Introduction
Introduction

Preamble  

1. The Kinetoplastid Protozoa  
2. Leishmania and Leishmaniasis: An overview  
   2.1. Classification of Leishmania  
   2.2. Morphology of Leishmania donovani  
   2.3. Life cycle of Leishmania parasite  
   2.4. Host Parasite interaction  
   2.5. Evasion of the host immune response during Leishmania infection  
   2.6. Clinical spectrum of Leishmaniasis  
   2.7. Diagnostics and plausible treatment for Leishmaniasis  

3. The Era of Genomics and Proteomics In Leishmania Biology  
4. Metabolism in Leishmania species  
   4.1. Purine metabolism: At a glance  
5. Adenosine Kinase: an overview  
   5.1. Adenosine Kinase from Leishmania donovani (LdAdK)  
6. Stage specific gene expression in Leishmania  
7. Immunosuppressant: An agent that decreases activity of immune system  
   7.1. Mechanism of rapamycin action.  
   7.2. Cyclosporin A (CsA): an antiparasitic drug  
   7.3. Immunophilins  
8. Novel rotamase or Peptidyl prolyl cis-trans isomerase activity of Cyclophilin  
9. Protein Aggregation: An outline  
   9.1. Abberation in protein folding leads to protein aggregation :  
   9.2. Nucleation and growth of protein Aggregates  
   9.3. Reversibility and specificity of physical aggregation  
   9.4. Thermodynamics and kinetics of protein aggregation  
   9.5. Factors influencing Protein Aggregation  
   9.6. Morphology and types of protein aggregates  
   9.7. Inhibition of protein aggregation  
10. Physiological role of Cyclophilins  
   10.1. Protein import to the mitochondria  
   10.2. Oxidative stress  
   10.3. Cell signalling  
   10.4. Apoptosis  
   10.5. Viral replication and activation  
11. Heat shock proteins as molecular chaperones  
12. Disaggregation Leads to Reactivation of Protein  
13. Folding in the endoplasmic reticulum  
14. Chaperones in parasites  
15. Protein aggregation may inhibit clearance of misfolded protein  
16. Protein folding and aggregation in biotechnology  
17. Rationale, objective and scope of the present work
Trypanosomatid parasites, that cause various neglected human diseases, have deadly effects on several million people worldwide. Visceral leishmaniasis or kala-azar, one of the disastrous trypanosomatid diseases, is spread by *Leishmania donovani*. *L. donovani*, a purine auxotroph, incapable of synthesizing purine bases *de novo*, fulfill the requirement of purines by salvaging purines from the host using its purine salvage pathway. Adenosine kinase, a key enzyme of this salvage pathway, phosphorylates adenosine to AMP. Previous reports from this laboratory suggested that AdK has an inherent tendency to form soluble aggregates, leading to inactivation of the enzyme. Further studies showed that a single domain cycliphilin (LdCyP) from *L. donovani* disrupted the aggregates with concomitant reactivation of the enzyme. The work presented in this dissertation revealed that ADP, one of the products of AdK reaction, facilitated aggregate formation. These ADP-induced aggregates were reversible in nature and were amenable to disaggregation in presence of LdCyP. In contrast, the naturally formed AdK aggregates were irreversible and therefore could not be reactivated by LdCyP-chaperone. Detailed structural and biochemical analysis of the two forms of aggregates revealed marked differences in the morphology between two forms of aggregates. Disaggregating and reactivating properties of LdCyP remained intact even after truncating 88-amino acids from the N-terminal end of the full-length LdCyP. A further search for smaller LdCyP fragment has resulted in identification of one LdCyP-decapeptide that appeared to have AdK stimulating activity. Although, the physiological significance of this phenomenon in *L. donovani* has not yet been firmly established, it is well known that like many other developmentally regulated parasitic proteins, AdK, along with adenosine deaminase (ADA), exhibits high specific activity during transformation from promastigote to amastigote. Hence, the importance of *in vivo* regulation of AdK activity, responsible for Ado uptake in purine auxotrophs, during multiplication of amastigote within the macrophage of hosts cannot be overlooked. Furthermore, in view of previous observation, that in stressed leishmanial cells retrograde translocation of LdCyP from the endoplasmic reticulum to cytoplasm is accompanied with increased uptake of Ado, makes it relevant. Hence, development of peptide-based antagonist as a tool to control the AdK activity might be a possible strategy to inhibit *in vivo* parasite multiplication. Studies on the molecular mechanism of AdK aggregation and the mechanism by which LdCyP exerts its unique chaperone function to reactivate the activity by disrupting only the ADP-induced aggregates and a detailed structural characterization and conformational analysis of the two forms of the aggregates is therefore clearly important.
A parasite lives in a close relationship with another organism, its host and for its life functions, it is entirely dependent on its host. The effects caused by parasitic diseases range from mild discomfort to death. Infection affects approximately 740 million people in the developing countries, including children and adults, of the tropics specifically in poor rural areas located in Africa, South-East Asia, Latin America and China. Three major groups of parasitic pathogens like protozoan parasites, parasitic helminths (worms), and arthropods, cause infection and resulting physiological damage to human. Protozoan parasites are unicellular organisms that cause a wide variety of diseases like malaria, amoebiasis, filariasis, trypanosomiasis (Chagas disease), toxoplasmosis, schistosomiasis, and leishmaniasis. Since parasites generally inhabit within the blood or internal organs of the host, they have logistical problems in terms of infecting a new host. To overcome this, vector transmission is the strategy used by protozoan parasites. This strategy involves a hematophagous (i.e., blood feeding) arthropod serving as an intermediary carrier between successive vertebrate hosts. Table 1 shows the list of several human diseases that are caused by parasitic protozoa and the corresponding arthropod vectors that transmit them.

<table>
<thead>
<tr>
<th>Diseases</th>
<th>Parasites</th>
<th>Vectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>African trypanosomiasis</td>
<td>Trypanosoma brucei</td>
<td>Glossina</td>
</tr>
<tr>
<td>Chagas' disease</td>
<td>Trypanosoma cruzi</td>
<td>Triatoma, Rhodnius</td>
</tr>
<tr>
<td>Leishmaniasis</td>
<td>Leishmania sp.</td>
<td>Phlebotomus, Lutzomyia</td>
</tr>
<tr>
<td>Malaria</td>
<td>Plasmodium sp.</td>
<td>Anopheles</td>
</tr>
<tr>
<td>Babesiosis</td>
<td>Babesia sp.</td>
<td>Ixodes</td>
</tr>
</tbody>
</table>

Table 1. Diseases caused by different parasitic protozoa and the corresponding vectors.

1. The Kinetoplastid Protozoa

Parasites of the genus *Leishmania, Trypanosoma, Crithidia* belong to the order Kinetoplastida, members of which all contain a unique cellular organelle situated at the base of the flagellum, called the kinetoplast (1). The kinetoplast is a rod shaped specialized mitochondrial structure that contains the extranuclear DNA (commonly known as kDNA) (Fig. 1) that can be divided into two groups: maxicircles and minicircles. The maxicircles are found at a copy number of 20-50 per kinetoplast whereas the minicircles are present at around 10,000 per kinetoplast. Maxicircles are homogeneous circular DNA molecules, 20-35 kb in size that encodes rRNAs and proteins involved in energy production. Minicircles, in contrast, are a heterogeneous group around 0.5-1.5 kb in size whose only known function is to encode guide RNAs (gRNAs) involved in RNA editing (2-4), a process confined only to the kinetoplastids.

Figure 1: Electron micrograph of kinetoplast DNA of *Leishmania*. 
Apart from RNA editing, the kinetoplastid protozoa possess an astonishing array of unusual features that are now paradigms in higher eukaryotes e.g. trans-splicing (3), glycoprophatidyl inositol anchors (5,6) gene rearrangements during antigenic variation (7,8), use of trypanothione (a glutathione-spermidine conjugate) rather than glutathione to regenerate thiols (9,10) and lack of de novo purine synthesis pathway (11-14). All these properties certainly make the kinetoplastids ‘organism of interest’ not only for drug development but also to gain important knowledge about some important biological phenomenon.

2. *Leishmania* and Leishmaniasis: An overview

The trypanosomatid parasite of the genus *Leishmania* is the etiological agent of a variety of disease manifestations, collectively known as Leishmaniasis (15). The parasite is named after "W.B. Leishman", who developed one of the earliest strains of *Leishmania* in 1901. They have an obligate digenetic life cycle, alternating between the insect vector and the vertebrate host. The mammalian vertebrate hosts include humans, dogs and some rodents. There are over 20 species and subspecies of *Leishmania* that are transmitted by about 30 species of phlebotomine sand flies, each causing a different spectrum of symptoms (16,17). Additionally, Leishmaniasis can be transmitted directly from person to another through the sharing of needles, as is often the case among intravenous drug users.

Leishmaniasis is prevalent throughout the tropical and sub-tropical regions of Africa, Asia, the Mediterranean, Southern Europe (old world) and South and Central America (new world). Despite enormous efforts, it has been proved difficult to predict the exact scale of the impact of the Leishmaniasis on public health, since many cases go unreported or misdiagnosed. About 2 million new cases occur annually, of which only 600,000 are officially declared. In the 1990s Sudan suffered a crisis with mortality of about 100,000 among 300,000 people at risk. The endemic areas include inter-tropical and temperate countries like Africa, Brazil, Sudan, Afghanistan, Saudi Arabia and India. According to a WHO report in 2001, leishmaniasis currently threatens 350 million men, women and children in 88 countries around the world, 72 of which are developing countries. The disease claims its victims among the most impecunious members of the community. If left untreated, the fatality rate can be as high as 100%. More recently, rapidly increasing number of fatality in cases of *Leishmania*/HIV co-infection has been reported, especially in southern Europe, where 25-70% of adult visceral leishmaniasis cases are related to HIV infection, and 1.5-9.5% of AIDS cases suffer from visceral leishmaniasis.

![Figure 2: Geographical distribution of Leishmania /HIV coinfection](image-url)
from newly acquired or reactivated visceral leishmaniasis (Fig. 2). For this reason, the need to study possible remedies of these diseases has gained immense importance in recent times (WHO www site 1997).

In India, Kala-azar is the most prevalent form of the disease and has been reported to occur in Assam, Meghalaya, Bihar, West Bengal, Uttar Pradesh, Jharkhand, Delhi, Gujrat, Pondicheriy and Tamilnadu. Outbreaks and epidemics of leishmaniasis have also been associated with urban development, deforestation, environmental changes and population migrations. Individual risk factors such as HIV, malnutrition, genetic factors, etc. are also responsible for epidemiological diversity of leishmaniasis. It is worth mentioning that the geographical distribution of the disease is limited to distribution of the sandfly, its susceptibility to environmental conditions, and its capacity to support the internal development of specific species of *Leishmania*.

### 2.1. Classification of Leishmania

Systematic position of *Leishmania* as classified by Levine *et al.* 1980 (18).

- **Kingdom:** Protista
- **Sub kingdom:** Protozoa
- **Phylum:** Sarcomastigophora
- **Sub phylum:** Mastigophora
- **Class:** Zoomastigophora
- **Order:** Kinetoplastida
- **Sub order:** Trypanosomatina
- **Family:** Trypanosomatidae
- **Genus:** *Leishmania*

Historically the different species of *Leishmania* were classified according to the clinical manifestation of the disease and geographical distribution. Recently different biochemical and immunological approaches like isozyme pattern, monoclonal antibody typing, DNA buoyant density, restriction mapping of the kinetoplast DNA, hybridization with total k-DNA or mini-circle probes, chromosomal DNA markers, karyotypic profiles and variation in size and DNA sequences of the mini-exon gene non-transcribed spacers have all permitted the selection of complex species, sub species and strains (19).

### Vectors across the world:

- **Indian vector:** *Phlebotomus argentipes*
- **Mediterranean vector:** *Phlebotomus perniciosus, Phlebotomus major*
- **Chinese vector:** *Phlebotomus chinensis, Phlebotomus sergenti*
- **Brazilian vector:** *Phlebotomus longipalpis / Lutzomyia longipalpis*

### 2.2. Morphology of Leishmania donovani

These unicellular parasitic protozoa exist in two distinct phenotypical forms, extracellularly as the promastigote in the alimentary tract of the insect vector (20) and intracellularly as the amastigote in the phagolysosomes of host mononuclear macrophage (21). Promastigotes (Fig. 3, Panel A) have a spindle shaped body, of about 14-20 μm in length and 1.5-3.5 μm at their widest part and possess a single anterior flagellum. Amastigotes (Panel B) are non-flagellated and hence non-motile; spherically shaped with a diameter of approximately 2-4 μm (Fig. 3). Amastigotes are the infective form of the parasite.
while promastigotes are less virulent. Both the forms harbour a single nucleus and a typical kinetoplast-mitochondrial complex with axoneme. The average length of DNA per cell is \(9 \times 10^6\) base pairs for \(L.\) donovani. During asexual cell division by binary fission, amastigotes typically take up a “nest” like and promastigotes adopt a “rosette” like conformation.

2.3. Life cycle of Leishmania parasite

The life cycle of \(Leishmania\) spp. (Fig. 4) starts when a parasitized female sand fly takes a blood meal from a human host. As the sand fly feeds, promastigote forms of the \(Leishmania\) parasite enter the human host via the proboscis. Within the human host, the promastigote form of the parasites is ingested by macrophage where they metamorphose into amastigote form and multiplies. They increase in number until the cell eventually bursts, then infect other phagocytic cells and continue the cycle. The infected host when bitten by another female sand fly, the fly picks up the parasites during the blood meal. Once ingested, the amastigotes first breaks out of the macrophages and develop flagella thereby transforming to the flagellated promastigote. The promastigotes stick to the sand fly midgut where they multiply in an extracellular environment and eventually the non-infective promastigotes develop into infective metacyclic promastigotes by a process known as metacyclogenesis. (Metacyclogenesis is accompanied by an increased ability to infect and survive in the vertebrate host, where the parasite is

**Figure 4**: Life cycle of Leishmania species
Introduction

attacked by the host's immune system. An \textit{in vitro} change in temperature from around 33-37°C to 22-28°C, results in the morphological transformation of the amastigotes into promastigotes. In the same way, on increasing the temperature of a promastigote culture from 22-28°C to 33-37°C they become amastigote-like in morphology (22,23). These two forms have been shown to differ in their surface antigens, coat proteins, rate of respiration, level of intracellular polyamines, metabolism of adenine and adenosine and also the levels of AMP and Ado metabolising enzymes (24-29). The optimal metabolic activities of the promastigotes are carried out at a neutral pH whereas the amastigotes are comfortable in a pH range of 4.0 to 5.5.

The metacyclic promastigotes then migrate to the proboscis of the insect vector and are inoculated into the next mammalian host when the sand fly takes another blood meal. Reproduction at all stages of the lifecycle is believed to occur by binary fission. No sexual stage has been identified.

2.4. Host Parasite interaction

Very little is known as to how the parasite interacts with the macrophages and parasitizes them. On the basis of information of the earlier work, it can be summarized that the promastigotes gain entry into the macrophage via receptor mediated endocytosis which help in establishing successful intracellular parasitism in phagolysosomes or parasitophorous vacuoles of macrophages (30). Although these mononuclear phagocytes are undoubtedly the principal host cells of \textit{Leishmania}, other experimental evidence exists for leishmanial infection of other cell types, e.g. dendritic cells (31) and fibroblasts (32,33). It has been shown that \textit{Leishmania spp} bind to complement receptors (CR1/CR3) (34-36), mannose-fructose receptor (MFR) (37,38) and fibronectin receptor (39,40) on macrophage surface. Evidence indicates that complement components fixed on the infective parasite fails to mediate cytolysis, but instead help them gain entry into macrophage (35). \textit{Leishmania} species may even play an active role by changing their surface to fix complement appropriately for binding to CR1, through which endocytosis is known to provoke less respiratory burst, thereby ensuring better intracellular survival of the parasites (41-43). Non-infective promastigotes may enter into the macrophage via CR3 receptor, which activates respiratory burst to a high level, resulting presumably in the destruction of the parasites (42). The surface of the \textit{Leishmania} promastigote (Fig. 5) is dominated by two highly abundant glycoconjugates that have been found to hold the key towards its survival and infectivity. The major glycoconjugates are glycosphatidyl inositol (GPI) anchored gp63 and two distinct classes of free GPI, the polydisperser lipophosphoglycans (LPGs) and low molecular mass glycosinositolphospholipids (GPIls). The other surface exposed enzymes, such as an acid phosphatase (AP) and 3' and 5'- nucleotidase and some other minor proteins have also been found.
2.5. Evasion of the host immune response during Leishmania infection

In order to evade the host defense machinery, the Leishmania species have evolved numerous strategies. For example, L. mexicana, L. major and L. braziliensis trigger the production of transforming growth factor β (TGF-β) and interleukin 10 (IL-10), which inhibit killing of intracellular organisms (44,45). There have been reports that infection with Leishmania promastigotes inhibits production of IL-12 and TNF-α (46) which are known to be crucial for recovery from infection (47). Various group of workers have shown that the parasite causes reduction of host MHC class II molecules available for binding to parasite antigens (48-51), which presumably prevent its detection in the macrophage by T cells. Another effect of the parasite on the macrophage described recently is inhibition of apoptosis through induction of 'pro-survival' cytokines such as macrophage colony-stimulating factor (M-CSF), tumour necrosis factor α (TNF-α) and IL-6 (52-54). This phenomenon may be responsible for the persistence of amastigotes by extending the life of the host cell. All these immune evasion mechanisms mean that many, if not all, infected macrophages remain immunologically silent. This provides an explanation for the slow development of the host protective mechanism and possibly for the long-term persistence of the parasite in an individual (55,56).

2.6. Clinical spectrum of Leishmaniasis

There are four major forms of the diseases viz. visceral, post Kala-azar dermal, cutaneous and muco-cutaneous leishmaniasis.

Visceral leishmaniasis (VL) or “Kala-azar” or “Dum Dum fever”:

The causative agents of this form of leishmaniasis are the different sub-species of L. donovani. They invade the phagocyte cells in the lymph nodes, liver, spleen and bone marrow. The disease is characterized by irregular bouts of fever, substantial weight loss, swelling of the spleen and liver, anaemia, cough, burning feet, insomnia, abdominal pain, joint pain, anorexia, epistaxis, diarrhoea and serious imbalance of serum proteins. In absence of treatment, 100% fatality is not unexpected.

Post Kala-azar dermal leishmaniasis (PKDL):

PKDL usually follows recovery from VL infection and begins with small measles-like lesions appearing on the face and gradually increasing in size. Eventually the lesions spread to the upper trunk, arms, forearms, thighs, legs, abdomen, the neck and the back. Then multiple lesions can coalesce to form larger lesions and can lead to the gross enlargement of facial features such as the nose and lips, giving an appearance similar to leprosy.

Cutaneous leishmaniasis (CL) or “Delhi boli”:

Cutaneous leishmaniasis of humans is associated with the members of L. major, and L. tropica. The disease occurs clinically as acute CL, chronic CL, recurrent CL and diffuse CL. The disease normally produces skin ulcers on the exposed parts of the body such as the face, arms and legs at the points where the parasites are introduced by the bite of the sand fly. It can produce a large number of lesions—sometimes up to 200, causing serious disability and invariably leaving the patient permanently scarred.
Muco-cutaenous leishmaniasis (MCL) or “Espundia”:

In mucocutaneous forms of leishmaniasis, lesions can lead to partial or total destruction of the mucosa membranes of the nose, mouth and throat cavities and surrounding tissues. It is initiated in the same manner as the cutaneous form and then metastasizes from skin lesion to the nasopharyngeal tissues and sometimes the buccal cavity, pharynx and larynx. The lesions produced are frequently very extensive and long lasting. The causative organism is mainly *L. braziliensis*.

2.7. Diagnostics and plausible treatment for Leishmaniasis

Diagnosis of the leishmanial diseases relies mainly on the detection of amastigotes by microscopy of spleen, bone marrow or lymph node aspirates. Physical exam may show signs of an enlarged spleen, liver, and lymph nodes. The patient may have been bitten by sandflies, or was in an area known for leishmaniasis.

Tests that may be done to diagnose the Visceral Leishmaniasis include:

- Biopsy of the spleen and culture
- Bone marrow biopsy and culture
- Direct agglutination assay
- Indirect immunofluorescent antibody test
- ELISA
- *Leishmania*-specific PCR test
- Liver biopsy and culture
- Lymphnode biopsy and culture
- Montenegro skin test
- Skin biopsy and culture
- Complete blood count
- Serologic testing
- Serum albumin
- Serum immunoglobulin

*Leishmania* infection can be prevented by avoidance of sandfly bites through use of fly-repellents. Vector control for leishmaniasis, based on spraying with residual insecticides can be effective where transmission occurs in and around the home. Animal reservoir control is based on identifying seropositive animals (e.g. stray dogs) and eliminating them either by shooting or by using poisoned baits. But the main control strategy of leishmaniasis involves case finding through proper diagnosis and treatment of the disease.

The first effective drug against leishmaniasis was urea stibamine discovered in 1912 by Prof. U. N. Brahmchari. This discovery saved millions of lives of poor Indians, for which Prof. Brahmchari was nominated for the Nobel Prize in 1929 (www.nobelprize.org). Later on, refinement and development of the pentavalent antimonials [Sb (v)] reduced its side effects and systemic or local use of these compounds is still the mainstay of treating all forms of leishmaniasis. The first line compounds includes
two pentavalent antimonials, sodium stibogluconate (Pentostam) and meglumine antimoniate (Glucantime). Failures and relapses occur in all forms of leishmaniasis hence in cases where these drugs are not effective, extremely toxic second line compounds e.g. pentamidine (Lomidine) and amphotericine B (Fungizone) are used. The emergence of drug resistance in Leishmania parasites is a major obstacle to their control and with no real promising prospects of an effective vaccine in the near future, there is an urgent need to develop new and better drugs. Amongst the new drugs that are currently under clinical evaluation are Ambisome (a formulation of amphotericine B in liposomes), Paromomycin (aminosidine), Ketoconazole (an inhibitor of synthesis of the membrane sterol) and Allopurinol. The latter function as an alternative substrate for the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRTase), thereby allowing the incorporation of allopurinol riboside into RNA, which leads to the inhibition of protein synthesis in the parasite (57). Two other pyrazolopyrimidines, thiopurinol and aminopurinol have also been shown to have antileishmanial activity. Aciclovir, a chlorinated amino acid antibiotic, is found to be remarkably effective in killing both the vector and the host form of the parasitic protozoa, Leishmania donovani (58). The drug irreversibly inactivates both in vitro and in vivo carbamoyl phosphate synthetase II, the first enzyme of the pyrimidine biosynthetic pathway and is thus a candidate for potential chemotherapy against increasing number of Kala-azar cases that are reported to be unresponsive to pentavalent antimonials. Also, the possibilities of using Sinefungin, a nucleoside antibiotic isolated from Streptomyces sp. as an inhibitor of

<table>
<thead>
<tr>
<th>DRUG</th>
<th>ORGANISM</th>
<th>TARGET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chloroquine</td>
<td>Plasmodium falciparum</td>
<td>Heme polymerase</td>
</tr>
<tr>
<td>Triclosan</td>
<td>Plasmodium falciparum</td>
<td>Enoyl ACP reductase</td>
</tr>
<tr>
<td>Ketoconazole</td>
<td>Trypanosoma cruzi</td>
<td>Enzymes in sterol biosynthesis</td>
</tr>
<tr>
<td></td>
<td>Trypanosoma brucei</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leishmania major</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leishmania amazonensis</td>
<td></td>
</tr>
<tr>
<td>Staurosporine</td>
<td>Leishmania spp.</td>
<td>Protein kinases in promastigote stage</td>
</tr>
<tr>
<td>Arsenicals like Melarsoprol</td>
<td>Trypanosoma spp.</td>
<td></td>
</tr>
<tr>
<td>DL-α difluoromethyl ornithine (DFMO)</td>
<td>Trypanosoma spp.</td>
<td></td>
</tr>
<tr>
<td>Deoxyadenosine analogues</td>
<td>Trypanosoma spp.</td>
<td></td>
</tr>
<tr>
<td>Analogues of hypoxanthine and inosine and allopurinol Mycophenolic acids</td>
<td>Trypanosoma spp.</td>
<td>s-adenosyl methionine (SAM)</td>
</tr>
<tr>
<td></td>
<td>Leishmania spp.</td>
<td>Hypoxanthine guanine phosphoribosyl transferase (HGPRT)</td>
</tr>
<tr>
<td>Amphotericin B</td>
<td>Trypanosoma cruzi</td>
<td>IMP dehydrogenase</td>
</tr>
<tr>
<td></td>
<td>Trypanosoma brucei</td>
<td>Adenosine Kinase</td>
</tr>
<tr>
<td></td>
<td>Toxoplasma gondii</td>
<td>Cell membrane</td>
</tr>
<tr>
<td></td>
<td>Leishmania spp.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Different Antiparasitic Drugs with Their Intracellular Targets.

the salvage pathway of the parasites is being investigated (59). Miltefosine, an ether-lipid analogue that blocks the ether-lipid metabolism in Leishmania has been reported to be the first and only orally
applicable antileishmanial compound so far (60). A promising approach towards improved chemotheraphy of leishmaniasis is the use to liposomes. Antimonials, pentamidine and amphotericin B encapsulated in liposomes are found to be more effective than free drug for the treatment of leishmaniasis (61,62).

A major advance has been the advent of liposomal formulations of amphotericin B, in which various alternative lipids have replaced deoxycholate. These formulations, which passively target amphotericin to macrophage-rich organs, are much more costly than conventional amphotericin B (making them cost-prohibitive in poor countries) but are associated with less nephrotoxicity and can be given in considerably shorter courses. Although visceral leishmaniasis is traditionally treated with multiple doses of amphotericin B deoxycholate, it appears, based on a single randomized trial, that a single dose of liposomal amphotericin B may be just as effective and cheaper.

Other parenteral alternatives that have merit include amphotericin B (not only in deoxycholate form but also in liposomal forms) and have generally replaced pentamidine. Miltefosine, a chemotherapeutic agent, is the first extremely effective oral agent for visceral leishmaniasis but is not currently available in the United States. Injectable paromycin has also been reported to be noninferior to amphotericin B but also is not currently FDA approved.

3. The Era of Genomics and Proteomics In Leishmania Biology

The determination of complete nucleotide sequence of an organism has always captured the imagination of the researchers, since it was expected to reveal the "keys of life". With this objective, scientists from the developing and developed countries, during 1993-1994, planned and initiated a number of parasite genome projects for mapping and sequencing of these medium-sized genomes. Thus unlike the situation only a few years ago, a large number of parasite-sequencing projects (e.g. Leishmania major, Plasmodium falciparum, Schistosoma mansoni, Trypanosoma cruzi, Trypanosoma brucei etc.) are currently underway generating an inundation of sequence data. The genomes of these parasites are around 30 - 40 Mb (63) and the areas of current activity is massively influenced by the very rapid accumulation of genome sequence information. The types of sequence data available, fall into three basic categories: complete or nearly complete genomic sequences, which are normally released as 'contigs' of overlapping sequence reads; genome-survey sequence (GSS) tags, which are generated by skimming genomic sequences (either random clones or bacterial artificial chromosome [BAC] ends); and expressed sequence tags (ESTs), which are generated from mRNAs expressed during one or more stages of the parasite life-cycle.

Very recently, the complete genome sequence for three trypanosomatid parasites (referred as "Tritryps") e.g. Leishmania major, Trypanosoma brucei, and Trypanosoma cruzi have been reported in a
Introduction

The 32.8-megabase haploid genome of Leishmania major (Friedlin strain) were predicted to contain 911 RNA genes, 39 pseudogenes, and 8272 protein-coding genes, 36% of which were ascribed a putative function (67). The T. brucei and T. cruzi genome encode about 8100 and 12000 protein-coding genes respectively. In all these parasite genomes, the organization of protein-coding genes into long, strand-specific, polycistronic clusters and lack of general transcription factors suggested that the mechanisms regulating RNA polymerase II-directed transcription are distinct from those operating in other eukaryotes. A comparison of gene content and genome architecture of T. brucei, T. cruzi, and L. major, three related pathogens with different life cycles and disease pathology, revealed a conserved core proteome of about 6200 genes in large polycistronic gene clusters (Fig. 6). However, 910 L. major genes have no orthologs in the other two Trityp genomes. These “Leishmania-restricted” genes, especially large surface antigen families, occur at chromosome-internal and subtelomeric regions.

Thus, with large amount of sequence information already available in the database, the major challenge now is to develop experimental methods for efficiently searching within this mass of information for genes that play crucial roles for the viability of the parasite. This task and its newly arisen field are often referred to as ‘functional genomics’. Molecular-genetic analysis of parasites has rapidly developed over the past decade and there is a reasonable collection of approaches available for post-genomic analysis of parasite genomes e.g. gene-knockouts, antisense RNA, RNA interference (RNAi), conditional expression etc. Apart from these, the emergence of proteomic technologies have allowed proteome-wide analyses of the Leishmania species in contrast to traditional biochemical analysis of single proteins (68). The application of these array of technologies will not only be important for academic studies of the parasite functions but will also be critical for assessment of global patterns of gene expression, particularly those occurring during stages of parasite life cycles that are refractory to normal experimental interrogation. Such a view might be key towards identification of likely drug and vaccine candidates and development of target-validation studies in drug-therapy initiatives. Hence it will be no exaggeration to mention that the success of Leishmania genome project represents one of the most important developments in our knowledge of the parasites. However, translation of this wealth of information into an understanding of the parasite biology and then on to drugs, vaccines and other healthcare developments for the diseases will need careful handling of these sequencing data and only then will the activity associated with post-genomics be turned from genome babble to real opportunities in gaining deeper insights into the biology of this parasite.

4. Metabolism in Leishmania species

The production of energy, in usable form, is a basic requirement of all organisms and the parasites are no exception to this. But kinetoplastid protozoa that probably diverged early in evolution from the higher eukaryotes are characterized by a number of unique features particularly with respect to their energy and carbohydrate metabolism. Among them the most peculiar aspect is the sequestration of a part of the glycolytic pathway and some other enzyme systems within peroxisome-like organelles called glycosomes (69-72). Owing to this distinctiveness, the metabolic pathways of the trypanosomatid protozoa have been the topic of enormous interest for the past several years in order to gain deeper insights into the parasite biology. In addition to this, comparative biochemistry between the host and the pathogen being an established route to identification of drug target, the novel biochemical pathways of the trypanosomatid parasites are explored with added endeavor.
In general, parasites are metabolically lazy and rely on the metabolism of the host both for a supply of prefabricated components such as purines, fatty acids, sterols and amino acids and for the removal of end products. Nonetheless, the parasite metabolism is considered to be highly sophisticated which allows them to adapt to the environments that their life cycles demand by retaining genetic capacity to induce many pathways, as and when needed.

4.1. Purine metabolism: At a Glance

Purines are of vital importance to all living organisms. They are essential for the synthesis of nucleic acids, proteins, and other metabolites as well as for energy requiring reactions. In view of their high rate of replication or overproduction, parasites require very active nucleic acid synthesis, necessitating large supplies of the indispensable purine nucleotides. Investigations on purine metabolism in *L. donovani* have shown that the parasite cannot synthesize purines de novo (12,13,73), though they have the machinery for de novo pyrimidine biosynthesis. Thus they have evolved a unique series of purine salvage enzymes for their survival. These salvage pathways are, therefore, important to parasite survival and hence could be prospective targets for selective drug designing. The purines are transported from host to the parasite through cell surface transporters that are indispensable for these purine auxotrophs. The presence of specific cell surface nucleobase/nucleoside transporters has been well documented for *Leishmania* species. In *L. donovani* mainly two types of nucleoside transporters (e.g. LdNT1 and LdNT2) have been reported. They mediate the uptake of purine nucleosides as well as some purine analogs but with differential specificities (Fig. 7). Since these transporters take up only nucleosides or nucleobases, the host nucleotides have to be converted into the respective nucleosides prior to uptake. This task is accomplished by unique cell surface 3'-nucleotidase/nucleases (74-77).

Mammalian cells salvage purines primarily by two pathways. The first provided by the enzymes of phosphoribosyl transferase family that transfers ribose and phosphate moieties to purines in one step:

\[
\text{Base} + \text{PP-ribose-P} \rightarrow \text{Base-ribose-P} + \text{PP}_i
\]

The second pathway comprises of two enzymes where, a phosphorylase first attaches a ribose to purines and then a nucleoside kinase attaches phosphate group to the nucleoside:

\[
\text{Base} + \text{ribose-1-P} \rightarrow \text{Base-ribose} + \text{P}_i
\]

\[
\text{Base-ribose} + \text{ATP} \rightarrow \text{Base-ribose-P} + \text{ADP}
\]

But the purine salvage pathways of *Leishmania species* (summarized in Fig. 8) are more complex and are different from those used by the mammalian cells. Owing to the absence of phosphorylase activity in some *Leishmania* species, it is believed that these organisms can salvage purines by means of phosphoribosyl transferase reaction and hence preferentially salvage purine bases instead of nucleosides (19). The key enzymes involved in this process are adenine deaminase and guanine deaminase, which convert adenine and guanine to hypoxanthine and xanthine respectively. The phosphoribosyl transferases of the parasite then convert hypoxanthine and xanthine to inosine.
monophosphate (IMP) and xanthine monophosphate (XMP) respectively. However, stage-specific differences exist for some *Leishmania* species, with promastigotes containing adenine deaminase and amastigotes containing adenosine deaminase, which converts adenosine to inosine that is subsequently converted to IMP by inosine phosphorylase. IMP and XMP can be converted to AMP and GMP by the following enzymes: adenylosuccinate synthase, adenylosuccinate lyase, IMP dehydrogenase and GMP synthase. In both promastigotes and amastigotes adenosine is also directly phosphorylated to AMP by the enzyme adenosine kinase but interestingly, the activity of the enzyme is fifty-fold higher in amastigotes than that reported for promastigotes (25). This enzyme has been isolated in pure form from *L. donovani* (78). Four other kinases e.g. inosine, xanthosine, AMP and GMP kinases have also been detected in low amounts.

Thus it appears that the *Leishmania* parasite possesses multiple routes for salvaging the purines, however, all the purine rings are interconvertible with an apparent branch point at IMP. Hence, unlike some other protozoa, the *Leishmania* species when culture in vitro, have no specific requirement for any particular purine bases.

**5. Adenosine Kinase: an overview**

\[
\text{Adenosine + ATP } \xrightarrow{\text{Mg}^{2+}} \text{AMP + ADP}
\]

AdK (ATP: adenosine 5’-phosphotransferase, EC.2.7.1.20), the most predominant nucleoside kinase presents in animal tissues, catalyses the phosphorylation of adenosine to its monophosphate in presence of ATP and divalent magnesium.

AdK was first identified independently by Caputto and Kornberg *et al.* in yeast and rabbit muscle (79,80). This was followed by extensive characterization of the enzyme from a number of sources that includes higher eukaryotes (81-83), parasitic protozoa (78,84-86), plant and bacteria (87,88) where the molecular weight of the enzyme varied between 23 to 56 kDa. In mammals, AdK has broad tissue distribution (89-91) where it plays important role in regulation of extracellular Ado levels and maintenance of intracellular adenylate pools (92). Decreased level of AdK causes resistance to anti-proliferative effects of a number of Ado analogs (93), increased purine excretion (94) and primary gout (95). In addition to this, by virtue of its broad substrate specificity (78,96,97), the enzyme is responsible for phosphorylation of numerous therapeutically useful purine nucleoside analogs (98,99). Several
nucleoside antiviral and anticancer drugs are AdK substrates and under \textit{in vivo} condition they undergo rapid phosphorylation to the 5'-monophosphate. In many cases, the monophosphate is subsequently converted by other kinases to the triphosphate which functions as the active metabolite. Examples include ribavirin \cite{100} and mizoribine \cite{101}.

Ado is the best-known substrate for the enzyme having $K_m$ values in the range of 0.2 to 20 $\mu$M. Besides the natural substrate, AdK also phosphorylate several other Ado analogs and is relatively non-specific with regard to its base moiety of ribonucleosides. The best substrates analogs are 8-azaadenosine, toyocamycin, tubercidin and formycin A \cite{78,99}. Inosine was also shown to be phosphorylated by rabbit liver AdK, although the substrate efficiency is several times lower than that of Ado \cite{99}. In contrast, the enzyme is comparatively specific for the ribosyl moiety. 2'-deoxyadenosine and arabinosyladenine are extremely poor substrates with substrate efficiencies of $10^{-4}$ to $10^{-6}$ that of adenosine \cite{99}. A pharmacological study with \textit{T. gondii} AdK also demonstrated that the hydroxyl groups at the 2' and 3' positions of adenosine are crucial for the enzyme-substrate interaction \cite{102}. Basically, these results support the suggestion of Bennett and Hill, that a 2'-hydroxyl group trans to the glycoside linkage is a prerequisite for appreciable binding to the enzyme \cite{103}. None of the pyrimidine ribonucleosides tested were substrates or inhibitors \cite{99}. As far as the phosphate donor is concerned, AdK has broad nucleotide triphosphate specificity but ATP is the most efficient phosphate donor. Like other kinases, the activity of the enzyme has been found to depend on a bivalent cation. Mg$^{++}$ is by far the best cation used and in most cases ATP-Mg$^{++}$ complex acts as the true substrate and the pH optima of the enzymatic reaction is a function of the concentrations of ATP and Mg$^{++}$ \cite{81}.

Regarding the mechanism of AdK catalyzed reaction there had been substantial debate among various groups. Using 6-methylmercaptopurine riboside as the substrate, Henderson \textit{et al}. claimed that the reaction carried out by the enzyme from Ehrlich ascites tumor cells proceeds by an ordered sequential mechanism in which ATP is the first substrate to bind and 6-methylmercaptopurine riboside monophosphate the last product to be released \cite{104}. Whereas, subsequent results from most other laboratories were consistent with a reaction sequence in which Ado is the first substrate to bind and AMP the last product to be released \cite{81,83,105}. Mechanistic studies showing net stereochemical inversion of the y-phosphate group suggested a direct in-line transfer of the phosphoryl group from ATP to Ado as opposed to a double displacement mechanism involving a phosphoenzyme intermediate \cite{106}. This was further established by the inhibition studies with the bi-substrate analogs of AdK \cite{107}.

\subsection{5.1. Adenosine Kinase from \textit{Leishmania donovani} (LdAdK)}

AdK from parasitic species have not been well characterized due to the lack of consensus in the biochemical as well as biophysical data available from various sources \cite{91,97,108,109} as well as unavailability of crystal structures. In parasitic protozoa like \textit{Leishmania donovani}, this crucial purine salvage enzyme becomes additionally important as these organisms lack the ability to synthesize purine nucleotides \textit{de novo} and hence depend entirely on their host for the salvage of these metabolites \cite{11}. Studies showed that unlike other higher eukaryotic AdKs, the parasitic enzyme possesses biochemical and immunological characteristics unique of its own \cite{78}. The salient features of the findings related to LdAdK are summarized below:

1. The enzyme is a 345-residue monomer of 38 KDa with $p_I$ of 8.8, which is sharply different from the $p_I$ (4.5-5.9) determined for AdK from other sources \cite{97,110}. The enzyme has a $p_I$ of 7.5 and the activity is dependent upon the optimum ATP-Mg$^{++}$ ratio. Studies showed that while the higher eukaryotic AdKs are
Introduction

prone to inhibition at high Ado and Mg$^{2+}$ concentrations (81,83), LdAdK is refractory to such inhibition. In contrast LdAdK is much more sensitive to inhibition by ATP (105). Apart from these, the enzyme is regulated by both of its products i.e. AMP and ADP (105).

2. The leishmanial AdK is also immunologically distinct from those of other sources. Polyclonal antibody raised against LdAdK did not cross react well with AdK from hamster, rat, rabbit and human (84).

3. Under a defined set of conditions, the enzyme exhibited an apparent $K_m$ of 16 and 50μM for Ado and ATP respectively. Of the nucleoside triphosphates tested, ATP and GTP were found to be the most efficient phosphate donors. The reaction proceeds through a sequential Bi-Bi mechanism, with Ado being the first substrate to bind to the enzyme and AMP the last product to be released (105).

4. Cytotoxic analogs of Ado e.g. tubercidin, 6-methylmercaptopurine riboside, formycin A, formycin B were found to inhibit parasite replication in vitro (111). Biochemical studies however indicate that only the first three analogs are substrate for the enzyme while formycin B is neither a substrate nor an inhibitor of the enzyme. Substrate competition experiments have revealed that tubercidin and 6-methylmercaptopurine riboside inhibit Ado phosphorylation in a competitive way and formycin A is a non-competitive inhibitor (105). Further studies have suggested that the enzyme possibly has two discrete catalytically active nucleoside interacting sites.

5. The information on the active site architecture of LdAdK is based on extensive chemical modification studies (112-114). With the use of group-specific chemical modifiers, it was demonstrated that the leishmanial enzyme, in contrast to analogous enzyme from other sources, harbors two conformationally vicinal cysteines (112-114) and one arginine residue (113) at or near the active site, modification of which led to substantial loss of catalytic activity. Fluorimetric studies have confirmed the presence of one tryptophan residue at or near the substrate binding sites of the enzyme. The studies further indicated the presence of a single Ado binding site in LdAdK as opposed to two, found in hamster and higher eukaryotic AdK (113).

6. By extensive biochemical and biophysical analysis of several site-specific mutants coupled with through structural study, established the functional role of R69, R131 and D299 in catalysis. This extensive work on LdAdK biochemistry have led us to develop a comprehensive model describing the phosphate transfer mechanism of LdAdK (115).

7. Further biochemical studies with site-specific mutagenesis revealed that though Ado and AMP occupy a virtually overlapping site of the enzyme, their mode of interaction with the enzyme are different. AMP binds through R131 whereas Ado binding is independent of this residue rather it binds to the D16. In addition to this a general mechanism has been proposed that can explain how AdK activity is regulated and the dual role of the R131 both in catalysis as well as in regulation (116).

8. During kinetic analysis, it was observed that purified LdAdK ceases to follow linear kinetics at the later part of the reaction. Investigations revealed that LdAdK has a tendency to aggregate thus causing inactivation. Interestingly, a cyclophilin from the same source (LdCyP) was found to disaggregate and reactivate LdAdK in an isomerase-independent fashion in vitro (117). This unique behavior was also observed in a simulated in vivo condition using AdK deficient E. coli strains (118).

6. Stage specific gene expression in Leishmania

The ability to undergo cellular differentiation is mandatory for the survival and pathogenicity of organisms. The Leishmanial promastigote, when transmitted from the sand fly to a mammalian host, experiences a rapid temperature elevation from 22-28°C to 35-37°C. This results in increased synthesis
of HSPs 100, 83 and 70. The Hsp 83 transcript rapidly degrades at 26°C while at 35°C it becomes more stable and is translated more efficiently (119). Hsp 100 is required in the mammalian stage of the parasite and is not expressed in the promastigote stage (120). This induction of translation is postulated to be coupled with the increased mRNA stability. Similarly transcripts of the A2 gene family are abundant in amastigote but rarely detectable in promastigotes (121).

Other genes which are solely expressed at the infective stage of the parasite include amastin (122), histone H1 (123), cph cysteine proteinases (124), P4 antigen (125) etc. β-tubulin synthesis increases 10 fold during differentiation from amastigote to promastigote with no change in the overall level of tubulin mRNAs suggesting a post transcriptional regulation (126). In L chagasi GP63 and GP46 increases to several fold upon transformation into infective form (127,128). According to the mechanism proposed for gene regulation in kinetoplastids, the 3'UTR sequences modulates transcript stability and translation efficiency either by forming a secondary structure or through trans-acting protein factors that bind to AU-rich sequence found in 3'UTR (129-131). However, identification of factors that regulate mRNA stability is yet to be identified. In addition to the 3'UTR sequences, the maturation of mRNA involving trans-splicing and polyadenylation, is also vital for the stability of mRNA and its subsequent expression.

7. Immunosuppressant: An agent that decreases activity of immune system

When an organ, such as a liver, a heart or a kidney, is transplanted from one person (the donor) into another (the recipient), the immune system of the recipient triggers the same response against the new organ it would have to any foreign material, setting off a chain of events that can damage the transplanted organ. This process is called rejection and it can occur rapidly (acute rejection), or over a long period of time (chronic rejection). Rejection can occur despite close matching of the donated organ and the transplant patient. Immunosuppressive drugs greatly decrease the risks of rejection, protecting the new organ and preserving its function. These drugs act by blocking the immune system so that it is less likely to react against the transplanted organ. A wide variety of drugs are available to achieve this aim but work in different ways to reduce the risk of rejection.

In addition to CsA, which is one of the most common immunosuppressive drug (discussed in subsequent section), FKS06, Azathioprine, Corticosteroids such as prednisolone and rapamycin are also used as drugs of choice to prevent organ / graft rejection in patients who have undergone transplant operations. Similar to CsA, FKS06 also inhibits the function of T lymphocytes by blocking signal transduction cascades required for T cell activation (132). FKS06 and rapamycin are all natural products produced by soil microorganism (133). Their relevant roles in nature may be reflected in their potent toxicity towards microorganism including fungi and parasites.

7.1. Mechanism of rapamycin action.

Rapamycin binds more strongly to FKBP compared to FK 506, although it has different immunosuppressive mechanism. In IL2 stimulated T-cells, rapamycin encumber cell cycle progression through G1/S transition of the proliferating cycle, resulting in "mid-to-late" G1 arrest. Two major biochemical alterations underlie this mode of action. First one affects the phosphorylation as well as activation of p70S6 kinase (p70S6K)—an early event of cytokine induced mitogenic response. By inhibiting this enzyme, whose major substrate is the 40S ribosomal subunit S6 protein, reduces the
translation of certain mRNA encoding ribosomal protein and elongation factor for protein synthesis, thereby decreasing protein synthesis. Second later effect of rapamycin in IL2 stimulated T-cells is an inhibition of enzymatic activity of the cyclin dependent kinase cdk2-cyclin E complex, which functions as a crucial regulator of G1/S transition. This inhibition results from a prevention of the decline of p27 cdk inhibitor that normally follows IL2 stimulation. To mediate these biochemical activities rapamycin needs to bind to FKBP, thereby forming a unique effector-molecular complex. However this rapamycin-FKBP complex directly causes neither p70S6K inhibition nor p27 induced cdk2-cyclin E inhibition unless and until this complex physically interacts with novel protein "mammalian target of rapamycin or mTOR", which has a sequence similarity with the catalytic domain of phosphatidylinositol kinase (134).

7.2. Cyclosporin A (CsA): an antiparasitic drug

CsA is a lipophilic, neutral, 11-amino acid cyclic peptide, isolated from the fungus *Tolypocladium inflatum*. It was discovered in 1973 and found to have anti-fungal effect (135). Later on, CsA was found to depress the antibody production in mice (136). In 1978, it was first used in human in a kidney transplant operation to prevent organ rejection (137). Initial work on the biochemical mechanism of CsA showed that it blocks the activation of T cells (138). An important advancement was the isolation of a tight binding protein receptor for CsA, from human spleen and bovine thymus cytosol. The 18-kDa receptor was named as cyclophilin (CyP)(139). The CsA-CyP complexes were then used to probe cellular extracts. The common target was found to be serine/threonine phosphatase, calcineurine (CN). CN is a heterodimer constituting a 59 kDa catalytic “A” subunit and a 19 kDa regulatory “B” subunit (140,141). The specific target for CN in the signal transduction cascade has not yet been properly characterized, but the favoured hypothesis suggests that dephosphorylation of the cytosolic component of the transcription factor NF-AT (nuclear factor for the activation of the T-cells) is required before it can enter into the nucleus (142,143). The CsA-CyP complex binds to CN and prevents its phosphorylation ability of NF-AT.

As a result of its potent T-cell immunosuppressive properties, CsA became an important tool for immunologists interested in dissecting immune response following infection. When applied with the same purpose to the field of immunoparasitology, the drug was unexpectedly found to possess anti-parasitic activity. Studies done on *Schistosoma spp* and *Plasmodium spp* infections (144,145) shown that this drug eliminated the parasites and / or extended the survival of the host. Since this initial discovery, CsA has been found to affect adversely a large number of protozoa and helminths both in vitro and in vivo. However, not all protozoa are susceptible, for example, *Brugia malayi* was found to be resistant to CsA (146). Structural damages has also been described in the nematode *Litomosoides carinii* as well as *Fasciola hepatica* (147,148). CsA binding CyPs have been demonstrated in *Schistosoma mansoni*, *Plasmodium falciparum* and *Toxoplasma gondii* (149). In *Leishmania spp* cutaneous lesions in susceptible mice caused by infection of *L. major* and *L. tropica* could be prevented with the administration of CsA or its non-immunosuppressive analogues. In contrast, distinctly exacerbated intracellular *L. donovani* (amastigote) parasite load in murine cells, suggesting possible differences in the status of CyPs in these two species of *Leishmania* (150). It has been shown that the species-specific toxicity of CsA in parasites is related to the presence of different CyP isoforms that bind CsA to varying degrees.
7.3. Immunophilins

Cellular targets, which bind immunosuppressants, are the immunophilins. Among these there are three major families of PPIase namely cyclophilin (CyP), FK506 binding protein (FKBP) and pervulins. Two such families of immunophilins are the CyPs, which bind to CsA, and the FKBPs that bind FK506 and rapamycin, whereas pervulins do not bind any known immunosuppressant although it contains PPIase domain despite their being very little sequence homology between these three families, they do share structural similarities that led to their conserve action. As previously discussed, these mediate immunosuppression by binding with their conjugate immunosuppressant drugs. Another interesting feature about CyP and FKBP is that both possess peptidyl-prolyl cis-trans isomerase activity (PPIase) or rotamase activity. This activity allows them to convert cis prolyl bonds into their trans conformer, which thought to be the rate-limiting step in protein folding. In addition, the immunophilins bind peptides and proteins and form coregulatory subunits of various molecular complexes (151). This multifunctionality places immunophilins at the crossroads of folding, assembly and trafficking of proteins in the cellular milieu, coregulation of molecular complexes and immunomodulation. However, it is a great enigma as to why nature utilizes two series of sequence unrelated PPIases, however, even more strangely, the PPIase activity of both the immunophilins are inhibited on binding with their conjugate immunosuppressant drugs (152,153). In this context, CyP consist of the largest group of PPIase exist as a multigenic family of proteins in eukaryotes. Immunosuppression is independent of of the PPIase activity of immunophilins as analogs of FK506 and CsA that inhibit the PPIase activity, fails to show any immunosuppressive effect (154,155). Six different isoforms of CyP have been described in human (132). These proteins bind to the immunosuppressive agent via a central highly conserved CsA-binding domain (156).

Based on their intracellular location and unique structural features, the cyclophilins have been divided into the following classes:

1. **Cyclophilin A (CyP A):** It is the most abundant isoform present in the cytosol of most cells. This isoform is the one that has the highest affinity for binding to CsA. The bound CsA-CyPA complex then, inhibits the function of CN. These are all usually around 18 kDa proteins. The level of CyP A in a particular cell depend upon the state of the cell. Such as the level of CyP A increased in proliferating MRC-5 cell, while the level decreased in SV 40 infected MRC-5 cells. A higher level of CyP A was observed in neoplastic tissue (157).

2. **Cyclophilin B (CyP B):** This isoform occurs in the endoplasmic reticulum of various cells and contains a hydrophobic signal sequence at the N terminus. On the C terminus part, some of the CyP B species contain endoplasmic reticulum retention signals while others have hydrophobic polypeptide sequence which may serve as anchors to membranes. *In vivo*, its amount is lesser than CyP A. Moreover, due to its non-cytosolic localisation, it is not accessible for binding to CsA although *in vitro* CyP B also shows high affinity for binding to the drug (158).

3. **Cyclophilin C (CyP C):** CyP C contains an endoplasmic reticulum signal sequence and was cloned from a mouse library (159).

4. **Cyclophilin D (CyP D):** This contains a mitochondrial signal sequence and was found in a human library (160).

5. **Cyclophilin 40 (CyP 40):** A 40 kDa protein, forms a component of the inactivated estrogen receptor complex along with the steroid receptor and HSP90 (161).
6. Cyclophilin S (CyP S): This form was isolated from human milk and its sequence was found in a chick embryo cDNA library. Later on it was shown to be a post transcriptionally modified form of CyP B with the signal sequence removed (162).

FK506 binding proteins:
FKBP isoforms belong to two general categories, small FKBP12s and large FKBP51s. The small FKBP isoforms, FKBP12, is involved in immunosuppression by inhibiting T-cell activation. In absence of FK506 and rapamycin, FKBP12 associates with receptors such as TGF-β type II or calcium channels such as ryanodine receptor (RyR) or the inositol-1,4,5-trisphosphate receptor (IP3R) (163). The IP3R is activated by phosphorylation with protein kinase A and inactivated by dephosphorylation with calcineurin (163).

Large FKBP12s have one or more FKBP12-like domains, a tetratricopeptide repeat (TPR) domain and a C-terminal domain that possibly binds calmodulin. FKBP51 and FKBP52 are the most well characterized large FKBP12s and both these proteins associate with hsp90 through their TPR domain in the native steroid receptor complex. Individually, FKBP52 has been shown to possess a chaperonic activity in vitro (164). Furthermore, FKBP52 associates with dynein-a motor protein that regulates movement along microtubules through its PPIase domain. It is suggested that the interaction of the chaperone complex with dynein through FKBP52 may allow steroid receptor (glucocorticoid receptor) to translocate along the cytoskeleton and to the nucleus (165,166).

Parvulins:
They do not bind to any known immunosuppressants but consist of one or two PPIase domains often accompanied by N- and C-terminal extensions. A group of parvulins were found to be involved in maturation, export or assembly of specific proteins with phosphorylated serine or threonine residues preceding proline (167,168). These form the subfamily of eukaryotic phosphate-specific parvulins that additionally contain an N-terminal WW-domain responsible for protein-protein interaction and were shown to be involved in cell-cycle regulation at the DNA replication check-point. The phosphate-specific parvulins are the only known PPIase with substrate specificity for phosphorylated serine or threonine side chain preceding proline. A parvulin Pin-1 was shown to interact with NIMA kinase and help in cell-cycle progression (153). CyP A in combination with a parvulin, Ess1, interacts with and regulates silencing of genes by Sin3-Rpd3 histone deacetylases (169). The E. coli periplasmic protein SurA, consisting of two iterative parvulin-like domains and is also known to be involved in the maturation of outer membrane porins (170).

8. Novel rotamase or Peptidyl prolyl cis-trans isomerase activity of Cyclophilin

In addition to their role in immunosuppression, every CyP examined to date possesses rotamase activity (Fig. 9), including the CyPA homologues present in crude extracts of protozoan parasites like Schistosoma mansoni, Toxoplasma gondii, Plasmodium falciparum etc with the CyP from higher organism. The PPIase activity was originally recorded as an 18-kDa protein isolated from a porcine kidney cortex (171). The activity is often assessed with α-
Introduction

chymotrypsin coupled enzyme assay. For example, linear peptides with the general formula Suc (Yaa) \( \alpha \)-Xaa-Pro-Phe-p-nitroanilide (Suc, succinyl) exist in equilibrium between the cis and trans forms. The C-terminal blocking group is cleaved off instantaneously by \( \alpha \)-chymotrypsin if the Xaa-Pro bond is exclusively in the trans form. Release of the chromogenic group (p-nitroaniline) is monitored by measuring the absorbance at constant wavelength of 405 nm with time. In the presence of a PPlase, the rate of the cis-trans isomerisation of peptides is accelerated. Biochemical studies have further shown that CyPs possess this activity and act both as catalysts and chaperones in protein folding events speeding up slow rate limiting steps in the folding of various proteins, including the protein rich collagens (172). It has been suggested that the biological activity of certain peptide hormones may be regulated by isomerisation (173). It has been shown through X ray crystallography that in a CyP-CsA complex, the CsA binding site is the PPIase active site (174). This is indeed true since in most of the CyPs, their isomerase activity is inhibited on the addition of exogenous CsA. The favored notion is that the drug binds and inhibits the enzyme because it contains a structure termed a 'twisted amide surrogate' which is a transition state mimic of a peptidyl-prolyl bond undergoing isomerisation (175). However, the affinity of the drug varies with the different isoforms of CyP. The X ray structure of a CyPB-CsA analogue complex contains an almost identical conformation of the bound CsA to that in the CyPA-CsA complex, implying that the inhibitory activities of CsA must be due to differences in accessibility of different isoforms of CyP by the drug (176).

X-ray crystallographic and site-directed mutagenesis studies indicated that 13 specific amino acid residues of CyP A are involved in the binding of CsA. A specific stretch "WLDGKHVV" is highly conserved among most of the CyP isoforms and their homologous (156,177). Liu and his co-workers compared human and E coli CyP and showed that the tryptophan 121 (W121) of human CyP A is particularly important for CsA binding. Contrastingly E coli CyP contain a phenylalanine (F) residue at this position and consequently bind CsA weakly. Substitution of this F with W totally reverses the affinity of the drug binding of this CyP.

Identification and cloning of a high affinity CsA-binding cyclophilin (LdCyP) from Leishmania donovani, a trypanosomatid parasite that is naturally resistant to CsA has been reported from our laboratory (178). The translated LdCyP consist of 187 amino acids with a cleavable 21 amino acid hydrophobic, ER directed N-terminal extension. Modelling studies confirmed that all the residues of human CyP responsible for interaction with CsA are sequentially and conformationally conserved in LdCyP. LdCyP displayed biochemical parameters comparable to human CyP. The sub-cellular organelles of L. donovani constitute the bulk of the CsA binding activity, whereas the binding activity of the cytosolic fraction was conspicuously low, indicated organellar localisation of the protein in the parasite. Insensitivity of L. donovani to CsA is probably due to the lack of CsA binding activity in the cytoplasm of the parasite. In turn it was suggested that LdCyP, which is located in the secretory pathway, might function as a chaperone by binding to membrane protein rather than as the mediator of calcineurin inhibition.
Proteins must fold into defined three-dimensional structures to have their functional activity. But, in reality, newly synthesized proteins are at great risk of aberrant folding and aggregation, thereby resulting in the possibility of forming potentially toxic species. Protein aggregation is one of the most troublesome manifestations of protein instability and uncontrolled protein aggregation has been proved as a major challenge in living system. Additionally, protein aggregation, along with other physical and/or chemical instabilities, remains to be one of the major barriers towards hindering commercialization of biotechnological products, especially protein drug candidates. Although a variety of methods have been used/or designed to prevent/or inhibit protein aggregation, the end results are often unsatisfactory for many proteins.

9.1. Abberation in protein folding leads to protein aggregation:

Protein folding is the process by which nascent polypeptide chain or misfolded protein is converted to biologically active/native conformation with a definite three-dimensional structure. Proteins, translated products of genes, are responsible for all physiological function in any biological system. Hence, defective genes, producing defective proteins, are mainly responsible for abnormal physiological responses leading to disease symptoms. Therefore, to become biologically active these proteins must fold into correct stable three-dimensional structure. In the final rate-limiting step of folding, the protein achieves its native conformation with the emergence of functional properties. This step includes: precise ordering of the secondary structures, the correct packing of the hydrophobic core, the correct domain pairing in multidomain proteins, the reshuffling of disulfide bonds, cis-trans proline isomerization, and subunit assembly in oligomeric proteins. During such intermediate transient processes there is always a chance to have a little deviation from the ordered sequence which ultimately leads to a misfolded product with loss of auto folding ability. These populations of misfolded transient products are neither being properly folded nor show any physiological activity; rather they are converted to aggregated molecules.

Protein folding/unfolding intermediates act as the precursors in protein aggregation, even though the intermediates are usually not stable (179). In contrast, completely folded or unfolded proteins do not aggregate easily as the hydrophobic side chains are either mostly buried out of contact with water, or randomly scattered (180). It is the patches of contiguous hydrophobic groups in the folding/unfolding intermediates that initiate the aggregation process. Aggregation of many proteins has been shown to be initiated by intermediates, such as scrapie amyloid (prion) protein (PrP27–30) (181), carbonic anhydrase B (CAB) (182,183), recombinant human growth hormone (rhGH) (184), insulin (185), human lysozyme variants (186), P22 tailspike polypeptide

Figure 10: A. a-helix structure, B. $\beta$-sheet structure
Introduction

Computer simulation studies also demonstrate that the aggregation process is originated primarily from interactions of partially folded intermediates (189). Thermal treatment can easily generate protein-unfolding intermediates, which can rapidly aggregate, such as ovalbumin (190). Another factor contributing to the rapid aggregation of intermediates is their high rate of diffusion relative to the folded state (191). The high diffusion rate of the folding intermediates can significantly increase the probability of oligomer formation.

Conceptualization of protein folding mechanism began when Anfinsen successfully refolded denatured and reduced ribonuclease into the fully active enzyme. He stated the fundamental principle of protein folding in 1973: the folding of a protein is determined by its amino acid sequence (119). Protein folding is driven by the need for a protein to sequester its hydrophobic side chains from water. This sequestering allows the polar groups of the protein’s backbone (NH and COOH) to hydrogen bond. The fold of a protein depends on the types of amino acids present within its protein sequence. The secondary structure of the α-helix is formed by the hydrogen bonding between the NH and CO groups of the same strand whereas the β-strand forms hydrogen bonding with other β-strands (Fig. 10).

These β-strands stabilize each other through intermolecular hydrogen-bond-forming β-sheets. Tertiary structures build on these secondary structures to form other higher-order structures such as β-sheets, ββ units, β-hairpins, and α-helix bundles. The fibrils of β-amyloid are composed of mainly β-sheets.

Protein misfolding is usually a transition from the normal secondary structure of the protein into a β-sheet conformation (192,193). This transition may occur through genetic factors and environmental factors where conditions such as temperature, pH, ionic strength, etc. play major role in the misfolding of the protein. Also, as in the case of prion disorders, infectious non-native β-sheet conformation proteins can induce misfolding of other proteins (192,193). The proposed lowering of activation energy necessary for protein/surface ordering and facilitation of an equilibrium shift toward the β-sheet conformation may help to explain the observed random coil to β-sheet transition of Aβ when interacting with lipids and phospholipids (194). However, recent data revealed the existence of α-helical predominant protein aggregates (195).

9.2. Nucleation and growth of protein Aggregates:

Based on the above analysis, the aggregation process can be described in the form of a scheme (1), where proteins form reversible unfolding intermediates, which then form reversible unfolded proteins or irreversible/reversible aggregates. The process from N to A can be considered as the nucleation step, which is usually rate limiting; in other words, the aggregation process is nucleation dependent.
Nucleation has been demonstrated or suggested to be the initial step for fibrillation of a sequence of *E. coli*. The initial protein aggregates are soluble but gradually become insoluble as they exceed certain size and solubility limits (180,196).

### 9.3. Reversibility and specificity of physical aggregation

Depending on the amino acid composition, thermodynamic stability and protein folding pattern, reversibility of any protein aggregation may alter. External factors like presence of excipients also may influence the process. The reversibility of protein aggregation is usually dependent on the stage of the aggregation process. The initial formation of soluble aggregates (nucleation) may be reversible but the subsequent formation of insoluble aggregates is usually irreversible. The two different stages correspond, respectively, to thermodynamically unfavorable and favorable processes (also see next section).

### 9.4. Thermodynamics and kinetics of protein aggregation

Patro and Przybycien (1994) demonstrated that protein aggregation initially led to an increase in the overall free energy of the system mainly due to the loss of certain number of monomer conformational and translational states (loss of entropy). This initial aggregation process is analogous to the nucleation step, where it is thermodynamically unfavourable and rate limiting. Therefore, it is predicated that proteins with low native energies have a higher energetic barrier for aggregation and are less likely to aggregate (197). After the initial step, further aggregation results in a decrease in the overall free energy of the system and the aggregation process is thermodynamically favoured (198). This is because mutual attraction of hydrophobic residues/patches (aggregation) minimizes the area of unfavourable protein–solvent interface (interaction).

### 9.5. Factors influencing Protein Aggregation

**9.5.1. Secondary structure of Proteins:**

Secondary structures in peptides or proteins may play a role in controlling aggregation (198) as well as stability (199). Generally, protein/peptide aggregation involves mostly β-sheets, while α-helix structures seem to be less likely to form aggregates. This may be due to a stronger dipole moment of α-helices than that of β-sheets (199). A few studies on peptide aggregation support the above contentions. The aggregation of amyloid peptides was shown to have two steps: random/helix-to-β-sheet transitions and aggregation of β-sheets (200). Incorporating proline, another amino acid of low β-sheet propensity, in Aβ 1–40 (or Aβ 1–42) is responsible for the reduction of the aggregation tendency (201). In the study of the association behaviour of water-soluble *de novo* β-sheet peptides (22–32 amino acids), Janek et al. (1999) demonstrated that the degree of association correlated with the stabilization of the β-sheet structure. Amyloid formation of several Aβ analogues directly correlated with the content of β-sheet conformation (201).

The β-sheet content in proteins often increases upon aggregation. If a protein has both α-helix and β-sheet structures, the aggregation-induced increase in β-sheet content is often accompanied by a drop in α-helix structure, such as insulin (202), human lysozyme variants (186,190), and ovalbumin (190). Therefore, protein aggregates are often composed of dominant β-sheet structures. Typical
examples include insulin fibrils consisting of extended β-chains lying perpendicular to the fibril axis (185) and amyloid fibrils consisting of cross-β conformation (203).

9.5.2. **Protein concentration** is another important factor in protein aggregation. The mean-field lattice model predicts that proteins will aggregate/precipitate at sufficiently high concentrations. This is because proteins may need a critical concentration to form the initial nucleus for initiation of the aggregation process (204).

9.6. **Morphology and types of protein aggregates**:

Protein aggregates come in different shapes and sizes, even for aggregates of a single protein (187). Protein can self-assemble to form several types of aggregates:

**Amyloid Fibrils**: Fibrils are of insoluble fibrous, stable, filamentous protein aggregates causing several deadly diseases in eukaryotes. Amyloid fibrils are characterized by core cross-β sheet structure in which continuous β-sheets are formed with β-strands running perpendicular to the long axis of the fibrils (Fig. 11).

Improper protein folding, leading to the formation of amyloid fibrils, is responsible for the development of spongiform encephalopathies. These severe pathologies include scrapie in sheep, mad cow disease in cattle and Creutzfeld-Jacob in humans. Alzheimer’s disease is also characterized by the presence of amyloid fibrilar deposits in the brain tissue. Table 3 summerizes several amyloid proteins and their related pathologies.

![Figure 11: Cross-β-sheets structure of protein](image.png)
<table>
<thead>
<tr>
<th>Clinical syndrome</th>
<th>Precursor protein</th>
<th>Fibril component</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer's diseases</td>
<td>Amyloid protein precursor</td>
<td>β-Peptide 1-40 to 1-43</td>
</tr>
<tr>
<td>Primary systemic amyloidosis</td>
<td>Immunoglobulin light chain</td>
<td>Intact light chain or fragments</td>
</tr>
<tr>
<td>Secondary systemic amyloidosis</td>
<td>Serum amyloid A</td>
<td>Amyloid A (76-residue fragment)</td>
</tr>
<tr>
<td>Senile systemic amyloidosis</td>
<td>Transthyretin</td>
<td>Transthyretin or fragments</td>
</tr>
<tr>
<td>Hereditary cerebral amyloid angiopathy</td>
<td>Cystatin C</td>
<td>Cystatin C minus 10 residues</td>
</tr>
<tr>
<td>Hemodialysis related amyloidosis</td>
<td>β₂ microglobulin</td>
<td>β₂ microglobulin</td>
</tr>
<tr>
<td>Familial amyloid polyneuropathy III</td>
<td>Apolipoprotein A1</td>
<td>Fragments of apolipoprotein A1</td>
</tr>
<tr>
<td>Finnish hereditary systemic amyloidosis</td>
<td>Gelsosin</td>
<td>71-aminoacid fragments of gelsosin</td>
</tr>
<tr>
<td>Type II diabetes</td>
<td>Islet amyloid polypeptide (IAPP)</td>
<td>Fragment of IAPP</td>
</tr>
<tr>
<td>Medullary carcinoma of the thyroid</td>
<td>Calcitonin</td>
<td>Fragments of calcitonin</td>
</tr>
<tr>
<td>Spongiform encephalopathies</td>
<td>Prion</td>
<td>Prion or fragments thereof</td>
</tr>
<tr>
<td>Atrial amyloidosis</td>
<td>Atrial natriuretic factor (ANF)</td>
<td>ANF</td>
</tr>
<tr>
<td>Injection localized amyloidosis</td>
<td>Insulin</td>
<td>Insulin</td>
</tr>
<tr>
<td>Hereditary renal amyloidosis</td>
<td>Fibrinogen</td>
<td>Fibrinogen fragments</td>
</tr>
</tbody>
</table>

Table 3: Amyloidogenic proteins and the corresponding diseases
**Amorphous aggregates:** In most cases, protein aggregates are amorphous. Amorphous aggregates are formed faster than fibrils. There is no special conformational prerequisite for amorphous aggregation to occur and many destabilized and partially folded proteins precipitate out of solution in a form of amorphous aggregates.

**Globular oligomer:** The term 'globular oligomer' or 'soluble oligomer' is used here to describe any nonmonomeric form of a protein that is soluble in aqueous solutions and remains in solution after high-speed centrifugation, indicating that it is not an insoluble fibrillar aggregated species. Depending on the amino acid sequence and peculiarities of the environment, proteins become self-assembled to produce soluble oligomer. Recent studies have highlighted the fact that small soluble oligomeric species are potentially more cytotoxic than mature fibrils (205). Invention of the toxicity of non-fibrillar amyloid and amylospheroid support the contention.

![Figure 12: Different types of oligomers formed in different pathways](image)

It is now well-established phenomenon that a given protein can self-assemble into various aggregated forms, depending on the peculiarities of its environment. In fact, the typical aggregation process only rarely results in the appearance of a homogeneous product where at the end of the reaction only one
aggregated species (amyloid fibrils, amorphous aggregates or soluble oligomers) is present. More often, heterogeneous mixtures of various aggregated forms are observed. Furthermore, each aggregated form can have multiple morphologies and monomers comprising morphologically different aggregated forms can be structurally different.

All these suggest that aggregation is not a simple reaction, but a very complex process with multiple related and unrelated pathways that can be connected or disjoined. However, regardless of the model or pathway considered, the appearance of a large aggregate inevitably involves the formation of some small oligomeric species.

9.7. Inhibition of protein aggregation:

9.7.1. Molecular chaperones

'Molecular chaperone' is a class of protein that facilitates proper folding of another proteins by binding to and stabilizing misfolded or unfolded or aggregated proteins. The basic paradigm of molecular chaperones is that they recognize and selectively bind nonnative proteins to form relatively stable complexes (206). In most cases, the complexes are dissociated by the binding and hydrolysis of ATP. It is likely that most, if not all, cellular proteins interact with chaperones at some stages of their lifetime. Chaperone act on another protein by two ways: one group of chaperone bind the unfolded or misfolded protein, producing properly folded native form; second group exert their chaperonic activity with the novel disaggregating property of the inactive protein aggregates by altering the environmental condition responsible for the aggregation or conformational modulation (207,208) (Fig. 13).

Chaperones act as catalysts, i.e., they transiently interact with their substrate proteins but are not present in the final folded product, and also in that they increase the yield of folded protein. The major classes of general chaperones are the HSP40, HSP60, HSP70, HSP90, HSP100 and the small heat shock proteins.

9.7.2. Cyclophilin as chaperone:

Since cyclophilin catalyses the cis-trans isomerisation of peptidyl prolyl bonds of certain proteins, they act as acceleration factor in protein folding and assembly. In proteins, the trans state of
peptide bond is favoured at least 100 times over cis, which means that in unfolded proteins the trans-cis isomerisation is extremely slow, more slow for peptidyl prolyl isomerisation. This is considered to be the rate limiting step in the folding of protein with cis peptide bond in their final conformation(209). While studying collagen folding in vitro in presence of CyP, Steinmann et al. (1991) demonstrated that inhibition of PPase activity by the addition of CsA resulted in delayed maturation of collagen(210). Similar results were obtained with RNase T1 (211).

Alteration of activities of enzymes and transcription factors are observed in presence of CyP (212,213). It has been suggested that multidomain large CyP perform chaperone like activities via their additional domain while the small single domain CyP is restricted to their PPase associated function only (214). A controversy exists regarding the chaperone function of small immunophilins (212,215,216). For example although immunophilin like CyP A and NinaA are known to form stable complex with their protein substrate HIV gag1 protein and Rh1 rhodopsin respectively (217), the interaction of CyP A with its client protein HIV gag1 does not lead to any increase in the yield of fully assembled virus particle, suggesting an unlikely role of CyP A as a chaperone (218,219).

However many observations suggest that CyP acts as chaperones (215,220) . Studies with guanidine-HCl treated human carbonic anhydrase II (HCA II) as well as creatine kinase show that in presence of CyP, early folding state can be enhanced by chaperone like activity of CyP, whereby preventing off-pathway reaction and the isomerase activity is performed later in the folding process. PPase independent chaperone activity of small protein Flpa of E coli was also reported (221) . Recently, chaperonic activity of single-domain CyP 18 was established on the substrate HCA II through hydrophobic condensation that enables rescue of transient misfolded molten globule intermediates (222) . More recently, a group of scientists showed the PPase independent chaperone-like function of recombinant human cyclophilin A on arginine kinase refolding (223). It was claimed that surface hydrophobicity of rhCyPA can suppress AK aggregation and binding to rhCyPA hydrophobic pocket is a prerequisite for chaperoning AK folding.Interestingly, some studies revealed that chaperones work inside the cell-organelles in an well-organised interactive pattern. Using complementary ER-specific methods, researchers mapped interactions between ER-lumenal chaperones and foldases and describe their organization in multiprotein complexes. It was proved from these novel scientific investigations that ERp72-cyclophilin B complex enhances the rate of folding of immunoglobulin G (224). In addition to ER-specific studies, multi-chaperones-interactors network and their role in carcinogenesis was also reported in mitochondria (MtCIN) (225). Thus due to their enzymatic properties, multiplicity, cellular localisation and the role in protein folding CyPs belongs to a diverse set of proteins family called molecular chaperones.

10. Physiological role of Cyclophilins

Studies reveal that apart from mediating CsA induced immunosuppressive effect, CyP are involved in various cellular processes including protein-protein interaction, cell division, protein folding, receptor maturation and several other processes. They have broad substrate specificity for nascent protein. Various isoforms of CyP are localised to specific organelles, implicating in specialised function in each of these isoforms. CyP A is capable of protecting cardiomyocytes against oxidative damage (226) . Immunophilins help in the in vitro folding process of polypeptides. For example, the folding and secretion of transferrin from HgpG2 cells (hepatocytes) is controlled by CyPs. Some CyP could function as both inflammatory proteins and cytokines and seem to co-regulate the expression of histamine and de
novo synthesis of the peptide leukotriene C4 \([\text{LTC4}]\) in human basophills (176). In yeast, the product of \(\text{Wis}^*\), a gene homologous to CyP, helps in progression of cell-cycle in strain, which are defective in cell-cycle regulating proteins (227). Both FKBPS2 and isoforms of CyP, CyP 40 are part of HSP90/70 complex that binds to steroid receptor. The C-terminal segment that include a 3-unit tetratricopeptide repeat domain (TPR) terminated by a potential site for calmodulin binding. This TPR domain is fundamentally important for the immunophilins to bind to the HSP90. In Xenopus oocyte, it was shown that the functional homooligomeric neuronal nicotinic (nAChR) and type-3 serotonin (5HT3R) receptor expression is dependent on CyP (113). CyP A is involved in the uptake of HIV-1 by macrophage and T-lymphocyte through its interaction with Gag-polypeptide (114).

Different forms of CyP, which have altered CsA binding domain and extended C and N-terminal domain, have been described in vertebrates (228), \(\text{Drossophila}\) (229) and in \(\text{Brugia malayi}\) (230). Such divergent forms may posses a more specific function. Consistent with this hypothesis, \(\text{NinaA}\), a divergent CyP isoforms identified in \(\text{Drossophila}\) was shown to be responsible for specific folding of the rhodopsin in the compound eye acting as specific chaperones (217,218).

10.1. Protein import to the mitochondria

It has been found that CyP 20 in \(\text{Neurospora crassa}\) is functionally linked with Hsp 70, Hsp 60 and Hsp 10 chaperone machinery (231). Refolding of proteins that are imported to mitochondria and folding of proteins that are synthesized on intramitochondrial ribosomes are ATP dependent and involves cooperation of Hsp 60 with mitochondrial Hsp 70 and Hsp 10 (232) as well as CyP 20. Another mitochondrial cyclophilin D is a component of the mitochondrial permeability transition pore (MPTP), a dynamic multiprotein complex located at the contact site between the outer and inner mitochondrial membrane, which is crucial for coordination between the mitochondrial intermembrane space and cellular matrix (233). The collapse of mitochondrial membrane potential is due to the CsA sensitive MPTP. CyP D binds to the inner mitochondrial membrane in a CsA sensitive manner. It was found that the PPIase activity is necessary for protein conformational changes associated with MPT pore formation.

10.2. Oxidative stress

Recent report suggested that the CyP A plays an antagonistic role in the cellular oxidative stress mediated by binding to the antioxidant protein Apol (213) or through the activation of the peroxidase activity of peroxidoxin (234). Another report suggested that mitochondrial CyP D has a deleterious role under oxidative stress by aiding the formation of permeability transition pores (PT) in mitochondria leading to necrotic death (235).

10.3. Cell signalling

The cell surface receptor CD147, involved in HIV-1 infection, is an extracellular receptor for CyP A and CyP B. In case of CyP A, its binding to the receptor initiates the cellular cascade that result in the activation of ERK and chemotaxis (236).

10.4. Apoptosis

Cyclophilin A, B and C have been shown to be involved in apoptosis which is independent of their PPIase activity. They posses calcium/magnesium-dependent DNA nuclease activity that is independent of their PPIase activity (237). This nuclease activity of cyclophilin is similar to the activity
Introduction

Introduction

of apoptotic endonucleases and is capable of degrading single, double and supercoiled DNA. It is believed that cyclophilin might involve in apoptotic genome degradation by their nuclease activity. On the other hand, CyP involvement in apoptosis as a part of the MPTP showed that the addition of CsA prevented apoptosis by blocking the opening of MPT, thus in turn stabilising the mitochondria and preventing the release of mitochondrial cytochrome C into the cytosol (238). Whether the anti-apoptotic activity of the CsA is due to the prevention of the MPT opening or that CsA is also effective in decreasing the endonuclease activity of the cyclophilin, remains to be explored.

10.5. Viral replication and activation

Cyclophilin A (CypA) is the main member of the immunophilin superfamily that has peptidylprolyl cis-trans isomerase activity. CypA plays critical roles in the replication and/or virion maturation for a number of different viruses. Upon influenza virus infection, CypA inhibits viral replication by interacting with the M1 protein. In addition, CypA is incorporated into the influenza virus virions (239). Cyclophilin A (CyPA) is a cellular protein that acts as a major factor in virus replication and/or virion maturation for a number of different viruses, including human immunodeficiency virus, hepatitis C virus, murine cytomegalovirus, influenza A virus and vaccinia virus. Cyclophilin A is required for efficient human cytomegalovirus DNA replication and reactivation (240). HCMV requires a stable level of CyPA for successful virus replication, virion production and reactivation from infected cells.

11. Heat shock proteins as molecular chaperones:

The small heat shock protein (HSP) and α-crystallin family consists of 12- to 43-kDa proteins that assemble into larger multimeric structures and contain a conserved C-terminal region termed the α-crystallin domain. Many of the small HSPs are induced only under stress conditions. They have been shown to function in vitro as chaperones by preventing protein aggregation in an ATP-independent manner. Extensive work has been done regarding structure and function of these chaperones (241-243). The small HSP exhibits high affinity for partially folded intermediates but show no apparent substrate specificity and are only functional in the oligomeric state (244).

Little is known about the mechanism of action of the small HSPs. It has been suggested that the hydrophobic interactions are critical in substrate binding.

HSP40 (DnaJ) family: The HSP40 or DnaJ family consists of over 100 members, characterised by the presence of a highly conserved J-domain of ~78 residues (244). Proteins in this family typically consist of several domains, eg. DnaJ contains after four conserved regions representing potential functional domains [the J-domain, which is linked by a Gly / Phe-rich region to a domain of unknown function, followed by a zinc-finger region and ending with the C-terminal domain, also of unknown function]. The best-studied examples are DnaJ from E. coli and several homologues from yeast such as Mdj1 and Ydj1 (245-247).

The best defined role for the HSP40 is a co-chaperone for HSP70; however, even this function is not well understood, and there is evidence to indicate that DnaJ and other members of the HSP40 family are chaperones in their own right, binding to at least some unfolded proteins and nascent chains (248). In E. coli, DnaK, DnaJ and GrpE cooperate synergistically in a variety of biological functions, including protein folding. The properties of DnaJ and its homologues have been studied in great detail for the past few decades (244,246,249,250).
Figure 14: GroEL reaction cycle. T and D represent ATP and ADP respectively. \( l_c \) represents an intermediate committed to fold to native state, whereas \( l_{oc} \) represents intermediates that are not so committed.

Significant specificity in the interactions between members of the HSP70, DnaJ and GrpE families has been observed (251). As noted, the interaction between a given HSP70 and DnaJ is determined by the J-domain (252). Recently, evidence for interactions between DnaJ homologues and HSP90 has been reported (253). It has also been suggested that DnaJ possesses an active dithiol / disulphide group and may catalyse protein disulphide formation, reduction and isomerisation (254).

HSP60 family: This family is also known as HSP60 or chaperonin family, which include both the GroE-mini chaperone network and TCP-1 ring complex families. GroEL and its homologues are found in prokaryotes, chloroplasts and mitochondria whereas TCP-1 and its homologues are found in the eukaryotic cytosol. Many of the HSP60 chaperones are also known as chaperonins (cpn60) and are ring shaped oligomeric protein complexes with a large central hydrophobic cavity in which nonnative proteins can bind. In bacteria, at least HSP60 require a cochaperonin, GroES (cpn10), for full function. In this context the term ‘chaperonin’ was originally coined by Ellis (206) to refer to non-heat-induced HSP60.

GroEL in combination with its cochaperonin GroES and ATP, facilitates protein folding, not only by preventing aggregation but also by simultaneously allowing partially folded intermediates to fold in a microenvironment that conducive to stabilize the native state (Fig. 14). Members of the HSP60 family are also involved in the assembly of the large mutiprotein complexes such as Rubisco by providing a favourable folding condition (255,256).

HSP70 family: The HSP70 is a family of molecular chaperones that are involved in protein folding and several other cellular functions and that exhibit weak ATPase activity. The HSP70 chaperones are composed of two major functional domains. The N-terminal highly conserved ATPase domain binds ADP and ATP very tightly (in the presence of Mg** and K+) and hydrolyses ATP, whereas the C-terminal domain is required for polypeptide binding. Cooperation of both domains is needed for protein folding as well as many of the functions of the E. coli HSP70, DnaK, involve two additional cofactors, DnaJ and GrpE. The component constituents of the HSP70 family are numerous and most of the organisms having multiple members.

Recently, a new pair of DnaK / DnaJ-like chaperones has been discovered in E.coli (257). Sequence differences between HSC66 and HSC20 compared with other HSP70 / HSP40 members...
Introduction

suggest that these chaperones may have different peptide binding specificity and is subjected to different regulatory mechanisms.

**HSP90 family:** The members of the HSP90 family are highly conserved essential proteins found in all organisms from bacteria to humans. Examples include the cytosolic form of eukaryotes, HSP90, the ER form, Grp94 and the *E. coli* homologue HtpG. Mammalian HSP90 exists as dimers. Although there are a number of similarities between the activities of HSP90 and HSP70, the former has several identified specific interactions, for example, with cytoskeletal elements, signal transduction proteins (including steroid hormone receptors) and protein kinases (such as the mitogen activated protein kinase system). HSP90 is frequently found in complexes with other chaperones. *In vitro*, HSP90 exhibits chaperone activity with diverse proteins, suggesting a general function. The properties of HSP90 have been extensively studied in several reviews (164,243,258,259). The chaperone property of HSP90 is found to be ATP-dependent (260).

**HSP100 family:** The heat inducible members of the HSP100 (or Clp) family of proteins have a number of very interesting properties and share a common function in helping organisms to survive under extreme stress (261). They perform a diverse set of functions, including proteolysis. They are highly conserved, present in all organisms and contain ATP and polypeptide binding sites. Both HSP104 and ClpA form six-membered ring complexes; the diameter of the interior of the rings is much smaller than in GroEL, making it unlikely that the HSP100 function analogously to HSP60. The basic mechanisms of these chaperones function are not understood. There is some suggestion that HSP104 may act in concert with HSP70 and DnaJ homologues to increase the yields of renatured proteins (261). It should be noted that no human analogues of HSP104 have been found.

Unlike HSP60 and HSP70, which are unable to resolubilize aggregated proteins *in vitro* (with the exception of RNA polymerase), HSP104 has been observed to solubilize thermally aggregated proteins both *in vitro* and *in vivo* (262). Interestingly, ClpA can substitute for the ATP-dependent chaperone function of DnaK and DnaJ in the *in vitro* activation of the plasmid "P1 repA" replication initiator protein (263). Another unusual feature of HSP104 is its role in triggering a prion like disorder in yeast, involving the extrachromosomal elements PSI+ and URE3 (264).

**Calnexin:** Calnexin, also referred to as IP90, p88 and p90, is an ~90 kDa integral membrane protein of the endoplasmic reticulum (ER). Many resident ER proteins act as molecular chaperones and participate in the proper folding of polypeptides and their assembly into multisubunit proteins. Studies indicate that calnexin associates with the major histocompatability complex (MHC) class I heavy chains, partial complexes of the T cell receptor and B cell membrane immunoglobulin, but not with completed receptor complexes. It has been shown that calnexin is a chaperone that retains incompletely or improperly folded proteins in the ER.

**Calreticulin:** Proteins undergo numerous modifications, including folding, translocation, and degradation. During such modifications, polypeptides are rarely in a native, stable state. Molecular chaperones are a diverse group of proteins that modulate polypeptide stability through ATP-dependent folding. Many chaperone proteins are found in the endoplasmic reticulum (ER). Calreticulin is a luminal ER protein that is 39% homologous to the ER chaperone protein, calnexin. Calreticulin contains C-terminal KDEL ER retention signal, and can bind Ca²⁺, Zn²⁺, ATP, and the proteins, ERP57 and protein disulfide isomerase. The chaperonic activities of calreticulin may include folding of both Asn-linked glycoproteins and non-glycosylated proteins. In addition, calreticulin is a component of MHC I/transporter associated with Ag presentation (TAP) complex where it may function in peptide assembly onto nascent class I molecules like calnexin. Calreticulin may also function in integrin signaling, since it
binds α3-integrin subunits and regulates integrin-mediated metalloprotease secretion. Thus, calreticulin may be involved in Ca\(^{2+}\) storage, cell adhesion, and protein folding (265).

**Protein disulphide isomerase:** Protein disulphide isomerase (PDI) is a critical cofactor in the folding of many proteins that are found in the ER (266-269). Many secreted proteins have multiple disulphide bonds, presenting potential limitation for correct disulphide pairing during folding. In vitro studies on the refolding of reduced proteins show that disulphide bond formation occurs rapidly and is followed much more slowly by thiol disulphide rearrangements leading to the correct disulphide pairings. Thus catalysis of oxidative folding is necessary in vivo to rapidly generate the correct disulphide bonds in newly synthesized proteins. In the eukaryotic ER, PDI performs this function. This is not surprising that PDI has been reported to have chaperone like activity at high concentrations (such as inhibition of aggregation) distinct from its disulphide bond interactions (270,271). PDI has two distinct catalytic sites situated in two domains homologous to thioredoxin, one near the N-terminus and the other near the C-terminus. The thioredoxin domains, by themselves, can catalyse disulphide formation but are unable to catalyse disulphide isomerizations (272).

**Specialized chaperones:** Some molecular chaperones may be highly specific in that they interact with only one or a very limited number of target proteins. PapD (273) which is involved in the assembly of bacterial pili and HSP47 (274), which is involved in the folding and processing of procollagen in the ER come under this group.

Receptor associated protein (RAP) is another example of a specialized molecular chaperone. It has the direct effect on the low density lipoprotein receptor-related protein (LRP), a large receptor that binds multiple ligands. The major role of RAP is to facilitate correct folding of LRP and to prevent the premature interaction of ligands with LRP (275).

There are also a number of chaperones involved in protein export, such as SecB from *E. coli* (276). SecB perform two functions; it maintains precursors of some exported proteins in the conformation which are compatible for export, by preventing them from aggregating or from folding to their native state in the cytoplasm and it delivers both nascent and completed precursors to SecA, one of the components of the export apparatus associated with the plasma membrane. Unlike the HSP70 and HSP60, in which hydrolysis of ATP is coupled to the binding and release of substrate proteins, SecB does not form stable complexes with substrate proteins. This may reflect the fact that secB does not mediate protein folding but is specialized for the export pathway.

In addition to DnaJ and GrpE, which function as cochaperones with DnaK, and have been known for several years, other protein cofactors that interact with HSP70 have been discovered. These include Hip (HSC70-interacting protein), BAG-1 and auxilin (277,278). The existence of these cofactors illustrates the complexity of the HSP70 chaperone machinery in cells.

**12. Disaggregation Leads to Reactivation of Protein**

Mechanisms by which aggregation and disaggregation of various proteins occur under various stress conditions within the complex cellular environment is under continuing investigation (207,279,280). For the first time, using heat-treated protein aggregates, in vitro experiments showed that Hsp104, together with Hsp70 and Hsp40, could disaggregate protein aggregates, thereby facilitating their refolding (261). The process of disaggregation was shown to be strictly ATP-dependent. Since then, several other chaperone complexes, having the capacity to disrupt protein aggregates, have been
discovered (208,281). Recently, a novel analytical study has been reported which explained the mechanistic pathway by which HSP 100 chaperone system played an essential role to disaggregate and reactivate aggregated proteins (282). DnaK-ClpB interaction, eventually, plays a dual role in both disaggregation and subsequent refolding of polypeptide chains as they emerge from the aggregate.

13. Folding in the endoplasmic reticulum

The ER is a key compartment in cells for protein export and contains many chaperones that are essential for the production of functional proteins for export (283). Folding begins with the insertion of a preprotein into the lumen of the ER and can occur either post-translationally, in which case the preprotein is completely synthesized on cytosolic ribosomes before being translocated, or cotranslationally, in which case membrane-associated ribosomes direct the nascent polypeptide chain into the lumen of ER concomitant with polypeptide elongation (284). The ER has excellent quality control mechanisms involving chaperones that recognize and selectively retain misfolded proteins, which are then either refolded or degraded in the cytosol through proteosome (285,286).

Although some of the major chaperones involved in protein folding in the ER are well studied, viz. BiP, Grp78, Grp94 and PDI, it is apparent that more have yet to be characterized. Because, the release of HSP90, PDI, BiP, Erp72, calreticulin and p50 is stimulated by Ca\(^{2+}\), these proteins appear to function as Ca\(^{2+}\) dependent chaperones (287). Evidences accumulating that the ER HSP70 chaperone machinery is similar to that in the cytosol and bacteria, in that at least two DnaJ homologues viz. Scjlp and Kar2p (DnaJ and HSP70 homologues of the yeast ER respectively) have been found in the yeast ER (254).

There are also evidences that peptidyl-prolyl isomerases bind to several nascent ER-proteins and regulates the protein secretion through ER, which in turn regulated itself via controlled aggregation-disaggregation of proteins (288,289).

14. Chaperones in parasites

Although, there are few reports regarding the expression of developmentally regulated chaperones such as HSPs (290-292), PPlases (293-296), p23 (297), DnaJ (298) and calreticulins (299) in parasites like *Plasmodium spp.*, *Trypanosoma spp.*, *Leishmania spp.* etc, but yet the importance of chaperones in parasite life cycle or in their virulence remains to be explored. Recently, the importance of HSP90 during transformation of *Leishmania donovani* promastigotes to amastigotes has been shown by proteomic study (300).

Though CyPs from different organisms are known to be involved in various cellular processes including protein-protein interaction, facilitation of protein folding, receptor maturation and several other mechanisms (158,215,217,221,231,301-304), in parasites, exact biological function of this protein is still unknown.
15. Protein aggregation may inhibit clearance of misfolded protein

Accumulation of aggregates inside the cell or extracellular deposition of misfolded protein may lead to neurodegenerative diseases. Very recent studies revealed that impairment in protein quality control (PQC) viz. ubiquitin-proteasome system (UPS) or the autophagosome-lysosome pathway inhibit the process of wasted and/or misfolded protein clearance (305). Under normal conditions, the cells are able to efficiently utilize their PQC system to handle the misfolded proteins and maintain the protein homeostasis. The molecular chaperones involved in the cellular PQC systems, such as heat shock proteins (Hsp), recognize misfolded proteins, assist in their refolding, prevent their aggregation, and help to repair or eliminate the damaged proteins. Deficiencies in PQC system causes abnormal accumulation of disease-associated proteins, resulting in cellular stress and finally neurodegeneration.

Dysfunction of PQC may also contribute to polyglutamine (Poly Q) expansion which abrogates the clearance mechanism of misfolded cytosolic protein by the ubiquitin-proteasome system (UPS) (306). PolyQ aggregation affect the cellular level of HSP 40 chaperone Sislp which interferes with the entrance of cytosolic proteins into the nucleus for degradation and form cytoplasmic inclusions.

16. Protein folding and aggregation in biotechnology

Two main applications concerning protein folding are involved in biotechnologies: protein engineering and the de novo design of proteins with novel functions. Recombinant proteins of pharmaceutical interest, such as growth hormone, insulin, and antihemophilic factor VIII, are commonly expressed and overproduced in E. coli or other cells. Since the overexpression of genes in foreign hosts often results in the formation of inclusion bodies, further processing such as unfolding and refolding are required for solubilising these proteins. These proteins can also be modified by genetic engineering to increase their stability for storage, which is an important industrial challenge. De novo protein design has recently emerged with the hope of constructing proteins with functions unprecedented in nature as well as increased solubility and stability (307). Genetic engineering provides a powerful methodology to redesign existing proteins. An understanding of folding is important for the analysis of many events involved in cellular regulation, the design of proteins with novel functions, the utilization of sequence information from the various genome projects and the development of novel therapeutic strategies for treating or preventing human diseases that are associated with the failure of proteins to fold correctly.

[Taken from the review entitled 'Protein folding: a perspective for biology, medicine and biotechnology' by Yon, J., M. (2001) Protein folding and its implications in biology and medicine Brazilian Journal of Medical and Biological Research 24, 419-435].

17. Rationale, objective and scope of the present work

Trypanosomatid diseases are a major cause of massive health hazard and death worldwide. Leishmania donovani, a purine auxoroph, causes leishmaniasis in human. Since AdK is the gate-way enzyme of the purine salvage pathway of the parasite, among all the purine salvage enzymes its role appears to be of prime importance in salvation of purines from the host due to the fact that phosphorylation of Ado to AMP by AdK is the driving force for the uptake of the purine base, especially in purine auxotrophs.
Introduction

However, studies suggest that AdK has an inherent tendency to form soluble aggregates (308). It is also known that for its aggregation to occur, AdK does not require heat exposure, pH alteration or treatment with co-solvents like trifluoroethanol. Moreover, the disaggregation and reactivation of AdK by LdCyP, which unlike other conventional nucleotide-binding chaperones viz. Hsp104, Gro E and Clp family of proteins, occurs in absence of ATP.

LdCyP, even, does not require any accessory protein or external unnatural factors for the chaperonic function. In addition to these, it has also been demonstrated that LdCyP acts on AdK in a manner that is completely peptidyl prolyl isomerase independent. Taken together, these facts made aggregation-disaggregation cycle of LdAdK quite inimitable and at the same time physiologically relevant. Clearly therefore, detailed studies may reveal as to whether LdCyP-mediated enzyme reactivation has any direct role in the Ado uptake of the cell.

Feedback inhibition of enzymatic reaction by its own product is well known. Results of the present work suggests that ADP, which is a reaction product of LdAdK, inhibits the enzyme by producing AdK-aggregates. Results also suggest that after binding to AdK, LdCyP withdraw ADP molecules from AdK-ADP complex, thereby effectively reducing AdK-aggregation resulting in significant enzymatic reactivation. Of the two forms of AdK-aggregates that are formed with or without ADP, experimental results demonstrated that there are structural and functional differences between the two forms of aggregates. Although various protein chaperones have been shown to modulate the function of various proteins by a number of means, reactivation of inactive enzyme aggregate, caused by its own reaction product, by CyP-chaperone may uncover a new chapter in chaperone literature. The principal goal of my work centers around studies on the characterization of AdK aggregates formed in absence and presence of the product-ligand (ADP) and analysis the stages of the reactivation process of the enzyme by LdCyP and a truncated fragment of LdCyP. Information generated from these studies, beside revealing structural insights into the reactivation mechanism of AdK at the molecular level, might be helpful while designing specific inhibitors for prevention of enzyme aggregates.

Finally, the present work on the mechanism of formation of leishmanial AdK aggregates and its disaggregation by a single domain CyP-chaperone and by a truncated LdCyP-fragment is of significance for the following reasons. Since the data presented in this dissertation show that N-del LdCyP (the truncated form of LdCyP) is as efficient as the full-length LdCyP in terms of its enzyme stimulating activity, it was of interest to investigate if there was still any shorter LdCyP-fragment that could perform the same enzyme reactivating function. In view of such a possibility, a study was undertaken to search for a LdCyP-peptide fragment of shortest chain length. Results presented showed that indeed a small decapeptide from LdCyP sequence stimulated AdK activity. In view of this finding, a rational strategy to develop peptide mimetics against protein reactivating ability of LdCyP could be envisioned. Apart from this, the structural characterization and detailed conformational analysis of the two forms of aggregates is expected throw more light on the detailed molecular mechanism by which LdCyP-chaperone exerts its unique function to disrupt ADP-facilitated aggregates of AdK and may provide some clue towards understanding as to how chaperone-mediated disaggregation of proteins occur. AdK-LdCyP interaction provides a unique model for this purpose.