A short history of inositol lipids

Robin F. Irvine

Department of Pharmacology, University of Cambridge, Cambridge CB2 1PD, United Kingdom

Abstract  The diverse family of inositol lipids is now known to be central to many aspects of cell biology. The route from the first discovery of inositol to our present day knowledge of inositol lipids spans more than 150 years and is long and complex. This is a brief account of some of the most important stages along that route.—Irvine, R. F. A short history of inositol lipids. J. Lipid Res. 2016. 57: 1987–1994.

Supplementary key words  phosphoinositides • phospholipases/C • phospholipids • phospholipids/phosphatidylinositol • receptors/hormone • history • phosphatidylinositol • polyphosphoinositol lipids

This is an enormous topic, were full justice to be done to it, because inositol lipids (and their derivatives/cousins, the inositol phosphates) impinge on a substantial proportion of cell biology. But this account will be short and informal, and in places personal. In the latter context, I have been fortunate enough to meet most of the pioneers mentioned below in person, and some of the tales and apocryphal stories come from discussions with them, often related to me after several beers. For those interested in further reading, Bob Michell has written a very full account of the history of inositol lipid signaling (1), and personal recollections from him (2), Clint Ballou (3), Michael Berridge (4), and Lowell Hokin (5) have also been published. As if that were not enough, I have recently published a personal account of the halcyon days of inositol trisphosphate (IP₃) (6).

So perhaps we have enough history in total, and I see my task here being to write a story that is complete insofar as I don’t knowingly miss out on anything really crucial, but short enough to be digestible and entertaining. So this is not a complete history, but a succint one, and I know that I will leave out many important papers and individuals; I hope no one is offended.

EARLY FOUNDATIONS

Inositol (myo-inositol, see below) was first isolated by Scherer (7), and called “inosite” because of its sweet taste. It was fully purified by Maquenne in 1887 (8–10) and its inert weight of its acetyl and benzoyl esters, and other chemical properties led him to establish its cyclohexanol structure by a beautiful series of classic papers; do read them if you can. Maquenne first isolated inositol from leaves, but later used large quantities of horse urine reduced by boiling, which apparently led to complaints from his neighbors.

There are nine different possible isomers of inositol, and it was largely the elegant work of Posternak that established the configuration of the principle inositol in eukaryotic tissues, myo-inositol [(11, 12) and see (13) for review]. Posternak was also responsible for the identification of phytic acid, the major store of organic phosphate in seeds, as being myo-inositol hexakisphosphate in the first decade of the 20th century (14).

Before continuing the main thread of this story, it is worth noting as an aside that other inositols (e.g., scylo, epi, and neo) occur in nature and isomerases that interconvert some of them (whose functions are as yet unknown) may be present in eukaryotes (15). In the soil, inositol hexakisphosphates are found based on isomers other than myo-inositol (16–18), and some eukaryotes, such as Entamoeba histolytica, make neo-inositol hexakisphosphate (19). It could be that there is an entire biochemistry out there of other inositols in other organisms of which we are currently almost completely ignorant.

PHOSPHATIDYLINOSITOL

The esterified existence of inositol in lipids (which from now on will be only the myo isomer) first emerged from bacteria (20), some of which (e.g., the tuberculosis pathogen) have complex inositol lipids [see (3)]. Later, esterified inositol was established as a component of eukaryotic lipids in plants (21) and mammals (22, 23). The pioneering

Abbreviations: DAG, diacylglycerol; PA, phosphatidic acid; PI, phosphatidylinositol; PIP, phosphatidylinositol phosphate; PIP₂, phosphatidylinositol phosphate (x denotes the isomeric position of monooester phosphate); PIP₃, phosphatidylinositol bisphosphate; PI(x,y)P₂, phosphatidylinositol bisphosphate (x,y denote the isomeric positions of monoester phosphates); PI-PLC, phosphoinositide-specific phospholipase C; PLC, phospholipase C.

1To whom correspondence should be addressed.

e-mail: rfi20@cam.ac.uk
chemistry of Clint Ballou and his coworkers led to the demonstration of phosphatidylinositol (PI) having the diester phosphate link between the 3-carbon of glycerol (in sn nomenclature) and the D-1-hydroxyl of the inositol (24). In the modern context of signaling, Ballou’s contributions to the structures of polyphosphoinositol lipids (discussed below) are the best known, but the extensive body of earlier work that included de novo synthesis of new enantiomers to establish PI’s structure (25) was an equally important and elegant piece of work. Parallel work by Dan Brown and his colleagues (26) made an additional crucial contribution to the establishment of PI structure.

ISOMERS AND NOMENCLATURE

This is a good point at which to refresh the reader’s memory on numbering and enantiomers. Anyone who has heard me lecture on inositides will be aware of my fondness for Bernie Agranoff’s turtle analogy (27) to clarify what can be a confusing issue. Figure 1 (adapted from (28)) illustrates this, and can be used by any isomerically challenged reader. The anti-clockwise numbering of the turtle who (like most of us) is right-flippered so his front right flipper is number 1, is then easy to remember. Note the axis of symmetry running through the turtle’s head to his tail, so inositol 1 and 3 phosphate, and inositol 4 and 6 phosphate, are enantiomeric pairs and inositol 2 and 5 phosphate are meso compounds.

Because this review is historical and might induce a reader to look at original papers from the past, it is important to remember that, around the 1970s, the official designation of what is D and what is L in inositol numbering was swapped over (to get L-numbering, start with the turtle’s front left flipper as 1 and number clockwise). Thus, for example, Clinton Ballou’s work (24) on inositol lipid structure established the phosphodiester link between glycerol and inositol as involving what they called the L-1 position of inositol, though it is what we now know as the D-1 position (the front right flipper on the turtle). A similar potential confusion can arise from reading Frank Eisenberg’s (29) classic designation of the two isomers of inositol monophosphate that arise from biosynthesis from glucose-6-phosphate versus alkaline hydrolysis of PI. These are an enantiomeric pair, in modern nomenclature D-inositol 3-phosphate and D-inositol 1-phosphate, respectively, but in the relevant paper (29), they are designated the other way around.

Finally, it is a rule of chemistry that the lowest isomeric numbering should be used if possible, so D-inositol 3-phosphate should more properly be called L-inositol 1-phosphate. This rule was relaxed in the 1990s to make the biochemistry of inositol phosphate metabolic pathways easier to follow, and throughout the rest of this review I will use modern rules and designations without further qualification or discussion.

PI4P AND PI(4,5)P₂

In the 1940s, Jordi Folch-Pi published a series of fractionations of bovine brain (30) that ultimately led to a fraction that was (as we know now) an approximately equal mixture of PI, PI4P [PI phosphate (PIP) (the number denotes the isomeric position of monoester phosphate)], and PI(4,5) P₂ [PI bisphosphate (PIP₂) (the parenthesized numbers denote the isomeric positions of monoester phosphates)] with a little PE (31); it is remarkable just how “pure” this mixture is, given that it has no chromatography in its protocol, only precipitations. Because the ratio of phosphorus to inositol ends up as about 2, Folch-Pi called it “diphosphoinositide” (31). It remained as such until the early 1960s, when two groups took it a stage further. Rex Dawson’s group purified PI(4,5)P₂ and, from its phosphate/inositol ratio of 3:1, called it “triphosphoinositide” (TPI) (32). Concurrently, Clinton Ballou’s group hydrolyzed Folch-Pi’s diphosphoinositide with alkali and identified a variety of inositol phosphates that led them to conclude that PI plus a PI4P and a PI(4,5)P₂ were present (33, 34). Later, Brown and Stewart (35) submitted a large quantity of triphosphoinositide (given to them by Dawson) to a similar analysis to confirm that Dawson’s purified “TPI” was indeed Ballou’s deduced PI(4,5)P₂.

Close reading of one of the works from Ballou’s laboratory (33) reveals a very warm acknowledgment to a Stephen Freer for “dedicated technical assistance.” This has a story behind it, in that isolating the many grams of Folch-Pi fraction that they worked with from more than 100 cow brains...
PI(4,5)P₂ FUNCTION: THE PLC STORY

Actually, it is functions in the plural we are talking about, as we now know PI(4,5)P₂ to be one of the most multi-functional molecules in biology (36). But from a historical viewpoint, the whole story of inositol lipid function focused on, and was inextricably linked to, its role as a generator of two second messengers, diacylglycerol (DAG) and inositol 1,4,5-trisphosphate, by virtue of its being split in half by members of the phosphoinositide-specific phospholipase C (PI-PLC) family (Fig. 2).

In the early 1950s, Lowell and Mabel Hokin were working in Sheffield investigating the control of RNA synthesis by incorporation of ³²P, and discovered a clear increase when pancreatic slices were stimulated with acetylcholine. They were due to move to Canada (McGill University), so they packed up the radioactive material and took it with them over the Atlantic (imagine doing that in the 21st century…) and proceeded to purify the RNA from the extracts. As they purified it, the radioactivity disappeared, so they analyzed the radioactive “junk” they had discarded and found that it was phospholipid. They published this seminal observation in 1953 (37). As Mabel later recalled, at this stage they could go no further because simple ways of separating and analyzing phospholipids from small quantities of tissue did not then exist. But the following year, Rex Dawson published his classic analytical method based on alkaline deacylation and two-dimensional separation of the glycerophosphoesters (38). A further observation he made, not usually noted by historians (because most people do not read the paper), was that if slices of tissue (brain or pancreas) were acutely labeled with ³²P, the majority of radioactivity was in phosphatidic acid (PA) and inositides, especially the diphosphoinositide fraction (38). The Hokinseized this method and confirmed Dawson’s observations for their acetylcholine-stimulated pancreatic slices: the increase in labeling they observed was almost all in PA and PI (39). The “phosphoinositide effect” was born.

The long and complex story of how this phenomenon was studied and discussed at length for the next 30 years is too much for relating here; Bob Michell has written an excellent account of these events (1, 2). There is an interesting theme that emerges from these studies, which illustrates the general principle of how you measure things can dictate what you see. For many years, the stimulated incorporation of ³²P into PA and PI was the easiest and, therefore, most widely used method, but this is of course indirect, and its sensitivity stems from the fact that the DAG generated by PI-PLC is phosphorylated into PA by DAG kinase, which is then incorporated directly (after activation to CMP-PA, also known as CDP-DAG (40)), so if the tissue is not labeled to equilibrium, this radiolabel incorporation exaggerates...
the actual mass rate of synthesis. Thus, even if there were a stimulation of a PC-specific PLC, because the phosphocho-
loline and CDP-choline pools would still be at a lower specific activity than the γ phosphate of ATP, the most prominent effect with acute 32P labeling could still be a specific increase in PA and PI labeling. Quantifying PI breakdown was challenging (41, 42), not least because under most circumstances when PI-PLC is activated there is no detectable drop in PI mass levels (the amount of PI consumed to re-
plish PIP2 levels is quite small); indeed, even PIP2 often does not go down detectably due to rapid compensatory resynthesis (see, e.g., (43) for an early report of this).

So the indirect method of measuring the “PI effect” was used almost exclusively and this unsurprisingly led to some false directions. Nevertheless, the discovery of a specific PI-
PLC (44, 45) and the PI decreases (41, 42) led to the gen-
eral consensus that a hydrolysis of inositol lipids was the likely primary reaction (46, 47). The PI-PLC family is now known to be a large one, with complex regulations by cell surface receptors as well as intracellular factors such as mono-
nemonic G-proteins. Many laboratories contributed to that story, but most people would agree that Sue-Goo Rhee’s extensive set of studies [see (48) for review] was the major clarifying factor.

With the benefit of hindsight, the earliest guess nearest to correctness as to inositol lipid function in the 1960s came from Jack Durell and his coworkers, who suggested that TPI hydrolysis was connected to acetylcholine recep-
tor activation, though they proposed that nicotinic (rather than muscarinic) receptors were responsible (49). The real breakthrough and cohesive synthesis of inositol function came with Bob Michell’s (47) seminal suggestion that PI-
PLC activation was a part of the signaling function of a group of receptors distinct from those coupled to adenyl cyclase. Moreover, he crucially suggested that increases in calcium influx were a part of the response to PI turnover (47).

For a while, there was a lot of argument as to which out of calcium rises and PI turnover was the chicken or the egg. Here again, the method of measurement could cause confu-
sion. For example, for a number of years arguments raged about whether the neutrophil, in which the increase in PA labeling was clearly dependent on extracellular calcium being present (50), was a “black swan,” which dis-
proved the hypotheses that all swans were white (i.e., if PI turnover preceded calcium in the pathway, it should always be calcium-independent). But Shamshad Cockcroft (50) eventually showed that the PA being quantified was at a lower specific activity than the PI, and indeed had a differ-
ent fatty acid profile (51). In fact, a large (extracellular calcium-dependent) increased activation of phospholipase D activity on PC was confusing the whole issue, and so the neutrophil was actually a rather dirty swan. The clearest experi-
mental link between calcium homeostasis and the PI effect came from the classic demonstration by Berridge and Fain (52) that blowfly salivary glands depended on a supply of inositol to support sustained calcium rises.

Arguments about calcium dependency also muddied the waters for a time as to whether PI or PIP2 was the primary substrate of PI-PLC in vivo [see e.g., (53)]. Not least of the arguments against PIP2 was the question, why do cells have so much PLC activity against PI if PIP2 is the major sub-
strate in the cell? This was answered by the demonstration that the various PI-PLCs could hydrolyze both substrates (54) and that they favor PIP2 under physiological assay condi-
tions (55, 56), which suggested that Dawson’s TPI-phos-
phodiesterase (57) and PI-phosphodiesterase (44, 45) were probably the same beast.

While on the subject of TPI-phosphodiesterase, this is a good place to mention that when it was discovered Dawson and Thompson (58) also described a phosphatase with differ-
ent properties from the phosphodiesterase (e.g., not calcium-dependent, but requiring magnesium), which re-
moved the monoester phosphate from PIP2 to generate PIP. This was the very beginning of the polyphosphoinosi-
tol lipid phosphatase story. This family of enzymes has grown enormously in the intervening years, and is now rec-
ognized as having major physiological and pathological sig-
nificance [see (59) for a recent review].

The final resolution of the PI/PIP2 argument came about when the products of hydrolysis were analyzed with [3H]inositol labeling of tissues. In particular, Mike Ber-
ridge’s demonstration that IP3 was clearly the product formed first (within seconds) after blowfly salivary glands were stimulated (60) was crucial, and he later showed with some elegant stop-flow and electrophysiological exper-
iments that this increase preceded any physiologically sign-
ificant detectable rise in calcium (61). The proposal (60) that IP3 might therefore be the mediator of the calcium rises was followed by the experimental demonstration that this was so by Streb et al. (62) in the same year.

As mentioned above, Mike Berridge (4) and I (6) have both written personal accounts of the events that led to the publication of Streb et al. (62), so I will not add to the length of this review by reiterating any of that. I also de-
scribe (6) how the concurrent discovery of 134IP3 started the proliferation of inositol phosphates, with their many functions [e.g., (63–65)]. Inositol phosphates are not lipids by any stretch of the imagination, so their story does not belong here, but it is important to remember that in most euksarotes (plants and slime molds being the most notable exceptions) the only route to the inositol phosphate family goes via the generation of 1,4,5-trisphosphate by PI-PLC, and not by phosphorylation all the way up from inositol; thus, the first three phosphorylations of the inositol ring take place in the lipids. So even in organisms such as yeast where IP3 has no known function itself, it is the essential precursor of all the other IPs (66).

The other product of PI-PLC is DAG, and that is indeed a lipid and is a centrally important second messenger. Nishizuka and colleagues’ seminal discovery of PKC as a DAG-activated protein kinase happened in 1979 (67), before IP3 was in any way suspected of having a function, drew many people into taking an interest in inositol lipids. This discovery was a typical example of impure compounds aid-
ing advances; when Nishizuka’s group switched over to an alternative commercial source of phosphatidylerine (re-
quired by PKC) that was pure, and was not contaminated
by DAG, the massive decrease in PKC activity led them to identify the missing stimulant. The double precursor role of PI(4,5)P₂ as a generator of two synergizing second messengers is a truly remarkable aspect of cell physiology. PI(4,5)P₂ is now known to have many other functions in its own right [see (36) for a comprehensive and magisterial review of this topic].

THE “OTHER” POLYPHOSPHOINOSITOL LIPIDS

PI4P, for a long time thought to be “just” a precursor to PI(4,5)P₂, is now recognized as having many physiological functions, in particular in intracellular membranes such as the Golgi [reviewed by Balla (36)], though it may also have a distinct crucial function in contributing to the high negative charge of the plasma membrane (68).

The remaining five polyphosphoinositol lipids (Fig. 3) are present in eukaryotic cells at one to two orders of magnitude lower than the levels of PI4P and PI(4,5)P₂, so they eluded discovery for a long time. The advent of HPLC analysis of deacylated [³H]inositol-labeled lipids was the major spur for discovery and analysis of these physiologically crucial cell components. The origins of the 3-phosphorylated family stem back to the discovery of what was thought to be a PI 4-kinase activity associated with oncogenes src and middle T antigen (69, 70). This activity was originally thought to be perhaps boosting PI4P levels to amplify the PI-PLC signals, but the continuation of this story by Cantley’s group showed that the oncogene-associated activity was molecularly distinct from the other PI kinases, and that it was differentially regulated by tyrosine kinases (71). Finally, they showed that their kinase was a PI 3-kinase (72). PI34P₂ and PIP₃ emerged as products produced in vitro by this activity if given the appropriate substrate (73), although PIP₃ was first reported in intact cells as a rapidly-produced lipid in stimulated neutrophils (74).

For a while there was considerable confusion, not unlike that earlier for PLC, as to which of the three 3-phosphorylated inositol lipids was produced first in vivo and thus which might be the primary functional molecule. This confusion was not helped by the identification by Waterfield and Emr’s groups of PI 3-kinase as a homolog of the yeast vps34 protein, which would only use PI as a substrate to generate PI3P (75) and would not phosphorylate PIP₂.

The clarification of PIP₃ as a molecule generated directly by 3-phosphorylation of PI(4,5)P₂ by a receptor-stimulated enzyme (thus echoing the generation of IP₃) came largely from Len Stephens and Phill Hawkins and their colleagues. They showed first, that we mammals have a PI-specific 3-kinase similar to yeast, but distinct from the enzyme(s) (there are four in vertebrates) that specifically phosphorylate PI(4,5)P₂, so PI3P is a separate story from PIP₃ (76). More importantly, elegant analysis of the individual phosphate moieties on the inositol ring under nonequilibrium labeling conditions established the receptor-stimulated activity in vivo as being 3-phosphorylation of PI(4,5)P₂ (77, 78). PIP₃ has proved to be a second messenger of immense physiological significance, with numerous effectors (the two inter-twined protein kinases, Akt and PDK, being the most famous), and I do not propose to tell that complex story here: Alex Toker (79) has written an excellent historical account of the phosphoinositide 3-kinases.

![Fig. 3. The polyphosphoinositol lipids. Adapted from (99). PI (A) and a schematic depiction of PI and the seven polyphosphoinositol lipids (B).](image-url)

A short history of inositol lipids 1991
PI3P is now appreciated as a crucial part of cellular trafficking, and PI(3,4)P_2 is finally coming into its own,orney to note in this historical context that PI3P may have been the first polyphosphoinositol lipid to appear during evolution. We eukaryotes make PI by combining inositol with activated PA (CMP-PA, also known as CDP-DAG) (40), but some bacteria, such as Mycobacterium, apparently use inositol 3-phosphate (synthesized from glucose 6-phosphate) instead to make PI3P, which is then dephosphorylated to PI (80). Archaea do something similar, though their lipids are based on backbones different from the 1,2,DAG [see, e.g., (81) for review]. Whether PI3P has a function in these organisms, other than as a synthetic precursor to PI, is not known.

PI(3,5)P_2 was discovered independently by Charles Brearley and his colleagues (82) and by the combined groups of Peter Parker and Bob Michell (83), who identified it on HPLC profiles as a peak greatly enhanced by osmotic stress in yeast; this lipid is again also now appreciated as a major player in cell trafficking regulation (36).

Finally, Lucia Rameh in Lew Cantley’s group was investigating the specificity of a PI(4,5)P_2 5-phosphatase, specificity being, in this context, how often it removed the 4-phosphate instead of the 5-phosphate. She quantified this side reaction (about 5% of the total activity) by the production of labeled PI5P from 32P-5-labeled PI2P, which had been prepared by incubating PHP with a recombinant PHP 5-kinase and 32P-ATP. Lucia ran out of recombinant PI P kinase, and used some enzyme purified from red cells that was in the laboratory deep freeze in its place, and was astonished to see a massive generation of radiolabeled PIP by the phosphatase. The reason, elegantly teased out, was that the PHP (from brain) was contaminated with some PI5P, and that the red cell enzyme, a type II PIP kinase, is actually a PI5 4-kinase, not a PHP 5-kinase (84). So this batch of 32PI(4,5)P_2 was instead labeled in the 4-position, and consequently the predominant 5-phosphate activity of the enzyme being studied now generated copious quantities of radiolabeled PIP! The identification in animal tissues of the natural substrate of type II PIP kinases (now known as PI5P4Ks), PI5P, followed (84), and this lipid is now also established as having multiple effectors in the cytoplasm and nucleus [see (85–87) for reviews].

Mention of the nucleus prompts the aside that nuclear inositol lipid metabolism and function, first established in the late 80s/early 90s (88, 89) is a growing aspect of the inositol lipid field (see, e.g., (90) for review) and reflects yet again the diversity and ubiquity of these minor, but crucially important, lipids.

PI GLYCAN ANCHORS

Finally, these should be included here as another remarkable diversification of inositol lipid chemistry and function. It was known by the 1970s that some snake venoms and bacterial extracts could cause release of enzymes into the bloodstream [see, e.g., (91) for early references] and it was Martin Low and his colleagues who showed that a bacterial PI-specific PLC liberated these proteins because they were bound to the outside surface of cells by a covalent link to an inositol lipid (91–93). The full structure of the complex inositol lipid (the 6-hydroxyl of the inositol ring is employed here) was worked out by Mike Ferguson and his colleagues [e.g., (94) and see (95) for a review of the complex chemistry of PI-glycans in parasitic protozoa]. In addition to being a releasable way of attaching proteins to the outside of our cells (of particular significance in immune cells), this link is employed by the parasites to rapidly shed their coat proteins and thus avoid immune surveillance.

CONCLUSION

When we look at just how many inositides there are, doing how many amazing things, it is not surprising that the emergence of our current knowledge was a complex path with many false leads and took decades. As with any such tale, it is much easier told with the benefit of hindsight, but if there is one lesson for us all, it is that almost all of the discoveries described above were made from what we would now call “blue-skies” research, with no clear aim or practical endpoint other than the uncovering of knowledge. How many of the studies described above would be funded now? No matter how many times we tell those who try to direct science that much of it cannot be directed, they don’t seem to listen.

The author thanks all those inositide pioneers whom he has met over 40 years for insights into how some of these events happened, and especially Bob Michell for his unmatched historical knowledge. The author’s involvement with inositol lipids and phosphates began when, as an undergraduate, he read a review by Rex Dawson (96), which led to the author joining his lab (97). The author is acutely conscious of the debt that he and the whole field owe Rex Dawson.

REFERENCES


