

Molecular cloning and biochemical characterization of a *Drosophila* phosphatidylinositol-specific phosphoinositide 3-kinase

Claude LINASSIER*‡, Lindsay K. MACDOUGALL*, Jan DOMIN* and Michael D. WATERFIELD*†§

*Ludwig Institute for Cancer Research, University College Branch, 91 Riding House Street, London W1P 8BT, and †Department of Biochemistry and Molecular Biology, University College London, Gower Street, London WC1E 6BT, U.K.

Molecular, biochemical and genetic characterization of phosphoinositide 3-kinases (PI3Ks) have identified distinct classes of enzymes involved in processes mediated by activation of cell-surface receptors and in constitutive intracellular protein trafficking events. The latter process appears to involve a PtdIns-specific PI3K first described in yeast as a mutant, *vps34*, defective in the sorting of newly synthesized proteins from the Golgi to the vacuole. We have identified a representative member of each class of PI3Ks in *Drosophila* using a PCR-based approach. In the present paper we describe the molecular cloning of a PI3K from *Drosophila*, PI3K_59F, that shows sequence similarity to Vps34.

PI3K_59F encodes a protein of 108 kDa co-linear with Vps34 homologues, and with three regions of sequence similarity to other PI3Ks. Biochemical characterization of the enzyme, by expression of the complete coding sequence as a glutathione S-transferase fusion protein in Sf9 cells, demonstrates that PI3K_59F is a PtdIns-specific PI3K that can utilize either Mg²⁺ or Mn²⁺. This activity is sensitive to inhibition both by non-ionic detergent (Nonidet P40) and by wortmannin (IC₅₀ 10 nM). PI3K_59F, therefore, conserves both the structural and biochemical properties of the Vps34 class of enzymes.

INTRODUCTION

PI3K activity was originally identified associated with activated tyrosine kinases in mammalian cells and implicated in processes such as mitogenesis and transformation. These processes were shown to be dependent on the 3-phosphorylated lipids produced. The p110 catalytic subunit of this PI3K can phosphorylate the D-3 position of the inositol ring of the phosphoinositides PtdIns, PtdIns4P and PtdIns(4,5)P₂ *in vitro*, to generate the phospholipids PtdIns3P, PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃. In intact cells, the rapid and transient elevation in the cellular levels of the membrane phospholipids PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃ in response to ligand stimulation have identified these products as potential signalling molecules. These findings suggest that, within the cell, PtdIns may not be accessible to this form of PI3K (reviewed in [1,2]).

Molecular cloning and PCR-based approaches have identified a large family of PI3K catalytic subunits among evolutionarily diverse species. These can be assigned, by sequence similarity within the kinase domain, to three separate classes. This classification appears to reflect both the catalytic activity, in particular the substrate specificity, and differences in domain structure throughout the entire molecule (reviewed in [3,4]).

The first class of PI3Ks are activated by cell-surface receptors and have a broad *in vitro* substrate specificity phosphorylating PtdIns, PtdIns4P and PtdIns(4,5)P₂. This includes the prototypic receptor tyrosine kinase-linked PI3K, which is composed of a regulatory p85 subunit and a catalytic p110 subunit [5–8]. Ligand

stimulation of receptor tyrosine kinases results in auto-phosphorylation of specific tyrosine residues; these phosphotyrosine residues then act as docking sites for the SH2 domains of p85, allowing translocation of the associated catalytic activity to lipid substrates at the membrane. A number of distinct p85 [5,9] and p110 subunits, p110 α [8] and p110 β [10], have been identified, but no differences in function have been ascribed to them (reviewed in [1]). This class also includes PI3K γ , which is activated in response to stimulation of G-protein-linked receptors. PI3K γ does not associate with the p85 adaptor, but the presence of a putative PH domain at the N-terminus could provide a regulatory role [11].

The prototypical member of the second class of PI3Ks, PI3K_68D, has a distinct substrate specificity restricted *in vitro* to PtdIns/PtdIns4P. So far, upstream and downstream interactions have not been identified, but these could be mediated by its C-terminal C2 domain or novel motifs within its N-terminal extension [3].

The third class of PI3K, homologues of *S. cerevisiae* Vps34, appear to have a substrate specificity restricted to PtdIns and form part of an evolutionarily conserved complex involved in protein-trafficking events. The VPS34 gene was initially identified through alleles defective in the sorting and delivery of soluble vacuolar hydrolases from the Golgi to the vacuole (the yeast equivalent of the mammalian lysosome), indicating an involvement of this form of PI3K in protein trafficking [12]. Wild-type VPS yeast genes regulate the active diversion of proteins from the secretory pathway to the vacuole. In mammalian cells a similar

Abbreviations used: PI3K, phosphoinositide 3-kinase [this term applies to all enzyme activities which phosphorylate the D-3 position of the inositol ring, thus including phosphorylation of PtdIns and PtdIns phosphate species, specifically PtdIns4P and PtdIns(4,5)P₂]; PI4K, phosphoinositide 4-kinase; NP40, Nonidet P40; GST, glutathione S-transferase.

‡ Present address: EA 1369, Faculté de Médecin, Université de Tours, 2 Boulevard Tonnelle, 37044 Tours, France.

§ To whom correspondence should be addressed, at the Ludwig Institute.

The nucleotide sequence of *Drosophila* PI3K_59F (Figure 1) has been deposited with the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X99912.

sorting mechanism exists. Lysosomal enzymes (analogous to the yeast hydrolases) are transported with secretory proteins through the early stages of the secretory pathway, then, in the late Golgi compartment, are actively segregated away from proteins destined for secretion (reviewed in [13]). Point mutations which alter residues in Vps34 that are highly conserved between lipid kinases result in a loss of PI3K activity and missorting of vacuolar proteins, suggesting that the lipid kinase activity is required for the sorting process [14]. Vps34 has been shown to be a PtdIns-specific PI3K [15] suggesting a diversity of function for distinct 3-phosphoinositides. Although yeast Vps34 shows considerable sequence similarity to p110 PI3Ks, particularly within the putative catalytic domain [8], the partial purification of a bovine PtdIns-specific PI3K [16] and the cloning of a human PtdIns-specific PI3K [17] indicates closer homologues exist. Furthermore, soybean (*Glycine max*) [18], *Arabidopsis thaliana* (thale cress) [19], the slime mould *Dictyostelium discoideum* [20] and the yeast *Schizosaccharomyces pombe* [21] Vps34 homologues have now been described. The human enzyme, termed 'PtdIns 3-kinase', interacts with a human homologue of the yeast protein Vps15 ([17]; C. Panaretou and M. D. Waterfield, unpublished work), a protein serine/threonine kinase which, in yeast, appears to constitutively localize Vps34 to the Golgi membrane and may play a role in activation of the lipid kinase [22]. Thus, like the p110-p85 heterodimeric PI3K, the Vps34-Vps15 holoenzyme forms an evolutionarily conserved complex of regulatory and catalytic subunits responsible for its distinct subcellular localization and substrate specificities.

A number of other proteins have been identified which show sequence similarity to the kinase domain of PI3Ks, but have not been shown to phosphorylate the D-3 position of phosphoinositides. These can be divided into three families, some of which may be protein, rather than lipid, kinases: TOR-related proteins, homologues of the *ataxia telangiectasia mutated* (ATM) gene and the PI4Ks (reviewed in [4,23]).

The molecular characterization of a number of PI3Ks has explained the nature of their association with upstream components and provided clues to differences in substrate specificity. However, downstream components have not been clearly defined despite the availability of recombinant proteins. Therefore one of the approaches we have taken is to use *Drosophila* as a genetically manipulatable, multicellular organism in which to delineate the signal-transduction pathways involving PI3Ks. As a first step towards this goal we have undertaken the molecular and biochemical characterization of PI3Ks in *Drosophila*.

We have previously reported the identification of a family of PI3Ks in *Drosophila* [3] using a PCR-based approach with degenerate primers to the core catalytic domain (Homology Region 1) conserved between mammalian p110 α and yeast Vps34, hereafter referred to as 'Sc_Vps34' to distinguish it from its homologues. Three sequences were identified, each of which corresponded to a distinct subgroup of PI3Ks and which were named after their unique and separate sites of hybridization to *Drosophila* polytene chromosomes. PI3K_92E/Dp110 [24] is related to the PI3Ks (p110 α and p110 β), which are linked to receptor tyrosine kinases. PI3K_68D defines a novel class of PI3Ks characterized structurally by a C-terminal C2 homology domain, and biochemically by an *in vitro* substrate specificity restricted to PtdIns and PtdIns4P [3]. The third sequence was most similar to Sc_Vps34 (56% identity over 128 amino acids). Additionally, the *Drosophila* and yeast proteins had a conserved 11-amino-acid deletion and a four-amino-acid insertion when aligned with the amino acid sequence of p110 α , further establishing their close similarity. In the present paper we report the complete amino acid sequence and the biochemical charac-

terization of the *Drosophila* Vps34 homologue, termed 'PI3K_59F'.

MATERIALS AND METHODS

General molecular-biology techniques were performed as described in [25]. cDNA probes were labelled by random priming using [α -³²P]dCTP. Oligonucleotides were labelled with [γ -³²P]ATP and polynucleotide kinase. DNA sequencing was performed using the *Taq* DyeDeoxyTerminator Cycle Sequencing Kit (Applied Biosystems) with T3-, T7- or PI3K_59F-specific primers and analysed on an Applied Biosystems 373A automated DNA sequencer. Data was analysed using the University of Wisconsin Genetics Computer Group software package [26].

Isolation of the PI3K_59F cDNA

The fragment corresponding to the core kinase domain of PI3K_59F [3] was excised from pGEM-T and used to screen 200000 plaque-forming units each of a total imaginal disc library and an embryonic library (provided by E. Hafen) in λ ZAP II (Stratagene). Filters were prehybridized in Church buffer (0.5 M sodium phosphate/7% SDS/1 mM EDTA, pH 7.2) at 65 °C and hybridized in the same solution containing 1 ng/ml radio-labelled probe. Filters were washed in 0.5 \times SSC/0.1% SDS at 65 °C prior to autoradiography (1 \times SSC is 0.15 M NaCl/0.015 M sodium citrate). Positive clones were purified by an additional two rounds of screening, and the products, excised using the helper phage, sequenced. Identity was confirmed with primers specific to the core kinase domain of PI3K_59F. The two longest clones, α 28 and α 29, obtained from the total imaginal disc library were selected and sequenced on both strands following subcloning of restriction fragments into pKS⁺ (Stratagene). Gaps in the sequence were filled using PI3K_59F-specific oligonucleotide primers.

Plasmid constructs

PI3K_59F₁₋₂₅₈

A PCR fragment generating novel *Eco*RI sites at positions -7 and 778 of the full-length cDNA was obtained using the α 29 cDNA template and the primers GCGCGAATTCAAATGGA-CCAGCCCACG (sense) and GGATTCGACCAAAGACAC-TTAAGGCGC (antisense) This was directly inserted into a T-vector (pT7 blue, AMS) to form pT7-PI3K_59F₁₋₂₅₈.

Glutathione S-transferase (GST)-PI3K_59F₁₋₉₄₉

To facilitate subcloning, a shuttle cassette was created encoding the complete PI3K_59F open reading frame by a double ligation of the 5' *Eco*RI-*Aat*II fragment from the pT7-PI3K_59F₁₋₂₅₈ construct and the 3' *Aat*II-*Eco*RI restriction fragment from the α 29 cDNA (see Figure 2B below) into *Eco*RI-digested pAcG2T vector [27]. All PCR-amplified sequences were confirmed by sequencing.

Expression of GST fusion proteins

Sf9 cells were cultured and infected using standard techniques [28].

GST-PI3K_59F₁₋₉₄₉ (full length) was co-transfected with BaculoGold DNA (Pharmlingen) into Sf9 cells using Lipofectin (Gibco). Typically Sf9 cells, at a density of 10⁶ cells/ml, were infected with baculoviruses for 48 h. Cells were then harvested by centrifugation, washed twice with PBS and lysed in TBS-T

[50 mM Tris (pH 7.5, 25 °C)/150 mM NaCl/1% (w/v) Triton X-100, containing 1 mM EDTA, 1 mM NaF, 0.1% (v/v) β -mercaptoethanol, and the protease inhibitors 1 mM benzamidine, 1 mM PMSF, 20 μ g/ml aprotinin and 2 μ g/ml leupeptin] for 20 min on ice. The lysates were cleared of particulate material by centrifugation at 15000 g for 20 min at 4 °C and then incubated with glutathione-Sepharose beads for 2 h at 4 °C with rotation. GST-tagged proteins bound to the beads were recovered by centrifugation and washed repeatedly with lysis buffer.

PI3K assay

GST-Sepharose-bound fusion protein and immunoprecipitates of *Drosophila* lysates on Protein A-Sepharose were assayed for PI3K activity, essentially as described previously [29]. Briefly, beads were washed twice in lysis buffer and twice in assay buffer [20 mM Tris/HCl (pH 7.5, 25 °C)/100 mM NaCl/0.1% β -mercaptoethanol]. Reactions were performed at 30 °C for 15 min, in a total volume of 50 μ l in the presence of either 3 mM Mg²⁺ or 3 mM Mn²⁺, 40 μ M ATP and 200 μ M lipid. In some experiments assays were performed in the presence of 0.5% Nonidet P40 (NP40; BDH) or various concentrations of wortmannin (Sigma). Reactions were terminated with 1 M HCl/chloroform/methanol (1:1:1, by vol.), mixed and centrifuged to separate organic and aqueous phases. The lipid products in the lower layer were extracted and resolved by TLC. For standard assays using PtdIns as substrate, the solvent used was chloroform/methanol/4 M NH₄OH (9:7:2, by vol.). For separation of PtdIns3P, PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃ the solvent was propan-1-ol/2 M acetic acid (13:7, by vol.). Radioactivity was analysed by autoradiography or quantified using a Phosphorimager (Molecular Dynamics). Alternatively, lipids in the lower phase were deacylated and analysed by HPLC [30].

HPLC characterization of lipid products

Assays of PI3K_{59F} were performed and products were extracted as described above, then dried and deacylated. HPLC analysis of glycerophosphoinositols was performed on a Beckman System Gold apparatus using two Partisphere SAX columns (Whatman) connected in series and eluted using a linear gradient of (NH₄)₂HPO₄, pH 3.8, in water at 0.5 ml/min. ³H-labelled, deacylated PtdIns4P (Amersham) was used as an internal standard. A PtdIns3P standard was prepared by phosphorylating PtdIns *in vitro*, with baculovirus-expressed p110 α , followed by deacylation as described above.

RESULTS

Sequence analysis of PI3K_{59F}

The PCR fragment encoding the core catalytic domain of PI3K_{59F} was used to screen several cDNA libraries (200000 plaque-forming units each) at high stringency. Three positive clones were obtained from an embryonic library and three from a total disc library. The two longest inserts, of 3.0 and 3.2 kb, which were obtained from the total disc library, were sequenced completely. Both contained an in-frame stop codon in the short untranslated sequence 5' to the initiating methionine codon and multiple in-frame stop codons within the 3' untranslated sequence. The sequences of the two clones were identical from nucleotide -3 (relative to the translation initiation start site) to nucleotide 3005, a region encompassing the complete coding sequence. The sequence in the region -3 to -9 (which was different between the two clones) was confirmed as identical with that of the longest clone by direct PCR sequencing of genomic

DNA. The sequence of the longer clone (shown in Figure 1A) additionally contained extensions at the 5' and 3' ends that included a short polypyrimidine tract at the 5' end and a poly(A) tract at the 3' end. The complete open reading frame of PI3K_{59F} (2847 nucleotides) encodes a protein of 949 amino acids with a predicted molecular mass of 108 kDa and a pI of 7.6 (Figure 1A). The sequence adjacent to the initiating methionine (G AAA ATG G) is in good agreement with the consensus sequence for translation initiation in *Drosophila* (C/A Aaa/c ATG g) as defined by Cavener and Ray [31].

A comparison of the amino acid sequence of PI3K_{59F} with other PI3Ks identified three regions of strong sequence similarity (Figure 1B and 1C). Within each of these regions of similarity, the sequence of PI3K_{59F} is most similar to other members of the Vps34 family. Homology region 1 encodes the putative core catalytic domain (amino acids 694–891 in PI3K_{59F}) and is located at the C-terminus. This region represents the region of greatest sequence identity with other PI3Ks (54% identity with Sc_Vps34; 38% identity with p110 α). Homology region 2 (amino acids 294–408 in PI3K_{59F}), previously referred to as the 'PIK domain', can be found in all PI3Ks and PI4Ks. Homology region 2 has also been identified in proteins, such as TOR2, which show sequence similarity to PI3Ks but for which lipid or protein kinase activity has not been demonstrated. Homology region 3 (amino acids 88–115 in PI3K_{59F}) appears to be specific to PI3Ks. The function of homology regions 2 and 3 is currently unknown. Class I (p110) PI3Ks additionally contain a ras-binding domain [32]. This region is not found within the Vps34 class of PI3Ks (Figure 1C).

A comparison with other members of the Vps34 subclass of PI3Ks demonstrates co-linearity, with sequence identity between these molecules throughout the sequence. PI3K_{59F} shows 56.5% identity (72% similarity) to human PI3K and 37.2% identity (59.5% similarity) to Sc_Vps34p. The alignment appears to reflect phylogenetic divergence; hence PI3K_{59F} most closely aligns with the human Vps34 homologue, PtdIns 3-kinase, whilst the two plant sequences At_Vps34 and Gm_Vps34 cluster (Figure 1D). Sequence similarity between PI3K_{59F} and other Vps34 homologues includes regions towards the N-terminus, the presumed site of interaction of Sc_Vps34 with Vps15 (Figures 1B and 1C). In contrast, PI3K_{59F} lacks identity with both the N-terminal 120 amino acids of p110 α that have been identified by deletion analysis [33] as the binding site for its adaptor, p85, which mediates association with activated tyrosine kinases and with the PH domain of PI3K γ [11], which may mediate membrane association and/or association with heterotrimeric G-protein receptors (Figure 1C).

A search of the database with the FASTA algorithm failed to reveal similarity with proteins other than PI kinases, even when the search involved the regions of lowest sequence similarity to these molecules.

Expression and biochemical characterization of PI3K_{59F}

In order to characterize the biochemical and pharmacological properties of PI3K_{59F}, we constructed a fusion protein that comprised the amino acid sequence of GST linked to that of the PI3K_{59F} protein (Figure 2A). After infection of Sf9 cells, the fusion protein was purified by glutathione-Sepharose affinity chromatography. Analysis by SDS/PAGE revealed a single protein of ~134 kDa consistent with the calculated size of the fusion protein (Figure 2B). The enzymic activity and substrate specificity *in vitro* were investigated and compared with those of mammalian p110 α , also expressed as a GST fusion protein using the baculovirus system. GST-PI3K_{59F} displayed activity to

(A)

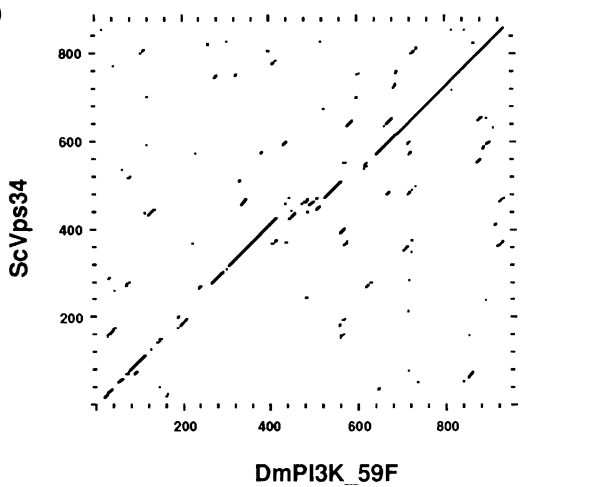
```

TTTTTTTTTTTTTTTTTTTTCTGTGTGGCAGCTGTGTGGTGTCCCGAAAATTCGCGAGAAAATAAAAGTGAATTGGCAGCAGGAAA      -1
ATGGACCAAGCCAGCACCATTTCCGCTACACAGCTCCCTCGCTCCAGAGCAGGTCAAATCAAGGTAGGAGAGCTCGAGGGGAAGAGCCGAGCTATGAGAACTGTA      120
M D Q P D D H F R Y I H S S S L H E R V Q I K V G T L E G K K R Q P D Y E K L L      40
GAGGATCCGATCTAGATTCCTGGGCTTACTCCAGGAGCACCCTCGTCCAGGTCGCTCGCAGGTTTCAACCAGGCGAGACCTACTGCTCCAGTACCAGTCTCTCAAG      240
E D P I L R S I S M F S E H P S F Q V R L Q V F N Q G R P Y C L P V Y T S S Y K      80
CGCTTCGCAAGCGATGGAGCTGAAACGAGTGGGTGACGCTGCCGCTGCAATCTCCGACCTGCCAGAGCGCCATGCTGGTGTGACCATCTGGACTGCTCGGCGCGGGCCAGACC      360
A F G K R W S W N E W V T L P L Q F S D L P R S A M L V L T I L D C S G A G Q T      120
ACGGTCATTGGAGGCACTCCATCTCAATGTCGCAAGGACGGCAGTTCGGGAGGGAATGTACGACTTGGAGTGTGGTGGGAGTCCGAGGCGATGGCAACTTCCAGTCGCAGC      480
T V I G G T S I S M F S E H P S F Q V R L Q V F N Q G R P Y C L P V Y T S S Y K      160
CCCGAAAGGGCAAGGAGTCGTCGAAGTCGAGATGCAGCCCTAGGAAAGTGGCCAAGAAACAGGAAACGCGCAGGTCGAAAAGGTGGACTGGCTGGACAGACTGACGTTTCCGGAA      600
P G K G K E S S K S Q M Q R L G K L A K K H R N G Q V Q K V D W L D R L T F R E      200
ATCGAGGTGATCAACGAGCGGCAAAAGCGTATGTCGCACTATATGTCCTTATGATTGAGTTCCGCAATCGTGGTGGATGACATGACAACCTACGCTGCTACTTCGAGCCGGAG      720
I E V I N E R E R K R M S D Y M F L M I E F P A I V V D D M Y N Y A V V H L Y F E P Q      240
GGTACGCTCAAATACAAGTCGCGCCAAACCTAAGTCGTTCTGTGCCGACTCCGAGATCCAAATGGAAAAATCTGGTCGAGCGCAAGCACCATCGATTGGCCCGTTCGAGCCGTTCA      840
G D V K Y K L P A K P K L V S V P D S E I Q M E N L V E R K H H R L A R S E R S      280
GGCATCTGGCAGAGATGCAACCAGCGACGACGATTCGCGATCAGCTCCACAGGATGTCTATAGTATCCACCACCCTATGTCTGAGCAGCAGGAGCAGGATCTGTTGTTGAAG      960
G I S D R D L S A T S S G L H G S V I P A N Q R A A S V L A A I K S D K S V S S G      320
TTCCGCTTTTACCTTTCCTCGCACAAGAAGCTCTGACCAAGTTCCTCAAGTCATCAATGGAAGTGGAGGACGAGGTTACGAGGCGCCTTTGGATGCTGGCCAACCTGGCGCCCATG      1080
F R F Y L S S H K K A L T K F L K C I N W K L E D E V T Q A L W M L A N W A P M      360
GACGTGGAGGATGCTTGAGCTGTAGTCAGCATTTTACGACCCGAGGTCGCAAAAGTACGCTGTAGTCCGATGGCCCAAGCGCCAGAGGATCTGCTGCTACTTGTCTTCCAG      1200
D V E I N E R E R K R M S D Y M F L M I E F P A I V V D D M Y N Y A V V H L Y F E P Q      400
CTGGTGAAGCGCTGAAGTACGAGGATCCGCGCCACATTGTTACCTGTCATGGTGCATCTTCCGAGCGGATGTTGGTGGCTCCATCCGAGCACAACCGCTCTCTGCTGGACCAG      1320
L V Q A L K Y E D P R H I V H L H G C I F P E R D V V R S I L D D N G S L L D Q      440
AGCAGCTATCGGATGAGCGCCAGCAGGACGAGTGGACTCCACGCTCGGTCATTCCGCGCAATCAGCGAGCGCGCAGTGTGTTGGCCGCGCATCAAGAGCGATAAGTCAGTGGATTCGGGT      1440
S S L S D L S A T S S G L H G S V I P A N Q R A A S V L A A I K S D K S V S S G      480
TCAGCGGTGGCAGTGGCAGTGGTGGTCCAGGATCGTTCGCTGCCAACCCCTCTGCTCCCGCCACTCCGGGAGCAGTTCGTTGCCCTGCGACTCAAACCTCAAATGCCCTCATGCTG      1560
S A G G S G S G G Q G S V A L P N P S A P A T P G S S S L P C D S N S N A L M L      520
GCGAGGGCATCAGCTTCGCAAGTTCGCAAGTTCGCAATCTGTCACATTCCTGATCCAGCGCTGTCACGAACCGCCAGCTGGCCAACCTACTTACTGTTGTTGTCATCGAGGTTGAG      1680
A E G I S F G S V P A N L C T F L I Q R A C T N A T L A N Y F Y W Y L S I E V E      560
GAAGTGGAGTCCGTGAGGAGCAGGACGAGAGGGCCACGACATGACGGTATGCTCAAGTGTGCTGAAGTCGTAGAGAACGGCACTTCAACCTGAGAGGCATCTTCTACAAC      1800
E V E S V R K Q D E R A H D M Y A M V L K M F L K V L E N G N F N L R G I F Y N      600
CTCCGGAAGCAGCTCGCTTTCATCGAGGAGCTGGTCAAGTGTGGCAACTGGCAGGAGCGGGAATCGCAATGAAGAGCGAGGAAAGTTCCAGAAGCTGTGGCCGAACAGGAC      1920
L R K Q R R F I D E L V K L V K L V A K E P G N R N K K T E K F Q K L L A E Q D      640
ATGTTCAAGGTGAACCTCAACCACTTCGAGCCCATTCGTTCCCTGGATCCGAGATCTACATCAACGATCGTGCCATGCGCATCGCTTTCAGAGCGCCCTGATCCGAGCC      2040
M F K V N F T T C F E P I P F P L D P E I Y I T K I V P M R T S L F K S A L M P A      680
AAGCTCACTTTTCACTGATTGCCATCAGGATACGCGCCATTTCGAAAGTCGTCAGGATGATCTGCGCAGGATCAGCTATCCTGAGATGATTTGATGGAGAAGTGGCTG      2160
K L T F V T S I A H H E Y A A I F K H G D D L R Q D Q L I L Q M I T L M D K L L      720
AGGCGGAGAACCTCGATCTGAAGTCGACTCCCTACAAGTGTGGCCACCAGCTCCAAGCAGGCTCTCTCCAGTACGTCGACTCTGACCGTGGCGGAGGTGCTGGCGGGGAGGCC      2280
R R E N L D L K L T P Y K V L A T S S K H G F L Q Y V D S C T V A E V L A R E T G      760
AACATCCACAATCTTCGCAAGCATTCCGTCGACCAATGAGCTTCCGCGGAGTTCGCGGAGTCTGACACCTATCAAGAGTCGCGGGCTACTGTGATACCTACTTTG      2400
N I H N F F R K H H P C D N G P Y G I S A E V M D T Y I K S C A G Y C V I T Y L      800
CTGGGTGGCGGACCGCACTGACAACTTGTGTCAGCACCAGCGGAAGCTTCCACATCGATTTCCGTACATCTGGACGCGATCCCAAGCCATGCCCGCCCAATGAAG      2520
L G V G D R H L D N L L L T T N G K L F H I D F G Y I L G R D P K P M F P P M K      840
CTGACCAAGAAATGGTGGAGCCATGGGGCGCATCGCTCCGAGCATTCCAGGATTCGCGAGCAGTGTACACCGCCATCTGACCTGCACTTGGCGGCGCATGCAAC      2640
L S K E M V E A M G G I S S E H H H E F R K Q C Y T A Y L H L R R H A N V H L N      880
CTGTTCCCTGATGGTGGAGCCACTGCGCGATATCGCTCAGAGCCGACAGCGGTCAAGAAGTGGAGGAGAACCTCAGCTGGCGCTACCGACGAGGAGCAGTGCAGCAC      2760
L F S L M V D A T V P D I A L E P D K A V K K V E E N L Q L G L T D E E A V Q H      920
TTGACAGCTCCCTGCGTGTCCATTCAGCGGTGATGCGGCTTTGGTGGAGCAGATTCATCGGTTTACCCEAATCTGGCGGAAGTGAAGGACATTGGCTGCAAGAATTTGTTGGA      2880
L Q S L L D V S I T A V M P A L V E Q I H R F T Q Y W R K *      949
GGAGCGAGTAACTCTAGTTGGTTCAGTGTGTTTAAATCAAAGTGTACATTAGATTAATACCTTATACTTATCAATATTATTAATATTATGATGAGTGTAGTGGCTGT      3000
AGCTAAAAGTCCGTTGGCCGAACCAATGTATCCACATATACATATATCTTGTCCACAAAAAAA          3071

```

(B)

(i)



(ii)

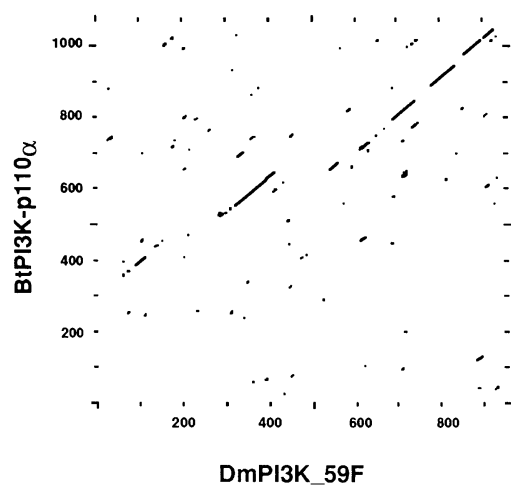


Figure 1 For legend see facing page.

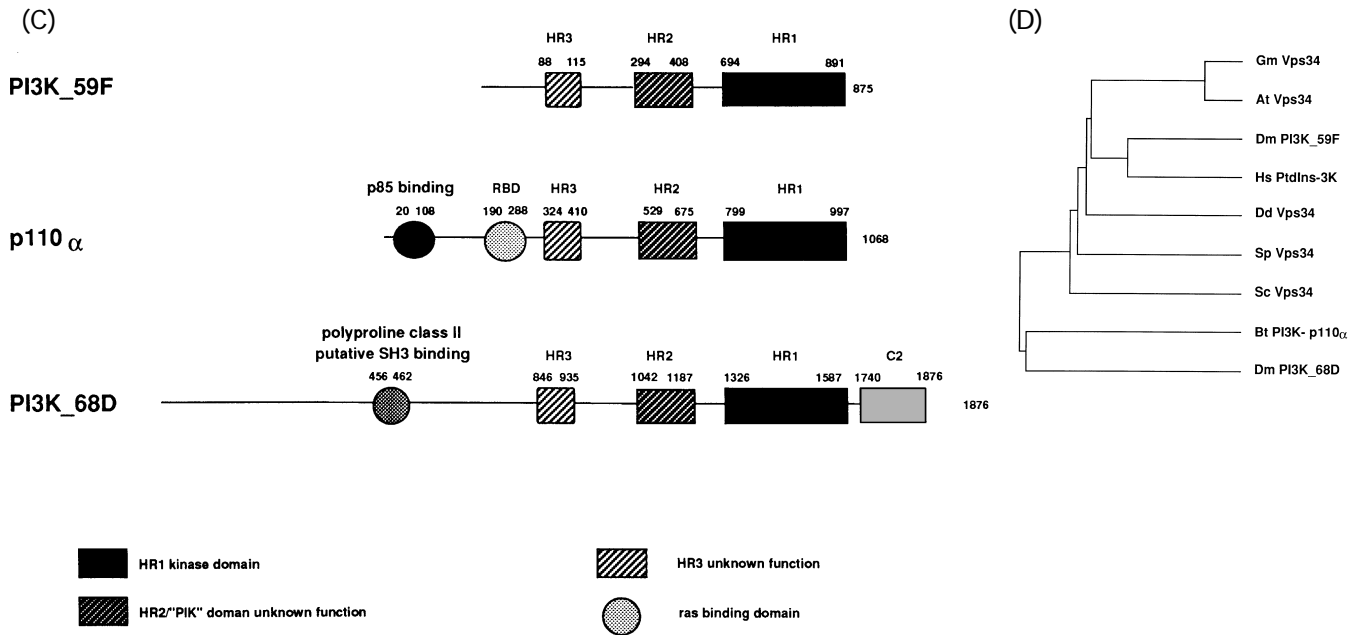


Figure 1 Sequence analysis and domain structure of PI3K_59F

(A) Nucleotide and derived amino acid sequence of *Drosophila* PI3K_59F. The putative wortmannin-binding site (Lys⁶⁹⁰) is underlined. (B) Dot-plot comparison of DmPI3K_59F (949 amino acids, horizontal axis) with (i) Sc_Vps34 (875 amino acids) and (ii) BtPI3K-p110 α (1068 amino acids), vertical axes, generated with the COMPARE programme (UWGCG package [26]) using a window of 30 and a stringency of 15. (C) Regions of sequence similarity between DmPI3K_59F (class III), BtPI3K-p110 α (class I) and DmPI3K_68D (class II). For details, see text and [4]. Numbers above modules refer to the amino acids predicted to form the boundaries of each domain. HR, PI3K homology domain; C2, C2 domain. (D) Dendrogram to show the evolutionary relationship between members of the Vps34 family (PI3K class III). Prototype members of class I (BtPI3K-p110 α) and class II (DmPI3K_68D) are also shown. The dendrogram was generated using the PILEUP programme (UWGCG package [26]). Abbreviations: Gm, *G. max*; At, *A. thaliana*; Dm, *D. melanogaster*; Hs, *H. sapiens* (man); Dd, *D. discoideum*; Sp, *Schiz. pombe*; Sc, *Saccharomyces cerevisiae* (baker's yeast); Bt, *Bos taurus* (ox).

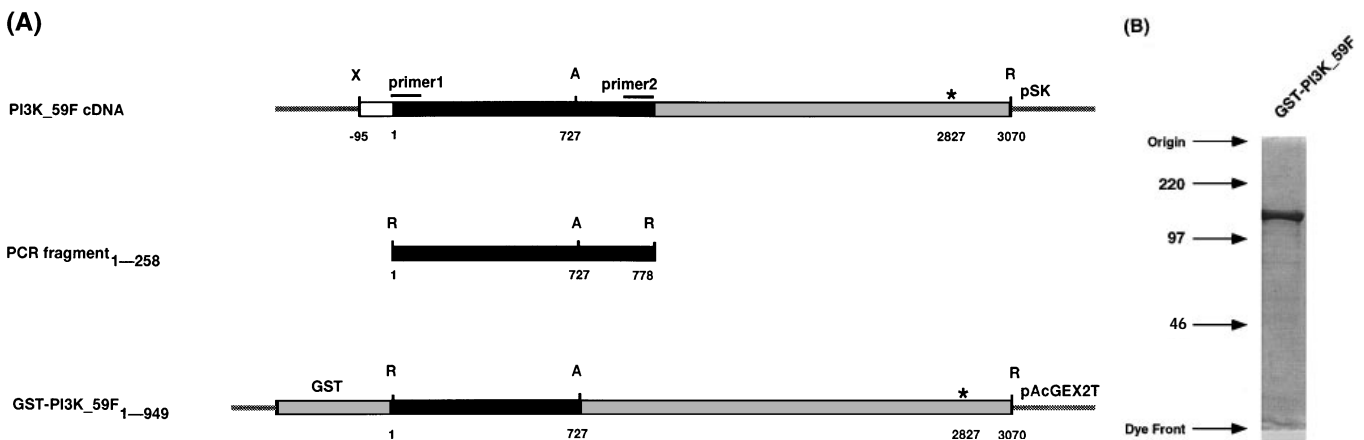


Figure 2 Expression of recombinant GST-PI3K_59F

(A) Plasmid constructs for expression of PI3K_59F as a GST fusion protein. For details, see Materials and methods section. Numbers refer to the nucleotide position in the open reading frame. Restriction enzyme sites: X, *Xho*I; A, *Aat*I; R, *Eco*RI. (B) Coomassie Blue staining of GST-PI3K_59F (~134 kDa). The molecular-mass markers are: myosin, 220 kDa; phosphorylase *b*, 97 kDa; and ovalbumin, 46 kDa.

wards PtdIns in the presence of either Mn²⁺ or Mg²⁺, but not towards PtdIns4P or PtdIns(4,5)P₂ in the presence of either cation (Figure 3A). In contrast, GST-p110 α displayed activity towards all three substrates, PtdIns, PtdIns4P and PtdIns(4,5)P₂, but only in the presence of Mg²⁺ (Figure 3A).

PI3K_59F is a member of the Vps34-related class of the PI3K family of enzymes, and the human PtdIns 3-kinase [17] and yeast Sc_Vps34 [15] enzymes have also been shown to have a substrate specificity restricted to PtdIns. The human homologue, however, has only a very weak activity in the presence of Mg²⁺ (Figure 3B,

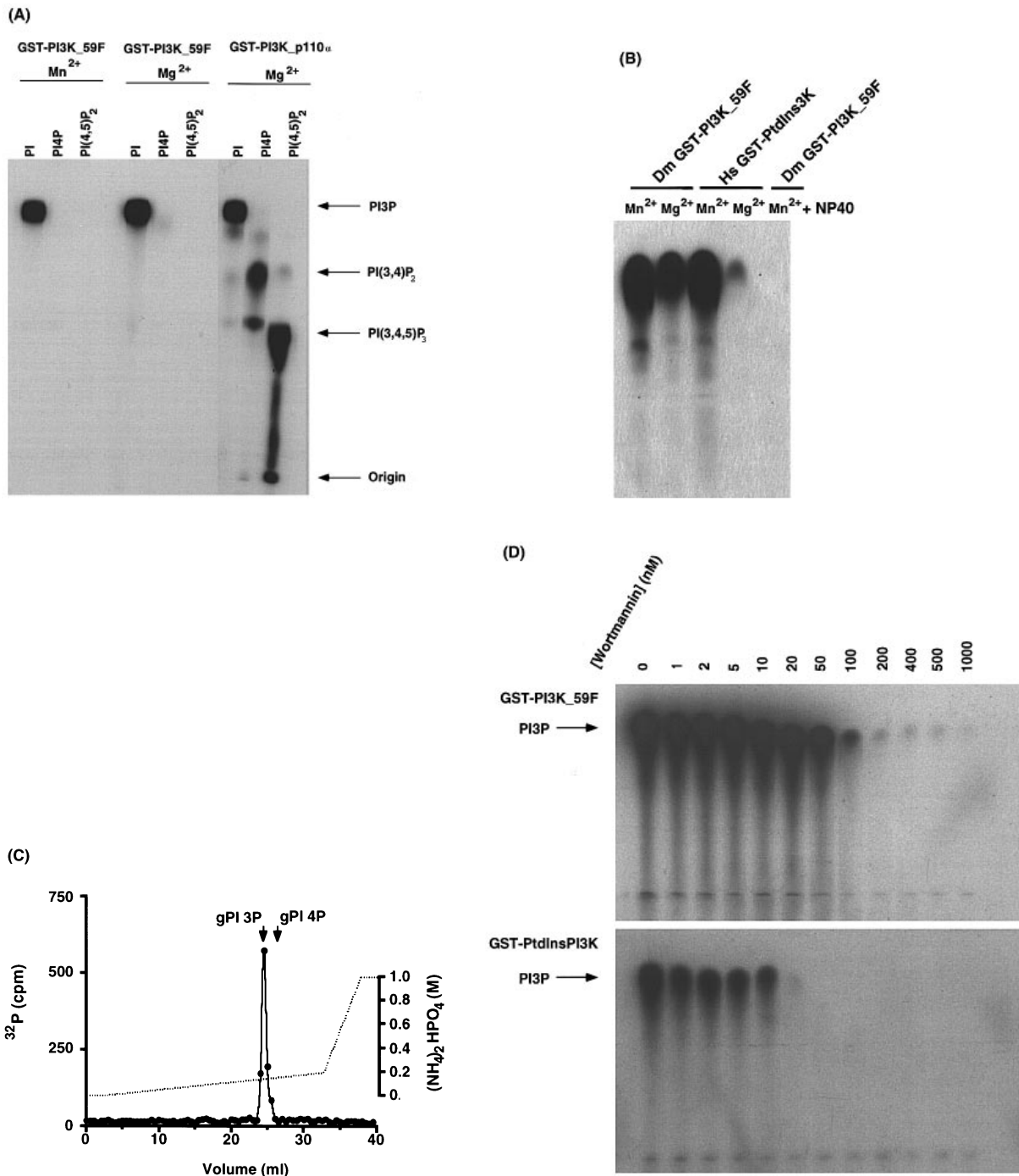


Figure 3 Biochemical characterization of recombinant PI3K_59F activity

The extracted lipids were resolved either by TLC and detected by autoradiography (**A**, **B** and **D**) or by HPLC (**C**). (**A**) *In vitro* substrate specificity of recombinant GST-PI3K_59F and GST-p110 α . PI3K assays were performed as described in the Materials and methods section using PtdIns, PtdIns4P and PtdIns(4,5)P₂ as substrates. (**B**) Effect of bivalent cations and NP40 on GST-PI3K_59F activity. Assays were performed using PtdIns as a substrate. The effect of bivalent cations on the human Vps34 homologue, PtdIns 3-kinase, expressed as a GST fusion protein is also shown. (**C**) HPLC analysis of the products of recombinant GST-PI3K_59F using PtdIns as a substrate. The broken line represents the (NH₄)₂HPO₄ gradient; gPI 3P and gPI 4P denote the respective positions of deacylated PtdIns3P and PtdIns4P standards. (**D**) Sensitivity of GST-PI3K_59F (top) and GST-PtdIns (bottom) activity to inhibition by wortmannin.

[17]). The more relaxed cation specificity of PI3K_59F indicates that a requirement for Mn²⁺ is not a general property of this class of PI3K and is consistent with the cation requirements of yeast PI3Ks. Both *Sc*_Vps34 [14,15] and *Sp*_Vps34 [21] immunoprecipitated from cells possess Mg²⁺-dependent PtdIns-specific

PI3K activity although a preference for Mn²⁺ has been reported for the *S.cerevisiae* enzyme [34].

A specificity for PtdIns is also a feature of PI4Ks but, unlike PI3Ks, these enzymes are characteristically insensitive to inhibition by non-ionic detergents (reviewed in [35]). PI3K_59F activity

was inhibited completely by 0.5% NP40 (Figure 3B), indicating that PI3K_{59F} is a PI3K. In order to formally confirm that PI3K_{59F} does encode a 3-kinase as implied from its close sequence similarity to other PI3Ks in the putative catalytic domain, anion-exchange HPLC was performed. The deacylated products of PtdIns phosphorylation yielded a single peak of radioactive glycerophosphate. This represented deacylated PtdIns3P as determined by its co-elution with the product of GST-p110 α . There was no detectable production of PtdIns4P (Figure 3C). These results confirm that PI3K_{59F} specifically phosphorylates the D-3 position of the inositol ring.

The antifungal compound wortmannin is a potent inhibitor of PI3Ks, particularly those of the p110 class, and has been used extensively in tissue-culture studies to investigate the role of the classical p110/p85 PI3K in intracellular signalling, although other classes of PI3Ks, such as PI3K γ and PI3K_{68D}, are also inhibited by wortmannin (reviewed in [36]). Indeed, some classes of PI4K are reported to be sensitive to wortmannin [37] as has the more distantly related protein kinase, DNA-dependent protein kinase, although with a considerably higher IC₅₀ value [38]. This inhibitor, therefore, cannot be used to implicate a specific PI3K in intracellular signalling, but has proved very useful in identifying processes controlled by this class of enzymes. Within the Vps34 family, considerable differences in sensitivity to wortmannin have been reported, although all the molecularly characterized forms conserve the lysine within the ATP-binding site that is targeted by this inhibitor [36]. Sc_Vps34, (IC₅₀ 3 μ M; [15]), Sp_Vps34 (IC₅₀ 600 nM; [21]) and a partially purified bovine PtdIns-specific PI3K (IC₅₀ 400 nM, [16]) have been reported to be relatively insensitive to wortmannin, whereas human PtdIns 3-kinase is inhibited by wortmannin with an IC₅₀ (2.5 nM) similar to that of GST-p110 α [17]. Recombinant PI3K_{59F} was inhibited by wortmannin with an IC₅₀ of 10 nM (Figure 3D).

DISCUSSION

We have described the molecular cloning and biochemical characterization of a PtdIns-specific PI3K from *Drosophila*, PI3K_{59F}, that shows similarity to Sc_Vps34, the yeast PI3K that mediates protein-sorting and -trafficking events. PI3K_{59F} encodes a protein of 108 kDa, co-linear with other Vps34 homologues, and with three regions of sequence similarity to other PI3Ks (Figure 1). Within each of these regions, sequence identity is highest with the Vps34 class of PI3Ks. Homology region I, the putative catalytic domain, is also distantly related to the catalytic domain of the protein kinase superfamily and contains motifs conserved in subdomains VIB (DXHXXN, Asp⁸⁰⁵-Asn⁸¹⁰ in PI3K_{59F}) and VII (DFG, Asp⁸²³, Phe⁸²⁴, Gly⁸²⁵ in PI3K_{59F}) of protein kinases [4,23]. In protein kinases these residues have been identified as being involved in substrate binding/product release, and point mutations within these motifs destroy the lipid kinase activities of p110 α [39] and Sc_Vps34 [14]. Additionally, lipid kinases conserve the lysine residue identified in protein kinases as essential for ATP binding. In p110 α , this residue, Lys⁸⁰², has been identified as the binding target for the inhibitor wortmannin and the non-hydrolysable ATP analogue 5'-[*p*-(fluorosulphonyl)benzoyl]adenosine [36]. The equivalent residue in PI3K_{59F}, Lys⁶⁹⁸, is presumably responsible for the observed inhibition of lipid kinase activity by wortmannin (Figure 3D). In protein kinases this invariant lysine residue is preceded by a GXGXXG motif within the P-loop, but this loop appears to be absent from the lipid kinases [4]. Considerable sequence variation between different classes of PI3Ks is apparent in the HR1 domain C-terminal to the DFG

motif. This region may contribute to substrate recognition [4,17]. For example, a conserved lysine-rich region present in p110 α/β and PI3K γ , but absent from the Vps34 family of PI3Ks is similar to the PtdIns(4,5)P₂ recognition motif described for gelsolin [40]. This may be responsible for the differences in substrate specificity between these PI3Ks, thus allowing the p110 family of enzymes to bind PtdIns(4,5)P₂ and use this lipid as a substrate. In the Vps34 family the equivalent location consists of a region rich in proline residues. The functions of homology regions 2 and 3 are currently unknown. Similarity between Vps34 family members extends beyond the three regions conserved with other PI3Ks, suggesting additional interactions specific to this class of the PI3K family.

The biochemical identification of PI3K_{59F} as a PtdIns-specific PI3K supports previous results with yeast and human homologues and distinguishes the catalytic activity of enzymes of the Vps34 family from other classes of PI3Ks which can additionally phosphorylate higher phosphorylated lipids. Thus different classes of PI3Ks appear to have distinct substrate specificities, allowing separate biological roles. This specificity is further facilitated by differences in subcellular location and upstream interactions. The generation of PtdIns3P by Sc_Vps34 appears to be specifically involved in regulating intracellular protein-sorting pathways between the *trans*-Golgi network and the yeast vacuole, although the exact mechanisms by which PtdIns3P regulates trafficking processes are currently unknown. It has been suggested that membrane patches enriched in PtdIns3P may act in the docking/fusion process through the recruitment or activation of specific vesicle coat proteins or other binding partners. Alternatively, PtdIns3P may play a biophysical role, promoting membrane curvature and hence bud formation (reviewed in [41]).

In contrast, PtdIns(3,4,5)P₃, but not PtdIns3P, is increased at the cell membrane following agonist stimulation and implicated in processes, such as mitogenesis, that have been linked to the stimulation of receptors at the cell surface and activation of p110 α/β . PI3Ks which generate PtdIns(3,4,5)P₃ have only been identified in multicellular organisms, indicating that they may have evolved to fulfil their more extensive signalling requirements. However, these PI3Ks may also retain an endocytic function required for cell-surface receptor function. The use of wortmannin and of PDGF- and colony-stimulating factor-receptor mutants lacking the binding site for PI3K, have indicated the involvement of p85/p110 PI3Ks in regulating the sorting of cell-surface receptors in the early endosome, either for recycling to the cell membrane or for degradation in the lysosome [42-44]. p85/p110 PI3Ks have also been shown to be involved in insulin-stimulated recruitment of the glucose transporters GLUT1 and GLUT4 to the plasma membrane. Furthermore, many downstream events resulting from cell-surface receptor activation of PI3K activity require fusion of vesicular structures, cytoskeletal rearrangements and morphological changes. However, a distinct role for a mammalian PtdIns-specific PI3K/Vps34 homologue in the sorting of intracellular proteins from the Golgi to the lysosome (rather than trafficking events proximal to the plasma membrane) is suggested by the missorting and secretion of newly synthesized cathepsin D, a hydrolase that is normally targeted to the lysosome, that occurs in the presence of wortmannin [45,46].

Our current knowledge of intracellular sorting and trafficking processes has largely been gained from both genetic experiments in yeast which have allowed the molecular identification of the proteins involved and mammalian cell-free assay systems which have aimed to reconstitute such events and explain them biochemically. The use of *Drosophila* as a genetically manipu-

latable model for protein trafficking may yield additional information on physiological processes in multicellular organisms. For example, studies on the *Drosophila shibire* mutant, yeast Vps1 mutants and mammalian dynamin, have all contributed to our understanding of the function of the dynamin protein. Analysis of Vps1 in yeast identified a role for this molecule in vesicle trafficking, whereas analysis of *shibire* in *Drosophila* established a function in neuronal secretion in multicellular organisms. A number of female sterile and lethal mutants such as *egalitarian*, *shutdown*, *retained* and *quit*, map to the same cytological region, 59F, as the *Drosophila* Vps34 homologue. These mutations affect oogenesis, a system which involves vesicle transport and in which it will be interesting to investigate whether the production of PtdIns3P may play a role.

We thank Ernst Hafen for provision of cDNA libraries, and B. Vanhaesebroeck and S. Volinia for baculoviruses encoding GST-PI3Kp110 α and GST-PtdIns-3K. C.L. acknowledges an European Union Research Training Fellowship (grant no. ERBBI02CT925107).

REFERENCES

- Stephens, L. R., Jackson, T. R. and Hawkins, P. T. (1993) *Biochim. Biophys. Acta* **1179**, 27–75
- Fry, M. J. (1994) *Biochim. Biophys. Acta* **1226**, 237–268
- MacDougall, L. K., Domin, J. and Waterfield, M. D. (1995) *Curr. Biol.* **5**, 1404–1415
- Zvelebil, M. J., MacDougall, L. K., Leever, S., Volinia, S., Vanhaesebroeck, B., Gout, I., Panayotou, G., Domin, J., Stein, R., Pagès, F. et al. (1996) *Philos. Trans. R. Soc. London B* **351**, 217–223
- Otsu, M., Hiles, I., Gout, I., Fry, M. J., Ruiz-Larrea, F., Panayotou, G., Thompson, A., Dhand, R., Hsuan, J., Totty, N. et al. (1991) *Cell* **65**, 91–104
- Skolnik, E. Y., Margolis, B., Mohammadi, M., Lowenstein, E., Fischer, R., Drepps, A., Ullrich, A. and Schlessinger, J. (1991) *Cell* **65**, 83–90
- Escobedo, J., Navankasattusas, S., Kavanaugh, W. M., Milfay, D., Fried, V. A. and Williams, L. T. (1991) *Cell* **65**, 75–82
- Hiles, I. D., Otsu, M., Volinia, S., Fry, M. J., Gout, I., Dhand, R., Panayotou, G., Ruiz-Larrea, F., Thompson, A., Totty, N. F. et al. (1992) *Cell* **70**, 419–429
- Pons, S., Asano, T., Glasheen, E., Miralpeix, M., Zhang, Y., Fisher, T. L., Myers, Jr., M. G., Sun X. J. and White, M. F. (1995) *Mol. Cell. Biol.* **15**, 4453–4465
- Hu, P., Mondino, A., Skolnik, E. Y. and Schlessinger, J. (1993) *Mol. Cell. Biol.* **13**, 7677–7688
- Stoyanov, B., Volinia, S., Hanck, T., Rubio, I., Loubtchenkov, M., Malek, D., Stoyanova, S., Vanhaesebroeck, B., Dhand, R., Nürnberg, B., Gierschik, P., Seedorf, K., Hsuan, J. J., Waterfield, M. D. and Wetzker, R. (1995) *Science* **269**, 690–693
- Herman, P. K. and Emr, S. D. (1990) *Mol. Cell. Biol.* **10**, 6742–6754
- Chapman, R. E. (1994) *Curr. Biol.* **4**, 1019–1022
- Schu, P. V., Takegawa, K., Fry, M. J., Stack, J. H., Waterfield, M. D. and Emr, S. D. (1993) *Science* **260**, 88–91
- Stack, J. H. and Emr, S. D. (1994) *J. Biol. Chem.* **269**, 31552–31562
- Stephens, L., Cooke, F. T., Walters, R., Jackson, T., Volinia, S., Gout, I., Waterfield, M. D. and Hawkins, P. T. (1994) *Curr. Biol.* **4**, 203–214
- Volinia, S., Dhand, R., Vanhaesebroeck, B., MacDougall, L. K., Stein, R., Zvelebil, M. J., Domin, J., Panaretou, C. and Waterfield, M. D. (1995) *EMBO J.* **14**, 3339–3348
- Hong, Z. and Verma, D. P. S. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 9617–9621
- Welters, P., Takegawa, K., Emr, S. D. and Chrispeels, M. J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 11398–11402
- Zhou, K., Takegawa, K., Emr, S. D. and Firtel, R. A. (1995) *Mol. Cell. Biol.* **15**, 5645–5656
- Takegawa, K., DeWald, D. B. and Emr, S. D. (1995) *J. Cell Sci.* **108**, 3745–3756
- Stack, J. H., Herman, P. K., Schu, P. V. and Emr, S. D. (1993) *EMBO J.* **12**, 2195–2204
- Hunter, T. (1995) *Cell* **83**, 1–4
- Leever, S. J., Weinkove, D., MacDougall, L. K., Hafen, E. and Waterfield, M. D. (1996) *EMBO J.* in press
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
- Davies, A. H., Jowett, J. B. M. and Jones, I. A. (1993) *Biotechnology* **11**, 933–936
- Summers, M. D. and Smith, H. E. (1987) *Tex. Exp. Station Bull.* **1555**, 1–56
- Whitman, M., Kaplan, D. R., Schaffhausen, B., Cantley, L. C. and Roberts, T. M. (1985) *Nature (London)* **315**, 239–242
- Clarke, N. G. and Dawson, R. M. (1981) *Biochem. J.* **195**, 301–309
- Caverner, D. R. and Ray, S. C. (1991) *Nucleic Acids Res.* **19**, 3185–3192
- Rodriguez-Viciana, P., Warne, P. H., Vanhaesebroeck, B., Waterfield, M. D. and Downward, J. (1996) *EMBO J.* **15**, 2442–2451
- Dhand, R., Hara, K., Hiles, I., Bax, B., Gout, I., Panayotou, G., Fry, M. J., Yonezawa, K., Kasuga, M. and Waterfield, M. D. (1994) *EMBO J.* **13**, 511–521
- Woscholski, R., Kodaki, T. and Parker, P. (1995) *Biochem. Soc. Trans.* **23**, 145
- Carpenter, C. L. and Cantley, L. C. (1990) *Biochemistry* **29**, 11147–11156
- Wymann, M. P., Bulgarelli-Leva, G., Zvelebil, M. M. J., Pirola, L., Vanhaesebroeck, B., Waterfield, M. D. and Panayotou, G. (1996) *Mol. Cell. Biol.* **16**, 1722–1733
- Nakanishi, S., Catt, K. J. and Balla, T. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5317–5321
- Hartley, K. O., Gell, D., Smith, G. C. M., Zhang, H., Divecha, N., Connelly, M. A., Admon, A., Lees-Miller, S. P., Anderson, C. W. and Jackson, S. P. (1995) *Cell* **82**, 849–856
- Dhand, R., Hiles, I., Panayotou, G., Roche, S., Fry, M. J., Gout, I., Totty, N. F., Truong, O., Vicendo, P., Yonezawa, K. et al. (1994) *EMBO J.* **13**, 522–533
- Yu, F.-X., Sun, H.-Q., Janmey, P. A. and Yin, H. L. (1992) *J. Biol. Chem.* **267**, 14616–14621
- De Camilli, P., Emr, S. D., McPherson, P. S. and Novick, P. (1996) *Science* **271**, 1533–1539
- Downing, J. R., Rousset, M. F. and Sherr, C. J. (1989) *Mol. Cell. Biol.* **9**, 2890–2896
- Carlberg, K., Tapley, P., Haystead, C. and Rohrschneider, L. (1991) *EMBO J.* **10**, 877–883
- Joly, M., Kazlauskas, A., Fay, F. S. and Corvera, S. (1994) *Science* **263**, 684–687
- Brown, W. J., DeWald, D. B., Emr, S. D., Plutner, W. and Balch, E. (1995) *J. Cell Biol.* **130**, 781–796
- Davidson, H. W. (1995) *J. Cell Biol.* **130**, 797–805