ORGANIC LETTERS

2007
Vol. 9, No. 26
5497−5500

Synthesis of Photoactive Analogues of a Cystine Knot Trypsin Inhibitor Protein

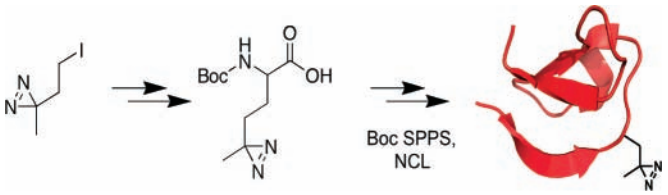
Thomas Durek,[†,‡] Junliang Zhang,[§] Chuan He,[§] and Stephen B.H. Kent*[,†,‡,§]

Institute for Biophysical Dynamics, Department of Chemistry, and Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois 60637

skent@uchicago.edu

Received October 9, 2007

ABSTRACT



We describe the preparation of a recently described photoactive amino acid analogue (photoMethionine) by two novel synthetic routes, one of which is flexible and enantiospecific, and the site-specific chemical incorporation of photoMethionine into a defined and functionally active protein using a combination of solid-phase peptide synthesis and modern chemical ligation methodology. Site-specific labeling of proteins with this amino acid analogue through chemical synthesis provides valuable probes for photoaffinity cross-linking studies.

Chemical cross-linking,[1] i.e., the formation of a covalent bond between spatially interacting protein molecules, has emerged as an important biochemical tool in the proteomics era. The cross-linking approach enables identification of interacting proteins under in vivo conditions and in conjunction with state of the art mass spectrometry is a useful way to elucidate the complex interaction networks that proteins are known to participate in. In contrast to the classical intrinsically reactive cross-linkers such as *N*-hydroxysuccinimide esters or maleimides, photoactivatable cross-linkers require UV illumination to initiate cross-link formation and offer the unique advantage of permitting temporal and spatial control over reactivity. Most of the photoactivatable reagents reported to date are based on arylazides, aryldiazirines, and arylketones, which upon irradiation generate highly reactive nitrene, carbene, or triplet biradical species, respectively.[2] These functional groups have been incorporated in the past into a number of biologically relevant molecules including peptides, DNA, lipids, and small molecule metabolites.[3] However, in the case of peptides and proteins, their bulky aryl-structure and limited relatedness to most naturally occurring amino acids can be disadvantageous, especially when sterically demanding protein−protein interaction interfaces are the subject of investigation.

Recently, novel alkyldiazirine amino acid analogues that closely resemble the aliphatic amino acids Leu, Ile, and Met have been described.[4] These amino acid analogues were shown to be incorporated into proteins by the biosynthetic machinery of mammalian cells, which underscores their structural similarity to the naturally occurring amino acids. In this paper we report on the efficient site-specific incorporation of one of these analogues ("photoMethionine" (phMet), Figure 1A) into a small protein using modern chemical synthesis methods.

[†] Institute for Biophysical Dynamics.
[‡] Department of Biochemistry and Molecular Biology.
[§] Department of Chemistry.

(1) Brunner, J. *Annu. Rev. Biochem.* **1993**, *62*, 483.

(2) (a) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25*, *663*. (b) Sadakane, Y.; Hatanaka, Y. *Anal. Sci.* **2006**, *22*, 209. (c) Chan, E. W. S.; Chattopadhaya, S.; Panicker, R. C.; Huang, X.; Yao, S. Q. *J. Am. Chem. Soc.* **2004**, *126*, 14435. (d) Sieber, S. A.; Niessen, S.; Hoover, H. S.; Cravatt, B. F. *Nat. Chem. Biol.* **2006**, *2*, 274.

(3) (a) Nassal, M. *J. Am. Chem. Soc.* **1984**, *106*, 7540. (b) Kauer, J. C.; Ericksonviitanen, S.; Wolfe, H. R.; Degrado, W. F. *J. Biol. Chem.* **1986**, *261*, *695*.

(4) Suchanek, M.; Radzikowska, A.; Thiele, C. *Nat. Methods* **2005**, *2*, 261.

10.1021/ol702461z CCC: $37.00
Published on Web 11/28/2007

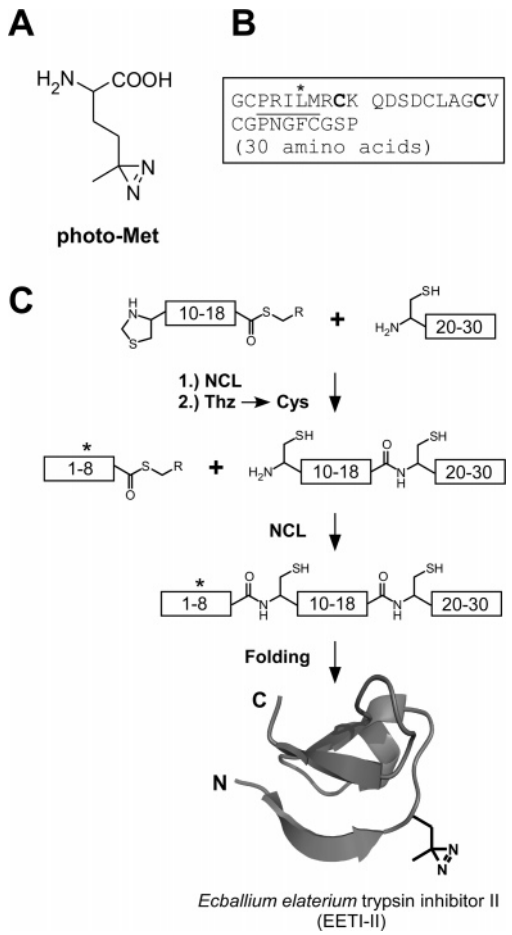


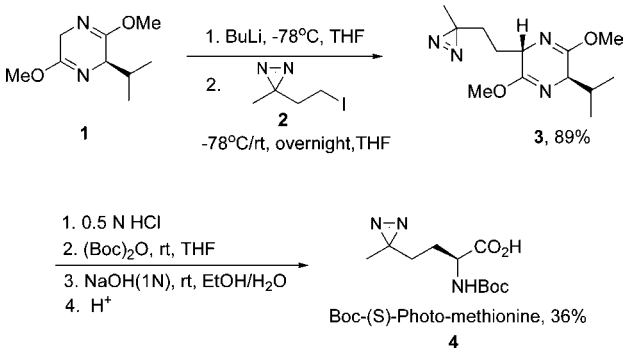
Figure 1. (A) PhotoMethionine and (B) amino acid sequence of EETI-II. Cysteines used as strategic sites for native chemical ligation of unprotected peptide segments are labeled in bold. The loop interacting in a substrate-like manner with the active site of trypsin is underlined. The site of incorporation of photoMethionine is marked with an asterisk (Leu[6]). (C) Synthetic scheme for the assembly of EETI-II from three unprotected peptide segments by sequential native chemical ligation.

As a target for incorporation of phMet we chose the small squash protein EETI-II (*Ecballium elaterium* trypsin inhibitor type II), which interacts with trypsin and other related serine proteases with high affinity (Kd in the range of pM-nM).[5] EETI-II is a so-called "microprotein",[6] comprising a 30 amino acid residue polypeptide chain that contains six cysteines that form three disulfides in the native protein molecule. The inhibitor binds its target enzyme in a substrate-like manner via an N-terminal loop comprising the amino acids −Pro[3]-Arg-Ile-Leu-Met[7]− forming a tight and highly complementary interaction interface (see the Supporting Information).[7] A crystal X-ray structure of the EETI-II: trypsin complex has been determined recently.[7b] We envisioned incorporation of phMet in place of Leu[6] in the binding loop because its side chain appears to be in close proximity to a number of functional groups near the active site of trypsin. The amino acid sequence of EETI-II is given in Figure 1B. Synthesis of EETI-II wild-type and phMet modified analogues was achieved by native chemical ligation[8a] of three unprotected peptide segments, starting at the C-terminus and extending the polypeptide chain toward the N-terminus (Figure 1C).[8b] Although polypeptide chains of this size can be made by optimized stepwise SPPS, we believe that the modular synthesis demonstrated in this work is particularly useful for the efficient preparation of analogues. The ligation-based synthetic strategy also serves as a prototype for the preparation of larger protein molecules containing photoactive moieties.

Racemic Boc-phMet[4] was synthesized in good yield by a novel route involving alkylation of a glycine Schiff base with 1-iodo-3,3′-azibutane; details can be found in the Supporting Information.

The racemic phMet was incorporated into the EETI-II segment Gly[1]-Arg[8]-$^{\alpha}$thioester in place of Leu[6] by optimized stepwise Boc SPPS chemistry.[9] Side chain deprotection and concomitant cleavage from the resin by treatment with anhydrous HF gave the two diastereomeric peptide-$^{\alpha}$thioesters, which were separated by RP-HPLC and isolated. Stereospecific synthesis of Boc-L-phMet was realized according to Schöllkopf's bislactim ether method, as described in the Supporting Information (Scheme 1). This enabled us to determine the absolute configuration of phMet in both diastereomeric peptides (see the Supporting Information). We found the diazirine-modified amino acid to be stable toward the repeated cycles of TFA treatment and HBTU/DIEA activation used in chain assembly by Boc chemistry SPPS; however, final HF treatment resulted in partial (∼30%) loss

Scheme 1. Synthesis of Boc-L-phMet



(5) Le Nguyen, D.; Nalis, D.; Castro, B. *Int. J. Pept. Protein Res.* **1989**, *34*, 492.
(6) Craik, D. J.; Clark, R. J.; Daly, N. L. *Exp. Op. Investig. Drugs* **2007**, *16*, 595.
(7) (a) Le Nguyen, D.; Heitz, A.; Chiche, L.; Castro, B.; Boigegrain, R. A.; Favel, A.; Coletti-Previero, M. A. *Biochimie* **1990**, *72*, 431. (b) Kratzner, R.; Debreczeni, J. E.; Pape, T.; Schneider, T. R.; Wentzel, A.; Kolmar, H.; Sheldrick, G. M.; Uson, I. *Acta Crystallogr. D: Biol. Crystallogr.* **2005**, *61*, 1255.
(8) (a) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776. (b) Bang, D.; Kent, S. B. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534.
(9) (a) Schnoelzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. *Int. J. Pept. Protein Res.* **1992**, *40*, 180. (b) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068.

of the diazirine functionality by addition of HF. These byproducts were readily removed by HPLC purification.

Peptide segments EETI-II(Thz^{9}-Gly^{18})-$^{\alpha}$thioester and EETI-II(Cys^{19}-Pro^{30}) were synthesized by Boc SPPS according to published procedures.[9] 1,3-Thiazolidine-4-carboxylic acid

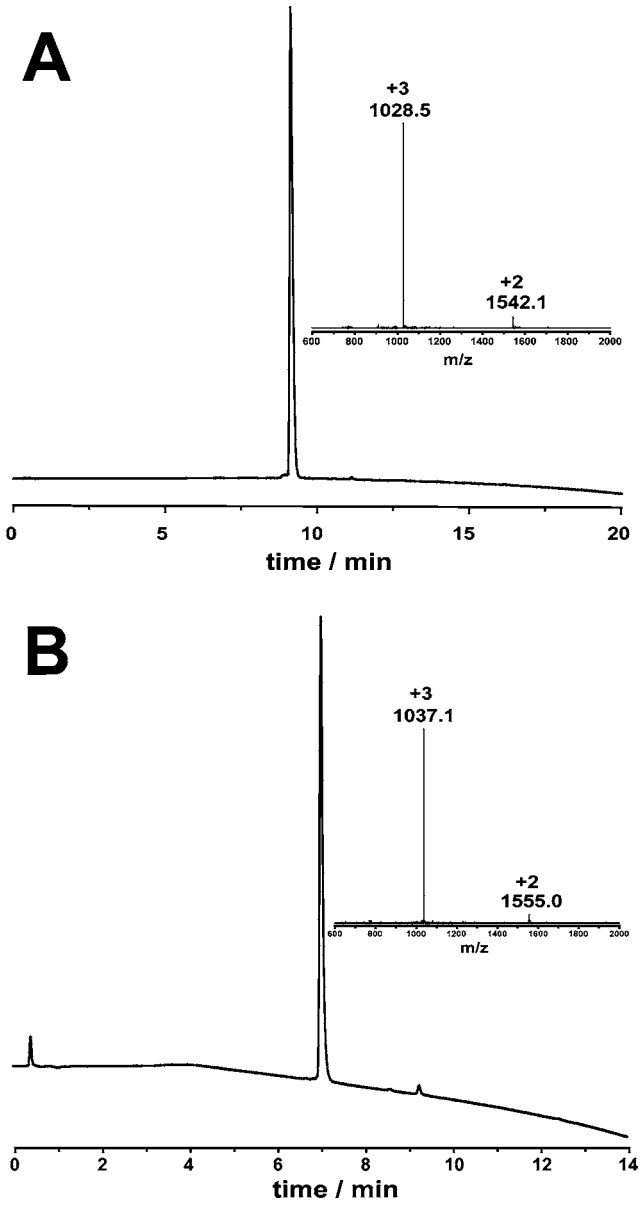


Figure 2. LC-MS analysis of the purified EETI-II proteins. The wild-type protein (A) and the Leu6L-phMet analogue (B) were of high purity [insets: ESI-MS of the major peaks; the observed mass for WT EETI-II was 3082.7 ± 0.1 Da; calculated mass 3082.6 Da (average isotopes); the observed mass for [Leu6L-phMet]EETI-II was 3108.4 ± 0.2 Da; calculated mass 3108.7 Da (average isotopes)].

(Thz) was used as a protected form of Cys^{9} in order to prevent side reactions during the first ligation reaction.[8b,10] The peptides EETI-II(Thz^{9}-Gly^{18})-$^{\alpha}$thioester and EETI-II(Cys^{19}-Pro^{30}) were covalently joined by native chemical ligation in 6 M guanidine hydrochloride, 0.2 M sodium phosphate, 30 mM mercaptophenylacetic acid, 20 mM tris(2-carboxyethyl)phosphine hydrochloride, pH 6.8 (see the Supporting Information for details), after which Thz^{9} was converted to Cys^{9} by treatment with methoxylamine hydrochloride at pH 4.0. Peptide EETI-II(Cys^{9}-Pro^{30}) was isolated in good yield (65%). Full-length wild-type and photoactivatable analogues of EETI-II were finally obtained by reacting unmodified and modified EETI-II(Gly^{1}-Arg^{8})-$^{\alpha}$thioester peptides with EETI-II(Cys^{9}-Pro^{30}) in separate native chemical ligation reactions using the conditions described above. Native chemical ligation of the diazirine-modified peptides proceeded smoothly, without accumulation of major side products (Supporting Information). We conclude that the diazirine functionality is fully compatible with native chemical ligation reaction conditions, including a reducing environment as well as the presence of thiol nucleophiles. Folding of the full-length polypeptides with concomitant disulfide formation was initiated in all cases simply by diluting the crude reaction mixture after the final ligation into buffer containing a redox system consisting of 10 mM oxidized glutathione and 2 mM reduced glutathione. Folding was complete within 2 h as determined in LC analysis by the appearance of a dominant product eluting at earlier retention time. Mass spectrometric analysis revealed the loss of 6 Da upon folding, which is in agreement with the formation of the three disulfides that are an integral part of the cystine-knot fold of EETI-II.[5] Three synthetic proteins, the two diastereomeric analogues ([Leu6L-phMet]EETI-II and [Leu6D-phMet]EETI-II) and the wild-type EETI-II, were isolated in good yield by reverse-phase HPLC (Figure 2).

We next determined the ability of the synthesized protein analogues to bind trypsin and inhibit its proteolytic activity. To this end, trypsin was assayed using N^{α}-benzoyl-Arg-(4-nitro)anilide as a chromogenic substrate in the presence of increasing concentrations of inhibitors (Figure 3).[11] [Leu6L-phMet]EETI-II displayed strong inhibitory activity against trypsin that was essentially indistinguishable from wild-type EETI-II (Figure 3). Although precise determination of the (very low) corresponding dissociation constants (K_d) describing this interaction is generally difficult under these conditions, inspection of the fitted data seems to suggest that they are in the lower nanomolar range (data fitting revealed a K_d of 0.7 ± 1.7 nM for both wild-type protein and the Leu6L-phMet analogue). In contrast [Leu6D-phMet]EETI-II showed only weak activity against trypsin, with a determined K_d of about 800 ± 120 nM. These results indicate that substitution of Leu^{6} by L-phMet^{6} in EETI-II does not significantly impair its ability to interact with and inhibit trypsin. Replacement of Leu^{6} by D-phMet,^{6} on the other hand, is much less tolerated, which underscores the high selectivity constraints imposed by trypsin's active site on EETI-II binding.

We next examined photo-cross-linking of [Leu6L-phMet]-EETI-II to trypsin upon illumination with near-UV light. The

(10) Villain, M.; Vizzavona, J.; Rose, K. *Chem. Biol.* **2001**, *8*, 673.

(11) Erlanger, B. F.; Cohen, W.; Kokowsky, N. *Arch. Biochem. Biophys.* **1961**, *95*, 271.

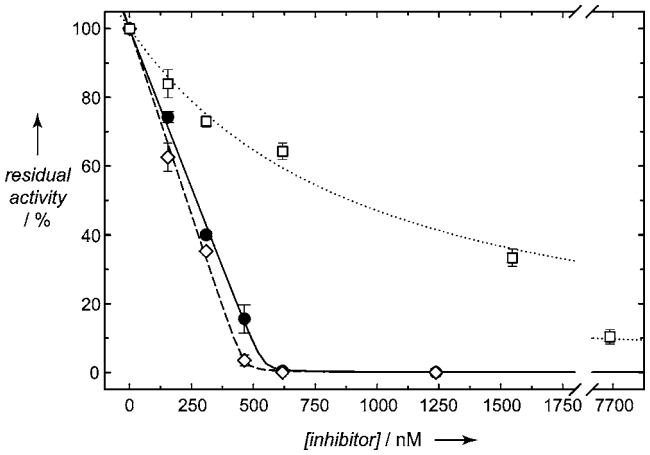


Figure 3. Inhibition of trypsin by wild-type EETI-II (●) and photoaffinity labeled analogues (◇ for Leu6L-phMet; □ for Leu6D-phMet). Data were fitted to the solution of a quadratic equation: (—) WT; (•••) Leu6D-phMet; and (---) Leu6L-phMet.

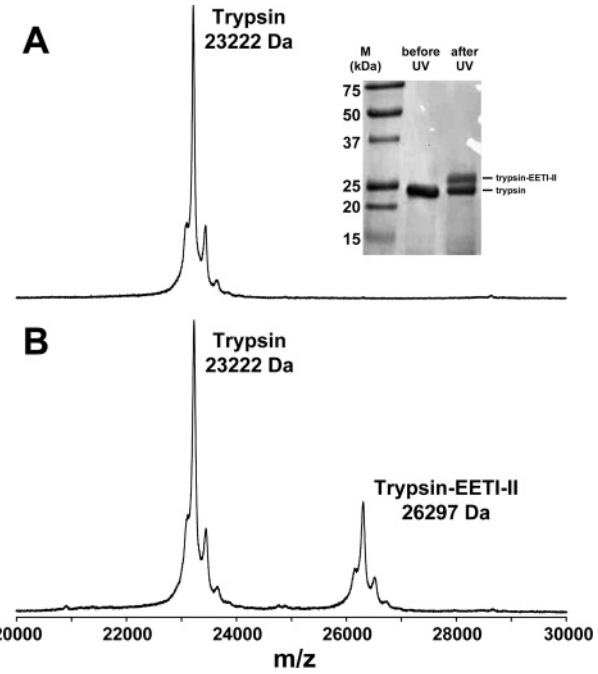


Figure 4. Photo-cross-linking of [Leu6L-phMet]EETI-II and trypsin. The two proteins were mixed in a 1:1 ratio at a concentration of ~100 μM. (A) MALDI-TOF MS spectrum before illumation. (B) MALDI-TOF MS spectrum after illumination of the sample with a 450 W mercury lamp for 5 min (λ > 300 nm). [Inset (upper panel): SDS-PAGE analysis of the cross-linking reaction.]

two proteins were mixed in a 1:1 ratio at a concentration of about 100 μM in aqueous buffer and the sample was illuminated with a 450 W mercury lamp for 5 min. MALDI-TOF analysis of the cross-linking reaction before and after UV illumination (Figure 4) revealed the appearance of a product with a mass 3.1 kDa higher than the calculated molecular mass of trypsin (23.2 kDa), reflecting formation of a chemical cross-link between trypsin and EETI-II (3.1 kDa).

We were unable to detect any species with higher molecular weight, which further underscores the high specificity of the cross-linking reaction. Photo-cross-linking efficiency was about 40% judged by SDS-PAGE analysis (Figure 4 (inset)); the efficiency of cross-linking was not increased by illumination times of up to 30 min.

In summary we have reported the efficient synthesis of Boc-protected analogues of photoMethionine by two new synthetic routes giving access to both D and L stereoisomers. The amino acid analogue is compatible with Boc chemistry solid-phase peptide synthesis and with the native chemical ligation of unprotected peptide segments. These methods will enable the preparation of photoactivatable variants of peptides and proteins and their use in photochemical "fishing" experiments to help dissect protein−protein and peptide−protein interactions in living cells.[12]

Acknowledgment. This research was supported by the Office of Science (BER), U.S. Department of Energy, Grant No. DE-FG02-07ER64501 (to S.B.H.K.).

Supporting Information Available: Experimental procedures and spectroscopic data for described compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

OL702461Z

(12) During preparation of this manuscript, Muir and co-workers reported on the semisynthesis of a phMet-containing analogue of the MH2 domain of the SMAD2 protein by expressed protein ligation: Vila-Perello, M.; Pratt, M. R.; Tulin, F.; Muir, T. W. *J. Am. Chem. Soc.* **2007**, *129*, 8068.