Evolution of the diverse biological roles of inositols

Robert H. Michell¹

School of Biosciences, University of Birmingham, Birmingham B15 2TT, U.K.

Abstract

Several of the nine hexahydroxycyclohexanes (inositols) have functions in Biology, with myo-inositol (Ins) in most of the starring roles; and Ins polyphosphates are amongst the most abundant organic phosphate constituents on Earth. Many Archaea make Ins and use it as a component of diphytanyl membrane phospholipids and the thermoprotective solute di-1,1′-Ins-1,1′′-phosphate. Few bacteria make Ins or use it, other than as a carbon source. Those that do include hyperthermophilic Thermotogales (which also employ di-1,1′-Ins-1,1′′-phosphate) and actinomycetes such as Mycobacterium spp. (which use mycothiol, an inositol-containing thiol, as an intracellular redox reagent and have characteristic phosphatidylinositol-linked surface oligosaccharides). Bacteria acquired their Ins₃P synthases by lateral gene transfer from Archaea. Many eukaryotes, including stressed plants, insects, deep-sea animals and kidney tubule cells, adapt to environmental variation by making or accumulating diverse inositol derivatives as 'compatible' solutes. Eukaryotes use phosphatidylinositol derivatives for numerous roles in cell signalling and regulation and in protein anchoring at the cell surface. Remarkably, the diradylglycerol cores of archaeal and eukaryote/bacterial glycerophospholipids have mirror image configurations: sn-2,3 and sn-1,2 respectively. Multicellular animals and amoebozoans exhibit the greatest variety of functions for PtdIns derivatives, including the use of PtdIns(3,4,5)P₃ as a signal. Evolutionarily, it seems likely that (i) early archaea made myo-inositol approx. 3500 Ma (million years) ago; (ii) archeaons brought inositol derivatives into early eukaryotes (approx. 2000 Ma?); (iii) soon thereafter, eukaryotes established ubiquitous functions for phosphoinositides in membrane trafficking and Ins polyphosphate synthesis; and (iv) since approx. 1000 Ma, further waves of functional diversification in amoebzoans and metazoans have introduced Ins(1,4,5)P₃ receptor Ca²⁺ channels and the messenger role of PtdIns(3,4,5)P₃.

¹email r.h.michell@bham.ac.uk
Introduction

Inositols (hexahydroxycyclohexanes) are key components of diverse biological molecules with a remarkable variety of functions. As the contributions to this Symposium illustrate, important signalling and regulatory events in animals have mainly dominated research on inositol derivatives. Early work is accessible by reference to reviews [1–5]. Most notably, PtdIns(4,5)P₂ stars in two widespread transmembrane signalling mechanisms: formation of the second messengers sn-1,2-diacylglycerol and Ins(1,4,5)P₃ by receptor-controlled PLC (phosphoinositidases C), and synthesis of the membrane-associated messenger PtdIns(3,4,5)P₃ by receptor-controlled PPIn3K-I (Type I phosphoinositide 3-kinases).

Each of the seven known PPIn (polyphosphoinositides), which are phosphorylated derivatives of the sn-1,2-diacyl-glycerol-3-phosphate-based membrane lipid PtdIns, has distinctive roles in regulating cellular processes, including exocytosis, membrane trafficking and cell movement. PtdIns (or, sometimes a physically similar inositol sphingolipid) is also the hallmark constituent of the GPI-(glycosylphosphatidylinositol) or GIPL (glycosylinositolphospholipid) anchors that tether many external proteins and complex glycans to eukaryote cells. Eukaryotes make a spectrum of myo-inositol polyphosphates and pyrophosphate derivatives thereof, for most of which functions remain to be found.

These widely studied compounds and processes, which are mainly restricted to eukaryotes, represent only part of the diverse spectrum of biological inositol compounds and functions. For example, inositol (and simple derivatives thereof) are widely distributed as compatible solutes that cells accumulate when their hosts encounter stresses such as drought, environmental salinity or over-heating. And myo-, scylo-, d-chiro- and neo-inositol polyphosphates contribute much of the organic phosphate of soils and of lake and estuarine sediments [4,6–12], hence the suggestion that ignorance about the origins and fates of environmental inositol polyphosphates constitutes “the greatest gap in our understanding of the global phosphorus cycle” [9]: see also a recent set of meeting abstracts (http://striweb.si.edu/inositol_conference/PDFs/Abstract_Book_Content.pdf).

The present article has two main aims. The first is to enumerate and briefly discuss the inositol-containing and/or inositol-related compounds that diverse organisms make and the astonishing variety of roles they fulfil: Figure 1 depicts some of these molecules. ‘New’ biological inositol derivatives are regularly added, and more almost certainly remain to be found. The second is to try to discover some clues as to how the remarkably versatile biological exploitation of inositols might have evolved, particularly during the early aeons of evolution during which the major classes of extant organisms (Bacteria, Archaea, Eukarya and subsets within each) evolved from their still-obscure progenitors.

The evolutionary context

Organisms that make inositol derivatives occur in all three Kingdoms of Life. Figure 2 offers a simplified Tree of Life, emphasizing organisms for which information about usage of inositols is available. This tree draws on recent models
Figure 1 A variety of the biological inositol derivatives discussed in this article. (A) The inositols, with myo-inositol and Agranoff’s mnemonic of a myo-inositol turtle both numbered in accordance with the ‘always D-numbering’ convention that is discussed in the text. (B) Ins3P3’Ins, used as a compatible solute by some archaea and a few bacteria. (C) Pinitol. (D) Gentobiosylcaldar-chaetidylinositol. (E) Archaetidylinositol. (F) Stearoyl.arachidonyl-PtdIns, with the three hydroxy groups that are phosphorylated in various PPIn highlighted. (G) A space-filling model of stearoyl.arachidonyl-PtdIns(4,5)P2. (H) Ceramide-phospho-1-Ins. (I) InsP6. (J) 5-PP-InsP6. (K) Mycothiol. (L) Axinelloside A, with the integral scyllo-inositol annotated.

© 2007 The Biochemical Society
of the contentious relationships between the various eukaryote groups: these envisage the metazoan, fungal and amoebozoan clades diverging during much the same period [13–17].

The context of the discussion that follows is that the Archaea/Bacteria divergence probably predated the present by approx. 3500 million years (Ma), the earliest divergences amongst primitive eukaryotes probably occurred before 2000 Ma, and the various phyla within the eukaryote ‘crown group’ probably diverged during the period 1200–700 Ma.

Annotated onto this simplified tree (Figure 2) are boxes that summarise which of the various roles currently played by inositols characterize extant members of each group of organisms. As discussed below, these suggest that myo-inositol was ‘invented’ very early by Archaea, and that evolution of major new functions has continued since the separation of the lineages leading to amoebozoans and metazoans around 1000 Ma ago.

**Myo-inositol and other inositols**

There are nine isomeric inositols (Figure 1A). Their all-carbon cyclohexane ring makes them remarkably chemically stable. Myo-inositol is the most widespread and versatile in Biology, but several others (including scyllo-inositol, neo-inositol, d-chiro-inositol, epi-inositol and muco-inositol), are found in
Evolution of the diverse biological roles of inositols

diverse places or organisms and in multiple forms. Later mentions simply of ‘inositol’ refer only to myo-inositol: other isomers are always named.

Cells probably make inositol isomers other than myo-inositol mainly by inverting the hydroxy group configuration around one or more of the ring carbons of myo-inositol [18–20], but there may be other routes at least to some isomers [21]. We are largely ignorant about the origins and functions of biological compounds containing ‘other’ inositols, and this should be fertile ground for future research.

Inositol biosynthesis

All organisms that make inositol use evolutionarily related Ins3P synthases (inositol monophosphate synthases; often termed MIPS), of which Ino1p of Saccharomyces cerevisiae is the prototype [22–28]. These mainly cytosolic [29] Ins3P synthases are NADH-dependent cycloaldolases that convert D-glucose-6-phosphate, via bound intermediates, into L-Ins1P.

L-Ins1P is also known as D-myoinositol 3-phosphate, or, simply and conveniently, as inositol 3-phosphate (abbreviated to Ins3P, with Ins meaning ‘myo-inositol with the numbering of the 1D configuration’). This convention of numbering all biological myo-inositol (Ins) derivatives in relation to 1D-myoinositol was introduced to obviate the difficulties that otherwise arise when rigid adherence to standard stereochemical numbering rules demands confusing switches between 1D- and 1L-numbering midway through metabolic pathways (see http://www.chem.qmul.ac.uk/iupac/cyclitol/myo.html).

Inositol in Archaea and Bacteria

The majority of archaean genomes encode an Ins3P synthase, and many archaeons incorporate the resulting inositol into their characteristic sn-1-phosphoryl-2, 3-diether-based membrane glycerolipids (discussed later). A subset of hyperthermophilic Archaea accumulates a di-myoinositol phosphodiester when they are exposed to temperatures higher than their normal growth conditions (see the section on compatible solutes). There have been indications, yet to be validated, that some archaeal surface proteins may be attached through eukaryote-like GPI-anchors [30–32].

By contrast, most bacterial genomes have no Ins3P synthase gene. Those that do are mostly in the actinobacterial clade (e.g. Mycobacterium spp.) or are obligate hyperthermophiles (e.g. members of the related Thermotoga and Aquifex clades). Bacteria of the Thermotoga/Aquifex grouping are genetically extraordinary, in that almost one-quarter of their genes have been imported from Archaea [33]. Closely related Ins3P synthases have also been noted sporadically in a few other bacteria and in the cyanobacterium Synechocystis PCC 6803 [25,34,35]. Sequence comparisons reveal that all bacterial Ins3P synthases are more closely related to archaean than to eukaryal Ins3P synthases [34,36]. This suggests that the ‘standard’ bacterial enzyme toolkit lacks an Ins3P synthase, and that bacteria

© 2007 The Biochemical Society
have through the ages acquired their Ins3P synthases (*Streptomyces coelicolor* has three!) by lateral gene transfer from co-existing Archaea [34].

A few gram-negative bacteria contain small amounts of PtdIns (see below), but most have none. I once made a total acid hydrolysate of approx. 5 g of *Escherichia coli*, trimethylsilylated the dried residue, subjected it to GLC and was disappointed to find no sign of trimethylsilyl Ins [3]. I thus realized that Ins is not a universal ingredient of living cells, as both Posternak [2] and I had imagined it to be!

So Ins is a core component of many Archaea, but not of most bacteria. A few bacteria have recruited Ins for particular purposes (see also the later comments on compatible solutes and on actinomycetes). For example, some *ice*” gram-negative bacteria have an inherited propensity to act as ice nucleators, with consequent damage to plants that they colonize. This ability is primarily conferred by ice-nucleating proteins, but some of the most effective ice nucleators have the highest, albeit small, PtdIns content [37]. Whether this might be achieved by ice-nucleating proteins being GPI-anchored on the bacterial surface remains uncertain [38,39]. A complex *chiro*-inositol-containing polysaccharide from *Bifidobacterium bifidum* is an enigmatic bacterial material [40].

The most parsimonious evolutionary interpretation of these observations is that Archaea invented Ins synthesis from glucose-6-phosphate and adopted Ins as a widespread membrane phospholipid headgroup early in their evolution, may be as long as 2500 Ma ago, and lateral gene transfer quite soon passed this ability to other organisms. Ins was probably first made, therefore, by some early, and relatively simple, primeval archaeon living in an environment over which it had little control. It seems that development of this ability to redirect a minor flux of carbon along a very short side pathway from central energy metabolism into a small and chemically very stable polar entity with versatile properties must have been a beneficial evolutionary stratagem.

**Inositol in eukaryotes**

Eukaryotes, with rare exceptions, always have an Ins3P synthase gene. All eukaryotes exploit inositol derivatives for numerous cellular roles (Table 1 and Figure 2), most notably as components of membrane phospholipids and protein anchors, and as polar hydrophilic solutes involved in cellular adjustment to environmental challenges (see below).

Even the few eukaryotes that cannot make their own Ins need it (*e.g.* *Schizosaccharomyces pombe*). They often live in Ins polyphosphate-rich environments from which Ins can readily be obtained [41]: an appropriate secreted phosphatase (phytase) and a transporter with which to pump Ins into the cell are all that they need. Few tissues in metazoans, the most complex eukaryotes, make their own Ins, notably the brain and testis in mammals [42,43]. Despite this universal requirement, it is hard to establish Ins deficiency in experimental mammals. Except in gerbils, this tends to become manifest only when animals are prevented from obtaining Ins from their food or from micro-organisms in their gut contents, or during lactation-driven nutrient stress [2,44–53].
### Table 1 Cyclitol-based compatible solutes

<table>
<thead>
<tr>
<th>Compound</th>
<th>Organisms, effects and behaviours</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Archaea</strong></td>
<td>Many hyperthermophilic archaea. Accumulate when temperature raised further, and in stationary phase.</td>
<td>[55,58,60,62,253]</td>
</tr>
<tr>
<td>Ins3P3Ins &amp; Man-2-β-O-Ins3P3Ins-O-2-β-Man</td>
<td>Hyperthermophilic Thermotoga spp., accumulate with raised temperature or salt.</td>
<td>[63,64,254]</td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scyllo-inositol</td>
<td>More in the tissues of echinoderms, gastropods and polychaetes that live in marine abyssal deeps.</td>
<td>[56,91,255]</td>
</tr>
<tr>
<td>Myo-inositol</td>
<td>Renomedullary cells</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dehydrated collembolans (springtails)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hibernating beetles</td>
<td></td>
</tr>
<tr>
<td>Scyllo-inositol and myo-inositol</td>
<td></td>
<td>[90]</td>
</tr>
<tr>
<td><strong>Animals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiro-inositol</td>
<td>Numerous species.</td>
<td>[2,68]</td>
</tr>
<tr>
<td>L-Quebrachitol [1L-(−)-2-O-methyl-chiro-inositol]</td>
<td>Oaks; rubber (Hevea brasiliensis) latex; maple (Acer saccharum) sap; some Proteacea</td>
<td>[2,69,76]</td>
</tr>
<tr>
<td>D-Pinitol (1D-3-O-methyl-chiro-inositol)</td>
<td>Some eucalypts; Limonium gmelini; Mesembryanthemum crystallina (ice-plant); some Proteacea; soybean (Glycine max).</td>
<td>[2,24,66,67,91,71,72,74,259]</td>
</tr>
<tr>
<td>(+)Ononitol (1D-4-O-methyl-myio-inositol)</td>
<td>Ice-plant (Mesembryanthemum crystallinum); transgenic tobacco; legumes, including rice bean (Vigna unguiculata); chickpea (Gicr arizanum); etc.</td>
<td>[2,24,65,70,78,260,261]</td>
</tr>
<tr>
<td>Sequoyitol (1D-5-O-methyl-myio-inositol)</td>
<td>Soybean; Aristolochia gigantea; can be epimerised to pinitol.</td>
<td>[262-264]</td>
</tr>
<tr>
<td>(−)Bomesitol (1D-1-O-methyl-myio-inositol)</td>
<td>Some Proteacea and legumes. Enantiomer in rubber latex</td>
<td>[2,69]</td>
</tr>
<tr>
<td>Viscumitol (1D-1,2-O-dimethyl-myo-inositol)</td>
<td>Mistletoe (Viscum album)</td>
<td>[265,266]</td>
</tr>
<tr>
<td>Liriodendritol (1D-1,4-di-O-methyl myo-inositol)</td>
<td>Litiodendron tulipifera (tulip tree)</td>
<td>[2]</td>
</tr>
<tr>
<td>Pinpollitol (0D-1,4-di-O-methyl-chiro-inositol)</td>
<td>All gymnosperms</td>
<td>[267,268]</td>
</tr>
<tr>
<td>(+)Quercitol (1D-1,3,4,2,5-cyclohexane-pentol)</td>
<td>Various plants, including sugar maple and many eucalypts.</td>
<td>[72,75,76]</td>
</tr>
<tr>
<td>Fagopyritols (pinitol, mono-, dis- or tris-galactosylated on 1 and/or 3 positions)</td>
<td>Buckwheat (Fagopyrum spp.): levels respond to desiccation stress.</td>
<td>[80,259]</td>
</tr>
<tr>
<td>Lathyritol [α-D-galactosyl-(1→3)-bomesitol]</td>
<td>Sweet Pea</td>
<td>[269]</td>
</tr>
</tbody>
</table>
Inositol derivatives as stress-adaptive compatible solutes

Many organisms make or accumulate substantial intracellular concentrations of small and polar, and often uncharged, cytosolic solutes, often termed compatible solutes, to protect themselves against environmental stresses such as extremes of heat, salt exposure, hydrostatic pressure or exposure to reactive oxygen species [24,54–59]. Compatible solutes include various polyols (glycerol, sorbitol, disaccharides such as trehalose etc.), methylamines [e.g. TMAO (trimethylamine N-oxide)], glycerophosphocholine, taurine and betaine, and also proline, glycine, β-alanine and other amino acids. The physically benign properties of these molecules allow cells to accumulate them to high concentrations and to tolerate large changes of concentration without major adverse effects on cell function. Their protective effects include protein stabilization, osmotic compensation and supercooling of tissue water [56].

Numerous organisms throughout Biology employ inositols and simple derivatives thereof, under the generic term cyclitols, as compatible solutes: these are listed in Table 1.

Certain archaeons, notably those that live in very hot environments such as volcanic springs, have evolved adaptations by which they survive transient exposure to even more extreme temperatures than those at which they normally grow. One such is accumulation of di-L-myo-inositol-1,1′-phosphate (sometimes abbreviated as DIP; Figure 1B) [60]. Using the all-D numbering, this compound can be systematically abbreviated as Ins₃P₁Ins. It is probably made by the transfer of Ins₃P from Ins₃P-CMP to C-3 of a free Ins molecule [61].

The estimated intracellular concentration of Ins₃P₁Ins increases from approx. 60 mM to approx. 200 mM after Thermococcus celer is shifted from its normal growth temperature of 88°C to 94°C [62]. The concentrations of other archael compatible solutes, such as mannosyglycerate, often rise when cells are exposed to thermal, hyperosmotic or salt stress, but the protective action of Ins₃P₁Ins seems to be fairly specific to heat stress [54,62], a notion supported by the ability of Ins₃P₁Ins to protect isolated Pyrococcus woesei glyceraldehyde-phosphate dehydrogenase against thermal inactivation [60]. Ins₃P₁Ins also accumulates during stationary phase [62].

As with Ins₃P synthase (see above), the hyperthermophilic bacteria Thermotoga maritima and Thermotoga neapolitana make Ins₃P₁Ins (plus the related stereoisomer Ins₃P₁Ins and the dimannosylated derivative Man-2-β-O-Ins₃P₁Ins-O-2-β-Man [63,64]), this time in response to both salt and heat stress. Given the remarkable genetics of these bacteria (see above), their ability to make Ins₃P₁Ins is likely to be a result of lateral gene transfer from a hyperthermophilic archaeon.

The best understood inositol-based compatible solutes are diverse molecules that accumulate in stressed (or stress-tolerant) eukaryotes, particularly plants. These include methyl-inositol and deoxy-inositol derivatives, sometimes with one or more attached sugars [2,24,65–73]. For example, pinitol (Figure 1C) is synthesized thus:

\[
\text{myo-Ins} \rightarrow \text{3-O-methyl-mylo-Ins (ononitol)} \rightarrow \\
\text{3-O-methyl-d-chiro-Ins (pinitol)}
\]
The accumulated cyclitols are often studied in leguminous plants, in which some are abundant. Changes in cytosolic concentrations of various methyl-inositols, especially pinitol, correlate with adaptation to arid or saline conditions. For example, soybean (Glycine max) cultivars from semi-arid regions of China constitutively make more pinitol than those from regions with higher rainfall, and water restriction of soybean plants promotes a graded accumulation of pinitol in their stems and leaves [71,74]. The cyclohexane-pentol quercitol is the predominant cyclitol in eucalypts from arid environments [72,75], and quebrachitol and/or bornesitol are the most abundant cyclitols in many Proteaceae [69]. Quebrachitol is, after sucrose, the dominant polyol of the winter sap of sugar maple (Acer saccharum) and of maple syrup [76].

Not only do some salt-tolerant plants accumulate cyclitols when saline-stressed [68], but the Ins3P synthases that support cyclitol accumulation may be optimized for activity in high salt conditions [26]. Transfer of the stress-regulated Ins methyltransferase of the ice-plant Mesembryanthemum crystallinum [77], a facultative halophyte that makes ononitol from myo-inositol, into tobacco plants that normally lack this activity makes the tobacco plants more tolerant of drought and salt stress, at least partly because of stress-induced myo-inositol and ononitol accumulation [78].

The raffinose series of sucrose-based oligosaccharides, which includes raffinose (trisaccharide), stachyose (tetrasaccharide) and verbascose (pentasaccharide), and the fagopyritols (galactosylated derivatives of pinitol) are structurally and functionally related. These soluble oligosaccharides accumulate in diverse plants as carbohydrate stores during seed development and they sometimes also confer stress-resistance [24,79–84] (See Table 1). Galactinol (1α-D-galactosyl-(1>3)-myo-inositol), made from myo-inositol and UDP-galactose by a family of stress-inducible galactinol synthases, is the galactose donor for their synthesis, but only some of them retain a cyclitol in their final structure. For example:

\[
\text{UDP-Gal + myo-Ins} \rightarrow \text{Galactinol + UDP} \\
\text{Galactinol + Sucrose} \rightarrow \text{Raffinose + myo-Ins} \\
\text{Galactinol + Raffinose} \rightarrow \text{Stachyose + myo-Ins}
\]

Less is known about the protective effects of cyclitols in animals, in which internal tissues are generally protected from osmotic or other stresses (but see the next paragraph). However, some eudaphic collembolans (springtails), small insects with permeable cuticles that live mainly in the water-saturated spaces between soil particles, lack this environmental security. Any substantial drying of the soil subjects them to extreme dehydration stress, and they respond with biosynthetic accumulation of remarkable amounts of myo-inositol and other polyols. This dramatically increases the osmotic potential of their body fluids, so that they become capable of rehydrating themselves by adsorbing atmospheric moisture [85–87]. Some beetles that must tolerate subzero temperatures during hibernation accumulate myo-inositol in the haemolymph as an antifreeze [88,89], and hibernating spiders accumulate myo-inositol and scyllo-inositol [90]. In a quite different setting, the concentrations of several
solutes, notably scyllo-inositol, escalate with increasing depth and pressure in some marine animals [56,91].

The renomedullary cells that line the mammalian kidney tubule do regularly encounter such environmental challenges. When the salt concentration in the glomerular filtrate rises, they face an hyperosmotic challenge to which they must adapt if they are to continue to function normally. They do so by increasing their intracellular concentrations of compatible solutes such as sorbitol and inositol [57,92]. Ins accumulation is achieved as a result of increased expression of two Na$^+$/Ins co-transporters. These are the sodium-dependent myo-inositol transporters SMIT1 and SMIT2 [92–97]. The regulatory region of the SMIT1 gene includes an osmotic response element and also responds to prostanoids produced by COX-2 (cyclooxygenase-2) [98,99]. If SMIT-catalysed Ins uptake is inhibited, blunting osmotic adaptation, the outer medullary cells suffer severe injury and acute renal failure follows [100]. Cerebral cortical astrocytes show a similar SMIT-dependent Ins accumulation when they are hyperosmotically stressed [101], and this protein is one candidate target for anti-bipolar drugs [102].

The protective value of the compatible solutes made by mammalian cells is illustrated by the fact that some of them, including myo-inositol, potentiate the capacity of the mutant form of CFTR, the cAMP-regulated Na$^+$ channel that has impaired function in cystic fibrosis, to mature and function correctly in the membranes of cultured epithelial cells [103].

How high concentrations of these cyclitols and other compatible solutes fulfill their presumed protective roles, serving as antifreezes, osmolytes, protection against reactive oxygen metabolites etc. is not clear. A key argument, summarized by Yancey, is that the manner in which they participate in water–solute–macromolecule interactions tends to stabilize proteins and membranes. Binding of destabilizers such as some salt ions and urea to proteins tends to expose groups that undergo thermodynamically favourable binding with the destabilizer, and pushes the proteins towards unfolding. By contrast, many compatible solutes do not bind to proteins and tend to be excluded from the hydration layer adjacent to a protein’s surface. Termed the ‘osmophobic’ effect, this exclusion arises from an apparent repulsion between stabilizers and the peptide backbone. This reduced exposure of the peptide backbone to thermodynamically unfavourable interactions with solutes allows the proteins to fold more compactly [56,104].

**Inositol in membrane phospholipids: sn-2,3-diphénylglycerol-1-phosphorylinositol (archaetidylinositol) and related lipids in Archaea**

Many Archaea contain the Ins-containing membrane glycerophospholipid ArcIns (archaetidylinositol; Figure 1E) and/or a glycosylated derivative thereof. As in all archaial glycerolipids, the membrane-embedded hydrophobic chains of ArcIns (typically $C_{20}$ or $C_{25}$ polyisoprenoid chains) are ether-linked to the 2- and 3-hydroxy groups of glycerol [105–107] (see Figure 1). This sn-2, 3-diradylglycerol-1-phosphoryl (archaetidyl) structure is the mirror image of the configuration of glycerol in the sn-1,2-diacetylglycerol-3-phosphoryl...
(phosphatidyl) glycerolipids of Eukarya and Bacteria (e.g. Figures 1F and 1G). Despite this difference, the Ins headgroups of ArcIns and of PtdIns are linked to the diradylglycerol-phosphate backbones through the same D-1-hydroxy group [108].

G1PDH (glycerol-1-phosphate dehydrogenases), make the sn-glycerol-1-phosphate core of archaeal lipids from DHAP (dihydroxyacetone phosphate). These are mainly found in Archaea, and bacterial G1PDHs again look like consequences of early lateral gene transfer [109]. The archaeal G1PDHs and the G3PDHs (glycerol-3-phosphate dehydrogenases) that convert DHAP into the sn-glycerol-3-phosphate of bacterial and eukaryal glycerolipids have evolved independently as near-basal members of two different dehydrogenase subfamilies [109–114].

In some archaeal phospholipids, the \( \omega \)-termini of the polyisoprenoid chains from the two membrane leaflets are covalently joined to form the even more remarkable caldarchaetidyl lipids [106,115,116]. Each bipolar caldarchaetidyl glycerolipid, of which Figure 1(D) shows an example, spans the membrane bilayer and unsymmetrically displays two headgroups, one on each side of the membrane. Strikingly, the orientation in archaeal membranes of this tetraether lipid with Ins and gentobiose headgroups, and also of the archaetidyl (diether) lipids that correspond to the two halves of this tetraether, mimics that of Ins lipids and glycolipids in eukaryotic plasma membranes: Ins headgroups face the cytosol and sugar headgroups are exposed to the external environment [117].

Some Archaea, particularly extreme halophiles (e.g. \textit{Halobacterium halobium}), lack both Ins3P synthase and Ins-containing phospholipids. This is apparently because bilayer membranes dominated by archaeal lipids with polyol headgroups (e.g. ArcIns and archaetidyglycerol) are unstable at very high salt concentrations. Instead, the membranes of extreme halophiles are dominated by archaetidyglycerol–methylphosphate, which is apparently specialized for bilayer formation in a high salt environment [118].

**Inositol(s) in membrane phospholipids: sn-1,2-diacylglycero-3-phosphoryl-1-inositol (phosphatidylinositol) and related lipids of Eukaryotes and Bacteria**

Inositol’s most studied cellular function is as an ubiquitous headgroup of eukaryotic membrane glycerophospholipids, in which an \( sn-1,2 \)-diacylglycerol-1-phosphoryl (phosphatidyl) core, which is characteristic of most bacterial and eukaryal glycerophospholipids, is linked to the axial 1-\( D \) hydroxy group of Ins (Figure 1F). PtdIns occurs in all Eukaryotes, as do at least some of its seven phosphorylated derivatives (Figures 1F and 1G).

As noted above, PtdIns is present, and occasionally essential, in a few bacteria, notably actinomycetes (\textit{Mycobacterium} spp., \textit{Corynebacterium} spp., \textit{Streptomyces} spp., etc; see below). These complex bacteria use autonomously synthesized Ins to make substantial amounts of PtdIns, and this serves as the precursor for a spectrum of PtdIns polymannosides and of GPI-linked
lipomannans: the latter display extended oligosaccharide structures that contribute to the diverse and characteristic mix of cell surface macromolecules on these cells [119–122] (see also Chapter 13).

Remarkably, it seems that a few bacteria can make membrane-spanning tetraether lipids with a superficial similarity to the caldarchaetidyl lipids of archaeons (see above). However these are based on the bacterial sn-1,2-diradyl configuration [123], and whether these ever include Ins is yet to be determined.

Approx. 2–30% of the phospholipid complement of most eukaryotic cells is PtdIns, which is, as has been shown in the plasma membranes of animal cells, mainly confined to the cytoplasmic leaflets of membrane lipid bilayers. Since early experiments on scyllo-inositol-rich tissues [124] it has generally been assumed that myo-inositol is the only inositol that is incorporated into phospholipids, but a few studies suggest that plant and animal cells may occasionally contain some phosphatidyl-scyllo-inositol or phosphatidyl-chiro-inositol [19,125–129].

**PPIn**

The seven phosphorylated derivatives of PtdIns (phosphoinositides, abbreviated as PPIn) have 1–3 monoester phosphate groups on some combination of the 3-, 4- and 5-hydroxy groups of PtdIns (shaded in Figure 1F). PPIn seem to be present only in eukaryotes. PtdIns4P and PtdIns(4,5)P2 are present in the plasma membranes [3] and several other compartments [130] of all eukaryotic cells [3,131,132]; and PtdIns3P and PtdIns(3,5)P2 are found in intracellular membranes of almost all nucleated eukaryotic cells [133–135]. By contrast, only metazoa and Dictyostelium (and other amoebzoans?) make PtdIns(3,4,5)P3 by the ‘standard’ PPIn3K-I route [136–139]. How many species employ PtdIns5P and PtdIns(3,4)P2, and with what functions, is little understood as yet [140–142] (see Chapter 11).

**PtdIns4P and PtdIns(4,5)P2, ubiquitous players in eukaryotes**

PtdIns(4,5)P2 (Figure 1G) is the substrate for activated PICs and plays several other roles in cells, for example, in exocytotic events and in cytoskeletal homoeostasis.

The ultimate origin of eukaryotic PICs, which have evolved into at least six subfamilies, is something of an enigma. PICδs constitute the only near-ubiquitous PIC subfamily, so they might be structurally and/or functionally most similar to the common ancestor of all extant PICs [143,144]. Despite much effort, however, we still do not understand at all clearly either the functions of PICδs or how they are regulated. What is established is that the Ins(1,4,5)P3 that is liberated when Plc1p, a yeast PICδ-like PIC, is activated does not hang around as a Ca2+-mobilizing second messenger. It is immediately converted into InsP6, and some of it thence to pyrophosphorylated derivatives of InsP6 [145]. Moreover, fungi and plants have no Ins(1,4,5)P3 receptors, the intracellular Ca2+-mobilizing targets of classical Ins(1,4,5)P3 signalling [146].

© 2007 The Biochemical Society
Evolution of the diverse biological roles of inositols

The biological functions of the other PICs, the β, γ, ε, ζ and η families; which are all restricted to metazoans, are much better understood [143] (and see Chapter 3).

Structural comparisons of bacterial PtdIns-specific PLCs and of eukaryotic PtdIns(4,5)P₂-selective PICs have revealed common features in their catalytic mechanisms [147–149]. This suggests evolution from a common ancestor, most likely a bacterial phospholipase, that became incorporated into a primitive ur-eukaryote [132].

**PtdIns3P and PtdIns(3,5)P₂ in membrane trafficking**

PtdIns3P and PtdIns(3,5)P₂, which play central roles in regulating membrane traffic within the endomembrane system of (all?) eukaryote cells, are discussed in detail in recent reviews [135,150,151] and elsewhere in this symposium (see Chapters 8,9 and 12).

**PtdIns(3,4,5)P₃ as a membrane-bound second messenger**

Conversion of PtdIns(4,5)P₂ into the plasma membrane-located messenger PtdIns(3,4,5)P₃ by PPI3K-I, in response to activation of receptor tyrosine kinases or G protein-coupled receptors or to activated Ras, seems to be confined to metazoans and amoebozoans (exemplified by Dictyostelium, see above; see Chapters 5 and 6). The trace of PtdIns(3,4,5)P₃ that is present in yeast seems to be formed by a different pathway and its function is not yet known [152,153].

**Inositol polyphosphates and polyphosphate diphosphates**

Phytate, a mixed cation salt of myo-inositol hexakisphosphate (phytic acid: Ins₆P₆, Figure 1I), was recognized as a major phospho-organic component of plant seeds more than a century ago: the crop seeds and fruit that are harvested each year contain more than 50 million metric tonnes of phytate [154].

Ins₆P₆ is an almost universal constituent of eukaryote cells, from plants and mammals to free-living (Dictyostelium) and parasitic (Entamoeba) amoebae [155,156]. There are multiple biosynthetic pathways to Ins₆P₆, both directly from free Ins [157] and from the Ins(1,4,5)P₃ liberated by PICs [158–161]. Remarkably little is known about the distribution or functions of Ins₆P₆ and related Ins polyphosphates within cells [155,156,162], but recent work has raised the possibility of a variety of unanticipated roles, especially in the nucleus [163–166] (see Chapters 16 and 18).

Not surprisingly, inositol polyphosphates are abundant in the environment, for example in soils and in freshwater and marine sediments. However, we do not know which organisms make these, and substantial and varying proportions are based on the ‘unconventional’ scylo-, d-chiro- and neo-inositol isomers [4,7–11,167–169].
At present, the only organisms known to make inositol polyphosphates other than InsP₆ are the parasitic amoeba *Entamoeba histolytica* and the free-living amoeba *Phreatamoeba balamuthi*, both of which make neo-inositol and scyllo-inositol polyphosphates by undefined routes [170]. Perhaps unwisely, most studies make the assumption that any inositol polyphosphates that are bio-synthetically labelled from [³H]myo-inositol contain myo-inositol rather than other inositol isomer(s); might these ‘unconventional’ inositol polyphosphates be more common than the published information suggests?

One of the most exciting advances of the last decade has been recognition that InsP₆, and at least some isomers of InsP₅ and maybe InsP₄, can undergo the addition of yet more phosphate groups, yielding inositol polyphosphates with one or two pyrophosphate groupings [164,171–175]: see, for example, Figure 1J. These compounds seem likely to be ubiquitous in eukaryotes. They are certainly present in metazoans, amoebozoans and fungi [175]; suggestive HPLC peaks in labelled barley and duckweed [176–178], and genomic prediction of an appropriate inositol phosphate multikinase in rice (*Oryza sativa*) and flax (*Linum usitatissimum*) points to their presence in plants [175]; and Irvine mentions a related putative kinase in *Giardia* [179]. Evolutionarily, these diphosphates must have appeared later than the monoesterified Ins polyphosphates. Their functions are yet to become clear (see [175]).

**Sphingolipids containing inositol**

It was realized long ago that many eukaryotes, notably fungi, plants and various protists, contain sphingolipids that have an InsPCer (Ins1P-ceramide; Figure 1H) moiety at their core [2,180–192]. The best characterized are the InsPCer-mannosides and InsPCer-anchored oligosaccharides that are displayed on the external surface of the plasma membrane in *S. cerevisiae* and other fungi [184,191,193,194]. The Ins-containing plant sphingolipids, typified by Carter’s ‘phytoglycolipid’, are complex and structurally less well characterized [180,188,195,196]. Although animals have generally been assumed to lack InsPCer-based lipids, there are recent reports of their presence in a feather star (*Comanthus japonica*), a helminth (*Ascaris suum*) and a polychaete (*Tylorrhynchus heterochetus*) [197–201].

Sphingolipids are mainly regarded as molecules of eukaryotes, but they also occur in a few, mainly or exclusively anaerobic, bacteria (*Sphingomonads, Flectobacillus*): there have so far been no reports of inositol-containing bacterial sphingolipids [193].

**Phosphoinositide-anchored proteins and oligosaccharides at cell surfaces**

Cells display molecules on their cell surfaces in a variety of ways, including through attachment to PtdIns-based anchors. The first hint of these structures came when exposure of kidney or bone cells to bacterial PtdIns-PLCs
(PtdIns-specific phospholipase Cs) caused the release of soluble alkaline phosphatase [202]. Later studies showed similar effects in many cells and with many proteins, and established that a substantial proportion of the externally displayed proteins, and oligosaccharides, on eukaryote cells are C-terminally linked (through a complex oligosaccharide and phosphoethanolamine) to a phosphoinositide molecule that is embedded in the external lipid leaflet of the plasma membrane [203–205].

Assembly of these structures, known either as GPI-anchors or GIPL-anchors, is a complex process that occurs in endoplasmic reticulum, Golgi complex and secretory membrane elements. It has been reported that chiro-inositol or scylo-inositol sometimes replaces myo-inositol in the GIPL anchors of some cell surface proteins [129,206], but it is possible that isomerization of myo-inositol to chiro-inositol will make this attribution artefactual [207].

Remarkably, there is substantial overlap between the last two groups of molecules that have been discussed, the inositol sphingolipids and GPI/GIPL-anchors, in that the cores of some of the membrane anchors of GIPL-anchored cell surface proteins and oligosaccharides are based on InsP Cer rather than GroPlns. The InsPCer structures are in the majority in yeasts (see Chapter 17). It seems likely that all GPI anchors are initially made with a diradylGroPIns (sn-1,2-diradylglycerophosphoinositol) core, though with a remarkable variety of different acyl and ether hydrophobic chains in different organisms. But then, at least in non-mammalian eukaryotes (including plants, fungi, trypanosomes and Dictyostelium), the phosphoinositol plus its linked glycans and/or proteins are transferred as a unit from the original diradylglycerol to ceramide, yielding an InsP Cer-anchored structure [194,208–215]. Confusingly, these InsP Cer-anchored and membrane-embedded molecules are often still referred to as GPI-anchored.

Chiro-inositol, pinitol and diabetes

Some years back, several groups observed that insulin-treated cells liberate mediator molecules that mimic, to varying degrees, the effects of insulin on cells. No definitive structures have been reported, but at least one of the active molecules seems to include a chiro-inositol glycoside structure, possibly derived from a cell surface glycophospholipid or from the GIPL anchors of cell surface proteins [216–221]. These observations were followed by indications of substantial changes in the metabolic handling of chiro-inositol in diabetic people and animals [20,222–227]. Most recently, several studies have indicated that dietary intake of myo-inositol, chiro-inositol or pinitol has therapeutic potential as a treatment for improving insulin-mediated glycaemic control, particularly in Type 1 diabetes [221,228–234], and the same molecules have beneficial effects on endothelial function in diabetics [235]. These observations have been accruing in parallel with recent major discoveries on the phosphoinositide 3-kinase/PtdIns(3,4,5)P₃-mediated actions of insulin, so they have received relatively little attention. They seem to offer the potential to illuminate one important aspect of how animals employ chiro-inositol.
Actinobacterial metabolites: mycothiol, antibiotics etc.

Many actinomycetes, including pathogenic and other mycobacteria, use the inositol containing sulfydryl peptide mycothiol [1-D-myoo-inositolyl-3-(N-acetyllysisteinyl)amido-α-D-glucopyranoside; Figure 1K] as their major intracellular redox buffer [236–239]. Moreover, mycobacterial pathogens appear to use mycothiol both as a protectant against electrophilic compounds and for the detoxification of reactive oxygen and nitrogen species. For example, Mycobacterium tuberculosis uses mycothiol to conjugate and detoxify antimycobacterial drugs (e.g. rifamycin S) in a manner similar to the formation of mercapturic acid derivatives in glutathione-based detoxification mechanisms. Mycothiol synthesis has therefore become a target for anti-mycobacterial drug development. It has also long been known that actinomycetes synthesize streptidine (a guanido derivative of scylo-inositol) and other myo-inositol-derived cyclitol derivatives as precursors of aminoglycoside antibiotics such as streptomycin [240–242].

Other complex biological products containing inositol

Another complex biological product containing inositol is axinelloside A, a telomerase inhibitor, isolated from a marine sponge: this is a complex, highly sulfated, lipopolysaccharide containing scylo-inositol lipopolysaccharide (Figure 1L) [243].

Can we now start to reconstruct early stages of the complex evolutionary progression that has led to the current multiplicity of functions for inositol-containing molecules in metazoans and other eukaryotes?

There have been recent attempts to reconstruct the evolutionary histories of kinase and phosphatase families involved in PPIn and inositol polyphosphate metabolism and function [e.g. 179], but little attention has yet been paid to the deeper history of the biology of inositols, particularly the transition from primitive Archaea and Bacteria to the ancestors of extant Eukarya.

Once Ins3P synthase had evolved in an early archaeon, the next major step will have been the incorporation of D-1-linked Ins into ArcIns by a still unidentified CMP-ArcOH (CMP-archaetidate) inositol transferase. This enzyme will likely be a member of the CMP-ArcOH alcohol transferase family, candidates for which have recently been identified in several ArcIns-containing archaeons [27], but we still await functional studies.

The passage of this type of biosynthetic pathway into eukaryotes, ultimately to make PtdIns rather than ArcIns, is likely to have required a series of developments. First, the synthesis of phosphatidate (sn-1,2-diacylglycerophosphate; PtdOH) by G3PDH and glycerophosphate acyltransferases, is likely to be achieved by enzymes with bacterial ancestry. Secondly, formation of CMP-PtdOH will have needed a cytosolic-type synthase, which may have evolved...
either from a bacterial progenitor or from an archaeal CMP-ArcOH synthase. And finally, transfer of the resulting phosphatidyl grouping to the D-1 hydroxy group of inositol may be catalysed by a CMP-PtdOH inositol transferase that is a descended from an archaeal CMP-ArcOH inositol transferase? When the archaeal CMP ArcOH synthetase and CMP-ArcOH inositol transferase just mentioned are biochemically examined, it will be particularly interesting to learn whether they exhibit any, even slight, activity against sn-2,3-diradylglycerophosphate precursors, particularly of diacyl types.

The difficult problem of how the archaeal and bacterial progenitors of eukaryotes brought together and blended the activities needed for synthesis of two stereochemically incompatible sets of lipids, and settled on the sn-1,2-diradyl configuration while retaining archaeal constituents such as inositol, remains to be solved. Various models are nicely discussed in a recent review [109].

Bacterial PtdIns-specific PLC and eukaryotic PICs, which are substrate-promiscuous but PtdIns(4,5)\(P_2\)-selective, have probably evolved from a common bacterial ancestor (see above). Hydrolysis by both proceeds either directly to Ins(1:2-cyclic)\(P\) [244] or to Ins1\(P\) via Ins(1:2-cyclic)\(P\) [or a phosphorylated Ins(1:2-cyclic)\(P\) derivative] as an intermediate [245–250]. A key difference lies in the preference by eukaryotic PICs for the phosphorylated forms of PtdIns. This property is, so far as we know, characteristic only of eukaryotic PICs. Intriguingly, however, a PIC from trypanosomes not only hydrolysates not only hydrolysates both PtdIns(4,5)\(P_2\) and PtdIns but also efficiently hydrolysates Ins\(P\)Cer \textit{in vitro} [251].

When the first PPIn were made, and in what organisms, are amongst the next key questions: PtdIns4\(P\) synthesis, my PhD thesis topic [252], is likely to have been an early eukaryotic innovation. And when did PLC first become capable of hydrolysing PPIn possessing a 4-phosphate group? Hints might come from carefully targetted examinations of extant phospholipases C and PtdIns 4-kinases.

\textbf{Note added in proof (received 2 October 2006)}

\textit{Porteresia coarctata} is a salt-resistant wild relative of rice, which has an Ins3\(P\) synthase that remains fully active even in 0.5 M NaCl. In a remarkable recent paper [75], the gene encoding this enzyme was introgressed into \textit{E. coli} (which normally contains no inositol), into \textit{S. pombe} (which requires inositol but cannot make its own), into a standard salt-sensitive rice cultivar and into Indian mustard plants (\textit{Brassica juncea}). This genetic modification had the same effects on all four organisms. Each made substantial amounts of intracellular myo-inositol even when salt-stressed, and each became able to survive and to grow in the presence of normally toxic amounts of salt. As the authors pointed out, an ability to make crop plants salt-tolerant simply by permitting them to sustain myo-inositol synthesis from glucose 6-phosphate in the face of salt stress is a very exciting prospect.

I am grateful to all who have, since the early 1960s, helped me to learn about the functions of inositol derivatives. In particular: my early mentors and supporters (Tim Hawthorne, Hal Coore, the late Georg Hübscher and Manfred Karnovsky); those
with whom I have shared the running of my research group through various periods (the late Roger Coleman and Bryan Finean, and then Chris Kirk and Steve Dove); and those who joined me in the hunt for the functions of inositol lipids in the early days when this was still unfashionable [Eduardo Lapetina, David Allan, Lynne Jones, Shamshad Cockcroft (formerly Jafferji), Martin Low, Motassim Billah, Peter Downes, Mike Hanley, Phill Hawkins and Len Stephens]. The MRC (Medical Research Council) has supported my research almost continuously since 1972, and to have held a Royal Society Professorship for almost two decades has been a great privilege.

References

© 2007 The Biochemical Society
Evolution of the diverse biological roles of inositols 241

© 2007 The Biochemical Society

© 2007 The Biochemical Society
Evolution of the diverse biological roles of inositols


© 2007 The Biochemical Society

© 2007 The Biochemical Society
Evolution of the diverse biological roles of inositols 245


© 2007 The Biochemical Society