Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 6128–6131, October 1982
Biochemistry

n-Tetradecanoyl is the NH_2-terminal blocking group of the catalytic subunit of cyclic AMP-dependent protein kinase from bovine cardiac muscle

(direct chemical ionization mass spectrometry/fast atom bombardment mass spectrometry/gas chromatographic mass spectrometry/HPLC/hydrophobic peptides)

STEVEN A. CARR*†, K. BIEMANN*‡, SHOZO SHOJI§¶, DAVID C. PARMELEE§, AND KOITI TITANI‡§

*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; and §Department of Biochemistry and Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195

Communicated by Hans Neurath, June 21, 1982

ABSTRACT **The unusual NH_2-terminal blocking group of the catalytic subunit of bovine cardiac muscle cyclic AMP-dependent protein was found to be amide-linked *n*-tetradecanoic acid by gas chromatographic-, direct chemical ionization-, and fast atom bombardment-mass spectrometry. In addition, fast atom bombardment mass spectrometry revealed the presence of an additional alanine which had been overlooked when the original sequence was determined. The corrected and completed NH_2-terminal sequence of the 350-amino acid catalytic subunit is $CH_3(CH_2)_{12}CONH$-Gly-Asn-Ala-Ala-Ala-Ala-Lys.**

Posttranslational covalent modification of proteins plays a central role in forming biologically active molecules from the end products of translation. Over 140 types of such alterations involving any of the amino acids occurring in proteins have been identified, and the diversity of functionalization is truly extraordinary (1). A subset of these involves acylation of the α-amino terminus of the protein and these blocking groups are usually formyl or acetyl or involve the free γ-carboxyl group of NH_2-terminal glutamic acid forming a pyrrolidone ring (pyroglutamic acid). However, longer chain fatty acids are rarely found at the NH_2 terminus of a protein and a prokaryotic membrane protein (2) is, to our knowledge, the only well-defined case. There are also other proteins known to contain covalently linked fatty acids (3–5) but, because these are amenable to Edman degradation, the acyl groups cannot be attached to the NH_2 terminus.

We have now found that the NH_2 terminus of the catalytic subunit of bovine cardiac muscle cyclic AMP-dependent protein kinase—the amino acid sequence of which was published recently (6)—is blocked by an *n*-tetradecanoyl (myristyl) group. It was this substituent that made the sequence analysis of the NH_2 terminus particularly difficult.

This unexpected acyl group was identified by using gas chromatographic mass spectrometry (GCMS) (7) and the recently introduced techniques of direct chemical ionization (DCI) (8–11) and fast atom bombardment (FAB) (12, 13) mass spectrometry. These data also led to a revision of the NH_2-terminal sequence which was found to contain an additional alanine.

MATERIALS AND METHODS

Isolation of the NH_2-Terminal Cyanogen Bromide Fragment. The catalytic subunit of cyclic AMP-dependent protein kinase was prepared from bovine heart as described by Peters *et al.* (14). The S-carboxymethylated protein (45 mg) was cleaved with cyanogen bromide (Eastman) at 20°C in 72% formic acid for 15 hr. The digest was separated on a column of Sephadex G-50 (superfine) in 9% formic acid. The third emergent peak contained two fragments, CB1 and CB4, as described by Shoji *et al.* (6). Fragment CB1 was purified on a μBondapak C_{18} column (Waters) by reversed-phase HPLC by using a system similar to that shown in Fig. 1.

Derivatization of Blocking Groups. Small peptides obtained after proteolytic cleavage of fragment CB1 were dissolved (≈40 nmol each) in 100 μl of trifluoroacetic acid (sequanol grade, Pierce) to which 1 ml of constant boiling HCl (≈6 M) (Pierce) was then added. The solutions were flushed with nitrogen, degassed, and heated under reduced pressure in Teflon sealed-glass hydrolysis bulbs (Pierce) at 110°C for 44 hr. Each of the hydrolysates was extracted three times with $CHCl_3$. The $CHCl_3$ extracts were taken to dryness under a stream of nitrogen and then were methylated (separately) with ethereal diazomethane [10% (vol/vol) MeOH] (15) for 30 min at 25°C. These solutions were blown down to dryness and taken up in 40 μl of $CHCl_3$ (each); 1- to 3-μl aliquots were used for gas chromatographic and GCMS analyses. Reagent blanks, solvent blanks from the HPLC, and standards of C_{12}, C_{14}, C_{16}, and C_{18} unbranched fatty acids (Sigma) were subjected to the same hydrolysis and diazomethylation conditions as the carboxylic acid fragment obtained from the peptides. The reagent and solvent blanks were found to be free of fatty acids and other potentially interfering contaminants. For 2 μg of standard (≈10 nmol) treated as described above, recovery of the fatty acids was >80% based on comparison of the gas chromatographic peak area with that obtained for 2 μg of standard which was diazomethylated directly.

Gas Chromatography and GCMS. Aliquots of the derivatized extracts of the peptide hydrolysates, solvent, and reagent blanks and fatty acid standards were subjected to gas chromatography on 30-m SE-54 and SE-30 fused-silica capillary columns (J&W Scientific, Orange Vale, CA) by using a Hewlett–Packard model 5840A gas chromatograph with a splitless injector and flame ionization detector. The temperature was linearly programed from 60°C to 310°C at 5°C/min. The GCMS computer system consists of a Varian model 3700 gas chromatograph coupled via an open-split interface to a Finnigan-MAT 212 double-focusing mass spectrometer operating in electron impact mode with an ionization potential of 70 eV. A Finnigan-MAT SS200 data system controls the instrument and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: GCMS, gas chromatographic mass spectrometry; DCI, direct chemical ionization; FAB, fast atom bombardment.

† Present address: Dept. of Biochemistry, Harvard Medical School, Boston, MA 02115.
‡ To whom reprint requests should be addressed.
¶ Present address: Dept. of Biochemistry, Faculty of Pharmaceutical Sciences, Kumamoto Univ., Kumamoto 862, Japan.

acquires, processes, and stores the data. For these experiments the mass spectrometer was scanned from mass-to-charge ratio m/z 40 to m/z 800 at 2 s per decade.

DCI and FAB Mass Spectrometry. DCI mass spectra were obtained as described (10, 11) by using 0.1–0.3 μg of the blocked tripeptide. FAB mass spectra (12, 13) were recorded on a Finnigan-MAT 731 double-focusing mass spectrometer equipped with an Ion Tech (Teddington, Middlesex, England) B-11 fine-beam saddle-field fast atom source; xenon was used as the reagent gas. Approximately 40 nmol of underivatized peptide was dissolved in 5–8 μl of glycerol containing 2–3 μl of trifluoroacetic acid; this was gently heated and sonicated briefly, and 1 μl of this solution was applied to a copper target attached to the field desorption insertion probe. The probe was then inserted into the ion source (combined electron impact/field ionization/field desorption/FAB) of the mass spectrometer and bombarded with the xenon beam (operating parameters: source pressure = 3×10^{-5} torr of Xe; FAB tube voltage = 6 kV; tube current = 3.5 mA).

RESULTS AND DISCUSSION

The determination of the amino acid sequence of the NH_2-terminal portion of the catalytic subunit of bovine heart protein kinase was hindered by the presence of a blocking group which prevented the direct use of Edman degradation on the intact protein. Proteolytic digestion of CB1 yielded two small blocked peptides from the NH_2-terminal portion of the protein (6). Digestion of these peptides by carboxypeptidases suggested the blocked sequence X-Gly-Asn-Ala-Ala-Ala-Lys.

To confirm this tentative NH_2-terminal sequence and also to identify the group blocking the NH_2-terminal glycine, the two small blocked peptides were reisolated by reversed-phase HPLC and subjected to mass spectrometric analysis. Fragment CB1, prepared as described, was digested with *N*-tosylphenylalanine chloromethyl ketone-treated trypsin (Worthington) in 0.1 M NH_4HCO_3 at pH 8.0 for 3 hr. The lyophilized digest was separated by HPLC as shown in Fig. 1*A*. Fraction CB1-T1 contained a blocked peptide that eluted at a very high concentration of acetonitrile (*ca.* 53%). Amino acid analysis showed that the peptide contains *ca.* 4 residues of alanine (1.06 Asx, 1.08 Gly, 3.76 Ala, and 1.00 Lys) rather than the three residues previously reported (6). After further purification of the peptide by rechromatography with the same HPLC system, a portion (*ca.* 200 nmol) was digested with pepsin at 37°C in 5% formic acid for 3 hr and subjected to HPLC (Fig. 1*B*). A new peptide (CB1-T1-P1) with a blocked NH_2 terminus and the composition 0.87 Asx, 1.02 Gly, and 1.00 Ala was recovered. The COOH-terminal portion of CB1-T1 was not recovered and is probably under the breakthrough peak of guanidine. Because both peptides require >50% acetonitrile for elution from this reversed-phase system—whereas synthetic acetyl-Ala-Ala-Ala elutes in 10% acetonitrile under analogous conditions (Fig. 1*A*)—it appeared probable that the blocking group was far more hydrophobic than an acetyl.

The structure of the blocking group was determined by a combination of mass spectrometric techniques. Hydrolysates of both X-Gly-Asn-Ala and the blocked tryptic fragment (6 M HCl, 44 hr, 110°C under reduced pressure) were extracted with chloroform, diazomethylated in ethereal diazomethane [10% (vol/vol) MeOH], and then subjected to gas chromatography on fused-silica capillary columns. A single sharp peak was obtained from the derivatized extracts of both the blocked tripeptide and tryptic fragment. This compound was found to coelute with authentic *n*-tetradecanoic acid methyl ester on two phases of differing polarity (SE-54 and SE-30 columns). Finally, the mass spectrum of this substance obtained by capillary column

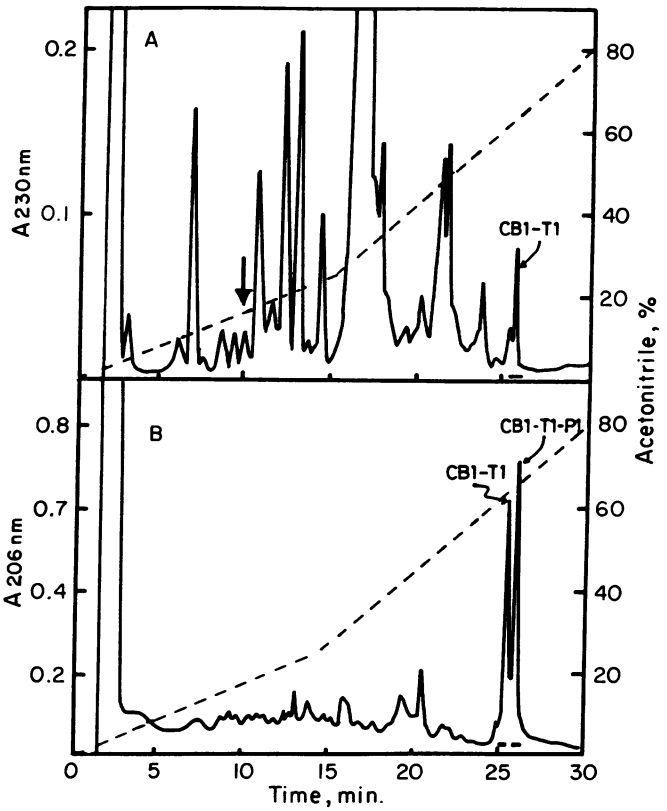

FIG. 1. Isolation of the NH_2-terminal blocked peptides by HPLC. (*A*) Separation of tryptic digest of fragment CB1 (*ca.* 500 nmol). The vertical arrow indicates the elution position of synthetic acetyl-Ala-Ala-Ala. (*B*) Separation of a peptic digest of peptide CB1-T1 (*ca.* 220 nmol). In both chromatograms the digests, in a small volume of 6 M guanidine HCl, were injected into a column of μBondapak C_{18} (0.39 × 30 cm) equilibrated in 0.1% trifluoroacetic acid and eluted at ambient temperature by a gradient (broken line) from 0.1% trifluoroacetic acid to 70% acetonitrile containing 0.08% trifluoroacetic acid. The flow rate was 2 ml/min.

GCMS was identical to the published spectrum of methyl myristate.

Additional confirmatory evidence for the structure of X and its attachment to the protein was obtained by analysis of the underivatized blocked tripeptide and the NH_2-terminal tryptic fragment by FAB (12, 13) and DCI with ammonia as the reagent gas (10, 11). Both FAB and DCI are soft ionization methods that produce cationized molecular ions and fragment ions from involatile and thermally labile substances. The FAB mass spectrum of X-Gly-Asn-Ala (not shown) exhibits an abundant protonated molecular ion $[M + H]^+$ at m/z 471 and sodium cationized molecular ion $[M + Na]^+$ at m/z 493. No sequence-related ions were present in the spectrum. The difference between the calculated weight of the peptide portion of the molecule and the experimentally determined molecular weight is 211, which corresponds to $CH_3(CH_2)_{12}CO$—.

The NH_3-DCI mass spectrum of X-Gly-Asn-Ala is dominated by fragment ions arising from the blocked portion of the molecule (Fig. 2). The ion at m/z 245 corresponds to the NH_4^+ cationized NH_2-terminal amide (i.e., $[CH_3(CH_2)_{12}CO—NH_2 + NH_4]^+$). The peak at m/z 302 that corresponds to $[CH_3(CH_2)_{12}CO—NH—CH_2—CO—NH_2 + NH_4]^+$, (= $[(A_1' + H) + NH_4]^+$; Fig. 2) confirms the assumption that the blocking group is attached to the NH_2-terminal Gly. Also present, although at very low relative abundance, are the ions $[M + H]^+ = m/z$ 471 and $[(M + H) - NH_3]^+ = m/z$ 454. A more general

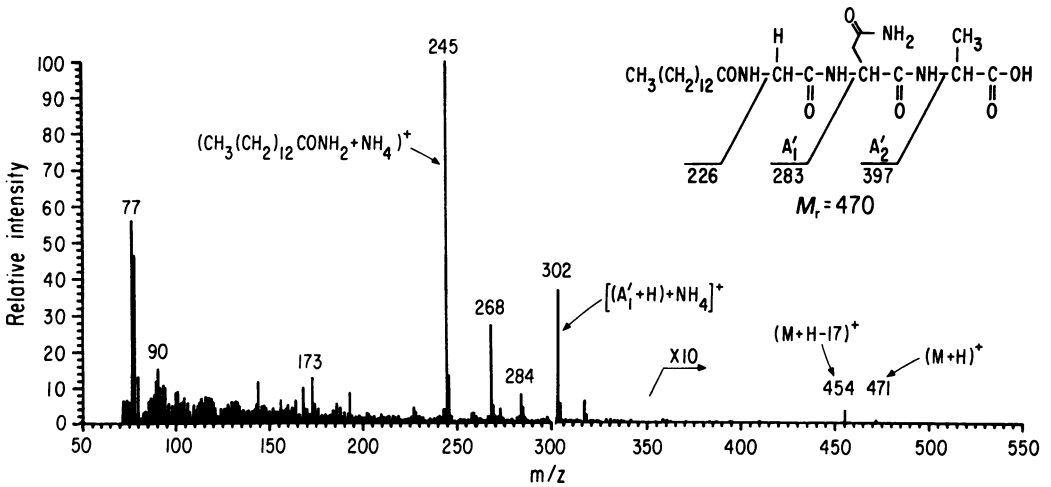

FIG. 2. NH_3-DCI mass spectrum of 0.3 μg of the underivatized blocked peptide *n*-$CH_3(CH_2)_{12}CO$-NH-Gly-Asn-Ala-OH.

discussion of peptide sequence analysis by NH_3-DCI has been presented elsewhere (10).

The FAB mass spectrum of the blocked NH_2-terminal tryptic fragment did not exhibit the expected protonated molecular ion at *m/z* 741; instead, a strong signal was observed at *m/z* 812 (Fig. 3). The mass difference of 71 corresponds to an Ala residue (—NH—$CHCH_3$—CO—), indicating that it had been omitted from the NH_2-terminal sequence proposed previously (6). The enhanced contribution to *m/z* 813 in the molecular ion cluster (Fig. 3) is due to partial conversion of Asn to Asp (mass = +1), which is probably due to the use of trifluoroacetic acid as a solvent. Therefore, the structure of the tryptic peptide must be as follows: $CH_3(CH_2)_{12}CO$-Gly-Asn-Ala-Ala-Ala-Ala-Lys. It is evident that the two partial sequences obtained by composition and carboxypeptidase digestion (see above) abut rather than overlap as originally concluded (6). The original amino acid analysis of the tryptic fragment gave an erroneously low value for the number of Ala (3.5), possibly due to minor contamination with a peptide ending with -Ala-Lys-Lys.

There have been few reports of fatty acids covalently bound to proteins, but in each case the protein appears to be associated with a membrane, such as in chicken embryo fibroblasts (3) and in virus-infected fibroblasts (4, 5). Perhaps the best documented analysis is that of the murein lipoprotein of the outer membrane of *Escherichia coli*, where three fatty acids are covalently attached (2). One of these acylated the α-amino group of the terminal cysteine residue and the other two are esterified to a glyceryl group which forms a thioether bond to the same cysteine. In the present case, there is evidence of the association of the type II protein kinase as a peripheral protein on the inner surface of the plasma membrane (16, 17). However, activation with cyclic AMP solubilizes the catalytic subunit that contains the myristyl group and leaves the regulatory subunit, and its NH_2-terminal acetyl group (18), associated with the membrane. Thus, a role for the myristyl group as a membrane anchor seems unlikely and one can only speculate alternative roles in subunit interactions, in substrate specificity, or possibly in intracellular compartmental translocation (19).

Finally, these results demonstrate that the combination of mass spectrometric techniques, such as DCI and FAB, with conventional Edman degradation or with the GCMS protein sequence analysis methodology (7) is far more useful in the determination of the complete and correct primary structure of a protein than any of the individual approaches alone.

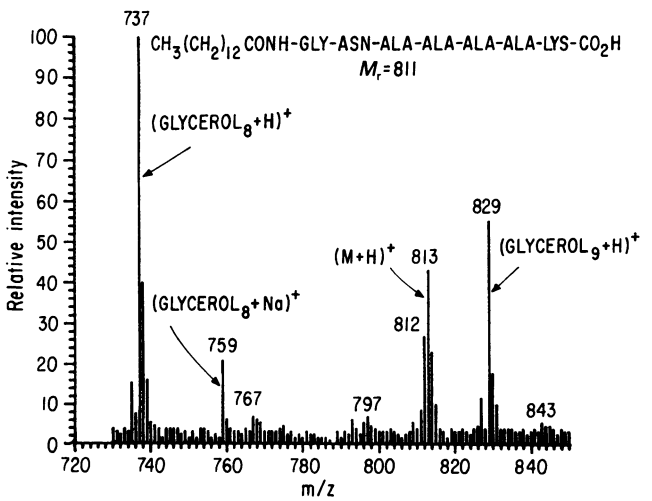

FIG. 3. FAB mass spectrum of the underivatized blocked NH_2-terminal tryptic peptide from bovine heart protein kinase. Only the molecular ion region is shown; the enhanced contribution of *m/z* 813 is due to partial conversion of Asn to Asp (see text). Ions arising from the glycerol sample matrix are also identified.

We thank Dr. Vernon Reinhold of Harvard Medical School for the kind use of the DCI equipment. We also thank Dr. Kenneth A. Walsh and Dr. Hans Neurath for their interest and valuable discussion and Mrs. Santosh Kumar for her technical assistance. This work was supported by grants from the National Institutes of Health at Massachusetts Institute of Technology (DRR00317 and GM 05472) and at the University of Washington (GM 15731).

1. Uy, R. & Wold, F. (1977) *Science* **198**, 890–896.
2. Hantke, K. & Braun, V. (1973) *Eur. J. Biochem.* **34**, 284–296.
3. Schlesinger, M. J., Magee, A. I. & Schmidt, M. F. G. (1980) *J. Biol. Chem.* **255**, 10021–10024.
4. Schmidt, M. F. G., Bracha, M. & Schlesinger, M. J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1687–1691.
5. Silberstein, A., Snider, M. D., Porterm, M. & Lodish, H. F. (1980) *Cell* **21**, 417–427.
6. Shoji, S., Parmelee, D. C., Wade, R. D., Kumar, S., Ericsson, L. H., Walsh, K. A., Neurath, H., Long, G. L., Demaille, J. G., Fischer, E. H. & Titani, K. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 848–851.
7. Biemann, K. (1980) in *Biochemical Applications of Mass Spectrometry*, eds. Waller, G. R. & Dermer, O. C. (Wiley, New York), 1st Suppl. Vol., pp. 469–525.
8. Baldwin, M. A. & McLafferty, F. W. (1973) *Org. Mass Spectrom.* **7**, 1353–1356.
9. Cotter, R. S. (1981) *Anal. Chem.* **52**, 1589A–1606A.

10. Carr, S. A. & Reinhold, V. N. (1982) in *Proceedings of the IV International Conference on Methods in Protein Sequence Analysis*, ed. Elzinga, M. (Humana, Clifton, NJ), pp. 263–270.
11. Reinhold, V. N. & Carr, S. A. (1982) *Anal. Chem.* **54,** 499–503.
12. Barber, M., Bordoli, R. S., Sedgwick, R. D. & Tyler, A. N. (1981) *Nature (London)* **293,** 270–275.
13. Barber, M., Bordoli, R. S., Garner, G. V., Gordon, D. B., Sedgwick, R. D., Tetler, L. W. & Tyler, A. N. (1981) *Biochem. J.* **197,** 401–404.
14. Peters, K. A., Demaille, J. G. & Fischer, E. H. (1977) *Biochemistry* **16,** 5691–5697.
15. Schlenk, H. & Gellerman, J. L. (1960) *Anal. Chem.* **32,** 1412–1417.
16. Rubin, C. S., Rosenfeld, R. D. & Rosen, A. M. (1973) *Proc. Natl. Acad. Sci. USA* **70,** 3735–3738.
17. Corbon, J. D., Sugden, P. H., Lincoln, T. M. & Keely, S. L. (1977) *J. Biol. Chem.* **252,** 3854–3861.
18. Takio, K., Smith, S. B., Krebs, E. G., Walsh, K. A. & Titani, K. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 2544–2548.
19. Jungmann, R. A. & Russell, D. H. (1977) *Life Sci.* **20,** 1787–1798.