CHAPTER 1

Introduction - Structural and functional diversities between members of the human HSPH, HSPA and DNAJ chaperone families

Jurre Hageman and Harm H. Kampinga

Department of Cell Biology, Section of Radiation and Stress Cell Biology, University Medical Center Groningen, University of Groningen, the Netherlands.

Parts of this chapter have been used for:

* Equal contribution
Abstract

Heat shock proteins (HSPs) were originally identified as stress-responsive proteins required to deal with proteotoxic stresses. Besides being stress-protective and possible targets for delaying progression of protein folding diseases, mutations in chaperones also have been shown to cause disease (chaperonopathies). The mechanism of action of the “classical”, stress-inducible HSPs in serving as molecular chaperones preventing the irreversible aggregation of stress-unfolded or disease-related misfolded proteins is beginning to emerge. However, the human genome encodes several members for each of the various HSP families that are not stress-related but contain conserved domains. Here, we have reviewed the existing literature on the various members of the human HSPH (HSP110), HSPA (HSP70), and DNAJ (HSP40) families. Apart from structural and functional homologies, several diversities between members and families can be found that not only point to differences in client specificity but also seem to serve differential client handling and processing. How substrate specificity and client processing is determined is far from being understood.
Introduction and aim of the thesis

Heat shock proteins (HSPs) were originally discovered as proteins that are upregulated upon and protective against proteotoxic stresses, i.e., situations that increase the fraction of proteins that are in a (partially) unfolded state, thereby enhancing their probability of forming intracellular protein aggregates that can lead to loss of cell function and eventually to cell death. We now appreciate that a variety of normal cellular processes (translation, transport over membranes) constantly challenge the cellular protein homeostasis and require protein quality control systems for assistance. In addition, diseases like Alzheimer’s and Parkinson’s disease, CAG-repeat diseases, and many heart diseases (e.g., atrial fibrillation) or physiological disturbances (e.g., hypoxia) are pathogenic because they disturb protein homeostasis. Finally, folding mutations may arise as a result of somatic mutations (aging) and genomic instability (cancer) requiring increased protein quality control.

The HSPs make up a group of structurally unrelated protein families (HSPA, HSPB, HSPC, HSPD, HSPH, and DNAJ) that play a prime role in protein homeostasis by binding to substrates at risk, thereby keeping them in a state competent for either refolding or degradation. As such, they belong to a much larger superfamily of molecular chaperones. The number of genes encoding the diverse HSP family members largely varies per organism. For HSPA, the number varies from three in *Escherichia coli*, 14 in Arabidopsis thaliana, and 12 in *Drosophila melanogaster* to 13 in *Homo sapiens*. For small HSP (sHSP), the number of genes is relatively high in plants and the same holds true for DNAJ (Table 1). Although originally identified as heat inducible proteins, many members are in fact not heat shock inducible. However, within each family, individual heat shock inducible proteins such as HSPB1, HSPA1, HSPH1, and DNAJB1 are found. Why the human genome contains so many members in most families (Table 2) sometimes with a high degree of sequence homology (HSPA and HSPH) but sometimes also with substantial sequence divergence in certain domains (HSPB and DNAJ) is largely unclear. Part of the redundancy may relate to the intra-compartmental distribution of the diverse family members and the requirement of their activities in these different compartments. Also, some HSPs exhibit tissue or development specific expression (Table 3). On one hand, this may reflect the ability to specifically regulate expression of the same activity and function. On the other hand, this suggests that chaperones are not merely promiscuous in terms of clients and indicates a strong need for specialized chaperones under these conditions. The fact that promiscuity may not be an essential feature of HSP/chaperone activity is further supported by the existence of different HSP families and family members within the same compartment (e.g., the cytosol) (Table 3). In this review, we focus on the structural, sequence, and functional divergence within the HSPA, HSPH and DNAJ families either in terms of client specificity or client processing.

The classical model of the HSP chaperone functions is primarily based on cell-free experiments with human HSPA1A/B (or HSPA8), HSPB1, HSPH1, and DNAJB1 and work on their orthologues in *E. coli*, yeast, and mouse. In this model, un- or misfolded proteins bind to these HSPs both directly or sequentially (Figure 1). In naive cells that were not stressed before, the instant chaperone action toward increases in proteotoxic stresses is obtained by constitutively expressed members such as cellular stores of small heat shock proteins (HSPB). It is thought that HSPB members are stored in oligomeric complexes that dissociate into smaller-sized molecules upon stress (1:2). This shift in oligomeric size allows for binding of unfolded proteins, which effectively neutralizes the chance for nonspecific interaction of the unfolded substrate with other proteins. HSPB members are ATP-independent chaperones and require other partners for further client processing (see below), which depending on the substrate...
and/or chaperone partners can be either refolding or degradation. How this distinction in client processing is regulated is not yet clear. Clearly, transfer to the HSPA machine has been suggested to promote folding in vitro as well as in living cells. This ATP-dependent chaperone constantly shuttles between an ATP-bound and an ADP-bound state in which it has different affinities for unfolded proteins (3). Substrates enter the HSPA complex in the ATP-bound configuration. In this configuration, HSPA has a high substrate on/off rate, meaning low substrate affinity. Upon binding, ATP is hydrolyzed which stabilizes the affinity of HSPA for its substrate, a reaction which is regulated by cofactors like DNAJs and CHIP. Subsequently, nucleotide exchange is stimulated (BAG-1, HSPBP1, and HSPH), resulting in an ATP-bound HSP70 complex followed by substrate release. Unfolded proteins may also directly enter the HSPA chaperone machine or with the assistance of HSPA cochaperones like DNAJ and HSPH. In fact, there are several modulators of the HSPA ATP cycle (HSPH, DNAJ, HIP, CHIP, BAG3, and HSPBP1), which not only modulate the cycle but also may confer client specificity to the HSPA machine and/or affect the fate of its client.

The HSP70 chaperoning machine

General introduction

Whereas, the HSPB chaperones (HSPB1, HSPB4, HSPB5) may form the primary line of defense under stress, the HSPA chaperone machine, especially HSPA1A/B in conjunction with its co-chaperone DNAJB1, are the strongest stress-inducible proteins. As an ATP-dependent chaperone machine, they can act on substrates bound to HSPB oligomers after being induced (see Figure 1). However, several members, including some stress induced members, are also expressed under non-stress conditions (HSPA8, DNAJA1, HSPA1A/B and DNAJB1). As such (or when induced) they can also directly interact with clients without collaborating with HSPB chaperones.

With the sequencing of the human genome and the computational annotation of its genes, it became apparent that the human genome encodes 13 different HSPA members, 4 HSPH members and 41 different DNAJ members (4;5). The majority of these HSPA and DNAJ proteins are thought to reside in the cytosol which suggests that chaperones have evolved with specialized functions.

HSPA proteins are highly homologous to the 4 members of the HSPH (HSP110) family. HSPH members differ from HSPA members by the presence of an extended linker domain of unknown function (6). In fact, HSPA4 and HSPA4L, are currently annotated as HSPA members in the NCBI gene database, but are more homologues to HSPH1 and therefore here referred to as HSPH2 and HSPH3. In addition, a 4th HSPH member (Grp170) is present in the ER and here referred to as HSPH4 (6).
Introduction and aim of the thesis

Typically, HSPA proteins consist of a N-terminal ATPase domain of 45 kDa and a C-terminal substrate binding domain of 25 kDa. The N-terminal domain can be subdivided in two large globular subdomains which are separated by a deep cleft where the nucleotide binds. The C-terminal domain can be subdivided into a binding pocket which has a beta sandwich structure and a highly variable α-helical subdomain. The ATPase and the C-terminal domain are separated by a small linker domain for HSPA members and a longer linker domain for HSPH members (Figure 2) (3). Although the details remain to be elucidated, this linker domain couples the nucleotide hydrolysis to the opening and closing of the substrate binding cavity (7). Why HSPH proteins have an extension of the linker domain is currently unknown. Clearly, like HSPA members, HSPH proteins can bind substrates. However, alone they cannot release substrates, which is typical for HSPA members.

Several important sites within the above described domains of the HSPA protein have been mapped with high precision. Different co-factors bind to different domains of the HSPA protein:

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Protein Name</th>
<th>Alternative Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSPA1</td>
<td>HSPA1</td>
<td>Heat-shock 70kD protein-2</td>
</tr>
<tr>
<td>HSPA2</td>
<td>HSPA2</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA3</td>
<td>HSPA3</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA4</td>
<td>HSPA4</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA5</td>
<td>HSPA5</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA6</td>
<td>HSPA6</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA7</td>
<td>HSPA7</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA8</td>
<td>HSPA8</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA9</td>
<td>HSPA9</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA10</td>
<td>HSPA10</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA11</td>
<td>HSPA11</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA12</td>
<td>HSPA12</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA13</td>
<td>HSPA13</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>HSPA14</td>
<td>HSPA14</td>
<td>Heat-shock 70kD protein 6 (HSP70B')</td>
</tr>
<tr>
<td>DNAJ1</td>
<td>DNAJ1</td>
<td>DNAJ1</td>
</tr>
<tr>
<td>DNAJ2</td>
<td>DNAJ2</td>
<td>DNAJ2</td>
</tr>
<tr>
<td>DNAJ3</td>
<td>DNAJ3</td>
<td>DNAJ3</td>
</tr>
<tr>
<td>DNAJ4</td>
<td>DNAJ4</td>
<td>DNAJ4</td>
</tr>
<tr>
<td>DNAJ5</td>
<td>DNAJ5</td>
<td>DNAJ5</td>
</tr>
<tr>
<td>DNAJ6</td>
<td>DNAJ6</td>
<td>DNAJ6</td>
</tr>
<tr>
<td>DNAJ7</td>
<td>DNAJ7</td>
<td>DNAJ7</td>
</tr>
<tr>
<td>DNAJ8</td>
<td>DNAJ8</td>
<td>DNAJ8</td>
</tr>
<tr>
<td>DNAJ9</td>
<td>DNAJ9</td>
<td>DNAJ9</td>
</tr>
<tr>
<td>DNAJ10</td>
<td>DNAJ10</td>
<td>DNAJ10</td>
</tr>
<tr>
<td>DNAJ11</td>
<td>DNAJ11</td>
<td>DNAJ11</td>
</tr>
</tbody>
</table>

Table 2: Alternative names and gene identifiers of the human HSP gene families

Structural features, oligomeric dynamics and functional regulation.

Typically, HSPA proteins consist of a N-terminal ATPase domain of 45 kDa and a C-terminal substrate binding domain of 25 kDa. The N-terminal domain can be subdivided in two large globular subdomains which are separated by a deep cleft where the nucleotide binds. The C-terminal domain can be subdivided into a binding pocket which has a beta sandwich structure and a highly variable α-helical subdomain. The ATPase and the C-terminal domain are separated by a small linker domain for HSPA members and a longer linker domain for HSPH members (Figure 2) (3). Although the details remain to be elucidated, this linker domain couples the nucleotide hydrolysis to the opening and closing of the substrate binding cavity (7). Why HSPH proteins have an extension of the linker domain is currently unknown. Clearly, like HSPA members, HSPH proteins can bind substrates. However, alone they cannot release substrates, which is typical for HSPA members.

Several important sites within the above described domains of the HSPA protein have been mapped with high precision. Different co-factors bind to different domains of the HSPA protein:
DNAJ binds to both the ATPase as well as the extreme C-terminal domain, whereas BAG-1 and HIP bind only to the ATPase domain and other co-factors such as Hop, Tpr-2, CHIP and DNAJ bind to the C-terminal domain (Figure 2) (8).

Interestingly, many cytosolic HSP70s from eukaryotes terminate in the EEVD motif (HSPA1A, HSPA1B, HSPA1L, HSPA2, HSPA6 and HSPA8) (9) while the HSPH proteins show deviations from this motif. The ER chaperones HSPA5 and HSPH4 terminate in the ER retention signals KDEL and NDEI respectively. Deletion or mutation of the EEVD motif causes a disruption of the interdomain communication resulting in an enhanced intrinsic ATPase activity (10), leading to a reduced binding of substrates. Also, this deletion resulted in complete abrogation of DNAJ binding to and regulation of HSPA (11).

A couple of (stress-dependent) posttranslational modifications have been reported for HSPA proteins such as phosphorylation (12;13) and acetylation (14) but the effects of these modifications on HSPA functionality remain largely unknown.

**Regulation of the HSPA chaperone cycle by co-factors**

HSPA proteins constantly shuttle between an ATP and an ADP bound state (Figure 1). Several co-factors have been identified that influence this process. DNAJ members stimulate the hydrolysis of ATP (15). HIP stabilizes the ADP bound HSPA complex (16) whereas HSPH (17;18), BAG-1 (19) and HSPBP1 (20) stimulate the ADP to ATP nucleotide exchange. The E3 ubiquitin ligase co-factor CHIP can inhibit the ATP hydrolyzing capacity of HSPA (21). Another co-factor, HOP is not acting on the HSPA ATPase cycle as such, but it is thought to link the HSPA chaperone to the HSP90/HSPC chaperone complex (22).

It has been suggested that the co-chaperones may determine the fate of a HSPA bound substrate. Co-factors such as DNAJ or HIP would tip the balance to refolding whereas BAG-1 and CHIP may specifically stimulate client degradation (23). Such a model is supported by the fact that BAG-1 has a ubiquitin-like domain and CHIP acts as an E3-ubiquitin ligase. However, CHIP shows a high degree of substrate specificity suggesting that this idea can not be generalized for all substrates. Moreover, in how far its action on the HSPA cycle is directly linked to its E3-ligase activity remains unclear. Clearly, CHIP can act on HSPA without a functional E3-ligase domain (24) and inversely, it can target proteins for degradation without its domain needed for HSPA interaction (25). Thus although a co-factor such as CHIP is probably not involved in determining the fate of the client, it provides the cell with a substrate binding co-factor capable of proteasomal targeting and hence, the clearance of nonfoldable substrates.

But, there are more arguments against the idea that co-chaperones determine the fate of HSPA-bound clients. For instance, it has been shown that DNAJB1 can stimulate client refolding (26;27) as well as client degradation (28). So, rather the state of the protein and its ability to correctly refold may determine its fate and HSPA may just provide a time window for this to occur, whilst preventing the irreversible aggregation of its client.

Interestingly, all of the HSPA co-factors except BAG-1 and possibly some DNAJ type 3 proteins, also can bind to unfolded substrates themselves by which they may confer specificity to the machine.

**HSPA family**

**Paradigm according to HSPA1A/B and HSPA8**

Crystallographic evidence shows that both HSPA and DNAJ resemble molecular clamps (29). However, the way by which they bind substrates is significantly different. The substrate bind-
The binding domain of HSPA proteins consists of a short beta sandwich motif which can be locked by the alpha-helical lid structure. The beta-sandwich contains a hydrophobic core of 4-5 amino acids with two flanking basic residues (30). Only a short linear polypeptide fits within the substrate binding domain of the monomeric HSPA protein and therefore it binds only a short stretch of peptides. Because HSPA recognizes very short hydrophobic peptides, it is thought that it can bind a wide variety of substrates. It must be mentioned, however, that these findings are all based on a small number of proteins like *E. coli* DNAK and DNAJ, Bovine HSP70/HSPA1A, and yeast DNAJ members Ydj1 and Sis1. As both the HSPA and the DNAJ families are quite diverse, it is likely that the proposed models do not account for all possible cellular HSPA/DNAJ complexes and that some of the different cytosolic HSPA complexes are adapted for a limited range of cellular substrates.

**Functional diversities and co-partners specificity**

It is yet unclear whether all HSPA members have the same mode of action, whether they are functionally interchangeable, if they act at different microenvironments or whether they require all the same co-factors. Features of different HSPA members are summarized in Table 3.
HSPA1A and HSPA1B differ by only two amino acids and are believed to be fully interchangeable proteins. This is confirmed by studies in knockout mice which only show a phenotype when both alleles were deleted (31). Both proteins have been referred to as HSP70i (or HSP72) (Table 2) and are the strongest stress-inducible HSPA members. They have been extensively investigated, and details are reviewed elsewhere (3). HSPA8, the cognate HSPA and previously referred to as Hsc70 (or HSP73) is expressed in all cell types. It is considered to be the essential “house-keeping” HSPA member that is involved in co-translational folding and chaperoning of proteins that require to be translocated over intracellular membranes or that need to be transferred to HSPC (HSP90). Its mechanism of actions and regulation of the ATP/ADP cycle seems highly similar to that of HSPA1A/B, but also some functional differences have been reported. For instance, the primate HSPA8 but not HSPA1A could complement the growth of a yeast strain which lacks all of the chromosomal HSP70 (SSA) genes. Domain swapping experiments revealed that the ATPase domain of HSPA8 was essential for this action, indicative that specificity may be mediated via specific DNAJ partnerships (32).

HSPA1L and HSPA2 are two cytosolic family members with high expression in the testis and HSPA2 has been shown to be required for spermatogenesis (33;34). Also, HSPA2 is elevated in mammary tumors (35). Their biochemical mode of action is currently unknown, but the fact that they are highly expressed in testis and that male knockout mice for HSPA2 are sterile suggest that they may differ in substrate or function specificity from HSPA1A/B or HSPA8. In line with this, knockdown of HSPA2 resulted in phenotypic and genotypic alterations in HeLa cells that were clearly distinct from those after HSPA1A knockdown (35).

HSPA6 is a poorly studied, stress-inducible protein that is lacking in rodents. It is not expressed under normal conditions and only induced upon severe heat stress, where it is thought to act as a final proteotoxic resistance buffer (36;37). HSPA7 is considered a pseudogene as its transcription product is terminated after 367 amino acids (38). As a complete conserved HSPA protein can originate by bypassing a frame shift at codon position 340 (4) it might also be a true gene which is highly homologues to HSPA6. For both HSPA6 and HSPA7, no data on their mode of action exist.

HSPA9, the mitochondrial HSPA member (HSP75), and HSPA5, the ER localized HSPA chaperone (BiP) are thought to act in a similar manner in their respective compartments as HSPA8 in the cytosol. In line with this, co-factors such as ER (DNAJC1, DNAJB11, DNAJB9 and DNAJC10) and mitochondrial (DNAJA3) specific DNAJs as well as ER (HSPH4/Grp170/HYOU1) and mitochondrial (HMGE) specific nucleotide exchange factors have been found (39;40). A small HSPA-like protein, STCH, which we propose to be referred to as HSPA13 (Table 2), has been found attached to microsomes (41) and may fulfill HSPA8-like actions here.

HSPA12A and HSPA12B are two distantly related proteins found in atherosclerotic lesions. HSPA12A is suggested to be involved in atherogenesis (42) whereas HSPA12B has been studied in zebrafish and it was found to be required for proper vessel formation (43). HSPA14 (HSP70L1) is the smallest HSPA protein and interacts with MPP11, the human ortholog of Zuo1, a cytosolic ribosome-associated chaperone that acts together with Ssz1p and the Ssb proteins in yeast as a chaperone for nascent polypeptide chains during translation (44;45).

So although biochemical details of many HSPA members have yet to be identified, recent systems biology approaches indicated that two distinct chaperone networks with specialized function exist. One molecular chaperone network may protect the proteome against environmental stress (HSP) and the second deals with protein translation (CLIPS) and is associated with ribosomes (46). In line with these findings, one may speculate that HSPA1A/B, HSPA1L and HSPA6/HSPA7 belong to the HSP network whereas HSPA5, HSPA8, HSPA9 and HSPA14 belong to the CLIPS network. However, direct evidence for this is yet lacking.
HSPH family

Paradigm according to HSPH1

In vitro, HSPH1 has been shown to suppress the aggregation of denatured luciferase resulting in enhanced refolding of luciferase. However, rabbit reticulocyte lysate was always required as a source of co-factors after the denaturation to stimulate refolding in these assays, indicating that HSPH members are good suppressors of irreversible aggregation but lack the release activity typical of HSPA proteins and necessary for the stimulation of protein refolding (47;48). Consistently, biochemical evidence from yeast Sse1 and mammalian HSPH2 showed that HSPH proteins are poor ATPases. (18)

Recently, it was found that HSPH members act as nucleotide exchange factors for both mammalian as well as yeast HSPA proteins (17;18). This is surprising as different nucleotide exchange factors (BAG-1, HSPBP1) were already identified in the mammalian cytosol (19;20). Interestingly, none of the nucleotide exchange factors shows significant primary sequence homology to the E. coli nucleotide exchange factor GrpE suggesting that they have evolved independently and it is currently unclear to what extent these nucleotide exchange factors functionally overlap. However, as HSPH proteins are known to have the capacity to hold (unfolded) proteins in a folding competent state (unlike BAG-1 and HSPBP1), they might act as coupling factors between substrate loading and nucleotide exchange for the refolding of specific substrates similar to the coupling of HSPA substrate loading and ATP hydrolysis by DNAJ proteins. Furthermore, the heat inducibility of at least HSPH1, but not BAG-1 or HSPBP1 (49) may serve specific functions (see below).

Like the other nucleotide exchange factors, HSPH members interact with HSPA members in the ADP configuration and stimulate the dissociation of ADP. The subsequent rebinding of ATP induces the dissociation of HSPH-HSPA complexes (18). By stimulating the nucleotide exchange of the HSPA complex, it was shown that HSPH accelerates the HSPA mediated folding of firefly luciferase (17;18). As both a substrate holder and nucleotide exchange factor, HSPH may be particularly relevant under (heat) stress conditions during which it may hold substrates (like HSPB) which can be passed on to HSPA for further handling after the stress. Interestingly, HSPH1 was already identified as a heat shock protein in the early 1980s and was shown to translocate to the nucleolus upon heat stress (50). The heat-inducible HSPA1 and DNAJB1 also translocate to the nucleolus upon heating (51;52) and we previously showed that the nucleolus functions as a temporal reservoir for nuclear unfolded proteins for HSPA1A/B dependent refolding (53). Although purely speculative, it is tempting to suggest that HSPH could fulfill the function of a “holder” chaperone in the nucleolus maintaining heat unfolded nuclear proteins in a folding competent state during stress conditions that, upon return to non-stressful situations, are transferred to HSPA for further processing. Such a model would imply that substrate binding and nucleotide exchange factor function have different temperature dependence, which remains to be demonstrated.

Functional diversities

Besides HSPH1 (HSP110), there are 3 other HSPH members in humans: HSPH2 (HSPA4/APG-2), HSPH3 (HSPA4L/APG-1) and HSPH4 (HYOU1/Grp170). Features of the different HSPH members are summarized in Table 3. Currently, it is unknown to what extent the different HSPH members functionally overlap. While one of them, HSPH4, the grp170 orthologue, is found in the ER (54) where it is likely to fulfill the role of a nucleotide exchange factor for HSPA5, the other 2 family members are found in the cytosolic/nuclear compartment. HSPH1 and HSPH2 are expressed ubiquitously, whilst HSPH3 is mainly expressed in testis
and HSPH3 knockout mice show defects in spermatogenesis (55), suggesting an unique role for this protein within the testis, maybe in conjunction with HSPA1L or HSPA2. That there also could be functional diversity between the human HSPH members may indirectly be deduced from what has been found for their orthologues (SSE) in yeast, where sse1Δ mutants are temperature sensitive and slow growing while sse2Δ mutants show no detectable phenotype (56).

The DNAJ superfamily

All eukaryotic cells contain DNAJ proteins which are known to stimulate the ATPase domain of HSPA chaperones. The common domain that defines this family is the J-domain that stimulates the HSPA ATPase domain. In the human genome, a striking number of 41 different DNAJ encoding genes have been identified (5). The exact protein partners of the different DNAJ proteins as well as the exact cellular functions are currently unknown for most of its members. DNAJ proteins are divided in three sub families (Figure 2); Type A proteins are the closest human orthologues of the E. coli DNAJ and contain, besides an extreme N-terminal J-domain, a glycine/phenylalanine-rich region, a cysteine rich region, and a variable C-terminal domain. Type B proteins contain all the above listed domains with the exception of the cysteine rich region and type C DNAJ proteins contain only the J domain that is not necessarily restricted at the N-terminus but can be positioned at any place within the protein (15).

Paradigm according to DNAJB1

The J-domain is highly conserved and folded in an α-helical secondary structure. A conserved sequence motif (HPD) in the J domain has been shown to be critical for accelerating the ATPase activity of HSP70. Adjacent to the J domain, a Glycine/Phenylalanine rich region is believed to function as a flexible spacer that separates the N-terminal J domain from the rest of the molecule. In the centre of the molecule, the Cysteine-rich domain contains four cysteine-rich repeats that fold around two zinc atoms. The C-terminal domain folds in a β-plated sheet structure and is involved in dimerization as well as in substrate binding and presentation (15).

In contrast to HSPA, which binds substrates as a monomer, DNAJ proteins dimerize in a V like structure (29;57;58). The binding motif of E. coli DNAJ consists of a hydrophobic core of eight residues enriched for arginine, aromatic amino acids and large aliphatic hydrophobic residues positioned in the middle of each monomer (59). Although each monomer contains only a short binding motif, dimer formation gives rise to a relatively large beta sheet projection. Each of the monomers binds part of the unfolded substrate and hold it in an extended conformation between the middle of the two monomers. HSP70 binds the DNAJ dimer at the tips of the V-like structure and takes over the substrate for binding and release cycles (58). Although both DNAJA and DNAJB members are known to dimerize, structural analysis of DNAJA1 and DNAJB4 revealed that there are significant structural differences between DNAJA and DNAJB members. DNAJA1 forms compact dimers in which the N and C-termini face each other. In contrast, DNAJB4 forms a dimer in which only the C-termini of the two monomers were in contact (60). These structural differences may very well relate to differences in substrate binding or selection.

Functional diversities

Informative data on the functional overlap and specificity between DNAJ members comes from detailed analyses in yeast in which 13 cytosolic J-proteins were compared (61). Interestingly, the J domains from a variety of J proteins from different classes of J proteins could com-
### Table 3: Properties of the human HSP families

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence identity</th>
<th>Molecular size (kDa)</th>
<th>Chromosome location</th>
<th>Tissue distribution</th>
<th>Subcellular localization</th>
<th>Clients/ Substrates</th>
<th>Disease association</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSPH1</td>
<td>100%</td>
<td>96.9</td>
<td>13q12.3</td>
<td>ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPH2</td>
<td>63%</td>
<td>94.3</td>
<td>5q31.1-5q31.2</td>
<td>ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPH3</td>
<td>58%</td>
<td>94.5</td>
<td>4q28</td>
<td>testis, brain, kidney, liver, lung, spleen</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPH4</td>
<td>25%</td>
<td>111.3</td>
<td>11q23.1</td>
<td>ubiquitous</td>
<td>ER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA1A</td>
<td>100%</td>
<td>70.0</td>
<td>6p21.3</td>
<td>ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA1B</td>
<td>99%</td>
<td>70.0</td>
<td>6p21.3</td>
<td>ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA1L</td>
<td>88%</td>
<td>70.4</td>
<td>6p21.3</td>
<td>testis</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA2</td>
<td>83%</td>
<td>70.0</td>
<td>14q24.1</td>
<td>testis/ ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA5</td>
<td>60%</td>
<td>71.0</td>
<td>9q33-q34.1</td>
<td>ubiquitous</td>
<td>ER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA6</td>
<td>81%</td>
<td>71.0</td>
<td>1q23</td>
<td>Brain, Liver, Ovary, Saliva</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA7</td>
<td>ND</td>
<td>ND</td>
<td>1q23.3</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA8</td>
<td>85%</td>
<td>70.9</td>
<td>11q24.1</td>
<td>ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA9</td>
<td>45%</td>
<td>73.7</td>
<td>5q31.1</td>
<td>B Cell Brain Liver Ovary Platelet Saliva</td>
<td>cytosol / nucleus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA12A</td>
<td>14%</td>
<td>141.0</td>
<td>10q26.12</td>
<td>endothelia, brain, heart, kidney, muscle, testis</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA12B</td>
<td>18%</td>
<td>75.7</td>
<td>20p13</td>
<td>endothelia, ubiquitous</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA13</td>
<td>20%</td>
<td>51.9</td>
<td>21q11</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSPA14</td>
<td>27%</td>
<td>54.8</td>
<td>10p14</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJA1</td>
<td>100%</td>
<td>44.9</td>
<td>9p13.1-p12</td>
<td>ubiquitous</td>
<td>cytosol / promiscuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJA2</td>
<td>54%</td>
<td>45.7</td>
<td>16q11.1-11q11.2</td>
<td>brain, heart, kidney, liver</td>
<td>cytosol / promiscuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJA3</td>
<td>24%</td>
<td>52.5</td>
<td>16p13.3</td>
<td>mammary gland, B cell</td>
<td>cytosol / promiscuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJA4</td>
<td>73%</td>
<td>44.7</td>
<td>15q24.1</td>
<td>brain</td>
<td>mitochondrial proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJB1</td>
<td>100%</td>
<td>38.2</td>
<td>19p13.2</td>
<td>ubiquitous</td>
<td>cytosol / promiscuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJB2</td>
<td>27%</td>
<td>35.630.6</td>
<td>2q32-q34</td>
<td>heart, muscle, brain</td>
<td>cytosol / ER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJB3</td>
<td>24%</td>
<td>26.7</td>
<td>1p D (MM)</td>
<td>testis</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAJB4</td>
<td>65%</td>
<td>37.8</td>
<td>1p31.1</td>
<td>ubiquitous</td>
<td>unknown</td>
<td></td>
<td>G protein beta subunit</td>
</tr>
<tr>
<td>DNAJB5</td>
<td>62%</td>
<td>39.1/26.9</td>
<td>9p13.2</td>
<td>Brain, Heart, Liver, Pancreas, Skeletal muscle, Spleen</td>
<td>cytosol / nucleus</td>
<td>keratin-18</td>
<td>unknown</td>
</tr>
<tr>
<td>DNAJB6</td>
<td>27%</td>
<td>36.1</td>
<td>7q36.3</td>
<td>ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td>keratin-18</td>
</tr>
<tr>
<td>DNAJB7</td>
<td>24%</td>
<td>35.4</td>
<td>22q13.2</td>
<td>ubiquitous</td>
<td>cytosol / nucleus</td>
<td></td>
<td>keratin-18</td>
</tr>
<tr>
<td>DNAJB8</td>
<td>23%</td>
<td>25.7</td>
<td>3q21.3</td>
<td>testis</td>
<td>cytosol / nucleus</td>
<td></td>
<td>keratin-18</td>
</tr>
<tr>
<td>DNAJB9</td>
<td>19%</td>
<td>25.5</td>
<td>7q31</td>
<td>ubiquitous</td>
<td>unknown</td>
<td></td>
<td>keratin-18</td>
</tr>
<tr>
<td>DNAJB10</td>
<td>23%</td>
<td>30.628.6</td>
<td>1p (Mm)</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td>keratin-18</td>
</tr>
<tr>
<td>DNAJB11</td>
<td>30%</td>
<td>40.5</td>
<td>3q28</td>
<td>ubiquitous</td>
<td>unknown</td>
<td></td>
<td>keratin-18</td>
</tr>
<tr>
<td>DNAJB12</td>
<td>18%</td>
<td>41.9</td>
<td>10q22.2</td>
<td>blood plasma</td>
<td>unknown</td>
<td></td>
<td>APOBEC1</td>
</tr>
<tr>
<td>DNAJB13</td>
<td>48%</td>
<td>36.1</td>
<td>11q3.4</td>
<td>fetus, spermatozoa, testis</td>
<td>unknown</td>
<td>keratin-18</td>
<td>unknown</td>
</tr>
<tr>
<td>DNAJB14</td>
<td>17%</td>
<td>42.533.5</td>
<td>4q2.3</td>
<td>unknown</td>
<td>unknown</td>
<td>keratin-18</td>
<td>unknown</td>
</tr>
</tbody>
</table>

**Table Notes:**
- **17**
- Introduction and aim of the thesis
- Table 3: Properties of the human HSP families
- **Introduction and aim of the thesis**
plement the severe growth defects in yeast lacking the DNAJ protein Ydj1. This demonstrates that the stimulation of the ATPase activity of the yeast HSPA is sufficient for many cellular processes. On the contrary, the phenotypes of 4 other DNAJ deletions (cwc23, sis1, jjj1 and jjj3) could only be rescued by expression of the deleted genes indicating that these proteins fulfill highly specialized and unique tasks.

For mammals, domain swapping experiments have confirmed that the J-domain can be exchanged between various DNAJ proteins with preservation of the biological function. For instance, a SV40 large T antigen fusion that contains the J domain from DNAJB1 is still functional in SV40 large T mediated p130 phosphorylation (62). Even more, the J domain from DNAJB1 could complement the J domain of yeast Ydj1 (63). This all indicates that the J-domain is highly conserved and not likely responsible for any functional diversity. Rather, it suggest that the J-domain may function only to recruit HSPA members to functions the specificity of which is determined in the C-termini of the DNAJ-members.

To date, it is also not known whether all DNAJ members act to stimulate the nucleotide cycle of HSPA machines. Clearly, in vitro, type A DNAJ molecules have substrate binding activity (59) and thus may function in substrate delivery to HSPA partners. It is still under debate if type B members also exhibit substrate binding activity. Whether DNAJ members have any preference for specific partners within the HSPA family is also yet unclear.

The DNAJA Family

As mentioned above, this family contains all the different identified domains and most closely resembles the E. coli DNAJ protein. The human genome contains four different members of the DNAJA family and general features are summarized in Table 3. Interestingly, it was found that whereas both DNAJA2 and DNAJA4 could stimulate the hydrolysis of ATP on HSPA1A, the DNAJA2-HSPA1A but not the DNAJA4-HSPA1A combination was able to support refolding of denatured luciferase. This indicates that at least DNAJA2 and DNAJA4 may have differential substrate specificity in that DNAJA4 is unable to bind and load denatured luciferase onto the HSPA1 chaperone (64).

The DNAJB Family

Although DNAJA1 shows the highest homology with the well-studied E. coli DNAJ, it is the DNAJB family (containing an N-terminal J-domain and a glycine/phenylalanine-rich region but lacking the zinc binding cysteine rich region), DNAJB1 in particular, that has been most extensively studied in mammalian cells. This is partly due to the fact that this member is the highest expressed DNAJ member in most if not all mammalian cell types. In addition, DNAJB1 is highly induced upon various stresses including heat shock and has been found to cooperate with both HSPA1A as well as HSPA8 in luciferase refolding in vitro (65) and in living cells (26;27). DNAJB4/HLJ1 and DNAJB5 are two DNAJB proteins with unknown function that are close paralogs of DNAJB1. DNAJB4 shows full length homology to DNAJB1 and is also known as HSC40, a non-heat inducible constitutively expressed member and proposed to act as a housekeeping HSP40/DNAJ protein just as HSC70/HSPA8 is proposed as the housekeeping equivalent of the stress inducible HSPA1.

DNAJB2/HSJ-1 is also relatively well-studied and is expressed as two isoforms. The long isoform is targeted to the cytosolic face of the ER by C-terminal geranylgeranylation while the short isoform is found in the cytosolic and nuclear compartment (66). Both proteins contain an ubiquitin interaction motif which is suggested to sort misfolded clients for proteasomal degradation. For this activity, DNAJB2 was dependent on its ability to cooperate with HSPA8 and a
Introduction and aim of the thesis

critical mutation in the J domain (HPD to QPD) abolished its activity (67). Whether DNAJB2 also can assist in HSP70 dependent protein refolding is currently unknown. The DNAJB members DNAJB9/ErdJ4 and DNAJB11/ErdJ3 are ER specific DNAJ members that collaborate with HSPA5 (68;69). Both proteins are induced upon ER stress (68;69). This implies that DNAJB9 and DNAJB11 likely fulfill equivalent roles to that of DNAJB1 in the cytosol.

DNAJB6/MRJ-1, DNAJB7 and DNAJB8 are three homologous proteins which share, besides a J domain, high sequence homology in the C-terminus. This C-terminal domain does not show any homology with known domains in the Pfam database (data not shown). As mentioned above, DNAJB6 binds keratin 18 (70) and it was found to be important to prevent toxic keratin aggregation which interferes with placental development (71). DNAJB6 has also been shown to suppress the toxic aggregation of mutant Huntington (72). Thus, the DNAJB6 protein seems to act as a generic anti-aggregation protein supporting the idea that (some) DNAJB members can bind substrates. Whereas this holds for DNAJB7 and DNAJB8 as well and whether and how DNAJB6 acts as a co-factor of HSPA in the prevention of protein aggregation remains to be elucidated.

Finally, there are 3 more diverse DNAJB members, DNAJB12, DNAJB14 and a testis-specific DNAJB13/Tsarg3 that have hardly been studied. DNAJB12 and DNAJB14 show a high sequence similarity in the C-terminus and contain the C-terminal DUF1977 (Domain of unknown function 1977) domain. DNAJB13 contains a C-terminal domain homologous to DNAJB1, DNAJB4 and DNAJB5 (Pfam, data not shown). General features of the DNAJB family are summarized in Table 3.

The DNAJC Family

With over 23 members, the DNAJC family represents the largest of the three subfamilies. The family is very diverse in both amino acid composition as well as protein length, the DNAJ domain being the only common feature (Figure 2). Protein length varies between 116 and 913 amino acids. Only a dozen of proteins have been studied. Among them are the DNAJC1/ErDJ1, SEC63/ErDJ2, DNAJC10/ErDJ5 proteins, all ER specific DNAJ proteins. All these proteins have been found to associate with HSPA5 (73-75). It is currently unknown to what extent these HSPA5 factors show functional overlap or specificity.

Another member, DNAJC19/TIMM14 is part of the mitochondrial TIM23 preprotein translocase. DNAJC19 stimulates the ATPase activity of the mitochondrial HSPA protein (HSPA9). The DNAJC19/HSPA9 complex converts the energy of ATP hydrolysis into movement of the unfolded polypeptide chain across the translocation pore resulting in the mitochondrial import of nuclear encoded proteins (76).

DNAJC21/ZRF1 is the orthologue of the yeast DNAJ protein called Zuo1, a ribosome associated DNAJ protein important for translation in yeast. Furthermore, several other DNAJC members such as DNAJC5/CSP (77), DNAJC6/auxilin (78) and DNAJC13/RME-8 (79) seem to function in endocytosis and exocytosis. Both DNAJC6 and DNAJC13 have been shown to collaborate with HSPA8 in the process of endocytosis (80).

Summary and perspectives

The human genome not only encodes a wide variety of HSP families but also a large number of individual proteins within each of these families. While the diversity is starting to be appreciated by a number of investigators, we yet have only faint clues about why such diversity
exists. Housekeeping members may be primarily involved in cotranslational folding of proteins and/or transport of proteins across membranes, whereas some (inducible) members may perform more stress-related functions. Substrate specificity for the HSPA machine(s) may be, in part, evoked by specific cofactors (DNAJJs and HSPH members). The different structures of the DNAJA and DNAJB C-termini may provide such possibilities for client specificities, but whether mammalian cells also make use of specific combinations between HSPA and DNAJ members remains to be elucidated. Determinations of the fate on client processing (toward folding or degradation) may not depend on this HSPA machine as such, although the presence of HSPA cofactors such as CHIP and BAG provides an easy link to the proteasome. Intriguingly, although binding of a client to some HSPB members can result in HSPA-dependent improved folding of (stress-denatured) clients, several HSPB members are linked to client degradation (proteasomal or autophagy) in both HSPA-dependent and -independent manners. The ability of nearly all HSPB members to associate with cytoskeletal elements, increasing their stability, suggests that HSPB members act as cytoskeleton specific chaperones. In addition, association of HSPBs with cytoskeletal elements may allow them to chaperone and transport un- or misfolded cytosolic proteins toward protein storage and/or degradation routes. However, much of this is still only speculative.

The ability of chaperones to handle unfolded proteins has challenged many researchers to test whether they could be used for prevention of the progression of protein folding diseases. Although this concept has had some support from in vitro work, so far studies with animal models have had limited success. More insight into the function of other individual HSP family members may help to elucidate better suppressors and/or strategies for these diseases. On the other hand, the finding that several neuropathies are related to mutations in HSP encoding genes (chaperonopathies) further supports their importance in neurodegeneration. Investigations into the molecular biochemical mechanisms by which these mutations lead to chaperonopathies will lead to an improved understanding of these neuropathologies and also improve our understanding of the normal function of the corresponding chaperones. Furthermore, searches for mutations in other family members or for alternative transcripts of individual HSP genes may lead to identification of causes for nonidiopathic cases of neuropathies or age-related decline in protein quality control and cellular aging.

Figure 2: Linear representation of HSPA/HSPH and DNAJ proteins.
Acknowledgements

This work was supported by Innovatierichtige Onderzoeksprogramma genomics Grant IGE03018.

References

Introduction and aim of the thesis


77. Braun, J. E. and Scheller, R. H. (1995) Cysteine string protein, a DNAJ family member, is
present on diverse secretory vesicles, Neuropharmacology 34, 1361-1369.


Aim of the thesis

In this thesis, we explore the diverse and mostly uncharacterized members of the HSPA and DNAJ families of heat shock proteins for their various activities to combat proteotoxic stresses which are fundamentally implicated in cellular ageing and protein misfolding diseases.

Outline of the thesis

Chapter 2
In this chapter, we used bioinformatics approaches to analyze various aspects of the HSPH/HSPA and DNAJ family such as expression patterns, stress inducibility, subcellular localization, and phylogenetic relationships. Thereafter, we started a systematic cloning approach to build a collection of HSPH/HSPA and DNAJ expression constructs in order to investigate their role in proteotoxic stress resistance.

Chapter 3
Here we build a library of luciferase reporters targeted to various cellular organelles. With this library, we measured the proteotoxic sensitivity of various cellular compartments. In addition, we quantitatively measured which cellular compartments can acquire thermotolerance, a transient condition of cellular resistance which is assumed to be caused by the up regulation of their repertoire of molecular chaperone proteins.

Chapter 4
Although molecular chaperone genes have certainly expanded during evolution as a result of cellular compartmentalization, most members are found in the cytosolic/nuclear compartment and have therefore been expanded for other, yet to be determined, reasons. In this chapter, we focus on the cytosolic and nuclear HSPH, HSPA, DNAJA and DNAJB families and assess its activity in heat denatured protein refolding and the suppression of heat induced protein aggregation.

Chapter 5
Polyglutamine disorders are caused by the progressive accumulation of toxic aggregates in various regions in the brain of affected individuals. Although molecular chaperones have been shown to suppress the progression of the disease, only a small subset of the HSPH/HSPA/DNAJ family has been analyzed. Therefore, we performed a systematic reverse genetic expression screen to search for HSP members with a high activity in the suppression of polyglutamine aggregation and toxicity.

Chapter 6
The results described in the different chapters are summarized and their implications for the research on molecular chaperones discussed. In addition, future perspectives and possible directions are debated.