Characterization of a human MHC class III region gene product with S-thioesterase activity

Begoña AGUADO[1] and R. Duncan CAMPBELL[1,2]

MRC Immunochemistry Unit, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, U.K.

Palmitoylated proteins contain a 16-carbon saturated fatty acyl group that is post-translationally attached by a labile thioester bond. These modified proteins are mainly membrane-bound; the lability of the thioester bond allows the process to be reversible, a unique property of this modification. We report here that the gene for G14, located in the class III region of the human MHC, encodes a polypeptide with significant sequence similarity to mammalian palmitoyl protein thioesterase (PPT1), an enzyme that removes palmitate from palmitoylated proteins. The gene for G14, also known as PPT2, is transcribed as at least five different transcripts, which are expressed in different cell lines of the immune system. Immunoprecipitation of these mammalian cells, with an anti-G14 antiserum, showed a specific band of approx. 42 kDa in cell extracts and supernatants. Expression of the G14 cDNA in the baculovirus system revealed that it encoded a secreted glycosylated polypeptide with S-thioesterase activity. The enzymic activity of the recombinant G14 protein was further characterized in quantitative spectrophotometric assays, which revealed that it had the highest S-thioesterase activity for the acyl groups palmitic and myristic acid followed by other long-chain acyl substrates. The S-thioesterase activity of the G14 protein was found to be considerably higher in supernatants than in cell extracts, which was consistent with the protein's being secreted. The G14 polypeptide contains, in addition to an N-terminal lipase domain, a C-terminal domain common to the cytokine receptor superfamily, which might determine the substrate specificity and/or the protein target of the G14 protein.

Key words: major histocompatibility complex, palmitoyl protein thioesterases 1 and 2.

INTRODUCTION

Multiple eukaryotic proteins are processed post-translationally by the addition of fatty acids or prenyl groups; this seems to have a key role in their being anchored to membranes, a process necessary for the diverse cellular functions of these modified proteins, including signal transduction, vesicle transport and maintenance of the cytoarchitecture (reviewed in [1–3]). Whereas myristoylation and prenylation are irreversible protein modifications, the addition of palmitate, via a labile thioester bond to cysteine residues, is a reversible event. Such a dynamic palmitoylation–depalmitoylation cycle suggests a regulatory role for this modification, which is emphasized by the fast turnover of the protein-bound palmitate. Many of the palmitoylated proteins have been implicated in signal-transduction pathways. For example, H-Ras undergoes prenylation and palmitoylation, both of which are required for its correct membrane attachment, and the α-subunits of heteromeric G (Gα) proteins are both myristoylated and palmitoylated. However, little is known about the enzymes responsible for palmitoylation and depalmitoylation. A more appropriate term for this modification would be S-acylation, because other fatty acyl chains can substitute for palmitoyl. It is likely that multiple acyltransferases exist with varying substrate specificities (reviewed in [1–3]).

A palmitoyl protein thioesterase (PPT) [4] was purified from bovine brain on the basis of its ability to remove palmitate from palmitoylated H-Ras produced in insect cells. In addition to H-Ras, the purified PPT accepted palmitoylated G_α and long-chain fatty acyl-CoA as its substrates *in vitro* [4]. However, cloning and expression of the rat and bovine proteins [5] revealed the presence of a signal peptide in the amino acid sequence and the secretion of the protein to the extracellular medium. For this reason it was suggested that intracellular proteins might not constitute the main substrates *in vivo* for the PPT enzyme [5]. Human PPT (hPPT) was also cloned [6] and the predicted polypeptide showed 91 % identity with the bovine PPT and 85 % identity with the rat PPT.

Mutations in the gene for hPPT have been identified as causing an autosomal recessive neurodegenerative disorder in childhood, infantile neuronal ceroid lipofuscinosis (INCL), which has a global incidence of 1 in 12 500 [6]. INCL is characterized by early vision loss and mental deterioration, and leads to a vegetative state in the patients by 3 years of age; death occurs at 8–11 years [6]. It has been reported that hPPT has a lysosomal localization and that the secreted protein is internalized into the lysosome through the mannose-6-phosphate receptor [7,8]. In contrast, a mutated hPPT is retained in the endoplasmic reticulum [6,8]. INCL patients display accumulation of lipid thioesters in immortalized lymphoblasts; this accumulation can be reversed in cultured cells by the addition of recombinant hPPT [9].

Recently, Soyombo and Hofmann [10] have described the molecular cloning and expression of another PPT (PPT2). In their preliminary characterization of PPT2 they reported that it does not remove palmitate groups from palmitoylated proteins such as H-Ras and albumin, which are substrates for the previously described PPT, PPT1. They also showed that, in

Abbreviations used: Nbs_2, 5,5′-dithiobis-(2-nitrobenzoic acid); FCS, foetal calf serum; FNIII, fibronectin type III; h, human; IFN-γ, interferon γ; INCL, infantile neuronal ceroid lipofuscinosis; LPAATα, lysophosphatidic acid acyltransferase α; LTBP, latent transforming growth-factor-β-binding protein; PPT, palmitoyl protein thioesterase; RT–PCR, reverse-transcriptase-mediated PCR.

[1] Present address: UK HGMP Resource Centre, Hinxton, Cambridge CB10 1SB, U.K.

[2] To whom correspondence should be addressed (e-mail rcampbel@hgmp.mrc.ac.uk).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number Y17958.

cross-correction experiments, PPT2 does not abolish the accumulation of protein-derived lipid thioesters in PPT-deficient cell lines.

Here we present a detailed characterization of PPT2 (*G14* gene product) and we also discuss structural features of the protein. We have found that the *G14* gene product is expressed in cells of the immune system as an approx. 42 kDa protein in cell extracts and supernatants, and is transcribed as at least five different transcripts. We have expressed the G14 cDNA in insect cells by using the baculovirus system and have found, with different acyl-CoAs as substrates, that it contains specific S-thioesterase activity. This activity was considerably higher in supernatants than in cell extracts, which is consistent with the finding that the protein is secreted.

The gene for G14 is located in the class III region of the human MHC, approx. 280 kb from the class II region, very close to the recently described lysophosphatidic acid acyltransferase (LPAATα) gene [11–13] and between the gene coding for a protein similar to the latent transforming growth-factor-β-binding protein (LTBP) (Genome Sequence Database accession number U89336) and the *Crep-rp* gene [14,15]. The MHC class III region contains several genes encoding proteins with potential roles in the immune system and in inflammation [16,17].

EXPERIMENTAL

cDNA sequence and computer analysis

Screening of a U937 cDNA library, with the use of the insert of the MHC class III region cosmid clone D3A as a probe, resulted in the isolation of 23 plasmid clones containing the G14 cDNA. Characterization of these clones by restriction-enzyme mapping revealed that pG14-6 contained a full-length cDNA insert of approx. 2.0 kb [18]. The insert of the pG14.6 plasmid was sequenced by the dideoxy chain-termination method after the cloning of random sonicated fragments in the size range 300–1000 bp by blunt-end ligation into *Sma*I-cut M13mp18. The sequence was assembled by using the Staden [19] programs and determined with a degeneracy of 8.1. Computer analyses were performed with the software package of the University of Wisconsin Genetics Computer Group (GCG) [20].

Cell culture

The suspension cell lines U937 (human promonocytic leukaemia), Molt4 and Jurkat J6 (human T-cell leukaemia), HL60 (human promyelocytic leukaemia), K562 (human erythroleukaemia), Raji (human Burkitt's lymphoma) and B1518 (human B lymphoblastoid) were grown in RPMI-1640 (Sigma). The adherent cell lines TK$^-$143B (143B human osteosarcoma, thymidine kinase negative) and HeLa (human epithelioid carcinoma) were grown in Dulbecco's modified Eagle's medium (Sigma). All the media were supplemented with 2 mM L-glutamine (Sigma), 100 i.u./ml penicillin, 100 μg/ml streptomycin (Sigma) and 10 % (v/v) heat-inactivated foetal calf serum (FCS). Suspension cells were stimulated with 10 ng/ml PMA (Sigma) and adherent cells were stimulated with 200 i.u./ml interferon γ (IFN-γ) (R&D Systems). The stimulation was performed by adding the compounds to the growth medium for 24 h at 37 °C. Sf21 cells were grown at 28 °C in TC100 medium (Gibco) supplemented with antibiotics and 10 % (v/v) heat-inactivated FCS (Gibco).

Transcriptional analysis

Total RNA was extracted by the RNAzol™ B method (Biotecs Laboratories) and poly(A)$^+$ RNA was prepared with the Quick Prep kit (Pharmacia) from the previously isolated total RNA. Samples of total RNA (20 μg) or poly(A)$^+$ RNA obtained from 20 μg of total RNA were fractionated in a 0.8 % agarose/formaldehyde denaturing gel and transferred to nitrocellulose membranes (Hybond C extra) [21]. Similar amounts of mRNA were detected for each cell line on the gels, by ethidium bromide staining, except for the cell line K562, which contained less mRNA, and for the HeLa cell line treated with IFN-γ, which contained more mRNA than untreated cells. Hybridization with ^{32}P-labelled probes was performed as described elsewhere [22]. The probe used was the G14 cDNA coding sequence insert from pG14Blsc (see below) labelled with ^{32}P with the Multiprime kit (Amersham) as described [22]. The Northern blot shown represents one experiment out of three performed. For reverse-transcriptase-mediated PCR (RT–PCR), the first-strand cDNA was synthesized by using 10 μg of total RNA and the reverse transcription kit (Promega). For PCR amplification, the specific 5′ oligonucleotide primer Ol1.G14 and the 3′ oligonucleotide primer Ol2.G14 were used (see below). PCR was performed in a final volume of 100 μl containing 1 μl of the newly reverse-transcribed cDNA. The amplified product (10 μl) was run directly on 1 % (w/v) agarose gels and Southern blot analysis was performed as described previously.

Baculovirus expression

The G14 coding sequence was first cloned into the plasmid pBluescript KS$^+$ (pBlsc) to create the master plasmid clone pG14Blsc. To remove the 5′ and 3′ flanking sequences of the G14 cDNA, a PCR copy of the open reading frame was generated by using oligonucleotide primers that also created *Xba*I sites adjacent to the initiating AUG codon and the stop codon. The oligonucleotide primers used were: forward Ol1.G14, 5′-GCGGTCTAGAGC**ATG**CTGGGGCTCTG-3′; reverse Ol2.G14, 5′-CGCCTCTAGATCC**TCA**GGAGAGCCAAGG-3′; the *Xba*I site is underlined and the initiation and stop (opposite orientation) codons are in bold. The PCR copy was gel-isolated and ligated into the *Xba*I-digested pBlsc vector and several clones were sequenced. To generate the G14 recombinant baculovirus, the insert of one clone (pG14Blsc.5), which did not contain any PCR errors, was excised, gel-isolated and ligated into the baculovirus transfer vector pAcCL29.1 [23] kindly provided by Dr. I. Jones (NERC Institute of Virology and Environmental Microbiology, Oxford, U.K.). The G14 insert was double-digested with *Xba*I–*Not*I, end-filled and cloned into *Sma*I-digested pAcCL29.1, to generate pG14Bac. *Spodoptera frugiperda* 21 (Sf21) cells were co-transfected with BacPAK6 DNA and pG14Bac to produce vG14Bac, as described [11]. The recombinant AcB15R has been described previously [24]. The metabolic labelling of proteins was performed as described [11].

Polyclonal antibodies

To raise polyclonal antibodies against the G14 polypeptide, approx. 10^8 Sf21 cells were infected, at a multiplicity of infection of 2 plaque-forming units per cell, with vG14Bac in a small volume of TC100 medium without FCS. After 4 days (or when the cytopathic effect was total), cells and medium were harvested and centrifuged to remove cells; the supernatant was concentrated 40-fold and dialysed against PBS at 4 °C with the use of Micro-ProDiCon (Bio-Molecular Dynamics) with PA-10 ProDiMen dialysis membranes (molecular mass cut-off 10 kDa). Two rabbits were immunized by multiple intradermal injections.

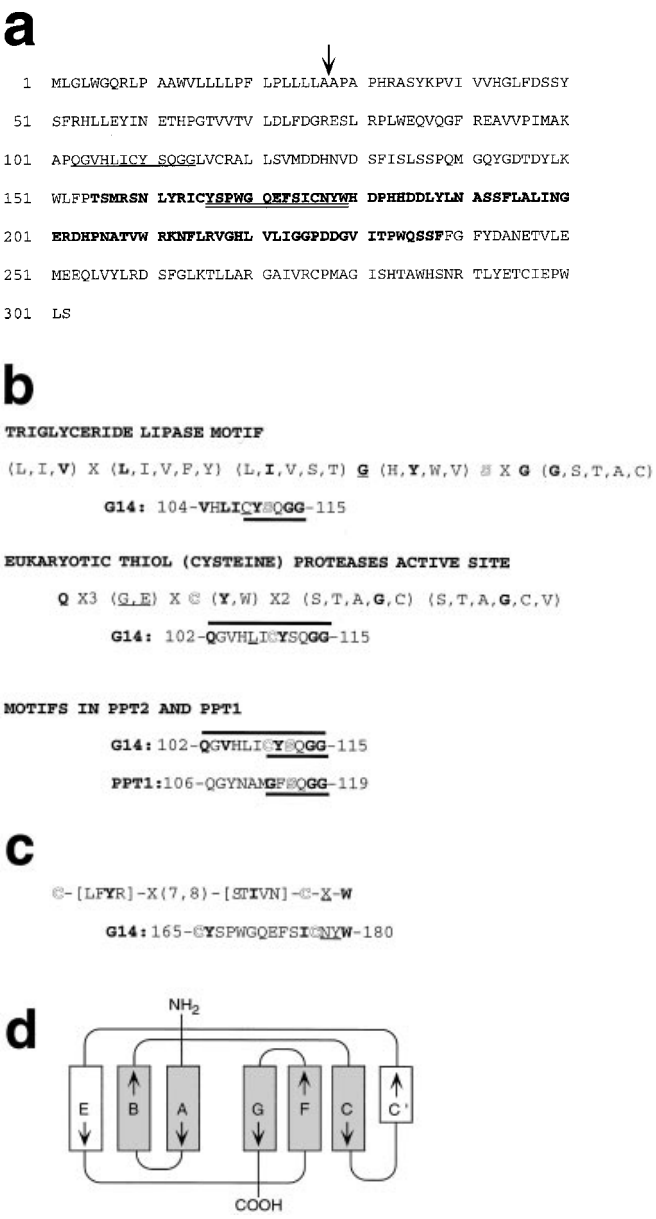

Figure 1 Sequence of the G14 polypeptide

(**a**) Amino acid sequence of the G14 polypeptide. The vertical arrow indicates the predicted cleavage site of the signal peptide. The lipase domain is underlined. The domain common to the cytokine receptor superfamily is represented in bold and the cytokine receptor superfamily motif is doubly underlined. (**b**) Triglyceride lipase motif and eukaryotic thiol (Cys) protease. The amino acid that constitutes the active centre, in each motif, is outlined. The amino acid that is different in the conserved motif is underlined. The pentanucleotide Gly-Xaa-Ser-Xaa-Gly (GXSXG) of the lipase motif is underlined. The alignments of the motifs present in the G14 polypeptide and PPT1 are shown, with the lipase motif underlined and the thiol motif overlined. (**c**) Growth factor and cytokine receptor family motif. The Cys residues conserved in the family are outlined. The two residues, which constitute the mismatch, between the last Cys and Trp residues are underlined. In (**b**) and (**c**) the G14 motifs are represented below the consensus motifs. (**d**) Topology of the N-terminal FNIII domain of the growth hormone receptor and the predicted topology of the G14 polypeptide. The seven β-strands are indicated by boxes with strands A, B and E forming one sheet and C, C′, F and G forming a second sheet. The five β-strands predicted in the G14 polypeptide are in grey.

Immunoprecipitation of cell extracts from insect and mammalian cell lines

Metabolic labelling of insect cells was performed as described elsewhere [11]. Similarly, for mammalian cell lines, 10^6 cells were labelled for 4 h at 37 °C with 500 μCi/ml Tran^{35}S-label (ICN Biomedicals; a mixture of approx. 80 % [^{35}S]methionine and approx. 20 % [^{35}S]cysteine, 1200 Ci/mmol) in Met-free minimal essential medium (Gibco). Medium or cells were incubated in immunoprecipitation (IP) buffer [10 mM Tris/HCl (pH 7.5)/150 mM NaCl/1 % (v/v) sodium deoxycholate/1 % (v/v) Nonidet P40/0.1 % SDS/2 mM unlabelled Met/1 mg/ml BSA/0.02 % NaN_3/1 mM PMSF]. Approx. $(0.5–10) \times 10^6$ c.p.m. (trichloroacetic acid-insoluble) of the metabolically labelled extracts in IP buffer were immunoprecipitated with Protein A–Sepharose [25] and analysed by SDS/PAGE [12 % (w/v) gel] followed by fluorography with Amplify (Amersham). Because the G14 polyclonal antibodies were obtained against the secreted protein expressed in the baculovirus system, these antibodies were preincubated with cell extracts or supernatants from infections with vG15Bac [11] to decrease the background.

Thioesterase assays

Sf21 cells were infected with the wild-type or vG14Bac baculovirus at a low multiplicity of infection (2 plaque-forming units per cell). The cells and media were harvested 72 h after infection (or when the cytopathic effect was total) and centrifuged to remove cells; the supernatant was concentrated 50-fold and dialysed against PBS at 4 °C with the Micro-ProDiCon as described above. The concentrated supernatant was stored at −70 °C. The cell pellet was resuspended in 50 mM Tris/HCl, pH 8, followed by four cycles of freeze–thaw and Dounce homogenization. These homogenates were centrifuged for 10 min at 400 ***g*** (2000 rev./min) to remove cell debris or large aggregates. The supernatant of this centrifugation, consisting of the cell extract, was divided into aliquots and stored at −70 °C.

A colorimetric enzyme assay for the measurement of G14 S-thioesterase activity was performed at room temperature essentially as described by Camp et al. [5], with minor modifications. In brief, the enzyme activity was assayed spectrophotometrically by measuring the reaction of the thiol group of the released CoA with 5,5′-dithiobis-(2-nitrobenzoic acid) (Nbs_2), giving an increase in A_{413}. A typical incubation mixture consisted of 40 mM Tris/HCl, pH 7.4, 0.5 mM Nbs_2, 4–100 μM acyl-CoA (Sigma) and 0.2–900 μg of cell extract homogenate or 1–500 μg of supernatant, in a volume of 1 ml. Nbs_2 was added as a 0.01 M solution in 0.1 M potassium phosphate buffer, pH 7.0. The reaction was initiated by the addition of the acyl-CoA after preincubation of the enzyme with all the other components at room temperature for 2 min. A molar absorbance coefficient of 13600 $M^{-1} \cdot cm^{-1}$ was used to calculate the activity. Different experiments with the same cell extract preparation and with myristoyl-CoA ($C_{14:0}$) as the acyl-CoA substrate were performed on different days; the results were consistent. To compare the specificity of the enzyme for different substrates, all the long-chain acyl-CoAs were used on the same day, all the short-chain acyl-CoAs were used on another day, and myristoyl-CoA was used on both days as a reference control.

A radiolabelled thioesterase assay was also performed. The cell lines TK$^-$143B and HeLa were radiolabelled for 8 h at 100 μCi/ml with [^{3}H]myristic acid (11.2 Ci/mmol; DuPont–New England Nuclear), for 4 h at 500 μCi/ml with [^{3}H]palmitic acid (56.5 Ci/mmol; DuPont–New England Nuclear), and for 16 h at 50 μCi/ml with [^{3}H]mevalonic acid. These radiolabelled cell extracts and supernatants were incubated in 50 mM Tris/HCl (pH 7.4)/1 mM EDTA for 1 h at 37 °C, with 15 μg of protein from cell extracts, or with 90 μg of protein from supernatants, of insect cells infected with wild-type baculovirus or vG14Bac in a

final volume of 20 μl. After incubation, the samples were centrifuged and sample buffer was added to stop the reaction. The samples were analysed by SDS/PAGE [12% (w/v) gel] followed by fluorography with En^{3}Hance (DuPont–New England Nuclear).

RESULTS

G14 protein

The putative 302-residue translation product of an almost full-length G14 cDNA was found to contain a hydrophobic leader peptide at the N-terminus and five potential N-linked glycosylation sites (Figure 1a), suggesting that the *G14* gene encoded a secretory glycoprotein. The predicted molecular mass of the G14 polypeptide was approx. 34.3 kDa including the signal peptide, and approx. 30 kDa after cleavage of the signal peptide. The FastA and BestFit programs from the GCG package revealed significant similarity between the predicted amino acid sequence of the *G14* gene product and the entire sequence of the human, rat and bovine enzyme PPT (PPT1) [5,6] (approx. 55% similarity and approx. 30% identity between the three species), suggesting that the *G14* gene product could encode a new member of this acyl-thioesterase enzyme family. Soyombo and Hofmann [10] recently described the sequence of a PPT2 enzyme and our sequence of the *G14* gene product is identical with theirs.

The G14 polypeptide was screened for amino acid motifs conserved in protein families by using the Motif program from the GCG package. Three motifs, with one mismatch each, were found. Of the two located at the N-terminus, one motif was common to triglyceride lipases and the other to eukaryotic thiol (Cys) proteases. Triglyceride lipases are lipolytic enzymes that hydrolyse the ester bond of triglycerides. The most conserved region in all these proteins is centred on a Ser residue (active-site residue) located in a conserved motif Gly-Xaa-Ser-Xaa-Gly (Figure 1b). The G14 protein contained a Cys residue instead of the first Gly in the motif. Other lipases with one mismatch in either of the consensus Gly residues have been described, such as subtilisin (Gly-Xaa-Ser-Xaa-Ala) [26]. The PPT1 amino acid sequence contains the Gly-Xaa-Ser-Xaa-Gly motif located at an equivalent amino acid position and it also contains a Gly-His-Asp motif at the C-terminus, characteristic of mammalian thioesterases, which was not present in the G14 polypeptide (Figure 1b [5,10], and results not shown).

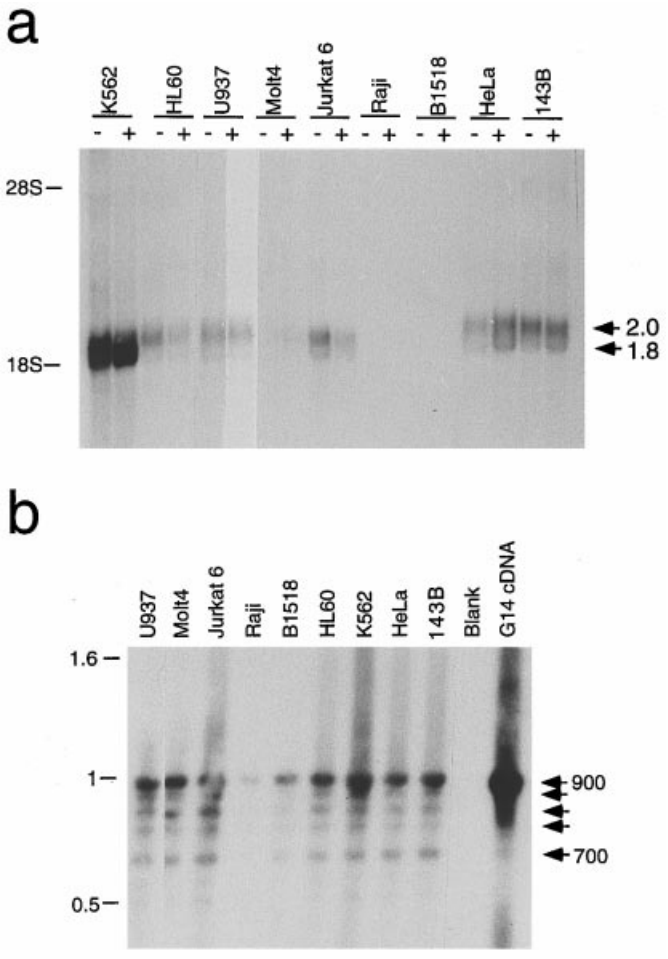

Figure 2 Transcriptional analysis of the *G14* gene by Northern blotting and RT–PCR

(**a**) Northern blot containing poly(A)$^+$ RNA extracted from the indicated cell lines treated (+) or not (−) with PMA (all except HeLa and 143B) or IFN-γ (HeLa and 143B). Note the two different G14 mRNA species of 1.8 and 2 kb (indicated with arrows). The positions of migration of 28 S and 18 S RNA are indicated at the left. (**b**) Southern blot analysis of the PCR amplification products of reverse-transcribed RNA obtained from the indicated cell lines. PCR controls were lacking DNA (blank) and plasmid pG14.6 DNA (G14 cDNA). The numbers at the left indicate the positions of DNA standards in kbp. The arrows at the right indicate the different G14 bands obtained with estimated lengths in bp.

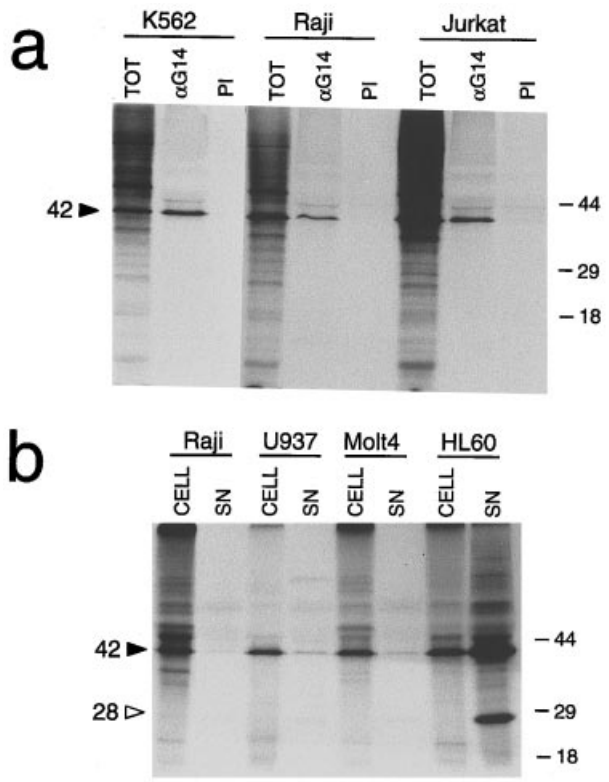

Figure 3 Immunoprecipitation of the G14 protein in mammalian cells

(**a**) Cell extracts (TOT) or the corresponding proteins immunoprecipitated with the anti-G14 antiserum (αG14) or with non-immune serum (PI). (**b**) Immunoprecipitations, with the anti-G14 antiserum, of radiolabelled cell extracts (CELL) or supernatants (SN) of the indicated cell lines. Proteins were analysed by SDS/PAGE and detected by fluorography. The molecular masses of the proteins (left) and molecular mass markers (right) are indicated in kDa. The black arrowhead indicates the 42 kDa specific G14 polypeptide and the white arrowhead the extra 28 kDa polypeptide found in supernatants.

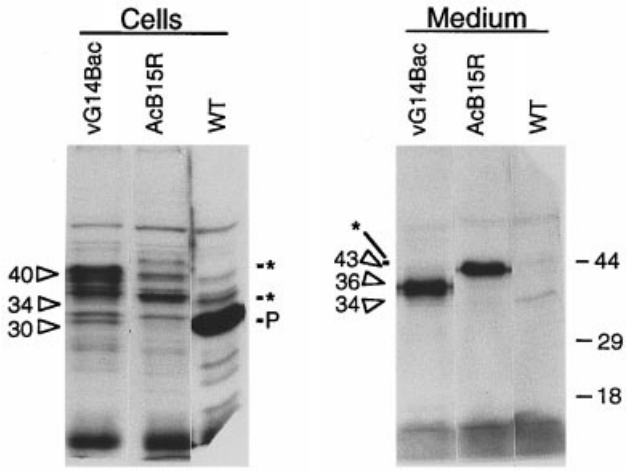

Figure 4 Expression of the *G14* gene product

(**a**) Sf21 cells infected with vG14Bac, AcB15R or wild-type (WT) baculovirus were pulse-labelled from 24 to 27 h after infection; proteins present in cells and in the medium were analysed by SDS/PAGE and detected by fluorography. The positions of the polyhedrin protein (P) and the vaccinia IL-1β receptor (*) are shown. The white triangles indicate the specific bands corresponding to the G14 protein. The molecular masses of the proteins (left) and the molecular mass markers (right) are indicated in kDa.

The G14 protein contained, in the same region as the lipase motif, a motif common to the active site of eukaryotic thiol (Cys) proteases (Figure 1b), a family of proteolytic enzymes that contain an active-site Cys residue. The residue at position 5 of the pattern is always Gly (with one exception); in the G14 polypeptide it was Leu (Figure 1b). This motif is not present in PPT1 because the Cys residue in the G14 polypeptide corresponds to the mismatched first Gly of the lipase consensus sequence (Figure 1b).

In addition, the amino acid sequence of the predicted G14 polypeptide showed, at the C-terminus, a motif common to the growth factor and cytokine receptors family (Figure 1c). The growth-factor receptors are type 1 membrane-spanning proteins and in the extracellular domain have two fibronectin type III (FNIII) domains, with the N-terminal FNIII domain containing four Cys residues spaced approx. 9, 27 and 15 residues apart. The C-terminal FNIII domain contains a Trp-Ser-Xaa-Trp-Ser (WSXWS) motif positioned just proximal to the membrane [27,28], which is not present in the G14 polypeptide because it lacked the complete C-terminal FNIII domain (results not shown).

Using the Peptidestructure program in the GCG package, the G14 polypeptide was predicted to have at the N-terminus α-helices and β-strands, and at the C-terminus (from residue 140) only β-strands (results not shown). The first five C-terminal β-strands predicted in the G14 polypeptide coincided with the position of five of the seven β-strands, which constitute a FNIII domain, in the cytokine receptor superfamily (Figure 1d) [27,28].

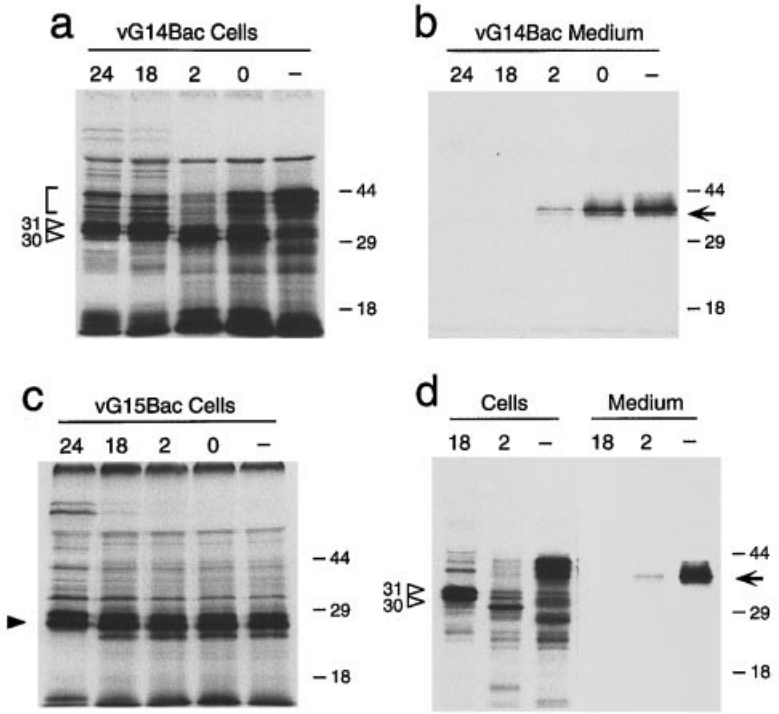

Figure 5 Effect of tunicamycin on the G14 protein expression

Sf21 cells infected with vG14Bac (**a**, **b**) or vG15Bac (**c**) were untreated (—) or treated with tunicamycin for 0, 2, 18 or 24 h (as indicated above each lane) before they were pulse-labelled from 24 to 27 h after infection in the presence of tunicamycin (2 μg/ml). Proteins present in cells (**a**, **c**) and in the medium (**b**) were analysed by SDS/PAGE and detected by fluorography. (**d**) Immunoprecipitation of the G14 polypeptide present in cells and in the medium of vG14Bac infections after treatment or not with tunicamycin for the indicated durations. The white triangles indicate the G14 specific bands and the black triangle in (**c**) the LPAATα specific bands in cell extracts. The arrows indicate the G14 specific secreted bands. The molecular masses of the proteins (left) and the molecular mass markers (right) are indicated in kDa.

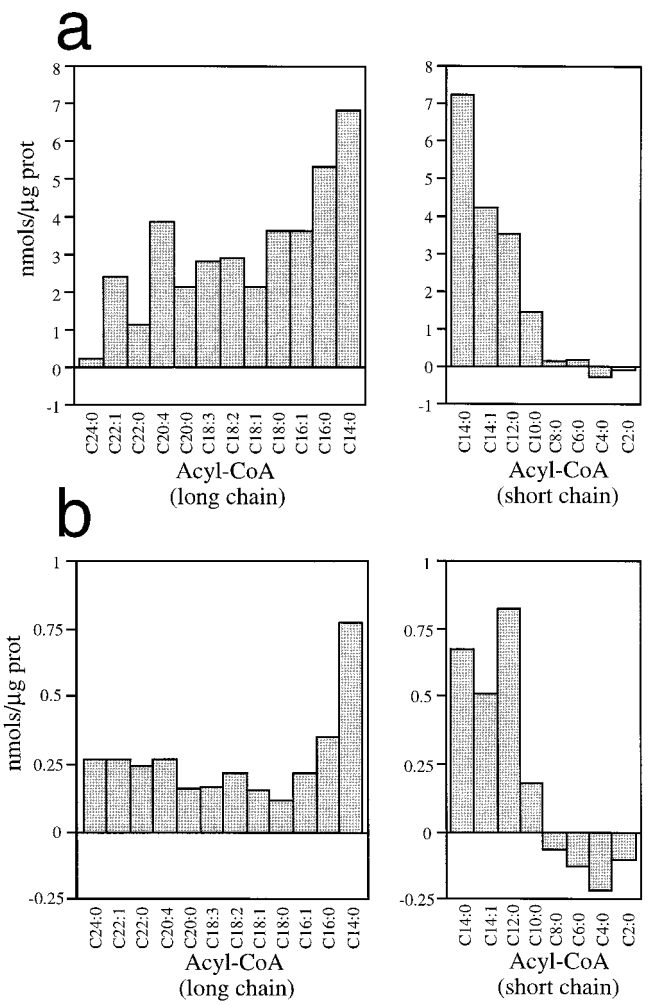

Figure 6 Spectrophotometric S-thioesterase assay for the G14 protein

Equivalent amounts (1.66 μl) of supernatants (**a**) or cell extracts (**b**) from vG14Bac infections were used. The difference, after 5 min of assay, between the activity obtained when using extracts from vG14Bac infections and extracts from wild-type baculovirus infections is shown. The long-chain acyl-CoAs [myristoyl ($C_{14:0}$), palmitoyl ($C_{16:0}$), palmitoleoyl ($C_{16:1}$), stearoyl ($C_{18:0}$), oleoyl ($C_{18:1}$), linoleoyl ($C_{18:2}$), linolenoyl ($C_{18:3}$), arachidoyl ($C_{20:0}$), arachidonoyl ($C_{20:4}$), behenoyl ($C_{22:0}$), eurocoyl ($C_{22:1}$) and lignoceroyl ($C_{24:0}$)] and short-chain acyl-CoAs [myristoyl ($C_{14:0}$), myristoleoyl ($C_{14:1}$), lauroyl ($C_{12:0}$), decanoyl ($C_{10:0}$), octanoyl ($C_{8:0}$), hexanoyl ($C_{6:0}$), butyryl ($C_{4:0}$) and acetyl ($C_{2:0}$)] used in the assay are shown.

Interestingly, PPT1 does not contain the cytokine receptor superfamily motif, and the Peptidestructure program did not predict all the β-strands found at the C-terminus of the G14 polypeptide (results not shown).

Analysis of the nucleotide sequence of the G14 cDNA showed that it contained an incomplete CTG repeat in the segment encoding a stretch of Leu residues in the signal peptide (Figure 1a). However, the same region in PPT1 contained a more incomplete repeat, which is reflected in fewer Leu residues in the signal peptide (results not shown).

Expression of the *G14* gene

The expression of the *G14* gene was analysed in different mammalian cell lines by transcriptional analysis (Northern blotting and RT–PCR) and by immunoprecipitation of the G14 polypeptide. Northern blot analysis of total RNA (results not shown) or poly(A)$^+$ RNA (Figure 2a), extracted from different mammalian cell lines, showed that the *G14* gene is expressed as different mRNA species. RNA from the cell lines K562, HeLa and TK$^-$143B contained 1.8 and 2 kb mRNA species. However, RNA from the cell lines Molt4 and HL60 seemed to contain only the 2 kb mRNA species. The predominant mRNA species in the cell lines U937 and Jurkat 6 was the 2 kb mRNA, but the 1.8 kb mRNA could also be detected (Figure 2a). Interestingly, the B-cell lines (Raji and B1518) showed no band at all, although in longer exposures a faint band of 2 kb cound be detected in the cell line B1518 (Figure 2a, and results not shown). No induction of G14 mRNA was apparent in cells treated with PMA or IFN-γ (Figure 2a). In addition, RT–PCR analysis of RNA from the same cell lines showed five different bands ranging in size from approx. 700 to approx. 900 bp (Figure 2b); the approx. 900 bp band was the most intense. The exact amplification product predicted from the use of this oligonucleotide primer pair on the complete cDNA is 934 bp. The same bands were observed in all the cell lines with similar intensities except the cell line B1518, for which all the products were slightly fainter, and the cell line Raji, which showed only a faint approx. 900 bp product (Figure 2b); however, much longer exposures (results not shown) showed the other bands. This pattern of expression in various cell lines was in agreement with the results obtained by Northern blot analysis with the approx. 700 to approx. 900 bp bands corresponding to the two bands detected by the latter method.

The expression of the G14 protein in different mammalian cell lines was analysed by immunoprecipitation of metabolically labelled cell extracts with G14 polyclonal antibodies. The human cell lines were chosen according to the G14 mRNA pattern of expression (Figure 2): high (K562), medium (U937, Jurkat 6 and HL60), low (Molt4) and very low (Raji). Immunoprecipitations of all the radiolabelled cell extracts showed a strong band at 42 kDa (Figure 3) that was not present in immunoprecipitates with the use of the preimmune serum (Figure 3a). The 42 kDa band was also detected in immunoprecipitates with the anti-G14 antiserum by using the radiolabelled supernatants, clearly in HL60 cells and, after long exposures of the autoradiographs, in the other cell lines (Figure 3b, and results not shown). Interestingly, a second band of approx. 28 kDa was detected in HL60 supernatants, but not in cell extracts, that was just visible in supernatants from other cell lines (Figure 3b) and might represent a product of one of the spliced forms of the *G14* gene.

Expression of the G14 polypeptide in insect cells by using the baculovirus system

To characterize the *G14* gene product, the recombinant baculovirus vG14Bac was constructed for expression of the protein in Sf21 cells. Pulse-labelling experiments with Tran^{35}S-label detected, in infected insect cell extracts, a 30 kDa G14 polypeptide and several G14 polypeptides with molecular masses between 34 and 40 kDa (Figure 4). The G14 protein was secreted into the medium by insect cells as 34 and 36 kDa polypeptides (Figure 4), although it was also possible to detect very faint additional bands up to 43 kDa, which might correspond to the G14 polypeptide with different levels of glycosylation. As a negative or positive control for expression and secretion, insect cells were infected with the wild-type baculovirus or a recombinant baculovirus expressing the vaccinia IL-1β receptor (AcB15R) [24] respectively (Figure 4).

To analyse the possible N-glycosylation of the polypeptide, which could account for the different bands, radiolabelling was performed in the presence of tunicamycin. Cell extracts from cells infected with vG14Bac and treated with tunicamycin for 2 h

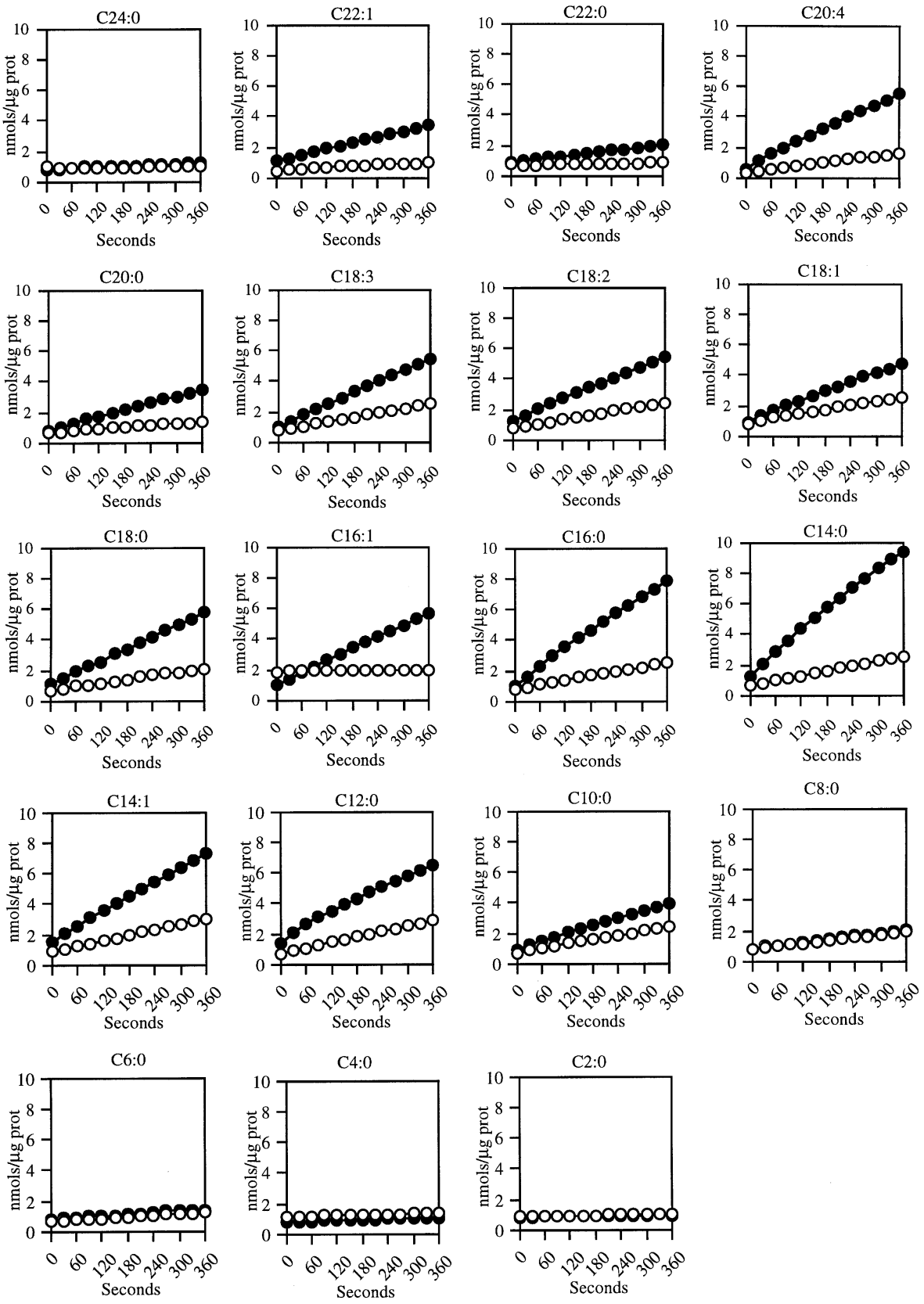

Figure 7 Spectrophotometric S-thioesterase assay with 1.66 μl (50 μg) of supernatants from vG14Bac infections (●) or wild-type virus infections (○)

The acyl-CoAs used were: lignoceroyl ($C_{24:0}$), eurocoyl ($C_{22:1}$), behenoyl ($C_{22:0}$), arachidonoyl ($C_{20:4}$), arachidoyl ($C_{20:0}$), linolenoyl ($C_{18:3}$), linoleoyl ($C_{18:2}$), oleoyl ($C_{18:1}$), stearoyl ($C_{18:0}$), palmitoleoyl ($C_{16:1}$), palmitoyl ($C_{16:0}$), myristoyl ($C_{14:0}$), myristoleoyl ($C_{14:1}$), lauroyl ($C_{12:0}$), decanoyl ($C_{10:0}$), octanoyl ($C_{8:0}$), hexanoyl ($C_{6:0}$), butyryl ($C_{4:0}$) and acetyl ($C_{2:0}$).

or more before being labelled showed a decrease in the intensity of the G14-specific bands with an increase in the intensity of 30 and 31 kDa polypeptides (Figure 5a). However, cells infected with vG15Bac, expressing the unglycosylated LPAATα [11], did not show any difference even after the longest duration of tunicamycin treatment (Figure 5c). Supernatants harvested from insect cells infected with vG14Bac showed a decrease in intensity of the G14 bands, especially the faint bands of higher molecular mass, even when the tunicamycin was added only during the labelling. A total lack of secretion of the G14 protein was observed in supernatants from infections performed in the presence of tunicamycin for 18 or 24 h (Figure 5b). These cell extracts and supernatants were also immunoprecipitated with anti-G14 polyclonal antibodies, raised against the secreted recombinant protein, indicating that the antibodies can recognize the unglycosylated polypeptide (Figure 5d).

Thioesterase activity of the G14 protein

The acyl-thioesterase enzymic properties of the G14 protein were assayed *in vitro* in spectrophotometric experiments with the use of the protein expressed in the baculovirus system. The assay used measures spectrophotometrically, in the presence of Nbs_2, the appearance of the CoA thiol group when the fatty acid is removed from the acyl-CoA. To define the acyl acceptor specificity of the G14 protein, in relation to the length and degree of saturation of the fatty acid, different acyl-CoAs were used. The cell extracts were resuspended in a volume equivalent to that of the corresponding concentrated supernatants, so that the activity present in supernatants (secreted) or cell extracts (cells) would correspond to the amount of protein produced from the same number of cells. In this case, 9 μg of cell extract was equivalent to 50 μg of supernatant, which corresponded in both cases to 1.66 μl of sample. These were chosen as the best conditions when different concentrations of cell extracts (0.2–900 μg) and supernatants (1–500 μg) were used in the assay (results not shown). The optimal concentration of acyl-CoA was also determined, by using 4–100 μM myristoyl-CoA (results not shown); 40 μM was selected for further experiments.

Under the selected conditions, using 1.66 μl of sample and 40 μM of acyl-CoA, the secreted form of the G14 protein showed acyl chain specificity depending on the length and degree of saturation of the fatty acid (Figure 6). The G14 protein showed very high S-thioesterase activity, after taking into account the background activity in supernatants from wild-type baculovirus-infected cells, towards the acyl chains $C_{14:0}$ (7.02 nmol/μg of protein) and $C_{16:0}$ (5.32 nmol/μg of protein) (Figure 6a). The G14 protein showed intermediate activity towards the acyl-CoAs $C_{14:1}$ (4.21 nmol/μg of protein), $C_{20:4}$ (3.86), $C_{16:1}$ (3.66), $C_{18:0}$ (3.63), $C_{12:0}$ (3.53), $C_{18:2}$ (2.91), $C_{18:3}$ (2.84), $C_{22:1}$ (2.44), $C_{18:1}$ (2.16) and $C_{20:0}$ (2.16) (Figure 6a). The G14 protein showed low activity towards the acyl-CoAs $C_{10:0}$ (1.47 nmol/μg of protein) and $C_{22:0}$ (1.15), and no activity towards the acyl-CoAs $C_{24:0}$, $C_{8:0}$, $C_{6:0}$, $C_{4:0}$ and $C_{2:0}$ (Figure 6a).

Interestingly, with the use of the same conditions and equivalent amount of cell extract as supernatant, the intracellular form of the G14 protein showed as little as 1/10 of the activity for all the substrates (Figure 6b); the best activities were found for the acyl-CoAs $C_{12:0}$ (0.82 nmol/μg of protein), $C_{14:0}$ (0.73) and $C_{14:1}$ (0.51) (Figure 6b). The next best activity was for the acyl-CoA $C_{16:0}$ (0.3 nmol/μg of protein) and the rest showed lower values, indicating no activity for those substrates (Figure 6b).

Figures 7 and 8 show the kinetic assays (used to produce the results shown in Figure 6) performed to detect S-thioesterase activity with different acyl-CoAs as substrates, with supernatants (Figure 7) or cell extracts (Figure 8) obtained from insect cells infected with wild-type baculovirus or vG14Bac. In all assays in which activity was detected, it was linear for at least the 6 min duration of the assay. The background activity of the cell extracts or supernatants from wild-type baculovirus infections was very low and constant for the duration of the assay (Figures 7 and 8). The acyl-CoAs that were found not to be substrates for cell extracts or supernatants containing the G14 protein showed similar values, independently of duration, to those obtained with cell extracts or supernatants from wild-type virus (Figures 7 and 8). In addition, to analyse the ability of the G14 protein to hydrolyse the thioester linkage of acylated protein substrates, different mammalian cell lines were labelled with [^{3}H]palmitic, [^{3}H]myristic and [^{3}H]mevalonic acids. These labelled extracts were treated with cell extracts or supernatants of insect cells containing the recombinant G14 protein. However, no difference in the labelled protein pattern was observed in comparison with control cell extracts or supernatants (results not shown).

DISCUSSION

The G14 polypeptide showed motifs common to triglyceride lipases and thiol proteases in its N-terminal region. In addition, the G14 polypeptide showed a motif common to the cytokine receptor superfamily at the C-terminus, which could be formed structurally by anti-parallel β-sheets. The crystal structure of pancreatic lipase has been resolved [29–31]: there are two structural domains, a catalytic N-terminus and a non-catalytic C-terminus. The interaction of the lipase with the lipid interface requires a small protein cofactor named colipase. Colipase binds to the C-terminal domain of the pancreatic lipase molecule. The N-terminal catalytic domain is composed of α-helices and β-sheets and contains the active site. However, the C-terminal non-catalytic domain is composed of anti-parallel β-sheets. The amino acid sequence of the G14 polypeptide showed that the lipase motif is located in the N-terminal region of the protein with predicted α-helices and β-sheets and that the cytokine receptor motif is in the C-terminal region with predicted anti-parallel β-sheets. This could be a unique feature of the G14 protein, not present in PPT1, that might determine the substrate specificity and/or the protein target of the G14 protein as an S-thioesterase. Resolution of the structure of the G14 protein would help in understanding these structural features.

The protein product of the gene for G14 has been expressed in the baculovirus system, showing a 30 kDa form in cell extracts, which could represent the unglycosylated processed protein, and 34–40 kDa forms, which might correspond to different glycosylated variants. The protein has also been found to be secreted to the extracellular medium as 34–36 kDa forms and a heterogeneous mixture up to 43 kDa, which could correspond to the mature glycosylated G14 protein. In the presence of tunicamycin, the protein was not secreted to the extracellular medium and the molecular mass of the intracellular protein was 30–31 kDa, indicating that the higher-molecular-mass polypeptides represent different glycosylation stages. Immunoprecipitations of radiolabelled cell extracts or supernatants from different mammalian cell lines with the anti-G14 antiserum showed a specific band of 42 kDa, which could correspond to the G14 protein. This 42 kDa band was detected at similar levels in all the cell lines analysed, independently of the level of expression of the G14 mRNA by Northern blot analysis, which suggested differences in the stabilities or the translation efficiencies of the mRNA species.

Northern blot analysis established that the *G14* gene expresses at least two different mRNA species of approx. 2.0 and approx.

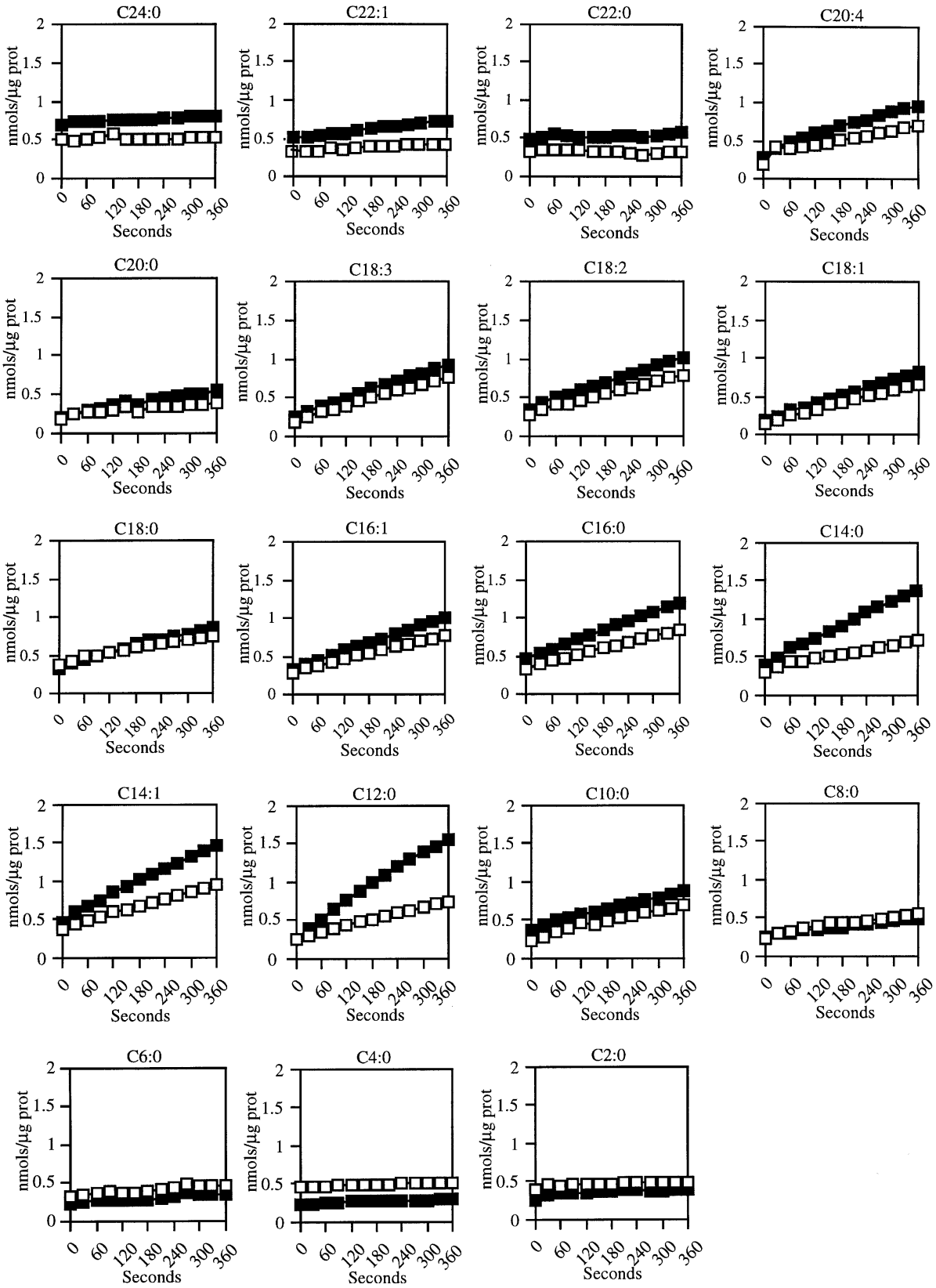

Figure 8 Spectrophotometric S-thioesterase assay with 1.66 μl (9 μg) of cell extracts from vG14Bac infections (■) or wild-type virus infections (□)

The acyl-CoAs used were: lignoceroyl ($C_{24:0}$), eurocoyl ($C_{22:1}$), behenoyl ($C_{22:0}$), arachidonoyl ($C_{20:4}$), arachidoyl ($C_{20:0}$), linolenoyl ($C_{18:3}$), linoleoyl ($C_{18:2}$), oleoyl ($C_{18:1}$), stearoyl ($C_{18:0}$), palmitoleoyl ($C_{16:1}$), palmitoyl ($C_{16:0}$), myristoyl ($C_{14:0}$), myristoleoyl ($C_{14:1}$), lauroyl ($C_{12:0}$), decanoyl ($C_{10:0}$), octanoyl ($C_{8:0}$), hexanoyl ($C_{6:0}$), butyryl ($C_{4:0}$) and acetyl ($C_{2:0}$).

1.8 kb. In comparison, Soyombo and Hofmann [10] showed by Northern blot analysis that PPT2 (gene for G14) in different human tissues is expressed as a major transcript of approx. 2.0 kb and as less abundant transcripts of 2.8 kb and 7.0 kb. However, they used as a probe an alternatively spliced form of PPT2 that contains six exons from the adjacent gene for LTBP. Using other probes they concluded that the 2.0 kb mRNA encodes the active form of the enzyme and the 2.8 kb transcript encodes the inactive, alternatively spliced PPT2-LTBP. The origin of the 7.0 kb transcript was not determined. We have seen in RT–PCR analysis, by amplification of the G14 coding sequence, the presence of at least five different bands. This indicated that the different transcripts are due to the alternative splicing of one or more coding exons rather than to differences in the length of the 3′ untranslated region. Nucleotide sequence analysis of the different RT–PCR products will be needed to confirm whether this is so. The finding of spliced forms that can lead to various molecular forms of the *G14* gene product might be relevant for the control of the enzymic activity and/or cellular localization of the G14 protein. Immunoprecipitation of HL60 cells with G14-specific antibodies revealed an approx. 28 kDa protein that was mainly secreted and poorly detected in cell extracts, which could represent a product of an alternative spliced transcript.

We have characterized the G14 protein activity in a quantitative assay and showed that supernatants obtained from vG14Bac infections of insect cells specifically contain S-thioesterase activity. The secreted G14 protein showed very high S-thioesterase activity towards the acyl chains $C_{14:0} > C_{16:0}$, moderate activity towards the acyl chains $C_{14:1} > C_{20:4} \approx C_{16:1} \approx C_{18:0} \approx C_{12:0} > C_{18:2} \approx C_{18:3} > C_{22:1} \approx C_{18:1} \approx C_{20:0}$, low activity towards the acyl chains $C_{10:0}$ and $C_{22:0}$, and no activity towards the acyl chains $C_{24:0}$, $C_{8:0}$, $C_{6:0}$, $C_{4:0}$ and $C_{2:0}$. However, purified recombinant bovine PPT1 showed, with the same enzymic assay, high activity towards the acyl chain $C_{14:0}$, moderate activity towards the acyl chains $C_{16:0}$, $C_{18:0}$ and $C_{18:1}$ [5], and only low activity towards the acyl chains $C_{12:0}$, $C_{20:4}$ and $C_{22:1}$ [5]. A comparison of the S-thioesterase activities of both enzymes indicates that the secreted G14 protein has a broader range of action than bovine PPT1, although both have a preference for long acyl chains (more than 12 or 14 carbons) over short acyl chains (less than 12 carbons). The secreted G14 protein showed a preference for saturated forms on the acyl chains of 14, 16 and 18 carbons; however, it showed a preference for the unsaturated forms when the acyl chains were 20 or 22 carbons long. This assay indicated S-thioesterase activity for the *G14* gene product (PPT2); we can conclude that the G14 polypeptide is an acyl-thioesterase. However, the assay cannot distinguish between esterase (thioesterase) and lipase activities; therefore the G14 protein could also have lipase activity. We prefer to name the *G14* gene product S-thioesterase (MHC class III S-thioesterase) instead of PPT2 because it has not yet been possible to demonstrate that it has PPT activity. Expression of the G14 polypeptide in the baculovirus system has been very successful because the wild-type baculovirus-infected cell extracts or supernatants have shown very little endogenous activity. This is in contrast with the results of Soyombo and Hofmann [10], who assayed only palmitoyl-CoA hydrolase activity of PPT2 by measuring the release of [^{3}H]palmitic acid from [^{3}H]palmitoyl-CoA by using cell homogenates of COS cells overexpressing PPT2. The assay that they used was performed after a 30 min incubation of the cell lysates in the presence of 5 mM *N*-ethylmaleimide, to suppress an endogenous *N*-ethylmaleimide-sensitive palmitoyl-CoA hydrolase activity that interfered with the activity above pH 6.0.

We have observed that, by using the same volume of cell extracts or supernatants, the activity of the enzyme detected was higher in supernatants (up to 10-fold) than in cell extracts. This indicated that the more active form of the protein is the secreted form or that very large amounts of the G14 protein are secreted. However, this is not in disagreement with an intracellular activity, because the secreted G14 protein could be internalized into the cell through a receptor and act on a target located in an intracellular organelle. This mechanism has been described for the secreted PPT1, which can be internalized into the cell by the mannose-6-phosphate receptor to act in the lysosome [7,8]. Soyombo and Hofmann [10] reported that PPT2 (G14 protein) binds to the mannose-6-phosphate receptor.

Mutations in hPPT1, which prevent the protein from being secreted, cause INCL, a severe brain disease [6]. Diseases associated with acyltransferases or esterases have also been described. Farnesyltransferase inhibitors have been shown to reverse the transformed phenotype of Ras-transformed cells and are currently being exploited for anti-tumorigenic potential. A mutation in Rab geranylgeranyltransferase produces the defect in choroideraemia, an X-linked retinal degenerative disease. Of clinical interest is the incorporation of synthetic myristate analogues into HIV-1 Gag, resulting in inhibition of the viral replication (see [3]).

The expansion of trinucleotide microsatellite repeats is an important mechanism of mutagenesis associated with various human genetic diseases, some of which are known to be neurogenic disorders. Interestingly, we have observed that the *G14* gene (MHC class III S-thioesterase) contains an incomplete CTG repeat in the 5′-coding region, which encodes a stretch of Leu residues in the signal peptide that are not present in PPT1. An alteration in the number of repeats on the *G14* gene could affect the processing of the signal peptide, with consequent abnormal protein localization or secretion and a lack of, or a defect in, its function. Here we have shown that the *G14* gene, which is located in the class III region of the MHC, contains S-thioesterase activity. A large number of diseases are associated with the products of genes located in the MHC, many of which are not fully explained by the classical class I and class II molecules. Further characterization of the gene for G14 (MHC class III S-thioesterase) at the genetic and molecular levels could elucidate its potential role in immunological or inflammatory diseases as well as in neurological diseases.

We thank Antonio Alcamí for helpful discussions and critical reading of the manuscript and for help with the baculovirus expression system, John Broxholme for help with computing, and Geoffrey Smith for access to specialized equipment. B. A. held a European Community Training and Mobility Bursary and a Florey-EPA Studentship from Queen's College, Oxford.

REFERENCES

1 Bhatnagar, R. S. and Gordon, J. I. (1997) Trends Cell Biol. **7**, 14–20
2 Casey, P. J. (1995) Science **268**, 221–225
3 Resh, M. D. (1996) Cell Signal. **8**, 403–412
4 Camp, L. A. and Hofmann, S. L. (1993) J. Biol. Chem. **268**, 22566–22574
5 Camp, L. A., Verkruyse, L. A., Afendis, S. J., Slaughter, C. A. and Hofmann, S. L. (1994) J. Biol. Chem. **269**, 23212–23219
6 Vesa, J., Hellsten, E., Verkruyse, L. A., Camp, L. A., Rapola, J., Santavuori, P., Hofmann, S. L. and Peltonen, L. (1995) Nature (London) **376**, 584–587
7 Verkruyse, L. A. and Hofmann, S. L. (1996) J. Biol. Chem. **271**, 15831–15836
8 Hellsten, E., Vesa, J., Olkkonen, V. M., Jalanko, A. and Peltonen, L. (1996) EMBO J. **15**, 5240–5245
9 Lu, J. Y., Verkruyse, L. A. and Hofmann, S. L. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 10046–10050
10 Soyombo, A. A. and Hofmann, S. L. (1997) J. Biol. Chem. **272**, 27456–27463

11 Aguado, B. and Campbell, R. D. (1998) J. Biol. Chem. **273**, 4096–4105
12 Stamps, A. C., Elmore, M. A., Hill, M. E., Kelly, K., Makda, A. A. and Finnen, M. J (1997) Biochem. J. **326**, 455–461
13 West, J., Tompkins, C. K., Balantac, N., Nudelman, E., Meengs, B., White, T., Bursten, S., Coleman, J., Kumar, A., Singer, J. W. and Leung, D. W. (1997) DNA Cell Biol. **16**, 691–701
14 Min, J., Shukla, H., Kozono, H., Bronson, S. K., Weissman, S. M. and Chaplin, D. D. (1995) Genomics **30**, 149–156
15 Khanna, A. and Campbell, R. D. (1996) Biochem. J. **319**, 81–89
16 Aguado, B., Milner, C. M. and Campbell, R. D. (1996) in HLA/MHC: Genes, Molecules and Function (Browning, M. and McMichael, A., eds.), pp. 39–75, Bios Scientific Publishers, Oxford
17 Campbell, R. D. and Trowsdale, J. (1997) Immunol. Today **18**, centrefold poster
18 Kendall, E., Sargent, C. A. and Campbell, R. D. (1990) Nucleic Acids Res. **18**, 7251–7257
19 Staden, R. (1987) in Nucleic Acid and Protein Sequence Analysis, a Practical Approach (Bishop, M. J. and Rawlings, C. J., eds.), pp. 173–217, IRL Press, Oxford
20 Pearson, W. R. and Lipman, D. J. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 2444–2448
21 Fourney, R. M., Miyakoshi, J., Day, R. S. I. and Paterson, M. C. (1988) Focus (BRL) **10**, 5–7
22 Aguado, B. and Campbell, R. D. (1995) Genomics **25**, 650–659
23 Livingstone, C. and Jones, I. (1989) Nucleic Acids Res. **17**, 2366
24 Alcamí, A. and Smith, G. L. (1992) Cell **71**, 153–167
25 Harlow, E. and Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
26 Blow, D. (1990) Nature (London) **343**, 694–695
27 Bazan, J. F. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 6934–6938
28 Wells, J. A. and de Vos, A. M. (1996) Annu. Rev. Biochem. **65**, 609–634
29 Winkler, F. K., D'Arcy, A. and Hunziker, W. (1990) Nature (London) **343**, 771–774
30 van Tilbeurgh, H., Egloff, M. P., Martinez, C., Rugani, N., Verger, R. and Cambillau, C. (1993) Nature (London) **362**, 814–820
31 van Tilbeurgh, H., Sarda, L., Verger, R. and Cambillau, C. (1992) Nature (London) **359**, 159–162

Received 20 November 1998; 20 April 1999; accepted 19 May 1999