Bioorganic & Medicinal Chemistry Letters 18 (2008) 4550–4553

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

N-terminal labeling of proteins by the Pictet–Spengler reaction

Tsubasa Sasaki, Koichiro Kodama, Hiroaki Suzuki, Seketsu Fukuzawa *, Kazuo Tachibana *

Department of Chemistry, School of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

ARTICLE INFO

Article history:
Received 18 June 2008
Revised 9 July 2008
Accepted 10 July 2008
Available online 15 July 2008

Keywords:
Pictet–Spengler reaction
Myoglobin
N-terminal
Biotin
Transamination reaction
Keto aldehyde

ABSTRACT

The Pictet–Spengler reaction was applied to the N-terminal labeling of horse heart myoglobin. This was performed in the following two steps: (1) conversion of the N-terminal glycine residue to an α-keto aldehyde by a transamination reaction and (2) condensation of the resulting activated myoglobin with tryptamine analogues by the Pictet–Spengler reaction. Ultraviolet (UV)/visible (vis) absorption and circular dichroism (CD) spectral data revealed that the tertiary structure of myoglobin was not altered by the Pictet–Spengler reaction.

Functionalization of proteins is increasingly important for biology studies.[1,2] A variety of chemistries have been used for protein functionalization, but some characteristics of the reactions often prevent their use in a wider variety of applications. For instance, the use of copper ions for Huisugen cycloaddition[3] and transition-metal catalyzed cross-coupling reactions[4] may cause oxidative modifications of proteins.[5] The diene group, a chemical tag for the Diels–Alder reaction, forms a sulfide bond with the thiols of cysteine residues.[6] Condensation products of an aldehyde (e.g., oxime[7,8] and hydrazone[9]) suffer from hydrolysis in aqueous solution. Therefore, there is a continuing demand for alternatives to the existing reactions.

The Pictet–Spengler reaction is a cyclization reaction, by which 3-(2-aminoethyl)indole or β-arylethylamine undergoes Mannich-type ring formation with an aldehyde (Scheme S1). A stable C–C bond can be formed on amino acid residues, but the typical Pictet–Spengler reaction requires strong Bronsted acids.[10] For instance, Li et al. used trifluoroacetic acid (1%, v/v in water) as the acid catalyst for chemoselective ligation of unprotected peptides.[11] Their conditions cannot be applied to protein modification, since proteins are denatured under acidic conditions.

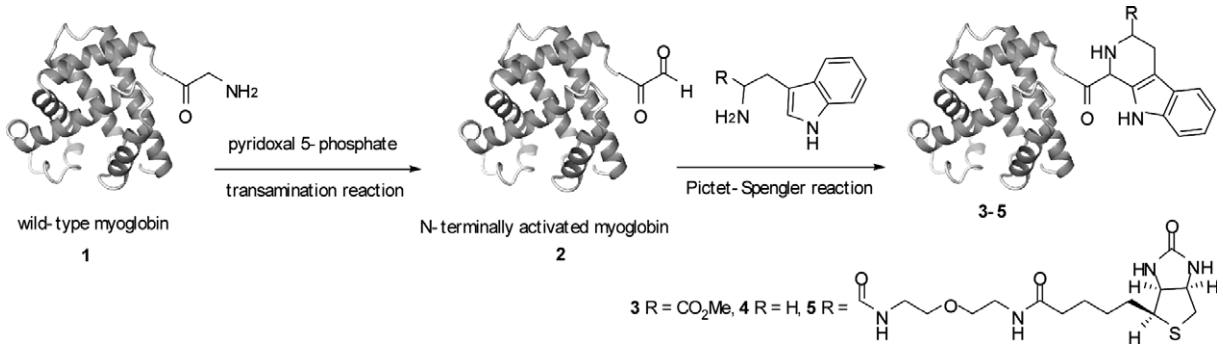
Several groups have developed methods for introducing an aldehyde into targeted positions of proteins.[7,8] For example, Gilmore et al. converted an N-terminal glycine residue to an aldehyde by a transamination reaction: the α-amino group was condensed with pyridoxal 5-phosphate to form an iminium cation, which was tautomerized and hydrolyzed to yield α-keto aldehydes. The transamination reaction proceeded without affecting the ε-amino groups of the lysine residues.[7]

We have applied the Pictet–Spengler reaction to the N-terminal labeling of horse heart myoglobin, a 153 amino acid residue heme-binding protein with an N-terminal glycine (Scheme 1). Tryptophan methyl ester and tryptamine were employed as the coupling partners of **2**. The N-terminally activated myoglobin **2** was incubated with either tryptophan methyl ester (50 mM) or tryptamine (50 mM) in phosphate buffer (pH 6.5) at 37 °C for 18 h. The resulting products were desalted and subjected to MALDI–MS. The singly charged ion peaks at *m/z* 16952.0 and 17151.4 corresponded to the theoretical mass values for **2** (Δ0.5 Da) and **3** (Δ−0.1 Da), respectively (Fig. 1A). In addition, the singly charged ion peaks at *m/z* 16952.6 and 17094.0 agreed with the theoretical mass values for **2** (Δ1.1 Da) and the modified myoglobin **4** (Δ−1.5 Da), respectively (Fig. 1B). The results showed that the Pictet–Spengler reaction proceeded successfully with myoglobin. Reaction product **3** was then analyzed by UV/vis absorption spectroscopy. The UV/vis spectra of the wild-type myoglobin **1** and the modified myoglobin **3** overlapped each other, except for the absorption peak at 280 nm (Fig. 2A). The peak absorption wavelength of tetrahydro-β-carboline is around 280 nm,[12] suggesting that reaction product **3** contained a 1,2,3,4-tetrahydro-β-carboline group.

Reaction product **3** was digested with *Achromobacter* protease I (Lys-C), and the Lys-C liberated peptides were subjected to MALDI–MS to detect the N-terminal peptide (K1-peptide; the region from residue 1 to K16) of **3**. The singly charged ion peaks at *m/z* 1817.5 and 2017.1 agreed with the calculated mass values for the K1-peptides of **2** (Δ1.6 Da) and **3** (Δ2.1 Da), respectively (Fig. S1).

* Corresponding authors. Tel./fax: +81 3 5841 8059.
E-mail addresses: seketsuf@chem.s.u-tokyo.ac.jp (S. Fukuzawa), ktachi@chem.s.u-tokyo.ac.jp (K. Tachibana).

0960-894X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2008.07.033



Scheme 1. N-terminal labeling of myoglobin by the Pictet–Spengler reaction.

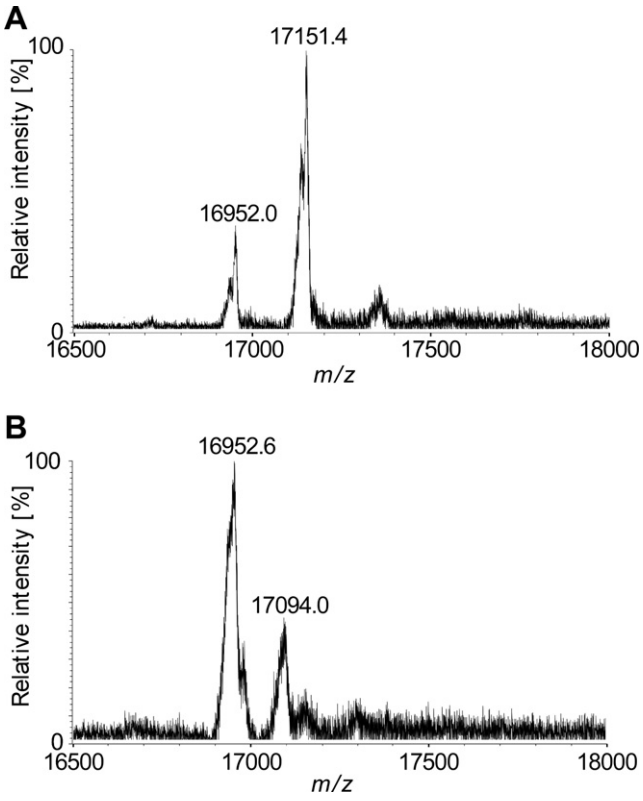


Figure 1. MALDI–MS spectra of the Pictet–Spengler reaction products condensed with (A) tryptamine and (B) tryptophan methyl ester. The calculated average mass for **2**, the modified myoglobin **3,** and **4** is 16951.5, 17151.5, and 17095.5, respectively.

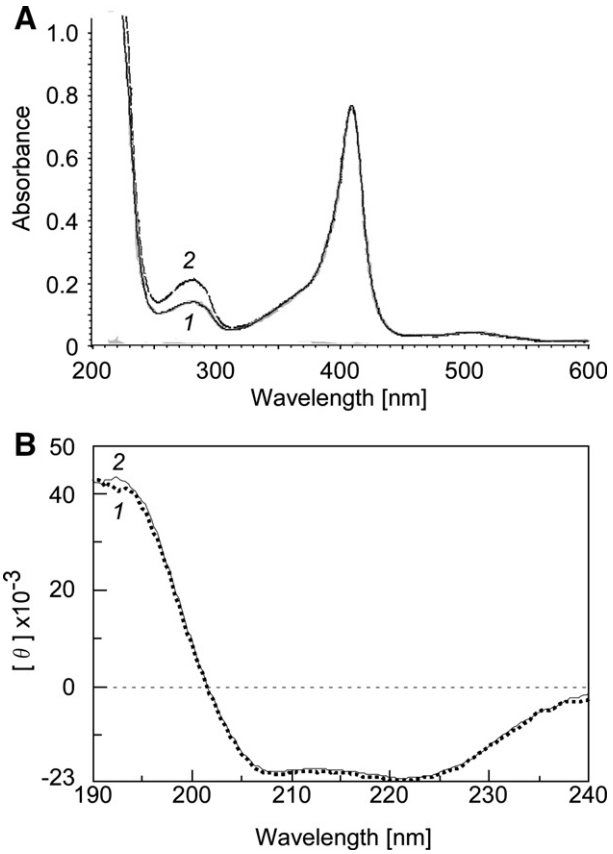


Figure 2. (A) UV/vis and (B) CD spectra of the wild-type myoglobin **1** (*1*) and the modified myoglobin **3** (*2*). The absorbance at around 410 nm is due to extinction coefficient of heme.

This result suggested that the N-terminal residue was modified by the Pictet–Spengler reaction. The K1-peptide of **3** was further purified by ODS-HPLC, using aqueous acetonitrile with 0.1% (v/v) trifluoroacetic acid, and was analyzed by MALDI–MS and MALDI postsource decay (PSD) fragmentation[13] to confirm the amino acid sequence. The singly charged ion peak at *m/z* 2010.9 (Δ0.9 Da) agreed with the calculated mass values for the K1-peptide containing a methyl β-carboline-3-carboxylate residue **6** (Figs. 3 and S2), suggesting the aromatization of the 1,2,3,4-tetrahydro-β-carboline-3-carboxylate residue. The ion peaks observed in the MALDI–PSD spectrum from the precursor ion of *m/z* 2011.8 were assigned along the amino acid sequence of **6** (Fig. 3). These results supported the conclusion that the N-terminal α-keto aldehyde was functionalized by the Pictet–Spengler reaction.

The 1,2,3,4-tetrahydro-β-carboline residue was not aromatized during the reaction. This view was supported by the following two findings: (1) the observed mass values for the K1-peptide of **3** agreed with the theoretical one when the K1-peptide was not purified by ODS-HPLC (Fig. S1), and (2) the UV absorption peak of **3** was observed at around 280 nm (Fig. 2A). Aromatization of tetrahydro-β-carboline shifts the UV absorption peak to higher values of the resonant wavelength at around pH 6.5.[14] The 1,2,3,4-tetrahydro-β-carboline residue might be spontaneously oxidized during HPLC purification or laser desorption ionization.

The modified myoglobin **3** was analyzed by circular dichroism (CD) spectroscopy. The CD spectral profiles of **3** and the wild-type myoglobin **1** were superimposable (Fig. 2B), which showed that the tertiary structure of the protein maintained in its native conformation throughout the reaction.[15] On the other hand, the incubation of the wild-type myoglobin **1** at 70 °C was accompanied by reductions in the intensities of the negative bands at 208 and 222 nm

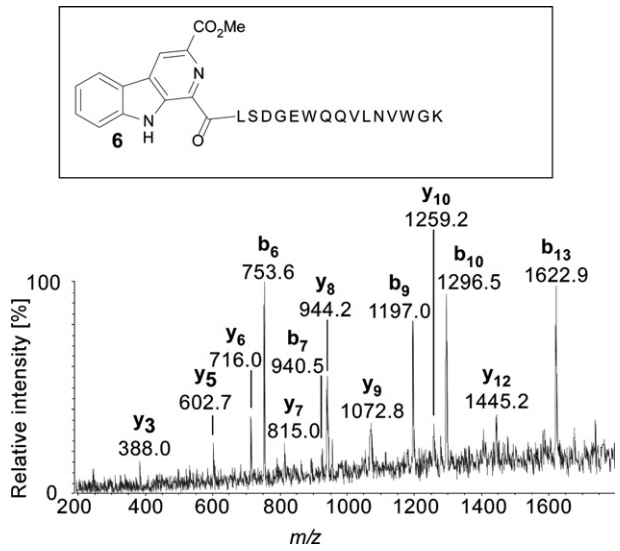
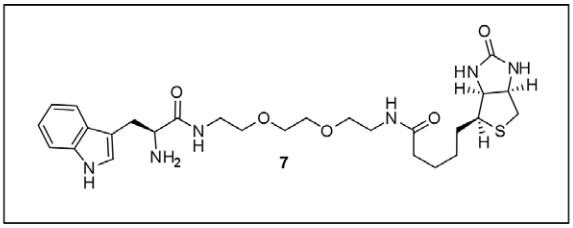
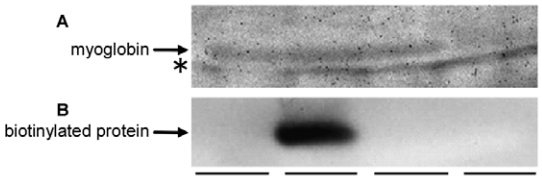


Figure 3. MALDI-PSD fragment spectrum of the N-terminal peptide (K1) from **3**. Inset shows the chemical structure of the oxidized K1-peptide **6**. The b (N-terminal) and y (C-terminal) fragment ions are labeled in the spectrum. The monoisotopic mass calculated from the sequence of **6** is given in Table S1.

(Fig. S3B), suggesting that the proportion of α-helix decreased, due to heat denaturation. In addition, UV/vis absorption spectroscopy indicated that the incubation of the wild-type myoglobin **1** at 70 °C reduced the intensity of the absorption peak at 410 nm against that at 280 nm (A_{410}/A_{280}) (Fig. S3A); the relative amount of the apo form of myoglobin **1** increased by heat denaturation.

Biotinylation of myoglobin was performed as an application of the Pictet–Spengler reaction to the site-specific functionalization of proteins. The biotinylation unit **7** (10 mM), consisting of tryptophan, an ethylene oxide linker, and biotin (Schemes 4 and S2), and the protein substrate **2** were mixed in phosphate buffer (pH 6.5) and incubated at 37 °C for 18 h. The resulting product **5** was subjected to SDS–PAGE and stained with fluorescent dyes to detect myoglobin (Fig. 4A). The proteins fractionated on the SDS–PAGE gel were also analyzed by Western blotting to detect biotin (Fig. 4B). The results indicated that myoglobin **2** was functionalized





lane:	1	2	3	4
N-terminally activated myoglobin **2** [μM]	60	60	60	-
modification unit **7** [mM]	10	10	-	10
methoxy amine [mM]	160	-	-	-

Figure 4. (A) Fluorescence detection of the protein moiety by staining with SYPRO Tangerine, and (B) chemiluminescence detection of the biotin moiety with streptavin-HRP. An asterisk indicates the dye front in a 15% SDS–PAGE gel. Inset shows the chemical structure of modification unit **7**.

with biotinylated tryptophan **7** (Fig. 4B, lane 2). In addition, no biotinylation was detected when myoglobin **2** was incubated with **7** in the presence of methoxy amine (Fig. 4B, lane 1), which competitively reacted with the aldehyde group on **2**. However, undesired biotinylation was detected after biotinylated tryptophan **7** was incubated with the wild-type myoglobin **1**, instead of using the N-terminal activated myoglobin **2** as a substrate (Fig. S4). The myoglobin **1** isolated from horse heart might contain additional aldehyde groups besides that at the N-terminus (e.g., lysine, arginine, and/or proline residues),[16] which underwent the Pictet–Spengler reaction to yield the non-specifically biotinylated myoglobin. Oxidized proteins containing aldehyde groups are known to accumulate in aging cells. However, the amount of oxidized myoglobin in the reaction mixture is negligible; the MALDI–MS spectrum indicated that the molecular ion for the myoglobin containing two tetrahydro-β-carboline groups (17.4 kDa.) was weakly observed, as compared with that for the modified myoglobin **3** (Fig. 1A), suggesting that the Pictet–Spengler reaction occurred predominantly at the N-terminal aldehyde.

In this study, the N-terminal residue of horse heart myoglobin was modified by the Pictet–Spengler reaction. The protein substrate is not restricted to myoglobin containing an N-terminal glycine residue; our approach is generally applicable to recombinant proteins. An N-terminal glycine residue can be generated by protease digestion of the cleavage site located between an N-terminal affinity tag and a glycine residue.

3-(2-Aminoethyl)indole was used as a substrate, and the Pictet–Spengler reaction proceeded under nearly physiological conditions: aqueous media, ambient temperature (<37 °C), and neutral pH (pH 6.5). Electron-rich 3-(2-aminoethyl)indole is a preferred substrate for the Pictet–Spengler reaction, as compared with β-arylethylamine.[17] Therefore, tryptophan, with a free carboxy group, will be a versatile coupling partner of aldehyde groups on proteins.

We used biotinylated tryptophan **7** to functionalize myoglobin with biotin. In addition, labeling with a fluorescent group is possible when a modification unit with a fluorescent dye is used instead of **7**. An advantage of N-terminal labeling with a low molecular-weight fluorescent group is that the activity of the modified proteins is less severely affected than when the proteins are fused with GFP (green fluorescent protein). The derivatization of the N-terminal residue with a fluorescent group can be an alternative approach to the labeling of proteins with fluorescent groups. This is possible by the Pictet–Spengler reaction and the following aromatization of the tetrahydro-β-carboline residue, since β-carboline derivatives (e.g., harmine, harmane, and norharmane) are fluorescent.[14] Thus, the application of the Pictet–Spengler reaction to protein functionalization will open up a new frontier of biotechnology.

Acknowledgments

We are grateful to Prof. Hiroshi Nishihara and Dr. Shoko Kume of the University of Tokyo, for help in measuring MALDI mass spectra and to Masaru Kinugawa of the University of Tokyo for help in measuring NMR spectra. Thanks are also due to Prof. Shigeyuki Yokoyama, Dr. Takuhiro Ito and Shinya Mimasu of the University of Tokyo, for help in HPLC purification and UV/vis absorption measurement and to Dr. Hiroshi Nakayama of RIKEN for valuable suggestion. This work was partly supported by a Grant-in-Aid for the Global COE program from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2008.07.033.

References and notes

1. Merkel, L.; Beckmann, H. S. G.; Wittmann, V.; Budisa, N. *ChemBioChem* **2008**, *9*, 1220.
2. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 19.
3. Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192.
4. Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. *ChemBioChem* **2006**, *7*, 134.
5. Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. *ChemBioChem* **2007**, *8*, 232–238.
6. de Araujo, A. D.; Palomo, J. M.; Cramer, J.; Seitz, O.; Alexandrov, K.; Waldmann, H. *Chem. Eur. J.* **2006**, *12*, 6095.
7. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed. Engl.* **2006**, *45*, 5307.
8. Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. *Nat. Chem. Biol.* **2007**, *3*, 321.
9. Zhang, Z.; Smith, B. A.; Wang, L.; Brock, A.; Cho, C.; Schultz, P. G. *Biochemistry* **2003**, *42*, 6735.
10. Saha, B.; Sharma, S.; Sawant, D.; Kundu, B. *Tetrahedron Lett.* **2007**, *48*, 1379.
11. Li, X.; Zhang, L.; Hall, S. E.; Tam, J. P. *Tetrahedron Lett.* **2000**, *41*, 4069.
12. Horiuchi, K.; Yonekawa, O.; Iwahara, K.; Kanno, T.; Kurihara, T.; Fujise, Y. *J. Biochem.* **1994**, *115*, 362.
13. Poulter, L.; Burlingame, A. L. *Methods Enzymol.* **1990**, *193*, 661.
14. Varela, A. P.; Burrows, H. D.; Douglas, P.; da Graca, M. M. *J. Photochem. Photobiol. A Chem.* **2001**, *146*, 29.
15. Shikama, K.; Matsuoka, A. *J. Mol. Biol.* **1989**, *209*, 489.
16. Levine, R. L.; Stadtman, E. R. *Exp. Gerontol.* **2001**, *36*, 1495.
17. Whaley, W. M.; Govindachari, T. R. *Org. React.* **1951**, *6*, 151.