Enzymatic tools for peptide ligation and cyclization

*Development and applications*

Schmidt, M.

Link to publication

*Creative Commons License* (see https://creativecommons.org/use-remix/cc-licenses):

*Other*

*Citation for published version* (APA):


*General rights*

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

*Disclaimer/Complaints regulations*

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
CHAPTER 4

Efficient Enzymatic Cyclization of Disulfide-Rich Peptides using Peptiligases

Part of this chapter have been published in


and

Abstract

Disulfide-rich macrocyclic peptides, e.g. cyclotides, represent a promising class of molecules with potential therapeutic use. However, their efficient synthesis at large scale still represents a major challenge. Here we report new chemo-enzymatic strategies using peptiligase variants, i.a. omniligase-1, for the efficient and scalable one-pot cyclization and folding of the native cyclotides MCoTI-II, kalata B1 and variants thereof as well as of the θ-defensin RTD-I. The synthesis of the kalata B1 variant T20K has been successfully demonstrated at multi-gram scale. Several ligation sites for each macrocycle render this approach highly flexible and facilitate both the larger scale manufacture and the engineering of bioactive, grafted cyclotide variants, therefore clearly offering a valuable and powerful extension of the existing toolbox of methodologies for peptide head-to-tail cyclization. In addition, the combination of two distinct peptiligase variants in a dual enzymatic approach enables the preparation of modified cyclotides or dimeric fusion variants thereof, potentially representing a useful approach in academic research.
Introduction

In the last decade, macrocyclic peptides have gained increased traction as a promising class of therapeutics. High metabolic stability and often higher potency compared to linear analogues render them excellent leads in drug design. Several mono- and multicyclic peptides are currently under clinical evaluation.[1] Disulfide-rich macrocycles such as cyclotides or θ-defensins are particularly promising classes of molecules for therapeutic application and cyclotides also hold potential as crop-protection agents in the agricultural industry.[2] Cyclotides such as the prototypic examples kalata B1 and MCoTI-II are characterized by a combination of a head-to-tail cyclic backbone and three conserved disulfide bonds forming a distinctive, knotted pattern, which makes them exceptionally stable (metabolically, thermally and chemically).[3–5] In contrast, θ-defensins such as the rhesus θ-defensin (RTD-1)[6] are cyclic peptides of 18 amino acids in length with two anti-parallel β-strands stabilized by a laddered arrangement of three disulfide bonds (I–VI, II–V, III–IV). θ-defensins are of mammalian origin and possess a range of antimicrobial activities.[6] Both cyclotides and θ-defensins are being used as pharmaceutical tools, particularly as scaffolds onto which bioactive epitopes can be fused.[3,7] Due to their highly conserved structure they can accommodate the introduction of bioactive epitopes (“grafting”) between the cysteines while retaining their native structural and biophysical characteristics. There are now several examples showing that the cyclotide framework can be used to design drug leads.[8–15]

Despite their pharmaceutical potential, the preparation of disulfide-rich macrocyclic peptides, especially at scale, still poses a significant challenge and novel tools for facilitating their design and (scalable) manufacture are highly desired. Although chemical approaches, such as native chemical ligation (NCL)[16] or carbodiimide-mediated couplings using entirely protected peptides,[17] are well established[17–20] and have been successfully applied for peptide head-to-tail cyclization in many studies, the requirement for careful optimization of the reaction conditions, lack of scalability and the often low yield of the synthesis limit their application on larger scale.[21] Thus, there is a clear need for novel, more efficient methods for peptide head-to-tail cyclization and, hence, the preparation of macrocyclic disulfide-rich peptides. In recent years, various research groups investigated enzymatic strategies to fill this unmet need.[22,23] Several examples have been described including the use of sortase A,[24–27] split inteins,[28–32] trypsin,[33,34] asparaginyl endopeptidases such as butelase 1[35–37] or OaAEP Ib[38] and, finally, engineered subtilisins[39] such as peptiligase variants.[22,40,41] (see chapter 1). In addition, recombinant expression of disulfide-rich peptides in bacteria has been reported.[42–47] Most of the enzymatic approaches investigated suffer from incomplete ligation reactions and low catalytic efficiency, and additionally leave a ligase “footprint”, an unavoidable enzyme recognition sequence at the coupling site. In contrast, peptiligase based enzyme variants such as omniligase-1 recently emerged as a very powerful tool for traceless (footprint free) enzyme-mediated peptide bond formation, both for inter- and intramolecular (cyclization) of unprotected peptides in aqueous solution (see chapter 2).[8]
Based on the basic cyclization reactions that have been described using omniligase-1 (see chapter 2) we sought to expand the scope of omniligase-1 mediated peptide cyclization to the class of disulfide-rich macrocycles, including the prototypical examples MCoTI-II, kalata B1 and RTD-1. Especially the presence of multiple cyclization sites per molecule to create increased synthetic flexibility as well as the possibility to cyclize pre-oxidized/ folded cyclotides would certainly expand the scope of this methodology for the benefit of the chemical biology community.

**Results and Discussion**

**MCoTI-II**

Our initial attempts to apply omniligase-1 catalyzed cyclization to disulfide-rich macrocyclic peptides focused on a member of the Möbius family of cyclotides, the 34-mer head-to-tail cyclic trypsin inhibitor MCoTI-II, naturally occurring in the plant *Momordica cochinchinensis*. Reduced MCoTI-II is known to form its cyclic cystine knot structure easily to adopt its native, oxidized conformation. Although several synthetic protocols for the synthesis of cyclotides have been developed, efficient head-to-tail cyclization still remains a challenge and a chemo-enzymatic approach bears great potential.

For MCoTI-II, based on a sequence-guided evaluation, we identified two potential ligation sites for omniligase-1 catalyzed cyclization: L1 (…VCPK // IL…) and L5 (…CRGN // GY…) (Figure 1). The coupling position was mainly chosen based on a hydrophobic amino acid in the P4 position, which is important for an efficient ligation, and in addition, proline was avoided in position P1’ and P2’. For a proof of principle study, ligation site L1 was chosen as a first target. The linear open chain (oc-) MCoTI-II-L1 precursor peptide Cam-
Efficient Enzymatic Cyclization of Disulfide-Rich Peptides using Peptiligases

ester (see Figure 2A) was assembled using standard Fmoc- solid phase peptide synthesis (SPPS). Subsequently, at a concentration of 1 mg/mL, the purified precursor peptide ester could be efficiently cyclized (Figure 2B) to reduced cyclic c-MCoTI-II-L1 within 30 minutes using omniligase-1. Stirring the reaction mixture for another 14 hours in a potassium phosphate solution (pH 8.5) containing 5 mM reduced glutathione (GSH) resulted in the formation of the intact, native cf-MCoTI-II. Although MCoTI-II is known to efficiently fold into its naturally occurring structure,[19] the correct disulfide connectivity as well as the overall structure was shown to be identical with native MCoTI-II using NMR structure determination. Both chemical shifts and NOESY spectra indicate the same native structure of MCoTI-II as determined before by means of X-ray crystallography[49] and NMR spectroscopy.[48–50] One single set of amide and amino resonances was observed, and most resonances appeared sharp in the various spectra.

Figure 2. A) Reaction scheme for the cyclization and oxidative folding of open chain oc-MCoTI-II-L1 in a one-pot reaction to yield cyclic folded cf-MCoTI-II-L1. B) HPLC traces of the ligation-cyclization-folding procedure after 0 min (oc-MCoTI-II-L1), 30 min (cyclic c-MCoTI-II-L1) and after 14 hours (cf-MCoTI-II).

This one-pot cyclization-folding procedure was shown to work equally well with crude, non HPLC purified, oc-MCoTI-II-L1-Cam-L-OH, highlighting the robustness of the enzymatic ligation step. Additionally, the scalability of the one-pot chemo-enzymatic approach was demonstrated by the preparation of the native cyclotide cf-MCoTI-II at 1 gram scale starting from crude oc-MCoTI-II-L1-Cam-L-OH. At a concentration of
4 mg/mL crude substrate the reduced intermediate c-MCoTI-II-L1 was formed efficiently without any multimerization observed. Upon subsequent oxidative folding for 20 hours in potassium phosphate buffer containing 5 mM reduced glutathione the final, native cf-MCoTI-II was obtained. Thus, due to the use of crude substrates, the number of steps required could be significantly reduced, i.e. only one final preparative HPLC-run was required instead of two separate purification steps for the precursor and product. The crude reaction mixture was purified via a single RP-HPLC purification and cf-MCoTI-II was obtained with over 95% HPLC purity in an overall yield of 8% based on the loading of the first amino acid on the resin. For a peptide of this length (34 amino acids) and complexity the yield is good but may be improved even further by optimizing the SPPS, cyclization/folding and purification protocol.

Various studies have been conducted in which bioactive epitopes have been grafted into flexible loops of the squash trypsin inhibitor cyclotide MCoTI-II, particularly into loops 1 and 6 (see cartoon representation of MCoTI-II: Figure 1).[7,8,12,15] Since the grafted epitopes typically vary in their sequence, a cyclization site distinct from the already established coupling site in loop 1 (ligation site L1: …VCPK // IL…; see before) and, moreover, outside loop 6, e.g. in loop 5, would be desired. In this way, the generation of chemically synthesized libraries would clearly be facilitated. Encouraged by this idea, we attempted to identify suitable ligation sites in loop 5 using a sequence-based computational evaluation. With this approach we identified a second ligation point (site L5: …CRGN // GY…) that should enable efficient omniligase-1 catalyzed cyclization. Following the preparation of the open-chain (oc) elongated carboxamidomethyl (Cam)-ester precursor MCoTI-II-L5-Cam-L-OH using classical SPPS,[40,51] our first cyclization attempts using omniligase-1 successfully resulted in the formation of c-MCoTI-II-L5. However, the linear C-terminal ester precursor MCoTI-II-L5-Cam-L-OH was not entirely stable during the course of the reaction as base-induced dehydration-type of reaction of P1-asparagine was observed (Figure 3).[52]

![Figure 3](image)

Figure 3. Asn dehydration-type of reaction.[53] B = base; X = OHMBA-L-OH.

Based on this finding, we decided to use a less activated, more stable C-terminal aryl-type of ester (hydroxymethylbenzoic acid (HMBA)-ester) that, in previous studies, was demonstrated to be well suited for omniligase-1 catalyzed chemo- enzymatic peptide synthesis (CEPS) on larger scale (see chapter 2) and can be easily introduced using classical SPPS (see experimental section). In addition, peptiligase variant #1015,
optimized for the recognition of small, polar amino acids in position P1, was expected to demonstrate a better performance than omniligase-1 in this particular ligation. Indeed, the HMBA-ester was found to be much less amenable to Asn dehydration and efficient conversion of the open chain oc-ester precursor MCoTI-II-L5 into its cyclic counterpart was observed within 30 min using peptilagase #1015. Overnight air oxidation in combination with the addition of 5 mM GSH\([40,54]\) yielded native (cyclic folded- cf) MCoTI-II (Figure 4) in a straightforward manner via a one-pot cyclization-folding procedure.

![Figure 4](image.png)

**Figure 4.** HPLC trace of the MCoTI-II-L5 one-pot cyclization-folding approach: oc-MCoTI-II-L5 HMBA-L-OH (0 min) is readily converted to cyclic c-MCoTI-II-L5 (30 min), followed by folding (2 mM GSH, pH 8.3) to adopt its native, oxidized confirmation (cf-MCoTI-II-L5; 16 h).

The correct disulfide connectivity was confirmed by the identical HPLC retention time (Rt) compared with chemically synthesized MCoTI-II. The trypsin inhibition activity of cf-MCoTI-II-L5 was subsequently determined and compared to that of chemically synthesized MCoTI-II (see experimental section: Figure 16). The Ki values were 0.0041 and 0.0050 nM, respectively, which is comparable to the previously reported Ki value of native MCoTI-II (0.0023 nM)\(^[9]\).

In general, the identification of this second ligation site in loop 5 of MCoTI-II (L5: …CRGN // GY…) represents a viable alternative to the one described in loop 1 (L1: …VCPK // IL…)\(^[40]\) and opens up the possibility for a modular synthesis of MCoTI-II. In addition, the identification of this second ligation site significantly expands the scope of using peptilagases for the rapid and simple preparation of MCoTI-II variants for grafting studies, i.e. the C-terminal part of MCoTI-II (loop 1-5) can be synthesized at larger scale and upon splitting of the resin derivatization with diverse sequences in loop 6 can be performed conveniently (see Figure 5).
Many disulfide-rich macrocycles adopt their native conformation only when formation of the correct disulfide pattern precedes backbone cyclization. Hence, using MCoTI-II as a model system, we also investigated whether the reaction order of enzymatic cyclization followed by oxidative folding could be reversed. This would significantly expand the scope of the CEPS macrocyclization platform and minimize the amounts of misfolded isomers, thus resulting in higher overall yields. Again, a C-terminal HMBA-L-OH ester was chosen, which proved to be stable under the basic and oxidative conditions required for oxidative folding. Oxidized open chain (oc-f) MCoTI-II-L1 HMBA-ester was obtained after applying oxidative folding conditions (pH 7.8, 5 mM GSH) for several hours. 1D-1H-NMR revealed that MCoTI-II-L1 HMBA ester adopted a native-like fold (I-IV, II-V, III-VI, see experimental section Figure 14). The well-dispersed peaks observed within the amide proton region (5-9 ppm) indicated that this peptide had a well-defined structure. Unfortunately, we were not able to obtain (oc-f) MCoTI-II-L5 HMBA-ester, because dehydration of Asn in position P1 was too severe under oxidative folding conditions. In contrast to our expectations, cyclization of (oc-f) MCoTI-II-L1 HMBA ester proceeded with high efficiency and the native product was obtained, as shown by NMR secondary shift comparison with native MCoTI-II (see Figure 6).

Figure 5. Schematic representation of a facilitated approach for grafting bioactive epitopes into loop 6 of MCoTI-II.
Figure 6. Comparison of Hα secondary chemical shifts of enzymatically cyclized MCoTI-II-L1 (green bars) and chemically cyclized MCoTI-II (grey bars) (below).

Efficient cyclization of the oxidized ester again exemplifies the broad substrate scope and flexibility of omniligase-1 mediated ligation. We anticipated the disulfide bridge Cys8-Cys25, which has to accommodate Cys8 in pocket S3, to render the peptide rigid and stericly hindered to adopt a conformation for binding that would be favorable for efficient cyclization. To rationalize the catalytic activity observed when we examined cyclization of the fully oxidized open chain MCoTI-II-L1 precursor (oc-f)-MCoTI-II-L1, we modeled the binding mode of cf-MCoTI-II in omniligase-1 using a combination of knowledge-guided docking and molecular dynamics (MD). We opted to model the binding mode of the fully oxidized and cyclized product cf-MCoTI-II because the 3 disulfide bonds and the head-to-tail cyclic peptide bond considerably restrict its conformational freedom, thus making it more challenging to bind to omniligase-1 than any partially reduced MCoTI-II. According to the model cf-MCoTI-II can bind omniligase-1 with loop 1, loop 6 and loop 3 at the peptide-enzyme interface, and with peptide’s amino acids P4, P2, P1 and P2’ positioned in the respective subsites. Val7 (P4) binds in the hydrophobic pocket formed by Leu78, Tyr86 and Leu126 and this appeared a very stable interaction when examined by MD simulations. Cys8 (P3), due to the disulfide bond with Cys25, is displaced from the S3 subsite and does not bind as deep in the omniligase-1 cleft as the P3 histidine from the propeptide present in the crystal structure of peptiligase variant thymoligase (pdb 5OX2). The S2 binding site is slightly altered due to the displacement of Leu126. This allows Pro9 (P2) to bind in the S2 subsite and form a hydrophobic contact with Leu126. Lys10 (P1) binds in the same site as proposed for thymosin-α1 in thymoligase (see chapter 3). However, in omniligase-1, Asp166 is mutated to a serine, thus there is no ionic interaction and the preferred Lys10 side-chain interactions are hydrogen bonds with the backbone of Gly127 and Cys8. We specifically searched for reactive binding modes, so the backbone of Lys10 is bound with the carbonyl oxygen within hydrogen-bond distance of the backbone of Cys221 and the side-chain of Asn155 (the oxyanion hole), while the carbonyl carbon is relatively close (3.9 Å) to the sulfur atom of the nucleophilic Cys221. The binding of the propeptide present in PDB 5OX2 shows that the substrate backbone from P6 to P2 binds in an extended conformation, forming an anti-parallel β-sheet with the β-strand from Val95 to Gln103. This is not observed in the binding mode we found for the acyl-donor part of cf-MCoTI-II, probably
due to three reasons. First, the acyl-donor part of MCoTI-II cannot adopt the required extended conformation due to the Cys8(P3)-Cys25 disulfide bond that displaces Cys8 and also because the P5 and P6 amino acids cannot interact with the omniligase-1 interface due to the compact folding of MCoTI-II. Third, the Cys8-Cys25 bond brings loop 3 from MCoTI-II to the interface with omniligase-1 and pushes Val95 to Gln103 away from the amino acids P1 to P4, precluding the formation of the typical anti-parallel β-sheet binding mode (Figure 7).

![Figure 7](image_url)

**Figure 7.** Comparison of the binding of the P6 to P2 amino acids in 5OX2 thymoligase (A- dark green) with the bound propeptide (A- light green) and the binding of cf MCoTI-II (B- yellow) in the modeled omniligase-1 (B- brown). The β-sheet structure formed by the β-strand Val95 to Gln103 and the P6-P2 amino acids (A) is not formed in the omniligase-1 with cf MCoTI-II. This is due to the opening and disruption of the β-strand Val95 to Gln103 caused by loop 3 of MCoTI-II, which has a disulfide bond to P3 Cys8.

Regarding the binding of the acyl-acceptor part of MCoTI-II, Ile11 (P1’) is oriented towards the inside of the cyclic peptide, with its side-chain shielded by hydrophobic interactions with MCoTI-II Pro9, Pro22, the two cystines formed by the pairs Cys8-Cys25 and Cys21-Cys33, and the aliphatic chain of Lys13. Leu12 (P2’) binds in the S2’ subsite interacting with the hydrophobic Trp189 of omniligase-1. This interaction is very stable during the whole MD simulation.

Comparing the conformation of thymoligase bound to its propeptide fragment (pdb 5OX2) to the conformation of omniligase-1 bound to cf-MCoTI-II as found by MD simulations, the most marked difference is in the β-strand formed by Val95 to Gln103. In the omniligase-1 structure complexed with cf-MCoTI-II, this β-strand is disrupted to accommodate loop 3 of cf-MCoTI-II (Figure 7). The disruption requires the shielding of the hydrophobic group of Leu126 in the Ser125-Ser132 loop from the solvent with formation of stabilizing hydrophobic interactions with Val95, Leu96 and Val107 (Figure 8). The MD simulations thus indicate that the structural rigidity of cf-MCoTI-II caused by the P3-Cys8-Cys25 disulfide would result in steric clashes of loop 3 with the omniligase-1 Val95-Gln103 β-strand, which is alleviated by disruption of the β-strand. The distortion of this strand is unusual. A superposition of 83 subtilisins available from the Protein Data Bank (PDB)[56] indicates strong structural conservation, with only 1ST2 and 2ST1 showing partial distortion of the affected β-strand (Figure 8).
Figure 8. A) Tube representation of the enzyme backbone for the superimposition of the modeled omniligase-1 with bound cf-MCoTI-II (only shown in panel B) as pink sticks) and 83 subtilisins available in the pdb. B) Zoomed-in panel showing the interaction between cf-MCoTI-II and regions Val95-Gln103 and Ser125-132 of omniligase-1, with hydrophobic shielding of Leu126. When compared to the other X-ray conformations, the largest difference of omniligase-1 (blue tube representation) is in the β-strand formed by Val95 to Gln103 and the nearby loop formed by Ser125 to Ser132 (bottom). Only pdb structures 1ST2 and 2ST1 showed a similar disruption of the β-strand formed by Val95 to Gln103 (red tubes), but not as extensive as our modeled omniligase-1 (blue tubes).

Superimposition of cf-MCoTI-II from this binding mode with 30 NMR structures (pdb 1HA9) reveals similar conformations, underlining the structural rigidity of cf-MCoTI-II. Although native MCoTI-II might bind omniligase-1, it does not appear to be the most favorable binding mode. Therefore, we hypothesize that transiently reducing the disulfide bond Cys8 (P3)-Cys25 would allow MCoTI-II to more easily adopt a favorable conformation for efficient cyclization. Interestingly, when the reaction was performed without any GSH present, the reaction still resulted in the exclusive formation of cf-MCoTI-II, although at a significantly reduced rate as compared to the reaction with GSH present (reaction time of 18 h vs. 2 h). This finding indicates that cyclization most likely proceeds via a partially unfolded version of (oc-f)-MCoTI-II, potentially promoted by GSH as a disulfide shuffling agent.

Kalata B1

Encouraged by the successful chemo-enzymatic synthesis of MCoTI-II using two different connection points, we focused next on another prototypic example of the Möbius family of cyclotides, namely kalata B1. Based on a sequence-guided evaluation we identified two potential ligation sites: L2 (...)VGGT // CN...) and L6 (...CTRN // GL...) (Figure 9A and B). Our initial studies focused on ligation site L2 and the use of omniligase-1.
Figure 9. A) Cartoon representation of kB1 (pdb 1NB1) and sequence of kB1 with potential ligation sites indicated. Disulfide bonds are shown in yellow. B) Linear C-terminal esters for omniligase-1 catalyzed cyclization of kB1 and structures of respective esters used.

After extensive screening of various omniligase-1 cyclization reaction conditions, c-kB1-L2 was obtained in a reasonable yield of 60% (HPLC), albeit, accompanied with 40% hydrolysis of the glycolate ester moiety. We identified Asn in P2’ position to be sub-optimal, since in this position hydrophobic amino acids are generally preferred by omniligase-1. Indeed, introduction of an N14L mutation into kB1 increased the ligation yield to over 85%. Based on this result we switched our focus to cyclization site L6 identical to the natural cyclization site in the biosynthesis of kalata B1.\cite{57} Again, using a C-terminal HMBA-ester, P1-Asn dehydration was minimized and head-to-tail cyclization was clearly dominant with efficient conversion of oc-kB1-L6 HMBA-ester to c-kB1-L6.

Based on these promising results we opted to establish a one-pot cyclization-folding procedure for the scalable manufacture of kB1-L6. For these scale-up studies we used the [T20K]kB1 variant of kB1, which is under preclinical evaluation as an immunosuppressive agent for the treatment of multiple sclerosis.\cite{58} After chemical synthesis of linear [T20K]kB1-L6-HMBA-ester on larger scale (5 mmol), we achieved efficient cyclization in a one-pot reaction using omniligase-1 and crude, non-lyophilized [T20K]kB1-L6-HMBA-ester. In order to minimize the occurrence of P1-Asn dehydration the pH of the ligation mixture was lowered to 8.0 and the amount of enzyme was increased to 0.006 molar equivalents with respect to the linear precursor peptide. Following cyclization, upon addition of 2-propanol to a final concentration of 50% complemented with 2 mM GSH/GSSG\cite{20} (GSSG = oxidized glutathione) and incubation for 15 h, native [T20K]kB1 was obtained (Figure 10A). Pure chemo-enzymatically synthesized [T20K]kB1 was shown by NMR spectroscopy (Figure 10B) and LC-MS analysis to be identical to fully chemically synthesized [T20K]kB1. Besides structural verification using NMR, the chemo-enzymatically synthesized [T20K]kB1 showed to have similar cytotoxicity (IC$_{50}$ = 2.21 µM; see experimental section: Figure 15) towards HeLa cells as its chemically synthesized counterpart (IC$_{50}$ = 2.42 µM).
In order to ease the processing of large quantities and to further reduce the amount of ligase required we investigated dosing of the crude linear \([T20K]kB1\)-L6 ester over time in a batch-wise manner to increase the concentration of final product and thus significantly facilitate the downstream processing. Although omniligase-1 is, in contrast to naturally occurring cyclases such as butelase-1, easily available via recombinant expression in large quantities, a further decrease in the quantity of catalyst will result in an additional cost-price reduction of the synthesis at scale. Whereas most omniligase-1-catalyzed cyclization reactions are performed in the 0.5-1 g/L (low \(\mu\)M) range, starting from 0.5 g of crude linear \([T20K]kB1\)-L6 ester at a concentration of 1 g/L, we were able to increase the concentration of linear ester by repeated dosing (five times 0.5 g every 30 min) to 6 g/L. The amount of omniligase-1 required could be reduced three-fold to \(\leq 0.002\) molar equivalents. After oxidative folding and direct HPLC purification, cf-[T20K]kB1 was obtained with a cyclization-folding-purification isolated yield of 35%. Clearly, this chemo-enzymatic strategy is a viable approach for the production of larger quantities of [T20K]kB1 or variants thereof.
Dual Enzymatic Ligation for the Synthesis of Side-Chain Conjugated kB1 Variants

There is an interest to couple cyclotides to other molecules, e.g. to nanobodies, for the targeted delivery of cyclotides, or to dyes. For this purpose, the introduction of bioorthogonal or specifically reacting groups to the side-chain of selected amino acid residues would represent a suitable approach, as they would allow regiospecific and facile derivatization. Specifically, we envisaged a dual enzymatic ligation strategy by which we could first head-to-tail cyclize a linear cyclotide precursor, followed by conjugation and subsequent oxidative folding to adopt its native fold. For this, the previously investigated [T20K]kB1 using the established ligation site L6 (…CTRN // GL…) represented a suitable model system. For the conjugation reaction following head-to-tail cyclization the ε-NH₂ group of Lys20 can be modified with two additional amino acids, i.e. H-Asp-Leu, which can subsequently be recognized as the P1’ and P2’ amino acids in a second enzymatic ligation. Lys20 is well suited for side-chain derivatization since substitutions in loop 4 of kB1 are considered to be tolerated without losing its overall fold and stability. However, for the dual enzyme strategy to be applicable an exquisite selectivity for cyclization is required, i.e. for enzymatic cyclization of the [T20K]kB1-L6-HMBA-L-OH the N-terminus ‘GL’ must be preferred over the side-chain handle N-terminus ‘DL’. For this, ‘DL’ has been chosen specifically, since it represents an N-terminal sequence that is well recognized by the (selective) peptiligase variant thymoligase (see chapter 3) but is less preferred by omniligase-1. In addition, the formation of a macrocycle via the side-chain ‘DL’ handle would result in an approximately 13 amino acid sized cycle, which would be less preferred than the formation of the larger head-to-tail cycle (see chapter 2). Indeed, we exclusively observed formation of the desired larger cycle c-[T20K(DL)]kB1-L6, leaving the ‘DL’ side-chain N-terminus untouched and available for a second ligation reaction using thymoligase (Figure 11). The formation of the desired kB1-cycle was confirmed via Edman degradation of c-[T20K(DL)]kB1-L6, which resulted in the formation of c-[T20K(L)]kB1-L6, from which the side-chain handle aspartic acid has been cleaved off (MW_{calc.}: 3036.3 Da, MW_{exp.}: 3034.5 Da). After the addition of 50% (v/v) 2-propanol and 2 mM GSH/GSSG followed by incubation overnight, oxidatively folded cf-[T20K(DL)]kB1 was obtained (see Figure 11), as confirmed by mass spectrometry (MW_{calc}= 3146.0 Da, MW_{exp} = 3146.0 Da). In addition, upon oxidative folding, the characteristic right shift of kB1 in the HPLC spectrum was observed together with a decreased solubility of the molecule, as is the case for native kB1.
Efficient Enzymatic Cyclization of Disulfide-Rich Peptides using Peptiligases

![Diagram of enzymatic cyclization](image)

Figure 11. Omniligase-1 catalyzed head-to-tail cyclization of oc-[T20K(DL)]kB1-L6-HMBA-L-OH (0 min). Efficient formation of c-[T20K(DL)] was observed (45 min). Subsequent addition of 50% (v/v) 2-propanol and 2 mM GSH/GSSG followed by incubation over-night yielded oxidatively folded cf-[T20K(DL)]kB1 (16 h).

Knowing that [T20K(DL)]kB1 still adopts its native fold, we attempted to couple a fluorescently labelled model peptide (Ac-K(FITC)KFSKK-Cam-L-OH) to the ‘DL’ side-chain handle, first to the reduced form c-[T20K(DL)]kB1. Using thymoligase as a catalyst we observed efficient and clean ligation to form c-[T20K(DL-’FITC’)]kB1 within 15 min (Figure 12A). Besides the desired product, only hydrolyzed Ac-K(FITC)KFSKK-OH resulting from the excess (7 eq.) of ester starting material was observed as a reaction byproduct. Upon applying oxidative conditions (50% (v/v) 2-propanol, 2 mM GSH/ GSSG) and subsequent incubation over-night, the folded conjugate [T20K(DL-’FITC’)]kB1 was obtained efficiently (see Figure 12A), as proven by mass spectrometry (MW<sub>calc</sub>: 4328.8 Da, MW<sub>exp</sub>: 4328.0 Da) and HPLC retention shift comparison. Interestingly, omniligase-1 catalyzed cyclization of oc-[T20K(DL)]kB1-L6-HMBA-L-OH to yield c-[T20K(DL)]kB1 could also be followed by thymoligase catalyzed side-chain conjugation in a one-pot reaction. However, significantly higher equivalents of the FITC labelled peptide ester were required for efficient ligation, since the intermolecular reaction is less favored under the high dilution conditions required for cyclization. Hence, an intermediate HPLC purification or concentration step is recommended to ensure efficient intermolecular ligation in the second side-chain handle conjugation step. Intermolecular ligation to the side-chain handle followed by enzymatic cyclization was not attempted, since the enzymatic discrimination between two different C-terminal esters and two different N-terminal sequences would have added an additional, unnecessary layer of complexity.
Chapter 4

Based on these successful initial experiments, we examined the general applicability using other peptides. Fast and efficient side-chain handle conjugation was also observed for a variant of the phage display selected α-helical PMI peptide (TSFAEYWNLSSK-Cam-L-OH) [8,63] (Figure 12B), which is known to antagonize protein-protein interaction (PPI’s) between p53 and Hdm2 or HdmX. To prevent oligomerization, the PMI sequence was capped with an N-terminal Pro residue, which is not tolerated by any peptiligase-variant.[64] Additionally a C-terminal Lys was added to ensure good binding affinity with thymoligase. Both ligation and folding proceeded very cleanly and, like with the FITC labelled peptide, the only ligation byproduct resulted from hydrolysis of the excess of PMI peptide ester. Using this concept, it could be envisaged to combine two distinct bioactive peptides in one molecule, e.g. by grafting one bioactive sequence into the [T20K(DL)]kB1 cyclotide scaffold and conjugating the second modality to the ‘DL’ side-chain, respectively. However, the extension towards grafting studies, the combination of different bioactive peptides and biological evaluation of these constructs fell beyond the scope of this research. A similar concept is the fusion of two [T20K(DL)]kB1 cyclotides using a double-headed Cam-ester (DHCE; see chapter 2). Again, efficient conversion to the corresponding ligated constructs as well as oxidative folding was observed (see Figure 12C), as confirmed by mass spectrometry.
Efficient Enzymatic Cyclization of Disulfide-Rich Peptides using Peptiligases

(MW_{calc}: 7688.6 Da, MW_{exp.}: 7689.6 Da) and the expected HPLC retention time shift. To ensure efficient conversion to the di-substituted DHCE fused construct and avoiding the monosubstituted construct, a high concentration and an excess of c-[T20K(DL)]kB1 was required. Interestingly, the ligation worked equally efficient when oxidatively folded cf-[T20K(DL)]kB1 was used. However, since ligation and oxidative folding can be combined in a one-pot procedure thymoligase-catalyzed ligation followed by oxidative folding is preferred.

In general, this approach, combining two specific ligations employing two different peptiligase variants, gives access to cyclic peptide conjugates and to the preparation of peptides (or proteins) of higher order structure. Besides the formation homo-dimers using a DHCE, also the formation of hetero-dimers, similar to bispecific antibodies[^65], using a DHCE with two distinct reactive Cam-ester moieties can be envisioned. Additionally, even the addition of a third cyclotide or peptide in general (or multiple copies of it) could be possible, if, for example click-reactive handles such as homoazidoalanine or azidolysine are used within the enzymatic recognition sequence of the DHCE. These amino acids are well tolerated by omniligase-1 (see chapter 2 and 6) and can be used for chemical follow-up ligations such as strain-promoted alkyne-azide cycloaddition (SPAAC).[^66] Moreover, it could also be conceived to apply the DHCE concept to the N-terminal homo-fusion of two proteins.

RTD-1

In addition to cyclotides, the non-mechanically interlocked θ-defensins, which possess a range of antimicrobial activities,[^6] are also being pursued as pharmaceutically exciting scaffolds, onto which bioactive epitopes can be grafted. Rhesus θ-defensin-1 (RTD-1) was the first θ-defensin discovered in leukocyte extracts from Rhesus macaques[^67] and can be considered as a prototypic example of this class of ultra-stable molecules. To date, several approaches have been described for the chemical synthesis of RTD-1, including chemical ring closure using standard coupling reagents,[^67,68] bacterial expression[^32,45,46] and the use of native chemical ligation.[^69] Although the latter represents a powerful approach for the generation of large libraries of RTD-1 variants, it is not suitable for their production on industrial scale. To overcome this limitation, we further explored the possibility of using peptiligase variants for the chemo-enzymatic one-pot synthesis of RTD-1. Based on a sequence-guided evaluation, we identified two potential ligation sites located in the respective β-turn loops of the molecule: A (…LCRR // GV…) and B (…ICTR // GF…; Figure 13A).
Figure 13. Stick representation of RTD-1 (pdb 1HVZ). Disulfide bonds shown in yellow. Ligation sites (A/B) are given together with the individual sequence of each linear ester precursor peptide. B/C) HPLC traces of omniligase-1 catalyzed cyclization of oc-RTD-1 A ester and oc-RTD-B ester, followed by folding into their native forms cf-RTD-1 A and cf-RTD-1 B. D) Comparison of secondary chemical shifts of enzymatically cyclized RTD-1 B (blue) and native RTD-1 (grey).

Following SPPS of both linear precursors as elongated C-terminal Cam-esters, namely RTD-1 A-Cam-L-OH and RTD-1 B-Cam-L-OH, we evaluated the possibility of cyclizing them using omniligase-1. As expected, both RTD-1 A-ester and RTD-1 B-ester were rapidly and efficiently cyclized in less than 30 min (Figure 13B and C). Since ligations can also be carried out starting from crude, non-lyophilized linear peptide, as demonstrated for [T20K]kB1 and MCoTI-II, the initial HPLC purification can be avoided if suitable crude qualities of the linear precursor are obtained. To ensure efficient cyclization in initial studies, reducing conditions were applied by the addition of 3.5 mM (tris(2-carboxyethyl)phosphine) (TCEP). Interestingly, we observed that RTD-1 has a very high propensity to rapidly adopt its native, oxidized fold. This led to the surprising discovery that after completion of the cyclization, the excess of TCEP present is entirely consumed and hence cannot maintain a reducing environment. Thus, as confirmed by LC-MS, native cf-RTD-1 is obtained after a reaction time of only 2-6 hours. In contrast to these observations, rapid oxidative folding of RTD-1 A/B ester was observed under non-reducing cyclization conditions; but no enzymatic ligation occurred, only fast hydrolysis of the C-terminal Cam-ester moiety. Identical observations were made when cyclization of oxidized, linear RTD-1 A/B was attempted. We expect that the excess of TCEP required for efficient cyclization can be lowered and delicately adjusted to the explicit batch of linear precursor used in order to obtain native RTD-1 in an even shorter time frame. The native folding of RTD-1 B was confirmed.
via one-dimensional $^1$H-NMR. The Hα chemical shifts of RTD-1 B were essentially identical to the chemical shifts of chemically synthesized RTD-1 reported by Craik et al (Figure 13D),[68] confirming that cf-RTD-1 B adopted a native θ-defensin fold. cf-RTD-1 A was unequivocally identified via LC-MS and both cf-RTD-1 A and cf-RTD-1 B showed identical retention times (Rt) and mass spectra (Figure 13B/C). In summary, we developed a new approach for the efficient synthesis of the θ-defensin RTD-1, which is characterized by its simplicity and high flexibility. The linear ester precursors can easily be obtained using Fmoc-SPPS, and the two ligation sites offer high flexibility for the generation of grafted RTD-1 variants, similarly to the flexible system for MCoTI-II grafting.

**Conclusion**

The work described demonstrates the applicability of CEPS using peptiligase variants, e.g. omniligase-1, for the cyclization of diverse classes of native disulfide-rich peptides without the need of adding an enzyme recognition sequence. In addition, the identification of multiple ligation sites per molecule together with the possibility to cyclize pre-folded disulfide-rich molecules using an increasingly stable HMBA-ester significantly increases the synthetic flexibility of this enzymatic approach. We have achieved the efficient chemo-enzymatic synthesis of the θ-defensin RTD-1 as well as the cyclotides MCoTI-II, kalata B1 and [T20K]kB1 in combination with concomitant folding in a one-pot procedure using two different ligation sites per molecule. This chemo-enzymatic strategy using peptiligase variants can be scaled to the multi-gram range and therefore represents a promising methodology for larger scale manufacture of this class of molecules, which are usually not easily accessible at large quantities. Compared to other previously reported methodologies, e.g. using sortase A, head-to-tail cyclization mediated by peptiligase variants such as omniligase-1 represents a valuable alternative and offers a powerful extension to the currently existing toolbox of enzymes for cyclization reactions. Our findings provide proof of concept for the potential broad applicability of enzymatic macrocyclization for the efficient preparation of disulfide-rich peptides. In addition, as exemplified for [T20K]kB1 the combination of two distinct peptiligase variants in a dual enzymatic approach enables the preparation of modified cyclotides or dimeric fusion variants thereof, potentially representing a valuable approach in academic research.
Expression and Purification of Peptiligase Variants (Omniligase-1, Peptiligase #1015)

Production and purification of the enzyme variants was carried out as described previously (see chapter 2). [55]

Solid Phase Peptide Synthesis

Peptides were assembled on Fmoc-Leu-Wang resin using standard Fmoc-SPPS procedures. Peptides containing glycolate-type of esters were introduced using Fmoc-glycolic acid as described earlier by Nuijens et al. [51] Peptides containing HMBA-type of esters were prepared using unprotected HMBA. HMBA (4 eq.) was coupled to a preloaded Leu-Wang resin using diisopropylcarbodiimide (DIC, 4 eq.) and OxymaPure (4 eq.) for 2x 20 min. Esterification was performed for 2x 40 min using Fmoc-amino acid (4 eq.), dimethylaminopyridine (0.1 eq.) and DIC (4 eq.). Esterification during scale-up studies of [T20K]kB1-L6 were performed using bromo-HMBA (bromomethyl benzoic acid) instead of HMBA. Differing from previous protocols, esterification was performed using Fmoc-AA-OH (4 eq.), OxymaPure (4 eq.), disopropylethylamine (DIPEA, 4 eq.) and potassium iodide (0.1 eq.) for 16 hours at 50°C. [70]

In general, for peptide synthesis Fmoc-Leu-Wang LL resin (0.31 mmol/ g) was used as a solid support. Coupling of amino acids was standardly carried out using Fmoc-amino acid (4 eq.), DIC (4 eq.) and OxymaPure (4 eq.) for 40 min in dimethylformamide (DMF). Fmoc-deprotection was performed for two times 10 min with 20% (v/v) piperidine in DMF. Final cleavage from the resin and sidechain deprotection was performed using an 87.5/5/2.5/2.5/2.5 (v/v/v/v/v) mixture of trifluoroacetic acid (TFA)/triisopropylsilane (TIS)/water/thioanisole/dithiothreitol (DTT) for 2 h. The crude peptide was precipitated using ice-cold (-20°C) diisopropylether, followed by drying under reduced pressure. Peptides were analyzed by HPLC-MS and if deemed necessary, purified using preparative HPLC followed by lyophilization. Synthesis was either carried out manually or using a CS Bio CS 136 peptide synthesizer (CS Bio Co., Menlo Park, CA, USA).

If not otherwise denoted, standard protocols were used (e.g. for the synthesis of the PMI peptide and RTD-1 variants). Differing from the protocols described above, the following protocols were used for the synthesis of particular peptide variants:
Efficient Enzymatic Cyclization of Disulfide-Rich Peptides using Peptiligases

**kB1 variants:** Fmoc-GG-OH and Fmoc-PG-OH (4 eq. each) were incorporated as dipeptides.

**MCoTI-II variants:** Fmoc-Gly-Ser(Psi(Me,Me)Pro)-OH, Fmoc-Asp(tBu)-Ser(Psi(Me,Me)Pro)-OH and Fmoc-Asp(OtBu)-(Dmb)Gly-OH were used during the synthesis.

**Double-headed Cam-ester:** The double-headed Cam-ester was prepared as described in chapter 2.

**[T20K(DL)]kB1-OHMBA-L-OH:** The side-chain handle derivatized kB1 was synthesized using Fmoc-Lys(Mtt)-OH (Mtt= methyltrityl) as building block. Mtt deprotection was performed by adding a mixture of 6.5/2/1/0.5 (v/v/v/v) DCM/HFIP/TFE/TIS (10 mL/ 500 mg resin) to dry resin (DCM= dichloromethane, HFIP= hexafluoroisopropanol, TFE= trifluoroethanol, TIS= triisopropylsilane). The mixture was gently stirred for 1 h, before it was washed with DCM (3x 2 min) and DMF (3x 2 min). Next, the following amino acid (Fmoc-Leu-OH, 4 eq.) was coupled to the ε-NH₂ of lysine using 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU; 4 eq.), OxymaPure (4 eq.) and DIPEA (8 eq.) for 60 min. Following Fmoc-deprotection using 20% (v/v) piperidine in DMF (2x 8 min). Fmoc-Asp(tBu)-OH was coupled identically to Fmoc-Leu-OH. Resin cleavage and work-up was performed as described above.

**Ac-K(FITC)KFSKK-Cam-L-OH:** The final amino acid (Lys) was coupled using Fmoc-Lys(Mtt)-OH as a building blocks, followed by Fmoc deprotection using 20% (v/v) piperidine in DMF (2x 8 min) and acetylation of the N-terminus with a mixture of 0.5/0.5/9 (v/v/v) acetic anhydride/DIPEA/DMF. After Mtt deprotection (described above) fluorescein isothiocyanate (FITC) isomer 1 (2 eq.) and 8 eq. DIPEA were coupled overnight at room temperature. After completion of the synthesis, resin cleavage was performed for 2 h using a 92.5/5/2.5 (v/v/v) mixture of TFA/TIS/water. After precipitation in ice-cold diisopropylether the crude peptide was dried in vacuo.

Alternatively to the protocols described, [T20K]kB1, MCoTI-II and RTD-1 were also synthesized and head-to-tail cyclized using a previously described method. Briefly, the peptide precursors were assembled on 2-chlorotrityl resin using Fmoc-SPPS on an automated peptide synthesizer and cleaved with 1% TFA in DCM. The side-chain-protected linear precursors were then cyclized in solution via standard amide bond formation. The protected cyclic peptides were isolated using preparative HPLC, followed by side-chain deprotection and disulfide bond formation. The cyclic and oxidized peptides were subsequently purified using HPLC on a semi-preparative C18 column to > 95% purity.

**Peptide Purification**

Crude peptide products were purified using preparative reversed-phase HPLC. Preparative HPLC was carried out on a Varian PrepStar solvent system configured with a Varian ProStar 330 PDA detector (Varian, Palo Alto, CA, USA).
Chapter 4

The purification process was monitored at a wavelength of $\lambda = 220$ nm. Products were purified using a Reprosil Gold C18 column (10 $\mu$m, 250x 50 mm, Dr. Maisch, Ammerbuch, Germany) eluting with a water acetonitrile (ACN) gradient, with a flow rate of 45 mL/min. As mobile phase a binary mixture of A (water + 0.05% (v/v) TFA) and B (ACN + 0.05% (v/v) TFA) was used by default. The gradient was chosen individually for each peptide. Samples were prepared by dissolving the crude peptide in ACN/water and filtration through a 0.22 $\mu$m filter unit.

Product containing fractions were pooled according to their purity, frozen in liquid N2 and lyophilized (-90°C). Products were stored at -20°C for further use.

Peptide Characterization

HPLC-MS:

The peptide purity was assessed using an Agilent 1260 Infinity HPLC system coupled with an Agilent 6130 quadrupole mass spectrometer (Agilent, Santa Clara, CA, USA) to determine the peptide mass. Separation was performed using a Waters XSelect® CSH C18 column (2.5 $\mu$m, 3.0 x 150 mm, Waters Corporation, Milford, MA, USA) or a Phenomenex Luna RP-C18 (10 $\mu$m, 4.6 x 250 mm, Phenomenex, Torrance, CA, USA) column, eluting with 0.05% (v/v) MSA in a water/ACN gradient, with a flow rate of 1 mL/min and a column temperature of 40°C. As mobile phase a binary mixture of solvent A (water + 0.05% (v/v) MSA) and solvent B (ACN + 0.05% (v/v) MSA) was used. A linear gradient from 5-60 % B in 7.5 min, followed by isocratic 95% solvent B for 3 min was used by default. For optimal results an appropriate gradient was chosen for each sample individually (see Table S 1). The purity of peptides was determined by automatically integrating product and impurity peaks of the relevant HPLC spectrum ($\lambda = 220$ nm).

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCoTI-II</td>
<td>15-40% ACN/ MQ-water (0.05% (v/v) MSA) in 7.5 min</td>
</tr>
<tr>
<td>Kalata B1 variants</td>
<td>25-45% ACN/ MQ-water (0.05% (v/v) MSA) in 7.5 min</td>
</tr>
<tr>
<td>RTD-1 variants</td>
<td>5/10 -60% ACN/ MQ-water (0.05% (v/v) MSA) in 7.5 min</td>
</tr>
</tbody>
</table>

Structural Characterization using Nuclear Magnetic Resonance (NMR):

Generally, peptides were dissolved in H$_2$O/D$_2$O (9:1, v/v) at a concentration of 1-2 mg/mL. The one- and two-dimensional spectra (TOCSY and NOESY) were acquired on a Bruker Avance-600 MHz NMR
spectrometer equipped with a cryoprobe and the mixing time for the 2D spectra was 80 and 200 ms at 298 K, respectively. All the spectra were referenced to 2,2-dimethyl-2-silapentane-5-sulfonic acid at 0 ppm and analyzed using Sparky.

**Figure 14.** The 1D 1H NMR spectrum of (oc-f)-MCoTI-II-L1-HMBA-L-OH. The well-dispersed peaks observed within the amide proton region (5-9 ppm) indicated that this peptide has a well-defined structure.

### Enzymatic Peptide Cyclization

All ligations were carried out at room temperature (20°C).

**MCoTI-II**

MCoTI-II-L1:

oc-MCoTI-II-L1-Cam-L-OH was dissolved in potassium phosphate buffer (1 M, pH 8.5) to the desired concentration (0.5- 4 mg/mL; 0.1- 0.8 mM) and after checking the pH (>8.0) omniligase-I (0.5 µM) was added. After completion of the ligation, GSH was added to a concentration of 5 mM and the reaction mixture was left stirring over-night to obtain native (folded) MCoTI-II.

MCoTI-II-L5:

For one-pot cyclization and folding oc-MCoTI-II-L5 HMBA-ester was dissolved in tricine buffer (100 mM, pH 8.5) containing 5 mM GSH to a concentration of 0.5 mg/mL (0.1 mM) and peptiligase #1015 (final concentration: 1.8 µM) was immediately added to start the reaction. After completion of the ligation the reaction mixture was left stirring over-night to obtain native (folded) MCoTI-II.
Kalata B1

kB1-L2:

kB1-L2 was dissolved at 0.3 mM in potassium phosphate buffer (1 M, pH 8.5) containing 3.5 mM TCEP. Omniligase-1 (0.5 µM) was added to initiate the reaction. In case of kB1-L6, 100 mM tricine pH 8.0 was used as a buffer and kB1-L6 was dissolved at 0.15 mM, while omniligase-1 was added to a final concentration of 1.8 µM.

[T20K]kB1-L6:

[T20K]kB1-L6-HMBA-L-OH (500 mg) was dissolved in 500 mL tricine buffer (100 mM, pH 8.0). Omniligase-1 (25 mg, 1.8 µM) was added to initiate the reaction. Additional 500 mg of crude, linear [T20K]kB1-L6-HMBA-ester was added every 30 min, up to a concentration of 6 mg/mL (1.7 mM). For one-pot cyclization-folding the reaction mixture 2-propanol was added to a concentration of 50% (v/v) supplemented with 2 mM GSH/ GSSG. The mixture was let stand over-night for oxidative folding to obtain (cf)-[T20K]kB1.

[T20K(DL)]kB1-L6:

oc-[T20K(DL)]kB1-L6-HMBA-L-OH (0.15 mM) was dissolved in 100 mM Tricine buffer pH 8.0 containing 3.5 mM TCEP and omniligase-1 was added to a concentration of 0.5 µM. The pH was ensured to be >8.0. The reaction was allowed to proceed at room temperature and was followed via HPLC-MS by direct injection of the reaction mixture. After completion of the cyclization (45 min) 2-propanol was added to a concentration of 50% (v/v) and the mixture was supplemented with 2 mM GSH/ GSSG (in this case no TCEP was added to the reaction at the beginning). After completion of the oxidative folding (over-night) the final product was purified via preparative HPLC.

RTD-1

RTD-1 A and B:

Both RTD-1 A and RTD-1 B were dissolved in tricine buffer (200 mM, pH 8.5) containing 3.5 mM TCEP to a final concentration of 0.5 mg/mL (0.17 mM). Omniligase-1 was added to a concentration of 0.7 µM. To obtain cyclic, oxidized RTD-1 the reaction was monitored till reduced RTD-1 was fully converted into its oxidized counterpart, followed by RP-HPLC purification.
Side-Chain Handle Conjugation to [T20K(DL)]kB1-L6

**Double-Headed Cam-ester:**

0.05 mg of DHCE (HO-L-Cam-KKSFDC-GDFSKK-Cam-L-OH, 28 nmol) and 0.5 mg of c-[T20K(DL)]kB1 (150 nmol, 5 eq.) were dissolved in 25 µL water followed by the addition of 25 µL 1 M aq. potassium phosphate buffer pH 8.5. To initiate the reaction, 10 µg (0.35 nmol) thymoligase was added. The reaction was allowed to proceed at room temperature. The reaction progress was followed via HPLC-MS by direct injection of the reaction mixture. After completion of the ligation reaction, 2-propanol was added to a concentration of 50% (v/v) and the mixture was supplemented with 2 mM GSH/GSSG. The reaction mixture was incubated over-night while continuously following the reaction by HPLC-MS.

**Coupling of the PMI-Peptide:**

0.2 mg (60 nmol) c-[T20K(DL)]kB1 was dissolved in 50 µL 200 mM aq. tricine pH 8.5 together with 0.6 mg PMI peptide ester (300 nmol, 5 eq.; H-PTSFAEYWNLLSK-Cam-L-OH). 10 µg thymoligase (0.35 nmol) was added to initiate the reaction, which was let proceed at room temperature. The reaction progress was followed via HPLC-MS by direct injection of the reaction mixture. After completion of the ligation, 2-propanol was added to a concentration of 50% (v/v) and the mixture was supplemented with 2 mM GSH/GSSG. The reaction mixture was incubated over-night while continuously following the reaction by HPLC-MS.

**Coupling of a Fluorescently Labelled Peptide:**

0.1 mg of c-[T20K(DL)]kB1 (30 nmol) was dissolved in 50 µL 100 mM tricine buffer pH 8.5 containing 3.5 mM TCEP to a concentration of 0.6 mM. 0.35 mg of Ac-K(FITC)KFSKK-Cam-L-OH (210 nmol, 7 eq.) was added together with 5 µg of thymoligase (0.18 nmol). The reaction was let proceed at room temperature and was followed by HPLC-MS. In case oxidative folding of the c-[T20K(DL-FITC)]kB1 was desired, TCEP was omitted from the reaction mixture. Instead, after completion of the ligation, 2-propanol was added to a concentration of 50% (v/v) and the mixture was supplemented with 2 mM GSH/GSSG. The reaction mixture was incubated over-night while continuously following the reaction by HPLC-MS.

**Edman Degradation of c-[T20L(DL)]kB1**

3 mg of c-KB1 T20K(DL) was dissolved in 500 µL deionized water and 500 µL pyridine. To this solution 25 µL phenylisothiocyanate (PITC) was added and the mixture was heated to 50°C for 20 min. Afterwards, the mixture was extracted with 6 mL of a mixture of 10:1 (v/v) heptane/ethyl acetate followed by centrifugation at 2000 x g. The aqueous layer (200 mL) was diluted with 200 µL TFA to cleave the N-terminal thiourea amino acids and was subsequently analyzed via HPLC-MS.
Bioactivity Assays

Cytotoxicity Assay:

The cytotoxicity of [T20K]kB1 against HeLa cells was evaluated using a resazurin-based cell viability assay as described previously. Briefly, HeLa cells were cultured in Dulbecco’s Modified Eagle Medium supplemented with 10% fetal bovine serum and Penicillin-Streptomycin (50 U/mL), in 175 cm² tissue culture flasks until 80% confluent. Cells were trypsinized and seeded in 96-well flat-bottom microplates (5 × 10³ cells/well) the day before the assay. A serial dilution of enzymatically or chemically cyclized [T20K]kB1 was prepared in Dulbecco’s Modified Eagle Medium (serum-free) and the peptide solutions were incubated with cells in triplicate for 2 h. After 2 h, peptide solutions were then removed from the plates, replaced with 0.005% (w/v) resazurin solution in fresh medium, and incubated with cells for 22 h at 37°C. Sterile H₂O and 0.01% (v/v) Triton X-100 were included as controls, representing 100% and 0% cell viability, respectively. The percentage of inhibition was calculated using the Triton X-treated cells as 100% inhibition.

![HeLa cytotoxicity](image)

*Figure 15, Inhibition curves of [T20K]kB1 (chem: chemically synthesized; enzy: chemo-enzymatically synthesized) against HeLa cells. The percentage of inhibition was obtained in triplicate and plotted as mean ± SEM using GraphPad Prism 7.*

Trypsin Inhibition Assay:

The inhibition constant (Kᵢ) of MCoTI-II against trypsin was evaluated as described previously. A serial dilution of both MCoTI-II (either cyclized enzymatically or chemically) in assay buffer (100 mM Tris-HCL, 100 mM NaCl, 10 mM CaCl₂, and 0.005% Triton X-100, pH 8) was incubated with 0.25 nM trypsin in non-binding surface 96-well plates (Corning) for 3 h at 25°C before the addition of 100 μM substrate (Boc-QAR-MCA; MCA = 4-Methyl-Coumaryl-7-Amide). The cleavage of substrate was monitored using a Tecan M1000...
microplate reader every 30 s. The $K_i$ value was calculated using Morrison equation in GraphPad Prism 7. The assay was conducted in triplicate.

**Figure 16.** Inhibition curves of MCoTI-II (chem: chemically synthesized; enzy: chemo-enzymatically synthesized) against trypsin. The percentage of activity was obtained from competitive inhibition assay in triplicate and plotted as mean ± SEM using GraphPad Prism 7.

**Computational Methods**

The numbering of MCoTI-II amino acids was adopted from [https://www.uniprot.org/](https://www.uniprot.org/):

![Superimposed Subtilisins (pdb):](image)

**Without ligands:** 1A2Q, 1AF4, 1AK9, 1AQN, 1AU9, 1AV7, 1AVT, 1BE6, 1BE8, 1BFK, 1BFU, 1C3L, 1DUI, 1GNV, 1GNS, 1S01, 1S02, 1SBC, 1SBH, 1SBI, 1SBT, 1SCA, 1SCB, 1SCD, 1SCN, 1SEL, 1ST2, 1SUB, 1SUC, 1SUD, 1SUE, 1SUP, 1UBN, 1VS, 1YJA, 1YJB, 1YJC, 2SBT, 2ST1, 2WUV, 2WUW, 3F49, 3UNX, 3VS, 4C3U, 4C3V,

**With ligands:** 1CSE, 1LW6, 1OYV, 1RO0, 1SBN, 1SIB, 1SPB, 1SUA, 1TM1, 1TM3, 1TM4, 1TM5, 1TM7, 1TMG, 1TO1, 1TO2, 1V5I, 1Y1K, 1Y33, 1Y34, 1Y3B, 1Y3C, 1Y3D, 1Y3F, 1Y4A, 1Y4D, 1Y6, 2SEC, 2SIC, 2SN, 3BGO, 3CNQ, 3CO0, 5OX2, 55IC.
System Preparation:

The structure of thymoligase (pdb: 5OX2)\(^{(72)}\) was used as template to obtain the structure of omniligase-1. We designed omniligase-1 with the Rosetta protocol PackRotamers\(^{(73,74)}\) and the following mutations were introduced: I107V, N156S, D166S, R217H and G222P.

To generate the initial guesses for the binding of MCoTI-II in omniligase-1 we used several thymoligase conformations. Besides the X-ray conformation found in pdb 5OX2\(^{(72)}\), we also used 32 more conformations sampled during a previous 100 ns unbiased MD simulation of thymoligase with its product thymosin-\(\alpha\)\(_1\). We selected these 32 conformations because they showed thymosin-\(\alpha\)\(_1\) bound in a reactive conformation, namely the carbonyl oxygen of amino acid P1 was within hydrogen-bond distance of the backbone of Cys221 and the side-chain of Asn155 (the oxyanion hole), the carbonyl carbon of amino acid P1 was within 3.5 Å of the sulfur atom from the nucleophilic thiol of Cys221 and the two hydrogen bonds between the catalytic triad Cys221, His64 and Asp32 were all present.

On these enzyme conformations we superimposed 30 NMR conformations of MCoTI-II from PDB 1HA9\(^{(50)}\) using the backbone of P1 to P2' amino acids and we selected as initial structures the pairs of omniligase-1/\(\text{cf-MCoTI-II}\) that presented the least steric clashes between the enzyme and the product.

Subsequently, the best four initial structures were subjected to an annealing protocol to refine the binding mode prediction. To prepare for the annealing protocol we first minimized and equilibrated the system during 100 ps at 300 K under constant number, volume and temperature (NVT) and 1 ns at 300K under constant pressure (NPT) conditions, while using harmonic constraints to keep the position of the omniligase-1/MCoTI-II atoms fixed. Next, we performed a 24 ns simulated annealing protocol as follows: first the system was heated from 300 K to 423 K during 2 ns and then cooled back to 300 K during 6 ns. This cycle was repeated for 3 times. During the entire annealing protocol, a harmonic potential with the force of 1000 kJ\(^{-1}\) mol\(^{-1}\) Å\(^2\) was applied to constraint the distance between the P1 carbonyl oxygen and the backbone of Cys221. From the final 3 ns of simulation in the annealing protocol we selected the conformation were the binding mode of \(\text{cf-MCoTI-II}\) was closer to a reactive conformation.

Finally, to assess the stability of these conformations, we relaxed the system for 20 ns at 300 K without any restraints or biasing potential and we searched for reactive conformation during the last 10 ns of simulation.

MD Details:

The MD simulations were conducted with GROMACS 5.0.7\(^{(75)}\) using the AMBER force field ff14SB\(^{(76)}\). The initial structures were placed in a dodecahedral box with a minimum distance between the solute and the box of 1.2 nm, then solvated with TIP3P water\(^{(77)}\) and counterions were added.
The temperature of the system was controlled via the V-rescale temperature coupling\textsuperscript{(78)} with a relaxation time of 0.1 ps in the NVT ensembles, and via the Nosé-Hoover thermostat\textsuperscript{(79,80)} with a relaxation time of 0.5 ps in the NPT ensembles. Isotropic pressure coupling was applied with a reference pressure of 1 bar and a relaxation time of 2.0 ps using the Berendsen barostat.\textsuperscript{(81)} The integration time step in all cases was 2 fs. Non-bonded interactions were calculated up to 1.2 nm using a neighbor list that was updated every 5 integration steps. Long-range electrostatic interactions were included with the Particle Mesh Ewald method.\textsuperscript{(82)}
References


Efficient Enzymatic Cyclization of Disulfide-Rich Peptides using Peptiligases


