In conclusion, our results show that inhibitors of acylation may be developed on the basis of analogues of myristic acid. Such inhibitors inhibit not only the N-myristoyl transferase activity, but also acylation of proteins in intact cells. They are also able to inhibit virus acylation. Preliminary results show that they inhibit HIV release from infected H9 cells in culture and therefore may be potential antiviral agents. A rather surprising result is that glucosamine may also interfere with acylation of viral proteins.


Received 11 May 1989

Structure determination of acylated proteins

ALASTAIR AITKEN
National Institute for Medical Research, Mill Hill, London NW7 1AA, U.K.

A wide range of eukaryotic proteins with diverse function are now known to be acylated. There are three main types of modification: at the N-terminus (on glycine residues [1]); internally attached to cysteine or hydroxyamino acids [2]; or at the C-terminus as a phosphatidylinositol containing glycan moiety [3]. In many cases, the function that has been ascribed to these post-translational covalent modifications is not simply to enhance membrane association.

Indeed the first two proteins that were shown to contain N-terminal myristoyl groups, the catalytic subunit of cyclic AMP-dependent protein kinase [4] and B-subunit of calcineurin [5], are not integral membrane proteins.

Fatty acids linked covalently to protein have been reported in the literature since Hankte & Braun [6] showed the attachment of fatty acids (mostly palmitate) to the amino group of the N-terminal cysteine residue in the murine lipoprotein of the Escherichia coli outer membrane. This protein also contains fatty acids in a diacylglyceride which is linked to the thiol of this cysteine residue. The precursor of the diacylglycerol appears to be one of the major phospholipid species and its fatty acid composition (mainly palmitate) reflects that of the phospholipids from E. coli.

Other proteins contain ester-linked fatty acids (first shown by Schlesinger et al. [7]). The best characterized of these are shown to be linked through the thiol group of cysteine, 70%–80% being palmitic acid and the rest oleic acid and stearic acid.

A number of peptides have been shown to contain an S-farnesyl thioether linked to a C-terminal cysteine residue. This amino acid is also esterified at the α-carboxyl with a methyl group. This trans-farnesyl (sesquiterpene, C20H30 unsaturated alkene) group has been identified in the mating hormone ‘a-factor’ from Saccharomyces cerevisiae.

Abbreviations used: h.p.l.c., high performance liquid chromatography; g.c., gas chromatography; f.a.b., fast atom bombardment; m.s., mass spectrometry.

and other fungi [8, 9]. These hormones have the C-terminal sequence Cys-AAA-Xaa where A is an aliphatic residue and Xaa is any amino acid. This is processed to leave the thioether and methyl esterified cysteine as the C-terminal amino acid. A similar process occurs in the family of G-protein-related p21ras that also contain farnesylated palmitoyl groups [10].

An essential aspect of these studies is the determination of the sites of acylation/alkylation and the precise chemical structure of the modification. Methods currently in use include mass spectrometry (fast atom bombardment (f.a.b.) and direct chemical ionization), high-sensitivity gas-phase protein sequencing and the use of radio-labelled precursors.

Chromatographic techniques include high-performance liquid chromatography (h.p.l.c.) and gas chromatography (g.c.), frequently combined with mass spectrometry (m.s.).

N.m.r. has also been widely used, particularly in the study of the C-terminal alklylation with farnesyl groups, where the terminal cysteine residue is also esterified with a methyl group [8, 9, 11, 12]. N.m.r. is particularly useful in determining the stereochemical configuration of this class of isoprenoid alkyl modification.

Analysis of N-terminal myristoyl blocking groups

The specific group of proteins that have been shown to contain myristic acid covalently attached to the N-terminus are listed in Table 1. These include the catalytic subunit of cyclic AMP-dependent protein kinase (C subunit) and the regulatory subunit of protein phosphatase 2B. The latter is a Ca2+-calmodulin-dependent protein phosphatase that is identical to a protein first isolated from bovine brain, termed calcineurin. The regulatory subunit of protein phosphatase 2B (calcineurin B) binds four calcium ions per mole with affinities in the micromolar range and determination of its complete amino acid sequence [13] has revealed extensive homology to calmodulin and other Ca2+-binding proteins such as troponin C. The discovery of myristic acid at the N-terminus of a protein phosphatase and a protein kinase suggests that this unusual blocking group may be involved in maintaining the subunit-subunit interactions between the
regulatory (B) and catalytic (A) subunits of calcineurin and between the regulatory (R) and catalytic subunits (C) of cyclic AMP-dependent protein kinase. It is unlikely that the role of myristic acid in either of these proteins is for membrane attachment, in contrast to the fatty acylation of some other known examples (see Table 1). The type II isoenzyme of cyclic AMP-dependent protein kinase is known to bind to the inner surface of the plasma membrane of some cells. However, for example of cyclic AMP releases the C subunit (with the myristoyl blocking group), while the regulatory subunit (containing an N-acetyl blocking group) remains membrane bound. It is also possible that the myristoyl group is involved in the translocation of the C subunit from the cytoplasm to the nucleus, across the nuclear membrane. The discovery of myristic acid at the N-terminus of a variety of viral proteins [14–16] has greatly increased the number of proteins known to have this N-terminal modification.

Some viral tyrosine-specific protein kinases have also been shown to contain myristoyl-N terminal groups ([14]; Table 1). The myristoylation in the capsid protein of all members of the picornaviruses has revealed a role for this post-translational modification in viral capsid assembly [15].

Studies on an 87 kDa phosphoprotein that is also myristoylated from macrophages have suggested a role for this modification in altering association of this protein with cellular membranes in response to lipopolysaccharide [17] which would enhance its phosphorylation by protein kinase C.

The finding that this N-terminal blocking group is almost exclusively myristic acid would indicate that the enzyme acylating these proteins has a high specificity for myristoyl-CoA. The N-terminal sequences of proteins known to contain myristoylated blocking groups are identical in having glycine as the first residue (Table 1). This myristoyl N-transferase [1] seems to recognize little consensus sequence beyond this. There must, however, be some other determinant, since many proteins contain N-terminal glycine residues. In some cases these are N-acetylated [18].

Isolation of peptides containing acyl and alkyl groups

The very hydrophobic nature of this type of post-translational modification was made quickly apparent by the technique of reverse-phase h.p.l.c. To isolate myristoylated peptides, for example, h.p.l.c. has been employed, at ambient temperatures, using Waters pBondapack C_18 columns ([4, 5]; see Fig. 1).

Vydac C_8 columns [19] have also been used where the myristoylated peptide eluted at an acetonitrile concentration of 37%.

The myristoylated peptides from NADH-cytochrome-b_5 reductase eluted as sharp peaks at an apparent concentration of 50% (v/v) propan-1-ol on a Synchronpak RP-P column [21].

In the case of the catalytic subunit of cyclic AMP-dependent protein kinase, the peptides were redissolved in a small volume of 6 M-guanidine/HCl, and eluted with water/acetoni-trile containing 0.1% (v/v) trifluoroacetic acid with an increase of 1% acetonitrile/min. The peptide was detected by u.v. absorption at 210 nm [20]. The predicted elution positions of identical undervirizated, acetylated and palmitoylated peptides are indicated, emphasizing the large effect of this modification on chromatographic properties.

![Fig. 1. Reverse-phase h.p.l.c. of myristoylated N-terminal peptide from calcineurin subunit B](image)

The decapeptide from cyanogen bromide digestion of this protein was subjected to h.p.l.c. on a μBondapack C_18 reverse-phase column (Waters) on a gradient of water/acetoni-trile [in 0.1% (v/v) trifluoroacetic acid] with an increase of 1% acetonitrile/min. The peptide was detected by u.v. absorption at 210 nm [20]. The predicted elution positions of identical undervirizated, acetylated and palmitoylated peptides are indicated, emphasizing the large effect of this modification on chromatographic properties.

<table>
<thead>
<tr>
<th>Protein function</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcineurin, subunit B (protein phosphatase)</td>
<td>myr-GNEASYPPL</td>
</tr>
<tr>
<td>Cyclic AMP-dependent protein kinase (Ser/Thr kinase)</td>
<td>myr-GNAAKK</td>
</tr>
<tr>
<td>Rous Sarcoma Virus p60^coy (tyrosine kinase)</td>
<td>myr-GSSKSKPK</td>
</tr>
<tr>
<td>Murine lymphoma p56 (tyrosine kinase)</td>
<td>myr-GCVSSNP</td>
</tr>
<tr>
<td>G protein, Go a-subunit (signal transduction/GTPase)</td>
<td>myr-GCTLSAEL</td>
</tr>
<tr>
<td>Torpedo 43 Da (acetylcholine receptor-associated protein)</td>
<td>myr-GDQTK</td>
</tr>
<tr>
<td>NADH-cytochrome-b_5 reductase (stearyl-CoA desaturase)</td>
<td>myr-GAQLSILG</td>
</tr>
<tr>
<td>gog Proteins (virus structure/assembly), e.g. MuLV p15^{env}</td>
<td>myr-GQTVTTPPL</td>
</tr>
<tr>
<td>Picornavirus VP4 (capsid assembly/viral entry), e.g. Poliovirus</td>
<td>myr-GAVSSQK</td>
</tr>
<tr>
<td>Hepatitis B pre-S1 (virus assembly/structure)</td>
<td>myr-GQQSTSNS</td>
</tr>
<tr>
<td>Polymavirus VP2 (virus assembly/structure)</td>
<td>myr-GAALTILV</td>
</tr>
<tr>
<td>HIV nef (reduces expression/GTP binding)</td>
<td>myr-GHKWSFS</td>
</tr>
</tbody>
</table>

Abbreviations: myr-, myristoyl-, i.e. CH_3(CH_2)_{14}CO—. Amino acids are given in the standard single letter code.

Table 1. Proteins known to be myristoylated

BIOCHEMICAL SOCIETY TRANSACTIONS
achieved with a Hypersil ODS (octadecylsilane) reverse-phase column eluting with 80\% (v/v) methanol [8].

In their study of the structure of the glycolipid tail in the Thy-1 glycoprotein, Tse et al. [23] used two-dimensional t.l.c. to purify the C-terminal peptide. This avoided subsequent problems of contamination with the Brij detergent necessary to solubilize the peptides.

Structural analysis by m.s.

F.a.b.-m.s. has proved the most useful mass spectrometric technique for the identification of myristic acid and other N-terminal blocking groups (acetyl, pyroglutamate, etc.) in peptides. This is particularly useful where sequence information is also required, since the Edman degradation chemistry cannot proceed without a free NH₂ group [24].

Mass spectrometers that have been widely used for study of peptides include Kratos MS50, VG ZAB-IF, VG70-250SE and Finnegan-MAT instruments, and details of the particular mode of operation will be found in the appropriate references. Iontech f.a.b. sources have been used to generate 6-8 kV xenon beams.

Fig. 2. F.a.b.-m.s. of myristoylated peptide from N-terminus of calcineurin subunit B

This is the actual spectrum obtained on a Kratos MS50 mass spectrometer in the negative ion mode [5]. A standard Kratos f.a.b. source was employed to generate a 4-6 kV xenon beam. Samples were dissolved in 1 \( \mu \)l of a \( \alpha \)-thioglycerol/diglycerol matrix (1:1, by vol.), and the mixture was introduced into the source on a copper probe tip. With this matrix, the sensitivity may be improved by about an order of magnitude compared with the use of a glycerol matrix, and has proved particularly suitable for non-polar peptides. The sequence of the peptide \((M-H)^- = 1270\) was deduced as follows:

\[
\text{CH}_3\text{(CH}_2\text{)}_{10}\text{Gly Asn Glu Ala Ser Tyr Pro Leu Glu Hsl} ^{582} ^{832} ^{929} ^{1042} ^{1171} ^{1270} ^{1003} ^{889} ^{760} ^{744} ^{689}
\]

Bond cleavages at each of the three possible positions in the peptide backbone may occur, i.e.

\[
\text{NH} - \text{CH-CO-NH-CH-CO-}
\]

With f.a.b.-m.s. (reviewed in Williams et al. [25]), structural information on peptides up to about 20-30 residues may be obtained. With the introduction of the caesium ion gun (VG Analytical Ltd), this has extended the possibility of obtaining molecular mass information on much larger fragments (106 amino acid residues for example; [26]). In f.a.b.-m.s., prior chemical derivatization is not necessary. Involatile and thermally labile compounds can be studied by this method and very small amounts of material can be used. Molecular mass determinations may in favourable cases be made with picomolar levels of peptides. Matrix compounds (used to introduce the sample to the probe) that have proved particularly useful for all types of peptides include glycerol, thioglycerol and mixtures of these. Frequently, a small amount of acetic or trifluoroacetic acid is included in this matrix to render the peptide more hydrophobic by 'ion suppression'. Depending on the mode of ionization and fragmentation of an individual peptide, identification of either positive or negative ions may yield more information.

An alternative matrix, 3-nitrobenzyl alcohol, has recently proved particularly useful for m.s. of peptides (for example,
in the structure determination of the farnesyl moiety in yeast a-factor [8].

In the example of calcineurin B-subunit, f.a.b. mass spectra of esterified and non-esterified myristoyl peptides were obtained on a triplequad from a Staphylococcus aureus V8 protease digest and on a decapeptide from eucalyptus bromide cleavage. Peptides were esterified with 15 mM methanolic HCl for 25 h at ambient temperature. When the number of carboxylic acid groups was determined from the increase in M$_r$ of the esterified peptide, the M$_r$ of the blocking group could be calculated from the known amino acid composition. In the case of calcineurin B and the C-subunit this was 211, corresponding to CH$_3$(CH$_2$)$_{14}$CO$^-$ The mass spectrum of the decapeptide from the B-subunit of calcineurin B containing the myristoyl group is shown in Fig. 2. From the fragment ions obtained, the complete sequence of the peptide could be deduced. This result emphasizes the great advantage of f.a.b.-m.s. in the study of N-terminally blocked peptides and the possibility of obtaining the sequence from the C-terminus of a peptide. In identifying the myristoylation-blocking group of the C-subunit, Carr et al. [4] also used direct chemical ionization with ammonia as the reagent gas.

Identification of the fatty acid

Initially, to identify the nature of the acylation, protein from H-labelled cell cultures is normally separated by SDS polyacrylamide-gel electrophoresis (PAGE), located by fluorography, and extracted by HCl hydrolysis. Proteins may be pretreated with NH$_2$OH to specifically identify N-acylations, since fatty acids in S- and O-linked esters are labile to 1 M-hydroxyamine (60-20 h, 23°C, pH 7, 8 or 10) or 0.1 M-methanolic KOH (90 min, 23°C). In contrast, amide-linked fatty acids are released by strong acid hydrolysis conditions, as indicated by the H$_2$SO$_4$ (88°C for 4 h) or 0.1 M-methanolic HCl. Hydrolysis of farnesyl thioethers is achieved by treatment with methyl iodide at room temperature, pH 3.5, followed by mild alkali treatment in sodium bicarbonate [12].

It is essential to confirm the exact nature of the fatty acid from [14] labelled cell cultures is normally separated by HPLC on silica plates, either as methyl esters or free acids, still finds some application. However, reverse-phase t.l.c. of the fatty acid on Whatman KC18, by ascending chromatography with acetic acid/acetone (1:1, by vol.) [32] is now more widely used.

The biological importance of the covalent modifications described above is now clearly established and it is certain that many more examples will be found as well as novel modifications. The isolation of the acylated and alkylated peptides is facilitated by techniques such as h.p.l.c. (where their hydrophobicity allows separation from unmodified peptides) and their structures determined by a combination of m.s. and microsequencing. Exciting developments include continuous flow (or ‘dynamic’) f.a.b.-m.s., which should permit great advances in sensitivity and the discovery of new examples of post-translational modifications.


1989
LIPID MODIFICATION OF PROTEINS


Post-translational processing of ras proteins

A. I. MAGEE
Laboratory of Eukaryotic Molecular Genetics, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

Introduction

Post-translational modification of proteins with lipid-derived substituents is now recognized as a widespread mechanism for changing the properties of proteins, often, but not always, with the result of increasing their affinity for membranes [1]. Among the range of proteins known to be modified by the covalent attachment of fatty acids are several transforming proteins, including the products of the src and ras oncogenes. We have engaged in the elucidation of the detailed mechanism of ras protein modification and the subsequent effects on protein properties and function. During the course of these studies, we have found that the previous concepts of ras protein processing were over-simplified and we have identified a novel hydrophobic modification additional to fatty acylation. A systematic study of the protein sequences responsible for signalling hydrophobic modification may be important in targeting many low molecular mass GTP-binding proteins and other proteins to distinct intracellular membrane locations.

The ras protein family

In mammals the ras oncogene family contains three members called N-, Harvey (H-) and Kirsten (K-) which code for highly homologous proteins of M, ~21 000 called p21 [2]. The K-ras gene can code for two products differing only in their C-terminal regions by virtue of splicing to two alternative fourth exons (4A and 4B); ras homologues have been found in all species tested, including yeasts. In addition, a number of related genes of varying degrees of sequence similarity have been detected by low-stringency hybridization. ras oncogenes are expressed in most adult tissues at varying levels and probably play a role in normal cellular proliferation and differentiation. However, they can become transforming either by point mutations in a number of key regions, notably around residues 12 and 59, or by overexpression of the normal gene. The detailed mechanism of action is unknown, but a key feature of ras and related proteins is that they can bind and hydrolyse GTP. Much effort has thus gone into studying the possibility that ras proteins act as membrane-bound 'G-proteins' which transduce signals from growth factor receptors into intracellular second messengers. This hypothesis has received conflicting evidence, but it is clear that ras proteins must be membrane-bound to exert their effects. It is the mechanism of this membrane binding which has been the subject of our studies.

Membrane binding of ras proteins and the CAAX Box

The primary structure of the ras proteins is not suggestive of a direct membrane binding. Sefton et al. [3] showed that the H-ras protein contained a covalently attached palmitic acid moiety which was subsequently shown to be in ester linkage, probably to a cysteine thiol. This suggested a role for acylation in membrane attachment and this idea was reinforced by the finding that removal by site-directed mutagenesis of Cys-186, three residues from the C-terminus, abolished acylation, membrane-binding and transformation. Thus it was believed that attachment of palmitic acid to Cys-186 was responsible for membrane association.

In 1986 Powers et al. [5] isolated a Saccharomyces cerevisiae mutant which suppressed the action of a transforming ras gene. It was found that ras protein in these cells was not acylated and failed to associate with membranes, and the gene responsible for this phenotype was called RAM. An allelic gene (dpr1) was isolated independently by Tamanoi and co-workers [6]. Interestingly, the same mutant cells failed to secrete the 'a mating factor although it was synthesized normally. The common element between the a-factor and ras proteins is the presence of a homologous cysteine residue near the C-terminus, Cys-186 in ras proteins. This cysteine occurs in the sequence Cys-A-A-Xaa (A, aliphatic; Xaa, any amino acid) which was proposed to signal the post-translational processing events and has been called the 'CAAAX Box' [7]. The effect of the RAM/dpr1 mutation is further evidence for the importance of Cys-186 in p21 ras processing.

Elucidation of the processing pathway

The simplest interpretation of the above data, that Cys-186 was the site of p21 ras acylation, was challenged when H-ras genes containing deletions upstream of Cys-186, but leaving the CAAX box intact, were found to be defective in acylation, but still showed partial membrane association and function [8]. A series of site-directed mutants which selectively removed one or both of two nearby upstream cysteines (positions 181 and 184) demonstrated that both of these residues were capable of acting as acylation sites (Fig. 1). N- and K(4A)-ras proteins contain single homologous cysteines (positions 181 and 180, respectively) which are also acylated, while K(4B)-ras does not.

Received 8 May 1989

Abbreviations used: Pal, palmitic acid; Far, farnesol.

Vol. 17