Organic LETTERS

Letter

pubs.acs.org/OrgLett

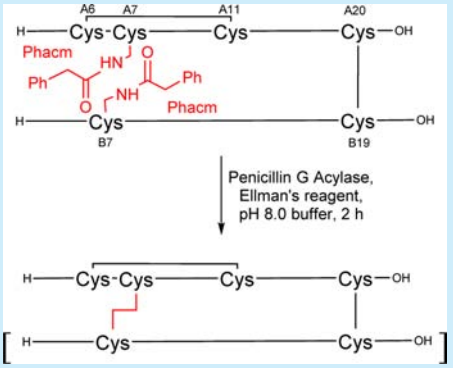
An Iodine-Free and Directed-Disulfide-Bond-Forming Route to Insulin Analogues

Fa Liu,* Qingyuan Liu, and Adam R. Mezo

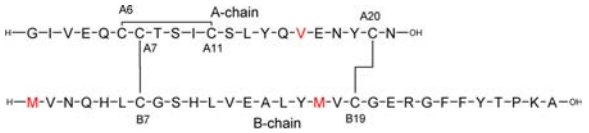
Lilly Research Laboratories, Indianapolis, Indiana 46285, United States

Supporting Information

ABSTRACT: An iodine-free synthetic route to insulin analogues has been established via a directed disulfide bond formation strategy. This method is completely compatible with oxidation-sensitive residues. The key step is constructing the third disulfide bond via a novel procedure involving phenylacetylaminomethyl group (Phacm), immobilized Penicillin G Acylase, and Ellman's reagent. We expect that this method could be broadly utilized for synthesizing insulin-like and other cysteine-rich peptides, in particular, where oxidation-sensitive residues are present in the sequence.



Insulin, relaxin, and other members of the insulin superfamily are biologically essential and structurally unique peptide hormones.[1] They consist of two individual chains, A-chain and B-chain, which are locked together by three disulfide bonds including one intra-A chain and two interchain (Figure 1).



ValA16, MetB1, MetB17-Porcine Insulin **1**

Figure 1. Sequence of porcine insulin analogue **1** (red color highlights the mutations).

Chemical synthesis of these peptides has been of long-standing interest due to the needs from a SAR (structure−activity relationship) perspective as well as the attractive nature of the challenging chemistry itself.[2] One of the main hurdles of these syntheses is the efficient formation of three native disulfide bonds. This has been achieved using two distinct strategies: protein folding and directed disulfide bond formation. The former approach relies on a thermodynamic process, which restricts its applications to analogues with foldability similar to or better than that of their native parents.[3] Conversely, the latter approach utilizes orthogonal protecting groups for each pair of cysteines, which permits unambiguous stepwise formation of each distinct disulfide bond.[4]

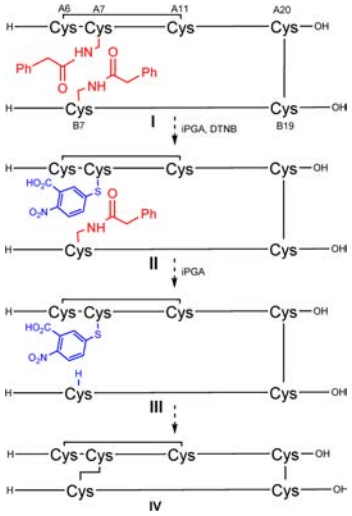
The directed-disulfide-bond-forming strategy can theoretically generate analogues with any possible mutation; therefore it is more suitable for SAR studies, where structural diversity is highly desired. However, in all reported existing routes using this strategy, a harsh I_2 (iodine) oxidation step has been employed once or twice.[4] Although this condition can efficiently remove Acm (acetylaminomethyl) groups and simultaneously form the corresponding disulfide bond, it may also modify oxidation-sensitive residues, such as methionine,[4h,5] tryptophan,[6] and tyrosine.[4a,7] In fact, all three of these residues are frequently found in the native sequences of peptides from the insulin superfamily.[1,8] To avoid possible side reactions induced by I_2 oxidation, we report herein an iodine-free, directed-disulfide-bond-forming procedure to synthesize this class of structurally unique peptides. A porcine insulin analogue **1** with three mutations, Leu to Val at A16, Phe to Met at B1, and Leu to Met at B17, was chosen as an illustrative target molecule (Figure 1).

In a recent communication, we reported that insertion of isoacyl dipeptide segments into the A- and B-chains of human insulin could significantly improve their solubility in aqueous acetonitrile.[4a] This modification enabled their purification via C18 reversed-phase HPLC, which provided clean A- and B-chains suitable for ligation. Upon consideration of how to synthesize insulin analogue **1**, a similar approach could provide A−B dimer **I** containing the A6−A11 and A20−B19 disulfide bonds (Scheme 1). Both of these disulfide bonds could be formed via a 2,2′-dithiobis(5-nitropyridine) (DTNP)-directed thiolysis where I_2-oxidation would not be involved. To construct the last disulfide bond (A7−B7) also under I_2-free conditions, we propose a novel procedure involving both the Phacm (phenylacetylaminomethyl) group[9] and DTNB (Ellman's reagent or 5,5′-dithiobis(2-nitrobenzoic acid)).[10] Phacm groups can be removed by using the enzyme Penicillin G Acylase (PGA) or its immobilized form (iPGA), which

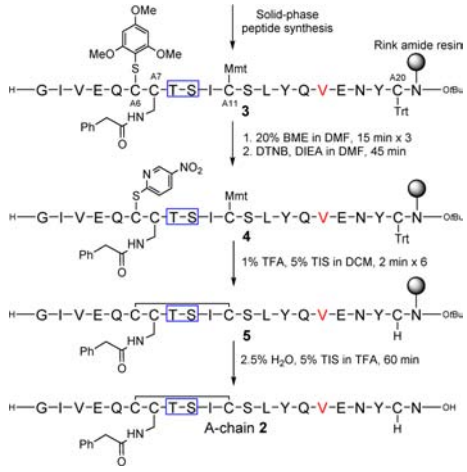
Received: May 1, 2014
Published: May 26, 2014

   dx.doi.org/10.1021/ol501252b | *Org. Lett.* 2014, 16, 3126−3129

Scheme 1. Proposed Phacm/DTNB/iPGA-Mediated Strategy for A7−B7 Disulfide Bond Formation



Scheme 2. Synthesis of (Disulfide A6−A11, Isoacyl A8−A9, CysA7-Phacm) A-Chain 2 (Isoacyl dipeptide segment is indicated by a blue box; red color indicates mutation)



selectively recognizes and cleaves the phenylacetyl moiety, followed by the release of formaldehyde and subsequent deprotection of the sulfhydryl group.[9]

In the new proposed procedure, both CysA7 and CysB7 in A−B dimer **I** are protected by Phacm groups (Scheme 1). When dimer **I** is treated with iPGA,[9] a sulfhydryl group will be liberated upon the removal of one Phacm group (on CysA7 or CysB7; CysA7 is shown as an example in Scheme 1). We hypothesized that, in the presence of DTNB, this newly generated free thiol will be immediately capped and concomitantly activated to provide intermediate **II**. This capping and activation step is expected to take place prior to potential disulfide bond scrambling, driven by the faster reaction kinetics of DTNB vs a Cys-Cys disulfide when exposed to sulfhydryl groups. Subsequent removal of the second Phacm group (on CysB7 or CysA7) results in intermediate **III**, which will quickly undergo A7−B7 disulfide bond formation via a DTNB-directed thiolysis to afford final product **IV**. The possible capping of the sulfhydryl group in intermediate **III** by DTNB, which would form a double-capped side product, is not expected to be of concern since formation of the A7−B7 disulfide bond, an intramolecular process, is expected to be much faster than capping by DTNB, an intermolecular process. However, one possibility which has not been ruled out is that the faster kinetics of this intramolecular reaction may be compromised by an unfavorable conformation of certain analogues, such as ones with less foldability.

To examine if protein foldability could potentially impact the utility of the proposed strategy, a folding-prohibiting mutation, Leu to Val at residue A16, was incorporated into porcine insulin analogue **1**. The placement of Val at A16 was previously reported to block both insulin chain combination and the *in vitro* refolding of proinsulin.[11] Moreover, to confirm its compatibility with oxidation-sensitive residues, two methionines were included in the sequence by mutating Phe to Met at B1 and Leu to Met at B17 (Figure 1).

(Disulfide A6-A11, Isoacyl A8-A9, CysA7-Phacm) A-chain **2** was prepared via a similar scheme as described previously.[4a] Each cysteine was protected with orthogonal protecting groups: CysA6-STmp (2,4,6-trimethoxyphenylthio),[12] CysA7-Phacm, CysA11-Mmt (4-methoxytrityl), and Cys20-Trt (Scheme 2). The synthesis began by loading Asp to Rink amide resin through the β-carboxyl of Fmoc-Asp-O*t*Bu, which would be converted to the native Asn of residue A21 after resin cleavage. The resulting resin-bound peptide **3** was treated with 20% BME (β-mercaptoethanol) to remove the STmp moiety on CysA6 and subsequently treated with DTNP and DIEA (*N*,*N*-diisopropylethylamine) to activate CysA6 as Cys-SNPy [(4-nitro-2-pyridinyl)thiol]. The Mmt group on CysA11 of resin **4** was then removed by 1% TFA, and the unmasked free thiol quickly reacted with CysA6-SNPy to form the intrachain disulfide A6−A11 to provide resin **5**. Peptide cleavage from resin **5** followed by purification by C18 RP-HPLC afforded A-chain **2** with an overall yield of 12% based on the substitution of the starting Rink amide resin.

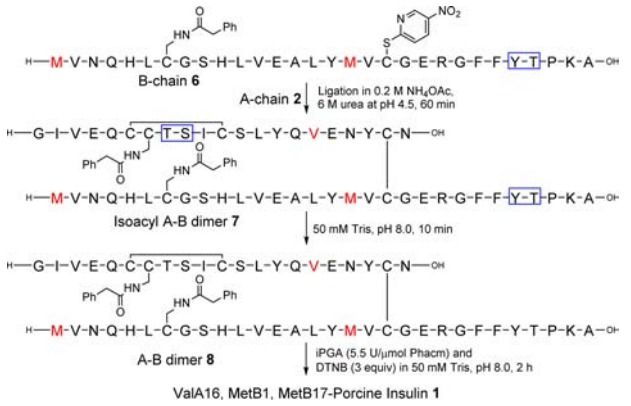
When preparing A-chain **2**, the S*t*Bu (*tert*-butylthiol) group was also tested as a protecting group of CysA6 in resin-bound peptide **3**; however, the complete removal of this S*t*Bu was not achieved even after overnight treatment with 25% BME in DMF. In addition, when preparing resin **4**, it was found that DIEA was needed to drive CysA6 activation (by DTNP) to completion. Both of these observations on CysA6 were different from those found in a previous procedure,[4a] presumably due to the bulky Phacm group attached on the adjacent residue CysA7. The bulkiness and hydrophobicity of the Phacm group may also explain why the yield of A-chain **2** was lower than that of a similar A-chain with CysA7 protected by an Acm group that was reported previously.[4a] The Dpm (Diphenylmethyl) group was also evaluated as a protecting group for CysA20 in resin-bound peptide **3**, since it was reported to be a better Mmt-orthogonal group than Trt when being treated with 1% TFA.[13] However, it was found that the purity of crude A-chain **2** that resulted from CysA20-Dpm was similar to that of CysA20-Trt.

(Isoacyl B26−B27, CysB7-Phacm, CysB19-SNPy) B-chain **6** was assembled on ChemMatrix resin[14] with cysteines protected as CysB7-Phacm and CysB19-Trt. After a simultaneous resin cleavage and CysB19 activation by DTNP,[4a,b,15] a purification step using a C18 RP-HPLC column provided B-chain **6** with a yield of 22%.

The chain ligation of A-chain **2** and B-chain **6** was completed within 60 min in an acetate buffer (pH 4.5) and afforded [Disulfide (A6−A11, A20−B19), Isoacyl (A8−A9, B26−B27)] A−B dimer **7** in 60% yield (Scheme 3). To form the A7−B7

dx.doi.org/10.1021/ol501252b | *Org. Lett.* 2014, 16, 3126−3129

Scheme 3. Synthesis of ValA16, MetB1, MetB17-Porcine Insulin 1 via an I_2-Free Route (Isoacyl dipeptide segment is indicated by a blue box; red color indicates mutation)



disulfide bond, Tris buffer (pH 8.0) was chosen since it provided sufficient solubility (>2 mg/mL) for isoacyl dimer **7**. Upon dissolution in Tris buffer (pH 8.0), the O-to-N shift of isoacyl dimer **7** was completed within 10 min and generated all-amide-backbone [Disulfide (A6−A11, A20−B19)] A−B dimer **8**. Subsequently, iPGA and DTNB were added to facilitate removal of the Phacm groups and the formation of the A7−B7 disulfide bond. 5.5 U of iPGA was used for every μmol of Phacm group,[9] and this ratio completely removed Phacm groups of A−B dimer **8** within 3 h.

To identify the optimal amount of DTNB needed to achieve the best yield of insulin analogue **1**, a variety of molar ratios of DTNB to isoacyl dimer **7** were evaluated, including 0, 0.3, 0.7, 1, 2, 4, and 8. It was found that 2 to 4 equiv of DTNB offered the best crude product profile as judged by analytical HPLC (Figure S4). Both desired product **1** and disulfide-bond-scrambled side products were observed when 0.7 or 1 equiv of DTNB was used. No product was identified when 0 or 0.3 equiv of DTNB was used. The hypothesized intermediate **II** (Scheme 1) was observed when more than 1.0 equiv of DTNB was used; comparatively, the DTNB-double-capped side product was only detectable when more than 4 equiv of DTNB was used. DMSO (10%, v/v) was also tested as a DTNB surrogate for forming disulfide bonds; however, only scrambled side products were observed (Figure S4). We speculate that this is due to the slower reaction kinetics for the DMSO oxidation as compared to competing disulfide bond scrambling at pH 8.0. 2,2′-Dithiodipyridine (DTDP) and DTNP were also attempted as thiol capping and activating agents; however, neither reagent possessed sufficient solubility in Tris buffer (pH 8.0) to enable their evaluation.

Interestingly, des-MetB1 porcine insulin **1** was detected by mass spectrometry as a side product in the amount of 10−20% of the desired product **1** (Figure S3). This may be due to the PGA enzyme, known as a hydrophobic-group-preferred N-terminal hydrolase,[16] possibly recognizing methionine at B1 as a weak substrate. To investigate if pH had an impact on the observed aminopeptidase-like activity, 100 mM phosphate buffer (pH 7.4) was also tested as the solvent; 10% DMSO was added to facilitate full dissolution of isoacyl dimer **7** at a concentration of ∼2 mg/mL. It was found that there was no significant difference in the ratio of Des-MetB1-**1** to **1**; however, the reaction kinetics in phosphate buffer (pH 7.4) was significantly slower than in Tris buffer (pH 8.0) (Figure S4).

Finally, in a moderate scale-up reaction, 24 mg of isoacyl dimer **7** afforded 8 mg of porcine insulin analogue **1** with a yield of 35% without detection of any oxidation of methionine or tyrosine (Scheme 3 and Figure 2). In contrast, direct

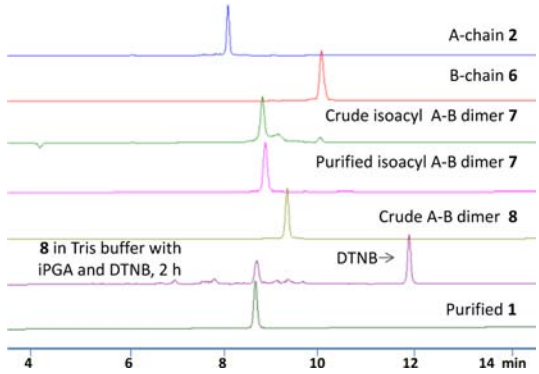


Figure 2. Synthesis of ValA16, MetB1, MetB17-porcine insulin **1** via an I_2-free route (HPLC traces at γ = 220 nm).

treatment of isoacyl dimer **7** or dimer **8** with I_2 only resulted in methionine-oxidized products, a portion of which was also mono- or bis-iodinated (Figures S5 and S6). This confirmed that the reported method is fully compatible with oxidation-sensitive residues. The overall yield of chain ligation and Phacm/DTNB/iPGA-mediated A7−B7 disulfide bond formation was 21%.

The native disulfide bonding pattern of insulin analogue **1** was unambiguously confirmed by Glu-C (*Staphylococcus aureus* Protease V8) digestion of **1**, with comparisons to all three possible disulfide isomers of the A(5−17)−B(1−13) fragment (Table S2 and Figures S7−S9). These three disulfide bond isomers were all generated by independent synthetic routes (Figures S10−S12). The protein structure of synthetic analogue **1** was also confirmed as nearly identical to human insulin by comparing their far UV-CD spectra (Figure S13). Its biological activity was found to be similar to human insulin as determined by their binding affinities to the insulin receptor (Table S3). Both the structural and biological data of synthetic insulin analogue **1** agreed well with previously reported data for the ValA16-insulin analogues.[11]

In conclusion, we have developed a concise I_2-free and directed-disulfide-bond-forming approach to synthesize insulin, and potentially relaxin and other insulin-like peptides. The key reaction involving the formation of the last interchain disulfide bond was accomplished by treating a precursor peptide containing two Cys-Phacm groups with the enzyme iPGA in the presence of Ellman's reagent. In this conversion, the first newly generated free thiol upon removal of one Phacm group by iPGA is immediately capped and activated by Ellman's reagent. Subsequently, unmasking of the second Cys-Phacm moiety concurrently facilitates formation of the last disulfide bond. The successful synthesis of an unfoldable porcine insulin analogue **1** suggests that this Phacm/DTNB/iPGA-mediated disulfide bond formation method is fully directed by chemical reactivity and independent of the protein foldability. This novel strategy adds an additional orthogonal approach to the existing collection of peptide disulfide bond formation methods. Due to the frequent occurrence of oxidation-sensitive residues in the native sequences of cysteine-rich peptides, we envision that this strategy could be broadly utilized for the preparation of these synthetically challenging peptides.

dx.doi.org/10.1021/ol501252b | *Org. Lett.* 2014, 16, 3126−3129

■ ASSOCIATED CONTENT

Supporting Information

The solid-phase peptide synthesis; the screening of optimal molar ratio of DTNB to isoacyl A−B dimer **7**; the disulfide bond mapping of insulin **1**. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: liufx@lilly.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Mr. Steven D. Kahl (Eli Lilly) for conducting the insulin receptor binding assay and Ms. Caitlyn M. Krukau (Eli Lilly) for performing CD analysis.

■ REFERENCES

(1) (a) Litwack, G.,; Ed. *Vitamins and Hormones: Insulin and IGFs*, 1st ed.; Elsevier: 2009. (b) Chan, L. J.; Hossain, M. A.; Samuel, C. S.; Separovic, F.; Wade, J. D. *Protein Pept. Lett.* **2011**, *18*, 220−229. (c) Schwabe, C.; Bullesbach, E. E. *FASEB J.* **1994**, *8*, 1152−1160.

(2) (a) Belgi, A.; Hossain, M. A.; Tregear, G. W.; Wade, J. D. *Immunol., Endocr. Metab. Agents Med. Chem.* **2011**, *11*, 40−47. (b) Wade, J. D.; Lin, F.; Hossain, M. A.; Shabanpoor, F.; Zhang, S.; Tregear, G. W. *Ann. N.Y. Acad. Sci.* **2009**, *1160*, 11−15. (c) Mayer, J. P.; Zhang, F.; DiMarchi, R. D. *Biopolymers* **2007**, *88*, 687−713.

(3) (a) Avital-Shmilovici, M.; Mandal, K.; Gates, Z. P.; Phillips, N. B.; Weiss, M. A.; Kent, S. B. *J. Am. Chem. Soc.* **2013**, *135*, 3173−3185. (b) Sohma, Y.; Hua, Q.-X.; Whittaker, J.; Weiss, M. A.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2010**, *49*, 5489−5493. (c) Sohma, Y.; Kent, S. B. H. *J. Am. Chem. Soc.* **2009**, *131*, 16313−16318. (d) Tofteng, A. P.; Jensen, K. J.; Schaffer, L.; Hoeg-Jensen, T. *ChemBioChem* **2008**, *9*, 2989−2996. (e) Marglin, A.; Merrifield, R. B. *J. Am. Chem. Soc.* **1966**, *88*, 5051−5052. (f) Katsoyannis, P. G.; Tometsko, A.; Zalut, C. *J. Am. Chem. Soc.* **1966**, *88*, 166−167. (g) Kung, Y. T.; Du, Y. C.; Huang, W. T.; Chen, C. C.; Ke, L. T. *Sci. Sin. (Engl. Ed.)* **1965**, *14*, 1710−1716. (h) Meienhofer, J.; Schnabel, E.; Bremer, H.; Brinkhoff, O.; Zabel, R.; Sroka, W.; Klostermayer, H.; Brandenburg, D.; Okuda, T.; Zahn, H. *Z. Naturforsch. B* **1963**, *18*, 1120−1121. (i) Zaykov, A. N.; Mayer, J. P.; Gelfanov, V. M.; DiMarchi, R. D. *ACS Chem. Biol.* **2014**, *9*, 683−691. (j) Kaur, Z. P.; Ochman, A. R.; Mayer, J. P.; Gelfanov, V. M.; DiMarchi, R. D. *ACS Chem. Biol.* **2013**, *8*, 1822−1829. (k) Tang, J.-G.; Wang, Z.-H.; Tregear, G. W.; Wade, J. D. *Biochemistry* **2003**, *42*, 2731−2739. (l) Barlos, K. K.; Gatos, D.; Vasileiou, Z.; Barlos, K. *J. Pept. Sci.* **2010**, *16*, 200−211.

(4) (a) Liu, F.; Luo, E. Y.; Flora, D. B.; Mezo, A. R. *Angew. Chem., Int. Ed.* **2014**, *53*, 3983−3987. (b) Liu, F.; Luo, E. Y.; Flora, D. B.; Mayer, J. P. *Org. Lett.* **2013**, *15*, 960−963. (c) Barlos, K. K.; Gatos, D.; Vasileiou, Z.; Barlos, K. *J. Pept. Sci.* **2010**, *16*, 200−211. (d) Hossain, M. A.; Belgi, A.; Lin, F.; Zhang, S.; Shabanpoor, F.; Chan, L.; Belyea, C.; Truong, H.-T.; Blair, A. R.; Andrikopoulos, S.; Tregear, G. W.; Wade, J. D. *Bioconjugate Chem.* **2009**, *20*, 1390−1396. (e) Samuel, C. S.; Lin, F.; Hossain, M. A.; Zhao, C.; Ferraro, T.; Bathgate, R. A. D.; Tregear, G. W.; Wade, J. D. *Biochemistry* **2007**, *46*, 5374−5381. (f) Bathgate, R. A. D.; Lin, F.; Hanson, N. F.; Otvos, L., Jr.; Guidolin, A.; Giannakis, C.; Bastiras, S.; Layfield, S. L.; Ferraro, T.; Ma, S.; Zhao, C.; Gundlach, A. L.; Samuel, C. S.; Tregear, G. W.; Wade, J. D. *Biochemistry* **2006**, *45*, 1043−1053. (g) Akaji, K.; Fujino, K.; Tatsumi, T.; Kiso, Y. *J. Am. Chem. Soc.* **1993**, *115*, 11384−11392. (h) Bullesbach, E. E.; Schwabe, C. *J. Biol. Chem.* **1991**, *266*, 10754−10761. (i) Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Riniker, B.; Rittel, W. *Helv. Chim. Acta* **1974**, *57*, 2617−2621. (j) Akhter, H. M.; Bathgate, R. A. D.; Kong, C. K.; Shabanpoor, F.; Zhang, S.; Haugaard-Jonsson, L. M.; Rosengren, K. J.; Tregear, G. W.; Wade, J. D. *ChemBioChem* **2008**, *9*, 1816−1822.

(5) (a) Chipiso, K.; Mbiya, W.; Morakinyo, M. K.; Simoyi, R. H. *Aust. J. Chem.* **2014**, *67*, 626−635. (b) Young, P. R.; Hsieh, L. S. *J. Org. Chem.* **1982**, *47*, 1419−1423.

(6) (a) Sieber, P.; Kamber, B.; Riniker, B.; Rittel, W. *Helv. Chim. Acta* **1980**, *63*, 2358−2363. (b) Hartdegen, F. J.; Rupley, J. A. *J. Am. Chem. Soc.* **1967**, *89*, 1743−1745.

(7) (a) Morikawa, S.; Kanatani, A.; Kobayashi, R.; Yoshimoto, T.; Tsuru, D. *Agric. Biol. Chem.* **1991**, *55*, 2751−2756. (b) Richards, H. R.; Speakman, J. B. *Nature* **1953**, *171*, 751.

(8) Conlon, J. M. *Peptides* **2001**, *22*, 1183−1193.

(9) (a) Gongora-Benitez, M.; Basso, A.; Bruckdorfer, T.; Royo, M.; Tulla-Puche, J.; Albericio, F. *Chem.—Eur. J.* **2012**, *18*, 16166−16176. (b) Royo, M.; Alsina, J.; Giralt, E.; Slomcyznska, U.; Albericio, F. *J. Chem. Soc., Perkin Trans. 1* **1995**, 1095−1102.

(10) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70−77.

(11) Liu, M.; Wan, Z.-L.; Chu, Y.-C.; Aladdin, H.; Klaproth, B.; Choquette, M.; Hua, Q.-X.; Mackin, R. B.; Rao, J. S.; De Meyts, P.; Katsoyannis, P. G.; Arvan, P.; Weiss, M. A. *J. Biol. Chem.* **2009**, *284*, 35259−35272.

(12) Postma, T. M.; Giraud, M.; Albericio, F. *Org. Lett.* **2012**, *14*, 5468−5471.

(13) Gongora-Benitez, M.; Mendive-Tapia, L.; Ramos-Tomillero, I.; Breman, A. C.; Tulla-Puche, J.; Albericio, F. *Org. Lett.* **2012**, *14*, 5472−5475.

(14) (a) de la Torre, B. G.; Jakab, A.; Andreu, D. *Int. J. Pept. Res. Ther.* **2007**, *13*, 265−270. (b) Garcia-Martin, F.; Quintanar-Audelo, M.; Garcia-Ramos, Y.; Cruz, L. J.; Gravel, C.; Furic, R.; Cote, S.; Tulla-Puche, J.; Albericio, F. *J. Comb. Chem.* **2006**, *8*, 213−220.

(15) Ghosh, A. K.; Fan, E. *Tetrahedron Lett.* **2000**, *41*, 165−168.

(16) (a) Oh, B.; Kim, K.; Park, J.; Yoon, J.; Han, D.; Kim, Y. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 486−492. (b) van der Mey, M.; de Vroom, E. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 345−348.