Chapter 1

INTRODUCTION
1. Protein synthesis

Protein synthesis is the process in which cells build proteins. Proteins are synthesized on RNA templates known as messenger RNAs (mRNAs) in a process called translation. Translation can be divided into three steps: initiation, elongation and termination. In both prokaryotes and eukaryotes the first step of the initiation stage is the binding of a specific initiator methionyl tRNA (Met-tRNA_i) and the mRNA to the small ribosomal subunit. The large ribosomal subunit then joins the complex, forming a functional ribosome on which elongation of the polypeptide chain proceeds. A number of nonribosomal proteins are also required for catalyzing various steps of the translation process.

Initiation step of protein synthesis in eukaryotes is more complex than the other steps, namely, elongation and termination. It requires at least ten proteins (each consisting of multiple polypeptide chains), which are designated as eukaryotic initiation factors (eIFs). The process of initiation is schematically shown in Fig. 1. The factors eIF1, eIF1A, and eIF3 bind to the 40S ribosomal subunit, and eIF2 (in a complex with GTP) associates with the Met-tRNA_i (Hershey, 1991; Merrick, 1992). The mRNA is recognized and brought to the ribosome by the eIF4 group of factors. The 5´ cap of the mRNA is recognized by eIF4E. Another factor, eIF4G, binds to both eIF4E and to the poly-A binding protein or PABP, associated with the poly-A tail at the 3´ end of the mRNA. Some of these initiation factors thus recognize both the 5´ and 3´ ends of mRNAs, accounting for the stimulatory effect of polyadenylation on translation. The initiation factors eIF4E and eIF4G, in association with eIF4A and eIF4B, then bring the mRNA to the 40S ribosomal subunit, with eIF4G interacting with eIF3. The 40S ribosomal subunit, in association with the bound Met-tRNA_i and eIFs, then scans the
mRNA to identify the AUG initiation codon. When the AUG codon is reached, eIF5 triggers the hydrolysis of GTP bound to eIF2. Initiation factors (including eIF2 bound to GDP) are then released, and a 60S subunit binds to the 40S subunit to form the 80S initiation complex of eukaryotic cells. After the initiation complex has formed, translation proceeds by elongation of the polypeptide chain. Protein synthesis is complete with the process of termination. Both these processes are catalyzed by various protein factors.

**Fig. 1.** Summary of initiation of protein synthesis in eukaryotes. (Adapted from Pal et al., 1996).
1.1. Regulation of protein synthesis

Although transcription is the primary level at which gene expression is controlled, the translation of mRNAs is also regulated in both prokaryotic and eukaryotic cells. Among the 3 steps of protein synthesis, regulation operates mostly at the initiation step. One mechanism of regulation at the level of initiation is the binding of repressor proteins to specific mRNA sequences, which block translation. However, the best characterized mechanism of translation initiation in eukaryotic cells, resulting in global effects on overall translational activity rather than on the translation of specific mRNAs, involves modulation of the activity of initiation factors, particularly eIF2 (reviewed in Hershey, 1991; Proud, 1992; Samuel, 1993; Wek, 1994; Kaufman, 1994; Clemens, 1994). During initiation of protein synthesis, eIF2 binds to GTP and Met-tRNA$_i$ to form a ternary complex, eIF2•GTP•Met-tRNA$_i$ and subsequently forms the 43S pre-initiation complex with the 40S ribosomal subunit. Upon binding of mRNA and joining of the 60S ribosomal subunit, eIF2•GTP is hydrolysed to eIF2•GDP (Fig. 1). eIF2, being rate limiting needs to be recycled for continued initiation of protein synthesis.

The recycling of eIF2 requires the exchange of bound GDP for GTP, which is catalyzed by initiation factor eIF2B (also called reversing factor). When eIF2 is phosphorylated at the $\alpha$-subunit (eIF2$\alpha$) by the eIF2$\alpha$ specific kinases, the binding affinity of eIF2 ($\alpha\bar{P}$)•GDP to eIF2B is much higher than that of eIF2•GDP to eIF2B. Since eIF2B is present in cells at a lower molar concentration with respect to eIF2, once the amount of phosphorylated eIF2 exceeds the amount of eIF2B, the GDP/GTP exchange activity of eIF2B is prevented, resulting thereby inhibition of protein synthesis (Hinnebusch, 1994; reviewed in Chen and London, 1995) (Fig. 2).
1.2. Role of eIF2α and eIF2α kinases

Eukaryotic cells recognize and process diverse stress signals to elicit programmes of gene expression that are designed to remediate cellular damage, or alternatively induce apoptosis. An important contributor to stress adaptation is a family of protein kinases that phosphorylate the α-subunit of eIF2. The phosphorylation of eIF2α was first detected in rabbit reticulocyte lysate deprived of hemin. The absence of hemin results in the activation of a highly specific eIF2α kinase, called the heme-regulated inhibitor (HRI) or EIF2AK1 (eIF2α kinase 1) (reviewed in Chen, 1993; Chen and London, 1995). In addition to HRI, three other eIF2 kinases have been identified in mammals. Each of these contains unique regulatory regions that recognize a different set of stress conditions (Fig. 3). For example, the eIF2 kinase PKR (EIF2AK2) participates in an anti-viral defence mechanism that is mediated by interferon (Barber,
Additional mammalian eIF2 kinases include PEK (pancreatic eIF2α kinase) or PERK [PKR (dsRNA-dependent protein kinase)-like ER (endoplasmic reticulum) kinase] or EIF2AK3 that is activated in response to misfolded protein in the ER (ER stress) (Kaufman, 2004) and GCN2 [general control non-derepressible-2] or EIF2AK4 which is induced during amino acid deprivation by a mechanism that involves uncharged tRNA binding to a regulatory region homologous with HisRS (histidyl-tRNA synthetase) enzymes (Hinnebusch, 2000; Narasimhan et al., 2004). GCN2 is also activated by other stresses that are not directly related to nutritional deprivation, including UV irradiation and proteasome inhibition, and genetic studies suggest a role for the HisRS-related domain and uncharged tRNA in response to diverse stresses. Phosphorylation of eIF2 impedes recycling of eIF2 to its active GTP-bound form (Fig. 3), and the accompanying reduction in the levels of eIF2•GTP reduces global translation, allowing cells to conserve resources and to initiate a reconfiguration of gene expression to effectively manage stress conditions.
Fig. 3. Protein kinases, PKR, HRI (heme-regulated inhibitor), PERK and GCN2 are activated by different stress conditions to regulate the levels of eIF2α•GTP via eIF2 phosphorylation. (Adapted from Wek et al., 2006).

The unique function of each of these eIF2α kinases in response to different stresses has been verified by generations of knock-out mice with each of the 4 eIF2α kinases (Yang et al., 1995; Han et al., 2001; Harding et al., 2001; Zhang et al., 2002). Pkr−/− mice are compromised in their ability to respond to viral challenges (Yang et al., 1995; Abraham et al., 1999) and Perk−/− mice develop diabetes between 2 and 4 weeks of age (Harding et al., 2001). Gcn2−/− mice have reduced viability upon amino acid starvation (Zhang et al., 2002) and the erythroid response to iron and heme deficiency was found to be abnormal in Hri−/− mice (Han et al., 2005).
1.3. Structural features of eIF2α

In eukaryotes, eIF2 is a heterotrimer composed of α (36 kDa), β (38 kDa), and γ (52 kDa) subunits, which appear to remain associated throughout the initiation cycle. Cross-linking and genetic studies have suggested that both β- and γ-subunits are implicated in guanine nucleotide and Met-tRNA<sub>i</sub> binding (Anthony et al., 1990; Gaspar et al., 1994; Erickson and Hannig, 1996). In addition, the β-subunit was shown to interact specifically with eIF5 during GTP hydrolysis and also to bind to mRNA (Das et al., 1997; Laurino et al., 1999; Thompson et al., 2000). The γ- subunit participates in the recognition of the start site for protein synthesis (Dorris et al., 1995). The α- subunit (eIF2α) binds to Met-tRNA<sub>i</sub> and delivers the ternary complex to the 40S ribosomal subunit in a GTP-dependent manner. Phosphorylation/dephosphorylation of eIF2α at Ser-51 is the major regulator of protein synthesis in eukaryotic cells. Structural studies have provided valuable information regarding the functioning of the eIF2α subunit.

1.3.1. Human eIF2α- crystal structure

The first structural analysis on eIF2α was reported by Nonato et al. (2002). The three-dimensional structure of a 22 kDa N-terminal portion of human eIF2α (heIF2α) was determined by X-ray diffraction at 1.9 Å. This structure comprised the residues 3-182, roughly the N-terminal two-thirds of full length heIF2α. The overall structure (Fig. 4) can be divided into two major domains: oligonucleotide binding domain (OB domain) and the helical domain.

The N-terminus is a β-barrel with five antiparallel β-strands in an OB domain fold. A β-hairpin connects β1 and β2; β2 and β3 are linked by a four-residue loop and β4 and β5 by a three-residue loop. The loop connecting β3 and β4 is longer by 17
residues, and it was not completely modeled due to a lack of interpretable electron density. The visible part, residues 48-50, begins with a turn of 3_10 helix that is stabilized by a hydrogen bond between the carbonyl group of Leu-46 and nitrogen of Glu-49. Ser-51 residue is situated immediately after this helix. Residues 63 and 64 at the end of the loop are also visible with a hydrogen bond between the carbonyl group of Arg-63 and the main chain nitrogen of Arg-66.

The standard features of the OB fold, like presence of a β-bulge at residue Val-23, and a left handed Gly-65 are seen in the OB domain. The eIF2α N-terminal domain is the latest example of the large oligonucleotide/oligosaccharide-binding fold. Although eIF2α shows very little sequence homology with the protein members of the nucleotide binding family, their structures are superimposable. Interestingly a major difference can be seen in the mode of RNA binding. The site proposed as the RNA binding is found in nearly the same position in all the members of the nucleic-acid binding superfamily; this is the β-barrel region, where three loops connecting β1 and β2, β3 and β4, and β4 and β5 come together (Fig. 4). In most common cases RNA binding seems to involve solvent-exposed aromatic residues and positively charged residues on the surface of the protein. The OB domain of heIF2α does not have any of the clustered positive charged residues observed for other members of this family. This is also consistent with the observation that, unlike other family members in which the biochemical studies and limited structural data support the idea of direct interaction with nucleic-acids (Ruff et al., 1991; Schindelin et al., 1993; Sette et al., 1997), there is little or no evidence that eIF2α binds RNA. Most cross-linking and genetic experiments suggest that β- and γ-subunits of eIF2 are the subunits involved in initiator tRNA- and ribosomal RNA binding (Gaspar et al., 1994).
Fig. 4. Overall structure of heIF2α. The phosphorylation site (Ser-51) is indicated. Part of the loop connecting β3 to β4 is missing in the final model and is shown as a thin green line. The disulfide bridge connecting the OB domain (red) and the helical domain (blue) is shown in ball-and-stick form. The 3_{10} helices are shown in light blue. (Adapted from Nonato et al., 2002).

A helical domain follows the OB domain, and contains residues 88-182. In contrast to the highly conserved architecture of the OB domain, the helical domain adopts a novel helical pattern consisting of a 28-residue-long α-helix (α1), a series of small α-helices (α2, α3, α4 and α5), and one 3_{10} helix folded into a very compact domain (Fig. 4). The first helix has extensive interactions, including a disulfide bridge, to fix its orientation with respect to the OB domain. Sequence alignment data of various eIF2α suggests that this disulfide bridge is present only in mammalian eIF2α. The two
domains meet along a negatively charged groove with highly conserved residues, indicating this site as the likely site for protein-protein interaction (Nonato et al., 2002).

A possible mechanism explaining the interaction of eIF2α and the regulatory domain of eIF2B was also proposed by Nonato et al. They proposed that eIF2B and eIF2 make direct contact with each other in the Ser-51 containing loop. The interesting feature of this highly conserved loop sequence is the predominance of positively charged residues (Arg-52, Arg-53, Arg-54, Arg-56, Lys-60, and Arg-63) following Ser-51. The crystal structure suggests an interaction of these two subunits in which the positively charged residues participate in the interaction that is enhanced by phosphorylation.

1.3.2. Human eIF2α- NMR structure

Solution structure of heIF2α was reported by Ito et al. (2004). The protein consists of two domains that are mobile relative to each other. The NMR structure of the N-terminal portion of heIF2α is consistent with the previously determined crystal structures of eIF2α (Nonato et al., 2002; Dhaliwal and Hoffman, 2003). The overexpressed heIF2α carried three mutations (A27Q, L46H, and V71K) in order to enhance solubility and improving expression {solubility enhanced heIF2α (se-heIF2α)}. These mutations did not alter the tertiary structure (Ito and Wagner, 2004).

The N-terminal domain (NTD) has an S1-type oligonucleotide/oligosaccharide binding-fold subdomain and an α-helical subdomain. The NMR structure of the NTD (Fig. 5) entirely matched the crystal structure of the NTD of heIF2α (Fig. 4) (Nonato et
An aromatic hydrophobic core between the S1 and α-helical subdomains, consisting of Phe-7, Tyr-8, Phe-101, Tyr-108, Phe-130 and Trp-135 (Fig. 6) is believed to stabilize the interaction between the two NTD subdomains (Ito et al., 2004).

Fig. 5. NMR structure of heIF2α. The ribbon presentation of heIF2α is shown with the secondary structural elements. While the α-helices and the β-sheets in the NTD are colored with blue and cyan, respectively, those in the CTD are colored with red and orange, respectively (Adapted from Ito et al., 2004).

The C-terminal domain (CTD) adopts an αβ-fold very similar to the CTD of elongation factor (eEF) 1Bα, the guanine nucleotide exchange factor for eEF1A. It contains five β-strands (β6-β10), two α-helices (α6 and α7), and one loosely associated C-terminal helix α8 that has no well defined orientation relative to the CTD or NTD (Fig. 5). The structural and functional similarities found between eIF2α/eIF2γ and eEF1Bα/eEF1A suggest a model for the interaction of eIF2α with eIF2γ, and eIF2 with Met-tRNAMet. The geometry of the putative eIF2α/eIF2γ interface suggests a possible role of eIF2α for the GDP/GTP exchange in eIF2γ (Ito et al., 2004).
Fig. 6. Key residues in the NTD of hεIF2α. The residues in the hydrophobic aromatic cluster which stabilizes the entire NTD structure are highlighted. (Adapted from Ito et al., 2004).

1.3.3. Yeast eIF2α

Crystal structure of eIF2α from *Saccharomyces cerevisiae* was reported by Dhaliwal and Hoffman (2003). In most respects, the structure reported (Fig. 7) was similar to the hεIF2α structure (Fig. 4). The *S. cerevisiae* eIF2α lacks a disulfide bridge that is present in the homologous protein in humans and some of the other higher eukaryotes. Interestingly, a conserved loop consisting of residues 51-65 and containing Ser-51, the putative phosphorylation site, is visible in the electron density maps of the *S. cerevisiae* eIF2α; most of this functionally important loop was not observed in the crystal structure of the human protein. This loop is relatively exposed to solvent, and contains two short $3_{\alpha}$ helices in addition to some extended structure. Ser-51 is located at the C-terminal end of one of the $3_{\alpha}$ helices and near several conserved positively
charged residues. The side-chain of Ser-51 is sufficiently exposed so that its phosphorylation would not necessitate a substantial change in the protein structure. However, a definitive answer as to whether the phosphorylation of Ser-51 induces a significant structural change in eIF2α must await direct investigation of the phosphorylated form using either X-ray crystallographic or NMR methods. The structures and relative positions of residues that have been implicated in kinase binding and in the interaction with guanine nucleotide exchange factor (eIF2B) are depicted in Fig. 8. (Dhaliwal and Hoffman, 2003).

![Structure of Yeast eIF2α](image)

**Fig. 7.** Structure of Yeast eIF2α. A diagram showing residues 1-175 of the yeast eIF2α structure, indicating the positions of some of the conserved side-chains in the vicinity of Ser-51. (Adapted from Dhaliwal and Hoffman, 2003).
Structure at the eIF2B interaction site

The guanine nucleotide exchange factor (eIF2B) contains five protein subunits. The regulatory subunits (α, β and δ) interact with eIF2α (Krishnamoorthy et al., 2001). This interaction is enhanced by the phosphorylation of Ser-51 of eIF2α. Mutational studies have mapped the residues in eIF2α that are important for this interaction (Vazquez de Aldana et al., 1993). These residues are Glu-49, in the 3₁₀ helix just before the phosphorylation site, Lys-79 and Gly-80, in the loop connecting β3 and β4, and Arg-
88 in the loop connecting β5 and α1. The electron density map reveals an intramolecular salt-bridge between Glu-49 and Arg-88. Lys-79 and Gly-80 are located 30 or more residues away from the phosphorylation site, suggesting that two non-contiguous segments within eIF2α may be involved in binding the regulatory subunits of eIF2B. Although Ser-51 and Arg-88 are far apart in the primary structure, a look at the tertiary structure reveals that the loop connecting β5 and α1, containing Arg-88, is in close proximity to the 3_10 helix located just before the phosphorylation site (Fig. 8A). Lys-79 and Gly-80, in the loop connecting β3 and β4, are on the surface of the protein; these loop residues are on the same face of the protein as the other residues involved in interaction with eIF2B (Fig. 8A).

**Structure at the kinase-binding site**

The residues of eIF2α that are involved in binding the specific kinases that phosphorylate Ser-51 have been mapped in the K3L protein, a structural mimic of eIF2α (Dar and Sicheri, 2002) that functions as a pseudo-substrate inhibitor of PKR encoded by vaccinia virus (Kawagishi-Kobayashi et al., 1997). Dar and Sicheri reported that PKR recognition motif and helix 1 of the helix insert region as important determinants for high-affinity binding to PKR. By mutational studies they showed that residues in the K3L protein corresponding to Lys-79 through Asp-83 in eIF2α (sequence KGYID) play a role in substrate recognition (Dar and Sicheri, 2002) (Fig. 9). Interestingly residues 79-83 are highly conserved in the eIF2α of all of the eukaryotes, which is consistent with their importance. In the yeast eIF2α structure, Lys-79 and Gly-80 form part of the loop connecting strands β4 and β5, and Tyr-81, Ile-82 and Asp-83 extend into the fifth β-strand. Four of these five residues are exposed on the surface of the protein (Fig. 8B); however, Ile-82 is part of the hydrophobic core of the β-barrel. The amino acid residues
flanking Ser-51 at the phosphorylation site are important for kinase recognition (Kawagishi-Kobayashi et al., 1997), and they presumably form a bipartite kinase binding surface. The distance from the substrate recognition motif (centered at Tyr-81) to Ser-51, the phosphorylation site, is about 23 Å. Interestingly, residues Lys-79 and Gly-80 are implicated in eIF2B binding (Vazquez de Aldana et al., 1993), which suggests an overlap between the binding domains of the eIF2α-specific kinases and eIF2B.

![Fig. 9. Conserved PKR Recognition Motif in K3L.](image)

The molecular surface of K3L is shown with K3L/eIF2α invariant residues colored green and the apex of the helical insert region, His 47, colored red. (B) and (C) are rotations of (A) by 120° and by 240° about the vertical axis, respectively (Adapted from Dar and Sicheri, 2002).

1.4. Structural features of eIF2α kinases

1.4.1. HRI

HRI or heme-regulated initiator of translation was discovered in reticulocytes under the conditions of iron- and heme deficiencies, which cause inhibition of protein synthesis at the level of initiation with disaggregation of polysomes (Waxman and Robinovitz, 1965; Waxman and Robinovitz, 1966; Grayzel et al., 1966; Zucker and Schulman, 1968; Maxwell and Rabinovitz, 1969; Legon et al., 1973). HRI undergoes
auto-phosphorylation and it specifically phosphorylates the α-subunit (38 kDa) of the eukaryotic initiation factor 2. Therefore it is called the heme-regulated eIF2α kinase (reviewed in Ochoa, 1983; London et al., 1987; Chen, 1993; Chen and London, 1995). HRI contains 12 conserved subdomains characteristic of all Ser/Thr Protein kinases (Hanks and Hunter, 1995).

HRI has been purified from rabbit reticulocyte lysate and has been extensively characterized (Kramer et al., 1976; Farell et al., 1977; Trachsel et al., 1978). Although the molecular mass of HRI predicted from the amino acid sequence is 72 kDa, previous reports have estimated that it is between 180 and 640 kDa (Chefalo et al., 1998; Bauer et al., 2001; Chen et al., 1989; Gross and Rabinovitz, 1972; Trachsel et al., 1978; Kudlicki et al., 1987; Yang et al., 1992). These various sizes are likely to be due to various oligomeric aggregates. Rabbit reticulocyte HRI cDNA was the first to be cloned and it recognizes a 3.1 Kb mRNA in reticulocytes (Chen et al., 1991, Anand and Pal, 1997). Although HRI is reported to sense the heme concentration in reticulocytes or red blood cells (Crosby et al., 1994), it is found in almost all tissues, including the brain (hippocampus, hypothalamus, and cerebellum), lung, heart, liver, spleen, kidney, thymus, stomach, pancreas, colon, testis, uterus and bone marrow, albeit in much lower quantity (Mellor et al., 1994; Berlanga et al., 1998; Omasa et al., 2002; Petrov et al., 2003). Therefore, it appears that HRI functions not only in the reticulocytes and red blood cells but also in most of the other cell types/tissues. The rabbit and rat HRI cDNAs encode for polypeptides of 626 and 621 amino acids, respectively, with a predicted subunit molecular mass of 70 kDa, although the apparent molecular size of each subunit is about 85-90 kDa. Both rabbit and rat HRI contain two heme regulatory
motifs (HRM), ACPYVM and RCPAQA that are located in the HRI kinase domain (reviewed in de Haro et al., 1996).

As illustrated in Fig. 10, HRI has three unique regions, in between the N-terminus and the C-terminus. Both the N-terminus and kinase insertion domain (KI) can bind heme, whereas the kinase catalytic domains (kinase I, kinase II) and the C-terminus cannot (Rafie-Kolpin et al., 2000). The N-terminus is necessary for stable high-affinity heme binding to HRI, and is also required for achieving higher eIF2α kinase activity, although it is not essential for the kinase activity of HRI. Importantly, the N-terminus is essential for the highly sensitive heme-regulation of HRI.

**Fig. 10. Protein structure of HRI.** HRI is divided into 5 domains as indicated. The amino acid sequence of mouse HRI is used here. Heme molecules are marked in red; S denotes the stable heme-binding site, while R denotes the reversible heme-binding site. *Histidine residues that coordinate the heme molecule. (Adapted from Chen, 2007).

**HRI activation- Biochemical and structural aspects**

**a) Activation during heme deficiency**

During heme deficiency, HRI is activated by multiple autophosphorylation and it is believed that multiple autophosphorylation is required for the formation of active and stable HRI that is regulated by heme. Newly synthesized HRI with heme
incorporated in its N-terminus domain rapidly dimerizes and undergoes intermolecular multiple autophosphorylation during heme deficiency in 3 stages (Fig. 11). In the first stage, autophosphorylation of newly synthesized HRI stabilizes HRI against aggregation. At this stage, HRI is an active autokinase, but is without eIF2α kinase activity. Additional autophosphorylation (denoted by P in Fig. 11) in the second stage is required for the formation of stable dimeric HRI that is regulated by heme. (Bauer et al., 2001). As shown in Fig. 11, during heme abundance, heme binds to this stable HRI and represses HRI activation (Rafie-Kolpin et al., 2000). During heme deficiency, HRI undergoes the final stage of autophosphorylation at Thr-485 (T485P) and attains its eIF2α kinase activity. This fully activated HRI is no longer regulated by heme, and is degraded (Rafie-Kolpin et al., 2003).

Both the autokinase and eIF2α kinase activities of purified homogeneous HRI were inhibited by hemin with an apparent K, of 0.2 μM (Chefalo et al., 1998; Bauer et al., 2001). Hemin inhibits ATP binding to HRI in a concentration-dependent manner (Chen et al., 1991) and thus blocks the kinase activities of HRI. Biochemical studies have demonstrated that HRI has 2 distinct types of heme-binding sites (Chefalo et al., 1998). One type of binding site is nearly saturated with stably-bound endogenous heme and copurifies with HRI, while the other binding site is available to bind exogenous hemin reversibly (Figs. 10, 11). This second reversible heme-binding site is likely to be responsible for the down-regulation of HRI activity by heme. The stoichiometry of 2 heme molecules per HRI monomer was established recently by direct measurement of heme chromophor (Bauer et al., 2001).
Fig. 11. **Role of heme in activating HRI.** When heme concentration is high, heme binds to the second heme-binding domain of HRI and keeps HRI in inactive state, thereby permitting globin protein synthesis and the formation of stable hemoglobin. During heme deficiency, HRI is activated by autophosphorylation. Activated HRI phosphorylates eIF2α and inhibits globin synthesis. (Adapted from Chen, 2007).

However, London and co-workers proposed a different mechanism for the activation of HRI. As per this model (Fig. 12), in the absence of hemin, HRI exists as a dimer, held by non-covalent interactions, and is active, whereas in the presence of hemin, HRI becomes an inactive, disulfide-linked dimer regardless of its phosphorylation state (reviewed in Pal et al., 1996).
Another model which Santoyo and colleagues postulated was that hemin regulates eIF2α kinase activity by promoting formation of an inactive heterodimer between HRI and hsp90 via disulfide bonds. The phosphorylation of the heme-reversible HRI promoted by Casein kinase-2 (CK2) activates HRI by preventing their interaction even in the presence of hemin. This constitutes the hemin-irreversible state of HRI. This could also be viewed as CK2 phosphorylating hsp90 and promoting the dissociation of the heterodimer HRI-hsp90. Once HRI is dissociated from hsp90, it would be phosphorylated by an autokinase activity (Mendez and DeHaro, 1994). This was later contradicted by Xu et al. (1997) who proposed a role for hsp90 in maintaining an active and stable conformation of HRI.
Structural basis of activation of HRI was put forth by Matts and co-workers (Yun et al., 2005). They emphasised on the inter-domain interactions that play an important role in HRI activation. The effect of hemin along with NO and CO on the interaction between the NTD and the catalytic domain of HRI were studied. Hemin stabilized the interaction of NTD and catalytic domain, and NO and CO respectively disrupted and stabilized this interaction. Those conditions which could stabilize the interaction were found to maintain the HRI in an inactive state and vice versa. These findings gave new insight into the mechanism of regulation of HRI and indicated that HRI’s kinase activity is regulated by interdomain interactions, with the NT-HBD being a structural or functional modulator of HRI’s catalytic activity (Yun et al., 2005).

The role of the N-terminal domain in oligomerization, kinase function, and heme sensing was reported by Miksanova et al., 2006. Multiangle light scattering (MALS) studies have been used to determine the state of oligomerization. Full length and various N-terminally truncated mutants were used in their study. These studies indicated that the full length HRI exists as a hexamer. The heme binding studies revealed that one molecule of heme binds to the full length HRI, probably with one axial ligand from the N-terminal domain and the other from the kinase domain, involving His and Cys residues, respectively (Fig. 13). This is not in agreement with the earlier observations (Bauer et al., 2001) that two molecules of heme bind to HRI. Kinetic studies showed that the HRI kinase reaction follows classical Michaelis-Menten kinetics with respect to ATP but sigmoidal kinetics and positive cooperativity between subunits with respect to the protein substrate (eIF2α) (Miksanova et al., 2006).
Recently, the molecular mechanism of heme-sensing by HRI has been studied (Igarashi et al., 2008). The optical and CD spectral findings on mutant proteins indicated that His-119/His-120 and Cys-409 are the axial ligands of the Fe (III) complex in the full length HRI protein. Cys-409 is part of the heme-regulatory Cys-Pro motif in the kinase domain. The study also demonstrated that the interactions between the N-terminal and kinase domains caused by global structural changes, in conjunction with coordination to the Fe (III) complex, play a key role in the heme-regulated catalytic function of HRI. A hypothetical model of the heme coordination structure of the full length HRI protein is presented in Fig. 14. The study proposed that heme regulation is induced by interactions between heme and the catalytic domain in conjunction with global tertiary structural changes at the N-terminal domain that accompany heme coordination and not merely by coordination of the heme iron with amino acids on the protein surface.
Fig. 14. A hypothetical model of the heme coordination structure of full-length HRI. Heme association/dissociation at the heme-sensing site of HRI regulates the eIF2α kinase reaction. Heme association with full-length HRI blocks catalysis (left), whereas heme dissociation opens the active site and allows catalysis (right). It is assumed that the heme binding site of HRI is situated away from the catalytic center and ATP binding site. Because the axial ligand(s) appears changed upon heme reduction, the redox-dependent ligand switch may additionally modulate catalysis. (Adapted from Igarashi et al., 2008).

b) Activation during various stresses

While each of the 4 eIF2α kinases has its specific stress signal in it, it is not clear which eIF2α kinase(s) is activated upon more general stresses such as oxidative stress, heat shock, and osmotic shock. In addition to HRI, both PKR (Farrell et al., 1977) and GCN2 (Lu et al., 2001) are expressed in erythroid precursors. Studies using Hri−/− erythroid precursors, have established that HRI was activated by arsenite-induced oxidative stress, osmotic shock, and heat shock, but not by ER stress, or amino acid or serum starvation in heme-sufficient reticulocytes and nucleated fetal liver erythroid precursor cells (Lu et al., 2001). These observations are consistent with the specific functions of PERK and GCN2 for ER stress and nutrient starvation, respectively. It is of
significant importance that HRI is the only eIF2α kinase activated by arsenite and is the major eIF2α kinase responsible for heat shock response in erythroid cells. Activation of HRI by arsenite involves reactive oxygen species and requires molecular chaperones hsp70 and hsp90 (Lu et al., 2001). Reports from few laboratories suggest that Hsp90 interacts with HRI both *in vivo* and *in vitro* in reticulocyte lysates to activate HRI. Hsp90 has been shown to co-purify with preparations of the kinase, and is found to modulate the activity of the enzyme. These results implicate hsp90 in the regulation of protein synthesis via its interaction with and perhaps regulation of HRI and phosphorylation of eIF2α (Rose et al., 1989). On the other hand Matts and Hurst (1989) reported an association of Hsp90 with latent HRI in hemin-supplemented reticulocyte lysates. They also showed that in reticulocyte preparations with different levels of Hsp90, the restoration of protein synthesis by the delayed addition of hemin is greater in lysates with higher levels of hsp90 (reviewed in Pal et al., 1996). Similarly, DeHaro and co-workers proposed a model suggesting that the binding of Hsp90 to HRI in the presence of hemin, inactivates HRI and the activation of HRI is caused by the disassociation of Hsp90 (Mendez et al., 1992; Mendez and DeHaro, 1994). This was later contradicted by Xu et al. (1997). They gave evidences supporting the fact that hsp90 and associated cohorts interact with inactive newly synthesized or partially denatured proteins to prevent their aggregation and chaperone the folding of the protein into a functional and active conformation. Hence, the principal role of the interaction of hsp90 with HRI appears to be stabilization of HRI in an activatable conformation which is protected from denaturation and aggregation while it is awaiting stimuli that induces HRI to undergo autophosphorylation and activation. Thus, it was clearly shown that dissociation of hsp90 from HRI is not the cause of HRI activation. Rather, it is proposed that hsp90 dissociation is the result of HRI attaining a conformation that is active and
stable but still regulated by heme in the absence of the hsp90 chaperone support (Xu et al., 1997). Recently, HRI has been reported to be activated by lead toxicity in \textit{in vitro} cultured K562 cells (Sarkar et al., 2005). As a hemoprotein, HRI could also be activated by NO and inhibited by CO (Uma et al., 2001). Activation of HRI by NO required its N-terminal domain that was loaded with heme (Igarashi et al., 2004, reviewed in Chen, 2007). The physiological significance of NO activation and CO inhibition of HRI remains to be further investigated.

1.4.2. PKR

The RNA-regulated protein kinase, (PKR) is an interferon-inducible enzyme of widespread occurrence in eukaryotic organisms. PKR is a member of the serine/threonine protein kinase family. PKR kinases are activated by an RNA-dependent autophosphorylation (Pestka et al., 1987; Samuel, 1991). cDNAs, encoding human PKR were first cloned in 1990 (Meurs et al., 1990). Subsequently, cDNA clones of PKR have been isolated from mouse and rat cells (reviewed in Samuel, 1993). The human PKR is a 551-amino acid protein with a molecular mass of about 62 kDa. This is somewhat lower than its apparent size (68 kDa) in polyacrylamide gels. Both mouse and rat PKR are smaller proteins of 518 and 514 amino acids, respectively (reviewed in Samuel, 1993). The PKR kinase possesses a short insert of approximately 30-40 amino acids in the subdomain IV. The conserved kinase catalytic region is represented in the carboxy-terminal half of PKR, whereas the amino-terminal half of the protein, which presumably plays a regulatory role, contains three clusters of basic amino acids. The first two basic regions contain two divergent copies of a dsRNA binding motif (Fig. 15) that are required for RNA binding, whereas the third basic region is dispensable for this function (Green and Mathews, 1992).
**Fig. 15. Protein structure of mouse PKR.** The 518 amino acid sequence of mouse PKR is illustrated in the figure. The amino terminal region contains two dsRNA binding domains. Kinase domain is interrupted by a large insert (KI domain). The numbers refer to the amino acid residues. (Adapted from Samuels, 1993).

The PKR gene is expressed in a wide range of cell types and it appears to be expressed constitutively at a low level under most physiological conditions. It is induced strongly at the transcriptional level within a few hours of interferon (IFN) treatment, and intracellular concentrations of the protein rise 5- to 10-fold under these circumstances. Studies on the subcellular distribution of PKR have shown that, although approximately 80% of the protein is cytoplasmic (mostly associated with ribosomes), a significant fraction is present in the nucleolus (Jeffrey et al., 1995). The significance of the later fraction needs to be established. The enzyme is phosphorylated extensively at multiple sites. The extend of phosphorylation is much less for the nucleolar form as compared to the cytoplasmic form of the kinase. PKR is a reasonably stable enzyme *in vivo* and so far there is no evidence for regulation of its concentration through changes in degradation rate (reviewed in Clemens, 1997).

Previous studies reported that there are two requirements for the activation of PKR. One is the binding of dsRNA to one or both motifs near the N-terminus; the other is the necessity for dimerization of the protein. The latter may come about as a
consequence of two (or more) molecules of PKR interacting with a single dsRNA molecule (Wu and Kaufman, 1997). However, there is also a strong evidence that direct protein-protein interaction can also bring about dimerization without the RNA ligand (reviewed in Clemens, 1997). This is further complicated by the observation that the same N-terminal regions of the protein that bind dsRNA are required for this protein-protein association. The detailed nature of the dimerization phenomenon needs to be established since it has implications for the mode of action of dominant negative mutants of PKR, which have been shown to cause tumorigenic transformation when stably expressed in NIH 3T3 cells (Meurs et al., 1993). Presently two models exist to describe how these mutants act to inhibit wild-type PKR activity. One idea is that the mutants sequester limiting amounts of cellular dsRNA activators, whereas an alternative suggestion is that there is direct heterodimerization between the mutant and the wild type proteins, which blocks the wild-type PKR activation (reviewed in Clemens, 1997).

The X-ray crystal structure of the catalytic domain of PKR in complex with eIF2α was recently reported by Dar et al. (2005). This structure gave details regarding the basis of substrate recognition and regulation of PKR. The structures reveal that eIF2α binds to the C-terminal catalytic lobe while catalytic-domain dimerization is mediated by the N-terminal lobe. In addition to inducing a local unfolding of the Ser-51 acceptor site in eIF2α, its mode of binding to PKR affords the Ser-51 site full access to the catalytic cleft of PKR.
**Fig. 16.** Structure of the PKR-eIF2α Complex: Ribbons representation of the PKR/eIF2α complex (trigonal P3$_2$121 space group) highlighting catalytic-domain dimerization mediated by the N lobe of PKR and eIF2α recognition mediated by the C lobe of PKR. The S1 subdomain (residues 3–90) and flanking helical subdomain (residues 91–175) of eIF2α are colored magenta and pink, respectively. The N and C lobes of PKR are colored purple and green (left molecule) and red and blue (right molecule), respectively. The activation segment (residues 432–458) and helix α0 (residues 260–266) of PKR are colored orange and yellow, respectively. The phospho-Thr446 side chain is shown in a ball-and-stick representation. Regions not modeled due to disorder and/or deletion in PKR (residues 338–351 corresponding to the eIF2α kinase characteristic insert) and eIF2α (residues 50–59 encompassing the Ser51 acceptor site) are shown as dashed lines. (Adapted from Dar et al., 2005).

**Structural overview of the PKR-eIF2α Complex**

The structure is remarkable in several regards. It sets a new standard by representing the first kinase to be crystallized with a full-length protein substrate. It also shows how dimerization contributes to kinase activation and highlights the role of the αG-helix of PKR (described below) as a docking motif for eIF2α. The structural predictions are supported by mutational studies by Dey et al. (2005). The coupling of dimerization and substrate docking, mediated by phosphorylation of the activation loop,
is revealed to be an ordered, highly dynamic, and extended allosteric process. The recent publication of a structure of GCN2 in the absence of eIF2α allows us to further appreciate the synergy between activation and substrate docking (Padyana et al., 2005). Like all eukaryotic protein kinases, PKR has a smaller, more dynamic amino-terminal lobe (N-Lobe) and a larger, stable, mostly helical carboxyl-terminal lobe (C-Lobe). Two events are necessary to stabilize the active conformation. In the N-lobe, correct positioning of the αC-helix is essential, whereas in the C-lobe the activation loop typically must be phosphorylated. This phosphate then interacts with a conserved His-Arg-Asp (HRD) motif that precedes the catalytic loop at the active site. This interaction of the phosphate with the HRD arginine stabilizes the active site (Johnson et al., 1996). Three very stable helices, (αE, αF, and αH), form the core of the C-lobe, whereas the αG-helix, in contrast, is more solvent exposed. Dar et al. (2005) described the crystal structures of two complexes of PKR and eIF2α. Both form a symmetrical dimer, in one case related by crystallographic symmetry. The dimer interface primarily involves the N-lobe of PKR and the importance of this interface for dimerization and activation was confirmed in the accompanying mutagenesis study from Dey et al. (2005). These studies demonstrate how essential it is to determine the structure of more complete kinase-substrate complexes in order to understand the detailed mechanistic features of protein kinase activation and protein phosphorylation. As the kinase core is highly conserved, each kinase is activated and interacts with its substrates in new ways. In the case of eIF2α, this structure of PKR: eIF2α explains why a peptide containing the phosphorylation site is a poor substrate, unlike full-length eIF2α. EIF2α, and many other protein substrates, docks to kinases by additional “tethering” sites that lie peripheral to the active site. The structure of PKR and eIF2α suggests that substrate docking is a highly dynamic process where eIF2α docking to a distal site may actually
Introduction

Contribute to the organization of the active site. Without the structure of the kinase-substrate complex, it is impossible to appreciate the complexity and synergy between these two regions. Additionally, the structure of eIF2α bound to PKR demonstrates the importance of the αG-helix of PKR as a substrate-docking motif, a function that is likely to be conserved in many protein kinases. The importance of the αG-helix for docking of proteins was demonstrated in three previous structures: the cdk2: KAP structure where cdk2 is the substrate for the KAP phosphatase, the RIα inhibitory subunit bound to the catalytic subunit of PKA (Song et al., 2001; Kim et al., 2005; Lei et al., 2000), and the auto-regulatory domain bound to PAK1. Interestingly, viruses also take advantage of this docking site. Dar et al. (2005) showed that by competing for the eIF2α docking site of PKR, K3L, a vaccinia protein, could prevent the shutdown of protein synthesis in response to viral entry. In PKR, unlike the other kinases, the αG-helix assumes an atypical position that is tightly coupled to the kinase activation loop, thus creating a new allosteric network that links substrate docking to dimerization (Fig. 17). Additionally, Arg-499, conserved in all eIF2α kinases at the C-terminus of the αG-helix, contributes to both positioning of this helix and coupling to the activation loop. An unresolved issue in this study is whether the αG-helix is always in an atypical position or whether this position is induced by eIF2α. In the GCN2 structure, with no bound eIF2α, the αG-helix is still in an atypical position indicating that this is likely to be an intrinsic feature of the eIF2α kinases (Padyana et al., 2005).

Another unusual feature revealed by the eIF2α-PKR complex is the region flanking the eIF2α phosphorylation site, Ser-51. Although this segment is poised in close proximity to the active site, it is nevertheless disordered in the crystal structure. This region, helical in free eIF2α, appears to have “melted” as a consequence of binding.
to PKR. Interestingly, phosphorylation can take place if the Ser-51 residue is replaced by a tyrosine. The structure, supported by mutagenesis analysis, explains this ambiguity by demonstrating that specificity is conveyed by a peripheral docking site. Once docked, the actual residue being phosphorylated can be rather promiscuous for PKR and perhaps for most dual-specific protein kinases. Furthermore, mutagenesis demonstrates that dimerization and auto-phosphorylation can be achieved even when the eIF2α docking site is non-functional indicating that the substrate binding is not required for kinase activation, at least for a nonspecific peptide substrate.

The two requirements for activation, phosphorylation of the Thr-446 in the activation loop and correct orientation of the αC-helix, are achieved for PKR by dimerization of the N-lobe. As shown in Fig. 17, by interacting with the HRD Arg-413, the phosphate couples the activation loop to the active site. Typically, several other basic residues interact with the phosphate. In PKR, Lys-304, and Arg-307 in the αC-helix fill this role thereby stabilizing the helix in its active conformation. This allows a conserved glutamate in the αC-helix (Glu-308) to interact with a conserved lysine (Lys-296) in β-strand 3, an interaction that is important for catalytic activity (reviewed in Taylor et al., 2005).
Fig. 17. Activation and coupling of distal sites by phosphorylation of the activation loop. (a) Coupling of the phosphorylated activation loop in eIF2α to the dimerization site through the αC-helix, to the substrate docking site through the αG-helix, and to the active site through the HRD Arg-413. (b) The auto-phosphorylated active form of PKR shows the multivalent coordination of phosphorylated Thr-446 (pT446) in the activation loop and the electrostatic pairing of Glu-308 and Lys-296. (Adapted from Taylor et al., 2005).

1.4.3. GCN2

GCN2 is the most widespread eIF2α kinase superfamily member, first identified as an inducer of GCN4, a transcriptional activator of amino acid biosynthetic genes in budding yeast. Yeast GCN2 is a 1590-amino acid protein with a molecular mass of about 180 kDa. The catalytic domain of GCN2 possesses a 110 amino acid large insert between subdomain IV and V. Adjacent to the kinase domain, GCN2 contains a 530...
amino acid sequence-related histidyl-tRNA synthetases (HisRS). This domain is required for its positive regulatory function \textit{in vivo}. The extreme carboxyl terminal segment of GCN2 is essential for its interaction with the 60S ribosomal subunit (Hinnebusch, 1994; Wek, 1994). Phosphorylation of eIF2\(\alpha\) by the GCN2 kinase, activated in response to amino acid starvation mediates gene-specific translational control of GCN4 in yeast. The GCN4 gene encodes a transcriptional activator of amino acid biosynthetic genes and GCN2 activity is required for increased translation of GCN4 mRNA in amino acid starved cells. Under these conditions, the translation of other yeast mRNAs is reduced. This unique response depends on four short upstream open reading frames in the leader region of GCN4 mRNA (Hinnebusch, 1994).

Similar to HRI and PKR, GCN2 also phosphorylates eIF2\(\alpha\) at Ser-51 residue. Substitution of Ser-51 in eIF2\(\alpha\) with Ala (the \textit{SUI2-S51A} allele) completely eliminates the increased phosphorylation in amino acid-starved cells and impairs derepression of \textit{GCN4} to the same extent as deletion of \textit{GCN2}. Immunopurified GCN2 specifically phosphorylated the \(\alpha\) subunit of eIF2 purified from rabbit or yeast but not yeast eIF2 containing the Ala-51 substitution (Dever \textit{et al.}, 1992). Thus, these results established that GCN2 stimulates \textit{GCN4} translation by phosphorylating eIF2\(\alpha\) on Ser-51. Low level expression of the mammalian eIF2\(\alpha\) kinases HRI and PKR in \textit{gcn2} mutants induces \textit{GCN4} translation in a manner completely dependent on Ser-51 in eIF2\(\alpha\) (reviewed in Hinnebusch, 1997).

GCN2 kinase has also been cloned and characterized from \textit{Drosophila melanogaster} (DGCN2) (Santoya \textit{et al.}, 1997) and from \textit{Neurospora crassa} (CPC3) (Sattlegger \textit{et al.}, 1998). Expression of DGCN2 is regulated developmentally and at
later stages becomes restricted to a few cells of the central nervous system (Santoyo et al., 1997). The first mammalian GCN2 homolog was cloned and characterized in mouse (Berlanga et al., 1999). Mouse GCN2 (MGCN2) has a conserved motif, N-terminal to the kinase subdomain V, and a large insert of 139 amino acids located between subdomains IV and V that are characteristic of the known eIF2α kinases. MGCN2 contains a class II aminoacyl-tRNA synthetase domain and a degenerate kinase segment, downstream and upstream of the eIF2α kinase domain, respectively, and both are singular features of GCN2 protein kinases (Fig. 18).

**Fig. 18. Protein structure of mouse GCN2.** The 1648 amino acid MGCN2 sequence is illustrated by a large box. Highlighted domains include N-terminal; the ‘degenerate kinase’ that is related to subdomains I-XI of eukaryotic protein kinase; the conserved two lobes of the eIF2α kinase domain, separated by a large insert (KI domain); and the HisRS-like domain that includes the three motifs (m1, m2 and m3) conserved among the class II aminoacyl-tRNA synthetases. The numbers refer to the amino acid residues. (Adapted from Berlanga et al., 1999).

The crystal structure of the Yeast GCN2 protein kinase domain was reported by Padyana et al. (2005). In order to examine the molecular mechanisms responsible for GCN2 autoinhibition and activation by uncharged tRNAs, they determined a series of X-ray structures of the dimeric catalytic domain of Saccharomyces cerevisiae GCN2 in apo- and ATP-bound forms at 3.0- and 2.75-Å resolution, respectively. The isolated PK
domain of GCN2 is completely inert \textit{in vitro}, but, remarkably, activity is rescued by single amino acid substitutions at Arg-794 (R794G) or Phe-842 (F842L). These constitutively activating (GCN2\textsuperscript{c}) mutations bypass the tRNA binding requirement for kinase activation \textit{in vivo}, and it was previously proposed by Padyana and co-workers that they alter the PK active site in a way that mimics conformational changes induced by interactions with the HisRS or RB/DD domains on tRNA binding. The crystal structures reveal partial closure of the active site cleft that restricts ATP binding by restraining the conformation of the hinge region between the N- and C-lobes of the PK domain. In addition, the X-ray structures of the R794G mutant PK domain in apo- and AMPPNP-bound forms at 1.95- and 2.0-Å resolution respectively were also reported. These two structures demonstrate that this activating mutation increases the flexibility of the hinge segment and opens a “molecular flap” that increases the inter-lobate space and accessibility of the enzyme active site. Hence, their work reveals a novel nucleotide gating mechanism via conformational modulation of the hinge region that controls kinase activity. They proposed a two-step activation mechanism in which tRNA binding to the HisRS domain leads to a comparable structural remodeling of the hinge region of wild-type GCN2 that facilitates ATP binding. Subsequent autophosphorylation of the activation loop is predicted to facilitate an additional realignment of active site residues necessary for substrate phosphorylation.

**Structural overview of the GCN2 PK domain**

The GCN2 PK has a typical, bi-lobate kinase fold with the ATP binding cleft positioned in the hinge region between the two lobes (Fig. 19a). Alignment of the crystal structure of GCN2PK\textsubscript{WT} with those of other protein kinase structures in the PDB showed conformational similarity to human cyclin-dependent kinase 2, human tyrosine
kinase c-Src, human protein kinase B, murine cAMP-dependent protein kinase and human hematopoietic cell kinase.

The smaller N-terminal lobe (N-lobe, residues 599–788) is preceded by a short α-helix (αA, residues 594–598) and contains a five-stranded (β1–β5), twisted β-sheet with two α-helices (αB and αC) linking β3–β4 (Fig. 19a). The position of αC is displaced as seen in the inactive conformation of CDK2 (38), c-Src, and hematopoietic cell kinases (39–41). The large insert characteristic of the eIF2α kinase family members occurs between β4 and β5. The nucleotide binding “P-loop” (606–611) linking β1 to β2 is poorly ordered in some copies of the enzyme within the asymmetric unit.

The larger C-terminal lobe (C-lobe, residues 795–982) containing amino acids implicated in catalysis, activation, and substrate recognition is predominantly α-helical (αD–αI), and is connected to the N-lobe by the hinge region (residues 790–794, Fig. 19a). The conformation of the Asp-Phe-Gly (DFG) motif of apo-GCN2PKWT closely resembles that of the cAPK in its active conformation (PDB code 1ATP). In both structures, the DFG motif projects into the active site, whereas the activation loop extends out into solvent (Fig. 19a, magenta and orange segments, respectively). The 42-residue activation loop in GCN2 (residues 853–894) is longer than is typically seen in protein kinases. Approximately 22 residues at the center of the activation loop (residues 861–882) appear disordered and were invisible in experimental electron density maps (Fig. 19a).

Consistent with previously published studies (Qiu et al., 2001; Narasimhan et al., 2004) indicating that GCN2 functions as a homodimer, the GCN2 PK domain
repeatedly crystallized as a symmetric homodimer (Fig. 19b, 19c) independent of lattice packing arrangements. On dimer formation, ~2,600 Å$^2$ of solvent-accessible surface area is buried, which is consistent with a stable homodimer (Lo Conte et al., 1999). Size exclusion chromatographic studies also demonstrated that GCN2 PK is dimeric in solution. The dimer interface is composed equally of hydrophobic and polar side chains, and is stabilized by ~26 amino acids from each monomer (within residues 594–830) that participate in 22 hydrogen bonding interactions. All but four of these residues are located in the N-lobe. Given this dimerization interface, the mode of dimerization observed in the crystals obtained almost certainly represents the PK domain dimer found within the dimer of full-length GCN2.

Fig. 19. Structure of wild-type GCN2 PK bound to Mg$_2$ATP. a, ribbon drawing of GCN2PK$_{WT}$·Mg$_2$ATP monomer. α-Helices are shown in cyan, 3$_{10}$-helices in blue, β-strands in green, connecting loops in gold, P-loop in red, catalytic loop in magenta, and activation segment in orange. ATP is shown as an atomic stick figure and the Mg$^{2+}$ ions are denoted with brown spheres. The deleted PK insert and the disordered activation loop are shown as broken lines. The phosphorylation site at Thr-887 is shown as an
atomic stick figure. The DFG motif of the kinase domain that occurs next to the β8-strand on the activation loop is obscured by the Mg\textsuperscript{2+} ions. b, surface representation of the GCN2PK\textsubscript{WT}·Mg\textsubscript{2+}ATP dimer (individual monomers in cyan and green, respectively) with the dimer axis perpendicular to the page. c, view in panel b rotated 90° about the vertical. The green monomer surface is rendered semitransparent and shown with an embedded protein ribbon. (Adapted from Padyana et al., 2005).

1.4.4. PERK

This is the most recently characterized mammalian eIF2α kinase. PERK cDNA was first identified and characterized from rat pancreatic islet cells by Shi \textit{et al.} (1998). This cDNA encoded a new related kinase, which was termed as pancreatic eIF2α kinase (PEK). PEK mRNA is expressed in all tissues, with highest level in pancreatic cells (Shi \textit{et al.}, 1998). PEK phosphorylates the α-subunit of eIF2 at residue Ser-51 in response to stresses that impair protein folding in the endoplasmic reticulum (ER). While the kinase domain of PEK is similar to those of eIF2α kinases, including the characteristic large insert between subdomains IV and V; the flanking 550-residue amino terminal sequences are distinct (Fig. 20). PEK was found to function in translation regulation in both the yeast and reticulocyte lysate model systems. Rat PEK protein is 1049 amino acid long having a predicted molecular weight of 128 kDa, whereas apparent molecular mass on SDS-PAGE was observed as 140 kDa. The mouse orthologue of PEK was identified and characterized by Harding \textit{et al.} (1999) and was named as PERK, on the basis of its similarity with PKR and its localization in endoplasmic reticulum membrane. Mouse PERK is a 1114 amino acid long protein. This was also found to be activated specifically in response to ER-stress. Activated mouse PERK could phosphorylate eIF2α at ser-51 residue.
Furthermore, PEK homologs from human, *Drosophila melanogaster* and *Caenorhabditis elegans* were also identified and characterized. Expression of human PEK mRNA was found in over 50 different tissues examined, with highest levels in secretory tissues. In mammalian cells subjected to ER stress, elevated eIF2\(\alpha\) phosphorylation was observed to be coincidental with increased PEK autophosphorylation and eIF2\(\alpha\) kinase activity (Sood *et al.*, 2000).

**Fig. 20. Protein structure of mouse PERK.** The 1114 amino acid sequence is illustrated in the figure. It has a distinct amino terminal sequence, which contains signal peptide (SP) and transmembrane (TM) domain. It also contains conserved two lobes of the eIF2\(\alpha\) kinase domain, separated by a large kinase insertion domain (black box). The numbers refer to the amino acid residues. (Adapted from Harding *et al.*, 1999).

Even though the detailed structural mechanism regarding the activation of PERK is not known presently, some studies on another type I transmembrane ER-localized protein kinase receptor namely Inositol-requiring kinase 1 (IRE1) has shed some light in this regard. The crystal structure of the luminal domain of human IRE1\(\alpha\) was reported by Zhou *et al.* (2006). The monomer of the luminal domain comprises a unique fold of a triangular assembly of \(\beta\)-sheet clusters. Structural analysis identified an extensive dimerization interface stabilized by hydrogen bonds and hydrophobic interactions. Dimerization creates an MHC-like groove at the interface. However, because this groove is too narrow for peptide binding and the purified luminal domain
forms high-affinity dimers \textit{in vitro}, peptide binding to this groove is not required for dimerization. Consistent with their structural observations, mutations that disrupt the dimerization interface produced IRE1\textalpha{} molecules that failed to either dimerize or activate the unfolded protein response (UPR) upon ER stress. In addition, mutations in a structurally homologous region within PERK also prevented dimerization. Their structural, biochemical, and functional studies \textit{in vivo} altogether demonstrated that IRE1 and PERK have conserved a common molecular interface necessary and sufficient for dimerization and UPR signalling (Zhou \textit{et al}., 2006).

1.4.5. Other eIF2\textalpha{} kinases

Apart from these four above mentioned eIF2\textalpha{} kinases, a number of protozoan eIF2\textalpha{} kinases have also been identified in different protozoan parasites. These eIF2\textalpha{} kinases play an important role in regulating protein synthesis in these parasites in response to environmental stresses, during invasion and differentiation. Even though the structural basis regarding the activation of these kinases is not available presently, various biochemical studies have provided valuable information regarding the regulation of protein synthesis in these parasites. A short description of these eIF2\textalpha{} kinases is given below:

\textbf{PfPK4}: PfPK4 is the first protozoan eIF2\textalpha{} kinase to be identified. Surolia and Padmanaban (1991) described the presence of heme-dependent eIF2\textalpha{} kinase activity in malarial parasite acting in a similar way to rabbit reticulocyte HRI. They observed an increase in eIF2\textalpha{} phosphorylation after chloroquine treatment of rabbit reticulocyte lysate, parasitic lysate and parasites, coinciding with a decrease in protein synthesis. Further, this eIF2\textalpha{} phosphorylation could be partially reversed by hemin. Hemin inhibited both autophosphorylation and substrate phosphorylation events. Eventually,
PfPK4 gene was cloned from the human malarial parasite *Plasmodium falciparum* by Mohrle *et al.* (1997). This gene encodes a protein of a predicted length of 1123 amino acids containing all the conserved regions characteristics of Ser/Thr protein kinases. The catalytic kinase domain possesses highest sequence identity (34-37%) with eIF2α kinases, especially HRI protein kinases. There are two kinase inserts in PfPK4, located at positions common to eIF2α kinases. The first insert separates kinase subdomains IV and VI by 559 amino acids, and the second insert separates subdomains VII and VIII by 41 amino acids. Both inserts are larger than their homologues in eIF2α kinases. The sequence of PfPK4 has one putative hemin-binding site. PfPK4 protein is expressed as two major forms of 80 and 90 kDa. Whereas the 80 kDa form is present throughout the intra-erythrocytic development and in merozoites, the two 90 kDa forms are only found in mature parasites. Thus, this eIF2α kinase helps the parasite in sensing the presence/absence of heme in the environment during invasion.

**TgIF2K-A:** This eIF2α kinase was first identified and characterized in *Toxoplasma gondii* by Sullivan’s group (Sullivan *et al.*, 2004), and designated as TgIF2K-A (*Toxoplasma gondii* initiation factor-2 kinase). They demonstrated that TgIF2K-A gets phosphorylated in response to heat shock and pH stress, and is known to induce differentiation to bradyzoites *in vitro*. Although the catalytic domain of TgIF2K-A contains sequence and structural features that are conserved among members of the eIF2 kinase family, TgIF2K-A has an extended N-terminal region that is highly divergent from other eIF2 kinases. TgIF2K-A participates in the adaptive response of the parasite to stress, as well as it contributes to the clinically relevant process of parasite differentiation to bradyzoites (Sullivan *et al.*, 2004).

**IfkA:** Two eIF2α kinases were identified in *Dictyostelium*. Neither of these two eIF2α kinases appeared to be involved in sensing amino acid starvation to initiate
development. However, one of the kinases, IfkA, was shown to phosphorylate eIF2α from 1 to 7 hours after the onset of development, resulting in a shift from polysomes to free ribosomes for bulk mRNA. IfkA is required for proper timing of aggregation and regulation of mound size in *Dictyostelium* (Fang *et al.*, 2003).

2. **Protein structure determination- a challenge in protein chemistry**

The basic problem in chemistry is to determine the structure of a molecule and to relate the properties and behaviour of the molecule to its structure. The detailed characterization of the structure, dynamics and folding process of a protein is crucial for understanding the biological functions it performs. The protein chemist therefore has the job, first, of determining the arrangements of the atoms within protein molecules, and then of understanding why these arrangements make it possible for proteins to do the things that they do (Anfinsen, 1972). Thus, the problem of protein structure determination has two aspects: (a) the determination of the structural formula and (b) the determination of molecular conformation. Starting with the work of Fisher and Hofmeister in 1902, great progress has been made in the determination of the classical organic structural formulas of proteins. In order to write the structural formula of a protein, it is necessary to determine the number of amino acid monomers that are incorporated into the protein molecule, and also the sequence in which they are joined together. Once this is done the next important aspect of the protein structure is determination of three-dimensional conformation of protein. The most direct method of dealing with this problem is by the method of X-ray crystallography. This method is, however, limited to study of molecules in the crystalline state, and therefore, protein chemists are also interested in developing other tools for investigating the structures of molecules, particularly in solution. Since most of the chemical processes of living
systems go on in solution, and not in the crystals, structural tools that can be used for solutions are especially interesting to biochemists. Several such tools are available, in addition to X-ray diffraction, for the study of the molecular structure of proteins. None of these other tools is capable of giving the detailed structural information that can be obtained using X-ray diffraction studies on crystals, but many of them have the advantage that they can be applied to proteins in solution and particularly to proteins for which crystals are not grown. Furthermore, the results are much easier to use and to interpret than is X-ray diffraction (Kauzmann, 1964).

Studies of protein molecules generally fall into one of the three broad categories: formation of polypeptide chain assembly from free amino acids, the folding of the chain into a unique three-dimensional object and the relationships between detailed geometry in solution and biological function. The last two, together with information on the molecular mechanism of enzymic catalysis, constitute the rationale of protein chemistry. Therefore the ultimate aim of the enzymologist and the protein chemist is to be able to synthesize an amino acid sequence that, when allowed to fold, will assume a stable predesigned three-dimensional arrangement of atoms capable of carrying out the desired catalytic act (Anfinsen, 1972).

The protein folding problem is the question of how a protein’s amino acid sequence dictates its three-dimensional atomic structure. The notion of a folding “problem” first emerged around 1960, with the appearance of the first atomic-resolution protein structures. Some form of internal crystalline regularity was previously expected (Kendrew, 1961), and α-helices had been anticipated by Linus Pauling and colleagues (Pauling and Corey, 1951; Pauling et al., 1951), but the first protein structures of the
globins had helices that were packed together in unexpected irregular ways. Since then, various aspects involved in protein folding have been treated as three different problems: (a) the folding code: the thermodynamic question of what balance of interatomic forces dictates the structure of the protein, for a given amino acid sequence; (b) protein structure prediction: the computational problem of how to predict a protein’s native structure from its amino acid sequence; and (c) the folding process: the kinetics question of what routes or pathways some proteins use to fold so quickly (reviewed in Dill et al., 2008). A well accepted model for describing protein functioning is the “Protein Quartet” model (Fig. 21), which attributes the biological activity of the protein to four unique conformations of the polypeptide chain (ordered forms, molten globules, premolten globules, and random coils) and transitions between any of them (reviewed in Uversky, 2002).

Fig. 21. (A) Protein Quartet Model proposed for generalization of the structure-function paradigm (Uversky, 2002). The ordered state is the natively folded structure of a protein that has a well-defined secondary and tertiary structure. Molten globule states are intermediates in the protein folding pathway with compact structures that exhibit a high content of secondary structure, non-specific tertiary structure and significant structural flexibility. Pre-molten globules are condensed, but not compact, and are the forms in which the unfolded chain accumulates before jumping over its folding barrier. Random
coils are highly unstructured protein denatured states. (B) A schematic energy landscape view of protein folding. The surface of the funnel represents a whole range from the multitude of denatured conformations to the unique native structure (Dill and Chan, 1997). (Adapted from Mohan and Hosur, 2009).

The structural know-how about the protein folding processes and the folded native protein are crucial aspects in understanding protein chemistry. Thus, the structural studies which involve solution studies and atomic-resolution studies by X-ray crystallography form an important source of information which provides insights into various aspects of structure-function relationship of proteins. The study presented in this thesis was carried out to look into the structural features of two important proteins: heme regulated eukaryotic initiation factor 2 alpha kinase (HRI) and its substrate; eukaryotic initiation factor 2 alpha (eIF2α) that take part in the control of translation initiation process during protein synthesis.

3. The Present Investigation

Various biochemical and biophysical studies from our laboratory and other laboratories have provided valuable information regarding the heme-regulated eIF2α kinase (HRI) and its substrate (eIF2α): its expression, regulation and activation. Crystal and solution structures of eIF2α have described the functioning of the substrate in detail (Nonato et al., 2002; Ito et al., 2004; Dhaliwal and Hoffman, 2003). However, these studies did not provide much information regarding structural transitions occurring under various biochemical conditions. No structural data to comprehend the mechanism of functioning of HRI is available till date. Thus biophysical characterization of HRI is indispensable to correlate the structure-function relationship of the kinase and its substrate. In order to perform the structural characterization of HRI and its substrate, we
undertook this study. Our main objective is to analyze the conformational transitions occurring under various denaturing/renaturing conditions and also determine the overall structure of these two proteins. The specific objectives of this study are as follows:

1. Cloning, overexpression and purification of HRI using rabbit HRI cDNA
2. Structural studies of HRI
3. Cloning, overexpression and purification of human eIF2α using K562 cells
4. Structural studies of heIF2α