A manual sequence method of peptides and phosphopeptides using 4-(1’-cyanoisoindoly1)phenylisothiocyanate

Takayuki Shibata\textsuperscript{a}, Moses N. Wainaina\textsuperscript{a}, Takayuki Miyoshi\textsuperscript{a}, Tsutomu Kabashima\textsuperscript{a}, Masaaki K\textsuperscript{a,b,*}

\textsuperscript{a}Faculty of Pharmaceutical Sciences, Graduate School of Biomedical Sciences, Nagasaki University, Bunkyo-Machi 1-14, Nagasaki 852-8521, Japan
\textsuperscript{b}Global Center of Excellence Program, Nagasaki University

\*Corresponding author. Tel. and fax: +81 95 819 2438. E-mail address: ms-kai@nagasaki-u.ac.jp
Abstract

A method for sequence analysis and identification of phosphoamino acids in peptides based on high performance liquid chromatography (HPLC) is described. The peptides were derivatized with an Edman type reagent, 4-(1'-cyanoisoindolyl)phenylisothiocyanate (CIPIC) and subsequently cleaved to generate stable and fluorescent 4-(1'-cyanoisoindolyl)phenylthiazolinone (CIP-TZ)-amino acids. Several experimental factors that affected derivatization on membranes were examined. Under the optimized conditions, the CIP-TZ derivatives of Try(p), Thr(p) and Ser(p) were obtained and separated from their parent amino acids with baseline resolution using an isocratic elution system. Up to the 4th residue of phosphorylated pentapeptides was successfully identified, whereas phosphoamino acid residues could not be detected by the conventional procedure using phenylisothiocyanate (PITC). The results demonstrated the potential of CIPIC as a derivatization reagent for peptide sequencing and the applicability of the method for the study and identification of phosphoamino acids in peptides.

Key words: CIPIC, phosphorylated peptides, sequencing, fluorescence detection
1. Introduction

Analysis of the amino acid sequence of proteins is highly important in the life sciences, because the covalent structures are responsible for the biological functions of individual proteins. Protein phosphorylation is an essential post-translational modification in proteins, and plays a central role in many biological processes. To understand protein functions and regulation at the cellular level, it is often necessary to identify the specific amino acid residues that are phosphorylated. Hence such characterization of protein phosphorylation sites is essential in many proteomics studies [1-3]. Conventional methods for analyzing O-phosphorylation sites involve the use of phospho-specific antibodies [4] and incorporation of $^{32}$P into cellular proteins followed by protease digestion and separation by 2D-gel electrophoresis. The phosphorylation site may also be determined by collecting the radioactive fractions using high performance liquid chromatography (HPLC) after radio-Edman degradation or eluting the peptides from the thin layer chromatography (TLC) plates [5,6]. Recently, mass spectrometry combined with immobilized metal affinity chromatography [7,8] has emerged as a non-radioactive method in the detection and characterization of protein phosphorylation. However, this technique is often hindered by several factors including substoichiometric population of phosphopeptides, ionization suppression relative to regular peptides and lability of the phosphate group. Another approach utilizes base-catalyzed elimination followed by Michael addition [9-11]. $\beta$-Elimination and Michael addition also occur at Cys and glycosylated Thr and Ser residues, rendering the method unspecific for phosphorylation sites alone. Even with the latest instrumentations and methodologies, analysis of phosphoproteins remains challenging.

In many procedures, identification and characterization of protein phosphorylation is conducted by determination of phosphoamino acids derived from hydrolysis of phosphorylated proteins. Several labeling reagents have been reported for the derivatization of phosphoamino acids and subsequent detection by HPLC or capillary electrophoresis. They include o-phthalaldehyde (OPA) [12,13], phenylisothiocyanate (PITC) [14], fluorescein isothiocyanate (FITC) [15], 9-fluorenylmethoxycarbonyl chloride (FMOC) [16], 5-(4,6-dichloro-s-triazin-2-ylamino)fluorescein (DTAF) [17], naphthalene-2,3-dicarboxaldehyde (NDA) [18], and 4'-dimethylaminoazobenzene-4-sulfonyl chloride (DABSCl) [19] and N-hydroxysuccinimidyl fluorescein-O-acetate (SIFA) [20]. Even though these reagents have found good use in
derivatization of phosphoamino acids, their application in peptide sequencing has not been reported. Edman degradation method remains the conventional procedure for N-terminal sequence analysis of proteins, but has not been successfully employed for the regular sequence analysis of phosphopeptides. The anilinothiazolinone (ATZ) and phenylthiohydantoin (PTH) derivatives of phosphotyrosine are hardly soluble, and the phosphoesters of Thr and Ser undergo β-elimination and degradation giving undefined products under the conditions used in standard Edman methods [21-23]. Therefore there is a need of suitable technique capable of sequencing low abundance phosphoproteins expressed in cells.

Recently we reported a fluorescent Edman type reagent, 4-(1’-cyanoisoindolyl)phenylisothiocyanate (CIPIC), in order to overcome the disadvantages of low sensitivity of the conventional Edman degradation, which was found to be highly reactive with amino acids and peptides with the lower detection limits ranging from 0.3 to 0.7 pmol [24]. We have also found that CIPIC could be used as a chemiluminescent reagent [25] in the presence of H$_2$O$_2$. The sensitivity of the fluorescence (FL) or chemiluminescence (CL) detection was about 20 times higher than that of ultra violet (UV) detection with PTH derivatives. In this work, we demonstrated the utility of CIPIC for the identification of phosphoamino acids in the regular sequence analysis of peptides. The peptides and phosphopeptides were derivatized with CIPIC to generate 4-(1’-cyanoisoindolyl)phenylthiocarbamoyl peptide (CIPTC-peptide) and thereafter cleaved with anhydrous trifluoroacetic acid (TFA) to generate fluorescent 4-(1’-cyanoisoindolyl)phenylthiazolinone (CIP-TZ) amino acids (Fig. 1). The CIP-TZ-amino acids are stable at room temperature for at least 1 h and do not require the conversion to their corresponding thiohydantoins with an acidic aqueous solution. Furthermore, the extra isoindolyl group in CIPIC increases the hydrophobicity of the CIP-TZ-amino acids, which enhances the solubility of CIP-TZ-phosphoamino acids and enables the extraction of them with organic solvents. The obtained CIP-TZ-amino acids were then detected and analyzed by HPLC with FL detection as the final product. The use of the filter disk as a reaction vessel, made from a glass tube and a Teflon bottom, allowed the simple manual sequencing of peptides. After optimization of the reaction conditions, the present method was successfully applied for the detection of phosphoamino acid existing in peptide samples and the sequencing of pentapeptides up to the 4th residue. In contrast, the phosphoamino acid residue was not identified by the sensitive fluorescent Edman degradation [26], although the 2nd, 3rd and 4th residues of the pentapeptide were
characterized. These results clearly indicated that CIPIC affords the first example of the sequencing of phosphopeptide based on Edman degradation protocol, and the applicability of the present method for direct determination of phosphorylated amino acid residue in phosphoproteins.

2. Experimental

2.1. Materials and apparatus

TFA, PITC, triethylamine (TEA), 2-methoxyethanol, TFA-treated glass fiber membrane were purchased from Wako (Osaka, Japan). All amino acid amides were purchased from Bachem (Bubendorf, Germany). Insulin chain-A (oxidized) and insulin chain-B (oxidized) were obtained from Sigma (St Louis, MO, USA). The phosphopeptide standards and the corresponding non-phosphorylated analogs were purchased from Sigma-Genosys, (Hokkaido, Japan). Acetonitrile, toluene, n-hexane and methanol were obtained from Kanto Chemicals (Tokyo, Japan). Chloroform, dichloromethane, n-propyl chloride, ethyl acetate, pyridine and acetic acid were purchased from Nacalai Tesque (Kyoto, Japan). Water was purified on a Milli-Q system WR 600 A from Millipore, (Molsheim, France). All chemicals were of analytical or guaranteed reagent grade and were used without further purification.

CIPIC was prepared in the laboratory [24] and working solution was freshly prepared in benzene: acetonitrile=1:1. Amino acid amides and insulin chain-A were dissolved in ultrapure water to the required concentrations and stored at 4°C. Insulin chain-B solution was also made in ultrapure water containing 1% TFA. Pyridine and TEA solutions were diluted to adequate concentrations with acetonitrile. The phosphopeptides and non-phosphorylated peptides were dissolved in 50% 2-methoxyethanol containing 1% TEA and 1% TFA, respectively. The glass-Teflon disk was prepared in the laboratory (Fig. 1, inset). To assemble the disk, a polytetrafluoroethylene (PTFE) filter membrane (1.0 μm pore size, 13 mm i.d) purchased from Advantec (Tokyo, Japan) was sandwiched between glass tube (15 mm height and 4 mm i.d) and Teflon support, and 2 pieces of GF membranes were then inserted into the unit.

The HPLC system consisted of a PU-2089 plus quaternary gradient pump, an FP 2020 Plus intelligent fluorescence detector, a UV 2070 detector from Jasco (Tokyo, Japan) and a recorder
from Rikadenki (Tokyo, Japan). The reversed phase column was TSK-Gel ODS 80Ts (150×4.6 mm i.d, particle size 5 µm) from Tosoh (Tokyo, Japan) with the column maintained at 40°C using a CO-8020 column oven from Tosoh. The derivatives were monitored simultaneously by UV detection at 280 nm and FL detection at 360 and 410 nm for excitation and emission, respectively.

2.2. General procedure for the preparation of standard CIP-TZ amino acids

A solution of amino acid amide (5 nmol) was applied onto the GF membranes in the glass-Teflon disk. After drying the membranes in vacuo at 37°C for 5 min, 5 µL of 50% pyridine and 5 µL of 100 mM CIPIC were successively added, and the disk was heated at 60°C for 15 min. The membranes were dried in vacuo at 37°C for 5 min and washed 2 times with 30 µL toluene. To a 250 µL eppendorf tube, 30 µL of TFA was added and the disk firmly placed over the tube, and then heated at 60°C for 2.5 min. The CIP-TZ derivative was then extracted from the membranes with 30 µL of methanol, and 10 µL of the eluent was injected into HPLC for FL detection.

2.3. Sequencing of peptides

The peptide solutions, XVFAL (X= Y, Y(p), T, T(p), S or S(p)) (10 nmol), were applied onto the GF membranes in the glass-Teflon disk. The peptides were reacted in the same manner as described in Section 2.2 except for the reaction time of 25 min. After drying the membranes in vacuo at 37°C for 5 min, the membranes were washed 4 times with 30 µL toluene. The coupled peptides were treated with TFA vapor for 2.5 min by heating 30 µL of TFA at 60°C as described above. The CIP-TZ derivatives were extracted with 30 µL of toluene:EtOAc=3:1. The eluent was then diluted with 30 µL of MeOH and a portion (10 µL) of the mixture was injected into HPLC for FL detection. At the same time, the membranes after extraction were dried in vacuo at 37°C for 5 min. The resulting peptides lacking the N-terminal amino acid residue was again placed for CIPIC reaction in the same manner as described above.
3. Results and discussion

3.1. Derivatization conditions with CIPIC

We have previously reported the manual sequencing method of peptide utilizing modified Edman degradation technique [26]. In this work, a small filter disk made of polypropylene was employed as a reaction vessel. However, heating of polypropylene with reagents and solvents produced many unexpected impurities which inhibited the following sequencing reactions. A similar inhibitory effect was observed when CIPIC was utilized for amino acid derivatization in our preliminary study. Therefore, we created an invulnerable reaction vessel using a grass tube and a Teflon support (Fig. 1, inset). The vessel was disconnectable throughout the Edman degradation procedure for peptide sequencing and did not produce the inhibitory artifacts. The vessel was proof against repeated use, whereas the polypropylene disk was disposable.

To optimize the conditions for the derivatization of CIPIC with peptides, several parameters of the reaction were investigated. These include the reagent concentration, solvents for extraction of CIP-TZ amino acids from the GF membranes, reaction temperature for the coupling and reaction time for the cleavage. Generally, the derivatization yield was based on the FL peak heights of the CIP-TZ derivatives. The concentration of CIPIC for peptide coupling was examined with insulin chain A. The yield of the obtained CIP-TZ amino acids increased rapidly, when the concentration of CIPIC was increased up to a ratio of 1:100 as shown in Fig. 2. Further increase of the reagent was accompanied by decrease in product formation and appearance of excessive by-product peaks. A 100 fold excess of the reagent was sufficient for the derivatization of peptides with less interference in chromatography.

![Fig. 2]

Derivatization of peptides with Edman type reagents are generally performed in weak alkaline medium. Bases such as TEA and pyridine facilitated the coupling reaction of CIPIC with amino acid amides (Asn-NH₂, His-NH₂ and Thr-NH₂) and peptides (insulin chain B and Y(p)VFAL). Fig. 3 A shows the effect of bases on coupling efficiency of CIPIC with the peptide Y(p)VFAL. It was found that 50% pyridine in acetonitrile provided the most intense peaks and
was used in the recommended procedure. The time course for the coupling reaction of CIPIC with peptides is shown in Fig. 3 B. The coupling products reached a plateau at 25 min for the 1st and 2nd cycle of phosphopeptides. The peak heights of CIP-TZ-Val at the 2nd cycle were almost the same for the three peptides, suggesting that the reactivity with CIPIC was ideally constant. Thus, heating at 60°C for 25 min was adopted for the coupling reaction. To remove the background interference due to excess reagents and impurities, the membranes were washed with different solvents after the CIPIC reaction. These impurities were found to inhibit the subsequent cleavage and coupling reactions. Apparently, when sequencing was performed without the washing procedure, low peak height of CIP-TZ-Val at the 2nd cycle was observed. Washing the membranes 4 times with toluene was found to remove most of the interfering materials without losing the coupled peptides.

\textless Fig. 3\textgreater

At the cleavage stage of the CIPIC reaction to generate CIP-TZ-amino acids, we applied a gas-phase method which was found to reduce the generation of TFA-related impurities on membranes [26]. The reaction vessel was tightly attached to a tube containing liquid TFA, and TFA vapor generated by heating the tube was allowed to pass through the membranes. The reaction kinetics was examined at 60°C and 80°C, and the results showed that the coupled peptides were readily cleaved to generate the corresponding CIP-TZ amino acids (Fig. 4). It was found that high temperature allowed the cleavage to proceed more rapidly, but the peaks of final derivatives were smaller in most occasions. Even though the formation of CIP-TZ-Tyr and CIP-TZ-Tyr(p) in the 1st cycle reached completion faster at 80°C (about 1 min), the products were found to be more stable at 60°C. The maximum yield of CIP-TZ-Val at the 2nd cycle for all the peptides reached a plateau at about 2.5 min. The peak heights of CIP-TZ-Val and Phe at the 2nd and 3rd cycles obtained after the cleavage of Y(p)VFAL were smaller than those of YVFAL, suggesting the difference in reactivity because of an electron withdrawal and/or a steric effect of the phosphate moiety. Taking into consideration of the possible degradation of the labile phosphoamino acids, the cleavage reaction time of 2.5 min and reaction temperature of 60°C provided the optimal yield of CIP-TZ derivatives for all the peptides.
Selection of solvents for the extraction of CIP-TZ amino acid was necessary due to the difference in polarity of CIP-TZ derivatives, especially for phosphoamino acids. Extractive loss of peptide from the membranes should also be minimized at this stage to carry out sequential analysis of the peptide. The extraction of CIP-TZ-Tyr(p), Thr(p) and Ser(p) was best achieved in EtOAc, MeOH and toluene: EtOAc=3:1 (Fig. 5). EtOAc was discouraged due to low retention of peptides in the second cycle. Surprisingly, the retention of the sample pentapeptides was sufficiently good using MeOH. However, MeOH extraction was found to be inefficient for a long sequencing procedure when insulin chain-B was used for sequence analysis (data not shown). Therefore, toluene: EtOAc=3:1 was chosen for regular sequencing since it could extract both the phosphorylated and non-phosphorylated CIP-TZ derivatives with high retention of the peptide.

3.2. Identification of CIP-TZ amino acids

The CIP-TZ derivatives of Tyr(p)/Tyr, Thr(p)/Thr and Ser(p)/Ser could be detected and well separated by HPLC based on retention times, even though the retention times of CIP-TZ-Thr(p) and Ser(p) were close. Using an acidic mobile phase, the phosphorylated derivatives generally had longer retention times. However, when neutral mobile phase comprising of PBS buffer was employed, the phosphorylated derivatives were poorly retained in the column and were eluted as dead volume. This is due to the high polarity of the phosphate anion as compared to the phosphoric acid group. As shown in the chromatograms (Fig. 6), all the derivatives were obtained as a single peak showing that CIPIC specifically reacted with the α-amino group. The detection limit of CIP-TZ-amino acids was 20 pmol with the signal-to-noise ratio (S/N) of 3. The detection limit of 0.3-0.7 pmol was achieved in our previous study regarding the derivatization of amino acids with CIPIC [24]. This difference in detection sensitivity should be caused by the efficiency of the reaction, loss of the peptides on the membranes and/or absorption of the CIP-TZ-amino acids on the membranes, since the present system involves solid-phase reaction, washing with and extraction steps. Most importantly, there was no evidence for the dephosphorylation or
degradation of phosphorylated CIP-TZ amino acids as has been reported in conventional Edman procedures [27].

Alternatively, modified Edman degradation was utilized in order to investigate the possibility of the detection of phosphoamino acid with PITC, in which peptides were reacted with PITC and the obtained ATZ-amino acid intermediates were treated with CIA to generate fluorescent PTC-CIA derivatives [26]. However, the PTC-CIA corresponding to the phosphoamino acids could not be detected and identified (data not shown). Furthermore, multiple product peaks appeared in the case of Thr(p) and Ser(p), and they contained identical peaks to the CIA derivatives of parent amino acids, possibly due to dephosphorylation. The reaction for fluorescence derivatization of ATZ-amino acids with CIA requires heating at 80°C for 20 min [26]. This harsh reaction condition should cause the dephosphorilation and/or the degradation of phospho-PTC-CIA. These findings suggested that CIPIC could be applied for the sequence determination of phosphopeptides without decomposing the phosphate groups.

3.3. Sequencing of peptides

A manual sequencing procedure established as described above was applied for sequencing of pentapeptides on GF membranes. Fig. 7 shows the results of the sequencing of T(p)VFAL and TVFAL. From the chromatograms, the CIP-TZ derivatives in the respective cycles were easily identified up to the 4th residue out of 5 amino acid residues. In particular, the Thr(p) residue at the 1st cycle was easily identified in the sequence. However, the 5th Leu was not detected, since the C-terminal amino acid produced at the 4th cleavage stage should be removed from the membranes simultaneously with CIP-TZ-Ala at the 4th extraction stage. The detection limit of CIP-TZ-amino acids at the 1st residue was almost same as that of CIP-TZ-amino acids derived from amino acid amides, and approximately 150 pmol of the 4th residue was detected with S/N=3. At the 3rd and 4th cycles, some CIP-TZ peaks derived from the incomplete reaction from previous cycles were observed. This might be caused by accumulated impurities which had been found to cause unfavorable interferences. This however did not affect the identification of amino acid residues. It should be noted that modified Edman degradation with CIA did not give
corresponding derivatives of phosphoamino acid residues, although the 2nd, 3rd and 4th residues of the pentapeptide were clearly identified. We also used VT(p)FAL, that has a different sequence from T(p)VFAL, for sequencing in order to compare the detection limit (data not shown). The peak height of CIP-TZ-Thr(p) at the 2nd cycle of VT(p)FAL was approximately 50% lower than that at the 1st cycle of T(p)VFAL. This might be due to the incomplete reaction of CIPIC with the peptide, loss of the peptide from the membranes and/or absorption of the CIP-TZ-Thr(p) on the membranes. On the other hand, the peak heights of CIP-TZ-Phe at the 3rd cycle and CIP-TZ-Ala at the 4th cycle of both peptides were almost identical. This indicated that almost the same sensitivity for sequencing would be obtained regardless of the position of phosphoamino acids in the pentapeptides. These results unambiguously proved that the present method is the first example that enables the direct sequencing of phosphopeptides.

<Fig. 7>

4. Conclusion

A manual sequencing method using CIPIC as a fluorescent Edman type reagent has been developed for peptide sequencing on solid-phase GF membranes in the glass-Teflon filter disk. CIPIC was found to be highly reactive with amino acids and peptides to generate stable fluorescent derivatives, CIP-TZ amino acids after cleavage. The excess of the reagents at the CIPIC coupling stage were washed out by a simple filtration with toluene, and CIP-TZ amino acids were easily eluted using toluene:EtOAc=3:1 solution without losing the peptides on the membranes. The phosphorylated CIP-TZ derivatives of phosphopeptides were identified and separated by HPLC. On the other hand, the ATZ derivatives of the Try(p), Thr(p) and Ser(p) could not be detected by the conventional procedure using PITC, even with the sensitive post-cleavage method with CIA. Dephosphorylation or degradation of the phospho-CIP-TZ derivatives was not observed, whilst dephosphorylation of corresponding phospho-PTC-CIA derivatives was observed. Under the procedure, up to four residues of pentapeptides could be identified with the detection limits of 20 pmol for the first residue and 150 pmol for the 4th residue, demonstrating the ability to sequence even short peptides on membranes. The results showed that CIPIC is the first example of an Edman-type reagent that allows the detection of
phosphoamino acid residues existing in peptide samples, and the glass-Teflon reaction vessel afforded a simple, sensitive and efficient sequencing of peptides. The proposed method could effectively be used for the direct analysis of primary sequence and phosphorylated position of phosphoproteins, even with a small amount of (phospho)proteins obtained from cell samples.

Acknowledgement

This work was supported by the Shimadzu Foundation (Japan), partly by Grants-in-Aid for Scientific Research from the Ministry of Health, Labor, and Welfare, Japan, and partly by the Asahi Glass Foundation.
References


**Figure captions**

**Fig. 1.** Derivatization of peptides with CIPIC. The peptide sample was applied and reacted with CIPIC on GF membranes supported in a glass-Teflon filter disk. The inset shows the design and assembly of the filter disk.
Fig. 2. Effect of CIPIC concentration on the coupling reaction of insulin chain A for the first three residues. Product yield was determined by the peak height of the final CIP-TZ amino acids at the selected cycles.
**Fig. 3.** Effect of (A) bases and (B) reaction time for the coupling reaction of peptides with CIPIC. The coupling yield was determined by the peak height of CIP-TZ derivatives after cleavage. In (A), the peptide (YpVFAL) was reacted with CIPIC in different basic conditions and the CIP-TZ derivatives at each cycle extracted with either toluene: EtOAc=3:1 or methanol. In (B), the peptides (Yp/Tp/SpVFAL) were reacted as described in the experimental section. The reaction efficiency was evaluated by the peak height of CIP-TZ-Val at the 2nd cycle for each peptide.
Fig. 4. Effect of reaction temperature and time on cleavage of peptides (A) Y(p)VFAL and (B) YVFAL. The peak height shows the yield of CIP-TZ amino acids at the 1st and 2nd cycles for each peptide.
Fig. 5. Selection of elution solvent for extraction of CIP-TZ amino acids from membranes after cleavage of phosphorylated pentapeptide Y(p)/T(p)/S(p)VFAL. The yield was determined by peak heights of the derivatives after each cycle.
Fig. 6. Separation and identification of CIP-TZ amino acids obtained by sequencing of pentapeptides XVFAL (X= Y, Y(p), T, T(p), S, or S(p)) and extracted from membranes using toluene:EtOAc=3:1 (30 μL). The eluents were diluted with MeOH (30 μL) and portions (10 μL each, equivalent to 1.67 nmol of the peptides used) were applied to HPLC. The HPLC chromatograms show the CIP-TZ amino acid derivative of each peptide at the 1st cycle. HPLC mobile phase: 70% MeOH, 10% 1.6 M AcOH, 20% H2O.
Fig. 7. HPLC chromatograms of CIP-TZ amino acids obtained by manual sequencing of pentapeptides (A) TVFAL, and (B) T(p)VFAL on membranes. The CIP-TZ amino acid derivatives at each cycle were extracted from membranes using toluene:EtOAc=3:1 (30 µL). The eluents were diluted with MeOH (30 µL) and portions (10 µL each, equivalent to 1.67 nmol of the peptides used) were applied to HPLC. The arrows show the position of the derivatives at the indicated cycle. HPLC mobile phase: 70% MeOH, 10% 1.6 M AcOH, 20% H2O.