Thermo-elasticity and adhesion as regulators of cell membrane architecture and function

2006 J. Phys.: Condens. Matter 18 R785

(http://iopscience.iop.org/0953-8984/18/45/R02)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **18** (2006) R785–R825 doi:10.1088/0953-8984/18/45/R02

TOPICAL REVIEW

Thermo-elasticity and adhesion as regulators of cell membrane architecture and function

Erich Sackmann

Physik Department E22, Technische Universitaet Muenchen, D-85748 Garching, Germany

E-mail: sackmann@ph.tum.de

Received 12 April 2006, in final form 23 August 2006
Published 26 October 2006
Online at stacks.iop.org/JPhysCM/18/R785

Abstract

Elastic forces and structural phase transitions control the architecture and function of bio-membranes from the molecular to the microscopic scale of organization. The multi-component lipid bilayer matrix behaves as a pseudo-ternary system. Together with elastically and electrostatically mediated specific lipid–protein interaction mechanisms, fluid–fluid phase separation can occur at physiological temperatures. This can drive the transient generation of micro-domains of distinct composition within multi-component lipid–protein alloys, enabling cells to optimize the efficiency of biochemical reactions by facilitating or inhibiting the access of enzymes by distinct substrates or regulatory proteins.

Together with global shape changes governed by the principle of minimum bending energy and induced curvature by macromolecular adsorption, phase separation processes can also play a key role for the sorting of lipids and proteins between intracellular compartments during the vesicle mediated intracellular material transport.

Cell adhesion is another example of mechanical force controlled membrane processes. By interplay of attractive lock and key forces, long range disjoining pressures mediated by repeller molecules or membrane undulations and elastic interfacial forces, adhesion induced domain formation can play a dual role for the immunological stimulation of lymphocytes and for the rapid control of the adhesion strength.

The present picture of the thermo-elastic control of membrane processes based on concepts of local thermal equilibrium is still rudimentary and has to be extended in the future to account for the intrinsic non-equilibrium situation associated with the constant restructuring of the cellular compartments on a timescale of minutes.

(Some figures in this article are in colour only in the electronic version)

Contents

1. Introduction

Cells are highly crowded colloidal assemblies of functional compartments enclosed by multi-component lipid–protein bilayers of distinct composition. Global order within the cytoplasmic space is established dynamically by continuous bi-directional transport of material from the endoplasmic reticulum (ER), the site of protein and lipid synthesis, to the plasma membrane, the site of material import and export, through vesicles [1]. Despite this rapid material turnover the characteristic composition of the lumen and the cell envelope of each compartment is astonishingly well preserved. Local order on the level of the membranes is transiently generated by the formation of micro-domains (often called rafts) or flask-like invaginations (caveolae and coated pits) with distinct composition. This lateral ordering can (i) optimize biochemical reactions by controlling the access of specific components to enzymes, (ii) accumulate proteins involved in basic membrane processes such as hormone signal transduction (cf [2a]) and (iii) help to maintain the specific composition of the intracellular compartments during the material turnover.

This article focuses on the control of cellular processes by the unique elastic properties and structural phase transitions of lipid bilayers. In the first part, the self-assembly of domains of distinct composition in multi-component lipid–protein bilayers by interplay of lateral phase separation, specific lipid–protein interaction mechanisms and coupling between membrane curvature and phase separation is discussed. Evidence is provided that support the hypothesis that the lipid matrix of membranes can be considered as a pseudo-ternary lipid alloy. One component (a mixture of lipids of the sphingomyelin family) exhibits a fluid to solid phase transition above physiological temperature. The second component (a mixture of glycerol

phospholipids) is fluid at physiological conditions. The third component (cholesterol) acts as a fluidizer of the solid phase. The formation of functional domains or the sorting of components between intracellular compartments can be mediated by phase separation between a cholesterol poor and a cholesterol rich (and more condensed) fluid phase in combination with selective elastic lipid–protein interaction mechanisms based on the thickness matching of hydrophobic domains of the lipid bilayer and the integral proteins. The second part deals with the control of the morphologies and shape transitions of vesicles, intracellular compartments and erythrocytes by the concept of minimum bending elasticity. In this context the generation of specific shapes and the fine tuning of the lateral distributions of membrane constituents by coupling of the lipid–protein bilayers to shear elastic intracellular macromolecular scaffolds is discussed. In the third part we discuss the control of vesicle mediated intracellular trafficking by induced bending moments (mediated by macromolecular adsorption) and curvature induced lateral phase separation of lipids and proteins. The important role of elastic mechanisms of selective lipid–protein interaction for sorting of lipids and proteins between compartments is stressed. A model of tension-induced membrane fusion is presented.

The last part deals with the control of cell adhesion by the interplay of specific attractive forces between conjugate receptor–ligand pairs, repulsive interfacial interactions mediated by repeller molecules or membrane undulations and adhesion induced elastic forces. A major consequence of this interplay of forces is the formation of micro-domains by receptor phase separation below a wetting transition point. This enables cells (i) to adhere strongly through the commitment of few receptors and (ii) to control the adhesion strength within seconds by internalization (endocytosis) or export (exocytosis) of receptors and repellers. The possible role of adhesion domains as confined reaction spaces (called immunological synapses) for the stimulation of immunological responses is pointed out.

The article focuses on our understanding of biological membranes based on experimental and theoretical model membrane studies which have contributed very substantially to our present knowledge of bio-membranes. Another purpose is to assess the state of the field 35 years after the birth of the bending elasticity concept of membranes. Main emphasis is laid on the presentation of the concepts and the discussion of a few distinct membrane processes, providing evidence for the control of biological functions by the elastic and thermodynamic properties of lipid–protein alloys. Therefore the list of references is restricted to a few basic publications, review articles or books. It is also assumed that the reader is familiar with the basic principles of the design of cells.

2. Molecular architecture of cell membranes

2.1. The erythrocyte cell envelope

The physicist's favourite prototype of a composite cell membrane is the envelope of erythrocytes (figure 1). This stratified shell is composed of two sub-shells: the 3 nm thick multi-component lipid–protein bilayer (called the plasma membrane (PM) in the following) and the (roughly 300 nm thick) two-dimensional macromolecular network (the membrane cytoskeleton) which is locally coupled to the cytoplasmic face of the PM. The extra-cellular side of the PM is covered by a densely packed (about 40 nm thick) macromolecular film formed by head groups of cell surface glycoproteins (such as cell adhesion molecules, cf figure 14 below). This so-called glycocalix serves the communication with the environment and acts as a filter controlling the access of macromolecules to the lipid–protein bilayer, but it also plays a key role for the control of cell adhesion, as will be shown below. The membrane-associated cytoskeleton is a rubber-like network composed of semi-flexible spectrin filaments that form

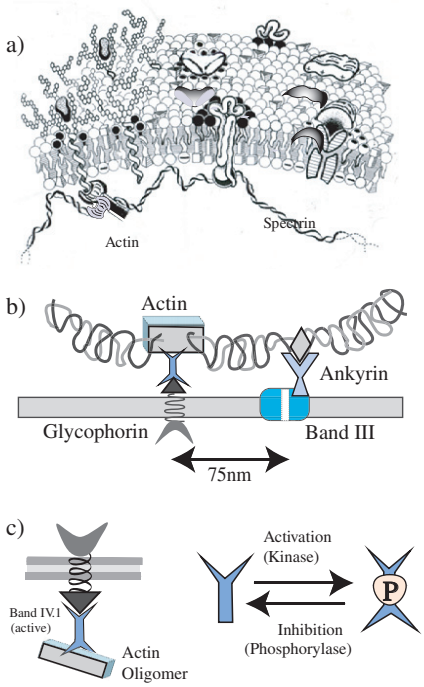

Figure 1. (a) Architecture of composite plasma membrane of erythrocyte made up of the (∼4 nm thick) lipid–protein bilayer, called the plasma membrane (PM), the (∼40 nm thick) glycocalix facing the extracellular space and the (∼300 nm thick) cell stabilizing spectrin–actin network. Note that the lipids are asymmetrically distributed between the two monolayers and that the proteins tend to be surrounded by clouds of specific lipids. Note second that the average thickness of the spectrin filaments (∼3 nm) is actually similar to that of the membrane. (b) Anchoring of spectrin–actin scaffold to PM by two linkers. Ankyrin couples the head-to-tail linkage of spectrin dimers to the intracellular domain of the anion channel band III, while the protein band IV.1 binds the actin oligomers to the intracellular domain of glycophorin, a glycoprotein acting also as a blood group antigen. (c) The spectrin–band IV.1 linkage is controlled by the state of phosphorylation of the linker, enabling cells to actively control the membrane elasticity through the density of active linkers (cf [2a, 3]).

a rather well defined triangular lattice. The sides (length $a \approx 75$ nm) consist of tetramers of spectrin that are formed by head-to-tail coupling of hetero-dimers of ≈100 nm contour length and ∼10 nm persistence length. These are interconnected at the vestiges of the network by 35 nm long oligomers of actin, the length of which is assumed to be determined by association with tropomyosin/troponin complexes [1]. The network is coupled to the plasma membrane by two types of linkers. One is band IV.1, which couples the actin oligomers to the intracellular domain of glycophorin, a membrane spanning receptor acting also as blood group antigen. The second linker is ankyrin. It couples the head-to-tail connection of the spectrin dimers to the intracellular domain of the band III, a membrane spanning protein that acts simultaneously as ion exchange protein.

The outstanding elastic properties of the cell envelope are controlled in two ways [2–5]: first by the density of defects of the network, and second by the lateral density of membrane–cytoskeleton linkages. Major defects of the triangular network are $+60°$ and $-60°$ disclinations, which may be paired (forming edge dislocations) or dissociated (cf figure 11 below). As indicated in figure 1(c), the linker IV.1 can be switched between an inactive

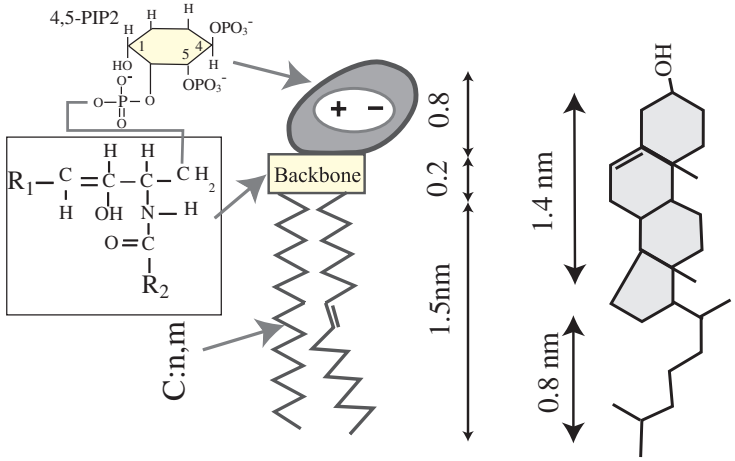

Figure 2. Basic lipid structures. Middle: Structure of phospholipids or glycolipids determined (i) by the size and electric charging of the head group, (ii) the structure of the backbone that couples the head group to the chains, (iii) the number, n, of C atoms of the hydrocarbon chains and (iv) the number, m, of double bonds. We denote specific lipids as X: n, m. Thus phosphatidylcholine with each chain composed of n saturated and m unsaturated bonds is abbreviated as PC: n, m. Left: structure of backbone of sphingomyelin (SPM), where R_i denote the chains. The head group corresponds to phosphatidyl inositol 4,5-biphosphate (PIP2), which plays a key functional role. Right: structure of cholesterol. Note the relative lengths of the two classes of components.

(closed) and an active (open) conformation by phosphorylation of a tyrosine group. This control mechanism of the density of active linkers is universal for a whole class of linkers (of the so-called band IV.1 family), which includes talin, the predominant actin-bilayer linker of nucleated cells (cf [1]). The coupling of a phosphate group to the linker is mediated by biochemical switches of the rho-family. These GTPases activate specific kinases (the enzymes transferring the phosphate groups to specific tyrosine side groups of the linker protein). Therefore, the mechanism shown in figure 1(c) provides a universal link between cell signalling and the structure and elastic behaviour of the cell envelope.

There is some evidence that the design principle of the composite erythrocyte shell is universal for most mammalian cells, in particular blood cells such as blood platelets. In this case the sub-membrane scaffold forms an interface between the plasma membrane and the actin cortex. It is often composed of spectrin analogues [1] such as fodrin (especially abundant in neural cells) or dystrophin (in muscle cells). Most interestingly, it turned out that intracellular compartments are also often covered by spectrin coats, which are assumed to play a key role for their intracellular transport [6].

2.2. Lipid bilayers as two-dimensional liquid crystalline matrix for proteins

Biomembranes of eukaryotic cells are made up of two major families of lipid molecules and the amphiphilic steroid derivative cholesterol (figure 2). The two lipid constituents, glycerol phospholipids (prototype phosphatidylcholine (PC)) and sphingolipids (prototype sphingomyelin = SPM) are each composed of two hydrophobic hydrocarbon chains, a semi-polar backbone and a hydrophilic head group, but are distinguished by the structure of the backbone. The lipids within each class are distinguished by three structural features: the chain length (which may vary between 16 and 24 carbon atoms), the number of double bonds (varying

between one and six) and the polarity of the head groups, which can be negatively charged (such as phosphatidyl serines and phosphatidyl inositols (cf figure 2)) or zwitterionic (such as PC and phosphatidyl ethanolamine (PE)). A specific component only present in mitochondria is the twofold negatively charged cardiolipin that exposes four hydrocarbon chains. There are no positively charged lipids.

The self-assembly of phospholipids in aqueous solutions into closed shells is a most important use of the hydrophobic effect by nature (cf [2b]). Due to the two hydrocarbon chains the solubility of natural lipids is well below 10^{-12} M (corresponding to a single lipid molecule in the volume of the erythrocyte).

Natural membranes are multi-component alloys composed of several tens of phospholipid components and cholesterol. However, a first step towards order formation in these multicomponent systems is mediated by three factors.

(1) The components are not evenly distributed between the intracellular compartments. Sphingomyelins and cholesterol are rare in the endoplasmic reticulum (5% and 6%, respectively) but most abundant in plasma membranes (20 and 50 mol%, respectively) and in lysosomes (30 and 25%), while the other phospholipids are nearly evenly distributed among the compartments. However, within one class of compartment the lipid composition is strictly controlled (a process called lipid homeostasis). This distinct distribution of lipids (and that of the membrane proteins) is maintained astonishingly well during the rapid intracellular trafficking through vesicles and is fine tuned by lipid exchange proteins [1]. Membranes are thus in stationary and not thermodynamic equilibrium and the lipid flux plays a key, but still unknown, role for the generation of order in cells.
(2) The lipids are asymmetrically distributed between the two leaflets. Thus, nearly all charged lipids (comprising about 10% of the total phospholipid content) reside in the inner monolayer of the PM, while SPM is located mainly in the outer leaflet. This asymmetry is supposed to be actively driven (under ATP consumption) by membrane proteins called 'flipases' (cf [1] for references). A second possible driving force for the charged lipid asymmetry is the electrical potential which is positive in the cytoplasm.
(3) Sphingomyelin is distinguished from other phospholipids by its rather high content of long (C18- and C20-) hydrocarbon chains that are saturated or exhibit only one double bond. Therefore, the chain melting transition of this lipid component lies above physiological temperatures (cf figure 4 below). Below we will show that this distinct hydrocarbon pattern of SPM together with selective lipid–protein interaction mechanisms can provide one mechanism for the generation of order in plasma membranes by lateral phase separation.

2.2.1. Integral and peripheral membrane proteins. Membrane proteins may be divided into two classes: integral proteins and peripheral proteins. Integral proteins penetrate the bilayer with one or several α-helices. Their functional domains may be located within the hydrophilic head group or within the hydrophobic core. Examples of the former class are cell adhesion molecules (cf figure 14), and of the second ion pumps, hormone receptors, photosynthetic reaction centres or enzymes generating extracellular matrix proteins such as hyaluronic acid synthase. Peripheral proteins are associated with membranes in various ways. The most common mechanism is anchoring by fatty acids or phospholipids (such as GPI anchors [1]). Examples include many receptors, and proteins involved in membrane based signal transduction processes, such as phospholipases or GTPases. A second major coupling mechanism is through electrostatic interaction. Basic protein segments also exposing fatty acids can directly couple to clusters of negative lipids (cf figure 3(a) left side), while Ca ions can form salt bridges between acidic protein segments and lipids (cf the case of annexin).

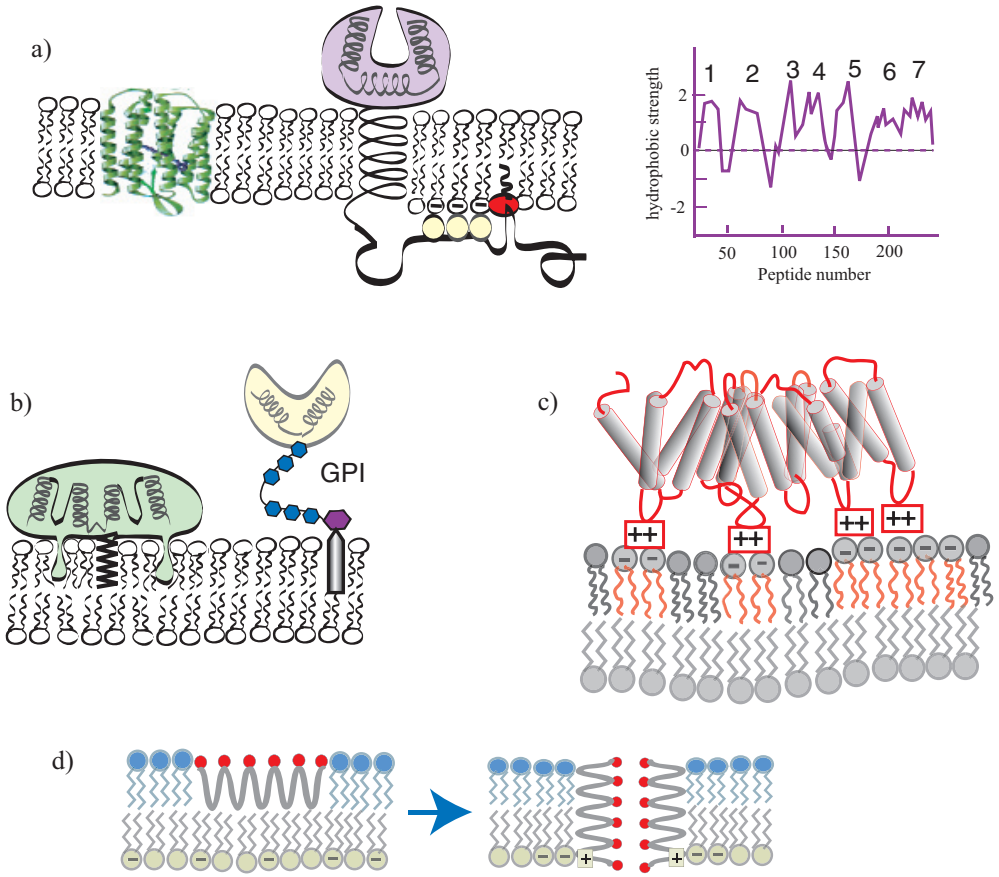

Figure 3. Schematic view of zoo of membrane proteins. (a) Integral proteins interacting predominantly through the hydrophobic effect. The hydrophobicity plot on the right side shows the distribution of hydrophobic and polar amino acids of bacteriorhodopsin, a protein that spans the membrane with seven α-helices. The hydrophobic sequences (1–7) form α-helices in a hydrophobic environment. The segment length of the hydrophobic stretches is adjusted to the thickness of the hydrocarbon core of the bilayers. Since the pitch of α-helices is $P = 0.54$ nm and comprises 3.6 amino acids, about 25 nm amino acids are required to span a bilayer of 4.0 nm. Note that proteins that are anchored with only one or two α-helices (such as transferrin receptors) often expose huge hydrophilic domains, which can interact with the bilayer both via electrostatic forces and fatty acid chains. (b) Left: interaction of peripheral protein with membrane by a fatty acid chain, which can be enforced by penetration of flexible peptide side-chain into the semipolar region of membrane (for example, the actin–membrane anchoring protein hisactophilin [8] or caveolin (cf figure 13). Right: anchoring of cell surface receptor through phosphoinositol anchor (cf figure 2) exposing oligosugar spacer (=GPI anchor). (c) Left: electrostatic binding of peripheral protein to charged lipid clusters by Ca^{2+} bridges. Prototypes of this coupling scheme are the annexins. (d) Model of two modes of interaction of proteins exhibiting amphiphatic helices (forming a rod with hydrophobic and hydrophilic faces). They can embed horizontally in one of the monolayers or switch to an integrated conformation, for instance by charging histidine groups. This mode of coupling has been established by NMR in model systems [7], but not in natural membranes yet.

A peculiar mode of binding is through amphiphatic α-helices, which exhibit a hydrophobic and hydrophilic face in the axial direction. They can switch between a horizontal and a normal position by charging the hydrophilic ends exhibiting histidine sequences [7]. The transition between membrane and cytoplasmic proteins is fuzzy. Thus, proteins involved

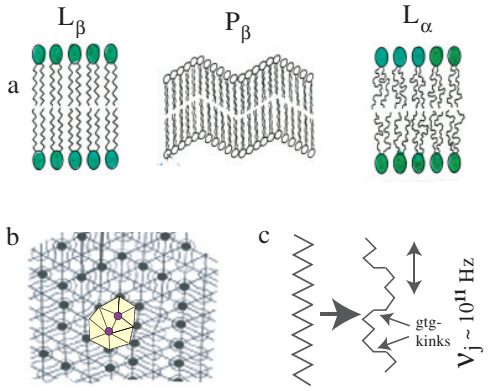

Figure 4. (a) Illustration of phase transitions of lipid bilayers from solid state (often called the gel phase) to the fluid state. The gel phase can be flat (L_β) or corrugated (P_β), depending on the ratio of the cross-sections of the chains and the head groups, respectively (cf [2b]). (b) Structure of gel phase showing that chains form a triangular lattice while the whole molecules (indicated by dumb-bells) form an orthorhombic lattice [2a]. In the centre an edge dislocation of the lattice consisting of a heptagon ($+60°$ disclination) and a pentagon ($-60°$ disclination) is shown. (c) The chain melting transition is associated with the spontaneous formation of gtg defects within the chains. Their rapid diffusion along the chains determines the fluidity of the bilayer.

in formation of coats such as clathrin and caveolin may also be considered as membrane associated proteins.

2.3. Self-assembly of functional domains in multi-component lipid–protein-alloys

It is generally accepted that the differentiation of cells is controlled by the expression of a specific pattern of proteins (the proteome). It becomes more and more evident that equally important is the distinct lipid composition of cell envelopes and intracellular compartments (the 'lipidome'). Below we argue that due to elastic mechanism of specific lipid–protein interaction the lipidome is most likely closely related to the proteome. This implies that lipids play a more active role for membrane functions than generally assumed. Below we discuss possibilities of the self-assembly of functional domains in two dimensional multi-component lipid–protein mixtures by interplay of thermodynamics and elastic forces. One major physical driving force for order formation in multicomponent membranes is the ability of lipid bilayers to undergo a manifold of thermotropic phase transitions. Most important is the chain melting transition of bilayers (corresponding to the smectic B to smectic A transition of liquid crystals). It is driven by thermal excitation of hindered rotations about C–C-bonds within the hydrocarbon chains resulting in an abrupt increase of conformational chain defects (g-t-g kinks) at $T > T_m$ (cf figure 4(a)). The membrane fluidity is determined by the rapid diffusion of the defects along the chain (according to quasi-elastic neutron scattering the defects move randomly with jump frequencies of $\nu = 10^{11}$ Hz [9]). The defect motion determines the passive trans-membrane transport of small molecules (cf [2b]). Cells can adjust the transition temperature to the environmental conditions by varying the number of double bonds and their position along the chains or by charging the head groups [2b]. The charge effect is due to the repulsive lateral electrostatic pressure π_{el} which (according to the Clausius–Clapeyron equation) lowers the transition temperature by $\Delta T_m = \Delta A \pi_{el} / \Delta S$. Plant physiologists make use of these mechanisms to grow cold resistive plants.

Until the discovery of triton insoluble membrane fractions (called rafts, cf [10]) the paradigm of biomembrane architecture for cell biologists was the fluid mosaic model first

proposed by Singer and Nicholson. The basic idea was that the lipid bilayer provides a homogeneous fluid matrix for the functional proteins but does not actively control the membrane structure and function. This is strange since judged from the lipid composition of biomembranes and our knowledge of thermodynamics of lipid mixtures it was evident that biomembranes should exhibit lateral phase segregation for several reasons. About 40% of the hydrocarbon chains are saturated C16 and C18 chains, exhibiting phase transitions above the physiological temperature. The plasma membrane contains about 20% sphingomyelin (SPM), which resides mostly in the outer leaflet (corresponding to a content of ∼40 mol%). SPM is much richer in long saturated hydrocarbon chains (C18:0–C22:0) than the other phospholipids (cf [2a] for details). Therefore, the natural mixture of brain SPM exhibits a broad fluid–solid transition regime with a maximum at 40 °C and the liquidus line extends up to 45 °C (figure 5(a)). Third, cholesterol is a fluidizer of lipid bilayers and has a preference for saturated lipids [11]. Taken together, the lipid moiety of biological membranes can thus be considered as a pseudo-ternary alloy that is composed of the high melting SPM fraction, a mixture of phospholipids with glycerol backbones forming the low melting component and cholesterol. Finally, lateral phase separation is expected to be mainly driven by selective lipid protein interaction mechanisms as will be discussed below.

Despite numerous experimental and theoretical studies the structural and thermodynamic properties of cholesterol–phospholipid mixtures are still poorly understood. One popular conception is that above about 25% cholesterol a homogeneous solution, the so-called fluid-condensed phase L_0, is formed (both for PC and SPM) up to a solubility limit of ∼50 mol% and at all temperatures. This simple model conflicts with several experimental findings. Considering numerous studies by calorimetry, molecular dynamic simulation and neutron scattering studies, the following more subtle phase diagram (shown in figure 5(a)) is suggested, which holds for the special case of cholesterol PC-C16:1 mixture.

(1) Up to 5 mol% (molar fraction $x_{ch} = 0.05$) the normal PC phase transitions and structural features, such as the ripple ($P_{\beta'}$) structure, are observed (cf [2b]).
(2) Between $x_{ch} \approx 0.08$ and 0.25 and at $T < T_m$, phase separation into a solid solution of cholesterol in PC and a stoichiometric mixture with $x_{st} = 0.25$ occurs. The formation of this specific phase (established by small angle neutron scattering (SANS) studies, cf [2b]) is often denoted as complex formation. At $T > T_m$ the fluid phase appears homogeneous (e.g. according to SANS and NMR studies), although calorimetric and spectroscopic studies [2b, 14] clearly indicate a transition line (L) that shifts to higher temperatures at increasing cholesterol content up to $x_{ch} = 42\%$.
(3) Between 25 and 42 mol% (the limit of solubility as suggested by SANS) a fluid-like phase is formed above and below T_m. Its lateral packing density is increased by about 20% compared to the L_α phase at $x_{ch} < 25\%$. Concomitantly the lateral diffusivity is reduced by a factor of two and the bending stiffness by a factor of four. However, this condensed phase exhibits a very broad thermotropic transition which defines the phase lines L and S (indicated by the dashed horizontal line) [17]. The associated heat of transition decreases linearly with the cholesterol content from $\Delta H \sim 20$ kJ M^{-1} at $x_{ch} \approx 0.1$ to zero at $x_{ch} \approx 0.45$. ΔH is larger by about a factor of two than the value for the chain melting transition [14] which suggests that the broad transition is associated with a drastic structural change of the membrane. One possibility is that it is associated with the dehydration of the semipolar interface.
(4) For the mixture of cholesterol with PC-C:14, 0, phase separation into the stoichiometric mixture and the saturated solution ($x_s = 0.24$) occurs at 10 °C below the chain melting transition as demonstrated by SANS studies (cf [2], chapter V) This miscibility gap is

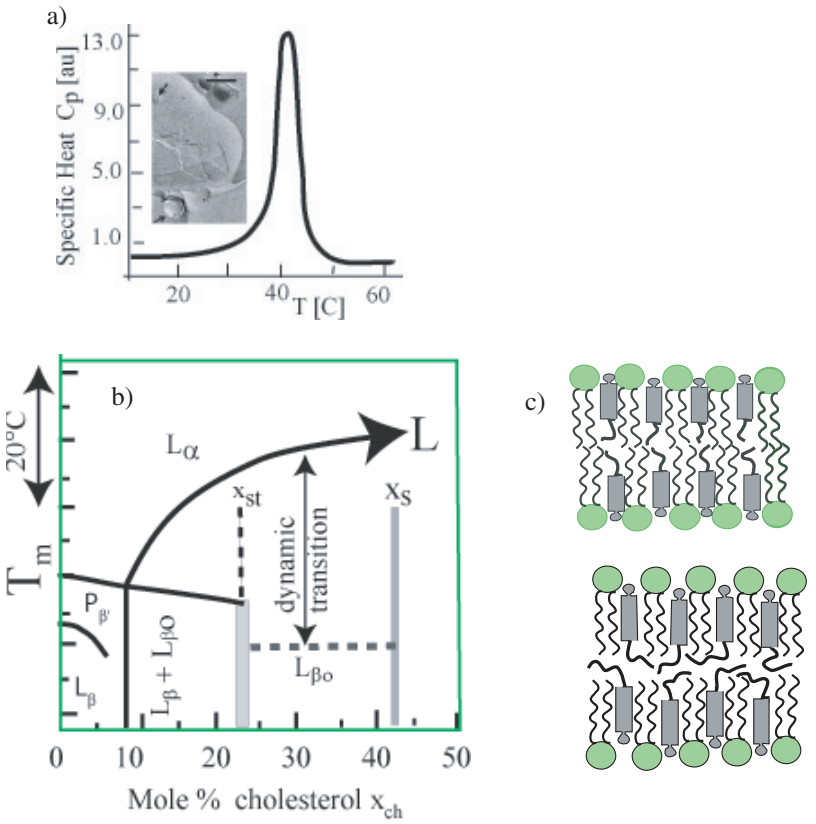

Figure 5. (a) Demonstration of phase transition of vesicles of brain sphingomyelin (SPM), reproduced from [12]. Composition: 2% C16:0; 58% C18:0; 6% C20:0; 9% C22:0; 7% C24:0 and 15% C24:1. The inset shows a freeze fracture electron micrograph. Closer inspection shows a wavy surface profile, indicating that the lipid is in the gel state. (b) Phosphatidylcholine–cholesterol phase diagram as suggested by NMR [13], calorimetry and spectroscopic methods (for references cf [2b, 14]) and dynamic and static neutron scattering [15, 17] studies. The absolute value of the chain-melting transition depends on the type of lipid. The compositions x_{st} and x_s correspond to a stoichiometric mixture and the saturation limit, respectively. Below the dashed line SANS shows phase separation into the stoichiometric and saturated mixtures. (c) Model of dynamic transition of condensed phase ($x > x_{st}$) as suggested by QENS. In the low temperature state, $L_{\beta o}$, cholesterol is shifted towards the centre of the bilayer with the hydrophobic chain possibly intercalated between the monolayers. Above line L the steroid is shifted towards the polar interface and the bilayer becomes thinner by 0.4 nm.

consistent with the phase rule if the mixture x_{st} is stoichiometric [2b]. At $x_{ch} > 0.42$ SANS experiments suggest decomposition of the mixture in a saturated phase L_s and cholesterol aggregates.

Combined studies by incoherent quasi-elastic neutron scattering (QELS), NMR and MD simulations [16] provide strong evidence that the broad calorimetric band is due to the following dynamic transition. Above the transition L ($T \geqslant T_L$) cholesterol exhibits very pronounced motions in the normal direction with root mean square amplitudes of ∼0.5 nm, and the position of the OH head groups at the polar interface is smeared out over 0.7 nm [16]. At decreasing temperature (below the line L) cholesterol is shifted towards the centre in such a way that its flexible chains penetrate into the opposite bilayer or even intercalate between the leaflets.

The membrane thickness increases by ∼0.4 nm and the out-of-plane motion of the steroid is strongly restricted. The transition may be associated with the shift of the hydrogen bonds between the OH group of the cholesterol and the carboxyl bond of the fatty acid by 0.6 nm towards the centre. This dynamic transition would be absent in monolayers, which may be the reason why it is very difficult to compare phase diagrams of mono- and bilayers [19]. A recent detailed mean field (regular solution) theory [18] suggests that mixtures of cholesterol and saturated lipids may exhibit a fluid–fluid coexistence (with one condensed phase) above the chain melting transitions.

Due to the high content of SPM and cholesterol ($x_{ch} > 42\%$), the lipid bilayer of the plasma membrane is expected to be in the condensed state $L_{o\beta}$, well below the phase line L. In contrast, the membranes of the ER, the Golgi and the nucleus contain less than 10 mol% of cholesterol and SPM and are therefore expected to be in the less condensed $L_{0\alpha}$ state. Together with selective lipid–protein interaction mechanisms discussed below, this difference in the phase state of the lipid bilayer moiety of different compartments can provide the major driving force for the sorting of proteins and lipids during intracellular trafficking or material export and import (cf appendix B). In order to understand these processes and the physical basis of micro-domain formation in natural membranes, more detailed studies of the lipid–protein interaction are necessary. Many studies of raft formation in cell membranes have been performed at room temperature, which is 15 °C below the transition temperature T_m of SPM, where SANS experiments predict two phase coexistence. Raft formation observed at room temperatures is therefore not necessarily biologically relevant.

2.3.1. Thermodynamics of lipid–protein mixtures. Similar to macromolecules in solutions, large integral proteins have a strong entropic tendency for segregation. This is predicted by the regular solution theory assuming, that the membrane can be considered as a lattice of lattice constant $\sqrt{A_L}$ (A_L area per lipid). By assuming that the protein occupies g sites in a lattice, the regular solution theory predicts for the free energy of mixing (cf [2b] for references)

$$\Delta G_{LP} = N k_B T \{\phi_L \ln \phi_L + (\phi_P/g) \ln \phi_P + \chi \phi_L \phi_P\} \tag{1}$$

ϕ_L and ϕ_P are the area fractions of the lipid and the protein, respectively, defined as

$$\phi_P = \frac{g n_P}{n_L + g n_P}; \qquad \phi_L = 1 - \phi_P. \tag{2}$$

The first two terms account for the entropy of mixing and the third for the heat of mixing. The entropy of mixing of lipid–protein alloys (and thus the tendency to mix homogeneously) is reduced compared to mixtures of lipids of equal size. This is due to the reduction in the number of configurations that can be generated by the random distribution of the components. By placing a macromolecule at a random site in the lattice, the number of sites occupied is $g = A_P/A_L$ and therefore the number of possibilities to deposit the other molecules is reduced. χ is the Bragg–Williams–Flory interaction parameter that accounts for the specificity of the lipid–protein interaction. It depends in very complicated way on the number of contacts between the protein surface and the lipids and should be considered as an empirical parameter that can be determined for instance from shifts of the chain melting transition temperature at low protein concentrations [2b]. Due to the reduction of the entropy of mixing, the critical concentration, ϕ_{Pc}, at which proteins start to segregate is shifted to low concentrations according to

$$\phi_{Pc} = \frac{1}{1 + \sqrt{g}}. \tag{3}$$

Consider a typical ion channel, such as (the 106 kDa) band III protein of erythrocytes. It spans the membrane with 13 helices and occupies an area of 4×8 nm^2. Therefore, $g \sim 50$ and

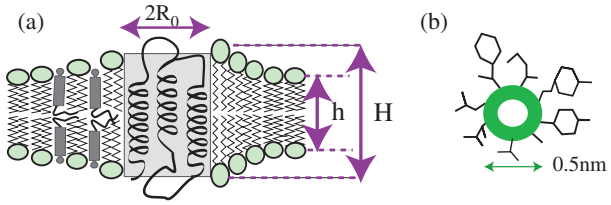

Figure 6. (a) Model of distortion of hydrocarbon chain near the surface of integral protein exhibiting larger thickness of hydrophobic domain than the bilayer. Note that both the conformation of the chains and the head groups are changed. On the left side the possible relaxation of the hydrocarbon chain stretching by cholesterol penetrating with chains between leaflets is indicated. (b) Cross section of α-helix exposing hydrophobic hairs, showing that it occupies roughly the area of a lipid molecule.

$\phi_{Pc} \sim 0.02$ (area per lipid: $A_L \approx 0.65$ nm^2). Protein segregation is thus expected to occur at a molar fraction of the order $x_P \sim 10^{-4}$.

This solubility limit would be even smaller if the extrinsic domains of the protein adsorbed to the membrane surface. An example is the transferrin receptor, which exhibits only one hydrophobic domain but affects the phase state of roughly 100 lipid molecules [2b].

2.3.2. Selective lipid–protein interaction. The above consideration shows that proteins have a strong tendency for segregation and are expected to induce lipid phase separation in fluid phases if they exhibit only a slight preference for one lipid component. Similarly, small concentrations of proteins interacting preferentially with a specific lipid component can change the phase diagrams of the lipid mixtures drastically. This effect can be used to estimate the number of lipids interacting with one protein by calorimetric studies [2b]. The question of microdomain formation in biomembranes is thus closely related to that of specific lipid–protein and protein–protein interaction mechanisms. Several selective lipid–protein interaction mechanisms based on electrostatic and elastic interactions have been established. Electrostatic mechanisms can induce the recruitment of extrinsic proteins. Two prominent biological examples are: (i) the Ca^{2+}-mediated coupling of proteins of the annexin family to negatively charged lipid membranes and (ii) the adsorption of protein kinase C to the inner leaflet of plasma membranes containing PS and PIP2 [20]. The adsorption of kinase C is mediated by 13 basic peptides and is assumed to be essential for the activity of the protein. Most importantly, the electrostatic protein binding is often accompanied by simultaneous anchoring of the extracellular domain through hydrocarbon tails (cf figure 3(a) right side). Neutron surface scattering studies of hisactophilin (an actin–membrane coupling protein) showed that the hydrophobic interaction through fatty acid chains can determine the binding energy and thus serve the control of the functional orientation of the proteins [8]. Moreover, the electrostatic coupling can be enforced by penetration of peptide sequences (cf [8] (figure 3(a))) or aromatic side chains [20] into the semipolar region of the bilayer.

The elastic selection mechanism (illustrated in figure 6) is based on the matching of the lengths of the hydrophobic domains of integral proteins and bilayers. It is sometimes called the mattress effect [21]. Experimental evidence for this mechanism was first established by calorimetric studies of photosynthetic reaction centres in membranes composed of lipids of different chain lengths by the Möhwald group (cf [2b] for references). A length mismatch of $\delta d \sim \pm 0.5$ nm (four CH_2-groups) leads to a shift in the transition temperature of $\Delta T_m \sim \pm 5\,^{\circ}$C. As expected, T_m is shifted to higher (lower) temperatures if the bilayer is thicker (thinner) than the protein, since the formation of chain defects leading to shortening

(lengthening) of the chains is impeded (favoured). From the freezing point depression one can estimate the energy costs of the hydrophobic mismatch δd [2b]. For $\delta d \sim 0.5$ nm a value of $\Delta G_{LP} \sim 1000\ k_B T$ was found. The mattress effect has been extensively studied by molecular statistical theories of hydrocarbon chain packing. The elastic free energy of deformation [23] has been calculated as a function of the mismatch δd by considering both the energy cost associated with the stretching (or compression) of the chains and with the change in surface tension due to the local variation of the head group density near the protein surface. The free energy per unit length of the protein perimeter for a thickness difference of $\delta d \sim 0.5$ nm is $\delta G \sim 5\ k_B T\ \text{nm}^{-1}$ [22]. This corresponds to a total elastic energy cost for the photosynthetic reaction centre (with $2\pi R_0 \sim 21$ nm) of $\Delta G \sim 200\ k_B T$ (which is five times smaller than the calorimetric result). The molecular statistical model also provides information on the persistence length ζ_L characterizing the extent of the elastic distortion of the bilayer in the lateral direction. The value of $\zeta_L \sim 1.5$ nm corresponds to two layers of distorted lipids which are denoted as boundary lipids. Experimental evidence for such boundary lipids is summarized in [2b]. As indicated in figure 6 the observed motional freedom of cholesterol along the lipid chains predicted by QENS provides a possible mechanism of relaxation of the hydrophobic mismatch since it can enlarge the bilayer thickness by 0.4 nm [15].

2.3.3. Short and long range protein–protein interaction. Several mechanisms of protein–protein interaction mediated by the lipid matrix are conceivable. These have been studied only by molecular statistical theories [23] and Monte Carlo simulation procedures (see [24]) as yet. If the hydrophobic thicknesses are matched, two types of short range interactions are predicted. At distances r (between the protein surfaces) smaller than the lipid diameter a_L, attractive depletion forces prevail, equivalent to those predicted for colloidal solutions. Interaction energies of the order of several tens ($\sim$30) $k_B T$ have been predicted for a 10 nm diameter protein [23]. These short range forces may be responsible for the formation of voltage sensitive ion channels, such as Ca^{2+} pumps or band III anion channels [25] by aggregation of protein monomers. At larger distances ($r > a_L$) interactions arise due to the perturbation of the lipid order parameter by the proteins. Exponentially decaying attractive forces are predicted by Monte Carlo simulations [24] with correlation lengths of the order of $\xi_l = 5\ a_L$. In contrast, the molecular statistical model predicts weak repulsions [23] for situations of hydrophobic matching, while attraction arises only if mismatch is introduced.

The protein–protein interaction mediated by perturbation of the chain order parameter can be also described by the Landau mean field theory of lipid order parameter fluctuations (for references cf [2b]). An advantage of this approach is that the strength and range of the inter-protein interactions can be related to thermodynamic and mechanical parameters of the bilayer. Direct experimental measurements of inter-protein interaction are very difficult and are still lacking. One possible approach is to analyse the random walk of proteins in dilute solutions by single molecule fluorescence techniques and to compare the result with Monte Carl simulations of random walks [24].

2.4. Thermo-elastic concept of microdomain formation in membranes

Order in cells is maintained by the constant bidirectional flow of material from the ER over the Golgi apparatus to the PM that is mediated by vesicles [1]. Budding and fission of vesicles from the donor compartment and their fusion with the acceptor compartments are thus processes of outstanding importance for the survival of cells. To maintain the distinct lipid and protein composition of the compartments, mechanisms of sorting of lipids and proteins must be implemented in these processes. It is evident that this requires that the budding and fission

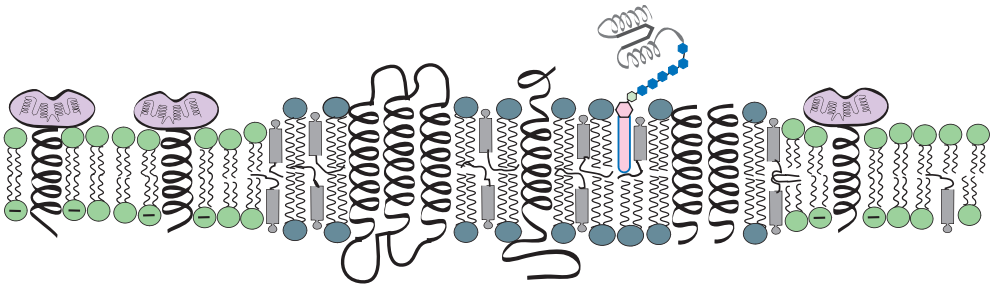

Figure 7. Schematic view of phase separation in bilayer of pseudo-ternary lipid mixture induced by matching of hydrophobic thicknesses of lipid bilayer and proteins. The situation of plasma membranes is shown with sphingomyelin and cholesterol forming the thick and non-saturated lipids the thin phases.

of vesicles is preceded by the formation of domains of distinct composition by lateral phase separation. Three types of such microdomains are currently considered: rafts, caveolae and coated pits. Only the second and third types are clearly observable by electron microscopy and are therefore well established.

In the following we provide evidence that microdomain formation is a natural consequence of the thermo-mechanical properties of the natural membranes which can be considered as pseudo-quaternary lipid–protein alloys. Due to the large elastic energy associated with the hydrophobic mismatch (several hundred k_BT for $\delta d \sim 0.5$ nm), lateral segregation of length adapted lipids and proteins is expected to occur in binary mixtures of lipids exhibiting different chain lengths, even if the lipid mixture alone would form a homogeneous phase. To minimize the elastic energy, the length adapted lipids and proteins are expected to segregate as indicated in figure 7.

A detailed evaluation of the structure of the hydrophobic domains of several hundred integral proteins in cellular compartments [27] showed that the integral proteins of the Golgi apparatus exhibit on average α-helical anchors composed of 15 residues. This corresponds to a hydrophobic length of $\langle H \rangle \sim 2.2$ nm. Proteins residing in the plasma membrane are 20 amino acids or $\langle H \rangle \sim 3.0$ nm long. This finding suggests that the elastic forces associated with the hydrophobic mismatch provide a major driving force for the sorting of proteins between these two compartments. Since the Golgi membranes contains much less of the long chain SPM and of the bilayer expander cholesterol than the PM it cannot accommodate the proteins with long hydrophobic domains. In membranes containing proteins with long and short α-helices (e.g. of the Golgi and the PM), lateral phase separation into a membrane fraction rich in SPM and cholesterol and a fraction rich in the low melting phospholipids and cholesterol (each containing the length adapted proteins) is expected to occur (cf figure 7). This situation holds for the compartments formed by fusion of the endosomes with the lysosomes or with vesicles of the Golgi (cf appendix B). If these compartments decay again into vesicles these will be enriched in one of the two phases. The hydrophobic matching could also provide a means to couple the biosynthesis of protein and lipids in such a way that the *de novo* synthesis of an integral protein engenders the synthesis of an appropriate number of length adapted lipids.

Despite numerous studies [26] a clear experimental proof of flat raft formation in biological membranes is still missing. Unfortunately, systematic studies of realistic model systems composed of raft-like mixtures and proteins of different hydrophobic thicknesses are still missing. Judged from triton solubilization experiments, their size is estimated to be of the order of 50 nm, and it is thus difficult to distinguish them from the budded microdomains. The

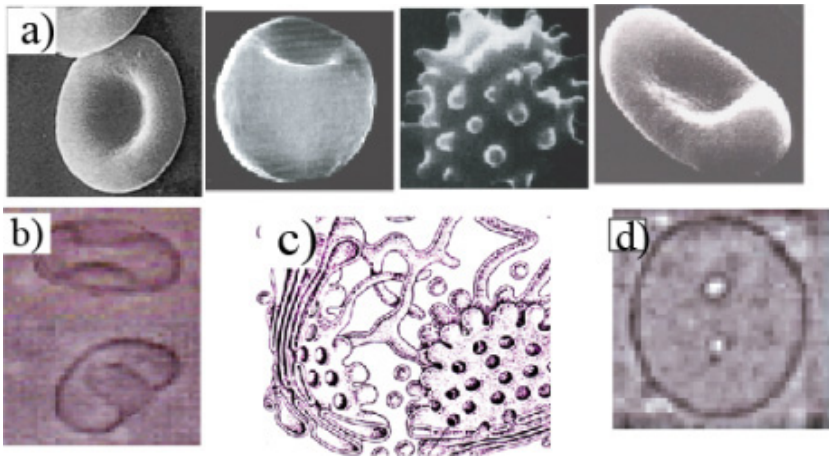

Figure 8. (a) Major shape transitions of erythrocytes. From left to right: discocyte, stomatocyte, echinocyte and elliptocyte. The discocyte to stomatocyte transition is often induced by cholesterol depletion; echinocyte formation is a consequence of aging while elliptocytes are formed by removal of the band IV.1 actin–membrane linker (cf figure 1(c)). (b) Discocyte to stomatocyte transition of single component vesicle induced by increasing the membrane excess area through heating by ∼1 °C after [33]. (c) Schematic view of ER forming collapsed vesicles with passages ('worm-holes') that are interconnected by branched tubes. (d) One component vesicle forming shape of toroidal topology after [34].

observed size could in fact be determined by the solubilization process itself. Considering the rapid turnover of PM and Golgi membranes, a possible explanation is that microdomains rich in SPM and glycolipids are manifestations of the fusion of vesicles split off the (late) endosomes or lysosomes with the PM. Some evidence for raft formation has been provided by elegant single molecule fluorescence studies of Brownian motion of different lipid probes [28]. The time dependence of the mean square displacement $\Delta X(t)^2$ of (SPM) lipids with long saturated chains showed non-Brownian behaviour $\sim: \Delta X^2 \sim t^{0.7}$ while short chain lipids showed normal Brownian behaviour ($\Delta X^2 \sim t$). It should be noted, however, that similar behaviour would also be observed if the SPM lipid probe resided preferentially in budded domains or if proteins were transiently immobilized by coupling to the actin cortex.

3. Global shape transitions of cells and cellular compartments

3.1. Basic concepts and model membrane studies

An impressive demonstration of the key role of general physical principles in biology is the control of the shape and the shape transitions of cells and intracellular compartments by the principle of minimum bending energy of soft shells [29–31]. Experimental evidence for this concept was provided by the finding that single component vesicles can undergo the same shape changes as erythrocytes (such as the discocyte to stomatocyte transition shown in figure 8(b)) by simply changing the area to volume ratio [33]. Even the complex shapes of the Golgi apparatus and the ER consisting of assemblies of flat vesicles perforated by passages (worm-holes) that are interconnected by branched tubes can be mimicked by single component vesicles (figure 8(c), [34]).

The manifold of observed shapes and routes of shape transitions have been predicted on the basis of the Helfrich (HE) model, the bilayer coupling (BC) model or the most general

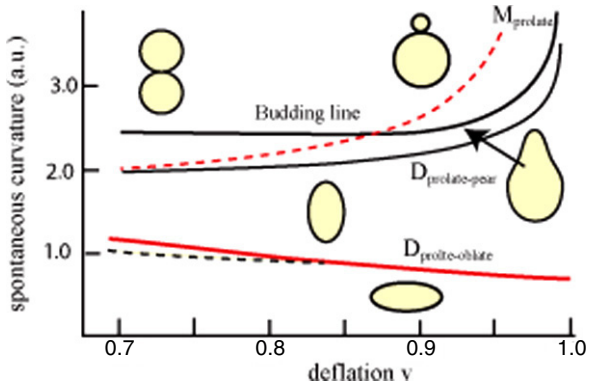

Figure 9. Phase diagram of shape transitions of vesicles (after [35]) predicted by the area difference elasticity (ADE) model with $\alpha = 1.4$ (predominance of BC model) in the phase space defined by v and $\bar{C}_0$. The drawn lines D_{op}, D_{pp} and D_{bud} denote first order shape transitions. The dashed lines define spinodal lines. At large degrees of deflations v (<0.85), the spinodal line has to be crossed for budding, which can slow down the transition drastically.

approach: the area difference elasticity (ADE) model, which interpolates between the two first approaches (cf appendix A).The energy functional of the ADE model can be expressed as [35, 36]

$$\Delta G_{\mathrm{ADE}} = \kappa/2 \left\{ \int\int \left[2H^2\right] \mathrm{d}O - 2\bar{C}_0 H + \right\} + \frac{\pi\kappa^*}{A} \left[\int\int 2H \,\mathrm{d}O \right]^2$$
$$\bar{C}_0 = C_0 + \frac{\pi\alpha\Delta A_0}{Ad}. \tag{4}$$

The shapes can be expressed in terms of three free parameters: (i) the ratio of the bending moduli $\alpha = \kappa^*/\kappa$ (which is about equal to one [32]); (ii) the reduced volume $v = 3V/4\pi R_{\mathrm{e}}^3$ (expressed as the ratio of the actual volume V of the shell to that of a sphere with the same area and radius R_{e}) and (iii) the effective spontaneous curvature $\bar{C}_0$. The latter depends on the spontaneous curvature C_0 and the initial area difference ΔA_0 defined in the appendix A. These do not enter independently and $\bar{C}_0$ should thus be considered as an empirical parameter. α is a measure for the relative contributions of the Helfrich and the area difference model.

Phase diagrams of shapes have been calculated with all three models by minimizing the free energy under the constraint of constant area and volume of the shells. All models yield essentially the same qualitative shape phase diagram and show that all observed symmetric shapes (including vesicles of toroidal topology) can be reproduced by changing the universal parameters $\bar{C}_0$ and v. However, the models differ (i) concerning the prediction of the order of the transitions, (ii) the existence of metastable phase lines and (iii) the trajectories within the phase diagram followed by the vesicles during changes of the control parameters. Figure 9 shows the phase diagram obtained by the general ADE model for $\alpha = 1.4$, which has been well verified experimentally by tedious measurement of the spontaneous curvature [35]. It provides some interesting insight into the budding process and the stability of erythrocyte shapes. First, the prolate and oblate shapes are stable over a large range of excess areas. Second, the transition from prolate ellipsoids to pear shapes and (discocyte-like) oblate ellipsoids, respectively, are discontinuous first order transitions, while the pear shape and the budded state may be separated by a spinodal line (SP). The budding transition may therefore be drastically delayed, rendering experimental studies tedious and time consuming. The first of the above findings provides an explanation for the stability of the discocyte shape over large range of osmotic pressures. The

Table 1. Geometric parameters (first row) and elastic constants (second row) of erythrocyte shell. Definition of elastic moduli (see appendix A).

Area: $A = 140\ \mu\text{m}^2$, Volume: $V_0 = 95\ \mu\text{m}^3$: Equivalent radius: $R_e = (A/4\pi)^{1/2} = 3.3\ \mu\text{m}$
$\kappa = 50\ k_B T \quad \mu = 2.5\ \mu\text{J m}^{-2} \quad K = 7.5\ \mu\text{J m}^{-2} \quad \kappa^* = 2\kappa/\pi$

spinodal slowing down of the budding may be the main reason why this first step of intracellular transport by vesicles is always induced by macromolecular adsorption. In particular, the budding of oblate shells exhibiting high negative spontaneous curvatures, such as the ER and the Golgi, could hardly be realized by changes of the global spontaneous curvature.

3.2. The shape transition of the composite shell of erythrocytes

The single shell model can explain the dominant discocyte–stomatocyte transition of erythrocytes astonishingly well, but it cannot account for the formation of echinocytes, irregularly shaped stomatocytes or elliptocytes. Three types of echinocytes can be distinguished: (i) oblate shapes with soft bumps protruding from the edge, (ii) quasi-spherical sea-urchin-like cells with up to 240 spicules and (iii) spherical cells with sharp spicules reminiscent of buds. As pointed out first by Stokke (cf [36] for references), spicule formation requires the introduction of a new length scale to define the size of the local protrusions. It comes in with the shear elasticity associated with the coupling of the bilayer to the cytoskeleton and is $\Lambda_s = \sqrt{\kappa/\mu}$. With the data of table 1 one expects a value of $\Lambda_s \sim 0.3\ \mu\text{m}$, in good agreement with the example shown in figure 8(a).

For a complete description of the mechanics of RBCs one has to consider the interaction between the two soft sub-shells. The simplest assumption is that the plasma membrane determines the bending elasticity of the composite shell, while the cytoskeleton controls the shear elasticity but does not contribute to the bending stiffness. The total elastic energy of the cell envelope is thus determined by the sum of the elastic free energies (cf appendix A).

$$\Delta G_{\text{Ery}} = \Delta G_{\text{Cyt}} + \Delta G_{\text{ADE}}. \tag{5}$$

In a detailed analysis of the variational problem performed by the Wortis group [36] a large number of shapes (including stomatocytes with triangular symmetry) were calculated as a function of the effective spontaneous curvature $\bar{C}_0$ and the degree of deflation v. The ground state of the erythrocyte is assumed to be the slightly crumpled baby cell (reticulocyte) which transforms into the discocyte during the maturing process (cf [2a]). This process is possibly associated with pre-stressing of the cytoskeleton. Some evidence for such an effect comes from freeze fracture EM studies, showing that the bilayer is locally undulated with amplitudes of 10 nm and wavelength of 100 nm corresponding to a relative excess area of about 1%. Computer simulations also suggest a smaller area of the cytoskeleton [4]. This reservoir of excess area is stored in the dynamic surface roughness associated with the thermally excited bending excitations. Fortunately, the effect of pre-stressing on the cell shape can be accounted for by renormalization of the elastic moduli (cf [36] for a detailed discussion). The MC simulations based on these ideas show that the discocyte is stable over a very large range of degrees of deflation and spontaneous curvatures ($\bar{C}_0 \approx -5$ to $+8\ \mu\text{m}^{-1}$), in agreement with numerous experiments. Stomatocytes form by a discontinuous transition at $\bar{C} \leqslant -10\ \mu\text{m}^{-1}$ while echinocytes form at positive spontaneous curvatures $\bar{C}_0 \geqslant +10\ \mu\text{m}^{-1}$. The width of the spicules decreases with increasing spontaneous curvature. They go over in buds exhibiting narrow necks when the width becomes smaller than the buckling length Λ_s. An important consequence of the cytoskeleton membrane coupling is the generation of local gradients

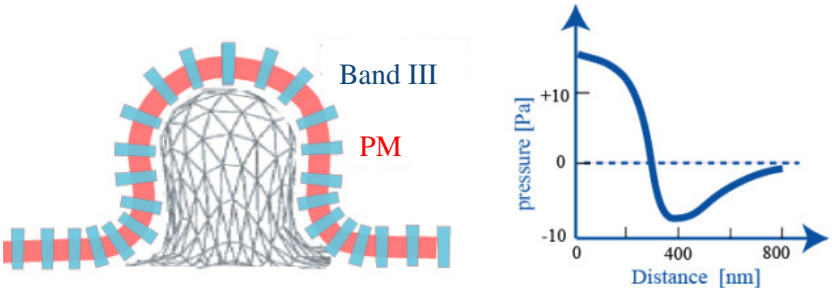

Figure 10. Generation of lateral gradients of membrane tension and normal pressure on membrane induced by bilayer cytoskeleton coupling (after Mukhopadhyay *et al* [36]). Note that the cytoskeleton pulls the PM towards the inside at the tip (lateral pressure positive). Note that the tension gradients are expected to induce a redistribution of the band III proteins coupled to the cytoskeleton, as indicated on the left side.

of lateral tension which is expansive in concave regions (such as the tips of spicules) and compressive at convex regions (such as the base of spicules). This local tension gradient leads to a locally varying osmotic pressure in the direction of the membrane normal $p_\perp$ (as predicted by the LaPlace equation). The cytoplasm pulls inward at the tip of the spicules with a pressure of $p_\perp = +10$ Pa and outward at the base (with $p_\perp = -5$ Pa) for a spicule formed at $\bar{C}_0 \approx +10\ \mu\text{m}^{-1}$. One possible consequence of the spatial variation of the tension and the associated density of the spectrin network is the non-isotropic distribution of proteins. Thus the band III density is expected to be higher at the concave central regions of the discocyte. Electron microscopy studies in the author's laboratory showed indeed that the density of band III is higher by a factor of 1.3 at the centre of the discocyte than at the rim (cf figure 10).

The lateral inhomogeneity introduced by the bilayer–cytoskeleton coupling may also help to explain the shape changes introduced by diseases. One example is the formation of ellipsoidal biconcave shapes by the genetic disease elliptocytosis. This break of cylindrical symmetry is caused by the lack of band IV and the decoupling of the actin oligomers from the PM. The original triangular symmetry of the cytoskeleton PM links is broken, which can give rise to a non-isotropic lateral tension with respect to the symmetry axis. Electron microscopic studies show indeed that the spectrin filaments tend to form elongated bundles and the membrane proteins form clusters [37]. Another example is spherocytosis caused by removal of the ankyrin linker. The cells assume spherical shapes, suggesting that this linker plays a key role for the generation of the negative spontaneous curvature.

In summary, complex non-isotropic shapes can be generated by coupling of shear elastic networks to single component bilayers. In composite membranes the shapes are further determined by coupling between curvature and lateral phase separation. In fact complex shapes such as echinocytes can also be generated by lateral phase separation [38]. The shapes of intracellular compartments are most likely determined by this mechanism. In fact, flat vesicles exhibiting passages (worm-holes) reminiscent of the morphology of the ER can form if two components have a strong tendency to segregate into two phases exhibiting different positive spontaneous curvatures [38]. The shape of cells and intracellular compartments is thus determined by both factors: their interaction with macromolecular networks and the coupling between curvature and phase separation.

The composite design of the envelope provides red blood cells with a unique combination of flexibility and mechanical stability that allows them to survive a several hundred kilometre long journey through the capillaries of the blood vessels during their 120 day lifetime

(cf [1, 2a]). The (shear force induced) discocyte-to-stomatocyte transition is essential to minimize the frictional resistance [39a], but the strain softening of the network may also be important [39b]. The loss of ions during the very strong deformations is avoided by the very small compressibility of the bilayer.

4. Curvature elasticity concept of intracellular trafficking and material export and import

The intracellular material transport between the ER and the plasma membrane or the material import (endocytosis and pinocytosis) and export (exocytosis) mediated by vesicle budding fission fusion processes are most prominent cellular processes controlled by coupling of curvature and phase separation. Although about 2% of the total PM can be turned over every minute, the distinct composition of the compartments is maintained. Nature uses two pathways for endocytosis: caveolae and coated pits. Both are formed by local curvatures induced by macromolecular adsorption but they transport materials to different targets (cf [1]). Vesicles internalized by coated pits deliver material to specific compartments (called early endosomes) or lysosomes (cf appendix B), while caveolae can transport the materials directly to the Golgi or the ER. This pathway is also used by pathogens (cf [40] for references) or toxins such as cholera toxins, enabling them to circumvent degradation by the lysosomes. Coated vesicles also play a key role for the intracellular transport between the ER and the Golgi [1].

In the plasma membrane coated pits are generated by coupling of clathrin to the intracellular domains of receptors, such as the receptor of the ion carrier transferrin (cf appendix B). The coupling is enabled by adaptor proteins (APi) such as AP2 in figure 11, which are recruited to the membrane by small helper proteins, the binding of which is initiated by exposure of a fatty acid chain triggered by exchange of GTP for GDP. By using adaptors as mediators between the membrane and clathrin, the same universal mechanism can be used for the vesicle transport from the trans-Golgi to the PM (with an adaptor AP1). Moreover, the same universal principle (but with a different coat protein) serves the bidirectional transport between the Golgi and the ER. The formation and fission of clathrin coated pits occurs in several steps that are controlled by different proteins. Initially, shallow indentations are formed by adsorption of patches of clathrin. The rapid formation of highly curved invaginations is driven, however, by binding of a second set of helper proteins such as dynamin and epsin [1, 41]. Below we provide evidence that these proteins serve the reduction of the elastic edge energy at the neck of budded vesicles during the final step of vesiculation. They can also induce the formation of long tube-like protrusions that penetrate deep into the cell.

Consider first the bending elasticity concept of clathrin coat formation (cf figure 11). Clathrin is composed of three 14 nm long leg-like protein filaments that are coupled to a central domain, thus forming a triskelion. The binding domain for the adaptor is located at the feet. Although the legs are rather stiff, they can fluctuate relatively freely about the central domain or the knee, which can facilitate the formation of highly curved networks. On planar surfaces large patches of hexagonal networks (of lattice constant $a \approx 10$ nm) are formed by parallel assembly of two legs, while fully developed coats exhibit icosahedron-like architectures. The formation of highly curved surfaces by clathrin networks can be understood in analogy to the formation of small vesicles by lipid bilayers in the lipid gel state. Highly curved (or crumpled surfaces) of 2D hexagonal crystals can be generated by dissociation of edge dislocations (shown in figure 4) into isolated $-60°$- and $+60°$-disclinations [42, 43]. A $-60°$ disclination is formed by cutting out a wedge of $60°$ ($\pi/3$ rad) from a flat lattice. The elastic energy cost of such a process in a flat surface of dimension L^2 is $\Delta E_{\text{disc}} = (\pi/3)^2 K_0 L^2$ (where K_0 is the Young modulus) and thus growth is quadratic with the size. It is reduced dramatically if the sheet

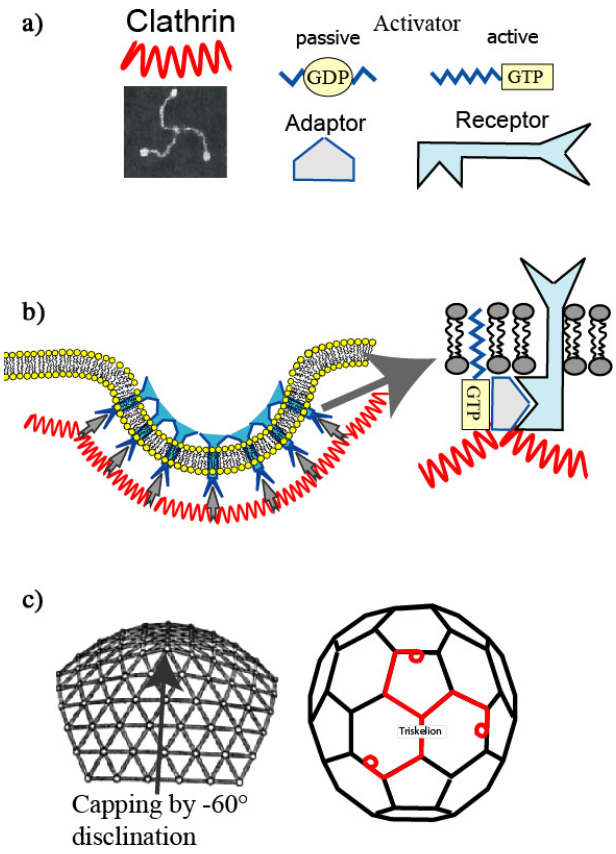

Figure 11. Model of local budding and vesiculation of plasma membrane by adsorption of clathrin coat. (a) Summary of major proteins involved in coat pit formations. (b) Binding of the three-legged clathrin molecule to the inner surface of the PM is mediated by adaptor proteins that bind the feet of the clathrin triskelion to the inner domain of receptors. Each adaptor is recruited to the membrane by an 'activator' protein that binds to the membrane through a fatty acid anchor. The initially hidden lipid anchor is exposed after replacement of guanosine diphosphate (GDP) by an energy rich GTP, a process triggered by an exchange factor [1]. The bending moment inducing curvature is generated by adaptor induced expansion of the inner monolayer (cf also figure 13). (c) Texican-hat-like caps are formed by isolated $-60°$ disclinations. resulting in the formation of an icosahedral spherical shell.

is allowed to bend to a value $\Delta E_{\rm disc} \approx \kappa \ln(R/a)$ and thus depends only weakly on the size of invagination. The Young modulus of 2D elastic plates is of the order of the shear modulus ($K_0 \approx 2\ \mu$). For coated membranes the shear modulus is determined by the clathrin coat (assumed to be of the order $\mu \sim 10^{-1}$ J m^{-2} [43]) and the bending modulus by the membrane ($\kappa \sim 0.50\ k_{\rm B}T$). The energy cost for the generation of a disclination in a flat clathrin plate of $L \sim 100$ nm is $\Delta E_{\rm disc} \sim 10^5\ k_{\rm B}T$, while for a curved surface of curvature L^{-1} it is reduced to $\Delta E_{\rm disc} \sim 100\ k_{\rm B}T$. This estimation suggests that clathrin coats exhibiting icosahedral symmetry can form on spontaneously curved surfaces above a threshold curvature C_0^*. A mean field theory based on the above ideas predicts that the minimum curvature increases with the radius, R, of the clathrin patch [43], while a Monte Carlo simulation study suggests that C_0^* decreases with R [44].

Apart from this discrepancy the above theoretical models can explain the formation of coated pits of icosahedral-like topology. They do not provide an explanation for the origin of the

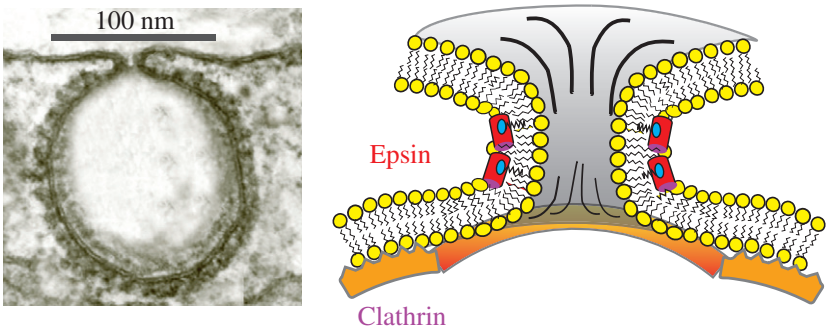

Figure 12. Left: electron micrograph of coated pit (adapted from [79]) showing the formation of a sharp edge at the transition from the coated membrane to the PM radius $\rho_n \sim 15$ nm. Right: model of formation of sharp neck by epsin. Note that without epsin the curvature at the neck would be determined by the persistence length of the bilayer deformation $\zeta_p \sim 100$ nm.

bending moment driving the curvature, which is a necessary step to dissociate the dislocations. A more likely explanation is that the strong bending moment is generated by binding of the adaptor proteins (such as AP2 in figure 11) that penetrate with fatty acid chains into the cytoplasmic monolayer. We can estimate the number of AP2 fatty acid anchors needed to induce a local radius of curvature of 50 nm. For a 1:1 cholesterol–PC mixture the compression modulus is 0.7 Pa m and $\kappa \approx 50\ k_B T$ (cf [2a]). The bending moment induced by incorporation of 1% of excess lipids in a monolayer is $M = \Delta\pi d_m = K d_m\ (\Delta A/A)$ and the radius of spontaneous curvature thus becomes $R_0 \approx \kappa/M \sim 10$ nm ($M \approx 10^{-11}$ N), demonstrating that the insertion of 1% of adaptor molecules (corresponding to the density of the vestiges of the clathrin network) is sufficient to induce the formation of coated pits of 60 nm diameter. In this model the major role of the clathrin coat would be to localize the adaptor molecules and the associated receptors.

4.1. On the problem of the neck formation and vesicle fission

As noted above, under real biological situations a second set of proteins such as dynamin and epsin [41, 45] is required to drive the formation of narrow necks. Moreover, dynein has been postulated to act as an ATP dependent pinching machine mediating vesicle fission [45]. But the following observations contradict this hypothesis and suggest that the main role of epsin and dynamin is to control the neck formation. The binding of epsin to membranes is mediated by a specific lipid: phosphatidylinositol-4,5-biphosphate (PtdIns(4,5)). An amphiphatic helix forms after binding to this specific phospholipid (cf figure 3). Its function as curvature inductor was demonstrated by the finding first that epsin and dynamin accumulates at the edge of the growing clathrin patches, and second that it induces tube-like protrusions from liposomes containing ∼10% of PtdIns(4,5). Most important, the binding is specific (and not due to charge effects) since no effect is observed if vesicles are doped with phosphatidylserine [41].

A major role of these neck forming molecules may thus be to allow for the change of curvatures at the edge between the rather stiff coat and the uncoated PM (cf figure 12), which requires the compensation and reversal of the clathrin induced spontaneous curvature. As shown in the chapter on adhesion, the spontaneous curvature at the edge of the clathrin coat exhibits a persistence length of ∼100 nm, while the local curvature at the neck changes sign over a distance of 10–20 nm (cf figure 12, left). Such sharp edges could be generated by the

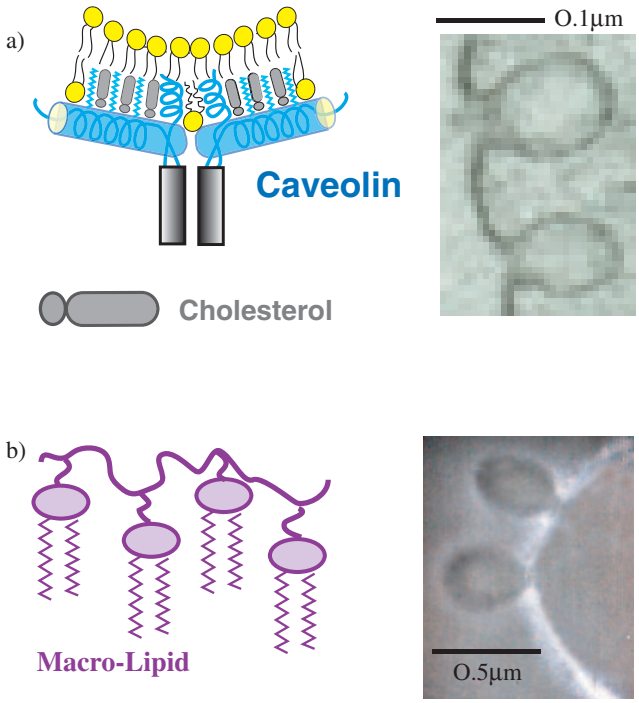

Figure 13. (a) Cartoon of induced bending of plasma membrane by penetration of fatty acids and peptide loop of caveolin molecules into the cytoplasmic monolayer of the PM. Right: typical shape of flask-like buds (adopted from [80]). (b) Formation of buds by the adsorption of macrolipids (after [81]).

penetration of the amphiphatic epsin helices into the outer leaflet of the PM (cf figure 12, right). The associated thinning of the membrane would reduce the bending stiffness drastically, thus facilitating the formation of sharp bends. An open question is the driving force of fission. Model membrane studies suggest that fission can be mediated by accumulation of solutes at the highly curved neck. Cholesterol induces fission of buds from the lipid vesicle at 50 mol% or higher, while at lower concentration fission is a very rare event [12]. More recent research suggests that dynamin generates fission promotors [46]. It activates an enzyme (of a protein family called endophilins) which transforms a lyso-phospholipid into a two chain phospholipid by coupling of arachidonic acid, (a bulky fourfold unsaturated fatty acid in *cis*-conformation). Similar to the above mentioned effect of cholesterol, this bulky lipid could also destabilize the narrow neck.

The second type of nano-invaginations of cell envelopes, the caveolae, are more flask-like. The inductor of the curvature [47–49] is caveolin. This protein exposes three hydrocarbon chains and a hydrophobic loop that penetrate into the cytoplasmic monolayer, thus generating a strong bending moment. This budding mechanism is more generic than that mediated by clathrin. It can be realized with pure lipid bilayers by adsorption of macrolipids (cf figure 13(b), [44]). An important specific property of caveolin is is the capacity to bind cholesterol [1]. Since the caveolin coated vesicles commute between the PM and the Golgi, they could play a key role for the uptake of cholesterol at the Golgi and its transport to the plasma membrane during exocytosis. Another specific feature of coated pits is the capacity to interact with actin filaments and microtubules to facilitate the intercellular transport [47].

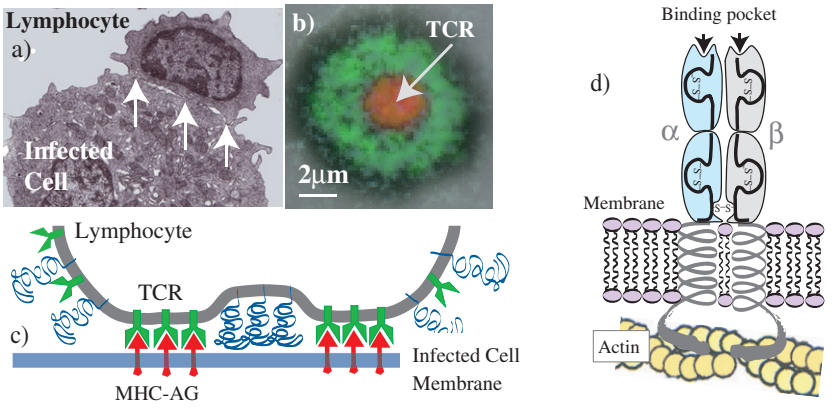

Figure 14. (a) A thin section electron micrograph of lymphocyte adhering on an infected antigen presenting cell (APC), clearly showing that the contact zone decays into domains of tight adhesion (indicated by arrows) that are separated by non-adhering regions (modified after [82]). The adhesion domains are formed by binding of antigens to T-cell receptor (TCR), a CAM belonging to the class of membrane bound antibodies. Note that antigens are only recognized if they are bound to cell surface molecules called major histocompatibility complex (MHC). (b) Fluorescence micrograph of adhesion zone formed by a lymphocyte of the T-cell family adhering on solid supported membranes that was doped with antigen carrying MHC proteins (modified after [83]). The red (dark grey) central region indicates a cluster of T-cell receptors. It is surrounded by a green (light grey) ring that is rich in CAMs acting as repellers. (c) Illustration of decay of zone of cell–cell contact into aggregates of receptor–ligand pairs that are separated by regions of weak contact. (d) Structure of T-cell receptor: a member of the immunoglobulin superfamily. The extracellular head groups of this family of CAMs are composed of (two to seven) domains exhibiting immunoglobulin structures while the intracellular domain exhibits an actin binding domain [51].

This work suggests.that actin binding could also be responsible for the assembly of caveolin in worm-like structures [50].

In summary, the formation of buds by the generic mechanism of induced curvature is well established but it is much more complex than generally assumed. Many proteins are involved and it is not clear yet which of these provides the main driving force for the generation of strongly curved necks. Despite great progress in recent years we are far from an understanding of the molecular mechanisms driving the detachment of the buds or their fusion. Another key question for further research concerns the question of sorting of lipids during the budding. Numerous theoretical and experimental studies of the physical basis of budding over the last 20 years were concerned with lipid bilayers, while experiments with lipid–protein bilayers are unfortunately still rather rare.

5. Cell adhesion as interplay of interfacial and elastic forces

In this section we show that the elasticity of soft shells also plays a key role for the control of cell adhesion and for the adhesion induced formation of functional domains. As an example we discuss the generation of immunological synapses which are supposed to play a key role for the stimulation of the lymphocyte. The adhesion between cells is mediated by specific interaction between integral proteins exposed by the plasma membranes, so called cell adhesion molecules (=CAM, [51]). In the case of cell tissue interaction, the CAMs (called receptors) recognize specific groups (called ligands) that are exposed by the macromolecules composing the tissue (such as collagen or fibronectin). As an example, figure 14 shows the recognition of a

virus infected cell exposing antigens (called an antigen presenting cell, APC) by a lymphocyte (belonging to the class of T-cells) through binding of the antigen (the ligand) to the membrane bound antibody (the T-cell receptor, TCR). The thin section electron micrograph shows that the area of contact between the cells decays into patches of tight contact (called adhesion domains in the following) that are separated by unbound regions. The adhesion domains form by aggregation of tightly bound receptor–ligand (R–L) pairs. Cell adhesion through adhesion domains is a universal concept of design of living matter. It serves the generation of layers of interconnected cells such as the endothelial cell monolayers lining the inner wall of blood vessels or epithelial cell layers generating the skin of organs. It is applied to connect these cell layers (e.g. endothelial cells) to the tissue surfaces (e.g. the basal membrane forming the inner surface of blood vessels) and plays a key role for cell crawling on surfaces. This suggests that adhesion through local contact formation offers some strategic advantages and that adhesion may be controlled by physical principles.

5.1. Modelling of cell adhesion processes and measurement of adhesion strengths

Numerous studies of the biochemical basis of cell adhesion suggest that contact formation is controlled by the interplay of the specific short range attraction forces (also called lock and key forces) mediated by the receptor–ligand bonds and the generic repulsion forces generated by CAMs that expose large head groups (extending up to 40 nm into the extra-cellular space such as the glycoprotein CD 43 [51]). In the case of cell tissue interaction, repulsion forces can also be generated by giant polysaccharides of the tissue such as the polysaccharide hyaluronic acid, that can exhibit hydrodynamic radii of up to 300 nm [1]. A powerful strategy to gain insight into the physical basis of cell adhesion is to establish model systems containing the major ingredients of adhesion such as shown in figure 15. Giant vesicles with reconstituted ligands serve as test cells, while solid supported membranes containing the conjugate receptors act as target cells or tissue. Owing to the planar design the time evolution of the adhesion process can be followed by high resolution micro-interferometry such as reflection interference contrast microscopy (shown in figures 15(b) and (c)). Adhesion induced contour changes of the adhering cells close to the substrate can be evaluated by reconstructing the cell surface profiles [52]. Numerous studies of various model systems by this technique showed that adhesion is indeed inevitably accompanied by receptor segregation and formation of tight adhesion domains (cf [52]). Below we provide evidence that the domain formation is an inevitable consequence of the interplay of short range attraction forces long range repulsion forces and membrane bending elasticity.

For a rigorous theoretical treatment of cell adhesion one has to consider the global and local deformation of the shells and the redistribution of cell adhesion molecules between adhering and free parts of the shell. Such models have been developed for the simple case of shells (e.g. vesicles) exhibiting only bending elasticity that adhere through van der Waals forces [53] and receptor ligand interaction while the effect of the repeller molecules have not been considered yet. One conclusion of this work is that the global bending deformation energy (per unit area) is small compared to the thermodynamic energy associated with the redistribution of the receptors. The adhesion process is, however, determined by the local interfacial forces and pronounced membrane deformations within the contact zone, while the free section of the adhering shell (comprising typically 80% of the total surface) serve as reservoir for receptors and repeller molecules. This is advantageous since the micro-interferometry technique RICM introduced in figure 15 enabled us to analyse the surface profile of the adhering shell in the contact zone with high resolution.

To gain insight into the physics of the adhesion process it is helpful to consider the analogy between an adhering shell and a partially wetting fluid droplet (shown in figure 16(a)). The

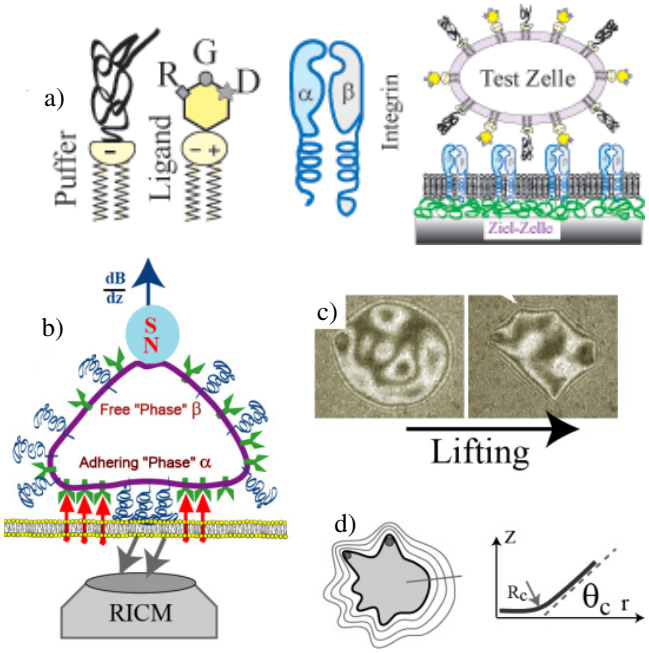

Figure 15. (a) Model of cell tissue interaction. Giant vesicles doped with lipid-coupled ligands composed of cyclic oligo-peptides containing an arginine (R)–glycine (G)–guanine (D) sequence of amino acids that is specifically recognized by integrins. To mimic the repulsion forces mediated by the glycocalix, about 1 mole % of lipids exposing macromolecular head groups (polyethylene oxide with $n = 2000$ monomers) are reconstituted into the outer monolayer. Supported membranes with reconstituted integrin receptors ($\alpha_i \beta_j$) act as target cells (or mimic the tissue). To create bio-analogue conditions and to avoid non-specific binding or denaturing of the receptors the supported membranes are separated from the solid surface by ultrathin polymer cushions. Most importantly, the receptors are fully mobile while they are fixed if the membrane is directly deposited on solid surfaces. (b) Visualization of adhering vesicle by reflection interference contrast microscopy (RICM). Images are formed by interference of light reflected from the cell surface and the top surface of the (transparent) substrate, respectively. The shell can be lifted by magnetic tweezers. (c) Interferogram of cell before and after application of lift force. The leopard-like texture of the adhering zone is due to the local height fluctuations caused by the bending excitations (after [84]). The tight adhesion domains are indicated as sharp lateral protrusions of the contact line L and are revealed subsequently by lifting the shell from the surface by magnetic tweezers [52b]. (d) Left: schematic view of adhesion zone with adhesion domains indicated by circles. Reconstruction of profile of adhering cell perpendicular to contact line L (by inverse cosine transformation of the interferogram) yields contact angle Θ_c and curvature R_c^{-1}. From these data the work of adhesion and the membrane tension can be determined.

contour of the fluid droplet at the contact line is determined by the equilibrium of the interfacial tensions

$$\sigma_{12} = \sigma_{13} + \sigma_{23} \cos \Theta_c. \tag{6}$$

Since surface tension σ_{ij} is a measure for the work (per unit area) associated with an incremental change, δA, of the area of contact, equation (1) also provides information on the energy balance of the adhesion. The energy per unit area gained by the partial wetting is $W = \sigma_{12} - \sigma_{13} - \sigma_{23}$ (note that the contacts between interfaces 2–3 and 1–3 are lost but gained between 1–2). Together with equation (1) this yields the well known Young law.

$$W = \sigma(1 - \cos \Theta_c). \tag{7}$$

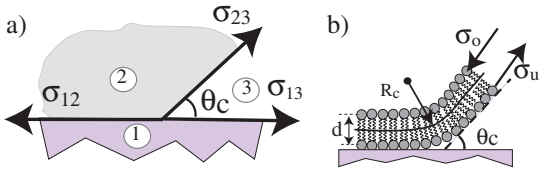

Figure 16. (a) Section through partially wetting fluid droplet at the contact line L. The edge is sharp on the atomic scale and the contour of the free area forms a contact angle Θ_c with the surface. σ_{ij} are the surface tensions between the materials i and j. (b) Section through adhering lipid bilayer showing the smooth transition between the free and adhering membrane at the contact line characterized by contact radius R_c. R_c is determined by the bending moment generated by the pair of tensions σ_u, σ_0 (cf appendix A). A contact angle Θ_c can be defined if the capillary length $\lambda_c = \sqrt{\kappa/\sigma}$ is of the order of a few micrometres, where σ is the adhesion induced membrane tension defined below [54].

Since W has also the dimension (N m^{-1}) of a two-dimensional pressure that determines the tendency for spreading it is also called spreading pressure.

Consider now the boundary conditions for the membrane. Since the membrane surface is also fully hydrated the Young equation still holds. In this case σ can be considered as the additional lateral membrane tension generated by the gain in adhesion energy. For fluid membranes σ is isotropic and equal over the whole shell surface. The elastic energy density associated with the finite contact curvature R_c is $\Delta g_{\text{elast}} \approx 1/2\,\kappa R_c^{-2}$, showing that a sharp edge would cost an infinite amount of elastic energy. A second equation for the work of unbinding W is gained by accounting for the boundary condition that the membrane left of the contact line L (in figure 16(b)) is parallel to the surface. The surface must exert a reaction force on the membrane that balances the bending moment generated by the pair of tensions σ_u and σ_0 at the two surfaces of the bilayer. An energy consideration yields[1]

$$W = \frac{\kappa}{2R_c^2}. \tag{8}$$

The boundary conditions, equations (7) and (8), provide the basis for the measurement of the adhesion strength W and membrane tensions σ, since the geometric parameters Θ_c and R_c can be determined by RICM and the bending modulus is often obtained through separate measurements (e.g. flicker spectroscopy). The bending stiffness can, however, also be measured simultaneously by analysing the changes of the surface contour by hydrodynamic shear [54] or magnetic lift forces [52b].

It should be noted that in the case of sharp protrusions of the contact line as revealed by lift forces the situation is more complex, since one is no longer in thermodynamic equilibrium. A second boundary condition accounting for the line force stabilizing the local protrusion of the contact line has to be considered, which is determined by the balance of the gradient of the bending moments. This opens the possibility for the measurement of unbinding forces of single R–L pairs under bioanalogue conditions [52b], which is not possible with single molecule spectroscopy. The behaviour of the contact zone under force depends on the mobility of receptors. If they are immobile, unbinding is achieved by relatively weak lift forces of some tens of piconewtons. If the receptors are mobile, the application of external forces leads to the growth of the adhesion domains [56].

[1] Note that the energy (per unit length of L) associated with the shift of the contact line perpendicular to L by an incremental distance δx is $M\frac{\partial \delta u}{\partial x} = M\frac{\partial^2 u}{\partial x^2}\delta x = \kappa R_c^{-2}\delta x$ (where δu is the deflection of the membrane in the vertical direction). This work is equal to the loss (or gain) in adhesion energy associated with the shift of the contact line, which is 2 $W\delta x$ (cf [55] (section 13) and [53]).

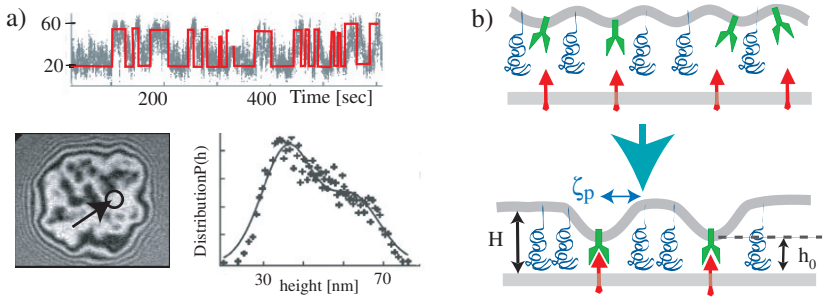

Figure 17. Demonstration of dynamic collisions of membrane with surface through local out-of-plane Brownian motions of membrane. (a) Top: local fluctuation of height of repeller doped membrane hovering over substrate. Measurement occurred at the circled area of the interferogram. The jumps between an unbound state at height $h \sim\leftarrow 60$ nm and a bound state at height of $h \sim\leftarrow 10$ nm are indicated by a thick dark line (after [85]). Bottom right: bimodal distribution of heights suggesting a double minimum of interfacial potential. (c) Model of local transient formation of R–L pairs induced by collisions of opposing membranes. Note local membrane deformation about the R–L pair extending over persistence lengths ζ_p (∼100 nm).

5.2. Cell adhesion as a first order wetting transition

The above consideration of the balance of forces at the contact line provided us with the tool to measure important physical parameters. For the understanding of the adhesion induced receptor segregation we have to consider the interfacial forces in the contact zone more closely. As illustrated in figure 17 the key point is the difference of the optimal interfacial distances enforced by the R–L pairs ($h \sim 10$ nm) and the repeller molecules ($H \sim 50$–100 nm), respectively. The formation of R–L bonds thus requires the local deformations of the membrane as indicated in figure 17. To understand the consequences of the deformation we consider the following energy balance.

$$\Delta G_{\mathrm{adh}} = \tfrac{1}{2}\kappa \int\int_{A_c} \Delta h(x,y)^2 \,\mathrm{d}O + w c_{RL} A_c + \tfrac{1}{2}\int\int_{Ac} V''(h-h_0)^2 \,\mathrm{d}O + G_{\mathrm{grav}}. \tag{9}$$

The first term accounts for the local bending deformation and the second for the gain in energy by the formation of c_{RL} R–L bonds per unit area, where w is the R–L-binding energy (typically 10 $k_B T$). The right integral accounts for the interfacial interaction, which is assumed to be determined by a harmonic potential of force constant $V'' = \mathrm{d}^2V/\mathrm{d}x^2$ [53]. G_{grav} is the gravitational energy [52b] that can provide a long range attractive potential. To determine the state of minimum energy one has to solve the Euler–Lagrange equation (as done in [58]). For our semi-quantitative consideration of the basic physical concepts it is sufficient to realize that through equation (11) a characteristic length $\zeta_p = \sqrt[4]{\kappa/V''}$ is introduced, as follows by a simple dimension analysis of the two integrals. This so-called persistence length is a measure for the lateral extension of the local deformation induced by the lock and key point force between R–L pairs (cf figure 17).

To estimate the contribution of the local deformation to ΔG_{ela} we have to consider the potential $V(h)$ more closely. It is determined by many contributions including electrostatic and dehydration forces (cf [59]). The experimental studies show, however, that the adhesion is controlled by two major contributions [52]: by the steric repulsion by the repeller molecules and by the undulation forces generated by the membrane bending fluctuations.

Consider first the repeller induced repulsion. Since the head groups of the repeller molecules are flexible as in the model systems (figure 15) or semi-flexible as in the case of CAMs, the repulsion force (per unit area) can be expressed in terms of the pressure generated

by macromolecules compressed between two rigid plates. If the inter-plate distance h is similar to the hydrodynamic radius R_g of the macromolecular head group the interfacial interaction potential can be approximated by the Dolan–Edwards law [60].

$$V_{DE}(h) \approx k_B T c_R (R_g/h)^2 \exp\{-3/2(h/R_g)^2\} \tag{10}$$

where c_R is the area density of repellers.

The disjoining pressure generated by membrane bending excitations is well known to play a key role for the adhesion of vesicles and erythrocytes. Recent studies show that they can also determine the adhesion of nucleated cells such as macrophages or endothelial cells lining the inner wall of blood vessels [61], suggesting a ubiquitous role of these entropic forces in cell biology. The local bending deformations $\delta h(x, y, t)$ decay in the tangential direction with persistence lengths $\zeta_p \sim 50$ nm and with correlation times $\tau \leqslant 0.1$ s. The mean square amplitudes are of the order of $\langle h(x, y)^2 \rangle \sim (k_B T/\kappa)\zeta_p^2$. Following an elegant image of Helfrich [57a], membranes can be considered to be composed of cushions of dimension $\zeta_p \times \zeta_p$ which exert Brownian motions in the normal direction. The collisions of the membrane with the wall induce a disjoining pressure of order $p_{disj} \sim k_B T/h\zeta_p^2$ as follows by assuming that the cushions behave as an ideal gas exchanging the mean energy $1/2\ k_B T$ during each collision. Since the maximum mean square amplitudes are of the order of the average interfacial distance, the disjoining pressure for tension free membranes becomes $p_{disj} = c_u\ (k_B T)^2/\kappa h^3$. Detailed renormalization group calculations and Monte Carlo simulations [57] yield a prefactor of $c_u \approx 0.115$. Note that for distances $h \leqslant 10$ nm the disjoining pressure is of the order of the Van der Waals forces. Note further that the undulations are impeded by membrane tensions and a more complex expression for p_{disj} is obtained [57]. Increasing the tension can thus induce first order transitions from free to bound states of vesicles and cells (cf [52a]). The bending excitations can be directly visualized by RICM. Figure 17 shows the adhesion of a vesicle doped with repeller molecules which is determined by van der Waals attraction. It is seen that the collisions of the membrane with the wall expel repeller molecules from the zone of tight contact and the membrane jumps randomly between a state of close contact (∼10 nm) and weak contact ($\langle h \rangle \sim 50$ nm). These collisions play a key role for the generation of nuclei of R–L pairs as shown in figure 17(b).

Consider now the consequences of the interplay between short range attraction forces and long range repulsion forces. Initially the shells hover over the surface at a distance $H \sim 100$ nm (determined by the balance of the undulations and gravity). Statistical collisions of the membranes induce the local formation of R–L pairs, which is associated with the local deformation of the membranes as indicated in figure 17(b). Due to the energy cost of the elastic deformation two statistically formed R–L pairs at a distance of order ζ_p attract each other and aggregate. It is obvious that the gain in deformation energy becomes

$$U_{ela} \approx cV''(H - h_0)^2 \zeta_p^2 \tag{11}$$

where c is a geometric factor. Monte Carlo simulations yield a value of $c = 8$ [58]. Note that the situation is very similar to the attraction of two proteins embedded in a bilayer if the hydrophobic thicknesses are not matched. Since the probability that two R–L pairs are formed is equal to the square of the molar fraction x_R, the total gain in internal energy by aggregation is $\Delta E_{ela} \approx z \cdot U_{ela} x_R^2$, where z is the coordination number of the lattice of R–L pairs. This gain in energy has to be paid for by loss of mixing entropy. If the gain in elastic energy by lateral aggregation of R–L pairs is larger than the loss in mixing entropy ($\Delta S = k_B T \ln x$) the cells will adhere by spontaneous decay of the adhesion zone into assemblies of R–L pairs separated by regions of very weak coupling. The situation is very similar to the phase separation processes in metal alloys or polymer solutions resulting in the formation of miscibility gaps. The transitions occur if the temperature is lowered below a critical value T_c, or at fixed T by

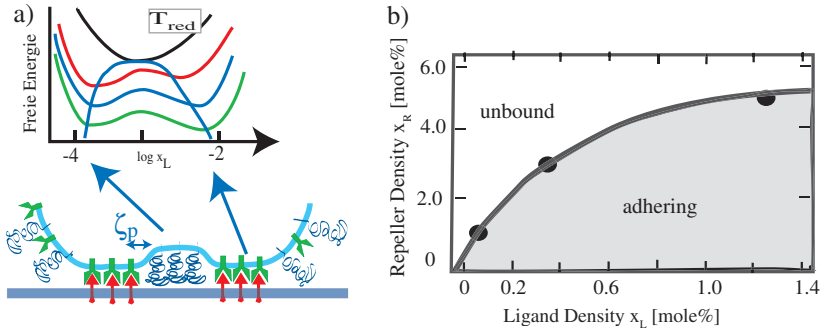

Figure 18. (a) Free energy of adhesion $\Delta F(x)$ as a function of the receptor concentration (expressed in terms of mole fraction) for various reduced temperatures $T_{\rm red} = k_{\rm B}T/zU_{\rm ela}$. (b) Example of phase diagram of wetting transition of vesicles containing RGD ligands adhering on supported membranes doped with integrin ($x_{\rm R} \sim 2 \times 10^{-3}$). The bilayer was deposited directly on glass, corresponding to immobilized receptors. The critical point is located at $x_{\rm L} = 0.01$ and $x_{\rm R} = 0.1$.

increasing the receptor (or ligand) concentration above a threshold $C_{\rm R}^*$ (cf figure 18(b)). Based on this analogy we can represent the situation in terms of the temperature dependence of the free adhesion energy $\Delta G_{\rm adh}$ by introducing a reduced temperature $T_{\rm red} \approx k_{\rm B}T/zU_{\rm ela}$ (cf figure 18). For $T_{\rm red} < T_{\rm red,c}\Delta G_{\rm adh}$ exhibits a double minimum and the adhesion zone decays into a phase of high (molar fraction x^+) and low (x^-) density of R–L pairs, corresponding to states of strong and weak adhesion, respectively. Since the states are separated by an activation barrier we deal with a first order transition.

A detailed calculation of the Helmholtz free energy based on the above considerations yields [52a]

$$F(x_{\rm R})/k_{\rm B}T \approx (1 - \Gamma/2)x_{\rm R}\ln x_{\rm R} + (1 - x_{\rm R})\ln(1 - x_{\rm R}) - \tfrac{1}{2}zU_{\rm ela}(x_{\rm R} - 1/2)^2. \qquad (12)$$

This equation resembles the free energy of mixed polymer solutions where the elastic energy $U_{\rm ela}$ plays the role of the Flory parameter. The factor Γ accounts for the entropy cost due to the suppression of long wavelength membrane bending undulations [58]. A similar reduction in mixing entropy would result if we assumed that the R–L pairs occupy an area that is g times larger than the area per lipid. The factor $(1 - \Gamma/2)$ would be replaced by g^{-1}. A more detailed analysis of equation (8) yields for the critical value of the reduced temperature $T^*_{\rm red,c} \approx 4/11$ [52]. The data obtained for the model system shown in figure 15 ($V'' \sim 10^8$ J m^{-4} [56], $(H - h_0) \sim 40$ nm, $\zeta_{\rm p} \sim 100$ nm) yield a value $T_{\rm red} \approx 0.1$; that is, at physiological temperatures we are well below the critical point.

From the point of view of cell biology the most important result is that the formation of adhesion domains sets in at very low receptor densities of the order $x_{\rm R} \geqslant 10^{-3}$. Therefore, cells need only some 10 000 of the costly receptors to form adhesion domains. The experiments show further that, provided the receptors are mobile, the adhesion domains can resist forces of several hundred piconewtons by force induced growth of the adhesion domains [56]. This is close to the forces cells transmit onto the surface during locomotion on surfaces driving cell crawling on substrates.

The domain structure is a metastable state. The large domains grow at the expense of the small ones, similar to the nucleation and growth processes of micro-crystallites in metal alloys undergoing phase separation. While metallurgists can control the microstructure of materials by freezing in of distinct microstructures, cells can regulate the size and distribution

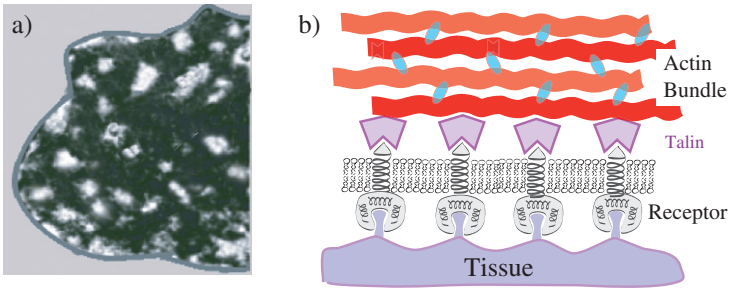

Figure 19. (a) Stabilization of adhesion domains by coupling of intracellular domains of integrin receptors to actin assemblies that can serve as stress fibres. A universal coupling factor is talin, that couples actin to intracellular domains of integrin receptors (after activation through phosphorylation). (b) Visualization of receptor clusters of adhering cells by fluorescence labelling of talin (modified after [86]).

of adhesion domains by coupling of the intracellular domains of the receptors to cross-linked actin assemblies as shown in figure 19 (forming focal complexes [54b, 62]).

5.3. Membrane elasticity and repeller osmotic pressure as regulators of adhesion strength

5.3.1. Membrane elasticity as control factor. The energy cost associated with the global elastic deformation of lipid bilayers is small compared to the gain in adhesion energy. However, membrane bending elasticity can control the adhesion strength through the elastic boundary conditions. The membrane curvature at the contact lines is typically 0.02 nm^{-1}. This corresponds to a bending energy density of $g_{ela} \sim \kappa/R_c^2 \sim 10^{-4}$ J m^{-2} (for $\kappa \sim 50\ k_BT$), which is one to two orders of magnitude larger than the typical values of the work of adhesion [52, 56]. Elasticity plays an even more dominant role for the control of adhesion strength of composite cell envelopes when the coupling of the plasma membrane to the actin cortex comes into play [54b]. A direct correlation between the membrane bending elasticity and the adhesion strengths was established by comparative studies of wild type cells of the slime mould *Dictyostelium discoideum* and mutants lacking the actin coupling protein talin, which is the dominant linker binding actin to the intracellular domain of cell surface receptors. The bending modulus of normal cells is $\kappa \sim 1000\ k_BT$. Knockout of talin reduces κ by a factor of 20 (to the value of the cholesterol rich plasma membrane), leading to a reduction of the work of adhesion W by nearly an order of magnitude [54b]. The control of the size and unbinding force of adhesion domains by coupling of the actin cortex to the intracellular domains of the clustered receptors (figure 19) plays a key role for tissue growth. Large supramolecular aggregates (called focal adhesion complexes) comprising several tens of types of proteins can form, which may serve as force sensors that monitor the stiffness of tissue or act as anchors of micro-muscles that drive the locomotion of cells [62].

5.3.2. Repeller osmotic pressure as control factor. One could naively think that the adhesion strength is mainly determined by the binding energy of the R–L pairs (that is by the term term $w\ c_R A_c$ in equation (11)). Quantitative measurements of adhesion strengths showed, however, that W is smaller by orders of magnitude than the bare receptor ligand binding energy. As noted first by Bell [63], this is a consequence of the osmotic pressure π_R exerted by the repeller molecules. For real situations of small receptor densities the work of adhesion is thus related to the binding energy ΔE by $W \approx \Delta E - k_BT \ln c_R$. This weakening of the adhesion strength by

the repeller molecules of the glycocalix is essential to prevent the collapse of soft shells such as vesicles or erythrocytes. Otherwise, vesicles would form spherical or ellipsoidal caps and explode, as happens if adhesion is for instance mediated by strong biotin streptavidin–biotin linkers.

5.3.3. A brief summary of the kinetics of adhesion. Very few experimental and theoretical studies of the kinetics of adhesion or de-adhesion have been done [64b]. The kinetics of de-adhesion is generally determined by the rate of unbinding of R–L bonds that depends on the loading rate of the applied force [64a]. As noted above, this is, however, fixed by the viscoelastic response time of the vesicle ($t \sim 0.1$ s, [52b]). Under certain conditions lag forces determined by the diffusion of CAMs can play a role [64a].

Two types of kinetic regimes of adhesion have been observed [64b]. At small concentrations of ligands the growing front $\xi(t)$ of the tight adhesion zones advances following a square root law: $\xi(t) \sim t^{1/2}$, while a cross-over to a linear regime $\xi(t) \sim t$ occurs at increasing ligand density. As shown experimentally and theoretically, the first law holds if the adhesion is diffusion limited while the linear law holds if the adhesion is dominated by the association rate of the R–L pairs.

5.4. Epilogue on adhesion

For many life processes the rapid formation and uncoupling of cell–cell contacts within timescales of seconds is essential. During crawling of cells on surfaces, for instance, cells have to constantly form new contacts at the leading front of pseudopods and unbind the local connections at the trailing end. The cadherin mediated connections between endothelium cells forming monolayers that line the inner wall of blood vessels have to unbind locally to enable the passage of white blood cells [66]. On the other side the adhesion strength must be strong enough to enable cells to transmit nanonewton forces to the substrate, that are required for the advancement of the cells. Adhesion through contact sites requiring few receptors is ideally suited to achieve this goal, since it opens many fast control mechanisms, some of which are depicted in figure 20. A fast mechanism is the adjustment of the receptor and repeller density in the plasma membrane by internalization (endocytosis) or export (exocytosis) of CAMs. A slower pathway (requiring several tens of minutes) is the *de novo* synthesis of receptor molecules or glycoproteins actin as repeller molecules. Other control mechanisms are based on the modification of cell and tissue surfaces such as the decomposition of repellers (e.g. the glycoprotein CD43) or extracellular matrix proteins (notably fibronectins and collagens) by proteases. The former mechanism enables lymphocytes or other white blood cells moving in the blood stream to dock onto endothelium cells and penetrate through the blood vessel into the tissue [66]. The regulation of adhesion by modification of the tissue surface plays a key role for the development of tissue (such as new blood vessels) but is also used by metastasis generating cancer cells to penetrate from the tissue into the blood stream. The cells secrete proteases that decompose the ligands on the surface of the wall of the blood vessel (the basal membrane), resulting in the generation of gaps within the endothelial cell layer by unbinding of cells [67]. Another mechanism is the unbinding of R–L bonds by antagonists competing with ligands for receptor binding domains [65].

In summary, the study of cell adhesion shows in an impressive way how cells use the laws of physics to explore their environment and adjust their material properties to the biological purpose. We hope that we can simultaneously show that it is possible to understand the function of complex living material on the basis of physics of soft materials. A major tool to achieve this goal is to establish more and more realistic model systems and to study systematically the physical properties of model wild type cells and mutants. In order to penetrate deeper into

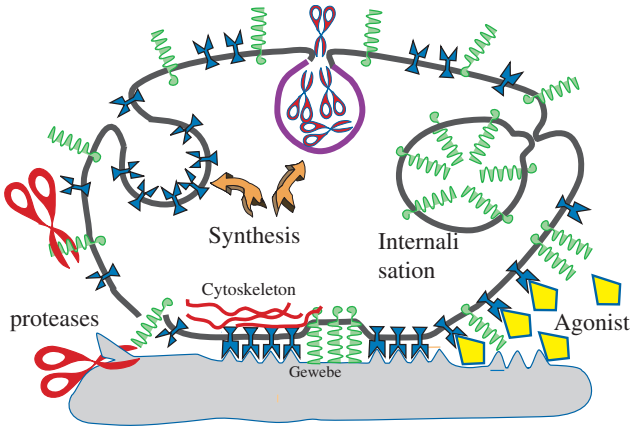

Figure 20. Summary of control of cell adhesion through variation of the density of receptors and CAMs acting as repeller molecules. Four regulation mechanisms are shown: (i) the increase of receptor concentration by exocytosis (top left), (ii) the reduction of the repulsive force by internalization of repeller molecules (top right), (iii) the unbinding of cell–tissue contacts by antagonists displacing ligands from receptors and (iv) die unbinding of cells from tissue by cleavage of ligands from tissue surfaces (e.g. basal membranes).

the secrets of adhesion it is necessary to generate more realistic mechanical models of cells. One important step is the reconstitution of actin cortices into giant vesicles. Recent studies in the author's laboratories have shown that this is feasible [68]. A second important task is the advancement of new micromechanical tools for high precision measurements of mechanical parameters of cell membranes and the development of the theory of elasticity of composite shells, leading fast to the non-linear theory of shells by Foeppl and Kapitza [55]. This is a necessary step towards the precise measurements of free adhesion energies and adhesion forces and their modification by mutations and diseases.

6. Adhesion induced functional domain formation and control of immune responses

Repeller molecules and receptors can exhibit enzymatic functions. In this case adhesion domains may play an additional role for the control of the efficiency of biochemical reactions by regulating the access of reactants or enzymes to confined reaction spaces. One example is the hormone amplification system discussed previously [2a]. Another example is the formation of immunological synapses discussed below. The immune response occurs in three steps (cf figure 21). First, the antigen presenting cells (APCs) are recognized by the lymphocyte through binding of T-cell receptors to antigens which are presented together with the major histocompatibility (MHC cf [69]) complex by the APC. Second, the adhesion leads to the formation of tight adhesion domains composed of receptor–ligand pairs (cf figure 14), which is associated with the recruitment of an effector protein (called co-receptor CD3) to the intracellular domain of the TCR. Third, the effector (CD3) is activated through coupling of phosphate groups to its intracellular domain by kinases, which trigger a sequence of events, resulting finally in the cell proliferation and clone formation [1]. In the resting state the activity of the effector (CD3) is constantly suppressed by rapid removal of the phosphate groups by an antagonistic phosphatase such as CD45. It exhibits an extracellular domain which is much longer than the TCR-AG complex (cf [51]). Therefore, this suppressor is expelled from the adhesion domain after contact formation and the CD3 protein is effectively activated by

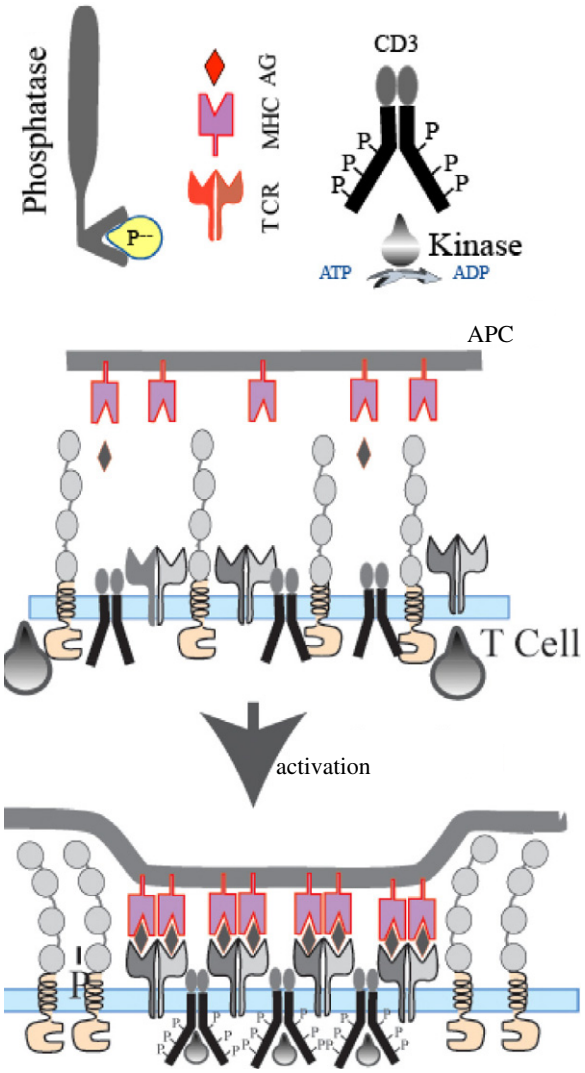

Figure 21. Simplified model of the activation of lymphocytes (T-cells) after adhesion on antigen presenting cells. Initially the co-receptor (CD3) activated by phosphorylation through kinases is deactivated again by the phosphatase CD 45. After formation of the tight adhesion domain the phosphatase is expelled from the domain and CD3 is effectively phosphorylated.

phosphorylation. Evidence for this model was provided in beautiful experiments by Choudhuri *et al* [70]. These authors changed the lengths of the extracellular domains of both the CD45 and the T-cell receptor and showed that the immune response is suppressed if the length of the CD45 is comparable to or shorter than that of the TCR-AG complex.

Acknowledgments

The present review is based on a series of lectures on the physics of cell membranes given at the Departments of Physics and Chemical Physics at the UCLA during a three month visit as

Regents Professor. The author gratefully acknowledges the hospitality of the colleagues and their helpful comments.

Appendix A. Elasticity of soft composite shells

The elastic deformation of biomembranes can be described by the three classical modes of deformation of shells: isotropic extensional deformations, shearing and bending, which are characterized by the three two dimensional (2D) elastic moduli K, μ and κ, respectively. Fluid membranes are shear free (unless fast processes are considered [71]). For comparison of material properties it is convenient to relate the two dimensional moduli of membranes to the elastic parameters of the bulk material: the Young modulus E and the Poisson ratio ν [55],

$$\kappa = \frac{Eh^3}{12(1-\nu^2)}; \qquad \mu = \frac{Eh}{2(1+\nu)}; \qquad K = \frac{Eh}{3(1-2\nu)} \tag{A.1}$$

K and κ are related by the empirical rule $K \approx \kappa/h^2$. To get an idea of the softness of fluid membranes it is helpful to compare some numbers. The Young modulus of lipid bilayers is (for $\kappa = 50\, k_B T$, $\nu = 1/2$) $E = 10^7$ Pa, of rubber $E \sim 10^9$ Pa and of gold $E \sim 3 \times 10^{11}$ Pa. A gold foil of $h \sim 0.2$ nm would have the same bending stiffness as a fluid bilayer.

The topology of shells can be characterized by two types of curvature, the mean curvature $H = 1/2(R_1^{-1} + R_2^{-1})$ and the Gaussian curvature $G = R_1^{-1} R_2^{-1}$, where R_i are the principle radii of curvature. If the contour of the membrane is defined as the z-coordinate of the neutral plane (called the Monge representation) one can express the curvatures by

$$R_1^{-1} = \partial^2 u/\partial x^2 \qquad \text{and} \qquad R_2^{-1} = \partial^2 u/\partial y^2.$$

Three models have been proposed to describe the curvature elasticity of fluid membranes. The Helfrich model is based on the concept of the splay elasticity of smectic phases, with the lipid bilayer being considered as a film of rod-like molecules. It accounts for the intrinsic asymmetry of natural membranes by introduction of a spontaneous curvature C_0. Thus the bending energy of a closed vesicle is expressed as [26]

$$\Delta G_{He} = \tfrac{1}{2}\kappa \int\int \mathrm{d}O(\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 - C_0)^2 + \tfrac{1}{2}\kappa_G \int\int \mathrm{d}O G^2 \tag{A.2}$$

κ_G is the Gaussian bending modulus, which has not been directly measured yet. Studies of membrane fusion suggest $\kappa_G \sim 0.3\kappa$ [72]. The Gaussian contribution can be ignored for fluid membranes of spherical topology but not for more complex topologies such as toroidal vesicles or intracellular compartments such as the ER (cf [32, 38] for details). It is often helpful to describe bending deformations in the force representations by assuming that the bending is induced by a pair of opposite surface tensions applied at the two surfaces (or more precisely a tension gradient $\mathrm{d}\pi/\mathrm{d}z$ in the normal direction). The induced curvature is then

$$C_0 = \frac{M}{\kappa} = \frac{1}{\kappa d} \int z \frac{\partial \pi}{\partial z}\, \mathrm{d}z. \tag{A.3}$$

For closed membranes the Helfrich energy functional has to be extended to account for the fact that two leaflets of the bilayer can mutually slide in the tangential direction and thus undergo different extensional deformations (cf [31]). Thus a local concave deformation of the stratified shell results in a global extension of the inner and a compression of the outer layer. This effect is considered in the so called bilayer coupling (BC) model [30, 32]. The area difference of the two leaflets can be easily related to the mean curvature as $\Delta A = d \oint \int 2H\, \mathrm{d}O$. Moreover, one has to account for the fact that ΔA depends on the initial area of the two leaflets. If a spherical vesicle of radius R is for instance generated from a flat piece, the two leaflets exhibit the same

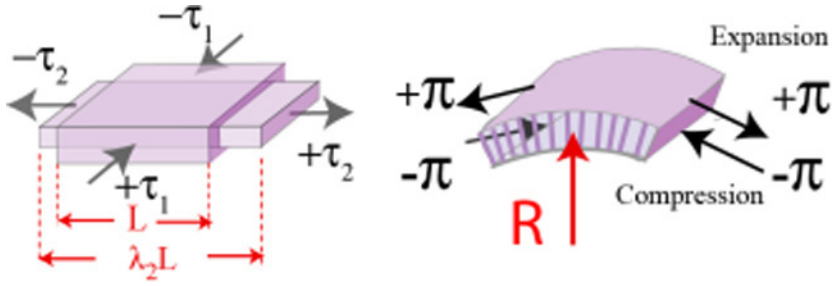

Figure A.1. Definition of deformation modes of shells. Left, shearing and lateral compression; right, bending. The lateral tensions τ and π are expressed as force/length. The strain of the bent shell is defined by the radius of curvature R. Note: it is not difficult to show that the above definition of shear deformation is equivalent to the familiar definition by considering the deformation of a square which is embedded in the original one but rotated by 45°.

number of lipids, while in the sphere there exists an area difference $\Delta A_0 = 8\pi R$. The bending energy is thus

$$\Delta G_{\mathrm{BC}} = \frac{\pi \kappa^*}{2Ad^2}(\Delta A - \Delta A_0)^2, \tag{A.4}$$

where d is the distance between the neutral planes of the two monolayers. κ^* (called the global bending modulus) is an effective bending modulus accounting for the coupling between the leaflets. It can be related to the area compressibility modulus of the two leaflets [31] and is about equal to κ. The total bending energy of a bilayer can thus be expressed as $\Delta G_{\mathrm{ela}} = \Delta G_{\mathrm{He}} + \alpha \Delta G_{\mathrm{BC}}$ (cf equation (4)). This approach, denoted as the area difference elasticity (=ADE) model, interpolates between both contributions to the deformation energy, where α accounts for the relative contributions. Note, however, that the above models fail for solid shells, strong deformations of membranes or closed spherical shells with small excess surface area. Here the Gaussian curvature is not negligible any more and one has to apply the rigorous Foeppel–Kapitza model of the elasticity of shells [42, 55].

For composite shells, such as cell envelopes, the elastic free energy has to be extended to account for the membrane associated cytoskeleton. This is only relatively simple for erythrocytes since the spectrin/actin network is isotropic and can be treated as a sub-shell of soft rubber without bending stiffness. For cells exhibiting actin cortices with bending stiffnesses of 1000 $k_{\mathrm{B}}T$ or larger [54], the rigorous shell theory has to be applied. Following the concept of rubber elasticity, the strain is traditionally represented by relative change of length λ_1 and λ_2 (defined in figure A.1). The planar deformation is generated by two pairs of tensions (dimension N m^{-1}) acting parallel to the neutral plane which deform a quadratic piece of membrane of area $A_0 = L \times L$ into rectangle of dimensions $\lambda_1 L$ and $\lambda_2 L$ (area $A = \lambda_1 \lambda_2 L^2$). If the orthogonal pairs of stresses τ_1 and τ_2 have equal amplitudes but opposite signs, $\lambda_1 = \lambda_2^{-1}$. The area is constant and one has pure shear. If the extensions differ ($\lambda_1 \neq \lambda_2$) the area changes and the strain (the relative area change) becomes $\Delta A/A_0 = \lambda_1 \lambda_2 - 1$. The total elastic energy cost of a network undergoing both shearing and extensional deformations is thus (cf [36])

$$\Delta G_{\mathrm{Cyt}} = \frac{1}{2}K \int\int \mathrm{d}O(1 - \lambda_1\lambda_2)^2 + \frac{1}{2}\mu \int\int \mathrm{d}O \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} - 2\right). \tag{A.5}$$

It should be noted that this classical model (by Mooney) of rubber elasticity is an intrinsically non-linear theory since the extensional ratios are quadratic in the strain tensor elements. A more general topological model of composite closed membranes that accounts for all modes of deformation in a general way was developed by Peterson [73].

Appendix B. Vesicle mediated iron import of cells

An important example of material trafficking is the import of Fe^{3+} ions (bound to the carrier transferrin) by endocytosis via clathrin coated pits as described in figure B.1. The uptake and release of the carrier and the ion are controlled in a smart way by the pH dependence of the binding constants. At neutral pH the binding constant of transferrin to the receptor is very small in the iron free state (= apotransferrin), while it is large if the ions are bound. The apotransferrin is thus replaced by Fe-transferrin after the recycling step when the receptors become exposed to the neutral extracellular fluid. In contrast, the binding constant of Fe to transferrin is high at neutral pH ($\ldots_D = 6 \times 10^{-9}$ M) but very (immeasurably) low at acidic pH. Therefore, the ion dissociates completely after fusion of the vesicle with the acidic endosome (also called the early endosome [1]). The cycle time of receptors is about 15 min and a cell can take up about 10^4 Fe^{3+} ions min^{-1}. The process can simultaneously serve the recycling of SPM and cholesterol to the PM, which may be due to a strong tendency of buds that are composed of natural SPM or are enriched in cholesterol to pinch off as demonstrated in [12].

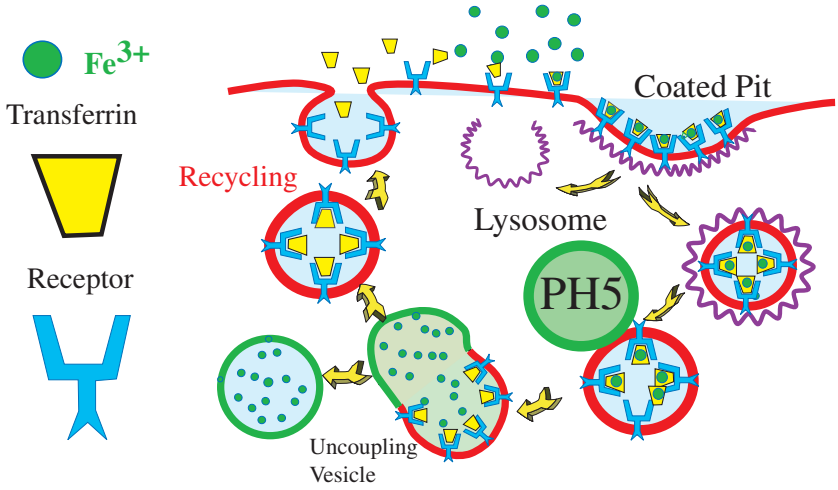

Figure B.1. Model of receptor mediated import of iron and membrane receptor recycling. Iron captured by transferrin binds to transferrin receptors which assemble in coated pits. After their detachment from the PM the clathrin coat is split off. This occurs by hydrolysis of the GTP coupled to the activator, resulting in the folding back of the hydrocarbon chain (cf figure 1(c)). The naked vesicle fuses with acidic endosomes (sometimes called early endosomes), resulting in the dissociation of the ion. The hybrid vesicle decays again into a vesicle enriched in transferrin receptors loaded with the carrier and vesicles filled with Fe^{3+}. The former fuses with the PM (thus recycling PM material) and the latter with lysosomes [87].

Appendix C. Tension induced switching of membrane fusion

Similar to budding and fission, fusion is a biological membrane instability that is controlled by a specific set of proteins, although pure lipid bilayers could also fulfil this task [74]. This suggests that the process is governed by the same general physical principle *in vivo* and *in vitro*. In both cases fusion appears to occur in three steps: (i) contact formation between the membranes that may be associated with dehydration of the interface; (ii) the merging of the outer monolayers (called hemi-fusion) and (iii) the breakage of the bilayer diaphragm (cf [74, 75] for references). The fusion of lipid membranes can be induced in multilayers by dehydration [74] and in vesicle

suspensions by depletion forces generated by adding polymers [75], or, in the presence of charged lipids, by Ca^{2+} or other multivalent ions. It is also well established that besides close membrane contacts a certain value of the membrane tension (that can be mediated by osmotic pressure) must be generated to catalyse the opening of the diaphragm between the membranes. At small tensions the opening can be accelerated by fusion promoters (such as surfactants) reducing the activation energy for pore formation [74].

In cells, fusion is mediated by specific sets of fusion proteins. The best studied example is the fusion of synaptic vesicles with the pre-synaptic membrane (denoted as plasma membrane (PM) as usual). It is mediated by a pair of membrane proteins, one of which (syntaxin) is bound to the PM (called the target membrane), and the other (synaptobrevin) is coupled to the vesicle. They are referred to as t-SNAREs and v-SNAREs (where the prefixes stand for target membrane and vesicle, respectively) [76]. Each protein is composed of a long rod-like extracellular fraction and a membrane spanning domain. The fusion is triggered by association of the extracellular domains of the two SNAREs with two associated rod-like cytoplasmic proteins, called SNAP, that are interconnected by a long flexible peptide loop. The associated SNAP components are bound to the membrane by one or two hydrocarbon chains. The four rod-like extracellular domains associate into an intertwined fourfold superhelix by forming typical coiled-coil structures, thus forming a cylindrical rod of about 12 nm length and 4 nm diameter. The coiled-coil structures form by association of periodic sequences of hydrophobic peptides in such a way that the inside of the rod is hydrophobic while the surface of the super-coil exhibits a negative surface charge [76]. Under *in vitro* conditions fusion of vesicles can be mediated by the SNARE–SNAP system without any additional protein [77]. Under physiological conditions the fusion of the synaptic vesicles with the pre-synaptic PM is accelerated by a specific ATPase, which induces the decay of the SNARE-SNAP superhelix. It is assumed to provide an additional driving force for fusion (cf [1] chapter 22). Most importantly, it also accelerates the recycling of the Ca^{2+}-filled vesicles which is essential for rapidly firing neurons.

The SNAREs have to fulfil two tasks. They must bring the membrane into close contact and generate the driving force for the fusion process. An important experimental finding is that complete fusion requires the trans-membrane peptide domains of both SNAREs. The replacement of this domain in only one SNARE component by a fatty acid chain impedes fusion [77]. Based on these findings, we can interpret the SNARE mediated fusion as a tension driven membrane instability following models of the osmotic tension driven fusion of lipid membranes [75a] and pore formation in lipid bilayer vesicles [75b]. The basic idea (illustrated in figure C.1) is that the driving force for the opening of passages between coupled membranes is provided by the free energy gained by adsorption of the extracellular SNARE–SNAP domains to the membranes. This creates a tension gradient between the region where the membranes meet and the membrane SNARE adhesion zone of

$$\delta\sigma = \sigma_0 - W,$$

where σ_0 is the membrane tension and W is the membrane–SNARE binding energy per unit area. The fusion of the membranes in the diaphragm can be described in terms of the formation of holes in bilayers. The free energy of this process can be expressed as [75]

$$f_p = f_0 + 2\pi\Gamma r - \pi r^2\sigma_{\text{eff}} \tag{C.1}$$

where r is the radius of the hole and σ_{eff} is the effective membrane tension. Γ is the line tension (the energy cost per unit length to generate a hole through the hydrophobic film). The second term accounts for the gain in energy due to the relaxation of the membrane tension by the growth of the opening. In the case of the hemi-fusion an additional energy gain term f_{ad} has to be added to account for the gain in energy by coupling of the two monolayers. It is easily verified that, according to equation (C.1), the energy cost first increases, reaches a maximum

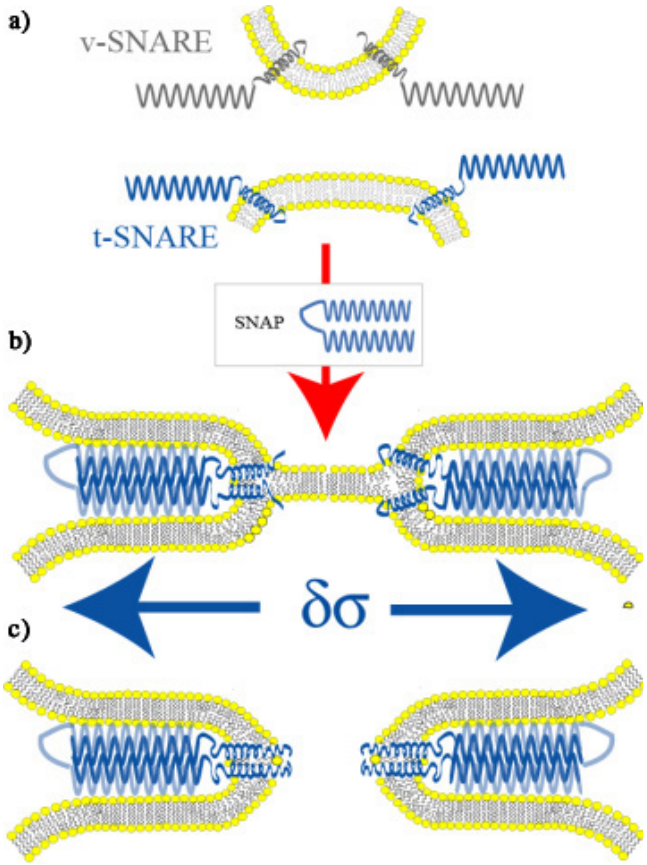

Figure C.1. Model of fusion of vesicles with plasma membrane mediated by lateral tension, $\delta\sigma$, generated by adsorption of SNARE–SNAP-complexes to membranes. Cross sections through the centre of the stalk are shown. The fusion is initiated by the aggregation of SNAREs through SNAP and is assumed to proceed via a hemi-fusion step (b). Note that in the case of hemi-fusion two possible configurations of the lipid–protein assembly at the edge of the diaphragm are shown on left and right sides. (c) The retraction of the bilayer during the second step is promoted by the hydrophilic end of the trans-membrane domains of the SNAREs. Similar to surfactants these can generate high curvatures at the edge of the pore. Note that these domains can be stretched considerably by an α-helix to β-sheet transition of the trans-membrane domain [77].

at the radius r^* and decreases again. The opening of the pore proceeds spontaneously if a critical pore of radius r^* is formed by thermal fluctuations. The situation is thus very similar to nucleation and growth processes in mixtures. The critical radius r^* and the activation energy f_{act} is obtained by minimizing f_p to

$$r^* = \frac{\Gamma}{\delta\sigma + f_{\mathrm{ad}}^*} \qquad f_{\mathrm{act}} = \frac{\Gamma^2}{\delta\sigma + f_{\mathrm{ad}}}. \tag{C.2}$$

These equations suggest that at a given tension gradient the two fusion steps can be promoted by different mechanisms. Hemi-fusion is promoted by the coupling of the monolayers, while the breakage of the bilayer diaphragm is favoured by the reduction of the line tension by the bilayer spanning trans-membrane domains of the SNAREs (cf figure C.1(c)).

Consider now the energies involved. The line energy for the hemi-fusion can be estimated from the data of [78] (figure 4) to $\Gamma \sim 10^{-10}$ J m^{-1} (300 $k_{\mathrm{B}}T$ for a pore radius of $r = 10$ nm).

The line tension of pores in bilayers was measured by model membrane studies to be $\Gamma \sim 10^{-11}$ J m^{-1} [75]. The gain in adhesion energy for hemi-fusion can be estimated from the hydrophobic chemical potential μ_L of lipids in bilayers (40 kJ Mole^{-1}) and the surface of the hydrocarbon chains ($A_L \sim 10^{-17}$ m^2 for PC:16,0) to be $f_{ad} \sim 7 \times 10^3$ J m^{-2}. Since no direct measurements of the adsorption energy of proteins on lipid membranes have been made yet, we estimate the tension difference $\delta\sigma$ (or the energy W) required to reduce the activation energy to be about 10 k_BT (4×10^{-20} J). For the case of hemi-fusion we find $W \sim 5 \times 10^{-2}$ J m^{-2} (or 5 k_BT/lipid) which is larger by a factor of three than the gain in hydrophobic interfacial energy. For the case of bilayer disruption $W \approx 7 \times 10^{-3}$ J m^{-2} ($\sim$1 k_BT/lipid). The corresponding critical coil radii are in both cases $r^* \sim 1$ nm.

In summary, the energy gain by adsorption of the extracellular SNARE domain to the membranes should be $W \geqslant 10^{-2}$ J m^{-2}. Unfortunately, no direct measurements of adsorption energies of proteins could be found. Measurement of the Gibbs free energy of protein adsorption to lipid monolayers suggests a value of $\sim 10^{-3}$ J m^{-2} ($\sim$0.5 k_BT per lipid, cf [8]). However, the above estimates of $\delta\sigma$ are upper limits. In the case of hemi-fusion the line tension can be reduced by the enrichment of long chain lipids at the rim of the hemi-fused diaphragm, and in case of the bilayer breakage by the membrane spanning domain of the SNAREs As indicated in figure C.1(c), the small extracellular domains of the SNAREs can generate a sharp edge. Model membrane studies suggest that this could reduce Γ by an order of magnitude. In this context, it is interesting to note that the trans-membrane domain of SNAREs can easily switch between an α-helix and β-sheet conformation. It can stretch by a factor of two and act like a spring, generating a pulling force on the bilayer. Evidence for such a role is provided by the finding that the fusion is impeded if this plastic behaviour is abolished [77b].

One reason for the above mentioned acceleration of synaptic vesicle fusion by ATPases may be that dynamics of fusion without such a promoter would be too slow. Since the ATPase dissociates the superhelix, their hydrophobic domains become exposed and the individual strands of the SNARE components would penetrate into the outer leaflets as an amphiphatic helix, yielding $W \sim f_{ad}$. It should finally be noted that the fusion process is very fast. Estimates of the response time τ_r based on the assumption that fusion is driven by the Marangoni effect predict $\tau_r \sim 0.1$ ms [75]. Thus, a transient increase of W by the dissociation of the SNARE complex is expected to last sufficiently long to drive the pore opening.

References

[1] Pollard T O and Earnshaw W C 2002 *Cell Biology* (Philadelphia, PA: Saunders)
[2a] Sackmann E 1995 Biological membranes architecture and function *Handbook of Biological Physics* vol 1A, ed R Lipowsky and E Sackmann (Amsterdam: Elsevier) chapter I
[2b] Sackmann E 1995 Physics of vesicles *Handbook of Biological Physics* vol 1A, ed R Lipowsky and E Sackmann (Amsterdam: Elsevier) chapter V
[3] Gov N and Safran S 2005 *Biophys. J.* **85** 1859–74
[4] Boal D 2002 *Mechanics of the Cell* (Cambridge: Cambridge University Press)
[5] Sackmann E, Bausch A and Vonna V 2002 Physics of composite cell membrane and actin based cytoskeleton *Nato Advanced Study Institute, Les Houches Session LXXV* ed H Flyvberg *et al* (Berlin: Springer)
[6] Lippincott-Schwartz J 1998 *Curr. Opin. Cell Biol.* **10** 52–9
[7] Bechinger B 2001 *Biophys. J.* **81** 2251–6
[8] Nauman Ch *et al* 1996 *Biophys. J.* **71** 811–23
[9] Koenig S *et al* 1992 *J. Physique* II **2** 1589–615
[10] Simons K and Ehehalt R 2002 *Clin Invest.* **110** 597–603
[11] Brzustowicz M *et al* 2002 *Biophys. J.* **82** 285–9
[12] Doebereiner H-G, Kaes J and Sackmann E 1991 *Biophys. J.* **60** 825–44
[13] Vist R and Davis J H 1990 *Biochemistry* **29** 451–64

[14] McMullen T and McElhaney R N 1995 *Biochim. Biophys. Acta* **1234** 90–8
Mabrey S, Mateo P and Sturtevant J 1978 *Biochemistry* **17** 2464–8
[15] Gliss Ch *et al* 1999 *Biophys. J.* **77** 331–40
[16] Klein M L and Tobias D 1998 *Biophys. J.* **75** 2147–56
[17] Reindl H, Brumm Th and Bayerl Th 1992 *Biophys. J.* **61** 1025–35
[18] Elliott R, Szleifer I and Schick M 2006 *Phys. Rev. Lett.* **96** 098181
[19] Keller S and McConnell H M 1999 *Phys. Rev. Lett.* **82** 1605–11
[20] Gambhir A *et al* 2004 *Biophys. J.* **86** 2188–207
[21] Bloom M, Evans E and Mouritsen O 1991 *Q. Rev. Biophys.* **24** 293–397
[22] Ben-Shaul A 1995 Molecular theory of chain packing elasticity and lipid protein interaction in lipid bilayers *Handbook of Biological Physics* vol 1A, ed R Lipowsky and E Sackmann (Amsterdam: Elsevier)
[23] May S and Ben-Shaul A 2000 *ChemPhysChem* **2** 4494–9
[24] Sintes T and Baumgaertner A 1997 *Biophys. J.* **73** 2251–9
[25] Wang D N 1994 *FEBS Lett.* **346** 26–31
[26] Anderson R and Jacobson K 2002 *Science* **296** 1821–30
[27] Bretcher M and Munro S 1993 *Science* **26** 1280–1
[28] Schuetz G, Kada V, Pastushenko H and Schindler J 2000 *EMBO J.* **19** 892–901
Tomishige M and Kusumi A 1999 *Mol. Biol. Cell* **10** 2475–9
[29] Helfrich W 1973 *Z. Naturf.* c **28** 693–703
Helfrich W and Servuss R M 1984 *Nuovo Cimento* **3D** 137–51
[30] Svetina S and Zeks B 1989 *Eur. Biophys. J.* **17** 101–10
[31] Evans E A 1974 *Biophys. J.* **14** 923–31
[32] Seifert U and Lipowsky R 1995 Morphology of vesicles *Handbook of Biological Physics* vol 1A, ed R Lipowsky and E Sackmann (Amsterdam: Elsevier)
[33] Kaes J and Sackmann E 2001 *Biophys. J.* **81** 2251–6
[34] Mutz M and Bensimon D 1991 *Phys. Rev.* A **43** 4525–8
[35] Doebereiner H-G *et al* 1997 *Phys. Rev.* E **49** 4458–74
[36] Mukhopadhyay R, Lim G and Wortis M 2002 *Biophys. J.* **82** 1756–77
[37] Yawata A *et al* 1997 *Blood* **90** 2471–90
[38] Kawakatsu T *et al* 1993 *J. Physique* II **3** 971–80
Gozdz W and Gompper G 1999 *Phys. Rev.* E **59** 4305–10
[39a] Noguchi H and Gompper G 2005 *Proc. Natl Acad. Sci.* **102** 14159–64
[39b] Lee J and Discher D 2001 *Biophys. J.* **81** 3178–92
[40] Gao H, Shi W and Freund L B 1995 *Proc. Natl Acad. Sci.* **102** 9469–73
[41] Ford G *et al* 2002 *Nature* **419** 361–6
[42] Nelson D and Peliti L 1987 *J. Physique* **48** 1085–92
[43] Mashl R and Bruisnma R 1998 *Biophys. J.* **74** 2862–72
[44] Kohyama T, Kroll D and Gompper G 2006 *Phys. Rev.* E at press
[45] Kozlov M 1999 *Biophys. J.* **77** 604–16
[46] Schmidt A *et al* 1999 *Nature* **401** 133–41
[47] Galbiati F *et al* 1999 *J. Cell Biol.* **274** 5843–50
[48] Schlegel A and Lisanti M 2000 *J. Biol. Chem.* **275** 21605–17
[49] Mundy D *et al* 2000 *J. Cell Sci.* **115** 4327–9
[50] Sens P and Turner M 2004 *Biophys. J.* **86** 2049–57
[51] Singer Th 1990 *Nature* **346** 425–43
[52a] Sackmann E and Bruinsma R 2002 *ChemPhysChem* **3** 262–9
[52b] Bruinsma R, Behrisch A and Sackmann E 2000 *Phys. Rev.* E **61** 4253–61
[52c] Guttenberg Z *et al* 2000 *Langmuir* **16** 8984–93
[53] Seifert U and Lipowsky R 1991 *Phys. Rev.* A **42** 4768–74
Smith A and Seifert U 2005 *Phys. Rev.* E **71** 061902–10
[54a] Raedler J *et al* 1995 *Phys. Rev.* E **51** 4526–36
[54b] Simson R *et al* 1998 *Biophys. J.* **74** 514–22
[55] Landau L D and Lifschitz E M 1970 *Theory of Elasticity* 2nd edn (Oxford: Pergamon)
[56] Sackmann E and Goennenwein S 2006 *Prog. Theor. Phys.* at press
[57a] Helfrich W 1978 *Z. Naturf.* a **33** 305–15
[57b] Lipowsky R 1995 Generic interaction of flexible membranes *Handbook of Biological Physics* vol I, ed R Lipowsky and E Sackmann (Amsterdam: Elseveier)
[57c] Netz R 1995 *Phys. Rev.* E **51** 2286–96

[58] Bruinsma R, Goulian M and Pincus P 1994 *Biophys. J.* **67** 746–50
[59] Swain P and Andelmann D 1999 *Langmuir* **15** 8902–12
[60] Dolan A and Edwards S F 1974 *Proc. R. Soc.* A **337** 509–14
[61] Zidovska A and Sackmann E 2006 *Phys. Rev. Lett.* **96** 048103–7
[62] Geiger B and Bershadsky A 2001 *Curr. Opin. Cell Biol.* **13** 584–92
[63] Bell G, Dembo M and Bongrand P 1984 *Biophys. J.* **45** 1051–60
[64a] Brochard-Wayart F and de Gennes P 2002 *Proc. Natl Acad. Sci.* **99** 7855–60
[64b] Boulbitch A, Guttenberg Z and Sackmann E 2001 *Biophys. J.* **81** 2743–52
[65] Smith A, Lorz B, Seifert U and Sackmann E 2006 *Biophys. J.* **90** 1064–80
[66] Shelley C, Da Silva N and Teodoridis J 2001 *Br. J. Haematol.* **115** 159–65
Cinamon G, Shinder V and Alon R 200 *Nat. Immunol.* **2** 515–22
[67] Stetler-Stevenson W, Aznavoorian S and Liotta L 1993 *Annu. Rev. Cell Biol.* **9** 541
[68] Limozin L, Roth A and Sackmann E E 2002 *Phys. Rev. Lett.* **89** 168103–7
[69] Watts Th, Gaub H and Mc Connell H M 1986 *Nature* **320** 179–81
[70] Choudhuri K *et al* 2005 *Nature* **436** 578–82
[71] Hildenbrand M and Bayerl Th 2005 *Biophys. J.* **88** 3360–7
[72] Kozlovsky Y, Efrat E, Siegel D and Kozlov M 2004 *Biophys. J.* **87** 2508–21
[73] Peterson M 1992 *Phys. Rev.* A **45** 4116–31
[74] Chermomordik L *et al* 1995 *Biophys. J.* **69** 922–9
Huang H, Chen F and Lee M-T 2004 *Phys. Rev. Lett.* **92** 198304
[75a] Safran S, Kuhl T and Israelachvili J 2001 *Biophys. J.* **81** 659–65
[75b] Karatekin E *et al* 2003 *Biophys. J.* **84** 1734–49
[76] Sutton R *et al* 1996 *Nature* **395** 347–53
[77a] McNew J *et al* 2000 *J. Cell Biol.* **150** 105–17
[77b] Langosch D *et al* 2001 *J. Mol. Biol.* **311** 709–21
[78] Koszlovsky Y, Chernomordik L and Kozlov M 2002 *Biophys. J.* **83** 2634–51
[79] Perry M *et al* 1979 *J. Cell Sci.* **39** 257–72
[80] Rothberg K *et al* 1992 *Cell* **68** 673–82
[81] Sackmann E 1994 *FEBS Lett.* **346** 3–16
[82] Zagury D *et al* 1975 *Eur. J. Immunol.* **5** 818–22
[83] Dustin M L *et al* 2001 *Trends Immunol.* **22** 192–6
[84] Goennenwein S *et al* 2003 *Biophys. J.* **85** 646–55
[85] Schilling J *et al* 2002 *Phys. Rev. Lett.* **88** 138102
[86] Tsujioka M *et al* 2004 *EMBO J.* **23** 2216–25
[87] Rothman J F and Schmid S L 1980 *Cell* **16** 5–9