Bioconjugate Chemistry

Article

pubs.acs.org/bc

Recombinant Production of Peptide C-Terminal α-Amides Using an Engineered Intein

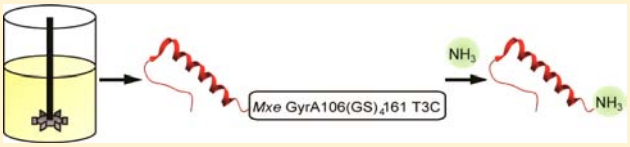
Louise Albertsen,*[,†,‡] Allan C. Shaw,[†] Jens C. Norrild,[†] and Kristian Strømgaard[‡]

[†]Novo Nordisk A/S, Novo Nordisk Park 1, DK-2760 Måløv, Denmark

[‡]Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark

Supporting Information

ABSTRACT: Peptides are of increasing interest as therapeutics in a wide range of diseases, including metabolic diseases such as diabetes and obesity. In the latter, peptide hormones such as peptide YY (PYY) and pancreatic peptide (PP) are important templates for drug design. Characteristic for these peptides is that they contain a C-terminal that is α-amidated, and this amidation is crucial for biological function. A challenge is to generate such peptides by recombinant means and particularly in a production scale. Here, we have examined an intein-mediated approach to generate a PYY derivative in a larger scale. Initially, we experienced challenges with hydrolysis of the intein fusion protein, which was reduced by a T3C mutation in the intein. Subsequently, we further engineered the intein to decrease the absolute size and improve the relative yield of the PYY derivative, which was achieved by substituting 54 residues of the 198 amino acid intein with an eight amino acid linker. The optimized intein construct was used to produce the PYY derivative under high cell density cultivation conditions, generating the peptide thioester precursor in good yields and subsequent amidation provided the target peptide.



INTRODUCTION

Peptides are a rapidly growing class of therapeutics with more than 50 peptide-based products currently on the market and even more in development, covering disease areas such as immunology, oncology, neurology, and endocrinology.[1,2] Peptides regulate a plethora of physiological functions, mainly by interactions with specific cellular receptors, whereby they induce cellular signaling events such as neurotransmission and release of hormones. Endogenous peptides have been associated with challenges as therapeutics due to their limited *in vivo* stability and bioavailability. However, the high specificity and low toxicity combined with improved ability to selectively modify and improve therapeutic properties of peptides has increased the relevance of peptides in drug development.[3]

In endocrinology, diseases are often caused by or associated with an imbalance of the level of peptide hormones, as seen in diseases such as diabetes and obesity. Notably, about half of the peptide hormones in the endocrine and nervous systems are α-amidated in their C-terminal[4] and the α-amide moiety is often crucial for biological activity and stability.[5] A particularly interesting class of C-terminal α-amidated peptides are those of the neuropeptide Y (NPY) family: NPY, pancreatic peptide (PP), and peptide YY (PYY).[6] These peptides mediate their physiological action by binding to four different subtypes of NPY receptors, designated Y_1, Y_2, Y_4, and Y_5, which are G protein-coupled receptors involved in appetite regulation. PYY is released from endocrine L cells in the intestine following meals, and upon activation of the Y_2 receptor subtype, food intake is reduced[7,8] and glucose disposal improved.[9,10] Thus, peptide derivatives of PYY have potential as therapeutic drugs for the treatment of obesity.[11,12] The major circulating form of human PYY is PYY(3−36),[13] in which the N-terminal residues Tyr-Pro has been cleaved by dipeptidyl peptidase IV.[14] PYY(3−36) is a selective Y_2 receptor agonist and the five C-terminal residues, including the terminal α-amide, are essential for receptor activation.[15]

The most widely used technologies for production of peptide therapeutics are microbial expression systems and chemical synthesis.[2] While a peptide C-terminal amide is easily achieved by chemical synthesis, it is not readily introduced into recombinant peptides derived from microbial hosts, which lack an α-amidating enzymatic machinery. In certain cases this has been overcome by subsequently employing peptidylglycine α-amidating monooxygenase[16,17] or introduction of an α-amidated amino acid by enzymatic condensation using thermolysin[18,19] or transacylation using carboxypeptidase Y.[20] An alternative can be to use intein fusion proteins, which in a few cases have been used for amidation of recombinant peptides.[21]

The intein-mediated approach combines advantages of molecular biology with chemistry, and enables the addition of a wide range of unnatural functionalities to a recombinant protein.[22] In this approach, a protein containing a C-terminal α-thioester can be generated by expression as an N-terminal fusion to the intein. The protein thioester can then be reacted with a synthetic peptide comprising an N-terminal cysteine[23,24] or the thioester can be used as a chemical handle for bioconjugation[25,26] or amidation.[21] The primary limitation of

Received: June 3, 2013
Revised: October 19, 2013
Published: October 19, 2013

using this technology for large scale production of *C*-terminally α-amidated peptides is the low yields generally observed, which may be ascribed to a combination of the large size of the intein[21] and hydrolytic instability of the intein fusion protein.[27−31]

We have investigated the perspectives of using the intein-mediated methodology for producing therapeutically relevant peptides containing a *C*-terminal α-amide. Specifically, we describe the generation of the PYY derivative, PYY(4−36)K4G-NH_2. We expressed this derivative as an *N*-terminal fusion to the Gyrase A intein from *Mycobacterium xenopi* (*Mxe* GyrA) (Figure 1) followed by thiolysis to the peptide α-thioester,

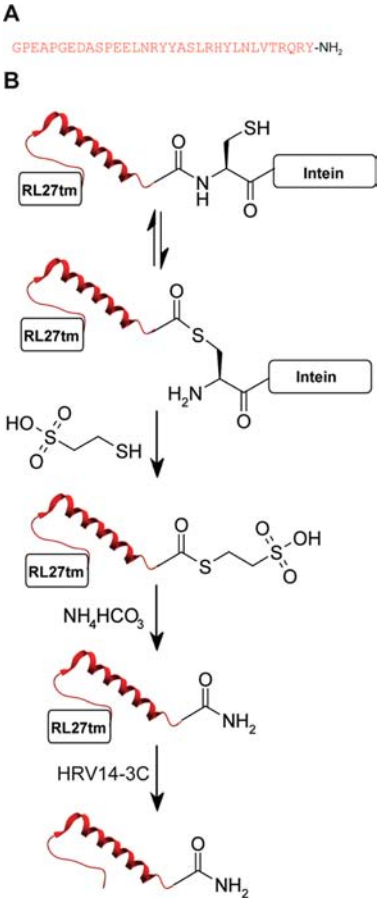


Figure 1. Intein-based strategy to generate *C*-terminal α-amidated PYY(4−36)K4G. (A) Sequence of PYY(4−36)K4G-NH_2. (B) Schematic representation of intein-based production of PYY(4−36)K4G-NH_2. PYY was genetically fused between an *N*-terminal purification tag and a *C*-terminal intein. An *N* to *S* acyl shift at the peptide−intein junction of the fusion protein resulted in the formation of a thioester intermediate, which was subsequently cleaved by trans-thioesterification using MESNa. The formed protein thioester was converted into the corresponding protein amide before removing the purification tag, which liberated the PYY(4−36)K4G-NH_2.

which was transformed to the *C*-terminal amide with ammonia. This process was optimized by addressing the two primary shortcomings of the methodology, namely, hydrolysis of the intein fusion protein and the intein size. Both these objectives were accomplished by engineering of the *Mxe* GyrA intein. The optimized methodology was used to generate the PYY(4−36)K4G-NH_2 under high cell density conditions in good yields, demonstrating the feasibility of the improved methodology for production of *C*-terminally α-amidated peptides.

EXPERIMENTAL PROCEDURES

Materials and Methods. All synthetic gene fragments were codon optimized for expression in *E. coli* and obtained from Geneart (Regensburg, Germany). Human rhinovirus 14−3C (HRV14−3C) protease and p1081 plasmid, a pET11a expression vector containing a gene fragment encoding a variant of ribosomal protein L27 from *Thermotoga maritima*[32] (in the following named RL27tm), were from Novo Nordisk (Måløv, Denmark). All restriction and ligase enzymes were purchased from New England Biolabs (Hertfordshire, UK) and DNA purification kits were obtained from Qiagen (Hilden, Germany). Competent *E. coli* TOP10 and BL21(DE3) cells, SeeBlue Plus2 prestained standard, NuPAGE Bis-Tris gels (4−12%) and 20x MES buffer were supplied by Invitrogen (Renfrew, UK) and isopropyl β-D-1-thiogalactopyranoside (IPTG) was purchased from AH Diagnostics (Aarhus, Denmark). Porcine trypsin was purchased from Promega (Madison, WI). 4× XT sample buffer for SDS-PAGE analysis was obtained from Bio-Rad Laboratories (Hercules, CA) and tris(2-carboxyethyl)phosphine (TCEP) 0.5 M bond breaker solution was purchased from Thermo Scientific (Rockford, IL). Ammonia bicarbonate, dithiothreitol (DTT), and ethylenediaminetetraacetic acid (EDTA) were obtained from Applichem GmBH (Darmstadt, Germany). Trifluoroacetic acid (TFA) and HPLC grade acetonitrile were purchased from Alfa Aesar (Ward Hill, MA) and Merck (Darmstadt, Germany), respectively. All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Molar extinction coefficients (280 nm) for determination of protein concentrations by UV_{280} were calculated according to Pace et al.[33]

All protein purifications were performed on an ÄKTA explorer 100 FLPC system from GE Healthcare. The proteins were analyzed by analytical RP-UPLC on Waters Acquity UPLC System using 0.1% TFA in water (solvent A) and 0.08% TFA in acetonitrile (solvent B) and UV detection at 214 nm. Three different UPLC methods were used; *UPLC method A*: 40 °C, C4 Acquity UPLC BEH300 column (1.7 μm; 2.1 × 100 mm) using a linear gradient of 20−60% B in 15 min at 0.4 mL/min. *UPLC method B*: 50 °C, C18 Acquity UPLC BEH column (1.7 μm; 2.1 × 150 mm) connected to a C18 BEH van guard pre column (1.7 μm; 2.1 × 5 mm) using a linear gradient of 28−33% B in 10 min, followed by a linear gradient of 33−53% B in 11 min at 0.4 mL/min. *UPLC method C*: As UPLC method B using a linear gradient of 28−33% B in 10 min. Unless otherwise specified, mass spectrometric analysis was performed on Agilent LC-MSD-TOF 1100 series instrument equipped with a C18 Zorbax 300SB rapid resolution column (3.5 μm; 2.1 × 50 mm) using water with 8.8 mM ammonium formate and 0.1% formic acid (solvent A) and 0.1% formic acid in acetonitrile (solvent B). *LC-MS method A*: 40 °C, linear gradient of 18−25% B in 1 min, followed by a linear gradient of 25−65% B in 10 min at 0.3 mL/min. *LC-MS method B*: 40 °C, linear gradient of 5−70% B in 15 min at 0.3 mL/min.

Construction of Plasmids. Plasmids p3292, p2948, and p2952 encoding proteins **1**, **3**, and **4**, respectively, were obtained by ligating synthetic gene fragments coding for human PYY(4−36)K4G [modified from UniProt P10082_3] fused between a 5′-end linker region (SSSEVLFQ) carrying a HRV14−3C[34] protease site and a 3′end *Mxe* GyrA intein variant (modified from UniProt P72065_2) into p1081 using

dx.doi.org/10.1021/bc4002689 | *Bioconjugate Chem.* 2013, 24, 1883−1894

T4 DNA ligase. The *Mxe* GyrA intein here referred to as wild-type (WT) carried an N198A mutation[24] to disable the self-splicing function. Plasmids p3606, p3608, p3609, p3610, p3617, p3619, and p3620 encoding proteins **5**-**11**, respectively, were constructed from synthetic gene fragments coding for PYY(14−36) fused to 3′-end minimized *Mxe* GyrA intein mutants, which were ligated into p3292. To obtain plasmid p3776 and p3803 encoding proteins **2** and **12**, respectively, a synthetic gene fragment coding for PYY(14−36) fused to a 3′end *Mxe* GyrA(1−5) T3C fragment was ligated into plasmid p3292 and p2952, respectively. All ligation products were used to transform competent TOP10 *E. coli* cells and the cells were cultured overnight at 37 °C on LB-amp (100 μg/mL) plates. Transformants carrying plasmids with the proteins of interest as verified by DNA sequencing (Eurofins MWG Operon, Germany) were used to transform *E. coli* BL21(DE3) cells.

Expression of Fusion Proteins in Shake Flasks. *E. coli* BL21(DE3) cells transformed with each fusion protein plasmid were grown in shaker flasks with baffles containing LB medium plus 100 μg/mL ampicillin at 37 °C and 230 rpm until OD_{600} = 0.4−0.6. For induction temperatures of 18 and 30 °C, the cells were cooled down to the desired temperature 20 min before induction. Protein expression was induced by addition of 0.5 mM IPTG for 20 h at 18 °C or for 3 h at 30 or 37 °C. After harvesting the cells by centrifugation in a Heraeus Multifuge 3 S-R centrifuge (4000 × *g*, 15 min), the wet pellets were stored at −20 °C until further use. For reducing SDS-PAGE analysis, pellets from 1 mL uninduced and induced culture were solubilized in 1× sample dye (Table 1). Moreover, to separate the soluble and insoluble protein fraction, pellets from 1 mL induced culture were sonicated in lysis buffer (Table 1), followed by centrifugation, and to the resulting pellets and supernatants were added 1× and 4× sample dye (Table 1), respectively. The extent of hydrolysis was calculated as $area_{Mxe\ GyrA}/(area_{Mxe\ GyrA} + area_{Fusion\ protein})$ after densitometric analysis of the SDS-PAGE gels using ImageJ 1.45s software. Harvested cells obtained from induction at 37 °C for 3 h were lysed by sonication on a Bandelin Sonopuls HD 3100 sonicator (40% power with pulses of 15 s on and off for 20 min on ice) in lysis buffer (Table 1). Lysis was performed at pH 5 as it resulted in a larger extent of insoluble protein for those of the fusion proteins, which were obtained in both fractions at neutral pH as analyzed by SDS-PAGE. The insoluble inclusion bodies (IBs) were spun down (4000 × *g*, 15 min) and the cell lysate decanted. The IBs were washed twice in Milli-Q water, divided into smaller fractions, and stored at −20 °C.

Table 1. Summary of Buffers Used in the Experimental Procedures

buffer name	composition
1× sample dye	1× XT sample buffer, 1 mM TCEP
4× sample dye	4× XT sample buffer, 4 mM TCEP
CD buffer	10 mM Tris-ClO_4 (pH 8)
Denaturation buffer 1	100 mM sodium phosphate, urea, 1 mM TCEP, 1 mM EDTA (pH 7.5)
Denaturation buffer 2	25 mM sodium phosphate, 8 M urea, 1 mM TCEP, 1 mM EDTA (pH 7.5)
Dilution buffer 1	100 mM sodium phosphate, 250 mM NaCl (pH 7.5)
Dilution buffer 2	100 mM sodium phosphate, 150 mM NaCl, 2 M urea (pH 7.5)
Dilution buffer 3	100 mM sodium phosphate, 150 mM NaCl, 1.3 M urea, 1 mM TCEP (pH 7.5)
Dilution buffer 4	25 mM sodium phosphate, 500 mM NaCl, 1 mM EDTA, 1 mM TCEP (pH 7.5)
Lysis buffer	25 mM sodium phosphate (pH 5)
Wash buffer	25 mM sodium phosphate, 1 mM EDTA, 1% Triton X-100 (pH 7)

Kinetic Studies of Thiol-Induced Intein Cleavages. IBs (from 10 mL culture induced at 37 °C) of fusion proteins were solubilized in denaturation buffer 1 (Table 1, 5 M urea), incubated on ice for 1 h, and filtered (0.45 μm). The proteins were characterized by LC-MS method A (Table S3) and protein concentrations were estimated by UPLC method A followed by dilution to a final concentration of 0.6 mg/mL with denaturation buffer 1 (5 M urea). For fusion proteins that were poorly extractable in a 5 M urea buffer, a similar denaturation buffer containing 8 M urea was used and the protein was diluted to 0.6 mg/mL and 5 M urea. The proteins were refolded by dilution into a mixture of dilution buffers 1 (2.5 volumes) and 2 (0.6 volumes) in a volume ratio of 1:1.5:0.6, giving a final concentration of 0.2 mg/mL protein in 100 mM sodium phosphate (pH 7.5), 2 M urea, 150 mM NaCl. Thiolysis was induced by addition of MESNa to 100 mM from a 2 M stock in water and pH was measured to ∼7.3. Typically, reactions were carried out in Eppendorf tubes in a volume of 1 mL with incubation at 4 °C overnight on a Thermomixer (300 rpm). Aliquots were collected over time and quenched with one volume of 1.7% HCl in 6 M GnHCl to pH ∼3. The progression of the thiol-induced cleavage was followed by UPLC method B and the different protein constituents were characterized by LC-MS method A. Reactions were performed in triplicate. Cleavage efficiencies were defined as $area_{product}/(area_{product} + area_{fusion\ protein}) \times 100\%$ with the product being RL27tm-PYY(4−36)K4G-MESH α-thioester. Efficiencies were plotted as a function of time. Pseudo-first-order reactions were achieved by using MESNa in high molar excess and rate constants (k_{obs}) were determined by fitting the data to a one phase exponential association equation [$P = P_0(1 - e^{-kt})$, where P is the percentage of formed RL27tm-PYY(4−36)K4G-MESH α-thioester at time t, P_0 was estimated as the maximum percentage of RL27tm-PYY(4−36)K4G-MESH α-thioester obtained, and k was the observed rate] using GraphPad Prism software.

Purification of RL27tm-PYY(4−36)K4G-MESH α-Thioester. Cells from 200 mL culture (induced at 37 °C) of each of proteins **1**, **2**, and **12** were lysed in lysis buffer by sonication on ice for 20 min (40% power with pulses of 15 s on and off for 20 min on ice) and the insoluble fraction obtained by centrifugation was washed twice in wash buffer (Table 1) followed by one wash in 100 mM sodium phosphate (pH 7.5). SDS-PAGE analysis indicated that no protein was present in the washes. The isolated IBs were solubilized in 20 mL denaturation buffer 1 (Table 1, 8 M urea) and incubated for 2 h at 4 °C with weak agitation. The expression yields of fusion proteins were estimated by UPLC method A using a purified standard of known concentration. The denatured protein sample was centrifuged (4000 × *g*, 40 min) and filtered (0.45 μm) before refolding by rapid dilution into 9 volumes of cold dilution buffer 3 (Table 1) resulting in a concentration of 135 mM NaCl and 2 M urea. Immediately after dilution, the cleavage was initiated by adding MESNa to a final concentration of 100 mM from a 2 M stock in dilution buffer 3. Reaction was allowed at 4 °C for 20 h at pH 7.3, followed by analysis using UPLC method B and LC-MS method A. The reaction mixture was diluted by three volumes of eluent A [50

dx.doi.org/10.1021/bc4002689 | *Bioconjugate Chem.* 2013, 24, 1883−1894

mM sodium phosphate (pH 7), 1 mM EDTA] to decrease the conductivity to ~12 mS/cm before loading onto a HiTrap SP Sepharose HP (16 × 25 mm; 5 mL, GE Healthcare) column at 5 mL/min. The RL27tm-PYY(4−36)K4G-MESH α-thioester was eluted using a linear gradient from 0% to 100% eluent B [50 mM sodium phosphate (pH 7), 1 mM EDTA, 1 M NaCl] over 20 column volumes at 3 mL/min. Fractions containing the RL27tm-PYY(4−36)K4G-MESH α-thioester were pooled and the amount was estimated by UPLC method C.

Purification and Biophysical Characterization of Inteins. The *Mxe* GyrA WT, 106$(GS)_4$161 and 106$(GS)_4$161 T3C intein variants were purified from IBs of fusion proteins **1**, **5**, and **12**, respectively. The IBs were solubilized in denaturation buffer 2 (Table 1) and refolded by rapid dilution into 9 volumes of dilution buffer 4 (Table 1). After overnight incubation at 4 °C with 100 mM MESNa (added from a 2 M stock in dilution buffer 4) and subsequent desalting in a stirred ultrafiltration cell with a 10K MWCO Millipore generated cellulose filter, the inteins were isolated by anion exchange on a HiTrap Q Sepharose HP column (16 × 25 mm; 5 mL) using gradient elution with eluent A [10 mM Tris (pH 7.5), 2 mM DTT] and eluent B [10 mM Tris (pH 7.5), 2 mM DTT, 1 M NaCl] from 0% to 100% B over 20 column volumes at 3 mL/min. The intein-containing fractions were loaded on a HiLoad Superdex 75 16/600 column and eluted in 10 mM Tris (pH 7.5), 150 mM NaCl, 2 mM DTT at 1 mL/min. Pooled fractions were desalted in Vivaspin 20 (10 K MWCO) PES columns (Sartorius) and analyzed by UPLC method A to purities of >96%. Far-ultraviolet circular dichroism (CD) spectra were recorded at 20 °C over the 190−260 nm range, using a J-815 CD spectropolarimeter (Jasco, UK) with a resolution of 0.5 nm in 1 mm path length cells. Protein samples of 10 μM were prepared by diluting the purified inteins with CD buffer (Table 1). Concentrations were estimated from OD_{280} after recording UV spectra on a Hewlett-Packard 8453 spectrophotometer, which confirmed that the three inteins all had absorption maxima of ~276 nm. Data were collected at 0.5 nm intervals at a scanning speed of 50 nm/min. The CD spectra were expressed as molar circular dichroism ΔE [$cm^{-1} \times M^{-1}$] after normalization (the collected data was normalized to the molar concentration of peptide bond and smoothed by the Savitzky-Golay algorithm).[35] Fluorescence measurements were made using a LS50B luminescence spectrometer (Perkin-Elmer, MA, USA) with the emission and excitation slit set to 2.5 nm. The fluorescence spectra were collected over the range of 300−500 nm at a scan speed of 20 nm/min and a protein concentration of 2 μM in CD buffer using an excitation wavelength of 280 nm.

Expression of Protein 12 in Fed-Batch Culture. As inoculum for the high-cell density fermentations, *E. coli* cells containing the plasmid p3803 was grown in LB medium supplied with 100 μg/mL ampicillin in shaker flasks for 6−8 h and inoculated into a bioreactor to reach an initial OD_{600} of about 0.2. Fermentations were carried out under aerobic conditions in 500 mL bioreactors (DasGib Technology) with an initial volume of 200 mL of a defined fermentation medium with glucose and ammonia as carbon and nitrogen sources, respectively, and with 100 μg/mL ampicillin added after sterilization (autoclavation at 121 °C for 30 min). The pH was maintained at 7.0 by addition of 5 N NH_4OH, the temperature was maintained at 37 °C, and an air flow rate of 0.4 L/min was sparkled through the culture broth throughout the fermentation period. The agitation speed (400−1200 rpm) was controlled to reach a dissolved oxygen level of at least 30% O_2 saturation. The initial glucose concentration in the fermentation medium was 10 g/L, and from 5 h of fermentation, a glucose feed supplied with magnesium and trace metals was added continuously in increasing steps until a final feed rate of 10 g glucose/L/h, which was kept until the fermentation stopped. The production of the fusion proteins was induced by the addition of 0.5 mM IPTG at an OD_{600} of 50−60. The fermentation was stopped after 4 h induction. Aliquots of 40 μL culture were lysed in 1× sample dye by sonication and the expression yields were estimated by densitometric analysis of SDS-PAGE gels using ImageJ. The cells were harvested by centrifugation in a Sorvall RC 6 Plus centrifuge with a F10−6 × 500y Rotor (13000 × *g*) and stored at −20 °C. The cells were resuspended in lysis buffer to OD_{600} of 40−50 and stirred 1 h at 4 °C. Cells were lysed on a constant cell disruption system E615 at a pressure of 1.36 Kbar. The lysed cells were spun down and the supernatant decanted. The insoluble IBs were washed in wash buffer and stirred at 4 °C for 2 h, followed by one wash in Milli-Q water. The IBs were divided equally into 20 smaller fractions.

Isolation of PYY(4−36)K4G-NH_2 from Fed-Batch Culture. IBs from 1/20 fed-batch culture were solubilized in denaturation buffer 2 (Table 1, 30 mL) and incubated at 4 °C for 3.5 h with gentle stirring. The protein concentration was estimated by UPLC method A and the sample was diluted to a concentration of ~1.2 mg/mL with denaturation buffer. The protein was refolded by adding the protein over 5 min into 8.5 volumes of dilution buffer 4 (Table 1) with stirring. Thiol-induced cleavage was initiated immediately by adding MESNa to 100 mM from a 2 M stock in dilution buffer 4. Reaction was allowed in the 0.8 M urea and 0.5 M NaCl solution at 4 °C and pH 7.3 overnight with stirring. Amidation was initiated by addition of solid NH_4HCO_3 to a final concentration of 1 M. The pH of the solution was adjusted to 8.5 and reaction was allowed at 4 °C overnight. Thiol-induced cleavage and amidation reactions were followed by UPLC method B and LC-MS method A. The amidation mixture was filtrated (0.22 um) and the filter washed with 8 M urea. The urea wash was loaded on a Phenomenex Luna column (15 μm and 300 Å; 10 × 250 mm) followed by the remaining amidation mixture. The RL27tm-PYY(4−36)K4G-NH_2 fusion protein was eluted using 0.1% TFA in water (solvent A) and 0.1% TFA in acetonitrile (solvent B) at a linear gradient of 25−45% B. After lyophilization the protein was dissolved in 25 mM sodium phosphate (pH 7.5) and the RL27tm extension was removed by enzymatic cleavage using HRV14 3C protease in an enzyme to substrate ratio of 1:12 (w/w) by incubation at room temperature for at least 48 h. The cleavage was followed by UPLC method C and LC-MS method B. After adjusting pH of the digest to 4.3, the PYY(4−36)K4G-NH_2 was recovered by loading at 3 mL/min onto a prepacked HiTrap SP sepharose HP column (5 mL) equilibrated with eluent A [10 mM NH_4HCO_3 (pH 5.5)], followed by a gradient of 0−100% eluent B [10 mM NH_4HCO_3 (pH 8.5)] over 5 column volumes, and the PYY(4−36)K4G-NH_2 was eluted using isocratic eluent B for 10 CV. The fractions containing the PYY(4−36)K4G-NH_2 were lyophilized and the peptide analyzed by UPLC method C and LC-MS method B. Yields were estimated by chemoluminescence based nitrogen detection.

Actone Functional Potency Assay. The Y_2 receptor is G_i-coupled and signals mainly through the cAMP dependent

pathway by inhibiting adenylate cyclase activity, which results in a decrease of cAMP production from ATP. The actone assay is based on a modified calcium channel that has a selective binding for cAMP, resulting in cellular calcium influx, detected by a calcium responsive dye. In order to measure decreased levels of cAMP, as result of NPY receptor activation, isoproterenol a $\beta1/\beta2$-adrenoreceptor agonist is added to activate adenylate cyclase and increase cAMP levels in the cell. Decreased cellular calcium concentrations, reflecting a decrease of cAMP levels due to NPY receptor activation, is detected as a decrease in fluorescence from the calcium sensitive dye. HEK-293 cells expressing the cAMP sensitive calcium channel and the Y_2 receptor (Codex Biosolution, Gaithersburg, MD, USA) were seeded into poly(lysine) coated 384 well plates (BD Biosciences, Franklin Lakes, NJ, USA) at a density of 14,000 cells/well in a volume of 25 μL in DMEM medium (Lonza, Verviers, Belgum) containing 10% FCS (Invitrogen, Carlsbad, CA, USA), 1% penicillin-streptomycin (Lonza), 250 μg/mL G418 (Invitrogen), and 1 μg/mL puromycin (Sigma-Aldrich, St. Louis, MO, USA). The cells were incubated overnight at 37 °C in a humidified milieu in 5% CO_2 followed by addition of 25 μL calcium dye buffer containing 1 vial Calcium 5 dye (part nr: 5000625 Molecular Devices, Sunnyvale, CA, USA) dissolved in 100 mL buffer containing 20 mM Hepes, 0.1% Ovalbumin, 0.005% Tween 20, 1.5 mM probenecid (Sigma-Aldrich), 250 μM cAMP phosphodiesterase inhibitor (4-(3-butoxy-4-methoxybenzyl)imidazolidin-2-one, Sigma-Aldrich), and 8 mM $CaCl_2$, and pH was adjusted to 7.4. Cells were incubated for 1 h with the calcium dye buffer and then placed in a FLIPR Tetra System (Molecular Devices) where the liquid handling system added analogue and isoproterenol (Sigma-Aldrich, 0.05 μM final concentration) simultaneously, directly followed by fluorescence signal measurement (Ex540/Em590) for 360 s with 30 s intervals. All measurements were performed in duplicates and EC50 values were calculated by nonlinear regression analysis of sigmoidal dose response curves using the GraphPad Prism. All data presented represent the mean value of three individual assays performed at separate occasions ± SD.

RESULTS

Design of Fusion Proteins. For the generation of α-amidated PYY, a recombinant dual fusion protein strategy was selected, in which the PYY peptide was inserted between an *N*-terminal purification tag and a *C*-terminal Gyrase A intein from *Mycobacterium xenopi* (*Mxe* GyrA) (Figure 1). The *Mxe* GyrA intein was selected as it tolerates a broad range of amino acids in the *C*-terminal of the target peptide.[36] In addition, Mxe GyrA can be efficiently refolded from inclusion bodies (IBs),[37] and allows effective thiolysis in up to 4 M urea.[38,39] The physiologically active PYY variant, PYY(3−36), contains an *N*-terminal isoleucine residue. An initiator methionine in the *N*-terminal of the PYY(3−36) sequence will not be efficiently removed by methionine aminopeptidase during expression in *E. coli*,[40] which makes an enzymatic step for correct processing of the *N*-terminal necessary. We therefore fused an alkaline, bacterial ribosomal protein, RL27tm, to the *N*-terminal of PYY via a protease site. The alkaline properties of an *N*-terminal RL27tm extension have previously been shown to be beneficial for expression and purification of recombinant proteins by cation exchange chromatography.[32] Initially, an enterokinase site was introduced between the *N*-terminal RL27tm extension and PYY(3−36); however, treatment with recombinant bovine enterokinase led to extensive unspecific degradation of PYY(3−36). Therefore, we instead used the PYY variant, PYY(4−36)K4G (Figure 1A), and fused the RL27tm extension via a sequence recognized by the more selective HRV14−3C protease. The strategy was to express the RL27tm-PYY(4−36)K4G-*Mxe* GyrA fusion protein, followed by thiol-induced cleavage of the intein, to generate the RL27tm-PYY(4−36)K4G α-thioester, which could subsequently be converted to the α-amidated target peptide PYY(4−36)K4G by amidation and removal of the RL27tm extension (Figure 1B).

Hydrolytic Stability of *Mxe* GyrA WT Fusion Protein. Fusion protein **1**, RL27tm-PYY(4−36)K4G-*Mxe* GyrA (33 kDa), was successfully expressed in *E. coli* BL21(DE3); however, products from hydrolysis of the fusion protein, corresponding to RL27tm-PYY α-acid (12 kDa) and *Mxe* GyrA WT (21 kDa) were observed (Figure 2, S1 and S2 as well as

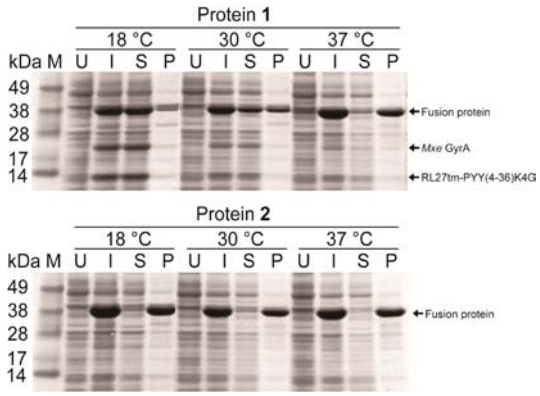


Figure 2. Expression of proteins **1** and **2** monitored by reducing SDS-PAGE. TCEP was used as reducing agent. M, molecular mass markers; U, uninduced; I, induced; S, soluble fraction; P, insoluble pellet fraction. Fusion protein refers to RL27tm-PYY(4−36)K4G-*Mxe* GyrA WT (**1**) or T3C (**2**) fusion proteins.

Table S1 and S2). Hydrolysis of intein fusion proteins has previously been reported[27−30] and is obviously detrimental for the overall yield of the target protein. Moreover, the hydrolysis product, the peptide carboxylic acid of PYY (Figure S1 and Table S1), will require separation from the desired thioester, which may be challenging. We further observed that the hydrolysis was highly dependent on the temperature during induction with lower degree of hydrolysis observed at higher temperatures. This was explained by accumulation of the fusion protein into insoluble IBs at higher temperatures, thereby avoiding hydrolysis of the fusion protein. The extent of hydrolysis was estimated to be 32%, 15%, and 5% after expression of protein **1** at 18, 30, and 37 °C, respectively (Figure 2), and the identity of the RL27tm-PYY(4−36)K4G (Figure S1, Table S1) and *Mxe* GyrA (Figure S2, Table S2) hydrolysis products was confirmed by in-gel trypsin digestion. In addition, attempts to refold and cleave protein **1** from IBs were hampered by precipitation, as well as formation of hydrolysis products. We anticipated that increasing concentrations of urea during refolding and cleavage would decrease precipitation[38,39] and hydrolysis. At high urea concentrations (≥5 M), the intein was denatured and inactive, whereas in 2 M urea, the intein was active and substantial hydrolysis was observed (Figure S3). However, when the fusion protein was solubilized in 5 M urea with subsequent dilution to 2 M immediately prior to adding thiol, hydrolysis was reduced.

dx.doi.org/10.1021/bc4002689 | *Bioconjugate Chem.* 2013, 24, 1883−1894

Thus, hydrolysis is not only occurring during expression, but is observed whenever the intein is on the active refolded form.

Reducing Hydrolysis by Introducing a T3C Mutation. In a previous attempt to avoid hydrolysis, introduction of a Thr-to-Cys mutation at position 3 in the *Mxe* GyrA intein reduced hydrolysis, which was considered to be due to the formation of a disulfide bond between the introduced Cys and Cys1 of the intein.[27] Following this approach, we introduced the T3C mutation into fusion protein **1**, resulting in RL27tm-PYY(4−36)K4G-*Mxe* GyrA T3C protein **2**. Expression demonstrated that hydrolysis was eliminated regardless of the induction temperature. For this construct we also observed that IBs formation was promoted at all three temperatures (Figure 2), which could explain the decreased hydrolysis. Also, for related fusion proteins of human PP an increased amount of IBs formation was observed for the *Mxe* GyrA T3C mutant compared to WT at all three tested induction temperatures, but in this case, hydrolysis was also highly reduced when the T3C containing mutant was soluble expressed after induction at 18 °C (Figure S4).

Having examined the beneficial effects for hydrolytic stability of the T3C mutation, we were interested in investigating its effect on intein activity. This was examined by following the thiol-induced cleavages of WT (**1**) and T3C (**2**) intein fusion proteins over time. Hence, both fusion proteins were isolated from IBs by solubilization in a denaturation buffer and refolded by rapid dilution into dilution buffer, resulting in a final concentration of 2 M urea. In preliminary experiments, the refolding was associated with precipitation, which could be reduced by including TCEP and a mixture of TCEP and NaCl in the denaturation and dilution buffers, respectively (Table S3). Immediately after dilution, the thiol-induced cleavages were induced by addition of excess MESNa. The reactions were monitored by UPLC (Figure 3A), resulting in pseudo-first-order reactions from which the apparent first-order rate constants (k_{obs}) were determined (Figure 3B). The T3C intein in protein **2** showed an approximate 4-fold decreased activity compared to WT intein in protein **1** with k_{obs} of $(2.5 \pm 0.2) \times 10^{-4}$ s^{-1} and $(1.0 \pm 0.1) \times 10^{-3}$ s^{-1} (Table 2), respectively. Next, the RL27tm-PYY(4−36)K4G-MESH thioester was isolated by cation exchange chromatography from IBs of fusion protein **1** and **2**, respectively (Figure S5), and the yields were estimated to 23 mg/L and 24 mg/L for the *Mxe* GyrA WT (**1**) and *Mxe* GyrA T3C (**2**) fusion proteins, respectively (Table 3). Further analysis showed that the content of the *C*-terminal α-acid byproduct was lower for protein **2**; however, this did not result in a correspondingly higher yield of RL27tm-PYY(4−36)K4G-MESH thioester.

Minimization of the *Mxe* GyrA Intein Size. A general challenge for the use of the intein-based expression system as a cost-efficient methodology for production of peptides and proteins is the size of the intein relative to the target peptide or protein. We observed expression yields of the intein fusion proteins **1** and **2** in the range of 119−130 mg/L; however, removing the intein reduced the absolute yield to 23 and 24 mg/L, respectively. A decreased intein size would reduce the overall construct size and potentially lead to increased peptide yields, and we therefore focused on minimization of the *Mxe* GyrA intein.

Inteins share a common Hedgehog/Intein (HINT) domain, which consist of a 12 β strand structural fold.[41] The X-ray crystal structure of the *Mxe* GyrA intein shows that two α-helices and a flexible loop region are placed between β-strand 9

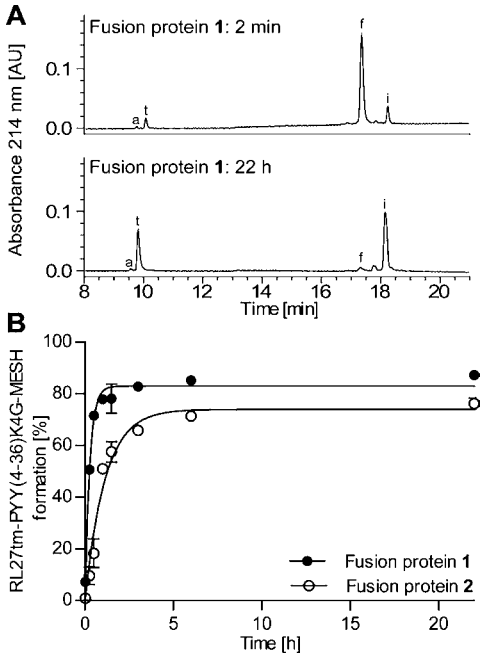


Figure 3. Thiol-induced cleavage profiles of protein **1** and **2**. (A) Representative RP-UPLC traces of the thiol-induced cleavage of protein **1** after 2 min and 22 h of reaction: a, RL27tm-PYY(4−36)K4G−OH; f, RL27tm-PYY(4−36)K4G-*Mxe* GyrA; i, *Mxe* GyrA intein; t, RL27tm-PYY(4−36)K4G-MESH thioester. (B) Thiol-induced cleavage curves. Error bars represent the standard deviations from three determinations. The calculated k_{obs} values are presented in Table 1

and 10,[42] and this region is projected away from the central HINT core. We speculated that this region or part of it could be excised from the intein without compromising the splicing activity. Initially, two constructs from protein **1** were designed, where either 46 (112−157, RL27tm-PYY(4−36)K4G-*Mxe* GyrA111Δ158, protein **3**) or 54 residues (107−160, RL27tm-PYY(4−36)K4G-*Mxe* GyrA106Δ161, protein **4**) were excised (Figure 4). In protein **4**, the entire linker region between β-strands 9 and 10 was removed, whereas protein **3** contains a few linker residues on each of the two connected strands to conserve flexibility (Figure 4A and B). The fusion proteins were obtained as IBs (Figure S6), and after isolation of the fusion protein from IBs by solubilization and refolding, intein activities were determined from the calculated k_{obs} values of the thiol-induced cleavages (Table 2, Figure 4C). Protein **4** was inactive, but protein **3** with remaining linker residues had only slightly reduced activity ($k_{obs} = (7.0 \pm 0.5) \times 10^{-4}$ s^{-1}) relative to the WT intein; thus, a short flexible linker between β-strands 9 and 10 seemed to be crucial for intein activity. Therefore, a flexible Gly-Ser linker (GSGSGSGS) was inserted into protein **4** providing RL27tm-PYY(4−36)K4G-*Mxe* GyrA106$(GS)_4$161 (protein **5**), which now had thiol-induced cleavage activity ($k_{obs} = (8.6 \pm 0.5) \times 10^{-4}$ s^{-1}) similar to the native intein. A range of other sequences were inserted to evaluate the effect of differences in structural integrity and charge of the linker region (LDRHGN, RDVETGE, ADNLALA, RRRRRRRR, GRGSGRGS, and GRGRGRGR), but these all had reduced activity relative to protein **5** (Figure S7A, Table S4). To examine if the intein in protein **5** could be even further minimized, residues at either side of the linker were systematically removed, resulting in six fusion proteins **6**-**11** (Table 2). While truncation *N*-terminally to the linker

Table 2. Properties of *Mxe* GyrA Intein Variants in RL27tm-PYY(4-36)K4G-Intein Fusion Proteins and Kinetic Characterization of their Thiol-Induced Cleavages

protein	*Mxe* GyrA intein mutant	deletion sites[a]	linker[b]	intein size	k_{obs}[c] [s^{-1}]
1	Wildtype			198	$(1.0 \pm 0.1) \times 10^{-3}$
2	T3C			198	$(2.5 \pm 0.2) \times 10^{-4}$
3	111Δ158	Ser111 Asp158		152	$(7.0 \pm 0.5) \times 10^{-4}$
4	106Δ161	Gln106 Phe161		144	n.d.
5	$106(GS)_4161$	Gln106 Phe161	GSGSGSGS	152	$(8.6 \pm 0.5) \times 10^{-4}$
6	$105(GS)_4161$	Ile105 Phe161	GSGSGSGS	151	$(4.2 \pm 0.1) \times 10^{-4}$
7	$104(GS)_4161$	Val104 Phe161	GSGSGSGS	150	$(5.4 \pm 0.3) \times 10^{-5}$
8	$103(GS)_4161$	Ala103 Phe161	GSGSGSGS	149	n.d.
9	$106(GS)_4162$	Gln106 Tyr162	GSGSGSGS	151	$(7.9 \pm 0.4) \times 10^{-4}$
10	$106(GS)_4163$	Gln106 Tyr163	GSGSGSGS	150	$(6.9 \pm 0.2) \times 10^{-4}$
11	$106(GS)_4164$	Gln106 Ala164	GSGSGSGS	149	$(5.1 \pm 0.3) \times 10^{-4}$
12	$106(GS)_4161$ T3C	Gln106 Phe161	GSGSGSGS	152	$(7.6 \pm 0.4) \times 10^{-5}$

[a]Residues between the given sites have been deleted in the respective intein mutants. [b]Linker inserted in between the deletion sites. [c]k_{obs} was calculated as described in the experimental procedures. Standard deviations are from nonlinear fits of triplicated data to the one-phase association equation. n.d.: not determined.

Table 3. Purification of RL27tm-PYY(4-36)K4G-MESH α-Thioester from Shake Flask Culture of Different Fusion Proteins[a]

protein	*Mxe* GyrA intein mutant	yield of fusion protein [mg/L]	yield of RL27tm-PYY(4−36)K4G-MESH[b] [mg/L]	recovery [mol %]	extent of α-acid byproduct[c] [%]
1	Wild-type	119	23	52	6.0
2	T3C	130	24	49	1.5
12	$106(GS)_4161$, T3C	147	43	69	1.5

[a]The yields obtained from 200 mL shake flask cultures were normalized to yield per liter culture. Data are presented as means from two independent experiments. [b]Amount of α-acid byproduct not included. [c]Estimated as $area_{RL27tm\text{-}PYY(4-36)K4G-OH}/(area_{RL27tm\text{-}PYY(4-36)K4G-OH} + area_{RL27tm\text{-}PYY(4-36)K4G-MESH})$ by RP-UPLC.

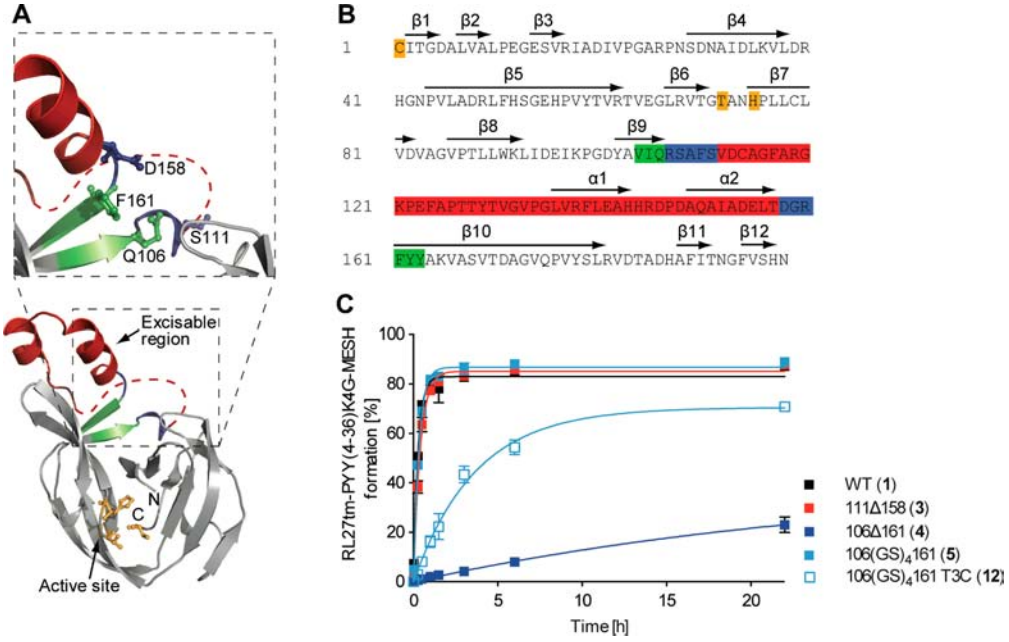


Figure 4. Minimization of the *Mxe* GyrA intein. (A) Ribbon presentation of the *Mxe* intein (PDB code: 1AM2) with the active site residues shown in yellow sticks and balls. The excisable region is colored according to the corresponding deletion site with the 111Δ158 deletion in red, the 106Δ161 deletion in blue, and stepwise truncation on either site of the 106Δ161 deletion in green. The disordered loop (residues 112−129) is represented by a dashed red stroke. Close-up view of the deletion sites with Ser111, Asp158, Gln106, and Phe161 shown in sticks. (B) Sequence of the *Mxe* intein with the position of secondary structure elements indicated by arrows. The deleted regions are highlighted in same color as in the structural representations. (C) Thiol-induced cleavage curves for WT and selected minimized *Mxe* GyrA intein variants. Error bars represent the standard deviations from three determinations. The calculated k_{obs} values are presented in Table 2

caused a dramatic decrease in intein activity, *C*-terminal truncation only led to a slight reduction in activity compared to protein **5** (Figure S7B and Table 2).

Next, the T3C mutation was introduced into protein **5** resulting in fusion protein RL27tm-PYY(4−36)K4G-*Mxe* GyrA$106(GS)_4161$ T3C (protein **12**), which reduced thiol-

induced cleavage activity approximately 13- and 11-fold compared to protein **1** and **5**, respectively (Figure 4C and Table 2). To determine whether the minimized intein led to increased peptide yields, the RL27tm-PYY(4−36)K4G-MESH α-thioester was isolated from protein **12**, resulting in a yield of 43 mg/L (Table 3), which is nearly twice the yield obtained from proteins **1** and **2**. After thiol-induced cleavage of proteins **1**, **5**, and **12** isolated from IBs the corresponding inteins, *Mxe* GyrA WT, *Mxe* GyrA106$(GS)_4$161, and *Mxe* GyrA106-$(GS)_4$161 T3C, respectively, were purified and subjected to CD and fluorescence spectroscopy. In the CD spectra, truncated inteins from proteins **5** and **12** showed similar spectral properties, indicating a similar secondary structure, while the full-length intein from protein **1** containing two additional α-helices was significantly different (Figure 5A). The

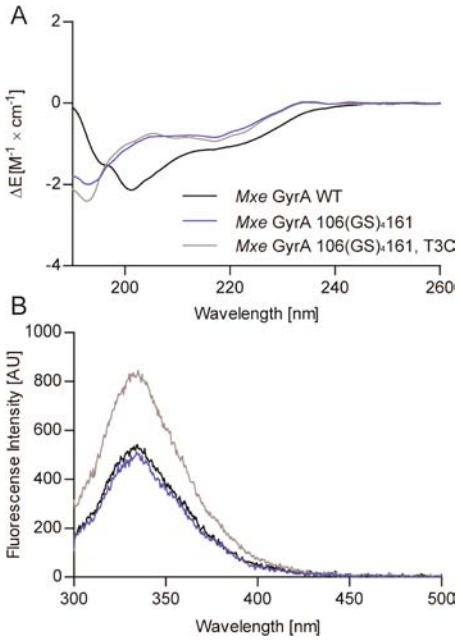


Figure 5. Biophysical characterization of *Mxe* GyrA intein variants. The *Mxe* GyrA WT, 106$(GS)_4$161, and 106$(GS)_4$161 T3C inteins were purified from protein **1**, **5**, and **12**, respectively, and subjected to CD (A) and fluorescence spectroscopy (B).

fluorescence properties of inteins from proteins **1** and **5** were comparable (Figure 5B), suggesting a conserved tertiary structure, while the intein from protein **12** containing the T3C mutation was markedly different, indicating that this mutation affected the tertiary structure.

Production of PYY(4−36)K4G-NH_2 in Large Scale. The applicability for production of α-amidated PYY (Figure 1) from the optimized fusion protein **12**, RL27tm-PYY(4−36)K4G-*Mxe* GyrA106$(GS)_4$161 T3C, on an industrial scale depends on its ability to express under high-cell density conditions. Protein **12** was therefore produced using controlled cultivation in a bioreactor and was as expected obtained as insoluble IBs, which were isolated and solubilized. Next, the protein was refolded, and when screening for refolding conditions for the fed-batch derived protein, we observed that highest protein recovery was obtained when adding 0.5 M NaCl in the dilution buffer (Figure S8). The protein was then immediately subjected to thiol-induced cleavage resulting in the RL27tm-PYY(4−36)K4G-MESH α-thioester, which was subsequently converted to the corresponding *C*-terminal α-amide by addition of 1 M NH_4HCO_3. The high conductivity caused by the presence of NaCl and NH_4HCO_3 rendered purification by cation exchange chromatography difficult, and we therefore replaced this step with a reverse-phase HPLC purification step. Finally, RL27tm was removed by treatment with HRV14−3C protease and the target PYY(4−36)K4G-NH_2 peptide was obtained after cation exchange chromatography. Each step of the purification process was monitored by UPLC (Figure S9) and SDS-PAGE (Figure 6A) and the final product was characterized by UPLC and LC-

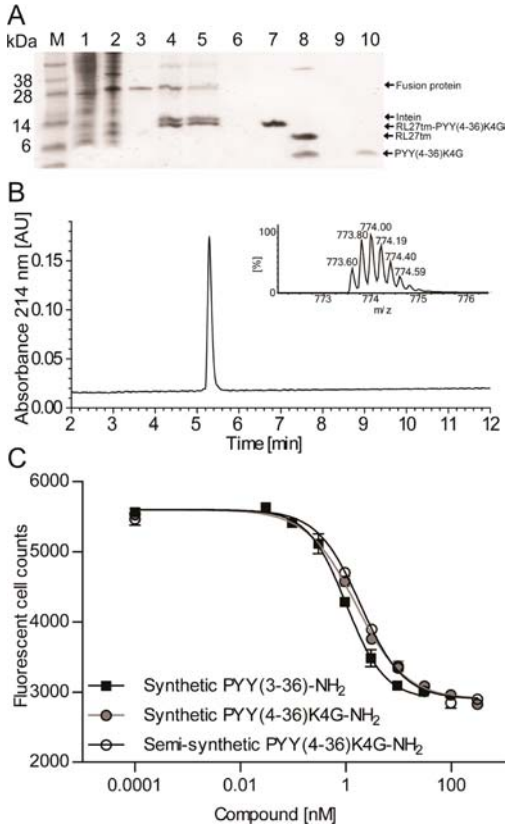


Figure 6. Purification of PYY(4−36)K4G-NH_2 after fed-batch cultivation of protein **12**. (A) SDS-PAGE analysis showing the different purification steps. Lane 1, uninduced cells; lane 2, induced cells; lane 3, solubilized inclusion bodies; lane 4, thiol-induced cleavage mixture; lane 5, amidation mixture; lane 6, flow-through from RP-HPLC; lane 7, pure RL27tm-PYY(4−36)K4G-NH_2; lane 8, HRV14−3C digestion; lane 9, flow-through from cation exchange chromatography; lane 10, pure PYY(4−36)K4G-NH_2; M, molecular mass markers. (B) RP-UPLC of the pure PYY(4−36)K4G-NH_2 with the insert showing ESI-MS of the $[M+5H]^{5+}$ ion (expected monoisotopic mass: 773.59). (C) Representative dose−response curves of semi-synthetic (EC_{50} = 1.37 ± 0.82) and synthetic PYY(4−36)K4G-NH_2 (EC_{50} = 1.38 ± 0.34) and synthetic PYY(3−36)-NH_2 (EC_{50} = 0.90 ± 0.10) in functional Y_2 receptor assays. For each concentration, the response was plotted as mean ± SEM, n = 2. The EC_{50} values were estimated from three independent experiments and giving as mean ± SD.

MS (Figure 6B). The average yield of fusion protein **12** from two independent cultivations was 4.4 g/L with a molar recovery of the desired peptide of ca. 30% after purification, corresponding to an estimated yield of 184 mg PYY(4−36)K4G-NH_2 per liter fed-batch culture. In order to verify the biological activity of PYY(4−36)K4G-NH_2, two reference peptides, PYY(4−36)K4G-NH_2 and PYY(3−36)-NH_2, were

dx.doi.org/10.1021/bc4002689 | *Bioconjugate Chem.* 2013, 24, 1883−1894

prepared by solid-phase peptide synthesis (Figure S10). The three peptides were tested on the human Y_2 receptor in a functional assay, showing that the semisynthetic analogue had an EC_{50} value that was identical to that of the synthetic analogues (Figure 6C).

DISCUSSION

Peptides and proteins are becoming increasingly relevant as drug candidates, with a concomitant interest in developing and improving methodologies to generate therapeutic peptides and proteins in large scale. A particular class of therapeutic peptides are those that are α-amidated in their *C*-terminal, such as pramlintide, an analogue of amylin. Pramlintide is produced by chemical synthesis in contrast to many other therapeutic peptides and proteins, which are generated by recombinant expression.[2] A specific challenge for recombinant expression of *C*-terminal α-amidated peptides and proteins is that the most frequently used expression systems, such as yeast and bacteria, are not capable of introducing the *C*-terminal amide.

Here, we investigated an intein-mediated strategy,[23] which can generate *C*-terminal α-amidated peptides by preparation of a peptide α-thioester and subsequent α-amidation by a nucleophilic attack. The *C*-terminal α-thioester can in principle react with any nucleophile and the strategy therefore allows a wider range of *C*-terminal modifications, which is not possible with other amidation strategies such as thermolysin[18,19] or carboxypeptidase Y[20] catalyzed ligation and peptidylglycine α-amidating monooxygenase.[16,17] Another advantage of the intein-mediated α-amidation strategy is that the peptide α-thioester is generated by thiolysis of a peptide-intein fusion protein, avoiding the need for processing enzymes. The intein-based approach for *C*-terminal α-amidation of peptides has previously been described for generation of hormones like calcitonin, parathyroid hormone, and luteinizing hormone–releasing hormone, but only on a laboratory scale due to very low yields of the target peptides.[21]

Here, we have explored the intein-mediated amidation approach to generate a PYY derivative in larger scale. PYY belongs to the class of NPY peptides, which are important for metabolic homeostasis, thereby making these peptides or analogues hereof potential obesity drug candidates.[12] Specifically, we examined production of PYY(4–36)K4G-NH_2, which has Y_2 receptor affinity similar to PYY(3–36)-NH_2, and generated a construct involving an *N*-terminal RL27tm extension, a HRV14–3C protease site, the target peptide, and an engineered *Mxe* GyrA intein. Two particular challenges for application of this strategy are potential hydrolysis of the peptide-intein fusion protein during expression and purification as well as the large size of the intein relative to the peptide.

One of the major determinants in the degree of hydrolysis of the peptide–intein fusion protein is the nature of the amino acid preceding the intein cleavage site, referred to as the −1 residue. The *C*-terminal residue of PYY(4–36)K4G-NH_2 is tyrosine that, as demonstrated by expression of maltose binding protein fused to *Mxe* GyrA intein in *E. coli*, should result in efficient thiol-induced cleavage and no hydrolysis during expression in *E. coli*.[36] However, expression of RL27tm-PYY(4–36)K4G-*Mxe* GyrA (**1**) resulted in extensive hydrolysis, which suggest that other factors than the −1 residue must play a role for hydrolysis of intein–fusion proteins. The hydrolysis of protein **1** could be reduced by increasing the induction temperature leading to increased accumulation into IBs, where the protein is typically found in a non-native folded and inactive conformation.[43] A number of studies have described similar challenges of hydrolysis of intein–fusion proteins during expression[27−30] or purification,[31] which in all cases led to substantially reduced yields. Previous attempts to circumvent hydrolysis include fragment complementation of split inteins[44] or introduction of a Cys residue in position 3 of the intein. The latter was demonstrated for the *Mxe* GyrA[27] and *Saccharomyces cerevisiae* vacuolar membrane ATPase (*Sce* VMA)[31] inteins, where hydrolysis during expression was reduced leading to increased yields of the target protein.

We therefore introduced a T3C mutation in the *Mxe* GyrA intein resulting in RL27tm-PYY(4–36)K4G-*Mxe* GyrA T3C (**2**) and observed a reduction of hydrolysis and increased amounts of protein in IBs. It was previously suggested that the decreased hydrolysis from the T3C mutation could be due to disulfide bond formation with the catalytically active Cys1, preventing generation of the labile thioester bond between the target protein and intein.[27] The formation of such a disulfide bond is not likely to occur in the reducing environment of wildtype *E. coli* cytoplasm. This has previously been demonstrated in attempts to generate a similar intramolecular disulfide bond between Cys1 of inteins and a preceding Cys within a flanking target protein, which was only formed after expression in *E. coli* origami, a strain lacking two oxidoreductases (ΔtrxB and ΔgorA) and having an oxidizing cellular environment.[45,46] An alternative to using oxidative expression strains involves secretion of the intein fusion proteins into the *E. coli* periplasm,[47] which is a naturally oxidizing compartment containing disulfide bond forming enzymes. However, we used neither an oxidizing expression strain nor secretion into the periplasm, and in preliminary experiments we were not able to obtain data confirming the existence of such a disulfide bond.

We envisaged that the reduced hydrolysis could be an effect of the increased IBs formation, but other factors may also contribute, since we also observed greatly suppressed hydrolysis for RL27tm-PP-*Mxe* GyrA T3C when soluble expressed after induction at 18 °C. The intein activity in thiol-induced cleavage was therefore examined and showed a decreased rate of thiolysis for fusion protein **2** compared to **1**, which could potentially contribute to the reduced hydrolysis observed during expression. The T3C mutation itself did not increase the yield of the desired RL27tm-PYY(4–36)K4G-MESH thioester from **2** relative to **1**, in contrast to previous studies.[27,31] Our results might be influenced by the fact that both fusion proteins **1** and **2** were isolated from IBs, where hydrolysis was already diminished, thereby reducing the potential effect of the T3C mutation. Thus, the T3C mutation may diminish hydrolysis by increased IB formation, alternatively by reduced intein activity, or a combination of these. How this influences the yields of target proteins and peptides may depend on whether the fusion proteins are obtained as soluble protein or in IBs.

The other major challenge in generation of therapeutic peptides by use of inteins is the relative size of the intein, and this may impact the absolute yield of the target peptide, which represents only a fraction of the entire fusion protein. In previous studies, inteins have been reduced in size by removing either large internal protein domains or smaller flexible linker regions, while preserving the 12 β strand HINT fold which is essential for intein activity.[48−51] The *Mxe* GyrA intein (198 amino acids) has two α-helices and an unstructured region (a total of 54 amino acids), which is not part of the catalytically active HINT core.[42] We showed that this region could be

dx.doi.org/10.1021/bc4002689 | Bioconjugate Chem. 2013, 24, 1883−1894

replaced with a flexible Gly-Ser linker, resulting in RL27tm-PYY(4−36)K4G-*Mxe* GyrA106$(GS)_4$161 (**5**), without affecting the thiol-induced cleavage activity. The removed region is located between β-strands 9 and 10 of the HINT core, and the intein activity was highly dependent on the presence and nature of linking residues between these two β-strands, as removal of the Gly-Ser linker blocked the thiol-induced cleavage. In addition, removal of residues from β-strands 9 and 10 of protein **5** reduced the intein activity, indicating that the intein could not be minimized further in this region. Thus, we reduced the size of *Mxe* GyrA intein by approximately 25% without compromising the activity of the intein.

The combination of the reducing hydrolysis by the T3C mutation, driving expression of the fusion protein toward IBs, and increasing yields by minimizing the intein has, to the best of our knowledge, not been explored before. When combining these modifications in RL27tm-PYY(4−36)K4G-*Mxe* GyrA106$(GS)_4$161 T3C (**12**), we observed a higher reduction in thiol-induced cleavage activity than when the T3C mutation was inserted in the full-length intein. As the minimized intein itself had activity similar to that of the WT protein, these data indicate that the minimized intein is more prone to potential structural perturbations made by inserting the T3C mutation. The minimized intein with the T3C mutation (**12**) had an altered fluorescence profile from both WT and minimized intein, suggesting that the T3C mutation might perturb the HINT intein fold and thereby reduce the activity. Regardless of these putative structural changes, the *Mxe* GyrA106$(GS)_4$161 T3C could be refolded from IBs and the thiol-induced cleavage was efficient in the presence of urea, two critical features of the *Mxe* GyrA intein.[37−39] Isolation of RL27tm-PYY(4−36)K4G-MESH α-thioester from protein **12** resulted in an increased yield and limited hydrolysis compared to protein **1**. Therefore, protein **12** was used in a larger scale production of a *C*-terminal α-amidated peptide, PYY(4−36)K4G-NH_2. In only a few cases, intein fusion proteins have been generated by high-cell density fed-batch fermentation in yields of 127−400 mg/L and there is only one example of a protein that was obtained by *N*-terminal intein cleavage, although without transformation into an α-thioester.[52,53] We have demonstrated that PYY(4−36)K4G-NH_2 could be obtained after fed-batch cultivation of protein **12** via generation of the intermediate RL27tm-PYY(4−36)K4G-MESH α-thioester, which was efficiently converted to the corresponding RL27tm-PYY(4−36)K4G-NH_2. The RL27tm extension was enzymatically removed, liberating the free PYY(4−36)K4G-NH_2 with a yield of 184 mg/L. Thus, a functionally engineered intein, which has been minimized and combined with a T3C mutation, shows potential for industrial production of α-amidated peptides. Future large scale studies will evaluate the general applicability of this engineered intein to generate α-thioester peptide fragments useful for native chemical ligation,[54] for α-amidation reactions, or for production of *C*-terminal bioconjugated proteins and peptides.

■ ASSOCIATED CONTENT

Supporting Information

Procedures for production of the PP-containing fusion proteins and protein S1−S7, and synthesis of PYY(4−36)K4G and PYY(3−36). Data from in-gel trypsin digest of hydrolysis products is enclosed in Figure S1, Table S1, Figure S2, and Table S2. Figure S3 shows the solubilization and thiol-induced cleavage profiles of protein **1** at different concentrations of urea. Figure S4 shows SDS-PAGE analysis after expression of the PP-containing *Mxe* GyrA WT and *Mxe* GyrA T3C fusion proteins. Table S3 shows refolding data for protein **2** obtained from shake-flask culture. Figure S5 shows an example on SDS-PAGE, UPLC, and MS data from purification of RL27tm-PYY(4−36)K4G-MESH thioester. Figure S6 represents SDS-PAGE of expression of engineered intein fusion proteins. Figure S7 and Table S4 shows intein activity data for minimized inteins. Table S5 contains MS analysis of all fusion proteins. Figure S8 contains data for refolding of protein **12** after fed-batch fermentation. Figure S9 shows an example on UPLC and MS data from the purification of PYY(4−36)K4G-NH_2 after expression under high-cell density conditions. Figure S10 shows UPLC and MS data of the synthetic peptide analogues. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: +45 24247140; e-mail: louise.albertsen@sund.ku.dk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank B. B. Hansen and Dr. D. B. Steensgaard (Novo Nordisk, Måløv, Denmark) for technical assistance and discussions in the biophysical characterization of *Mxe* GyrA intein mutants. M. Mørkenborg and Dr. S. Østergaard (Novo Nordisk, Måløv, Denmark) are acknowledged for preparation of the reference peptides. We also thank Dr. J. Petersen (Novo Nordisk, Måløv, Denmark) for handling the high-cell density cultivations. Finally, H. Jensen (Novo Nordisk, Måløv, Denmark) is acknowledged for collecting MS analysis of the in-gel trypsin digests. This project was supported by the Danish Agency for Science, Technology and Innovation (Ministry of Science, Innovation and Higher Education).

■ ABBREVIATIONS

DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HRV14−3C, human rhinovirus14−3C; IBs, inclusion bodies; IPTG, isopropyl β-D-1-thiogalactopyranoside; MESNa, sodium 2-mercptoethanesulfonate; *Mxe* GyrA, *Mycobacterium xenopi* Gyrase A; PYY, peptide YY; RL27tm, ribosomal protein L27 from *Thermotoga maritime*; *Sce* VMA, *Saccharomyces cerevisiae* vacuolar membrane ATPase; TCEP, tris(2-carboxyethyl)-phosphine; WT, wild-type

■ REFERENCES

(1) Bellmann-Sickert, K., and Beck-Sickinger, A. G. (2010) Peptide drugs to target G protein-coupled receptors. *Trends Pharmacol. Sci. 31*, 434−441.

(2) Vlieghe, P., Lisowski, V., Martinez, J., and Khrestchatisky, M. (2010) Synthetic therapeutic peptides: science and market. *Drug Discovery Today 15*, 40−56.

(3) Frokjaer, S., and Otzen, D. E. (2005) Protein drug stability: a formulation challenge. *Nat. Rev. Drug Discovery 4*, 298−306.

(4) Eipper, B. A., and Mains, R. E. (1988) Peptide α-amidation. *Annu. Rev. Physiol. 50*, 333−344.

(5) Merkler, D. J. (1994) C-terminal amidated peptides: production by the in vitro enzymatic amidation of glycine-extended peptides and the importance of the amide to bioactivity. *Enzyme Microb. Technol. 16*, 450−456.

(6) Eipper, B. A., Stoffers, D. A., and Mains, R. E. (1992) The biosynthesis of neuropeptides: peptide alpha-amidation. *Annu. Rev. Neurosci. 15*, 57−85.
(7) Chelikani, P. K., Haver, A. C., Reeve, J. R., Jr., Keire, D. A., and Reidelberger, R. D. (2006) Daily, intermittent intravenous infusion of peptide YY(3−36) reduces daily food intake and adiposity in rats. *Am. J. Physiol. Regul. Integr. Comp. Physiol. 290*, R298−R305.
(8) Batterham, R. L., Cowley, M. A., Small, C. J., Herzog, H., Cohen, M. A., Dakin, C. L., Wren, A. M., Brynes, A. E., Low, M. J., Ghatei, M. A., Cone, R. D., and Bloom, S. R. (2002) Gut hormone PYY(3−36) physiologically inhibits food intake. *Nature 418*, 650−654.
(9) van den Hoek, A. M., Heijboer, A. C., Corssmit, E. P., Voshol, P. J., Romijn, J. A., Havekes, L. M., and Pijl, H. (2004) PYY3−36 reinforces insulin action on glucose disposal in mice fed a high-fat diet. *Diabetes 53*, 1949−1952.
(10) Ortiz, A. A., Milardo, L. F., DeCarr, L. B., Buckholz, T. M., Mays, M. R., Claus, T. H., Livingston, J. N., Mahle, C. D., and Lumb, K. J. (2007) A novel long-acting selective neuropeptide Y2 receptor polyethylene glycol-conjugated peptide agonist reduces food intake and body weight and improves glucose metabolism in rodents. *J. Pharmacol. Exp. Ther. 323*, 692−700.
(11) Karra, E., Chandarana, K., and Batterham, R. L. (2009) The role of peptide YY in appetite regulation and obesity. *J. Physiol. (Lond.) 587*, 19−25.
(12) Greenwood, H. C., Bloom, S. R., and Murphy, K. G. (2011) Peptides and their potential role in the treatment of diabetes and obesity. *Rev. Diabet. Stud. 8*, 355−368.
(13) Batterham, R. L., Heffron, H., Kapoor, S., Chivers, J. E., Chandarana, K., Herzog, H., Le Roux, C. W., Thomas, E. L., Bell, J. D., and Withers, D. J. (2006) Critical role for peptide YY in protein-mediated satiation and body-weight regulation. *Cell Metab. 4*, 223−233.
(14) Medeiros, M. D., and Turner, A. J. (1994) Processing and metabolism of peptide-YY: pivotal roles of dipeptidylpeptidase-IV, aminopeptidase-P, and endopeptidase-24.11. *Endocrinology 134*, 2088−2094.
(15) Beck-Sickinger, A. G., Wieland, H. A., Wittneben, H., Willim, K. D., Rudolf, K., and Jung, G. (1994) Complete L-alanine scan of neuropeptide Y reveals ligands binding to Y_1 and Y_2 receptors with distinguished conformations. *Eur. J. Biochem. 225*, 947−958.
(16) Tamburini, P. P., Jones, B. N., Consalvo, A. P., Young, S. D., Lovato, S. J., Gilligan, J. P., Wennogle, L. P., Erion, M., and Jeng, A. Y. (1988) Structure-activity relationships for glycine-extended peptides and the α-amidating enzyme derived from medullary thyroid CA-77 cells. *Arch. Biochem. Biophys. 267*, 623−631.
(17) Bradbury, A. F., Finnie, M. D., and Smyth, D. G. (1982) Mechanism of C-terminal amide formation by pituitary enzymes. *Nature 298*, 686−688.
(18) Morihara, K. (1991) Thermolysin catalyzed semisynthesis of peptide hormones by introduction of Phe-NH_2 or Tyr-NH_2 at the carboxyl termini. *Biomed. Biochim. Acta 50*, S15−S18.
(19) Sakina, K., Kawazura, K., Morihara, K., and Yajima, H. (1988) Thermolysin-catalyzed synthesis of peptide amides. *Chem. Pharm. Bull. 36*, 4345−4354.
(20) Henriksen, D. B., Rolland, M., Jakobsen, M. H., Buchardt, O., and Breddam, K. (1992) C-terminal amidation of calcitonin by carboxypeptidase Y catalyzed transpeptidation with a photocleavable nucleophile. *Pept. Res. 5*, 321−324.
(21) Cottingham, I. R., Millar, A., Emslie, E., Colman, A., Schnieke, A. E., and McKee, C. (2001) A method for the amidation of recombinant peptides expressed as intein fusion proteins in Escherichia coli. *Nat. Biotechnol. 19*, 974−977.
(22) Muralidharan, V., and Muir, T. W. (2006) Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods 3*, 429−438.
(23) Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U. S. A. 95*, 6705−6710.
(24) Evans, T. C., Jr., Benner, J., and Xu, M. Q. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci. 7*, 2256−2264.
(25) Wood, R. J., Pascoe, D. D., Brown, Z. K., Medlicott, E. M., Kriek, M., Neylon, C., and Roach, P. L. (2004) Optimized conjugation of a fluorescent label to proteins via intein-mediated activation and ligation. *Bioconjugate Chem. 15*, 366−372.
(26) Lue, R. Y., Chen, G. Y., Hu, Y., Zhu, Q., and Yao, S. Q. (2004) Versatile protein biotinylation strategies for potential high-throughput proteomics. *J. Am. Chem. Soc. 126*, 1055−1062.
(27) Cui, C., Zhao, W., Chen, J., Wang, J., and Li, Q. (2006) Elimination of in vivo cleavage between target protein and intein in the intein-mediated protein purification systems. *Protein Express. Purif. 50*, 74−81.
(28) Machova, Z., Muhle, C., Krauss, U., Trehin, R., Koch, A., Merkle, H. P., and Beck-Sickinger, A. G. (2002) Cellular internalization of enhanced green fluorescent protein ligated to a human calcitonin-based carrier peptide. *ChemBioChem 3*, 672−677.
(29) Tan, L. P., Lue, R. Y., Chen, G. Y., and Yao, S. Q. (2004) Improving the intein-mediated, site-specific protein biotinylation strategies both in vitro and in vivo. *Bioorg. Med. Chem. Lett. 14*, 6067−6070.
(30) Zhao, Z., Lu, W., Dun, B., Jin, D., Ping, S., Zhang, W., Chen, M., Xu, M. Q., and Lin, M. (2008) Purification of green fluorescent protein using a two-intein system. *Appl. Microbiol. Biotechnol. 77*, 1175−1180.
(31) Zhao, W., Zhang, Y., Cui, C., Li, Q., and Wang, J. (2008) An efficient on-column expressed protein ligation strategy: application to segmental triple labeling of human apolipoprotein E3. *Protein Sci. 17*, 736−747.
(32) Shaw, A. C., Bang, S., Su, J. Purifying recombinant proteins expressed in eukaryotic or prokaryotic host cells comprises using highly basic protein from thermophilic bacteria as purification tag of recombinant protein in cation-exchange chromatography purification. Patent application WO06108826.
(33) Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci. 4*, 2411−2423.
(34) Cordingley, M. G., Callahan, P. L., Sardana, V. V., Garsky, V. M., and Colonno, R. J. (1990) Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. *J. Biol. Chem. 265*, 9062−9065.
(35) Savitzky, A., and Golay, M. J. E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem. 36*, 1627−1639.
(36) Southworth, M. W., Amaya, K., Evans, T. C., Xu, M. Q., and Perler, F. B. (1999) Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. *Biotechniques 27*, 110−120.
(37) Valiyaveetil, F. I., MacKinnon, R., and Muir, T. W. (2002) Semisynthesis and folding of the potassium channel KcsA. *J. Am. Chem. Soc. 124*, 9113−9120.
(38) Flavell, R. R., Kothari, P., Bar-Dagan, M., Synan, M., Vallabhajosula, S., Friedman, J. M., Muir, T. W., and Ceccarini, G. (2008) Site-specific ^{18}F-labeling of the protein hormone leptin using a general two-step ligation procedure. *J. Am. Chem. Soc. 130*, 9106−9112.
(39) Sydor, J. R., Mariano, M., Sideris, S., and Nock, S. (2002) Establishment of intein-mediated protein ligation under denaturing conditions: C-terminal labeling of a single-chain antibody for biochip screening. *Bioconjugate Chem. 13*, 707−712.
(40) Hirel, P. H., Schmitter, M. J., Dessen, P., Fayat, G., and Blanquet, S. (1989) Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. U. S. A. 86*, 8247−8251.
(41) Perler, F. B. (1998) Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. *Cell 92*, 1−4.
(42) Klabunde, T., Sharma, S., Telenti, A., Jacobs, W. R., Jr., and Sacchettini, J. C. (1998) Crystal structure of GyrA intein from

Mycobacterium xenopi reveals structural basis of protein splicing. *Nat. Struct. Biol. 5*, 31−36.
(43) Rudolph, R., and Lilie, H. (1996) In vitro folding of inclusion body proteins. *FASEB J. 10*, 49−56.
(44) Volkmann, G., Sun, W., and Liu, X. Q. (2009) Controllable protein cleavages through intein fragment complementation. *Protein Sci. 18*, 2393−2402.
(45) Callahan, B. P., Stanger, M., and Belfort, M. (2013) A redox trap to augment the intein toolbox. *Biotechnol. Bioeng. 110*, 1565−1573.
(46) Callahan, B. P., Topilina, N. I., Stanger, M. J., Van, R. P., and Belfort, M. (2011) Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications. *Nat. Struct. Mol Biol. 18*, 630−633.
(47) Deschuyteneer, G., Garcia, S., Michiels, B., Baudoux, B., Degand, H., Morsomme, P., and Soumillion, P. (2010) Intein-mediated cyclization of randomized peptides in the periplasm of Escherichia coli and their extracellular secretion. *ACS Chem. Biol. 5*, 691−700.
(48) Elleuche, S., Doring, K., and Poggeler, S. (2008) Minimization of a eukaryotic mini-intein. *Biochem. Biophys. Res. Commun. 366*, 239−243.
(49) Hiraga, K., Derbyshire, V., Dansereau, J. T., Van, R. P., and Belfort, M. (2005) Minimization and stabilization of the Mycobacterium tuberculosis recA intein. *J. Mol. Biol. 354*, 916−926.
(50) Shingledecker, K., Jiang, S. Q., and Paulus, H. (1998) Molecular dissection of the Mycobacterium tuberculosis RecA intein: design of a minimal intein and of a trans-splicing system involving two intein fragments. *Gene 207*, 187−195.
(51) Chong, S. R., and Xu, M. Q. (1997) Protein splicing of the Saccharomyces cerevisiae VMA intein without the endonuclease motifs. *J. Biol. Chem. 272*, 15587−15590.
(52) Fong, B. A., and Wood, D. W. (2010) Expression and purification of ELP-intein-tagged target proteins in high cell density E. coli fermentation. *Microb. Cell Fact. 9*, 77.
(53) Sharma, S. S., Chong, S., and Harcum, S. W. (2006) Intein-mediated protein purification of fusion proteins expressed under high-cell density conditions in E. coli. *J. Biotechnol. 125*, 48−56.
(54) Dawson, P. E., Muir, T. W., Clark-Lewis, I., and Kent, S. B. (1994) Synthesis of proteins by native chemical ligation. *Science 266*, 776−779.