Dissertation
zur Erlangung der Doktorwürde
an der
Gesamtfakultät für Mathematik,
Ingenieur- und Naturwissenschaften
der
Ruprecht-Karls-Universität Heidelberg

Thema:

The WW domain scaffold as model system for *de novo* design of miniaturized phosphate receptors and phosphatases

vorgelegt von:

Christina Lindner

Heidelberg 2024

Gutachter

1. *Professor Dr. Christian Klein*
Institut für Pharmazie und Molekulare Biotechnologie, Ruprecht-Karls-Universität Heidelberg

2. *Jun. Professor Dr. Franziska Thomas*
Organisch-Chemisches Institut, Ruprecht-Karls-Universität Heidelberg

Die in der Dissertation beschriebenen Arbeiten wurden unter der Anleitung und Betreuung von Frau Jun.-Prof. Dr. Franziska Thomas im Zeitraum von Juni 2020-Dezember 2023 am Organisch-Chemischen Institut der Ruprecht-Karls-Universität Heidelberg durchgeführt.

Selbstständigkeitserklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit mit dem Titel " The WW domain scaffold as model system for de novo design of miniaturized phosphate receptors and phosphatases" selbstständig verfasst habe und nur die von mir angegebenen Quellen und Hilfsmittel verwendet habe. Wörtlich oder sinngemäß aus anderen Werken entnommene Stellen habe ich unter Angabe der Quellen kenntlich gemacht.
Weiterhin erkläre ich, dass diese Abschlussarbeit, auch nicht in Auszügen, an dieser oder an einer anderen Universität oder Hochschule eingereicht wurde.

Heidelberg, der 23.08.2024

Christina Lindner

To my parents Elisabeth and Wilhelm,
my sister Julia,
and my husband Gino

Abstract

The current goal of research in the field of synthetic chemistry is the design and production of peptide and protein structures that function as biological synthesis machines.[1] Initially, the *de novo* design of peptide catalysts focused on α-helical structural motifs, such as the coiled coil.[2,3] Given the relatively rigid and self-assembling nature of these scaffolds[4,5], scientific efforts are concentrated on the design of smaller single-chain β-sheet motifs, such as the WW domain.[6–9] Due to its properties as a small, independently folding protein motif with 34-40 amino acid residues, as well as its function as a protein interaction module with a flexible binding site, the WW domain was selected as a scaffold for the design of miniaturized proteins in this work.[10]

To establish the WW domain as a scaffold for miniprotein design, an iterative sequence analysis approach was initially employed to identify sequence-structure-stability-relationships, starting from a consensus sequence of 90 native WW domains. In accordance with the identified sequence-structure-stability relationships, three highly thermostable WW domain core scaffolds were designed, displaying melting temperatures of 89 and 93 °C in thermal CD denaturation experiments. Multi-dimensional high-resolution NMR experiments of one highly thermostable scaffold peptide demonstrated the typical WW domain structure, comprising three anti-parallel β-sheets and additional structural features that contribute to its high thermostability. As a proof of concept for the design of functional miniproteins, binding motifs of the three different WW domain groups were grafted onto the thermostable WW domain core scaffold to obtain thermostable WW domains with group-specific binding properties. Binding studies on the designed group-specific thermostable WW domains revealed micromolar-to-sub-micromolar binding affinities to their respective ligands. In addition to the introduction of native functions into the thermostable WW domain core scaffold, binding motifs for non-native functions, such as metal- and organophosphate binding, were grafted onto a scaffold peptide. This resulted in the design of *de novo* thermostable WW domains with micromolar-to-submicromolar binding affinities to their respective ligands. Following the successful design of an adenosine-triphosphate (ATP)-binding thermostable variant (WW-HS-ATP), further peptide variants were designed with diverse sequence modifications to study the influence on binding affinity to ATP, ADP and AMP. All *de novo* designed variants, as well as WW-HS-ATP, demonstrated micromolar binding affinities to ATP, ADP and AMP. To ascertain whether one of the ATP-binding peptides exhibits enzymatic activity in terms of phosphatase activity, exploratory studies were conducted, which identified certain candidates that showed minimal phosphatase-like activity. To obtain thermostable WW domain miniproteins with enzymatic activity, further sequence redesign and additional studies are required. In conclusion, this work successfully demonstrated that the establishment of sequence-structure-stability relationships of native WW domains, resulted in the design of highly thermostable WW domain core scaffolds, which could be further functionalized with native and non-native functions, while maintaining thermostable properties. The designed thermostable WW domain scaffolds could serve as a template for future β-sheet miniprotein design with diverse receptor, sensor or catalytic properties.

Kurzzusammenfassung

Das derzeitige Ziel der Forschung auf dem Gebiet der synthetischen Chemie ist der Entwurf und die Herstellung von Peptid- und Proteinstrukturen, die als biologische Synthesemaschinen fungieren.[1] Ursprünglich konzentrierte sich das Design von Peptidkatalysatoren auf α-helikale Strukturmotive, wie z. B. das Coiled Coil.[2,3] In Anbetracht der relativ starren und selbstfaltenden Natur dieser Gerüste[4,5] konzentrieren sich unsere wissenschaftlichen Bemühungen auf das Design kleinerer, einzelsträngiger β-Faltblatt-Motive, wie das der WW-Domäne.[6–9] Aufgrund ihrer Eigenschaften als kleines, unabhängig faltendes Proteinmotiv mit 34-40 Aminosäureresten, sowie ihrer Funktion als Proteininteraktionsmodul mit einer flexiblen Bindungsstelle wurde die WW-Domäne in dieser Arbeit als Gerüst für das Design von miniaturisierten Proteinen ausgewählt.[10]

Um die WW-Domäne als Gerüst für das Design von Miniproteinen zu etablieren, wurde zunächst ein iterativer Sequenzanalyse Ansatz angewendet, um Sequenz-Struktur-Stabilitäts-Beziehungen zu identifizieren, ausgehend von einer Konsensussequenz von 90 nativen WW-Domänen. In Übereinstimmung mit den Sequenz-Struktur-Stabilitäts-Beziehungen wurden drei außergewöhnlich thermostabile WW-Domänen-Kerngerüste entworfen, die in thermischen CD-Denaturierungsexperimenten Schmelztemperaturen von 89 und 93 °C aufwiesen. Multidimensionale, hochauflösende NMR-Experimente eines der hoch thermostabilen Kerngerüste zeigten die typische WW-Domänenstruktur, die drei antiparallele β-Faltblätter und zusätzliche strukturelle Merkmale umfasst, die zu seiner hohen Thermostabilität beitragen. Als Konzeptnachweis für das Design funktioneller Miniproteine wurden Bindungsmotive der drei verschiedenen WW-Domänengruppen in das thermostabile Gerüst eingebracht, um thermostabile WW-Domänen mit gruppenspezifischen Bindungseigenschaften zu erhalten. Bindungsstudien an den entworfenen gruppenspezifischen thermostabilen WW-Domänen zeigten mikromolare bis submikromolare Bindungsaffinitäten zu ihren spezifischen Liganden. Zusätzlich zur Einführung nativer Funktionen in das Kerngerüst der thermostabilen WW-Domänen wurden Bindungsmotive für nicht-native Funktionen, wie Metall- und Organophosphat-Bindung, in ein Gerüstpeptid eingebracht. Dies führte zu neuartigen thermostabilen WW-Domänen mit mikromolaren bis submikromolaren Bindungsaffinitäten zu ihren jeweiligen Liganden. Nach dem erfolgreichen Design einer Adenosintriphosphat (ATP)-bindenden thermostabilen Variante (WW-HS-ATP) wurden weitere Peptidvarianten mit verschiedenen Sequenzmodifikationen entworfen, um den Einfluss auf die Bindungsaffinität zu ATP, ADP und AMP zu untersuchen. Alle Peptidvarianten sowie WW-HS-ATP zeigten mikromolare Bindungsaffinitäten zu ATP, ADP und AMP. Um zu bewerten, ob eines der ATP-bindenden Peptide enzymatische Aktivität im Sinne einer Phosphatase-Aktivität aufweist, wurden explorative Studien durchgeführt, die einige Kandidaten identifizierten, die eine minimale Phosphatase-ähnliche Aktivität zeigten. Um zukünftig thermostabile WW-Domänen-Miniproteine mit enzymatischer Aktivität zu erhalten, sind weitere Sequenzmodifikationen und Studien erforderlich. Zusammenfassend wurde in dieser Arbeit erfolgreich gezeigt, dass die Etablierung von Sequenz-Struktur-Stabilitäts-Beziehungen von nativen WW-Domänen zum Design von hochgradig thermostabilen WW-Domänen Kerngerüsten geführt hat, die mit nativen und nicht-nativen Funktionen weiter funktionalisiert werden konnten, wobei die thermostabilen Eigenschaften erhalten blieben. Die thermostabilen WW Domänen Kerngerüste können zukünftig als Vorlage für die Entwicklung von β-Faltblatt-Miniproteinen mit verschiedenen Rezeptor-, Sensor- oder katalytischen Eigenschaften dienen.

Table of Contents

1 Introduction

Proteins are ubiquitous macromolecules that are involved in countless biological processes in living organisms, serving as catalysts, signaling molecules, metabolites, or structural components.[11,12] They are composed of 20 native amino acids that are assembled in a linear sequence called the primary structure. Proteins consist of a polypeptide backbone with attached side chains, which provide different properties to the protein like polarity, hydrophobicity, charge and reactivity.[12] In addition to chemical characteristics, the order of the amino acids influences protein folding, thus determining its function. Many variables specify the folded state, including sequence, backbone topology and sidechain conformation.[13] To generate proteins with novel or enhanced function protein engineering is widely used. It is a method of altering the sequence of a protein to discover and obtain the desired property, such as a change in substrate specificity, increased stability to temperature, organic solvents, and/or extreme pH.[14]

A current goal of research in the field of synthetic biology, which includes several research areas such as drug discovery, chemical biology, and engineering, is the design and production of peptide- and protein structures that function as biological synthesis machines or have novel properties and functions.[1] An important protein engineering technique is the directed evolution of natural proteins, which is used to design e.g. catalytically active antibodies or modified enzymes.[15–17] Although the evolutionary design serves as a powerful starting point, designed native enzymes may also exhibit undesirable properties, such as a lack of structural or thermal stability.[18] Thus, bottom-up enzyme design aims at the miniaturization of suitable protein scaffolds, the characterization of sequence-structure-function relationships, and easily adaptable substrate specificity of peptide -or protein catalysts.[1] Many rationally designed peptide catalysts are based on α-helical structural motifs such as the coiled-coil.[2,5,19] Due to the fact that these are relatively rigid and often self-assembling scaffolds, current scientific efforts are focused on smaller β-sheet motifs such as the Src homology 3 (SH3) or the rsp5 (WW) domain.[20,21] In general, peptide scaffolds used for rational *de novo* enzyme design are often found as protein-protein interaction (PPI) modules in nature.[15–18] *De novo* enzyme design approaches focus on native protein scaffolds that are equipped with desired physicochemical properties and onto which active sites are grafted.[18,21] Many of these peptide scaffolds derive from large globular proteins and when comprising ≤ 100 amino acids are referred to as miniproteins, which can consist of two or more secondary structure elements that are stabilized by either noncovalent interactions, covalent cross-linking or metal binding. Prominent representatives of miniprotein structures are the polyproline-II-loop-α-helix fold of the avian pancreatic peptide hormone (aPP), the ββα-fold structure of DNA-binding zinc fingers or β-sheet-based structures such as β-hairpins or Trp-zippers.[3]

In addition to their promising application in synthetic biology, miniproteins have great potential in medicinal chemistry as drug candidates and can serve as adequate alternatives to small molecule or protein therapeutics. Due to their medium size, they can be easily functionalized with specific binding motifs to interact with high selectivity and biological activity with targets such as receptors or protein-protein interactions (PPIs), which are usually challenging for small molecules.[8,9,22] Furthermore, miniproteins can access intracellular targets or tissues that are inaccessible to protein therapeutics such as antibodies.[8,9]

The aim of this work is to first generate sequence-structure-stability relationships for WW domains, in order to establish this β-sheet motif for the subsequent design of thermostable and functional miniproteins.

2 State of the Art

2.1 Miniproteins

2.1.1 Definition & Classification

As the name already implies, miniproteins are characterized by their small size of about 1 – 10 kDa and are known to be stably folded molecules with versatile functions. The small size can be advantageous in terms of facilitated chemical synthesis and the possibility to introduce modifications for tailored functionalization, such as drug conjugation, and their generally wide accessibility to diverse protein targets in diverse and challenging biological compartments.[8,22]

Miniproteins can be classified into three subcategories based on the physicochemical properties of their scaffolds (cf. Figure 1). One subcategory are the hydrophobic core miniproteins, which consist of α-helical or β-sheet based structural motifs that are located around the hydrophobic core, which determines the folding process upon solvent exclusion. Prominent representatives of hydrophobic miniproteins are affibodies, adnectins, nanofittins, and fynomers. Another subcategory are the Cys-Cys disulfide (cystine)-reinforced miniproteins, whose folding and structural stability is based on cystines. Cystine-dense peptides (CDPs), such as animal and plant defense peptides in nature, cystine knots, Kunitz domains, and avimers belong to this highly diverse category of cystine-stabilized scaffolds. The third subcategory includes chemically stabilized miniproteins, which are small secondary structural motifs that all require chemical crosslinking for stabilization. Examples of this category include β-hairpins, stapled peptides consisting mostly of α-helices, and bicyclic peptides, composed of multiple loop regions.[8,9]

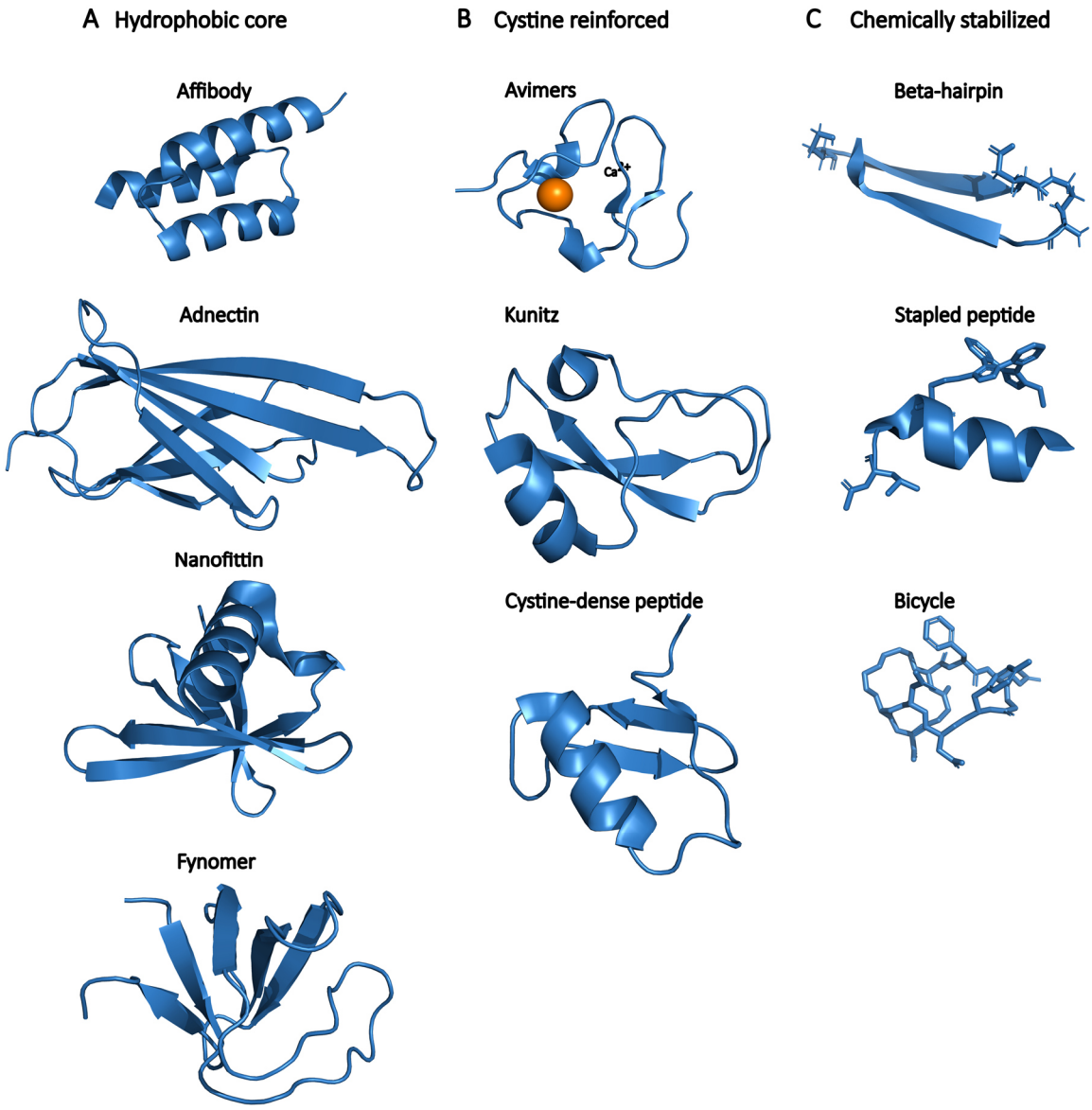


Figure 1: The three classes of miniproteins and their representatives. A) Hydrophobic core miniproteins with affibodies (pdb code: 3mzw)[23], adnectins (pdb code: 3qwq)[24], nanofittins (pdb code: 2xiw)[25] and fynomers (pdb code: 4afq)[26] as representative scaffolds. **B**) Cystine reinforced miniproteins include avimers (pdb code: 1ajj)[27], Kunitz domains (pdb code: 5c67)[28], and cystine-dense peptides (pdb code: 6atw)[29]. **C**) Chemically stabilized peptides with β-hairpins (pdb code: 5v63)[30], stapled peptides (pdb code: 5afg)[31] and bicyclic peptides (pdb code: 5eel)[32]. PDB files were accessed with RCSB PDB.[33] This image is based on the classification scheme and illustration of Crook *et al.* 2020.[8]

2.1.2 Approaches for (mini)protein design

In general, protein design involves two main strategies, the bottom-up and the top-down approach. The bottom-up approach implies *de novo* design of mostly functional proteins from scratch by considering different structural motifs and their respective sequence-to-structure-to-function relationships to build stabilized scaffolds with tailored functions, such as protein-based biosensors for antibody recognition or signaling molecules to modulate synthetic receptors.[34] It often includes the design of a hydrophobic packing, the incorporation of linker regions, charged partners, and compatible H-bonding groups or disulfide bonds to stabilize structure-specific intramolecular interactions and probably also specific functional residues.

Computational modeling, energy calculations, and visualization are also part of this approach.[35]

The top-down approach aims at functional engineering of existing sequence motifs of known proteins or peptides (cf. Figure 2).[35,36] This strategy also includes computational protein engineering, which relies on algorithms that enable structure prediction, protein backbone, and binding pocket design.[1] Protein surface design of folded proteins is studied to reveal the importance of hydrogen bonding and salt bridge interactions between surface groups.[37] Furthermore, it is often used to optimize the binding pocket and thus the ligand-binding affinity of enzymes. Fragmentation studies that generate subdomains to determine whether the resulting miniprotein folds independently or can assemble as oligomers to reconstitute the parent architecture are also part of the top-down protein engineering.[35]

Novel functions are often introduced by mutations within a rational or combinatorial design approach. The underlying principle of the rational design approach is to understand sequence-structure relationships by studying mutants. An example of mutational analysis for *in vitro* and *in vivo* is the alanine scan, where amino acids are mutated individually to alanine, resulting in a loss or reduction of function, including binding, catalysis, or stability.[35] Rational protein design represents a promising approach not only for the analysis of protein-protein interactions or the kinetics and thermodynamics of protein folding, but also for the design of miniprotein scaffolds with non-native functions, such as the development of a CD4-mimetic miniprotein that serves as an inhibitor of human immune deficiency virus type 1 (HIV-1).[9,35,38] Combinatorial protein design is a complementary approach to directed design, facilitating the understanding of sequence-structure compatibility and the discovery of novel sequences capable of folding into a specific structure.[13] It is a powerful tool for proteins where the understanding of molecular properties is incomplete. Two key elements of the approach are the generation of a library with the desired degree of diversity and the identification of sequences exhibiting “protein-like” properties in terms of structure and function, which can lead to a large number of sequences, up to 10^{12}. This approach can be used to identify “hits”, i.e. a few sequences that exhibit desirable characteristics, such as enhanced stability or strong binding properties. Due to the often extensive size of the protein library, combinatorial experiments frequently possess a high degree of complexity and are typically constrained to sequence and qualitative chemical properties. An advantage of this approach is the potential to expand the protein sequence database and to investigate proteins that are decoupled from the evolutionary requirements of natural proteins, as for example ubiquitin variants and proteins with special properties, such as self-assembly or metal-binding proteins.[13]

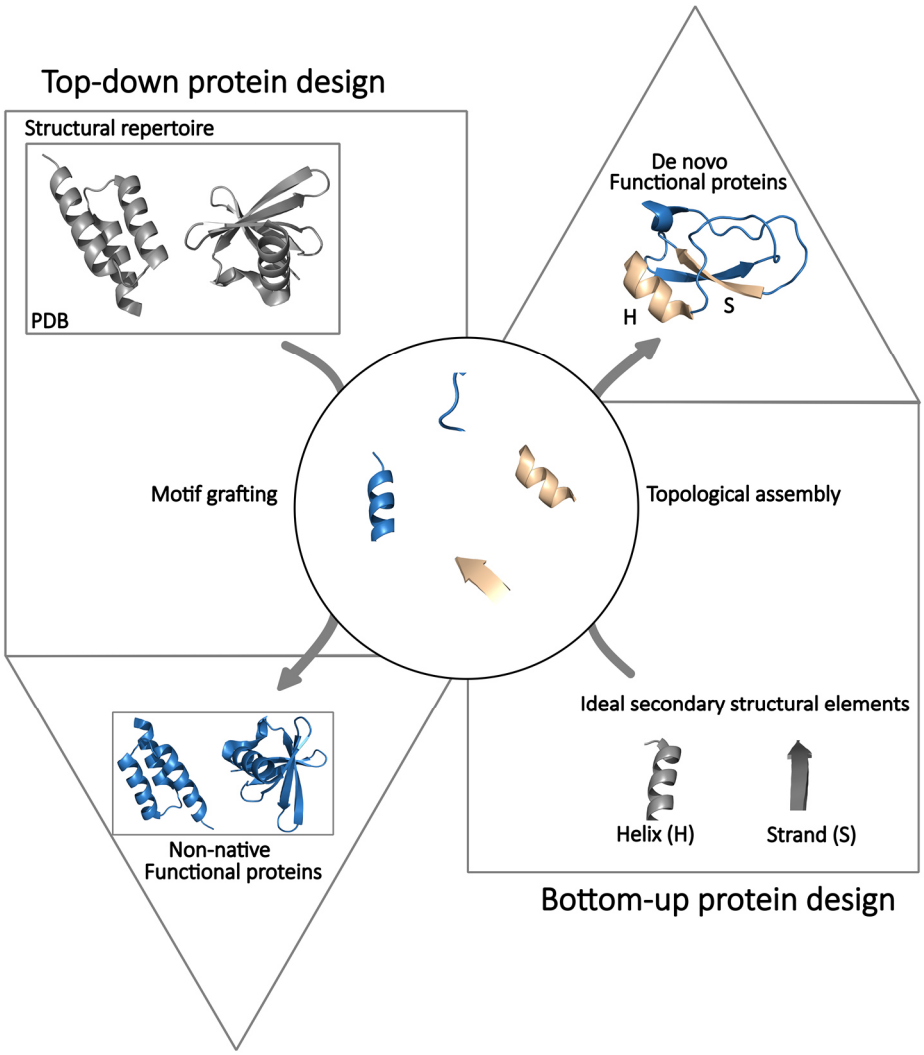


Figure 2: Top-down and bottom-up protein design approaches. Top-down approaches include the grafting of functional motifs onto pre-existing and characterized protein structures, resulting in non-native functional proteins. In bottom-up strategies idealized secondary structural elements (SSEs) are merged for generating *de novo* functional proteins. This scheme is adapted and based on the concept of Yang *et al*.[34]

2.1.2.1 Design of thermostable miniproteins

Thermostability, defined as the ability of proteins to maintain their stability and functionality at temperatures above 50 °C, is a highly favorable property in (mini)proteins.[39,40] This property expands the applicability of proteins beyond their use in basic research in the fields of synthetic biology and chemistry, where they are required for several methods, such as the polymerase chain reaction (PCR) for instance or enzyme catalytic studies. Additionally, thermostable (mini)proteins have also been found to exhibit a longer shelf life in drug therapeutics and chemical industrial applications.[40]

Some thermostable miniproteins with different structural core scaffolds have already been published. One of the smallest representatives of stably folded miniproteins are the Trp-cage peptides with 18 to 20 amino acid residues.[36] The first Trp-cage peptide, TC5b, was designed by Neidigh *et al.* in 2002 and demonstrated a stable and characteristic structure consisting of an N-terminal α-helix, a short 3_{10}-helix, a C-terminal polyproline II region, a hydrophobic core region (Tyr3, Trp6, Gly11, Pro12, Pro18 and Pro19) and a structure-stabilizing salt bridge between Asp9 and Arg16(cf. Figure 3 A).[36,41–43] In CD thermal denaturation studies, TC5b

revealed a melting temperature of 43 °C.[41] A highly thermostable Trp-cage miniprotein was designed by introducing several modifications, such as the addition of Gly residues at both peptide termini, the substitution of Lys8 with an Ala residue and the final peptide cyclization, resulted in the generation of the highly stable cyclo-TC1 miniprotein, with a T_m value of 95 °C (cf. Figure 3 B).[41,44]

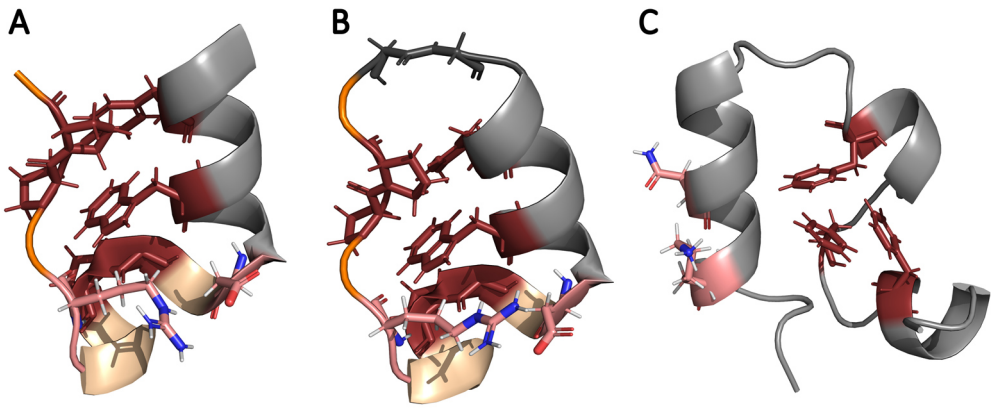


Figure 3: NMR structures of TC5b, the thermostable cyclo-TC1 and HP36. A) NMR structure of TC5b of Neidigh *et al.* (pdb code: 1l2y)[43]. The N-terminal α-helix is depicted in grey, the 3_{10}-helical structure is shown in beige, and the C-terminal poly-proline region is displayed in orange. In both structures the residues of the hydrophobic core are depicted in red and the residues Asp9 and Arg16 that form a structure-stabilizing salt-bridge are represented as pink sticks. **B**) NMR structure of cyclo-TC1 of Scian *et al*. (pdb code: 2ll5)[41]. **C**) NMR structure of HP36. The three Phe residues, which are part of the hydrophobic core region, are displayed as red sticks. The amino acids Asn68 and Lys70 of HP36 (pdb code: 1vii)[45] were mutated to generate the highly thermostable HP36 double mutant, published by Bi *et al.* 2007.[46] PDB files were accessed with RCSB PDB.[33]

Another highly thermostable miniprotein was generated by rational design of the independently and stably folding C-terminal subdomain of the villin headpiece HP36, which is part of an F-actin-binding domain that plays a role in the cytoskeleton.[46,47] The HP36 miniprotein consists of three α-helices and a hydrophobic core (cf. Figure 3 C) that contains three Phe residues among other hydrophobic and structure-stabilizing amino acids.[9,45,48] The introduction of two amino acid mutations, respectively Asn68Ala and Lys70Met, led to the design of the highly thermostable HP36 double mutant, which exhibits a melting temperature of 90 °C in thermal denaturation studies.[46]

2.1.2.2 Functional miniproteins

In addition to structural and thermal stability, several efforts have been made to generate functional miniproteins that either exhibit receptor-like properties, serve as pharmacologically active molecules, or act as synthetic catalysts.[9] Due to the availability of highly diverse scaffolds that are used to generate miniproteins with tailored application, their functions are also diverse. In drug development, miniproteins have been designed to exert various pharmacological effects on target molecules, including agonist or antagonist activities, as well as to function as inhibitors of certain disease-associated PPIs.[8,9]

A miniprotein belonging to the group of cyclotides and used as a common scaffold for functionalization, is the kalata B1 cyclotide.[49] Cyclotides are a subgroup of cystine-dense peptides derived from plants that have a circular backbone and are characterized by their six

conserved Cys residues, which form structure-stabilizing disulfide bridges (cf. Figure 4 A). Due to the highly stable structure of cyclotides, which also confers them high proteolytic resistance, they have been subject of extensive research covering their potential application as insecticidal molecules in agriculture and promising therapeutic agents in drug discovery.[8,9,49] The kalata B1 cyclotide peptide has been identified as a suitable miniprotein scaffold for the development of stable, bioactive peptides, such as the melanocortin receptor 4 (MCR4) agonist kB1 for the treatment of obesity or the myelin oligodendrocyte glycoprotein (MOG)-based peptide MOG3 as a promising agent for the treatment of multiple sclerosis (MS).[50,51]

MCR4 is a G protein-coupled receptor that plays a role in metabolism. Upon its activation by endogenous melanocyte-stimulating hormones (MSH), MCR4 has been shown to induce appetite reduction and an increase in energy expenditure,[52] which makes it a promising target for obesity treatment.[50] The MRC4-stimulating MSH's share the His-Phe-Arg-Trp sequence motif, which was grafted onto the highly structurally stable kalata B1 cyclotide scaffold, resulting in the design of the kB1(GHFRWG;23-28) analog of Eliasen *et al.* in 2012 (cf. Figure 4 B). The MCR4 agonist demonstrated high affinity and selectivity for the receptor, as well as high proteolytic stability.[50]

Another functional miniprotein that was obtained by grafting the epitope of myelin oligodendrocyte glycoprotein (MOG) (residues of the position 35-55) onto the cyclotide kalata B1 scaffold. This was achieved through rational design of the MOG3 peptide by Wang *et al*. MOG3, which contains the grafted sequence pattern RSPFSRV in Loop 5 at position 41 to 47, was found to exhibit a stable structure and to actively prevent disease progression by immune response modulation, thereby reducing myelin and axonal damage in an *in vivo* MS mouse model.[9,51]

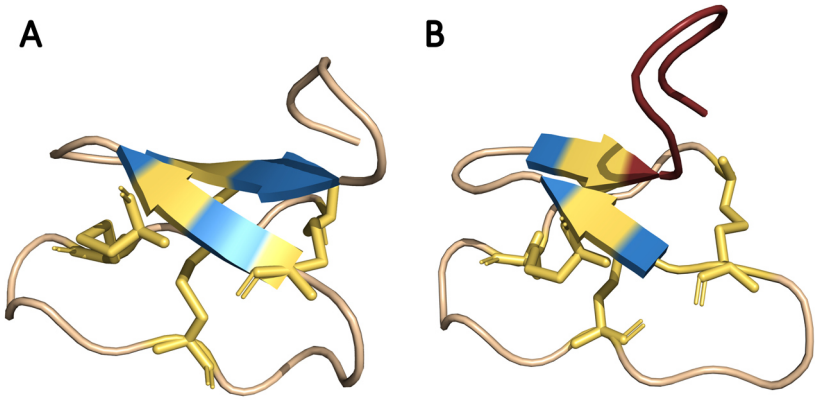


Figure 4: NMR structures of the cyclotides kalata B1 and the kalata B1 (GHFRWG, 23-28) variant. A) kalata B1 scaffold (pdb code: 1nb1)[53] is depicted with Cys residues represented as yellow sticks, beta sheets are shown in blue and loop regions and peptide termini are displayed in beige. **B**) kalata B1 mutant (GHFRWG, 23-28) (pdb code: 2lur)[54] is shown with the grafted sequence pattern displayed in red. PDB files were accessed with RCSB PDB.[33]

A functional miniprotein that belongs to the group of hydrophobic core peptides, is a CD4-binding adnectin, which was identified as a highly potent HIV-1 inhibitor, exhibits a nanomolar binding affinity to human CD4 and thus blocks the fusion of the virion with T-cell membranes. This adnectin-based HIV-1 inhibitor also demonstrated high activity against numerous viral

HIV-1 envelope proteins, indicating a promising broad-spectrum function.[8,55] In addition to the development of adnectins with antiviral properties, several efforts have been made to design an adnectin-based miniprotein with anti-angiogenic effects for cancer therapy. Another adnecting variant, has been identified that inhibits the conversion of low-density lipoprotein (LDL), leading to cholesterol reduction and providing a promising pharmaceutical agent for cardiovascular disease.[56]

The third group of miniproteins, the chemically stabilized stapled miniproteins, have been extensively studied for their application in cancer research. One promising candidate, designated as aStAx-35, is an α-helix-based stapled peptide that was found to inhibit the activation of transciption factor 4 (TCF4) by binding directly to β-catenin, thereby preventing the PPI interaction of TCF4 and β-catenin. The interaction of aStax-35 with β-catenin results in the inhibition of the oncogenic and cancer-progressing Wingless and INT-1 (Wnt) signal pathway, which is deregulated in various types of cancer.[8,57] The aStAx-35 peptide consists of an α-helix based on the axin peptide and an all-hydrocarbon staple that is introduced at Ser471 and Ser475. The structural characterization of the aStAx-35 miniprotein in complex with β-catenin in the crystal structure (cf. Figure 5) revealed that two introduced Trp residues at the positions 468 and 481 were responsible for the nanomolar binding affinity of aSTAx-35 to β-catenin by forming short-range Trp-Trp interactions with β-catenin.[57]

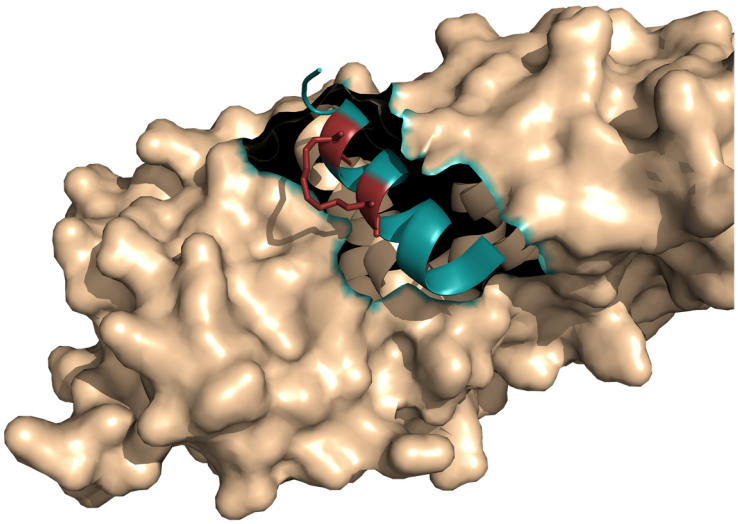


Figure 5: Crystal structure of the stapled miniprotein aStax-35 with β-catenin. The β-catenin protein is displayed in surface mode (beige). The aStax-35 peptide (blue) (pdb code: 4djs)[58] with the introduced staple (red). PDB files were accessed with RCSB PDB.[33]

2.1.2.3 Chances and limitations of miniproteins

To date, the application of miniproteins in medicinal chemistry and drug design, has been extensively studied in recent years, resulting in promising candidates. In general, miniproteins have several advantages over small molecules or antibodies due to their medium size of up to 10 kDa.[3,8,9] As discussed in previous chapters, miniproteins can serve as scaffolds for the introduction of various functional groups and non-canonical amino acids, leading to the design of high-affinity and highly potent therapeutic agents with diverse tailored pharmaceutical effects. Miniproteins can function as receptor agonists or antagonists, as well as steric blockers to prevent disease-associated PPIs.[8] Additionally, they can access biological

compartments that are unavailable for large-sized antibodies. However, while miniproteins offer several advantages, their application is limited by the availability of appropriate scaffolds with sufficient size, shape, and structural and thermal stability. [8,9,36]

The structural and thermal stability of (mini)proteins is one important factor that determines their applicability. Nevertheless, when mutations are introduced to a protein scaffold with the objective of improving or introducing a novel function, this top-down engineering approach frequently results in structural destabilization, which is defined as the stability-function trade-off in proteins.[59] As a consequence of structural instability, proteins often reveal the tendency to aggregate, thereby limiting or even hindering their intended application. In the publication by Teufl *et al.*, it was reported that every folded protein has a threshold robustness. Within this margin, mutations can be introduced without leading to the formation of structurally unstable proteins. Once this threshold is exceeded, the protein is unable to maintain its native structure, resulting in the formation of destabilized and unfolded proteins. For highly thermostable proteins, this threshold robustness was found to be greater than for proteins with low structural and thermal stability.[59] This stability-function trade-off is not exclusive to proteins; it also applies to miniproteins. Moreover, it has been identified in the context of several protein engineering approaches. To overcome this phenomenon, various strategies have been developed, including the use of highly thermostable protein scaffolds to introduce novel functions.[59]

In the last decades of miniprotein design, there was a predominant focus on the extensively studied α-helical coiled coils due to their well-understood sequence-structure relationships, which are essential for bottom-up engineering approaches.[60,61] A variety of α-helical scaffolds were employed in the design of functional miniproteins, including an artificial esterase mini-enzyme based on the α-helical bovine pancreatic polypeptide (bPP) and an α-helical miniprotein that provides cell permeability using the avian pancreatic peptide as the initial design scaffold.[3,62] In addition to α-helical core scaffolds, considerable effort has been made to design functional β-sheet-based miniproteins, which were previously neglected in protein design due to their reported high tendency to aggregate, as well as low solubility, and poor folding properties.[9,61,63] Despite these limitations, β-sheet-based miniproteins have already been successfully designed to function as minireceptors for various target proteins, or served as important motifs to study folding kinetics.[9,61,64]

2.2 *De novo* design of WW domain scaffolds

2.2.1 Classification and abundance of the WW domain in nature

The WW domain belongs to the group of miniproteins, and is considered as the smallest native, independently folding protein module, comprising 30-40 amino acid residues. The WW domain contains three antiparallel β-sheets that are connected by two loops, and exhibits a flexible binding site (cf. Figure 6 A).[10,65–67] The name of the WW domain refers to the two highly conserved Trp residues, which are usually located 20 to 22 amino acids apart within the sequence and have been reported to exhibit structural and functional roles.[68] The structural stabilization of the WW domain is mediated by the highly conserved hydrophobic core residues, which include: Leu2 and Pro3 of the N-terminus; Trp6 of β-Sheet 1 (referred to as

BS1); Phe/Tyr19 of β-Sheet 2 (BS2); and Pro32 of the C-terminal region (cf. Figure 6 B).[6,65,66,68–70]

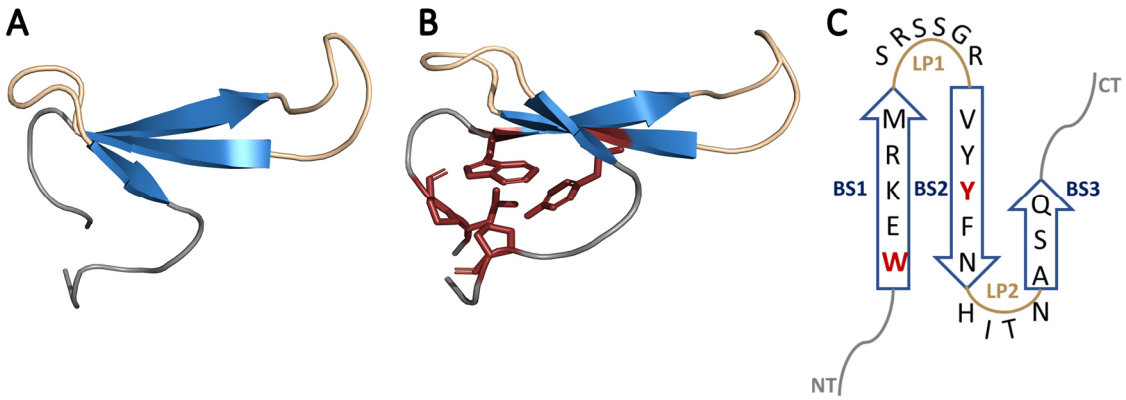


Figure 6: Characteristic structure of a WW domain. A) NMR structure of *h*Pin1$_{WW}$ (pdb code: 1i8g)[71] with beta sheet regions in blue, loops in beige and the peptide termini in grey. **B**) NMR structure of *h*Pin1$_{WW}$ (pdb code: 1i8g)[71] conserved residues of the hydrophobic core of *h*Pin1$_{WW}$ are presented as red sticks. **C**) Two-dimensional schematic representation of *h*Pin1$_{WW}$, with beta-sheets indicated as BS, loop regions as LP, and the peptide termini shown as NT and CT. PDB files have been accessed with RCSB PDB.[33]

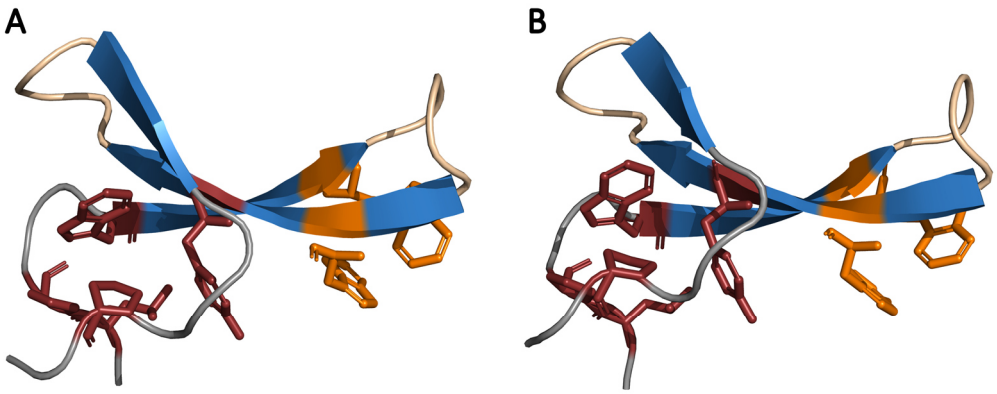
In nature, WW domains are highly abundant, predominantly functioning as protein-protein interaction modules that bind to proline-rich sequence motifs.[69,72] Ligand recognition was found to be mostly mediated by the two loop regions.[64] Native proteins that contain WW domains have versatile functions, as they play a role in signal transduction, transcription, the regulation of the cell cycle, and in disease-associated PPIs, such as in cancer pathogenesis and neuronal disorders.[65–67,73] A well-studied native WW domain in literature belongs to the human peptidylprolyl cis-trans isomerase "protein interacting with never in mitosis A 1" (*h*Pin1).[74] In general, WW domains are classified into four groups, each of which binds to a different ligand sequence motif. WW domains of Group I a/b bind to proline-rich motifs of the type `PPxY/LPxP` (with *x* is a random amino acid), Group II recognizes ligands with a `PPLP` sequence pattern, Group III binds to `PPRP` motifs, and Group IV WW domains bind to phosphorylated ligands of the type `pS/pTP`.[65–67,75] Although WW domains were traditionally divided into four groups, Kato *et al.* proposed to merge WW domain Group II and III, as they share similar binding motifs.[72]

2.2.2 Early studies on WW domains

The WW domain has been the subject of extensive research over the past two decades, with the aim of elucidating the folding thermodynamics and kinetics of β-sheet proteins.[64,76–78] In 1994, Bork *et al.* first identified the WW domain in the Yes-kinase-associated protein (YAP) in several species, including human, mouse and chicken.[79] In 1996, Macias *et al.* reported the typical β-sheet structure of the WW domain *h*YAP65$_{WW}$ upon binding to the proline-rich Group I-specific ligand `GTPPPPYTVG` by nuclear magnetic resonance (NMR) spectroscopy.[76] Subsequent studies revealed the existence of highly conserved residues that contribute to the structural stability of WW domains, leading to the classification of WW domains into different groups.[77] In the early 2000s, thermodynamic folding studies were conducted on *h*YAP$_{WW}$ and *h*Pin1$_{WW}$ by Jäger *et al*, which demonstrated the importance of LP1 for the folding of a WW domain and its two-state folding behavior in thermal denaturation experiments.[64,80]

2.2.3 Design of thermostable WW domains

The most thermostable WW domain-based peptide to date is WWst34, with a melting temperature of 94 °C.[81] The rational design approach employed for the highly thermostable *h*Pin1$_{WW}$ variant involved the introduction of a Trp-Trp cross-strand pair at a non-hydrogen-bonded (NHB) site, as previously demonstrated in a study by Jäger *et al*. Additionally, further redesign of the loop regions LP1 and LP2 was conducted (cf. Figure 7 B).[81,82] The peptide variant 2, originally designed by Jäger *et al*., exhibits two tryptophan residues that were introduced at sequence position 10 of BS1 and position 17 of BS2 (cf. Figure 7 A). The indole groups of the side chains of the two Trp residues form a cross β-strand structure-stabilizing interaction, which results in a melting temperature of 75 °C.[82] In light of the findings of the highly structure-stabilizing Trp-Trp cross-strand interaction, Kier *et al.* proceeded to mutate the amino acids, primarily in the LP1 region. The Ser11 was substituted with a charged aspartate residue, in order to enhance solubility. Furthermore, a flexible Gly residue was incorporated at sequence position 13 to facilitate further stabilization of the folding. Further redesign of structural segments, including BS2, LP2, BS3, and the peptide's C-terminus, was performed, resulting in the final amino acid sequence of WWst34, which is presented in Figure 8 C for comparison with the peptide variant 2 of Jäger *et al*.[81,82]



C

WW domain	Sequence ----NT----BS1----LP1----BS2--LP2--BS3----CT------- 1 - 10 - 20 - - 30	T_m / °C
Peptide variant 2[a]	H-KLPPG-**WEKRW**-SRSSGR-**WYYF**-NHIT-**NAS**-QWERPSG-OH	75.0
WWst34[b]	H-KLPPG-**WEKRW**-DRGSGR-**WFYF**-NRIT-**GKR**-QFERPSD-OH	94.0

Figure 7: Predicted structures and sequences of the published highly thermostable peptides variant 2 and WWst34. A) *AlphaFold2* prediced structure of peptide variant 2 by Jäger *et al*. **B**) *AlphaFold2* predicted structure of WWst34 by Kier *et al*. Amino acid residues depicted in blue indicate beta sheet regions, loop regions are shown in beige and the peptide termini are displayed in grey. Residues that are part of the hydrophobic core are shown in red, residues in orange display the introduced Trp residues. **C**) Sequences that are presented were extracted from the publications of Jäger *et al.* (a)[82] and Kier et al. (b)[81].

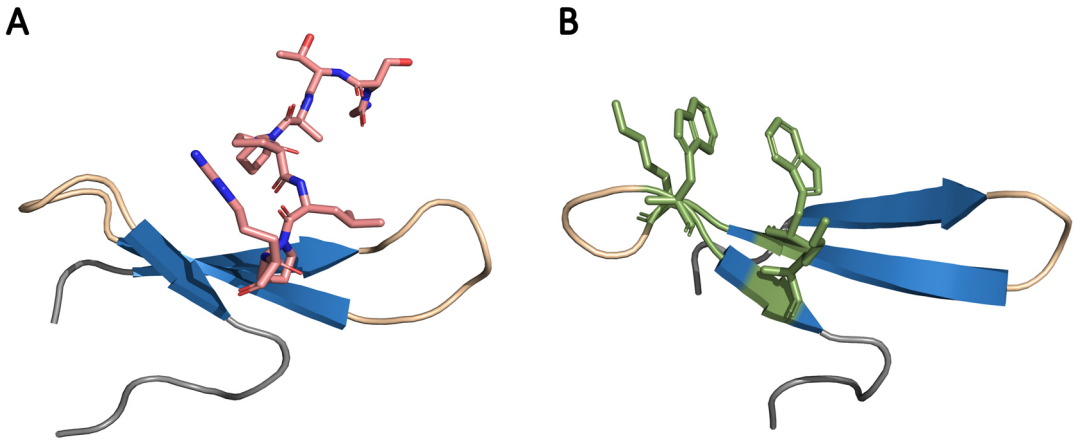
2.2.4 Design of functional WW domains

In addition to studies on WW domain structure and stability, several advances have been made in the design of functional WW domains over the past few years. These recent developments have led to the design of minireceptors that are capable of binding to single-

stranded (*ss*) deoxyribonucleic acid (DNA), as well as metals such as Zn(II), and a variety of organophosphates, including adenosine triphosphate (ATP) and phosphorylcholine (PC).[83–85]

2.2.4.1 Single-strand DNA binding WW domain

The interactions between proteins and *ss*DNA play a crucial role in numerous biological pathways and processes, including DNA recombination, replication, and DNA repair mechanisms.[86] The first *ss*DNA-binding WW domain, which is based on the native formin binding protein 11 (FBP11) WW1 domain sequence (cf. Figure 8 A), was introduced by Stewart *et al.* in 2011 and is referred to as Mut1. The design of a *ss*DNA-binding WW domain minireceptor was conducted by grafting an oligonucleotide/oligosaccharide binding motif (OB-fold), which is typically located on *ss*DNA and recognized by proteins, onto the FBP11 native sequence with the sequence pattern WKWK (cf. Figure 8 B).[83] In studies previously conducted on nucleotide binding with β-hairpin peptides by Butterfield *et al.*, the aforementioned sequence motif was identified to be essential for the peptide-DNA interaction.[87] Moreover, the Mut1 peptide exhibited a mutated Asn-Gly pattern within the turn sequence region, thereby enabling the formation of a stably folded turn segment.[83]



C

WW domain	Sequence
	---NT---BS1---LP1---BS2---LP2--BS3---CT----
	1 - 10 - 20 - - 30
FBP11 WW1[a]	Ac-SE-**WTEHK**-SPDGR-**TYYY**-NKVN-**GOW**-QSTWEKPG-NH_2
Mut1[b]	Ac-SR-**WTEHK**-SNGR--**TYYW**-NKVN-**GOW**-QKTWEKPG-NH_2

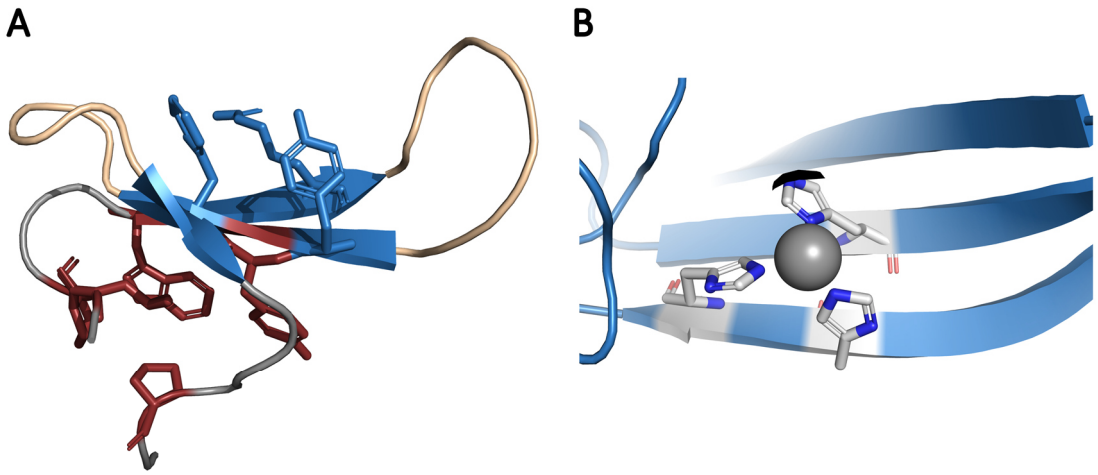
Figure 8: Structures and sequences of the native FBP11 WW1 domain and the *ss*DNA-binding Mut1 variant. A) NMR structure of the native FBP11 WW1 domain (pdb code: 1zr7)[7] and its native poly-proline ligand (rose sticks) by Kato *et al*. PDB file was accessed with RCSB PDB.[33] **B**) Predicted *AlphaFold2* structure of Mut1 by Stewart *et al*.[83] The peptide termini are depicted in grey, beta Sheet regions in blue, loop/turn segments in beige, and the WKWK sequence pattern residues in green. **C**) Sequences extracted from publications of FBP11 WW1 (a)[7] and Mut1 (b)[83] are illustrated. Beta sheet regions are displayed in blue, while the grafted WKWK sequence pattern in Mut1 is shown in green.

The binding studies of Mut1 revealed a highly selective binding to *ss*DNA in the micromolar range, which was also two-fold stronger than that observed for previously designed β-hairpin monomers. This suggests that the β-Sheet 1 plays an important role in the development of a

high binding affinity to *ss*DNA. Additionally, the Mut1 peptide, which contains the introduced OB-fold motif, has been demonstrated to lack the capacity to bind to native FBP11 proline-rich ligand sequences. Instead, it exhibits a selective binding affinity for *ss*DNA. In contrast the FBP11 wild-type WW domain peptide has been shown to bind exclusively to its native proline-rich ligand.[83] The Mut1 peptide design represents one of the first functional miniproteins based on the triple β-sheet structure of a WW domain.

2.2.4.2 Metal-binding WW domain minireceptors

In a recent study, another functional WW domain peptide that is capable of reversibly switching its conformation upon Zn(II) binding or an induced change in pH, was designed and reported by Pham *et al*.[84] The rational design approach was based on the WW domain of *h*Pin1, a well-studied and independently folding natural WW domain. *h*Pin1$_{WW}$ comprises three anti-parallel β-sheets and has been reported to exhibit an inner and outer hydrophobic core (cf. Figure 9 A), which contribute to the peptide's structural stability. The highly conserved residues Leu2, Pro3, Trp6, Tyr/Phe19 and Pro32 form the inner hydrophobic core, which was maintained. The outer hydrophobic core, located on the surface of the three β-sheets, is comprised of the three residues Arg9, Tyr18, and Phe20.[64,80,88,89,90] These residues were mutated to a His_3 site that was identified in the active site of the carbonic anhydrase (CA) enzyme II (cf. Figure 9 B) and was postulated to provide Zn(II)-binding properties, as well as pH responsiveness.[84,91] Three additional modifications were introduced to obtain the WW-CA peptide. These included the substitution of Arg16 with Gln and Asn25 with Gln, in order to enlarge the hydrogen bonding pattern, and the mutation of His21 to Gln was introduced to prevent unspecific metal binding (cf. Figure 9 C).[84]



C

WW domain	Sequence
	----NT----BS1----LP1----BS2--LP2--BS3----CT-------
	1 - 10 - 20 - - 30
*h*Pin1$_{WW}$[a]	H-K**LP**PG-**WEKRM**-SRSSGR-**VYYF**-NHIT-**NAS**-QWER**P**SG-OH
WW-CA[b]	H-K**LP**PG-**WEKHM**-SRSSGQ-**VHYH**-NSIT-**QAS**-QWER**P**SG-OH

Figure 9: Structures and sequences of *h*Pin1$_{WW}$ and WW-CA. **A**) NMR structure of *h*Pin1$_{WW}$ by Wintjens *et al.* (pdb code: 1i6c)[92] is shown with beta sheets depicted in blue, loop regions displayed in beige and the peptide termini illustrated in grey. Residues presented as red sticks are part of the inner hydrophobic core, and residues

shown as blue sticks belong to the outer hydrophobic core. **B**) Active center of carbonic anhydrase II (pdb code: 1ca2)[93] is illustrated with the His_3 site depicted as sticks and Zn(II) presented in grey. **C**) Sequences of *h*Pin1$_{WW}$ by Wintjens *et al.* (a)[92] and WW-CA by Pham *et al.* (b)[84] are presented with the residues of the inner hydrophobic core displayed in bold red, beta-Sheet regions are shown in dark blue, and the grafted His_3 site depicted in light blue. PDB files were accessed with RCSB PDB.[33] This figure is based on Figure 1 in the publication of Pham *et al*.[84]

Structural characterization by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy demonstrated changes in the conformation of WW-CA following the binding of Zn(II) from a weakly folded apo state to a stably folded holo state. In the presence of Zn(II), WW-CA displayed a cooperative folding-to unfolding behavior in CD thermal denaturation studies with a melting temperature of 34.3 °C. Furthermore, conformational changes of the WW-CA-Zn(II) complex were found to be inducible by changes in the pH, leading to the conclusion that WW-CA is a Zn(II)- and pH-responsive switchable peptide. Binding studies of WW-CA with Zn(II) revealed a high binding affinity of the designed peptide to Zn(II) with a K_d value of 1.2 µM. Moreover, the binding of WW-CA to Zn(II) was shown to be highly specific, as several other tested divalent cations, such as Ca(II), Mg(II) and Fe(II), did not bind to the His_3 site of WW-CA. In conclusion, the rationally *de novo* designed WW-CA represents a functional and switchable β-sheet peptide.[84]

2.2.4.3 Organophosphate WW domain minireceptors

In addition to a metal-binding *h*Pin1$_{WW}$-derived peptide, organophosphate-binding minireceptors have been designed based on a combinatorial screening approach of split-WW domains in our laboratory, as detailed in the publication by Neitz *et al*.[85] The model peptide *h*Pin1$_{WW}$ was initially used to generate fragment peptides that were linked to a coiled coil (CC) sequence at their N-terminus, resulting in the B_{CC} peptide series. Alternatively, they were linked to a CC sequence at their C-terminus, leading to the generation of A_{CC} peptides (cf. Figure 10 A). Upon the association of the antiparallel heterodimeric CC segments ($A_{CC}B_{CC}$), one of the fragmented peptides, designated as B_{CC}-*h*Pin1$_{WW}$-L1, was observed to effectively reconstitute with the corresponding *h*Pin1$_{WW}$-L1-A_{CC} peptide fragment, exhibiting the characteristic WW domain structure. The functionality in terms of the binding affinity of the reconstituted CC-*h*Pin1$_{WW}$-L1 variant to the *h*Pin1$_{WW}$ native CTD ligand was examined by fluorescence titration spectroscopy. The binding studies revealed a comparable binding affinity of the reconstituted CC-*h*Pin1$_{WW}$-L1 peptide (K_d = 2.0 µM) to that of the native *h*Pin1$_{WW}$ (K_d = 2.2 µM). Based on the observation of successful reconstitution of *h*Pin1$_{WW}$-derived split peptides, a library of split-WW domain variants was generated, comprising four different N-terminal fragments and six different C-terminal fragments with linked CC segments and resulting in a total of 24 peptide combinations (cf. Figure 10 B).[85]

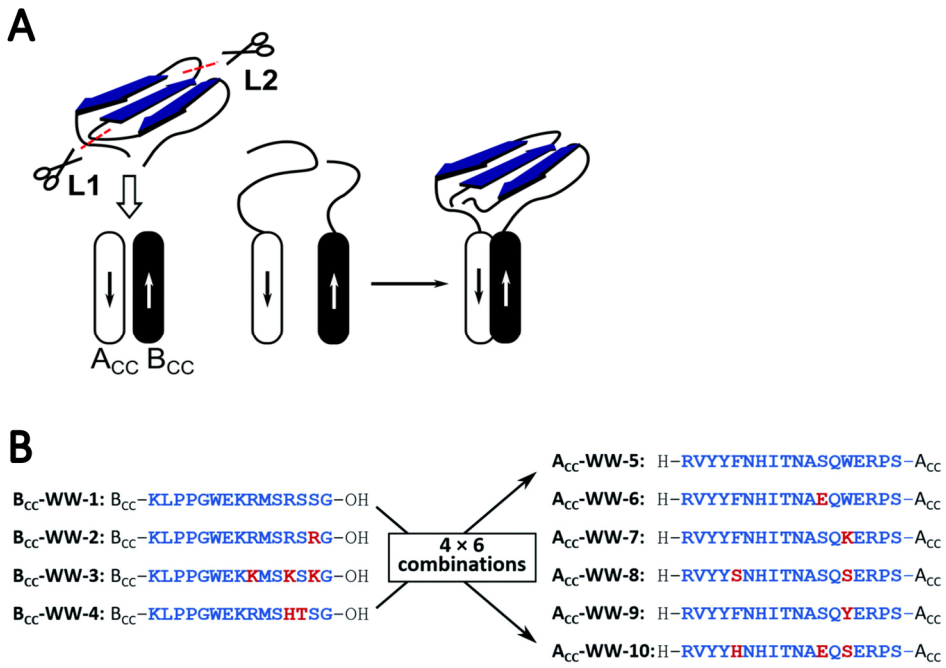


Figure 10: Combinatorial screening of a split-WW domain peptide library for organophosphate binding. A) Illustration of the reconstitution concept of fragmented *h*Pin1$_{WW}$ split sequences upon coil coil (CC) association (A_{CC}-acidic coil strand; B_{CC}-basic coil strand; L1 - Loop1; L2 – Loop2). **B**) Sequences of the B_{CC}-*h*Pin1$_{WW}$-L1 variants (B_{CC}-WW-1 – B_{CC}-WW-4) and *h*Pin1$_{WW}$-L1-A_{CC} variants (A_{CC}-WW-5 – A_{CC}-WW-10) utilized for the split-WW domain library. In total 24 split-WW domain peptides were generated when combining the N -and C-terminal fragments. This figure was adapted and reused from Neitz *et al.*[85], with the permission from © 2022 the Royal Society of Chemistry.

The peptide library was screened for organophosphate binding, with a particular focus on binding to adenosine triphosphate (ATP), cyclic adenosine monophosphate (cAMP), phosphorylcholine (PC), or inositol hexaphosphate (IP6). The results of the binding studies demonstrated that the peptide fragment combination CC-WW-2-10 exhibited a specific binding affinity for ATP, whereas the peptide variant CC-WW-1-8 was identified as a selective PC binder. Following the identification of selective organophosphate-binding CC-fusion peptide variants, WW-2-10 and WW-1-8 were synthesized without CC sequence fragments, resulting in a complete WW domain peptide sequence. The structural analysis of both peptides revealed characteristic spectra of a WW domain, exhibiting the signal of the exciton coupling at 227 nm. Thermal denaturation studies showed that WW-2-10 and WW-1-8 displayed low thermostabilities in comparison to the native *h*Pin1$_{WW}$ peptide. Nevertheless, the binding studies demonstrated that both peptides possess micromolar binding affinities to their respective organophosphate ligands. The binding experiments of WW-2-10 with ATP revealed a K_d value of 14 µM, while WW-1-8 demonstrated binding to PC in the low micromolar range, with a K_d of 1.3 µM. The identified highly selective binding of WW-2-10 to ATP and WW-1-8 to PC represent promising functional properties of *h*Pin1$_{WW}$-derived minireceptors.[85]

3 Aim of Research

In protein or peptide engineering, the introduction of mutations in native amino acid sequences to gain additional or improved function often leads to an overall destabilization of the structure, referring to the stability-function trade-off previously described by Teufl *et al*. As a result, the ultimate application of engineered miniproteins in research fields, where temperature stability is required, is often limited or precluded.[59]

In order to establish the β-sheet-based WW domain scaffold in miniprotein design, general design rules need to be implemented as guidelines. One major goal of this work is to identify sequence-structure relationships in WW domains. To achieve this aim, it is planned to perform sequence alignments of naturally occurring WW domains to identify the most prevalent amino acids at specific sequence positions, which are considered for the design of a WW-domain based consensus sequence (WW-CS) (cf. Figure 11, step1). The WW-CS will serve as a starting point for the segment-wise sequence iteration, which will be initiated in the structural segment of β-Sheet 1 (BS1) within WW-CS. In total six rounds of sequence iterations will be conducted, and the results will be used to formulate WW domain-specific sequence-structure-stability relationships.

Based on the sequence-structure-stability relationships, a thermostable WW domain scaffold will be designed to serve as a core scaffold, on which functions can be grafted (cf. Figure 11, step 2 and 3). The underlying rationale of this approach is that once the scaffold is thermostable, amino acid mutations can be introduced to install a specific function, such as binding to a well-defined class of molecules (receptor function), thereby enabling the design of miniproteins with improved thermostability compared to top-down designs derived from evolutionary optimized natural WW domains. To prove this hypothesis, it is planned to design WW domains with different receptor properties, such as peptide, metal and small molecule recognition. As a proof of concept for the rational design, the four different WW domain group-specific binding motifs will be grafted onto the hyperstable WW domain core scaffold. To increase the complexity, WW domains with metal- and organophosphate-binding properties will be designed. The final objective is to use the thermostable WW domain core scaffold to design a miniprotein with phosphatase activity. The hydrolytic activity of the designed phosphate-binding peptide variants towards the phosphoester bond of two different molecules, respectively *para*-nitrophenol phosphate (*p*-NPP) and ATP or its di- or monophosphate derivatives, will be tested to identify a potential enzymatic function (cf. Figure 11, step 4).

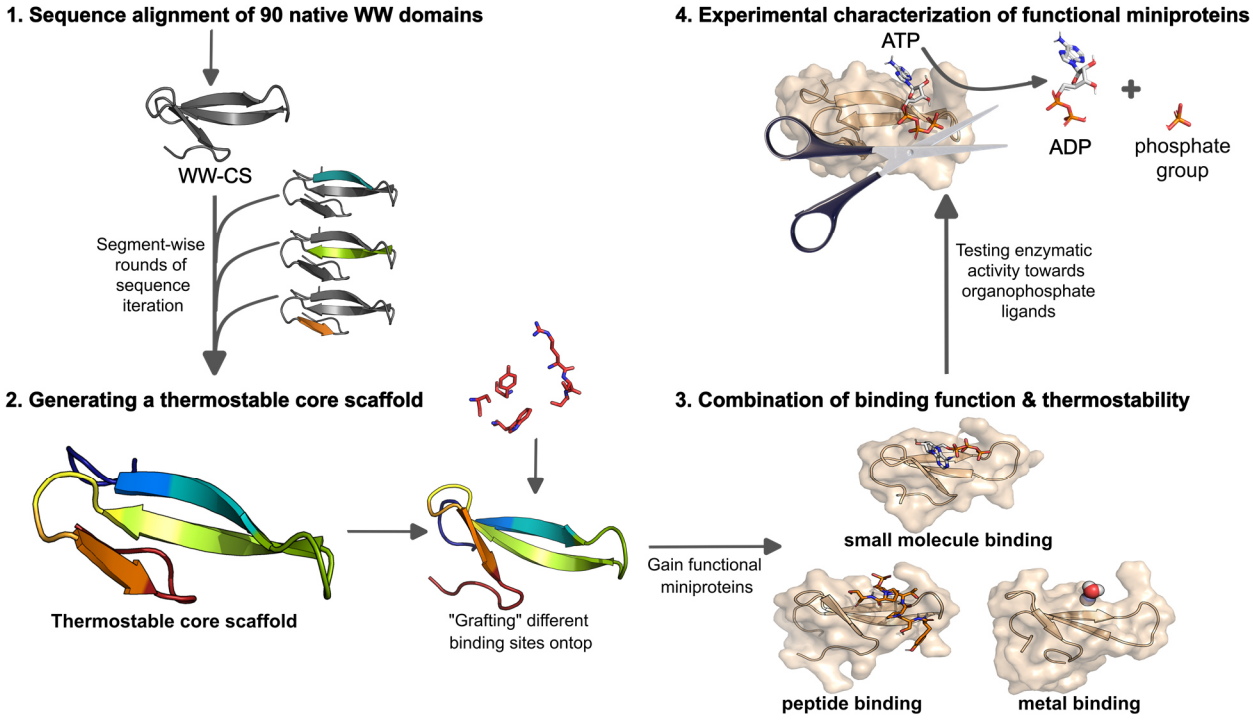


Figure 11: Overview of the four central research goals that are addressed in this work. Parts of the figure were adapted with permission from Lindner *et al*.[94], ©2024 American Chemical Society.

4 Results and Discussion

4.1 Design of a thermostable basic WW domain scaffold

Most of the data that are presented in this chapter were previously published in the Journal of the American Chemical Society (JACS) under the title "Thermostable WW domain scaffold to design functional β-sheet miniproteins".[94] Permission to reuse or modify the tables and figures within this dissertation was granted by the American Chemical Society.

4.1.1 Design concept and generation of a WW domain consensus sequence (WW-CS)

All previous design approaches for chemically synthesized WW domains have in common that they do not focus exclusively on the WW domain structure, but rather on the ligand-protein-domain complex. In this context, conserved group-specific amino acid residues of WW domains that are relevant for ligand binding have been identified by statistical coupling analysis (SCA) within the method of multiple sequence alignments performed by several research groups.[77,95,96] In 2000, Macias *et al*. presented a WW domain prototype with moderate thermostability of 44.2 °C, designed after evolutionary structure analysis of selected 22 different WW domains to identify structurally important amino acids at specific sequence positions.[77] Based on the fact that protein as well as miniprotein design is mostly relying on multiple sequence alignments (MSAs) of existing protein structures[36,97], available in the Protein Data Bank (PDB), a similar design approach was utilized in this research, combined with an iterative and experimental assessment of the structure and thermostability of the generated variants.

Sequence alignment of 90 WW domains that are found in the Protein Data Bank[33] or have been published in previous studies[77,96], was applied to identify the degree of conservation of amino acid residues at specific sequence positions (cf. Appendix Table S1).[94] The results from Macias *et al.* and Socolich *et al.* showed the presence of the three universal β-sheets at the sequence positions 6-10 (β_1), 17-20 (β_2) and 25-27 (β_3), which are stabilized by inter- and intra-strand hydrogen bonds. In addition, the hydrophobic core formed by the highly conserved residues Leu2, Pro3, Trp6, Tyr/Phe19 and Pro32 contributes to the overall structure and stability.[77,96] Additional sequence alignments conducted in this work (cf. Appendix Table S1) further indicate a high degree of conservation of Gly15 and Asn21. To visualize the statistical abundance of amino acids at certain sequence positions, a WebLogo[98,99] was created (cf. Figure 12 A).

The statistically most abundant residues were chosen for WW-CS for most of the sequence positions. Nevertheless, to prevent aspartimide formation in solid-phase peptide synthesis (SPPS), Asp was not selected at positions, where possible. However, the Asp residue at sequence position 11 was maintained, because it was hypothesized to be involved in structure-stabilizing electrostatic interactions with charged residues. All of these considerations resulted in the amino acid sequence for WW-CS presented in Figure 12 B.

In silico structure prediction of WW-CS with *AlphaFold2*[100] revealed a typical WW domain structure of three β-sheets (BS, cf. Figure 12 C and D, blue) that are connected by two loops

(LP, beige). The peptide termini are depicted in grey. The general WW domain-specific conserved amino acid residues that form the structure-stabilizing hydrophobic core are shown in red (cf. Figure 12 D).

WW-CS was synthesized by automated microwave-assisted SPPS and the peptide's secondary structure and thermal stability were characterized by circular dichroism (CD) spectroscopy. As displayed in Figure 12 E, WW-CS exhibits a CD spectrum characteristic of a WW domain with a maximum at 228 nm, a minimum at 205 nm, and a shoulder in the short wavelength region around 218 nm.[101] The characteristic maximum signal at 228 nm and the minimum at 205 nm originate from the exciton coupling between the two aromatic amino acids Trp6 and Tyr19.[82] The thermal denaturation profile of WW-CS, revealed a melting temperature of 53.5 °C (cf. Figure 12 F), indicating good thermal stability.

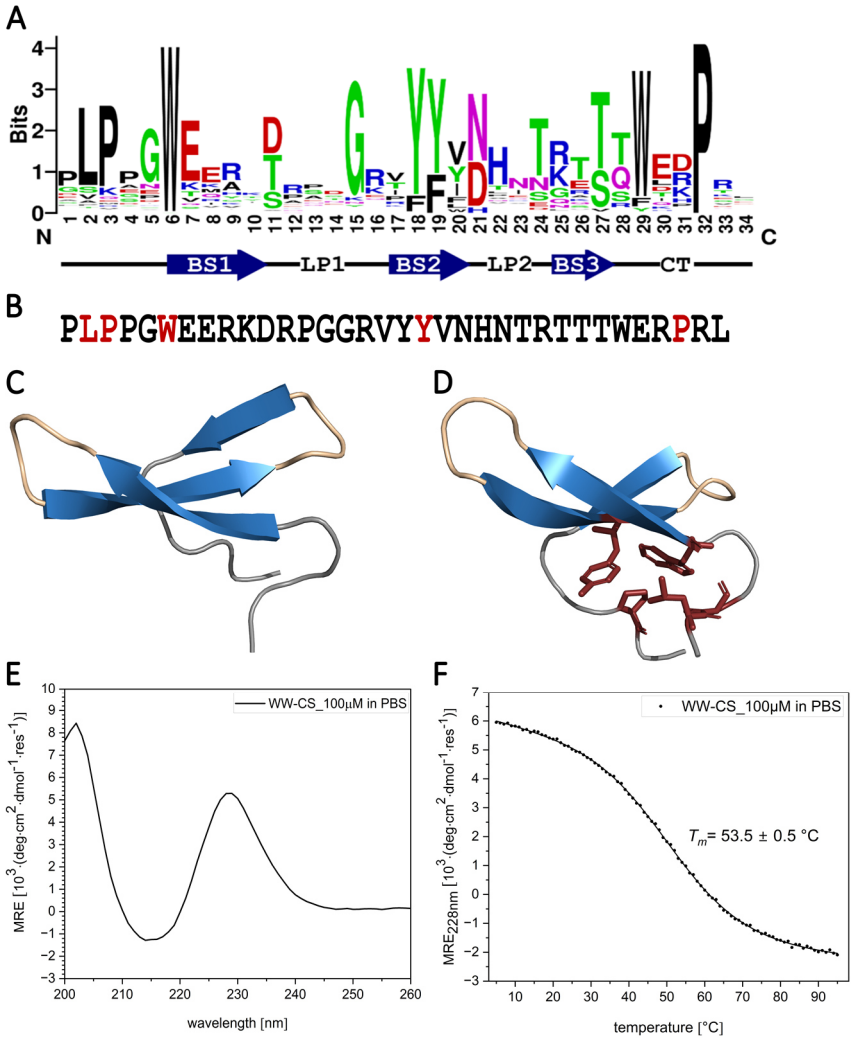


Figure 12: The design of WW-CS. A) WebLogo[98,99] of 90 aligned native WW domains, with β-sheets (blue arrows), loop regions and the peptide termini. **B**) Amino acid sequence of WW-CS. Highly conserved residues are displayed in red. **C**), **D**) Predicted secondary structure of WW-CS, with β-sheet regions displayed in blue, loop segments in wheat and the termini in grey. Structure predictions were performed with *AlphaFold2*[100] and visualized with *PyMol*. **E**), **F**) CD spectrum and thermal denaturation profile of WW-CS. CD data represent mean values of triplicate measurements conducted from 190 – 260 °nm at 20 °C. Thermal denaturation curves represent mean values of triplicate measurements. Signals for denaturation curves were determined at a fixed wavelength of 228 nm and variable temperatures, ranging from 5-95 °C. The melting temperature T_m of 53.5 °C was obtained by applying the two-state transition model of a monomer from folded to unfolded state.[101] Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

4.1.2 Analysis of structure-stabilizing sequence patterns

Based on WW-CS, iterative sequence modifications, were performed to identify structure-stabilizing residues. Due to the high degree of conservation, the N-terminus (NT) was not considered for further sequential modification. The direction for iterative sequence modification was from the N-to the C-terminus, since Loop 1 (LP1) has been identified as the initial nucleation site for folding.[80,102] In addition, the results of an alanine and glycine scanning of *h*Pin1$_{WW}$ conducted by Jäger *et al*. in 2009[89] were taken into account for the iterative screening studies. The iterative approach was implemented segment-wise, i.e. starting from the first structural segment, β-Sheet 1 (BS1), followed by LP1, BS2, LP2, BS3 and the C-terminus. The most structure-stabilizing amino acid pattern of one segement was transferred to the subsequent round of iteration. Consequently, six rounds of iteration were performed with 38 different peptide variants in total, all derived from WW-CS.

β-Sheet 1 (BS1) series

The BS1 (6-10) was the initial segment investigated for structure-stabilizing amino acid residues. In addition to the highly conserved Trp6, BS1 is reported to exhibit an accumulation of charged residues. Despite this observation, a sequence pattern of alternating polar and hydrophobic residues was chosen, which is often found in β-sheet structures.[103–105] The introduction of hydrophobic residues in BS1 was performed to investigate whether an enlargement of the hydrophobic core leads to a greater thermostability of the respective peptide variant. Moreover, polar, charged and uncharged amino acid residues were introduced, in order to identify the influence of charges in BS1 on the overall peptide stability. Thus, acidic glutamate residues, as well as basic lysine or uncharged glutamine residues were introduced in the polar positions. Following the described design rules, seven different peptide variants were synthesized and characterized in terms of structure and thermal stability (cf. Table 1).

Table 1: **Overview of the sequence variations of the BS1 series.** The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Introduced sequence modifications are displayed in blue. The experimentally determined melting temperatures are listed as T_m. The sequence pattern highlighted in grey was taken to the next round of iteration. This table has already been published in minorly modified form in Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Iteration	WW domain	Sequence ----NT-------BS1----LP1-----BS2---LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
	WW-CS	Ac-PLPPG-WEERK-DRPGGR-VYYV-NHNT-RTT-TWERPRL-NH_2	53.5 ± 0.5
1	**WW-BS1-1**	Ac-PLPPG-**WEIRI**-DRPGGR-IYFY-NHNT-KTT-TWERPRL-NH_2	69.5 ± 0.5
	WW-BS1-2	Ac-PLPPG-**WEIRL**-DRPGGR-IYFY-NHNT-KTT-TWERPRL-NH_2	66.0 ± 0.5
	WW-BS1-3	Ac-PLPPG-**WEIRT**-DRPGGR-IYFY-NHNT-KTT-TWERPRL-NH_2	65.5 ± 0.5
	WW-BS1-4	Ac-PLPPG-**WEIKI**-DRPGGR-IYFY-NHNT-KTT-TWERPRL-NH_2	65.5 ± 0.5
	WW-BS1-5	Ac-PLPPG-**WEIEI**-DRPGGR-IYFY-NHNT-KTT-TWERPRL-NH_2	24.5 ± 1.0
	WW-BS1-6	Ac-PLPPG-**WKIKI**-DRPGGR-IYFY-NHNT-KTT-TWERPRL-NH_2	49.5 ± 0.5
	WW-BS1-7	Ac-PLPPG-**WQIQI**-DRPGGR-IYFY-NHNT-KTT-TWERPRL-NH_2	51.0 ± 0.5

In general, all seven BS1 variants, showed CD spectra characteristic of a WW domain with a signal maximum at around 228 nm (cf. Figure 13 A). Hence, it can be concluded that all seven peptides are properly folded at 20 °C. When the BS1 variants are compared in detail, the highest signal intensities at 228 nm are found for WW-BS1-1 and WW-BS1-3, which differ only in the amino acid at hydrophobic sequence position 10. WW-BS1-1 was designed with an Ile and WW-BS1-3 with a Thr, both classified as β-branched residues with a high beta propensity.[103,105] The lowest exciton signal at 228 nm is detected for WW-BS1-5, where two Glu residues are introduced in BS1 at polar sequence positions 7 and 9, indicating structural destabilization by charge repulsion.

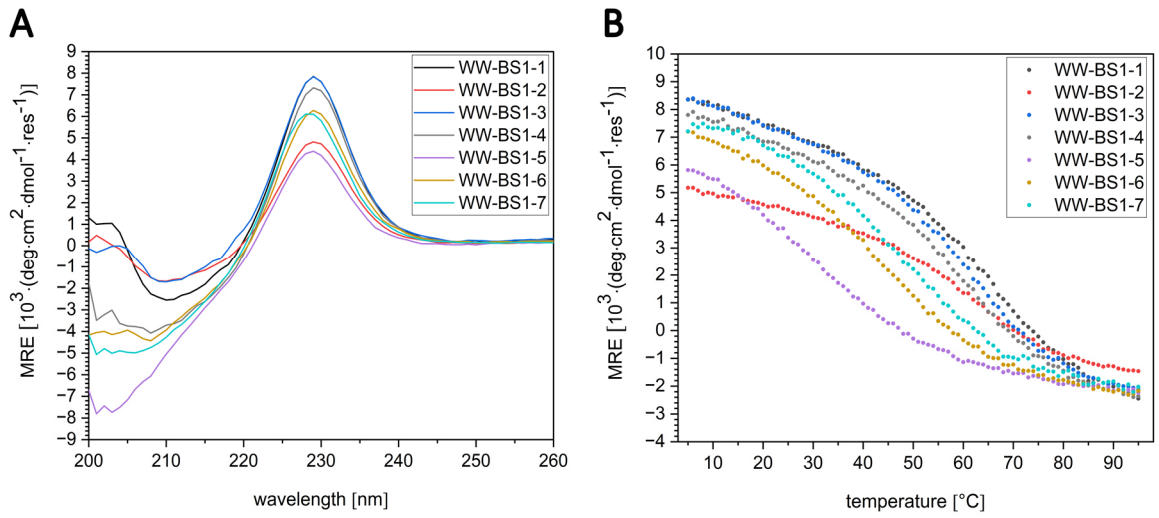


Figure 13: CD spectra and thermal denaturation profiles of the BS1 series. A) CD spectra were recorded in duplicates at 20 °C. **B**) Thermal denaturation curves determined at a fixed wavelength of 228 nm from 5-95 °C. For thermal denaturation curves the mean out of duplicates is shown. Conditions: 100 µM peptide, PBS buffer 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

CD thermal denaturation experiments were performed to investigate thermal stabilities of the BS1 variants. All peptide variants reveal sigmoidal denaturation curves, indicating a cooperative folding-to-unfolding behavior (cf. Figure 13 B).[101] Consistent with the results from CD spectra measurements, showing one of the highest exciton coupling signals at 228 nm, WW-BS1-1 with the `WEIRI` sequence pattern showed the highest thermostability with a melting temperature of 69.5 ± 0.5°C (n = 2), which supports the hypothesis that an extension of the hydrophobic core by introducing an hydrophobic residue into BS1, leads to thermostable WW domain peptide variants.

The hydrophobic amino acid at position 10 seems to have a decisive influence on the thermostability of the peptide, since it is the only amino acid that differs between the peptide variants WW-BS1-1, WW-BS1-2 and WW-BS1-3, which all have high but varying melting temperatures (T_m: WW-BS1-1 = 69.5 °C; T_m: WW-BS1-2 = 66.0 °C; T_m: WW-BS1-3 = 65.5 °C). Thus, it seems that the presence of Ile in the BS1 segment has a greater structure-stabilizing effect on thermostability than Thr. Nevertheless, there is no significant difference between the melting temperatures of WW-BS1-2 and WW-BS1-3, although they show different thermal denaturation profiles. This indicates that Thr and Leu do not exert different effects on structural stability.

A good thermostability, albeit lower if compared to WW-BS1-1, with a melting temperature of 65.5 °C can also be detected for WW-BS1-4 with WEIKI as sequence pattern. The positively charged lysine at sequence position 9 seems to have a similar structure-stabilizing effect as arginine for the variants WW-BS1-2 and WW-BS1-3, but this is not the case when arginine is combined with Ile at position 10 as in WW-BS1-1. This suggests that Lys and Arg are not interchangeable at this sequence position. Equally charged residues at positions 7 and 9 (EE for WW-BS1-5 and KK for WW-BS1-6) resulted in significant reduction of thermal stability (WW-BS1-5: T_m = 24.5 °C and WW-BS1-6: T_m = 49.5 °C) due to charge repulsion.

The variant WW-BS1-7 was chosen to study the uncharged but polar glutamine residues at the positions 7 and 9 within the BS1 segment. The introduction of two Gln residues results in good but reduced thermostability, with a melting temperature of 51.0 °C. This highlights the structure stabilizing or destabilizing effect of charges.

The influence of the different amino acids used in the BS1 segment design on the overall thermostability of the peptide is summarized in Table 1 and Figure 13, clearly indicating that WW-BS1-1 is the most thermostable WW-BS1 peptide. After studying the different sequence combinations for the BS1 segment, the WEIRI motif was selected for the subsequent round of iterations.

Loop 1 (LP1) series

Multiple sequence alignments did not reveal a high degree of conservation for the LP1 region, therefore a simple and flexible loop sequence[106], consisting of the highly flexible Gly and Ser residues with the two different patterns $(GS)_n$ or $(SG)_n$ was examined. As the Gly at position 15 was found to be highly abundant, its relevance for the peptide's structural integrity was investigated. In addition, there is a high prevalence of oppositely charged residues at positions 11 and 16, which could form a structure-stabilizing electrostatic interaction. In total, eight different LP1 peptide variants were investigated for their structural and thermal stability.

Table 2: **Overview of designed and characterized LP1 sequence variants.** The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Introduced sequence modifications are displayed in blue. The experimentally determined melting temperatures are listed as T_m. The sequence pattern highlighted in grey was taken to the next round of iteration. This table has been adapted from Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Iteration	WW domain	Sequence ------NT------BS1----LP1-----BS2--LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
2	**WW-LP1-1**[a]	Ac-PLPPG-WEIRI-**SGSGSG**-IYFY-NHNT-KTT-TWERPRL-NH_2	51.5 ± 0.5
	WW-LP1-2[a]	Ac-PLPPG-WEIRI-**DGSGSK**-IYFY-NHNT-KTT-TWERPRL-NH_2	60.5 ± 0.5
	WW-LP1-3[a]	Ac-PLPPG-WEIRI-**DGSGSR**-IYFY-NHNT-KTT-TWERPRL-NH_2	62.5 ± 0.5
	WW-LP1-4[a]	Ac-PLPPG-WEIRI-**EGSGSR**-IYFY-NHNT-KTT-TWERPRL-NH_2	57.5 ± 0.5
	WW-LP1-5	Ac-PLPPG-WEIRI-**GSGSGS**-IYFY-NHNT-KTT-TWERPRL-NH_2	53.5 ± 0.5
	WW-LP1-6	Ac-PLPPG-WEIRI-**DSGSGR**-IYFY-NHNT-KTT-TWERPRL-NH_2	85.0 ± 0.5
	WW-LP1-7	Ac-PLPPG-WEIRI-**ESGSGR**-IYFY-NHNT-KTT-TWERPRL-NH_2	79.5 ± 0.5
	WW-LP1-8	Ac-PLPPG-WEIRI-**QSGSGQ**-IYFY-NHNT-KTT-TWERPRL-NH_2	62.0 ± 0.5

[a] Peptides showed aggregation tendency over time.

The CD spectra at 20 °C were characteristic of a WW domain, with the most pronounced exciton coupling observed for WW-LP1-8 (cf. Figure 14 A). Thermal denaturation studies revealed distinct curve patterns and varying melting temperatures for all peptide variants (cf. Figure 14 B). The LP1 variants 1 to 4, displaying a Ser at position 15, showed a high tendency to aggregate over time. This indicates the importance of Gly at position 15 for the structural integrity of the WW domain scaffold.

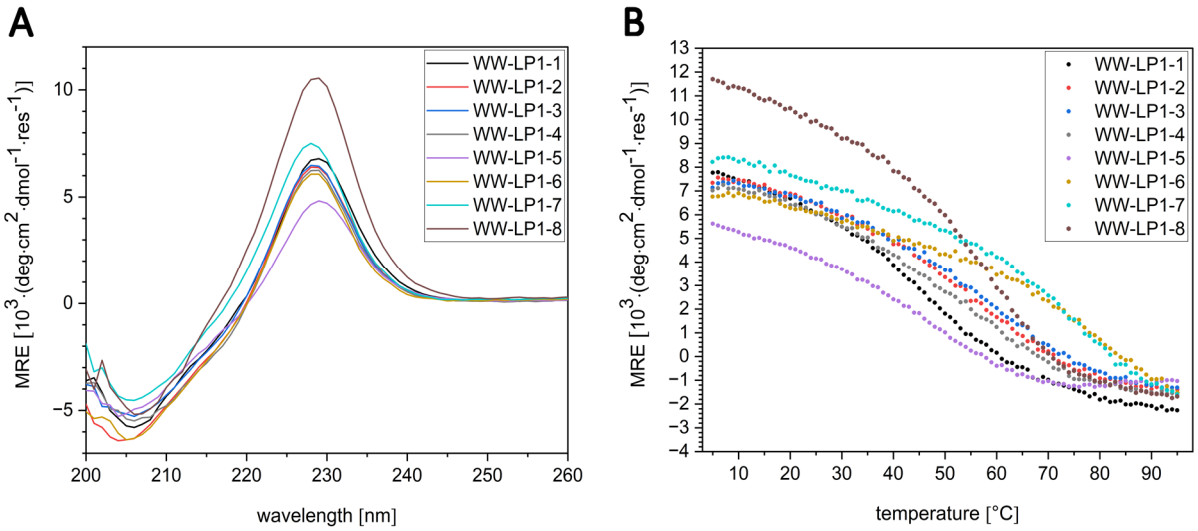


Figure 14: CD spectra and thermal denaturation curves of LP1 series. A) CD spectra were recorded in duplicates at 20 °C. **B**) Thermal denaturation curves determined at a fixed wavelength of 228 nm from 5-95 °C. For thermal denaturation curves the mean out of duplicates is shown. Conditions: 100 µM peptide, PBS buffer 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

The most thermostable LP1 peptides were WW-LP1-6 (T_m = 85.4 °C) and LP1-7 (T_m = 79.5 °C), both with oppositely charged residues at positions 11 and 16. WW-LP1-5 and WW-LP1-8 demonstrated moderate thermostability

After studying the LP1 series, it was decided to select the variant with the simplest loop with respect to its amino acid sequence, respectively WW-LP1-5 that revealed a T_m value of 53.5 °C. As it is revealed in Figure 15, WW-LP1-5 shows a complete denaturation in the performed measurement window, leading to a lower error value for the determined melting temperature, when compared with the two variants WW-LP1-6 and WW-LP1-7. Since the aim was to investigate further structure-stabilizing sequence patterns in the following structural segments, the sequence pattern, exhibiting a GSGSGS loop with moderate thermostability was considered for subsequent round of iterations.

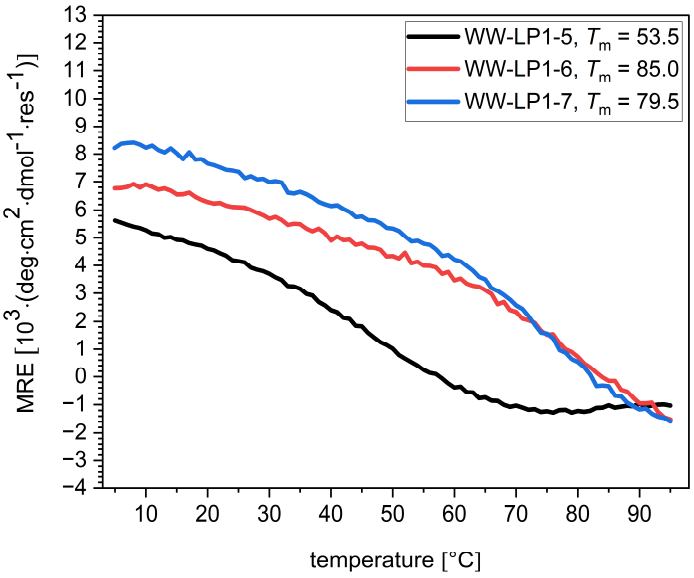


Figure 15: Thermal denaturation profiles of WW-LP1-5, WW-LP1-6 and WW-LP1-7. Thermal denaturation curves of WW-LP1-6 and WW-LP1-7 do not reach the lower baseline within the measurement range (5 – 95 °C), leading to higher error values, when a two-state model is fitted. In contrast, the fitting for WW-LP1-5 is more accurate. Conditions: 100 µM peptide, PBS buffer 7.5. This figure is reused from Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

β-Sheet 2 (BS2) series

The β-Sheet 2 (BS2) segment (17-20) is rich in aromatic residues. According to Macias *et al.*, especially Tyr residues are involved in cross-strand hydrogen bonds, the Tyr/Phe at sequence position 19 is considered to be highly conserved, since it is part of the hydrophobic core.[77] The high prevalence of aromatic residues in BS2, was also elucidated by alanine and glycine scanning by Jäger *et al.*, which showed a high degree of destabilization or even unfolding of *h*Pin1$_{WW}$, when Tyr19 was replaced by Ala.[89] Taking into account the results of those previously reported mutagenesis studies, peptide variants containing mainly polar and aromatic residues at the sequence positions 18-20 were considered and tested for structural and thermal stability. For position 17, Ile was selected as a β-branched amino acid. Finally, four different peptide variants were selected for BS2 engineering (cf. Table 3).

Table 3: **Overview of designed and characterized BS2 sequence variants.** The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Introduced sequence modifications are displayed in blue. The experimentally determined melting temperatures are listed as T_m. The sequence pattern highlighted in grey was taken to the next round of iteration. This table has already been published in minorly modified form in Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

Iteration	WW domain	Sequence ------NT------BS1----LP1-----BS2--LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
3	**WW-BS2-1**	Ac-PLPPG-WEIRI-GSGSGS-**IYYY**-NHNT-KTT-TWERPRL-NH_2	58.0 ± 0.5
	WW-BS2-2	Ac-PLPPG-WEIRI-GSGSGS-**ISYY**-NHNT-KTT-TWERPRL-NH_2	46.5 ± 0.5
	WW-BS2-3	Ac-PLPPG-WEIRI-GSGSGS-**IFFY**-NHNT-KTT-TWERPRL-NH_2	51.5 ± 0.5
	WW-BS2-4	Ac-PLPPG-WEIRI-GSGSGS-**ISFY**-NHNT-KTT-TWERPRL-NH_2	37.0 ± 1.0

All four BS2 variants displayed a CD spectrum, which was characteristic of a WW domain with the most pronounced exciton signal found in WW-BS2-1 (cf. Figure 16 A). The other three BS2 peptides show a more distinct minimum at 215 nm. However, all four designed peptides were stably folded at 20 °C. When analyzing the thermal stabilities of all BS2 variants, the high initial mean residue ellipticity (MRE) signal for WW-BS2-1 and its slow decrease with increasing temperature, compared to the other three BS2 variants is noticeable (cf. Figure 16 B). In addition, the steepness and general shape of the curves, including the melting temperature, were different for all four variants (cf. Table 3).

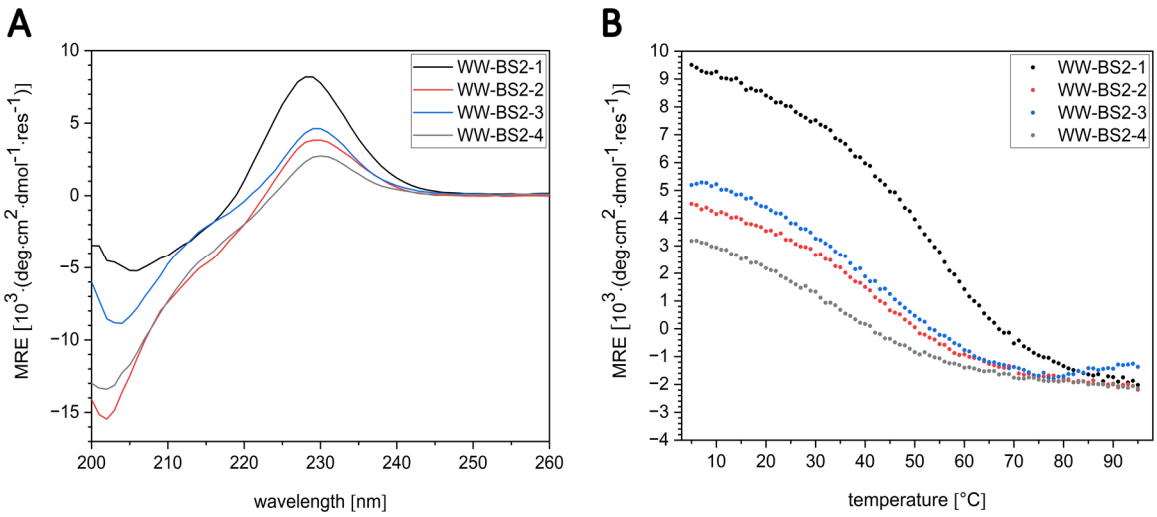


Figure 16: CD spectra and thermal denaturation curves of the BS2 series. A) CD spectra were recorded in duplicates at 20 °C. **B**) Thermal denaturation curves determined at a fixed wavelength of 228 nm from 5-95 °C. For thermal denaturation curves the mean out of duplicates is shown. Conditions: 100 µM peptide, PBS buffer pH 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

The BS2 variant with the sequence pattern `IYYY` exhibited the highest thermostability of all variants with a melting temperature of 58.0 °C. When Tyr is replaced by Phe at sequence position 19 the T_m value is reduced by almost 5 K, showing the importance of Tyr for the hydrophobic core packing. In addition to the impact of aromatic amino acids at sequence position 19, the influence of the Tyr18Phe mutation was examined in WW-BS2-3. This amino acid substitution resulted in a thermal destabilization, with a decrease of 6 K (cf. Table 3). This phenomenon may be attributed to the fact that Phe is unable to function as a hydrogen bond donor.[107] The impact of non-aromatic amino acids that can function as hydrogen bond donors, such as Ser, have also been studied. A comparison of the melting temperatures of the two variants with Ser at position 18 (cf. Table 3), WW-BS2-2 and WW-BS2-4, reveals that both peptides exhibit significantly lower melting temperatures than WW-BS2-1, which contains the aromatic Tyr. For WW-BS2-2, a decrease of approximately 11 K is observed, while for WW-BS2-4, the decrease is more than 21 K.

The results of the BS2 segment screening demonstrated that `IYYY` exhibited the greatest propensity for structural stabilization. This sequence pattern was subsequently employed in the LP2 (21-24) round of iterations.

Loop 2 (LP2) series

The LP2 sequence contains a highly conserved structure-stabilizing Asn at position 21, which was previously reported by Jäger *et al*.[89] and computationally confirmed by the multiple sequence alignment presented above (cf. Appendix, Table S1). Consequently, N21 was not altered in the LP2 series. With regard to the remaining sequence positions, the approach employed to incorporate flexible (Ser, Gly, Asn) and β-branched (Ile, Thr) amino acids, in a manner analogous to the iterative round of LP1, was utilized. A total of nine distinct sequence variations were tested in order to assess their impact on structural and thermal stability (cf. Table 4).

Table 4: **Overview of the sequences and melting temperatures of the LP2 series.** The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Introduced sequence modifications are displayed in blue. The experimentally determined melting temperatures are listed as T_m. The sequence pattern highlighted in grey was taken to the next round of iteration. This table has already been published in minorly modified form in Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Iteration	WW domain	Sequence ------NT------BS1----LP1-----BS2--LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
4	**WW-LP2-1**	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NSGS**-KTT-TWERPRL-NH_2	42.5 ± 0.5
	WW-LP2-2	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NEGS**-KTT-TWERPRL-NH_2	36.5 ± 0.5
	WW-LP2-3	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NSNS**-KTT-TWERPRL-NH_2	49.0 ± 0.5
	WW-LP2-4	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NSNT**-KTT-TWERPRL-NH_2	57.0 ± 0.5
	WW-LP2-5	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NSIT**-KTT-TWERPRL-NH_2	64.5 ± 0.5
	WW-LP2-6	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NSLT**-KTT-TWERPRL-NH_2	63.5 ± 0.5
	WW-LP2-7	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NGSG**-KTT-TWERPRL-NH_2	< 0
	WW-LP2-8	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NGNG**-KTT-TWERPRL-NH_2	3.0 ± 5.0
	WW-LP2-9	Ac-PLPPG-WEIRI-GSGSGS-IYYY-**NGIT**-KTT-TWERPRL-NH_2	54.0 ± 0.5

Structure analysis based on CD spectral measurements revealed the characteristic maximum of the WW domain at 228 nm for all LP2 variants. However, the two peptides WW-LP2-7 and WW-LP2-8 show a significantly reduced exciton signal compared to all other LP2 sequences (cf. Figure 17 A), all of which have an average residue ellipticity signal above 5. Consistent with the observations from the CD spectral measurements, the thermal denaturation profiles of WW-LP2-7 and WW-LP2-8, both with a sequence pattern of NGXG, reveal melting temperatures < 3 °C (cf. Figure 17 B), indicating that Gly as a flexible residue located at the sequence positions 22 and 24 leads to a loss in structural integrity.

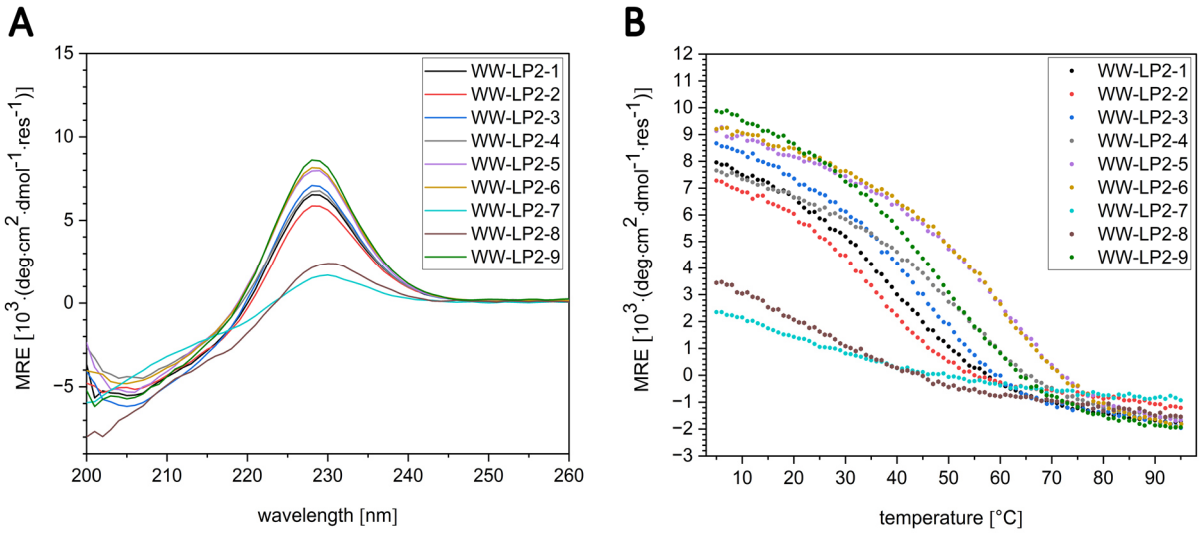


Figure 17: CD spectra and thermal denaturation curves of LP2 variants. A) CD spectra were recorded in duplicates at 20 °C. **B**) Thermal denaturation curves determined at a fixed wavelength of 228 nm from 5-95 °C. For thermal denaturation curves the mean out of duplicates is shown. Conditions: 100 µM peptide, PBS buffer 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

A Gly residue combined with the amino acids Ile and Thr with high β-propensity[103] at the remaining two amino acid positions in WW-LP2-9 leads to a thermostable variant with a melting temperature of 54 °C. Apart from Gly, the influence of the flexible and polar Ser, the charged amino acid Glu, as well as the hydrophobic Ile residue on the structural stability were investigated. All peptide variants with a Ser at position 22, exhibited melting temperatures that were higher than the variant with a charged Glu, WW-LP2-2 (cf. Table 4), indicating that Ser is the most structure-stabilizing amino acid at this position. Although Glu22 might form a salt bridge with Lys25 in WW-LP2-2, the overall thermostability of this variant is significantly lower than that of all variants with a Ser at position 22.

In contrast to the destabilizing effect, mediated by the introduction of Asn at position 23, Ile provides more flexibility for the hydrophobic core residues and increases the probability of interactions with them. Thermal denaturation studies with CD spectroscopy confirmed this hypothesis, as WW-LP2-5 revealed a significantly higher melting temperature of 64.5 °C than WW-LP2-4 with 57.0 °C, clearly indicating that an introduced Ile residue at sequence position 23 is structure-stabilizing.

The round of LP2 iteration included an investigation of the influence of the different polar residues Ser and Thr at sequence position 24 and their hydrogen bonding ability on the overall peptide stability. In general, LP2 variants with Thr at position 24, showed increased melting temperatures compared to WW-LP2-3 (with Ser24) (cf. Table 4). Analysis of the predicted secondary structures of WW-LP2-5 (with Thr24), the most thermostable LP2 peptide variant, compared to WW-LP2-3, showed that the variant with a threonine residue provides two polar interactions: an intra-loop polar interaction with Asn21 and one with Thr26 in the BS3 segment, leading to a higher structural stability than the Ser24-containing WW-LP2-3 variant.

In conclusion, the round of LP2 iterations revealed the importance of polar residues at positions 22 and 24 due to the hydrogen bonding ability. Ser at position 22 was structurally preferred, whereas the introduction of a Thr residue at position 24, combined with an

hydrophobic Ile at position 23, resulted in the most thermostable peptide variant. As a consequence, the NSIT sequence pattern was chosen for the next round of iterations.

β-Sheet 3 (BS3) series

The third β-sheet region (sequence positions 25-27) in WW domains is known to be threonine-rich.[77] Previous results from multiple sequence alignments (cf. Appendix, Table S1) indicated a high prevalence of positively charged amino acids, such as Lys or Arg at position 25, as depicted in the WebLogo[98,99] (cf. Figure 12 A). Therefore, several sequence variations with either Lys or Arg at the first BS3 position were tested. In addition, the Thr residue at position 26 was substituted with Ile in one variant. For position 27, either polar residues, such as Ser and Thr, were introduced or the influence of the hydrophobic amino acid Ala and the polar Asn was investigated. In total, seven different variants were studied with regard to their structure and thermostability (cf. Table 5).

Table 5: **Overview of the sequence variations of the BS3 series.** The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Introduced sequence modifications are displayed in blue. The experimentally determined melting temperatures are listed as T_m. The sequence pattern highlighted in grey was taken to the next round of iteration. This table has already been published in minorly modified form in Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Iteration	WW domain	Sequence ------NT------BS1----LP1-----BS2--LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
5	**WW-BS3-1**	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-**KTS**-TWERPRL-NH_2	63.0 ± 0.5
	WW-BS3-2	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-**KTN**-TWERPRL-NH_2	50.5 ± 0.5
	WW-BS3-3	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-**KTA**-TWERPRL-NH_2	64.0 ± 0.5
	WW-BS3-4	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-**RTT**-TWERPRL-NH_2	62.5 ± 0.5
	WW-BS3-5	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-**RTS**-TWERPRL-NH_2	61.5 ± 0.5
	WW-BS3-6	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-**RTA**-TWERPRL-NH_2	62.5 ± 0.5
	WW-BS3-7[a]	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-**RIT**-TWERPRL-NH_2	48.5 ± 0.5

[a] Peptides showed aggregation tendency over time.

The structure analysis of the BS3 peptide variants by CD spectroscopy revealed spectral characteristics of a WW domain with a maximum at 228 nm, which can be attributed to the exciton coupling between Trp6 and Tyr19 (cf. Figure 18 A). The variant WW-BS3-7 shows decreased intensity of the maximum and a less pronounced minimum compared to all other BS3 peptides. In accordance with the findings of the structural analysis conducted at 20 °C, WW-BS3-7 exhibits a low thermal stability, with a T_m of 48.5 °C and a high tendency to aggregate over time. Consequently, the sequence pattern of RIT in the BS3 segment, was not further considered in the iterative sequence analysis. However, the structure and the thermal denaturation profiles of the other BS3 variants display a high degree of similarity, as well as resemblance in melting temperatures, particularly for WW-BS3-1 and WW-BS3-3 to WW-BS3-6 (cf. Figure 18 B).

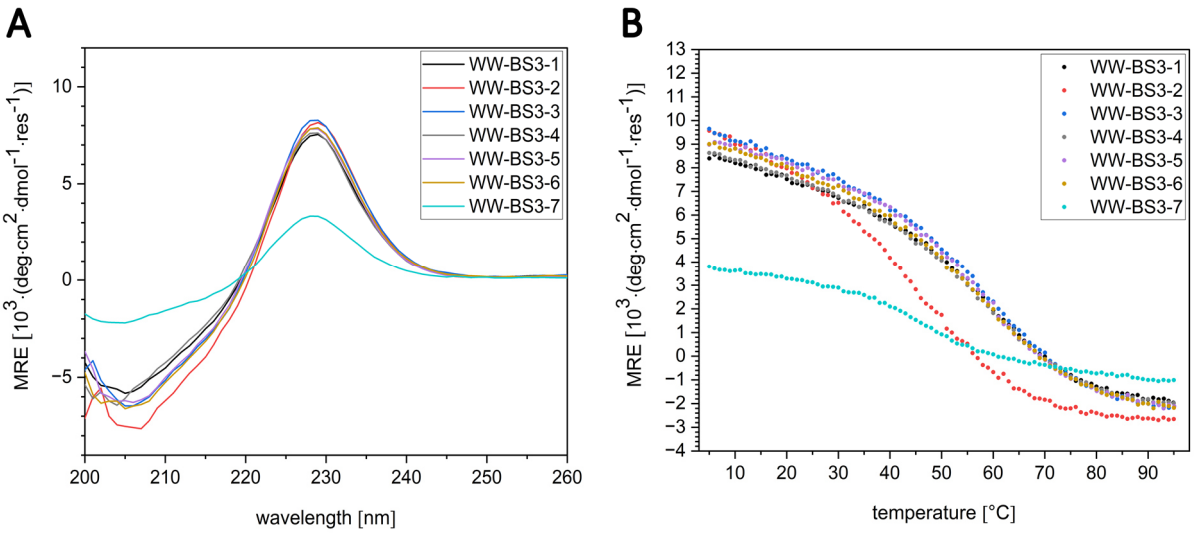


Figure 18: CD spectra and thermal denaturation curves of peptide variants of the BS3 series. A) CD spectra were recorded in duplicates at 20 °C. **B**) Thermal denaturation curves determined at a fixed wavelength of 228 nm from 5-95 °C. For thermal denaturation curves the mean out of duplicates is shown. Conditions: 100 µM peptide, PBS buffer 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

The previously obtained result of the WW domain-based sequence alignments, summarized in the WebLogo[98,99], indicates a preference for postitively charged amino acids, such as Arg or Lys, at the sequence position 25 in the BS3 segment. To ascertain whether a preferred residue exists, analogous sequence patterns were examined, including WW-LP2-5 and WW-BS3-4, WW-BS3-1 and WW-BS3-5, or WW-BS3-3 and WW-BS3-6 (cf. Table 5). Upon analysis of the melting temperatures of the aforementioned sequence patterns, it was observed that peptide variants exhibiting a Lys residue at position 25 displayed only a 1 K higher melting temperature than those with Arg, suggesting that both residues are interchangeable.

However, an exchange of threonine at position 26 results in a destabilization of the peptide structure and aggregation, as it was previously observed for WW-BS3-7. The Thr residue forms structurally important interstrand polar interactions with Asn21 and Thr24, both of which are located in the LP2 segment. These interactions are of significant importance for the overall structural integrity of the peptide.

At sequence position 27 of the short BS3 segment, the influence of polar residues, including Thr, Ser, and Asn, as well as the hydrophobic Ala, on the peptide's structural stability was studied. In general, all tested amino acids led to thermostable variants with melting temperatures ranging from 50.5 °C for WW-BS3-2 to 61.5 °C for WW-BS3-5 (cf. Table 5). This indicates a low degree of conservation at the final BS3 sequence position.

Nevertheless, introducing a polar Thr at sequence position 27 that displays a high β-propensity[103], resulted in the most thermostable variant, WW-LP2-5 with a sequence pattern of KTT, which was taken to the final round of iterations.

C-terminal (CT) series

C-terminal modifications were conducted to investigate the postulated relevance of Trp29 in WW domains.[77,80,89,92,108] This was achieved by substituting the Trp with an Ala. Furthermore, the impact of replacing the charged residue Glu with the polar, but uncharged Gln was studied. Another objective was to examine the effect of a C-terminal truncation. In total, three different peptide variants (cf. Table 6) were studied with regard to their structure and thermal stability.

Table 6: Overview of the sequences of the CT series. The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Introduced sequence modifications are displayed in blue. The experimentally determined melting temperatures are listed as T_m. The sequence pattern highlighted in grey was taken to the next round of iteration. This table has already been published in minorly modified form in Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

Iteration	WW domain	Sequence ------NT------BS1----LP1-----BS2--LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
6	WW-CT-1	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-KTT-**TAERPRL**-NH_2	61.5 ± 0.5
	WW-CT-2	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-KTT-**TWQRPRL**-NH_2	63.0 ± 0.5
	WW-CT-3	Ac-PLPPG-WEIRI-GSGSGS-IYYY-NSIT-KTT-**TWERP** -NH_2	63.5 ± 0.5

All three peptide variants exhibited similar CD spectra with the characteristic exciton signal at around 228 nm (cf. Figure 19 A). Thermal denaturation studies revealed that the C-terminally modified variants showed very good thermostability (cf. Figure 19 B), with melting temperatures ranging from 61.5 to 63.5 °C (cf. Table 6). This indicates that the studied mutations do not greatly influence the thermal and thus structural stability of the WW domain scaffold.

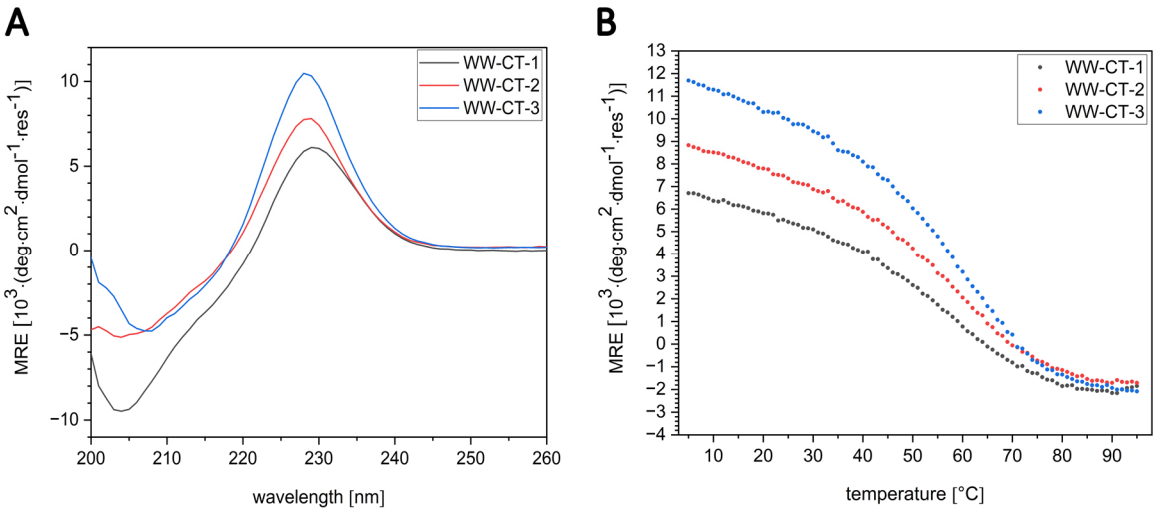


Figure 19: CD spectra and thermal denaturation profiles of the CD series. A) CD spectra were recorded in duplicates at 20 °C. **B**) Thermal denaturation curves determined at a fixed wavelength of 228 nm from 5-95 °C. For thermal denaturation curves the mean out of duplicates is shown. Conditions: 100 µM peptide, PBS buffer 7.5.Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

Previously Ala scanning from Jäger *et al.* revealed that the C-terminal Trp residue is highly conserved due to its suggested role in the structure and function of WW domains.[89] The results of thermostability studies, presented in Figure 19 B, indicate that the exchange of the Trp residue to Ala does not significantly influence thermostability. This suggests that Trp29 plays a more important role in the function of a WW domain, such as the ligand binding. Since Ala belongs to the group of aliphatic, hydrophobic amino acids, it can only form stabilizing hydrophobic interactions with neighboring side chains.[109] In contrast, tryptophan, which is classified as an aromatic amino acid, is able to interact via polar interactions, aromatic π-π or cation-π stacking, in addition to its ability to form hydrophobic interactions.[110]

The mutation of Glu30 to the uncharged, hydrophilic glutamine residue in WW-CT-2 was investigated with the objective of studying the influence of a charged amino acid in the C-terminal segment on the overall peptide structural and thermal stability. The Glu30Gln mutation appears to have a minimal destabilizing impact on the overall peptide‘s thermostability, resulting in a reduction of approximately 1 K in melting temperature.

The C-terminally truncated variant, WW-CT-3, displays a broader thermal denaturation profile compared to the other two peptide variants (cf. Figure 19 B). Nevertheless, thermal denaturation of WW-CT-3 revealed a very good melting temperature of 63.5 °C, a decrease of approximately 1 °C, compared to WW-LP2-5 with 64.5 °C.

The study of C-terminal modifications in WW domain-based peptides revealed that the position 29 is tolerant to mutations. However, the C-terminal segment of WW-CS was maintained without variations for the generation of sequence-structure-stability relationships (cf. Chapter 4.1.3) and the subsequent design of a thermostable scaffold (cf. Chapter 4.1.4).

4.1.3 Sequence-structure-stability relationships of a WW domain core scaffold

The reported results of the iterative sequence analysis indicate that the 38 investigated WW domain-derived peptides showed a high degree of structural robustness with respect to amino acid mutations. Despite the occurrence of destabilized and aggregation-prone variants when some sequence positions were varied, the overall structural stability and folding propensity of the variants remained intact.

The most important findings with respect to structural stability are as follows: 1. Oppositely charged amino acids at sequence positions 7 and 9 in BS1 and positions 11 and 16 in LP1 for intra-segment stabilization by electrostatic interactions; 2. Enlargement of the hydrophobic core by inserting hydrophobic amino acids, specifically Ile, into the BS1 and LP2 segment; 3. Tyr at sequence position 19 in BS2; 4. Positively charged residues at the first position of BS3 and the structural importance of Thr at position 26 in BS3.

These sequence-structure-relationships are visualized in Figure 20 and were employed in the design of highly thermostable WW domains.

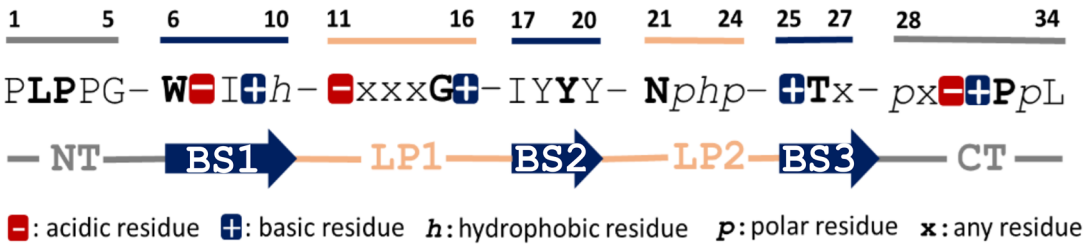


Figure 20: Sequence-structure relationships in WW domains identified by an iterative sequence analysis. The amino acid sequence to obtain thermostable peptide variants is displayed with highly conserved residues in bold. Sequence positions of the different segments are indicated with numbers. Positively charged residues are represented with a "+" and negatively charged residues with a "-" symbol. This figure has already been published in Lindner *et al.*[94], and was minorly adapted with permission from ©2024 American Chemical Society.

4.1.4 Design of highly thermostable WW-domains

4.1.4.1 Synthesis, structure and stability

After the successful identification of structure- and stability-relevant amino acids in the WW domain scaffold, the next goal was to combine these sequence motifs to design a highly thermostable variant that could serve as a core scaffold for subsequent functionalization. To identify the most suitable candidate, three different variants were designed and characterized. One variant, WW-HS-1, contains an aspartate residue at position 11 in LP1. Due to the fact that in peptides with one or more Asp residues unwanted aspartimide formation can occur during SPPS[111], a CSY-protected Asp was used in SPPS with an optimized protocol from Pham *et al.*[112] In addition, another peptide variant, WW-HS-2, containing the Glu at position 11, was studied. To also test the influence of capping of the termini on the thermostability of the peptide, WW-HS-2 with free N-terminus and C-terminal acid, named WW-HS-2_{OH}, was synthesized and characterized.

Table 7: Overview of designed HS variants. The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Amino acids or terminal modifications that differ in-between the variants are displayed in red and BS regions are shown in blue. This table has already been published in modified form in Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

WW domain	Sequence ------NT------BS1----LP1-----BS2--LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
WW-HS-1	Ac-PLPPG-WEIRI-DSGSGS-IYYY-NSIT-KTT-TAERPRL-NH_2	93.5 ± 0.5
WW-HS-2	Ac-PLPPG-WEIRI-ESGSGS-IYYY-NSIT-KTT-TWQRPRL-NH_2	89.0 ± 0.5
WW-HS-2_{OH}	H-PLPPG-WEIRI-ESGSGS-IYYY-NSIT-KTT-TWERP -OH	88.5 ± 0.5

CD spectroscopy of all three HS variants revealed the characteristic spectrum of a WW domain with an exciton signal at 228 nm (Figure 21 A). In comparison to WW-CS, the maximum in the CD spectra of the three designed HS peptides is significantly increased. Moreover, the hyperstable peptides reveal distinct thermal denaturation profiles in comparison to WW-CS. The denaturation curves of the WW-HS series exhibit a broad upper plateau, extending almost until 80 °C, followed by a steep transition and an indicated lower plateau above 100 °C. In

contrast, WW-CS shows a rather fast transition from a folded to an unfolded state (cf. Figure 21 B). WW-HS-1 displays a melting temperature of 93.5 °C (cf. Table 7). However, WW-HS-2 and the uncapped variant WW-HS-2_{OH} also reveal melting temperatures of approximately 89 °C, which represents an increase of about 36 K compared to WW-CS that displayed a T_m of 53.5 °C.

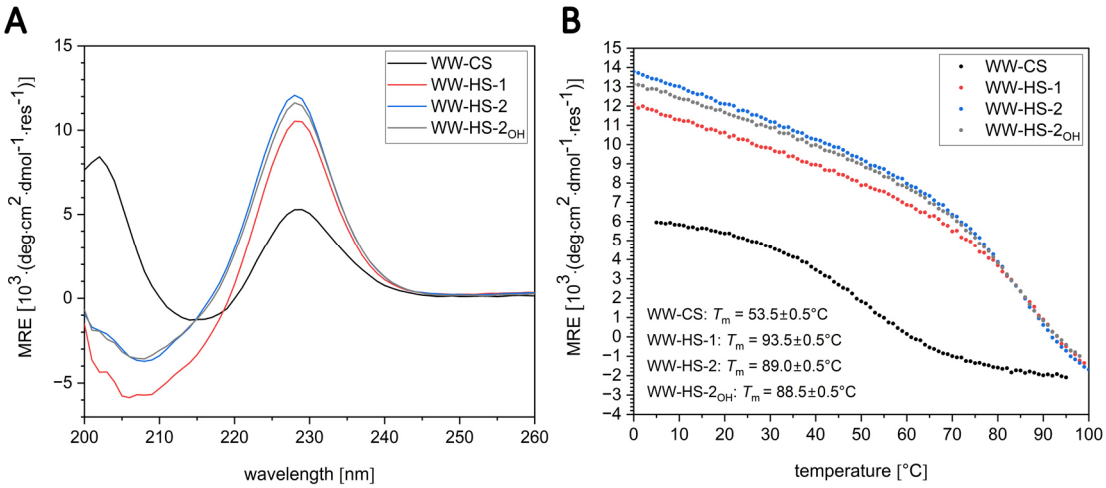


Figure 21: CD spectra and thermal denaturation curves of WW-HS variants. A) CD spectra were recorded in duplicates at 20 °C. **B**) Thermal denaturation curves were determined at a fixed wavelength of 228 nm from 0-98 °C. For thermal denaturation curves the mean out of duplicates is shown. Conditions: 100 µM peptide, PBS buffer 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

4.1.4.2 *In silico* structure analysis

To obtain an initial understanding of the impact of polar and hydrophobic interactions on the secondary structures of WW-HS-1 and WW-HS-2 in comparison to WW-CS, both hyperstable variants were subjected to structural prediction using *AlphaFold2* and *OmegaFold*. Two distinct prediction tools were employed to ascertain the reliability of the outcomes and to identify the optimal approach for the structural prediction of WW domain-based peptides.

Both prediction tools display the structural segments as the three anti-parallel β-sheets and the loop regions for WW-HS-1 and WW-HS-2 with a high degree of similarity. Nevertheless, discrepancies can be observed in the predicted orientation of the BS2 and the disordered N- and C-terminal segments (cf. Figure S39, Appendix). These minor differences can be attributed to the distinct methodologies and algorithms employed by the two prediction tools *AlphaFold2*[100] and *OmegaFold*[113].

When the predicted structures of both WW-HS peptide variants are analyzed in comparison with WW-CS, it becomes evident that there are substantial differences in the orientation of the loop regions, in addition to the BS2 segment and the termini (cf. Figure S39, Appendix). The distinct orientations of the described structural segments are the result of the introduced amino acid mutations and the changed interaction pattern. One of the stabilizing interactions that were considered after performing iterative sequence analysis was the extension of the hydrophobic core by introducing isoleucine at sequence position 8 in BS1 and position 17 in BS2 in both hyperstable designs. The mutation of Glu8 in WW-CS to an isoleucine residue in WW-HS-1 and WW-HS-2 enabled hydrophobic interactions between Ile8 and the highly

conserved residues Leu2 and Tyr19 of the hydrophobic core. In contrast to the polar and charged residue Glu8 in WW-CS, the hydrophobic Ile8 is situated in close proximity to Leu2 and Tyr19 (cf. Figure S39, Appendix), resulting in the anticipated structure-stabilizing enlargement of the hydrophobic core.

A more detailed analysis of the respective interaction patterns of WW-HS-1 and WW-HS-2 revealed the presence of structure-stabilizing inter- and intra-segment interactions in the predicted structures. Starting with the analysis of WW-HS-1, the hydrophobic core (red) is enlarged by the previously described introduction of Ile (rose) in BS1 and LP2 (cf. Figure 22 A). Furthermore, an extensive polar hydrogen bond pattern was identified within LP1 between Asp11 and Gly13, Ser14 and Gly15, as well as an electrostatic interaction between Asp11 and Arg16. In addition, two electrostatic interactions within BS1 between Glu7 and Arg9, an inter-strand stabilizing cation-π interaction between Trp29 and Arg16, as well as an intrasegment stabilization of the C-terminus with polar interactions between Thr28 with Trp29 and Glu30 were found (cf. Figure 22 B). The combination of these inter- and intrasegment interactions of varying natures, resulted in the remarkable stability of the WW-HS-1 variant, with a melting temperature of 93.5 °C.

Some interactions, such as the electrostatic interaction within LP1 between the oppositely charged amino acids Asp11 in WW-HS-1 or Glu11 in WW-HS-2 and Arg16, as well as the salt bridge between Glu7 and Arg9 in BS1, are identical in both HS peptide variants. Nevertheless, a single amino acid mutation at position 11 from Asp in WW-HS-1 to Glu in WW-HS-2 results in a distinct cation-π interaction, hydrogen bond pattern, and electrostatic interaction (cf. Figure 22 D). In terms of cation-π interactions, WW-HS-2 shows the presence of three of these. Two cation-π interactions were identified between Trp29 and Arg16, whereas WW-HS-1 only displays one aromatic interaction between the two residues in total. The third cation-π interaction in WW-HS-2 is formed between Arg9 of BS1 and Tyr20 of BS2. In contrast to the extensive intersegment aromatic interactions, WW-HS-2 displays a reduced number of polar hydrogen bonds within the LP1 segment compared to WW-HS-1. In the LP1 region, the polar interactions of Glu11 in WW-HS-2 were observed to be limited to those with Ser14 and Gly15, with no interactions identified for Gly13, which was not present in the peptide WW-HS-1, containing an aspartate amino acid at sequence position 11. Referring to the intra-segment BS1 electrostatic interaction between Glu7 and Arg9, WW-HS-1 exhibits two salt bridges (cf. Figure 22 B), whereas WW-HS-2 displays only one.

In conclusion, the two highly thermostable peptides WW-HS-1 and WW-HS-2, which were generated from sequence-structure-stability relationships based on iterative segment analysis, represent some notable differences in their intra- and inter-segment connections. Although the two designs differ only in one single amino acid at sequence position 11 and when structurally aligned show little variation in their side chain orientation (cf. Figure S39, Appendix), they display significant differnces in their interaction patterns. The results of the interaction analysis of the *in silico* predicted structures indicate that WW-HS-1 exhibits a greater number of electrostatic interactions and a more extensive hydrogen bond pattern.

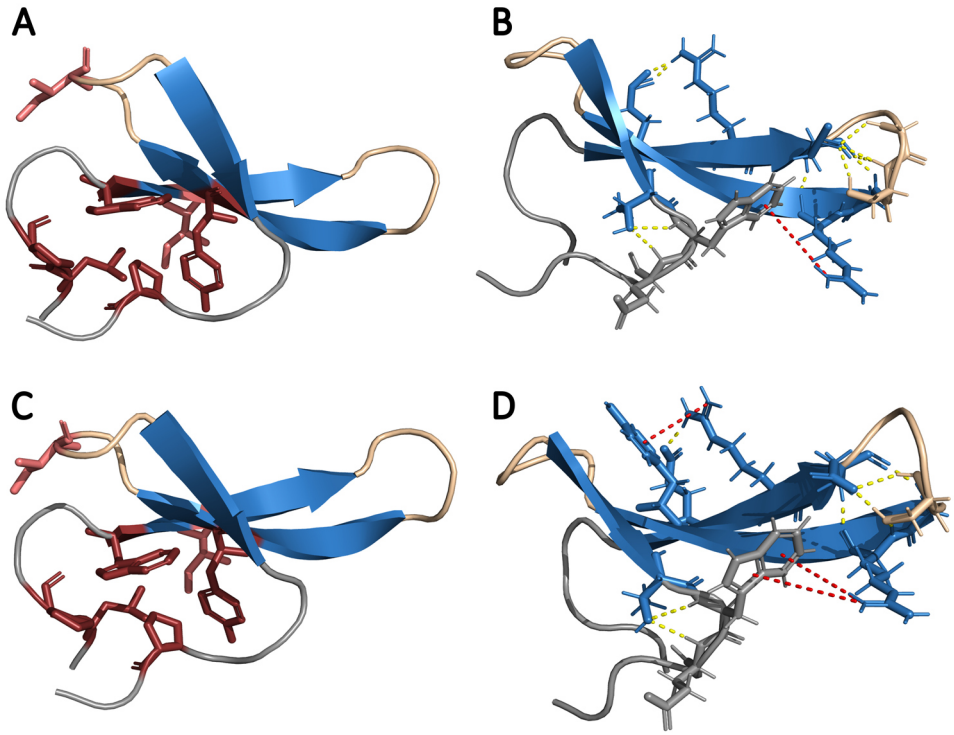


Figure 22: Visualization of the *in silico* predicted thermostable variants WW-HS-1 and WW-HS-2. A), **B**) Predicted structure of WW-HS-1 with *AlphaFold2*[100]. **C**), **D**) Predicted structure WW-HS-2 with *AlphaFold2*[100] . BS regions are shown in blue, LP regions in beige and the peptide termini in grey. **A**), **C**) Hydrophobic core residues are illustrated in red, the two Ile residues are displayed as rose sticks. **B**), **D**) Residues important for polar interactions (yellow), and electrostatic interactions (red) are represented.

4.1.4.3 *In silico* structure analysis

To investigate the reversibility of unfolding of the WW-HS variants, cooling curves were recorded subsequent to thermal denaturation experiments and the CD spectra were recorded at 20°C before and after the thermal de- and renaturation.

All three hyperthermostable variants show significantly decreased signals in the CD spectra following thermal renaturation (cf. Figure 23 A, C and E). This result indicates that thermal denaturation is not reversible.

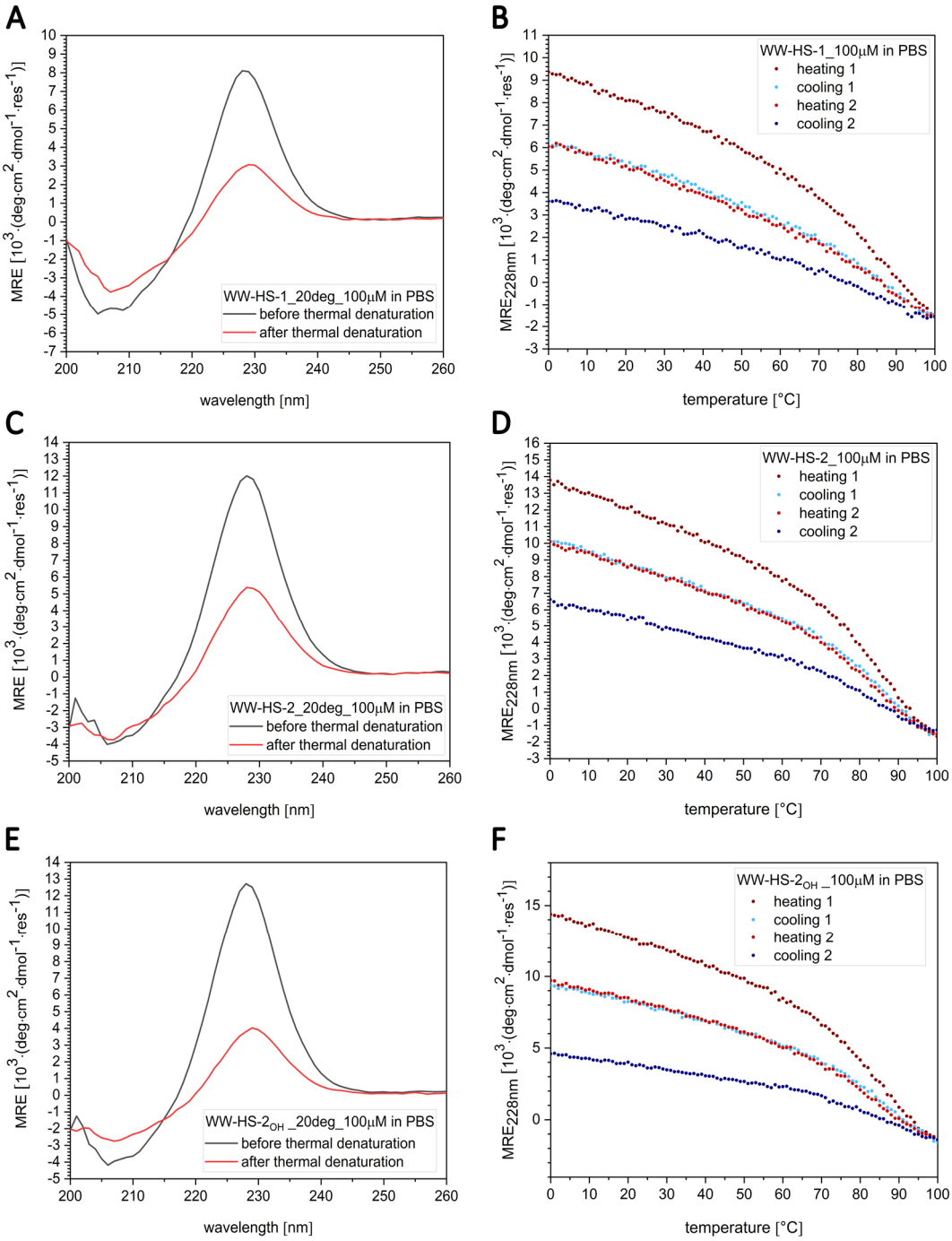


Figure 23: CD spectra before and after thermal denaturation and heating and cooling curves of the WW-HS series. CD spectra for **A**) WW-HS-1, **C**) WW-HS-2, **E**) WW-HS-2_{OH} were recorded in duplicates at 20 °C before and after thermal denaturation experiments. Thermal denaturation curves for **B**) WW-HS-1, **D**) WW-HS-2, **F**) WW-HS-2_{OH} were determined at a fixed wavelength of 228 nm from 0-98 °C. For thermal denaturation curves the mean out of duplicates is shown. For each variant two heating and cooling curve measurements were performed. Conditions: 100 µM peptide, PBS buffer, pH 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

In order to investigate whether the refolding property can indeed be attained by a rational redesign of WW-HS, we performed a redesign of BS1, since we assumed that the extension of the hydrophobic core in this segment might lead to the not fully reversible thermal unfolding of the WW-HS variants. For the BS1 redesign charged or polar residues were introduced, leading to three redesigned WW-HS peptide variants with different BS1 sequences (cf. Table

8). All three redesigned peptide variants were characterized by CD spectroscopy and thermal denaturation experiments, including thermal heating and cooling.

Table 8: Overview of redesigned HS variants. The peptide termini are abbreviated as NT and CT, β-sheet sequence regions as BS and loop segments as LP. Residues of the BS regions are displayed in blue. The experimentally determined melting temperatures are listed as T_m values. This table has already been published in minorly modified form in Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

WW domain	Sequence ------NT------BS1----LP1-----BS2--LP2--BS3---CT-- 1 - 10 - 20 - 30	T_m / °C
WW-HS-3	Ac-PLPPG-WEQRI-ESGSGS-IYYY-NSIT-KTT-TWERPRL-NH_2	81.0 ± 0.5
WW-HS-4	Ac-PLPPG-WEQRT-ESGSGS-IYYY-NSIT-KTT-TWERPRL-NH_2	76.5 ± 0.5
WW-HS-5	Ac-PLPPG-WEERK-ESGSGS-IYYY-NSIT-KTT-TWERPRL-NH_2	78.5 ± 0.5

First, the impact of the amino acid modifications in the BS1 segment on its structural and thermal stability was investigated by CD spectroscopy. As illustrated in Figure 24 A, all three redesigned variants show the characteristic spectrum of a WW domain. Notably, WW-HS-4 and WW-HS-5 exhibit the most pronounced exciton signals, even if compared to the first generation of WW-HS (WW-HS-1, WW-HS-2, and WW-HS-2_{OH}). With regard to thermal stability, WW-HS-3 is the most thermostable peptide, exhibiting a melting temperature of 81.0 °C within the second generation of WW-HS peptides (cf. Table 8). However, all three redesigned WW-HS variants show decreased thermostabilities in comparison to the first WW-HS generation (cf. Table 8), which emphasizes the importance of the hydrophobic core extension on thermostability.

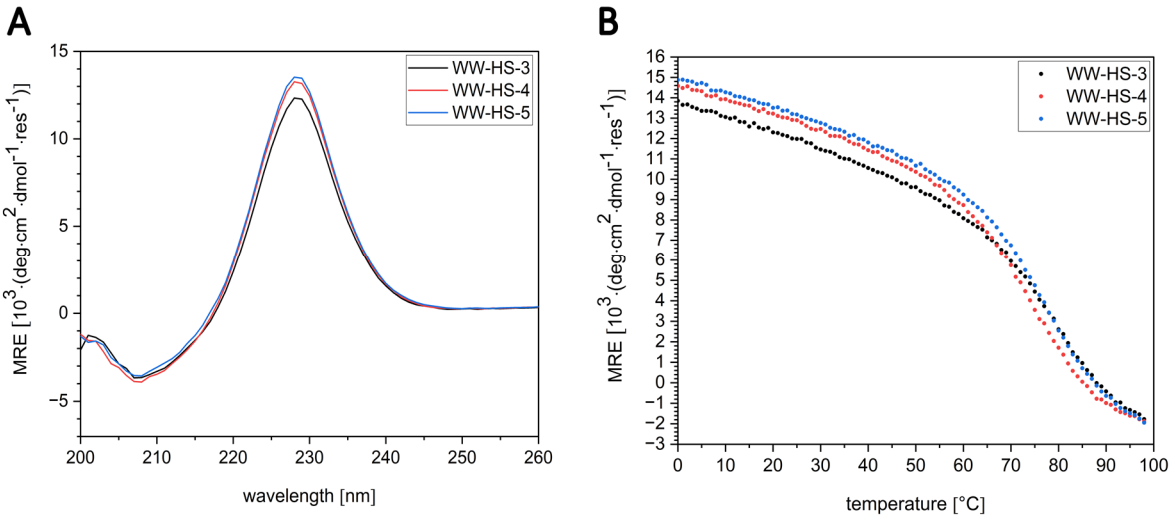


Figure 24: CD spectra and thermal denaturation curves of redesigned WW-HS variants. A) CD spectra were recorded at 20 °C. **B**) Thermal denaturation curves were determined at a fixed wavelength of 228 nm from 0-98 °C. For CD spectra and thermal denaturation curves the mean of triplicate is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

Apart from the decreased thermostability in the second WW-HS generation as a result of the BS1 redesign, the primary objective was to identify hyperthermostable peptide variants that show reversible denaturation. Consequently, all three variants were subjected to a

comprehensive analysis of their CD spectra, both before and after thermal denaturation, as well as their melting and cooling profiles. The CD spectra of all three redesigned peptides were characteristic of a WW domain (cf. Figure 25 A, C, E). However, the CD signals exhibited a significant decrease following thermal denaturation. Analyses of the heating and cooling curve profiles of the redesigned peptides revealed a signal decrease between heating and cooling curves (cf. Figure 25 B, D, F), indicating non-reversible denaturation.

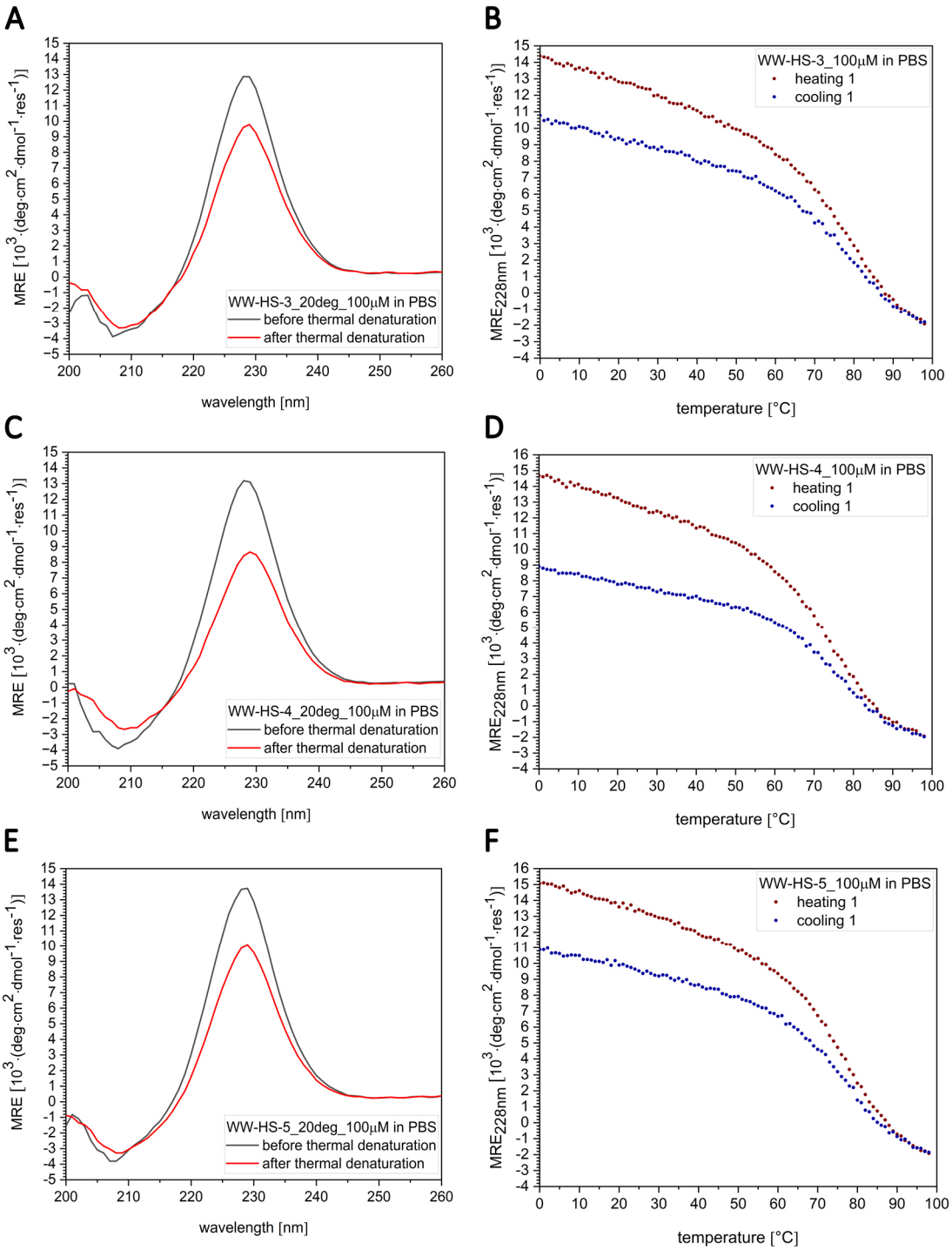


Figure 25: CD spectra before and after thermal denaturation and heating and cooling curve profiles of all three redesigned WW-HS variants. CD spectra for **A**) WW-HS-3, **C**) WW-HS-4, **E**) WW-HS-5 were recorded in duplicates at 20 °C before and after thermal denaturation experiments. Thermal denaturation curves for **B**) WW-HS-3, **D**) WW-HS-4, **F**) WW-HS-5 were determined at a fixed wavelength of 228 nm from 0-98 °C. For thermal denaturation curves the mean out of duplicates is shown. For each variant one heating and cooling curve measurement was

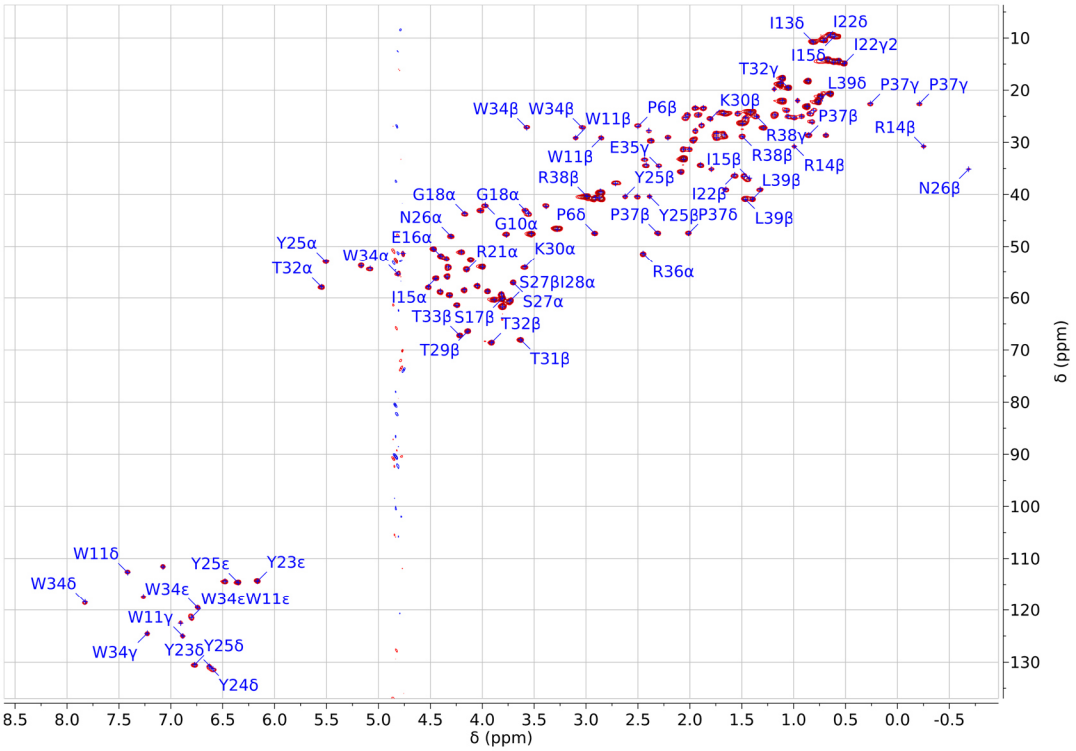
performed. Conditions: 100 µM peptide, PBS buffer, pH 7.5. Parts of the figure were adapted with permission from Lindner *et al*.[94], ©2024 American Chemical Society.

In summary, all three variants show decreased thermostability in comparison to WW-HS-1, WW-HS-2, and WW-HS-2_{OH}, with reduced melting temperatures of more than 7 K. Furthermore, the redesigned variants do not display reversible folding-to-unfolding behavior. Consequently, the variants WW-HS-3, WW-HS-4, and WW-HS-5 were excluded from further consideration for the selection of the most suitable thermostable core scaffold.

4.1.4.4 NMR structure analysis

Subsequently, one of the three most thermostable designed WW-HS peptide variants, WW-HS-2_{OH}, was studied by nuclear magnetic resonance (NMR) spectroscopy to determine the three-dimensional structure on molecular resolution (cf. Figure 26 A and B).

A



B

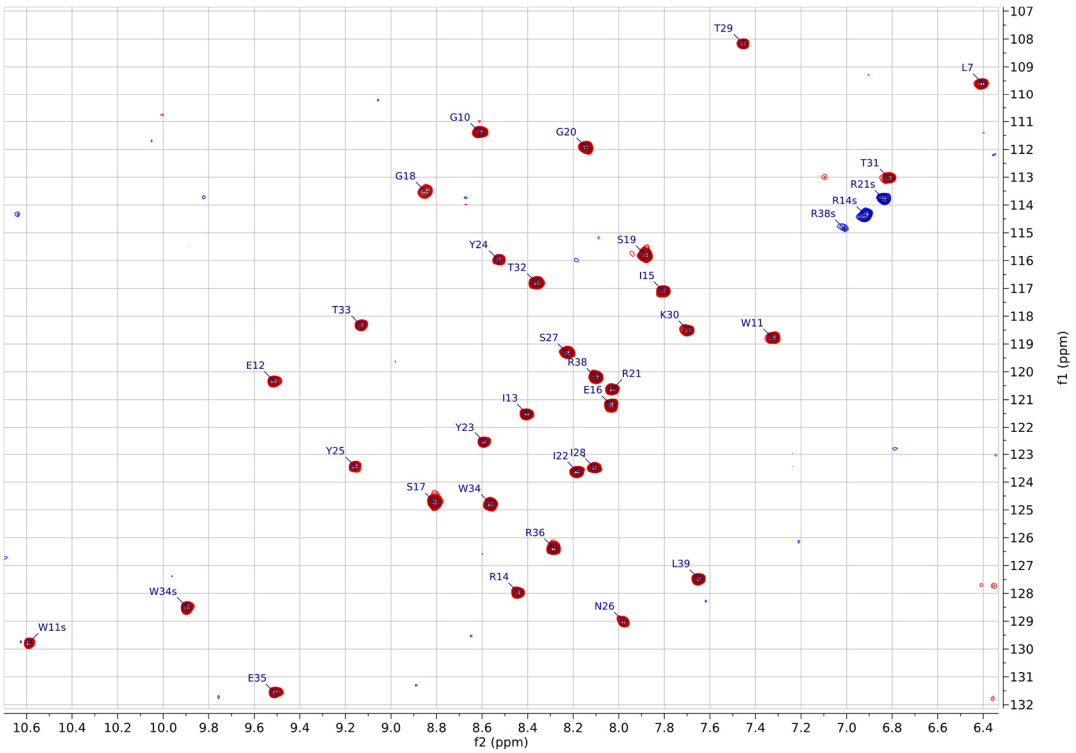


Figure 26: Two-dimensional heteronuclear NMR spectra of WW-HS-2$_{OH}$. A) Two-dimensional ^{1}H-^{13}C HSQC NMR spectrum with the allocated cross peaks obtained from combined 2D ^{1}H-^{1}H NOESY, 2D ^{1}H-^{1}H TOCSY and 2D ^{1}H-^{13}C TOCSY-HSQC NMR spectra. CH, CH$_2$ and CH$_3$ correlations are shown followed by the respective residue at the sequence position, starting with number 6 for the first residue, and the carbon atom of the side chain. **B**) Two-dimensional ^{1}H-^{15}N HSQC NMR spectrum with allocated cross peaks obtained from combined 2D ^{1}H-^{1}H NOESY, 2D ^{1}H-^{1}H TOCSY and 2D ^{1}H-^{13}C TOCSY-HSQC NMR spectra. Amide protons at their respective sequence position

are shown. For some cross peaks “s” indicates the respective side chain. Positive signal heights are displayed in red, negative in blue. Both spectra were aquired at T = 293 K and B_0 = 18.8 T by Prof. Kovermann and coworkers. Both figures are reprinted with permission from Lindner *et al*.[94], ©2024 American Chemical Society.

The data obtained from the multi-dimensional NMR experiments were then integrated to generate a three-dimensional structure, comprising ten conformers with the lowest overall energy. Upon analysis of the structural bundle of the ten conformers of WW-HS-2_{OH}, the WW domain‘s typical secondary structure, consisting of three β-sheets, connected via two loop regions, was identified (cf. Figure 27 A). However, the orientation of the β-sheets towards each other is different if compared to other WW domains, particularly in the case of the amino acids located in the C-terminal β-sheet positions. Furthermore, some of the conformers show distinct positions of the N-and C-terminal regions.

When analyzing the structural bundle of the ten conformers of WW-HS-2_{OH}, it was observed that the orientation of Glu11 is well conserved. The energetically most favored position of the amino acid Arg16 differs slightly when comparing the ten conformers of WW-HS-2_{OH}. Furthermore, an intersegment electrostatic interaction between Arg16 of LP1 and Glu30 of the C-terminus was identified (cf. Figure 27 B). This long-range interaction leads to a strong structure stabilization and thus a high thermostability of WW-HS-2_{OH}.

In the CD spectra measurements of WW-HS-2_{OH} (cf. Figure 21 A), a high exciton signal was identified, which can be explained by the interaction of Trp6 and Tyr19. Both aromatic residues are in close proximity to each other and their side chain orientation is well conserved within the ten conformers (cf. Figure 27 C).

Another intersegment structure-stabilizing interaction is found between Trp29 of the C-terminus and the flexible LP1 region, respectively Ser12. The side chain of the tryptophan at position 29 is highly conserved among the conformers and is situated in close proximity to the LP1 region (cf. Figure 27 D), thus facilitating the formation of polar interactions between both segments.

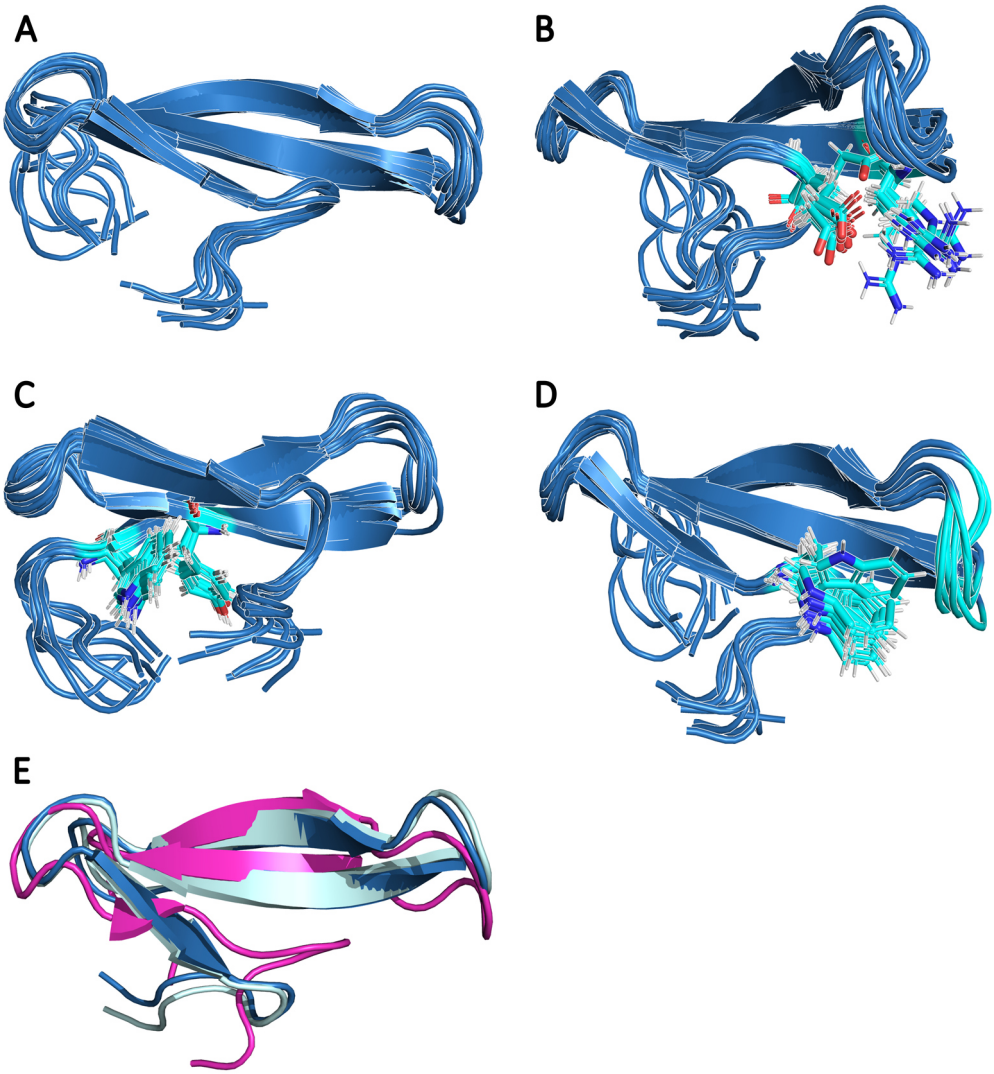


Figure 27: Three-dimensional structure of WW-HS-2$_{OH}$ and structural alignment. A) Structural alignment of the ten different NMR conformers (blue). **B**) Intersegment salt bridge between Arg16 and Glu30 (light blue sticks). **C**) Aromatic residues Trp6 and Tyr19 (blue sticks). **D**) Proximity of Trp29 (light blue sticks) and LP1 residues (light blue). **E**) Aligned NMR structure (pink, pdb code: 8qtq)[94] with the predicted structures of *AlphaFold2*[100] (blue) and *OmegaFold*[113] (light blue). Some of the presented figures have already been published in modified form in Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Upon comparison of the experimentally determined structure of WW-HS-2$_{OH}$ with the predicted structures obtained from *AlphaFold2* and *OmegaFold*, some minor differences can be identified. One noteable discrepancy is the distinct representation of the BS3 segment, which is displayed as a short and twisted β-sheet (cf. Figure 27 E). Additionally, the length of the BS2 region is shorter in the experimentally determined structure than in the predicted ones. Furthermore, the orientation of the BS2 and BS3 region is depicted as coplanar in the structure obtained from NMR measurements, in contrast to the twisted representation of the predicted β-sheet segments. Nevertheless, the two loop regions are displayed in a comparable orientation and length in all three structures.

In summary, the NMR structure of the highly thermostable variant WW-HS-2$_{OH}$ constituted a pivotal step in the characterization of the WW domain-based scaffold. This was evidenced by the identification of several highly structure-stabilizing interactions that collectively

contribute to the exceptional thermostability observed. These experimentally determined electrostatic interactions are incorporated into the previously presented sequence-structure-stability relationships, which are displayed in Figure 28 below.

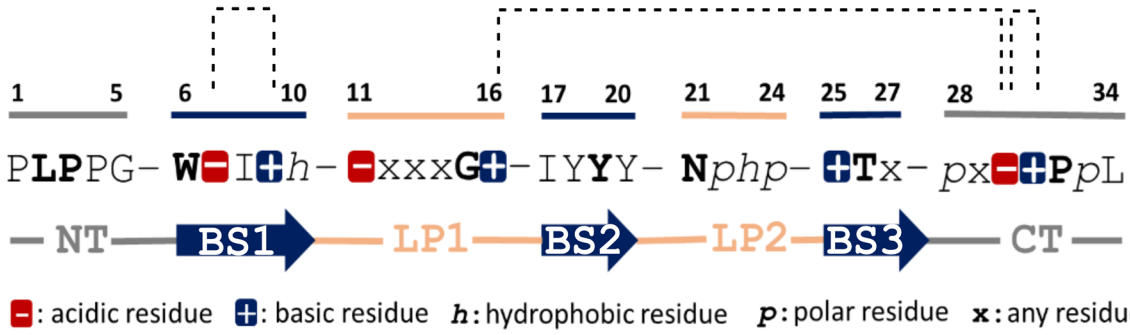


Figure 28: Sequence-structure relationships in WW domains identified by an iterative sequence analysis. The amino acid sequence to obtain thermostable peptide variants is displayed with highly conserved residues in bold. Sequence positions of the different segments are indicated with numbers. Positively charged residues are represented with a "+" and negatively charged residues with a "-" symbol. Dashed lines indicate electrostatic interactions that were found in two-dimensional NMR measurements. Parts of the figure were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

The final combination of the identified sequence-structure-stability relationships with intra- and intersegment interactions, which were gained by structural characterization on the atomic level, is essential for further peptide design when functional amino acids are introduced in the thermostable basic scaffold.

4.2 Application of the hyperthermostable scaffold

4.2.1 On WW domain group-specific motifs

WW domains are protein-protein interaction modules that bind to short proline-rich sequences.[69,72] In general, they are divided into four different groups, depending on the ligand binding motif recognized (cf. Table 9). Although, all ligand motifs consist of different amino acid sequences, they all share the folding of a typical PPII-helix. The four traditionally classified groups of WW domains are: Group I, which binds to proline-rich motifs of the type PP*x*Y (where *x* is a random amino acid), Group II/III, which are often considered as one because of their similar binding pattern, bind to PPLP (II)/PPRP (III) motifs, and Group IV WW domains recognize phosphorylated ligand sequence motifs of the type pSP/pTP.[65,67,69,72,75,114]

Table 9: Ligand sequence motifs of the three WW domain groups. Ligand recognition motifs indicated in the one-letter amino acid code for each WW domain group with *x* as a random amino acid.[65,67,69,72,75,114]

WW domain Group	Ligand recognition motif
WW-I	PP*x*Y
WW-II/III	PPLP (II)/PPRP (III)
WW-IV	pSP/pTP

These previously published findings on group-specific ligand recognition of WW domains have been taken into account for the generation of group-specific thermostable WW domains. For this purpose, WW-HS-2 was used as a general platform to graft the respective ligand-binding relevant amino acids ontop, while maintaining thermostability.

4.2.1.1 Design principle

To successfully graft ligand binding sites onto the thermostable scaffold, group-specific binding patterns have to be identified. The analysis of binding-relevant amino acids for the presented four WW domain groups was performed by computational alanine scanning on published NMR spectroscopic and X-ray crystal structure data and is summarized in Table 10.[115] In the widely used alanine scanning approach, binding-relevant amino acids at the respective positions, the so-called "hot spot" residues at the interaction interface, are identified based on the binding energy that differs when alanine mutations are introduced.[116] The results of the *Robetta* alanine scanning revealed a high conservation of His22, Thr27 and Trp29 for Group I of the WW domain, as well as the importance of Glu7 for weak interaction. In Group II/III of WW domains, Tyr residues within the BS2 segment are highly conserved. In Group IV WW domains, Arg residues in the BS1 and LP1 regions appear to play an important role in the recognition of phosphorylated ligands.

Table 10: Overview of the group-specific WW domains used for the analysis of ligand-binding relevant positions and residues. For residue identification *Robetta* alanine scanning was performed.[115] Amino acids displayed in blue indicate weak interactions, residues in red are essential for ligand binding. PDB files have been accessed with RCSB PDB.[33] Parts of the table were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

Group	WW domain	Sequence	PDB identifier
I	*h*YAP-WW1	PLPAGWEMAKT-SSGQRYFLNHIDQTTTWQDPRK	2ltw
	*h*YAP-WW2	PLPDGWEQAMT-QDGEIYYINHKNKTTSWLDPRL	2law
	*m*Itch-WW3	PLPPGWEKRTD-SNGRVYFVNHNTRITQWEDPRS	2jo9
	*m*Nedd4-2-WW3	FLPPGWEMRIA-PNGRPFFIDHNTKTTTWEDPRL	2mpt
	*h*Smurf1-WW2	PLPPGWEVRST-VSGRIYFVDHNNRTTQFTDPRL	2lb1
	*h*Nedd4-1-WW3	FLPKGWEVRHA-PNGRPFFIDHNTKTTTWEDPRL	2kq0
	*h*Smurf2-WW3	PLPPGWEIRNT-ATGRVYFVDHNNRTTQFTDPRL	2djy
	*d*Su(dx)-WW4	PLPPGWEIRYT-AAGERFFVDHNTRRTTFEDPRP	2jmf
II/III	*h*WWFE65	DLPAGWMRVQD-TSG-TYYWHIPTGTTQWEPPGR	2oei
	*h*HypA/FBP11-WW1	GAKSMWTEHKS-PDGRTYYYNTETKQSTWEKPDD	2dyf
	*m*FBP28-WW2	TAVSEWTEYKT-ADGKTYYYNNRTLESTWEKPQE	2rly
IV	*h*Pin1-WW	KLPPGWEKRMSRSSGRVYYFNHITNASQWERPSG	1i8g
	*h*Smurf1-WW1	ELPEGYEQRTT-VQGQVYFLHTQTGVSTWHDPRI	2laz
	*m*Nedd4-2-WW2	GLPSGWEERKD-AKGRTYYVNHNNRTTTWTRPIM	2lb2

One representative of each WW domain groups was illustrated in *PyMol* and, in accordance with the published literature, important group-specific amino acids that play a role in ligand-binding of Pro-rich sequence motifs were identified. For the WW domain Group I representative *h*YAP-WW1, the three amino acids Leu20 in BS2, His22 in LP2 and Gln27 in the BS3 segment were identified to bind ligands with a PPxY sequence motif. As shown in Figure 29 A, the Tyr residue of the Group I ligand is bound in the cavity, formed by the three group-

specific amino acids (green). The binding site for the polyproline-based ligand sequence of group I (orange) is located between Tyr18 of BS2 and the C-terminal Trp29 (blue).[6] In *h*WWFE65, a typical Group II/III WW domain peptide, a tryptophan residue at position 20 in LP2 and a threonine residue at position 27 in BS3 interact with the Group II/III l ligand backbone. The Group II/III WW domain ligand with a sequence motif of `PPLP` forms interactions with the common proline binding site Tyr18 and Trp29 in WW domains (cf. Figure 29 B).[69] A prominent representative of Group IV WW domains is *h*Pin1, which recognizes phosphorylated threonine residues in group-specific ligands with a sequence motif of `pSP/pTP`. In particular, the two positively charged Arg residues at positions 12 and 16 within the LP1 segment bind to the negatively charged phosphate group of the phosphorylated Thr of the ligand (cf. Figure 29 C). For Group IV WW domains, the proline binding site of the proline-rich ligand is located between Tyr18 and Trp29.[70]

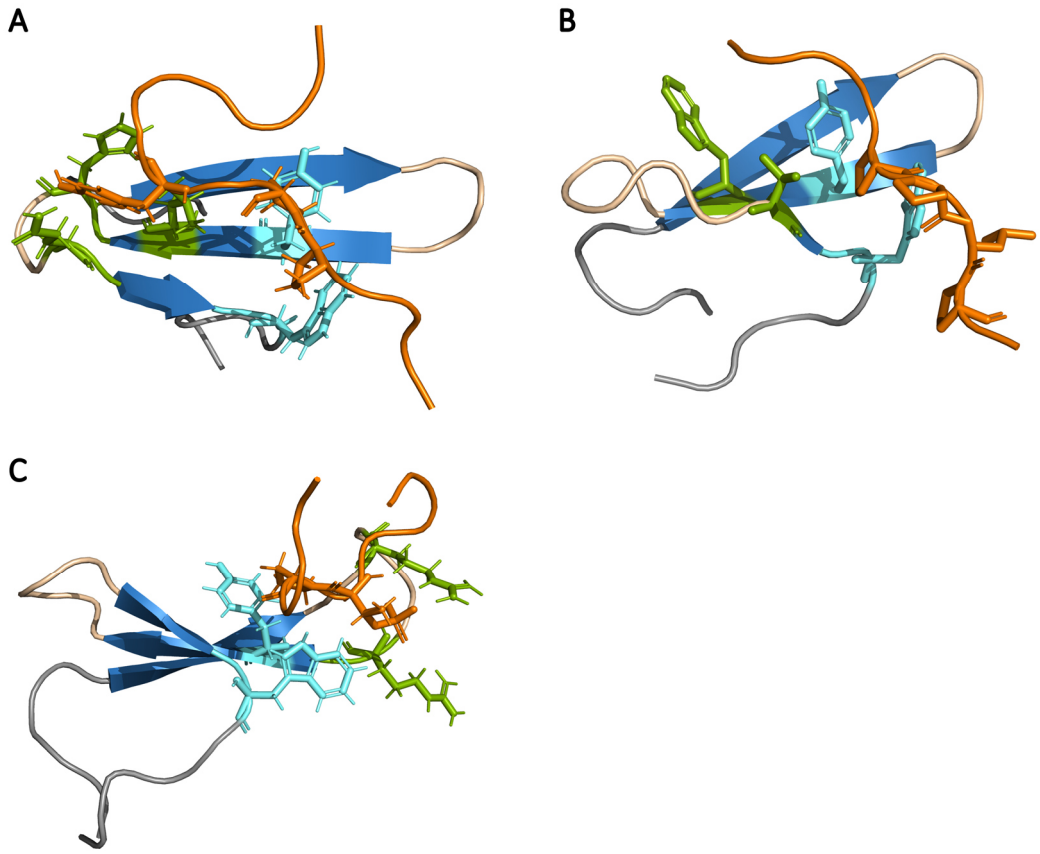


Figure 29: Specific sequence motifs of all three WW domain groups. BS segements are displayed in blue, LP regions beige, peptide termini grey, ligands orange, group-specific residues as green sticks, and residues of the poly-proline binding site as light blue sticks **A**) Structure of WW domain Group I *h*YAP1$_{WW}$ (pdb code: 2ltw).[117] **B**) Structure of WW domain Group II/III peptide *h*WWFE65 (pdb code: 2oei).[118] **C**) Structure of the WW domain Group IV peptide *h*Pin1$_{WW}$ (pdb code: 1i8g).[92] PDB files have been accessed with RCSB PDB.[33]

For the design of thermostable, group-specific WW domain-based peptides with native ligand-binding functions, the results of the *Robetta* alanine scanning (cf. Table 10) were combined with the previously published data on ligand-binding relevant amino acids of the three WW domain groups.[65–67,75]

Although some amino acids within a WW domain group are highly conserved at certain sequence positions, such as the glutamic acid at position 7 within Group I, some amino acid

variations have also been identified at binding-relevant positions. Consequently, the most suitable residue was selected for these positions, i.e. the amino acid has to be highly abundant at the respective sequence position within the WW domain group and should not interfere with the structure-stabilizing interactions intentionally introduced into the highly thermostable scaffold of WW-HS-2. This engineered peptide serves as a platform to graft group-specific amino acids onto that are also relevant for ligand binding. Besides the rationale of successfully transferring binding-relevant amino acids onto the WW-HS-2 peptide, the goal was to maintain the high thermostability by retaining as many amino acids as possible and only exchanging group-specific amino acids that appear to be relevant for gaining group-specific ligand-binding properties.

Thus, to design a WW domain-based Group I, WW-HS-I, a Gln residue was introduced at position 13 within LP1, an Ile residue at position 20, and a His residue at position 22 of LP2. Due to the fact that in Group I WW domain peptides glutamine residues are found in the LP1 region, it was decided to also introduce this amino acid in a WW-HS-I peptide variant. The Ile at position 20 was also identified in some native Group I representatives and was preferred over a leucine or valine, since Ile residues have a high β-sheet propensity[103] and consequently could lead to structural stabilization of the BS2 segment. Since the histidine at position 22 of the LP2 region is highly conserved throughout the WW domain Group I[6], it was selected for the WW-HS-I design (cf. Table 11).

To design a highly thermostable WW domain Group II/III peptide, WW-HS-II/III, several mutations were introduced into the WW-HS-2 scaffold. First, the highly abundant Thr7 and Tyr9, both located within the BS1 region, were adopted. For a high number of representative Group II/III peptides, the LP1 region is shorter compared to other WW domain groups. Therefore, it was decided to use a shorter LP1 in WW-HS-II/III. To preserve the possibility of a highly structure-stabilizing electrostatic interaction within LP1, the glutamate residue was retained at position 11. The Ser12 was removed to obtain a shortened loop of only five amino acids instead of the regular amount of six residues. Regarding the remaining sequence positions of LP1, a sequence pattern with DG was frequently found. Sequences containing an aspartate residue followed by a glycine are known to be highly prone to aspartimide formation during SPPS.[111,119] To avoid the unwanted side reaction and maintain a straightforward peptide synthesis process, the aspartate residue was replaced by an asparagine, resulting in the sequence pattern ANG, which was previously reported in a design approach of metal-binding WW-CA variants by Pham *et al.* in 2023.[120] In BS2, the highly abundant Thr17 was introduced to obtain WW-HS-II/III (cf. Table 11). In addition, a glutamate residue, which was found to be abundant in Group II/III WW domains, was exchanged with the threonine at position 26 of WW-HS-2.

The highly thermostable variant with WW domain Group IV binding properties, WW-HS-IV (cf. Table 11), was designed by grafting the SRS sequence pattern of LP1 with the phosphate binding-relevant Arg 12, identified in *h*Pin1$_{WW}$[70], onto the WW-HS-2 scaffold. In addition, a Ser was introduced at position 27 of the BS3 segment, as this residue was also found to be essential for ligand-binding in Group IV peptides. In the C-terminal region, the Thr28 was exchanged by a Gln residue, which was also found to be part of the ligand binding interaction in *h*Pin1.

Table 11: Amino acid sequences of four group-specific designed WW-HS variants based on WW-HS-2. Residues displayed in blue indicate BS regions, and residues in red were identified as group-specific sequence patterns. Parts of the table were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

WW domain	Sequence ------NT-----BS1------LP1----BS2--LP2--BS3---CT-- 1 - 10 - 20 - - 30
WW-HS-2	Ac-PLPPG-WEIRI-ESGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2
WW-HS-I	Ac-PLPPG-**WEIRI**-**E**S**Q**SG**R**-**IYYI**-N**H**IT-**KT**-T**W**ERPRL-NH_2
WW-HS-II/III	Ac-PLPPG-**WTIYI**-**E**-ANGR-**TYYY**-NSIT-**KET**-**TW**ERPRL-NH_2
WW-HS-IV	Ac-PLPPG-**WEIRI**-**SRSS**GR-**IYYY**-NSIT-**KTS**-Q**W**ERPRL-NH_2

In addition to the designed group-specific WW domain-based peptides presented above, a native WW domain peptide was selected for each group as a reference for the subsequent studies (cf. Table 12). In order to facilitate peptide synthesis as far as possible, native WW domains were chosen based on their amino acid sequence, which either completely lack Asp residues or did not contain Asp-Gly sequence patterns, which are highly prone to aspartimide formation.[119]

Table 12: Selected native WW domain group-specific reference peptides. Sequences were extracted from indicated pdb entries.[92,121] PDB files have been accessed with RCSB PDB.[33] Residues depicted in blue display BS regions. Parts of the table were adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

WW domain	Sequence 1 10 20 30	PDB Identifier
*y*YAP1$_{WW2}$	Ac-GPLPEG**WEQAI**T-PEGE**IYYI**NHKN**KTT**SWLDPRL-NH_2	2l4j
*h*GAS7$_{WW}$	Ac-VILPPG**WQSYL**S-PQGR**RYYV**NTTT**NET**TWERPSS-NH_2	2ysh
*h*Pin1$_{WW}$	Ac-KLPPG-**WEKRM**SRSSGR**VYYF**NHIT**NAS**QWERPSG-NH_2	1i8g

4.2.1.2 Structure and thermal stability

After the successful synthesis of the designed and native peptides of the three WW domain groups, the structures and thermostability were investigated by CD spectroscopy to determine whether the introduction of binding-relevant amino acids onto the thermostable WW-HS-2 scaffold would still yield a thermostable variant.

Starting with the secondary structure analysis of the WW domain Group I native (*y*YAP1$_{WW2}$) and designed peptide (WW-HS-I), the characteristic WW domain structure with a maximum signal at a wavelength of about 228 nm was identified in both variants (cf. Figure 30 A). However, the exciton signal was more pronounced in the designed WW-HS-I peptide, indicating that the two aromatic amino acids, forming the exciton coupling might be closer in proximity in the designed group-specific peptide compared to *y*YAP1$_{WW2}$. Thermal denaturation profiles revealed a distinct curve pattern for both peptides (cf. Figure 30 B). The native peptide exhibited a more sigmoidal curve, whereas the designed WW-HS-I variant showed a broad transition and a rather flat curve profile. The native *y*YAP1$_{WW2}$ peptide

revealed a moderate melting temperature of 66.5 °C, while the designed variant WW-HS-I, showed high thermostability with a T_m of 75.5 °C.

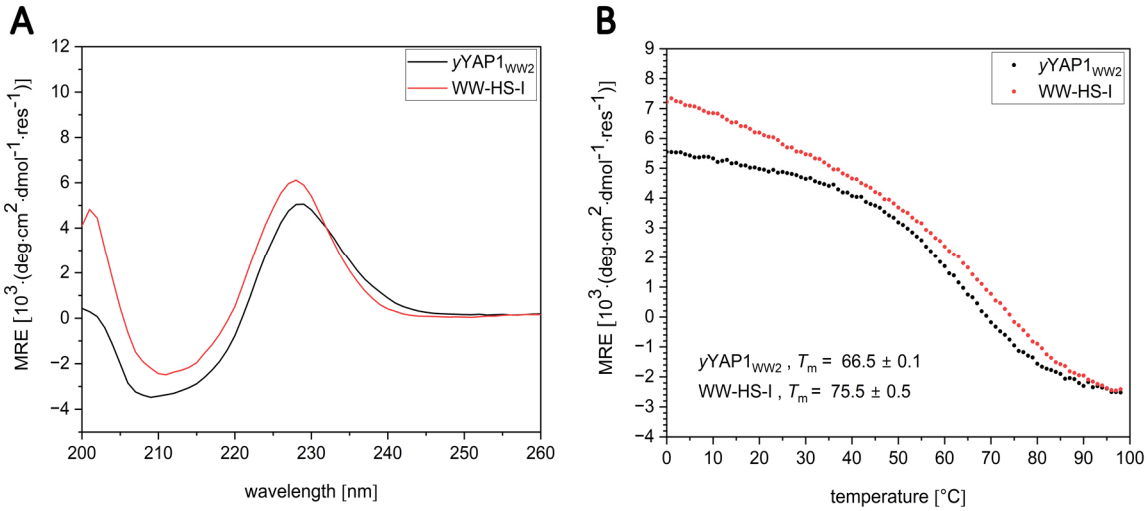


Figure 30: Structure and thermal denaturation of the native and designed WW domain Group I peptides. A) Secondary structures of *y*YAP1$_{WW2}$ and WW-HS-I, measured in triplicate at 20 °C, are shown as mean values. **B**) Thermal denaturation curves, using mean values of triplicate measurements at a fixed wavelength of 228 nm from 0 -98 °C, are displayed. Conditions: 100 µM peptide, PBS buffer, pH 7.5. This figure was adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

The native and designed peptides of Group II/III were also investigated for structural and thermal stability. The results of CD spectroscopy revealed the characteristic WW domain maximum at approximately 228 nm for both variants (cf. Figure 31 A), but a higher signal intensity was identified for the designed WW-HS-II/III peptide. Thermal denaturation of both Group II/III peptides showed sigmoidal curve shapes for both, but with a higher initial signal and longer transition from the folded to the unfolded state for WW-HS-II/III (cf. Figure 31 B), resulting also in a higher melting temperature of 78.0 °C. In contrast, the native *h*Gas7$_{WW}$ variant revealed a melting temperature of 64.0 °C, indicating a less thermostable peptide if compared to the designed variant.

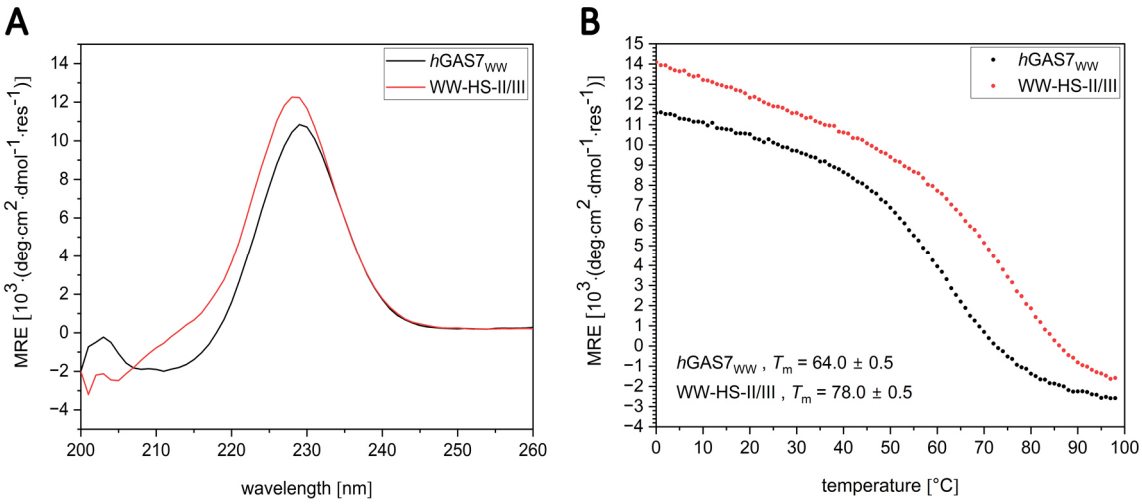


Figure 31: Structure and thermal denaturation of the native and designed WW domain Group II/III peptides. A) Secondary structures of *h*Gas7$_{WW}$ and WW-HS-II/III, measured in triplicate at 20 °C, are shown as mean values. **B**) Thermal denaturation curves, using mean values of triplicate measurements at a fixed wavelength of 228 nm from 0 -98 °C, are displayed. Conditions: 100 µM peptide, PBS buffer, pH 7.5. This figure was adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

The recorded CD spectra and thermal denaturation profiles were also acquired and analyzed for the WW domain Group IV native (*h*Pin1$_{WW}$) and designed peptide (WW-HS-IV), as displayed in Figure 32 A and B. The elucidation of the secondary structure and CD signal intensities of *h*Pin1$_{WW}$ and WW-HS-IV, did not reveal any significant differences. However, a small, notable shift of the maximum by approximately 1 nm was observed for WW-HS-IV. Although, the initial CD signals of both thermal denaturation curves were in a similar range, the curve profiles varied at temperatures above 35 °C (cf. Figure 32 B), indicating different denaturation behavior and thermostabilities. The melting curve of *h*Pin1$_{WW}$ shows a sigmoidal shape with an upper and lower plateau, resulting in a melting temperature of 55 °C. WW-HS-IV reveals a slower decrease in signal intensity with increasing temperature, with a lower plateau starting at a temperature of 90 °C. This rather flat signal decrease of the denaturation curve of WW-HS-IV, can be explained by the high thermostability of the designed peptide with a calculated melting temperature of 77.5 °C.

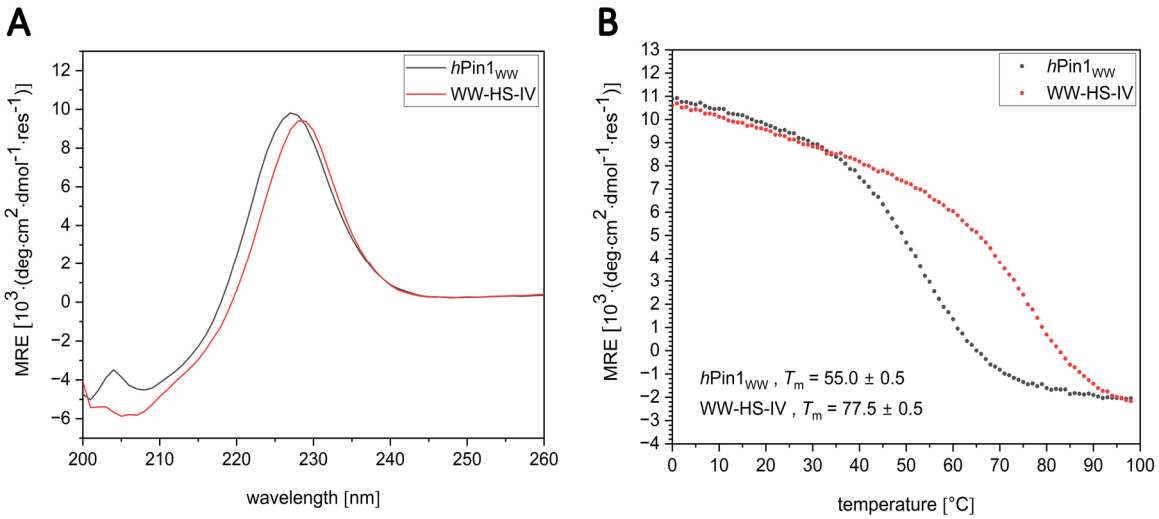


Figure 32: Structure and thermal denaturation of the native and designed WW domain Group IV peptides. A) Secondary structures of *h*Pin1$_{WW}$ and WW-HS-IV, measured in triplicate at 20 °C, are shown as mean values. **B**) Thermal denaturation curves, using mean values of triplicate measurements at a fixed wavelength of 228 nm from 0 -98 °C, are displayed. Conditions: 100 µM peptide, PBS buffer, pH 7.5. This figure was adapted with permission from Lindner *et al.*[94], ©2024 American Chemical Society.

In general, the goal of obtaining group-specific highly thermostable peptides by grafting binding sites of the respective WW domain group onto the WW-HS-2 scaffold, was successfuly achieved. This was demonstrated by CD denaturation experiments, which showed a significant increase in melting temperatures for all designed group-specific peptides compared to native representatives. Nevertheless, the question remains whether transferring the group-specific amino acids to the thermostable scaffold will also result in functional peptide variants with respect to ligand binding. This will be addressed in the following chapter.

4.2.1.3 Studies on binding affinity

Following the successful grafting of ligand-binding specific amino acids onto the WW-HS-2 scaffold and the generation of moderately to highly thermostable group-specific peptide variants, the binding affinities of the native and designed peptides were examined and compared. The primary objective was to retain thermostable peptides that are additionally capable of binding to their respective ligands (cf. Table 13), thereby combining thermostability with function. However, the applied design approach did not aim at optimized binding.

Table 13: Amino acid sequences of WW domain group-specific ligands and the respective K_d values. Ligand-specific sequence patterns are indicated in red. This table has been previously published in minorly modified form in Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

WW domain group peptide	Ligand sequence	K_d [µM]
WW-HS-I *y*YAP1$_{WW2}$	Ac-GTP**PPPY**TVG-NH_2	1.1 ± 0.2 108 ± 56
WW-HS-II/III *h*GAS7$_{WW}$	Ac-GSTA**PPLP**R--NH_2	3.8 ± 0.4 0.9 ± 0.4
WW-HS-IV *h*Pin1$_{WW}$	Ac-YSPT-**pS**-**P**S-NH_2	2.5± 0.9 1.2 ± 0.2

The respective binding affinities of the designed group-specific WW domains and their native peptides towards their ligands were investigated by measuring the changes in intrinsic tryptophan fluorescence upon titration of the ligand into the phosphate-buffered peptide solution. The respective dissocation constant of each peptide was determined based on the individual saturation binding curve (detailed description c.f. Chapter 6 Materials and Methods) and subsequently the K_d values of the designed variants were compared with their natural reference peptides.

All three group-specific designed thermostable peptides revealed binding to their respective ligands (cf. Figure 33 A) with dissociation constants in the low micromolar range (cf. Table 13). Compared to the native WW domain-based reference peptides, the designed variants displayed equal binding affinities to their respective ligands, except for the WW domain Group I. The native *y*YAP1$_{WW2}$ showed a comparatively weak binding to ligand WW-I with a K_d value of 108 µM, whereas the designed peptide displayed a binding affinity that was two orders of magnitude stronger than the native reference. However, this moderate ligand binding of *y*YAP1$_{WW2}$ was also reported in literature by Webb *et al.* in 2011.[122]

The other three reference peptides exhibited similar binding to their respective ligand as the designed group-specific HS peptides (cf. Table 13).

After an extensive literature search, no binding affinity values of *h*GAS7$_{WW}$ for ligands with a sequence pattern of Group II/III WW domains have been published so far, so that a comparison of the native peptide reference with the literature was not possible.

For the native WW domain Group IV reference peptide *h*Pin1$_{WW}$, a K_d value of 34 µM was reported in binding studies with an N-terminally fluorescently labeled Group IV ligand.[123] Verdecia *et al.* used fluorescence anisotropy to determine the binding affinity of *h*Pin1$_{WW}$ to its characteristic ligand, which appears to show a weaker peptide-ligand interaction by an order of magnitude[123] compared to the K_d of 1.2 µM identified in this work (cf. Table 13). The

use of two different methods may explain the reported discrepancy in K_d values for the native Group IV WW domain peptide.

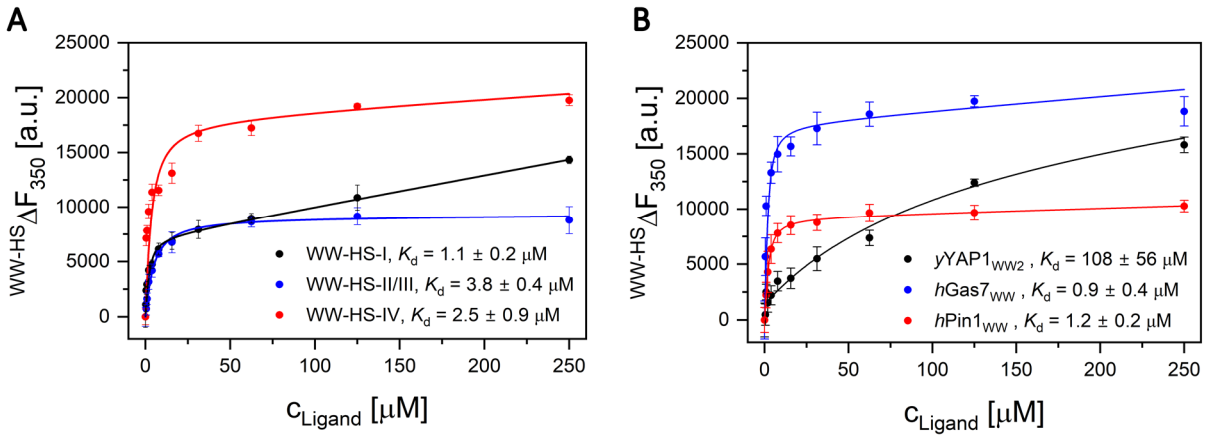


Figure 33: Saturation binding curves of native and designed group-specific WW domain-based peptides to their group-specific ligands. Binding curves obtained from **A**) designed group-specific WW-HS peptides (2 μM), and **B**) native reference WW domain peptides (2 μM), obtained by monitoring the change in intrinsic tryptophan fluorescence upon ligand titration (0-250 μM). Measurements were performed at room temperature in PBS (pH 7.5). K_d values were obtained by non-linear least squares fitting of the respective saturation binding curves. The data have already been published in a minorly modified form in Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

In conclusion, the approach of grafting group-specific ligand binding sites onto the highly thermostable WW-HS-2 core scaffold, resulted in thermostable to highly thermostable WW domains that additionally exhibited ligand-binding properties with similar or even stronger binding affinity (WW-HS-I) compared to native reference peptides. The successful combination of thermostability and function, validated the inital hypothesis that once a highly thermostable WW domain scaffold has been designed, it can be further functionalized by grafting specific binding motifs onto it, while maintaining its thermostability. Moreover, the robustness of the WW-HS-2 core scaffold to mutations introduced to allow binding to corresponding ligands was confirmed.

4.2.2 On a metal-binding minireceptor

Following the successful grafting of native binding motifs onto the designed thermostable core scaffold WW-HS-2, the objective was to expand the scope of functionalities to include metal binding. This was achieved by grafting the His_3 site of a Zn(II) binding WW domain peptide, WW-CA, from Pham *et al.*[84] onto the highly thermostable WW-HS-1, thus obtaining a thermostable Zn(II) binding WW domain. To investigate the secondary structure of this redesigned WW-HS-CA peptide, CD spectra were recorded in the presence and absence of Zn(II). Subsequently, the binding affinity of WW-HS-CA was examined by conducting a competitive titration assay based on UV/Vis spectroscopy.

4.2.2.1 Design principle

Based on the amino acid sequence of *h*Pin1$_{WW}$, Pham *et al*. employed a top-down design approach to engineer a peptide with a His_3 site, which is originally found in the active site of the enzyme carbonic anhydrase II (CA).[84] The binding of Zn (II) to the His_3 site is stabilized by coordinative bonds that further form hydrogen bonds between two His and two Gln residues

of both loops and an intra-β-strand interaction of a histidine and glutamate.[91,93] This design approach resulted in a peptide variant with Zn(II)-binding properties in the low micromolar range. Moreover, it was demonstrated that upon the binding of Zn(II), the characteristic WW domain structure with the exciton coupling at approximately 228 nm was stabilized, as well as the overall improved thermostability of WW-CA in thermal denaturation curves. Although the thermostability of WW-CA was increased in its holo state, the melting temperature was depicted as rather low, with a T_m of 34.3 °C.[84]
Consequently, the rationale was to further stabilize the WW-CA peptide by grafting the His_3 site on the highly thermostable core scaffold WW-HS-1, resulting in the peptide WW-HS-CA (cf. Table 14), which was further characterized with respect to its structural and thermal stability, as well as its ability to bind Zn(II), as described in the following sections.

Table 14: Sequence of the identified Zn(II) binding peptide WW-CA from Pham *et al.* 2022[84] and the redesigned WW-HS-CA. Residues depicted in red represent the His_3 site in WW-CA and WW-HS-CA. Residues displayed in blue indicate BS regions. The sequence of WW-HS-CA has already been published in Lindner *et al.*[94]

WW domain	Sequence
	-------NT------BS1------LP1-----BS------LP2--BS3----CT--
	1 - 10 - 20 - - 30
WW-CA	H-KLPPG-**WEKHM**-SRSSGR-**VHYH**-NSIT-**QAS**-QWERPSG-OH
WW-HS-CA	H-PLPPG-**WEIHI**-DSGSGR-**IHYH**-NSIT-**KTT**-TWERPRL-OH

4.2.2.2 Characterization of structure and thermal stability

The impact of sequence modifications, was investigated with respect to the structure and thermostability of the resulting peptide, WW-HS-CA. Structure analysis and thermal denaturation measurements were performed both in the presence and absence of Zn(II).

The CD spectra analysis revealed a spectrum that is characteristic of a WW domain (cf. Figure 34 A). A comparison of the CD spectra of WW-HS-CA in the presence and absence of Zn(II) revealed a more pronounced exciton coupling at 228 nm in the apo state than in the sample with Zn(II) (1:1 ratio). This indicates that the structure of WW-HS-CA in the apo state is already stable. In contrast, the published WW-CA from Pham *et al.* exhibited a more pronounced exciton coupling when bound to Zn(II), suggesting that the peptide's fold was stabilized in its holo state.[84]

Thermal denaturation experiments revealed two distinct denaturation profiles for the apo and holo states of WW-HS-CA (cf. Figure 34 B). In accordance with the published results of Pham *et al.*[84], the denaturation profile of WW-HS-CA with Zn(II) displays a broad transition. Furthermore, the thermal stability of WW-HS-CA is increased when Zn(II) is bound in its holo state, with a melting temperature of 75.5 ± 0.5 °C. Although, the melting temperature of WW-HS-CA in the apo state is significantly decreased (T_m = 43.0 ± 0.5 °C) in comparison to the holo state, it still exhibits a good thermostability, which is already higher than that of the previously

published WW-CA-Zn(II) complex of Pham *et al*., which revealed a melting temperature of 34.3 °C.[84]

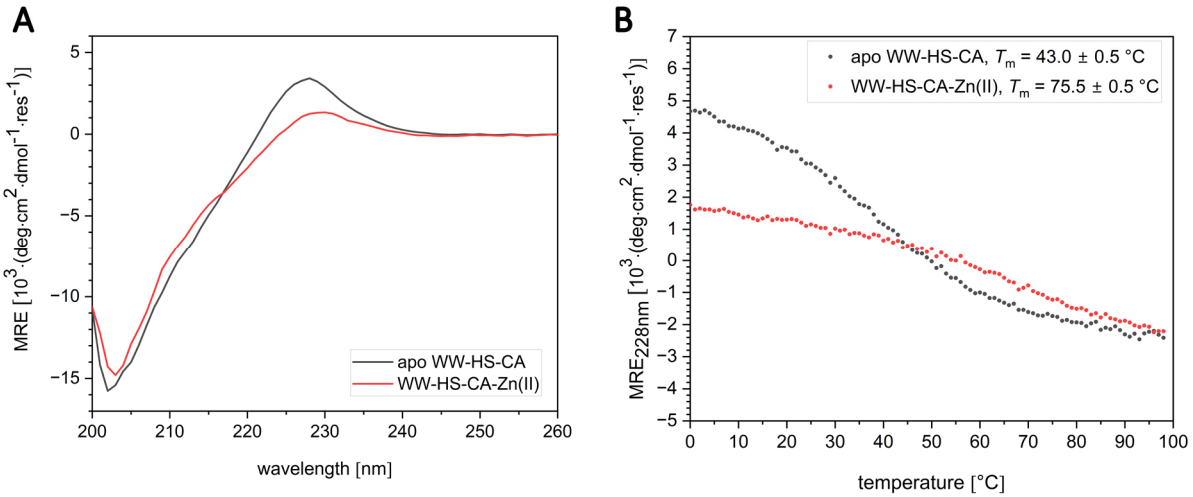


Figure 34: CD spectra and thermal denaturation curves of WW-HS-CA in the apo and holo state. A) CD spectra of apo and holo WW-HS-CA incubated overnight in MOPS buffered-saline (pH 7.2) at 20 °C, the ratio of peptide to $ZnSO_4$ was 1:1 for the holo peptide. The mean of duplicates is shown. **B**) Thermal denaturation curves, monitored at a fixed wavelength of 228 nm from 0 to 98 °C, are shown as mean of duplicates. The experiments were performed by Niklas Schwegler and the data have already been published in modified form in Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

In conclusion, the transfer of a His_3 site onto the WW-HS-1 core scaffold resulted in a thermostable WW-HS-CA peptide with high thermostabiliy in the presence of Zn(II). This result, supports the hypothesis that grafting metal-binding motifs onto the highly thermostable core scaffold is possible and leads to a highly structurally and thermostable peptide.

4.2.2.3 Studies on binding affinity

In addition to the successful *de novo* design of WW-CA, which resulted a highly thermostable WW-HS-CA peptide variant when occurring in its holo state, the binding affinity to Zn(II) was examined with a comparative titration experiment.

The comparative titration was conducted in the presence of the dye MagFura2, which is excitable by UV light and a competitor for Zn(II) binding. The binding of MagFura2 to Zn(II) can be monitored by absorbance measurements.[124] The obtained data demonstrated a high binding affinity of WW-HS-CA to Zn(II), with a K_d value of 0.1 µM (cf. Figure 35). In comparison to WW-CA of Pham *et al*., which was reported to exhibit a K_d value of 1.2 µM[84], this is an increase of one order of magnitude, indicating strong binding. Despite this, the recently redesigned peptide variant WW-CA-ANG of Pham *et al*., which has a shortened LP1 region, displayed a five-fold higher K_d value to Zn(II) than WW-HS-CA.[120]

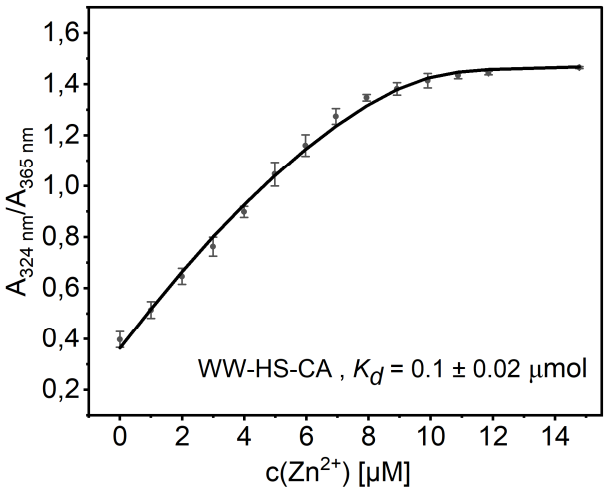


Figure 35: Competitive titration experiment of WW-HS-CA with Zn(II) and MagFura2. The mean of three measurements is presented. Absorbance values were measured at two wavelengths: λ_1 = 324 nm and λ_2 = 365 nm and the K_d value was determined by fitting a single-site model to the data using the *DynaFit* software.[125] The experiment was conducted by Niklas Schwegler and has been previously published in Lindner *et al.* 2024. The figure was adapted from Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

In conclusion, transferring a His_3 site onto WW-HS-1, resulted in the thermostable WW-HS-CA peptide with increased Zn(II) binding affinity in the submicromolar range, which is superior to previously reported peptide designs of Pham *et al*. This successful application of a highly thermostable core scaffold for the grafting of non-native functions, such as metal binding properties onto it, is one example of the broad applicability of the thermostable core scaffold for the design of functional and thermostable WW domain miniproteins.

4.2.3 On organophosphate-binding minireceptors

Functional mini-receptors that exhibit organophosphate binding properties, were identified by a combinatorial screening of split WW domains published by Neitz *et al*. (for a detailed description refer to Chapter 2, State of the Art). Here, *h*Pin1$_{WW}$-based minireceptors were designed that showed selective binding properties to phosphorylcholine (PC) or adenosine triphosphate (ATP).[85] Following the successful grafting of binding sites relevant for peptide and metal binding onto one of the WW-HS core scaffolds, the subsequent objective was to engineer thermostable peptides that exhibit selective binding to organophosphates, such as PC or ATP, while maintaining thermostability.

4.2.3.1 Organophosphate-binding peptides

As previously stated, a selective PC-binding peptide, designated WW-1-8, was identified by a combinatorial approach by Neitz *et al*. This peptide exhibited only low thermostability, with a melting temperature of 14 °C, but demonstrated high binding affinity to PC, with a K_d value of 1.3 ± 0.3 µM.[85] Since the applicability of this peptide as a PC minireceptor is limited by its low thermal stability, the goal was to generate a PC-binding peptide with good structural and thermal stability. Apart from the selective PC-binding WW-1-8 peptide, Neitz *et al*. also identified a selective ATP binding minireceptor, referred to as WW-2-10.[85] Thus, in addition to the transfer of binding motifs relevant for PC binding onto the highly thermostable scaffold

WW-HS-1, another objective was to engineer a highly thermostable peptide with ATP-binding properties.

4.2.3.1.1 Design principle

In order to obtain a PC-binding and an ATP-binding WW domain with sufficient structural and thermal stability, the WW-HS-1 peptide served as a scaffold onto which identified PC-binding or ATP-binding amino acid residues were grafted. While the goal was to increase the thermostability of the peptides by using the highly thermostable WW-HS-1 peptide as a core scaffold, the binding affinities to their organophosphate ligands should be maintained. By performing *in silico* docking studies of PC to WW-1-8 and ATP to WW-2-10 with *AutoDockVina*[126,127], potential binding-relevant sequence patterns of the precursor peptides WW-1-8 and WW-2-10 (cf. Table 15) were identified. These amino acid residues were combined with the WW-HS-1 peptide, resulting in two promising candidates (cf. Table 17, section below), whose secondary structures were modeled with *AlphaFold2*[100]. Their binding affinities to PC or ATP were predicted *in silico* with *AutoDockVina*[126,127], and the most promising peptide candidate in terms of the lowest predicted binding energy was selected for chemical synthesis. Subsequently, the thermal stability was examined by CD spectroscopy and the binding affinity to PC or ATP was determined by intrinsic tryptophan fluorescence spectroscopy.

Table 15: The amino acid sequence of the identified PC-binding peptide WW-1-8 and the ATP-binding variant WW-2-10 published in Neitz *et al*. 2022.[85]

WW domain	Sequence
	--------NT------BS1-----LP1------BS2—LP2—BS3---CT-- 1 - 10 - 20 - - 30
WW-1-8	H-KLPPG-WEKRM-SRSSGR-VYYS-NHIT-NAS-QSERPSG-OH
WW-2-10	H-KLPPG-WEKRM-SRSRGR-VYYH-NHIT-NAE-QSERPSG-OH

4.2.3.2 *In silico* prediction of promising candidates

4.2.3.2.1 Docking studies of WW-1-8 with PC

In silico modeling of WW-1-8 with PC using *AutoDockVina*[126,127] were performed three times and revealed nine different ligand binding positions with varying predicted binding affinities (cf. Table 16). The distinct PC binding positions were visually analyzed in *PyMol* and binding-relevant sequence patterns of the most favorable binding poses were extracted for subsequent WW-1-8 peptide *de novo* design (cf. Table 16). Interestingly, the sequence patterns for positions 2 and 9, as well as for positions 3 and 4 are identical, but when analyzing the amount of polar interactions formed between the binding partners, crucial differences were identified (cf. Figure 36), resulting in varying predicted binding affinities.

Table 16: Sequence patterns of the respective structural elements obtained from different predicted poses of PC with WW-1-8. Results of one representative docking round, including the ligand docking pose and the respective predicted binding affinity with binding-relevant residues depicted in red. *In silico* modeling was performed in triplicates with *AutoDockVina*[126,127] and the results were analyzed in *PyMol*.

Number of the predicted ligand poses	Predicted binding affinity [kcal/mol]	Sequence regions						
		NT (1-5)	BS1 (6-10)	LP1 (11-16)	BS2 (17-20)	LP2 (21-24)	BS3 (25-27)	CT (28-34)
1	- 3.2	KLPPG	WEKRM	**S**RS**S**GR	VYYS	NHIT	NA**S**	Q**S**ERPSG
2	- 3.2	KLPPG	WEK**R**M	SRSSGR	V**Y**Y**S**	N**H**IT	NAS	QSERPSG
3	- 3.2	KLPPG	WEKRM	**SRS**SGR	VYYS	NHIT	NAS	QSERPSG
4	- 3.2	KLPPG	WEKRM	**SRS**SGR	VYYS	NHIT	NAS	QSERPSG
9	- 3.1	KLPPG	WEK**R**M	SRSSGR	V**YYS**	N**H**IT	NAS	QSERPSG

As illustrated in Table 16, the ligand binding poses obtained from the modeling studies revealed similar or even identical sequence patterns. In general, the LP1 region appears to play an important role in most of the predicted and energetically most favorable poses of PC for the WW-1-8 peptide. In the publication by Neitz *et al*. two binding pockets were postulated in WW-1-8, the polar pocket P1 and the positively charged binding groove P2.[85] These two binding grooves were also found when four of the predicted ligand poses were visualized in *PyMol* (cf. Figure 36 A-D). In the case of the proposed ligand binding poses 2 and 9 phosphorylcholine binds to the polar pocket P1 (cf. Figure 36 A and B). Here, the choline moiety is mainly stabilized by polar interactions with Arg9 in BS1, while the phosphate residue is bound to Tyr18 and Ser20 of BS2 and His22 in LP2. In contrast, the predicted ligand positions 3 and 4 suggest that PC binds to the binding groove P2 (cf. Figure 36 C and D). There, the ligand is mainly stabilized by binding to Ser11, Arg12 and Ser13 of LP1 with its phosphate moiety.

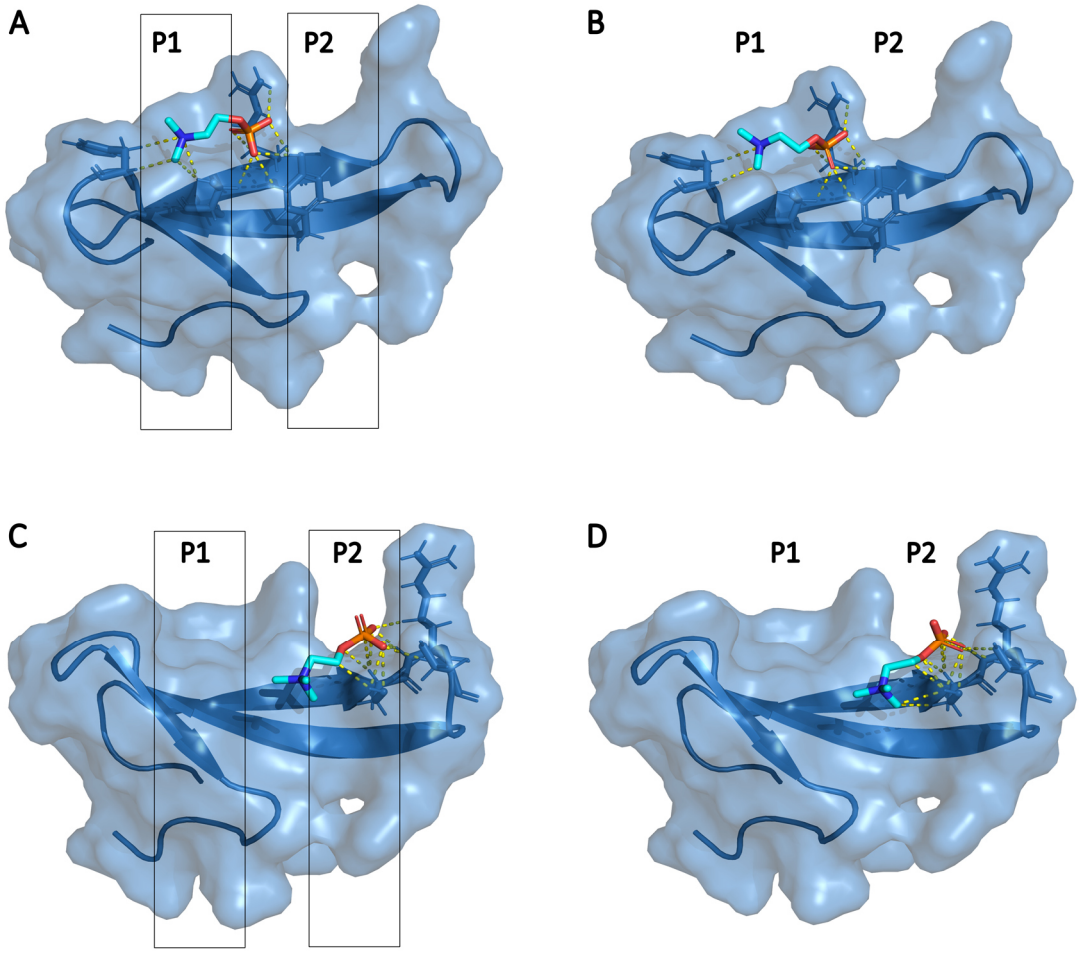


Figure 36: *In silico* modeled interaction of PC with the PC-binding minireceptor WW-1-8.[85] Docking studies were conducted with *AutoDockVina*[126,127]. Postulated binding pockets are indicated as P1 and P2 in each panel, and polar interactions are depicted as yellow dashed lines. The phosphate moiety of PC is illustrated in orange, and the choline moiety in blue. **A**) Predicted ligand pose 2. **B**) Predicted ligand pose 9. **C**) Predicted ligand pose 3. **D**) Predicted ligand pose 4. Three rounds of docking studies were performed on WW-1-8 with the following settings: Grid Box in point size (80, 80, 80), exhaustiveness = 24. *AlphaFold2*[100] was used to prepare the structure prediction of WW-1-8 prior to the docking studies.

Taking into account all the results of the *in silico* docking studies with PC and WW-1-8 in AutoDockVina, and additionally considering the importance of a positively charged amino acid at position 11 for the thermostability of WW-HS-1, the peptide variant WW-HS-PC was designed as shown in Table 17. This variant contains Ser residues at positions 20 and 27, which were identified as functionally relevant in WW-1-8 and additionally the positively charged Arg at position 25, which was found to be interchangeable with Lys in the WW domain iterative sequence analysis, but was observed to bind well to negatively charged phosphate groups.[85,128]

Table 17: Sequence design of the thermostable PC-binding WW domain WW-HS-PC. The sequence of WW-1-8 was extracted from Neitz *et al.*[85] WW-HS-PC has already been published in Lindner *et al.*[94] Residues depicted in blue represent BS regions, important amino acid residues identified by ligand docking studies with *AutoDockVina*[126,127] are displayed in red.

WW domain	Sequence
	-------NT------BS1------LP1-----BS2—-LP2—BS3----CT--
	1 - 10 - 20 - - 30
WW-1-8	H-KLPPG-WEKRM-SRSSGR-VYYS-NHIT-NAS-QSERPSG-OH
WW-HS-1	Ac-PLPPG-WEIRI-DSGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2
WW-HS-PC	H-PLPPG-**WEIRI**-DSGSGR-**IYYS**-NSIT-**RTS**-TWERPRL-OH

WW-HS-PC was subjected to *in silico* docking with *AutoDockVina*[126,127] and compared to WW-1-8 in terms of the predicted binding energies to PC (cf. Table 18). The ligand modeling studies revealed a predicted binding energy of WW-HS-PC to PC of -3.5 kcal/mol, which is minorly increased compared to the predicted affinity of WW-1-8.

Table 18: Predicted binding energies for the binding of PC to WW-1-8 and WW-HS-PC. Ligand pose number is also indicated. *In silico* modeling was performed in triplicates and one representative output is displayed.

WW domain peptide	Number of the predicted ligand position	Predicted binding affinity [kcal/mol]
WW-1-8	2	- 3.2
WW-HS-PC	2	- 3.5

The analysis of *in silico* docking of WW-HS-PC with PC revealed that PC appears to bind to the positively charged pocket P2 (cf. Figure 37). There, polar contacts of the amino acid residues Arg9 of BS1, Ser20 of BS2 and Ser22 of LP2 to PC were identified.

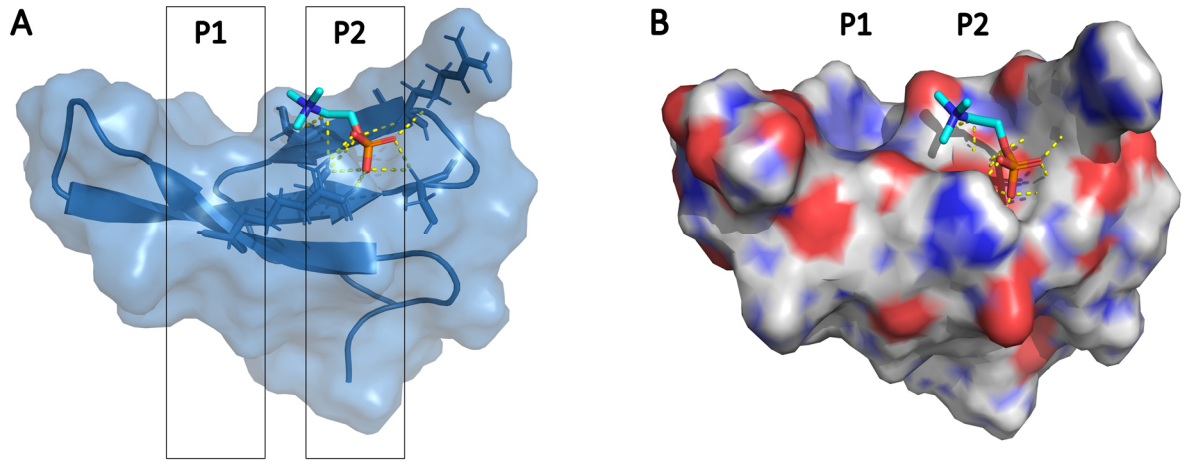


Figure 37: *In silico* docking studies of PC with the designed peptide WW-HS-PC. Two different visualization modes are shown with ligand pose 2. **A**) Transparent surface mode with residues interacting with PC displayed as blue sticks. **B**) Surface mode illustrated by elements. The two postulated binding pockets P1 and P2 from *Neitz et al.*[85] are shown in all panels. The phosphate moiety of PC is depicted in orange, the choline moiety in blue. Three rounds of docking studies were performed with the following settings: Grid Box in point size (80, 80, 80), exhaustiveness = 24. *AlphaFold2*[100] was used to prepare the structure prediction of the peptide prior to the docking studies.

4.2.3.2.2 Docking studies of WW-2-10 with ATP

First, structure prediction of WW-2-10 was performed with *AlphaFold2*[100] and the ATP ligand was docked to the peptide *in silico* with *AutoDockVina*[126,127]. As output, several binding poses of ATP with their respective binding affinity values were generated and the most energetically favored binding poses of ATP were visualized in *PyMol*. For the different structural elements, such as all three BS segments and the LP1 region, which is described as a phosphate binding site,[85] a high occurance of certain amino acids in the respective structural segments was observed, which are summarized in Table 19.

Table 19: Sequence patterns of the respective structural elements obtained from the different predicted poses of ATP to WW-2-10. The sequence of WW-2-10 was extracted from Neitz *et al*.[85] Results of one representative docking round, including the ligand docking pose and the respective predicted binding affinity are shown with binding-relevant residues depicted in red. *In silico* modeling studies were performed in triplicates with *AutoDockVina* [126,127] and the results were analyzed in *PyMol*.

Number of the predicted ligand pose	Predicted binding affinity [kcal/mol]	Sequence regions						
		CT (1-5)	BS1 (6-10)	LP1 (11-16)	BS2 (17-20)	LP2 (21-24)	BS3 (25-27)	NT (28-34)
1	- 6.4	KLPPG	WE**KR**M	**SRSR**GR	VYYH	NHIT	NAE	QSERPSG
2	- 6.3	KLPPG	WE**KRM**	S**RS**RGR	V**Y**YH	NHIT	NAE	QSERPSG
3	- 6.2	KLPPG	WE**KRM**	**SRSR**GR	V**Y**YH	NHIT	NAE	QSERPSG
4	- 6.0	KLPPG	WEKRM	**SRSR**GR	V**Y**YH	NHIT	NAE	QSERPSG
8	- 5.8	KLPPG	WEK**R**M	**SRSR**GR	V**Y**YH	NHIT	NA**E**	QSERPSG

In total, nine different binding poses of ATP to WW-2-10, the *h*Pin1$_{WW}$-derived ATP-binding minireceptor, were found by computational modeling using *AutoDockVina.* The predicted positions 1-3 of ATP exhibited the lowest binding affinity and thus the most energetically favored binding poses, all of which shared some common sequence patterns, such as the presence of Lys and Arg at sequence positions 8 and 9 in BS1 and an alternating pattern of SRSR in LP1 (cf. Table 19). As already postulated in the publication by Neitz *et al*., the positively charged Arg residues, which form a positively charged binding pocket P2, interact with the negatively charged triphosphate tail of ATP via electrostatic interactions (cf. Figure 38 A and C). In addition to the positively charged binding pocket P2, which mainly stabilizes the triphosphate moiety, a polar pocket P1 was identified.[85] This polar binding pocket was also found in the computational modeling presented below and stabilizes the nucleoside moiety of ATP by hydrogen bonding in the predicted and illustrated ligand pose 4 (cf. Figure 38 C and D). In contrast, the illustrated ligand pose 1, with the energetically best predicted affinitiy value, appears to bind ATP exclusively in binding pocket 2.

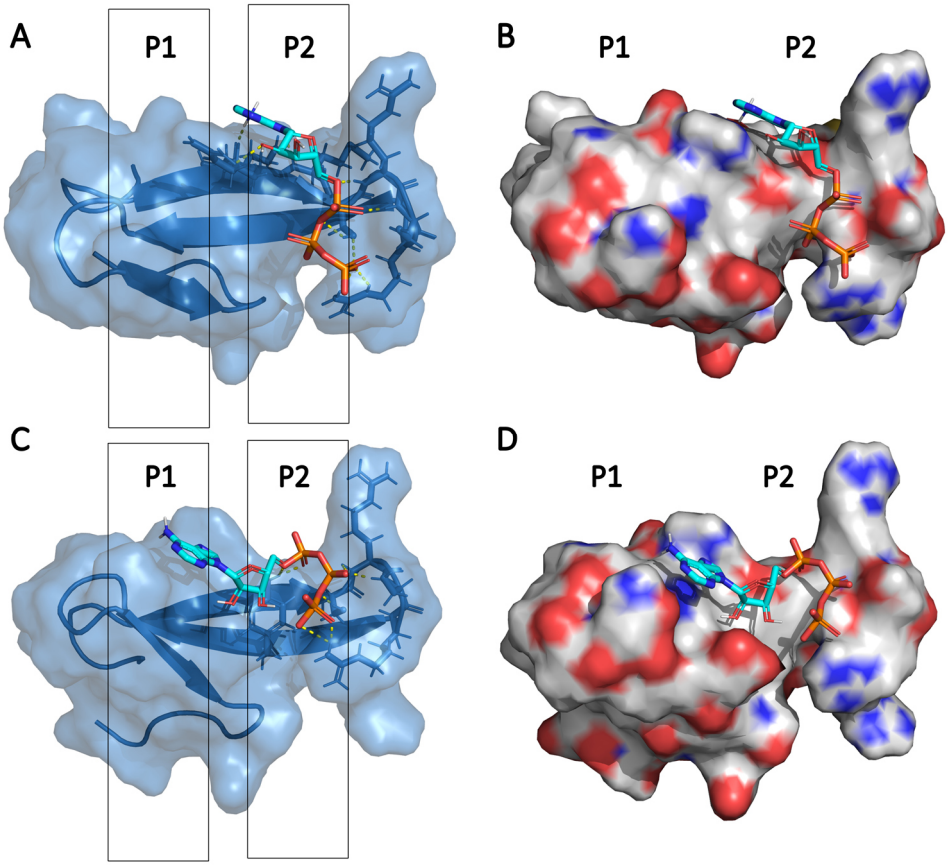


Figure 38: *In silico* modeled interaction of ATP with the ATP-binding minireceptor WW-2-10. Docking studies were conducted with *AutoDockVina*[126,127]. Postulated binding pockets from Neitz *et al.*[85] are indicated as P1 and P2 in each panel, and polar interactions are depicted as yellow dashed lines. The phosphate moiety of ATP is illustrated in orange, and the nucleotide moiety in blue. **A**), **B**) Predicted ligand pose 1 is displayed in two different surface representation modes. **C**), **D**) Predicted ligand pose 4 is illustrated in two different surface representation modes. Three rounds of docking studies were performed on WW-2-10 with the following settings: Grid Box in point size (80, 80, 80), exhaustiveness = 24. *AlphaFold2*[100] was used to prepare the structure prediction of WW-2-10 prior to the docking studies.

The identification of commonly shared sequence patterns, presented above, which derived from ligand docking of the ATP-binding minireceptor WW-2-10, was subsequently combined with the WW-HS-1 core scaffold, leading to the WW-HS-ATP variant (cf. Table 20).

Table 20: Sequence design of a thermostable ATP-binding WW domain WW-HS-ATP. WW-HS-ATP has already been published in Lindner *et al*.[94] Residues depicted in blue represent BS regions, important amino acid residues identified by ligand docking studies with *AutoDockVina* are displayed in red.

WW domain	Sequence
	--------NT-----BS1------LP1-----BS2—-LP2—BS3--- CT--
	1 - 10 - 20 - - 30
WW-2-10	H-KLPPG-WEKRM-SRSRGR-VYYH-NHIT-NAE-QSERPSG-OH
WW-HS-1	Ac-PLPPG-WEIRI-DSGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2
WW-HS-ATP	H-PLPPG-**WEIRI**-D**RSR**GR-**IYYH**-NSIT-**KTE**-TWERPRL-OH

Subsequently, WW-HS-ATP was subjected to *in silico* docking with *AutoDockVina*[126,127] and compared to WW-2-10 in terms of the predicted binding energies to ATP (cf. Table 21). The

ligand modeling studies revealed a predicted binding energy of WW-HS-ATP to ATP of -6.0 kcal/mol, which is minorly decreased compared to the predicted affinity of WW-2-10.

Table 21: Predicted binding energies for the binding of ATP to WW-2-10 and WW-HS-ATP. Ligand pose number is also indicated. *In silico* modeling was performed in triplicates and one representative output is displayed.

WW domain peptide	Number of the predicted ligand position	Predicted binding affinity [kcal/mol]
WW-2-10	1	- 6.4
WW-HS-ATP	2	- 6.0

The analysis of *in silico* docking studies of WW-HS-ATP with ATP showed that ATP appears to bind to the positively charged pocket P2 (cf. Figure 39 A and B). In addition, polar contacts of the amino acid residues Arg9 of BS1, Arg12 of LP1 and Tyr18 of BS2 to ATP were identified.

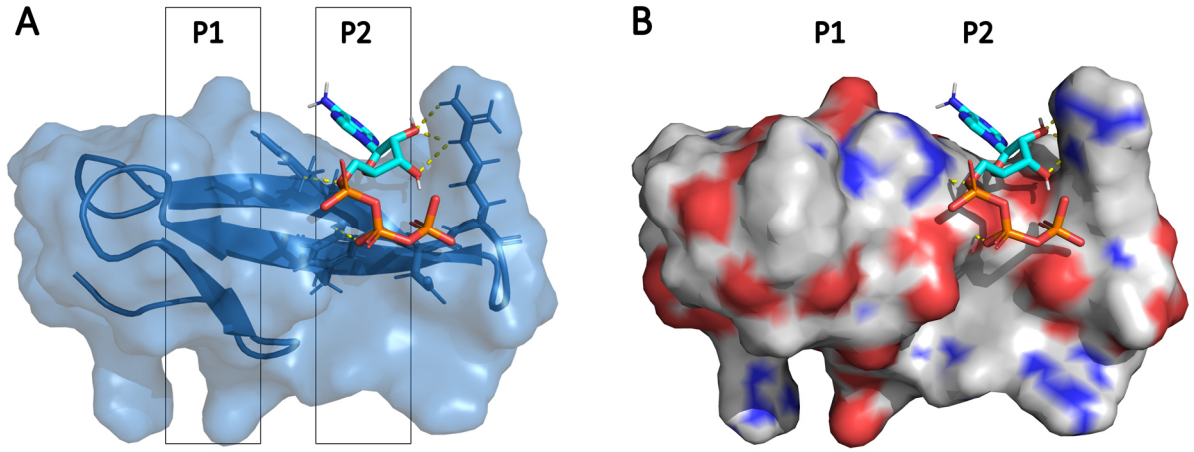


Figure 39: *In silico* docking of ATP with the designed peptide WW-HS-ATP. Two different visualization modes are shown with ligand pose 2. **A**) Transparent surface mode with residues interacting with PC displayed as blue sticks. **B**) Surface mode illustrated by elements. The two postulated binding pockets P1 and P2 from *Neitz et al.*[85] are are shown in all panels. The phosphate moiety of ATP is depicted in orange, the nucleotide moiety in blue. Three rounds of docking studies were performed with the following settings: Grid Box in point size (80, 80, 80), exhaustiveness = 24. *AlphaFold2*[100] was used to prepare the structure prediction of the peptide prior to the docking studies.

The following two sections summarize the most important results of WW-HS-PC and WW-HS-ATP regarding their structure and thermal stability as well as the binding affinity to the respective organophosphates, in comparison to the precursor peptides WW-1-8 and WW-2-10 of Neitz *et al.*[85] Some of the results presented below have already been published in the publication from Lindner *et al.* 2024.[94]

4.2.3.1.3 CD spectroscopic characterization

After the successful synthesis of the designed peptides WW-HS-PC and WW-HS-ATP, structure and thermal stability of both variants was studied with CD spectroscopy. The obtained data were compared with the precursor peptides WW-1-8 and WW-2-10 of Neitz *et al.* (c.f. Figure 40 A, C). As already described by Neitz *et al.*, the exciton coupling of WW-1-8 is only minorly pronounced[85], whereas the redesigned peptide WW-HS-PC shows a high signal intensity in the characteristic region of the WW-domain exciton coupling at 228 nm, indicating a stably folded structure (cf. Figure 40 A). The thermal denaturation curves also revealed significant

differences between the two peptides in terms of curve shapes and melting temperatures. Regarding the shape of the thermal denaturation curve, WW-HS-PC displayed a more sigmoidal curve, in contrast to WW-1-8, which shows a rather flat curve (cf. Figure 40 B). The previously published peptide WW-1-8 was reported to exhibit a melting temperature of 14 °C using a different fitting method.[85] In order to obtain data comparable to those of WW-HS-PC, the raw data of WW-1-8 were reanalyzed using an equation proposed by Norma Greenfield (detailed procedure cf. Chapter 6 Materials and Methods).[101] This reanalysis resulted in a melting temperature of 10.5 °C for WW-1-8, while WW-HS-PC showed an increased melting temperature of 47.5 °C. The structure and thermal stability of WW-HS-ATP was also examined and compared with WW-2-10 from Neitz *et al*.[85] Similar to WW-1-8, WW-2-10 exhibits a low exciton signal at 228 nm, while WW-HS-ATP shows a pronounced signal (cf. Figure 40 C). Thermal denaturation of WW-2-10 revealed a relatively flat curve and a low melting temperature of 16.5 °C, while WW-HS-ATP exhibits a broad transition, with a T_m of 58 °C (cf. Figure 40 D).

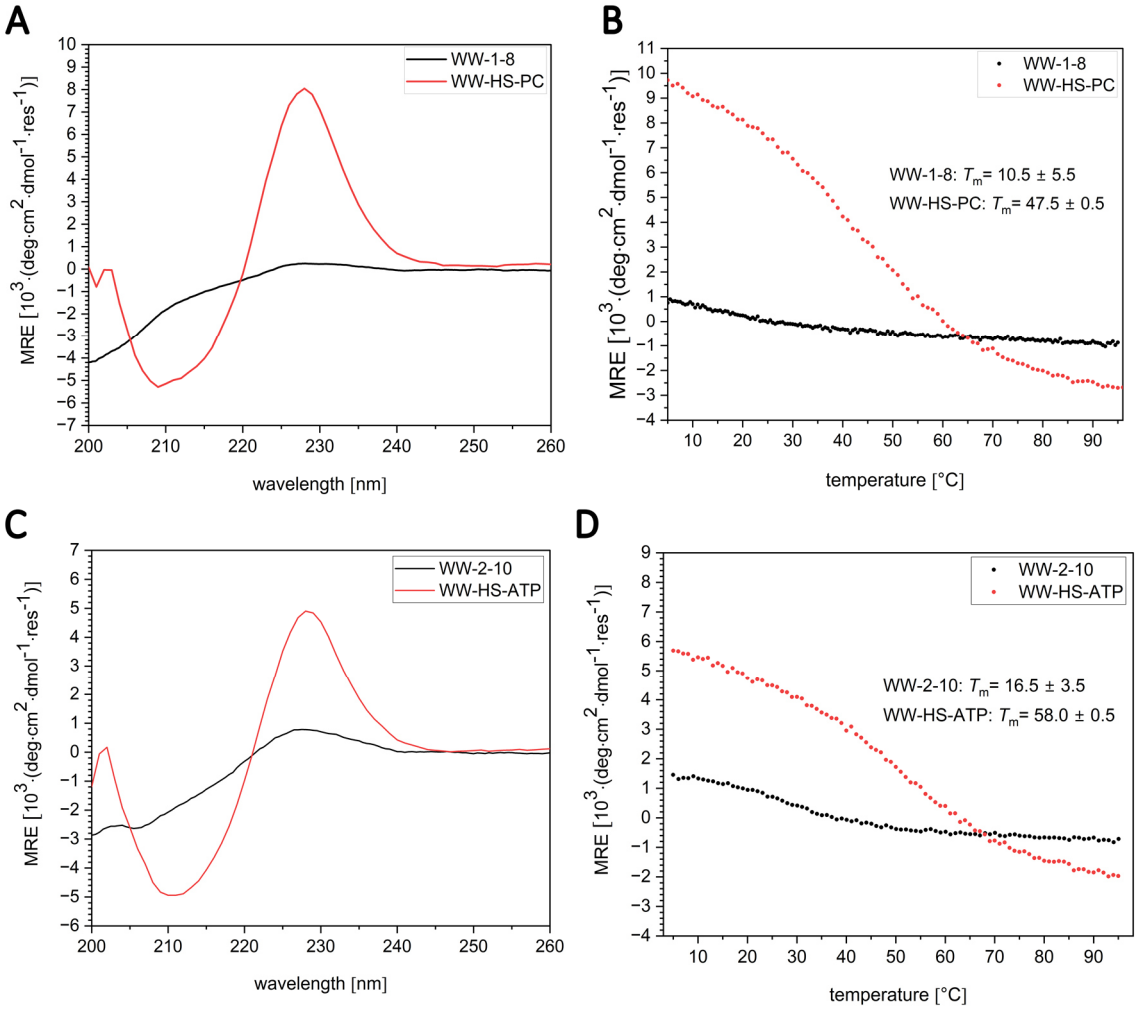


Figure 40: CD spectra and thermal denaturation curves of WW-1-8, WW-HS-PC, WW-2-10 and WW-HS-ATP. A) CD spectra of WW-1-8 and WW-HS-PC at 20 °C are shown as mean of duplicates. **B**) Thermal denaturation curves of WW-1-8 (black) and WW-HS-PC (red) are displayed as mean of duplicates. **C**) CD spectra of WW-2-10 and WW-HS-ATP at 20 °C are depicted as mean values of duplicates. **D**) Thermal denaturation curves of WW-2-10 (black) and WW-HS-ATP (red) are displayed as mean of duplicates recorded from 0 – 95 °C, at a fixed wavelength of 228 nm. Conditions: 100 µM peptide, PBS buffer, pH 7.5. Data of WW-1-8 and WW-2-10 were reused from Neitz *et al*.[85], with permission from © 2022 the Royal Society of Chemistry, and data of WW-HS-PC and WW-HS-ATP were reused in modified form from Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

In summary, the grafting of PC- and ATP-binding sites onto the WW-HS-1 peptide scaffold resulted in the two peptides WW-HS-PC and WW-HS-ATP, which exhhibited significantly higher melting temperatures than their precursor peptides WW-1-8 and WW-2-10.

4.2.3.2.3 Studies on binding affinity

Besides the thermostability of the *de novo* designed WW domain peptides WW-HS-PC and WW-HS-ATP, the objective was to maintain the high binding affinities to their respective organophosphate ligands. Neitz *et al*. previously reported a K_d value of 1.1 ± 0.3 °C for WW-1-8 to PC, and a K_d value of 14 ± 3.4 °C for WW-2-10 to ATP.[85] To investigate the binding affinities of WW-HS-PC to PC, and WW-HS-ATP to ATP, the change in the intrinsic tryptophan fluorescence upon ligand titration was examined. A high binding affinity of WW-HS-PC to PC was identified with a K_d value of 3.0 ± 1.4 µM (cf. Figure 41), which is in a similar micromolar range as WW-1-8. For the binding experiment of WW-HS-ATP to ATP, a K_d value of 4.9 ± 0.9 µM was identified (cf. Figure 41).

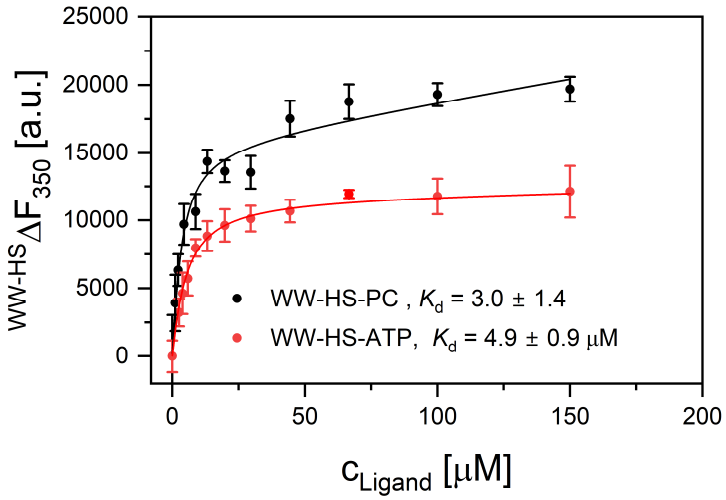


Figure 41: Saturation binding curves of WW-HS-PC titrated with PC, and WW-HS-ATP titrated with ATP. The change in the intrinsic tryptophan fluorescence signal was monitored at 350 nm and fitted to a quadratic equation. This experiment was performed by Jun. Prof. Franziska Thomas. The data have been previously published in Lindner *et al*.[94], the figure was adapted with permission © 2024 the American Chemical Society.

In conclusion, grafting the PC-binding site in WW-1-8, as well as the ATP-binding site in WW-2-10 to the highly thermostable scaffold WW-HS-1 resulted in WW-HS-PC and WW-HS-ATP. The two *de novo* designed peptides exhibited significantly increased melting temperatures compared to their precursor peptides WW-1-8, and WW-2-10. WW-HS-PC, as well as WW-HS-ATP both exhibited similar binding affinities to their respective organophosphate ligands. This successful application of adding non-native functions to thermostable WW-domains further confirms the initial design rationale that grafting ligand binding motifs onto a thermostable core scaffold leads to functional and thermostable peptides.

4.3 Rational *de novo* design of ATP-binding peptides

In addition to the thermostable ATP-binding peptide WW-HS-ATP, another object was to design *de novo* ATP-binding peptide variants that are either based on the other WW-HS-2 core scaffold or are top-down designs of WW-2-10. First, the structure of the designed peptide

variants was predicted using *AlphaFold2*[100], binding energies to ATP were predicted with *in silico* ligand docking in *AutoDockVina*[126,127]. Subsequently, these variants were chemically synthesized, and examined regarding the influence of sequence modifications on structural and thermal stability via CD spectroscopy, as well as on the binding affinity to ATP and its derivatives ADP and AMP by fluorescence spectroscopy. As a final step, the potential catalytic activity of the ATP-binding peptides in the hydrolysis of *para*-nitrophenyl phosphate (*p*-NPP) was investigated. In addition, the possible phosphatase-like properties of all ATP-binding candidates towards ATP and its derivatives ADP and AMP were studied using an analytical RP-HPLC assay.

4.3.1 *In silico* prediction of ATP-binding candidates

The identification of common binding motifs relevant for ATP binding, presented above, derived from ligand docking of the ATP-binding minireceptor WW-2-10, were combined with the WW-HS-1 or WW-HS-2 scaffolds, and several sequence modifications, such as the introduction of Cys in WW-2-10 to induce potential hydrolytic activity, resulted in several possible peptide candidates (cf. Table 22 *vide infra*).

In WW-2-10-HS1-1, the oxidation-sensitive Met at position 10 was replaced by an isosteric norleucine (Nle) residue. In WW-2-10-HS1-2, an Ile was introduced at this position to increase thermostability. To further increase stability, Glu was introduced at position 11 in WW-2-10-HS2, as it was found to be highly structure stabilizing in the iterative sequence analysis (c.f. Chapter 4.1.2). This peptide additionally exhibits a LP2 and BS3 sequence pattern similar to WW-HS-2, but with a Thr residue at sequence position 26 in BS3. Two peptide variants, namely WW-2-10-R9C-sb1 and WW-2-10-R9C-sb2, were engineered with a Cys residue at position 9 of BS1. For these two peptides the aim was to study the influence of the Cys residue on a potential hydrolytic activity (cf. Chapter 4.3.4). The rationale for introducing a Cys residue is based on the fact that this amino acid is frequently present within the active site of hydrolases.[129]

Table 22: Sequences of *de novo* designed ATP-binding peptide variants. Sequences of WW-2-10 and WW-HS-1 have already been published in Neitz *et al.*[85] and Lindner *et al.*[94] Residues depicted in blue represent BS regions, important amino acid residues identified by ligand docking studies with *AutoDockVina*[126,127] are displayed in red.

WW domain	Sequence
	--------NT------BS1--------LP1------BS2—LP2—BS3---CT--
	1 - 10 - 20 - - 30
WW-2-10	H-KLPPG-**WEKRM** -S**RSR**GR-**VYYH**-NHIT-**NAE**-QSERPSG-OH
WW-HS-1	Ac-PLPPG-**WEIRI** -DSGSGR-**IYYY**-NSIT-**KTT**-TWERPRL-NH_2
WW-2-10-HS1-1	H-PLPPG-**WEKRNle**-S**RSR**GR-**IYYH**-NSIT-**KTE**-TWERPRL-OH
WW-2-10-HS1-2	H-PLPPG-**WEKRI** -S**RSR**GR-**IYYH**-NSIT-**KTE**-TWERPRL-OH
WW-2-10-HS2	H-PLPPG-**WEIRI** -E**RSR**GR-**IYYH**-NHIT-**NTE**-TWERPRL-OH
WW-2-10-R9C-sb1	H-KLPPG-**WEICI** -S**RSR**GR-**VYYH**-NHIT-**NTE**-QSERPSG-OH
WW-2-10-R9C-sb2	H-KLPPG-**WEICI** -E**RSR**GR-**VYYH**-NHIT-**NTE**-QWERPSG-OH

Subsequently, structure predictions of all designed peptide variants were conducted with *AlphaFold2*[100] and were examined regarding their predicted ATP binding affinity by *in silico*

docking with *AutoDockVina*[126,127]. For all variants the identical settings were utilized as employed for the ligand docking with WW-2-10 and WW-HS-ATP, in order to obtain comparable predicted binding energy values, which are presented in Table 23. As the utilized version of *AlphaFold2*[100] was unable to predict a secondary structure for peptides containing non-proteinogenic amino acids, such as the Nle present in WW-2-10-HS1-1, no binding energies could be predicted for this peptide. In conclusion, the peptide designs exhibited comparable *in silico* binding energy values, within a range of –5.4 to –6.6 kcal/mol. However, the WW-2-10-HS1-2 displayed slightly elevated binding energy values, while the remaining peptide variants demonstrated reduced binding affinity to ATP in comparison to the previously designed ATP minireceptor WW-2-10.

Table 23: Predicted binding energies for the binding of ATP to the *de novo* designed variants. Ligand pose number is also indicated. *In silico* modeling was performed in triplicates and one representative output is displayed. For WW-2-10-HS1-1 no predicted binding energy values could be determined, designated as "n. a.".

WW domain peptide	Number of the predicted ligand pose	Predicted binding affinity [kcal/mol]
WW-2-10	1	- 6.4
WW-HS-ATP	2	- 6.0
WW-2-10-HS1-1	n.a.	n.a.
WW-2-10-HS1-2	1	- 6.6
WW-2-10-HS2	6	- 5.4
WW-2-10-R9C-sb1	1	- 6.2
WW-2-10-R9C-sb2	4	- 5.8

The analysis of the potential ATP poses within the binding pockets of the respective peptides, revealed decisive differences in the orientation of the triphosphate of ATP (cf. Figure 42 A-D). The ATP appears to be elongated when binding to the pronounced binding pocket P2 in WW-2-10-HS1-2 (cf. Figure 42 A). In all other WW-2-10-variants, the ATP adopts a rather bent structure due to the electrostatic interactions, which are predominantly formed with amino acids located in the P2 pocket. Although the ligand appears to bind most preferentially to the positively charged binding pocket P2 in all peptide variants, the triphosphate moiety seems to extend to the polar pocket for both Cys variants (cf. Figure 42 C and D).

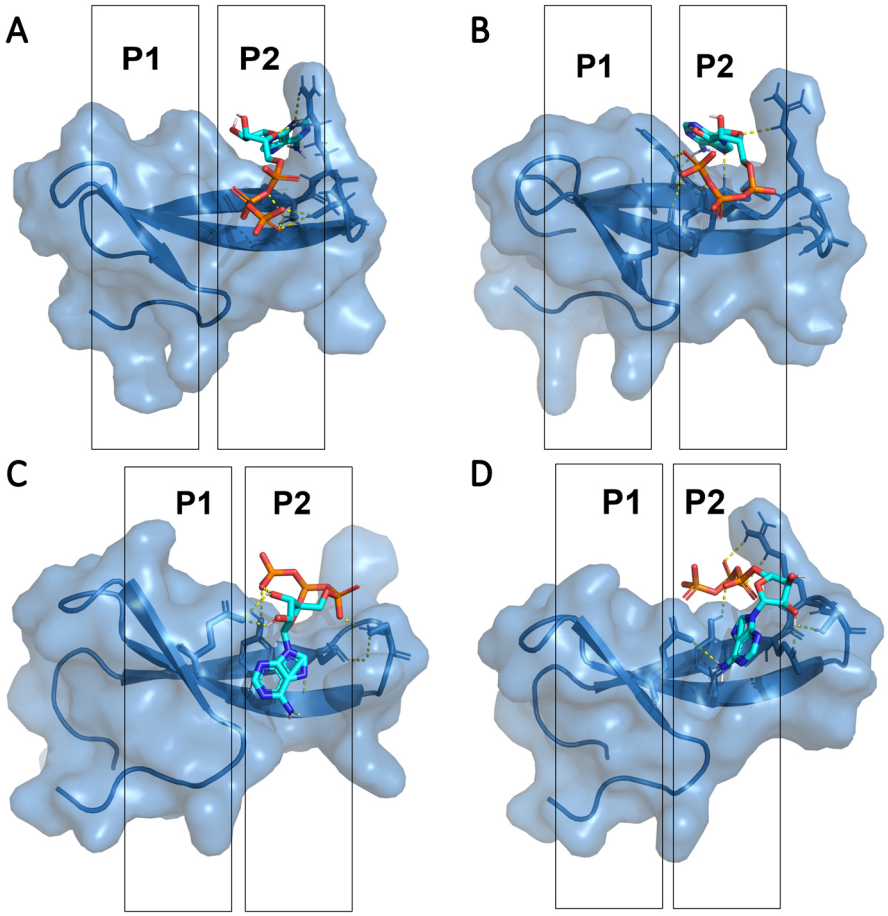


Figure 42: *In silico* docking of ATP with the *de novo* designed peptide variants. Docking studies were conducted with *AutoDockVina*. Postulated binding pockets are indicated as P1 and P2 in each panel, and polar interactions are depicted as yellow dashed lines. The phosphate moiety of ATP is illustrated in orange, and the nucleotide moiety in blue. **A**) Predicted ATP pose 1 with WW-2-10-HS1-2. **B**) Predicted ATP pose 6 of WW-2-10-HS2. **C**) Predicted ATP pose 1 with WW-2-10-R9C-sb1. **D**) Predicted ATP pose 4 with WW-2-10-R9C.

The *in silico* docking studies indicated that all *de novo* designed peptide variants are predicted to bind to ATP. The computational results were experimentally assessed by employing fluorescence titration spectroscopy to elucidate the binding properties of the peptide designs to ATP in greater detail (cf. Chapter 4.3.3).

4.3.2 CD spectroscopic characterization

Following the successful synthesis of the peptide variants, structure analysis and thermal denaturation were performed using CD spectroscopy. In addition to the peptide solutions prepared in PBS buffer (pH 7.4) directly prior to the CD measurements, peptide samples incubated overnight were analyzed to determine whether there was a change in structure and thermal stability induced by incubation. Furthermore, the impact of ATP and Mg^{2+} on the peptide's structure and overall stability was investigated. The two published peptides WW-2-10 and WW-HS-ATP served as control peptides for the *de novo* designed variants.

First, CD spectra of all five redesigned peptides were compared to that of the ATP-binding minireceptor WW-2-10[85] and WW-HS-ATP (cf. Figure 43 A).[94] The peptides WW-2-10 and WW-2-10-R9C-sb1 exhibited the lowest signal intensity at 228 nm in their CD spectra. (cf. Figure 43 A). The highest exciton signals were identified for WW-2-10-HS2 and WW-HS-ATP. The analysis of thermal denaturation profiles revealed differences in the respective curve shapes (cf. Figure 43 B). The ATP-binding peptide WW-2-10 exhibits a relatively flat curve shape and a low melting temperature of 16.5 °C (cf. Table 24). In the paper by Neitz *et al.*, a melting temperature of 36.5 °C was reported for WW-2-10.[85] This discrepancy can be attributed to the different two-state folding model employed for data analysis, which assumes a two-state transition between a folded dimer and an unfolded monomer and displays the data as a fraction of folded molecules.[130] The Cys peptide variant WW-2-10-R9C-sb1 displayed a comparable melting temperature to WW-2-10, with a value of 18.0 °C (cf. Table 24). The peptide variant with the highest thermostability is WW-HS-ATP, with a T_m of 58.0 °C and a broad transition region (cf. Table 24 and Figure 43 B).

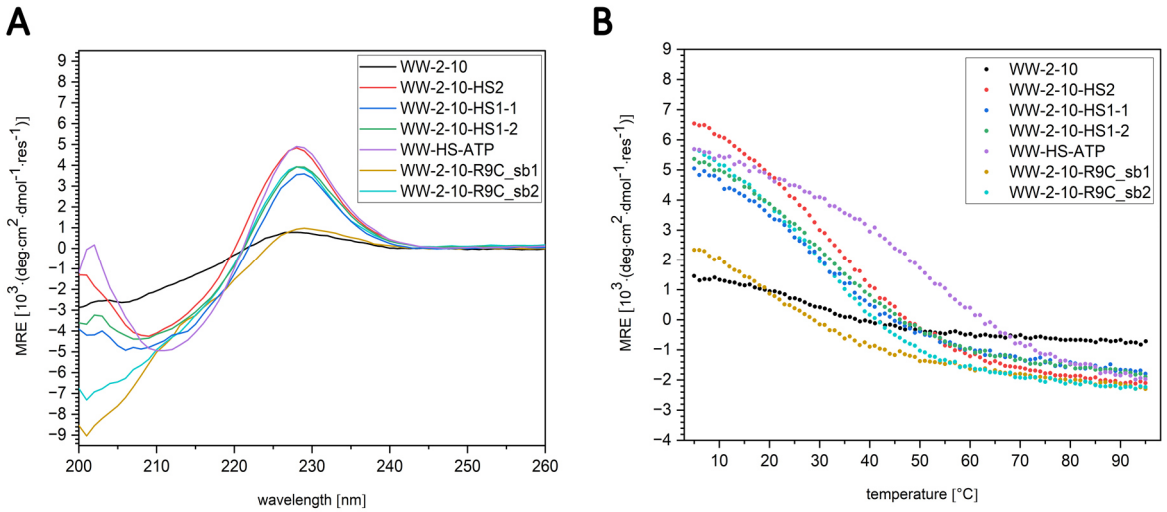


Figure 43: CD spectra and thermal denaturation curves of the *de novo* designed ATP variants. A) CD spectra of, displayed as mean values of triplicate measurements (except for WW-2-10-R9C-sb1, duplicate measurement), performed at 20 °C. **B**) Thermal denaturation curves of the peptide variants, depicted as mean values of triplicate measurement (except for WW-2-10-R9C-sb1, duplicate measurement), measured at a fixed wavelength of 228 nm from 5 -95 °C. Conditions: 100 µM peptide, PBS buffer, pH 7.4.

In conclusion, the peptide WW-2-10-R9C-sb1 exhibited the least stable folded structure at 20 °C and the lowest thermostability. All other peptides showed moderate thermal stabilities, with T_m values summarized in Table 24.

Table 24: Melting temperatures of the de novo designed ATP variants. Mean values of triplicate measurements are shown, except for WW-2-10-R9C-sb1, marked with an "*".

WW domain	Melting temperature [°C]
WW-2-10	16.5 ± 3.5
WW-HS-ATP	58.0 ± 0.5
WW-2-10-HS2	34.0 ± 0.8
WW-2-10-R9C-sb1*	18.0 ± 3.7*
WW-2-10-R9C-sb2	30.5 ± 0.9
WW-2-10-HS1-1	29.5 ± 1.5
WW-2-10-HS1-2	30.0 ± 1.0

As a next step, CD experiments were conducted to test the influence of overnight incubations (stored at 4 °C) of the four selected variants WW-2-10, WW-HS-ATP, WW-2-10-HS2, and WW-2-10-HS1-1 on their secondary structure and thermal stability. Due to the fact that the two R9C peptide variants did not reveal high signal intensities in their CD spectra and thermal denaturation curves, as well as their primary design reason was for exploratory studies of enzymatic activity, these two peptides were not considered. Among the studied peptides, WW-2-10-HS2 and WW-HS-ATP were selected because they both exhibited the highest thermal stabilities and differed in the HS core scaffold used for their design (cf. Table 24 and Table 22). In addition, WW-2-10-HS1-1 was selected for the CD incubation studies because it contains Nle at position 10, in order to avoid structural changes due to Met oxidation during long-term incubation.[131]

In addition to incubations containing only the peptide in PBS (pH 7.5), the effect of ATP and ATP with Mg^{2+} as incubation additives was investigated. First, the published ATP-binding WW-domain WW-2-10 was studied. The CD spectrum analysis of WW-2-10 under the four different conditions, including the CD spectrum of WW-2-10 without incubation (measurement data from Neitz *et al*. paper[85]), peptide incubation overnight in PBS buffer, peptide incubation with ATP (1:1 ratio), and peptide incubation with ATP (1:1 ratio) and 1 mM Mg^{2+} (1:1:10), was performed. The structure analysis revealed different intensities for the exciton signal between the four conditions (cf. Figure 44 A). The highest signal at 228 nm was identified for the sample of overnight incubation without additives, whereas the lowest intensity was identified for the sample without incubation. In general, it appears that all incubated samples revealed a more stably folded structure than the CD sample directly measured after preparation. Nevertheless, the results obtained for the peptide incubation with ATP and Mg^{2+} are based on a single measurement and require confirmation through a second meausurement in the future. The melting temperatures were obtained by fitting the data of the thermal denaturation curve to Eq. 2 and 3 (cf. Material and Methods, section 6.2.8), which was applied for all four peptides and their different conditions. The subsequent fitting method was employed to reduce the fitting error and to accomodate the disparate curve profile patterns, resulting in some different melting temperatures (cf. Table 24), obtained for the previously presented values of the peptide samples without incubation. Although, the thermal denaturation profiles of the peptide variants exhibit variability regarding their shape and signal intensities (cf. Figure 44 B), the melting temperatures are within a similar range (cf. Table 25). The incubation of the peptide with ATP did not affect the stability. It remains to be seen whether an incubation of

the peptide with ATP and Mg^{2+} will result in a significant structural stabilization, which will be confirmed by a second measurement in the future.

The CD spectra analysis of WW-2-10-HS2 revealed a slight decrease in the exciton signal at 228 nm for the incubated peptide sample and the one that also contained ATP (cf. Figure 44 C). The single measurement result indicates a possible stabilizing effect of Mg^{2+} on the incubated peptide sample with ATP. However, this effect requires confirmation through a second measurement. Thermal denaturation studies of the four different conditions revealed similar curve patterns with a sigmoidal shape for all samples, although there was some variation in the intensities (cf. Figure 44 D), resulting in similar melting temperatures. In conclusion, the incubation period did not significantly impact the structural and thermal stability of WW-2-10-HS2. The CD spectrum for WW-2-10-HS1-1, indicated a slight increase in the exciton signal for the peptide sample incubated overnight without additives (cf. Figure 44 E). All other examined conditions appear to have no impact on the stability of the structure. Thermal denaturation studies revealed highly similar curve shapes for all conditions (cf. Figure 53 F). Only the signal intensities for the peptide sample after incubation were found to be slightly elevated compared to the other samples, yet this did not affect the melting temperature, which remained consistent with the other conditions (cf. Table 35). In conclusion, the incubation of WW-2-10-HS1-1 overnight and the addition of ATP did not significantly impact the structural and thermal stability.

The previously identified most thermostable variant, WW-HS-ATP, demonstrated no significant difference in the exciton signal at 228 nm across the different conditions. However, a minor increase was observed in the sample with ATP and Mg^{2+} after incubation(cf. Figure 44 G). The analysis of the thermal denaturation curves revealed a high degree of similarity in all curve shapes (cf. Figure 44 H), resulting in similar melting temperatures (cf. Table 25). These findings indicate that the structure and high thermal stability of WW-HS-ATP was not influenced by the applied incubation conditions.

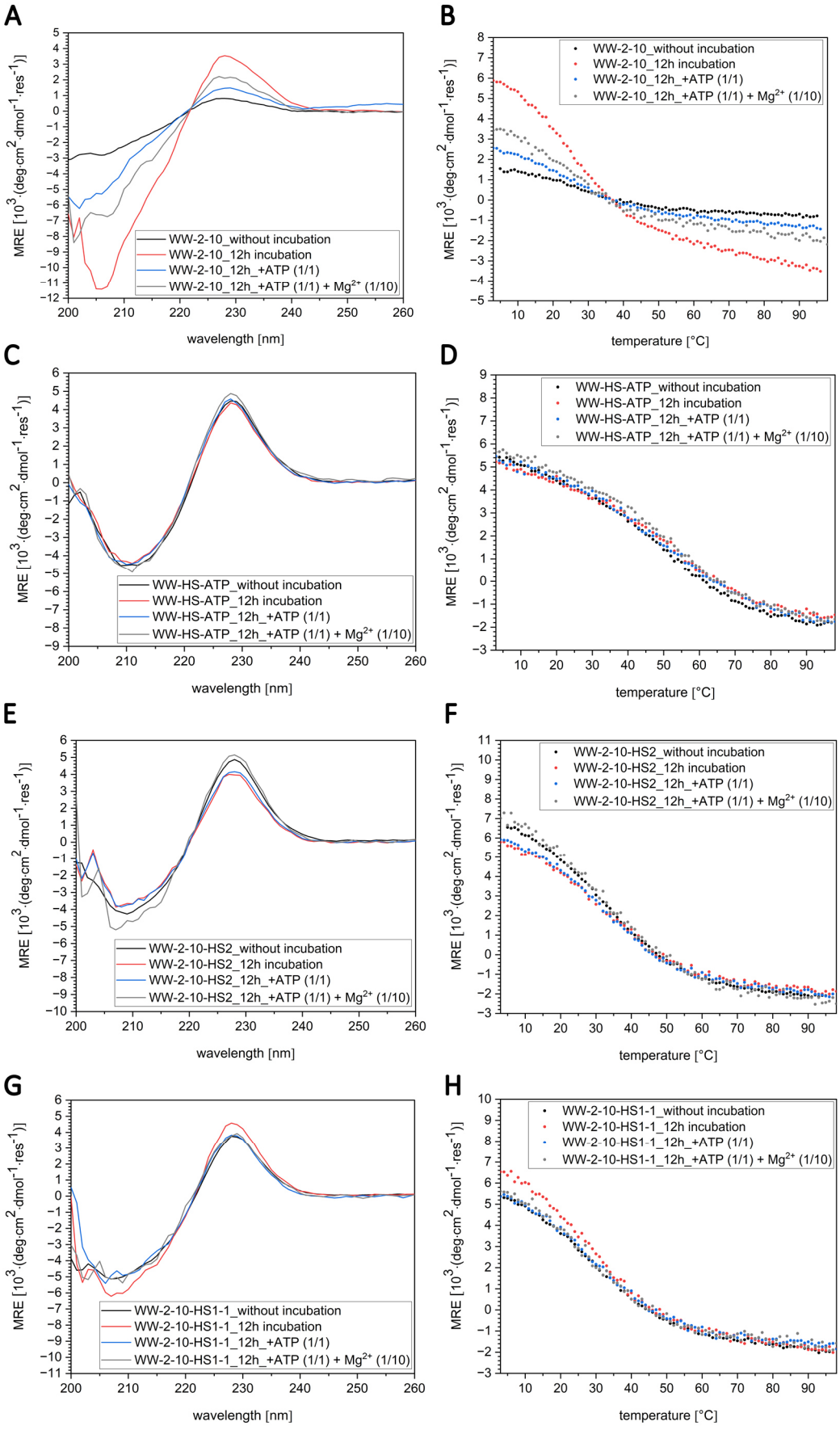


Figure 44: CD spectra and thermal denaturation curves of four selected ATP-binding peptide variants after applying different incubation conditions. CD spectra of **A**) WW-2-10, **C**) WW-HS-ATP, **E**) WW-2-10-HS2, and **G**) WW-2-10-HS1-1. CD spectra were recored from 190 – 260 nm at 20 °C. With the exception of the incubations

with ATP and $Mg^{2+}$for WW-2-10, WW-2-10-HS2, and WW-2-10-HS1-1, the mean of duplicates is shown. Thermal denaturation curves of **B**) WW-2-10, **D**) WW-HS-ATP, **F**) WW-2-10-HS2, and **H**) WW-2-10-HS1-1, recorded from 0 -98 °C at a fixed wavelength of 228 nm. Here, also mean of duplicate data are shown, except for the incubations with ATP and $Mg^{2+}$for WW-2-10, WW-2-10-HS2 and WW-2-10-HS1-1. Conditions: 100 µM peptide, 100 µM ATP if present, 1 mM Mg^{2+} if present, PBS buffer, pH 7.4.

In summary, the designed peptides demonstrated no significant impact of incubation or the addition of ATP, as well as mixtures of ATP with Mg^{2+}, on the structure or thermal stability. This suggests that the design of all *de novo* designed peptides except for the two Cys variants, which was based on two distinct highly thermostable scaffolds with grafted ATP-binding motifs, provides enhanced robustness to prolonged incubation conditions.

Table 25: List of melting temperature values of all *de novo* designed ATP-binding peptide variants at different conditions. The determination of melting temperatures was performed by using a two-state model with linear correction for the upper and lower curve plateau. Melting temperatures that have been previously presented (WW-2-10, WW-HS-ATP) and differ due to data reanalysis with the mentioned fitting method are marked with an *.

WW domain	Incubation method	Melting temperature [°C]	Number of measurements (n)
WW-2-10	No incubation	29.0 ± 2.0*	n=2
	12 h incubation	27.5 ± 1.0	n=2
	12 h incubation with ATP (1:1)	28.5 ± 1.5	n=2
	12 h incubation with ATP (1:1) and Mg^{2+} (1:10)	31.5 ± 1.5	n=1
WW-HS-ATP	No incubation	57.0 ± 1.0*	n=2
	12 h incubation	59.5 ± 1.5	n=2
	12 h incubation with ATP (1:1)	56.0 ± 2.0	n=2
	12 h incubation with ATP (1:1) and Mg^{2+} (1:10)	58.0 ± 1.0	n=2
WW-2-10-HS2	No incubation	36.0 ± 1.0*	n=2
	12 h incubation	36.5 ± 1.5	n=2
	12 h incubation with ATP (1:1)	35.5 ± 1.0	n=2
	12 h incubation with ATP (1:1) and Mg^{2+} (1:10)	35.0 ± 2.5	n=1
WW-2-10-HS1-1	No incubation	32.5 ± 1.0*	n=2
	12 h incubation	32.0 ± 1.5	n=2
	12 h incubation with ATP (1:1)	34.0 ± 1.5	n=2
	12 h incubation with ATP (1:1) and Mg^{2+} (1:10)	31.2± 2.5	n=1

4.3.3 Studies on binding affinity

Following the performance of experiments to assess the structural and thermal stability of the designed peptides, their binding affinity to ATP, ADP and AMP were examined. Furthermore, the impact of 2 mM Mg^{2+} on the ATP binding was investigated, as the divalent cation Mg^{2+} was

previously identified to form a complex with the negatively charged triphosphate moiety of ATP in nature.[132] Binding studies were performed using a fluorescence titration approach, whereby the change in intrinsic tryptophan fluorescence upon ligand titration (0 – 250 µM) was monitored, with a constant peptide concentration of 2 µM.

For all peptide variants, an enhancement of the tryptophan fluorescence intensity was observed with increasing ligand concentrations. However, the degree of increase was contingent upon the quantity of tryptophan residues present in the amino acid sequence. The two peptide variants, WW-2-10 and WW-2-10-R9C-sb1, both contained a single tryptophan residue and exhibited the smallest change in fluorescence intensity at 350 nm in three ligand binding experiments. Only for the binding data obtained from ADP and AMP ligand titration, the difference in fluorescence intensity was comparable to that observed for the peptide variants comprising two Trp residues (cf. Figure 45 C and D).

In general, all of the six designed peptides demonstrated binding affinity to ATP in the low micromolar range (cf. Figure 45 A and Table 26). Additionally, the published ATP-binding minireceptor WW-2-10 of Neitz *et al*.[85] was examined as a positive control. Upon comparison of the binding affinities of all seven studied peptides to ATP, no significant differences were identified in their K_d values (cf. Table 26). Although there were great differences in the change of intrinsic fluorescence intensity, the saturation binding curves revealed similar curve shapes (cf. Figure 45 A).

The presence of Mg^{2+} in the incubation mixtures of ATP, resulted in alterations to the binding curves (cf. Figure 45 B) and their binding affinity to ATP for some peptides. One peptide, WW-2-10-R9C-sb1, exhibited a slight increase in the K_d value of 1.7 µM (cf. Table 26). It is likely that the Gln residue in the C-terminal region plays a role in this process, as it has been demonstrated in the literature that conserved Gln residues in ATP-binding-cassette (ABC) transporters interact with the divalent cation cofactor Mg^{2+}, thereby influencing Mg^{2+}-ATP binding by interacting with the divalent cation cofactor.[133] The four peptides, WW-2-10, WW-2-10-HS2, WW-2-10-HS1-2, and WW-HS-ATP, exhibited decreased binding affinities to ATP, when Mg^{2+} was present in the incubations (cf. Table 26). In contrast, for all other variants tested, the addition of Mg^{2+} did not result in any change to their ATP binding affinity (cf. Table 26).

The results of the binding studies towards ADP (cf. Figure 45 C) revealed that all tested peptides showed micromolar binding affinities. In comparison to the binding affinity observed for ATP, three variants (WW-2-10, WW-2-10-HS1-1, and WW-2-10-HS1-2) exhibited slightly increased binding affinities (cf. Table 26), whereas the other peptides revealed similar K_d values. One possible explanation for the increased binding affinity is that ADP is a smaller molecule than ATP. This may facilitate access to the positively charged binding pocket P2, where it can interact with the positively charged residues within the peptide's binding site.

Binding studies with AMP (cf. Figure 45 D) revealed that all of the six tested peptide variants (WW-HS-ATP was not measured due to temporary stock shortage) bind to AMP in the micromolar range (cf. Table 26). For AMP binding, similar affinities were identified for the peptides WW-2-10, WW-HS-2, WW-2-10-HS1-2, and WW-2-10-R9C-sb2, as were observed in the ATP studies. However, WW-2-10-HS1-1 appears to exhibit a weaker binding affinity to

AMP (K_d = 4.2 ± 2.1) compared to ATP (K_d = 1.2 ± 0.2). In contrast, WW-2-10-R9C-sb1 binds more strongly to AMP than to ATP (cf. Table 26).

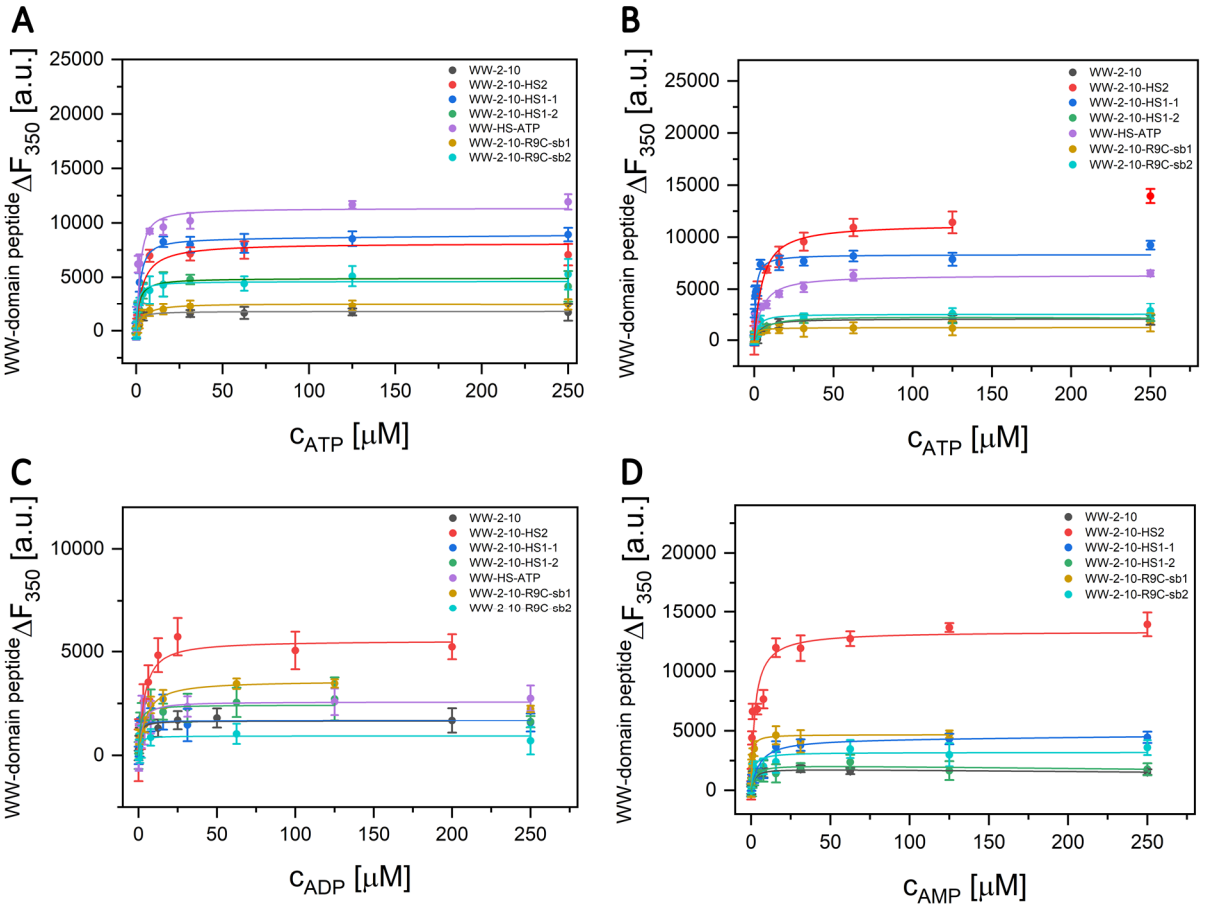


Figure 45: Saturation binding curves of the *de novo* designed ATP-binding peptides to organophosphate ligands. A) Binding curves for the seven variants with ATP. **B**) Binding curves for the seven variants with ATP in the presence of 2 mM Mg^{2+}. **C**) Binding curves for the seven variants with ADP. **D**) Binding curves for six variants with AMP. Experimental conditions: intrinsic tryptophan fluorescence signal was monitored at 350 nm, 2 µM peptide concentration, ligand concentration (0 – 250 µM), in 1xPBS (pH 7.4). Ligand-peptide mixtures were stored at 4 °C under constant shaking at 200 rpm. For determining the K_d values, the binding curves were fitted to a hyperbolic or quadratic equation, with the standard error (SE) as error bars.

As previously described for WW-2-10-R9C-sb1, which revealed a higher binding affinity to ATP in the presence of Mg^{2+}, some of the observed differences in the binding affinity towards ATP, ADP and AMP among the studied peptides may be attributed to the sequence modifications that were implemented. In contrast, some amino acid exchanges were conducted, which were presumed to potentially influence the binding affinity. These included the substitution of Ser11 with a negatively charged amino acid, such as Asp or Glu. However, these modifications did not appear to affect the binding properties of the peptide variants. It is possible that these substitutions exclusively led to an increase in structural stabilization, as was observed in the case of WW-HS-ATP.

Table 26: Binding affinities of the different *de novo* designed peptides and ATP, ADP and AMP. K_d values were obtained by fitting the binding curves to a quadratic equation, those with an * were obtained by fitting the data to a hyperbolic function.

Peptide variant	Ligand	K_d [µM]
WW-2-10	ATP	1.5 ± 0.3 *
	ATP (Mg^{2+})	3.2 ± 1.0
	ADP	0.8 ± 0.2 *
	AMP	1.3 ± 0.5
WW-2-10-HS2	ATP	2.4 ± 1.2
	ATP (Mg^{2+})	4.0 ± 0.9
	ADP	2.8 ± 0.7
	AMP	2.4 ± 0.8 *
WW-2-10-HS1-1	ATP	1.2 ± 0.2
	ATP (Mg^{2+})	0.8 ± 0.2 *
	ADP	0.5 ± 0.2 *
	AMP	4.2 ± 2.1
WW-2-10-HS1-2	ATP	1.6 ± 0.4 *
	ATP (Mg^{2+})	5.4 ± 0.8
	ADP	0.9 ± 0.2 *
	AMP	1.6 ± 0.6
WW-HS-ATP	ATP	1.2 ± 0.4
	ATP (Mg^{2+})	4.0 ± 1.4
	ADP	1.5 ± 0.6
	AMP	No data
WW-2-10-R9C-sb1	ATP	3.1 ± 1.3
	ATP (Mg^{2+})	1.4 ± 0.2 *
	ADP	3.7 ± 1.3
	AMP	0.5 ± 0.1 *
WW-2-10-R9C-sb2	ATP	0.9 ± 0.2 *
	ATP (Mg^{2+})	1.6 ± 0.7
	ADP	0.9 ± 0.4 *
	AMP	1.4 ± 0.5 *

4.3.4 Studies on hydrolytic activity

After a detailed characterization of the *de novo* designed ATP-binding variants, with respect to their structural and thermal stability, as well as their binding affinity for ATP, ADP and AMP, two exploratory studies were performed to test whether the peptides exhibit phosphatase-like activities. A colorimetric kinetic study was conducted to investigate a general phosphatase activity towards phosphorylated molecules, such as *para*-nitrophenyl phosphate (*p*-NPP). As a second experiment, an RP-HPLC based assay was performed to study possible ATPase or ADPase enzymatic activity of the peptide variants.

First, a colorimetric assay was conducted to examine the general ability of the peptides to cleave a phosphate bond. The *p*-NPP molecule was utilized as a substrate, and its degradation to the colorful nitrophenolat ion was monitored at 405 nm. The selected absorption wavelength was based on the specifications provided by the manufacturer and for the

purpose of comparability with the published *de novo* designed ATPase, designated as AltTPase from Wang *et al*.[134] As a positive control enzyme, an acid phosphatase from potato was utilized, with a pH optimum at 4.8. Given that the activity of native phosphatases is pH-dependent and they are classified as acid or alkaline phosphatases[12], the kinetic measurements were performed in phosphate-buffered saline with two different pH values (4.8 and 7.4) for 30 min at room temperature. In addition to the seven ATP-binding peptides and the phosphatase, the autohydrolysis of *p*-NPP over time was studied as a reference. Following the completion of the experiment, the mean values of the obtained duplicate or triplicate absorbance data were calculated and fitted to a linear regression model to identify the slopes m_{obs} for the determination of the rate constants k_{obs} and the catalytic constants k_2 (calculations cf. Material and Methods, Chapter 6.2.9). The results of the exploratory phosphatase activity test are shown in Figure 47. The positive control phosphatase demonstrated high hydrolytic activity in the two PBS buffers with different pH values, despite the pH optimum of the enzyme being indicated as 4.8 by the manufacturer.

The hydrolytic activity screening experiment performed at pH 4.8 revealed that three peptides, namely WW-2-10-HS2, WW-HS-ATP, and WW-2-10-R9C-sb1 exhibited hydrolytic activity above the *p*-NPP autohydrolysis background signal (cf. Figure 47 A). The highest phosphatase activity under acidic conditions was identified for WW-2-10-HS2, with a k_2 value of 1.1 ± 0.004 10^{-3} s^{-1}, which represents 0.13 % of the determined enzymatic activity of the potato phosphatase. The WW-HS-ATP peptide also exhibited hydrolytic activity towards *p*-NPP with a k_2 value of 0.83 ± 0.09 10^{-3} s^{-1}, but with lower catalytic efficiency compared to WW-2-10-HS2. *In silico* binding studies of *p*-NPP with both peptide variants were conducted with *AutoDockVina* and visualized with *PyMol* (cf. Figure 46 A and B). The results revealed that only Asp11 in the LP1 region of WW-2-10-HS2 was different to WW-HS-ATP. In WW-HS-ATP polar interactions between Tyr17, Glu27, and Trp29 were identified, which are also found in WW-2-10-HS2 in addition to the negatively charged Asp11. In native acid phosphatases, His residues frequently play a pivotal role in the active site of enzymes. A study performed by Pfeiffer *et al.* in 2022 revealed that these residues contribute to enhanced phosphatase activity when compared to a variant with a substituted Asp nucleophile.[135] However, the His residue at position 19 in BS2 appears to be in a moderate distance to the predicted favorable binding site in WW-2-10-HS2 and WW-HS-ATP, forming no close interaction with *p*-NPP (cf. Figure 46 A and B). It is necessary to examine whether the His19 residue plays a role in the hydrolytic activity of both peptides in future studies. The peptide variant WW-2-10-R9C-sb1 also exhibited low hydrolytic activity towards *p*-NPP at pH 4.8. The introduced Cys residue at position 9 in BS1 and the Ser at position 11 may be part of the active site. This latter residue was identified as part of the binding site for *p*-NPP in *in silico* modeling with *AutoDockVina*[126,127] (cf. Figure 46 C). In addition, Ser11, Ser13, and Arg14 in LP1 were identified as residues binding to *p*-NPP. Furthermore, the three residues Tyr17, Glu27, and Ser29 were part of the binding site of WW-2-10-R9C-sb1, similar to the other two hydrolytic peptide variants, as previously mentioned.

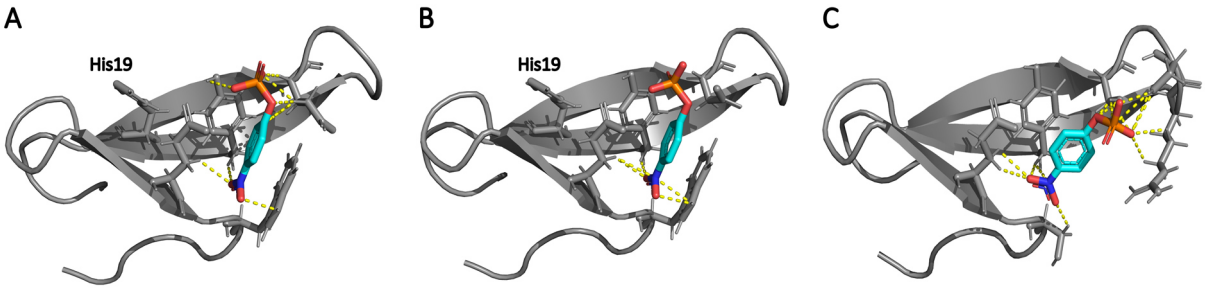
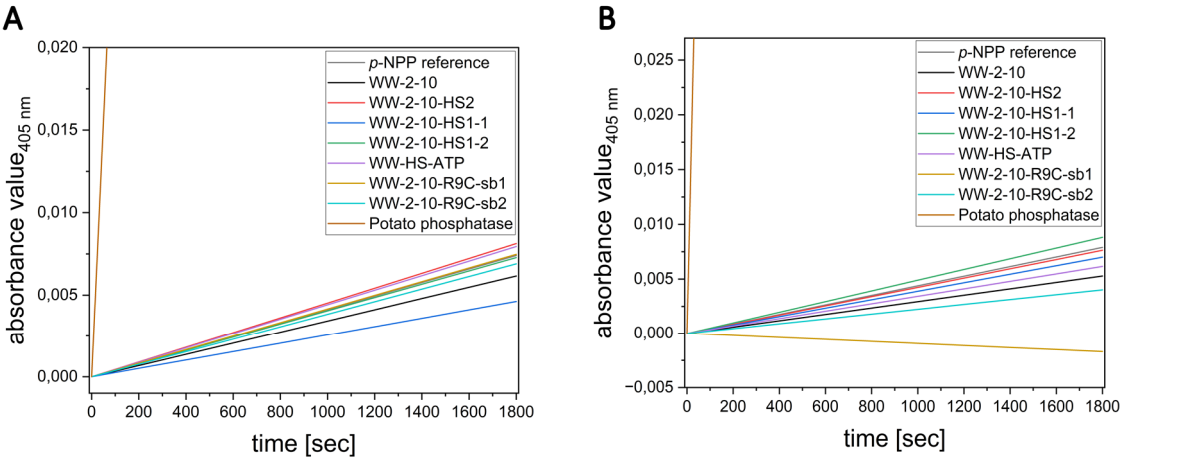


Figure 46: *In silico* docking studies of *p*-NPP with the three minorly hydrolytically active variants WW-2-10-HS2 WW-HS-ATP and WW-2-10-R9C-sb1 in acidic pH. A) Predicted docking pose of *p*-NPP in WW-2-10-HS2. **B**) Predicted docking pose of *p*-NPP in WW-HS-ATP. **C**) Predicted docking pose of *p*-NPP in WW-2-10-R9C-sb1. The His19 residue is highlighted in **A** and **B**. Residues interacting with *p*-NPP (blue: nitrophenol part, orange: phosphate moiety) via polar interactions (yellow dashed lines) are displayed as grey sticks.

The other four designed peptides revealed decreased absorbance values at 405 nm in comparison to the *p*-NPP reference sample, indicating a lack of hydrolytic activity towards *p*-NPP at pH 4.8 (cf. Figure 47 A). At pH 7.4, only one variant, respectively WW-2-10-HS1-2, exhibited hydrolytic activity towards *p*-NPP, with a k_2 value of 1.42 ± 0.11 x 10^{-3} s^{-1} (cf. Figure 47 B). This represents 0.06 % of the hydrolytic activity of the studied control phosphatase. The peptides identified as catalytically active at pH 4.8 did not exhibit any enzymatic activity at physiological pH. In future studies a potential hydrolytic activity of the peptides towards *p*-NPP should be elucidated under basic conditions with another buffer.



pH 4.8

Peptide	m_{obs} [10^{-6}]	k_{obs} [µM/s]	k_2 [10^{-3}/s]	k_2 [10^{-2}/min]
WW-2-10	3.4240 ± 0.1318	-1.9205 ± 0.0539	-1.9205 ± 0.0539	-11.5230 ± 0.3232
WW-2-10-HS2	4.5152 ± 0.1498	1.1107 ± 0.0038	1.1107 ± 0.0038	6.664 ± 0.02302
WW-2-10-HS1-1	4.4160 ± 0.1818	-4.3065 ± 0.1649	-4.3065 ± 0.1649	-25.8390 ± 0.9896
WW-2-10-HS1-2	4.1549 ± 0.1467	-0.19739 ± 0.0079	-0.1974 ± 0.0079	-1.1843 ± 0.0477
WW-HS-ATP	5.1144 ± 0.1866	0.8349 ± 0.0853	0.8349 ± 0.0853	5.0097 ± 0.5116
WW-2-10_R9C_sb1	2.5650 ± 0.0918	0.1098 ± 0.0123	0.1098 ± 0.0123	0.6590 ± 0.0738
WW-2-10_R9C_sb2	4.0443 ± 0.1483	-0.7855 ± 0.0603	-0.7855 ± 0.0603	-4.7128 ± 0.3619
Potato Phosphatase	309.5680 ± 4.4487	848.4795 ± 11.9376	848.4795 ± 11.9376	509.0877 ± 71.6254

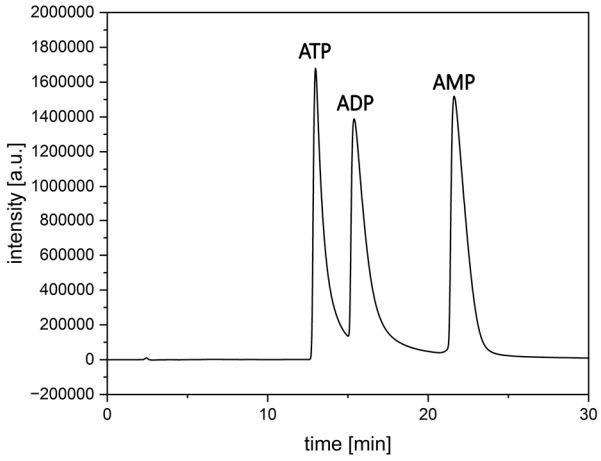
pH 7.4

Peptide	m_{obs} [10^{-6}]	k_{obs} [µM/s]	k_2 [10^{-3}/s]	k_2 [10^{-2}/min]
WW-2-10	2.9337 ± 0.11118	-4.0353 ± 0.0942	-4.0353 ± 0.0942	-24.2120 ± 0.5655
WW-2-10-HS2	4.2416 ± 0.1948	-0.4022 ± 0.1381	-0.4022 ± 0.1381	-2.4133 ± 0.8289
WW-2-10-HS1-1	3.8875 ± 0.1643	-1.3859 ± 0.0535	-1.3859 ± 0.0535	-8.3157 ± 0.32077
WW-2-10-HS1-2	4.8960 ± 0.1861	1.4157 ± 0.1140	1.4157 ± 0.1140	8.4942 ± 0.6840
WW-HS-ATP	3.4215 ± 0.1350	-2.6802 ± 0.0279	-2.6802 ± 0.0279	-16.0810 ± 0.1676
WW-2-10_R9C_sb1	-0.9416 ± 0.1271	-14.8000 ± 0.0499	-14.8000 ± 0.0499	-88.8001 ± 0.2996
WW-2-10_R9C_sb2	2.2302 ± 0.1494	-5.9895 ± 0.0121	-5.9895 ± 0.0121	-35.9368 ± 0.0725
Potato Phosphatase	916.2140 ± 1.7651	2532.8545 ± 4.5003	2532.8545 ± 4.5003	15197.1268 ± 27.0016

Figure 47: Screening of *de novo* designed ATP-binding variants for hydrolytic activity towards *p*-NPP. Measurement of nitrophenolate ion absorbance at 405 nm in peptide incubations with *p*-NPP for 30 min in PBS with **A**) pH 4.8 and **B**) pH 7.4. Absorbance values were collected in an interval of 10 sec. For determining the slope of the linear graph m_{obs}, a linear regression model was fitted to the mean value of triplicate measurement data (except for WW-2-10-R9C-sb2 (pH 4.8), WW-2-10-HS2 (pH 7.4), duplicate measurements). Based m_{obs}, the rate constant k_{obs}, the catalytical constant k_2 were determined and listed below the graphs, including the respective error values.

A second exploratory study was conducted based on a RP-HPLC method to separate and quantify the adenosine phosphates ATP, ADP, and AMP. The objective was to examine whether one of the seven peptides exhibited hydrolytic activity towards one of the adenosine phosphate molecules. In accordance with the published procedure by Juarez-Facio *et al.*, a 50 mM potassium hydrogen phosphate buffer (pH 6.8) was utilized as the mobile phase.[136] The separation of the adenosine phosphate molecules was performed on a Shimadzu HPLC system, equipped with an analytical C18 column (cf. Materials and Methods, Chapter 6.2.9 XY for detailed procedure) applying an isocratic setting for 30 min at a flow rate of 0.6 ml/min. The intensity was monitored at 254 nm. For quantification of the obtained peaks, the area under the curve (AUC) was integrated and compared with the different sample mixtures. As a reference sample, 1 mM of the incubated adenosine phosphate molecule was examined prior to the incubation of peptide-adenosine phosphate mixtures (ratio of 1:100) for study.
In order to facilitate differentiation between the three adenosine phosphate molecules, a mixture, containing 1 mM of each molecule was examined. As illustrated in Figure 48, the retention times (R_t) of ATP, with 13.0 min, and for ADP, with 15.4 min are relatively close, thus the two peaks were not baseline separated. However, the AMP peak, with a retention time of 21.6 min, was clearly distinguishable from the other two adenosine phosphate molecules. The calculated AUC in percent is presented in Figure 48 B.

A



B

Peak	Retention time (R_t) [min]	Peak area [%]
ATP	13.0	27.3
ADP	15.4	36.0
AMP	21.6	36.7

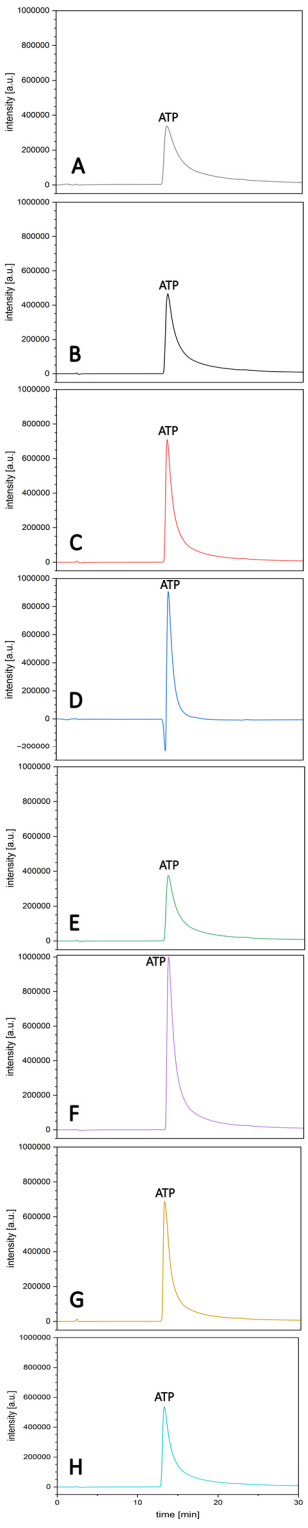
Figure 48: Chromatographic separation of an overnight incubated organophosphate mixture. A) RP-HPLC chromatogram of a 1 mM ATP, ADP, AMP overnight incubation (ratio 1:1:1) **B**) Retention times R_t and yield of the respective area under the curve (AUC) of all three organophosphate peaks. Absorbance at 254 nm was monitored for 30 min with an isocratic setting using a 50 mM KH_2PO_4 buffer (pH 6.8) and a flow rate of 0.6 ml/min. The AUC of the peaks was integrated using the *OriginPro* peak analysis software.

Following the identification of the individual retention times of ATP, ADP, and AMP, the hydrolytic activity of all seven peptides was tested towards the three adenosine phosphate molecules. The peptide-adenosine phosphate samples were incubated at 4 °C overnight (minimum of 12 h), and allowed to equilibrate at room temperature for 30 min prior to the injection of 100 µl of the sample volume. The results of the peptide incubations with 1 mM ATP are presented in Figure 49 1 A-H.
In general, a single visible peak with a broad backtailing was identified for the different peptide incubations with ATP (cf. Figure 49 1 B-H) and the ATP reference sample (Figure 49 1 A). This

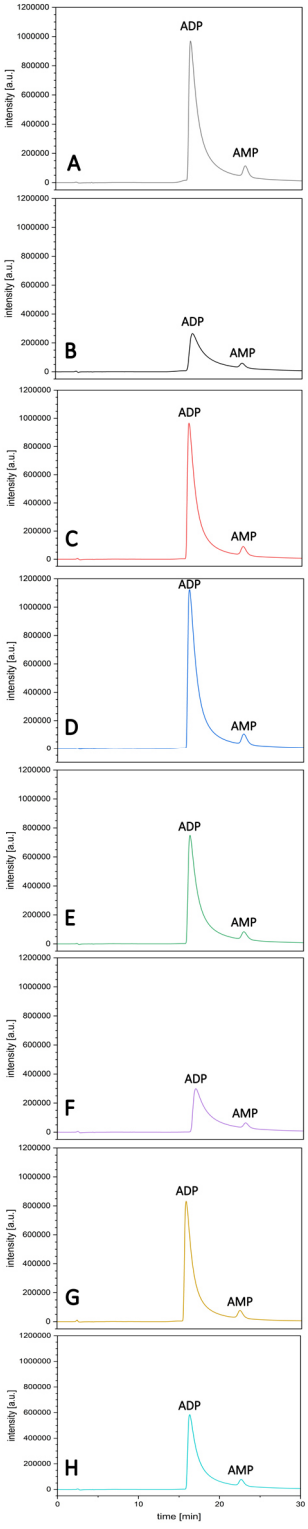
prevented the differentiation between an ATP and a potential ADP peak. In addition to the ATP peak, a signal was identified at the retention time observed for AMP in all incubations by the peak analysizer software from *OriginPro*, albeit with varying intensity (cf. Figure 49 3A). This identified peak was only slightly detectable by visual inspection. The highest yield, with 13.3 %, was found in the ATP reference sample compared to all incubations, leading to the assumption that none of the tested peptides exhibited hydrolytic activity towards ATP, at least in the described setting of the exploratory study. It seems more probable that ATP undergoes a minor hydrolysis over time, resulting in the formation of AMP and pyrophosphate during incubation. Notably, the quantity of AMP generated is significantly lower in all peptide-ATP mixture samples relative to the ATP reference sample (cf. Figure 49 3 A). This observation may be attributed to a binding interaction between ATP and the respective peptide, leading to a stabilization of the adenosine phosphate molecule.

The analysis of the ADP incubations revealed AMP peaks for all eight studied samples, in addition to the expected ADP-specific peak (cf. Figure 49 2 A-H). Although an AMP peak was observed in all samples, the yield of AMP was significantly increased in the two samples, WW-2-10 and WW-HS-ATP, by over 50 % compared to the ADP reference (cf. Figure 49 3 B), which leads to the conclusion that both peptides exhibit hydrolytic activity towards ADP. Furthermore, the incubation of WW-2-10-R9C-sb2 with ADP also revealed higher yields of AMP than the ADP reference sample. Although these findings indicate phosphatase-like properties for the peptides, further studies are necessary to confirm the results in the future. In addition, studies with AMP were conducted and only one AMP peak was identified for all samples during the isocratic HPLC run of 30 min (cf. Figure S40, Appendix). To further assess the possibility of AMPase activity, this exploratory study needs to be optimized in order to enable monitoring of the single phosphate group and, with respect to the ADPase activity, the pyrophosphate group.

1



2



3

A

Incubation sample	Peak	Retention time (R_t) [min]	Peak area [%]
1 mM ATP	ATP	13.5	85.4
	AMP	21.6	13.3
WW-2-10 + ATP	ATP	13.5	91.4
	AMP	21.8	8.6
WW-2-10-HS2 + ATP	ATP	13.5	93.7
	AMP	21.7	6.3
WW-2-10-HS1-1 + ATP	ATP	13.6	98.0
	AMP	22.5	2.0
WW-2-10-HS1-2 + ATP	ATP	13.5	90.8
	AMP	21.8	9.3
WW-HS-ATP + ATP	ATP	13.6	95.1
	AMP	22.1	4.9
WW-2-10-R9C-sb1 + ATP	ATP	13.2	94.1
	AMP	21.5	5.6
WW-2-10-R9C-sb2 + ATP	ATP	13.3	91.8
	AMP	21.5	8.2

B

Incubation sample	Peak	Retention time (R_t) [min]	Peak area [%]
1 mM ADP	ADP	16.4	87.3
	AMP	23.1	12.7
WW-2-10 + ADP	ADP	16.7	79.9
	AMP	22.7	20.0
WW-2-10-HS2 + ADP	ADP	16.1	89.2
	AMP	22.8	10.8
WW-2-10-HS1-1 + ADP	ADP	16.1	89.5
	AMP	22.7	10.5
WW-2-10-HS1-2 + ADP	ADP	16.2	87.5
	AMP	22.7	12.5
WW-HS-ATP + ADP	ADP	16.9	80.2
	AMP	22.9	19.4
WW-2-10-R9C-sb1 +ADP	ADP	15.8	89.2
	AMP	22.4	10.7
WW-2-10-R9C-sb2 + ADP	ADP	16.3	85.8
	AMP	22.7	14.2

Figure 49: RP-HPLC chromatograms of peptide-adenosine phosphate incubations. 1 A) to **H**) Chromatograms of peptide-ATP incubations (1:10 ratio). **1 A**) ATP reference sample without peptide. ATP with **1 B**) WW-2-10, **1 C**) WW-2-10-HS2, **1 D**) WW-2-10-HS1-1, **1 E**) WW-2-10-HS1-2, **1 F**) WW-HS-ATP, **1 G**) WW-2-10-R9C-sb1, **1 H**) WW-2-10-R9C-sb2. **2 A**) to **H**) Chromatograms of peptide-ADP incubations (1:10 ratio). **2 A**) ADP reference sample without peptide. ADP with **2 B**) WW-2-10, **2 C**) WW-2-10-HS2, **2 D**) WW-2-10-HS1-1, **2 E**) WW-2-10-HS1-2, **2 F**) WW-HS-ATP, **2 G**) WW-2-10-R9C-sb1, **2 H**) WW-2-10-R9C-sb2. **3 A**) Retention time R_t of the peaks of ATP incubations and the peak yield in percent. **3 B**) Retention time R_t of the peaks of ADP incubations and the peak yield in percent. Peak yields were determined by integrating the AUC of the peaks using the *OriginPro*

Peakanalysis software. Following settings were used: an injection of 100 µl of the sample was studied by monitoring the absorbance at 254 nm for 30 min with an isocratic setting using a 50 mM KH_2PO_4 buffer (pH 6.8) and a flow rate of 0.6 ml/min.

In conclusion, the two exploratory studies indicated low phosphatase-like activity for some peptide candidates. In both experiments, WW-HS-ATP demonstrated hydrolytic activity towards *p*-NPP and ADP as well. The colorimetric assay revealed that other peptides, such as WW-2-10-HS2, WW-2-10-R9C-sb1, and WW-2-10-HS1-2, were capable of converting *p*-NPP to the colorful nitrophenolate ion above background signal, provided that the selected pH value was appropriate. The results of the RP-HPLC-based assay, which was performed to study and quantify possible degradation products of peptide with ATP, ADP, and AMP, revealed that the amount of AMP present in the WW-2-10 and WW-2-10-R9C-sb2 incubations with ADP was greater than that observed in the reference sample. This leads to the assumption that these two peptides also possess low ADPase-like activity. It should be noted that the presented results in this section were exploratory studies, indicating minimal catalytic activity for some peptide variants. To verify the phosphatase-like activities for the aforementioned candidates, further intensive investigations employing optimized methodologies are required in the future.

5 Summary and Outlook

5.1 Summary

In the context of this work, the WW domain, a β-sheet based, protein-protein interaction module that is highly abundant in nature[61,77,80,89,90,137], was analyzed for its sequence-structure-stability relationships, which were subsequently used to design thermostable WW domain core scaffolds. These thermostable core scaffolds were functionalized by grafting different binding motifs onto the surface, leading to various WW-domain based minireceptors (cf. Figure 50).

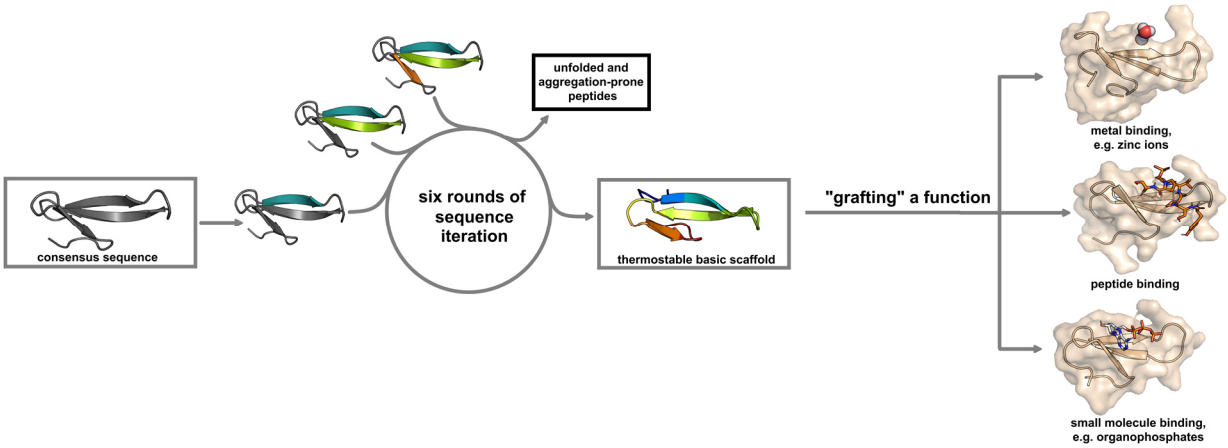


Figure 50: Concept of the iterative sequence analysis to generate functional thermostable WW-domain peptides. This figure is adapted from Lindner *et al.*[94], with permission from © 2024 the Americal Chemical Society.

To address the initial research objective of identifying sequence-structure-stability relationships in WW domains, a sequence alignment of 90 native WW domains was performed. In accordance with the literature, highly conserved residues, such as Leu2, Pro3, Trp6, Tyr/Phe19 and Pro32, which are part of the hydrophobic core, and contribute to the peptide's overall structural stability[6,76,77], were preserved in the design of the consensus sequence WW-CS. Based on WW-CS, six rounds of iterative sequence analysis were performed, resulting in the analysis of 38 WW domain variants. Except for the highly conserved N-terminus, the sequence iterations encompassed all other structural segments of WW-CS, including BS1, LP1, BS2, LP2, BS3, and CT. The iterative sequence analysis resulted in the identification of sequence-structure-stability-relationships (SSSRs) of WW-domain based peptides (cf. Figure 51).

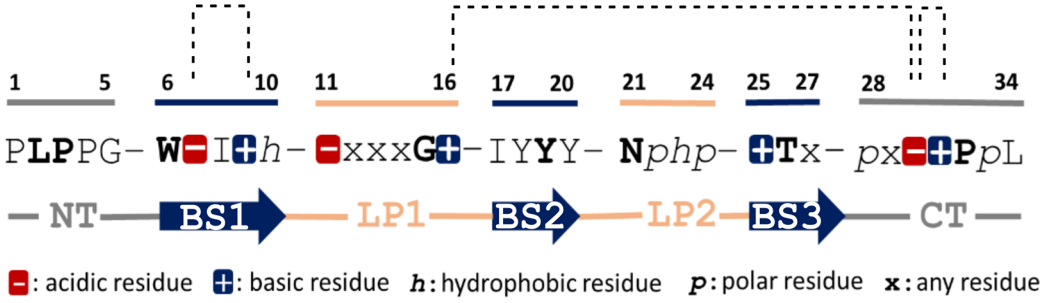


Figure 51: Identified sequence-structure-stability relationships in WW domain-based peptides. Residues displayed in bold were identified as highly conserved, as they provide structural stability. Plus and minus symbols

represent residues with positive and negative charge, which form electrostatic interactions as observed in NMR spectroscopic studies. This figure is adapted from Lindner *et al.*[94], with permission from © 2024 the Americal Chemical Society.

Three of the most important findings that have led to the development of highly thermostable WW-domain core scaffolds and sequences are summarized in Figure 51. One of the structure-stabilizing sequence features is the extension of the hydrophobic core by introducing the hydrobobic, high β-propensity exhibiting Ile residue at the positions 8 in BS1 and position 23 in LP2. Moreover, the presence of oppositely charged residues at positions 7 and 9 in BS1 and positions 16 in LP1 and 30 in CT, as well as the introduction of the aromatic Tyr residue at position 19 of BS2, have a highly structure-stabilizing effect. Three highly thermostable WW domains (WW-HS) were designed (cf. Figure 52 A). CD spectroscopic and thermal denaturation analysis revealed that all three WW-HS peptides exhibited WW domain structures (cf. Figure 52 B), and very high melting temperatures, ranging from 88.5 to 93.5 °C (cf. Figure 52 C). The characteristic structure of the WW domain with three antiparallel β-sheets of WW-HS2$_{OH}$ (cf. Figure 52 D), was confirmed by multi-dimensional high-resolution NMR spectroscopy performed in collaboration with Michael Kovermann from the University of Konstanz.

A

WW domain	Sequence ----NT---BS1---LP1--BS2--LP2--BS3---CT-- 1 - 10 - 20 - - 30	T_m / °C
WW-HS-1	Ac-PLPPG-WEIRI-DSGSGS-IYYY-NSIT-KTT-TAERPRL-NH_2	93.5 ± 0.5
WW-HS-2	Ac-PLPPG-WEIRI-ESGSGS-IYYY-NSIT-KTT-TWQRPRL-NH_2	89.0 ± 0.5
WW-HS-2$_{OH}$	H-PLPPG-WEIRI-ESGSGS-IYYY-NSIT-KTT-TWERP -OH	88.5 ± 0.5

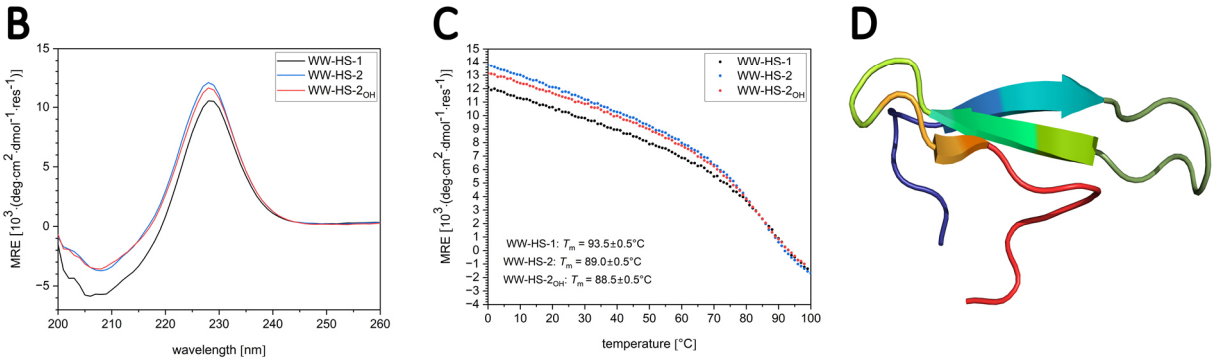


Figure 52: Sequence, CD spectra, thermal denaturation curves and NMR structure of the WW-HS series. A) Sequences of the WW-HS peptides are shown with modifications between the variants displayed in red. **B**) Mean values of the CD spectra are presented. **C**) Mean values of the thermal denaturation curves and the respective melting temperatures are presented. **D**) NMR structure of WW-HS-2$_{OH}$ (pdb code: 8qtq) is shown. NMR experiments were performed in collaboration with Michael Kovermann and coworkers. This figure was adapted with permission from Lindner *et al.*[94], © 2024 the American Chemical Society.

Following the design of highly thermostable WW-HS core scaffolds, the subsequent objective was to graft the binding motifs of the three WW domain groups onto an WW-HS core scaffold to obtain peptide candidates exhibiting native-like functions, as well as thermal stability. Based on the results of *Robetta*[115] alanine scanning, ligand-binding sites of all groups were identified and transferred onto the WW-HS-2 core scaffold. The structural and thermal stability of all designed group-specific WW-HS peptides was evaluated using CD spectroscopy (cf. Figure 53 A and B). All group-specific WW-HS peptides exhibited stable folding and high

melting temperatures in comparison to the studied native reference peptides. Investigations on binding affinity demonstrated that the designed candidates possess comparable or enhanced binding affinity as their native counterparts (cf. Figure 53 C and D).

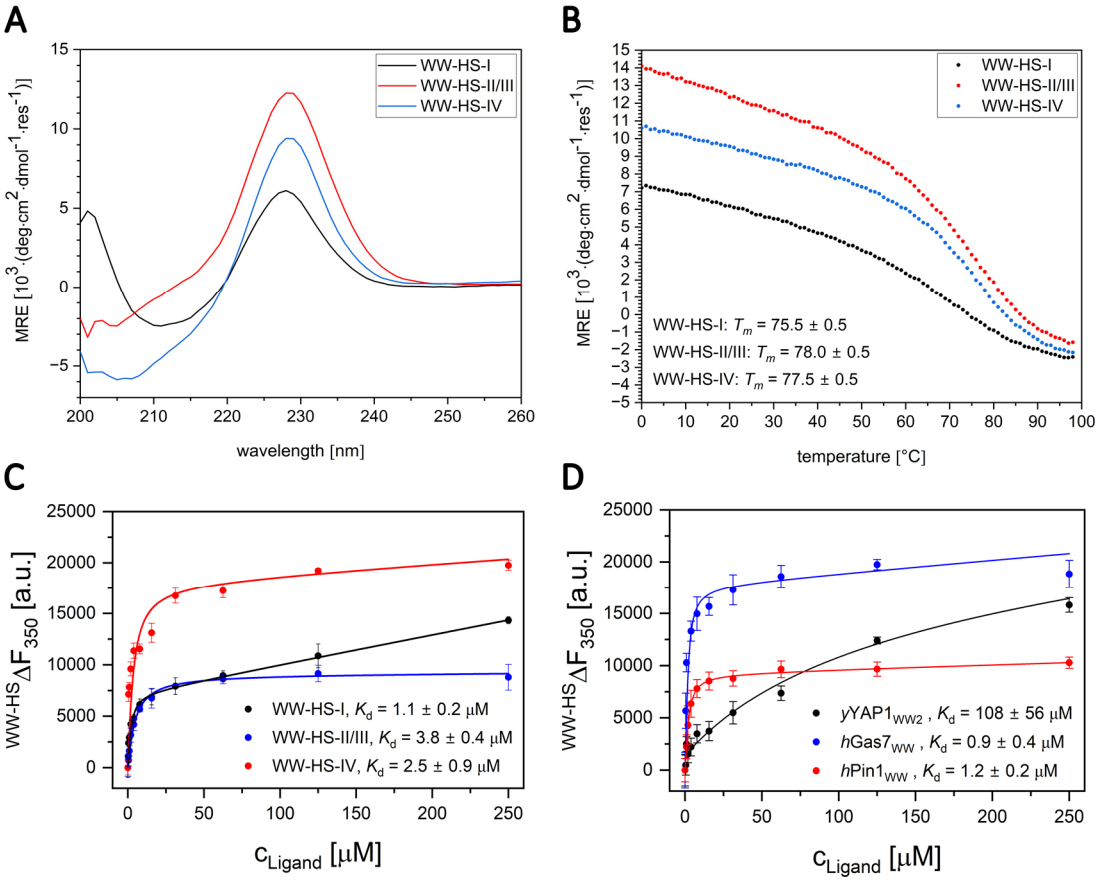


Figure 53: Designed thermostable group-specific WW-domain based peptides. A) CD spectra and **B**) thermal denaturation cruves are represented as mean of triplicates. **C**) Binding saturation curve data of group-specific WW-HS peptides and **D**) the native reference peptides obtained after ligand titration. The change of intrinsic tryptophan fluorescence was measured at 350 nm. Conditions: 2 µM peptide, PBS buffer, pH 7.5. This figure was adapted with permission from Lindner *et al.*[94], © 2024 the American Chemical Society.

Another objective was to utilize the design approach of grafting various ligand-binding sites onto a WW-HS core scaffold to generate miniproteins with non-native binding properties, such as metal and organophosphate binding. To engineer a metal-binding thermostable peptide, the rationale of transferring the His_3 site to a WW domain peptide scaffold, as described by Pham *et al.* in the design of the Zn(II)-binding WW-CA peptide[84], was applied to generate WW-HS-CA. CD thermal denaturation experiments revealed a higher melting temperature of the WW-HS-CA-Zn(II) complex than the peptide in its apo state (cf. Figure 54 B). Competitive titration experiments of WW-HS-CA with Zn(II) revealed a Zn(II) binding affinity in the submicromolar range (cf. Figure 54 E), which was improved by at least a factor of five compared to previously reported peptide designs by Pham *et al.*[84,120] Apart from the metal-binding WW-HS-CA, thermostable organophosphate-binding WW domains were designed by grafting the relevant PC and ATP binding motifs onto the WW-HS-1 core scaffold. Thermal denaturation with WW-HS-PC revealed an increase in T_m of approximately 36 K compared to WW-1-8 from Neitz *et al.*[85]. Binding studies of WW-HS-PC with PC resulted in a micromolar binding affinity (K_d = 3.0 ± 1.4 µM). Thermal denaturation of WW-HS-ATP displayed a good T_m of 58.8 °C, while WW-2-10 of Neitz *et al.* exhibited a lower thermostability

with a T_m of 16.3 °C. Binding studies of WW-HS-ATP and ATP revealed a high binding affinity of WW-HS-ATP to ATP (K_d = 4.9 ± 0.9 µM). The design of thermostable WW domain metal- and organophosphate minireceptors demonstrates the utility of thermostable core scaffolds for the design of miniproteins with non-native functions.

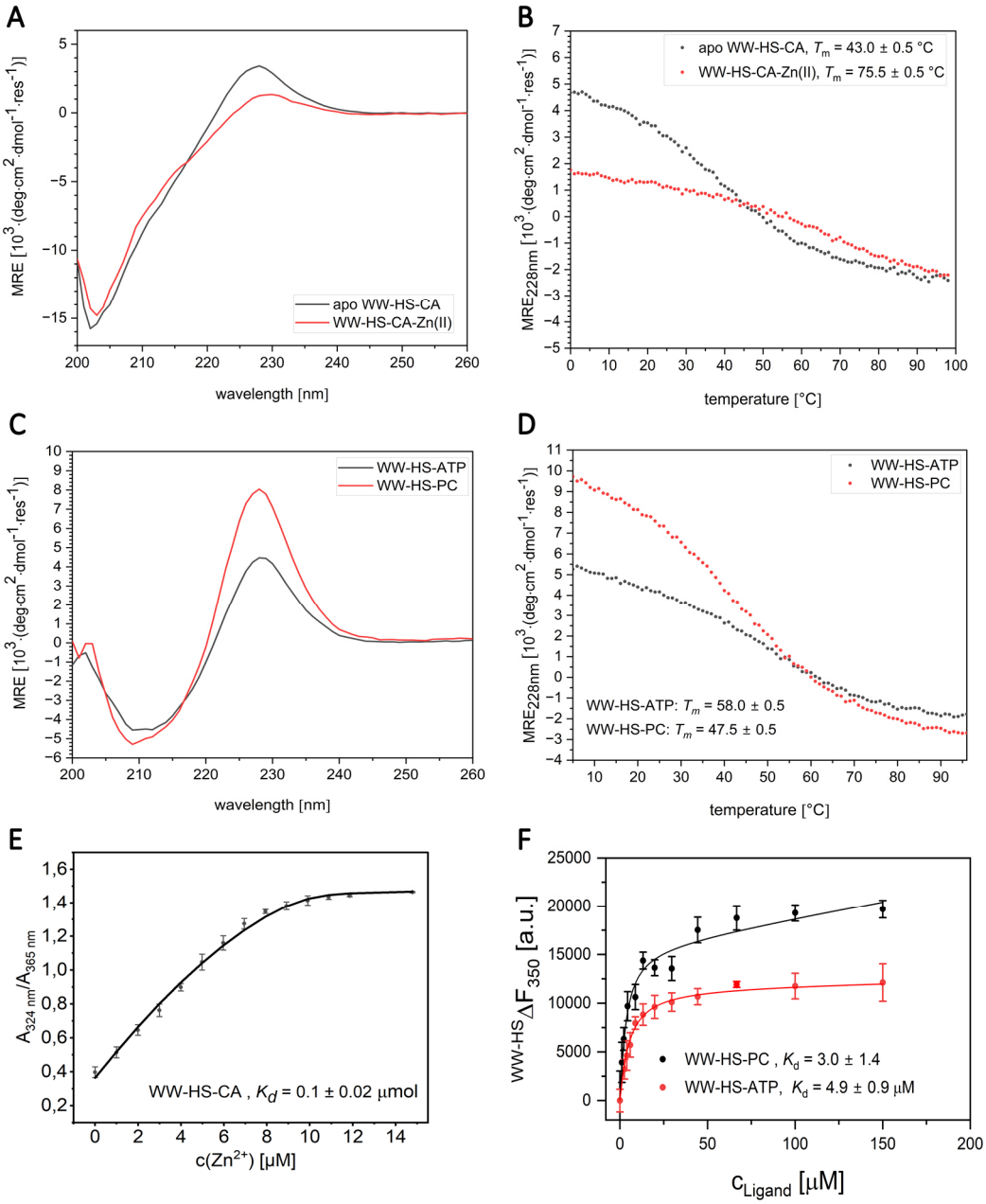


Figure 54: *De novo* designed functional and thermostable WW domain peptide variants. A) CD spectra of apo and holo WW-HS-CA with Zn(II). **B**) Thermal denaturation curves of apo and holo WW-HS-CA with Zn(II). **C**) CD spectra of WW-HS-ATP and WW-HS-PC. **D**) Thermal denaturation curves of WW-HS-ATP and WW-HS-PC. **E**) Saturation binding curve of WW-HS-CA upon binding to Zn(II). **F**) Saturation binding curves of WW-HS-ATP with ATP and WW-HS-PC with PC upon ligand titration. This figure was minorly adapted from Lindner *et al*.[94], with permission from © 2024 the American Chemical Society.

In addition to the designed WW-HS-ATP peptide, other WW domains with organophosphate binding activity were designed in a top-down approach from *h*Pin1$_{WW}$ and studied for their ATP, ADP and AMP binding properties. All peptides exhibited similar binding affinities to ATP in the low micromolar range. The influence of overnight incubations with or without the

addition of either ATP or ATP and 1 mM Mg^{2+} on the structural and thermal stability of WW-2-10 and the three most thermostable peptide variants were examined. In conclusion, incubations with and without the addition of ATP or ATP mixtures with Mg^{2+} did not significantly influence the structural or thermal stability of the tested peptides. In addition, the binding affinities of the peptides to ATP, ADP and AMP, were examined. The presence of Mg^{2+} in the peptide incubation mixtures, led to different results, for some peptides the binding affinity to ATP remained unaltered, some exhibited increased, and others decreased binding affinity to ATP. The results of the binding studies with the peptides to ADP, and AMP, revealed micromolar binding affinities for all tested peptides. Two final exploratory studies were conducted to investigate the potential phosphatase-like activity of the WW domain peptides. However, some of the studied peptides showed at best minimal catalytic activity.

5.2 Outlook

The designed highly thermostable WW domain core scaffolds, presented in this work, have the potential to be functionalized with native and non-native binding properties, thereby providing a platform for the future design of functional β-sheet based miniproteins with further receptor, sensor and catalytic functions. The knowledge gained from the identified sequence-structure-stability-relationships that led to the design of the thermostable WW domain core scaffolds, will be highly valuable for future miniprotein design, as it provides information about which amino acid residues can be modified to install specific binding motifs for a desired function, while maintaining thermostability.

One potential objective for future research can be the design of thermostable WW domain peptides with additional metal-binding properties, such as Mg^{2+} or Mn^{2+} binding. In the literature, Asp and Glu residues have been predominantly found in protein sites that coordinate Mg^{2+} and Mn^{2+} binding, both independently and in the presence of additional ligands, including ATP and its derivatives.[138] Moreover, it has been demonstrated that these two divalent cations can be interchangeable in the active centers of certain proteins, such as phosphatases and kinases.[138,139] These findings could serve as a valuable source of information that can be further utilized in the design of thermostable Mg^{2+} or Mn^{2+}-binding WW domain peptides.

In addition to their metal-binding properties, thermostable WW domain core scaffolds can be functionalized with carbohydrate-binding motifs. These miniproteins can be employed in the investigation of carbohydrate-protein interactions and in the further design of miniproteins with diverse biological functions. In nature, carbohydrate-protein interactions play an important role in regulating cellular processes[140], including cell-cell interaction, cell adhesion[141], and immune response[142]. This illustrates the relevance of designing thermostable miniproteins with tailored carbohydrate-binding functions in the future.

A further objective for future research could be to design of thermostable WW domain peptides with specific binding properties that can be switched by environmental factors, such as a change in pH or the presence or absence of a ligand. In this way, thermostable miniproteins could function as advanced sensory systems with broad future applications.

Furthermore, the prospective investigations of the phosphatase-like activities of the adenosine phosphate molecule-binding WW domain peptides will be the subject of detailed future studies, which will facilitate the natural-like β-sheet-based design of thermostable minienzymes. A long-term objective may be to develop WW domain peptide scaffolds that are catalytically active and mediate specific reactions with broad substrate specificity, while retaining the advantages of enzyme catalysis, such as stereo-selectivity and mild physiological reaction conditions. In general, thermostable minienzymes with high specificity would also be applicable in various industrial fields, including the pharmaceutical, chemical, detergent, and food industries.[39]

6 Materials and Methods

6.1 Materials

6.1.1 Amino acids, chemicals and resins for SPPS

For microwave-assisted automated solid-phase peptide synthesis with the *Liberty Blue CEM* peptide synthesizer standard Fmoc-protected amino acids were acquired from *Iris Biotech GmbH* (Marktredwitz, Germany). The special amino acid Fmoc-Ser(HPO_3Bzl)-OH (98 % purity) was purchased from *Th. Geyer GmbH & Co KG* (Renningen, Germany), Fmoc-Nle-OH (98 % purity) was supplied by *Carbolution Chemicals GmbH* (St. Ingbert, Germany) and Fmoc-Asp(OMpe)-OH by *BLD Pharmatech GmbH* (Reinbek, Germany).
For peptide amides, H-Rink amide ChemMatrix® (loading density: 0.40-0.60 mmol g^{-1}) from *Sigma Aldrich* (Taufkirchen, Germany) was used until the production was discontinued in October 2022. H-Rink amide 4-methylbenzylhydrylamine (MBHA) resin (loading density: 0.67 mmol g^{-1}) was then utilized as substitute, acquired from *Iris Biotech GmbH* (Marktredwitz, Germany). Peptide acids were synthesized either on H-Ser or H-Gly-preloaded HMPB resin (loading density: 0.30-0.65 mmol g^{-1}) for the $hPin1_{WW}$ split peptides and for the WW-HS peptides on H-Leu-preloaded ChemMatrix® resin (loading density: 0.30-0.65 mmol g^{-1}), all purchased from *Sigma Aldrich* (Taufkirchen, Germany).
The chemicals for SPPS were acquired from different suppliers: *N,N*-diisopropylcarbodiime (DIC) was purchased from *Iris Biotech GmbH* (Marktredwitz, Germany), whereas Oxyma, *N,N*-diisopropylamine (DIPEA), piperidine, trifluoroacetic acid (TFA, HPLC grade) and triisopropylsilane (TIPS) were supplied by *Merck KGaA* (Darmstadt, Germany). *N,N*-Dimethylformamide (DMF, peptide synthesis grade) and dichloromethane (DCM, peptide synthesis grade) were obtained from *Fisher Scientific* (Loughborough, United Kingdom).

6.1.2 Buffers and solutions

For all experiments and buffer preparations, ultrapure water, purified with a *Sartorius arium® mini* lab water system and degassed by sonication in an ultrasonic bath (*Ultrasonic Cleaner USC-T, VWR*) for at least 10 min was used. The pH of one-fold concentrated buffers was measured with a *HANNA instruments HI 2210* pH-meter, equipped with a *HI 1330* pH electrode and set to the respective value by using 1 M HCl or 1 M NaOH solution. All buffers were thorougly mixed during pH adjustment with a stirring bar on *a Heidolph® Hei-Plate Mix'n'Heat* magnetic plate at 550 -600 rpm mixing speed to ensure a homogenous distribution.

Biophysical experiments, unless stated otherwise, were performed with phosphate-buffered saline (PBS) of either pH 7.4 (8.2 mM Na_2HPO_4, 1.8 mM K_2HPO_4, 137 mM NaCl, 2.7 mM KCl) or 7.5 (10 mM Na_2HPO_4, 1.8 mM K_2HPO_4, 137 mM NaCl, 2.7 mM KCl). Both PBS buffers were prepared as tenfold stock solutions in graduated flasks, filtered with a cellulose acetate filter (pore size: 0.2 µM) and degassed, before being aliquoted and stored at -20 °C. All samples were dissolved in either 10xPBS and further diluted with ultrapure water or directly in pre-diluted 1xPBS.

For the enzymatic activity screening assay *para*-nitrophenylphosphate (*p*-NPP), PBS with additional pH values of 4.8 and 9.8 were used. After defrosting the 10xPBS (pH 7.4), it was diluted in degassed ultrapure water. After adjusting the pH-value the remaining amount of water for the final dilution was added and the 1xPBS buffers were stored at room temperature until usage.

WW-HS-CA was measured in a 10 mM 3-(*N*-morpholino) propanesulfonic acid (MOPS)-buffered saline, containing 150 mM NaCl with a pH of 7.2. For sample preparation a five-fold stock solution (stored at 4 °C) was diluted with the respective amount of degassed, ultrapure water.

For the HPLC-based enzymatic activity assay of peptide samples towards selected organophosphate molecules (ATP, ADP and AMP), a 50 mM K_2HPO_4 buffer (pH 6.8) was prepared in degassed ultrapure water and stored at room temperature.

6.1.3 Chemicals

All reagents and solvents, if not otherwise stated (see section 6.1.1), were at least *pro analysis* grade quality and were acquired from the following companies: *Carbolution Chemicals GmbH* (St. Ingbert, Germany), *Carl Roth GmbH + Co. KG* (Karlsruhe, Germany), *Honeywell* (Seelze, Germany), *neoFroxx GmbH* (Einhausen, Germany), *Th. Geyer GmbH & Co. KG* (Renningen, Germany) and *VWR International* (Fontenay-sous-Bois, France).

6.1.4 Expendable supplies

Special expendable supplies, such as low-binding eppendorf tubes (*Fisher Scientific International*) as well as low-binding pipette tips (*VWR International*) were exclusively used for the sample preparation of peptides and ligands that were further investigated in binding experiments, in order to prevent the absorbance of peptides to the plastic. Fluorescence titration experiments were performed in non-binding, black, F-bottom-shaped 96-well plates from *Greiner Bio-One*.

6.1.5 Synthesized peptides

All peptides presented in this work, were synthesized on the microwave-assisted automated peptide synthesizer *Liberty Blue* from *CEM* (details see Chapter 6.2.1). The amino acid sequences for the different project-related peptides and their respective masses are summarized in Tables 1-10.

Table 1: **Sequences of WW domain-based basic screening peptides.** The peptide termini are displayed as "NT" for N-terminus and "CT" for C-terminus. Beta-sheets are indicated as BS and loop regions as LP. All peptide sequences have already been published in Lindner *et al.* 2024, except for WW-LP1-8. This list was minorly adapted and reused from Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Peptide name	Sequence	$[M+H]^+$ (calc.)	$[M+H]^+$ (found)
	--------NT------BS1------LP1------BS2-----LP2------BS3--------CT		
WW-CS	Ac-PLPPG-WEERK-DRPGGR-VYYV -NHNT- RTT -TWERPRL-NH_2	4175.1	4174.1
WW-BS1-1	Ac-PLPPG-**WEIRI**-DRPGGR-IYFY -NHNT- KTT -TWERPRL-NH_2	4178.2	4176.6
WW-BS1-2	Ac-PLPPG-**WEIRL**-DRPGGR-IYFY -NHNT- KTT -TWERPRL-NH_2	4178.2	4177.4
WW-BS1-3	Ac-PLPPG-**WEIRT**-DRPGGR-IYFY -NHNT- KTT -TWERPRL-NH_2	4166.1	4165.1
WW-BS1-4	Ac-PLPPG-**WEIKI**-DRPGGR-IYFY -NHNT- KTT -TWERPRL-NH_2	4150.2	4149.0
WW-BS1-5	Ac-PLPPG-**WEIEI**-DRPGGR-IYFY -NHNT- KTT -TWERPRL-NH_2	4151.1	4150.4
WW-BS1-6	Ac-PLPPG-**WKIKI**-DRPGGR-IYFY -NHNT- KTT -TWERPRL-NH_2	4149.2	4148.0
WW-BS1-7	Ac-PLPPG-**WQIQI**-DRPGGR-IYFY -NHNT- KTT -TWERPRL-NH_2	4149.2	4148.1
WW-LP1-1	Ac-PLPPG-WEIRI-**SGSGSG**-IYFY -NHNT- KTT -TWERPRL-NH_2	3972.0	3971.6
WW-LP1-2	Ac-PLPPG-WEIRI-**DGSGSK**-IYFY -NHNT- KTT -TWERPRL-NH_2	4071.1	4070.1
WW-LP1-3	Ac-PLPPG-WEIRI-**DGSGSR**-IYFY -NHNT- KTT -TWERPRL-NH_2	4099.1	4098.2
WW-LP1-4	Ac-PLPPG-WEIRI-**EGSGSR**-IYFY -NHNT- KTT -TWERPRL-NH_2	4113.1	4112.4
WW-LP1-5	Ac-PLPPG-WEIRI-**GSGSGS**-IYFY -NHNT- KTT -TWERPRL-NH_2	3972.0	3971.0
WW-LP1-6	Ac-PLPPG-WEIRI-**DSGSGR**-IYFY -NHNT- KTT -TWERPRL-NH_2	4099.1	4098.0
WW-LP1-7	Ac-PLPPG-WEIRI-**ESGSGR**-IYFY -NHNT- KTT -TWERPRL-NH_2	4113.1	4113.2
WW-LP1-8	Ac-PLPPG-WEIRI-**QSGSGQ**-IYFY -NHNT- KTT -TWERPRL-NH_2	4084.1	4084.0
WW-BS2-1	Ac-PLPPG-WEIRI-QSGSGQ-**IYYY** -NHNT- KTT -TWERPRL-NH_2	3988.0	3987.0
WW-BS2-2	Ac-PLPPG-WEIRI-GSGSGS-**ISYY** -NHNT- KTT -TWERPRL-NH_2	3912.0	3911.6
WW-BS2-3	Ac-PLPPG-WEIRI-GSGSGS-**IFFY** -NHNT- KTT -TWERPRL-NH_2	3956.0	3954.8
WW-BS2-4	Ac-PLPPG-WEIRI-GSGSGS-**ISFY** -NHNT- KTT -TWERPRL-NH_2	3896.0	3895.5
WW-LP2-1	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NSGS**- KTT -TWERPRL-NH_2	3867.0	3865.7
WW-LP2-2	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NEGS**- KTT -TWERPRL-NH_2	3909.0	3907.9
WW-LP2-3	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NSNS**- KTT -TWERPRL-NH_2	3924.0	3923.2
WW-LP2-4	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NSNT**- KTT -TWERPRL-NH_2	3938.0	3937.7
WW-LP2-5	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NSIT**- KTT -TWERPRL-NH_2	3937.0	3936.1
WW-LP2-6	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NSLT**- KTT -TWERPRL-NH_2	3937.0	3936.0
WW-LP2-7	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NGSG**- KTT -TWERPRL-NH_2	3836.9	3837.3
WW-LP2-8	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NGNG**- KTT -TWERPRL-NH_2	3864.0	3862.9
WW-LP2-9	Ac-PLPPG-WEIRI-GSGSGS-IYYY -**NGIT**- KTT -TWERPRL-NH_2	3907.0	3906.3
WW-BS3-1	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- **KTS** -TWERPRL-NH_2	3923.0	3921.8
WW-BS3-2	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- **KTN** -TWERPRL-NH_2	3950.0	3948.8
WW-BS3-3	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- **KTA** -TWERPRL-NH_2	3907.0	3906.3
WW-BS3-4	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- **RTT** -TWERPRL-NH_2	3965.0	3964.0
WW-BS3-5	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- **RTS** -TWERPRL-NH_2	3951.0	3950.1
WW-BS3-6	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- **RTA** -TWERPRL-NH_2	3935.0	3933.9
WW-BS3-7	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- **RIT** -TWERPRL-NH_2	3977.1	3977.3
WW-CT-1	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- KTT -**TAERPRL**-NH_2	3822.0	3821.2
WW-CT-2	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- KTT -**TWQRPRL**-NH_2	3936.0	3935.0
WW-CT-3	Ac-PLPPG-WEIRI-GSGSGS-IYYY -NSIT- KTT -**TWERP** -NH_2	3667.8	3666.8

Table 2: Sequences of WW-domain based thermostable variants. The peptide termini are displayed as "NT" for N-terminus and "CT" for C-terminus. Beta-sheets are indicated as BS and loop regions as LP. All peptide sequences have already been published in Lindner *et al.* 2024. This list was minorly adapted and reused from Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Peptide name	Sequence --------NT------BS1------LP1-------BS2---LP2----BS3------CT	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
WW-HS-1	Ac-PLPPG-WEIRI-DSGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2	4064.1	4062.9
WW-HS-2	Ac-PLPPG-WEIRI-ESGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2	4078.1	4077.0
WW-HS-2$_{OH}$	H -PLPPG-WEIRI-ESGSGR-IYYY-NSIT-KTT-TWERPRL-OH	4037.1	4036.1
WW-HS-3	Ac-PLPPG-WEQRI-ESGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2	4093.1	4092.1
WW-HS-4	Ac-PLPPG-WEQRT-ESGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2	4081.1	4080.6
WW-HS-5	Ac-PLPPG-WEERK-ESGSGR-IYYY-NSIT-KTT-TWERPRL-NH_2	4109.1	4108.1

Table 3: Sequences of WW-domain group-specific thermostable variants. The peptide termini are displayed as "NT" for N-terminus and "CT" for C-terminus. Beta-sheets are indicated as BS and loop regions as LP. All peptide sequences have already been published in Lindner *et al.* 2024. This list was minorly adapted and reused from Lindner *et al*.[94], with permission from ©2024 American Chemical Society.

Peptide name	Sequence --------NT-------BS1------LP1-------BS2---LP2----BS3------CT	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
WW-HS-I	Ac-PLPPG--WEIRI-ESQSGR-IYYI-NHIT-KTT-TWERPRL--NH_2	4149.2	4148.2
WW-HS-II/III	Ac-PLPPG--WTIYI-E-ANGR-TYYY-NSIT-KET-TWERPRL--NH_2	4027.0	4026.3
WW-HS-IV	Ac-PLPPG--WEIRI-SRSSGR-IYYY-NSIT-KTS-QWERPRL--NH_2	4148.2	4146.9

Table 4: Sequences of WW-domain group-specific reference peptides. The peptide termini are displayed as "NT" for N-terminus and "CT" for C-terminus. Beta-sheets are indicated as BS and loop regions as LP. All peptide sequences have already been published in Lindner *et al.* 2024, except for WW-LP1-8. This list was minorly adapted and reused from Lindner *et al*.[94] with permission from ©2024 American Chemical Society.

Peptide name	Sequence --------NT-------BS1-------LP1------BS2---LP2----BS3-----CT	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
***y*YAP1$_{WW2}$**	Ac-GPLPEG-WEQAI--TPEGE-IYYI-NHKN-KTT-SWLDPRL--NH_2	3994.0	3993.2
***h*Gas7$_{WW}$**	Ac-VILPPG-WQSYL--SPQGR-RYYV-NTTT-NET-TWERPSS--NH_2	4025.0	4024.1
***h*Pin1$_{WW}$**	Ac-KLPPG--WEKRM-SRSSGR-VYYF-NHIT-NAS-QWERPSG--NH_2	4063.0	4062.1

Table 6: Sequences of WW-domain group-specific peptide ligands. All peptide sequences have already been published in Lindner *et al.* 2024. This list was minorly adapted and reused from Lindner *et a*l.[94] with permission from ©2024 American Chemical Society.

Peptide name	Sequence	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
Ligand WW-I	Ac-GTPPPPYTVG-NH_2	1026.5	1048.7 [M + Na+]
Ligand WW-II/III	Ac-GSTAPPLPR-NH_2	936.5	936.6
Ligand WW-IV	Ac-YSPTpSPS-NH_2	859.3	857.3

Table 7: Sequences of *h*Pin1$_{WW}$-based split peptides. All peptide sequences have already been published in Neitz *et al.* 2022[85] and are reused with permission from © 2022 the Royal Society of Chemistry.

Peptide name	Sequence	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
B$_{CC}$-WW-1	Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMSRSSG-OH	5197.1	5197.9
B$_{CC}$-WW-2	Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMSRSRG-OH	5266.2	5264.5
B$_{CC}$-WW-3	Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKKMSKSKG-OH	5182.1	5180.8
B$_{CC}$-WW-4	Ac-AQLKKKLQANKKELAQLKKKLQALKKKLAQ-KLPPGWEKRMSHTSG-OH	5192.1	5190.1
A$_{CC}$-WW-5	H-RVYYFNHITNASQWERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-NH$_2$	5769.0	5769.3
A$_{CC}$-WW-6	H-RVYYFNHITNAEQWERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-NH$_2$	5811.0	5809.5
A$_{CC}$-WW-7	H-RVYYFNHITNASQKERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-NH$_2$	5711.0	5710.5
A$_{CC}$-WW-8	H-RVYYSNHITNASQSERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-NH$_2$	5609.9	5609.7
A$_{CC}$-WW-9	H-RVYYFNHITNASQYERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-NH$_2$	5746.0	5745.1
A$_{CC}$-WW-10	H-RVYYHNHITNAEQSERPS-GAQLEKELQALEKKLAQLEKENQALEKELAQ-NH$_2$	5701.0	5700.4

Table 8: Sequences of *h*Pin1$_{WW}$-based split combined peptides. All peptide sequences have already been published in Neitz *et al.* 2022[85] and are reused with permission from © 2022 the Royal Society of Chemistry.

Peptide name	Sequence	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
WW-2-10	H-KLPPGWEKRMSRSRGRVYYHNHITNAEQSERPSG-OH	4023.0	4023.0
WW-1-8	H-KLPPGWEKRMSRSSGRVYYSNHITNASQSERPSG-OH	3862.9	3862.8

Table 9: Sequences of further designed ATP-binding peptides based on the WW-2-10 peptide.

Peptide name	Sequence	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
	------NT-------BS1---------LP1------BS2---LP2----BS3-----CT		
WW-2-10-HS2	H-PLPPG-WEIRI -ERSRGR-IYYH-NHIT-NTE-TWERPRL-OH	4243.2	4241.3
WW-2-10_R9C_sb1	H-KLPPG-WEICI -SRSRGR-VYYH-NHIT-NTE-QSERPSG-OH	3968.0	3967.3
WW-2-10_R9C_sb2	H-KLPPG-WEICI -ERSRGR-VYYH-NHIT-NTE-QWERPSG-OH	4109.0	4108.0
WW-2-10-HS1-1	H-PLPPG-WEKRNle-SRSRGR-IYYH-NSIT-KTE-TWERPRL-OH	4180.2	4181.1
WW-2-10-HS1-2	H-PLPPG-WEKRI -SRSRGR-IYYH-NSIT-KTE-TWERPRL-OH	4180.2	4180.8

Table 10: Sequences of a further developed organophosphate-binding and Zn^{2+}-binding WW domain peptide. The peptide termini are displayed as "NT" for N-terminus and "CT" for C-terminus. Beta-sheets are indicated as BS and loop regions as LP. All peptide sequences have already been published in Lindner *et al.* 2024. This list was minorly adapted and reused from Lindner *et al.*[94], with permission from ©2024 American Chemical Society.

Peptide name	Sequence	[M+H]$^+$ (calc.)	[M+H]$^+$ (found)
	------NT-------BS1------LP1------BS2---LP2----BS3-----CT		
WW-HS-ATP	H-PLPPG-WEIRI-DRSRGR-IYYH-NSIT-KTE-TWERPRL-OH	4193.2	4193.6
WW-HS-PC	H-PLPPG-WEIRI-DSGSGR-IYYS-NSIT-RTS-TWERPRL-OH	3961.0	3963.5
WW-HS-CA	H-PLPPG-WEIHI-DSGSGR-IHYH-NSIT-KTT-TWERPRL-OH	3952.0	3952.8

6.2 Methods

6.2.1 Solid-phase peptide synthesis (SPPS)

All peptides were synthesized on a *Liberty Blue CEM* microwave-assisted automated solid-phase peptide synthesizer using a standard *Fmoc/tBu* orthogonal protection group strategy. Different solid support resins were utilized (cf. Table 11). For the peptides synthesized for the

WW domain screening project, the five thermostable WW domain variants, the group-specific WW-HS peptides, the reference peptides and the group-specific ligands, synthesized as peptide amides, H-Rink Amide ChemMatrix® resin (loading density: 0.40-0.60 mmol g^{-1}) was used. As a result of the discontinuation of ChemMatrix®resin production, the peptide WW-II/III and further re-synthesized peptides were synthesized on H-Rink amide 4-methylbenzylhydrylamine (MBHA) polystyrene resin (loading density: 0.67 mmol g^{-1}).

A glycine or serine-preloaded HMPB ChemMatrix® resin (loading density: 0.30-0.65 mmol/g), depending on the final amino acid at the peptide's C-terminus, was utilized for generating the four N-terminal split peptides B_{CC}-WW-1 to B_{CC}-WW-4 (C-terminal acids), and the later combined split peptides WW-2-10 and WW-1-8. The six C-terminal split fragment peptides, bearing a C-terminal amide, were synthesized on H-Rink Amide ChemMatrix® resin. The six different *h*Pin1-derived C-terminal split peptides were synthesized using a split approach. This implies that the C-terminal coiled coil sequence part was synthesized once in a normal synthesis scale (0.1 mmol), subsequently the generated resin was then divided into six parts for the different modified *h*Pin1 fragments that were all attached to the identical coiled coil part by subsequent SPPS.

The uncapped hyperthermostable variant WW-HS-2_{OH}, as well as the refined ATP-binding peptide variants and the thermostable application peptides (WW-HS-ATP, WW-HS-PC and WW-HS-CA) were synthesized as peptide acids on a leucine-preloaded HMPB-ChemMatrix® resin (loading density: 0.30-0.65 mmol g^{-1}).

The peptide synthesis scale was based on the project purpose. Peptides used for screening projects as the iterative sequence screening for generating a thermostable variant and the combinatorial screening approach for identifying organophosphate binders, were synthesized on a 0.05 mmol scale. For all other peptides that were intensively studied concerning structure, stability, binding behavior and/or enzymatic activity, a synthesis scale of 0.1 mmol was applied.

Table 11: List of utilized solid support resins for SPPS.

Peptide group	Solid support resin	Synthesis scale
WW domain screening	H-Rink Amide ChemMatrix® resin or H-Rink amide 4-MBHA polystyrene resin (after the discontinuation of ChemMatrix®)	0.05 mmol
WW-HS-1 to 5	H-Rink amide 4-MBHA polystyrene resin (after the discontinuation of ChemMatrix®)	0.1 mmol
WW-HS-2_{OH}	Leu-preloaded HMPB-ChemMatrix® resin	
WW-HS group-specific variants **WW domain group reference peptides**	H-Rink Amide ChemMatrix® resin or H-Rink amide 4-MBHA polystyrene resin (after the discontinuation of ChemMatrix®)	0.1 mmol
***h*Pin1$_{WW}$-based split B_{cc} variants**	Gly-preloaded HMPB-ChemMatrix® resin	0.05 mmol
***h*Pin1$_{WW}$-based split A_{cc} variants**	Ser-preloaded HMPB-ChemMatrix® resin	
***De novo* designed ATP-binding variants**	Gly-preloaded HMPB-ChemMatrix® resin or Leu-preloaded HMPB-ChemMatrix® resin	0.1 mmol
Organophosphate-binding peptides	Leu-preloaded HMPB-ChemMatrix® resin	0.1 mmol

Zn(II)-binding WW-HS-CA peptide	Leu-preloaded HMPB-ChemMatrix® resin	0.1 mmol

The required amount of amino acids (0.2 M in DMF), activator, activator base, deprotection and DMF was determined based on the *CEM* usage calculator software and adjusted accordingly with the reagent calculator software for individual synthesis purpose. First, the amount of amino acids was weighed and dissolved in DMF. Amino acids that were not sufficiently dissolved in DMF by vortexing, were additionally placed in an ultrasonic bath (*Ultrasonic Cleaner USC-T, VWR)* for 2 min. As activator base solution OxymaPure® (1.0 M) and DIPEA (0.1 M) in DMF was used. Whereas DIC (0.5 M) in DMF was prepared as activator solution. When the synthesis scale was reduced to 0.05 mmol, the final concentration of the before mentioned reagents was reduced by half accordingly.

Before SPPS was performed on the *Liberty Blue CEM* microwave-assisted device, the utilized solid support resin was pre-swollen in DMF for 20 min, to ensure a proper diffusion of amino acids and reagents into the resin's pores and thus facilitate the solvation of the growing peptide chain.[143]

For deprotecting the temporary Fmoc group, a 20 % (*v/v*) piperidine solution in DMF was used. In case of utilizing the OMpe protected Asp amino acid (exclusively for the synthesis of WW-HS-CA and WW-HS-PC), 5 % (*v/v*) of formic acid was added to the deprotection, in order to prevent aspartimide formation by acidifying the deprotection solution.[144]

After adding the pre-swollen resin to the reaction vessel, the automated synthesis was started with another automatic resin swelling step (addition of 15 ml DMF, incubation time: 300 s). The generally utilized synthesis methods, summarized in Table 11, are based on the CarboMAX™ method of *CEM*[145], named as single coupling (HS) Carbo CAM or double coupling (HS) Carbo CAM.[146] These methods were only minorly modified concerning the introduction of a higher amount of washing steps and a longer draining time of the reaction vessel. For the amino acid His, a milder temperature setting was used and for Arg double couplings were performed. Since the coupling efficiency during SPPS is known to be reduced with increased length of peptide chain[147], double couplings were generally performed after the 18th amino acid, concerning the last 16 amino acids for all WW-domain based peptides (synthesis direction from C-to N-terminus), except for the hPin1-derived split variants.

One of the hyperthermostable peptides, namely WW-HS-2$_{OH}$, contains a DSG sequence, which is known to be highly prone to aspartimide formation. In order to prevent this side reaction, a CSY-protected Asp residue in combination with a modified synthesis method at room temperature (25 °C), beginning with Ser at position 12, was used for SPPS (cf. Table 12, Single CSY Coupling (HS) Carbo CAM Cool Down or Double CSY Coupling (HS) Carbo CAM Cool Down).

Table 12: Standard methods used for automated SPPS on a Liberty Blue peptide synthesizer from *CEM*.

Method	Procedure	Parameter	Temperature [°C]	Microwave power [W]	Time duration[sec]
	a. Deprotection	a. Volume: 4 ml	75 90	115 30	15 50

Standard Deprotection	b. 3 x Wash	b. Each with: Volume: 4 ml, Drain Time: 5 s			
Standard Single Coupling	a. Deprotection	a. Volume: 4 ml	75 90	115 30	15 50
	b. Wash	b. Volume: 4 ml, Drain Time: 10 s			
	c. 3 x Wash	c. Each with: Volume: 4 ml, Drain Time: 5 s			
	d. Coupling	d. Volume AA: 2.5 ml, Volume Act.: 2 ml, Volume Act. B.: 1 ml	75 90	170 30	15 110
	e. Wash	e. Volume: 4 ml, Drain Time: 10 s			
	f. Wash	f. Volume: 4 ml, Drain Time: 5 s			
Standard Double Coupling	a. 2 x Deprotection	a. Each with: Volume: 4 ml	75 90	115 30	15 50
	b. Wash	b. Volume: 4 ml, Drain Time: 10 s			
	c. 4 x Wash	c. Each with: Volume: 4 ml, Drain Time: 5 s			
	d. Wash	d. Volume: 4 ml, Drain Time: 10 s			
	e. Coupling	e. Volume AA: 2.5 ml, Volume Act.: 2 ml, Volume Act. B.: 1 ml	75 90	170 30	15 110
	f. 2 x Wash	f. Each with: Volume: 4 ml, Drain Time: 10 s			
	g. Coupling	g. Volume AA: 2.5 ml, Volume Act.: 2 ml, Volume Act. B.: 1 ml	75 90	170 30	15 110
	h. 4 x Wash	h. Each with: Volume: 4 ml, Drain Time: 10 s			
Standard Coupling for His	a. Deprotection	a. Volume: 4 ml	75 90	115 30	15 50
	b. 4 x Wash	b. Each with: Volume: 4 ml, Drain Time: 5 s			
	c. Coupling at 50 °C for 10 min	c. Volume AA: 2.5 ml, Volume Act.: 2 ml, Volume Act. B.: 1 ml	25 50	0 35	120 480
	d. 2 x Wash	d. Each with: Volume: 4 ml, Drain Time: 5 s			
Single CSY Coupling (HS) Carbo CAM Cool Down	a. 2 x Deprotection	a. Each with: Volume: 4 ml	25	0	300
	b. Wash	b. Volume: 4 ml, Drain Time: 10 s			
	c. 3 x Wash	c. Each with: Volume: 4 ml, Drain Time: 5 s			
	d. Coupling	d. Volume AA: 2.5 ml, Volume Act.: 2 ml, Volume Act. B.: 1 ml	25	0	3600
	e. 4 x Wash	e. Each with: Volume: 4 ml, Drain Time: 10 s			

Double CSY Coupling (HS) Carbo CAM Cool Down	a. 2 x Deprotection	a. Each with: Volume: 4 ml	25	0	300
	b. Wash	b. Volume: 4 ml, Drain Time: 10 s			
	c. 4 x Wash	c. Each with: Volume: 4 ml, Drain Time: 5 s			
	d. Wash	d. Volume: 4 ml, Drain Time: 10 s			
	e. Coupling	e. Volume AA: 2.5 ml, Volume Act.: 2 ml, Volume Act. B.: 1 ml	25	0	3600
	f. 2 x Wash	f. Each with: Volume: 4 ml, Drain Time: 10 s			
	g. Coupling	g. Volume AA: 2.5 ml, Volume Act.: 2 ml, Volume Act. B.: 1 ml	25	0	3600
	h. 2 x Wash	h. Each with: Volume: 4 ml, Drain Time: 10 s			

6.2.2 Acetylation of a peptide's N-terminus

Experimental procedure

The N-termini of all WW-domain based peptides used for the thermostable scaffold project, as well as all designed group-specific hyperstable peptides, except for WW-HS-2_{OH}, were acetylated with acetic anhydride in pyridine (1:9, v/v). The mixture was added directly to the peptide resin and incubated for 10 min under shaking (*Orbital Shaker*, *VWR*) at room temperature. Subsequently, the solution was discarded, the peptide resin was washed 20 times with 5 ml DCM and dried *in vacuo* in a desiccator.

6.2.3 Cleavage of a peptide from it's solid support resin

Experimental procedure

As a result of applying the *Fmoc/tBu* synthesis strategy for the SPPS of all peptides presented in this research work, the simultaneous removal of the synthesized peptide from its attached linker-solid support resin at the C-terminus and the cleavage of the permanent protecting groups from the respective functional side-chains, was performed under acidic conditions using a TFA-based mixture TFA/TIPS/ultrapure H_2O (90:5:5, (*v/v/v*)).

The cleavage solution was added to the peptide resin in a 10 ml *BD* syringe and incubated for three hours on a shaker (*Orbital Shaker*, *VWR*) at room temperature. Then, the peptide, dissolved in the mixture, was collected in a 50 ml *Falcon* tube and the resin was washed 2 x with 4 ml TFA. Subsequently, the peptide-containing liquid was concentrated to a fifth of the initial volume under a constant flow of nitrogen. By adding 25-30 ml of ice-cold diethylether to the reaction tube and centrifuging it for 10 min at 3000 rpm (*Mega Star 600 centrifuge*, *VWR*), the crude peptide was precipitated. The supernatant was discarded and the peptide was washed twice with ice-cold ether. Afterwards, the peptide pellet was dried for one hour in a desiccator, dissolved in a 10-15 ml mixture of acetonitrile/ultrapure H_2O, frozen in liquid

nitrogen and freeze-dried with a lyophilizer (*Alpha 2-4-LSCbasic*, *Christ*), in order to obtain a fine white solid.

6.2.4 Removal of a CSY-protection group

Experimental procedure

The Fmoc-Asp (CSY)-OH amino acid was used for the peptide WW-HS-2_{OH}. The deprotection of the CSY permanent protection group was performed based on the publication of Pham *et al*.[112] After SPPS, the CSY protected peptide resin was cleaved off the solid support, according to the procedure described in chapter 1.2.2. The CSY protected crude peptide was purified by semi-preparative Reverse-Phase (RP) High-Performance Liquid Chromatography (HPLC) before the removal of the CSY protection group.

For CSY deprotection 250 nmol of the CSY protected peptide was dissolved in HPLC buffer A (ultrapure H_2O + 0.1 % TFA), containing 10 % of *1,1,1,3,3,3*-Hexafluoroisopropanol (HFIP) and the cleavage of the CSY-group was achieved by adding a solution of 1.25 mM *N*-chlorosuccinimide (NCS) in HPLC buffer B (acetonitrile + 0.1 % TFA) with 10 % HFIP. Once the NCS solution was added (2.2 equivalents per CSY group), the mixture was gently vortexed (*VV3*-Vortex mixer, *VWR*) and equilibrated for 10 min. The step of adding NCS solution was repeated another three times. The cleavage reaction was then stopped with a 7.5 mM sodium ascorbate solution in 10 % of HPLC buffer B, dissolved in HPLC buffer A. The mixture was instantly frozen in liquid nitrogen and freeze-dried with a lyophilizer (*Alpha 2-4-LSCbasic*, *Christ*).

6.2.5 Semi-preparative reversed phase high-performance liquid chromatography (RP-HPLC)

Experimental procedure

As semi-preparative reversed-phase HPLC chromatographic devices the Jasco chromatography system (pump *PU-4180*, detector *UV-4070* and column oven *CO-4060*) and the *Shimadzu* chromatography system (pump *LC-40D*, degasser *DGU-405*, photo diode detector *SPD-M40*, column oven *CTO-40S*) were utilized. All peptide purification steps were either performed on a *VDS optilab VDSpher® PUR* 100 C18-SE (250 x 10 mm, 100 Å, 5 µm) column and a *Macherey-Nagel VP Nucleodur* C18-EC (250 x 10 mm, 100 Å, 5 µm) column at a flow rate of 3 ml/min and a column oven temperature of 50 °C.

For both systems the purification was performed by using different linear gradients, depending on the peptide's polarity/hydrophobicity, that are summarized in Table 13. Before a peptide was injected, the chosen HPLC column was equilibrated for 20 min with the starting composition of the selected HPLC gradient. As solvent system, ultrapure water with 0.1 % TFA was used as polar Buffer A and acetonitrile with 0.1 % TFA as Buffer B. After starting the HPLC method, the starting composition was kept constant for 5 min. Then, the linear gradient was increased to the final value within 30 min, followed by an increase to 100 % Buffer B within 1 min and an incubation time of 5 min. At the end of the method, the Buffer B content was reduced to the starting gradient within 1 min and kept constant for the remaining 5 min.

Prior to HPLC purifications lyophilized crude or pre-purified peptide fractions were freshly dissolved in an appropriate mixture of HPLC buffer A and B, whereas the final content of the hydrophobic solvent was chosen to be less than or similar to the starting composition of the program.

Table 13: List of utilized linear gradients for peptide purification, all with a gradient duration of 30 min.

Gradient Number	Percentage of HPLC Buffer B in A (*v/v*)
1	5-20 %
2	10-30 %
3	10-35 %
4	10-40 %
5	10-50 %
6	15-50 %
7	20-40 %
8	20-50 %
9	20-60 %
10	20-70 %
11	25-40 %
12	25-45 %
13	25-50 %
14	25-60 %
15	30-60 %

Prior to the injection, 1.5 -1.8 ml of the peptide solution was filtered with a H-PTFE syringe filter (0.2 µM) from *Macherey-Nagel* directly into the 2 ml loop. For first crude peptide purification runs, several peaks were collected in multiple fractions and the product was subsequently identified by MALDI-TOF mass spectrometry (cf. Chapter 6.2.7). For re-purification rounds, exclusively the main peak containing the peptide, was collected in four to six different fractions. Peptide fractions that showed a purity of ≥ 95 % in the analytical HPLC chromatogram were pooled in ultrapure H_2O and freeze-dried using a lyophilizer (*Alpha 2-4-LSCbasic*, *Christ*). Subsequently, the concentration of the pure peptide fractions was determined by UV/Vis spectroscopy (cf. Chapter 6.2.8.1).

6.2.6 Analytical HPLC

Experimental procedure

Lyophilized samples peptides were dissolved in 20 % HPLC Buffer B in Buffer A and further diluted (1:1) in Buffer A in the HPLC vial. A total sample volume of 20 µl was chosen for analysis. Pre-purified and already dissolved fractions were also further diluted in HPLC Buffer A, in order to have a final hydrophobic content less than that of the starting composition of the respective program. To enforce homogenous mixture in the sample vial, the peptide solution was pipetted up and down up to 10 times, when added to Buffer A. The vials were then equipped with a lid and placed in the autosampler of the *VWR Hitachi Primaide* chromatography system, equipped with a *VDSPher PUR* 100 C18-SE (5 µm, 250 mm by 4.6 mm) reversed-phase column from *VDS Optilab* that was equilibrated at 50 °C at a flow rate of 1 ml min^{-1}. The linear gradients that were used are summarized in table 15. Automatically

recorded chromatograms were examined concerning the presence of side products and general purity of the sample. Peptide fractions with a purity of ≤ 95 % were further used for experiments.

Table 15: List of utilized linear gradients for peptide purification, all with a gradient duration of 30 min.

Gradient Number	Percentage of HPLC Buffer B in A (*v/v*)
1	5-20 %
2	10-30 %
3	10-35 %
4	10-40 %
5	10-50 %
6	15-50 %
7	20-40 %
8	20-50 %
9	20-60 %
10	20-70 %
11	25-40 %
12	25-45 %
13	25-50 %
14	25-60 %
15	30-60 %

6.2.7 Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) spectrometry

Experimental procedure

For the identification of the synthesized peptides by their monoisotopic mass, MALDI-TOF measurements were performed. A saturated *2,5*-dihydroxy benzoic acid (DHB) solution (129.8 mM, 20 mg DHB solid dissolved in a 1 ml solvent mixture of 70:30 (*v/v*) of HPLC Buffer A in B) was chosen as matrix for all peptides described in this work. The samples were dissolved in either ultrapure H_2O or HPLC buffers. First, 1 µl of DHB matrix solution was applied on a MALDI target plate, followed by rapid addition of 1 µl of the peptide solution. By pipetting up and down three times, the solution was properly mixed and the spot was dried at room temperature. For all peptides with a mass between 2000-4400 Da, the reflector positive mode was used at the *autoflex Speed* from *Bruker* for mass analysis. Peptide ligands, such as Ligand WW-IV, with a high negative net charge were measured in linear negative mode.

6.2.8 Spectroscopic methods

6.2.8.1 UV/Vis spectrometry

Experimental procedure for concentration determination

The concentration of pure peptide fractions was examined at an *IMPLEN NP80* nanophotometer using the wavelength program with the following settings (λ: either 280 or 214 nm, baseline correction at 750 nm, smoothing: set off). Before the absorbance at the respective wavelength was measured in a 1 cm quartz cuvette from *Hellma Analytics*, the

lyophilized sample was dissolved in degassed ultrapure H_2O. A dilution of 1:10 (*v/v*) of the sample with a total volume of 200 µl was prepared in an Eppendorf tube, vortexed, centrifuged and transferred to the cuvette. As a blank value, the absorbance of 200 µl ultrapure H_2O was taken. In between of different measurements, the cuvette was cleaned according to the following standard washing procedure: 3 x ultrapure H_2O, 3 x 2 %Hellmanex mixture(2% *v/v* in ultrapure H_2O), 4-6 x ultrapure H_2O and 2 x isopropanol (*p.a.* grade). To dry the cuvette, a constant flow of nitrogen was applied. Diluted peptide mixtures were recovered after each measurement and after aliquotation transferred back to the pure peptide fraction. The peptide concentration was determined using the Beer-Lambert Law (cf. Eq. 1).

$$A = log \cdot \frac{I0}{I} = \varepsilon(\lambda) \cdot c \cdot d \qquad \textbf{Eq. 1}$$

with ***A*** as the absorbance, $\boldsymbol{I_0}$ as the input light intensity, *I* the output/transmitted light intensity, the extinction coefficient ε at a specific wavelength $\boldsymbol{\lambda}$, the analyte's concentration ***c*** and the pathlength ***d*** of the cuvette.

Specific extinction coefficients, listed in Table 15, were calculated from the weighted sum of chromophores (Trp and Tyr) in the amino acid sequence. The concentration of peptide ligands that did not contain chromophores, was identified using the absorbance value at 214 nmn (cf. Table 16).[148]

Table 16: List of specific extinction coefficients. The specific extinction coefficients indicated with an "a" were utilized for peptide variants belonging to the WW basic screening project and the WW-HS peptide series.[149] The extinction coefficients indicated with a "b" were used for the *h*Pin1$_{WW}$-based split combi project in Neitz *et al.* and the WW-HS organophosphate-binding and WW-HS-CA peptides.[150]

Analyte	Specific extinction coefficient [L · mol^{-1} · cm^{-1}]	Wavelength [nm]
Tryptophan[a]	5690	280
Tyrosine[a]	1280	280
Tryptophan[b]	5540	280
Tyrosine[b]	1480	280
Peptide bond	923	214
Alanine	32	214
Arginine	102	214
Glycine	21	214
Leucine	45	214
Proline	30	214
Serine	34	214
Threonine	41	214
Valine	43	214

6.2.8.2 Circular dichroism (CD) spectroscopy

Experimental procedure

All CD spectra and thermal denaturation curves were recorded on a *JASCO J-810* CD spectrometer with *JASCO PTC423S* Peltier thermostat cell holder under constant nitrogen flow of 6-7 ml/min. Peptides designed for the screening project and all WW-HS peptides, except for WW-HS-CA, were measured in PBS buffer with a pH of either 7.4 (WW-2-10 and its

variants) or 7.5. The zinc-binding hyperstable variant WW-HS-CA was dissolved in ultrapure H_2O and 5XMOPS with 150 mM NaCl (pH 7.2). All peptides were dissolved in degassed ultrapure H_2O then, 10xPBS buffer was added and the mixture was vortexed with a *VV3-Vortex mixer* from *VWR*, centrifuged and stored on ice prior to the experiment. The final concentration of the peptide used for all measurements was 100 µM. Measurements were performed in a 1 mm spectrosil quartz cuvette from *Starna*. As blank, the respective buffer without peptide was recorded. Before sample measurement, the mixture was transferred to the cuvette, placed in the cell holder and was equilibrated for 3 min at the respective temperature.

CD spectra were recorded at a wavelength range of 190-260 nm at 20 °C with a scanning speed of 100 nm · min^{-1}, a time response of 2 s and a 1 nm bandwidth with a low sensitivity setting (cf Table 18). The obtained CD spectra were averaged from five consecutively recorded individual spectra, CD signals were corrected by the blank (Buffer) measurement and converted to mean residue ellipticity (MRE) (deg·cm^2·$dmol^{-1}$·res^{-1}) using the following Equation 2.

$$MRE = \frac{100 \cdot \theta}{l \cdot c \cdot n} \quad \textbf{Eq. 2}$$

with θ being the ellipticity (mdeg), l the optical path length (cm), c the peptide concentration (mmol · L^{-1}) and n (res) the number of peptide backbone bonds.[101]

Table 18: Parameter settings for CD spectrum measurements.

Parameter name	Setting
Jasco program	Spectrum Measurement
Sensitivity	Low (1000 mdeg)
Wavelength start	260 nm
Wavelength end	190 nm
Data pitch	1 nm
Scanning mode	Continuous
Scanning speed	100 nm/min
Response	1 sec
Band width	2 nm
Accumulation number for spectra	5

The overall peptide stability was examined by thermal denaturation. For denaturation curves the *Jasco* program *Variable Temperature Scan* with the following settings was used: fixed wavelength at 228 nm, temperature range for screening peptides: 5-95 °C and for examining promising hyperthermostable candidates: 0-98 °C, both with a heat rate of 1 °C · min^{-1}, a 1 nm bandwidth, a 16 s response time and a data point collection interval of 1 °C (cf. table 19). The respective melting temperature T_M was determined by fitting Equation 3 to the respective thermal denaturation profiles:

$$[\Theta]_{228\,\mathrm{nm}}(T) = \frac{\left(m_f \cdot (T + 273.15) + \theta_f\right) + \theta_u - \left(m_f \cdot (T + 273.15) + \theta_f\right)}{1 + \exp\left(\left({}^{-\Delta H}/_{R \cdot ((T + 273.15) - (T_m + 273.15))}\right)\right)} \quad \textbf{Eq. 3}$$

with θ_f representing the ellipticity of the fully folded peptide, θ_u the ellipticity of the unfolded peptide, m_f the slope of the pre-transition, ΔH the molar enthalpy, T the temperature in Kelvin, T_m the melting temperature in Kelvin and R the universal gas constant.[130]

To ensure accurate fitting for redesigned ATP-binding peptide variants, the lower data plateau of the unfolded peptide θ_u was further described as a linear function with the form $y = mx + b$.

All hyperthermostable peptides (WW-HS series) were also tested on a possible reversibility of thermal denaturation by recording a cooling curve subsequent to the denaturation curves with the same settings as stated above. A visual check for potential formed peptide aggregates was performed, followed by another CD spectrum measurement.

In between different peptide measurements the following washing procedure of the CD cuvette was applied: First, the cuvette was rinsed 3 x with 1 mM HCl, continued by 4 x ultrapure H_2O, 3 x 2 % Hellmanex solution, 4-6 x ultrapure H_2O and 3 x isopropanol (*p.a.* grade). The cuvette was dried under constant nitrogen flow.

Table 19: Parameter settings for CD thermal denaturation curve measurements.

Parameter name	Setting
Jasco program	Variable Temperature Measurement
Fixed Wavelength	228 nm
Sensitivity	Low (1000 mdeg)
Starting temperature	0 °C or 5 °C
Final temperature	98 °C or 95 °C
Temperature slope	1 °C/min
Response	2 sec
Band width	2 nm

For three different ATP-binding peptides, namely WW-HS-ATP, WW-2-10 and WW-2-10-HS2, the influence of the ATP binding on the secondary structure, as well as on the melting temperature was determined by preparing overnight incubations of 100 µM peptide with 100 µM ATP in 1xPBS (pH: 7.4). The samples were stored at -4 °C until the measurements were performed on the next day using the settings described above. Furthermore, a possible influence of Mg^{2+} on the structure and stability of peptide-ATP complexes was investigated for overnight incubations of 100 µM peptide with 100 µM ATP and 1 mM Mg^{2+} in 1xPBS (pH: 7.4).
Besides the CD measurement of WW-HS-CA in MOPS-buffered saline (pH 7.2), the influence of Zn^{2+} was also examined. Therefore overnight incubations of 100 µM WW-HS-CA with 100 µM Zn^{2+} were prepared and measured on the following day with the standard settings as mentioned in table 17 and 18.

6.2.8.3 Fluorescence titration spectroscopy

Experimental procedure

To study the binding affinity of the designed peptides towards WW domain group-specific ligands or selected organophosphates (ATP, ADP, AMP, PC and IP6), the change of the intrinsic tryptophan fluorescence intensity was examined with a *BMG Labtech* microplate reader (*ClarioStar*) in non-binding, F-bottom-shaped 96-well plates from *Greiner Bio One* at room temperature (ranging from 23 to 25 °C). As buffer, a 1xPBS (pH 7.4) was used, except for investigating WW-CA-HS towards its $Zinc^{2+}$ binding affinity. Here, 1xMOPS buffer with 150 mM NaCl and a pH value of 7.2, was utilized.

Incubations of the respective peptide (2 µM final concentration) with increasing ligand concentrations (0 to 250 µM) were prepared from tenfold stocks and in non-binding eppendorf tubes using non-binding pipette tips for pipetting the exact volume and incubated for at least 12 h at 4 °C either at the fridge (for the screening of split combination peptides for organophosphate binding and the examination of WW domain HS group-specific peptides towards their respective ligands) or equilibrated at 4 °C on a *Hettich BENELUX* thermomixer under constant shaking at 200 rpm and light-protection. For investigating the influence of Mg^{2+} on the ATP binding affinity of all variants, 2 mM MgCl in 1xPBS was added to the peptide incubations.

Fluorescence spectra were recorded from 335 – 400 nm with 20 flashes per well, a gain of 1800, a focal height set to 4.4 mm, a top-to-bottom reading and a setting time of 0.2 s. An excitation filter was used to set the excitation wavelength λ_{ex} to 295 nm. Acquired fluorescence intensity signals were corrected by 1xPBS buffer signal and the tryptophan fluorescence signal of a peptide incubation without ligand was used as zero data point. In total, 11 different ligand concentrations were chosen (0.25; 0.5; 1.0; 2.0; 3.9; 7.8; 15.6; 31.3; 62.5; 125; 250 µM) per binding experiment and each concentration was measured in octuplicate. After background signal correction, the resulting binding curves were plotted from the change of fluorescence signals and fitted to the quadratic equation, Eq. 4.

$$F_{350} = Y_0 + S \cdot \frac{(c + P + K_d) - \sqrt{(c + P + K_d)^2 - 4 \cdot P \cdot c}}{2 \cdot P} \quad \textbf{Eq. 4}$$

with *c* being the ligand concentration, K_d the binding dissociation constant, Y_0 the fluorescence at zero ligand concentration, *S* the maximum change of fluorescence intensity and *P* the peptide concentration.

Error bars shown in the respective binding curves were calculated as standard error (*SE*), using the following equation, Eq. 5:

$$SE = \frac{SD}{\sqrt{n}} \quad \textbf{Eq. 5}$$

with *SD* representing the standard deviation and *n* the amount of included data points (ranging from 4 to 8).

For some saturation binding curves obtained from the organophosphate binding screening, Eq.5 was not applicable, thus the hyperbolic fitting function (Hyperbl) of the *OriginPro* software with the following equation, Eq. 6, was utilized.

$$y = \frac{P1 \cdot x}{P2 + x} \quad \textbf{Eq. 6}$$

with P1 as asymptote of the y-axis, x as ligand concentration and P2 as K_d value.

6.2.9 Enzymatic activity tests

6.2.9.1 Kinetic measurements of *p*-NPP hydrolysis

Experimental procedure for kinetic measurements of *p*-NPP hydrolysis

Kinetic measurements were performed on an *IMPLEN NP80* nanophotometer in order to follow the conversion of *p*-NPP to the colorful nitrophenolat ion, that absorbs at a wavelength of 405 nm by enzymatically active molecules. Thus, the absorbance at 405 nm was recorded for 30 min with a data interval of 10 s in a 1 cm quartz cuvette from *Hellma Analytics*. Measurements were performed in PBS buffer of the following two different pH values (4.8 and 7.4) to test peptides for acidic and neutral phosphatase activity in triplicates. As background measurement the absorbance of 1 mM *p*-NPP in PBS was examined over time. As positive control, Potato phosphatase, provided by *Sigma Aldrich*, was utilized. The control enzyme and all tested peptide solutions were prepared as 200 µM stock solutions in PBS with a final concentration of 20 µM in the cuvette. The light-sensitive substrate solution of *p*-NPP was prepared as tenfold stock in an eppendorf tube (equipped with aluminium foil), the concentration was determined by its hydrated weight in PBS. The substrate solution was continously protected from light throughout the experiment and its storage at -4°C. Background measurements with 1 mM *p*-NPP were always performed prior to the experiment. When preparing the mixtures in Eppendorf tubes, PBS buffer was put first. Then, either the enzyme or peptide solution was added and mixed only by pipetting the liquid up and down 10 times. Directly before the measurement was started, the substrate solution was added and mixed by pipetting up and down 10 times. The solution was then transferred to the cuvette. The first data point of every mixture was used as blank and subsequently the kinetic measurement was performed.

Data analysis

The mean of triplicate measurements was determined and fitted by a linear regression model of the *OriginPro* software to obtain the slope m_{obs}. Based on the m_{obs} parameter, the reaction rate v_{obs} (cf. Eq. 7), the constant k_{obs} (cf. Eq. 8) and the catalytic efficiency k_2 (cf. Eq. 9) were calculated.[151] To determine k_{obs} for all studied peptides, the m_{obs} rates were corrected by the m_{ref} value obtained for the measurements of *p*-NPP autohydrolysis in PBS buffer.

$$v_{\mathrm{obs}} = \frac{m_{\mathrm{obs}} - m_{\mathrm{ref}}}{d \cdot \varepsilon_{405\,nm}} \quad \textbf{Eq.7}$$

with m_{obs} as the slope of the reaction with peptide and m_{ref} as the *p*-NPP autohydrolysis reaction, *d* as the cuvette pathlength [cm] and $\varepsilon_{405\ nm}$ as the specific extinction coefficient of *p*-NPP at 405 nm ($\varepsilon_{405\ nm}$ = 18000 $M^{-1}cm^{-1}$).[134,152]

$$k_{obs} = \frac{V_{obs}}{c_p} \qquad \textbf{Eq.8}$$

with c_p as the concentration of the peptide (catalyst).

$$k_2 = \frac{k_{obs}}{c_s} \qquad \textbf{Eq.9}$$

with c_s as the concentration of the substrate (*p*-NPP).

The error Δk_{obs} was calculated with the Gauss-error-propagation (cf. Eq. 10).

$$\Delta k_{obs} = \sqrt{\left(\frac{\Delta m_{obs}}{c_p \cdot d \cdot \varepsilon_{405\ nm}}\right)^2 + \left(\frac{\Delta m_{ref}}{c_p \cdot d \cdot \varepsilon_{405\ nm}}\right)^2} \qquad \textbf{Eq.10}$$

6.2.9.2 RP-HPLC method for the determination and quantification of adenosine phosphate-peptide mixtures

Experimental procedure for an analytical set-up on a semi-preparative RP-HPLC

In order to study typical chromatographic patterns of different organophosphates, such as adenosine triphosphate (ATP), adenosine diphosphate (ADP) and adenosine monophosphate (AMP), and additionally to investigate possible enzymatic activity of ATP-and ADP-binding peptides towards the respective organophosphate ligands, an analytical C18 ec column (250 x 4.6 mm, 100 Å, 5 µm) from *Macherey Nagel* was used at the *Shimadzu* chromatography system.

Frozen, solid organophosphate stocks (either ATP, ADP or AMP), aquired from *Sigma Aldrich,* were equilibrated on room temperature for at least 1 h, before a tenfold concentrated stock solution (10 mM) was prepared in 1xPBS (pH: 7.4). The hydrated mass of the respective molecule was determined by weight and used to determine concentrations of the stock solution. The final concentration of the respective organophosphate ligand in the incubation mixtures was 1 mM.

Preparation of peptide-ligand mixtures

Prior to the preparation of peptide (10 µM) ligand (1 mM) mixtures (ratio peptide to ligand: 1:100), frozen and lyophilized peptide aliquots were equilibrated for at least 30 min at room temperature and freshly dissolved in 1xPBS (pH: 7.4). A total volume of 300 µl for peptide-ligand mixtures was prepared in Eppendorf tubes, gently vortexed, centrifuged and stored at 4 °C overnight.

Enzymatic assay

The basic separation principle of the three different organophosphate molecules, ATP, ADP and AMP *via* HPLC was based on the publication of Juarez-Facio *et al*,[136] the used isocratic

method as well as enzymatic activity testing of peptides used in this work was substantially modified.

The *Shimadzu* chromatographic system pump *LC-40D*, degasser *DGU-405*, photo diode detector *SPD-M40*, column oven *CTO-40S*) was used with an analytical column (C18ec, Size: 250 x 4.6 mm, 100 Å, 5 µm, *Macherey Nagel*), at a flow rate of 0.6 ml/min, an column oven temperature of 20 °C and an isocratic method of 100 % 50 mM KH_2PO_4 (pH: 6.8) for 30 min, followed by 20 min washing with 70:30 (v/v) methanol: ultrapure H_2O and an equilibration of KH_2PO_4 buffer for 20 min prior to the next run.

Before injecting the first peptide-ligand mixture, the HPLC system was equilibrated on the 50 mM KH_2PO_4 buffer for at least 30 min. A volume of 100 µl of incubated peptide-organophosphate mixture was injected and the intensity was monitored at a wavelength of 254, 280 and 220 nm. As background measurements, the peak pattern of utilized organophosphates was examined for pure compounds and in mixtures of ATP:ADP (1:1) and ATP:ADP:AMP (1:1:1). Peptide-ligand incubations contained only one organophosphate molecule at a time and were examined after a 100 µl injection of the respective mixture.

6.2.10 Software and Data analysis

All data presented in this work were analyzed with *OriginPro® 2020*, *2022* and *2023* and Microsoft Excel Version *2022/2023*. Illustrated schemes and figures were generated using the *InkScape®* software. Chemical structures and reaction mechanisms were created with *ChemDraw22-2-0.* The secondary structure of all designed and synthesized peptides was visualized with the *PyMOL* software version *4.6.0*.

6.2.10.1 Prediction tools for peptide secondary structure

Computational procedure

Structure Prediction using *Alpha Fold2*

AlphaFold2 was accessed through Google Colaboratory (*ColabFold*) version *1.5.5* for basic open-access structure prediction using sequence alignments generated through *MMseqs2*.[100] First, *ColabFold* was connected to the Graphical Processing Unit (GPU) of the Google Compute Back-End in python3 to powerfully execute the structure prediction and enable MSA's with the databases.

As **query_sequence**, the respective peptide sequence was entered, as well as its name as **jobname**. The amount of relaxation steps (**num_relax**) of the predicted structure were set to five. As output, five different relaxed and unrelaxed ranked *pdb* files, as well as *png* files that show sequence coverage and local distance difference test (lDDT) of the five predicted structures. The best-ranked *pdb* file was then further used for visual analysis in *PyMol* or *in silico* docking studies with *AutoDockVina*.

Structure Prediction using *OmegaFold*

Before peptide structure prediction was performed with *OmegaFold*[113], the environment was prepared based on the procedure presented by *Github*[153] by using the two following commands in Linux:

To install *OmegaFold* from source:

```
pip install git+https://github.com/HeliXonProtein/OmegaFold.git
```

and to generate a repository:

```
git clone https://github.com/HeliXonProtein/OmegaFold
cd OmegaFold
python setup.py install
```

Then, the input file was created. The headline started with >, followed by the peptide's name and the amino acid sequence in the next line. The file was saved with the peptide's name as normal *fasta* file.

For running *OmegaFold*, a Linux shell was opened in the generated directory named OmegaFold and the following command was used:

```
omegafold Input_file.fasta Output_directory
```

After a successful prediction run, a pdb file was generated with the predicted peptide structure. This was visualized with *PyMol*.

6.2.10.2 Modelling of peptide-ligand interactions with *AutoDockVina*

Computational procedure

For molecular docking calculations *AutoDockVina*[126,127] *v. 1.5.7* was utilized. To prepare the ligand and receptor files, the programs *ChemDraw Professional 22.2.0, ChemDraw3D* and *AutoDockTools* (ADT) *v. 1.5.7* were applied. All predicted positions of the docked and receptor-bound ligand complex were visualized with *PyMol.* In principle, the computational procedure for the molecular docking followed the steps and suggestions of the published protocol of Forli *et al.* 2016 with only small alterations.[154]

Preparation of the ligand molecule

The respective ligand molecule used for docking was generated as two-dimensional structure in *ChemDraw*, saved as *cdxml* file, re-opened in *Chem3D* to check the three-dimensional structure and subsequently saved as *pdb* file. The *pdb* file was then loaded in the *AutoDockTool* software, defined as ligand structure and rotatable bonds were chosen. The default setting representing amide bonds and equivalents as non-rotatable, was maintained. The readily prepared ligand file was finally saved as *pdbqt* file for its further use in *AutoDockVina*.[126,127]

Preparation of the receptor molecule

For the docking studies, the secondary structure of *de novo* rationally designed peptides was previously predicted with *AlphaFold2*[100] and the generated *pdb* file of the best ranked one was loaded in the *AutoDockTool* program by choosing the tab "Grid->Macromolecule ->Open". After ADT has processed the input file, it was saved as *pdbqt* file for further use.

Further individual settings for the molecular docking studies

After having prepared the ligand and receptor molecule in suitable file formats for *AutoDockVina*, the receptor molecule was re-loaded in ADT under the option choose macromolecule and the ligand was selected by entering the tab “Grid ->SetMapTypes->OpenLigand“.

In *AutoDockVina*[126,127] the molecular docking of a ligand with its receptor molecule takes place in a size-defined three-dimensional search space (x, y and z), referred to as GridBox. Within the pre-defined search space, the ligand can be docked within suitable binding pockets of the receptor molecule. The larger the user-defined grid box is set, the more computational power is demanded as the number of possible binding sites is increased. For molecular docking studies performed in this work, a GridBox that covers the entire size of the peptides (all displaying around 34 amino acids), used as receptor macromolecule, was chosen (grid point size: x = 80, y = 80, z = 80).

After choosing the suitable parameters for the grid box, the changes were saved in configuration files, referred to as *config* files.

The amount of computational effort, in AutoDock Vina is called exhaustiveness. This parameter can be utilized to improve the consistency of the docking results by performing additional simulations. The default setting of the exhaustiveness is 8. Since Forli *et al* reported a better consistency with an increased value set to 24, this parameter value was utilized for the docking studies presented in this work with the following command:

```
./vina –config config.txt –exhaustiveness=24
```

Running AutoDock Vina

Subsequently, the *AutoDockVina*[126,127] program was launched. The calculated binding affinities of all found docked positions of the ligand were saved as *txt* files. As output file, a *pdbqt* file was generated that was visualized with *PyMol*. All molecular docking studies were performed with the same parameters at least three times to encourage reproducible and reliable data.

7 List of Literature

[1] P.-S. Huang, S. E. Boyken, D. Baker, *Nature* **2016**, *537*, 320.

[2] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature* **1996**, *382*, 525.

[3] E. G. Baker, G. J. Bartlett, K. L. Porter Goff, D. N. Woolfson, *Accounts of chemical research* **2017**, *50*, 2085.

[4] W. M. Park, *International journal of molecular sciences* **2020**, *21*.

[5] Z. Gáspári, L. Nyitray, *Biomolecular concepts* **2011**, *2*, 199.

[6] J. R. Pires, F. Taha-Nejad, F. Toepert, T. Ast, U. Hoffmüller, J. Schneider-Mergener, R. Kühne, M. J. Macias, H. Oschkinat, *Journal of molecular biology* **2001**, *314*, 1147.

[7] Y. Kato, Y. Hino, M. Tanokura, *Solution structure of the first WW domain of FBP11*, **2006**.

[8] Z. R. Crook, N. W. Nairn, J. M. Olson, *Trends in biochemical sciences* **2020**, *45*, 332.

[9] A. Ciesiołkiewicz, J. Lizandra Perez, Ł. Berlicki, *Bioorganic & medicinal chemistry letters* **2022**, *71*, 128806.

[10] L. Otte, U. Wiedemann, B. Schlegel, J. R. Pires, M. Beyermann, P. Schmieder, G. Krause, R. Volkmer-Engert, J. Schneider-Mergener, H. Oschkinat, *Protein science : a publication of the Protein Society* **2003**, *12*, 491.

[11] A. Aggeli, N. Boden, S. Zhang, *Self-Assembling Peptide Systems in Biology, Medicine and Engineering*, Springer Netherlands, Dordrecht, **2002**.

[12] J. M. Berg, J. L. Tymoczko, G. J. Gatto jr., L. Stryer, *Biochemistry 8/e*, Springer Spektrum, Berlin, Heidelberg, **2018**.

[13] J. G. Saven, *Current opinion in structural biology* **2002**, *12*, 453.

[14] A. K. Patel, R. R. Singhania, A. Pandey in *Biotechnology of microbial enzymes*. *Production, biocatalysis and industrial applications* (Eds.: G. Brahmachari, J. L. Adrio, A. L. Demain), Elsevier, Amsterdam, Netherlands, **2017**, pp. 13–41.

[15] M. J. Harms, J. W. Thornton, *Nature reviews. Genetics* **2013**, *14*, 559.

[16] S. G. Peisajovich, D. S. Tawfik, *Nature methods* **2007**, *4*, 991.

[17] W. Xiong, B. Liu, Y. Shen, K. Jing, T. R. Savage, *Biochemical Engineering Journal* **2021**, *174*, 108096.

[18] G. P. Pinto, M. Corbella, A. O. Demkiv, S. C. L. Kamerlin, *Trends in biochemical sciences* **2022**, *47*, 375.

[19] F. Thomas in *Peptide Bionanomaterials: From Design to Application* (Ed.: M. A. Elsawy), Springer, Cham, **2023**, pp. 87–119.

[20] a) J. A. Camarero, J. Pavel, T. W. Muir, *Angewandte Chemie International Edition* **1998**, *37*, 347; b) I. Coin, M. Beyermann, M. Bienert, *Nature protocols* **2007**, *2*, 3247; c) F. Mende, M. Beisswenger, O. Seitz, *Journal of the American Chemical Society* **2010**, *132*, 11110; d) S. Vishwanath, N. Srinivasan, *Progress in biophysics and molecular biology* **2014**, *116*, 151.

[21] A. Zanghellini, *Current opinion in biotechnology* **2014**, *29*, 132.

[22] M. Erak, K. Bellmann-Sickert, S. Els-Heindl, A. G. Beck-Sickinger, *Bioorganic & medicinal chemistry* **2018**, *26*, 2759.

[23] C. Eigenbrot, M. H. Ultsch, *HER2 extracelluar region with affinity matured 3-helix affibody ZHER2:342*, **2010**.

[24] S. Sheriff, *Crystal structure of the extracellular domain of the epidermal growth factor receptor in complex with an adnectin*, **2012**.

[25] M. Bellinzoni, S. Colinet, G. Behar, P. M. Alzari, F. Pecorari, *Crystal structure of the Sac7d-derived IgG1-binder C3-C24S*, **2011**.

[26] D. Schlatter, S. Brack, D. W. Banner, S. Batey, J. Benz, J. Bertschinger, W. Huber, C. Joseph, A. Rufer, A. van der Kloosters et al., *Human Chymase - Fynomer Complex*, **2012**.

[27] D. Fass, S. C. Blacklow, P. S. Kim, J. M. Berger, *LDL RECEPTOR LIGAND-BINDING MODULE 5, CALCIUM-COORDINATING*, **1997**.
[28] O. Kayode, B. Sankaran, E. S. Radisky, *Human Mesotrypsin in complex with amyloid precursor protein inhibitor variant APPI-M17G/I18F/F34V*, **2016**.
[29] M. M. Gewe, P. Rupert, R. K. Strong, *Exploring Cystine Dense Peptide Space to Open a Unique Molecular Toolbox*, **2018**.
[30] R. K. Spencer, P. J. Salveson, J. S. Nowick, *Crystal structure of macrocycles containing Abeta 16-22 (KLV(PHI)FAE) and Abeta 30-36 (AII(SAR)L(ORN)V)*, **2017**.
[31] Y. H. Lau, Y. Wu, M. Rossmann, P. de Andrade, Y. S. Tan, G. J. McKenzie, A. R. Venkitaraman, M. Hyvonen, D. R. Spring, *Structure of the Stapled Peptide Bound to Mdm2*, **2016**.
[32] G. M. Watson, M. J. Gunzburg, M. Wilce, J. A. Wilce, *Grb7 SH2 with bicyclic peptide inhibitor*, **2016**.
[33] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic acids research* **2000**, *28*, 235.
[34] C. Yang, F. Sesterhenn, J. Bonet, E. A. van Aalen, L. Scheller, L. A. Abriata, J. T. Cramer, X. Wen, S. Rosset, S. Georgeon et al., *Nature chemical biology* **2021**, *17*, 492.
[35] M. Blaber, J. Lee, *Current opinion in structural biology* **2012**, *22*, 442.
[36] K. Ożga, Ł. Berlicki, *ACS bio & med chem Au* **2022**, *2*, 316.
[37] C. M. Kraemer-Pecore, A. M. Wollacott, J. R. Desjarlais, *Current opinion in chemical biology* **2001**, *5*, 690.
[38] F. Curreli, K. Haque, L. Xie, Q. Qiu, J. Xu, W. Yong, X. Tong, A. K. Debnath, *Bioorganic & medicinal chemistry* **2015**, *23*, 7618.
[39] C. H. A. Che Hussian, W. Y. Leong, *Journal, genetic engineering & biotechnology* **2023**, *21*, 37.
[40] R. E. S. Thomson, S. E. Carrera-Pacheco, E. M. J. Gillam, *The Journal of biological chemistry* **2022**, *298*, 102435.
[41] M. Scian, J. C. Lin, I. Le Trong, G. I. Makhatadze, R. E. Stenkamp, N. H. Andersen, *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 12521.
[42] H. Meuzelaar, K. A. Marino, A. Huerta-Viga, M. R. Panman, L. E. J. Smeenk, A. J. Kettelarij, J. H. van Maarseveen, P. Timmerman, P. G. Bolhuis, S. Woutersen, *The journal of physical chemistry. B* **2013**, *117*, 11490.
[43] J. W. Neidigh, R. M. Fesinmeyer, N. H. Andersen, *Nature structural biology* **2002**, *9*, 425.
[44] J. C. Lin, M. Scian, N. H. Andersen, *Cyclo-TC1 Trp-cage*, **2012**.
[45] C. J. McKnight, P. T. Matsudaira, P. S. Kim, *Nature structural biology* **1997**, *4*, 180.
[46] Y. Bi, J.-H. Cho, E.-Y. Kim, B. Shan, H. Schindelin, D. P. Raleigh, *Biochemistry* **2007**, *46*, 7497.
[47] a) S. Xiao, Y. Bi, B. Shan, D. P. Raleigh, *Biochemistry* **2009**, *48*, 4607; b) L. E. Packer, B. Song, D. P. Raleigh, C. J. McKnight, *Biochemistry* **2011**, *50*, 3706.
[48] W. Meng, B. Shan, Y. Tang, D. P. Raleigh, *Protein science : a publication of the Protein Society* **2009**, *18*, 1692.
[49] K. J. Rosengren, N. L. Daly, M. R. Plan, C. Waine, D. J. Craik, *The Journal of biological chemistry* **2003**, *278*, 8606.
[50] R. Eliasen, N. L. Daly, B. S. Wulff, T. L. Andresen, K. W. Conde-Frieboes, D. J. Craik, *The Journal of biological chemistry* **2012**, *287*, 40493.
[51] C. K. Wang, C. W. Gruber, M. Cemazar, C. Siatskas, P. Tagore, N. Payne, G. Sun, S. Wang, C. C. Bernard, D. J. Craik, *ACS chemical biology* **2014**, *9*, 156.
[52] W.-J. Shen, T. Yao, X. Kong, K. W. Williams, T. Liu, *Biochimica et biophysica acta. Molecular basis of disease* **2017**, *1863*, 2477.
[53] K. J. Rosengren, N. L. Daly, M. R. Plan, C. Waine, D. J. Craik, *High resolution solution structure of kalata B1*, **2003**.
[54] N. L. Daly, R. Eliasen, D. J. Craik, *NMR solution structure of Kb1[ghrw;23-28]*, **2012**.

[55] D. Wensel, Y. Sun, Z. Li, S. Zhang, C. Picarillo, T. McDonagh, D. Fabrizio, M. Cockett, M. Krystal, J. Davis, *Antimicrobial agents and chemotherapy* **2017**, *61*.
[56] S. Aghaabdollahian, R. Ahangari Cohan, D. Norouzian, F. Davami, M. R. Asadi Karam, F. Torkashvand, G. Vaseghi, R. Moazzami, S. Latif Dizaji, *Scientific reports* **2019**, *9*, 2978.
[57] T. N. Grossmann, J. T.-H. Yeh, B. R. Bowman, Q. Chu, R. E. Moellering, G. L. Verdine, *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 17942.
[58] B. R. Bowman, T. N. Grossmann, J.-H. Yeh, G. L. Verdine, *Structure of beta-catenin in complex with a stapled peptide inhibitor*, **2012**.
[59] M. Teufl, C. U. Zajc, M. W. Traxlmayr, *ACS synthetic biology* **2022**, *11*, 1030.
[60] D. N. Woolfson, *The Journal of biological chemistry* **2023**, *299*, 104579.
[61] T. L. Pham, F. Thomas, *Chembiochem : a European journal of chemical biology* **2024**, *25*, e202300745.
[62] a) I. V. Korendovych, W. F. DeGrado, *Quarterly reviews of biophysics* **2020**, *53*, e3; b) A. J. Nicoll, R. K. Allemann, *Organic & biomolecular chemistry* **2004**, *2*, 2175; c) B. A. Smith, D. S. Daniels, A. E. Coplin, G. E. Jordan, L. M. McGregor, A. Schepartz, *Journal of the American Chemical Society* **2008**, *130*, 2948.
[63] A. J. Doig, *Chem. Commun.* **1997**, 2153.
[64] M. Jäger, H. Nguyen, J. C. Crane, J. W. Kelly, M. Gruebele, *Journal of molecular biology* **2001**, *311*, 373.
[65] M. Sudol, T. Hunter, *Cell* **2000**, *103*, 1001.
[66] J. L. Ilsley, M. Sudol, S. J. Winder, *Cellular signalling* **2002**, *14*, 183.
[67] Z. Salah, A. Alian, R. I. Aqeilan, *Frontiers in bioscience (Landmark edition)* **2012**, *17*, 331.
[68] V. Uversky, *Chembiochem : a European journal of chemical biology* **2008**, *9*, 2750.
[69] M. Meiyappan, G. Birrane, J. A. A. Ladias, *Journal of molecular biology* **2007**, *372*, 970.
[70] R. Wintjens, J. M. Wieruszeski, H. Drobecq, P. Rousselot-Pailley, L. Buée, G. Lippens, I. Landrieu, *The Journal of biological chemistry* **2001**, *276*, 25150.
[71] R. Wintjens, J.-M. Wieruszeski, H. Drobecq, G. Lippens, I. Landrieu, *SOLUTION STRUCTURE OF PIN1 WW DOMAIN COMPLEXED WITH CDC25 PHOSPHOTHREONINE PEPTIDE*, **2001**.
[72] Y. Kato, K. Nagata, M. Takahashi, L. Lian, J. J. Herrero, M. Sudol, M. Tanokura, *The Journal of biological chemistry* **2004**, *279*, 31833.
[73] a) E. Bayer, S. Goettsch, J. W. Mueller, B. Griewel, E. Guiberman, L. M. Mayr, P. Bayer, *The Journal of biological chemistry* **2003**, *278*, 26183; b) S.-S. Huang, L.-J. Hsu, N.-S. Chang, *FASEB bioAdvances* **2020**, *2*, 234.
[74] K. P. Lu, X. Z. Zhou, *Nature reviews. Molecular cell biology* **2007**, *8*, 904.
[75] Y. Kato, M. Ito, K. Kawai, K. Nagata, M. Tanokura, *The Journal of biological chemistry* **2002**, *277*, 10173.
[76] M. J. Macias, M. Hyvönen, E. Baraldi, J. Schultz, M. Sudol, M. Saraste, H. Oschkinat, *Nature* **1996**, *382*, 646.
[77] M. J. Macias, V. Gervais, C. Civera, H. Oschkinat, *Nature structural biology* **2000**, *7*, 375.
[78] M. J. Macias, S. Wiesner, M. Sudol, *FEBS letters* **2002**, *513*, 30.
[79] P. Bork, M. Sudol, *Trends in biochemical sciences* **1994**, *19*, 531.
[80] M. Jäger, Y. Zhang, J. Bieschke, H. Nguyen, M. Dendle, M. E. Bowman, J. P. Noel, M. Gruebele, J. W. Kelly, *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 10648.
[81] B. L. Kier, J. M. Anderson, N. H. Andersen, *Journal of the American Chemical Society* **2014**, *136*, 741.
[82] M. Jäger, M. Dendle, A. A. Fuller, J. W. Kelly, *Protein science : a publication of the Protein Society* **2007**, *16*, 2306.
[83] A. L. Stewart, J. H. Park, M. L. Waters, *Biochemistry* **2011**, *50*, 2575.

[84] T. L. Pham, M. Kovermann, F. Thomas, *ACS synthetic biology* **2022**, *11*, 254.
[85] H. Neitz, N. B. Paul, F. R. Häge, C. Lindner, R. Graebner, M. Kovermann, F. Thomas, *Chemical science* **2022**, *13*, 9079.
[86] a) A. Bochkarev, R. A. Pfuetzner, A. M. Edwards, L. Frappier, *Nature* **1997**, *385*, 176; b) M. S. Wold, *Annual review of biochemistry* **1997**, *66*, 61; c) D. L. Theobald, R. M. Mitton-Fry, D. S. Wuttke, *Annual review of biophysics and biomolecular structure* **2003**, *32*, 115; d) A. Bochkarev, E. Bochkareva, *Current opinion in structural biology* **2004**, *14*, 36.
[87] S. M. Butterfield, W. J. Cooper, M. L. Waters, *Journal of the American Chemical Society* **2005**, *127*, 24.
[88] a) M. Jager, S. Deechongkit, E. K. Koepf, H. Nguyen, J. Gao, E. T. Powers, M. Gruebele, J. W. Kelly, *Biopolymers* **2008**, *90*, 751; b) J. A. Kowalski, K. Liu, J. W. Kelly, *Biopolymers* **2002**, *63*, 111.
[89] M. Jäger, M. Dendle, J. W. Kelly, *Protein science : a publication of the Protein Society* **2009**, *18*, 1806.
[90] P. Tian, R. B. Best, *Biophysical journal* **2017**, *113*, 1719.
[91] A. E. Eriksson, T. A. Jones, A. Liljas, *Proteins* **1988**, *4*, 274.
[92] R. Wintjens, J.-M. Wieruszeski, H. Drobecq, G. Lippens, I. Landrieu, *SOLUTION STRUCTURE OF PIN1 WW DOMAIN*, **2001**.
[93] A. E. Eriksson, T. A. Jones, A. Liljas, *REFINED STRUCTURE OF HUMAN CARBONIC ANHYDRASE II AT 2.0 ANGSTROMS RESOLUTION*, **1990**.
[94] C. Lindner, A. Friemel, N. Schwegler, L. Timmermann, T. L. Pham, V. Reusche, M. Kovermann, F. Thomas, *Journal of the American Chemical Society* **2024**, *146*, 16590.
[95] W. P. Russ, D. M. Lowery, P. Mishra, M. B. Yaffe, R. Ranganathan, *Nature* **2005**, *437*, 579.
[96] M. Socolich, S. W. Lockless, W. P. Russ, H. Lee, K. H. Gardner, R. Ranganathan, *Nature* **2005**, *437*, 512.
[97] T. Wang, C. Liang, Y. Hou, M. Zheng, H. Xu, Y. An, S. Xiao, L. Liu, S. Lian, *Biotechnology letters* **2020**, *42*, 1305.
[98] G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, *Genome research* **2004**, *14*, 1188.
[99] G. E. Crooks, S. E. Brenner, "WebLogo - About", can be found under https://weblogo.berkeley.edu/, **2008**.
[100] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, *Nature methods* **2022**, *19*, 679.
[101] N. J. Greenfield, *Nature protocols* **2006**, *1*, 2876.
[102] K. Dave, M. Jäger, H. Nguyen, J. W. Kelly, M. Gruebele, *Journal of molecular biology* **2016**, *428*, 1617.
[103] D. L. Minor, P. S. Kim, *Nature* **1994**, *367*, 660.
[104] R. V. Ulijn, A. M. Smith, *Chemical Society reviews* **2008**, *37*, 664.
[105] J. S. Richardson, D. C. Richardson, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 2754.
[106] V. P. Reddy Chichili, V. Kumar, J. Sivaraman, *Protein science : a publication of the Protein Society* **2013**, *22*, 153.
[107] A. L. Cuff, R. W. Janes, A. C. R. Martin, *Bioinformatics* **2006**, *22*, 1464.
[108] a) C. M. Kraemer-Pecore, J. T. J. Lecomte, J. R. Desjarlais, *Protein science : a publication of the Protein Society* **2003**, *12*, 2194; b) O. Staub, D. Rotin, *Structure (London, England : 1993)* **1996**, *4*, 495.
[109] M. H. AL Mughram, C. Catalano, N. B. Herrington, M. K. Safo, G. E. Kellogg, *Frontiers in Molecular Biosciences* **2023**, *10*, 1116868.
[110] a) D. E. Schlamadinger, J. E. Gable, J. E. Kim, *The journal of physical chemistry. B* **2009**, *113*, 14769; b) J. Shao, B. P. Kuiper, A.-M. W. H. Thunnissen, R. H. Cool, L. Zhou, C. Huang, B. W. Dijkstra, J. Broos, *Journal of the American Chemical Society* **2022**, *144*, 13815; c) S. Majumder, S.

Khamrui, R. Banerjee, P. Bhowmik, U. Sen, *Biochimica et biophysica acta* **2015**, *1854*, 55; d) H. Zhao, C. Liu, W. Ding, L. Tang, Y. Fang, Y. Chen, L. Hu, Y. Yuan, D. Fang, S. Lin, *Journal of the American Chemical Society* **2022**, *144*, 6742.

[111] R. Subirós-Funosas, A. El-Faham, F. Albericio, *Tetrahedron* **2011**, *67*, 8595.

[112] T. L. Pham, J. Zilke, C. C. Müller, F. Thomas, *RSC chemical biology* **2022**, *3*, 426.

[113] R. Wu, F. Ding, R. Wang, R. Shen, X. Zhang, S. Luo, C. Su, Z. Wu, Q. Xie, B. Berger et al., *High-resolution de novo structure prediction from primary sequence*, **2022**.

[114] F. Chen, Q. Wang, Y. Mu, S. Sun, X. Yuan, P. Shang, B. Ji, *Journal of molecular recognition : JMR* **2022**, *35*, e2947.

[115] T. Kortemme, D. E. Kim, D. Baker, *Science's STKE : signal transduction knowledge environment* **2004**, *2004*, pl2.

[116] T. Kortemme, D. Baker, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 14116.

[117] M. J. Macias, E. Aragon, N. Goerner, Q. Xi, T. Lopes, S. Gao, J. Massague, *YAP WW1 in complex with a Smad7 derived peptide*, **2012**.

[118] M. Meiyappan, G. Birrane, J. Ladias, *Crystal structure of human FE65-WW domain in complex with human Mena peptide*, **2007**.

[119] R. Behrendt, P. White, J. Offer, *Journal of Peptide Science* **2016**, *22*, 4.

[120] T. L. Pham, M. R. Conde González, S. Fazliev, A. Kishore, P. Comba, F. Thomas, *Chembiochem : a European journal of chemical biology* **2024**, *25*, e202300715.

[121] a) C. Webb, A. Upadhyay, M. Furutani-Seiki, S. Bagby, *Yap ww2*, **2010**; b) S. Ohnishi, N. Tochio, M. Sato, S. Koshiba, T. Harada, S. Watanabe, T. Kigawa, S. Yokoyama, *Solution structure of the WW domain from the human growth-arrest-specific protein 7, GAS-7*, **2007**.

[122] C. Webb, A. Upadhyay, F. Giuntini, I. Eggleston, M. Furutani-Seiki, R. Ishima, S. Bagby, *Biochemistry* **2011**, *50*, 3300.

[123] M. A. Verdecia, M. E. Bowman, K. P. Lu, T. Hunter, J. P. Noel, *Nature structural biology* **2000**, *7*, 639.

[124] a) T. J. Simons, *Journal of biochemical and biophysical methods* **1993**, *27*, 25; b) A. Meni, E. T. Yukl, *Biomolecules* **2020**, *10*.

[125] P. Kuzmic, *Analytical biochemistry* **1996**, *237*, 260.

[126] J. Eberhardt, D. Santos-Martins, A. F. Tillack, S. Forli, *Journal of chemical information and modeling* **2021**, *61*, 3891.

[127] O. Trott, A. J. Olson, *Journal of computational chemistry* **2010**, *31*, 455.

[128] C. Arter, L. Trask, S. Ward, S. Yeoh, R. Bayliss, *The Journal of biological chemistry* **2022**, *298*, 102247.

[129] a) C. S. Yuan, D. B. Ault-Riché, R. T. Borchardt, *The Journal of biological chemistry* **1996**, *271*, 28009; b) G. Santiago, M. Martínez-Martínez, S. Alonso, R. Bargiela, C. Coscolín, P. N. Golyshin, V. Guallar, M. Ferrer, *Biochemistry* **2018**, *57*, 2245; c) J. Munoz-Munoz, A. Cartmell, N. Terrapon, B. Henrissat, H. J. Gilbert, *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114*, 4936.

[130] N. J. Greenfield, *Nature protocols* **2006**, *1*, 2527.

[131] J. C. Aledo, *Protein science : a publication of the Protein Society* **2019**, *28*, 1785.

[132] a) D. Lüthi, D. Günzel, J. A. S. McGuigan, *Experimental Physiology* **1999**, *84*, 231; b) C. Andreini, I. Bertini, G. Cavallaro, G. L. Holliday, J. M. Thornton, *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry* **2008**, *13*, 1205; c) J. J. Lacapère, N. Bennett, Y. Dupont, F. Guillain, *Journal of Biological Chemistry* **1990**, *265*, 348; d) A. Bojovschi, M. S. Liu, R. J. Sadus, *The Journal of chemical physics* **2014**, *140*, 115102.

[133] R. Yang, Y. Hou, C. A. Campbell, K. Palaniyandi, Q. Zhao, A. J. Bordner, X. Chang, *Biochimica et biophysica acta* **2011**, *1808*, 1790.

[134] M. S. Wang, M. H. Hecht, *Journal of the American Chemical Society* **2020**, *142*, 15230.
[135] M. Pfeiffer, R. M. Crean, C. Moreira, A. Parracino, G. Oberdorfer, L. Brecker, F. Hammerschmidt, S. C. L. Kamerlin, B. Nidetzky, *ACS catalysis* **2022**, *12*, 3357.
[136] A. T. Juarez-Facio, V. Martin de Lagarde, C. Monteil, J.-M. Vaugeois, C. Corbiere, T. Rogez-Florent, *Molecules (Basel, Switzerland)* **2021**, *26*.
[137] A. M. Fernández-Escamilla, S. Ventura, L. Serrano, M. A. Jiménez, *Protein science : a publication of the Protein Society* **2006**, *15*, 2278.
[138] V. V. Khrustalev, E. V. Barkovsky, T. A. Khrustaleva, *Journal of Theoretical Biology* **2016**, *395*, 174.
[139] a) P. Ciancaglini, J. M. Pizauro, F. A. Leone, *Journal of inorganic biochemistry* **1997**, *66*, 51; b) R. González, N. Carvajal, A. Morán, *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **1984**, *78*, 389.
[140] a) R. C. Diehl, R. S. Chorghade, A. M. Keys, M. M. Alam, S. A. Early, A. E. Dugan, M. Krupkin, K. Ribbeck, H. J. Kulik, L. L. Kiessling, *JACS Au* **2024**; b) W. Chen, S. Enck, J. L. Price, D. L. Powers, E. T. Powers, C.-H. Wong, H. J. Dyson, J. W. Kelly, *Journal of the American Chemical Society* **2013**, *135*, 9877.
[141] Z. Zhao, X. Xu, H. Cheng, M. C. Miller, Z. He, H. Gu, Z. Zhang, A. Raz, K. H. Mayo, G. Tai et al., *Proceedings of the National Academy of Sciences of the United States of America* **2021**, *118*.
[142] Y. van Kooyk, G. A. Rabinovich, *Nature immunology* **2008**, *9*, 593.
[143] a) D. Hudson, *Journal of combinatorial chemistry* **1999**, *1*, 333; b) F. García-Martín, M. Quintanar-Audelo, Y. García-Ramos, L. J. Cruz, C. Gravel, R. Furic, S. Côté, J. Tulla-Puche, F. Albericio, *Journal of combinatorial chemistry* **2006**, *8*, 213.
[144] T. Michels, R. Dölling, U. Haberkorn, W. Mier, *Organic letters* **2012**, *14*, 5218.
[145] a) CEM Corporation (Ed.), **2018**; b) CEM (Ed.), **2018**.
[146] J. M. Collins, K. A. Porter, S. K. Singh, G. S. Vanier, *Organic letters* **2014**, *16*, 940.
[147] S. B. Kent, *Annual review of biochemistry* **1988**, *57*, 957.
[148] B. J. H. Kuipers, H. Gruppen, *Journal of agricultural and food chemistry* **2007**, *55*, 5445.
[149] H. Edelhoch, *Biochemistry* **1967**, *6*, 1948.
[150] H. Mach, C. R. Middaugh, R. V. Lewis, *Analytical biochemistry* **1992**, *200*, 74.
[151] J. A. Verpoorte, S. Mehta, J. T. Edsall, *Journal of Biological Chemistry* **1967**, *242*, 4221.
[152] B. Zhou, Z. Y. Zhang, *Journal of Biological Chemistry* **1999**, *274*, 35526.
[153] GitHub, "GitHub - HeliXonProtein/OmegaFold: OmegaFold Release Code", can be found under https://github.com/HeliXonProtein/OmegaFold, **2024**.
[154] S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell, A. J. Olson, *Nature protocols* **2016**, *11*, 905.

8. Appendix

8.1 List of Abbreviations

a.u.	arbitary units
Ac	acetyl
ADP	adenosine diphosphate
Ala	alanine
AMP	adenosine monophosphate
aPP	avian pancreatic polypeptide
Arg	arginine
Asn	asparagine
Asp	aspartate
ATP	adenosine triphosphate
AUC	area under the curve
bPP	bovine pancreatic polypeptide
BS	beta sheet
c	concentration
CA	carbonic anhydrase
cAMP	cyclic adenosine monophosphate
CC	coiled coil
CD	Circular dichroism
CD4	Cluster of differentiation 4
CDPs	Cystine-dense peptides
CT	C-terminus
Cys	cysteine
DCM	Dichloromethane
DIC	N,N'-Diisopropylcarbodiimide
DIPEA	N,N-Diisopropylamine
DMF	N,N-Dimethylformamide
DNA	Deoxyribonucleic acid
Eq.	equation
et al.	*et alia*
FBP11	formin binding protein 11
Fmoc	Fluorenylmethoxycarbonyl
Gas7	growth arrest specific 7
Gln	glutamine
Glu	glutamate
Gly	glycine
H	helix
h	hour
h	*human*
HFIP	*1,1,1,3,3,3*-Hexafluoroisopropanol
His	histidine
HIV	human immune deficiency virus type 1

HP36	villin headpiece subdomain 36
*h*Pin1	protein interacting with never in mitosis A 1
HPLC	High performance liquid chromatography
HS	highly stable
HSQC	Heteronuclear Single Quantum Correlation
Ile	isoleucine
JACS	Journal of the American Chemical Society
K	kelvin
k_2	catalytical efficiency
kB1	kalata B1
kcal	kilocalorie
K_d	dissociation constant
kDa	kilo Dalton
KH_2PO_4	potassium dihydrogen phosphate buffer
l	optical path length
LDL	low-density lipoprotein
Leu	leucine
LP	Loop
Lys	lysine
M	molar
MALDI-TOF	matrix-assisted laser desorption ionization time-of-flight
MBHA	4-methylbenzylhydrylamine
MCR4	melanocortin receptor 4
MeCN	Acetonitrile
min	minute
ml	milliliter
mM	milli molar
mmol	millimole
MOG	myelin oligodendrocyte glycoprotein
MOPS	3-(N-Morpholino)propanesulfonic acid
MOPS	3-(*N*-morpholino) propanesulfonic acid
MRE	mean residue ellipticity
MS	multiple sclerosis
MSAs	multiple sequence alignments
MSH	melanocyte-stimulating hormones
n	amount
Na_2HPO_4	Disodium hydrogen phosphate
NHB	non-hydrogen-bonded
Nle	norleucine
nm	nanometer
NMR	nuclear magnetic resonance
NT	N-terminus
OB	oligonucleotide/oligosaccharide binding
Oxyma	Ethyl-2-cyano-2-(hydroxyimino) acetate
PC	phosphorylcholine

PCR	polymerase chain reaction
PDB	protein data bank
Phe	phenylalanine
p-NPP	*para*-nitrophenyl phosphate
PPI	protein-protein interaction
Pro	proline
RP-HPLC	Reversed-phase high performance liquid chromatography
rpm	revolutions per minute
s	second (time)
S	strand
SCA	statistical coupling analysis
SD	standard deviation
SE	standard error
Ser	serine
SH3	Src homology 3
SPPS	solid-phase peptide synthesis
ss	single-stranded
SSEs	secondary structural elements
SSSRs	sequence-structure-stability-relationships
T	*Tesla*
tBu	*tert*-Butyl
TCF4	transciption factor 4
TFA	Trifluoroacetic acid
Thr	threonine
TIPS	Triisopropylsilane
T_m	melting temperature
TOCSY	Total correlation spectroscopy
Trp	tryptophan
Tyr	Tyrosine
v_{obs}	reaction rate
Wnt	Wingless and INT-1
WW-CA	WW-carbonic anhydrase
WW-CS	WW-consensus sequence
WW-HS	WW-highly stable
y	*yeast*
YAP	Yes-kinase-associated protein
Δ*H*	molar enthalpy
(*v/v*)	*volume percent*
°C	degree Celsius
µM	micro molar
2D	Two dimensional

8.2 Sequence alignment of native WW domains

Table S1. Sequence alignment of the listed 90 WW domains used for the design of WW-CS. This table was reused from Lindner *et al.*[94], with permission from © 2024 the American Chemical Society.

WW domain	Sequence	Identifier
*h*APB3-WW	GLPPGWRKIHD-AAG-TYYWHVPSGSTQWQRPTW	2YSC[a]
*h*WWFE65	DLPAGWMRVQD-TSG-TYYWHIPTGTTQWEPPGR	2HO2[a]
*m*SAV1-WW2	GLPPGWERVES-SEFGTYYVDHTNKRAQYRHPCA	7BQF[a]
*h*HypB-WW	VLPPNWKTARD-PEGKIYYYHVITRQTQWDPPTW	2MDI[a]
*h*MAGI-1-WW2	ELPAGWEKIED-PVYGIYYVDHINRKTQYENPVL	2YSE[a]
*m*MAGI-2-WW2	ELPYGWEKIDD-PIYGTYYVDHINRRTQFENPVL	6JJZ[a]
*h*Gas7-WW	-LPPGWQSYLS-PQGRRYYVNTTTNETTWERPSS	2YSH[a]
*m*Gas7-WW	ILPPGWHSYLS-PQGRRYYVNTTTNETTWERPSS	B1ATI9[b,1]
*m*SAV1-WW1	PLPPGWSVDWT-MRGRKYYIDHNTNTTHWSPLES	7BQF[a]
*m*Nedd4-2-WW2	GLPSGWEERKD-AKGRTYYVNHNNRTTTWTRPIM	1WR4[a]
*m*Nedd4-2-WW1	PLPPGWEEKVD-NLGRTYYVNHNNRSTQWHRPSL	1WR3[a]
*h*Nedd4-2-WW3	PLPPGWEEKVD-NLGRTYYVNHNNRTTQWHRPSL	7LP3[a]
*h*Nedd4-1-WW1	PLPPGWEERQD-ILGRTYYVNHESRRTQWKRPTP	P46934[b,2]
*r*Nedd4-WW1	PLPPGWEERQD-VLGRTYYVNHESRTTQWKRPSP	2N8S[a]
*h*Itch-WW1	GLPPGWEQRVD-QHGRVYYVDHVEKRTTWDRPSG	2DMV[a]
*m*Itch-WW1	PLPPGWEQRVD-QHGRVYYVDHVEKRTTWDRPEP	5XMC[a]
*d*Su(dx)-WW1	PLPAGWEIRLD-QYGRRYYVDHNTRSTYWEKPTP	Q9Y0H4[b,1]
*h*Nedd4-1-WW2	GLPPGWEEKQD-ERGRSYYVDHNSRTTTWTKPTV	P46934[b,2]
*r*Nedd4-WW2	GLPPGWEEKQD-DRGRSYYVDHNSKTTTWSKPTV	2N8U[a]
*d*Su(dx)-WW2	PLPPGWEIRKD-GRGRVYYVDHNTRKTTWQRPNS	Q9Y0H4[b,1]
*h*WWP2-WW2	PLPPGWEKRTD-PRGRFYYVDHNTRTTTWQRPTA	6J1Z[a]
*m*Itch-WW2	PLPPGWERRVD-NMGRIYYVDHFTRTTTWQRPTL	5XMC[a]
*h*WWP1-WW1	PLPPGWERRVD-DRRRVYYVDHNTRTTTWQRPTM	6J1X[a]
*sp*Dre4-WW2	PNNDSWVVVFT-KKNRYFFHNLKSHESYWEPPLE	Q09685[b,1]
*h*WWOX-WW2	DLPYGWEQETD-ENGQVFFVDHINKRTTYLDPRL	1WMV[a]
*h*BAG-3-WW1	PLPPGWEIKIDPQTGWPFFVDHNSRTTTWNDPRV	O95817[b,1]
*h*HECW1-2-WW	ELPRGWEIKTD-QQGKSFFVDHNSRATTFIDPRI	3L4H[a]
*h*Dmd-WW	SVQGPWERAIS-PNKVPYYINHETQTTSWDHPKM	1EG3[a]
*h*DRP2-WW	SVQVPWERAIS-PNKVPYYINHQAQTTCWDHPKM	Q13474[b,1]
*h*STXBP4-WW	-LPYGWEEAYT-ADGIKYFINHVTQTTSWIHPVM	2YSG[a]
*m*KIBRA-WW2	ELPEGWEEAYD-PQVGDYFIDHNTKTTQIEDPRV	6JJW[a]
*m*KIBRA-WW1	PLPEGWEEARD-FDGKVYYIDHRNRTTSWIDPRD	6JJW[a]
*h*MAGI-1-WW1	PLPENWEMAYT-ENGEVYFIDHNTKTTSWLDPRC	2YSD[a]
*m*MAGI-2-WW1	PLPDNWEMAYT-EKGEVYFIDHNTKTTSWLDPRL	6JJZ[a]
*h*YAP-WW2	PLPDGWEQAMT-QDGEIYYINHKNKTTSWLDPRL	2LAW[a]
*y*YAP1-WW2	PLPEGWEQAIT-PEGEIYYINHKNKTTSWLDPRL	2L4J[a]
*h*Smurf1-WW1	ELPEGYEQRTT-VQGQVYFLHTQTGVSTWHDPRI	2LB0[a]
*h*Smurf2-WW2	DLPEGYEQRTT-QQGQVYFLHTQTGVSTWHDPRV	2KXQ[a]
*h*YAP-WW1	PLPAGWEMAKT-SSGQRYFLNHIDQTTTWQDPRK	2LAX[a]
*d*Su(dx)-WW3	PLPDGWEKKIQ-SDNRVYFVNHKNRTTQWEDPRT	1TK7[a]
*m*Itch-WW3	PLPPGWEKRTD-SNGRVYFVNHNTRITQWEDPRS	1YIU[a]
*h*WWP1-WW2	PLPPGWEKRVD-STDRVYFVNHNTKTTQWEDPRT	6J1X[a]
*h*Nedd4-2-WW4	PLPPGWEERIH-LDGRTFYIDHNSKITQWEDPRL	Q96PU5[b,1]
*h*Nedd4-1-WW4	PLPPGWEERTH-TDGRIFYINHNIKRTQWEDPRL	P46934[b,1]
*r*Nedd4-WW3	PLPPGWEERTH-TDGRVFFINHNIKKTQWEDPRM	1I5H[a]
*d*Nedd4-WW3	PLPPRWAMQVA-PNGRTFFIDHASRRTTWIDPRN	2EZ5[a]
*y*HECT-E3-WW	PLPSGWEMRLT-NTARVYFVDHNTKTTTWDDPRL	3OLM[a]
*h*Smurf1-WW2	PLPPGWEVRST-VSGRIYFVDHNNRTTQFTDPRL	2LB1[a]
*h*Smurf2-WW3	PLPPGWEIRNT-ATGRVYFVDHNNRTTQFTDPRL	2DJY[a]
*h*Itch-WW4	GLPEGWEMRFT-VDGIPYFVDHNRRTTTYIDPRT	2YSF[a]
*pa*Itch-WW4	KLPEGWEMRFT-VDGIAYFVDHNRRTTTYIDPRT	A0A096NUS9[b,1]
*h*Nedd4-1-WW3	FLPKGWEVRHA-PNGRPFFIDHNTKTTTWEDPRK	4N7F[a]
*m*Nedd4-2-WW3	FLPPGWEMRIA-PNGRPFFIDHNTKTTTWEDPRL	1WR7[a]
*sp*Pub1-WW2	ELPPGWEQRYT-PEGRPYFVDHNTRTTTWVDPRR	Q92462[b,1]
*d*Su(dx)-WW4	PLPPGWEIRYT-AAGERFFVDHNTRRTTFEDPRP	1TK7[a]
*h*WWP1-WW3	PLPEGWEIRYT-REGVRYFVDHNTRTTTFKDPRN	6J1X[a]
*h*WWP2-WW4	ALPPGWEMKYT-SEGVRYFVDHNTRTTTFKDPRP	6RSS[a]
*h*CMTR1-WW	TVNEPWTMGFSKSFKKKFFYNKKTKDSTFDLPAD	Q8N1G2[b,2]
*t*Db10-WW	TLPKPWKGLVDGTTGFIYFWNPETNDTQYERPVP	P46942[b,2]

*ca*ESS1-WW	GLPPNWTIRVSRSHNKEYFLNQSTNESSWDPPYG	1YW5[a]
*y*ESS1-WW	GLPTPWTVRYSKSKKREYFFNPETKHSQWEEPEG	7KKF[a]
*an*PinA-WW	GLPAGWEVRHSNSKNLPYYFNPATRESRWEPPAD	2JV4[a]
*nc*Spp-1-WW	GLPEDWEVRHSQSKNLPYYFNSATKTSRWEPPSG	O60045[b,1]
*h*PCIF1-WW	LVHAGWEKCWSRRENRPYYFNRFTNQSLWEMPVL	2JX8[a]
*d*Dod-WW	QLPDGWEKRTSRSTGMSYYLNMYTKESQWDQPTE	P54353[b,1]
*h*Pin1-WW	KLPPGWEKRMSRSSGRVYYFNHITNASQWERPSG	2KCF[a]
*y*YJQ8-WW	RLPPGWEIIHE-NGRPLYYNAEQKTKLHYPPYGY	1E0N[a]
*h*TCERG1-WW	IPGTPWCVVWT-GDERVFFYNPTTRLSMWDRPDD	2DK7[a]
*ca*Yo61-WW1	SVESDWSVHTN-EKGTPYYHNRVTKQTSWIKPDV	P34600[b,2]
*h*DGCR8-WW	PLPDGWIMTFH-NSGVPVYLHRESRVVTWSRPYF	3LE4[a]
*ac*L8H7M2-WW	ASVDGWKQYFT-AEGNAYYYNEVSGETSWDPPSS	L8H7M2[b,2]
*sp*Dre4-WW1	PLPPGWTEHKA-PSGIPYYWNAELKKSTYQRPSF	Q09685[b,1]
*h*HypA/FBP11-WW2	LSKCPWKEYKS-DSGKTYYYNSQTKESRWAKPKE	2L5F[a]
*m*HypA/FBP11-WW2	LSKCPWKTYKS-DSGKPYYYNSQTKESRWAKPKE	Q9R1C7[b,2]
*pa*PRPF40B-WW2	LSQCPWKEYKS-DTGKPYYYNNQSKESRWTRPKD	A0A096MRB5[b,1]
*ca*Yo61-WW2	PQQGQWKEFMS-DDGKPYYYNTLTKKTQWVKPDG	P34600[b,2]
*m*FBP21-WW1	ASKGGWVEGVT-ADGHCYYYDLITGASQWEKPEG	Q61048[b,1]
*h*FBP21-WW1	PSKGRWVEGIT-SEGYHYYYDLISGASQWEKPEG	2JXW[a]
*h*WBP-4-WW	----RWVEGIT-SEGYHYYYDLISGASQWEKPEG	2DK1[a]
*y*YPR152C	-MRGEWQEFKT-PAGKKYYYNKNTKQSRWEKPNL	Q06525[b,2]
*pa*PRPF40B-WW1	PPRALWSEHVA-PDGRIYYYNADDKQSVWEKPSV	A0A096MRB5[b,1]
*y*Prp40-WW2	LRENGWKAAKT-ADGKVYYYNPTTRETSWTIPAF	1O6W[a]
*m*CA150-WW1	-TEEIWVENKT-PDGKVYYYNARTRESAWTKPDG	2YSI[a]
*m*FBP28-WW2	TAVSEWTEYKT-ADGKTYYYNNRTLESTWEKPQE	1E0L[a]
*y*Prp40-WW1	--MSIWKEAKD-ASGRIYYYNTLTKKSTWEKPKE	1O6W[a]
*h*FBP21-WW2	AVKTVWVEGLS-EDGFTYYYNTETGESRWEKPDD	2JXW[a]
*m*FBP21-WW2	AAKAVWVEGLS-EDGYTYYYNTETGESKWEKPED	Q61048[b,1]
*h*HypA/FBP11-WW1	GAKSMWTEHKS-PDGRTYYYNTETKQSTWEKPDD	2L5F[a]
Prototype	GLPPGWDEYKT-HNGKTYYYNHNTKTSTWTDPRM	1E0M[a,c]
CC45	PLPPGWERRTD-VEGKVYYFNVRTLTTTWERPTI	1YMZ[a,d]

8.3 Analytical HPLC chromatograms and MALDI-TOF-MS spectra

8.3.1 Generation of a WW domain consensus sequence and iterative sequence modification

Table S2: WW domain screening peptides. Calculated mass and identified mass of MALDI-TOF MS spectrum, the RP-HPLC linear gradient, and characteristic retention time t_R of each peptide is summarized.

Peptide name	$[M+H]^+$ (calc.)	$[M+H]^+$ (found)	Percentage of HPLC Buffer B in A (*v*/*v*)	t_R [min]
WW-CS	4175.1	4174.1	20-50 %	16.0
WW-BS1-1	4178.2	4176.6	25-50 %	16.1
WW-BS1-2	4178.2	4177.4	25-50 %	16.8
WW-BS1-3	4166.1	4165.1	25-50 %	13.1
WW-BS1-4	4150.2	4149.0	25-50 %	15.7
WW-BS1-5	4151.1	4150.4	25-50 %	17.4
WW-BS1-6	4149.2	4148.0	25-50 %	13.6
WW-BS1-7	4149.2	4148.1	25-50 %	17.4
WW-LP1-1	3972.0	3971.6	25-50 %	19.1
WW-LP1-2	4071.1	4070.1	25-50 %	15.9
WW-LP1-3	4099.1	4098.2	25-50 %	17.6
WW-LP1-4	4113.1	4112.4	25-50 %	17.9
WW-LP1-5	3972.0	3971.0	25-50 %	19.2
WW-LP1-6	4099.1	4098.0	25-50 %	16.3
WW-LP1-7	4113.1	4113.2	25-50 %	17.4
WW-BS2-1	3988.0	3987.0	25-50 %	17.7
WW-BS2-2	3912.0	3911.6	25-50 %	17.1
WW-BS2-3	3956.0	3954.8	25-50 %	20.3
WW-BS2-4	3896.0	3895.5	25-50 %	18.7
WW-LP2-1	3867.0	3865.7	25-50 %	18.1
WW-LP2-2	3909.0	3907.9	25-50 %	18.6
WW-LP2-3	3924.0	3923.2	25-50 %	18.0
WW-LP2-4	3938.0	3937.7	25-50 %	18.0
WW-LP2-5	3937.0	3936.1	25-50 %	21.1
WW-LP2-6	3937.0	3936.0	25-50 %	20.6
WW-LP2-7	3836.9	3837.3	25-50 %	19.1
WW-LP2-8	3864.0	3862.9	25-50 %	18.6
WW-LP2-9	3907.0	3906.3	25-50 %	17.5
WW-BS3-1	3923.0	3921.8	25-50 %	20.3
WW-BS3-2	3950.0	3948.8	25-50 %	20.7
WW-BS3-3	3907.0	3906.3	25-50 %	21.3
WW-BS3-4	3965.0	3964.0	25-50 %	21.3
WW-BS3-5	3951.0	3950.1	25-50 %	20.4
WW-BS3-6	3935.0	3933.9	25-50 %	21.0
WW-BS3-7	3977.1	3977.3	25-50 %	23.6
WW-CT-1	3822.0	3821.2	25-50 %	19.5
WW-CT-2	3936.0	3935.0	25-50 %	20.9
WW-CT-3	3667.8	3666.8	25-50 %	21.0

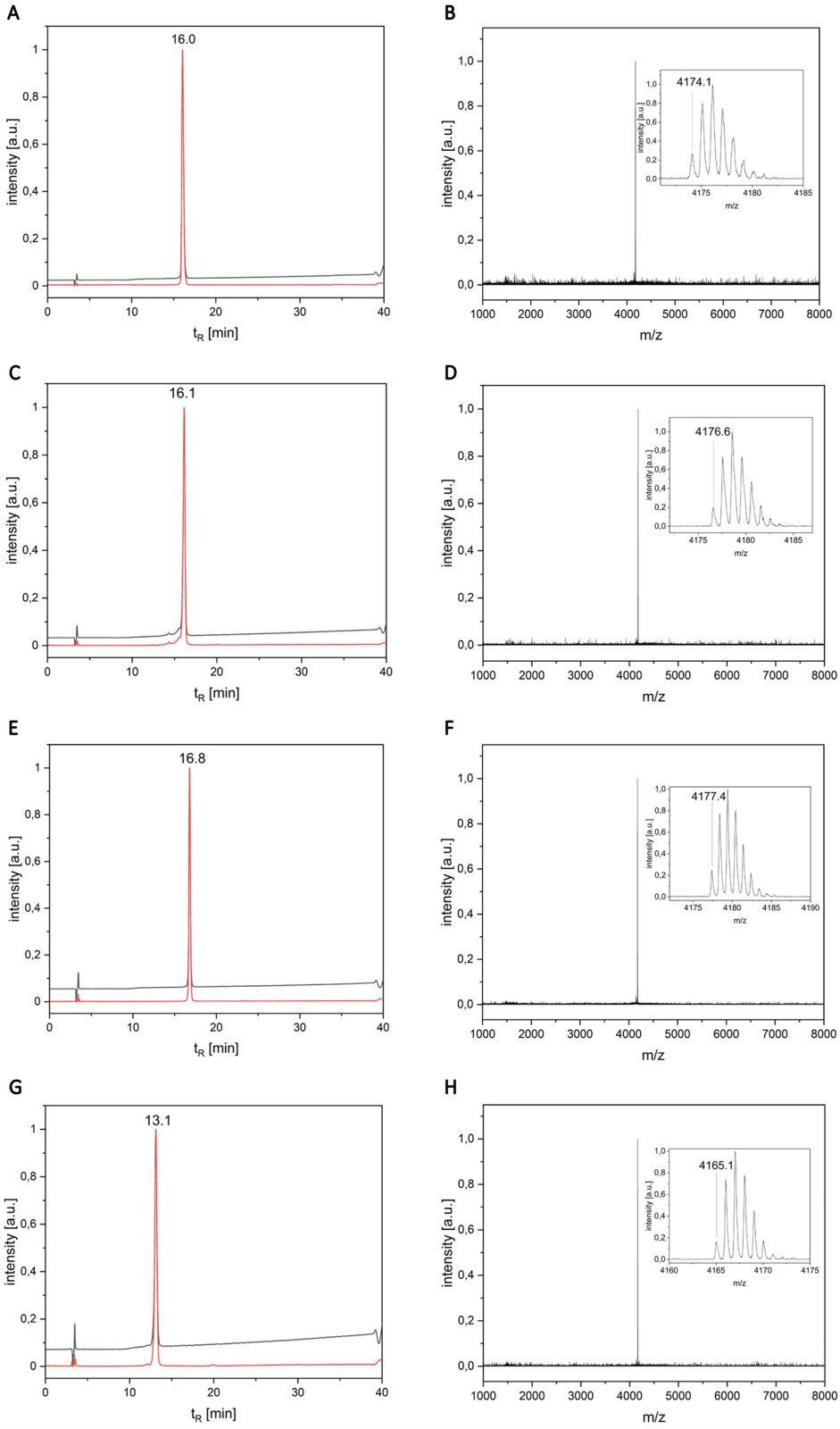


Figure S1: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-CS, **C**) WW-BS1-1, **E**) WW-BS1-2, and **G**) WW-BS1-3. MALDI-TOF-MS spectra of **B**) WW-CS, **D**) WW-BS1-1, **F**) WW-BS1-2, and **H**) WW-BS1-3. This figure is reprinted with permission from Lindner *et al.*[94], © 2024 the American Chemical Society.

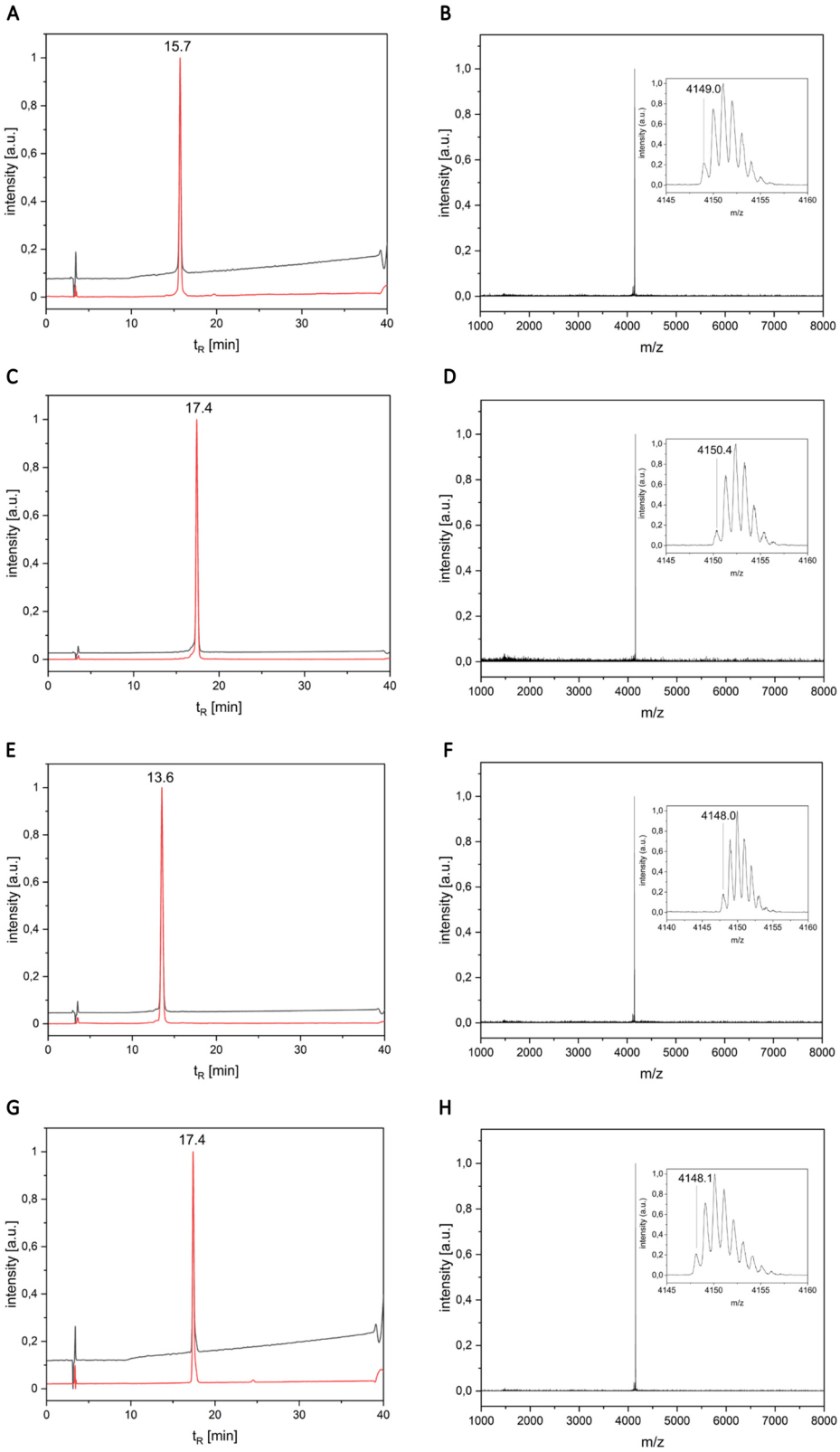


Figure S2: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-BS1-4, **C**) WW-BS1-5, **E**) WW-BS1-6, and **G**) WW-BS1-7. MALDI-TOF-MS spectra of **B**) WW-BS1-4, **D**) WW-BS1-5, **F**) WW-BS1-6, and **H**) WW-BS1-7. This figure is reprinted with permission from Lindner *et al.*[94], © 2024 the American Chemical Society.

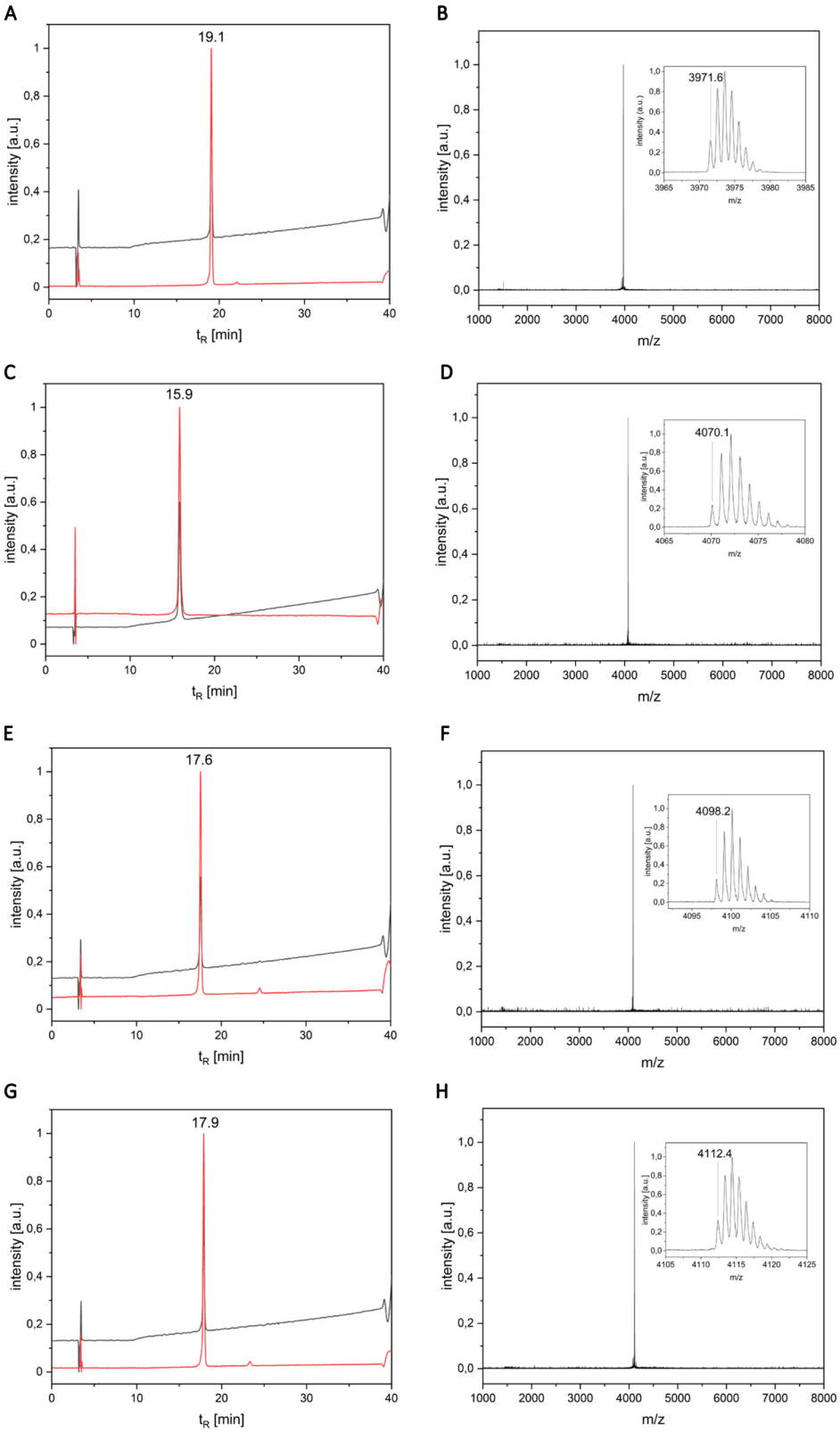


Figure S3: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-LP1-1, **C**) WW-LP1-2, **E**) WW-LP1-3, and **G**) WW-LP1-4. MALDI-TOF-MS spectra of **B**) WW-LP1-1, **D**) WW-LP1-2, **F**) WW-LP1-3, and **H**) WW-LP1-4. This figure is reprinted with permission from Lindner *et al.*[94], © 2024 the American Chemical Society.

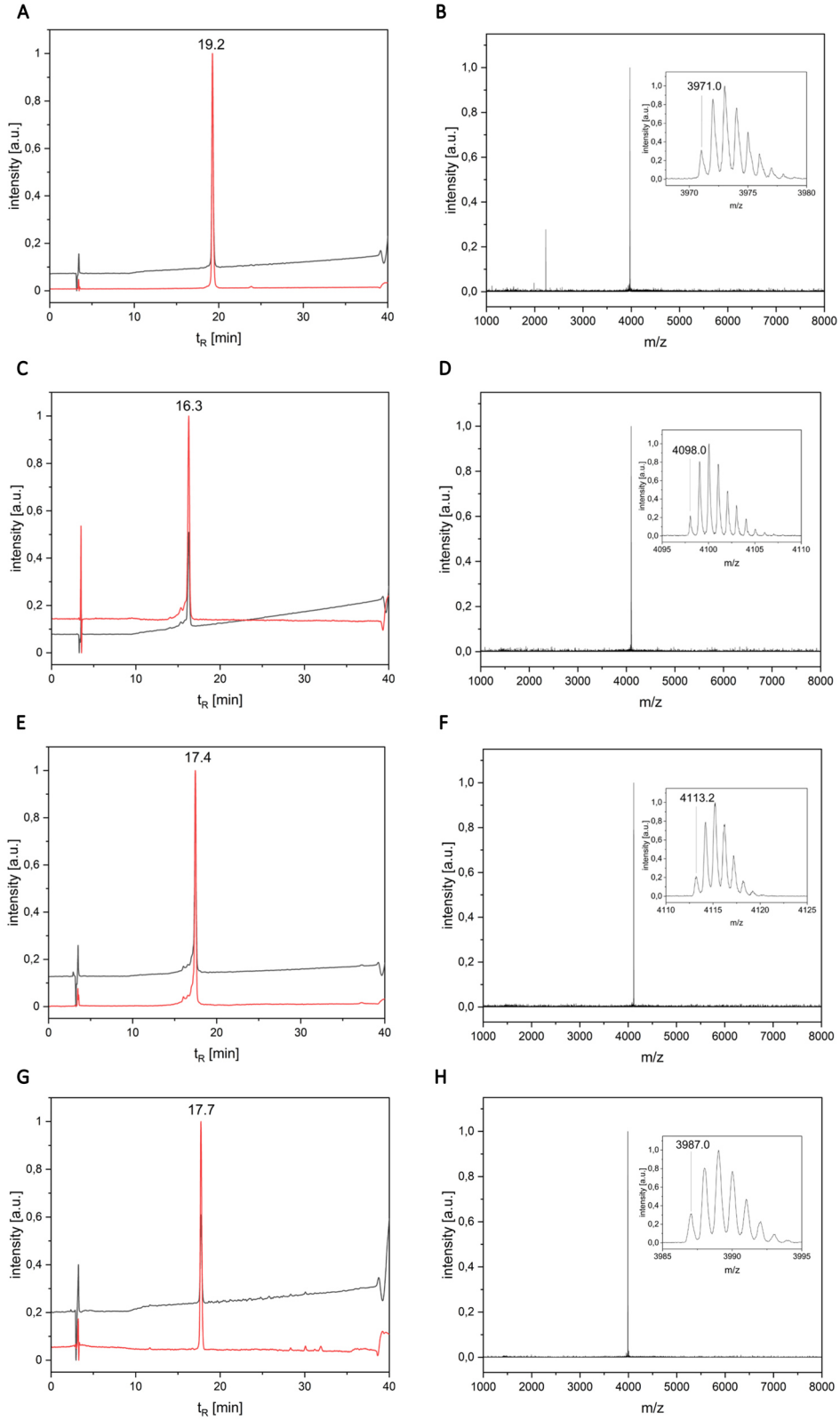


Figure S4: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-LP1-5, **C**) WW-LP1-6, **E**) WW-LP1-7, and **G**) WW-BS2-1. MALDI-TOF-MS spectra of **B**) WW-LP1-5, **D**) WW-LP1-6, **F**) WW-LP1-7, and **H**) WW-BS2-1. This figure was adapted and reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

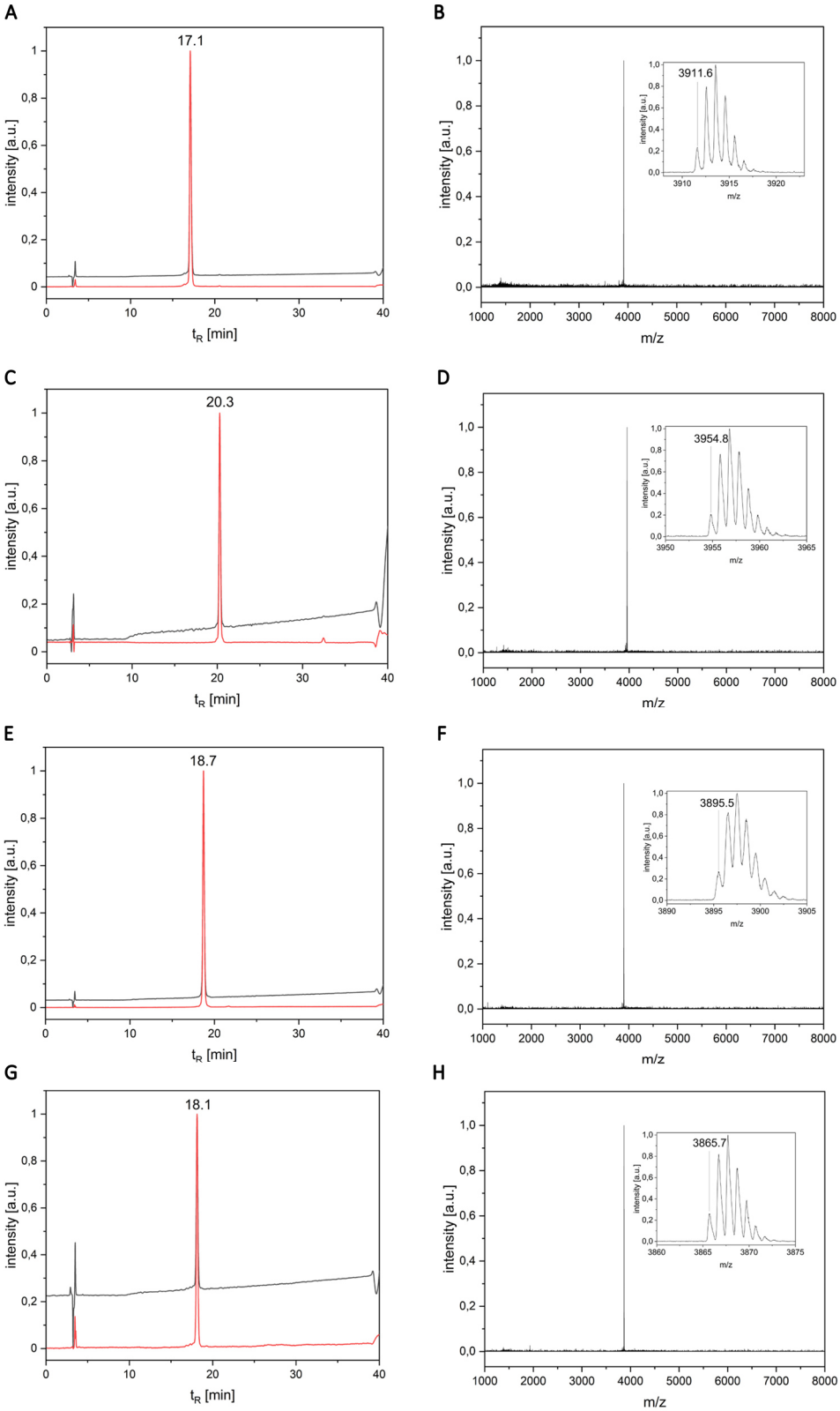


Figure S5: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-BS2-2, **C**) WW-BS2-3, **E**) WW-BS2-4, and **G**) WW-LP2-1. MALDI-TOF-MS spectra of **B**) WW-BS2-2, **D**) WW-BS2-3, **F**) WW-BS2-4, and **H**) WW-LP2-1. This figure was adapted and reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

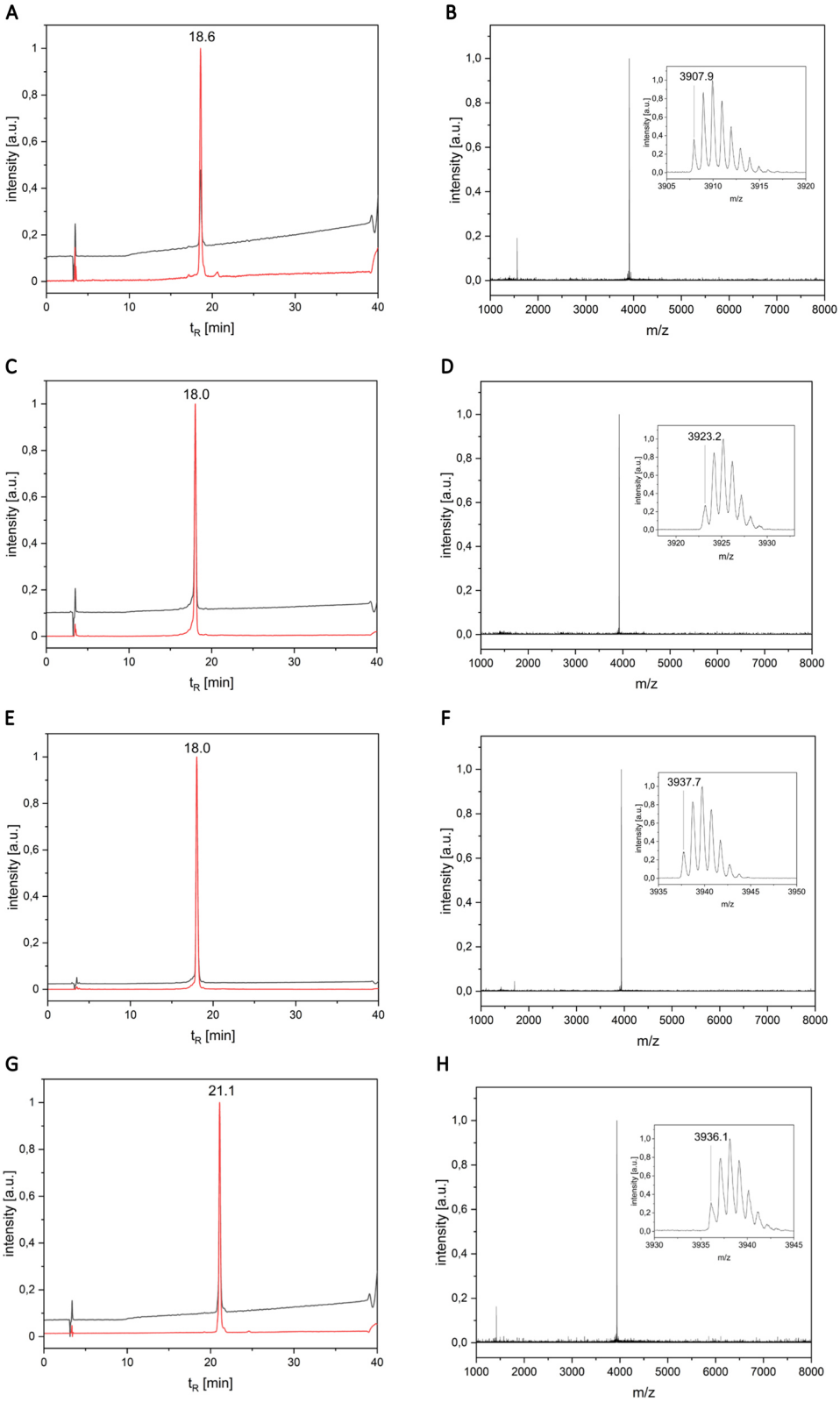


Figure S6: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-LP2-2, **C**) WW-LP2-3, **E**) WW-LP2-4, and **G**) WW-LP2-5. MALDI-TOF-MS spectra of **B**) WW-LP2-2, **D**) WW-LP2-3, **F**) WW-LP2-4, and **H**) WW-LP2-5. This figure was adapted and reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

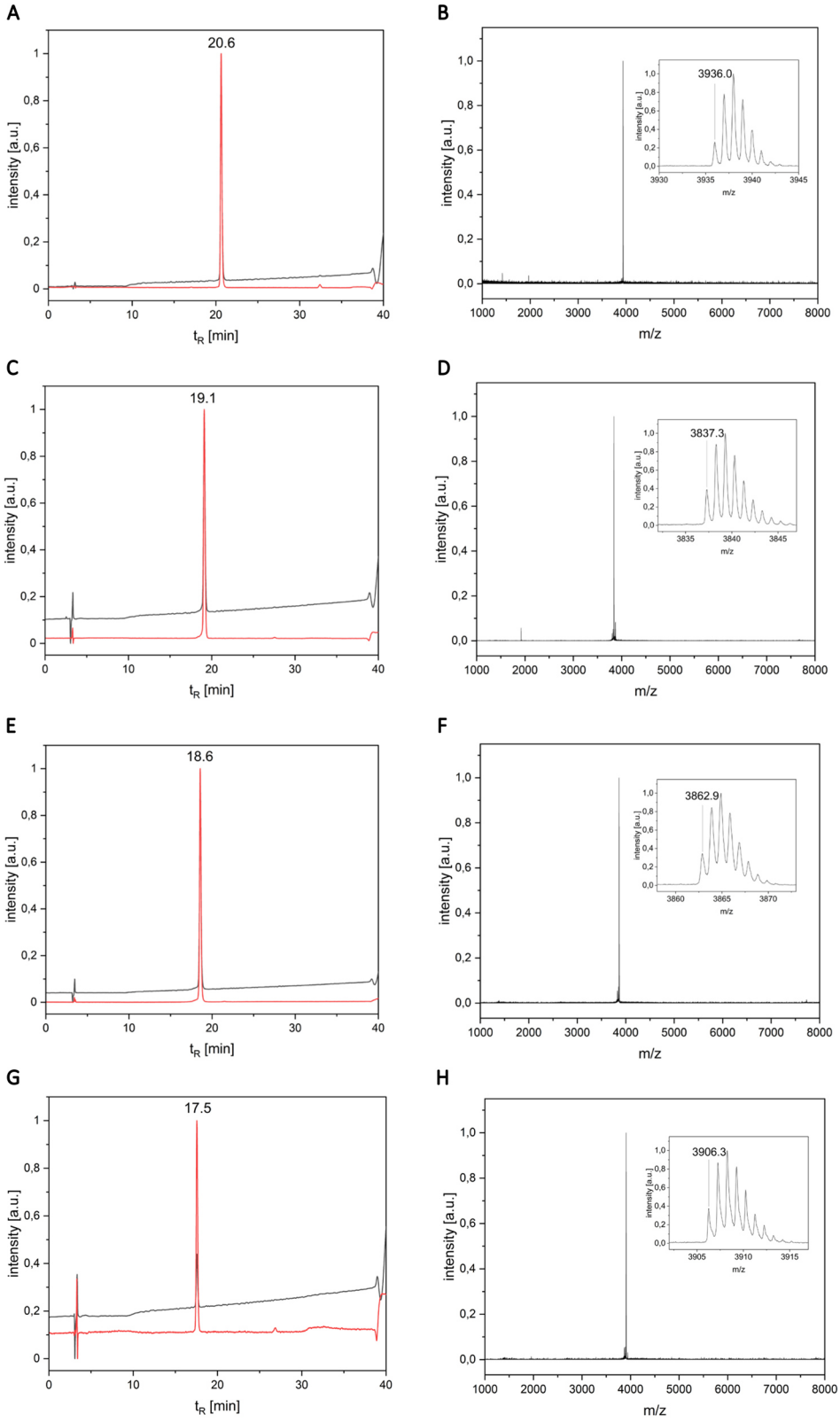


Figure S7: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-LP2-6, **C**) WW-LP2-7, **E**) WW-LP2-8, and **G**) WW-LP2-9. MALDI-TOF-MS spectra of **B**) WW-LP2-6, **D**) WW-LP2-7, **F**) WW-LP2-8, and **H**) WW-LP2-9. This figure was adapted and reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

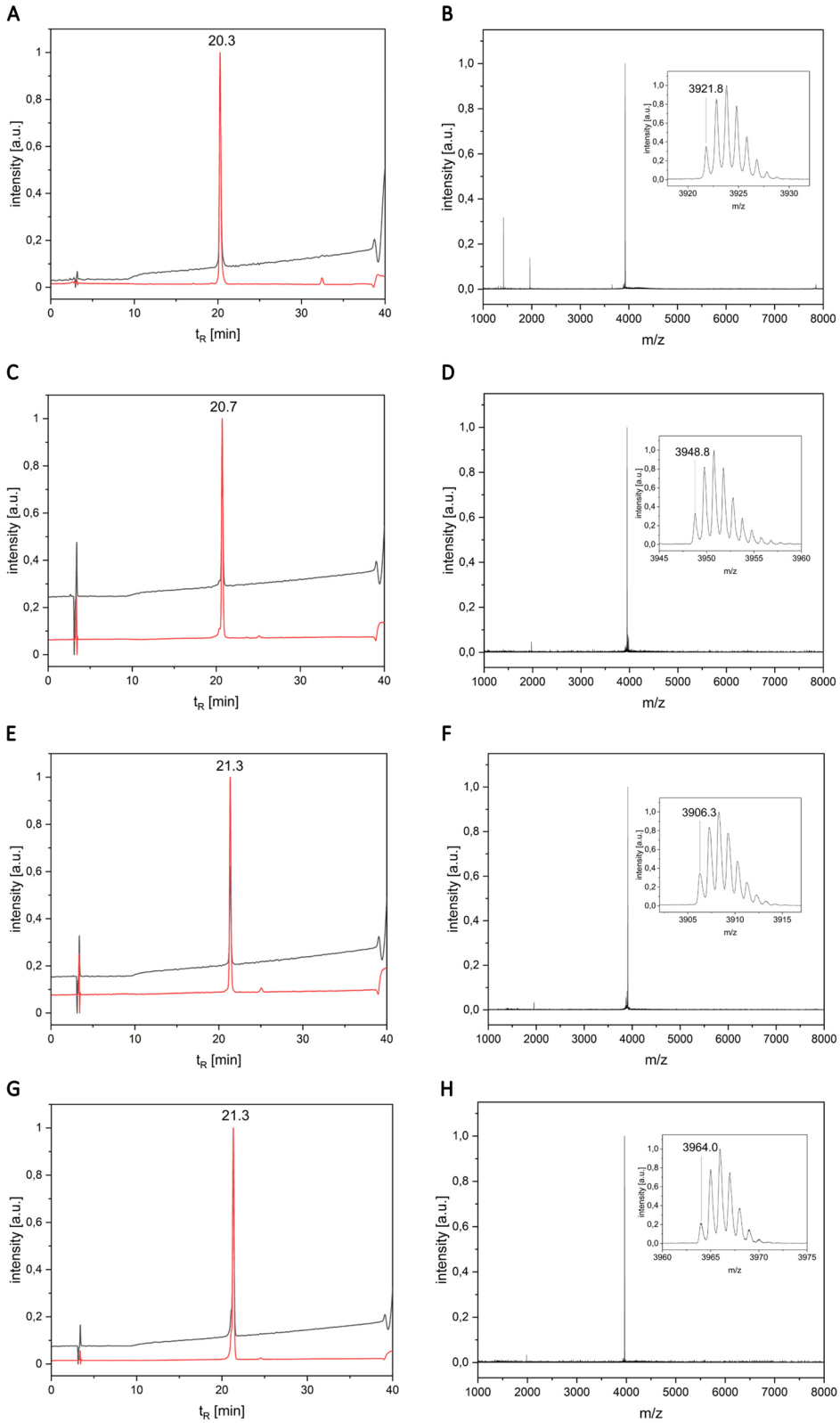


Figure S8: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-BS3-1, **C**) WW-BS3-2, **E**) WW-BS3-3, and **G**) WW-BS3-4. MALDI-TOF-MS spectra of **B**) WW-BS3-1, **D**) WW-BS3-2, **F**) WW-BS3-3, and **H**) WW-BS3-4. This figure was adapted and reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

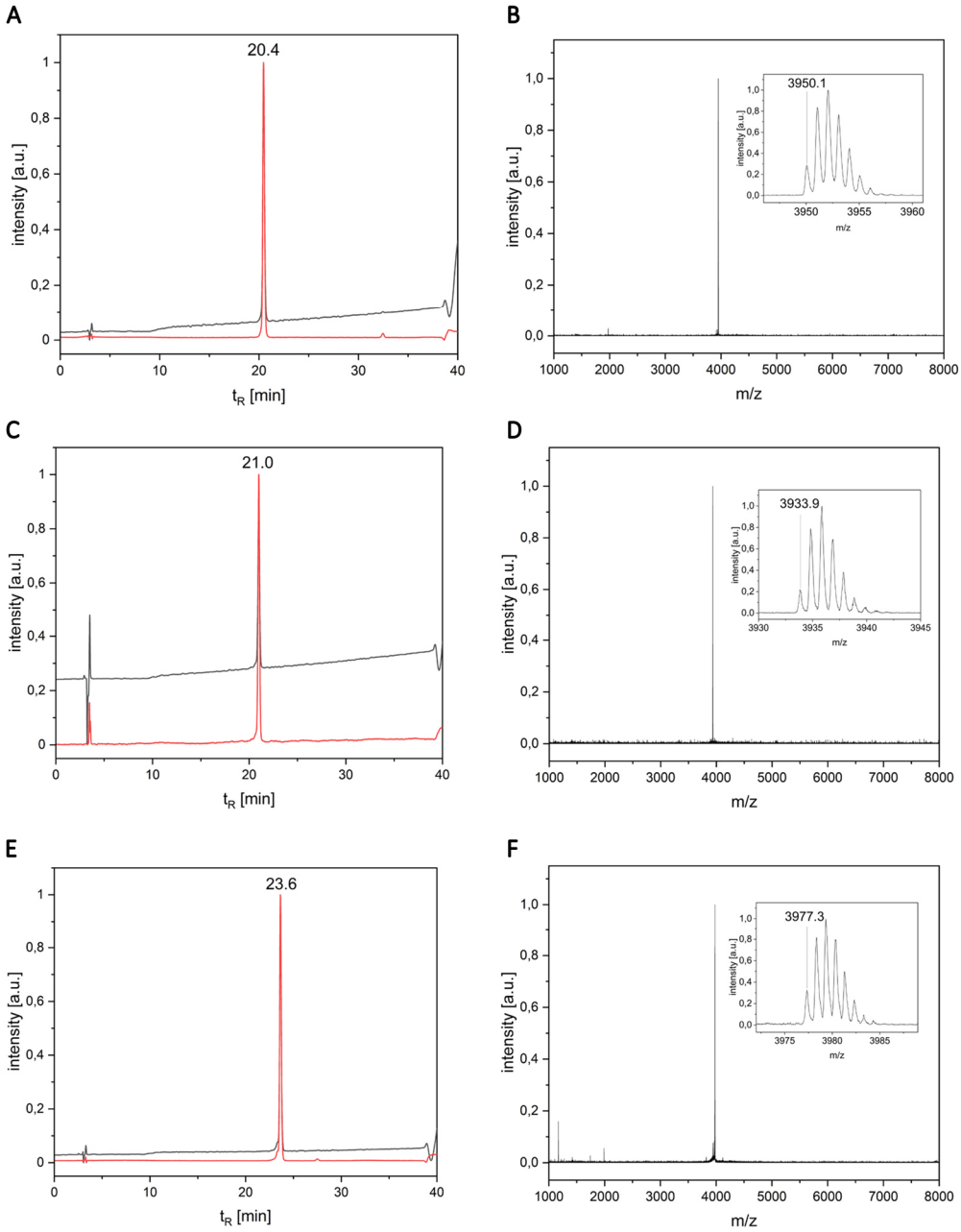


Figure S9: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-BS3-5, **C**) WW-BS3-6, **E**) WW-BS3-7. MALDI-TOF-MS spectra of **B**) WW-BS3-5, **D**) WW-BS3-6, **F**) WW-BS3-7. This figure was adapted and reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

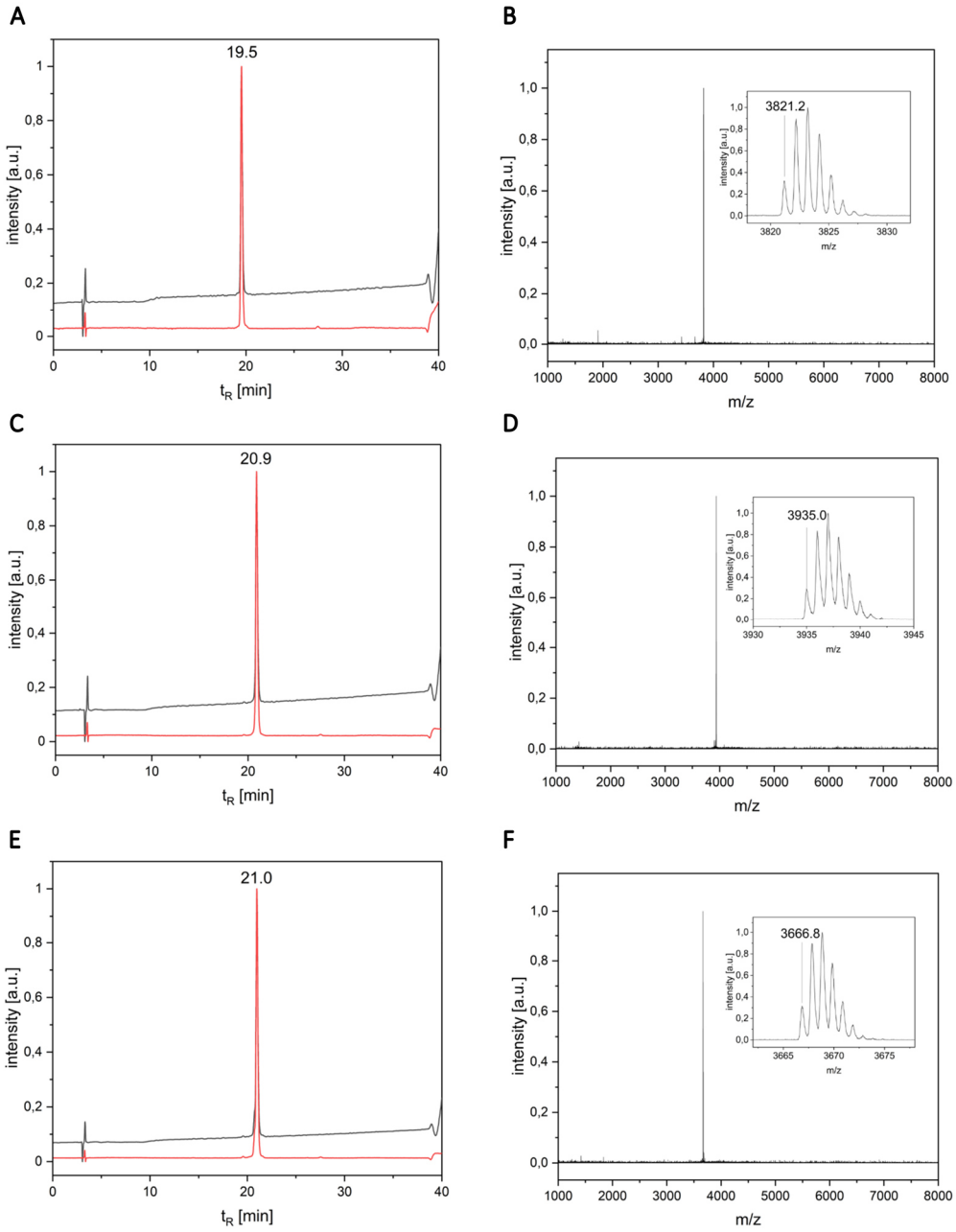


Figure S10: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-CT-1, **C**) WW-CT-2, **E**) WW-CT-3. MALDI-TOF-MS spectra of **B**) WW-CT-1, **D**) WW-CT-2, **F**) WW-CT-3. This figure was reprinted with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

8.3.2 Generation of highly thermostable WW domain scaffold peptides and functionalized thermostable WW domain peptides

Table S3: Highly thermostable WW domain scaffold peptides and functionalized thermostable WW domain peptides. Calculated mass and identified mass of MALDI-TOF MS spectrum, the RP-HPLC linear gradient, and characteristic retention time t_R of each peptide is summarized.

Peptide name	$[M+H]^+$ (calc.)	$[M+H]^+$ (found)	Percentage of HPLC Buffer B in A (*v*/*v*)	t_R [min]
WW-HS-1	4064.1	4062.9	25-50 %	18.4
WW-HS2	4078.1	4077.0	25-50 %	19.3
WW-HS-2$_{OH}$	4037.1	4036.1	25-45 %	15.9
WW-HS-3	4093.1	4092.1	25-50 %	17.2
WW-HS-4	4081.1	4080.6	25-50 %	13.2
WW-HS-5	4109.1	4108.1	25-50 %	9.8
WW-HS-I	4149.2	4148.2	25-50 %	19.1
WW-HS-II/III	4027.0	4026.3	25-60 %	20.2
WW-HS-IV	4148.2	4146.9	25-50 %	17.4
***y*YAP1$_{WW2}$**	3994.0	3993.2	25-50 %	20.2
***h*Gas7$_{WW}$**	4025.0	4024.1	25-50 %	17.8
***h*Pin1$_{WW}$**	4063.0	4062.1	15-50 %	20.0
Ligand WW-I	1026.5	1048.7 [M + Na^+]	10-40 %	18.9
Ligand WW-II/III	936.5	936.6	10-40 %	16.8
Ligand WW-IV	859.3	857.3	5-20 %	14.9
WW-HS-CA	3952.0	3952.8	20-45 %	16.8
WW-HS-ATP	4193.2	4193.6	20-50 %	16.3
WW-HS-PC	3961.0	3963.5	20-50 %	18.5

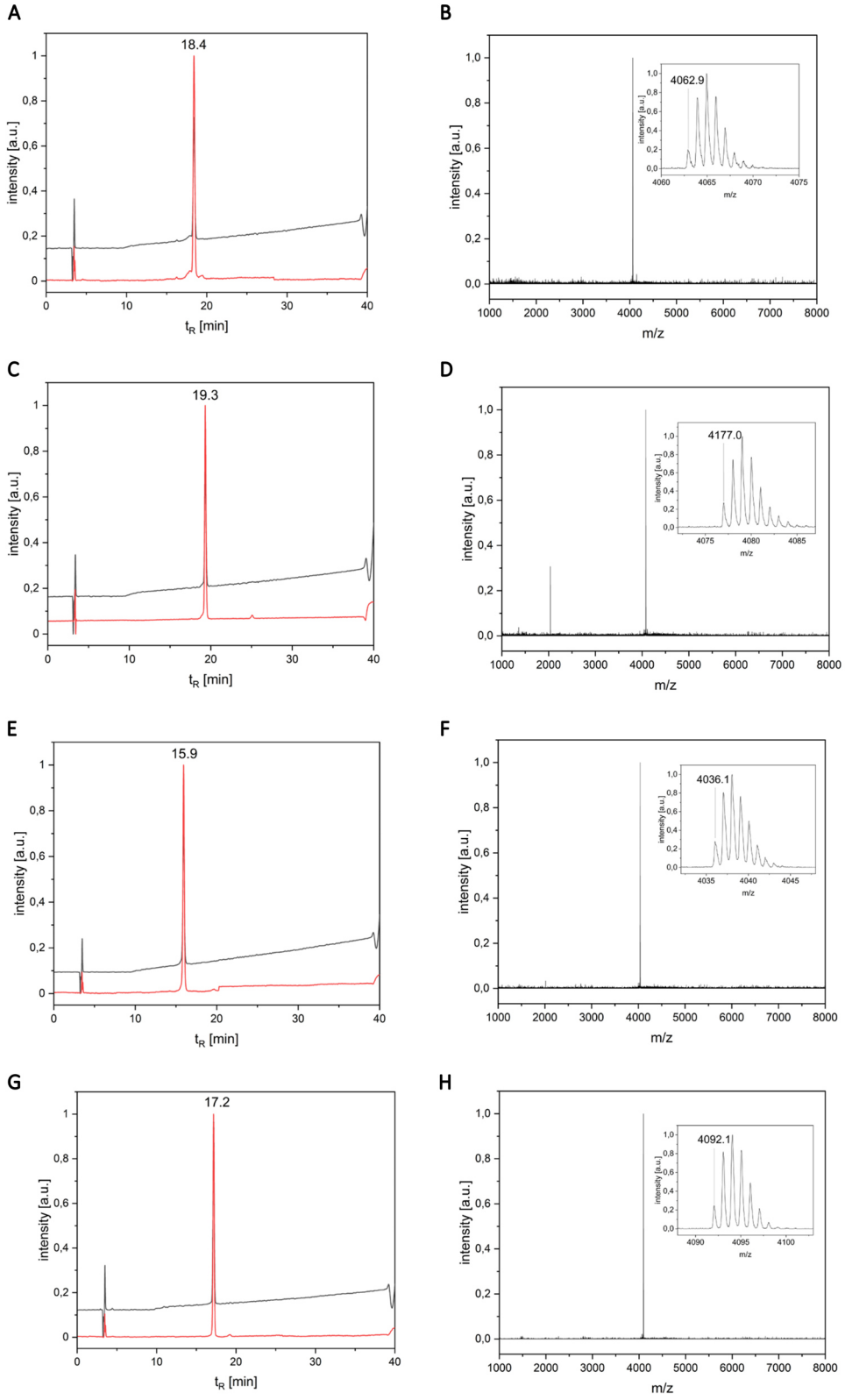


Figure S11: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-HS-1, **C**) WW-HS-2, **E**) WW-HS-2_{OH}, and **G**) WW-HS-3. MALDI-TOF-MS spectra of **B**) WW-HS-1, **D**) WW-HS-2, **F**) WW-HS-2_{OH}, and **H**) WW-HS-3. This figure was reprinted with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

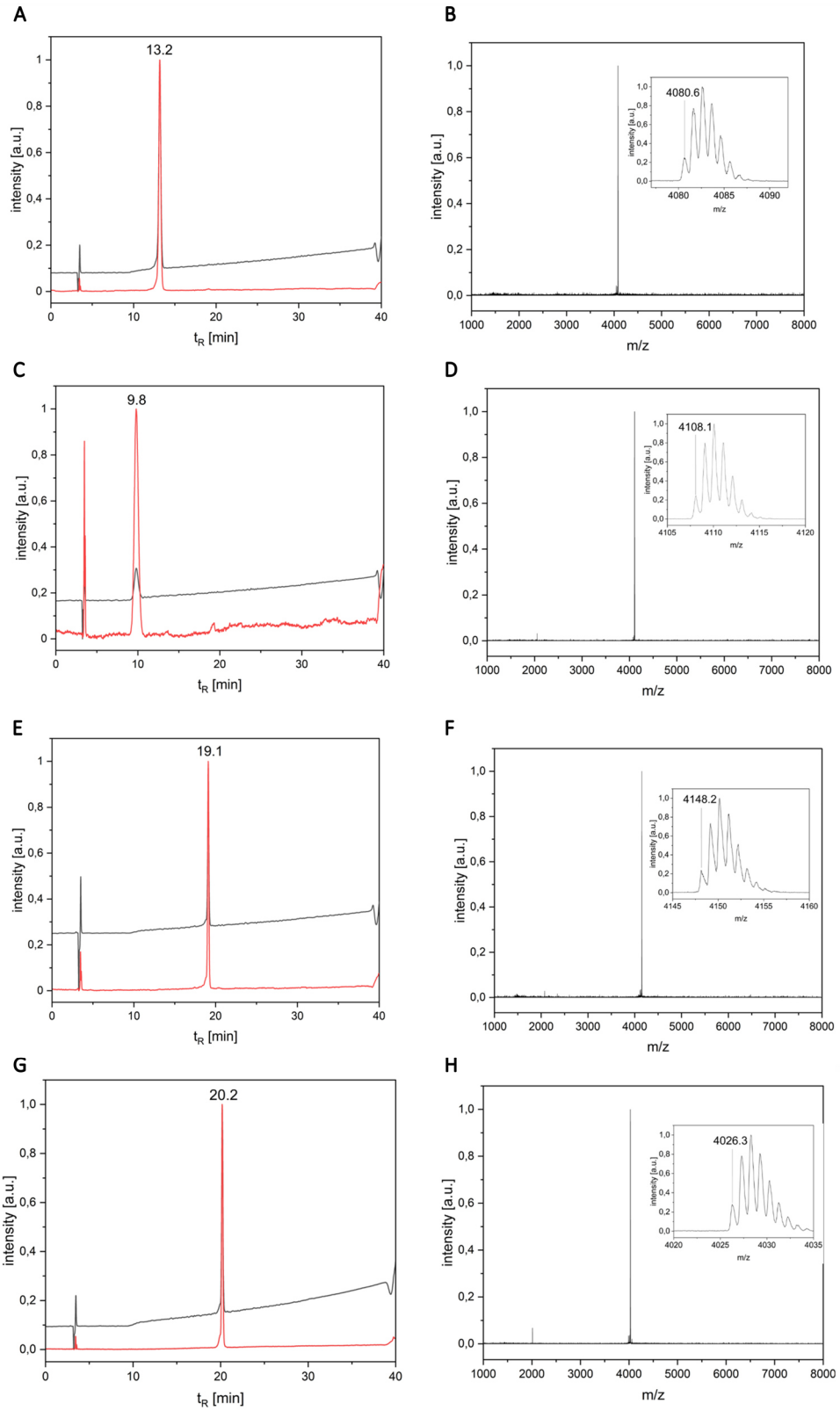


Figure S12: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-HS-4, **C**) WW-HS-5, **E**) WW-HS-I, and **G**) WW-HS-II/III. MALDI-TOF-MS spectra of **B**) WW-HS-4, **D**) WW-HS-5, **F**) WW-HS-I, and **H**) WW-HS-II/III. This figure was adapted with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

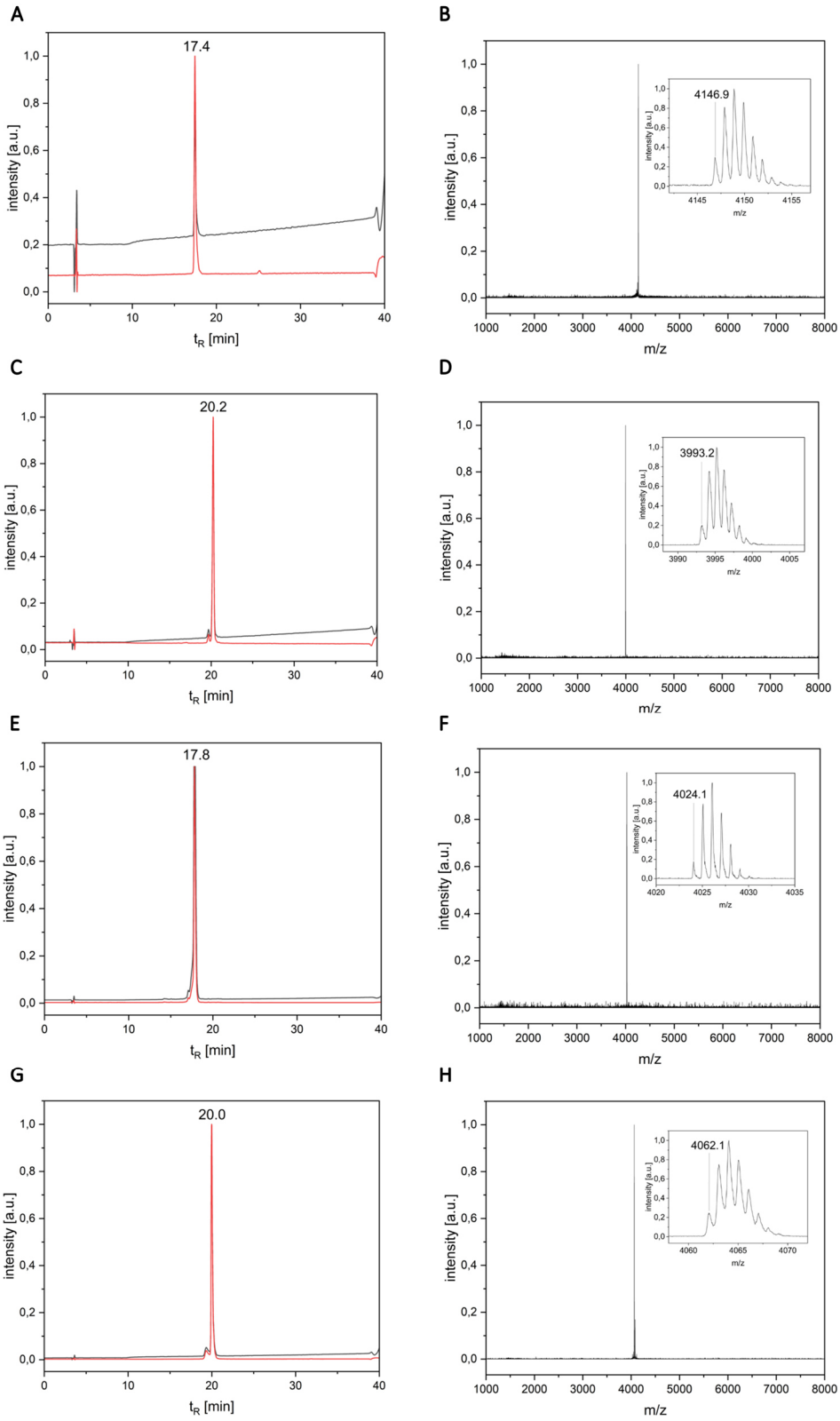


Figure S13: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-HS-IV, **C**) *y*YAP1$_{WW2}$, **E**) *h*Gas7$_{WW}$, and **G**) *h*Pin1$_{WW}$. MALDI-TOF-MS spectra of **B**) WW-HS-IV, **D**) *y*YAP1$_{WW2}$, **F**) *h*Gas7$_{WW}$, and **H**) *h*Pin1$_{WW}$. This figure was adapted with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

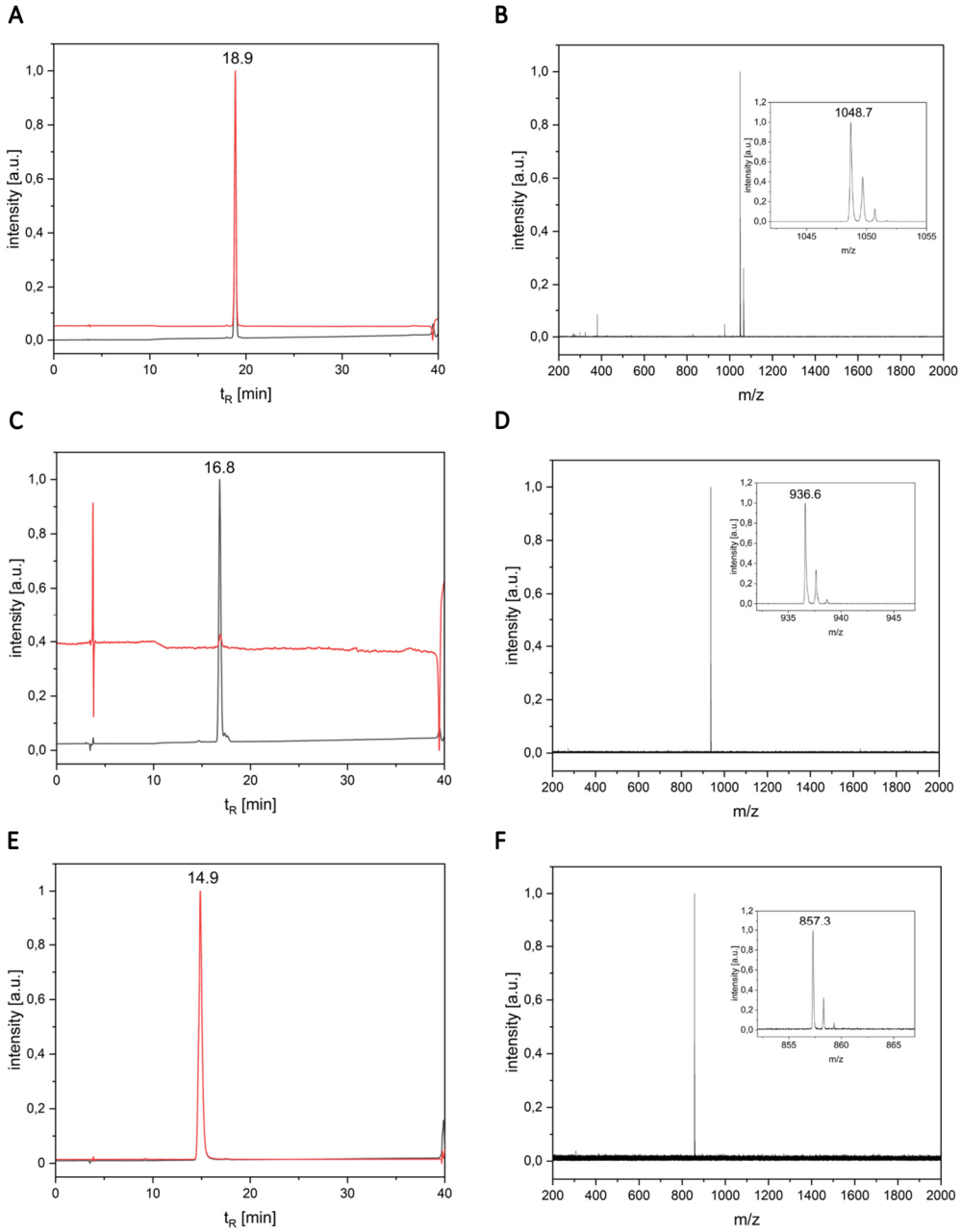


Figure S14: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) Ligand-WW-HS-I, **C**) Ligand-WW-HS-II/III, **E**) Ligand-WW-HS-IV. MALDI-TOF-MS spectra of **B**) Ligand-WW-HS-I, **D**) Ligand-WW-HS-II/III, **F**) Ligand-WW-HS-IV (mass found: $M+Na^+$). This figure was adapted with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

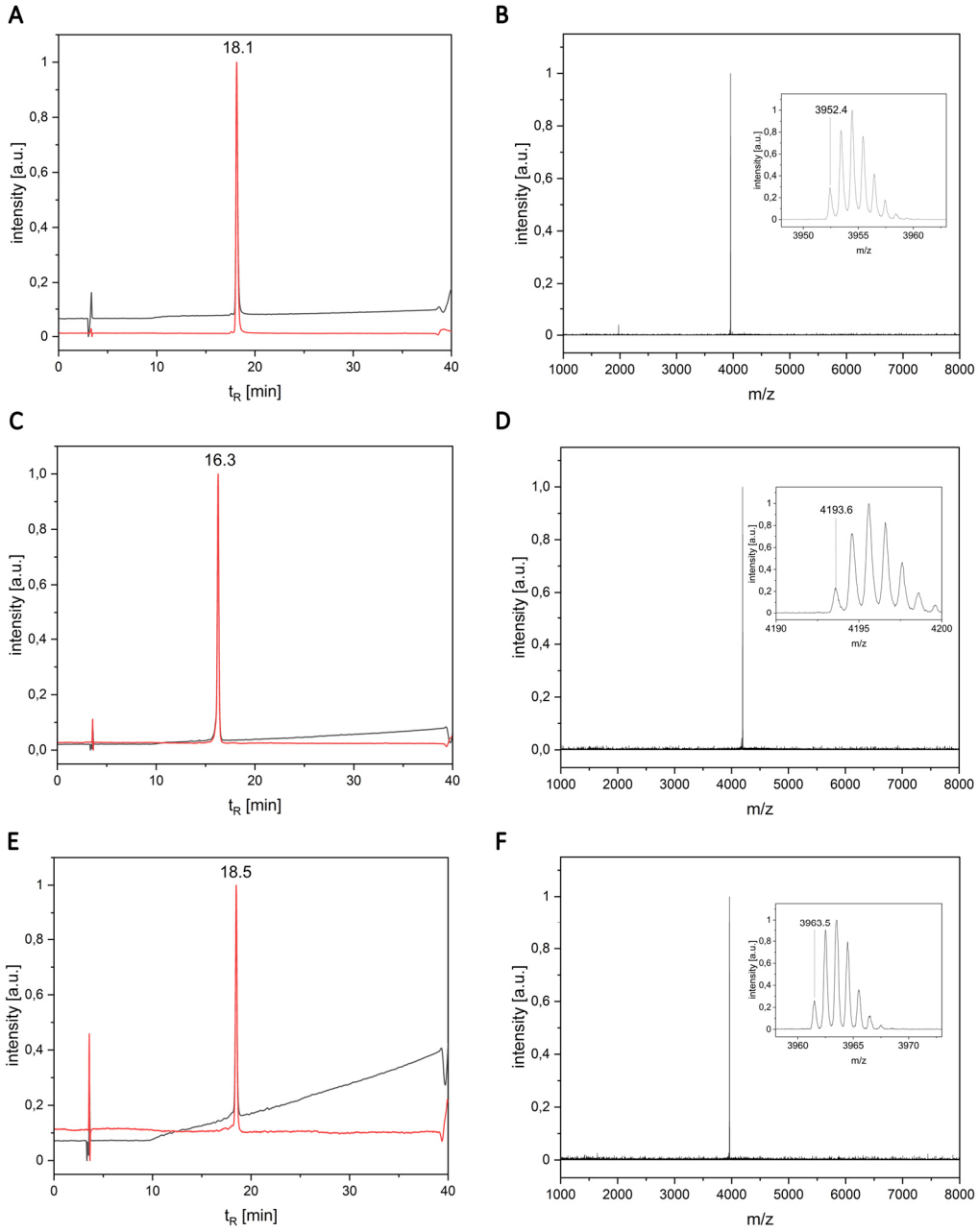


Figure S15: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-HS-CA, **C**) WW-HS-ATP, **E**) WW-HS-PC. MALDI-TOF-MS spectra of **B**) WW-HS-CA, **D**) WW-HS-ATP, **F**) WW-HS-PC. This figure was reprinted with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

8.3.3 Design of *de novo* ATP-binding variants

Table S4: *De novo* designed ATP-binding peptides. Calculated mass and identified mass of MALDI-TOF MS spectrum, the RP-HPLC linear gradient, and characteristic retention time t_R of the respective peptide variant is summarized.

Peptide name	$[M+H]^+$ (calc.)	$[M+H]^+$ (found)	Percentage of HPLC Buffer B in A (*v*/*v*)	t_R [min]
WW-2-10	4023.0	4023.0	10-40 %	19.1
WW-2-10-HS2	4243.2	4241.3	15-50 %	13.7
WW-2-10-HS1-1	4180.2	4181.1	20-50 %	14.2
WW-2-10-HS1-2	4180.2	4180.8	20-50 %	13.7
WW-2-10-R9C-sb1	3968.0	3967.3	20-50 %	13.7
WW-2-10-R9C-sb2	4109.0	4108.0	20-50 %	16.0

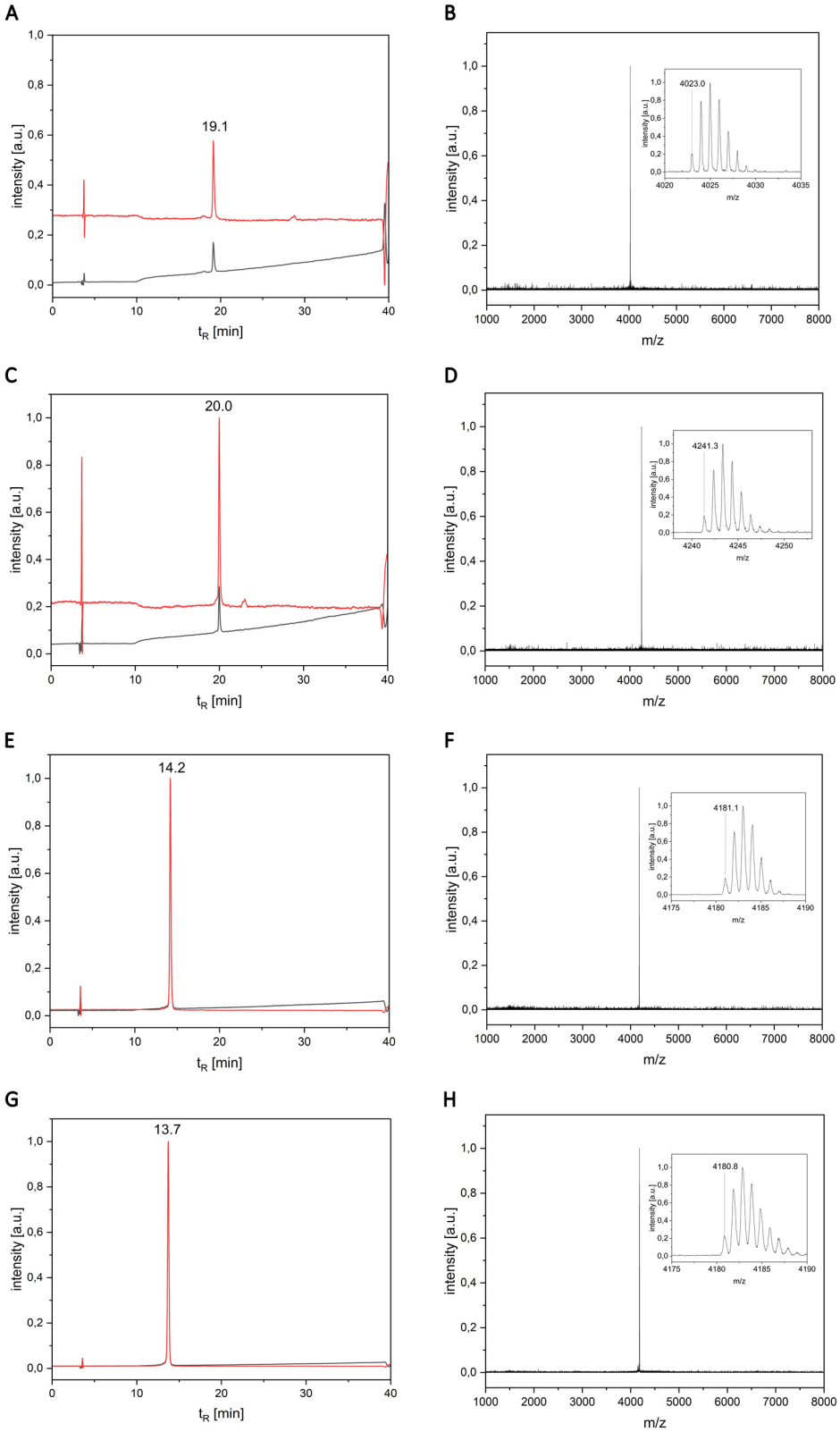


Figure S16: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-2-10, **C**) WW-2-10-HS2, **E**) WW-2-10-HS1-1, and **G**) WW-2-10-HS1-2. MALDI-TOF-MS spectra of **B**) WW-2-10, **D**) WW-2-10-HS2, **F**) WW-2-10-HS1-1, and **H**) WW-2-10-HS1-2.

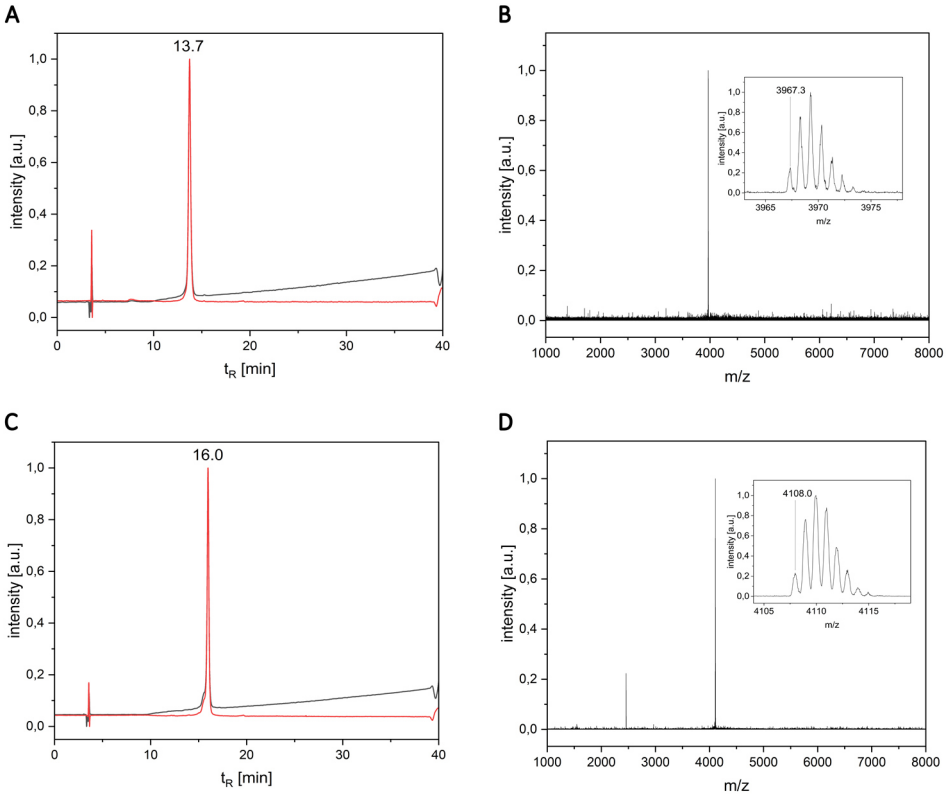


Figure S17: Analytical HPLC chromatograms at 220 nm (black) and 280 nm (red) and MALDI-TOF mass spectra of the synthesized peptides. HPLC traces of **A**) WW-2-10-R9C-sb1, **C**) WW-2-10-R9C-sb2. MALDI-TOF-MS spectra of **B**) WW-2-10-R9C-sb1, **D**) WW-2-10-R9C-sb2.

8.4 CD spectra

8.4.1 CD measurements of WW domain screening peptides

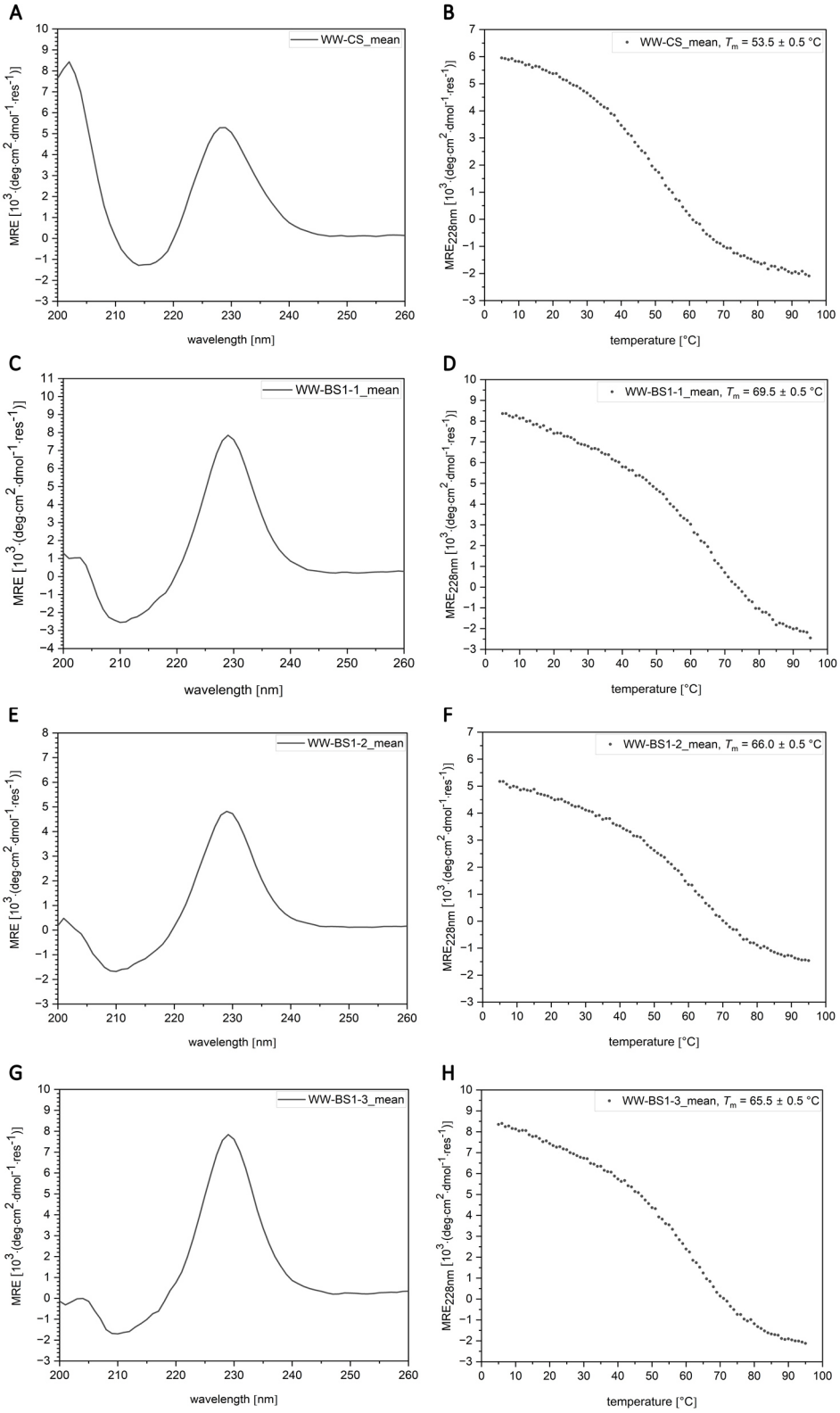


Figure S18: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-CS, **C**) WW-BS1-1, **E**) WW-BS1-2, **G**) WW-BS1-3. Thermal denaturation measurements of **B**) WW-CS, **D**) WW-BS1-1, **F**) WW-BS1-2, **H**) WW-BS1-3. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

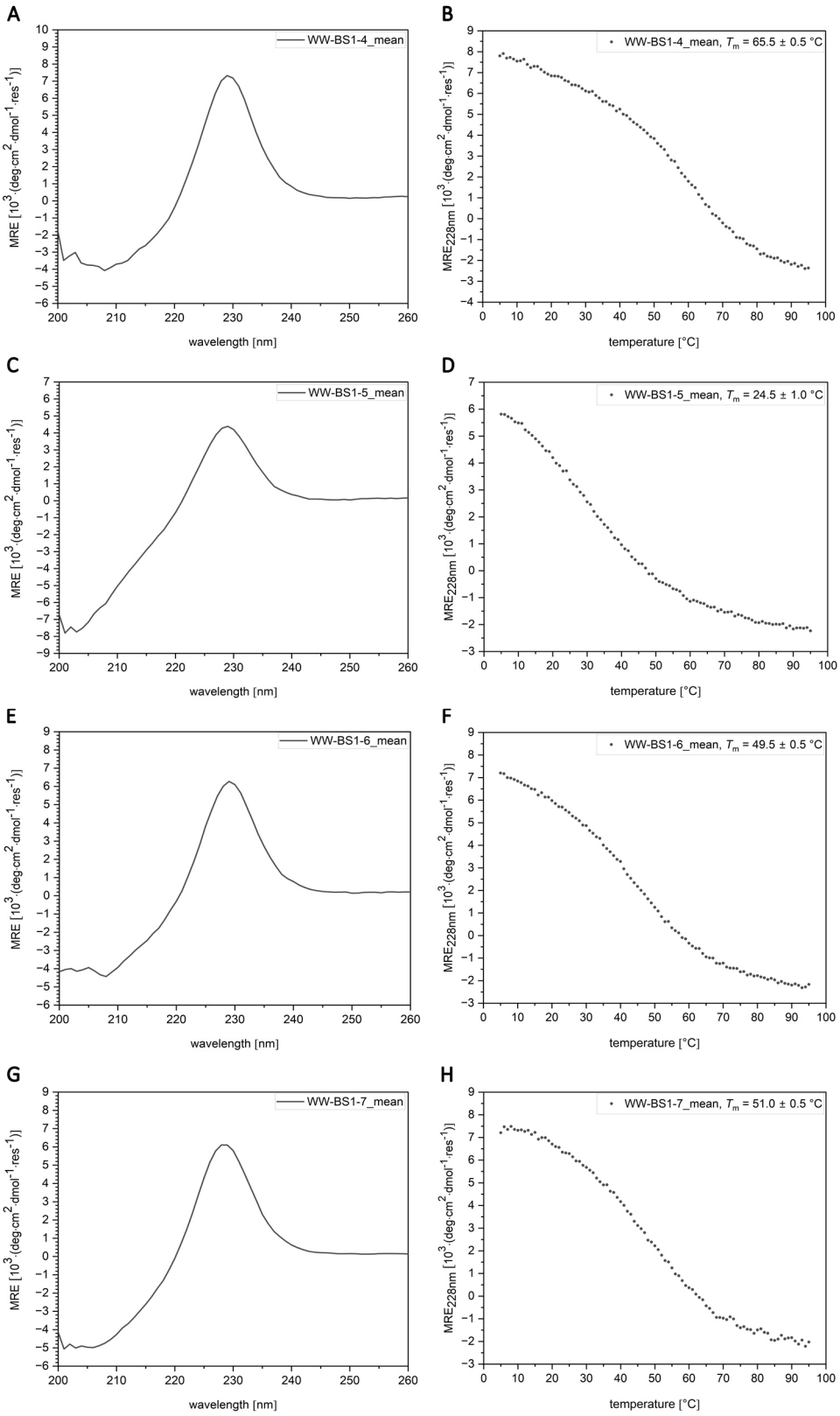


Figure S19: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-BS1-4, **C**) WW-BS1-5, **E**) WW-BS1-6, **G**) WW-BS1-7. Thermal denaturation measurements of **B**) WW-BS1-4, **D**) WW-BS1-5, **F**) WW-BS1-6, **H**) WW-BS1-7. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

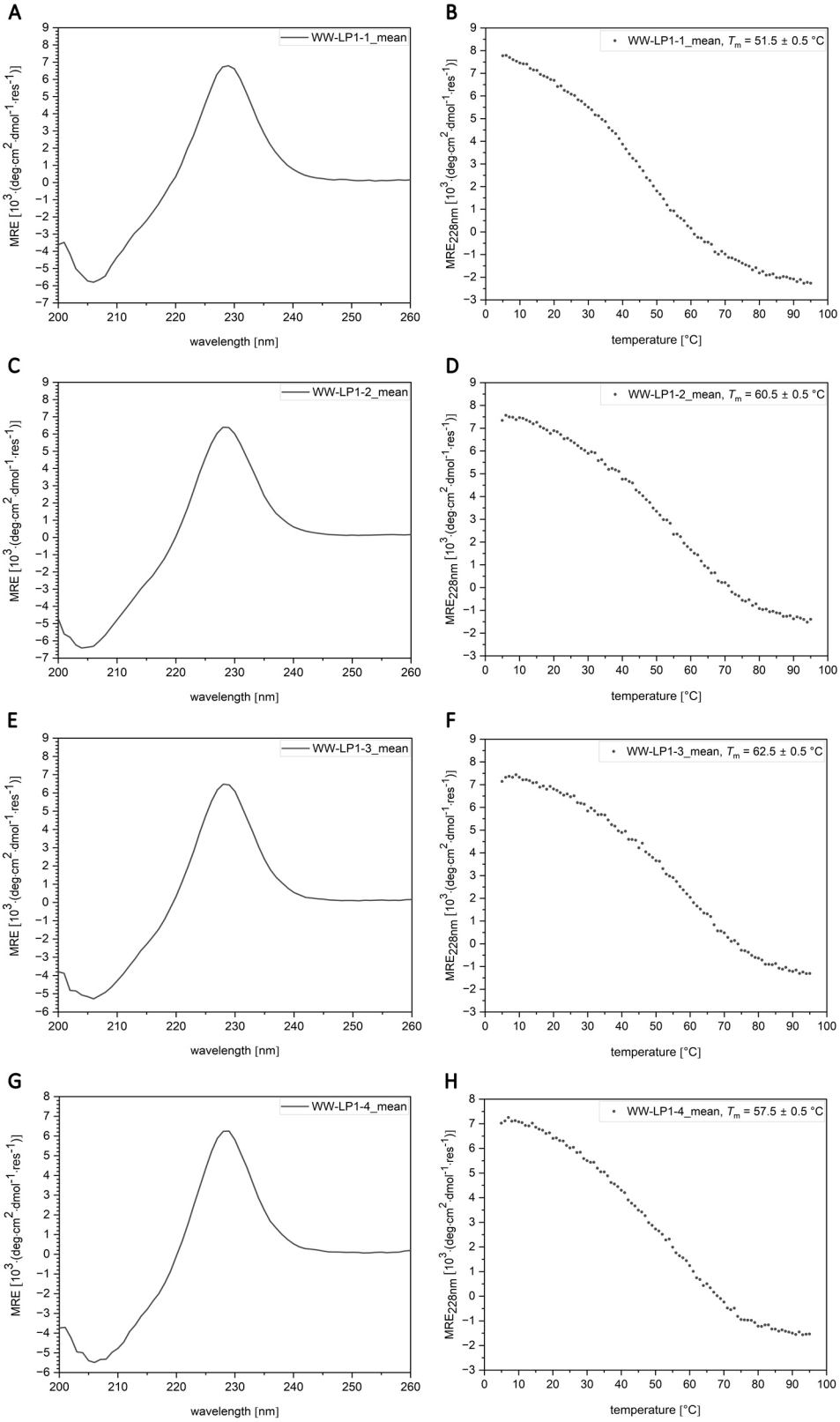


Figure S20: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-LP1-1, **C**) WW-LP1-2, **E**) WW-LP1-3, **G**) WW-LP1-4. Thermal denaturation measurements of **B**) WW-LP1-1, **D**) WW-LP1-2, **F**) WW-LP1-3, **H**) WW-LP1-4. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

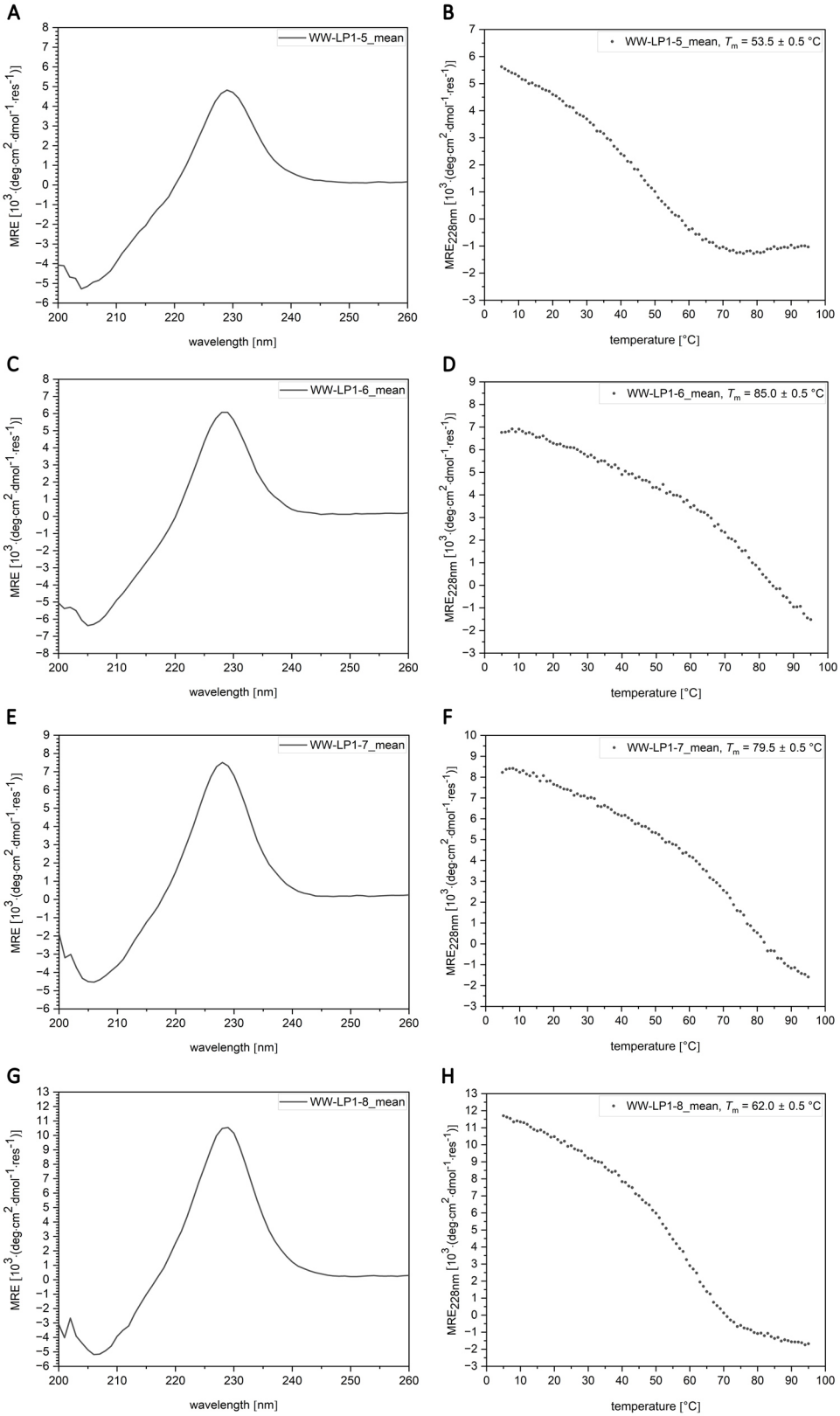


Figure S21: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-LP1-5, **C**) WW-LP1-6, **E**) WW-LP1-7, **G**) WW-LP1-8. Thermal denaturation measurements of **B**) WW-LP1-5, **D**) WW-LP1-6, **F**) WW-LP1-7, **H**) WW-LP1-8. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

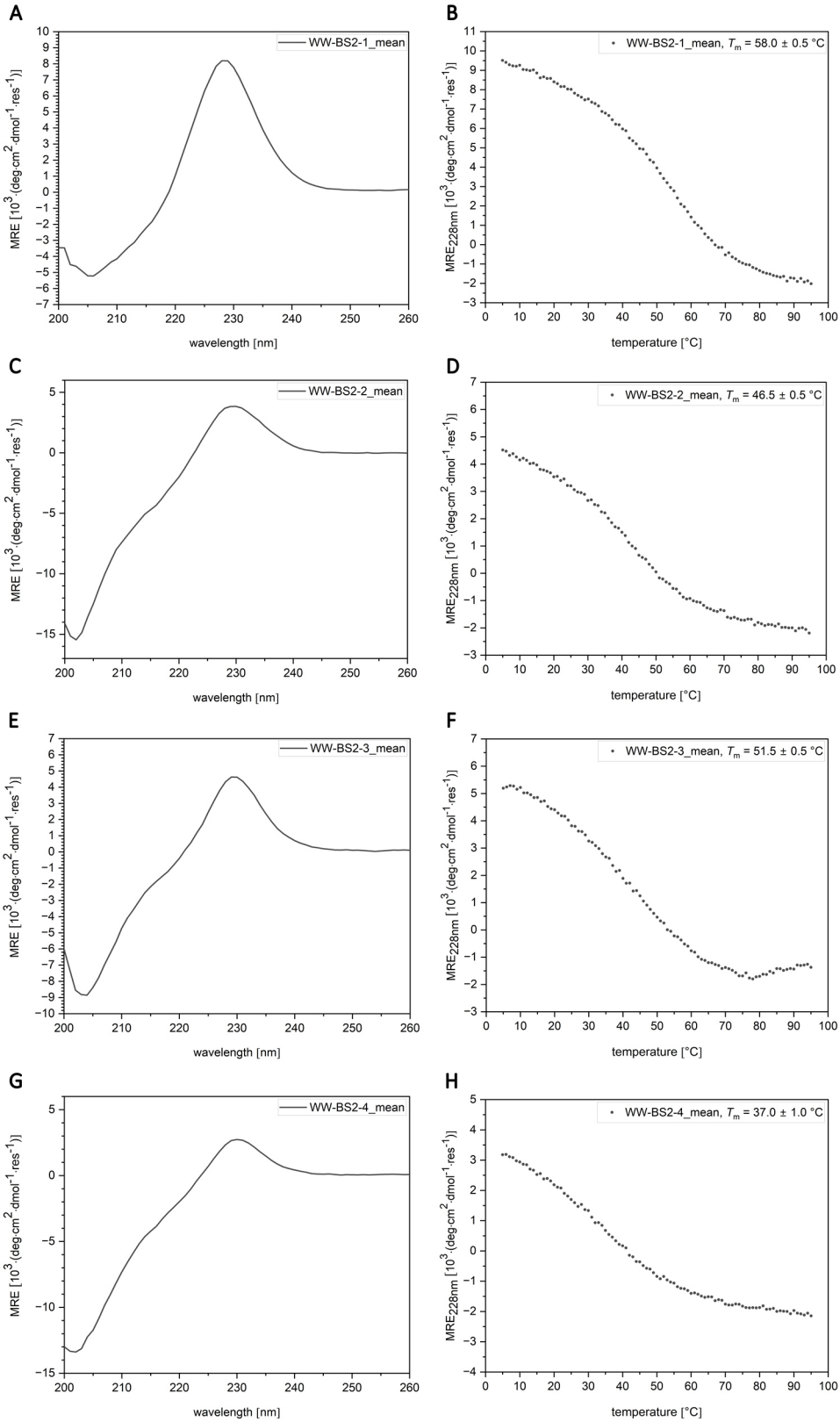


Figure S22: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-BS2-1, **C**) WW-BS2-2, **E**) WW-BS2-3, **G**) WW-BS2-4. Thermal denaturation measurements of **B**) WW-BS2-1, **D**) WW-BS2-2, **F**) WW-BS2-3, **H**) WW-BS2-4. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

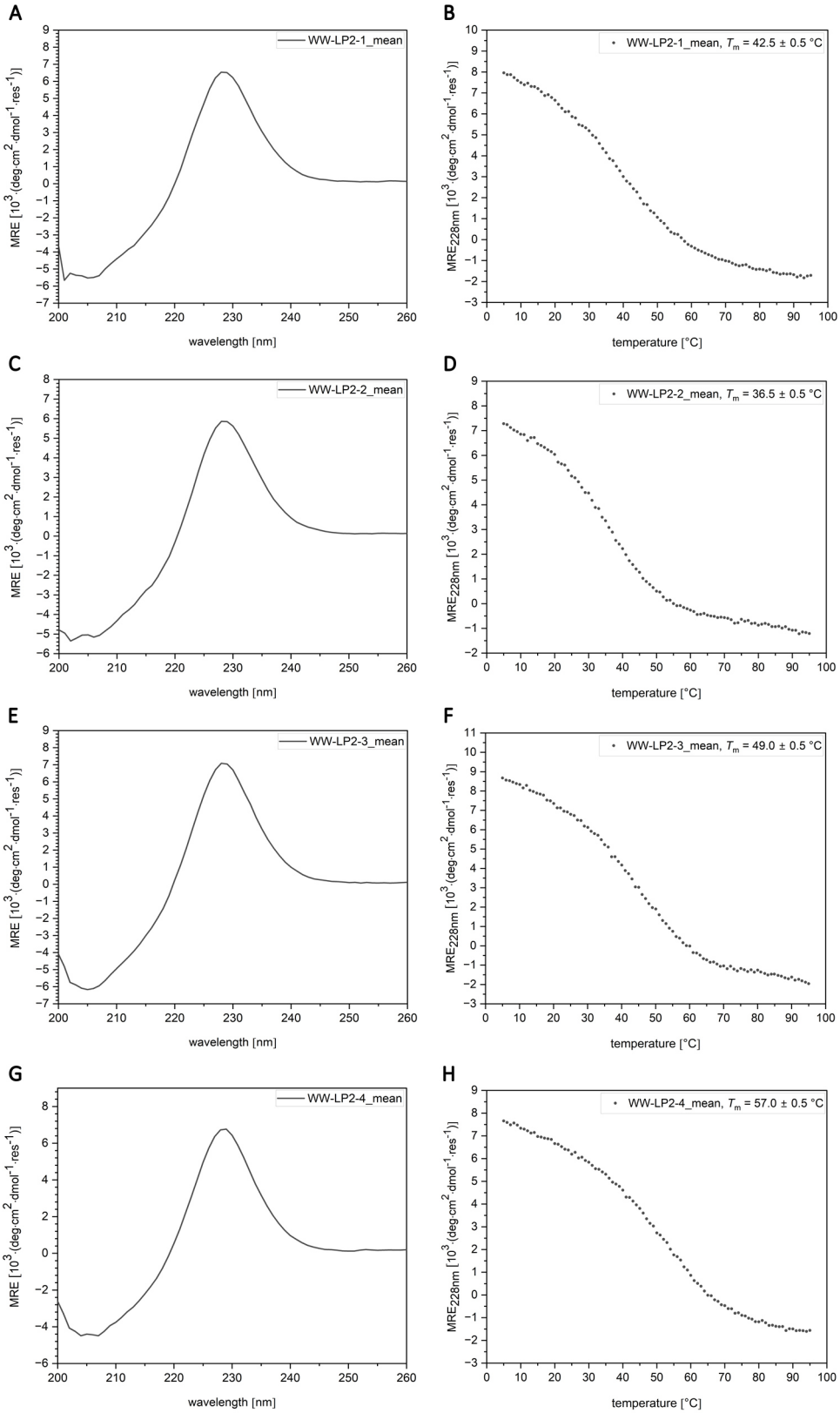


Figure S23: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-LP2-1, **C**) WW-LP2-2, **E**) WW-LP2-3, **G**) WW-LP2-4. Thermal denaturation measurements of **B**) WW-LP2-1, **D**) WW-LP2-2, **F**) WW-LP2-3, **H**) WW-LP2-4. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

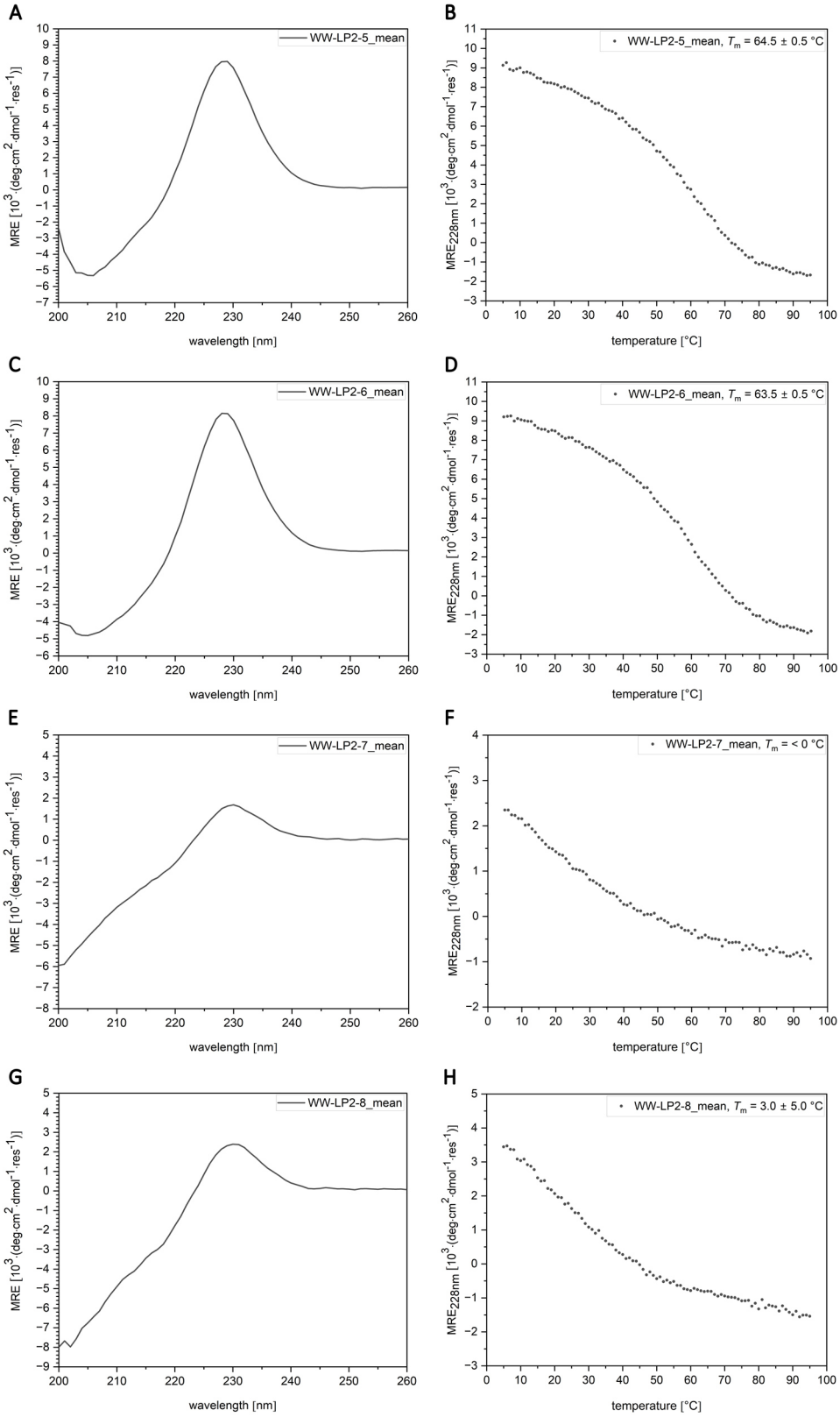


Figure S24: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-LP2-5, **C**) WW-LP2-6, **E**) WW-LP2-7, **G**) WW-LP2-8. Thermal denaturation measurements of **B**) WW-LP2-5, **D**) WW-LP2-6, **F**) WW-LP2-7, **H**) WW-LP2-8. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

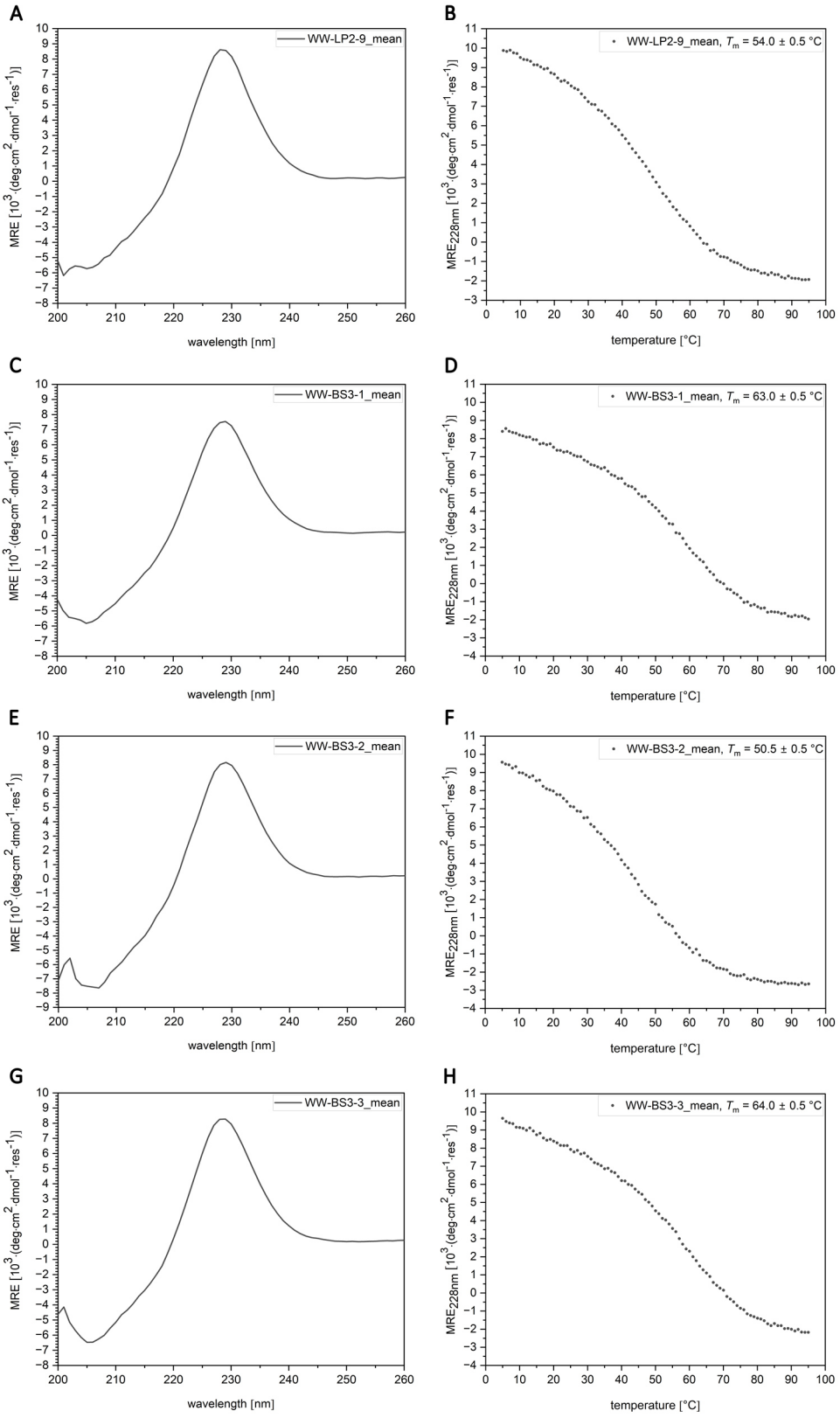


Figure S25: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-LP2-9, **C**) WW-BS3-1, **E**) WW-BS3-2, **G**) WW-BS3-3. Thermal denaturation measurements of **B**) WW-LP2-9, **D**) WW-BS3-1, **F**) WW-BS3-2, **H**) WW-BS3-3. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

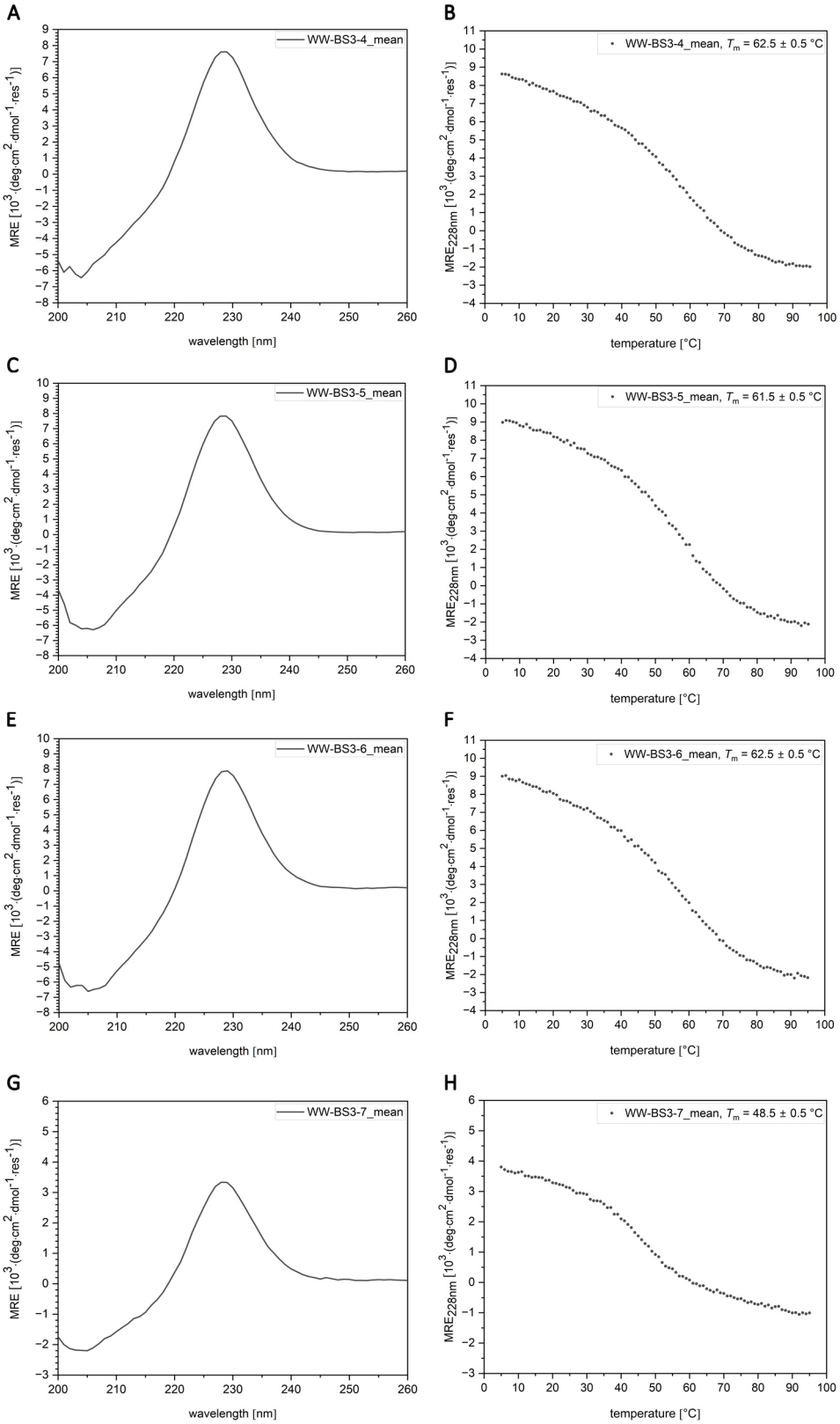


Figure S26: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-BS3-4, **C**) WW-BS3-5, **E**) WW-BS3-6, **G**) WW-BS3-7. Thermal denaturation measurements of **B**) WW-BS3-4, **D**) WW-BS3-5, **F**) WW-BS3-6, **H**) WW-BS3-7. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

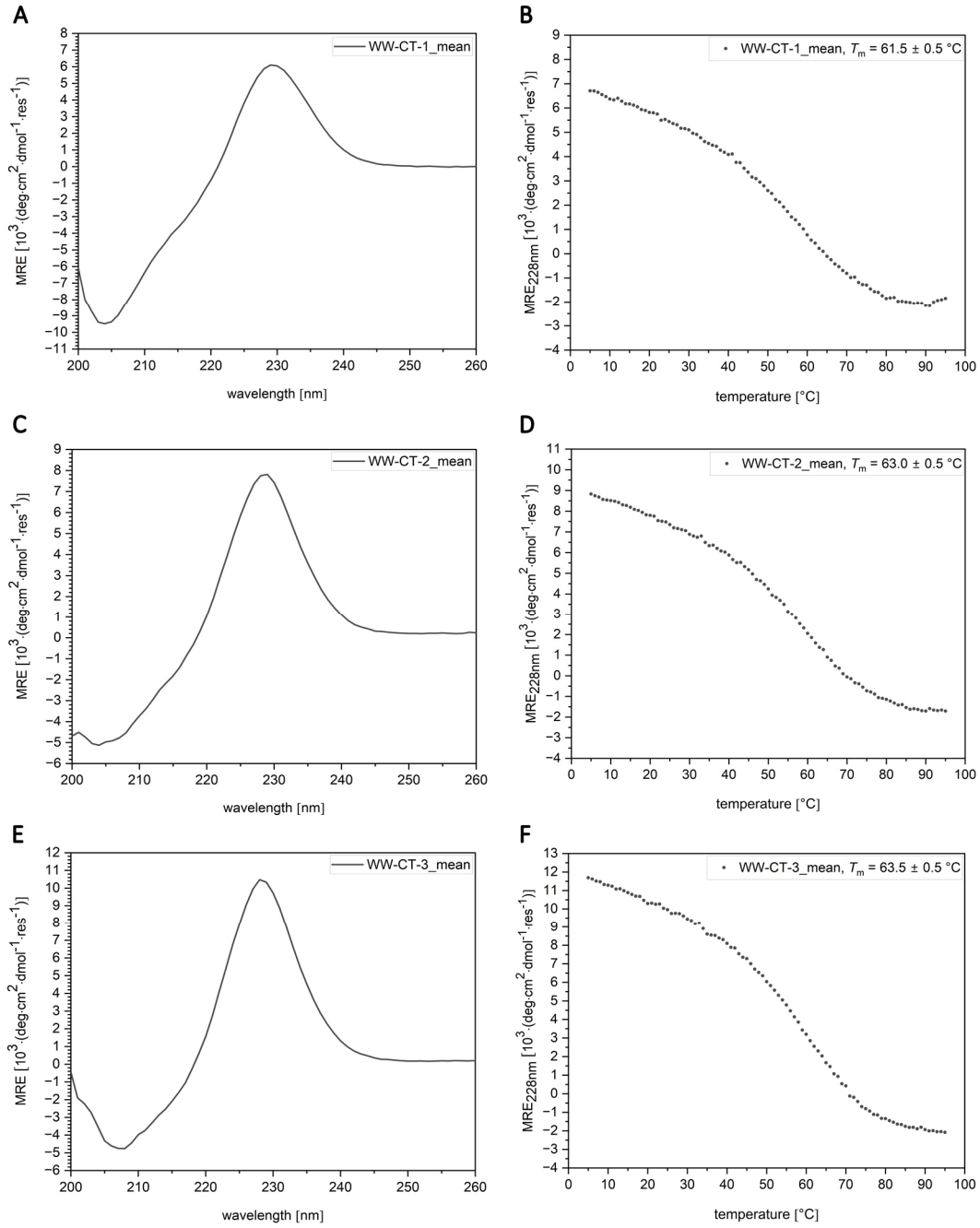


Figure S27: CD spectra and thermal denaturation curves of WW domain screening peptides. CD spectra measurements of **A**) WW-CT-1, **C**) WW-CT-2, **E**) WW-CT-3. Thermal denaturation measurements of **B**) WW-CT-1, **D**) WW-CT-2, **F**) WW-CT-3. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

8.4.2 CD measurements of WW-HS peptides

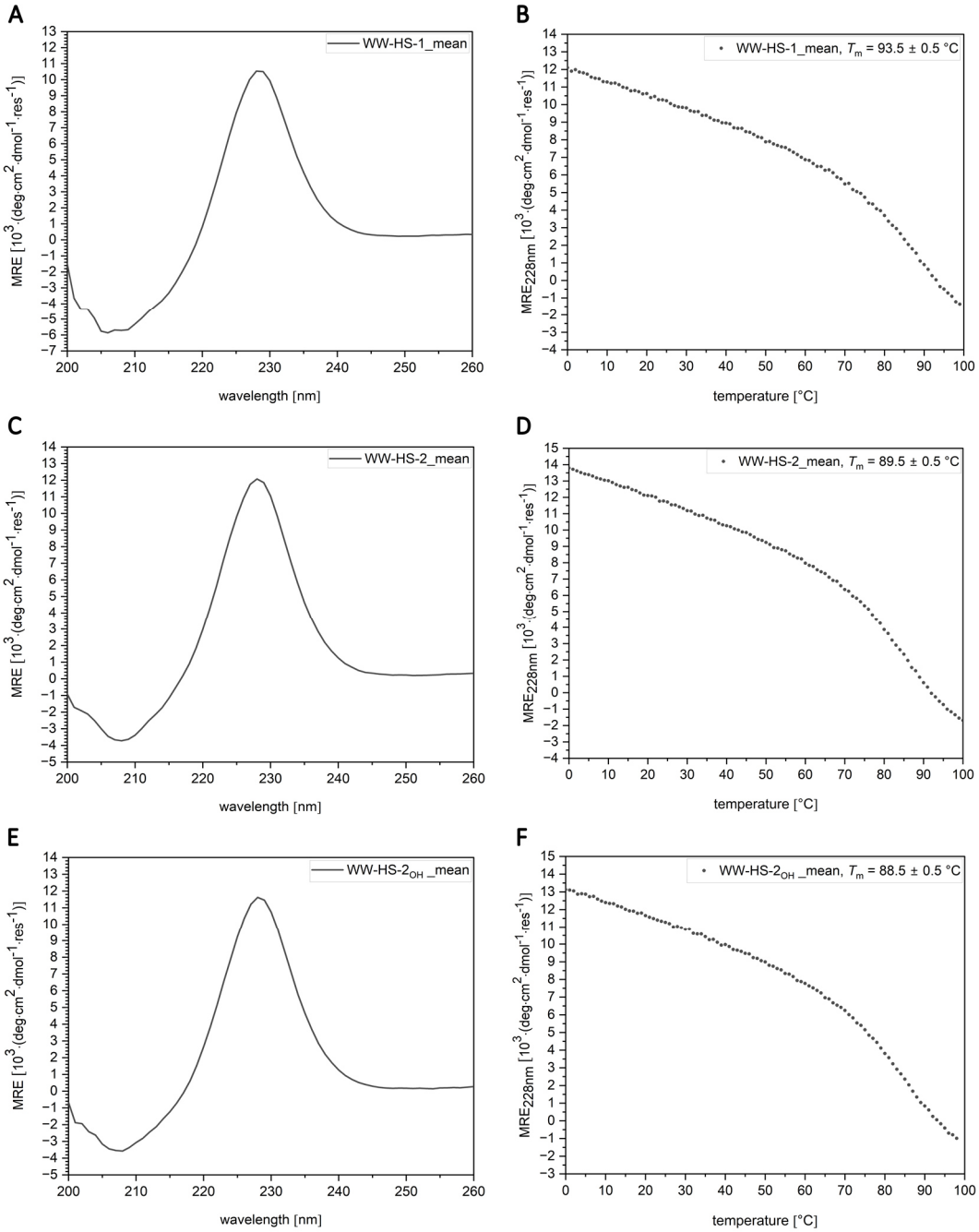


Figure S28: CD spectra and thermal denaturation curves of WW domain HS peptides. CD spectra measurements of **A**) WW-HS-1, **C**) WW-HS-2, **E**) WW-HS-2_{OH}. Thermal denaturation measurements of **B**) WW-HS-1, **D**) WW-HS-2, **F**) WW-HS-2_{OH}. Conditions: 100 µM peptide, PBS buffer, pH 7.5. For CD spectra and thermal denaturation curves the mean of triplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

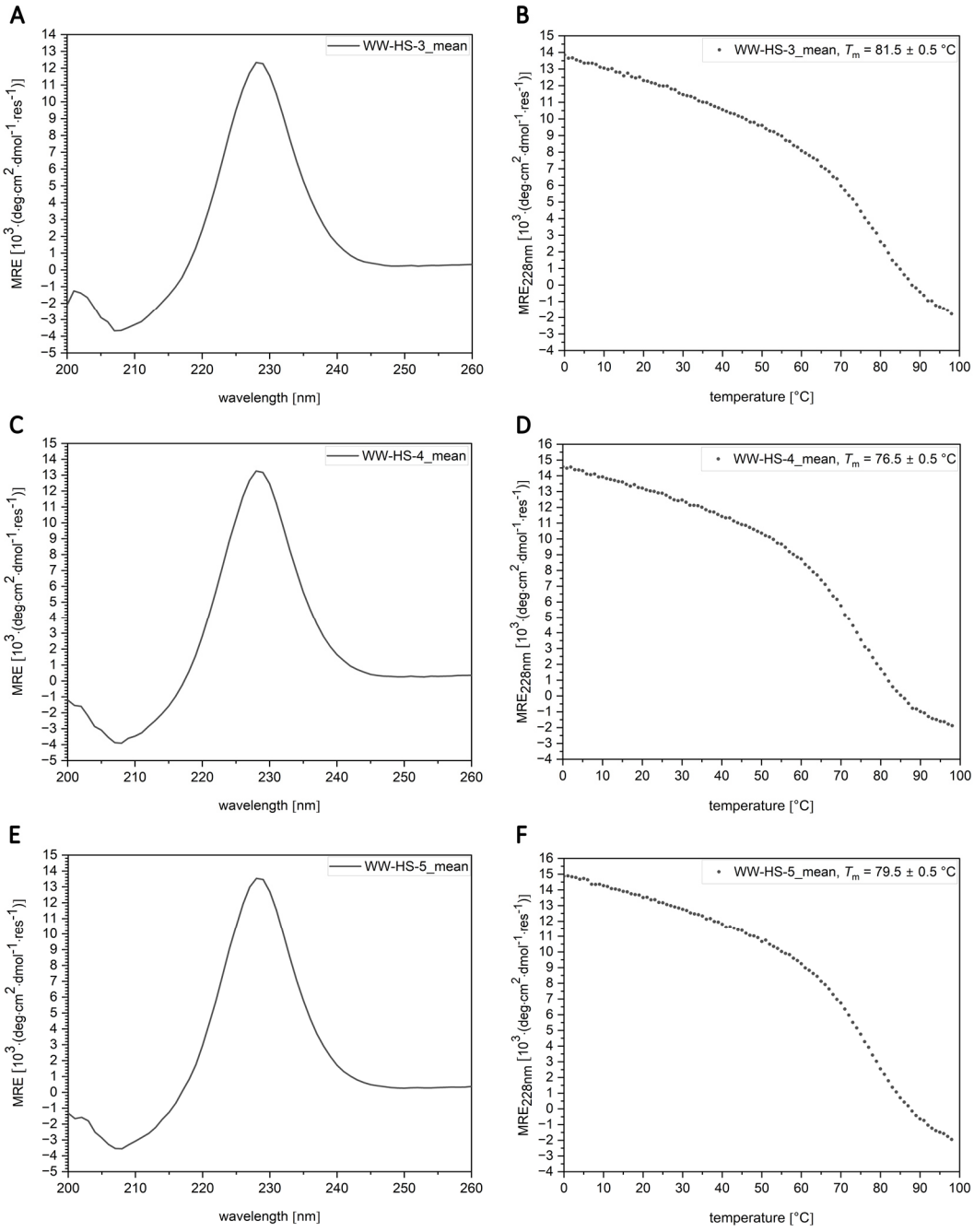


Figure S29: CD spectra and thermal denaturation curves of WW domain HS peptides. CD spectra measurements of **A**) WW-HS-3, **C**) WW-HS-4, **E**) WW-HS-5. Thermal denaturation measurements of **B**) WW-HS-3, **D**) WW-HS-4, **F**) WW-HS-5. For CD spectra and thermal denaturation curves the mean of triplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

8.4.2 CD measurements of functionalized WW-HS peptides

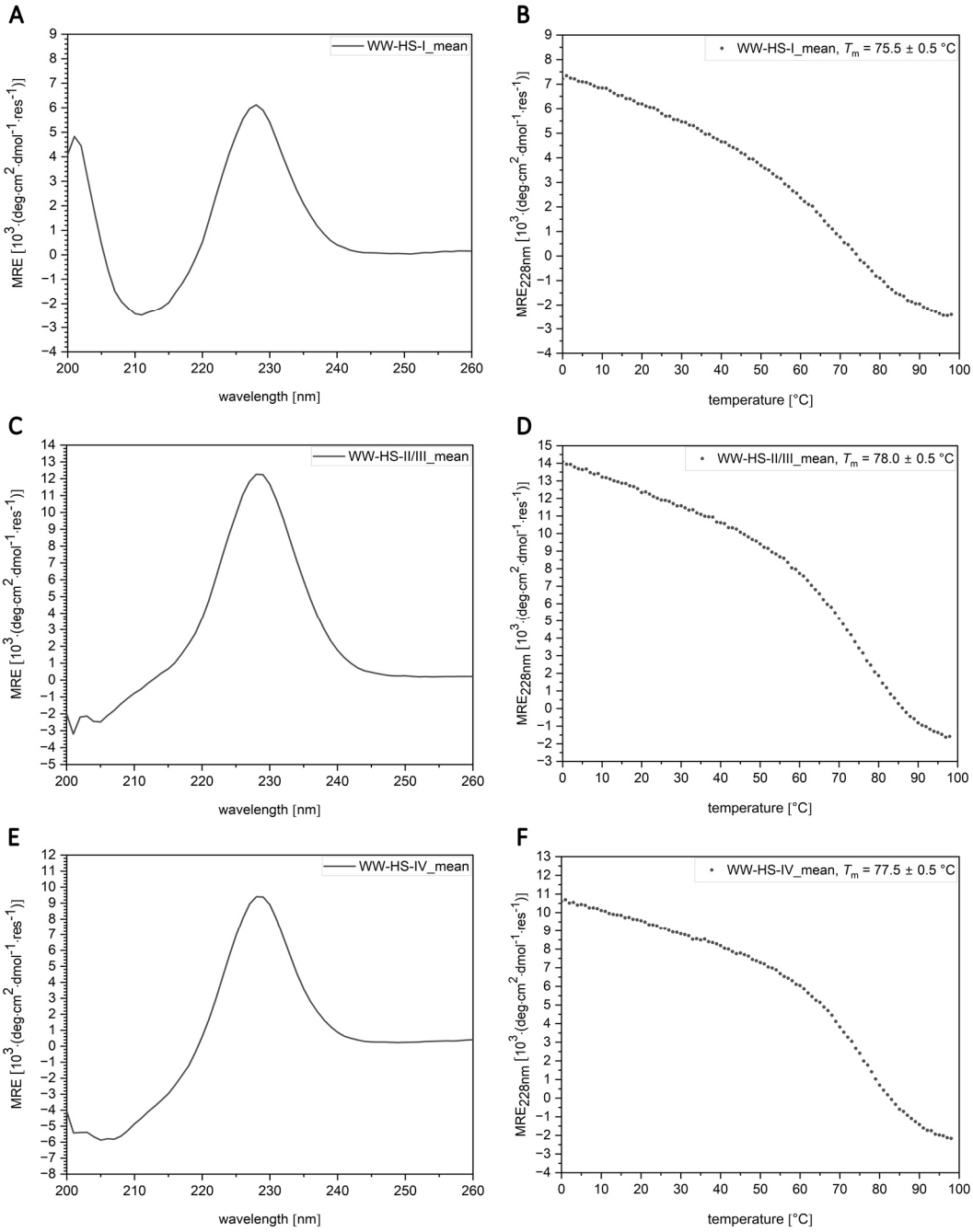


Figure S30: CD spectra and thermal denaturation curves of functionalized WW domain HS peptides. CD spectra measurements of **A**) WW-HS-I, **C**) WW-HS-II/III, **E**) WW-HS-IV. Thermal denaturation measurements of **B**) WW-HS-I, **D**) WW-HS-II/III, **F**) WW-HS-IV. For CD spectra and thermal denaturation curves the mean of triplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

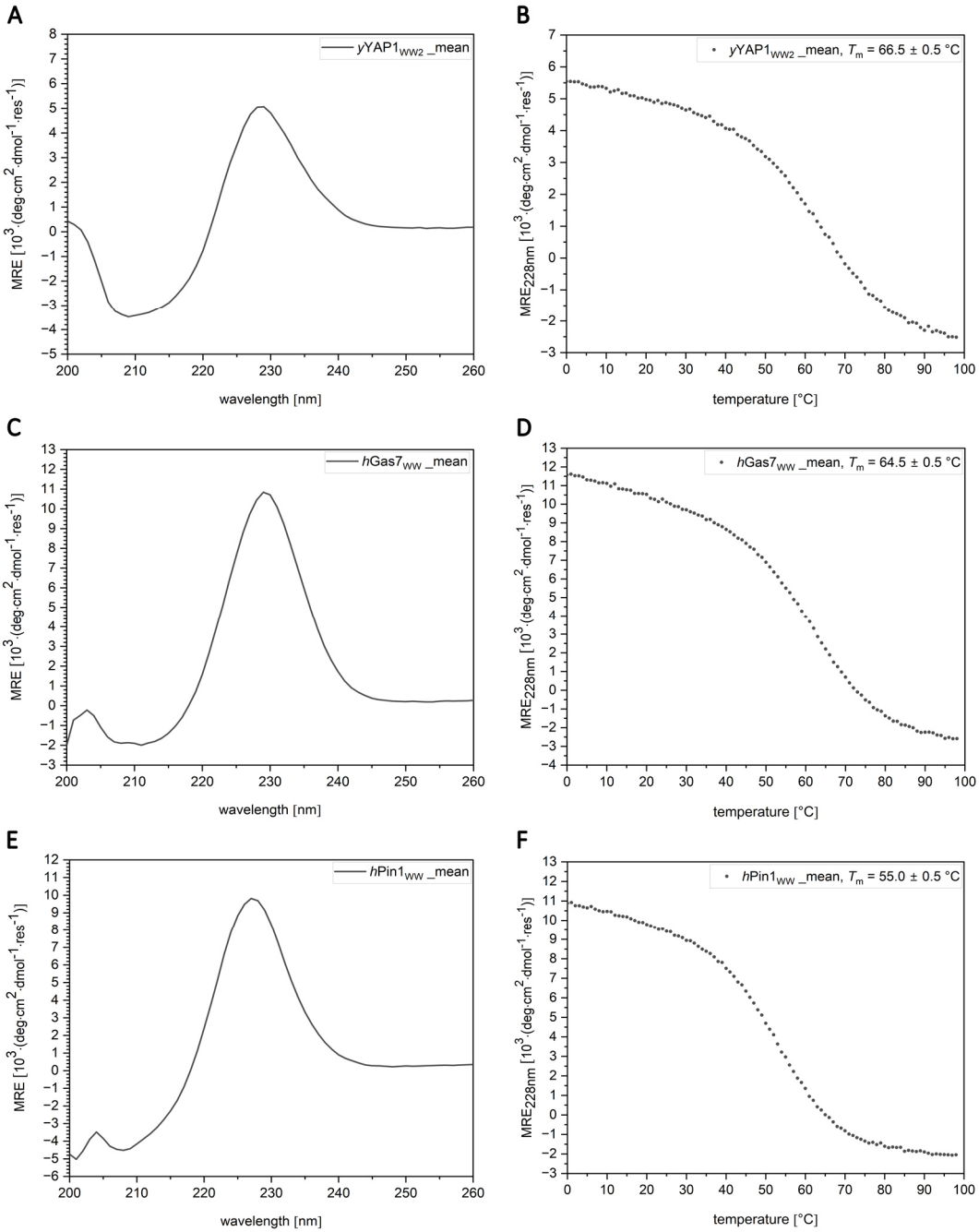


Figure S31: CD spectra and thermal denaturation curves of reference WW domain peptides. CD spectra measurements of **A**) *y*YAP1WW, **C**) *h*Gas7WW, **E**) *h*Pin1WW. Thermal denaturation measurements of **B**) *y*YAP1WW, **D**) *h*Gas7WW, **F**) *h*Pin1WW. For CD spectra and thermal denaturation curves the mean of triplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.5. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

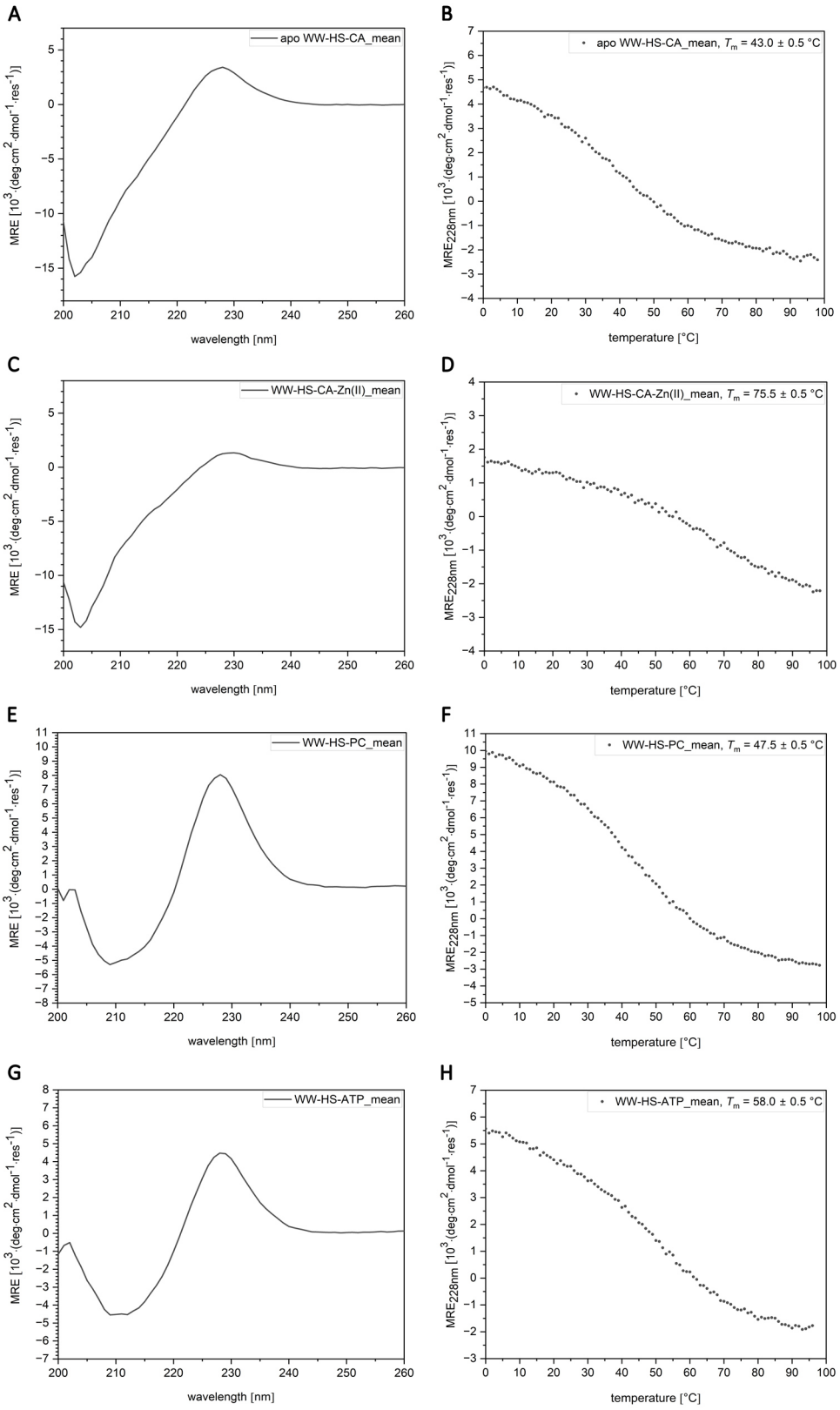


Figure S32: CD spectra and thermal denaturation curves of metal-binding and organophosphate-binding WW-HS peptides. CD spectra measurements of **A**) apo WW-HS-CA, **C**) WW-HS-CA-Zn(II), **E**) WW-HS-PC, **G**) WW-HS-ATP. Thermal denaturation measurements of **B**) apo WW-HS-CA, **D**) WW-HS-CA-Zn(II), **F**) WW-HS-PC, **H**) WW-HS-ATP. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, MOPS buffer (pH 7.2) and 100 µM $ZnSO_4$ for WW-HS-CA, PBS buffer (pH 7.5) for WW-HS-PC and WW-HS-ATP. These CD data have been published in modified form in Lindner *et al.* [94], © 2024 American Chemical Society.

8.4.3 CD measurements of d*e novo* designed ATP-binding peptides

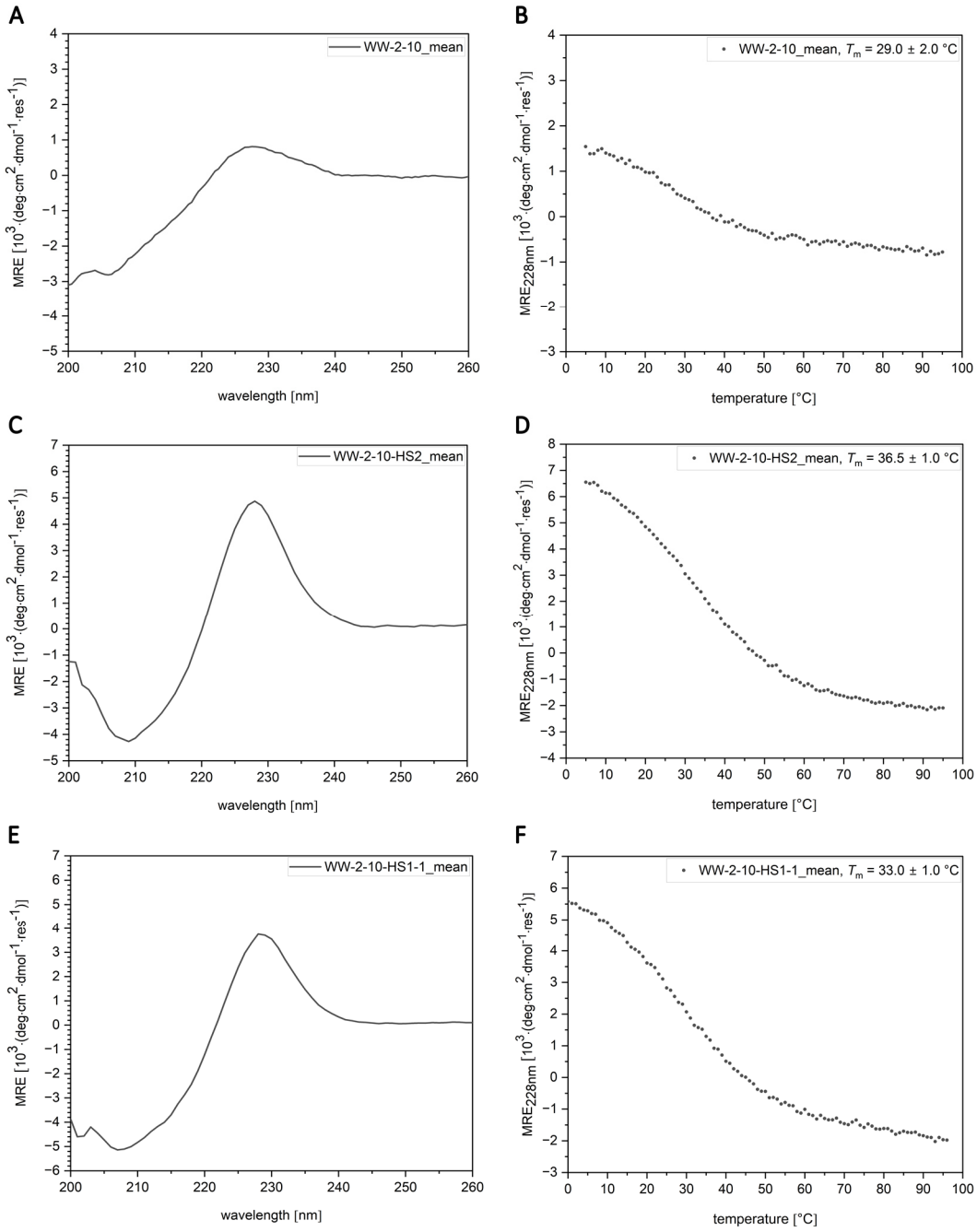


Figure S33: CD spectra and thermal denaturation curves of *de novo* designed ATP-binding peptides. CD spectra measurements of **A***)* WW-2-10 from Neitz *et al.* [85], **C**) WW-2-10-HS2, **E**) WW-2-10-HS1-1. Thermal denaturation measurements of **B**) WW-2-10 from Neitz *et al.* [85], **D**) WW-2-10-HS2, **F**) WW-2-10-HS1-1. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.4.

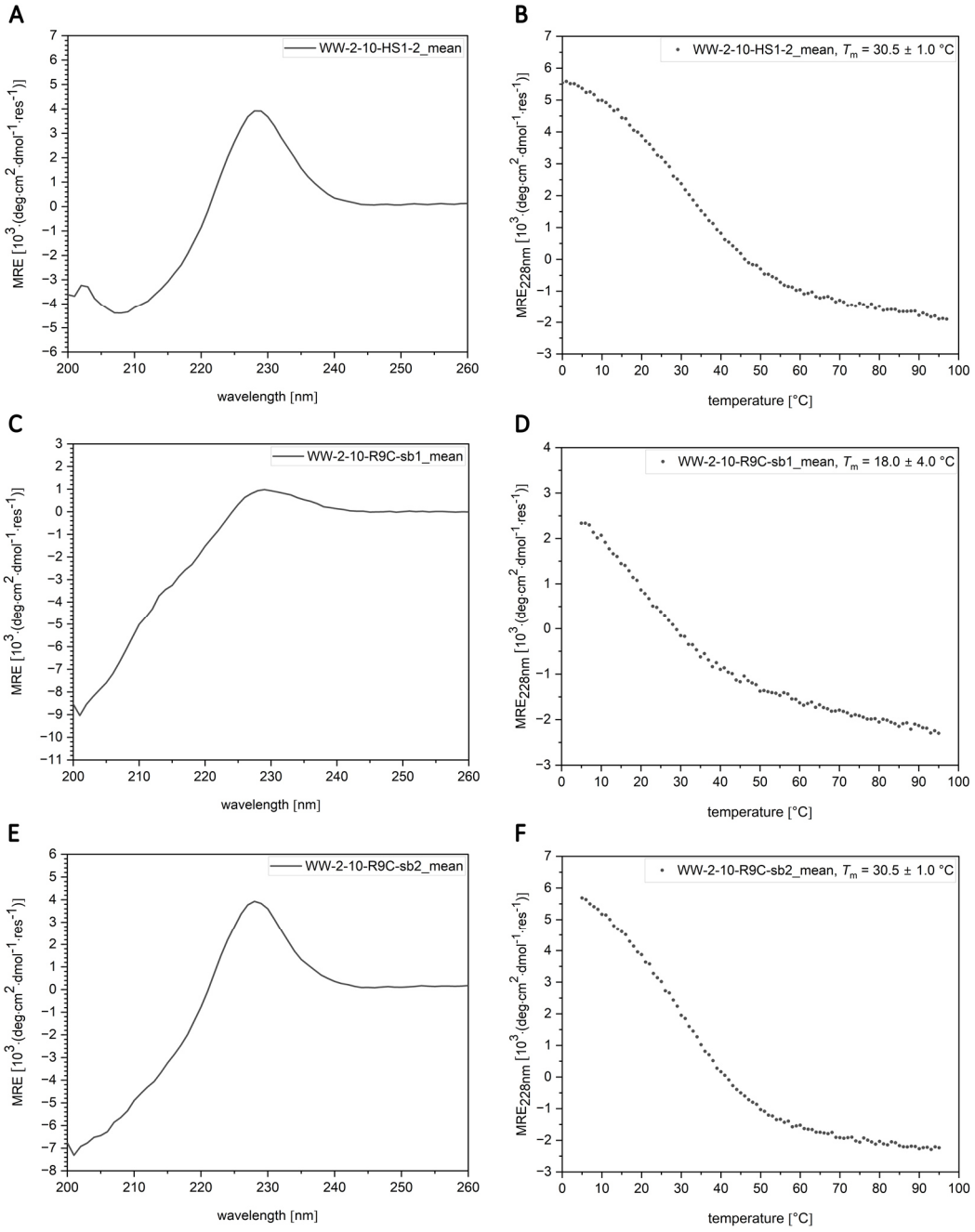


Figure S34: CD spectra and thermal denaturation curves of *de novo* designed ATP-binding peptides. CD spectra measurements of **A***)* WW-2-10-HS1-2, **C**) WW-2-10-R9C-sb1, **E**) WW-2-10-R9C-sb2. Thermal denaturation measurements of **B**) WW-2-10-HS1-2, **D**) WW-2-10-R9C-sb1, **F**) WW-2-10-R9C-sb2. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.4.

Peptide overnight incubation

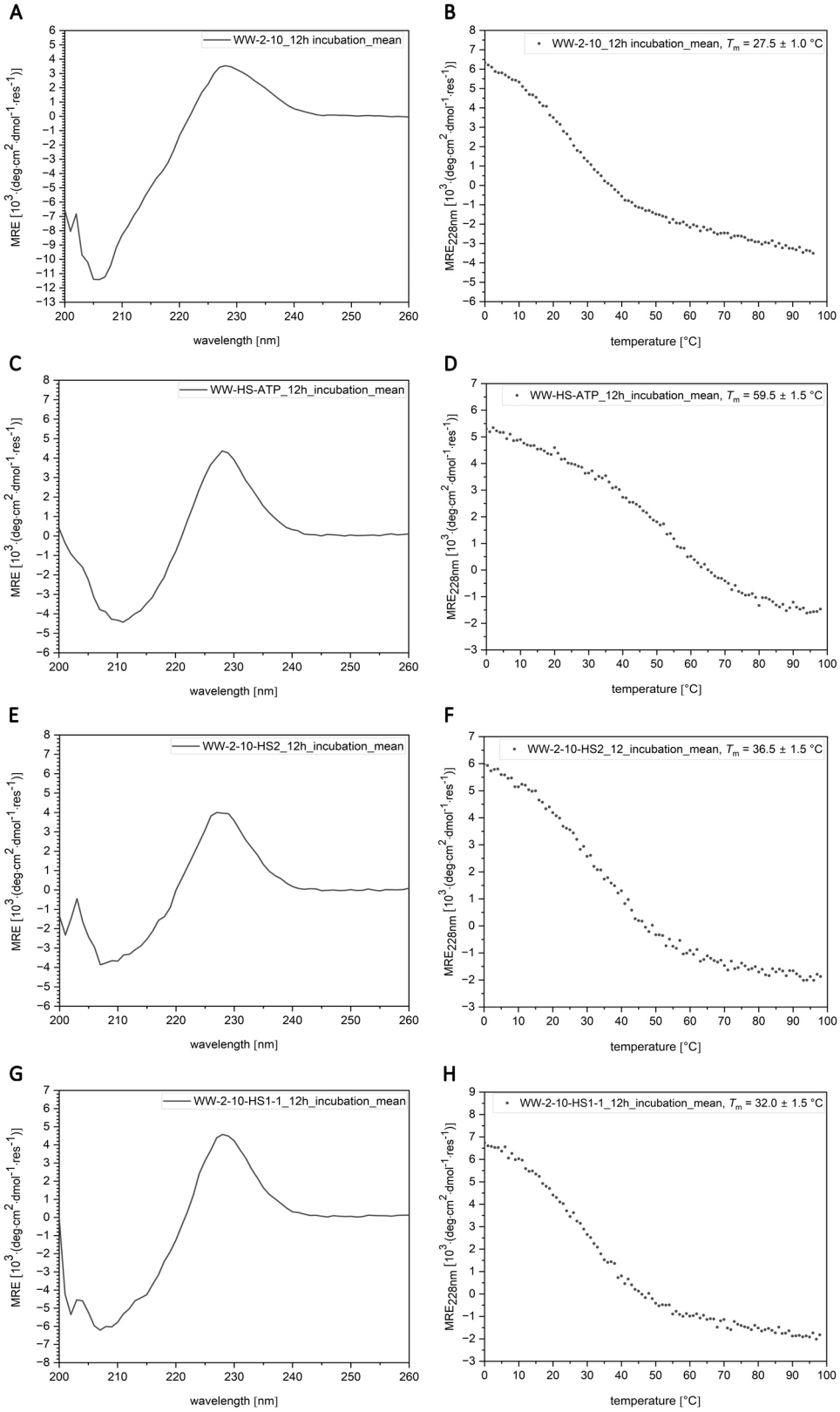


Figure S35: CD spectra and thermal denaturation curves of *de novo* designed ATP-binding peptides. CD spectra measurements of overnight incubations of **A***)* WW-2-10, **C**) WW-HS-ATP, **E**) WW-2-10-HS2, **G**) WW-2-10-HS1-1. Thermal denaturation measurements of overnight incubations of **B**) WW-2-10, **D**) WW-HS-ATP, **F**) WW-2-10-HS2, **H**) WW-2-10-HS1-1. For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, PBS buffer, pH 7.4, incubation time at least 12 h.

Peptide overnight incubation with ATP (1/1)

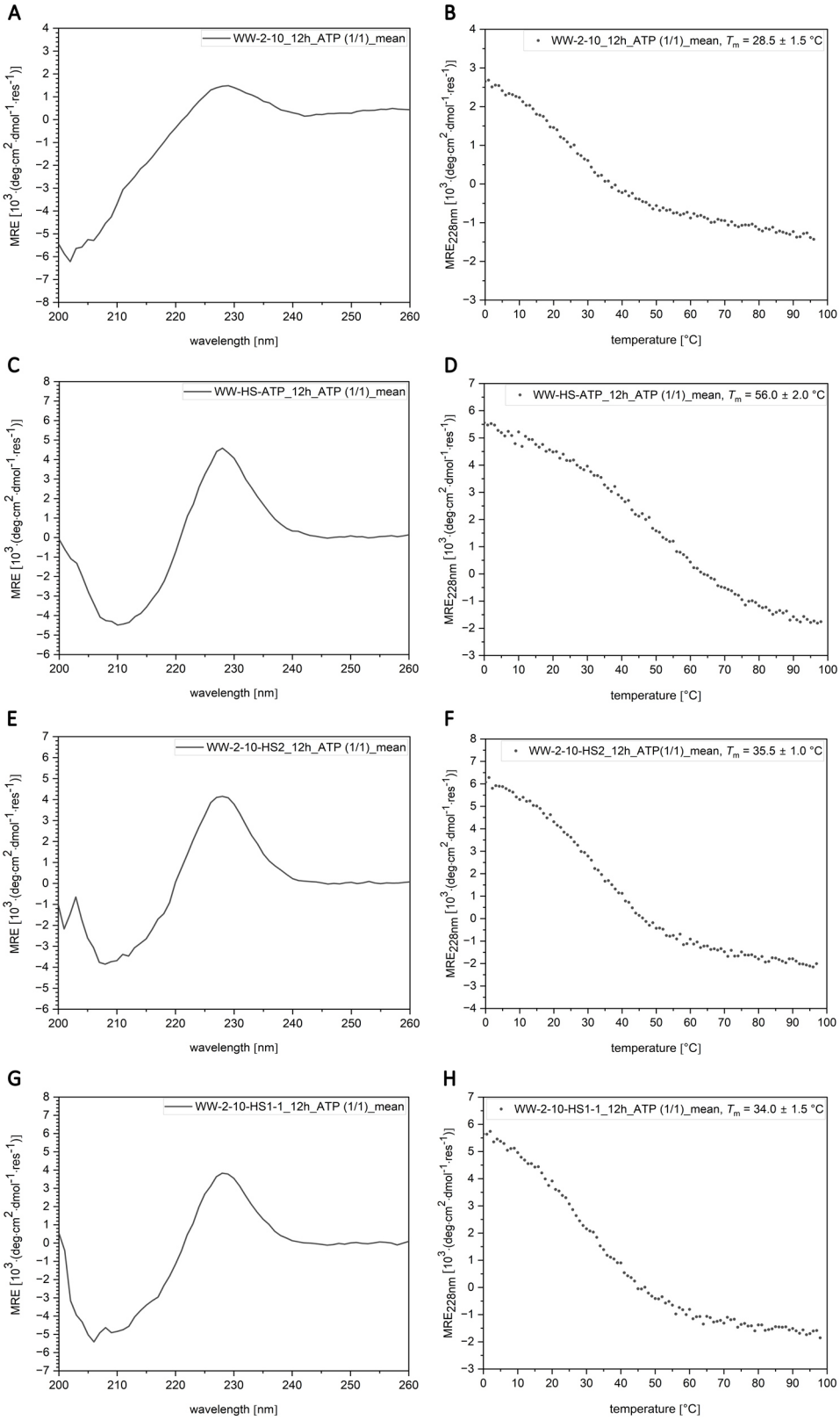


Figure S36: CD spectra and thermal denaturation curves of *de novo* designed ATP-binding peptides. CD spectra measurements of overnight incubations of **A***)* WW-2-10, **C**) WW-HS-ATP, **E**) WW-2-10-HS2, **G**) WW-2-10-HS1-1, with ATP (ratio 1/1). Thermal denaturation measurements of overnight incubations of **B**) WW-2-10, **D**) WW-HS-ATP, **F**) WW-2-10-HS2, **H**) WW-2-10-HS1-1, with ATP (ratio 1/1). For CD spectra and thermal denaturation curves the mean of duplicates is shown. Conditions: 100 µM peptide, 100 µM ATP, PBS buffer, pH 7.4, incubation time at least 12 h.

Peptide overnight incubations with ATP (1/1) and Mg^{2+} (1/10)

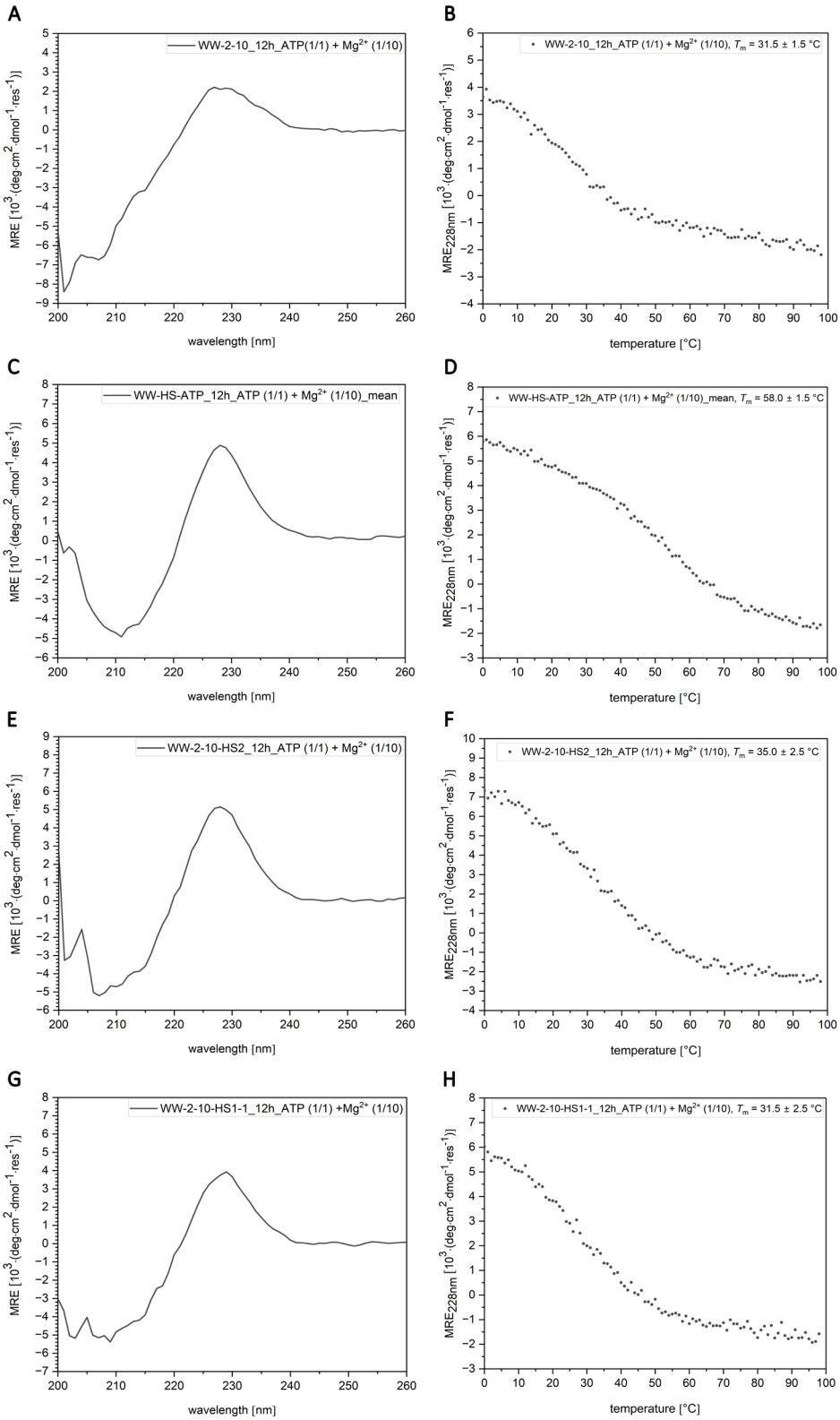


Figure S37: CD spectra and thermal denaturation curves of *de novo* designed ATP-binding peptides. CD spectra measurements of overnight incubations of **A***)* WW-2-10, **C**) WW-HS-ATP, **E**) WW-2-10-HS2, **G**) WW-2-10-HS1-1, with ATP (ratio 1/1) and Mg^{2+} (1/10). Thermal denaturation measurements of overnight incubations of **B**) WW-2-10, **D**) WW-HS-ATP, **F**) WW-2-10-HS2, **H**) WW-2-10-HS1-1, with ATP (ratio 1/1) and Mg^{2+} (1/10). For CD spectra and thermal denaturation curves a single measurement is shown, except for WW-HS-ATP, for which mean of duplicates are displayed. Conditions: 100 µM peptide, 100 µM ATP, 1 mM MgCl, PBS buffer, pH 7.4, incubation time at least 12 h.

8.5 NMR spectra

Table S5: NMR statistics of the analysis of the obtained ten conformers that exhibit the lowest energy. The NMR experiments were conducted by Prof. Michael Kovermann and coworkers of the University of Konstanz in collaboration. The table is reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

NMR distance and dihedral constraints	
Distance constraints	424
Total unambiguous NOE	378
intraresidue	143
sequential	100
medium	28
long	107
Total ambiguous NOE	46
Total dihedral angle restraints	27
phi	14
psi	13
Structure statistics	
Violations (mean and SD)	
Distance constraints, Å	0.022±0.002
Max. distance constraint violation, Å	0.026
Dihedral angle constraints, °	1.6±0.3
Max. dihedral angle violation, °	2.0
Ramachandran analysis	
Most favored, %	72.3
Additionally allowed, %	18.6
Generously allowed, %	4.4
Disallowed, %	4.6
Deviations from idealized geometry	
Bonds lengths, Å	0.002
Bond angles, °	0.568
Impropers, °	0.354
Average pairwise rmsd, Å	
Heavy, all	0.54±0.17
Backbone, all	0.53±0.18
Heavy, 2nd structure	0.28±0.10
Backbone, 2nd structure	0.27±0.08

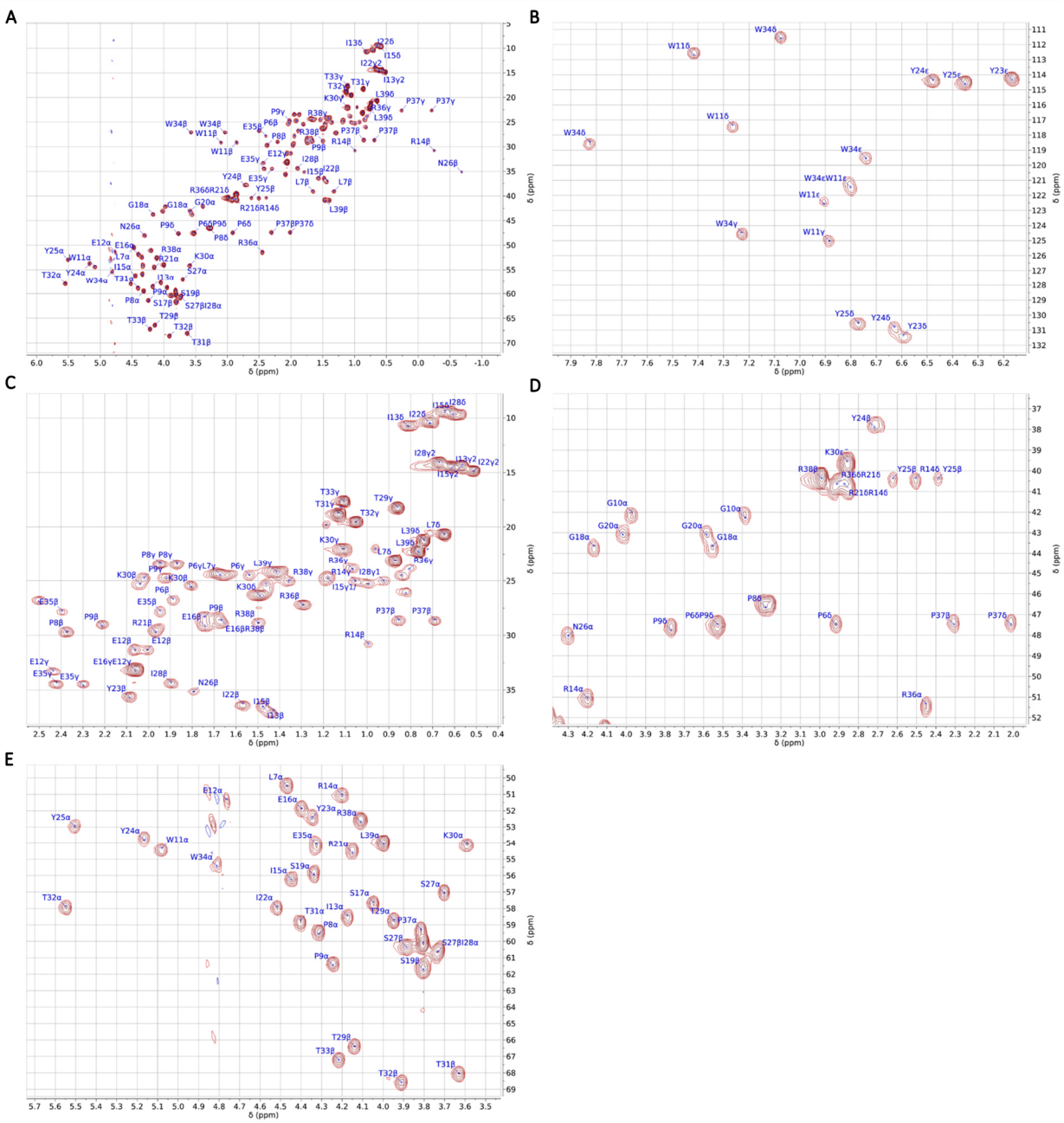


Figure S38: NMR spectra of two-dimensional ^{1}H-^{13}C HSQC NMR experiments conducted for WW-HS-2$_{OH}$. Cross peaks were fully assigned by analyzing 2D ^{1}H-^{1}H NOESY, 2D ^{1}H-^{1}H TOCSY and 2D ^{1}H-^{13}C TOCSY-HSQC NMR spectra. Correlations of CH, CH_2 and CH_3 are indicated with the one letter amino acid code, followed by the sequence position and the cabon atom side chain. Correlations of **A**) aliphatic, **B**) aromatic rings, including Trp and Tyr residues, **C**) CH_2 and CH_3, **D**) and **E**) CH and CH_2, are illustrated. NMR spectra measurements were conducted by Prof. Michael Kovermann and coworkers from the University of Konstanz in collaboration. This Figure was reused with permission from Lindner *et al.* [94], © 2024 American Chemical Society.

8.6 Additional *PyMol* graphs to WW-HS-1 and WW-HS-2

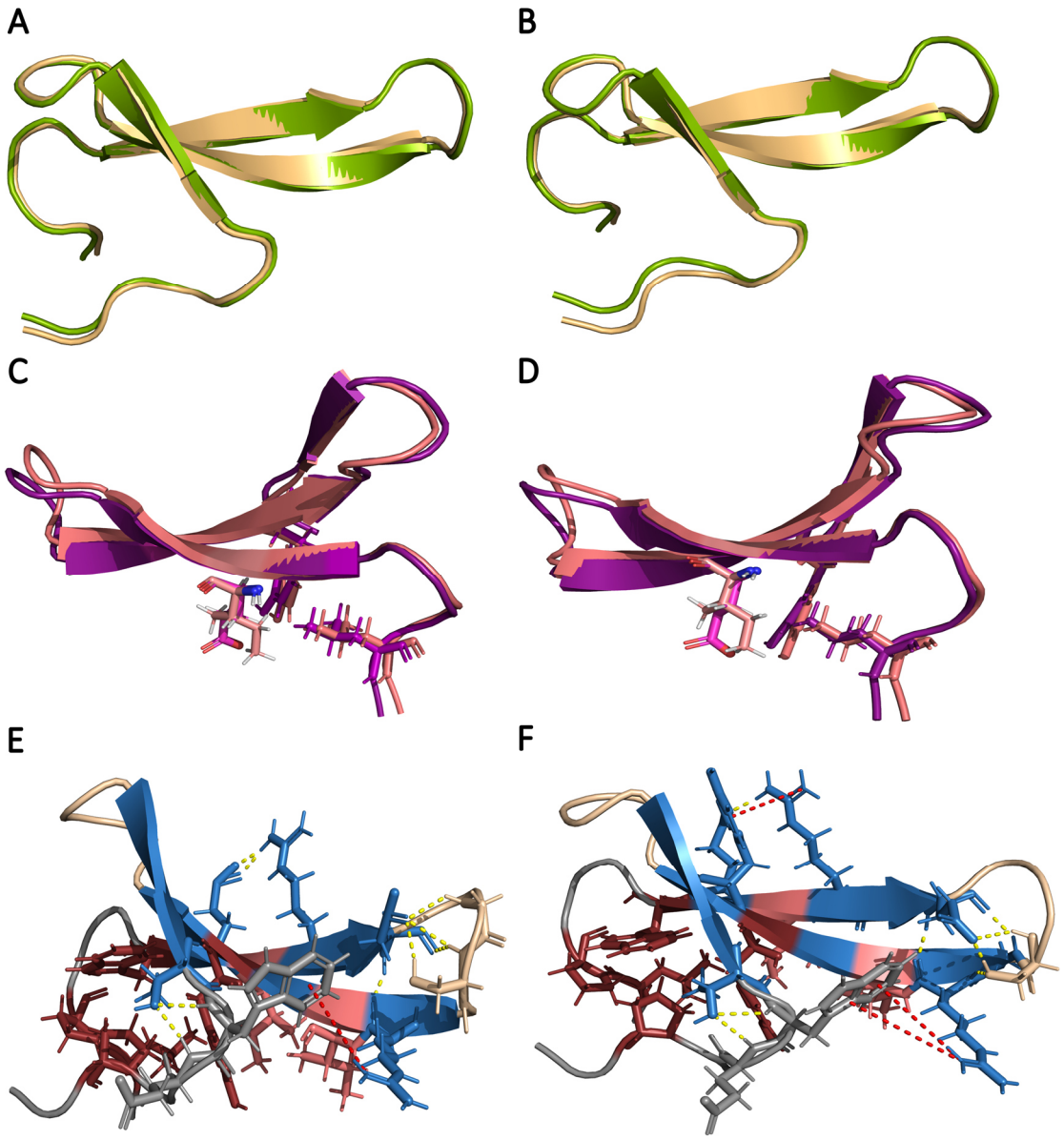
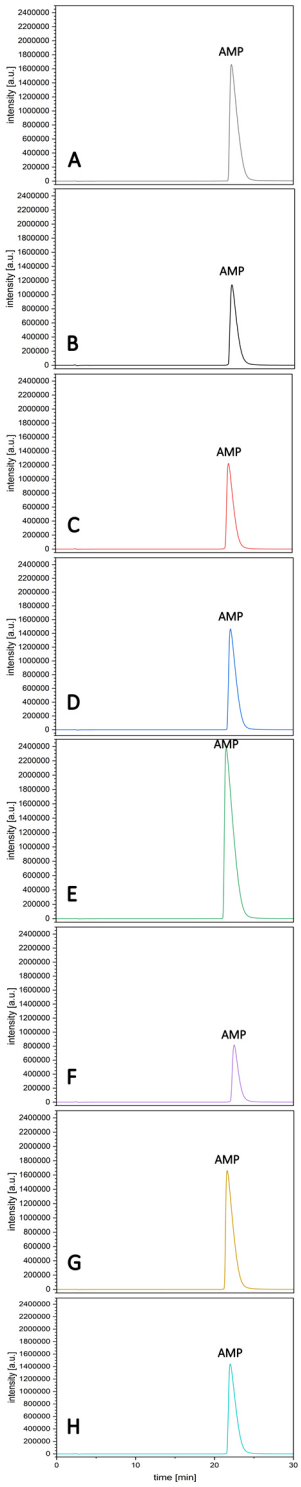


Figure S39: *In silico* structure predictions of WW-HS-1 and WW-HS-2. Comparison of structure predictions with *AlphaFold2* (green) and *OmegaFold* (beige) of **A**) WW-HS-1, and **B**) WW-HS-2. Comparison of *AlphaFold2* structures of WW-CS (purple) aligned with **C**) WW-HS-1 (rose), and **D**) WW-HS-2 (rose), displaying the residues Leu2, Tyr19 and Ile8, and Glu8 in WW-CS as sticks. Predicted *AlphaFold2* structures of **E**) WW-HS-1, and **F**) WW-HS-2, with residues of the hydrophobic core (red), Ile8 and Ile17 (rose), the BS segments (blue), the LP regions (beige), peptide termini (grey), aromatic interactions (red, dashed lines), and polar interactions (yellow, dashed lines).

8.7 Examinations on enzymatic activity

RP-HPLC Chromatogram of peptide-AMP incubations

1



2

A

Incubation sample	Peak	Retention time (R_t) [min]	Peak area [%]
1 mM AMP	AMP	22.2	100
WW-2-10 + AMP	AMP	22.1	100
WW-2-10-HS2 + AMP	AMP	21.9	100
WW-2-10-HS1-1 + AMP	AMP	22.1	100
WW-2-10-HS1-2 + AMP	AMP	21.5	100
WW-HS-ATP + AMP	AMP	22.5	100
WW-2-10-R9C-sb1 +AMP	AMP	21.6	100
WW-2-10-R9C-sb2 + AMP	AMP	22.0	100

Figure S40: RP-HPLC chromatograms of peptide-AMP incubations. 1 A) to **H**) Chromatograms of peptide-AMP incubations (1:10 ratio). **1 A**) AMP reference sample without peptide. AMP with **1 B**) WW-2-10, **1 C**) WW-2-10-HS2, **1 D**) WW-2-10-HS1-1, **1 E**) WW-2-10-HS1-2, **1 F**) WW-HS-ATP, **1 G**) WW-2-10-R9C-sb1, **1 H**) WW-2-10-R9C-sb1. 2 A) Retention time t_R of the peaks of AMP incubations and the peak area in percent. The AUC of the peaks were determined by integration with the *OriginPro* Peakanalysis software. Following settings were used: 100 µl injection of the sample, absorbance wavelength was monitored at 254 nm, isocratic setting for 30 min using a 50 mM KH_2PO_4 buffer (pH 6.8), and a flow rate of 0.6 ml/min.

Danksagung

Zu Anfang möchte ich mich ganz herzlich bei Frau *Jun.-Prof. Dr. Franziska Thomas* für die tolle Betreuung, das ehrliche Feedback und die stetige Unterstützung während meiner Promotion bedanken. Zudem bin ich sehr dankbar dafür, dass ich dieses spannende und vielversprechende Thema bearbeiten durfte und mich neben der Aneignung von fachlicher Expertise, im Rahmen der Promotion auch persönlich weiterentwickeln konnte.

Des Weiteren gilt mein Dank Herrn *Prof. Dr. Christian Klein*, für die stetige organisatorische Unterstützung, sowie die Möglichkeit das CD-Spektrometer im Institut uneingeschränkt und exzessiv nutzen zu dürfen. Auch möchte ich mich für das großzügige Angebot das Praktikum für den Masterstudiengang Molekulare Biotechnologie mitorganisieren und betreuen zu dürfen, bedanken.

Ich möchte mich natürlich auch bei dem gesamten *Arbeitskreis Thomas* für die vielen schönen Jahre und insbesondere für die gegenseitige Unterstützung und gute Arbeitsatmosphäre, den guten Zusammenhalt und die vielen humorvollen Momenten bedanken. *Vanessa Reusche* und *Truc Lam* danke ich für die vielen Ratschläge und ihre Hilfsbereitschaft bei auftretenden Schwierigkeiten. Ich möchte mich auch herzlich bei *Florian Häge* und *Niklas Schwegler* für ihre fachliche und emotionale Unterstützung, sowie eure Bereitschaft mit mir auch mal eine dringend benötigte Kaffeepause zu machen und uns bei Bedarf gegenseitig aufzubauen, bedanken. *Roman Graebner*, *Thomas Heim* und *Marius Werner* danke ich für ihre humorvolle Art, dadurch konnten wir oft zusammen scherzen und gemeinsam über unsere auch ab und zu sehr flachen Witze lachen. Herzlichen Dank auch an *Laura Brüchle* für die guten Gespräche und die lustige Zeit beim German Peptide Symposium in Jena. Den neuen Doktoranden *Marcos Conde González*, *Agi Kishore*, *Chiara Keil* und *Jonas Huber* bin ich dankbar für die vielen interessanten Gespräche und eure Begeisterung und Mithilfe bei der Organisation von Gruppenaktivitäten, wie beispielsweise unserem legendären Karaoke-Abend.

Neben dem AK Thomas, möchte ich mich natürlich auch bei dem *AK Klein* für die herzliche Aufnahme in die Gruppe in meiner Zeit der Praktikumsbetreuung und den vielen Stunden vor dem CD-Spektrometer danken. Herrn *Heiko Rudy* danke ich für seine große Hilfsbereitschaft bei technischen Problemen mit dem CD-Spektrometer und den guten Gesprächen während meiner Messzeiten. Zudem möchte ich mich bei *Johannes Lang* für die guten Unterhaltungen, Kaffee Treffen im Botanik Café und die gute Freundschaft bedanken. Bei *Mila Leuthold* möchte ich mich für die gute Zeit und Hilfe während meiner Praktikumsbetreuung im AK Klein danken. Ich freue mich sehr, dass ich *Franziska Grün* im wöchentlichen IPMB Seminar kennenlernen konnte und durch sie Mitglied in der Vereinigung für Chemie und Wirtschaft geworden bin. Dort haben wir uns neben der Universität regelmäßig bei interessanten Stammtischtreffen mit vielen Gleichgesinnten ausgetauscht und angefreundet.

Ich danke auch all meinen Studenten, die ich während ihrer Praktikumszeit, Hilfswissenschaftlertätigkeiten oder Abschlussarbeiten betreuen durfte. Die Betreuung hat mir immer viel Freude bereitet und auch ich habe viel aus der Betreuungszeit lernen und die Zukunft mitnehmen können.

Zudem möchte ich auch meiner guten Freundin *Jules* aus der Schulzeit dafür danken, dass wir unsere Freundschaft und gegenseitige Unterstützung die vielen Jahre trotz seltener Treffen nie vernachlässigt haben und es sich auch nach Wochen oder Monaten immer so angefühlt hat, als hätten wir uns erst gestern gesehen.

Mein besonderer Dank gilt natürlich auch meiner Familie, meinen Eltern *Elisabeth* und *Wilhelm*, die mich in meiner gesamten Studienzeit und auch der Promotion in jeder erdenklichen Weise unterstützt haben, hinter mir standen, und immer an mich geglaubt haben. Ohne euch wäre ich nicht so weit gekommen! Dass ich mich immer auf euch verlassen kann, bedeutet mir sehr viel. Auch meiner großen Schwester *Julia* bin ich unendlich dankbar für ihre emotionale Unterstützung, ihre aufbauenden Worte und ihren Rückhalt, wann immer ich ihn gebraucht habe. Ich bin überglücklich dich als meine Schwester zu haben.

Nicht zuletzt möchte ich meinem Ehemann *Gino* für seine Unterstützung danken. Auch wenn ich oft, viele zusätzliche Stunden im Labor verbracht habe, hast du mir immer den Freiraum gegeben, meine beruflichen Ziele zu verfolgen, auch wenn das für dich bedeutet hat zurückstecken zu müssen. Dank dir habe ich bei gemeinsamen Unternehmungen oder Urlauben wieder Kraft und Motivation schöpfen können. Du bist mein Ruhepol in stürmischen Zeiten und ich freue mich auf die gemeinsame Zukunft mit dir.