University of Groningen

Engineering a Diverse Ligase Toolbox for Peptide Segment Condensation

Nuijens, Timo; Toplak, Ana; Quaedflieg, Peter J. L. M.; Drenth, Jeroen; Wu, Bian; Janssen, Dick B.

Published in:
Advanced Synthesis & Catalysis

DOI:
[10.1002/adsc.201600774](https://doi.org/10.1002/adsc.201600774)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Nuijens, T., Toplak, A., Quaedflieg, P. J. L. M., Drenth, J., Wu, B., & Janssen, D. B. (2016). Engineering a Diverse Ligase Toolbox for Peptide Segment Condensation. *Advanced Synthesis & Catalysis*, 358(24), 4041-4048. <https://doi.org/10.1002/adsc.201600774>

Copyright

Other than for strictly personal use, 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), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Engineering a Diverse Ligase Toolbox for Peptide Segment Condensation

Timo Nuijens,^{+a} Ana Toplak,^{+a,b} Peter J. L. M Quaedflieg,^a Jeroen Drenth,^b Bian Wu,^{b,c,*} and Dick B. Janssen^{b,*}

^a EnzyPep B.V., Brightlands Campus, Urmonderbaan 22, 6167 RD Geleen, The Netherlands

^b Biotransformation and Biocatalysis, Groningen Biomolecular Science and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Fax: (+31)-50-363-4165; phone: (+31)-50-376-34008; e-mail: d.b.janssen@rug.nl

^c Present address: CAS Key Laboratory of Microbial Physiology and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, China, and State Key Laboratory of Transduction Technology, Chinese Academy of Sciences, China

Fax: (+86)-10-6480-7468; phone: (+86)-10-6480-6035; e-mail: wub@im.ac.cn

⁺ These authors contributed equally.

Received: July 19, 2016; Revised: October 2, 2016; Published online: December 8, 2016

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/adsc.201600774>.

Abstract: The substrate profile of peptidylase, a stable enzyme designed for peptide ligation in aqueous environments, was mapped using six different peptide libraries. The most discriminating substrate binding pocket proved to be the first nucleophile binding subsite (S1'), which is crucial for the peptide ligation yield. Two important amino acids shaping the S1' pocket are M213 and L208. A site-saturation library of the M213 position yielded two variants with a significantly broadened substrate profile, i.e., M213G and M213P. Next, examination of two libraries with M213G+L208X and M213P+L208X (with X being any proteinogenic amino acid) resulted in a toolbox of enzymes which can accommodate any proteinogenic amino acid in the S1' pocket, except proline. The applicability of a particu-

lar enzyme variant in chemoenzymatic peptide synthesis was demonstrated by coupling at the gram scale of two peptide segments to yield exenatide, a 39-mer therapeutic peptide used in the treatment of diabetes type II. The overall yield of 43% is at least 2-fold higher than yields reported for conventional syntheses of exenatide by full solid-phase peptide synthesis; large-scale production costs are expected to be significantly reduced if the enzymatic coupling process is employed to manufacture this peptide.

Keywords: chemoenzymatic peptide synthesis; exenatide; genetic engineering; peptide ligase; segment condensation

Introduction

Application of peptides (and proteins) as therapeutic agents or prodrugs,^[1,2] nutritional additives or cosmetic ingredients is gaining importance.^[3] By 2018 the market for peptide therapeutics (i.e., with 5–50 amino acids) is expected to reach over \$25 billion, mainly driven by their use in the treatment of cardiovascular and metabolic diseases.^[4] With many peptides in (pre)clinical development, the prospects for new applications look bright. Not only are the numbers of therapeutic peptides in development increasing but also their complexity and quantity requirements.

Whereas in the past many therapeutic peptides contained less than 10 amino acids, nowadays most peptides entering the market are over 25 amino acids long.^[5] Therapeutic peptides containing unnatural amino acids, multiple disulfide bridges and cyclic structures are becoming more common.

The solution- and solid-phase chemical synthesis of medium to long peptides (containing 20–50 amino acids) is time-consuming, often with modest yields and high production costs. Currently, most peptides are synthesized by solid-phase peptide synthesis (SPPS). However, during SPPS side-products accumulate and purification requires tedious preparative

HPLC procedures to separate the target peptide from a complex mixture of by-products. In general, for medium to long peptides the overall yield of SPPS is below 25% and two or sometimes more preparative HPLC runs are needed to obtain the product with sufficient purity.^[6]

Higher yields can be obtained with a convergent synthetic strategy involving chemical condensation of peptide segments in solution.^[7] However, this is often not feasible due to racemization of the C-terminal amino acid of the acyl segment. As a result, chemical segment condensation is only applicable if Gly or Pro residues are present at strategic positions within the peptide sequence, which is mostly not the case. Furthermore, all side-chain functionalities must be protected, resulting in very hydrophobic, poorly soluble peptide segments which are difficult to purify.

On the other hand, by using enzymatic peptide segment condensation in aqueous solution, racemization can be totally avoided and side-chain protecting groups are not required. To achieve ligation, a peptide activated as a C-terminal ester (acyl donor) is coupled to a peptide segment with a free nucleophilic N-terminus (acyl acceptor) using a serine protease for catalysis.^[8] Such enzymatic ligations only proceed efficiently under kinetic control, and hydrolytic side-reactions (i.e., hydrolysis of the ester functionality and/or of the peptide backbone of the segments) often cannot be avoided. However, by replacing the active site serine of a protease by a cysteine the “synthesis to hydrolysis” ratio (S/H ratio) can be dramatically improved.^[9] The acyl-enzyme thioester which is formed as a covalent intermediate in the thiol variant is more prone to nucleophilic attack by the amine nucleophile of a peptide segment (synthesis) than by water (hydrolysis). Recently we reported the discovery of peptiligase (Ptl), a calcium-independent subtilisin BPN’ variant that catalyzes peptide segment condensations in aqueous environments both in high yields and with superior S/H ratios.^[11]

Peptiligase could hold the key to the synthesis of long peptides *via* a convergent segment condensation approach, combining classical SPPS for the synthesis of relatively short peptide segments (<20 amino acids) and biocatalysis for the subsequent segment ligation in aqueous medium. In this publication, we report variants of peptiligase with significantly broader substrate profiles and with even higher S/H ratios. Using one of these improved peptiligase variants, we demonstrate the convergent synthesis of the therapeutic drug exenatide from two segments of similar length in excellent overall yield and purity.

Results and Discussion

Substrate Profile of Peptiligase

Preferably, a ligase for peptide segment condensation should have a broad substrate profile. However, most proteases found in nature accept a limited range of substrates, as has been shown for subtilisins.^[13] Peptiligase (Ptl) has six substrate recognition subsites; four involved in acyl donor binding (S1–S4), and two forming the binding region for the nucleophilic peptide (S1’ and S2’).^[10] Subsites S1, S4, S1’ and S2’ are clearly defined pockets in the active site groove.

To determine the substrate profile of each binding pocket of peptiligase and its variants, we performed an extensive series of reactions with six peptide libraries synthesized by SPPS. The libraries 1–4 listed below served to map the acyl donor sites, whereas the libraries 5–6 mapped the nucleophilic peptide sites:

- 1: Ac-Asp-Phe-Ser-Lys-Xxx-OCam-Leu-OH
- 2: Ac-Asp-Phe-Ser-Xxx-Leu-OCam-Leu-OH
- 3: Ac-Asp-Phe-Xxx-Lys-Leu-OCam-Leu-OH
- 4: Ac-Asp-Xxx-Ser-Lys-Leu-OCam-Leu-OH
- 5: H-Xxx-Leu-Arg-NH₂
- 6: H-Ala-Xxx-Arg-NH₂

where Xxx represents any proteinogenic amino acid except cysteine and OCam-Leu-OH is the carboxamidomethyl (glycolamide) ester coupled to leucine (obtained by standard SPPS as reported previously).^[14] The -Leu-OH residue was added to the Cam ester to improve enzyme recognition and minimize background chemical hydrolysis. Libraries 1–4 were used in enzymatic coupling reactions with the standard nucleophilic peptide H-Ala-Leu-Arg-NH₂ and libraries 5–6 were used in combination with the standard activated acyl donor Ac-Asp-Phe-Ser-Lys-Leu-OCam-Leu-OH (Figure 1).

For mapping the substrate profile of peptiligase, the acyl donor and acceptor substrates were deliberately used at low concentrations (10 and 15 mM, respectively), resulting in rather low S/H ratios with profound differences between the various substrates. The P1–P4 side chains (Figure 1) bind into the S1–S4 pockets, influencing the acylation reaction that leads to acyl-enzyme formation. The specificity of the S1 pocket appeared to be relatively broad, but reactions with tryptophan, aspartic acid and glycine were slow and no reaction at all was observed with proline. The S2 pocket mainly preferred positively charged amino acids and reaction rates were lower when aspartic or glutamic acid were present at P2. At the P3 position, all amino acids were well accepted, in agreement with the S3 pocket being poorly defined. On the contrary, the S4 pocket, which is furthest away from the coupling site, appeared to be very discriminating and

Figure 1. Peptiligase substrate profile of the acyl donor binding pockets (S1–S4). The black bars indicate conversion of acyl donor to synthetic product (yield) as determined by LC-MS; the grey bars represent conversion to hydrolytic product.

only accepted hydrophobic or slightly polar amino acids, in agreement with literature data on subtilisin BPN'.^[15] Although reaction rates differed substantially, in general all P1–P4 substrates were accepted as long as the P4 residue was hydrophobic or only slightly polar, and no proline was present at the P1 position (Figure 1).

Figure 2. Peptiligase substrate profile of the amine nucleophile binding pockets (S1', S2'). The black bars indicate the yield of synthetic product as determined by LC-MS; the grey bars represent conversion to the hydrolytic product.

Once the acyl-enzyme complex is formed, the binding and activation of the nucleophilic peptide accepted in the S1' and S2' pockets will determine the yield of synthetic product (Figure 2). If the amine nucleophile is poorly accepted, more thioester intermediate will be hydrolyzed and the S/H ratio will decrease. From our screening, the S1' pocket appeared to be the most discriminating one, as also observed with subtiligase.^[10] Only the small amino acids serine, glycine and alanine were well accepted. The ligation yield for all other amino acids was <25%. The S2' pocket showed a pronounced preference for hydrophobic amino acids. Peptides with residues at P2' offering electrostatic binding (Asp, Glu, Lys, Arg) or no interactions (Gly) were poorly accepted as nucleophiles. Proline was neither accepted at P1' nor at P2'.

For a generally applicable enzymatic peptide ligation technology, a peptiligase variant with a broader P1' substrate profile is highly desired. Therefore, we investigated the possibility to widen the substrate profile of the S1' pocket by mutagenesis.

Engineering the S1' Pocket

To identify mutations that might broaden the nucleophile acceptance of peptiligase, the S1' pocket of the peptiligase model was inspected for interactions with the P1' residue of a docked eglin inhibitor. The structures suggest that the P1' aspartic acid residue of the docked substrate has prominent interactions with resi-

Figure 3. Peptiligase model with the protease inhibitor eglin bound to the active site. The Pro-Val-Thr-Arg-Asp segment of eglin is shown. The residues of the S1' binding pocket M213 and L208 targeted for mutagenesis are highlighted in cyan. The hybrid model was constructed with Yasara using 1GNV and 1SBN^[8] as templates.

dues M213, L208 and G210, which form the S1' pocket^[16] and may play a discriminating role in P1' recognition (Figure 3).

First, we constructed a peptiligase M213X site-saturation library (with X representing all proteinogenic amino acids) by site-directed mutagenesis. The M213 position proved to be very sensitive to changes and out of the 19 mutants only 6, i.e., M213A, M213E, M213G, M213N, M213P, and M213Q, were expressed in a satisfactory quantity. These six variants were produced at a larger scale and after isolation and purification their S1' substrate profiles were mapped using Ac-Asp-Phe-Ser-Lys-Leu-OCam-Leu-OH as acyl donor and 20 H-Xxx-Leu-Arg-NH₂ peptides as acceptors. Replacing M213 with polar or acidic residues (E, N, Q) narrowed the S1' substrate profile, but replacing M213 with smaller amino acids (A, G) or proline broadened the substrate profile, with M213G and M213P being the optimal variants (Figure 4).

With the original peptiligase only four amino acids at P1' provided good substrates with a >30% cou-

Figure 4. P1' substrate profile of peptiligase (Ptl) and its M213G and M213P mutants. Synthesis refers to conversion of acyl donor to synthetic product, which was determined by LC-MS. In all cases, the Cam-ester starting material was completely consumed and the difference between synthesis and 100% reflects conversion to hydrolytic product.

pling yield under the assay conditions, whereas in the case of the Ptl+M213G and Ptl+M213P variants ten P1' amino acids allowed peptides that scored a synthetic yield of 30% or higher. By performing a coupling reaction with four-fold more concentrated substrates at pH 8.5, the synthetic yield was further improved, e.g., to 91% for the coupling of H-Glu-Leu-Arg-NH₂ to the standard Cam ester using the M213P mutant, which scored only 36% product with these substrates in the initial screening (Figure 4). Therefore, we reasoned that all peptides with P1' amino acid residues scoring >30% product can be used as efficient acyl acceptors in enzymatic ligation reactions, indicating that with the two new enzyme variants around 50% of all amino acids are well accepted at the P1' position in peptide coupling reactions.

To investigate the influence of the second residue forming the S1' pocket (L208), we produced L208X site-saturation libraries from both the Ptl+M213G and the Ptl+M213P variants of peptiligase. With respect to the expression of the mutants, the L208 position proved to be much less sensitive to substitutions than the M213 position and all 40 variants, i.e., the M213G+L208X and M213P+L208X Ptl mutants, were obtained in good quantity and purity (30 mg L⁻¹ culture, >95% purity).

The P1' substrate profile of these 40 enzymes was mapped by coupling the standard acyl donor (Ac-Asp-Phe-Ser-Lys-Leu-OCam-Leu-OH) to the peptides of the same peptide library (H-Xxx-Leu-Arg-NH₂) that was used to identify the M213G/P parents, again resulting in many variants with a unique P1' substrate profile. Substitution of L208 with the small amino acids (G, A, S, N) broadened the substrate profile for both the M213G and M213P peptiligase variants, especially by allowing coupling of nucleophilic peptides with more bulky groups (Trp, Met, Asn, Leu, Phe and His) in the P1' position. The variants with the broadest substrate profile proved to be Ptl+M213G+L208F/G and Ptl+M213P+L208H/A, accepting in total 17 amino acids at P1' giving >30% synthesis (Figure 5). The M213G+L208F and M213P+L208H mutants of peptiligase are distinctive enzyme variants with an extraordinarily broad P1' substrate profile.

Some amino acids, including glutamic acid, are still poorly accepted at the S1' pocket by the L208X variants discussed above. In contrast, introduction of some polar L208 mutations into the M213G or the M213P template increased the selectivity of nucleophile acceptance. For example, substitution of L208 by a positively charged amino acid resulted in enzyme variants with a high activity for glutamic and aspartic acids at P1' (Figure 6). Still other enzyme variants, e.g. with the L208D mutation, proved to be very efficient for coupling peptides with an asparagine at the P1' position (Figure 7). Thus, we have created a tool-

Figure 5. P1' substrate profile of peptiligase variants Ptl+M213G+L208F/G and Ptl+M213P+L208H/A. Conversion of Cam-ester starting material was complete in all cases. Formation of synthetic product was determined by LC-MS.

Figure 6. P1' substrate profile of peptiligase Ptl+M213G+L208K/R. Conversion of acyl donor to synthetic product was determined by LC-MS. Consumption of Cam-ester starting material was complete in all reactions.

Figure 7. P1' substrate profile of peptiligase Ptl+M213G+L208D and Ptl+M213P+L208D. Conversion of acyl donor to synthetic product was determined by LC-MS. Consumption of Cam-ester starting material was complete in all reactions.

box of ligases with complementary selectivity that can be used for a diverse range of coupling reactions.

Chemoenzymatic Synthesis of Exenatide on a Gram Scale

To demonstrate the versatility of our enzyme platform and peptide ligation technology, we embarked on the synthesis of the 39-mer pharmaceutical peptide exenatide (Figure 8). This peptide is challenging to synthesize by chemical segment condensation, since no glycine or proline residues are present in the 5–28 sequence. An approach was chosen based on synthesis and coupling of two segments of similar length, combining the best of SPPS with acceptable sequences for the variant Ptl+M213P+L208H.

The H-1-21-OCam-Leu-OH peptide segment (H-His-Gly-Glu-Gly-Thr-Phe-Thr-Ser-Asp-Leu-Ser-Lys-Gln-Met-Glu-Glu-Glu-Ala-Val-Arg-Leu-OCam-Leu-OH) was synthesized at a gram scale by SPPS on a Wang resin with an overall yield of 56%, and with 97.1% purity after a single preparative HPLC step. The synthesis of the second segment H-22-39-NH₂ (H-Phe-Ile-Glu-Trp-Leu-Lys-Asn-Gly-Gly-Pro-Ser-Ser-Gly-Ala-Pro-Pro-Pro-Ser-NH₂) was performed at a gram scale by SPPS on a Rink resin with an overall yield of 65%, and with 96.2% purity after a single preparative HPLC step. Due to the phenylalanine at the P1' site this H-22-39-NH₂ exenatide fragment is not recognized by the original peptiligase (Figure 2), but the new variant Ptl+M213P+L208H can couple peptides with phenylalanine at P1' very efficiently (Figure 5). The mutations in the S1' pocket endow this variant with a strong preference for the H-Phe-Leu sequence (P1' and P2' residues) present at the *N*-terminus of the H-22-39-NH₂ segment over the H-His-Gly pair at the *N*-terminus of the H-1-21-OCam-Leu-OH acyl donor segment. As a consequence, we reasoned that the *N*-terminal amine functionality of the H-1-21-OCam-Leu-OH peptide could be left unprotected and that the redesigned P1'–P2' preference of the M213P+L208H peptiligase should allow synthesis of exenatide without the use of any protecting groups. The final active pharmaceutical ingredient was indeed obtained on a gram scale in a single enzymatic step as described below.

The Ptl+M213P+L208H variant (1.7 mg) was added to a phosphate buffered solution (1 M, pH 8.1) of H-1-21-OCam-Leu-OH·4TFA (22 mg mL^{−1}) and H-22-39-NH₂·2TFA (17 mg mL^{−1}, 1.1 equivalents). The reaction mixture, with a total volume of 30 mL (and a theoretical H-1-39-NH₂·5TFA yield of 1.05 g), was stirred at room temperature and monitored by LC-MS over time (Figure 9). After full conversion of the ester segment (90 min), only 12.9% of ester hydrolysis was observed (H-1-21-OH), and no cleavage of the

Figure 8. Synthesis of exenatide from two segments using peptiligase Ptl + M213P + L208H.

peptide bonds (by protease activity) was detected. The exenatide product was obtained in 86.0% crude yield, and after acidification to pH 5 with TFA the crude mixture was directly separated by preparative HPLC from the starting materials and some hydrolyzed Cam-ester. After lyophilization 821 mg H-1-39-NH₂·5TFA were obtained (78.2% yield based on the Cam-ester). After this convenient isolation procedure, the overall yield of the purified exenatide based on the loading of the first amino acid of the Cam-ester segment was 43.5%, which is significantly higher than the reported yields for full SPPS of exenatide (10–25%).^[6] The final exenatide product was 99.6% pure and the chiral purity of all amino acids was verified (see the Supporting Information). Although the number of preparative HPLC steps was not reduced (three in total), the overall cost reduction for this enzymatic process for the production of exenatide was estimated at 40–50% compared to processes based on full SPPS. The enzyme consumption was only around 2 mg per g exenatide product, so that the enzyme costs are almost neglectable.

Conclusions

In summary, we have mapped the substrate profile of peptiligase which revealed that the S1' pocket is most discriminating when compared to the S1–S4 and S2'

pockets, which limits the applicability of peptiligase in peptide ligation. Using structure-inspired protein engineering the substrate profile of the S1' pocket was radically broadened. Moreover, we obtained selective peptiligase variants with redesigned substrate profiles that allow the use of new peptide ligation sites and avoid the necessity of *N*-terminal protection. The applicability of these engineered peptide ligases was demonstrated by the chemoenzymatic synthesis of exenatide on a gram scale from two unprotected segments. Currently, we are further expanding the enzyme platform by enzyme engineering methods.

Experimental Section

Construction of Peptiligase Variants

The peptiligase M213X library was prepared by the QuikChange method (Stratagene) using a pBS42-based *E. coli*-*B. subtilis* shuttle vector harboring the peptiligase gene as described before.^[11] After screening in *B. subtilis* six active variants were re-cloned into the shuttle vector pBE-S DNA (Takara Bio Inc.) using MluI-BamHI restriction sites. The construct pBE-S-Ptl included the *aprE* promoter sequence, the peptiligase signal peptide sequence, the coding sequence for the mature peptiligase and a *C*-terminal hexahistidine tag, and a terminator sequence. This plasmid was also used for preparation of the M213P/G + L208X libraries, again by the QuikChange mutagenesis. DpnI-digested and purified

Figure 9. HPLC analysis of chemoenzymatic exenatide synthesis reactions using the Ptl+M213P+L208H as the coupling enzyme. Samples were analyzed after different reaction times.

plasmids were transformed to the competent *E. coli* TOP10 strain and transformants were selected on LB-agar plates containing $100 \mu\text{g mL}^{-1}$ ampicillin. The mutant genes were confirmed by DNA sequencing.

Expression and Purification of Peptilgase Variants

The shuttle plasmids carrying mutants of interest were transformed into *B. subtilis* GX4935, a strain deficient in extracellular neutral and serine proteases that was a kind gift from Prof. P. N. Bryan (University of Maryland, USA).^[19] Colonies were picked and grown overnight in 5 mL LB with kanamycin ($10 \mu\text{g mL}^{-1}$) at 37°C in a shaking incubator. To the 30 mL Terrific Broth supplemented with antibiotic (kanamycin $10 \mu\text{g mL}^{-1}$) and amino acids (100 mg L^{-1} Trp, 100 mg L^{-1} Met and 100 mg L^{-1} Lys) 0.6 mL of the overnight culture were added. The cells were grown for 48 h at 37°C in a shaking incubator (200 rpm). The cells were harvested by centrifugation (15 min, 4,000 rpm, 4°C). The medium (30 mL) was decanted and concentrated on an Amicon centrifugal unit (15 mL, 10 kDa cut-off) in two centrifugation steps (15 min, 4000 rpm, 4°C). The concentrated medium (0.5 mL) was then exchanged for buffer A (25 mM Tricine, pH 7.5, 0.5 M NaCl, 20 mM imidazole) in three washing/concentration steps (14 mL buffer A, 10 min, 4,000 rpm, 4°C).

For His-tag purification Talon resin (2.5 mL, Clontech) was added to a plastic column cartridge. The resin was washed with 5 mL MilliQ water and equilibrated with 5 mL of buffer A. The crude enzyme was loaded on the column and washed with 5 mL buffer A. The enzyme was eluted with 5 mL buffer B (25 mM Tricine, pH 7.5, 0.5 M NaCl, 200 mM imidazole). The eluate was concentrated on an Amicon-centrifugal unit (5 mL, 10 kDa cut-off) by centrifugation (15 min, 4,000 rpm, 4°C) and buffer was exchanged to 25 mM Tricine, pH 7.5, in three washing/concentrating steps (5 mL buffer, 10 min, 4,000 rpm, 4°C). The protein concentration was measured using NanoDrop (A280 nm) and the purity was checked with SDS-PAGE and estimated to be 90%. The purified enzyme was flash-frozen in liquid nitrogen and stored at -80°C until further use.

Oligopeptide-OCam-Leu-OH Ester Synthesis

1 g of Fmoc-Leu-Wang resin (with a loading of 0.72 mmol/g) was washed with dichloromethane (DCM) (2×2 min, 10 mL) and DMF (2×2 min, 10 mL) and Fmoc-deprotected using piperidine/DMF (1/4, v/v, 2×8 min, 10 mL). After washing with DMF (2×2 min, 10 mL), DCM (2×2 min, 10 mL) and DMF (2×2 min, 10 mL), iodoacetic acid (4 equiv.) was coupled to the resin using DCC (4 equiv.) and HOAt (4 equiv.) in DCM (45 min, 10 mL). After washing with DMF (2×2 min, 10 mL), DCM (2×2 min, 10 mL) and THF (2×2 min, 10 mL), the resin was loaded with an Fmoc-protected amino acid using 4 equiv. Fmoc-Xxx-OH and 10 equiv. DIPEA in DMF/THF (1/1, v/v, 10 mL) at 50°C for 20 h. After washing with DMF (2×2 min, 10 mL), DCM (2×2 min, 10 mL) and DMF (2×2 min, 10 mL), standard SPPS protocols were followed to elongate the peptide.^[20] Optionally the *N*-terminal amine function was acetylated using a mixture of Ac_2O (10 vol%), DIPEA (5 vol%), HOBt (0.2 wt%) in DMF (2×10 min, 10 mL). Cleavage from the resin and side-chain deprotection was performed using a mixture of trifluoroacetic acid (TFA), triisopropylsilane (TIS) and water (95/2.5/2.5, v/v/v, 15 mL) for 120 min. The crude peptide was precipitated using methyl *tert*-butyl ether (MTBE)/*n*-heptanes (1/1, v/v, 50 mL). The precipitated peptide was collected by centrifugation and washed twice with MTBE/*n*-heptanes (1/1, v/v, 50 mL) followed by lyophilization from acetonitrile/water (1/1, v/v, 50 mL). The crude peptide Cam-Leu-OH esters were purified by preparative HPLC and pure fractions were lyophilized. For the exenatide H-1-21-OCam-Leu-OH segment, 5 g of resin were used.

Oligopeptide Amine Nucleophile Synthesis

1 g of Rink resin {4-[(2,4-dimethoxyphenyl)(Fmoc-amino)-methyl]phenoxyalkyl linker, with a loading of 0.64 mmol/g} was washed with DCM (2×2 min, 10 mL) and DMF (2×2 min, 10 mL) and Fmoc-deprotected using piperidine/DMF (1/4, v/v, 2×8 min, 10 mL). Standard SPPS protocols were followed to elongate the peptide.^[20] Cleavage from the resin and side-chain deprotection were performed using a mixture of TFA/TIS/water (95/2.5/2.5, v/v/v, 15 mL) for 120 min. The crude peptide was precipitated using MTBE/*n*-heptanes (1/1, v/v, 50 mL). The precipitated peptide was collected by centrifugation and washed twice with MTBE/*n*-heptanes (1/1, v/v, 50 mL) followed by lyophilization from acetonitrile

(Acn)/water (1/1, v/v, 50 mL). The crude P1' and P2' peptide libraries (H-Xxx-Leu-Arg-NH₂ and H-Ala-Xxx-Arg-NH₂, respectively) were used as such for the enzymatic screening reactions. For exenatide, the crude peptide amide nucleophile was purified by preparative HPLC and pure fractions were lyophilized. For the exenatide H-22–39-NH₂ segment, 5 g of resin were used.

Substrate Scope Screening

The ligation reactions were typically performed at 20 °C in a 96-well plate. Three different stock solutions were prepared: (i) 100 mM Cam-ester (library 1–4) in milliQ water, (ii) 200 mM amine nucleophile (library 5 and 6) in milliQ water, and (iii) 2 µg mL⁻¹ of ligase in 1 N phosphate buffer, pH 8.0. To each well, 100 µL Cam-ester stock solution, 100 µL amine nucleophile stock solution and 800 µL ligase stock solution were added. Aliquots of 200 µL were quenched in 800 µL acetonitrile/water/MSA (48/48/4, v/v/v) and analyzed by HPLC. For screening the Cam-ester libraries 1–4 the reaction time (before quenching and analysis) was 15 min. For screening of the amine nucleophile libraries 5 and 6 the reaction time was 30 min (full conversion of the Cam-ester starting material to ligation product and hydrolyzed by-product). Conversions and synthesis are expressed as % of acyl donor starting material. All screening experiments with peptide libraries included the original peptidylase (Ptl) and one identical peptide as a control. Synthetic yields calculated from such replicates varied by less than 10%.

Synthesis of Exenatide (H-1–39-NH₂)

The Cam-ester fragment H-1–21-OCam-Leu-OH (573.8 mg, 0.192 mmol) and amine nucleophile fragment H-22–39-NH₂ (434.3 mg, 0.21 mmol, 1.1 equiv.) were dissolved in 26.361 g of phosphate buffer (0.76 M, final pH: 8.12). The Ptl M213P L208H solution was added (3.940 g stock solution, corresponding to 1.7 mg pure enzyme) and the mixture was stirred at ambient temperature for 90 min. The reaction was monitored by LC-MS in time (1 µL of reaction mixture directly injected into the HPLC after 10, 30, 60 and 90 min). The crude reaction mixture was acidified to pH 5.01 using TFA (1.908 g). Part of the reaction mixture (20.685 g) was diluted with ACN/H₂O (47.117 g, 1/3, v/v) and 51.077 g of this mixture were directly purified by preparative HPLC and freeze-dried. Preparative HPLC loading: 0.1 wt%; conditions: gradient from 22% ACN/H₂O to 60% ACN/H₂O (containing 0.1% TFA) in 40 min. A Varian Dynamax Pursuit C18 column with 10 micron particles was used.

Acknowledgements

The authors thank Dr. J.-M. van der Laan and Dr. R. Lax for useful discussions and Dr. H. J. Wijma for assistance with the molecular modelling.

References

- [1] R. A. Copeland, *Enzymes*, Wiley-VCH, **2000**.
- [2] H. Korhonen, A. Pihlanto, *Curr. Pharm. Des.* **2007**, *13*, 829–843.
- [3] D. Hans, P. R. Young, D. D. P. Fairlie, *Med. Chem.* **2006**, *2*, 627–646.
- [4] E. Farkas, M. Ryadnow, *Amino acids, Peptides and Proteins*, Royal Society of Chemistry, **2013**.
- [5] V. Vergote, C. Burvenich, C. Van de Wiele, B. Spiegel-eer, *J. Pept. Sci.* **2009**, *15*, 697–710.
- [6] a) P. Mi, J. Liu, Y. Ma, J. Yuan, (Hybio Pharmaceutical), WO Patent WO2014067084 A1, **2014**; 22.0% yield. b) P. Mi, H. Li, L. Qin, Y. Ma, J. Yuan, (Shenzhen Hanyu Pharmaceutical), Chinese Patent CN101538324, **2009**; 23.8% yield. c) T. M. Ivanovich, E. I. Ivanovich, G. I. Gluzdikov, I. A. Aleksandrovich, N. S. Konstantinovna, (St. Petersburg University), Russian Patent RU2458066, **2012**; 33.0% crude yield.
- [7] T. Nuijens, P. J. L. M. Quaedflieg, H.-D. Jakubke, *Hydrolysis and synthesis of peptides*, in: *Enzyme catalysis in organic synthesis*, Wiley-VCH, **2011**.
- [8] F. Bordusa, *Chem. Rev.* **2002**, *102*, 4817–4868.
- [9] D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243–247.
- [10] A. C. Braisted, J. K. Judice, J. A. Wells, *Methods Enzymol.* **1994**, *289*, 298–313.
- [11] A. Toplak, T. Nuijens, P. J. L. M. Quaedflieg, B. Wu, D. B. Janssen, *Adv. Synth. Catal.* **2016**, *358*, 2140–2147.
- [12] O. Almog, T. Gallangher, M. Tordova, J. Hoskins, P. Bryan, G. L. Gilliland, *Proteins* **1998**, *31*, 21–32.
- [13] E. K. Bratovanova, D. D. Petkov, *Biochem. J.* **1987**, *248*, 957–960.
- [14] a) T. Nuijens, A. Toplak, M. B. A. C. van de Meulereek, M. Schmidt, M. Goldbach, P. J. L. M. Quaedflieg, *Tetrahedron Lett.* **2016**, *57*, 3635–3638; b) R. J. A. C. De Beer, T. Nuijens, L. Wiermans, P. J. L. M. Quaedflieg, F. P. Rutjes, *Org. Biomol. Chem.* **2012**, *10*, 6767–6775.
- [15] L. M. Bech, S. B. Slarensen, K. Breddam, *Biochemistry* **1993**, *32*, 2845–2852.
- [16] a) W. J. Moree, P. Sears, K. Kawashiro, K. Witte, C.-H. Wong, *J. Am. Chem. Soc.* **1997**, *119*, 3942–3947; b) P. Sears, M. Schuster, P. Wang, K. Witte, C.-H. Wong, *J. Am. Chem. Soc.* **1994**, *116*, 6521–6530; c) S. L. Strausberg, R. Biao, K. E. Fisher, P. A. Alexander, P. N. Bryan, *Biochemistry* **2005**, *44*, 3272–3279.
- [17] D. A. Estell, T. P. Graycar, J. A. Wells, *J. Biol. Chem.* **1985**, *260*, 6518–6521.
- [18] P. A. Bury, R. T. Carr, J. E. Crider, B. S. Dehoff, R. A. Gabel, B. T. King, (F. Hoffmann-La Roche), WO Patent WO2009053315, **2009**.
- [19] S. R. Fahnestock, K. E. Fischer, *Appl. Environ. Microbiol.* **1987**, *5*, 379–384.
- [20] W. C. Chan and P. D. White, *Fmoc solid phase peptide synthesis: a practical approach*, Oxford University Press, **2000**.