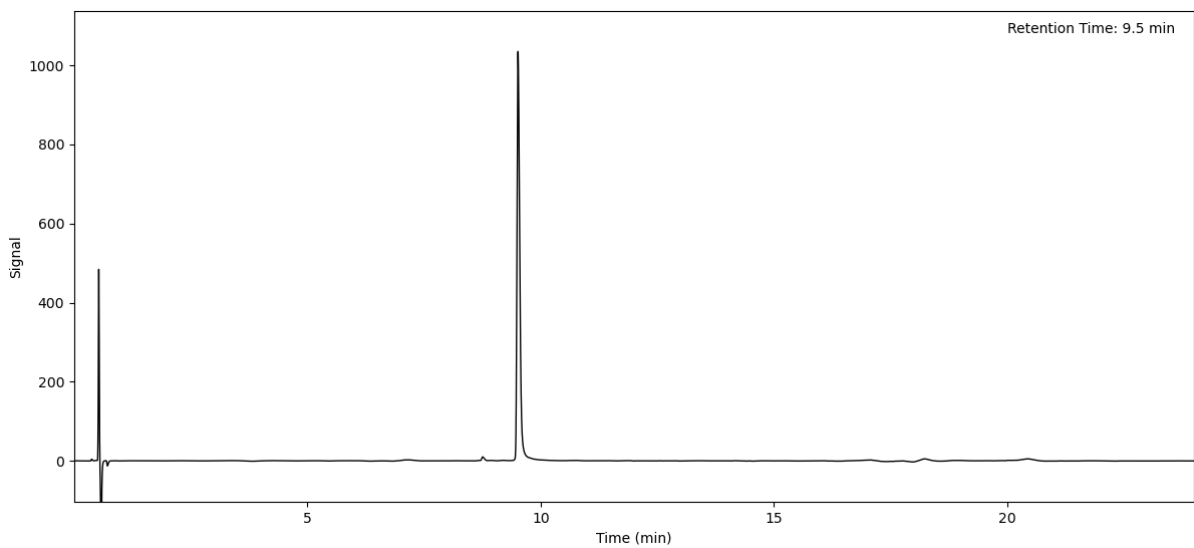
Peptide 85 (Method A)



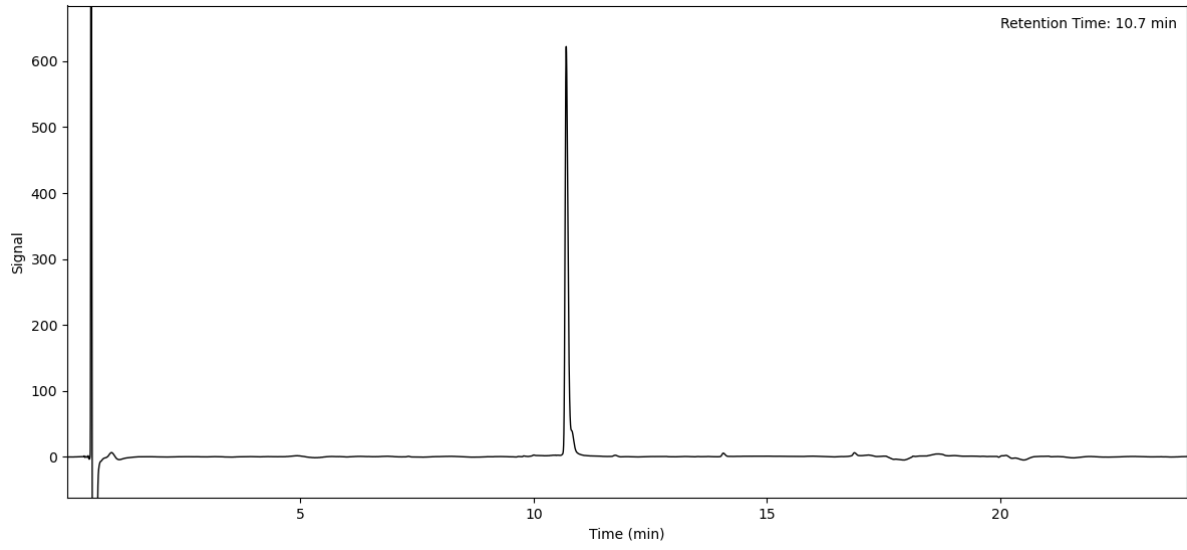
Peptide 86 (Method A)



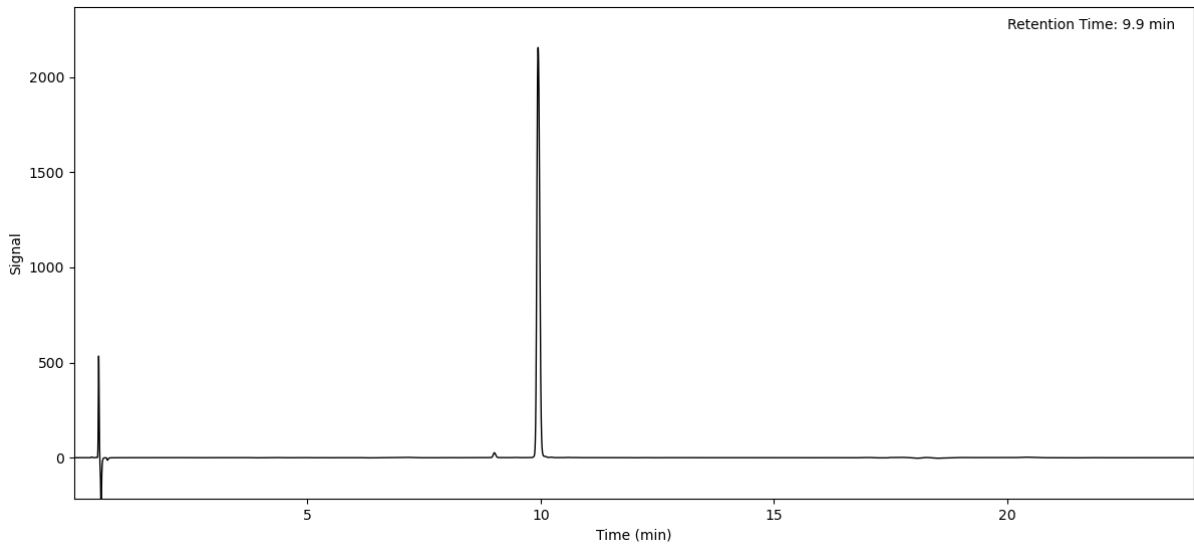
Peptide 87 (Method A)



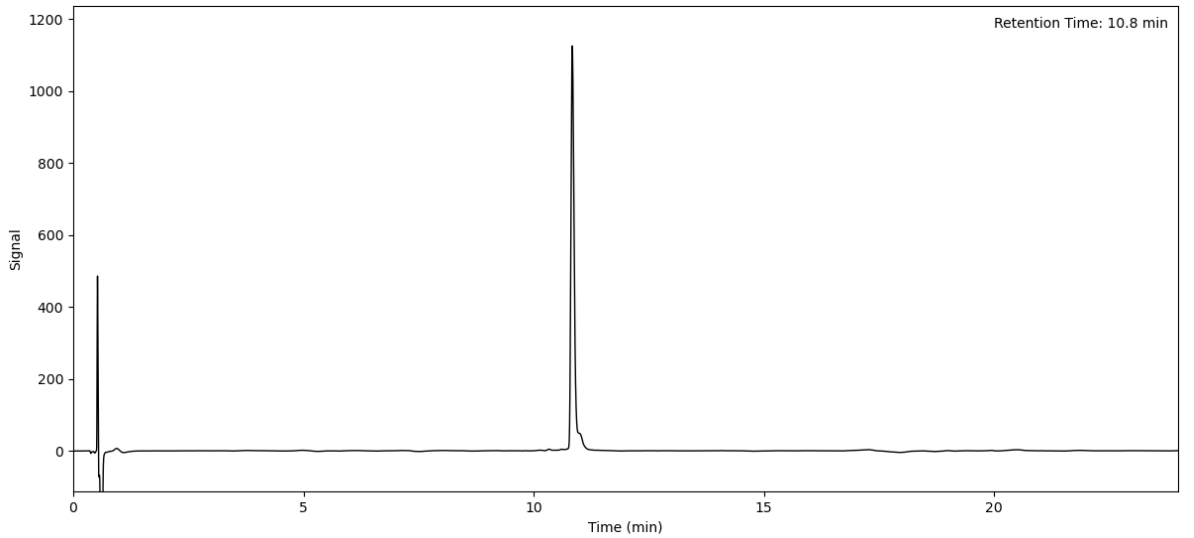
Peptide 88 (Method A)



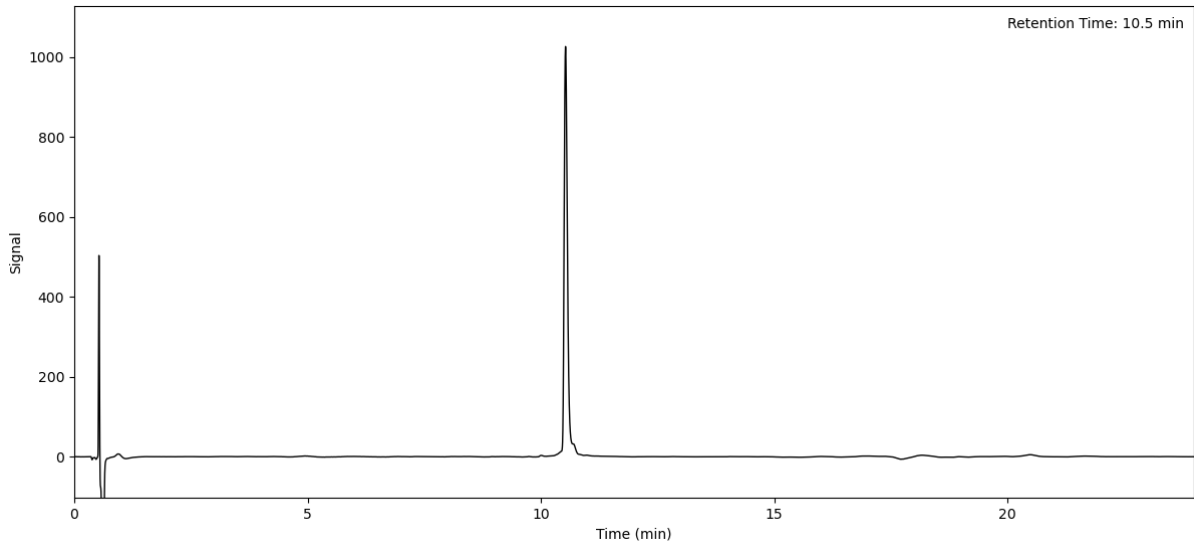
Peptide 89 (Method A)



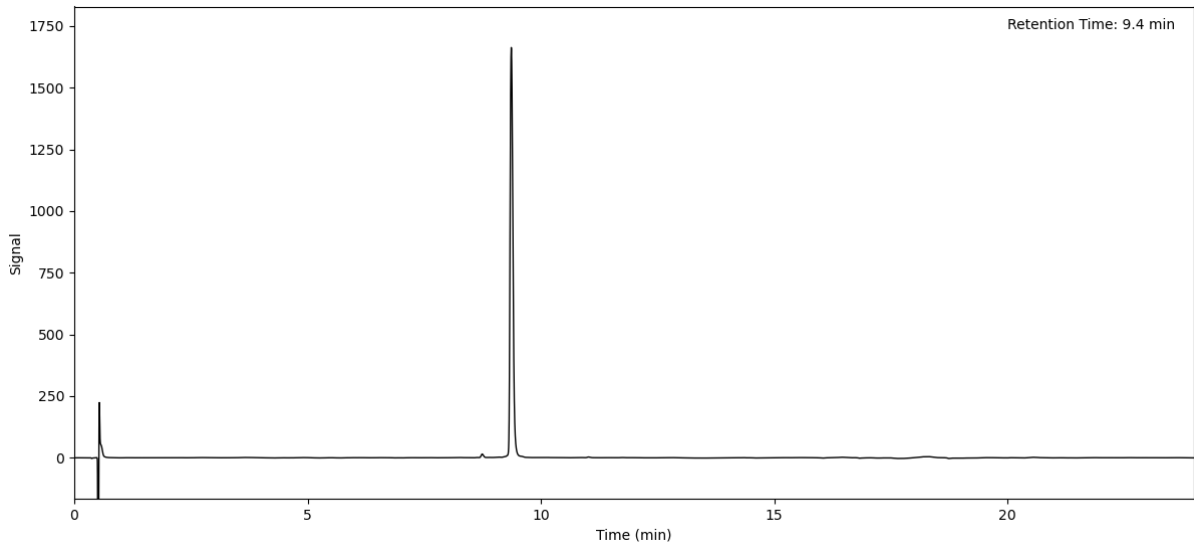
Peptide 90 (Method A)



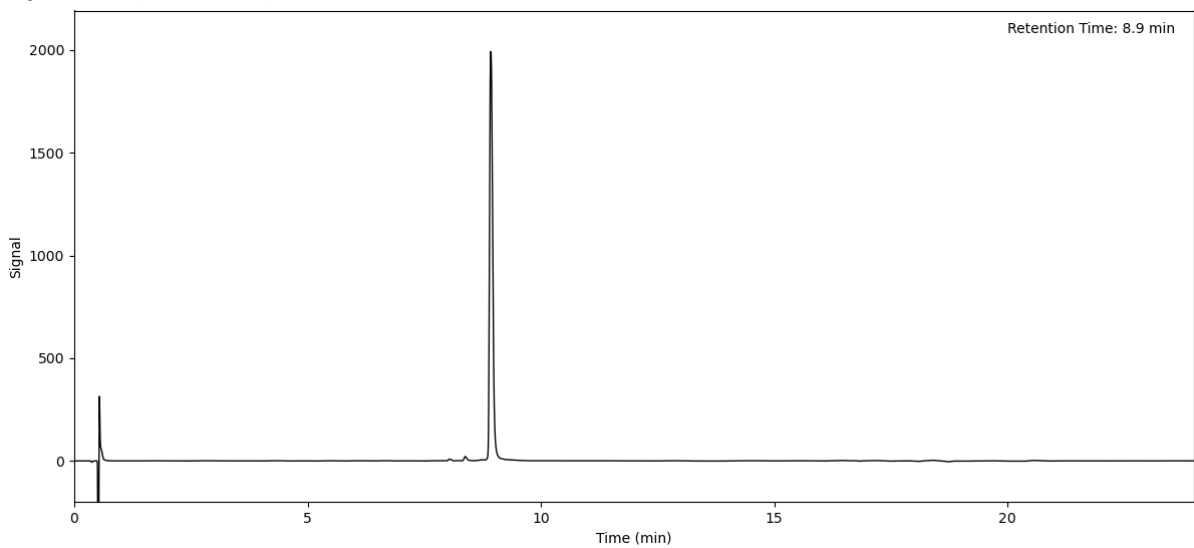
Peptide 91 (Method A)



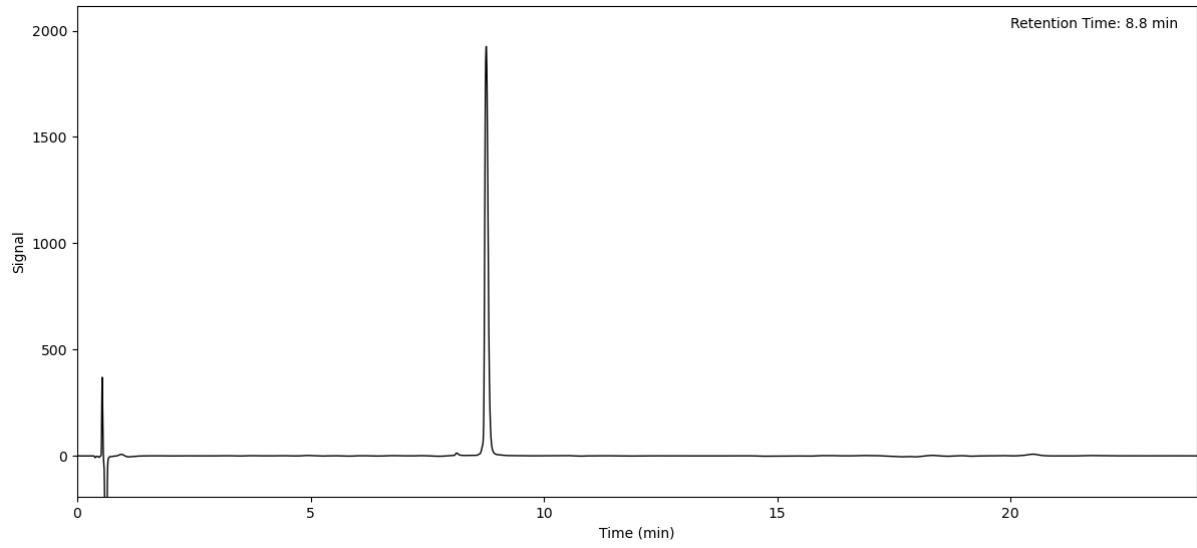
Peptide 92 (Method A)



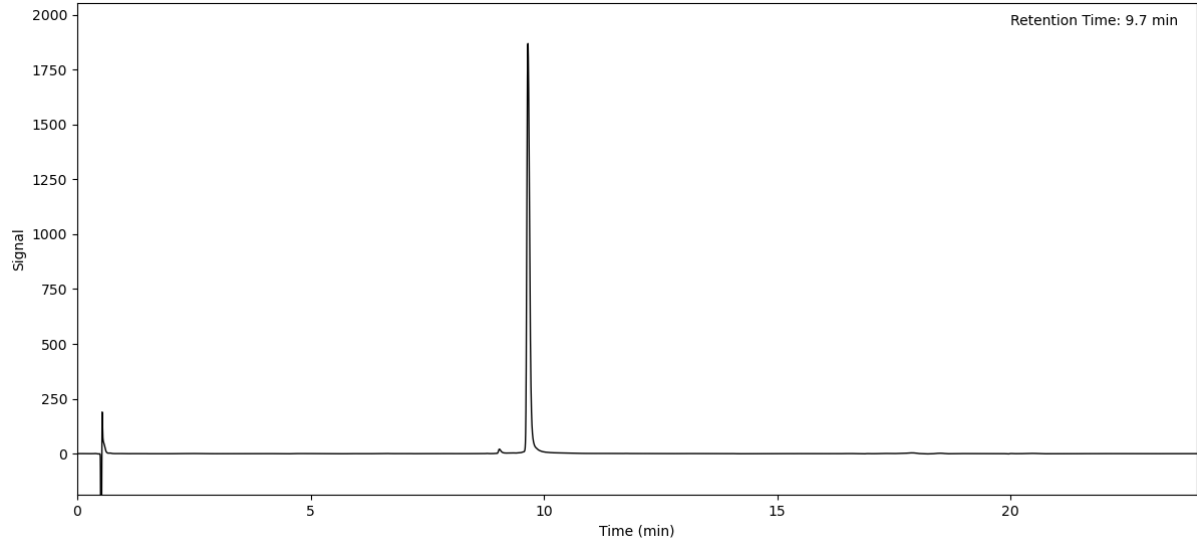
Peptide 93 (Method A)



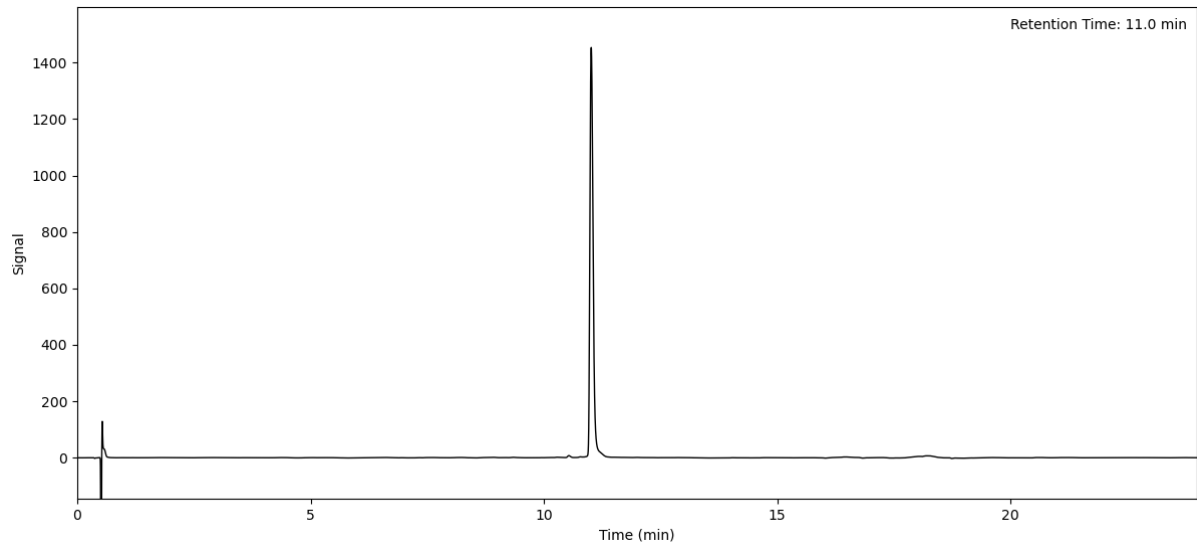
Peptide 94 (Method A)



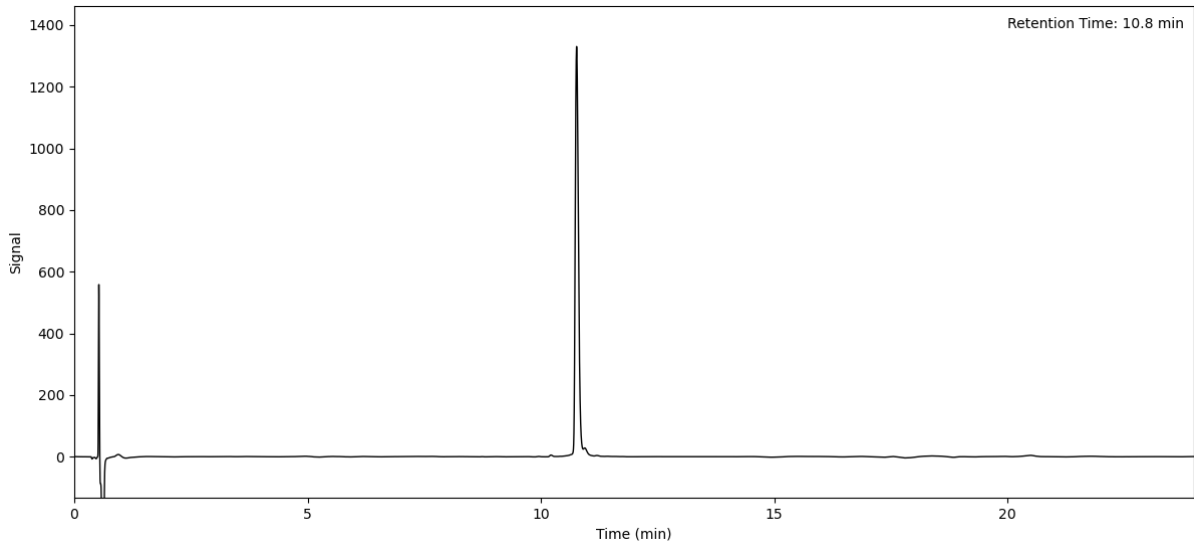
Peptide 95 (Method A)



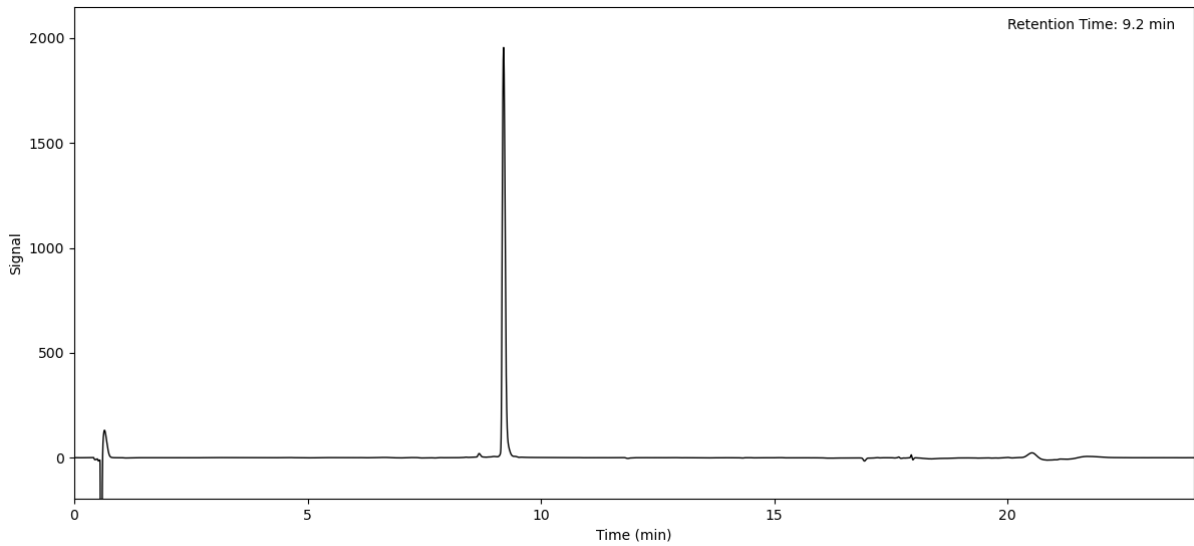
Peptide 96 (Method A)



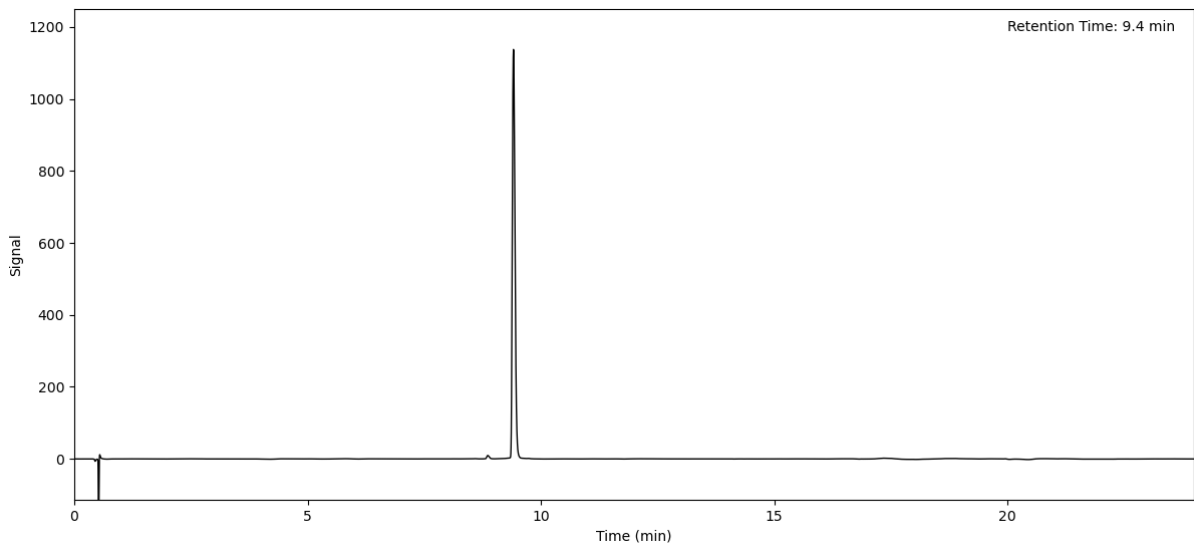
Peptide 97 (Method A)



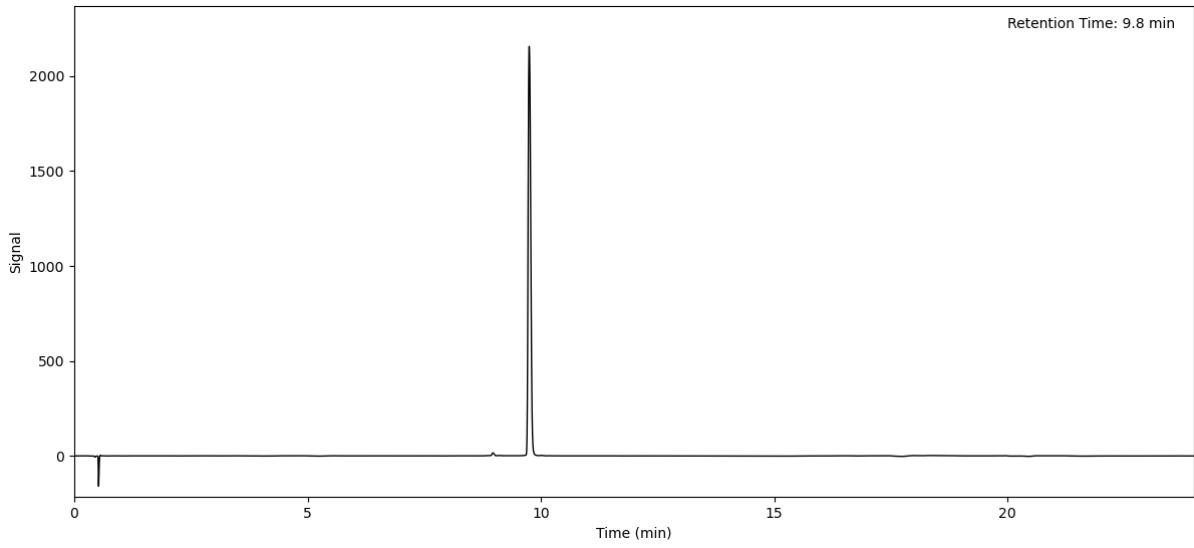
Peptide 98 (Method A)



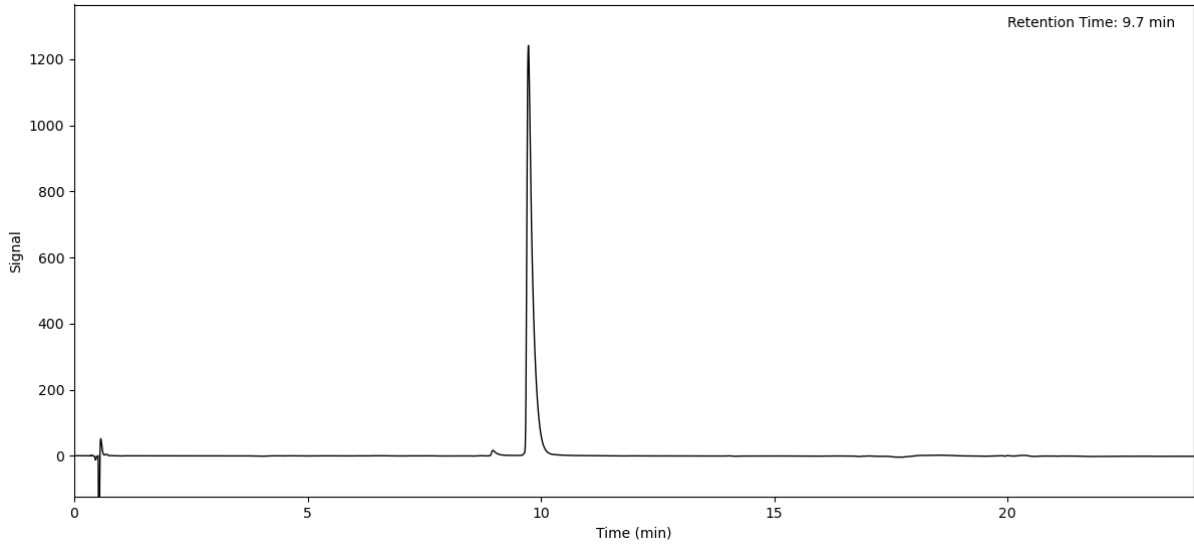
Peptide 99 (Method A)



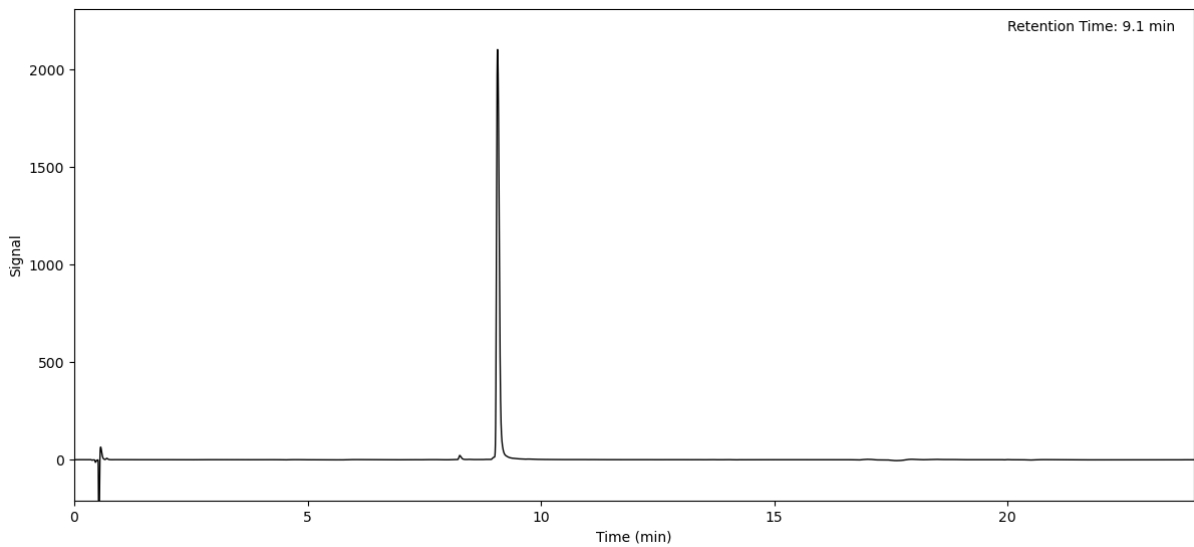
Peptide 100 (Method A)



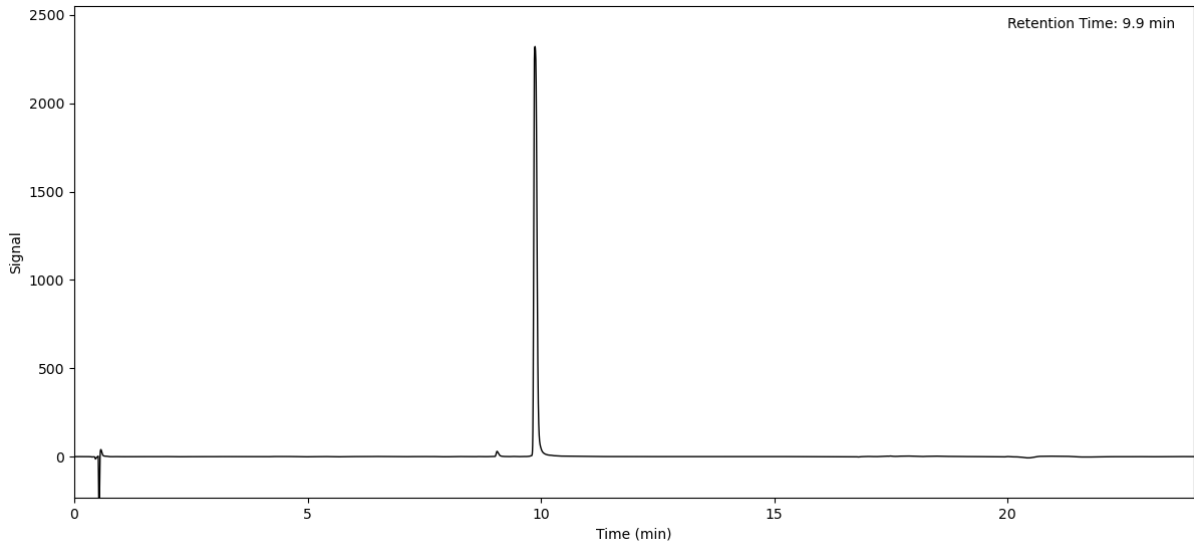
Peptide 101 (Method A)



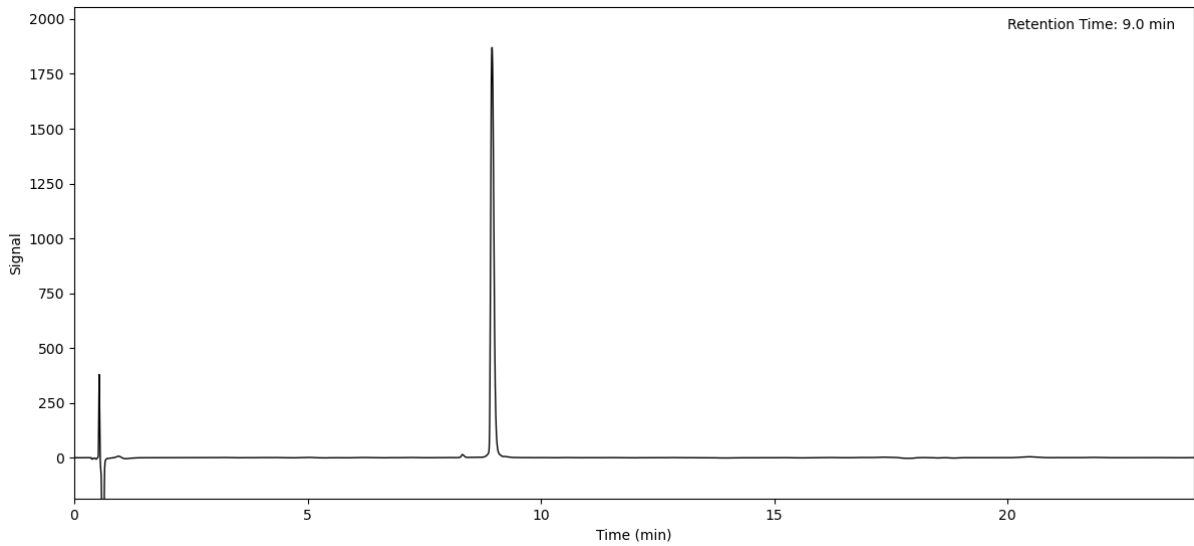
Peptide 102 (Method A)



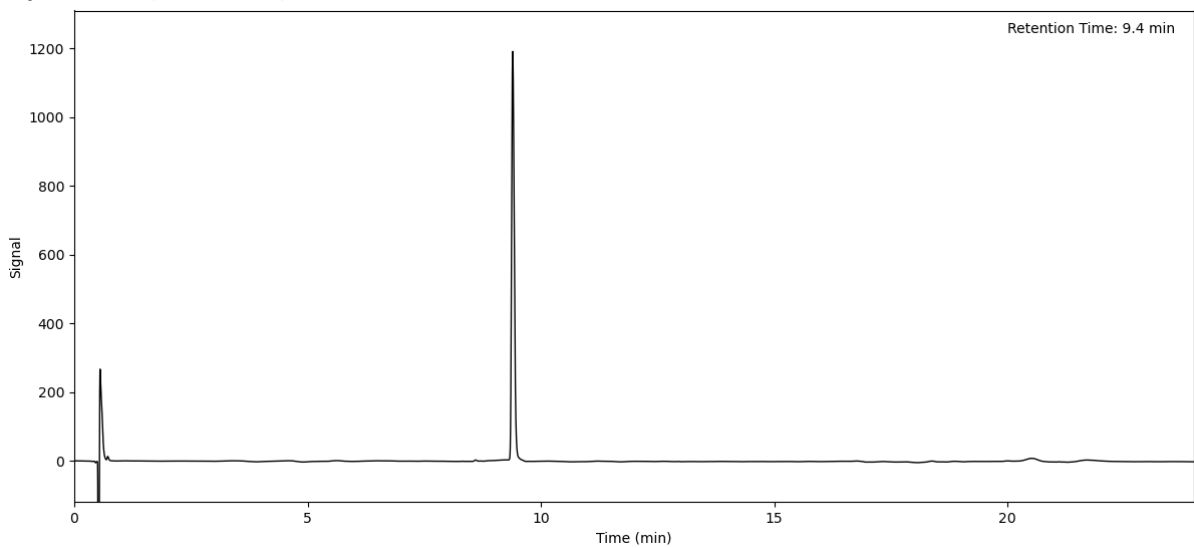
Peptide 103 (Method A)



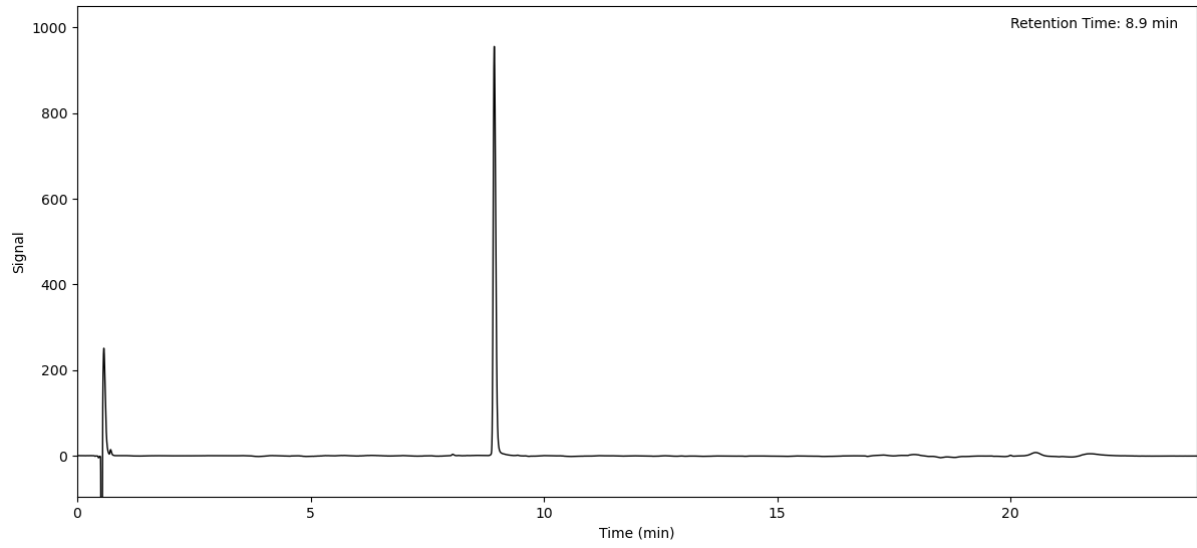
Peptide 104 (Method A)



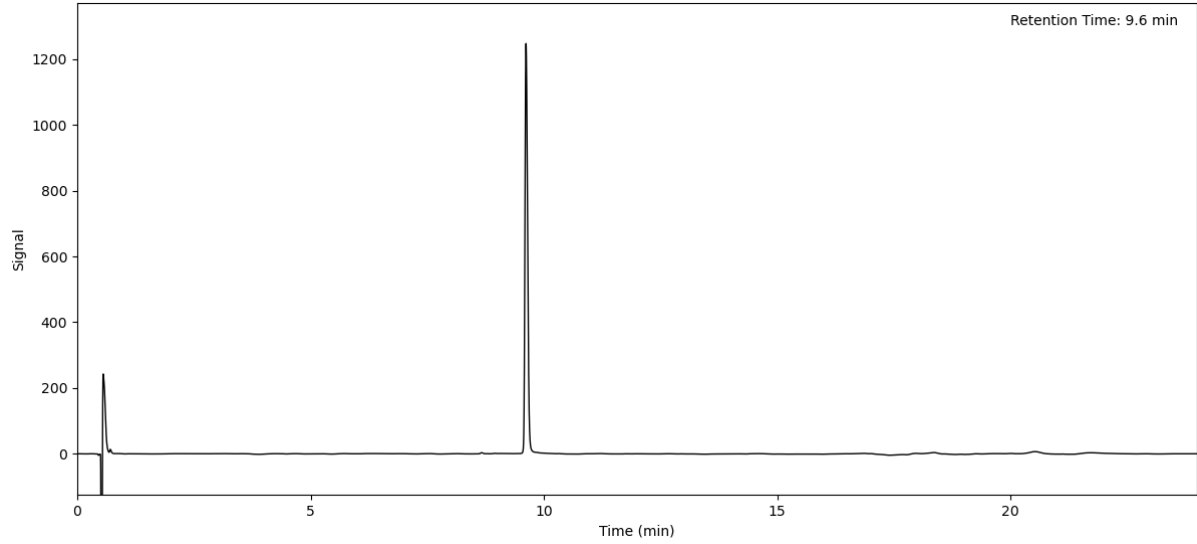
Peptide 105 (Method A)



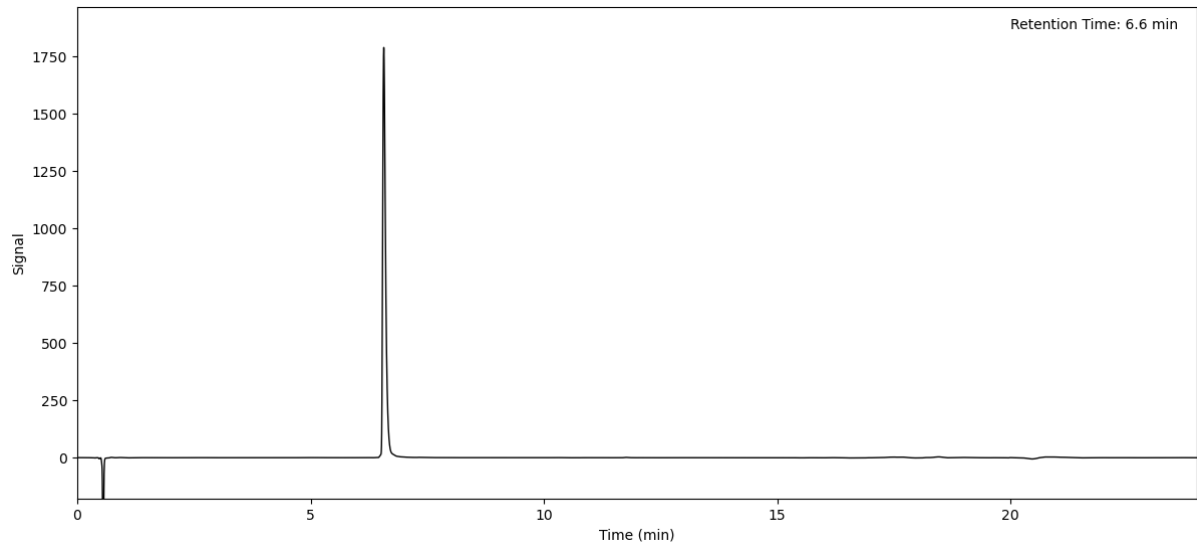
Peptide 106 (Method A)



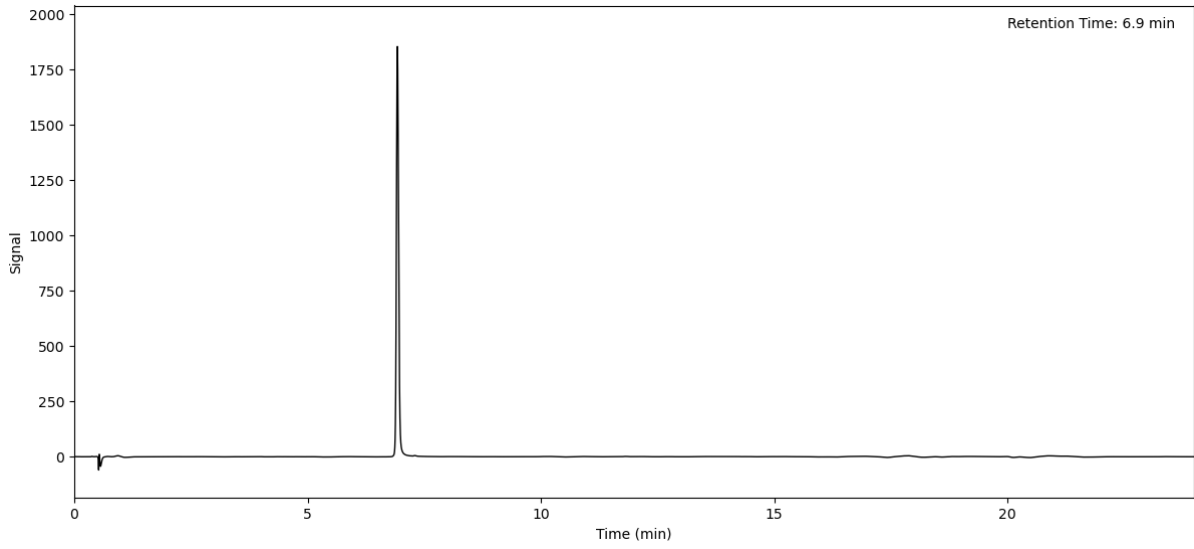
Peptide 107 (Method A)



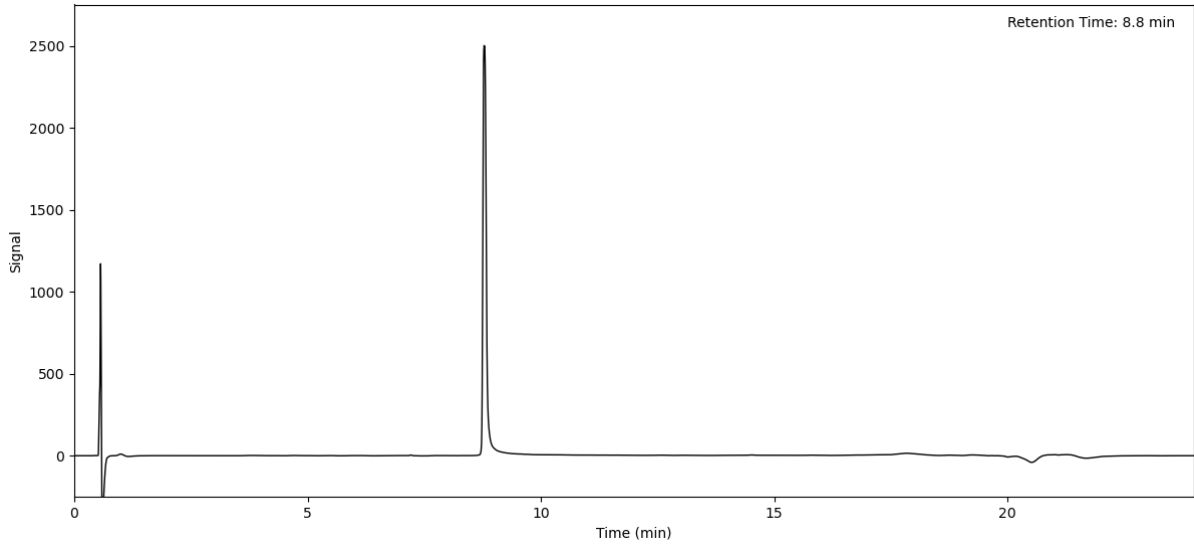
Peptide 108 (Method A)



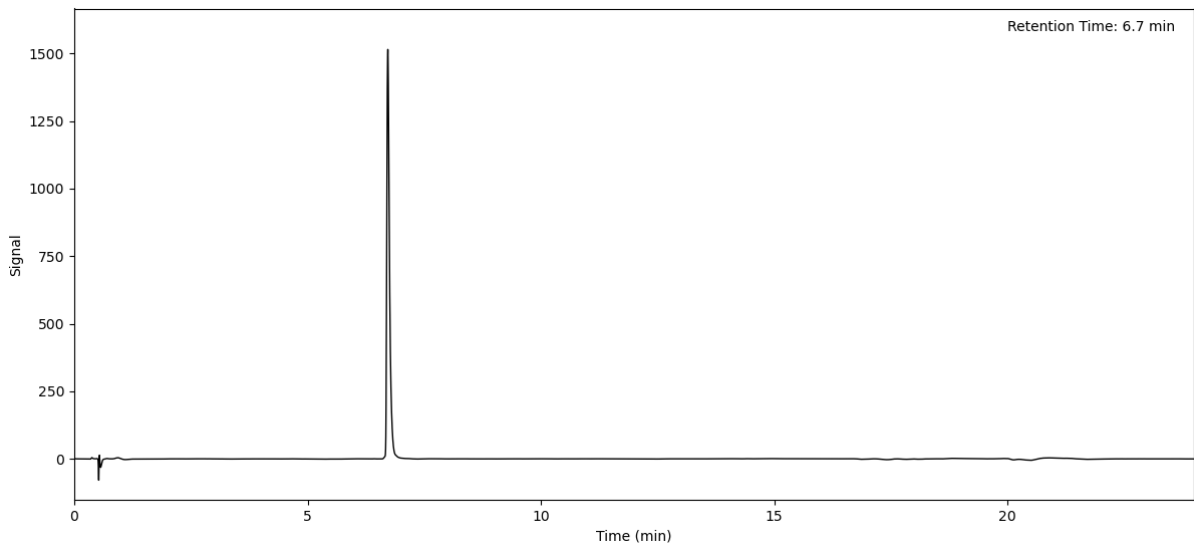
Peptide 109 (Method A)



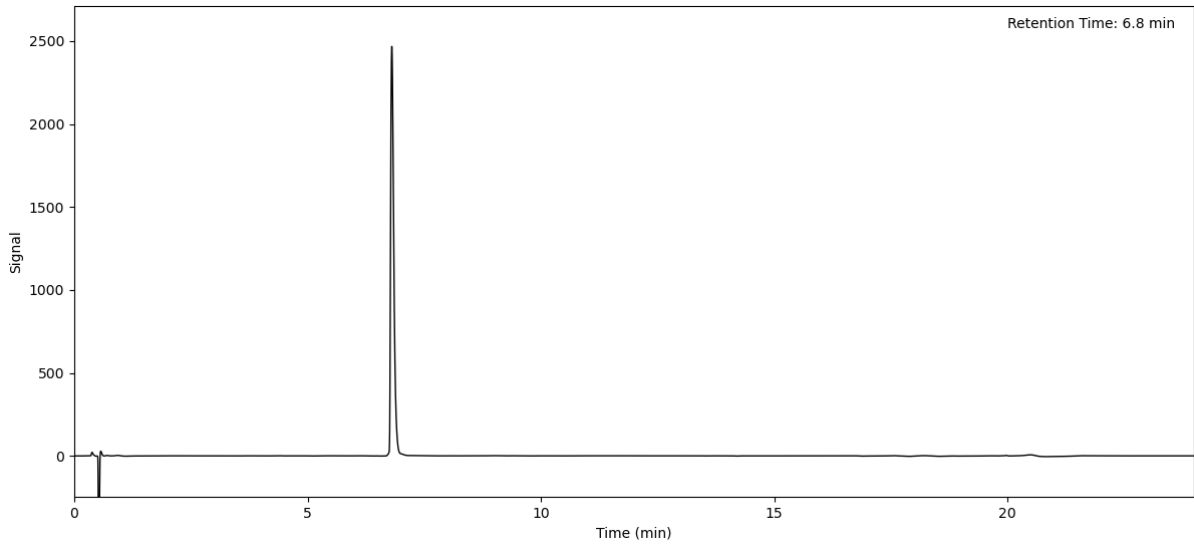
Peptide 110 (Method A)



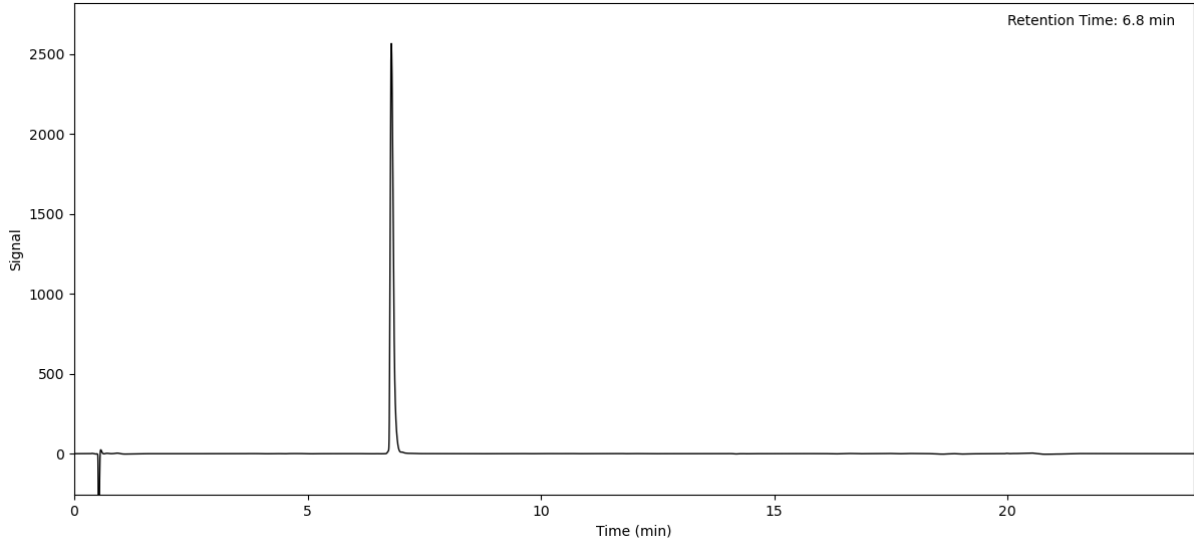
Peptide 111 (Method A)



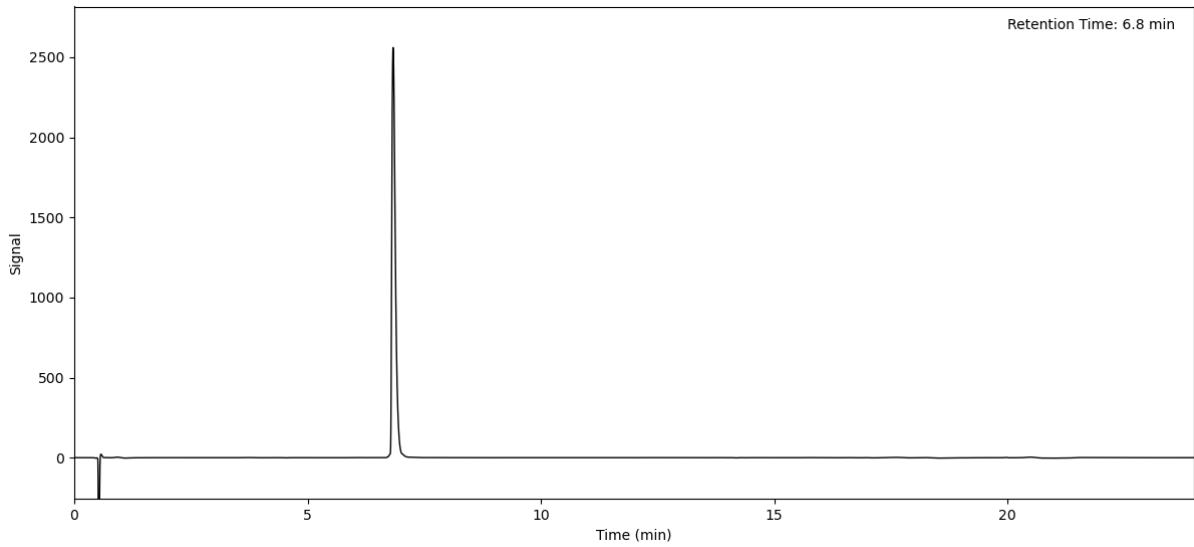
Peptide 112 (Method A)



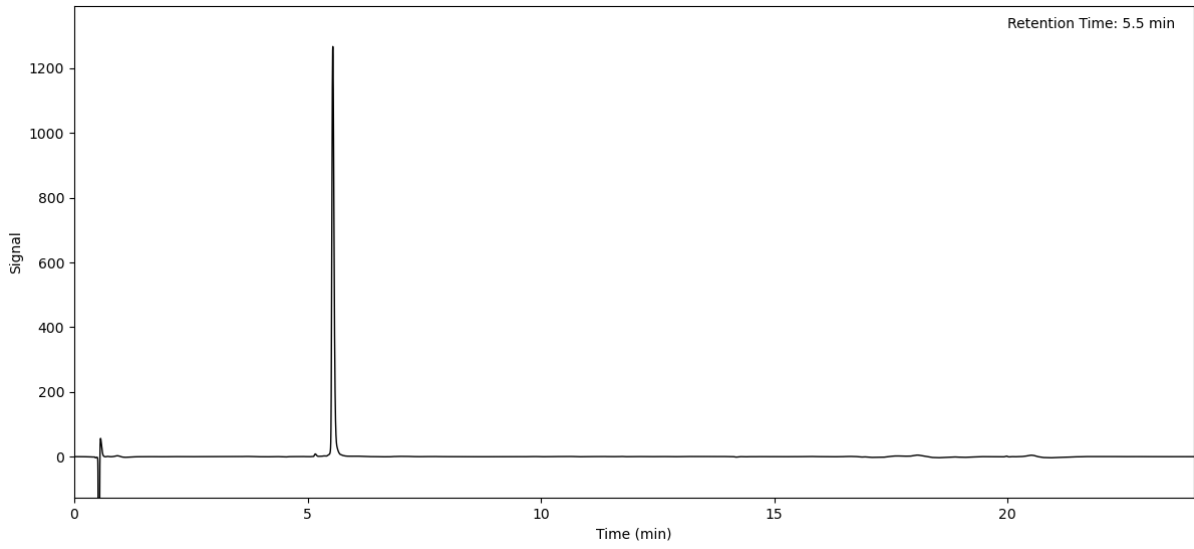
Peptide 113 (Method A)



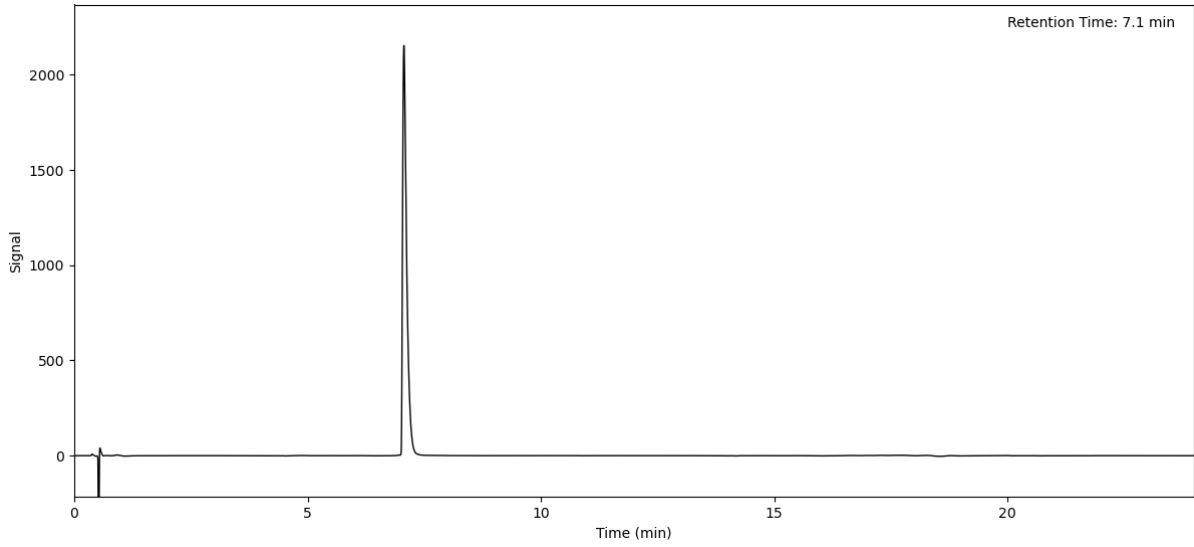
Peptide 114 (Method A)



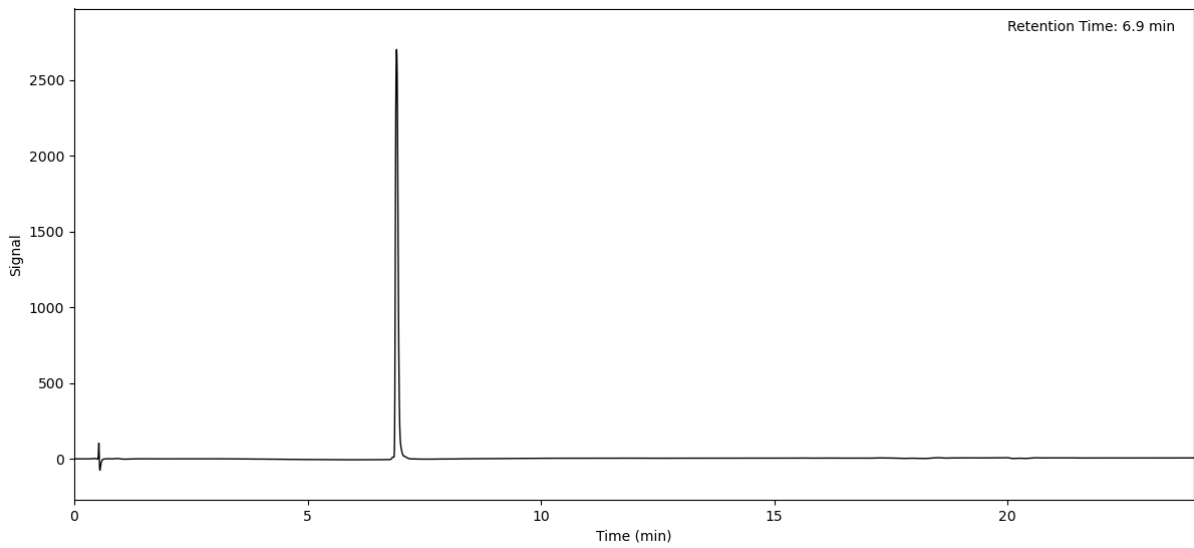
Peptide 115 (Method A)



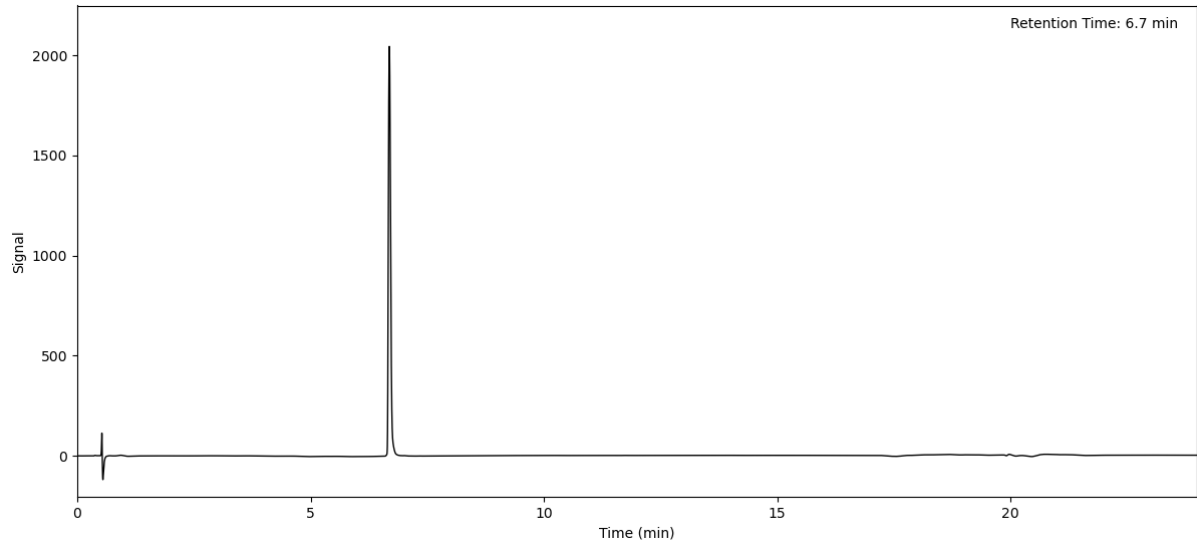
Peptide 116 (Method A)



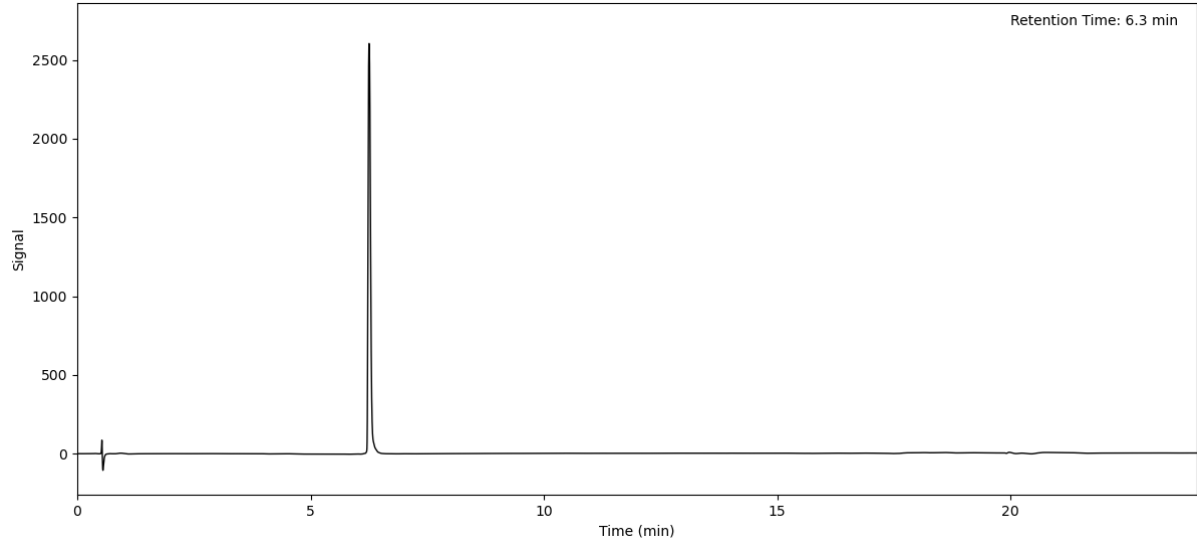
Peptide 117 (Method A)



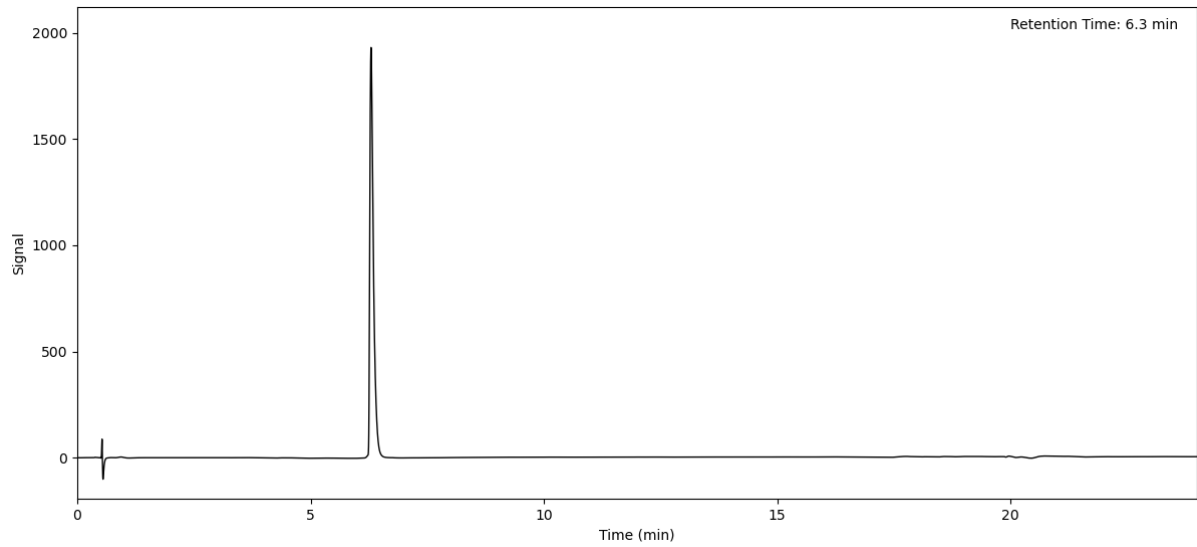
Peptide 118 (Method A)



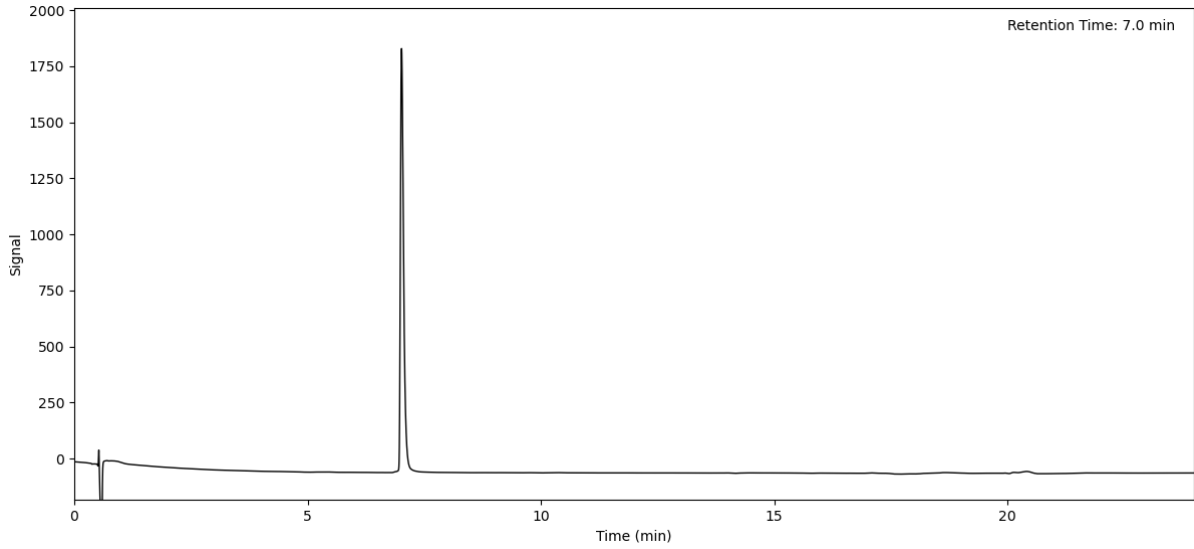
Peptide 119 (Method A)



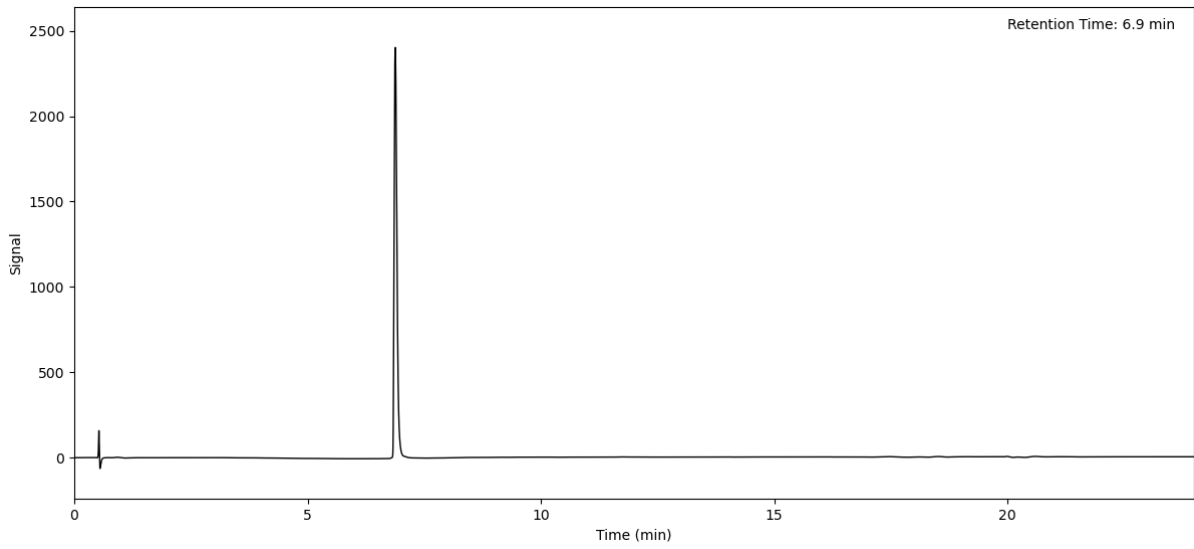
Peptide 120 (Method A)



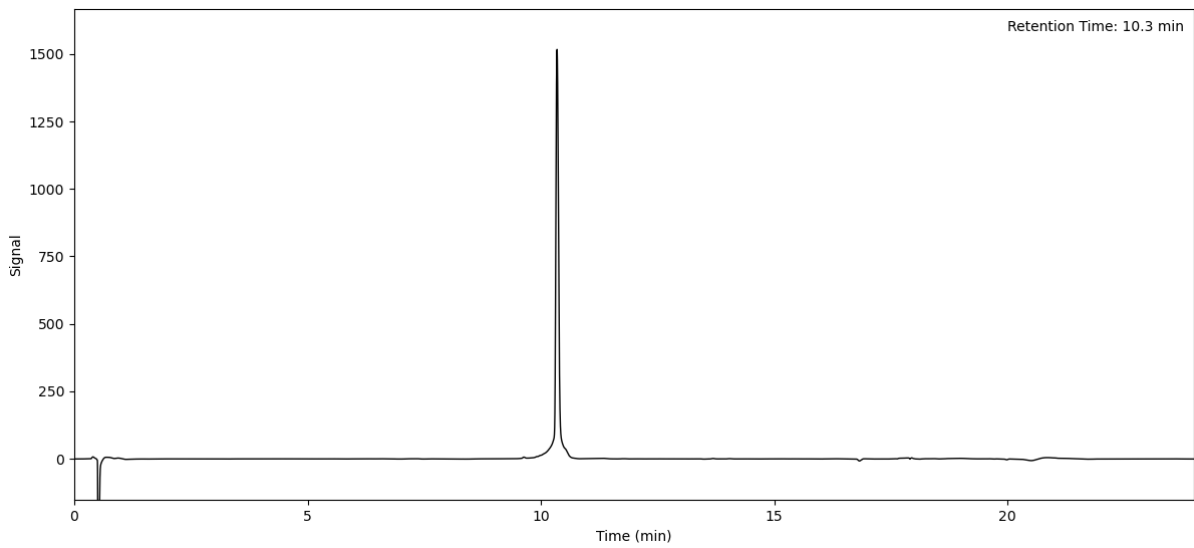
Peptide 121 (Method A)



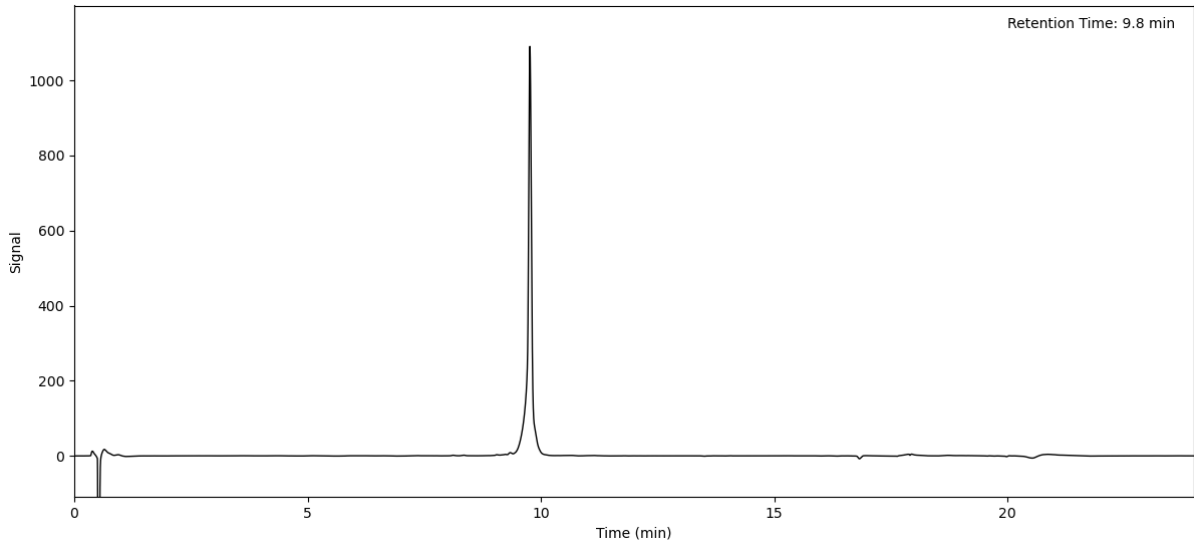
Peptide 122 (Method A)



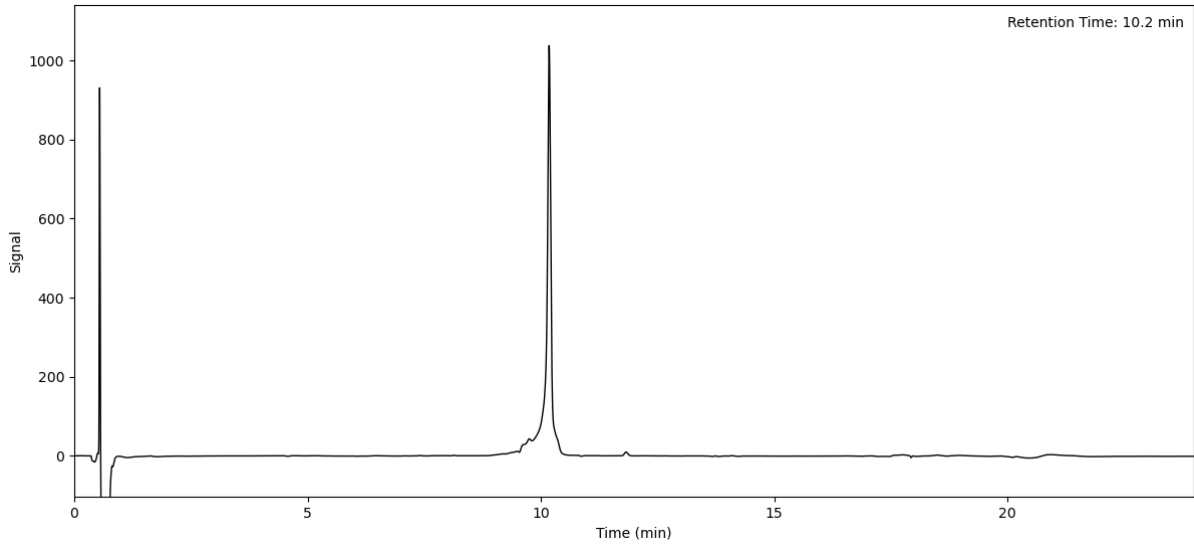
Peptide 123 (Method A)



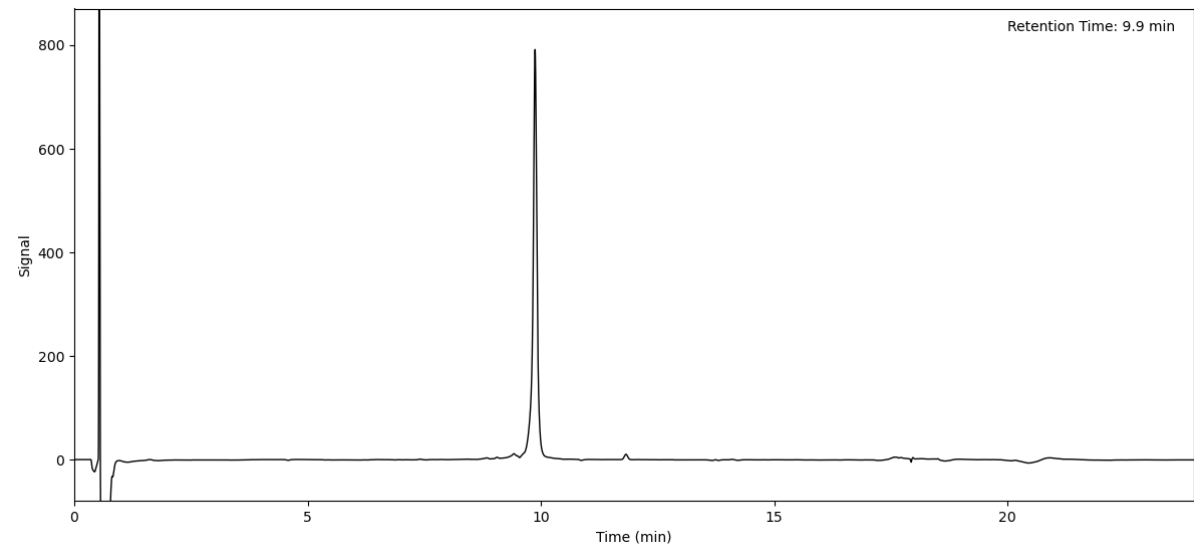
Peptide 124 (Method A)



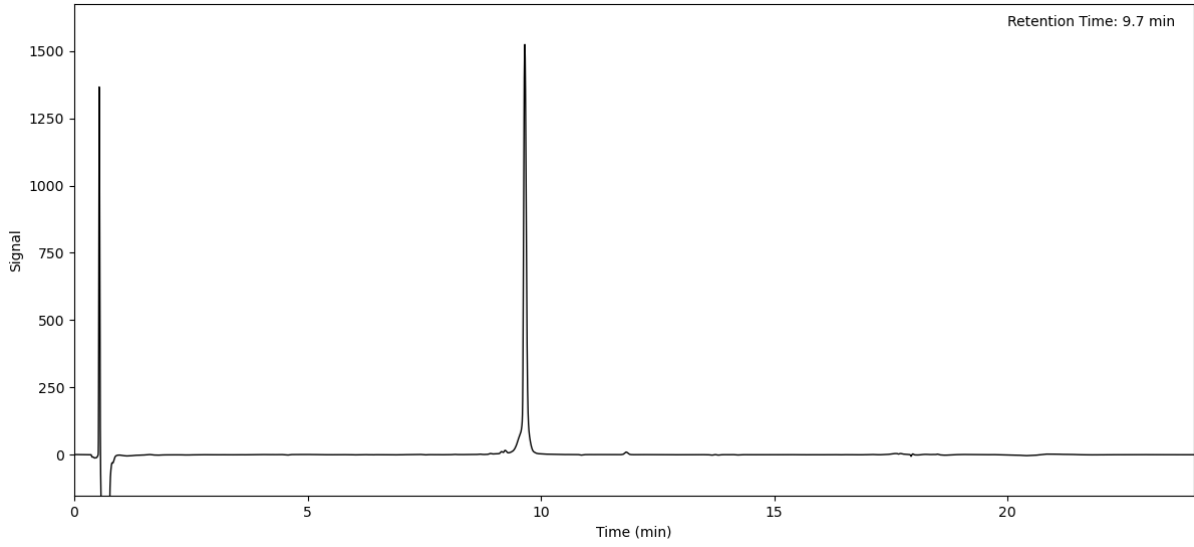
Peptide 125 (Method A)



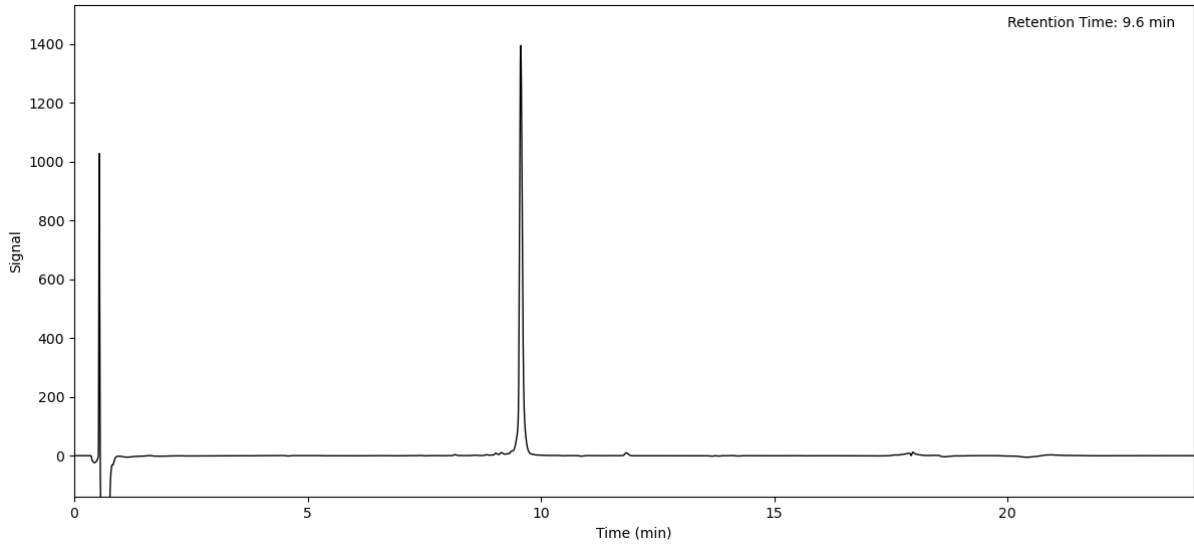
Peptide 126 (Method A)



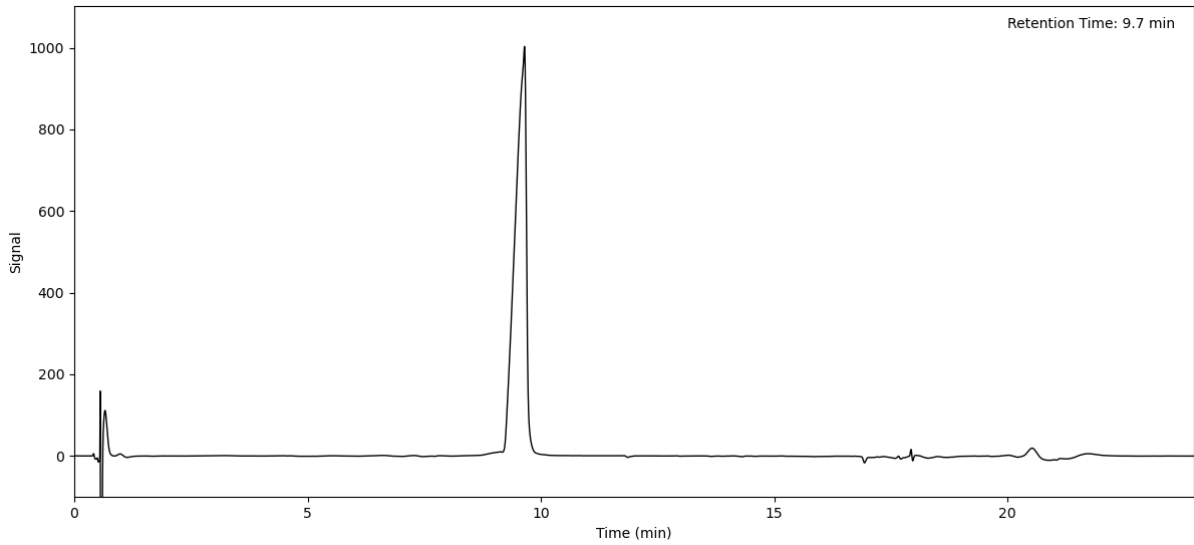
Peptide 127 (Method A)



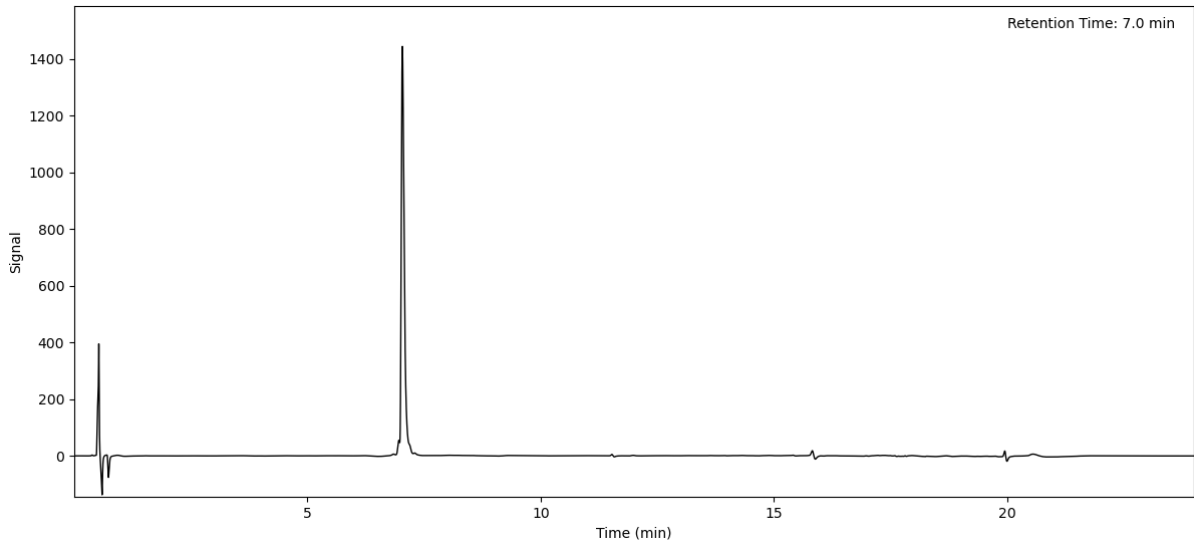
Peptide 128 (Method A)



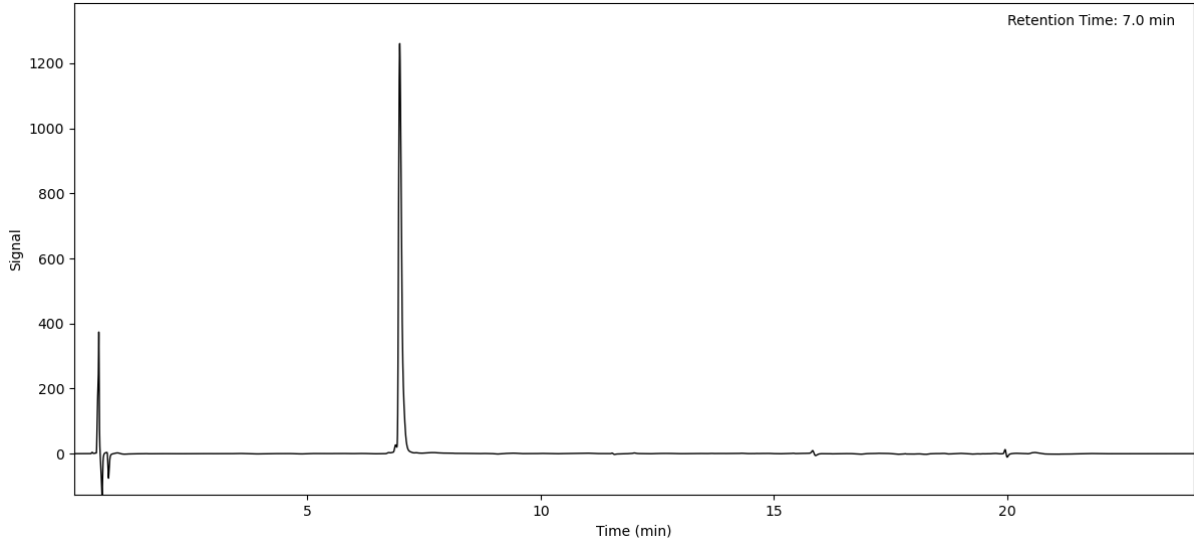
Peptide 129 (Method A)



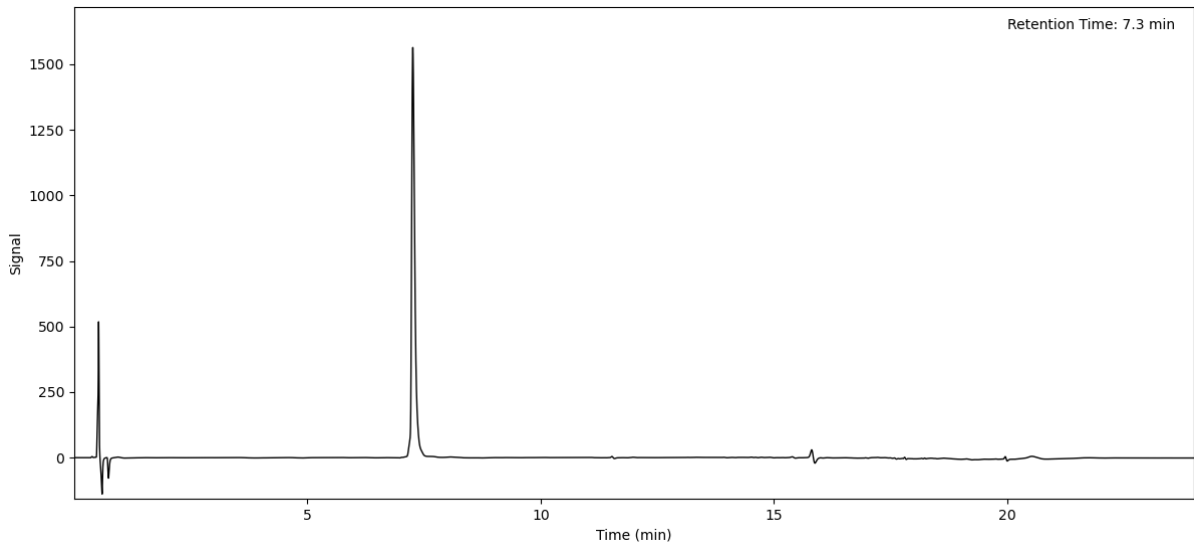
Peptide 130 (Method A)



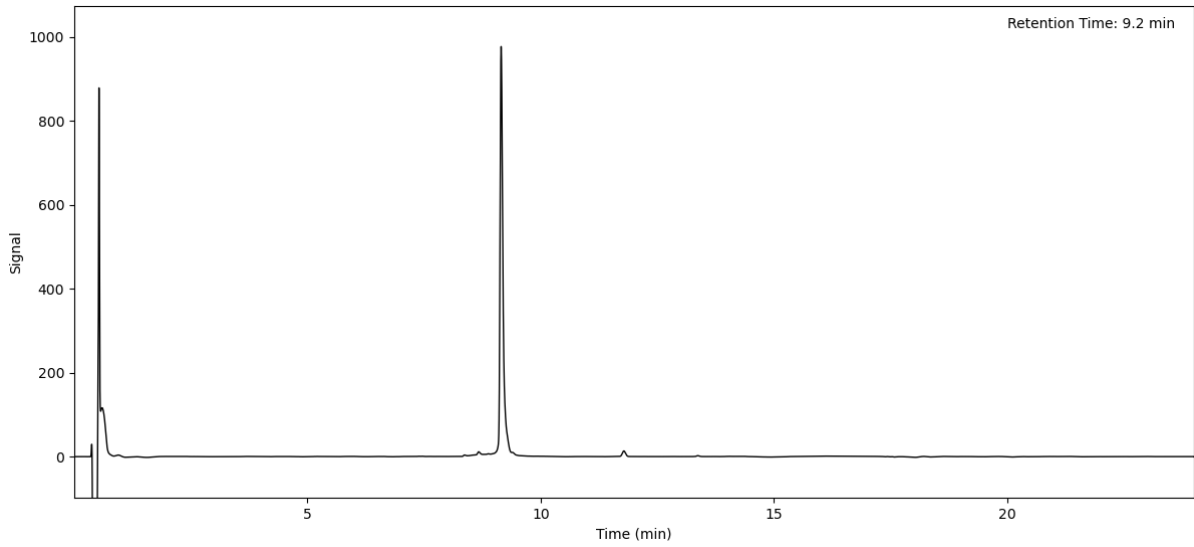
Peptide 131 (Method A)



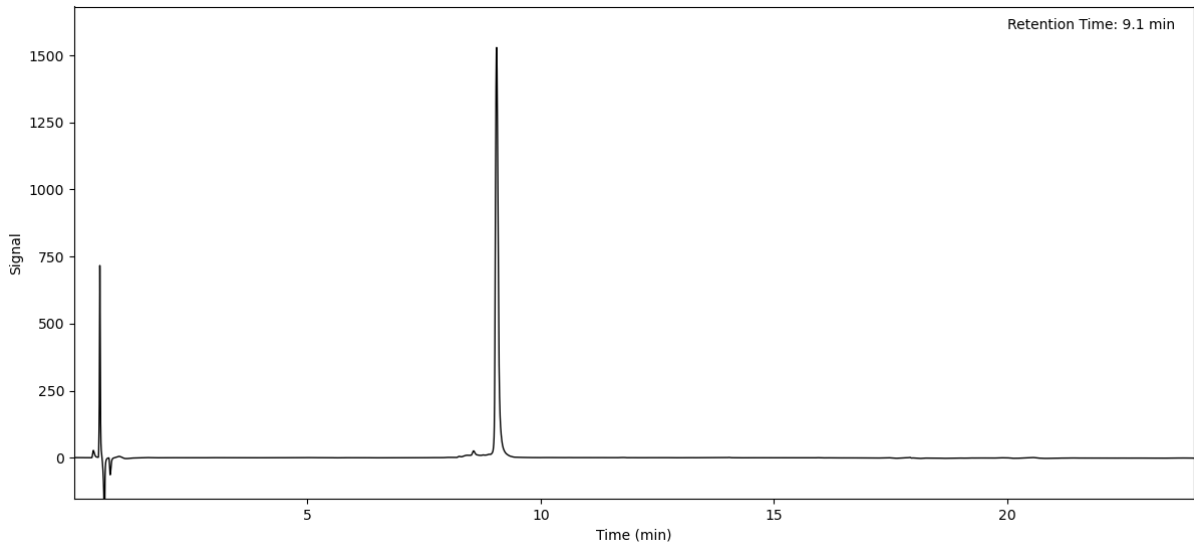
Peptide 132 (Method A)



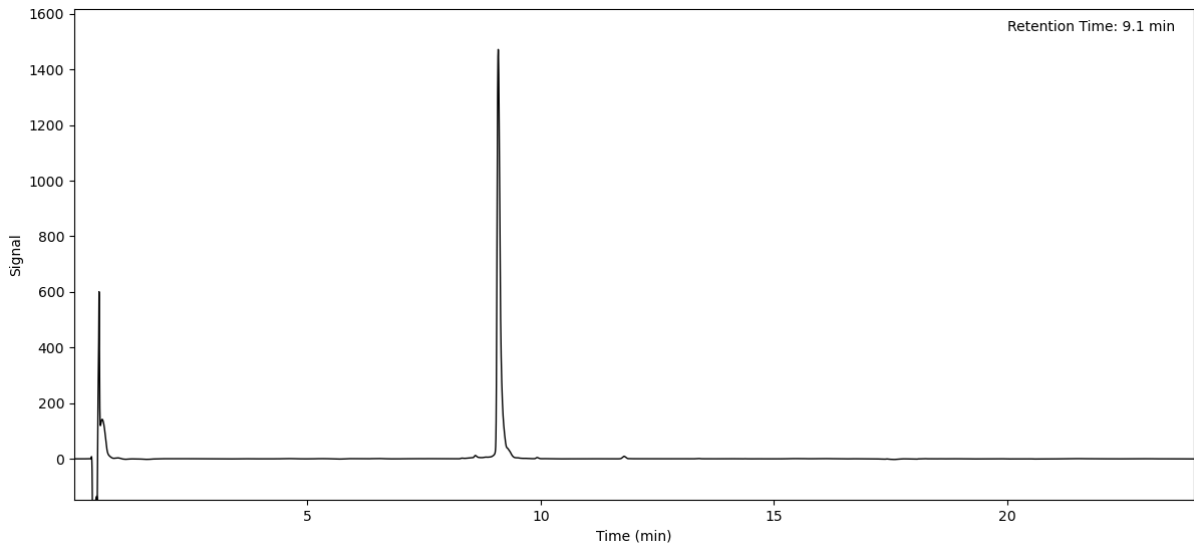
Peptide 133 (Method A)



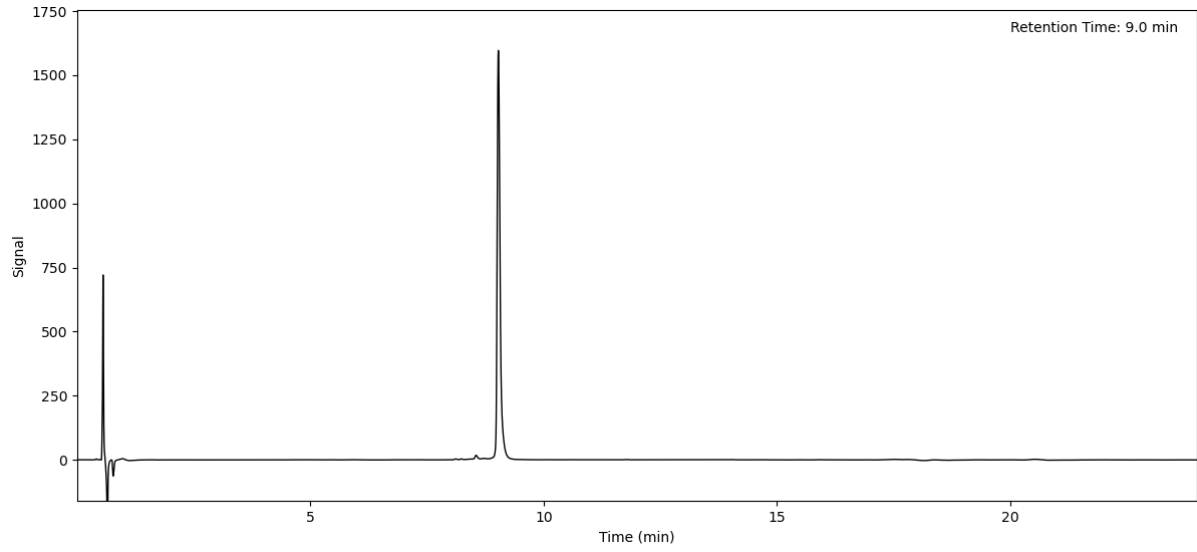
Peptide 134 (Method A)



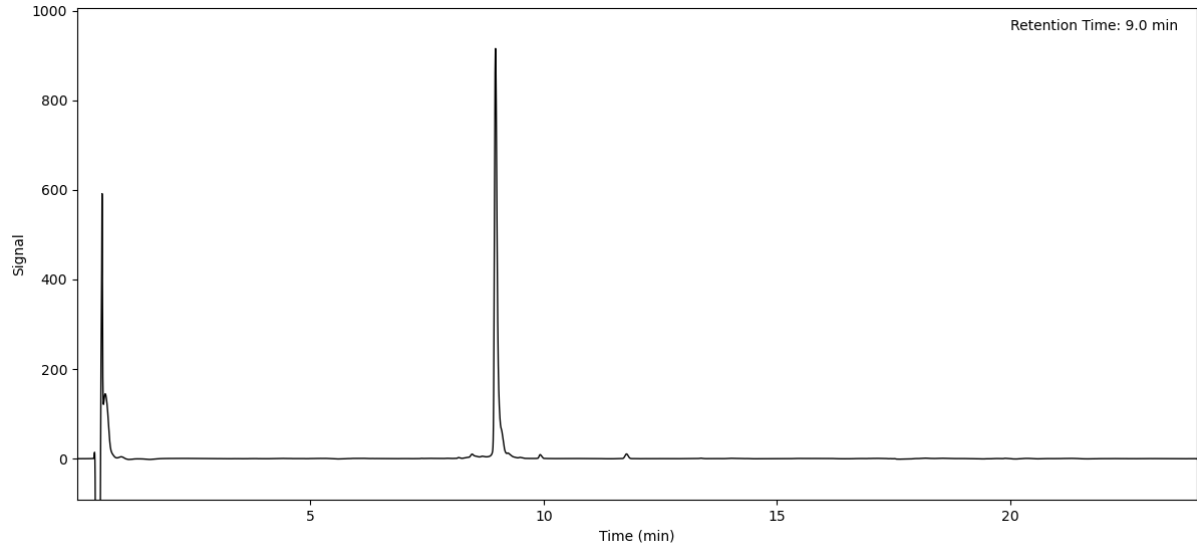
Peptide 135 (Method A)



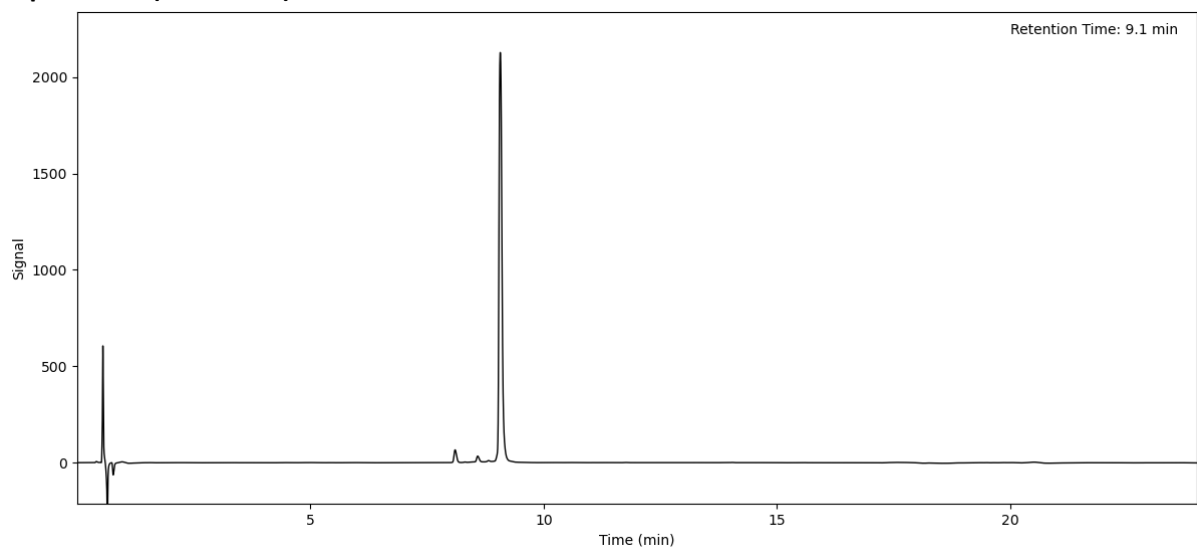
Peptide 136 (Method A)



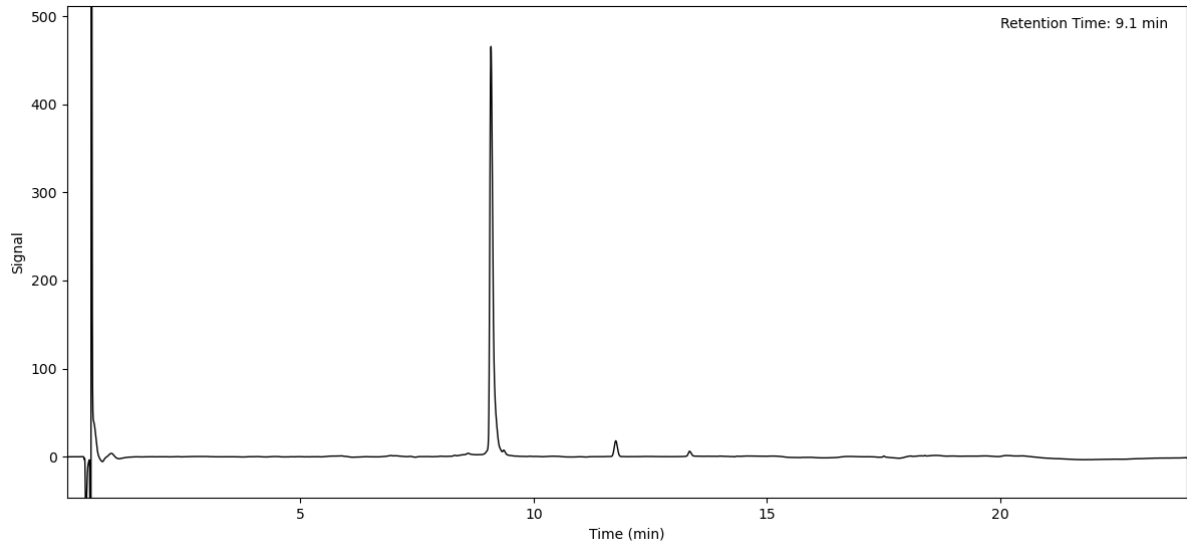
Peptide 137 (Method A)



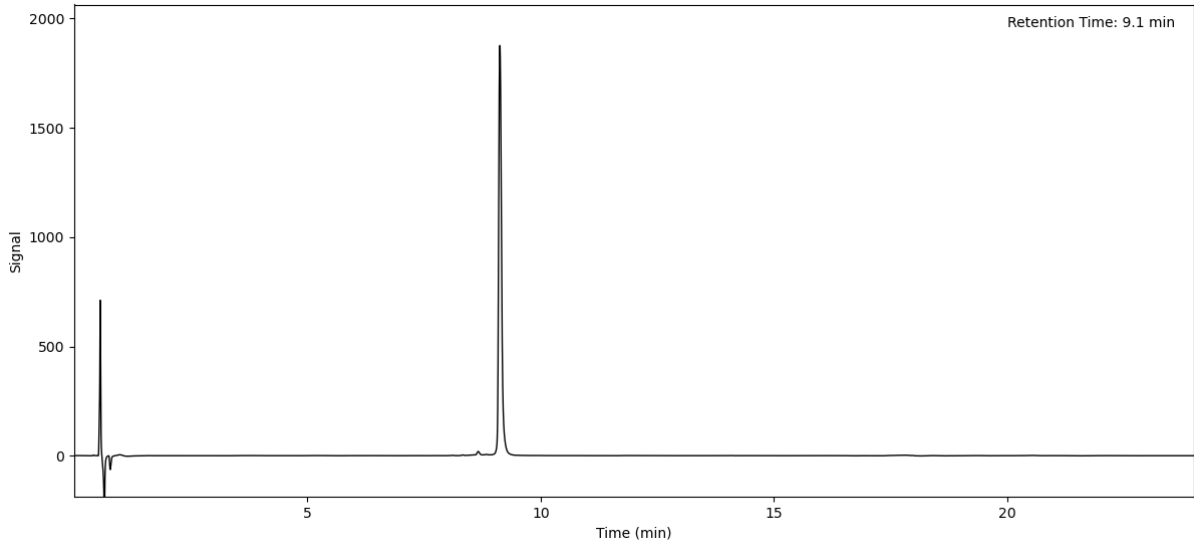
Peptide 138 (Method A)



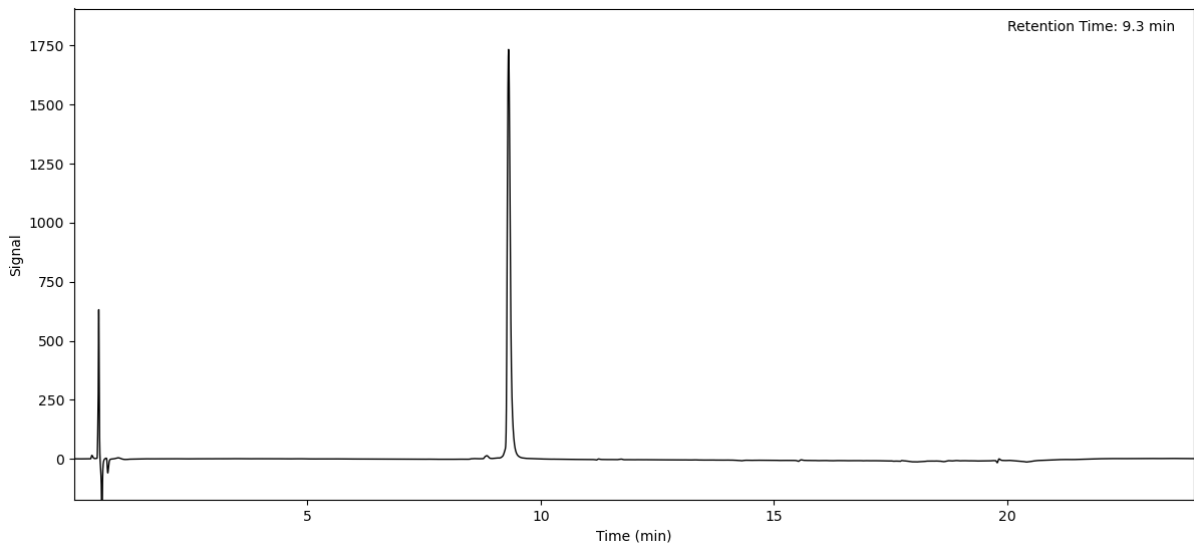
Peptide 139 (Method A)



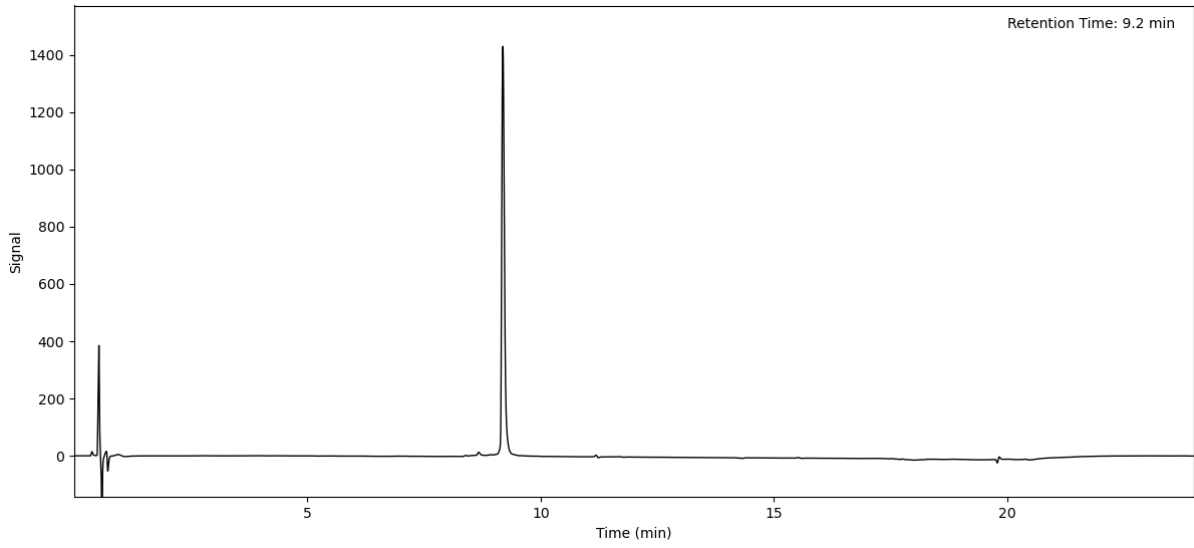
Peptide 140 (Method A)



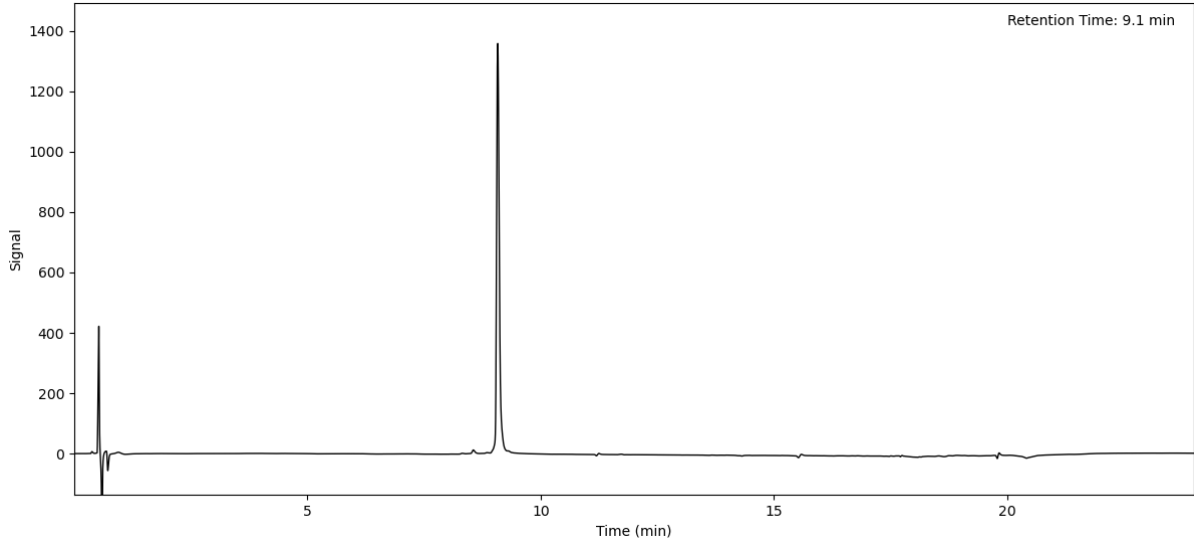
Peptide 141 (Method A)



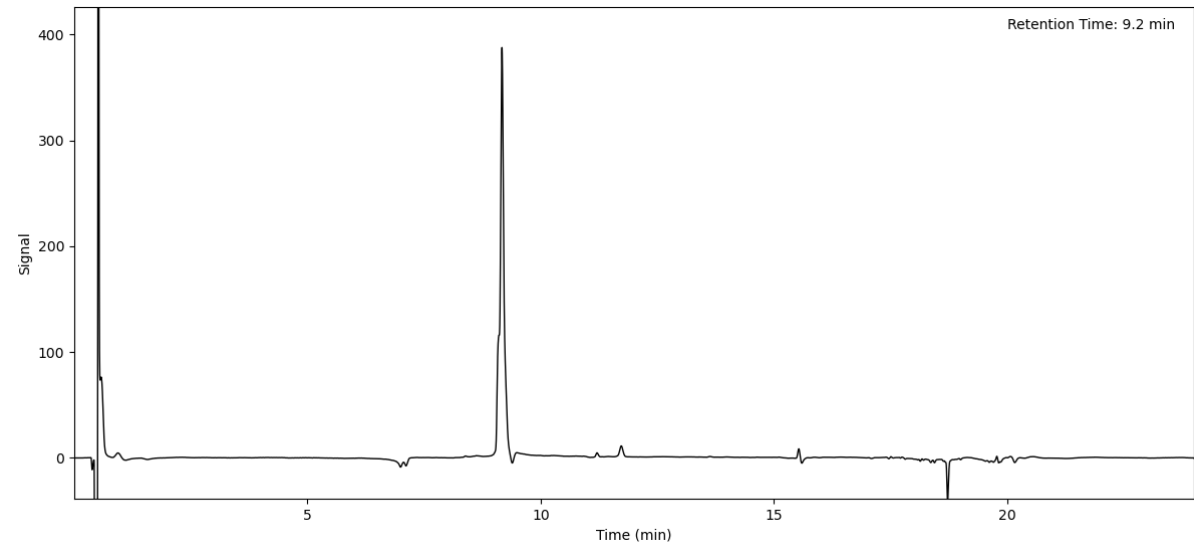
Peptide 142 (Method A)



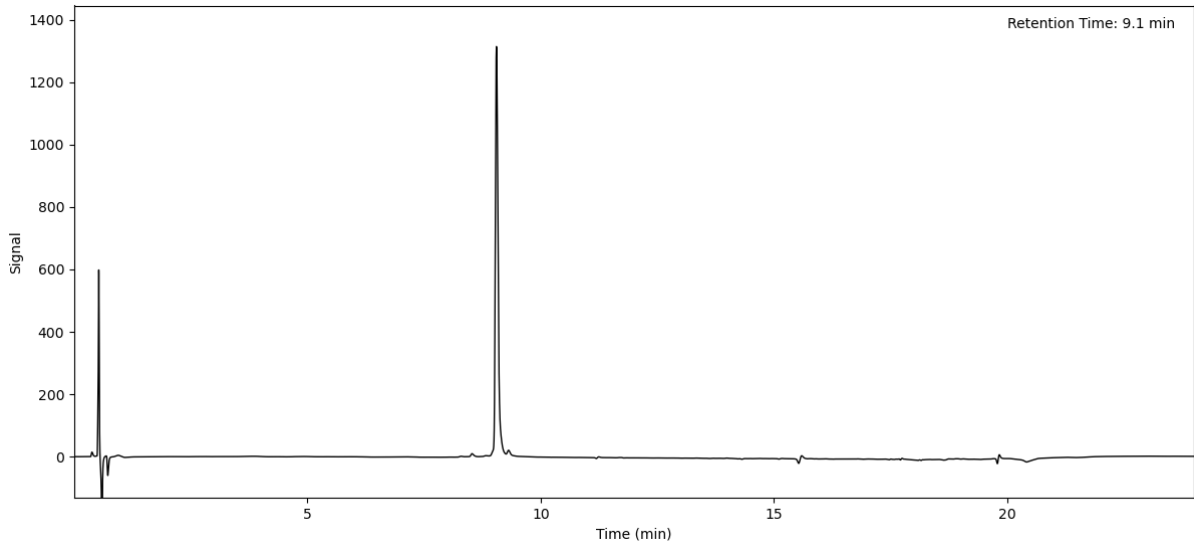
Peptide 143 isomer 1 (Method A)



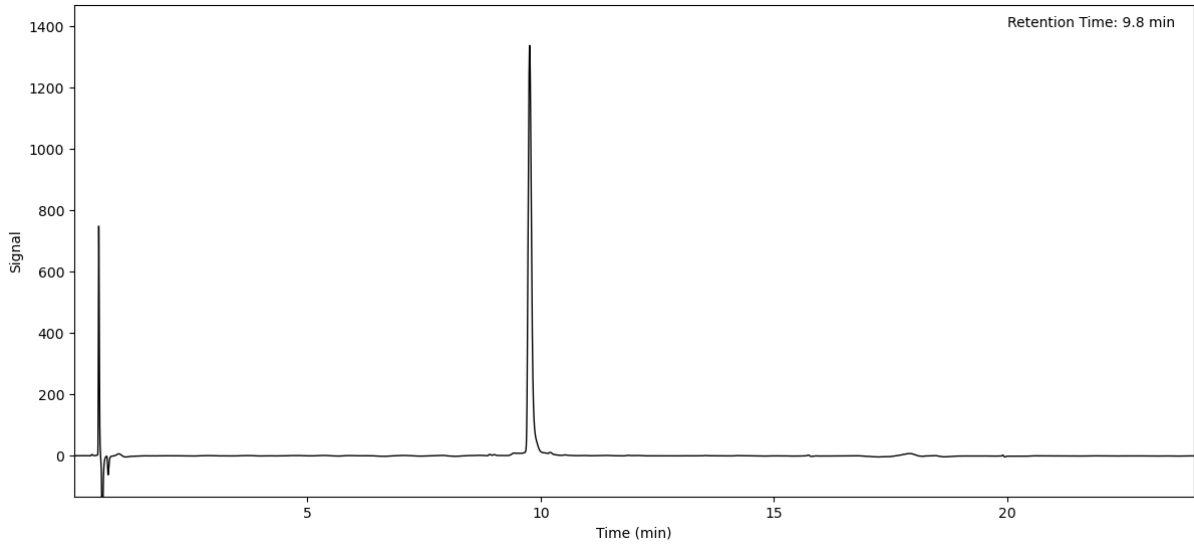
Peptide 143 isomer 2 (Method A)



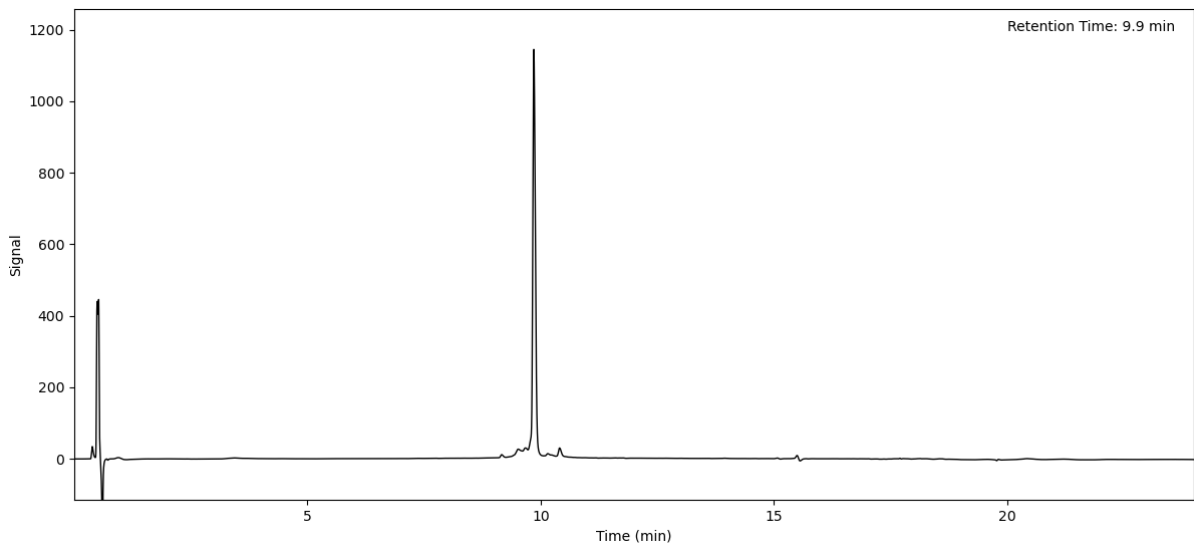
Peptide 144 (Method A)



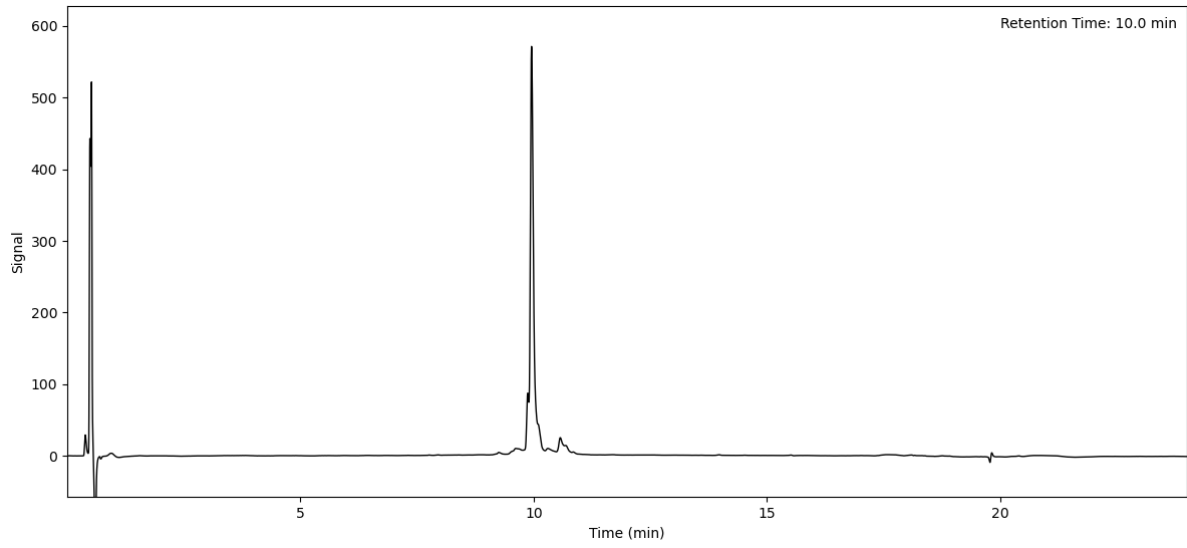
Peptide 145 (Method A)



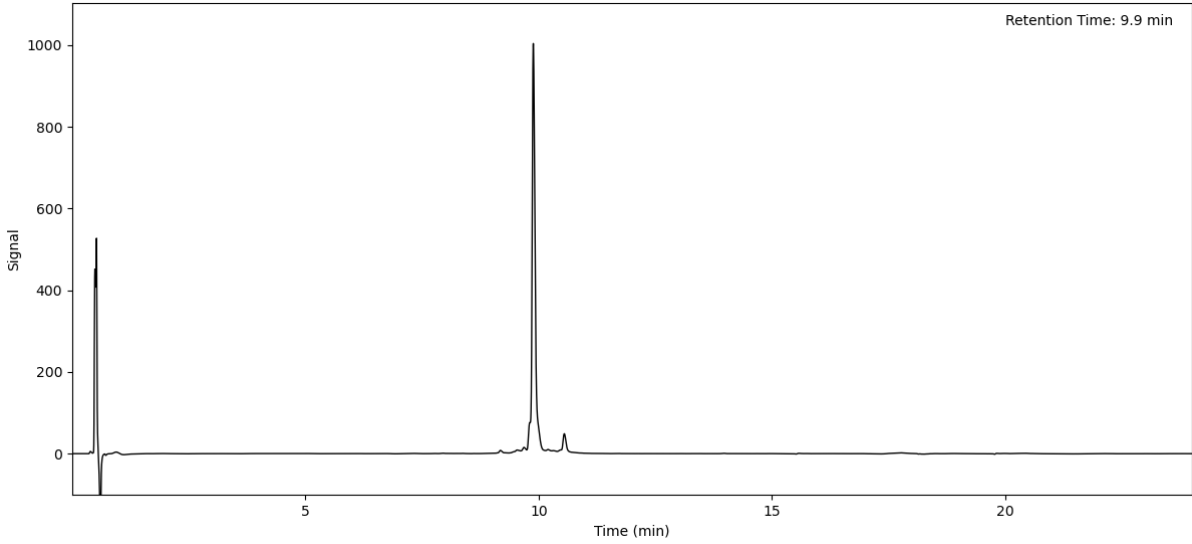
Peptide 146 isomer 1 (Method A)



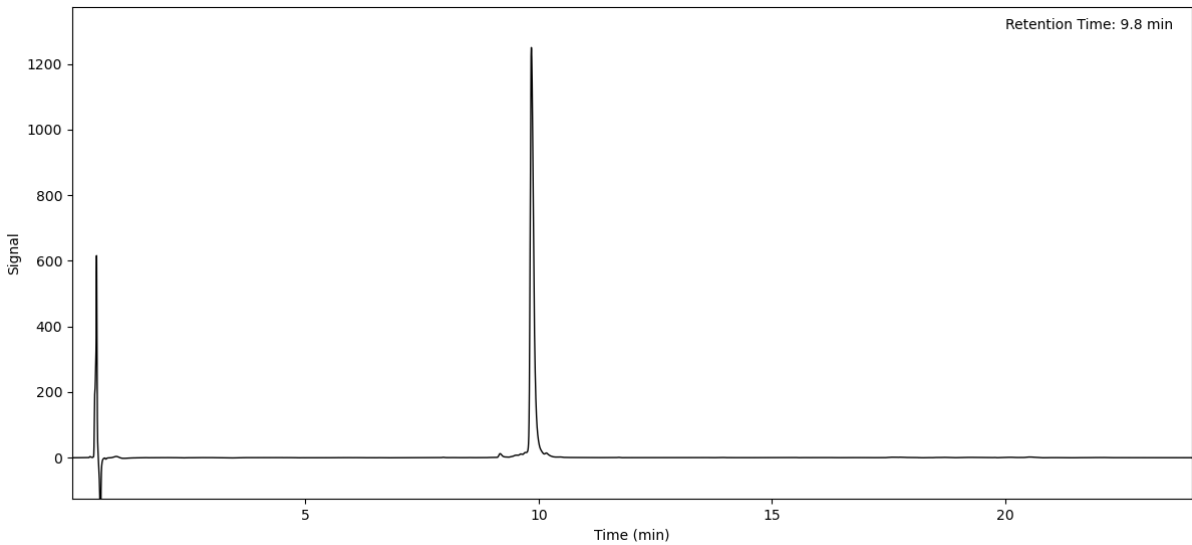
Peptide 146 isomer 2 (Method A)



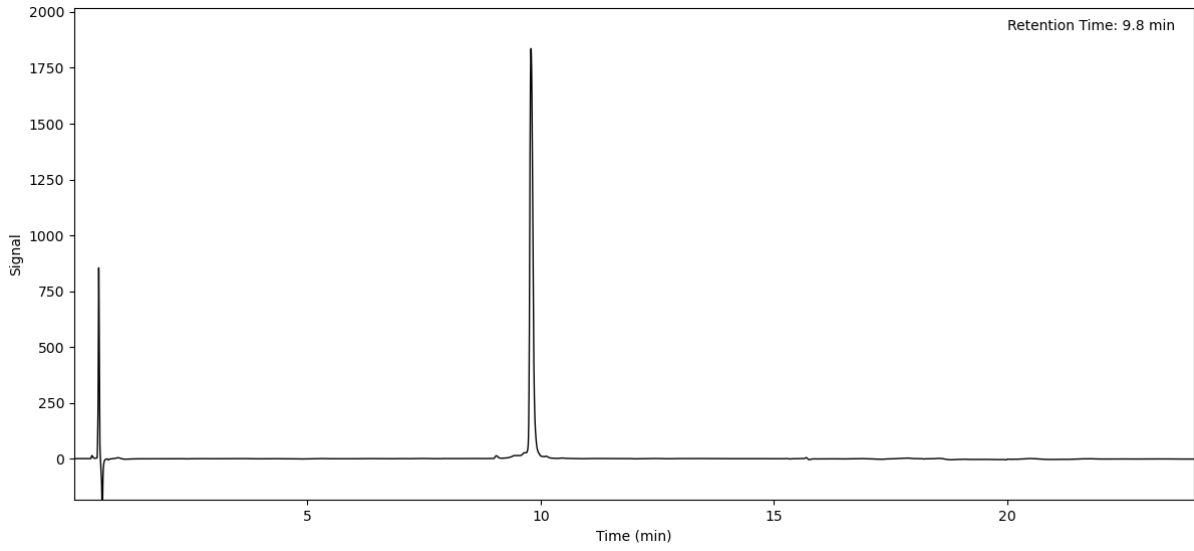
Peptide 147 (Method A)



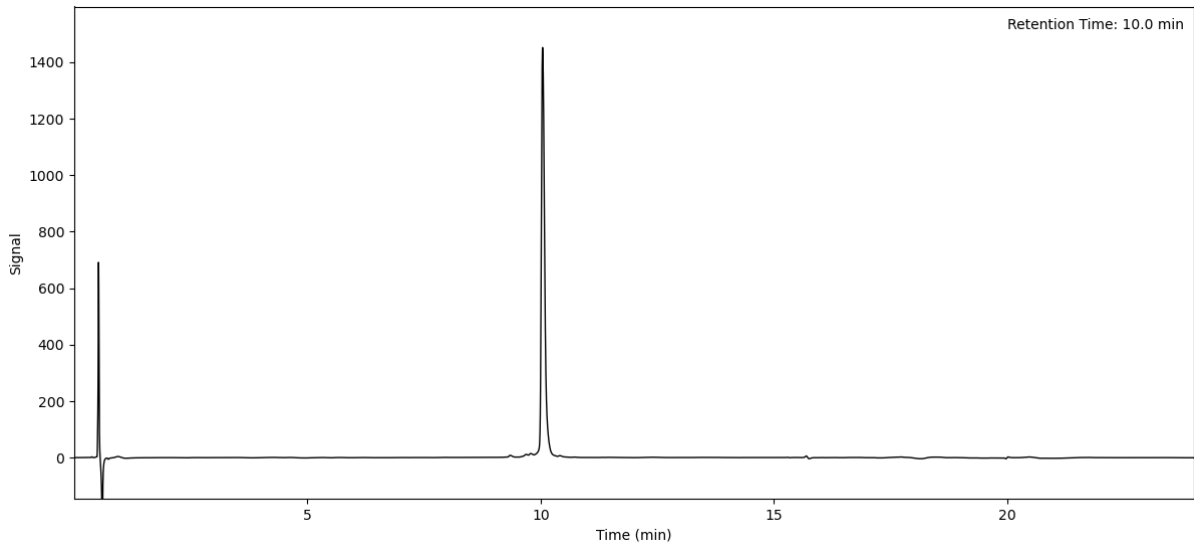
Peptide 148 (Method A)



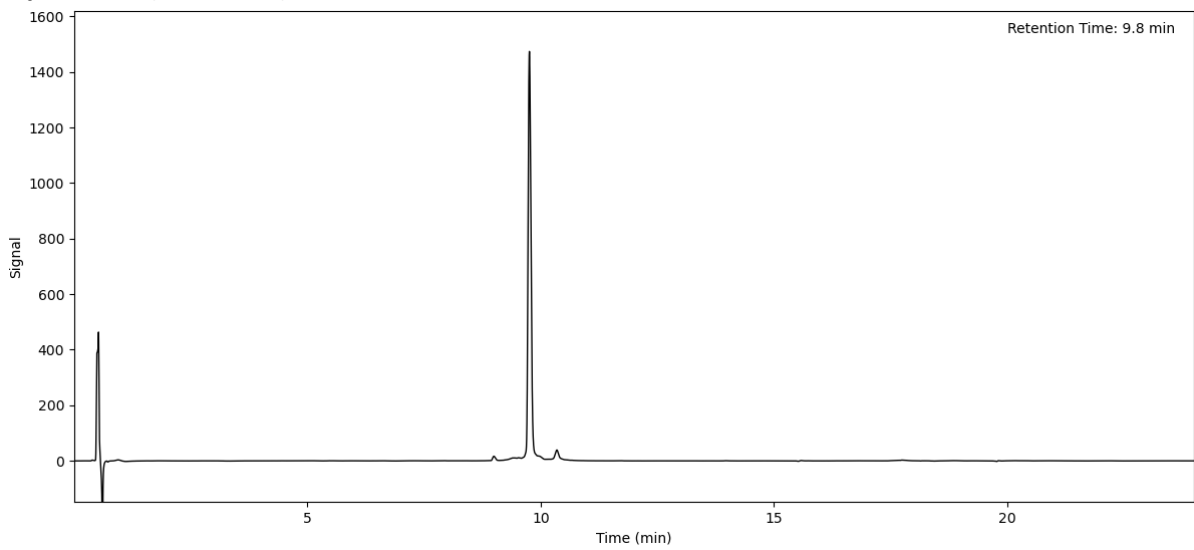
Peptide 149 (Method A)



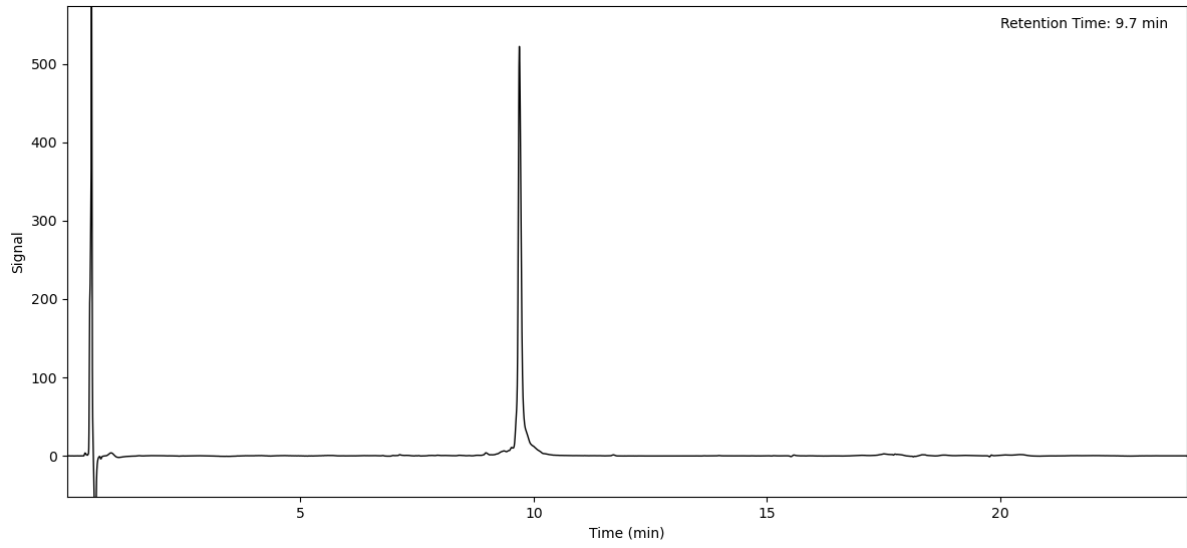
Peptide 150 (Method A)



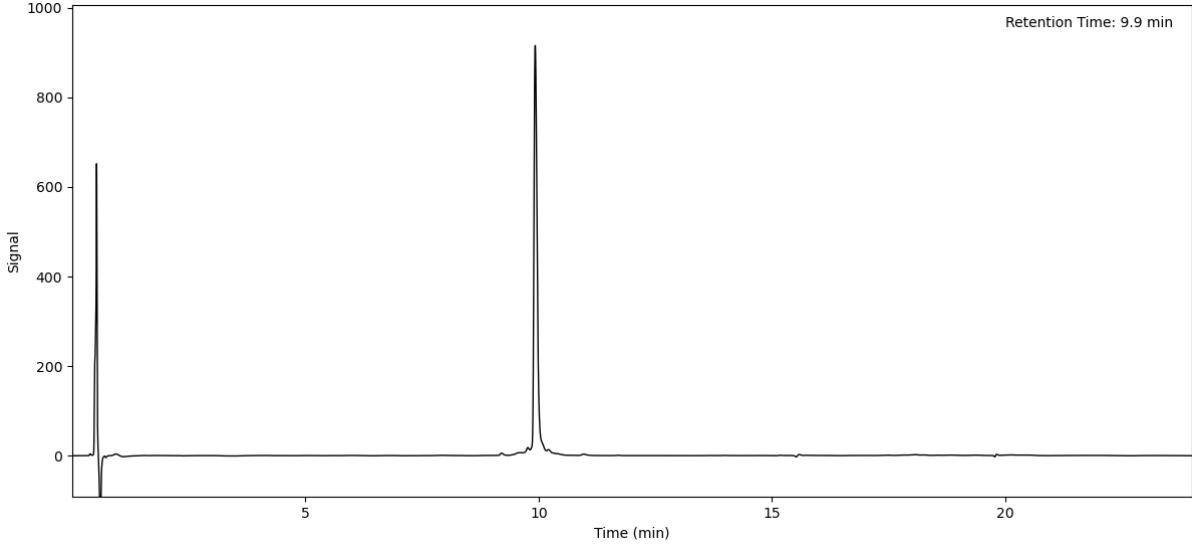
Peptide 151 (Method A)



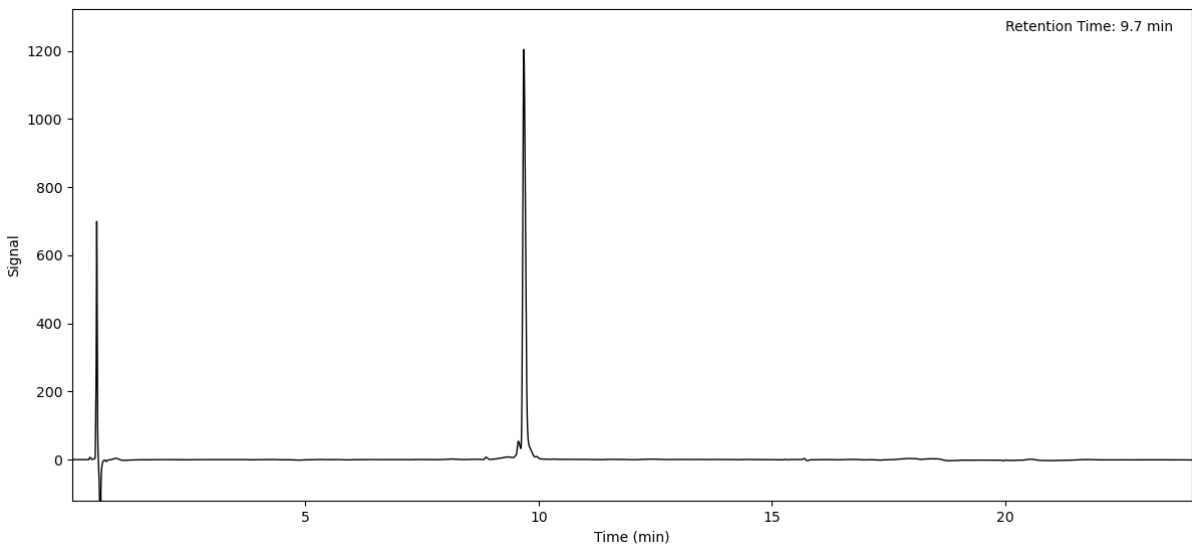
Peptide 152 (Method A)



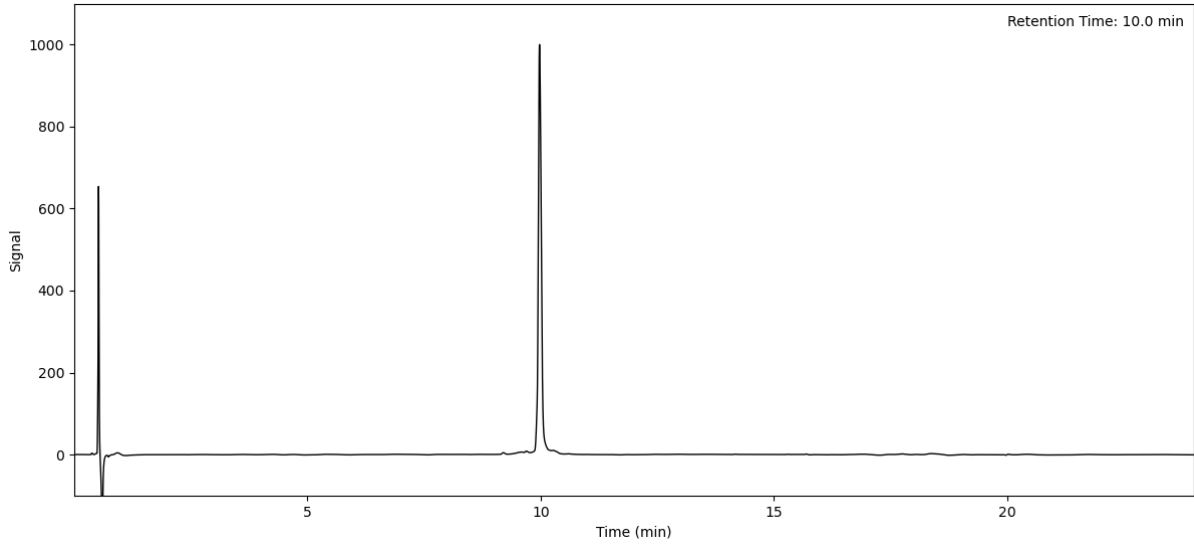
Peptide 153 (Method A)



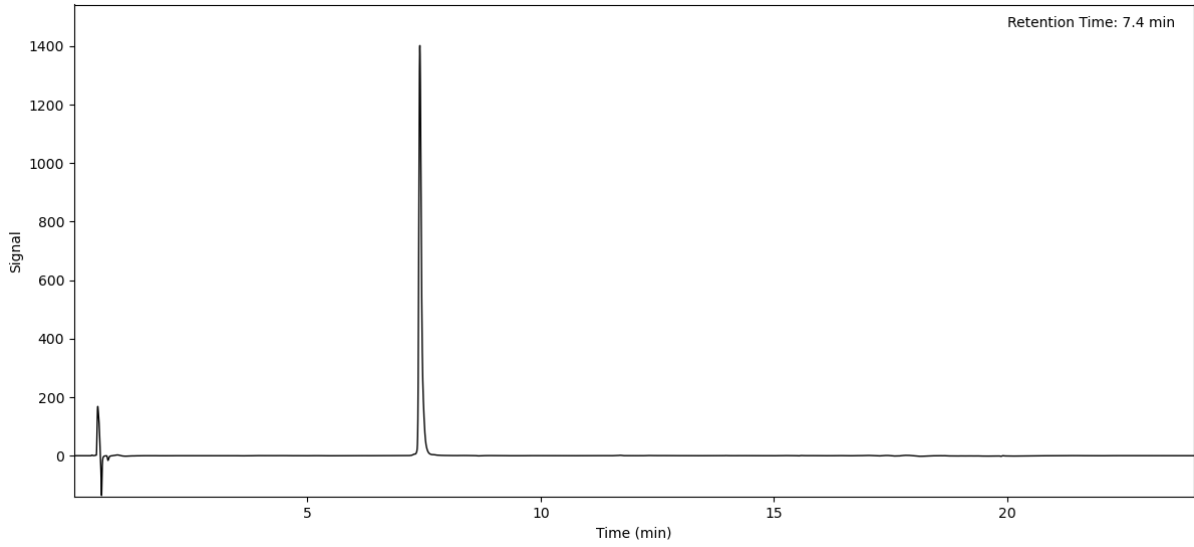
Peptide 154 (Method A)



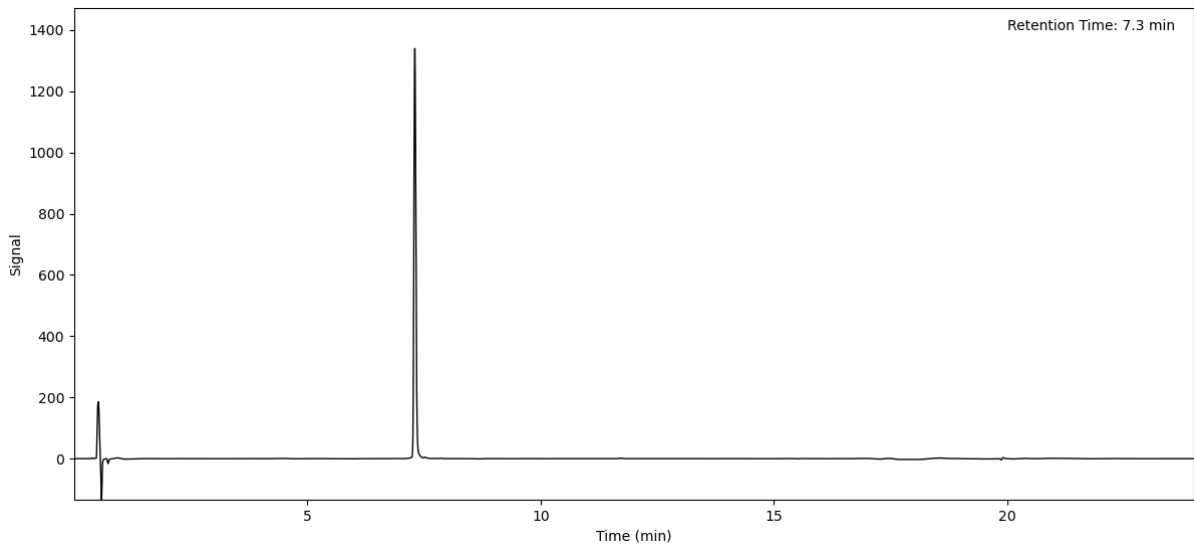
Peptide 155 (Method A)



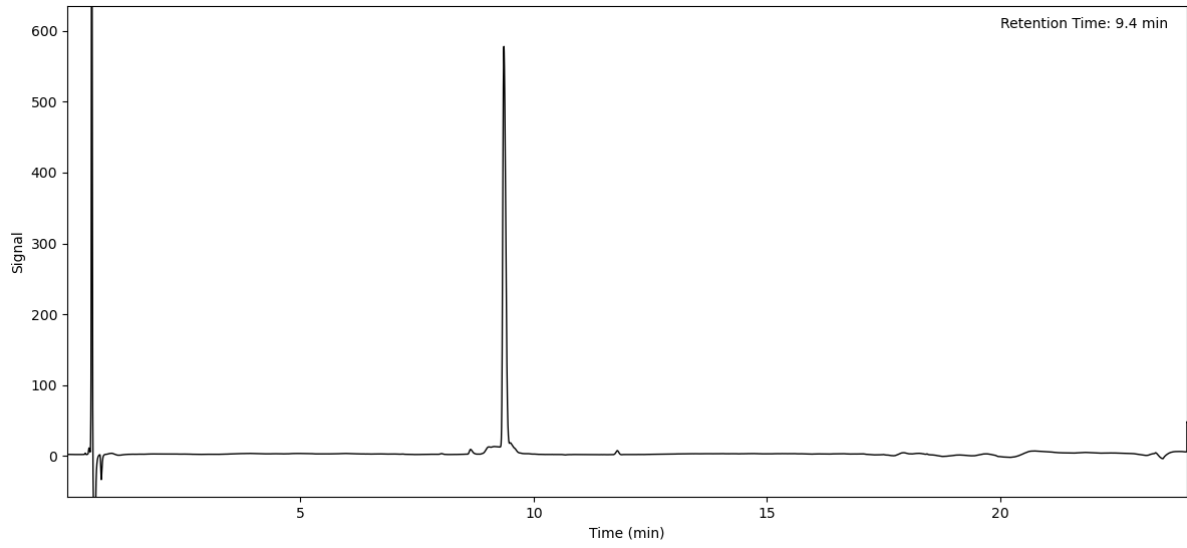
Peptide 156 (Method A)



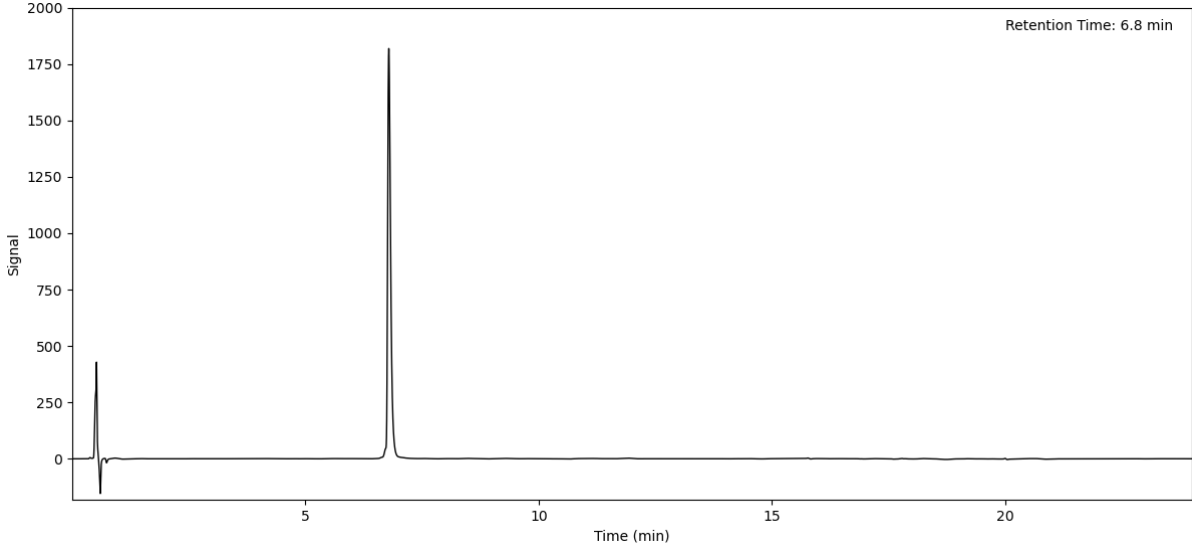
Peptide 157 (Method A)



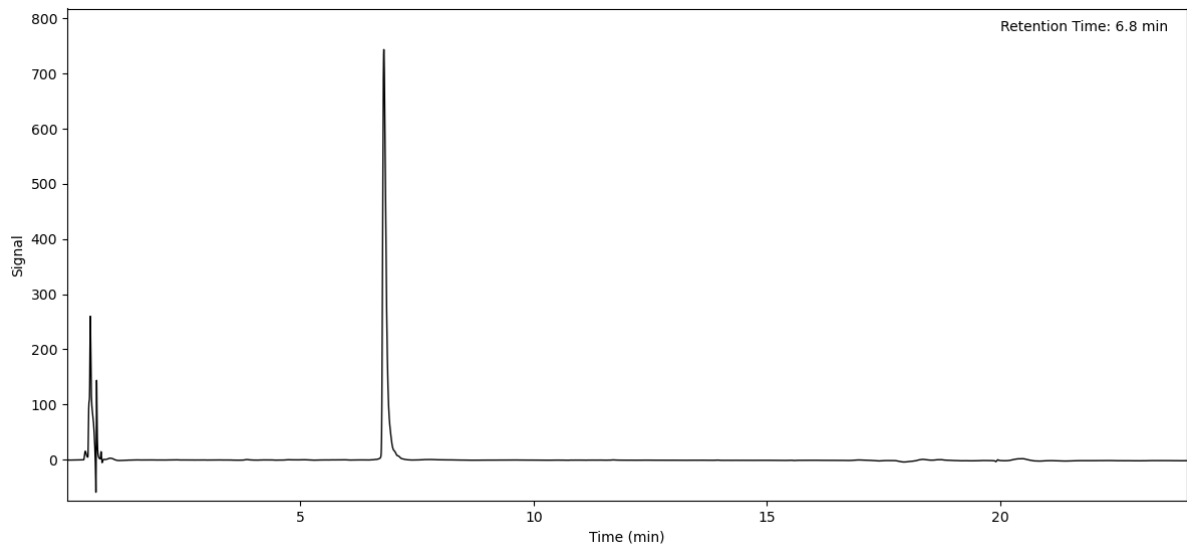
Peptide 158 (Method A)



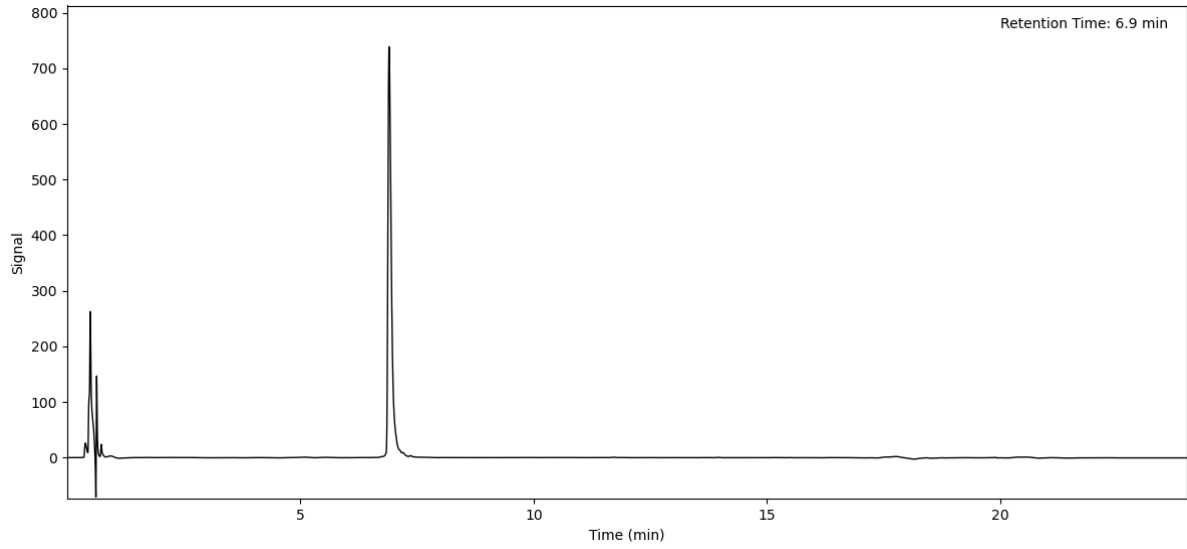
Peptide 159 (Method A)



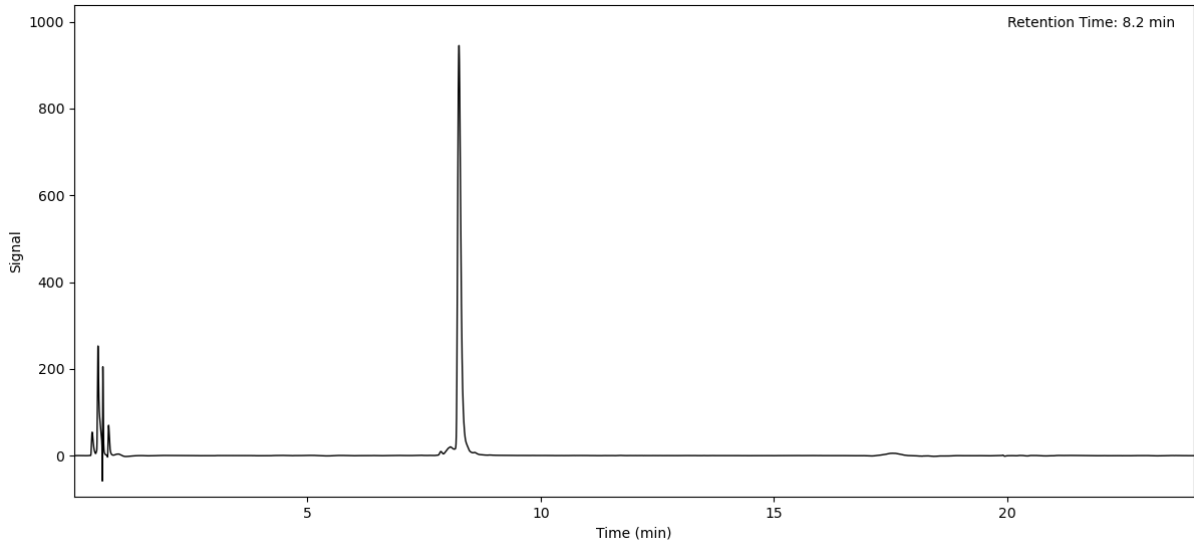
Peptide 160 (Method A)



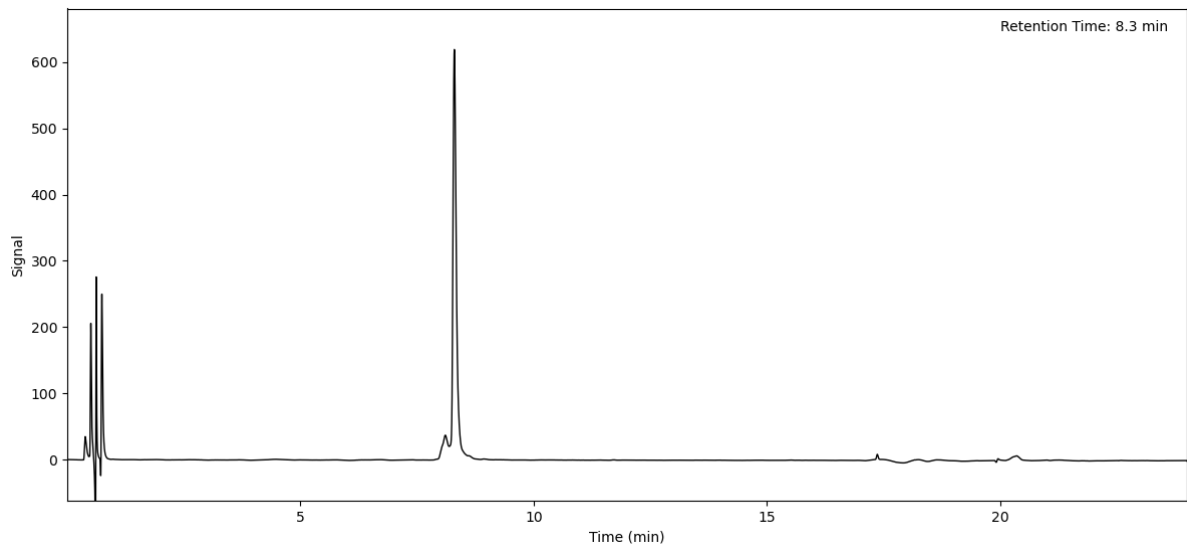
Peptide 161 (Method A)



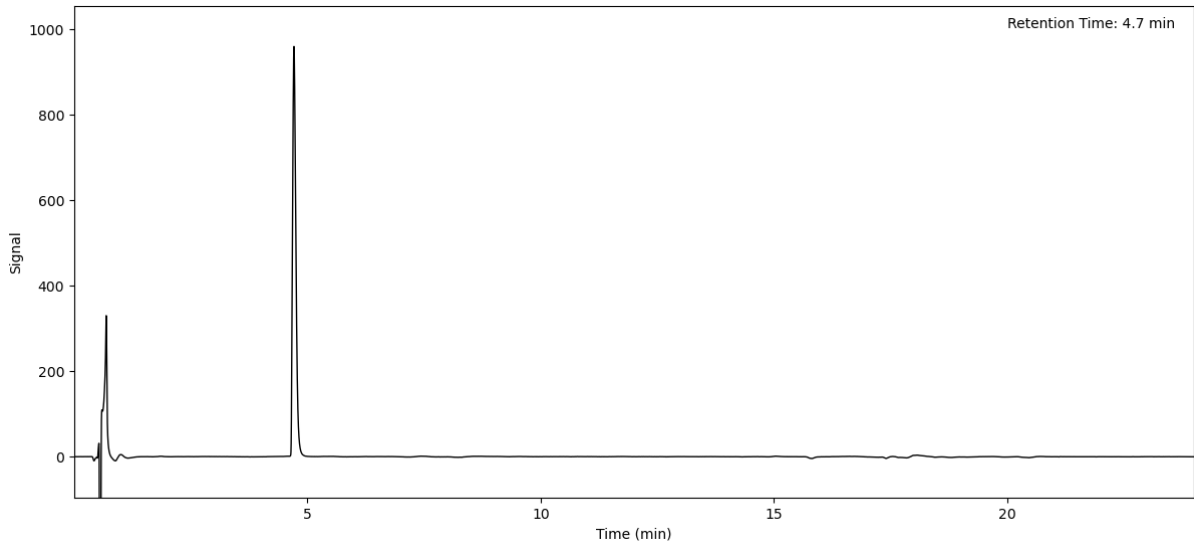
Peptide 162 (Method A)



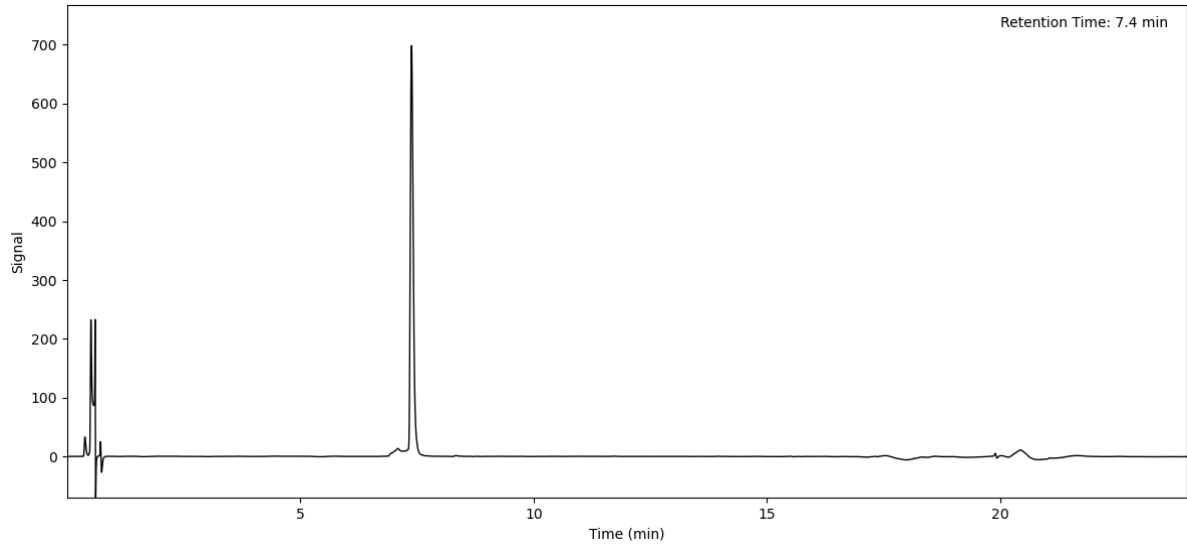
Peptide 163 (Method A)



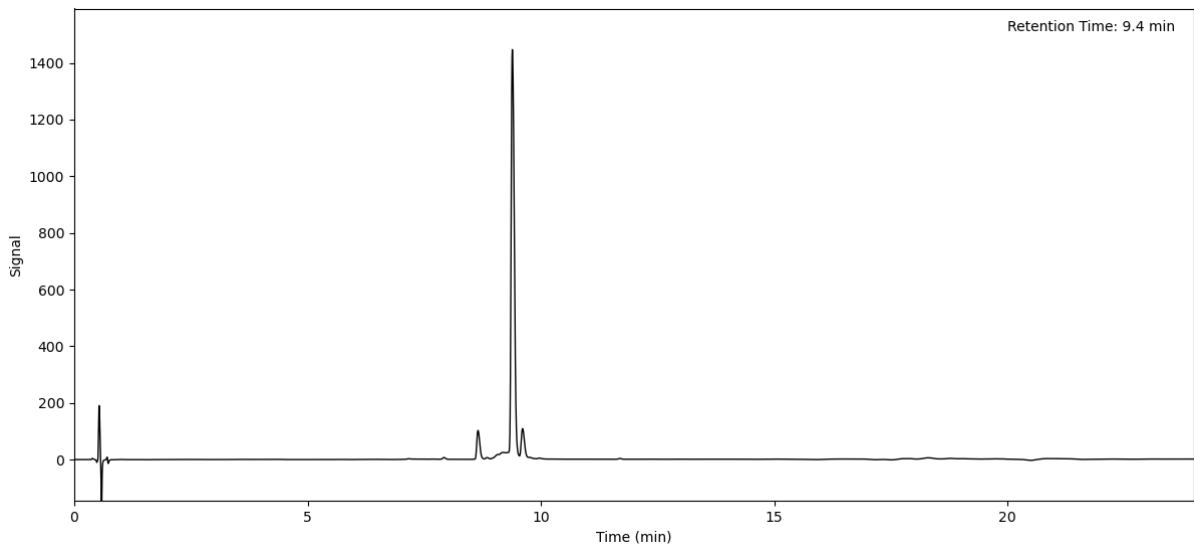
Peptide 164 (Method A)



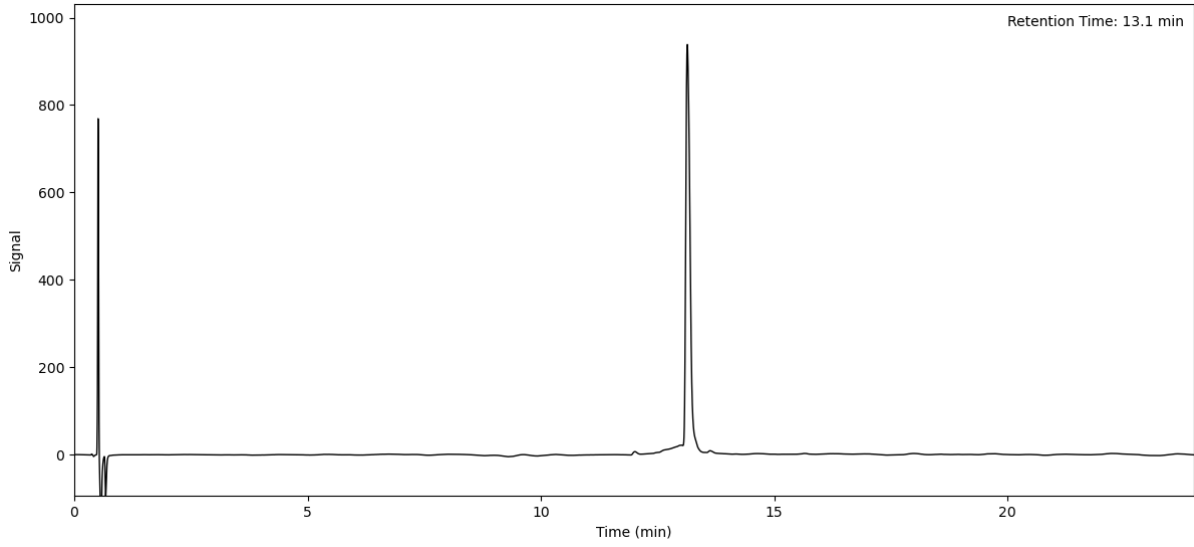
Peptide 165 (Method A)



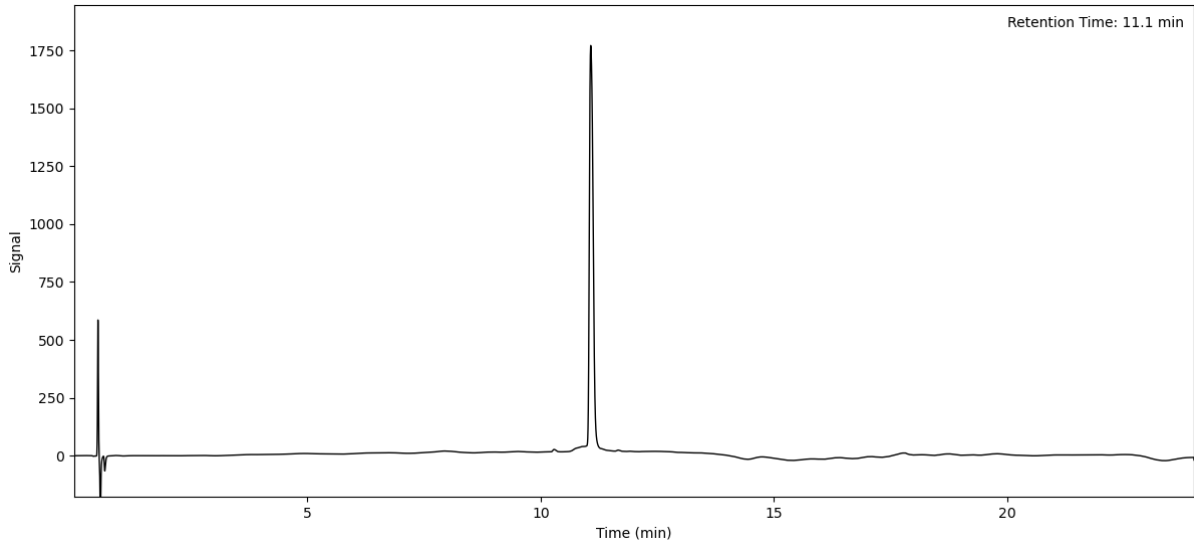
Peptide 166 (Method A)



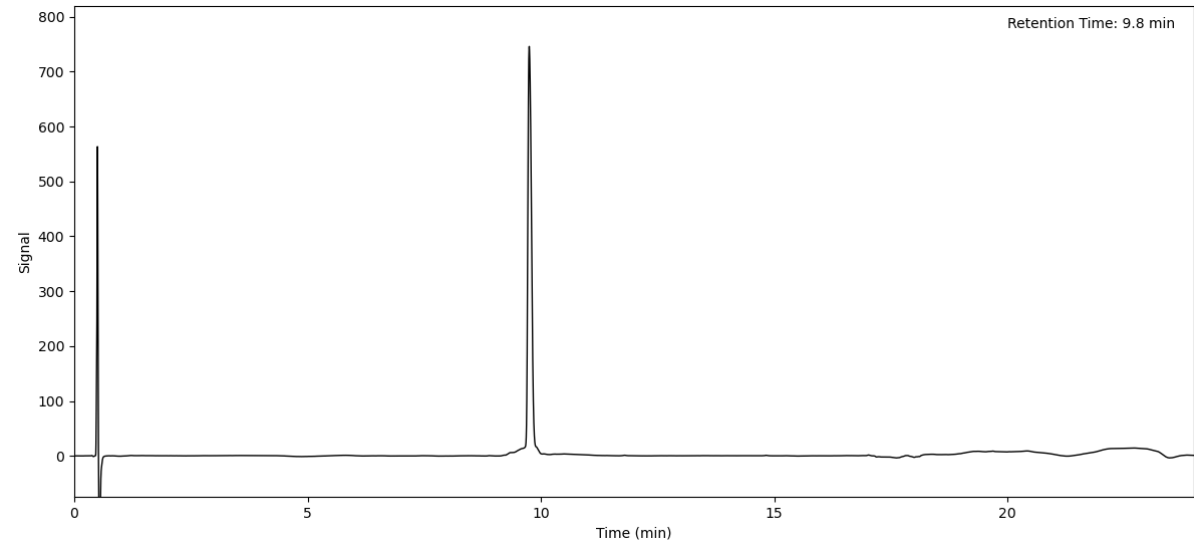
Peptide 167 (Method A)



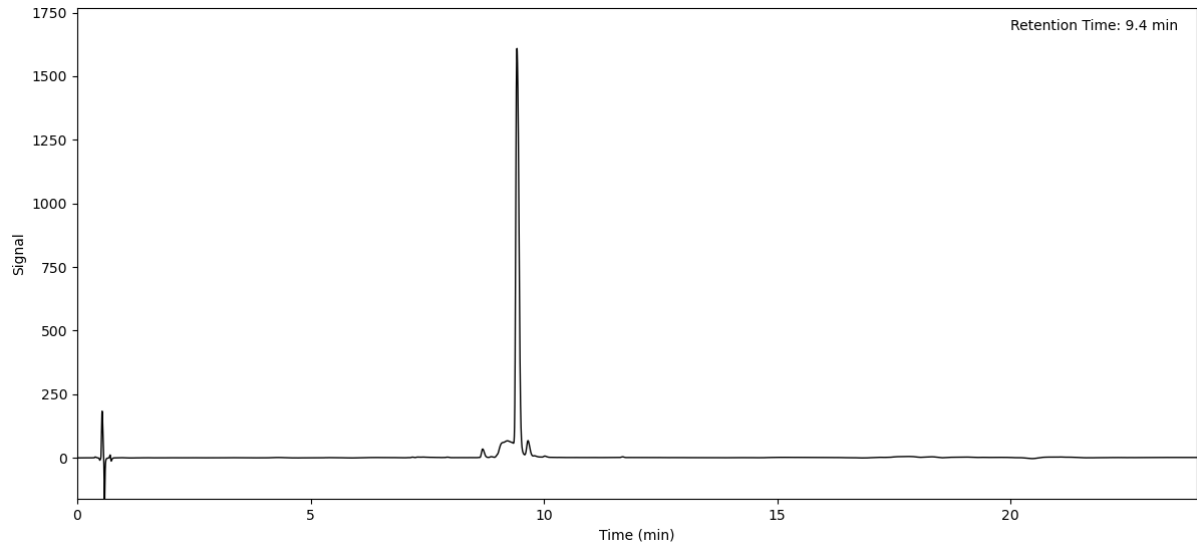
Peptide 168 (Method A)



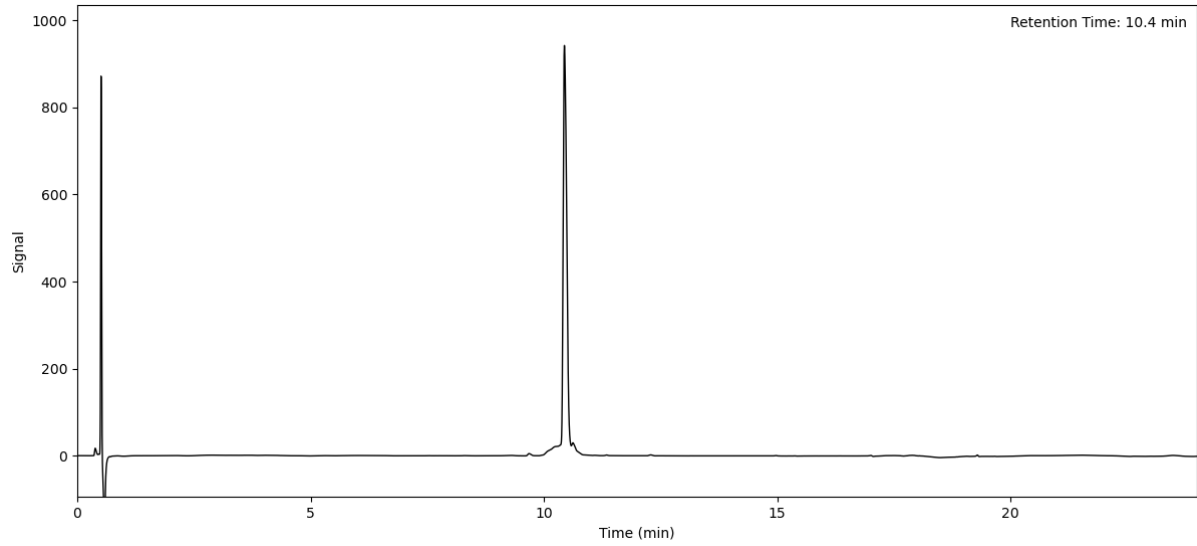
Peptide 169 (Method A)



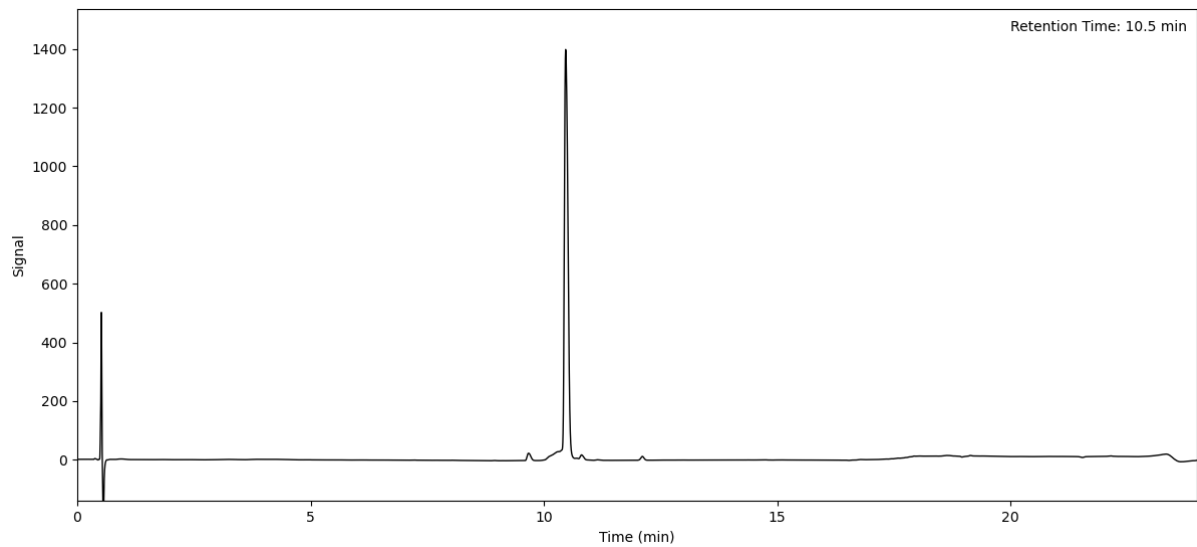
Peptide 170 (Method A)



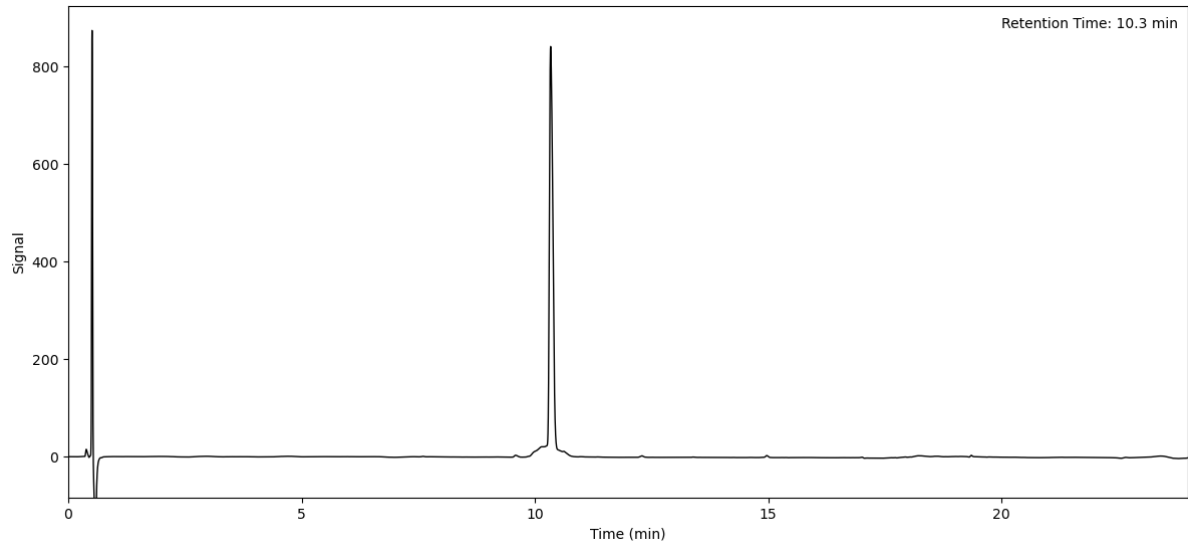
Peptide 171 (Method A)



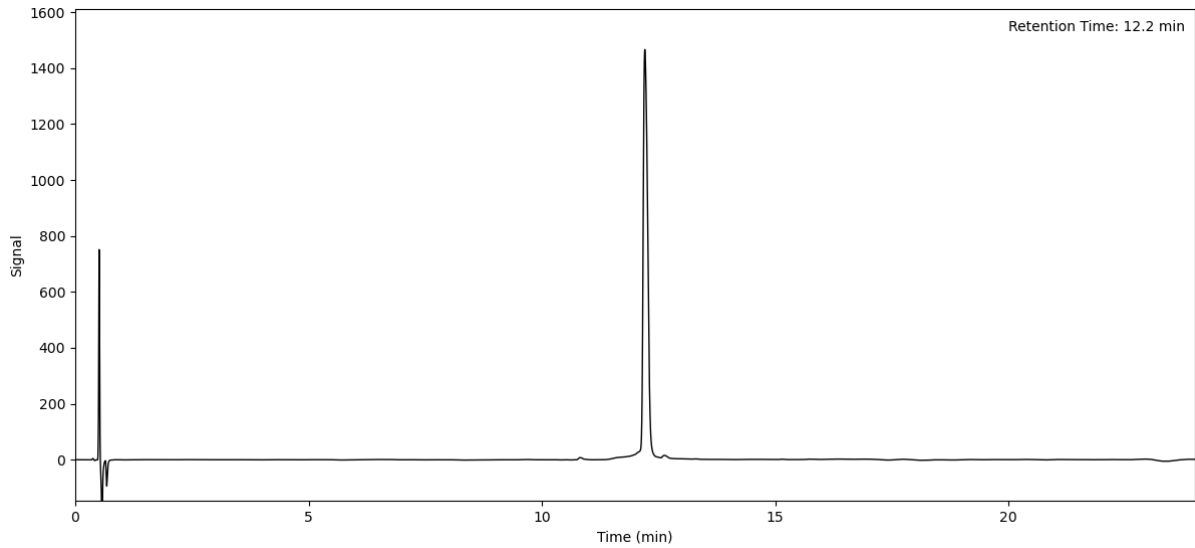
Peptide 172 (Method A)



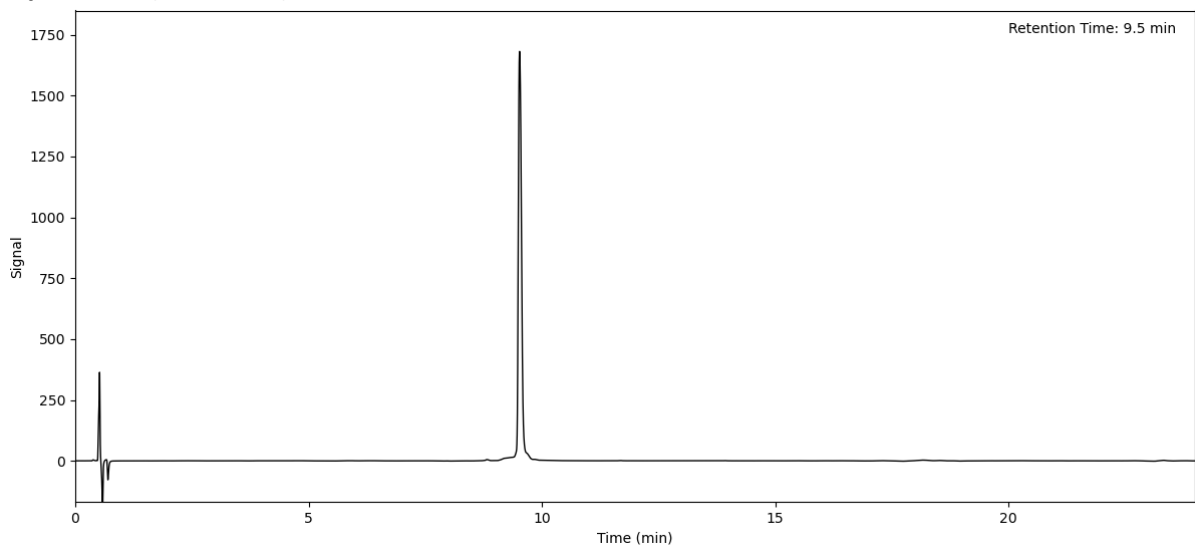
Peptide 173 (Method A)



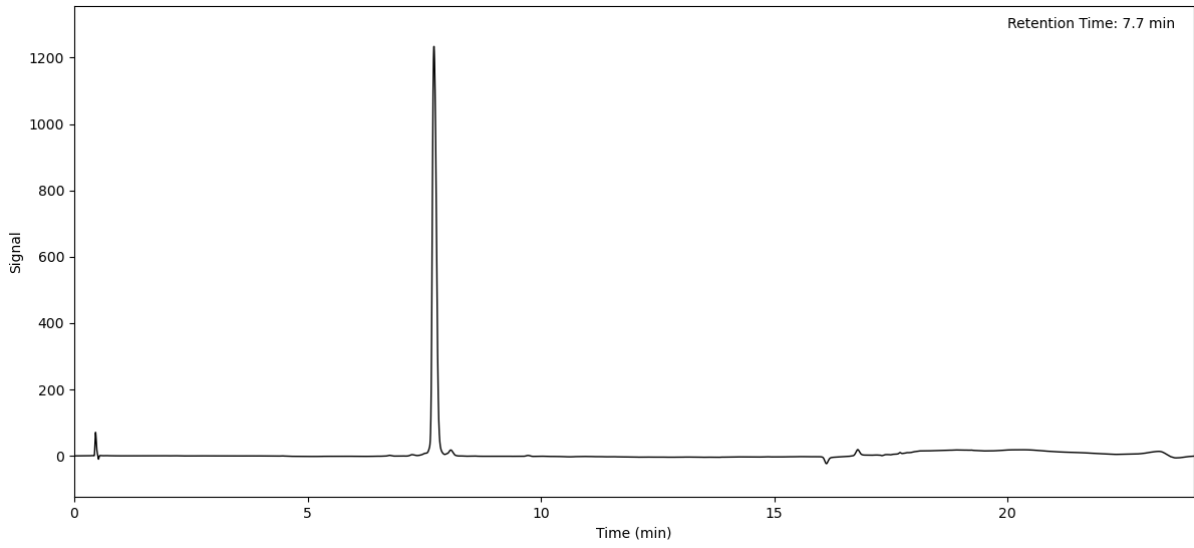
Peptide 174 (Method A)



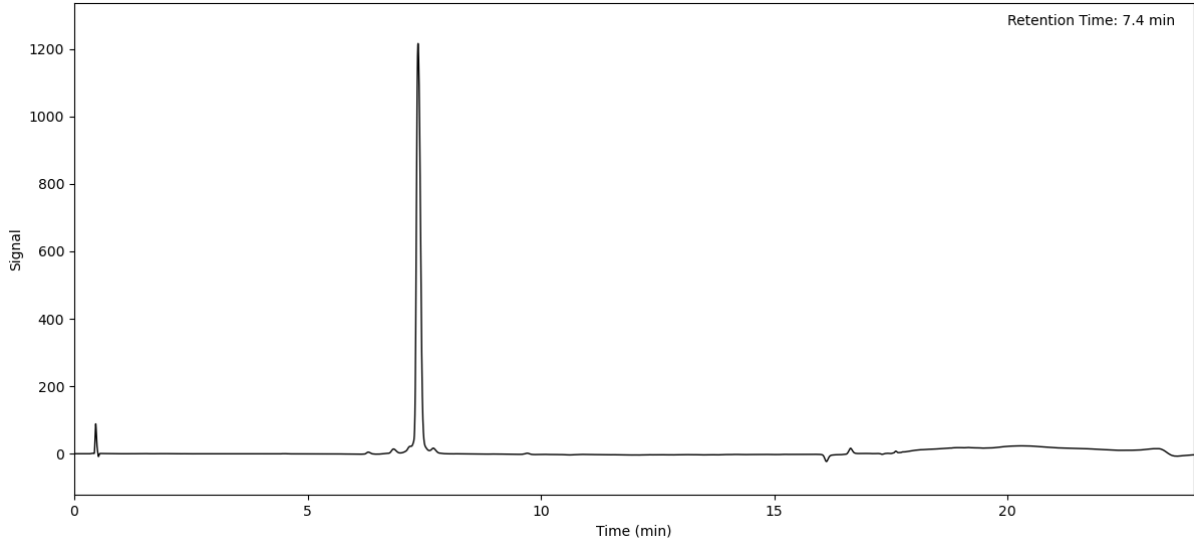
Peptide 175 (Method A)



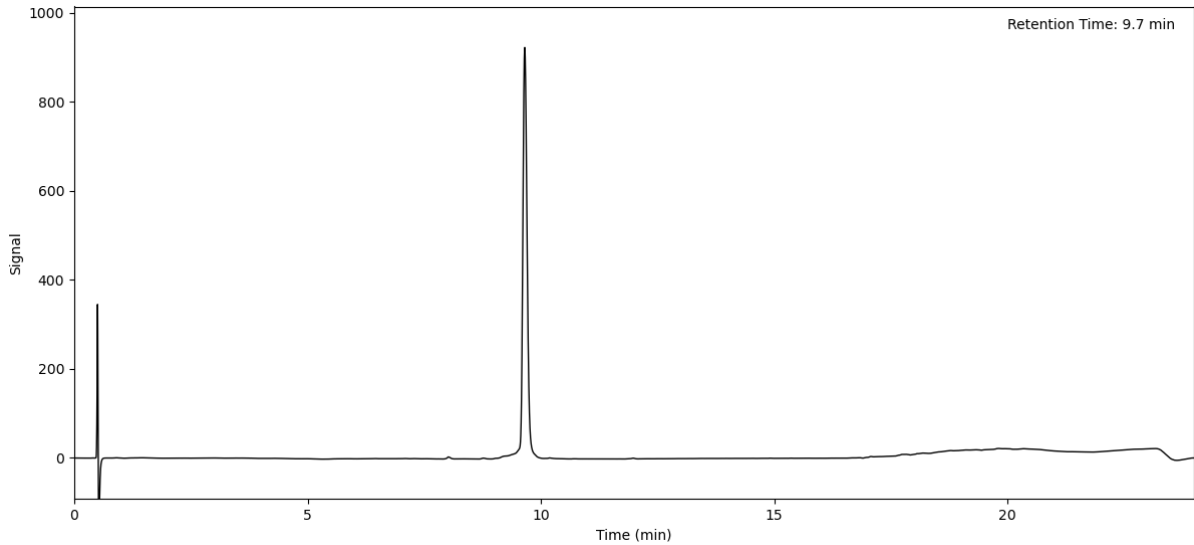
Peptide 176 (Method A)



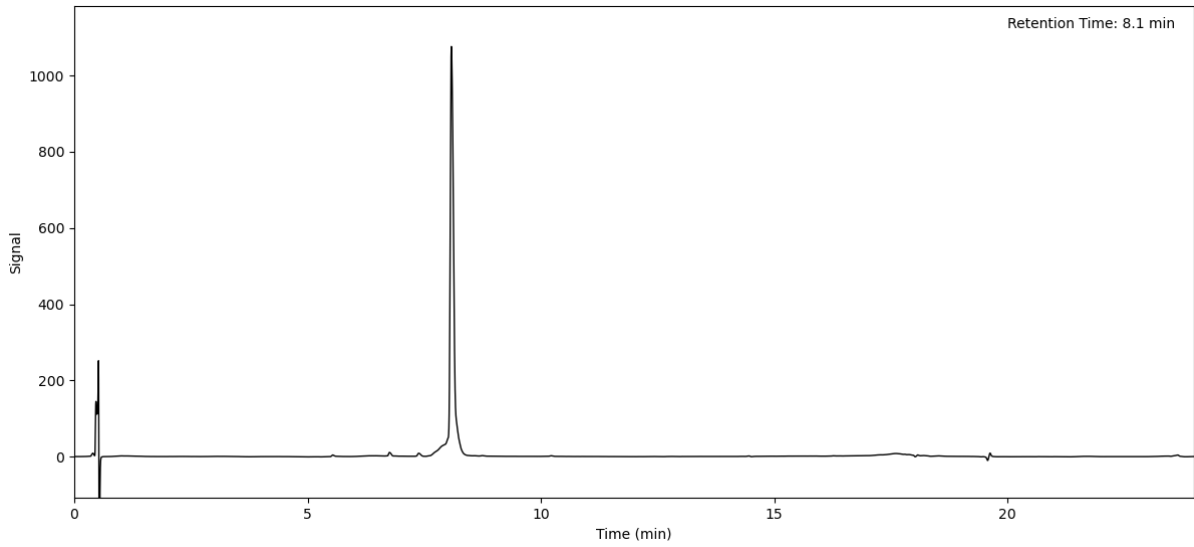
Peptide 177 (Method A)



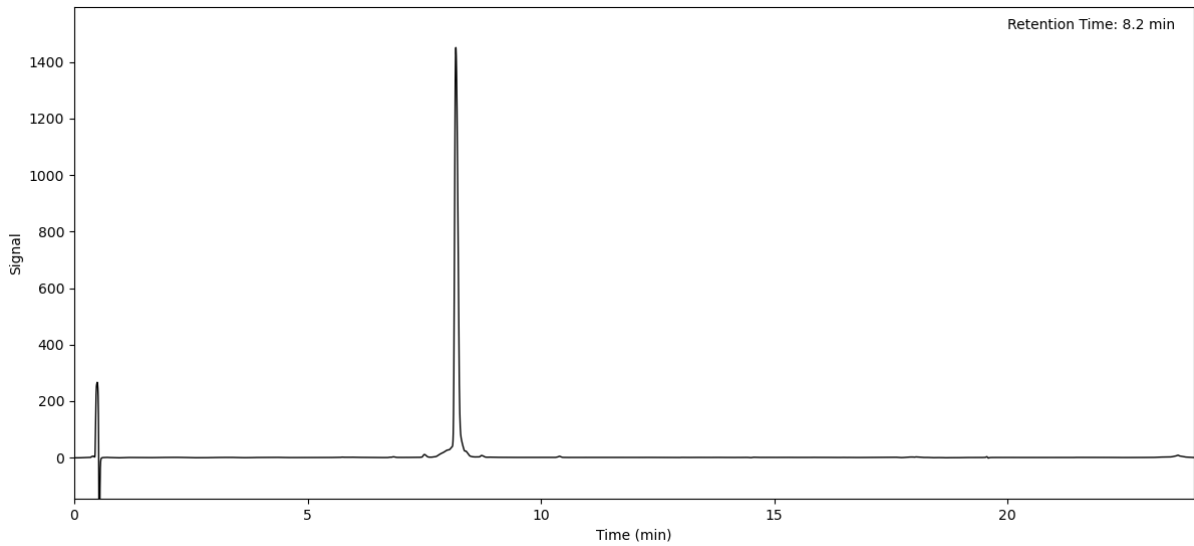
Peptide 178 (Method A)



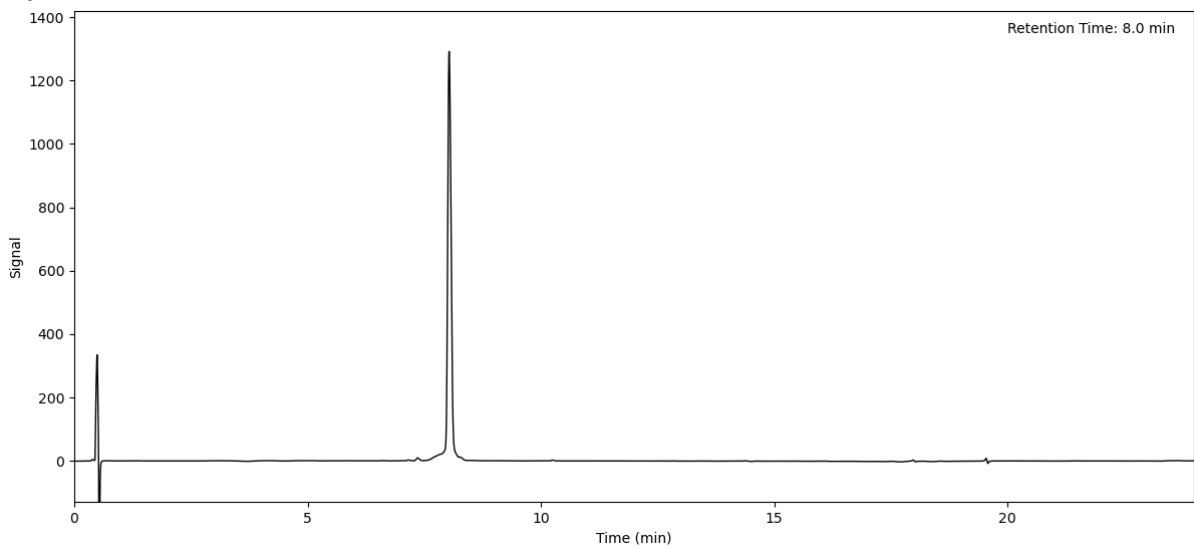
Peptide 179 (Method A)



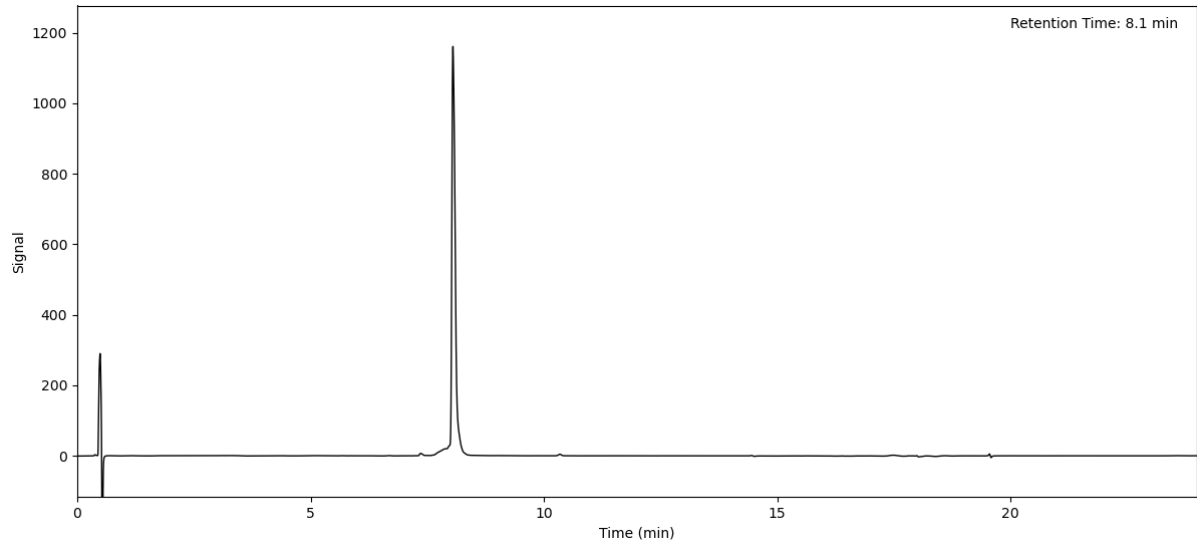
Peptide 180 (Method A)



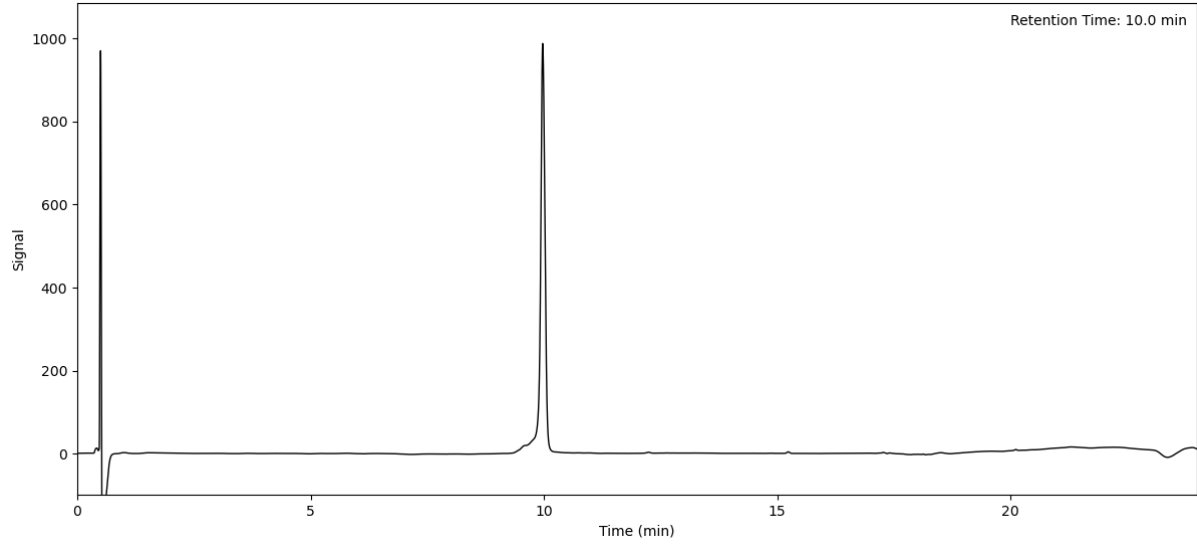
Peptide 181 (Method A)



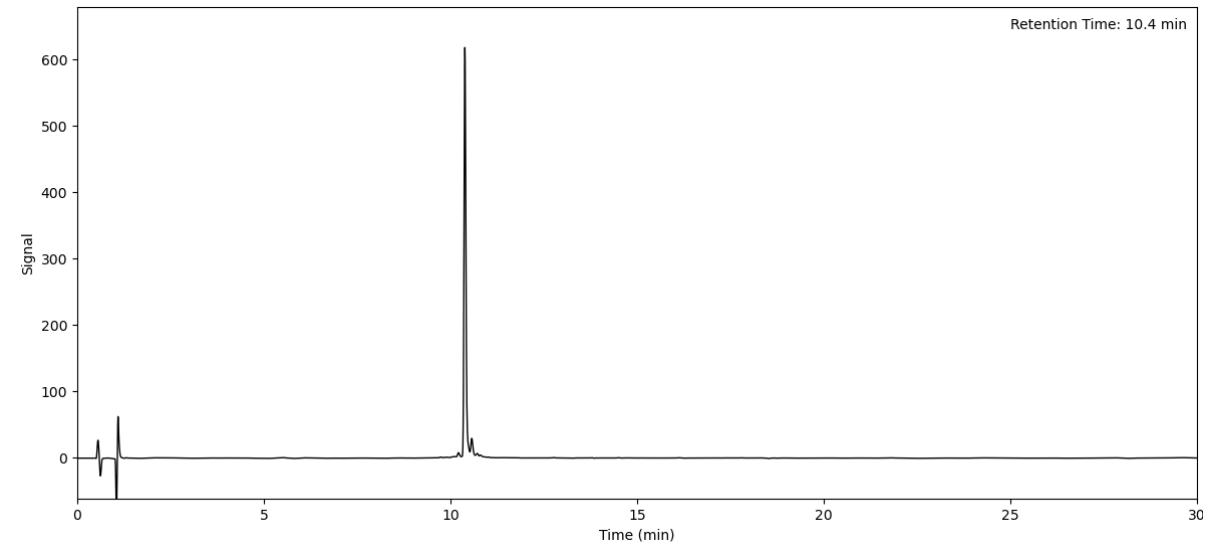
Peptide 182 (Method A)



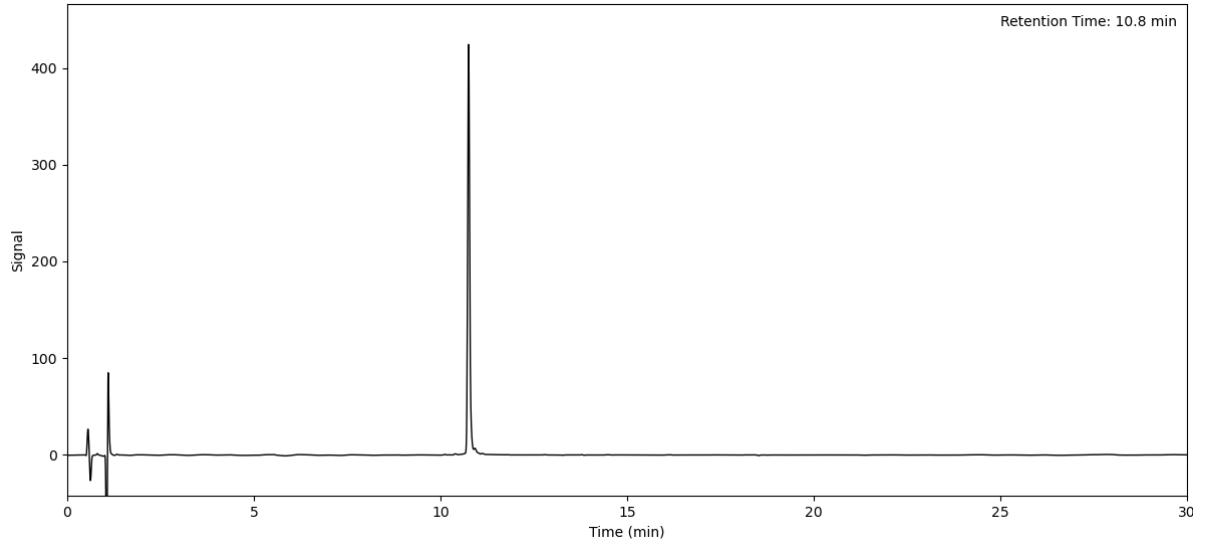
Peptide 183 (Method A)



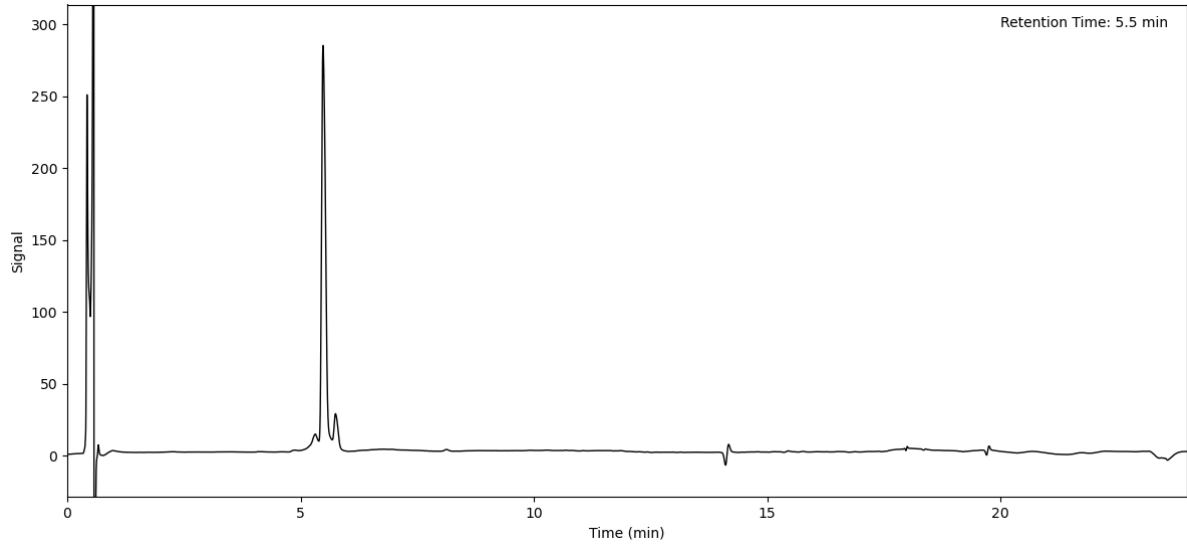
Peptide 184 (Method C)



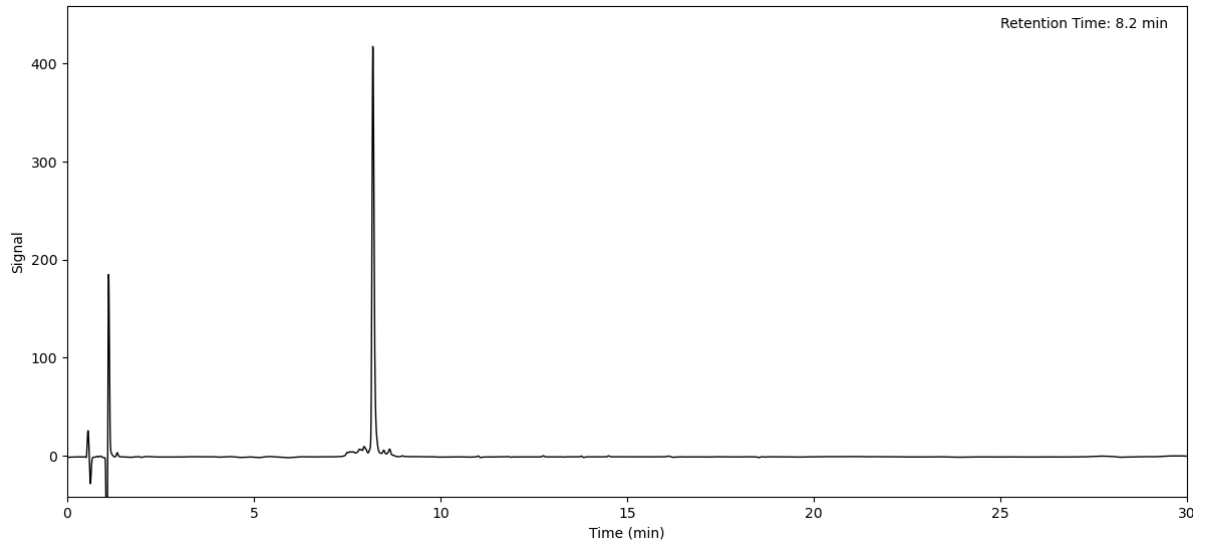
Peptide 185 (Method C)



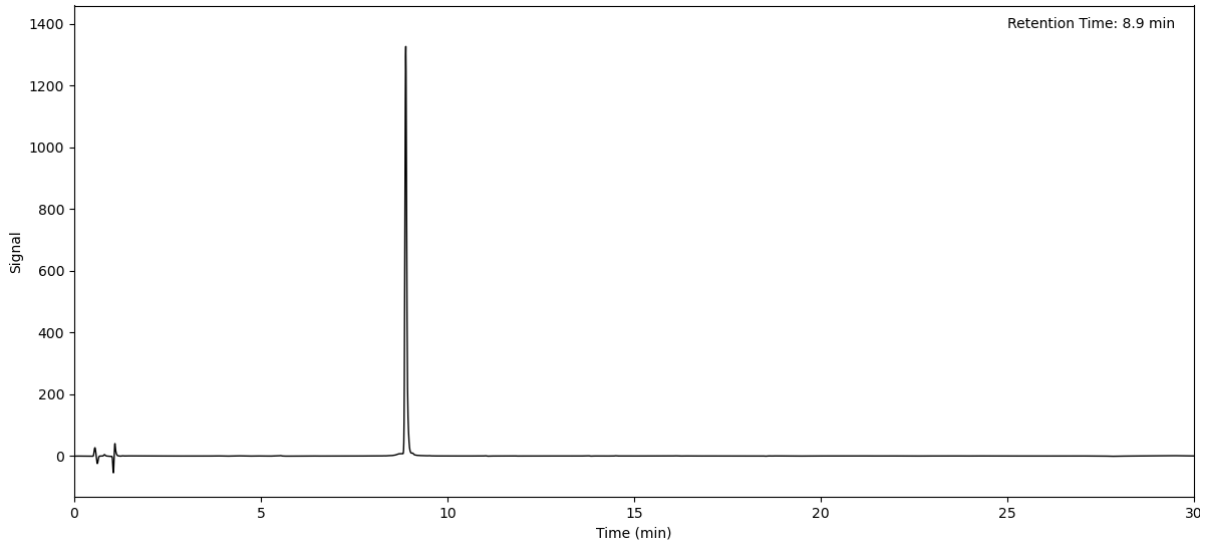
Peptide 186 (Method A)



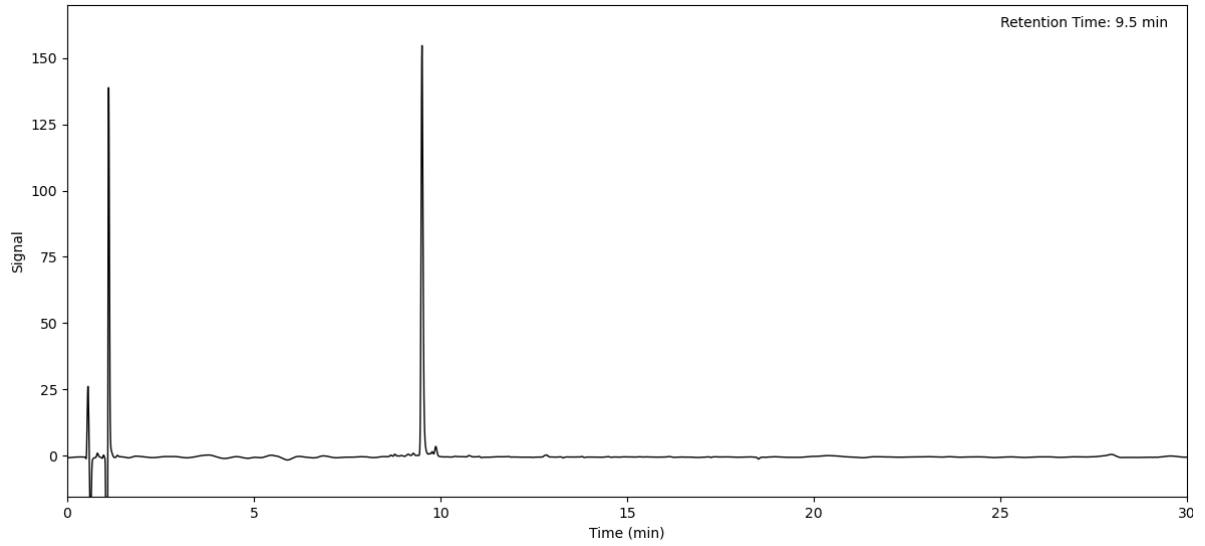
Peptide 187 (Method C)



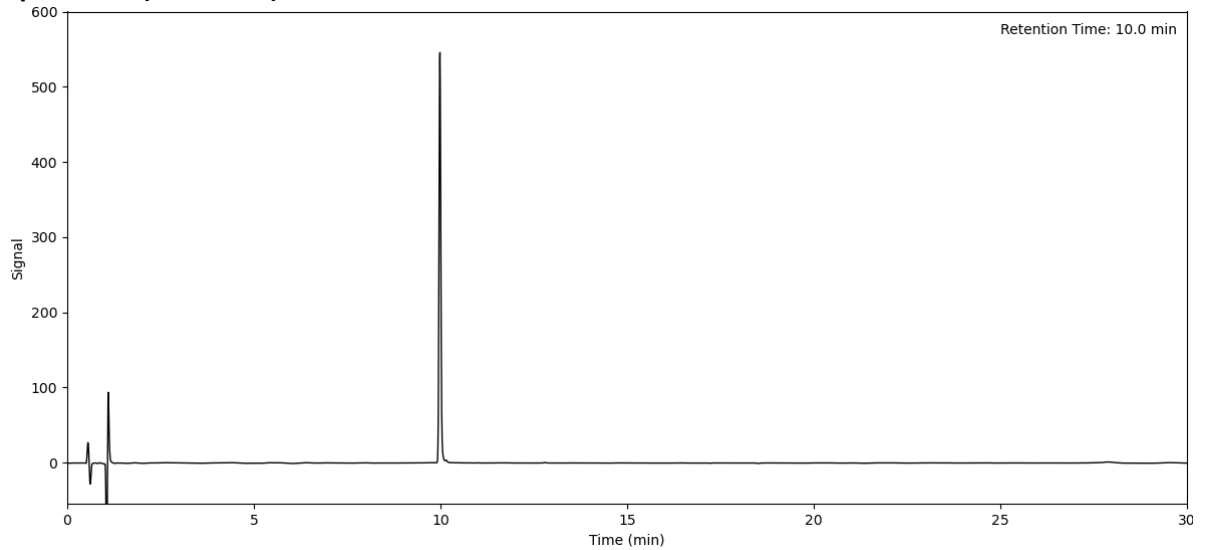
Peptide 188 (Method C)



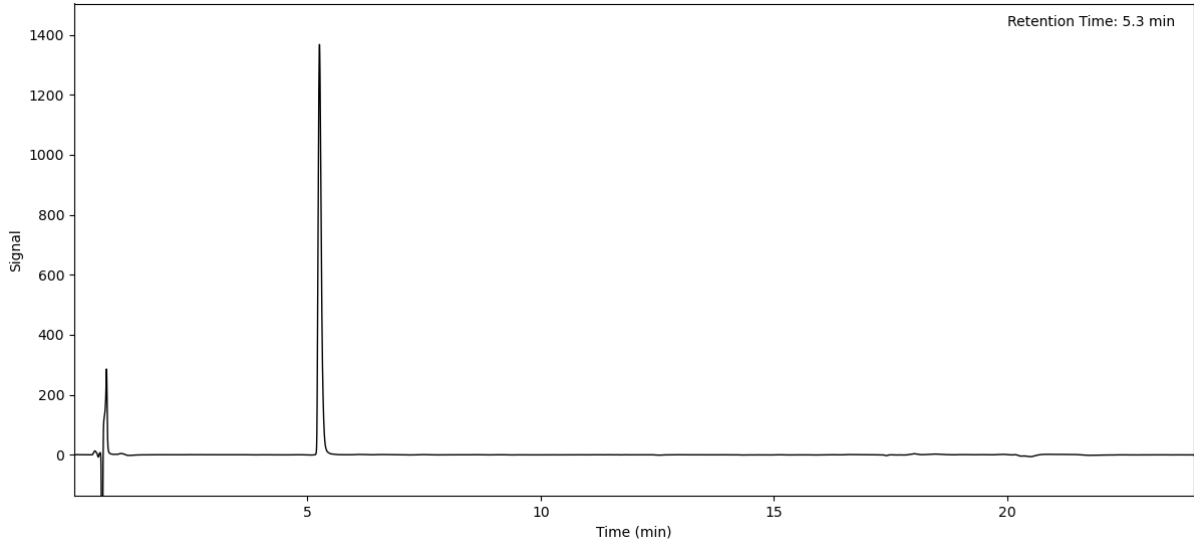
Peptide 189 (Method C)



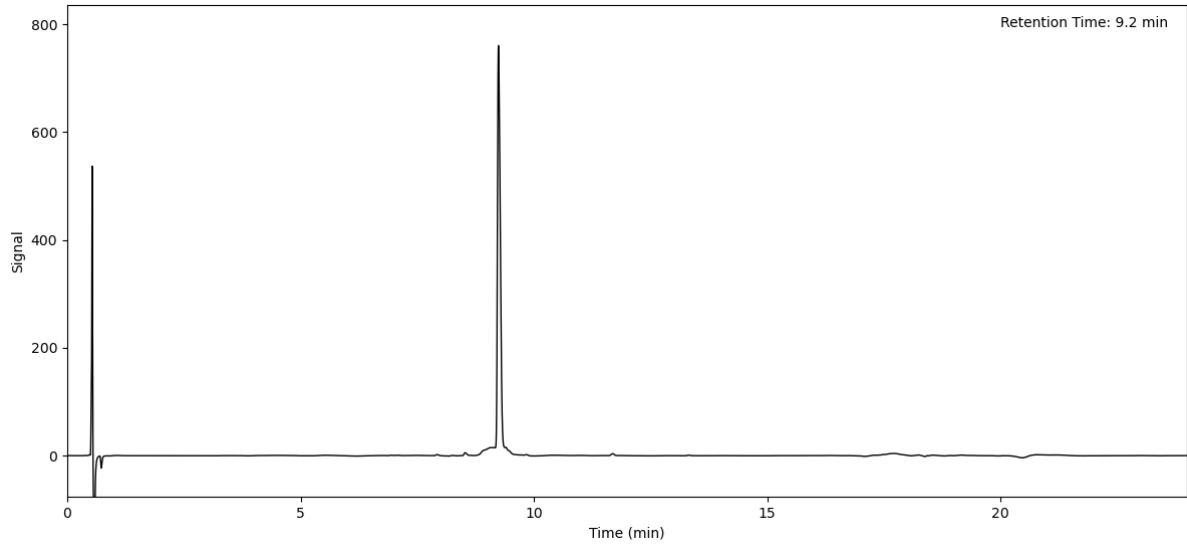
Peptide 190 (Method C)



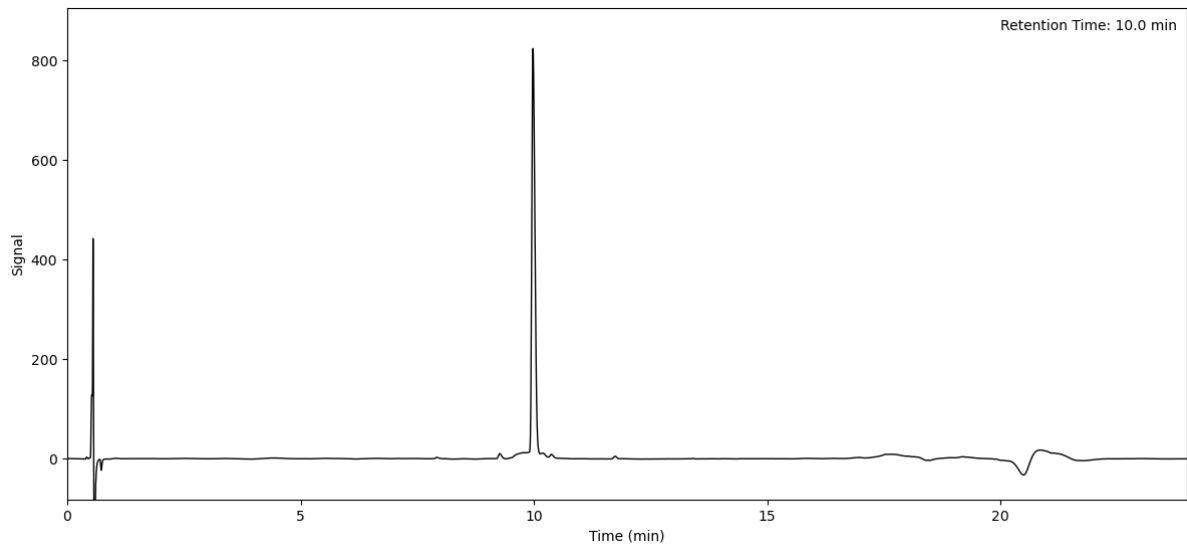
H4 Histone Tail Peptide (Method A)



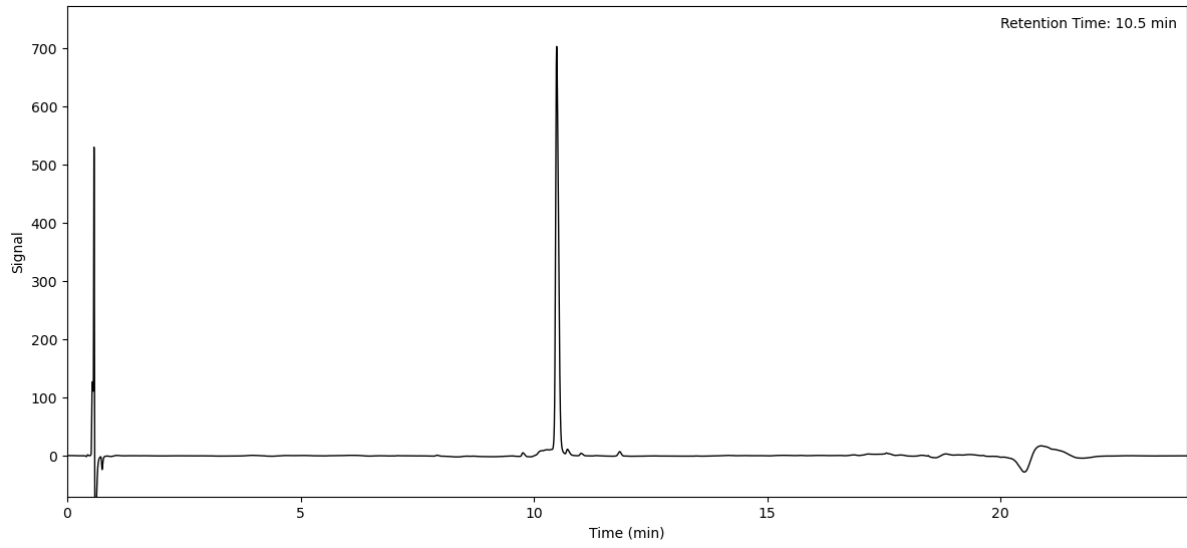
Peptide S1 (Method A)



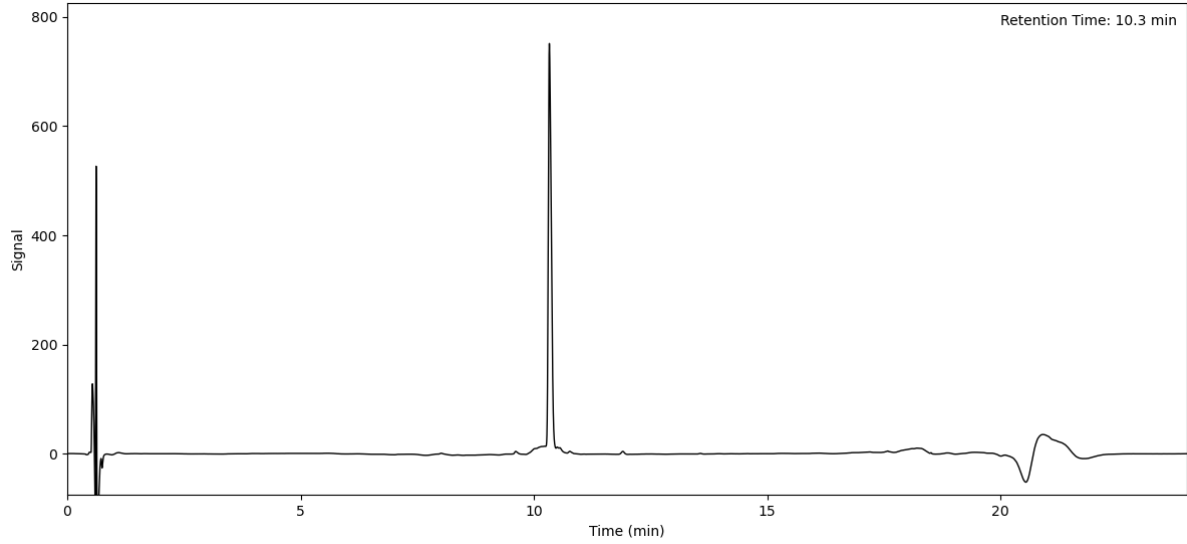
Peptide S2 (Method A)



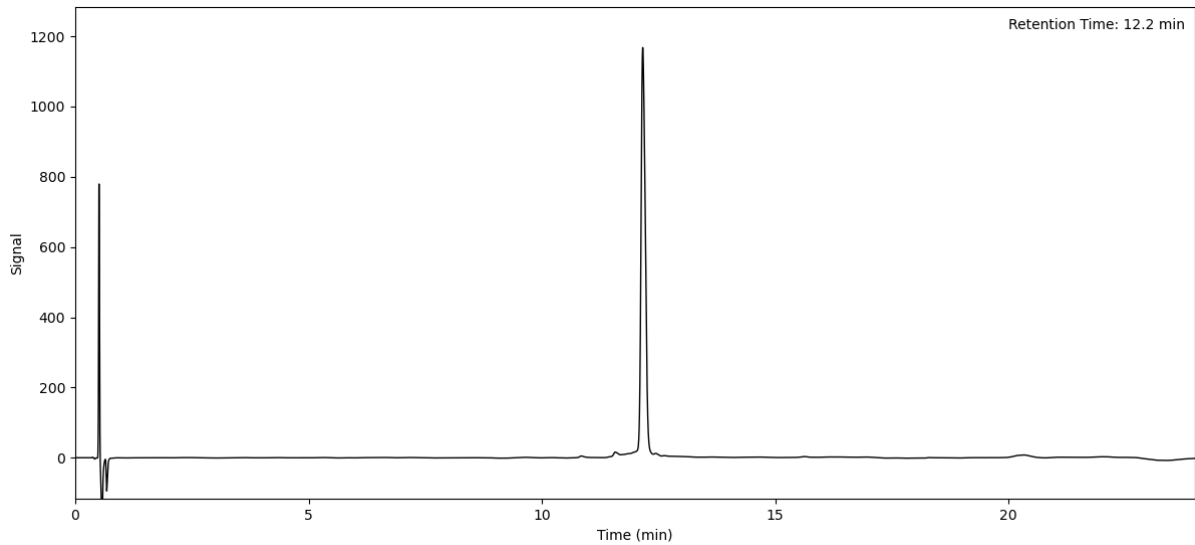
Peptide S3 (Method A)



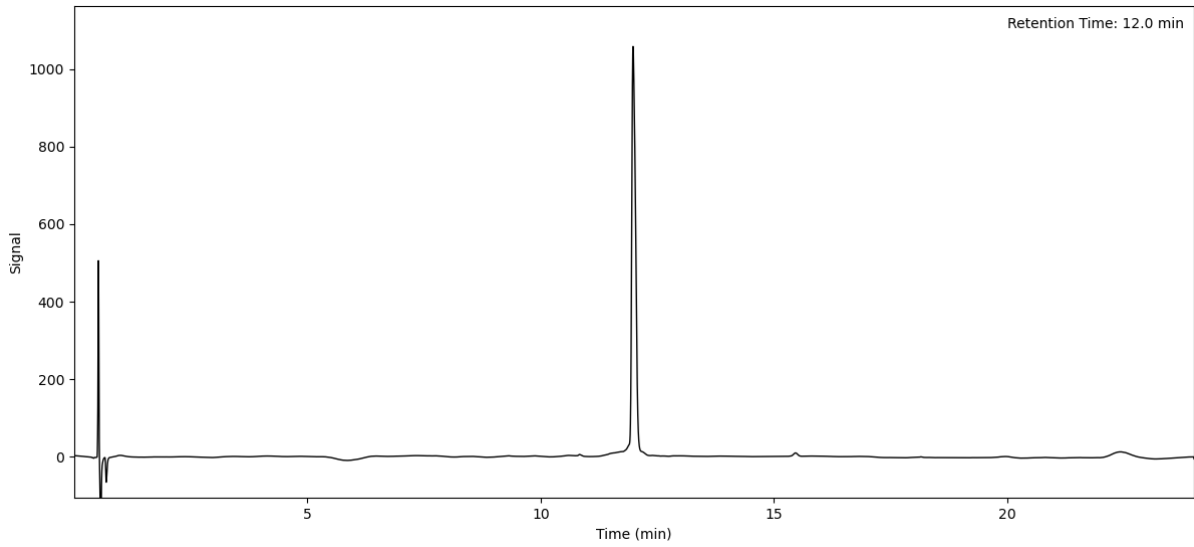
Peptide S4 (Method A)



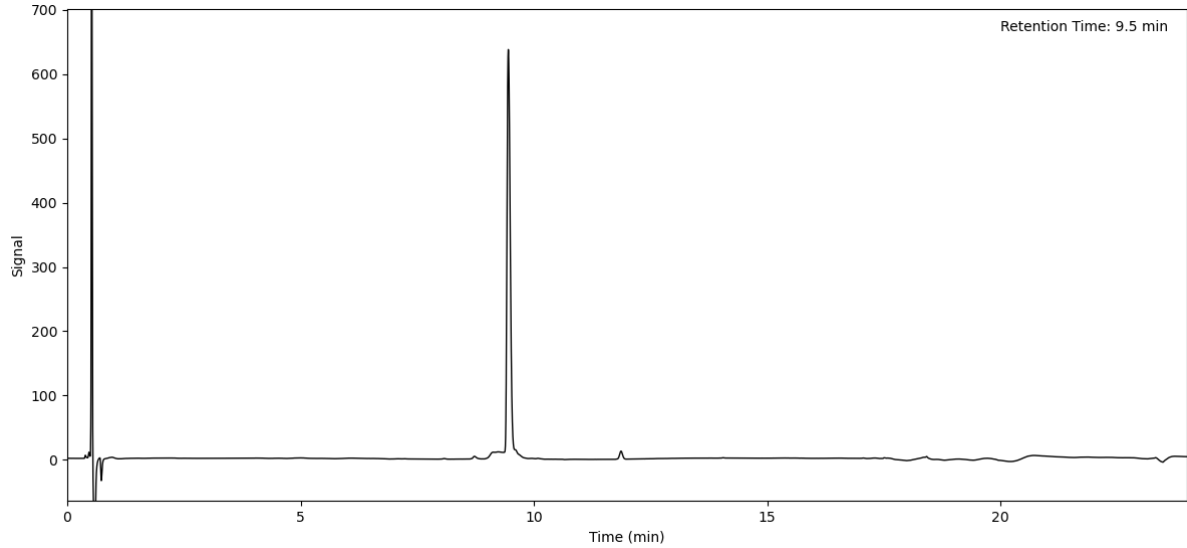
Peptide S5 (Method A)



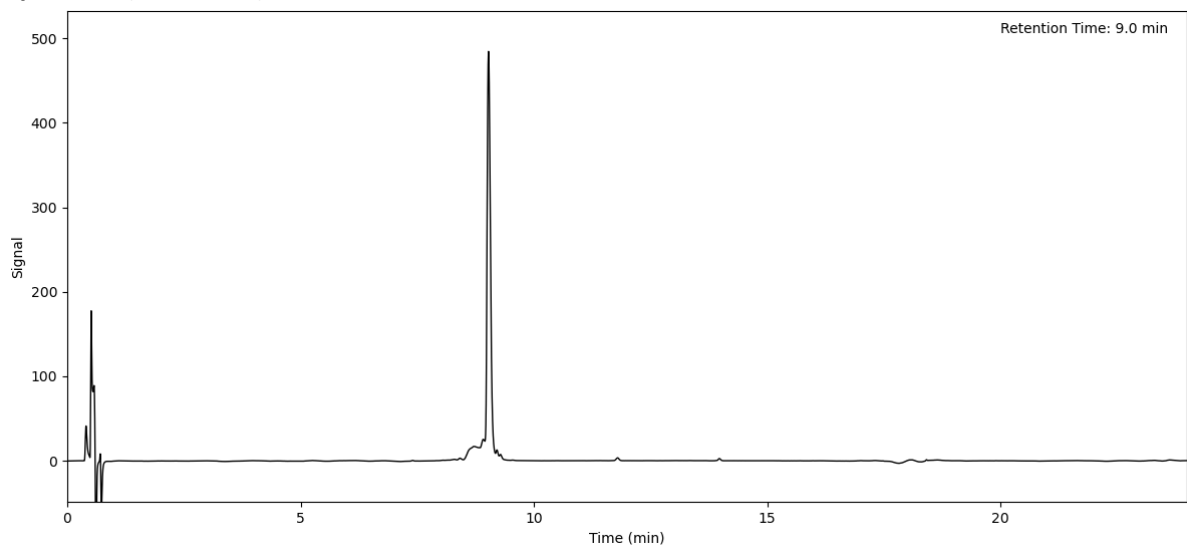
Peptide S6 (Method A)



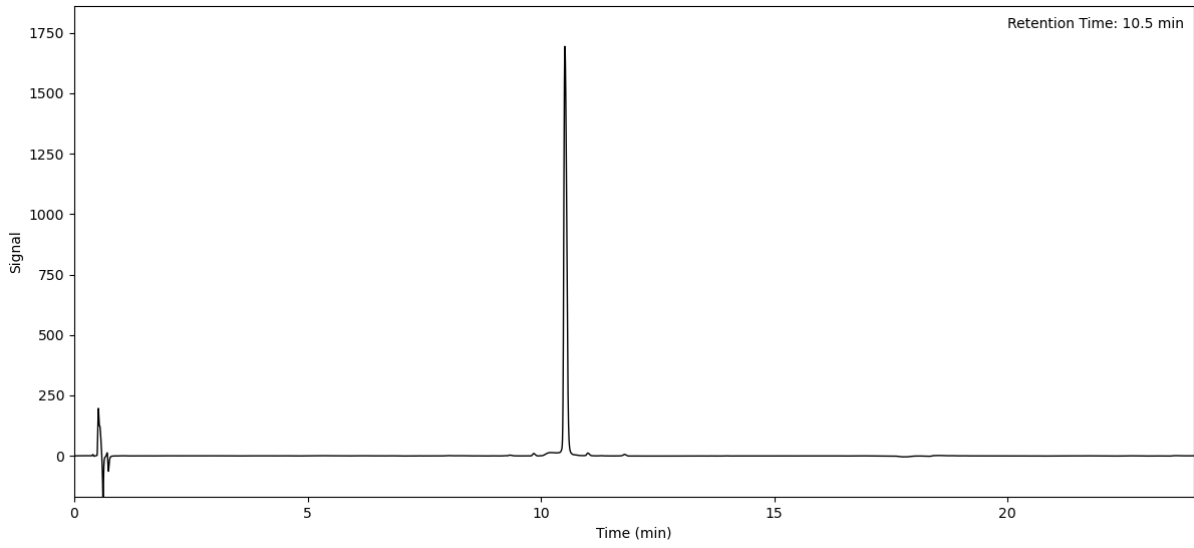
Peptide S7 (Method A)



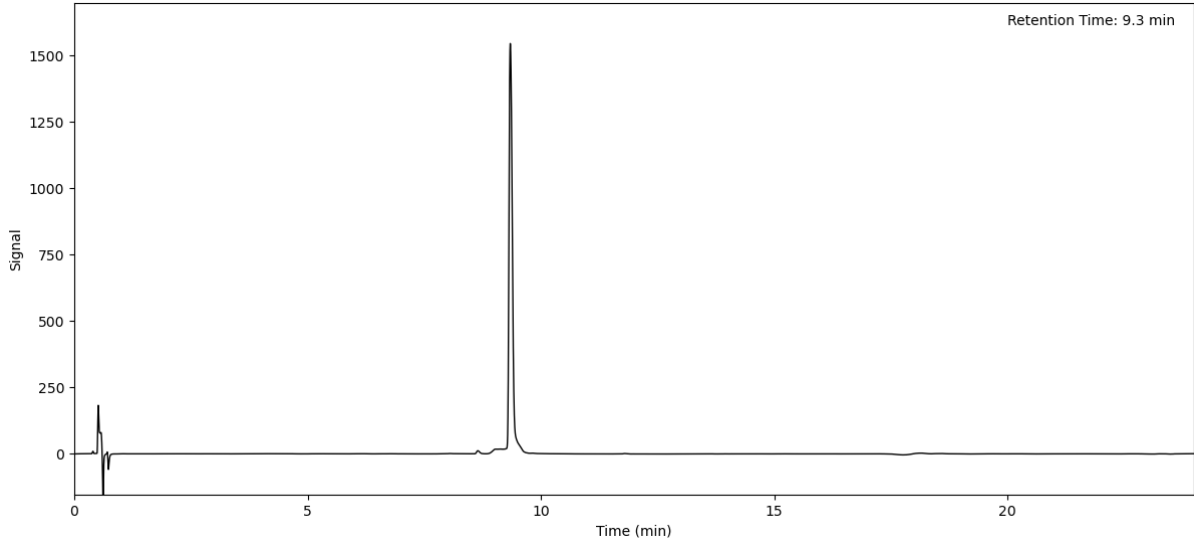
Peptide S8 (Method A)



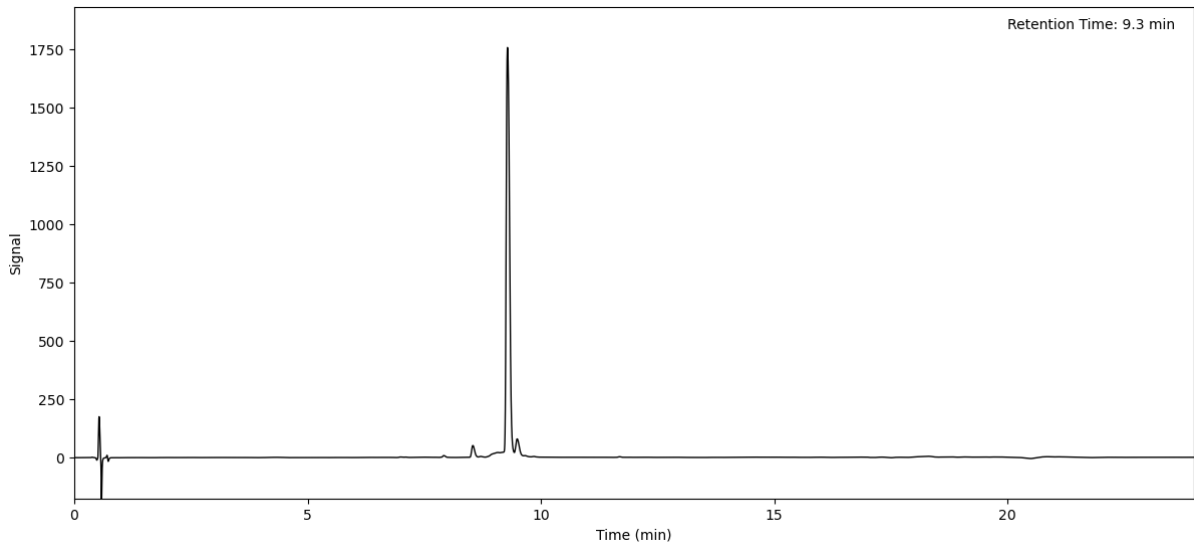
Peptide S9 (Method A)



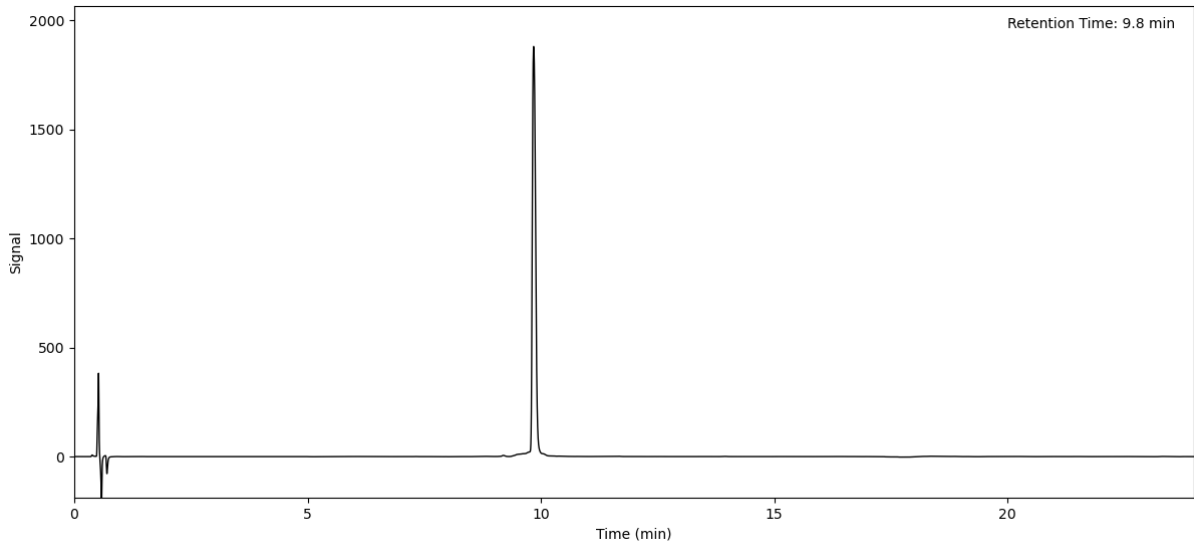
Peptide S10 (Method A)



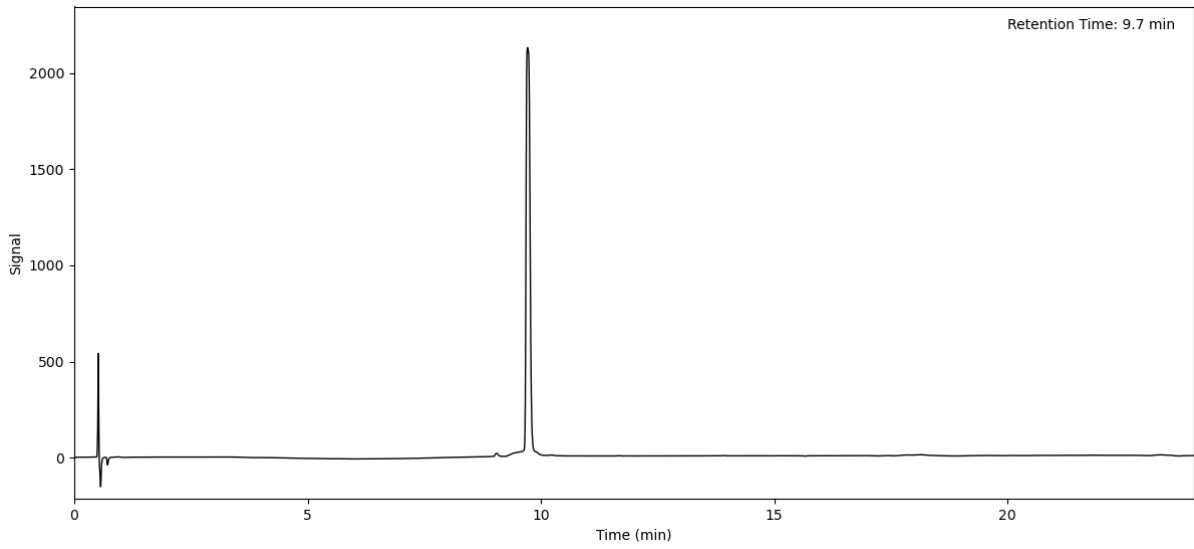
Peptide S11 (Method A)



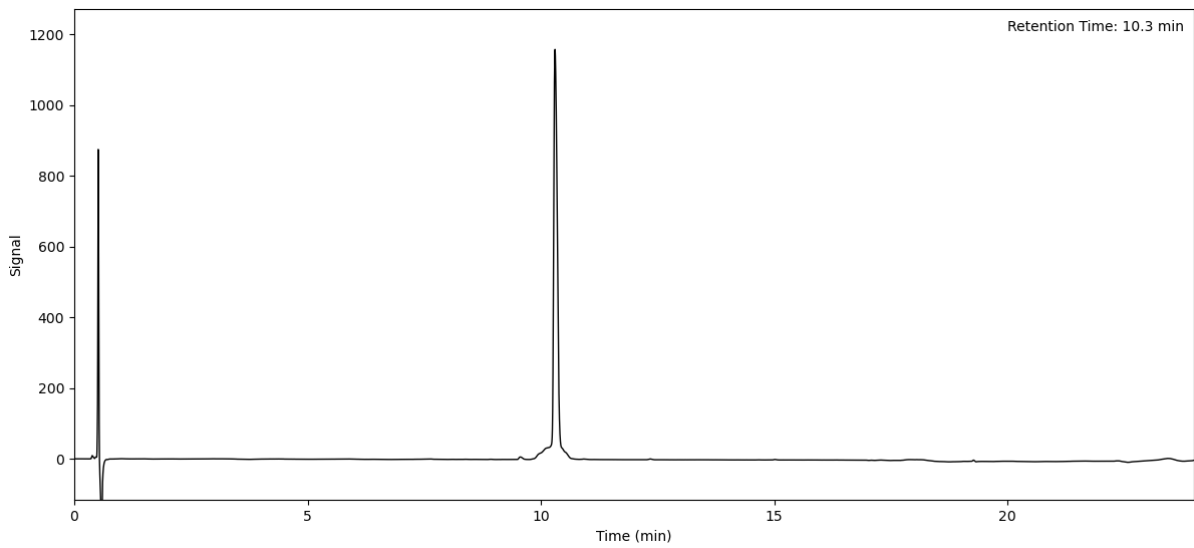
Peptide S12 (Method A)



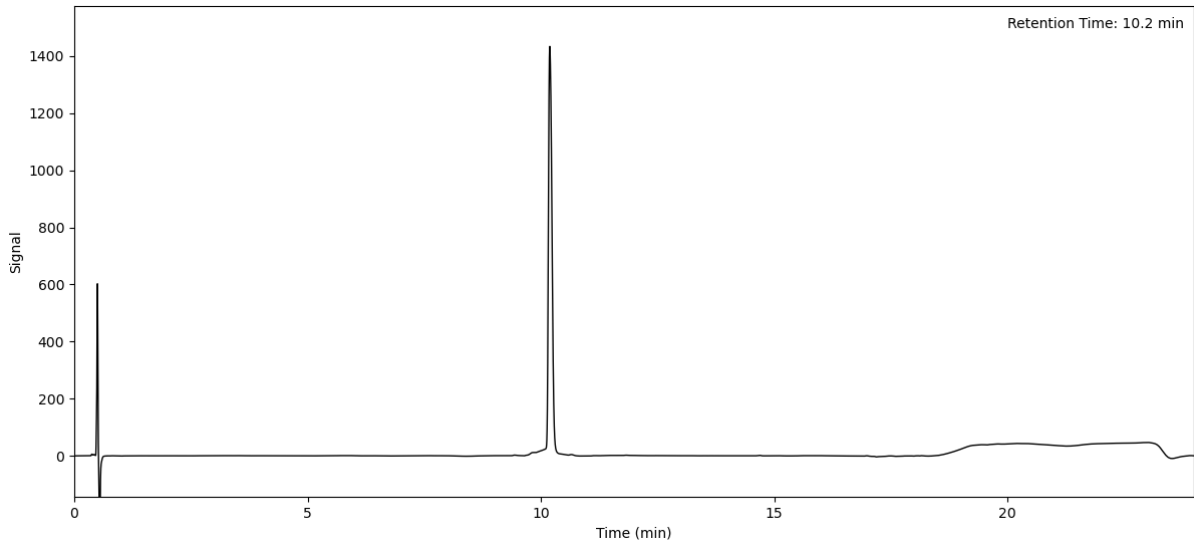
Peptide S13 (Method A)



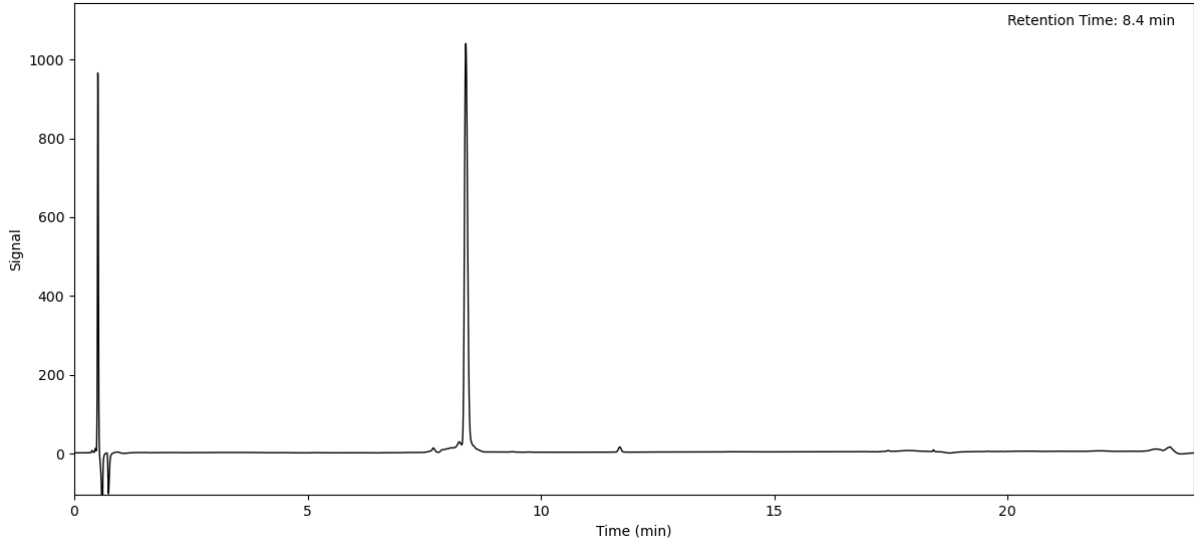
Peptide S14 (Method A)



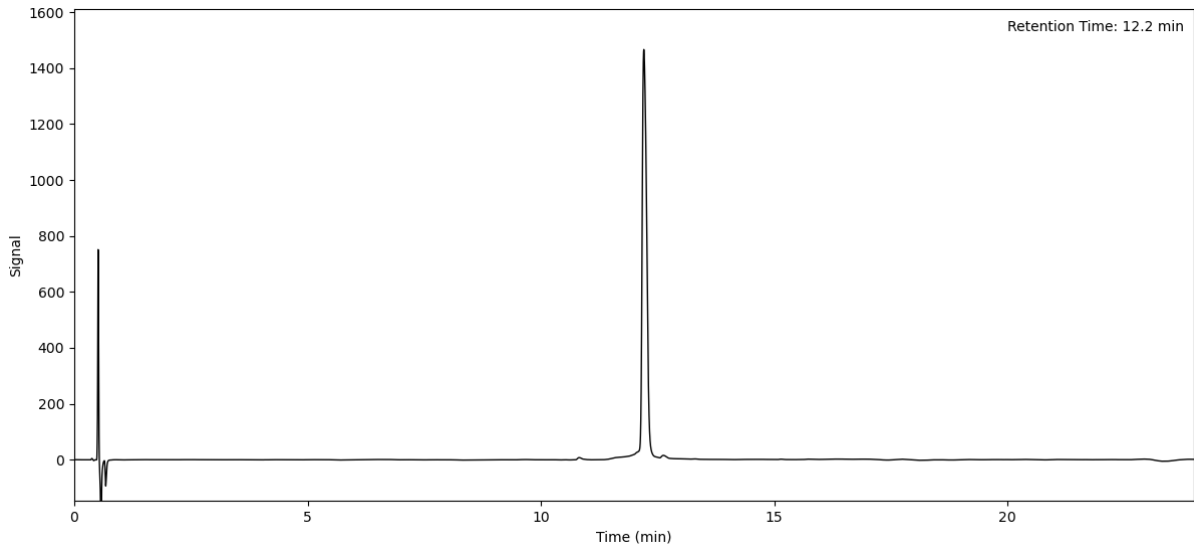
Peptide S15 (Method A)



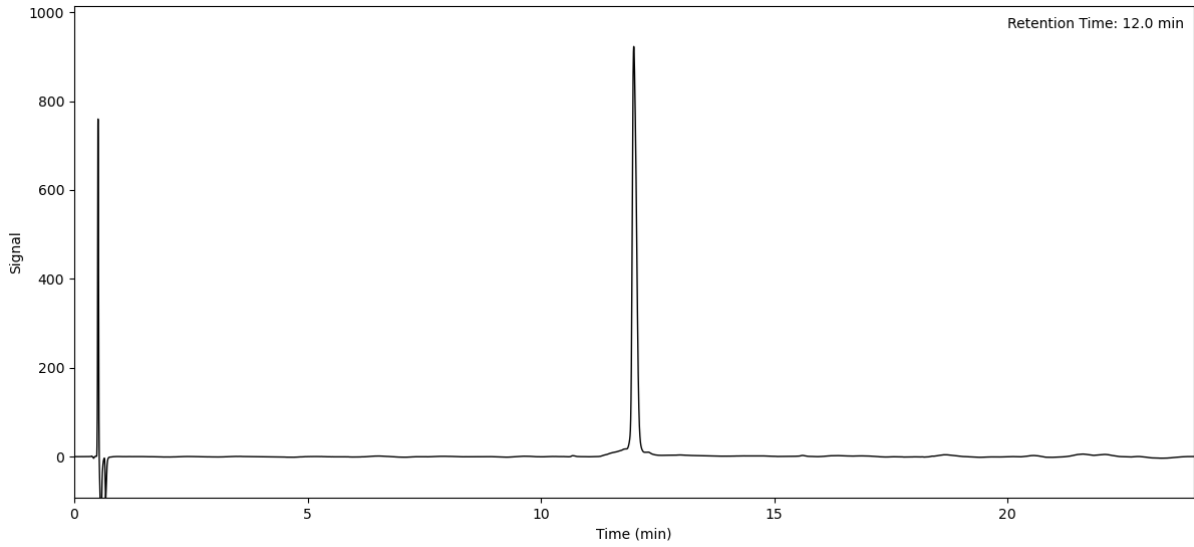
Peptide S16 (Method A)



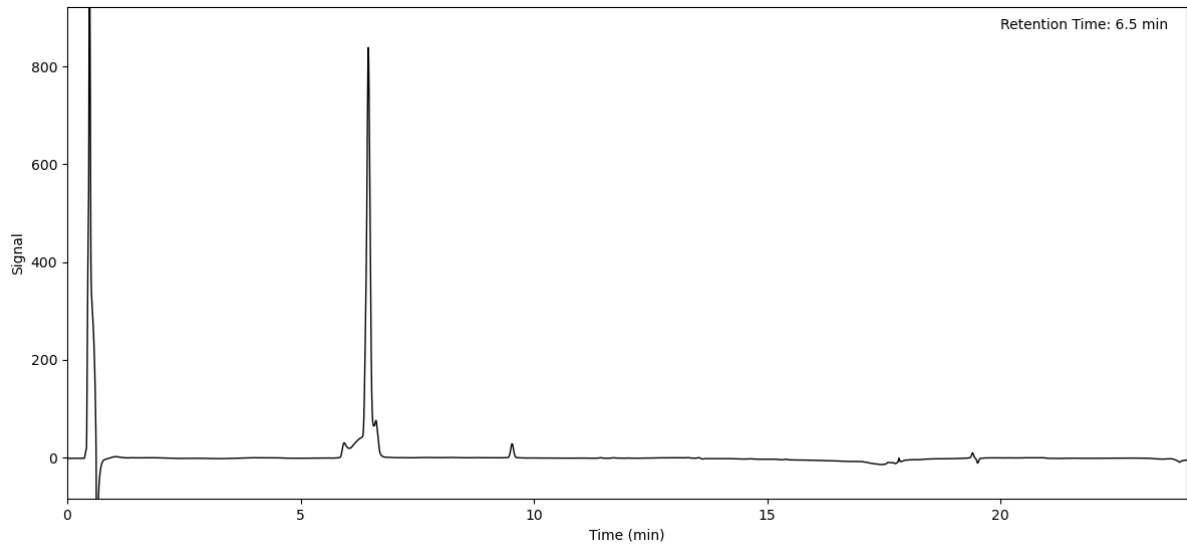
Peptide S17 (Method A)



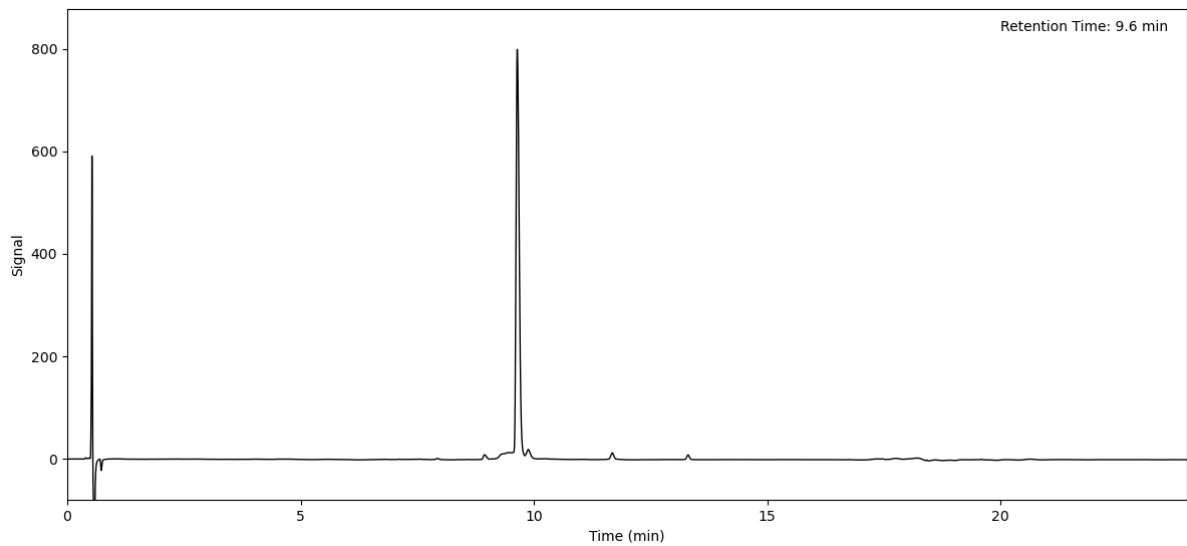
Peptide S18 (Method A)



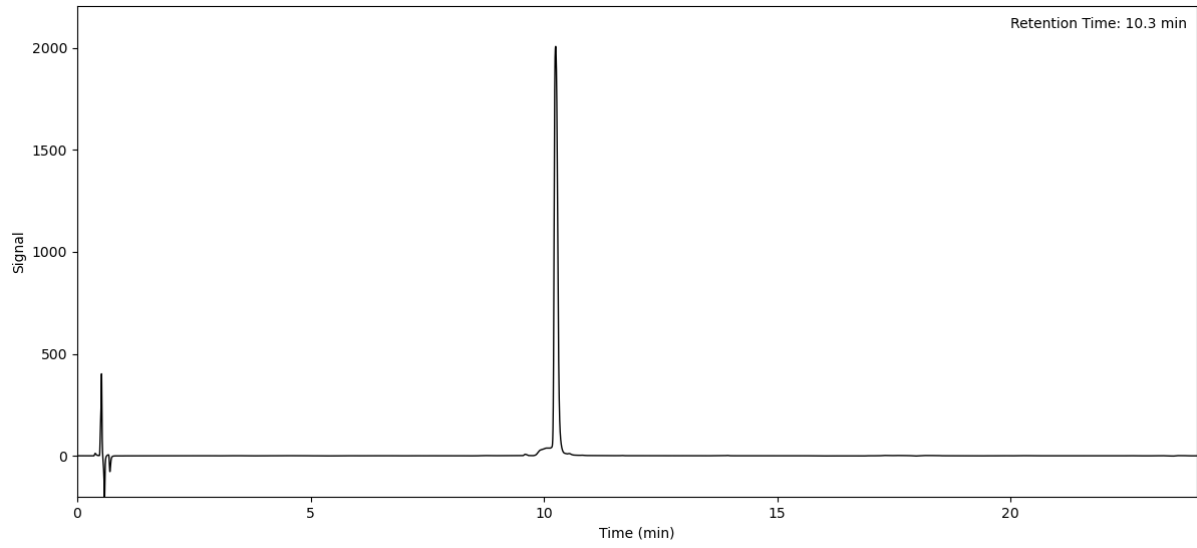
Peptide S19 (Method A)



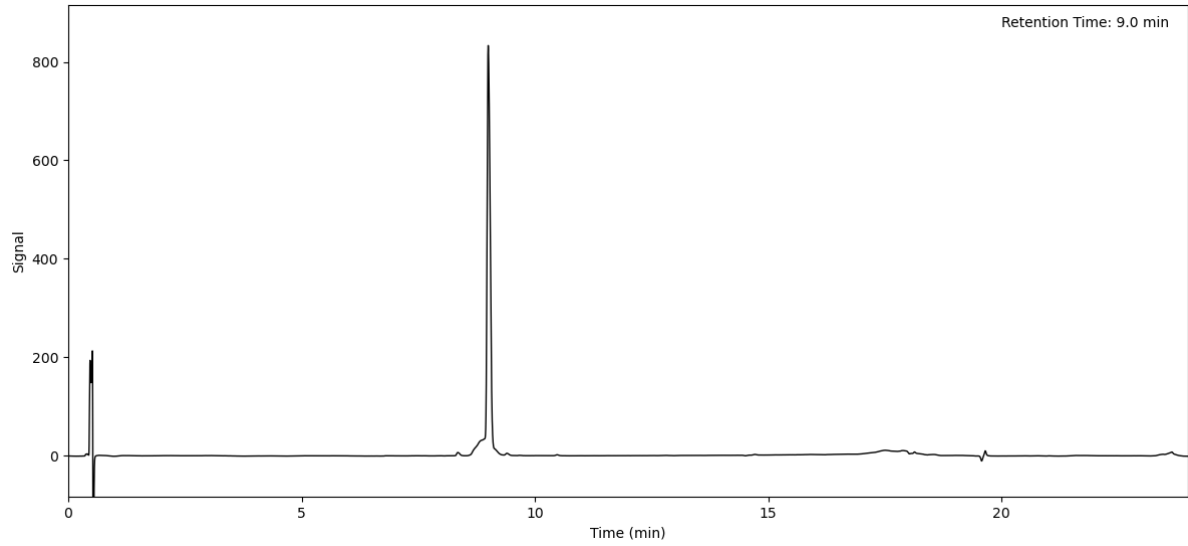
Peptide S20 (Method A)



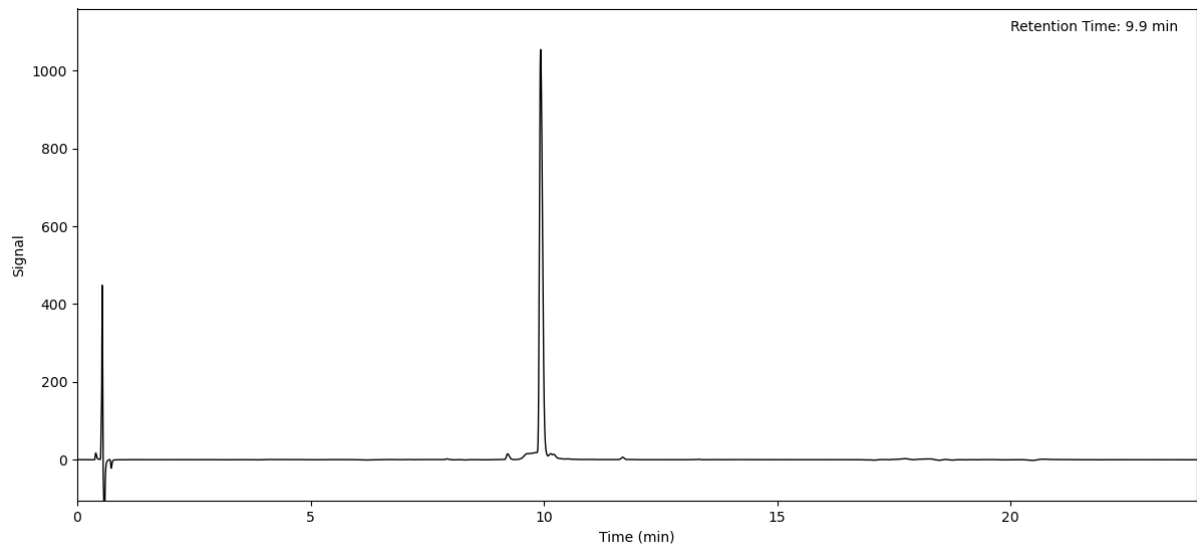
Peptide S21 (Method A)



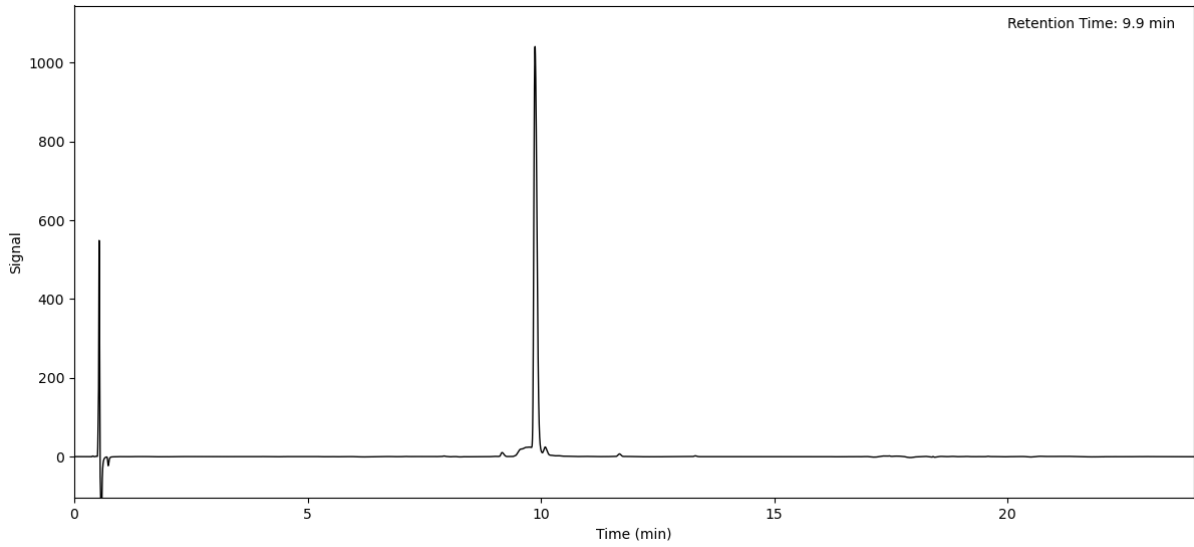
Peptide S22 (Method A)



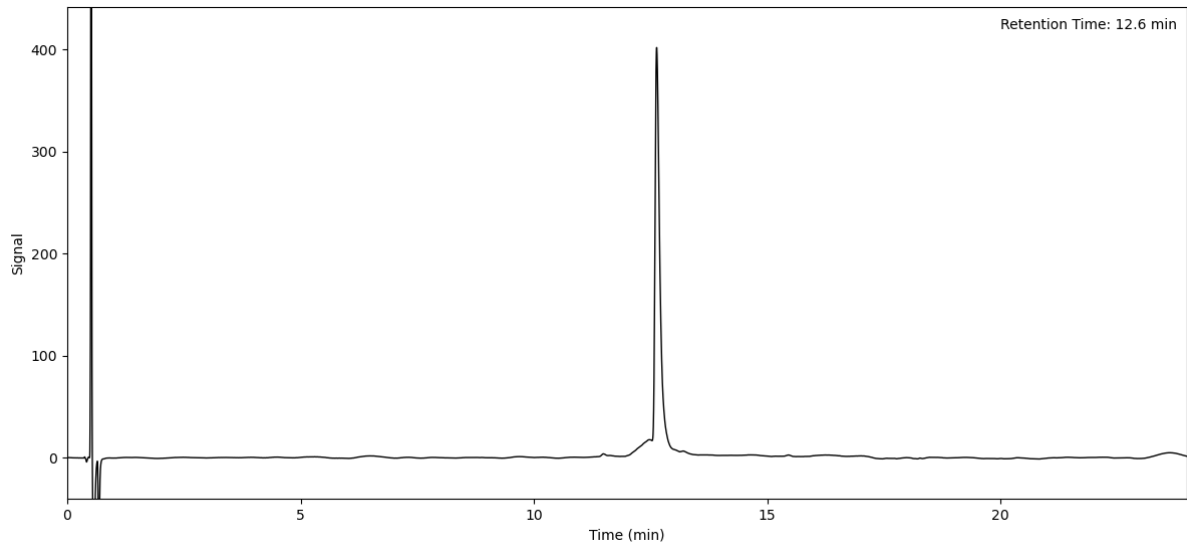
Peptide S23 (Method A)



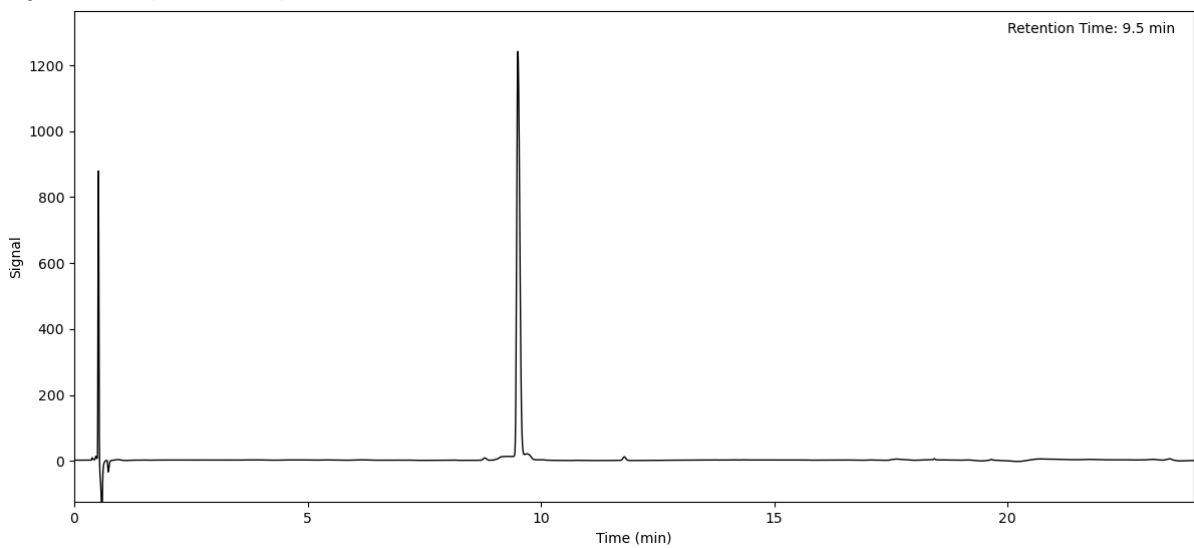
Peptide S24 (Method A)



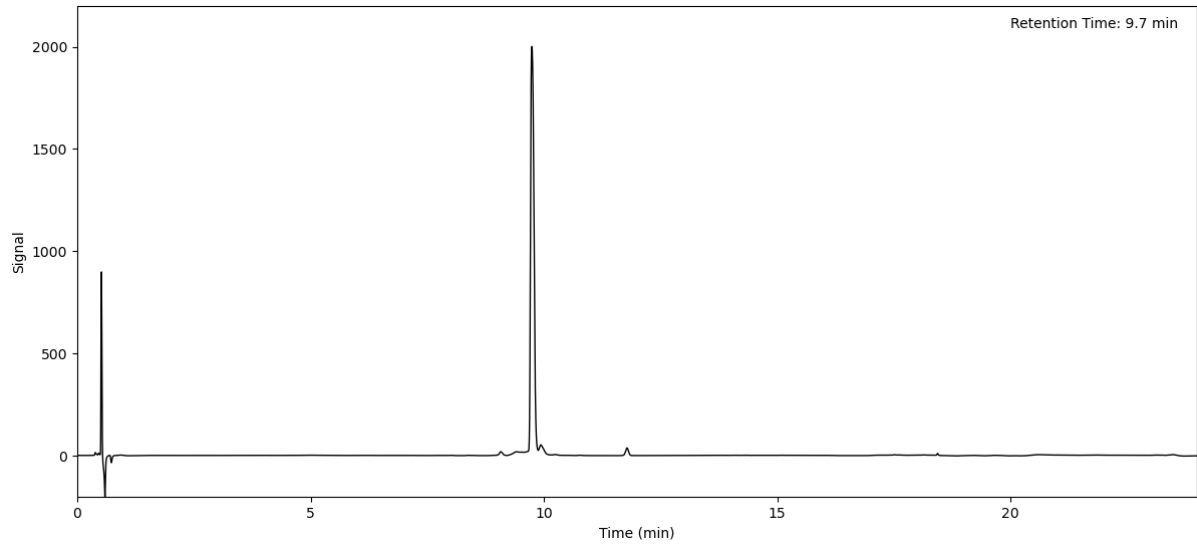
Peptide S25 (Method A)



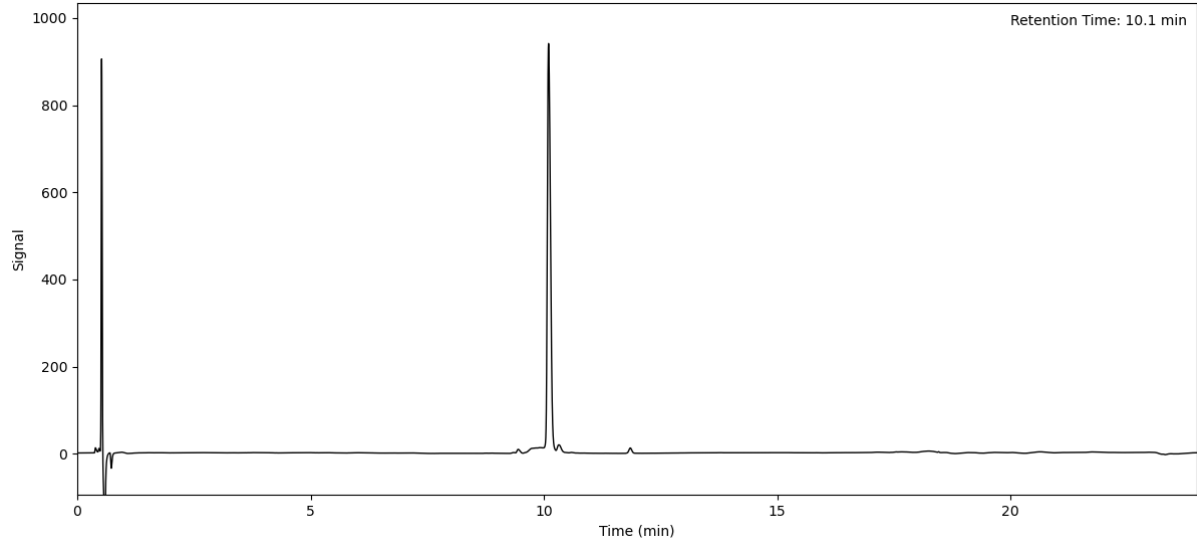
Peptide S26 (Method A)



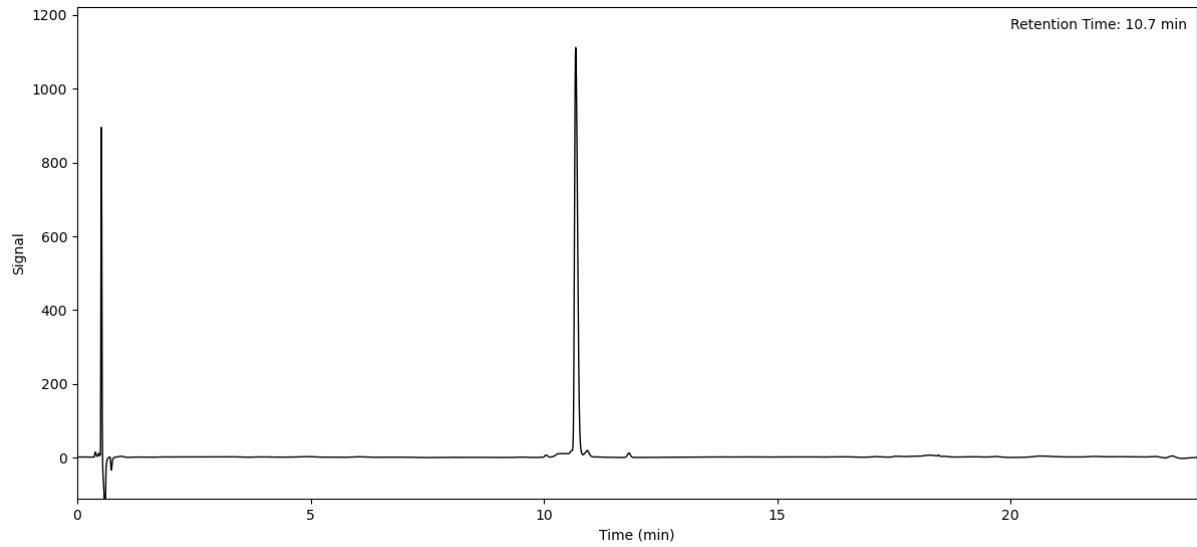
Peptide S27 (Method A)



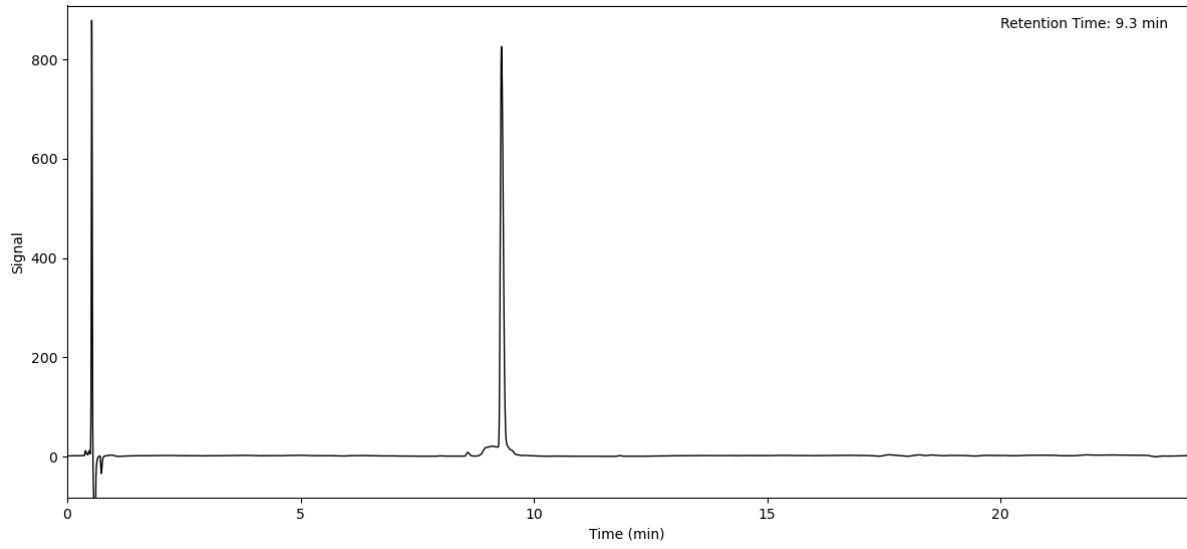
Peptide S28 (Method A)



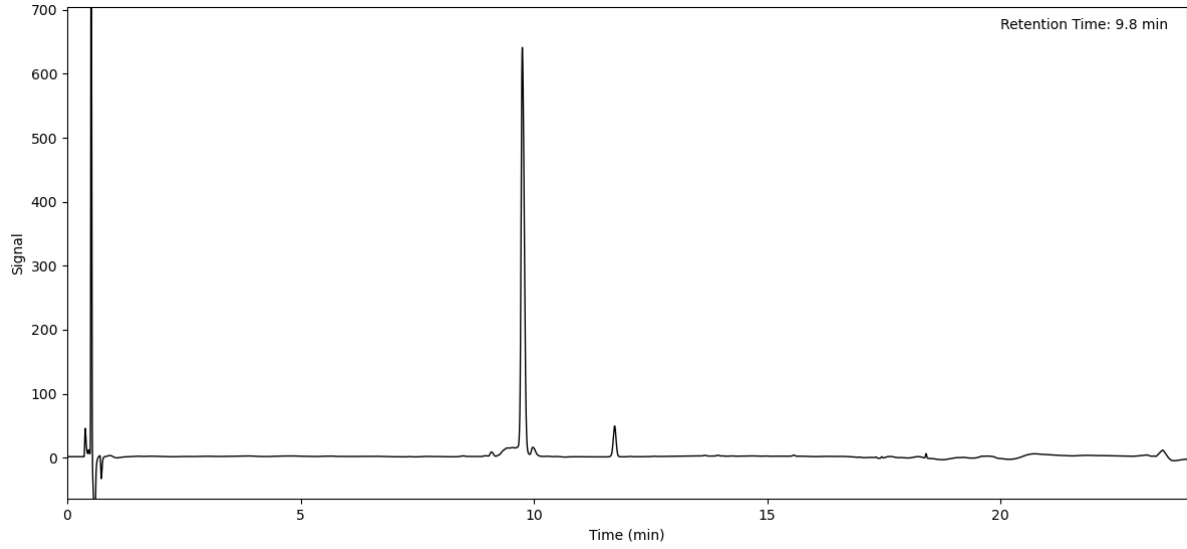
Peptide S29 (Method A)



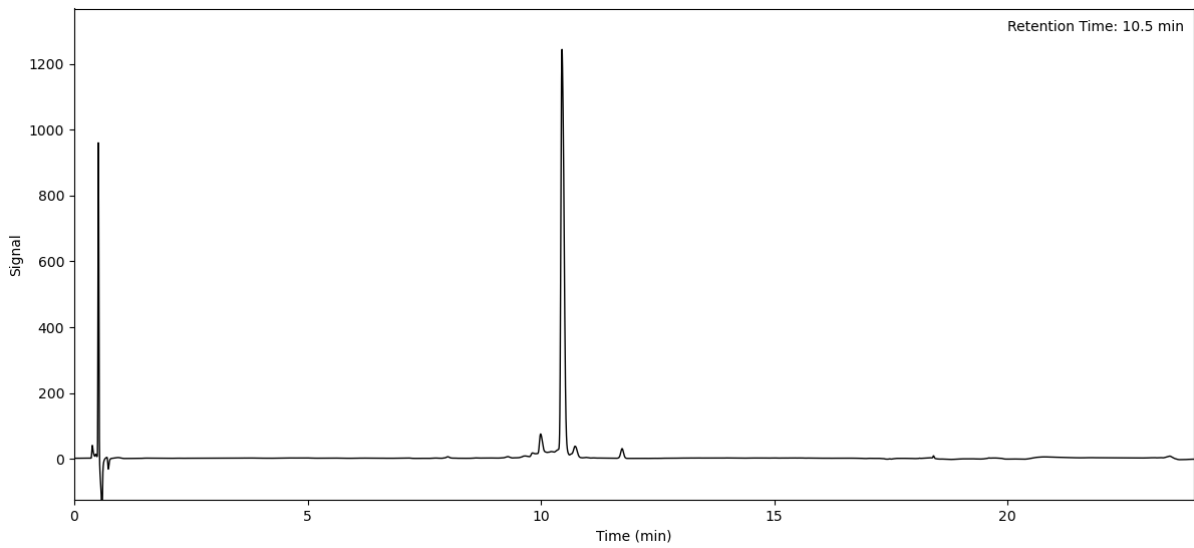
Peptide S30 (Method A)



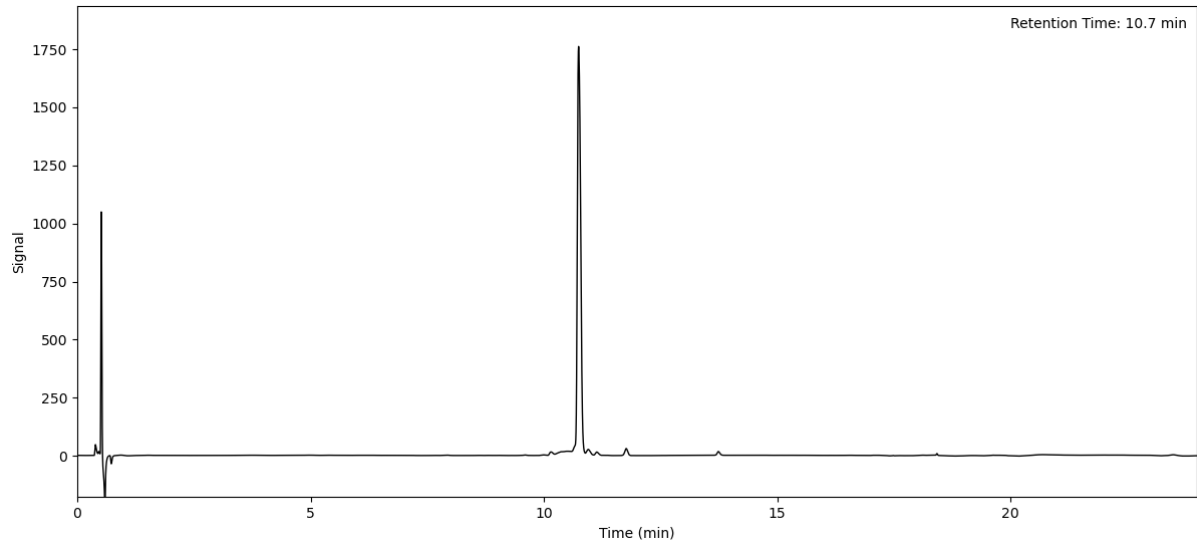
Peptide S31 (Method A)



Peptide S32 (Method A)



Peptide S33 (Method A)



8.2. Publication List

The results presented in this thesis contributed to the following publications:

Krzyzanowski, A., Esser, L.M., Willaume, A., Prudent, R., Peter, C., 't Hart P. and Waldmann H., 2022. Development of Macrocyclic PRMT5 Adaptor Protein Interaction Inhibitors. *Journal of Medicinal Chemistry*, 65(22), pp.15300-15311.

Krzyzanowski, A., Gasper, R., Adihou, H., 't Hart, P. and Waldmann, H., 2021. Biochemical Investigation of the Interaction of pICln, RioK1 and COPR5 with the PRMT5–MEP50 Complex. *ChemBioChem*, 22(11), pp.1908-1914.

The author's work performed during the doctoral studies also contributed to the following articles:

Flegel, J., Shaaban, S., Jun, Z.J., Schulte, B., Lian, Y., **Krzyzanowski, A.**, Metz, M., Schneidewind, T., Wesseler, F., Flegel, A., Reich, A., Brause, A., Xue, G., Zhang, M., Dötsch, L., Stender, I.D., Hoffmann, J.E., Scheel, R., Janning, P., Rastinejad, F., Schade, D., Strohmann, C., Antonchick, A.P., Sievers, S., Moura-Alves, P., Ziegler, S. and Waldmann H., 2022. The Highly Potent AhR Agonist Picoberin Modulates Hh-Dependent Osteoblast Differentiation. *Journal of Medicinal Chemistry*, DOI: 10.1021/acs.jmedchem.2c00956.

Grigalunas, M., Patil, S., **Krzyzanowski, A.**, Pahl, A., Flegel, J., Schölermann, B., Xie, J., Sievers, S., Ziegler, S. and Waldmann, H., 2022. Unprecedented Combination of Polyketide Natural Product Fragments Identifies the New Hedgehog Signaling Pathway Inhibitor Grismonone. *Chemistry–A European Journal*, e202202164.

Akbarzadeh, M., Flegel, J., Patil, S., Shang, E., Narayan, R., Buchholzer, M., Kazemein Jasemi, N.S., Grigalunas, M., **Krzyzanowski, A.**, Abegg, D., Shuster, A., Potowski, M., Karatas, H., Karageorgis, G., Mosaddeghzadeh, N., Zischinsky, M.L., Merten, C., Golz, C., Brieger, L., Strohmann, C., Antonchick, A.P., Janning, P., Adibekian, A., Goody, R.S., Reza Ahmadian, M., Ziegler, S. and Waldmann, H., 2022. The Pseudo-Natural Product Rhonin Targets RHOGDI. *Angewandte Chemie International Edition*, 61(18), e202115193.

't Hart, P., Hommen, P., Noisier, A., **Krzyzanowski, A.**, Schüler, D., Porfetye, A.T., Akbarzadeh, M., Vetter, I.R., Adihou, H. and Waldmann, H., 2021. Structure based design of bicyclic peptide inhibitors of RbAp48. *Angewandte Chemie International Edition*, 60(4), pp.1813-1820.

't Hart, P., Openy, J., **Krzyzanowski, A.**, Adihou, H. and Waldmann, H., 2019. Hot-Spot Guided Design of Macrocyclic Inhibitors of the LSD1-CoREST1 Interaction. *Tetrahedron*, 75(48), p.130685.

8.3. Eidesstattliche Versicherung (Affidavit)

Krzyzanowski, Adrian

214929

Name, Vorname
(Surname, first name)

Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

“Targeting PRMT5 Protein-Protein Interactions with Peptidic Modulators”

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.

Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Ort, Datum
(Place, date)

Unterschrift
(Signature)