TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät Chemie

Lehrstuhl für Biotechnologie

Genetic and transcriptional Regulation of HSP-90 and Structural analysis of HSP-90 system in complex with kinase client

Siyuan Sima

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Johannes Buchner

Prüfer der Dissertation: 1. Priv.-Doz. Dr. Klaus Richter

2. Prof. Dr. Matthias Feige

Die Dissertation wurde am 05.08.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 02.10.2019 angenommen.

Publications

Part of this thesis is published or is currently in the process of being published. The equal contribution of authors is indicated with a star (*).

- *Papsdorf K*,* ***Sima S****, Richter G, Richter K.* Construction and evaluation of yeast expression networks by database-guided predictions. Microbial Cell, Vol.3 (No. 6), 236-247 (2016)

- *Eckl Julia*,* ***Sima Siyuan****, Marcus Katrin, Lindemann Claudia, Richter Klaus.* HSP-90-downregulation influences the heat-shock response, innate immune response and onset of oocyte development in nematodes. PLoS One, 12 (10), e0186386 (2017)

- ***Sima Siyuan****, Klaus Richter.* Regulation of the HSP-90 system**.** Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1865 (6), 889-897 (2018)

- ***Sima Siyuan****, Lukas Schmauder, Klaus Richter.* Genome-wide analysis of yeast expression data based on a priori generated co-regulation cliques. Microb Cell. 6(3):160-176 (2019)

- *Katharina Papsdorf*,* ***Sima Siyuan*****, Lukas Schmauder*, Sebastian Peter, Lisa Renner, Patrica Hoffelner, Klaus Richter*. *head-bent* resistant Hsc70 variants show reduced Hsp40 affinity and altered protein folding activity. Scientific reports, 11955(9), (2019)

- (in preparation)
  Sima Siyuan*, Katalin Barkovits, Stephan Hacker, Nils Hellwig, Nina Morgener & Klaus Richter*. sB-Raf complexes with the HSP-90 system are stabilized by the presence of the large PPIase FKB-6.

Content

Publications 4

Abstract 8

Zusammenfassung 10

1 Introduction 13

1.1 HSP-90 and its chaperon system* 13

1.2 The HSR regulating transcriptional factor: HSF-1 22

1.3 Kinase client of HSP-90 and the role of HSP-90 in the regulation of their clients 23

1.4 Cellular expression response of yeast 24

2 Aim of the work 26

3 Material and methods 28

3.1 Materials used in the thesis 28

3.2 Experimental Methods 40

4 Results 54

4.1 Genome-wide analysis of yeast expression data based on co-expression clusters* 54

4.2 HSP-90 down regulation has an impact on the heat shock response, innate immune response and intestinal development of the nematodes.* 67

4.3 Binding of the HSF-1 DNA-binding domain to multimeric *C.elegans* consensus HSEs is governed by cooperativity. 81

4.4 sB-Raf complexes with the HSP-90 system are stabilized by the presence of KFB-6 92

4.5 Oncogenic sB-Raf mutants alter its kinase stability and the ATP-dependent release from the HSP-90 chaperon system. 106

5 Discussions 117

5.1 Yeast genome-wide expression data analysis* 117

5.2 Consequences of the HSP-90 down regulation * 120

5.3 Interaction of HSF-1 DBD and regulation of HSF-1's transcriptional activity in *C.elegans* 123

5.4 Cochaperon FKB-6 contributes to the stabilization of kinase•HSP-90 complex 125

6 Supplemental figures 130

7 References 153

8 Acknowledgement 161

9 Eidesstattliche Erklärung 162

Abstract

With the fast development of technologies, biochemical researches focused on cellular signaling and regulation has been expanded from tracking single signal pathways to generating a massive overview of all the involved parameters and analysis of the hidden relationship among these processes. In this thesis, starting with the analysis of yeast expression data, genome wide changes are observed and genes are clustered based on co-regulated expression changes. The genes were assigned to cliques based on co-regulation expression level by applying an unbiased classification algorithm. The whole genome was involved without losing any non-top hits. The visualization of these obtained clusters simplifies the understanding of significant genome-wide regulation changes.

As one of the key molecules in regulation and maturation of kinases and transcription factors, HSP-90 is especially focused in the rest part of this thesis. The changes in expression level induced by HSP-90 down regulation in *C. elegans* are analyzed with the help of microarrays, which reveals that the innate immune response pathway and genes related to life stage development are mainly influenced. The responding factors match on the gene level as predicted from the expression data analysis.

Under stress conditions the expression level of HSP-90 is elevated to assist cells to fight against damages induced by unfolded or misfolded proteins. The regulating pathway is triggered by the transcription factor HSF-1. As the initial step of HSP-90 regulation, HSF-1 is investigated in terms of DNA recognition and binding. To what extent the sequence of heat shock element influences the occupancy of heat-shock promoter sequences and to what extent the binding reaction is influenced by the N-terminal extension of HSF-1 are investigated by comparing binding behavior of natural promotors of HSF-1 regulated genes. The factors that impact the DNA-protein interaction are identified. This sheds light on the very early steps in the HSF-1 regulation and contributes to understand the recognition and selectivity of HSF-1 targets.

Under normal conditions the main task of HSP-90 is to help its clients to fulfill their function by assisting their folding and by preventing unexpected misfolding, a process which is mainly regulated and supported by co-chaperons of HSP-90. One major type of HSP-90's clients - the kinase - is analyzed in terms of interaction and complex formation with HSP-90. In this thesis, using sB-Raf as a kinase model, it has been observed that with the help of its cochaperon CDC-37 and FKB-6, HSP-90 gains better affinity to the kinase domain, and thus improves the stability in the kinase processing procedure. When the kinase is mutated, HSP-90 still is able to form a complex with a slightly decreased affinity as compromise to the changes induced by mutants in kinase structure and activity. The presented results offer information on the formation and the

structural arrangement of Raf-chaperone complexes and provide insight into the association of kinase client with molecular chaperones in general.

Zusammenfassung

Durch die rapide Entwicklung neuer Technologien wurden biochemische Forschungen, welche sich auf zelluläre Signalgebung und Regulation konzentrieren, ausgehend von der Verfolgung einzelner Signalwege bis hin zur Erzeugung eines umfassenden Überblicks aller beteiligter Parameter sowie deren Untersuchung, erweitert. Beginnend mit der Analyse von Expressionsdaten der Hefe, werden in dieser Arbeit genomweite Veränderungen beobachtet und die entsprechenden Gene anhand ihrer koregulierten Expressionsniveaus gruppiert, welche anschließend durch die Anwendung eines wertefreien Klassifizierungsalgorithmus sogenannten Cliquen zugewiesen werden. Hierfür wurde das gesamte Genom involviert, ohne einen einzigen Nicht-Top-Treffern zu verlieren. Die Visualisierung dieser erhaltenen Cluster vereinfacht somit das Verständnis von signifikanten Änderungen innerhalb der genomweiten Regulierung.

HSP-90 tritt hierbei als eines der Schlüsselmoleküle für die Regulierung und Reifung von Kinasen und Transkriptionsfaktoren auf. Die durch die Herunterregulation von HSP-90 induzierten molekularen Veränderungen in *C. elegans* werden mit Hilfe von Mikroarrays analysiert, welche zeigen, dass die angeborene Immunantwort sowie Gene die mit dem Entwicklungsstadium verbunden sind, hauptsächlich beeinflusst werden. Obwohl sie von einem anderen Organismus stammen, stimmen die reagierenden Faktoren, wie durch die Expressionsdatenanalyse vorhergesagt, auf der Genomebene überein.

Unter Stressbedingungen ist das Expressionsniveau von HSP-90 erhöht, um die Zellen bei der Bekämpfung von Schäden zu unterstützen, die durch entfaltete oder fehlgefaltete Proteine hervorgerufen werden. Die regulierende Signaltransduktion wird durch den Transkriptionsfaktor HSF-1 ausgelöst. Als Initiationsschritt der HSP-90-Regulation wird HSF-1 hinsichtlich der DNA-Erkennung und -Bindung untersucht. Inwieweit die Sequenz der HSE die Belegung der Hitzeschockpromotorsequenzen beeinflusst und die Bindungsreaktion durch die N-terminale Verlängerung von HSF-1 beeinflusst wird, wird durch den Vergleich des Bindungsverhaltens von nativen Promotoren der durch HSF-1 regulierten Gene untersucht. Außerdem werden die Faktoren, welche die DNA-Protein-Interaktion beeinflussen, untersucht. Dies wirft Licht auf die sehr frühen Schritte der HSF-1-Regulierung und trägt zum Verständnis der Erkennung und Selektivität von HSF-1-Zielen bei.

Unter Normalbedingungen besteht die Hauptaufgabe von HSP-90 darin, seinen Substratproteinen bei der Erfüllung ihrer Funktion zu helfen, indem es deren korrekte Faltung unterstützt die durch Ko-Chaperone von HSP-90 reguliert und unterstützt wird. Kinasen, als ein Haupttyp des HSP-90-Substratproteins, werden hinsichtlich ihrer Interaktion und Komplexbildung mit HSP-90 analysiert. Erstmals konnte beobachtet werden, dass HSP-90 mit Hilfe der Ko-Chaperone CDC-37 und FKB-6 eine stärkere Affinität zur Kinase-Domäne aufweist

und somit die Stabilität bei der Regulierung des Kinaseprozesses verbessert. Auch bei Mutation der Kinase, ist HSP-90 in der Lage einen Komplex mit einer leicht abnehmenden Affinität zu bilden, auch wenn sich die Kinase-Struktur/-Aktivität ändert. Die präsentierten Ergebnisse liefern Informationen über die strukturelle Anordnung der Raf-Chaperon-Komplexe und geben allgemeine Einblicke in die Assoziation des Kinase-Substratproteins mit molekularen Chaperone.

1 Introduction

1.1 HSP-90 and its chaperon system*

*Part of this chapter has been published in Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2018

HSP-90 is a highly conserved and abundant chaperone, which participates in essential cellular activities by supporting and regulating the folding, modification and maturation process of its clients. The client processing is achieved via extensive conformational changes within the dimeric chaperone. This requires its ATP hydrolysis activity that is controlled by auto-inhibitory mechanisms and several structurally diverse cofactors. When HSP-90 is handling different types of clients, it relies dramatically on client-specific cofactors, which can adapt HSP-90's activities to meet the client's requirement at different conditions as well as in different cell types. Additionally, post-translational modifications can influence almost every aspect of HSP-90's interactions and activities. In the following chapter, the basic function, the correlated regulatory principle, impacting factors of HSP-90 and the mechanisms for regulating the HSP-90 machinery are discussed.

1.1.1 HSP-90 and its hydrolysis cycle

In all eukaryotes HSP-90 is an essential protein with observable increase in the expression under thermal stress. In proteomic analysis HSP-90 is present among the 50 highest expressed proteins [1]. Many eukaryotes contain multiple copies of cytosolic HSP-90 and additional homologs, which encode mitochondrial or ER-localized variants. These are called Trap1 (Tumor Necrosis Factor Receptor-Associated protein-1) and Grp94, respectively. Obviously HSP-90's functions diversified from a dispensable bacterial heat-shock protein and evolved into a mandatory chaperone for many essential clients. Concomitantly, the system has expanded to include about 20 cofactors that regulate the functions of HSP-90. These proteins bind to the HSP-90 dimer and prime HSP-90 for client-specific activities [2, 3].

All HSP-90-homologs are ATP-binding proteins with a C-terminal dimerization site and about 700 amino acids per subunit. The dimeric chaperone has two ATP binding sites in its N-terminal domains (NTD). Nucleotide binding and hydrolysis regulate the conformational cycle of HSP-90. Evidence suggests that HSP-90's conformational changes induced by ATP hydrolysis enable it to modify client conformations. This apparently supports important structural rearrangements in its client proteins contributing to client maturation and preventing their degradation via the

proteasome. HSP-90's ATPase cycle has been assumed to be symmetric with both ATP molecules being hydrolyzed simultaneously in a conformation, where two N-terminal ATP-binding domains are interacting with each other. This evidence was mostly based on the symmetric X-ray structures of nucleotide-complexed forms of yeast Hsp82 and Grp94 [4-7]. In particular the recent structural and biochemical analysis of the mitochondrial HSP-90-protein Trap1 has revealed detailed steps of these conformational changes in a homolog, which unlike the cytosolic HSP-90, is not regulated by the cofactor interaction. Trap1 instead performs its ATP hydrolysis in the absence of regulatory cofactors and surprisingly with clear signs of an asymmetric reaction (Figure 1.1).

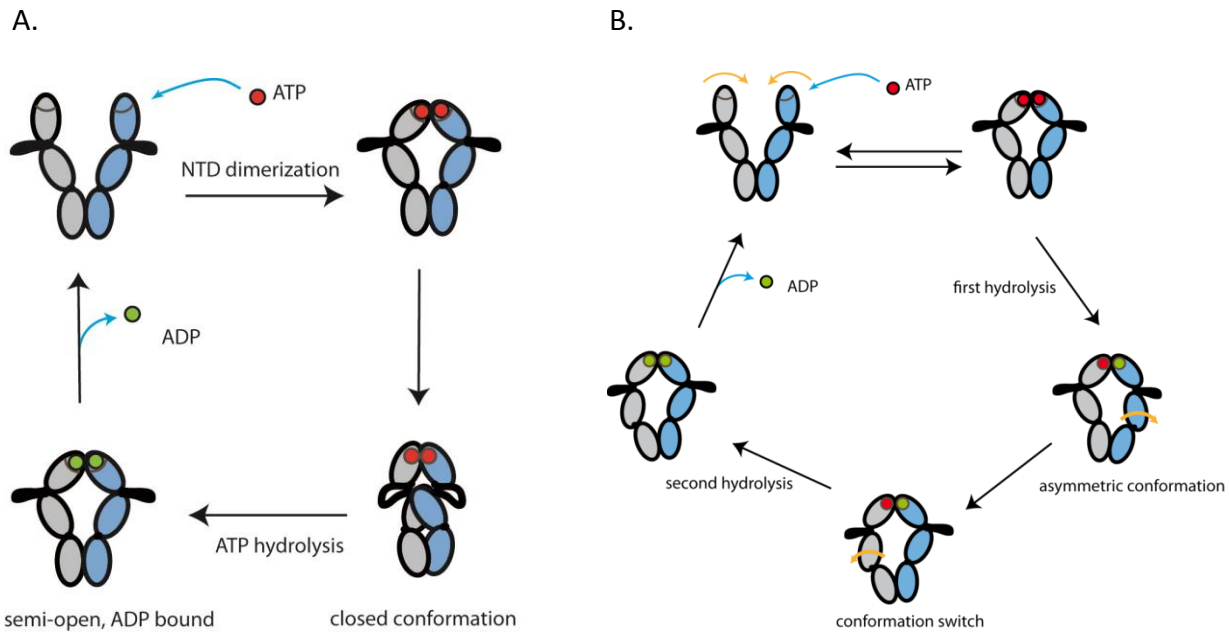


Figure 1.1 Schematic comparisons of (A) yeast HSP-90 hydrolysis cycle vs (B) asymmetric Trap1 cycle. The blue arrows indicate movements of nucleotides and the orange arrows present conformational changes of the protein.

1.1.2 Co-chaperons of HSP-90

The cytosolic HSP-90 is regulated by about 20 structurally diverse cofactors. Most of these regulatory proteins bind to the C-terminal MEEVD motif of HSP-90 via their TPR-domains. While the cytosolic HSP-90 protein itself interacts with a large number of different client proteins, its cofactors are much more specific. Some of these proteins simultaneously interact with HSP-90

and specific client proteins, bringing them together in multimeric protein complexes. For these the cofactors regulating HSP-90's functionality are well described and concepts regarding the client turnover mechanisms have been established.

The co-factor of HSP-90 can be mainly divided into two groups, namely the TPR containing cochaperon and the TPR-free cochaperons [8]. TPR domain is known as a binding site of the HSP-90 MEEVD motif and thus acts as a linker of the cochaperon to HSP-90. The TPR co-chaperons include FKB proteins (FKBP51, FKBP52 and AIP), which contributes to the client chaperon maturation; HOP which assists the binding of Hsp70 to HSP-90; CHIP which tags client proteins of HSP-90 for degradation; PP5 which acts as a phosphatase; Sgt1/Unc45/Tah1 which are related to the formation or assembly of HSP-90 complexes and Tom70/Toc64 which are involved in the protein import. Other cochaperons which don't contain the TPR domain include CDC-37, which is important for client transport to HSP-90; Aha-1 which stimulates the HSP-90 activity and induces its conformational change and p23 which is involved in the client maturation of HSP-90.

As one of the best known chaperone cofactors, the function of CDC-37 has been extensively studied in the last decades. Its N-terminal domain is thought to be a client and HSP-90 binding side, the M-domain is recognized by HSP-90 and the function of the C-terminus of CDC-37 remains unclear to date. CDC-37 has been reported to directly bind the kinase domain of B-Raf and form a heterocomplex with HSP-90 [9], which is required for B-Raf function [10]. The CDC-37 client binding site between residues 181 and 200 interacts with the glycine-rich loop of the kinase client [11], where it partially unfolds the client protein at the segment and forms a stable complex with the client via multidomain interfaces [12]. The binding of client proteins is regulated by the N-terminus of CDC-37 [13, 14]. The stabilization of the client kinase by CDC-37 does not depend on the presence of HSP-90, but can be performed by CDC-37 alone [15]. Interaction of CDC-37 with the N-terminus of HSP-90 generates an HSP-90 dimeric conformation suitable for client recognition [16]. After the activation of B-Raf the protein complex is remodeled, visible based on the remodeling of the quaternary structure of the complex [17].

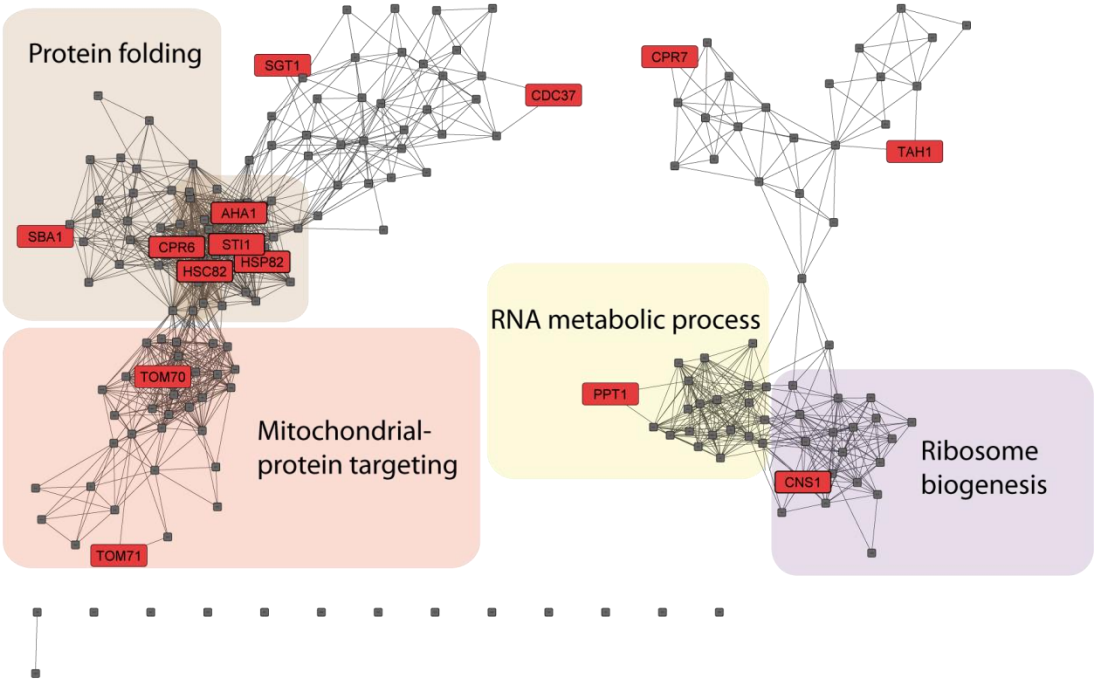
Many cofactors of cytosolic HSP-90 have been identified and for most of them the client-specificity is far from clear. Plotting the whole transcriptional regulation network of yeast HSP-90 and its cofactors (Figure 1.2 A) shows how different the expression conditions of these cofactors are, implying that some cofactors are only present under certain conditions. The two HSP-90-isoforms HSP82 and HSC82 and a core set of their cofactors, including AHA1, CPR6 and STI1 are expressed together with other heat-shock proteins and other components of the heat-shock response [18]. Under heat-stress the HSP-90 concentrations are increased, but not nearly as much as that of heat-inducible Hsp70 proteins or small heat-shock proteins. Some HSP-90-cofactors are regulated entirely different from HSP82 to HSC82. The HSP-90-associated phosphatase PPT1 is expressed together with RNA modifying proteins and CNS1 is expressed

together with components of the ribosome assembly pathway. TOM70, together with its homolog TOM71, is part of a coexpression clique that is enriched in proteins of the mitochondrial protein trafficking pathway. These specific cofactors may involve HSP-90 in very specific tasks, as shown for the mitochondrial import regulator and HSP-90-cofactor Tom70. In these cases, certain expression conditions may alter the presence of one specific cofactor and thereby shift the focus of the available HSP-90 molecules towards the newly expressed cofactor's clients.

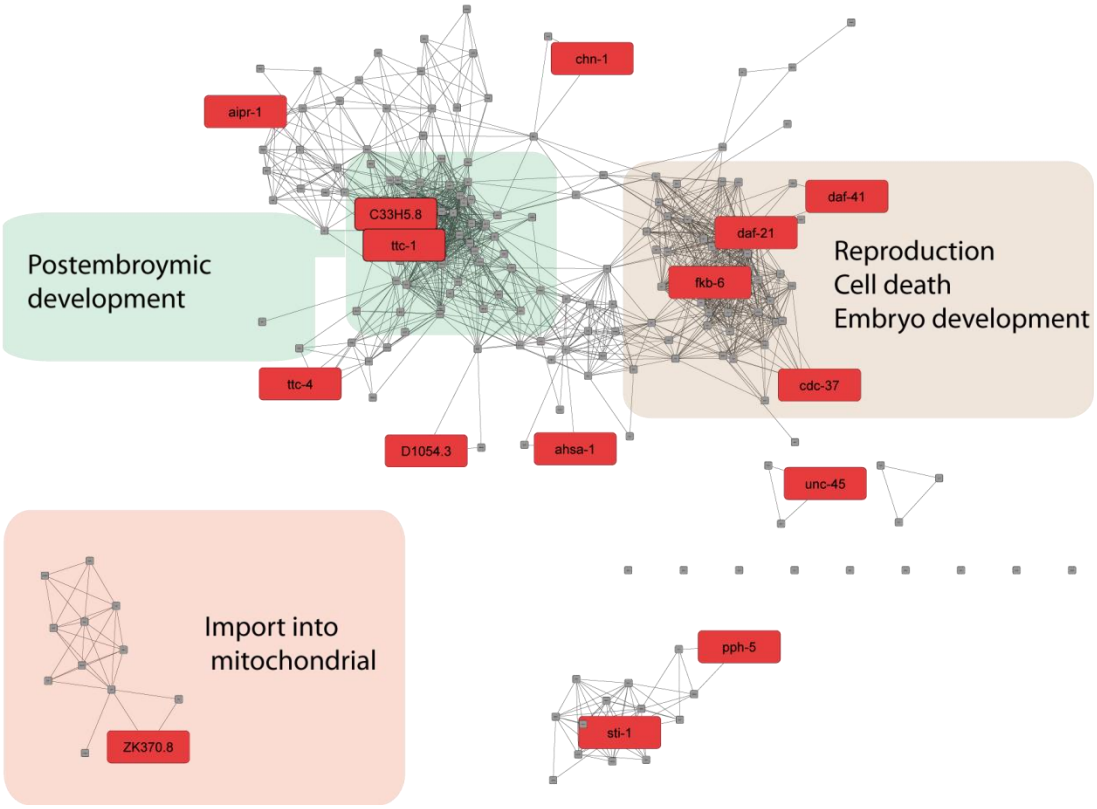
In all multicellular systems homolog cofactors of the HSP-90 system are encoded and binding studies indicate that the interaction properties of these cofactors are mostly conserved. In all three metazoan systems (Figure 1.2 B, C and D), like in yeast, the HSP-90 proteins are highly expressed and coregulated with other heat-shock proteins. As in yeast, this pattern is shared by several cofactors, such as Ahsa1, Stip1 and several others [18]. The correlation of these cofactors with the core HSP-90 system seems to be preserved during evolution. However, in particular in the mammalian systems, it is obvious that in some cell types HSP-90 is confronted with tissue-specific cofactors. E.g. UNC45B expresses with an entirely specific gene clique that is present in skeletal muscle tissue, while the expression of AIPL1 correlates strongly with photoreceptors and other proteins from the visual sensory system. This is independently evident from mouse coexpression databases and from human ones. Interesting in this respect is the regulation of apparent gene duplications in the mammalian system, where one homolog apparently is supporting a highly specialized function in a specialized tissue, while the other homolog is retaining its coexpression relationship with the core HSP-90 machinery. This is evident for AIP and AIPL1, or CDC-37 and CDC-37L1, or UNC45A and UNC45B. Thus in the higher eukaryotic systems, where tissue specificity is most developed, a diversification of the HSP-90 system becomes apparent and several cofactors show a remarkable tissue specificity that correlates with known details about their function.

Figure 1.2 Expression of HSP-90 and its cofactors in different model organisms (see next page)**.** Expression networks of the HSP-90 systems for yeast (A), *C.elegans* (B), mouse (C) and human (D) were generated from the connections within the coXpress data base, modSEEK and SPELL databases [18-20]. The Top 20 coexpressed genes were used for HSP-90 and each cofactor to generate the networks [21]. The functional labelling of the clusters was performed after GO enrichment analysis of the colored genes [22, 23]. GO enrichment analysis is only shown in those cases, where the significance levels were high enough. Only the top three GO terms were included in the figures, but very general GO terms were omitted at the expense of more informative and specific ones. The current names of the cofactors were used, even though some of them differ strongly between the organisms. As such, ttc-4 in *C. elegans* represents CNS1 in yeast, CPR6 and CPR7 represent Cyp40/PPID, ZK370.8 represents TOM70, C33H5.8 represents TAH1 in yeast and RPAP3 in mammals, D1054.3 represents SGT1 or SUGT1 and chn-1 represents STUB1.

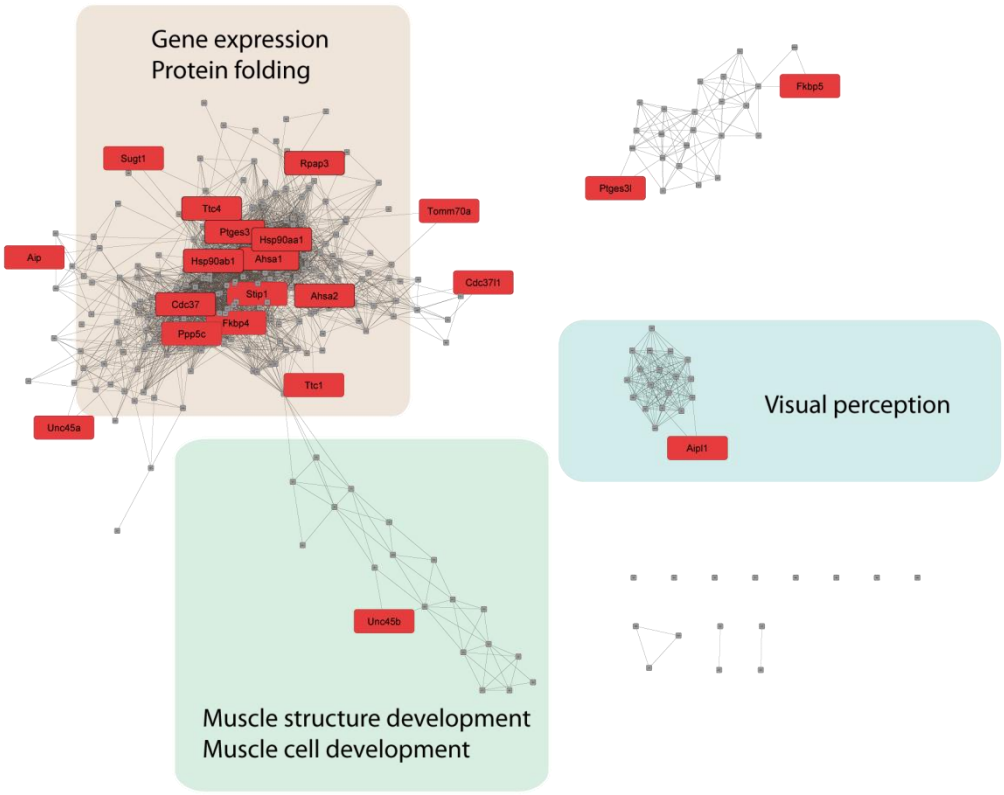
A.


Protein folding
SGT1
CDC37
AHA1
SBA1
CPR6
STI1
HSP82
HSC82
TOM70
TOM71
Mitochondrial-
protein targeting
CPR7
TAH1
RNA metabolic process
PPT1
CNS1
Ribosome
biogenesis

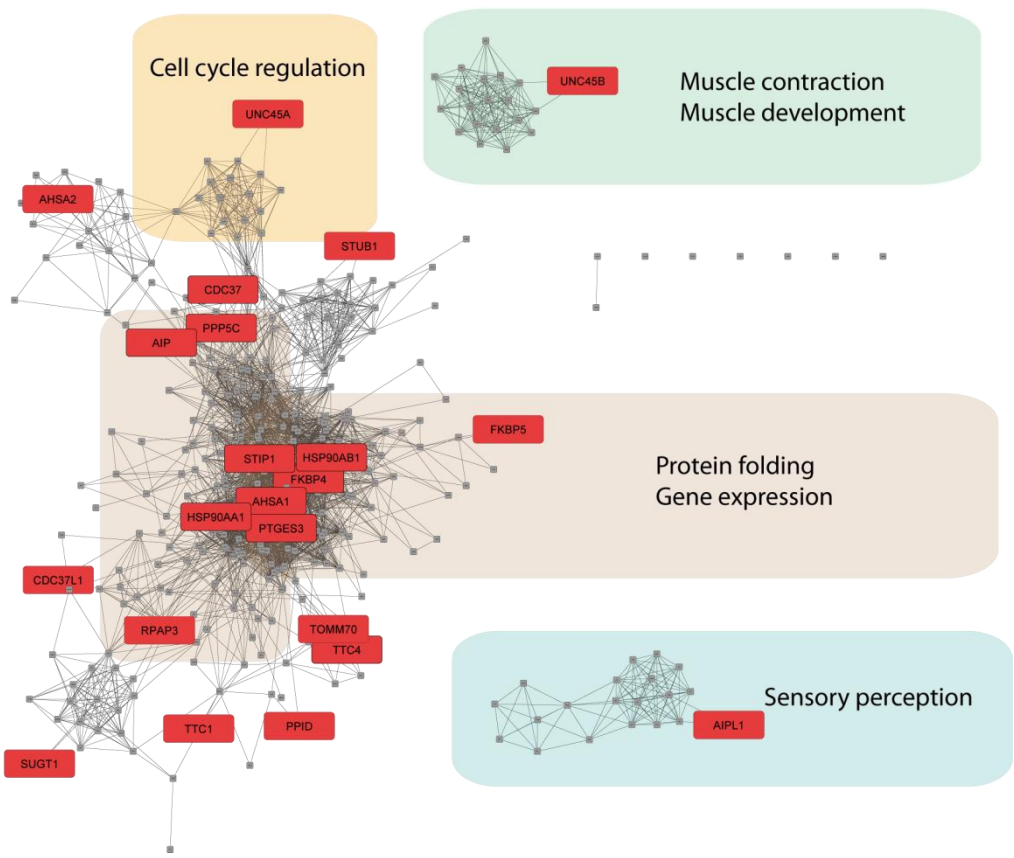
B.


chn-1
aipr-1
C33H5.8
ttc-1
Postembroymic
development
daf-41
daf-21
fkb-6
Reproduction
Cell death
Embryo development
ttc-4
cdc-37
D1054.3
ahsa-1
unc-45
Import into
mitochondrial
ZK370.8
pph-5
sti-1

C.


Gene expression
Protein folding
Sugt1
Rpap3
Ttc4
Tomm70a
Ptges3
Hsp90aa1
Aip
Hsp90ab1
Ahsa1
Cdc37l1
Stip1
Cdc37
Ahsa2
Fkbp4
Ppp5c
Ttc1
Unc45a
Fkbp5
Ptges3l
Visual perception
Aipl1
Unc45b
Muscle structure development
Muscle cell development

D.


Cell cycle regulation
UNC45A
AHSA2
STUB1
CDC37
PPP5C
AIP
FKBP5
STIP1
HSP90AB1
FKBP4
AHSA1
HSP90AA1
PTGES3
CDC37L1
RPAP3
TOMM70
TTC4
TTC1
PPID
SUGT1
UNC45B
Muscle contraction
Muscle development
Protein folding
Gene expression
Sensory perception
AIPL1

Several new cochaperones for HSP-90 have been found recently, such as the tumor suppressor proteins Tsc1 and Tsc2, which are related to the tuberous sclerosis complex (TSC) syndrome [24]. Furthermore the tumor suppressor FLCN has been found to be a client of HSP-90 together with its cochaperones FNIP1 and FNIP2 (folliculin-interacting proteins 1 and 2) [25]. Directly interacting with the M-domain of HSP-90, these cochaperones compete with the activating cochaperon Aha1 and thus influence the balance between activation and inhibition of HSP-90's ATPase activity, while utilizing the HSP-90 machinery for their specific clients.

1.1.3 Post translational modification of the HSP-90 complex

Besides nucleotides and cochaperones, another layer of HSP-90 regulation is associated with post-translational modifications (PTM). Varies of modifications of HSP-90 have been reported mainly including the phosphorylation, acetylation and methylation at HSP-90 side chains. Many PTMs directly modify HSP-90 itself; others modify its cochaperones and clients. These modifications generally improve the HSP-90 functionality towards the individual need of certain clients.

1.1.3.1 Phosphorylation of HSP-90

Phosphorylation is the most frequently detected post translational modification of HSP-90. A large number of phosphorylation sites of HSP-90α and HSP-90β are compiled in the PhosphoSitePlus database [26] and likewise for the other proteins of the system. Well-documented phosphorylation sites, which have been identified more than five times, are distributed over the surface of human HSP-90α. Sequence comparisons show that more than half of those sites are conserved between the two mammalian HSP-90 homologs, implying that the same regulatory mechanisms apply in both homologs. Also recent biochemical studies show that phosphorylation sites regulate most of HSP-90's activities.

Several phosphorylation sites have been identified that either regulate the mechanism of HSP-90 or affect the interaction between HSP-90 and its cochaperones and clients. Most biochemical data on phosphorylation sites are available for HSP-90α/HSP-90AA1. Some sites affect a wide range of activities. Phosphorylation of Thr90, as performed by protein kinase A, affects the binding affinity of ATP and modifies the association of several HSP-90 cochaperones, such as Aha1, p23, Pp5, Chip/STUB1, CDC-37, Hop/STIP1 and Hsp70 [27]. Other sites seem to regulate specific cofactors: Tyr197-phosphorylation weakens the interaction with CDC-37, Tyr313-phosphorylation promotes the binding of Aha1, driving the kinase-chaperoning cycle forwards [28]. Phosphorylation of Thr725 or Ser726, which can be performed by CK1, CK2 or GSK3β,

regulate the differential binding status for Hop/STIP1 and Chip/STUB1 at the TPR-domain acceptor site [29]. In general, phosphorylation at residues close to the MEEVD motif of HSP-90 can regulate the affinity of TPR-domain containing cochaperones [30]. Also client interaction can be regulated by phosphorylation. As such, the phosphorylation at Thr115, Thr425 or Thr603 regulates client binding by either 'switching on' the binding site or reducing its affinity for the kinase PKCγ, which itself can attach the phosphates [31]. While several sites are shared between HSP-90α and HSP-90β, some characterized sites also are isoform-specific. One specific phosphorylation site, Ser365 in human HSP-90β, has also been shown to affect the interactions with HSP-90's clients and CDC-37 [32].

Very detailed studies have also been performed in the yeast model system, investigating the consequences of Hsp82's phosphorylation status. Phosphorylation of Thr22 in yeast HSP-90 by casein kinase 2 (CK2), a phosphorylation site conserved also in the human system, was found to reduce the interaction with Aha1 and CDC-37 [33]. In an extensive study Soroka et al. have generated phosphomimetic variants of many more phosphorylation sites and characterized several aspects of this modified Hsp82s. Phosphorylation at Ser379 in the M-domain of HSP-90 reduces the ATPase activity and chaperone function and modifies the formation of the contacts between NTD and M-domain during the ATPase cycle [34]. Ser602 and Ser604 in the C-terminal domain also slow down the ATPase cycle, but they do so by regulating the inter-domain interactions. While apparently many different movements and interactions of HSP-90 are tunable by phosphorylation, it will be interesting to learn, to what extent this part of the regulation is conserved in HSP-90-homologs from different species.

1.1.3.2 Phosphorylation of cofactors

Also cofactors of HSP-90 are subject to phosphorylation reactions. As such, the PhosphoSitePlus database contains 6 well-documented phosphorylation sites for human CDC-37 and 7 well-documented phosphorylation sites for human Hop/STIP1. Some of those have been characterized regarding their influences.

As such, phosphorylation on Tyr4 and Tyr298 was found to disrupt the association of CDC-37 with client kinases [28]. Beside this, in particular one phosphorylation site is highly conserved in CDC-37 proteins from different species. Phosphorylation at Ser13 alters the structure of the highly conserved N-terminal region and affects processing of kinase clients [35]. This site can be phosphorylated by CK2 also in other metazoans, like *C.elegans*, or even in yeast. Interestingly, phosphorylation at this site can be removed by the HSP-90-associated phosphatase Pp5 [36]. There is growing evidence that this site controls the kinase-chaperoning cycle. Ser13 is phosphorylated prior to client kinase binding and Ser13-phosphorylated CDC-37-kinase complexes are binding to HSP-90. The phosphate on Ser13 is removed by Pp5 prior to Aha1 or

ATP-binding to HSP-90, which sets the stage for the second part of the hydrolysis cycle and the release of the kinase. This mechanism couples the phosphorylation of CDC-37 to its interaction with kinase and HSP-90 [36]. This mechanism may complement the otherwise observed dissociation of the CDC-37-kinase complex after nucleotide-binding to the kinase [37]. Thus it is becoming clear that also cofactors are subject to regulation by phosphorylation and contribute in this way to the regulation of the system. While evidence for CDC-37 phosphorylation is furthest ahead, phosphorylation sites on other cofactors are also getting identified and await their biochemical characterization.

1.1.3.3 Other modifications of the HSP-90 system

Beyond phosphorylation, HSP-90's activity is directly regulated by reversible acetylation mediated by histone acetyltransferases (HAT) and histone deacetylases (HDAC6) [38, 39]. Detailed analyses, using HDAC-inhibitors, revealed that under basal condition, acetylated HSP-90s form complexes with their co-chaperones Hop, Hsp70 and Hsp40. Deacetylation by HDAC6 at Lys294 [40] generates HSP-90 molecules that are ready to form complexes with the co-chaperone p23 and GR. Deacetylation thus supports GR-activation by stabilizing the matured GR-HSP-90 complex. This may be achieved by increasing HSP-90's ATP-binding affinity and thus shifting the HSP-90 conformation towards formation of the HSP-90-GR heterocomplex [41, 42]. Beyond the client GR, HSP-90 acetylation also regulates other pathways such as estrogen receptor signaling [43].

HSP-90 complex formation was also found to be regulated by methylation and this modification may be relevant for the maintenance and function of skeletal muscle. HSP-90 in this case forms a complex with titin. Deficiency in the methyl transferase Smyd2 results in the loss of HSP-90 methylation, impaired titin stability, and altered muscle function [44]. The interaction of HSP-90α with the methyltransferase Smyd2 and Smyd3 leads to the methylation of Lys209 in the nucleotide-binding domain and Lys615 in the C-terminal dimerization domain of HSP-90. The methylation is influenced by the co-chaperones of HSP-90 and by demethylation through the lysine amine oxidase Lsd1 [45]. Methylation of Lys531 and Lys574 on HSP-90β results in higher dimerization affinity and substitutions of Lys531 and Lys574 with alanine reduced HSP-90β's interaction with Hop or CDC-37. Methylated HSP-90β accelerated the proliferation of cancer cells *in vivo*, indicating that these modifications have relevance for several diseases associated with HSP-90 [46].

Mediated by nitric oxide (NO), S-nitrosylation is a reversible post-translational modification of HSP-90, which occurs in the C-terminal domain at Cys597 [47]. It has been observed that S-nitrosylation at this location can lead to inhibition of the ATPase activity of HSP-90 [48]. This site is located in an important switch region, which controls conformational changes in the C-

terminal domain and the communication with the N-terminal domain during the ATPase cycle [49].

SUMOylation of HSP-90 can take place at the N-terminal domain at Lys178 of yeast HSP-90 or Lys191 of human HSP-90α. It stimulates the ATPase activity and is supportive for the interaction with the cochaperon Aha1. The SUMOylation in the N-terminal domain affects the heat shock response induced by elevated temperatures and the binding of HSP-90 inhibitors [50]. Finally, SUMOylation of the HSP-90 cochaperons Fkbp51 regulates its binding to HSP-90-GR complexes, which has an impact on the GR signaling pathway [51].

1.2 The HSR regulating transcriptional factor: HSF-1

The heat shock transcription factor (HSF-1) is an essential transcriptional activator of heat-shock genes in the heat-shock response and the immune response [52, 53]. HSF proteins are further involved in many developmental processes like the embryonic placenta development [54], the female meiotic division [55], and the general transcription [56] by upregulating specific target genes. HSF proteins are also negative regulators of RNA polymerase II promoters and have an important role in the ubiquitin proteolytic pathway, the cellular proliferation [57, 58] and the regulation of multicellular organism growth [59, 60]. It activates genes of the classical HSR like Hsp70, small HSPs and Hsp40s [52, 61].

HSF-1 in nematodes is a protein of 671 amino acids, which mainly contains three domains. From the N-term to the C-term are the DNA-binding domain (DBD), the oligomerization domain and the carboxyl-terminal regulatory domain. Nematode HSF-1 further contains 82 amino acids extension at the N-terminus of unknown function. Under normal conditions HSF proteins are monomeric when they interact with HSP-90 and Hsc70 to form multi-chaperone complexes, which prevent HSF-1 from forming trimers and being activated. Under heat-stress or other inducing conditions, Hsf proteins are released from the protecting chaperon and trimerized. In most cases HSF-1 binds as a trimeric protein to HSE-containing DNA sequences. The phosphorylation of HSF proteins triggers the translocation to the nucleus and initiates the transcription [62-64]. Despite these regulatory cycles, the interaction of HSF proteins with consensus dsDNA is observable also for non-activated forms. In this respect, it is mostly unclear, how HSF proteins distinguish its various HSE-containing target genes.

It has been observed that young larvae at L2 stage show a reduced expression of Hsf-1 compared to later life stages, when Hsf-1 is induced much stronger. The aging adult then again reduces the inducibility of the response. This implies a complex behavior of the HSF-1 regulation during aging and may allow the nematode to ensure that its reproductive phase is best

protected from stressful events. Interesting in connection with the heat-shock response is the observation that several thousand HSEs are present in promoter regions of the genome, even though the canonical heat-shock protein only belongs to a small cluster of inducible genes, which could be induced the HSR. Despite this, it is still unclear how HSF-1 distinguishes numerous potential target genes. The smallest DNA motif that can be recognized and bound to Hsf-1 is called the heat shock element (HSE, sequence nGAAxxTTCn). Given that more than 4000 consensus HSEs can be found in nematode promoter regions, apparently many of those genes are not induced during the heat-shock response and may genes contribute to the HSF-1's functions as an essential transcription factor under non-stressed conditions. How does HSF-1 selectively bind to the HSEs which is able to induce heat shock proteins becomes an interesting question.

1.3 Kinase client of HSP-90 and the role of HSP-90 in the regulation of their clients

Protein kinases are important contributors to the cellular signal transduction. During this procedure they bind ATP and phosphorylate their target proteins. Protein kinases are regulated by multiple means, including internal domain rearrangements and several subtle loop movements. Nearly a third of the known kinase in human is identified to be potential client of HSP-90. However the factor that determines whether a kinase is a client of HSP-90 has not been found to date. Many protein kinases possess complicated life cycles that include HSP-90-dependent steps during their maturation and regulated degradation. The mechanism of the HSP-90 involvement is not well understood, even though the importance of the HSP-90-cofactor CDC-37 for kinase maturation is well established to date. Kinase complexes of CDC-37•HSP-90 have been confirmed for several kinases, including B-sB-Raf, v-Src, Akt.

Being observed side by side in living cells, the activation of kinase involves molecular chaperons. It has been proposed that these chaperones, in particular HSP-90 and its cofactor CDC-37 influence the binding of small molecules and nucleotides to the active center of sB-Raf kinase [65]. Observations reveal a depletion of nucleotides on the formation of protein complexes. In the meantime the nucleotide stabilizes the sB-Raf kinase strongly and prevents its interaction with other chaperones, in particular with CDC-37. In the HSP-90 chaperon system, kinase clients are brought to HSP-90 by the co-chaperon CDC-37.

The regulation pathway of sB-Raf has been postulated to include the N-terminal regulatory domain that can act as inhibitors of the kinase activity and loop regions surrounding the active center of the protein [66]. The kinase domain, which is conserved among all cellular protein kinases, appears to have also conserved regulatory mechanisms. In many cases loops controlling the active center are phosphorylated, which elevates the activity of the kinase. In its active form

the sB-Raf-kinase then forms a protein complex with 14-3-3 proteins and the MEK-kinase [67, 68]. Its activity is further controlled by the Ras-protein, which binds to the most N-terminal regulatory domain of sB-Raf [69].

B-Raf belongs to one of the most commonly mutated proteins in the human genome and its mutations are found in many cancers, including malignant melanomas. Recent studies which focused on the signal pathway that involve the RAF kinase have classified oncogenic mutants of RAF into three groups based on its activating mechanisms: The first type is the highly independent mutant V600, which is the most frequently observed variant of B-Raf that covers more than 90% of the known case. It shows an increased kinase activity towards its clients from the MAPK cascade [70], which is active as a monomer and doesn't require the RAS upstream stimulation; the second group includes relatively independent mutants at amino acids G464, G469 and K601 etc. which require dimerization for the activation. However the second type doesn't rely on the RAS signaling; and the third group including G464, N581 and D594 which are relative strictly regulated both by the dimerization and the RAS upstream signaling [71]. The most frequently identified oncogenic mutations locate either in the activation segment or in the glycine-rich loop — motifs which are thought to alter the inactive state to favor the active conformation [72]. The kinase B-raf is a well-known client of HSP-90 and it has been suggested that the HSP-90 chaperon system can stabilize meta-stable activated states of its kinase client [73].

The HSP-90-cofactor CDC-37 binds to kinases and targets them to the HSP-90-chaperone machinery [74, 75]. This was initially shown for the sB-Raf and v-Src. Many protein kinases have been identified as binding partners for CDC-37 or HSP-90•CDC-37 [76]. It has been shown that the inhibition of HSP-90 with radicicol or geldanamycin, two naturally occurring competitive inhibitors of ATP-binding to HSP-90, leads to degradation of HSP-90-dependent protein kinases [77]. This observation is applied therapeutically to deplete oncogenic HSP-90-dependent protein kinases, such as mutated versions of sB-Raf or Akt, to reduce oncogenic growth of tumors, where these kinases are important for malignant growth [78-80]. Other cochaperons of HSP-90 e.g. Aha-1 [81] and Bag-1 [82] are also reported to have influences in the activity and complex formation of sB-Raf kinase.

1.4 Cellular expression response of yeast

The single cell organism Saccharomyces cerevisiae is one of the best studied eukaryote model organisms [83] with a well annotated genome of approximately 5800 genes [84], of which 23% are homologous to humans [85]. It has more than 100 genes encoding regulatory proteins influencing directly the expression of other genes [86]. Combined with a short life cycle, S.

cerevisiae therefore is widely used for biological network studies based on microarray data. Regulated networks can be based on protein-protein interactions (PPIs), signaling or metabolic connections or co-expression relationships. These networks usually are visualized as nodes for genes, connected by edges, representing interactions between these nodes [87]. In most cases this approach leads to the grouping of connected nodes and the formation of co-regulated cliques or clusters. Often co- regulated genes have been found to share also functional properties or even form protein complexes directly [88].

Several methods and organisms have been used to construct genomic or proteomic networks in recent years [85, 87, 89-92]. In some cases training sets of selected micro-array experiments are used to derive connections between nodes and to make a network from them. In other cases databases, like STRING [93], COXPRESdb [18], modSEEK [20] or SPELL [19], are used, which contain information on co-regulation properties, but also may contain other types of interactions, like PPIs and co-naming in articles or abstracts [93]. After network construction, the identification of clusters, cliques or modules is the next important step. It is widely recognized that in most cases a change in the transcriptional program is not only affecting one or two clusters, but due to the connected nature of cellular responses it affects several transcriptional modules simultaneously and to variable extent. These clusters need to be well separated, especially if further analyses on isolated gene groups are planned, as in general enrichment analyses profits from high quality gene sets.

The genome-wide co-expression database is used to generate networks with fairly high connection density from hits of microarray experiments. This approach was applied to the identification of differentially induced gene clusters after polyglutamine expression [94] or differentially expressed gene clusters after HSP-90-depletion in C. elegans [1]. These networks were constructed from the Top 100 or Top 200 hits per experiment and in all cases several co-regulation clusters could be separated from each other [1, 21, 94]. Nevertheless, in all cases some genes could not be connected within these networks even though they showed strongly altered expression behavior. Also, when trying to separate clusters within these networks, for some gene groups no significant gene ontology (GO)-term or transcription factor (TF) assignments could be obtained in enrichment analyses [21]. This could be caused by the limitation to 100 or 200 Top-genes and the resulting exclusion of many important genes from these networks.

2 Aim of the work

The aim of the thesis was to study how Hsp90 is involved in organismic processes and how it fulfills its functions on the molecular level. To target the first part, a simplified system to investigate genomic and proteomic responses should be established initially. Using yeast as a model, the co-regulation response is observed on the genomic level based on the analysis of a full genomic co-expression database. The resulting cliques would offer valuable assistance in analyzing microarray data in the context of the genome-wide response. On the other hand, in *C.elegans*, the consequences of HSP-90 down-regulation are investigated both on genomic and proteomic level to gain knowledge on how HSP-90 is involved and influences the entire development of the nematodes.

Meanwhile on the molecular level, the interaction and binding properties of the upstream HSP-90-regulating transcriptional factor HSF-1 and the interaction and regulation of HSP-90's kinase client B-Raf is investigated.

Aimed at understanding how HSF-1 differentiates between different nematode HSEs, a set of pre-defined coregulation cliques was used to identify, which genes are part of the canonical heat-shock response and which HSE-containing genes group in other coregulation cliques. The comparison between the entire N-terminal region and the isolated DNA binding domain reveals the role of the first 82 amino acids containing motif in *C.elegans* HSF-1. The differences in complex formation with promotor DNAs, which are regulated by HSF-1, are observed and the partial concentration of each complex species is calculated. This offers hints on how the transcriptional factor recognizes its target as well as how HSF-1 associates on the DNA.

As for the client of HSP-90, the complex of a stabilized kinase domain of B-Raf (sB-Raf) with CDC-37 and HSP-90 was reconstructed. The influence of nucleotide or cochaperon binding is tested. The formation process of the kinase-chaperon complex is observed in SV-AUC stepwise. The structure of the kinase complex is investigated by X-linking/mass spectrometry and by 3D reconstitution of negatively stained quaternary complexes, which provides a model with structural information. Combing all information obtained from these analyses, the process of the kinase complex formation as well as the mechanisms of this procedure could be explained. Moreover, the role of FKB-6 in this complex is discussed.

Furthermore, the interaction of the HSP-90 system with mutants of sB-Raf was observed. Using the solubilized kinase domain of B-Raf (sB-Raf) as a model, influences of various regulating factors such as nucleotides and molecular chaperones on different classes of B-Raf mutants, which locate in the glycine-rich loop and activation segment, are observed. This may offer a clue to understand the mechanistic differences of oncogenic mutants and the role of HSP-90 system in regulating its kinase client.

3 Material and methods

3.1 Materials used in the thesis

Name	Source
1,4-Dithiothreitol (DTT)	Roth (Karlsruhe, Germany)
5-(and-6)-Carboxyfluorescein succinimidyl ester (FAM)	Invitrogen (La Jolla USA)
ADP	Roche (Basel, Switzerland)
Agar	Serva (Heidelberg, Germany)
Agarose, ultra-pure	Roth (Karlsruhe, Germany)
Ammonium peroxide sulfate (APS)	Roche (Mannheim, Germany)
Ammonium sulfate	Merck (Darmstadt, Germany)
Ampicillin	Roth (Karlsruhe, Germany)
AMP-PNP	Roche (Basel, Switzerland)
ATP	Roche (Basel, Switzerland)
ATPγs	Roche (Basel, Switzerland)
Bacto Yeast Extract	Difco (Detroit, USA)
Bradford	Serva (Heidelberg, Germany)
Bromphenolblue S	Serva (Heidelberg, Germany)
Bis-Sulfo Succinimidyl Suberate (BS3)	Creative Molecules Inc.
Cholesterol	Sigma (St. Louis, USA)
Coomassie Protein Assay Reagent	Pierce (Rockford, USA)
Di-Sulfo Succinimidyl Glutarate (DSSG)	Creative Molecules Inc.
ECL-Westernblot Detection System	GE Healthcare (Munich, Germany)
Ethanol	Merck (Darmstadt, Germany)
Ethidium bromide	Sigma (St. Louis, USA)
Ethylene diamine tetra acidic acid (EDTA)	Merck (Darmstadt, Germany)
Glucose	Sigma (St. Louis, USA)
Glycine	Roth (Karlsruhe, Germany)

Imidazole	Sigma (St. Louis, USA)
Isopropanol	Roth (Karlsruhe, Germany)
Isopropyl-β-D-thiol galakto pyranosid (IPTG)	Roth (Karlsruhe, Germany)
Kanamycin	Roth (Karlsruhe, Germany)
lactate dehydrogenase	Roche (Basel, Switzerland)
LiO_2Ac	Roth (Karlsruhe, Germany)
Milk powder	Roth (Karlsruhe, Germany)
N-(2-Hydroxyethyl)-piperazin-N´2-Ethansulfonic acid (HEPES)	ICN (Costa Mesa, USA)
N,N,N´,N´-Tetraethyl ethylene diamine (TEMED)	Roth (Karlsruhe, Germany)
NADH	Roche (Basel, Switzerland)
Phosphoenolpyruvate (PEP)	Sigma (St. Louis, USA)
Poly ethyl imine (PEI), linear	Polysciences (Illinois, USA)
Protease Inhibitor Mix HP	Serva (Heidelberg, Germany)
pyruvate kinase	Roche (Basel, Switzerland)
Roti Mark pre-stained	Roth (Karlsruhe, Germany)
SDS-PAGE Standard Low Weight	BioRad (Hercules, USA)
Sodium dodecyl sulfate (SDS)	Roth (Karlsruhe, Germany)
Stain G	Serva (Heidelberg, Germany)
SYBR™ Green Nucleic Acid Gel Stain	Invitrogen (La Jolla USA)
SYPRO™ Orange Protein Gel Stain	Invitrogen (La Jolla USA)
Tetracycline	Roche (Basel, Switzerland)
Tris-(Hydroxymethyl)-amino methane (Tris)	ICN, Costa Mesa, USA
Yeast Nitrogen Base w/o phosphate w/o amino acids	MP Biomedicals (Santa Ana, USA)
β-Mercaptoethanol, pure	Merck (Darmstadt, Germany)
$CaCl_2$	Roth (Karlsruhe, Germany)
$MgSO_4$	Roth (Karlsruhe, Germany)
K_2PO_4	Roth (Karlsruhe, Germany)
KCl	Roth (Karlsruhe, Germany)

$MgCl_2$	Roth (Karlsruhe, Germany)
KH_2PO_4	Roth (Karlsruhe, Germany)
Na_2HPO_4	Roth (Karlsruhe, Germany)
NaCl	Roth (Karlsruhe, Germany)
NADH	Roche (Basel, Switzerland)
salmon sperm	Promega (Fitchburg, USA)
Trypsin	Promega (Fitchburg, USA)
NH4HCO3	Roth (Karlsruhe, Germany)
acetonitrile (ACN)	Sigma (St. Louis, USA)
Guanidine	Roth (Karlsruhe, Germany)
HCl	Roth (Karlsruhe, Germany)
3-Indoleacetic acid (IAA)	Sigma (St. Louis, USA)
Trifluoroacetic acid (TFA)	Roth (Karlsruhe, Germany)
α-cyano-4-hydroxyl cinnamic acid	Sigma (St. Louis, USA)
LC-MS/MS Peptide Reference Mix	Promega (Fitchburg, USA)
Coomassie Brilliant-Blue R-250	Serva (Heidelberg, Germany)
Glutaraldehyde, 25% in water	Serva (Heidelberg, Germany)
Formaldehyde, 37% p.A.	Roth (Karlsruhe, Germany)
Acrylamide (38%, 2% Bisacrylamide)	Roth (Karlsruhe, Germany)
Glycerol, 99 %	ICN, Costa Mesa, USA
peqGOLD 1 kb DNA ladder	Peqlab (Erlangen, Germany)
5,5’ Dithio-bis-Nitrobenzoic acid (DTNB)	Sigma (St. Louis, USA)
$15NH_4SO_4$	Euriso-top (Saarbrücken, Germany)
Polyoxyethylen Sorbitan monolaurat (Tween 20)	Merck (Darmstadt, Germany)
Alexa Fluor™ 488 NHS Ester (Succinimidyl Ester)	Invitrogen (La Jolla USA)
Alexa Fluor™ 488 C5 Maleimide	Invitrogen (La Jolla USA)
PEG 4000	Roth (Karlsruhe, Germany)

Buffers

Protein purification:

Ni-NTA A	20 mM Tris, pH 7.3 1 mM EDTA 1 mM DTT
Ni-NTA B	20 mM Tris, pH 7.3 1 mM EDTA 1 mM DTT 300 mM Imidazole
Ni-NTA C	40 mM HEPES, pH 7.5 20 mM KCl, 1 mM EDTA, 1 mM DTT,
Ni-NTA D	40 mM HEPES, pH 7.5 20 mM KCl, 1 mM EDTA, 1 mM DTT, 500 mM KCl
Res. Q A	20 mM Tris, pH 7.3 1 mM EDTA 1 mM DTT
Res. Q B	20 mM Tris, pH 7.3 1 mM EDTA 1 mM DTT 500 mM NaCl
SEC	20 mM Tris, pH 7.3 1 mM EDTA 1 mM DTT
Res. S A	25 mM Tris, pH 7.5 1 mM DTT 10% (v/v) Glycerin
Res. S B	25 mM Tris, pH 7.5 1 mM DTT 500 mM NaCl 10% (v/v) Glycerol
SEC raf	25 mM Tris, pH 7.5 1 mM DTT 10% (v/v) Glycerin
Dialysis buffer	40 mM HEPES/KOH, pH 7.5 20 mM KCl 1 mM EDTA 1 mM DTT

Storage buffer	20 mM K_2HPO_4, pH 7,5 300 mM KCl, 1 mM DTT, 2 mM EDTA
GST buffer	20 mM K_2HPO_4, pH 7,5 300 mM KCl, 1 mM DTT, 10 mM GST (reduced) 2 mM EDTA

MALDI sample preparation

Wash buffer 1 (W1)	10 mM ammonium hydro carbonate solution
Wash buffer 2 (W2)	1:1 ACN/ 10 mM NH_4HCO_3
Wash buffer 3 (W3)	1 mM β-mercaptoethanol
Reduce buffer 1 (R1)	0.5 mM Guanidine/HCl, 0.8 mM EDTA, 0.1 mM Tris/HCl, 1 mM DTT, pH 8.2
Reduce buffer 2 (R2)	92.5 mg IAA in 1 mL 0.1 M Tris/HCl
Matrix	300 µL 40% ACN, 200 µL 1:1 ACN/TFA, 10 mg α-cyano-4-hydroxyl cinnamic acid
Standard peptide mixture	400 µL Matrix, 5 µL Arycotensin 2, 5 µL ACTH 1:24

SDS-gel preparation:

Stacking-gel	125 mM Tris/HCl, pH 6.8 0.2 (w/v) SDS 5 % (w/v) Acrylamide Polymerization is started by the addition of TEMED, 2.5 µl/gel and APS 25 µl/gel
Separation-gel	62.5 mM Tris/HCl, pH 8.8 0.2 % (w/v) SDS 12.5% (w/v) Acrylamide Polymerization is started by the addition of TEMED, 5 µl/gel and APS,50 µl/gel 1 mM DTT
Running buffer	25 mM Tris/HCl, pH 6.8 200 mM Glycine 0.1 (w/v) SDS
Fairbanks A	2.5 g Coomassie Brilliant Blue R

	250 mL Ethanol 80 mL Acetic acid ad 1 L ddH_2O
Fairbanks D	250 mL Ethanol 80 mL Acetic acid ad 1 L ddH_2O

Standard buffer used in AUC, CD and TSA measurements:

M-buffer	40 mM Hepes, pH 7,5 50 mM KCl, 5 mM $MgCl_2$
H-buffer	40 mM Hepes, pH 7,5 50 mM KCl

Electrophoretic mobility shift assay:

10X Binding buffer	100 mM Tris, pH 7.5 10 mM EDTA, 1 M KCl, 1 mM DTT, 50% v/v glycerol, 0.10 mg/mL BSA
DNA buffer	10 mM HEPES, pH 7.0 100 mM NaCl 2.5 mM MgCl2
Loading buffer (10X)	50% Glycerin 10 mM EDTA 0.5% Bromphenolblue
10X Tris-acetate EDTA (TAE) electrophoresis buffer	400 mM Tris, 25 mM EDTA, brought to pH 7.8 with acetic acid.

Yeast transformation:

PLATE mix	40% PEG 4000 100 mM LiAc 10 mM Tris/HCl (pH 7.5) 1 mM EDTA

Nematode growth and maintenance

M9	3 g KH_2PO_4 6 g Na_2HPO_4 5 g NaCl 1 mL 1 M $MgSO_4$ ad 1 L, autoclaved ddH_2O
PBS	5.44 g KH_2PO_4 7.11 g Na_2HPO_4 2.34 g NaCl Ad 1 L ddH_2O
PBST	PBS + 0,05% (v/v) Tween 20
Bleach	3.5 mL ddH_2O 0.5 mL 5M NaOH 1 mL 5% Natriumhypochloride

Mass spectrometry with Q-Exactive

solvent A	0.1% TFA in ddH_2O
solvent B	0.1% TFA, 50% ACN in ddH_2O
solvent C	0.1% formic acid (FA) in ddH_2O
solvent D	0.1% FA, 84% ACN in ddH_2O

Media

NGM	3g NaCl 2.5g Peptone 17g Agar Add 1L ddH_2O After autoclave add 1M $CaCl_2$ 1mL 5 mg/mL cholesterol 1mL 1M $MgSO_4$ 1mL 1M K_2PO_4, 25 mL
RNAi Plates	Add extra to NGM: 1mM IPTG 6 µg/mL Tetracycline 50 µg/mL Ampicillin
YPD	Sigma-Aldrich (Hamburg, Germany) 50g/L
LB0 Media	Sigma-Aldrich (Hamburg, Germany) 20 g/L
DMEM	Sigma-Aldrich (Hamburg, Germany) D5030
opti-MEM	ThermoFisher, 31985062

Enzyme and kits

Antarctic phosphatase	New England Biolabs (Ipswich, USA)
DNase	Sigma Aldrich (St. Louis, USA)
GoTaq polymerase	New England Biolabs (Ipswich, USA)
Restriction enzymes	New England Biolabs (Ipswich, USA)
RNase A	Sigma Aldrich (St. Louis, USA)
T4 ligase	New England Biolabs (Ipswich, USA)
Aconitase Activity Assay Kit	Sigma (St. Louis, USA)
NAD/NADH Quantification Kit	Sigma (St. Louis, USA)
Western Blot detection system, Western-BrightTM ECM spray	Advansta (Menlo Park, USA)
Wizard® Plus SV Minipreps DNA Purification Systems	Promega (Fitchburg, USA)
Wizard® SV Gel and PCR Clean-Up System	Promega (Fitchburg, USA)
SV Total RNA Isolation System	Promega (Fitchburg, USA)
Brilliant III SYBR® Master Mix Kit	Agilent (Santa Clara, USA)
Q5® High-Fidelity DNA Polymerase	New England Biolabs (Ipswich, USA)
AlamarBlue™ Cell Viability Reagent	ThermoFisher, DAL1025

Strains and Organisms

AU0010	Dod24-GFP	Prof. Frederick Ausubel (Harvard University)
AU0185	Clec60-GFP	Prof. Frederick Ausubel (Harvard University)
BC12422	Zip-8a	CGC, Minneapolis, USA
HZ455	sepq1_GFP	CGC, Minneapolis, USA
MQD 396	PUD-1.1	Prof. Dong Meng-Qiu (National Institute of Biological Sciences, Bejing)
OP201	PQM1-GFP	CGC, Minneapolis, USA
RB711	PQM1	CGC, Minneapolis, USA
RT688	Cav-1-GFP	CGC, Minneapolis, USA
RW10375	His72-GFP, Tbx9-cherry	CGC, Minneapolis, USA
SD1583	myo-3-GFP, C50F7.5p-cherry	CGC, Minneapolis, USA
SHK207	Skr-5	Prof. Sivan Henis-Korenblit (Bar-Ilan University, Israel)

Plasmids

Hsf-1 DBD	Amp	pGATE
Hsf-1-N	Amp	pGATE
BRaf-WT	Amp	pet-28
BRaf-464	Amp	pet-28
BRaf-469	Amp	pet-28
BRaf-509	Amp	pet-28
BRaf-581	Amp	pet-28
BRaf-600	Amp	pet-28
BRaf-WT	Amp	pcDNA3
BRaf-464	Amp	pcDNA3
BRaf-469	Amp	pcDNA3
BRaf-509	Amp	pcDNA3
BRaf-581	Amp	pcDNA3
BRaf-600	Amp	pcDNA3
FKB-6	Kan	pet-28b
FKB-6 ΔFK1	Kan	pet-28b
FKB-6 ΔFK2	Kan	pet-28b
Cdc-37 *C.elegans*	Kan	pet-28b
Daf-21	Kan	pet-28b

Primers

FK2-Fwd	AGG AAGCGGAGGAGGAGGAAGCCCAGCGACTTGGGAAATG
FK2 rev	CCTCCTCCGCTTCCTCCTCCTCCCGGAGAGATATCCTCGGC
FK1 Fwd	GGAATTATGGACAGGGATGGGACGAT
FK1 rev	GGAATTCTAAGCATTGGAAGTAGAGCT
A481F fwd	TGGTGATGTTttcGTTAAAATGCTG
A481F rev	TGCCATTTACCTTTATAAAC
V600E fwd	TCTGGCAACCgaaAAAAGCCGTT
V600E rev	CCAAAATCACCAATTTTCACGG
R509H fwd	TATTCACATGatgGGTTTTACGC
R509H rev	TTCTGCTGTTTATGGGTTATAG

Hsf-1 (Nde)	TTAAGG CATATG cagc caacagggaa tcaaatacaa
Hsf-82 (Nde)	TTAAGG CATATG aacaaag ttgacgacga caa
Hsf-310 (XhoI)	attttcctcgag ttattc agagttcattcgagcacgtt ttgagttggg
Hsf-334(XhoI)	attttcctcgag ttaacggg aaaatggttc ctgagtctcg cgttgaagtg
Hsp-1fw:	tgacgaaattcgaattttctagaatcccgccacgc
Hsp-1 rev:	gcgtggcgggattctagaaaattcgaatttcgtca
F44E5.4 fw	gcagtggaatattccagaagcttctagaagaagtt
F44E4.5 rev	aacttcttctagaagcttctggaatattccactgc
Bag-1 fw	aaaagtgtcgagaatgttcacgaaaaatcgttaga
Bag-1rev	tctaacgatttttcgtgaacattctcgacactttt
Bag-1-m-fw	aaaagtgtcgagaatgtctacgaaaaatcgttaga
Bag-1-m-rev	tctaacgatttttcgtagacattctcgacactttt
F44E5.4-m-fw	gcagtggaatattccaagagcttctagaagaagtt
F44E5.4-m-rev	aacttcttctagaagctcttggaatattccactgc
DNJ-13_fwd	agtaaatagaacgctctggaaagttccgcactctt
Dnj-13-rev	aagagtgcggaactttccagagcgttctatttact
Unc-23_fwd	acggagcctcggattcttccagaaaattgagtctc
Unc-23-rev	gagactcaattttctggaagaatccgaggctccgt
Dnj-12_fwd	aaaagtgtcgagaatgttcacgaaaaatcgttaga
Dnj-12-rev	tctaacgatttttcgtgaacattctcgacactttt
Hsp-16.2a	gccttacagaatgttctagaaggtcctagatgcat
Hsp-16.2a-rev	atgcatctaggaccttctagaacattctgtaaggc
Hsp-16.2b	acaagcagctcgaatgttctagaaaaaggtggaaa
Hsp-16.2b-rev	tttccacctttttctagaacattcgagctgcttgt
Hsp-70_fwd	agtaaattgtagaaggttctagaagatgccagagg
Hsp-70-rev	cctctggcatcttctagaaccttctacaatttact

Equipment

Beckman Proteome Lab XLA ultracentrifuge	Beckman Coulter, Brea
fluorescence detection system	Aviv Biomedical (Lakewood, NY)
Äkta FPLC	GE Healthcare (Little Chalfont, UK)
JEM-1400	JEOL (USA)
Agilent 2100 Bioanlayser system	Agilent (Santa Clara, USA)
Mixer Mill MM400	Retsch
UltiMate® 3000 UPLC system	Thermo Fisher Scientific
Q Exactive Hybrid-Quadropole-Orbitrap mass spectrometer	Thermo Fisher Scientific
75 µm x 20 mm Acclaim® PepMap100 C18 column with 5 µm particle size	Thermo Fisher Scientific
Leica MZ-16FA fluorescence microscope	Leica (Wentzler, Germany)
Ultra-Flex III	BRUKER
Light cycler	Roche
Carry-100 UV-VIS spectrometer	Varian (Palo Alto, USA)
Nano Drop ND-2000	Peqlab (Erlangen, Germany)
HisTrap HP 5 ml	GE Healthcare (Little Chalfont, UK)
Resource Q/S 6 ml	GE Healthcare (Little Chalfont, UK)
Superdex 200 Prep grade (130 ml)	GE Healthcare (Little Chalfont, UK)
Superdex 200 Prep grade (240 ml)	GE Healthcare (Little Chalfont, UK)
Superdex 75 prep grade (130 ml)	GE Healthcare (Little Chalfont, UK)
Superdex 75 Prep grade (240 ml)	GE Healthcare (Little Chalfont, UK)

Computer programs and databases

Mascot Distiller	Matrix Science Inc., USA
Cystoscape	National Resource for Network Biology
GEO microarray repository	NCBI
RMAExpress	Written by Ben Bolstad, bmb@bmbolstad.com.
pLink	Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China.
PANTHER	Paul Thomas
YEASTRACT	http://www.yeastract.com/index.php
PyMol	Schrödinger
Chimera	University of California, San Francisco
MaxQuant 1.5	https://www.maxquant.org/
Microsoft Office 2010	Microsoft (Redmond, USA)
OriginPro 8.6 G	OriginLab Corperation (Northampton, USA)
SerialCloner 2.6.1	Franck Perez (Serial Basic)
UltraScan	(Demeler et al, 2010)
HADDOCK 2.2	BonvinLab
EMAN 2.0	University of California San Francisco
CryoSparc	https://cryosparc.com/
RELION	developed in the group of Sjors Scheres, MRC Laboratory of Molecular Biology
IMAGIC	http://www.imagescience.de/imagic.html

3.2 Experimental Methods

Yeast

Transformation of plasmids

Transformation of yeast is based on an optimized lithium acetate transformation method. 1 mL two-days culture were spin down at 6 krpm, washed with water, resuspend in 150 μL PLATE-mix and mixed with 4 μL 1M DTT, 2 μL salmon sperm and 200 μg plasmid. Cells were cultured at 30°C for 30 min and then heat shocked at 42 °C for 20 min. After centrifugation, yeast cells were resuspended in 100 μL water and transformed onto selective plates. Colonies were obtained after incubation at 30 °C for 2-4 days.

MALDI identification of proteins

Protein bands were picked and cut into ca. 1 mm^2 before use. Samples were washed with 100 μL W1 for 10 min, the supernatant was discarded and the washing step was repeated once. After samples were washed with 100 μL W2 for 10 min for the third time, the supernatant was discarded. Samples were incubated in 100 μL freshly made buffer R1 at 37 °C for 30 min. 10 μL buffer R2 was added and samples were incubated in dark at RT for 15 min before the supernatant was discarded. After addition of 100 μ L W1 and 5 μL W3, samples were washed for 10 min. After washed with 100 μL W2 at RT for 10 min and with 100 μL for another 10 min twice, samples were dried by speedvac at 30 °C for 5 min, followed by overnight digestion in trypsin. Samples were loaded on MS target and analyzed with RP-proteomics method. Raw data of MALDI is analyzed via Flex analysis and Biotech. Finger print peaks are identified by loading the data to www.matrixsience.com.

Co-expression data base

The co-expression database which is specific for the gene chip yeast genome 2.0 arrays is generated from the 3120 GPL2529 datasets from the GEO microarray repository with experiments for S. pombe excluded. The normalization is done by the software RMAExpress, followed by the Pearson correlation calculation for each gene-gene-pair with the tools from www.sciencecode.net. For 5813 probe sets, all the correlation coefficients for each yeast gene and the rankings were stored a systematic database. After converting the gene probe IDs into the commonly used yeast gene names with the info's from the GPL2529 platform. However for

most of the affymetrix control probes, a respective gene name can't be identified, in this case the probe set IDs were remained.

Generation and evaluation of gene clusters

The top 10 coregulated gene from the database for each gene is used to generate a genome-wide coexpression network. The network matrix for the whole genome was sorted by the strongest gene-gene coexpression, from which clusters were isolated following these rules: New cluster is generated when none of the genes in the coexpression pair has been included in an existing cluster yet. Genes in the same coexpression pair belong to the same cluster. Small clusters with a total gene number less than six were fused in to large one, which can be linked to it via a new gene pair. If both clusters contain more than five genes and can be linked by a new assigned gene pair, then both genes remain as part of the previously assigned cluster.

Upon separating all the gene coexpression pairs, the isolated clusters were named after the gene pair with the highest number of intra-cluster connections. The clusters together with the genome-wide network are available from https://github.com/klarichter/clusterSeparator.

Analysis of microarray sample

Expression values of experimental microarray data sets from the SPELL-server and the previous measurement were exported to Cystoscape, where the networks were visualized with corresponding coloring. The thresholds were set at 0.25 steps from -1 (in green) to 1 (red). Average expression values and the UpRegScore as defined in Papsdorf et al. [21] were calculated for each cluster.

The top 200 genes in each directions were used to determine the GO and TF- enrichment, whose results is described in (–log10(p)) for each cluster. 20 randomly genes were used as control in the analysis. With the YEASTRACT and the PANTHER web service the transcriptional factor which regulates the clusters and the biological processes which involve these clusters were analyzed. The significance of parameters for each clusters were described with averages (z-score) and standard deviation (p-values). RMSD and MAE for each gene were calculated to evaluate the genome-wide performance.

To allow general use of these 72 expression cliques and the described analysis methods, this microarray data analysis are integrated into the webserver at www.clusterex.com.

C.elegans

Maintenance and growth

Nematodes were handled according to established protocols. For normal growth, nematodes were cultivated on NGM plates and fed with OP50 bacteria at 20 °C. During RNAi experiments nematodes were maintained on NGM plates supplemented with ampicillin, tetracycline and 1 mM IPTG. HTT15 (DE3) bacteria were transformed with the plasmids L4440 or L4440-daf-21 and were grown in 5 mL LB medium (Amp, Tet). At an OD600 of 0.6 dsRNA expression in these cultures was induced with 1 mM IPTG for four hours. After that the dsRNA-expressing bacteria were placed on RNAi plates and synchronized L1 nematodes – prepared by bleaching and hatched overnight in M9-buffer – were added onto plates with induced bacteria. RNAi experiments were performed at 20 °C until phenotypes became obvious after 2.5-3 days. Nematodes were then harvested for proteomic or transcriptomic analysis or for imaging. Small heterogeneity in RNAi-treated nematodes was observed at the time of harvesting in different biological replicates. Thus control nematodes were cultivated for one further day to confirm the full development of the expected phenotypes.

Protein expression and purification

Human sB-Raf and all other involved chaperons were expressed from a pET vector with His6-tag. Plasmids were transformed in the E.coli strain BL21. Cultures were incubated at 37 °C, induced with 1 mM IPTG at an OD of 0.6 and harvested 16 hours post induction at 18 °C. For sB-Raf purification, cells were lysed in buffer A (25 mM Tris/HCl, pH 8, 20 mM KCl, 1 mM DTT and 10% v/v glycerol). Lysate was clarified by centrifugation (20,000 rpm for 45 min). Then the lysate was purified via a Resource S column with elution in buffer B (25 mM Tris/HCl, pH=8, 250 mM NaCl, 1 mM DTT, and 10% v/v glycerol). Proteins were further purified by Superdex 200 size-exclusion chromatography in buffer A. CDC-37, HSP-90, STI-1, FKB-6, CDC-37, human HSP-90 and PPH-5 were purified first via His-trap 5 mL FF column and then further by gel filtration chromatography in buffer C (40 mM Hepes/KOH, pH 7.5, 20 mM KCl). Expression plasmids for ΔFK1 and ΔFK2 were generated using the Q5 mutagenesis kit with primers designed by NEBaseChanger. FKB-6 mutants were purified via a His trap column in buffer (40 mM Hepes/KOH, pH 7.5, 1 mM EDTA, 1 mM DTT.) followed by a Resource Q column. The loading on the Resource Q column is done in buffer D and elution from the resource Q column is done with buffer E (40 mM Hepes/KOH, pH 7.5, 1 mM EDTA, 1 mM DTT and 500 mM NaCl). ΔFK mutants were purified via a Superdex 70 size-exclusion chromatography in buffer D. The purity and molecular mass of the protein were verified by SDS-PAGE and MALDI mass spectrometry of an ultra-Flex III (Bruker).

Hydropathic plots and sequence alignments of different Hsf proteins from diverse species were compared to define the N-terminal and C-terminal residues of the DNA-binding domain of *C. elegans* HSF-1. Based on these results, a fragment starting at the 82nd amino acid and extending to a hydrophilic linker region at the 198th amino acid was subcloned into the pGATE vector (HSF-1 DBD) and fused with a GST-tag. HSF-1 DBD was expressed and purified using a GST-tag column followed by cleavage of the GST-tag via TEV-protease. The protein fragment was further purified via ion exchange chromatography and size-exclusion chromatography (all columns from GE Amersham Biotech). The purity of the fragment was assessed by SDS-PAGE and the identity of the protein was confirmed by peptide fingerprinting with mass spectrometry on a Bruker ultra-flex III MALDI-TOF/TOF instrument (Bruker).

The N-terminus of HSF-1 contains a nematode specific extension of 82 amino acids. The fragment was subcloned into the pGATE vector (HSF-1 NBD) and expressed and purified it in a manner similar to HSF-1 DBD and confirmed its purity and identity as described.

Thermo shift assays

TSA assays were performed with 0.2 mg/mL protein mixed with 2 µL 1:100 v/v SYPRO-Orange solution (Life Technologies, Carlsbad, USA) and 2 mM nucleotide in 20 µL TSA buffer (20 mM Tris/HCl, pH 7.5, 50 mM KCl, 1 mM EDTA, 5 mM MgCl2, 1 mM DTT, 10% v/v glycerol). Measurements were carried out with the real-time qPCR System Mx3000P (Agilent Technologies Inc., Santa Clara, USA) at an excitation wavelength of 470 nm and an emission wavelength of 570 nm. During the experiment, the heating temperature was increased at a rate of 1 °C per cycle in 71 cycles from 25 °C to 95 °C. Experiments were set up in triplicates.

CD measurements

Proteins were measured at a final concentration of 0.2 mg/mL. 300 µL samples are transferred to a cuvette and CD spectrum is recorded at 20 °C in the wave length range of 210 nm to 280 nm. The thermo transition is then measured at 220 nm between 20 °C to 90 °C.

EMSA shift assays

Single strain DNA (ssDNA) of identical length, but containing sequences representing the promoter regions of *F44E5.5, Hsp-16.2a, Hsp-16.2b, hsp-1, hsp-70* and *Dnj-13* were obtained (MWG Biotech, Ebersberg). Equal amounts of ssDNA were mixed with the reverse complement sequence, heated up to 95 °C and allowed to align at room temperature. EMSA shift assays were performed to observe the interaction between HSF-1 DBD and HSE-containing dsDNA. To this end HSF-1 DBD was added to the dsDNA and the resulting sample was analyzed on native PAGE. To visualize the DNA, the polyacrylamide-gel was incubated with SYRR green dye and analyzed in a Typhoon Fluorescence Scanner (General Electric) using the Alexa Fluor filter at 532 nm. The strength of binding was qualified based on the intensity of the stained complex bands in ImageJ.

Aggregation assay

The absorbance of protein is recorded at 360 nm at 30 °C with a Carry-100 bio UV-VIS spectrometer. 3 μM protein is measured either alone or mixed with 1 mM nucleotides. The total sample volume is 120 μL. The data was recorded for 2 hours.

ATPase activity assays

The ATP turnover of HSP-90 alone and in the presence of cochaperones was analyzed by ATPase activity assays. To this end, an ATP-regenerating system containing pyruvate kinase, phosphoenolpyruvate, NADH, and lactate dehydrogenase was used. Measurements were performed at 25 °C using 3 μM HSP-90 and 1.5 μM cochaperons in 40 mM HEPES/KOH, pH 7.5, 20 mM KCl, 1 mM DTT, and 5 mM $MgCl_2$. Reactions were started by adding 2 mm ATP, and the absorbance of NADH was recorded at 340 nm. Background activities were detected after adding the HSP-90-specific inhibitor Radicicol (Sigma-Aldrich). The activity of HSP-90 was calculated using the following equation with $\epsilon(NAD^+) - \epsilon(NADH) = -6200\ M^{-1}\ cm^{-1}$.

$$\text{Activity} = \frac{\frac{\Delta A_{340}}{\Delta t} - \left(\frac{\Delta A_{340}}{\Delta t}\right)_{\text{Background}}}{(\epsilon(NAD^+) - \epsilon(NADH)) \cdot c(\text{ATPase})}$$

Fluorescence labelling of protein

Cysteine residues of human Cdc37 and nematode CDC-37 was labeled using ATTO 488 maleimide (ATTO-Tec, Germany). A 2-fold molar excess of the label was added to 0.5 mg protein in a buffer of 40 mM HEPES/KOH, pH 7.5, 20 mM KCl and 10% (v/v) glycerol. After an incubation time of 3 h at 20 °C, the reaction was stopped by adding 20 mM DTT and free label was separated from the labeled protein via dialysis in the same buffer. The degree of labeling and the protein concentration were determined by UV/VIS spectroscopy using following equations:

$$A_{Protein} = A_{280} - A_{500} \cdot (CF_{280})$$

$$DOL = A_{500} \cdot MW / [protein] \cdot \varepsilon_{dye}$$

Where CF_{280} = 0.09 and ε_{dye} = 90,000 M^{-1} cm^{-1} according to the manufacturer (Invitrogen). The degree of labeling for sB-Raf was found to be 1.12.

Analytical ultracentrifugation (Kinase interaction)

AUC experiments were performed in a Beckman Proteome Lab XLA ultracentrifuge (Beckman Coulter, Brea) equipped with a fluorescence detection system (Aviv Biomedical, Lakewood, NY) and a Ti-50 rotor (Beckman Coulter, Brea) at 20 °C and at 42,000 rpm. 300 nM of the labeled protein were analyzed in the absence and presence of 0.5-1 µM of its unlabeled putative binding partners or nucleotides at indicated concentrations. Measurements were performed in the storage buffer (40 mM Hepes/KOH, pH 7.5, 20 mM KCl, 1 mM DTT, and 5 mM $MgCl_2$). dc/dt-plots were calculated by subtracting scans within a certain range and averaging over the different. Plots were normalized against the starting fluorescence intensity and were fit to bi-Gaussian functions to get good estimates of the sedimentation coefficient $s_{20,w}$ and the peak amplitude.

Analytical ultracentrifugation and data analysis (HSF-1 Titration)

Analytical ultracentrifugation was performed to quantify the binding of HSF-1 DBD to dsDNA sequences. To this end single strain DNA sequences from different promoter regions of the same length were mixed with their reverse complement strand in storage buffer, heated up to 95 °C and then slowly cooled to RT to generate the double stranded DNA product that matches the promoter region. To 1.5 µM dsDNA HSF-1 DBD was added at different stoichiometries (0 µM, 2.25 µM, 4.5 µM, 7.5 µM, 10.5 µM, 15 µM and 22.5 µM) and the absorbance of these samples was recorded in analytical ultracentrifugation sedimentation velocity experiments at 260 nm and 280 nm at 42,000 rpm.

Data analysis of individual samples was initially performed with UltraScan. To this end for all experiments 2DSA-analyses with the same setting (s-value range from 0 to 10 and f/f0 range from 1 to 4) were performed. This way a distribution of complex species was obtained for each experiment twice, once at 280 nm and once at 260 nm. The complexity in these results requires that for a unifying solution a strong reduction in search space must be enforced to focus Ultrascan's results and to allow a further analysis based on the species concentrations. The reduced model contains the most abundant species of the binding reaction, represented by HSF-1 DBD, ssDNA, dsDNA, dsDNA+1 HSF-1, dsDNA + 2 HSF-1, dsDNA + 3 HSF-1, dsDNA + 4 HSF-1 and dsDNA + 5 HSF-1 were derived. The direct determination of sedimentation parameters was possible for HSF-1 DBD, dsDNA and ssDNA and to some extent for dsDNA + 5 HSF-1 at full saturation of the F44E5.4 promoter dsDNA. Given that all DNA strands were of the same size, a unique value was assigned to these species independent of the dsDNA used. Estimating the numbers for the partially occupied species a sedimentation coefficient ($s_{20,w}$) and a diffusion coefficient ($D_{20,w}$) was obtained for them.

A custom grid containing only these species and their respective values was generated in UltraScan and used to repeat all analyses performed in unconstrained manner before. RMSD values of the unconstrained fit and the custom grid constrained fit were compared in tabulated form to ensure that the fit quality reduced due to the constraints is acceptable. Based on the constrained 2DSA analysis, percentages for each species were obtained and converted to concentrations using the extinction coefficients calculated for each species. Values for the 260 nm and the 280 nm experiments were calculated for final representation and then plot against the concentration of DBD and fit to KD-values for each step. This approach was used to obtain concentrations for each species in each of the experiments at both wavelengths.

Further, the specific volume of each species was calculated based on the following equation and used to confirm the MW of each species:

$$\bar{v}_c = \sum_{i=1}^{N} f_i \bar{v}_i = f_p \bar{v}_p + \sum f_{np} \bar{v}_{np}$$

The following value pairs for $D_{20,w}$ and $s_{20,w}$ were used and the extinction coefficients, specific volumes and molecular weights were calculated for each species and saved as 20190526-0949-CustomGrid.

Table 1 Settings of the custom grid applied in AUC analysis.

	$s_{20,w}$ (S)	$D_{20,w}$ (*1E-06)	Specific volume (estimated)	MW (kD)
HSF-1 DBD	1.5	1.2	0.72	13.921
dsDNA	3.0	0.92	0.54	21.500
ssDNA	1.7	1.4	0.54	10.843
dsDNA+1 DBD	3.8	0.82	0.61	35421
dsDNA+2 DBD	4.2	0.74	0.64	49342
dsDNA+3 DBD	5.4	0.65	0.66	63263
dsDNA+4 DBD	5.6	0.58	0.67	77184
dsDNA+5 DBD	5.8	0.54	0.68	91105

An alternatively approach based on global fitting of dc/dt plots derived with customized scripts from nearby scans was employed. dc/ct-plots were analysed by fitting experimental series of seven protein concentrations to the same set of Gaussian functions in the in-house software diffUZ. These Gaussian functions represent the free HSF-1 protein, the free DNA, the singly complexed species, the doubly complexed DNA, the DNA containing three, four and finally five molecules of HSF-1. In cases, where only three HSF-1s could bind as judged from the DNA sequence, the amplitudes of the Gaussian functions representing these higher occupied species were set to 0. The species containing HSF-1 were assigned $s_{20,w}$ values within reasonable boundaries. Error bars on the amplitudes were derived from varying these $s_{20,w}$ values and the resulting changes in amplitudes were evaluated.

RNAi assays

During RNAi experiments nematodes were maintained on NGM plates supplemented with ampicillin, tetracycline and 1 mM IPTG. HTT15 (DE3) bacteria were transformed with the plasmids L4440 or L4440-daf-21 and were grown in 5 mL LB (Amp, Tet). At an OD600 of 0.6 dsRNA expression in these cultures was induced with 1 mM IPTG for four hours. After that the dsRNA-expressing bacteria were placed on RNAi plates and synchronized L1 nematodes –

prepared by bleaching and hatched overnight in M9 buffer – were added onto the plates with induced bacteria. RNAi experiments were performed at 20 °C until phenotypes became obvious after 2.5 days. Nematodes were then harvested for proteomic or transcriptomic analysis or for imaging.

qRT-PCR

Nematodes were harvested 2.5 days after placement of L1 larvae on RNAi plates. Isolation of total RNA from the nematodes was performed with the SV Total RNA Isolation System following the manufacturer's instructions. Primers used in the qRT-PCR experiments were designed with Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The final primer sequences were: *act-1*_fwd (AATCCAAGAGAGGTATCCTTA), *act-1*_rev (GATGGCGACATACATGGCT), *skr-5*_fwd (AATTGGTGCTGGCAGCCAC), *skr-5*_rev (GTTACCCAAGTTGAAAACGGCACG). qRT-PCR samples were prepared with the Brilliant III SYBR® Master Mix Kit. The PCR program on the Agilent Mx3000p was selected as described in the user manual. After the PCR reaction, the amplified DNA was tested to confirm the purity and the correct length of the product. This was done by on-chip gel electrophoresis with the Agilent 2100 Bioanlayser system. The expression values were only used when a specific product is found at the right size. The experiments were performed in triplicates and averaged.

Native mass spectrometry/LILBID-MS (laser induced liquid bead ion desorption mass spectrometry)

A piezo-driven droplet generator (MD-K-130 from Microdrop Technologies GmbH, Norderstedt, Germany) was used to produce droplets of 30 µm diameter with a frequency of 10 Hz at a pressure of 100 mbar. Samples were loaded directly into the droplet generator and the generated droplets were subsequently transferred to high vacuum and irradiated by a mid-IR laser directly in the ion source. The laser employed was a Nd:YAG laser operating at 10 Hz, the wavelength being tuned by a LiNbO3 optical parametric oscillator to 2.94 µm ± 5 nm, the absorbing wavelength of the symmetric and asymmetric O-H stretching vibration of water. The pulse length was 6 ns with a maximum energy of 23 mJ. The laser power was measured by an optical power meter (PM100D, Thorlabs, Munich, Germany).

Droplet irradiation leads to an explosive expansion of the droplet containing the sample and solvated ions are released and analyzed in an in-house built time-of-flight setup operating at 10-6 mbar. The ion source is based on a Wiley-McLaren type design. The ions are accelerated into the grounded flight tube and guided towards the detector via a reflectron. The detector setup is based on a Daly-type detector optimized for the detection of high m/z ions. The voltage of the

first (repeller) and second plate was set to -4 kV in the ion source. The third plate was grounded. The repeller was pulsed to – 6.6 kV for 370 µs after droplet irradiation. The pulse was applied between 2-20 µs after the droplet irradiation (delayed extraction time). The einzel lenses were set to -3.0 kV. The reflectron was set to -7.2 kV. Post-acceleration was set to +17 kV at the MCP impact surface. Spectra processing was done by using the software Massign [95] based on LabVIEW.

All Proteins were used in a 25 mM Tris, 20 mM KCl, pH 7.5 buffers except for RAF-wt, which was in 25 mM Tris, 250 mM NaCl, 1 mM DTT and 10% Glycerol. The cross-linked sample was measured directly from the storage buffer, while the single proteins were subsequently mixes in a ratio of sbRaf-WT:Cdc37:FKB-6:Daf-21 = 1:1:1:1.

Crosslink reactions and purification of crosslinked products via HPLC

To allow efficient detection of crosslinked peptides, an isotopically labelled BS^3-crosslinker was used at a ratio of 1:1 with non-labelled BS^3. 10 µM of sB-Raf, CDC-37, FKB-6 and HSP-90 were pre-incubated in the crosslink buffer (40 mM Hepes/KOH, pH 7.5, 50 mM KCl) for 5 min. The crosslinking reaction was initiated by adding 500 µM H_{12}/D_{12}-BS^3 (Creative Biomolecules). After 30 min incubation at 25 °C, reactions were quenched by adding 5x SDS-PAGE sample buffer (250 mM Tris/HCl, pH 6.8, 10% SDS, 30% (v/v) Glycerol, 10 mM DTT, 0.05% (w/v) Bromophenol Blue). Samples were then separated in SDS-gel and complex bands were further analyzed in the MS.

Isotopic labeling of protein

To perform mass spectrometric analysis of nematodes, nematodes were labeled with 15N. 1 g/L $^{15}NH_4SO_4$ was used instead of the normal $^{14}NH_4SO_4$. Nematode growth was not affected by the 15N containing nitrogen source. Nematodes were harvested after 2.5 days by washing them off plates with M9 buffer. They were separated from residual bacteria by gravity sedimentation in a 15 mL plastic tube. The supernatant was removed and after three such washing steps nematodes were lysed in a Retsch Mixer Mill MM400 in 1 mL M9 buffer. The lysate was cleared via centrifugation and the protein concentration was determined by Bradford reagent. The isotope-tagged lysates were then mixed at the calculated ratio to ensure that equal amounts of each sample were combined. This sample was frozen and stored at -80 °C until mass spectrometric analysis was performed.

Mass spectrometry (HSF-1)

After a tryptic digest of the proteins, 15 µL of the sample peptides were analyzed by nano LC-MS/MS on an UltiMate® 3000 UPLC system online coupled to a Q Exactive Hybrid-Quadropole-Orbitrap mass spectrometer with electrospray ion source. Initially, the samples were desalted and concentrated on a 75 µm x 20 mm Acclaim® PepMap100 C18 column with 5 µm particle size with a flow rate of 30 µL/min using 95% solvent A and 5% solvent B for 7 minutes. Peptides were subsequently loaded onto a 75 µm x 50 cm Acclaim® PepMap100 RSLC C18 column with 2 µm particle size using a mixture of 95% solvent C and 5% solvent D. The elution of peptides was performed at a constant flow rate of 400 nL/min in a linear gradient of 5% to 40% solvent D over 120 minutes. Mass spectrometry was performed in the data-dependent acquisition mode with Full scan Fourier transform mass spectrometry acquired in an m/z range of 350 to 1400 with a resolution of 70,000. The ten most intense peptides (charge range ≥ +2) in the FTMS-scan were selected for higher-energy collisional dissociation with the collision energy set to 27. The dynamic exclusion time was set on 30 seconds with an MS/MS resolution of 35.000 within an m/z range of 300 to 2000.

Mass spectrometry (Kinase complex)

Crosslinked protein bands were hashed and destained by three times alternating 10-min treatments with 10 mM ammonium hydrogen carbonate (pH 8.3, buffer A) and buffer B (buffer A:100% acetonitrile a ratio of 50:50 (v/v)). After the second incubation with 50 mM ammonium bicarbonate, samples were treated with 50 µL 10 mM DTT for 1 h at 56 °C and with 50 µL 50 mM IAA for 45 min at room temperature, before the destaining protocol was continued. Finally, gel pieces were dried in a vacuum. Digestion was initiated by adding 8 µL of trypsin solution and was performed overnight. The digestion was stopped, and peptides eluted by incubating the gel pieces two times for 15 min with 30 µL of a 1:1 solution containing 100% acetonitrile and 0.1% (v/v) TFA in an ice-cooled ultrasonic bath. Samples were dried in a vacuum concentrator and resuspended in 20 µL 0.1% (v/v) TFA. Afterwards, the peptide con-centration was determined by amino acid analysis (AAA) as described by Plum et al. [96].

200 ng digested sample were measured by nano LC-ESI-MS/MS. An UltiMate 3000 RSLC nano LC system was utilized for nano HPLC analysis using the following solvent system: (A) 0.1% TFA; (B) 84% ACN, 0.1% FA. Samples were initially loaded on a trap column with a flow rate of 30 µl/min with 0.1% TFA. After sample concentration and washing, the trap column was serially connected with an analytical C18 column, and the peptides were separated with a flow rate of 400 nL/min using a solvent gradient of 4% to 40% B for 95 min at 60 °C. After each sample measurement, 1

h of column washing was performed for equilibration. The HPLC system was on-line connected to the nano-electrospray ionization source of a LTQ Orbitrap Elite mass spectrometer. The mass spectrometer was operated in a data-dependent mode with the spray voltage set to 1,600 V in positive mode and a capillary temperature of 275 K. Full scan MS spectra (mass range 350-2000 m/z) were acquired in the Orbitrap analyzer at a mass resolution of 60,000. The twenty most intensive ions per spectra were subsequently fragmented using collision-induced dissociation (35% normalized collision energy) and scanned in the linear ion trap. The m/z values triggering MS/MS were set on a dynamic exclusion list for 30 seconds.

Analysis of the MS results

The raw data obtained from the MS measurements were analyzed first with the software MaxQuant 1.5 [https://maxquant.org/maxquant/] using the fasta sequences of the crosslinked proteins. The generated peak lists from MaxQuant which contain the intensity and elution time were then analyzed with the in-house software xMass, which searches and reports the possible crosslinked peptide pairs. To this end, all MaxQuant-derived peaks were imported and xMass determined all potential pairs of peaks, which deviated in mass by 12.076 Da and could be assigned to a potential crosslinked product from the simulated product library. The threshold was set to 6 ppm, 2 miss-cleavages were allowed in each crosslinked peptide. The obtained hit list then was filtered based on the type of the crosslinked pair and the peak intensity in the parent scans. All of the peptide pairs that are identified as intermolecular pairs and have a similar parent scan intensity were collected and further checked, whether they were identified as both, heavy and light crosslinked products.

The MS results were further confirmed with the software pLink [http://pfind.ict.ac.cn]. The analysis was performed with the following settings: Peptide mass range: 400 – 6000 Da; Peptide length: 4-60 AA; precursor tolerance: ± 10 ppm; fragment tolerance: ± 10 ppm; filter tolerance: ± 5 ppm; FDR < 5% at PSM level. Reported crosslinked spectra were filtered to a mass error < 2 ppm and then compared to the xMass outputs.

Electron microscopy of protein complexes

Negative stain was performed with a 1:50 dilution of 10 μM sB-Raf complexes or with sB-Raf complexes purified by SEC-HPLC after crosslinking. The complex was stained for 1 min on carbon-coated copper grids with 2% (w/v) uranyl acetate after glow-discharging. Samples were imaged in a JEM-1400 microscope (JEOL USA Inc.) operating at 120 kV and were recorded in low-dose mode at a cs-value of 3.4 and 60,000 X magnification. The pixel size was 2.75 angstrom.

Image analysis

Defocus and astigmatism were determined using EMAN2.0. Boxed particles with a box size of 128 pixels from 135 micrographs were imported into IMAGIC for further processing. A total amount of 7913 particles were centered using three rounds of translational alignment to the rotationally averaged SELF. Particles were reference free aligned using MRA and hierarchical ascendant classified to generate reference classes to which all particles were aligned. This process was repeated iteratively. Well-defined classes representing distinct views of the complex were chosen for an angular reconstitution to generate an initial model in RELION. Additional classes were incorporated into the model with increasing Euler angle precision in three rounds of anchor set refinement. At this point, the reconstruction was 3D masked and refined by projection matching. The angle assignment was sorted by re-projection errors (agreement between the average for each orientation and the corresponding re-projection) and visually checked to ensure that the classes were acceptable. A 3D map was then created from acceptable classes. At this stage, there was no visible change in the reconstruction between rounds of refinement. The final reconstruction was generated using 5177 particles, containing 65% of the data.

Model Building

In EM the anticlockwise twist of the crystal structure of full-length HSP-90 [7] was used to assign the hand of the EM map. All crystal structures used for model building (HSP-90, PDB code 2CG9; HSP-90-CDC-37-Cdk4, PDB code 5FWL) were filtered to 20 Å using the mol-map command in chimera to allow a direct comparison with the EM re-construction.

For the complex structure, model docking was performed with the HADDOCK 2.2 webserver [97] with the PDB structure 5FWL and 1KT1 as initial structure. Docking is performed with default settings of the server, where the intermolecular crosslinked amino acids (FKB-6 (288)-HSP-90 (407) and FKB-6(71)-HSP-90 (585)) were used as restraints. In total nine structures were obtained from the server and the corresponding results were shown in the figure 4.4.8.

Mammalian cell

Cell culture and handling

Frozen cells were thawed and transferred into 5 mL fresh DMEM +10% FBS medium, spin down at 1 krpm for 5 min. After the medium change, the cells were resuspended in 10 mL growth medium in a T25 culture flask. Cells were maintained at 37 °C, 5% CO_2.

When the cells reached up to 80% confluent, the medium is discarded. Cells were then washed with fresh medium. 10^5 cells were inoculated in a new flask; 10 mL medium was added and incubated at 37 °C, 5% CO2.

Cell transfection

PEI solution at 1 µg/µl in sterile water is prepared, neutralized with HCl to pH 7.2 and sterilize filtered. The cells were washed 3 – 4 times with media shortly before transfection. DNA is mixed with Opti-MEM before adding PEI to the diluted DNA and vortexed for 10 seconds immediately. After 15 minutes of incubation at room temperature, the DNA/PEI mix was added into 6 well plates, well mixed and then incubated at 37°C, 5% CO_2. Cells were harvested 24h post transfection.

Cell viability test

The viability of cell culture was tested with the Alamar Blue kit. The ready-to-use Alamar Blue solution was added to cells, incubated for 2 hours at 37°C between pH 6.8 and pH 7.4. Cell respiration was measured according to the absorbance (570 nm and 600 nm).

Imaging

The images were taken by the HAMAMATSU camera from the Leica MZ-16FA fluorescence microscope at 40 X magnification, both under normal light or with filtered fluorescent light.

4 Results

4.1 Genome-wide analysis of yeast expression data based on co-expression clusters*

*this chapter has been published in microbial cell, 2019

4.1.1 Yeast expression cliques derived from platform-specific coregulation data

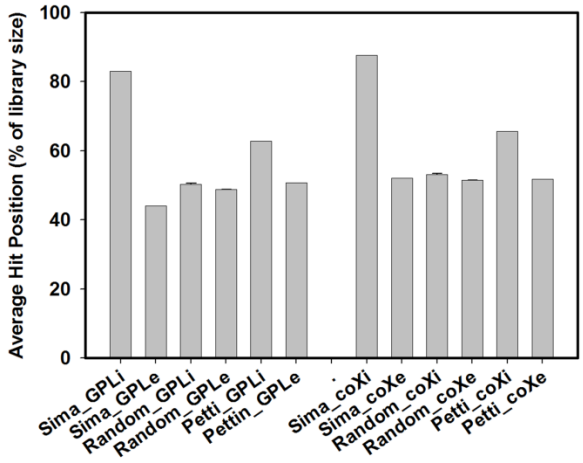
To obtain genome-wide co-expression cliques for all yeast genes analyzed on Gene Chip Yeast Genome 2.0 arrays the publicly available information from previously reported experiments was used. A co-regulation database based on the correlations within 3196 *S. cerevisiae* microarray experiments available for the GPL2529 platform in the GEO repository (Suppl. Figure 1.1) was generated. Using this database, a network termed 'GPL2529full' was generated, which contains 151,676 gene-gene connections in the full genome (Suppl. Figure 1.2) and then separated 72 expression cliques based on an optimized classification algorithm as described in the Material and Methods section (Suppl. Figure 1.3). The size of the cliques varied considerably with the smallest clique being 6 genes (SAN1-SKN7) and the largest clique being 775 genes (CUP5-TEF1). Most expression cliques were in the range of 15-100 genes and every yeast gene analyzed on the GPL2529 microarray platform was included exactly once.

To what extent the isolated cliques indeed contain genes with high rankings in the co-expression database 'GPL2529full' by calculating the average intra-clique ranking and the average inter-clique ranking. The first value is much higher, showing that indeed the cliques accumulate preferentially coexpressed genes (Figure 4.1.1 A). To exclude bias from using the same database for construction and evaluation of the cliques, another database is used, which is also publicly available (COXPRESdb) [18]. Here a similar spread is observable. In both cases this spread is eliminated, if randomly scrambled gene cliques are used (Figure 4.1.1 A). Finally this test is used to compare this clique set to another publicly available genome-wide clique set obtained with a different approach [98]. For this clique set the spread between intra- and inter-clique rankings is much smaller, suggesting that the clique set determined from the high-density networks could indeed show improved classification of the co-regulated genes.

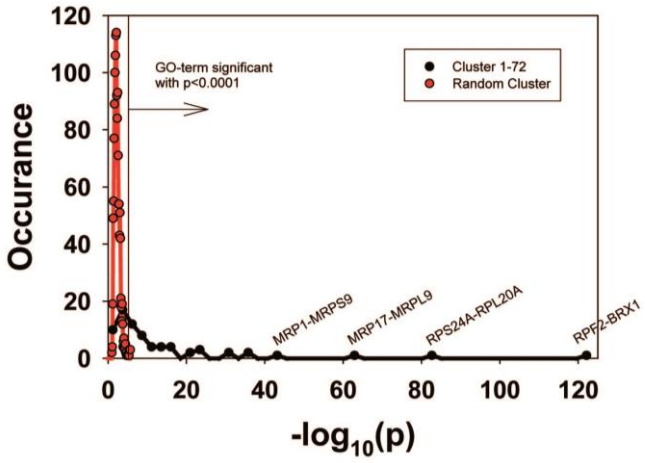
The isolated cliques were analyzed to evaluate functional correlations between the genes in each clique with determined enriched GO-terms for each of the isolated gene groups (the most prominent results are summarized in Table 1.1). Many of the determined $p_{Enriched}$-values were in the (-log10)-range of +7 to +13, some even reached a (-log10($p_{Enriched}$)) value of +124 (RPF2-BRX1, "nucleolus"). These values imply with very high confidence that the isolated co-expression cliques group genes of similar cellular functions. To evaluate the employed GO-term selection method, the scrambled expression cliques with the same enrichment evaluation method was tested. The random cliques observed mostly -log10($p_{Enriched}$) between +1 and +3 and even after

20 such scrambled clique-tests, the outstanding functional grouping of the genes in most co-expression cliques was obvious.

A.



B.



C.

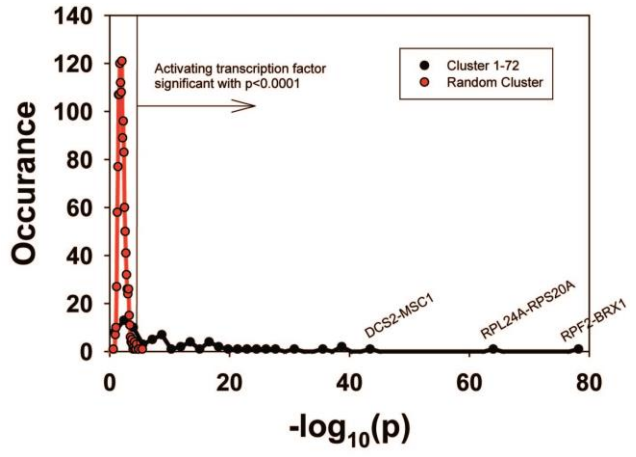


Figure 4.1.1 Evaluation of Genome-wide expression cliques. (A) Comparison between the calculated clique set, random genes and the clique set from Petti et al. The comparison is based either on the own co-expression database (left side) or on the co-expression database from COXPRESdb (right side). (B) GO-term assignments for the 72 identified cliques or for random cliques. The significance threshold for p<0.0001 is indicated as vertical line in the figure. As comparison to the calculated cliques (red plot) random gene cliques were used (black plot) (C) Assignment of activating TFs for the identified 72 cliques or for random cliques based on the YEASTRACT database. The significance threshold for p<0.0001 is indicated as vertical line in the figure.

It's assumed that at $p_{Selection}$<0.0001 (-log 10 ($p_{Selection}$)>4), corresponding to a Z-score larger than 3.72, sufficiently high significance is achieved and this Z-score requires a -log 10 ($p_{Enriched}$)>4.831 (Figure 4.1.1 B). Many cliques in fact gave highly significant assignments, like the cluster RPS24A-RPL20A ((-log 10 ($p_{Enriched}$) =+82, "cytoplasmic translation") or the cluster MRP17-MRPL9 ((-log 10($p_{Enriched}$)=+65, "mitochondrion organization") or MRP1-MRPS9 ((-log 10 ($p_{Enriched}$)=+42, "mitochondrion"). Many other cliques also were assigned with functions that obviously are correct. Altogether 4623 of the genes are included in cliques with very high significance (p<0.0001) and these are based on 1457 direct GO-term hits. The assignment of the most significant GO-terms to the separated gene cliques are listed in table 1.1 with the respective p-values and the significance thresholds derived from the control experiments.

Cluster No	Cluster Name	Cluster Size	Best_GO (2.2792 ± 0.6861)	Best YeastTF Binding (1.932 ± 0.58379) Second_YeastTF (1.5418 ± 0.3868) Third_YeastTFBind (1.3388 ± 0.3202)	Best YeastTF Activating (2.0476 ± 0.5872) Second_YeastTF (1.6848 ± 0.4249) Third_YeastTFActiv (1.5010 ± 0.36498)	Best YeastTF Inhibitory (2.0532 ± 0.6589) Second_YeastTF (1.6832 ± 0.5321) Third_YeastTF (1.5233 ± 0.3758)
4	RPF2-BRX1	263	Nucleolus (123.362639103465)	MSS11(4.96158410225936) SKO1(2.39561077378327) POG1(1.32824772405911)	SWI5(79.0064177435691) YOX1(76.2966688166691) YHP1(71.2869171493983)	RLM1(60.5970903703291) MSN2(38.7948863611814) SFP1(34.5169886341715)
25	RPS24A-RPL20A	56	cytoplasmic translation (81.8674670251124)	IFH1(59.9796314084271) SFP1(53.8519237725248) FHL1(43.6090002844126)	IFH1(63.9977426330421) GCR1(48.6642087928913) SDS3(38.0960111550602)	FKH1(63.3399033514049) MGA2(39.6691556613836) RAP1(37.5213597230813)
19	MRP17-MRPL9	83	mitochondrion organization (64.0751096675982)	YRM1(1.77552412453368) TOS4(1.7656386999132) SIP4(0.999543003592693)	CUP9(15.149061571022) TYE7(14.7210368513215) HAP2(13.7611923991662)	HIR3(5.17301569802421) HIR1(5.08840028676645) CST6(3.44234377674376)
21	MRP1-MRPS9	65	Mitochondrion (42.0640043506021)	YAP7(2.42325555137621) RTG3(2.30863082995615) NDT80(1.69806404354601)	RPH1(5.41200704215856) HSF1(5.01548488321318) MSN2(4.71255936866622)	HIR3(8.25294755608941) HIR1(6.19784626147915) SKO1(5.37119132106321)
8	PRE10-RPT1	180	proteolysis involved in cellular protein catabolic process (36.5081599380838)	RPN4(18.335578991259) REB1(3.89457980466554) HSF1(2.22503340073335)	RPN4(7.98857704983206) SPT3(3.45468103007458) SPT20(3.37948222425901)	AFT2(12.213409085956) ERT1(9.3908851787614) RDR1(8.02026573388743)
11	YOR365C-GAS4	148	meiotic cell cycle (35.6041918861595)	SUM1(22.5665807254708) THI2(1.77764591716636) PDC2(1.61942966950795)	HST1(36.3325111467661) SUM1(30.3114558239157) UME6(7.80995919285569)	SUM1(11.497860109969) RSF2(3.43950284594182) TEC1(3.43609155894238)
1	CUP5-TEF1	775	endomembrane system (29.9776604998377)	FHL1(7.43106059883068) SFP1(7.11958918929006) IFH1(5.01451226496547)	SAS3(24.7688170844467) HAP2(22.6293107600953) PGD1(22.2672173572115)	CST6(53.8577107846705) GCR1(48.3623105170302) RAP1(40.072659558028)
32	ERG10-ERG1	45	lipid metabolic process (29.9550335896853)	HAP1(24.7809537603695) INO2(7.24688246682872) MGA2(5.88473168479062)	INO2(11.3282531099314) OPI1(8.26942290457504) TDA9(8.02505203391379)	RAP1(12.7651209263292) HAP1(11.4195765020269) SFP1(10.133537194265)
23	TCM62-GUF1	59	Mitochondrion (24.1684221280945)	RTG1(1.87684121666229) YHP1(1.86290878841167) MIG2(1.60849351548175)	PPR1(2.20432875523954) MSN4(1.98978193793454) MSN2(1.80480991081465)	TEC1(1.96136945548008) GAT4(1.57244438417038) RSC2(1.05618266848047)
40	GPA1-STE4	34	Conjugation (23.8353164069696)	TEC1(5.45191180064392) STE12(4.9802549376848) MCM1(4.15118720237767)	YHP1(16.8505947854273) MGA1(8.38389450373479) YOX1(8.32095271295989)	SDS3(12.9199062928831) STP2(10.2342189258178) SIR3(9.26747627089548)
26	QCR7-COX6	53	mitochondrial envelope (22.8235114747937)	HAP4(23.6692603170055) HAP2(22.3084195499722) HAP3(21.5196736220755)	HAP4(20.6057590842643) SWI3(15.3780803775645) GCR1(12.0523799924958)	SUT1(19.39174012583) SOK2(15.3078454593917) GIS1(13.1574378776752)
43	LYS21-LYS12	28	cellular amino acid metabolic process (21.9730687277123)	GCN4(16.8953791952509) ARG81(4.21344427741765) MSN2(4.13647166196768)	GCN4(16.2595769412032) EDS1(14.5547125693737) LYS14(14.250805551449)	LEU3(7.26189850097212) SET2(6.71030441094563) RTG3(6.37005023041399)

Table 1.1. Most prominent results from GO-enrichment. The cliques with p-values <10^{-20} are ranked according to their p-values.

A similar type of analysis is performed to obtain information on potential transcription factors for the 72 expression cliques, where the information from YEASTRACT is used to determine the three most likely transcription factors in the categories of "Binding", "Activating" and "Inhibiting". Here for most cliques the strongest enriched transcription factor is supported by a (-log 10($p_{Enriched}$)) in the range of +5 to +30 (Table 1.1 and Figure 4.1.1 C). When performing an analysis on 20 scrambled control separations, these values are considerably lower, implying that the applied separation method leads to gene cliques, whose regulation apparently can be linked to specific sets of transcription factors. Based on average and standard deviation of the random sets, a Z-score of 3.72 is required and thereby the high significance threshold for the best transcription factor requires a -log 10($p_{Enriched}$)>4.103. Best (-log 10($p_{Enriched}$))-values in the category "activating transcription factor" were obtained for the cluster RPF2-BRX1 and RPS24A-RPL20A. Altogether 4474 genes were included in cliques with at least one highly significant transcription factor in the category of "activating TF", with 1444 direct gene hits in the YEASTRACT database. Further transcription factors were assigned from the two other categories. Thus the ability to assign common functional properties and specific transcription factors to most of the clustered genes suggests that the separation of the cliques correlates well with the transcriptional logic encoded in the yeast genome.

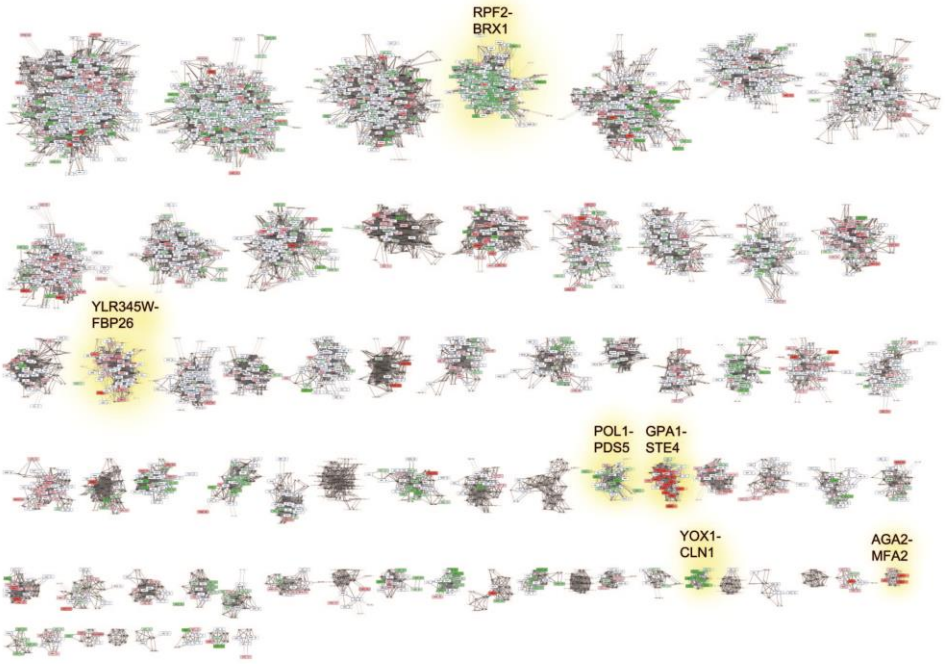
4.1.2 The a priori clusters provide detailed genome-wide transcriptional response information

Microarray experiments performed on other platforms were tested to prove, whether the separation into these 72 cliques generally reflects the experimental realities in isolated experiments.

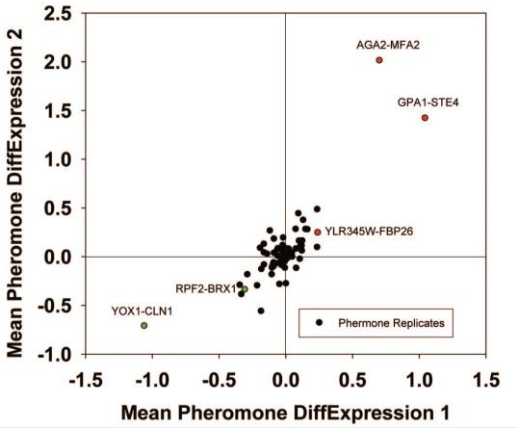
Starting with a reported response to α-pheromone (GSE7525, [99]), accumulation of certain cliques was observed when the expression difference of each gene was plotted onto the clustered network (Figure 4.1.2 A). The Top 200 hits in each direction was used to see, whether these preferentially fall into some of the 72 cliques. This clearly is the case as judged from the enrichment factors and p-values derived from this analysis and in particular the clique GPA1-STE4 ((-log 10(p) =+27) performs outstandingly, followed with a larger distance by the clique DCS2-MSC1 ((-log 10(p)=+9). The cliques RPF2-BRX1 ((-log 10(p) =+44), YOX1-CLN1 ((-log 10(p) =+18) and SER1-ADE12 ((-log 10(p) =+4.2) are significantly enriched for downregulated genes (Table 1.2 with the most prominent results). To confirm the up- or downregulation of these cliques relative to the other gene cliques, the average expression differences and the UpRegScore were calculated. The random clique analysis offers information on stochastic variations in these parameters and was able to derive significance parameters for them. The average expression a difference in RPF2-BRX1 is -0.30375 and upregulation in GPA1-STE4 is

+1.046. YOX1-CLN1 is significantly downregulated (-1.0574, (-log 10(p) =+15)). Several other cliques were also significantly shifted regarding their expression changes at lower, but still highly significant levels (-log 10(p) >+4). The two replicates were then compared. Here the similarity between these two experiments is very high and the correlation of expression differences (Figure 4.1.2 B) and UpRegScore (Figure 4.1.2 C) for the 72 cliques is almost linear. The strong correlation between the replicates shows that these two replicates yield very similar results in respect to the cliques induced or suppressed.

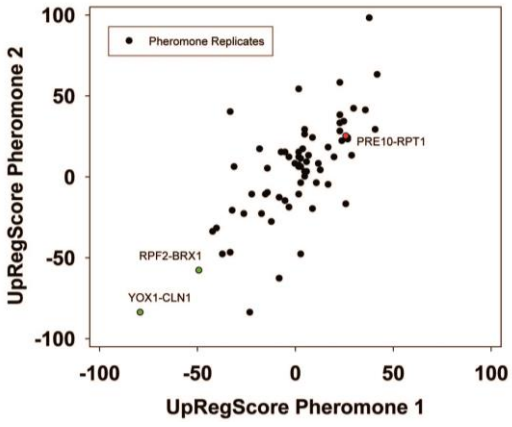
A.



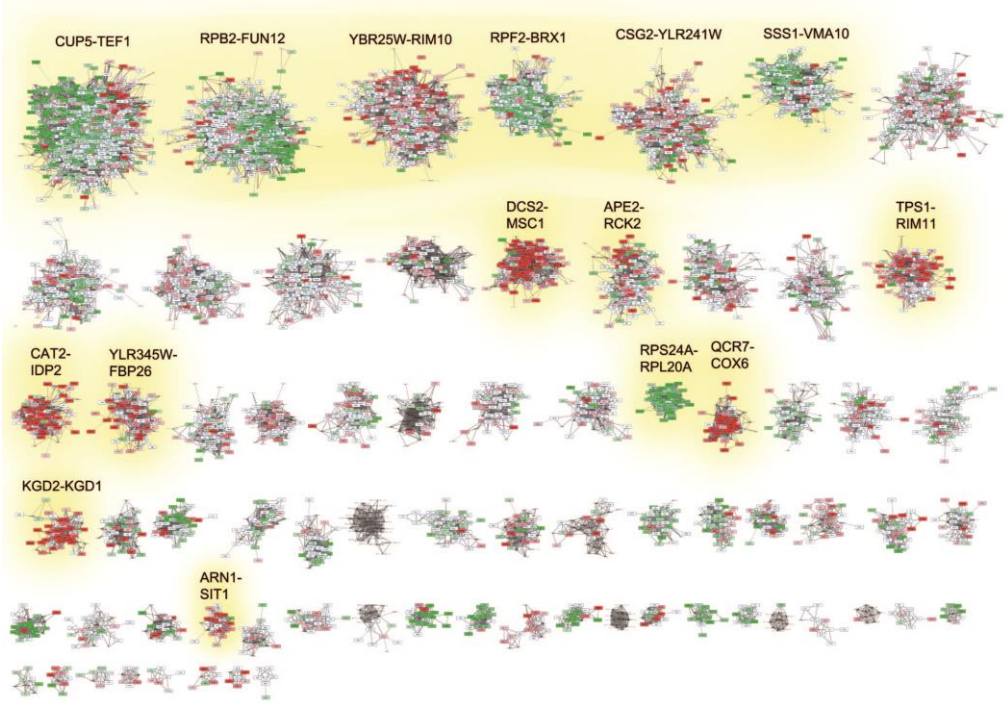
B.



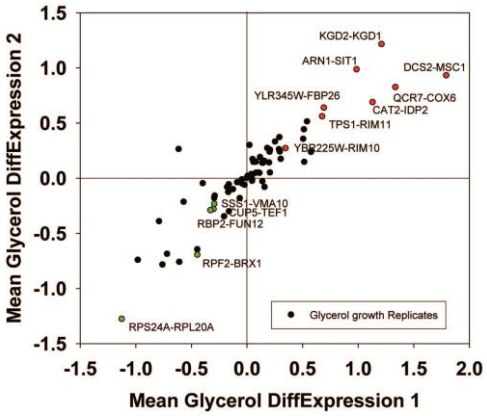
C.



D.



E.



F.

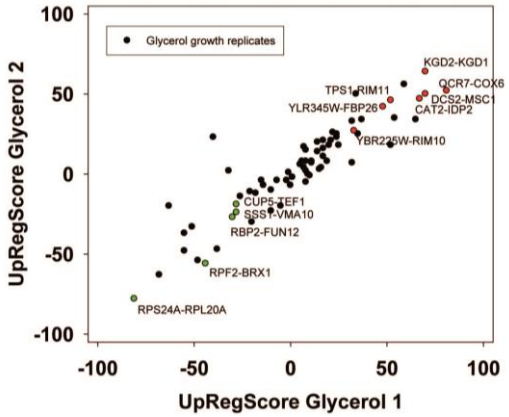


Figure 4.1.2 Analysis of microarray data sets on pheromone response and growth conditions. (A) Response to α-pheromone with data from [99]. Downregulation is shaded in four levels of green, upregulation in four levels of red. Genes, which are in the clustered network, but received no values in the experiment, were blanked out. Significantly changed cliques in replicate 1 are highlighted with yellow background and labelled with their clique name. (B) Comparison between two replicates in respect to average expression changes in the 72 cliques. The significantly changed cliques are marked in respect to average expression in both experiments ($p<0.0001$) in the plot with red (upregulation) and green (downregulation) and labelled them accordingly. (C) UpRegScore for each cluster in the two replicates. Cliques, where

both replicates showed significant shifts in the same direction were labelled in red (upregulation) or green (downregulation) and named accordingly. (D) Differences between glycerol-induced growth and glucose supported growth as described by [100]. Downregulation is indicated by four levels of green, upregulation by four levels of red. Significantly changed cliques of replicate 1 are highlighted with yellow background and labelled. (E) Correlation between two replicates and indication as to which values show significant upregulation in both replicates (red dots). The same is applied for downregulation (green labelling). (F) Comparison between two replicates regarding the UpRegScore of each clique. The labelling is performed as in B.

Clique Name	Clique Size	Genome Size	$Up_{Selected}$ Replicate 1	$Down_{Selected}$ Replicate 1	Differ. Exp Replicate 1	Differ. Exp Replicate 2	UpRegScore Replicate 1	UpRegScore Replicate 2
GPA1-STE4	23	3474	0		1.43	1.04	72	42
AGA2-MFA2	4	3474	2.09E-05		1.08	0.70	60	38
YLR345W-FBP26	45	3474	1.31E-07		0.25	0.24	42	41
TPS1-RIM11	62	3474	4.52E-08		0.20	0.10	44	23
PRE10-RPT1	139	3474	7.49E-07		0.14	0.11	32	26
RPB2-FUN12	392	3474		0.00082	-0.11	-0.087	-15	-15
POL1-PDS5	25	3474		0.613	-0.33	-0.28	-45	-42
RPF2-BRX1	208	3474		0	-0.35	-0.30	-53	-49
YOX1-CLN1	8	3474		4.51E-151	-1.29	-1.06	-94	-79

Table 1.2. Most prominent hits from differential expression after pheromone induction within the 72 expression cliques for replicate 1 and replicate 2. Results significantly above or below baseline ($p<10^{-4}$) are colored in the categories for Top 200-Enrichment of replicate 1, clique differential expression in both replicates or UpRegScore in both replicates.

The same analysis was performed with a data set reporting on the differences between glycerol and glucose based growth (GSE6302, [100]). Visual inspection of the expression differences in the 72 cliques show that this response is producing much stronger expression changes than the response to pheromone and many more cliques appear systematically affected (Figure 4.1.2 D). Similar as before, the identification of influenced cliques is possible based on their expression

differences, the ranks of the genes in upregulation lists (UpRegScore) and the enrichment of the Top 200 genes within the cliques (most prominent results in Table 1.3). Given that the response involves many more genes, the calculation of average expression differences for each clique appears very rewarding in addition to the Top 200-enrichment. The cliques DCS1-MSC1 (+1.79, (-log10(p) =+139)) and QCR7-COX6 (+1.339, (-log10(p) =+27)) and CAT2-IDP2 (+1.133, (-log10(p) =+20)) give the most significant upregulation and the cliques CUP5-TEF1 (-0.289, (-log10(p) =+18)), RPS24A-RPL20A (-1.12, (-log10(p) =+18)) and RPB2-FUN12 (-0.32, (-log10(p) =+16)) give the most significant downregulation response (Figure 4.1.2 E). Also the UpRegScore yields highly significant p-values for each of those cliques (all (-log10(p) >+12)). Meanwhile the reproducibly of these cliques was tested in other biological replicates of this sample condition. The two replicates strongly correlate in plots where the average expression differences (Figure 4.1.2 E) or the UpRegScore (Figure 4.1.2 F) of each clique are directly compared.

Clique Name	Clique Size	Genome Size	$Up_{Selected}$ Replicate 1	$Down_{Selected}$ Replicate 1	Differ. Exp Replicate 1	Differ. Exp Replicate 2	UpRegScore Replicate 1	UpRegScore Replicate 2
DCS2-MSC1	93	4591	0	0.18	1.80	0.93	70	50
QCR7-COX6	33	4591	2.21E-148		1.34	0.82	81	52
KGD2-KGD1	41	4591	0		1.21	1.21	70	64
CAT2-IDP2	70	4591	0	0.15	1.13	0.69	67	47
ARN1-SIT1	19	4591	1.29E-07		0.99	0.98	59	56
TPS1-RIM11	82	4591	4.50E-10		0.70	0.64	52	46
YLR345W-FBP26	61	4591	3.40E-17		0.68	0.56	48	42
YBR225W-RIM101	317	4591	0.48	0.19	0.35	0.27	33	27
CUP5-TEF1	651	4591	0.16	9.87E-256	-0.29	-0.24	-28	-19
SSS1-VMA10	185	4591		0.66	-0.30	-0.28	-28	-24
RPB2-FUN12	476	4591	0.22	3.13E-12	-0.32	-0.29	-30	-27
RPF2-BRX1	236	4591		0.0045	-0.44	-0.70	-44	-56
RPS24A-RPL20A	53	4591		0	-1.12	-1.28	-81	-78

Table 1.3. Most prominent hits from differential expression in glucose versus glycerol based growth within the 72 expression cliques for replicate 1 and replicate 2. Results significantly above or below baseline ($p<10^{-4}$) are colored in the categories for Top 200-Enrichment of replicate 1, clique differential expression in both replicates or UpRegScore in both replicates.

The strong correlation between genes within one clique in these experiments shows that many cliques generated by the connection and separation method are indeed regulated as transcriptional units. This confirms that the evaluation of individual microarrays with this type of clustered analysis nicely reflects the transcriptional response. Also it confirms that this analysis approach can help to investigate strong and weak responses alike based on p-values on all derived parameters.

4.1.3 Full-genome analysis of the response to polyglutamine overexpression plasmids

The 72 cliques present a good way to analyze the genome-wide responses from expression data and they seem to work even in the treatment of single biological experiments, which is helpful in extracting relevant information from the microarray samples. These arrays report on the overexpression of polyglutamine proteins of different length, with one of the constructs (Q56) producing a slow growth phenotype, while the other (Q30) does not induce growth defects.

A. B.

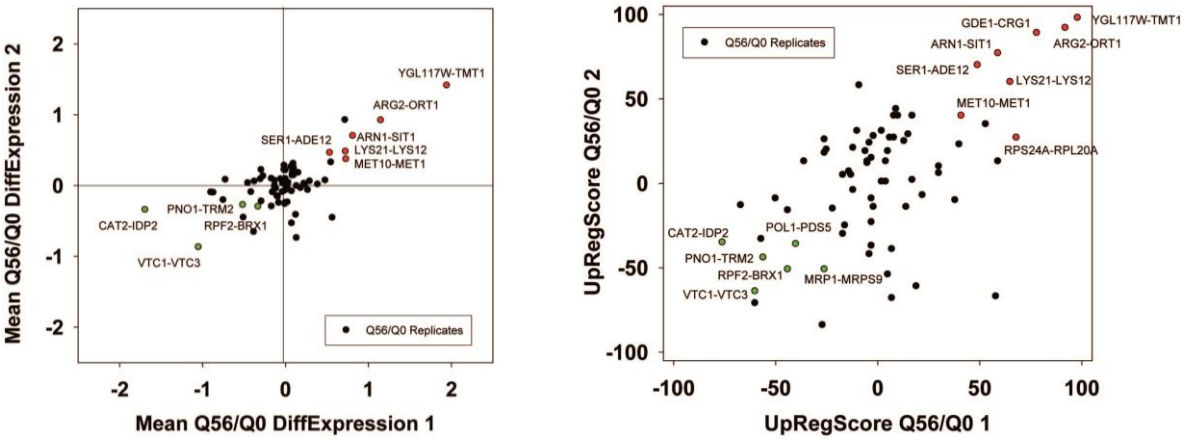


Figure 4.1.3 Analysis of the toxic effects of polyglutamine expression with pQ56. Correlation between two experiments of Q56-expression induced toxicity based on the average expression differences (A) or the UpRegScore (B) the clusters, which are significantly up- or downregulated in both experiments, are highlighted in the respective color.

Experiments had been performed independently and maintained on the plates for 2-4 days before analysis to compensate for the different growth rates [94]. Employing the three analysis methods Top 200 Enrichment, Differential Expression and UpRegScore on Q56/Q0, significant

up- or downregulations in the first experiment are observed (most prominent in Table 1.4). These also match the visual inspection of the response (Figure 4.1.3 A). In general these cliques represent the large network, which had been assigned to the response to the nutritional status before [94], but now these genes are forming separate cliques. The upregulated genes include two large cliques with only mild upregulation: CUP5-TEF1 (0.271, -log 10(p) =+30) and RPB2-FUN12 (0.186, -log 10(p) =+20), which achieve significance based on their large size (775 genes and 524 genes) despite the small expression changes. There also are several smaller, but strongly affected cliques including YGL117W-TMT1 (1.94, -log 10(p) =+16), LYS21-LYS12 (0.73, -log 10(p) =+13), ARG2-ORT1 (1.15, -log 10(p) =+13), ARN1-SIT1 (0.814, -log 10(p) =+10) and MET10-MET1 (0.73, -log 10(p) =+10). Most genes included in these cliques were identified before, but here likewise large parts of the response could not be assigned to GO-term or transcription factor groups. This is possible now that the hits are embedded into the context of their cliques.

Clique Name	Clique Size	Genome Size	$Up_{Selected}$ Experiment 1	$Down_{Selected}$ Experiment 1	Differ. Exp 1	Differ. Exp 2	UpReg 1	UpReg 2	Aver Exp	Aver UpReg
YGL117W-TMT1	7	5811	1.40E-243		1.95	1.41	98	98	1.68	98
ARG2-ORT1	9	5811	2.51E-43		1.15	0.92	92	92	1.04	92
GDE1-CRG1	7	5811	0.016		0.72	0.92	78	89	0.82	83.5
ARN1-SIT1	21	5811	1.93E-225		0.81	0.70	59	77	0.76	68
LYS21-LYS12	28	5811	6.76E-47		0.73	0.48	65	60	0.60	62.5
MET10-MET1	30	5811	4.47E-36		0.73	0.37	41	40	0.55	40.5
SER1-ADE12	19	5811	5.74E-11		0.54	0.46	49	70	0.50	59.5
HAL5-AYT1	24	5811	4.34E-12		0.55	0.33	53	35	0.44	44
RPS24A-RPL20A	56	5811			0.48	0.07	68	27	0.28	47.5
MRP1-MRPS9	65	5811			-0.17	-0.30	-26	-51	-0.23	-38.5
SPT10-PRP18	77	5811	0.53	0.38	-0.41	-0.09	-44	-16	-0.25	-30
POL1-PDS5	35	5811	0.63		-0.29	-0.22	-40	-36	-0.26	-38
RPF2-BRX1	263	5811	0.10	0.13	-0.33	-0.30	-44	-51	-0.31	-47.5
PNO1-TRM2	52	5811		0.21	-0.51	-0.28	-56	-44	-0.39	-50
KGD2-KGD1	48	5811	0.088	5.35E-57	-0.75	-0.21	-57	-33	-0.48	-45
1770541_at-CGR1	9	5811			-0.50	-0.45	-60	-71	-0.48	-65.5
QCR7-COX6	53	5811		4.62E-30	-0.87	-0.10	-67	-13	-0.48	-40
DMA2-THI20	21	5811	0.15	1.28E-18	-0.90	-0.097	-50	-9	-0.50	-29.5
ZPS1-ZRT1	7	5811	0.0008	2.62E-11	-0.38	-0.66	-27	-84	-0.52	-55.5
VTC1-VTC3	16	5811		4.68E-25	-1.05	-0.87	-60	-64	-0.96	-62
CAT2-IDP2	89	5811	0.17	0	-1.69	-0.34	-76	-35	-1.02	-55.5

Table 1.4. Most prominent hits from differential expression between Q56 and Q0 within the 72 expression cliques. Results significant above or below baseline ($p<10^{-4}$) are colored in the categories for Top 200-Enrichment of experiment 1, clique differential expression in both experiments or UpRegScore in both experiments.

When comparing the two experiments incubated on agar plates for different times, a correlation for Q56 is still clearly visible, where the average expression differences of the cliques produce a roughly straight line, hinting to consistent differences at least in the strongly affected cliques. So for the intoxicated sample, obviously the recorded response is consistent independently of the incubation time and sample condition.

For Q30, where no toxicity is observed, the expression differences are much smaller. This lead to difficulties when generating networks from the Top 100-differentially expressed genes. Interestingly, while many cliques yield significant enrichments or expression shifts, only few of them show this behavior consistently in both experiments. Clearly, no obvious correlation between the two experiments is observable (Figure 4.1.4) and only few strongly altered cliques show consistent differential expression in both experiments. Nevertheless, each experiment shows its own significant difference between its Q30 and its Q0 control sample.

A. B.

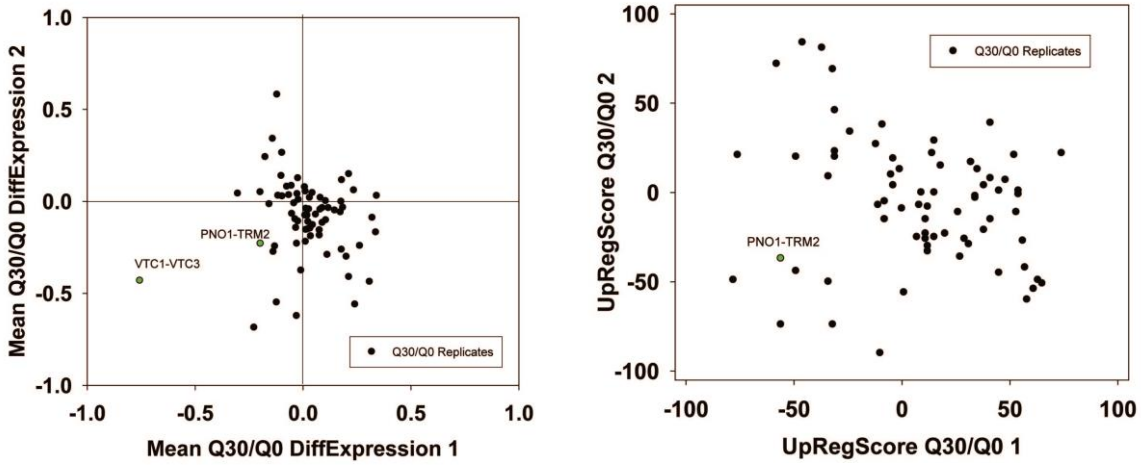


Figure 4.1.4. Analysis of non-toxic effects of polyglutamine expression with pQ30. (A) Correlation based on the expression differences and (B) the UpRegScore with the cliques being labelled, if significantly shifted in the same direction in both experiments.

Concerning that due to the small influence of Q30 versus Q0 even small differences in growth conditions on the plates (e.g. different colony density) are masking the specific response. Apparently this effect can be stronger than the influence of the Q30-construct itself. Nevertheless, both experiments – based on the highly significant and visually observable shifts in some of their cliques - provide accurate information on the differences between the Q30 and Q0 samples in each experiment. Combining the analyses of both experiments, only the VTC1-VTC3 clique, the ZPS1-ZRT1 clique and the PNO1-TRM2 clique remain as candidates for a consistent influence from the overexpression of polyglutamine Q30. Interestingly, the PNO1-TRM2 clique is even downregulated in Q56 and Q30 samples alike. This wasn't observed in the earlier analysis, possibly due to PNO1-TRM2 not having enough highly affected genes to generate a cluster of its own in the previously used methods [94].

4.1.4 GSEA yields gene sets with similarities to the strongest affected cliques

To compare with other enrichment methods, the same microarray experiments were applied and evaluated with the GSEA software. The cliques acquired from above were added as additional gene sets in the gene set database to obtain the enrichment scores as well. For the response to α-pheromone several gene sets with high enrichment scores were obtained. Among them the GPA1-SPE4 is present as second ranked and the clique YJL052C-A-YOR268C is ranked as number 20. When analyzing the 20 Top scoring gene sets, eight of those sets are sharing at least 30% of their genes with GPA1-SPE4, implying that many of the identified gene sets are related to the clique GPA1-SPE4. The same procedure was employed to compare Glycerol and Glucose based growth for the heat-shock response and for the experiments comparing either Q56 to Q0 or Q30 to Q0. In all cases many of the best gene sets from GSEA-analysis contain the genes also enriched in the relevant cliques from the clique set. In the other direction also the Top scoring cliques are represented directly by the Top scoring gene sets from GSEA. In contrast the less strongly affected cliques from the clique set, which still show significant deviations from baseline, are usually not represented in the Top 20 cliques of the GSEA-analysis. This implies that the clique set derived here may have advantages in uncovering significant but weakly responding sets due to its limitation to 72 static sets representing the full genome.

In general with an elevated number of co-regulators in building the matrix, the number of involved hits and identified connections per hit increases as well, which is valid in both case for the Q30 and the Q56 data sets (Figure 4.1.2 A and B). In the meantime this is applicable for the random hit, which implies the numbers of co-regulators used in building the networks require to be selected reasonably. The influences of initial hit list number on the connectivity parameters are investigated by comparing the number of networks generated from the polyglutamine data

sets and the random hits with varied number of hits. From 30 to 300 hits are selected for the network construction and the corresponding number of involved hits and number of connections per hit were plotted in figure 4.1.3. With elevated number of hits the percentage of hits involved in the network increases for the random hits list. At a number of 300 randomly selected genes, which covers 5% of the entire yeast genome, the possibility of co-regulated gene partner being found in the cluster is getting bigger.

4.2 HSP-90 down regulation has an impact on the heat shock response, innate immune response and intestinal development of the nematodes.*

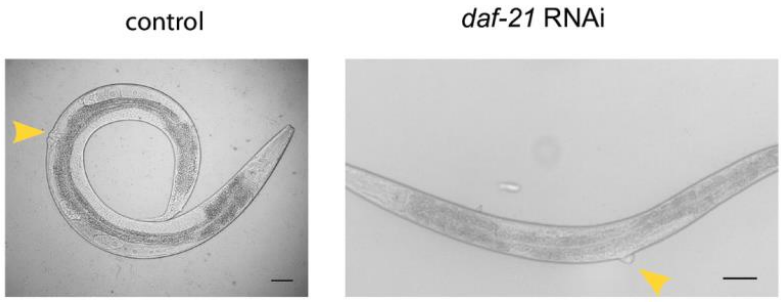
*this chapter has been published in PLoS One. 2017. PMID: 29078207

4.2.1 HSP-90-downregulation induces chaperons' accumulation and depletion of certain lectins.

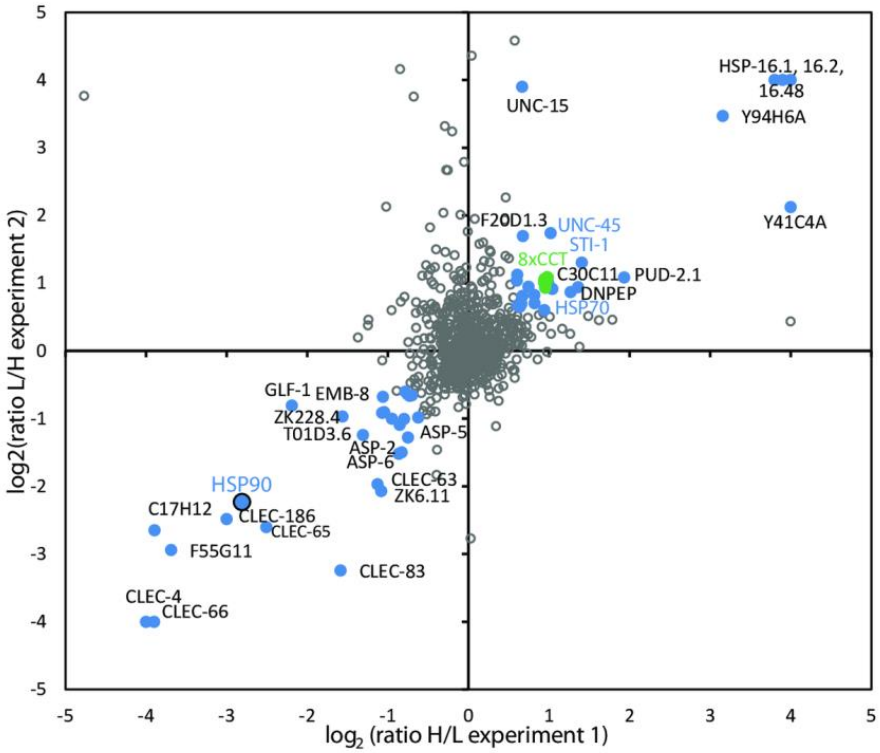
The proteomic response of the nematodes upon HSP-90 depletion is investigated via mass spectrometric comparison of HSP-90 depleted and control nematodes. This was achieved based on RNAi experiments against HSP-90 with 15N-labeled nitrogen source for the depleted nematodes and normal 14N nitrogen source for the control growth. Arrested nematodes showed a slower development and sterility compared to the control nematodes at the L4 stage (moment of harvest, Figure 4.2.1 A). The nematodes were lysed and equal amount of isotope-specific samples were combined before analyzed via the mass spectrometry system, which offered a direct comparison of the identified peptides within a single experiment. To avoid the potentially impact of nitrogen source on the results, the experiment was repeated with converted settings (14N nitrogen source is used for HSP-90 depletion and 15N for the control).

In the data analysis these two experiments were combined and about 400 proteins were quantified. Comparing the relative expression levels showed that after HSP-90 depletion, many known heat-shock proteins and the associating chaperone systems were up-regulated (Figure 4.2.1 B), which indicates that the induction of heat-shock response is significant. Among these proteins, the small heat-shock proteins HSP-16.2, HSP-16.1 and HSP-16.48 were highly overexpressed with an increase around 32 fold in expression level. Cochaperons as STI-1 and UNC-45 and protein folding machinery CCT/TriC showed an increase of 2 fold (Figure 4.2.1 B). Other enriched proteins include UNC-15, Y94H6A.10, Y41C4A.11 and F20D1.3, whose functions are not known to date. Interestingly, with the information from a transcriptional database most of these proteins can be clustered into a network (Figure 4.2.1 C left), which implies that these proteins are regulated in a similar way. The GO-term enrichment of these proteins gives enrichment in "protein folding" and "unfolded protein binding", which fits nicely to the nature of these proteins.

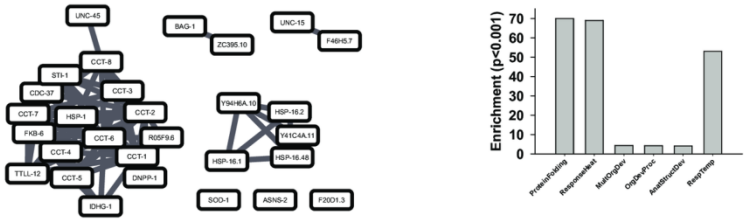
A.



B.



C.



D.

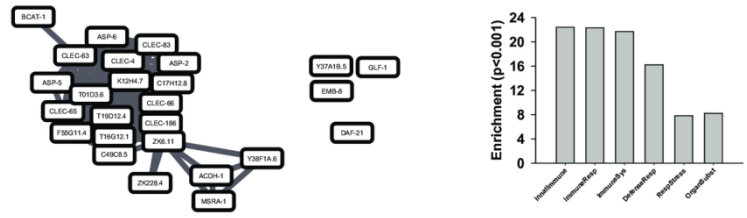


Figure 4.2.1 Influence of HSP-90-depletion on the *C. elegans* proteome. (A) Control and HSP-90 RNAi treated N2 nematodes at the moment of harvesting for analysis. The developing phenotype can be seen at the vulva site (yellow arrow) and the HSP-90-depleted nematodes are slightly smaller at that stage. (B) Proteomic changes based on isotope-labelled mass

spectrometry. The two independent experiments are plotted in the two dimensions. Blue spots indicate genes, for which the identification scores based on the number of detected peptides imply a reliable quantification. (C) Network of upregulated proteins and corresponding GO enrichment analysis. The network was generated based on hit-to-hit relationships in a coexpression database as described in the Materials and Methods section. D) Network of downregulated proteins and corresponding GO enrichment analysis. The network was constructed with identical settings as in C.

On the other side the knock-down experiment is confirmed to be successful with log2 of the expression level of DAF-21/HSP-90 being -2.8. Besides that, the expression level of several proteins was found to be reduced, including the lectin-like genes (such as CLEC-4, CLEC-65, CLEC-66 and CLEC-83) and aspartate proteases (such as ASP-2, ASP-5 and ASP-6), which can be clustered based on their coexpression and the GO-enrichment analysis indicates that these proteins are coregulated as part of the innate immune response. Other unassigned protein that are down-regulated upon HSP-90 depletion such as K12H4.7, T19D12.4, F55G11.4, C49C8.5, ZK6.11 and ZK228.4 can be associated with innate immune response according to Wormbase [101]. Taken together the response of *C.elegans* to HSP-90 depletion on proteomic level mainly involves proteins related to the heat-shock response and the innate immune response.

4.2.2 Up-regulations in gene expression referring to developmental processes

To further define the networks from the proteomic response, the full transcriptional responses were investigated under similar conditions by Affymetrix DNA microarrays. The experiments were carried out in biological triplicates and averaged in analysis. The top 250 genes which were most strongly up or down-regulated from the 26,959 sampled transcripts were clustered with the help of public coexpression database to calculate to transcriptional networks.

The significance of the obtained networks was qualified by comparing the connectivity with networks generated from 100 random genes. Here the parameters describing connectivity for both up and down-regulated clusters were increased (Figure 4.2.2 A) and the number of connections per gene for these clusters are 748 and 952 respectively, which is much higher than that of the random clusters that have 12 and 16 connections per gene ($p<0.001$, Figure 4.2.2 B).

Knowing the correspondence between the response on transcription and co-expression level upon HSP-90 depletion, the completeness of the transcriptional response is tested by checking whether up or down-regulated genes aside of the top 250 genes can be successfully predicted based on the coexpression database. This is qualifies by a score ranged from -100, which

indicates the least possibility for a gene to be found as top 250-300, to +100 which indicates the evaluated gene is most likely to be found as between the top 250-300. The predicted score for both up and down regulated clusters are over 90 while for the random gene cluster it's close to 0 (Figure 4.2.2 C). In this way the accuracy of the prediction for the transcriptional database is confirmed based on the coexpression datasets, which is valid in the averaged hit list and for each single biological replicate.

The connections of the hits were visualized in figure 4.2.3. The up-regulated genes upon HSP-90 depletion can be arranged into two clusters, which are labeled as Daf-21up_1 and Daf-21up_2 in the following text. The cluster Daf-21up_1 mainly contains genes of the heat-shock response, including hsp-16.2 and hsp-70. Besides this, GO-term enrichment indicates that genes related in the innate immune response exist in the Daf-21up_1 cluster as well. The strongest up-regulated genes (Table 2.1), such as genes arf-1.1 (60-fold), Y41C4A.11 (50-fold), Y47H10A.5 (32-fold) and C50F7.5 (8-fold) are found in this cluster. The cluster Daf-21up_2 is mainly involved with the slow development of the nematodes, where based on GO-term enrichment genes related to molding cycle, cuticle development, and also to the development of individual tissues like pharynx and intestine were found (figure 4.2.3 A).

To qualify these data, the regulation scores were calculated for both clusters in the three replicates. In all cases, the innate immune/heat-shock response cluster Daf-21up_1 is strongly upregulated with a score of 86, 98 and 100 for each individual replicate. The upregulation of Daf-21up_2 is slightly weaker in the first experiment with a regulation score of 70 while the other two replicate score for 97 and 97 respectively. As control the corresponding score for random gene sets is around 4 ($p<0.001$, Figure 4.2.4 A). Even though the three biological replicates differs from each other due to variable RNAi response or slightly difference of the developmental state, the clusters generated after the transcriptional response analysis still showed similar response so as the identified top hits in all three replicates (Suppl. Figure 2.1-2.3).

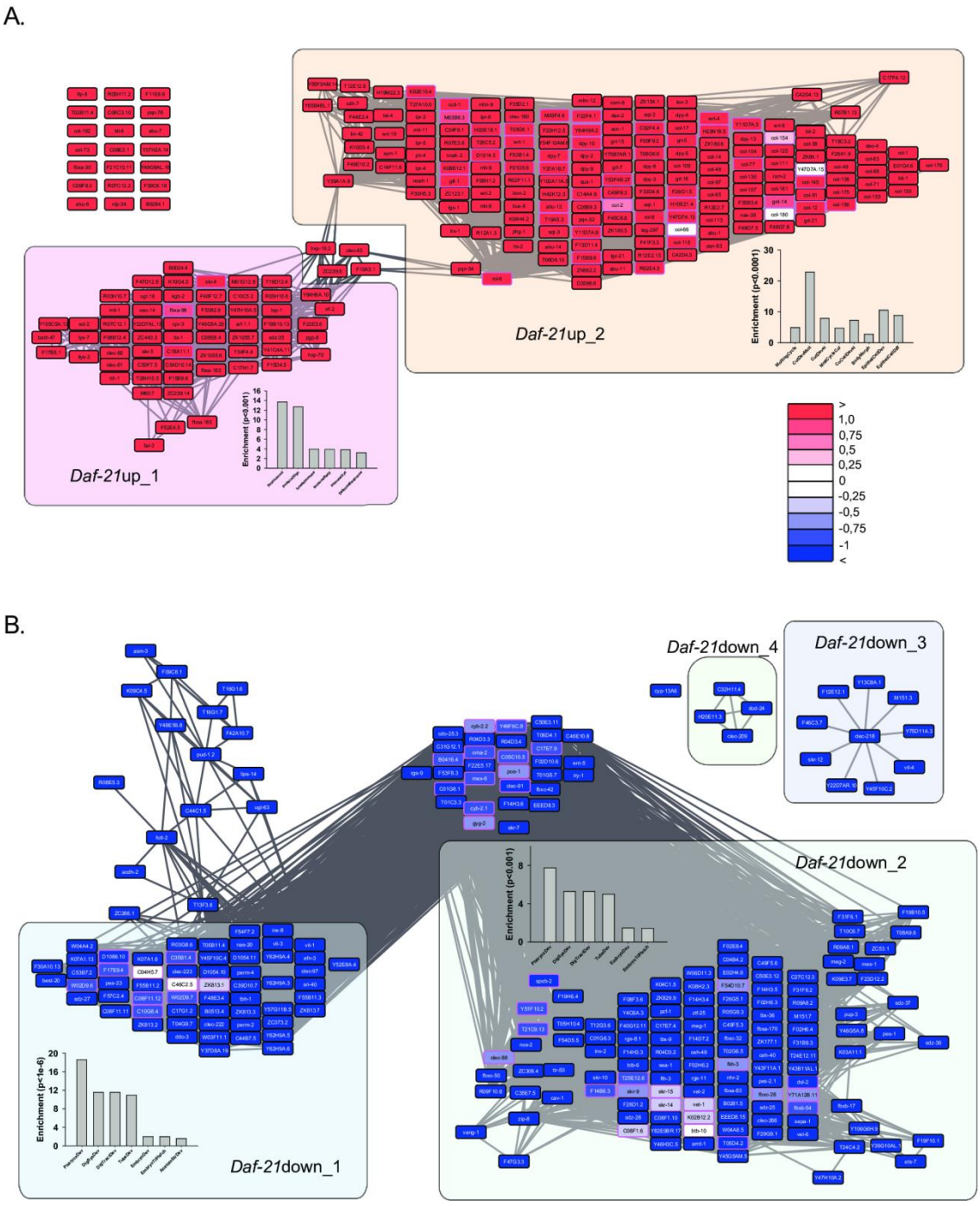


Figure 4.2.2 Transcriptional networks affected by RNAi against *daf-21*. Networks for upregulated (A) and downregulated (B) genes after HSP-90-depletion. The networks were constructed as described in the Materials and Methods section. The color code uses four shadings of blue and red to indicate the expression differences with darkest blue being $\log_2(\text{DiffExp}) < -1$, lightest blue being $\log_2(\text{DiffExp}) < -0.25$, darkest red being $\log_2(\text{DiffExp}) > 1$

and lightest red being $\log_2$(DiffExp) > 0.25. In between 0.25 and -0.25 the nodes are white as indicated in the legend. Large clusters have been subjected to GO-term enrichment analysis and the results are depicted adjacent to the respective cluster in the network figure.

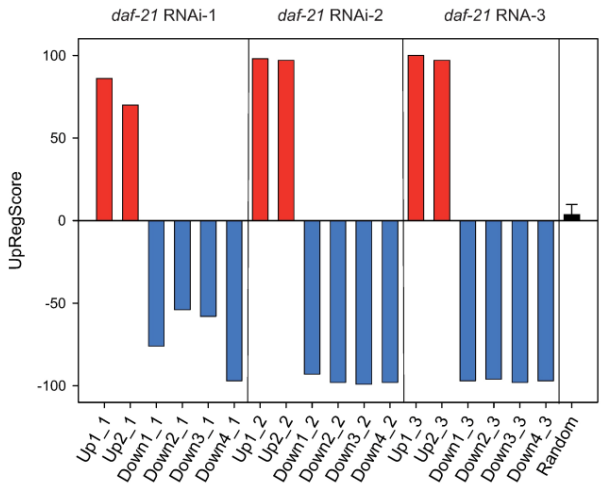
4.2.3 Downregulated genes are involved in germline development and innate immune response.

Upon visualization of the down-regulated genes after HSP-90 depletion, they form two larger clusters (Daf 21down_1 and Daf 21down_2) based on the coexpression relationships, which are closely connected and other two smaller clusters (Daf 21down_3 and Daf 21down_4) which are independently regulated (Figure 4.2.3 B). The gene cyp-13A6 is the only gene which is isolated from the networks and doesn't have any connections identified with other genes. The qualification of these results showed high similarities among the three replicates even though the first replicate showed a weaker response. The Regscores for the clusters are between 54 and 59 while for the random hits its only 3.7 ± 6 ($p<0.001$).

Based on Wormbase information, many genes in the Daf 21down_1 and Daf 21down_2 clusters are related to the development of different tissues from pharynx and epithelial cells to embryo development, which is validated by GO-enrichment analysis and single-gene queries. For example, genes associated with processes in the germ line including C39D10.7, Y62H9A.4, Y62H9A.6, ZC373.2 and C17G1.2; genes found related to the spermatheca contains clec-223, ZK813.3 and F55B11.3, and to the early embryo prior to gastrulation like pes-2.1, T04G9.7, vet-6 and F55B11.2. So the main response represented by the Daf 21down_1 and Daf 21down_2 clusters reflects a slower developmental progress and a stalled development of its gonads and embryos of the HSP-90 depleted nematodes.

Interestingly, in the first replicate the response is shifted towards the innate immune response and some highly inducible genes are found in the Daf-21down_4 cluster. As examples here, the gene dod-24 is identified as regulator of DAF-16 and is related to the innate immune response [102] and response to stress conditions [103]. Other genes connected with dod-24 like H20E11.3 clec-209 and C32H11.4 are found to be related to the innate immune response as well [104, 105]. Taken together, in the down-regulated clusters the genes involved in the innate immune response can be separated from other connected genes in the coexpression data.

A.



B.

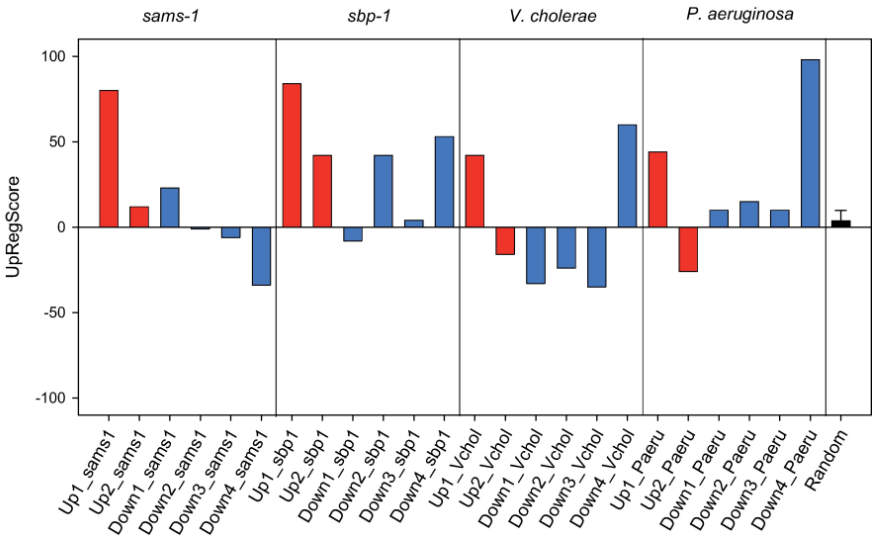


Figure 4.2.3 Regulation of separated clusters during *daf*-21 RNAi-response and other conditions. (A) Comparison between the three replicates regarding the up- or downregulation of specific clusters. The included genes are used to determine the UpRegScore of the respective cluster in each of the experiments as described in Materials and Methods. (B) The score was calculated as described in the Materials and Methods section. Calculation of the UpRegScore for each gene cluster in the microarray experiments for RNAi against *sams-1* or *sbp-1* (left side), *V. cholerae* VC109/VC110 exposure and *P. aeruginosa* exposure (right side). The corresponding results on the single-gene basis are shown in S 5-8 Fig.

Upregulated Gen	Transgene Strain	Average	STD	*Daf-21*Up
arf-1.1	-	6.01837	0.978103	1
Y41C4A.11	-	5.68503	0.463387	1
Y47H10A.5	-	4.5152	1.60171	1
C50F7.5	(SD1583)	4.19376	1.28643	1
F15B9.6	-	3.95603	1.10137	1
skr-5	SHK207	3.8018	1.12279	1
Y37H2A.14	-	3.76196	0.726931	0
sdz-35	-	3.59986	0.951894	1
clec-60	AU185	3.53321	1.9834	1
T26H5.9	-	3.51657	1.14567	-
fbxa-163	-	3.45604	1.59722	1
R03H10.6	-	3.35083	0.623473	1
K08D10.10	-	3.30692	0.91793	-
ZK355.8	-	3.23513	2.23024	-
clec-61	-	3.19526	1.54801	1
F22E5.6	-	3.08391	0.635536	1
B0563.9	-	3.07138	2.39947	-
Y68A4A.13	-	2.99054	2.03498	-
col-176	-	2.94957	2.66572	2
R07C12.1	-	2.93562	0.611104	1
C08E8.4	-	2.79468	0.755811	1
F29G9.14	-	2.78478	0.923283	-
C08E3.1	-	2.78009	1.5147	0
C17F4.12	-	2.75546	1.97348	2
dct-3	-	2.71538	0.490552	-
bli-1	(CH1445)	2.69926	2.15158	2
sqt-1	-	2.68336	2.36365	2
col-73	-	2.68196	1.31729	0
tts-1	-	2.62731	1.48224	1
clec-180	-	2.58966	2.50659	2
dao-2	-	2.57777	2.30527	2
pqn-63	-	2.56242	1.92395	2
dpy-13	-	2.55089	1.87921	2
Y105C5A.12	-	2.54054	1.38919	1
F13D11.4	-	2.53557	1.98456	2
col-38	-	2.53333	1.71205	2
B0024.4	-	2.53134	1.13112	1
ugt-18	-	2.49466	0.887937	1
tag-297	-	2.49185	1.78713	2
F33H12.7	-	2.48457	1.03716	-
F44E2.4	-	2.47906	2.21763	2
F48G7.8	-	2.47772	1.01412	2
col-138	-	2.46287	1.70462	2
lpr-3	-	2.46111	2.68673	2
mlt-11	-	2.43061	2.54662	2
B0284.1	(BC15640)	2.42306	0.804259	0
Y94H6A.10	-	2.41241	1.07078	1
R02F11.1	-	2.41111	2.02839	2
col-175	-	2.40976	1.29287	2
tsp-1	-	2.39753	1.1672	1
T28H10.3	(RW10819)	2.39438	1.05975	1
M60.7	-	2.39325	1.40075	1
F10A3.1	(BC13458)	2.38407	0.650951	1+2
rol-1	-	2.38016	1.67645	2
E01G4.6	-	2.37237	2.05367	2
acn-1	-	2.36373	1.87353	2
ptr-4	-	2.35218	2.04175	2
Y65B4BL.1	-	2.33269	2.28006	2
H42K12.3	-	2.331	2.50187	2
col-169	-	2.33062	3.05638	-
hsp-70	-	2.33044	1.01536	1
C54D10.14	-	2.32684	0.40607	1
col-49	-	2.31983	1.81372	2
B0462.5	-	2.31955	1.40587	-
ZK84.1	-	2.31004	2.09149	2

Downregulated Gen	Transgene Strain	Average	STD	*Daf-21*Down
tbh-1	(MT9971)	-2.56878	2.57384	1
Y57G11B.5	-	-2.49476	2.46702	1
clec-223	-	-2.4418	2.47944	1
F59D8.3	-	-2.41413	1.70888	-
ZK813.3	-	-2.3442	2.07643	1
F49E12.1	-	-2.33443	2.44208	-
clec-209	-	-2.2619	1.69682	4
W04A4.2	-	-2.25409	2.23416	1
cyp-13A6	-	-2.24297	3.3709	0
C39D10.7	-	-2.23089	2.65949	1
cav-1	RT688	-2.21036	1.0788	2
vit-4	-	-2.20557	2.55327	3
Y62H9A.6	-	-2.16989	2.44449	1
dod-24	AU10	-2.1679	1.88533	4
C25A8.4	-	-2.10898	2.4159	-
K10C2.8	-	-2.07235	1.45343	-
T04G9.7	-	-2.06123	2.48285	1
Y62H9A.4	-	-2.05896	2.06525	1
sepa-1	HZ455	-2.05658	1.49506	2
C35E7.5	-	-2.0464	1.21709	2
skr-7	-	-2.04313	1.10403	1+2
pes-2.1	-	-2.03482	1.36011	2
sri-40	-	-2.02298	1.54135	1
D1054.11	(SD1981)	-2.01782	2.16539	1
C44B7.5	-	-2.01272	2.04742	1
ZC373.2	-	-2.0072	2.26472	1
efn-3	-	-2.00689	1.79793	1
ets-7	-	-1.99975	1.08082	2
Y48E1B.8	-	-1.99174	1.2993	0+1
T27A10.8	-	-1.98758	1.45337	-
F48E3.4	-	-1.98397	2.74176	1
ZK813.2	-	-1.94439	1.6098	1
vet-6	-	-1.94433	1.06462	2
Y39G10AL.1	-	-1.94291	0.778773	2
sdz-25	-	-1.94233	1.38382	2
vit-3	-	-1.9359	2.34383	1
D1086.6	-	-1.93088	1.66361	-
F55B11.3	(SRS86)	-1.92928	2.06282	1
ZK813.7	-	-1.92374	1.98332	1
Y62H9A.5	-	-1.9195	2.61316	1
B0513.4	-	-1.88055	2.34603	1
K11D12.13	-	-1.87013	1.85672	-
vit-1	-	-1.8571	2.33645	1
nas-20	-	-1.8557	1.43733	1
K07A1.6	-	-1.84355	1.80729	1
F57C2.4	-	-1.83249	1.67229	1
D1054.10	-	-1.81174	1.8609	1
acdh-2	-	-1.79298	2.37892	0+1
F15E11.15	-	-1.77703	2.61117	-
C17G1.2	-	-1.77345	2.15736	1
ttr-42	-	-1.7662	1.31818	-
clec-266	-	-1.75624	1.65512	2
linc-121	-	-1.75054	1.21151	-
sdz-28	(RW10371)	-1.74133	0.828096	2
C04B4.2	(BC15374)	-1.73929	1.57356	2
skr-10	-	-1.73509	1.33333	2
D1086.11	-	-1.71749	1.85678	-
F54F7.2	-	-1.71409	1.33282	1
clec-218	-	-1.71321	0.870155	3
F02E8.4	-	-1.71192	1.65181	2
T12B5.15	-	-1.70195	0.941428	-
R04D3.3	-	-1.67453	1.71792	1+2
Y62H9A.3	-	-1.66364	1.89092	1
pud-1.1	MQD396	-1.65744	2.70827	-
cdc-25.3	(BC12401)	-1.65662	1.28574	1+2

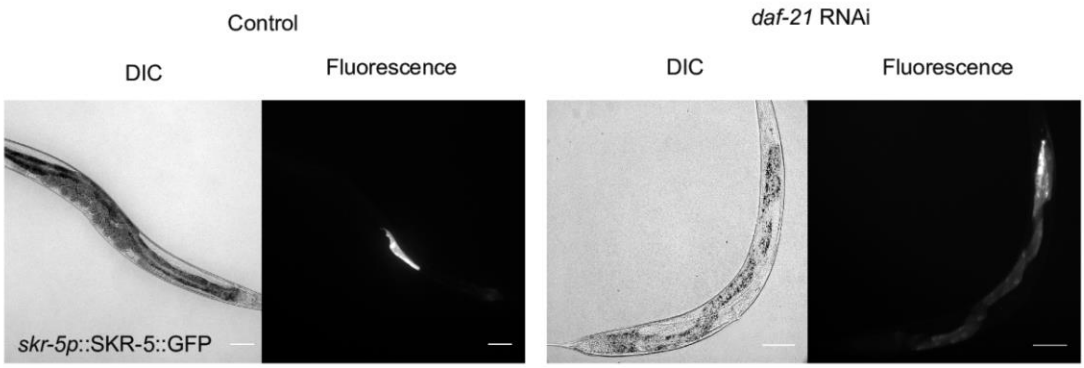
Y34F4.4	-	2.29565	1.02139	1	F55B11.2	-	-1.6509	1.4361	1
col-65	-	2.29107	1.22273	2	col-135	-	-1.63889	1.0186	-
scl-2	-	2.27978	1.3477	1	R03G8.6	-	-1.631	1.34914	1
abu-14	-	2.27971	1.84813	2	flh-3	-	-1.6308	1.60568	2
T01B7.13	-	2.27827	1.93391	-	F13G11.3	(SD1973)	-1.62962	1.36361	-
bli-2	-	2.27603	2.26035	2	ent-5	-	-1.62676	1.54476	1+2
ZC239.14	-	2.27466	1.25068	1	C49F5.6	-	-1.62265	1.44651	2

Table 2.1 Averaged expression differences of strongly affected genes after *daf-21* RNAi. The genes depicted are the most strongly affected genes. The three experiments were averaged to obtain the average expression change and the standard deviation. These genes are part of the network structures of Fig 2. The weaker hits are truncated in this table, but included in the networks in Fig 2. The assignment to the clusters is based on the cluster description in Fig 2. "0" implies that the gene is outside of all marked clusters, "+" connections imply that the gene is in-between marked clusters and "–'' implies that this particular gene was not found in the used coexpression database. In the column "strain" all currently available fluorescent reporter strains for these hits are shown, but strains in brackets were not used in the study.

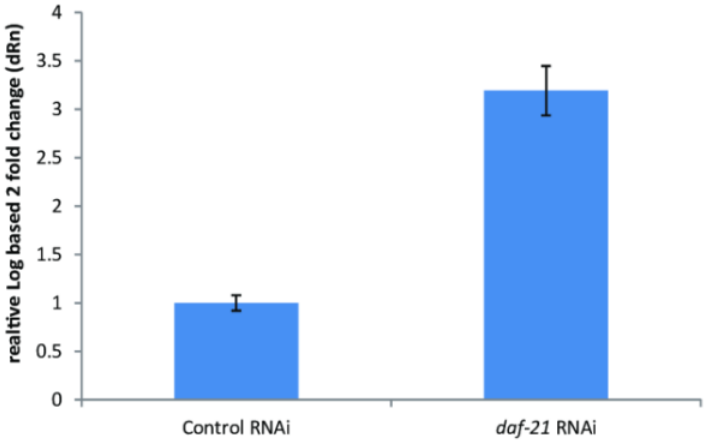
4.2.4 Expression changes after HSP-90-depletion overlap with described immune responses.

To further identify the responding pattern, the experiments with similar expression patterns were compared to the HSP-90 depleted coexpression response. Impressively the most prominent genes identified in this experiment have been equally reported to be most prominent in other microarray analysis on the innate immune response [106]. Following this the response upon depletion of SAM-1 and SBP-1 [106, 107] were compared to the HSP-90 depletion. The RNAi experiment datasets were obtained from publicly available database and the responding clusters were generated and plotted onto the HSP-90 depletion clusters, which showed an overlap mainly in the Daf-21up_1 cluster in both cases. The Regscores for each of the clusters were calculated (Figure 4.2.4 B, Suppl. Figure 2.4-2.5), which showed high similarity to the Daf-21up_1 cluster for these two experiments with Regscores around 80-85 ($p<0.001$) while no obvious overlap can be observed for other clusters. This implies that the Daf-21up_1 and Daf-21up_2 are regulated independently and the HSP-90 depletion induces certain response in a similar manner to those genes (SAMS-1 and SBP-1) connected to the normal regulation of the innate immune response.

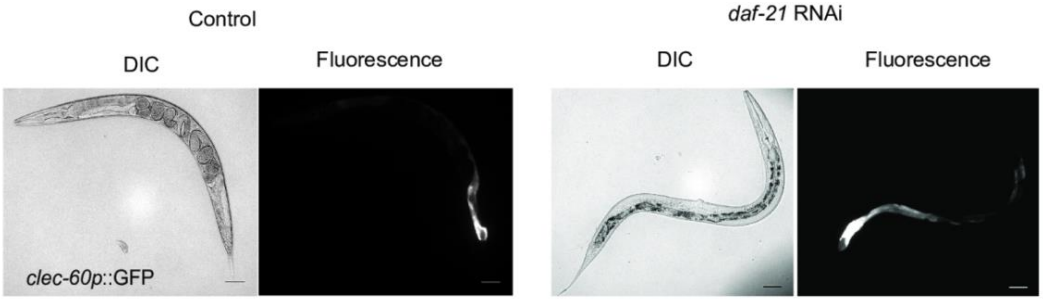
A.


Control
DIC
Fluorescence
daf-21 RNAi
DIC
Fluorescence
skr-5p::SKR-5::GFP

B.


realtive Log based 2 fold change (dRn)
4
3.5
3
2.5
2
1.5
1
0.5
0
Control RNAi
daf-21 RNAi

C.


Control
DIC
Fluorescence
daf-21 RNAi
DIC
Fluorescence
clec-60p::GFP

D.

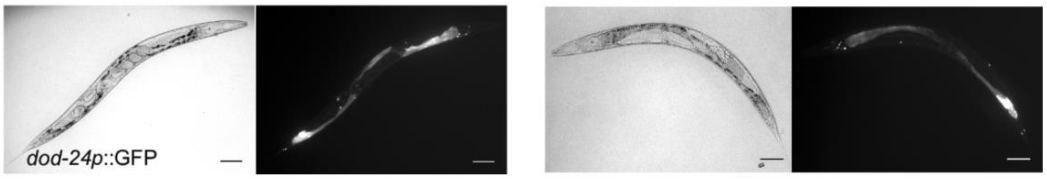

dod-24p::GFP

Figure 4.2.4 Immune response genes respond to *daf-21*-RNAi in intestinal tissues. DIC and fluorescence images of HSP-90-treated (right panels) and control nematodes (left panels) are shown. (A) SHK207 nematodes were imaged to visualize SKR-5::GFP protein. Control nematodes showed rare expression in Int5, Int6 and Int7 cells. HSP-90-depleted nematodes generally showed induction in the full intestine. (B) qRT-PCR has been performed for *skr-5* in control nematodes and in HSP-90-depleted nematodes as described in Materials and Methods. (C) AU185 nematodes were imaged to visualize *clec-60p*-driven GFP expression. Control nematodes express in Int8 and Int9 cells and rarely in more anterior intestinal cells (left panel). HSP-90-depleted nematodes show induction in most parts of the intestine (right panel). (D) AU10 nematodes were imaged to visualize expression of *dod-24p*-driven GFP. Control nematodes show induction in all parts of the intestine (left panel). Expression is substantially reduced in HSP-90-depleted nematodes (right panel). The scale bar represents 100 µm.

Whether the response presented by Daf-21up_1 cluster overlaps with the pathogen-induced immune response has been tested in the meantime for the bacteria Vibrio cholerae and Pseudomonas aeruginosa [104, 108] (Figure 4.2.4 B, Suppl. Figure 2.6-2.7). The Regscore of these experiments ranges in 41-44 with $p<0.001$ and the overlap in the Daf-21up_1 cluster is less than the SAMS-1 and SBP-1 induced responses. No significant similarities are observed for the other clusters. Taken together the HSP-90 depletion induces the innate immune response in an independent way to the impaired development of the nematodes.

4.2.5 HSP-90-depletion induced changes to the immune response are evident in intestinal tissues

Upon identification of genes differentially affected in the HSP-90 depletion, the main hits were traced with fluorescent reporter strains to shine light on the potential tissues/organs in which the HSP-90 depletion occur. The publicly available reporter strains containing the most prominent hits found in the immune response clusters were obtained, which contains a GFP-translational fusion. RNAi experiments against HSP-90 were performed to observe the changes in the florescence intensity and localization upon HSP-90-depletion.

For the Skp-1 homolog protein SKR-5, which is upregulated 6-30 fold in the microarray data, the expression was observed during later larval stages in posterior intestine as reported previously [109]. The HSP-90-depleted worms have much stronger fluorescence intensity than the control worms and the entire intestine is visible (Figure 4.2.5 A). In parallel the up-regulation is evident by the qRT-PCR assays with a threefold increase (Figure 4.2.5 B). Another example is the C-type-

lectin clec-60, which is known to be regulated as part of the innate immune response. In the control RNAi worms, it's expressed in the larval intestine and the Int8 and the Int9 cells [110, 111]. Upon HSP-90 depletion, the florescence intensity has increased strongly (Figure 4.2.5 C). The third tested reporter strain contains DOD-24, which is an epoxide hydrolase related to the innate immune response [103]. For the DOD-24, the expression level is quiet high throughout the entire intestine under normal growth conditions (Figure 4.2.5 D, left). Upon HSP-90 depletion, the expression in the Int3 to Int7 intestinal cells is strongly reduced and only in the posterior part of the intestine the expression level is conserved (Figure 4.2.5 D, right).

Taken together, all of the most prominent reporter strains showed clear changes in the expression level as part of the immune response. The fluorescent level correlates to the expression analysis, which further confirms the influence of the HSP-90 depletion on the nematode's innate immune response.

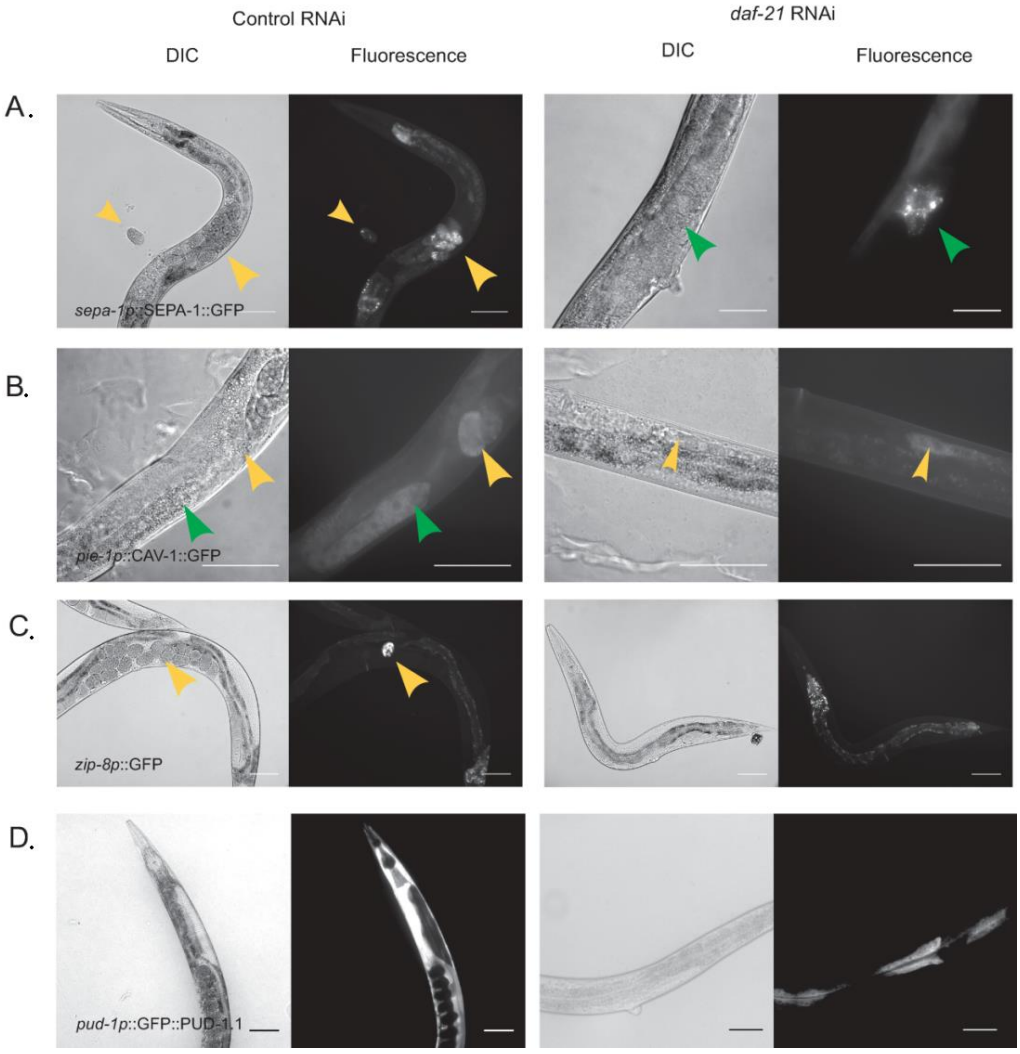


Figure 4.2.5 Genes from cluster *daf-21* Down_2 are suppressed by HSP-90 RNAi in gonads and embryos. (A) HZ455 nematodes were imaged to visualize the expression and subcellular localization of SEPA-1 protein. Yellow arrows indicate the position of developed embryos, while green arrows indicate the fluorescence prior to passage through the spermathecum. (B) RT688 nematodes were imaged to visualize subcellular localization of CAV-1 protein. The CAV-1 protein in the meiotic zone of the gonad arm is indicated in both panels with a green arrow, while yellow arrows indicate the position of developed embryos. (C) BC12422 nematodes were imaged to visualize the cells expressing *zip-8*. The scale bar represents 100 µm. Yellow arrows show the position of a *bzip-8*::GFP expressing embryo. (D) GFP::PUD-1.1 can be observed in intestinal tissues in later larval stages. After HSP-90 RNAi this expression is strongly diminished. The scale bar represents 100 µm.

4.2.6 Genes from developmental clusters are affected in the nematode's gonads and embryos.

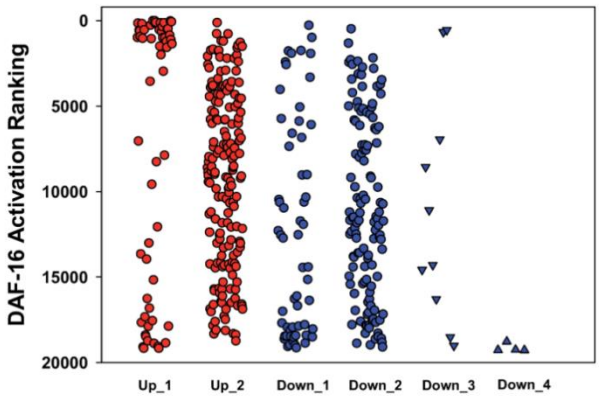
For the cluster of the early developmental genes, reporter strains were similarly tested to confirm the changes brought by the HSP-90 depletion. The SPEA-1 is a protein which mediates autophagy of the P-granules in the early embryos and is degraded once the embryo reaches 100 cells [112, 113]. Fluorescent reporter of SPEA-1 was observed to localize in the gonad arms and in the embryos next to the spermathecum. HSP-90 depletion induces a strong reduction of the fluorescent signal and defect in gonad arms. Beside this the embryos next to the spermathecum doesn't show any fluorescent signal which likely indicates that defects are shown in the fertilized oocytes as well.

The translational reporter for the caveolin homolog CAV-1 [114] tested in RNAi is clustered in the Daf-21 down-2, which is required during the early cell division stage and will vanish soon in the embryo afterwards. Upon HSP-90 depletion the oocytes approaching the spermatheca showed a reduced fluorescent signal (Figure 4.2.6 B). Another hit in the same cluster is the bZIP-like transcriptional factor ZIP-8 [115], which is activated in the early embryos and can be observed in the intestine of adults. Upon HSP-90 depletion the fluorescent signal in the oocytes was disappeared (Figure 4.2.6 C). Taken together the reduction of these genes in expression indicates that the development of functional oocytes is regulated by the HSP-90.

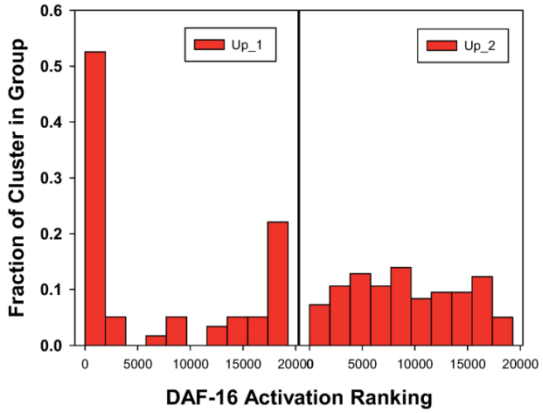
In contrast, reporter strain of the gene pud-1.1, which is only loosely connected to the down-regulated clusters instead of located in the center of the cluster, is compared in the RNAi experiments. Function of the PUD-1.1 is yet remaining as unknown while it's found to be regulated during aging and development of the nematode [116, 117]. Fluorescence of the PUD-1.1 was observed in the intestine at a late larval stage, which is reduced in the HSP-90-depleted

worms (Figure 4.2.6 D). With these RNAi experiments of the reporter strains it's confirmed that the depletion of HSP-90 has an impact on the development of the oocytes and intestinal tissues, where heat shock response and innate immune response are triggered.

A.



B.



C.

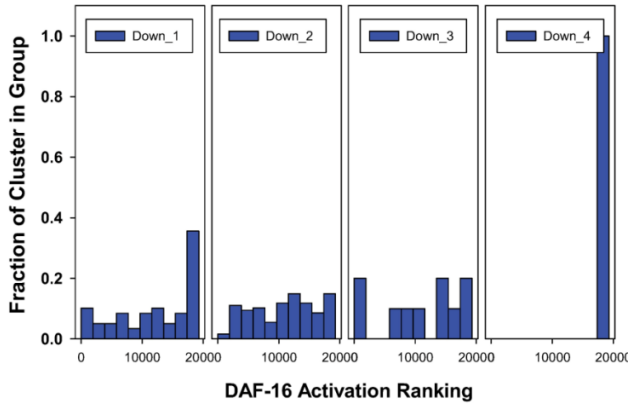


Figure 4.2.6 DAF-16 targets and their distribution in HSP-90 RNAi induced expression clusters. All the genes distributed into the clusters of Figure 4.2.2 are evaluated, whether they are ranked DAF-16 Class I (upregulated) or DAF-16 Class II (downregulated) targets. To this end the ranking was employed from Tepper *et al.* [118]. (A) For each cluster the position of the genes is marked by a point at the respective ranking position. The top 1663 genes represent targets considered upregulated by DAF-16; the bottom 1733 genes represent targets downregulated by DAF-16 according to this ranking. (B) Histograms for the two clusters *daf-21*Up_1 and *daf-21*Up_2, where the bar on the left side roughly represents the 1663 Class I targets and the bar on the right side the 1733 Class II targets. (C) Histograms for the clusters downregulated after HSP-90 RNAi.

4.3 Binding of the HSF-1 DNA-binding domain to multimeric *C.elegans* consensus HSEs is governed by cooperativity.

HSP-90 is expressed under normal condition as well as under stressed conditions. The expression level is under tight regulation of transcriptional factors, which are responsible for the heat shock response, mostly the HSF-1. Upon activation, HSF-1 is translocated into the nucleus and binds to the promotor region of heat shock proteins, where the main interaction takes place on the Hsf-1 binding region of DNA known as HSE (sequence = nGAAn). Furthermore HSEs are often found in the promotor region of genes which are not closely related to heat shock proteins. Hence the mechanism of HSF-1 specifically recognizing its target promotor under stress condition remains elusive. By comparing the binding behavior of HSF-1 to HSE containing promotors originated from the Hsf-1 regulated and non-regulated genes, this chapter reveals factors influencing the HSF-1 DNA binding procedure and thus brings a new aspect on how the expression of SHR related genes are regulated.

4.3.1 Distribution of nematode HSE is not limited in the heat shock response related genes in coexpression cliques

Using the same method for the yeast genome response clustering, the *C.elegans* genes related to the heat shock response were identified and clustered, which includes typical HSR genes *hsp-16.1*, *hsp-16.2*, *hsp-16.48*, *hsp-16.41*, *hsp-70*, *F44E5.4* and *F44E5.5.* These results are reproducibly upregulated significantly in three microarray analysis among the other genes distributed in 307 coexpression clusters. Other weaker upregulated genes including *unc-23* and *lact-4* were not found belonging to the HSR genes set.

Previous report had highlighted a large number of HSEs in nematode genomes [119]. To this end 1000 bp promoter regions upstream of all *C.elegans* genes are tested. Around 4000 HSEs were identified containing a consensus sequence for HSF-1 in their promoter region. These HSE containing genes enrich mostly in a cluster which is known as the heat shock response cluster, while the rest distribute stochastically over the other cliques. Surprisingly, also many genes related to the heat-shock response, like *dnj-12*, *dnj-13*, *bag-1*, *hsp-110* and *hsp-1* contain HSE-sequences in their promoters, which apparently are not inducible by HSF-1. Other genes containing HSEs enrich in different cliques, which obviously are not upregulated under heat-stress conditions. Given that the genes with these HSE in their promotor region apparently are not induced on the same level upon heat-shock, they were classified as non-inducible, normal inducible and strongly inducible HSE-regions controlled by HSF-1. This extent of regulation resulting from HSF-1's actions is well beyond the induction of stress genes under stress

conditions and reaches far into the normal growth cycle of the nematode under non-stressed conditions.

4.3.2 Chaperone-gene derived HSEs show different affinity to HSF-1 DBD

Aimed at understanding the differential expression from these HSEs, investigation of the binding affinity of HSF-1 DBD to several chaperone-genes derived HSE-containing probes reveals whether the binding correlates with the observed induction strength upon heat-shock. To this end probes from the heat-shock responsive cluster, e.g. *hsp-70* (only one trimeric HSE-site), *hsp-16.2* (high consensus score tetrameric site plus additional trimeric site) and *F44E5.4* (highest consensus score pentameric site) as well as probes with identical length of the non-induced heat-shock related proteins *dnj-13* (tetrameric HSE-site) and *dnj-12* (trimeric HSE-site), *hsp-1* and *bag-1* were generated. Only *F44E5.4*p contains further HSEs in its sequence. These sequences are directly taken from respective promoter regions which locate usually within 250 bps upstream of the transcription starting point (Table 3.1). For all genes, the probes with the highest consensus score were synthesized. The probes for *F44E4.5*p also represent *F44E5.5*p and those for *hsp-16.2*p represent *hsp-16.11*p, *hsp-16.41*p, *hsp-16.48*p. These genes have identical sequences in their promoter regions at similar positions.

Gene	forward sequence	Trimeric consensus scores	Orientation	Start Position relative to ATG
Bag-1	cttctcgtcatttctagaagtttcttaattttcta	12.42	Rev	-118
DNJ-12	aaaagtgtcgagaatgttcacgaaaaatcgttaga	9.66	Fwd	-221
DNJ-13	agtaaatagaacgctctggaaagttccgcactctt	13.01, 9.1	Fwd, Rev	-205
F44E5.4	gcagtggaatattccagaagcttctagaagaagtt	11.59, 12.43, 13.07	Fwd, Rev, Fwd	-124 (also *F44E5.5*)
HSP-1	tgacgaaattcgaatttctagaatcccgccacgc	10.24	Fwd	-150
Hsp-16.2a	gccttacagaatgttctagaaggtcctagatgcat	9.41, 12.81	Rev, Fwd	-117 (*hsp-16.49* and *hsp-16.41*) -253 (*hsp-16.11* and *hsp-16.2*)
Hsp-16.2b	acaagcagctcgaatgttctagaaaaaggtggaaa	11.86	Fwd	-104
Hsp-70	agtaaattgtagaaggttctagaagatgccagagg	12.71	Fwd	-112

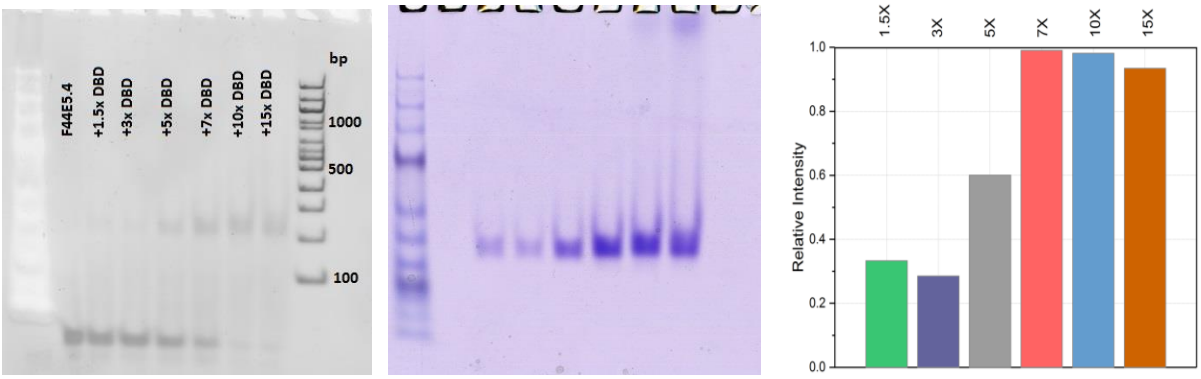
Table 3.1 HSE-containing probes designed from the promoter sequences of chaperone genes and used in binding studies. The relationship to cliques is indicated in the last column. The HSEs (nGAAn) are highlight in green in forward sequence.

EMSA was performed to test the interaction between the purified HSF-1 DBD and the dsDNAs. Since the promotor of F44E5.4 contains the most amounts of HSEs, it is selected for testing the saturation in Hsf-1 DBD titration (Figure 4.3.1 A). Hsf-1 DBD was titrated with 0x, 1.5x, 3x, 5x, 7x, 10x and 15x molar excess to 1.5 µM DNA. The formation of complex bands and the reduction of free DNA were observed in EMSA. At an excess of 10X DBD, no further reduction of free DNA was observed and the existence of free Hsf-1 DBD becomes clearly visible on the gel. Hence to this end, a 10-fold excess of HSF-1 DBD over each DNA probes was used to reach the saturated level of protein-DNA complex. At this stage the affinity of DNA probes obtained from the promotor region of genes regulated by HSF-1 and negative control were tested (Figure 4.3.1 B). The reduction in migration speed differs widely for each probe. While the *dnj-13, bag-1, unc-23* and *hsp-1* promoters barely interacted with HSF-1 DBD, others like dsDNA derived from the promoter of *F44E5.4, hsp-70, hsp16.2* and *dnj-12* bind strongly and form an intense band representing the DNA-protein complex. This indicates that regardless of the HSF amount, the DNA originated from Hsf-1 inducible genes have higher affinity to Hsf-1.

Cooperativity in binding to multimeric HSEs can be revealed by SV-AUC.

The interaction of HSF-1 DBD to these promotor-derived sequences was further quantified based on an analytical ultracentrifugation experiment. To this end, a titration in SV-AUC was performed by adding different concentrations of HSF-1 DBD to DNA probes. 1.5 µM dsDNA was supplemented with different molar excess of HSF-1 DBD (0x, 1.5x, 3x, 5x, 7x, 10x and 15x) over the DNA. This resulted in binding isotherms that can be analyzed regarding the affinity, stoichiometry and cooperativity of HSF-1 DBD on the DNA. In the titration, the shift of peaks with increasing protein concentration indicates the formation of complexes at different ratios. The complex with *F44E5.E*p reaches a saturated level when the dsDNA to HSF-1 DBD ratio is about 1:10. At this point, the signal of remaining unbound HSF-1 DBD is clearly visible, which implies that the complex formation is a saturatable equilibrium process. The detailed binding site occupancy was then determined based on the sedimentation profile in SV-AUC experiments as described in the Materials and Methods section.

A.



B.

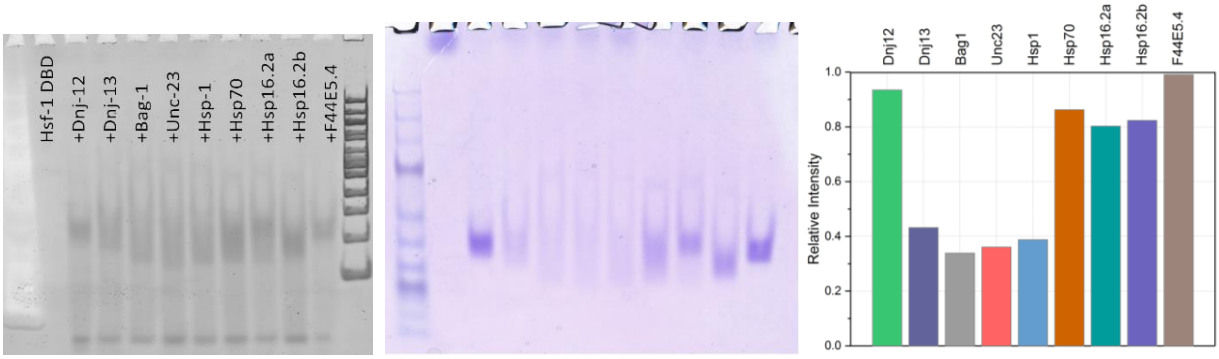
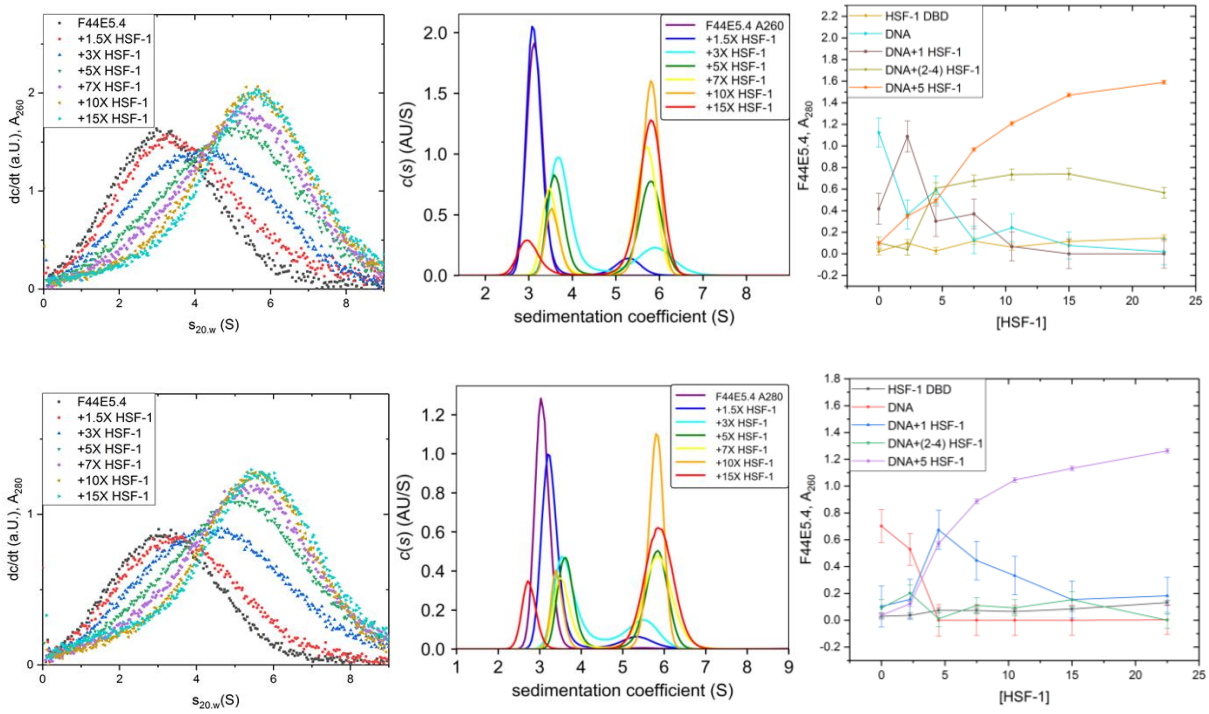


Figure 4.3.1 EMSA analysis of DNA•Hsf-1 complex formation. DNAs were stained with the DNA stain (SERVA, grey colored images in first row) and proteins were stained with Coomassie Blue (blue colored images in second row). The relative intensity (third row) was measured via ImageJ. (A) The titration of HSF-1 DBD to the promotor F44E5.4; (B) Comparison of selected DNA promotors.

The analysis of several distinct HSE-sequences contributes to understand different interactions of HSF-1 DBD with these genes. A fitting method based on the UltraScan custom grid was used to determine partial concentrations of the individual species during the assembling process as described in the Materials and Methods section. This method offers the opportunity to fit concentration dependent changes to the species occupancy and thereby distinguish estimated values for affinities and stoichiometries for HSF-1 DBD binding to each of the probes. The titration data fitted to this model yielded values for each species' concentration. As a reference, single components including ssDNA, ds DNA and the purified Hsf-1 DBD were measured under the same condition in the SV-AUC. The dc/dt curves don't show many differences (Suppl. Figure 3.1) among the single strained- and double strained DNA. With this measurement, initial data points of the custom grid were obtained. Since the F44E5.4 contains the largest amount of HSEs

it was selected for the titration experiment, from which data points of the 2DSA analysis were obtained and the partial concentration of each complex species in the same mixture was calculated. Starting from the dc/dt plots the partial concentration of each complex species is fitted with multiple methods to invalidate the 2DSA analysis. For the sample containing F44E5.4, fitting results are given in the figure 4.3.2 A (obtained from Sedfit/Sedphat or Gaussian function) and the figure 4.3.2 B (obtained from 2DSA). The comparison of 2DSA results of each complex species formed with different DNAs is shown in figure 4.3.3 and the corresponding fitting with Gaussian function are given in the Suppl. Figure 3.2-3.6. The fitting results reveal a stepwise binding procedure of Hsf-1 DBDs onto the multimeric HSE containing DNA. At a lower HSF-1 concentration, complexes containing monomeric or dimeric HSF-1 were identified, which was upgraded to complexes containing multimeric HSF-1 DBD when the amount of available protein was elevated.

A. **dc/dt** **Fitted by Sedfit/Sedphat** **Fitted by Gaussian function**



B.

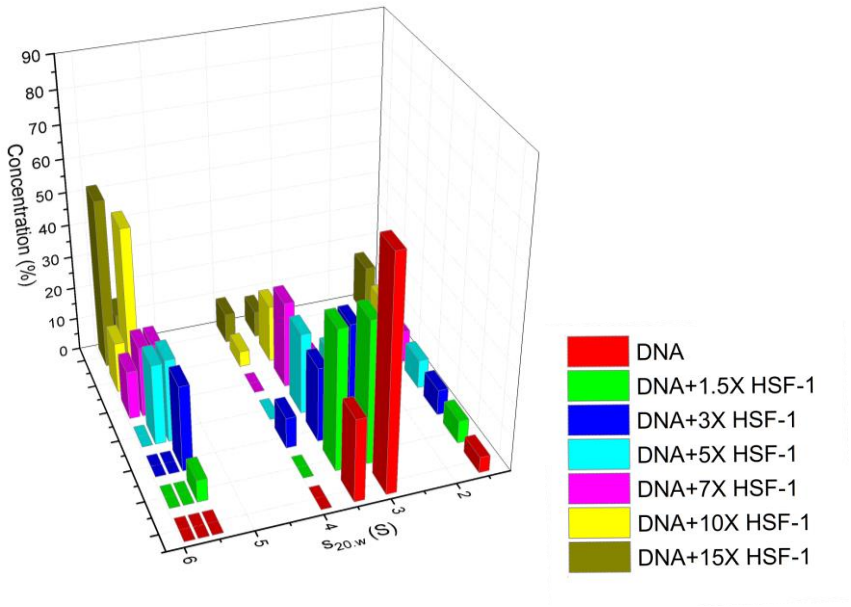


Figure 4.3.2 Analysis of the interaction between promotor F44E5.4 and Hsf-1 DBD via titration in SV-AUC. (A) dc/dt plot of the absorbance measured at 260 nm (up) and at 280 nm (down) with corresponding fitting analysis performed via the Sedfit/Sedphat or Gaussian function. (B) Partial concentration of complex species obtained via the custom grid fitting in the F44E5.4 sample.

F44E5.4 contains pentameric HSEs and is one of the highest HSF-1 induced genes. The HSEs are more often to be found in trimeric form in the DNA region. To understand how HSF-1 DBD interacts with other genes containing less HSEs (Table 3.1), which differ in the sequence and consensus score, the HSF-1 affinity of each probe was tested. Results of the experiment seem to correlate well with the predicted consensus score for HSF-1. The tetrameric HSEs were also tested at constant concentrations of DNA to derive, whether binding at these sites is favored over binding to trimeric HSEs (Figure 4.3.3). Indeed, differences were observed, when identical concentrations of HSF-1 DBD were used, suggesting a preferential binding of HSF-1 DBD to the larger HSEs. Further the binding stoichiometry of trimeric, tetrameric and pentameric DNA probes are different. The comparison of titration results shows a higher sedimentation coefficient after binding tetrameric probe and a further increased $s_{20,w}$ after binding pentameric probes. It is very clear from these data that different assembly mechanisms are applied in different probes and different stoichiometries must be assumed. Furthermore, the models imply some degree of cooperativity during the assembly process, implying that the binding of one Hsf-1 DBD to the DNA facilitates the binding of the next and the third.

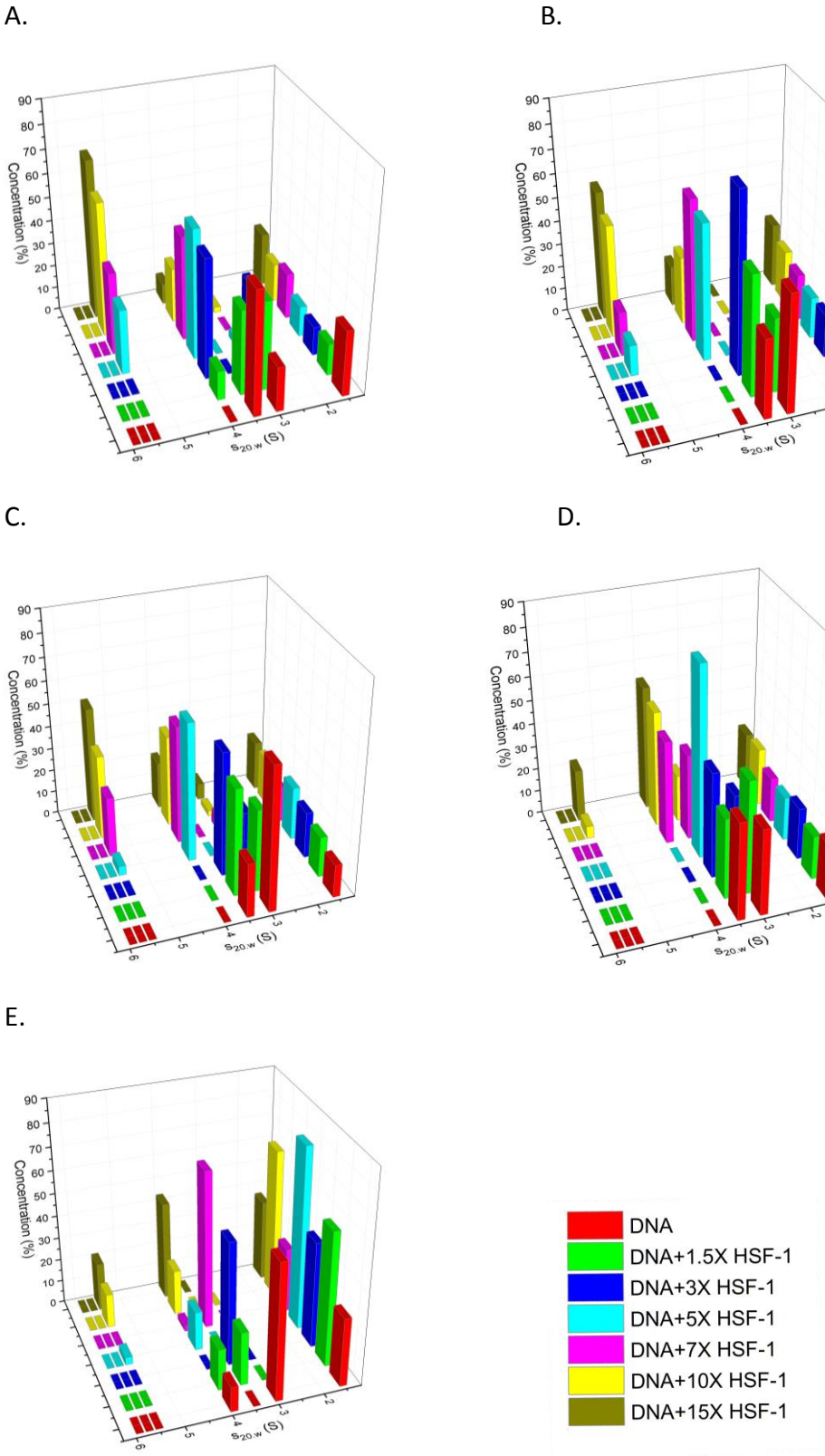


Figure 4.3.3. Partial concentration of each species in DNA-Hsf-1 samples measured in AUC. DNA = (A) HSP16.2a; (B) HSP16.2 b; (C) HSP70; (D) HSP1; (E) DNJ13.

To gain further knowledge on how Hsf-1 interacts with DNAs, crosslink reactions of HSF-1 DBD with the promotor *F44E5.4* (highest consensus score pentameric site) was performed. Previously researchers built the crystal structure of HSE with the human HSF-1 DBD (PDB: 5D5U), where the distance between the two HSF-1 DBD binding to the anti-parallel DNA double string is between 8-13 Å. Thus the crosslinker BS3 with a spacer length of 11 Å was used. The protein alone is unable to form sufficient crosslinked products while in the presence of DNA promotors the crosslinked product accumulate to clear bands, whose intensity correlates proportionally to the amount of HSEs in the promotor (Figure 4.3.4 A, first lane). To understand how HSF-1 DBDs are arranged on the DNA, the purified 15-N isotopic labeled HSF-1 was used to analyze the potential interaction site with the DNA promotor E44F5.4. The crosslink reaction was performed using 14N HSF-1, 14N:15N = 1:1 and 15N HSF-1 (Figure 4.3.4 A, sample 1-3). Corresponding crosslinked bands were analyzed in MS. On the gel, crosslink bands for the dimeric Hsf-1 so as bands containing oligomeric Hsf-1 are visible, which indicates that HSF-1 DBD binds in multimeric units onto the DNA. Due to limited yield, only the dimeric crosslinked bands were further analyzed in MS.

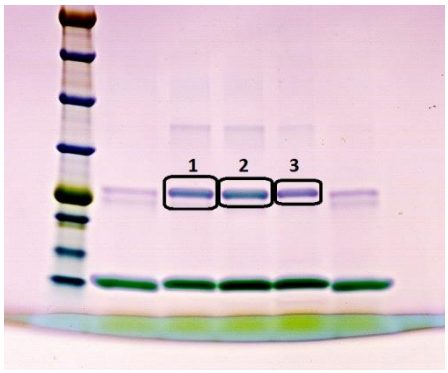
The precursors found from MS scan were plotted in the figure 4.3.4 B-C with the majority distributed in the range of ±5 ppm. In total 55 spectra were found containing crosslinked pairs, from which 14 spectra are identified for intermolecular crosslinked peptide pairs in Hsf-1 DBD samples (figure 4.3.4 D and Table 3.2). 11 of the 14 spectra are assigned for the peptide pair KMTPLSQGGLTR (1) – KQSAR (1), containing peptide pairs of both 14N and 15N labeled protein as well as for light and heavy crosslinked pairs. The remaining three spectra are assigned for the peptide pair PELLSQIKR (8) – KQSAR (1), which only contain the light crosslinked 14N-15N peptide pair and the heavy crosslinked 14N- peptide pair. The heavy crosslinked 15N peptide pair is not found. The mass error for all of these 14 precursors is less than 2 ppm. The spectra of both identified peptide pairs with light or heavy crosslinker are shown in the Suppl. Figure 3.8 as an example.

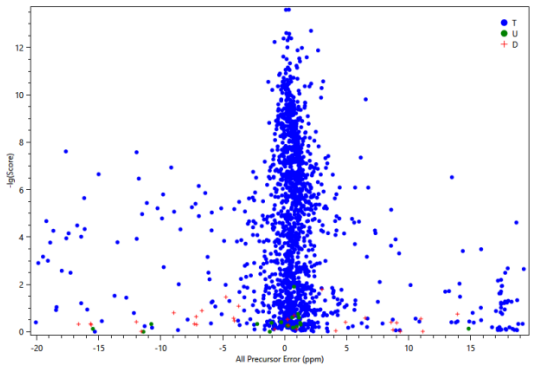
The identified peptide pairs located at the terminal region of Hsf-1 DBD are labelled in yellow in the figure 4.3.4 E. MS results indicate that the interaction of Hsf-1 DBDs in the DNA-protein complex mainly is a head to tail (KMTPLSQGGLTR (1) – KQSAR (1)) contact to the neighboring Hsf-1 DBD. On the other side the existence of the tail to tail (PELLSQIKR (8) – KQSAR (1)) crosslinked peptide indicates the existence of another arrangement option which is less preferred.

Table 3.2 Crosslinked peptide pairs of Hsf-1 DBD identified in MS.

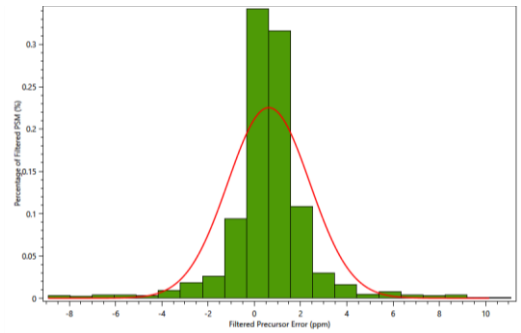
Precursor Mass	Peptide	Linker	Peptide Mass	Score	Precursor Mass Error (Da)	Precursor Mass Error (ppm)
2.03E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3_d12	2.03E+03	9.06E-02	-0.00035	-0.172653
2.03E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3_d12	2.03E+03	9.23E-03	-0.000225	-0.110992
2.03E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3_d12	2.03E+03	6.56E-02	0.000358	0.1766
2.03E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3_d12	2.03E+03	1.03E-01	0.000664	0.327548
2.02E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3	2.02E+03	4.50E-02	0.001236	0.613367
2.02E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3	2.02E+03	1.32E-02	0.001336	0.662992
2.03E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3_d12	2.03E+03	8.43E-02	0.001445	0.712812
2.05E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3_d12	2.05E+03	2.94E-01	0.002107	1.025752
2.02E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3	2.02E+03	4.74E-01	0.002789	1.384046
2.04E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3	2.04E+03	1.23E-02	0.0032	1.567071
2.04E+03	KMTPLSQGGLTR (1)-KQSAR(1)	BS3	2.04E+03	6.23E-02	0.003521	1.724268
1.83E+03	PELLSQIKR(8)-KQSAR(1)	BS3	1.83E+03	2.30E-01	-0.001406	-0.766637
1.82E+03	PELLSQIKR(8)-KQSAR(1)	BS3_d12	1.82E+03	5.62E-01	-0.00137	-0.751867
1.81E+03	PELLSQIKR(8)-KQSAR(1)	BS3	1.81E+03	7.09E-02	0.000963	0.532028

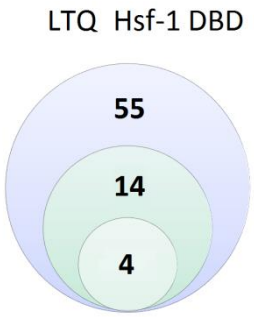
A. B.





C. D. E.





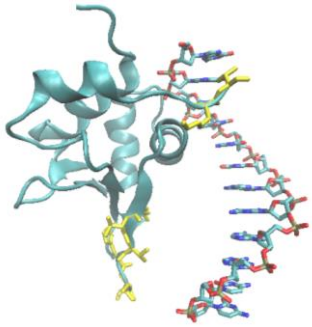


Figure 4.3.4 Identification of Hsf-1 DBD interaction site. (A) BS^3 crosslinked Hsf-1 DBD and DNA promotor F44E5.4. Lane 1: 14N-labeled HSF-1; lane 2: 14N:15N-labelled Hsf-1 = 1:1; lane 3: 15N-labeled HSF-1. These gel bands were then analyzed in MS. (B) Precursor error distribution of identified crosslinked bands via MS. (C) Filtered precursor error distribution. (D) Venn graphic of the crosslinked peptide pairs. Purple indicates the total amount of identified spectra; green indicates the intermolecular crosslink site; white indicates the amount of confirmed intermolecular crosslinked pairs in both 14N and 15N labelled Hsf-1 DBD. (E) Identified crosslinked sites were labeled in yellow in Hsf-1 DBD (Based on PDB structure: 5D5U).

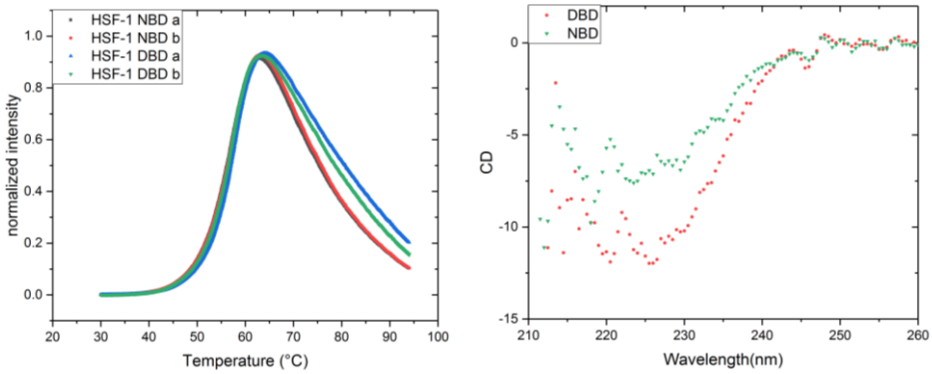
4.3.3 Additional 82 amino acids in the N-terminal domain of the nematode HSF-1 blocks the DNA-protein complex formation

The nematode HSF-1 is characterized by additional 82 amino acids at its N-terminus. It is interesting to know whether these 82 N-terminal amino acid residues have an influence on physical properties and the functionality of HSF-1 DBD. To this end a Hsf-1 fragment containing this additional amino acid stretch and the HSF- 1 DBD region was generated (Hsf-1 AA 1-198, short as NBD). TSA assay of the HSF-1 NBD was performed to check the influence of the N-

terminal amino acids on the stability of Hsf-1 DBD (Figure 4.3.5 A). In the TSA no obvious differences were observed between the Hsf-1 BDB and NBD. The circular dichroism (CD) spectra (Figure 4.3.5 B) and unfolding transitions recorded by CD between 10 and 90 °C (Suppl. Figure 3.9) for both Hsf-1 domains are comparable, implying that the N-terminal residues do not influence the structure of HSF-1 DBD core.

To investigate the functionality of the N-terminal amino acid stretch, EMSA of the Hsf-1 NBD was performed with different DNA probes (Figure 4.3.5 C-D). Unlike the DBD, the Hsf-1 NBD showed very weak interaction with all tested DNA probes, which implies that the N-terminal amino acids, although not altering physical properties, has a reduced affinity to HSEs compared to Hsf-1 DBD.

A. B.



C. D.

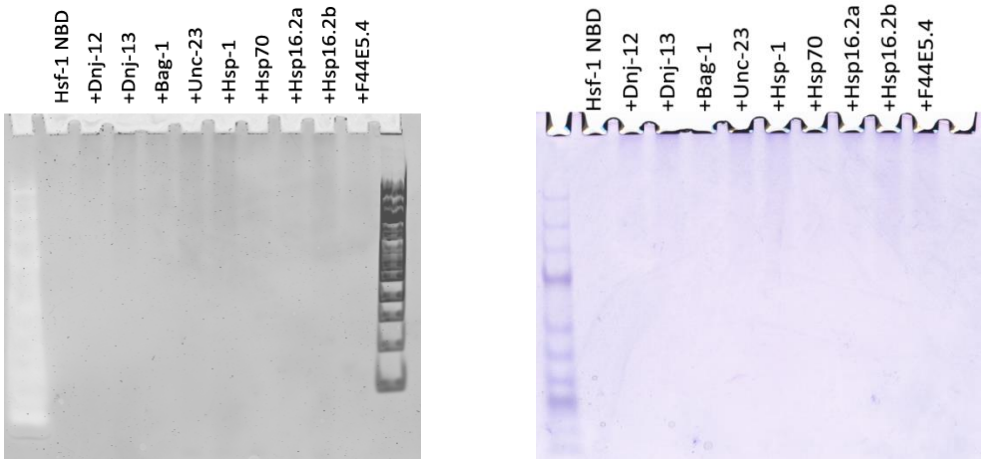


Figure 4.3.5 Comparison of the N-term Hsf-1 DBD vs Hsf-1 DBD. (A)Thermal shift assay recorded between 30-95°C. (B) CD spectrum recorded between 210-260 nm. (C) DNA stain and (D) protein stain of the Hsf-1 NBD with different DNAs in EMSA.

4.4 sB-Raf complexes with the HSP-90 system are stabilized by the presence of KFB-6

4.4.1 Nucleotides antagonize against protein chaperons in complex formation

HSP-90 and CDC-37 are reported to form protein complexes with the stabilized kinase domain of B-Raf (sB-Raf) and the complexes are sensitive to the presence of ATP [37]. To gain further knowledge on the interaction between kinase and chaperons, ATTO488-labelled sB-Raf (*sB-Raf) is used to investigate the complex formation within the kinase-chaperon system. Addition of CDC-37 initiates the complex formation, increasing the $s_{20,w}$ of sB-Raf from 2.8 S to 4.3 S, which corresponds roughly to the heterodimeric complex observed before with labelled *CDC-37 [120].

To address the influences of different nucleotides on the formation of binary complex, crosslink reactions and analytical ultra-centrifugation are applied (Figure 4.4.1 A-B). The complex is disrupted in the presence of ATP and *sB-Raf sediments again as a monomeric protein. In terms of sedimentation coefficient changes, a similar effect can be observed for ADP, but AMP-PNP and ATPγs do not influence the interaction within the complex. Unlabeled proteins were crosslinked to test, whether the release of kinase also can be observed under these conditions and whether the binding of the nucleotide to the kinase influences the complex formation. Surprisingly, the influence of the nucleotide ADP, ATP and AMP-PNP in preventing sB-Raf from forming a chaperone complex depends on the order of addition. The protection of ADP weakens binary complex bands in crosslinking experiments and fully dissociates the binary complex in AUC experiments. In contrast, in the presence of ATPγs, the binary complex formation is blocked when ATPγs is added priory to the reaction mixture (in crosslink) and the complex is formed slowly at the equilibrium stage (in AUC).

It is well described that CDC-37 delivers HSP-90 client proteins to HSP-90. It was tested next whether the presence of HSP-90 has an effect on influences of nucleotides (Figure 4.4.1 C-D and Suppl. Figure 4.1). The formation of full ternary protein complexes with an $s_{20,w}$ of 6.5 S has been observed [65]. This complex appeared less sensitive to nucleotides compared to the binary complex with CDC-37. Addition of ADP leads to a strong dissociation of the complex. Addition of ATP, AMP-PNP and ATPγs does not prevent the formation of ternary complex and ATPγs even induces an increase in the sedimentation coefficient to 7.0 S, which implies that a closing of HSP-90 dimer occurs. This highlights the complexity of interactions in this protein complex, when both, the kinase and the chaperone can bind nucleotide. These data imply that the sB-Raf can be complexed by the chaperon system as well as by nucleotides. In the binary complex the presence of ATP, AMP-PNP and ATPγs induces the disassociation of the CDC-37•sB-Raf complex while the ternary complex is fairly stable in the presence of these nucleotides. ADP in general has no influence on the binary complex but supports an open conformation of HSP-90, which causes the disassociation of ternary complex. On the other side the order of addition is relevant:

If ATP is added to the kinase prior to the addition of chaperone, the complexes cannot be assembled, if ATP is added afterwards it still obviously reduces the extent of complex formation in cross-linking experiments (Figure 4.4.1 D), but may not have the same efficiency.

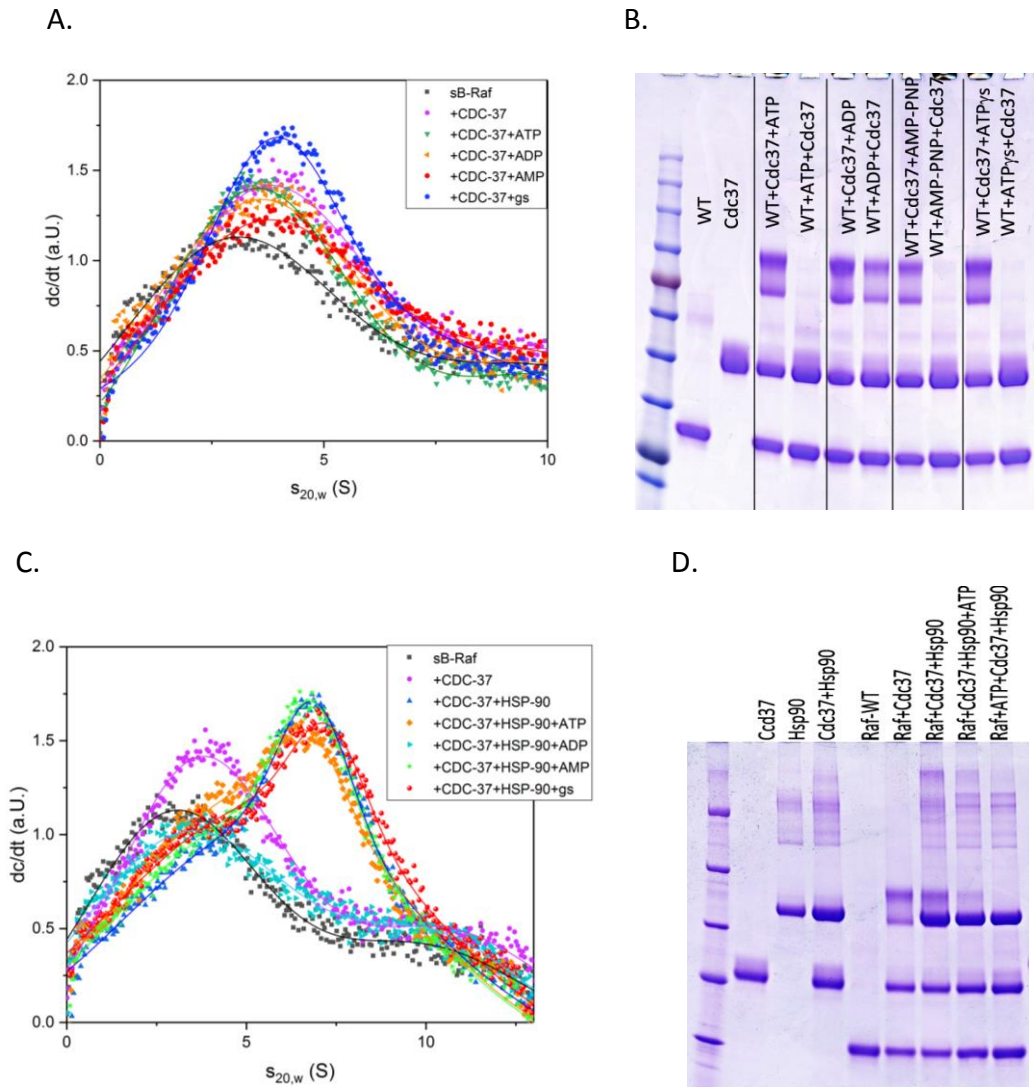


Figure 4.4.1 Influence of nucleotides on kinase complex formation. (A) Formation of the sB-Raf•CDC-37 binary complex and influence of ATP (green), ADP (orange), AMP-PNP (purple) and ATPγs (blue). (B) Influence of nucleotides on the binary complex formation when added at a different order. (C) Formation of sB-Raf•CDC-37•HSP-90 ternary complex and influence of ATP (orange), ADP (dark green), AMP-PNP (light green) and ATPγs (red). (D) Influence of ATP on the formation of the ternary complex when added at a different order.

4.4.2 FKB-6 stabilizes HSP-90•CDC-37•sB-Raf complexes

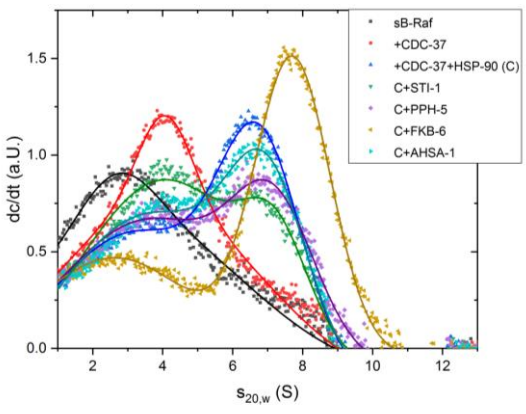
HSP-90 relies on its co-chaperons to fulfill its function. Thus it is interesting to know whether the presence of additional cofactors can influence the complex formation. To this end the complex formation with the HSP-90 co-factor AHSA-1, STI-1, PPH-5 and FKB-6 was compared as these proteins have been reported to interact with HSP-90 in the nematode chaperone system [121-124].

Addition of AHSA-1 reduced the amount of sB-Raf•CDC-37•HSP-90 ternary complex and so does STI-1, but to a lower extent (Figure 4.4.2 A). PPH-5 reduced the amount of CDC-37•HSP-90 bound to sB-Raf. Surprisingly, FKB-6 stabilizes the sB-Raf•CDC-37•HSP-90 complex by forming a stable complex at almost 8 S. The striking increase in sedimentation coefficient implies that an interaction between the large PPIase and the rest of the complex could be relevant for the interaction between kinases and the chaperone system.

Next the influence of ATP on the formation of these quaternary complexes was tested (Figure 4.4.2 B). For all HSP-90 co-factors aside of FKB-6, addition of nucleotides leads to dissociation of the chaperone complex, where a reduction of the ternary peak can be observed. In the case of FKB-6, the quaternary complex is not destabilized by the presence of the ATP, but shifts towards a slightly larger sedimentation coefficient ($s_{20,w}$ = 9S), which possibly is a sign of the closure of the HSP-90 N-terminal domains.

Different types of nucleotides were tested as well (Suppl. Figure 4.2). In the presence of ADP and partially in the presence of ATP, where large chaperone-kinase complexes could be observed before, the addition of nucleotides partially dissociates the quaternary complex. On the other hand the addition of AMP-PNP and ATPγs leads to complexes with a slight shift to higher sedimentation coefficients, confirming that these nucleotides in principle are able to close the HSP-90 chaperone even in the presence of client and two cofactors.

A.



B.

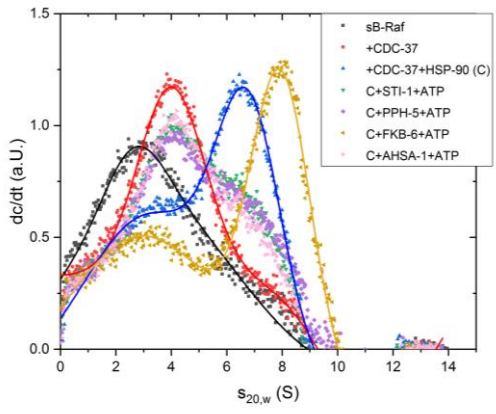


Figure 4.4.2 Influence of HSP-90 co-chaperons on the ternary complex of sB-Raf•CDC-37•HSP-90. (A) Changes of sB-Raf•CDC-37•HSP-90 ternary complex induced by the presence of HSP-90 co-chaperons, measured with fluoresce labeled sB-Raf. (B) Influences of nucleotides on HSP-90 co-chaperons complexes.

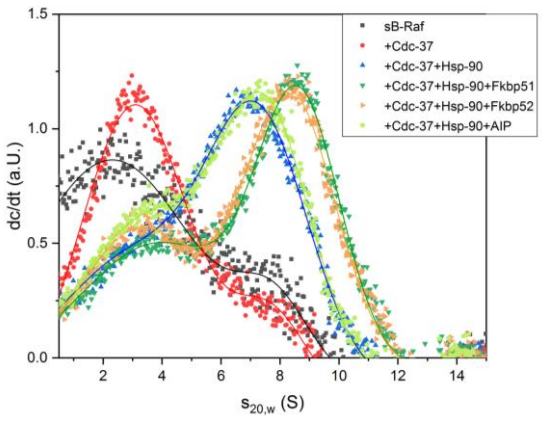
4.4.3 Human Fkbp-like proteins integrate into Cdc-37·sB-Raf·Hsp-90 complexes

Previously observed results are obtained from the nematode chaperone system, which only contains one type of large PPIase, FKB-6. To test whether the result is generally applicable in the human system, the human FKB-6 homolog Fkbp51, Fkbp52 and Aip were tested in similar experiments with labelled sB-Raf.

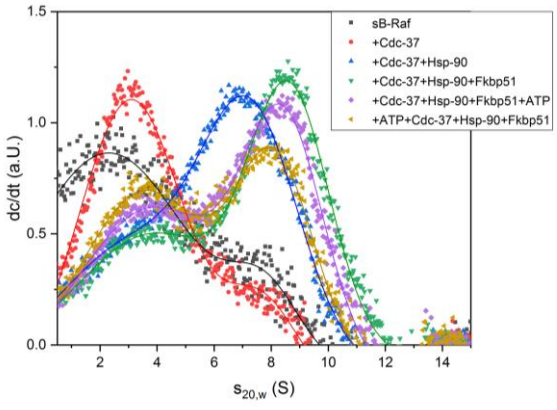
Ternary complexes between sB-Raf, Hsp90β and human Cdc37 were formed similarly to the nematode protein complex described before. Addition of the two large PPIases Fkbp51 and Fkbp52 shows their ability to integrate into the complex. Both Fkbp51 and Fkbp52 form a stable quaternary complex with labeled sB-Raf•Cdc37•Hsp90β, which increased the sedimentation coefficient to 9 S. In contrast Aip showed barely any interaction with the sB-Raf•Cdc37• Hsp90β ternary complex (Figure 4.4.3 A).

The influence of ATP was tested in the human system as well (Figure 4.4.3 B-D). Unlike in the *C.elegans* system, the presence of ATP partially disassociates the quaternary complex formed with Fkbp51 and Fkbp52 and leads to a reduced $s_{20,w}$ value. In the case of Aip the ternary complex is partially disassociated after addition of ATP. Taken together, the interaction between kinase, Hsp90 and cofactors is mostly conserved in the human system. However the effect of the large PPIase and the influence of nucleotides are weakened in the human system, while the stability of the complex also is lower in the human system.

A.



B.



C. D.

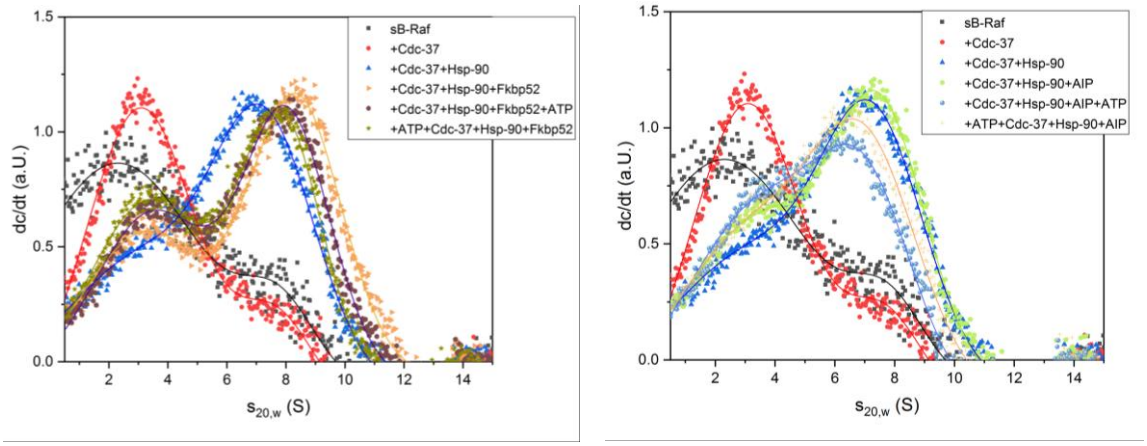


Figure 4.4.3 Formation of the quaternary complex in the human system. (A) Formation of the quaternary complex with the human homolog chaperon Fkbp51 (dark green), Fkbp52 (orange) and AIP (light green). (B) Influence of nucleotides on the quaternary complex with Fkbp51. (C) Influence of nucleotides on the quaternary complex with Fkbp52. (D) Influence of nucleotides on the quaternary complex with AIP.

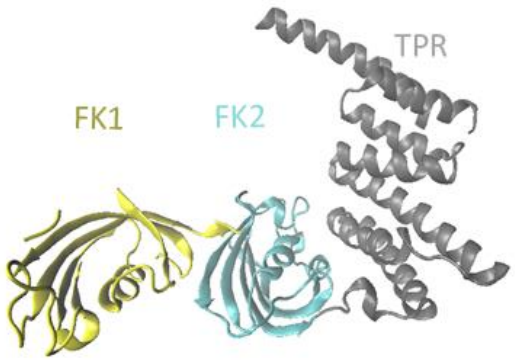
4.4.4 Fkbp-like domains in FKB-6 influence the complex stability in different ways

In both *C.elegans* and human system, FKB-6 or its homologues are able to stabilize the sB-Raf containing complex. To further understand how FKB-6 fulfills this function, the structure of FKB-6 is observed in detail. FKB-6 mainly contains two FKBP domains and a C-terminal TPR domain. Aimed to understand the role of FKBP domains in the complex formation, constructs of FKB-6 with deleted domains are generated and the performance of these FKB-6 mutants in quaternary complex formation is observed. While the structure of FKB-6 has not been solved to date, the structure of its human homolog Fkbp51 is known [125] and the 42% shared identity between the related proteins ensured a generation of variants based on the human homolog. Thus, two constructs were generated: the deletion of the first FKBP-domain in FKB-6 is named as ΔFK1-FKB-6 in the following text and the deletion of the second FKBP-domain correspondingly is named as ΔFK2-FKB-6. The deletion of the second FKBP-domain is achieved by linking the first FKBP domain and the C-terminal TPR region with a $(GGGGS)_3$ linker (Figure 4.4.4 A, alignment shown in Suppl. figure 4.3).

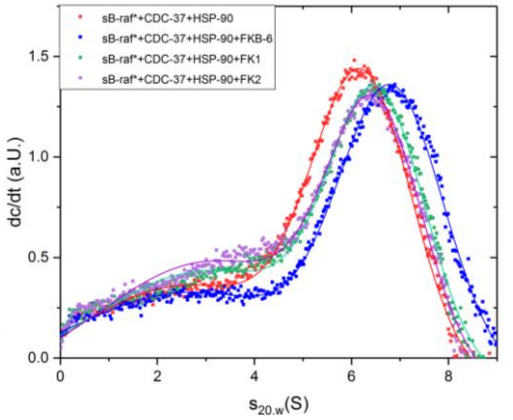
AUC results on the complex formation with these deletion constructs imply that both FKB-6 mutants are still able to bind to the HSP-90 chaperone system on a comparable level, with a slightly smaller $s_{20,w}$ compared to the complex formed with full length FKB-6 (Figure 4.4.4 B).

This indicates that the deletion of either FKBP-domain has no influence on the binding of FKB-6 to HSP-90. This is to be expected as FKB-6 binds to HSP-90 via its TPR domain. How FKBP-depleted versions affect the stability of the kinase-chaperon complex was tested by adding different nucleotides to the complex. In the presence of ATP, sB-Raf is released from the complex and more free sB-Raf can be detected with a sedimentation coefficient around 2.0 S. The ability of FKB-6 variants to maintain sB-Raf in the chaperon complex nevertheless is dramatically reduced compared to full-length FKB-6 (Figure 4.4.4 C). The same is observed in the presence of ADP, where likewise a reduced stability is observed (Figure 4.4.4 D). Thus, while less important for the complex formation in general, the ΔFK1 and ΔFK2 fragment nevertheless regulate the stability of the kinase-chaperon complex, possibly by different mechanisms. Despite the need to further evaluate the contribution of FKBP-domains in the client recognition; apparently both FKBP-domains can supplement each other when forming these complexes in different ways.

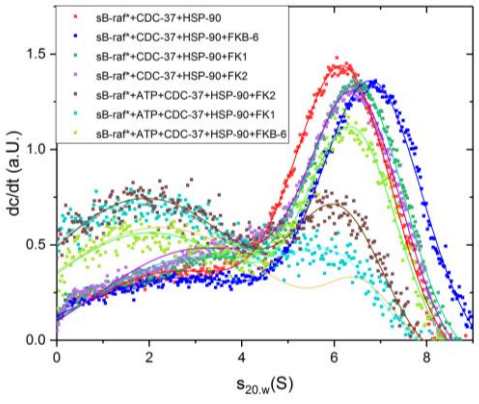
A.



B.



C.



D.

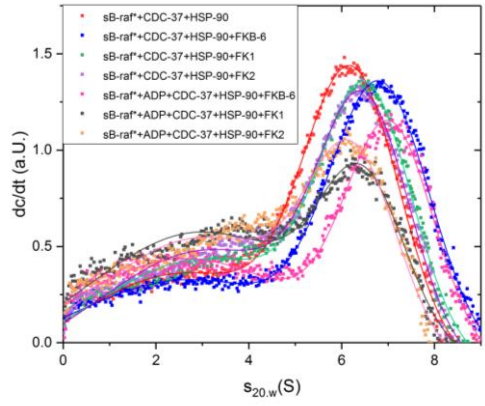
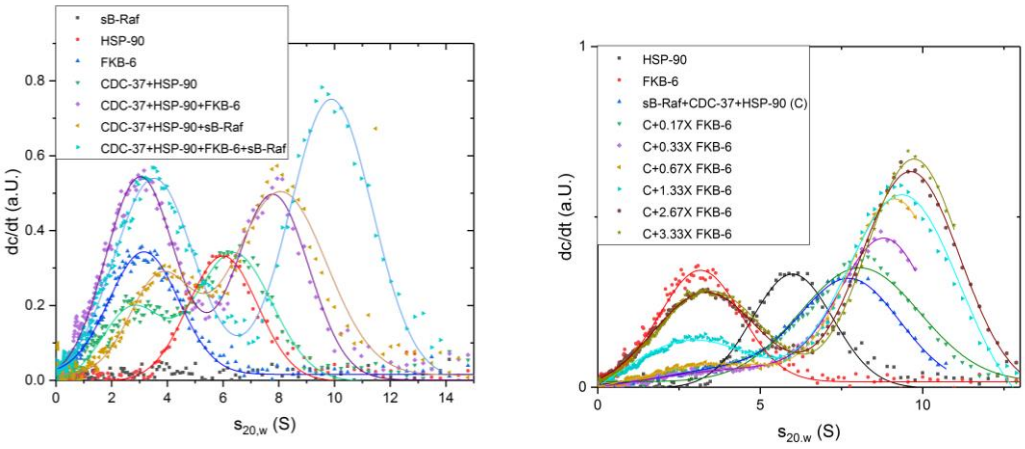


Figure 4.4.4 Influence of Fkbp domains on the complex formation. (A) Structure of the human homolog Fkbp51 (PDB: 1KT1) is used as reference to mutant FKB-6. The first Fkbp domain is colored in yellow, the second Fkbp domain is colored in light blue and the TPR domain is colored in grey. (B) Comparison of the complex formation with FKB-6 and its mutants. (C) Influence of ATP on the complex formation. (D) Influence of ADP on the complex formation.

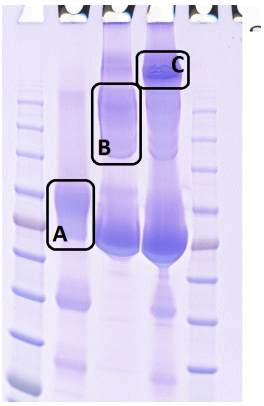
4.4.5 Structural investigations of the kinase-chaperon complex

The stoichiometry of the Cdk4•Cdc37•Hsp90 complex has been determined as 1:1:2 [126]. Aimed at further understanding how the kinase and cochaperons interact with HSP-90 in the sB-Raf containing system, the complex formation is validated by analytical ultracentrifugation and native mass spectrometry. Using UV/VIS absorbance optics in the AUC allows observing all proteins in a sample and evaluating the stoichiometric relationships. Here the absorbance at 280 nm was measured for each single component as well as for the complex mixture. The stoichiometry of FKB-6 is tested by titration of FKB-6 to the pre-formed sB-Raf•CDC-37•HSP-90 complex. In the complex sample the peak at 6.3 S representing the HSP-90 dimer was almost abolished. Instead large complexes at ca. 10 S indicated a complete saturation of HSP-90 and its entire inclusion into larger protein complexes, when all cofactors and client are in excess over the chaperone (Figure 4.4.5 A, light blue curve). The titration reveals saturation at a ratio of 1:1 (Figure 4.4.5 B) indicating that only one FKB-6 molecule is associated in a complex molecule. The stoichiometry was also measured by native mass spectrometry [127, 128]. This is validated by the comparison of the native MS data of the crosslinked complex (Figure 4.4.5 C-D). Here clearly a dominant formation of the complex sB-Raf•CDC-37•FKB-6•HSP-90 at a stoichiometry of 1:1:1:2 can be observed, containing 1 client, one set of cofactors and a dimeric HSP-90.

A. B.



C.



D.

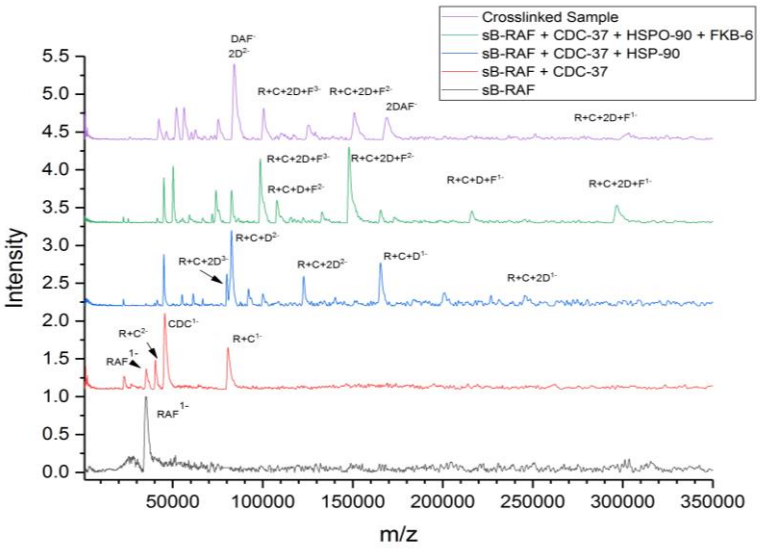


Figure 4.4.5. Kinetics of the kinase-complex. (A) Stepwise formation of the complex measured by UV absorbance at 280 nm. (B) Titration of the FKB-6 to sB-Raf•CDC-37•HSP-90 ternary complex. (C) Crosslink of the complex vs single components. Sample A = sB-Raf with CDC-37; B = Daf-21 dimer; C = crosslinked complex. (D) Native MS of the crosslinked sample from (C) vs single components.

It is well known that the function of CDC-37 is to deliver a client to HSP-90. In contrast the function of FKB-6 in the kinase complex is not known to date. To find a clue on the complex assembly and the potential function of FKB-6 in the complex, the crosslinked complex is digested and analyzed via MS.

As qualification of the analysis, all of the spectra obtained from the two replicates were shown in the Suppl. Figure 4.4 A-B. The spectra were filtered with an error range of ±6 ppm and histograms of the filtered spectrum are shown in the Suppl. Figure 4.4 C-D. The majority of precursors detected are found within an error range of ±6 ppm conforms that both experiments for this complex are of sufficient quality. Compared to the LTQ replicate, the QE replicate delivered more spectra which are identified for intermolecular peptide pairs. Although using different algorithm, the two analysis programm reported comparable results. Due to the diversity of typsin cutting combinations, different peptide pairs were reported, after plotting all crosslinked peptide pairs in the protein sequence, it is observed that the main interface in the complex are similar in both replicates. Data analysis was performed based on the filtered spectra with an error range of ±2 ppm and the amount of spectra is given in the Suppl. Figure 4.4 E. With the strategy outlined in the Method section, both intramolecular and intermolecular crosslinked peptide pairs were detected. While the intramolecular crosslinks were used to confirm the rationality of results, the intermolecular crosslinked sites were used to describe the

relatively arrangement of single particles in the complex. The numbers of spectra identified for intermolecular crosslinked peptide pairs in each sample are shown in the Suppl. Figure 4.4 F.

The intermolecular crosslinked peptides identified in the three experiments (sB-Raf•CDC-37•FKB-6•HSP-90 and as control sB-Raf•CDC-37, dimeric HSP-90) are listed in the table 4.1. Examples of spectra of identified peptide pairs for the sB-Raf•CDC-37 complex as well as the crosslinked sites in each protein component are given in the figure 4.4.6 and the Suppl. Figure 4.5.

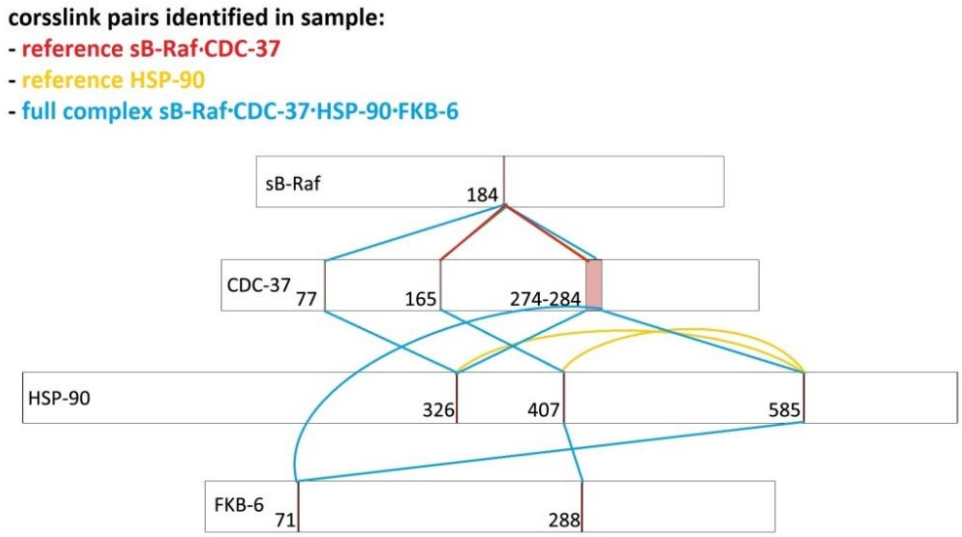


Figure 4.4.6 Structural information obtained from MS. Schematic diagram of key crosslinked pairs identified from MS samples. Red lines indicate crosslink sides in the sample sB-Raf•CDC-37; Yellow lines indicate pairs identified in the HSP-90 sample; Blue lines indicate pairs from the complex sample.

In total there are two intermolecular crosslinked pairs (IGDFGLATVKSR (sB-Raf, 184) – KPQAPK (CDC-37, 165) and IGDFGLATVKSR (sB-Raf, 184) – QFFKK (CDC-37, 273)) identified in the sB-Raf•CDC-37 complex which locate in the M-domain of CDC-37 and the N-and C-lobe interface of the kinase domain. In the HSP-90 sample, plenty of peptide pairs are identified. Of all these peptides, two pairs of crosslinked peptide pairs were identified to build new contacts in the complex, which are mainly located within the M-domain: FYEQFGKNLK (HSP-90, 407) – IMKAQALR (HSP-90, 585) and APFDLFENKK (HSP-90, 326) – IMKAQALR (HSP-90, 585). These are the dimer interaction sites, but it is unclear whether these crosslinks took place within the subunit or between two subunits. In total 37 unique intermolecular crosslinked peptide pairs in the complex were identified (Table 4.1). In the following, changes detected in the fully

assembled complex were compared to the references, which mainly involve these crosslinked pairs (Figure 4.4.6, blue curve): IGDFGLATVKSR (sB-Raf, 184) – MAEKKMEQEK (CDC-37, 77) – APFDLFENKK (HSP-90, 326); KPQAPK (CDC-37, 165) – KFYEQFGK (HSP-90, 407) and AEEVLEYEKSTDPEK (FKB-6, 288) – KPQAPK (CDC-37, 165); IGDFGLATVKSR (sB-Raf, 184) – KFEAAEPVYMK (CDC-37, 274) – APFDLFENKK (HSP-90, 326), KFEAAEPVYMKHYQDEVK (CDC-37, 284) – IMKAQALR (HSP-90, 585) and FEAAEPVYMKHYQDEVK (CDC-37, 284) – GNVIKGWDLGVATMTK (FKB-6, 71). Comparison to the reference samples showed that in the complex molecule, previous sB-Raf•CDC-37 contact sites were separated and they build new connections in the complex with HSP-90 and FKB-6 separately.

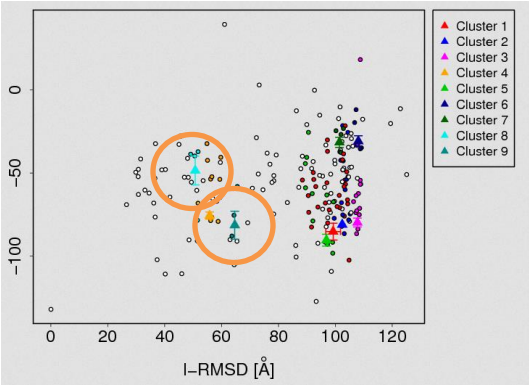
The most interesting aspect revealed by the crosslinking result is the behavior of FKB-6 in the complex. In total two crosslinked peptide pairs are identified containing FKB-6 motifs. One motif which locates in the TPR region of FKB-6 is crosslinked to HSP-90 and the M-domain of CDC-37 (AEEVLEYEKSTDPEK (FKB-6, 288) – KPQAPK (CDC-37, 165) – KFYEQFGK (HSP-90, 407)), where the CDC-37 motif is involved in the sB-Raf•CDC-37 interface in the reference sB-Raf•CDC-37 complex. The other motif is found in the first FKBP-domain of FKB-6 and is crosslinked to CDC-37 (FEAAEPVYMKHYQDEVK (CDC-37, 284) – GNVIKGWDLGVATMTK (FKB-6, 71)) via HSP-90 (KFEAAEPVYMKHYQDEVK (CDC-37, 284) – IMKAQALR (HSP-90, 585)) at the position right next to the sB-Raf•CDC-37•HSP-90 interface. These data reveal that FKB-6 directly participates in the kinase-chaperon interface in the complex.

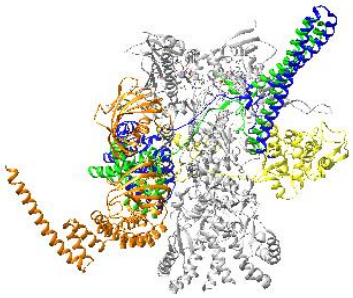
Furthermore, model docking was performed based on LC-MSMS results of the full complex and published kinase•HSP-90 PDB structures. In total nine models were generated based on the PDB structure 5FWL and 1KT1 [129, 130]. The RMSD distribution of all the models was compared to reveal those with the highest level of confidence (Figure 4.4.7 A, where the top two are highlighted with orange circle). The top two docked models with the highest score are shown in the figure 4.4.7 B-C together with corresponding qualifying parameters. Based on these results, electrical density map of the full complex containing the co-chaperone FKB-6 was generated in the figure 4.4.7 D, which directly offers a potential view of the complex assembly, where FKB-6 binds to the middle domain of HSP-90 and reaches out to the interface between HSP-90 and CDC-37•kinase.

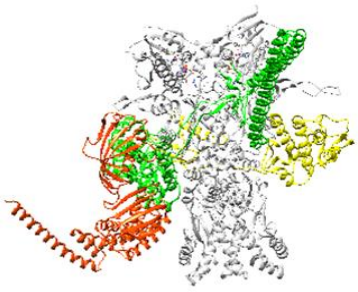
Table 4.1 Intermolecular crosslinked peptide pairs identified in the kinase-chaperon complex via MS (See next page)**.** Peptide pairs identified in the LTQ_replicate are marked in orange and peptide pairs identified in the QE_replicate are colorless.

Peptide1	Peptide2	Proteins
AEEVLEYEKSTDPEK	KPQAPK	FKB-6 (288)-CDC-37 (165)
APFDLFENKK	MAEKK	HSP-90 (326)-CDC-37 (77)
EAELEEKER	KNLVK	CDC-37 (137)-HSP-90 (385)
EKYFEDEELNK	KMEELEK	HSP-90 (254)-CDC-37 (94)
FEAAEPVYMKHYQDEVK	GNVIKGWDLGVATMTK	CDC-37 (284)-FKB-6 (71)
FEAAEPVYMKHYQDEVK	IMKAQALR	CDC-37 (284)-HSP-90 (585)
GVAQKMDEEVFK	MKTLGR	CDC-37 (359)-b-Raf (22)
IGDFGLATVKSR	KMEQEK	b-Raf (184)-CDC-37 (78)
IGDFGLATVKSR	APFDLFENKK	b-Raf (184)-HSP-90 (326)
ITEKKPQAPK	KFYEQFGK	CDC-37 (164)-HSP-90 (407)
KFEAAEPVYMK	IGDFGLATVKSR	CDC-37 (274)-b-Raf (184)
KFEAAEPVYMKHYQDEVK	DSSTMGYMAAKK	CDC-37 (284)-HSP-90 (601)
KFYEQFGK	KPQAPK	HSP-90 (407)-CDC-37 (165)
KHSQFIGYPIK	KMEELEK	HSP-90 (197)-CDC-37 (94)
KMEQEK	SKNSIK	CDC-37 (78)-HSP-90 (329)
KMEQEK	ILKVIR	CDC-37 (78)-HSP-90 (381)
KMEQEK	KNLVK	CDC-37 (78)-HSP-90 (385)
MAEKK	KNLVK	CDC-37 (77)-HSP-90 (385)
MEQEKIDKEK	NSIKLYVR	CDC-37 (83)-HSP-90 (333)
NSIKLYVR	QFFKK	HSP-90 (333)-CDC-37 (273)
NSIKLYVR	KMEQEK	HSP-90 (333)-CDC-37 (78)
NSIKLYVR	IDKEK	HSP-90 (333)-CDC-37 (86)
NSIKLYVR	KMEELEK	HSP-90 (333)-CDC-37 (94)
RAEEVLEYEKSTDPEK	KPQAPK	FKB-6 (288)-CDC-37 (165)
SKNSIK	MAEKK	HSP-90 (329)-CDC-37 (77)
SLNAVATNTTVQKQFFK	APFDLFENKK	CDC-37 (269)-HSP-90 (326)
WKDIEVSDDEDDTHPNIDTPSLFR	KNLVK	CDC-37 (42)-HSP-90 (385)
YFEDEELNKTKPIWTR	KLAAADVTDK	HSP-90 (263)-CDC-37 (101)
YQALTEPSELDTGKELFIK	KLAAADVTDK	HSP-90 (62)-CDC-37 (101)
YQALTEPSELDTGKELFIK	SDIQKQIDEVK	HSP-90 (62)-CDC-37 (115)
Peptide1	Peptid2	Proteins
APFDLFENKK	QFFKK	HSP-90 (326)-CDC-37 (273)
KFEAAEPVYMK	APFDLFENKK	CDC-37 (274)-HSP-90 (326)
KFEAAEPVYMK	IGDFGLATVKSR	CDC-37 (274)-b-Raf (184)
KMEELEK	EIVKK	CDC-37 (94)-HSP-90 (196)
KVEKVGVSNR	AQTKR	HSP-90 (555)-CDC-37 (304)
NSIKLYVR	GTTSKK	HSP-90 (333)-CDC-37 (93)
NSIKLYVR	IDKEK	HSP-90 (333)-CDC-37 (86)
NSIKLYVR	QFFKK	HSP-90 (333)-CDC-37 (273)
QFFKK	SKNSIK	CDC-37 (273)-HSP-90 (329)

A. B. C.







Total score	-90.4 +/- 7.1	-85.3 +/- 10.2
Cluster size	11	29
RMSD from the overall lowest-energy structure	38.0 +/- 1.1	37.9 +/- 1.7
Van der Waals energy	-28.3 +/- 8.5	-33.3 +/- 4.3
Electrostatic energy	-200.1 +/- 40.6	-237.1 +/- 40.6
Desolvation energy	-22.2 +/- 10.3	-4.6 +/- 7.3
Restraints violation energy	0.1 +/- 0.14	0.5 +/- 0.34
Buried Surface Area	1078.3 +/- 190.6	1322.9 +/- 107.0
Z-Score	-1.0	-0.8

D.

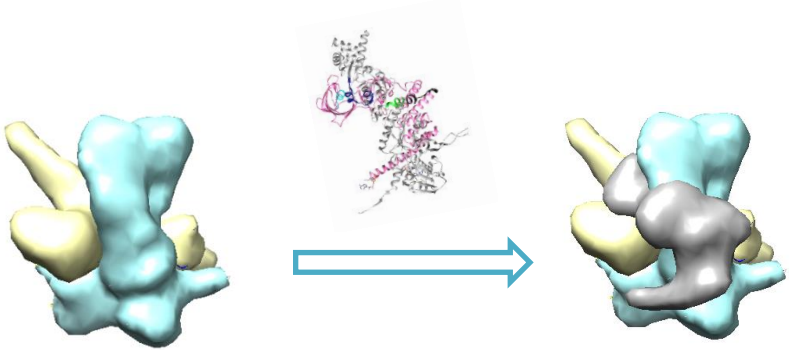
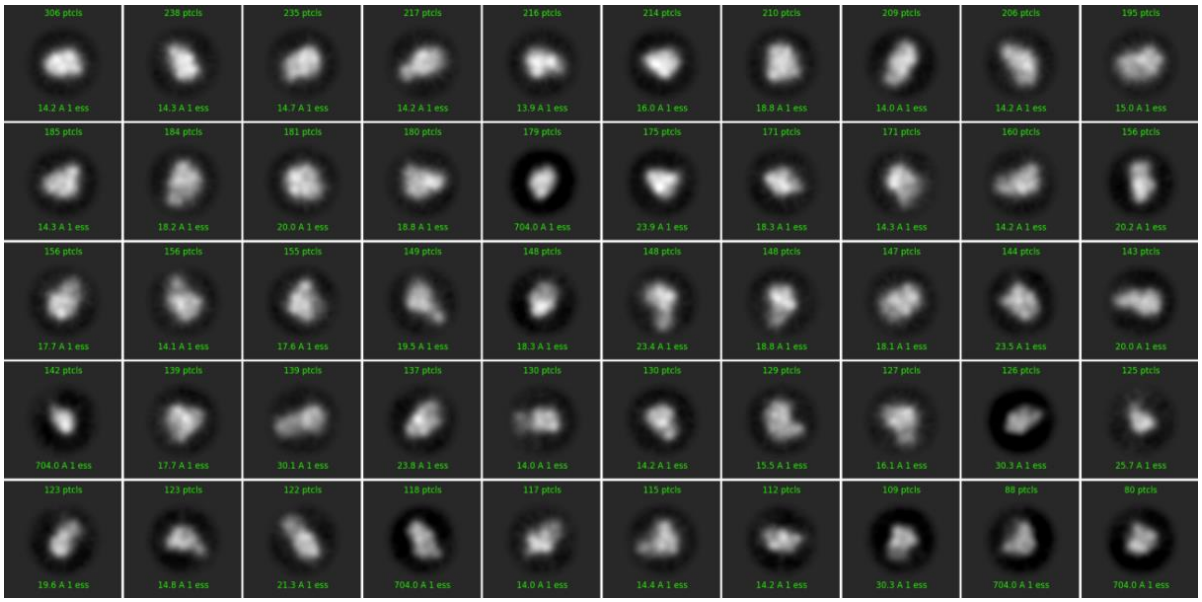


Figure 4.4.7 Model docking of the kinase chaperon complex. A. RMSD of nine docked models. The ones with the top scores (model 8 and 9) are labeled in orange circle. B. Model 8 in A; C. Model 9 in A; D. Association of the structure based on the MS crosslink results. Crosslinked sites are marked in blue and green. Structure is adopted from the PDB structure 5FWL.

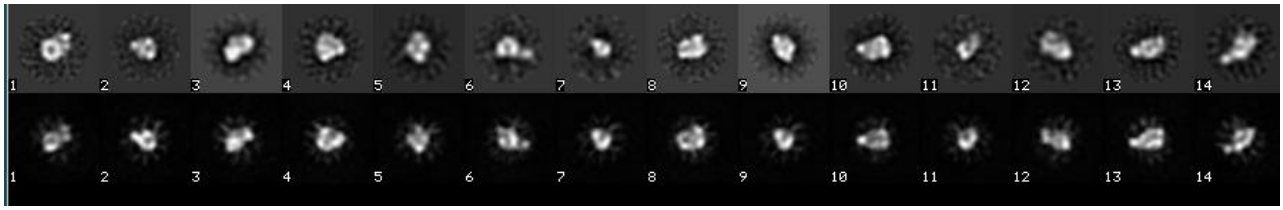
As next the crosslinked complex purified from the HPLC is negatively stained, images of the complex were taken by EM and the structure of the complex is reconstructed based on

micrographs taken from the sample. Starting with around 10 thousands of particles, the best ones were selected for classification. In total 7913 particles were used in the classification and obtained classes are shown in the figure 4.4.8 A. The classes were averaged and used in 3D-reconstruction. The 2D class averages were compared to the reprojection of the reconstructed 3D structure in the figure 4.4.8 B. The selected 2D average images were obtained based on 5177 particles from 7913 filtered particles with the best contrast, which covered 65% of all the particles used in the analysis. In the meantime 3D reconstruction was performed with two softwares based on different algorithms to prove the reliability of the obtained 3D structure. In the figure 4.4.8 C the reconstructed 3D structure from Relion (shown in blue) is compared to the structure obtained from cyrosparc (shown in pink). Both structures possess comparable detailed features at the broader side and the narrower side fits nicely to the well- established HSP-90 C-terminal structures. Single components of the complex originated from the PDB structure 5fwl were isolated and fitted into the Relion generated 3D structure in the figure 4.4.8 D. To date the structure of FKB-6 is still unclear, thus FKB-6 is not fitted in the figure 4.4.8 D. Considering that FKBPs bind to the C-terminal of HSP-90, the empty density close to the HSP-90 terminal region is likely to be FKB-6.

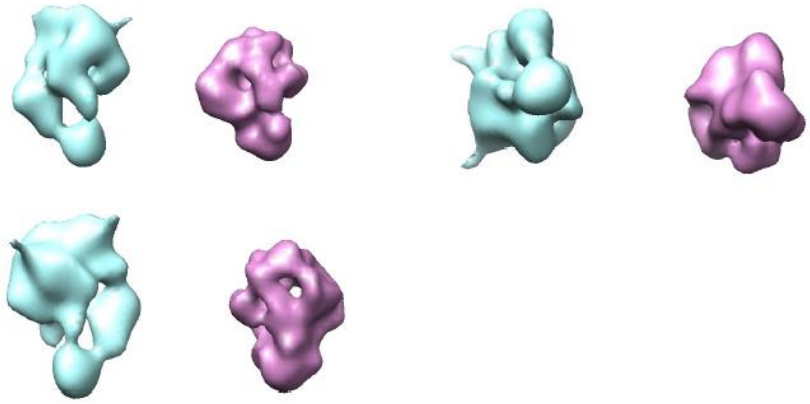
A.



B.



C.



D.

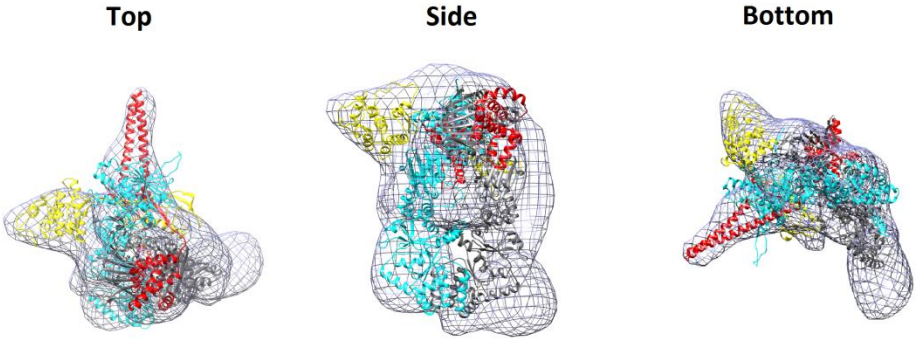


Figure 4.4.8 Structural information of the kinase•Co-chaperon•HSP-90 complex. (A) 2D classes of negative stain images generated by Cryosparc. (B) Comparison of 2D class averages and 3D-reprojection of the obtained structure. (C) The Re-constructed 3D structure at a resolution of 24 Angstrom. (D) Fitted domains in the blue-colored 3D reconstruction of (C), compared to the structure from Verba et.al. [129]. Cdc-37 is colored in red; Kinase domain is colored in yellow; HSP-90 monomers are colored in blue or dark grey.

4.5 Oncogenic sB-Raf mutants alter its kinase stability and the ATP-dependent release from the HSP-90 chaperon system.

It has been reported that the cancerogenic mutant B-Raf V600E requires HSP-90 for its stability and HSP-90 prevents this mutant to form aggregates. On the other side, the typical kinase pathway doesn't always require the assistance of protein chaperons. Mechanisms for the regulation of protein chaperones on the RAF kinase is not well known to date. Considering one of the essential function of HSP-90 chaperone system is to stabilize its client protein and helps it to regain its function, it's speculated that the requirement of chaperone only occurs under certain conditions, when cells are under stress or when mutants require specific support to maintain its stability. In this chapter, by comparing multiple parameters which involve the regulation of kinase properties, key factors so as the role of molecular chaperons in regulating the kinase's stability and activity are discussed.

4.5.1 The stabilizing effect of nucleotides is reduced in cancerogenic mutants.

As important structural units of a kinase domain, the p-loop and the activation segment participate in the activation of the kinase and sequential changes within these motifs could potentially influence the kinase activation [68, 131]. To develop a set of modified sB-Raf variants, several reported oncogenic mutants of B-Raf located within these motifs were selected and used to test the influence of ATP and chaperones on variant sB-Raf kinase. The class II mutants G464E and G469A were chosen to investigate changes of interactions in the P-loop, and the class I mutant V600E and the class III mutant D594G were chosen to test influences of changes in the activation segment respectively (Figure 4.5.1).

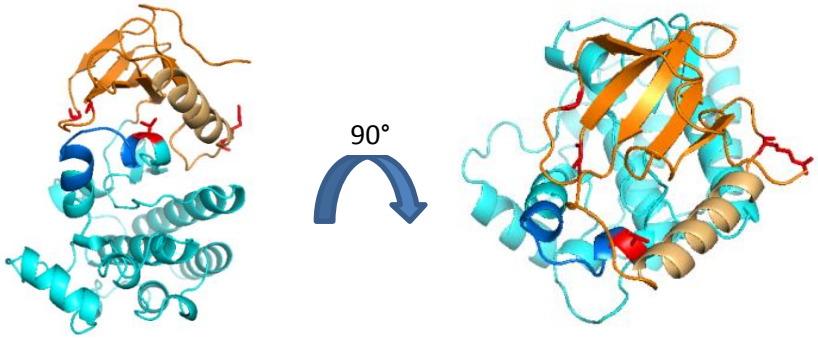


Figure 4.5.1 Structure of the WT sB-Raf kinase domain with positions of mutants indicated.
Left: side view, right top view. Cartoon diagram illustrates the structure of B-Raf kinase domain

based on the PDB entry 4EHE. N-lobe is shown in orange; C-lobe is shown in light blue; Mutants are labeled as licorice in red; αC-helix is colored in beige; Activation segment is colored in dark blue.

Aimed at understanding the influence of mutants on the functionality of B-Raf, changes were tested first *in vivo*. To this end the selected sB-Raf mutants were subcloned into a pcDNA3 vector and fused with an YFP-tag. The protein expression level was determined via the YFP-fluorescence intensity 48 hours post transfection in HEK293 cells (Figure 4.5.2).

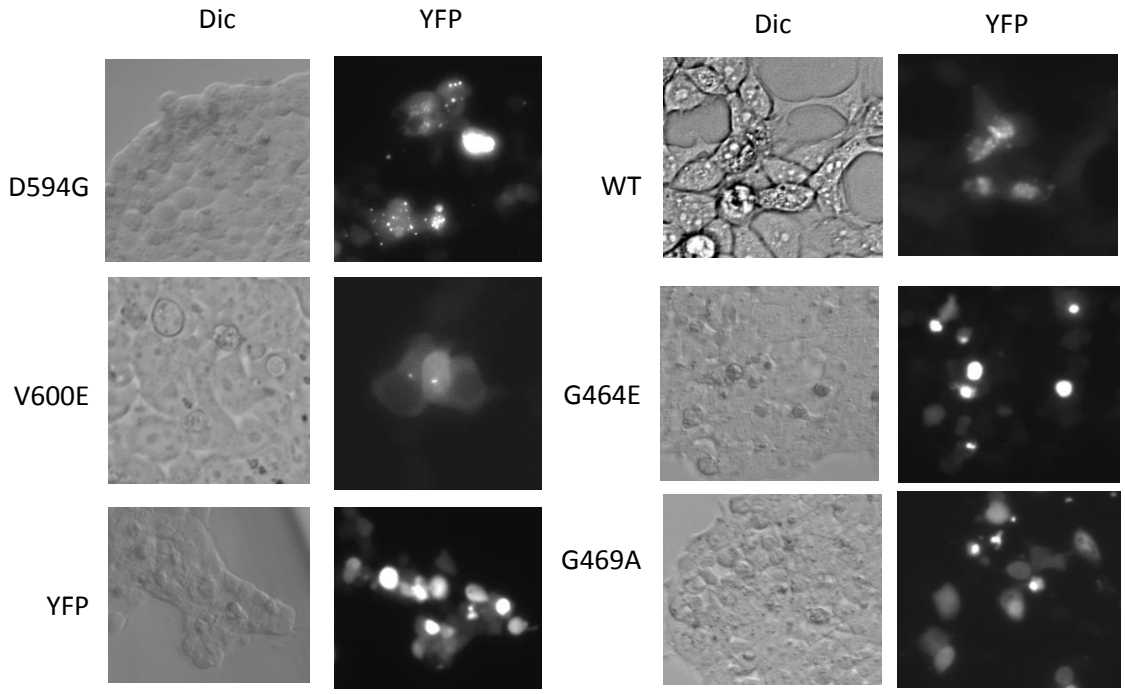


Figure 4.5.2 Expression of B-Raf proteins in HEK cells. Mutated raf-constructs were fused to YFP in pcDNA3 vector. Cells were transfected with corresponding RAF WT or mutants and the fluorescence intensity was measured 48h post transfection. Left panel: normal image of transfected cells; right panel: fluorescence images taken under an YFP filter.

In the cell, WT sB-Raf is expressed mostly in the cytosol and is equally distributed. The mutant G464E and G469A in general have a slightly elevated expression level compared to sB-Raf. All of the three mutants observed are free of any obvious aggregation. The tendency of the mutant G464E and G469A to accumulate in the cell nucleus is slightly increased, resulting in bright spots in the vicinity of the nucleus. Interestingly, the point mutant V600E and D594G form aggregates close to the cell nucleus while the expression level of these mutants is apparently unchanged.

Altogether the behavior of sB-Raf variants differs from the WT protein *in vivo* and the formation of aggregates observed in some of these mutants implies a change in their stability.

To confirm the influence on the protein stability, the mutants were investigated *in vitro*. To this end, the mutants were purified after recombinant expression via His-tag. The stability of sB-Raf WT protein and its variants was analyzed by TSA. Since the very aggregation prone variant D594G could not be purified to sufficient amount, this mutant was omitted from *in vitro* study. WT sB-Raf shows a melting temperature of 41 °C while the melting temperature of the other tested mutants varied in the range of 37-43 °C (Figure 4.5.3).

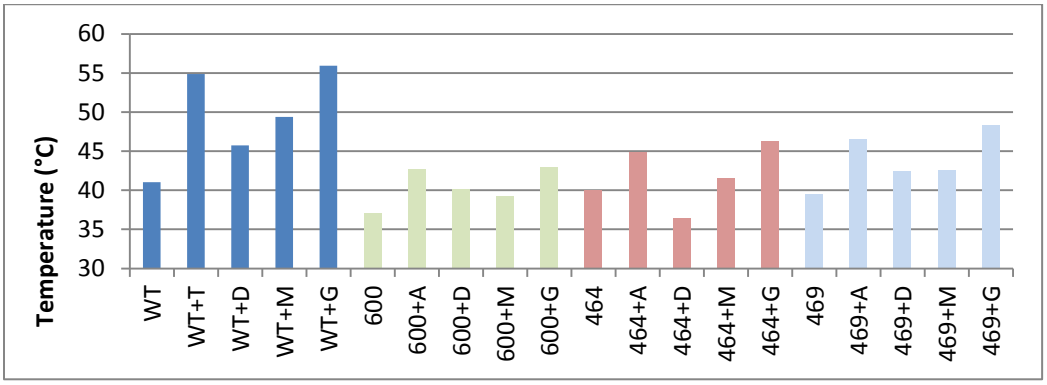


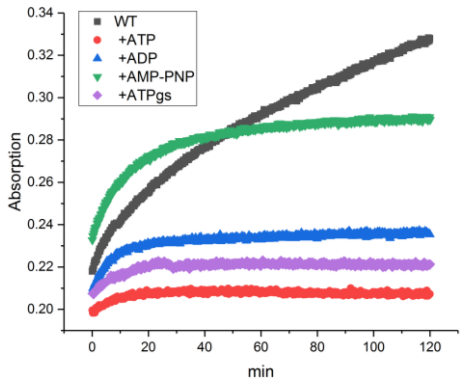
Figure 4.5.3 Reduced stabilizing effect of nucleotides on sB-Raf mutants. Thermal shift assays were performed with 0.2 mg/mL Raf WT kinase domain or with sB-Raf mutants, in the presence or absence of 2 mM nucleotides respectively. T is short for ATP, D: ADP; M: AMP-PNP and G: ATPγs.

It has been observed that the presence of nucleotides stabilizes the sB-Raf kinase domain [37]. Here changes in the stability upon addition of different nucleotides were tracked (Suppl. Figure 5.1 A-D). ATP, ADP, AMP-PNP and ATPγs were tested and all of them influence the stability of sB-Raf WT strongly, as in the presence of each type of nucleotide, the melting temperature of sB-Raf shows a strong increase. With the presence of AMP-PNP and ADP, the stability of sB-Raf WT increased 6 to 9 degrees, while the presence of the ATP and ATPγs generates an increase of 15 degrees and the unfolding temperature reaches 55 °C. In the case of mutants the increase is much less pronounced. The presence of the ATP shifts the measured melting temperature 4-7 degrees higher than samples without ATP, resulting in the range of 42-46 °C, which is almost 10 degrees lower than the effect brought by ATP to the WT. The influence of other mutants is in a comparable level to ATP, namely generating a small increase in the range of 3-6 °C (Figure 4.5.3). The stabilizing effect of ATP & ATPγs is stronger than ADP & AMP-PNP and increased the melting

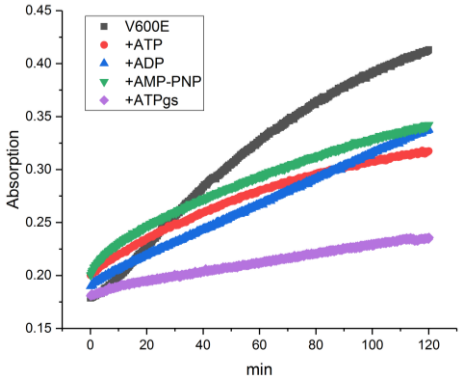
temperature in a range of 2-6 degrees. ATPγs has the highest stabilizing effect, generating a shift of 6-9 degrees in case of mutants. Taken together the ability of nucleotides to stabilize the investigated sB-Raf variants is sharply decreased compared to the WT protein.

To further prove the influence of nucleotides on the stability of the sB-Raf structure, protein aggregation assays were performed with sB-Raf WT and mutants. Aggregation assays were performed in the presence of all four nucleotides (Figure 4.5.4). All measurements were performed under the same condition; however the observed absorbance of mutant V600E is almost twice as high as that for WT sB-Raf. The same trend of aggregate formation can be observed for the other two mutants. The presence of nucleotides efficiently prevents the aggregate formation of WT and mutants. In the WT, the formation of aggregation is almost abolished in the presence of nucleotides. For the sB-Raf mutants, the formation of aggregates is slowed down when nucleotides were added. In the presence of ATP and ATPγs, the formation of aggregates is abolished in the mutant G464E and G469A while in V600E it's slowed down. In the presence of ADP and AMP-PNP, the formation of aggregates is obviously reduced. The efficacy of ATP and ATPγs in preventing the formation of aggregates is obviously higher than that of the ADP and AMP-PNP, which reinforces the conclusion of the TSA analysis.

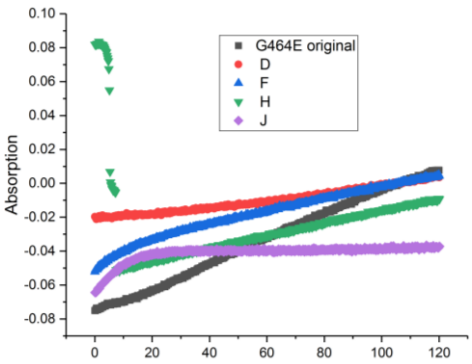
A.



B.



C.



D.

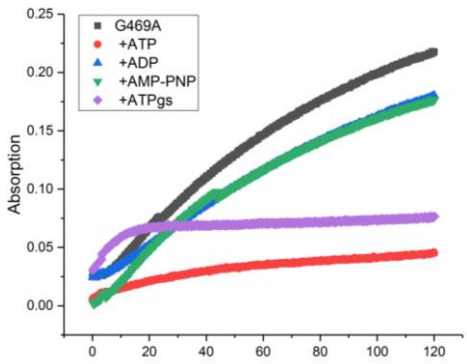


Figure 4.5.4 Aggregation assay of Raf WT and mutants. All experiments were performed under the same condition. (A) Absorbance of 3 µM WT at 360 nm. (B) Absorbance of 3 µM V600E at 360 nm. (C) Absorbance of 3 µM G464E at 360 nm. (D) Absorbance of 3 µM G469A at 360 nm.

4.5.2 Chaperon complexes of mutated kinase are less sensitive to the presence of nucleotides

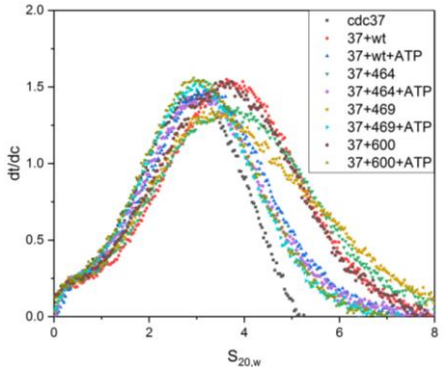
Same as most of the Hsp90-dependent protein kinases, B-Raf requires the assistance of CDC-37 to bind to HSP-90 [65, 132]. It has been reported that protein chaperons antagonize with nucleotides for binding to a kinase [133]. Here the influence of ATP on the complex formation with sB-Raf mutants was observed in analytical ultracentrifugation with ATTO488-labeled CDC-37 (*CDC-37) (Figure 4.5.5 A). In the absence of ATP the binary complex is observed with an $s_{20,w}$ value of 4.2 for WT kinase and for the mutant V600E. Binary complexes formed with the mutant G464E and G469A have a slightly smaller sedimentation coefficient around 3.9, which may indicate a weaker interaction between *CDC-37 and mutants. In the presence of ATP, formation of the binary complex with sB-Raf WT and all mutants is completely blocked by ATP and no binary complex is visible. All of the sB-Raf variants are able to form stable ternary complex with sedimentation coefficients around 6.2 (Figure 4.5.5 B), which were also weakened in the presence of ATP, but full disruption of ternary complexes is not observed, as *CDC-37 still interacts with HSP-90, even though the kinase might be displaced.

It's assumed that the exchange between CDC-37 and nucleotide-binding to sB-Raf is slow. It thus is interesting to analyze whether these mutants influence the assembly of the complexes. To address this issue, crosslinking reactions of CDC-37 and sB-Raf proteins in the presence or absence of ATP were performed (Figure 4.5.5 C). Both sB-Raf WT and mutants are able to form complexes with chaperons in the absence of ATP. The binary complexes split into two bands on SDS gel with a molecular weight difference of 30 kDa, which could imply that a second sB-Raf molecule can be crosslinked to the CDC-37•sB-Raf heterocomplex.

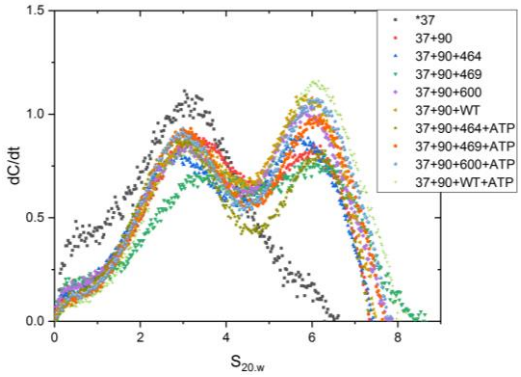
In the presence of ATP, competition between nucleotide binding and CDC-37 binding is observed for sB-Raf WT but much weaker in the mutant complexes. The sB-Raf V600E•CDC-37 complex is barely influenced by the addition of ATP. For the mutant G464E and G469A, the binary complexes were reduced only partially. This implies that the influence of ATP on the complex stability is weakened. As the binary complex formation with mutants is altered, it is interesting to see to what extent the ternary complexes are influenced (Figure 4.5.5 D). For the WT sB-Raf again addition of ATP prevents the formation of ternary complexes. As the mutant V600E doesn't bind to HSP-90 in the absence of nucleotides, the presence of ATP has no effect on its complex formation. Both mutant G464E and G469A bind to HSP-90 and CDC-37 in a

similar manner as sB-Raf WT. In the ternary complex it is nevertheless obvious that the addition of ATP has no detectable influence on the ternary mutant complexes.

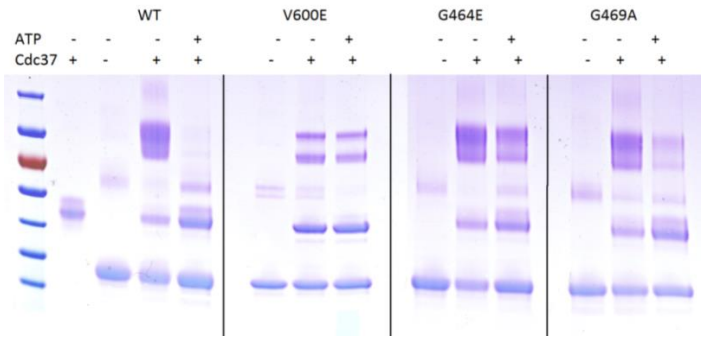
A.



B.



C.



D.

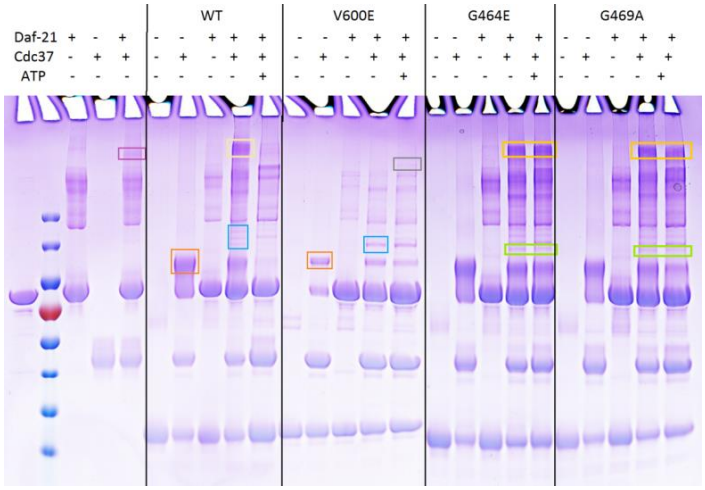


Figure 4.5.5 Formation of kinase-chaperon complex and influence of ATP on its stability. A. Formation of binary complex in AUC at a concentration of 500 nM CDC-37 and 3 µM sB-Raf kinase domain. 2 mM ATP is added to test the influence on the complex stability. B. Formation of ternary complex in AUC measured under the same condition as in A. C. Influence of ATP on crosslink reaction between CDC-37 and sB-Raf mutants. D. Crosslink reaction of the ternary complex and the influence of ATP.

4.5.3 Nucleotide affinity and kinase dimerization influence the kinase-chaperon complex formation

As several observed effects relate to nucleotide binding or dimerization of sB-Raf, further sB-Raf variants were derived from known mechanistic mutations in B-Raf to get more information on these processes. The R509H mutation generates a preferentially monomeric version of B-Raf and the A481F mutation is reported to have sharply lower affinity to nucleotides. These variants should be able to provide mechanistic information on the interaction in the complex. The mutants' effects were firstly tested in thermal shift assay. Indeed, sB-Raf A481F was inert to extra nucleotides since no obvious differences in melting temperature were observed when nucleotides were added (Figure 4.5.6 A). As for the sB-Raf R509H mutation the stability is reduced under all conditions (Figure 4.5.6 B). When investigating the interaction in the binary complex in analytical ultracentrifugation experiments, the behavior of sB-Raf A481F and sB-Raf R509H is comparable in the absence of nucleotides.

Previous results indicate that the ability to form dimers and to bind nucleotides is important for the regulation of kinase•chaperone complex. Thus mechanistic mutations were used to verify these interpretations. The nucleotide inert mutant sB-Raf A481F can be used to distinguish the influence of nucleotides on the complex formation, which due to the mutation in sB-Raf can only influence the Hsp90 protein. Nucleotides have direct impact on the chaperon, which results in a shift in the s-value in AUC measurements. Even with a loss in nucleotides binding activity as in the mutant A481F, complex bands of the kinase with chaperons are still clearly visible as main crosslink products on the gel (Figure 4.5.6 C). On the other side the dimerization state of the kinase seems to play an important role in the complex formation. For the monomeric mutant R509H, association to CDC-37 is much more sensitive to the presence of nucleotides. The amount of formed complex is reduced compared to other mutants or the WT (Figure 4.5.6 D).

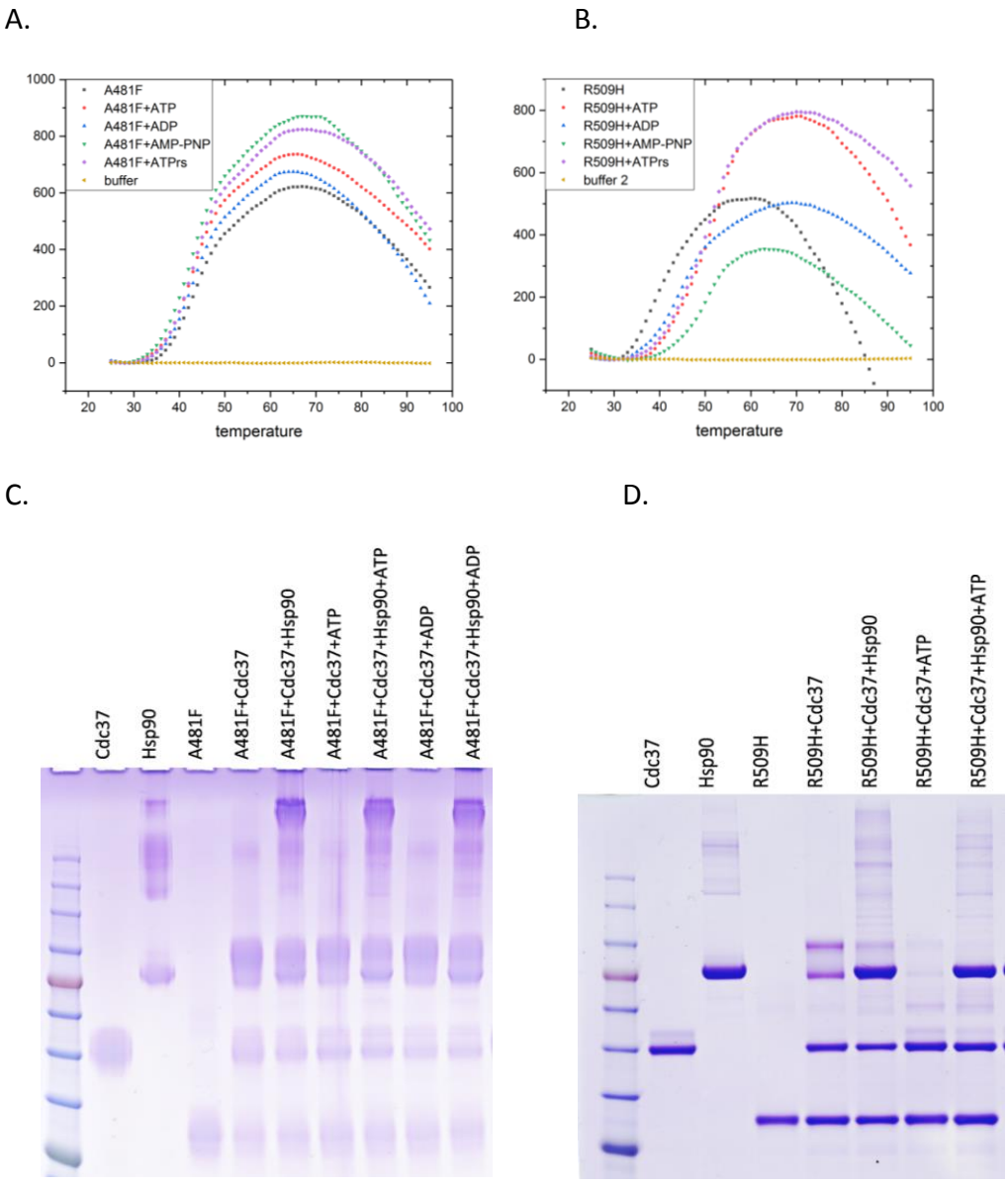


Figure 4.5.6 Regulation factors of the kinase-chaperon complex. A. Influence of nucleotides on the stability of the mutants A481F measured in TSA. B. Influence of nucleotides on the stability of the mutants R509H measured in TSA under the same condition as in A. C. Influence of ATP on crosslink reaction between CDC-37 and sB-Raf A481F. D. Influence of ATP on crosslink reaction between CDC-37 and sB-Raf R509H.

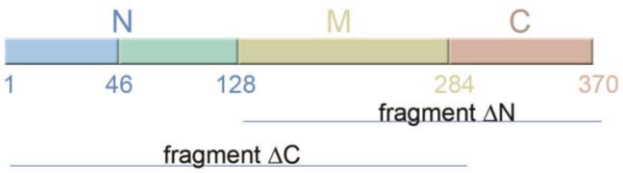
4.5.4 sB-Raf interaction with the N-terminal interaction site of CDC-37 is affected by mutations

Previous studies had shown that the M-domain of CDC-37 interacts with sB-Raf while the N- and C-terminal regions regulate this interaction [65]. Further it was observed that the isolated N-terminal domain (AA1-128) of CDC-37 interacts with kinases and HSP-90, while the C-terminus (AA284-370) does not interact with kinases directly [134]. Knowing the interaction between the sB-Raf kinase domain and CDC-37 is less efficient for mutants; it's interesting to find out whether the sensitive part on CDC-37 can be identified. To address this issue, interaction of the N-terminal region of CDC-37 (AA1-283, ΔC CDC-37) and fragments of CDC-37 lacking the N-terminal domain (AA128-370, ΔN CDC-37) with variant sB-Raf proteins were observed (Figure. 4.5.7 A).

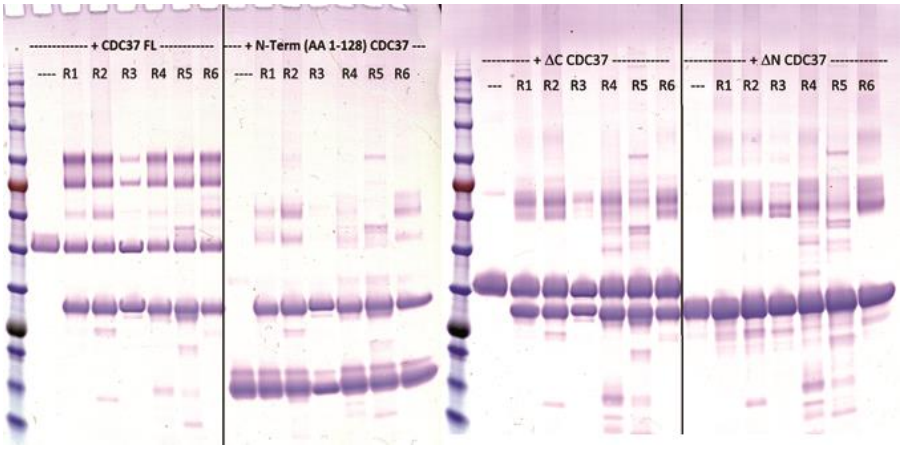
With both CDC-37 fragments the interaction of sB-Raf mutants are decreased compared to the full-length CDC-37 protein (Figure 4.5.7 B). This is likely due to omitting interacting parts that contribute affinity to the complex.

Further the interaction was tested by crosslinking sB-Raf proteins to CDC-37 fragments. Here likewise the crosslinking efficiency is sharply reduced compared to the full-length CDC-37 protein (Figure 4.5.7 B). All sB-Raf variants can form two crosslinked products with CDC-37 differentiate by about 35 kDa in the molecular weight. This implies that sB-Raf can bind either monomeric or dimeric onto the full-length CDC-37. The very N-terminus of CDC-37 (AA 1-128) has the strongest impact on complex formation, while the other CDC-37 fragments form comparable complexes with sB-Raf mutants. In particular, sB-Raf V600E does not interact with the N-terminus of CDC and binds to CDC-37 ΔC- and ΔN- fragments with lower affinity, forming a much weaker band. The amount of crosslink products are qualified with ImageJ and plotted in the figure 4.5.7 C. Interestingly, the mutant A481F and R509H bind on a similar level as WT with full length CDC-37, but the interaction of the ΔC and ΔN fragment of CDC-37 with the mutant A481F is much weaker while the interaction with the mutant R509H gets even stronger. In general all SB-RAF mutants prefer to bind as dimer in the interaction with ΔC CDC-37 and the majority binds only as monomer when faced with the ΔN CDC-37. In terms of the chaperon, the binding of kinase constructs is mainly regulated by the N-term (AA 1-128) CDC-37.

A.



B.



C.

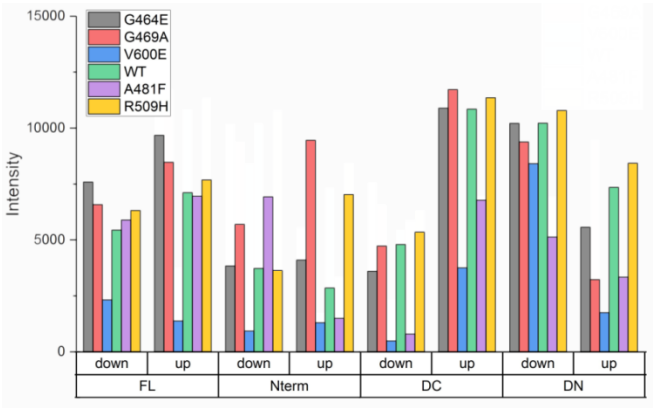


Figure 4.5.7 Interaction of the sB-Raf mutants with CDC-37 fragments. A. Structure of the CDC-37 fragments used in the experiment. B. Crosslink reaction of the sB-Raf mutants with different CDC-37 constructs. R1 = sB-Raf G464E; R2 = sB-Raf G469A; R3 = sB-Raf V600E; R4 =sB-Raf WT; R5 = sB-Raf A481F; R6 = sB-Raf R509H. C. Qualification of the crosslink results via ImageJ. The intensity of the bands is used to present amount of corresponding crosslinked protein. Up indicates the higher band of the crosslink products, which potentially contains two sB-Raf molecules and down indicates the lower band of the crosslink products which potentially contains one sB-Raf molecule.

5 Discussions

5.1 Yeast genome-wide expression data analysis*

*this chapter has been published in microbial cell, 2019

The yeast genome-wide expression level is analyzed by generating co-expression cliques representing the full genome and then used in the statistical performance of those in genome-wide expression analyses. Indeed, in every tested microarray experiment, the enrichment of upregulated and downregulated cliques is observable. This confirms that in all these experiments, even if expression changes are very low, e.g. for the Q30 expression in comparison to the Q0 in yeast, still the expression differences that are significantly above the noise threshold are analyzed.

Application of the a priori clustered genome-wide co-expression cliques

The 72 co-expression cliques representing the full genome allow the visualization of even very small shifts of individual cliques and thus provide significant information that may be lost otherwise. Also the analysis suggests that it's beneficial to analyze repeated experiments also individually without averaging the expression values prior to analysis as the individual analysis may capture the influence of small differences better between these repetitions. This level of detail could also not be achieved, if genes are analyzed only as individual responders. Instead, in context of their clique, the concerted response of the clique enables significance tests on several levels and thus exposes even weak but concerted changes.

Significance here has been achieved without the need to pre-filter the microarray data, without loss of any probes with low signals, without manual readjustments to the cliques and without weighing factors on the probes. While defined parameter settings were used to produce the network (connections with the 10 Top hits from the co-expression ranking) and to excise the cliques from it (protection threshold at 5 genes producing cliques of minimally 6 genes), an iterative optimization procedure could potentially improve the clique set quality further. It's noticeable that producing smaller cliques will impact the ability to obtain significant results as the standard deviation increases strongly for smaller clique sizes (Suppl. Figure 1.4). Thus the *a priori* clustered genome as presented here could be a good resource to perform a fast genome-wide analysis regarding the status of most of the important expression cliques encoded in the yeast genome. To enable general use of this analysis method, the network-files including the analysis method are available at www.clusterex.de.

Beyond the analysis of the expression data, the correlation of the 72 cliques with GO-terms and transcription factors provides interesting information on the molecular events that happen in the yeast cell. Several thousand genes could be directly related to the most prospective GO-terms or transcription factors for the 72 cliques. Potentially uncharacterized genes are assigned within this genome-wide clustering, placing them in cliques with well characterized genes and thereby providing a functional correlation at least in those cliques, where the GO-term assignment is very clear. These assignments are also potentially valuable to target the most prospective transcription factors: at least for the transcription factors IFH1 and GCR1, which are assigned to the clique RPL24A-RPS20A, there is indeed strong evidence that they are involved in regulation of ribosomal protein expression [135, 136].

In general the main results compare well with other enrichment methods, like the gene-set-enrichment analysis GSEA. In comparison to this approach the 72 clique sets developed are static and represent the full genome, with one gene being assigned to one clique. In most cases tested here by us, the GSEA finds the gene sets with the highest enrichment and the top scoring cliques perform comparably. While GSEA is performed on more than 2000 gene sets, the cliques represent the entire genome within 72 cliques and make all its potentially observable responses accessible in a fast and efficient way. Even cliques with lower expression changes, which would show up only after a large number of gene sets when using the entire GSEA gene set database are readily observable here.

Identification of cliques reporting on the presence of toxic and non-toxic polyglutamines

The group of co-expression cliques is applied to visualize and re-evaluate the polyglutamine microarray experiments. PQ56-induced expression changes highlight strongly affected cliques, which contain many of the genes described before. Clearly significant in both experiments is the upregulation of the ARN1-SIT1 clique and the MET10-MET1 clique, which were previously identified as “iron-responsive” and “sulfur-responsive” [94]. The many genes not assigned to specific clusters in the study of the Top 100 genes, now are enriched in the cliques ARG2-ORT1, YGL117W-TMT1, LYS21-LYS12 and SER1-ADE12. Besides, the functions and potential TFs can be assigned to the gene cliques that could not be assigned before. The occurrence of those upregulated gene cliques in two different experiments (Experiment 1: Both samples three days on plate, Experiment 2: Q0 two days, Q56 four days on plates to compensate for the slow growth) implies that this result could be relatively stable over a broader range of incubation conditions. Similarly the downregulated cliques include the VTC1-VTC3 cluster previously assigned as “phosphor-related”, but also here the large number of genes, which previously were called “diauxic shift related” now can be assigned to the three clusters PNO1-TRM2, RPF2-BRX1 and CAT2-IDP2, separating these genes into nucleus, nucleolus and metabolism-related gene

groups. In general this analysis provides a more detailed description, even though this description does not include the fine structure of the genes within the cliques yet. These results indicate that several pathways are affected by the expression of the longer form of the polyglutamine constructs.

For the non-toxic Q30-YFP construct, significant expression differences in the two experiments have been observed as well. In general these yeasts, which are not intoxicated, show much milder expression differences compared to Q56. Furthermore only few clusters show reproducible responses in the two experiments. Nevertheless each of the two experiments produces its own significant response. The strongest overlap between the experiments is the VTC1-VTC3 cluster, which is downregulated in both samples. None appears significantly upregulated in both experiments. This corresponds to the previous analysis based on the Top 100 hits, where the VTC1-VTC3 cluster was identified, but no further significant changes could be extracted from the Top 100 genes. In the response of the cliques as presented here instead, additional cliques are significantly downregulated in both experiments: this is the ZPS1-ZRT1 clique and the PNO1-TRM2 cluster. Given that these cliques reacted twice in the Q56/Q0 experiments and twice in the Q30/Q0 experiments, it may well be systematically affected by the presence of polyglutamine proteins, but due to the very weak expression differences, this will require further experiments. Similarly, the clique GL117W-TMT1 and the clique GDE1-CRG1 are upregulated in both experiments and might be require further reproduction due to their low expression changes.

Based on these examples, this analysis method is generally applicable for analyzing and comparing single experiments regarding their up- or downregulated expression cliques on a genome-wide basis with an ability to detect weak, concerted and reproducible expression changes.

5.2 Consequences of the HSP-90 down regulation *

*this chapter has been published in PLoS One. 2017. PMID: 29078207

Depletion of the HSP-90 by RNA interference leads to various morphological and transcriptional changes in the *C. elegans*. The combined phenotype of the HSP-90 knock-down includes developmental changes to gonad and vulval structures, induction of the heat-shock response, changes in the muscle ultrastructure, as well as the induction of the innate immune response are observed. Based on the function of the HSP-90 as a regulator of signaling kinases and transcription factors, these diverse changes could be caused by multiple HSP-90-dependent processes. Here the whole genome expression data was used to define pathways that are linked to the HSP-90-depletion.

Separate transcriptional pathways are induced by RNAi against HSP-90

To understand the induction of separate pathways, it is important to organize the differentially expressed genes into their respective expression clusters and then attempt guesses on transcription factors for the coexpressed gene groups. With help of the gene clustering approach that based on the use of genome-wide coexpression data, information-rich networks in yeast were generated which clearly separated independent clusters and thus divided hit lists into separate transcriptional units [21]. Potential transcription factors could be derived from lists of target genes as available from YEASTRACT [137]. Appling the same clustering strategy to the multicellular nematode *C. elegans*, information-rich networks are obtained with all tested parameters proving a highly significant network and cluster formation. In this study, the *C. elegans* response to the HSP-90 knock-down can be separated into multiple distinct clusters that can further be defined as independent based on the very low interconnection-numbers. This is supported by the independent microarray experiments, which trigger only part of the HSP-90-depletion response. The GO-term enrichment and individual searches further show that this clustering approach also separates the genes according to distinct functional processes. The reporter strains used in this study validate the organization strategy since genes from the same cluster behave similar regarding their tissue-specific expression and their fluorescence changes after HSP-90 RNAi. By this approach the response to HSP-90-depletion can be separated into one part that reflects a strongly induced innate immune response and into another part that reflects the slower and incomplete development of the nematodes and their reproductive structures.

HSPs are upregulated both on the proteome and transcriptome level

The induction of the heat shock response after HSP-90-depletion has been observed in reporter strains before [138, 139] and confirmed on the protein and mRNA levels. The transcriptional networks overlap with the proteomic response in particular for the heat-shock proteins HSP-16.1, HSP-16.2 and HSP-70. On the protein level the wider chaperone network is elevated, including many of the HSP-90 cofactors such as CDC-37, UNC-45, STI-1, ZC395.10 and FKB-6. Even though hsp-16.2 and hsp-70 are included in the transcriptional networks, the expression changes for most chaperones are barely evident on the mRNA level. This also is true for HSP-90 itself, whose mRNA is only weakly suppressed at the moment of harvesting (log2(DiffExp) =-0.15). This may imply that the chaperone system in the harvested nematodes has already built up a new balance at the beginning of phenotype-development. The heat-shock response certainly is not the strongest response on the transcriptional level at the moment of harvesting as other transcriptional clusters are more prominent among the top 1.2% protein-coding genes.

HSP-90-depletion induces the innate immune response in intestinal cells

The induction of the immune response is much more evident at the harvesting stage. Here the upregulation of certain genes, such as Y41C4A.11 (log2 of 5.69) and Y94H6A.10 (log2 of 2.42) from cluster daf 21 Up_1 (innate immune response/heat-shock response) can also be observed on the proteomic level (log2 of 3.06 and log2 of 3.31, respectively). While Y41C4A.11 is a distant *C. elegans* homolog of the lipopolysaccharide-induced TNF-factor alpha, the function of Y94H6A.10 is unknown but its general coexpression with the heat-shock response and the innate immune response is evident in all experiments of this study. In microarrays and mass spectrometry approaches applied in this study, immune response genes are found to be both upregulated and downregulated. As such, several C-type lectins of the innate immune response are reduced after RNAi-treatment against HSP-90 [105, 110, 140]. In general, most of the proteins found at reduced levels in the mass spectrometry experiments share a functional or coexpression connection to components of the innate immune response. This result holds true at the transcriptomic level, as many upregulated genes from cluster daf-21 Up_1 (C08E8.4, Y47H10A.5, B0024.4, F15B9.6) can also be found upregulated in other cases where the innate immune response is activated.

As expected, the expression in reporter strains for skr-5, clec-60 and dod-24 is most affected by the HSP-90 RNAi in the intestine, the place where most reactions of the immune response take place. It is interesting that several prominent hits of this response group are regulated by the transcription factor DAF-16, a known regulator of the immune response [141-143]. skr-5 and dod-24 are direct target genes of DAF-16, along with hsp-16.2, fbxa-163, C08E8.4, F13D11.4, and Y105C5A.12 [103, 109], which also are among the top induced genes. Given that DAF-16

target genes are well described, whether certain clusters of the networks are strongly enriched for DAF-16 targets was tested, which would suggest that DAF-16 activity is influenced in the HSP-90 depleted nematodes. To this end, a recently published genome-wide ranking of DAF-16 targets was used, which assigned 1663 Class I targets (upregulated by DAF-16) and 1733 Class II targets (downregulated by DAF-16 and its co-regulator PQM-1) [118]. Using this resource, the ranking number for each clustered gene (figure 4.2.6 A) is determined and tested whether any of the clusters show significant DAF-16 correlation. It is evident that cluster daf-21 Up_1 in particular contains genes which are regulated by DAF-16 (Figure 4.2.6 B). Altogether, 71% of the genes in this cluster are assigned as DAF-16 targets, which are significantly enriched over the 18% expected by normal distribution. Clusters daf-21 Down_1 and daf-21 Down_4 also show enrichment of DAF-16 targets, but they are formed from genes repressed by DAF-16/PQM-1 (Figure 4.2.6 C). These data suggest that the DAF-16 pathway is highly active after HSP-90 RNAi. It is interesting in this respect that the induced DAF-16/PQM-1 response differs from the published target ranking, as a group of 12 genes typically repressed by DAF-16/PQM-1 are now activated in daf-21 Up_1. This may suggest that the balanced regulation between DAF-16 and its transcription suppressing co-regulator PQM-1 could be altered by HSP-90 RNAi, leading to the induction of the innate immune response as observed here.

Germline-specific markers confirm effects of HSP-90-depletion on events before fertilization

Germline development is disrupted in HSP-90-depleted nematodes. This is evident from the sporadic failure of HSP-90-depleted nematodes to reach fertility and develop correct gonadal or vulval structures. The germline-specific changes were not visible on the proteomic level, likely because most of these proteins are not among the 400 proteins which could be quantified. However the transcriptional response does show very consistent genome-wide changes. In cluster Daf-21down_1 and Daf-21down_2, many genes have a relationship to germline and embryo development. This matches the disruptive influence of HSP-90 RNAi on development. It nevertheless has to be kept in mind that GFP-expression from genomic integrated plasmid arrays may not entirely correlate with the native context of this protein. Thus, a more detailed analysis of the affected cell types would need to involve detection methods for endogenous mRNA or protein, like fluorescence in situ hybridization or antibody staining. Despite this the tendency is obvious that most tested target genes from this cluster are coexpressed within the gonads and developing embryos. Thus, reporters for the promoter activity and protein localization of cav-1, zip-8 and sepa-1 will be helpful to get information on the delayed processes during development. How these changes originate is still enigmatic but the identification of defined developmental alterations related to the observed phenotype could serve as a starting point to study the failures during development after HSP-90-knock down in more detail.

5.3 Interaction of HSF-1 DBD and regulation of HSF-1's transcriptional activity in *C.elegans*

Hsf-1 is the key regulator of HSR which initiates the elevated expression of HSPs in the cell post heat stress. Most of the studies on the regulation of Hsf-1 mainly focused on the molecular signaling pathway and revealed a strictly regulation by the chaperon system and the post-translational modification. In this chapter the regulation of Hsf-1 activity is observed in terms of DNA interaction and the mechanism of how Hsf-1 disentangles multiple DNA targets is studied.

Orientation and amount of HSE in the promotor region influence the binding of HSF-1

Over 4000 genes have been identified to contain Hsf-1 interacting HSEs in their promotor regions in *C.elegans*, however only 8 of these genes belong to heat shock response genes induced by Hsf-1. Hence how Hsf-1 recognizes and regulates its target genes from the massive HSE containing genome remains elusive. To understand this, HSF-1 interaction with DNA promotors that encode genes either induced by HSF-1 or those are not transcriptional regulated by HSF-1 was compared, where the affinity of HSF-1 to DNA is related to the amount and orientation of HSEs in DNA. Genes that are dramatically induced upon Hsf-1 induction often contain more HSEs in the promotor region. It's known that HSF-1 DBD directly interacts with HSEs, considering that HSF-1 DBD forms a trimeric structure in the active form, it's possible that multimeric HSEs would contribute to a higher affinity to the active HSF-1, which correlates to the observation in the Hsf-1 titration in AUC experiments, where the binding affinity elevates with increasing HSE amounts in DNA promotor regions (BAG-1 vs DNJ13 vs F44E5.4 in the figure 4.3.3). The affinity is further more affected by the relative orientation in DNA sequence. When the DNA promotors contain the same amount of HSEs, the affinity of HSF-1 still could be influenced by short linking regions between HSEs. The closer to each other they are, the more stable is the protein-DNA complex.

On the other side the arrangement of Hsf-1 DBD on DNA is controlled by the nucleotide sequence. The crosslink-MS analysis showed that Hsf-1 DBD binds preferentially in a 'head to tail' form on the sequence 'GAA'. At the converted sequence the orientation of DBD is likely converted as well, which is further confirmed in the modelling approach. The binding to HSE is likely blocked by the presence of N-terminal 82 amino acids in *C.elegans*, which further showed a potential way of regulating HSF-1•DNA interaction.

Modelling approaches predict the binding behavior on trimeric DBD-binding sites.

Interested in the interaction strength as well as the binding orientation, modelling approaches were performed in order to understand the principle of how order and relative arrangement of HSEs in the promotor region influences the interaction with Hsf-1. To this end nematode HSF-1 DBD was generated based on the human homolog structure (PDB entry: 5D5U) and its DNA-binding dynamics with different DNA probes was simulated. With the help of the software FoldX, stability contributions based on the binding mode is calculated. The simulation of the Protein-DNA contact indicates that the interaction mainly takes place between the DNA g2-t3-t4 motif and the protein S68, R71 and Q72 backbone. The DNA base t4 builds specific contacts with the protein S68 back bone. The sidechain of the S68 gets in contact with the DNA g2 as well. In the meantime the side chain of the protein R71 and Q72 get in contact with the DNA base t3 and t4. This indicates HSE in the DNA promotor have certain preferred orientation of Hsf-1 during the binding procedure, namely a head to tail arrangement, which agrees with the crosslink analysis. Other unspecific contact sides have been detected as well, which may support the binding of protein to DNA by reducing its entropy in forming hydrogen bounds with neighboring sidechains or backbones. When the DNA sequence is mutated to AGA, the interaction with Hsf-1 is completely abolished. This proves that the GAA sequence indeed is fundamental for the Hsf-1 binding and all three nucleic acid contribute to the interaction by either directly bind to the backbone of the protein AS68 or form hydrogen bonds with proteins' side chain.

The interaction simulation showed that the modification of the consensus sequence from GAA to AGA results in a markedly higher dissociation constant as indicated by the movement of the mass center of the α-helix away from the DNA double strand. This mutation is introduced into the center position in the dsDNA to test to what extent the introduced change modifies the interaction pattern. The interaction efficiency and cooperativity is reduced in observation if this modification is included in the F44E5.4p promoter sequence. Nevertheless still 5 HSF-1 molecules can bind, implying that the drastic change to DNA consensus sequence can be compensated by cooperativity in the system. This enables a reduced, but measurable interaction even at this position.

These results indicate that the interaction between dsDNA and HSF-1 is governed by the affinity towards GAA triplet and by the cooperativity resulting from the binding of multiple HSF-1 molecules to the same strand of DNA.

5.4 Cochaperon FKB-6 contributes to the stabilization of kinase•HSP-90 complex

HSP-90 together with its cochaperons assists its client in the folding and activation. In this chapter identification of other HSP-90 cochaperon beyond the well-known CDC-37 that involved in assisting client in forming complex with HSP-90 is described. Besides that the influence of nucleotides on the kinase client of HSP-90 also has been observed.

Nucleotides and protein chaperons antagonize against each other in stabilizing sB-Raf kinase.

It is a well described feature of HSP-90 to interact with nucleotides and its cochaperons. In the context of kinase cycle the importance of this interaction is far less clear than in the context of other HSP-90 clients like the steroid receptors, where ATP-binding and hydrolysis by HSP-90 contributes to the generation of steroid binding competent receptor. The data presented here give a glimpse on the importance of nucleotides and cochaperons for protein kinases in HSP-90 complexes.

Interestingly, as long as the HSP-90-kinase complex is at low affinity, the nucleotide binding to kinase disrupts the complex formation with CDC-37 or even with CDC-37•HSP-90. The formation of binary complex can be prevented by adding nucleotides prior to chaperons. On the other hand, when the binary complex is pre-formed, extra addition of nucleotides is still able to trigger disassociation of the binary complex.

The impact of nucleotides is getting weaker in the ternary complex with HSP-90. Even when ATP is present, HSP-90 managed to bind to sB-Raf via CDC-37, same as the case for AMP-PNP and ATPγs. Only in the presence of ADP, where HSP-90 is likely kept at the open conformation, the formation of ternary complex is blocked.

FKB-6 stabilizes the kinase•CDC-37•HSP-90 complex with antagonistic binding to ATP.

In all of the tested HSP-90 cochaperons, FKB-6 is the only one which is able to bind onto the ternary kinase•CDC-37•HSP-90 complex. This is a very interesting observation as FKB-6 is reported to predominately locate in neural cells and has a lower expression level in hypodermal and somatic tissues [124], where Daf-21 is predominately expressed. Hence FKB-6 is unlikely to fulfill the same function as the Cyp40 or other FKBP homologs. For the first time it is observed that FKB-6 participates in the kinase•HSP-90 complexing *in vitro*, where conformational changes of the whole complex is potentially inducible by the presence of ATP.

If ATP is added to sB-Raf•CDC-37•FKB-6•HSP-90 complexes, instead of triggering complex disassociation as observed in the complex without FKB-6, it owns an increased sedimentation coefficient. This implies that with the presence of FKB-6, the HSP-90 machinery is capable of closing and the closed, twisted HSP-90 complex induced by ATP sediments faster than the more open, semi-closed ring-like structure without ATP. Thus FKB-6 may be critically required to reach an efficient HSP-90 ATPase cycle for kinases. This feature is shared mostly by Fkbp51 and Fkbp52 however is abolished by AIP.

Since both Fkbp51 and Fkbp52 contain two Fkbp domains and AIP contains only one Fkbp domain, whether individual Fkbp domains play a different role in stabilizing the kinase complex becomes an interesting question. To address this issue FKB-6 structures with deleted Fkbp domains were generated and the corresponding interaction with the sB-Raf•CDC-37•HSP-90 complex has been observed in AUC. With the deletion of Fkbp domains a similar level of complex formation compared to WT FKB-6 was observed. A clear shift in the s-value indicates successfully assembly of FKB mutants onto the sB-Raf•CDC-37•HSP-90 complex. However this is strongly affected by the presence of nucleotides, which interferes the association of FKB mutants. This implies that the cooperation of Fkbp domains in FKB-6 is necessary in stabilizing the kinase•HSP-90 complex.

Even though the full-length FKB-6 is able to stabilize the complex in the presence of ATP, lacking one of the Fkbp domains clearly arises difficulties to keep the complex at a stable stage. One possible mechanism could be that the binding of kinase client to HSP-90 via CDC-37 induce changes in HSP-90 conformation, which favors further FKB-6 binding. The presence of FKB-6 in contrast, binds contact with sB-Raf and keeps it stable in the complex, likely via binding of Fkbp domains to sB-Raf kinase.

Deletion of the first Fkbp domain seems to induce a bigger loss of sB-Raf in the complex. The presence of ADP does not induced the loss of bound sB-Raf in the complex, instead caused a shift towards larger s-value with WT FKB-6, which indicates HSP-90 in the complex is at a more compact conformation. This cannot be observed in the complex of FKB mutants, especially for the FK2 mutant. The first Fkbp domain affects the affinity of sB-Raf and the second Fkbp domain has an influence on HSP-90 in the complex. The complex formation is measured via labeled CDC-37 as well, which gives a comparable result (Suppl. Figure 4.6).

Structural information indicate Fkb-6 supports the client transfer in the complex

Several cochaperons of HSP-90 are reported to build direct contact with HSP-90's client to transfer it to HSP-90 while other cochaperons participate HSP-90's cycle without directly contacting HSP-90's clients but rather by adjusting the ATPase activity or inducing certain

conformational changes. How FKB-6 is assembled within the complex and whether it has direct contact with B-Raf or has an impact on HSP-90's function, conformation or activity remain elusive.

Intermolecular crosslinked peptides hits identified in MS offer a clue on the structural information of the kinase-chaperon complex. The majorities of crosslinked peptide pairs identified in the complex link CDC-37 and HSP-90 together. Contacts between HSP-90 and sB-Raf are found under the assistance of FKB-6. The comparison to reference samples showed that during the complex formation, the structure of HSP-90 dimer is rearranged to adopt the newly arrived co-chaperon-kinase complex by hugging it with the M- and C-domain. This is confirmed by the disassociation of three previously self-crosslinked motifs in the HSP-90 dimer sample and the formation of new interface between kinase and CDC-37 in the complex as well as plenty of new CDC-37•HSP-90 contacts. Differences of the crosslinked partner in the reference vs the complex sample indicate a transfer of B-raf from CDC-37 to HSP-90 under the help of FKB-6. This may offer a new aspect of the function interpretation of Fkb-6 that FKB-6 is involved in the kinase client transferring procedure where it holds the kinase at the HSP-90 M- and C-domain with its first Fkbp domain and the TPR region, thus stabilizes the complex by reducing its flexibility and entropic energy.

The most attractive knowledge delivered by crosslink results is the behavior of Fkb-6 in the complex. Interestingly, compared to the revealed kinase-HSP-90 structure based on electron microscopy (PDB: 5FWL, [129, 144]), two of the hotspots found in this structure are also detected as crosslinked pairs in this study. One is the CDC-37 M/C motif which gets in contact with HSP-90. This is supposed to be the main binding side of HSP-90 and B-Raf kinase. The exact motif is found to be crosslinked to HSP-90 in the MS analysis. The other one is a HSP-90-kinase interface with the kinase N-lobe, where the protection of a kinase flexible region is achieved through the interaction with HSP-90. HSP-90 regions were found to be linked to B-sB-Raf, CDC-37 and FKB-6, which implies that these motifs together form the quaternary interface of the kinase•chaperon complex.

On the other side the model docking delivered a much clear and straight forward overview of the complex assembly. When fitted to the model structure, FKB-6 wraps around HSP-90 (Figure 4.4.7 D). Those two crosslinked site in FKB-6 offer a clue to the role of FKB-6 in the complex, which is to assistant the transfer of kinase from CDC-37 to HSP-90 as well as stabilizing the whole complex by linking all components together.

Molecular Dynamics simulations indicate increased flexibility in the mutated kinase domain

Knowing that mutants of sB-Raf behave differently in interacting with chaperons, the mechanism behind this is explored by dynamic simulation. To this end, simulations of sB-Raf kinase domain mimicking its active (PDB: 4MNE) and inactive state (PDB: 4EHE and 4MBJ) were performed. As a control the B-factor of the simulated WT and the mutant V600E are calculated and plotted in the Suppl. Figure 5.2 A. Aside from the relative flexible terminal region, the strongest structural changes were mainly observed at three positions: the alpha-C-helix, the activation segment and the flexible region between the N- and C-lobe.

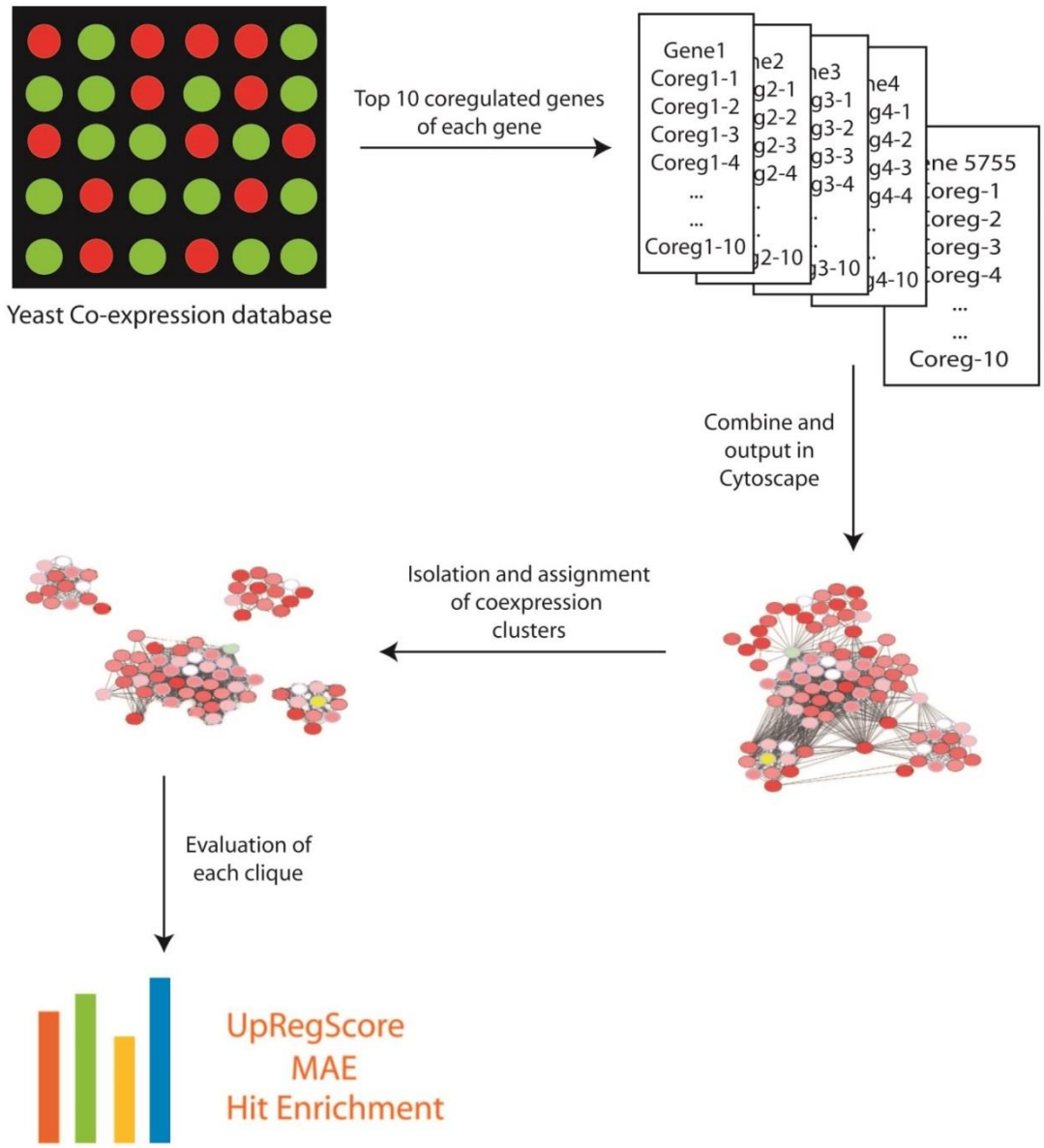
In the simulation, regions that show clearly differences in the active/inactive state were mainly focused. The beginning step of the kinase activation is ATP binding, which requires opening of the ATP-binding cleft in the kinase domain. The ATP-binding cleft in the kinase domain includes a flexible region between the N-and C-lobe, thus the movement of the region induces changes in the relative volume of the binding pocket. Suppl. figure 5.2 B shows the diameter of the ATP-binding pocket in kinase domain. Here clearly two state of the binding pocket are visible, namely a relatively opened state with a diameter of 43-48 Å and a relatively closed state with a diameter of 38-41 Å. A shift of 6-7 Å from inactive conformation to active conformation is observed. The ATP-binding step is followed by the conformational rearrangement of the activation segment, which occurs at the α-C-helix and the activation loop. In the Suppl. Figure 5.2 C the distance between alpha-C-helix and the activation loop was used to represent open and closed conformation of the catalytic cycle (Suppl. figure 5.2 C). The loop is more opened in the active state and a shift around 5 Å is clearly visible in analysis. In the active state, the kinase activation loop domain prefers the opened state and in the inactive conformation it prefers the closed state, which has a destabilizing effect on the molecule.

In mutants the simulation was performed both in active and inactive mode and differences are compared (Suppl. figure 5.2 D-E). In terms of the ATP-binding pocket, it's observed that in the active form, the mutant G464E showed a similar pocket size as the mutant G469A in a range of 35-43 Å; the mutant V600E however has a smaller binding pocket, which is in the range of 25-40 Å with a clearly two-stage distribution. In WT the binding pocket size measured is in the range of 32-38 Å, which is comparable to the size of the mutant V600E. In the inactive state, all the kinase constructs have an ATP-binding pocket size of 28-34 Å while the mutant V600E opens even further and ends at a range of 36-43 Å. In terms of catalytic form formation, the distance between the active loop and the α-C helix varies in mutants. In WT the distance is in the same range as in the mutant V600E (30-35 Å). However in the mutant G469A it increased to 33-43 Å. The mutant G464E has the largest distance which is in the range of 36-46 Å. In contrast, in the inactive state of kinase domain, the distance measured for the mutant G464E and G469A are similar to the distance in WT, which is in the range of 43-50 Å. However the distance in the mutant V600E has no obvious change in comparison to the active state, it still showed a

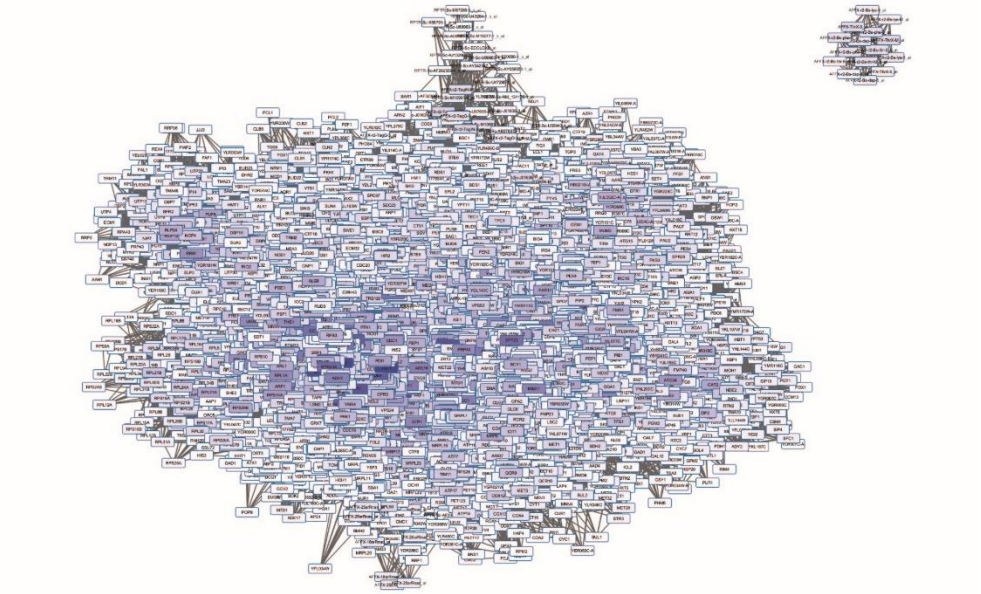
distance between 30-36 Å. Taken together, the mutant V600E is active in both active-mimicking and inactive mimicking state. The other two mutants showed potentially elevated activity in the simulation in the active mimicking state and comparable activity to WT in the inactive mimicking state.

6 Supplemental figures

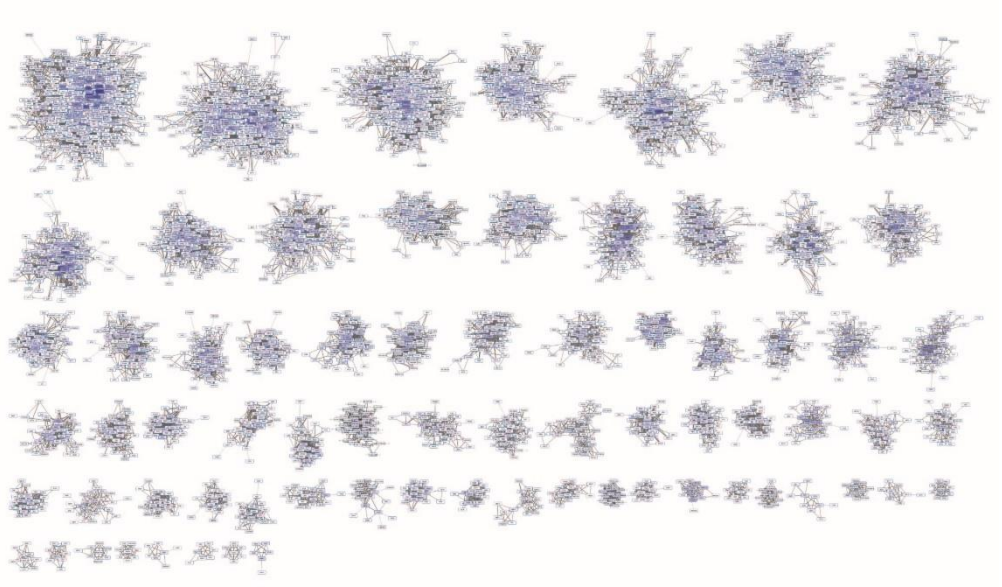
Suppl. Figure 1.1 Flow diagram of clusterEx. Edges in the network are generated by connections within the co-expression database. For each gene the Top 10 co-regulated genes from the database are used to generate 121 edges and 11 nodes. Edges are summed up and the resulting matrix file is used to draw the network. It further is used to cut the network in the more dense modules by preferentially maintaining the edges, which have higher numbers of occurrence as described in Material and Methods.



Suppl. Figure 1.2 Cluster of the entire yeast genome as represented on GeneChip Yeast 2.0 arrays based on 3120 microarray experiments and the observed co-expression relationships.

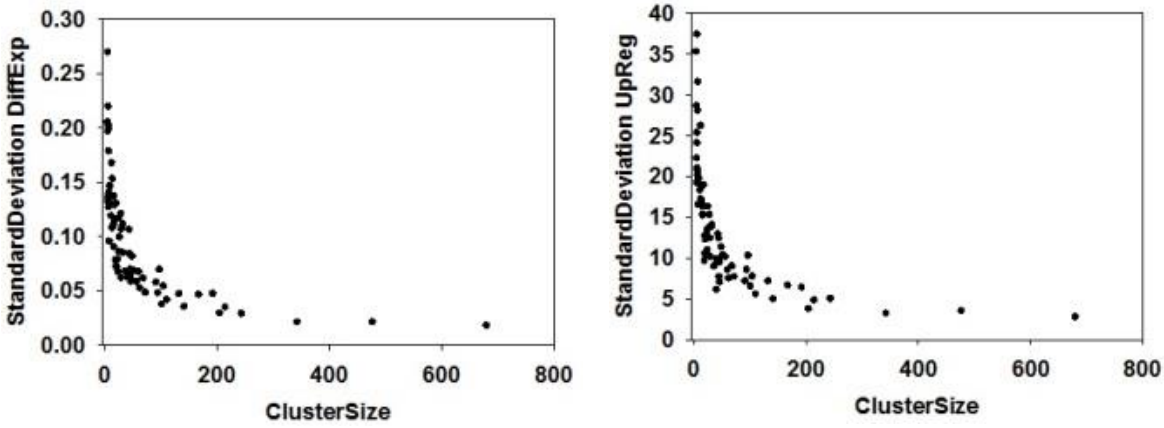


Suppl. Figure 1.3. Clusters separated from the network in supplemental figure 2. The bluish color shows the number of connections within the genome-wide cluster for the respective gene. This value varies between 8 for the Affymetrix Probe RPTR-Sc-A00196-1_s_at (white) and 3328 for TEF1 (blue). Clusters are arranged according to size and numbers will be assigned to them.



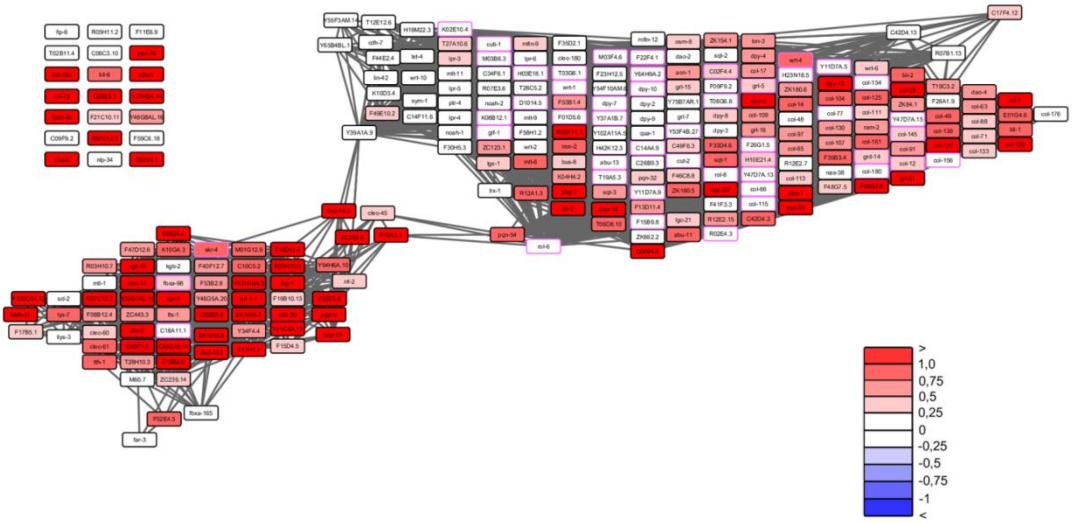
Suppl. Figure 1.4 Plot of the standard deviation of of the clusters. (A) The averaged expression changes in random experiments against the size. (B) The standard deviation of UpRegScores in random experiments. The relationship is as expected and fits to STD_0/sqrt(ClusterSize) are reasonably good for ClusterSize > 10.

A. B.

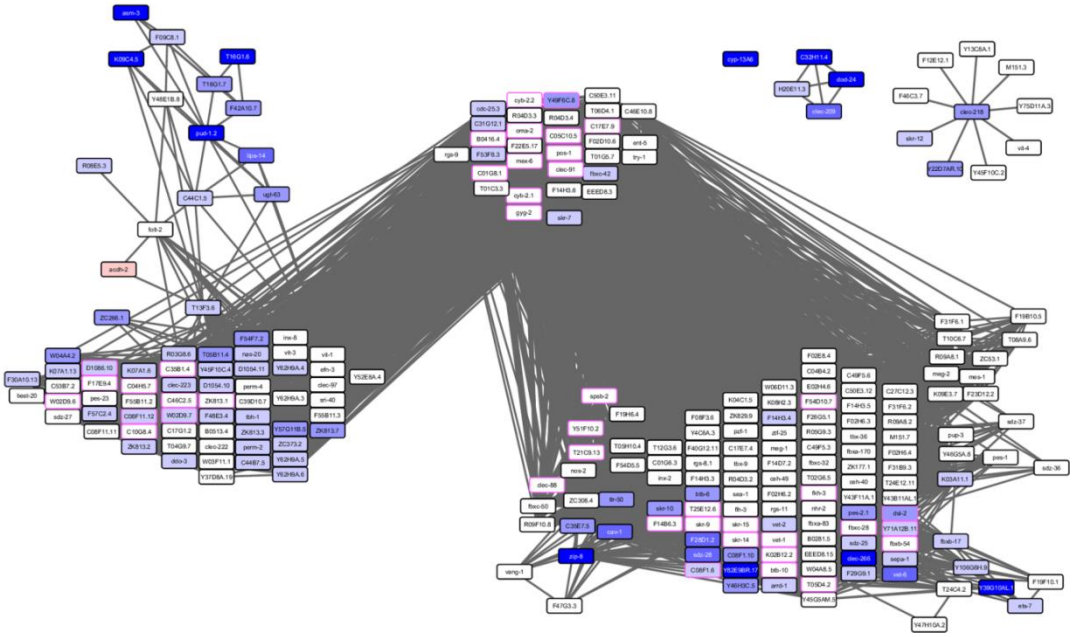


Suppl. Figure 2.1 Up- and downregulation of the HSP-90-responsive clusters in the first RNAi experiment. Blue indicates the different levels of downregulation, shadings of red highlight upregulation.

A.



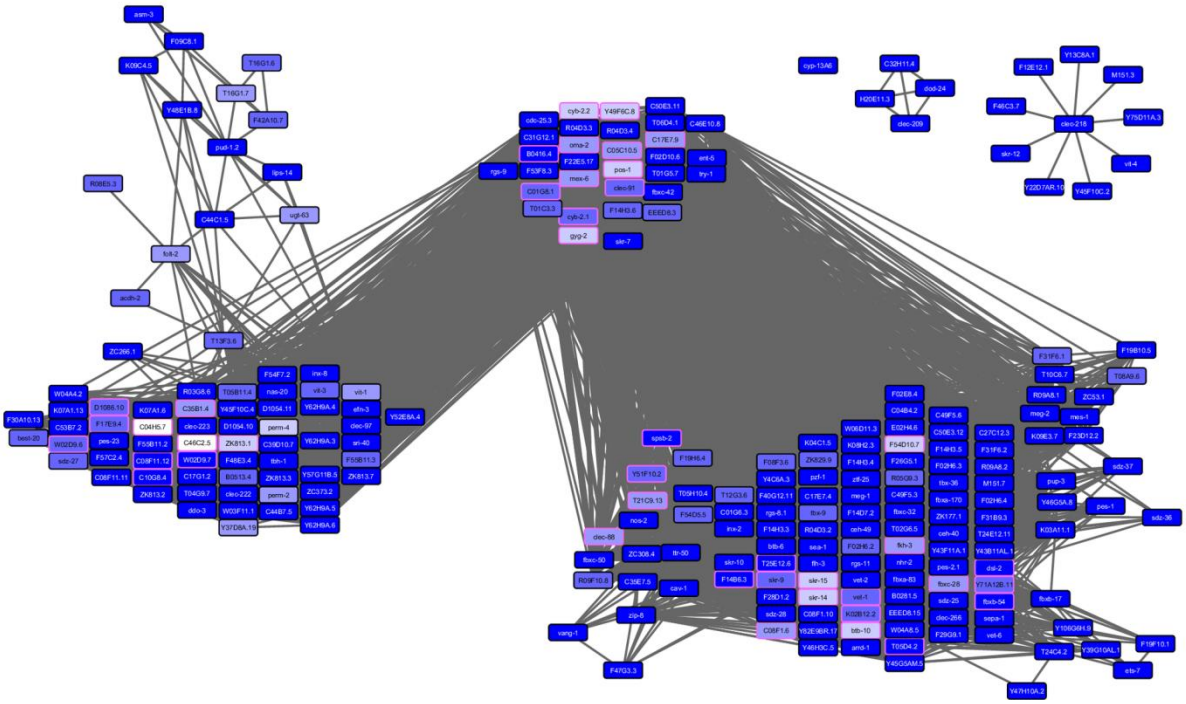
B.



Suppl. Figure 2.2 Up- and downregulation of the HSP-90-responsive clusters in the second RNAi experiment. Blue indicates the different levels of downregulation, shadings of red highlight upregulation.

A.

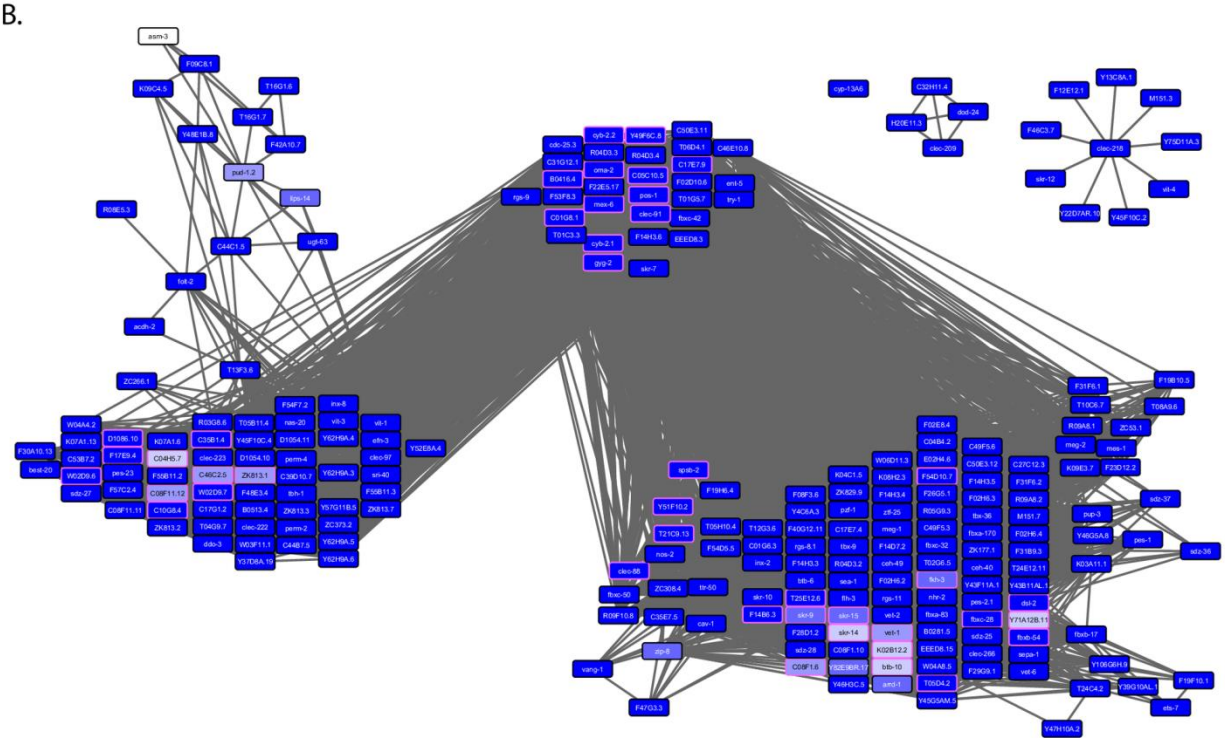
B.



Suppl. Figure 2.3 Up- and downregulation of the HSP-90-responsive clusters in the third RNAi experiment. Blue indicates the different levels of downregulation, shadings of red highlight upregulation.

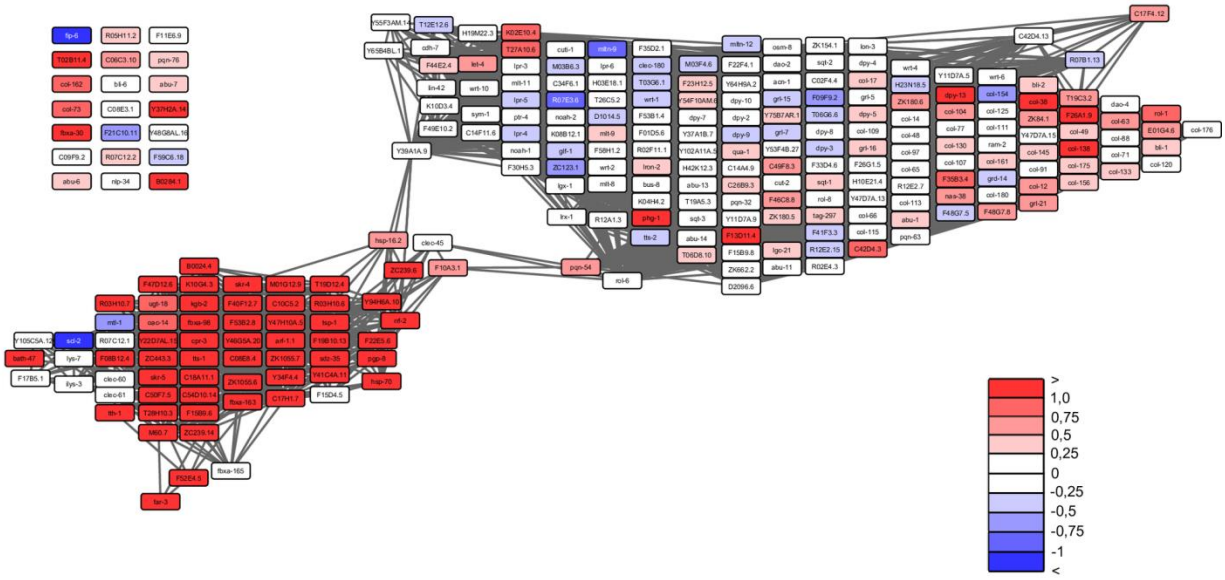
A.

B.

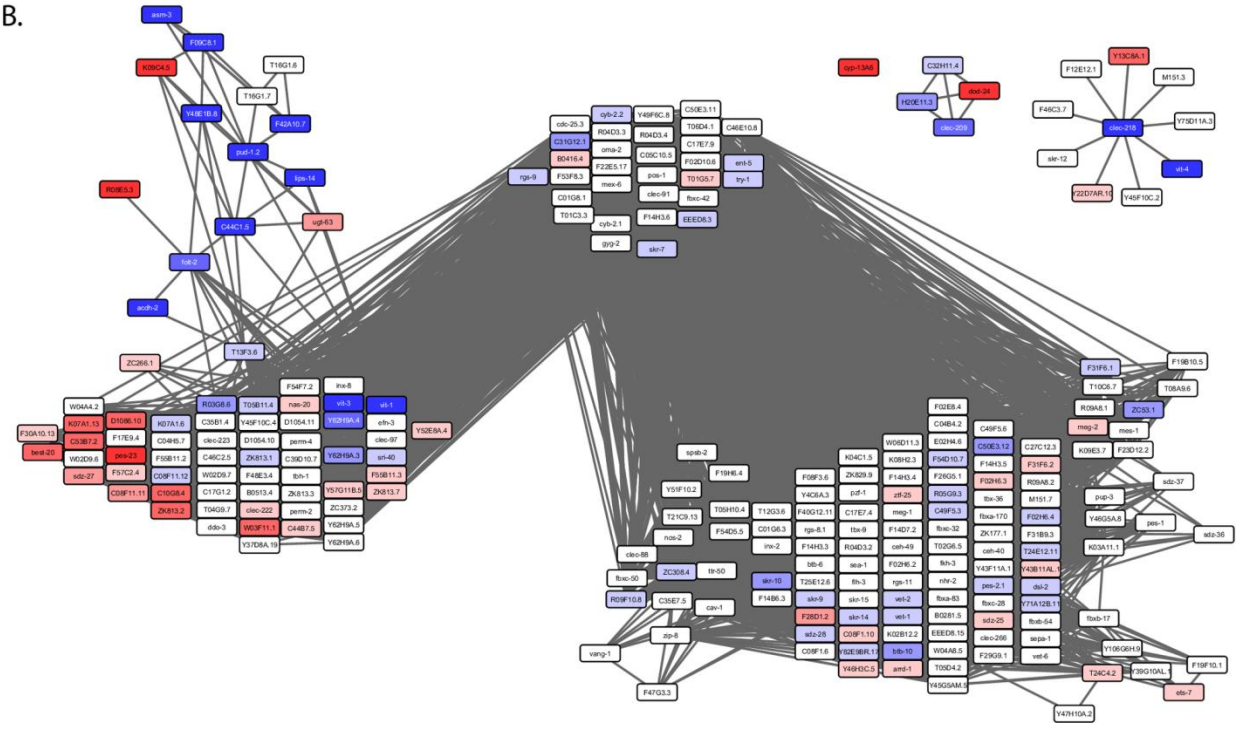


Suppl. Figure 2.4 Up- and downregulation of the HSP-90-responsive clusters in the RNAi experiments with depletion of *sams-1*. The three replicates of this experiment were averaged to obtain the depicted expression values. Blue indicates the different levels of downregulation, shadings of red highlight upregulation.

A.

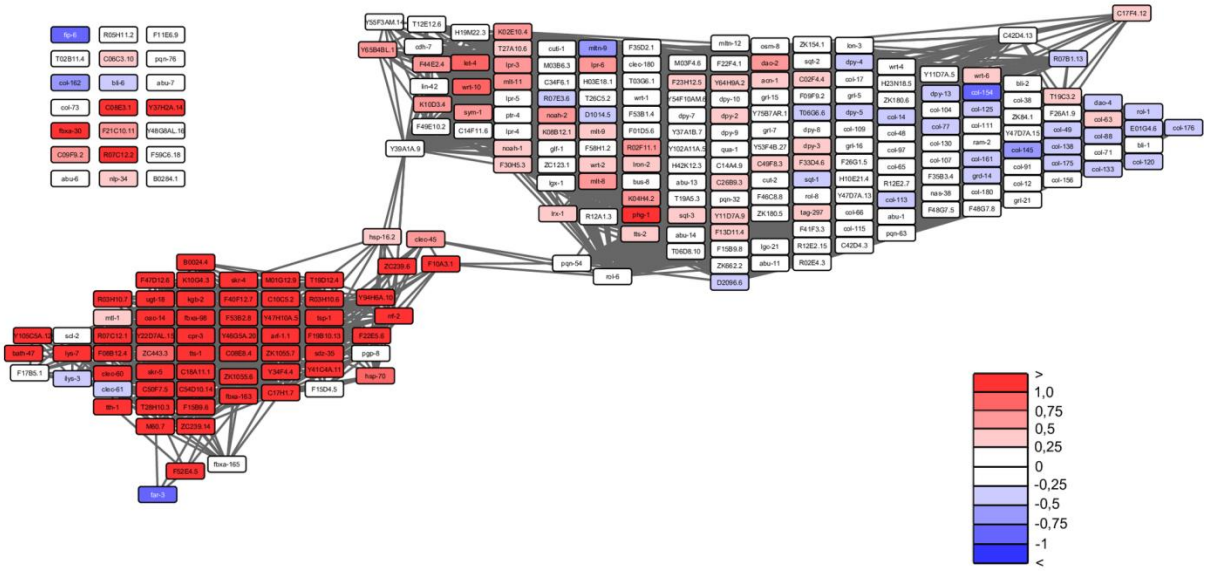


B.

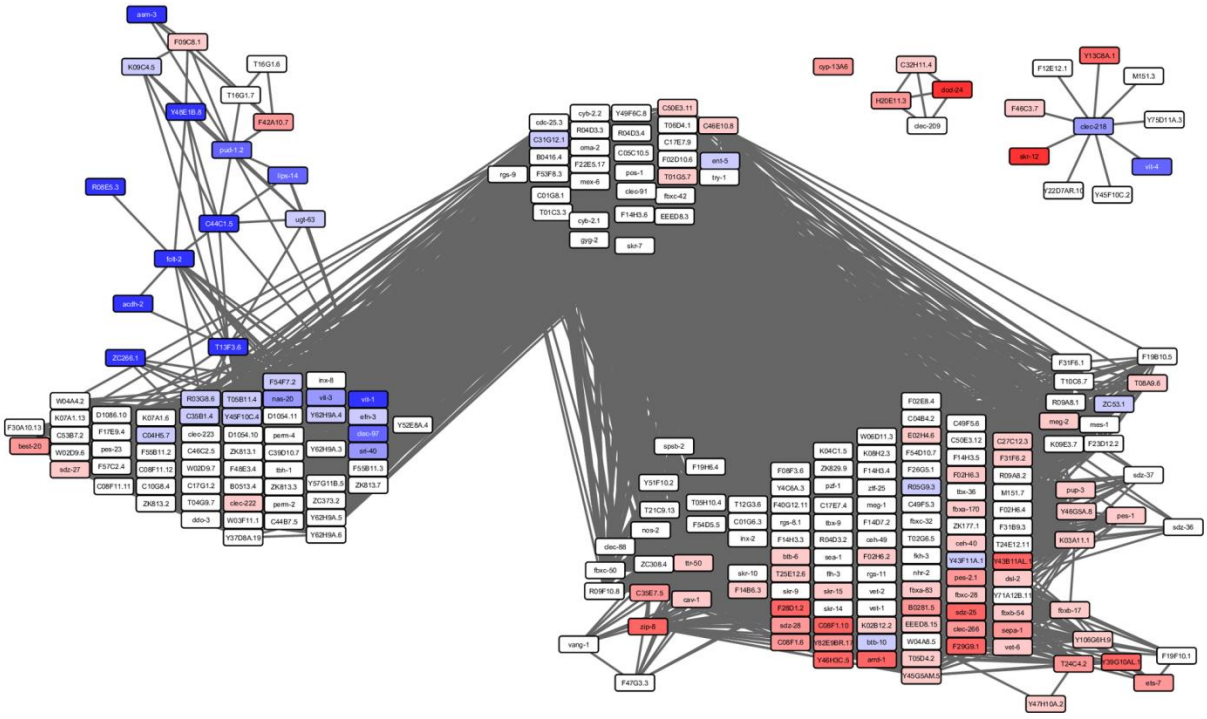


Suppl. Figure 2.5 Up- and downregulation of the HSP-90-responsive clusters in the RNAi experiments with depletion of *sbp-1*. All three replicates of this experiment were averaged here. Blue indicates the different levels of downregulation, shadings of red highlight upregulation.

A.

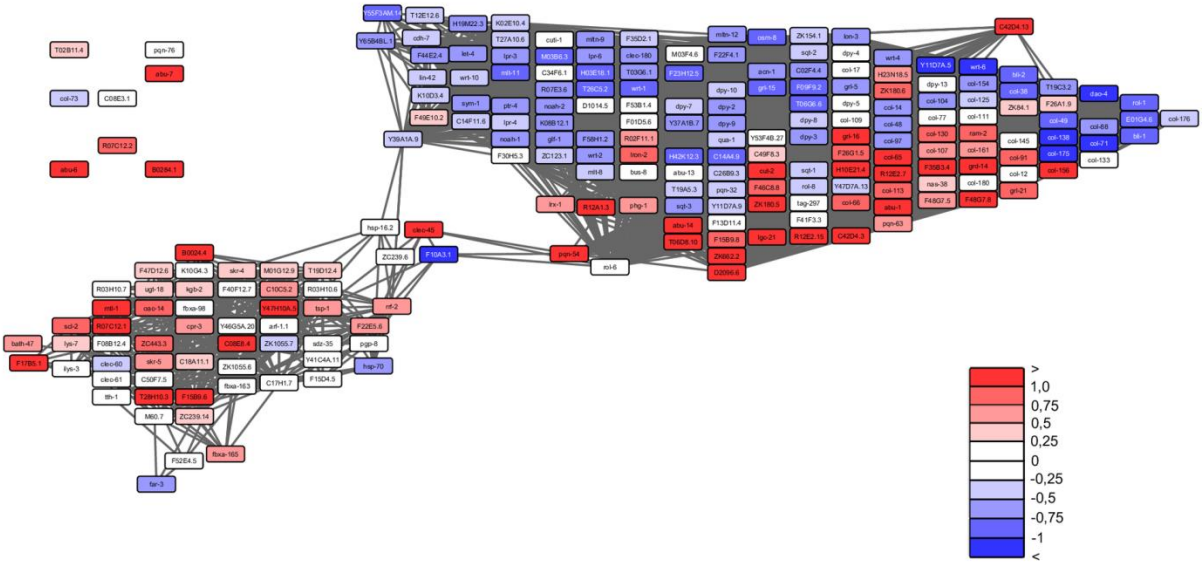


B.

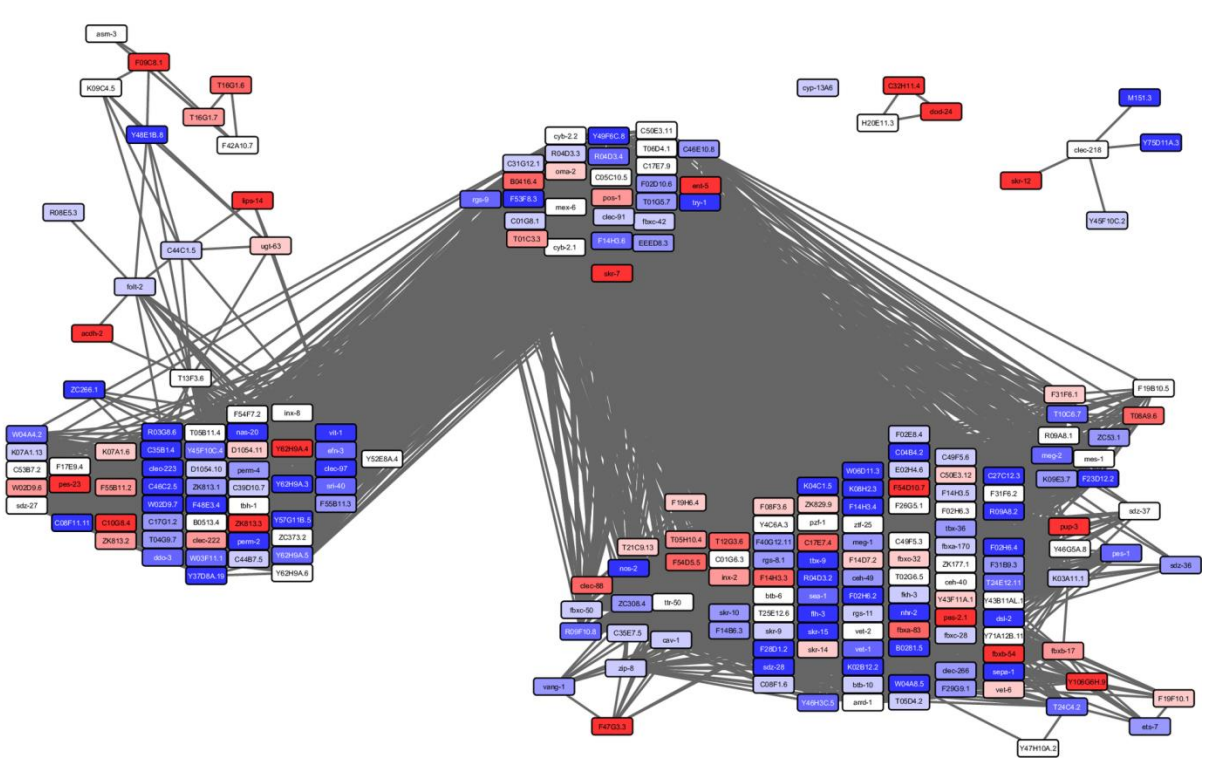


Suppl. Figure 2.6 Up-and downregulation of the HSP-90-responsive clusters in the response to *Vibrio cholerae*. Strain VC109, which induces the immune response was compared with VC110. Blue indicates the different levels of downregulation, shadings of red highlight upregulation. Genes, which were not tested in this microarray experiment, were omitted from the figure.

A.

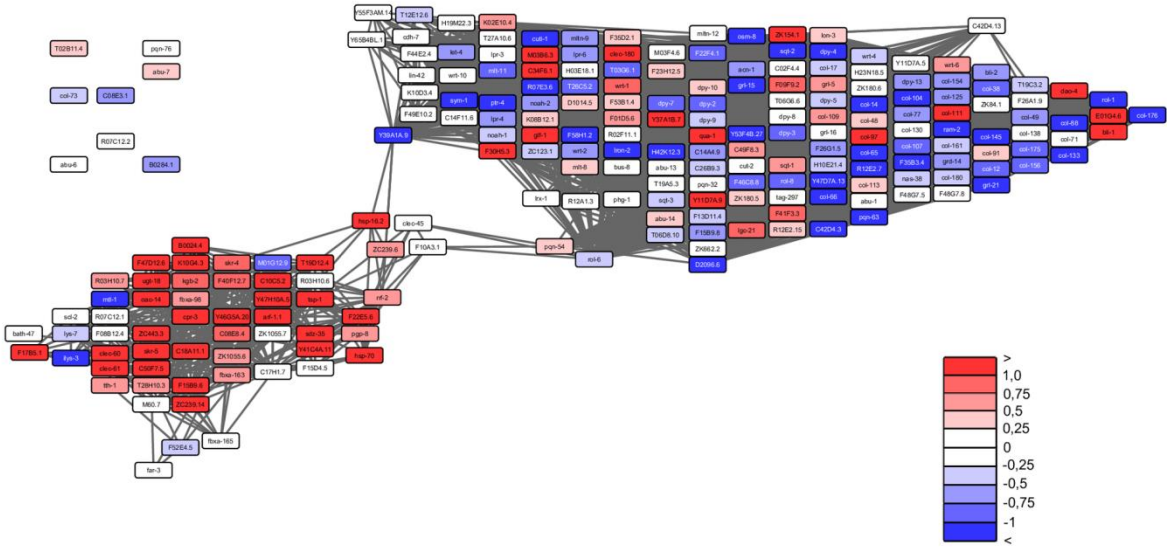


B.

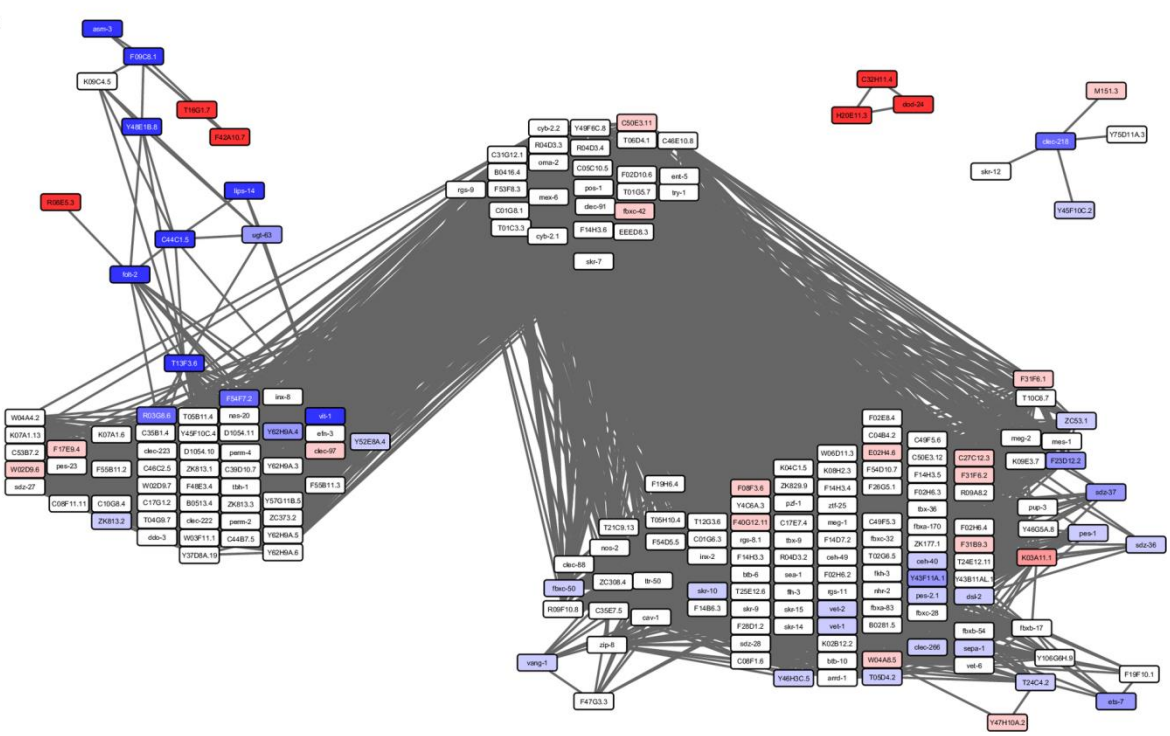


Suppl. figure 2.7 Up-and downregulation of the HSP-90-responsive clusters in the response to *Pseudomonas aeruginosa*. Blue indicates the different levels of downregulation, shadings of red highlight upregulation. Genes not tested in this microarray experiment were omitted from the figure.

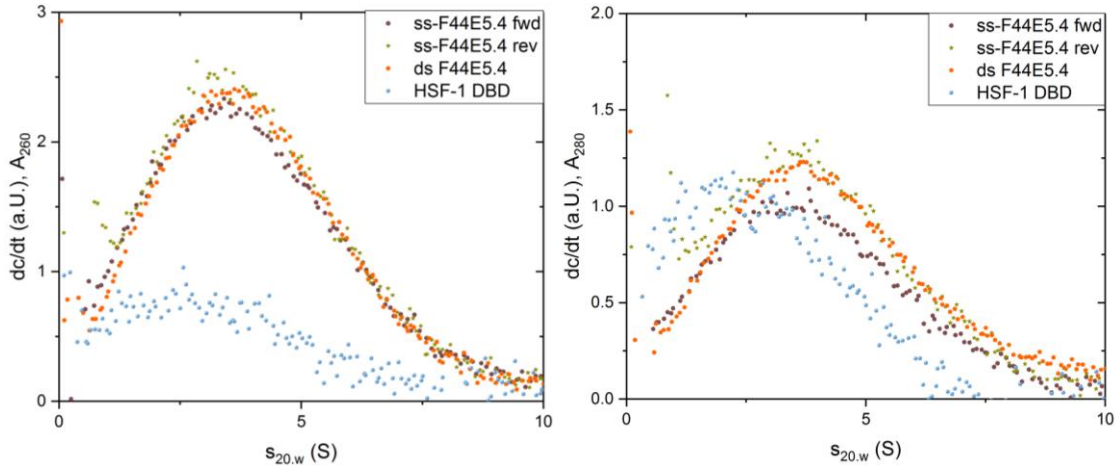
A.



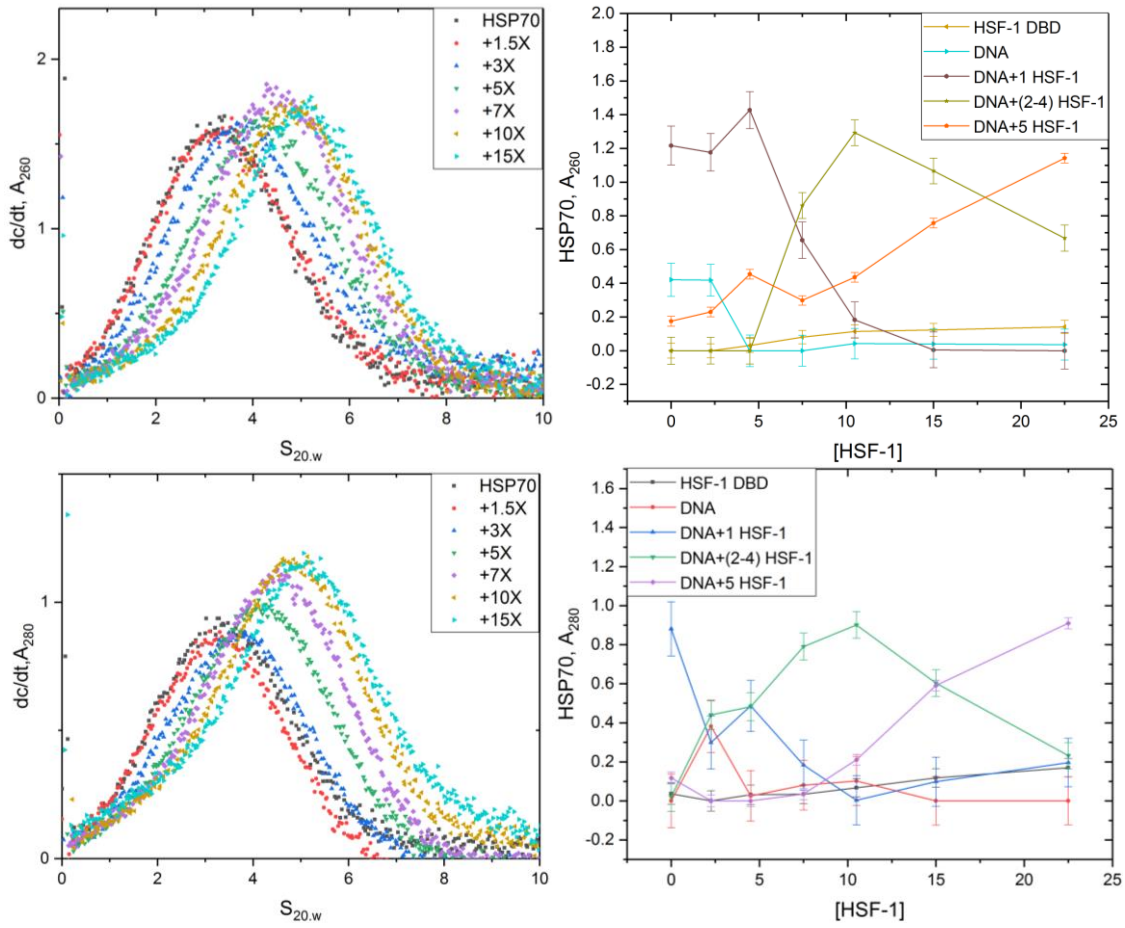
B.



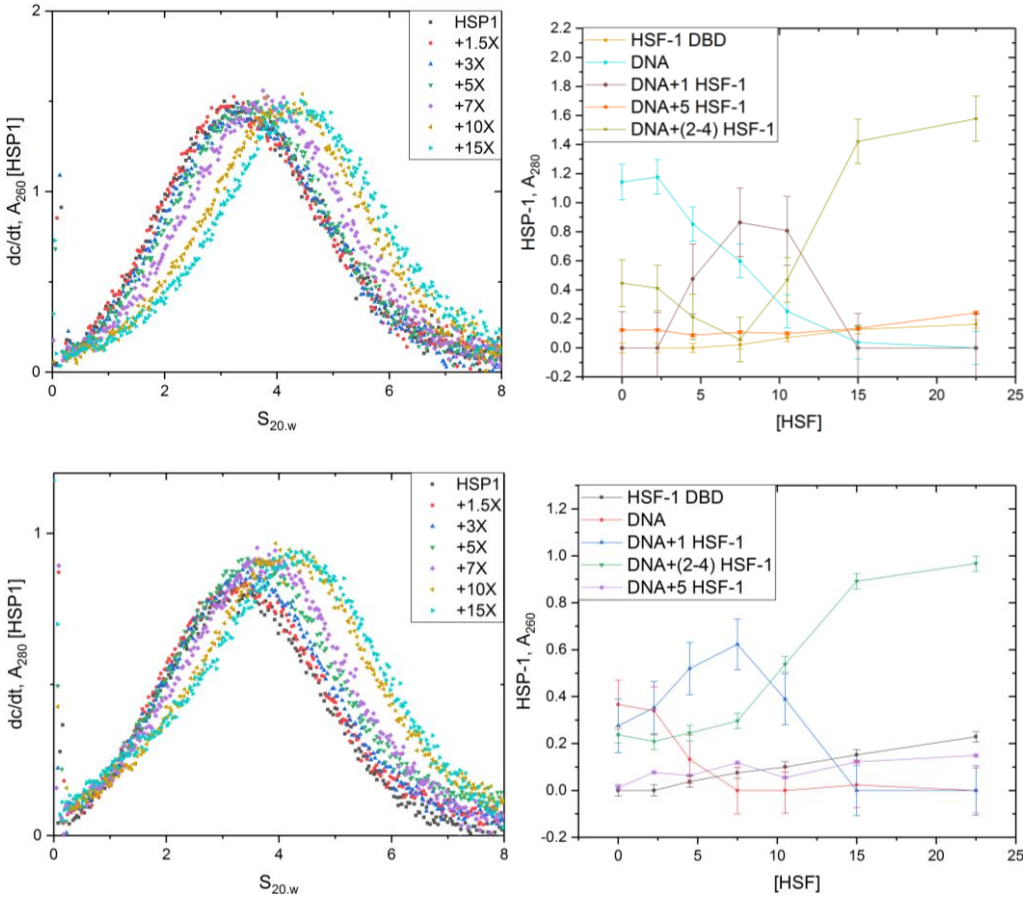
Suppl. Figure 3.1 Absorption of single components at 260 nm (left) and 280 nm (right) recorded in SV-AUC.



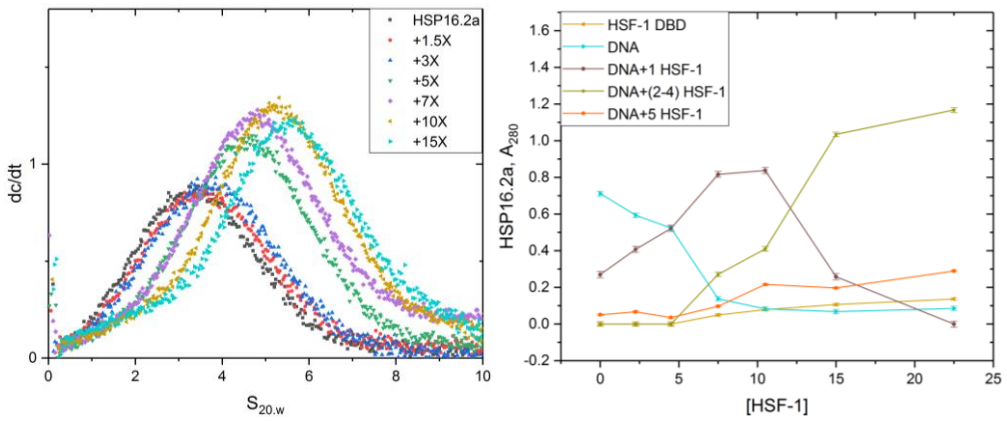
Suppl. Figure 3.2 dc/dt plots of promotor Hsp70 and corresponding partial concentration of complex species obtained via Gaussian fitting.

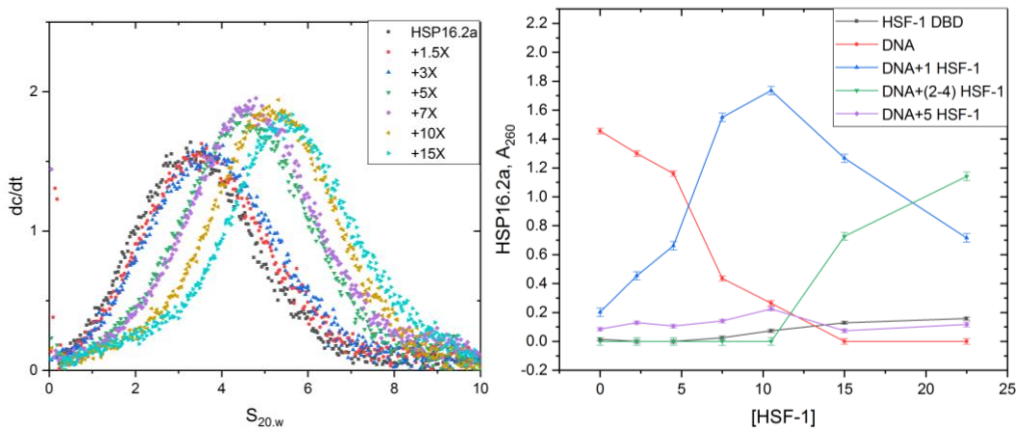


Suppl. Figure 3.3 dc/dt plots of promotor Hsp1 and corresponding partial concentration of complex species obtained via Gaussian fitting.

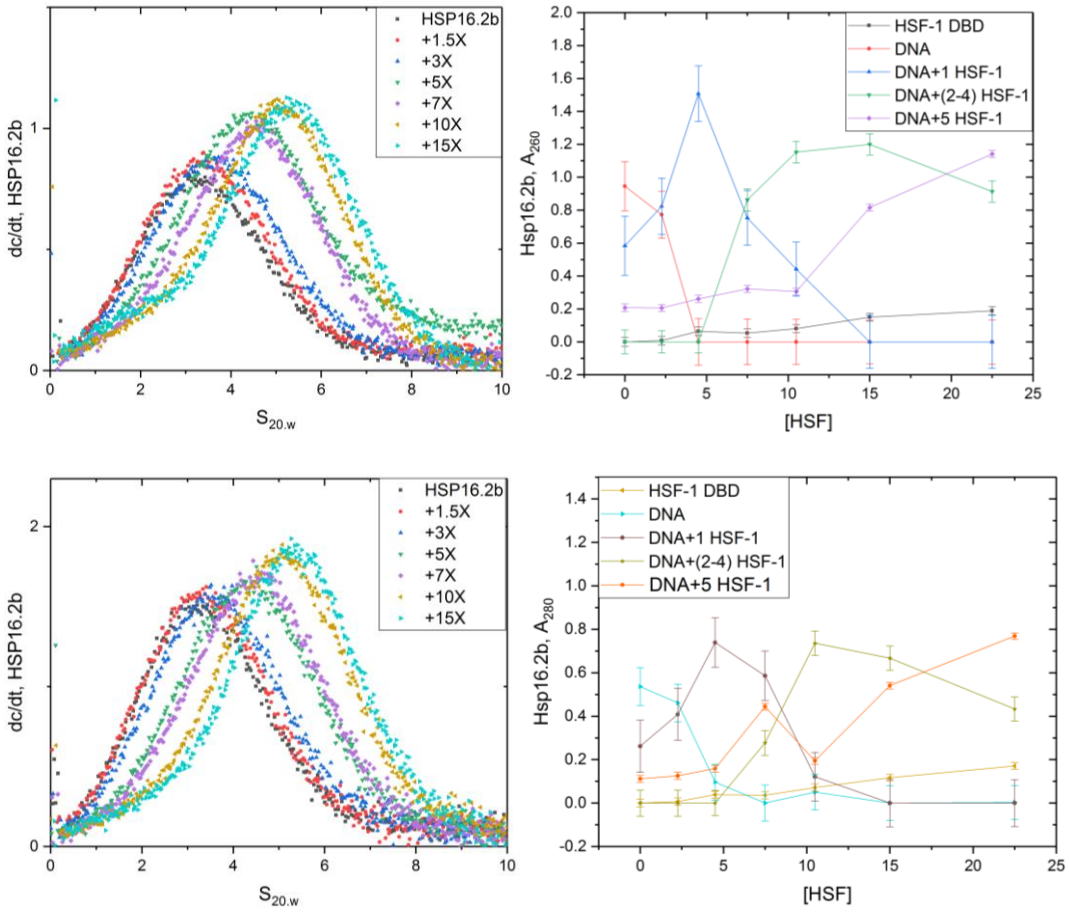


Suppl. Figure 3.4 dc/dt plots of the promotor Hsp16.2a and corresponding partial concentration of complex species obtained via Gaussian fitting.

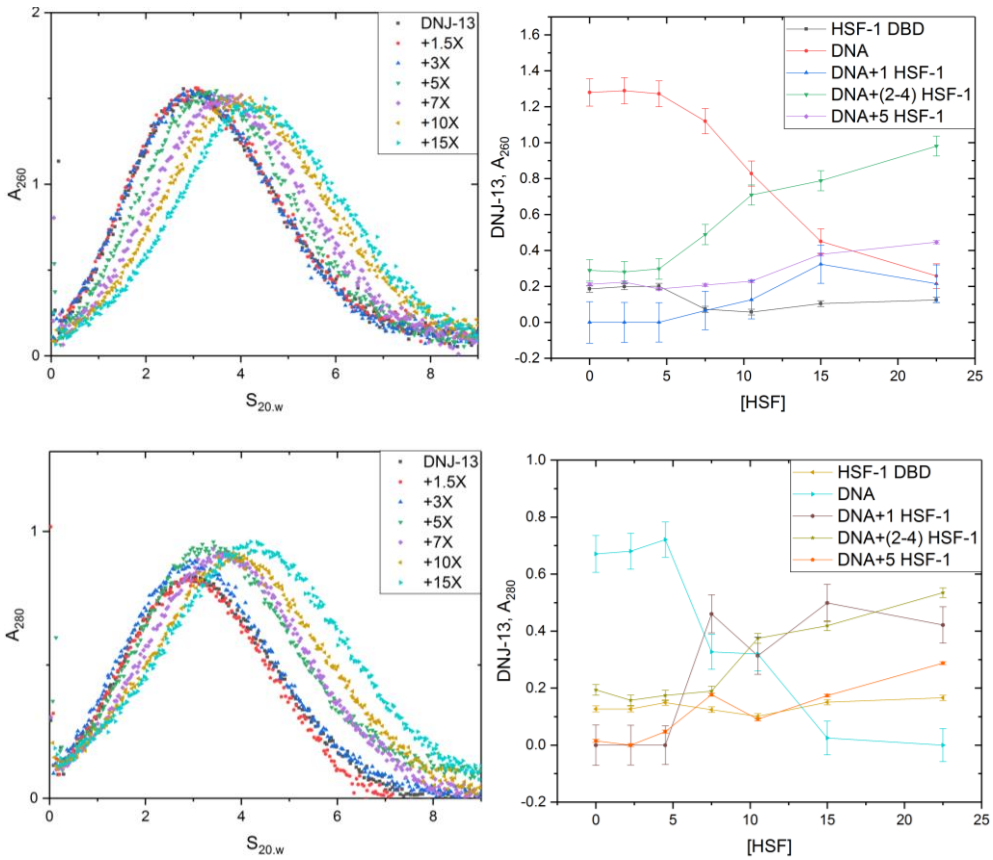




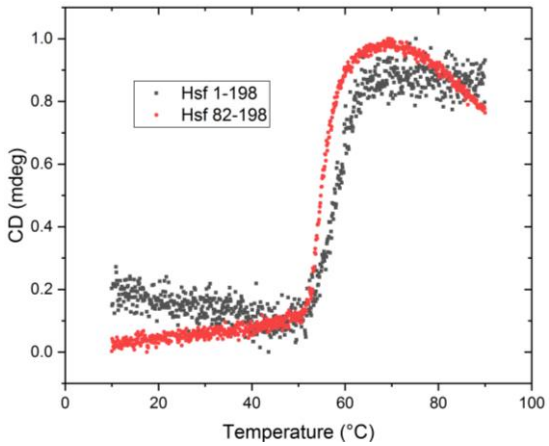
Suppl. Figure 3.5 dc/dt plots of promotor Hsp16.2b and the corresponding partial concentration of complex species obtained via Gaussian fitting.



Suppl. Figure 3.6 dc/dt plots of promotor Dnj-13 and the corresponding partial concentration of complex species obtained via Gaussian fitting.

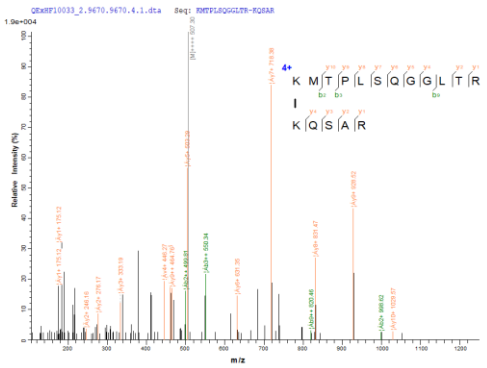


Suppl. Figure 3.7 Comparison of thermal transition recorded in CD of Hsf-1 NBD (AA 1-198) and DBD (AA 82-198).

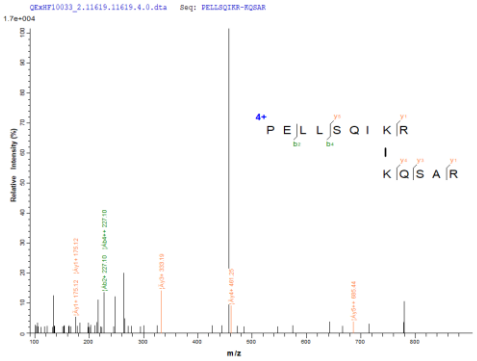


Suppl. Figure 3.8 Spectra of identified Hsf-1 DBD intermolecular crosslinked peptide pairs. A-B. Heavy BS3 crosslinked peptide pair. C-D. Light BS3 crosslinked peptide pair.

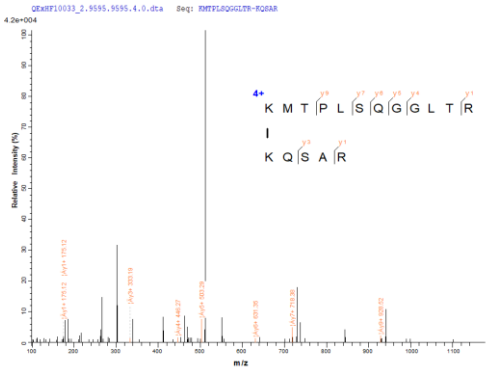
A.



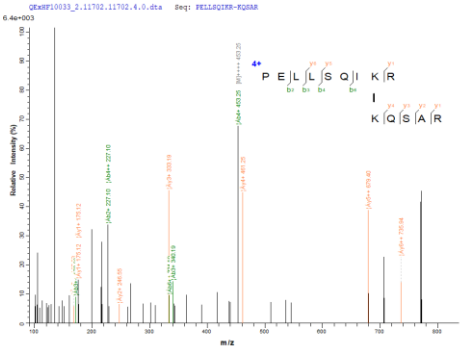
B.



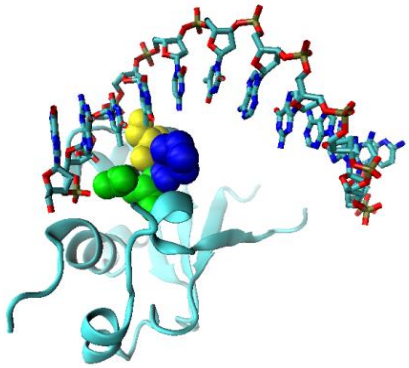
C.



D.

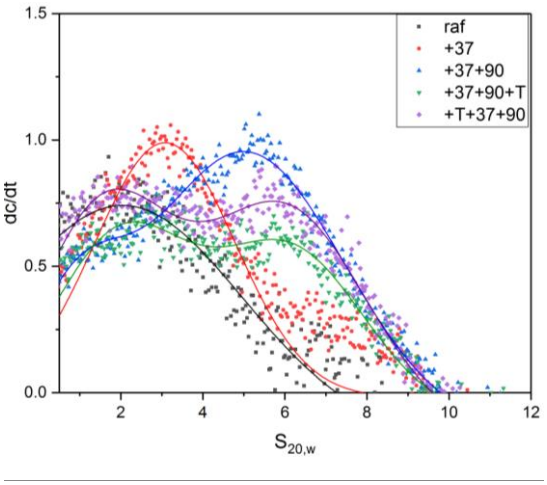


Suppl. Figure 3.9 Simulated interaction site between Hsf-1 and HSE. This structure is obtained based on homology modelling of the structure PDB: 5D5U. Amino acids involved in the interaction site are labeled in VDW model in the VMD. AA68 is shown in yellow, AA71 in blue and AA 72 in green.

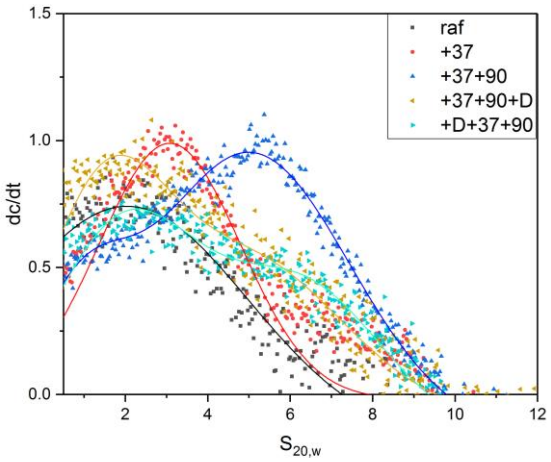


Suppl. Figure 4.1 Influence of different nucleotides on the Raf·CDC-37·HSP-90 complex formation. (A) T = ATP; (B) D = ADP; (C) M = AMP-PNP; (D) rs = ATPγs. The order of added components is the same as the order written in legend, which plays an important role in this case.

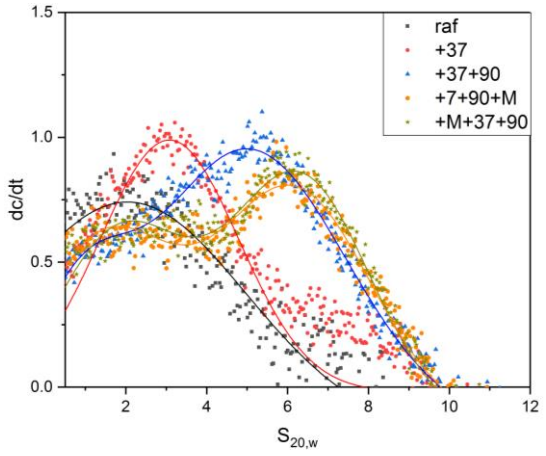
A.



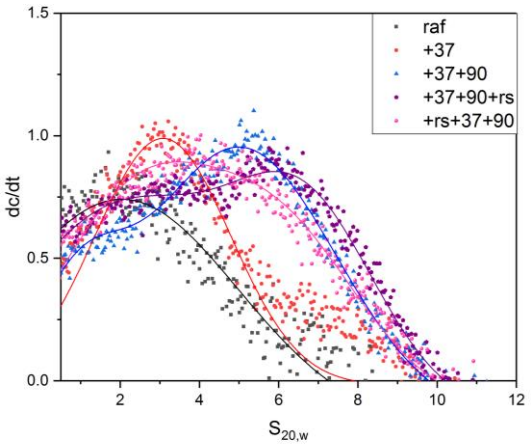
B.



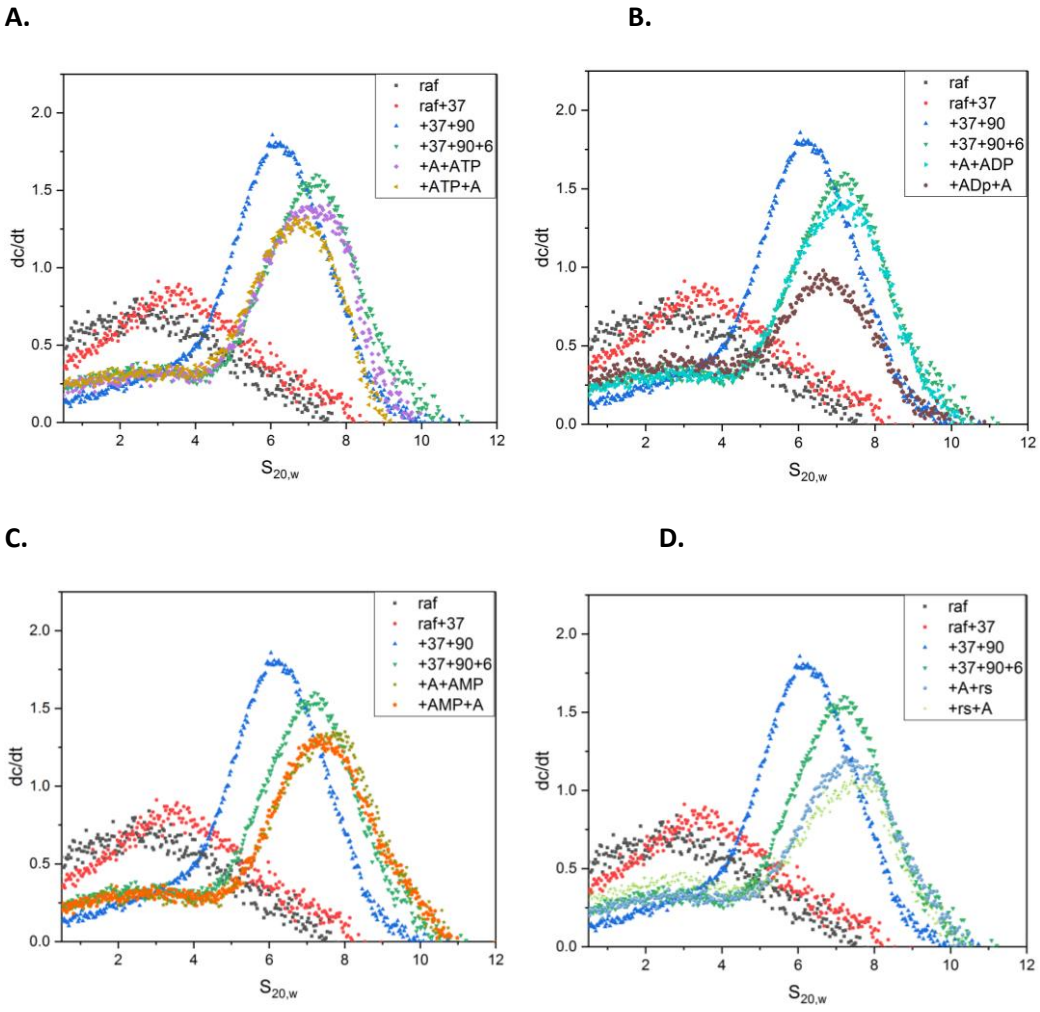
C.



D.



Suppl. Figure 4.2 Influence of nucleotides on the formation of sB-Raf•CDC-37•HSP-90•FKB-6 complex. The letter A is short for the whole complex: sB-Raf•CDC-37•HSP-90•FKB-6. (A) Nucleotide used is ATP. (B) Nucleotide used is ADP. (C) Nucleotide used is AMP-PNP. (D) Nucleotide used is ATPγs.



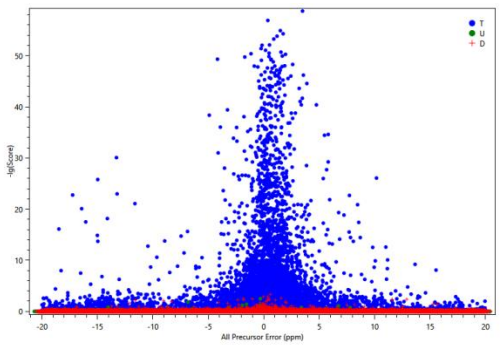
Suppl. Figure 4.3 Alignment of *C.elegans* FKB-6 and the human homolog Fkbp51 obtained from NCBI blast and the corresponding FK deletion sequence. The linker region of FKB-6 is labeled in blue while the linker region of Fkbp51 is labeled in yellow.

```
Query  4    EKIDITPKKDGGVLKLIKKEGQGVVKPTTGTTVKVHYVGTLENGTKFDSSRDRGDQFSFN  63
            E +DI+PK+D GVLK+IK+EG G   P  G  V VHY G L +GTKFDSS DR D+FSF+
Sbjct  21   EGVDISPKQDEGVLKVIKREGTGTEMPMIGDRVFVHYTGWLLDGTKFDSSLDRKDKFSFD  80

Query  64   LGRGNVIKGWDLGVATMTKGEVAEFTIRSDYGYGDAGSPPKIPGGATLIFEVELFEWSAE  123
            LG+G VIK WD+ +ATM  GEV   T + +Y YG AGSPPKIP  ATL+FEVELFE+  E
```

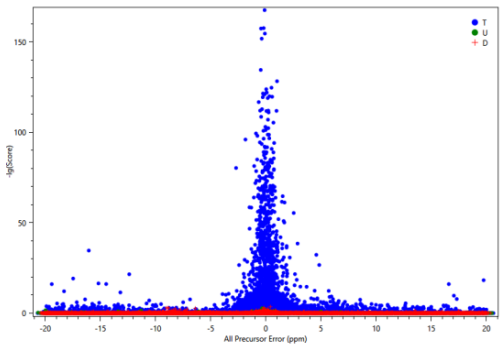
```
Sbjct  81   LGKGEVIKAWDIAIATMKVGEVCHITCKPEYAYGSAGSPPKIPPNATLVFEVELFEFKGE  140

Query  124  DISPDRDGTILRTIIVEGSKNSFPNDTSKVLAHCVGTYQGTEFYNREVNFHIGEGSEEGL  183
            D++ + DG I+R I   G   + PN+ + V     G Y+   F  RE+ F IGEG    L
Sbjct  141  DLTEEEDGGIIRRIQTRGEGYAKPNEGAIVEVALEGYYKDKLFDQRELRFEIGEGENLDL  200

Query  184  PEGVERALRRFQLGEKSKIEIRGHKYTYGNSPPAGSNIPVNATLEFTIFLKEFEKVPATW  243
            P G+ERA++R + GE S + ++   Y +G+       IP NA L++ + LK FEK   +W
Sbjct  201  PYGLERAIQRMEKGEHSIVYLKP-SYAFGSVGKEKFQIPPNAELKYELHLKSFEKAKESW  259

Query  244  EMTAEEKLDAAKQAKDRGTMYLQKGNLKLAYNKYKRAEEVLEYEKSTDPEKMAERETILN  303
            EM +EEKL+ +   K+RGT+Y ++G  K A  +YK+    LEYE S   E+  + + +
Sbjct  260  EMNSEEKLEQSTIVKERGTVYFKEGKYKQALLQYKKIVSWLEYESSFSNEEAQKAQALRL  319

Query  304  GAYLNLSLVCSKQNEQLECIKWCDKVLETKPGNVKALYRKATALLTMNEVRDAMKLFEKI  363
             ++LNL++   K       I+ C+K LE    N K L+R+  A L +N+   A   F+K+
Sbjct  320  ASHLNLAMCHLKLQAFSAAIESCNKALELDSNNEKGLFRRGEAHLAVNDFELARADFQKV  379

Query  364  VEVEPENKAAAQQIIVCRNTIREQNERDKKRFKNLFAKISTEE----------DKPTNTV  413
            +++ P NKAA  Q+ VC+  IR Q  R+KK + N+F +++ EE           D PT+T
Sbjct  380  LQLYPNNKAAKTQLAVCQQRIRRQLAREKKLYANMFERLAEEENKAKAEASSGDHPTDT-  438

Query  414  EDEDEIVASTSGSST  428
            E ++E  ++T+GS +
Sbjct  439  EMKEEQKSNTAGSQS  453
```

Suppl. Figure 4.4 Qualification of MS analysis. Error distribution of all precusors in (A) the LTQ replicate and (B) the QE replicate; Error distribution of filtered precursors in (C) LTQ replicate and (D) in the QE replicate; (E) Total amounts of spectra reported by the two analysis programms in different replicates. a stands for results obtained from the analysis programm xMass; b for results obtained from the analysis programm pLink. (F) Comparison of the amount of reported and unique peptide pairs in MS analysis. Larger circle on the left side indicates results obtained from the xMass and small circle on the right side indicates reported results obtianed from the pLink.

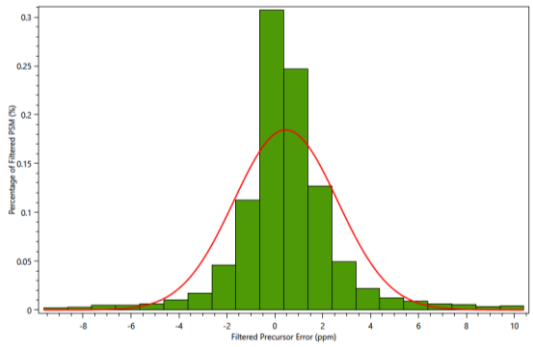
A.



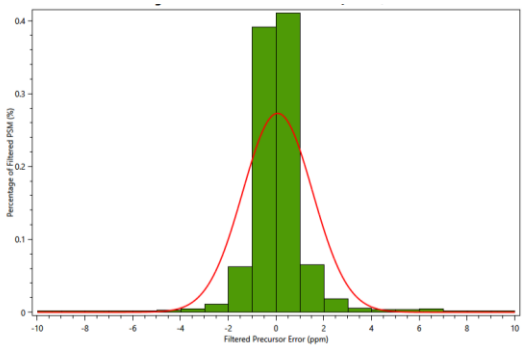
B.



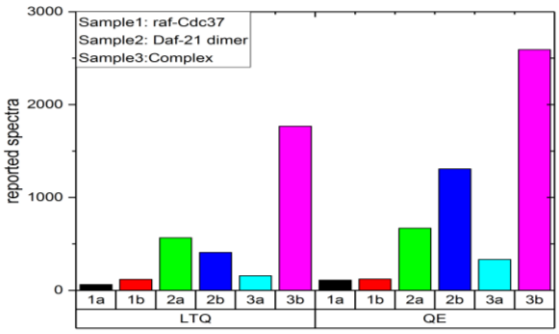
C.


Percentage of Filtered PSM (%)
Filtered Precursor Error (ppm)

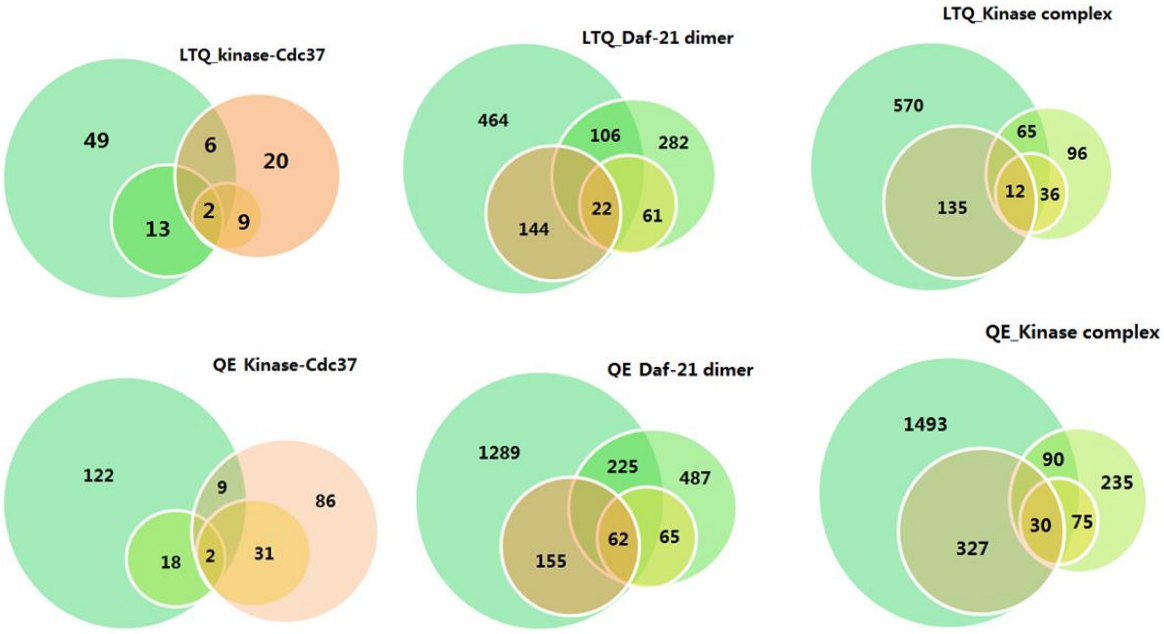
D.


Percentage of Filtered PSM (%)
Filtered Precursor Error (ppm)

E.

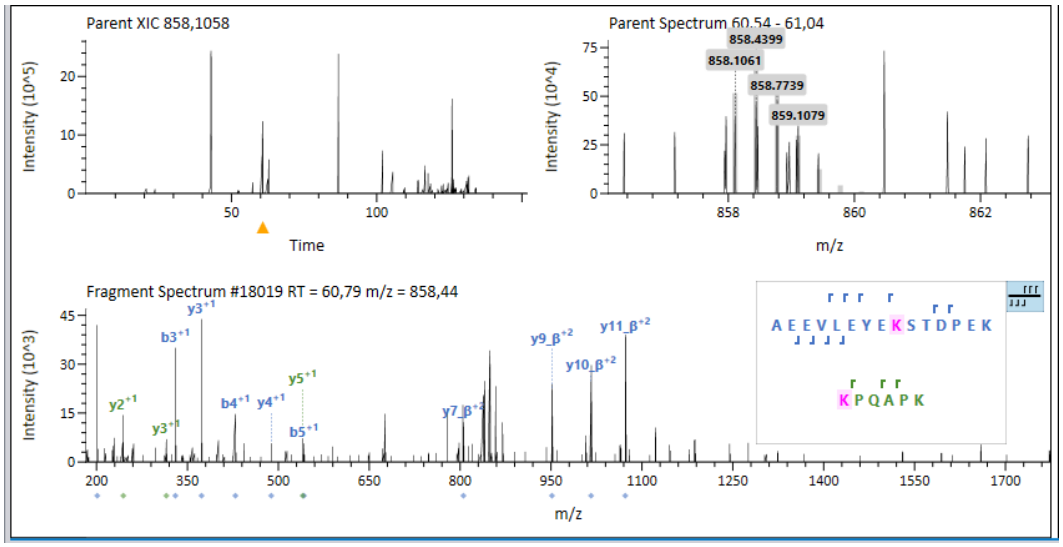

Sample1: raf-Cdc37
Sample2: Daf-21 dimer
Sample3:Complex
reported spectra
3000
2000
1000
0
1a
1b
2a
2b
3a
3b
LTQ
QE

F.

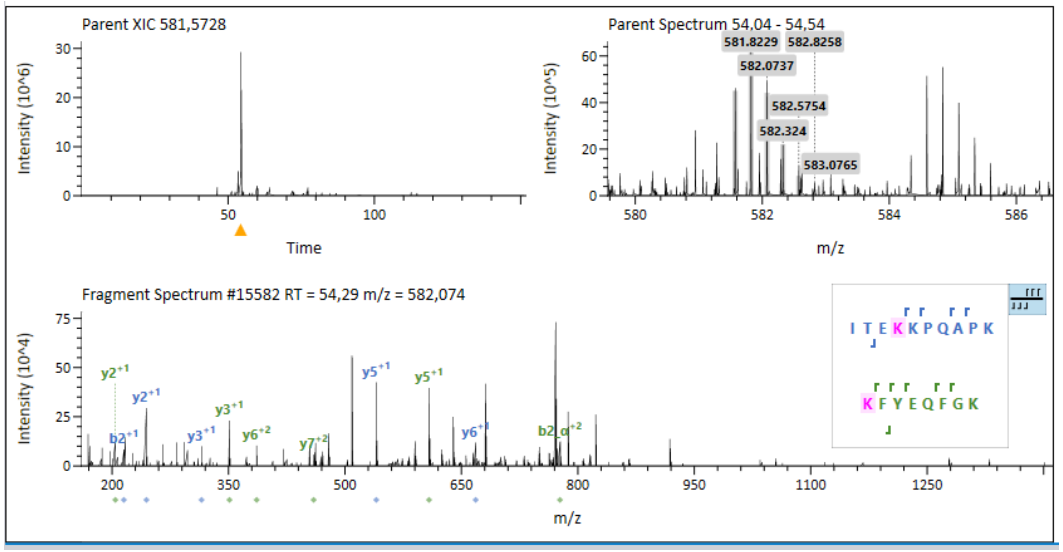

LTQ_kinase-Cdc37
49
6
20
13
2
9
LTQ_Daf-21 dimer
464
106
282
144
22
61
LTQ_Kinase complex
570
65
96
135
12
36
QE Kinase-Cdc37
122
9
86
18
2
31
QE Daf-21 dimer
1289
225
487
155
62
65
QE_Kinase complex
1493
90
235
327
30
75

Suppl. figure 4.5 Most interesting Identified crosslink sites in the kinase complex. Here the spectra of crosslinked peptide pairs are given: A-B. The contact site of CDC-37, HSP-90 and FKB-6. C-E. The contact side among Cdc37, sB-Raf, HSP-90 and FKB-6.

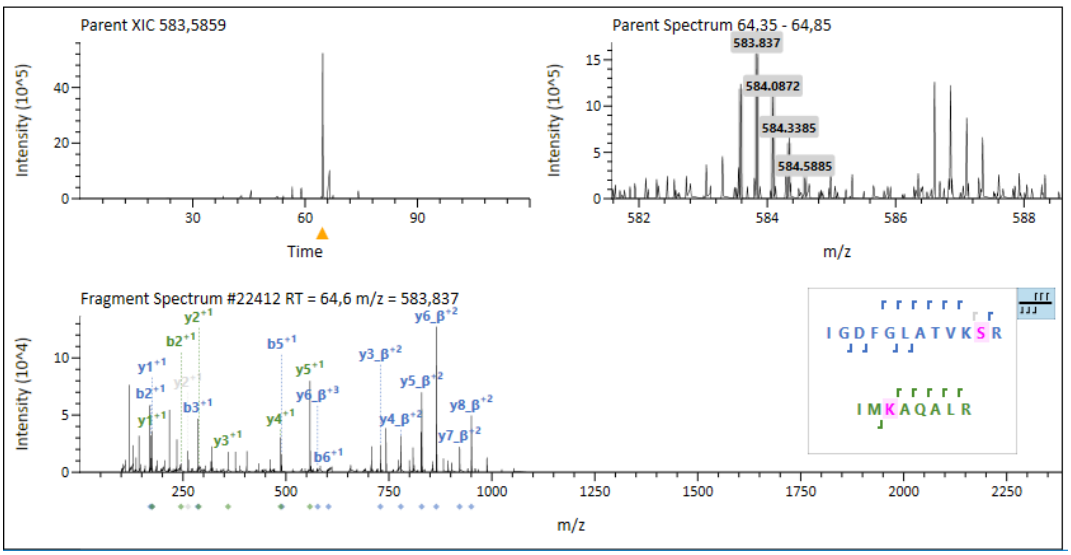
A.



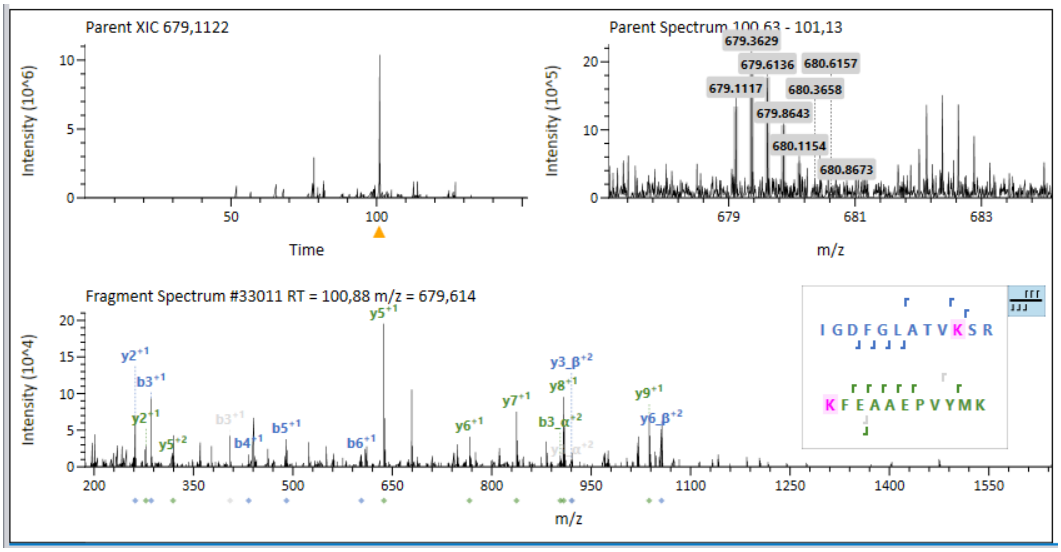
B.



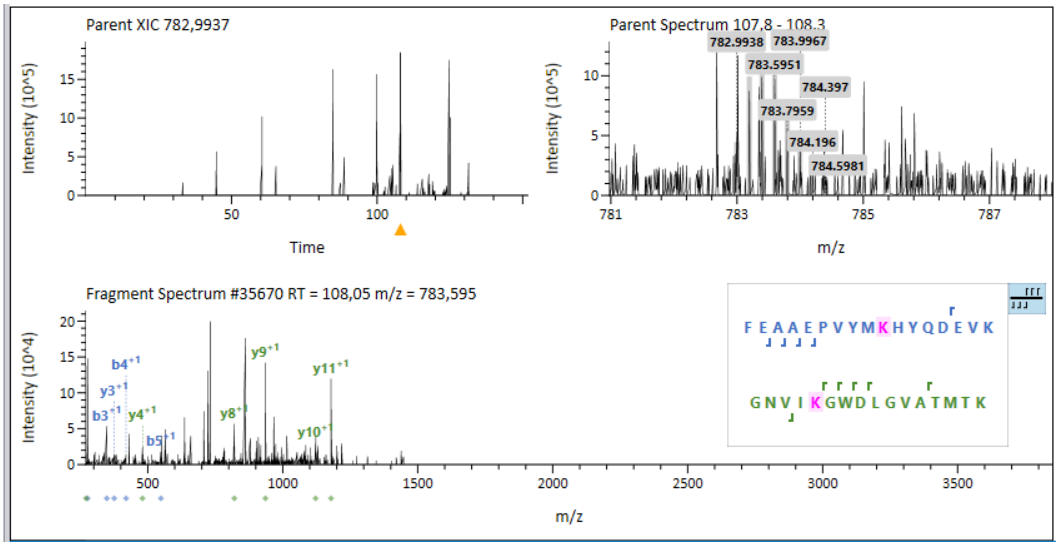
C.



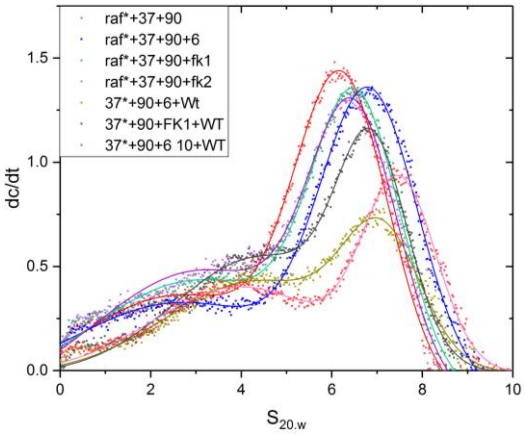
D.



E.

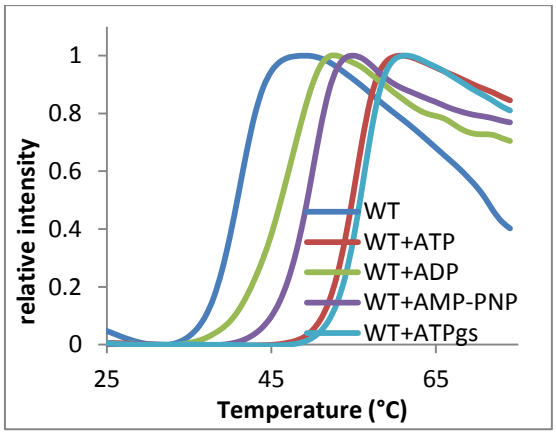


Suppl. Figure 4.6 Raf-complex formed with different labeled proteins. The labeled protein is marked with a star.

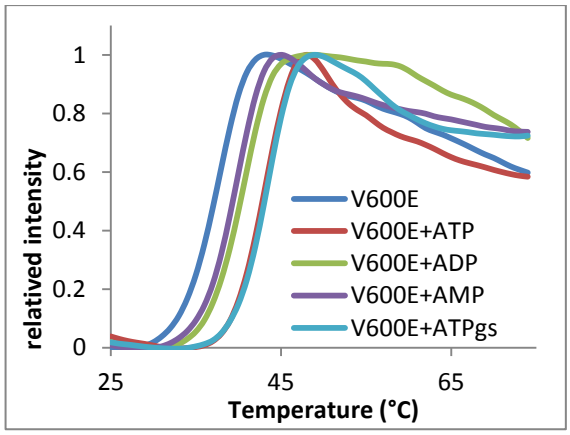


Suppl. Figure 5.1 Comparison of the normalized TSA results of raf kinase domain WT vs mutants. (A) WT-Raf kinase domain measured with different nucleotides. (B) Raf mutant V600E measured with different nucleotides. (C) Raf mutant G464E measured with different nucleotides. (D) Raf mutant G464E measured with different nucleotides.

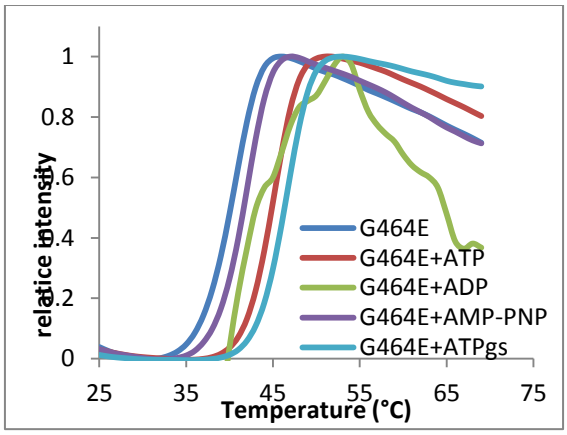
A



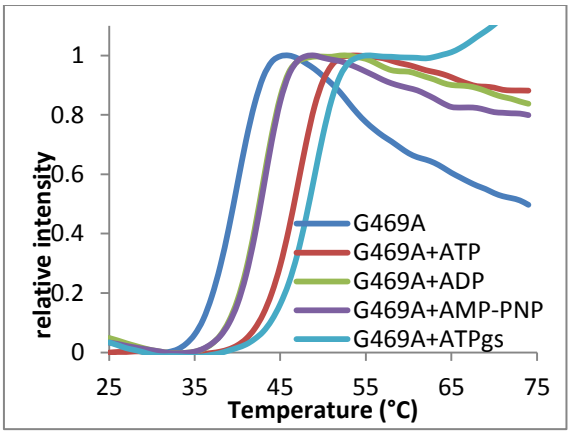
B



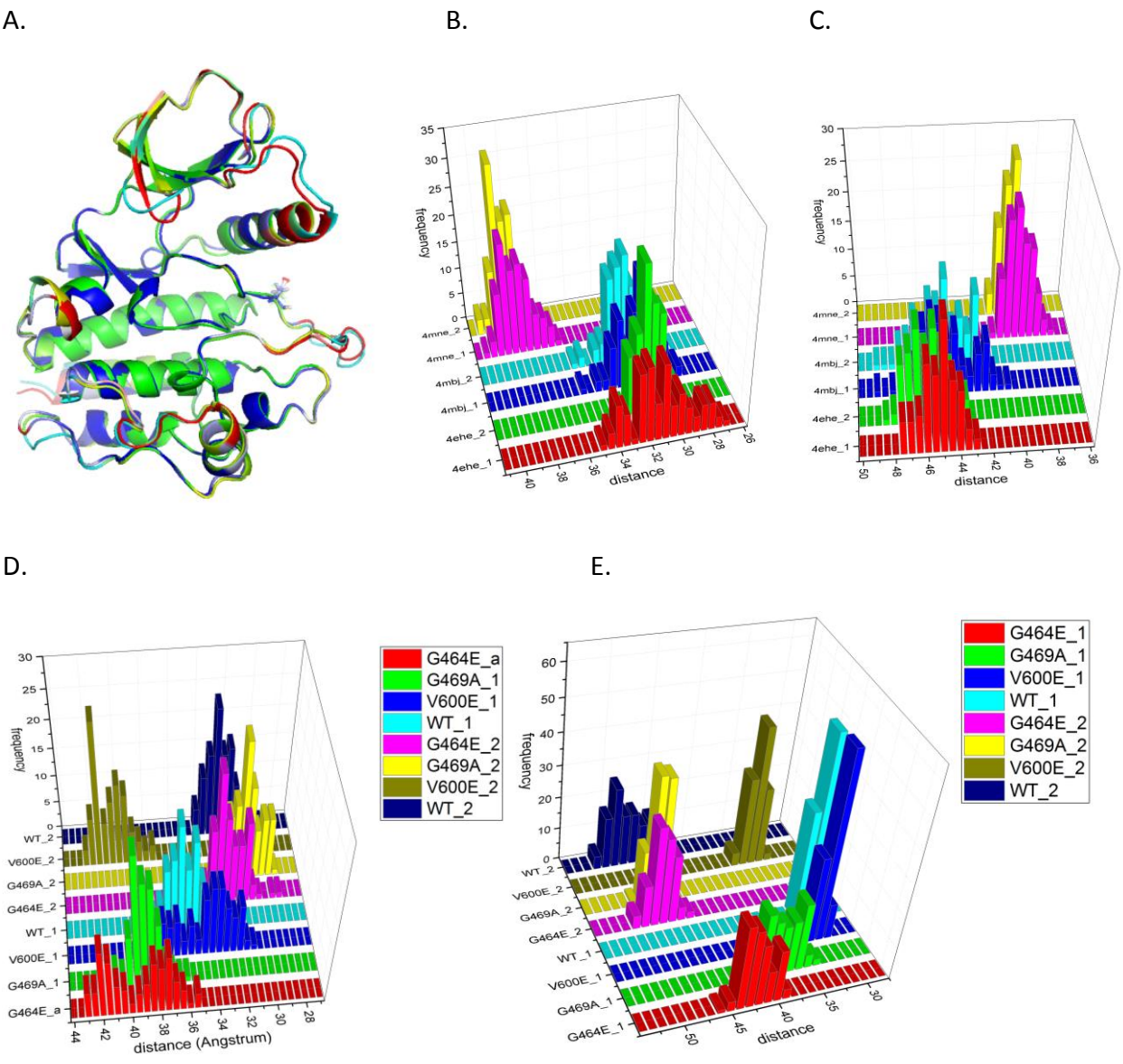
C



D



Suppl. Figure 5.2 Molecular dynamics simulation of sB-Raf kinase domain and its corresponding mutants in the active/ inactive/ ATP-bound state. A. B-factor map of simulated sB-Raf structures. WT is showed in green-yellow-cyan and the mutant V600E is showed in blue-white-red. B. Size of the ATP-binding pocket (residue 602 and 1441) of the sB-Raf kinase domain in the inactive form (4ehe and 4mbj) or in the active form (4mne). C. Distance between the activation loop and active center (residue 734 and 3463) in the kinase domain. D. The size of the ATP-binding pocket in mutants at different state. 1 indicates the active state while 2 indicates the inactive state of the kinase domain. E. Distance between the active loop and the activation center in the mutants at different state.



7 References

1. Eckl, J., et al., *Hsp90-downregulation influences the heat-shock response, innate immune response and onset of oocyte development in nematodes.* PLoS One, 2017. **12**(10): p. e0186386.
2. Schopf, F.H., M.M. Biebl, and J. Buchner, *The HSP90 chaperone machinery.* Nat Rev Mol Cell Biol, 2017. **18**(6): p. 345-360.
3. Prodromou, C., *Regulatory Mechanisms of Hsp90.* Biochem Mol Biol J, 2017. **3**(1): p. 2.
4. Dollins, D.E., et al., *Structures of GRP94-nucleotide complexes reveal mechanistic differences between the hsp90 chaperones.* Mol Cell, 2007. **28**(1): p. 41-56.
5. Immormino, R.M., et al., *Ligand-induced conformational shift in the N-terminal domain of GRP94, an Hsp90 chaperone.* J Biol Chem, 2004. **279**(44): p. 46162-71.
6. Dollins, D.E., R.M. Immormino, and D.T. Gewirth, *Structure of unliganded GRP94, the endoplasmic reticulum Hsp90. Basis for nucleotide-induced conformational change.* J Biol Chem, 2005. **280**(34): p. 30438-47.
7. Ali, M.M., et al., *Crystal structure of an Hsp90-nucleotide-p23/Sba1 closed chaperone complex.* Nature, 2006. **440**(7087): p. 1013-7.
8. Li, Y., et al., *Sulforaphane inhibits pancreatic cancer through disrupting Hsp90-p50(Cdc37) complex and direct interactions with amino acids residues of Hsp90.* J Nutr Biochem, 2012. **23**(12): p. 1617-26.
9. Silverstein, A.M., et al., *p50(cdc37) binds directly to the catalytic domain of Raf as well as to a site on hsp90 that is topologically adjacent to the tetratricopeptide repeat binding site.* J Biol Chem, 1998. **273**(32): p. 20090-5.
10. Grammatikakis, N., et al., *p50(cdc37) acting in concert with Hsp90 is required for Raf-1 function.* Mol Cell Biol, 1999. **19**(3): p. 1661-72.
11. Terasawa, K., et al., *Cdc37 interacts with the glycine-rich loop of Hsp90 client kinases.* Mol Cell Biol, 2006. **26**(9): p. 3378-89.
12. Keramisanou, D., et al., *Molecular Mechanism of Protein Kinase Recognition and Sorting by the Hsp90 Kinome-Specific Cochaperone Cdc37.* Mol Cell, 2016. **62**(2): p. 260-71.
13. Terasawa, K., et al., *Client binding of Cdc37 is regulated intramolecularly and intermolecularly.* Biosci Biotechnol Biochem, 2006. **70**(6): p. 1542-6.
14. Terasawa, K. and Y. Minami, *A client-binding site of Cdc37.* FEBS J, 2005. **272**(18): p. 4684-90.
15. Smith, J.R., et al., *Restricting direct interaction of CDC37 with HSP90 does not compromise chaperoning of client proteins.* Oncogene, 2015. **34**(1): p. 15-26.
16. Zhang, W., et al., *Biochemical and structural studies of the interaction of Cdc37 with Hsp90.* J Mol Biol, 2004. **340**(4): p. 891-907.
17. Diedrich, B., et al., *Discrete cytosolic macromolecular BRAF complexes exhibit distinct activities and composition.* EMBO J, 2017. **36**(5): p. 646-663.
18. Okamura, Y., et al., *COXPRESdb in 2015: coexpression database for animal species by DNA-microarray and RNAseq-based expression data with multiple quality assessment systems.* Nucleic Acids Res, 2015. **43**(Database issue): p. D82-6.
19. Hibbs, M.A., et al., *Exploring the functional landscape of gene expression: directed search of large microarray compendia.* Bioinformatics, 2007. **23**(20): p. 2692-9.
20. Zhu, Q., et al., *Targeted exploration and analysis of large cross-platform human transcriptomic compendia.* Nat Methods, 2015. **12**(3): p. 211-4, 3 p following 214.
21. Papsdorf, K., et al., *Construction and evaluation of yeast expression networks by database-guided predictions.* Microb Cell, 2016. **3**(6): p. 236-247.

22. Mi, H., et al., *PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements.* Nucleic Acids Res, 2017. **45**(D1): p. D183-D189.
23. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.
24. Woodford, M.R., et al., *Tumor suppressor Tsc1 is a new Hsp90 co-chaperone that facilitates folding of kinase and non-kinase clients.* EMBO J, 2017. **36**(24): p. 3650-3665.
25. Woodford, M.R., et al., *The FNIP co-chaperones decelerate the Hsp90 chaperone cycle and enhance drug binding.* Nat Commun, 2016. **7**: p. 12037.
26. Hornbeck, P.V., et al., *PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse.* Nucleic Acids Res, 2012. **40**(Database issue): p. D261-70.
27. Wang, X., et al., *Thr90 phosphorylation of Hsp90alpha by protein kinase A regulates its chaperone machinery.* Biochem J, 2012. **441**(1): p. 387-97.
28. Xu, W., et al., *Dynamic tyrosine phosphorylation modulates cycling of the HSP90-P50(CDC37)-AHA1 chaperone machine.* Mol Cell, 2012. **47**(3): p. 434-43.
29. Muller, P., et al., *C-terminal phosphorylation of Hsp70 and Hsp90 regulates alternate binding to co-chaperones CHIP and HOP to determine cellular protein folding/degradation balances.* Oncogene, 2013. **32**(25): p. 3101-10.
30. Assimon, V.A., D.R. Southworth, and J.E. Gestwicki, *Specific Binding of Tetratricopeptide Repeat Proteins to Heat Shock Protein 70 (Hsp70) and Heat Shock Protein 90 (Hsp90) Is Regulated by Affinity and Phosphorylation.* Biochemistry, 2015. **54**(48): p. 7120-31.
31. Lu, X.A., et al., *The regulatory mechanism of a client kinase controlling its own release from Hsp90 chaperone machinery through phosphorylation.* Biochem J, 2014. **457**(1): p. 171-83.
32. Nguyen, M.T.N., et al., *Isoform-Specific Phosphorylation in Human Hsp90beta Affects Interaction with Clients and the Cochaperone Cdc37.* J Mol Biol, 2017. **429**(5): p. 732-752.
33. Mollapour, M., et al., *Threonine 22 phosphorylation attenuates Hsp90 interaction with cochaperones and affects its chaperone activity.* Mol Cell, 2011. **41**(6): p. 672-81.
34. Soroka, J., et al., *Conformational switching of the molecular chaperone Hsp90 via regulated phosphorylation.* Mol Cell, 2012. **45**(4): p. 517-28.
35. Liu, W. and R. Landgraf, *Phosphorylated and unphosphorylated serine 13 of CDC37 stabilize distinct interactions between its client and HSP90 binding domains.* Biochemistry, 2015. **54**(7): p. 1493-504.
36. Vaughan, C.K., et al., *Hsp90-dependent activation of protein kinases is regulated by chaperone-targeted dephosphorylation of Cdc37.* Mol Cell, 2008. **31**(6): p. 886-95.
37. Eckl, J.M., et al., *Nucleotide-Free sB-Raf is Preferentially Bound by Hsp90 and Cdc37 In Vitro.* J Mol Biol, 2016. **428**(20): p. 4185-4196.
38. Aoyagi, S. and T.K. Archer, *Modulating molecular chaperone Hsp90 functions through reversible acetylation.* Trends Cell Biol, 2005. **15**(11): p. 565-7.
39. Cohen, T. and T.P. Yao, *AcK-knowledge reversible acetylation.* Sci STKE, 2004. **2004**(245): p. pe42.
40. Scroggins, B.T., et al., *An acetylation site in the middle domain of Hsp90 regulates chaperone function.* Mol Cell, 2007. **25**(1): p. 151-9.
41. Murphy, P.J., et al., *Regulation of the dynamics of hsp90 action on the glucocorticoid receptor by acetylation/deacetylation of the chaperone.* J Biol Chem, 2005. **280**(40): p. 33792-9.
42. Bali, P., et al., *Inhibition of histone deacetylase 6 acetylates and disrupts the chaperone function of heat shock protein 90: a novel basis for antileukemia activity of histone deacetylase inhibitors.* J Biol Chem, 2005. **280**(29): p. 26729-34.
43. Suuronen, T., et al., *Regulation of ER alpha signaling pathway in neuronal HN10 cells: role of protein acetylation and Hsp90.* Neurochem Res, 2008. **33**(9): p. 1768-75.

44. Donlin, L.T., et al., *Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization.* Genes Dev, 2012. **26**(2): p. 114-9.
45. Abu-Farha, M., et al., *Proteomic analyses of the SMYD family interactomes identify HSP90 as a novel target for SMYD2.* J Mol Cell Biol, 2011. **3**(5): p. 301-8.
46. Hamamoto, R., et al., *SMYD2-dependent HSP90 methylation promotes cancer cell proliferation by regulating the chaperone complex formation.* Cancer Lett, 2014. **351**(1): p. 126-33.
47. Martinez-Ruiz, A., et al., *S-nitrosylation of Hsp90 promotes the inhibition of its ATPase and endothelial nitric oxide synthase regulatory activities.* Proc Natl Acad Sci U S A, 2005. **102**(24): p. 8525-30.
48. Retzlaff, M., et al., *Hsp90 is regulated by a switch point in the C-terminal domain.* EMBO Rep, 2009. **10**(10): p. 1147-53.
49. Morra, G., G. Verkhivker, and G. Colombo, *Modeling signal propagation mechanisms and ligand-based conformational dynamics of the Hsp90 molecular chaperone full-length dimer.* PLoS Comput Biol, 2009. **5**(3): p. e1000323.
50. Mollapour, M., et al., *Asymmetric Hsp90 N domain SUMOylation recruits Aha1 and ATP-competitive inhibitors.* Mol Cell, 2014. **53**(2): p. 317-29.
51. Antunica-Noguerol, M., et al., *The activity of the glucocorticoid receptor is regulated by SUMO conjugation to FKBP51.* Cell Death Differ, 2016. **23**(10): p. 1579-91.
52. Skaggs, H.S., et al., *HSF1-TPR interaction facilitates export of stress-induced HSP70 mRNA.* J Biol Chem, 2007. **282**(47): p. 33902-7.
53. Williams, R.S. and I.J. Benjamin, *Protective responses in the ischemic myocardium.* J Clin Invest, 2000. **106**(7): p. 813-8.
54. Xiao, X., et al., *HSF1 is required for extra-embryonic development, postnatal growth and protection during inflammatory responses in mice.* EMBO J, 1999. **18**(21): p. 5943-52.
55. Kallio, M., et al., *Brain abnormalities, defective meiotic chromosome synapsis and female subfertility in HSF2 null mice.* EMBO J, 2002. **21**(11): p. 2591-601.
56. Kaitsuka, T., K. Tomizawa, and M. Matsushita, *Transformation of eEF1Bdelta into heat-shock response transcription factor by alternative splicing.* EMBO Rep, 2011. **12**(7): p. 673-81.
57. Kubota, H., et al., *Transcriptional activation of mouse cytosolic chaperonin CCT subunit genes by heat shock factors HSF1 and HSF2.* FEBS Lett, 1999. **461**(1-2): p. 125-9.
58. Cahill, C.M., et al., *Transcriptional repression of the prointerleukin 1beta gene by heat shock factor 1.* J Biol Chem, 1996. **271**(40): p. 24874-9.
59. Pirkkala, L., P. Nykanen, and L. Sistonen, *Roles of the heat shock transcription factors in regulation of the heat shock response and beyond.* FASEB J, 2001. **15**(7): p. 1118-31.
60. Gidalevitz, T., V. Prahlad, and R.I. Morimoto, *The stress of protein misfolding: from single cells to multicellular organisms.* Cold Spring Harb Perspect Biol, 2011. **3**(6).
61. Vihervaara, A. and L. Sistonen, *HSF1 at a glance.* J Cell Sci, 2014. **127**(Pt 2): p. 261-6.
62. Wang, X., et al., *Phosphorylation of HSF1 by MAPK-activated protein kinase 2 on serine 121, inhibits transcriptional activity and promotes HSP90 binding.* J Biol Chem, 2006. **281**(2): p. 782-91.
63. Xing, H., et al., *HSF1 modulation of Hsp70 mRNA polyadenylation via interaction with symplekin.* J Biol Chem, 2004. **279**(11): p. 10551-5.
64. Guo, Y., et al., *Evidence for a mechanism of repression of heat shock factor 1 transcriptional activity by a multichaperone complex.* J Biol Chem, 2001. **276**(49): p. 45791-9.
65. Eckl, J.M., et al., *Hsp90.Cdc37 Complexes with Protein Kinases Form Cooperatively with Multiple Distinct Interaction Sites.* J Biol Chem, 2015. **290**(52): p. 30843-54.
66. Kornev, A.P., et al., *Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism.* Proc Natl Acad Sci U S A, 2006. **103**(47): p. 17783-8.
67. Fantl, W.J., et al., *Activation of Raf-1 by 14-3-3 proteins.* Nature, 1994. **371**(6498): p. 612-4.

68. Hu, J., et al., *Allosteric activation of functionally asymmetric RAF kinase dimers.* Cell, 2013. **154**(5): p. 1036-1046.
69. Jaiswal, R.K., et al., *The mitogen-activated protein kinase cascade is activated by B-Raf in response to nerve growth factor through interaction with p21ras.* Mol Cell Biol, 1994. **14**(10): p. 6944-53.
70. Moretti, S., et al., *Biochemical and molecular characterization of the novel BRAF(V599Ins) mutation detected in a classic papillary thyroid carcinoma.* Oncogene, 2006. **25**(30): p. 4235-40.
71. Dankner, M., et al., *Classifying BRAF alterations in cancer: new rational therapeutic strategies for actionable mutations.* Oncogene, 2018. **37**(24): p. 3183-3199.
72. Davies, H., et al., *Mutations of the BRAF gene in human cancer.* Nature, 2002. **417**(6892): p. 949-54.
73. Boczek, E.E., et al., *Conformational processing of oncogenic v-Src kinase by the molecular chaperone Hsp90.* Proc Natl Acad Sci U S A, 2015. **112**(25): p. E3189-98.
74. Stepanova, L., et al., *Mammalian p50Cdc37 is a protein kinase-targeting subunit of Hsp90 that binds and stabilizes Cdk4.* Genes Dev, 1996. **10**(12): p. 1491-502.
75. Stancato, L.F., et al., *Raf exists in a native heterocomplex with hsp90 and p50 that can be reconstituted in a cell-free system.* J Biol Chem, 1993. **268**(29): p. 21711-6.
76. Taipale, M., et al., *Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition.* Cell, 2012. **150**(5): p. 987-1001.
77. Xu, W. and L. Neckers, *The double edge of the HSP90-CDC37 chaperone machinery: opposing determinants of kinase stability and activity.* Future Oncol, 2012. **8**(8): p. 939-42.
78. Butler, L.M., et al., *Maximizing the Therapeutic Potential of HSP90 Inhibitors.* Mol Cancer Res, 2015. **13**(11): p. 1445-51.
79. Miyata, Y., H. Nakamoto, and L. Neckers, *The therapeutic target Hsp90 and cancer hallmarks.* Curr Pharm Des, 2013. **19**(3): p. 347-65.
80. Paraiso, K.H., et al., *The HSP90 inhibitor XL888 overcomes BRAF inhibitor resistance mediated through diverse mechanisms.* Clin Cancer Res, 2012. **18**(9): p. 2502-14.
81. Holmes, J.L., et al., *Silencing of HSP90 cochaperone AHA1 expression decreases client protein activation and increases cellular sensitivity to the HSP90 inhibitor 17-allylamino-17-demethoxygeldanamycin.* Cancer Res, 2008. **68**(4): p. 1188-97.
82. Wang, H.G., et al., *Bcl-2 interacting protein, BAG-1, binds to and activates the kinase Raf-1.* Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7063-8.
83. Botstein, D., S.A. Chervitz, and J.M. Cherry, *Yeast as a model organism.* Science, 1997. **277**(5330): p. 1259-60.
84. Miklos, G.L. and G.M. Rubin, *The role of the genome project in determining gene function: insights from model organisms.* Cell, 1996. **86**(4): p. 521-9.
85. Liu, W., et al., *From Saccharomyces cerevisiae to human: The important gene co-expression modules.* Biomed Rep, 2017. **7**(2): p. 153-158.
86. Guelzim, N., et al., *Topological and causal structure of the yeast transcriptional regulatory network.* Nat Genet, 2002. **31**(1): p. 60-3.
87. Serin, E.A., et al., *Learning from Co-expression Networks: Possibilities and Challenges.* Front Plant Sci, 2016. **7**: p. 444.
88. Stuart, J.M., et al., *A gene-coexpression network for global discovery of conserved genetic modules.* Science, 2003. **302**(5643): p. 249-55.
89. Jiang, J., et al., *Construction and application of a co-expression network in Mycobacterium tuberculosis.* Sci Rep, 2016. **6**: p. 28422.
90. Yu, H., et al., *NetMiner-an ensemble pipeline for building genome-wide and high-quality gene co-expression network using massive-scale RNA-seq samples.* PLoS One, 2018. **13**(2): p. e0192613.

91. Saelens, W., R. Cannoodt, and Y. Saeys, *A comprehensive evaluation of module detection methods for gene expression data.* Nat Commun, 2018. **9**(1): p. 1090.
92. Liu, W., et al., *Construction and Analysis of Gene Co-Expression Networks in Escherichia coli.* Cells, 2018. **7**(3).
93. Szklarczyk, D., et al., *The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible.* Nucleic Acids Res, 2017. **45**(D1): p. D362-D368.
94. Papsdorf, K., et al., *Polyglutamine toxicity in yeast induces metabolic alterations and mitochondrial defects.* BMC Genomics, 2015. **16**: p. 662.
95. Morgner, N. and C.V. Robinson, *Massign: an assignment strategy for maximizing information from the mass spectra of heterogeneous protein assemblies.* Anal Chem, 2012. **84**(6): p. 2939-48.
96. Plum, S., et al., *Combined enrichment of neuromelanin granules and synaptosomes from human substantia nigra pars compacta tissue for proteomic analysis.* J Proteomics, 2013. **94**: p. 202-206.
97. van Zundert, G.C.P., et al., *The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes.* J Mol Biol, 2016. **428**(4): p. 720-725.
98. Petti, A.A., et al., *Survival of starving yeast is correlated with oxidative stress response and nonrespiratory mitochondrial function.* Proc Natl Acad Sci U S A, 2011. **108**(45): p. E1089-98.
99. Tirosh, I., et al., *On the relation between promoter divergence and gene expression evolution.* Mol Syst Biol, 2008. **4**: p. 159.
100. Levy, S., et al., *Strategy of transcription regulation in the budding yeast.* PLoS One, 2007. **2**(2): p. e250.
101. Harris, T.W., et al., *WormBase: a comprehensive resource for nematode research.* Nucleic Acids Res, 2010. **38**(Database issue): p. D463-7.
102. Bolz, D.D., J.L. Tenor, and A. Aballay, *A conserved PMK-1/p38 MAPK is required in caenorhabditis elegans tissue-specific immune response to Yersinia pestis infection.* J Biol Chem, 2010. **285**(14): p. 10832-40.
103. Murphy, C.T., et al., *Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans.* Nature, 2003. **424**(6946): p. 277-83.
104. Sahu, S.N., et al., *Genomic analysis of immune response against Vibrio cholerae hemolysin in Caenorhabditis elegans.* PLoS One, 2012. **7**(5): p. e38200.
105. Shapira, M., et al., *A conserved role for a GATA transcription factor in regulating epithelial innate immune responses.* Proc Natl Acad Sci U S A, 2006. **103**(38): p. 14086-91.
106. Ding, W., et al., *s-Adenosylmethionine Levels Govern Innate Immunity through Distinct Methylation-Dependent Pathways.* Cell Metab, 2015. **22**(4): p. 633-45.
107. Walker, A.K., et al., *A conserved SREBP-1/phosphatidylcholine feedback circuit regulates lipogenesis in metazoans.* Cell, 2011. **147**(4): p. 840-52.
108. Irazoqui, J.E., et al., *Distinct pathogenesis and host responses during infection of C. elegans by P. aeruginosa and S. aureus.* PLoS Pathog, 2010. **6**: p. e1000982.
109. Safra, M., et al., *The FOXO transcription factor DAF-16 bypasses ire-1 requirement to promote endoplasmic reticulum homeostasis.* Cell Metab, 2014. **20**(5): p. 870-881.
110. Schulenburg, H., et al., *Specificity of the innate immune system and diversity of C-type lectin domain (CTLD) proteins in the nematode Caenorhabditis elegans.* Immunobiology, 2008. **213**(3-4): p. 237-50.
111. Sivamaruthi, B.S. and K. Balamurugan, *Physiological and Immunological Regulations in Caenorhabditis elegans Infected with Salmonella enterica serovar Typhi.* Indian J Microbiol, 2014. **54**(1): p. 52-8.
112. Zhang, Y., et al., *SEPA-1 mediates the specific recognition and degradation of P granule components by autophagy in C. elegans.* Cell, 2009. **136**(2): p. 308-21.
113. Tian, Y., et al., *C. elegans screen identifies autophagy genes specific to multicellular organisms.* Cell, 2010. **141**(6): p. 1042-55.

114. Sato, K., et al., *Dynamic regulation of caveolin-1 trafficking in the germ line and embryo of Caenorhabditis elegans.* Mol Biol Cell, 2006. **17**(7): p. 3085-94.
115. McKay, S.J., et al., *Gene expression profiling of cells, tissues, and developmental stages of the nematode C. elegans.* Cold Spring Harb Symp Quant Biol, 2003. **68**: p. 159-69.
116. Dong, M.Q., et al., *Quantitative mass spectrometry identifies insulin signaling targets in C. elegans.* Science, 2007. **317**(5838): p. 660-3.
117. Ding, Y.H., et al., *Characterization of PUD-1 and PUD-2, two proteins up-regulated in a long-lived daf-2 mutant.* PLoS One, 2013. **8**(6): p. e67158.
118. Tepper, R.G., et al., *PQM-1 complements DAF-16 as a key transcriptional regulator of DAF-2-mediated development and longevity.* Cell, 2013. **154**(3): p. 676-690.
119. Li, X., et al., *Activation of the Na+/H+ exchanger in isolated cardiomyocytes through beta-Raf dependent pathways. Role of Thr653 of the cytosolic tail.* J Mol Cell Cardiol, 2016. **99**: p. 65-75.
120. Eckl, J.M., et al., *Cdc37 (cell division cycle 37) restricts Hsp90 (heat shock protein 90) motility by interaction with N-terminal and middle domain binding sites.* J Biol Chem, 2013. **288**(22): p. 16032-42.
121. Lotz, G.P., et al., *Aha1 binds to the middle domain of Hsp90, contributes to client protein activation, and stimulates the ATPase activity of the molecular chaperone.* J Biol Chem, 2003. **278**(19): p. 17228-35.
122. Lee, P., et al., *Sti1 and Cdc37 can stabilize Hsp90 in chaperone complexes with a protein kinase.* Mol Biol Cell, 2004. **15**(4): p. 1785-92.
123. Haslbeck, V., et al., *The activity of protein phosphatase 5 towards native clients is modulated by the middle- and C-terminal domains of Hsp90.* Sci Rep, 2015. **5**: p. 17058.
124. Richardson, J.M., et al., *Cloning, expression and characterisation of FKB-6, the sole large TPR-containing immunophilin from C. elegans.* Biochem Biophys Res Commun, 2007. **360**(3): p. 566-72.
125. Linnert, M., et al., *The FKBP-type domain of the human aryl hydrocarbon receptor-interacting protein reveals an unusual Hsp90 interaction.* Biochemistry, 2013. **52**(12): p. 2097-107.
126. Vaughan, C.K., et al., *Structure of an Hsp90-Cdc37-Cdk4 complex.* Mol Cell, 2006. **23**(5): p. 697-707.
127. Morgner, N., et al., *A novel approach to analyze membrane proteins by laser mass spectrometry: from protein subunits to the integral complex.* J Am Soc Mass Spectrom, 2007. **18**(8): p. 1429-38.
128. Peetz, O., et al., *LILBID and nESI: Different Native Mass Spectrometry Techniques as Tools in Structural Biology.* J Am Soc Mass Spectrom, 2019. **30**(1): p. 181-191.
129. Verba, K.A., et al., *Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase.* Science, 2016. **352**(6293): p. 1542-7.
130. Sinars, C.R., et al., *Structure of the large FK506-binding protein FKBP51, an Hsp90-binding protein and a component of steroid receptor complexes.* Proc Natl Acad Sci U S A, 2003. **100**(3): p. 868-73.
131. Roskoski, R., Jr., *RAF protein-serine/threonine kinases: structure and regulation.* Biochem Biophys Res Commun, 2010. **399**(3): p. 313-7.
132. da Rocha Dias, S., et al., *Activated B-RAF is an Hsp90 client protein that is targeted by the anticancer drug 17-allylamino-17-demethoxygeldanamycin.* Cancer Res, 2005. **65**(23): p. 10686-91.
133. Polier, S., et al., *ATP-competitive inhibitors block protein kinase recruitment to the Hsp90-Cdc37 system.* Nat Chem Biol, 2013. **9**(5): p. 307-12.
134. Shao, J., et al., *Functional dissection of cdc37: characterization of domain structure and amino acid residues critical for protein kinase binding.* Biochemistry, 2003. **42**(43): p. 12577-88.
135. Sasaki, H. and H. Uemura, *Influence of low glycolytic activities in gcr1 and gcr2 mutants on the expression of other metabolic pathway genes in Saccharomyces cerevisiae.* Yeast, 2005. **22**(2): p. 111-27.

136. Wade, J.T., D.B. Hall, and K. Struhl, *The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes.* Nature, 2004. **432**(7020): p. 1054-8.
137. Teixeira, M.C., et al., *The YEASTRACT database: an upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae.* Nucleic Acids Res, 2014. **42**(Database issue): p. D161-6.
138. Gaiser, A.M., et al., *Downregulation of the Hsp90 system causes defects in muscle cells of Caenorhabditis elegans.* PLoS One, 2011. **6**(9): p. e25485.
139. Guisbert, E., et al., *Identification of a tissue-selective heat shock response regulatory network.* PLoS Genet, 2013. **9**(4): p. e1003466.
140. Mallo, G.V., et al., *Inducible antibacterial defense system in C. elegans.* Curr Biol, 2002. **12**(14): p. 1209-14.
141. Singh, V. and A. Aballay, *Regulation of DAF-16-mediated Innate Immunity in Caenorhabditis elegans.* J Biol Chem, 2009. **284**(51): p. 35580-7.
142. Evans, E.A., W.C. Chen, and M.W. Tan, *The DAF-2 insulin-like signaling pathway independently regulates aging and immunity in C. elegans.* Aging Cell, 2008. **7**(6): p. 879-93.
143. Marudhupandiyan, S. and K. Balamurugan, *Intrinsic JNK-MAPK pathway involvement requires daf-16-mediated immune response during Shigella flexneri infection in C. elegans.* Immunol Res, 2017. **65**(3): p. 609-621.
144. Czemeres, J., K. Buse, and G.M. Verkhivker, *Atomistic simulations and network-based modeling of the Hsp90-Cdc37 chaperone binding with Cdk4 client protein: A mechanism of chaperoning kinase clients by exploiting weak spots of intrinsically dynamic kinase domains.* PLoS One, 2017. **12**(12): p. e0190267.

8 Acknowledgement

First of all I would like to thank Prof. Dr. Johannes Buchner for offering me the chance to finish my doctorate studies in his open and creative group, which inspired and supported me over the years of studies.

The next I would like to thank my supervisor Dr. Klaus Richter for providing attractive projects, for all intensive, productive and creative progress discussions; for offering suggestions when things didn't work as expected and meanwhile supporting the freedom to explore unknown scientific world.

Many thanks to all the labs and organizations I cooperated with during the doctorate studies. Thanks to Dr. Thomas Stempfl of the Kompetenzzentrum Fluoreszente Bioanalytik (KFB) for the microarray measurement; to Wang Chen and Qian Zhu from modSEEK (Princeton University) for providing coexpression relationships to feed into the ClusterEx program; to Theresa Stiernagle from CGC who sent out the strains for this study and further to the groups of Prof. Sivan Henis-Korenblit (Bar-Ilan University, Israel), Prof. Dong Meng-Qiu (National Institute of Biological Sciences, Bejing) and Prof. Frederick Ausubel (Harvard University) for sharing the nematode strains used in this study; to Dr. Katalin Barkovits, Dr. Stephan Hacker, Dr. Nils Hellwig and Prof. Nina Morgener for performing native/ crosslinked MS measurement; to Dr. Christoph Kaiser and Prof. Sevil Weinkauf for helping with electron microscope imaging and offering advises in data analysis.

Furthermore I would like to thank my lovely colleagues and all practical and Bachelor students for their supporting during my studies. Julia and Katha, thanks for your advises and tips which enlightened my first years; Lukas, thanks for organizing a orientated and functional lab and your help in improving my German level; Anna, thanks for creative ideas and diverse discussions in daily research and many thanks to all colleagues who offered advises and help within these years.

Also I want to thank my family and friends for their support, encourage and love. I enjoyed the time spent with you.

9 Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Diese Arbeit wurde bisher keiner Prüfungskommission vorgelegt. Teile dieser Arbeit wurden oder werden in wissenschaftlichen Journalen veröffentlicht.

Garching,

Siyuan Sima