Pier Luigi Luisi
Pasquale Stano *Editors*

The Minimal Cell

The Biophysics of Cell Compartment and the Origin of Cell Functionality

Springer

The Minimal Cell

Pier Luigi Luisi • Pasquale Stano
Editors

The Minimal Cell

The Biophysics of Cell Compartment
and the Origin of Cell Functionality

Editors
Pier Luigi Luisi
Biology Department
University of Roma Tre
Rome
Italy
luisi@mat.ethz.ch

Pasquale Stano
Biology Department
University of Roma Tre
Rome
Italy
stano@uniroma3.it

Preface

The main reason behind this book was the necessity to give a proper account to the considerable interest that arose in the last ten years on the notion of minimal cell. This is broadly defined as the cell containing the minimal and sufficient number of components to be defined as alive, or at least capable of displaying some of the fundamental functions of a living cell. Obviously such a definition encompasses entire families of semi-synthetic cells, where the term semi-synthetic indicates that most of the basic components (DNA, enzymes) are extant natural components, whereas the compartments are usually ad hoc constructed vesicles (generally liposomes, i.e. vesicles made out of lipids). Today this research program is seen as part of the broader scenario of synthetic biology, and in particular of "chemical synthetic biology" (Luisi 2007), as indeed the operational procedures attending their production in the laboratory are exquisitely chemical (and not so much genetic manipulations as most of the engineering synthetic biology).

The story of the minimal cell on the basis of liposomes started in the early 1990s mostly in my laboratory at the ETH Zürich, where we set up methods to perform complex molecular biology reactions inside liposomes, for example the polymerase chain reaction (Oberholzer et al. 1995a), or the incorporation of the entire ribosome machinery inside liposomes with the production of the first polypeptide chain (Oberholzer et al. 1999). I believe the term "minimal cell", related to the synthetic biology using liposomes, appeared in that 1995 paper with Oberholzer (Oberholzer et al. 1995b). At that time we were practically the only ones in the field, but at the turn of the century several valid groups begun to work with liposomal minimal cells, among which Yomo (Yu et al. 2001; Ishikawa et al. 2004), Noireaux and Libchaber (2004), Yoshikawa and coworkers (Nomura et al. 2003; Tsumoto et al. 2001); and well known people like David Deamer dealt with the subject (Chakrabarti et al. 1994, Pohorille and Deamer 2002). When I moved to Rome in 2003, I took with me the research projects in chemical synthetic biology, and had the good fortune to be accompanied by Cristiano Chiarabelli and Pasquale Stano, whereby the latter has been supervising in Rome all the work with liposomes, including the minimal cell work.

While progress in this field was progressing, it became clear to several of us that a number of basic concepts involving the physical chemistry of cells were necessary

to advance the work with the minimal cell/minimal life, concepts which were by no means clarified yet in the literature. The entrapment of components into a liposome was in fact re-proposing the general problem of the local concentration in biological compartments, with keywords such as molecular crowding, the state of water in a restricted volume, the anomalous internal pH, and the general question whether and to what extent the thermodynamic and kinetic patterns of the bulk solution were still valid in a small, crowded environment.

These fundamental questions relating to the physical and chemical state of components inside a living cell, and to their behaviour, form the second and complementary reason for the production of this book.

Thus, this book is formally divided into three related and partly overlapping parts. It starts with a discussion on the physical aspects, beginning with the question of the physical state of the cytoplasm (Keighron and Keating), the mechanical properties of the cell (Boal and Forde), the internal cellular crowding (Acerenza and Graña). The next question is another classic one: what are the minimal dimensions of a cell (Moore), a question which has been recently addressed in the literature with the minimal size of functional liposomes (Souza et al. 2009). The conclusive chapter of the first part deals with a discussion of confinement and crowding (Minton and Rivas).

After this physico-chemical prelude, the book moves towards biological functionality, and we will find the contribution of Lancaster and Adams where the question of the minimal physical size is seen also in terms of the minimal genome size. The two minimal sizes are related by another classic question: how can the long, rigid duplex of DNA be entrapped into a small compartment having dimensions which are an order of magnitude smaller. This question has been investigated in the literature also with reverse micelles, obtaining evidence for the super-condensed "psi" form of DNA (Osfouri et al. 2005).

And how does water present in the crowded environment play in all that? The contribution by Pollack and collaborators addresses this point.

And than the book moves towards more and more biological aspects, with the contribution by Monnard and Deamer , who tackle the question of the assembly of bio-membranes, and Gardner and Davies who address the constitution of cells, still from a chemical point of view.

And now, with the third part of the book, we are in business with the minimal cell. We start with an analysis of Yoshikawa, Nomura and collaborators on how to construct an in vitro model of a living cellular system, followed by a review article by our group where most of the basic concepts and experimental procedures to make a liposome minimal cell are summarized and critically evaluated. From here we move towards the most advanced operational procedure of Kumura to produce two different membrane proteins inside liposomes. With the contribution of Petra Schwille we find another classic question in the liposome field, namely, what can we do with giant vesicles, those which can reach the dimensions of real biological cells? A modelistic approach to the minimal cell is presented by Fabio Mavelli, in the very interesting case of a RNA minimal cell – no proteins, no DNA. And we

end up with a contribution by Yomo and collaborators, where we are finally taken up to the complexity of biological evolution.

It is clear, I believe, from this rich display that the research on the minimal cell(s) represents an important sector of the life science activity. Without forgetting that this project has a target of historical, almost epochal value: the construction of life in the laboratory.

Rome   Pier Luigi Luisi

References

Chakrabarti AC, Breaker RR, Joyce GF, Deamer DW (1994) Production of RNA by a polymerase protein encapsulated within phospholipid vesicles. J Mol Evol 39:555–559

Ishikawa K, Sato K, Shima Y, Urabe I, Yomo T (2004) Expression of a cascading genetic network within liposomes. FEBS Lett 576, 387–390

Luisi PL (2007) Chemical aspects of synthetic biology. Chem Biodivers 4:603–621

Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. PNAS 101:17669–17674

Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. ChemBioChem 4:1172–1175

Oberholzer T, Albrizio M, Luisi PL (1995a) Polymerase chain reaction in liposomes. Chem Biol 2:677–682

Oberholzer T, Nierhaus KH, Luisi PL (1999) Protein expression in liposomes. Biochem Biophys Res Comm 261:238–241

Oberholzer T, Wick R, Luisi PL, Biebricher CK (1995b) Enzymatic RNA replication in self-reproducing vesicles: an approach to a minimal cell. Biochem Biophys Res Comm 207:250–257

Osfouri S, Stano P, Luisi PL (2005) Condensed DNA in Lipid Microcompartments. J Phys Chem B 109:19929–19935

Pohorille A, Deamer D (2002) Artificial cells: prospects for biotechnology. Trends Biotechnol 20:123–128

Souza T, Stano P, Luisi PL (2009) The minimal size of liposome-based model cells brings about a remarkably enhanced entrapment and protein synthesis. ChemBioChem 10:1056–1063

Tsumoto K, Nomura SM, Nakatani Y, Yoshikawa K (2001) Giant liposome as a biochemical reactor: transcription of DNA and transportation by laser tweezers. Langmuir 17:7225–7228

Yu W, Sato K, Wakabayashi M, Nakatshi T, Ko-Mitamura EP, Shima Y, Urabe I, Yomo T (2001) Synthesis of functional protein in liposome. J Biosci Bioeng 92:590–593

Contents

Part I
The Physical Aspects

Chapter 1
Towards a Minimal Cytoplasm

Jacqueline D. Keighron and Christine D. Keating

Abstract Cytoplasm is common to all living cells and is increasingly recognized for its active role in facilitating and regulating intracellular reactions. Models for the crowded, dynamic, and compartmentalized cytoplasm should therefore be considered for incorporation in minimal cells. This chapter begins with an introduction to biological cytoplasm and then describes several approaches to primitive models of the intracellular environment. Key aspects of the cytoplasm that have been incorporated in model systems thus far include restricted volumes, macromolecular crowding, and colocalization or compartmentation of biomolecular components.

In this chapter, we focus on experimental models for an often overlooked but ubiquitous and critically important cellular component, the cytoplasm. The cytoplasm was once viewed as a relatively homogeneous background solution that contained the organelles and cytoskeletal elements. This view has been challenged by recent work demonstrating the substantial impact that high concentrations of macromolecules, small volumes, and dynamic spatial organization can have on biomacromolecular conformations, interactions, and reactivity. Many aspects of the intracellular environment cannot be reconstituted in simple aqueous buffers, due to lost organization upon dilution and/or homogenization. However, increased understanding of the composition and properties of biological cytoplasm is leading to the generation and analysis of biomimetic cytoplasm-like environments. Such work could help elucidate the physical basis of intracellular organization and its consequences both for living cells and their prebiotic ancestors.

   This chapter begins with a general description of the cytoplasm and its properties. We then introduce the basic components of bulk and microscale model cytoplasms that have been considered to date, with particular attention focused on two classes

J.D. Keighron and C.D. Keating (✉)
Department of Chemistry, Pennsylvania State University, University Park, 16802, PA, USA
e-mail: keating@chem.psu.edu

of model system: (1) enzyme assemblies, which mimic the multienzyme complexes found in cells, and (2) aqueous phase domains, which mimic intracellular microcompartments. Finally, we briefly discuss the possible evolutionary importance of a minimal cytoplasm.

1.1 Cytoplasm

Cytoplasm refers to the cytosol and any organelles or other inclusions. It makes up a significant fraction of the cell volume (Goodsell 1991), has on average 200–300 mg/ml of protein (Zimmerman and Trach 1991), and is the site of protein production/degradation and much of cellular metabolism (Alberts et al. 2002). The highly crowded liquid or gel-like phase of the cytoplasm is referred to as the cytosol or cytoplasmic matrix, which contains high concentrations of macromolecules and is heterogeneous, with spatially varying composition (Walter and Brooks 1995; Alberts et al. 2002; Ellis and Minton 2003). Three important aspects of the intracellular environment – macromolecular crowding, small volume, and compartmentation – are of particular interest. These will be discussed in the following sections.

1.1.1 Macromolecular Crowding

Although cells are approximately 70% water by weight, macromolecules occupy most of the cytoplasmic volume. Consequently, the cytosol is a thermodynamically nonideal solution, in which the volume of solutes cannot be ignored, and their chemical activities cannot be accurately estimated as equal to their concentrations (Minton 2008; Zimmerman and Minton 1993). Figure 1.1a depicts an *E. coli* cell at ×5,000,000 magnification, such that the volume occupied by macromolecules, small molecules, and water in the cell can be visualized. Because no two molecules can occupy the same physical space at a given time, any volume occupied by a macromolecule in the cytoplasm is unavailable for subsequent macromolecules to sample. As illustrated in Fig. 1.1b, this leads to an increase in effective concentration for macromolecules, which cannot access much of the volume due to the presence of other macromolecular solutes. Thus, volume exclusion due to macromolecular crowding can have a substantial effect on the behavior of biomacromolecules, altering reaction rates and equilibria in the cytosol (Ellis 2001; Zimmerman and Trach 1991).

The concentration of background macromolecules within a cell has implications for protein-protein interactions. Multiprotein structures observed and/or postulated to occur in living cells often cannot be recovered from lysed cells, presumably because the weak binding interactions responsible for assembly are lost in dilute solution. As Srere and Mathews pointed out, for a weakly associated enzyme complex with dissociation constant $K_d = 10^{-6}$ M and intracellular concentration 5×10^{-6} M, dilution
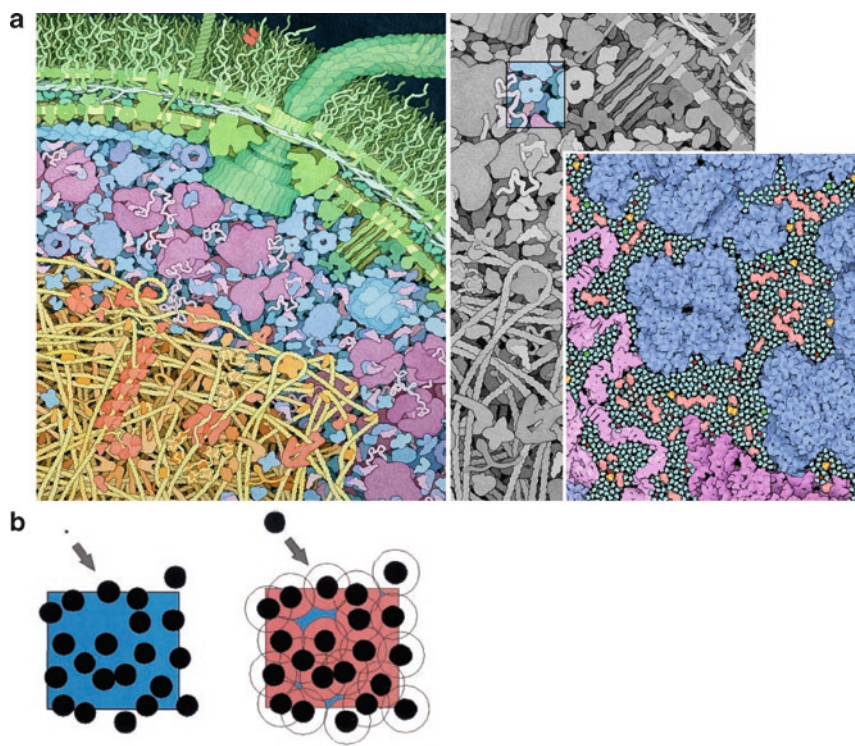
Fig. 1.1 (**a**) Drawings depicting macromolecular crowing in a portion of an *E. coli* cell at increasing magnification from left to right. At the highest magnification individual water molecules are shown (Adapted with permission from Goodsell 2009). (**b**) Volume exclusion impacts larger solutes (*left*) to a greater extent than smaller solutes (*right*). The shaded portion of the solution represents the volume available to a molecule of interest (*black*)

by 5× during cell lysis will reduce the percentage of complexes from 60% to just 10% (Srere and Mathews 1990). Strategies for preventing dissociation due to dilution when cells are lysed include reducing the amount of solution used and adding crowding agents such as polyethyleneglycol (PEG) during lysis. Addition of a nonbiological crowding agent such as PEG has been used to reconstitute enzyme complexes outside of the cell that are thought to exist in the cytoplasm (Datta et al. 1985; Rohwar et al. 1998).

The composition of macromolecules in the cytoplasm can change over time, suggesting that protein structure and/or association state may be altered as the cytosolic environment changes (Batra et al. 2009). Changes in the concentration of macromolecules can occur by osmotic shrinkage or swelling, and can regulate activities within the cell. Increased macromolecular crowding upon exposure to higher external osmolality has been proposed to protect *E. coli* cells by favoring macromolecular

associations, which has several consequences including the stabilization of protein-DNA interactions (Cayley and Record 2004). Mammalian cells also respond to osmotic stress. For example, the rate of proteolysis and protein synthesis in rat liver cells could be correlated with the osmotic pressure. Under hypotonic conditions, water entered the cell, decreasing the concentration of macromolecules, and proteolysis was inhibited. Under hypertonic conditions, protein synthesis was inhibited and proteolysis was stimulated (Haussinger and Lang 1991; Stoll et al. 1992).

While this chapter focuses on the cytoplasm, it is important to note that the effects of macromolecular crowding are not confined to the cytosol. Crowding has also been shown to affect nuclear compartments. When cell nuclei were placed in hypotonic solution where water was absorbed through the membrane, increasing the total volume, compartments such as nucleoli and PML bodies dissociated. Activity of the nucleoli dropped to ~15%. Addition of a crowding agent, PEG, caused these compartments to reform and improved nucleoli activity. When this process was carried out within an agarose gel, which confined the total volume of the nucleus and prevented dilution, no disassembly occurred (Hancock 2004).

1.1.2 Microvolumes

The small volume of biological cells also impacts intracellular reactions. Size varies depending on the type of cell, with the smallest bacterial cells having volumes as low as 1×10^{-17} L and some of the largest oocytes as high as 0.5 μL (Luby-Phelps 2000). Biomolecules such as enzymes can react more favorably due to the increase in collisional frequency in a confined volume (Chiu et al. 1999a, b; Walde and Ichikawa 2001; Chen et al. 2009). It is important to consider not only the concentrations of molecules within the cell, but also the absolute numbers of molecules. Many biomolecules have limited copy numbers in the cell, for example an enzyme present at 10 nM in a bacterial cell may only translate to a single copy in the sub-femtoliter volume of a single cell. In E. coli, the metabolite pyruvate is present at a concentration of 390 μM (Albe et al. 1990); this equates to about 1,400 molecules in the entire cell (cytoplasmic volume of 6×10^{-16} L). Even in the much larger volume of mammalian cells [~2 pL (Short et al. 2005)], a 1 nM concentration of a biomolecule would only equate to ~1,000 copies in the entire cell. With low numbers of copies present, fluctuations can limit the availability of substrates for a reaction to take place (Luby-Phelps 2000). Within the microvolume of the cell, limited substrate availability, the diffusional barrier of the cytoplasm and the possibility of metabolite scavenging increase the benefits of enzyme association (Milani et al. 2003) and favor the formation of compartments.

1.1.3 Compartmentation

The intracellular environment is not only small and crowded with biomacromolecules, it is also spatially compartmentalized. Cytoplasm exhibits several types of organization,

the most obvious of which are the distinct environments provided by the organelles. Additionally, the cytosol can maintain differences in composition at different sites via complex formation, and ion and small molecule gradients are common (Aw 2000). Even single bacterial cells show internal heterogeneity (Fig. 1.1a, left). The nonuniform distribution of molecules within the cell, often termed microcompartmentation, is thought to be important for cellular function, yet in many cases how it is maintained and why it is necessary remain unclear (Walter and Brooks 1995).

In eukaryotic cells, membranes often serve as intracellular boundaries, compartmentalizing various portions of the cell within distinct organelles. Each carries out functions vital to survival. For example, mitochondria and chloroplasts generate ATP for cellular energy requirements, the endoplasmic reticulum translates and folds proteins, and the golgi apparatus modifies proteins after translation. The contents of these compartments are non-homogenous; the nucleus contains nucleoli and subnuclear structures such as the PML bodies and Cajal bodies (Dundr and Mistelli 2001), while the individual cisternae of the golgi apparatus contain different enzymes (Alberts et al. 2002). It should also be noted that the organelles themselves are non-randomly distributed within the cytoplasm at considerable energetic cost to the cell (Luby-Phelps 2000). Mitochondria, for example, will localize near sites of high ATP consumption in order to overcome diffusion barriers in the cell that could reduce the efficiency of ATP utilization (Aw 2000).

Small molecules and metabolites are also non-uniformly distributed. Clusters of mitochondria cause gradients of O_2, ATP, and pH in the cytoplasm. ATP that is generated by the mitochondria is found in higher concentrations around the mitochondrial surface, and O_2 and pH levels are lower here than elsewhere in the cytoplasm (Aw 2000). Pools of metabolites are thought to build up in the vicinity of enzymes in metabolic pathways, which may have a rate enhancing effect on activity since the local substrate concentration is higher around the enzyme (Milani et al. 2003; Aw 2000). For example, in prokaryotes the pooling of dNTPs aids in DNA replication because the K_M of the replication machinery is low; saturation conditions require 0.2–1.0 mM dNTPs, while intracellular concentrations of dNTPs are ~100 µM (Aw 2000).

Components of metabolic pathways come together in a variety of ways to maximize metabolic efficiency (Srere 1987; Islam et al. 2007). Multienzyme complexes have been reported for the Krebs cycle, glycolysis, and nucleic acid production. The enzymes of the Krebs or citric acid cycle provide an example of a static enzyme complex, found in the cytoplasm of prokaryotes and the mitochondrial matrix of eukaryotes. This complex is involved in cellular respiration and regulated by substrate availability, product inhibition, and allosteric inhibitors (Nelson and Cox 2000; Ovadi and Srere 2000). In contrast, the de novo purine biosynthetic pathway is a transient assembly in which the enzymes colocalize when purines are not available to the cell. Here, proximity between the enzymes may play a role in regulating the activity of the pathway (An et al. 2008). To take advantage of the anticipated benefits of colocalization (Conrado et al. 2007), Dueber et al. recently constructed an artificial metabolon within a living cell that enhanced the production of a desired product, mevalonate (Dueber et al. 2009). This was achieved by expressing modified versions of the three biosynthetic enzymes, each with a different

binding domain for sites on a scaffold protein that was also expressed in the cells. The scaffold served to spatially organize the three enzymes within the *E. coli* cell, and to control the stoichiometry of the assembly. This strategy resulted in a 77-fold increase in product generated, suggesting that the desired assemblies did form in the cell, and that the colocalization of these sequential enzymes substantially improved efficiency of the pathway (Dueber et al. 2009).

In addition to organelles, molecular gradients, and metabolic assemblies, still other types of intracellular organization have been observed. For example, in prokaryotes nucleic acids are condensed to form a nucleoid, which is spatially distinct from metabolic components (Goodsell 1991). Recently, researchers have determined that P granules in *C. elegans* embryos are liquid droplets. P granules are thought to play a role in germ cell specification and are uniformly distributed throughout the cytoplasm before the embryo develops polarity. Once symmetry is broken and cell division begins the granules in the anterior of the cell dissolve while granules in the posterior of the cell condense. This is proposed to be driven by concentration gradients of the granule components related to the concentrations of specific proteins in the cell (Brangwynne et al. 2009). The presence of liquid droplets as distinct intracellular domains is consistent with hypotheses by several groups that aqueous phase separation should occur in cytoplasm due to high concentrations of macromolecules (Walter and Brooks 1995; Edmond and Ogston 1968 and Runnstrom 1963). Walters and Brooks proposed that if phase separation did occur, it could lead to microcompartmentation of the cytoplasm, and could explain many observations that have been made in biological cells (Walter and Brooks 1995).

1.2 Experimental Models for the Intracellular Environment

Models for cytoplasm can come in a variety of forms depending on what properties are under consideration. Perhaps the simplest minimal cytoplasms are bulk aqueous solutions that mimic the thermodynamically nonideal solution chemistry of the cell by including high concentrations of background macromolecules. Models for multiprotein complexes have been prepared using protein fusions or assembly scaffolds to colocalize sequential enzymes; this enables study of the kinetic consequences of colocalization outside of the cell. The small volume of the cell interior has been mimicked by conducting reactions inside "cell-sized" vesicles or liquid droplets. Droplets can be easier to prepare with desired internal contents, while vesicles, especially lipid vesicles (liposomes) may be considered more cell-like due to their semipermeable lipid bilayer membranes. Recent work by several groups has sought to add complexity to the interiors of cell-sized lipid vesicles. For example, organization of vesicle interiors has been realized by encapsulation of smaller vesicles, hydrogels, or aqueous two-phase systems. These provide primitive models for microcompartmentation of the cytoplasm. The following subsections introduce experimental models of cytoplasm, and are organized based on whether the model systems are bulk solutions or microvolumes.

1.2.1 Bulk Cytoplasm Models

In vitro experiments can be become more realistic models of the intracellular environment even by very simple changes to standard bulk solutions. For example, enzyme and substrate concentrations can be selected to more accurately reflect those found in the cell or intracellular compartment where the reaction occurs in vivo (Luby-Phelps, 2000; Albe et al. 1990). Since background macromolecule concentrations exert a substantial impact on protein conformation, intramolecular interactions, diffusion, etc., it is important to include them in any model of the cytoplasm. Addition of even simple polymers can provide excluded volume effects, bringing the test tube environment closer to that found intracellularly. More elaborate models can also be adopted, including colocalization of biomolecular components.

1.2.1.1 Polymer Solutions Can Provide Volume Exclusion

Several excellent reviews have appeared which cover experimental and theoretical work on the effect of macromolecular crowding/volume exclusion on biochemical conformations, associations, and reactions (Minton 2007; Minton 2008; Ellis 2001; Zimmerman and Minton 1993). We will introduce this subject briefly here, and refer the reader to these reviews for further information.

Addition of several weight percent of a polymer such as PEG, dextran, or Ficoll can provide a simple model for the volume exclusion effects of macromolecular crowding. The magnitude of these effects is large. This can be seen for example in the effect of volume exclusion on protein association. The association constant for dimerization of a spherical, 40 kDa protein has been estimated to be about an order of magnitude greater in a crowded solution mimicking *E. coli* cytoplasm as compared to dilute solution. The effect becomes several orders of magnitude for tetramer formation (Ellis 2001). Larger molecules drive association more effectively than smaller ones; PEGs with higher molecular weight (MW) have been found to induce fibrillation of α-synuclein faster than lower MW PEGs (Munishkina et al. 2004). Snoussi and Halle have studied protein self-association as a function of increasing dextran volume fraction, and found a 30-fold increase in the formation of decamers at a volume fraction of 14% dextran, which is still lower than intracellular excluded volume (Snoussi and Halle 2005). The identity of crowding agent used can impact results. For example, PEG 3.3 kDa was more effective than Ficoll 70 kDa or larger molecular weight proteins (10× as compared to 3× and 6× over control, respectively) (Kozer and Schriber 2004). Minton and coworkers have recommended that PEG be avoided as a crowding agent due to additional effects observed for this polymer that cannot be quantitatively described by volume exclusion alone. They suggest polysaccharides (dextran, Ficoll) or proteins (BSA) as attractive alternatives to PEG (Zhou et al. 2008).

Because the cytosol has many different macromolecular components rather than a single species, model systems that incorporate multiple crowding agents may more accurately mimic the intracellular environment. Mixtures of different crowding agents Ficoll 70 kDa and dextran 6 kDa have been used in studies of protein folding; researchers found that the mixture of two crowding agents exerted a greater stabilizing force on the enzyme structure than the additive effects of the two molecules by ~0.5 kcal/mol (Batra et al. 2009). The ratio of two crowding agents has been predicted and experimentally observed to impact the effect on protein folding (Zhou 2008; Batra et al. 2009; Du et al. 2006).

1.2.1.2 Enzyme Assemblies Can Provide Colocalization

There are a variety of proposed benefits to the multienzyme complexes described in Section 1.1.3. These include increased local concentration of intermediates, protection of intermediates from scavenging by other pathways, safeguarding labile intermediates, and increased rate of production due to the reduced proximity between enzymes (Milani et al. 2003; Srere 1987). Perhaps most significantly, *changes* in colocalization could provide a mechanism for regulating pathways (Keleti et al. 1989; An et al. 2008). Quantitative assessment of these anticipated advantages is difficult in vivo due to the complexity of the biological systems. Model systems offer a promising route to experimentally investigating the potential consequences of colocalization. Enzymes have been assembled artificially on scaffolds such as metal and latex nanoparticles, or expressed as fusion proteins.

Fusion proteins containing two or more activities. Fusions have been prepared via genetic engineering such that a single polypeptide chain is expressed that contains what would have been two (or more) distinct proteins, with a short linker sequence connecting them. In this way, sequential enzymatic activities can be incorporated in close proximity as a single molecule (Bulow and Mosbach 1991; Conrado et al. 2008). Fusion proteins are attractive models for multienzyme complexes because they provide control over enzyme stoichiometry and fix the relative position of the component enzymes (i.e., either at the C- or N-terminus of adjacent enzymes). They do not require adsorption to a solid support, and thus avoid denaturation concerns common to surface adsorption experiments. Some sequential enzyme pathways, for example the purine de novo biosynthesis pathway, appear to have taken a similar approach to colocalizing some of the activities during evolution: in *E. coli*, each of the ten different steps in the pathway is catalyzed by a separate protein, while in higher organisms several of the steps are found in the same protein, which in terms of sequence homology appears to be essentially a fusion of the proteins that were present in *E. coli* (Zalkin and Dixon 1992).

Fusion proteins that colocalize sequential enzymatic activities can increase the sequential reaction rate and decrease the transient time before a steady state activity is achieved (Ljungcrantz et al. 1989; Lindbladh et al. 1994; Mao et al. 1995). Proximity also enhances the rate of sequential reaction as compared to a competing

reaction (Bulow 1987; Lindbladh et al. 1994; Mao et al. 1995). An example of this can be found with a fusion protein of the glucose phosphotransferase system from *E coli*. Here subunits from four enzymes were expressed together and compared to the activity of an equimolar mixture of isolated subunits (Mao et al. 1995). They found that activity of the fusion protein was 3–4× greater than the isolated subunits, and that a competing pathway did not affect activity of the fusion protein, but did compete effectively with the isolated subunits. Multiple copies of each subunit were also incorporated into the same fusion protein. More copies of the early steps in the pathway was found to stimulate the total pathway production, while more copies of enzyme subunits occurring later in the pathway had a negligible or inhibitory effect on activity. Similar experiments with isolated subunits had little to no effect on the activity of the pathway (Mao et al. 1995). In the presence of a background macromolecule such as PEG the sequential activity of fusions is enhanced, most likely due to reduced diffusion coefficients for intermediates (Datta et al. 1985).

Fusion proteins are not without challenges. Although gene fusion orients the C and N termini of the proteins, the orientation of the active site cannot be easily controlled, and aggregation can be common for multi-subunit enzymes. The individual enzymes in the fusion can have reduced activity, and fusion can affect the K_M, optimal pH, and stability of the proteins (Bulow and Mosbach 1991). In some cases, it may not be clear whether enhanced activities are due to channeling or just proximity/diffusion (Pettersson and Pettersson 2001; Shatalin et al. 1999). Despite these potential difficulties, fusion proteins have provided a good model system for multienzyme complexes without potential deleterious effects of adsorption to a solid support.

Enzymes attached on or in scaffolds. Intracellular surfaces such as membranes and the cytoskeleton are thought to serve as scaffolds for assembly of some multiprotein complexes (Keleti et al. 1989; Campanella et al. 2005). Adsorption or chemical attachment of enzymes to solid supports such as metallic, polymeric, or silica particles can serve as a simple model system for these types of assemblies (Mosbach 1976). Methods for preparation and analysis of protein-coated scaffold particles are relatively advanced due to longtime use of protein-coated gold nanoparticles as biospecific stains in electron microscopy (Hayat 1989), and more recent bioanalytical and biotechnological application for biomolecule-coated nano- and microparticles (Wang 2006). Unfortunately most of this work has not focused on quantification of enzymatic activity for bound proteins. There are, however, a growing number of studies that have compared activity for bound vs. free enzymes and for colocalized vs. non-colocalized sequential reactions (Beeckmans et al. 1993; Cans et al. 2007; Abad et al. 2005; Srere et al. 1973; Mukai et al. 2009; Pescador et al 2008).

Adsorption to a scaffold can alter the enzymatic activity for several reasons, including denaturation of the protein on the surface, steric hindrance due to the presence of the particle and/or the adjacent adsorbed protein, and/or adsorption orientations that may block the active site. It is therefore important to compare activities for adsorbed enzymes with an equivalent amount of enzyme in free solution. Determination of specific activity requires quantification of the number of

enzyme molecules bound to the particles, which can be challenging for irreversibly bound molecules. This can be done by subtractive methods where the amount bound is determined based on what is left in solution, or by removing the enzymes from the particles (Horisberger and Vauthey 1984; Xie et al. 2003). When adsorption is not reversible, as in the case of proteins directly adsorbed to gold nanoparticles, it may be possible to dissolve the particles to release the bound biomolecules (Cans et al. 2007).

Enzymes can be attached to a scaffold in a single layer using covalent or noncovalent approaches; reversible immobilization is particularly attractive because they enable removal of bound enzyme molecules for quantification and determination of specific activity. For example, HRP has been expressed with a hexahistidine tag, which was then bound to Ni(II)-nitrilotriacetic acid groups on a gold particle. The activity for these HRP bound in a single layer was essentially the same as free in solution (Abad et al. 2005). This general approach has been used to prepare surfaces coated with mixtures of two sequential glycolytic enzymes, hexokinase and glucose-6-phosphate isomerase. Sequential activity of the surface-tethered enzyme pair could then be measured (Fig. 1.2a) to demonstrate rate enhancement due to colocalization (Mukai et al. 2009).

Biological multienzyme assemblies can have protein cores, and in some cases the interior proteins are enzymes responsible for catalysis of one of the reaction steps (Alberts et al. 2002). Model systems for enzyme multilayers can be assembled by direct adsorption. For example, Cans et al. (Fig. 1.2b) quantified HRP enzymatic activity for enzyme adsorbed to the nanoparticle under different conditions. When sufficient enzyme was available in solution, HRP multilayers formed on the particles. Multilayer formation impacted the activity of the adsorbed molecules. As the number of enzymes per particle increased, the specific activity of HRP decreased and the total activity per particle remained relatively constant. This suggests that only the outermost layers of enzyme were active, perhaps due to steric hindrance caused by the surrounding enzyme, or the inability of substrates to diffuse into the complex (Cans et al. 2007). Since biological complexes can have many layers of enzymes, the steric and diffusion concerns raised by the model system may be relevant intracellularly. For example, the pyruvate dehydrogenase complex is a three-enzyme pathway that can be isolated from cells as 50 nm assemblies (Wagenknecht et al. 1991). The first and third enzymes in the pathway are associated with the surface of a core made up to 60 copies of the second enzyme (Fig. 1.2c) (Alberts et al. 2002).

An attractive model system for controlled fabrication of enzyme multilayers is the incorporation of the pathway enzymes during layer-by-layer deposition of polyelectrolytes. These are constructed by sequential adsorption of polycations and polyanions on a latex bead or other support. Enzymes can be incorporated in desired layers during assembly (Caruso and Schuler 2000; Balabushevich et al. 2005). Although most efforts in this area have been directed towards biotechnological or sensor applications, these types of assemblies are also interesting as models for intracellular organization. Pescador et al. recently used this approach to immobilize the enzymes onto the scaffold particle with
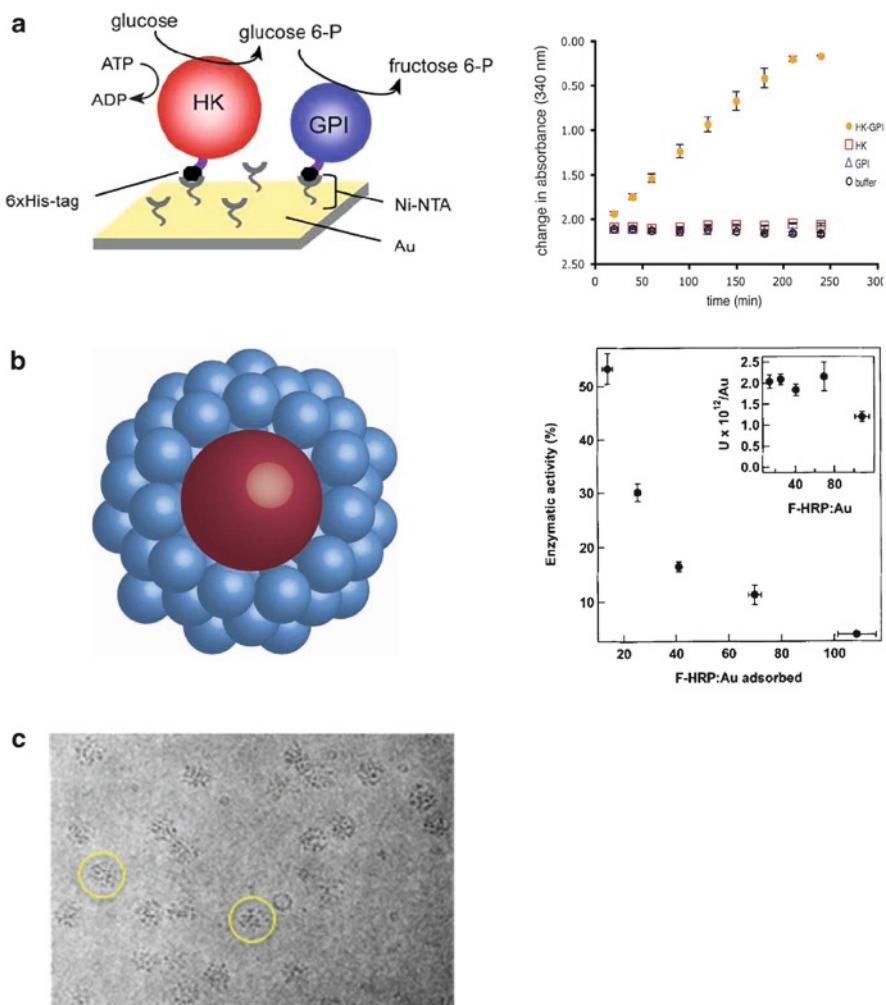
Fig. 1.2 (**a**) Enzymatic activity for two sequential enzymes bound to a planar support (Adapted with permission from Mukai et al. 2009). (**b**) Effect of multilayer formation on HRP enzymatic activity (Adapted with permission from Cans et al. 2007). (**c**) Transmission electron microscope image of mammalian pyruvate dehydrogenase complex. Scale bar = 100 nm (Image adapted with permission from Wagenknecht et al. 1991)

controlled geometry: the two enzymes were assembled either into the same or different layers of the coating. These authors not only confirmed that pathway efficiency was greater for the artificial complex than in dilute solution, they also showed that even within the complex, enzymes work more efficiently when they are closer together and there are fewer barriers for the diffusion of intermediates (Pescador et al. 2008).

The properties of the scaffold itself can be an important factor in these types of experiments (Katchalski et al. 1971). When Koch-Schmidt et al. compared activities for separate and colocalized malate dehydrogenase (MDH) and citrate synthase (CS), with colocalization achieved either as fusion proteins or by adsorption to a sepharose bead scaffold, they found that fusion had no effect on the efficiency of the pathway in dilute solution, and the activity of the immobilized enzymes were 10–20% higher than that of enzymes in dilute solution (Koch-Schmidt et al. 1977). Sepharose beads provided a unique support for the enzyme system in this case because of large pores that enabled diffusion and concentration build up into the bead, and an unstirred layer of solution associated with the surface of the bead. In a related study, Srere and coworkers observed rate enhancements for a three-enzyme system (MDH, CS, and lactate dehydrogenase) immobilized together that depended both on the ratio of the three enzymes and the type of immobilizing matrix (Srere et al. 1973). Systems that include three sequential enzymes generally show an increased advantage to colocalization as compared with two-enzyme systems (Mattiasson and Mosbach 1971; Srere et al. 1973).

1.2.2 "Cell-Sized" Volumes

A variety of approaches have appeared for conducting reactions inside lipid vesicles and other microscale containers, including surface-based strategies that can monitor many liposome reactors simultaneously (Christensen and Stamou 2007), and liposome-nanotube networks that offer sophisticated control over reactant mixing (Lizana et al. 2009). Several enzymatic processes have been carried out within vesicles, including protein/nucleic acid expression, PCR, production of lipid used for vesicle growth, and polymerization of polypeptides and sugars (Nasseau et al. 2001; Apel et al. 2002; Murtas et al. 2007; Hanczyc et al. 2003; Chen and Szostak 2004; Monnard et al. 2007; Oberholzer et al. 1999a; Oberholzer et al. 1999b; Oberholzer et al. 1995). It is unfortunately not common to directly compare reactivity inside the vesicles with that in bulk solution, partly due to uncertainties in the encapsulated concentrations of reactants.

Depending on the system under study, enzymatic reactions in restricted volumes can be similar to or different from reactions carried out in bulk solution. Chen et al monitored the production of a fluorescent enzymatic product to determine kinetic constants for trypsin in attoliter volumes. They found that the turnover number (k_{cat}) in restricted volume was always higher than that found in free solution, and that the Michaelis-Menton constant (K_M) decreased, leading to a two orders magnitude increase in enzyme efficiency (k_{cat}/K_M) (Chen et al. 2009). This increase was attributed to an increased frequency of collision between enzyme and substrate molecules with each other in a microcompartment (Chiu et al. 1999b; Chen et al. 2009). In contrast, Vogel and coworkers measured enzymatic parameters for alkaline phosphatase in micron-scale liposomes that were quite similar to values for

bulk solution (Bolinger et al. 2008). Even for micron-scale liposomes, volume can be used to regulate reaction rates by varying the volume to dilute or concentrate encapsulated reagents (Lizana et al. 2008).

Several recent studies have incorporated additional aspects of cytoplasm within the "cell-sized" volumes of lipid vesicles, bringing us closer to a minimal cell/minimal cytoplasm combination. These efforts will be described in some detail in the following section.

1.3   Incorporation of Model Cytoplasm into the Minimal Cell

Membranes composed of lipid and protein serve as a selectively permeable barrier between the exterior and interior of the cell (plasma membrane), as well as the boundary between cellular components in higher order cells (nuclear, organelle membranes). The boundary between the inside and outside of a model cell can be composed of a variety of materials, both biological and non-biological, that serve to enclose the interior of the cell from the external environment. Non-biological materials used as microvolume boundaries include polyelectrolyte shells (Liu et al. 2005), di-block copolymer vesicles (Chen et al. 2009), and microemulsions (Pietrini and Luisi 2004). Considerable effort has also focused on vesicles formed from lipid membranes, which can be functionalized with additional membrane components such as modified lipid headgroups, cholesterol, and proteins (Walde and Ichikawa 2001; Luisi and Walde 2000; Luisi et al. 1999). Finally, other amphiphiles such as fatty acids and their glycerol esters have been explored as a means of increasing the permeability of the membrane for improved transport of reactants (Mansy et al. 2008; Mansy 2009). In the following sections, we will focus on the organization of aqueous interiors within any of these boundary membranes.
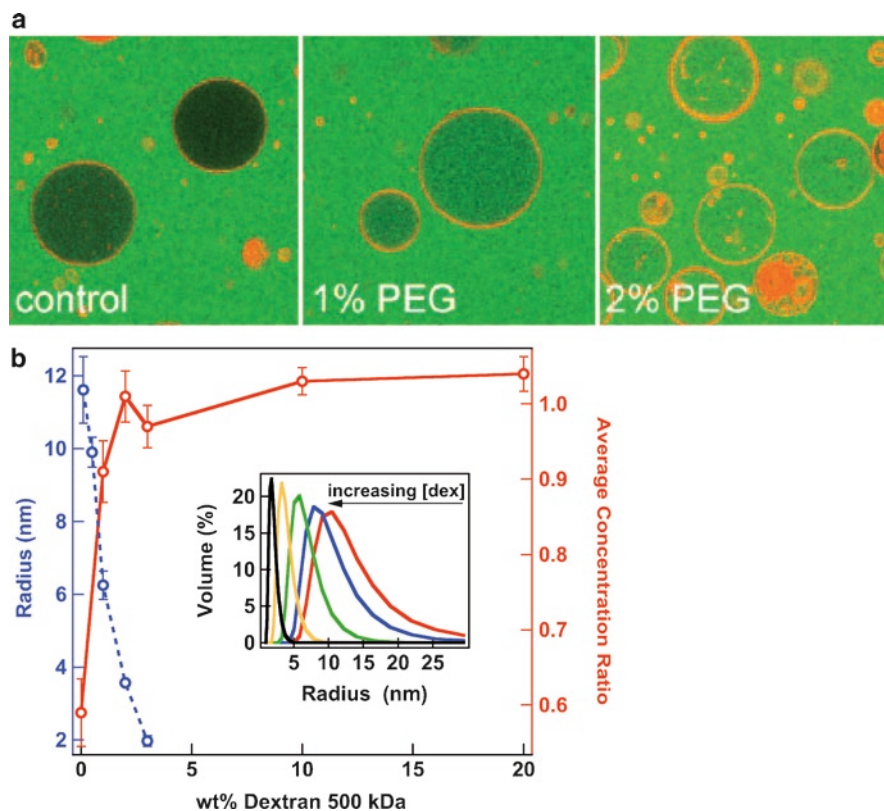
1.3.1   Macromolecules and Macromolecular Crowding in Model Cells

The inclusion of biomacromolecules such as DNA, RNA, and enzymes into a vesicle makes it possible to perform biochemical reactions in restricted volumes and provides a basic model of a protocell (Monnard et al. 1997). To date, a variety of species have been encapsulated within vesicles, including polymers, nucleic acids, enzymes, proteins, small molecules, and gels (Dominak and Keating 2007; Fischer et al. 2002; Oberholzer et al. 1999a, b; Walde and Ichikawa 2001). In contrast to bulk model systems, crowding efforts in microvolumes have thus far focused on the process and effectiveness of encapsulating macromolecules rather than the physico-chemical consequences of their presence. Several factors have been found

to influence the encapsulation of a solute, including its molecular weight, the size of the vesicle, the concentration and charge of both solute and vesicle, and the salt concentration (Monnard et al. 1997). In addition to making small-volume biochemical reactions possible, encapsulating macromolecular species into model cells can provide a more cell-like environment where microcompartmentation, macromolecular crowding, and restricted substrate concentrations can have measurable impacts on the outcome of a reaction.

An important parameter for both model cell and medical applications is the encapsulation efficiency (EE). Two general approaches have been taken to determination of encapsulation efficiency, with the selection of method depending on the experimental system under investigation. For submicroscopic vesicles, it is common to measure the bulk encapsulation efficiency (BEE), which is the concentration of encapsulated species in a population of vesicles after they have been disrupted and their contents released into bulk solution (Szoka and Papahadjopoulos 1980). For example, if a solute were encapsulated in liposomes such that the concentration inside each liposome was equal to the external concentration, then the BEE would depend entirely on the volume fraction encapsulated, which is generally small. BEE can be improved via increasing the lipid concentration, which increases the volume fraction encapsulated, or by solute-specific approaches such as electrostatic attractions between the solute and the vesicles. In contrast, the individual encapsulation efficiency (EE_{ind}), refers to the relative efficiency of solute encapsulation in individual vesicles in a population (Dominak and Keating 2007). EE_{ind} is generally used for giant vesicles over 1 μm in diameter, which can be individually visualized via optical microscopy, interrogated via flow cytometry, or optically trapped prior to lysing (Nishimura et al. 2009; Sun and Chiu 2005). In this case, if a solute was encapsulated such that the internal and external concentrations were equal, the EE_{ind} would be considered 100%. Both methods can provide information on solute encapsulation, and can be used to compare preparation approaches to determine which provides superior encapsulation. For performing reactions inside model cells, the internal concentration information provided by EE_{ind} is of particular interest.

Solute encapsulation, especially for macromolecules above 40 kDa, can be quite heterogeneous between vesicles in the same batch, which could present difficulties in analyzing experiments if not accounted for (Szoka and Paphadjopoulos 1978; Dominak and Keating 2007). Yomo and coworkers have used flow cytometry to compare populations of vesicles prepared in different ways; these authors found that the water-in-oil emulsion method gave superior lamellarity and encapsulated volume (Nishimura et al. 2009). The homogeneity of encapsulation was improved for vesicles prepared by gentle hydration by performing encapsulation under conditions of macromolecular crowding (Dominak and Keating 2008). Figure 1.3a demonstrates how the encapsulation of a fluorecently-labeled 500 kDa dextran molecule is affected by the presence of a background macromolecule. As the concentration of 8 kDa PEG increases, the fluorescence inside the vesicle increases, indicating increased internal concentration of the fluorescently labeled dextran. The mechanism for improved encapsulation was crowding-induced solute condensation. Figure 1.3b

Fig. 1.3 Macromolecular crowding increases polymer encapsulation in giant lipid vesicles. (**a**) Confocal fluorescence shows the location of labeled dextan 500 kDa after formation of giant vesicles by gentle hydration in a solution of this solute. Increasing concentrations of a nonfluorescent polymeric cosolute (PEG) leads to improved encapsulation of the labeled dextran. (**b**) Light scattering shows condensation of dextran as volume exclusion is increased (in this case by increasing the dextran concentration); increased encapsulation efficiency results from crowding-induced condensation of the solute. The solid line represents the average concentration ratio fluorescence inside to outside the vesicle and the dashed line represents the average diameter of the dextran (Adapted with permission from Dominak and Keating 2008)

shows how the size and encapsulation of the dextran is altered by the addition of a crowding species, in this case more dextran. As the solution becomes more crowded (higher wt% dextran), the hydrodynamic radius of the dextran becomes smaller, which leads to an increase in the encapsulation efficiency (Dominak and Keating 2008).

1.3.2   Compartments in Model Cells

As described in Section 1.3, living cells exhibit intracellular organization, with various components found in localized domains and a range of anticipated advantages

that can be difficult to examine in vivo (Milani et al. 2003). Several approaches have been developed to provide compartmentalization of cytomimetic models. These include compartments formed by: (i) internal boundaries, such as small vesicles entrapped in larger vesicles, (ii) hydrogels, such as poly(*N*-isopropylacrylamide), and (iii) aqueous phase separation. Each of these will be discussed below.

1.3.2.1 Compartments Formed by Interior Vesicles

In eukaryotic cells, organelles serve as physically bounded compartments where macromolecules can be sequestered from the surrounding environment. Artificial compartments with similarly distinct boundaries have been formed to subdivide the interior volume of vesicles. For example, small unilamellar lipid vesicles (SUVs, less than 50 nm in diameter) containing a cargo of carboxyfluoroscein were encapsulated within large unilamellar lipid vesicles (LUVs, up to 1 μm in diameter). This is roughly analogous to a vacuole in a eukaryotic cell where a membrane separates the interior volume from the cytoplasm. Lipid compositions for the interior SUVs and the encapsulating larger vesicles were selected to enable SUVs to withstand the formation of the larger vesicle without loss of their contents, and to enable temperature-controlled release of the SUV contents without disruption of the larger vesicles (Bonlinger et al. 2004).

Similar compartments have been formed using layer-by-layer adsorption of polyelectrolytes to form semi-permeable shells around an interior and exterior compartment, containing HRP and glucose oxidase respectively. This was accomplished by using sacrificial scaffold of calcium carbonate to build the shells. The polyelectrolyte shells that served as the physical barrier between the inner and outer compartment, and between the outer compartment and the surrounding environment were permeable to small molecules but not the enzymes themselves. This enabled the substrate, glucose, to diffuse into the vesicle from the external solution. The glucose undergoes an enzymatic reaction in the exterior compartment to produce hydrogen peroxide, which diffuses into the interior compartment and undergoes another enzymatic reaction with Amplex Red to produce a fluorescent product, which can then diffuse out of the interior compartment (Kreft et al. 2007; Tong and Gao 2008). This is analogous to the metabolic processes that occur within the cytoplasmic environment and the mitochondrial environment, where small molecules pass between the two during metabolism.

1.3.2.2 Compartments Formed by Hydrogels

Hydrogels have been encapsulated within giant vesicles to form interior compartments that have an independent shape and structure from the vesicle. In some cases these gels can be compared to the cytoskeleton. Viallet et al. used agarose to examine the relationship between membrane deformation and the viscoelastic properties of the cytoskeleton upon osmotic shrinking and swelling.

They found that the vesicle shape upon shrinking was different from what has been reported for solution filled vesicles, and depended on the amount of agarose encapsulated (Viallet et al. 2004). The cytoskeletal proteins actin and tubulin have also been encapsulated and polymerized within vesicles; this provides a route to increased "intracellular" complexity in model cells, and can help model the cytoskeleton (Merkle et al. 2008; Nomura et al. 2002). Unlike the nonbiological polymers discussed here, however, these do not create microcompartments.

Responsive hydrogels have also been incorporated within lipid membranes. Needham and coworkers prepared a synthetic secretory granule from pH-responsive microgels that were loaded with the fluorescent anticancer drug doxorubicin, then coated with a lipid bilayer. The drug could be released by pH-induced swelling of the microgel (Kiser et al. 1998). The responsive hydrogel PNIPAAm has been used to form compartments within larger lipid vesicles (Campillo et al. 2007; Campillo et al. 2008; Campillo et al. 2009; Jesorka et al. 2005; Markstrom et al. 2007). When encapsulated at low concentrations – less than 0.3 volume fraction of the vesicle – the polymer chains formed individual aggregates. At higher encapsulated concentrations, distinct PMIPAAm microcompartments could be observed within the vesicle (Campillo et al. 2008). The hydrogel compartments were temperature-responsive (Fig. 1.4b), contracting
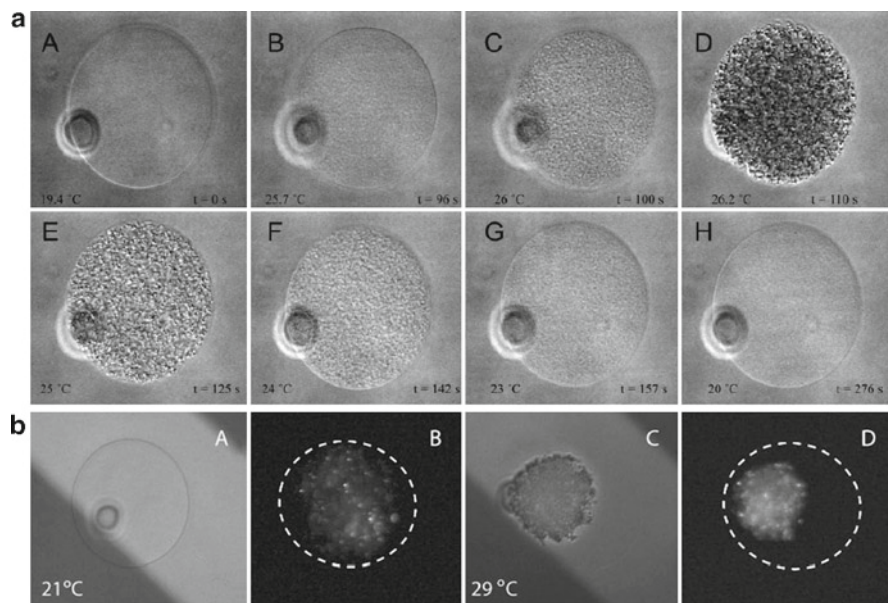
Fig. 1.4 Responsive hydrogel microcompartments in giant lipid vesicles. (**a**) PNIPAAm undergoes reversible phase transitions upon changes in temperature (Adapted with permission from Jesorka et al 2005). (**b**) Fluorescent microparticles are compartmentalized within the hydrogel. Brightfield and fluorescence images are shown before and after heating to shrink the microcompartment, concentrating their contents (Adapted with permission from Markstrom et al 2007)

markedly when heated (Jesorka et al. 2005). This process has been used to reversibly localize latex beads within the hydrogel compartment, as seen in Fig. 1.4b (Markstrom et al. 2007). Multiple distinct compartments were formed by fusing multiple vesicles at elevated temperatures, creating complex compartmentalization, when the temperature was lowered these phases combine to form a single PNIPAAm gel (Markstrom et al. 2007). By crosslinking the polymer to varying degrees the internal compartment of the vesicle can demonstrate viscosity and shear modulus on par with that observed in a living cell. When this polymer is crosslinked to form gel inside the vesicle the membrane shrinks and swells along with the PNIPAAm, changing the size of the membrane bound compartment (Campillo et al. 2007; Campillo et al. 2008; Campillo et al. 2009).

1.3.2.3   Compartments Formed by Aqueous Phase Separation

As described above for bulk solutions, biological or nonbiological polymers can provide volume exclusion. When two or more different polymers are added, it becomes possible to generate distinct coexisting aqueous phases, which can serve as microcompartments. Aqueous phase separation has been postulated to account for many types of intracellular organization (Walter and Brooks 1995), and has recently been observed in living cells (Brangwynne et al. 2009; Ge et al. 2009).

Solutions of two hydrophilic polymers, such as PEG and dextran, form two distinct aqueous phases, each predominantly composed of one type of polymer and water. Phase separation occurs with relatively low polymer concentrations, such that the overall system is predominantly water (Hatti-Kaul 2000). Figure 1.5a shows a phase diagram for PEG 4.6 kDa and dextran 10 kDa in water. Here, the lines represent coexistence curves at different temperatures. As depicted in the inset, when the polymer concentrations fall below the curve the solution exists as one phase, while compositions above the coexistence curve exhibit phase separation. Addition of a third polymer or other solute at a low concentration will result in its partitioning between the two phases; for example, proteins generally accumulate in the dextran-rich phase of a PEG/dextran aqueous two-phase system (ATPS) (Zaslavsky 1995). If a third polymer is added at higher concentration, another phase can form. Multiphase systems containing up to 18 phases have been prepared (Albertsson 1981).

Aqueous two- or three-phase systems are often used for separations of biomolecules or cellular components such as organelles (Rogers and Eiteman 1995). They provide a gentle environment for biomolecules, where pH and ionic strength can be varied as desired. Additionally, the polymers can have a stabilizing influence on structure and biological activity. Figure 1.5b shows a suspension of one polymer phase within the other where a fluorescently labeled protein, α1-antitrypsin, has been partitioned into the dextran-rich phase droplets. Partitioning can be tailored by manipulating the composition of the system, for example by increasing polymer concentration, altering polymer identity, changing pH, or adding an affinity tag to one of the polymers (Hatti-Kaul 2000; Zaslavsky 1995;
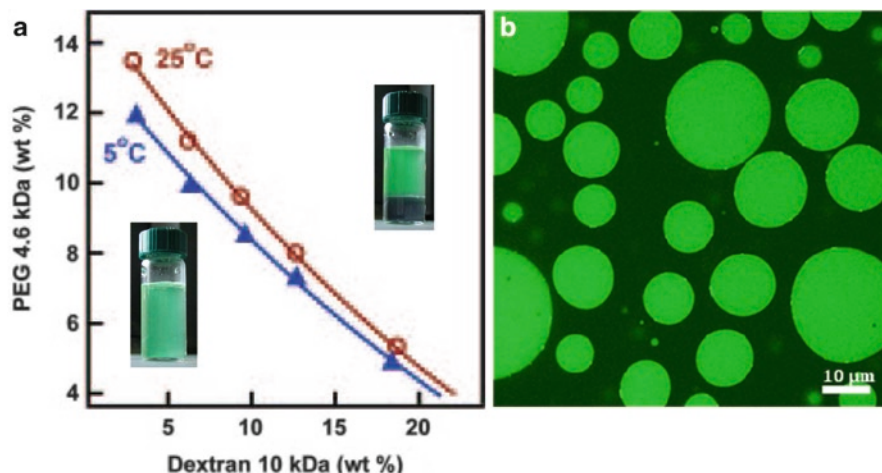
Fig. 1.5 Aqueous two-phase systems. (**a**) Phase diagram for a PEG/dextran aqueous two-phase system (ATPS) showing coexistence curves at two temperatures. The solution exists as a single phase at compositions below the curve, and phase separates into a PEG-rich top phase (labeled with AlexaFluor 488-labeled PEG) and a dextran-rich bottom phase at higher polymer concentrations (Adapted with permission from Long et al. 2005) (**b**) Protein partitioning in an ATPS. Here, droplets of the dextran-rich phase surrounded by the PEG-rich phase were formed by shaking a bulk ATPS. Fluorescence indicates the location of AlexaFluor 488-labeled α1-antitrypsin, which has partitioned strongly into the dextran-rich phase

Albertsson 1981). These effects are enhanced by increasing the size of the solute. Effectively increased size can be achieved by for proteins by conjugating them to a scaffold such as a gold nanoparticle, which has been shown to substantially improve partitioning for several proteins (Long and Keating 2006).

PEG/dextran ATPS have been encapsulated within giant lipid vesicles by forming the vesicles under conditions where the ATPS exists as a single phase (Long et al 2005; Long et al 2008; Li et al. 2008). The resulting vesicles have two internal compartments, corresponding to the PEG-rich and dextran-rich aqueous phases. Optical microscopy images of an ATPS-containing vesicle are shown in Fig. 1.6a. On the left is a transmitted light image with differential interference contrast, and on the right two confocal fluorescence images of the same vesicle. The vesicle membrane has been labeled with fluorescent lipid, and is shown in red (center). Note that there is no membrane present between the PEG and dextran phases, which can be distinguished by the localization of a fluorescently labeled streptavidin, which is partitioned to the PEG-rich aqueous phase (right). Biomolecules such as proteins and nucleic acids have been localized to one of the two coexisting aqueous compartments via simple partitioning into one of the phases (Long et al. 2005). As will be described below, the local distribution of protein within the vesicle interior can be altered by inducing a phase transition, or coupled to patches of distinct lipid membrane phase domains to provide a simple model of cellular polarity (Long et al. 2005; Long et al. 2008; Cans et al. 2008).
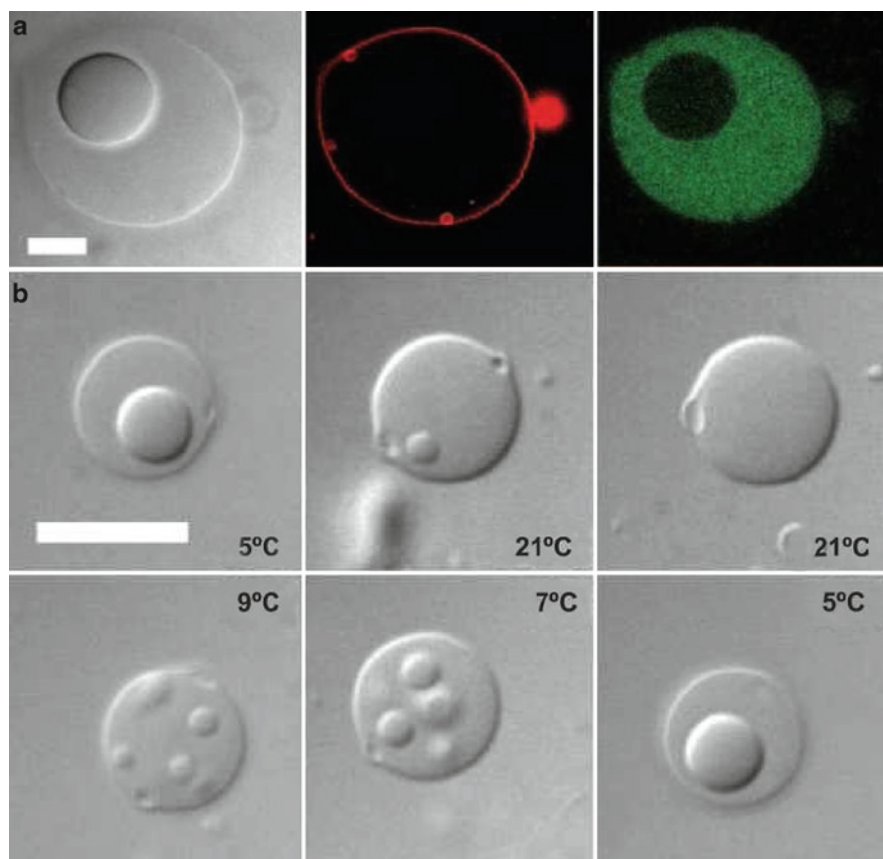
Fig. 1.6 Aqueous two-phase systems encapsulated within giant lipid vesicles. (**a**) Local protein concentration is controlled by partitioning into one of the aqueous phases, which serves as a microcompartment. (*Left*) DIC image, (*center*) fluorescence of rhodamine tagged lipid, (*right*) fluorescence of fluorescein tagged streptavidin localized to the PEG phase. (**b**) Microcompartments due to phase separation can be dissolved and reformed by changing the temperature, scale bars = 10 microns (Adapted with permission from Long et al 2005)

Dynamic protein compartmentation. PEG/dextran solutions encapsulated within lipid vesicles as described above can be converted between a single- and two- phase system with slight changes in temperature or vesicle volume (i.e. polymer concentration). This is possible only for certain polymer concentrations: too little polymer would never phase separate, while too much would not convert to a single phase at elevated temperature. Desirable polymer compositions must be determined by examination of the phase diagram.

Polymer compositions that fall between two coexistence curves at different temperatures (Fig. 1.5a) can phase separate reversibly to form distinct compartments. Proteins that partition into one phase of the ATPS can be dynamically localized to one compartment of the vesicle. As seen in Fig. 1.6b, at low temperatures an

ATPS encapsulated vesicle contains two phases, each of which can serve as microcompartments for proteins or other molecules. When the temperature is increased the polymer composition lies below the relevant coexistence curve and the solution becomes homogeneous, distributing the protein evenly throughout the vesicle. If the temperature is decreased again the polymer composition falls above the relevant coexistence curve and the solution phase separates, forming regions of higher and lower local protein concentration (Helfrich et al. 2002; Long et al. 2005). Local protein concentration differences of tenfold could be maintained indefinitely in these microcompartments (Long et al. 2008).

Changes in osmotic pressure have been used to change the morphology of an ATPS-containing vesicle. Addition of sucrose to the surrounding solution resulted in removal of water, decreasing the vesicle volume because membrane tension was reduced and the concentration of both PEG and dextran polymers was increased, changing the contact angle between the dextran phase and membrane allowing for increased wetting (Li et al. 2008). As the excess membrane area increased, it became possible for the two aqueous phases to change position within the vesicle, minimizing contact between the PEG-rich and dextran-rich phase domains. Figure 1.7 shows optical microscope images of this process, which results in nonspherical vesicles that have one or more "buds" containing one of the aqueous phases attached to a central vesicle body containing the other phase. Addition of
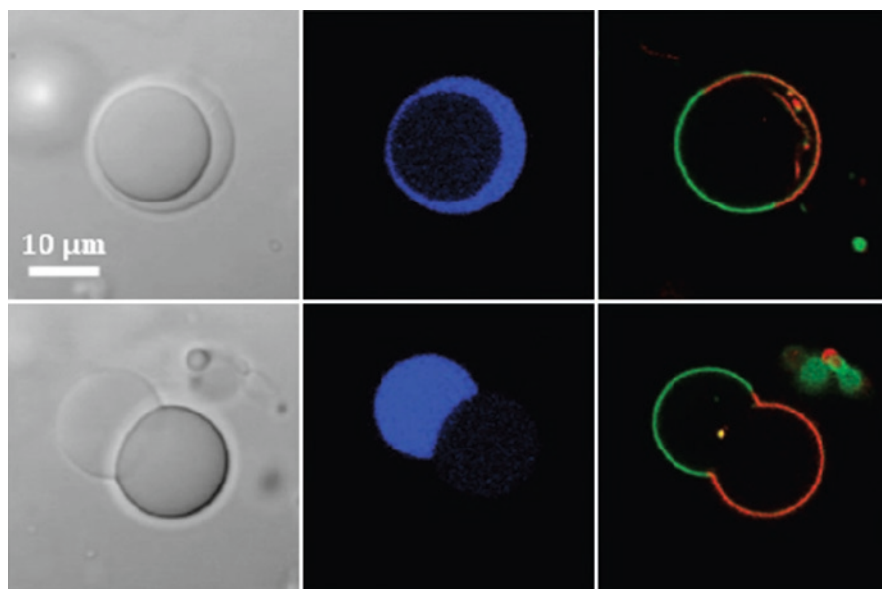
Fig. 1.7 Coupling internal microcompartments with heterogeneity in the lipid membrane that surrounds the vesicle. *Top*: lipid domains are labeled in red (L_d) and green (L_o), and the PEG-rich aqueous phase has been labeled with alexa 633. *Bottom*: osmotic stress-induced budding results in coupling of the PEGylated L_o membrane domain with the PEG-rich interior aqueous domain, leaving the dextran-rich aqueous domain to be coated by the L_d membrane (Adapted with permission from Cans et al 2008)

water to the outside of the vesicle reverses this process as the polymers are diluted (Long et al. 2008). These budded model cells exhibit structural and chemical/biochemical polarity, and protein solutes are nonuniformly distributed to the vesicle bud or body region. Simple experimental models such as this could provide an opportunity to study the effects a polar intracellular environment in a simplified experimental model.

Membrane Interactions. Organization and heterogeneity exist not only in the cytoplasm but also within the lipid membranes of biological cells. Patches of the membrane called domains or rafts are thought to concentrate certain membrane proteins and to enable signaling from the exterior to the interior of the cell (Alberts et al. 2002; Veatch et al. 2007). In model cells the membrane can demonstrate heterogeneity and organization even in the absence of membrane proteins (Veatch and Keller 2005). For example, when two lipid molecules with dissimilar structure are mixed with cholesterol, micron-scale domains are observed in model membranes (Cicuta et al. 2007; Veatch et al. 2007; Collins and Keller 2008). These lipid membrane domains have been coupled to internal aqueous phase domains in giant vesicles that contained a PEG/dextran ATPS cytoplasm mimic (Cans et al. 2008). This was accomplished by incorporating lipids with PEGylated headgroups preferentially into one of two coexisting micron-scale lipid domains. When vesicles containing a PEG/dextran ATPS were subjected to a hypertonic sucrose solution, the resulting budding transition led to vesicles in which the PEG-rich internal aqueous phase was coated by the PEGlyated lipid domain, leaving the dextran-rich aqueous phase coated with the other lipid domain (Cans et al. 2008). An image is shown in Fig. 1.7, where the green represents the PEGylated liquid ordered phase domain (L_o) of the lipid membrane, red indicates the liquid disordered (L_d) membrane phase, and blue shows the location of dye-labeled PEG in the aqueous interior. Such vesicles are a primitive model for coupling of cytoplasmic and membrane microcompartmentation.

1.4 The Role of Cytoplasm in the Evolution of the Cell

The earliest stages of cell formation are hypothesized to have occurred in dilute solutions (Lazcano and Miller 1999). Computational evidence suggests that as the cell evolved the concentration of its encapsulated components could play a crucial role in growth rate and viability (Acerenza and Grana 2006). Non-covalent interactions between macromolecules such as proteins and nucleic acids have been suggested to improve macromolecule stability and activity (Spitzer and Poolman 2009), leading to a more crowded cytoplasm becoming favorable. This increase in concentration would increase the opportunity for macromolecules to interact with each other and form favorable interactions to increase fitness, such as self-organization (Misteli 2001). Self-organized structures like enzyme complexes and the cytoskeleton do not require complex mechanisms to maintain and regulate

their structure, but have the flexibility to change readily when necessary. Self-organization allows the cell to carry out its functions through the most thermodynamically favorable method, while the non-covalent interactions it is based on allows for the flexibility to adjust to outside forces. Even in the earliest precursors to the living cell, simple organization mechanisms such as phase separation or coacervation to form microcompartments may have been key to concentrating organic molecules, facilitating the production and retention of polymers, and enhancing the reaction rate to produce precursor molecules (Oparin 1976; Fox 1976). Experimental models that incorporate minimal cytoplasms can provide valuable insight into both modern and prebiotic cells.

1.5   Conclusions

Just as the plasma membrane was once thought of as protein floating in a sea of lipid, but has more recently been found to harbor many levels of organizational complexity thought to be crucial for normal cell function, our understanding of the cytoplasm has changed much in recent years. There is increasing appreciation for the active role of the cytoplasmic environment and its organization in intracellular processes. It will be important to include minimal cytoplasms in experiments meant to model living cells. From cytoplasm models of the greatest simplicity it has already been possible to incorporate physical phenomena such as macromolecular crowding that are known to exert substantial effects on biochemical reactions. By combining the minimal cytoplasms discussed in this review with other components such as lipid membranes and the transcriptional, translational, and metabolic machinery of the cell, it will be increasingly possible to prepare minimal cells that approach the complexity and functionality of their living counterparts.

References

Abad JM, Mertens SFL, Pita M et al (2005) Functionalization of thioctic acid-capped gold nanoparticles for specific immobilization of histidine-tagged proteins. J Amer Chem Soc 127:5689–5694

Acerenza L, Grana M (2006) On the origins of a crowded cytoplasm. J Mol Evol 63:583–590

Albe KR, Butler MH, Wright BE (1990) Cellular concentrations of enzymes and their substrates. J Theor Biol 143:163–195

Alberts B, Johnson A, Lewis J et al (2002) Molecular biology of the cell, 4th edn. Garland Science, New York

Albertsson P-A (1981) Partition of cell particles and macromolecules, 2nd edn. Wiley, New York

An S, Kumar R, Sheets ED et al (2008) Reversible compartmentalization of de novo purine biosynthetic complexes in living cells. Science 320:103–106

Apel CL, Deamer DW, Mautner MN (2002) Self-assembled vesicles of monocarboxylic acids and alcohols: conditions for stability and for the encapsulation of biopolymers. Biochim Biophys Acta 1559:1–9

Aw TY (2000) Intracellular compartmentation of organelles and gradients of low molecular weight species. Internatl Rev Cytol 192:223–253

Balabushevich NG, Sukhorukov GB, Larionova NI (2005) Polyelectrolyte multilayer microspheres as carriers for bienzyme system: preparation and characterization. Macomol Rapid Commun 26:1168–1172

Batra J, Xu K, Zhou HX (2009) Nonadditive effects of mixed crowding on protein stability. Proteins 77:133–138

Beeckmans S, Van Driessche E, Kanarek L (1993) Immobilized enzymes as tools for the demonstration of metabolon formation. A short overview. J Mol Rec 6:195–204

Bolinger P-Y, Stamou D, Vogel H (2004) Integrated nanoreactor systems: triggering the release and mixing of compounds inside single vesicles. J Amer Chem Soc 126:8594–8595

Bolinger P-Y, Stamou D, Vogel H (2008) An integrated self-assembled nanofluidic system for controlled biological chemistries. Angew Chem Int Ed 47:5544–5549

Brangwynne CP, Eckmann CR, Courson DS et al (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science 324:1729–1732

Bulow L (1987) characterization of an artificial bifunctional enzyme, β-galactosidease/galactokinase, prepared by gene fusion. Eur J Biochem 163:443–448

Bulow L, Mosbach K (1991) Multienzyme systems obtained by gene fusion. Trends Biotechnol 9:226–231

Campanella ME, Chu H, Low P (2005) Assembly and regulation of a glycolytic enzyme complex on the human erythrocyte membrane. Proc Natl Acad Sci U S A 102:2402–2407

Campillo CC, Schoder AP, Marques CM et al (2008) Volume transition in composite poly(NIPAM)-giant unilamellar vesicles. Soft Matter 4:2486–2491

Campillo C, Pepin-Donat B, Viallet A (2007) Responsive viscoelastic giant lipid vesicles filled with a poly(N-isopropylacrylamide) artificial cytoskeleton. Soft Matter 3:1421–1427

Campillo CC, Schroder AP, Marques CM et al (2009) Composite gel-filled giant vesicles: membrane homogeneity and mechanical properties. Mater Sci Eng C 29:393–397

Cans A-S, Andes-Koback M, Keating CD (2008) Positioning lipid membrane domains in giant vesicles by micro-organization of aqueous cytoplasm mimic. J Amer Chem Soc 130:7400–7406

Cans AS, Dean SL, Reyes FE et al (2007) Synthesis and characterization of enzyme-Au bioconjugates: HRP and fluorescein-labeled HRP. Nanobiotechnol 3:12–22

Caruso F, Schuler C (2000) Enzyme multilayers on colloid particles: assembly, stability, and enzymatic activity. Langmuir 16:9595–9603

Cayley S, Record MT Jr (2004) Large changes in cytoplasmic biopolymer concentration with osmolality indicate that macromolecular crowding may regulate protein-DNA interactions and growth rate in osmotically stressed *Escherichia coli* K-12. J Mol Recognit 17:488–492

Chen IA, Szostak JW (2004) A kineic study of the growth of fatty acid vesicles. Biophys J 87:988–998

Chen Q, Schonherr H, Vancso GJ (2009) Small Block-copolymer vesicles as nanoreactors for enzymatic reactions. Small 5:1436–1445

Chiu DT, Wilson CF, Karlsson A et al (1999a) Manipulating the biochemical nanoenvironment around single molecules contained within vesicles. Chem Phys 247:133–139

Chiu DT, Wilson CF, Ryttsen F et al (1999b) Chemical transformations in individual ultrasmall biomimetic containers. Science 283:1892–1895

Christensen SM, Stamou D (2007) Surface-based lipid vesicle reactor systems: fabrication and applications. Soft Matter 3:828–836

Cicuta P, Keller SL, Veatch SL (2007) Diffusion of lipid domains in lipid bilayer membranes. J Phys Chem B 111:3328–3331

Collins MD, Keller SL (2008) Tuning lipid mixtures to induce or suppress domain formation across leaflets of unsupported asymmetric bilayers. Proc Natl Acad Sci 105:1224–1228

Conrado RJ, Varner JD, DeLisa MP (2008) Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. Curr Opin Biotechnol 19:1–8

Conrado RJ, Mansell TJ, Varner JD et al (2007) Stochastic reaction-diffusion simulation of enzyme compartmentalization reveals improved catalytic efficiency of a synthetic metabolic pathway. Metab Eng 9:355–363

Datta A, Merz JM, Spivey HO (1985) Substrate channeling of oxalacetate in solid-state complexes of malate dehydrogense and citrate synthase. J Biol Chem 260:15008–15012

Dominak LM, Keating CD (2007) Polymer encapsulation within giant lipid vesicles. Langmuir 23:7148–7154

Dominak LM, Keating CD (2008) Macromolecular crowding improves polymer encapsulation within giant lipid vesicles. Langmuir 24:13565–13571

Du F, Zhou Z, Mo Z-Y, Shi J-Z et al (2006) Mixed macromolecular crowding accelerates the refolding of rabbit muscle creatine kinase: implications for protein folding in physiological environments. J Mol Biol 364:469–482

Dueber JE, Wu GC, Malmirchegini GR et al (2009) Synthetic protein scaffolds provide modular control over metabolic flux. Nature Biotechnol 27:753–759

Dundr M, Mistelli T (2001) Functional architecture in the cell nucleus. Biochem J 356:297–310

Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26:597–604

Ellis RJ, Minton AP (2003) Join the Crowd. Nature 425:27–28

Fischer A, Oberholzer T, Luisi PL (2000) Giant vesicles as models to study the interactions between membranes and proteins. Biochim Biophys Acta 1467:177–188

Edmond E, Ogston AG (1968) An approach to the study of phase separation in ternary aqueous systems. Biochem J 109:569–576

Fischer A, Franco A, Oberholzer T (2002) Giant vesicles as microreactors for enzymatic mRNA synthesis. ChemBioChem 3:409–417

Fox SW (1976) The evolutionary significance of phase-separated microsystems. Orig Life 7:49–68

Ge X, Conley AJ, Brandle JE et al (2009) In vivo formation of protein based aqueous microcompartmetns. J Amer Chem Soc 131:9094–9099

Goodsell DS (1991) Inside a living cell. Trends Biol Sci 16:203–206

Goodsell DS (2009) Machinery of life, 2nd edn. Springer, New York

Hancock RJ (2004) A role for macromolecular crowding effects in the assembly and function of compartments in the nucleus. Struct Biol 146:281–290

Hanczyc MM, Fujikawa SM, Szostak JW (2003) Experimental models of primitive cellular compartments: encapsultion, growth, and divison. Science 302:618–622

Hatti-Kaul R (2000) Aqueous two-phase systems – methods and protocols. Humana Press, Totowa NJ

Haussinger D, Lang F (1991) Cell-volume in regulation of hepatic-function – a mechanism for metabolic control. Biochim Biophys Acta 1071:331–350

Hayat MA (1989) Colloidal gold principles, methods, and applications, vol 1–3. Academic Press, San Diego

Helfrich MR, Mangeney-Slavin LK, Long MS et al (2002) Aqueous phase separation in giant vesicles. J Amer Chem Soc 124:13374–13375

Horisberger M, Vauthey M (1984) Labeling of colloidal gold with protein – a quantitative study using beta-lactoglobulin. Histochem 80:13–18

Ishikawa K, Sato K, Shima Y et al (2004) Expression of a cascading genetic network within liposomes. FEBS Lett 576:387–390

Islam MM, Wallin R, Wynn RM et al (2007) A novel branched-chain amino acid metabolon - protein-protein interactions in a supramolecular complex. J Biol Chem 282:1189–11903

Jesorka A, Markstrom M, Orwar O (2005) Controlling the internal structure of giant unilamellar vesicles by means of reversible temperature dependent sol-gel transition of internalized poly(N-isopropyl acrylamide). Lanmuir 21:1230–1237

Katchalski E, Silman I, Goldman R (1971) Effect of the microenvironment on the mode of action of immobilized enzymes. In: Nord FF (ed) Advances in Enzymology – and related areas of molecular biology. Wiley, New York

Keleti T, Ovadi J, Batke J (1989) Kinetic and physic-chemical analysis of enzyme complexes and their possible role in the control of metabolism. Prog Biophys molec Biol 53:105–152

Kiser PF, Wilson G, Needham D (1998) A synthetic mimic of the secretory granule for drug delivery. Nature 394:459–462

Koch-Schmidt A-C, Mattiasson B, Mosbach K (1977) Aspects on microenvironmental compartmentation – an evaluation of the influence of restricted diffusion, exclusion effects, and enzyme proximity on the overall efficiency of the sequential two-enzyme system malate dehydrogenase – citrate synthase in its soluble and immobilized form. Eur J Biochem 81:71–78

Kozer N, Schriber G (2004) Effect of crowding on protein-protein association rates: fundamental differences between low and high mass crowding agents. J Mol Biol 336:763–774

Kreft O, Prevot M, Mohwald H et al (2007) Shell-in-shell microcapsules: a novel tool for integrated, spatially confined enzymatic reactions. Angew Chem Int Ed 46:5605–5608

Lazcano A, Miller SL (1999) On the origin of metabolic pathways. J Mol Evol 49:424–431

Li Y, Lipowsky R, Dimova R (2008) Transition from complete to partial wetting within membrane compartments. J Am Chem Soc 130:12252–12253

Lindbladh C, Rault M, Haggland C et al (1994) Preparation and kinetic characterization of a fusion protein of yeast mitochondrial citrate synthase and malate dehydrogenase. Biochem 33:11692–11698

Liu XY, Niu ZW, Xu HF et al (2005) Crosslinkable composite spheres and capsules synthesized by heterocoagluation. Macromol Rapid Commun 28:1002–1007

Liu AP, Fletcher DA (2009) Biology under construction: in vitro reconsistution of cellular function. Nature Rev 10:644–650

Lizana L, Bauer B, Orwar O (2008) Controlling the rates of biochemical reactions and signaling networks by shape and volume changes. Proc Natl Acad Sci USA 105:4099–4104

Lizana L, Konkoli Z, Bauer B et al (2009) Controlling chemistry by geometry in nanoscale systems. Annu Rev Phys Chem 60:449–468

Ljungcrantz P, Carlsson H, Mansson M-O et al (1989) Construction of an artificial bifunctional enzyme, β-galactosidase/galactose dehydrogenase, exhibiting efficient galactose channeling. Biochem 28:8786–8792

Long MS, Keating CD (2006) Nanoparticle conjugation increases protein partitioning in aqueous two-phase systems. Anal Chem 78:379–386

Long MS, Jones CD, Helfrich MR et al (2005) Dynamic microcompartmentation in synthetic cells. Proc Natl Acad Sci 102:5920–5925

Long MS, Cans A-S, Keating CD (2008) Budding and asymmetric protein microcompartmentation in giant vesicles containing two aqueous phases. J Amer Chem Soc 130:756–762

Luby-Phelps K (2000) Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Internatl Rev Cytol 192:189–221

Luisi PL, Walde P, Oberholzer T (1999) Lipid vesicles as possible intermediates in the origin of life. Curr Opin Colloid Surface Sci 4:33–39

Luisi PL, Walde P (2000) Giant Vesicles. Wiley, Chichester, England

Markstrom M, Gunnarsson A, Orwar O et al (2007) Dynamic microcompartmentalization of giant unilamellar vesicles by sol-gel transition and temperature induced shrinking/swellingof poly(N-isopropyl acrylamide). Soft Matter 3:587–595

Mansy SS (2009) Model protocells from single-chain lipids. Internatl J Mol Sci 10:835–843

Mansy SS, Schrum JP, Krishnamurthy M et al (2008) Template-directed synthesis of a genetic polymer in a model protocell. Nature 454:122–125

Mao Q, Schunk T, Gerber B et al (1995) A string of enzymes, purification and characterization of a fusion protein comprising the four subunits of the glucose phosphotransferase system of Escherichia coli. J Biol Chem 31:18295–18300

Mattiasson B, Mosbach K (1971) Studies on a matrix-bound three-enzyme system. Biochim Biophys Acta 235:253–257

Merkle D, Kahya N, Schwille P (2008) Reconstitution and anchoring of cytoskeleton inside giant unilamellar vesicles. ChemBioChem 9:2673–2681

Milani M, Pesce A, Bolognesi M et al (2003) Substrate channeling – molecular basis. Biochem Mol Biol Ed 31:228–233

Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. Annu Rev Biophys 37:375–397

Minton AP (2007) The effective hard particle model provides a simple, robust, and broadly applicable description of nonideal behavior in concentrated solutions of bovine serum albumin and other nonassociating proteins. J Pharm Sci 96:3466–3469

Minton AP (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J Biol Chem 278:10677–10680

Misteli T (2001) The concept of self-organization in cellular architecture. J Cell Biol 155:181–185

Monnard P-A, Oberholzer T, Luisi PL (1997) Entrapment of nucleic acids in liposomes. Biochem Biophys Acta 1329:39–50

Monnard P-A, Luptak A, Deamer DW (2007) Models of primitive cellular life: polymerases and templates in liposomes. Phil Trans Royal Soc B 362:1741–1750

Mosbach K (1976) Immobilized enzymes. FEBS Lett 62(suppl):E80–E95

Mukai C, Bergkvist M, Nelson JL et al (2009) Sequential reactions of surface-tethered glycolytic enzymes. Chem Biol 16:1013–1020

Munishkina LA, Cooper EM, Uversky VN et al (2004) The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. J Mol Recogn 17:456–464

Murtas G, Kuruma Y, Bianchini P et al (2007) Protein synthesis in liposomes with a minimal set of enzymes. Biochem Biophys Res Commun 363:12–17

Nasseau M, Boublik Y, Meier W et al (2001) Substrate-permeable encapsulation of enzymes maintains effective activity, stabilizes against denaturation, and protects against proteolytic degradation. Biotechnol Bioeng 75:615–618

Nelson DL, Cox MM (2000) Lehninger principles of biochemistry. Worth Publishers, New York

Nishimura K, Hosoi T, Sunami T et al (2009) Population analysis of structural properties of giant liposomes by flow cytometry. Langmuir 25:10439–10443

Nomura F, Honda M, Takeda S et al (2002) Morphological and topological transformation of membrane vesicles. J Biol Phys 28:225–235

Oberholzer T, Albrizio M, Luisi PL (1995) Polymerase chain reaction in liposomes. Chem Biol 2:677–682

Oberholzer T, Meyer E, Amato I et al (1999a) Enzyme reactions in liposomes using the detergent-induced liposome loading method. Biochim Biophys Acta 1416:57–68

Oberholzer T, Nierhaus KH, Luisi PL (1999b) Protein expression in liposomes. Biochem Biophys Res Commun 261:238–241

Oparin AI (1976) Evolution of the concepts of the origin of life, 1924–1974. Orig Life 7:3–8

Ovadi J, Srere PA (2000) Macromolecular compartmentation and channeling. Int Rev Cytol 192:255–280

Pescador P, Katakis I, Toca-Herrera JL et al (2008) Efficiency of a bienzyme sequential reaction system immobilized on polyelectrolyte multilayer-coated colloids. Langmuir 24:14108–14114

Pettersson H, Pettersson G (2001) Kinetics of the coupled reaction catalyzed by a fusion protein of β-galactosidase and galactose dehydrogenase. Bichim et Biophys Acta 1549:155–160

Pietrini AV, Luisi PL (2004) Cell-free protein synthesis through solubilisate exchange in water/oil emulsion compartments. ChemBioChem 5:1055–1062

Rohwar JM, Postma PW, Kholodenko BN et al (1998) Implications of macromolecular crowding for signal transduction and metabolic channeling. Proc Natl Acad Sci USA 95:10547–10552

Rogers RD, Eiteman MA (1995) Aqueous biphasic separations biomolecules to metal ions. Plenum Press, New York

Runnstrom J (1963) Sperm-induced protrusions in sea urchin oocytes: a study of phase separation and mixing in living cytoplasm. Dev Biol 7:38–50

Shatalin K, Lebreton S, Rault-Leonardon M et al (1999) Electrostatic channeling of oxaloacetate in a fusion protein of porcine citrate synthase and porcine mitochondrial malate dehydrogenase. Biochem 38:881–889

Short KW, Carpenter S, Freyer JP et al (2005) Raman spectroscopy detects biochemical changes due to proliferation in mammalian cell cultures. Biophys J 88:4274–4288

Snoussi K, Halle B (2005) Protein self-association induced by macromolecular crowding: a quantitative analysis by magnetic relaxation dispersion. Biophys J 88:2855–2866

Spitzer J, Poolman B (2009) The role of biomacromolecular crowding, ionic strength, and physicochemical gradients in the complexities of life's emergence. Microbiol Mol Biol Rev 73:371–388

Srere PA (1987) Complexes of sequential multienzyme enzymes. Annu Rev Biochem 56:89–124

Srere PA, Mathews CK (1990) Purification of multienzyme complexes. Meth Enzymol 182:539–551

Srere PA, Mattiasson B, Mosbah K (1973) Immobilized 3-enzyme system – model for microenvironmental compartmentation in mitochondria. Proc Natl Acad Sci USA 70:2534–2538

Stoll B, Gerok W, Lang F et al (1992) Liver cell volume and protein synthesis. Biochem J 287:217–222

Sun BY, Chiu DT (2005) Determination of the encapsulation efficiency of individual vesicles using single-vesicle photolysis and confocal single-molecule detection. Anal Chem 77:2770–2776

Szoka F, Papahadjopoulos (1978) Procedure for preparation of liposomes with large internal aqueous space and high capture by reverse-phase evaporation. Proc Natl Acad Sci USA 75:4194–4198

Szoka F, Papahadjopoulos (1980) Comparative properties and methods of preparation of lipid vesicles (liposomes). Ann Rev Biophys Bioeng 9:467–508

Tong W, Gao C (2008) Multilayer microcapsules with tailored structures for bio-related applications. J Mater Chem 18:3799–3812

Veatch SL, Keller SL (2005) Seeing spots: complex phase behavior in simple membranes. Biochim Biophys Acta 1746:172–185

Veatch SL, Soubias O, Keller SL et al (2007) Critical fluctuations in domain-forming lipid mixtures. Proc Natl Acad Sci USA 104:17650–17655

Viallet A, Dalous J, Abkarian M (2004) Giant lipid vesicles filled with gel: shape instability induced by osmotic shrinkage. Biophys J 86:2179–2187

Wagenknecht T, Grassucci R, Radke GA (1991) Cryoelectron microscopy of mammalian pyruvate dehydrogenase complex. J Biol Chem 266:24650–24656

Walde P, Ichikawa S (2001) Enzymes inside lipid vesicles: preparation, reactivity and applications. Biomol Eng 18:143–177

Walter H, Brooks DE (1995) Phase separation in cytoplasm, due to macromolecular crowding, is the basis of microcompartmentation. FEBS Let 15231:135–139

Wang P (2006) Nanoscale biocatalyst systems. Curr Opin Biotechnol 17:574–579

Xie H, Tkachenko AG, Glomm WR et al (2003) Critical flocculation concentrations, binding isotherms, and ligand exchange properties of peptide-modified gold nanoparticles studied by UV-visible, fluorescence, and time-correlated single photon counting spectroscopies. Anal Chem 75:5797–5805

Zalkin H, Dixon JE (1992) De novo purine nucledotide biosynthesis. Prog Nucleic Acid Res Mol Biol 42:259–287

Zaslavsky B (1995) Aqueous two-phase partitioning. Marcel Dekker, New York

Zhou H-X (2008) Effect of mixed macromolecular crowding agents on protein folding. Proteins 72:1109–1113

Zhou H-X, Rivas G, Minton AP (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. Annu Rev Biophys 37:375–397

Zimmerman SB, Minton AP (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu Rev Biophys Biomol Struct 22:27–65

Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentration and excluded volume effects for the cytoplasm of Escherichia coli. J Mol Biol 222:599–620

Chapter 2
Evolution of the Cell's Mechanical Design

David Boal and Cameron Forde

Abstract The mechanical properties of the cell's structural components influence the size, shape, and functionality of the cell throughout its division cycle. For example, a combination of the plasma membrane's edge tension and bending resistance sets a lower bound on cell size, while the membrane's tear resistance sets a pressure-dependent upper bound on the size of cells lacking a cell wall. The division cycle of the simplest cells may be dominated by one or two principles such as the maximization of entropy, or the minimization of energy or structural materials. By studying colonies of cells, modern and fossilized, with techniques from classical and statistical mechanics, a partial history can be charted for the appearance and properties of the simplest cell designs.

Keywords Cell mechanics • membrane elasticity • microfossils • cell division cycle

2.1   Introduction

The sizes and shapes of cells have become more diverse with the passage of geological time, built around a core of micron length scales and morphologies such as rods, diplococci and filaments that appeared early in the Earth's history. Yet the incredibly slow change in cell dimensions is suggestive of optimization – that certain structural designs are most appropriate for the complete chemical and physical environment in which cells grow. The factors that must influence the design include ease of construction and repair, appropriate strength and permeability, availability of a mechanically feasible division cycle etc. Further, there may be physical principles at play, such as energy minimization or materials minimization, and the relative importance of each principle may depend on cell morphology or design.

D. Boal (✉) and C. Forde
Department of Physics, Simon Fraser University, V5A 1S6 Burnaby, BC, Canada
e-mail: boal@sfu.ca

The maximization of entropy provides a simple example of how a physical principle can drive the division cycle. If the area of the cell's boundary (its membrane and cell wall, if present) grows as fast as its volume, then at some point the surface will form entropy-producing arms and channels, as shown in the computer simulation of Fig. 2.1, a category of shapes known as branched polymers. If the arms are physically able pinch off to form individual cells, then the entropy-driven formation of branched polymer shapes can be the basis of a very simple cell division cycle. Of course, this design is not at all efficient in the usage of materials: there is a metabolic cost to producing cell boundary material and the boundary area of branched polymer shapes is rather large before cell division is achievable. Among other cell shapes, diplococci, rods, filaments have division cycles that are more materials efficient than branched polymers.

In this chapter, we examine the mechanical properties of the structurally simplest cells. In Section 2.1, the most important structural components are identified and their properties as a function of cell size are analyzed using results from continuum mechanics. For example, we describe the surface stress experienced by a cell under elevated interior pressure and examine the deformation energy of a lipid bilayer. Section 2.2 addresses the question of how the construction of cells has changed in the past three billion years (3 Ga). In this section, we analyze the bending resistance of 2–3 Ga biofilaments using a technique originating in statistical mechanics, and we demonstrate the consistency of design in both ancient and modern filamentous cyanobacteria, as well as estimate some bounds on the mechanical properties of these filamentous cells. In Section 2.3 we examine a number of models for the cell division cycle that focus on the changes in cell shape during growth and division. We show how to extract the time evolution of a system by measuring the instantaneous properties of an ensemble of cells with steady-state growth. This methodology is then applied in Section 2.4 to diplococci in order to study the division cycles of 2 Ga microfossils and modern cyanobacteria. Our conclusions are summarized in Section 2.5.
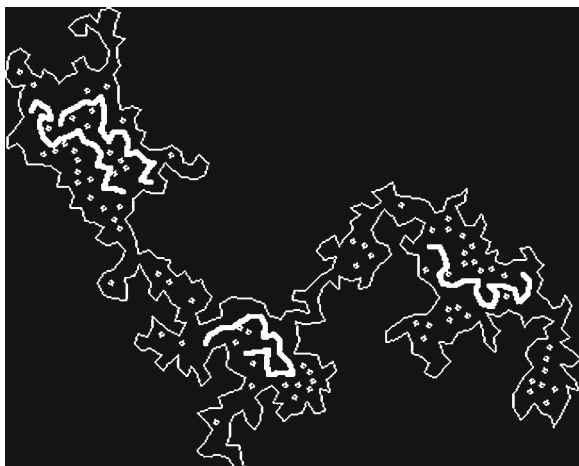
Fig. 2.1 Computer simulation of entropy-driven cell division in two dimensions. Enclosed within this cell are four genetic polymers (*linked spheres*) as well as numerous solvent spheres. Entropy-laden arms only appear when the perimeter of the cell becomes large (Boal and Jun unpublished)

2.2 Mechanical Features of a Simple Cell

Consider as a model system a very simple cell design in which there are only one or two structural components: (1) a fluid membrane that bounds the cell and isolates its components from the environment, and perhaps (2) a shear-resistant wall attached to the fluid membrane. By the word "fluid" we mean a two-dimensional structure that has no shear resistance in its plane: for example, the chocolate coating on a cherry freshly dipped in liquid chocolate is effectively a two-dimensional fluid in that the chocolate can flow to adapt to the shape of the cherry. In contrast, a child's balloon resists shear in the plane of the rubber membrane, even though that membrane can be highly deformed by the pressure of the balloon. The interior of this simple cell may be under elevated osmotic pressure, just like modern bacteria, and the larger the cell is, the more likely the fluid membrane must be augmented by a structure like a cell wall with high tensile strength. We now examine several mechanical properties of fluid membranes.

2.2.1 Bending Resistance of a Membrane

The lipid bilayer that forms the (two-dimensional) fluid boundary of the cell is a self-assembled structure whose equilibrium configuration is spatially flat if the molecular composition is the same within both leaflets of the bilayer. Efforts to bend an initially flat bilayer require the outlay of an energy cost per unit area E whose simplest parametrization is (Helfrich 1973)

$$E = (\kappa / 2)(1 / R_1 + 1 / R_2)^2 + \kappa_{\mathbf{G}} / R_1 R_2, \tag{2.1}$$

where the constants κ (bending rigidity) and κ_{G} (Gaussian curvature modulus) have units of energy. The quantities R_1 and R_2 are the two principal radii of curvature; for example, a sphere of radius R has $R_1 = R_2 = R$, while a cylinder of radius R has $R_1 = R$ around the circumference and $R_2 = \infty$ (i.e. no curvature) along the axis of symmetry. To find the bending energy of a particular surface, one simply integrates E over the area of the surface; hence the deformation energy of a spherical shell of radius R is $8\pi\kappa + 4\pi\kappa_{\mathrm{G}}$, independent of R. Lipid bilayers in conventional cells are found to have $\kappa = 10\text{–}25 \, k_{\mathrm{B}}T$, where k_{B} is Boltzmann's constant and T is the absolute temperature (Evans and Rawicz 1990). The value of κ_{G} is less well known, but expected to have a similar value to κ. Thus, the deformation energy of a bilayer formed into a spherical shell is $12\pi\kappa = 250\text{–}600 \, k_{\mathrm{B}}T$ when $\kappa = \kappa_{\mathrm{G}}$. Although this is not a huge amount of deformation energy, why would lipid bilayers spontaneously deform at all to form a simple cell? To answer this question, we next look at the so-called edge tension of a bilayer.

2.2.2 Edge Tension of a Bilayer

A fluid membrane not only resists bending, it also resists stretching and will rupture once its area has been stretched by more than a few percent from its unstressed value.

The creation of a hole in a membrane likely involves reconfiguring the lipid molecules around the boundary of the hole in order to reduce contact between the aqueous medium surrounding the bilayer and the water-avoiding hydrocarbon chains of the lipid molecules that are normally hidden within the bilayer. In general, the orientation of the lipids at the hole boundary is energetically unfavorable compared to their orientation in an intact bilayer: that is, there is an energy penalty for creating a hole. The boundary of the hole can be characterized by an edge tension λ (energy per unit length along the boundary), such that the energy of the hole is equal to λ times its perimeter. Measured values of λ are in the 10^{-11} J/m range; which is larger than the minimum λ required for membrane stability as estimated from computer simulations of membrane rupture (Boal and Rao 1992).

2.2.3   Minimal Cell Size to Close a Bilayer into a Sphere

The energy of the membrane boundary and the energy of membrane bending have a different dependence on the physical size of the membrane, with the result that a flat membrane must reach a minimum size before it becomes energetically favorable for the membrane to close up into a sphere. In detail, consider a membrane in the shape of a flat disk of radius R_{disk}, perimeter $2\pi R_{disk}$, and consequently, total edge energy $2\pi R_{disk}\lambda$. This shape will be energetically favored over a closed sphere with bending energy $12\pi\kappa$ (when $\kappa = \kappa_G$) so long as $R_{disk} < 6\kappa/\lambda$. Since we are more interested in the dimensions of closed spheres than flat disks, we replace R_{disk} by $2R_{sphere}$ which applies when the disk and sphere have the same area. Thus, the minimum radius of a closed sphere within this description of membrane energetics is $R_{sphere} > 3\kappa/\lambda$ (after Fromhertz 1983). Typical values of $\kappa \sim 15 \, k_BT$ and $\lambda \sim 10^{11}$ J/m lead to the condition $R_{sphere} > 20$ nm, which is somewhat less than the minimal size found for pure bilayer vesicles obtained in laboratory studies. Once the membrane has adopted the shape of a sphere, the configuration could be further stabilized by the addition of lipids to the outer leaflet of the bilayer, thus reducing the strain created by the bending deformation.

2.2.4   Maximal Size for Wall-Less Cells Under Pressure

In a child's balloon or a bicycle tire, the pressure from the confined gas creates a stress within the rubber membrane that forms the boundary of the system. The rubber membrane can be regarded as an effectively two-dimensional system because its thickness is much smaller than its lateral dimensions. Within the plane of the membrane, then, there is a (two-dimensional) surface stress Π having units of energy per unit area. For a spherical shell supporting a pressure difference P across the shell, the surface stress is given by

$$\Pi = PR / 2, \tag{2.2}$$

where R is the radius of the sphere and P has units of energy per unit volume, as usual for a three-dimensional stress. Equation (2.2) tells us that for a fixed pressure difference, the smaller the radius of the sphere, the smaller the surface stress. This is the reason why a simple bacterium can support an osmotic pressure of several atmospheres without needing a cell wall as thick as a tire.

When subjected to a surface stress, a membrane first stretches and then ruptures: depending on their composition, lipid bilayers typically rupture at Π around 1×10^{-2} J/m² on laboratory time scales (Needham and Hochmuth 1989). For a spherical cell of radius $R = 1$ μm and no cell wall, rupture occurs at a fairly low osmotic pressure: Equation (2.2) predicts that the pressure at a failure stress of $\Pi = 10^{-2}$ J/m² would be 2×10^4 J/m³ = 0.2 atm. Thus, a bacterium requires a cell wall to support an osmotic pressure of several atmospheres, which is more than the lipid bilayer of the plasma membrane can withstand. However, such is not the case for smaller cells: the same type of calculation shows that a pure bilayer vesicle of just 100 nm radius (or diameter 0.2 μm) could operate at an osmotic pressure of 2 atm. without needing a cell wall for additional strength (Boal 2002).

2.2.5 Bending and Packaging of DNA

Modern cells carry their genetic blueprint as DNA, which has a contour length that well exceeds the linear dimension of the cell itself. As a double-stranded helix, DNA is considerably stiffer than a simple flexible polymer like a saturated alkane, such that for eukaryotic cells such as our own, the packaging of DNA inside the cell is a challenge. The stiffness of a linear filament is often characterized through its bending rigidity κ_f or its persistence length ξ; we use the latter representation in this section and introduce κ_f in Section 2.3. Mathematically, the persistence length is a measure of the length scale over which the orientation of a curve undergoes a significant change in direction. For molecules whose variation in shape is governed by thermal fluctuations, ξ and κ_f are directly proportional to each other through $\xi = \kappa_f/k_BT$, from which one sees that the stiffer the filament (larger κ_f) the longer the persistence length.

If a filament with a specific value of ξ undergoes random changes in direction all along its contour length L_c, then the root mean square value of the displacement $\mathbf{r}_{ee}$ between the positions of the two ends of the filament is given by

$$< r_{ee}^2 > = 2\xi L_c, \tag{2.3}$$

where the notation $< ... >$ implies than an ensemble average has been made from a selection of all observed configurations. Equation (2.3) tells us that the stiffer the filament (larger ξ) the greater the end-to-end displacement for a fixed contour length. Let's apply this to the DNA of *E. coli*, whose DNA contains 4.7 million base pairs; at 0.34 nm per base pair, this strand of DNA has a contour length of 1.6×10^6 nm. The persistence length of DNA is quoted as about 53 nm (Bustamante et al. 1994), so Eq. (2.3) predicts that the root mean square end-to-end displacement of an open strand of DNA of *E. coli* is given by $< r_{ee}^2 >^{1/2} = 13$ μm, not that much larger than the physical size of the bacterium itself.

A related measure of the size of a flexible filament is its radius of gyration, R_g, which is defined by $<R_g^2> = N^{-1} \sum_{i=1,N} r_i^2$, where the filament has been appropriately sampled at N points with displacements $\mathbf{r}_i$ from the center-of-mass position of the filament. If the physical overlap of remote sections of the filament is permitted, then randomly oriented filaments are governed by $<R_g^2> = <r_{ee}^2>/6 = \xi L_c / 3$. For example, we would expect $<R_g^2>^{1/2} = 5.3\,\mu m$ for the DNA of *E. coli* given the value for $<r_{ee}^2>$ calculated in the *E. coli* example above. A similar calculation for mycoplasma, a very small cell with 800,000 base pairs of DNA, yields $<R_g^2>^{1/2} = 2.2\,\mu m$. For both of these structurally simple cells, the effective size of a ball of their DNA is roughly the same linear dimension as the cell itself. However, this is not the case for eukaryotic cells: human DNA is much longer than bacterial such that it takes up far more volume in the cell as a random coil. Consequently, advanced cells have developed a packaging technique in which their DNA is wrapped around barrel-shaped proteins called histones, with a diameter of 11 nm, in order to organize and sequester their long genetic blueprints.

The examples in this section (the constraints on cell size, the need for cell walls to maintain cell integrity, the packaging of DNA) all illustrate the influence of cell mechanics and construction on the stability and function of the cell.

2.3   Structural Evolution of Filamentous Cells

We now turn our attention to how the design and construction of cells has changed over time, using as our guide a comparison between modern cyanobacteria and microfossils more than two billion years old. The approach is not so much to make statistical comparisons between cell shapes, but rather to determine, where feasible, mechanical characteristics of cells before they were fossilized and chart the evolution of these characteristics. In this section, we focus on the elastic properties of filamentous cells; in Section 2.4 we investigate the cell cycles of diplococci and rod-like cells.

First appearing more than three billion years ago, filament-forming cells have been present throughout much of the Earth's history (Cloud 1965; Barghoorn and Schopf 1966; Walsh and Lowe 1985; Schopf and Packer 1987; Schopf 1993; Rasmussen 2000). Three examples of two-billion-year-old filamentous structures are displayed in Fig. 2.2: parts (a) and (b) are *Gunflintia minuta* Barghoorn and *Gunflintia grandis* Barghoorn, respectively (Barghoorn and Tyler 1965; author's specimens from Lake Superior, Canada) and part (c) is *Halythrix* Schopf (Schopf 1968; specimen from Belcher Islands, Canada, in Hofmann 1976). Even older examples of filamentous structures include 3.23 Ga pyritic replacement filaments (Rasmussen 2000). The original internal construction of these filaments has been destroyed or modified by the fossilization process, but that doesn't mean that their native mechanical properties cannot be probed by other means. Based solely on the analysis of static images, several techniques are available for determining the mechanical behavior of cells, and these approaches are equally adaptable to microfossils as they are to living cells.

The technique that we examine most closely in this section is the tangent correlation length ξ_t that can be related to the resistance of a filament against

Fig. 2.2 Examples of *Gunflintia minuta* (**a**) and *Gunflintia grandis* (**b**) from GSC 10913c (Schreiber, Ontario, Canada; Boal and Ng 2010); *Halythrix* (**c**) from GSC 42769 (Belcher Islands, Canada; reported by Hofmann 1976). Scale bar is 10 μm in all images
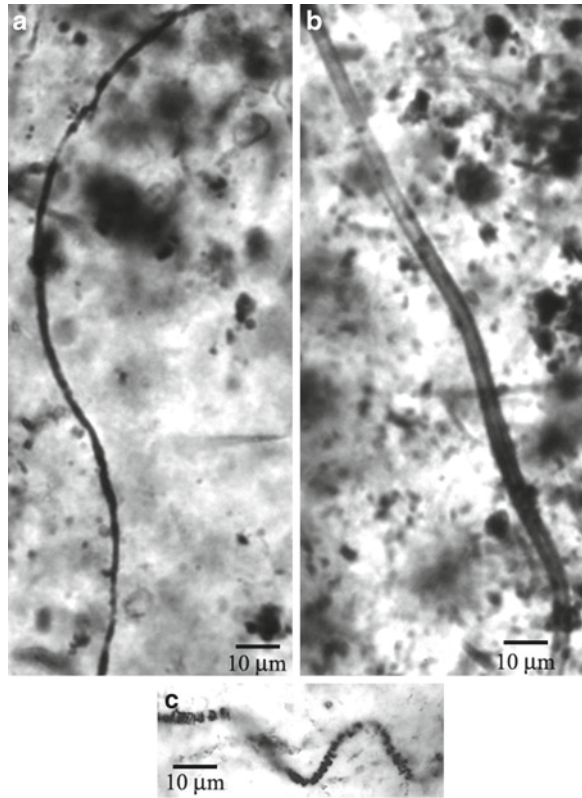
Fig. 2.3 Unit tangent vectors ($\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_n$) at arc lengths (s_1, s_2, s_n) along a sinuous curve. Separations between the locations $\Delta s = |s_2 - s_1|$ are distances, not displacements
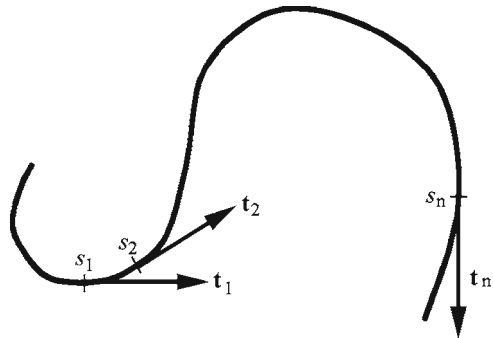
bending. To see how this works, consider the changes in the local orientation of the sinuous curve illustrated in Fig. 2.3 as recorded through the behavior of the tangent vector $\mathbf{t}(s)$ at location s along the curve, where $\mathbf{t}$ has unit length according to the dot product $\mathbf{t} \bullet \mathbf{t} = 1$. If two location s_1 and s_2 are close to each other on the curve, then $\mathbf{t}(s_1)$ and $\mathbf{t}(s_2)$ have similar directions and their dot product is close to unity. On the other hand, if s_1 and s_2 are far apart along the curve (even though they may be close spatially) their tangent vectors may point in quite different directions and $\mathbf{t}(s_1) \bullet \mathbf{t}(s_2)$

may vary between -1 and $+1$. The average behavior of $\mathbf{t}(s_1) \bullet \mathbf{t}(s_2)$ is contained in the tangent correlation function $C_t(\Delta s)$,

$$C_t(\Delta s) \equiv\, <t(s_1)\bullet t(s_2)>.\qquad\qquad(2.4)$$

The ensemble average indicated by the brackets $< \ldots >$ on the right hand side of this equation is performed over all pairs of points s_1 and s_2 subject to the constraint that $|s_2 - s_1|$ is equal to a particular Δs specified on the left hand side. When s_1 and s_2 are nearby ($\Delta s \cong 0$), the ensemble average over $\mathbf{t}(s_1) \bullet \mathbf{t}(s_2)$ is necessarily close to unity, whereas when Δs is so large such that the tangent orientations are random with respect to each other, the average is close to zero. For a variety of very general situations, this behavior of $C_t(\Delta s)$ at small and large Δs is described by exponential decay in Δs (see Doi and Edwards 1986 or Boal 2002):

$$C_t(\Delta s) = \exp(-\Delta s\,/\,\xi_t).\qquad\qquad(2.5)$$

The length scale for the correlations is provided by the tangent correlation length ξ_t: the more sinuous the curve, the smaller is ξ_t. For filaments whose shapes are governed by thermal fluctuations in their deformation energy, the correlation length ξ_t of Eq. (2.5) is the same as the persistence length of Eq. (2.3). As a technical aside, it should be mentioned that correlation function depends on the dimensionality of the system: the true correlation length ξ_3 of a filament in three dimensional space is related to the correlation length ξ_{2p} of the same filament whose shape is projected into two dimensions via $\xi_3 = (3\pi/8)\ \xi_{2p}$.

Now, the magnitude of the deformation of a filament in response to a shear stress is inversely proportional to the filament's stiffness or, properly speaking, its *flexural rigidity* κ_f: the stiffer the filament, the smaller the deformation. To be specific, the deformation energy per unit length for bending a uniform rod is equal to $\kappa_f/2$ multiplied by the square of the rate of change of the tangent direction along the filament contour. It can be established that the flexural rigidity of a uniform solid cylinder of radius R is given by (see Boal 2002)

$$\kappa_f = \pi Y R^4\,/\,4,\quad\text{(solid cylinder)}\qquad\qquad(2.6)$$

where Y is the Young's modulus of the material; typically $Y \sim (1\text{--}5) \times 10^8$ J/m^3 for soft biomaterials. For a hollow cylinder of outer radius R bounded by a wall of thickness t, the flexural rigidity can be approximated by

$$\kappa_f = \pi Y R^3 t,\quad\text{(hollow cylinder)}\qquad\qquad(2.7)$$

when $t << R$. Thus, the flexural rigidity grows as R^3 or R^4 for these two simple shapes. Given that the energetic cost of the deformation is proportional to κ_f, it would not be surprising if the tangent correlation length ξ_t is also proportional to R^3 or R^4, as a benchmark.

Two correlation functions obtained from microfossils are shown in Fig. 2.4 for projected filament trajectories in two dimensions, leading to the determination of the correlation length ξ_{2p}; the figure shows both the raw data as well as their fit with an

exponential function (Boal and Ng 2010). Two taxa from the Gunflint Formation are displayed in the figure: *G. minuta*, the narrower of the two, is found to have $\xi_{2p} = 670 \pm 40$ μm, while the wider *G. grandis* has a shorter ξ_{2p} of 330 ± 30 μm for a particular subset of the *G. grandis* filaments (there may be two populations of filaments in the group that are now collectively assigned as *G. grandis*, the *S* subset has a smaller mean diameter than the genus as a whole, while the *L* subset has a larger mean diameter). In both cases, ξ_{2p} is a remarkable two orders of magnitude larger than the diameter of the filament itself, a ratio of $\xi_t{:}R$ that is common among both microfossils and modern cells. For comparison, Fig. 2.5 shows the tangent correlation lengths of three genera of modern cyanobacterial filaments that represent three very different cell geometries (*Geitlerinema*, *Pseudanabaena*, and *Oscillatoria*, from smallest to largest filament diameter). Both *Geitlerinema* and *Oscillatoria* exhibit values of ξ_2 that rise with filament diameter among species of the genus.

Equations (2.6) and (2.7) demonstrate that κ_f depends most strongly on the filament radius, as R^3 or R^4 for the two idealized systems that we have considered.

Fig. 2.4 Tangent correlation function $C_t(\Delta s)$ as a function of separation Δs (in micrometer) obtained by weighted average for *G. minuta* (*disks*) and the *S* group of *G. grandis* (*diamonds*). The solid curves are the exponential decays predicted by Eq. (2.5) with $\xi_{2p} = 670$ and 330 μm for *G. minuta* and the *S* group of *G. grandis*, respectively (Boal and Ng 2010)
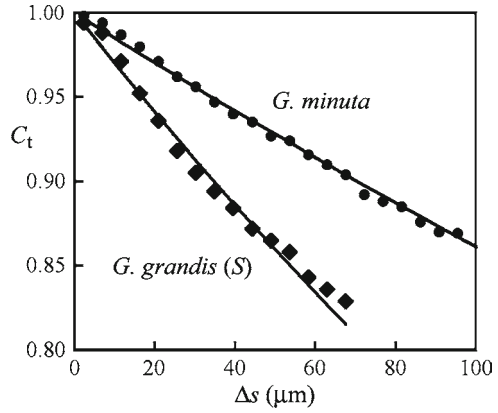
Fig. 2.5 Measured ξ_2 (in micrometer) for filamentous cyanobacteria as a function of their mean diameter *D*. The correlation functions are approximately described by $4.3 \cdot D^{3.3 \pm 1}$ for *Oscillatoria* and $62 \cdot D^{5.1 \pm 1}$ for *Geitlerinema* (*solid and dot-dashed lines*, respectively; both *D* and the result are in micrometer). The cyanobacteria are *Geitlerinema* (*crosses*), *Pseudanabaena* (*square*), and *Oscillatoria* (*disks*) (Boal and Ng 2010)
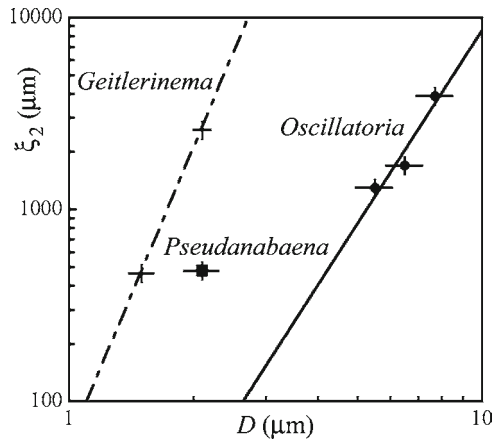
There are structural differences among even the three genera in Fig. 2.5, so the most likely behavior of κ_f is that the species within a given genus obey a particular R^n scaling, but the proportionality constant will vary from one genus to another. For a filament subject only to thermal fluctuations in its deformation energy, the correlation length ξ_t is linearly proportional to the flexural rigidity κ_f. We assume that this proportionality is also valid here, so that the anticipated functional form of the correlation length is $\xi_t = CR^n$ (where the proportionality constant C varies with the genus). Hence, a log-log plot of ξ_t versus R should be a straight line with a slope of 3 or 4 associated with the power-law dependence of κ_f on R.

The two straight lines shown in Fig. 2.5 are the functions

$$\xi_2 = 62 \bullet D^{5.1 \pm 1} \quad (Geitlerinema) \tag{2.8}$$

$$\xi_2 = 4.3 \bullet D^{3.3 \pm 1}, \quad (Oscillatoria) \tag{2.9}$$

where D is the filament diameter quoted in microns, and the result for ξ_2 is also in microns (Boal and Ng 2010). The exponents in these functional forms, 5.1 ± 1 and 3.3 ± 1, are seen to be in good agreement with the expectations from continuum mechanics for the R-dependence of the flexural rigidity. Yet Eqs. (2.8) and (2.9) are obviously not identical, indicating that there is a genus-dependence to the behavior of the correlation length ξ_2. One characteristic that distinguishes among the three genera of Fig. 2.5 and that might contribute to the difference between Eqs. (2.8) and (2.9) is the mean length-to-width ratio of the individual cells: roughly four for *Geitlerinema*, 1.5–2 for *Pseudanabaena* and 0.7 for *Oscillatoria*. Thus, at a given filament diameter, ξ_2 increases with the length-to-width ratio of the cell in Fig. 2.5. Given the very large difference between the cell length-to-width ratios of *Geitlerinema* and *Oscillatoria*, it may be that these two genera lie near two distinct soft limits for the range of tangent correlation lengths available to cellular filaments. That is, with its large length-to-width ratio of individual cells, *Geitlerinema* may represent one limit, while the small length-to-width ratio of *Oscillatoria* represents the opposite limit.

Let's now compare the behavior of modern filamentous cyanobacteria in Fig. 2.5 with the measured correlation lengths of microfossils as displayed in Fig. 2.6. The first observation is that the tangent correlation lengths of the microfossil taxa *Gunflintia* and *Eomycetopsis* are easily in the same range as modern filamentous cyanobacteria. At 700 ± 100 μm, ξ_{2p} of *E. filiformis* is not far from Eq. (2.8) for ξ_2 of *Geitlerinema* represented by the dot-dashed line on the figure. These two types of filaments also have a similar visual appearance as tube-like structures. In addition, *E. filiformis* is not that far removed from $\xi_2 = 480 \pm 50$ of modern *Pseudanabaena* PCC 7403, although *Pseudanabaena* possesses marked indentations at the cell division planes while *E. filiformis* does not. The three variants of *Gunflintia* in Fig. 2.6 have correlation lengths in the 300–700 μm range for populations with apparent diameters of 1–4 μm: *G. minuta* lies near Eq. (2.8) for *Geitlerinema* while the *L* subgroup of *G. grandis* lies near Eq. (2.9) for *Oscillatoria*. All of the *Gunflintia* microfossils lie within the soft boundaries provided by Eqs. (2.8) and (2.9) for the most likely domain of correlation lengths. At less than 50 μm, the very short tangent

Fig. 2.6 Measured ξ_{2p} (in micrometer) for microfossil filaments as a function of their mean diameter. The curves $4.3 \cdot D^{3.3 \pm 1}$ and $62 \cdot D^{5.1 \pm 1}$ (both D and the result are in micrometer) are drawn for reference and also appear in Fig. 2.5. The filaments are *Gunflintia* (*disks*), *Halythrix* (*cross*), and *E. filiformis* (*square*); pyritic replacement filaments are indicated by the diamond near the *y*-axis (Boal and Ng 2010)
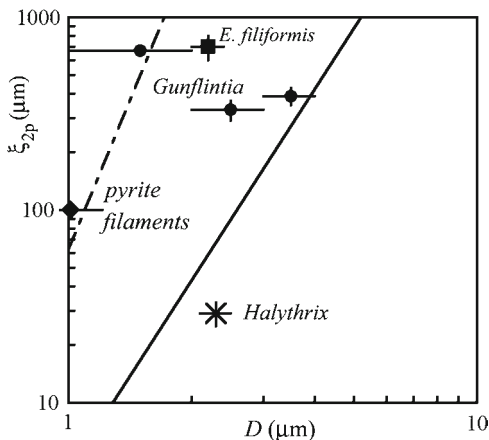
correlation length of the taxon *Halythrix* is much lower than the range found among the selection of modern cyanobacteria in Fig. 2.6, although it is not that far removed from the extrapolated fit to *Oscillatoria* in the figure.

The figure also displays ξ_{2p} from 3.23 Ga pyritic replacement filaments observed in a volcanogenic massive sulphide deposit (Rasmussen 2000). The magnitude of ξ_{2p} determined for these objects is below *Gunflintia* and *E. filiformis* by factors of three or more, although it is still larger than *Halythrix*. However, pyritic replacement filaments are also relatively narrow, and the combination of their width and tangent correlation length is completely consistent with an extrapolation of the empirical description of ξ_2 of *Geitlerinema*, as can be seen from Fig. 2.6, although this is not proof that the pyritic filaments had a biological origin.

The analysis of filament shapes demonstrates the similarity of the rigidity of both modern filamentous cyanobacteria and filamentous microfossils reaching back billions of years: most of the filaments exhibit a tangent correlation length that is one or two orders of magnitude greater than their radius. However, the similarity of their tangent correlation lengths by itself does not imply that these microfossil taxa must be cyanobacteria, as the correlation lengths of eukaryotic green algae with similar diameters also lie in this range; for example, $\xi_2 = 900 \pm 100$ nm for the green alga *Stichococcus* with a mean diameter of 3.5 ± 0.2 μm (Boal and Forde 2010). What this analysis does demonstrate is that the some aspects of the design of filamentous cells probably emerged fairly early in the history of life, and that filaments represent a robust and adaptable cell design.

2.4 Models for the Cell Division Cycle

At the molecular level, the materials suitable for the construction of cells must satisfy constraints imposed by the need for at least some of the cell's structural components to self-assemble and the need for the cell to change shape during the division cycle. At its simplest level, the division cycle only requires that the mean

volume V and area A of the cell double independently during the cycle once a steady-state growth pattern has been achieved. These changes in the area and volume of the cell can be accommodated through the use of a (two-dimensional) fluid boundary such as a lipid monolayer or bilayer, which has the flexibility to change shape as needed with a minimal cost in deformation energy. An illustration of a division cycle that even a fluid *mono*layer can support is displayed in Fig. 2.1: the monolayer grows by the random addition of new molecules and ultimately forms wavy arms that can pinch off to form new cells, assuming that the cell's genetic blueprint has been replicated on the same time scale as the growth in cell size. The main physical principle that drives this model cycle is one that nature adores: the maximization of entropy (simulation by Boal and Jun 2010; see Luisi 2006 and Zhu and Szostak 2009 for overviews of related experimental work).

The cell shape displayed in Fig. 2.1 belongs to a family of random shapes that obey branched polymer scaling, where the surface area is proportional to the enclosed volume, unlike a spherical balloon where $A \sim V^{2/3}$. However, these shapes are not efficient in the usage of materials: it isn't so much that the surface area grows so fast, but rather that the cell must be sufficiently large before (i) the branched polymer shapes emerge, and (ii) its genetic blueprints have replicated and separated. Given that cells must expend metabolic energy to produce the molecules composing the cell boundary, other routes to cell division may be more appropriate. One possible design is based on the use of molecules that, perhaps because of their spatial conformation, generate a membrane that is naturally deformed. For instance, a sphere of radius R has a surface with curvature $C = 1/R$. Suppose a membrane is made from molecules that favor surfaces that spontaneously deform to some particular curvature C_o. As more molecules are added to this membrane, it grows at constant curvature in the form of two overlapping spherical caps linked together at a ring with radius less than $1/C_o$, until the ring closes, leaving just two touching spheres. Now, this design is not flawless, in that the membrane curvature in the intersection region of the ring has the wrong sign – if the surface is concave (inwardly curved) over most of the linked spheres, it is convex (outwards) along the intersection ring itself, like the shape of an old-fashioned hour-glass.

Another group of materials-efficient cell designs is based on systems with two mechanical components. These components need not possess distinct molecular composition, but rather need to move independently of one another over some range of shapes. Even a bilayer composed of only one type of lipid is sufficient, so long as the time scale for lipid molecules to migrate between leaflets of the bilayer is sufficiently long. Consider, for example, the bilayer structures with symmetric molecular composition shown in Fig. 2.7. If the inter-leaflet migration of molecules is slow, then material produced within the cell can be added to the inner leaflet without immediately transferring to the outer one. For a short time, the inner leaflet can accommodate more molecules without increasing its area, resulting in an increase in its molecular density and, correspondingly, its state of strain. However, this strain can be relieved through buckling, which adds a ring of material to the inner leaflet, as shown in cross-section in Fig. 2.7b. Although there is deformation energy associated with the region where the ring joins the spherical shape of the inner leaflet,
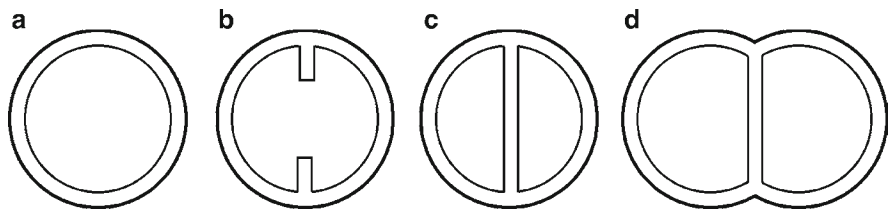
Fig. 2.7 One possible model for the division cycle of a cell with a boundary having two mechanical components. In panels (**a–c**), the inner layer grows and buckles to form two separate chambers, then both layers grow at constant curvature as in panel (**d**) until the original area and volume have doubled

for the most part the bilayer neck formed by the ring is flat and can be extended at no cost in deformation energy; in fact, the "hole" in the flat part of the membrane created by the ring possesses an edge tension that favors the contraction of the ring to form two separate chambers as in Fig. 2.7c. Of course, this is not a complete cell cycle, as the configuration in Fig. 2.7c still has the original volume and outer leaflet area as Fig. 2.7a; in fact, even the area of the inner leaflet has not doubled yet. However, if the time scale for the transfer of material from the inner to outer leaflets is not too long, the outer layer may now start to grow, allowing the enclosed volume to do likewise as in Fig. 2.7d.

Leaving aside the entropy-driven approach to cell division, we have described two "toy" models for the cell division cycle: (1) growth at constant curvature and (2) independent growth of bilayer leaflets. The time evolution of the length (L), surface area, and enclosed volume of the cell are different for each of these models: for growth at constant surface curvature, the cell elongates continuously, while in the independent leaflet model, the cell length is initially constant while the inner leaflet grows and buckles. Thus, there is a particular time dependence $L(t)$, $A(t)$, and $V(t)$ associated with the cell shape within each model for the division cycle. For living cells, this time dependence can be measured in the lab by photographing the growth of a single cell. Of course, this technique fails for microfossils, requiring the development of an alternate means of determining their division cycle. One such approach is based on the measurement of an ensemble of cells undergoing steady-state growth.

Suppose that we have an ensemble of n_{tot} cells whose shape we measure one by one. We assume that each cell started to grow at a random initial time, so the shapes of the cells in the sample are uncorrelated. Choosing a particular variable β (for example, length, area, volume...) we count that there are dn_β cells having a value of β between β and $\beta + d\beta$. Now, dn_β is a number, which necessarily depends on the total size of the sample n_{tot}. One can remove this dependence on the size of the experimental sample by constructing the probability density P_β (the probability per unit β) from the definition.

$$dn_\beta = n_{tot} \, P_\beta \, d\beta. \qquad (2.10)$$

By integrating Eq. (2.10) over β, one finds that P_β is normalized to unity: $\int P_\beta \, d\beta = 1$. Note that P_β has units of β^{-1}, whereas dn_β is simply a number. The link between P_β and the time-dependence $\beta(t)$ is that under steady state conditions, P_β is given by

$$P_\beta = (\partial\beta / \partial t)^{-1} / T_2, \tag{2.11}$$

where T_2 is the doubling time of the cell cycle. For illustration, suppose the cell has the shape of a uniform cylinder that increases in length L from ℓ to 2ℓ (at fixed radius) whereupon it divides symmetrically at time T_2. If $L(t)$ grows linearly with time as $L(t) = (1 + t/T_2)\,\ell$, then $P_\ell = 1/\ell$. In words, the physical meaning of Eq. (2.11) is that the more rapidly the quantity β changes, (i.e. larger $\partial\beta/\partial t$) the less time the cell spends in that range of β because $(\partial\beta/\partial t)^{-1}$ is small. This is familiar in the simple pendulum, which moves the fastest through its vertical position and slowest through its turning points, such that it is least likely to be found in the vertical position and most likely to be found at the turning points.

Next, consider the changes in shape of a diplococcus, which we represent as two intersecting spheres with the same radius R, like the outlines in Fig. 2.7. The cells in a single-species colony for either cyanobacteria or microfossils with the diplococcus shape are found to have uniformly similar radii, from which we conclude that the cells grow at constant width $2R$; their surfaces also appear to have curvature close to $1/R$. Thus, the length, area and volume of this family of shapes depend on only one geometrical quantity, which we choose to be the separation s between the centers of the intersecting spheres. Through the division cycle, the diplococcus grows from $s = 0$ (a single spherical cell) to $s = 2R$ (two spheres in contact), with the length L, area A, and volume V of the cell depending on s as

$$L = 2R(1+\beta) \tag{2.12a}$$

$$A = 4\pi R^2(1+\beta) \tag{2.12b}$$

$$V = (4\pi R^3/3) \bullet [1 + \beta(3 - \beta^2)/2] \tag{2.12c}$$

where $\beta \equiv s/2R$. Once the time dependence of just one of L, A, or V is known, the time dependence of the remainder is determined and the probability densities P can be calculated from Eq. (2.11). It turns out that the volume appears to have the simplest functional dependence on time, with a linear increase in time or exponential increase in time being the most likely (Bennett et al. 2007):

Linear volume increase During the doubling time T_2, the rate of change of the volume is constant at $dV/dt = (4\pi R^3/3)/T_2$. The time dependence of the overall cell length can be found from this form, which then yields

$$p_\beta = 3(1 - \beta^2)/2. \quad \text{(linear in } t\,) \tag{2.13}$$

Exponential volume increase In this situation, the rate at which the volume increases is proportional to the instantaneous value of the volume, or $dV/dt = V \ln 2/T_2$. From this,

$$p_\beta = [3(1 - \beta^2)/\ln 2]/[2 + \beta(3 - \beta^2)]. \quad \text{(exponential in } t\,) \tag{2.14}$$

Although it might appear to be somehow more consistent to calculate P_V rather than P_β, the cell's length can be more accurately measured than its volume, so that P_β is the more useful quantity.

2.5 Evolution of the Division Cycle of Rod-Like Cells and Diplococci

The methodology described in Section 2.3 permits the determination of the time evolution of a quantity (such as the cell length) through the measurement of an ensemble of cells undergoing steady state growth. Being a statistical technique, a reasonably good size sample must be taken to ensure accuracy: this is achievable for colonies with 200 or more identifiable microfossils, although in the lab, samples with more than 600 cells are preferred. Analyses of, and comparisons between, two morphologies of modern and ancient cells have been performed thus far: rod-like cells and diplococci (also referred to as dyads in some contexts). For theoretical reasons, the diplococcus morphology is the more useful of the two, and is the focus of this section (which follows Bennett et al. 2007).

Three taxa of microfossils with the diplococcus morphology are displayed in Fig. 2.8, along with a rod-like taxon in the upper left-hand corner of the figure
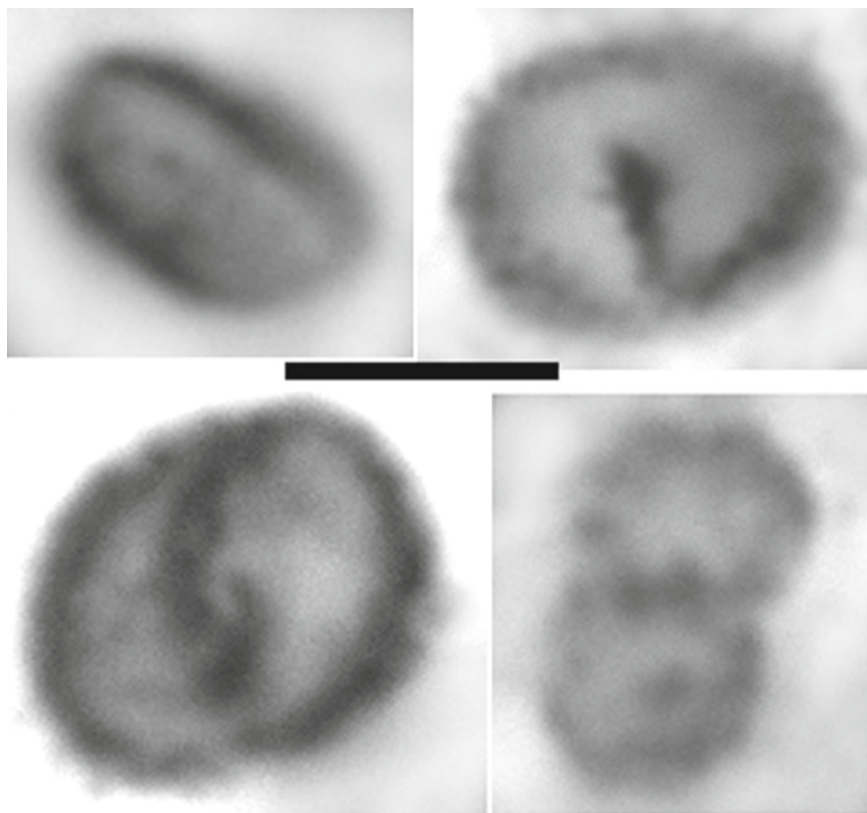
Fig.2.8 Examples of 2 Ga non-filamentous microfossils from the Belcher Island, Canada. Clockwise from the upper left: bacillus-like *E. moorei* and dyads *EB* (unclassified colony), *S. parvum*, and *E. belcherensis capsulata*. Scale bar, 5 μm (Bennett et al. 2007)

(microfossil specimens reported in Hofmann 1976). Modern diplococcal cyanobacteria have a very similar shapes, although the boundary is much crisper. In the strains described here, the smallest cells in a population are spherical, although there are some species where the smallest cells are slightly elongated, yet not so long as to be classified as bacillus-shaped. Colonies of ancient or modern diplococci all possess distributions in cell width that are very narrow: the standard deviation in cell width is less than 10% of the width itself. This suggests that the cells grow at fairly constant width, which is confirmed by scatter plots of cell width against cell length where only weak correlations are found to exist between the two variables. Thus, we are confident that Eq. (2.12) for L, A, and V of intersecting spheres with constant radius captures the shape of the cells to a good approximation.

The probability density P_β for the dimensionless separation $\beta = s/2R$ is shown in Fig. 2.9 for two species of the modern diplococcus *Synechocystis* from the Pasteur Culture Collection, PCC 6804 and PCC 6714 (Boal and Forde 2010). The strains have spherical initial configurations such that there is complete overlap of the two mathematical surfaces describing the general diplococcal shape: *i.e.*, $s = 0$ and consequently $\beta = 0$. As a result, P_β is non-vanishing in the smallest measurable range of β. In fact, P_β is peaked around $\beta = 0$, above which it declines and eventully vanishes as $\beta \rightarrow 1$.
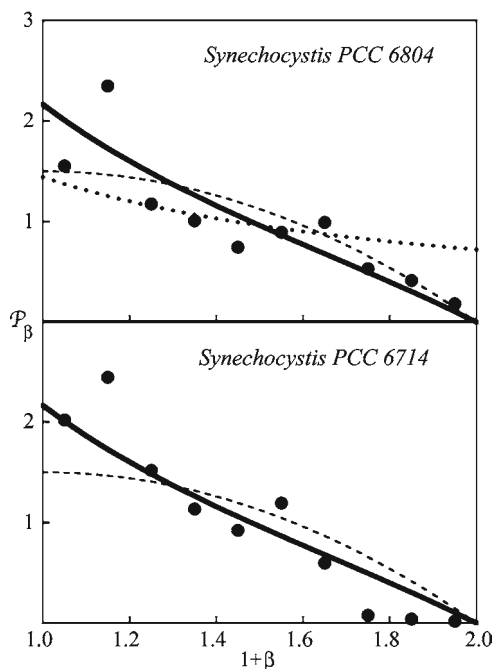
Fig. 2.9 Probability density P_β as a function of β for *Synechocystis* diplococci PCC 6804 and PCC 6714. Shown for comparison are predictions from models based on exponential (*solid curve*) or linear (*dashed curve*) volume growth, as well as exponential growth of cell length or area (*dotted curve*, *top panel* only) (Boal and Forde unpublished)

This behavior of P_β has the physical interpretation that the cell grows most slowly at the start of its division cycle (large P_β) and most rapidly at its end (small P_β).

How do the measurements of P_β in Fig. 2.9 compare with models of cell growth? If the length or area of a diplococcus grows linearly with time from $\beta = 0$ to $\beta = 1$, then P_β is constant at unity for all β, clearly in disagreement with the measurements. Exponential growth based on cell length or area should obey $P_\beta = [\ln2 (1 + \beta)]^{-1}$, which decreases from 1.44 at $\beta = 0$ to 0.72 at $\beta = 1$. These values are well below the data in Fig. 2.9 at $\beta = 0$ and well above them at $\beta = 1$. In other words, neither of the length/area-based growth models with linear or exponential time dependence agrees with the observed behavior of P_β. The predicted P_β based upon volume growth with linear or exponential time dependence is also plotted in Fig. 2.9. The differences between the theoretical curves are obviously not large, which is expected because the exponential function e^x is approximately linear in x at small x. The intercept of P_β at $\beta = 0$ is predicted to be $3/(2 \ln2) = 2.16$ for exponential volume growth and $3/2$ for linear volume growth. The higher (exponential) value of P_β ($\beta = 0$) is mildly preferred by experiment in both panels of the figure, but linear volume growth is not ruled out. This preference for exponential growth in volume is seen for many other cyanobacteria, including other diplococci as well as rod-like cells (Boal and Forde unpublished).

The lab specimens reported in Fig. 2.9 were prepared such that the symmetry axis of each cell was made to lie in the observational plane. In contrast, fossilized cells may have random orientations, so that either their shape must be reconstructed in three dimensions, or the analysis must take into account the orientation if the measured shapes are projections onto a plane. Here, we work with the unitless probability density P_Λ, where Λ is the projected cell length divided by the cell width (the width is unchanged by projection); Λ runs from 1 when $\beta = 0$, to 2 when $\beta = 1$. The correspondence between β and Λ is not unique, in that a range of configurations in β can contribute to a given Λ: for example, cells of any length may have $\Lambda = 1$ if their symmetry axis points towards the observer. Figure 2.10 displays P_Λ for three taxa of fossilized diplococci and the probability density is seen to rise rapidly to a maximum exceeding $P_\Lambda = 2.5$ at Λ near 1 before falling more gently as Λ approaches its upper limit of two for linked spheres. This behavior is very similar to that of P_β for modern diplococci in Fig. 2.9.

In terms of mathematical models for cell growth, it is not difficult to obtain P_Λ from P_β, and Fig. 2.10 contains the predictions of two models for the cell division cycle: exponential growth in volume (solid curve) and exponential growth in area (dashed curve). Exponential growth in area is a poor representation of the measured P_Λ, particularly at Λ near one, just as this same model failed to capture the behavior of P_β for modern diplococci. Linear growth in volume also underpredicts P_Λ at $\Lambda = 1$, where it approaches $3\pi/4 = 2.36$; P_Λ then exceeds the data as $\Lambda \to 2$ in this model. In contrast, exponential growth in volume corresponds most closely to both the quantitative and qualitative features of P_Λ seen in Fig. 2.10, and predicts $P_\Lambda(1) = 3.34$.

Let's now return to the principles behind the models for cell growth. Models where the growth of a cell is linear in time assume that change occurs at the same rate throughout the division cycle no matter what the contents of the cell. Examples
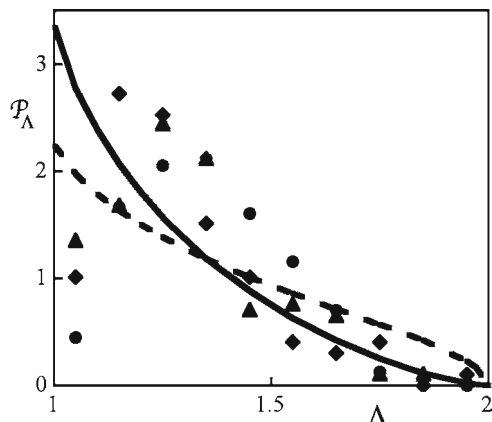
Fig. 2.10 Combined data for the probability density P_Λ of the dyads *S. parvum* (*diamonds*), *E. belcherensis capsulata* (*circles*), and an unclassified colony labeled *EB* (*triangles*) compared to the expectations of growth at constant curvature and exponential increase in volume (*solid curve*) or area (*dashed curve*). The largest values of P_Λ have about 30 cells per data bin for a statistical uncertainty of about 20% per individual datum (Bennett et al. 2007)

of linear models can be found in some eukaryotic cells, where cell mass grows linearly with time (Killander and Zetterberg 1965). Here, we observe that the only linear model not immediately ruled out by data is the linear rise in volume, for which agreement with data is marginal. Exponential growth may arise from several different mechanistic origins. Exponential growth in area corresponds to new surface being created at a rate proportional to the area available to absorb new material – a logical possibility but not supported by Fig. 2.10. Lastly, exponential growth in volume arises if new volume is created at a rate proportional to the cell's contents, which is the only scenario to comfortably describe the data. What we conclude from Figs. 2.9 and 2.10 is that exponential growth in volume has very likely been a guiding principle for the division cycle that was established at least two billion years ago for bacteria with the diplococcus morphology.

2.6 Summary

This article has examined the role that continuum and statistical mechanics plays in determining the size, shape and functionality of the simplest cell designs, focusing on cell morphologies such as diplococci and filaments that have at the most two important structural components. We described the bending resistance, edge tension and rupture resistance of lipid bilayers and showed the constraints that these elastic and mechanical properties place on the size of the simplest cells. For example, the rupture resistance of the bilayer generates a pressure-dependent cell radius beyond which a bilayer requires a cell wall for reinforcement. Similarly, the bending

resistance and edge tension of the bilayer set a minimal membrane area for the cell to spontaneously close into a spherical topology. Mechanical principles also may dominate the simplest or earliest forms of the cell division cycle. For example, the maximization of entropy can lead to a division cycle in which the membrane grows until it produces entropy-rich arms that can pinch off to form new cells if DNA replication and separation are appropriately choreographed. The minimization of deformation energy or of the consumption of materials also favors specific forms of the division cycle.

We investigated the mechanical features of cells more than two billion years old using a combination of statistical mechanics and comparisons between modern cyanobacteria and microfossils. In Section 2.2, we characterized the sinuous behavior of filamentous cells by means of a tangent correlation length ξ_t, demonstrated its power law dependence on filament diameter as D^n within a given genus, and provided an argument from continuum mechanics that this power law should have an exponent n in the range of 3–4. We also obtained soft bounds on the relationship between ξ_t and D for modern filamentous cyanobacteria according to the length-to-width ratio of individual cells within a filament, and then demonstrated that filamentous microfossils and pyritic replacement filaments satisfied these bounds. From this, we argued that the general mechanical features of filamentous cells were probably established relatively early in the development of life. Finally, in Section 2.4, we examined the division cycles of non-filamentous cells, focusing heavily on diplococci, in which we applied a technique that extracts the time dependence of a geometrical observable such as cell volume from an analysis of a colony of cells under steady-state growth conditions. It's found that that modern cyanobacteria and microfossils with rod-like or diplococcal shape are most consistent with exponential volume growth (although linear growth in volume cannot be ruled out). This argues that the volume of a cell increases with the volumetric contents of the cell, a division cycle that dates back at least two billion years.

References

Barghoorn ES, Schopf JW (1966) Microorganisms three billion years old from the Precambrian of South Africa. Science 152:758–763

Barghoorn ES, Tyler SA (1965) Microorganisms from the Gunflint chert. Science 147:563–577

Bennett S, Boal DH, Ruotsalainen H (2007) Growth modes of 2-Ga microfossils. Paleobiology 33:382–396

Boal DH (2002) Mechanics of the cell. Cambridge University Press, Cambridge

Boal DH, Forde CE (2010) Unpublished

Boal DH, Jun S (2010) Unpublished

Boal DH, Ng J (2010) Shape analysis of filamentous Precambrian microfossils and modern cyanobacteria. Paleobiology 36:555–572

Boal DH, Rao M (1992) Topology changes in fluid membranes. Phys Rev A46:3037–3045

Bustamante C, Marko JF, Siggia ED, Smith S (1994) Entropic elasticity of Λ-phage DNA. Science 265:1599–1600

Cloud PE Jr (1965) Significance of the Gunflint (Precambrian) microflora. Science 148:27–35

Doi M, Edwards SF (1986) The theory of polymer dynamics. Oxford University Press, Oxford

Evans E, Rawicz W (1990) Entropy-driven tension and elasticity in condensed-fluid membranes. Phys Rev Lett 64:2094–2097

Fromhertz P (1983) Lipid-vesicle structure: size control by edge-active agents. Chem Phys Lett 94:259–266

Helfrich W (1973) Elastic properties of lipid bilayers: theory and possible experiments. Z Naturforsch 28c:693–703

Hofmann HJ (1976) Precambrian microflora, Belcher Islands, Canada: significance and systematics. J Paleo 50:1040–1073

Killander D, Zetterberg A (1965) A quantitative cytochemical investigation of the relationship between cell mass and initiation of DNA synthesis in mouse fibroblasts in vitro. Experimental Cell Research, 40:12–20

Luisi PL (2006) The emergence of life: from chemical origins to synthetic biology. Cambridge University Press, Cambridge

Needham D, Hochmuth RM (1989) Electromechanical permeabilization of lipid vesicles. Biophys J 55:1001–1009

Rasmussen B (2000) Filamentous microfossils in a 3.235-million-year-old volcanogenic massive sulphide deposit. Nature 405:676–679

Schopf JW (1968) Microflora of the bitter springs formation, Late Precambrian, central Australia. J Paleo 42:651–688

Schopf JW (1993) Microfossils of the early Archean Apex chert: new evidence of the antiquity of life. Science 260:640–646

Schopf JW, Packer BM (1987) Early Archean (3.3-billion to 3.5 billion-year-old) microfossils from Warrawoona Group, Australia. Science 237:70–73

Walsh MM, Lowe DR (1985) Filamentous microfossils from the 3, 500 Myr-old Onverwacht Group, Barberton Mountain Land, South Africa. Nature 314:530–532

Zhu TF, Szostak JW (2009) Coupled growth and division of model protocell membranes. J Am Chem Soc 131:5705–5713

Chapter 3
On the Minimal Requirements
for the Emergence of Cellular Crowding

Luis Acerenza and Martín Graña

Abstract A plausible route to highly concentrated cellular compartments is sketched. The main driving force comes from the selective advantage of a higher macromolecular content. The evolution from a diluted to a crowded state is described with a minimal model of growing bacteria. Additional simplifications to the model, aimed to mimic ancestral protocellular growth and division, lead to a model still rendering a selective advantage to increased macromolecular concentration.

3.1 Introduction

The state of modern cells is characterised by a high macromolecular concentration, which remains in a narrow range of values (Fulton 1982; Zimmerman and Trach 1991). In *Escherichia coli*, this state holds across many physiological conditions and through evolution (Bremer and Dennis 1996; Lenski et al. 1998; Schaechter et al. 1958). Recently, the degree of macromolecular crowding in different mammalian cell lines has been shown to be conserved (Guigas et al. 2007). In contrast, as runs for many aspects of 'life as we don't know it' (Wachtershauser 1992; Wachtershauser 2000), polymer concentrations in pre-biotic and early life environments are largely unknown, most arguments in the literature suggesting that it was rather low (Lazcano and Miller 1999).

L. Acerenza (✉)
Systems Biology Laboratory, University of the Republic, Montevideo, Uruguay
e-mail: aceren@fcien.edu.uy

M. Graña (✉)
Bioinformatics Unit, Pasteur Institute of Montevideo, Montevideo, Uruguay
e-mail: mgrana@pasteur.edu.uy

The consequences of a high total macromolecular concentration and associated low free water availability on cytoplasm organisation and function, was the subject of many scientific contributions. For instance, effects of macromolecular crowding have been reported in phenomena belonging to different levels of cellular organisation: enzyme kinetics, protein folding, biochemical pathways, signal transduction processes, cell volume regulation, active transport, DNA conformation and function and protein-protein association (Laurent 1971; Minton 1981; Bray 1998; Ellis 2001; Ellis and Minton 2003; Garner and Burg 1994; Goobes et al. 2003; Kornberg 2000; Rohwer et al. 1998; van den Berg et al. 2000; Zimmerman and Minton 1993; Zimmerman and Murphy 1996; Walter and Brooks 1995; Kozer et al. 2007). These effects depend on the existence of non-specific macromolecular interactions operating, at the physiochemical level, on the thermodynamic and kinetic properties of the system (Hall and Minton 2003; Zimmerman and Minton 1993).

For what concerns the potential advantageous effect of macromolecular crowding on fitness, several considerations have been previously made. For example, cellular crowding has been invoked to participate in mechanisms stabilising enzyme complexes and macromolecular interactions (Goobes et al. 2003; Minton 1983; Record et al. 1998). In a treatment going from molecule interactions to physiology, macromolecular crowding was suggested to provide a sensor mechanism of cell volume change (Burg 2000; Minton et al. 1992). Furthermore, macromolecular crowding, together with diffusion, was proposed as one of the fundamental constraint in the evolution of cell signal transduction pathways (Bray 1998). At the metabolic level, the high cellular concentration of macromolecules was associated to a cytoplasm organisation where substrate channelling would take place (Srere 1987; Srere and Ovadi 1990). These studies suggest that some physicochemical properties of proteins placed in crowded cellular environments, could potentially confer the organism an evolutionary advantage.

In spite of the above-described arguments, suggesting that evolving from a diluted to a crowded cytoplasm could be beneficial, the path followed during evolution to achieve the crowded cellular state is by no means evident. On one hand, we know that the increase of protein content in an already crowded cell could negatively impact on fitness (Cayley et al. 1991). More importantly, in a putative diluted ancestral scenario the non-specific macromolecular interactions operating at the physiochemical level would be absent. Therefore, the non-specific interactions cannot be invoked as the only driving force leading to a high cellular macromolecular concentration.

In order to gain insight into the origins of a crowded cytoplasm, we moved from a molecular-level to a cellular-level description, that is, from macromolecular interactions to physiological processes. This systemic point of view allowed us to show that the increase in concentration of proteins, maintaining their physicochemical properties unchanged, suffices to explain the evolutionary advantage of having highly concentrated intracellular milieus (Acerenza and Graña 2006). First, the effect of increasing cellular protein concentration was tested in a mathematical

model of *E. coli*, describing cell cycle dynamics and long-term experimental evolution of cell volume and growth rate (Graña and Acerenza 2001). This minimal bacterial model has some features probably absent in the early protocells. Further simplifications of this model, to mimic protocellular structures with very primitive mechanisms of growth and division, still support a selective advantage of protocells with higher macromolecular contents.

3.2   Minimal Bacterial Model

The mathematical model of *E. coli* used here was originally developed to explain a rather striking result coming from experimental evolution (Graña and Acerenza 2001). Evolving populations of *E. coli* under controlled conditions, founded by a single clone, were monitored through thousands of generations (Lenski 2009); all of them showed a tandem increase in mean cell volume and relative fitness (Lenski and Travisano 1994). Using scaling arguments, one would expect that a larger organism, having a smaller surface to volume ratio and a smaller input of nutrient per unit volume, would grow slower, which is in sharp contrast with experiments. To explain this type of behaviour we found it necessary to complement the model with processes operating at two levels: organism physiology and population genetics (Graña and Acerenza 2001).

We wanted to depict how cell volume, v, and relative fitness, μ / μ_0 (growth rate of derived population relative to ancestor), change during evolution. Therefore, the physiology relevant to the evolutionary question was: the relationship between cell volume and growth rate, v *vs* μ (Donachie and Robinson 1987) and the growth rate response to changes in external nutrient concentration, μ *vs* X (Monod 1949). Such physiological responses were implemented in a modular model. The nutrient incorporated to the cell is consumed in two modules. The "growth module" includes all the metabolic processes degrading nutrient for cellular maintenance and growth, and the "adaptation module" uses resources for adaptation to changes in the environment. It is worth noting that the criterion to define the modules is functional; it is not directly related to metabolic structure. For instance, if cells grow consuming glucose as the only carbon source then processes transforming other sugars belong to the adaptation module and not to the growth module (as would be the case if these other sugars were degraded for growth). The growth module is in charge of the production of cell protein and, therefore, of the increase in cell volume during growth. But, as we shall see, the rate of volume increase would also depend on the activity of the adaptation module, which competes with the growth module for the nutrient.

The minimal model may be formulated using three differential equations, one for each of the following variables: number of nutrient molecules inside the cell, n, number of protein molecules, P, and number of signal molecules for initiation of DNA replication, s. In the first differential equation,

$$\frac{dn}{dt} = k_i \frac{X}{K_X + X}\left((\pi d^2 / 2) + 4v / d\right) - (k_g + k_a)n,$$

the nutrient enters the cylindrical cell through its surface (first term in the r.h.s.) and is consumed by the growth and adaptation modules (second and third terms in the r.h.s., respectively). v is the cell volume. The second differential equation describes protein synthesis and degradation:

$$\frac{dP}{dt} = k_g n - k_d P.$$

The dynamic behaviour of cell volume, v, under the constraint that protein is produced and degraded remaining its total concentration (C_T) constant, is simply given by:

$$v = \frac{P}{C_T}$$

The synthesis and degradation of signal molecules is described by a third differential equation:

$$\frac{ds}{dt} = k_{gs} n - k_{ds} s.$$

The system of differential equations is integrated until a time t_u, at which s reaches a threshold value s_u. At this point signal production stops and DNA replication starts. The number of signal molecules remains unchanged from this moment until cell division ends. The cell divides a time t^* after initiation of replication, this interval of time being approximately constant for many *E. coli* strains (Helmstetter 1996). Time intervals are obtained by the expression $t^* = 60 / \mu^*$ where $\mu^* = (k^* X / (K^* + X)) (2000 / C_T))$ (for details see Acerenza and Graña 2006). The doubling time t_d and growth rate μ are calculated with: $t_d = t_u + t^*$ and $\mu = 60 / t_d$.

Finally, population-level processes were considered. Many metabolic functions have been found to decay through evolution in controlled conditions, such as the ability to metabolise substrates actually absent in the medium (Cooper and Lenski 2000). We implemented this genetic decay through a simple population genetics framework. The resulting model, combining cell physiology and population genetics, succeeded in quantitatively reproducing the parallel increase in cell volume and fitness, in the evolution of *E. coli* during 10,000 generations, which was the original aim motivating the development of the model (Graña and Acerenza 2001).

The core message conveyed by this minimal model is as follows. The decay of unused functions during evolution lowers nutrient consumption for physiological adaptation purposes, increasing nutrient availability for protein production. Since the organism keeps a roughly constant composition through cell cycle and

evolution, the surplus of protein produced leads to increased volumes. The rate of signal biosynthesis also rises with increased resources allocated to growth, shortening doubling time and thus enhancing growth rate.

The model can also explain other known experimental data from *E. coli* physiology and evolution. The fact that some division mutants of *E. coli* (e.g., *fts Z* mutants) form filaments of several hundred times the normal bacterial size is explained by a key property of the model, namely, that volume grows indefinitely if cell division is not imposed (Lutkenhaus and Mukerjee 1996). Incorporating in *E. coli* an inducible plasmid carrying the *dnaA* gene and overproducing its product, the signal protein DnaA, results in smaller and faster growing cells what in the model corresponds to obtaining a smaller cell volume and higher growth rate when increasing k_{gs} (Løbner-Olesen et al. 1989). Experimental evidence favouring antagonistic pleiotropy (i.e. detrimental mutations in a previous environment that are beneficial in a new one) as a mechanism overriding neutral mutation accumulation for the decay of unused catabolic functions during controlled evolution, is in agreement with a property of the model consisting in that the decay of the adaptation module in the new environment increases growth rate (Cooper and Lenski 2000). Experimental curves representing the physiology of cell volume and growth rate, obtained at different stages of evolution, have the same shape as those calculated with the model (Mongold and Lenski 1996). Finally, the increase in yield of processes converting glucose into cell mass during *E. coli* controlled evolution is also explained by the model, because as a consequence of the genetic decay of the adaptation module, nutrient previously consumed for adaptation is used for growth (Lenski and Mongold 2000). In fact, we are not aware of any experimental results relating to the physiology and evolution of cell volume and relative fitness of *E. coli*, testable in the model, that the model could not explain.

The minimal bacterial model described above was used to test whether (and how) changing protein concentration would affect growth rate (Acerenza and Graña 2006). During the time scales of physiological responses and short term evolution, protein concentration (C_T) in the model remains bounded. Therefore, C_T is taken as a parameter. Yet, on long term evolutionary time scales, protein concentration could change, what may be thought of as a slow change in time of this parameter. Would the parameter change be advantageous, i.e. lead to increased relative population fitness, the new value of protein concentration and its phenotypic consequences would be retained and fixed by selection.

In the model, all rates are proportional to the protein concentrations sustaining them. This model formulation represents a putative diluted ancestral scenario with protein association-dissociation processes being shifted to dissociation, promoting independent functioning of macromolecules. A change in protein concentration will produce a proportional change in the activities of all the proteins, leaving the functions of the individual molecules unaffected. Therefore, implementing a change in protein concentration by a certain factor in the model requires changing C_T and the five parameters describing the activities of the processes taking place (k_i, k_g, k_a, k_{gs} and k^*) by the same factor. The effect of protein concentration on the growth rate is obtained studying the effect of a simultaneous change in these six parameters.
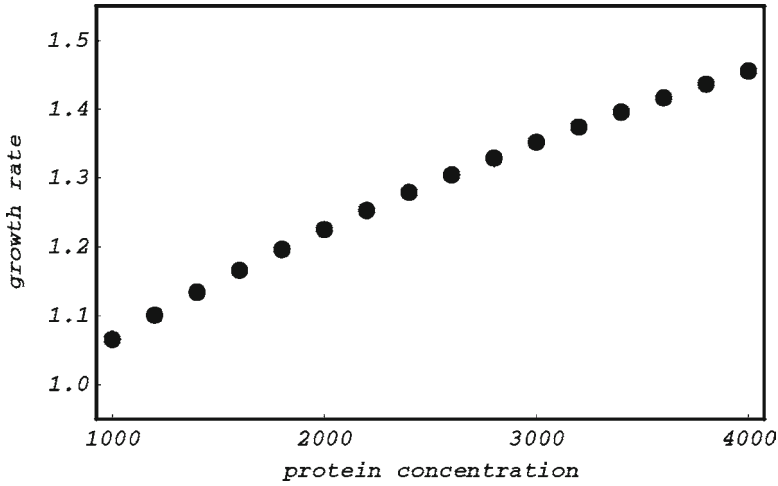
Fig. 3.1 Minimal bacterial model. Growth rate versus protein concentration. Parameter values are: $X = 5$, $k_i = 40$, $k_g = 1.5$, $k_a = 4.5$, $k_{gs} = 2.5\,10^{-5}$, $k^* = 3$, $K_X = 5$, $K^* = 1$, $k_d = 10^{-6}$, $k_{ds} = 10^{-6}$, $s_u = 0.02$ and $d = 1$. For details see text (Adapted from: Acerenza and Graña 2006)

In Fig. 3.1, we plotted the growth rate as a function of protein concentration, C_T (k_i, k_g, k_a, k_{gs} and k^* were simultaneously changed by the same factor as C_T). Increasing the six parameters results in an increased growth rate. Hence, these results suggest that an increased protein concentration under diluted conditions (with unchanged proportions and specific properties of the proteins) increases the fitness of the organism (Acerenza and Graña 2006). Next we give additional results showing why the model behaves in this way.

The increase in activity of the growth module leads to an augmented production rate of the signal. Therefore, at higher protein concentrations, the threshold value for the signal (s_u) is reached at smaller volumes (Fig. 3.2a). Smaller volumes give higher surface to volume ratios, and therefore higher inputs of nutrient per unit time and per unit volume, what contributes to growth rate increase. In Fig. 3.2b, we plotted growth rate over surface to volume ratio as a function of protein concentration. The constant relationship obtained suggests that increased mean surface to volume ratios, appearing for higher values of protein concentrations, would be the main contributors to a growth rate increase.

Despite being a highly simplified version of reality, the minimal bacterial model shows features found in rather evolved cellular organisms. For instance, the implemented cell division mechanism depends on the existence of a signal reaching a given threshold for initiation of replication. In order to study the effect of macromolecular content increase under conditions that could represent an earlier scenario of protocellular growth and division, we will further simplify the minimal bacterial model in the next section.
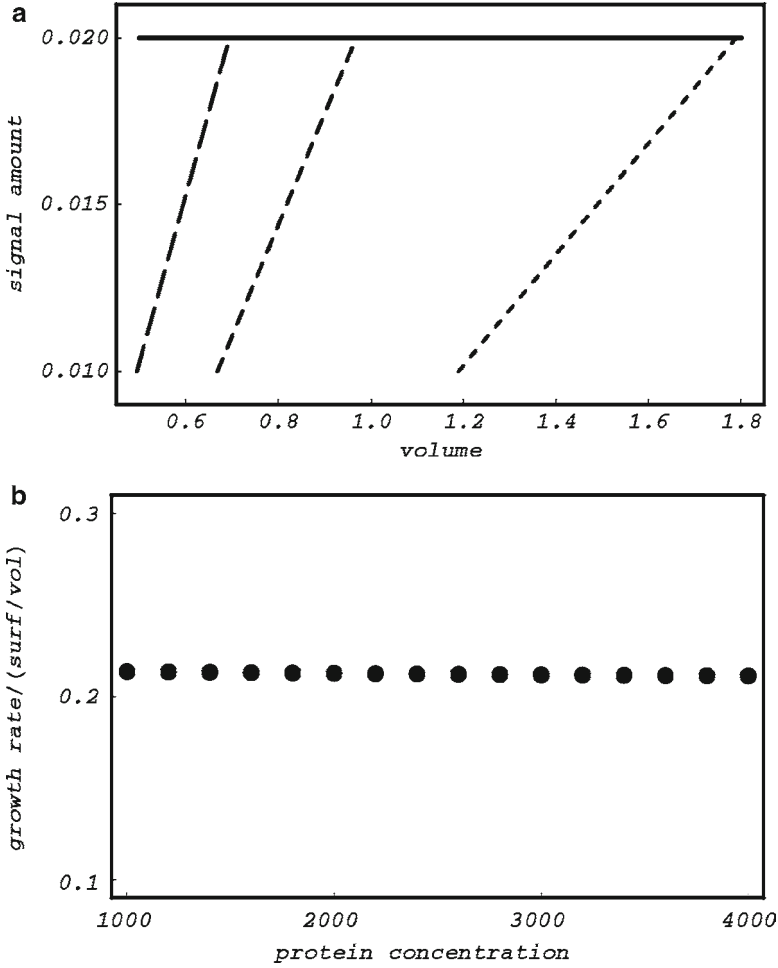
Fig. 3.2 (**a**) Minimal bacterial model. Parametric plot $s(t)$ versus $v(t)$. Dashed curves for $C_T =$ 1,000, 2,000 and 3,000 (longer dashes represent higher values of C_T). Threshold value for the signal ($s_u = 0.02$) shown in thick line. At higher C_T, the threshold value of the signal is reached for smaller volumes. (**b**) Minimal bacterial model. (Growth rate)/(surface/volume) versus protein concentration. For details see text (Adapted from: Acerenza and Graña 2006)

3.3   Minimal Protocellular Model

The minimal bacterial model from the previous section can be simplified to represent an ancestral scenario of protocellular evolution. The simplified model is described next.

Two modules consume the nutrient entering the protocell: the "growth module" and the "waste module". Part of the resources is used for growth and maintenance by the "growth module", as was the case for the bacterial model. But, in the protocellular

model, the rest of nutrient is assumed to be consumed by a "waste module". In an early evolutionary stage, many processes are supposed to operate in an inefficient way. For example, macromolecules should often show very low specificity resulting in useless end-products and useful metabolic intermediates showing inadequate chemical properties are significantly lost by diffusion to the exterior of the protocell. Additionally, sophisticated adaptation mechanisms intended to correct the effects of environmental fluctuations should still be absent. Therefore, we assume that a substantial part of the nutrient incorporated is consumed, by a "waste module", with no apparent benefit to the protocell. To keep nomenclature as homogeneous as possible, we also use the term "protein" for macromolecular species catalysing protocellular processes.

The minimal protocellular model is formulated using two differential equations. The variables are: number of nutrient molecules inside the protocell, n, and number of protein molecules, P. The first differential equation,

$$\frac{dn}{dt} = k_i \frac{X}{K_X + X} (4\pi)^{1/3} (3v)^{2/3} - (k_g + k_w)n,$$

shows an input of nutrient similar to the bacterial model, the difference being that the protocell is a sphere and not a cylinder. In this equation, v is the protocell volume and $(4\pi)^{1/3} (3v)^{2/3}$ is the surface expressed in terms of the volume. Nutrient is consumed by growth and waste modules (second and third terms in the r.h.s.), being k_g and k_w the activity constants associated to these modules, respectively. The second differential equation describes protein synthesis and degradation:

$$\frac{dP}{dt} = k_g n - k_d P.$$

Here we also assume that proteins are being produced and degraded, while their total concentrations (C_T) remain constant, v being given by:

$$v = \frac{P}{C_T}$$

The two differential equations are integrated until a time t_d, when the protocell reaches a critical volume, v_c. At this moment the protocell splits into two 'identical' daughters. v_c is the volume beyond which membranes become unstable, increasing the probability of spontaneous division (Rashevsky 1960; Sole et al. 2009).

Note that the parameter values used in the bacterial and protocelluar models (see legends to Figs. 3.1 and 3.3) reflect the reasonable assumption that in an ancestral scenario the activity of the processes are much smaller than in present organisms.

Next we test in the model the effect of protein concentration on growth rate. A change in protein concentration will produce a proportional change in the activities
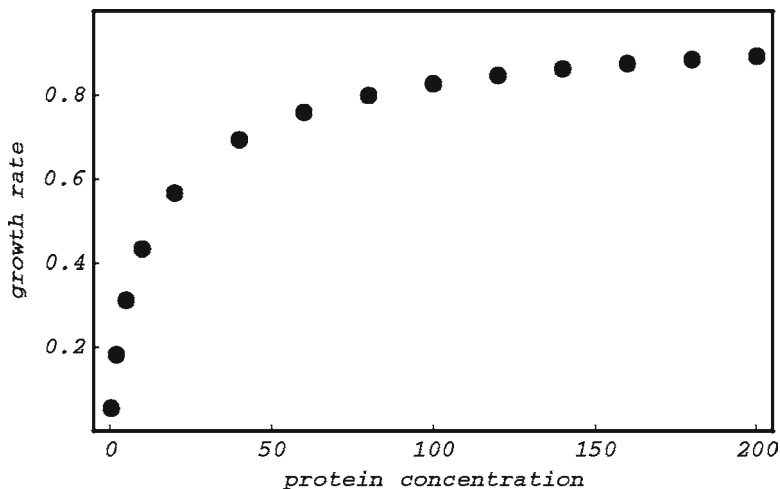
Fig. 3.3 Minimal protocellular model. Growth rate versus protein concentration. Parameter values are: $X = 5$, $k_i = 10.52 \, 10^{-3}$, $k_g = 1.5 \, 10^{-3}$, $k_w = 4.5 \, 10^{-3}$, $K_X = 5$, $k_d = 10^{-6}$ and $v_c = 2$. For details see text

of all the proteins. In this protocell we assume that there are two processes catalysed by proteins: nutrient incorporation (k_i) and growth (k_g). Therefore, implementing in the model a change in protein concentration by a certain factor requires changing C_T, k_i and k_g simultaneously by the same factor.

In Fig. 3.3, growth rate is plotted as a function C_T (k_i and k_g simultaneously changed by the same factor as C_T). Increasing these three parameters results in an increased growth rate. Therefore, in the ancestral scenario described by the model, an increased protein concentration still increases protocellular fitness. Next we show additional results suggesting how the model generates this property.

The increase in activity of the input module (k_i) produces an increase in the nutrient internal to the protocell, n. This tends to increase the flux of the growth module (j_g) and the flux of the waste module (j_w). But since the activity of the growth module (k_g) also increases with protein concentration, the increase in j_g is greater than the increase in j_w, being the ratio of the fluxes, j_g / j_w, an increasing function (Fig. 3.4). Therefore, the increase in growth rate with protein concentration (Fig. 3.3) could be explained by the reassignment of resources from waste to growth.

3.4   Final Remarks

We argue that the origin of a highly concentrated cytoplasm found in contemporary cells could have a plausible explanation in terms of fitness gain produced by the effects that increasing protein concentrations, in diluted conditions, have on cellular
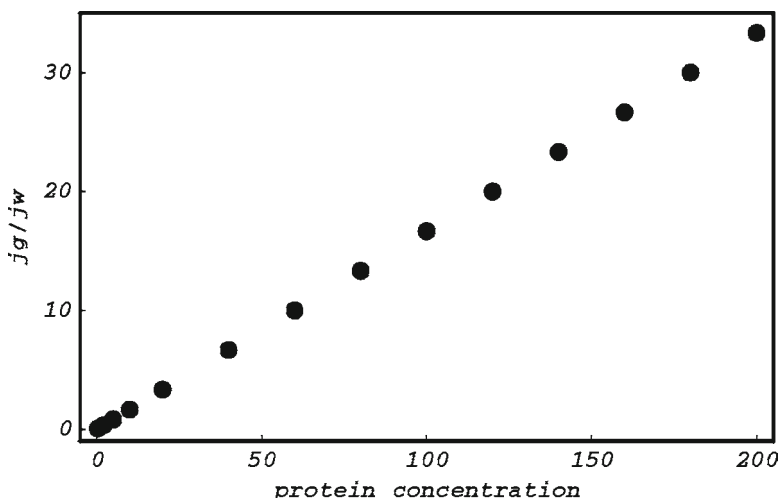
Fig. 3.4 Minimal protocellular model. (Growth flux/Waste Flux) versus protein concentration

physiology. Contributions of macromolecular crowding rooted in non-specific interactions, appearing when relatively high protein concentrations are attained, do superimpose to this effect.

Two models were used to test the effect of protein concentration on growth rate. The first is a minimal bacterial model previously proposed to describe the parallel increase in cell volume and fitness in the course of *E. coli* experimental evolution. In this model the input of nutrient is consumed by growth and adaptation modules. The rate of the three processes, input of nutrient, growth and adaptation, increase proportionally with protein concentration. Division is imposed with a mechanism similar to the one used by present bacteria, i.e., a signal is produced until it reaches a threshold value, at this moment signal production stops and DNA replication starts and, finally, division takes place at an approximately fixed time after initiation of replication. The second model is a minimal protocellular model obtained simplifying the first one. In this model the adaptation module has been replaced by a waste module, representing the loss of nutrient through processes that confer no benefit to the protocell. This process is assumed to be independent of protein concentration (in practice it could show a lower dependence on protein concentration than the other processes, what was neglected in the model). In addition, the elaborated genetic mechanism of replication and division is replaced by a biophysical mechanism of division based on the relationship between membrane stability and protocell size (Rashevsky 1960; Sole et al. 2009). In both models, increasing protein concentration enhances growth rate, conferring a selective advantage to the organism. Interestingly, if in the minimal protocellular model the input of nutrient does not depend on protein concentration, a significant increase in growth rate with protein concentration appears to be difficult to obtain.

It is important to note that the mechanisms underlying the selective advantage of higher protein concentrations in the bacterial and protocellular models are very

different. In the bacteria, the change in mean surface to volume ratio with protein concentration appears to be a central determinant of the selective advantage (Fig. 3.2b). In the protocell, division occurs after reaching a critical volume, v_c, which is assumed to be independent of protein concentration. Therefore, the mean surface to volume ratio does not change with protein concentration. On the other hand, in the protocellular model, the increase on the fluxes ratio, j_g / j_w, with protein concentration, appears to make an important contribution to produce higher growth rates. But, in the bacterial model the ratio of the fluxes of the growth and adaptation modules, j_g / j_a, does not change with protein concentration (results not shown).

Although the minimal bacterial model is a highly simplified representation of bacterial functioning, it includes a rather elaborated cell division mechanism with similar features to those found in present organisms. Therefore, if used to represent an ancestral organism, this should be at a sufficiently advanced stage of evolution where cellular metabolism is based on the co-ordinated action of nucleic acids and proteins. On the other hand, the minimal protocellular model is intended to capture earlier growth and division modes. It could, for example, represent a protocellular structure having a metabolic organisation based on autocatalytic sets of proteins or nucleic acids (Dyson 1985; Kauffman 1986; Luisi 2002; Luisi et al. 2002; Luisi et al. 2006; Oberholzer and Luisi 2002; Stano et al. 2009; Szostak et al. 2001). Such systems would display metabolic networks with catalytic closure, whose self-propagation would fuel growth and division processes of a statistical nature (Segre et al. 2000).

In our context, a mechanism of evolution of total cellular protein concentration could be as follows. During pre-biotic evolution, where rather diluted conditions likely prevailed, protein concentration increases were driven solely by their effect on cellular physiology. This effect could be present in protocells devoid of nucleic acids or in early cellular organisms with a metabolic functioning based on proteins and nucleic acids. At some point during evolution, proteins reached concentrations sufficiently high for macromolecular non-specific interactions to enter into operation. These interactions could have positive or negative effects on fitness. Advantageous interactions reinforce the positive effect on physiology and are retained. In contrast, detrimental interactions may prevent additional increases in protein concentration by neutralizing beneficial physiological effects, and are replaced by mutation/selection of protein variants not promoting these unfavourable interactions. Consequently, further increases in protein concentration driven by the effects on cell physiology, or by newly evolved favourable non-specific molecular interactions, could be obtained. This process would continue until reaching a highly crowded cytoplasm. At this stage, increases in total macromolecular concentration, even small ones, are expected to strongly impact thermodynamic/kinetic properties of intracellular reactions. Therefore, the evolutionary trend from low to high concentrations would slower its pace. Paralleling the total macromolecular concentration increase, physiological mechanisms to maintain and restore this concentration – as found in extant organisms, – would appear and be fixed to counteract harmful effects of external perturbations (Minton 2001). In addition to "well-mixed" macromolecular crowding,

described above, elaborated forms of spatial architecture would appear with protein concentration increase through evolution. For instance, some proteins would start to reversibly associate with cellular structural components, forming different association-dissociation states with different activities and/or functions. Other proteins would form complexes, promoting the channelling of metabolic intermediates between active sites, with minor diffusion into the bulk phase, thus enhancing the overall flexibility of metabolic regulation (Clegg 1984; Winkel 2004). Recent findings show that widespread reorganisation of metabolic enzymes into reversible assemblies could play a major role in yeast response to nutrient starvation (Narayanaswamy et al. 2009). The appearance of these complex forms of organisation during evolution would be facilitated by a previously reached internal milieu with high protein concentration, promoting macromolecular association. Therefore, the driving force here proposed could have paved the way for several elaborated forms of spatial organisation found in present organisms, and could eventually be a universal trait of protobiological settings. Decoding fundamental rules has been claimed central to study 'early cell ancestors', which actually fill the gap between inanimate and animate matter (Pohorille 2009).

Experimentally assessing physiological effects of protein concentration increase on growth rate would require measurements in systems devoid of non-specific interactions. This poses a problem, because in the highly crowded cytoplasm of present organisms the effects of non-specific interactions are dominant. For example, increasing external osmolarity, cytoplasmic protein concentration of *E. coli* K-12 could be increased by more than 50%, leading to a linear decrease in the growth rate (Cayley et al. 1991). This behaviour, which seems conflicting with the minimal bacterial model described above, might be explained by a disadvantageous effect of non-specific interactions appearing under the new conditions. The effect of protein concentration on growth rate could ultimately be tested in experimental protocellular systems, aimed to capture essential properties of life, such as self-reproduction and homeostasis (Zepik et al. 2001). Perhaps more elaborated systems than those built so far would be required. Still, we believe adequate synthetic systems will be obtained within the challenging and fascinating scientific programme of synthesising life (Szostak et al. 2001; Stano et al. 2009).

References

Acerenza L, Graña M (2006) On the origins of a crowded cytoplasm. J Mol Evol 63:583–590
Bray D (1998) Signaling complexes: biophysical constraints on intracellular communication. Annu Rev Biophys Biomol Struct 27:59–75
Bremer H, Dennis PP (1996) Modulation of chemical composition and other parameters of the cell by growth rate. In: Neidhart FC et al (eds) Escherichia coli and salmonella: cellular and molecular biology. ASM Press, Washington DC
Burg MB (2000) Macromolecular crowding as a cell volume sensor. Cell Physiol Biochem 10:251–256

Cayley S, Lewis BA, Guttman HJ, Jr Record MT (1991) Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity. Implications for protein-DNA interactions in vivo. J Mol Biol 222:281–300

Clegg JS (1984) Properties and metabolism of the aqueous cytoplasm and its bounderies. Am J Physiol 246:R133–R151

Cooper VS, Lenski RE (2000) The population genetics of ecological specialization in evolving Escherichia coli populations. Nature 407:736–739

Donachie WD, Robinson AC (1987) Cell division: parameter values and the process. In: Neidhart FC et al (eds) Escherichia coli and salmonella: cellular and molecular biology. ASM Press, Washington DC

Dyson F (1985) Origins of life. Cambridge University Press, Cambridge

Ellis RJ (2001) Macromolecular crowding: obvious but underappreciated. Trends Biochem Sci 26:597–604

Ellis RJ, Minton AP (2003) Cell biology: join the crowd. Nature 425:27–28

Fulton AB (1982) How crowded is the cytoplasm? Cell 30:345–347

Garner MM, Burg MB (1994) Macromolecular crowding and confinement in cells exposed to hypertonicity. Am J Physiol 266:C877–C892

Goobes R, Kahana N, Cohen O, Minsky A (2003) Metabolic buffering exerted by macromolecular crowding on DNA-DNA interactions: origin and physiological significance. Biochemistry 42:2431–2440

Graña M, Acerenza L (2001) A model combining cell physiology and population genetics to explain Escherichia coli laboratory evolution. BMC Evol Biol 1:12

Guigas G, Kalla C, Weiss M (2007) The degree of macromolecular crowding in the cytoplasm and nucleoplasm of mammalian cells is conserved. FEBS Lett 581:5094–5098

Hall D, Minton AP (2003) Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. Biochim Biophys Acta 1649:127–139

Helmstetter C (1996) Timing of synthetic activities in the cell cycle. In: Neidhart FC et al (eds) Escherichia coli and salmonella: cellular and molecular biology. ASM Press, Washington DC

Kauffman SA (1986) Autocatalytic sets of proteins. J Theor Biol 119:1–24

Kornberg A (2000) Ten commandments: lessons from the enzymology of DNA replication. J Bacteriol 182:3613–3618

Kozer N, Kuttner YY, Haran G, Schreiber G (2007) Protein-protein association in polymer solutions: from dilute to semidilute to concentrated. Biophys J 92:2139–2149

Laurent TC (1971) Enzyme reactions in polymer media. Eur J Biochem 21:498–506

Lazcano A, Miller SL (1999) On the origin of metabolic pathways. J Mol Evol 49:424–431

Lenski RE (2009) Available at http://myxo.css.msu.edu/ecoli/summdata.html

Lenski RE, Mongold JA (2000) Cell size, shape and fitness in evolving populations of bacteria. In: Brown JH, West GB (eds) Scaling in biology. Oxford University Press, London

Lenski RE, Mongold JA, Sniegowski PD, Travisano M, Vasi F, Gerrish PJ, Schmidt TM (1998) Evolution of competitive fitness in experimental populations of E. coli: what makes one genotype a better competitor than another? Antonie Van Leeuwenhoek 73:35–47

Lenski RE, Travisano M (1994) Dynamics of adaptation and diversification: a 10, 000-generation experiment with bacterial populations. Proc Natl Acad Sci U S A 91:6808–6814

Løbner-Olesen A, Skarstad K, Hansen FG, von Meyenburg K, Boye E (1989) The DnaA protein determines the initiation mass of Escherichia coli K-12. Cell 57:881–889

Luisi PL (2002) Towards the engineering of minimal living cells. Anat Rec 268:208–214

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 93:1–13

Luisi PL, Oberholzer T, Lazcano A (2002) The notion of a DNA minimal cell: a general discourse and some guidelines for an experimental approach. Helvetica Chimica Acta 85:1759–1777

Lutkenhaus J, Mukerjee A (1996) Cell division. In: Neidhart FC et al (eds) Escherichia coli and salmonella: cellular and molecular biology. ASM Press, Washington DC

Minton AP (1981) Excluded volume as a determinant of macromolecular structure reactivity. Biopolymers 20:2093–2120

Minton AP (1983) The effect of volume occupancy upon the thermodynamic activity of proteins: some biochemical consequences. Mol Cell Biochem 55:119–140

Minton AP (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J Biol Chem 276:10577–10580

Minton AP, Colclasure GC, Parker JC (1992) Model for the role of macromolecular crowding in regulation of cellular volume. Proc Natl Acad Sci USA 89:10504–10506

Mongold JA, Lenski RE (1996) Experimental rejection of a nonadaptive explanation for increased cell size in Escherichia coli. J Bacteriol 178:5333–5334

Monod J (1949) The growth of bacterial cultures. Annu Rev Microbiol 3:371–394

Narayanaswamy R, Levy M, Tsechansky M, Stovall GM, O'connell JD, Mirrielees J, Ellington AD, Marcotte EM (2009) Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation. Proc Natl Acad Sci USA 106:10147–10152

Oberholzer T, Luisi PL (2002) The use of liposomes for constructing cell models. J Biol Phys 28:733–744

Pohorille A (2009) Early ancestors of existing cells. In: Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF (eds) Protocells. The MIT Press, Cambridge, MA

Rashevsky N (1960) Mathematical biophysics: physico-mathematical foundations of biology, vol 1, 3rd edn. Dover Publications, Mineloa, NY

Jr Record MT, Courtenay ES, Cayley S, Guttman HJ (1998) Biophysical compensation mechanisms buffering E. coli protein-nucleic acid interactions against changing environments. Trends Biochem Sci 23:190–194

Rohwer JM, Postma PW, Kholodenko BN, Westerhoff HV (1998) Implications of macromolecular crowding for signal transduction and metabolite channeling. Proc Natl Acad Sci USA 95:10547–10552

Schaechter M, Maaloe O, Kjeldgaard NO (1958) Dependency on medium and temperature of cell size and chemical composition during balanced grown of Salmonella typhimurium. J Gen Microbiol 19:592–606

Segre D, Ben-Eli D, Lancet D (2000) Compositional genomes: prebiotic information transfer in mutually catalytic noncovalent assemblies. Proc Natl Acad Sci U S A 97:4112–4117

Sole RV, Macia J, Fellermann H, Munteanu A, Sardanyes J, Valverde S (2009) Models of protocell replication. In: Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF (eds) Protocells. The MIT Press, Cambridge, MA

Srere PA (1987) Complexes of sequential metabolic enzymes. Annu Rev Biochem 56:89–124

Srere PA, Ovadi J (1990) Enzyme-enzyme interactions and their metabolic role. FEBS Lett 268:360–364

Stano P, Murtas G, Luisi PL (2009) Semisynthetic minimal cells: new advances and perspectives. In: Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF (eds) Protocells. The MIT Press, Cambridge, MA

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390

van den Berg B, Wain R, Dobson CM, Ellis RJ (2000) Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell. EMBO J 19:3870–3875

Wachtershauser G (1992) Groundworks for an evolutionary biochemistry: the iron-sulphur world. Prog Biophys Mol Biol 58:85–201

Wachtershauser G (2000) Origin of life. Life as we don't know it. Science 289:1307–1308

Walter H, Brooks DE (1995) Phase separation in cytoplasm, due to macromolecular crowding, is the basis for microcompartmentation. FEBS Lett 361:135–139

Winkel BSJ (2004) Metabolic channeling in plants. Annu Rev Plant Biol 55:85–107

Zepik HH, Blochliger E, Luisi PL (2001) A chemical model of homeostasis. Angew Chem Int Ed Engl 40:199–202

Zimmerman SB, Minton AP (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu Rev Biophys Biomol Struct 22:27–65

Zimmerman SB, Murphy LD (1996) Macromolecular crowding and the mandatory condensation of DNA in bacteria. FEBS Lett 390:245–248

Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. J Mol Biol 222:599–620

Chapter 4
How Small is Small?

A Biophysical Chemists Thoughts about the Lower Limits of Cell Sizes

Peter B. Moore

Abstract Arguments based on the physical sizes of genomes indicate that autonomously replicating cells significantly smaller than those of *Mycoplasma genitalium* (radius: ~0.15 μm; volume: ~1 × 10⁻² μm³) are unlikely to exist. Moreover, microorganisms having cell volumes in that range, or below, are certain to be parasites, as *Mycoplasma genitalium* is, because their genomes will necessarily be so small that they cannot include genes for more than a handful of the enzymes required for intermediary metabolism.

4.1 Introduction

In the fall of 1998 a workshop was held under the auspices of the Space Studies Board of the National Research Council, the proceedings of which were published under the title, "Size Limits of Very Small Organisms" (NRC Space Studies Board 1999). A paper had appeared 2 years earlier announcing the discovery of carbonate-rich granules in a meteorite of Martian origin that its authors proposed were biological, and might even be the fossilized remains of microorganisms (McKay et al. 1996). The plausibility of that proposal was the focus of the 1998 workshop, and one of the many reasons for doubting its validity was that no cells that small (diameter: ~0.1 μm) have ever been found on Earth. In effect the participants in the workshop were asked to consider a question that has long intrigued biologists. How small can an organism be? The rise of the discipline of synthetic biology has endowed this question with renewed relevance.

In 1998, I argued that organisms the size of the granules found in the Martian meteorite by McKay and coworkers are physically impossible (Moore 1999). This chapter revisits those arguments in light of recent developments that have raised

P.B. Moore (✉)
Departments of Chemistry, and Molecular Biophysics and Biochemistry,
Yale University, USA
e-mail: peter.moore@yale.edu

doubts about the interpretation of one set of observations that in 1998 appeared to indicate that organisms the size of the meteorite granules in question actually exist on Earth.

4.2 What is an Organism?

Any discussion of the minimum size of organisms must start with a definition for the word "organism". In order for an object/system to qualify as an organism it must have two properties. First, it must be surrounded by a physical boundary, e.g. for unicellular organisms, a cell membrane, that separates it from its environment. Further that boundary must remain intact topologically for its entire life; the object must at all times have a definite inside and a definite outside. Second, the object must be capable of replicating autonomously. Here the adjective "autonomously" means that the object/system must synthesize internally all the molecules required for its replication that have molecular weights above ~1,000, and in addition, the biochemistry required for replication must occur inside its limiting boundary. The reader will note that by this definition, viruses are *not* organisms, a point on which dictionary definitions of the word "organism" are often vague.

4.3 On the Sizes of Extant Bacteria

The smallest organisms on Earth are all unicellular prokaryotes, either eubacteria or archaebacteria. For most biologists *Escherichia coli* is the prototypical prokaryote. It is found in the intestinal tracts of mammals, where the temperature is a comfortable 37°C, and its hosts provide it a rich diet. *E. coli* is capable of living in far more Spartan surroundings. Media consisting of mineral salts and glucose suffice, and it will even grow on similar media in which the carbon source is acetate. Only photosynthetic microorganisms, which can grow on water, salts, CO_2, and light, are able to make do with significantly less biochemical help.

A rapidly growing *E. coli* cell is a cylinder ~2 μm long and ~1 μm in diameter. Its volume is thus about 1.6 μm³, and its mass roughly 1.7×10^{-15} kg. About 70% of its mass is water, which is typical for active cytoplasm. The cell contains about four copies of a DNA genome that has a molecular weight of 3×10^9 per genome. Thus DNA accounts for about 1% of the cell mass (~2×10^{-17} kg). The *E. coli* genome encodes about 4,300 different kinds of proteins (Blattner et al. 1997), although it never makes all of them at the same time, and protein accounts for 15% of its total mass. RNA accounts for another 6% of its cell mass, and most of that is transfer RNA or ribosomal RNA. (Such a cell contains about 30,000 ribosomes). Small organic molecules and ions of various sorts make up the remaining 7% of the non-aqueous mass (see Watson 1965).

Compared to some prokaryotes, *E. coli* is huge. The smallest unicellular organisms identified so far are all mollicutes (Pollack et al. 1997), a group that includes *Mycoplasma genitalium*, which is of particular interest here because the sequence of its genome is known. The linear dimensions of mollicute cells are about a quarter those of *E. coli* cells, which means their interior volumes are about 1% as large. Simple arithmetic shows that cells that small cannot accommodate four copies of a genome the size of *E. coli*'s; they are not big enough. Even a single copy of a genome as large as *E. coli*'s would be too much, assuming that active cytoplasm must be ~70% water. It is not surprising, therefore, that mollicutes have much smaller genomes than *E. coli,* and fewer genome copies per cell. *M. genitalium* cells contain, on average, two copies of a genome about a tenth the size of the *E. coli* genome. Even so, the DNA concentration in these cells is still seven to eight times higher than it is in *E. coli* (see Table 4.1).

Compared to *E. coli, M. genitalium* is biochemically crippled. Its genome encodes about a tenth as many proteins as the genome of *E. coli* (4,288 proteins (Blattner et al. 1997) versus 471 proteins (Fraser et al. 1995)), and among the genes it lacks are most of the genes for the enzymes required for intermediary metabolism. For that reason *M. genitalium* can only grow in environments in which most of the small organic molecules that *E. coli* is capable of making for itself are already present. That is why mollicutes are obligatory parasites.

Clearly, cells having volumes 1% that of *E. coli* cells are feasible. Are there any unicellular organisms with cell volumes 0.01% that of *E. coli*? As Table 4.1 indicates, there are genome-containing, biological objects of that size in the biological world: for example, the head of bacteriophage T4, which is the part of that virus's structure that contains its genome. T4 is one of the more complicated viruses that infect *E. coli*, and by viral standards, its genome is big. It is a quarter the size of the genome of *M. genitalium*, but instead of encoding the ~120 genes you might anticipate (=470/4), it encodes about 300 (Miller et al. 2003). Thus the gene density is very high in the T4 genome, as it is in other viral genomes.

What is important here about the T4 head is its interior DNA concentration, which is ~0.69 g/ml. The partial specific volume of DNA is about 0.54 ml/g, which implies that if DNA could be packed so that were no interstices between strands,

Table 4.1 DNA concentration as a function of organism size

Organism	Radius (μm)	Volume (μm)³	Genome MW	No. of copies/cell	gDNA/ml
E. coli	0.73	1.6	3×10^9	4	0.013
M. genitalium	0.15	1.3×10^{-2}	4×10^8	2	0.100
T4 head	0.04	2.4×10^{-4}	1×10^8	1	0.690

Data for *E. coli* from (Watson 1965) (Data for *M. genitalium* from Morowitz 1992; Data for T4 from (Stryer 1981). No allowance was made for the contribution made by membrane/cell wall thickness to the size of *E. coli* when computing its interior volume. For *M. genitalium*, interior volume was computed allowing 0.005 μm for its membrane thickness, and the volume of the T4 head provided by Stryer is its internal volume. For the sake of argument, all of the objects considered are assumed to be spheres.

which it cannot, the DNA concentration of in the packed mass that resulted would be ~1.85 g/ml. Thus the (double-stranded) DNA in the head of bacteriophage T4 must be almost close-packed.

The DNA in T4 heads, like the genomes in all viruses, is metabolically inert, and it is meant to be so. Viruses don't do anything; they are simply genome transport modules. Not only that, it is impossible for the DNA in T4 heads to be replicated or expressed in situ because there isn't room for the macromolecules that would be needed to make it happen, let alone room to accommodate the two copies of the genome that would have to be resident there at the end of a single round of DNA replication. T4 DNA becomes biologically active only when injected into bacterial cells, where the total DNA concentration is much lower, and the host-encoded, host-synthesized macromolecules the virus requires for its replication are available. T4 is incapable of replicating autonomously; it is even more parasitic than *M. genitalium*. I take it as axiomatic that DNA packed as densely as it is in bacteriophages and other viruses will always be inert metabolically. That being so, organisms having cellular volume 1% those of *M. genitalium* cells are impossible, and cells having volumes 10% that size look improbable.

The mollicutes are not the only prokaryotes that have small cell sizes. Marine ultramicrobacteria are also unusually small. However, their "design specifications" are consistent with the arguments made above. *Sphinogomonas sp.,* which is a typical member of that group, has a cell volume about four times larger than that of *M. genitalium,* and a genome that is twice as big (Schut et al. 1997). A genome that large can encode enough enzymes to carry out the intermediary metabolism those organisms require, and their intracellular DNA concentrations are reasonable.

At the time of the 1998 workshop, the objects that were most difficult to accommodate were the so-called "nanobacteria", which were believed to be free-living organisms capable of autonomous replication (Kajander et al. 1997). Objects of that class are rather variable in size, but some are reported to have *outside* radii as small as ~0.05 μm, which is of the same order as the *inside* radius of the head of bacteriophage T4!

In 1998 it was very hard to understand how anything as small as a nanobacterium could be an organism. It has been argued that autonomously replicating organisms need genomes that encode at least 250 proteins (Mushegian and Koonin 1996), which implies that they must have genomes no smaller than that of T4. However, if a double-stranded DNA that large were stuffed into a cell that has an outside diameter of 0.05 μm, which may not be physically possible given that some of volume of such a cell is taken up by its limiting membrane, there would surely be no room for anything else. The resulting object would have to be a metabolically inert, virus-like entity.

A lot of work has been done on nanobacteria over the last decade, and it now seems unlikely that nanobacteria are organisms (Martel and Young 2008). It appears probable that they are, in fact, inanimate calcium carbonate granules. One cannot help wondering if the processes that give rise to "nanobacteria" might not be related to those that produced the objects found in the Martian meteorite that precipitated the NRC meeting held in 1998.

4.4   Expedients for Reducing Cell Size

It is clear that the physical size of DNA molecules large enough to encode a plausible set of proteins sets a lower bound on cell sizes. One way you could think of alleviating that constraint would be to use an information storage polymer that is less massive than DNA, per unit of information stored. A base pair in a DNA duplex has a molecular weight of 600–650, and it encodes two bits of sequence information. Certainly one could think of using single-stranded DNA, rather than a double-stranded nucleic acid, for the genome of a cell. Some viruses have genomes that are single-stranded nucleic acids. The only obvious downside is that single-stranded genomes are more vulnerable to mutation than double-stranded genomes. This expedient would reduce the molecular weight per bit by a factor of 2, i.e. from ~300 to ~150.

What about replacing nucleic acid genomes with genomes that have completely different, but more economical chemical structures? In my estimation the benefits to be gained by a change in polymer chemistry are unlikely to be large enough to make a real difference. Nucleotides may look bigger than they need to be, but in fact, Watson-Crick base pairs are remarkably efficient devices for storing sequence information. The reason they work so well is not that Watson-Crick pairs are intrinsically superior to anything else in some deep, fundamental way, but that Watson-Crick pairs are easily distinguished from non-Watson-Crick base pairs chemically. Thus an information storage system based on Watson-Crick pairs is "good" because it can be accurately read out and replicated. It would be a challenge to come up with a storage/read-out system based on an alternative chemistry that works as efficiently, and even if such a system were devised, it would be astonishing if the polymer it used had a molecular weight per bit of encoded information ratio much lower than that of single-stranded nucleic acid.

Given that one is "stuck" with DNA genomes, what can be done to reduce genome bulk beyond using single-stranded DNA? Experiments done in the 1970s suggest that modern proteins may be larger than their functions absolutely require. For the sake of argument, assume that the genes in our putative nanocell genome are half the size of the genes in the genome of T4. If this could be done, our organism might be able to survive with a single-stranded DNA genome having a molecular weight of about 0.25×10^8 that encodes about 300 proteins, each of which is roughly half the molecular weight of the proteins that perform the corresponding functions in other organisms. To make things even more favorable, we will assume that only a partial copy of the Watson-Crick complement of its genome is generated as the cell grows, and that growth is so slow that the average number of copies of the genome present per cell is 1.0. Furthermore, we will allow the intracellular concentration of DNA to increase to 0.2 g/ml, which is twice (!) that found in *M. genitalium.* The estimate for interior cell volume that emerges for this nanocell is ~10^{-4} μm^3, which is similar to the interior volume of the T4 head, and in the range characteristic of nanobacteria, and the granules in the Martian meteorite.

Several comments are in order about the construct envisioned above.

1. In order to arrive at a "design" for this ~0.05 μm radius cell, a biochemistry has been invoked that is unlike that found in any modern life form.
2. Whatever else this nanocell can do, it is not going to grow on a medium consisting of mineral salts and glucose. Something else, which is far larger and more complicated, will have to do the biochemical "heavy lifting" that makes the survival of such nanocells possible. To put it another way, a biosphere consisting entirely of organisms of the size and metabolic complexity of *E. coli is* easy to envision; it may once have existed on this planet (Schopf 1983). A biosphere consisting of organisms the size and complexity of *M. genitalium* looks impossible, and thus a biosphere based on our putative nanocells is out of the question. It follows that those looking for evidence of life on Mars (or elsewhere) should seek fossils of microorganisms that have cells the size of *E. coli*, or larger because if there ever was life on that planet, it will have depended on organisms at least that complicated.
3. All reports of "cells" significantly smaller than *M. genitalium* should be treated with skepticism.
4. Wise biologists never say "never." Our knowledge of life on Earth is still fragmentary. Maybe there really are cells out there as small as the nanocell just described. If such things exist, their biochemistry will surely be unlike that of any organism we know today, and their discoverers will have the privilege of exploring an entirely new world.

References

Blattner FR, Plunkett G, Bloch CA et al (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1474

Fraser CM, Gocayne JD, White O et al (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270:397–403

Kajander EO, Kuronen I, Akerman K, et al (1997) Nanobacteria from blood: the smallest culturable autonomously replicating agent on Earth. Proceedings of SPIE 3111:420–428.

Martel J, Young JD-E (2008) Purported nanobacteria in human blood as calcium carbonate nanoparticles. Proc Natl Acad Sci USA 105:5549–5554

McKay DS, Gibson EK Jr, Thomas-Keprta KL et al (1996) Search for life on Mars: possible relic of biogenic activity in Martian meteorite ALH84001. Science 273:924–930

Miller ES, Kutter E, Mosig G et al (2003) Bacteriophage T4 genome. Microbiol Mol Biol Rev 67:86–156

Moore PB (1999) A biophysical chemist's thoughts on cell size. Size limits of very small microorganisms: proceedings of a workshop, pp 16–20. Steering Group of Astrobiology of the Space Studies Board, National Research Council, Washington, DC

Morowitz HJ (1992) Beginnings of cellular life. Yale University Press, New Haven

Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci USA 93:10268–10273

NRC Space Studies Board (1999) Size limits of very small mircoorganisms. National Academy Press, Washington, DC

Pollack JD, Williams MV, McElhaney RN (1997) The comparative metabolism of the mollicutes (Mycoplasmas): the utility for taxonomic classification and the relationship of putative gene annotation and phylogeny to enzymatic function in the smallest free-living cells. Crit Rev Microbiol 23:269–354

Schopf JW (ed) (1983) Earth's earliest biosphere: its origins and evolution. Princeton University Press, Princeton, NJ

Schut F, Gottshal JC, Prins RA (1997) Isolation and characterization of the marine ultramicro-bacterium Sphingomonas sp. srain RB2256. FEMS Microbiol Rev 20:363–369

Stryer L (1981) Biochemistry, 2nd edn. W.H. Freeman, San Francisco, CA

Watson JD (1965) Molecular biology of the gene. W.A. Benjamin, New York

Chapter 5
Biochemical Reactions in the Crowded and Confined Physiological Environment: Physical Chemistry Meets Synthetic Biology

Allen P. Minton and Germán Rivas

FOREWORD

Proteins and nucleic acids constitute at least 20–30% of the total mass (and volume) of all living organisms without exception. Although local composition may vary widely with location within a given cell and between cells, it is evident that much of the chemistry of life – as opposed to laboratory biochemistry – takes place within media containing a substantial volume fraction of macromolecules. These media are termed "crowded" or "volume-occupied", rather than "concentrated", as no single macromolecular species need be concentrated. Moreover, many biological compartments do not consist of a continuous fluid phase, but rather a series of small interstitial elements of fluid, or "pores", bounded by membranes or other relatively immobile structural elements such as cytoskeletal filaments. Such interstitial volume elements might be likened to the holes in a sponge, except that the characteristic sizes of the "holes" are of the order of tens of nanometers. The soluble macromolecules within these pores are termed "confined" to reflect the discontinuous nature of the fluid phase and the small dimensions of the pores.

During the last decades it has gradually become recognized that the reactivity of an individual macromolecular probe species in crowded and confined media may be qualitatively different from that of the same species in a medium that is dilute in all macromolecular solutes. The volume excluded by macromolecules to each other in crowded and confined media has been predicted to have major qualitative effects upon a broad range of biochemical, biophysical, and physiological phenomena, including – but not limited to – protein stability and folding, macromolecular

A.P. Minton (✉)
Section on Physical Biochemistry, Laboratory of Biochemistry and Genetics,
NIDDK, National Institutes of Health, Bethesda, MD
e-mail: minton@helix.nig.gov

G. Rivas
Chemical and Physical Biology Program, Centro de Investigaciones Biológicas,
CSIC, Madrid, Spain
e-mail: grivas@cib.csic.es

association equilibria and kinetics, enzymatic activity, formation of supramolecular assemblies, genome structure and function, as well as phase separation phenomena and compartmentation inside cells and artificial vesicles. Many of these effects have been observed experimentally (Zhou et al. 2008).

Although cells are likely to contain a variety of localized microenvironments that are more or less crowded/confined than suggested by cell-average measurements, it is difficult to envisage any microenvironment within a living cell that is totally free of excluded volume effects, since these effects are independent of the chemical composition, detailed structure, and specific interactions of the particular macromolecules within the microenvironment. Excluded volume effects depend only upon gross molecular size and shape, and the fractional volume occupancy of the most abundant species. It is our thesis that in order to narrow the gap between laboratory biochemistry and the chemistry of living organisms, it is necessary to determine experimentally and to understand theoretically the extent to which crowding and confinement influence the rates and equilibria of macromolecular reactions involved in biological processes.

Crowding and confinement must also be considered in the emerging areas of synthetic biology and the developing minimal cell systems, in which quantitative global models of intracellular reactions are needed. In order to attain or optimize functionality, the reconstitution in vitro of dynamic multi-protein complexes essential to critical biochemical processes will have to be performed under controlled physicochemical conditions that mimic the physiological environment in which these complexes have evolved to function in vivo.

In the article "How can biochemical reactions within cells differ from those in test tubes?" (Minton 2006), that follows this short comment, it is discussed how macromolecular crowding, macromolecular confinement, and macromolecular adsorption affect reaction rates and equilibria, and the consequent effects on cellular biochemistry. The aim of this chapter is therefore to introduce the reader to the world of crowded and confined macromolecular systems; we are confident that such considerations will have a prominent role in the next developments of the complex cell-mimicking molecular systems.

HOW CAN BIOCHEMICAL REACTIONS WITHIN CELLS DIFFER FROM THOSE IN TEST TUBES?[1]

5.1 Introduction

Although the time when a living cell was regarded naively as a 'bag of enzymes' is long past, the extent to which the high internal concentration of macromolecules and the constraints of cellular architecture (see Fig. 5.1) can influence intracellular

---

[1] Originally published as: Minton 2006. Reproduced with permission
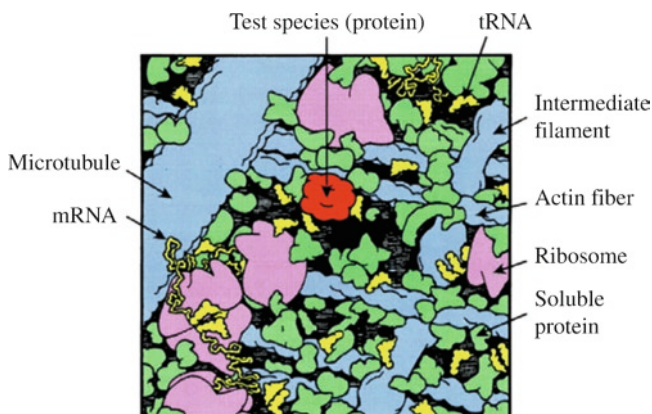
Fig. 5.1 Cartoon of eukaryotic cytoplasm magnified ×10^6 (Reproduced from Minton (Minton 2001a) and modified from Goodsell (Goodsell 1993)

biochemical reactions is still not generally appreciated. In this Commentary, I briefly describe three mechanisms by which features of the cellular interior can significantly alter, in a largely nonspecific manner, the equilibria and rates of biochemical reactions relative to the equilibria and rates of the same reactions in dilute solution. The three mechanisms correspond to three different types of interactions between a particular macromolecular reactant (X) and constituents of the local environment (the background).

5.2 Types of Background and Background Interactions

5.2.1 Macromolecular Crowding

In some intracellular compartments, X may be surrounded by a variety of other soluble macromolecules that cumulatively occupy a significant fraction of the total volume (Fulton 1982). Since no single macromolecular species need be present at high concentration, such media are referred to as crowded or volume occupied rather than concentrated (Minton and Wilf 1981). Since macromolecules cannot interpenetrate, the fraction of volume into which a macromolecule can be placed at any instant is much less than the fraction of volume into which a small molecule can be placed (Fig. 5.2a, b). The total free energy of interaction between X and all of the other molecules in the crowded medium is inversely related to the probability of successful placement of X at a random location within the crowded medium (Lebowitz et al. 1965). Hence the extra work required to transfer X from a dilute to a crowded solution resulting from steric repulsion between X and background
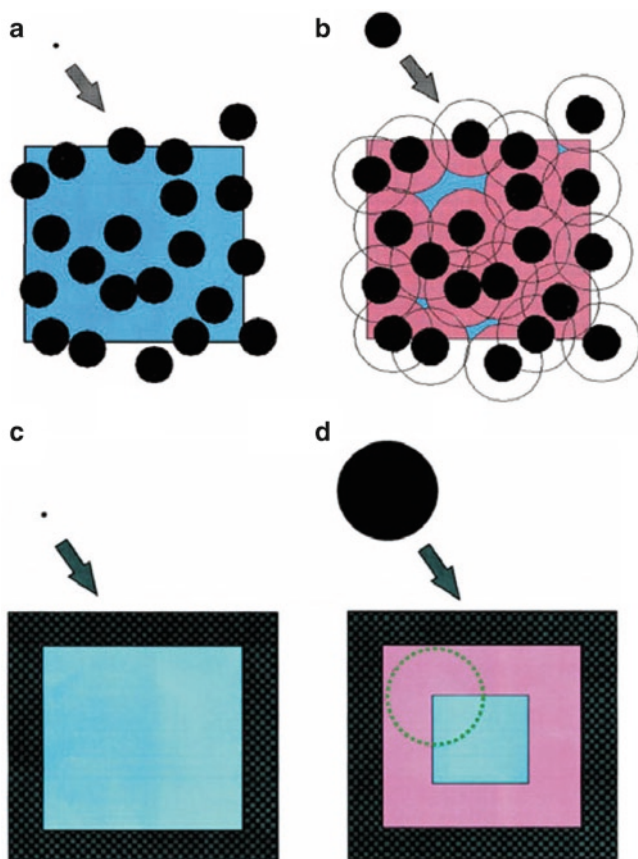
Fig. 5.2 Schematic illustration of volume available (*blue*) and volume excluded (*pink* and *black*) to the center of a (**a**, **c**) small and (**b**, **d**) large spherical molecule added to a solution containing a approximately 30% volume fraction of large spherical molecules (**a**, **b**) and the interior of a pore shown in (*square*) cross-section (**c**, **d**) (Reproduced from Minton (Minton 2001a)

molecules depends upon the relative sizes and shapes of X and background species, and increases in a highly nonlinear fashion with increasing size of X and with fractional volume occupancy (Minton 1998).

5.2.2 Macromolecular Confinement

In eukaryotic cells, much of the fluid phase of cytoplasm exists within interstices between a variety of large fibrous and membranous structures (Knull and Minton 1996). It has been estimated that between 10% and 100% of the fluid phase of the cytoplasm lies within one macromolecular diameter of the surface of one of

these structures (Minton 1989). By analogy with size exclusion columns, we may generically treat these interstitial elements of fluid volume as 'pores'. When the size of a pore is not much larger than that of an enclosed macromolecule reactant X, steric-repulsive interactions between X and the pore boundaries result in a reduction of the volume available to X (Fig. 5.2c,d). Hence work (free energy) is required to transfer a molecule of X from an element of unbounded solution into a pore of equal volume (Giddings et al. 1968). The magnitude of excess work depends strongly upon the relative sizes and shapes of both the confined macromolecule and the pore (Giddings et al. 1968; Minton 1992).

5.2.3   Macromolecular Adsorption

When X bears a net charge opposite to that of the surface of a nearby fiber or membrane, X may be reversibly non-specifically adsorbed onto the surface (Cutsforth et al. 1989; Knull and Walsh 1992; Lakatos and Minton 1991). Similarly, if X is a post-translationally modified protein bearing a lipid side chain, it may also reversibly and nonspecifically adsorb onto lipid bilayer membranes (Arbuzova et al. 1998). When surface adsorption is spontaneous, the associated free energy change is negative, but its magnitude depends on entropic factors that vary nonspecifically with the size and shape of X (Minton 1995).

5.2.4   Influence of Background Interactions upon Reaction Equilibria and Rates

Figure 5.3 illustrates how nonspecific interactions between reactants and the background can influence the rate and/or equilibrium of a particular reaction – for example the association of two globular proteins, A and B, to form a heterodimer. In this reaction, the 'reactants' are A and B separated by a distance sufficiently large that they do not interact, and the 'product' is the heterodimer. In the 'transition state', A and B are close to each other and orientated such that they are poised to form the heterodimer, but they are still not bound to each other by short-range attractive interactions.

The free energy profile of a reaction taking place in bulk solvent may be altered when the same reaction takes place within a medium such as a cellular interior that is filled with other macromolecules, even if these macromolecules are nominally inert (i.e. they do not participate directly in the reaction). Individual pairwise interactions between a single molecule within the reaction and a molecule from the background need be neither strong nor specific. The density of macromolecules within cells is sufficiently high (~500 g/l in some compartments) that the sum of a large number of weak, nonspecific interactions can contribute substantially to the standard free energy of each state of the reaction system.
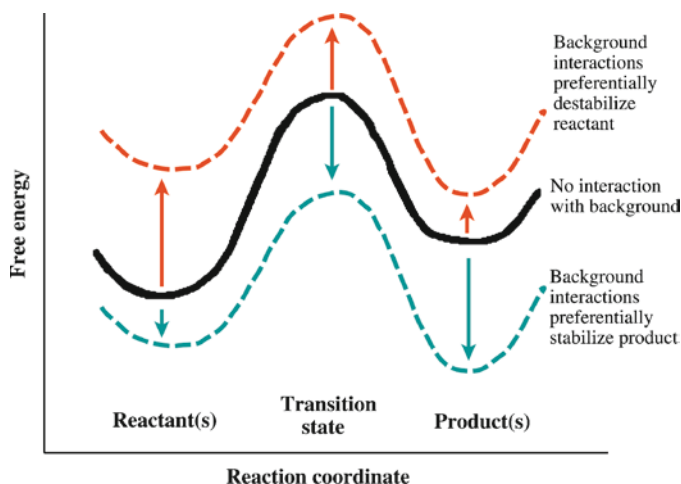
Fig. 5.3 Free energy profiles of a chemical reaction in bulk solution (*black curve*), in the presence of stabilizing background interactions (*green*), and in the presence of destabilizing background interactions (*red*)

The free energy profile in green in Fig. 5.3 has been shifted downwards because all three states of the system interact with the background in an attractive (free energy lowering) fashion. Such nonspecific intermolecular attraction can be due to weak electrostatic or hydrophobic effects and often results in the formation of weak, nonspecific complexes (e.g. the interaction of proteins with urea). The important feature of this profile is that the relative free energies of the three states of the system have also been altered: the background interactions stabilize the transition state and products more than they do the reactants. These background interactions should therefore push the equilibrium state towards product formation, primarily by enhancing the forward reaction rate.

By contrast, the free energy profile in red in Fig. 5.3 has been shifted upward because all three states of the system interact with the background in a repulsive (free energy raising) fashion. Nonspecific intermolecular repulsion can be due to volume exclusion (steric repulsion) or electrostatic effects and does not lead to formation of complexes between the mutually repelling species. Nevertheless, because the repulsive interactions in this example destabilize the reactant state more than they do the transition and product states, the overall effect upon the relative free energies of the three states is identical to that above. The equilibrium is shifted towards product formation, because of a preferential increase in the forward reaction rate. Note that even though the mechanisms underlying these two perturbations of the free energy profile are different, one cannot distinguish between them solely by measuring changes in reaction equilibria or kinetics.

There are many other combinations of repulsive and attractive background interactions that can lead to the same relative shifts in the free energies of the three reaction states and hence the same changes in reaction rates and equilibria. By considering the molecular composition and environmental variables characterizing

a particular reaction system, one may discern whether dominant background interactions are likely to be attractive or repulsive. But in a medium as complex and heterogeneous as the cytoplasm for example, the task of identifying these and their influence on any specific reaction is extremely challenging.

5.3   A Common Energetic Formalism

We may analyze the effects of all background interactions on a particular reaction in the context of a uniform thermodynamic picture. Consider, for example, the dimerization reaction introduced above. The apparent equilibrium constant[2] for association, $K_{AB} \equiv c_{AB}/c_A c_B$, is related to the standard free energy of association, ΔF_{AB}, by the thermodynamic relationship:

$$\Delta F_{AB} = -RT \ln K_{AB} \tag{5.1}$$

where R denotes the molar gas constant and T the absolute temperature.[3] We may construct thermodynamic cycles that relate the free energies of association in the absence and presence of background interactions (Fig. 5.4), because free energy is conserved around the cycle:

$$\Delta F_{AB} - \Delta F_{AB}^0 = \Delta F_{AB} - \left(\Delta F_{I,A} + \Delta F_{I,B} \right) \tag{5.2}$$

where ΔF^0_{AB} denotes the standard free energy of association of A and B in the absence of background interaction and $\Delta F_{I,X}$ denotes the standard free energy of interaction between X and the background. It follows from Eqs. (5.1) and (5.2) that:

$$\frac{K_{AB}}{K_{AB}^0} = \exp\left(\frac{\Delta F_{I,A} + \Delta F_{I,B} - \Delta F_{I,AB}}{RT} \right) \tag{5.3}$$

It is evident that background interactions will increase the apparent equilibrium constant for association when the sum of free energies of the individual background interactions of A and B is more positive than the free energy of background interactions of the heterodimer AB. This result is independent of whether the absolute value of any individual $\Delta F_{I,X}$ is positive or negative (see Fig. 5.3).

Note that, although Eqs (5.1, 5.2 and 5.3) are entirely general, each type of background interaction results in a distinctly different dependence of $\Delta F_{I,X}$ upon the

---

[2] The apparent equilibrium constant, defined as a function of the concentrations of solute species, is distinguished from the thermodynamic equilibrium constant, which is defined as a function of the thermodynamic activities of solute species. The relations presented in this section describe the effect of background interactions upon the apparent equilibrium constant for the selected reaction

[3] Free energy changes denoted by ΔF may refer to either Gibbs or Helmholtz free energies since the relationships described here hold equally in constant pressure and constant volume systems
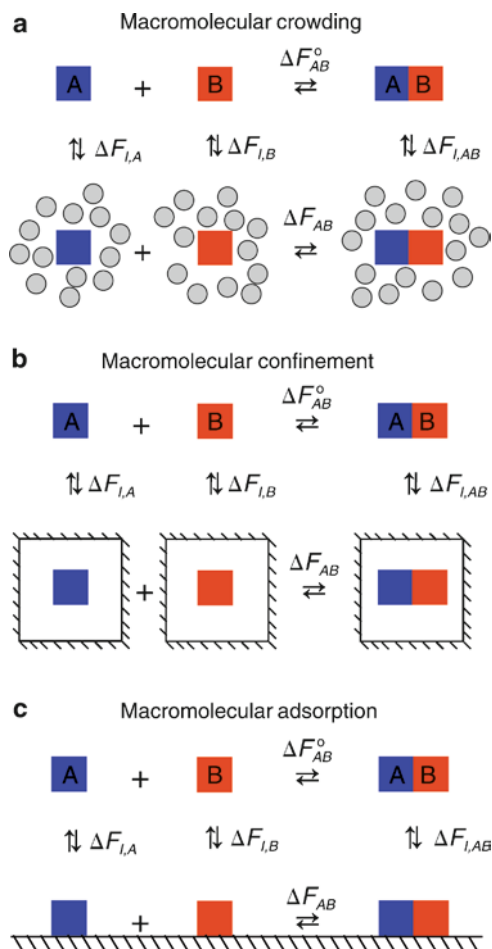
Fig. 5.4 Schematic illustration of how a single set of free energy relationships comprising a thermodynamic cycle may be used to analyze the influence of (**a**) macromolecular crowding, (**b**) macromolecular confinement, and (**c**) macromolecular adsorption upon macromolecular association equilibria

size and shape of X and the sizes, shapes and fractional volume occupancies of those background constituents with which X interacts.

5.4   Predictions and Observations

Approximate statistical-thermodynamic models of the effects of the background interactions upon macromolecular reactions have been developed (e.g. Minton 1981; Minton 1992; Minton 1995; Zhou and Dill 2001; Zhou 2004). These are

based upon mesoscopic or coarse-grained models that take into account the molecular nature of the individual reactant species, but adopt simplified representations of the molecules and the interactions between them. Such models cannot account quantitatively for the behavior of macromolecules in environments as complex and intrinsically variable as the cytoplasm or nucleoplasm. However, they successfully predict large effects of 'inert' background substances upon a variety of reaction rates and equilibria in vitro (see Tables 5.1, 5.2, and 5.3). In favorable cases, the magnitude of the observed effects agrees quantitatively with theoretical prediction. Data demonstrating the effects of crowding are shown in Fig. 5.5. Each of the effects was predicted qualitatively by one of the theories cited above.

5.5 Relevance to Cell Biology

Most predicted effects of macromolecular crowding, confinement, and adsorption on macromolecular isomerization equilibria and association equilibria and rates have been confirmed qualitatively (and in favorable cases quantitatively) in vitro,[4] but are such effects similarly important in a living cell? On the basis of results obtained from partitioning and size-exclusion experiments conducted with concentrated cell lysates, Zimmerman and Trach estimated that excluded volume effects in the cytoplasm of *E. coli* are comparable to those obtained in a 35% solution of a ~70 kDa globular protein, such as bovine serum albumin or hemoglobin (Zimmerman and Trach 1991). Although this estimate might provide a useful starting point, it is clear that the answer is far more complex and elusive. Any cell is extremely large relative to any particular macromolecule of interest and is likely to contain several micro-environments, within each of which a particular macromolecule X will be subject to a different set of background interactions. For example, the cytoplasm of even an organism as simple as *E. coli* contains at least three such micro-environments: the immediate vicinity of the inner plasma membrane, within which X will encounter a high local concentration of membrane phospholipids and proteins; the interior and immediate vicinity of the nucleoid, within which X will encounter an extremely high local concentration of DNA; and the remaining cytoplasm, within which X will be subject primarily to the influence of other soluble proteins. One would expect the relative contributions of macromolecular crowding, confinement and adsorption to macromolecular reactivity to be rather different within each of these three micro-compartments. That complexity acknowledged, it is nevertheless true that large and unambiguous background effects have indeed been observed within certain specialized, simplified cells.

The interior of an erythrocyte is essentially a highly concentrated solution of hemoglobin. The affinity of erythrocytes containing sickle cell hemoglobin (HbS) for oxygen depends significantly upon the intracellular hemoglobin concentration,

---

[4] Predicted effects of macromolecular confinement on association equilibria have not yet been tested

Table 5.1 Predicted and observed effects of macromolecular crowding by high concentrations of inert macromolecules

Predicted effect	Relevant observations[a]
Enhancement of equilibrium association of dilute macromolecules (Minton 1981; Nichol et al. 1981)	Large enhancement in the extent of sickle hemoglobin polymerization (Bookchin et al. 1999); large increase in affinity of DNA-binding proteins for DNA (Zimmerman and Harrison 1987; Jarvis et al. 1990); enhanced self-association of spectrin (Lindner and Ralston 1995), actin (Lindner and Ralston 1997), fibrinogen (Rivas et al. 1999), tubulin (Rivas et al. 1999) and FtsZ (Rivas et al. 2001)
Acceleration of slow (transition state limited) protein associations (Minton 1983; Minton 2001a)	Large increases in the rate of fiber formation by sickle cell hemoglobin (Rotter et al. 2005), actin (Drenckhahn and Pollard 1986), tubulin (Herzog and Weber 1978) and fibrin (Wilf et al. 1985). Large increases in the rate of amyloid formation by apoCII (Hatters et al. 2002) and alpha-synuclein (Shtilerman et al. 2002; Uversky et al. 2002). Large increases in the rate of self-assembly of HIV capsid protein (del Alamo et al. 2005)
Deceleration of rapid (diffusion limited) protein associations (Zimmerman and Minton 1993)	Reduction in rate of diffusion-limited association of TEM and BLIP (Kozer and Schreiber 2004)
Stabilization of proteins against denaturation by heat or chaotropes (Cheung et al. 2005; Minton 2000a)	Elevated temperatures for half-denaturation of actin (Tellam et al. 1983) and lysozyme (Sasahara et al. 2003). Elevated urea concentration for half-denaturation of FK-506-binding protein (Spencer et al. 2005). Partial restoration of catalytic activity of ribonuclease in the presence of high urea concentration (Tokuriki et al. 2004)
Acceleration of protein refolding to the native state (Cheung et al. 2005)	
Enhancement of aggregation of proteins that are partially or fully denatured (Hall and Minton 2002; Hall and Minton 2004)	Proteins that refold spontaneously in dilute solution aggregate in crowded solution and require chaperones to refold (Martin 2002; van den Berg et al. 1999)

[a]Additional relevant observations reported prior to 2003 have been compiled in Minton 1983; Zimmerman and Minton 1993; Minton 2001a; Hall and Minton 2003

because oxygen binding is linked to the polymerization of HbS (May and Huehns 1975). The concentration dependence of oxygen affinity may be quantitatively accounted for if, and only if, steric repulsion between hemoglobin molecules is taken into account (Minton 1976). Ferrone has recently reviewed a large body of work demonstrating that the extensively characterized kinetics of cell sickling subsequent to deoxygenation may be well accounted for by models that take into account the substantial effect of macromolecular crowding within erythrocytes upon the thermodynamic activity (effective concentration) of monomers and each oligomer in the hemolyzate (Ferrone 2004). By means of cleverly designed

Table 5.2 Predicted and observed effects of macromolecular confinement

Predicted effect	Relevant observations
Enhancement or inhibition of the association of confined macromolecules, depending upon complementarity of the shapes of oligomers and the confining cavity[a] (Minton 1992)	
Stabilization of native proteins against unfolding by heat or chaotropes. Magnitude of effect will increase as the size of the protein approaches the smallest dimension of the cavity (Zhou and Dill 2001; Klimov et al. 2002)	Encapsulation of a variety of proteins in hydrated silica-sol-gel glasses(Eggers and Valentine 2001) or in polyacrylamide gels (Bolis et al. 2004) increases the temperature of thermal denaturation. Encapsulation ofribonuclease T1 in reverse micelles increases temperature of thermal denaturation, and the denaturation temperature increases as the water content (volume?) of the micelle decreases (Shastry and Eftink, 1996)
Acceleration of refolding to the native state (Klimov et al. 2002)	

[a]For example, a rod-like oligomer will fit more easily into a tubular cavity than a large spherical oligomer of the same stoichiometry (total volume)

Table 5.3 Predicted and observed effects of macromolecular adsorption

Predicted effect	Relevant observations
Enhancement of equilibrium self- and/ or hetero-association of adsorbed macromolecules (Minton 1995)	Proteins that do not self-associate in solution form 2D crystals when nonspecifically adsorbed to surfaces (Darst et al. 1988)
Association of adsorbed macromolecules enhances adsorption capacity and may result in cooperative equilibrium adsorption isotherms (Chatelier and Minton 1996; Minton 2000b)	Equilibrium adsorption isotherms of proteins onto various types of surfaces exhibit positive cooperativity (Blanco et al. 1989; Cutsforth et al. 1989)
Association of adsorbed macromolecules can accelerate the maximum rate of adsorption and may result in cooperative adsorption progress profiles (Minton 2001b)	Time-dependent adsorption of several proteins onto supported lipid bilayers and other surfaces exhibits a rate that increases with increasing surface occupancy (Fernandez and Berry 2003; Nygren and Stenberg 1990; Ramsden et al. 1994)
Adsorption can destabilize the solution conformation of a protein relative to alternate conformations that interact more strongly with the adsorbing surface (Minton 1995; Cheung and Thirumalai 2006)	Surface denaturation is a widely recognized obstacle to the purification and preparation of proteins at very low concentration (e.g. Edwards and Huber 1992)

experiments, volume regulatory mechanisms in dog erythrocytes were shown to respond nonspecifically to changes in the intracellular concentration of macromolecules rather than changes in volume per se (Colclasure and Parker 1992).
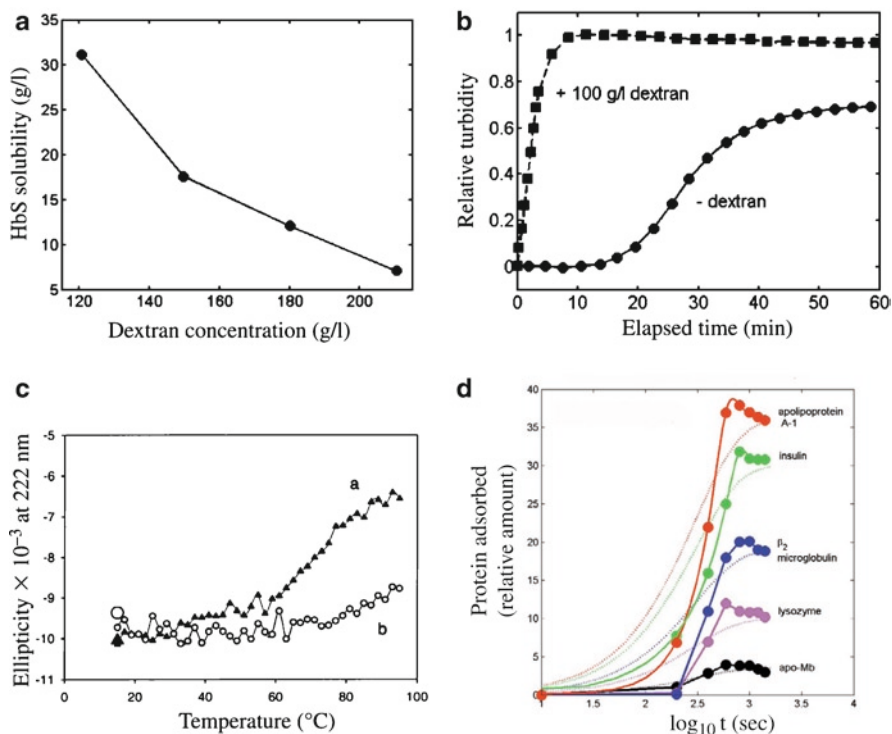
Fig. 5.5 (**a**) Dependence of the solubility of deoxy HbS upon dextran concentration (Bookchin et al. 1999). A 1.75-fold increase in dextran concentration results in a greater than fourfold increase in the equilibrium tendency of deoxy HbS to polymerize. (**b**) Time course of assembly of HIV capsid protein in the absence (●) and presence (■) of 100 g/l Ficoll 70 (del Alamo et al. 2005). The half-time for assembly of the protein is decreased ten-fold in the presence of Ficoll. (**c**) Temperature dependence of the ellipticity of alpha-lactalbumin in bulk solution (▲) and encapsulated in hydrated silica sol-gel glass (O). The confined protein behaves normally at low temperature, but exhibits only partial unfolding even at temperatures approaching 100°C (Reproduced with permission from Eggers and Valentine (Eggers and Valentine 2001)). (**d**) Time dependence of adsorption of several unrelated proteins to a supported phospholipid bilayer (*symbols* and *solid curves*) (Fernandez and Berry 2003). The dotted curves indicate the time dependence that would be expected in the absence of interaction between adsorbed proteins. The enhanced steepness of the experimentally observed curves (*solid*) relative to the reference curves (*dotted*) indicates self-association resulting in clustering of the adsorbed proteins (Minton 2001b)

The interior of an eye lens cell consists essentially of a solution of a small number of crystallins at very high concentration (total protein concentration >500 g/l). These proteins are not translated postnatally. Thus, the crystallin molecules that an animal is born with must remain structurally intact over much of its lifetime to preserve lens transparency. The thermal stability of these proteins increases with concentration (Steadman et al. 1989), and the extraordinary thermal stability of an intact lens has been attributed in part to the stabilizing effects of macromolecular crowding inside the lens cell (Bloemendal et al. 2004).

Excluded volume theory predicts that at the high levels of fractional occupancy characterizing almost all biological fluid media, even small changes in cellular hydration will result in disproportionately large changes in the reactivity of a broad spectrum of macromolecular reactants. This prediction is consistent with and may account for the following general observations: (1) Relatively modest changes of cellular volume in animal cells are associated with changes in a wide variety of diverse intracellular processes, such as the polymerization/depolymerization of cytoskeletal filaments or the activation/deactivation of membrane ion channels, that are much too large to be accounted for on the basis of simple mass action (reviewed by Lang et al. 1998). (2) Complex and very diverse systems for maintaining the concentrations of cellular contents within narrow limits have evolved within all life forms, be they bacterial, plant or animal (Somero et al. 1992). (3) The age-related onset of a variety of protein-aggregation-linked diseases (Koo et al. 1999) correlates with significant loss of cellular and tissue hydration in the elderly and concomitant increases in the volume excluded to protein (Parameswaran et al. 1995; Hatters et al. 2002).

During the last few years several experimental techniques have been developed that enable certain aspects of the behavior of selected macromolecules – typically labeled and/or overexpressed – to be monitored within living cells (e.g. Ghaemmaghami and Oas 2001; Ignatova and Gierasch 2004; McNulty et al. 2006). It is hoped that techniques like these, or others yet to be devised, can be used to investigate the influence of background interactions upon the behavior of the selected species within their native environments. However, one must be extremely cautious about interpreting the results of such experiments. In particular, it is necessary to design and carry out control experiments that clearly indicate whether or not the processes of labeling and/or overexpression result in abnormal distribution of the selected protein within the cell, induction of artifactual associations, or disruption of the normal background interactions that one is attempting to investigate.

This Commentary bears as its title the provocative question "How can biochemical reactions within cells differ from those in test tubes?" When one can specify the composition of the local environment of a particular reaction at a specific intracellular location and at a particular time point in the cell cycle, studies such as those cited here will help to provide answers to the equally interesting and perhaps more biological question "By how much do biochemical reactions within cells differ from those in test tubes?"

References

Arbuzova A, Murray D, McLaughlin S (1998) MARCKS, membranes, and calmodulin: kinetics of their interaction. Biochim Biophys Acta 1376:369–379

Blanco R, Arai A, Grinberg N, Yarmush DM, Karger BL (1989) Role of association on protein adsorption isotherms. Beta-lactoglobulin A adsorbed on a weakly hydrophobic surface. J Chromatogr 482:1–12

Bloemendal H, de Jong W, Jaenicke R, Lubsen NH, Slingsby C, Tardieu A (2004) Ageing and vision: structure, stability and function of lens crystallins. Prog Biophys Mol Biol 86:407–485

Bolis D, Politou AS, Kelly G, Pastore A, Temussi PA (2004) Protein stability in nanocages: a novel approach for influencing protein stability by molecular confinement. J Mol Biol 336:203–212

Bookchin RM, Balasz T, Wang Z, Josephs R, Lew VL (1999) Polymer structure and solubility of deoxyhemoglobin S in the presence of high concentrations of volume-excluding 70-kDa dextran. Effects of non-S hemoglobins and inhibitors. J Biol Chem 274:6689–6697

Chatelier RC, Minton AP (1996) Adsorption of globular proteins on locally planar surfaces: models for the effect of excluded surface area and aggregation of adsorbed protein on adsorption equilibria. Biophys J 71:2367–2374

Cheung MS, Thirumalai D (2006) Nanopore-protein interactions dramatically alter stability and yield of the native state in restricted spaces. J Mol Biol 357:632643

Cheung MS, Klimov D, Thirumalai D (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins. Proc Natl Acad Sci USA 102:4753–4758

Colclasure GC, Parker JC (1992) Cytosolic protein concentration is the primary volume signal for swelling-induced [K-Cl] cotransport in dog red cells. J Gen Physiol 100:1–10

Cutsforth G, Whitaker R, Hermans J, Lentz B (1989) A new model to describe extrinsic protein binding to phospholipid membranes of varying composition: application to human coagulation proteins. Biochemistry 28:7453–7461

Darst SA, Ribi HO, Pierce DW, Kornberg RD (1988) Two-dimensional crystals of E. coli RNA polymerase holoenzyme on positively charged lipid layers. J Mol Biol 203:269–273

del Alamo M, Rivas G, Mateu MG (2005) Effect of macromolecular crowding agents on human immunodeficiency virus type 1 capsid protein assembly in vitro. J Virol 79:14271–14281

Drenckhahn D, Pollard TD (1986) Elongation of actin filaments is a diffusion limited reaction at the barbed end and is accelerated by inert macromolecules. J Biol Chem 261:12754–12758

Edwards RA, Huber RE (1992) Surface denaturation of proteins: the thermal inactivation of beta-galactosidase (Escherichia coli) on wall-liquid surfaces. Biochem Cell Biol 70:63–69

Eggers D, Valentine J (2001) Molecular confinement influences protein structure and enhances thermal protein stability. Protein Sci 10:250–261

Fernandez A, Berry RS (2003) Proteins with H-bond packing defects are highly interactive with lipid bilayers: implications for amyloidogenesis. Proc Natl Acad Sci USA 100:2391–2396

Ferrone F (2004) Polymerization and sickle cell disease: a molecular view. Microcirculation 11:115–128

Fulton AB (1982) How crowded is the cytoplasm? Cell 30:345–347

Ghaemmaghami S, Oas TG (2001) Quantitative protein stability measurement in vivo. Nat Struct Biol 8:879–882

Giddings JC, Kucera E, Russell CP, Myers MN (1968) Statistical theory for the equilibrium distribution of rigid molecules in inert porous networks. Exclusion chromatography. J Phys Chem 72:4397–4408

Goodsell DS (1993) The machinery of life. Springer, New York

Hall D, Minton AP (2002) Effects of inert volume-excluding macromolecules on protein fiber formation I. Equilibrium models. Biophys Chem 98:93–104

Hall D, Minton AP (2004) Effects of inert volume-excluding macromolecules on protein fiber formation II. Kinetic models for nucleated fiber growth. Biophys Chem 107:299–316

Hatters D, Minton AP, Howlett GJ (2002) Macromolecular crowding accelerates amyloid formation by human apolipoprotein C-II. J Biol Chem 277:7824–7830

Herzog W, Weber K (1978) Microtubule formation by pure brain tubulin in vitro. The influence of dextran and polyethylene glycol. Eur J Biochem 91:249–254

Ignatova Z, Gierasch LM (2004) Monitoring protein stability and aggregation in vivo by real time fluorescent labeling. Proc Natl Acad Sci USA 101:523–528

Jarvis TC, Ring DM, Daube SS, von Hippel PH (1990) "Macromolecular crowding": thermodynamic consequences for protein-protein interactions within the T4 DNA replication complex. J Biol Chem 265:15160–15167

Klimov DK, Newfield D, Thirumalai D (2002) Simulations of beta-hairpin folding confined to spherical pores using distributed computing. Proc Natl Acad Sci USA 99:8019–8024

Knull HR, Walsh JL (1992) Association of glycolytic enzymes with the cytoskeleton. Curr Top Cell Regul 33:15–30

Knull H, Minton AP (1996) Structure within eukaryotic cytoplasm and its relationship to glycolytic metabolism. Cell Biochem Funct 14:237–248

Koo EH, Lansbury PT, Kelly JW (1999) Amyloid diseases: abnormal protein aggregation in neurodegeneration. Proc Natl Acad Sci USA 96:9989–9990

Kozer N, Schreiber G (2004) Effect of crowding on protein-protein association rates: fundamental differences between low and high mass crowding agents. J Mol Biol 336:763–774

Lakatos S, Minton AP (1991) Interactions between globular proteins and F-actin in isotonic saline solution. J Biol Chem 266:18707–18713

Lang F, Busch GL, Ritter M, Völkl H, Waldegger S, Gulbins E, Häussinger D (1998) Functional significance of cell volume regulatory mechanisms. Physiol Rev 78:247–306

Lebowitz JL, Helfand E, Praestgaard E (1965) Scaled particle theory of fluid mixtures. J Chem Phys 43:774–779

Lindner R, Ralston G (1995) Effects of dextran on the self-association of human spectrin. Biophys Chem 57:15–25

Lindner RA, Ralston GB (1997) Macromolecular crowding: effects on actin polymerization. Biophys Chem 66:57–66

Martin J (2002) Requirement for GroEL/GroES-dependent protein folding under nonpermissive conditions of macromolecular crowding. Biochemistry 41:5050–5055

May A, Huehns ER (1975) The concentration dependence of the oxygen affinity of haemoglobin. S Br J Haematol 30:317–335

McNulty BC, Young GB, Pielak GJ (2006) Macromolecular crowding in the Escherichia coli periplasm maintains alpha-synuclein disorder. J Mol Biol 355:893–897

Minton AP (1976) Quantitative relations between oxygen saturation and aggregation of sickle-cell hemoglobin: analysis of oxygen binding data. In: Hercules JI, Cottam GL, Waterman MR, Schechter AN (eds) Proceedings of the symposium on molecular and cellular aspects of sickle cell disease, pp 257–273. U.S. Department of Health, Education and Welfare, Bethesda, MD

Minton AP (1981) Excluded volume as a determinant of macromolecular structure and reactivity. Biopolymers 20:2093–2120

Minton AP (1983) The effect of volume occupancy upon the thermodynamic activity of proteins: some biochemical consequences. Mol Cell Biochem 55:119–140

Minton AP (1989) Holobiochemistry: an integrated approach to the understanding of biochemical mechanism that emerges from the study of proteins and protein associations in volume-occupied solutions. In: Srere P, Jones ME, Mathews C (eds) Structural and Organizational Aspects of Metabolic Regulation. Alan R. Liss, New York, pp 291–306

Minton AP (1992) Confinement as a determinant of macromolecular structure and reactivity. Biophys J 63:1090–1100

Minton AP (1995) Confinement as a determinant of macromolecular structure and reactivity. II. Effects of weakly attractive interactions between confined macrosolutes and confining structures. Biophys J 68:1311–1322

Minton AP (1998) Molecular crowding: analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion. Meth Enzymol 295:127–149

Minton AP (2000a) Effect of a concentrated "inert" macromolecular cosolute on the stability of a globular protein with respect to denaturation by heat and by chaotropes: a statistical-thermodynamic model. Biophys J 78:101–109

Minton AP (2000b) Effects of excluded surface area and adsorbate clustering on surface adsorption isotherms I. Equilibrium models. Biophys Chem 86:239–247

Minton AP (2001a) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J Biol Chem 276:10577–10580

Minton AP (2001b) Effects of excluded surface area and adsorbate clustering on surface adsorption of proteins II. Kinetic models. Biophys J 80:1641–1648

Minton AP (2006) How can biochemical reactions within cells differ from those in test tubes? J Cell Sci 119(14):2863–2869

Minton AP, Wilf J (1981) Effect of macromolecular crowding upon the structure and function of an enzyme: glyceraldehyde-3-phosphate dehydrogenase. Biochemistry 20:4821–4826

Nichol L, Ogston A, Wills P (1981) Effect of inert polymers on protein self association. FEBS Lett 126:18–20

Nygren H, Stenberg M (1990) Surface-induced aggregation of ferritin: kinetics of adsorption to a hydrophobic surface. Biophys Chem 38:67–75

Parameswaran S, Barber BJ, Babbit RA, Dutta S (1995) Age-related changes in albumin-excluded volume fraction. Microvasc Res 50:373–380

Ramsden JJ, Bachmanova GI, Archakov AI (1994) Kinetic evidence for protein clustering at a surface. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics 50:5072–5076

Rivas G, Fernandez JA, Minton AP (1999) Direct observation of the self association of dilute proteins in the presence of inert macromolecules at high concentration via tracer sedimentation equilibrium: theory, experiment, and biological significance. Biochemistry 38:9379–9388

Rivas G, Fernandez JA, Minton AP (2001) Direct observation of the enhancement of non-cooperative protein self-assembly by macromolecular crowding: indefinite linear self-association of bacterial cell division protein FtsZ. Proc Natl Acad Sci USA 98:3150–3155

Rotter M, Aprelev A, Adachi K, Ferrone F (2005) Molecular crowding limits the role of fetal hemoglobin in therapy for sickle cell disease. J Mol Biol 347:1015–1023

Sasahara K, McPhie P, Minton AP (2003) Effect of dextran on protein stability and conformation attributed to macromolecular crowding. J Mol Biol 326:1227–1237

Shastry M, Eftink M (1996) Reversible thermal unfolding of ribonuclease T1 in reverse micelles. Biochemistry 35:4094–4101

Shtilerman M, Ding PT, Lansbury PT (2002) Molecular crowding accelerates fibrillization of alpha-synuclein: could an increase in the cytoplasmic protein concentration induce Parkinson's disease? Biochemistry 41:3855–3860

Somero GN, Osmond CB, Bolis CL (1992) Water and life. Springer, Berlin

Spencer D, Xu K, Logan T, Zhou H (2005) Effects of pH, salt, and macromolecular crowding on the stability of FK506-binding protein: an integrated experimental and theoretical study. J Mol Biol 351:219–232

Steadman BL, Trautman PA, Lawson EQ, Raymond MJ, Mood DA, Thomson JA, Middaugh CR (1989) A differential scanning calorimetric study of the bovine lens crystallins. Biochemistry 28:9653–9658

Tellam RL, Sculley MJ, Nichol LW, Wills PR (1983) Influence of polyethylene glycol 6000 on the properties of skeletal-muscle actin. Biochem J 213:651–659

Tokuriki N, Kinjo M, Negi S, Hoshino M, Goto Y, Urabe I, Yomo T (2004) Protein folding by the effects of macromolecular crowding. Protein Sci 13:125–133

Uversky V, Cooper M, Bower K, Li J, Fink A (2002) Accelerated alpha synuclein fibrillation in crowded milieu. FEBS Lett 515:99–103

van den Berg B, Ellis R, Dobson C (1999) Effects of macromolecular crowding on protein folding and aggregation. EMBO J 18:6927–6933

Wilf J, Gladner JA, Minton AP (1985) Acceleration of fibrin gel formation by unrelated proteins. Thromb Res 37:681–688

Zhou HX (2004) Protein folding and binding in confined spaces and in crowded solutions. J Mol Recognit 17:368–375

Zhou HX, Dill KA (2001) Stabilization of proteins in confined spaces. Biochemistry 40:11289–11293

Zhou HX, Rivas G, Minton AP (2008) Annu. Rev. Biophys. 37:375–397

Zimmerman SB, Harrison B (1987) Macromolecular crowding increases binding of DNA polymerase to DNA: an adaptive effect. Proc Natl Acad Sci USA 84:1871–1875

Zimmerman SB, Trach SO (1991) Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. J Mol Biol 222:599–620

Zimmerman SB, Minton AP (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. Annu Rev Biophys Biomol Struct 22:27–65

Part II
Steps Towards Functionality

Chapter 6
The Influence of Environment and Metabolic Capacity on the Size of a Microorganism

W. Andrew Lancaster and Michael W.W. Adams

Abstract The environment a microorganism inhabits dictates the metabolic capacity necessary for it to survive, and ultimately the minimum size which an organism can achieve. Nutrient rich environments such as those experienced by parasitic bacteria can accommodate organisms with restricted metabolic capacities with relatively few genes, perhaps as few as 250. Nutrient poor environments, such as those experienced by autotrophs, provide only minerals and gases and require high biosynthetic capacity to synthesize all cellular carbon from CO_2. This high biosynthetic capacity requires at most 1,500 (an actual value) and perhaps as few as 750 genes. Calculations show that as theoretical minimal cell size is decreased, the cellular volume devoted to the DNA required to encode the minimum gene complement becomes a limiting factor in further reduction. Assuming composition of 50% water, 20% protein, 10% ribosomes and 10% DNA, a spherical cell with minimum biosynthetic capacity (250 genes) would be at least 172 nm in diameter. A cell with high biosynthetic capacity (750 genes) of the same composition would be at least 248 nm in diameter. It is concluded that cells with biochemical requirements for growth, metabolism and reproduction similar to those of known organisms cannot be smaller than 172 nm.

6.1 Introduction

The question of minimum microbial size was dramatically brought to the fore several years ago by the report of McKay and coworkers (McKay et al. 1996) in which objects with upper dimensions of 20 by 100 nm were postulated to be of cellular origin. These claims have remained controversial more than a decade after their initial publication (Thomas-Keprta et al. 2002; Golden et al. 2004). Nevertheless, the discovery of possible extraterrestrial microbial fossils has spurred interest in ongoing efforts to detect past or current conditions suitable for life on the surface of Mars.

W.A. Lancaster (✉) and M.W.W. Adams
Department of Biochemistry and Molecular Biology, University of Georgia, 30602 Athens, GA
e-mail: adams@bmb.uga.edu

In addition, so-called ultramicrobacteria have been isolated from marine environments that can pass through a 200 nm filter and have cell volumes of 0.03–0.08 μm^3 (Eguchi et al. 1996). Recently one of these ultramicrobacteria, *Herminiimonas glaciei*, was isolated from an ice core taken at a depth of 3 km from a Greenland glacier (Loveland-Curtze et al. 2009). The organism was successfully grown in laboratory culture after lying dormant for approximately 120,000 years. Entities known as "nanobacteria" having diameters as low as 80 nm and implicated in human disease were reported to have been cultured from blood (Kajander et al. 1997), which suggested a very low limit on microbial cell size. Subsequent study has shown that these entities are most likely calcium carbonate precipitates (Martel and Young 2008), leaving the question of the lower limit on microbial cell size very much open. In light of these and other studies, it is important to estimate a reasonable limit for the theoretical minimum size of microorganisms. Would it be possible for a cell with 50 nm diameter to contain sufficient biological material to sustain a free-living lifestyle? This immediately raises the question of what minimum quantity of biological material is sufficient. One measure that has been increasingly informative as the quantity of sequence data has exploded is genome size, which determines diversity of proteins (enzymes) that an organism has at its disposal to support growth, metabolism and reproduction.

The minimum complement of genes necessary to maintain life is intensely dependant on the biosynthetic capacity of a given organism, which is in turn largely determined by the environment in which the organism is found. It is no surprise that a wide range of biosynthetic capacities exists among organisms given the extent to which dependence on the environment varies. All organisms must satisfy the same basic biochemical requirements, but some environments provide only basic compounds necessary for life as we know it, while other provide a variety of complex molecules, obviating the need for their synthesis by an organism inhabiting such an environment. Minimum requirements for human survival include ten or so amino acids, various minerals, a variety of biological cofactors (vitamins), and a continual supply of oxygen. Like us, the vast majority of microorganisms require a fixed carbon source, which is usually a carbohydrate although some can utilize lipids, nucleotides, or various simple organic compounds. Microbes, however, differ from humans in their ability to synthesize most of the amino acids as well as essential cofactors which we must obtain from our environment. In contrast, some microorganisms are intensely dependent upon their environments. For example, some microbial parasites do not synthesize any amino acid or lipid, and only a few vitamins and nucleotides; rather, they obtain all of these compounds from their host. One interesting class of organisms completely dependent on a host organism is the mutualistic endosymbionts, one example of which is *Carsonella rudii* which provides essential amino acids to their insect hosts and lack genes for many essential synthetic pathways. The genome of *Carsonella rudii* is only 160 kilobases, which encodes 182 open reading frames, making it the smallest sequenced bacterial genome (Nakabachi et al. 2006). The extreme reduction of the *C. rudii* genome and the degree of its interaction with the host have led to the proposal that it represents a transition from a symbiotic cell to an organelle of the host cell (Tamames et al. 2007).

Such reduced life-forms approach the environmental dependence of a virus, which consists of a protein coat that surrounds a defined amount of nucleic acid (DNA or RNA). It has recently been proposed that biological entities should be grouped into two major categories, Ribosome-encoding organisms (REOs) and Capsid-encoding organisms (CEOs) (Raoult and Forterre 2008). In this scheme, REOs include free living cells, parasites, endosymbionts and organelles, CEOs include viruses, with plasmids and viroids considered "orphans", encoding neither ribosomes nor capsids. The CEOs rely on their host organism not only for basic biological compounds, but for the transcriptional and translational machinery of REOs necessary for their own reproduction. Along with the plasmids and viroids, these entities are able to achieve very small size by encoding minimal information necessary for their replication and exploiting the host environment to provide the complex machinery and metabolic activity necessary to support it.

Plasmids, Viruses and parasitic Bacteria obviously depend greatly on their environments to provide complex and varied components necessary for their survival. Other organisms however, thrive in quite sparse environments. Such organisms require only simple chemicals, such as CO_2, O_2, H_2, and NH_3 and synthesize all amino acids, cofactors, nucleotides, etc., with CO_2 as the sole carbon source, and ammonia (or even N_2 gas) as the nitrogen source. These autotrophs may obtain the energy used to synthesize these compounds from light, as in green plants and cyanobacteria, or from the oxidation of abundant compounds present in their environments. Substances oxidized by aerobic autotrophs include H_2, CO, CH_4, NH_3, or H_2S. Similarly, anaerobic autotrophs growing on H_2 and CO_2 also conserve energy during the reduction of CO_2, either by the production of methane or acetate as accomplished by methanogens and acetogens, respectively. There is obviously a spectrum of environmental dependence with a rich diversity of organisms represented at the aforementioned extremes and everywhere in between.

Given the vast differences in the dependence of organisms on their environments, it should be clarified what is meant by "free-living" in the context of elucidating the minimal hypothetical cell. All life forms must at a minimum undergo self-directed reproduction when supplied with the molecules they need in sufficient quantities and, in some cases, non-molecular sources of free energy in the form of electromagnetic radiation. This minimum requirement is met by entities which do not construct the machinery necessary for the execution of this replication. The environment of a CEO, a cell with an array of replicative machinery which can be usurped, clearly allows these entities to attain a very small size indeed while allowing replication. In fact, an even greater reduction is possible for plasmids which rely on the host cell not only for the copying of genetic information, but with its maintenance and transmission. There is a major distinction to be made between such replicators and even the most diminutive parasitic bacterium which, while relying on an array of complex molecules from the environment, does not reproduce via a transfer of genetic material into the environment. Although parasitic cellular life may depend on the complex substances provided by a living host, these nutrients, if provided to the organism, replace the host and allow the parasite to grow and reproduce. This is not the case with viruses and plasmids

which rely on the intact functioning system of the host. These entities will not therefore be considered in determining the lower size limit for life.

As has been discussed, life forms occupy a wide range of environments which greatly affect their metabolic capacities. In determining the lower limit on microbial cell size, it will be useful to consider two extremes, the rich environment experienced by parasitic life forms, and the sparse environment providing only gasses and minerals experienced by many autotrophs. In determining these limits, it is vital to estimate the number of proteins required by our two classes of organisms, those with high and low biosynthetic capacities. The ever increasing availability of genome sequences for a variety of microorganisms enables quantitative estimations to be made.

6.2 Organisms with Low Biosynthetic Capacity

Those organisms that are most dependent upon their environment are the parasitic bacteria, the prototypical example of which are the *Mycoplasma*. The genome of *Mycoplasma genitalium* was the second to be completely sequenced (Fraser et al. 1995). At 0.58 Mb, this still represents the smallest known genome of any free-living organism. The genome contains 470 predicted protein coding regions, and these include those required for DNA replication, transcription and translation, DNA repair, cellular transport, and energy metabolism. An early comparative genomic study utilizing the genome of the first organism to be completely sequenced, *Haemophilus influenzae* (Fleischmann et al. 1995), led to the conclusion that the "minimal gene set that is necessary and sufficient to sustain the existence of a modern-type cell" is (only) 256 genes, or about half of the genome of *M. genitalium* (Mushegian and Koonin 1996). Subsequent comparative genomic and experimental studies have yielded a range of estimates for the minimal genome complement (Carbone 2006) and the concept of what constitutes a minimal gene set has evolved as greater amounts of sequence data have become available. A recent study involving 573 bacterial genomes identified an "extended core" set of ~250 protein families present in almost all species (Lapierre and Gogarten 2009) which may be considered as necessary but not necessarily sufficient for a modern minimal bacterial cell. It should be noted, however, that while some parasitic organisms such as *Mycoplasma* species grow in the absence of their hosts, to do so they require an extremely rich medium which mimics the availability of nutrients encountered in the infected host. These organisms maintain a minimal biosynthetic capacity, a capacity that could reasonably be satisfied by approximately 250 different proteins.

6.3 The Most Slowly Evolving Microorganisms

In determining the "minimum" set of genes that a minimal-size microbe might contain, we must also consider what is meant by the term "modern-type cell" quoted above (Mushegian and Koonin 1996). Are all present-day organisms highly sophisticated cells

with a range of metabolic capabilities, only some of which are utilized and then under very specialized conditions? For example, *E. coli* could be regarded as highly evolved because it exhibits a range of metabolic modes, including growth under aerobic and anaerobic conditions, the utilization of a wide variety of different carbon sources, etc. Indeed, such a large metabolic capacity might be reflected in its genetic content of 4.64 Mb encoding 4,288 genes (Blattner et al. 1997). Similarly, metabolically diverse species such as *Bacillus subtilus* and *Pseudomonas putida* have genomes of comparable size (≥4 Mb). The average genome length of 1,113 sequenced bacterial genomes including large genomes such as *Sorangium cellulosum* (13 Mb) is 3.6 Mb, (Liolios et al. 2008). In other words, it is not unusual for well characterized microorganisms to contain 4,000 or more genes, in some cases many more. Hence, are there more-slowly-evolving organisms, and do they contain less genetic material and have fewer metabolic choices?

By phylogenetic analyses based on 16S rRNA sequence comparisons, the most-slowly-evolving microorganisms are the deepest branches, the first to have diverged within the two major lineages corresponding to the Bacteria and the Archaea (Woese et al. 1990; Stetter 2006). Remarkably, in both domains these are the hyperthermophiles, organisms that grow optimally at temperatures of 80°C and above. Within the bacteria domain this includes few genera, the most well studied of which are *Thermotoga* and *Aquifex*, while there are more than 20 genera of hyperthermophilic archaea (Liolios et al. 2008). In fact, one of the two major branches within the archaeal domain consists almost entirely of hyperthermophiles, while in the other branch the hyperthermophiles are the most slowly evolving of the known genera. A great deal is known about the genome contents of these hyperthermophilic organisms as representatives were among the first sequenced genomes, and there is continuing interest in the potential applications of thermostable enzymes. There are currently 31 completely sequenced hyperthermophilic archaeal and seven hyperthermophilic bacterial genomes, with many more ongoing sequencing projects (Liolios et al. 2008). Interestingly, most of these organisms have genomes only about half the size of that found in *E. coli*, with those of *Sulfolobus solfataricus* and *Ignicoccus hospitalis* being the largest (2.99 Mb) and smallest (1.30 Mb) of the free-living members of this group, respectively with an average of 2.04 Mb for all the hyperthermophilic organisms in this set. Thus, the most slowly evolving organisms known (at least as determined by 16S rRNA analyses) do indeed have relatively small genomes, although they are still highly complex life-forms. Interestingly, *Ignicoccus hospitalis*, the free-living hyperthermophile with the smallest sequenced genome is host to the smallest hyperthermophilic parasite, *Nanoarchaeum equitans* (Huber et al. 2002). *N. equitans* has a diameter of only 400 nm, and its 491 kilobase genome containing 552 predicted open reading frames (Waters et al. 2003) is very close in size to that of *M. genitalium*. Experiments have shown that organisms closely related to the sequenced *Ignicoccus* and *Nanoarchaeum* species are among the first to colonize nascent hydrothermal vents (McCliment et al. 2006). Nanoarchaeal sequences have been identified in a wide range of environments, but whether all members are obligate parasites, or there are free-living representatives remains unknown (Forterre et al. 2009).

6.4   Organisms with High Biosynthetic Capacity

The genome sequences of autotrophs can provide us with information on the upper limit of the number of different protein encoding genes necessary for organisms living in sparse environments providing only gases and minerals. To date, the genomes of several hyperthermophilic autotrophs have been sequenced. The first to be sequenced was the archaeon, *Methanococcus jannaschii*, (later reclassified as *Methanocaldococcus jannaschii)*, (Bult et al. 1996), which is a methanogen that grows up to 90°C using H_2 and CO_2 as energy and carbon sources and generates methane as an end product. The genome of *Methanocaldococcus jannaschii* is 1.7 Mb with approximately 1,700 protein encoding genes. The average genome length of eight sequenced hyperthermophilic autotrophs, seven of which are archaea, is 1.7 Mb (Liolios et al. 2008). This group includes anaerobic sulphur, nitrate and arseneic reducers and methanogens. The one bacterium in this group is the microaerophilic *Aquifex aeolicus* (Deckert et al. 1998), which grows up to 95°C on H_2 and CO_2. The genome of *Aquifex aeolicus* is 1.5 Mb with approximately 1,500 protein encoding genes. It should be noted that the pathway of CO_2 assimilation and the biochemistry of energy conservation across these sequenced examples is quite varied, yet their metabolic requirements are met by genomes of similar size. On the other hand, these genomes are much larger than those of the parasitic organisms discussed above. Presumably, these organisms require many more genes because they must synthesize all of their cellular components from CO_2. Hence they contain about three times the genetic information of *M. genitalium*. This seems appropriate considering that the latter organism is supplied with all of its amino acids, nucleotides, fatty acids, "vitamins," and an energy source (glucose). From this direct comparison we might conclude that about two thirds of the DNA in *A. aeolicus* and *M. jannaschii*, or approximately 1,000 genes, encodes proteins that function to carry out these biosynthetic tasks and produce all of these compounds from CO_2.

6.5   The Smallest Cell

From the above discussion it can be concluded that a cell with minimal biosynthetic capacity that is growing in a nutrient-rich medium requires between 250 (a calculated value) and 500 genes (the approximate number in M. genitalium) to grow. At the other extreme is a cell that synthesizes all of its cellular material from CO_2, and this requires at most 1,500 genes (the approximate number in *A. aeolicus* and *I. hospitalis*) and probably closer to 750 genes (half of the actual value). With these values in mind, it is instructive to consider the amount of biological material which can be contained within a hypothetical small cell of 50 nm diameter. Assuming a 5 nm thick lipid bilayer, a spherical cell of 50 nm diameter would have an internal volume of 33,500 nm^3. An *E. coli* cell, with dimensions of about 1.3 by 4.0 μm, has an internal volume of about 5×10^9 nm^3, or almost two million times the volume of the 50 nm diameter cell.

What quantities of the various biochemical structures found in a typical prokaryotic cell can be accommodated within this hypothetical cell with six orders of magnitude less space than well characterized real cells? A ribosome has a diameter of about 20 nm, and ribosomes are typically 25% of the mass (dry weight) of a bacterial cell (although this varies considerably depending on the growth rate). Assuming a similar percentage of the volume of a 50 nm diameter cell is devoted to them, such a cell could contain only two ribosomes (4,200 nm^3 each). Whether only two would limit cell growth to any extent is unknown; nevertheless, the cell is certainly large enough to contain ribosomes, albeit only two. On the other hand, proteins usually constitute about half of the dry weight of bacterial cells. Let us assume that they also occupy approximately 50% of the volume of the 50 nm diameter cell, and that, in general, proteins have an average molecular weight of 30 kDa, which corresponds to a diameter of about 4 nm per protein. If a cell of 50 nm diameter were 50% protein by volume, then this would correspond to about 520 such molecules (average 30 kDa) per cell.

Are two ribosomes and 520 "average-sized" protein molecules sufficient to support the growth of a cell? This amount of protein represents on average, only two copies of each protein for a cell with minimal biosynthetic capacity (predicted to contain 250 genes). Next, the amount of DNA such a cell could contain must be considered. DNA typically occupies only about 3% of the total mass (dry weight) of well characterized bacterial cells, so upon first consideration, it would seem that the volume of genetic material especially in an organism with a minimum gene content, would be a major limiting factor on cellular volume. For example, with a diameter of 2 nm and length of 0.34 nm/bp, the 4.64 Mb of E. coli has a volume of 4.9×10^6 nm^3, which is less than 1% of the cell volume. The 0.49 Mb *N. equitans* genome comprises only about 1.5% of the volume of these 400 nm diameter cells. Surprisingly, however, simple calculations show that DNA is a determining factor in much smaller cells. Thus, the hypothetical 50 nm diameter cell, 75% of which (by volume) is occupied by proteins and ribosomes, could contain, even if the remaining 25% of the cell were devoted to DNA, only eight genes (of 1,000 bp each)!

6.6   DNA Content Determines Minimal Cell Size

Since our hypothetical 50 nm cell has revealed the major volume constraint imposed by DNA, it is clear that the minimum necessary amount of genetic material can serve as a basis for computing a minimum cell size. Even if a hypothetical cell could devote half of its total internal volume to DNA, a cell with minimal biosynthetic capacity would need to have a diameter of 110 nm to contain its 250 genes. This increase in volume also provides space for additional components other than DNA. Assuming that half of the remaining volume (25%) of the 110 nm diameter cell is devoted to protein, and 12.5% to ribosomes, the cell could contain 4,000 kDa proteins (eight per gene) and 15 ribosomes. A hypothetical minimal autotroph with 750 genes and the same proportion of components (50% DNA, 25% protein, 12.5% ribosomes) would have a volume of 156 nm. This 156 nm cell could contain 12,400

(16 per gene) proteins and 48 ribosomes. Cells with this proportion of components would have 12.5% of their volume remaining for other cellular components such as lipids, metabolites, cofactors and inorganic substances. These minimum volumes have neglected one vital component, water, which typically occupies the greatest volume of microbial cells. If it is assumed that the hypothetical minimum cell is comprised of 50% water, the minimum computed volumes rise to 136 and 194 nm for cells with minimal and maximal biosynthetic capacity, respectively.

It is clear from these calculations that DNA represents a critical factor in the minimal limit on cell size. The calculations made the unrealistic assumption that 25% of the cell's volume could be devoted to DNA. If a more accurate estimate of minimal cell size is to be reached, a more reasonable figure for the amount of DNA which can be contained in a cell should be used. As previously mentioned, E. coli DNA occupies less than 1% of cellular volume, and it does not seem likely that more than 10% of the volume of any viable functioning cell could be devoted to the chromosome. While there are additional serious constraints on the packing of DNA such as strand bending and negative charge repulsion and the necessity of unwinding the strand for transcription and replication which may further limit the maximum proportion, assuming a content of 10% DNA should yield an informative estimate of a reasonable lower limit on cell size. Thus, a cell that contains 250 genes that occupy 10% of its volume would be 172 nm in diameter (Fig. 6.1b), while one containing 750 genes occupying the same relative volume would be 248 nm in diameter (Fig. 6.1c). Assuming such cells contain 20% by volume protein and 10% by volume ribosomes, the 172 nm cell could accommodate 64 ribosomes and over 16,000 proteins, or 65 per gene, and the 248 nm cell could contain three times as much.

It may be concluded, therefore, that the minimum theoretical size for a cell is 172 nm in diameter. Since such a cell has a minimal biosynthetic capacity and a reduced genome of 250 genes, it would need to be supplied with all amino acids, fatty acids, nucleotides, cofactors, and other components to survive. The cell would have no cell-wall and would consist, by volume, of 10% DNA, 10% ribosomes, 20% protein, and 50% water. It would contain approximately 65 proteins per gene as well as 65 ribosomes. In contrast, an autotrophic cell with a much higher biosynthetic capacity, such that it could synthesize all cellular components from CO_2, would be 248 nm in diameter, assuming that its DNA is limited to 10% of the cell volume. Note that these calculations assume a theoretical minimum gene content, which is about half of that present in known life-forms. The amount of DNA in a known autotrophic organism (approximately 1,500 genes in *A. aeolicus*) would require a cell of at least 314 nm in diameter, assuming that it occupied 10% of the cell by volume. Hence, depending on the biosynthetic capacity of a cell, and the extent to which the calculated minimum gene content is realistic, its minimum diameter is between 172 and 314 nm. Overall, one can conclude that microorganisms cannot have diameters significantly less than 200 nm if they have the same basic biochemical requirements for growth as all other extant life-forms, but even then they would require very specialized environments. More likely, the absolute minimum size is closer to 250 nm where the cell is able to grow on simple compounds commonly found in various natural environments.
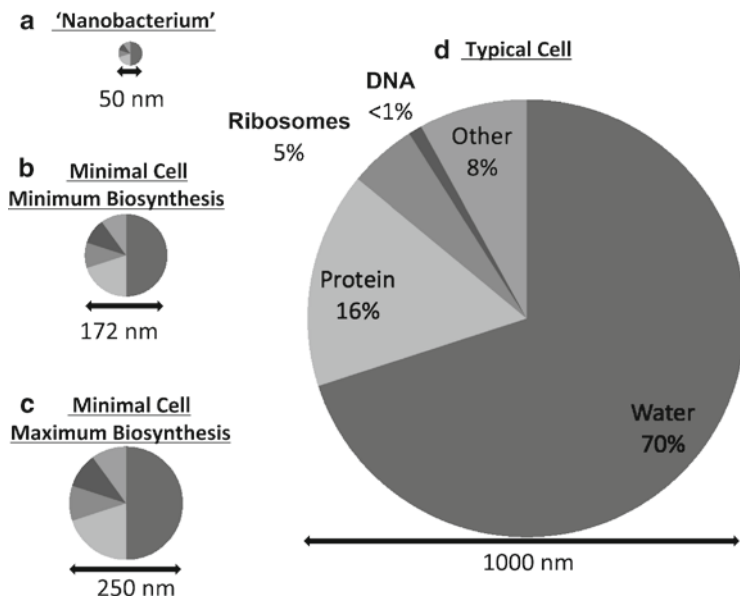
Fig. 6.1 A comparison of the diameters (drawn to scale) and compositions of (**a**) a 50 nm hypothetical 'nanobacterium', (**b**) a minimal cell with minimal biosynthetic capacity, (**c**) a minimal cell with maximal biosynthetic capacity, and (**d**) a typical 1 μm prokaryotic cell with a composition based on that reported for *E. coli,* which includes ≤1% of the total volume donated to DNA (Neidhardt 1996). The minimal cell sizes (**a–c**) are based on compositions of 10% DNA, 50% water, 20% protein, 10% ribosomes and 10% other components

   To date, no free-living organism close to the hypothetical 250 nm minimum diameter has been identified. What place would such a minimal organism have in the modern world? In natural environments, an organism must cope with a myriad of challenges such as constantly changing metabolite concentration, changing temperature, the intrusion of toxic compounds and radiation, as well as competition from other organisms. The ability to deal effectively with such situations will undoubtedly outweigh the metabolic cost of maintaining some number of genes in addition to the minimal core. This may not have been the case in archaic ecosystems with lower diversity of cellular life and viruses, as well as potentially reduced replicative fidelity relative to that present in modern systems. Whether or not ancestral microbes resembled our concept of minimal life or not, there is a place for minimal life in the modern world. The field of synthetic biology is burgeoning, and attempts are being made to construct a minimal cell (Glass et al. 2006; Gibson et al. 2008) both as a tractable perturbable system in which to study the fundamentals of life, and as a platform for creating custom organisms with known behavior to produce useful substances and perform specific tasks. In the latter regard, the minimal cell has two primary advantages. The first is that such a system can potentially achieve greater efficiency and more predictable behavior by adding only that which is necessary to perform the desired task relative to engineering a system

based on a versatile organism such as *E. coli*. The other is that the metabolic and environmental fragility becomes a distinct advantage in certain cases. In producing biological substances in a laboratory environment, stable conditions can be maintained, and the real or perceived danger of engineered microbes escaping into natural ecosystems causing damage is reduced. The utility of such organisms built from the "ground up" remains to be seen, but the potential to further our understanding of fundamental life processes as well as perform useful tasks and produce useful compounds is staggering.

References

Blattner FR, Plunkett G III, Bloch CA et al (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1462

Bult CJ, White O, Olsen GJ et al (1996) Complete genome sequence of the Methanogenic Archaeon, *Methanococcus jannaschii*. Science 273:1058–1073

Carbone A (2006) Computational prediction of genomic functional cores specific to different microbes. J Mol Evol 63:733–746

Deckert G, Warren PV, Gaasterland T et al (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature 392:353–358

Eguchi M, Nishikawa T, Macdonald K et al (1996) Responses to stress and nutrient availability by the marine Ultramicrobacterium *Sphingomonas* sp. Strain RB2256. Appl Environ Microbiol 62:1287–1294

Fleischmann RD, Adams MD, White O et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496–512

Forterre P, Gribaldo S, Brochier-Armanet C (2009) Happy together: genomic insights into the unique *Nanoarchaeum/Ignicoccus* association. J Biol 8:7

Fraser CM, Gocayne JD, White O et al (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270:397–403

Gibson DG, Benders GA, Andrews-Pfannkoch C et al (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. Science 319:1215–1220

Glass JI, Assad-Garcia N, Alperovich N et al (2006) Essential genes of a minimal bacterium. Proc Natl Acad Sci USA 103:425–430

Golden DC, Ming DW, Morris RV et al (2004) Evidence for exclusively inorganic formation of magnetite in Martian meteorite ALH84001. Am Mineral 89:681–695

Huber H, Hohn MJ, Rachel R et al (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature 417:63–67

Kajander EO, Kuronen I, Akerman KK et al (1997). Nanobacteria from blood: the smallest culturable autonomously replicating agent on Earth. Instruments, methods, and missions for the investigation of extraterrestrial microorganisms. SPIE, San Diego, CA, USA

Lapierre P, Gogarten JP (2009) Estimating the size of the bacterial pan-genome. Trends Genet 25:107–110

Liolios K, Mavromatis K, Tavernarakis N et al (2008) The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res 36:D475–D479

Loveland-Curtze J, Miteva VI, Brenchley JE (2009) *Herminiimonas glaciei* sp. nov., a novel ultramicrobacterium from 3042 m deep Greenland glacial ice. Int J Syst Evol Microbiol 59:1272–1277

Martel J, Young JD (2008) Purported nanobacteria in human blood as calcium carbonate nanoparticles. Proc Natl Acad Sci USA 105:5549–5554

McCliment EA, Voglesonger KM, O'Day PA et al (2006) Colonization of nascent, deep-sea hydrothermal vents by a novel Archaeal and Nanoarchaeal assemblage. Environ Microbiol 8:114–125

McKay DS, Gibson EK Jr, Thomas-Keprta KL et al (1996) Search for past life on Mars: possible relic biogenic activity in martian meteorite ALH84001. Science 273:924–930

Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci USA 93:10268–10273

Nakabachi A, Yamashita A, Toh H et al (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314:267

Neidhardt FC (1996) Chemical composition of *Escherichia coli*. In: Neidhardt FC, Umbarger HE (eds) *Escherichia coli* and *Salmonella*, cellular and molecular biology, vol 1, pp 13–16. America Society for Microbiology, Washington DC

Raoult D, Forterre P (2008) Redefining viruses: lessons from Mimivirus. Nat Rev Microbiol 6:315–319

Stetter KO (2006) Hyperthermophiles in the history of life. Philos Trans R Soc Lond B Biol Sci 361:1837–1842, discussion 1842–1843

Tamames J, Gil R, Latorre A et al (2007) The frontier between cell and organelle: genome analysis of Candidatus Carsonella ruddii. BMC Evol Biol 7:181

Thomas-Keprta KL, Clemett SJ, Bazylinski DA et al (2002) Magnetofossils from ancient Mars: a robust biosignature in the martian meteorite ALH84001. Appl Environ Microbiol 68:3663–3672

Waters E, Hohn MJ, Ahel I et al (2003) The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. Proc Natl Acad Sci USA 100:12984–12988

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576–4579

Chapter 7
The Minimal Cell and Life's Origin: Role of Water and Aqueous Interfaces

Gerald H. Pollack, Xavier Figueroa, and Qing Zhao

Abstract The cell is rich with interfaces. But the role of these interfaces with water has received little attention among biologists, who generally consider water to be a mere background carrier of the more important molecules of life. Hydrophilic surfaces do impact water, and it has been recently shown that the impact is larger than anticipated. Surfaces order water to substantial distances. The ordered water excludes solutes and separates charge. These features not only contribute to the gel-like nature of the cell, but also lead to an inevitability that pre-cells will form out of simple environmental constituents. Hence, an experimentally based mechanism is advanced to explain both life's origin its requirement for water.

7.1   Introduction

Even before the appearance of the classic book by Frey-Wyssling more than a half-century ago (Frey-Wyssling 1953) the cytoplasm's gel-like consistency should have been evident to any who ventured to crack open a raw egg. Such consistency would have been immediately obvious. From this gel foundation grew the long debated idea of the "gel-sol" transition as a central biological mechanism (Jones 1999; Berry et al. 2000), as well as various other consequences of the cytoplasm's gel-like nature (Janmey et al. 2001; Hochachka 1999).

Such gel-based phenomena are well studied by physical scientists, but the fruits of their understanding have made little headway in the biological arena. Most

G.H. Pollack (✉), X. Figueroa, and Q. Zhao
Department of Bioengineering 355061, University of Washington, 98195 Seattle, WA, USA
e-mail: ghp@u.washington.edu

Q. Zhao
State Key Laboratory for Mesoscopic Physics, and Electron Microscopy Laboratory,
School of Physics, Peking University, 100871 Beijing, China

biologists have had little exposure to the physical chemistry of gels, interfaces, or even water. And fewer have considered how these features might relate to the minimal cell or to life's origin.

Perhaps it is for this reason that virtually all cell biological mechanisms build on the notion that the cell is an aqueous solution – or, more specifically that molecules freely diffuse within the cell. Diffusional steps are required in current understanding of many stimulus-to-action processes; these include: ions diffusing into and out of membrane channels; ions diffusing into and out of membrane pumps; ions diffusing through the cytoplasm; proteins diffusing toward other proteins; substrates diffusing toward enzymes; etc. A cascade of diffusional steps underlies virtually every intracellular process – notwithstanding the cytoplasm's character as a gel, where diffusion can sometimes be slow enough to be biologically irrelevant.

This odd dichotomy between theory and evidence has grown unchecked, in large part because scientists with only limited familiarity with the physical nature of gels and interfaces have carried the flag of cell-biological understanding.

What havoc has such misunderstanding wrought?

7.2 Problems with the Aqueous-Solution-Based Paradigm

Consider the consequences of presuming that the cytoplasm is an aqueous solution. To keep this "solution" and its solutes constrained, a membrane, which is presumed impervious to most solutes, is presumed to surround this liquid-like milieu. But solutes need to pass into and out of the cell – to nourish the cell, to effect communication between cells, to exude waste products, etc. In order for solutes to pass into and out of the cell, the membrane requires openable pores. Well over 100 solute-specific channels have been identified, with new ones emerging regularly.

The same goes for membrane pumps: Since ion concentrations inside and outside the cell are rarely in electrochemical equilibrium, the observed concentration gradients are thought to be maintained by active pumping, mediated by specific entities lodged within the membrane. The concepts of channels and pumps will certainly be familiar to biological readers; for others, the cell-biology text by Alberts et al. (1994) provides a detailed review of this foundational paradigm – along with the manner in which this paradigm accounts for many basics of cell function. In essence, solute partitioning between the inside and the outside of the cell is assumed to be a product of a solute-impermeable membrane containing membrane pumps and membrane channels.

If partitioning requires a continuous, impermeant barrier, then violating the barrier should collapse the concentration gradients. Metabolic processes should grind to a halt, enzymes and fuel should dissipate as they diffuse out of the cytoplasm, and the cell should be quickly brought to the edge of death.

7.2.1 Does this Really Happen?

To disrupt the membrane experimentally, scientists have concocted an array of implements not unlike lances, swords and guns:

- *Microelectrodes*. These are plunged into cells, in order to measure electrical potentials between inside and outside, or to pass substances into the cell cytoplasm. The microelectrode tip may seem diminutive by conventional standards, but to the 10-μm cell invasion by a 1-μm probe is roughly akin to the reader being invaded by a fence post.
- *Electroporation* is a widely used method of effecting material transfer into a cell. By shot-gunning the cell with a barrage of high-voltage pulses the membrane becomes riddled with orifices large enough to pass genes, proteins and other macromolecules – certainly large enough to pass ions and small molecules easily.
- The *patch-clamp* method involves the plucking of a 1-μm patch of membrane from the cell for electrophysiological investigation; the cell membrane is grossly violated.

Although such insults may sometimes inflict fatal injuries they are not necessarily consequential. Consider the microelectrode plunge and subsequent withdrawal. The anticipated surge of ions, proteins and metabolites might be thwarted if the hole could be kept plugged by the microelectrode shank – but this is not always the case. Micropipettes used to microinject calcium-sensitive dyes at multiple sites along muscle cells require repeated withdrawals and penetrations, each withdrawal leaving multiple micron-sized injuries. Yet, normal function is observed for up to several days (Taylor et al. 1975). The results of patch removal are similar. Here again, the hole in question is more than a million times the cross-section of the potassium ion. Yet, following removal of the 1-μm patch, the 10 μm isolated heart cell is commonly found to live on and continue beating (Tung and Vassort, personal communications).

Similarly innocuous is the insult of electroporation. Entry of large molecules into the cell is demonstrable even when molecules are introduced into the bath up to many hours after the end of the electrical barrage that initially created the holes (Xie et al. 1990; Klenchin et al. 1991; Prausnitz et al. 1994; Schwister and Deuticke 1985; Serpersu et al. 1985). Hence, the pores must remain open for long periods without resealing. Structural studies following membrane disruption in muscle and nerve cells also show no membrane re-growth (Cameron 1988; Krause et al. 1984). Thus, notwithstanding long-term membrane orifices of even macromolecular size, with expected leakage of molecules critical for life, the cell does not perish.

If the examples above seem too technical, consider the common alga *Caulerpa*, a single cell whose length can grow to several meters. This giant cell contains stem, roots, and leaves in one cellular unit undivided by any internal walls or membranes (Jacobs 1994). Although battered by pounding waves and gnawed on by hungry fish, such breaches of integrity do not impair survival. In fact, deliberately cut

sections of stem or leaf will grow back into entire cells. Severing of the membrane is unexpectedly devoid of serious consequence.

Yet another example of major insult lies within the domain of experimental genetics, where cells are routinely sectioned in order to monitor the fates of the respective fragments. When cultured epithelial cells are sectioned by a sharp micropipette, the non-nucleated fraction survives for 1–2 days, while the nucleated, centrosome-containing fraction survives indefinitely and can go on to produce progeny (Maniotis and Schliwa 1991). Sectioned muscle and nerve cells survive similarly (Yawo and Kuno 1985; Casademont et al. 1988; Krause et al. 1984), despite the absence of membrane resealing (Cameron 1988; Krause et al. 1984).

Finally, and perhaps not surprisingly in light of all that has been said, ordinary cells in the body are continually in a "membrane-wounded" state. Cells that suffer mechanical abrasion in particular, such as skin cells, gut endothelial cells, and muscle cells are especially prone to membrane wounds, as confirmed by passive entry into the cell of large tracers that ordinarily fail to enter. Yet, such cells appear structurally and functionally normal (McNeil and Ito 1990; McNeil and Steinhardt 1997). The fraction of wounded cells in different tissues is variable. In cardiac muscle cells it is ~20%, but the fraction rises to 60% in the presence of certain kinds of performance-enhancing drugs (Clarke et al. 1995). Thus, tears in the membrane occur commonly and frequently even in normal, functioning tissue, possibly due to surface abrasion. Membrane ruptures are perfectly normal.

7.3 Cells as Gels

Evidently, punching holes in the membrane does not wreak havoc with the cell even though the holes may be monumental in size relative to an ion. It appears we are stuck on the horns of a dilemma. If a continuous barrier envelops the cell and is consequential for function, one needs to explain why breaching the barrier is not more consequential than the evidence indicates. On the other hand, if we entertain the possibility that the barrier may be non-continuous – so that creating yet another opening makes little difference – we then challenge the dogma on which all mechanisms of cell biological function rest, for the continuous barrier concept has become axiomatic.

7.3.1 Is There an Escape?

If the cytoplasm is not an aqueous solution after all, then the need for a continuous barrier (with pumps and channels) becomes less acute. If the cytoplasm is a gel, for example, the membrane could be far less consequential. This argument does not necessarily imply that the membrane is absent – only that its continuity may not be essential for function. Such a paradigm could go a long way toward explaining the

aforementioned membrane-breach anomalies, for gels can be sliced with relative impunity. The gel-like cell might or might not tolerate major insults, depending on the nature of the insult and the degree of cytoplasmic damage inflicted. Death is not obligatory. A continuous barrier is not required for gel integrity – just as we have seen that a continuous barrier is not required for cell integrity. A critical feature of the cytoplasm, then, may be its gel-like consistency.

7.4   Cells, Gels and Water

Gels are built around a scaffold of long-chain polymers, often cross-linked to one another and invested with solvent. The cytoplasm is much the same. Cellular polymers such as proteins, polysaccharides, and nucleic acids are long chained molecules, frequently cross-linked to one another to form a matrix. The matrix holds the solvent (water) – which is retained even when the cell is de-membranated: "Skinned" muscle cells, for example, retain water in the same way as gels. Very much, then, the cytoplasm resembles an ordinary gel – as textbooks assert.

How the gel/cell matrix holds water is a matter of some importance (Rand et al. 2000), and there are at least two mechanistic possibilities. The first is osmotic: charged surfaces attract counter-ions, which draw in water. In the second mechanism, water-molecule dipoles adsorb onto charged surfaces, and subsequently onto one another to form multilayers. The first mechanism is unlikely to be the prevailing one because: (1) gels placed in a water bath of large-enough size should eventually be depleted of the counter-ions on which water retention depends; yet, the hydrated gel state is retained; and (2), cytoplasm placed under high-speed centrifugation loses ions well before it loses water (Ling and Walton 1976).

The second hypothesis, that). The thesis is that water can build layer upon layer (Fig. 7.1). This charged surfaces attract water dipoles in multilayers, is an old one (Ling 1965) view had been controversial at one time, but it has been given support by several groundbreaking observations. The first is the now-classical observation by Pashley and Kitchener that polished quartz surfaces placed in a humid atmosphere will adsorb films of water up to 600 molecular layers thick (Pashley and Kitchener 1979); this implies adsorption of a substantial number of layers, one upon another. The second observations are those of Israelachvili and colleagues, who measured the force required to displace solvents sandwiched between closely spaced parallel mica surfaces (Horn and Israelachvili 1981; Israelachvili and McGuiggan 1988; Israelachvili and Wennerström 1996). The overall behavior was largely classical, following so-called DLVO theory. However, superimposed on the anticipated monotonic response was a series of regularly spaced peaks and valleys (Fig. 7.2). The spacing between peaks was always equal to the molecular diameter of the sandwiched fluid. Thus, the force oscillations appeared to arise from a layering of molecules between the surfaces.

Although the Israelachvili experiments confirm molecular layering near charged surfaces, they do not prove that the molecules are linked to one another in the
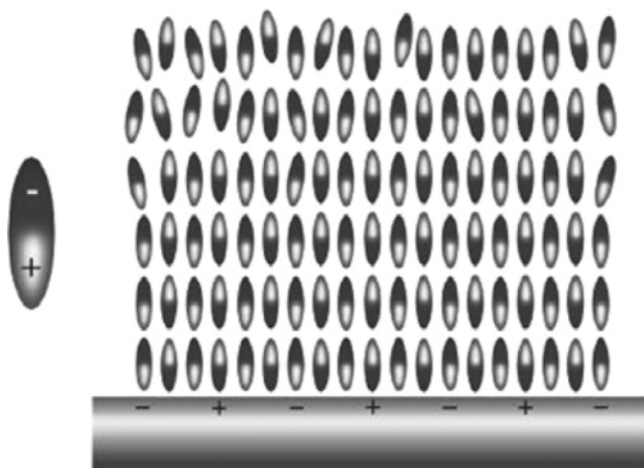
Fig. 7.1 Organization of water molecules adjacent to charged surface
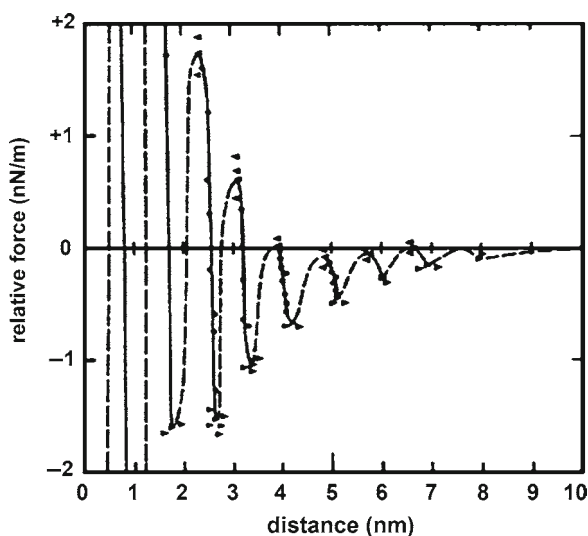
Fig. 7.2 Effect of separation on force between closely spaced mica plates. Only the oscillatory part of the response is shown (After Horn and Israelachvili 1981)

manner implied in Fig. 7.1. However, more recent experiments using carbon nano-tube-tipped AFM probes approaching flexible monolayer surfaces in water show similar layering (Jarvis et al. 2000), implying that the ordering does not arise merely from packing constraints; and, the Pashley/Kitchener experiment implies that many layers are possible. Hence, the kind of layering diagrammed in Fig. 7.1 is collectively implied by these experiments.
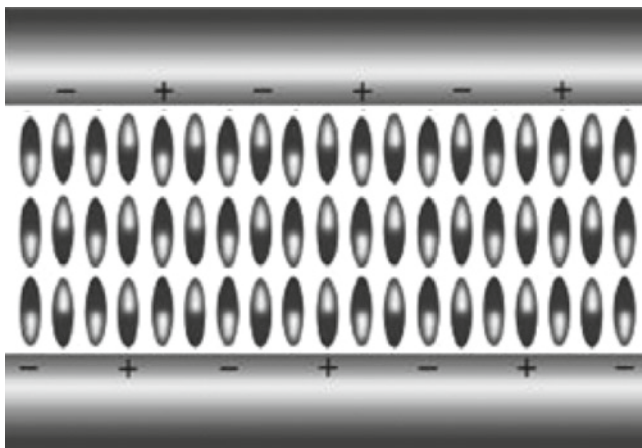
Fig. 7.3 Structured water dipoles effectively "glue" charged surfaces to one another

When two charged polymeric surfaces lie in proximity of one another, the interfacial water layers can bond the surfaces much like glue (Fig. 7.3). This is revealed in common experience. Separating two glass slides stacked face-to-face is no problem; however, when the slides are wet separation is formidable – sandwiched water molecules cling tenaciously to the glass surfaces and to one another, preventing separation. A similar principle holds in sand: A foot will ordinarily sink deeply into dry sand at the beach, leaving a large imprint; but in wet sand, the imprint is shallow. Water clings to the sand particles, bonding them together with enough strength to support one's full weight.

The picture that emerges, then, is that of a cytoplasmic matrix very much resembling a gel matrix. Water molecules are retained in both cases because of their affinity for the charged (hydrophilic) surfaces and their affinity for one another. The polymer matrix and adsorbed water largely make up the gel. This explains why de-membranated cells retain their integrity.

Embodied in this gel-like construct are many features that have relevance for cell function. An important one is ion partitioning. According to the prevailing view, the explanation for the ion gradients found between extracellular and intracellular compartments lies in a balance between passive flow through channels and active transport by pumps. Thus, the low sodium concentration inside the cell relative to outside is presumed to arise from the activity of sodium pumps, which transport sodium ions against their concentration gradient from the cytoplasm across the cell membrane. The gel construct invites an alternative explanation. It looks toward differences of solubility between extracellular bulk water and intracellular layered, or "structured" water – as well as differences of affinity of various ions for the cell's charged polymeric surfaces (Ling 1992). Na^+ has a larger hydrated diameter than K^+, and is therefore more profoundly excluded from the cytoplasm than K^+; and, because the hydration layers require more energy to remove from Na^+ than from K^+, the latter has higher affinity for the cell's negatively charged polymeric surfaces.

Hence, the cytoplasm has considerably more potassium than sodium. A fuller treatment of this fundamental biological feature is given in the recent book by this author (Pollack 2001).

Similarly, the gel construct provides an explanation for the cell potential. The cell is filled with negatively charged polymers. These polymers attract cations. The number of cations that can enter the cell (gel) is restricted by the cations' low solubility in structured water. Those cations that do enter compete with water dipoles for the cell's fixed anionic charges. Hence, the negative charge in the cell is not fully balanced by cations. The residual charge amounts to approximately 0.3 mol/kg (Wiggins 1990). With net negative charge, the cytoplasm will have a net negative potential. Indeed, depending on conditions, membrane-free cells can show potentials as large as 50 mV (Collins and Edwards 1971). And, gels – with no membrane at all – made of negatively charged polymers show comparable or larger negative potentials, while gels built of positively charged polymers show equivalent positive potentials (Fig. 7.4). Membranes, pumps, and channels evidently play no role in these latter potentials because they are absent. Yet, the electrical potential is similar to that of the cell.
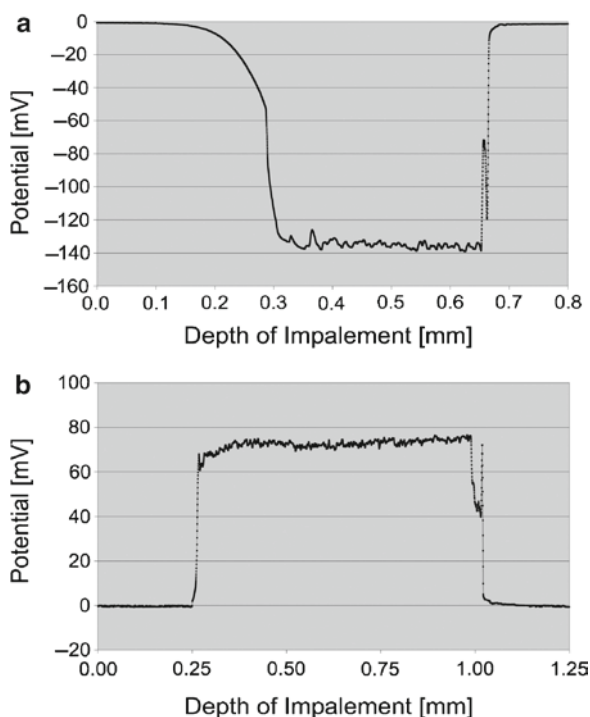
Fig. 7.4 KCl-filled microelectrodes stuck into gel strips at slow constant velocity, and then withdrawn. (a) Typical anionic gel, polyacrylamide/polypotassiumacrylate, shows negative potential. (b) Typcial cationic gel, polyacrylamide/ polydialyldimethylammonium chloride, shows positive potential (Courtesy, R. Gülch)

Hence, the gel paradigm can go quite far in explaining the cell's most fundamental attributes – distribution of ions, and the presence of an electrical potential. The book *Cells, Gels and the Engines of Life* (Pollack 2001) goes on to show that the gel paradigm shows good promise of explaining many of the cell's basic functional and dynamic features as well. All of them depend centrally on the role of interfacial water.

7.5   Interfacial Water and Exclusion Zones

Following up the seemingly critical role of interfacial water in cell function, we found more recently that this water had some unexpected and highly implicative features. We found that it profoundly excluded colloidal and molecular solutes and that under certain circumstances the zone containing this water could extend to distances as much as several hundred micrometers from the nucleating surface (Zheng and Pollack 2003). Numerous controls showed that this was not the trivial result of some unsuspected artifact. By now experiments showing such long-range exclusion have been repeated in many laboratories, and it recently came to our attention that similar results had been reported 4 decades ago (Green and Otori 1970). These extensive solute-free regions are termed "exclusion zones."

Exclusion zones are observed next to many surfaces including hydrogels, hydrophilic polymers, monolayers, ion-exchange beads and biological tissues. They exclude a diverse array of solutes of various size, type, and polarity. Hence, the exclusion phenomenon is a general feature of water adjacent to hydrophilic surface (Zheng et al. 2006; Zheng and Pollack 2006).

To test whether the physical properties of exclusion-zone water differ from those of bulk water, six methods have thus far been applied. They include the following: NMR, infrared, and birefringence imaging; and, measurements of electrical potential, viscosity, and UV-VIS absorption spectra. The results confirm that the interfacial zone differs physically from the bulk zone, and that the former is a distinct, less mobile, more ordered phase of water that can coexist with the contiguous bulk phase (Zheng et al. 2006; Pollack and Clegg 2008).

In all likelihood, then, these ordered exclusion zones are the same as the structured water inside cells. Both occur next to charged or hydrophilic surfaces, both exclude solutes, and both can extend rather far, although those inside the cell are obviously restricted from growing as large as those that can be seen next to flat polymeric surfaces in large chambers.

7.6   Charge Separation and Energy

Of the aforementioned exclusion-zone features, a particularly significant one is charge separation. Although the overall net charge may remain zero, water in the exclusion zone is negatively charged, while the bulk-water region beyond the

exclusion zone contains positive charges. The potential difference between the two regions is on the order of 100–200 mV, depending on the particular hydrophilic surface nucleating the zone. Electrodes inserted into the respective zones and connected through a resistor show ample, persisting, current flow (Pollack and Clegg 2008).

Charge separation within water itself may seem counterintuitive, but it is surprisingly common. In cloud water for example, evidence of charge separation lies in the consequent lightning discharges, 80% of which occur from cloud to cloud. In the laboratory, the famous Kelvin water-dropper experiment shows electric discharge between juxtaposed bodies of pure water, again indicating water-based charge separation.

Charge separation in water has also been demonstrated in our laboratory. In a chamber filled with pure water, electrodes placed at either end of the chamber and driven at low voltage to pass current through the water, create ample charge buildup: the half of the water bath nearest the anode remains positive, while the half nearest the cathode remains negative, even well after the electrical driver has been disconnected (Klimov and Pollack 2007). Such stored water-based charge turns out to be largely recoverable (Ovchinnikova and Pollack 2009).

These three examples offer precedent for charge separation in water, and give credence to the similar phenomenon observed in water near hydrophilic surfaces.

For building these charged exclusion zones, the required energy apparently comes from light. We found that incident radiant energy, including UV, visible, and near-infrared wavelengths, induces exclusion-zone growth. IR is especially effective: at wavelength 3.1 μm, a 10-min exposure to incident radiation – weak enough to elevate water temperature by no more than 1°C – causes the exclusion-zone width to increase by up to four times (Chai et al. 2009). This is an impressive increase for a relatively modest energy input, and it goes without saying (see below) that such solar-based or geothermal-based energy is likely to have relevance for the origin of life.

The molecular mechanism by which the exclusion zone expands is not fully understood. Incident photons must interact in some way with bulk water to build this zone. Water is naturally dissociated to some extent into H^+ and OH^- (the basis of pH measurement); and, it is possible that incident photons enhance this natural dissociation near hydrophilic surfaces. The negatively charged moiety would then go on to build the negatively charged exclusion zone, leaving the bulk zone positively charged with protons or hydronium ions. In effect, this process amounts to a light-energy driven proton pump – a process surely useful for an incipient cell.

Photons from sunlight or geothermal sources, then, may exert an effect that goes well beyond mere heating. Solar energy incident on water evidently separates charge. Such light-induced near-surface water splitting is reminiscent of what happens in reaction centers as the first step of photosynthesis. Hence, the charge-separation phenomenon identified here may be something akin to an initial step in photosynthesis – a kind of "generic" first step that occurs next to many hydrophilic surfaces rather than only next to those specific to green plants or

bacteria. This process could represent not only a potentially significant energetic pathway that may be broadly relevant for nature, but also as argued below, a central protagonist for the origin of life.

7.7 Exclusion Zones and Protons

A prime attribute of the interfacial exclusion zone is separation of charge. The zone itself is ordinarily negative, while the region beyond is positive. This positive potential is measured directly and is also consistent with pH measurements, which show an extreme drop of pH immediately beyond the exclusion zone, often to less than pH 3. This drop is observed not only with pH-sensitive dyes, but also with miniature pH probes (Chai et al. 2009). Hence, so long as the exclusion zone builds, proton concentration will grow in the bulk water beyond.

But negative and positive zones are physically different. The negative charges are found within the exclusion zone, which is fixed in space. The positive charges, by contrast, are unconstrained and free to diffuse in the region beyond: These protons (or hydronium ions) will move depending on local electrical gradients. Essentially, these free positive charges arise as a byproduct of the environmental energy that builds the exclusion zone, and are free to accomplish whatever tasks their presence may require.

What might those tasks include? Might they relate to life's origins?

7.8 Like-Likes-Like

In water, similarly charged particles paradoxically attract one another. Feynman referred to this counterintuitive yet long-recognized phenomenon as "like-likes-like," theorizing that the paradoxical attraction of like-charged particles lay in abundant "unlikes" situated between the like (Feynman et al. 1963). Thus, for negatively charged particles, the attraction would occur because of interposed positive charges, whereas for positively charged particles the required interposed charges would be negative.

Sogami and Ise (1984), took a similar theoretical route, mainly based on the extensive experimental work by Ise and colleague (Ise and Okubo 1980; Ise and Okubo 1983). Ise's experiments showed not only that like-charged colloidal particles attract one another, but also that as the particles drew closer, the attraction was increasingly balanced by an inter-particle repulsion; this balance yielded stable colloid crystals, within which elements were regularly spaced (Ise 2007; Ito et al. 1994). Again, the attractive force between like-charged entities was considered to arise as a result of unlike charges in between. And, an equally crucial feature of this attraction was the development of order.

like-likes-like (Feynman)
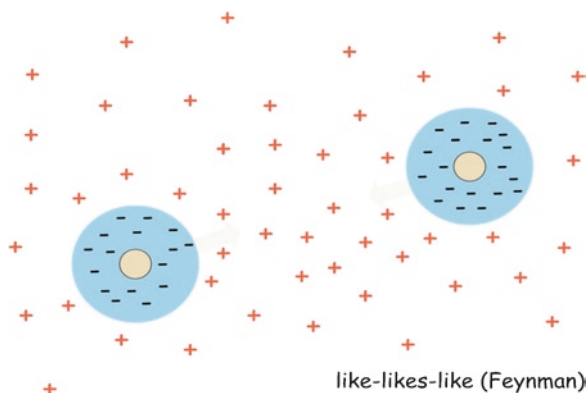
Fig. 7.5 High concentration of protons in region between negatively charged entities creates an attractive force, which draws the entities together

On the other hand, the source of the required unlike charges has remained uncertain. Counter-ions have been considered as the likely source, but whether they might occur in adequate concentration has been less clear. A possibility that we investigated is that they might come from exclusion-zone formation. In the case of negatively charged entities such as the majority of biological substances, coalescence would require interposed positive charges. Thus, protons arising out of exclusion-zone formation could thus constitute the required "unlikes."

How such a mechanism might work is as follows (Fig. 7.5). Consider a suspension of negatively charged entities such as microspheres or other colloidal particles in water. As a result of incident radiant energy, each microsphere develops a negatively charged exclusion zone, thereby generating free protons in bulk water. The exclusion zones take the form of microsphere-enveloping shells, while the protons are free to diffuse from these shells. The positive charges will then accumulate in between the microspheres, constituting the required unlikes that draw the microspheres toward one another. Coalescence continues until the attraction becomes balanced by inter-particle repulsion, at which point an ordered array is formed.

If light is the ultimate energizing agent, then by increasing the light intensity, more protons should be generated, and elements of the ordered crystal should draw nearer to one another. That is what the experimental results show: incident light diminishes the microsphere-to-microsphere distance (Zhao et al. 2008).

A direct test of the like-likes-like through-an-intermediate-of-unlikes hypothesis was carried out using ion-exchange beads (Nagornyak et al. 2009). Conducting this test was possible because attraction between beads could be seen even on large, experimentally accessible, scales. Thus, suspended in water, gel beads on the order of 0.5 mm diameter attracted one another at surface-to-surface separations as much as 400 μm. In the case of negatively charged beads, pH measurements showed a high proton concentration in between the beads, supported by microelectrode measurements that showed a relatively positive potential in the inter-bead region.

Substituting positively charged beads for the negative ones gave a result that was polarity-opposite.

Hence, fresh experimental evidence confirms Feynman's theoretical expectation: like-likes-like through an intermediate of unlikes. Like-charged entities attract; and, they become ordered.

7.9   Biological Coalescence and Origin of Life

We return now to the issue of the origin of life. The first step is likely to have been some kind of coalescence of substances. Substances initially widely dispersed must somehow have been brought together to form a gel-like or colloid-like proto-cell. How then might such coalescence happen?

A plausible explanation comes from the like-likes-like mechanism. Envision simple, identical molecules scattered widely about, dissolved in shallow water or in deeper water near a hydrothermal vent. These molecules could be simple sugars or short-chain amino acids – the latter being extremely primitive and found even in space and on the moon (Bernstein et al. 2002; Fox et al. 1981). Radiant energy from the sun (or from a geothermal vent) beats down upon these primitive molecules. This energy builds exclusion zones around each one. If the exclusion zones are negatively charged, then their growth will generate free protons. These protons will then draw these molecules together, forming larger, ordered entities.

This kind of coalescence is experimentally confirmed. Radiant energy (heat) demonstrably causes amino acids to coalesce into highly ordered polyamino acids, called "thermal proteins." These protein-like structures in turn self assemble into distinct, microsphere-like entities (Fox 1986), which easily classify as coascervates or gels. Such globular entities exhibit a striking array of cell-like features such as growth, breaking apart, selective accumulation and exclusion of organic substances, and even electrical action potentials (Fox 1986; Przybylski et al. 1982; Nakashima and Fox 1980; Fox 1980; Fox 1986; Fox 1991). Hence, they qualify as proto-cells.

Indeed, the polyamino-acid filaments of these protocells could well serve as primitive substrates for inheritance – as could the ordered water immediately surrounding them. Information is available, and could be passed along.

At the same time, proto-cells with differing polyamino-acid content could aggregate to form multi-unit entities. Such entities would be similar to multi-cellular entities. Thus, a simple energy-based mechanism (solar or geothermal), operating through the like-likes-like principle, can produce coalescence, order, and even function at multiple levels of organization.

If these entities are produced by the like-likes-like principle, then they ought to contain large volumes of negatively charged, ordered water. The books by Pollack and Ling (Ling 1992; Pollack 2001) review the evidence that both gels and cells are filled with ordered water, or equivalently, exclusion-zone water.

Also of interest is the observation that the potential difference between the inside and outside of a gel/coaservate is similar to the potential difference between inside and outside of a cell. This was mentioned above. Given the negative electrical potential typical of most exclusion zones, a possibility is that the water itself contributes to the cell's negative potential. Hence, ordered water might play roles beyond that of bringing molecules into coalescence.

The emphasis so far has been on entities of like charge, which attract one another. This narrow focus should not be interpreted as suggesting that the ordinary attraction between unlike charges is not significant; in fact, it is central to the like-likes-like mechanism. The point is that whether entities are of opposite polarity or the same polarity, the evidence indicates that they will attract. Thus, counter-intuitively, sun's energy tends to drive all entities into coalescence.

For driving the coalescence necessary for life's origin, then, a simple and primitive mechanism is available. Solar or geothermal radiation separates charge in water, which then draws dissolved or suspended entities into ordered coalescence.

7.10 Conclusion: Is Life's Origin a One-Time Event?

Although the exact date has been subject to occasional revision, the prevailing view is that life began some 3.5 billion years ago. A unique coalescence of events is thought to have taken place at that time, which gave rise to the first pre-cellular entity. Out of that entity grew life, as we know it.

Another possibility is that life continues to originate. If the above-mentioned condensation mechanism is indeed the first step, then that step might be occurring continually – for the required conditions are omnipresent. All that is necessary are primitive molecules, water, and the sun's energy.

In such a scenario, the biosphere consists of living entities at various stages of evolutionary development. The most sophisticated we know well. The simplest are proto-cells that are no more complex than primitive gels or coaservates. Such proto-cells would then evolve with time into progressively more complex entities.

Quite remarkable is the extent to which such entities can resemble cells. Gels/coaservates can have negative electrical potentials just like cells; they absorb energy from the environment just like cells; and, like cells, their charged constituents move, here driven by energy from the environment. These constituents, including both water and dissolved/suspended substances, can be ordered, much like the cell. And, they can also grow like the cell: the like-likes-like mechanism expands their size so long as environmental energy is present to drive the expansion. Hence, these primitive gels/coaservates are more cell-like than one might initially suppose. The line of distinction between gels/coaservates and cells is somewhat blurred.

Perhaps the most central feature of all of the developments outlined above is water. Cell water is generally thought of as the background carrier of the more important molecules of life. The developments outlined here imply otherwise.

Water is a molecule that participates centrally in cellular chemistry. Its role is so central as to explain not only the origin of the pre-cell, but also many features of more sophisticated and more highly differentiated cells.

References

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994) Molecular biology of the cell, 3rd edn. Garland, NY

Bernstein MP et al (2002) Racemic amino acids from the ultraviolet photolysis of interstellar ice analogues. Nature 416(6879):401–403

Berry H, Pelta J, Lairez D, Larreta-Garde V (2000) Gel-sol transition can describe the proteolysis of extracellular matrix gels. Biochim Biophys Acta 1524(2–3):110–117

Cameron I (1988) Ultrastructural observations on the transectioned end of frog skeletal muscles. Physiol Chem Phys Med NMR 20:221–225

Frey-Wyssling A (1953) Submicroscopic morphology of protoplasm. Elsevier, Amsterdam

Casademont J, Carpenter S, Karpati G (1988) Vacuolation of muscle fibers near sarcolemmal breaks represents T tubule dilation secondary to enhanced sodium pump activity. J Neuropath Exp Neurol 47:618–628

Chai B, Yoo H, Pollack GH (2009) Effect of radiant energy on near-surface water. J Phys Chem B 113:13953–13958

Clarke MSF, Caldwell RW, Miyake K, McNeil PL (1995) Contraction-induced cell wounding and release of fibroblast growth factor in heart. Circ Res 76:927–934

Collins EW Jr, Edwards C (1971) Role of Donnan equilibrium in the resting potentials in glycerol-extracted muscle. Am J Physiol 22(4):1130–1133

Feynman R, Leighton R, Sands M (1963) The Feynman lecture on Physics. Addison-Wesley, Reading, MA

Fox SW (1980) Metabolic microspheres: origins and evolution. Naturwissenschaften 67(8):378–383

Fox SW (1986a) Molecular selection in a unified evolutionary sequence. Int J Quantum Chem Quantum Biol Symp 13:223–235

Fox SW (1986) Molecular selection and natural selection. Quart Rev Biol 61(3): 375–386. CR – Copyright © 1986. The University of Chicago Press, Chicago, IL

Fox S (1991) Synthesis of life in the lab? Defining a protoliving system. Quart Rev Biol 66(2):181–185

Fox SW, Harada K, Hare PE (1981) Amino acids from the moon: notes on meteorites. Subcell Biochem 8:357–373

Green K, Otori T (1970) Direct measurements of membrane unstirred layers. J Physiol 207(1):93–102

Hochachka PW (1999) The metabolic implications of intracellular circulation. Proc Natl Acad Sci U S A 96(22):12233–12239

Horn RG, Israelachvili JN (1981) Direct measurement of astructural forces between two surfaces in a nonpolar liquid. J Chem Phys 75(3):1400–1411

Ise N, Okubo T (1980) "Ordered" distribution of electrically charged solutes in dilute solutions. Acc Chem Res 13:303

Ise N, Okubo T (1983) Ordered structure in diluted solutions of highly charged polymer latices as studied by microscopy. Chem Phys 78:536

Ise N (2007) When, why and how does like-like-like? Jpn Acad Ser B(83)

Israelachvili JN, McGuiggan PM (1988) Forces between surfaces in liquids. Science 241:795–800

Israelachvili JN, Wennerström H (1996) Role of hydration and water structure in biological and colloidal interactions. Nature 379:219–225

Ito K, Yoshida H, Ise N (1994) Void structure in colloidal dispersions. Science 263(5143):66–68

Jacobs WP (1994) Caulerpa. Sci Amer 100–105

Janmey PA, Shah JV, Tang JX, Stossel TP (2001) Actin filament networks. Results Probl Cell Differ 32:181–199

Jarvis SP et al (2000) Local solvation shell measurement in water using a carbon nanotube probe. J Phys Chem B 104:6091–6097

Jones DS (1999) Dynamic mechanical analysis of polymeric systems of pharmaceutical and biomedical significance. Int J Pharm 179(2):167–178

Klenchin VA, Sukharev SI, Serov SM, Chernomordik LV, Chizmadzhev YA (1991) Electrically induced DNA uptake by cells is a fast process involving DNA electrophoresis. Biophys J 60(4):804–811

Klimov A, Pollack GH (2007) Visualization of charge-carrier propagation in water. Langmuir 23(23):11890–11895

Krause TL, Fishman HM, Ballinger ML, Ballinger GD, Bittner GD (1984) Extent mechanism of sealing in transected giant axons of squid and earthworms. J Neurosci 14:6638–6651

Ling GN (1965) The physical state of water in living cell and model systems. Ann NY Acad Sci 125:401

Ling GN, Walton CL (1976) What retains water in living cells? Science 191:293–295

Ling GN (1992) A revolution in the physiology of the living cell. Krieger, Malabar, Fl

Maniotis A, Schliwa M (1991) Microsurgical removal of centrosomes blocks cell reproduction and centriole generation in BSC-1 cells. Cell 67:495–504

McNeil PL, Ito S (1990) Molecular traffic through plasma membrane disruptions of cells in vivo. J Cell Sci 67:495–504

McNeil PL, Steinhardt RA (1997) Loss, restoration, and maintenance of plasma membrane integrity. J Cell Bio 137(1):1–4

Nagornyak K, Yoo H, Pollack GH (2009) Mechanism of attraction between like-charged particles in aqueous solution. Soft Matter 5:3850–3857

Nakashima T, Fox SW (1980) Synthesis of peptides from amino acids and ATP with lysine-rich proteinoid. J Mol Evol 15(2):161–168

Ovchinnikova K, Pollack GH (2009) Can water store charge? Langmuir 25:542–547

Pashley RM, Kitchener JA (1979) Surface forces in adsorbed multilayers of water on quartz. J Colloid Interface Sci 71:491–500

Prausnitz MR, Milano CD, Gimm JA, Langer R, Weaver JC (1994) Quantitative study of molecular transport due to electroporation: uptake of bovine serum albumin by erythrocyte ghosts. Biophys J 66(5):1522–1530

Pollack GH (2001) Cells, gels and the engines of life: a new, unifying approach to cell function. Ebner and Sons, Seattle

Pollack GH, Clegg J (2008) Unsuspected linkage between unstirred layers, exclusion zones and water. In: Pollack GH, Chin W-C (eds) Phase transitions in cell biology, p 183. Springer, New York

Przybylski AT et al (1982) Membrane, action, and oscillatory potentials in simulated protocells. Naturwissenschaften 69(12):561–563

Rand RP, Parsegian VA, Rau DC (2000) Intracellular osmotic action. Cell Mol Life Sci 57(7):1018–1032

Schwister K, Deuticke B (1985) Formation and properties of aqueous leaks induced in human erythrocytes by electrical breakdown. Biophys Acta 816(2):332–348

Serpersu EH, Kinosita K Jr, Tsong TY (1985) Reversible and irreversible modification of erythrocyte membrane permeability by electric field. Biochim Biophys Acta 812(3): 779–785

Sogami I, Ise N (1984) On the electrostatic interaction in macroionic solutions. J Chem Phys 81:6320

Taylor SR, Shlevin HH, Lopez JR (1975) Calcium in excitation-contraction coupling of skeletal muscle. Biochem Soc Transact 7:759–764

Wiggins PM (1990) Role of water in some biological processes. Microbiol Rev 54(4):432–449

Xie TD, Sun L, Tsong TY (1990) Studies of mechanisms of electric field-induced DNA transfection. Biophys J 58:13–19

Yawo H, Kuno M (1985) Calcium dependence of membrane sealing at the cut end of the cockroach giant axon. J Neurosci 5:1626–1632

Zhao Q et al (2008) Unexpected effect of light on colloidal crystal spacing. Langmuir 24(5):1750–1755

Zheng JM, Chin W-C, Khijniak E, Khijniak E Jr, Pollack GH (2006) Surfaces and interfacial water: evidence that hydrophilic surfaces have long-range impact. Adv Coll Inter Sci 127(1):19–27

Zheng JM, Pollack GH (2003) Long-range forces extending from polymer-gel surfaces. Phys Rev E 68(3 Pt 1):031408

Zheng J, Pollack GH (2006) Solute exclusion and potential distribution near hydrophilic surfaces. In: Pollack GH, Cameron IL, Wheatley DN (eds) Water and the cell, pp 165–174. Springer, New York

Chapter 8
Membrane Self-Assembly Processes: Steps Toward the First Cellular Life*

Pierre-Alain Monnard and David W. Deamer

Abstract This review addresses the question of the origin of life, with emphasis on plausible boundary structures that may have initially provided cellular compartmentation. Some form of compartmentation is a necessary prerequisite for maintaining the integrity of interdependent molecular systems that are associated with metabolism, and for permitting variations required for speciation. The fact that lipid-bilayer membranes define boundaries of all contemporary living cells suggests that protocellular compartments were likely to have required similar, self-assembled boundaries. Amphiphiles such as short-chain fatty acids, which were presumably available on the early Earth, can self-assemble into stable vesicles that encapsulate hydrophilic solutes with catalytic activity. Their suspensions in aqueous media have therefore been used to investigate nutrient uptake across simple membranes and encapsulated catalyzed reactions, both of which would be essential processes in protocellular life forms.

8.1 Introduction

This chapter addresses the question of the origin of life, with emphasis on plausible boundary structures that initially provided cellular compartments. The emergence of life on the early Earth required the presence of at least three different substances

and related physical properties: liquid water, a source of free energy, and organic compounds capable of self-assembly. Liquid water is essential for all life today, and it is highly implausible that life, as known on Earth, can exist in its absence. Possible energy sources include sunlight, if life began on the Earth's surface, or energy arising from chemical disequilibria in submarine or subterranean sites. Self-assembling compounds must have been available to provide building blocks for polymer synthesis and formation of boundary structures.

The first forms of life were represented by self-assembled molecular systems having a specific set of chemical and physical properties, and it is worth listing these here to provide a foundation for later discussion. First, the system must have defined boundaries that separate it from its environment (Deamer et al. 1994; Lazcano 1994a, b; Tawfik and Griffiths 1998; Luisi et al. 1999; Szostak et al. 2001; Deamer et al. 2002; Luisi 2006). A specific set of catalyzed metabolic and polymerization reactions must occur in the encapsulated volume, which implies effective exchange of nutrients and energy from the environment. Perhaps most central to the definition of life is that the entire molecular system must be able to reproduce itself using self-assembly of components and genetically coded polymerization reactions. This capacity implies information transfer between molecules within a cell, then from one generation to the next as the cell reproduces (Varela et al. 1974; Luisi and Varela 1989). Finally, it must be possible for small changes to be introduced into the general components that direct polymerization. In a population of reproducing systems that compete for energy and nutrients, the changes produce variations in individual molecular systems that affect the efficiency of growth and reproduction. A population of bounded molecular systems can then undergo a variety of selective processes required for Darwinian evolution.

Our research has focused on the nature of boundary structures that defined the first cellular life. Such structures are required for speciation, for energy capture and transduction, and for development of the complex network of catalytic reactions associated with metabolism (Deamer and Oro 1980; Deamer 1997; Szostak et al. 2001). In contemporary cells, a fundamental role of membrane boundaries is to provide a selective permeability barrier that is necessary for separating the cytoplasm from the external environment. The transmembrane transport of nutrients and ionic solutes is mediated by a variety of membrane-associated proteins that act as channels, carriers and active transporters (pumps). Membrane receptors provide a sensor mechanism that permits communication between the intracellular milieu and the outside world. Membranes also capture light energy and redox energy by using pigment systems and electron transport to generate electrochemical proton gradients as a source of free energy.

All of these functions require membrane-associated proteins which were presumably absent in the first forms of cellular life. It therefore seems likely that the membrane boundaries of the earliest cells simply provided a selective permeability barrier that permitted the permeation of essential nutrients but retained polymeric products of primitive biosynthesis. This concept has guided our research over the past decade, and is the main theme of what we will discuss here.

8.2 Models of Protocellular Compartments

The concept that life began on the early Earth as self-assembled structures of organic material was first presented by Oparin (1924). Laboratory investigations of such structures began in the 1950s within the framework of "chemical evolution as a transition to life" (Oparin 1957). At that time the role of membranes as boundary structures had not yet been established. Instead it was believed that a living cell could be understood as a collection of aggregated colloidal particles. Therefore, Oparin and his coworkers prepared heterogeneous, spherical aggregates from macromolecular components, such as gum Arabic, and gelatin or histone which could provide localized sites for enzymatic reactions (Oparin et al. 1976). These aggregates, called coacervates, were not intrinsically stable, and their molecular compositions were highly variable. Furthermore, coacervates have no permeability barrier, so they lacked the capacity for encapsulated metabolism and accumulation of biosynthetic products.

The first suggestion that membranes played a role in the origin of life was put forward by Haldane (1929) who wrote that "The cell consists of numerous half-living chemical molecules suspended in water and enclosed in an oily film. When the whole sea was a vast chemical laboratory the conditions for the formation of such films must have been relatively favourable…" Goldacre (1958) proposed that the first membranes could have been produced by wave action disturbing films of lipid-like surfactants.

As we learned more about the role of membranes in defining cell structures, it became clear that all membranes incorporated lipid bilayers as the primary permeability barrier, and that phospholipid is a nearly ubiquitous amphiphilic component of the bilayer. Bangham and co-workers (1965) first demonstrated that phospholipids spontaneously form bilayer vesicles having dimensions in the range of bacterial cells. Lipid bilayer vesicles are commonly referred to as liposomes, and such self-assembled membrane structures can be used as models of the earliest cell membranes. The first question we will address concerns the nature of the lipid-like compounds available on the early Earth. One possibility is that phospholipids were synthesized during prebiotic chemical evolution. In fact, several early papers demonstrated that phospholipids could be synthesized under simulated prebiotic conditions from mixtures of fatty acids, glycerol, and phosphate (Hargreaves et al. 1977; Oro et al. 1978). However, the simultaneous presence of all three components on the early Earth seems improbable, and we have therefore turned our attention to simpler membranogenic amphiphiles.

8.3 Stability and Permeability of Amphiphile Vesicles

Although the ability of phospholipids to self-assemble into membranous vesicles is common knowledge, it is less well known that a variety of membranous structures can also be prepared from single-chain amphiphiles such as fatty acids (see Fig. 8.1),
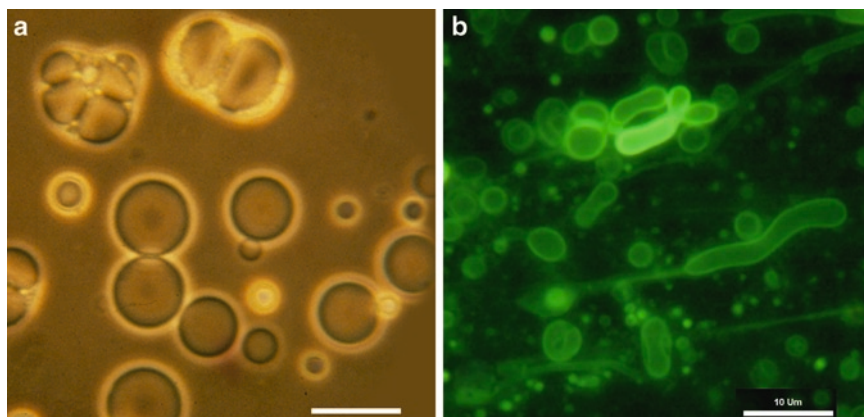
Fig. 8.1 Primitive membrane structures visualized by light microscopy. (**a**) Amphiphilic compounds extracted from the Murchison meteorite form membranous vesicles when exposed to dilute aqueous salt solutions at pH > 7.0. The probable components of the vesicles are monocarboxylic acids ranging from 8–11 carbons in length together with admixtures of polycyclic aromatic hydrocarbon derivatives. (**b**) Monocarboxylic acids in pure form also self-assemble into membranous vesicles, as shown here for decanoic acid:decanol (37 mM:3mM, C10, pH 7.4) stained with rhodamine 6G and observed by epifluorescence microscopy. This dye inserts in the bilayer membranes which appear green. Bar = 10 μm

fatty alcohols and monoglycerides. We will argue that such vesicles are plausible models for the formation of early cellular compartments. An important aspect of this argument is that the prebiotic availability of such amphiphiles has been established. Carbonaceous meteorites contain a rich mixture of organic compounds that were synthesized abiotically in the early solar system, and this mixture can be used as a guide to the kinds of organics likely to be available on the early Earth, either delivered during late accretion or synthesized at the Earth's surface. For example, Miller first demonstrated that amino acids are synthesized in mixtures of reduced gases that are chemically activated by impinging sources of free energy such as electrical discharge (Miller 1953). The conjecture that similar reactions could occur in the early solar system was confirmed by the discovery of a variety of amino acids in the Murchison meteorite (Kvenvolden et al. 1970).

A similar argument can be made for monocarboxylic acids. These have been synthesized under a variety of simulated prebiotic conditions (Deamer and Oro 1980; McCollom et al. 1999; Rushdi and Simoneit 2001). Furthermore, monocarboxylic acids ranging from 2 to 12 carbons in length are abundant components of the organic mixture present in the Murchison meteorite (Lawless and Yuen 1979; Komiya et al. 1993; Mautner et al. 1995). It has also been established that certain components of the Murchison organics are amphiphiles and have the capacity to assemble into membranous vesicles (Deamer 1985; Deamer and Pashley 1989). Figure 8.1 shows several examples of such vesicles by light microscopy, and it is clear that certain organic components have the capacity to assemble into recognizable membranes. The presence of a permeability barrier is confirmed by the fact that

such vesicles can capture and maintain concentration gradients of pyranine, an anionic fluorescent dye marker (Apel et al. 2002).

Although the composition of the membrane-forming amphiphiles present in the Murchison organic mixture has not yet been established in detail, it is clear that substantial amounts of monocarboxylic acids are present (Mautner et al. 1995). For this reason we have begun to investigate the physical properties of self-assembled structures produced by monocarboxylic acids of various chain lengths, and of mixtures with other simple amphiphilic compounds. Gebicki and Hicks (1973; 1976) first established that a fatty acid, oleic acid, forms vesicular structures. Since this discovery, the bilayer-forming potential of fatty acids with shorter hydrocarbon chains (C8-C11) has been investigated (Hargreaves and Deamer 1978; Apel et al. 2002; Monnard et al. 2002; Maurer et al. 2009). The length and the degree of unsaturation of the hydrocarbon chains play an important role in determining the bilayer-membrane properties, such as permeability, and stability, which would have been essential for primitive life forms.

In suspensions of fatty acids at concentrations above the critical bilayer concentration (CBC, by analogy to the critical micelle concentration, CMC), fatty acid bilayer membranes (see (1) in Fig. 8.2) are stabilized by van der Waal interactions between their hydrocarbon chains and by hydrogen bonds formed between deprotonated and protonated acid molecules (Rosano et al. 1969; Haines 1983; Apel et al. 2002). For this reason the formation of bilayer vesicles is highly sensitive to pH (see Fig. 8.3). Below the apparent pK_a of the acid inserted in a structure (this pK_a is higher than that of single acid molecule), the fatty acid droplets will replace the bilayer structures, and above it micelles will form. In addition, CBC increases with decreasing length and degree of unsaturation of hydrocarbon chains. For instance, octanoic acid (C8:0, the shortest vesicle-forming carboxylic acid) decanoic acid (C10:0), oleic acid (C18:1) at pH 7 form bilayers at amphiphile concentrations higher than 145, 10, and 0.19 mM, respectively (Maurer et al. 2009).

This tendency becomes significant when the prebiotic availability of amphiphiles is considered. As discussed earlier, GC-MS analysis of the fatty acid content of the Murchison meteorites shows that fatty acids with hydrocarbon chain length from C8-C12 are present. Reaction products of Fischer-Tropsch type synthesis produce mixtures of alkanoic acids with hydrocarbon chains as long as 22 carbons, yet the main products were again heptanoic (C7:0), octanoic (C8:0), and nonanoic acids (C9:0) (Rushdi and Simoneit 2001). If fatty acids were primary constituents of protocellular membranes, the relative distribution of fatty acids would result in membranes composed of short-chain amphiphiles that have a very high CBC, implying that high concentrations of short fatty acids would be required to trigger vesicle formation. This raises the issue of availability in the early Earth environment, in which short chained fatty acids were likely to be present in low concentrations, so that an efficient concentrating mechanism was needed to reach their CBC. One can reasonably speculate that vesicles of short chain fatty acids could have existed in transient pools where water evaporation concentrated the amphiphiles, together with other solutes. It follows that this concentrating mechanism would not only have helped form the first membranes, but also increased the probability that a
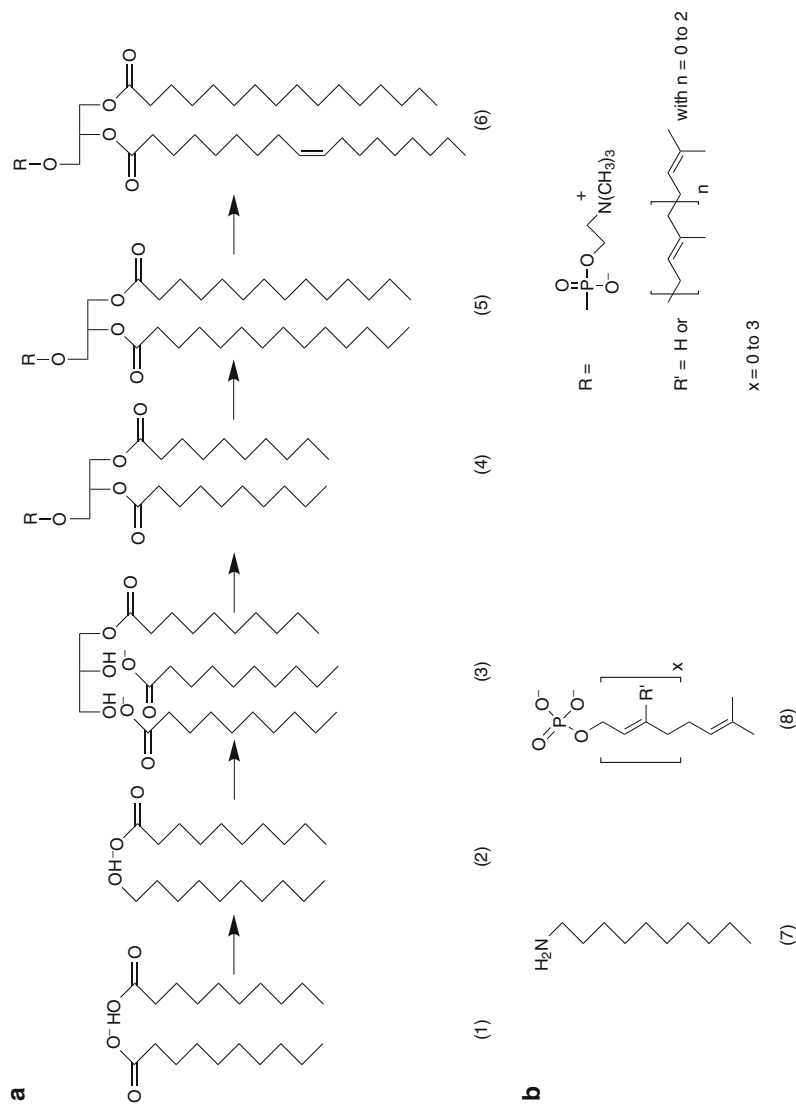
Fig. 8.2 Chemical structure of amphiphiles and amphiphile mixtures forming bilayer vesicles. (**a**) Fatty acid derivatives: (1) Pure fatty acid (decanoic acid, DA); (2) DA/decyl alcohol or (3) DA/glycerol decanoate mixtures; (4) didecyl-sn-glycerol-3-phosphocholine (DCPC); (5) dimysristoyl-sn-glycerol-3-phosphocholine (DMPC); (6) 1-palmitoyl-2-oleyl-sn-glycerol-3-phosphocholine (POPC). (**b**) Alternate amphiphiles: (7) decylamine (8) polyprenyl phosphates
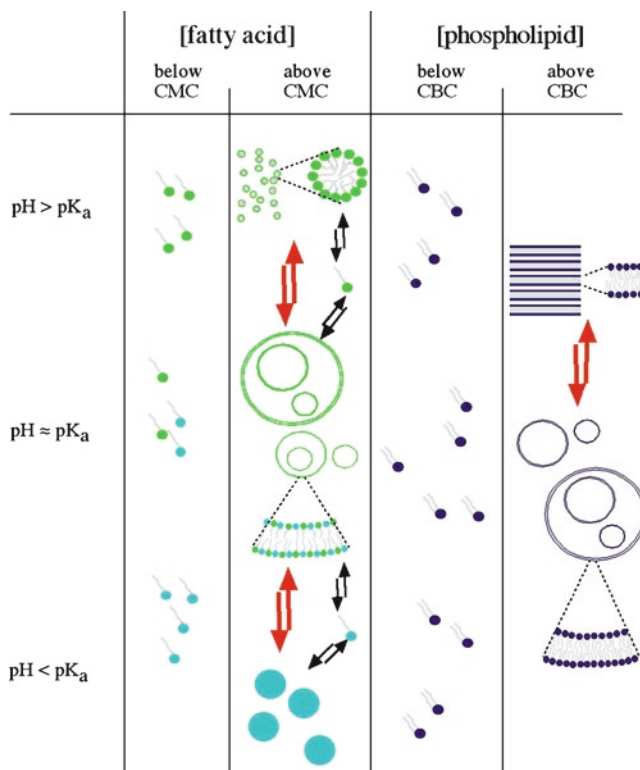
Fig. 8.3 Structures formed by amphiphilic molecules suspended in aqueous solutions. Fatty acids and phospholipids assemble into a variety of structures when suspended in aqueous media. The self-assembly process depends on the amphiphile concentration, the pH, and ionic content. Pure fatty acids are deprotonated at pH values higher than the apparent pK_a, and form micellar aggregates at concentrations higher than the critical micelle concentration (CMC). When the pH of a micellar suspension of fatty acids is slowly lowered to their apparent pK_a ($\pm$ 0.5–1 unit), vesicles become the predominant structure as the number of protonated molecules increases. Upon further acidification, fatty acid molecules become entirely protonated, and droplets of free acid form. At all pH ranges, fatty acid aggregates are always in dynamic equilibrium with single molecules. Pure phospholipids (PC) below their CBC (critical bilayer concentration) are present as single molecules. As their concentration increases above CBC, two structures are possible: At low water content, PC form stacked planar bilayer sheets which transform into vesicular structures at high water content

newly formed vesicle would encapsulate catalytic species already present in the environment along with an initial supply of substrates.

Permeability of primitive membranes is also a significant factor to consider. Previous studies of cellular membranes clearly demonstrated the role of the lipid bilayer as the primary permeability barrier to free diffusion of polar and ionic solutes. More recent observations have shown that the bilayer is not simply a barrier, but also exhibits a surprising degree of selectivity. For instance, Sacerdote and Szostak (2005) found that a lipid bilayer was relatively permeable to ribose when compared to

other pentoses and hexoses. Further investigations of vesicle bilayers have shown that permeability can be modulated by varying bilayer composition. In general, the permeability of bilayers to ionic solutes is inversely proportional to the length and the degree of unsaturation of the hydrocarbon chains. It would therefore be advantageous for primitive life forms to have membrane boundaries composed of short chain amphiphiles. Such bilayers would have facilitated exchange of small solutes between the protocell and its environment, providing a steady supply of nutrients, while still maintaining larger interacting polymers in a cellular compartment.

Apel et al. (2002) investigated the permeability of bilayers formed from short chain monocarboxylic acids (8–11 carbons), and established that these vesicles could efficiently retain polymeric material such as nucleic acids (see Fig. 8.4) or proteins. Other large ionic molecules, such as ADP, can also slowly diffuse across oleic acid bilayers (C18:1 Δ9), and serve as both as an energy source and substrate for RNA polymerization by polynucleotide phosphorylase (PNPase) (Walde et al. 1994a).

We must also take into account the effects on bilayer stability of environmental conditions such the salinity, ionic strength, pH, temperature, and divalent cations. For instance, if the emergence of life occurred in a marine environment, small sodium chloride concentration gradients would be expected to develop osmotic
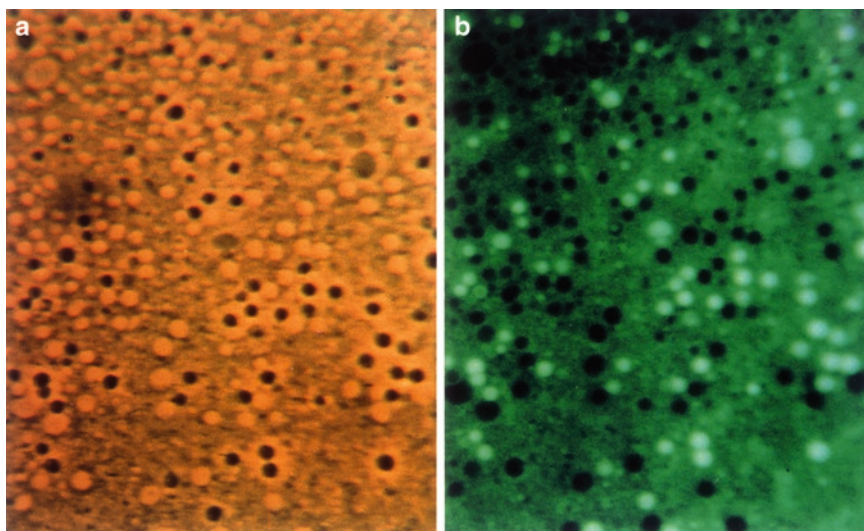
Fig. 8.4 Encapsulation of macromolecules in fatty acid bilayers visualized by epifluorescence microscopy. Phase (**a**) and epifluorescence (**b**) micrographs of vesicles produced from n-dodecanoic acid, n-dodecanol (5:1 molar ratio) at pH 8. The mixed vesicles formed by pH vesiculation were mixed in a 2:1 mass ratio with sonicated salmon testis DNA (approximately 600 bp in length) then dried. A dilute solution of acridine orange dye was added, and the newly formed vesicles were photographed by phase and fluorescence. Bar = 20 μm. Vesicles with entrapped DNA that are dark in (**a**) contain DNA which become fluorescent under UV-illumination in (**b**). Phase-light vesicles that do not contain DNA do not take up dye, and therefore do not fluoresce (**b**). The dark phase contrast is produced by the refractive index difference in those vesicles containing DNA

pressure across membranes (Wilson and Maloney 1976). We have investigated this effect and established that fatty acid vesicles are unstable in sodium chloride at concentrations near that of sea water (Monnard et al. 2002). The self-assembly of decanoic acid into vesicles is markedly inhibited at molar ratios of lipid to NaCl exceeding 3:1 (see Fig. 8.5). Furthermore, millimolar concentrations of divalent cations in marine salts cause fatty acid vesicles to precipitate as non-membranous aggregates (Monnard et al. 2002). The presence of divalent cations in early seawater would presumably inhibit the self-assembly of fatty acid membranes. These observations suggest an important constraint on aqueous sites related to the origin of cellular life. Unless future research demonstrates a plausible mixture of amphiphiles that can produce stable bilayer membranes in the presence of typical marine salts, cellular life would more likely have first been established in a fresh water environment. This implies that landmasses were required for the origin of life, so that fresh water could accumulate in the form of ponds, rivers and shallow inland seas.
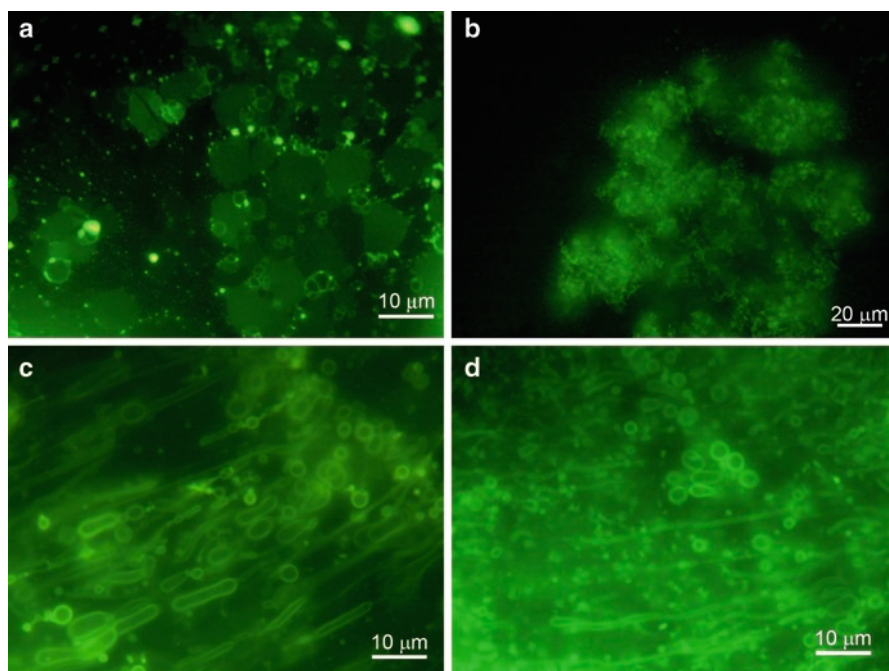
Fig. 8.5 Stability of fatty acid vesicles compared to fatty acid/glycerol alkanoate mixed vesicles in the presence of ionic solutes. (**a**) Decanoic acid (64 mM) in the presence of sodium chloride (636 mM) at a molar ratio DA to NaCl of 1:10. (**b**) Decanoic acid (64 mM) in the presence of calcium chloride at a molar ratio DA to divalent cation of 2:1. (**c**) A mixture of GMD/DA (7.5 mM/15 mM) in the presence of sodium chloride (447 mM) at a molar ratio DA to NaCl of 1:33. (**d**) A mixture of GMD/DA (7.5 mM/15 mM) in the presence of calcium a molar ratio DA to divalent cation of 2:1. DA vesicle suspensions (see Fig. 8.1b) when exposed to either NaCl or CaCl$_2$, precipitate whereas GMD/DA mixed vesicles remain intact (DA = decanoic acid; GMD = glycerol monodecanoate)

The prevailing temperature must also have affected primitive membranes as the Earth slowly cooled from above boiling to approximately 50°C in the Achaean Era when the first cell-like entities emerged (~3.5 Ga) (Knauth 2005). Elevated temperatures would have allowed reactions requiring high activation energy to occur more readily, but would also have enhanced degradation of thermolabile molecules. Furthermore, high temperatures can also destabilize membranes, leading to increased bilayer permeability followed by fatty acid precipitation.

Recent analyses of lipid composition in extremophiles have established that the membrane lipid composition of an organism surviving at high temperatures is quite different from that of a mesophile. Thermophiles today use branched chain hydrocarbons and membrane spanning lipids (De Rosa et al. 1991) to increase the stability of their membranes. The question of how first life overcame the temperature difficulties remains open.

Single unsaturated long hydrocarbon-chain amphiphiles, such as oleic acid that are known to form stable structures at elevated temperatures, are metabolic products in modern cells and are not synthesized in plausible prebiotic reactions. For this reason, saturated fatty acids are considered more plausible models of prebiotic amphiphiles. Saturated fatty acids with hydrocarbon chain lengths longer than 11C can form stable vesicles when the temperature is increased above 30°C (Skurtveit et al. 1989; Maurer et al. 2009). Thus, Archaean membranes composed of such amphiphiles could have persisted. The relative lack of pertinent studies clearly stems from fact that performing experiments at elevated temperatures is difficult. Vesicles can be stable at higher temperatures (like lauric acid) but tend to precipitate at lower temperatures (less than 30°C) so most investigations have been carried out with amphiphiles that are fluid at room temperature.

The effect of long-term temperature exposure was so far investigated for some saturated short-chain fatty acids as well as oleic acid (Mansy and Szostak 2008; Maurer et al. 2009). As might be expected, even oleic acid vesicles were disrupted by 24 h incubation at temperatures >65°C and decanoic acid vesicles were similarly degraded at 45°C. However, lauric acid and oleic acid vesicles remained stable at 45°C, which is a conservative temperature in the context of the emergence of life.

Vesicle-vesicle interactions triggered by temperature are interesting from a genetic perspective. While individuality of modern cells underlies evolutionary processes on a relatively long time frame, in primitive cellular systems this quality may not have been as essential. The interchange of internal contents as well as bilayer components could have led to increased complexity so that the emergence of new functions was promoted. Indeed, the bilayer could be made more stable by exchanging components between cells, thereby enhancing non-covalent interactions. Meanwhile, the internal contents could have been shared so that metabolic catalysts could interact with information carriers that were developed in separate containers, and then mixed without being exposed to the external media.

One of the simplest bilayer forming systems, decanoic acid, does not form vesicles that are stable as individual entities even at room temperature (20–25°C). The interchange of membrane materials between two vesicle populations was significant over 24 h time periods and was enhanced with increasing temperature, probably due to

fusion events. The length of the hydrocarbon chain plays a critical role, because it was observed that membrane mixing is reduced by 10–30% for lauric acid vesicles (12 carbons instead of ten) (Maurer et al. 2009).

8.4   Lipid Bilayer Membranes

Phosphatidylcholine (PC) is a highly evolved lipid species that is synthesized by a series of enzyme-catalyzed energy-dependent reactions. As such, PC provides a marked contrast with fatty acid as a membrane component. PC vesicles, also called liposomes, spontaneously form at CBC as low as 10^{-9} – 10^{-12} M. These structures are very stable to pH and divalent cations, and can withstand wide temperature variations (Oberholzer et al. 1995b). Although stable vesicles are formed by PC molecules having relatively long hydrocarbon chains, liposomes with hydrocarbon chains shorter than ten carbons are unstable, see (4) in Fig. 8.2. The synthesis of PC under prebiotic conditions seems implausible because of its molecular complexity, so these vesicles are unlikely candidates for self-assembly of early membranous boundary structures. Nonetheless, their relative stability also makes PC vesicles a useful model system in the laboratory.

The permeability of PC-bilayers to large ionic solutes is generally lower than that of fatty acid bilayers. The permeability properties of PC membranes have been extensively investigated for small ionic solutes (Kanehisa and Tsong 1978; Rosenquist et al. 1981; Deamer and Bramhall 1986; Chakrabarti et al. 1994; Paula et al. 1996) and small solutes (Langner and Hui 1993) in part because these lipids represent a significant portion of the amphiphiles forming biological membranes. The diffusion of charged solutes is clearly dependent on the gel-fluid phase transition (Mouritsen et al. 1995), and the length of the hydrocarbon chains. Chakrabarti et al. (1994) have further established that bilayer permeability to larger ionic molecules, such as ADP, has three distinct regimes. If the lipid chains were shorter than 12 carbons in length no selective permeability could be observed. Bilayers composed of 16-, 18-carbon chains were relatively impermeable, and could maintain concentration gradients of ADP for hours to days. On the other hand, dimysristoyl-sn-glycerol-3-phosphocholine bilayers (C14:0, DMPC) are sufficiently permeable so that diffusion of solutes the size of ADP or even ATP can provide substrates for entrapped enzymes. For this reason DMPC is often chosen for the experimental model systems to be described later in this review.

8.5   Mixed Amphiphile Systems

We have so far considered membranes composed of a single amphiphilic species, and these systems are very interesting as experimental models. However, it is unlikely that membranous boundaries composed of a pure amphiphile could exist

under prebiotic conditions, so it is reasonable to propose that primitive cellular life forms relied on mixed bilayers from abiotic sources of amphiphiles. In fact, all organisms today use lipid mixtures to construct cell membranes and modulate their properties. Some also require self-assembling amphiphiles from the environment. Examples include tocopherols and carotenoids (vitamins E and A) that are functional components of membranes, and essential fatty acids such as linoleic and arachidonic acid. Admixture of cholesterol is also used by eukaryotic cells to modulate properties of bilayers such as permeability and robustness.

Studies of mixed bilayers have demonstrated that similar modulating effects can be achieved by mixing prebiotically plausible amphiphiles. For instance, adding small amounts of a second amphiphile such as a fatty alcohol to pure fatty acid bilayers can substantially lower their CBC and mitigate disruptive effects of inorganic solutes on fatty acid bilayers (see Fig. 8.5). Such amphiphiles are considerably less complex than cholesterol, and can be synthesized in Fischer-Tropsch type reactions that are plausible for early Earth organic synthesis.

Adding alcohols with the same hydrocarbon chain length to fatty acid vesicles (Apel et al. 2002) not only reduced the CBC of the fatty acid, but also dramatically increased the stability of vesicles in alkaline conditions (see (2) in Fig. 8.2). When a micellar solution of nonanoic acid, at pH 11, was mixed with 1-nonanol, which by itself does not form vesicles, vesiculation was immediately induced upon vortexing. An aqueous solution of nonanoic acid required a concentration at least 85 mM to form vesicles, but only 20 mM acid when mixed with 2 mM nonanol. Longer chained alcohols exhibit the same property when mixed with the corresponding fatty acids. When these vesicles containing encapsulated nucleic acid, or proteins are exposed to digestive enzymes (Apel et al. 2002), the encapsulated species remained functional, a result that establishes the relative stability of the mixed structures.

The mechanism underlying increased vesicle stability upon addition of an alcohol presumably involves hydrogen bonding between fatty acids and the hydroxyl groups of alcohols, which are not affected by pH, thereby enhancing the stability of the structures. This interaction does not inhibit the electrostatic interaction between carboxylic headgroups and divalent cations, so that fatty acid/fatty alcohol vesicles are not significantly more stable in the presence of divalent cations than those composed of pure fatty acids. On the other hand, polyols, such as monoglycerides (Hargreaves et al. 1977) are able to mitigate this disruptive interaction, and allow formation of more robust membranes (see (3) in Fig. 8.2). Mixed vesicles composed of glycerol monodecanoate (GMD)/decanoic acid (DA) at a molar ratio of GMD/DA near to 1:2 were exposed to increasing concentration of NaCl, and divalent cations (Monnard et al. 2002). We found that these vesicles could cope with up to ten times more concentrated ionic inorganic solutions than the pure fatty acid counterparts.

Admixtures of glycerol monoacyls and fatty acids of same hydrocarbon chain length also significantly improved thermal stability of the resulting vesicles as determined by encapsulation efficiency and bilayer mixing experiments (Mansy and Szostak 2008; Maurer et al. 2009). In these systems, encapsulation of short oligomers of nucleic acids could be preserved for incubation periods of less than

10 h at temperatures up to 100°C or for weeks at temperature up to 45°C. At the molecular level, a significant decrease (up to 60%) of membrane component mixing between two vesicle populations was observed even at elevated temperatures.

Another mixed amphiphilic bilayer system was recently reported by Namani and Deamer (2008) (see (7) In Fig. 8.2), who found that mixtures of alkyl amines and fatty acids form vesicles at strongly basic and acidic pH ranges which are resistant to the effects of divalent cations up to 0.1 M. For instance, vesicles formed by 1:1 mole ratio mixtures of decylamine and decanoic acid were stable at acid pH of 3 and alkaline pH of 11, but crystallized at neutral pH. The vesicles were relatively permeable to pyranine, a fluorescent anionic dye, but permeability could be reduced by adding 2 mol% of a polycyclic aromatic hydrocarbon such as pyrene, which apparently had a stabilizing effect similar to that of polycyclic cholesterol in biological membranes. Permeability to the dye was also reduced by increasing the chain length of the amphiphiles. The authors concluded that primitive cell membranes were likely to be composed of mixtures of amphiphilic and hydrophobic molecules that manifested increased stability over pure fatty acid membranes.

We note that the addition of fatty acids or fatty alcohols to phospholipid vesicles markedly increases their permeability to large ionic solutes, in marked contrast to the stabilizing alcohol effects on fatty acid bilayers. For example, addition of myristic acid (C14:0; MA) to dimyristoyl-sn-glycero-3-phosphocholine (C14:0, C14:0; DMPC) increases the permeability of the bilayers towards large charged molecules, such as NTPs by one order of magnitude (Monnard and Deamer 2001). Similarly, the addition of amphiphilic detergents to phospholipid membranes at sublytic concentrations permits the passive diffusion of molecules even as large as proteins across phosphatidylcholine bilayers which are otherwise totally impermeable (Oberholzer et al. 1999).

8.6 Prebiotic Plausibility of Various Model Membranes

To summarize, it is likely that the components of early membranes have undergone a considerable evolution as the first forms of life slowly acquired new catalytic capacities. Figure 8.2a summarizes the chemical structures of six different amphiphile systems, all capable of forming vesicles, and the increase in molecular complexity is readily apparent. Understanding the emergence of life and its early evolution is closely related to the origin and evolution of cellular boundaries. At first, the requirements imposed on early membranes by the lack of protein transport assemblies, and of stabilizing structures, such as cytosomal support proteins or cell walls suggests relatively permeable membrane boundaries which are best modeled by fatty-acid vesicles (see Table 8.1). As early cells began to synthesize polymeric molecules such as RNA oligomers (Khvorova et al. 1999; Vlassov et al. 2001), small polypeptides (Ghadiri et al. 1994; Oliver and Deamer 1994; Clark et al. 1998; Kim et al. 1998) or other molecules that could mediate simple transport processes, more robust mixed amphiphile membranes would have replaced them.

Table 8.1 Summary of the properties of various compartment models

Compartment type	Amphiphiles	Properties	
		Limitations	Strengths
Coacervates		Aggregates of proteins, gelatine	Localized sites for enzymatic reactions
		Low stability	
		No selective permeability	
Vesicle	Fatty acid	High CBC	Prebiotic synthesis
		Sensivity to salts and temperature	Entrapment efficiency
			Selective permeability
			Membrane growth
	Phosphatidylcholine	Complex molecule	Stability (temperature, salts)
		Low permeability	Encapsulation efficiency
	Mixed amphiphiles	Sensivity to salts and temperature	Prebiotic synthesis
			Encapsulation efficiency
			Selective permeability
	Isoprenoids	Synthesis/availability	Encapsulation efficiency
			Selective permeability
Mineral		No enclosed volume	Availability
			Selectivity by absorption

Finally, as populations of microbial organisms increasingly were able to produce their own specific amphiphiles, membranes of more homogenous composition would have appeared.

We note that amphiphile candidates other than fatty acids have been proposed, such as polyprenyl derivatives (Nomura et al. 2001; Gotoh et al. 2006; Streiff et al. 2007) (see (8) in Fig. 8.2). These molecules also form vesicles that can capture and retain macromolecules, such DNA, but are limited by the lack of a prebiotically plausible synthetic pathway (Ourisson and Nakatani 1999).

8.7 Mineral Surfaces and Compartments

Bernal (1951); Wächtershäuser (1988); Cody et al. (2000) Martin and Russell (2003) and Baaske et al. (2007) have proposed a different approach to the self-assembly of organic compounds in which mineral surfaces act as templates or containers for assembling organic solutes into ordered structures. Indeed, a variety of polar substances are adsorbed with considerable specificity on clays, followed by polymerization

reactions (Ferris 1994, 1999; Ferris et al 1996). The adsorption on clays, although effective, does not have clear continuity with an evolutionary process leading to cellular life. Assuming that newly produced polymers can be released from binding sites on mineral surfaces, the lack of true cellular compartmentation would prevent their accumulation and significantly reduce further reactions that are essential for the evolution of increasingly complex catalytic reaction networks. Although it is highly probable that mineral surfaces played a role in the synthesis of complex organic molecules, they do not intrinsically possess the properties of compartmentation provided by amphiphile vesicles.

8.8   Compartmentalization of Catalytic Molecules

The membranous boundaries of early forms of cellular life, even though more permeable than those of contemporary cells, would have been sufficiently impermeable to polymeric materials so that complex networks of catalyzed reactions could develop. It follows that encapsulation of primitive catalytic molecular assemblies assumed to be present in the environment would have required a reversible process by which the bilayers were first disrupted, allowing the entry of these assemblies, and then resealed. The encapsulation of single molecules may have been fairly common on the early Earth in vesicles with an internal volume ranging from 5×10^{-13} to 5×10^{-18} liters for vesicular structures with diameters of 10–0.2 μm diameter. Encapsulation of specific multiple-component catalytic systems would have been increasingly difficult as the number of components increased. This suggests that the metabolism of early cells might have been necessarily simple until protein-mediated transport systems evolved.

Under prebiotic conditions, three plausible entrapment processes can be envisioned (1) simultaneous dehydration/rehydration of the vesicles and solutes (Deamer and Barchfeld 1982; Shew and Deamer 1985; Monnard et al. 1997), (2) production of amphiphile molecules, from non bilayer-forming precursors, in an environment containing solutes followed by amphiphile self-assembly (Walde et al. 1994b), and (3) aerosol-based vesicle formation (Dobson et al. 2000).

Two of the above processes have been extensively used in experimental studies of catalytic assemblies entrapped in amphiphile vesicles. The dehydration/rehydration process is based on the simultaneous drying of vesicles and macromolecules, which results in the formation of multilayered structures of lipid bilayers with intercalated solutes. Solute molecules are captured upon rehydration when the lipid bilayers reseal into vesicles, and it seems likely that multiple cycles of dehydration/rehydration could have occurred in intertidal zones on the early Earth. The second process was proposed by Luisi and co-workers (Bachmann et al. 1992; Walde et al. 1994b) in which membranous structures were produced at the expense of chemical energy stored in water-insoluble amphiphile precursors, or delivered to them from external sources. As the concentration of amphiphiles increased to values above the CBC, vesicles would have formed, entrapping macromolecules.

Once the catalytic species has been encapsulated in vesicles, access to nutrients and energy sources becomes crucial. As stated previously, early life forms presumably lacked specialized membrane transport systems, so that simple uptake mechanisms such as passive diffusion would play an essential role for the nutrient transport across boundary membranes. Could transmembrane diffusion be fast enough to keep up with the demands of a primitive metabolism? The metabolic rates of early cellular life are difficult to estimate, but it is likely that nutrient requirements may have been less than those of contemporary cells. For instance, if in fact an RNA world was a stepping stone toward contemporary life, it is clear that reactions catalyzed by ribozymes proceed much more slowly than those catalyzed by similar enzymes (Kagen et al. 1999; McKay and Wedekind 1999).

As soon as a recognizable metabolism is established within cellular boundaries, an energy supply becomes paramount. Several aspects of primitive metabolism have been studied, including energy uptake, compartmentalization of catalytic species pertaining to genetic-code information transfers (see Fig. 8.6), and general issues related to metabolism. (See Walde and Ichikawa (2001) or Monnard et al. (2008) for a review of catalytic species entrapped in vesicular structures.) Three sources of energy could be harvested by primitive cells: chemical energy in the form of chemical bonds or oxidation-reduction reactions, and light energy (see Deamer 1997 for review). Abundant sources of chemical energy are highly plausible components of the prebiotic environment (Morowitz 1992). However, as the membranous boundaries evolved and became less permeable to free diffusion of solutes, the uptake of chemical energy would have decreased due to slower permeation rates.

At some point light energy became the most abundant source of energy, as it is today, but how could it have been harvested by primitive cellular life? To capture light energy, photons must first be absorbed by a pigment system and then transduced into usable forms of chemical energy. It is reasonable to assume that photosynthetic assemblies comparable to those of contemporary life were absent, so that pigment systems such polycyclic aromatic hydrocarbons (PAHs) organic iron complexes, porphyrins, and proteinoids, could have been incorporated into the structure of bilayer membranes (Deamer 1992; Deamer 1997). The presence of PAH as an abundant organic component of carbonaceous meteorites materials has been established (Cronin et al. 1988) and it seems likely that PAH and their derivatives would have been among the most common organic compounds in the early Earth environment. A few preliminary studies of PAH as pigment systems have been reported. For instance, upon illumination of amphiphile membranes containing small amounts of pyrene carboxylaldehyde or other PAH derivatives, substantial pH gradients can be established, acidic inside (Deamer 1992).

Departing from the usual compartmentalization concept (the internal aqueous volume as the actual container), a team at Los Alamos National laboratory and the center for Fundamental Living Technology in Denmark (FLinT) investigated an alternative co-location (or interfacial) design (Rasmussen et al. 2003, 2004) where the bilayers themselves serve as compartments for the various components that constitute a photocatalytic, amphiphile producing metabolism.
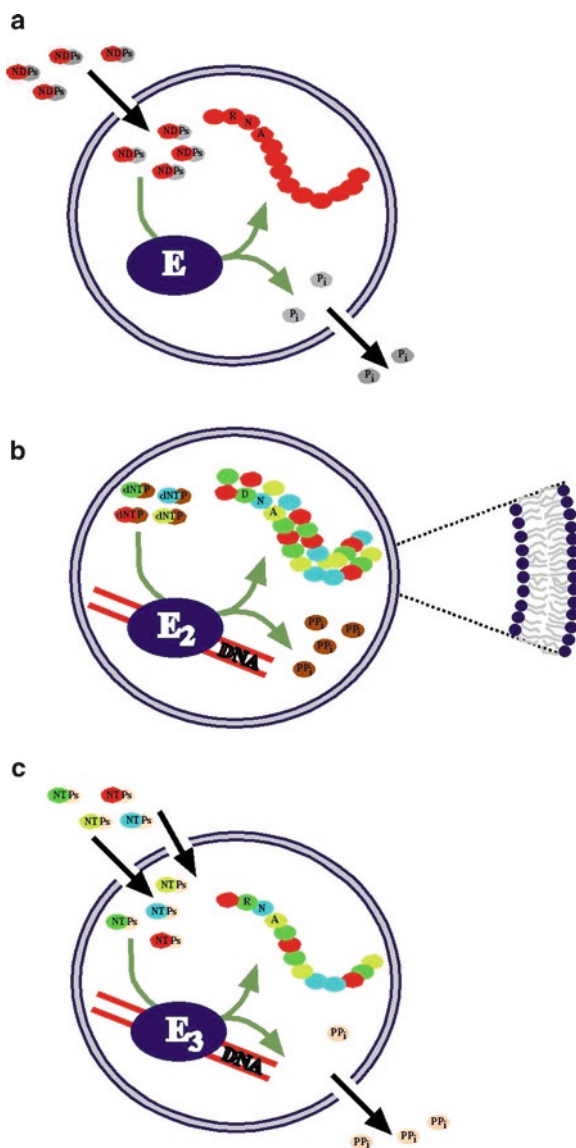
Fig. 8.6 Schematic representation of enzymatic reactions in vesicles. RNA polymerization mediated by polynucleotide phosphorylase. E represents the PNPase enzyme encapsulated within liposomes. ADP is added in the external medium, and must passively diffuse across the amphiphile bilayers to be processed by the enzyme. (**b**) PCR in POPC liposomes. The DNA polymerase enzyme (E_2) with its template (the primers are omitted), and its substrates were encapsulated simultaneously. (**c**) RNA transcription mediated by T7 RNA polymerase. The T7 RNA polymerase enzyme (E_3) and its template are encapsulated. NTPs are added externally as an energy source and substrates for the enzyme

This model (see Fig. 8.7a) is based on the interconnected work of a phosphosensitizer, i.e. a ruthenium complex, an information component, i.e., a nucleic acid system, both anchored by hydrocarbon chains into the compartment, i.e., a single chain amphiphile bilayer. The non-enzymatically replicable information molecule is not envisioned as a true encoding molecule, but rather as a sequence dependent actuator of the metabolic process that produces fatty acid from an oil-like precursor. The nucleic acids play the role of an electron donor/relay in the amphiphile-precursor photo-fragmentation catalyzed by a ruthenium complex. The effectiveness of the electron relay function should depend on the base sequence of the nucleic acid polymer, as the investigation of electron transfer on DNA strands has established (Giese 2002).

Experiments with a simplified system version (see Fig. 8.7b) have established the coupling of information molecule, photo-metabolism and container assembly (DeClue et al. 2009). When a simple chemical system composed of a fatty acid precursor material (but no fatty acid), a hydrogen donor, and a simplified version of the nucleic acid-ruthenium complex, (i.e., a single nucleobase, 8-oxoguanine, which is tethered to one bipyridine ligand of the metal center) in an aqueous solution is exposed to white light, the initial mixture devoid of any stable structure is over time transformed to a system full of membranous structures, vesicles, and tubes. The 8-oxo-guanine acts as an electron donor/relay in the system, stabilizing the photo-excited ruthenium complex long enough for the cleavage of the amphiphile precursor to take place efficiently. Its absence or its replacement by guanine (the natural nucleobase with the next most favorable, but insufficient redox potential) resulted in an almost complete lack of production of amphiphiles, illustrated by the absence of membranous structure formation. Furthermore, the amphiphile precursor and the nucleobase-ruthenium complexes both have an affinity for the resulting decanoic acid structures. These results demonstrate a required interconnection between a specific type of information molecule and the metabolism. When established, the result is the formation of a protocell that is capable of supporting its subsequent life cycles using both nutrients, as well as a primary energy source to produce membrane forming amphiphiles.

8.9 Transfer of Encapsulated "Genetic" Information

Even the earliest form of life must have possessed some sort of genetic apparatus that could direct the synthesis of polymers and pass genetic information from one generation of cells to the next. Although it is unlikely that nucleic acids and proteins together were components of the apparatus, analogous polymers must have been synthesized by an as yet unknown pathway, which were capable of the linked interactions leading to evolutionary selection. For this reason, we have been exploring simple "genetic" information transfer within vesicles.

A minimal transcription system should be composed of a molecule with the dual functions of catalysis and information storage (Joyce 1998; Rogers and Joyce 1999). RNA seems to be a plausible candidate as first proposed by Gilbert in the "RNA world" conjecture (Woese 1968). Even though the search for an RNA fragment
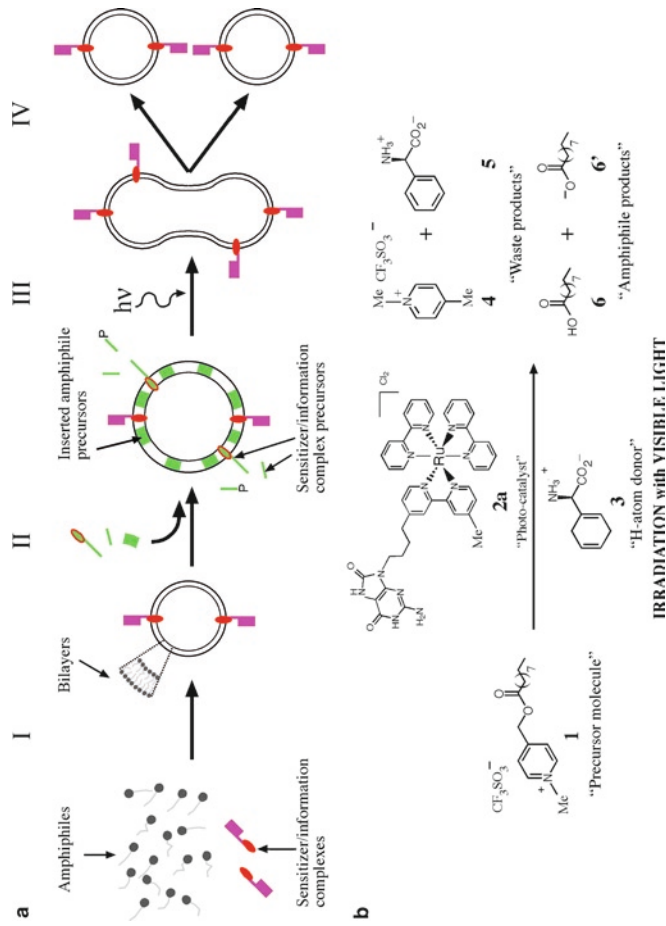
Fig. 8.7 Co-location or interfacial design. (**a**) This design is based on the idea that all chemical compounds that form the protocell have a hydrophobic moiety and thus can participate to the self-assembly of the compartment (phase I). In the resulting system, the true compartment is the bilayer structure and its interface with the solvent. The chemicals that are required to synthesize more protocell building blocks having the same hydrophobic character can once added to the external medium spontaneously insert (phase II) into the protocell compartment where they will be converted (phase III) using light energy into either amphiphiles or information molecule copies. This production of new building blocks leads to the growth of the systems and once a threshhold size is reached, should induce the division of the grown protocell into new systems having similar or enhanced properties (phase IV). (**b**) Nucleobase-mediated photochemical production of decanoic acid from its picolyl ester precursor. The precursor **1** is dispersed in an aqueous buffered solution containing the electron donor linked photocatalyst (8-oxo-G-Ru photocatalyst, a derivative of [tris(2,2'-bipyridine)-Ru(II)]²⁺) **2a**, and the dihydrophenylglycine hydrogen donor **3**. Upon irradiation, fatty acids (**6** protonated and **6'** deprotonated) form and self-assemble into membranous structures once the critical vesicle concentration of the precursor/fatty acid mixture is reached. Two "waste" compounds: N-methyl picolinium **4** and phenylglycine **5** are also produced

with polymerase activity has progressed rapidly (Johnston et al. 2001; Zaher and Unrau 2007), an efficient RNA-dependent RNA polymerase remains elusive, and no working experimental system is yet available. Therefore, most of the research on genetic information transfer within amphiphile vesicles has been conducted using RNA/DNA polymerase enzymes, PNPase (Chakrabarti et al. 1994; Walde et al. 1994a), Taq polymerase (Oberholzer et al. 1995b), Qβ replicase (Oberholzer et al. 1995a, Kita et al. 2008), and T7 RNA polymerase (Monnard et al. 2007). These model systems can help in determining which mechanisms might have been involved in the early information transfer.

Polynucleotide phosphorylase (PNPase) in bacterial cells normally functions to hydrolyze RNA to monomeric species, but if NDPs are in excess, it will produce random RNA polymers thousands of nucleotides in length. To demonstrate that polynucleotides can be synthesized by encapsulated PNPase using transmembrane transport of substrates, two groups independently encapsulated PNPase within different vesicular systems, dimyristoyl-sn-glycero-3-phosphocholine (DMPC, C14:0) (Chakrabarti et al. 1994), a phospholipid, and oleic acid (C18∆9) (Walde et al. 1994a), an unsaturated fatty acid. In both experiments, ADP was the substrate and RNA in the form of poly(A) was produced after 1–5 days incubation, and remained within the amphiphile-bound compartment (see Fig. 8.6a). The reaction rate with the DMPC-encapsulated enzyme was determined to be approximately 20% of that with PNPase in an aqueous buffer, showing that the bilayer was a substantial barrier to substrate permeation. These systems demonstrated that ADP could permeate across lipid bilayers at rates sufficient to support polymer synthesis by PNPase.

The next experimental step involved encapsulation of a more complex enzyme system capable of catalyzing replication or transcription. Two such reactions include the amplification of RNA template by Qβ replicase entrapped in oleic acid vesicles (Oberholzer et al. 1995a) and the amplification DNA by PCR within phospholipid vesicles (Oberholzer et al. 1995b) (See Fig. 8.6b). Both reactions are template dependent and require metal ions as cofactors, as well as primers for the PCR experiment. All components of the reaction (three for the Qβ replicase, and five for the PCR) must be captured simultaneously in a single vesicle with their respective substrates because of the low permeability of the bilayer membranes to NTPs and dNTPs. In both cases, the expected products were formed in low yields. Oberholzer et al. calculated that the average aqueous vesicular volume was 3.3×10^{-18} l in the PCR experiments, which represents approximately 8000 dNTP molecules per reaction volume, enough to produce ten or so double stranded DNA products.

Polymerization is limited by the low permeability of PC bilayers to substrate molecules in both studies described above. However, the PNPase experiments using DMPC vesicles established that passive diffusion of substrate under certain conditions may be fast enough to allow template-directed enzymatic amplification of encapsulated genetic material. We have attempted to develop this approach further by designing a system capable of relatively high rates of substrate permeation. Because of their stability and permeability properties, DMPC liposomes were chosen to encapsulate a template-directed T7 RNA polymerase (Monnard et al. 2007). Even though DMPC liposomes are more permeable to NTPs than the POPC vesicles used for the PCR reactions, their permeability coefficient is in the range of 10^{-11} cm s^{-1} at

37°C (Monnard and Deamer 2001), a rate that permits only a few NTPs per second to enter a given vesicle. To increase permeability, the reactions were cycled between 23.3°C and 37°C. Under these conditions the vesicles undergo a fluid – gel state phase transition at 23.3°C, a temperature at which they are orders of magnitude more permeable than in the fluid phase. As a result substrate uptake by a vesicle increases to approximately 1,500 molecules during a 5-min incubation at 23.3°C, followed by a 1-min incubation at 37°C to induce optimal T7 RNA polymerase activity. RNA product was detected within the liposomes, and could be reverse-transcribed. More recently, Mansy et al. (2008) reported that the bilayer boundary of lipid vesicles composed of fatty acids and glycerol monoesters are sufficiently permeable so that activated nucleotides can enter the vesicle and take part in a non-enzymatic, template-directed polymerization of DNA derivatives. These results are enlightening in several ways. First, they demonstrate that the bilayer stability allows the encapsulation of a complex enzymatic assembly in an aqueous volume as small as 5×10^{-18} l. Second, passive diffusion of substrate as large as nucleotides provided enough substrate for gene amplification to take place. The fact that the products remain inside the vesicles, where they are protected from degradation by proteases and nucleases, underlines the importance of the compartment in the early evolution of cellular life. That is, membrane-bounded compartments can accumulate polymeric products which in turn can further interact and produce more complex systems.

8.10   Self-Reproducing Compartments

As discussed earlier, even the earliest cellular forms of life must have had membranous boundaries, a specific set of catalyzed metabolic and polymerization reactions, and the capacity for self-reproduction. Self-reproduction entails not only the replication of a genetic material and catalysts, but also the production of additional membrane surface to accommodate growth and to permit the budding of daughter cells. Figure 8.8 illustrates a hypotheticical minimal cell of an RNA world. The genetic material must code for three essential components of the cell. (1) A ribozymal polymerase activity is required for transcription of the gene into other active ribozymes, including replication of the genetic RNA itself. (2) A ribozymal acyltransferase activity synthesizes new membrane molecules from their precursors. Acyl CoA and lysophophatidylcholine are shown as examples of such precursors, and in fact have been demonstrated to generate new membranes from soluble precursors (Gavino and Deamer 1982). (3) An RNA fragment, "bud RNA", is required to trigger the budding of the compartment boundaries once these exceed a threshold surface area.

Luisi and co-workers at the ETH in Zürich have approached the question of self-reproducing compartments by using vesicles formed from fatty acids (Walde et al. 1994a; Walde et al. 1994b; Walde et al. 1994c; Oberholzer et al. 1995a; Morigaki et al. 1997) which are plausible amphiphile candidates for self-assembling boundary structures of early cellular life. Starting with preformed oleic acid vesicles, a source of amphiphile precursors was added in the form of a water-immiscible anhydride of oleic acid. The anhydride does not form vesicles but instead is present as
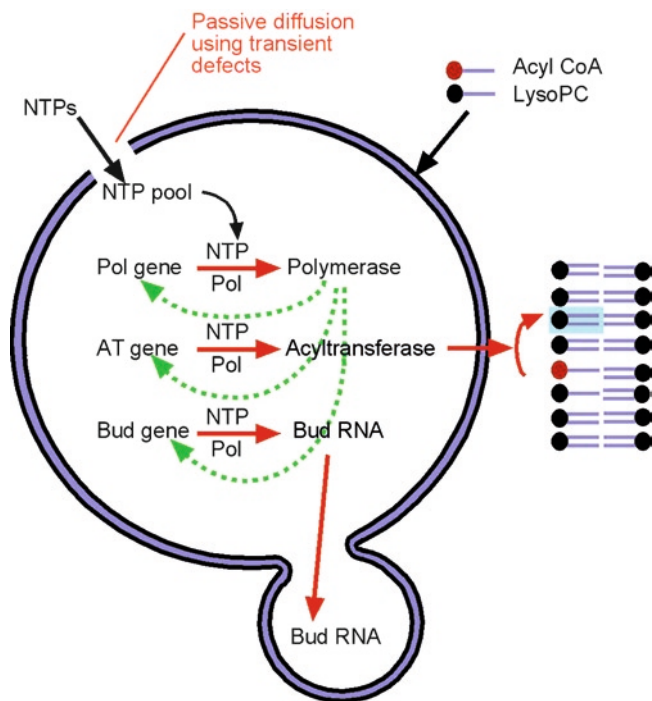
Fig. 8.8 A minimal cell based on catalytic/structural RNA. A laboratory version of a minimal cell depicted here would possess a rudimentary genetic library coding for three catalytic/structural RNA required for its growth and self-reproduction: Pol gene (polymerase for replication and transcription), AT gene (acyltransferase for lipid synthesis) and bud RNA. Substrates in form of NTPs passively diffuse across the bilayer membranes using transient defects, while precursors of bilayer compounds (Acyl CoA and lysophosphatidylcholine, LysoPC) partition into the existing bilayers (*black arrows*) were the AT RNA synthesizes phospholipids. The red arrows represent the actual reaction network of the minimal cells: RNA polymerization, lipid formation, and structural promotion of membrane budding. The green dashed arrows represent the secondary activity of the polymerase ribozyme that produces copies of the genes. Note that the polymerase, acyltransferase, as well as three genes would also be present in the aqueous volume of the budding membrane

fluid droplets. As the anhydride hydrolyzes, it provides additional membrane building blocks so that the number and the size of vesicles in the aqueous medium increase. Significantly, when the hydrolysis is carried out without vesicles in the aqueous phase, two distinct kinetic regimes are observed, an initial slow phase before vesicles are present, and an autocatalytic fast phase as the number of vesicles in the suspension increases.

In two additional series of experiments using oleic acid vesicles with entrapped PNPase and Qβ replicase, it was possible to demonstrate vesicle self-reproduction and simultaneous polymerization of RNA from substrates added to the external medium. These represent the first model systems to incorporate a catalyzed membrane growth in concert with a catalyzed synthesis of a nucleic acid (Walde et al. 1994a; Oberholzer et al. 1995a).

Szostak and coworkers (Hanczyc et al. 2003; Hanczyc and Szostak 2004; Zhu and Szostak 2009) were able to demonstrate not only growth of fatty acid vesicles containing encapsulated RNA, but also a kind of division imposed by shear forces when the grown vesicles were passed through a filter containing microscopic pores or submitted to an external force. Chen and Szostak (2004) further established a detailed kinetic analysis of the growth, and also demonstrated that vesicles could compete for membrane components (Chen et al. 2004) present in their environment and undergo a kind of evolutionary selection.

It is interesting to consider these experiments in terms of the prebiotic environment. The anhydride droplets are models for an external source of membrane precursors, and the hydrolysis of the anhydride linkage proceeds at the expense of chemical energy stored in the anhydride bond. Furthermore, the fatty acid bilayers themselves catalyze the reaction by which their components are produced. A system similar to the self-reproducing oleic acid vesicles might have preceded a lipid synthesis coded in the genetic material of early cells, even though such systems lack feedback regulation and could result in the production of empty vesicles. Indeed if the rate of amphiphile formation substantially exceed the reproduction rates of the metabolic network elements (genetic codes and catalytic species), these species will be rapidly diluted in the expanding internal aqueous compartment, and so that newly formed vesicles lose the characteristics of the parent cell.

Attempts have been recently made to closely link the replication of containers to an endogenous production of amphiphiles either using simple metal catalytic systems (DeClue et al. 2009) or complex transcription and translation machinery encapsulated within lipid vesicles, so called minimal cell concept (Luisi 2006).

In the former system, the catalytic rates and the fact that catalysis occurred in the presence of vesicles (later stage of the reported reaction) indicate that in principle an internalized production of amphiphiles at rates compatible with the replication of the compartment is achievable. However, the system still lacks the possibility to replicate all its components, in particular the information part.

This minimal cell method uses existing genes to produce proteins, which in turn will catalyze the synthesis of bilayer-forming molecules. Although a highly evolved set of genes employed in this type of experiments is clearly not prebiotic, the use of an existing genetic/metabolic network simplifies the process of creating a synthetic system. Using phospholipid membranes, the Luisi group has encapsulated a functional cell-free genetic/metabolic system, which can produce proteins involved in the synthesis of phosphatic acid (Kuruma et al. 2009). Although these synthesized proteins were not yet capable of producing membrane-forming material at a rate sufficient to measure the container growth, this experiment established the feasibility of internal amphiphile production and thus represents a first significant step in linking directly amphiphile production to an internal catalytic system.

In all systems studied to date, the main element still missing remains a regulatory feedback mechanism that would, if incorporated into early cellular system models, ensure a concerted replication of metabolic and information components and compartment.

8.11 Conclusions and Future Directions

The study of carbonaceous meteoritic material and laboratory models of plausible Fischer-Tropsch type reactions show that vesicle-forming amphiphiles were likely to be present on the early Earth, and therefore could have participated in the formation of boundary membranes required by early cellular life. Moreover, short-chain amphiphile-based vesicles have properties similar to those of liposomes formed from phospholipids that are primary components of contemporary cellular membranes. They tend to be less stable and more permeable to ionic solutes, but as discussed here, higher permeability can be an advantage in the absence of specialized transport proteins.

Recent investigations have established that encapsulation procedures such as dehydration/rehydration or pH vesiculation permit simultaneous trapping of all components of complex catalytic systems, which remain active in the vesicular compartment. Furthermore, the membrane boundaries of such vesicles can protect the encapsulated catalysts from agents such as proteases and nucleases. On the early Earth, it seems likely that a membrane-protected microenvironment would be conducive to survival of spontaneous molecular systems on the evolutionary path to the first forms of life.

Another interesting aspect of amphiphile vesicles is their selective permeability, which permits the passive diffusion of smaller substrate molecules and prevents the release of larger metabolic products, thereby leading to accumulation of polymeric products. This observation demonstrates that an early form of life could have relied on a similar mechanism to take up nutrients and energy from the environment. The fact that enzymes having high turnover rates can be supplied with substrate by this simple mechanism supports the idea that a rather complex, perhaps ribozyme-based metabolism could be sustained in a closed compartment without need for transmembrane protein carrier systems. We also note that accumulation of polymeric products would permit and promote additional reactions that lead to increased metabolic complexity in cellular compartments.

Experiments have also demonstrated that simple vesicles composed of fatty acids could have undergone membrane growth using external amphiphile precursors, and thus undergo a kind of self-reproduction, in the sense that new vesicles are produced by growth and dispersion of preformed vesicles. This concept has been discussed in detail as a lipid world scenario (Segre et al. 2001).

However, the vesicle-based model systems described here also demonstrate the limitations of membrane-encapsulated reactions that must be overcome before we can design a model system endowed with the properties of a minimal cell (see Fig. 8.8). Even though single-component vesicles are useful laboratory models, they may not accurately reflect the compartment boundaries of early protocells in terms of mixed lipid components, permeability and substrate transport. When we attempt to model a protocell using highly evolved polymerase enzymes with high turnover rates, the low permeability of model bilayer membranes imposes significant limitations on successful outcomes. Recent studies tend to emphasize one

aspect of a protocell, such as RNA polymerization, and lack the interacting metabolic pathways characteristic of living systems. Amphiphile vesicular systems with entrapped catalytic species possess no feedback regulation, a hallmark of life. Even though early cells may have incorporated simplified metabolic pathways, it would still be necessary to control their metabolism in some still unknown way to survive and undergo further evolutionary development.

In terms of future research directions, it is necessary to develop new membrane compositions to enhance interactions between the compartment and its environment. A steady supply of nutrient/energy relying on transport mechanisms such as passive diffusion will be a milestone in designing a plausible model for the minimal cell. Furthermore, as our knowledge of ribozyme chemistry increases (Voytek and Joyce 2007), it may be possible to bypass the requirement of protein enzymes as catalysts. This will close an enormous gap in our understanding of the origin of life, in that ribosomes are necessary for protein synthesis. It seems unlikely that we will reach a point at which it would be possible to include a full complement of DNA, RNA, ribosomes and protein synthesis in a defined minimal cell system and replicate all of them. Finally, by investigating both the synthesis of encapsulated polymeric molecules and the growth of membrane boundaries, a mechanism may be discovered that provides essential insights into the origin of feedback regulation.

References

Apel CL, Mautner MN, Deamer DW (2002) Self-assembled vesicles of monocarboxylic acids and alcohols: conditions for stability and for encapsulation of biopolymers. Biochim Biophys Acta 1559:1–9

Baaske P, Weinert FM, Duhr S, Lemke KH, Russell MJ, Braun D (2007) Extreme accumulation of nucleotides in simulated hydrothermal pore systems. Proc Natl Acad Sci USA 104: 9346–9351

Bachmann PA, Luisi PL, Lang J (1992) Autocatalytic self-replicating micelles as models for prebiotic structures. Nature 357:57–59

Bangham AD, Standish MM, Miller N (1965) Diffusion of univalent ions across the lamellae of swollen phospholipids. J Mol Biol 13:238–252

Bernal JD (1951) The physical basis of life. Routledge, London

Chakrabarti AC, Breaker RR, Joyce GF, Deamer DW (1994) Production of RNA by polymerase protein encapsulated within phospholipid vesicles. J Mol Evol 39:555–559

Chen IA, Roberts RW, Szostak JW (2004) The emergence of competition between model protocells. Science 305:1474–1476

Chen IA, Szostak JW (2004) A kinetic study of the growth of fatty acid vesicles. Biophys J 87:988–998

Clark TD, Buehler LK, Ghadiri MR (1998) Self-assembling cyclic b^3-peptide nanotubes as artificial transmembrane ion channels. J Am Chem Soc 120:651–656

Cody GD, Boctor NZ, Filley TR, Hazen RM, Scott JH, Sharma A, Yoder HS Jr (2000) Primordial carbonylated iron-sulfur compounds and synthesis of pyruvate. Science 289:1337–1340

Cronin JR, Pizzarello S, Cruickshank DP (1988) Organic matter in carbonaceous chondrites, planetary satellites, asteroids and comets. In: Matthews MS (ed) Meteorites and the early solar system. University of Arizona Press, Tucson, pp 819–857

Deamer DW, Oro J (1980) Role of lipids in prebiotic structures. Biosystems 12:167–175

Deamer DW, Barchfeld GL (1982) Encapsulation of macromolecules by lipid vesicles under simulated prebiotic conditions. J Mol Evol 18:203–206

Deamer DW (1985) Boundary structures are formed by organic components of the Murchison carbonaceous chondrites. Nature 317:792–794

Deamer DW, Bramhall J (1986) Permeability of lipid bilayers to water and ionic solutes. Chem Phys Lipids 40:167–188

Deamer DW, Pashley RM (1989) Amphiphilic components of the Murchison carbonaceous chondrite: surface properties and membrane formation. Orig Life Evol Biosph 19:21–38

Deamer DW (1992) Polycyclic aromatic hydrocarbons: primitive pigment systems in the prebiotic environment. Adv Space Res 12(4):183–189

Deamer DW, Mahon EH, Bosco G (1994) Self-assembly and function of primitive membrane structures. In: Bengtson S (ed) Early life on earth. Columbia University Press, New York/Chichester/West Sussex, pp 107–123

Deamer DW (1997) The first living systems: a bioenergetic perspective. Microbiol Mol Biol Rev 61:230–261

Deamer DW, Dworkin JP, Sandford SA, Bernstein MP, Allamandola LJ (2002) The first cell membranes. Astrobiology 2:371–382

DeClue MS, Monnard P-A, Bailey JA, Maurer SE, Collis GE, Ziock H-J, Rasmussen S, Boncella JM (2009) Nucleobase mediated, photocatalytic vesicle formation from an ester precursor. J Am Chem Soc 131:931–933

De Rosa M, Trincone A, Nicolaus B, Gambacorta A (1991) Archeabacteria: lipids, membrane structures, and adaptations to environmental stresses. In: Di Prisco G, Federation of European Biochemical Societies. Meeting (eds) Life under extreme conditions: biochemical adaptation, pp vii, 144. Springer, Berlin/New York

Dobson CM, Ellison GB, Tuck AF, Vaida V (2000) Atmospheric aerosols as prebiotic chemical reactors. Proc Nat Acad Sci USA V97:11864–11868

Ferris JP (1994) The prebiotic synthesis and replication of RNA oligomers: the transition from prebiotic molecules to the RNA world. In: Fleischaker GR (ed) Self-production of supramolecular structures, pp 89–98. Kluwer, Dordrecht

Ferris JP (1999) Prebiotic synthesis on minerals: bridging the prebiotic and RNA worlds. Biological Bull 196:311–314

Ferris JP, Hill AR Jr, Liu R, Orgel LE (1996) Synthesis of long prebiotic oligomers on mineral surfaces. Nature 381:59–61

Gavino V, Deamer DW (1982) Purification of acyl CoA: 1-acyl-sn-glycerophosphorylcholine acyltrasnferase. J Bioenerg Biomembr 14:513–526

Gebicki JM, Hicks M (1973) Ufasomes are stable particles surrounded by unsaturated fatty acid membranes. Nature 243:232–234

Gebicki JM, Hicks M (1976) Preparation and properties of vesicles enclosed by fatty acid membranes. Chem Phys Lipid 16:142–160

Ghadiri MR, Granja JR, Buehler LK (1994) Artificial transmembrane ion channels from self-assembling peptide nanotubes. Nature 369:301–304

Giese B (2002) Electron transfer in DNA. Curr Opin Chem Biol 6:612–618

Goldacre RJ (1958) Surface films: their collapse on compression, the shapes and sizes of cells, and the origin of life. In: Danielli JF, Pankhurst KGA, Riddiford AC (eds) Surface phenomena in biology and chemistry. Pergamon Press, New York, pp 12–27

Gotoh M, Miki A, Nagano H, Ribeiro N, Elhabiri M, Gumienna-Konteck E, Albrecht-Gary A-M, Schmutz M, Ourisson G, Nakatani Y (2006) Membrane properties of branched polyprenyl phosphates, postulated as primitive membrane constituents. Chem Biodiv 3:434–455

Haines TH (1983) Anionic lipid headgroups as a proton-conducting pathway along the surface of membranes: a hypothesis. Proc Natl Acad Sci USA 80:160–164

Haldane JBS (1929) The origin of life. Rat Ann 148:3–10

Hanczyc MM, Fujikawa SM, Szostak JW (2003) Experimental models of primitive cellular compartments: encapsulation, growth, and division. Science 302:618–622

Hanczyc MM, Szostak JW (2004) Replicating vesicles as models of primitive cell growth and division. Curr Opin Chem Biol 8:660–664

Hargreaves WR, Mulvihill SJ, Deamer DW (1977) Synthesis of phophosplipids and membranes in prebiotic conditions. Nature 266:78–80

Hargreaves WR, Deamer DW (1978) Liposomes from ionic, single-chain amphiphiles. Biochemistry 17:3759–3768

Johnston WK, Unrau PJ, Lawrence MS, Glasner ME, Bartel DP (2001) RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. Science 292: 1319–1325

Joyce GF (1998) Nucleic acid enzymes: playing with a fuller deck. Proc Natl Acad Sci USA 95:5845–5847

Kagen P, Cech TR, Golden BL (1999) Building a catalytic active site using only RNA. In: Atkins JF (ed) The RNA world, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 321–349

Kanehisa MI, Tsong TY (1978) Cluster model of lipid phase transition with application to passive permeation of molecules and structures relaxations in lipid bilayers. J Am Chem Soc 100:424–432

Khvorova A, Kwak Y-G, Tamkun M, Majerfeld I, Yarus M (1999) RNAs that bind and change the permeability of phospholipid membranes. Proc Nat Acad Sci USA 96:10649–10654

Kim HS, Hartgerink JD, Ghadiri MR (1998) Oriented self-assembly of cyclic peptide nanotubes in lipid membranes. J Am Chem Soc V120:4417–4424

Kita H, Matsuura T, Sunami T, Hosoda K, Ichihashi N, Tsukada K, Urabe I, Yomo T (2008) Replication of genetic information with self-encoded replicase in liposomes. Chem Bio Chem 9:2403–2410

Knauth LP (2005) Temperature and salinity history of Precambrian ocean: implications for the course of microbial evolution. Palaeogeogr Palaeoclimatol Palaeoecol 219:53–69

Komiya M, Shimoyama A, Harada K (1993) Examination of organic compounds from insoluble organic matter isolated from some Antarctic carbonaceous chondrites by heating experiments. Geochim Cosmochim Acta 57:907–914

Kuruma Y, Stano P, Ueda T, Luisi PL (2009) A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. Biochim Biophys Acta 1788:567–574

Kvenvolden KA, Lawless J, Pering K, Peterson E, Flores J, Ponnamperuma C, Kaplan IR, Moore C (1970) Evidence for extraterrestrial amino acids and hydrocarbons in the Murchison meteorite. Nature 228:923–926

Langner M, Hui SW (1993) Dithionite penetration through phospholipid bilayers as a measure of defects in lipid molecular packing. Chem Phys Lipids 65:23–30

Lawless JG, Yuen GU (1979) Quantitation of monocarboxylic acids in the Murchison carbonaceous meteorite. Nature 282:431–454

Lazcano A (1994a) The transition from nonliving too living. In: Bengtson S (ed) Early life on earth. Columbia University Press, New York/Chichester/West Sussex, pp 60–69

Lazcano A (1994b) The RNA world, its predecessors, and its descendants. In: Bengtson S (ed) Early life on earth. Columbia University Press, New York/Chichester/West Sussex, pp 70–80

Luisi PL, Varela FJ (1989) Self-replicating micelles - a chemical version of a minimal autopoietic system. Orig Life Evol Biosph 19:633–643

Luisi PL, Walde P, Oberholzer T (1999) Lipid vesicles as possible intermediates in the origin of life. Curr Opin Colloid Interface Sci 4:33–38

Luisi PL (2006) The emergence of life: from chemical origins to synthetic biology. Cambridge University Press, Cambridge

Mansy SS, Schrum JP, Krishnamurthy M, Tobé S, Treco DA, Szostak JW (2008) Template-directed synthesis of a genetic polymer in a model protocell. Nature 454:122–125

Mansy SS, Szostak JW (2008) Thermostability of model protocell membranes. Proc Natl Acad Sci 105:13351–13355

Martin W, Russell MJ (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. Philos Trans R Soc Lond B Biol Sci 358:59–83

Maurer SE, Deamer DW, Boncella JM, Monnard PA (2009) Chemical evolution of amphiphiles: glycerol monoacyl derivatives stabilize plausible prebiotic membranes. Astrobiology (in press)

Mautner M, Leonard RL, Deamer DW (1995) Meteorite organics in planetary environments: hydrothermal release, surface activity and microbial utilization. Planetary Space Sci 43:139–147

McCollom TM, Ritter G, Simoneit BRT (1999) Lipid synthesis under hydrothermal conditions by Fischer-Tropsch-type reactions. Orig Life Evol Biosph 29:153–166

McKay DB, Wedekind JE (1999) Small ribozymes. In: Atkins JF (ed) The RNA world, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 265–286

Miller SL (1953) Production of amino acids under possible primitive Earth conditions. Science 117:528–529

Monnard P-A, Oberholzer T, Luisi PL (1997) Encapsulation of polynucleotides in liposomes. Biochim Biophys Acta 1329:39–50

Monnard P-A, Deamer DW (2001) Nutrient uptake by protocells: a liposome model system. Orig Life Evol Biosph 31:147–155

Monnard P-A, Apel CL, Kanavarioti A, Deamer DW (2002) Influence of ionic inorganic solutes on self-assembly and polymerization processes related to early forms of life: implications for a prebiotic aqueous medium. Astrobiology 2:139–152

Monnard P-A, Luptak A, Deamer DW (2007) Models of primitive cellular life: polymerases and templates in liposomes. Philos Trans R Soc B Biol Sci 362:1741–1750

Monnard P-A, Ziock H-J, Declue MS (2008) Organic nano-compartments as biomimetic reactors, and protocells. Curr Nanosci 4:71–87

Morigaki K, Dallavalle S, Walde P, Colonna S, Luisi PL (1997) Autopoietic self-reproduction of chiral fatty acid vesicles. J Am Chem Soc 119:292–301

Morowitz HJ (1992) Beginnings of cellular life. Metabolism recapitulates biogenesis. Yale University Press, New Haven, CT

Mouritsen OG, Jorgenseb K, Honger T (1995) Permeability of lipid bilayers near the phase transition. In: Simon SA (ed) Permeability and stability of lipid bilayers. CRC Press, Boca Raton, pp 137–160

Namani T, Deamer DW (2008) Stability of model membranes in extreme environments. Orig Life Evol Biosph 38:329–341

Nomura S-iM, Yoshikawa K, Dannenmuller O, Chasserot-Golaz S, Ourisson G, Nakatani Y. 2001. Towards proto-cells: "Primitive" lipid vesicles encapsulating giant DNA and its histone complex. Chembiochem:457–459.

Oberholzer T, Wick R, Luisi PL, Biebricher CK (1995a) Enzymatic RNA replication in self-reproducing vesicles: an approach to a minimal cell. Biochim Biophys Res Commun 207:250–257

Oberholzer T, Albrizio M, Luisi PL (1995b) Polymerase chain reaction in liposomes. Chem Biol 2:677–682

Oberholzer T, Meyer E, Amato I, Lustig A, Monnard P-A (1999) Enzymatic reactions in liposomes using the detergent-induced liposome loading method. Biochim Biophys Acta 1416:57–68

Oliver AE, Deamer DW (1994) a-Helical hydrophobic polypeptides form proton-selective channels in lipid bilayers. Biophys J 66:1364–1379

Oparin AI (1924) The origin of life. Moscow: Izd. Moskovshii Rabochii. English translation in: J. D. Bernal. 1967. The origin of life, pp 199–234. Weidenfeld and Nicolson, London

Oparin AI (1957) The origin of life on the earth. Academic Press, New York

Oparin AI, Orlovskii AF, Bukhlaeva VY, Gladilin KL (1976) Influence of the enzymatic synthesis of polyadenylic acid on a coacervate system. Dokl Akad Nauk SSSR 226:972–974

Oro J, Sherwood E, Eichberg J, Epps D (1978) Formation of phospholipids under primitive earth conditions and roles of membranes in prebiological evolution. In: Deamer DW (ed) Light transducing membranes. Academic Press, London, pp 1–22

Ourisson G, Nakatani Y (1999) Origins of cellular life: molecular foundations and new approaches. Tetrahedron 55:3183–3190

Paula S, Volkov AG, Van Hoek AN, Haines TH, Deamer DW (1996) Permeation of proton, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness. Biophys J 70:339–348

Rasmussen S, Chen L, Nilsson M, Abe S (2003) Bridging nonliving and living matter. Artif Life 9(3):269–316

Rasmussen S, Chen L, Deamer DW, Krakauer DC, Packard NH, Stadler PF, Bedau MA (2004) Evolution. Transitions from nonliving to living matter. Science 303:963–965

Rogers J, Joyce GF (1999) A ribozyme that lacks cytidine. Nature 402:323–325

Rosano HL, Christodolou AP, Feinstein ME (1969) Competition of cations at charged micelle and monolayer interfaces. J Colloid Interface Sci 29:335–344

Rosenquist K, Gabran T, Rydhag L (1981) Studies of permeability across bilayers of lecithin. In: Sfflrat L (ed) 11th Scandinavian symposium on lipids, pp 85–89. Lipidforum, Scandinavian forum for lipid research and technology, Bergen, Norway

Rushdi AI, Simoneit BR (2001) Lipid formation by aqueous Fischer-Tropsch-type synthesis over a temperature range of 100 to 400 degrees C. Orig Life Evol Biosph 31:103–118

Sacerdote MG, Szostak JW (2005) Semi-permeable lipid bilayers exhibit diastereoselectivity favoring ribose; implications for the origins of life. Proc Nat Acad Sci 102:6004–6008

Segre D, Ben-Eli D, Deamer DW, Lancet D (2001) The lipid world. Orig Life Evol Biosph 31: 119–145

Shew RL, Deamer DW (1985) A novel method for encapsulation of macromolecules in liposomes. Biochim Biophys Acta 816:1–8

Skurtveit R, Sjoblom J, Hoiland H (1989) Emulsions under elevated-temperature and pressure conditions. 1. The model system water hexadecanoic acid sodium hexadecanoate decane at 70-degrees-C. J Colloid Interf Sci 133:395–403

Streiff S, Ribeiro N, Wu Z, Gumienna-Konteck E, Elhabiri M, Albrecht-Gary A-M, Ourisson G, Nakatani Y (2007) "Primitive" membrane from polyprenyl phosphates and polyprenyl alcohols. Chem Biol 14:313–319

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390

Tawfik DS, Griffiths AD (1998) Man-made cell-like compartments for molecular evolution. Nat Biotech 16:652–656

Varela FJ, Maturana HR, Uribe R (1974) Autopoiesis: the organization of living systems, its characterization and a model. Biosystems 5:287–296

Voytek SB, Joyce GF (2007) Emergence of a fast-reacting ribozyme that is capable of undergoing continuous evolution. Proc Natl Acad Sci 104:15288–15293

Vlassov A, Khvorova A, Yarus M (2001) Binding and disruption of phospholipid bilayers by supramolecular RNA complexes. Proc Nat Acad Sci USA V98:7706–7711

Wächtershäuser G (1988) Before enzyme and template: theory of surface metabolism. Microbiol Rev 52:452–484

Walde P, Goto A, Monnard P-A, Wessicken M, Luisi PL (1994a) Oparin's reaction revisited: enzymatic synthesis of poly(adenyl acid) in micelles and self-reproducing vesicles. J Am Chem Soc 116:7541–7547

Walde P, Wick R, Fresta M, Mangone A, Luisi PL (1994b) Autopoietic self-reproduction of fatty acid vesicles. J Am Chem Soc 116:11649–11654

Walde P, Ichikawa S (2001) Review. Enzyme inside lipid vesicles: preparation, reactivity and applications. Biomol Eng 18:143–177

Walde P, Wick R, Fresta M, Mangone A, Luisi PL (1994c) Autopoietic self-reproduction of fatty acid vesicles. J Am Chem Soc 116:11649–11654

Wilson TH, Maloney PC (1976) Speculations on the evolution of ion transport mechanisms. Fed Am Soc Exp Biol Proc 35:2174–2179

Woese C (1968) in *The Genetic Code. Harper & Row*

Zaher HS, Unrau PJ (2007) Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. RNA 13:1017–1026

Zhu TF, Szostak JW (2009) Coupled growth and division of model protocell membranes. J Am Chem Soc 131:5705–5713

Chapter 9
Approaches to Building Chemical Cells/Chells: Examples of Relevant Mechanistic 'Couples'

Paul M. Gardner and Benjamin G. Davis

Abstract Operationally functional couples of Container (C), Metabolism (M) and Information (I) and their potential translation into protocellular models are explored through illustration by select examples.

9.1 Introduction: Grounds for Focusing on Container, Metabolism and Information

Considerable research effort in recent years has gone into creating a so-called 'minimal' artificial cell; one with components and characteristics that might be described as 'alive' (Bedau et al. 2000; Deamer 2005; Rasmussen et al. 2004). Such constructs have attracted the use of several, essentially interchangeable, names: artificial cells, (Pohorille and Deamer 2002) protocells, (Rasmussen et al. 2008) minimal cells (the title of this book) or even chemical cells/chells (Cronin et al. 2006; Gardner et al. 2009). Amongst them all, though, are central, and arguably useful, ambiguities surrounding what defines such constructs. Each construct has in essence been predicated on a certain set of instinctive goals and in this chapter we seek simply to outline reflections around our own biased thoughts and starting points in this area. We should emphasise that this work has developed with and has benefited from wonderful collaborative projects with many people and reflects the influences of many discussions (see acknowledgements).

The construction of such an artificial cell has traditionally followed one of two approaches: a so-called 'top-down' approach, in which genes are sequentially knocked out from organisms until only the minimal genome required for sustaining life or a prescribed function is left (Glass et al. 2006), and a 'bottom up' approach, which aims to create and assemble the components of a cell from scratch. One of our central starting points in considering *de novo* approaches was to ask "From *what* new?"; this

P.M. Gardner and B.G. Davis (✉)
Department of Chemistry, University of Oxford, UK
e-mail: Ben.Davis@chem.ox.ac.uk

fundamental questioning of starting point was intended to allow, in principle, the discovery of cell-like systems from varied building blocks and with true breadth of approach. From our personal perspective that was greatly facilitated by a partitioning into functional contributors or parts. For us, this allowed a 'freedom of thought' perhaps more free of conventionality (where the convention may perhaps be considered the current forms of 'life'). This then allowed a choice of contributing functions that have full form in isolation (and good fitness for that isolated purpose) but perhaps unknown or eventually weak putative use when in combination. In essence, we intended that a creatively productive approach, which, having identified one function (such as **C**, containment – see below) would allow us to ask what should or could this container or 'box' be made of.

These functions, for us, have become reduced to just three: **C**, container; **M**, metabolism and **I**, information. We know full well from our own discussions that many will see these three functions as being insufficient for a form of life or to define what might be termed cellularity. These have emerged as those functions that, to us, serve to encapsulate putative emergent outcome with highest strategic orthogonality; in essence these notions of function, **C M I**, are far enough apart and display sufficiently small functional overlap so as to minimise both a) our own conceptual confusion and b) prevent designed or unintentional duplication of function that might lead to processes inhibitory to an eventual couple. For example, in our schema the inclusion of, perhaps, **R** for reproductive capacity, would (and has) led to functional redundancies (e.g., designs for **C** predicated and ultimately weakened by a 'need' for utility in **R**). This philosophical artifice of 'parts' that might or might not work together has certainly helped us to focus more cleanly on these component parts. Their subsequent failure in couples (or even triple) then can stand-or-fall on the basis of some form of pressure, a fitness test (that itself can be varied). This too then stands alone as a failure of that couple against that pressure (a death in a given environment, no more no less) without causing incorrect generalizations about inherent component function utility.

It is therefore not our intention here to justify the adoption of these parts vis-à-vis others and detailed analysis of their possible role and peer assessment of their utility will be covered elsewhere (Bedau et al. in preparation). Instead we will briefly outline here identities and the basis for our favouring these particular components.

Container (**C**), identified by others early (Oparin 1953; Luisi 1996), we see as necessary more than simply as a physical simile to the use of the term cell. Since life is, in essence, a far-from-equilibrium process (kinetic-product control in chemical terms) such barriers vitally maintain differences in Gibbs free energy (G). The form of this barrier or 'chemical potential gradient' therefore could in principle be of any making, even non-traditional e.g., photonic. That said, typical approaches, while therefore not theoretically limited to biological materials or concepts, still typically use lipid membranes as a material for delineation of the cellular system (Murtas et al. 2007); this reflects more than just an interest in pre-biotic chemistry but reflects too the power of self-assembly approaches in protocell chemistry. The kinetics for formation and alteration of such boundaries mediated by non-covalent interactions are sufficiently varied to allow, in turn, a rapid sampling of many

container permutations. This is, as it were, an instinctively dynamic combinatorial solution to generating container fitness based on sound chemical (Curtin-Hammett) principles.

This compartmentalisation is important if the other commonly cited (Ganti 2003; Luisi 2006) requirements for a living system are to be met, those of metabolism (**M**) and information transfer (**I**). Metabolism firstly essentially fuels the differences in G; both by being the source of chemical potential but also by being mechanistically fit for purpose. This means that whilst a constant 'drip feed' of G is not necessary (a difference could be created by a 'bolus' of feedstock, see for example (Bachmann et al. 1992; Gardner et al. 2009) there will either be a limited 'lifetime' for the process or there will be a need to ensure that this bolus is replenished for continued metabolism. Secondly, the units of metabolism (starting products, intermediates, products) will likely play a role (of possible varying directness) in other functions **C** and **I**; the products of **M** may act, in part as material sources for these functions. Together these notions can be delineated more prosaically under terms traditionally associated with biology e.g., trophism, allowing phenotypic analysis of feedstock source and utilization. However, these analyses might overlook some lynchpin concepts surrounding the very different roles of e.g., extraction of feedstock from environments as compared with 'fit' utilization. Analyses of the latter have highlighted the likely pivotal role of catalytic and autocatalytic reactions or networks. These may be achieved through systems of linked reactions or even through linked mechanistic manifolds of a single (auto) catalytic reaction (or more). It is of course near-inevitable that feedstock type will dictate reaction type and may prevent effective molecular surrogacy; access to e.g., phosphorus, in our own biology valuably allows access to higher symmetries in transition states and intermediates that can increase reactivity, compatibility and even selectivity in sets of parallel manifolds.

Information (**I**) or its transmission in some form, it could be argued if one were to anthropomorphise, is perhaps the 'purpose' of life; in many ways the measurement of emergent informational complexity is one of the ways in which we, as a species, have learned to instinctively judge life and the value of it when it is present in given organisms. It can be argued in evolutionary terms that if a 'unit of information', a 'gene', survives and is enhanced then this is emergent success when judged in the context of that gene. Transmission of information with concomitant emergence requires a delicate balance to be struck: fidelity or stability of transmission on one hand (minimal corruption) with kinetic flexibility/sensitivity on the other. In chemical terms informational density can certainly be transmitted with greater 'atom economy' then via polynucleotides, where only the base is present as a 'bit' in the form of a glorified chemical functional group. Ironically RNA's backbone (with four contiguous stereogenic centres per nucleotide and therefore potentially 16 permutations) is a much more informationally-rich unit; yet the fact that these permutations are not harnessed perhaps speaks volumes about the importance of encoding and transmission of information rather than simply its creation (simply noise). Nonetheless, it does not mean that such **I**-storage methods or molecules cannot be usefully considered in *de novo* functional couples; decoding of e.g., sugar

stereogenic (perhaps one could talk of "*stereogenes*") information does take place in some organisms but just not at such a pivotal level of informational hierarchy.

As we hint in the title of this chapter, these functions are, to our mind, more mechanistic couples, some of which may work whilst some may not. However, our principle guiding philosophy has been that the failure of these mechanistic components in particular couples should not disqualify them. It is the wonderful interdependence of this work that is important. Just because emergent or recognisably useful properties might not be obtained from a given couple e.g., C_1–M_4, that result says nothing about the utility of container C_1 or metabolism M_4 in other couples. All functions may have use and this chapter therefore serves to act as list or repository of those couples that have inspired our approaches.

These functions have, in our own limited research examples, (Gardner et al. 2009) focused on chemical principles like those before them (Ganti 2003). Hence we favour the term chell to reinforce a design principle that looks back to properties that may be abstracted with chemical language and philosophy: container construction (might) = self assembly; (emergent) information (might) = functional group or stereochemical diversity; metabolism (might) = competent catalysis. However, this reflects largely our own fascinations and is not intended to represent a judgement on validity of other approaches based on e.g., a synthesis (or antithesis) of biotic strategy or even computational complexity (membrane computing).

9.2 The 'Turing Test' for Artificial Life

Despite these central, functional concepts and their useful role in our thinking the lack of a universally accepted theory of living systems can make it difficult to assess degrees of progress or success towards this goal. This clouding or obscuring of research progress, as measured against functional goals, by a somewhat philosophical barrier parallels that experienced by the Artificial Intelligence community in the 1940s and 1950s (Turing 1950). In order to provide a realisable target within the field of artificial cells, we and others proposed 'an approach to the recognition of 'living' artificial chemical systems based on chemical cells ('chells') as a *Gedankenexperiment* that exploits a cellular imitation game' (Cronin et al. 2006).

We noted that a similar fundamental problem was encountered in the area of artificial intelligence (AI), wherein researchers wrestled with the question of whether machines can think. Alan Turing effectively circumvented this problem by proposing the so-called 'imitation game' (Turing 1950) which later became known as the 'Turing test'. The question 'can machines think?' was dismissed by Turing as not being worthy of useful consideration and was instead redefined by an operational scenario for AI: can machines imitate the act of thinking such that an interrogator cannot distinguish between a machine and a person? This was an elegant operational way of assessing whether a machine can think, regardless of what was meant by thought; the question 'can machines think?' was, in this way,

replaced by 'are there imaginable digital computers that would do well in the imitation game?' Since its introduction, the concept of the Turing test has helped advance understanding of AI and represented a powerful shift in the research culture in this area.

We therefore proposed to use a similar approach, that of an 'imitation game' that could help answer the analogous question 'Are there artificial chemical systems that would do well in the cellular imitation game?' (Table 9.1). We proposed that the separation of physical embodiment and ability, (as in the original Turing test for AI) should be preserved. Hence the synthetic cell could be built as a chemoton (Ganti 2003) the size of a family car or as a microscopic lipid membrane. This we hoped would avoid the prejudice that life must be based upon our own biomolecules such as DNA and proteins. Although it should be acknowledged that owing to our more detailed knowledge of such systems and the sometimes difficulty of conceiving of other forms of life, biomimetic systems will be the most common approach adopted.

The concept of the test is that the artificial cell in question should be able to communicate with natural cells (directly or via a relay) in such a way that, as in the spirit of the original imitation game, it can be interrogated by the latter, in a language sufficiently sophisticated so as to appropriately distinguish between alternative outputs from realizable experiments. We suggested, as an example, a cellular imitation game set-up (Fig. 9.1) where the chell might imitate a natural cell and where the latter plays the role of interrogator through properly configured screening and life-support devices (Wagler et al. 2004). In this idealized cellular imitation game, interrogation could take place following any of a series of increasingly more complex and sophisticated mechanisms, starting perhaps with a relatively simple quorum-sensing type of language based on low molecular weight signalling molecules (Winzer et al. 2002) (see below), and moving perhaps toward mechanical transduction, (Bao and Suresh 2003) bio-film formation or swarming patterns of behaviour (Ben-Jacob et al. 1994).

We envisaged that the 'Turing test' concept might allow measurable and achievable goals within the field of artificial cells. As knowledge advances, more refined versions of the cellular imitation game could be implemented. This mechanism

Table 9.1 Comparison of Turing tests for intelligence and life (Figure adapted from Cronin et al. in Nature Biotechnology 2006 (Cronin et al. 2006). Used with permission)

	Turing test for intelligence	Turing test for life
Imitated emergent property	Thought	Cellular functions (e.g., metabolism, evolution and containment)
Embodiment of property	Computational digital machine	Chemical system (e.g., artificial cell or 'chell')
Probing mechanism	Questions/answers mediated by natural language	Questions/answers mediated by (natural) physico-chemical language (e.g. interconversion of chemical potentials)
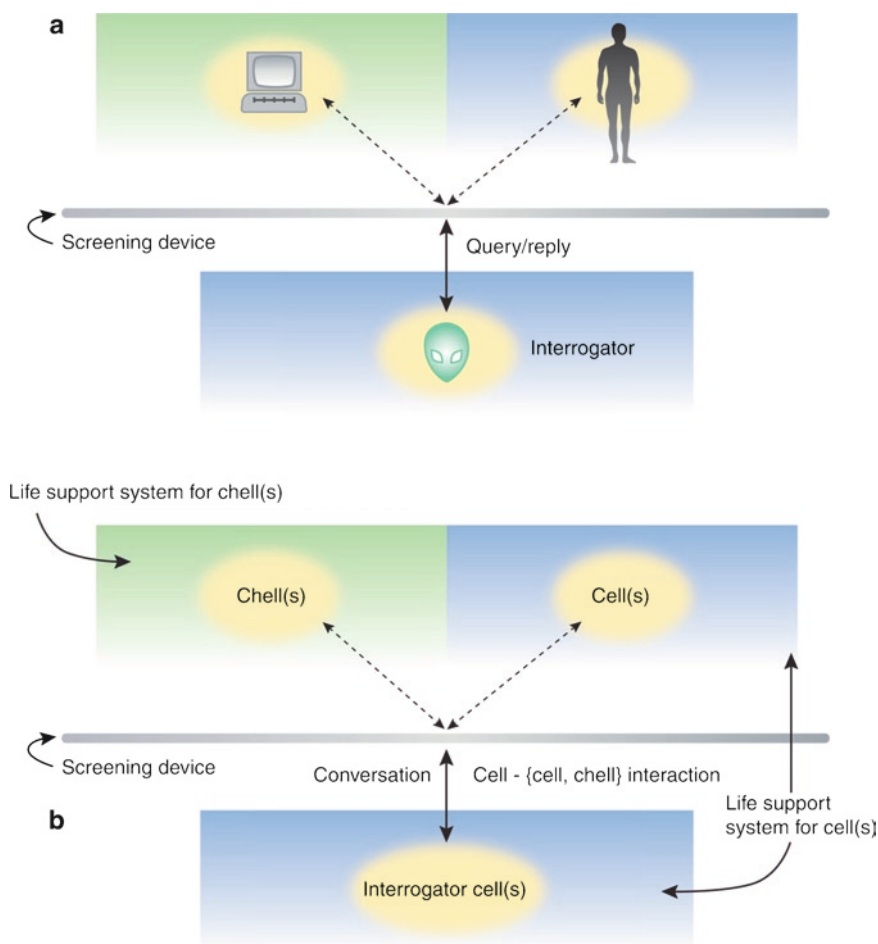
Fig. 9.1 Different takes on Turing. (**a**) Representation of the classic Turing test with an intelligent interrogator (that is, a person, interacting with two compartments, each containing either a computer or a person, the location of which is unknown to the interrogator). (**b**) The extension of this interrogator or interaction between cells and chells (see text) is depicted. It is suggested that cell signalling could be used as the medium for conveying the interactions but other, perhaps simpler, mechanisms could be used (Taken from Cronin et al. in Nature Biotechnology 2006 (Cronin et al. 2006). Used with permission)

strengthens the imitation game by implementing what Harel (who proposed a similar test for systems biology modelling) (Harel 2005) has called a "Popperian twist." Thus, even if a chell successfully passed a test based on quorum sensing, it might still fail to succeed on a more sophisticated version of the imitation game where the interrogator cell uses mechanical transduction at the cell membrane or horizontal gene transfer as a probing language. Thus, the refinement of the system to pass the next test would represent another advancement in the field.

9.3   Paired Couplings of Components for Life

While no clear demonstrations of the three-way coupling of **C–M–I** have yet been achieved, progress is being made towards this goal, by way of the stepping-stone of pair-wise coupling of these components. Considered below are some of the more notable efforts towards coupling each pair, with the aim of highlighting future possibilities for advances in these fields. Discussions have led to the proposition that a competition might be set based on assessing the success of such couplings (Bedau et al. in preparation). The details of this intriguing idea will be published elsewhere and so it is not the purpose of this chapter to revisit this area. Instead, we hope here to pique interest in the notions of coupling **C**, **M**, **I** by listing some select examples that have inspired us.

9.4   Container and Metabolism Coupling (C–M)

The coupling of a container with a chemical metabolism is perhaps conceptually the most simple of the three systems to be considered. The difficulty comes in determining exactly what constitutes a 'metabolism' in such a context. A simplistic definition of a metabolism is a system of chemical reactions required to maintain a system, but it could be argued that such a system, if capable of complete self-maintenance (autopoiesis) would be alive in and of itself (Luisi 2006). Clearly this requires some simplification. As such, the notion of a metabolism in these pairwise couplings considered below will be one in which there is a series of repeating chemical reactions – the repetition typically achieved by autocatalysis.

   So how might a container and a metabolism be coupled? The simplest idea, and one which has been more extensively explored, is that of self-replicating containers. That is, a container, in or around which reactions are occurring that create more of the container component(s), and over time these reactions (the metabolism) allow self-replication of containers. Because of the link with research on artificial cells and that of the origin of life, the majority of research in this field has been on so-called 'prebiotically plausible' lipid containers (Segre et al. 2001). These have-typically been fatty acids, which have a chemically simple structure, and form vesicles at pH values close to their pK_a, at which point hydrogen bonding stabilises the interaction between head-groups, although the vesicles formed are still comparatively unstable compared to those formed from phospholipids and the like. The important role of fatty acids in primary and secondary metabolism of extant living cells lends weight to the argument that such fatty acids were incorporated into cells in early stages of life on Earth.

   A crucial breakthrough in research on fatty-acid vesicles was described by Luisi and co-workers in 1992, who demonstrated autocatalytic self-replication of caprylate vesicles (Bachmann et al. 1992). Alkaline hydrolysis of ethyl caprylate produces sodium caprylate, which slowly aggregates. The aggregates then catalyse the hydrolysis of further ethyl caprylate on their surface, rapidly increasing the rate
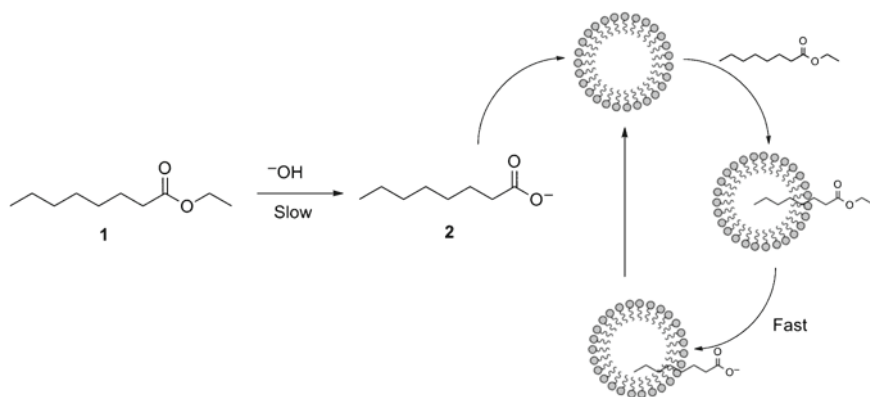
Scheme 9.1 Scheme for the autocatalytic formation of caprylate vesicles described by Luisi and co-workers(Bachmann et al. 1992). Hydrolysis of ethyl caprylate **1** is followed by aggregation into vesicles, which catalyse further hydrolysis of the caprylate

of vesicle formation as sodium caprylate is incorporated into the membranes. The vesicles grow, bud and divide, in a potentially infinitely self-sustaining (autopoetic) manner (Scheme 9.1).

How is the system autocatalytic? A detailed theoretical model (Mavelli and Luisi 1996) considered the reaction must occur at either the organic/water macroscopic interface, or at the microscopic surface of the aggregates, owing to the hydrophobic nature of the anhydride. As long as this surface hydrolysis is the rate-limiting step, it can be shown that the rate of hydrolysis is dependent upon the area of the amphiphilic surface present. As such, the initial slow-step of the reaction takes place only at the organic/water interface before vesicles have formed. Then, as further amphiphile is produced and aggregates to form vesicles, the increased surface area increases the reaction rate of the hydrolysis reaction further. It is worth noting that the time of the 'slow' initial step is dependent on the 'critical aggregation concentration' of the lipid, which in turn is dependent on the lipid structure, in particular the length of the carbon chain. This has potential implications for the choice of lipid for an artificial cell – in theory the rate of division could be usefully tuned by careful choice of lipid.

Fatty acid-vesicle replication by hydrolysis has been extensively studied, to the extent that the use of such vesicles could be considered paradigmatic (Sole et al. 2007). The work of Luisi et al. has been followed by work on other fatty acids. Using a similar reaction system, the self-replication of vesicles formed form the 18-carbon oleic acid was reported shortly after the caprylate vesicles were described (Walde et al. 1994b). The longer chain length causes the vesicles formed to be considerably more kinetically stable, owing to the poorer solubility of the tail groups in aqueous solution, and thus these lipids are far more likely to be of use in the realisation of artificial cells. The self-replication of unsaturated chains fatty acids has also been reported (Chen and Szostak 2004). Although a detailed discussion of lipid dynamics is beyond the scope of this article, unsaturated fatty acids

can have considerably different properties from their saturated counterparts (Cevc 1993). It is likely that membranes sustaining metabolism and information are liable to require a variety of different lipid derivatives for sufficient stability (and flexibility in emergence – see above), (Mansy 2009) so the demonstration that different types of lipid can replicate in these ways is an important one.

Other self-replicating systems have also been described. Takakura et al. described a system based on an imine-forming reaction, using amphiphiles that are completely non-biological. In this system an aniline and a dioxalane are the components used. Under acidic conditions, and in the presence of a third synthetic catalytic amphiphile, the dioxalane is hydrolysed to the aldehyde, and the aniline and aldehyde then form an imine, linking the two components into one vesicle-forming amphiphile. The amphiphilic nature of the catalyst requires this process to take place within the membrane; however the catalyst itself is not replicated, preventing this process from being autocatalytic, and vesicle division is eventually halted by dilution of the catalyst between the newly-formed vesicle membranes. Nonetheless, this remains an important demonstration that non-biological molecules can and indeed should be considered as a basis for artificial cellular compartments.

Another system worthy of note is one of replicating reverse micelles described by Bachman et al. (Bachmann et al. 1991). Reverse micelles are water droplets in organic solvent, with a monolayer of amphiphile around each droplet to stabilise it. In a similar mechanism to that described above for fatty-acid vesicles, reverse micelles formed of octanoate can be made to replicate by the alkaline-catalysed hydrolysis of octyl octanoate at the organic solvent/lipid interface. Owing to the constant volume of water present, the daughter reverse micelles are smaller, however, and do not grow in the same manner as the fatty-acid vesicles, so only a few rounds of replication were possible. However, this is an important demonstration of the possibility of non-aqueous solvents and reaction systems in artificial cell research. Owing to biases both biological and chemical, water is often viewed as the only solvent in which artificial cells may be constructed, (Benner et al. 2004) but from a chemical point of view, a system in an organic solvent passing the Turing test would be just as valid a demonstration of an artificial cell as one in a 'prebiotically plausible' environment.

9.5   C–I

The coupling of container and information is, depending on definition, either the hardest of the pairwise couplings to be considered, or already solved. Such a paradox comes about from the question, once again, of what exactly constitutes a 'pairwise' coupling in this context. If it is sufficient, for example, to demonstrate information transfer *within* a bounded container, then several examples of such a system already exist (see below). On the other hand, if the two are required to be linked by chemical reaction, then how would this be achieved without the intervention of a metabolism

of reactions as an intermediary? Does such linking thereby imply metabolism by definition? Clearly this can only be the case if the metabolism is complete, such that it maintains itself, the container and the information system. As such, this section will briefly consider the advances in demonstrating information encoding or transfer within a bounded container, and move on to discussing examples of systems in which the two are moving towards a chemical link.

The principal examples of information transfer within bounded containers have explored nucleic acid replication and protein expression within vesicles. An ultimate aim is to introduce a DNA cassette capable of catalysing its own replication and protein expression (Luisi et al. 2006). The production of poly(A) in vesicles has been described independently by two groups (Chakrabarti et al. 1994; Walde et al. 1994a). Polynucleotide phosphorylase was entrapped within relatively permeable vesicles (with short chain amphiphiles). ATP could then diffuse through the vesicle membranes and was polymerised by the enzyme, whereupon it could be visualised by fluorescence of intercalators or gel electrophoresis. Building upon these reports, the polymerase chain reaction (PCR) in a lipid vesicle was also described (Oberholzer et al. 1995a). In this case the vesicles were stable and relatively impermeable, and the components of the PCR reaction (T7 polymerase, dNTPs and a plasmid template) were encapsulated and the reaction initiated by temperature change.

RNA synthesis using all four NTPs and Qβ replicase has also been described (Oberholzer et al. 1995b). This used a similar system to that for the ATP replication described above, although the more complex MDV-1 template was used. Significantly in this case, however, incorporation of the system into the self-replicating oleate/caprylate vesicle system allowed RNA replication and vesicle replication to occur simultaneously. However, these two reactions were independent, and the vesicles replicated considerably faster than the RNA, resulting in 'death by dilution', in that the majority of vesicles in the final solution contain no RNA. As such, this system cannot be considered coupled; the two processes were occurring in the same vessel, but not in a way that affected one-another.

Gene expression in vesicles has also been described, the ultimate aim of such experiments being to increase the complexity of the expression system such that enzymes for replication are encoded. However only single protein expression has so far been reported. The first protein to be expressed in this way was poly-L-phenylalanine, (Oberholzer et al. 1999) utilising a poly(U) RNA encapsulated with phenylalanine, ribosomes, Phe-tRNA and elongation factors. Since then, research has focussed on optimising expression of single proteins, usually GFP since this protein allows easy and quantitative fluorescent readout of the level of its own expression. The usual strategy is to entrap the RNA template, tRNAs, ribosomes, and amino acids etc within the liposome. GFP expression for example has been reported in lechithin vesicles (Yu et al. 2001) (analysed by flow cytometry), small unilamellar vesicles using the ethanol-injection method (Oberholzer and Luisi 2002) and within giant unilamellar vesicles, with expression directly observable by laser-excited fluorescence microscopy (Nomura et al. 2003).

In all of these examples, however, the presence of a membrane is somewhat incidental. While the experiments described represent significant technical breakthroughs in the field of vesicle chemistry, the vesicle itself could be thought of as just a reaction chamber, and there may be limited difference in the actual chemistry of the reactions due to the membrane, as opposed to, say, carrying out the reaction in a glass test tube. However, a few examples in which the membrane does play a more prominent role in the control of the reactions have been published. A particularly interesting example combining the replication of genetic polymers in vesicles and diffusion across membranes was recently reported by Szostak and co-workers (Mansy et al. 2008). It was shown that small molecules such as modified-dNTPs can cross myristoleate fatty-acid based vesicle membranes. These dNTPs could then be incorporated into a template-directed poly(C)-DNA-mimic synthesis (Scheme 9.2). Whilst the monomers were freely diffusible, once incorporated into polymers they were effectively encapsulated within the vesicles. It was hypothesised that the resultant increase in osmotic pressure might cause initiation of vesicle replication if a self-replicating fatty-acid system with similar properties could be found. This would allow vesicle division to be connected to polymer replication (Chen et al. 2004) (Scheme 9.2).

The crucial role of the membrane in this example was to control diffusion of the feedstock dNTPs into and out of the cell. By careful control of the membrane composition, the authors were able to ensure that the permeability of the membrane was such that the dNTPs could cross, whereas large molecules and template could not. As such, the membrane acted as a selectively permeable barrier, in a manner conceptually akin to some living cells. The resultant hypothesis that osmotic pressure could couple DNA replication to vesicle replication is an intriguing one, and if successful would be a true coupling of container and information.

A similar coupling was demonstrated by Noireaux et al., (Noireaux et al. 2005), who also used a semi-permeable membrane system to allow passage of small-molecule solutes whilst retaining macromolecules. In this case, however, the vesicle

Small molecules diffuse in   DNA Replication   Swelling under osmotic pressure   Vesicle division
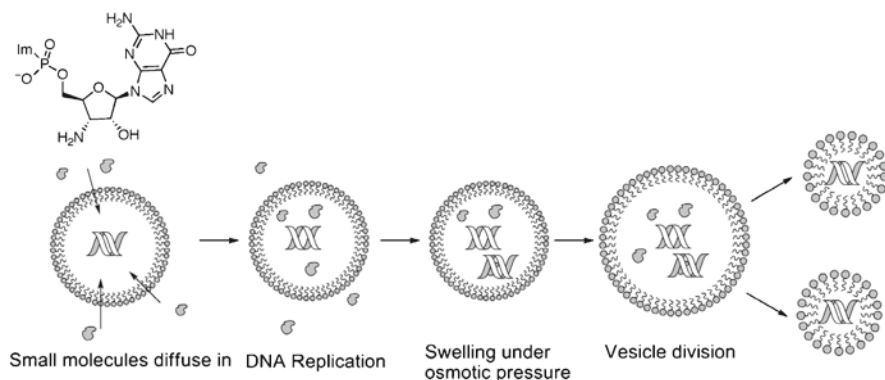
Scheme 9.2 Heterotroph vesicle system suggested by Szostak and co-workers (Mansy et al. 2008)

was composed of relatively impermeable phospholipids, and the permeability was conferred by the insertion of the pore protein α-hemolysin (Walker et al. 1992) into the membrane. The protein forms a beta-barrel structure that spontaneously inserts into lipid bilayers, and the pore itself is of sufficient diameter to allow the passage of dNTPs. In contrast to the previous example, gene expression rather than DNA replication was performed inside the vesicle, and the robust expression of GFP was demonstrated. Critically, the α-hemolysin was not present in the system initially, but was expressed by the DNA entrapped in the cell. This gave rise to a form of two way interaction – the information modified the container, which in turn allowed the further expression of the information by allowing additional dNTPs into the vesicle system, extending the observable gene expression by several days. Although requiring a relatively advanced genetic system and highly evolved protein, this was an important demonstration of the possibilities for affecting the membrane by the genetic component, and a possible next step might be to couple this with the vesicle self-replication as described above.

9.6   M–I

The union of metabolism and information is critical to the existence of life – so much so that it is often suggested that one of the natural outcomes of this union, self replication, is the only criterion for life (Cleland and Chyba 2002). Self-replication of a chemical entity is inevitably a combination of metabolism and information – the replicated object itself providing the 'information' and the reactions that produce it the metabolism. As such, it may even be argued that some of the systems discussed above contain metabolism, information and containers, given that the containers house replicating DNA, although these systems clearly are not coupled.

DNA and RNA remain the prominent examples of template-driven replicating molecules capable of undergoing replication by complete polymerisation from monomer units. Efforts are focussed on creating new such molecules, (Wittung et al. 1994; Bean et al. 2006). Attempts to create DNA and RNA sequences that can self-replicate without the intervention of enzymes increase the plausibility of the RNA world hypothesis. Although some success has been attained in this route with DNA, (Li and Nicolaou 1994) the 'holy grail' of the approach is templated replicating RNA, which has not yet been demonstrated.

Other approaches to such information-metabolism coupling use peptides as the replicating units. Such replication was first shown by Ghadiri et al. in 1996; (Lee et al. 1996) in this example a 32-mer polypeptide unit was shown to be capable of autocatalytically catalysing its own formation from constituent 17 and 15-mer units, acting as a template (Fig. 9.2). The peptide and its constituents form helices, and hydrophobic interactions between helices are the cause of the templating effect. The helical requirement has an interesting effect: that of allowing amplification of a chiral product from a heterochiral pool of reagents (Saghatelian et al. 2001). The 'handedness' of
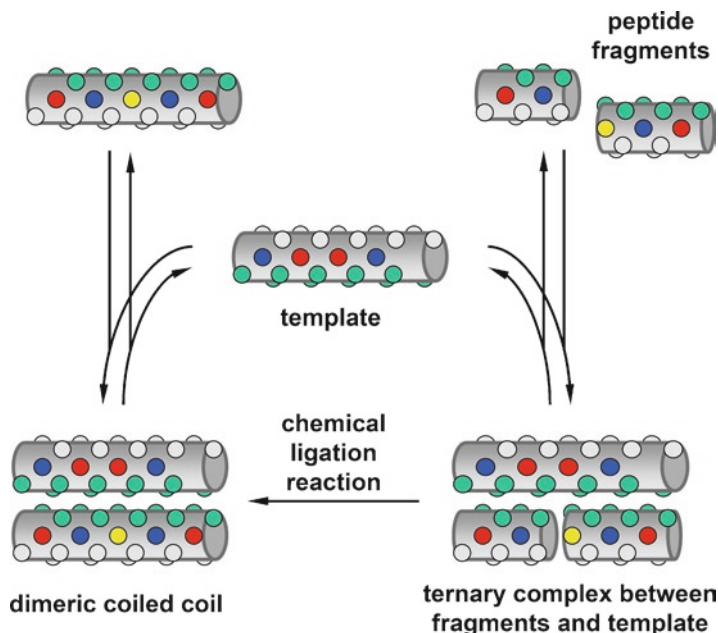
Fig. 9.2 Ghadiri's autocatalytic cycle for a self-replicating coiled-coil peptide. Template binding promotes the joining of the smaller peptides. The product extended peptide then templates and catalyzes product formation in the cycle (Taken from (Issac et al. 2001). Permissions requested)

the helix (or indeed ability to form a helix at all – mixed DL amino acid chains do not form regular helices) is crucial for the catalysis to occur, and as such only the product with the same chirality as the template is preferentially formed.

The limitations of this approach are clear, though; the peptides were only capable of catalysing the formation of amide bonds from activated polypeptides, and only a single sequence, or small number of sequences were used. There have been a few reports in which multiple sequences are used: for example dynamic error correction to eliminate mutations has been demonstrated (Kay et al. 1998). In this case, 'non-self' peptides were coupled at a slower rate than the template component peptides, and as such the 'mutated' peptides were selected against. Further developments in this field will hopefully allow the construction of templated peptides from much smaller components.

Moving away from biological molecules, a few less biotic, non-organic systems have demonstrated the potential for the coupling of metabolic reactions with information, although here the 'information' may require a broader definition (see above). The Soai reaction (Soai et al. 1995), is one of the most significant discoveries, not only with respect to the topic in hand, but with regard to the origin of biological chirality and asymmetric catalysis. The theoretical basis for this reaction was described over 40 years before a laboratory example was described: in an analysis of 'spontaneous asymmetric synthesis', Frank (Frank 1953)
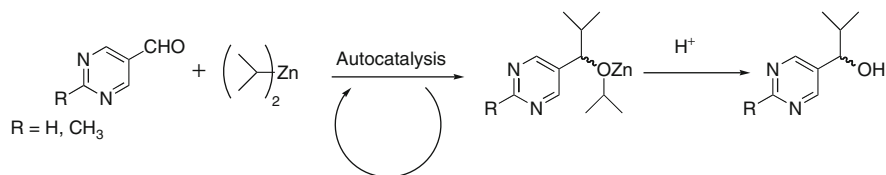
Scheme 9.3 The Soai reaction

described the autocatalytic production of an entantiomerically enriched product from racemic mixture. In this case, the 'information' is the chirality of the product, which then acts as a catalyst for its own production, favouring itself over its enantiomer.

In the Soai reaction (Scheme 9.3), diisopropyl zinc converts pyrimidyl aldehydes to chiral alcohols with considerable enantiomeric amplification – a 10% ee in the starting material results in over 70% ee in the product mixture. This product can then be used as catalyst in another reaction, giving still further enantiomeric enrichment up to over 90%. Initially the reaction was investigated only with the product present as a chiral bias, but other chiral molecules at as low as 0.1% ee have been shown to induce high enantiomeric excess (Shibata et al. 1998), as have inorganic minerals such as chiral quartz crystals (Soai et al. 1999) and even plane-polarised light (Shibata et al. 1998). Indeed, detailed kinetic investigations (which reveals dimers to be the active catalytic agents) have shown that stochastic processes forming catalysts can result in enantiomeric enrichment by themselves (Gridnev et al. 2003). While the reaction is, in its current form at least, unlikely to be relevant in artificial cells (and likely not in origin of life discussions), the principles demonstrated and the kinetic analysis that has resulted (Saito et al. 2006; Islas et al. 2005; Gridnev et al. 2003) has been critical in our understanding of this information-reaction coupling.

9.7 Future Directions: Towards the Union of CMI

Our own efforts (Gardner et al. 2009) have tried to build in some small way on these motivating examples. The use of another, rare example of a single, multi-manifold, autocatalytic reaction has allowed us to move towards notions of autotrophy that might allow interaction with container and produce information that could be selectively interpreted or decoded. In this system, we were able to employ an aqueous but abiotic sugar-synthesizing reaction, the formose reaction, (Butlerow 1861) to generate sugars within a phospholipid protocell from basic feedstock. Charging of the cell/chell with a bolus of formaldehyde and Ca(II), triggering of this autocatalytic reaction as a pseudo-metabolism by pH change followed by ion-channel enabled release of the sugars allowed the stimulation of a quorum response in a neighbouring, natural bacterial partner cell population (as judged by bioluminescence). Intriguingly,
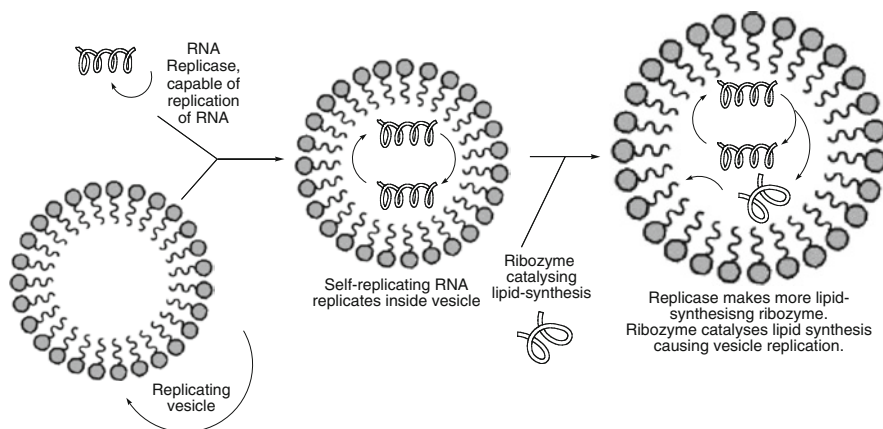
Scheme 9.4 Minimal RNA cell based on two ribozymes described by Luisi, Szostak and Bartel(Szostak et al. 2001)

the sugar products formed by the protocell differed from those formed in bulk for-mose reactions and suggest influence of container upon **M** as a source of **I**.

A particularly elegant synthesis was described by Luisi, Szostak and Bartel (Szostak et al. 2001) who delineate a powerful, possible artificial cell containing just two ribozymes (Kruger et al. 1982). These ribozymes are an RNA replicase (capable of templated RNA replication), and a lipid-synthesising ribozyme. Replication of the latter would result in faster production of lipids which would result in cell division coupled to RNA replication (Scheme 9.4).

References

Bachmann PA, Luisi PL, Lang J (1991) Self-replicating reverse micelles. Chimia 45:266–268
Bachmann PA, Luisi PL, Lang J (1992) Autocatalytic self-replicating micelles as models for prebiotic structures. Nature 357:57–59
Bao G, Suresh S (2003) Cell and molecular mechanics of biological materials. Nat Mater 2:715–725

Bean HD, Anet FAL, Gould IR, Hud NV (2006) Glyoxylate as a backbone linkage for a prebiotic ancestor of RNA. Orig Life Evol Biosph 36:39–63

Bedau MA, Krasnogor N, Davis BG, Cronin L, Deamer D, Hanczyc M, Harel D, Lancet D, Packard N, Rasmussen S, Schroeder S, Stano P, Whitaker B, Valencia A Critical assessment of artificial chemical cells: milestones in the quest for artificial life (in prep)

Bedau MA, Mccaskill JS, Packard NH, Rasmussen S, Adami C, Green DG, Ikegami T, Kaneko K, Ray TS (2000) Open problems in artificial life. Artif Life 6:363–376

Ben-Jacob E, Schochet O, Tenenbaum A, Cohen I, Czirok A, Vicsek T (1994) Generic modelling of cooperative growth patterns in bacterial colonies. Nature 368:46–49

Benner SA, Ricardo A, Carrigan MA (2004) Is there a common chemical model for life in the universe? Curr Opin Chem Biol 8:672–689

Butlerow A (1861) Bildung einer zuckerartigen Substanz durch Synthese. Annalen 120:295–298

Cevc G (ed) (1993) Phospholipids handbook. Marcel Decker, New York

Chakrabarti AC, Breaker RR, Joyce GF, Deamer DW (1994) Production of RNA by a polymerase protein encapsulated within phospholipid-vesicles. J Mol Evol 39:555–559

Chen IA, Roberts RW, Szostak JW (2004) The emergence of competition between model proto-cells. Science 305:1474–1476

Chen IA, Szostak JW (2004) A kinetic study of the growth of fatty acid vesicles. Biophys J 87:988–998

Cleland CE, Chyba CF (2002) Defining 'Life'. Orig Life Evol Biosph 32:387–393

Cronin L, Krasnogor N, Davis BG, Alexander C, Robertson N, Steinke JHG, Schroeder SLM, Khlobystov AN, Cooper G, Gardner PM, Siepmann P, Whitaker BJ, Marsh D (2006) The imitation game-a computational chemical approach to recognizing life. Nat Biotechnol 24:1203–1206

Deamer D (2005) A giant step towards artificial life? Trends Biotechnol 23:336–338

Frank FC (1953) On spontaneous asymmetric synthesis. Biochim Biophy Acta 11:459–463

Ganti T (2003) The principles of life. Oxford University Press, New York

Gardner PM, Winzer K, Davis BG (2009) Sugar synthesis in a protocellular model leads to a cell signalling response in bacteria. Nat Chem 1:377–383

Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA III, Smith HO, Venter JC (2006) Essential genes of a minimal bacterium. Proc Nat Acad Sci U S A 103:425–430

Gridnev ID, Serafimov JM, Quiney H, Brown JM (2003) Reflections on spontaneous asymmetric synthesis by amplifying autocatalysis. Org Biomol Chem 1:3811–3819

Harel D (2005) A Turing-like test for biological modeling. Nat Biotechnol 23:495–496

Islas JR, Lavabre D, Grevy J-M, Lamoneda RH, Cabrera HR, Micheau J-C, Buhse T (2005) Mirror-symmetry breaking in the Soai reaction: A kinetic understanding. Proc Natl Acad Sci U S A 102:13743–13748

Issac R, Ham Y-W, Chmielewski J (2001) The design of self-replicating helical peptides. Curr Opin Struct Biol 11:458–463

Kay S, David HL, Jose AM, Michael V, Ghadiri MR (1998) Dynamic error correction in autocata-lytic peptide networks. Angew Chem Int Ed 37:126–128

Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR (1982) Self-splicing RNA – auto-excision and auto-cyclization of the ribosomal-RNA intervening sequence of tetrahymena. Cell 31:147–157

Lee DH, Granja JR, Martinez JA, Severin K, Ghadiri MR (1996) A self-replicating peptide. Nature 382:525–528

Li T, Nicolaou KC (1994) Chemical self-replication of palindromic duplex DNA. Nature 369:218–221

Luisi PL (1996) Self-reproduction of micelles and vesicles: models for the mechanisms of life from the perspective of compartmented chemistry. Adv Chem Phys 92:425–438

Luisi PL (2006) The emergence of life. from chemical origins to synthetic biology. Cambridge University Press, Cambridge

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 93:1–13

Mansy SS (2009) Model protocells from single-chain lipids. Int J Mol Sci 10:835–843

Mansy SS, Schrum JP, Krishnamurthy M, Tobe S, Treco DA, Szostak JW (2008) Template-directed synthesis of a genetic polymer in a model protocell. Nature 454:122–125

Mavelli F, Luisi PL (1996) Autopoietic self-reproducing vesicles: a simplified kinetic model. J Phys Chem 100:16600–16607

Murtas G, Kuruma Y, Bianchini P, Diaspro A, Luisi PL (2007) Protein synthesis in liposomes with a minimal set of enzymes. Biochem Biophys Res Commun 363:12–17

Noireaux V, Bar-Ziv R, Godefroy J, Salman H, Libchaber A (2005) Toward an artificial cell based on gene expression in vesicles. Phys Biol 2

Nomura SIM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. ChemBioChem 4:1172–1175

Oberholzer T, Albrizio M, Luisi PL (1995a) Polymerase chain reaction in liposomes. Chem Biol 2:677–682

Oberholzer T, Luisi PL (2002) The use of liposomes for constructing cell models. J Biol Phys 28:733–744

Oberholzer T, Nierhaus KH, Luisi PL (1999) Protein expression in liposomes. Biochem Biophys Res Commun 261:238–241

Oberholzer T, Wick R, Luisi PL, Biebricher CK (1995b) Enzymic RNA replication in self-reproducing vesicles: an approach to a minimal cell. Biochem Biophys Res Commun 207:250–257

Oparin AI (1953) The origin of life. Dover, New York

Pohorille A, Deamer D (2002) Artificial cells: prospects for biotechnology. Trends Biotechnol 20:123–128

Rasmussen S, Bedau MA, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF (2008) Protocells: bridging nonliving and living matter. The MIT Press, Cambridge

Rasmussen S, Chen L, Deamer D, Krakauer DC, Packard NH, Stadler PF, Bedau MA (2004) Transitions from nonliving to living matter. Science 303:963–965

Saghatelian A, Yokobayashi Y, Soltani K, Ghadiri MR (2001) A chiroselective peptide replicator. Nature 409:797–801

Saito Y, Sugimori T, Hyuga H (2006) Stochastic approach to enantiomeric excess amplification and chiral symmetry breaking. Los Alamos National Laboratory, Preprint Archive, Condensed Matter, pp 1–24, arXiv:cond-mat/0612385

Segre D, Ben-Eli D, Deamer DW, Lancet D (2001) The lipid world. Orig Life Evol Biosph 31:119–145

Shibata T, Yamamoto J, Matsumoto N, Yonekubo S, Osanai S, Soai K (1998) Amplification of a slight enantiomeric imbalance in molecules based on asymmetric autocatalysis: the first correlation between high enantiomeric enrichment in a chiral molecule and circularly polarized light. J Am Chem Soc 120:12157–12158

Soai K, Osanai S, Kadowaki K, Yonekubo S, Shibata T, Sato I (1999) d- and l-quartz-promoted highly enantioselective synthesis of a chiral organic compound. J Am Chem Soc 121:11235–11236

Soai K, Shibata T, Morioka H, Choji K (1995) Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. Nature 378:767–768

Sole RV, Munteanu A, Rodriguez-Caso C, Macia J (2007) Synthetic protocell biology: from reproduction to computation. Philos Trans Royal Soc B Biol Sci 362:1727–1739

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390

Turing AM (1950) Computing machinery and intelligence. Mind 59:433–460

Wagler PF, Tangen U, Maeke T, Chemnitz S, Juenger M, Mccaskill JS (2004) Molecular systems on-chip (MSoC) steps forward for programmable biosystems. Proc SPIE Int Soc Opt Eng 5389:298–305

Walde P, Goto A, Monnard PA, Wessicken M, Luisi PL (1994a) Oparin's reactions revisited: enzymatic synthesis of poly(adenylic acid) in micelles and self-reproducing vesicles. J Am Chem Soc 116:7541–7547

Walde P, Wick R, Fresta M, Mangone A, Luisi PL (1994b) Autopoietic self-reproduction of fatty acid vesicles. J Am Chem Soc 116:11649–11654

Walker B, Krishnasastry M, Zorn L, Kasianowicz J, Bayley H (1992) Functional expression of the alpha-hemolysin of Staphylococcus aureus in intact escherichia-coli and in cell lysates – deletion of 5 c-terminal amino-acids selectively impairs hemolytic-activity. J Biol Chem 267:10902–10909

Winzer K, Hardie KR, Williams P (2002) Bacterial cell-to-cell communication: sorry, can't talk now – gone to lunch. Curr Opin Microbiol 5:216–222

Wittung P, Nielsen PE, Buchardt O, Egholm M, Norden B (1994) DNA-like double helix formed by peptide nucleic acid. Nature 368:561–563

Yu W, Sato K, Wakabayashi M, Nakaishi T, Ko-Mitamura EP, Shima Y, Urabe I, Yomo T (2001) Synthesis of functional protein in liposome. J Biosci Bioeng 92:590–593

Part III
Steps Towards Minimal Life

Chapter 10
Construction of an In Vitro Model of a Living Cellular System

K. Yoshikawa, S.M. Nomura, K. Tsumoto, and K. Takiguchi

Abstract The use of materials that are fully compatible with living cells may be the most reasonable plan for producing artificial models of cells. In this contribution, we present the results of some recent studies toward the construction of artificial cells, and explain methods for emulating several important specific functions of real cells in artificial liposomes. We discuss several aspects of the current experimental approaches, from the optimal construction of giant vesicles (GVs) to the realization of complex vesicle-based molecular system. The entrapment of DNA and enzymes inside GVs are discussed, as well as the synthesis of water soluble proteins inside vesicles. Emphasis is given on the new approaches to synthesize membrane-soluble proteins on the vesicle membrane. The case of connexin 43 (Cx43)-containing liposomes is described in details, as well as the extension of the proteoliposome technology to giant lipsomes, suitable for direct microscopy imaging. Application of the current in vitro synthetic approaches to the study of cell morphology is also discussed, by referring to the partial reconstitution of cytoskeleton inside GVs. In our recent study, G-actin and BBMI were simultaneously entrapped inside liposomes, through the method of natural swelling, resulting in efficient GV transformation. We conclude with some general consideration about the future work on cell models using vesicles.

K. Yoshikawa (✉)
Graduate School of Science, Kyoto University, 606-8502, Sakyo-ku, Kyoto, Japan
e-mail: yoshikaw@scphys.kyoto-u.ac.jp

S.M. Nomura
Institute for Integrated Cell-Material Sciences, JST PRESTO, Kyoto University, 606-8501, Sakyo-ku, Kyoto, Japan

K. Tsumoto
Graduate School of Engineering, Mie University, 514-8507, Tsu-shi, Mie, Japan

K. Takiguchi
Graduate School of Science, Nagoya University, 464-8602, Chikusa-ku, Nagoya, Japan

10.1   General Introduction

Over the past several decades there have been great advances in our understanding of biomolecules. Studies on individual biomolecules have provided useful information to approach the long-standing problem of "What is life?" Life is definitely an autonomous system, i.e., the many different molecules that play essential roles in life exhibit dynamic organization. Thus, if we only consider the current reductionist approach, it is impossible to understand how cells maintain their lives. Therefore, a constructive approach is needed to answer this question. The world is full of living things with tens of millions of cell types. The use of materials that are fully compatible with living cells may be the most reasonable plan for producing artificial models of cells. In this section, we present the results of recent studies toward the construction of artificial cells, and explain methods for emulating several important specific functions of real cells in artificial liposomes. We focus our explanation on the experimental results from our research group.

Phospholipid vesicles, which are called "vesicles" or "liposomes", are composed of self-assembled lipid bilayer membranes. Liposomes have been widely used in the fields of molecular biology and medicine as drug carriers (Lasic 1993; Kamps et al. 2003). As described in the other chapters of this book, giant vesicles (GVs) with diameters of up to tens of micrometers have received considerable attention (Luisi 2000). They can be directly observed by the use of phase-contrast or fluorescent (including confocal) microscopy. Their internal chemical reactions, as well as just their shape, have been observed, due to the remarkable progress in fluorescence molecular observation technology.

By using such GVs, we have reported the following research topics with the aim of constructing artificial cell models.

10.2   Construction of the Structure and Function of a Model Cell

Figure 10.1a shows the successful entrapment of T4 phage in a giant vesicle (Nomura and Yoshikawa 2000). Conventional liposomes with a size of several hundreds of Angstroms are too small to incorporate such giant molecules. By adapting the procedure for the natural swelling of phospholipid film with an aqueous solution containing giant DNAs, we obtained GVs that entrapped giant DNAs (Fig. 10.1a). Entrapped DNA molecules were directly and accurately observed at the level of single molecules by fluorescence microscopy. Interestingly, the proportion of trapped giant DNA was greater than that in the external solution.

Sato et al. measured the efficiency of the entrapment of giant genetic DNA molecules in GVs obtained by the spontaneous natural swelling method. They counted
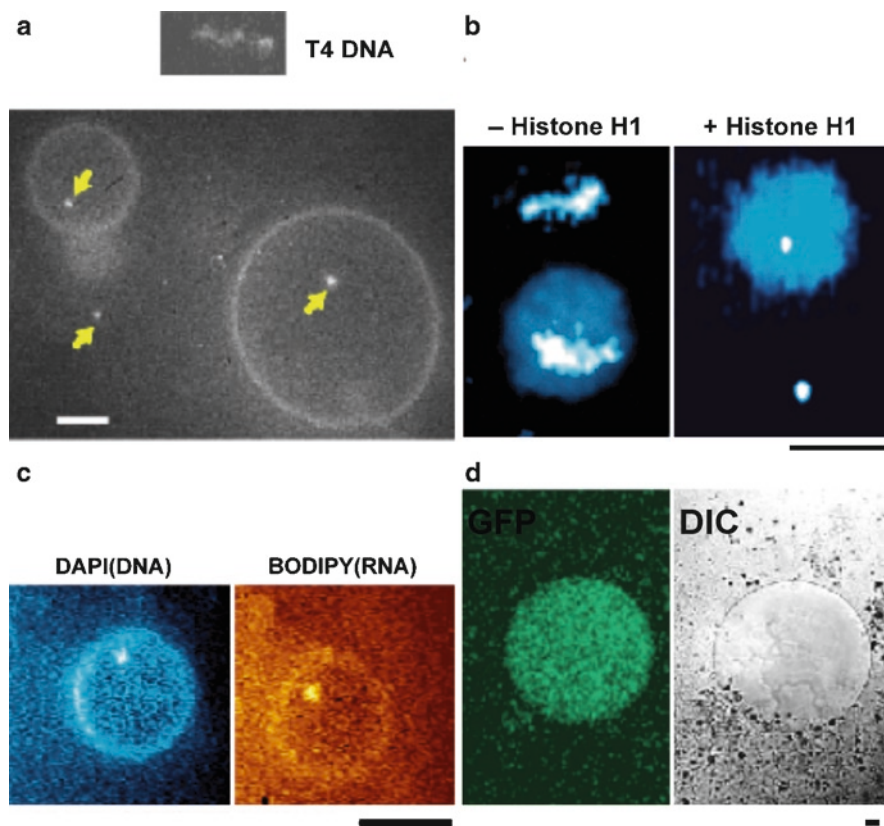
Fig. 10.1 Basic models of an artificial cell. (**a**) Entrapment of T4 phage DNA in a giant vesicle. Arrows show collapsed single long DNA molecules. (**b**) Giant vesicle made from a geranylgeranylphosphate membrane entrapping T4 DNA with or without Histone H1 protein. In the presence of the Histone, a long DNA molecule is folded into a collapsed state. (**c**) Gene transcription reaction observed in giant vesicles. In the presence of T7 RNA polymerase with fluorescence-labeled UTP, the internal RNA signals of individual GVs were observed (*right*, internal bright spot). (**d**) Coupled transcription/translation reaction in a giant vesicle. Green fluorescence proteins were observed by the use of laser scanning microscopy. Bar = 5 μm (These are Originally published in Nomura and Yoshikawa 2000 (**a**), Nomura et al. 2001 (**b**) and Tsumoto et al. 2001 (**c**))

the number density of individual DNA molecules inside and outside GVs. In the presence of Mg^{2+} as a counter cation, the number density of entrapped DNA molecules was up to tenfold greater than that in the outside solution (Sato et al. 2003). This enhancement of DNA uptake into GVs was discussed in terms of the interaction of a divalent ion and phospholipid molecules.

Next, we tried to incorporate DNA with binding proteins into GVs. We used the membrane molecule geranylgeranylphosphate (Nomura et al. 2001). This molecule

is a primitive lipid that can be synthesized by a prebiotic reaction (Fig. 10.1b). In the presence of Histone H1, giant DNA molecules assumed a globular conformation. The lipid could provide a very flexible membrane under physiological conditions for entrapping entire DNA molecules and proteins. In addition, with the use of optical tweezers, folded DNA molecules can be forced from the external solution into a GV. Such a non-invasive technique for introducing specific molecules into a GV should be useful for controlling or analyzing artificial cell models, and particularly for setting the time-zero of their internal reactions.

This system can be extended with an internal enzymatic reaction. Tsumoto et al. used a gene transcription reaction for encapsulation in GVs (Tsumoto et al. 2001). In the presence of T7 RNA polymerase with fluorescence-labeled UTP, the internal RNA signal of individual GVs was observed (Fig. 10.1c). They also confirmed that a cell-sized vesicles acts as a barrier that can prevent the attack of RNase in the bathing solution surrounding it.

We should be able to insert all of the processes that are central to biology into GVs.

The insertion of in vitro gene expression in a liposome can be regarded as a bootstrap sequence for making a cell model. Yu et al. observed GFP synthesis in a liposome solution (Yu et al. 2001). Nomura et al. also reported that functional protein synthesis was observed in cell-sized lipid vesicles following encapsulation of a gene-expression system (Nomura et al. 2003). Interestingly, at the early stage of the reaction, the expression efficiency inside the vesicle was remarkably higher than that in the outside solution. This fact supports the hypothesis that the internal space of a cell-sized liposome is a suitable environment for protein-synthesis reactions. We sometimes refer to such a system as in lipo protein synthesis.

Recently, Yamaji et al. reported another GFP-expressing model. An in vitro translation system has been developed from cell components of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. The availability of its genome sequence, along with the development of gene manipulation techniques, has made T. kodakaraensis an attractive model organism for studying various biological mechanisms in the hyperthermophilic archaea. Giant vesicles were prepared from water-in-oil cell-sized emulsions that incorporated the translation system/ GFP mRNA mixture. After incubation at 40°C for 90 min, a significant increase in average fluorescence intensity was observed in giant vesicles with GFP mRNA, but not in those without mRNA (Yamaji et al. 2009). The *in lipo* translation system of *T. kodakaraensis* allows the synthesis of proteins at high temperatures (optimal 40–45°C) and should be useful for the preparation of artificial cellular models.

However, the compartmented space of a lipid bilayer membrane is strictly isolated from the external environment. All living organisms on Earth use membrane proteins for communication with their environment. Thus, functional membrane protein synthesis with liposomes is a crucial step toward a self-maintaining model. Based on the above research, we started to construct new models for membrane protein synthesis.

Despite the rapid advancement with in vitro gene expression systems, most of the research has been devoted to the synthesis of water-soluble proteins (Shimizu et al. 2006; Vinarov et al. 2006; Nakano et al. 2004; Endo and Sawasaki 2003). The synthesis of membrane proteins by a cell-free system is still limited (Katzen et al. 2009; Kuruma et al. 2009; Kalmbach et al. 2007; Noireaux and Libchaber 2004; Nomura et al. 2008; Kaneda et al. 2009; Robelek et al. 2007).

Therefore, one of the present authors (Nomura) and co-workers reported the construction of proteo-liposomes by an in situ-synthesized membrane protein within a liposome (Nomura et al. 2008; Kaneda et al. 2009).

Apo cytochrome b5 (b5) and its fusion proteins were synthesized and directly localized onto the liposome membrane. Giant liposomes were present in a translation reaction cocktail of a wheat germ cell-free protein translation system. A single transmembrane protein b5 was directly displayed on the liposome surface (Nomura et al. 2008). This method also uses a hydrophobic tag for the immobilization of specific hydrophilic molecular parts on the liposomal surface.

Connexin 43 (Cx43)-containing liposomes were prepared by using cell-free transcription/translation systems with plasmids encoding Cx43 (Fig. 10.2a). The liposomes were spontaneously formed by use of the natural swelling method. The expressed membrane protein, Cx43, was directly constituted to the liposome membrane upon in vitro synthesis, which gave pure membrane protein-containing liposomes (Fig. 10.2b). Connexin is a membrane protein that makes a gap-junction (GJ) between neighboring mammalian cells. GJ is composed of a pair of six docked connexin proteins from two cells. Small molecules such as ions, second messengers, and metabolites in cytosol can be directly transferred from one cell to a neighboring cell through GJs. The driving force for this transfer is considered to be simple diffusion. The molecular mass cut-off for a GJ channel is generally considered to be below 1–1.8 kDa.

Figure 10.2c shows the function of the Cx43-proteoliposome. The hydrophilic dye calcein (633 Da) was efficiently transferred from Cx43-expressing liposomes to cultured cells (Cx43-expressing). This transfer is significantly blocked in the presence of a gap junction inhibitor (18β-glycyrrhetinic acid) and in the presence of cells expressing another type of connexin (Cx32). The results show that calcein entered the cell through a connexin-mediated pathway. Cx43 liposomes that contained a soluble NEMO binding domain peptide suppressed the intracellular signaling cascade IL-1β-induced NF-kB activation and cyclooxygenase-2 expression in Cx43-expressing cells, which confirmed effective peptide transfer into the cell (Kaneda et al. 2009).

The studies described here raise several problems that should be solved, such as the nature of the membrane protein's insertion, a qualitative analysis of the efficiency of gene expression, the qualitative production of giant "unilamellar" liposomes with a gene expression system, etc. However, they do seem to offer the potential for artificial cell models that can communicate with real cells through the use of their compatible protocols.
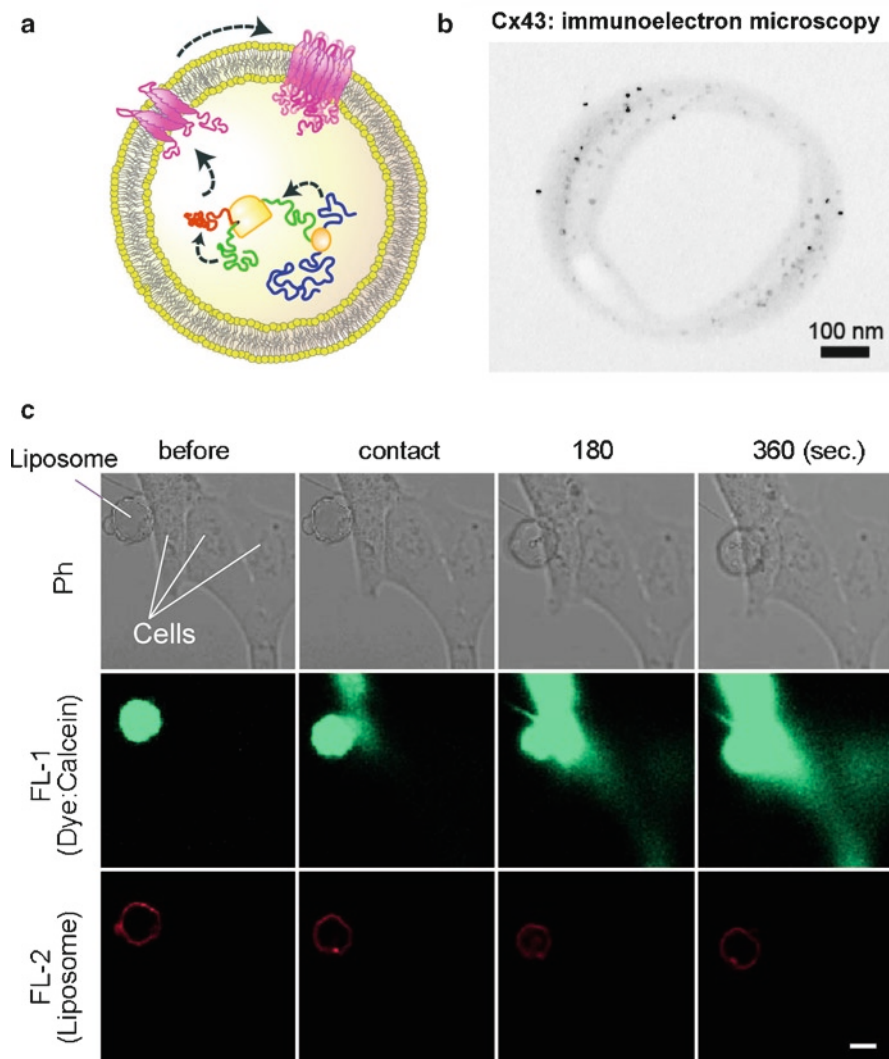
Fig. 10.2 Membrane protein expression with liposomes. (**a**) Schematic illustration. (**b**) Immunoelectron microscopic image. Liposomes were fixed (*gray circle*) and treated by anti-connexin with immunogold (*black dots*). (**c**) Time-course images of the transfer of calcein dye from Cx43-expressing GVs to cells. This small dye diffused slowly into the cellular internal space after the liposome had attached to U2OS cells (Cx43-expressing). The bottom images show the results with the GV membrane dye TexasRed-DHPE. Bar = 10 μm (This is Originally published in Kaneda et al. 2009)

10.3   Recombinant Proteoliposomes

Proteoliposomes are defined as liposomes with membranes that possess membrane proteins with biological functions. Since membrane proteins are usually related to signal transduction, cell response, and cell communication at cellular interfaces, they are useful for expanding the functionality of a model cell system. Conventionally, proteoliposomes are most often prepared using a detergent method (Rigaud and Lévy 2003), according to the following experimental procedure: First, the membrane protein of interest is obtained from a biological resource, which is, in general, a specific tissue with an abundance of the protein, or transfected cultured cells that express the protein from a recombinant gene. The membrane protein is dissolved with detergent and mixed with similarly dissolved phospholipid. The mixture of dissolved membrane protein and lipid is then dialyzed against a buffer, which leads to the spontaneous organization of liposomes with the membrane protein reconstituted on their membrane, i.e., proteoliposomes. The resulting proteoliposome usually measures on the order of ~100 nm, which corresponds to the size of large unilamellar vesicles (LUVs). Since this conventional method has good controllability and reproducibility, it is widely used in investigations of membrane proteins. Therefore, to obtain giant proteoliposomes (or proteo-GVs), researchers often start with this method. For example, Kahya et al. (2001) demonstrated a novel method for the preparation of proteo-GVs based on fusion between GVs and conventional proteo-liposomes that are both prepared in advance. They reconstituted bacteriorhodopsins (BRs) in small liposomes (LUVs) conjugated with the fusogenic peptide WAE, and converted GVs into proteo-GVs by inducing liposome fusion, which results in the diffusive migration of BRs throughout the GV membranes. Girard et al. (2004) presented another method based on conventional proteoliposomes. They modified the electroformation method that is frequently used to prepare giant unilamellar vesicles (GVs) (Angelova 2000). In that method, small proteoliposomes containing reconstituted membrane proteins (BRs or Ca^{2+}-ATPase) are dehydrated on an electrode to form phospholipid lamellar films. These films are then rehydrated using electroformation to generate proteo-GVs. While both methods have been popular recently, they have serious problems inherent to the method for preparing conventional proteoliposomes. Thus, when the desired membrane protein is not expressed in abundance, it may be prone to degradation (probably through the dissociation of multiple subunits) in the presence of detergent.

On the other hand, as in the above-mentioned method based on the direct insertion of membrane proteins expressed in cell-free systems, a few strategies have been developed in which the functions of membrane proteins are bestowed on GVs instead of via the preparation of conventional proteoliposomes. Recently, Yoshimura and colleagues presented a novel method based on a baculovirus gene expression system for preparing proteoliposomes without the use of detergent (Fukushima et al. 2008; Fukushima et al. 2009; Tsumoto and Yoshimura 2009). This baculovirus system consists of *Autographa californica* nucleopolyhedrovirus (AcNPV) and cultured insect cells (*Spodoptera frugiperda* [Sf9] cells). It is a popular tool for producing

recombinant proteins in eukaryotes and is commercially available from well-known suppliers (Kost et al. 2005). AcNPV is an enveloped virus that has two alternative viral forms in its life cycle: an occluded virion (OV) and a budded virion (BV). In OVs, enveloped virus nucleocapsids are embedded in a large amount of polyhedrin proteins that are strongly expressed in the late stage of the viral infection. However, the polyhedrin protein is not a requisite for AcNPV propagation in cultured cells, and thus, once a recombinant cDNA of the desired protein is constructed under the polyhedrin promoter sequence on a transfer vector and transferred into insect cells with an AcNPV genome DNA, a recombinant baculovirus budded virion with the gene is generated. Sf9 cells infected with the recombinant BVs strongly express the desired protein and reproduce BVs, which bud from Sf9 cell membranes. Interestingly, when a recombinant gene is a transmembrane protein or a membrane-anchored protein, reproduced BVs also have the protein on their own envelopes with an intact structure and function. This means that if BVs are collected and purified from the supernatant of Sf9 cell culture medium using a common method such as ultracentrifugation, the desired membrane protein can be easily collected in a fraction of the BVs. With this procedure, some pioneering researchers have prepared recombinant BVs with membrane proteins, such as a β_2-adrenergic receptor (Loisel et al. 1997), leukotriene B_4 receptor (Masuda et al. 2003), γ-secretase (Hayashi et al. 2004), GPI-anchored acetylcholine esterase, etc. Since AcNPV BVs enter cells through absorptive endocytosis during infection, they naturally have the fusogenic envelope glycoprotein GP64, which is activated under acidic conditions (pH 4.0–5.5) similar to those in endosomes. By taking advantage of this characteristic of baculovirus, Yoshimura et al. designed a protocol to prepare proteoliposomes through membrane fusion between recombinant BVs and plain conventional liposomes, LUVs and multilamellar vesicles (MLVs). Although it is not yet clear what membrane factor serves as a receptor for GP64, it has been shown that an acidic lipid is required for GP64 entry. In fact, by using BVs labeled with octadecylrhodamine B (R18), they showed that a steep increase in R18 fluorescence arose in the region of pH 4.0–5.0 when the BVs were added to liposomes consisting of phosphatidylcholine (PC) with phosphatidylserine (PS), phosphatidylglycerol (PG), phosphatidic acid (PA) or phosphatidylinositol (PI), indicating that exposure to an acidic condition causes BVs to fuse with liposome membranes. They constructed recombinant AcNPV BVs with the transmembrane protein thyroid stimulating hormone receptor TSHR, which is a member of the G protein coupled receptor (GPCR) family, and nicotinic acetylcholine receptor α subunit (AChRα), which is a component of a pentameric ligand-gated ion channel. By collecting liposomes fused with recombinant BV envelopes, they showed that a portion of the membrane protein appeared in a collected fraction of liposomes, and this depended on the pH at which the BVs were treated. Furthermore, biotinylated liposomes fused with the BVs were immobilized on enzyme-linked immunosorbent assay (ELISA) plates through biotin-streptavidin interaction, and the liposomes then specifically reacted with the anti-TSHR or the anti-AChRα antibody. This result suggests that the membrane proteins migrate by diffusion over liposome membranes upon BV-liposome fusion, and then proteoliposomes are prepared: they refer to these as "recombinant proteoliposomes." In addition, recombinant
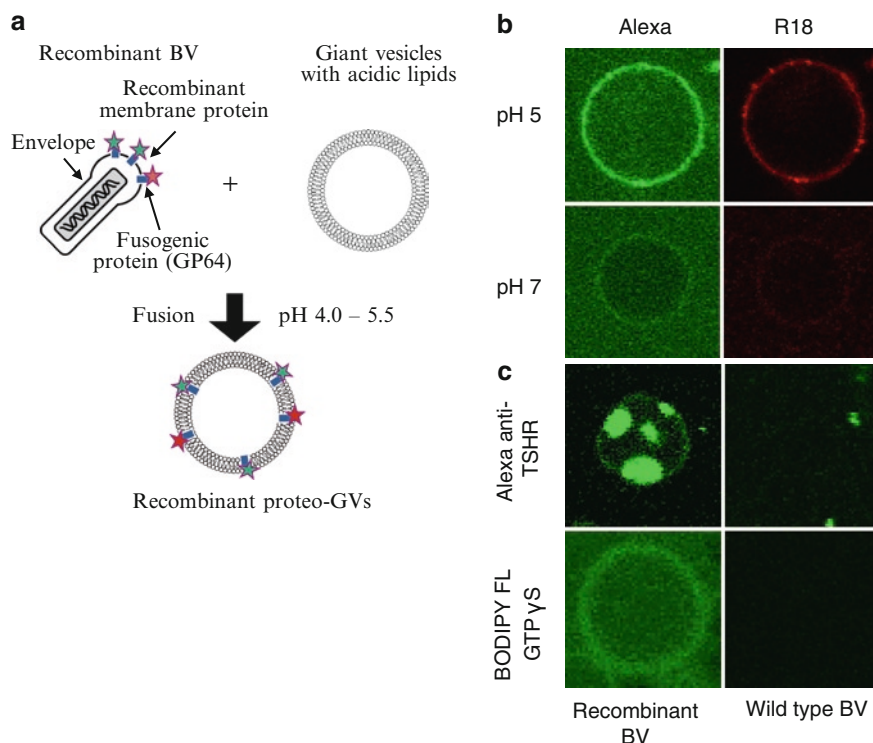
Fig. 10.3 Preparation of recombinant proteo-GVs using the baculovirus liposome fusion method. (**a**) A schematic illustration depicting the strategy. (**b**) A typical fluorescence image of a single GV fused with BVs labeled with Alexa 488 carboxylic acid, succinimidyl ester for proteins and R18 for lipids. Fusion between GVs and BVs is induced by a decrease in the pH of the solution. Both fluorescent dyes diffused over the GV homogeneously in this case. (**c**) Visualization of TSHR (*top*) and Gαs (*bottom*) on proteo-GVs. Giant vesicles were fused with the indicated BVs, and then treated with the reagents

proteoliposomes can react with their own ligands such as thyroid stimulating hormone (TSH) and α-bungarotoxin (a specific toxin for AChR). It is not clear whether any membrane protein can be expressed on BVs once its cDNA is incorporated into a baculovirus vector. However, this strategy may represent a tool with which we can reconstitute complex membrane protein systems on liposomes at will, since we consider a recombinant AcNPV BV as a fusing unit that possesses the desired protein for loading step by step. This method can also be applied to the production of proteo-GVs by analogy to Khaya's method. As shown in Fig. 10.3, BVs and GVs consisting of PC and PS are mixed at a pH below 5.5 (4.0–5.5), and they fuse with each other. The fusion properties seem somewhat different between small liposomes and GVs; for example, PA exhibits strong fusogenic activity in GVs, and the optimal pH for inducing fusion is different. Along with fusion, the migration of recombinant proteins over GV membranes is observed. We tried to reconstitute TSHR on recombinant proteo-GVs as the first step in constructing a multiple membrane protein

system. As shown in the fluorescence microscopic image in Fig. 10.3, Alexa Fluor 488-conjugated anti-TSHR antibody can be used to visualize TSHR on proteo-GVs. Fluorescent patches were frequently observed on proteo-GVs, and this might imply the inhomogeneous mixing of protein components from BVs with GV membranes, but the actual reason remains to be clarified. Furthermore, according to the procedure, G proteins, which are anchored to the inner cell membranes through isoprenyl/acyl chains, can also be reconstituted on proteoliposomes. Stimulatory G protein α subunits were visualized using BODIPY FL GTPγS. Currently, we are trying to reconstitute a multiple recombinant protein system on single proteo-GVs with functional signal transduction.

10.4 Efficient Construction of Giant Vesicles

As the number of studies using GVs, or especially GUVs, increases, novel preparation methods are occasionally presented. However, methods based on lipid hydration are most often used. In general, the electroformation method and natural swelling (gentle hydration) can be classified into this category (Rodriguez et al. 2005); the former is most popular because of its high efficiency and reproducibility, whereas the latter is also desirable because of the simplicity of the preparation procedure. While GUVs may be produced and examined in a physiological salt solution, the number, size distribution and quality of GUVs are basically deteriorated. This trend becomes particularly evident when GUVs are made up of a neutral phospholipid such as PC, which is the chief component of biological membranes. It may be possible to avoid this problem by characterizing the conditions for and the mechanism of the hydration of lipid films. In the case of electroformation, the details with respect to the species of buffer and phospholipid have been elucidated (Fischer et al. 2000; Shimanouchi et al. 2009). Here, we use the natural swelling method to realize the efficient formation of PC-rich GUVs in saline. Interlayer repulsion is believed to be necessary to generate GUVs, but this is weakened in the presence of salt. Therefore, repulsive interaction can be reinforced by adding a small amount of (1) acidic phospholipid (Akashi et al. 1996), (2) divalent alkaline earth metal ions such as Ca^{2+} and Mg^{2+}, which bind to a head group of PC with some affinity (Magome et al. 1997; Akashi et al. 1998), or (3) PEGylated phospholipid, resulting in steric hindrance due to a highly hydrated hydrophilic polymer (Yamashita et al. 2002). In addition, it has been suggested that osmotic pressure within interlayer spaces plays a crucial role during the hydration and swelling of phospholipid films (Magome et al. 1997). If we consider this suggestion, it is easy to design a novel protocol in which dry phospholipid films sandwich osmolytes (sugar) in the interlayer spaces, as depicted in Fig. 10.4 (Tsumoto et al. 2009). In this method, methanol-soluble hexose such as glucose, mannose, fructose, etc. and dioleoylphosphatidylcholine (DOPC) are solved in a mixed organic solvent (chloroform/methanol), and the mixture is then evaporated to form lipid films. These DOPC films can be hydrated very efficiently and produce very giant unilamellar
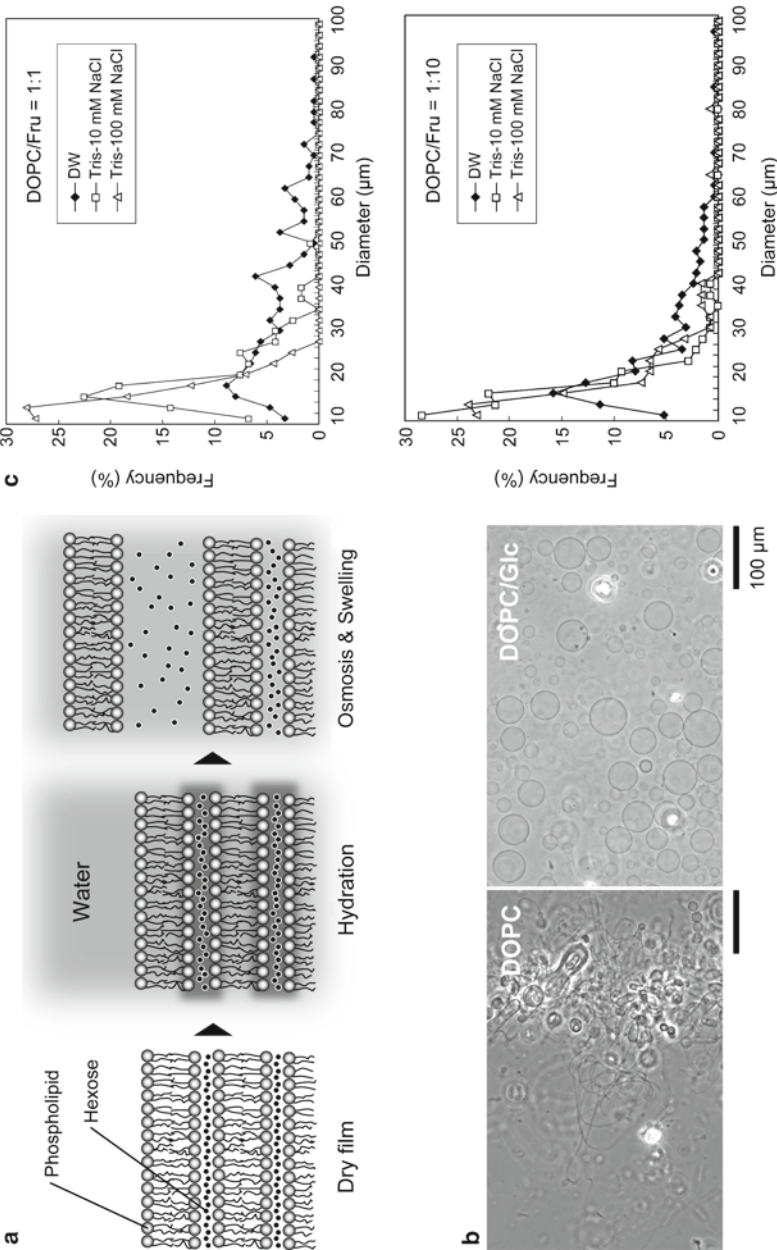
Fig. 10.4 (**a**) Schematic representation of the hydration and swelling of sugar-sandwiched phospholipid films. (**b**) Microscopic images of typical giant vesicles generated from DOPC films without (*left*) and with (*right*) sugar (glucose). (**c**) Size distribution of GUVs generated from DOPC films with fructose in the absence/presence of salt (This is Originally published in Tsumoto et al. 2009)

vesicles even if the hydration buffer contains up to ~100 mM salt; generally, giant spherical and unilamellar vesicles rarely appear under such conditions (Fig. 10.4). The properties (size distribution and population) of prepared GUVs depend on both the sugar and salt concentrations. GUVs formed in the absence of salt have a very wide size distribution, ranging from ~10–100 μm in diameter, whereas the distribution becomes sharper and a peak appears at around ~20 μm. GUV formation is inhibited by high salt concentrations, but this trend can be reduced by increasing the sugar content. This implies that for GUV formation, a certain sugar content is required to give a high osmotic pressure upon hydration so that swelling proceeds against the exterior hypotonic solution. The existence of a threshold can be explained by an unbinding transition (Yamada et al. 2007), in which osmotic pressure over a threshold value causes sufficient repulsion to make the interlayer distance infinite. However, excess sugar (the molar ratio of DOPC/hexose = 1:>100) inhibits GUV formation. If we consider that the efficiency of GUV formation is strongly correlated with the geometry of the initial lipid film (Hishida et al. 2005), this is probably due to the fact that the lamellar stacking structure of a dry lipid film is destroyed and, thus, it cannot induce effective osmosis during hydration. As shown here, the natural swelling method can be improved by modifying the hydration conditions, and this may be revisited by researchers who understand the scientific benefit of using GVs.
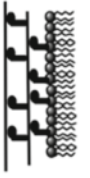
10.5   Construction of a Model for Studying Changes in Cell Morphology

Living cells and their organelles are compartmentalized by biomembranes and each has a specific morphology depending on its role and function. Their morphologies are believed to be determined and maintained by cytoskeletal networks (Hotani et al. 2003; Rodriguez et al. 2003). Actin filament (F-actin) is a major component of the cytoskeleton, and is involved in a variety of cellular functions, including the extension of microspikes from cells (Albrecht-Buehler and Lancaster 1976), the movement of filopodia in neural growth cones (Mitchison and Cramer 1996; Mitchison and Kirschner 1988), the extension or retraction of pseudopods during amoeboid movement (Taylor and Condeelis 1979), and the contraction of contractile rings during cell division (Schroeder 1973). F-actins also provide mechanical support, including stress fibers (Byers and Fujikawa 1982). In most situations, actin works together with myosins, which are molecular motors. Therefore, it is important to study the role of the actin cytoskeleton and myosin motors in relation to membrane morphogenesis and cell movement.

Several studies have investigated living cells and found many potential key proteins or genes and core interactions between these factors. To date, based solely on the results of the current analytical methodology which uses a "top-down" approach, it is impossible to elucidate the principles that underlie the life-events in living cells because cells are typically very complex systems. To obtain a better

understanding of the essential mechanism, an artificial cell model that consists of cytoskeletal proteins (actin and myosins) and giant liposomes has been developed using methods such as natural swelling, electroformation, and spontaneous transfer, as a consequence of a "bottom-up" approach (Bormann et al. 1992; Cortese et al. 1989; Fygenson et al. 1997; Häckl et al. 1998; Limozin et al. 2003, 2005; Limozin and Sackmann 2002; Miyata and Hotani 1992; Miyata et al. 1999; Takiguchi 1991; Takiguchi et al. 2008). Giant liposomes, which consist of closed lipid bilayer membranes in aqueous solution with sizes equal to those of living cells, are expected to serve as a useful model in the real world, and should allow us to make real-time observations using optical microscopy (Bangham 1995; Cortese et al. 1989; Hotani 1984; Hotani et al. 2003; Lasic 1995; Lipowsky 1991; Miyata et al. 1999).

Over 20 classes of different types of myosins have been identified to date. Myosins can be classified into two groups according to their head-structure, i.e. double- or single-headed myosins. Myosin of the double-headed type, such as myosin-II, and a double-headed derivative of myosin-II, such as heavy meromyosin (HMM), which has been frequently studied as a representative double-headed myosin (Fig. 10.5), can transform actin bundles or actin gels (Takiguchi 1991; Tanaka-Takiguchi et al. 2004). On the other hand, subfragment 1 (S-1) has often been studied as a representative simple single-headed myosin. S-1 has only one actin-binding motor domain and can not crosslink F-actins (Fig. 10.5). In living cells, F-actin is frequently anchored to membranes through mediating proteins, which

Fig. 10.5 Model showing the arrangement of F-actins in the presence of each myosin (from *left* to *right*, HMM, S-1, or BBMI). The upper row illustrates the molecular structure of each protein. The lower row illustrates the arrangement of F-actins. F-actin is depicted as a line, while HMM, S-1, and BBMI are drawn as double-headed V-shaped dimers, single-headed structures, and single-headed musical note-like structures, respectively. HMM consists of two heavy chains possessing an actin-sliding motor domain (head), an accompanying coiled-coil region that is responsible for dimerization, and several light chains (not shown here). Therefore, HMM can cross-link F-actins into bundles in the absence of ATP. On the other hand, S-1 consists of only a single heavy chain possessing a single motor domain, so that it can not cross-link F-actins. BBMI consists of a single motor domain and a tail domain, and since the tail domain can bind both F-actin and a lipid membrane, BBMI can crosslink F-actin and the membrane, as well as different F-actins

probably regulate a specific membranous morphology. Myosin-I isolated from brush border (BBMI) exhibits a unique property compared to conventional myosin and its derivatives, HMM and S-1. BBMI is a single-headed myosin. However, it can crosslink F-actin and membrane, as well as different F-actins, since its tail domain can bind both F-actin and a lipid membrane (Fig. 10.5) (Coluccio 1997).

Accordingly, we compared the effects of HMM, S-1 and BBMI on the morphogenesis of liposomes caused by actin assembly, by co-encapsulating these myosins together in liposomes.

Giant vesicles made from PC by spontaneous transfer and from PC or PE and PG by natural swelling are spherical or otherwise unstable tubules and do not show any significant transformation (Fig. 10.6a).

Recently, several attempts have been made to use water-in-oil (W/O) droplets coated by phospholipids as a precursor of GVs (Hamada et al. 2008; Hase et al. 2007; Hase and Yoshikawa 2006; Noireaux and Libchaber 2004; Pautot et al. 2003a, 2003b; Pontani et al. 2009; Takiguchi et al. 2008; Yamada et al. 2006). This spontaneous transfer method enabled us to encapsulate desired amounts of F-actin with or without myosins into GVs (Takiguchi et al. 2008). With this method, we prepared GVs containing only F-actin (at 200 μM), which had already polymerized before encapsulation. However, they showed no transformation (Fig. 10.6b), and F-actins inside GVs showed a uniform distribution (Fig. 10.6b, left). Note that 200 μM is comparable to the actin concentration in living cells (Janson et al. 1991; Pollard et al. 2000), and is the upper limit of concentration that allows handling due to its very high viscosity. On the other hand, GVs containing monomeric actin (G-actin) prepared by natural swelling transformed into a rigid stable shape as the encapsulated actin polymerized. Until the encapsulated actin polymerized, GVs containing G-actin assumed spherical or unstable tubular shapes as with empty GVs. Once the actin polymerized, the GVs assumed flat disk or spoon shapes (Fig. 10.6c) (Honda et al. 1999; Miyata and Hotani 1992). F-actin that had polymerized in GVs spontaneously aligned along the periphery of the flattened GVs (Fig. 10.6c, left). Consistent with this finding, Miyata and colleagues showed that GVs encapsulating 200 μM of G-actin prepared by the natural swelling method grew protrusive structures as the actin polymerized (Miyata et al. 1999).

Figure 10.6d shows GVs that have entrapped F-actin together with HMM. In contrast to the images with only F-actin (Fig. 10.6b), in the presence of HMM, F-actin assemblies, such as bundles and networks, are generated (Fig. 10.6d, left). Deformed GVs tended to appear reproducibly as the HMM concentration increased, although a large fraction of GVs remained spherical.

On the other hand, when S-1 rather than HMM was co-encapsulated with F-actin into giant vesicles, the GVs maintained their spherical shapes and showed a uniform distribution of F-actins, even under conditions in which excess S-1 was co-encapsulated (Fig. 10.6e).

We have prepared GVs that simultaneously contained G-actin and BBMI, a membrane-associable myosin, through the method of natural swelling. After the actin polymerized, the GVs effectively transformed into protruded shapes even at a lower concentration of actin compared to the cases with HMM or other well-known

actin-crosslinking proteins (Fig. 10.6f) (Honda et al. 1999). In the presence of BBMI, the actin bundles formed are promptly recruited and bound to the inner surface of the liposome membrane, resulting in efficient GV transformation (Fig. 10.6f, left).

In this section, by studying the morphogenesis of GVs, we have demonstrated that the mechanical force generated by the assembly of actin network structures itself is strong enough to transform cell-sized giant vesicles.

G-actin-encapsulating liposomes grow protrusive structures as the actin polymerizes. On the other hand, liposomes that encapsulate only F-actin that has already polymerized show no transformation. Together, these findings suggest that the assembly of G-actins at the end of membrane-pushing F-actin, i.e., the mechanical force generated during actin polymerization and elongation, is required for liposomal transformation (Hotani et al. 2003; Maemichi et al. 2008; Miyata et al. 1999).

HMM and BBMI can crosslink F-actins into actin bundles or actin gels and thereby transform these bundles or gels (Tanaka-Takiguchi et al. 2004, and Honda et al. 1999, unpublished data), while S-1 can not crosslink F-actins (Fig. 10.5). Liposomes that have co-encapsulated HMM but not S-1 show morphological changes (Fig. 10.6d–f), indicating that the simple bundling or redistribution of F-actin, i.e., inter-filament physical interactions, can be a motive force for organizing actin networks and liposomal morphogenesis, as well as the polymerization of G-actin and the subsequent bundling of F-actin. Moreover, BBMI can connect actin bundles to a lipid membrane (Fig. 10.5). GVs encapsulating BBMI assumed protruded shapes even when the concentration of encapsulated actin was comparably low. Since the actin bundles formed with BBMI were efficiently bound to the liposome membrane by BBMI, the force generated by actin assembly may have been directly transferred to the membrane. These results indicate that manipulation of the membrane morphology may be possible through the regulation of local interactions between actomyosin systems and membranes.

This simple and very efficient system which uses giant vesicles containing both F-actin and myosins might represent the first critical step for developing a motile artificial cell model, i.e. for achieving the spontaneous motion of giant vesicles containing active cytoskeletal networks.

An important problem that remains is how to supply sufficient amounts of ATP, the fuel that living cells use for, among other things, their motility, or regulatory factors to an actomyosin system within GVs. The use of proteins that form a membrane pore is the likeliest tool for passing such molecules across lipid bilayer membranes. Interestingly, proteins in the band 4.1 superfamily can open large stable holes in lipid bilayers (Saitoh et al. 1998; Takeda et al. 2006). Membrane holes opened by these proteins are stable and the diameters of the holes (up to several micrometers) depend on the protein concentration. Recently, α-hemolysin, a toxin from bacteria, has been shown to be useful for molecular transport into vesicles (Gibrat et al. 2008; Noireaux and Libchaber 2004; Pontani et al. 2009). Soluble α-hemolysin penetrates into membranes and spontaneously forms channels (about 3 nm in diameter). This phenomenon may be useful for supplying ATP to actomyosin-containing giant vesicles.

No Actin (Control)

a

+ Actin

b

+ Actin

c

+ Actin & HMM

d

+ Actin & S-1

e

+ Actin & BBMI

f
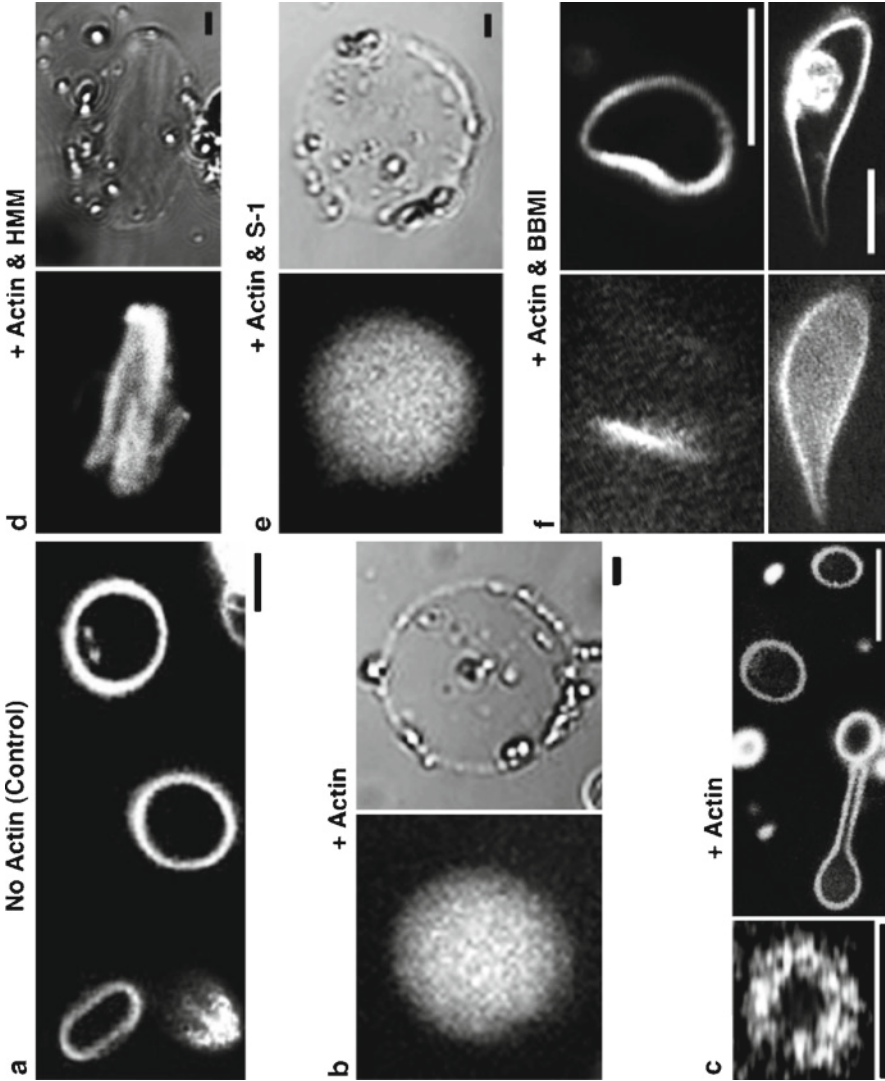
Fig. 10.6 Morphological changes in actin-containing giant vesicles with or without each myosin. (**a**) Dark-field microscopic image of empty GVs (i.e., containing neither actin nor myosin). The lipid composition was PC and PG (1:1 mol:mol). (**b**) Morphology of GVs that encapsulated F-actin that had already polymerized formed by the spontaneous transfer method (left: fluorescence, right: transmission, confocal microscopic images). The concentration of F-actin inside the GVs is 50 μM. The fluorescence images show the distribution of F-actin that was labeled with rhodamine-phalloidin. The lipid composition was PC. (**c**) Transformed GVs obtained by the polymerization of encapsulated G-actin into F-actin. G-actin-encapsulating GVs were formed by natural swelling at 4°C. The G-actin inside the GVs was polymerized to form F-actin by raising the temperature to 25°C. (right) Four disk-shaped and one flat spoon-shaped GVs are shown (dark-field image, viewed from the top). The lipid composition was PE and PG (1:1 mol:mol), and the concentration of encapsulated actin was 100 μM. (left) A fluorescent micrograph that shows the F-actin localization within a disk-shaped GV obtained by actin assembly (top view). The assembled F-actin was labeled with rhodamine-phalloidin. The lipid composition was PC and PG (1:1 mol:mol), and the concentration of the encapsulated G-actin was 50 μM. Note that we can distinguish a disk-shaped GV from a spherical one by shifting the focus plane of the microscope, or by the continuous observation of each free-tumbling GV in a microscope specimen. (**d**) Transformed GVs obtained by the simultaneous encapsulation of F-actin and HMM by the spontaneous transfer method (left: fluorescence, right: transmission, confocal microscopic images). The concentrations of F-actin and HMM inside the GVs are 50 and 7.5 μM, respectively. The fluorescence images show the distribution of F-actin. The lipid composition was PC. A giant vesicle with a non-spherical irregular shape is shown. (**e**) Morphology of a GV simultaneously encapsulating F-actin and S-1 formed by the spontaneous transfer method (left: fluorescence, right: transmission, confocal microscopic images). The concentrations of F-actin and S-1 inside the GV are 50 and 60 μM, respectively. The fluorescence images show the distribution of F-actin. The lipid composition was PC. (**f**) Transformed GVs obtained by the simultaneous encapsulation of G-actin and BBMI by the natural swelling method (left: fluorescence, right: dark-field). 10 (upper) or 25 μM (bottom) G-actin and 5 μM BBMI were encapsulated at 4°C and the G-actin inside the GVs was then polymerized to form F-actin by raising the temperature to 25°C. The fluorescence images show the distribution of F-actin. The lipid composition was PC and PG (1:1 mol:mol). The shape of each transformed GV did not change (**c**, **d** and **f**). Transmission images show the existence of small oil droplets around the GVs that were formed by spontaneous transfer, which were squeezed out from the oil phase (**b**, d and **f**). Bar = 5 μm (These are Originally published in Honda et al. 1999 (**c**) and Takiguchi et al. 2008 (**b**, **d** and **e**))

10.6 Conclusion

In this chapter, we have described some possible approaches to constructing cell models based on giant vesicles. If we consider this problem from the perspective of the diversity of cellular functions, it is very easy to understand that there is no single method for constructing a cell model. This modeling research needs to be integrated with some essential research area, such as synthetic biology, physical chemistry, theoretical biology, biomaterials, micro-nano engineering etc. Such cross-disciplinary studies should elevate this subject to a major research area.

References

Akashi K, Miyata H, Itoh H, Kinosita K Jr (1996) Preparation of giant liposomes in physiological conditions and their characterization under an optical microscope. Biophys J 71:3242–3250

Akashi K, Miyata H, Itoh H, Kinosita K Jr (1998) Formation of giant liposomes promoted by divalent cations: critical role of electrostatic repulsion. Biophys J 74:2973–2982

Albrecht-Buehler G, Lancaster RM (1976) A quantitative description of the extension and retraction of surface protrusions in spreading 3T3 mouse fibroblasts. J Cell Biol 71:370–382

Angelova MI (2000) Liposome electroformation. In: Luisi PL, Walde P (eds) Giant vesicles. Wiley, New York

Bangham AD (1995) Surrogate cells or Trojan horses. The discovery of liposomes. Bio Essays 17:1081–1088

Bormann M, Kos J, Kurzmeier H et al (1992) In: Lipowsky R, Richter D, Kremer K (eds) The structure and conformation of amphiphilic membranes. Springer, Heidelberg

Byers HR, Fujikawa K (1982) Stress fibers in cells in situ: immunofluorescence visualization with antiactin, antimyosin, and anti-a-actinin. J Cell Biol 93:804–811

Coluccio LM (1997) Myosin I. Am J Physiol 273:C347–C359

Cortese JD, Schwab B III, Frieden C (1989) Actin polymerization induces a shape change in actin-containing vesicles. Proc Natl Acad Sci USA 86:5773–5777

Endo Y, Sawasaki T (2003) High-throughput, genome-scale protein production method based on the wheat germ cell-free expression system. Biotechnol Adv 21(8):695–713

Fischer A, Luisi PL, Oberholzer T, Walde P (2000) Formation of giant vesicles from different kinds of lipids using the electroformation method. In: Luisi PL, Walde P (eds) Giant vesicles. Wiley, New York

Fukushima H, Mizutani M, Imamura K, Morino K, Kobayashi J, Okumura K, Tsumoto K, Yoshimura T (2008) Development of a novel preparation method of recombinant proteoliposomes using baculovirus gene expression systems. J Biochem 144:763–770

Fukushima H, Matsuo H, Imamura K, Morino K, Okumura K, Tsumoto K, Yoshimura T (2009) Diagnosis and discrimination of autoimmune Graves' disease and Hashimoto's disease using thyroid-stimulating hormone receptor-containing recombinant proteoliposomes. J Biosci Bioeng 108:551–556

Fygenson DK, Marko JF, Libchaber A (1997) Mechanics of microtubule-based membrane extension. Phys Rev Lett 79:4497–4500

Gibrat G, Pastoriza-Gallego M, Thiebot B et al (2008) Polyelectrolyte entry and transport through an asymmetric a-hemolysin channel. J Phys Chem B 112:14687–14691

Girard P, Pécréaux J, Lenoir G, Falson P, Rigaud JL, Bassereau P (2004) A new method for the reconstitution of membrane proteins into giant unilamellar vesicles. Biophys J 87:419–429

Häckl W, Bärmann M, Sackmann E (1998) Shape changes of self-assembled actin bilayer composite membranes. Phys Rev Lett 80:1786–1789

Hayashi I, Urano Y, Fukuda R, Isoo N, Kodama T, Hamakubo T, Tomita T, Iwatsubo T (2004) Selective reconstitution and recovery of functional γ-secretase complex on budded baculovirus particles. J Biol Chem 279:38040–38046

Hishida M, Seto H, Yoshikawa K (2005) Smooth/rough layering in liquid-crystalline/gel state of dry phospholipid film, in relation to its ability to generate giant vesicles. Chem Phys Lett 411:267–272

Hamada T, Miura Y, Komatsu Y et al (2008) Construction of asymmetric cell-sized lipid vesicles from lipid-coated water-in-oil microdroplets. J Phys Chem B 112:14678–14681

Hase M, Yamada A, Hamada T et al (2007) Manipulation of cell-sized phospholipid-coated microdroplets and their use as biochemical microreactors. Langmuir 23:348–352

Hase M, Yoshikawa K (2006) Structural transition of actin filament in a cell-sized water droplet with a phospholipid membrane. J Chem Phys 124:Art. No. 104903

Honda M, Takiguchi K, Ishikawa S et al (1999) Morphogenesis of liposomes encapsulating actin depends on the type of actin-crosslinking. J Mol Biol 287:293–300

Hotani H (1984) Transformation pathways of liposomes. J Mol Biol 178:113–120

Hotani H, Inaba T, Nomura F et al (2003) Mechanical analyses of morphological and topological transformation of liposomes. Biosystems 71:93–100

Janson LW, Kolega J, Taylor DL (1991) Modulation of contraction by gelation/solation in a reconstituted motile model. J Cell Biol 114:1005–1015

Kahya N, Pécheur EI, de Boeij WP, Wiersma DA, Hoekstra D (2001) Reconstitution of membrane proteins into giant unilamellar vesicles via peptide-induced fusion. Biophys J 81:1464–1474

Kalmbach R et al (2007) Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes. J Mol Biol 371:639–648

Kamps JAAM, Scherphof GL, Sullivan S, Gong Y, Hughes J (2003) In: Torchilin VP, Weissig V (eds) Liposomes, a practical approach, 2nd edn. Oxford University Press, New York, pp 267–301

Kaneda M, Nomura SM, Ichinose S, Kondo S, Nakahama K, Akiyoshi K, Morita I (2009) Direct formation of proteo-liposomes by in vitro synthesis and cellular cytosolic delivery with connexin-expressing liposomes. Biomaterials 30:3971–3977

Katzen F, Peterson TC, KudlickiKatzen W (2009) Membrane protein expression: no cells required. Trends Biotechnol 27(8):455–460

Kost TA, Condreay JP, Jarvis DL (2005) Baculovirus as versatile vectors for protein expression in insect and mammalian cells. Nat Biotechnol 23:567–575

Kuruma Y, Stano P, Ueda T, Luisi PL (2009) A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. Biochim Biophys Acta 1788:567–574

Lasic DD (1993) Liposomes: from Physics to applications. Elsevier, Amsterdam, pp 243–458

Lasic DD (1995) Applications of liposomes. In: Lipowsky R, Sackmann E (eds) Structure and dynamics of membranes, vol 1a, pp 491–519. Elsevier, Amsterdam

Limozin L, Bärmann M, Sackmann E (2003) On the organization of self-assembled actin networks in giant vesicles. Eur Phys J E 10:319–330

Limozin L, Roth A, Sackmann E (2005) Microviscoelastic moduli of biomimetic cell envelopes. Phys Rev Lett 95:Art. No. 178101

Limozin L, Sackmann E (2002) Polymorphism of cross-linked actin networks in giant vesicles. Phys Rev Lett 89:Art. No. 168103

Lipowsky R (1991) The conformation of membranes. Nature 349:475–481

Loisel TP, Ansanay H, St-Onge S, Gay B, Boulanger P, Strosberg AD, Marullo S, Bouvier M (1997) Recovery of homogeneous and functional β2-adrenergic receptors from extracellular baculovirus particles. Nat Biotechnol 15:1300–1304

Luisi PL (2000) Why gaiat vesicles? In: Luisi PL, Walde P (eds) Giant vesicles. Wiley, New York

Maemichi H, Shikinaka K, Kakugo A et al (2008) Morphogenesis of liposomes caused by polycation-induced actin assembly formation. Langmuir 24:11975–11981

Magome N, Takemura T, Yoshikawa K (1997) Spontaneous formation of giant liposomes from neutral phospholipids. Chem Lett 26:205–206

Masuda K, Itoh H, Sakihama T, Akiyama C, Takahashi K, Fukuda R, Yokomizo T, Shimizu T, Kodama T, Hamakubo T (2003) A combinatorial G protein-coupled receptor reconstitution

system on budded baculovirus. Evidence for $G\alpha_i$ and $G\alpha_o$ coupling to a human leukotriene B_4 receptor. J Biol Chem 278:24552–24562

Mitchison TJ, Cramer LP (1996) Actin-based cell motility and cell locomotion. Cell 84:371–379

Mitchison TJ, Kirschner M (1988) Cytoskeletal dynamics and nerve growth. Neuron 1:761–772

Miyata H, Hotani H (1992) Morphological changes in liposomes caused by polymerization of encapsulated actin and spontaneous formation of actin bundles. Proc Natl Acad Sci USA 89:11547–11551

Miyata H, Nishiyama S, Akashi K et al (1999) Protrusive growth from giant liposomes driven by actin polymerization. Proc Natl Acad Sci USA 96:2048–2053

Nakano H, Kawarasaki Y, Yamane T (2004) Cell-free protein synthesis systems: increasing their performance and applications. Adv Biochem Eng Biotechnol 90:135–149

Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. Proc Natl Acad Sci USA 101:17669–17674

Nomura S, Yoshikawa K (2000) Giant phospholipid vesicles entrapping giant DNA. In: Luisi PL, Walde P (eds) Giant vesicles. Wiley, New York

Nomura SM, Yoshikawa Y, Yoshikawa K, Dannenmuller O, Chasserot-Golez S, Ourisson G, Nakatani Y (2001) Towards proto-cells:"primitive" lipid vesicles encapsulating giant DNA and its histone complex. Chembiochem 2(6):457–459

Nomura SM, Kondoh S, Asayama W, Asada A, Nishikawa S, Akiyoshi K (2008) Direct preparation of giant proteo-liposomes by in vitro membrane protein synthesis. J Biotechnol 133(2):190–195

Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. Chembiochem 4(11):1172–1175

Pautot S, Frisken BJ, Weitz DA (2003a) Engineering asymmetric vesicles. Proc Natl Acad Sci USA 100:10718–10721

Pautot S, Frisken BJ, Weitz DA (2003b) Production of unilamellar vesicles using an inverted emulsion. Langmuir 19:2870–2879

Pollard TD, Blanchoin L, Mullins RD (2000) Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. Annu Rev Biophys Biomol Struct 29:545–576

Pontani L-L, van der Gucht J, Salbreux G et al (2009) Reconstitution of an actin cortex inside a liposome. Biophys J 96:192–198

Rigaud JL, Lévy D (2003) Reconstitution of membrane proteins into liposomes. Methods Enzymol 372:65–86

Robelek R, Lemker ES, Wiltschi B, Kirste V, Naumann R, Oesterhelt D, Sinner E-K (2007) Incorporation of in vitro synthesized gpcr into a tethered artificial lipid membrane system. Angew Chem Int Ed Eng 46:605–608

Rodriguez N, Pincet F, Cribier S (2005) Giant vesicles formed by gentle hydration and electroformation: a comparison by fluorescence microscopy. Colloids Surf B Biointerfaces 42:125–130

Rodriguez OC, Schaefer AW, Mandato CA et al (2003) Conserved microtubule-actin interactions in cell movement and morphogenesis. Nat Cell Biol 5:599–609

Saitoh A, Takiguchi K, Tanaka Y et al (1998) Opening-up of liposomal membranes by talin. Proc Natl Acad Sci USA 95:1026–1031

Sato Y, Nomura SM, Yoshikawa K (2003) Enhanced uptake of giant DNA in cell-sized liposomes. Chem Phys Lett 380:279–285

Schroeder TE (1973) Actin in dividing cells: contractile ring filaments bind heavy meromyosin. Proc Natl Acad Sci USA 70:1688–1692

Shimanouchi T, Umakoshi H, Kuboi R (2009) Kinetic study on giant vesicle formation with electroformation method. Langmuir 25:4835–4840

Shimizu Y, Kuruma Y, Ying BW, Umekage S, Ueda T (2006) Cell-free translation systems for protein engineering. FEBS J 273:4133–4140

Takeda S, Saitoh A, Furuta M et al (2006) Opening of holes in liposomal membranes is induced by proteins possessing the FERM domain. J Mol Biol 362:403–413

Takiguchi K (1991) Heavy meromyosin induces sliding movements between antiparallel actin filaments. J Biochem 109:520–527

Takiguchi K, Yamada A, Negishi M et al (2008) Entrapping desired amounts of actin filaments and molecular motor proteins in giant liposomes. Langmuir 24:11323–11326

Tanaka-Takiguchi Y, Kakei T, Tanimura A et al (2004) The elongation and contraction of actin bundles are induced by double-headed myosins in a motor concentration-dependent manner. J Mol Biol 341:467–476

Taylor DL, Condeelis JS (1979) Cytoplasmic structure and contractility in amoeboid cells. Int Rev Cytol 56:57–144

Tsumoto K, Matsuo H, Tomita M, Yoshimura T (2009) Efficient formation of giant liposomes through the gentle hydration of phosphatidylcholine films doped with sugar. Colloids Surf B Biointerfaces 68:98–105

Tsumoto K, Nomura SM, Nakatani Y, Yoshikawa K (2001) Giant liposome as a biochemical reactor: transcription of DNA and transportation by laser tweezers. Langmuir 17(23):7225–7228

Tsumoto K, Yoshimura T (2009) Recombinant proteoliposomes prepared using baculovirus expression systems. Methods Enzymol 465:95–109

Vinarov DA, Loushin Newman CL, Markley JL (2006) Wheat germ cell-free platform for eukaryotic protein production. FEBS J 273(18):4160–4169

Yamada A, Yamanaka Y, Hamada T et al (2006) Spontaneous transfer of phospholipid-coated oil-in-oil and water-in-oil micro-droplets through an oil/water interface. Langmuir 22:9824–9828

Yamada NL, Hishida M, Seto H, Tsumoto K, Yoshimura T (2007) Unbinding of lipid bilayers induced by osmotic pressure in relation to unilamellar vesicle formation. EPL 80:48002

Yamaji K, Kanai T, Nomura SM, Akiyoshi K, Negishi M, Atomi H, Yoshikawa K, Imanaka T (2009) Protein synthesis in giant liposomes using the in vitro translation system of *Thermococcus kodakaraensis*. IEEE trans Nanobiosci 8:325–331

Yamashita Y, Oka M, Tanaka T, Yamazaki M (2002) A new method for the preparation of giant liposomes in high salt concentrations and growth of protein microcrystals in them. Biochim Biophys Acta 1561:129–134

Yu W, Sato K, Wakabayashi M, Nakaishi T, Ko-Mitamura EP, Shima Y, Urabe I, Yomo T (2001) Synthesis of functional protein in liposome. J Biosci Bioeng 92(6):590–593

Chapter 11
New and Unexpected Insights on the Formation of Protocells from a Synthetic Biology Approach: The Case of Entrapment of Biomacromolecules and Protein Synthesis Inside Vesicles

Pasquale Stano, Tereza Pereira de Souza, Matteo Allegretti, Yutetsu Kuruma, and Pier Luigi Luisi

Abstract In this chapter we present our "semi-synthetic" approach to the construction of minimal living cells.

Firstly, we shortly review the advancements carried out in the last few years, especially in the field of reactions inside liposomes. Then we discuss our recent study on the investigation of minimal cell size, carried out from the perspective of synthetic biology. In particular, we examine the question of the minimal physical size of cells by using liposomes with entrapped the complete ribosomal machinery for protein expression (green fluorescence protein), and making the assumption that this size would also correspond to a full fledged cell. We have found that liposomes with radius ca. 100 nm – which is the smallest size ever considered in the literature for protein expression – are still capable of protein expression. Finally, we show how this study has provided insightful results that actually broaden our perspective and pave the way to more extensive studies on vesicle formation and encapsulation of solutes. In fact, since classic statistical analysis gives zero or negligible probability for the simultaneous entrapment of many different molecular components in one single 100 nm radius spherical compartment at the given bulk concentration, the agreement between theoretical statistical predictions and experimental data is possible by assuming that the concentration of solutes in the liposomes becomes larger by at least of a factor 20.

These aspects have a great relevance in origins of life scenarios, where the formation of functional protocells is a decisive – yet not fully understood – event. We provide a final discussion on possible mechanisms and on the experimental investigations we would like to carry out in the near future.

P. Stano, M. Allegretti, and P.L. Luisi (✉)
Biology Department, University of RomaTre, Viale Guglielmo Marconi 446, 00146, Rome, Italy
e-mail: luisi@mat.ethz.ch

T.P. de Souza
Institute of Pharmacy, Friedrich-Schiller-Universität Jena, Lessingstraße 8, 07743, Jena, Germany

Y. Kuruma
Department of Medical Genome Sciences, Graduate School of Frontier Sciences,
The University of Tokyo, Kashiwanoha 5-1-5, Kashiwa-shi, Chiba, Japan

11.1   Minimal Cells and Synthetic Biology

As stated in a recent review, "we have been much better at taking cells apart than putting them together" (Liu and Fletcher 2009). In recent years, however, there has been a growing interest in a new approach for understanding biological systems. The emerging field of synthetic (constructive) biology is characterized by attempting to build complex bioinspired molecular systems, starting from their components. Synthetic biology encompasses several aspects of modern biological research, from the classical bioengineering approaches (Endy 2005; Drubin et al. 2007) that often use the concept of combination of "standard biological parts", to the in vitro projects aimed to construct biomolecular systems of increasing complexity from chemical or biological components (Benner and Sismour 2005; Forster and Church 2006; Luisi 2007; Schwille and Diez 2009). In a recent review, Victor de Lorenzo and Antoine Danchin (2008) define three pillars of synthetic biology as: (1) molecular biology, evolutionary genomics and biotechnology; (2) engineering, computing, modeling; (3) origins of life, artificial life, orthogonal life.

The "constructive" aspects of synthetic biology point to the generation of a new perspective in basic science, a perspective that has also stimulated epistemic discussions (Morange 2009; Luisi in press). The approach can be shortly described as follows: rather than taking cells apart to understand cellular functions, scientists now attempt to mimic biological functions by building them from basic parts. This approach has implications in basic research since it puts existing models of biological functions to experimental verification: if we really understand how a process works, and know all its components, we should be able to build a minimal physical model of it from scratch. Furthermore, if we can indeed reconstitute biological functions in the lab, a wide field of applications in healthcare, energy and material science comes into reach.

Our group has been involved in this research since the early 1990s, especially by focusing on the concept of autopoiesis and autopoietic systems (Luisi and Varela 1990; Luisi 2003). In particular, after the early results on autopoietic self-reproduction of synthetic compartments as micelles (Bachmann et al. 1992), reverse micelles (Bachmann et al. 1990) and lipid vesicles (Walde et al. 1994), we have introduced the notion of minimal cells (Oberholzer et al. 1995; Luisi et al. 2002).

In the last few years the notion of the "minimal cell", as a form of minimal life, has gained considerable attention both from the theoretical and experimental point of views (Szostak et al. 2001; Luisi et al. 2006; Forster and Church 2006). This concept is important for assessing the minimal and sufficient conditions for cellular life, and also to gain an insight of the early cells, conceivably much simpler than the modern cells. The research on minimal cells may be useful to develop simple cellular reactors for biotechnological applications (Pohorille and Deamer 2002; Forster and Church 2007; Zhang et al. 2008), and in the emerging field of "wetware" information and communication and technology (Cronin et al. 2006; Hadorn et al. 2009).

In this chapter, we would like to discuss the issue of the minimal cell size, a question that can be indeed effectively approached by synthetic (constructive) biology. Most of this chapter is based on our recent publication (Souza et al. 2009) but it has

been widened by some recent considerations based on preliminary results with ferritin-containing liposomes.

11.2   The Minimal Size of Cells

The issue of size limits of small microorganisms has been discussed in the literature by focusing on the complexity of modern and early cells, its relation to viability, biochemical requirements, as well as physical and evolutionary constrains (Knoll 1999). Microorganisms that receive many basic nutrients and metabolites from their "environment" need fewer genes, as observed in intracellular mutualists as *Buchnera aphidicola str. Cc* (357 genes), or host-associated parasites as *Mycoplasma genitalium* (482 genes). In contrast, free-living prokaryotes as *Bacillus subtilis* or *Escherichia coli*, whose metabolism needs to produce all kind of low molecular weight compounds, exceed 4,000 genes (4,099 and 4,289, respectively) (Fehér et al. 2007). The dimensions of such microorganisms range from about 0.3 to 1.5 μm, which correspond to a volume of about 0.013 and 1.6 μm^3 for *M. genitalium* and *E. coli*, respectively (Moore 1999). Together with a reduction of number of molecular species, a reduction in size is possible only if the number of copies of each species is reduced. It follows that the efficiency and the reproduction rate of hypothetical small cells should decrease. According to stringent assumption (e.g. one ribosome, one t-RNA set, one m-RNA for each of 100 non-ribosomal proteins, each present in ten copies), the diameter of a spherical cell compatible with a modern system of genome expression would be between 200 and 300 nm (Knoll 1999). The reduction in size is possible only accompanied by a parallel reduction of genes number, as evident by the fact that long double stranded DNA cannot be packed in small compartments, an argument that suggests the use of RNA, or of single stranded, flexible DNA, by hypothetical small cells (Boal 1999). Small cells, on the other hand, may be favoured in diffusion-limited growth conditions, thanks to their high surface-to-volume ratio, and also by the fact that – at a given membrane composition – small cells can sustain typical values of bacterial osmotic pressure and therefore avoid the construction of cell walls (this in turn would reduce the amount of DNA, the part namely devoted to the construction of the cell walls) (Boal 1999).

In addition to these theoretical considerations, we should mention the controversial reports on "nanobacteria" (Kajander and Ciftcioglu 1998) which are particles found in human and cow blood, with dimensions around 200 nm (diameter). It has been recently pointed out that the bacterial status of nanobacteria is still lacking satisfactory evidence, and the term "calcifying nanoparticles" has been recently adopted to describe such bodies (Kajander 2006). Other authors have reported on similar tiny corpuscles, but generally it is not clear whether they are "living" in the common sense of bacterial life (Ciftcioglu et al. 2006).

It is clear from this analysis that the question about the minimal size of a cell is still open – and that is a relevant question. The clarification of such a question would be important in the field of the origin of life, as it is conceivable that the origin of cell started with minimal protocells (radius <100 nm?) along the pathway

of evolution – and it is then interesting to assess whether and to what extent small compartments may permit life. The question becomes particularly timely in the present era of synthetic biology, as it has become possible to construct in the laboratory molecular systems that display living (or living-like) properties, such in the case of the semi-synthetic minimal cells.

We have recently tackled the question of the minimal cell size (Souza et al. 2009). Our approach is based on the use of lipid vesicles (liposomes) as models for minimal biological cells. In particular, we have used vesicles that contain the whole transcription and translation macromolecular machinery. This is based on the argument that the minimal size of such vesicles, which are able to display protein synthesis, may also characterize the minimal physical size for a viable cell. Protein biosynthesis, though representing only a subset of cellular minimal metabolism (~65% in term of number of genes in the minimal genome, see Gil et al. 2004), is already a process of considerable complexity, and, moreover, is – to date – the only experimentally feasible synthetic model.

11.3 In Vitro Protein Expression with a Minimal Set of Enzymes

Cell-free protein synthesis system is a biotechnological tool to perform protein synthesis in a test tube. Although conventional cell-free systems are based on cell-extracts such as *E. coli* (Liu et al. 2005) or wheat germ (Kohno and Endo 2007), a unique cell-free system has been developed by Ueda's research group (Shimizu et al. 2001). The system, called as PURE system, is composed of 31 translational factors (21 aminoacyl-tRNA synthatases, three initiation factors, three elongation factors, and four release factors), four energy recycling factors, T7 RNA polymerase, highly purified 70S ribosome, tRNAs, and some low molecular weight components (Fig. 11.1). All these enzymes and factors are essential for the transcription and translation reactions and represent the minimal set of macromolecules to accomplish protein synthesis in vitro. By using the PURE system, protein synthesis can take place within a totally defined condition concerning the number of components and those concentrations, whereas the extract-based systems are considered just as black-boxes. The composition of the PURE system is almost minimized for fulfilling protein synthesis reaction. This feature is a great advantage for the study of minimal cells (Luisi et al. 2006). In fact, within the framework of the synthetic biology, the first man-made cell (synthetic cell) should be constituted with a minimal number of factors. This is because all the constituting factors must be all reproduced within the synthetic cell, which is one of the characteristics of life. Regarding this point, it is evident that the number of the components is a critical element for the construction of a fully (core-and-shell) self-reproducing cell. If the PURE system can synthesize its all components, this would be considered a self-reproducing system and becomes a backbone of the self-reproducing cell.

From the practical viewpoint, the target protein is produced within some hours by the cell-free systems by simply adding the DNA gene in the mixture. It has been
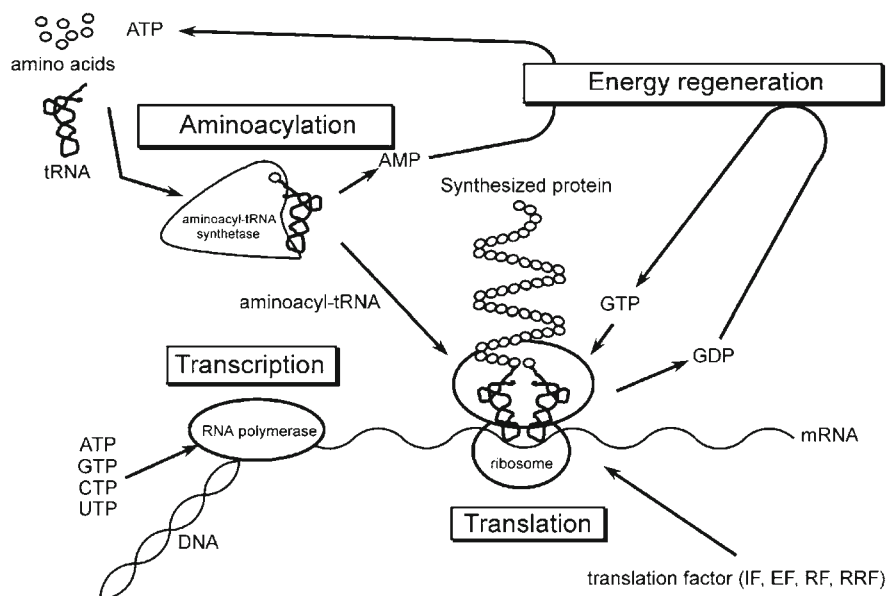
Fig. 11.1 Schematic of all components and reactions in the PURE system (Reproduced from Shimizu et al. 2005. With permission)

demonstrated that the cell-free mixture can be assembled even in a very small volume (of the order of 10 μl) and it is only weakly affected by liposome encapsulation. In this respect, it has been shown that PURE system can be easily encapsulated by film-hydration method (Sunami et al. 2010) or ethanol injection (Souza et al. 2009). Surprisingly, the cell-free system seems relatively stable and flexible for some experimental processes, e.g., a certain level of sonication, incubation at 42°C, addition of 250 mM sucrose (or other sugars), and several freeze-and-thaw cycles. These techniques are commonly employed for the liposome preparation, and the PURE system is stable with respect to these conditions. Moreover, liposome processing by membrane extruder (Hope et al. 1985; MacDonald et al. 1991) is also acceptable, which is an effective method to make homogeneous size and single lamellar liposomes. In fact, the internal protein synthesis using the extruded liposomes encapsulating PURE system is shown in Fig. 11.2. It can be concluded that by coupling liposome technology with the PURE system cell-free technology it will be indeed possible to design and construct minimal yet functional semi-synthetic cells.

11.4   Results

In this work we have used vesicles with size (diameter) in the range 100–200 nm. The preparation of 200 nm vesicles is possible by two classical methods: (1) extrusion of larger vesicles through polycarbonate membranes having 200 nm pores, and

Liposomes	non		+			
Extruding	non		non		400nm	
RNase	–	+	–	+	–	+
GPAT						

Size (diameter)		600-2000nm	400nm
lamellarity		Multi-lamellar	Unilamellar
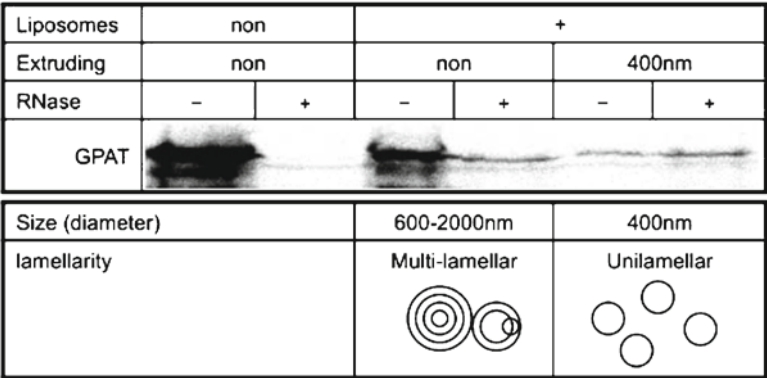
Fig. 11.2 Protein synthesis inside extruded liposome. The PURE system was encapsulated in liposomes and passed a micro-syringe extruder with 400 nm pore size membrane. GPAT protein was synthesized in the presence of [^{35}S]methionine and visualized by SDS-PAGE and autoradiograph. External protein synthesis of liposomes were blocked by addition of RNase which cannot enter liposome inside. Therefore, the band at the lanes indicated as "RNase+" demonstrate internal synthesis

(2) spontaneous vesicle formation via injection method (Batzri and Korn 1973; Domazou and Luisi 2002). As in vitro protein expression kit we have used: (a) commercial *E. coli* cell extracts, which are not well defined in terms of number and concentration of components but provide higher yields; and (b) the above-mentioned reconstructed transcription-translation kit from purified components. The composition of PURE system used in this study collectively counts 82 distinct macromolecular elements, can be found in several publications (Shimizu et al. 2001, 2005; Souza et al. 2009). In order to readily follow protein expression, as in previous approaches (for a review, see Chiarabelli et al. 2009), we have introduced in the translation system the *egfp* gene under T7 promoter, focusing then on the expression of enhanced green fluorescent protein (EGFP).

11.4.1 Protein Expression in Small Liposomes

In order to construct micro-compartments able to host the coupled transcription and translation reactions, we form lipid vesicles in situ in a solution containing all molecular components needed to perform the reaction. This is accomplished by injecting, into an aqueous phase, a concentrated solution of POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine) in ethanol.

The formation of liposome is immediate and their size and morphology depends on the concentration of POPC in ethanol. In particular, higher concentrations bring about the formation of large vesicles (typically multilamellar and/or multivesicular, as well as vesicle clusters. Starting from this heterogeneous population of vesicles, small vesicles (VET$_{200}$) can be obtained by extrusion. On the other hand, when the injection is carried out at a lower POPC concentration a homogeneous population

of small spherical vesicles can be obtained (VEI) We have preliminarily optimised the nature of the solvent (ethanol or methanol), the POPC concentration, and the injection volume, in order to get the smallest lipid vesicles, at the highest lipid concentration, and the lowest alcohol content in the samples. It follows that extruded vesicles are better obtained by injecting 500 mM POPC in a solution containing all the components of the transcription-translation kit, followed by extrusion through 100 or 200 nm-pore membranes. The injection of 150 mM POPC in ethanol represents instead the best choice to prepare small vesicles without extrusion. Additional details can be found in our original report (Souza et al. 2009). The particle size distributions of VET_{200} and VEI vesicles are shown in Fig. 11.3, where number-weighted dynamic light scattering (DLS) profiles are sketched. Notice that VEI, obtained spontaneously by the ethanol injection method, are indistinguishable from the VET_{200}, obtained in two steps (injection + extrusion). Both consist in a narrowly distributed population around 100 nm in radius.

After the formation of vesicles, in order to prevent the EGFP synthesis in the solution outside the vesicles, inhibitors such as EDTA, RNase or proteinase K are immediately added. These substances are unable to cross the lipid bilayer and therefore do not inhibit the reaction inside the liposomes. We have then tested the production of EGFP inside vesicles by using cell-free kits, as commercially available *E. coli* extracts or the above mentioned purified components, i.e., the PURE system. In Fig. 11.4a it is shown the time course of green fluorescence in 100 nm (radius) POPC extruded vesicles, in the absence (curve a) or in the presence (curves b, c) of externally added EDTA. In the first case, EGFP synthesis outside as well as inside liposomes is observed, whereas in the latter cases, the fluorescence increase corresponds to the synthesis of EGFP inside vesicles.

When EDTA is added before the formation of vesicles, no fluorescence is observed (curve d). The emission spectrum of vesicle samples, when the reaction is completed, has been recorded in the presence of sodium cholate (to eliminate the vesicle scattering), and it is shown in Fig. 11.4b. The emission maximum, i.e., 511 nm, and the peak shape further confirm the occurrence of well-expressed EGFP.

Fig. 11.3 DLS size distribution (CONTIN, 90°) of vesicles prepared by the extrusion method (VET200) and by the ethanol injection method (VEI). Under conditions optimized in this work, VET200 and VEI are almost undistinguishable by DLS analysis (Adapted from Souza et al. 2009. With permission)
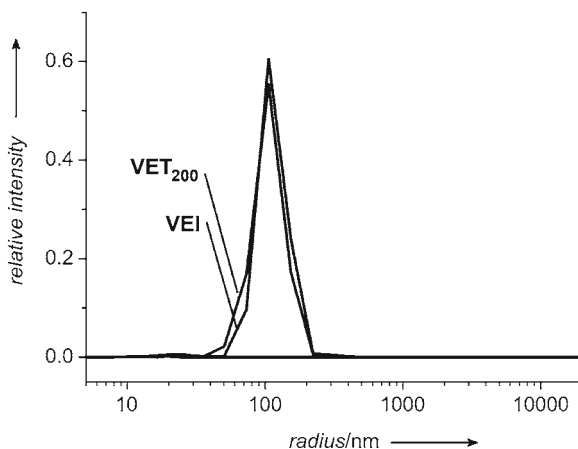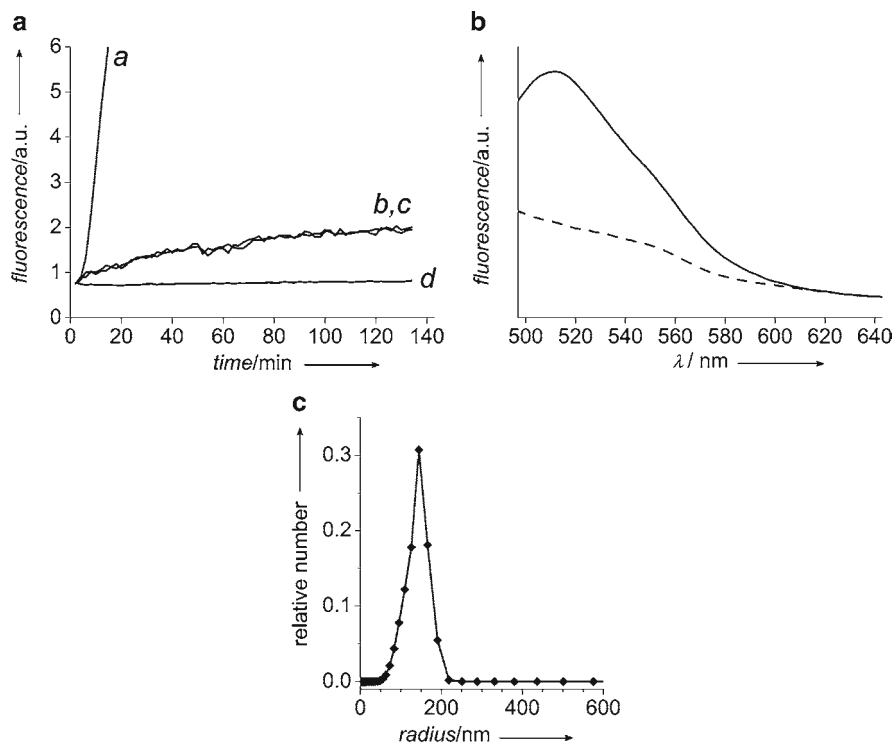
Fig. 11.4 EGFP Expression inside 200 nm extruded vesicles VET200. (**a**) Time profile of EGFP production: (**a**) EGFP expression in the absence of externally added EDTA; (**b**, **c**) EGFP expression inside vesicles, EDTA being added externally; (**d**) negative control, EDTA being added before the formation of vesicles. (**b**) Fluorescence emission spectra of EGFP-expressing vesicles after 3 h: EGFP-expressing vesicles (*continuous line*); negative control (*dashed line*). Spectra are recorded in the presence of sodium cholate in order to eliminate the scattering due to vesicles. (**c**) Number-weighted DLS size distribution of extruded vesicles. DLS analysis was carried out at the end of incubation (Adapted from Souza et al. 2009. With permission)

Finally, the size of vesicles has been measured by DLS analysis (Fig. 11.4c), revealing a sharp monomodal distribution with no particles larger than 200 nm radius. These data demonstrate that EGFP is effectively produced inside small vesicles. We can therefore draw a first conclusion. Ribosomal protein synthesis can take place in compartments as small as 100 nm radius, either by forming large vesicles which are then reduced to small ones, either by directly forming small vesicles by the injection method. Being capable of hosting protein biosynthesis, we suggest that this can be taken as the first experimental evidence that a dimension of ca. 200 nm may be compatible also with full-fledged cells.

This first conclusion stems from the qualitative assessment of protein synthesis (yes/no) into small vesicles. We can, however, improve our understanding of the system under study by performing a more detailed analysis.

In order quantify EGFP production into vesicles, it is possible to express the EGFP yield in relative and absolute terms. The relative yield (yield (%)) is calculated as the ratio between fluorescence intensity – due to EGFP synthesis – produced inside

vesicles (internal EGFP) versus that one produced inside and outside vesicles (total EGFP). This value follows the tradition of expressing the entrapment of solutes into vesicles as "entrapped volume" (see Appendix). On the other hand, the absolute amount of EGFP is calculated by means of a calibration line, and normalized with respect to the lipid concentration (that may vary among experiments), giving the "entrapment efficiency" (see Appendix). These two measures of EGFP yields are shown in Table 11.1, together with other measured and calculated values. Small vesicles, despite their small individual volume and irrespective of the transcription/translation kit used, produce EGFP with relative yield of about 10%. The cumulative internal volume of all vesicles, calculated under the hypothesis of spherical shape, and at the concentration used in these experiment (POPC 15 mM), is also about 10% in both cases. Apparently, protein expression inside vesicles follows the expetations based on the ratio between internal (encapsulated) volume and total available volume. This conclusion is however based on average quantities. Concerning the absolute yield, it is possible to estimate that PURE system and *E. coli* extracts produce, inside small vesicles, respectively, 103 and 309 pmol EGFP/mmol POPC. These yields correspond to 5–10 EGFP molecules/100 vesicles (considered unilamellar). The higher yields obtained with the *E. coli* extracts is in all likelihood due to the fact that we have a much greater number of enzymes and other components with respect to the minimal number of components in the PURE system. Similar results have been obtained recently in a systematic comparative study on cell-free systems in solution (Hillebrecht and Chong 2008). As evident from the large standard deviation found in Table 11.1, experiments were often characterized by an quite variable output, probably due to the intrinsically stochastic nature of solute entrapment and confined reactivity.

Table 11.1 Summary of EGFP yields in lipid vesicles

	1	2	3
Transcription-translation kit	E. coli extracts	E. coli extracts	PURE system
[POPC] (mM)	15	8.5	15
Liposome type[a]	VET200	VEI	VET200
Radius[b] (nm)	102 ± 23	124 ± 21	112 ± 11
Polydispersity[c]	0.10 ± 0.03	0.16 ± 0.05	0.12 ± 0.02
EGFP relative yield (%)[d]	10.0 ± 7.0	0.3 ± 0.1	9.9 ± 5.2
Internal vesicle volume (%)[e]	10.2	7.2	11.3
pmol EGFP/mmol lipid	309 ± 211	27 ± 6	103 ± 45
EGFP molecules/vesicle	0.11 ± 0.10	0.01 ± 0.006	0.05 ± 0.02
Number of experiments (replicates)[f]	3	3	3

[a]VET200 = vesicles obtained by extrusion of pre-formed vesicles; VEI = vesicles obtained directly by injecting lipids in ethanol
[b]Hydrodynamic radius as calculated by cumulant analysis of dynamic light scattering (DLS) data
[c]Polydispersity index as calculated by the cumulant analysis of DLS data
[d]Relative EGFP yield (%) is calculated as the ratio between fluorescence due to internal synthesis and fluorescence due to total (internal + external) synthesis
[e]Internal vesicle volume has been calculated by considering: (1) unilamellar spherical vesicles; (2) bilayer thickness: 3.8 nm; (3) POPC headgroup area: 0.72 nm^2
[f]Number of experiments (replicates). Data reported as mean value ± standard deviation are based on the measure of replicates
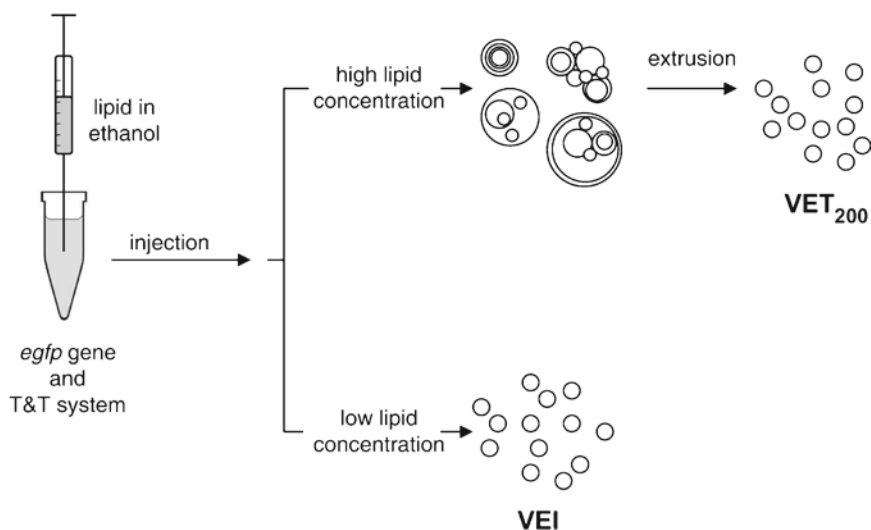
Fig. 11.5 Experimental approach to the preparation of small vesicles (100 nm in radius). (**a**) A small amount (3% v/v) of a POPC/ethanol solution is injected into a vial containing the *egfp* gene and the translation-transcription kit. Depending on the POPC stock concentration, a population of large (multilamellar, heterogeneous, clusterized) vesicles, or small (generally unilamellar) vesicles (VEI) is obtained. Small vesicles (VET200) can be obtained from a population of larger vesicles by extrusion. After vesicle formation and processing, an external inhibitor is added, so that EGFP can be produced only inside vesicles. Vesicles not drawn to scale (Adapted from Souza et al. 2009. With permission)

The preparation method described above starts with the encapsulation of the transcription-translation components in large vesicles, which are then forced to divide by shearing forces (extrusion). Alternatively, small vesicles can be prepared directly in one step (Fig. 11.5) by the injection method, yielding vesicles VEI, which are indistinguishable from extruded vesicles VET_{200}. Figure 11.6 shows the fluorescence time course vesicles prepared by multiple injection method, in the presence of external inhibitor (curve a). When the inhibitor is added to the transcription-translation kit before the formation of liposomes, the lower profile (curve b) is obtained. The size of vesicles was measured by DLS analysis, indicating that the size of vesicles is always below 220 nm. The amount of EGFP produced in this case is lower. In particular, the relative yield of *E. coli* cell extracts entrapped in vesicles (radius: 124 nm) is $0.3 \pm 0.1\%$, corresponding to about 27 pmol EGFP/mmol POPC; in addition, the EGFP produced by the reconstituted PURE system cannot be detected. Therefore, the maximal yield we could measure corresponds to one EGFP molecule/100 vesicles. Very probably, there are few vesicles synthesizing EGFP, so that a realistic picture based on individual vesicles may substantially differs from average values. For example, the figure of one EGFP molecule/100 vesicles may turn to actually be ten EGFP molecules synthesized inside one vesicle out 1000.

We can also compare the efficiency of protein expression inside liposomes with respect to the bulk water (Table 11.1) and make comparison between compartmentalized and bulk EGFP synthesis as well as between liposomes prepared in different

Fig. 11.6 EGFP Expression inside 200 nm spontaneously formed vesicles VEI. Time profile of EGFP production: (**a**) EGFP expression inside vesicles, EDTA being added externally; (**b**) negative control, EDTA being added before the formation of vesicles. Error bars represent the standard deviation of three independent samples (**a**) and two negative control ones (**b**), respectively (Adapted from Souza et al. 2009. With permission)
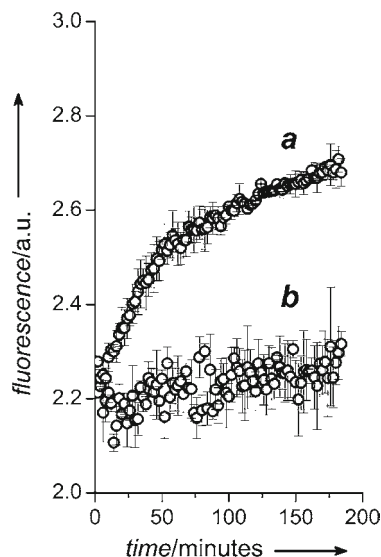
way. The typical EGFP production in bulk water under similar conditions (in the presence of liposomes and 3% v/v ethanol) is around 4.6 nM, as determined by a control experiment. In contrary, the EGFP production inside compartments corresponds to an internal concentration of about 15–46 nM. We reach the conclusion that compartmentalized protein synthesis occurs about 3.3–10 times more effectively than in the bulk. This calculation is based on the assumption that all liposomes contribute equally to the EGFP expression. Most likely, only a fraction of the liposomes will be fully viable (containing all 83 macromolecular components plus small molecules), and accordingly, the efficiency referred only to this fraction will be higher than the factors 3.3–10 mentioned above. By similar calculation it can be said that VEI gives an internal EGFP concentration of about 3.2 nM which is in line with the bulk value, but again this yield refers to an average EGFP content inside liposomes.

Due to these anomalous values of EGFP production yields and entrapped volume ratioes, we claim that very probably only few vesicles are effectively able to produce proteins, and that such "special" vesicles can also produce a large amount of EGFP, well above the value obtained for experiments in bulk phase. What is therefore the reason of such behavior?

11.5 The Conundrum of the Multiple Entrapment and the Hypothesis of "Superconcentration"

Data presented in the previous section provide clear evidence of the ribosomal synthesis of protein inside liposomes with radius ca. 100 nm. And whereas this is simply a fact, the existence of this fact from a mere theoretical point of view is not at all clear. We would like now to dwell upon this "entrapment conundrum", which

represents indeed one of the most interesting aspects of the work with liposomes. The starting consideration is that, since POPC liposomes do not fuse each other (Cheng and Luisi 2003) and since nucleic acids, proteins, and small charged molecules cannot cross the liposome membrane, all required components must be present inside every EGFP-producing vesicle from the moment of its formation. The PURE system, as the simplest transcription-translation kit capable of protein synthesis, is composed by 83 different components (enzymes, ribosomes and t-RNAs, plus the *egfp* gene, see Souza et al. 2009), so that the inescapable conclusion from our work is that we are in the presence of vesicles that have entrapped more than 80 different macromolecular components. *E. coli* cell extracts certainly contain more components, but we can base our considerations of the minimal number of PURE system.

The internal volume of a 100 nm (inner radius) vesicle is around 4.2×10^6 nm^3, so that there is enough space to accommodate several bulky ribosomes (~1,000 nm^3), and all enzymes such as the complete set amino acyl-tRNA synthase, translation factors, low molecular-weight compounds, etc., all in several copies. However, the key point is not the lack of sufficient space, rather the adverse co-entrapment statistics of all components in the same lipid vesicle. We are dealing with the probability of entrapping 83 different molecular components (i.e., the PURE system) inside one single compartment, whereby under our set of experimental conditions only random statistics should govern the entrapment. Let us calculate such a probability by assuming that the entrapment events follow the Poisson distribution (justification and discussion are given in Souza et al. 2009). The probability $\wp$ of finding a lipid vesicles that contains at least one molecule of each of the 83 molecular species is given by Eq. (11.1),

$$\wp_c(R) = \prod_{k=1}^{N}\left[1 - \exp\left(-C_k V(R)\right)\right] \qquad (11.1)$$

where the product is extended over the $N = 83$ species, and C_k are the numerical concentrations of the k-th species, whereas V is the vesicle volume. The dependence of the probability and of the volume from the vesicle radius R has been shown explicitly. The probability function (Eq. 11.1) gives indeed intruiguing insights into the protein expression in the small vesicles used in this study. The probability versus radius functions are illustrated in Fig. 11.7, where the three different solid curves show the probability of formation of a "viable" liposome as a function of its size, under the hypothesis that at least one (curve a), five (curve b), or ten (curve c) molecules of each molecular species are present inside a liposome of given size. The calculation clearly shows what we have called the "entrapment conundrum". In fact, according to these curves, for vesicles with radius below ca. 250 nm, the probability of a simultaneous co-entrapment of the transcription-translation molecular components should be vanishing small. For example, in the case of 100 nm radius vesicles, the probability of independent entrapment of at least one copy of the 83 components (each present at the concentrations used in this study) is 10^{-26}. In contrast, larger vesicles – for example those with a radius of 1 μm – have probability ~1 in the three cases (notice that the expression of EGFP in *extruded* VET$_{200}$, despite the fact that
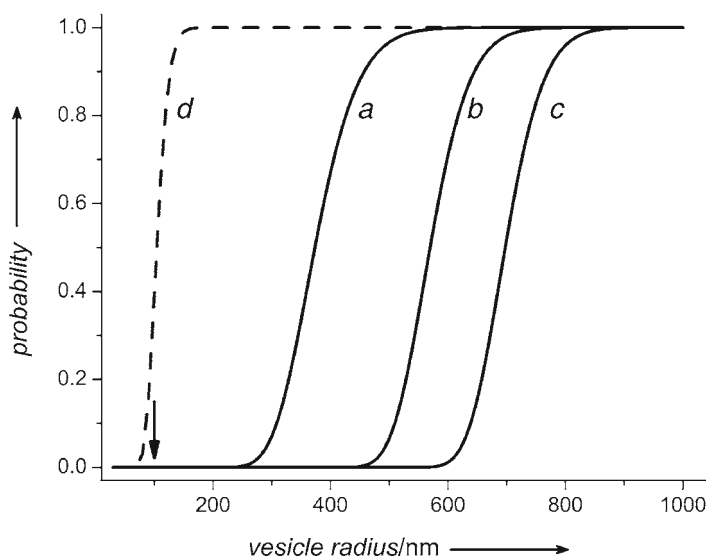
Fig. 11.7 Probability of co-entrapment of all macromolecular components of transcription-translation kit inside lipid vesicles of a given radius. The entrapment of each molecule is modeled as a poissonian process, and the cumulative probability is calculated as product of probabilities of independent events. The concentrations of enzymes, t-RNA, and other factors correspond to those of the PURE system. The arrow points to the vesicle size used in the experiments. The three solid curves indicate the probability of entrapping at least one (**a**) five (**b**) or ten (**c**) copies of each molecular specie inside the same vesicle. The dashed curve (**d**) indicate the probability of entrapping at least one copy of each molecular species under the hypothesis that their concentrations are all 50 times higher than the nominal (*bulk*) concentrations (Reproduced from Souza et al. 2009). With permission

they derive from large vesicles, cannot be easily justified too; for a short discussion, see supplementary material in Souza et al. 2009).

The argument of association between molecules does not hold in this case. In fact, we excluded the possibility of massive association of enzymes, t-RNA, and other components in solution, by measuring the size of particles present in the reacting mixture. DLS measurement of the bulk PURE system initial solution shows no aggregates larger than the ribosomes (Souza et al. 2009). At any rate, we have also calculated the probability curve under the assumption that the 83 components were associated in 50, 20, or 10 molecular clusters. In no case the probability reaches experimentally significant values (Souza et al. 2009). What can be then the explanation of the contradiction between experimental facts, and the curves in Fig. 11.7? It is interesting to notice that this question has not been arisen before in the literature, except by us in a very dubitative way (Luisi 2006).

The most interesting possibility is that the entrapment of the components is attended by a concentration enrichment, so that the concentration of the components inside the vesicles is significantly larger than bulk one. This would mean that the C_k values in the Eq. (11.1) are larger than we assumed to be. How much larger should they be, in order that the probability of total entrapment reaches finite

probability values? For example, Eq. (11.1) gives probability values equal to 1% only if we assume local enzyme concentrations 20 times higher than nominal bulk concentrations. Higher probability values correspond to higher enhancement factors, up to 50 × (~40%) or 100 × (~90%). This very high local concentration would also provide the simplest rationale for the enhanced protein expression yield inside vesicles (see Table 11.1). The mechanism of concentration enrichment inside vesicles cannot be explained with present data. Our working hypotheses, which guides our current experimental approach, focus on the possible cooperative effects arising during the formation of vesicles in the presence of solutes. Such effects might bring about a spontaneous accumulation of solutes into vesicles, with possible consequent expulsion of water during solutes entrapment, so that the local solutes' concentrations increase. This is also related to possible depletion effects (Minton 2001) and consequent internal crowding, in view of the fact that the 100 nm (inner radius) vesicles have a confinement volume of only 4.2 aL (4.2×10^{-18} L).

11.6   Investigating Protein Entrapment into Vesicles

Stimulated by our working hypothesis on superconcentration of solutes inside vesicles, we have recently started a systematic investigation on the entrapment of proteins inside lipid vesicles. To carry out this study, clearly, the classical methods for determining the entrapment yields are not suitable. This is made clear by considering that our data on cell-free proteins synthesis inside vesicles tell us that most likely only a minor fraction of vesicles are "viable", i.e., contain all the required macromolecules needed for EGFP production. Clearly, we are not interested in mean values, that can be obtained by classical approaches. It is useful to shortly recall here how the average entrapment is measured. When liposomes are prepared in the presence of a water-soluble compound, the amount of entrapped solute can be easily determined by measuring the concentration of the probe after removal of unentrapped solute. In the ideal case, the amount of entrapped solute should be proportional to the "entrapped volume", i.e. the fraction of the solution volume that correspond to the sum of internal volumes of all liposomes. The amount of entrapped probe is often normalized to take into account the lipid (or liposome) concentration, in order to express the "entrapment efficiency". This is a convenient way of expressing entrapment because the entrapped volume depends on the lipid concentration. Clearly, all these procedures give average values, because they are based on the measure of all solutes entrapped in all liposomes. From these average values, it is possible to calculate the average number of solutes per liposome, and compared with the expected value, calculated by keeping into account the nominal solute concentration (i.e., its concentration in the aqueous phase) and the liposome internal volume.

It is evident, however, that different liposomes can contain different number of solutes. This depends on two factors: firstly, liposomes can follow different formation pathways, and therefore contain a different amount of solutes for mechanistic reasons. Secondly, spontaneous fluctuations always occur in natural systems and they

account for a kind of natural heterogeneity in solute concentration, even if the mechanism of liposome formation is uniform. Fluctuation theory predicts that the magnitude of number fluctuations scales as $1/\sqrt{N}$, where N is the expected number of solutes. Large stochastic fluctuations are therefore associated to systems characterized by small numbers.

There are very few reports on the study of solute entrapment for *each* vesicles. The first report came from Chiu and coworkers, who used single-vesicle photolysis and confocal single molecule detection to determine the carboxyfluorescein concentration within individual giant vesicles (GVs) (Sun and Chiu 2005). The second report was published by the lab of Keating (Dominak and Keating 2007). By using GVs again, they have shown that the majority of GVs had high internal concentrations of polymer or small-molecule encapsulants equal to or slightly greater (×1.5) than the external concentration. Finally, a recent report by Stamou (Lohse et al. 2008) shows how the solute concentration in single vesicles with diameters down to 100 nm can be measured by fluorescent methods, revealing: (i) the existence of empty vesicles and (ii) an inverse relation between internal concentration and vesicle diameter. The latter result is quite surprising and goes against the expectation (independency from vesicle diameter). Another technique that can be used to determine the individual vesicle entrapment is electronmicroscopy. In particular, since submicrometric liposomes can be visualized by electronmicroscopy, the technique is based on the direct counting of molecules inside each liposome. This is indeed possible by the cryo-transmission electronmicroscopy (cryo-TEM), that allow direct visualization of liposomes and of their content. The target molecule should be itself visible by TEM, or made visible by proper labeling. Ferritin is a iron-storage protein that has been extensively used in electronmicroscopy; its scattering power derives from the iron phosphate core composed by about 4,500 iron atoms. It is therefore possible to entrap ferritin inside liposome and directly count the number of ferritin found in each liposome. In the past, we have already exploited this tool by investigatin the fate of ferritin entrapped inside liposomes during liposome self-reproduction (Berclaz et al. 2001a, b). Currently, we are investigating the entrapment of ferritin inside small liposomes by cryo-TEM in order to directly count the individual vesicle content, and make a comparison with theoretically expected value, obtained on the basis of average concentration. Moreover, it is also possible to check the *number distribution* of ferritin inside vesicles with the same internal volume, in order to verify the Poisson distribution. By varying the lipid concentration, the ferritin concentration and the preparation method we will obtain detailed and unique information about the process of ferritin entrapment. These results will be used as a basis for the development of a general model of solute entrapment inside vesicle under the hypothesis that no specific interactions occur between ferritin molecules and between ferritin molecules and lipids.

We can show here only preliminary results. A more detailed analysis is reported in our recent publication (Luisi et al. 2010). We have investigated the system composed by POPC/oleate (4/1 mol/mol) vesicles prepared by film hydration in the presence of ferritin solution (4, 8 and 16 μM). Vesicles are not extruded and therefore represent the outcome of spontaneous processes of solute entrapment

regulated only by the physico-chemical features of molecular components and by their self-assembling/self-organization. Firstly we compute the average number of ferritin molecules that should be found in vesicles with radius 50 nm. The value is 1.26, 2.52 and 5.04 ferritin/vesicle for liposomes prepared by using 4, 8 or 16 μM ferritin solution, respectively. It is apparent from the preliminary analysis that the actual number of ferritin molecules inside vesicles does not correlate with the average expected value. In particular, we have found a very large number of empty vesicles or some vesicles containing only one or two ferritin molecules, whereas the vesicles with more ferritins become less probable as the number of ferritin molecules increases. Surprisingly, however, we have found a few vesicles which have instead entrapped a quite large number of ferritin molecules, well above the average value computed on the basis of bulk concentration. Figure 11.8 illustrate four unilamellar vesicles with different ferritin content. The comparison between measured and expected ferritin content is reported in Table 11.2.

For example, vesicle *a* (radius 39 nm), which contains no ferritin, should contain instead 1.19 ferritin. However, the probability of finding a empty vesicle with such size is about 30%. Its existence is therefore expected on the basis of Poisson distribution. Similar arguments hold for vesicle *d* (68 nm), containing five ferritin molecules versus the expected value of 6.34. Poisson distribution imply that about 15% of such vesicles may indeed contain five ferritin molecules. Vesicles *b* and *c*, on the other hand, have similar radius (59 versus 58 nm) but have quite different content. Whereas vesicle *c* is empty, vesicle *b* contains 38 ferritin molecules. This corresponds to a ×10 enhancement of local ferritin

Fig. 11.8 Cryo-transmission electronmicrograph showing four unilamellar POPC/oleate (4/1) vesicles prepared in the presence of 8 μM ferritin (visible as an intense black spot). Unentrapped ferritin has been removed by size-exclusion chromatography. Liposomes have been prepared by direct film hydration
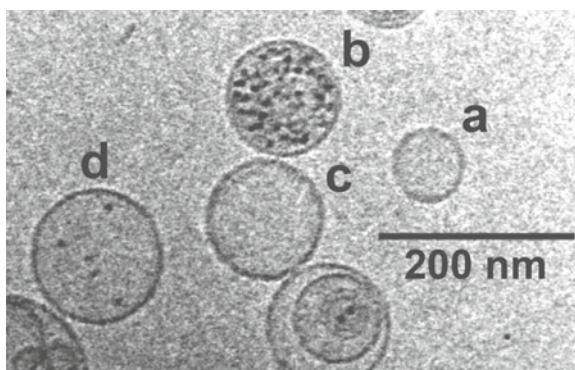
Table 11.2 Comparison between the expected and the actual number of ferritin molecules in the vesicles shown in Fig. 11.8

Vesicle	Radius (nm)	N ferritin (expected)	N ferritin (found)	Found/expected concentration ratio	Poisson probability (N_found)
a	39	1.19	0	0	0.304
b	59	4.14	38	9.18	8.5×10^{-24}
c	58	3.93	0	0	0.020
d	68	6.34	5	0.79	0.151

concentration. The probability of occurrence of vesicle *c* is low but finite (2%), whereas the probability of finding a vesicle *b* as in Fig. 11.8 is essentially zero (one vesicle out a population of 10^{23} vesicles). Since we have found such kind of vesicle in the ratio of about 1% (data not shown), we conclude that the formation of "superconcentrated" vesicles is a rare event but occurring with a probability ca. 10^{20} higher than expected on the basis of random entrapment of water-soluble molecules.

At the present stage of investigation it is not easy to find a physical explanation for these observations. Clearly, the relevance of such facts increases if similar effects could be found when different probes are entrapped in different kind of vesicles, so that a general theory about the encapsulation of macromolecules can be outlined. Among the elements that have to be taken into account for explaining the observations, it is probably important to consider: (1) the mechanism of liposome formation; (2) chemical nature of liposome interface and the chemical structure of the solute; (3) the solute concentration and the solute/lipid ratio; (4) the stochasticity of events at the molecular level; (5) the abundance of "empty" vesicles and the apparent violation of Poisson distribution with the consequent establishment of another distribution law. As anticipated, one plausible hypothesis focuses on non-linear (cooperative) effects between locally high protein concentration, generated for example by stochastic fluctuations or by spontaneous accumulation near the membrane, and the dynamics of liposome formation (closure), e.g. by closure of open-ended bilayers, or by pearling-like mechanisms (Lasic 1993) (all diverse liposome formation mechanisms can be reduced – after all – to these two basic morphological transformation). Further studies are currently devoted in our lab to obtain a more detailed analysis of ferritin entrapment inside liposomes.

Coming back to the case of PURE system entrapment inside vesicles, it is clear that the study of ferritin entrapment represents only the first step of a more complex approach. In particular, results from ferritin entrapment refer to the encapsulation of only one kind of solute (ferritin), and reveal the number distribution of it inside vesicles. In contrary, when two or more solutes have to be entraped, the additional issue of diverse possible *combinations* arises (e.g., ten "A" molecules and five "B" molecules, versus five "A" and ten "B" molecule). Due to the large number of molecules involved in the entrapment of the transcription and translation kit, the progress in our undestanding must necessarily pass through simpler cases of increasing complexity (one solute, two solutes, a small metabolic route, etc.).

11.7 Concluding Remarks

All these considerations open new perspectives to the involvement of a closed membrane to the reactivity of cell models and most probably, by inference, of biological cells; and may therefore shed a new light to the importance of (micro)

compartmentation effects in the origin of life scenarios – a notion which, on the basis of self-reproducing micelles and vesicles, had been emphasized since long (Luisi et al. 1999). In the specific case of this study, the relevant question is about the timing at which membrane compartments came into picture as host for the first form of metabolism. If we assume first the development of a relatively complex metabolism outside the compartment, then we have a real difficulty to conceive how all these molecular components would have been all entrapped into a single compartment. On the other hand, the hypothesis that metabolism originated from the inside of a compartment and developed inside meets the difficulty that we would have to conceive a semi-permeable highly sophisticated membrane from the very prebiotic beginning, which does not appear plausible.

As a final remark, we would like to emphasize that the constructive approach used to verify that small vesicles can host complex reaction networks, as the protein synthesis (taken in this work as a paradigm of the whole cellular metabolism), has brought about the identification of a basic biophysical mechanism. Such finding can be very relevant for basic science, and more in particular to approach the question of origins of life, as well as biotechnological applications of liposomes as microreactors, sensors, and drug delivery system. This is one of the powerful aspects of synthetic biology that is perhaps not yet realized by the growing community of scientists working in this novel field.

11.8 Experimental Section

Details about material and methods used for demonstrating the protein expression inside small vesicles are given elsewhere (Souza et al. 2009; Stano et al. 2010). We report here the summary of the procedure. The mix of transcription/translation components, preparation of vesicles, and vesicle manipulations were all done in the cold room (4°C), and samples were kept constantly in ice, in order to inhibit early start of transcription-translation reactions. All equipments, which are used to manipulate or process vesicles were place in the cold room at least 30 min before experiments.

The components of the cell-free protein expression kit (Promega or PURE system) were mixed in a Eppendorf tube placed in ice, and then liposomes were formed in situ by the ethanol injection method. Following the indications of the provider, we mixed: S30 extracts (20 μL), T7S30 extracts (15 μL), amino acid complete mixture (10 μL) and pWM-T7-EGFP plasmid (2.0 μg), (Promega kit); solution A (25 μL), solution B (10 μL), and pWM-T7-EGFP plasmid (1 μg), (PURE system). In both cases, the final volume was adjusted with distilled water to 48.5 μL. In all experiments, the stock concentration of pWM-T7-EGFP plasmid in pure water was adjusted to 1 μg/μL, so that 1 or 2 μL were added to the reacting mixture.

Vesicles were formed by injecting a solution of POPC in the transcription-translation mixture, and processed as described by Souza et al. (2009) to form 100 nm radius vesicles. In order to inhibit the protein expression, we added, after the liposome formation, three different kind of inhibitors: (1) 45 mM EDTA; or (2)

0.18 mg/mL RNase A; or (3) 0.18 mg/mL protease K. The resulting EGFP expression was measured by fluorometrically. Real-time measurement of green fluorescent is conveniently performed by incubating the samples in a RT-PCR Corbett Rotor-Gene 6000, used as sensitive fluorometer, by running it at constant temperature. Such instrument improves reproducibility, enhances sensibility, allows the use of small sample volumes (25 or 50 mL), and provides homogeneous heating. Vesicle size was assessed by DLS analysis at the end of the incubation time.

Appendix

Entrapment of water-soluble compounds inside liposomes

The entrapment of water-soluble compounds inside liposomes can be described by different approaches. The first (phenomenological) one is the most direct and simply accounts for the moles of solute per mole of lipid. Such value is called *entrapment efficiency* (**EE**), it can be always measured, also for lipid-soluble or membrane-associated molecules. In the case of water-soluble compounds, however, it is easy to define a theoretical entrapment value that is based on the expected concentration of the solutes in the aqueous core of liposomes.

Under the hypothesis that the solute concentration inside liposomes is equal to the solute concentration of the solution used to prepare liposomes ($C_{in} = C_{out} = C_0$), it is expected that liposomes with volume V_1 will entrap, on average, $N_1 = C_0 V_1$ solute molecules. The *individual entrapment yield (or ratio)* can be therefore defined as $EE_1 = X/N_1$, where X is the actual number of solute molecules found in a liposome that should contain istead N_1 molecules. When $X > N_1$, then $EE_1 > 1$, and vice versa; when $X = N_1$ then $EE_1 = 1$. In order to measure EE_1 it is needed to employ a technique that gives the number or the concentration of solute in each liposome, as well as their volume.

Another way to express the expected entrapment yield is to consider the *overal entrapped volume ratio*, defined as $R = V_L/V_{tot}$, where V_L is the sum of internal volumes of all liposomes, and V_{tot} is the volume of the solution. Such ratio depends in turn from the liposome size and concentration (and therefore from the lipid concentration, that is the easily measurable quantity).

In the case of water-soluble molecules, these three ways to express the entrapment of solutes inside liposomes are related with each other, as shown in the numerical example below.

Consider 1 mL of a monodisperse suspension of spherical unilamellar POPC liposomes having 100 nm radius. The POPC concentration (C_{lipid}) is 30 mM and the

solute concentration is 20 μM. Molecular parameters of POPC molecule are: head group area 0.72 nm^2, tail length 1.9 nm. The **EE** is calculated as 1.34 10^{-4} solute molecule/lipid molecule, whereas the overall entrapment ratio **R** is around 0.2, i.e., 20% of the solution volume actually corresponds to internal liposome volume, so that 20% of solutes present in 1 mL should therefore be found inside liposomes. Since there are 5.38 10^{13} liposome/mL and 1.2 10^{16} solute molecule/mL it is expected to found on average 45 molecule/vesicle. The same value (45) is easily calculated by considering the solute concentration (20 μM) and the volume of one vesicle (3.73 10^6 nm^3), by simply applying $N_1 = C_0 V_1$. By considering that the aggregation number (g) for 100 nm radius vesicles is 3.36 10^5, the product **EE** $\times g$ also gives the expected value for N_1 (45 solute molecule/vesicle).

From the knowledge of chemical composition and vesicle size, we can therefore derive:

- The expected number of solute molecules/vesicle (N_1)
- The expected entrapment ratio, on the basis of the overall vesicle entrapped volume (**R**)
- The expected entrapment efficiency **EE** as C_0 **R** /C_{lipid}. The latter value is then checked against the experimentally measured quantity, so that conclusions can be drawn about the entrapment process

The **EE** value is an average quantity, that consider all liposomes together. Only in special cases it is possible to measure N_1 and therefore **EE$_1$**, a quantity that is related to each liposome instead.

In addition, from the expected (calculated) value N_1 it is possible to calculate the probability of finding liposomes containin any number X of solute molecules by modeling the process as a random event, and using the Poisson probability:

$$\wp(X) = e^{-N_1} \frac{N_1^X}{X!}$$

Such probability is then compared, if experimentally feasible, with the frequency of observed vesicles with X solute molecules inside.

References

Bachmann PA, Walde P, Luisi PL, Lang J (1990) Self-replicating reverse micelles and chemical autopoiesis. J Am Chem Soc 112:8200–8201

Bachmann PA, Luisi PL, Lang J (1992) Autocatalytic self-replicating micelles as models for prebiotic structures. Nature 357:57–59

Batzri S, Korn ED (1973) Single bilayer liposomes prepared without sonication. Biochim Biophys Acta 298:1015–1019

Benner SA, Sismour A (2005) Synthetic biology. Nature Rev 6:533–543

Berclaz N, Blöchliger E, Müller M, Luisi PL (2001a) Matrix effect of vesicle formation as investigated by Cryotransmission electron microscopy. J PhysChem B 105:1065–1071

Berclaz N, Müller M, Walde P, Luisi PL (2001b) Growth and transformation of vesicles studied by Ferritin Labeling and Cryotransmission electron microscopy. J Phys Chem B 105:1056–1064

Boal D (1999) Mechanical characteristics of very small cells. In: Size limits of very small microorganisms. National Academic Press, Washington DC, pp 26–31

Cheng Z, Luisi PL (2003) Coexistence and mutual competition of vesicles with different size distributions. J Phys Chem B 107:10940–10945

Chiarabelli C, Stano P, Luisi PL (2009) Chemical approaches to synthetic biology. Curr Opin Biotechnol 20:492–497

Ciftcioglu N, McKay DS, Mathew G, Kajander EO (2006) Nanobacteria: fact or fiction? Characteristics, detection, and medical importance of novel self-replicating, calcifying nanoparticles. J Investig Med 54:385–394

Cronin L, Krasnogor N, Davis BG, Alexander C, Robertson N, Steinke JHG, Schroeder SLM, Khlobystov AN, Cooper G, Gardner PM, Siepmann P, Whitaker BJ, Marsh D (2006) The imitation game – a computational/chemical approach to recognizing life. Nature Biotech 24:1203–1206

De Lorenzo V, Danchin A (2008) Synthetic biology: discovering new worlds and new words. EMBO Reports 9:9

Domazou AS, Luisi PL (2002) Size distribution of spontaneously formed liposomes by the alcohol injection method. J Lipos Res 12:205–220

Dominak LM, Keating CD (2007) Polymer encapsulation within giant lipid vesicles. Langmuir 23:7148–7154

Drubin DA, Way JC, Silver PA (2007) Designing biological systems. Genes Dev 21:242–254

Endy D (2005) Foundations for engineering biology. Nature 438:449–453

Fehér T, Papp B, Pá C, Pósfai G (2007) Systematic genome reductions: theoretical and experimental approaches. Chem Rev 107:3498–3513

Forster AC, Church GM (2007) Synthetic biology projects in vitro. Genome Res 17:1–6

Forster AC, Church GM (2006) Towards synthesis of a minimal cell. Mol Systems Biol 2:45

Gil R, Silva FJ, Peretó J, Moya A (2004) Determination of the core of a minimal bacteria gene set. Microbiol Mol Biol Rev 68:518–537

Hadorn M, Burla P, Eggenberger Hotz P (2009) Towards tailored communication networks in assemblies of artificial cells. In: Korb KB, Randall M, Hendtlass T (eds) Artificial life: borrowing from biology, p 126–135

Hillebrecht JR, Chong S (2008) A comparative study of protein synthesis in in vitro systems: from the prokaryotic reconstituted to the eukaryotic extract-based. BMC Biotechnol 8:58

Hope MJ, Bally MB, Webb G, Cullis PR (1985) Production of large unilamellar vesicles by a rapid extrusion procedure. Characterization of size distribution, trapped volume and ability to maintain a membrane potential. Biochim Biophys Acta 812:55–65

Kajander EO (2006) Nanobacteria – propagating calcifying nanoparticles. Lett Appl Microbiol 42:549–552

Kajander EO, Ciftcioglu N (1998) Nanobacteria: an alternative mechanism for pathogenic intra- and extracellular calcification and stone formation. PNAS 95:8274–8279

Knoll A (ed) (1999) Size limits of very small microorganisms. National Academic Press, Washington DC

Kohno T, Endo Y (2007) Production of protein for nuclear magnetic resonance study using the wheat germ cell-free system. Methods Mol Biol 375:257–272

Lasic DD (1993) Liposomes: from Physics to applications. Elsevier, Amsterdam

Liu AP, Fletcher DA (2009) Biology under construction: in vitro reconstitution of cellular function. Nature Rev 10:644–650

Liu DV, Zawada JF, Swartz JR (2005) Streamlining Escherichia coli S30 extract preparation for economical cell-free protein synthesis. Biotechnol Prog 21:460–465

Lohse B, Pierre-Yves B, Dimitrios S (2008) Encapsulation efficiency measured on single small unilamellar vesicles. J Am Chem Soc 130:14372–14373

Luisi PL (2003) Autopoiesis: a review and a reappraisal. Naturwissenschaften, Springer 90:49–59

Luisi PL (in press) The synthetic approach in biology: epistemic notes for synthetic biology. Naturwissenschaften

Luisi PL (in press) Epistemology notes on synthetic biology, in: *Chemical Synthetic Biology*, C Chiarabelli and PL Luisi (eds.), Wiley, New York

Luisi PL, Varela FJ (1990) Self-replicating micelles – a chemical version of minimal autopoietic systems. Orig Life Evol Biosph 19:633–643

Luisi PL, Oberholzer T, Lazcano A (2002) The notion of a DNA minimal cell: a general discourse and some guidelines for an experimental approach. Helvetica Chimica Acta 85(6):1759–1777

Luisi PL (2006) The emergence of life: from chemical origin to synthetic biology. Cambridge University Press, Cambridge

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissensch 93:1–13

Luisi PL (2007) Chemical aspects of synthetic biology. Chem Biodivers 4:603–621

Luisi PL, Walde P, Oberholzer T (1999) Lipid vesicles as possible intermediates in the origin of life. Curr Opi Colloid Interface Sci 4:33–39

Luisi PL, Allegretti M, Souza TP, Steininger F, Fahr A, Stano P (2010) Spontaneous protein crowding in liposomes: a new vista for the origin of cellular metabolism. ChemBioChem DOI 10.1002/cbic.201000381

MacDonald RC, MacDonald RI, Menco BP, Takeshita K, Subbarao NK, Hu LR (1991) Small-volume extrusion apparatus for preparation of large, unilamellar vesicles. Biochim Biophys Acta 1061:297–303

Minton AP (2001) The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. J Biol Chem 276:10577–10580

Moore PB (1999) A biophysical chemist's thoughts on cell size. In: Size limits of very small microorganisms. National Academic Press, Washington DC, pp 16–20

Morange M (2009) A critical perspective on synthetic biology. HYLE Int J Philos Chem 15:21–30

Oberholzer T, Wick R, Luisi PL, Biebricher CK (1995) Enzymatic RNA replication in self- reproducing vesicles: an approach to a minimal cell. Biochem Biophys Res Comm 207:250–257

Pohorille A, Deamer D (2002) Artificial cells: prospects for biotechnology. Trends Biotech 20:123–128

Schwille P, Diez S (2009) Synthetic biology of minimal systems. Crit Rev Biochem Mol Biol 44:223–242

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755

Shimizu Y, Kanamori T, Ueda T (2005) Protein synthesis by PURE translation systems. Methods 36:299–304

Souza T, Stano P, Luisi PL (2009) The minimal size of liposome-based model cells brings about a remarkably enhanced entrapment and protein synthesis. ChemBioChem 10:1056–1063

Stano P, Kuruma Y, Souza TP, Luisi PL (2010) Biosynthesis of proteins inside liposomes. Meth Mol Biol 606:127–145

Sun B, Chiu D (2005) Determination of the encapsulation efficiency of individual vesicles using single-vesicle photolysis and confocal single-molecule detection. Anal Chem 77:2770–2776

Sunami T, Hosoda K, Suzuki H, Matsuura T, Yomo T (2010) Cellular compartment model for exploring the effect of the lipidic membrane on the kinetics of encapsulated biochemical reactions. Langmuir 26(11):8544–8551

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390

Walde P, Roger Wick, Fresta M, Mangone A, Luisi PL (1994) Autopoietic self-reproduction of fatty acid vesicles. JACS 116:11649–11654

Zhang Y, Ruder WC, LeDuc PR (2008) Artificial cells: building bioinspired systems using small-scale biology. Trends Biotech 26:14

Chapter 12
Liposomes Mediated Synthesis of Membrane Proteins

Yutetsu Kuruma

Abstract Synthetic biology is an emerging field that aims at constructing artificial biological systems by combining engineering and molecular biology approaches. One of the most ambitious research line concerns the socalled semi-synthetic minimal cells, which are liposome-based system capable of synthesizing the lipids within the liposome surface. This goal can be reached by reconstituting membrane proteins within liposomes and allow them to synthesize lipids. This approach, that can be defined as biochemical, was already reported by Schmidli et al. (in 1991). In more advanced models, however, a full reconstruction of the biochemical pathway requires (1) the synthesis of functional membrane enzymes inside liposomes, and (2) the local synthesis of lipids as catalyzed by the in situ synthesized enzymes. Here we show the synthesis and the activity – inside liposomes – of two membrane proteins involved in phospholipids biosynthesis pathway. The proteins, sn-glycerol-3-phosphate acyltransferase (GPAT) and lysophosphatidic acid acyltransferase (LPAAT), have been synthesized by using a totally reconstructed cell-free system (PURE system) encapsulated in liposomes. The activities of internally synthesized GPAT and LPAAT were confirmed by detecting the produced lysophosphatidic acid and phosphatidic acid, respectively. Through this procedure, we have implemented the first phase of a design aimed at synthesizing phospholipid membrane from liposome within – which corresponds to the autopoietic growth mechanism.

Keywords Liposomes • Membrane protein • Cell-free translation system • Lipid synthesis • Minimal cells

Y. Kuruma (✉)
Department of Medical Genome Sciences, Graduate School of Frontier Sciences,
The University of Tokyo, Kashiwanoha 5-1-5, 277-8562 Kashiwa, Chiba, Japan
e-mail: kuruma@molbio.t.u-tokyo.ac.jp

12.1 Introduction

Several research groups are currently attempting the construction of synthetic cells by forming each single component (Luisi 2007; Luisi et al. 2006). The vast amount scientific knowledge and the growing databases could provide a chance to construct whole cellular systems. The ongoing projects in Synthetic Cells research are related to various fields of science, from the conventional sciences (biology, chemistry, and mathematics) to the progressed sciences (biotechnology, polymer science, bioinformatics, etc.) (Fig. 12.1).

Attempts to construct Synthetic Cells with a minimal number of molecules are almost equal to understanding the ancient early cells. During the construction of the Minimal Cells, there are some important checkpoints that must be considered. To say that such a constructed cell is alive, the cells must satisfy at least three properties, i.e. self-maintenance, self-reproduction, and evolvability (Luisi et al. 2006). Although it is a fact that the present cells have realistically covered all of these three properties, the total construction of the Synthetic Cells containing such properties is still far from reality, at present. In Synthetic Biology, a constructive approach has attracted attention that aims to construct the Synthetic Cells by the use of an in vitro gene expression system inside liposome (Nomura et al. 2003; Yu et al. 2001). By performing protein synthesis within a limited compartment, such as a liposome, we can model the Synthetic Cells that mimic the living cells.

In this chapter, we present the current achievements about protein synthesis inside liposomes, especially the synthesis of functional membrane proteins inside liposomes.
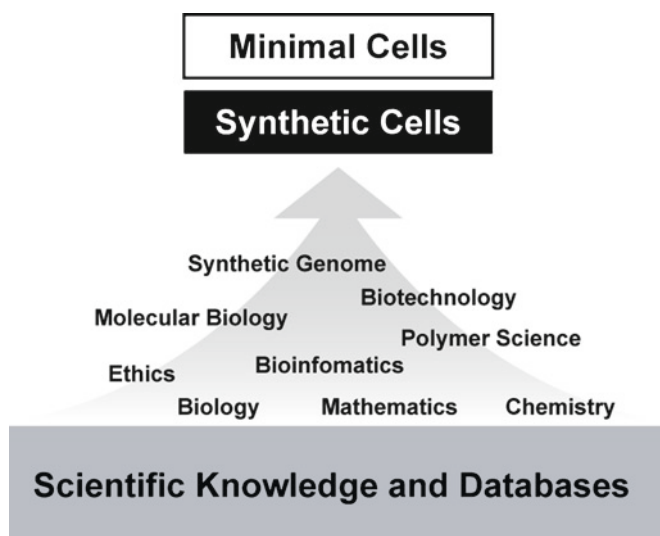
Fig. 12.1 Schematic representation of synthetic cells and minimal cells

12.2   Construction of Minimal Cells by Synthetic Approaches

12.2.1   Minimal Cells

In modern cells, cellular processes are conducted by highly complicated biochemical systems, involving the coordinated actions of several molecular agents. Such cellular systems are regarded as the consequences of a great number of evolutionary stages that have occurred over long periods of time. When we think about the very early stage of cells in evolution, such as the form of original cells, the question arises as to how they formed themselves. Did they have the complicated metabolic systems found in our modern cells, or did they have very simplified systems that could barely keep the cell alive? One of the common arguments for the question is that the early cells might have only contained indispensable molecules and genes (Luisi 2006), but they were not completely autotrophic cells. This is the fundamental concept of the Minimal Cell. Therefore, the definition of the number of really essential molecules and genes needed to keep the cell alive is an important task for the construction of Minimal Cells. Such attempts are almost equal to finding how much the number of genes within the modern cells can be reduced while keeping the cell alive.

12.2.2   Minimal Genome

A Minimal Genome is a theoretical gene set that consists of the minimal number of genes needed to keep the cell alive (Mushegian and Koonin 1996). Studies of the Minimal Genome have adopted two approaches, i.e. bottom-up and top-down. The bottom-up approach is exemplified by the work of Moya's group (Gil et al. 2004; Moya et al. 2009). They have tried to determine the minimal set of bacterial genes from the viewpoint of comparative genomics, by comparing several genomes within different species to find highly conserved genes. In the report by Gil et al., the minimal number of genes was defined as 206, which encode proteins (Gil et al. 2004). Although this defined number is hypothetical, what is interesting about this work is that the definition of the minimal genes is based on the definition of the minimal imperative metabolic systems. For example, almost the entire gene sets for transcription and translation, which are definitely core mechanisms for cellular life, are included in their list. Additionally, an energy metabolic system is considered. Since the defined number contains only protein-coding genes, the 206 genes may be too small of an estimation. In other words, a certain number of genes responsible for RNAs should be considered for the minimal genome. However, it is also true that the estimation of the core of a minimal bacterial gene set by Gil et al. has been regarded as a landmark in studies of the Minimal Genome.

In contrast, the top-down approach has been implemented by Venter's group (Hutchison et al. 1999; Glass et al. 2006). Their strategy is that, based on the genome of *Mycoplasma genitalium*, which has the smallest genome of any known organism,
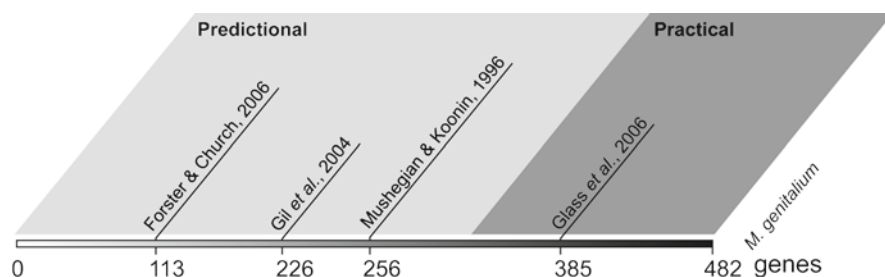
Fig. 12.2 Schematic of the studies on the minimal genome and their numbers. The predictional work and the practical work are highlighted in light gray and deep gray, respectively. The genome size of *M. genitalium* is represented as 482

they have deleted the genes by transposon mutagenesis (Hutchison et al. 1999). Based on such experimental results, the number of really essential genes was identified as 382, by disrupting the nonessential protein-coding genes from the 482 genes of the *M. genitalium* genome. According their conclusion, the 385 protein-coding genes (including three additional phosphate transporter genes) and the 43 RNA-coding genes are logically essential for the growth of the minimal cell (Glass et al. 2006).

The other predictions from minimal genome research are summarized in Fig. 12.2. From these studies, a realistic number of genes for the minimal genome would conceivably be between 250 and 350.

12.3 Protein Synthesis Inside Liposomes

12.3.1 Cell-Free Translation Systems

Protein synthesis inside liposomes can be performed by the encapsulation of a cell-free translation system within phospholipid-based liposomes. Several kinds of cell-free systems have been developed thus far, and most of them are commercially available from pharmaceutical companies (Table 12.1). Especially, *E. coli*- and wheat germ-based cell-free systems have been commonly employed in the studies of Synthetic Cells. In addition to these systems, there is a reconstructed cell-free translation system which called the PURE system (Shimizu et al. 2001). The PURE system contains a minimal set of enzymes that are essential for transcription and translation reactions. All of the enzymes are individually over-expressed in *E. coli* cells and are highly purified by metal affinity column chromatography. The purified enzymes are combined in a tube with precise regulation of each concentration.

Both types of cell-free systems, the cell-extract-based system and the PURE system, have their respective merits and pitfalls. The advantage of an extract-based system, such as *E. coli* or wheat germ, is that the productivity of the cell-free system is relatively high, and is generally maximal at 500 μg/ml (see Table 12.1). However, since the extract systems contain a large number of unrecognized molecules (perhaps also membrane crumbs), unexpected errors might occur at a stage of

Table 12.1 Types of cell-free translation system. The described yields are the maximal values of each system.

Name	Source	Yield	Company
S30 T7 High-Yield System	*E. coli* extract	500 μg/ml	Promega
EasyXpress System™	*E. coli* extract	500 μg/ml	QIAgen
Expressway Plus™	*E. coli* extract	300 μg/ml	Invitrogen
Rapid Translation System	*E. coli* extract	400 μg/ml	Roche
ENDEXT®	Wheat germ extract	5 mg/ml	Cell-Free Science
TnT (R) Coupled Wheat Germ Extract Systems	Wheat germ extract	10 mg/ml	Promega
TnT (R) Coupled Reticulocyte Lysate Systems	Rabbit reticulocyte extract	10 mg/ml	Promega
Transdirect™ insect cell	Insect cell extract	50 μg/ml	SHIMADZU BIOTECH
PURESYSTEM	Reconstructed system (from *E. coli*)	200 μg/ml	New England Biolabs

the activity measurement of the produced protein. The other problem is that the extract system strays far from the requirements of the Minimal Cells. On the other hand, the PURE system (a reconstructed cell-free system) perfectly matches the concept of Minimal Cells, since the system contains only the necessary factors for protein synthesis. Moreover, this property allows a practically nought background value in the activity measurement. The only weakness is the lower productivity, as compared with that of the extract system. Although the productivity of the PURE system is not very high, the PURE system has provided big advantages for the construction of Synthetic Cells and Minimal Cells.

12.3.2   *Liposomes*

Another important aspect for the inside synthesis is the lipid composition for liposome formation. Lipids possessing a phosphatidylcholine (PC) head are known to form very stable liposomes in aqueous solutions. Therefore, based on the PC lipid, several variations of the lipid composition have been explored for applicable liposomes in previous research. For instance, Yu et al. and Sunami et al. used DSPE-PEG5000 in the composition, in order to prevent liposomal fusion and disruption (Sunami et al. 2006). As another example, Kuruma et al. adopted the lipid composition of the *E. coli* inner cellular membrane within 1-palmitoyl-2-oleyl-phosphatidylcholine (POPC) liposomes (Kuruma et al. 2009). Although there are no standard lipid compositions for the inside synthesis so far, there are some fundamental properties that we must consider in liposome formation. First, the lipids must not interrupt the protein synthesis by the cell-free system. It is not clear which kinds of lipids are harmful for the cell-free system, but a certain amount of a cell membrane lipid extract (e.g. *E. coli* total lipid extract) inhibits protein synthesis. Second, the liposome must sustain a substantial internal space for encapsulation of the molecular

Table 12.2 Protein synthesis inside liposomes

Year	Protein	Cell-free system	Lipid(s)	Size (diameter)	References
2002	EGFP	E. coli extract	POPC	Unknown	Oberholzer et al. [25]
2003	rsGFP	Wheat germ extract	DOPC/DOPG (10:1)	5 μm	Nomura et al. [3]
2004	T7 RNA polymerase and GFP	E. coli extract	EYPC/cholesterol/ DSPE-PEG5000 (1.5:1.0:0.08)	400 nm	Ishikawa et al. [14]
2004	α-Hemolysin and GFP	E. coli extract	Egg lecithin	15-30 μm	Noireaux et al. [15]
2006	GFP	PURE system	POPC/PLPC/SOPC/SLPC/ cholesterol/DSPE-Peg5000 (129:67:48:24:180:14)	400 nm	Sunami et al. [12]
2008	β-Subunit of Qβ replicase and β-Galactosidase	PURE system	POPC/cholesterol/DSPE-PEG5000 (58:39:3)	400 nm	Kita et al. [26]
2008	Glycerol-3-phosphate acyltransferase and lysophosphatidic acid acyltransferase	PURE system	POPC/POPE/POPG/cardiolipin (50.8:35.6:11.5:2.1)	400 nm	Kuruma et al. [13]

Abbreviations: GFP, green fluorescent protein; EGFP, enhanced GFP; rsGFP, red-shifted GFP; DHFR, dihydrofolate reductase; DOPC, dioleoylphosphatidyl-choline; DOPG, dioleoylphosphatidylglycerol; EYPC, egg yolk phosphatidylcholine; POPC, 1-palmitoyl-2-oleoyl-phosphatidylcholine; PLPC, 1-palmitoyl-2-linoleoyl-phosphatidylcholine; SOPC, 1-stearoyl-2-oleoyl-phosphatidylcholine; SLPC, 1-stearoyl-2-linoleoyl-phosphatidylcholine; POPE, 1-palmitoyl-2-oleoyl-phosphatidylethanolamine; POPG, 1-palmitoyl-2-oleoyl-phosphatidylglycerol.

components of the cell-free system. Third, the lamellarity of the lipid bilayer of the liposome is also important; ideally, it should be unilamellar. Additionally, when an integral membrane protein is synthesized inside the liposome, the lipid composition fulfilling the above three conditions must contribute to the activity of the membrane protein, by providing a good environment on the lipid bilayer.

In summary, the coordination of the enclosed cell-free system and the outer-shell liposome is important to achieve internal protein synthesis.

12.3.3   Model Proteins

In the past few years, some research groups have demonstrated internal protein synthesis (Table 12.2). The most general model protein is green fluorescent protein (GFP) and its derivatives (red-shifted GFP (rsGFP) and enhanced GFP (EGFP)), because of the ease of observation by microscopy or fluorometric analysis (Nomura et al. 2003; Ishikawa et al. 2004). By taking advantage of GFP, a couple of chimeric proteins have been genetically constructed and utilized in studies of inside synthesis. Noireaux and Libchaber have constructed a chimeric protein of α-hemolysin and EGFP (Noireaux and Libchaber 2004). Using this construct, they performed protein synthesis inside a giant liposome and, furthermore, they observed the membrane localization of the synthesized protein by means of the nature of α-hemolysin, which forms a nano-sized pore on the lipid bilayer of the liposome. Nomura et al. conjugated apo cytochrome b_5 (b5)-GFP and b5-dihydrofolate reductase (DHFR), and detected membrane localization of the synthesized b5-GFP (Nomura et al. 2008). Moreover, the enzymatic activity of the b5-DHFR, which was localized at the liposome membrane, was also detected. This particular study was not a case of inside synthesis; however, their results suggested further applications for inside synthesis.

12.4   Liposome-Mediated Membrane Protein Synthesis

There are only a few cases of membrane protein synthesis inside liposomes thus far. In contrast to a soluble protein, like GFP, membrane proteins possess hydrophobic fragments in their structures, and thus easily aggregate in an aqueous solution. For this reason, the presence of lipid (or detergent) during the translation reaction is essential for the synthesis of an active membrane protein. Membrane proteins play many important roles in the cell, and some of them are considered to be necessary for the Minimal Cells (Gil et al. 2004); e.g., a membrane channel or transporter, to provide the capability of membrane permeability. Through such membrane proteins, the Minimal Cells can exchange various molecules between the inside and outside of the liposome and achieve communication with the exterior environment, as a semi-open system. This is a very important affair from the viewpoint of nutrient intake, which is eventually connected to the energy cycle within the cell to maintain self-maintenance.

The following section introduces the inside synthesis of α-hemolysin, which can form a pore on the liposome membrane. Additionally, a unique study that demonstrates the synthesis of an integral membrane protein and a periplasmic protein (hypothetical) inside liposomes is introduced.

12.4.1   α-Hemolysin

Noireaux and Libchaber successfully synthesized α-hemolysin (αHL) pore protein inside lecithin-based liposomes prepared as giant unilamellar vesicles (GUVs), which have a size of 5–10 μm (Noireaux and Libchaber 2004). αHL is a bacterial toxin that can form a 1.5 nm size pore on the cell membrane. The pore possesses selective permeability by size, which enables the passage of even a single stranded polynucleotide chain (Kawano et al. 2009). Therefore, the idea of the study is that, by synthesizing the αHL inside the liposome and forming the selective pores on the membrane, the liposomes can assimilate energy and substrates from the outside environment, to prolong the internal protein synthesis reaction. In fact, the lifetime of the internal protein synthesis was prolonged by more than 4 days due to the synthesized αHL conjugated eGFP, whereas the internal synthesis of eGFP alone was observed for only a few hours (Fig. 12.3). According their calculations, the synthesized pores
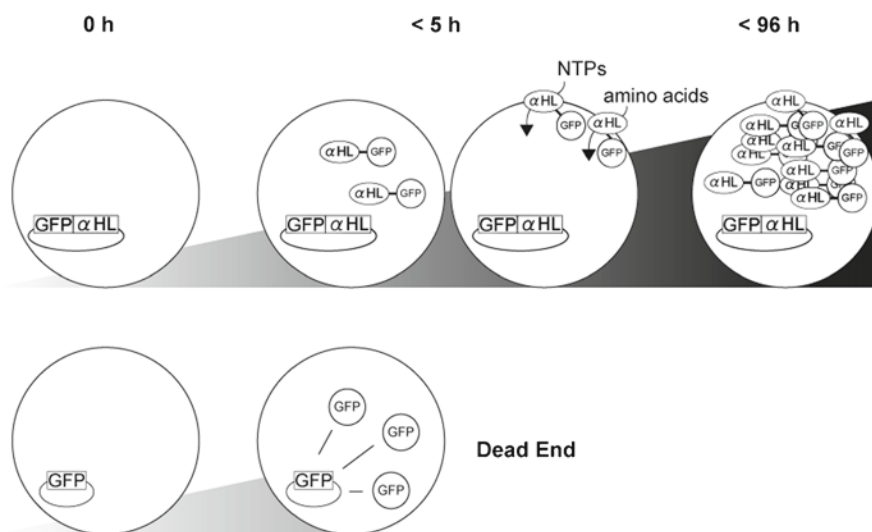
Fig. 12.3 Schematic of long-term protein synthesis inside the liposome. Alpha hemolysin pore protein (αHL), conjugated with enhanced GFP (eGFP), was synthesized inside lecithin-based vesicles. The αHL synthesizing vesicles can form size-limiting pores on the vesicle membrane that are suitable for NTP and amino acid uptake, and thus prolong the life-time of internal synthesis up to 96 h (*upper*). In contrast, the vesicles synthesizing only eGFP reach the limitation of the internal substrates and energy within 5 h (*lower*), and protein production stops

occupied 20% of the liposome membrane at the ten timepoint, and that increased to 75% of the membrane after 4 days.

The other important point is that the spontaneous membrane localization of the αHL conjugated GFP was occurring inside the liposomes. This fact indicates that, during the synthesis, a certain amount of the synthesized membrane protein is able to become localized onto the membrane and to retain the functional form, whereas the membrane insertion of a protein is well-regulated by responsible cytosol proteins in bacteria; i.e. signal recognition particle (SRP) and its receptor (SR) (Kuruma et al. 2005).

12.4.2 Membrane Enzymes Involved in Lipid Biosynthesis

Although αHL was successfully synthesized inside the liposome, the αHL pore protein cannot be considered as an authentic membrane protein, because the protein can be soluble in aqueous solution (Jayasinghe et al. 2006). Generally, the solubility of a membrane protein (especially multi-spanning membrane proteins) is extremely low in an aqueous solution. For instance, when a membrane protein was synthesized in the PURE system that deletes chaperons or detergents, large amount of the products were collected as precipitates after centrifugation, whereas the soluble proteins were present in the supernatants (Kuruma et al. 2005).

An example of a membrane protein can be found in the work of Kuruma et al., who performed the internal synthesis of two membrane proteins within liposomes encapsulating the PURE system and the responsible genes. What is interesting about these proteins is that they are involved in a phospholipid biosynthesis pathway (Cronan 2003), i.e. sn-glycerol-3-phosphate acyltransferase (GPAT) and lysophosphatidic acid acyltransferase (LPAAT). GPAT is an integral membrane protein possessing multi-spanning fragments. LPAAT is a hypothetical periplasmic membrane protein anchored on the cellular membrane. GPAT catalyzes the synthesis of lysophosphatidic acid (LPA) from sn-glycerol-3-phosphate (G3P) and acyl-CoA, and LPAAT catalyzes phosphatidic acid (PA) synthesis from LPA and acyl-CoA (Fig. 12.4). Therefore, the aim of this research was to construct lipid-synthesizing liposomes that depended on the activities of the internally synthesized GPAT and LPAAT. The accumulation of the new lipid within the membrane will stimulate the liposome self-growth and, eventually, self-division. The final goal of this research is the construction of a self-reproducible Minimal Cell. In the report by Kuruma et al., the internal protein synthesis of GPAT and LPAAT was accomplished by the encapsulated PURE system from each corresponding gene (Fig. 12.4). After the protein synthesis, the activity assays for the GPAT and LPAAT were further analyzed by encapsulating their substrates. The acyltransferase activity of the internal GPAT was evaluated by observation of the resulting LPA product. Although a low percentage of the synthesized GPAT enzyme was thought to be active, this is the first example where an integral membrane protein was synthesized inside liposomes and its enzymatic activity was detected only in the interior of the liposomes. Another membrane enzyme, LPAAT, was also
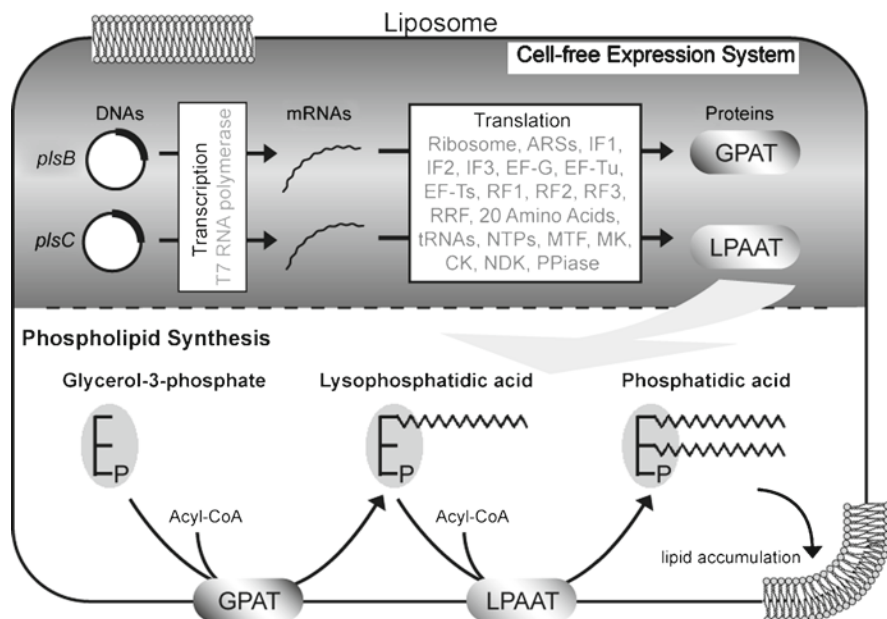
Fig. 12.4 Illustration of the lipid synthesizing liposome. The phospholipid-based liposome contains a reconstructed cell-free system consisting of T7 RNA polymerase and translation factors, in order to synthesize GPAT and LPAAT from the corresponding genes plsB and plsC, respectively. The synthesized GPAT and LPAAT are integrated into the liposome membrane and thus gain their enzymatic activities. Based on the integrated GPAT and LPAAT, new phospholipids are produced from glycerol-3-phosphate and acyl-CoA. Finally, the newly synthesized phospholipids are incorporated in the lipid bilayer of the liposome and stimulate liposome self-division, which is regarded as a prototype of a self-reproduction system in the Minimal Cell

analyzed for its activity. The activity of the synthesized LPAAT was initiated by the encapsulation of the LPA and acyl-CoA substrates into the liposomes, and was detected by observation of the resulting PA product. Based on the fact that LPAAT is not functional in the presence of a reducing agent, such as β-mercaptoethanol, LPAAT is thought to require some disulfide bounds in its active structure. This is a typical property of a periplasmic protein. To solve this problem, oxidized glutathione was added to the system after the quenching of protein synthesis, which requires reductive conditions. Again, the activity of the internal LPAAT was confirmed by detecting the resulting PA product.

12.5 Future Developments

As shown by the aforementioned experiments, membrane protein synthesis is feasible by a cell-free translation system inside liposomes. However, the membrane localization of the proteins thus produced relies on spontaneous membrane integration, which

is a hydrophobic interaction between the membrane protein and the lipid bilayer. In order to regulate membrane targeting and integration in the present systems, we must create a membrane integration or translocation system for the efficient production of active membrane proteins.

In the cell, most of the integral membrane proteins are biosynthesized by coupling with translation (co-translational) and integration via the Sec translocon (Luirink et al. 2005). Although the Sec translocon has several components (Gold et al. 2007), the minimum set from the bacterial system is considered to be SecY and SecE for membrane integration (Tsukazaki et al. 2008). Therefore, to reconstruct the Sec machinery within liposomes, only the SecYE translocon proteins and a suitable lipid composition, which contributes to the proper conformation of the translocon, are theoretically needed. With coupling to the Sec translocon, two more cytosolic factors (SRP and SR, see Section 12.4.1) are effective to guide a nascent peptide during the translation reaction (Kuruma et al. 2005). In addition to these components, Nishiyama et al. suggested the importance of factor X, which assists in the membrane integration of proteins (Kawashima et al. 2008). By developing the complete Sec machinery, all of the synthesized membrane proteins can be integrated via the Sec translocon in a co-translational manner, with the proper membrane topology needed for their functions.

12.6   Conclusions and Remarks

The challenges of constructing Synthetic Cells (or Minimal Cells) are now less daunting, thanks to the cell-free translation systems and liposomes. For the scenario of the Minimal Cell, the cell must have the smallest set of genes, and this policy requires that the cells contain sequential reactions of transcription and translation. The proteins thus produced must be functional inside the compartments (Nomura et al. 2003) and, moreover, they must facilitate more complex biochemical reactions. In order to develop such advanced systems, the liposome membrane must possess selective permeability to obtain materials (such as energy and substrates) from the outside and thus allow the constant internal reactions to be achieved. In this sense, the pioneering work of Noireaux and Libchaber (Noireaux and Libchaber 2004) is a major breakthrough, from the viewpoints of both the prolongation of the internal protein synthesis reaction and the generation of membrane proteins on the liposome membrane.

The next challenge for the inside protein synthesis is the generation of large molecular complexes on the liposome membrane. Some transporters or channels are composed of a number of integral membrane proteins. For example, ATP synthase contains three different integral membrane proteins and five different cytosolic proteins. Although the biosynthetic pathway of the ATP synthase complex has not been fully elucidated, a number of in vitro studies might provide a foundation to construct ATP synthase within liposomes. In other words, by using a cell-free system and liposomes, we can develop energy-producing liposomes as an autonomous system.

References

Cronan JE (2003) Bacterial membrane lipids: where do we stand? Annu Rev Microbiol 57:203–224

Gil R, Silva FJ, Peretó J, Moya A (2004) Determination of the core of a minimal bacterial gene set. Microbiol Mol Biol Rev 68:518–537

Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA 3rd, Smith HO, Venter JC (2006) Essential genes of a minimal bacterium. Proc Natl Acad Sci USA 103:425–430

Gold VA, Duong F, Collinson I (2007) Structure and function of the bacterial Sec translocon. Mol Membr Biol 24:387–394

Hutchison CA, Peterson SN, Gill SR, Cline RT, White O, Fraser CM, Smith HO, Venter JC (1999) Global transposon mutagenesis and a minimal Mycoplasma genome. Science 286:2165–2169

Ishikawa K, Sato K, Shima Y, Urabe I, Yomo T (2004) Expression of a cascading genetic network within liposomes. FEBS Lett 576:387–390

Jayasinghe L, Miles G, Bayley H (2006) Role of the amino latch of staphylococcal alpha-hemolysin in pore formation: a co-operative interaction between the N terminus and position 217. J Biol Chem 281:2195–2204

Kawano R, Schibel AE, Cauley C, White HS (2009) Controlling the translocation of single-stranded DNA through alpha-hemolysin ion channels using viscosity. Langmuir 25:1233–1237

Kawashima Y, Miyazaki E, Müller M, Tokuda H, Nishiyama K (2008) Diacylglycerol specifically blocks spontaneous integration of membrane proteins and allows detection of a factor-assisted integration. J Biol Chem 283:24489–24496

Kita H, Matsuura T, Sunami T, Hosoda K, Ichihashi N, Tsukada K, Urabe I, Yomo T (2008) Replication of genetic information with self-encoded replicase in liposomes. Chembiochem 9:2403–2410

Kuruma Y, Stano P, Ueda T, Luisi PL (2009) A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. Biochim Biophys Acta 1788:567–574

Kuruma Y, Nishiyama K, Shimizu Y, Müller M, Ueda T (2005) Development of a minimal cell-free translation system for the synthesis of presecretory and integral membrane proteins. Biotechnol Prog 21:1243–1251

Luirink J, von Heijne G, Houben E, de Gier JW (2005) Biogenesis of inner membrane proteins in Escherichia coli. Annu Rev Microbiol 59:329–355

Luisi PL (2006) The emergence of life. From chemical origins to synthetic biology. Cambridge University Press, Cambridge

Luisi PL (2007) Chemical aspects of synthetic biology. Chem Biodivers 4:603–621

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells. Naturwissenschaften 93:1–13

Mushegian AR, Koonin EV (1996) A minimal gene set for cellular life derived by comparison of complete bacterial genomes. Proc Natl Acad Sci USA 93:10268–10273

Moya A, Gil R, Latorre A, Peretó J, Pilar Garcillán-Barcia M, de la Cruz F (2009) Toward minimal bacterial cells: evolution vs. design. FEMS Microbiol Rev 33:225–235

Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. Chembiochem 4:1172–1175

Nomura SM, Kondoh S, Asayama W, Asada A, Nishikawa S, Akiyoshi K (2008) Direct preparation of giant proteo-liposomes by in vitro membrane protein synthesis. J Biotechnol 133:190–195

Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. Proc Natl Acad Sci USA 101:17669–17674

Oberholzer T, Luisi PL (2002) The use of liposomes for constructing cell models. J Biol Phys 28:733–744

Schmidli PK, Schurtenberger P, Luisi PL (1991) Liposome-mediated enzymatic synthesis of phosphatidylcholine as an approach to self-replicating liposomes. J Am Chem Soc 113:8127–8130

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755

Sunami T, Sato K, Matsuura T, Tsukada K, Urabe I, Yomo T (2006) Femtoliter compartment in liposomes for in vitro selection of proteins. Anal Biochem 357:128–136

Tsukazaki T, Mori H, Fukai S, Ishitani R, Mori T, Dohmae N, Perederina A, Sugita Y, Vassylyev DG, Ito K, Nureki O (2008) Conformational transition of Sec machinery inferred from bacterial SecYE structures. Nature 455:988–991

Yu W, Sato K, Wakabayashi M, Nakaishi T, Ko-Mitamura EP, Shima Y, Urabe I, Yomo T (2001) Synthesis of functional protein in liposome. J Biosci Bioeng 92:590–593

Chapter 13
Giant Unilamellar Vesicles: From Minimal Membrane Systems to Minimal Cells?

Petra Schwille

Abstract Giant unilamellar vesicles have been studied by membrane and lipid researchers for more than two decades. With their comfortable dimensions between single and hundreds of micrometers that are easily accessible to optical imaging and manipulation techniques, they have been proven ideal model systems to study membrane morphology and mechanical parameters, such as surface tension, elasticity, and local curvature, relevant for membrane structure and transformations. Moreover, since the advent of the raft hypothesis (Simons and Ikonen 1997) in cell biology, there has also been rising interest from the biological community to better understand the relevance of local lipid order for the lateral sorting and induction of functionality of membrane proteins. It is now widely accepted that the quantitative representation and local order of specific lipids in membranes of various organelles, in tight concert with the respective proteins inserted or attached to them, accounts to a large extent for their functionality. However, since the exact relationships and also the structural features in live cells are often too complex or too small to be resolved quantitatively, minimal systems with reduced complexity, such as GUVs, pave the way to a more fundamental understanding of lipid-lipid and lipid-protein interactions of physiological importance. Inspired by the success of these minimal systems approaches to cell biological phenomena, many researchers nurture strong hopes that such a bottom-up approach does not stop at the membrane or at membrane-related processes, but that the GUV model system can be worked into more elaborate models of biomolecular self-organization. Next obvious steps would be to include factors that are able to controllably induce transformations, such as division, of these cell-like compartments, and to then combine the transformable compartments with information units that could be reproduced during division. In other words, to convert the GUV membrane shell into a minimal system for cellular reproduction. In this article,

P. Schwille (✉)
Biophysics/BIOTEC, TU Dresden, Tatzberg 47-51, D-01307 Dresden,
and
Max-Planck-Institute for Molecular Cell Biology and Genetics, Pfotenhauerstr. 108,
D-01307 Dresden
e-mail: schwille@biotec.tu-dresden.de

I will briefly discuss the past accomplishments and the future perspectives of the GUV model system, highlight its specific virtues, but also mention its obvious shortcomings with respect to a GUV-based synthetic biology.

13.1 Short History of the GUV Model System

A vesicle is the biological representation of a liposome, i.e., in principle nothing else than a bubble made up from a double layer of amphiphilic lipid molecules, which usually forms spontaneously under specific conditions, due to their tendency to self-organize. In the zoo of lipid vesicles, ranging from several 10 nm to hundreds of micrometers, GUVs (giant unilamellar vesicles) are, as the name already indicates, the largest representatives. They are very fragile entities, as in the same way as for toy soap bubbles, the largest are the hardest to make and to keep intact over time. However, they display the essential advantage over all other vesicles that they can be observed in real time by light microscopy without having to destroy them, providing access to features like shape and shape dynamics, as well as large-scale lipid organization within their membranes. Due to their gigantic sizes, the membranes of giant vesicles can be assumed to be almost flat with vanishing curvature.

Consequently, GUVs have in the past years become extremely attractive systems in the study of model membranes, particularly because of their mentioned compatibility with optical methods such as light and fluorescence microscopy, and because of the large unperturbed areas of free-standing membranes, void of any support, which they provide. Quite early, they were frequently employed in the biophysical investigation of membrane surface and shape transformations under various external and lipidic conditions, following the theoretical models of fluid crystal mechanics (Helfrich 1973; Lipowsk 1991; Lipowsky and Sackmann 1995; Seifert 1997).

Later, following the release of easily reproducible protocols for their formation (Angelova et al. 1992; Akashi et al. 1996), they were proposed and demonstrated to be perfect platforms for the investigation of lipid domain or phase formation in binary or ternary lipid mixtures (Korlach et al. 1999; Bagatolli and Gratton 2000; Baumgart et al. 2003; Veatch and Keller (2003a,b)). Protein insertion into the bilayer was successfully accomplished (Kahya et al. 2001; Girard et al. 2004) to better understand the mutual relationship between lipids and local lipid structure, and protein function. Following the tracks of earlier mechanical studies, GUVs have in recent applications been subject to mechanical manipulations by microcapillaries and optical tweezers, membrane tubes being pulled (Roux et al. 2002; Koster et al. 2003), or large tube-vesicle networks being assembled from them (Davidson et al. 2003), and they are even amenable to patch clamping. Due to the easy accessibility of their surfaces, they are used as templates for membrane sculpting or membrane transforming proteins (Roux et al. 2006; Römer et al. 2007; Wollert et al. 2009), as discussed below.

The breakthrough of the GUV model system in biology came shortly after the acknowledgment of the potential biological relevance of detergent resistant membranes (DRMs), or, as their more popular connotation, so-called lipid rafts

(Simons and Ikonen 1997; Simons and van Meer 1988). The underlying concept is that the heterogeneous lipid distribution in cellular membranes has functional relevance in various biological processes. Under certain conditions, microdomains with different physical properties form within the membrane and act as platforms for the sorting, enrichment, and activation of membrane-associated proteins. Being enriched in sphingomyelin and cholesterol, these domains proved for a long time impossible to be observed directly in live cells, due to their presumed sub-resolution sizes of less then 100 nm (Day and Kenworthy 2009). Although different indirect evidence for local membrane organization and its relevance for protein activity has been accumulated (Varma and Mayor 1998; Rajendran and Simons 2005; Wawrezinieck et al. 2005), only the advent of super-resolution microscopy seems to provide first direct experimental proof for their – at least transient – existence in live cellular membranes (Eggeling et al. 2009).

13.2 How to Make GUVs?

Although lipids are prone to self-organization into micelles and liposomes, depending on the size and shape of the respective lipids, the creation of giant vesicles is not quite as straightforward. GUVs can in principle be generated by coalescence and fusion of large unilamellar vesicles of several hundred nanometers size, e.g. formed by extrusion, but this process takes up to days to reach acceptable yields, competing with degeneration processes, and the reachable sizes are also limited to up to around 10 μm. Thus, most of the current protocols used for production of GUVs with high yield refer back to the seminal work by Akashi et al. 1996, and Angelova et al. 1992. The underlying principle is to dehydrate a lipid film on a surface and then slowly rehydrate using different geometries and external conditions. Most prominent is the rehydration of the lipids in the presence of a slowly alternating electrical field, either between two adjacent platinum wires, or sandwiched between two metal-coated coverslips, usually using ITO as the metal substrate. However, many recent protocols, particularly when using more physiological conditions including salt, prefer the so called "gentle hydration" or "natural swelling", where no electric fields are applied. The main drawbacks are the lower quality of the GUVs, and that the procedure takes very long. I will briefly comment on two variants of the electroformation method in more detail, because this is definitely the simplest way to obtain nice GUVs with reliable yield.

13.2.1 Electroswelling on Wires

Required for preparation are lipid stock solutions made by dissolving lyophilized lipids in chloroform or a mixture of chloroform and methanol, as well as, for the hardware, home-made Teflon and Pt chambers, heating devices, a Hamilton syringe, a function

generator, connecting clamps and cables, and observation chambers, e.g., Lab-Tek chambered cover glasses. First, the Pt chamber is cleaned with water, ethanol and chloroform dried. Then, the lipid solution is spread on each of the Pt wires. The wires are put under vacuum at room temperature for ca. 2 h to allow complete evaporation of the solvents. A sucrose solution of the desired osmolarity is prepared, and a small fraction of the solution warmed up above the transition temperature of the lipid mixture under study. The electroformation chamber is filled with warmed sucrose solution and closed. The chamber is placed in a heating block above the transition temperature of the lipid mix, and the cables of the function generator are connected to each of the two Pt-wires, avoiding contact between them. The function generator is operated at a sinus wave with a voltage that depends on the distance of the electrodes (ca. 2 V RMS) and a frequency of 10 Hz. Electroformation under these conditions is allowed for ca. 1.5 h. Then the frequency is decreased to 2 Hz for around 15 min, so that the formed GUVs detach from the Pt wires. Together with glucose or buffer solution matching the sucrose solution used to prepare the GUVs in osmolarity, they are gently removed from the formation device by a tip-cut pipette filled in a (Lab-Tek) observation chamber, and allowed to sediment for around 5 min.

13.2.2   Electroswelling Between ITO-Coated Coverslips

The preparation steps are very similar to those described above, with the major difference in the geometry of the formation chamber, which is now an arrangement of two ITO-coated coverslips opposing each other in a (usually home-built) Teflon chamber. ITO coverslips can be fabricated by various facilities and companies, e.g., in our case, GeSiM, (Groβerkmannsdorf, Germany). Between them, there is an acrylic spacer, and polyethylene tubing is used matching the size of the inlets. To attach the wires, adhesive copper tape is attached to the conductive side of the coverslips. Clamps and cables are connected to the function generator as described above. For GUV preparation, the ITO-coverslips are cleaned with water, ethanol and chloroform and dried, and then the stocks of desired lipid mixtures are added to pre-heated coverslips above the transition temperature of the lipid mixture, such that a lipid film is formed after solvent evaporation. The ITO coverslips are put under vacuum for around 2 h to allow complete evaporation of the solvents. Then, the ITO-coverslips are assembled to each side of the plastic spacer, taking into account that the conductive side of the coverslips must face the inside of the spacer, and that the lipids are positioned at the bottom of the chamber. The assembled chamber is heated above the transition temperature in an oven or a heating block. A sucrose solution of the desired osmolarity is prepared, warmed up at the same temperature as the electroformation chamber and degased by sonication in a bath. The solution is injected through the bottom inlet of the chamber avoiding the presence of bubbles inside the chamber. The chamber is placed in an oven or a heating block above the transition temperature of the lipid mixture, and the copper strips are connected to the function generator. The function generator is operated at a

sinus wave with a voltage of 1.4 V (RMS) and a frequency of 10 Hz for around
1.5 h. After turning it off, the chamber is allowed equilibrate to room temperature
before GUVs can be extracted as explained above.

13.2.3   GUV Production at Physiological Conditions

When proteins are to be inserted into the GUV membranes or to the vesicle interior,
as described in many applications below, electroformation becomes more critical.
Even if purified proteins withstand the treatment of dehydration and rehydration,
the ionic (in particular, salt) conditions of the buffer that allows vesicle electrofor-
mation in slowly alternating fields are usually considerably below physiological
concentrations. Recent publications (Pott et al. 2008; Montes et al. 2007) report
successful strategies how to arrive at higher salt concentrations. This can be done
by using the natural swelling method, but also employing electroformation. They
show that deposits for dehydration on the electrodes do not have to be made from
lipid solutions in organic solvents, but can also be aqueous, which then better sup-
port electroformation. However, these protocols are still not generally compatible
with complex membrane protein reconstitution and have to be adapted for every
specific membrane and protein preparation in a "trial and error" procedure.

13.2.4   Reverse Emulsion

Unlike the methods discussed above that are all variations of the common theme
of dehydration-rehydration, creating GUVs by reverse emulsion (Pautot et al.
2003) has a different underlying principle, which makes it a true alternative in
many ways. In particular, the aspect under which it was originally advertised, i.e.
generation of asymmetric bilayers with different leaflet compositions, deserves
attention. On the other hand, to date, there have not been too many success stories
reported using this technique (Noireaux and Libchaber 2004; Pontani et al. 2009),
which can be a hint on some underlying problems regarding reproducibility and
yield. The idea is to first create water droplets in lipid-saturated oil, and then to
force those droplets through an oil-water interface, in order to create a true lipid
bilayer. The fluids are arranged in the way that the emulsion phase is placed on a
so-called intermediate phase of the same oil, containing the lipids for the outer
leaflet, and an aqueous phase at the bottom. The intermediate phase allows con-
trolling the composition of each leaflet independently. There are many ways to
realize droplet generation, and to make the droplets cross the interface. The first
and most prominent realization was to do this by centrifugation. In the meantime,
other more technically advanced ways have been provided by various groups,
including piezo pumping of water droplets through interfaces, similar to the
approach described below.

13.2.5   Jetting

The idea is to concert the principle of blowing soap bubbles to the aqueous environment. This is most conveniently realized in microfluidic flow chambers (Stachowiak et al. 2008), where a precisely piezo-controlled pulsatile liquid jet is directed into a unilamellar lipid bilayer, established at a small orifice. For every pulse, a single vesicle is formed from the bilayer by a small water droplet forced through the interface. This has the advantage that GUVs can be produced and filled at the same time, and also yields amazingly monodisperse size distributions. The average diameters created by this technique were between 40 and 200 μm, quite convenient for studying them by optical techniques. So far, small particles as well as the pore protein α-hemolysin were included in the fluid phase, the latter allowing the transport of dye molecules across the membrane. This formation method might prove to be the most attractive for applications of synthetic biology with GUVs, no follow-up reports using this technique have so far been published.

13.3   GUVs with Membrane Domains: A Success Story

It has long been known from membrane thermodynamics that lipids can form different order structures, dependent on the length and assembly of their tails and head groups, which determine the Van der Waals interactions between adjacent lipid molecules. These order structures fulfil the thermodynamic requirements to call them phases; in particular they display discontinuous changes of enthalpy and entropy when measured in a calorimeter. The most prominent representations of phases in membranes were the so-called fluid (or liquid disordered) phase, characterized by a relatively low order in the tail region, with correspondingly high lateral mobility of the lipids, and the gel phase, displaying high order in the tail region and low mobility of the lipids. However, it has long been suspected that intermediate phases, such as the so-called liquid-ordered phase could exist. The liquid ordered phase is understood as a phase with higher order in the tail region of the lipids, but nevertheless rather large diffusional mobility. However, full attention was brought to this special state of membrane order when Simons and coworkers first postulated the existence of functional lipid domains, termed lipid rafts, in the cell membrane, which, at present knowledge, seem to exhibit many properties of what one considers a liquid ordered state of the membrane. These lipid rafts were at first indirectly postulated, through the existence of detergent resistant membranes (DRMs) withstanding even harsh purification steps to extract certain membrane proteins.

Because sphingomyelin and cholesterol are enriched in DRMs, they were quickly proposed to be essential ingredients of lipid rafts. Consequently, to better understand the thermodynamic properties of membranes enriched in these two native membrane constituents, GUVs were created composed of what is today called the "canonical raft mixture", i.e. unsaturated PC, sphingomyelin, and cholesterol in about equal
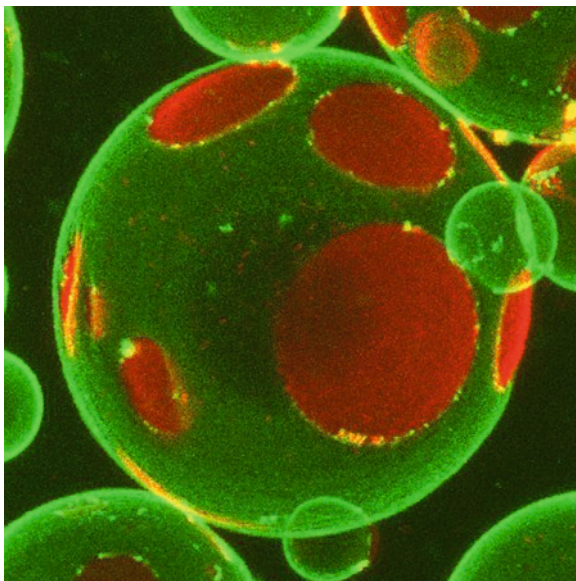
Fig. 13.1 Phase-separating GUVs composed of DOPC/SM/Chol (1:1:1) and labeled with 0.1% DiD (*red*), EqtII-V22CAl488 (*green*). The pore-forming toxin EqtII-V22CAl488 largely prefers the liquid ordered phase (Schön et al. 2008)

amounts (Dietrich et al. 2001; Kahya et al. 2003;) Veatch and Keller (2003a). In GUVs as minimal models for domain-forming membranes (Fig. 13.1), the co-existence of a so-called liquid disordered, or fluid, membrane phase, and a more viscous liquid ordered phase, could be easily demonstrated by a combination of optical imaging and fluorescence correlation spectroscopy (FCS), using dyes with different partition coefficients in the different membrane environments (Bacia et al. 2004). Spreading vesicles of the same lipid mixture on solid supports such as mica, AFM could at the same time reveal that the so-called liquid-ordered phase, enriched in sphingomyelin and cholesterol, was about 1 nm thicker than the fluid phase, which corresponds well with the assumption of higher order of the carbohydrate tails in these domains (Saslowsky et al. 2002; Chiantia et al. 2006).

Although much fundamental membrane thermodynamics could be performed with the GUV model system, like mapping exact phase diagrams of specific ternary lipid mixtures, including different sterols (Scherfeld et al. 2003, Veatch and Keller 2003; Feigenson 2006; Bacia et al. 2005), one of its most attractive aspects is the possibility of quantitatively investigating the relevance of lipid composition and local lipid order on protein functionality, both for membrane-associated proteins carrying specific linkers, but also for true transmembrane proteins. Not only the raft-association or partitioning of the proteins into a specific membrane domain can be tested, but also the presumed mutual relationship between the functional state of a protein induced by, e.g., aggregation, conformational change, or chemical modification, and its respective lipid environment. It is well conceivable that the functional

state determines the partitioning of the protein into either the liquid ordered or the liquid disordered phase. At the same time, the specific protein function may determine the local membrane order around it. Both effects could be demonstrated in the GUV model system, after a number of protocols have been released how to efficiently transfer functional proteins into GUV membranes (Kahya et al. 2001; Bacia et al. 2004; Girard et al. 2004; Kalvodova et al. 2005). An exciting outcome of these studies was that the primary association of many raft-associated proteins with liquid ordered domains was not necessarily evident, however, upon activation or clustering, the affinity towards the raft phase became prevalent. Hammond et al. (Hammond et al. 2005) demonstrated that the clustering of GM1 in membranes upon binding of cholera toxin even induced domain formation, and the re-sorting of a transmembrane peptide. In a similar assay with more complex membrane vesicles derived from full cell membranes, Lingwood et al. (Lingwood et al. 2008) showed that a considerable fraction of lipid was actually following this re-sorting, being drawn to a more ordered state by the aggregation of protein around it. On the other hand, proteins with membrane anchors such as GPI, or that specifically bind to lipids enriched in rafts (like Cholera toxin and Equinatoxin II), could indeed be proven to preferentially segregate into the liquid ordered domains (Chiantia et al. 2008; Schön et al. 2008).

13.4 GUVs Being Transformed by Proteins

In the previous section, it was shown that GUVs can indeed tremendously aid our understanding of the relationship between proteins and lipids, particularly of how local membrane structure may influence the sorting and functionality of membrane associated or transmembrane proteins. In fact, the variability of cellular membranes in terms of lipid composition and lateral assembly indicates that local heterogeneities induce domain structures within membranes that can be understood as individually confined, two-dimensional functional units, almost like organelles in the three-dimensional aqueous environment. However, looking at the membrane as a two-dimensional entity only partly reflects its physiological role. Membranes are constantly being transformed into spatial structures such as lobes, cristae, stacks, tubes, and vesicles. Consequently, the protein machinery responsible for membrane transformations, and their action on freestanding membranes as GUVs, is another very interesting study object for quantitative biologists and biophysicists concerned with membrane-associated processes.

Prominent membrane transformations in living cells involve the uptake and release of molecules, their packaging and transport from and to distinct sites, the transformation of whole organelles and finally, the large-scale restructuring of the cell membrane during cell division. All of these transformations are tightly regulated and catalyzed by specific protein machineries, presumably triggered through a sophisticated interplay between the local lipid environment and the specifically

adapted protein structure and function (McMahon and Gallop 2005). Although force-inducing motor proteins, as discussed below, are often involved in large-scale membrane transformations, most of the intracellular membrane traffic, e.g. between the Golgi network and the endoplasmic reticulum, actually relies on the recruitment of cytosolic protein machineries, called "coats", prone to cooperatively form higher order structures and thus, inducing buds with specific membrane curvature, which can then be easily transformed into three-dimensional structures such as tubes or vesicles. Many of these proteins share as a structural motif amphipathic peptides or helixes that can efficiently penetrate the membrane (Farsad and De Camilli 2003). The best known coat machineries include the clathrin system (Kirchhausen 2000), as well as the lately very prominent BAR domain proteins (McMahon and Gallop 2005; Antonny 2006).

In recent years, evidence has accumulated that this recruitment of cytoplasmic coats occurs predominantly at specific sites, e.g. with already existing curvatures through lipid asymmetry, or on domains with higher membrane fluidity. Here, in vitro reconstitution of this machinery onto GUVs, especially with domains, is a valuable tool to better understand and characterize the physical parameters governing coat protein recruitment and the subsequent membrane transformation under defined conditions (Sens et al. 2008). In such an approach, Manneville et al. (Manneville et al. 2008) have reconstituted the Arf1-dependent assembly of the COPI coat on GUVs. They could show that membrane recruitment of Arf1-GTP alone occurs exclusively on disordered lipid domains without inducing membrane deformation. Self-assembly of coatomer, as a next step, induces extensive membrane deformation, but only at low membrane tension, and with different lipid composition from the parental membrane, suggesting that the COPI coat does not only introduce local curvature, but may also promote lipid sorting.

If large proteins bind and assemble to one side of a membrane, curvature will result, depending on the mode of assembly or aggregation. This phenomenon has been recently observed with the Gb3 (glycolipid)-binding B-subunit of bacterial Shiga toxin (Römer et al. 2007). Upon Shiga toxin binding to cell and model membranes, narrow membrane tubes were induced, which were supposed to occur through protein-induced lipid reorganization favouring negative membrane curvature. Interestingly, depending on the mode of assembly or aggregation, membrane deformations can lead to protrusions or invaginations, as shown by Saarikangas et al. for different BAR proteins (Saarikangas et al. 2009).

The relevance of the local lipid environment in membrane deformations and budding was highlighted in another recent study with the focus on on intraluminal vesicles of multivesicular endosomes (Trajkovic et al. 2008) In a combined cellular and GUV study, it was found that cargo is segregated into distinct subdomains on the endosomal membrane, and that the transfer of exosome-associated domains into the lumen of the endosome required the sphingolipid ceramide. Purified exosomes were enriched in ceramide, and the release of exosomes was reduced after the inhibition of neutral sphingomyelinases. On the GUVs, conversion of sphingomyelin into ceramide resulted in the spontaneous, protein-free budding of small vesicles away from the GUV membrane.

An intriguing recent study on the in vitro reconstitution of the ESCRT-III proteins in GUVs shows membrane bending by the complex in an unconventional way, where bending and budding away from the mother membrane can actually be induced against the direction of coat assembly, i.e. to the inside if the binding occurs from the outside (Wollert et al. 2009; Barelli and Antonny 2009).

In addition to coat-forming proteins which induce membrane deformations and transformations through self-assembly during membrane attachment, several mechanisms of motor-induced membrane transformations, requiring energy input in form of ATP, can be studied in cell free GUV model systems. Dynamic interactions between the lipid membranes of organelles and the cytoskeleton are critical for intracellular trafficking. Performing in vitro experiments using cell-free extracts, it has been shown that motor proteins, moving along cytoskeletal filaments, exert forces on the membranes to which they are attached (Dabora and Sheetz 1988; Vale and Hotani 1988). To better understand the formation of membrane tubes by motor proteins from free-standing membranes, a tube pulling system was reconstituted by Bassereau and co-workers, by attaching purified kinesin-1 motors to the membranes of GUVs via micron-sized beads (Roux et al. 2002). In an alternative approach, the formation of membrane tubes and tubular networks was observed when kinesin-1 motors were directly linked to the lipids of giant unilamellar vesicles (Koster et al. 2003). The involved forces could be measured directly with optical tweezers, and it was hypothesized that the cooperation of multiple motors, but not their rigid linkage to each other, is necessary for tube formation. The existence of an accumulation of motors at the tip of the tubes was confirmed by fluorescence intensity measurements on the motor density along the tubes, and it was determined that the number of motors pulling a tube can range from four to a few tens (Leduc et al. 2004).

Another interesting mechanism of motor-induced membrane deformation in GUV model systems was found with the GTPase dynamin (Roux et al. 2006), involved in membrane fission in endocytosis. Addition of dynamin and GTP to membrane tubules previously formed from GUVs by the motor activity of kinesin on microtubules, as described above, resulted in twisting of the tubules and supercoiling, suggesting a rotatory movement of the helix turns relative to each other during GTP hydrolysis.

13.5 Splitting GUVs

In the quest of transforming minimal membrane systems, such as GUVs, into minimal cells, one of the most crucial tasks is the reconstitution of spontaneous splitting or division. As the simplest solution to this, membrane transformations of vesicles can be easily induced through variations of the physical parameters that control shape and size. There have been numerous reports about the splitting of vesicles (Bachmann et al. 1992; Luisi et al. 2004), e.g. their breakup into two or more daughter vesicles if lipid mass is constantly added, thereby destabilizing the

surface (Bozic and Svetina 2004), or if osmotic pressure is increased, potentially also through reactions of encapsulated material, as beautifully demonstrated by the Szostak group (Chen et al. 2004) with replicating RNA molecules. Many other chapters within this volume refer explicitly to the phenomenon of splitting vesicles, therefore I will not go into much detail here. In our own work, we demonstrated that splitting of phase-separating vesicles can occur along the phase boundaries of liquid ordered and liquid-disordered membrane domains, if the salt concentration outside the vesicle is slightly altered with respect to the inside (Bacia et al. 2005). Figure 13.2 shows impressively the budding of the liquid-ordered domain away from the larger part of the vesicle, being in the liquid-disordered state. The mechanism behind this process is the larger increase of the line tension along the domain boundary with respect to the increase of surface tension that is against a budding transition. Although this kind of vesicle splitting can be rather controllably realized in heterogeneous membrane systems, and the perspective of sorting membranes (and the correspondingly attached or inserted proteins at the degrees of their specific segregation) is attractive in many ways, the fact that the daughter vesicles are distinct in composition renders this phenomenon rather inappropriate to realize vesicle division. Therefore, the next section will discuss the potential of employing the same kind of protein machinery which is used in "modern" live cells to control the splitting process, namely, a tight interplay between structural elements, termed the cytoskeleton, and active elements, such as molecular motors.
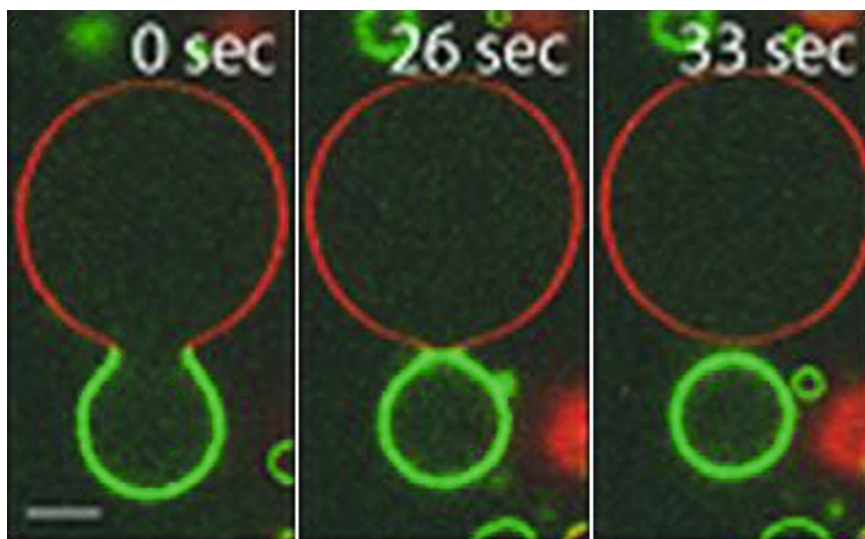
Fig. 13.2 Budding of a liquid ordered domain (*green*) away from a mother vesicle of liquid disordered phase (*red*) upon increased hypertonic conditions (Bacia et al. 2005)

13.6   More Than Membrane: GUVs with a Cytoskeleton/Cortex

13.6.1   General Actin Cortex Assembly

In contrast to other protocols envisioned mainly by "artificial-life" researchers, I will here discuss the possibility of arriving at a controllable splitting mechanism by involving a rather sophisticated (and certainly not primordial) protein machinery based on cytoskeletal elements such as actin filaments. In cells, the interplay between filaments anchored to membranes and active motor proteins, inducing the required forces to transform these membranes, tightly regulates the division and opens up the control of these processes by providing interfaces for other proteins or protein machineries that activate and deactivate the actual division process.

A general requirement to couple the activity of cytoskeletal motors to model membranes is the establishment of a proper interface, i.e., the reconstitution of a cytoskeleton- or cortex-like structure on the membrane, through stationary or transient anchors. Although minimal systems based on filaments and molecular motors haven been extensively studied over the past decades, involving many beautiful in vitro assays mainly on microtubule-kinesin systems, the linkage of these systems to membranes has so far not been studied in detail.

On the other hand, actin, presumably one of the major constitutents of cellular cortices, has been polymerized within GUVs quite early, but without stable attachment between the membrane and the filaments (Limozin and Sackmann 2002). There are many actin-based superstructures interacting with the plasma membrane at different cellular locations, such as the highly branched, polymerizing actin network at the leading edge of migratory cells, stress fibers (long bundles of actin cables) that are attached to sites of adhesion, long actin filaments in filopodia and actin-rich structures found at invaginations in endocytic and phagocytic structures. Actin polymerization is essential for cell locomotion and is thought to generate the force responsible for cellular protrusions. Bacteria like Listeria and Shigella use actin assembly to propel themselves in living cells (Theriot et al. 1992). Cell migration as well as intra- and intercellular motility of bacterial and viral pathogens and endocytic vesicles require a polymerizing actin filament network, exerting protrusive and attachment forces on membranes. Nucleation of actin networks is spatially and temporarily coordinated by a complex interplay of several proteins, i.e. small GTPases in concert with lipid second messengers such as phosphatidyl-inisitol 4,5 biphosphate (PIP2). A minimal system for in vitro reconstitution of this propulsion machinery on GUVs, containing PIP2 using purified components, was established to determine the minimal requirements of actin-based motility (Delatour et al. 2008). This way, it was found that in addition to actin, an activated Arp2/3 (actin related protein 2 and 3) complex for enhanced nucleation, actin depolymerizing factors, capping proteins, and ATP were required for sustained motility. Propulsion is then due to actin polymerization at one end of the actin filaments, and depolymerization at the other end, in a so called 'treadmilling' process (Loisel et al. 1999).

Specifically for this anchoring system, PIP2, in collaboration with the small GTPase Cdc42, binds N-WASP (Neuronal Wiskott-Aldrich Syndrom Protein) and triggers a conformational change that allows binding to and activation of the Arp2/3 complex, which in turn nucleates the formation of a branched actin network in the lamellipodium, with filament barbed ends facing the leading-edge membrane (Chhabra and Higgs 2007; Pollard and Borisy 2003). In a recent approach to anchor an actin network to domain-exhibiting GUVs, Liu and Fletcher (Liu and Fletcher 2007) used this combination of N-WASP, bound to PIP2 and activated Arp2/3 to polymerize actin on the outer surface of GUVs. On phase-separated vesicles, N-WASP, Arp2/3 and actin only formed actin networks on TMR-PIP2 enriched domains. In analyzing the domain melting temperature, they found that the actin network on the surface of the membrane can lead to both: induction of new domains, and the stabilization of existing domains. Furthermore, the actin network seems to spatially bias the location of domain formation after cycling the temperature above and below the melting temperature. This work nicely illustrates how a dynamic actin cytoskeleton can organize the cellular membrane, not only by restricting lipid and protein diffusion, but also by actively organizing membrane domains. This system was further reconstituted onto GUVs together with activated ARF1 to verify actin-based propulsion of ARF1 vesicles and establish a minimal system for the dynamics of Golgi membranes (Heuvingh et al. 2007). Recently, an elegant protocol based on reverse emulsions was released, allowing actin polymerization to nucleate and assemble at the inner membrane of a GUV, resulting in shell thicknesses of the cortices of up to 0.4 μm, 10% of the vesicle size (Pontani et al. 2009).

A completely different approach to tightly anchor cytoskeletal elements to GUV membranes was chosen by our group (Merkle et al. 2008). Using porcine total brain lipid extracts rather than synthetic lipids, static linkage of actin filaments through the ankyrin/spectrin machinery bound to functional ion channels in the GUV membrane could be accomplished (Fig. 13.3). This will be more extensively discussed below.

13.6.2   Stable Filament Anchoring

True to the idea of a minimal cytoskeleton tightly anchored to a GUV membrane, we have reconstituted actin filaments in GUVs made from isolated membrane fractions from porcine brain. The use of membrane fractions maintains the complex lipid composition found in a native brain membrane state, in addition to containing the necessary integral membrane proteins for anchoring the cytoskeleton. For stable attachment, we have also isolated spectrin and ankyrin, added them to the GUVs and anchored the actin filaments to the inner walls of the porcine GUVs. This was the first demonstration of a quaternary-protein system reconstituted in and anchored to the interior walls of GUVs made of native membrane. This work demonstrated the ability to use GUVs as a model 'cell-like' compartment, in which multiprotein systems can be reconstituted and examined in the
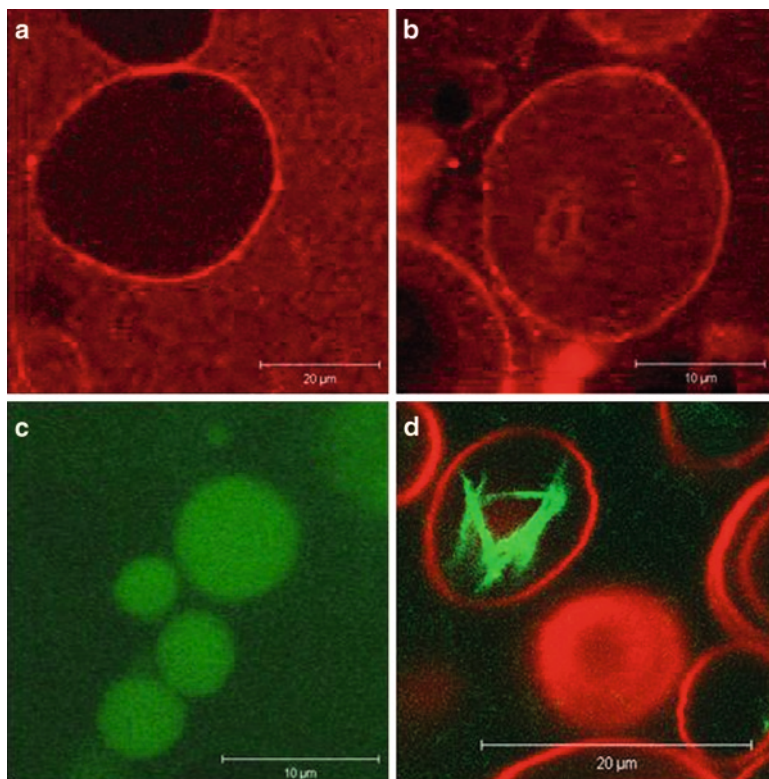
Fig. 13.3 Stable anchoring of actin filaments within the lumen of GUVs made from procine brain membranes, through active ion channels and the spectron/ankyrin protein system. Antibody staining of Na/K ATPase (**a**) and the Ca^{2+} ATPase (**b**), influx of Ca upon addition of ATP and $CaCl_2$, evidenced by fluorescence of Ca green (**c**), stable anchoring of actin filaments when co-incubated with purified spectrin and ankyrin (**d**) (Merkle et al. 2008)

presence of complex lipid mixtures, but may also signify a critical step towards achieving the requirements for actin-induced division of a minimal cell system.

13.6.2.1 GUVs Containing Functional Ion Channels

With the goal of mimicking a physiologically relevant environment for the anchoring of cytoskeleton inside GUVs, and due to the fact that there already exists a thorough characterization of the spectrin-based cytoskeleton in brain, membrane fractions were isolated from porcine brain to form GUVs consisting of a native lipid composition and containing integral membrane proteins capable of binding ankyrin/spectrin. Following protocols by Davis and Bennett (Davis and Bennett 1984) the brains were homogenized, membrane fractions were separated via centrifugation, demyelinated, and the membrane associated proteins were separated from the membrane fractions

containing integral proteins. The extracted membrane fraction was spotted onto ITO coverslips and GUVs were grown via electro-swelling as previously described (Grzybek et al. 2006). GUVs grown from these preparations resulted in various vesicles both multi and unilamellar ranging in size between 5 and 100 μm. They could be visualized either with the lipid dye DiD-C18, or via immunostaining with specific antibodies targeting proteins known to be integral to brain membranes. Figure 13.3 demonstrates a typical GUV from porcine brain, stained with a mono-clonal antibody against Na/K ATPase (a) and the Ca^{2+} ATPase (b), detected via a secondary antibody labeled with Cy5 fluorescent dye. It can be seen that the antibody homogeneously localizes along the surface of the GUV. Immunoblots against actin and tubulin demonstrated the absence of cytoskeletal filaments in these membrane preparations.

To check whether the functionality of the channels was preserved, GUVs were prepared in the presence of ATP and Calcium Green (a sensitive calcium probe that becomes more fluorescent in the presence of Ca^{2+}), such that ATP and Calcium ended up inside the GUVs. We then washed the GUVs extensively with buffer void of Calcium Green. At first, the GUVs showed virtually no fluorescence signal (data not shown). Upon addition of $CaCl_2$ and ATP, the interior of the GUVs became bright fluorescent green (Fig. 13.3c) while GUVs grown and washed in the absence of ATP remained at a nearly non-detectable fluorescent level (not shown). These results demonstrate that the integral membrane ion transporters reconstituted in our GUVs from porcine brain remained active.

13.6.2.2   Attachment of Actin Filaments

GUVs were grown as described above in the presence of purified actin, labeled with phalloidine-Alexa 488. After electroswelling, the GUVs were washed extensively with buffer in order to remove the majority of the actin filaments on the outside of the GUVs. The resulting GUVs contained actin filaments dispersed throughout their interior (Fig. 13.3d). When GUVs were prepared from highly enriched protein fractions containing spectrin and ankyrin, the filaments within the GUVs were no longer dispersed, but displayed dense packing near the walls of the GUVs. In the majority of cases, the filaments appeared anchored to the interior wall of the GUV. After essential controls with total brain lipid extracts not including the ion chan-nels, these results indicated the first reconstitution of stably anchored cytoskeleton to the interior walls of GUVs via the spectrin-ankyrin proteins, which bind to func-tional transmembrane ion channels (such as Na/K ATPase). The spectrin-based membrane skeleton is a network of cytoplasmic structural proteins, first investi-gated in erythrocytes, that underlies regions of the plasma membrane in diverse cells and tissues (Kalvodova et al. 2005). Within this complex architecture, spectrin is thought to connect certain membrane proteins at the cell surface with actin and microtubules in the cytoplasm, and thereby affect the topography and dynamic behavior of these proteins.

13.7 GUVs as Containers

Clearly, the most attractive task of using GUVs as containers is to establish something like an artificial cell, or protocell, using a chemical system encapsulated that is able to self-replicate and evolve (Luisi 2007; Deamer 2005). There are fascinating perspectives of how systems composed entirely of RNA and specific membranes could indeed self-replicate and evolve in the sense that multiple vesicles are generated from a single one, competing for resources (Szostak et al. 2001). However, in these systems, the control over the compartment itself, its size and timing of division, would not or only indirectly be encoded on and regulated by the encapsulated information. As already outlined above, it may thus be an even more fascinating perspective to allow the inclusion of proteins to the GUVs, and ask for a putative protein machinery that could directly couple the replication of nucleic acids to the shape transformations and division of the compartment as such. Introducing proteins to the minimal system naturally enlarges the number of components considerably, although the machinery for in vitro transcription and translation is still supposed to be of modest size (Shimizu et al. 2001). There have been fascinating reports how to realize in vitro protein expression in GUVs, integrating purified machineries for cell-free protein translation (Luisi et al. 2006), as briefly discussed below.

13.7.1 *Cell Free Protein Expression in Vesicles*

A great breakthrough in the quest for protein expression by the machinery of modern cells as a module on its own, to be combined with genes of choice and transplanted into any closed compartment that could resemble the cellular environment, was published in 1998 by Tawfik and Griffiths (Tawfik and Griffiths 1998). They utilized water-in-oil-emulsions to encapsulate different genes into small water droplets separated by mineral oil, achieving true genotype-phenotype coupling. This droplet approach is a potent tool for evolutionary biotechnology, and appealing also for minimal system design, but has the significant drawback that it is an almost completely closed system, where the exchange between different compartments, or between the compartment and its environment, is basically impossible, due to the mineral oil around the droplets. Thus, the sustainability of these systems is limited by the difficulty of delivering new nutrients and energy after the compartment is created and closed. A very attractive solution to this problem, utilizing GUVs, was published a few years ago by the Libchaber group (Noireaux and Libchaber 2004). In this study, the compartmentalization was first achieved by reverse emulsification as outlined above (Pautot et al. 2003), i.e. by centrifuging water droplets through an oil-water-interface and thus creating a fatty acid bilayer from what was first a simple monolayer. The so generated vesicles proved to be stable over days, and, through the incorporation of pore-forming proteins such as alpha-hemolysin, they could also entertain a constant energy supply to the interior,

allowing the protein expression machinery to continually function by creating new protein. Over 3 days, a continuous accumulation of green fluorescent protein could be documented.

Proof-of-principle of cell-free protein expression in membrane vesicles, first published by Nomura et al. (2003) using the gentle hydration technique to generate the GUVs in presence of cell-free extract, was particularly exciting, because it constitutes a potent minimal cell system. It was later intensified using smaller liposomes (Murtas et al. 2007; Luisi et al. 2006), probably because the yield of stable large vesicles using the reverse emulsion technique is still not satisfactory. Thus, if the attractive goal of expressing proteins inside large vesicles (that can be conveniently studied by optical microscopy) is further pursued, other ways of filling vesicles need to be considered. Here, the above mentioned protocol of micro-jetting (Stachowiak et al. 2008), forming vesicles through deforming a planar lipid bilayer into a vesicle, is particularly attractive because the vesicles can in this way be filled with any solution from the jet. Although this procedure still remains to be tested with different lipid and buffer compositions, it presently appears to be most attractive and versatile approach to create large numbers of giant vesicles with basically no limitation of inserting complex mixtures or biomolecules.

13.8   Perspective: Bacterial Cell Division Realized in GUVs?

Clearly, controlled division using a protein-based divisome machinery is one of the most spectacular visions when it comes to converting GUV into minimal cells. Since the simplest divisomes that we know of are the prokaryotic ones, avoiding the double task of nuclear division and cytokinesis, it makes sense to speculate about reconstituting bacterial divisomes in vesicles. One particularly well studied system is the divisome of *E.coli* bacteria, based around a contractile ring, called the Z ring. This division ring is formed by a set of at least ten proteins that assemble at mid-cell to drive cytokinesis (Vicente et al. 2006). Initially three proteins, FtsZ, FtsA and ZipA assemble together forming a proto-ring, into which the other components are added. To ensure the correct assembly of this ring at mid-cell, two main positioning mechanisms are required: the Min system and nucleoid occlusion, which select the constriction site (Yu and Margolin 1999).

The Min system, which is probably the best characterized divisome sub-system in quantitative terms, consists of MinC, D and E proteins, which oscillate repeatedly from one pole of the cell to the opposite pole (Raskin and de Boer 1999). MinC inhibits FtsZ polymerization (Hu and Lutkenhaus 1999) and thereby presumably prevents the ring from forming away from the cell center. MinD, a membrane-bound ATPase, activates MinC's inhibitory activity and directs it to the membrane, while MinE activates MinD's ATPase to drive the oscillation. Their combined action results in FtsZ assembly into a ring at the midcell, where the local time-averaged concentration of MinC inhibitor is lowest. The other spatial regulatory system, nucleoid occlusion, is an additional mechanism to avoid septation at places

occupied by the bacterial nucleoid (bulk chromosomal DNA), and therefore acts as a fail-safe system.

The relationship between MinD and MinE can be considered a classical energy consuming self-organized system as already suggested by Turing in 1952 (Turing 1952), where energy is consumed by the ATPase MinD. When bound to ATP, MinD dimerizes and exposes its membrane targeting sequence (MTS), attaching the previously soluble protein to the inner bacterial membrane. MinE binds to membrane bound MinD, stimulating the ATP turnover. Subsequently, both proteins detach from the membrane and become soluble again. The dynamic pattern formation is ascribed to dynamical instability driven by the hydrolysis of ATP (Howard et al. 2001; Kruse et al. 2007).

The beauty of the Min system is that its self-organization behaviour can be easily reconstituted in minimal membrane-protein systems, as shown by our group using purified and fluorescently-labelled MinD and MinE, and supported membranes made of E.coli lipids PE, PG and cardiolipin (Loose et al. 2008). Upon addition of ATP, impressive travelling wave patterns on the membrane could be observed (Fig. 13.4), retaining the essential features of cellular oscillations, i.e., MinE following MinD at the trailing edge of the wave, and the relationship between wavelength, propagation speed, and diffusional mobility of the proteins in solution being consistent with the cellular situation.

The next logical steps toward reconstitution of a divisome machinery into vesicles are now to (1) find a similar minimal system solution for in vitro assembly of the
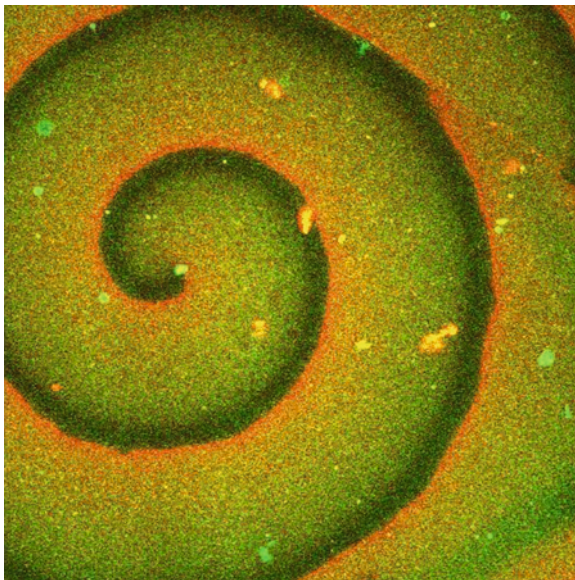
Fig. 13.4 Snapshot from travelling waves of the Min proteins MinD (*green*) and MinE (*red*) on a supported lipid bilayer from E.coli lipid extract, after addition of ATP. In E.coli bacteria, these proteins exhibit pole-to-pole oscillations to determine the mid-cell location (Loose et al. 2008)

Z ring, and (2) observe how the Min waves react to the closing of the membrane surface, by injecting the reaction-diffusion system to vesicles. These are the topics of ongoing work in our laboratory, in order to identify the minimal set of proteins required to divide a vesicle. As a first step towards ring assembly, the attachment of FtsZ filaments to GUV surfaces was realized with a mutant of FtsZ that exhibits a membrane targeting sequence itself (Osawa et al. 2008), avoiding the complete FtsA/ZipA membrane anchoring machinery in order to target the protein to the GUV surface (Fig. 13.5). Due to the absence of MinC gradients, the filaments are still isotropically attached to the surface. Nevertheless, the image shows a minimal realization of a prokaryotic cytoskeleton (FtsZ being a tubulin homologue) on a membrane vesicle, and thus a potentially very interesting study object for a better understanding of the mutual relationship between membranes and membrane-deforming protein machineries.
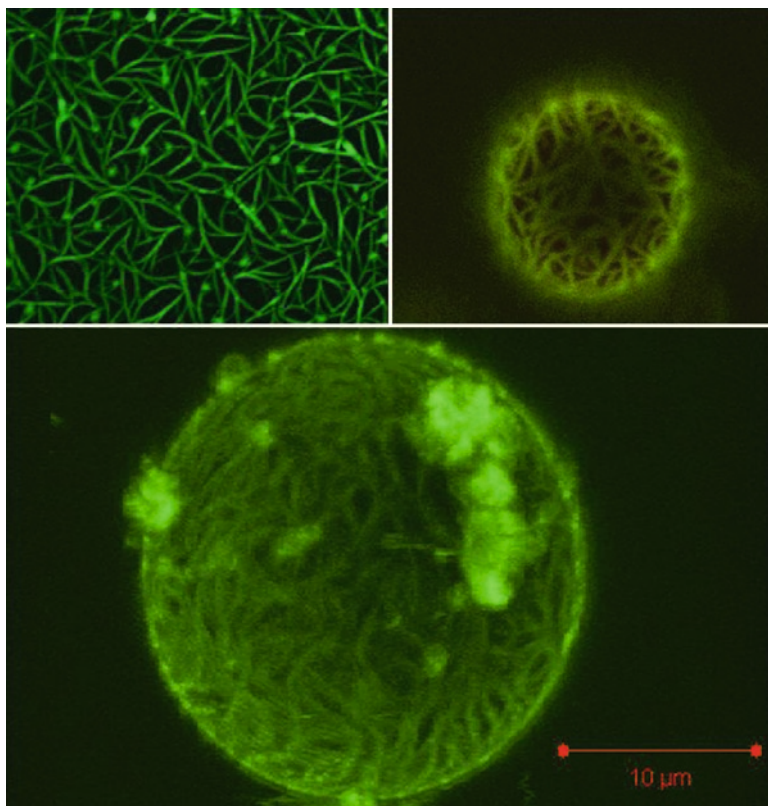
Fig. 13.5 Fluorescently labelled FtsZ protein with membrane targeting sequence (FtsZ-MTS, Osawa et al. 2008) when reconstituted and polymerized on supported membranes (*upper left*) and on GUVs (*upper right*, and *lower panel*) (Chiantia and Arumugam unpublished data)

13.9 Conclusion and Outlook

Giant unilamellar vesicles have been valuable study objects for fundamental understanding of various lipid systems, in particular mixtures prone to phase separation and thus, domain formation. They have greatly aided our understanding of membrane-protein interactions, both in the sorting and partitioning of proteins into particular lipid phases, but also in the quantitative characterization of protein activity in transforming membranes. Due to their comfortable size, amenable for all kinds of opical microscopy, they are much more attractive for biophysical characterizations of minimal biological systems than those based on small liposomes. In the context of realizing minimal cells, GUVs are still not widely used due to remaining problems with stability and reproducibility of their generation. However, recent protocols how to generate them at physiological conditions, and, more importantly, to deliver functional proteins and nucleic acids to their interior, raise strong hopes for a more prominent role in minimal cell assembly. A particularly interesting perspective is the realization of controlled vesicle division by a minimal divisome, e.g. taken from prokaryotic systems.

References

Akashi K, Miyata H, Itoh H, Kinosita K (1996) Preparation of giant liposomes in physiological conditions and their characterization under an optical microscope. Biophys J 71:3242–3250

Angelova MI, Soleau S, Meleard P, Faucon JF, Bothorel B (1992) Preparation of giant vesicles by external AC electric fields. Kinetics and applications. Progr Colloid Polym Sci 89:127–131

Antonny B (2006) Membrane deformation by protein coats. Curr Op Cell Biol 18(4):386–394

Bacia K, Schuette CG, Kahya N, Jahn R, Schwille P (2004) SNAREs prefer liquid-disordered over "raft" (liquid-ordered) domains when reconstituted into giant unilamellar vesicles. J Biol Chem 279:37951–37955

Bachmann PA, Luisi PL, Lang J (1992) Autocatalytic selfreplicating micelles as models for prebiotic structures. Nature 357:57–59

Bacia K, Schwille P, Kurzchalia T (2005) Sterol structure determines the separation of phases and the curvature of the liquid-ordered phase in model membranes. Proc Natl Acad Sci USA 102:3272–3277

Bagatolli LA, Gratton E (2000) Two photon fluorescence microscopy of coexisting lipid domains in giant unilamellar vesicles of binary phospholipid mixtures. Biophys J 78:290–305

Barelli H, Antonny B (2009) Detached membrane binding. Nature 458:159–160

Baumgart T, Hess ST, Webb WW (2003) Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. Nature 425:821–824

Bozic B, Svetina S (2004) A relationship between membrane properties forms the basis of a selectivity mechanism for vesicle self-reproduction. Eur Biophys J 33:565–571

Chen IA, Roberts RW, Szostak JW (2004) The emergence of competition between model protocells. Science 305:1474–1476

Chhabra ES, Higgs HN (2007) The many faces of actin: matching assembly factors with cellular structures. Nat Cell Biol 9(10):1110–1121

Chiantia S, Ries J, Kahya N, Schwille P (2006) Combined AFM and two-focus SFCS study of Raft-exhibiting model membranes. ChemPhysChem 7:2409–2418

Chiantia S, Ries J, Chwastek G, Carrer D, Li Z, Bittman R, Schwille P (2008) Role of ceramide in membrane protein organization investigated by combined AFM and FCS. Biochim Biophys Acta 1778:1356–1364

Dabora SL, Sheetz MP (1988) The microtubule-dependent formation of a tubulovesicular network with characteristics of the Er from cultured-cell extracts. Cell 54:27–35

Davidson M, Karlsso M, Sinclair J, Sott K, Orwar O (2003) Nanotube-vesicle networks with functionalized membranes and interiors. J Am Chem Soc 125(2):374–378

Davis JQ, Bennett V (1984) Brain ankyrin. A membrane-associated protein with binding sites for spectrin, tubulin, and the cytoplasmic domain of the erythrocyte anion channel. J Biol Chem 259:13550–13559

Day CA, Kenworthy AK (2009) Tracking microdomain dynamics in cell membranes. Biochim Biophys Acta Biomembranes 1788:245–253

Deamer D (2005) A giant step towards artificial life? Trends Biotechnol 23:336–338

Delatour V, Helfer E, Didry D, Le KHD, Gaucher JF, Carlier MF, Romet-Lemonne G (2008) Arp2/3 controls the motile behavior of N-WASP-functionalized GUVs and modulates N-WASP surface distribution by mediating transient links with actin filaments. Biophys J 94:4890–4905

Dietrich C, Bagatolli LA, Volovyk ZN, Thompson NL, Levi M, Jacobson K, Gratton E (2001) Lipid rafts reconstituted in model membranes. Biophys J 80:1417–1428

Eggeling C, Ringemann C, Sandhoff K, Schönle A, Hell SW et al (2009) Direct observation of the nanoscale dynamics of membrane lipids in a living cell. Nature 457:1159–1162

Farsad K, De Camilli P (2003) Mechanisms of membrane deformation. Curr Op Cell Biol 15:372–381

Feigenson GW (2006) Phase behavior of lipid mixtures. Nat Chem Biol 2:560–563

Girard P, Pecreaux J, Lenoir G, Falson P, Rigaud JL, Bassereau P (2004) A new method for the reconstitution of membrane proteins into giant unilamellar vesicles. Biophys J 87:419–429

Grzybek M, Chorzalska A, Bok E, Hryniewicz-Janowska A, Czogalla A, Diakowski W, Sikorski AF (2006) Spectrin-phospholipid interactions. Existence of multiple kinds of binding sites? Chem Phys Lipids 141:133–141

Hammond AT, Heberle FA, Baumgart T, Holowka D, Baird B, Feigenson GW (2005) Crosslinking a lipid raft component triggers liquid ordered-liquid disordered phase separation in model plasma membranes. Proc Natl Acad Sci USA 102:6320–6325

Helfrich W (1973) Elastic properties of lipid bilayers – theory and possible experiments. Z Naturforsch C 28:693–703

Heuvingh J, Franco M, Chavrier P, Sykes C (2007) ARF1-mediated actin polymerization produces movement of artificial vesicles. Proc Natl Acad Sci USA 104:16928–16933

Howard M, Rutenberg AD, de Vet S (2001) Dynamic compartmentalization of bacteria: accurate division in E-coli. Phys Rev Lett 87:278102

Hu ZL, Lutkenhaus J (1999) Topological regulation of cell division in Escherichia coli involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE. Mol Microbiol 34:82–90

Kahya N, Pecheur EI, de Boeij WP, Wiersma DA, Hoekstra D (2001) Reconstitution of membrane proteins into giant unilamellar vesicles via peptide-induced fusion. Biophys J 81:1464–1474

Kahya N, Scherfeld D, Bacia K, Poolman B, Schwille P (2003) Probing lipid mobility of raft-exhibiting model membranes by fluorescence correlation spectroscopy. J Biol Chem 278:28109–28115

Kalvodova L, Kahya N, Schwille P, Ehehalt R, Verkade P, Drechsel D, Simons K (2005) Lipids as modulators of proteolytic activity of BACE: involvement of cholesterol, glycosphingolipids, and anionic phospholipids in vitro. J Biol Chem 280:36815–36823

Kirchhausen T (2000) Clathrin. Clathrin Annu Rev Biochem 69:699–727

Korlach J, Schwille P, Webb WW, Feigenson GW (1999) Characterization of lipid bilayer phases by confocal microscopy and fluorescence correlation spectroscopy. Proc Natl Acad Sci USA 96:8461–8466

Koster G, VanDuijn M, Hofs B, Dogterom M (2003) Membrane tube formation from giant vesicles by dynamic association of motor proteins. Proc Natl Acad Sci USA 100:2351–2356

Kruse K, Howard M, Margolin W (2007) An experimentalist's guide to computational modelling of the Min system. Mol Microbiol 63:1279–1284

Leduc C, Campas O, Zeldovich KB, Roux A, Jolimaitre P, Bourel-Bonnet L, Goud B, Joanny JF, Basserea P, Prost J (2004) Cooperative extraction of membrane nanotubes by molecular motors. Proc Natl Acad Sci USA 101:17096–17101

Limozin L, Sackmann E (2002) Polymorphism of cross-linked actin networks in giant vesicles. Phys Rev Lett 89:168103.1–168103.4

Lingwood D, Ries J, Schwille P, Simons K (2008) Plasma membranes are poised for activation of raft phase coalescence at physiological temperature. Proc Natl Acad Sci USA 105: 10005–10010

Lipowsk R (1991) The conformation of membranes. Nature 349:475–481

Lipowsky R, Sackmann E (1995) Structure and dynamics of membranes. Elsevier, North Holland

Liu AP, Fletcher DA (2007) Actin polymerization serves as a membrane domain switch in model lipid bilayers. Biophys J 92:697–697

Loisel TP, Boujemaa R, Pantaloni D, Carlier MF (1999) Reconstitution of actin-based motility of Listeria and Shigella using pure proteins. Nature 401:613–616

Loose M, Fischer-Friedrich E, Ries J, Kruse K, Schwille P (2008) Spatial regulators for bacterial cell division self-organize into surface waves in vitro. Science 320:789–792

Luisi PL (2007) Chemical aspects of synthetic biology. Chem Biodivers 4(4):603–621

Luisi PL, Stano P, Rasi S, Mavelli F (2004) A possible route to prebiotic vesicle reproduction. Artif Life 10:297–308

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 93: 1–13

Manneville JB, Casella JF, Ambroggio E, Gounon P, Bertherat J, Bassereau P, Cartaud J, Antonny B, Goud B (2008) COPI coat assembly occurs on liquid-disordered domains and the associated membrane deformations are limited by membrane tension. Proc Natl Acad Sci USA 105:16946–16951

McMahon HT, Gallop JL (2005) Membrane curvature and mechanisms of dynamic cell membrane remodelling. Nature 438:590–596

Merkle D, Kahya N, Schwille P (2008) Reconstitution and anchoring of cytoskeleton inside giant unilamellar vesicles. Chembiochem 9:2673–2681

Montes LR, Alonso A, Goñi FM, Bagatolli LA (2007) Giant Unilamellar Vesicles electroformed from native membranes and organic lipid mixtures under physiological conditions. Biophys J 93:3548–3554

Murtas G, Kuruma Y, Bianchini P, Diaspro A, Luisi PL (2007) Protein synthesis in liposomes with a minimal set of enzymes. Biochem Biophys Res Comm 363:12–17

Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. Proc Natl Acad Sci USA 101:17669–17674

Nomura SM, Tsumoto K, Hamada T, Akiyoshi K, Nakatani Y, Yoshikawa K (2003) Gene expression within cell-sized lipid vesicles. ChemBioChem 4:1172–1175

Osawa M, Anderson DE, Erickson HP (2008) Reconstitution of contractile FtsZ rings in liposomes. Science 320:792–794

Pautot S, Frisken BJ, Weitz DA (2003) Engineering asymmetric vesicles. Proc Natl Acad Sci USA 100:10718–10721

Pollard TD, Borisy GG (2003) Cellular motility driven by assembly and disassembly of actin filaments. Cell 112:453–465

Pontani LL, van der Gucht J, Salbreux G, Heuvingh G, Joanny JF, Sykes C (2009) Reconstitution of an actin cortex inside a liposome. Biophys J 96:192–198

Pott T, Bouvrais H, Meleard P (2008) Giant unilamellar vesicle formation under physiologically relevant conditions. Chem Phys Lipids 154:115–119

Rajendran L, Simons K (2005) Lipid rafts and membrane dynamics. J Cell Sci 118:1099–1102

Raskin DM, de Boer PA (1999) Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli. Proc Natl Acad Sci USA 96:4971–4976

Römer W, Berland L, Sens P, Bassereau P, Johannes L et al (2007) Shiga toxin induces tubular membrane invaginations for its uptake into cells. Nature 450(7170):670

Roux A, Cappello G, Cartaud J, Prost J, Goud B, Bassereau P (2002) A minimal system allowing tubulation with molecular motors pulling on giant liposomes. Proc Natl Acad Sci USA 99: 5394–5399

Roux A, Uyhazi K, Frost A, De Camilli P (2006) GTP-dependent twisting of dynamin implicates constriction and tension in membrane fission. Nature 441(7092):528–531

Saarikangas J, Zhao HX, Pykalainen A, Laurinmaki P, Mattila PK, Kinnunen PKJ, Butcher SJ, Lappalainen P (2009) Molecular mechanisms of membrane deformation by I-BAR domain proteins. Curr Biol 19:95–107

Saslowsky DE, Lawrence J, Ren XY, Brown DA, Henderson RM, Edwardson JM (2002) Placental alkaline phosphatase is efficiently targeted to rafts in supported lipid bilayers. J Biol Chem 277:26966–26970

Scherfeld D, Kahya N, Schwille P (2003) Lipid dynamics and domain formation in model membranes composed of ternary mixtures of unsaturated and saturated phosphatidylcholines and cholesterol. Biophys J 85:3758–3768

Schön P, Garcia-Saez AJ, Malovrh P, Bacia K, Anderluh G, Schwille P (2008) Equinatoxin II permeabilizing activity depends on the presence of sphingomyelin and lipid phase coexistence. Biophys J 95:691–698

Seifert U (1997) Configurations of fluid membranes and vesicles. Adv Phys 46:13–137

Sens P, Johannes L, Bassereau P (2008) Biophysical approaches to protein-induced membrane deformations in trafficking. Curr Op Cell Biol 20:476–482

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755

Simons K, Ikonen E (1997) Functional rafts in cell membranes. Nature 387:569–572

Simons K, van Meer G (1988) Lipid sorting in epithelial-cells. Biochemistry 27:6197–6202

Stachowiak JC, Richmond DL, Li TH, Liu AP, Parekh SH, Fletcher DA (2008) Unilamellar vesicle formation and encapsulation by microfluidic jetting. Proc Natl Acad Sci USA 105:4697–4702

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390

Tawfik DS, Griffiths AD (1998) Man-made cell-like compartments for molecular evolution. Nat Biotechnol 16:652–656

Theriot JA, Mitchison TJ, Tilney LG, Portnoy DA (1992) The rate of actin-based motility of intra-cellular listeria-monocytogenes equals the rate of actin polymerization. Nature 357:257–260

Trajkovic K, Hsu C, Chiantia S, Rajendran L, Wenzel D, Wieland F, Schwille P, Brügger B, Simons M (2008) Ceramide triggers budding of exosome vesicles into multivesicular Endosomes. Science 319(5867):1244–1247

Turing AM (1952) The chemical basis of morphogenesis. Phil Trans Royal Soc London B Biol Sci 237:37–72

Vale RD, Hotani H (1988) Formation of membrane networks invitro by Kinesin-Driven microtubule movement. J Cell Biol 107:2233–2241

Varma R, Mayor S (1998) GPI-anchored proteins are organized in submicron domains at the cell surface. Nature 394:798–801

Veatch SL. Keller SL (2003a) A closer look at the canonical 'Raft Mixture' in model membrane studies. Biophys J 84:725–726

Veatch SL. Keller SL (2003b) Separation of liquid phases in giant vesicles of ternary mixtures of phospholipids and cholestrol. Biophys J 85:3074–3083

Vicente M, Rico AI, Martinez-Arteaga R, Mingorance J (2006) Septum enlightenment: assembly of bacterial division proteins. J Bacteriol 188:19–27

Wawrezinieck L, Rigneault H, Marguet D, Lenne PF (2005) Fluorescence correlation spectroscopy diffusion laws to probe the submicron cell membrane organization. Biophys J 89:4029–4042

Wollert T, Wunder C, Lippincott-Schwartz J, Hurley JH (2009) Membrane scission by the ESCRT-III complex. Nature 458:172–177

Yu XC, Margolin W (1999) FtsZ ring clusters in *min* and partition mutants: role of both the Min system and the nucleoid in regulating FtsZ ring localisation. Mol Microbiol 32:315–326

Chapter 14
Theoretical Approaches to Ribocell Modeling

Fabio Mavelli

Abstract The so-called Ribocell (ribozymes-based cell) is a theoretical cellular model that has been proposed some years ago as a possible minimal cell prototype. It consists in a self-replicating minimum RNA genome coupled with a self-reproducing lipid vesicle compartment. This model assumes the existence of two hypothetical ribozymes one able to catalyze the conversion of molecular precursors into lipids and the second one able to replicate RNA strands. Therefore, in an environment rich both of lipid precursors and activated nucleotides, the ribocell can self-reproduce if the genome self-replication and the compartment self-reproduction mechanisms are somehow synchronized. The aim of this contribution is to explore the feasibility of this model with in silico simulations using kinetic parameters available in literature.

14.1 Introduction

A minimal cell, or protocell, is a minimum supra molecular bounded structure that can be considered alive. The adjectives minimum/minimal stress the fact that this molecular machinery must be set up by the lowest number of constituents and it must involve the lowest number of processes in the internal metabolism. On the other hand, the requirement to be bounded guaranties that its individuality is kept throughout its life, or in other words, that it will be easy to distinguish the living cell from the surrounding environment.

In order to be considered alive, a protocell must be able to self-maintain, i.e. it must stay in a steady state where all its constituents are continuously synthesized and refurbished (Luisi 2003). The internal metabolism must be able to transforms the nutrients available in the environment into fresh molecular constituents and continuously replace the old with the new ones. *Self-Maintenance* is a necessary

F. Mavelli (✉)
Chemistry Department – University of Bari, Via Orabona 4, 70125, Bari, Italy
e-mail: mavelli@chimica.uniba.it

condition that must be fulfilled by protocells to be alive; nevertheless it is not sufficient to implement the cellular life as we actually know it (Luisi 1998). In fact, another important requisite is the capability to *Self-Reproduction*. This can be seen as a consequence of the internal metabolism that can keep the minimal cell in a stationary state or bring it to a continuous growing regime that carry the organism to divide in order to maintain its internal coherence,[1] i.e. its stability. It is worthwhile to remark, as both Self-Maintenance and Self-Reproduction are individual features of a cell, i.e. properties of a single living being that can be observed during its life time. On the other hand, *Evolvability* is a collective property that can be exhibited only by a population of cells and in a time range of several generations according to a Darwinian selection mechanism (Luisi 1998).

Some years ago, Szostak and colleagues proposed a minimal cell prototype called the *Ribocell*: the RNA based cell (Szostak et al. 2001) that in principle can exhibit all the three properties to be considered a living protocell. This theoretical cellular model consists in a self-replicating minimum genome coupled with the self-reproduction of the lipid vesicular container (Fig. 14.1). These authors
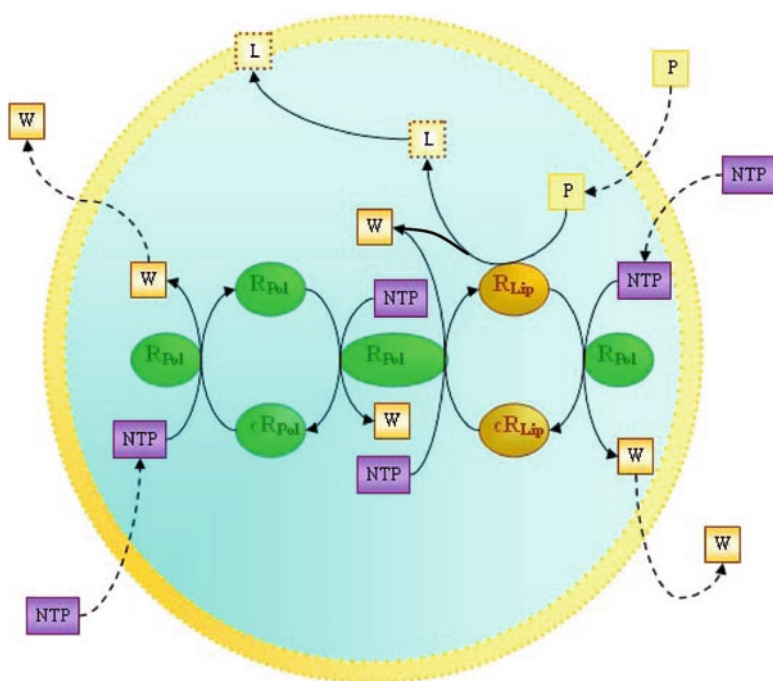
Fig. 14.1 The Ribocell model: graphical representation of membrane transport processes (*dashed lines*) and internal reactions (*solid lines*)

---

[1]Of course a cell can also die for not optimal external conditions.

envisaged the existence of two hypothetical ribozymes, one (R_{Lip}) able to catalyze the conversion of molecular precursors (P) into membrane lipids (L) and the other one (R_{Pol}) able to duplicate RNA strands. Therefore, in an environment rich of both lipid precursors (P) and activated nucleotides (NTPs), the Ribocell can self-reproduce if both processes: the genome self-replication and the membrane reproduction (growth and division), are somehow synchronized.

In order to be able of self-replicating the entire genome, the ribocell must contain a minimum genetic kit composed of three RNA strands: one R_{Pol} that catalyzes the base pair transcription of a RNA template with one $R_{Lip}/_cR_{Lip}$ and one $R_{Pol}/_cR_{Pol}$ both that work as templates for the transcription. On the other hand, an R_{Lip} filament is necessary to catalyze the lipid precursor conversion so that the membrane can grow by taking up the fresh amphiphiles. The optimal minimum three-ribozymes kit: R_{Pol}, $_cR_{Pol}$ and R_{Lip}, should be at least doubled before the division to have a chance that both the daughter vesicles continue to be active as the mother. In fact, in the most optimistic case both daughters will include the correct kit of ribozymes after division. In a more realistic case, if the vesicle self-reproduction is not synchronized with the genome self-replication or if a random distribution of RNA filaments takes place after the vesicle division, ribozymes segregation in the two different compartments may occur bringing to *death for dilution*. Different scenarios can be sketched after the cell division and the RNA strands partition. In particular, the death for dilution is reached if vesicles are produced without any ribozymes (*empty vesicles*) or containing only one lone R_{Pol} or many filaments of $_cR_{Pol}$ and/or $_cR_{Lip}$ (*inert vesicles*). Vesicles that encapsulate R_{Lip} strands are *self-reproducing vesicles*: they are able to synthesize lipids and then can growth and divide producing in turn self-replicating and/or empty vesicles. On the other hand, vesicles containing more than one molecule of R_{Pol} or both R_{Pol} and $_cR_{Pol}$ are able to self-replicate this reduced genome (*self-replicating genome*) but they can not self-reproduce the membrane. Moreover, a reduced version of the ribocell consists in a lipid aggregate that contains one R_{Pol} and $R_{Lip}/_cR_{Lip}$ strands. As a consequence of this, *reduced ribocells* are able to replicate the $R_{Lip}/_cR_{Lip}$ genetic staff and at the same time to synthesize lipids. Therefore they can growth and divide producing in turn reduced ribocells and/or self-replicating, inert or empty vesicles.

In this chapter we tried to explore all these scenarios presenting an *in silico* implementation of the ribocell and simulating the time behavior of a single proto-cell or a population of them with a deterministic and stochastic approach. Deterministic calculations are used to follow the average time course of a vesicle population by solving the ordinary differential equation set (ODE set) associated with the ribocell metabolism. With this approach we can quickly explore the model parameter space to find sets of parameters that show the attainment of a stationary regime of growth and division alternating stages where genome self-replication and vesicle reproduction are perfectly synchronized. On the other hand, the robustness of the ribocell to random fluctuations will be tested by using ENVIRONMENT, a software platform recently developed for stochastically simulate compartmentalized reacting systems

(Mavelli and Ruiz-Mirazo 2007, 2010). This Monte Carlo program allows us to follow a population of ribocells simulating a random partition of ribozymes after any division and taking into account the intrinsic stochasticity of chemical processes.

14.2 In Silico Ribocell

In this section all the assumptions and the kinetic mechanism assumed in order to describe the Ribocell time behavior in silico will be illustrated and discussed in details. Whenever possible, all the physical and chemical processes will be modeled by using experimental data available in literature in order to get an in silico model as more realistic as possible. In the following subsections the Ribocell kinetic model is decomposed into the dynamics of the lipid bilayer shell (the lipid vesicle) described as an open, volume-changing, reacting system, and the internal reaction mechanism (the metabolism), i.e. the complete list of reactions taking place in the vesicle aqueous core.

14.2.1 Reacting Vesicle Dynamics

A reacting vesicle can be described as a homogeneous compartmentalized reacting aqueous domain encapsulated into a close lipid bilayer. Metabolites and water can be exchanged with the external environment through the membrane driven by the metabolite concentration gradient and the osmotic pressure respectively. The membrane can release and take up amphiphilic molecules towards and from both the external and internal aqueous solutions, while chemical reactions can occur in the vesicle water pool. Table 14.1 compares the concentration change rates with the event propensity probability density functions for all the mentioned processes. The deterministic rates will be used for building up the deterministic ODES according to the classical kinetic theory, while the propensity probability density functions are used to carry out the stochastic simulations according to the Gillespie's Direct method (Gillespie 1977, 1976).

In both these approaches, assuming for sake of simplicity that the vesicle membrane is composed by a single amphiphilic molecule X_L, the state of the vesicle is described by the array:

$$\mathbf{x}^T = \left(n_1^C, n_2^C, \ldots n_N^C, n_L^\mu, V_C \right)$$

where n_i^C are the molecular numbers of species X_i ($i = 1,2 \ldots N$) present in the vesicle aqueous core and n_L^μ is the number of amphiphiles X_L in the membrane. V_C is the water internal volume. In the stochastic approach, all n_i^C and n_L^μ are discrete integer numbers, while in the deterministic analysis they are continuous real values

Table 14.1 Deterministic Rates against propensity density probability for reacting and transport events

Event	Deterministic rate $(Ms)^{-1}$	Propensity density probability (s^{-1})
Internal chemical reactions[a]	$k_\rho \prod_j^N \left(\dfrac{n_j^C}{V_C N_A} \right)^{a_{j,\rho}}$	$\dfrac{k_\rho}{(V_C N_A)^{M_\rho - 1}} \prod_j^N \left(\begin{array}{c} n_j^C \\ a_{j,\rho} \end{array} \right)$

$$a_{1,\rho} X_1 + \ldots a_{N,\rho} X_N \xrightarrow{\;r_\rho\;} b_{1,\rho} X_1 + \ldots b_{N,\rho} X_N$$

Event	Deterministic rate	Propensity density probability
Solute X_n membrane transport[b]	$P_n \dfrac{S_\mu}{V_C} \left(C_n^E - C_n^C \right)$	$D_n S_\mu \dfrac{\left| \left(C_n^E - C_n^C \right) \right|^{(C)}}{\lambda_\mu}$
Membrane lipid release	$\dfrac{k_{out} n_L^\mu}{N_A V_C}$	$k_{out} n_L^\mu$
Membrane lipid uptake	$\left(\dfrac{k_\rho\, in S_\mu [X_L^C]}{N_A V_C} \right)$	$k_\rho\, in S_\mu [X_L^C]$
Water flux[d]	$v_{aq} P_{aq} S_\mu \left(C_T^C - C_T^E \right)$	$V_C = \sum_{i=1}^N n_i^C \Big/ \left(N_A C_T^E \right)$

[a] **a** and **b** ($N \times R$) stochiometric matrixes, N_A Avogadro's number, r_ρ reaction rate, k_ρ kinetic constant, M_ρ γ-th reaction molecularity (Mavelli and Piotto 2006)
[b] The relationship between the macroscopic permeability P_n and the molecular diffusion coefficient D_n is: $D_n = P_n \lambda_\mu N_A$ (Mavelli and Ruiz-Mirazo 2010), λ_μ being the membrane thickness
[c] The absolute value guarantees that the propensity density probability is positive and the molecules are moved in the opposite direction respect to the concentration gradient
[d] v_{aq} is the water molar volume, while C_T^E and C_T^C are the total concentrations in the external and internal aqueous solutions respectively and Paq is the membrane water permeability

that give the average time behaviour of a reacting vesicles population.[2] Another main difference in the two approaches is that while in the deterministic calculation the water flux through the membrane is explicitly considered as driven by an osmotic pressure gradient, in the stochastic simulation this flux of water will be considered instantaneous. In fact, at each step throughout the simulation the aqueous vesicle core volume is rescaled in order to keep the vesicle in an osmotic balanced state by using the formula reported in Table 14.1. As showed elsewhere (Mavelli and Ruiz-Mirazo 2010), the main discrepancy between the two approaches is in

---

[2] The underlying assumption of the deterministic approach is that all vesicles have the same time behavior in average.

describing the time course of vesicles shrinking or swelling due to an abrupt change of the osmolite external concentration as after a fast dilution or a large solute addition. As a consequence of this, the difference between the two approaches will be negligible since the ribocell will be studied starting from osmotic balance conditions.

14.2.1.1 Membrane Stability

In both approaches, the vesicle surface is estimated by the formula: $S_\mu = (\alpha_L n_L^\mu)/2$ where α_L is the amphiphile X_L head area and 1/2 takes into account the double layer nature of the membrane. Since the membrane surface and the aqueous core volume follow two different dynamics, this may bring the vesicles toward unstable conditions. The stability of the vesicle membrane can then be monitored by means of the reduced surface Φ:

$$\Phi = \frac{S_\mu}{S_{V_C}^\varnothing} = \frac{S_\mu}{\sqrt[3]{36\pi V_C^2}}$$

that is, the ratio of the actual membrane surface S_μ divided by the spherical area that would perfectly wrap the actual core volume V_C. Assuming that, for a given size, the spherical shape ($\Phi = 1$) represents the minimum energy state, swollen ($\Phi < 1$) and deflated ($\Phi > 1$) vesicles are in high energetic conditions due to the elastic and the bending tension respectively. Therefore, vesicles are stable only in a small range of Φ values around 1:

$$(1-\varepsilon) \le \Phi \le \sqrt[3]{2}\,(1+\eta)$$

ε and η being the osmotic and dividing tolerance coefficients respectively. In fact, vesicles in hypotonic solutions can swell, stretching the membrane until they reach a critical state: $\Phi = 1 - \varepsilon$. The osmotic tolerance ε can be experimentally determined by measuring the maximum difference in osmolite aqueous concentrations -between the internal core (C_T^C) and the external environment (C_T^E) – that vesicles can bear. For oleic acid and POPC vesicles ε was found equal to 0.21 and 0.59 respectively (Chen et al. 2004; Mavelli et al. 2008). In our model, at the critical bursting point vesicles are assumed simply to break down, releasing all their internal content into the environment and remaining in solution as flat bilayers, since bilayer sealing processes are not considered in this model. On the other hand, deflated vesicles are supposed to be able to divide in order to minimize the bending energy. The dividing condition is reached when they can form two equal volume spherical daughters ($\Phi = 2^{1/3}$). So η introduces a tolerance that is linked to the relative flexibility exhibited by any membrane, however as a simplifying assumption $\eta = 0$ will be supposed.

After the splitting all the molecular content of the mother vesicle is halved in the deterministic approach, while it is randomly distributed between the twin daughters in the stochastic simulations.

$$R_{Pol} + {}_cR_{Pol} \underset{k_S}{\overset{k_{SS}}{\rightleftharpoons}} R_cR_{Pol}$$

$$R_{Lip} + {}_cR_{Lip} \underset{k_S}{\overset{k_{SS}}{\rightleftharpoons}} R_cR_{Lip}$$

(1)

$$R_{Pol} + S \xrightarrow{k_{R@S}} R@S$$

$$R@S_cS_n + NTP_{n+1} \xrightarrow{k_{NPT}} R@S_cS_{n+1} + W$$

$$R@S_cS \xrightarrow{k_{R@SS}} R_{Pol} + S_cS$$

(2)

$$P + R_{Lip} \xrightarrow{k_L} L + R_{Lip} + W$$

(3)
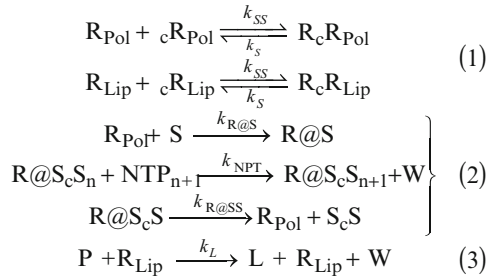
Fig. 14.2 The Ribocell internal metabolism: (1) reversible association of RNA polymerase (R_{Pol}) and RNA-synthase (R_{Lip}) strands with the respective complement ${}_cR_{Pol}$ and ${}_cR_{Lip}$; (2) catalytic cycle of the RNA replication (S = R_{Pol}, ${}_cR_{Pol}$, R_{Lip} and ${}_cR_{Lip}$), (3) conversion of the precursor P into the membrane lipid L catalyzed by the ribozyme R_{Lip}

14.2.2   Internal Metabolism

The Ribocell internal metabolism proposed in this work is illustrated in Fig. 14.2. Both pairs of RNA strands reversibly associate (1) and these processes are shifted towards the dimer formation and strongly dependent on temperature. The replication of any RNA strand is catalyzed by the polymerase R_{Pol} according to the steps in bracket (2). This process is described as a catalytic template-direct addition of mononucleotides with high fidelity and processivity.

It starts with R_{Pol} binding any monomeric strand S (S = R_{Pol}, ${}_cR_{Pol}$, R_{Lip} and ${}_cR_{Lip}$) to form the complex R@S. Then this complex will initiate the polymerization of the conjugate strand ${}_cS$, by coupling and binding iteratively the complementary bases and releasing the by-product W. When the strand ${}_cS$ has been completely formed, the polymerase releases the new dimer. In order to duplicate the entire genome, the replication cycle should operate on all the four RNA strands at least once. Finally, the precursor P is converted into the lipid L by the assistance of the ribozyme R_{Lip} (3)

14.2.3   Kinetic Parameters and Assumptions

Table 14.2 shows the values of all kinetic constants and membrane permeability used in this work with the respective references from literature. The two ribozymes are assumed both 20 nucleotides long (the minimum number for observing a folded RNA conformation) with a random sequence of bases, and with a similar kinetic behavior for sake of simplicity. The kinetic constants of formation k_{SS} and dissociation k_S of both dimers were set equal to the values experimentally measured for a sequence of ten nucleotides (Christensen 2007). The kinetic constants for both complex formation R@S and complex dissociation $R@S_cS$ (S = R_{Pol}, ${}_cR_{Pol}$, R_{Lip} and ${}_cR_{Lip}$) were set equal to those measured for the human enzyme β-polymerase (Tsoi and Yang 2002) that binds a DNA single strand and dissociates from a DNA double helix, respectively. The rate constant for the catalytic synthesis of lipid k_L was taken

Table 14.2 Kinetic constants and permeability of the Ribocell in silico model at room temperature (S = R_{Pol}, $_cR_{Pol}$, R_{Lip} and $_cR_{Lip}$)

Kinetic patameters	Values	Process description	Reference
k_{SS}[s⁻¹M⁻¹]	8.8·10⁶	Formation of dimers $R_c R_{Pol}$ and $R_c c R_{Lip}$	Christensen 2007
k_S[s⁻¹]	2.2·10⁻⁶	Dissociation of dimers $R_c R_{Pol}$ and $R_c c R_{Lip}$	Christensen 2007
$k_{R@S}$[s⁻¹M⁻¹]	5.32·10⁵	Formation of R@S	Tsoi and Yang 2002
$k_{R@SS}$[s⁻¹]	9.9·10⁻³	Dissociation of complexes R@S_cS	Tsoi and Yang 2002
k_{NTP}[s⁻¹M⁻¹]	0.113	Nucleotide polymerization in Oleic Vesicle	Mansy et al 2008
k_L [s⁻¹M⁻¹]	0.017	Lipid precursor conversion	Stage-Zimmermann and Uhlenbeck 1998
k_{in} [dm²s⁻¹]	7.6·10¹⁹	Oleic acid association to the membrane	Mavelli et al. 2008
k_{out} [dm²s⁻¹]	7.6·10⁻²	Oleic acid release from the membrane	Mavelli et al. 2008
P_P [cm s⁻¹]	4.2·10⁻⁹	Membrane permeability to lipid precursor	
P_{NTP} [cm s⁻¹]	1.9·10⁻¹¹	Membrane permeability to nucleotides	Mansy et al 2008
$P_W = P_S$	0.0	Membrane permeability to W and genetic staff	
P_{aq} [cm s⁻¹]	1.0·10⁻³	Oleic acid membrane permeability to water	Sacerdote and Szostak 2005

equal to that of the splicing reaction, catalyzed by Hammerhead ribozyme (Stage-Zimmermann and Uhlenbeck 1998). The kinetic behavior of different nucleotides is assumed to be the same and to k_{NTP} is assigned a single value derived by simulating data not shown the DNA template directed synthesis in fatty acid vesicles (Mansy et al. 2008). In the same way was also estimated the common value of the membrane permeability to activated nucleotides, while the membrane is assumed impermeable to genetic material. The kinetic constants of the membrane/aqueous solution lipid exchange are taken for oleic acid vesicles from a previous work (Mavelli et al. 2008) where they were obtained by simulating the competition between isotonic and osmotically swollen oleic vesicles (Chen et al. 2004). The amphiphile head area $\alpha_L = 0.3$ nm² and the osmotic tolerance $\varepsilon = 0.21$ are defined according to data reported in literature for oleic acid vesicles (Chen et al. 2004).

The only two parameters assigned arbitrarily are then the membrane permeability to the byproduct: $P_W = 0.0$ cm/s, based on the assumption that W is a charged species, and the permeability to the precursor: $P_P = 0.42·10⁻⁸$ cm/s, corresponding to the oleic acid membrane permeability to Arabitol (Sacerdote and Szostak 2005) and comparable to those of similar organic compounds.

A common simplifying assumption to both approaches is to consider the external concentrations of nucleotides (NTPs), lipid precursor (P), byproduct (W) and osmotic buffer (B) to be constant throughout the process time course thanks to an incoming flux of material in the reactor vessel: *continuous stirred tank reactor approximation*. For all these compounds, except the inert osmotic buffer B, the internal aqueous concentrations are zero at the beginning. The internal initial concentrations of the four nucleotides are set 0.5 mM in analogy with the experimental set up of the DNA

template directed synthesis in vesicles (Mansy et al. 2008). The same value was used also for the lipid precursor internal concentration while the W concentration is null. The internal concentration of the osmotic buffer is set 0.3 M, while the external concentration is properly adjusted in order to counterbalance the presence of the genetic material inside the vesicle and to start from isotonic conditions.

14.3 Theoretical Approaches

In this section the deterministic and the stochastic theoretical approaches will be briefly introduced and discussed.

14.3.1 *Deterministic Analysis*

The deterministic time evolution of the Ribocell is obtained by solving the following differential ODE set:

$$
\begin{cases}
\dfrac{dn_i^C}{dt} = N_A V_C \sum_{\rho=1}^{R} \left(b_{i,\rho} - a_{i,\rho} \right) r_\rho + N_A P_i S_\mu \left([X_i]^{Ex} - \dfrac{n_i^C}{N_A V_C} \right) \\
\qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2 \ldots N, \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad i \neq L \\
\dfrac{dn_L^C}{dt} = N_A V_C \sum_{\rho=1}^{R} \left(b_{i,\rho} - a_{i,\rho} \right) r_\rho - k_{in} S_\mu \dfrac{n_L^C}{N_A V_C} + k_{out} n_L^\mu \\
\dfrac{dn_L^\mu}{dt} = k_{in} S_\mu \dfrac{n_L^C}{N_A V_C} - k_{out} n_L^\mu \\
\dfrac{dV_C}{dt} = v_{aq} P_{aq} S_\mu \left(C_T^C - C_T^E \right)
\end{cases}
$$

where R is the number of internal metabolic reactions. The solution of the ODE set gives the time course of the average molecular numbers of species in the vesicle aqueous core (n_i^C, $i = 1,2,\ldots N$) and in the lipid membrane (n_L^μ) along with the volume of the core V_C.

Since all the intermediate complexes $R@S_cS_n$ are considered explicitly, the number of equations in the ODE set is strictly dependent on the length of the two ribozymes, while their base sequence does not affect the ribocell dynamics because we neglect difference in the nucleotides time behavior for simplicity. A specific program, called RIBOCELL, has been then developed in order to build up the correct ODE set on the basis of the R_{Pol} and R_{Lip} assumed sequences. This program can also solve the ODE set numerically by using the algorithms and imposing the following constraints:

$$n_i^C \geq 0 (i = 1, 2, ...N)$$

$$n_L^\mu \geq 0$$

$$1 - \varepsilon \leq \Phi \leq \sqrt[3]{2}$$

When the condition $\Phi > \sqrt[3]{2}$ is verified then the vesicle undergoes a division into two identical daughters and all the variables of the system state array **x** are then halved. On the other hand, the calculation is stopped when $\Phi < 1 - \varepsilon$ because an osmotic crisis occurred or if the two inequalities $n_i^C \geq 0 (i = 1, 2, ...N)$ and $n_L^\mu \geq 0$ are not satisfied).

14.3.2   Stochastic Simulations

The stochastic approach to chemical kinetics associates a density probability function w_j instead of a deterministic rate to each event occurring in a chemically reacting system (Van Kampen 1981, Gillespie 1992). $w_j\, dt$ gives the probability that the j-th event, i.e. both a reaction or a transport process, can take place in the next infinitesimal time interval $[t,\ t + dt]$ and for this reason is called the event propensity density probability. All these functions depend on the system state, i.e. on the molecular populations of the involved species present in the system, as reported in Table 14.1. In fact, this approach takes into account explicitly the discrete nature of the real molecular system state on the contrary of the deterministic one. The average time evolution of the chemically reacting system has to be determined by solving a finite-difference partial differential equation called the Kinetic Master Equation that gives the rate of change of the system state density probability[3] in terms of the w_j functions. Without going into technical details[4], the KME solutions give also insights on the deviations from the system average time behavior, but it can be solved analytically only in few cases (McQuarry 1975). In the middle of 1970s, Gillespie presented a general Monte Carlo procedure that exactly simulates the stochastic time evolution of a homogeneous chemically reacting system (Gillespie 1976, 1977, 2007). For high populated reacting systems, this method can be used as a robust Monte Carlo algorithm for integrating the ODE set associated with the kinetic mechanism although it is not very efficient when compared with the ODE numerical methods (Gillespie 2009). On the other hand, in the case of scarce populated vesicles, it can be used to elucidate the role of intrinsic stochasticity and state discreteness of the real chemically reacting systems.

---

[3] The system state density probability $\wp(\mathbf{x}, t)$ is defined so that the product $\wp(\mathbf{x}, t)dt$ gives the probability to find the system in the state **x** at the time interval $[t,\ t + dt]$.

[4] For further details on the Stochastic Kinetic Theory the reader can refer to the van Kampen's or Gillepie's books (van Kampen 1981; Gillespie 1992).

In recent years, this approach is getting a growing attention in particular for the description of biochemical systems (McAdams and Arkin 1997; Lu et al. 2004; Samoilov et al. 2005; Ao 2005; Lecca 2007) where the concentration of bio-molecules can be very low. Different improvements have been also proposed for the original procedure (Gillespie 2007; Li et al. 2008). In this framework, we have recently presented and tested a computational platform called REACTOR (Mavelli and Piotto 2006), specifically developed to study stochastically homogeneous systems where chemical reactions can take place along with self-assembly processes (Mavelli and Stano 2010). This program has been also improved to deal with volume-changing, globally heterogeneous, compartmentalized reacting systems and to make it suitable for designing and testing realistic proto-cell models and possible explanations for their spontaneous emergence in pre-biotic conditions (Mavelli and Ruiz-Mirazo 2007, 2010; Ruiz-Mirazo and Mavelli 2007, 2008; Mavelli et al. 2008).

14.4   Results and Discussion

In this section the results obtained with both the theoretical approaches are presented and discussed. Fast deterministic calculations are used to scan the parameter space in order to find the range of the Ribocell stability, while time consuming stochastic simulations are carried out in order to test the Ribocell robustness to random fluctuations. Three different long time regimes are in principle expected: (1) the stationary grow-division regime (GD), i.e. a stable steady state where the protocell grows in size and divides keeping the same radius after any division; (2) the Ribocell burst due to an osmotic crisis (OC) when the core volume V_C increases more rapidly than the membrane surface S_μ; (3) the death for dilution regime (DD) if the membrane surface S_μ increases rapidly and the protocell divides before the genome duplication is completed.

14.4.1   Deterministic Curves

In all the studied cases, we start from a single 100 nm-radius spherical vesicle containing only two pairs of ribozymes in form of duplex R_cR_{Lip} and R_cR_{Pol} and the calculations are carried out at room temperature using the parameter values reported in Table 14.2. The Ribocell deterministic time evolution is reported in Fig. 14.3, where all the three expected regimes are shown as a function of k_L the kinetic constant of the precursor conversion cathalized by the R_{Lip} rybozyme.

In particular, when k_L is equal to the reference value 1.7×10^{-2} s^{-1}M^{-1}, that is the catalytic constant k_{cat} estimated for the splicing reaction catalyzed by the Hammered ribozyme (Stage-Zimmermann and Uhlenbeck 1998), an osmotic crisis takes place (OC), see Fig. 14.3a. Since the rate of lipid production is very slow the membrane

surface S_μ remains practically constant while the vesicle core volume rapidly increases as the self-catalytic genome replication becomes to speed up. This is due to the increase of the internal concentration of byproducts W, released during the nucleotides polymerization and it is confirmed by the decreasing trend of the relative surface Φ. When Φ crosses the lower stability limit of the oleic acid membrane: $(1 - \varepsilon)$ = 0.79 (lower dashed line), the Ribocell bursts and the calculations are stopped.

Figure 14.3b shows the case of a growth-division stationary regime. In this case V_C raises along with S_μ and Φ and as soon as the reduced surface Φ crosses the upper stability limit ($\Phi = 2^{1/3} = 1.26$) the mother cell divides halving all its internal molecular content with an identical daughter. After few generations, a stationary cyclic regime is attained at a larger radius (112.4 nm) than the initial one, and this value remains constant after any further division. In fact, at the steady regime, the Ribocell exhibits a perfect synchronization between the genome self-replication and the vesicular compartment reproduction. During the growth phase, a synchronized increase of all the internal molecular content along with the membrane surface and the core volume takes place and the Ribocell doubles all its constituents before dividing ($\Phi = 1.26$). It is worthwhile to remark that all the metabolite internal concentrations remain constant throughout the growth phase (data not shown) and this guaranties the Ribocell to be in a genuine steady state fed by the substrates incoming flux from the external environment.

For k_L values greater equal to $1.7 \times 10^{+3}$ s^{-1}M^{-1} the GD steady state is always reached as shown by Table 14.3. The increasing efficiency of the R_{Lip} catalytic activity reflects in decreasing the life time cycle, the radius and the genetic strand amount of the Ribocell at steady GD regime respectively. Actually, when k_L is greater equal to $1.7 \times 10^{+6}$ the death for dilution phenomenon takes place. In fact in this case the surface area of the lipid membrane increases too much quickly and the synchronization with the genome replication occurs only for a total molecular average amount of ribozymes R_{Pol} and R_{Lip} equal to 0.081 after 20 generations, see Fig. 14.3c. Although the value obtained is physically reasonable, since it represents a cell population average, it means that in the most optimistic case only one protocell on 37 can encapsulate all the three ribozymes needed for its self-reproduction. The other 36 must completely free from any RNA polymer. Of course, since this scenario is statistically implausible it should be clear that the deterministic analysis tends to overestimate the range of stability of the Ribocell.

Table 14.3 Deterministic results for different R_{Lip} catalytic efficiency k_L: all the reported data are determined after 20 generations

k_L (s^{-1}M^{-1})		Δt_{div} (days)	Radius (nm)	Total Number or RNA strands	R_{Lip}	$_cR_{Lip}$	R_{Pol}	$_cR_{Pol}$
					(%) Fraction			
1.7×10^{-2}	OC							
$1.7 \times 10^{+2}$	OC							
$1.7 \times 10^{+3}$	GD	127.8	157.3	41384	28.43	28.43	26.13	17.01
$1.7 \times 10^{+4}$	GD	81.1	112.4	790	26.08	26.08	25.82	22.03
$1.7 \times 10^{+5}$	GD	80.3	110.7	8	25.00	25.00	25.00	25.00
$1.7 \times 10^{+6}$	DD	79.8	110.7	0				

a

b
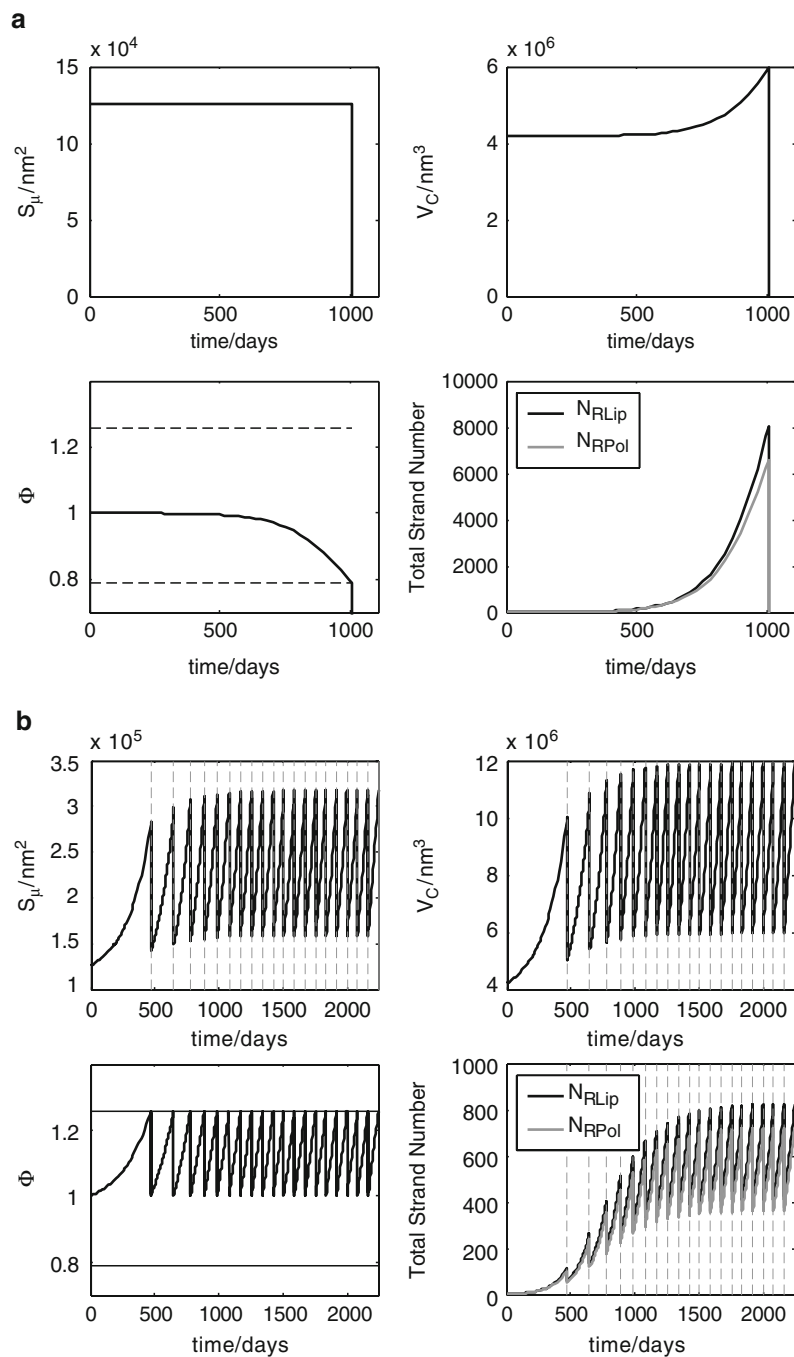
Fig. 14.3 Time Evolution of the Ribocell depending on the catalytic efficiency k_L of the rybozime R_{Lip}; see text for further details. (**a**) Osmotic Crisis: $k_L = 1.7 \times 10^{-2}$ s^{-1}M^{-1}; (**b**) Growth-Division regime: $k_L = 1.7 \times 10^{+4}$ s^{-1}M^{-1}
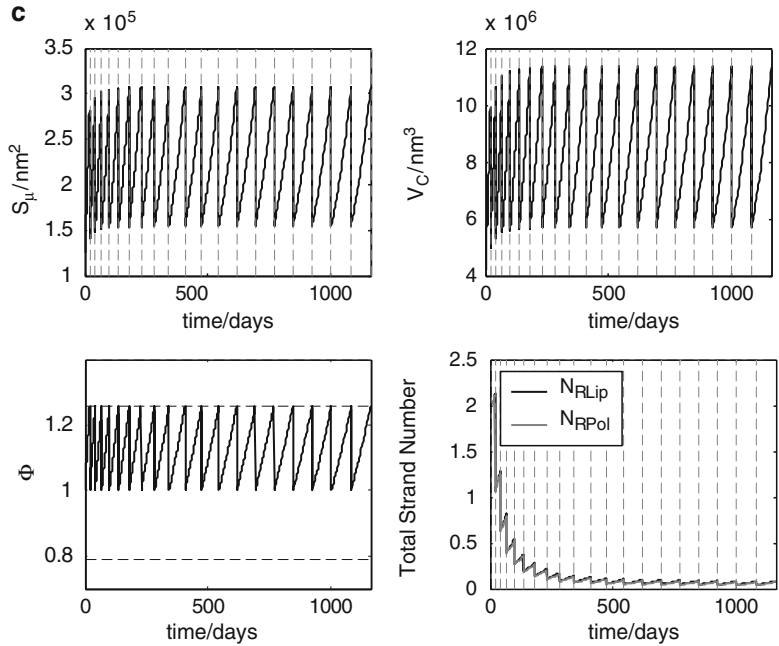
Fig 14.3 (continued) (c) Death for Dilution: $k_L = 1.7 \times 10^{+6}$ s^{-1}M^{-1}
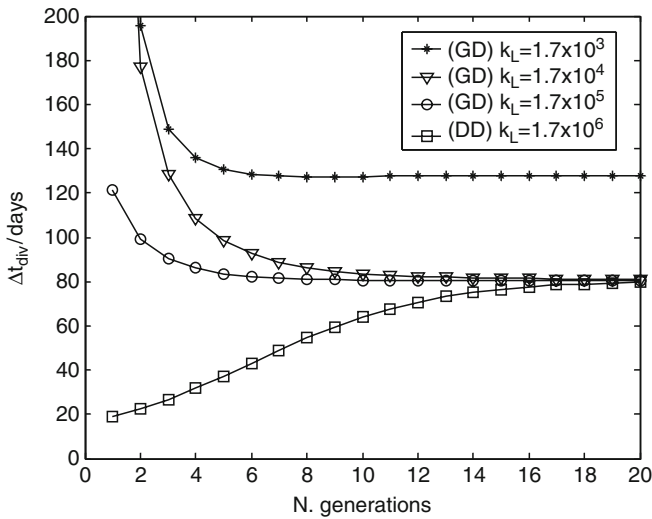
Fig. 14.4 Ribocell life time Δt_{div} against the number of generations for different k_L values

Finally in Fig. 14.4, the Ribocell time life Δt_{div}, i.e. the interval of time between two subsequent divisions, are reported for different k_L values against the number of generations. This plot shows that in all the studied cases after 20 generations a stationary state is attained.

14.4.2   Simulation Data

Stochastic simulation are carried out for the two cases $k_L = 1.7 \times 10^{+4}$ s^{-1}M^{-1} and $k_L = 1.7 \times 10^{+5}$ s^{-1}M^{-1} starting from 50 spherical vesicles of 100 nm radius with a genetic staff composition near to the steady state regime as reported in Table 14.3. As for the deterministic calculations, after each division only one of the two daughters is kept, in order to perform simulations at a constant number of protocells and to avoid a huge increase of the running time.

The simulated time course of Φ (gray line with error bars) reported in Fig.14.5a clearly shows that at the beginning the average time behavior of the entire Ribocell population is statistically equivalent to the deterministic time evolution (black line), however, as soon as the division time is approached, a dispersion of the time behavior of any single Ribocell takes place. This is partially due to random fluctuations in the reacting events, although the main effect is due to the stochastic distribution of the genetic staff between the two daughters after any division. In fact, while the random displacements of the internal nucleotides concentrations are in some degree counterbalance by molecular transport processes from the external environment, in the case of genetic molecules, since is has been assumed they have a negligible membrane permeability, concentration displacements after division can really differentiate the cell population. Moreover, these differences in the genetic staff amount can be amplified by the self-catalytic genome replication determining divergent time behaviors as reflected in the diverse division time.

This effect is more pronounced when $k_L = 1.7 \times 10^5$ s^{-1}M^{-1} as clearly shows by Fig. 14.5b. In this case the disagreement between the stochastic simulation and the deterministic behavior appear evident from the beginning. This can be ascribed to the small amount of the initial genetic staff that is of two orders of magnitude lower than in the previous case. Therefore the cell is much more sensitive to the random fluctuations of reacting processes and this produces a large dispersion of the division time Δt_{div}. On the other hand the simulated trend of the internal concentration of the precursor P is in agreement with deterministic curve for both the studied cases, while the byproduct W seems to be underestimated by the simulations. This can be ascribed to the presence of some self-replicating, empty and inert vesicles in which the concentration of W remains lower than in the genuine Ribocells. In fact, since these vesicles do not contain R$_{Pol}$ they can not produce W as a byproduct of the nucleotide polymerization. Figure 14.6 reports the composition of the cell population against generations, i.e. after each division, showing that at long time only a population fraction around 70% is made of genuine Ribocells.
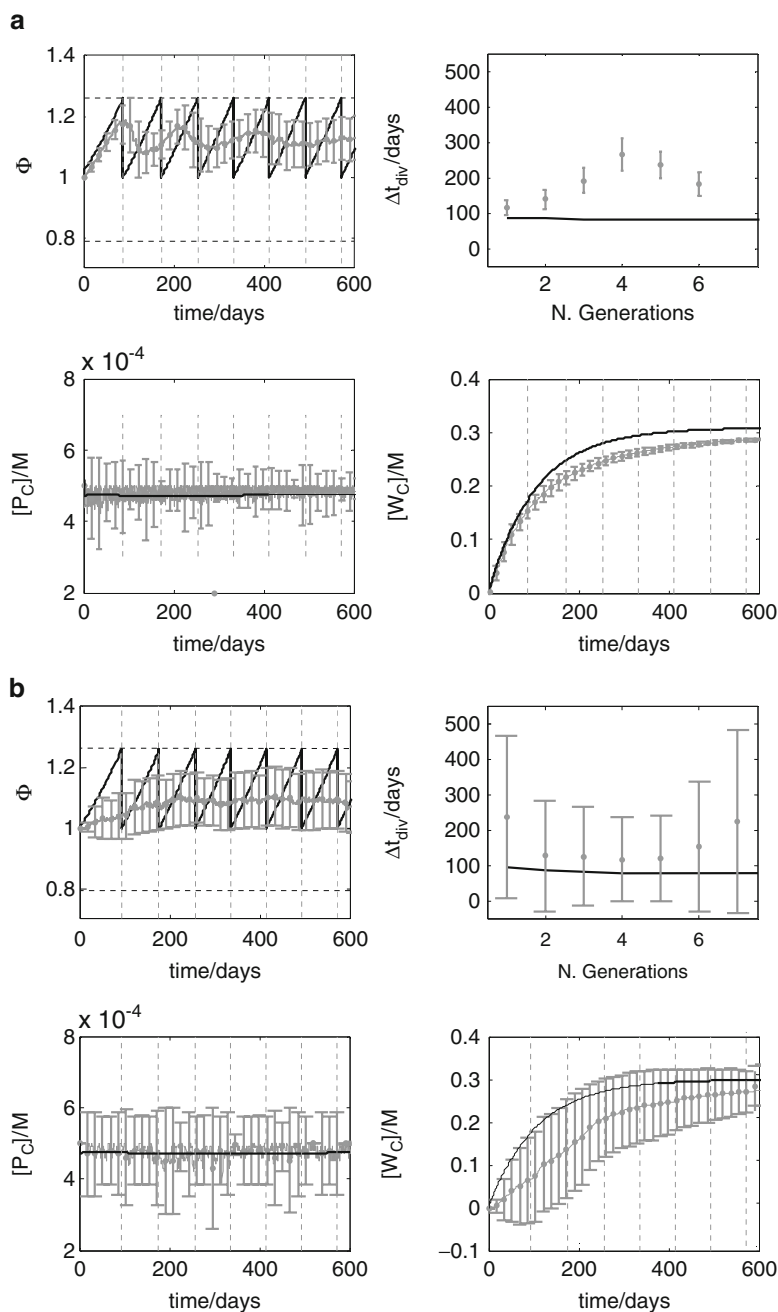
Fig. 14.5 Comparison between deterministic curves (*black lines*) and stochastic simulation data (*gray lines with error bars*) of the Ribocell time behavior obtained setting (**a**) $k_L = 1.7 \times 10^4$ s^{-1}M^{-1} and (**b**) $k_L = 1.7 \times 10^5$ s^{-1}M^{-1}. Vertical dashed lines are the deterministic division times
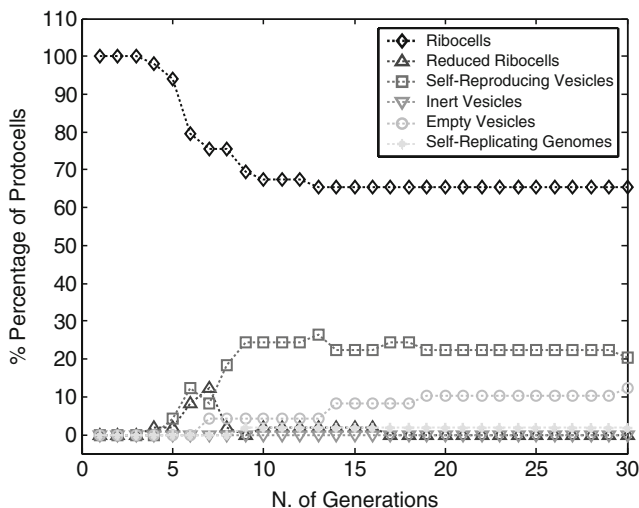
Fig. 14.6 Different composition of the Protocell population against the generation number: $k_L = 1.7 \times 10^5$ s^{-1}M^{-1}

14.5 Conclusions

In this chapter a detailed kinetic model for the minimal RNA based-cell: the Ribocell has been presented and discussed. Two possible approaches to the study of its time evolution have been illustrated and compared. In particular, the deterministic analysis allows a more efficient exploration of the kinetic parameter space, while stochastic simulations better describe the chemical process at molecular level taking into account the role of fluctuations. Keeping in mind that the kinetic parameters used herein for both the analyses were all taken from experimental data reported in literature at room temperature, the emergence of a spontaneous synchronization between the genome self-replication and the lipid membrane self-reproduction is quite a surprising result, within the approximations and the kinetic parameters used. However, the scale of the protocell time life is really too large in order to be able to consider the Ribocell as a feasible, although still hypothetical, minimal cell model. The high values of division times, about 100 of days, compared to that of real cells, can be ascribed mainly to the low value used for the dissociation constant k_S of RNA dimers at room temperature. In fact, an increase of the work temperature could greatly increase the efficiency of reproduction as shown recently by the process of self-catalyzed replication of RNA strands (Lincoln and Joyce 2009) and, on the other hand, fatty acid vesicles are quite stable up to 90°C (Mansy and Szostak 2008).

Therefore, the Ribocell time behavior needs a more deep theoretical analysis and some improvements of the in silico model are also necessary. First of all, kinetic parameters at higher temperature need to be estimated and also the kinetic behavior

of nucleotides and RNA strands with a diverse base sequence has to be differentiated. Conversely, some conclusions can be drawn from these preliminary results. The random nature of reacting events (intrinsic stochasticity) can highly differentiated the time course of each single protocell in the population, since the effect of fluctuations is enlarged by the autocatalytic character of genome replication. Moreover, another source of time course dispersion is the casual distribution of the cell internal content after each division (extrinsic stochasticity). Also in this case, displacement from the deterministic equality of the genetic staff amount in both the daughter cells is amplified by the nature of the internal metabolism. However, in the case of Ribocells, while intrinsic stochasticity can determine equivalent behaviors with different time scales, the extrinsic randomness can produce completely different outcomes bringing to the death for dilution if a complete segregation of ribozymes in diverse lipid compartments takes place. As a consequence of this, stochastic simulations give a better description of a ribocell population time course taking into account both the intrinsic and extrinsic stocasticity and also the discreteness of the reacting system state at molecular level.

Finally, it is worthwhile to remark that extrinsic stocasticity also effects the initial conditions and starting from an initial ideal population of identical protocells, in terms of sizes and molecular content, merely circumvents this problem. Therefore, the theoretical analysis has another level of complexity that is due to the uncertainty of the starting conditions while the experimental implementation of the future ribocells will require an appropriate and reproducible preparation procedure for the system set up.

References

Ao P (2005) Metabolic network modelling: including stochastic effects. Comput Chem Engin 29:2297–2303

Chen IA, Roberts RW, Szostak JW (2004) The emergence of competition between model protocells. Science 305:1474–1476

Christensen U (2007) Thermodynamic and kinetic characterization of duplex formation between 2'-O, 4'-C-Methylene-modified Oligoribonucleotides. DNA RNA Biosci Rep 27:327–333

Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22:403–434

Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81:2340–2361

Gillespie DT (1992) Markov processes: an introduction for physical scientists. Academic Press, San Diego

Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58:35–55

Gillespie DT (2009) Deterministic limit of the stochastic chemical kinetics. J Phys Chem 113:1640–1644

Lecca P (2007) Simulating the cellular passive transport of glucose using a time-dependent extension of Gillespie algorithm for stochastic pi-calculus. Int J Data Mim Bioinf 1:315–336

Li H, Cao Y, Petzold LR, Gillespie DT (2008) Algorithms and software for stochastic simulation of biochemical reacting systems. Biotechnol Prog 24:56–61

Lincoln TA, Joyce GF (2009) Self-sustained replication of an RNA enzyme. Science 323:1229–1232

Lu T, Volfson D, Tsimring L, Hasty J (2004) Cellular growth and division in the Gillespie algorithm. Syst Biol 1:121–128

Luisi PL (1998) About various definitions of life. Orig Life Evol Biosph 28:613–622

Luisi PL (2003) Autopoiesis: a review and a reappraisal. Naturwissensch 90:49–59

Mansy SS, Szostak JW (2008) Thermostability of model protocell membranes. PNAS 105:13351–13355

Mansy SS, Schrum JP et al (2008) Template directed synthesis of a genetic polymer in a model protocell. Nature 454:122–126

Mavelli F, Lerario M, Ruiz-Mirazo K (2008) 'Environment': a stochastic simulation platform to study protocell dynamics. In: Arabnia HR et al (eds) BIO-COMP'08 proceedings, vol II. CSREA Press, New York

Mavelli F, Piotto S (2006) Stochastic simulations of homogeneous chemically reacting systems. J Mol Struct 771:55–64

Mavelli F, Ruiz-Mirazo K (2007) Stochastic simulations of minimal self-reproducing cellular systems. Phil Trans Royal Soc B 362:1789–1802

Mavelli F, Ruiz-Mirazo K (2010) Environment: a computational platform to stochastically simulate self-replicating reacting vesicles. Phys Biol 7:036002

Mavelli F, Stano P (2010) Kinetic models for autopoietic chemical systems: role of fluctuations in homeostatic regime. Phys Biol 7:16010

McAdams HH, Arkin AP (1997) Stochastic mechanisms in gene expression. Proc Natl Acad Sci U S A 94:814–819

McQuarry DA (1975) Probability theory and stochastic processes. In: Henderson D (ed) Physical chemistry, an advanced treatise XIB. Academic Press, New York

Ruiz-Mirazo K, Mavelli F (2007) Simulation model for functionalized vesicles: lipid-peptide integration in minimal protocells. In: Almeida F (ed) Advances in artificial life. Springer, Berlin

Ruiz-Mirazo K, Mavelli F (2008) On the way towards basic autonomous systems: stochastic simulations of minimal lipid-peptide cells. Biosystems 91:374–387

Sacerdote MG, Szostak JW (2005) Semipermeable lipid bilayers exhibit diastereo-selectivity favoring ribose. PNAS 102:6004–6008

Samoilov M, Plyasunov S, Arkin AP (2005) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. Proc Natl Acad Sci USA 102:2310–2315

Stage-Zimmermann TK, Uhlenbeck OC (1998) Hammerhead ribozyme kinetics. RNA 4:875–889

Szostak JW, Bartel DP and Luisi PL (2001) Synthesizing life. Nature 409:387–390.

Tsoi PY, Yang M (2002) Surface plasmon resonance study of human polymerase β binding to DNA. Biochem J 361:317–325

Van Kampen NG (1981) Stochastic processes in physics and chemistry. Elsevier, Amsterdam

Chapter 15
Evolvability and Self-Replication of Genetic Information in Liposomes

Tomoaki Matsuura, Norikazu Ichihashi, Takeshi Sunami, Hiroshi Kita, Hiroaki Suzuki, and Tetsuya Yomo

Abstract To realize the minimal cell by a bottom-up approach, increasingly complex biochemical reactions are being encapsulated in lipid vesicles (liposomes). Here, we describe the encapsulation of one of the general properties of living organisms into liposomes, i.e., replication of genetic information with self-encoded replicase, and evolvability. We also discuss the possibility of realizing a minimal cell that can proliferate autonomously.

15.1 Top-Down and Bottom-Up Approaches Toward Minimal Cell Synthesis

Current living organisms show high degrees of complexity. For example, *Escherichia coli*, a relatively simple organism, is known to encode more than 4,000 genes (Riley et al. 2006), the products of which interact with each other constituting a highly complex network (Butland et al. 2005; Rual et al. 2005; Uetz et al. 2000). Living systems exhibit a number of specific properties, including self-reproducibility, evolvability, and adaptability. To date, it has not been possible to create artificial systems

---

T. Yomo (✉)
Graduate School of Frontier Biosciences, Osaka University, 1-5 Yamadaoka,
565-0871 Suita, Osaka, Japan
e-mail: yomo@ist.osaka-u.ac.jp

T. Yomo, T. Matsuura, N. Ichihashi, and H. Suzuki
Department of Bioinformatics Engineering, Graduate School of Information Science and
Technology, Osaka University, 1-5 Yamadaoka, 565-0871 Suita, Osaka, Japan

T. Yomo, T. Matsuura, N. Ichihashi, H. Suzuki, T. Sunami, and H. Kita
Exploratory Research for Advanced Technology (ERATO), Japan Science and
Technology Agency (JST), 1-5 Yamadaoka, 565-0871 Suita, Osaka, Japan

that possess all of these properties by mixing the constituent parts, indicating a lack of knowledge regarding the design of living systems. On the other hand, comprehensive analysis indicated that large fractions of genes can be knocked out without affecting the fitness of various organisms (Baba et al. 2006; Giaever et al. 2002; de Visser et al. 2003; Kitano 2004). Single gene knockout experiments of the yeast genome indicated that more than 80% of the genes are not essential, and a similar value was obtained in *E. coli*. These observations raise the question of the minimum number of components and the requirements for a system to be considered alive, i.e., to exhibit properties including self-reproducibility and evolvability.

A number of groups have proposed synthesis of minimal cells to address the above questions (Deamer 2005; Szostak et al. 2001; Forster and Church 2006). To achieve this, two contrasting approaches have been reported (Luisi 2002), i.e., the top-down and bottom-up approaches (Fig. 15.1). The top-down approach aims to realize the minimum set of genes that will allow the organism to grow by continuously reducing the size of the genome (Hutchison et al. 1999; Hashimoto et al. 2005; Yu et al. 2002). The top-down approach provides information and insights in addition to the minimum size of the genome. For example, *E. coli* with a 15% reduced genome was found to exhibit a significantly reduced error rate and higher electroporation efficiency (Posfai et al. 2006). Unexpected properties seem to emerge in these top-down approaches and further investigations are expected to provide insight into the behavior of minimal cells.

The top-down approach has contributed significantly to the realization that extant living systems can be simplified to certain extent, and is also expected to identify the necessary conditions for living systems. The bottom-up approach, the counterpart of the top-down approach, aims to construct a biological system with
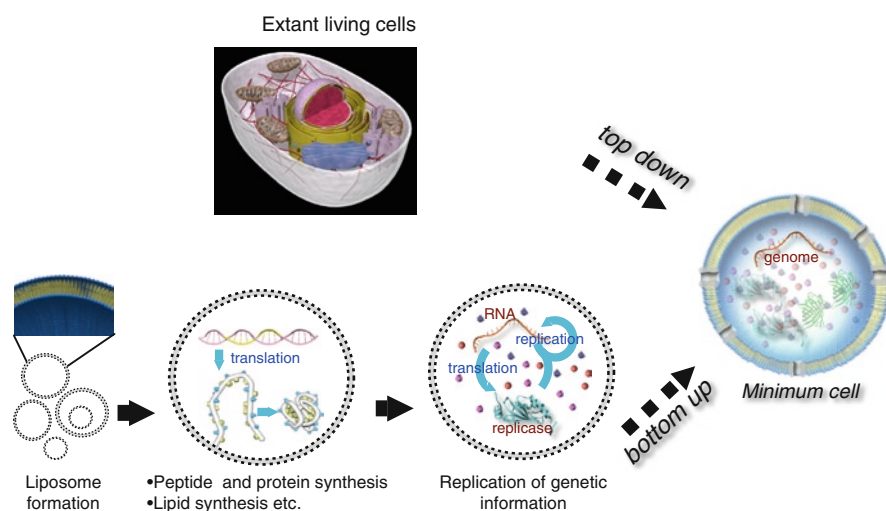
Fig. 15.1 Top-down and bottom-up approaches toward synthesis of a minimal cell. See text for details

defined components, and is expected to identify sufficient conditions for living systems. As simple mixing of the components does not lead to the emergence of living systems, the bottom-up approach aims to elucidate the design principle of biological systems during the construction process (Benner and Sismour 2005; Simpson 2006). This knowledge will not be obtained only by investigating the properties of the various elements of the systems. While bottom-up approaches are used to understand various biological systems, we have focused on the approaches that primarily involve the use of lipid vesicles for synthesizing minimal cells.

15.2 Achievements by the Bottom-Up Approach Toward Minimal Cell Synthesis

The ultimate goal of the bottom-up approach is the synthesis of a living cell. In practice, progress toward this goal has been proceeding in discrete steps (Fig. 15.1). While there are various definitions of a minimal cell (Luisi 2002; Luisi et al. 2006), the least requirements are the presence of lipid vesicles and the presence of internal reactions, such as gene replication and metabolism. A number of groups have been assembling the elements that partially fulfill the properties of a living system. For example, it was shown to be possible to generate artificial lipid vesicles (liposomes) of the same size as small bacteria from amphiphilic molecules (Bangham and Horne 1964). Liposomes were also shown to be capable of autocatalytic growth, and to even be able to undergo repeated cycles of growth and division (Hanczyc et al. 2003; Takakura and Sugawara 2004). Various types of biological reaction (nucleic acid and protein synthesis, integration of pore proteins, and two-stage genetic cascade reaction) have been conducted successfully within the environment provided by liposomes (Noireaux and Libchaber 2004; Nomura et al. 2003; Yu et al. 2001; Oberholzer et al. 1995; Ishikawa et al. 2004). These studies represent significant steps toward assembly of an artificial minimal cell. However, one of the fundamental reactions, self-replication of the genetic information, has not been achieved. In the following section, we describe our recent progress in this respect.

15.3 Replication of Genetic Information in Liposomes

All living organisms replicate their genome using a set of self-encoded enzymes, and can therefore be designated as self-replicating systems (Paul and Joyce 2004) (Fig. 15.2a). This self-replication is different from reactions such as polymerase chain reaction (PCR), which also enables the exponential amplification of genes. In the case of PCR, the catalyst (replicase) that replicates the genes is added externally, while in the self-replication catalysts are supplied internally. The replicase, decoded by the translation machinery, is encoded on the gene that is being replicated.
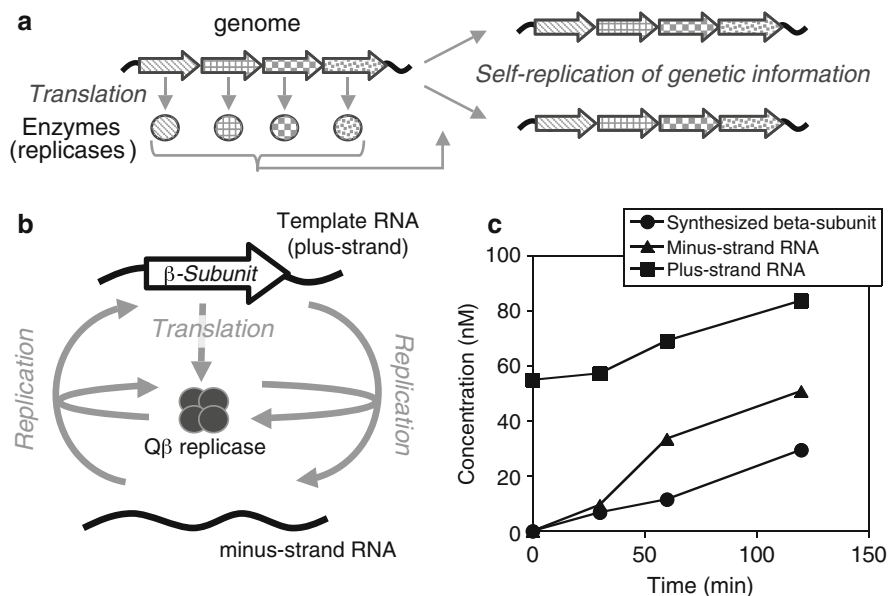
Fig. 15.2 Self-replication of genetic information. (**a**) In extant living cells, the genome encodes a large number of proteins. Proteins that are decoded from the genome consequently generate a copy of the genome. As the genome is replicated by the self-encoded proteins, this process can be termed self-replication of the genetic information. (**b**) The simplified self-replication system constructed in the present study. The template RNA (plus-strand) plays two roles: it acts as a template for RNA replication reaction by Qβ replicase and encodes the β-subunit (catalytic subunit) of Qβ replicase. When the template RNA is added to the in vitro translation system, translated β-subunit assembles with all three host proteins (ribosomal protein S1, elongation factor Tu (EF-Tu), and Ts (EF-Ts)) that are originally included in the in vitro translation system to generate mature Qβ replicase. The replicase then copies the RNA that encodes itself. (**c**) Time course of the reaction shown in (**b**)

Hence, the self-replicating system can also be referred to as a self-encoding system (Kita et al. 2008).

We assembled a simplified self-replication system consisting only of defined components, one template RNA sequence as a gene and an in vitro translation system reconstructed from purified translation factors (Shimizu et al. 2001) as the machinery for decoding (Fig. 15.2b), and encapsulated it in liposomes (Kita et al. 2008). In the in vitro translation system, the β-subunit (catalytic subunit) of Qβ replicase (Haruna and Spiegelman 1965; Hosoda et al. 2007) is synthesized from the template RNA (Fig. 15.2b). The replicase then replicates the template RNA used for its production. Figure 15.2c shows the time course of the self-replication reaction where the increases in the plus and minus strand RNAs and the replicase concentrations can be seen. To monitor the self-replication reaction by fluorescence, we introduced the β-galactosidase sequence into the minus strand (Fig. 15.3). After minus strand synthesis, β-galactosidase is translated and produces the green fluorescent molecule CM-fluorescein. When this reaction was carried out in
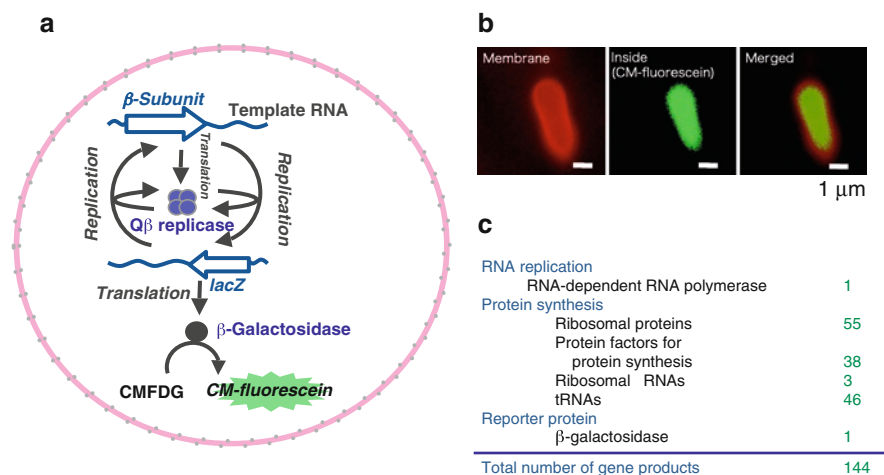
Fig. 15.3 Self-replication of genetic information in liposomes. (**a**) Schematic representation of the self-replication reaction modified by insertion of the *lacZ* gene into the minus strand. The Qβ replicase β-subunit was encoded on the plus strand RNA, and β-galactosidase was encoded on the minus strand RNA (complement of the plus strand RNA). Nonfluorescent CMFDG was hydrolyzed by β-galactosidase to yield green fluorescent CM-fluorescein. (**b**) The reaction within the liposomes observed under fluorescence microscopy. *Left*, red image (membranes); *Middle*, green image (hydrolyzed CMFDG); *Right*, overlay of left and middle images. Scale bar indicates 1 μm. (**c**) The self-replication system constructed consists of only 144 gene products

liposomes, green fluorescence was observed inside the liposomes (Fig. 15.3). Overlaying the images of membrane (red) and indicator of the internal reaction (green) clearly showed that the reaction occurred inside the liposomes. The reaction system consists of 144 gene products: 3 rRNAs, 46 tRNAs, 55 ribosomal proteins, 38 proteins for protein synthesis, one reporter protein, and one RNA replicase. This number is comparable with the proposed minimal complement of cellular life (approximately 150), which can grow with only small molecule nutrients (Deamer 2005; Luisi et al. 2006).

15.4 Effects of Compartmentalization of the Self-Replication Reaction

In all living cells, cytosolic biochemical reactions are enclosed within a lipid vesicle. The number of molecules encapsulated is substantially reduced in such small volumes, and stochastic behavior of the reaction should be apparent. This raises the question of whether we could see the effect of compartmentalization of the self-replication reaction in small vesicles. Thus, the time course of the reaction in liposomes was determined using a fluorescence activated cell sorter (FACS) to investigate the statistical nature and the dynamics of the reaction in liposomes
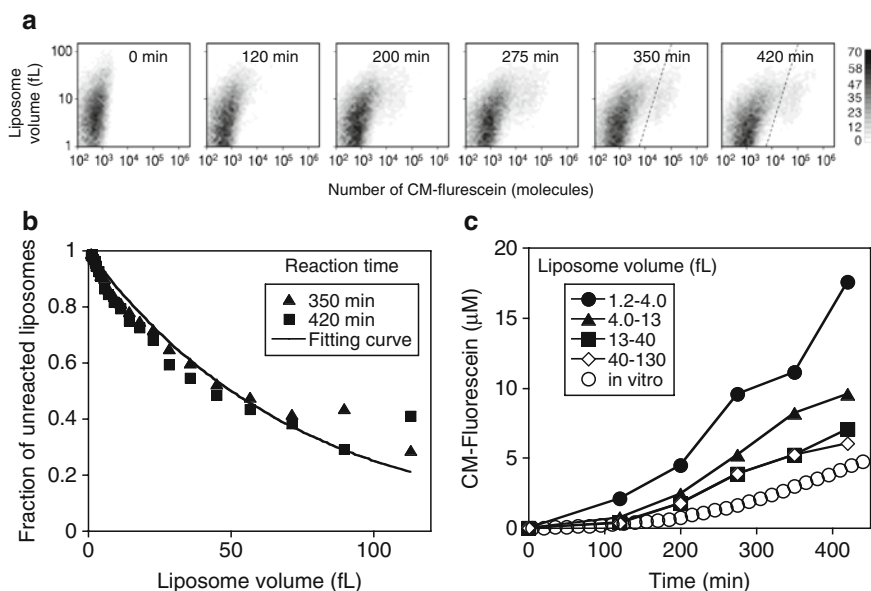
Fig. 15.4 Dynamics and stochastic nature of self-replication in liposomes. (**a**) Time course of the reaction analyzed by FACS. Contour maps of the relationship between the number of CM-fluorescein (*horizontal*) and internal aqueous volume (*vertical*) are shown. Liposomes present at the right (*left*) of the dashed lines at 350 and 420 min were defined as reacted (non-reacted) liposomes. (**b**) Dependency of the fractions of non-reacted liposomes ($P_{\text{non-react}}$) on the liposome internal volume. The solid line shows the results of curve fitting with equation $P_{\text{non-react}} = e^{-\gamma V}$, where V (fL) is the liposome internal volume, and γ (events/fL) is the average frequency of successful replication reaction per fL. (**c**) Time courses of the reaction in liposomes with different internal volumes and in vitro. For details, see Kita et al. (2008)

(Fig. 15.4). Below, we describe the effects of compartmentalization observed in our experiments.

Figure 15.4a shows the time course of the self-replication reaction in liposomes. The vertical axis shows the internal volume of each liposome ranging from 1 to 100 fL, which is typical for multilamellar liposomes prepared by the freeze-drying method as described previously (Hosoda et al. 2008; Sato et al. 2006; Sunami et al. 2006). The distribution of liposomes moved toward the right over time, indicating an increase in the number of fluorescent products encapsulated inside the liposomes, and thus the occurrence of the self-replication reaction. Liposomes were assigned as reacted when the self-replication reaction succeeded at least once inside, equivalent to the production of more than a single molecule of β-galactosidase. Defining the reacted liposomes as those on the right of the dashed lines of Fig. 15.4a, reactions occurred in only 13% of liposomes at the end of the reaction. This fraction did not change between 350 and 420 min, indicating the reaction no longer proceeded at these time points. Furthermore, the fractions of non-reacted liposomes ($P_{\text{non-react}}$) at 350 and 420 min were dependent on the liposome internal volume (Fig. 15.4b). Assuming that the number of successful self-replication reactions per unit volume is

a Poisson process, the dependency can be described as: $P_{\text{non-react}} = e^{-\gamma V}$, where V (fL) is the liposome internal volume, and γ (events/fL) is the average frequency of successful replication reactions per fL. We found that this equation fitted well with the data and obtained $\gamma = 0.014$ (events/fL). Thus, only 1.4% of the liposomes with a volume of 1 fL exhibited the green fluorescent signal, while the others did not. This low reaction efficiency caused quantization of the reaction: the reaction succeeded in only a fraction of the liposomes. Furthermore, the smaller the liposomes, the lower the probability that the complete self-replication reaction will occur. This observation is one of the consequences of compartmentalization. While there are many possible reasons for the low reaction frequency (Ichihashi et al. 2008), including degradation of RNAs, inactivation of enzymes, accumulation of inhibitory products, or competition between translation and replication, the established system is the starting point and further optimization is expected to lead to the emergence of highly efficient self-replication systems.

The liposomes prepared by the freeze-drying method vary in their internal volume ranging from 1 to 100 fL (Hosoda et al. 2008; Sato et al. 2006; Sunami et al. 2006). From the FACS data shown in Fig. 15.4a, we thus determined the reaction time course in liposomes with different internal volumes (Fig. 15.4c). We first found that the reaction proceeds faster in smaller liposomes. Successful self-replication reaction is equivalent to the production of β-galactosidase, and a single successful reaction yields a fixed number of β-galactosidase molecules. Due to the low reaction efficiency, $\gamma = 0.014$ (events/fL), the number of successful reactions (events) in liposomes smaller than 71.4 (= 1/0.014) fL was almost 1 or 0. Therefore, the number of β-galactosidase molecules produced should be the same in these liposomes. As a consequence, the same number of CM-fluorescein molecules was produced regardless of the volume. This gave rise to the observation that smaller liposomes exhibit a faster increase in CM-fluorescein concentration. On the other hand, the reaction kinetics in larger liposomes became increasingly similar to that in vitro (Fig. 15.4c), indicating that the *in liposome* self-replication system was functioning as designed.

The result of rate acceleration in small liposomes gives insight into the minimal cell synthesis. When the number of enzymes is limited in the minimal cells, the smaller cells should benefit from being small because smaller cell size is associated with faster increases in product concentration, which results in faster response to for example to environmental changes.

15.5   Evolvability in Liposomes

All living organisms have the potential to evolve, and thus evolvability is a common property of living systems. Evolution proceeds by consecutive rounds of mutation and selection, and as a consequence the fittest individuals are selected from among the population. For the self-replication system described above as well as extant cells to evolve, the genomes (or genes) must be encapsulated into a compartment

and only a small number of them should be present in each compartment. These conditions are required to link the phenotype with the genotype; the translated replicase needs to contribute to amplifying only the gene encoding itself. Assume a pool of mutant genes, where minor fractions of the genes encode replicases with higher activities than the others. Without compartmentalization, the highly active replicases will amplify other genes as well as their own (Matsuura et al. 2002), and thus the enrichment of genes encoding highly active replicases among the population will not occur. On the other hand, with compartmentalization, the highly active replicase amplify only the gene that encode itself, and thus the enrichment of gene encoding highly active replicase among the population is achieved. Thus, compartmentalization of the genome is necessary for realizing the evolvability. While number of strategies has been developed to realize the evolvability (Matsuura and Yomo 2006), here in this section, we describe the experimental demonstration of evolvability using liposomes (Sunami et al. 2006).

To demonstrate evolvability in liposomes, we used two green fluorescent protein (GFP) genes, GFPuv5 and GFPuv2, the former of which shows an eightfold stronger fluorescent signal than the latter. Two plasmids encoding the respective genes were mixed (GFPuv5:GFPuv2 = 0.85:0.15), and encapsulated into liposomes together with the in vitro transcription – translation system. The plasmid concentration was 1 ng/μL, corresponding to the concentration required to achieve a single copy of the plasmid on average per liposome 2 μm in diameter. Figure 15.5a shows the results of FACS analysis of these liposomes. Six different regions (R1–R6) exhibiting higher green fluorescence signal but with different internal aqueous
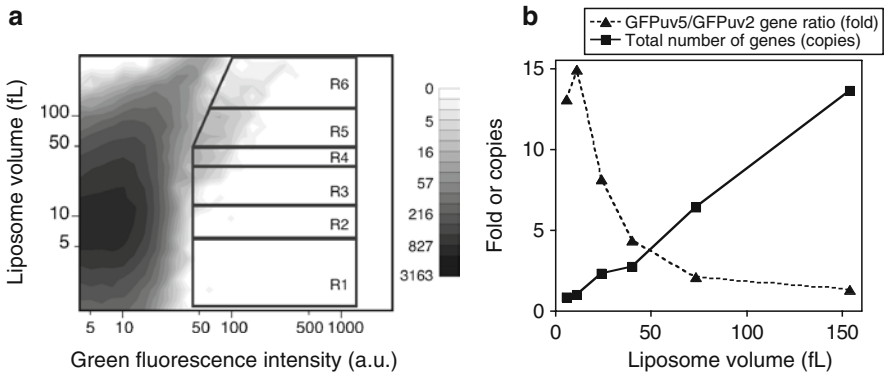
Fig. 15.5 Experimental demonstration of evolvability in liposomes. (**a**) Contour maps of green fluorescence intensity (*horizontal*) and internal aqueous volume (*vertical*) of liposomes after protein synthesis reaction with cell-free system using a mixture of GFPuv2 and GFPuv5 genes (GFPuv2:GFPuv5 gene = 0.85:0.15) as the template. Regions R1–R6 exhibiting higher green fluorescence signal but with different internal aqueous volume were sorted independently by FACS to obtain plasmids for further analysis. (**b**) Relationship between internal aqueous volume of liposomes and enrichment factor or total number of genes. Each point in the plot represents the values obtained from regions R1–R6, respectively

volume were sorted independently by FACS, and the plasmids contained were recovered for further evaluation.

Using quantitative PCR, the numbers of genes encoding GFPuv2 and GFPuv5 were determined from a mixture of the two. Using these numbers, the average number of plasmids in each liposome and the enrichment factor of the GFPuv5 gene over the GFPuv2 gene for each of regions R1–R6 were estimated (Fig. 15.5b). As expected, much higher enrichment factors were seen in the regions with liposomes possessing lower plasmid copy number. That is, higher evolvability was achieved in liposomes with genes present in lower copy numbers.

In the above examples, we showed that the size of the compartment is an important parameter for achieving efficient reaction and evolvability. These aspects must also be considered in synthesizing a minimal cell. The ability to evolve can be achieved when the genetic information molecule is compartmentalized to link the phenotype with the genotype. However, if the compartment size is relatively large, the concentrations of the metabolic products as well as genes are diluted, resulting in poor reaction efficiency. Thus, there should be an appropriate range of compartment size. A reasonable assumption is that a single copy of the gene is used to synthesize 10^3 protein molecules. This would produce a protein concentration of about 1 μM or 1 nM for a compartment on the order of 1 or 10 μm in diameter, respectively. Ordinary dissociation constants for protein – protein interactions and protein – DNA interactions range from 1 μM to 1 nM, but not to 1 pM. Therefore, the size of the minimal cell should be of the order of 1–10 μm in diameter for enabling the regulation of metabolic and genetic networks through interactions among biopolymers. Thus, compartment size is important for both evolvability and for biopolymer interactions.

15.6   Future Prospects

We have described reconstitution of the replication of genetic information (Kita et al. 2008) and experimental demonstration of the evolvability (Sunami et al. 2006) in liposomes. These are important steps toward the construction of minimal cells, although the compartment is still static. In this section, we will discuss the future challenges in synthesis of dynamic and autonomous minimal cells by the bottom-up approach.

Extant cells utilize exogenous nutrients to supply energy (self-maintenance), and thereby reproduce the entire contents of the cell (self-reproduction). These two properties, self-maintenance and self-reproduction, are essential for autonomous growth of the cells. Our system still lacks the machinery to incorporate nutrients from the external environment, which may be overcome by regulating the permeability of the membrane. Indeed, Noireaux and Libchaber (2004) showed that the internal protein synthesis reaction in liposomes can be prolonged by the presence of the α-hemolysin pore protein from *Staphylococcus aureus*. Integration of such molecules into our system may facilitate the development of "self-maintenance."

Another challenge is to encode the entire macromolecular components on the genome to realize "self-reproduction." The genome (RNA) in our system encodes only the RNA replicase. Nevertheless, we have shown that the addition of another gene (i.e., the β-galactosidase gene) is possible. Further increasing the number of encoded genes may lead to the realization of "self-reproduction" in our system.

Above we have described a possible strategy to synthesize a minimal cell by the bottom-up approach, which requires knowledge in terms of designability. On the other hand, extant cells have emerged without a designer; emerging through the evolutionary processes involving iteration of mutation and selection. Therefore, we can mimic what nature has done to synthesize a minimal cell. We are currently engaged in studies to develop an evolutionary system as shown in Fig. 15.6.

In the first step, the self-replication reaction is carried out in liposomes by encapsulating a small amount of genetic information together with the protein translation machinery. This results in amplification of the genetic information but the internal reaction will eventually terminate because of inactivation of the components or depletion of substrates. The liposomes are then divided into smaller ones by mechanical forces, e.g., by extrusion, so that the amount of genetic information in each vesicle becomes less than one on average. At this stage, the number of vesicles will increase, and only a portion will be selected at random for the next steps. The selected vesicles are then fused with larger vesicles (Pantazatos and
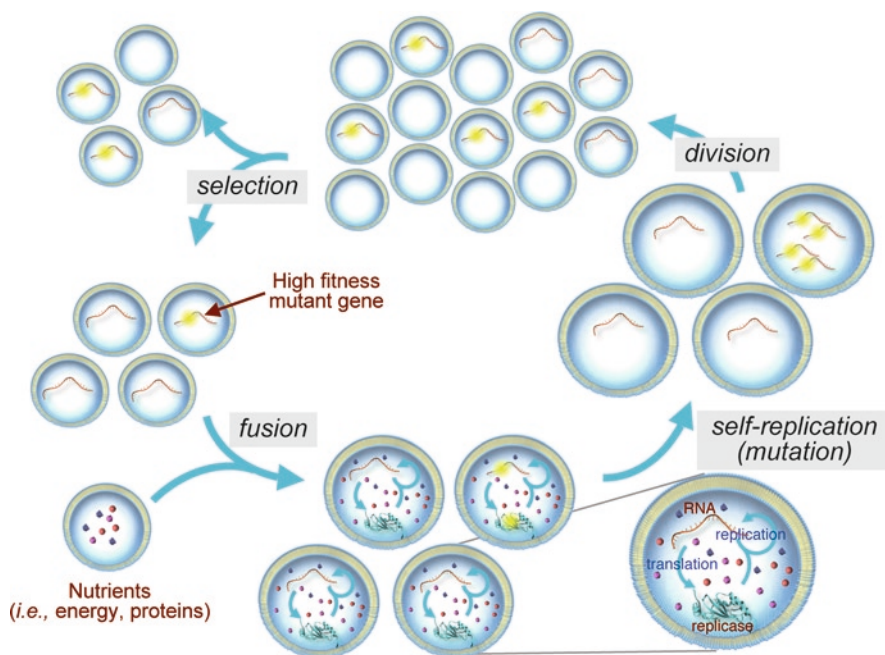
Fig. 15.6 Scheme of evolutionary programming toward synthesis of autonomously evolving minimal cell. See text for details

MacDonald 1999; Stamatatos et al. 1988), which contain the protein translation machinery and nutrients (NTPs and amino acids) required for the self-replication reaction. The self-replication reaction is conducted within the vesicles again. In this way, iterative rounds of division and fusion of vesicles will lead to continuous proliferation of the genetic information, and evolution will take place simultaneously. Qβ replicase is known to be error prone (Drake 1993), and therefore mutations will be introduced into the genetic information. If genes with higher fitness emerge, these are likely to dominate the population during the proliferation process. Now, we can begin to consider utilizing this autonomous optimization process for synthesizing a minimal cell.

We can first evolve a gene with higher self-replication efficiency. We then can incorporate a set of genes encoding the components required for lipid synthesis, and carry out the evolutionary process. During this process, we select the vesicles that have grown autonomously and have also divided into smaller ones. Repeating these steps will lead to the emergence of genes that encode not only high self-replication efficiency but also efficient lipid synthesis and vesicle division. Alternatively, genes encoding channel proteins may be incorporated into the genetic information, followed by the evolutionary cycles and then selecting for the emergence of self-maintenance. As soon as an evolvable self-replicating system is constructed, it will no longer be necessary to improve our design, and instead we will be able to allow evolution to design the minimal cell. Such *evolutionary programming* is one of the promising strategy toward the synthesis of a minimal cell.

References

Baba T, Ara T, Hasegawa M et al (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2(2006):0008

Bangham AD, Horne RW (1964) Negative staining of phospholipids and their structural modification by surface-active agents as observed in the electron microscope. J Mol Biol 8:660–668

Benner SA, Sismour AM (2005) Synthetic biology. Nat Rev Genet 6:533–543

Butland G, Peregrin-Alvarez JM, Li J et al (2005) Interaction network containing conserved and essential protein complexes in *Escherichia coli*. Nature 433:531–537

de Visser JA, Hermisson J, Wagner GP et al (2003) Perspective: evolution and detection of genetic robustness. Evolution 57:1959–1972

Deamer D (2005) A giant step towards artificial life? Trends Biotechnol 23:336–338

Drake JW (1993) Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci USA 90:4171–4175

Forster AC, Church GM (2006) Towards synthesis of a minimal cell. Mol Syst Biol 2:45

Giaever G, Chu AM, Ni L et al (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418:387–391

Hanczyc MM, Fujikawa SM, Szostak JW (2003) Experimental models of primitive cellular compartments: encapsulation, growth, and division. Science 302:618–622

Haruna I, Spiegelman S (1965) Autocatalytic synthesis of a viral RNA *in vitro*. Science 150:884–886

Hashimoto M, Ichimura T, Mizoguchi H et al (2005) Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. Mol Microbiol 55:137–149

Hosoda K, Matsuura T, Kita H et al (2007) Kinetic analysis of the entire RNA amplification process by Qbeta replicase. J Biol Chem 282:15516–15527

Hosoda K, Sunami T, Kazuta Y et al. (2008) Quantitative study of the structure of multilamellar giant liposomes as a container of protein synthesis reaction. Langmuir

Hutchison CA, Peterson SN, Gill SR et al (1999) Global transposon mutagenesis and a minimal Mycoplasma genome. Science 286:2165–2169

Ichihashi N, Matsuura T, Kita H et al (2008) Importance of translation-replication balance for efficient replication by the self-encoded replicase. Chembiochem 9:3023–3028

Ishikawa K, Sato K, Shima Y et al (2004) Expression of a cascading genetic network within liposomes. FEBS Lett 576:387–390

Kita H, Matsuura T, Sunami T et al (2008) Replication of genetic information with self-encoded replicase in liposomes. Chembiochem 9:2403–2410

Kitano H (2004) Biological robustness. Nat Rev Genet 5:826–837

Luisi PL (2002) Toward the engineering of minimal living cells. Anat Rec 268:208–214

Luisi PL, Ferri F, Stano P (2006) Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften 93:1–13

Matsuura T, Yamaguchi M, Ko-Mitamura EP et al (2002) Importance of compartment formation for a self-encoding system. Proc Natl Acad Sci U S A 99:7514–7517

Matsuura T, Yomo T (2006) *In vitro* evolution of proteins. J Biosci Bioeng 101:449–456

Noireaux V, Libchaber A (2004) A vesicle bioreactor as a step toward an artificial cell assembly. Proc Natl Acad Sci U S A 101:17669–17674

Nomura SM, Tsumoto K, Hamada T et al (2003) Gene expression within cell-sized lipid vesicles. Chembiochem 4:1172–1175

Oberholzer T, Albrizio M, Luisi PL (1995) Polymerase chain reaction in liposomes. Chem Biol 2:677–682

Pantazatos DP, MacDonald RC (1999) Directly observed membrane fusion between oppositely charged phospholipid bilayers. J Membr Biol 170:27–38

Paul N, Joyce GF (2004) Minimal self-replicating systems. Curr Opin Chem Biol 8:634–639

Posfai G, Plunkett G 3rd, Feher T et al (2006) Emergent properties of reduced-genome *Escherichia coli*. Science 312:1044–1046

Riley M, Abe T, Arnaud MB et al (2006) Escherichia coli K-12: a cooperatively developed annotation snapshot–2005. Nucleic Acids Res 34:1–9

Rual JF, Venkatesan K, Hao T et al (2005) Towards a proteome-scale map of the human protein-protein interaction network. Nature 437:1173–1178

Sato K, Obinata K, Sugawara T et al (2006) Quantification of structural properties of cell-sized individual liposomes by flow cytometry. J Biosci Bioeng 102:171–178

Shimizu Y, Inoue A, Tomari Y et al (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755

Simpson ML (2006) Cell-free synthetic biology: a bottom-up approach to discovery by design. Mol Syst Biol 2:69

Stamatatos L, Leventis R, Zuckermann MJ et al (1988) Interactions of cationic lipid vesicles with negatively charged phospholipid vesicles and biological membranes. Biochemistry 27:3917–3925

Sunami T, Sato K, Matsuura T et al (2006) Femtoliter compartment in liposomes for *in vitro* selection of proteins. Anal Biochem 357:128–136

Szostak JW, Bartel DP, Luisi PL (2001) Synthesizing life. Nature 409:387–390

Takakura K, Sugawara T (2004) Membrane dynamics of a myelin-like giant multilamellar vesicle applicable to a self-reproducing system. Langmuir 20:3832–3834

Uetz P, Giot L, Cagney G et al (2000) A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature 403:623–627

Yu BJ, Sung BH, Koob MD et al (2002) Minimization of the *Escherichia coli* genome using a Tn5-targeted Cre/loxP excision system. Nat Biotechnol 20:1018–1023

Yu W, Sato K, Wakabayashi M et al (2001) Synthesis of functional protein in liposome. J Biosci Bioeng 92:590–593

Index