INTRODUCTION
Entamoeba histolytica: the parasite

Protozoa have become adapted to practically all types of habitat on the face of the earth, including the human body. Though fewer than 20 genera of protozoa cause disease in humans, their impact is formidable. For example, about 500 million people are infected with Entamoeba histolytica and as many as 40,000 die directly or indirectly due to E. histolytica infection each year, making it a major cause of parasitic death worldwide. The parasitic protozoa E. histolytica is the causative agent of human invasive amebiasis, a common health problem in some developing countries, characterized by potentially fatal intestinal and liver lesions. The motile form of the parasite, the trophozoite, usually lives as a harmless commensal in the lumen of the large intestine, where it multiplies and differentiates into cysts, the resistant form, responsible for the transmission of the infection. As a commensal, E. histolytica induces no signs or symptoms in this condition. Normally, it causes the disease condition in about 10% of the infected individuals; virulent amebas invade the intestinal mucosa and produce dysentery or ameboma, and occasionally through blood it can spread giving rise to extra intestinal lesions, mainly liver abscess. Most cases of intestinal amebiasis manifest as diarrhoea or dysentery and have a self-limited course. Amebic liver abscess, however, is potentially fatal, unless promptly diagnosed and properly treated. Entamoeba dispar, an ameba morphologically similar to E. histolytica that also colonises the human gut, displays no invasive potential and is the underlying cause of most of the asymptomatic ameba infections reported worldwide.

Life Cycle and Infection:
The life cycle of E. histolytica consists of various consecutive stages, namely, trophozoite, cyst and metacyst (Figure 1). Trophozoites (average diameter 10-60 um) dwell in the colon, where they multiply by binary fission and encyst, producing typical four-nucleated cysts after two successive nuclear divisions. The dynamic activity and pleomorphism of the trophozoites are based on a simple
Figure 1: Human infection by *Entamoeba histolytica*. Infection begins with ingestion of cysts via contaminated food or water. Cysts undergo excystation (2) in the intestine giving rise to trophozoites (3). Trophozoites migrate to the large intestine by binary fission (4) and produce cysts (1), which are passed in the feces.
cytoplasmic layout which lacks a number of organelles usually found universally; for example, there is no structured cytoskeleton nor cytoplasmic microtubules; no defined Golgi complex nor endoplasmic reticulum, no mitochondria nor any system of primary and secondary lysosomes (Martinez-Palomo, 1987). Cysts are found in the formed stools of carriers as round or slightly oval hyaline bodies, 8-20 um in diameter, with a rigid wall that protects the amebas outside the human gastrointestinal tract.

Infection of *E. histolytica* is ubiquitous, but the highest endemic incidence is usually found in communities with inadequate sanitation. The main reservoir of *E. histolytica* is man, although morphologically similar amebas may be found in primates, dogs and cats. Human susceptibility to infection appears to be general, but most individuals harbouring the parasite do not develop disease. Some host properties that favour invasive amebiasis are intestinal microflora, certain nutritional habits, deficient immune response, alcoholism and pregnancy (Walsh, 1988).

**The organism:**

The successful cultivation of axenic *E. histolytica* by Diamond (1968) has been a major contribution to research in amebiasis, and prompted much of subsequent research on the biology of the parasite. *E. histolytica* appears to one of the most primitive eukaryotes. Scanning electron microscopy reveals the extreme pleomorphism of the trophozoites of pathogenic *E. histolytica*: lobopodia, endocytic stomata, filopodia and uroids may be found on the surface (Gonzalez-Robles et al., 1983). Much of the cytoplasm consists of vacuolar system; the vacuoles (diameters of 0.5-9 µm) are mainly a result of endocytosis, but other functions remain to be characterized. The cytoplasm of trophozoites also contains a lattice of tubules and vesicles superficially resembling smooth endoplasmic reticulum. The lattice is made up of extremely thin tubules of diameter ~ 20nm forming irregular whorls or parallel arrays. Ribosomes occur
singly, but are often ordered in helical arrays - these may aggregate in large crystalline inclusions several microns long, constituting the "chromatoid body" often seen in cysts. Whether ribosomes in helices and in chromatoid bodies are functionally mature ribosomes or ribosomal precursors remains to be seen. The cytoplasm contains a large amount of actin that usually is not polymerized as microfilaments (Meza et al., 1983). Microtubules have been characterized in the nuclei of dividing trophozoites (Martinez-Palomo, 1992). The study of the cytoskeleton of the ameba is of interest in view of the involvement of cytoskeletal components in the motile processes related to the cytopathic properties of the pathogen, such as adhesion, phagocytosis, and exocytosis. The nucleus is \( \sim 4-7 \) \( \mu \)m in diameter and has a bilayered membrane interrupted by numerous pores. The total DNA content is \( \sim 0.5 \) pg DNA per cell, all of which are present in the nucleus; there are no DNA-containing organelles (Byers, 1986). Chromatin clumps are uniform and evenly distributed. The ploidy level is still not clear; the consensus is about 4 (Willhoeft et al, 1999). In *E. histolytica*, in contrast to other unicellular eukaryotes, rDNA is present exclusively as extrachromosomal circular molecules (Bhattacharya et al., 1989).

**Biochemistry of *Entamoeba histolytica***:

*E. histolytica* is a facultative anerobe and some of its glycolytic enzymes are present only in other microaerophilic bacteria (Weinbach, 1981). This may relate to the apparent ability of the pathogenic strains to go from the environment of the intestinal lumen at low oxygen pressure to that encountered upon invasion of solid organs with an abundant blood supply. In many other aspects, *E. histolytica* resembles anaerobic and microaerophilic bacteria rather than typical eukaryotes, as demonstrated by the fact that it lacks glutathione metabolism (Fahey et al., 1984). In other organisms, glutathione protects against oxygen toxicity. High concentrations of cysteine and other thiols have been found in Entamoeba that carry out the functions of glutathione and glutathione-dependent enzymes (Fairlamb, 1989). Also, the presence of a disulphide oxido-reductase that binds
FAD as a cofactor can have a role in protection against oxygen toxicity (Bruchhaus et al., 1995). A unique feature of *E. histolytica* glycolytic pathway is the utilization of inorganic pyrophosphate instead of ATP as an energy source. In addition, the presence of multifunctional NAD* and NADP* -linked alcohol dehydrogenases, previously reported only in anaerobes and facultative anaerobes, has been reported (Yang et al., 1994).

**The nucleus and genome of *Entamoeba histolytica***:
The haploid genome size of *E. histolytica* strain HK-9 is $3 \times 10^7$ bp (Gelderman et al., 1971) based on replication kinetics. The *Entamoeba* genome is low (~22.4%) in [G+C] content. Variations in DNA content of cells in culture have been ascribed to the level of ploidy, repeated DNA and/or genomic heterogeneity in cultured axenic cells (Lopez-Revilla et al., 1978). The genome of *Entamoeba* consists of both linear chromosomes and plasmid-like circular DNA molecules; among the latter class are the rDNA circles (Dhar et al., 1995; Dhar et al., 1996). The ribosomal episomes are located in the vicinity of the nuclear membrane and are not associated with the chromosomes, and their segregation appears to be precede the chromosomal separation (Willhoeft et al., 2000). An improved separation of *E. histolytica* chromosomes by PFGE has been reported; most of the bands represented linear chromosomes as these were sensitive to Bal31 treatment (Bagchi et al., 1999). Willhoeft et al. (1999) demonstrated the presence of 14 linkage groups in *E. histolytica* with the ploidy being four. The *Entamoeba* genome does not condense to allow visual observation of mitosis and meiosis stages by traditional methods. FISH analysis of methanol:acetic acid spreads of genomic DNA also indicate that there are 14 chromosomes with a ploidy of 4 in *E. histolytica* (Willhoeft et al., 2000).

**Lipids and membrane of *Entamoeba histolytica***:
The membrane lipid composition is also unusual. The plasma membrane of *E. histolytica* is ~10nm thick and consists of an array of glycoproteins. An unusual
phospholipid, ceramide aminoethyl phosphonate (CAEP) has been shown to be present in internal vesicles and also in the plasma membrane (Aley et al., 1980). This may be physiologically important as CAEP is resistant to hydrolysis and may protect the parasite against the action of its own phospholipase. Interaction of the trophozoite plasma membrane with specific ligands induces a dramatic redistribution of surface components that accumulate at the uroid and are later released into the medium. This capping of surface molecules occurs through a sliding mechanism that involves both actin and myosin and is regulated by calmodulin and a myosin light chain kinase (Espinosa-Cantellano et al., 1994). Capping, resistance to complement (Reed et al., 1986) and masking with host proteins have been postulated as means to evade the host humoral response. Several E. histolytica surface molecules that mediate adhesion to cells or the intestinal mucosa include:

a. 260 kDa N-acetyl-D-galactosamine inhibitable lectin (Petri et al., 1989);
b. 220 kDa N-acetylglucosamine inhibitable lectin (Rosales-Encina et al., 1987);
c. 120 kDa surface adhesin (Arroyo et al., 1987) and
d. 110 kDa Proteophosphoglycan (PPG) (Bhattacharya et al., 1992)
Antibodies against these surface molecules abolish amoeba adhesion and phagocytosis of target cells.

Pathogenesis: Involvement of signalling pathways:
E. histolytica derives its name from its ability to lyse virtually every tissue in the human and other experimental animals. Lysis can take place via contact-dependent as well as contact independent mechanisms. Initial contact may be mediated by the various surface lectins especially the Gal/GalNAc lectin (Petri et al., 1995). Once attached, it has been suggested that the parasite releases an active 77 amino acid peptide, the amebapore, that is inserted into the host cell membrane as ion channels, resulting in bacteriocidal and cytolytic activities (Leippe et al., 1994). Tissue invasion is followed by degradation of the extracellular matrix by a multi-component pathway. Collagenases (secreted in
the so-called "electron-dense granules") and secreted cysteine proteases have been shown to act on a variety of ECM proteins viz., collagen, fibronectin, laminin and some proteoglycans. Cysteine proteinases, widely reported to be involved in ameba pathogenicity, are the most abundant proteases present in the cell (Keene et al., 1990). The ameba cysteine proteases (coded by at least six distinct genes, EhCP1 through to EhCP6) are secreted out but some have been noted to be present on the cell surface also. Interestingly, the expression of these proteases is 10- to 100-fold higher in the pathogenic *E. histolytica* than in the non-pathogenic *Entamoeba dispar* (Tannich et al., 1991). The non-pathogenic E. dispar trophozoites have only four cysteine proteinase genes, but apparently two of these are expressed (Brucchaus et al, 1996).

The involvement of signalling pathways as also the importance of the cytoskeleton in the pathogenesis of *E. histolytica* has been under detailed investigation (Guillen, 1996; Meza, 2000). Many cell surface receptors e.g., the 170 kDa Gal/GalNAc lectin, laminin receptor and a homolog of the leucocyte adhesion molecule, LECAM3, may initiate signalling after interaction with bacterial or mammalian cell surface molecules or ECM components. Transduction of these signals may be mediated through heterotrimeric G proteins (Soid-Raggi et al., 1998) and subsequently, GTPases, multifunctional 14-3-3 kinases and protein phosphatases may be involved.

Following the interaction of the trophozoites with the ECM components, especially fibronectin (FN), there are major structural and metabolic changes (Manning-Cela et al, 1997; Vazquez et al, 1995). There is a marked reorganization of the actin cytoskeleton with the formation of adhesion plaques, the signals being transduced via the activation of protein kinase C pathway. Another pathway seems to take place via a focal adhesion kinase, pp125*FAK*, reported to be a part of the complex of proteins known to be phosphotyrosine substrates for the src family of tyrosine kinases. Activation of this kinase might initiate a signal transduction pathway that activates the MAP Kinase cascade (Perez et al., 1996). Incubation of trophozoites with FN produces a sustained rise
in intracellular \([\text{Ca}^{2+}]\) levels, which promotes stabilization of adhesion plaques and focal contacts by shifting soluble, monomeric G-actin to the polymerized F-actin configuration. External \(\text{Ca}^{2+}\) influx has been shown to be responsible for this increased \([\text{Ca}^{2+}]\) levels (Carbajal et al., 1996). An 80 kDa CD44 cross reactive protein on the surface of \(E.\ histolytica\) that also may mediate the interaction between intestinal epithelial cells and the parasite (Renesto et al., 1997).

Pathogenicity of \(E.\ histolytica\) may be regulated by the intestinal bacterial flora e.g., mucin degrading bacterial glycosidases and colonic luminal proteases together, but not alone, degrade the key adherence lectin on \(E.\ histolytica\) trophozoites resulting in decreased epithelial cell adherence (Variyam, 1996). These in vitro findings suggest that a potential novel host defence mechanism exists in the human colon wherein the invasiveness of a pathogen can be controlled by the combined action of bacterial and host hydrolases.

**Role of Calcium in Entamoeba:**

Essential components of the already documented \(\text{Ca}^+\) signalling pathway, e.g. IP\(_3\) and IP\(_4\) mobilizable internal calcium stores and P type \(\text{Ca}^+\) ATPases have been shown to be present in \(E.\ histolytica\) (Figure 2). Raha et al (1994, 1995) showed that the intracellular \(\text{Ca}^+\) can be stored in internal vesicles; these vesicles also contain the IP\(_3\) and IP\(_4\) receptors (Giri et al., 2001). Calmodulin (CaM), a highly conserved ubiquitous calcium-binding protein, has been shown to be present in \(E.\ histolytica\) by partial purification and functional analysis (Munoz et al., 1992). It was demonstrated that \(\text{Ca}^{2+}/\text{CaM}\) might have a role in the release of collagenases through degranulation of electron-dense granules (Munoz et al., 1991; de Lourdes et al, 2001).

The P type \(\text{Ca}^+\) ATPase of \(E.\ histolytica\) is homologous to a human plasma membrane \(\text{Ca}^{2+}\) transporter and \(S.\ cerevisiae\) vacuolar \(\text{Ca}^{2+}\) transporting ATPase (Descoteaux et al, 1995). This might couple the \(\text{Ca}^{2+}\) signal to the moiety required to activate protein kinase C (PKC), with the consequent stimulation of adhesion plate formation and adhesion to fibronectin (FN) (Santiago et al, 1994).
Figure 2: Different Calcium signalling pathways in *Entamoeba histolytica*
This process is essential for the parasite interaction with ECM and its subsequent degradation. It is possible that after activation of the PKC or adenylyl cyclase pathways, actin genes in amoeba could be regulated by phosphorylation of transcription factors that bind to regulatory elements located in the actin gene promoter. The formation of adhesion plates, a pre-requisite for host cell invasion, requires not only the synthesis of actin, but also the recruitment of F-actin to the sites of contact and its interaction with several F-actin-binding proteins, many of which are also activated by phosphorylation by PKC, PKA and other kinases.

Ravdin et al (1988) reported oscillatory regional increases in parasite intracellular [Ca$^{2+}$] independent of the presence of target cells in the vicinity; this may be related to the motility, chemotaxis and normal cytoskeletal functions of the parasite. However, on contact with a parasite, there is an immediate increase in the host cell Ca$^{2+}$. Amebic cytolytic activity can be blocked by either Ca$^{2+}$ channel blockers or in the presence of EGTA. Thus, Ca$^{2+}$ is speculated to play a direct role in the *E. histolytica*-induced target cell death.

Nickel et al (2000) reported the molecular characterization of two novel Ca$^{2+}$ dependent granule proteins, named grainin1 and 2 in *E. histolytica*. Both of these proteins possess three EF hands each. The authors postulated that these granule proteins may be implicated in functions vital for the primitive phagocytosis processes, control of endocytotic pathways and Ca$^{2+}$ dependent granular discharge.

The changes in the calcium profile is also related to the cell cycle and the developmental stages of the parasite i.e., the cyst or the trophozoite stages, implicating its role in the cyst forming activity of *E. invadens* (Ganguly et al, 2001). Furthermore, extracellular Ca$^{2+}$, amebic intracellular calcium flux, bepridil-sensitive calcium channels and a putative CaM-dependent signal transduction pathway have been implicated in the growth and encystations of *Entamoeba* (Makioka et al, 2001).
Calcium as a universal second messenger:
Ionized calcium is one of the most universal agents for signal transduction throughout the living forms controlling diverse cellular processes (Berridge, 2000). Ca$^{2+}$ is able to accommodate 4-12 oxygen atoms in its primary coordination sphere, but coordination numbers of 6-8 are most commonly seen (McPhalen et al., 1991). Proteins bind Ca via six oxygen atoms, often provided by Glutamate and Aspartate residues. Quiescent or resting cells have an intracellular Ca$^{2+}$ concentration of 100nM; activated cells can rapidly take up Ca$^{2+}$ to levels of 1uM. This can regulate different cellular processes due to variations in the speed, amplitude and spatio-temporal patterns of the Ca$^{2+}$ signal. Two important facts regarding Ca$^{2+}$ signalling are:

- (1) Ca$^{2+}$ cannot be metabolized, unlike other second messenger molecules; this requires a tight regulation of the intracellular Ca$^{2+}$ with the help of ion channels and specialized binding proteins;
- (2) Ca$^{2+}$ can precipitate phosphates, the universal energy currency. It also has lower affinity for water than other divalent cations like Mg$^{2+}$. Therefore, the cells initially evolved strategies for binding Ca$^{2+}$, perhaps at first simply to reduce its cytoplasmic levels, but later for signal transduction (Clapham, 1995).

Generation of calcium signals:
(a) Generation of Ca$^{2+}$ signal is initiated via Ca$^{2+}$ channels or transporters that are located either on the plasma membrane or in the internal ER/SR membranes. The Ca$^{2+}$ channels have been classified according to their mode of action:

i) Receptor Operated Channels:
Non-excitable cells like blood cells, hepatocytes and endothelia take up Ca$^{2+}$ via their slow IP$_3$ (Inositol trisphosphate) mediated pathway. IP$_3$, the second messenger involved, is released by predominantly two receptors: the G protein-coupled receptor class of seven transmembrane-spanning receptors (GCRs) and the receptor tyrosine kinases (RTKs). GCRs activate phospholipase β whereas RTKs activate phospholipase γ to convert phosphatidylinositol (4,5) biphosphate
into IP₃ and DAG (Berridge et al., 1989), leading to an increase in the intracellular Ca²⁺.

ii) Voltage Operated Channels (VOC):
Another Ca²⁺ entry pathway in cells is mediated by membrane hyperpolarization. Open K⁺ channels force the membrane potential to more negative potentials, drawing Ca²⁺ more rapidly across the membrane. Specialized voltage-independent Ca²⁺ selective channels activated by second messenger molecules also operate in cells.

iii) Mechanically Operated Channels:
The Plasma Membrane Calcium Pump (PMCA) is a P type ATPase (Pederson et al., 1987) having ten transmembrane domains and constitutes an important Ca²⁺ entry route in tissues other than excitable tissues like the heart, where the more powerful Na⁺/Ca²⁺ exchanger predominates. However, in cells where the latter predominates, the PMCA may play a role in fine-tuning of cytosolic Ca²⁺, operating in a concentration range where the low-affinity exchanger cannot operate efficiently. Recent study by Guerini et al (2000) on cultured cerebellar neurons has shown that the expression of the isoforms of this pump is transcriptionally regulated by Ca²⁺.

In various tissues like neurons, skeletal muscle and endocrine cells, Na⁺-dependent Ca²⁺ efflux is mediated by a carrier that exchanges three Na⁺ for two Ca²⁺. This exchanger extrudes Ca²⁺ against its large electrochemical gradient using the energy provided by the inward movement of Na⁺ down its steep electrochemical gradient. This exchange is electrogenic and is regulated by the membrane potential.

iv) Store Operated Channels (SOC):
The so-called "store operated" Ca²⁺ channels refill empty intracellular Ca²⁺ stores in the ER following agonist-elicited Ca²⁺ release, a process called "capacitative calcium entry". This is established via the generation of an inward Ca²⁺ current (I_{CRAC/DAC}, calcium release activated or depletion activated current; Fasolato et al,
1994). $I_{CRAC}$ has an extremely low conductance of $\sim 20 \text{fS}$, with the net current passing being $\sim 5 \text{pA}$; in contrast, the net current observed in neurons in presence of activated voltage-dependent Ca$^{2+}$ channels is hundreds of pA. However, this capacitative Ca$^{2+}$ influx is necessary to maintain a normal secretory pattern (Neher, 1988) in mast cells and basophils.

(b) On the other hand, release of Ca$^{2+}$ from the internal stores like ER/SR and mitochondria occur through the following major pathways:

i) **IP$_3$ Receptors (IP$_3$Rs) and Ryanodine Receptors (RyRs):**
The principal activator of these internal Ca$^{2+}$ channels is Ca$^{2+}$ itself via the principle of Ca$^{2+}$-induced Ca$^{2+}$ release (CICR). When external stimuli interact with their extracellular receptors, Ca$^{2+}$ mobilizing second messengers are generated. One such messenger is IP$_3$ that diffuses into the cell to bind to IP$_3$R whereas the RyRs act via cyclic ADP ribose (cADPR). The IP$_3$Rs and RyRs are coexpressed in numerous cell types, from neurons to smooth muscle cells. Varying combinations of their expression in different cell types may be important in the regulation of specific functions. Both of these are several times larger than the voltage-gated Ca$^{2+}$ channels and their Ca$^{2+}$ conductance is about 10 times greater (100pS).

ii) **Nicotinic Acid Adenine Dinucleotide Phosphate (NAADP$^+$) Receptor:**
NAADP$^+$ is the most recently characterized second messenger that interacts with an unknown receptor and releases Ca$^{2+}$ from internal stores apart from ER (Lee and Aarhus, 1995). This Ca$^{2+}$ release is antagonized by L-type Ca$^{2+}$ channel blockers, whereas that by IP$_3$ and cADPR is not.

iii) **SCaMPER (sphingolipid Ca$^{2+}$ release mediating protein of SR):**
SCaMPER is a 181 amino acid ER protein with two putative membrane-spanning domains having ion-channel activities activated by sphingosyl phosphocholines (Mao et al, 1996). Because SCaMPER mediated Ca$^{2+}$ release is insensitive to either ryanodine or La$^{3+}$, this forms a novel Ca$^{2+}$ permeable channel activity distinct from either RyRs or IP$_3$Rs.
Elementary calcium signals: local and global patterns:

Intracellular Ca\(^{2+}\) gradients function to initiate cell migration, exocytosis, lymphocyte killer cell activation, acid secretion, transcellular ion transport, neurotransmitter release, gap junction regulation and a variety of other activities (Tsien et al., 1990). The most complex wave patterns of Ca\(^{2+}\) signals, exhibiting hot spots and spherical and planar waves have been shown in Xenopus oocytes (Lechleiter et al., 1991). The early reports on Ca\(^{2+}\) signalling characterized elementary Ca\(^{2+}\) entry and release signals: “quantum emission domains” (QEDs) in giant squid synapses and “bumps” in Drosophila photoreceptors. Recently, elementary Ca\(^{2+}\) signals associated with the release of Ca\(^{2+}\) from internal stores have been referred to as “quarks”, “sparks”, “blips” and “puffs”. These names reflect whether the event is produced by RyRs (quarks and sparks; Cheng et al., 1993) or by IP\(_3\)Rs (blips and puffs; Yao et al., 1995) and whether it is associated with the opening of a single channel or a group of channels. The large variety of cellular Ca\(^{2+}\) binding proteins (CaBPs) with unique Ca\(^{2+}\) binding rates and affinities dictates that Ca\(^{2+}\) waves and oscillations will have differential effects in cells (Clapham et al., 1995).

Elementary local Ca\(^{2+}\) signals control exocytosis in synaptic and secretory vesicles, activation of ion channels, mitochondrial energy metabolism and the generation of nuclear-specific Ca\(^{2+}\) signals. It has a limited spatial range (1-6 \(\mu\)m) and allows Ca\(^{2+}\) to exert a highly specific effect. Also, these can cause rapid cellular effects at relatively low energy costs to the cells. On the other hand, globalized events like Ca\(^{2+}\) waves and oscillations set up slow waves spreading in a regenerative manner throughout the cells at 10-100\(\mu\)m/s. For Ca\(^{2+}\) waves to occur, most of the IP\(_3\)Rs and RyRs must be sufficiently sensitive to Ca\(^{2+}\) to respond to each other via CICR. One group of Ca\(^{2+}\) channels release Ca\(^{2+}\), which then diffuse to neighbouring receptors to excite further Ca\(^{2+}\) release, thereby initiating the process. When gap junctions connect cells, these intracellular waves can spread to neighbouring cells to create intracellular waves capable of coordinating the activity of many cells (Lansley et al, 1999). CICR generates an...
oscillatory \( \text{Ca}^{2+} \) release pattern combined with its periodic blockade by the depletion of \( \text{Ca}^{2+} \) stores and with the direct inhibition of the ER/SR receptor channels by high \( \text{Ca}^{2+} \) (Parker et al, 1992). This oscillatory pattern may be a means to allow \( \text{Ca}^{2+} \) signalling to occur avoiding the potentially deleterious permanent increase of \( \text{Ca}^{2+} \) in the cell. This may be essential for processes that demand longer periods of signalling than can be executed by single brief \( \text{Ca}^{2+} \) transient: repetitive spiking may ensure proper functioning of the cells while at the same time avoiding \( \text{Ca}^{2+} \) death. It is interesting to note that the amplitude and frequency of oscillations (that may vary from a few seconds to minutes) have very specific effects on gene transcription (Dolmetsch et al, 1997). For decoding the frequency-modulated signalling system, cells have developed "molecular machines" such as CaMKII and PKC. Frequency coding is used to control various cellular processes such as liver metabolism, smooth muscle contractility and differential gene expression. Dolmetsch et al (1997) have demonstrated that a low frequency \( \text{Ca}^{2+} \) spike activates the NF-\( \kappa \)B transcription factor while higher frequencies activate NF-\( \kappa \)AT. It has been shown that CaMKII can immediately integrate or decode frequency-encoded intracellular \( \text{Ca}^{2+} \) signal. Regardless of the \( \text{Ca}^{2+} / \text{CaM} \) pulse duration, it was conclusively shown that autonomous activation of CaMKII increased steeply as a function of frequency. Expectedly, shorter pulse durations required greater frequencies for activation, but once the threshold has been achieved, the steepness of the frequency activation curve was much greater. In summary, it is possible to prime the system with \( \text{Ca}^{2+} \) bursts of a particular frequency and then subsequently to maintain the response levels with signals of substantially lower frequency. This might be correlated with a cell's ability to distinguish between an "intentional" \( \text{Ca}^{2+} \) signal from potentially spurious and artefactual \( \text{Ca}^{2+} \) changes.
Functions of Ca\(^{2+}\) signalling:

The most important feature of Ca\(^{2+}\) signalling is its unique versatility. As Berridge et al. (1999) pointed out that Ca\(^{2+}\) occurs as a fundamental moiety throughout the lifecycle of an organism. It triggers the initiation of life at fertilization, it mediates various developmental cascades and in differentiated cells it controls diverse functions like contraction, secretion, metabolism, proliferation and cell cycle, learning and memory and finally cell death.

i) Fertilization:

During fertilization, Ca\(^{2+}\) spikes abut for about two hours in mammalian eggs. Elevation of egg [Ca\(^{2+}\)] at fertilization causes exocytosis of cortical granules and this ensures the elevation of the fertilization envelope and prevents polyspermy. A G-protein linked sperm receptor (Shilling et al., 1994) or a 57 kDa src family kinase (Abassi et al., 2000) may be involved in triggering Ca\(^{2+}\) transients. In hamster sperm, an oligomeric protein (monomer MW 33 kDa) called oscillin has been implicated (Parrington et al., 1996) in from spermatozoa cytoplasm enters the oocyte after fusion causing. Oscillations in Ca\(^{2+}\) levels after fertilization cause loss of MPF activity in oocytes and initiate a cascade of dephosphorylation events. The Ca\(^{2+}\) spikes following fertilization stimulate CaMKII which acts via CDC25 to dephosphorylate CDK1 (cyclin-dependent kinase 1), resulting in cyclin B activation and completion of meiosis. As the embryo approaches mitosis, a series of Ca\(^{2+}\) transients cause nuclear envelope breakdown and cell cleavage.

ii) Post developmental cell proliferation/cell cycle/tumorigenesis:

Depletion of intracellular Ca\(^{2+}\) activates eIF-2\(\alpha\) kinase or protein kinase R (PKR). PKR phosphorylates and thereby inhibits eIF2\(\alpha\), a rate limiting step in translation initiation. In addition, using Ca\(^{2+}\) channel blockers like nifedipine and verapamil, it has been shown that these inhibit cell proliferation by a cell cycle arrest in Go/G1 phase (Zeitler, h. et al., 1997).

Ca\(^{2+}\) waves cause anaphase onset and mitosis exit by inducing the degradation of the cyclin subunit of cdc2. It also inactivates the mos/MAPK pathway, thus
preventing its action on the subsequent steps of meiosis. However, the importance of Ca\textsuperscript{2+} signals in mammalian somatic cell mitosis is still being debated; it is involved in chromosome disjunction, but not in spindle elongation or chromosome decondensation. Ca\textsuperscript{2+} is one of the key regulators of cell proliferation, functioning in conjunction with other signalling pathways such as those controlled by MAPK and phosphatidylinositol-3-OH kinase (PI(3)K) (Lu et al., 1993). Ca\textsuperscript{2+} can interact with the MAPK signalling pathway by activating a proline-rich tyrosine kinase 2 (PYK2) that acts via the GTPase ras to induce the MAPK cascade.

iii) NO in Ca\textsuperscript{2+} signalling:

NO (Nitrogen monoxide) is a molecular messenger controlling neurotransmission, blood clotting, arterial pressure and macrophage defense mechanisms. The different isoforms of NO synthase are either associated tightly with a CaM subunit or have CaM binding sites. NO enhances cGMP production; which then decrease cellular Ca\textsuperscript{2+} either by phospholamban-activated SERCA pump activity or phosphorylation of InsP\textsubscript{3}Rs.

iv) Intracellular Proteolysis and Ca\textsuperscript{2+}:

A Ca\textsuperscript{2+}-dependent protease known as the "nuclear scaffold protease" can degrade the family of nuclear matrix proteins, the lamins (Clawson et al., 1992). Moreover, a family of ubiquitous cysteine proteases known as calpains possess Ca\textsuperscript{2+} binding domains (Carafoli et al., 1998). Disruption of skeletal muscle specific calpain has been shown to be associated with limp girdle muscular dystrophy type 2A where the defective calpain cannot degrade IκB proteins, thus sequestering NFκB in the cytoplasm. Excessive calpain activity is involved in several pathological conditions like Alzheimer's disease, cataract and multiple sclerosis.

v) Ca\textsuperscript{2+} and secretion:

Secretion of biologically active compounds from intracellular vesicles requires the presence of Ca\textsuperscript{2+}. Ca\textsuperscript{2+} causes the fusion of vesicles by acting in conjunction with N-ethyl-maleimide-sensitive factor (NSF) and the soluble NSF attachment protein.
(SNAP). The SNARES (SNAP receptors) are activated by Ca\(^{2+}\) when the VOCS manage to create subplasma membrane microdomains of high [Ca\(^{2+}\)]. Ca\(^{2+}\) controls many steps in exocytosis like vesicle recruitment from the membrane and its subsequent docking and fusion with the help of proteins like synaptobrevins (VAMPS) and syntaxins. CaBPs like annexin, CaM and S-100 proteins act as Ca\(^{2+}\) sensors necessary for secretion of granules.

vi) Ca\(^{2+}\) and Muscle Contraction:
Troponin is the Ca\(^{2+}\) sensor present universally in skeletal and cardiac muscle myofibrils. Troponin along with tropomyosin, represses the contractile interaction of actin and myosin, a state which is removed by binding of troponin to Ca\(^{2+}\). In skeletal muscle, L-type VOC in the plasma membrane interacts directly with the cytoplasmic head of RyRs of SR whereby membrane depolarization is transduced to Ca\(^{2+}\) secretion from SR. In smooth muscle, Ca\(^{2+}\) activates CaM that phosphorylates MLCK and causes actin-myosin contractility.

vii) Ca\(^{2+}\) in Memory and Learning:
Memory storage can be of two types: LTP, which is the sustained increase of the efficiency of synaptic transmission caused by brief pulses of high-frequency stimulation and LTD, which is the sustained depression caused by brief activation of an excitatory pathway. CaMKII is the most abundant protein in the postsynaptic density and has been implicated to "remember" the synaptic activity that had triggered the Ca\(^{2+}\) induced conversion of CaMKII from a Ca\(^{2+}\)-dependent to a Ca\(^{2+}\)-independent autophosphorylated form (Fukunaga et al., 1993). It has been shown that targeted disruption of the CaMKII genes inhibits LTP in hippocampal slices and impairs spatial memory.

viii) Ca\(^{2+}\) and apoptosis:
Programmed cell death or apoptosis involves the concerted action of sphingomyelin signalling pathway, the redox system, the stress-activated protein kinase cascade and the Ca\(^{2+}\) signalling pathway. The continuous flow of Ca\(^{2+}\) between the mitochondria and ER can be adjusted by the presence of pro-
apoptotic molecules like ceramides; these two signals i.e., altered Ca\(^{2+}\) dynamics and ceramides can act simultaneously to help in the formation of permeability transition pores (PTP) in mitochondria. Ca\(^{2+}\) also helps in the expression of Fas. In addition, a low molecular weight Ca\(^{2+}\) dependent nuclease, NUC18, activated in apoptotic lymphoid cells (Gaido et al., 1991) and a Ca\(^{2+}\) dependent DNase I (Peitschle et al., 1993) have been characterized in apoptotic cells. Death agonists (Bax, Bad, Bak) and death antagonists (Bcl-2, Bcl-XL) interact with the Ca\(^{2+}\) flow between the ER and mitochondria. Recent reports indicate that Bcl-2 enhances the storage of Ca\(^{2+}\) perhaps by upregulating the SERCA gene expression (Kuo et al., 1998). Yano et al (1988) have recently demonstrated that CaMKK can activate protein kinase B/Akt, a serine/threonine protease that can prevent apoptosis by phosphorylating the pro-apoptotic protein Bad.

\[\text{ix)} \text{Ca}^{2+} \text{ and Transcription:}\]

During cell proliferation, Ca\(^{2+}\) controls the activity of transcription factors like NF-AT, NF-kB and CREB and thus control cell proliferation. These transcription factors activate numerous target genes e.g. interleukin 2 system genes (switches on DNA synthesis) and Fas ligands that trigger apoptosis. Ca\(^{2+}\) stimulates the Ca\(^{2+}\) dependent protein phosphatase calcineurin which dephosphorylates NF-AT, enabling the latter to enter the nucleus. Interestingly, in Sachharomyces cerevisiae, CaN regulates the transcription factor Tcnlp/Crzlp (Matheos et al., 1997) and down regulates the transcription of a PMCA pump (Guerini et al., 2000) and a Na\(^+\)/Ca\(^{2+}\)exchanger (Liet et al., 2000) in cerebellar granules neurons.

A CaM inhibitory peptide targeted to the nucleus can block DNA synthesis and cell cycle progression (Wang et al., 1996). Ca\(^{2+}\) also activates gene transcription directly by antagonizing the activity of DREAM (Downstream regulatory element of prodynorfin and C-fos genes), a tetrameric EF hand protein (Carrion et al., 1999). Prodynosfin is involved in neuronal processes like memory acquisition and pain sensation. In the presence of Ca\(^{2+}\), DREAM dissociates from DRE.
(down stream regulatory element of the gene) releasing the transcription repression. Interestingly, DREAM is homologous to colsenilin which binds presenilin involved in Alzheimer's disease. In *E. histolytica*, a similar EF-hand containing transcriptional factor, URE3-BP, has been characterized which regulates transcription of the virulence factor Gal/GalNAc-inhibitable lectin. Calcium decreases the affinity of URE3-BP for its target DNA; however, the amino acid sequence of this novel protein has little similarity to other known DNA-binding proteins (Gilchrist et al., 2001). Ca$^{2+}$ regulation has been shown to exist for other bHLH transcription factors e.g. these affecting the AML 1 family (Comeliussen et al., 1994).

**x) Ca$^{2+}$ in Protozoan Biology:**

Several cellular and physiological processes regulated by calcium has been well documented in protozoa (Scheibel, 1992). In many aspects, the protozoan system differs from already established mammalian pathways, e.g. many CaM-regulated enzymes, such as Ca$^{2+}$ ATPase and adenylyl cyclase are CaM-independent in protozoa under *in vitro* conditions. CaM has been characterized in many protozoa, e.g., *Tetrahymena*, *Trypanosoma*, *Leishmania*, *Giardia* and *Plasmodium*. PKC has also been identified in many protozoa (Munoz et al., 1987). However, in many instances, these enzymes have been described only biochemically and their role in the so-called calcium signal transduction pathway are yet to be well documented.

In *Plasmodium falciparum*, the PfCDPK1 is CaM-independent and is involved in active membrane biogenesis and thus plays an important role in erythrocyte invasion (Zhao et al., 1994). Another Ca$^{2+}$ dependent kinase, PfKIN is a homolog of the yeast SNF which is involved in stress response. The expression of CaM, a housekeeping gene, has been shown to be stage-specific in Plasmodium (Rojas et al., 1995). It is involved in erythrocyte invasion as well as schizont maturation (Matsumoto et al., 1987). Interestingly, CaM levels in infected RBCs come down by as much as 30%. In *Plasmodium knowlesi*, invasion by merozoites can be
Introduction

prevented by chelation of Ca\(^{2+}\), thus implicating a role of Ca\(^{2+}\) in parasite invasion (Johnson et al, 1981).

The role of calcium in schistosome biology is of paramount importance (Modha et al, 1999). Calcium not only inhibits protease activity in host glands but also facilitates parasitic glycocalyx removal, thus limiting damage to the parasite membrane that might result from complement attack.

In Trypanosomatids, various cellular processes such as Ca\(^{2+}\) dependent fragmentation of nuclear DNA, loss of motility, environmental sensing, host cell invasion and cell death are mediated by numerous cellular calcium-binding proteins, including CaM (Pereira et al, 1997). Other CaBPs include isoforms of adenylate cyclase, protein kinases, endonucleases, and flagellar EF hand CaBPs. In T. cruzi, CaM has been shown to be involved in regulation of cell growth; deletion of the CaM gene slows the procyclic parasite growth by \(
\text{\sim}50\%\) (Eid et al, 1991). CaM binding motifs occur in many variant surface glycoproteins. T. cruzi trypomastogotes invade a variety of mammalian cells triggering an increase in the intracellular calcium, an effect not observed with the non-infective epimastigote stage (Tardieux et al, 1994). Treatments that decrease or increase cytoplasmic calcium (Yakubu et al, 1994) modifies T. cruzi infectivity downward or upward respectively.

In Leishmania, Olivier et al (1992) showed an increase in intracellular calcium in human monocytes following infection by amastigotes. However, the late changes in intracellular calcium in Leishmania-infected host cells seem to be more probably related to the toxicity of the parasites to the host cells rather than to the invasion process. A Leishmania Ca\(^{2+}\) ATPase has also been identified as a potential virulence factor. Lu et al (1997) demonstrated a significant link between the expression of the Ca\(^{2+}\) ATPase, intracellular Ca\(^{2+}\) pool and virulence. In addition, centrin, a Ca\(^{2+}\) dependent cytoskeletal protein was found to be essential for overcoming the G2/M point, thus implicating a role of calcium in Leishmania growth (Selvapandiyan et al, 2001).
CaM has been identified in *Giardia lamblia* as a factor involved in its excystation process. This is crucial for the initiation of infection by this protozoa (Bernal et al., 1998). Two types of Ca\(^{2+}\) ATPases have been identified in this protozoon: one, a membrane-associated Ca\(^{2+}\) activated one and the other, a soluble, cytoplasmic enzyme activated by Mg\(^{++}\). It has been shown that the membrane-associated Ca\(^{2+}\) ATPase is involved in intracellular Ca\(^{2+}\) transport in a CaM-dependent manner.

**Calcium Binding Proteins:**
CaBPs are the mediators of Ca\(^{2+}\) signals. Intracellular CaBPs can be broadly subdivided into two categories:

(A) proteins which bind Ca\(^{2+}\) to modulate its concentration in the cellular context:

1. **EF-hand proteins:** A classical EF hand (Figure 3) is a helix-loop-helix motif characterized by a sequence of ~12 residues that participate in metal coordination. The three dimensional arrangement of these domains is reminiscent of the thumb, index and middle fingers of a hand, hence the name. These proteins, broadly respond to Ca\(^{2+}\) in two ways: - one group (e.g., parvalbumin and calbindin) do not undergo a significant conformational change on binding Ca\(^{2+}\) and essentially function as Ca\(^{2+}\) buffers whereas the second group, the Ca\(^{2+}\) sensors (e.g., calmodulin, troponin C, neuronal protein recoverin, the S100 proteins) undergo Ca\(^{2+}\)-induced conformational changes and transduce the signal to downstream signalling molecules.

2. **CaM binding domains in target proteins:** proteins e.g., MLCK which bind to Ca\(^{2+}\)-CaM through an amphipathic helix. This is a Ca\(^{2+}\)-dependent process. However, a novel Ca\(^{2+}\)-independent CaM binding motif named IQ has been characterized (Rhoads et al., 1997) in proteins like neuromodulin.

3. **Non EF hand proteins:** these proteins bind Ca\(^{2+}\) with less affinity than EF hand proteins. The best characterized modules are the C2 domains (present in
Figure 3: A typical EF hand bound to calcium (blue circle) which is coordinated through hydrogen bonds with acidic residues as well as backbone oxygen atoms. The EF hand consists of an alpha helix E, a loop around the calcium ion and a second helix F.
PKC, cPLA2, PLC, calpain) and the endonexin fold (present in the annexin family).

(4) Ca$^{2+}$ storage proteins with non EF hand motifs e.g., calsequestrin, calreticulin, chromogranins etc. These typically bind Ca$^{2+}$ with low affinity and high capacity and regulate Ca$^{2+}$ homeostasis.

(B) Transporter proteins:

Ca$^{2+}$ transporters: The function of these proteins is to regulate the intracellular Ca$^{2+}$ concentrations. These are located in the plasma membrane and organellar membranes and bind Ca$^{2+}$ with widely differing affinities. Some of these are organized as membrane channels (Voltage-Operated, Receptor-Operated and Store-Operated channels of plasma membrane; InsP$_3$R and RyR and NAADP+ gated channels of the ER/SR) that interact loosely with Ca$^{2+}$ and transport Ca$^{2+}$ across membranes. They do so in response to gating events operated by potential changes, by the interaction of messengers with specific domains of channel proteins or by releasing of intracellular Ca$^{2+}$ stores.

Other proteins interact with Ca$^{2+}$ with high affinity (i.e. the Ca$^{2+}$ ATPases) that transport Ca$^{2+}$ against a steep concentration gradient using ATP-derived energy. Other proteins (e.g. Na+/H+ exchanges or Na$^+$/Ca$^{2+}$ exchanger) bind Ca$^{2+}$ with intermediate affinity and transport Ca$^{2+}$ in exchange of other ions transport in the opposite direction.

Calmodulin:

CaM is a ubiquitously expressed highly conserved, Ca$^{2+}$ binding protein (148 residues in vertebrates) comprising of four EF hands. The first two EF hands combine to form a globular N-term domain that is separated by a short flexible linker from a highly homologous C-terminal domain containing two more EF hands. The C terminal EF hands have a 3-5 times higher affinity for Ca$^{2+}$ than the N-terminal ones. In the absence of Ca$^{2+}$, the N-terminal domain of the apo CaM
Introduction

molecule adopts a 'closed' conformation in which the four EF hands are packed together; in the presence of Ca\(^{2+}\), the C-terminal domain adopts a 'semi open' conformation whereby there is a concerted exposure of hydrophobic groups that may interact with target proteins. The binding of target proteins by CaM increases the affinity of CaM for Ca\(^{2+}\) by \(~10\) times (Peersen, et al., 1997) and sensitizes the CaM-effector complex to changes in Ca\(^{2+}\) concentration. Ca\(^{2+}\)/CaM is involved in the interaction with a variety of cellular effectors. The crystal structures of Ca\(^{2+}\) saturated CaM reveal that the protein has a dumbbell shape comprising the N- and C-terminal globular domains connected by a long \(\alpha\)-helix, the so-called 'central helix' (Babu et al., 1985). A range of biophysical studies including crosslinking experiments and small-angle x-ray and neutron scattering indicates that upon complexation with certain target peptides, CaM adopts a globular conformation. CaM uses its two solvent-exposed hydrophobic surfaces to engulf the aromatic and long aliphatic side chains of residues on target proteins (Ikura et al., 1992). In addition, the proper location of negatively charged residues in the vicinity of the hydrophobic surfaces is critical for the binding of CaM to its targets. The modulation of the inherent plasticity of CaM by different ligands is a key element of molecular recognition and the mechanism of signal transduction. NMR studies (Ikura et al., 1991) have demonstrated that the middle portion (residues 77-81) of the 'central helix' is non-helical and flexible in solution. Upon binding to target region, this flexible linker becomes further extended which allows CaM to accommodate targets of different sizes. The mode of binding of the regulatory light chain seen recently in the S1 myosin crystal structure (Rayment et al., 1993) has features that resemble those of the bound Ca\(^{2+}\)/CaM structures. The central helix of Ca\(^{2+}\)/CaM plays a key role by acting as a variant "expansion joint", allowing different relative positioning of the lobes as different target enzymes are recognized. Another study by Ikura et al. (1992) demonstrated the binding of Ca\(^{2+}\)/CaM with the M13 peptide (derived from rabbit MLCK) in which the target peptide is sequestered into a hydrophobic channel formed by the two domains of CaM with interactions involving 19
residues of M13. A key requirement appears to be the presence of two long chain hydrophobic or aromatic residues separated by 12 residues in order to anchor the peptide to the two domains of CaM (Ikura et al., 1995); the rope (i.e., the CaM binding domain of the target) has to be long enough and have two knots at each end for the two hands (domains) of CaM to grip it. Helices of varying length can be accommodated by the flexibility of the interconnecting linker region of CaM, which acts as a hinge (Gerstein et al., 1994), joining the globular domains as they fold around the target peptide. Ca\textsuperscript{2+} serves to organize and stabilize the domain structure in a conformation that can bind the target, while the central helix remains flexible, an essential condition for target recognition. Target binding further stabilizes the domain structure, which increases the Ca\textsuperscript{2+} affinity. Thus, the Ca\textsuperscript{2+} signal restricts the available conformational states of CaM to those that are most favourable to target recognition and activation (Meador et al., 1993).

Biochemically, it has been demonstrated that various mutations and deletions in the flexible linker can be accommodated without destroying CaM activity because these residues form few contacts with the target domain and may thus be important in allowing the globular domains to come closer together upon complexation (Kretsinger, 1992).

**Calmodulin-effector coupling:**

**Use of peptides to study binding and activation:**

The Ca\textsuperscript{2+} controlled exposure of hydrophobic groups in the two domains of CaM releases a considerable amount of biochemical energy that is accepted by the target leading to a change in the affinity of CaM for the effectors and/or an alteration in the effectors functions. A short peptide for ~20 residues responsible for binding Ca\textsuperscript{2+}/CaM, designated as CaM binding domain has been identified in many CaM regulated proteins. CaM binding domains in typical target proteins do not have any sequence homology. They often comprise a stretch of 20-25 residues with a potential to form a positively charged amphipathic helix (Crivici et
The crystal structure of CaM kinase I reveals that the CaM binding domain directly interacts with and sterically obstructs the putative substrate-binding site of the inactive enzyme (Goldburg et al., 1996). A homologous hydrophobic residue is conserved in other CaM kinases. Spectroscopic and crystallographic studies of Ca$^{2+}$/CaM complexed with these CaM binding domains have shown that this conserved hydrophobic residue interacts exclusively with the Methionine rich hydrophobic pocket in the C terminal domain of Ca$^{2+}$/CaM. Recently determined 3-D structures of Ca$^{2+}$/CaM bound to peptides from the plasma membrane Ca$^{2+}$-ATPase pump and a CaMKK revealed additional modes of interaction. These studies indicate that the C-terminal domain of Ca$^{2+}$/CaM might confer binding energy on the intact enzymes. Results from crystallographic studies show that hydrophobic residues in the N-terminal domain of Ca$^{2+}$/CaM mainly interact with the C terminal part of the CaM-binding peptides of smMLCK and CaMKII respectively. These mechanistic studies indicate that the regulation of enzymes by Ca$^{2+}$/CaM is a highly ordered, cooperative and complementary process that contributes to both the affinity and specificity for targets. A surprising and welcome observation is the discovery that the structures of Ca$^{2+}$/CaM/peptide complexes are relevant to their corresponding enzymes. However, there are mounting speculations that peptides might not be entirely suitable for studying other forms of a full-length target protein. For example, Ca$^{2+}$ has a considerable higher affinity for the CaM-MLCK-peptide complex than for the corresponding CaM-MLCK-enzyme complex (Peerson et al, 1997). One explanation for the adaptability and the higher affinity exhibited by the shorter CaM-binding domains is the conformational flexibility inherent in isolated peptides. However, peptides can be and have been successfully used to model effector functions in the Ca$^{2+}$/CaM system (Table1).
Table 1: Alignment of calmodulin binding peptides SK MLCK, SM MLCK, Mastoparan, and Melittin. Relative physical properties of the peptides have been designated as polar $S^\circ$, $T^\circ$, $Q^\circ$, $N^\circ$, $H^\circ$; anionic, $R^+$, $K^+$. Residues $W^\circ$, $F^\circ$, $I^\circ$, $L^\circ$, $P^\circ$ and $A^\circ$ are assigned solid squares corresponding to their relative degree of hydrophobicity. The coloured residues indicate alignment of amino acids with common properties.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>R' R' W K' K' N\circ F' I' A' V' S\circ A' A' N\circ R' F' K' K'</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK MLCK</td>
<td>R' K' W Q\circ K' T\circ G H\circ A' V' R' A' I' G R' L' S\circ S\circ</td>
</tr>
<tr>
<td>SM MLCK</td>
<td>I' N\circ L' K' A' L' A' A' A' L' A' K' K' I' L'</td>
</tr>
<tr>
<td>Mastoparan</td>
<td>K' I' W S\circ I' L' A' P' L' G T\circ T\circ L' V' K' L' V' A'</td>
</tr>
</tbody>
</table>
Introduction

Phage display: the technique
The foundation for epitope libraries had its origin from two parallel developments. Marion Geysen (1984) pioneered the concepts that (1) short peptides bearing critical binding residues (mimotopes) can chemically mimic the folded antigenic determinants on proteins (epitopes) and (2) in many cases, the non-covalent bonds formed between a few critical residues of an epitope and its binding molecule or ligate may make a major contribution to the total energy of binding. Simultaneously, Stephen Parmsley and George Smith (1988) developed a bacteriophage expression vector that could display foreign epitopes on its surface and named it the so-called “fusion phage”. This vector was used to construct a random peptide library, with each bacteriophage displaying one each of a six amino acid stretch from a repertoire of virtually all 6 mers possible to be synthesized. They also developed the method of “biopanning”, a phage display library where affinity purification of specific phage peptides binding to an epitope is possible. Thus, biologically active molecules can be selected from large populations of randomly generated sequences in an epitope library.

As a follow-up to this, Smith, McCutchan and de la Cruz postulated that it should be possible to use the recently developed expression vector and biopanning technique to identify mimotopes. This led to the idea of identifying peptide ligands for antibodies by biopanning epitope libraries, which could be used for vaccine design, epitope mapping, finding mimetics for receptors/ enzymes and/or the identification of interacting proteins. Cwirla et al (1990) and Scott et al (1990) expanded George Smith’s idea by creating a large and diverse oligonucleotide library for biopanning. One of the advantages of phage display illustrated by this study is the linkage between the peptide and the genetic information for it being provided by the phage itself. While determining the sequence of hundreds of peptides is difficult, if not impossible, the sequencing of an equal number of phage insert DNA coding for a consensus sequence is very much feasible. The capacity of this technique to correlate protein structure and function in an interactive fashion makes possible new methods of finding novel drugs with
Introduction

therapeutic potential. Phage display is also an attractive alternative to the traditional methods of antibody studies, hybridoma technology and animal immunization, enabling the building of antibodies in bacteria by mimicking characteristics of immune selection. Biopanning can be performed on plastic ELISA plates or culture dishes, glass beads, agarose beads and even on magnetic particles. While initial panning tends to select low affinity ligands, successive rounds of selection can give peptides with increasingly higher affinity (Smith et al, 1993). The success of M13 phage display has prompted the development of alternative display systems e.g., in λ phage (Santini et al, 1998), T4 phage (Ren et al, 1998) and baculoviral systems (Possee, 1997). The basic principle of a peptide library is depicted in Figure 4. Phage display describes a selection technique in which a peptide or a protein is expressed as a fusion with a coat protein of a bacteriophage, resulting in display of the fused protein on the surface of the virion, while the DNA encoding the fusion resides within the virion. This creates a physical linkage between a vast library of random peptide sequences to the DNA encoding each sequence, allowing rapid identification of peptide ligands for a variety of target molecules e.g., antibodies, enzymes, cell surface receptors, etc. The Ph.D.-12 Phage Display Library (NEB) is based on a combinatorial library of random 12 mer peptides fused to a minor coat protein (pIII) of M13 phage. The displayed 12 mer peptides are expressed at the N terminal of pIII i.e., the first residue of the mature protein is the first randomized position. The peptide is followed by a short spacer (G-G-G-S) and then the wild type pIII sequence. The library consists of \( \sim 1.9 \times 10^9 \) electroporated sequences, amplified once to yield \( \sim 20 \) copies of each sequence in 10\( \mu \)l of the supplied phage. Extensive sequencing of the naïve library has revealed a wide diversity of sequences with no obvious positional bias. The most straightforward method of biopanning involves directly coating a plastic surface with the target of interest (by non specific hydrophobic and electrostatic interactions), washing away the excess and passing the pool of phage over the target-coated surface. Depending on the target, direct coating can occasionally result in inaccessible
Figure 4: Methodology of biopanning of a 12 mer phage peptide library followed in this study.
ligand binding sites, either due to steric hindrance or partial denaturation of the target along the surface. As an alternative to directly coating the plate with the target molecule, the target can be reacted with the phage in solution, followed by affinity capture of the phage-target complexes. Depending on the target, binding in solution can result in improved kinetics compared to surface binding and can bypass problems associated with partial denaturation of the target on plastic surfaces. Affinity capture requires some sort of affinity tag on the target; this is typically accomplished by biotinylating the target and capturing the complexes with immobilized streptavidin.

**Uses of phage display library:**
Scott *et al* (1992) used a library of hexapeptides to compete with a carbohydrate ligand for binding to Con A. This highlights the versatility of peptides as ligands for a wide range of binding sites. Phage display has made possible the development of nonmacromolecular species that possess the favourable molecular recognition characteristics of antibodies but that which can be identified quickly and easily and synthesized in large amounts. Barbas *et al* (1992) pioneered the first semisynthetic or synthetic antibody as a "chemical solution to the diversity problem". This report was followed by others (Zebedee *et al*, 1992; Barbas *et al*, 1993; Lerner *et al*, 1992) who tried to combine the techniques of combinatorial libraries and phage display. It has been shown that phage-displayed peptides can be phosphorylated *in vitro* and this property has been utilized by Dente *et al* (1997) to isolate ligands for SH2 and phosphotyrosine containing peptides. Another interesting report (Matthews *et al*, 1992) detailed the utility of phage display for applications other than a binding interaction; the authors used a monovalent phage display library as a protease substrate in order to define the optimal substrates for the enzyme.

Directed evolution of a desired catalytic activity from libraries of random mutants represents a powerful route to enhancing the stability and substrate specificity of enzymes (Katz *et al*, 1997; Stemmer, 1994). Mutant library construction using
targeted random mutagenesis and/or DNA shuffling followed by colony-based plate assays has led to the discovery of enzyme variants from relatively small libraries (~10^3 - 10^4 moieties). Rapid screening of very large libraries, however, has been possible for only a few enzymatic reactions in which the desired function can be linked to a selectable phenotype. Phage display offers a possible alternative to colony-based plate screening approaches because it is amenable to screening large libraries; however, enrichment is generally based on ligand binding rather than catalytic turnover (Demartis et al., 1999). However, recently, Olsen et al. (2000) demonstrated a high throughput screening of large libraries on the basis of catalytic turnover. The authors enriched E. coli expressing the serine protease ompT from cells expressing an inactive ompT variant by over ~5000 times in a single round. By screening a library of 6 X 10^5 random ompT variants, variant proteases with catalytic activities enhanced by as much as 60 times could be isolated.

Recently, constrained peptide libraries where the randomized sequence is flanked by a pair of cysteine residues have been used for various purposes. Under non-reducing conditions, the cysteines will spontaneously form a disulphide cross talk, resulting in phage display of cyclized peptides. Such libraries have proven useful in the identification of structural epitopes (Schumacher et al., 1996), mirror-image ligands for D-amino acid targets (Wrighton et al., 1996) and peptide-based therapeutics.

**Using phage display to screen for CaM-binding peptides:**
The phage display technique has also been used in conjunction with the ubiquitous calcium sensor molecule, calmodulin in elucidating its binding motifs and structural biology. Studies using NMR, X-ray crystallography, chemical modifications of CaM, site-directed mutagenesis of CaM and also of its targets have provided us with valuable clues as to the specificity of interaction between CaM and its target molecules. Yet, keeping in mind the variability in primary amino acid sequences of CaM-binding proteins, many questions arise regarding
the detailed mechanism that allows CaM to recognize and discriminate between more than twenty target molecules. Resolution of this issue is likely to depend on characterization of specific CaM-target complexes, as well as on defining novel CaM-binding sequences and analyzing their interaction with CaM, particularly in the context of the intact target proteins. The use of random peptide libraries of large diversity represents a strategy to identify structural sequences that bind a specific protein. Dedman et al (1993), in an effort to elucidate the structural interaction of CaM with its cellular target proteins, isolated CaM binders using CaM-sepharose affinity chromatography. However, it is interesting to note that the consensus binding motifs discovered in this study ("WP" motif) has not been selected during natural selection of the CaM transduction of intracellular calcium signal. As an extension of this study, Nevalainen et al (1997) identified 10 mer peptides from a phage library which displayed specificity with respect to inhibition of CaM-dependent enzyme activity.

Limitations of the phage display technique:
One crucial aspect in display technology concerns the nature of screening experiments and the interpretation of screening results. Failure to identify an interacting ligand in a screen is essentially a negative result; however, it is ambiguous as to whether it is because of the inherent complexity of the library, the procedures used to retrieve the binders or the format of the receptor-target presentation. Even in cases where authentic binders have been identified, it is still questionable whether the clones with the most desirable properties or of physiological significance either are present in the library or have been isolated in the screen. In addition, biological display is inherently under certain selection pressure independent of those imposed by the target presentation. This negative selection pressure eliminates or reduces proteins that are toxic to the host, disruptive in folding, sensitive to proteolytic degradation, poorly translated and processing/secretion-defective (Li, 2000). Because phage displays almost always involve repeated steps of amplification, the only limited positive selections are
sequences that have advantages in the form of expression and/or host growth. However, in ideal situations, the lasting objectives are to functionally display all proteins and to minimize any selection pressure or expression bias. Improvements in display technologies in these aspects could potentially provide an essential interface between genomic proteins and physiologically significant or therapeutic biomolecules.

**Aim and Scope of the Present Study:**

In order to elucidate the molecular mechanisms of the mode of Ca\(^{2+}\) action during the pathogenesis of amebiasis, it would be worthwhile to identify and study Ca\(^{2+}\)-binding proteins involved in signal transduction pathways in *Entamoeba histolytica*. A gene encoding a novel EF-hand calcium binding protein (EhCaBP1 or CaBP1) has been cloned and characterized from *Entamoeba histolytica* (Prasad, 1993; Prasad et al., 1992, 1993). The protein has been overexpressed in *E.coli* and partially characterized. It shares some biophysical properties with CaM, such as small size, negative charge, heat stability and Ca\(^{2+}\)-dependent mobility shift in SDS-PAGE (Figure 5A). The 3D structure of Ca\(^{2+}\)-bound EhCaBP has been derived using multidimensional nuclear magnetic resonance (NMR) spectroscopic techniques (Atreya et al., 2001). The study reveals the presence of two globular domains connected by a flexible linker region spanning 8 amino acid residues (Figure 5B). Each domain consists of a pair of helix-loop-helix motifs similar to the canonical EF-hand motif of calcium-binding proteins. EhCaBP binds to four Ca\(^{2+}\) with high affinity (two in each domain), and it is structurally related to calmodulin (CaM) and troponin C (TnC) despite its low sequence homology (approximately 29%) with these proteins. Earlier, biochemical studies showed that EhCaBP is involved in a novel signal transduction mechanism, distinct from CaM (Yadava et al., 1997). A possible reason for such a functional diversity is revealed by a detailed comparison of the 3D structure of EhCaBP with that of CaM and TnC. The studies indicate a more open C-terminal domain for EhCaBP with larger water exposed total hydrophobic surface area as compared to CaM and TnC. Further dissimilarities between the structures include the presence of
Far UV CD spectra of apo (A) and holo (B) CaBP at pH 7.0 and 20°C

Calcium binding affinities of the four calcium binding domains of EhCaBP

<table>
<thead>
<tr>
<th>LIGAND</th>
<th>SITE</th>
<th>$K_b$ (M$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calcium in 0.05 M KCl @20°C</td>
<td>I</td>
<td>$(3.13 \pm 0.4) \times 10^2$</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>$(4.614 \pm 0.8) \times 10^3$</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>$(7.098 \pm 1.2) \times 10^2$</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>$(6.75 \pm 0.9) \times 10^3$</td>
</tr>
</tbody>
</table>

Figure 5A: Kinetics of calcium binding to CaBP1
Figure 5B: Ribbon structure of Ca\(^{2+}\)-CaBP1
two Gly residues (G63 and G67) in the central linker region of EhCaBP, which seem to impart it a greater flexibility compared to CaM and TnC and also play crucial role in its biological function. Thus, unlike in CaM and TnC, wherein the length and/or composition of the central linker have been found to be crucial for their function, in EhCaBP, both flexibility as well as amino acid composition is required for the function of the protein. The CaBP1 crystals belong to the hexagonal space group P6122 with unit-cell dimensions of $a = b = 96.21$, $c = 65.48$ Å. Preliminary molecular-replacement computations suggest that the structure of this protein is likely to be similar to that of calmodulin (Gopal et al., 1998). However, EhCaBP1 and CaM have been shown to be not only interacting with unique sets of *Entamoeba histolytica* proteins, but they are functionally distinct (Yadava et al., 1997). A 50 kDa CaBP1-dependent protein kinase has been shown in the *E. histolytica* lysate; in addition, several CaBP1-binding proteins have been characterized using CaBP1 affinity chromatography. Under conditions of stress, e.g., during serum starvation, CaBP1BP synthesis increases about 3 fold, suggesting that the CaBP1-dependent signal transduction pathway may be related to the growth of the parasite. In the presence of Ca$^{2+}$, EhCaM can activate cAMP-PDE whereas EhCaBP1 cannot. We have tried to characterize CaBP1 by creating deletion fragments and testing the function of each fragment. We have also cloned and partially characterized a natural variant of this protein, named EhCaBP2; these two proteins are nearly identical except in the 75-nucleotide stretch spanning the central linker domain. This study envisages functional characterization of these two proteins using different molecular techniques including phage display.