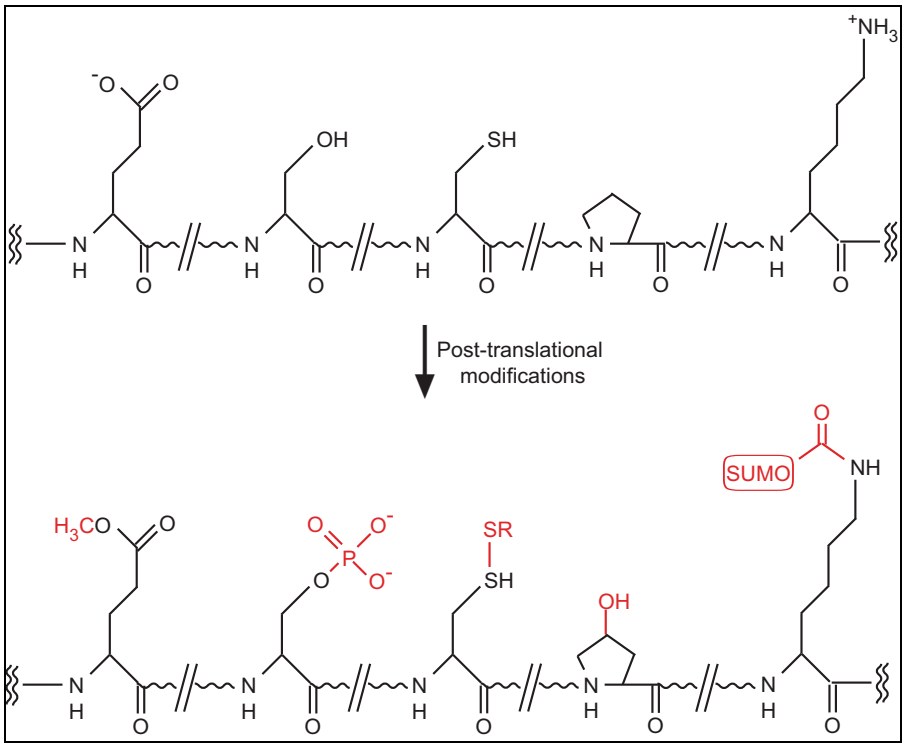
Universität Bremen

Post-translational modifications of Hepatoma-derived growth factor (HDGF)



Ketan Thakar

Universität Bremen
2010

Post-translational modifications of Hepatoma-derived growth factor (HDGF)

Dissertation submitted as a partial fulfillment for procuring the degree Doctor of Natural Science (*Dr. rer. nat.*)

Submitted to

Fachbereich 2 Biologie/Chemie
Universität Bremen

Submitted by

Ketan Thakar

M.Sc. Biochemistry and Molecular Biology

Bremen 2010

Examination Committee

Reviewers:

1. **Prof. Dr. Sørge Kelm**
Universität Bremen, Fachbereich 2, Biochemie
Postfach 330440, 28334 Bremen

2. **Dr. Kathrin Mädler**
Universität Bremen, Fachbereich 2, Biochemie
Leobener Straße NW2, B2060, 28359 Bremen

Examiners:

1. **Dr. Frank Dietz**
Universität Bremen, Fachbereich 2, Biochemie
Postfach 330440, 28334 Bremen

2. **Prof. Dr. Reimer Stick**
Universität Bremen, Fachbereich 2, Institut für Zellbiologie
Leobener Straße NW2 A3290, 28359 Bremen

Date of Public defense:

28 May, 2010

Table of Contents

I. Summary

Post-translational modifications (PTM's) are modifications that occur during or after protein translation. Nascent or folded protein can be subjected to an array of specific enzyme-catalyzed modifications on the amino acid side chains or the peptide backbone. Two broad categories of protein PTM's occur; the first includes all enzyme-catalyzed covalent additions of different lower molecular chemical groups up to complex proteins to amino acid side chains in the target protein, whereas the second category comprises structural changes and the cleavage of peptide backbones in proteins either by action of proteases or, less commonly, by autocatalytic cleavage. PTM's can modulate the function of proteins by altering their activity state, localization, turnover, and interactions with other proteins.

Hepatoma-derived growth factor (HDGF) is the prototype of a family of six proteins comprising HDGF, the four HDGF-related proteins (HRP-1–4), and the lens epithelium-derived growth factor (LEDGF). HDGF exhibits growth factor properties and has been implicated in organ development and tissue differentiation of the intestine, kidney, liver, and cardiovascular system. Recently, the role of HDGF in cancer biology has become a main focus of its research. HDGF was found to be over-expressed in a large number of different tumor types. Although a direct influence of HDGF on tumor biology is still unclear, its expression is correlated with metastasis and tumor recurrence in multiple studies. HDGF appears to be a novel prognostic marker for different types of cancer. Growth promoting as well as other activities of HDGF, like the suppression of differentiation; possible role in apoptotic processes; or its angiogenic properties have been suggested to play a role in tumor induction and/or cancer progression.

Interestingly, until now, very little is known about how PTM's are involved in modulating HDGF function. Therefore, the main aim of the thesis was the identification of HDGF PTM's and their consequence on its function. At first, we have identified that HDGF is post-translationally modified by SUMO-1 at a non-consensus site and SUMOylated HDGF does not associate with chromatin in contrast to the unSUMOylated form. Further, we show that HDGF secretion is regulated by the presence or absence of a serine phosphorylation site and loss in secretion is attributed to N-terminal processing of the protein. Additionally, the two cysteine residues in HDGF are involved in formation of intra-and inter-molecular disulfide bonds and N-terminal processing favors dimer formation. Furthermore, we demonstrated the presence of HRP-2 isoforms with their developmentally regulated expression in different rat brain regions and were able to show that HDGF can interact with HRP-2. Moreover, a new HRP-2 isoform can exclusively interact and enrich a processed form of HDGF. Finally, we identified the presence of a functional caspase cleavage site in the C-terminal region of murine HDGF. The results complied provide substantial new knowledge concerning PTM's of HDGF and affected mechanisms, and will help to add new perspectives for research in the field of HDGF and its related proteins.

II. Zusammenfassung

Posttranslationale Modifikationen (PTM's) sind Modifikationen, die während oder nach erfolgter Proteintranslation auftreten. Das entstehende oder schon gefaltete Protein kann an verschiedenen Aminosäureseitenketten oder den Peptidbindungen über eine Vielzahl von Enzym-katalysierten Reaktionen modifiziert werden. Im wesentlichen können zwei Kategorien von PTM's auftreten; die erste umfasst alle Enzym-katalysierten, kovalenten Bindungen verschiedener niedermolekularer, chemischer Gruppen bis hin zu komplexen Proteinen an Aminosäureseitenketten von Zielproteinen, während die zweite Kategorie Reaktionen umfasst, die zu Strukturänderungen oder zur Spaltung von Peptidbindungen durch Proteasen oder weniger häufig, durch Autokatalyse führen. PTM's können die Funktion von Proteinen modulieren indem sie deren Aktivitätstatus, die Lokalisation, Halbwertszeit und z.B. die Interaktion mit anderen Proteinen beeinflussen.

Der Hepatoma-derived growth factor (HDGF) ist der Prototyp einer Proteinfamilie, die sechs Mitglieder umfasst und zu denen neben HDGF die vier HDGF-ähnlichen Proteine (HRP-1–4) und der lens epithelium-derived growth factor (LEDGF), gehören. HDGF zeigt Wachstumsfaktor-Eigenschaften und ist in Zusammenhang mit der Organ-Entwicklung und Gewebe-Differenzierung im Dünndarm, Niere, Leber und dem Kardiovaskulären-System, gebracht worden. Ein weiterer Fokus der HDGF-Forschung befasst sich mit der Rolle des Proteins in der Krebsbiologie. Es konnte gezeigt werden, dass HDGF in einer Vielzahl verschiedener Tumor-Typen überexprimiert wird. Obwohl der direkte Einfluss von HDGF auf die Tumorentwicklung noch immer unklar ist, korreliert die Expression des Proteins mit dem Auftreten von Metastasen und dem erneuten Auftreten von Tumoren in mehreren Studien. HDGF scheint damit ein neuer prognostischer Marker für verschiedene Krebsformen zu sein. Seine Wachstums-fördernde, als auch andere Eigenschaften, wie die Unterdrückung der Differenzierung, eine mögliche Rolle in apoptotischen Prozessen oder seine angiogenen Eigenschaften wurden mit einer möglichen Rolle bei der Tumor-Induktion und /oder bei dem Fortschreiten der Krebsentwicklung, in Zusammenhang gebracht.

Weil bis heute wenig aufgeklärt ist, wie PTM's die Funktionalität des HDGF beeinflußen können, lag der Fokus dieser Arbeit in der Identifikation von HDGF PTM's und deren Konsequenzen auf die Protein-Funktion. Zunächst haben wir gezeigt, dass HDGF posttranslational über SUMO-1 an einem nicht-Konsensusmotiv modifiziert werden kann und das SUMOyliertes HDGF im Gegensatz zu der nicht-SUMOylierten Form, nicht mehr an Chromatin bindet. Ferner, konnten wir zeigen, dass HDGF-Sekretion über das Vorhandensein einer vorausgesagten Serin-Phosphorylierungsstelle reguliert wird und das ein Verlust der Sekretion mit einer N-terminalen Prozessierung des Proteins zusammenhängt. Außerdem, sind die zwei Cystein-Reste im HDGF in der Bildung von intra- und intermolekularen Disulfidbrücken involviert und N-terminale Prozessierung unterstützt die Bildung von Homodimeren. Desweiteren, konnten wir das Vorhandensein von verschiedenen HRP-2 Isoformen zeigen

und deren entwicklungsabhängige Expression in verschiedenen Rattengehirn-Regionen. In diesem Zusammenhang, konnten wir nachweisen, dass HDGF mit HRP-2 interagiert und das eine neu identifizierte HRP-2 Isoform exklusiv mit einer prozessierten Form von HDGF interagiert und diese spezifisch anreichern kann. Darüber hinaus, haben wir das Vorhandensein einer funktionellen Caspase Schnittstelle in der C-terminalen Region von HDGF identifiziert.

Die Ergebnisse dieser Arbeit tragen wesentlich zu einer Erweiterung des Wissens um die Funktion von posttranslationalen Modifikationen des HDGF und dadurch beeinflusster Mechanismen bei und helfen damit, in diesem Feld, die Forschung an dem HDGF und den HDGF-ähnlichen Proteinen um neue Ansätze zu bereichern.

1 Introduction

1.1 The HDGF related protein (HRP) family

Hepatoma-derived growth factor (HDGF) and HDGF related proteins (HRPs) belong to a gene family with a well-conserved amino acid sequence at the N-terminus. HDGF forms the prototype of a new family of growth factors called HDGF-related proteins (HRPs), which includes HRP-1, HRP-2, HRP-3, HRP-4 and lens epithelium-derived growth factor (LEDGF/p75) (Table 1.1). For all HRP members the following common features are described: (1) homology in the first N-terminal 98 amino acid residues or the *hath* region (*h*omologous to *a*mino *t*erminus of *H*DGF) (2) a PWWP domain in the *hath* region (3) a canonical bipartite nuclear localization signal in the *non-hath* C-terminal region (4) lack of a hydrophobic signal peptide and (5) altered electrophoretic running behavior.

Table 1.1 Hepatoma-derived growth factor related proteins (HRPs). The number of amino acids shown for the HRPs is representative of human forms.

HRP member	Number of amino acids	Description	Tissue distribution	Reference
HDGF	240	Small, acidic	Ubiquitous	[1]
HRP-1	286	Small, acidic	Only in testis	[2]
HRP-2	672	Large, basic	Ubiquitous	[2]
HRP-3	203	Small, basic	Brain and testis	[3]
LEDGF/p75	530	Large, basic	Ubiquitous	[4]
HRP-4	235	Small, acidic	Only in testis	[5]

Despite the uncertainty surrounding the various functional details, the mitogenic action of HRP proteins has attracted considerable attention and parallels were drawn between the proliferative action of HDGF and its developmental as well as cancer-related roles. In this section, we get an overview on the current state of knowledge about HRP proteins; HDGF structure-function relationship; and the physiological roles as well as pathological relevance of HDGF.

1.1.1 Hepatoma-derived growth factor

Hepatoma-derived growth factor (HDGF) was originally isolated from conditioned medium of the human hepatoma-derived cell line, HuH-7, as a heparin-binding and proliferation promoting factor [6]. Recombinant cloning of HDGF derived from HuH-7 cDNA and protein sequencing showed that HDGF shares homology (23.4% identity and 35.6% homology) with the high mobility group (HMG)-1 protein [1]. HDGF is an acidic polypeptide with mitogenic activity for fibroblasts [1], endothelial cells [7], vascular smooth muscle cells [8], and hepatoma cells [9].

Nuclear targeting of HDGF was reported to be required for mitogenesis [10, 11]. HDGF contains a nuclear localization signal (NLS); ^{155}KRRAGLLEDSPKRPK^{170} (basic residues underlined) {NLS2}, homologous to the reported consensus sequences for a bipartite NLS which consists of two clusters of basic residues separated by 10-12 amino acids including proline residues [1]. Additionally, another NLS like basic amino acid-rich region, ^{75}KPNKRK^{80} (basic residues underlined) {NLS1}, is found in the *hath* region [12]. The cellular trafficking of HDGF is considered atypical, as compared to the majority of described growth factors, due to the fact that HDGF lacks a classical secretion signal but is still able to exit the cell by a yet unknown route. Extracellular HDGF undergoes receptor mediated internalization after which it is targeted to the nucleus via NLS2. The extracellular HDGF receptor and HDGF sub-nuclear target(s) are however not fully described so far.

1.1.2 Hepatoma-derived growth factor related protein-1 (HRP-1)

Izumoto et al. (1997) screened a mouse testis cDNA library with the 0.9 kb cDNA of human HDGF [2]. Isolated clones were grouped into three distinct classes named, HDGF, HRP-1 and HRP-2 based on sequence relationship. HRP-1 was observed to be 46 amino acid longer as compared to HDGF, showing a significantly high negative net charge (Lys + Arg: 8.4%, Glu +Asp: 26.4%) and is rich in proline residues (9.5%). The tissue specific expression pattern of HRP-1 shows expression restricted only to testis. Interestingly, nothing is known about the function of this HRP family member.

1.1.3 Hepatoma-derived growth factor related protein-2 (HRP-2)

With an open reading frame of 2007 bp, encoding a protein of 669 amino acids, HRP-2 is the largest member of the HRP family and has a predicted M_r of ~74 kDa and pI 9.14 [2]. HRP-2 shows significant homology to the N-terminal region of HDGF (68%) within the 98-amino acid *hath* region. Found on chromosome 17 in mice, the HRP-2 gene encodes 16 exons and is expressed in a wide variety of tissues. It carries a high level of basic and acidic residues, with Lys + Arg making nearly 20.6% and Glu + Asp nearly 19.7% of the total protein. HRP-2 is also rich in proline and serine residues which is characteristic of nuclear proteins [2] and hints towards a nuclear site of action of the protein. Another feature of HRP-2 is the presence of a mixed charge cluster, a segment of the protein with alternating positive and negative residues. Eukaryotic regulatory proteins, including transcription and replication factors, are characterized

by the presence of charge clusters; and it may well be possible that HRP-2 protein functions as a regulatory protein [2]. A novel domain, the integrase binding domain (IBD), has been discovered in HRP-2. The ~80 amino acid residue spanning domain is necessary to bind HIV-1 integrase (HIV-1 IN). Another HRP family member, LEDGF (p75) contains the homologous sequence, and the IBDs of HRP-2 as well as LEDGF/p75 are capable of binding and stimulating HIV-1 IN *in vitro* [13]. However, unlike LEDGF, HRP-2 lacks the ability to tether HIV-1 IN to chromatin. Apart from this function, the molecular functional mechanisms; specificity or regulation; and physiological relevance of HRP-2 is largely unknown.

1.1.4 Hepatoma-derived growth factor related protein-3 (HRP-3)

Discovered by Ikegame et al. (1999), HRP-3 shows 81.4% identity to the HDGF *hath* region [3]. Mapped to chromosome 15, region q25 in humans, the HRP-3 gene encodes a protein of 203 amino acid residues with a basic iso-electric point. HRP-3, like HDGF, was found to have growth-stimulating activity. With its expression mainly restricted to nervous tissue, HRP-3 may have a specialized function [14]. El-Tahir et al. (2009) showed that HRP-3 is the first member of the HRP family to interact with the cytoskeleton. The results suggest that HRP-3 is involved in neurite growth and exerts its effects through interaction with tubulin and microtubules. Also, HRP-3 is redistributed from an extra-nuclear to a nuclear localization during development suggesting a role of HRP-3 during neuritogenesis. The neuritogenic effect is due to the interaction of HRP-3 with tubulin dimers and assembled microtubules. It promotes tubulin assembly into microtubules and stabilizes microtubules once they have been formed. In addition, HRP-3 promotes bundling of microtubules. The stabilization and bundling of cytoskeletal components was attributed to the dimerisation of HRP-3 [15].

1.1.5 Hepatoma-derived growth factor related protein-4 (HRP-4)

HRP-4 was discovered by Dietz et al. (2002) from a bovine cDNA library and codes for a 36 kDa protein with 235 amino acids [5]. HRP-4, like HDGF, can bind to the glycoaminoglycans heparin and heparan sulphate and shows growth promoting activity. Amino acid sequence comparison of the N-terminal 91 amino acids of bHRP-4 to other HRP members showed that it shares 86 and 89% amino acids with HDGF and mHRP-1 respectively, but only 70 and 69% amino acids with mHRP-2/hHRP-3 and LEDGF respectively. Like HRP-1, HRP-4 is exclusively expressed in testis but its function still remains unclear.

1.1.6 Lens epithelium derived growth factor (LEDGF)

LEDGF is an ubiquitously expressed nuclear protein, tightly associated with chromatin throughout the cell cycle. Chromatin association is primarily mediated by three conserved sequence elements within the N-terminal half of the protein: the PWWP domain, nuclear localization signal (NLS), and a dual copy of the AT-hook DNA binding motif. Recently, study revealed that the association with chromatin is essential for LEDGF function during HIV-1 infection, highlighting its significance in chromatin binding [16]. Like HRP-2

LEDGF contains a second evolutionarily conserved domain within its extended C-terminus. It is this domain (IBD) that mediates the interaction with HIV-1 integrase. Two alternatively spliced isoforms, LEDGF/p75 and LEDGF/p52 are expressed from the same gene (human *PSIP1*). The smaller p52 isoform lacks the IBD and fails to engage HIV-1 IN *in vitro* or in living cells. The cellular functions of LEDGF remain largely uncharacterized, although initial reports have indicated a role for LEDGF/p75 in transcriptional regulation. LEDGF is not essential for cell survival, although the majority of LEDGF-null mice died soon after birth or showed a range of developmental abnormalities in adulthood [17]. LEDGF also binds to heparin which protects it from heat, acid-base deactivation, and proteolytic degradation with trypsin and chymotrypsin. Heparin-bound LEDGF greatly potentiates survival of mouse LECs in culture [18].

1.2 HDGF structure-function relationship

Analysis of the HDGF structure showed that it contains two structurally independent regions, (i) a N-terminal PWWP domain within the *hath* region, and (ii) an unstructured C-terminal part [19].

1.2.1 HDGF is a modular protein with two structurally independent domains

The few functional data available on HDGF gave a clue that it might be a modular protein in which the N- and C-terminal regions have distinctive roles. The N-terminal region was shown to be responsible for cell surface binding, promoting proliferation through triggering a membrane receptor-mediated pathway [20] and HDGF internalization [11], while the C-terminal region was found to be responsible for mediating mitogenic effects after nuclear entry [11]. A functional distinction between the N-and C-terminal regions of HDGF was supported by a recent study demonstrating, that the N-terminal *hath* region of human HDGF is a well structured domain while the C-terminal *non-hath* region is disordered and does not pursue any defined secondary structure [21]. Moreover, the ordered and disordered structures of the N- and C-terminal regions, respectively, were shown to be independent of linkage of the two regions within the full length HDGF or their presence in two separate constructs [21].

1.2.2 HDGF PWWP domain

HRPs are members of a family of proteins with a conserved PWWP domain. The PWWP domain was first characterized from the *WHSC1* gene. It contains a conserved 70 amino acid sequence and has been found in around 60 eukaryotic proteins. Initially, the PWWP domain was hypothesized to be a site for protein-protein interaction [22]. However, the PWWP domain of DNA-methyltransferase 3b (Dnmt3b) has been shown to interact with DNA [23-25]. In addition to the PWWP domain, proteins in this family are frequently known to contain chromatin association domains such as the bromodomain, chromodomain, SET domain, and the Cys-rich Zn-binding domains [22]. This strongly suggests a role in chromatin

regulation or modification of these proteins. The structures of three PWWP-domain containing proteins, Dnmt3b (PDB ID: 1KHC) [23], *S. pombe* protein SPBC215.07c (PDB ID: 1H3Z) [26], and mouse HRP-3 (mHRP-3) (PDB ID: 1N27) have been previously determined. In HDGF, the identified PWWP domain spans from amino acid residues Cys_{12} to Ser_{68} and the corresponding motif here is PHWP (one letter code for the sequence ^{24}Pro-His-Trp-Pro^{27}) [22]. Sequence alignment (Figure 1.1) and structural comparison showed that these four domains are folded similarly, all containing a five-strand anti-parallel β-barrel and α-helical structures.

The secondary structure of the HDGF PWWP domain (Figure 1.2a) consists of five β-strands, followed by two α-helices. The stereo pair of the ribbon structure of the domain is shown in Figure 1.2b. A high-resolution solution structure of the N-terminal PWWP domain of HDGF using NMR spectroscopy has been determined [19, 27]. The five anti-parallel β-strands form a highly conserved β barrel (β1, 15-18; β2, 26-32; β3, 45-49; β4, 53-58; β5, 63-65). Strands β1 and β2 are linked by a β-β arch (MKGYP), which allows the side chains of Met_{20} and Tyr_{23} to point to the core of the barrel. The PHWP motif constitutes the N-terminal part of the strand β2. Pro_{24} forms a β-bulge, which along with the bend created by Pro_{27}, is necessary for the formation of the β-barrel. In addition, the indole ring of the Trp_{26} packs against the helix A, while the side chain of His_{25} packs against the side chain of highly conserved Lys_{19} residue. Strands β2 and β3 are linked by the long, poorly defined L2 loop, which is also the most variable region in all the HRP's [5]. Pro_{60} in the L3 loop assumes a trans-conformation to rotate the backbone orientation by 90º, thus permitting the formation of hydrogen bonds between strands β5 and β1. Helix A connects to β5 by another trans-proline residue at position 65. Helix B, which is oriented approximately anti-parallel to helix A, is connected to helix A by a seven residue linker.

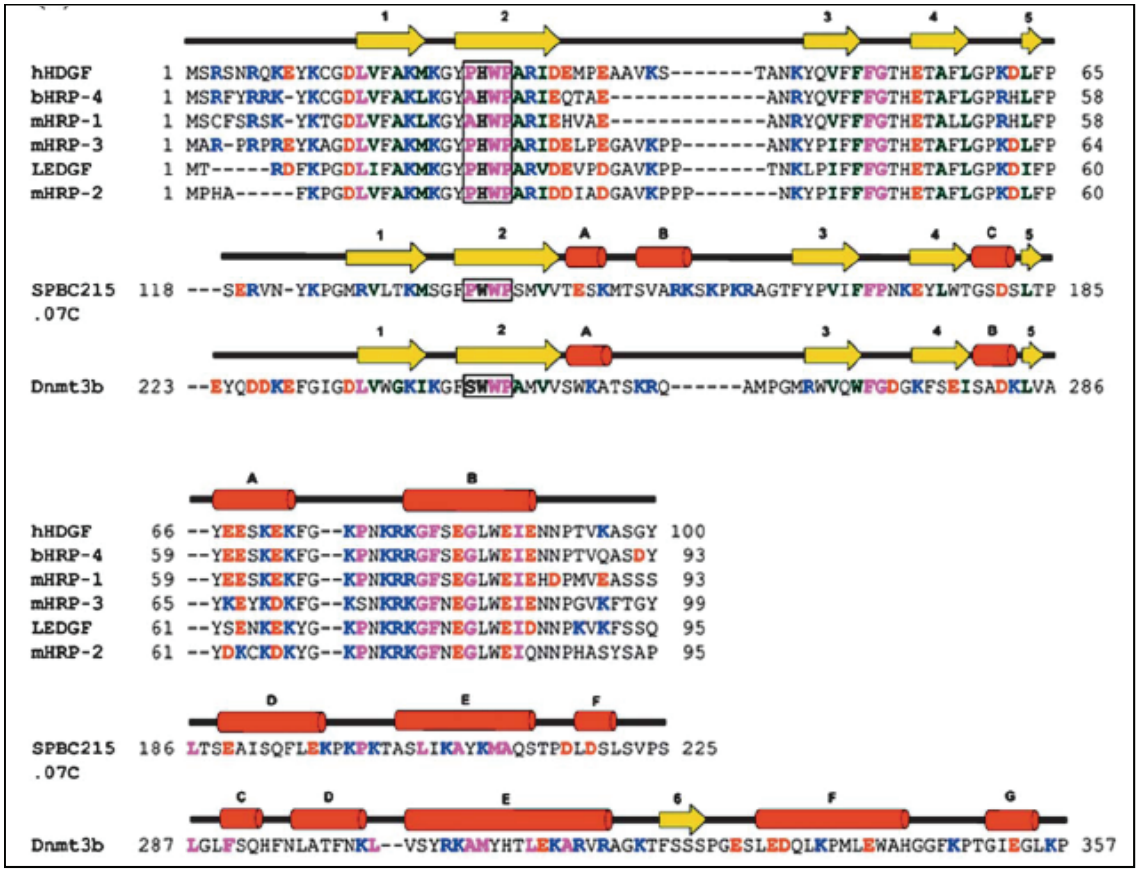


Figure 1.1 Sequence alignment of the *hath* regions of HRPs (hHDGF, mHRP-1, mHRP-2, mHRP-3, bHRP-4, and LEDGF), *S. pombe* protein SPBC215.07C (residues 118–225), and Dnmt3b (residues 223–357). The color codes are: blue, positively charged residues; red, negatively charged residues. Deletions and insertions are indicated by dashes. The characteristic PWWP motifs are boxed. The numbers and letters above the secondary structure elements are shown as arrows for strands and cylinders for helices. (Adapted from [19] without permission)

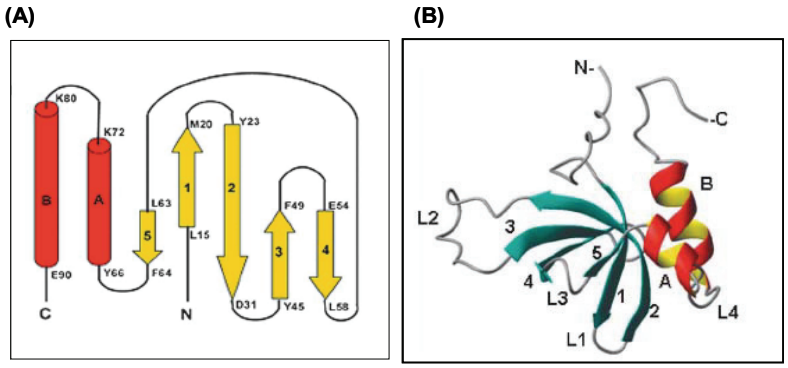


Figure 1.2 (a) Secondary structure topology of the *hath* region of human HDGF. (b) The stereo pairs of the ribbon representation of the structure of the *hath* domain of human HDGF (Adapted from [19] without permission).

1.2.2a HDGF binds to DNA through the N-terminal PWWP domain

Functionally, most of the PWWP family proteins are involved in chromatin remodeling. The role of the PWWP domain in HDGF was largely unknown until recently, it has been reported that only the domain is necessary and sufficient to mediate DNA binding [28]. The study demonstrated that DNA/HDGF PWWP domain interaction is unique, because the minimum required DNA binding element is only 37 bp. Furthermore, HDGF functions in the nucleus as a direct DNA binding protein to repress the expression of specific target genes involved in cell proliferation and differentiation.

1.2.2b HDGF PWWP domain as a potential protein-protein interaction domain

Due to its position at either the N- or C-terminus, the composition of amino acids close to the PWWP motif, and the pattern of further domains, it has been suggested that the PWWP domain plays a role in protein-protein interaction [22]. This may have an influence on regulating the specificity of a protein to its substrate. The PWWP domain may play an important role in the interaction among HRP family members as well as their interaction with other proteins. These protein interactions might control a variety of cellular processes including protein transport and signaling.

1.2.3 HDGF and heparin binding specificity

Heparin and heparan sulfate are complex linear polymers consisting disaccharide repeats of alternating uronic acid and glucosamine. NMR based approach showed that the heparin-derived hexasaccharide appears to bind to the *hath* region with lower affinity compared to low molecular mass heparin in SPR experiments [19]. The binding site identified by chemical-shift perturbation mapped a highly positively charged patch, formed by the β1-β2 turn and the L3, L4 loops. The positively charged residues proposed to be involved in heparin binding are Lys_{19}, Lys_{61}, Lys_{72}, Lys_{78}, Arg_{79} and Lys_{80} and are well conserved in the *hath* region of all HRP's. A recent study showed that the bacterial expressed *hath* region of human HDGF can form dimers which bind to heparin with two orders of magnitude higher affinity than the monomer [29]. Figure 1.3 shows the proposed secondary structure topology of the domain swapped PWWP module dimer. The dimer is formed by exchanging the N-terminal β1-β2 hairpins of the two monomers which are linked by two long L2 loops. The PWWP dimer uses the same binding region to interact with heparin oligomers, and two heparin binding regions in dimer can form a contiguous binding site. Thus, the enhanced heparin binding affinity can be attributed to a multi-valence effect due to binding of longer heparin oligomer. In heparin free "open conformation" (Figure 1.4a), each monomer is capable of binding to the heparin molecule. Oligosaccharide molecule binds to one monomer (Figure 1.4b) and places itself in the vicinity of the other monomer, the flexible linker allows the free protomer to adjust its binding site in a favorable orientation for interacting with free oligosaccharide units, resulting in a "closed conformation", with the binding sites of both monomers lined up in a contiguous manner (Figure 1.4c, d).

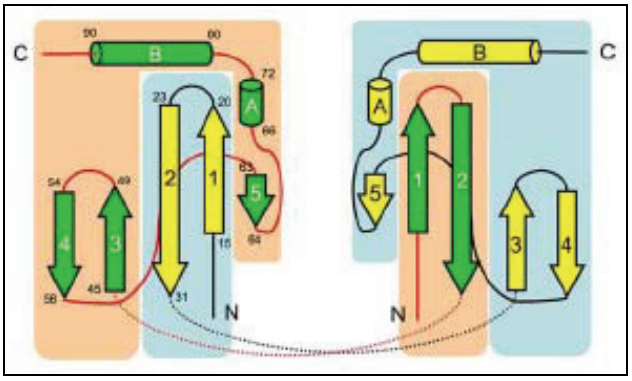


Figure 1.3 Secondary structure topology of the domain swapped PWWP dimer. The domain swapping involves exchange of the two strands β1 and β2. Individual monomer molecules are represented by green and yellow ribbon diagrams. (Adapted from [29] without permission).

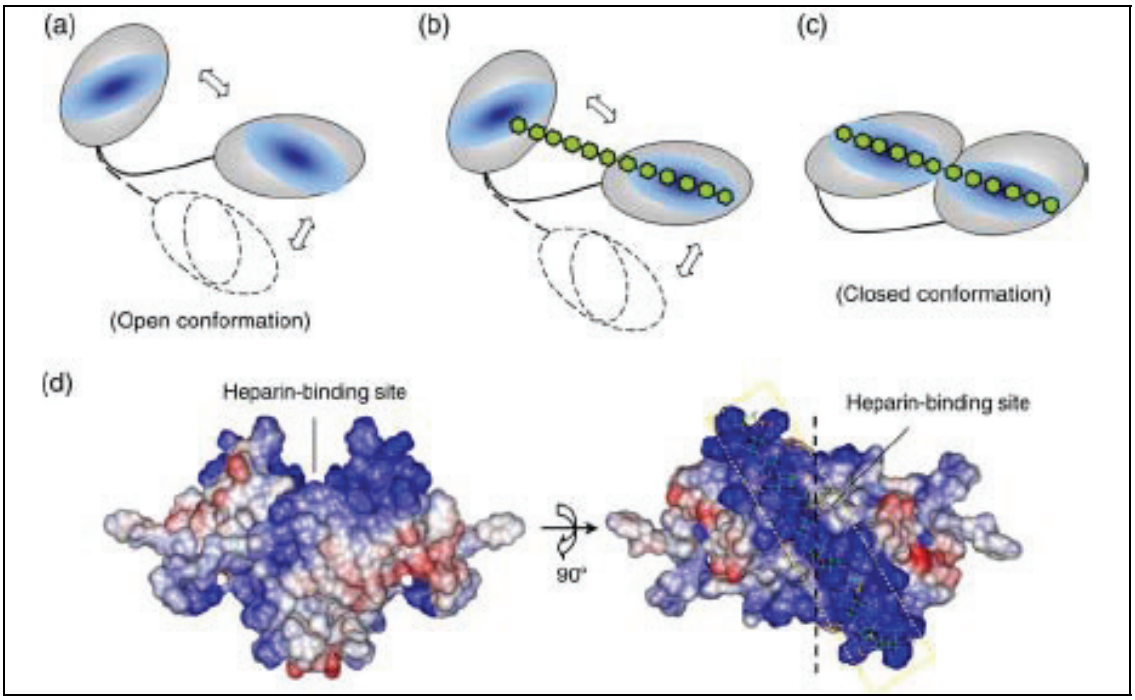


Figure 1.4 Model illustrating the enhanced heparin binding affinity for the PWWP dimer. (a) The "open conformation" of the free PWWP dimer, which possesses one heparin binding site capable of binding a heparin molecule of six units long on each of the monomers. (b) The intermediate conformation of the PWWP dimer with a heparin molecule binding to one protomer. (c) The closed conformation of the PWWP dimer, with both protomers bound to the heparin molecule. (d) Two views of the surface charge distributions of the PWWP dimer in a computer-modeled closed conformation. The positive and negative potentials are colored blue and red, respectively. A heparin-derived dodecasaccharide is docked onto the putative heparin-binding site (dotted box) to indicate the size of the binding pocket (right). (Adapted from [29] without permission).

1.3 Biological roles of HDGF

The first biological effect described for HDGF was its ability to stimulate DNA synthesis in Swiss 3T3 cells as assessed by the rate of incorporation of (^{3}H) thymidine [1]. Due to the fact that HDGF was first isolated from a cancerous cell line, led to intensive investigations that mainly dealt with; (i) links between HDGF under pathological conditions that are related to abnormal cellular proliferation such as tumorigenesis and neoplastic angiogenesis, (ii) the relation between the proliferative action of HDGF and development (iii) its role in organ remodeling after injury (iv) its involvement in apoptosis. Other biological functions were also attributed to HDGF like serving as a neurotrophic factor preventing neuron death [30] or as trophic factor for motor neurons up-regulated in the spinal cord of transgenic mice before onset of degeneration [31].

1.3.1 HDGF in cancer development, prognosis and diagnosis

Several studies implicated HDGF in the development and progression of neoplasias. A strong correlation was observed between the up-regulation of HDGF and hepatic malignancy [32, 33]. Elevated level of HDGF was detected in the highly proliferative hepatocellular carcinoma compared to adjacent normal hepatocytes, and was associated with loss in differentiation [32]. It was suggested that HDGF exerts such an effect on hepatic cells through a combination of autocrine and paracrine actions [33]. Such data are also supported by an earlier study showing that HDGF antisense oligonucleotides inhibit the growth of cultured hepatoma cells [9]. Such a correlation, similar to that with hepatocellular carcinoma, was also found between HDGF and melanoma [34]. Using a functional proteomics approach, HDGF was found to be one of eight candidates that are differentially regulated in transformed cells, where the levels of HDGF expression increases in early and late stage melanomas compared to low levels in normal melanocytes [34]. The correlative behavior between the increased HDGF levels and cancer development encouraged efforts attempting to establish HDGF as a diagnostic and prognostic tool. It was possible to show that HDGF is a strong prognostic predictor for patients with non-small cell lung cancer (NSCLC). Two research groups demonstrated that there is a direct and strong correlation between the increased levels of HDGF expression (immunohistochemically assessed) and the poor overall survival of NSCLC patients after curative therapy [35, 36]. In another study it was shown that HDGF is one of the proteins that can be useful in the determination of the regressive (where HDGF is down-regulated) versus progressive status of tumors [37]. Furthermore, HDGF expression significantly increases in human colorectal cancers, especially in tumors capable in DNA mismatch repair, and thus represented a novel marker for specific tumor subtypes [38]. Different studies established association of HDGF with the recurrence and prognosis of gastric carcinoma and pancreatic ductal carcinoma [39, 40].

The ability of HDGF to induce tumorigenesis *in vivo* is attributed direct angiogenic activity and stimulates the potent angiogenic factor VEGF (vascular endothelial growth factor). Through this combined action, HDGF is supposed to assist in the formation of new blood vessels that are required for tumor

development and progression [41, 42]. Such findings, on the angiogenesis-mediated role of HDGF in tumorigenesis, were also supported by data showing HDGF stimulated proliferation of endothelial cells, including those of the kidney and the cardiovascular systems [7, 43]. However, the molecular mechanism of HDGF in cancer incidence and development are still unknown. One study illustrated that the independent abnormal expression of HDGF can promote cell proliferation but was insufficient to cause cancer. Increasing rates in carcinogenesis were attributed to the cumulative effects produced by aberrant cell signal transduction activated by gene mutation and epigenetic modification. Supporting this hypothesis it has been shown that HDGF over-expression activates Erk1/2 and promotes growth of human gastric cancer AGS cells in an anchorage-independent manner [44].

1.3.2 HDGF and its role in developmental regulation

Many antigens are highly and broadly expressed during embryogenesis and then normally show lower levels and limited distribution after birth, to be only again highly expressed in case of cancer development or in response to tissue damage. HDGF was observed to be developmentally regulated and plays a putative role in renal, cardiovascular, hepatic and intestinal development. HDGF is widely distributed in the kidney at early stages of development but eventually becomes confined to sites of active morphogenesis and, except for renal tubules, disappears from the adult kidney [7]. In another study, HDGF was detected in the nuclei of smooth muscle cells (SMC) and endothelial cells from 19-day fetal, but not in the adult, rat aorta and thus concluded that HDGF helps to regulate SMC growth during development [10]. HDGF is also expressed early in embryonic heart and fetal gut suggesting that it may play a role in cardiovascular growth and differentiation and represented the first described nuclear-targeted mitogen in the developing heart [43]. HDGF expression in hepatocytes decreased with differentiation suggesting that HDGF might also participate in the regulation of hepatocyte proliferation during liver development [45]. Also, HDGF is highly expressed in early gut tissue with dramatically reduced levels after villous epithelial differentiation [38].

1.3.3 HDGF in organ remodeling after injury

Different studies also demonstrated that HDGF play a role in processes of organ repair and remodeling after injury. For example, HDGF helps to regulate the proliferation of SMC in response to vascular injury [8]. Smooth muscles normally show low proliferative rate and their rapid proliferation and migration after vascular injury might result in narrowing of the vascular lumen leading to vascular stenosis [46]. Many growth factors are released in the process of vascular remodeling. Interestingly, HDGF is not expressed in the vascular wall until it is injured and thus represents an attractive target for therapeutic intervention [47]. A similar role of HDGF in lung remodeling was provided by stimulating lung epithelial growth after injury [48]. Furthermore, HDGF seems to play a role in the renewal of intestinal epithelium and that the presence of auto anti-HDGF antibodies in patients with chronic ulcerative colitis might be a reason for their delayed

mucosal healing [49]. HDGF expression is induced in hepatocytes before induction of DNA synthesis in two liver regeneration models suggesting that it might also be involved in liver regeneration possibly as an autocrine factor. In addition, HDGF participates in human liver regeneration after drug-induced liver damage and surgical resection thus making it a candidate for the protection and enhancement of the recovery from liver damage as a result of surgical resection, viral infection or drug intoxicity [50].

1.3.4 HDGF involvement in apoptosis

Induction of apoptosis via the TNF receptor requires the release of the IAP antagonist Smac/Diablo from the mitochondria, which subsequently disrupts the TRAF2-cIAP1 complex and permits activation of caspase-8 [51]. RNAi-induced silencing of the HDGF gene prevented the release of Smac/Diablo upon TNF-α treatment, suggesting that lack of HDGF interferes with the release of pro-apoptotic factors from the mitochondria [52]. Clermont et al. (2008) showed that triggering of endothelial cell apoptosis by TNFα and cycloheximide leads to an early dephosphorylation of HDGF which occurs before mitochondrial membrane permeabilization and downstream from an initiator caspase [53]. Knock-down of HDGF not only induced apoptosis in human cancer cells through the Bad-mediated intrinsic apoptotic pathway [54], but also the Fas-mediated extrinsic apoptotic pathway, eventually suppressing anchorage-independent growth of cancer cells [55].

1.4 Concept of the thesis

The aim of the present thesis is to improve the knowledge about HDGF and its functional regulation. The thesis specifically centers on post-translational modifications and protein-protein interactions for HDGF. Post-translational modifications are known to modulate cellular pathways by causing dynamic and reversible alterations in the interaction profile of proteins. More detailed insights in this area would greatly increase our understanding about molecular mechanisms and their impact on HDGF function.
In particular the following questions were addressed and elaborated:

- **Is HDGF modified by SUMOylation and does it affect HDGF function?**

The fact that many nuclear/DNA interacting proteins are post-translationally modified by small ubiquitin-like modifier (SUMO) thereby regulating their function makes it plausible that HDGF might also be a candidate for SUMO modification. In this part of the thesis we investigate if HDGF is SUMO modified and if so, where is the site of modification. We were able to show that HDGF is SUMOylated at a non-consensus motif and the modification negatively influences the ability of HDGF to bind to chromatin.

- **Does post-translational modification play a role in regulating HDGF secretion?**

The cellular trafficking of HDGF is considered atypical due to the fact that HDGF lacks a classical secretion signal but is still able to exit the cell by a yet unknown route. In this part of the thesis we investigated regulation of HDGF secretion. We show for the first time that the HDGF N-terminal region mediates secretion and that loss of a potential serine phosphorylation site in HDGF regulates N-terminally processing and secretion of the protein.

- **Does HDGF interact with other members of the HRP family of proteins?**

Protein-protein interactions play a pivotal role in regulation of function. Recently, HDGF dimer formation via a novel domain swapping mechanism through interaction of the N-terminal *hath* region of the protein was demonstrated. The fact that all HRP members share structural homology in the N-terminal region led to the basic hypothesis of interaction of HDGF with other members of the HRP family. We discovered a new HRP-2 splice variant and showed specific interaction of the different HRP-2 splice variants with HDGF.

2

Materials & Methods

2.1 Cell lines

Cell line	Provider
COS-7 cells (African green monkey kidney cells)	DSMZ, Braunschweig, Germany
HEK (Human embryonic kidney cells)	Dr. Rainer Niedenthal Hannover, Germany.
NIH 3T3 (Mouse embryonic fibroblast cells)	Dr. Sebastian Franken Bonn, Germany

2.2 Plasmid vectors and Primers

Item	Provider
pCDNA3 eukaryotic expression vector	Amersham
pCMV3b eukaryotic expression vector	Clontech
pEGFP-C3 eukaryotic expression vector	Clontech
pEGFP-N3 eukaryotic expression vector	Clontech
pEXPR-IBA5 eukaryotic expression vector	IBA BioTAGnology
Primers	Eurofins MWG Operon

2.3 Kits

Kit name	Manufacturer
GeneJET ™ Plasmid Miniprep kit	Fermentas
NucleoBond® Xtra Midi kit	Macherey-Nagel
QIAquick PCR Purification kit	Qiagen
QIAquick Gel Extraction kit	Qiagen
BCA Protein assay kit	Pierce/Thermo
ECL advanced Western blotting detection kit	Amersham
Pierce ECL Western blotting substrate	Pierce/Thermo
RevertAid H Minus First strand cDNA synthesis kit	Fermentas

2.4 Antibodies

Antibody	Stock concentration	Provider/Catalogue#
Affinity purified polyclonal sheep anti-mouse HDGF	0.40 mg/mL	Dr. Frank Dietz, Bremen, Germany
Affinity purified polyclonal rabbit anti-HRP-2	0.10 mg/mL	Dr. Frank Dietz, Bremen, Germany
Affinity purified polyclonal sheep anti-*Strep* Tag	0.40 mg/mL	Dr. Frank Dietz, Bremen, Germany
Polyclonal rabbit anti-SUMO-1	0.20 mg/mL	Santa Cruz, USA/ Sc 9060
Polyclonal mouse anti-EGFP	1.00 mg/mL	BD Biosciences 632381
Monoclonal mouse anti-Myc	2.40 mg/mL	ATCC (Myc 1-9E10.2) CRL1729
Peroxidase-conjugated affinity purified donkey anti-sheep IgG (H+L)	0.45 mg/mL	Dianova 713-035-147
Peroxidase-conjugated affinity purified donkey anti-mouse IgG (H+L)	0.40 mg/mL	Dianova 715-035-151
Peroxidase-conjugated affinity purified donkey anti-rabbit IgG (H+L)	0.40 mg/mL	Dianova 711-035-152
Cy3-conjugated affinity purified donkey anti-sheep IgG (H+L)	0.75 mg/mL	Dianova 713-165-003
Cy2-conjugated affinity purified rabbit anti-mouse IgG (H+L)	1.50 mg/mL	Dianova 315-225-003
Cy3-conjugated affinity purified donkey anti-rabbit IgG (H+L)	0.50 mg/mL	Dianova 711-165-152
Cy2-conjugated affinity purified donkey anti-sheep IgG (H+L)	0.75 mg/mL	Dianova 713-225-003

2.5 Growth media

Media*	Composition
DMEM (Dulbecco's Modified Eagle's Medium)	13.37 g DMEM 03.70 g $NaHCO_3$ add to 1 L with ddH_2O
LB medium (Liquid Broth)	15.00 g Tryptone 05.00 g Yeast extract 05.00 g NaCl add to 1 L with ddH_2O
LB ampicillin agar	15.00 g agar 1 L LB medium 50 µg/mL ampicillin
LB kanamycin agar	15.00 g agar 1 L LB medium 25 µg/mL kanamycin
SOC medium	20.00 g tryptone 05.00 g yeast extract 00.50 g NaCl 02.50 mL KCl (1M) add to 1 L with ddH_2O

2.6 Solutions and buffers

Solutions/buffer	Composition
Agarose gel electrophoresis loading buffer (6X)	00.09% Bromophenol blue 60.00% Glycerol 60.00 mM EDTA
Ampicillin stock solution	500 mg/mL in ddH_2O
Ethidium bromide stock solution	10 mg/mL in ddH_2O
Glycine Solution	0.10 M Glycine HCl, pH 3.0
Kanamycin stock solution	50 mg/mL in ddH_2O
PBS buffer (Phosphate Buffer Saline)	37.00 mM NaCl 02.70 mM Na_2HPO_4 01.50 mM K_2HPO_4
SDS-PAGE electrophoresis (running) buffer (10X)	30.30 g Tris base 144.0 g Glycine 10.00 g SDS to 1 L with ddH_2O

* All growth media listed here were sterilized by autoclaving.

SDS-PAGE stacking gel buffer	0.5 M Tris-HCl, pH 6.8 0.4% SDS in ddH_2O
SDS-PAGE separation gel buffer	1.5 M Tris-HCl, pH 8.8 0.4 % SDS in ddH_2O
SDS-PAGE sample buffer (2x)	0.125 M Tris-HCl, pH 6.8 4% SDS 20% glycerol (v/v) 0.02% bromophenol blue
TAE buffer (10x) (Agarose gel electrophoresis running buffer)	0.4 M Tris acetate 10 mM EDTA, pH 8.3
TBS (Tris Buffer Saline)	10 mM Tris HCl, pH 7.4 150 mM NaCl
TNE lysis buffer (NP-40 lysis buffer)	20 mM Tris 150 mM NaCl 5 mM EDTA 1% NP-40, pH 7.4
Trypsin/EDTA	0.5 g Trypsin 0.2 g EDTA to 1 L using sterile 1xPBS
Western blotting buffer	25 mM Tris-HCl 192 mM glycine 20% (v/v) Methanol
Protease inhibitor cocktail-tablet	(1 tablet/10 mL lysis buffer)

Chromatin binding assay buffers:

CSK I	10 mM Pipes, pH 6.8 100 mM NaCl 1 mM EDTA 300 mM Sucrose 1 mM $MgCl_2$ 1 mM DTT 0.5% Triton X-100 Protease inhibitor -tablet 1 mM PMSF
RIPA	150 mM Tris-HCl, pH 8.0 150 mM NaCl 0.5% Sodium deoxycholate 0.1% SDS 1% (v/v) NP-40
CSK II	10 mM Pipes, pH 6.8 50 mM NaCl 300 mM Sucrose 6 mM $MgCl_2$ 1 mM DTT

2.7 Molecular cloning and plasmid isolation†

2.7.1 Bacterial Transformation

Around 12 μL of the ligation mix were added to 200 μL half thawed chemically competent *E. coli* XL1 blue cells obtained from -80°C storage. After keeping the mix for 10 min on ice, the tube was transferred to a 37°C heat block for 90 s (heat shock) and then back on ice for 2 min. Pre-warmed SOC medium (1 mL) was then added to the transformation mixture and incubated in the 37°C shaker for 45 min to allow the expression of the resistance gene under non-selective conditions. The culture was then aseptically spread on a pre-warmed agar plate containing the specified antibiotic (depending on the resistance gene included in the host plasmid). The plate was incubated overnight at 37°C to allow the growth of transformed bacteria.

2.7.2 Colony PCR

Colony PCR was performed to distinguish colonies arising from bacteria in which only the re-ligated plasmid had been transformed from those into which the plasmid containing the insert of interest had been transferred. Each 25 μL PCR reaction mix contained: a portion of the colony under test , 0.25 μL sense primer, 0.25 μL antisense primer (primers concentration 100 pmol/μL), 1.5 μL $MgCl_2$ (25 mM), 2.5 μL of 10x *Taq* buffer, 0.5 μL dNTPs (10 mM each), 0.10 μL *Taq* polymerase (5 U/μL). The PCR reaction took place in a thermocycler (Eppendorf) using following conditions: 3 min at 95°C, 29 cycles of 30 s at 95°C (denaturation), 1 min at 56°C (annealing) and 1 min at 72°C (extension). The 29 cycles were followed by a final extension step of 3 min at 72°C. 10 μL of the PCR product were then mixed with 2 μL of 6x agarose gel loading buffer and loaded to a 1% (w/v) agarose gel containing ethidium bromide (0.1μL/mL). The samples were allowed to run at constant voltage (100 V) for adequate time. The bands were visualized under UV light and the size of the insert was determined in comparison to the DNA mass ruler ladder mix (Fermentas).

2.7.3 Plasmid Purification

A positive clone for the insert was aseptically transferred to 5 mL or 200 mL sterile LB medium containing the antibiotic which depends on the resistance gene in the plasmid vector. The clones were allowed to grow overnight in a 37°C incubator with shaking at 250 revolutions min^{-1}. Plasmids were purified from bacterial cultures using the NucleoBond Xtra Midi kit (Macherey-Nagel) for the 200 mL cultures or the GeneJET Plasmid Miniprep kit (Fermentas) for the 5 mL cultures.

† This section only contains methods not described or elaborated in publications

2.7.3a Plasmid purification using the NucleoBond Xtra Midi kit (Macherey-Nagel)

A 250 mL overnight bacterial culture was centrifuged at 6000 *g* for 15 min at 4ºC. The supernatant was decanted and residual medium was removed by placing the tubes in an inverted position. The pellet was re-suspended in 8 mL of the provided Buffer RES (ice cold with RNase A) and the homogenized lysate was transferred to a 50 mL tube. Alkaline lysis was done by adding 8 mL of the Buffer LYS followed by incubating the mix for 5 min at RT. The lysate was neutralized by adding 6 mL of the Buffer NEU and the fluffy precipitate was allowed to flow through the column, pre-equilibrated with 12 mL Buffer EQU, by gravity. The column filter was first washed twice with 5 mL Buffer EQU, filter then discarded and the column now washed with 8 mL of buffer WASH. The bound DNA was eluted using 5 mL Buffer ELU. For DNA precipitation, 3.5 mL isopropanol was added to the eluate and the mix was incubated for 5 min. The eluate/isopropanol mix was centrifuged at 15000 *g* for 30 min at 4ºC, the pellet obtained was washed with 2 mL 70% ethanol and re-centrifuged at 15000 *g* for 15 min at RT and left to dry. Finally the DNA pellet was reconstituted in 1 mL buffer TE.

2.7.3b Plasmid purification using the GeneJET Plasmid Miniprep kit (Fermentas)

A 5 mL overnight bacterial culture was centrifuged at 12000 *g* for 10 min at 4ºC, the supernatant was decanted and residual medium was removed. The pellet was re-suspended in 250 µL of the provided resuspension buffer (ice-cold) and suspension was made homogenous by continuous vortexing for 30 s. Following this, 250 µL of lysis solution was added to the tube and mixed thoroughly by inverting the tube 4-6 times. Neutralization solution (350 µL) was then added to the tube and mixed thoroughly by inverting the tube 4-6 times. This was followed by a centrifugation step for 10 min on a table top centrifuge. The resulting supernatant is applied to GeneJET spin column and centrifuged again at 13000 rpm for 30-60 s. The GeneJET spin column was then washed using 500 µL of wash solution, centrifuged for 30-60 s. The flow through is discarded and the assembly spun once again to remove the residual wash buffer. The column was then transferred to a new tube and the DNA eluted with 50 µL of elution buffer by centrifugation for 1 min at full speed to obtain plasmid DNA.

2.7.4 Measurement of plasmid concentration and level of purity

The plasmid concentrations were determined by measuring the absorbance of the eluates at 260 nm using the Eppendorf photometer. The purity (level of protein contamination) was assessed from the A 260/280 value.

2.7.5 Agarose gel electrophoresis

Agarose gels (1%) containing ethidium bromide was pre-casted. For sample preparation, 20 µL of each sample under test was mixed with 5 µL of the agarose gel loading buffer and loaded to the wells. 10 µL of

the molecular weight marker (Mass Ruler DNA Ladder Mix, Fermentas) was loaded in parallel. The electrophoresis was performed at 100 volts for approximately 30 min after which the DNA was visualized under the UV transilluminator (Herolab).

2.8 Plasmid construction

2.8.1 Production of untagged mHDGF D205G and hHDGF G205D mutants

Plasmids were constructed using standard recombinant cloning techniques and all changes were verified by DNA sequencing. HDGF mutants were generated by site-directed mutagenesis with the Quickchange Multi-Site Direct Mutagenesis kit (Stratagene, La Jolla, CA, U.S.A.) according to the manufacturer's protocol. Untagged wild type mouse HDGF and human HDGF cloned in pcDNA3 Amp was used as a template and amplified using the primers listed below in Table 2.1.

Table 2.1 List of primers used for site directed mutagenesis. In case a restriction site is created, such site is shown in parentheses in the description lane and in bold in sequence lane. The base changes leading to the appearance or disappearance of restriction sites are underlined. The shaded triplets represent changed codons in mutants. The mutated base leading to mutations are underlined in shaded codons.

Name	Description	Sequence (5´→3´)
OSK 1290	sense primer mHDGF D205G BamHI (G^GATCC) appears	CCCTCTGAGCCA**GGATCC**GGCCAGGGACCT
OSK 1289	antisense primer mHDGF D205G BamHI appears	AGGTCCCTGGCC**GGATCC**TGGCTCAGAGGG
OSK 1292	sense primer hHDGF G205D (silent mutation)	CCCTCTGAGCCCGACTCTGGCCGGGGGCCT
OSK 1291	antisense primer hHDGF G205D (silent mutation)	AGGCCCCCGGCCAGAGTCGGGCTCAGAGGG

Plasmids encoding for HA-SUMO-2 and HA-SUMO-3 were provided by Dr. Rainer Niedenthal (Hannover, Germany). Other cloning strategies, plasmids used and mutants produced during the course of the entire study have been elaborated under the materials and methods section of individual publications in the next chapter of this thesis.

2.8.2 Site-directed mutagenesis

The insertion of point mutations was performed based on the QuickChange site-directed mutagenesis method (Stratagene). The method is based on nonstrand-displacing action of the *Pfu* DNA polymerase, as well as on the selective action of the *DpnI* endonuclease on methylated and hemimethylated DNA. Each reaction tube contained 100 ng of the plasmid with the wt HDGF insert as a template, 1.25 µL of the sense and antisense primer (100 pmol/µL each) containing the nucleotide exchanges for the mutation which has to be introduced, 1 µL *Pfu* polymerase (5U/µL), 5µL of 10x *Pfu* polymerase buffer and 1 µL dNTPs (10 mM each). The following profile was used for amplification: 5 min at 95°C followed by 18 cycles of 1 min

denaturation at 95°C, 1 min primer annealing at 63°C and an elongation step for 20 min at 72°C. The 18 cycles were followed by a final elongation step at 72°C for 25 min. For digestion of wild type template, 1µL *DpnI* (10 U/µL) was immediately added to the amplification product, and the mixture was incubated at 37°C for 3 h. The *DpnI* digest was directly used to transform chemically competent bacteria (*E. coli* XL1 blue). The mix was processed as described in section 2.7.1 and the bacteria were plated on LB Amp Agar plates. Following 24 h incubation at 37ºC, colonies obtained were subjected to colony PCR using specific primers.

2.8.3 Preliminary assessment of the site-directed mutagenesis

Preliminary assessment of the success of site-directed mutagenesis was done by checking for presence of silent mutations which led to a loss or gain of a restriction site introduced through the mutagenesis primers. To check for the presence or absence of a restriction site by silent mutation, 1 µL (0.5-3U) of the required restriction enzyme was directly added to 20 µL of the colony PCR product and positive clones were selected comparing their restriction digest pattern to a similar digest done on the wild type fragment. Plasmids were prepared from clones showing positive results for the introduced mutation and the mutations were subsequently confirmed by sequencing.

2.8.4 DNA Sequencing PCR reaction

The sequencing PCR reaction was performed by using the Big Dye method. The Big Dye consists of fluorescent labeled dNTP and dideoxy NTP (ddNTP's) which are incorporated onto the amplified DNA fragment. Fragments of varying length are amplified which are analyzed using a DNA-sequencer. Each 5 µL reaction mix contained: 1 µL template DNA, 1 µL PCR water, 0.8 µL Big Dye, 1.2 µL reaction buffer (2.5x) and 1µL of sequencing primer (5 pmol/µL). The PCR reaction conditions were: 96°C for 20 min (denaturation) followed by 60 cycles of 10°C for 10 min, 55°C for 5 min, 60°C for 4 min followed by a final holding step at 20°C. 20 µL of PCR amplified plasmid DNA was added to a rehydrated Sephadex powder G-50 Superfine plate and centrifuged at 900 *g* for 5 min at RT. The purified and PCR amplified DNA was sequenced at the Max Planck Institute for Marine Microbiology, Bremen, Germany.

2.9 Recombinant protein expression in eukaryotic cell lines

2.9.1 Culturing eukaryotic cell lines

Eukaryotic cell lines were used throughout this study to produce various recombinant proteins that were investigated. NIH 3T3 cells (mouse embryonic fibroblast cells), COS-7 cells are fibroblasts from kidney of the African green monkey (*Cercopithecus aethiops),* and HEK 293 cells (human embryonic kidney cells) were used during this study. All cells grow as an adherent monolayer. NIH 3T3, COS-7 and HEK 293 were maintained on 10 cm tissue culture plates (Sarstedt) covered with DMEM medium containing

10% fetal calf serum (FCS) and 10 µg/mL antibiotic gentamycin. Cells grown for *in vitro* transfections were transfected when adherent cells had reached a subconfluent state (~ 70 % coverage of the plate surface).

2.9.2 *In vitro* transfection

Transient transfections were done by adding DNA plasmids, containing an insert coding for the protein of interest, together with the ExGen 500 gene delivery reagent (Fermentas) to adherent cultured cells. Due to charge interaction the ExGen 500 and DNA form complexes that settle by gravity onto the cell and are absorbed by endocytosis. Transfections were carried out according to manufacturer's protocol. The average post-transfection incubation time was 24 h, after which the cells were either harvested or processed for immunofluorescence.

2.9.3 Cell Harvesting

Twenty four hours after transfection, culture medium was removed using a pasteur pipette connected to a bench aspirator. The cells were briefly washed twice using 1xPBS to remove residual culture medium containing fetal calf serum (FCS). The specified volume of TNE lysis buffer with protease inhibitor tablet (1 tablet/10 mL of lysis buffer) with or without 20 mM each of NEM and IAA, according to the setup of experiment, was then applied and plates were placed for 10 min on ice to allow cell lysis. The cells were then scraped off the plate using a rubber scrapper, disaggregated using a pipette tip and then transferred to 1.5 mL cap. Following 10 min incubation on ice, the cell lysates were centrifuged (15 min, 4°C, 14000 *g*) to pellet unlysed cells and debris. The resulting supernatants were transferred into fresh 1.5 mL tubes and used for subsequent investigations.

2.10 Protein Analysis

2.10.1 Sodium dodecyl sulphate-Polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is an electrophorectic technique used for separation of proteins on the basis of their molecular weights irrespective of their native charge or their 3D structure. Introduction of the anionic surfactant SDS and treatment of the protein sample with reducing agents like dithiothreitol (DTT) causes the linearization of proteins at 95ºC and imparts a uniform negative charge on the protein. The separation of the proteins takes place due the difference in their masses. 10 µL of the protein sample to be analyzed was mixed with 10 µL of the 2x SDS-PAGE sample buffer (with or without DTT). The mix was heated for 5 min at 95°C and loaded to the discontinuous SDS polyacrylamide gel casted in miniprotean 3 system (Bio-Rad). A pre-stained molecular weight marker (Fermentas) was loaded in a separate lane. The electrophoresis was performed at constant current of 20 mA with the voltage set to maximum of 200 V

2.10.2 Western blotting

Western blotting is a 3-step process in which SDS-PAGE as a first step is used to resolve a mixture of proteins by size. The second step is the transfer of the resolved proteins from the gel to a membrane support (E.g. polyvinylidene difluoride (PVDF)) via electroelution. The third and final step is processing of the blot for detection of specific proteins with an antibody.

2.10.2a Transferring resolved proteins on PVDF membrane

Following SDS PAGE, the gels were equilibrated for 10 min in chilled (4°C) Western blotting buffer. A cut-to-size PVDF membrane was briefly activated by immersing in methanol for few seconds; washed briefly in ddH_2O and then equilibrated in blotting buffer. The gel and activated membrane were sandwiched between a fiber pad and two filter papers from each side. The assembled gel/blot sandwich was placed in the blotting cassette, installed into the cassette holder and placed into the blotting tank (Bio-Rad, Miniprotean III tank blot). The tank was filled with blotting buffer and electro-elution was allowed to occur at constant voltage (100V) for 60 min.

2.10.2b Immunodetection

Disassembling the blot sandwich, the PVDF membrane was briefly washed in 1xTBS/Tween 0.05% in order to remove residual methanol and gel pieces. The membrane was then incubated on the shaker with 5% skimmed milk or 5% BSA (in 1xTBS/Tween 0.05%) for 1 h at RT to block unspecific binding. The membrane, covered with the primary antibody solution, was sealed with parafilm and incubated at RT for 3 h or 4 °C overnight (1 mL solution at the dilutions specified in Table 2.2). The primary antibody solution was washed 3 times, 10 min each, using 1xTBS/Tween 0.05%. The membrane was incubated with 20 mL of the POD-conjugated secondary antibody solution (at the dilutions specified in Table 2.2). The secondary antibody solution was washed with the same procedure as described for the primary antibody. For detection, 750 µL ECL solution A/solution B (1:1 mix) were applied to each membrane and allowed to stand for 5 min. The membrane was sealed in plastic and exposed to an X-ray film (Amersham). The protein bands were visualized after developing the X-ray film (Agfa developer).

Table 2.2 Primary and secondary antibodies used for detecting proteins of interest after western blotting.

Antigen/Tag	Primary antibody /dilution	Secondary antibody/dilution
mHDGF	Polyclonal sheep anti-mouse HDGF /1:1000	POD Donkey anti-sheep /1:20000
SUMO-1	Polyclonal rabbit anti-SUMO-1 /1:1000	POD Donkey anti-rabbit /1:10000
*Strep*Tag	Polyclonal sheep anti-*Strep*Tag/1:1000	POD Donkey anti-sheep /1:20000
MycTag	Monoclonal mouse anti-Myc/1:500	POD Donkey anti-mouse /1:20000
EGFP	Monoclonal mouse anti-EGFP/1:500	POD Donkey anti-mouse /1:20000
HRP-2	Polyclonal rabbit anti-HRP-2/ 1:500	POD Donkey anti-rabbit /1:20000
Actin	Polyclonal mouse anti-Actin/ 1:2000	POD Donkey anti-mouse /1:20000
α-tubulin	Polyclonal mouse anti-human α-tubulin/ 1:1000	POD Donkey anti-mouse /1:20000

3
Publications

List of publications and my contribution towards them

- ### Publication 1

Ketan Thakar, Rainer Niedenthal, Elwy Okaz, Sebastian Franken, Astrid Jakobs, Shivangi Gupta, Sørge Kelm and Frank Dietz (2008)

SUMOylation of the hepatoma-derived growth factor negatively influences its binding to chromatin

FEBS Journal **275:** 1411–1426

I developed the concept of the study in co-operation with FD. I conducted the experiments supported by FD, EO, and SG. RN provided SUMO-1 expression plasmid used in work. SF and AJ provided the MALDI-TOF data for the study. Along with FD, I analyzed the data and was involved in writing and revision of manuscript together with FD. SK, RN and SF also provided assistance in revision the manuscript

- ### Publication 2

Ketan Thakar, Tim Kröcher, Soniya Savant, Doron Gollnast, Sørge Kelm, Frank Dietz (2010)

Secretion of hepatoma-derived growth factor is regulated by N-terminal processing

Biological Chemistry (submitted)

I developed the concept of the study and designed the experiments in co-operation with FD. I carried out the experiments supported by FD, TK, DG and SS. I analyzed the data and wrote the manuscript, which was revised together with FD and SK.

- ### Publication 3

Ketan Thakar , Dipti Kelkar, Teja Shidore, Shivangi Gupta, Sørge Kelm, Frank Dietz (2010)

Interaction of HRP-2 splice variants with Hepatoma-derived growth factor

Biological Chemistry (submitted)

I developed the concept of the study, designed and carried out the experiments supported by FD, DK, SG, and TS. I analyzed the data and wrote the manuscript, which was revised together with FD and SK.

3.1 Publication 1

SUMOylation of the hepatoma-derived growth factor negatively influences its binding to chromatin

K Thakar, R Niedenthal, E Okaz, S Franken, A Jakobs, S Gupta, S Kelm and F Dietz

2008

FEBS Journal
275: 1411–1426

the FEBS Journal

SUMOylation of the hepatoma-derived growth factor negatively influences its binding to chromatin

Ketan Thakar[1], Rainer Niedenthal[2], Elwy Okaz[1], Sebastian Franken[3], Astrid Jakobs[2], Shivangi Gupta[1], Sørge Kelm[1] and Frank Dietz[1]

1 Department of Biochemistry, Centre for Biomolecular Interactions Bremen (CBIB), University of Bremen, Germany
2 Department of Biochemistry, Hannover Medical School, Germany
3 Department of Physiological Chemistry, Rheinische-Friedrich-Wilhelm University of Bonn, Germany

Keywords
HDGF related protein (HRP); nuclear localization; PWWP domain; SUMOylation

Correspondence
F. Dietz, Department of Biochemistry, Centre for Biomolecular Interactions Bremen (CBIB), University of Bremen, Leobener Strasse im NW2, 28359 Bremen, Germany
Fax: +49 421 218 2981
Tel: +49 421 218 4324
E-mail: fdietz@uni-bremen.de

(Received 16 May 2007, revised 7 December 2007, accepted 16 January 2008)

doi:10.1111/j.1742-4658.2008.06303.x

Hepatoma-derived growth factor is a nuclear targeted mitogen containing a PWWP domain that mediates binding to DNA. To date, almost nothing is known about the molecular mechanisms of the functions of hepatoma-derived growth factor, its routes of secretion and internalization or post-translational modifications. In the present study, we show for the first time that hepatoma-derived growth factor is modified by the covalent attachment of small ubiquitin-related modifier 1 (SUMO-1), a post-translational modification with regulatory functions for an increasing number of proteins. Using a basal SUMOylation system in *Escherichia coli* followed by a MALDI-TOF-MS based peptide analysis, we identified the lysine residue SUMOylated located in the N-terminal part of the protein adjacent to the PWWP domain. Surprisingly, this lysine residue is not part of the consensus motif described for SUMOylation. With a series of hepatoma-derived growth factor mutants, we then confirmed that this unusual location is also used in mammalian cells and that SUMOylation of hepatoma-derived growth factor takes place in the nucleus. Finally, we demonstrate that SUMOylated hepatoma-derived growth factor is not binding to chromatin, in contrast to its unSUMOylated form. These observations potentially provide new perspectives for a better understanding of the functions of hepatoma-derived growth factor.

Hepatoma-derived growth factor (HDGF) is the ubiquitously expressed prototype of a family of proteins called the HDGF-related proteins (HRPs) [1]. To date, four HRPs (HRP-1 to HRP-4) and a protein called the lens epithelium-derived growth factor (LEDGF) have been described [2,3]. HDGF was identified as a 25 kDa heparin binding protein, purified from conditioned media of the human hepatocarcinoma cell line Huh7 [1]. Upon overexpression, HDGF shows mitogenic activity and this growth promoting activity depends on its nuclear localization mediated by the presence of two functional nuclear localization signals (NLS) within its primary amino acid sequence [4,5]. HRPs and LEDGF share a highly conserved N-terminal region of approximately 100 amino acids called the *hath* region (homologous to the amino terminus of HDGF). This region includes a PWWP domain found in an increasing number of proteins [6,7].

Structural data available for the *hath* region of HDGF and HRP-3 revealed a characteristic fold made up of a five-stranded β barrel followed by α-helical elements [8–10]. The PWWP domain shares similarities

Abbreviations
DAPI, 4′-6-diamino-2-phenylindole HCL; Dnmt, DNA methyltransferase; EGFP, enhanced green fluorescence protein; HA, hemagglutinin epitope; *hath*, homologous to the amino terminus of HDGF; HDGF, hepatoma-derived growth factor; HRP, HDGF-related protein; IAA, iodacetamide; LEDGF, lens epithelium-derived growth factor; NLS, nuclear localization signal; SUMO-1, small ubiquitin-related modifier 1; wt, wild-type.

with the well known Tudor and Chromo domain and, like these domains, it has been proposed to play a role in DNA-binding and/or protein–protein interactions. In the case of HDGF, the PWWP domain may have a dual function in binding double-stranded DNA as well as the glycosaminoglycan heparin [2,8–10]. DNA-binding via the PWWP domain of HDGF appears to be specific for a region covering approximately 40 bp found in potential target genes of HDGF [11], although it has not been clarified whether further specificity may be mediated by the C-terminal portion of the protein [8]. A recent study by Sue *et al.* [12] demonstrated that dimerization of the PWWP domain by an unusual domain-swapping leads to an increased binding affinity for heparin. However, the physiological role of this phenomenon is unclear.

Hepatoma-derived growth factor is secreted from cells. The mechanism for externalization remains unclear because HDGF, like the other HRPs, has no obvious signal peptide. Extracellular HDGF appears to be internalized by binding to heparan sulfate or other mechanisms [5]. Recent studies have provided evidence for a potential receptor specifically binding extracellular HDGF, leading to the activation of intracellular signalling cascades [13].

The expression of HDGF changes during development, as shown for kidney, liver, heart and vascular tissue [14–20]. In addition, recent studies have demonstrated that HDGF is differentially expressed in the brain [21] and also can function as a potent neurotrophic factor [22,23]. Furthermore, different studies have shown that HDGF can serve as a prognostic marker in a variety of human cancers [24–31] and that it probably promotes angiogenesis and tumor progression [32].

Phosphorylation prediction programs have identified HDGF as a good candidate for phosphorylation on several serine and threonine residues, but only one mass spectroscopy based approach has confirmed the use of serine residues S132, S133 and S165 [33], although no evidence for functional relevance was provided.

A post-translational modification found in several nuclear proteins comprises the attachment of the small ubiquitin-related modifier 1 (SUMO-1). The modification by this 11 kDa protein is mechanistically related to that of ubiquitin, with which it shares a high degree of structural similarity. Like ubiquitination, SUMOylation is a dynamic process that is mediated by activating (E1), conjugating (E2) and ligating (E3) enzymes and can be reversed by the action of SUMO specific proteases [34,35]. Despite these similarities, the functions of both modifications differ. SUMOylation of target proteins usually occurs on lysine residues in the context of a highly conserved recognition motif ΨKxE/D (where Ψ stands for a large hydrophobic amino acid, K is the lysine modified, x is any amino acid and E/D are the negatively charged amino acids glutamate or aspartate). Well documented functions of SUMOylation are the regulation of subcellular distribution, DNA repair, transcriptional regulation, stabilization, RNA metabolism and cell signalling [34–38]. SUMO itself can further serve as a docking site for the binding of other proteins containing SUMO binding motifs [39–42].

Based on the knowledge that HDGF is a nuclear targeted mitogen with DNA binding capacity, we investigated whether HDGF is also modified by the addition of SUMO-1. In the present study, we show for the first time that HDGF serves as a template for SUMO-1 conjugation, although it does not contain a suitable consensus site for SUMOylation. Using a basal SUMOylation system in *Escherichia coli* [43] followed by a MALDI-based peptide analysis of the SUMOylated HDGF wild-type (wt) and a series of HDGF mutants, we identified an unusual SUMOylation site located in the N-terminal *hath* region. Furthermore, we discovered that SUMOylated HDGF does not bind to chromatin, in contrast to its unSUMOylated form.

Results

SUMOylation of HDGF

For most SUMOylated proteins, overexpression of the target protein together with SUMO-1 is necessary to detect SUMOylation. When untagged HDGFwt (apparent molecular mass = 40 kDa) is overexpressed together with enhanced green fluorescence protein (EGFP)-SUMO-1 in COS-7 cells (Fig. 1A) or human embryonic kidney cells (HEK293) (data not shown) we use the advantage of the much higher molecular mass shift of the EGFP-SUMO-1 SUMOylated proteins. We detected an extra protein band reacting with a specific anti-HDGF serum in the molecular weight range expected for EGFP-SUMO-1 conjugated HDGF (apparent molecular weight = 100 kDa). This observation suggests that HDGF can be modified by SUMOylation. In several systems, it has been shown that SUMOylation is a highly dynamic, reversible modification, which is sensitve to the action of specific isopeptidases. Since these are cysteine proteases, deSUMOylation can be partially blocked by lysing the cells in the presence of *N*-ethylmaleimide and iodacetamide (IAA) to alkylate the free SH-group in the active centres of the isopeptidases. In agreement with this

phenomenon, we found that most of the suspected HDGF-EGFP-SUMO-1 band migrating at 100 kDa is lost if the cell lysates are incubated for 15 min on ice in the absence of *N*-ethylmaleimide and IAA (Fig. 1B), whereas almost no loss occurs if these SH-alkylating reagents are included.

To further investigate whether the additional protein bands are indeed due to the covalent modification of HDGF with SUMO-1, we used C-terminally *Strep*-tagged HDGF, which can be precipitated using StrepTactin® beads. Similar to the experiment with untagged HDGFwt, the C-terminal *Strep*-tag-labelled HDGF appears to be SUMOylated (Fig. 1C) because, in StrepTactin® precipitates, a similar upshifted HDGF-band occurs, which reacts with antibody against SUMO-1 (Fig. 1B). As negative controls, we cotransfected either EGFP alone or an EGFP-SUMO-1 chimera lacking the C-terminal di-glycine motif (EGFP-SUMO-1ΔGG) required for the isopeptide bond formation of SUMO-1 with the acceptor lysine.

Only EGFP-SUMO-1 including the terminal di-glycine was covalently attached to HDGF, whereas EGFP alone or EGFP-SUMO-1ΔGG was not (Fig. 1C).

To overcome problems in the detection of HDGF mutants due to the restricted specificity of the polyclonal anti-HDGF antiserum, in all further experiments, *Strep*-tag-labelled HDGF constructs were employed, providing the possibility of using identical StrepTactin® precipitation and detection assays for all HDGF mutants. For all these experiments, similar results were obtained if HEK293 cells were utilized instead of COS-7 cells, supporting the notion that SUMOylation of HDGF is not a COS-7 cell restricted modification.

Overexpression of the conjugating enzyme Ubc-9 (E2) is commonly used to obtain detectable amounts of SUMOylated products. However, this was not required for the SUMOylation of HDGF because omission of the hemagglutinin epitope (HA)-Ubc-9 encoding plasmid did not reduce the level of SUMOylation (Fig. 1C). These data suggested that endogenous levels of Ubc-9 in COS-7 cells are sufficient to generate detectable levels of SUMOylated HDGF.

SUMOylation site of HDGF in *E. coli*

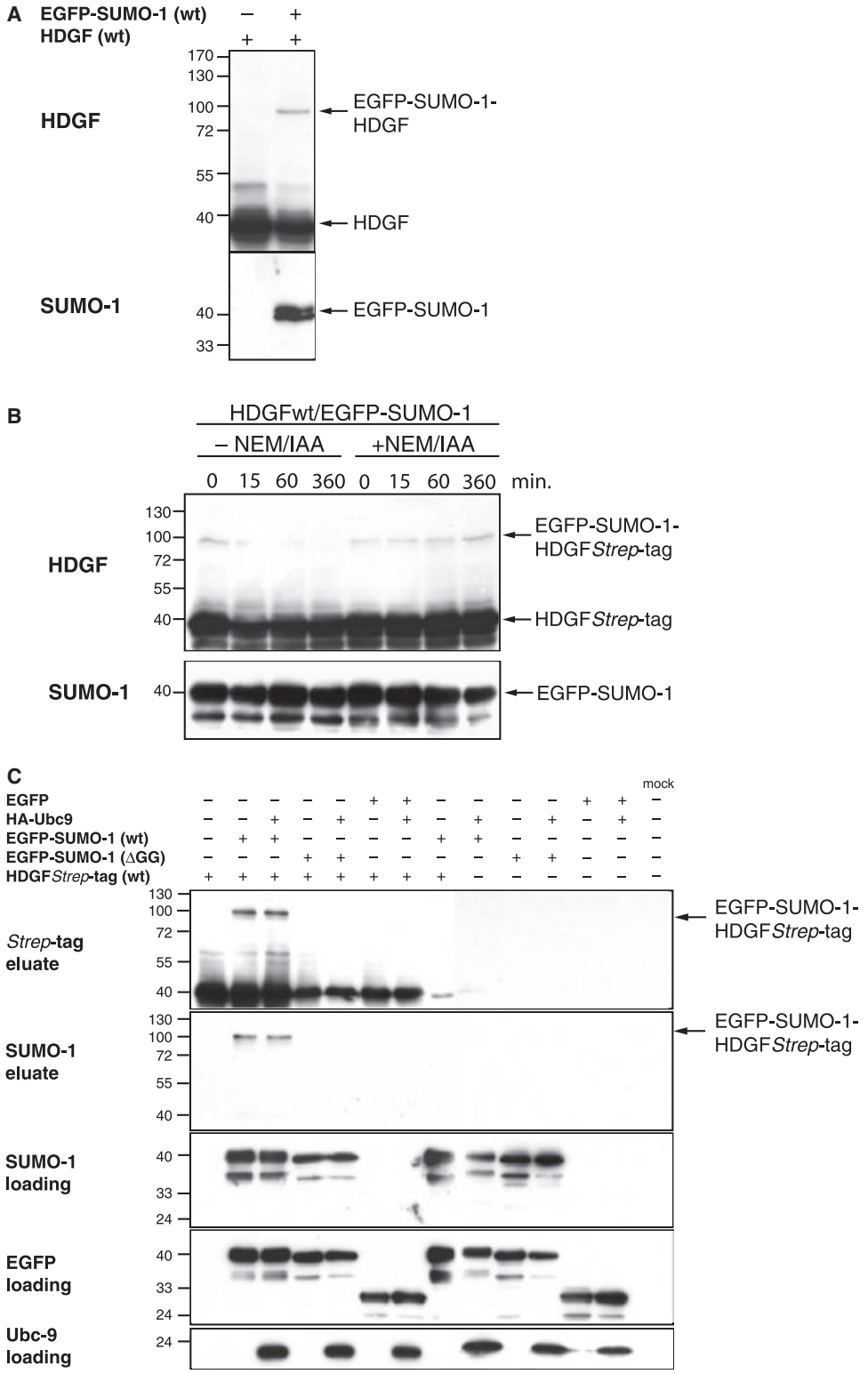
Screening the primary amino acid sequence of HDGF from different mammalian species (Fig. 2) using the prediction programs for SUMOylation motifs SUMOplot™ (http://www.abgent.com/doc/sumoplot) and the SUMOsp–SUMOylation sites PREDICTION program (http://bioinformatics.lcd-ustc.org/sumosp/), we found only three motifs that weakly match the postulated consensus motif ΨKxE/D (Fig. 2). The highest score by SUMO site prediction was obtained for K223 in the motif $AK_{223}EE$ of mouse HDGF (> 0.8), which is TKED in human and chimpanzee or AKED in bovine HDGF (Fig. 2). Other potential motifs predicted with low scores (0.5) are $EK_{148}NE$ and $PK_{167}RP$. Single K to R mutations in these three motifs and the expression of the mutants together with EGFP-SUMO-1 in COS-7 cells did not lead to any detectable loss in SUMOylation compared to HDGFwt (Fig. 3). This strongly suggested that other lysine residues than these are SUMOylated in HDGF.

Since no obvious consensus motif was found in HDGF, we aimed to identify the SUMOylation site(s) of HDGF by MS after tryptic digestion. In order to obtain sufficient amounts of SUMOylated HDGF for MALDI-TOF-MS analysis, we expressed the protein in *E. coli*. Essential compounds of the SUMOylation machinery, such as the activating (E1) and conjugating (E2) enzymes but not the ligating enzyme (E3) [43], were coexpressed with *Strep*-tagged mHDGFwt in *E. coli* strain BL21 DE3. This artificial bacterial SUMOylation system enabled us to purify high levels of SUMOylated and unmodified HDGF from bacterial lysates via StrepTactin® precipitation. After 2D electrophoresis of the purified protein, spots representing unmodified HDGF*Strep*-tag appearing at a molecular weight of 40 kDa (Fig. 4A, spot 1) and a protein spot appearing at a molecular weight of 63 kDa (Fig. 4A, spot 2) were cut out, digested with trypsin and used for MALDI-TOF-MS analysis. The protein spot at 63 kDa is only observed if HDGF is coexpressed with the SUMOylation machinery (data not shown). In the spectra of both protein spots, we found peptide masses perfectly matching the expected peptides from HDGF. However, only in spot 2 did we recognize additional peptide masses corresponding to peptides derived from huSUMO-1 (Fig. 4C).

Furthermore, comparison of the obtained peptide spectra of both spots clearly showed an almost complete loss of two mass peaks in the chromatogram derived from the digest of spot 2, most probably corresponding to monoSUMOylated HDGF (Fig. 4D,E). These mass peaks perfectly match the HDGF peptides K_{80}-K_{96} (K_{80}GFSEGLWEINNPTVK_{96}) and G_{81}-K_{96} (G_{81}FSEGLWEINNPTVK_{96}) of the *hath* region (Fig. 2). This observation strongly suggested that either K_{80} or K_{96} is modified by SUMOylation.

SUMOylation of HDGF at Lys_{80} in mammalian cell lines

To investigate whether K_{80} or K_{96} is SUMOylated in mammalian cells as suggested by the bacterial system,



we generated and expressed K80R and K96R mutants of HDGF in COS-7 cells together with EGFP-SUMO-1. Interestingly, expression of the mutants in comparison to HDGFwt clearly showed an almost complete loss of the signal for SUMOylated HDGF only in the case of the mutant K80R (Fig. 5A).

Fig. 1. HDGF is SUMOylated in mammalian cells. COS-7 cells transfected with plasmids coding for the indicated proteins were lysed and analysed by SDS/PAGE and western blotting with HDGF specific, *Strep*-tag specific, SUMO-1 specific, EGFP specific and Ubc9 specific antibodies as indicated. (A) COS-7 cells were transfected with plasmids coding for the expression of HDGF untagged (wt), alone or together with EGFP-SUMO-1. (B) COS-7 cells were transfected with plasmids coding for the expression of HDGF untagged (wt), alone or together with EGFP-SUMO-1 and were lysed in TNE buffer either in the presence or absence of *N*-ethylmaleimide and IAA (20 mM each). In the absence of *N*-ethylmaleimide and IAA, the higher molecular weight band starting to disappear after 15 min of incubation of the cleared cell lysate on ice. (C) COS-7 cells were transfected with empty plasmids (mock) or plasmids coding for HDGF*Strep*-tag (wt), together with EGFP-SUMO-1, EGFP-SUMO-1ΔGG, HA-Ubc9, or EGFP as indicated. All cell lysates were treated with StrepTactin® beads for specific precipitation of HDGF and SUMO-1 conjugated HDGF (see Experimental procedures). SUMOylated HDGF could only be detected in the eluates of co-expressed HDGF*Strep*-tag and EGFP-SUMO-1wt with or without overexpression of Ubc-9 using antibodies directed against the *Strep*-tag or SUMO-1 as indicated.

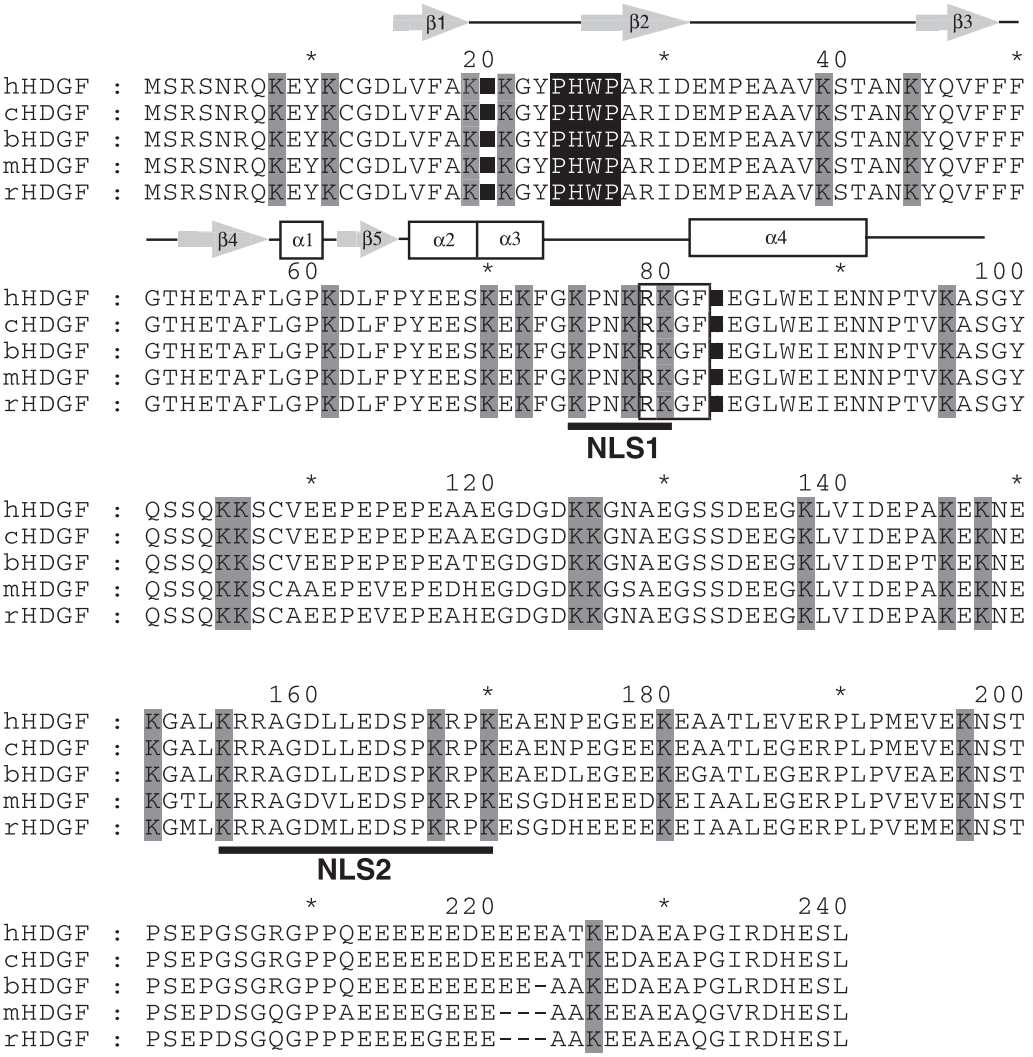


Fig. 2. Amino acid alignment of HDGF from different mammalian species. Sequence alignment of HDGF from human (hHDGF, *Homo sapiens*, Genbank accession CAI16347), chimpanzee (cHDGF, *Pan troglodytes* NCB accession XP_513894), bovine (bHDGF, *Bos taurus*, NCB accession CAB40626), murine (mHDGF, *Mus musculus*, NCB accession BAB30979) and rat (rHDGF, *Rattus norvegicus*, NCB accession AAL47132) origin, respectively. Gaps introduced to generate this alignment are indicated by dashes. Lysine residues within potential SUMOylation motifs are highlighted in gray. K80 in the identified nonconsensus SUMOylation motif $RK_{80}GF$ is boxed. The PWWP motif is indicated by a black background. Sequences for NLS1 and 2 are underlined. The numbers and letters within the secondary structure elements of the *hath* region (amino acids 1–98) are shown as arrows (β-strands) and rectangles (α-helices) according to PDB file 2B8A. Amino acid residues for each sequence are numbered from the initiation methionine.

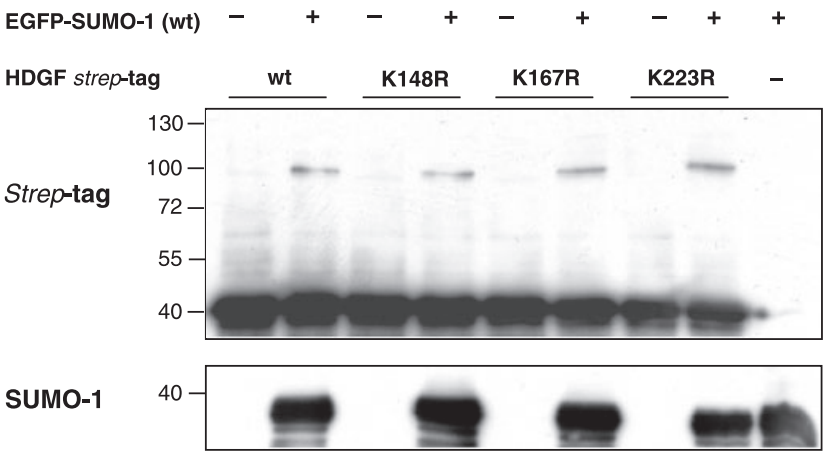


Fig. 3. SUMOylation analysis of HDGF mutants. COS-7 cells were transfected with plasmids coding for different HDGF*Strep*-tag variants (wt, K148R, K167R, K223R), alone or together with EGFP-SUMO-1 as indicated. After precipitation with StrepTactin®, SUMOylated HDGF was detected by western blotting using polyclonal anti-*Strep*-tag serum. To verify equal expression of EGFP-SUMO-1, identical volumes of the cleared cell lysates were analysed by western blotting with a SUMO-1 specific antibody.

K80 in the sequence $RK_{80}GF$ is inconsistent with the postulated consensus motif ΨKxE/D because it contains a positively charged instead of a large hydrophobic first amino acid, and a hydrophobic instead of a negatively charged fourth amino acid. Furthermore, K80 is part of the monopartite NLS1 ($K_{75}PNKRK_{80}$) in the *hath* region of the protein [5] and, thus, the nonexistent SUMOylation of the K80R mutant could have been due to an altered subcellular distribution. To exclude this possibility, we analysed transfected COS-7 cells by immunocytochemistry either in the presence or absence of overexpressed EGFP-SUMO-1 (Fig. 5B). The data obtained clearly demonstrate that HDGF K80R is predominately located in the nucleus very similar to the wild-type protein and that overexpression of SUMO-1 does not obviously alter the location of the mutant protein.

We next investigated whether SUMOylation of HDGF requires nuclear localization of the protein. The previously described NLS2 mutants K155N, K170N and K155/170N [44] are mainly located in the cytoplasm, with almost no HDGF detectable in the nucleus in case of mutant K155/170N (Fig. 6B). Therefore, the SUMOylation status of these HDGF mutant proteins was investigated. Coexpression of the NLS2 mutants together with EGFP-SUMO-1 revealed that SUMOylation levels decrease with increasing amounts of HDGF located in the cytoplasm (K170N < K155N < K155/170N; Fig. 6A). These data are in good agreement with the hypothesis that HDGF SUMOylation occurs in the nucleus.

SUMOylated HDGF is not bound to chromatin

HDGF has been proposed to bind DNA, with this function being mediated by its N-terminal PWWP domain [8,9]. To investigate whether HDGF also binds to chromatin and whether SUMOylation of this protein has any influence on this association, similar chromatin binding assays were performed with extracts from COS-7 cells transfected with HDGFwt alone or co-transfected with HDGFwt and EGFP-SUMO-1. HDGF can be detected in both the Triton-soluble fraction S1 (Fig. 7A) and the Triton-insoluble fraction P1 (Fig. 7A). Most of the HDGF from P1 is solubilized by DNase and a high salt concentration (Fig. 7A). This treatment specifically disrupts Triton-resistant chromatin binding. Only a small amount of HDGF occurs in the cytoskeletal and nuclear matrix fraction P2 (Fig. 7A).

If the cells were co-transfected with EGFP-SUMO-1, the unSUMOylated HDGF shows the same distribution, whereas the SUMOylated HDGF is only found in the Triton-insoluble fractions P1 and P2 (Fig. 7B) and is excluded from the chromatin-bound fraction S2 (Fig. 7B). Interestingly, the signal for SUMOylated HDGF occurs as a double band in fractions P1 and P2 (Fig. 7B). By contrast, mainly the upper band is detected in the total lysate (Fig. 7B) or the Triton-soluble fraction S1 (Fig. 7B). Both EGFP-SUMO-1 and EGFP are exclusively found in the Triton-soluble fraction S1 (Fig. 7C,D). The distribution of the control proteins LEDGF/p75 (chromatin; Fig. 7E), Lamin B1 (nuclear matrix; Fig. 7F) and protein α-tubulin (cytosol; Fig. 7G) was used to confirm the specificity of the assays.

Discussion

To date, besides mass spectrometric evidence for phosphorylation of Ser [33], post-translational modifications of HDGF have not been characterized. In the present study, we demonstrate that HDGF is modified by the attachment of SUMO-1 in mammalian cells and that

the endogenous SUMOylation machineries in COS-7 and HEK293 cells are sufficient to modify HDGF. Similar to other published systems [37], SUMOylation of HDGF is rapidly lost in cell lysates by the action of SUMO proteases upon cell lysis. Although blocking these isopeptidases using *N*-ethylmaleimide and IAA helps to recover a fraction of the SUMO modified protein, it is difficult to estimate the actual degree of SUMOylation in a cell at a given time point. So far, the detection of SUMOylated HDGF from an endogenous source has not been possible. Most likely, this is due to the limitations of the detection systems and the much lower levels of HDGF naturally expressed in the cells studied to date. This is similar to most other published systems [37], where verification of SUMOylation under physiological conditions has been a tremendous problem.

Our search for the SUMO-1 acceptor sites in the primary structure of HDGF did not reveal the clear SUMOylation consensus sequence ΨKxE/D. The best candidates were $EK_{148}NE$ and $PK_{167}RP$ and $AK_{223}KE$ (Fig. 2), which turned out not to be SUMOylated (Fig. 3). Thus, our further attempts to identify the SUMOylation site in HDGF required guidelines other than simply looking for a consensus sequence.

A general problem of identifying nonconsensus SUMOylation sites in substrate proteins with a high number of lysine residues like HDGF is the time consuming preparation of mutant proteins. To overcome this problem, we followed a strategy using MS to identify SUMOylated peptides from proteins expressed in bacteria together with compounds of the SUMO-1 conjugating system [43], without the problem of isopeptidases deconjugating SUMO. Although this system

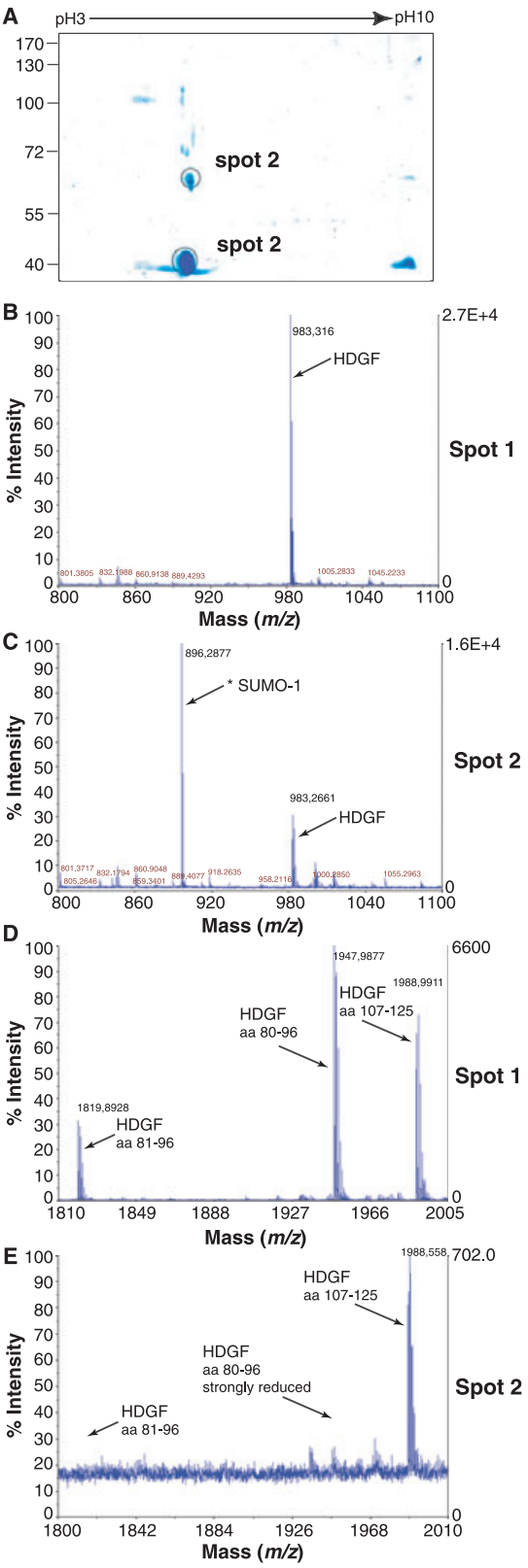


Fig. 4. MALDI-TOF analysis of SUMOylation of HDGF expressed in *E. coli* and identification of K80 and K96 as possible SUMOylation sites. SUMOylated HDGF*Strep*-tag was expressed in *E. coli* and purified using StrepTactin® as described. (A) UnSUMOylated (spot 1) and SUMOylated HDGF (spot 2) were separated by 2D SDS/PAGE. (B, C) MALDI-TOF spectra were obtained from protein spots 1 and 2, respectively, shown in (A). In addition to HDGF specific peptides such as the one comprising amino acids 22–29 (983.49 Da; arrows in B, C), spot 2 contains peptides derived by the tryptic digest of human SUMO-1 (amino acids 64–70; 896.46 Da; asterisk in C). (D, E) MALDI-TOF spectra were obtained from unSUMOylated (spot 1) and SUMOylated (spot 2) HDGF. Whereas the HDGF peptide corresponding to amino acids 107–125 could be obtained from both spots, the spectrum obtained from SUMOylated HDGF (spot 2) shows a specific loss of peptides comprising amino acids 80–96 (1947.9877 Da) and 81–96 (1819.8928 Da). aa, amino acids.

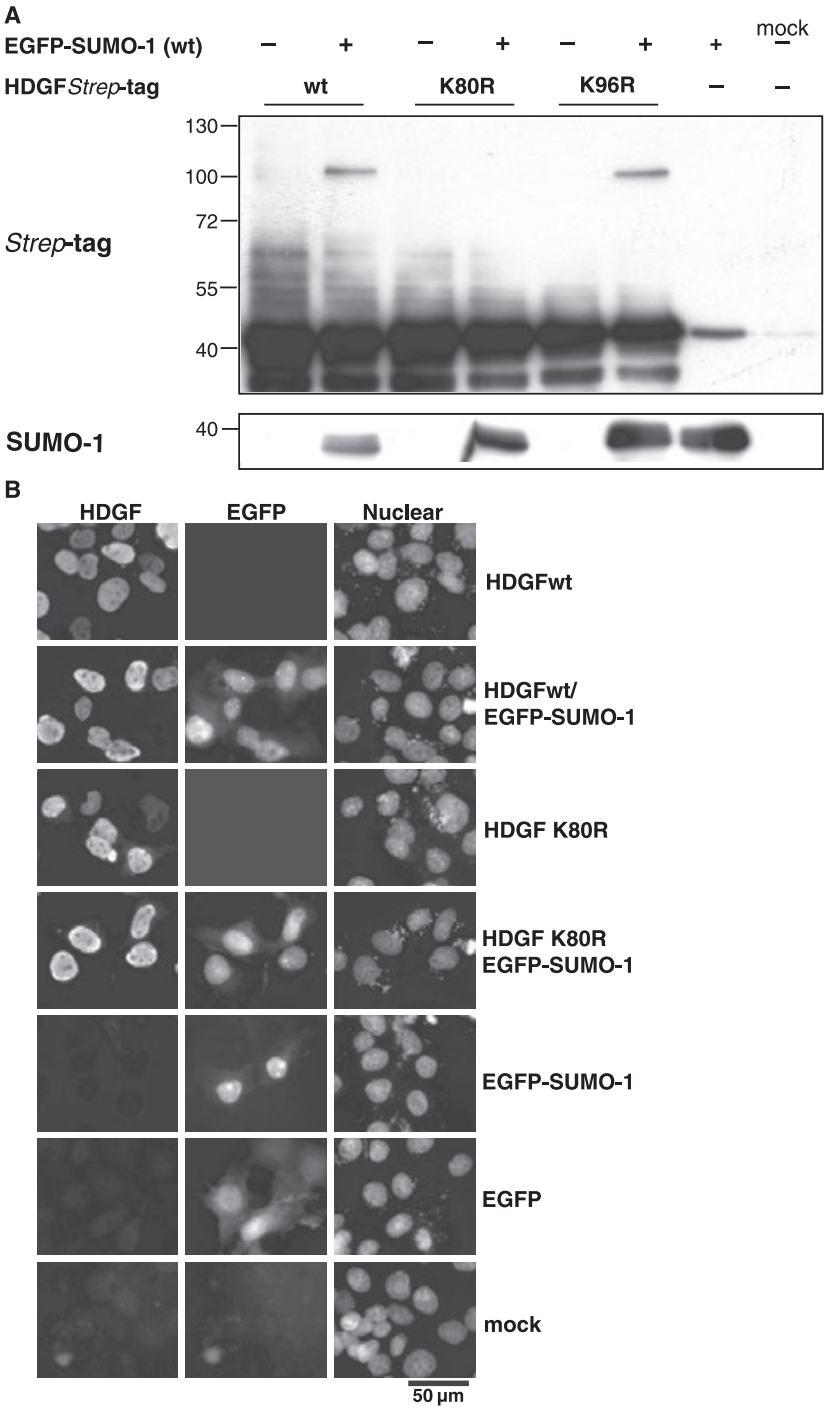


Fig. 5. Identification of lysine 80 as the major HDGF SUMOylation site in mammalian cells. (A) The mammalian cell line COS-7 was transfected with plasmids coding for HDGF*Strep*-tag variants (wt, K80R or K96R), alone or together with a plasmid coding for EGFP-SUMO-1 as indicated. After precipitation with StrepTactin®, SUMOylated HDGF was detected by western blotting using polyclonal anti-*Strep*-tag serum. To verify equal expression of EGFP-SUMO-1, identical volumes of the cleared cell lysates were analyzed by western blotting with a SUMO-1 specific antibody. (B) To demonstrate subcellular localization of HDGF, COS-7 cells on coverslips were transfected with plasmids coding for HDGF*Strep*-tag (wt or K80R) alone or together with EGFP-SUMO-1 as indicated. Paraformaldehyde-fixed cells were permeabilized with methanol and were stained using polyclonal anti-HDGF serum. Immunostaining for HDGF and the EGFP signal for EGFP-SUMO-1 was analyzed by fluorescence microscopy. DAPI was used for nuclear staining. Scale bar = 50 µm.

previously had not been used for SUMOylation of proteins with nonconsensus motifs, we applied it successfully for the preparation of SUMOylated HDGF in sufficient quantities to identify the lysine residue used for the conjugation of SUMO-1 as K80 in the nonconsensus motif $RK_{80}GF$. Lacking the expression of any of the SUMO ligases, overexpression of the conjugating enzyme Ubc-9 in the bacterial SUMOylation system is sufficient to recognize the nonconsensus motif in HDGF. Thus, we are convinced that the bacterial SUMOylation system also is a helpful tool for the identification of SUMO substrates for nonconsensus SUMOylation sites. However, besides K80, at least one other site can be used in the bacterial system

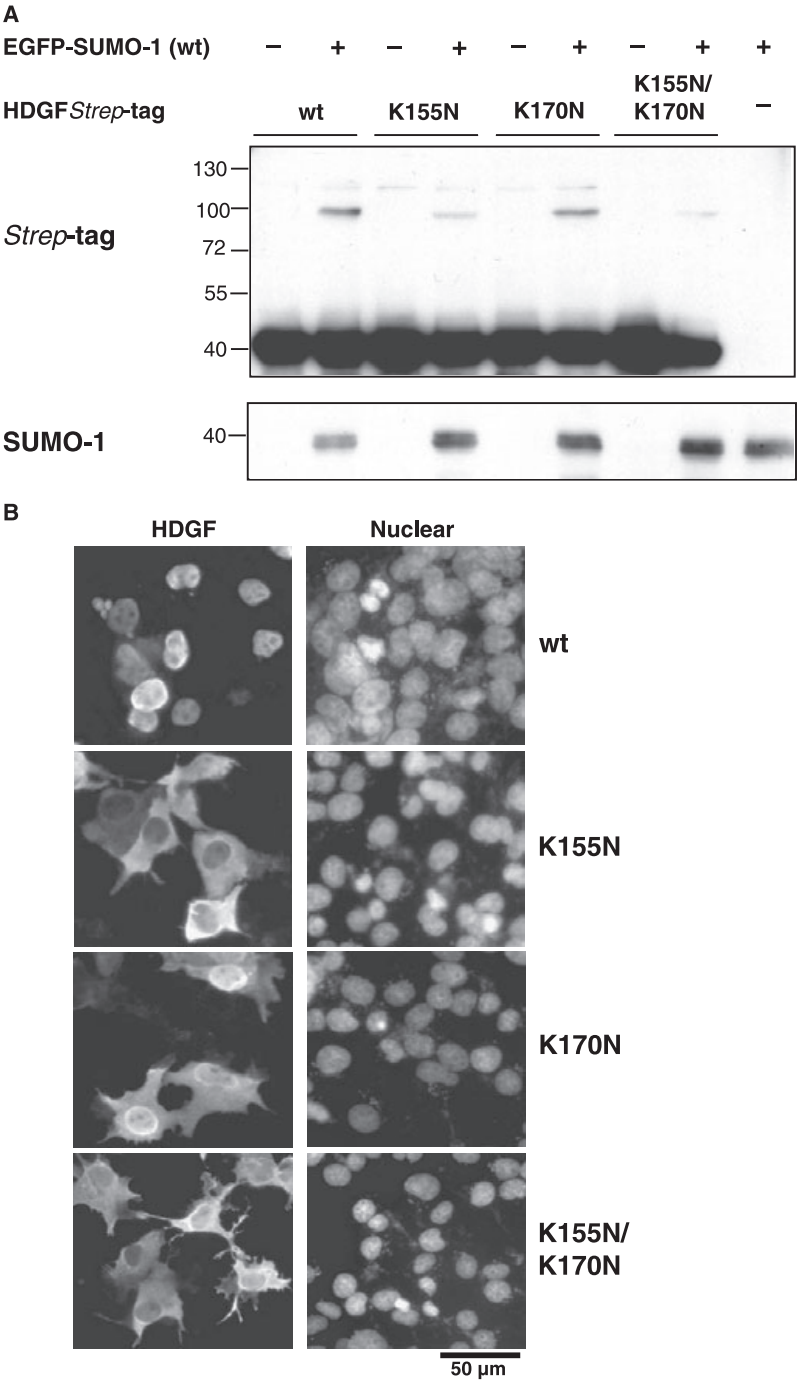


Fig. 6. SUMOylation of HDGF NLS2 mutants. (A) COS-7 cells were transfected with plasmids coding for HDGFStrep-tag variants (wt, K155N, K170N and K155/170N) alone or together with EGFP-SUMO-1 as indicated. After precipitation with StrepTactin®, SUMOylated HDGF was detected by western blotting using polyclonal anti-*Strep*-tag specific serum. To verify the equal expression of EGFP-SUMO-1, identical volumes of the cleared cell lysates were analyzed by western blotting with SUMO-1 specific antibody. (B) To investigate the subcellular localization of HDGF, COS-7 cells on coverslips were transfected with plasmids coding for HDGF*Strep*-tag (wt or NLS 2 mutants) alone or together with EGFP-SUMO-1 as indicated. Paraformaldehyde fixed cells were permeabilized by methanol and were stained using polyclonal anti-HDGF serum. Immunostaining for HDGF and the EGFP signal for EGFP-SUMO-1 was analyzed by fluorescence microscopy. DAPI was used for nuclear staining. Scale bar = 50 µm.

because a K80R HDGF mutant can still be SUMOylated (data not shown). Until now, this site could not be identified by MALDI-TOF-MS analysis. SUMOylation of the K80R mutant in the bacterial system might suggest that a potential second SUMOylation motif is used in mammalian cells only under specific circumstances (e.g. cell cycle state, developmental state, cell type). This possibility is supported by the observation that a very faint band in the range of SUMOylated HDGF is detected sometimes for the K80R HDGF mutant expressed in mammalian cells if the film is overexposed for a very long time.

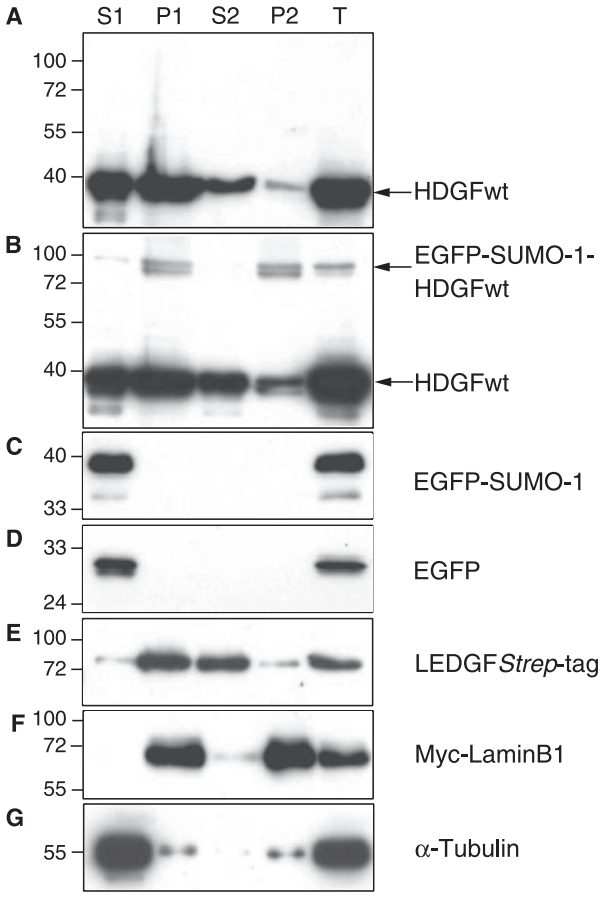


Fig. 7. Chromatin binding assay. COS-7 cells were used to study Triton-resistant chromatin-binding of HDGF and SUMOylated HDGF. Cell fractions were prepared as described in the Experimental procedures. S1, Triton-soluble fraction; P1, Triton-insoluble fraction; S2, DNase and high salt released fraction from P1 (chromatin-bound); P2, Triton-insoluble fraction after DNase and high salt treatment (non-chromatin-bound); T, total lysate. All proteins were detected by western blotting using polyclonal anti-*Strep*-tag specific serum for LEDGF*Strep*-tag, UnSUMOylated and SUMOylated HDGF*Strep*-tag; monoclonal anti-Myc-tag serum for MycLaminB1; and monoclonal anti-α-tubulin specific serum for α-tubulin. (A) Transfection of HDGF*Strep*-tag. (B) Co-transfection of HDGF*Strep*-tag and EGFP-SUMO-1. (C–E) Fractionation patterns of control proteins. LEDGF, LaminB1 and α-tubulin.

The sequence $RK_{80}GF$ with a positively charged amino acid in the first position and a hydrophobic in position four does not match the postulated consensus SUMOylation motif ΨKxE. SUMOylation at nonconsensus motifs has also been shown for proteins such as proliferating cell nuclear antigen [45] and the ubiquitin-conjugating enzyme E2-25K [46]. The SUMOylation site $FK_{14}EV$ in E2-25K is only recognized by the SUMOylation machinery in the structural context of an α-helix whereas, in an unstructured peptide, neighbouring consensus motifs were employed that were not used for conjugation in the correctly folded protein. The same structural context could also be important for K80 in HDGF because of its position adjacent to the α-helix α4 (PDB 2B8A; Fig. 2) spanning amino acids 82–91 [8].

SUMOylation of HDGF correlates with its localization in the nucleus achieving the most prominent effects with NLS2 mutants K155N and K155/170N, respectively (Fig. 6A). This observation is in agreement with the SUMOylation of K80 occuring either during the nuclear translocation or after HDGF has already entered the nucleus. SUMOylation in position 80 does not appear to be essential for the nuclear import or the retention of HDGF in the nucleus because HDGF K80R is located in the nucleus as HDGFwt. At present, it is not possible to investigate specifically the subnuclear distribution of SUMOylated HDGF because there are no appropriate tools to distinguish it from free HDGF and free SUMO-1 present in large excess.

In principle, SUMOylation can have different functional consequences for HDGF. At its most simple, SUMO-1 itself can act as a docking site for proteins containing a SUMO-binding motif [42,47]. Lukasik *et al.* [8] and, more recently, Yang *et al.* [11] showed that the PWWP domain of HDGF is sufficient to bind DNA *in vitro*. By performing chromatin binding assays with cell lysates (Fig. 7A–E), we were able to show that HDGF binds to chromatin in the cell because it occurs in the fraction comprising predominately chromatin-bound proteins. By contrast, SUMOylated HDGF is excluded from this fraction and was only found in the fractions containing cytoskeletal and nuclear matrix proteins. It has been shown for other transcription factors that SUMOylation can have a negative influence on their DNA-binding ability. Several different mechanisms for these effects have been identified. For example, SUMOylation of heat shock factor 2 within its binding domain directly interferes with the DNA-binding site [48], whereas the release of the thymine DNA glycosylase from DNA upon SUMOylation is mediated by a conformational change of a protruded interfering α-helix outside the DNA-binding site of the protein [49]. Furthermore, Tsuruzoe *et al.* [50] showed the negative influence of SUMOylation on DNA-binding and transactivation of the high mobility group protein Sox2. Sox2 is SUMOylated outside the high mobility group domain so that the loss of binding might be explained by conformational changes. Beside this direct influence, SUMO modification could probably also act as a docking site for the recruitment of other proteins that negatively regulate DNA-binding or association with DNA-bound proteins. For DNA methyltransferases

(Dnmt) 3a and 3b and the HDGF family member LEDGF/p75, a PWWP domain mediated DNA-binding activity and chromatin targeting function has been described [51–53]. Very recently, Li *et al.* [54] were able to show SUMOylation of Dnmt3a within an extended PWWP region depending on the interaction with the polycomb group protein chromobox 4 acting as an E3 ligase. The responsible lysine for SUMOylation in Dnmt3a has not yet been identified but, in a similar sequence context to that of the $RK_{80}GF$ motif in HDGF, a $RK_{363}AI$ motif can be found in Dnmt3a. Chemical shift changes in NMR spectra of the HDGF PWWP domain indicated that the putative DNA-binding site overlaps with a patch built by a number of solvent-exposed Lys residues, including K80. Thus, SUMOylation of K80 might have a direct influence on the ability of HDGF to bind to DNA. In the model structure of PDB file 1RI0 [10], the helix is placed in a groove between β2 and the β3/β4 loop in a fixed direction. Interestingly, W26 of the PHWP motif vertically stacks against the aromatic ring of F82 within the helix, whereas K80 is solvent accessible. It is possible that SUMOylation at K80 might alter this conformation or it might just sterically hinder the interaction with the DNA.

Although the data presented in the present study clearly demonstrate that K80 is the main SUMOylation site in HDGF, we cannot exclude with certainty that other SUMOylation sites are used under specific circumstances (e.g. cell cycle state, cell type, development, stress conditions, etc.). Initial evidence for this possibility comes from the observation that, upon overexposure of the films, we noticed a very faint band in the range of SUMOylated HDGF in the case of the K80R mutant in mammalian cells (data not shown).

The C-terminal region of the PWWP domain including α-helix 82–91 of HDGF differs strongly between PWWP domain containing proteins [6,7]. It could be speculated that this region of the HDGF protein is involved in specific protein–protein interactions responsible for targeting the proteins to an appropriate region of DNA. In this case, SUMOylation at position K80 could modify such protein–protein interaction, either by inhibiting or promoting it. For example, the structural data suggests that the decrease in ubiquitin conjugation of E2-25K after SUMO-1 binding is mediated by sterically blocking the binding site of the activating enzyme (E1) instead of affecting the cysteine responsible for ubiquitin conjugation directly [46].

It should be noted that the RKGF motif is also found in other HRP members, such as HRP-2, HRP-3 and LEDGF, whereas, in the testis specific HRP-4 and HRP-1, the corresponding Lys residues are changed to an Arg. Thus, it would be interesting to investigate other HRP family members for their ability to be SUMOylated.

In summary, the present study provides a new perspective for our future understanding of the molecular mechanisms involved in the biological functions of HDGF, and possibly other HRPs, through SUMOylation.

Experimental procedures

Cell culture and transfections

COS-7 (kidney fibroblasts from the African green monkey *Cercopithecus aethiops*) and HEK293 (human embryonic kidney cells) were grown in DMEM supplemented with 10% fetal bovine serum, 100 $U \cdot mL^{-1}$ penicillin and 50 $\mu g \cdot mL^{-1}$ streptomycin at 37 °C in a humidified atmosphere of 5% (vol/vol) CO_2. For transient transfection, cells were grown to a density of 80% confluence. Directly before transfection, the media was changed from 10% fetal bovine serum to 2% fetal bovine serum. Transfection was carried out with the polyethylenimine ExGEN 500 according to the manufacturer's instructions (Fermentas, St Leon-Rot, Germany). For protein expression, COS-7 and HEK293 cells were transfected with the indicated amount of the different plasmids and were harvested after 24 h.

Plasmids

Plasmids were constructed by standard recombinant cloning techniques and all changes were verified by DNA sequencing. For the addition of a *Strep*-tag II peptide (NH_2-WSHPQFEK-COOH) [55] to the C-terminus of HDGF, sense (5′-AATTCGATATCGGTACCTGGAGCCACCCG CAGTTCGAAAAATAAGC-3′) and antisense (5′-GGC CGCTTATTTTTCGAACTGCGGGTGGCTCCAGGTAC CGATATCG-3′) oligonucleotides containing the sequence coding for the Tag and *Eco*RI and *Not*I sites were annealed and cloned into *Eco*RI/*Not*I digested pcDNA3 Amp vector (Invitrogen, Carlsbad, CA, USA).

The coding sequence of HDGFwt (GenBank accession number BC021654) was amplified using murine adult brain cDNA as a source and subcloned in frame into *Hind*III/*Bam*HI digested pcDNA3 Amp *Strep*-tag N-terminal. The primers used for amplification were: 5′-CGAAGCTTAT GTCGCGATCCAACCGGCAG-3′ and 5′-CGGGATCCC TACAGGCTCTCATGATCTCT-3′.

For the expression of untagged HDGF, we amplified mouse adult brain cDNA with the same sense primer and an antisense primer containing an internal Stop codon (5′CGGGATCCTCACTACAGGCTCTCATGATCTCT).

The coding sequence HDGF was cloned into *Hind*III/*Bam*HI digested pcDNA3 Amp.

HDGF mutants were generated by site-directed mutagenesis with the Quickchange Multi-Site Direct Mutagenesis kit (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions using the plasmid HDGF*Strep*-tag wt cloned in pcDNA3 Amp as a template. Positive clones were identified by inserting or removing restriction sites by silent mutations within the following primer pairs: HDGF*Strep*-tag K80R (*Hha*I appears) 5′-GGCA AGCCCAACAAGCGCAGAGGGTTCAGCGAG-3′, 5′-CTCGCTGAACCCTCTGCGCTTGTTGGGCTTGCC-3′; HDGF*Strep*-tag K96R (*Eco*RV appears) 5′-CTACAGT CAGGGCCTCTGGATATCAGTC-3′, 5′-GACTGATAT CCAGAGGCCCTGACTGTAG-3′; HDGF*Strep*-tag K148R (*Sty*I disappears) 5′-AACCAGCCAAAGAGCG GAACGAAAAGGGC-3′, 5′-TTGGTCGGTTTCTCGCC TTGCTTTTCCCG-3′; HDGF*Strep*-tag K167R (*Taq*I disappears) 5′-GAGGACTCCCCCCGACGTCCCAAGGA G-3′, 5′-CTCCTGAGGGGGGCTGCAGGGTTCCTG-3′; and HDGF*Strep*-tag K223R (*Sma*I appears) 5′-GAG GAAGAGGCTGCCCGGGAAGAGGGCTGAAGCC-3′, 5′-GGCTTCAGCCTCTTCCCGGGCAGCCTCTTCCTC-3′.

The primers used for mutagenesis of lysine residues in the NLS 2 were: HDGF*Strep*-tag K155N (*Eco*57I disappears) 5′-GAAAAGGGCACGCTGAATAGGAGAGCAG GG-3′, 5′-CCCTGCTCTCCTATTCAGCGTGCCCTTTT C-3′ and HDGF*Strep*-tag K170N (*Sma*I appears) 5′-CCTAAACGTCCCCGGGAGTCAGGAGAC-3′, 5′-GTC TCCTGACTCCCGGGGACGTTTAGG-3′.

For the double mutant HDGF*Strep*-tag K155/170N, K155N was used as a template for mutagenesis.

For the expression of HDGF*Strep*-tag in *E. coli*, the coding sequence was cloned into the bacterial expression vector pET28a (Merck, Schwalbach, Germany) digested with *Nco*I and *Xho*I using the plasmid encoding HDGF*Strep*-tag as a template and the primers: 5′-CGCCATGGCGCGATCC AACCGGCAG-3′ and 5′-CGCTCGAGTTATTTTTCGA ACTGCGGGTGGC-3′.

Mouse SUMO-1wt, mouse LEDGF/p75wt and mouse LaminB1wt were amplified using mouse adult brain cDNA as a template and the following primers derived from the published cDNA sequences for SUMO-1 (GenBank accession number BC083158): 5′-CGGGATCCATGTCTGACC AGGAGGCA-3′, 5′-ACGCGTCGACCTAAACCGTCG AGTGACC-3′; for LEDGF/p75 (GenBank accession number AF339083): 5′-CGGGATCCATGACTCGCGATTT CAAACCT-3′, 5′-CGGAATTCGTTATCTAGTGTAGA CTC-3′; and for LaminB1 (GenBank accession number NM_010721): 5′-CGAAGATCTATGGCGACCGCGACC CCCGTGC-3′, 5′-CGCTCGAGTCACATAATGGCAC AGCTT-3′.

Mouse SUMO-1ΔGG was amplified using the specific anti sense primer: 5′-ACGCGTCGACCGTTTGTTCCTG ATAAAC-3′.

The fragments for SUMO-1wt and SUMO-1ΔGG were cloned C-terminal in frame to EGFP into *Bg*lII and *Sa*lI digested pEGFP-C1 (Clontech, Palo Alto, CA, USA). The fragment for LEDGF/p75wt was cloned N-terminal in frame to the *Strep*-tag into *Bam*HI and *Eco*RI digested pcDNA III *Strep*-tag vector and LaminB1wt was cloned C-terminal in frame to the Myc-tag into *Bam*HI and *Xho*I digested pCMV3B (Clontech). For the expression of EGFP, the eukaryotic expression plasmid pEGFP C3 (Clontech) was used. N-terminal hemagglutinin HA-tagged murine Ubc9 cloned in pcDNA III Amp was a gift from R. Bernards (Netherlands Cancer Institute, Amesterdam, the Netherlands).

Cell extracts

Cells were washed twice with NaCl/P_i before treatment with ice cold TNE lysis buffer containing 20 mM Tris/Cl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 1% NP-40 and freshly added 20 mM *N*-ethylmaleimide (Sigma-Aldrich, Steinheim, Germany), 20 mM IAA (Sigma-Aldrich) + protease inhibitors (Complete Mini, EDTA free; Roche, Mannheim, Germany) on ice. After 10 min on ice, cells were scraped from the dish, transferred to a 1.5 mL reaction tube and incubated on ice for an additional 10 min. Cell debris was separated by centrifugation at 20 000 ***g*** at 4 °C for 20 min and the supernatant transferred to a new 1.5 mL reaction tube. The cleared lysate was used directly for SDS/PAGE and western blot or for the purification of the recombinant *Strep*-tagged HDGF using StrepTactin® (IBA, Göttingen, Germany).

StrepTactin® purification

Cleared lysate (900 μL) derived from transiently transfected COS-7 cells on a 10 cm diameter cell culture dish (Sarstedt, Nümbrecht, Germany) was incubated in a batch procedure with 30 μL of StrepTactin® Sepharose beads on a rotor shaker for 6 h at 4 °C. After incubation, beads were centrifuged at 240 ***g*** at 4 °C for 15 s and the supernatant was transferred to a new 1.5 mL reaction tube. Beads were washed four times with 1 mL of TNE lysis buffer. After the final washing step, bound proteins were eluted by directly boiling the beads at 95 °C for 10 min in 80 μL 2 × SDS/PAGE sample buffer. After centrifugation, the supernatant was transferred to a fresh 1.5 mL reaction tube. A 20 μL fraction of the eluate was used for SDS/PAGE.

Chromatin binding assay

The method was performed as previously described [53] with slight modification. Twenty-four hours after transfection, COS-7 was lysed for 10 min on ice in 1 mL of ice cold CSK I buffer (10 mM Pipes, pH 6.8, 100 mM NaCl, 1 mM EDTA, 300 mM sucrose, 1 mM $MgCl_2$, 1 mM dithithreitol) supplemented with 0.5% Triton X-100, protease inhibitors (Complete Mini; Roche). Ten percent of the lysate was

mixed with RIPA buffer [150 mM Tris–HCl, pH 8.0, 150 mM NaCl, 0.5% DOC, 0.1% (w/v) SDS, 1% (v/v) NP-40] and kept for further analysis as total lysate (T). The remaining cell lysate was divided into two equal portions, which were centrifuged at 450 ***g*** at 4 °C for 3 min. The supernatants where pooled and represent the Triton-soluble protein fraction (S1). The pellets contain chromatin-bound, nuclear matrix bound and insoluble proteins. One of the pellets was resuspended in RIPA buffer (P1). The second pellet was resuspended in CSK II buffer (10 mM Pipes, pH 6.8, 50 mM NaCl, 300 mM sucrose, 6 mM $MgCl_2$, 1 mM dithiothreitol), treated with 6 $U \cdot \mu L^{-1}$ DNase (Boehringer, Mannheim, Germany) for 30 min followed by extraction with 250 mM NH_4SO_4 for 10 min at 25 °C. The sample treated with DNase and salt was then centrifuged at 1200 ***g*** for 6 min at 4 °C and the supernatant containing DNase-released chromatin-associated proteins (S2) and the pellet containing insoluble, cytoskeletal and nuclear matrix proteins (P2) were both collected. P2 was resuspended in 100 μL RIPA buffer. A total of 20 μg protein of each fraction was loaded for SDS/PAGE.

Immunofluorescence

For transfection and subsequent immunofluorescence, cells (3×10^4) were plated on glass coverslips in a 24-well plate and grown overnight. Cells were washed twice in NaCl/P_i and then fixed in 4% paraformaldehyde in NaCl/P_i for 10 min. For permeabilization, cells were treated for 2 min with ice-cold methanol on ice and subsequently washed twice with NaCl/P_i/2% BSA before blocking unspecific binding sites with DMEM/10% fetal bovine serum for 30 min. HDGF was detected using polyclonal sheep anti-HDGF serum as described previously [56]. Cells were incubated with primary antibody for 90 min in NaCl/P_i/2% BSA at room temperature. After washing twice with NaCl/P_i/2% BSA, cells were treated for 90 min in NaCl/P_i/2% BSA at room temperature in the dark with the corresponding fluorophor-conjugated secondary antibody (Dianova, Hamburg, Germany). Cells were washed in NaCl/P_i/2% BSA and then treated with the nuclear stain 4′-6-diamino-2-phenylindole HCL (DAPI) (Roche) for 1 min at room temperature. After additional washing, cells were embedded in mounting media (Sigma, Deisenhofen, Germany) and images were captured on a fluorescence microscope (Olympus, Hamburg, Germany). EGFP-SUMO-1 was detected by autofluorescence of the EGFP-tag.

Expression of SUMOylated HDGF from *E. coli*

The system of overexpression of SUMOylated proteins in *E. coli* has been described previously [43]. The chemical competent *E. coli* BL21 DE3 (Pharmacia, Uppsala, Sweden) transformed with the plasmid containing the SUMOylation machinery (pE1E2SUMO-1) was then transformed in a second step with the plasmid coding for HDGFwt*Strep*-tag. Overnight cultures were used to inoculate fresh medium (1 : 100, vol/vol); the new cultures were subsequently grown until D_{600} of 0.5 was reached to be induced with a final concentration of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (Sigma) and grown for an additional 4 h. Cells were collected by centrifugation at 5400 ***g*** at 4 °C for 15 min and resuspended in 2% of the original volume with 100 mM Tris–HCl (pH 8.0 at 4 °C), 150 mM NaCl, 1 mM EDTA. The bacteria were sonicated on ice until complete lysis and lysates were clarified by centrifugation at 20 000 ***g*** for 15 min at 4 °C. The recombinant *Strep*-tagged HDGF was affinity purified using StrepTactin® beads according to the manufacturer's instructions.

Western blot analysis

Samples were separated by SDS/PAGE and transferred onto poly(vinylidene difluoride) (Millipore, Schwalbach, Germany) membranes (MiniProtean III; BioRad, München, Germany). The membranes were blocked with 5% nonfat dry milk powder or 5% BSA in Tris-buffered saline 0.15% Tween 20™ for 1 h. Immunodetection was performed by incubating the membranes with the different primary antibodies diluted in blocking buffer overnight at 4 °C. After four washes with Tris-buffered saline 0.15% Tween 20™, the membranes were incubated with secondary antibody conjugated with horseradish peroxidase for 2 h. After four washes, blots were developed with the enhanced chemiluminescence system (Amersham, Braunschweig, Germany) and membranes were exposed to X-ray film (Amersham).

The following antibodies were used: polyclonal sheep anti-HDGF (1 : 1000), rabbit polyclonal anti-*Strep*-tag (1 : 1000; IBA, Göttingen, Germany), polyclonal rabbit anti-SUMO-1 (sc-9060; 1 : 500; Santa Cruz Biotechnology, Santa Cruz, CA, USA), mouse monoclonal anti-EGFP (JL-8; 1 : 5000; Clontech), polyclonal rabbit anti-Ubc9 (sc-10759; 1 : 500; Santa Cruz Biotechnology), polyclonal mouse anti-human α-tubulin (DM1A; 1 : 1000; Sigma) and monoclonal mouse anti-Myc-tag (9E10; 1 : 500; Sigma).

2D electrophoresis

IEF was carried out according to the manufacturer's instructions (BioRad) using immobilized pH gradient strips (7 cm; BioRad) with a linear pH gradient of 3–10. For IEF, 30 μg aliquots of the *Strep*Tactin® purified HDGF*Strep*-tag and SUMO-1 conjugated HDGF*Strep*-tag from the bacterial expression system were acetone precipitated by adding acetone (four times the sample volume) and incubated at −20 °C for 4 h. Precipitated proteins were collected by centrifugation at 20 000 ***g*** at 4 °C for 15 min and dissolved in 125 μL 8 M urea, 2% CHAPS, 50 mM dithiothreitol, 0.2% Bio-Lyte ampholytes (Bio-Rad, Hercules, CA, USA).

MALDI-TOF-MS

Following 2D SDS/PAGE, Coomassise Blue-stained protein spots were excised from gels and processed using a Trypsin Profile IGD Kit according to the manufacturer's instructions (Sigma). After extraction, peptides were premixed with a saturated matrix solution of α-cyano-4-hydroxycinnamic acid in 60% acetonitrile with 0.1% formic acid for spotting onto the target plate. MALDI-TOF-MS analysis was performed on an Applied Biosystem 4700 Proteomics Analyzer mass spectrometer in reflectron positive ion mode (Applied Biosystems, Darmstadt, Germany). For each spectrum, a minimum of 500 shots were accumulated in a mass range of 700–5000 Da. For calibration, a standard peptide mixture procured from Applied Biosystem was spotted next to each sample and used as an external control. After identifying HDGF-derived peptides via a search against the SWISS-PROT database, internal recalibration was performed using these peptides to reach highest mass accuracy.

Acknowledgements

Our thanks go to Dr Yasuhiro Uchimura from the Graduate School of Science and Technology of the Department of Biological Science at the University of Kumamoto for kindly providing the pE1E2S1 plasmid and to Dr R. Bernards from the Netherlands Cancer Institute for kindly providing the HA-Ubc9 expression plasmid. We would also like to thank Mrs Nazila Isakovic for her excellent technical support. This work is supported by the University of Bremen (BFK-Nr.02/608/3 and 02/100/1) and the Tönjes-Vagt-Stiftung.

References

1 Nakamura H, Izumoto Y, Kambe H, Kuroda T, Mori T, Kawamura K, Yamamoto H & Kishimoto T (1994) Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein. *J Biol Chem* **269**, 25143–25149.

2 Dietz F, Franken S, Yoshida K, Nakamura H, Kappler J & Gieselmann V (2002) The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. *Biochem J* **366**, 491–500.

3 Izumoto Y, Kuroda T, Harada H, Kishimoto T & Nakamura H (1997) Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. *Biochem Biophys Res Commun* **238**, 26–32.

4 Everett AD, Stoops T & McNamara CA (2001) Nuclear targeting is required for hepatoma-derived growth factor-stimulated mitogenesis in vascular smooth muscle cells. *J Biol Chem* **276**, 37564–37568.

5 Kishima Y, Yamamoto H, Izumoto Y, Yoshida K, Enomoto H, Yamamoto M, Kuroda T, Ito H, Yoshizaki K & Nakamura H (2002) Hepatoma-derived growth factor stimulates cell growth after translocation to the nucleus by nuclear localization signals. *J Biol Chem* **277**, 10315–10322.

6 Slater LM, Allen MD & Bycroft M (2003) Structural variation in PWWP domains. *J Mol Biol* **330**, 571–576.

7 Stec I, Nagl SB, van Ommen GJB & den Dunnen JT (2000) The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett* **473**, 1–5.

8 Lukasik SM, Cierpicki T, Borloz M, Grembecka J, Everett A & Bushweller JH (2006) High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. *Protein Sci* **15**, 314–323.

9 Nameki N, Tochio N, Koshiba S, Inoue M, Yabuki T, Aoki M, Seki E, Matsuda T, Fujikura Y, Saito M *et al.* (2005) Solution structure of the PWWP domain of the hepatoma-derived growth factor family. *Protein Sci* **14**, 756–764.

10 Sue SC, Chen JY, Lee SC, Wu WG & Huang TH (2004) Solution structure and heparin interaction of human hepatoma-derived growth factor. *J Mol Biol* **343**, 1365–1377.

11 Yang J & Everett AD (2007) Hepatoma derived growth factor binds DNA through the N-terminal PWWP domain. *BMC Mol Biol* **8**, 101.

12 Sue SC, Lee WT, Tien SC, Lee SC, Yu JG, Wu WJ, Wu WG & Huang TH (2007) PWWP module of human hepatoma-derived growth factor forms a domain-swapped dimer with much higher affinity for heparin. *J Mol Biol* **367**, 456–472.

13 Abouzied MM, El Tahir HM, Prenner L, Haberlein H, Gieselmann V & Franken S (2005) Hepatoma-derived growth factor. Significance of amino acid residues 81-100 in cell surface interaction and proliferative activity. *J Biol Chem* **280**, 10945–10954.

14 Enomoto H, Yoshida K, Kishima Y, Okuda Y & Nakamura H (2002) Participation of hepatoma-derived growth factor in the regulation of fetal hepatocyte proliferation. *J Gastroenterol* **37**(Suppl. 14), 158–161.

15 Enomoto H, Yoshida K, Kishima Y, Kinoshita T, Yamamoto M, Everett AD, Miyajima A & Nakamura H (2002) Hepatoma-derived growth factor is highly expressed in developing liver and promotes fetal hepatocyte proliferation. *Hepatology* **36**, 1519–1527.

16 Everett AD, Lobe DR, Matsumura ME, Nakamura H & McNamara CA (2000) Hepatoma-derived growth factor stimulates smooth muscle cell growth and is expressed in vascular development. *J Clin Invest* **105**, 567–575.

17 Everett AD (2001) Identification, cloning, and developmental expression of hepatoma-derived growth factor in the developing rat heart. *Dev Dyn* **222**, 450–458.
18 Everett AD, Narron JV, Stoops T, Nakamura H & Tucker A (2004) Hepatoma-derived growth factor is a pulmonary endothelial cell-expressed angiogenic factor. *Am J Physiol Lung Cell Mol Physiol* **286**, L1194–L1201.
19 Mori M, Morishita H, Nakamura H, Matsuoka H, Yoshida K, Kishima Y, Zhou Z, Kida H, Funakoshi T, Goya S *et al.* (2004) Hepatoma-derived growth factor is involved in lung remodeling by stimulating epithelial growth. *Am J Respir Cell Mol Biol* **30**, 459–469.
20 Oliver JA & Al Awqati Q (1998) An endothelial growth factor involved in rat renal development. *J Clin Invest* **102**, 1208–1219.
21 Abouzied MM, Baader SL, Dietz F, Kappler J, Gieselmann V & Franken S (2004) Expression patterns and different subcellular localization of the growth factors HDGF (hepatoma-derived growth factor) and HRP-3 (HDGF-related protein-3) suggest functions in addition to their mitogenic activity. *Biochem J* **378**, 169–176.
22 Marubuchi S, Okuda T, Tagawa K, Enokido Y, Horiuchi D, Shimokawa R, Tamura T, Qi ML, Eishi Y, Watabe K *et al.* (2006) Hepatoma-derived growth factor, a new trophic factor for motor neurons, is up-regulated in the spinal cord of PQBP-1 transgenic mice before onset of degeneration. *J Neurochem* **99**, 70–83.
23 Zhou Z, Yamamoto Y, Sugai F, Yoshida K, Kishima Y, Sumi H, Nakamura H & Sakoda S (2004) Hepatoma-derived growth factor is a neurotrophic factor harbored in the nucleus. *J Biol Chem* **279**, 27320–27326.
24 Hu TH, Huang CC, Liu LF, Lin PR, Liu SY, Chang HW, Changchien CS, Lee CM, Chuang JH & Tai MH (2003) Expression of hepatoma-derived growth factor in hepatocellular carcinoma. *Cancer* **98**, 1444–1456.
25 Iwasaki T, Nakagawa K, Nakamura H, Takada Y, Matsui K & Kawahara K (2005) Hepatoma-derived growth factor as a prognostic marker in completely resected non-small-cell lung cancer. *Oncol Rep* **13**, 1075–1080.
26 Lepourcelet M, Tou L, Cai L, Sawada J, Lazar AJ, Glickman JN, Williamson JA, Everett AD, Redston M, Fox EA *et al.* (2005) Insights into developmental mechanisms and cancers in the mammalian intestine derived from serial analysis of gene expression and study of the hepatoma-derived growth factor (HDGF). *Development* **132**, 415–427.
27 Ren HN, Tang XM, Lee JJ, Feng L, Everett AD, Hong WK, Khuri FR & Mao L (2004) Expression of hepatoma-derived growth factor is a strong prognostic predictor for patients with early-stage non-small-cell lung cancer. *J Clin Oncol* **22**, 3230–3237.
28 Yamamoto S, Tomita Y, Hoshida Y, Takiguchi S, Fujiwara Y, Yasuda T, Doki Y, Yoshida K, Aozasa K, Nakamura H *et al.* (2006) Expression of hepatoma-derived growth factor is correlated with lymph node metastasis and prognosis of gastric carcinoma. *Clin Cancer Res* **12**, 117–122.
29 Yoshida K, Nakamura H, Okuda Y, Enomoto H, Kishima Y, Uyama H, Ito H, Hirasawa T, Inagaki S & Kawase I (2003) Expression of hepatoma-derived growth factor in hepatocarcinogenesis. *J Gastroen Hepatol* **18**, 1293–1301.
30 Yoshida K, Tomita Y, Okuda Y, Yamamoto S, Enomoto H, Uyama H, Ito H, Hoshida Y, Aozasa K, Nagano H *et al.* (2006) Hepatoma-derived growth factor is a novel prognostic factor for hepatocellular carcinoma. *Ann Surg Oncol* **13**, 159–167.
31 Zhang J, Ren H, Yuan P, Lang W, Zhang L & Mao L (2006) Down-regulation of hepatoma-derived growth factor inhibits anchorage-independent growth and invasion of non-small cell lung cancer cells. *Cancer Res* **66**, 18–23.
32 Okuda Y, Nakamura H, Yoshida K, Enomoto H, Uyama H, Hirotani T, Funamoto M, Ito H, Everett AD, Hada T *et al.* (2003) Hepatoma-derived growth factor induces tumorigenesis in vivo through both direct angiogenic activity and induction of vascular endothelial growth factor. *Cancer Sci* **94**, 1034–1041.
33 Shu HJ, Chen S, Bi Q, Mumby M & Brekken DL (2004) Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line. *Mol Cell Proteom* **3**, 279–286.
34 Hay RT (2005) SUMO: A history of modification. *Mol Cell* **18**, 1–12.
35 Johnson ES (2004) Protein modification by SUMO. *Annu Rev Biochem* **73**, 355–382.
36 Dohmen RJ (2004) SUMO protein modification. *BBA-Mol Cell Res* **1695**, 113–131.
37 Melchior F, Schergaut M & Pichler A (2003) SUMO: ligases, isopeptidases and nuclear pores. *Trends Biochem Sci* **28**, 612–618.
38 Muller S, Ledl A & Schmidt D (2004) SUMO: a regulator of gene expression and genome integrity. *Oncogene* **23**, 1998–2008.
39 Hannich JT, Lewis A, Kroetz MB, Li SJ, Heide H, Emili A & Hochstrasser M (2005) Defining the SUMO-modified proteome by multiple approaches in *Saccharomyces cerevisiae*. *J Biol Chem* **280**, 4102–4110.
40 Hecker CM, Rabiller M, Haglund K, Bayer P & Dikic I (2006) Specification of SUMO1- and SUMO2-interacting motifs. *J Biol Chem* **281**, 16117–16127.
41 Minty A, Dumont X, Kaghad M & Caput D (2000) Covalent modification of p73 alpha by SUMO-1 – two-hybrid screening with p73 identifies novel SUMO-1-interacting proteins and a SUMO-1 interaction motif. *J Biol Chem* **275**, 36316–36323.
42 Song J, Zhang ZM, Hu WD & Chen Y (2005) Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif – a reversal of the bound orientation. *J Biol Chem* **280**, 40122–40129.

43 Uchimura Y, Nakamura M, Sugasawa K, Nakao M & Saitoh H (2004) Overproduction of eukaryotic SUMO-1- and SUMO-2-conjugated proteins in Escherichia coli. *Anal Biochem* **331**, 204–206.

44 Everett AD & Bushweller J (2003) Hepatoma derived growth factor is a nuclear targeted mitogen. *Curr Drug Targets* **4**, 367–371.

45 Hoege C, Pfander B, Moldovan GL, Pyrowolakis G & Jentsch S (2002) RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **419**, 135–141.

46 Pichler A, Knipscheer P, Oberhofer E, van Dijk WJ, Korner R, Olsen JV, Jentsch S, Melchior F & Sixma TK (2005) SUMO modification of the ubiquitin-conjugating enzyme E2-25K. *Nat Struct Mol Biol* **12**, 264–269.

47 Song J, Durrin LK, Wilkinson TA, Krontiris TG & Chen Y (2004) Identification of a SUMO-Binding motif that recognizes SUMO modified proteins. *Mol Biol Cell* **15**, 119A.

48 Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugasawa K, Hanaoka F, Tochio H, Hiroaki H & Shirakawa M (2005) Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. *Nature* **435**, 979–982.

49 Anckar J, Hietakangas V, Denessiouk K, Thiele DJ, Johnson MS & Sistonen L (2006) Inhibition of DNA binding by differential sumoylation of heat shock factors. *Mol Cell Biol* **26**, 955–964.

50 Tsuruzoe S, Ishihara K, Uchimura Y, Watanabe S, Sekita Y, Aoto T, Saitoh H, Yuasa Y, Niwa H, Kawasuji M *et al*. (2006) Inhibition of DNA binding of Sox2 by the SUMO conjugation. *Biochem Biophys Res Commun* **351**, 920–926.

51 Ge YZ, Pu MT, Gowher H, Wu HP, Ding JP, Jeltsch A & Xu GL (2004) Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *J Biol Chem* **279**, 25447–25454.

52 Qiu C, Sawada K, Zhang X & Cheng XD (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Biol* **9**, 217–224.

53 Llano M, Vanegas M, Hutchins N, Thompson D, Delgado S & Poeschla EM (2006) Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J Mol Biol* **360**, 760–773.

54 Li B, Zhou J, Liu P, Hu J, Jin H, Shimono Y, Takahashi M & Xu G (2007) Polycomb protein Cbx4 promotes SUMO modification of de novo DNA methyltransferase Dnmt3a. *Biochem J* **405**, 369–378.

55 Schmidt TGM, Koepke J, Frank R & Skerra A (1996) Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin. *J Mol Biol* **255**, 753–766.

56 El Tahir HM, Dietz F, Dringen R, Schwabe K, Strenge K, Kelm S, Abouzied MM, Gieselmann V & Franken S (2006) Expression of hepatoma-derived growth factor family members in the adult central nervous system. *BMC Neurosci* **7**, 6.

3.2 Publication 2

Secretion of hepatoma-derived growth factor is regulated by N-terminal processing

K Thakar, T Kröcher, S Savant, D Gollnast, S Kelm, F Dietz

2010

Biological Chemistry
(Submitted; MS no. BIOLCHEM-10-181)

BIOLOGICAL CHEMISTRY

Secretion of hepatoma-derived growth factor is regulated by N-terminal processing

Ketan Thakar, Tim Kröcher, Doron Gollnast, Soniya Savant, Sørge Kelm, Frank Dietz [1]

Centre for Biomolecular Interactions Bremen (CBIB), Department of Biochemistry, University of Bremen, Leobener Str. NW2, 28359 Bremen, Germany

[1]To whom correspondence should be addressed:
Centre for Biomolecular Interactions Bremen (CBIB)
Department of Biochemistry, University of Bremen
Leobener Str. im NW2, 28359 Bremen, Germany,
Phone ++4942121863223
FAX ++4942121863221
E-mail: fdietz@uni-bremen.de

15 text pages (including references), 6 Figure pages;
Total word count: 5830; Abstract: 160; Total no. of characters (no spaces): 34,012

Abstract

Hepatoma-derived growth factor (HDGF) was first purified as a growth factor secreted by hepatoma cells. It promotes angiogenesis and has been related to tumorigenesis. To date, little is known about the molecular mechanisms of HDGF functions and especially its routes or regulation of secretion. Here we show that secretion of HDGF requires the N-terminal 10 amino acids and that this peptide can mediate secretion of other proteins like enhanced green fluorescent protein (EGFP) if fused to their N-terminus. Our results further demonstrate that cysteine residues at positions 12 and 108 are linked via an intra-molecular disulfide bridge. Surprisingly, phosphorylation of serine 165 in the C-terminal part of HDGF plays a critical role in the secretion process. If this serine is replaced by alanine, the N-terminus is truncated, the intra-molecular disulfide bridge is not formed and the protein is not secreted. In summary, these observations provide a model how phosphorylation, a disulfide bridge and proteolytic cleavage are involved in HDGF secretion.

Key words

Intra-molecular disulfide bridge; non-classical secretion; N-terminal processing; phosphorylation.

Introduction

Hepatoma-derived growth factor (HDGF) is a heparin binding factor originally purified from culture supernatants of the human hepatoma cell line Huh-7 [Nakamura et al., 1989] or from the rat metanephrogenic mesenchymal cell line 7.1.1 [Nakamura et al., 1994]. HDGF forms the prototype of a growth factor family called *H*DGF-*r*elated *p*roteins (HRPs), which includes HRP-1 [Kuroda et al., 1999], HRP-2 [Izumoto et al., 1997], HRP-3 [Ikegame et al., 1999], HRP-4 [Dietz et al., 2002] and lens epithelium-derived growth factor (LEDGF/p75/p52) [Singh et al., 2000]

HDGF stimulates growth of several cell types including fibroblasts [Nakamura et al., 1994], endothelial cells [Oliver and Al-Awqati, 1998], smooth muscle cells [Everett et al., 2000] and lung epithelial cells [Mori et al., 2004]; and is highly expressed in fetal hepatocytes promoting their growth, suggesting its role in liver development [Enomoto et al., 2002]. Furthermore, HDGF is tumorigenic and a prognostic factor in various cancer forms like in human hepatocellular carcinoma (HCC) [Yoshida et al., 2003], non-small-cell lung cancer (NSCLC) [Ren et al., 2004], gastric carcinoma [Yamamoto et al., 2006], gastrointestinal stromal tumor [Chang et al., 2007], colorectal stromal tumors [Hu et al., 2009]. As reported previously, HDGF may act as a survival factor for cancer cells at least through the regulation of the apoptotic pathways. Knock down experiments of HDGF have demonstrated that it not only influences the Bad-mediate intrinsic apoptotic pathway but also the extrinsic Fas-mediate apoptotic pathway, eventually suppressing anchorage dependent growth of cancerous cells [Tsang et al., 2008; Tsang et al., 2009].

HDGF trafficking is complex. Its translocation into the nucleus is mediated by its mono- and bipartite nuclear localization signals and is required to stimulate DNA replication [Everett et al., 2001; Kishima et al., 2002]. Furthermore, HDGF is secreted by a yet unknown route as it lacks a classical secretion signal. Extracellular HDGF appears to be internalized by binding to heparin sulfate or other mechanisms [Kishima et al., 2002]. Evidence for a potential receptor which specifically binds to extracellular HDGF leading to activation of an intracellular signal cascade has also been demonstrated [Abouzied et al., 2005]. Little is know about post-translational mechanisms regulating HDGF function. Recently, we demonstrated that HDGF can be modified by SUMO-1 [Thakar et al., 2008] and that this occurs at a non-consensus motif leading to a strongly reduced binding of HDGF to chromatin. Phosphorylation of three serine residues (S132, S133 and S165) was identified by mass spectroscopy [Shu et al., 2004]. These modifications may regulate HDGF function since it is rapidly dephosphorylated in response to different apoptotic inducers in endothelial cells [Clermont et al., 2008].

The N-terminal region (residues 1-100) of HDGF is structurally well defined, contains a PWWP domain and is called *hath* region as it is similar in all HRPs. It is made up of a five stranded β-barrel followed by α-helical elements. [Nameki et al., 2005; Lukasik et al., 2006; Sue et al., 2007] and binds double stranded

DNA as well as the glycosaminoglycan heparin. If expressed in bacteria, the *hath* region can dimerise by an unusual domain-swapping mechanism leading to an increased binding affinity for heparin [Sue et al., 2004]. However, the physiological relevance of this phenomenon has remained unclear. The structure of the C-terminal part has remained unknown so far.

In the present study, we demonstrate that the N-terminal peptide of HDGF is essential for its secretion and provide evidence for the participation of serine phosphorylation; an intramolecular disulfide bridge; and proteolytic cleavage in the N-terminus.

Results

Serine 165 is essential for the secretion of HDGF

As no functions have been described for the phosphorylation of HDGF, we wanted to investigate the hypothesis that it regulates secretion of this protein using ectopic expression in fibroblasts. Comparing murine NIH3T3 cells, human HEK 293 cells and monkey COS-7 cells high levels of secreted HDGF was detected by a polyclonal anti-murine HDGF antibody [El-Tahir et al., 2006] only in NIH3T3 cells (Figure 1A) whereas none was detected in the tissue culture supernatants of the other two cell lines (data not shown). Upon transfection of COS-7 or HEK 293 cells with plasmids encoding murine HDGF, secreted protein can easily be detected in cell supernatant (Figure 1B). This is not due to increased cell death as determined by LDH assay (Figure 1C).

We generated HDGF variants in which potential serine phosphorylation sites had been mutated to alanine and expressed these in HEK 293 cells. The expression pattern and levels of secretion for HDGF wt and these mutants are shown in Figure 2A. All mutants were expressed and secreted similar to HDGF wt with the exception of HDGF S165A which was not detected in cell culture supernatants. In addition, most of this mutant appeared to have a lower molecular mass. Next we addressed the question whether the molecular mass change was due to either N- or C-terminal processing. For this, untagged, C-terminal *Strep*-tagged or N-terminal *Strep*-tagged HDGF wt and S165A mutants were expressed in HEK 293 cells. The results shown in Figure 2B clearly indicate that the molecular mass change is due to a processing of the N-terminal region of HDGF. Interestingly, the presence of a C-terminal *Strep*Tag seems to protect the S165A mutant from this processing. Therefore, it was necessary to use untagged proteins in all further secretion studies.

The N-terminus of HDGF mediates secretion

The above described results raised the interesting question whether the N-terminal stretch of HDGF is required for secretion. To explore this idea, we created N-terminally truncated HDGF variants (ΔN11, ΔN13 and ΔN15, indicating the number of N-terminal amino acids deleted) of HDGF wt and analysed their expression and secretion. The results shown in Figure 3 provide clear evidence that the truncation of HDGF wt by 11 or more amino acids prevents its secretion. Furthermore, the data indicate that the N-terminal truncation of HDGF S165A corresponds to the loss of the first 11 or more amino acids. N-terminal amino acid sequencing of the processed S165A mutant (performed by Proteome Factory AG, Berlin, Germany), provided evidence for heterogeneous N-termini starting with ^{13}GDL^{15}, ^{17}AK^{18} or ^{22}GY^{23} suggesting the loss of the 12 to 21 N-terminal amino acids.

The observation that the N-terminal stretch of HDGF is essential for its secretion led to the idea that this part may be sufficient to mediate secretion also of other proteins. To investigate this possibility, we prepared reporter protein (EGFP and SNAP-*Strep*Tag) chimeras extended at their N-termini with peptides representing the N-terminal 10 or 19 amino acids of HDGF. Expression and secretion of these protein chimeras (EGFP, HDGF N10-EGFP, HDGF N19-EGFP, SNAP-*Strep*, HDGF N10-SNAP-*Strep* and HDGF N19-SNAP-*Strep*) by transfected HEK 293 cells was investigated. The results (Figure 4A-B) indicated that the first 10 amino acids of HDGF can mediate the secretion of reporter proteins, while tagging these with the first 19 amino acids of HDGF strongly reduces the amounts of secreted protein.

Disulfide bridges in HDGF

Previously, dimerisation of the HDGF *hath* region (residues 1-100) by an unusual domain-swapping mechanism within the PWWP domain has been reported [Sue et al., 2004]. The involvement of a disulfide bridge via cysteine 12 had been excluded in that study. However, the protein investigated did not include cysteine 108. Since our data suggested that cysteine 12 is lost in HDGF S165A, we reconsidered the possibility of intra- or inter-molecular disulfide bonds in HDGF wt and the S165A mutant expressed in mammalian cells. To investigate this, C12A and C108A mutants of these proteins were expressed in HEK 293 cells and analyzed by SDS-PAGE under reducing and non-reducing conditions. The data presented in Figure 5A suggest that intra- and inter-molecular disulfide bonds occur in HDGF. Under non-reducing conditions the HDGF wt monomer shows a higher electrophoretic mobility than the corresponding C12A and the C108A mutants, whereas under reducing conditions, wt and the mutants migration is identical. Moreover, under non-reducing conditions a distinct band at a molecular mass of about 85kDa is clearly detected for the HDGF C12A but not the C108 mutant. This band most likely represents the homodimer linked via Cys_{108}. Furthermore, N-terminal truncations of the HDGF S165A or C12A/S165A mutants appear to enhance dimer formation. Interestingly, the HDGF ΔN15 variant behaves exactly like the S165A mutant, strongly supporting the idea that N-terminal truncation favors dimer

formation. It appears that disulfide-linked HDGF dimers are not secreted, even if they are not truncated, since no dimers of the C12A mutant was detected in cell supernatants (Figure 5A). Interestingly, secreted HDGF wt, C12A or C108A mutants appear as two bands if analyzed in the presence of DTT, whereas only a single band is detected for HDGF wt under non-reducing conditions. This suggests that some protein molecules are proteolytically cleaved at their N-terminus. The resulting fragments are held together by the Cys_{12}-Cys_{108} disulfide bridge in HDGF wt but not in the cysteine mutants. This cleavage seems not to be an artifact of over-expression as it is also observed for HDGF endogenously expressed in NIH3T3 cells (Figure 5B).

Discussion

The cellular trafficking of HDGF is considered to be *a*typical due to the fact that HDGF lacks a classical secretion signal but is still secreted by a yet unknown pathway.
Here we show for the first time that the N-terminal stretch of HDGF is necessary for secretion. Furthermore, we demonstrate that a peptide comprising the first 10 N-terminal amino acids of HDGF are sufficient to mediate secretion of other proteins as shown for EGFP or SNAP. Surprisingly, extending this "secretion-peptide" to the first 19 amino acids of HDGF strongly reduces this effect. According to the known structure of the HDGF *hath* region, the first 10 amino acids are unstructured and exposed to the solvent while the first structural element of HDGF begins with a β-sheet at Leucine 15 [Sue et al., 2004]. Possibly, the first 10 amino acids of HDGF fused to the reporter protein function as freely accessible peptide as in HDGF, whereas the additional 9 amino acids may lead to a structural element together with the reporter proteins used interfering with the secretion process.

An unexpected observation of this study was that the HDGF S165A mutant is not secreted but truncated at the N-terminus. Phosphorylation of HDGF at serine residues, including S165, has been described previously [Shu et al., 2004] without providing evidence for functions of these modifications. Here we demonstrate that loss of one of these phosphorylation sites, S165, interferes with secretion probably by promoting N-terminal truncation. Furthermore, in this study we demonstrate that HDGF is able to form intra- and inter-molecular disulfide bonds. Indirect evidence for disulfide bonds being essential for HDGF function was already provided in the original description of the mitogenic activity of HDGF enriched from cell culture supernatants [Nakamura et al., 1989].

Based on our analysis of the band pattern observed for intracellular and extracellular HDGF wt and mutants, we propose a model for the possible mechanisms involved in secretion and N-terminal processing as illustrated in Figure 6. According to this model, a conformational change requiring the presence of S165 is necessary to build an intra-molecular disulfide bond between C12 and C108. Some, but not all of the HDGF molecules, are proteolytically cleaved in the N-terminus and the two resulting

fragments are held together by the internal disulfide bridge (Figure 6A). This proteolytic step does not seem to be essential for secretion, since processed HDGF molecules are found both inside and outside of the cell. The intramolecular disulfide bridge itself appears not to be essential for secretion, since both cysteine mutants (HDGF C12A and HDGF C108A) are secreted as monomers (Figure 6B). Surprisingly, also for these mutants the N-terminally truncated forms are found in the supernatant despite the fact that the N-terminus is essential for secretion. Besides the obvious explanation that the cleavage occurred later in the secretion process, it is possible that the N-terminal fragment remains bound non-covalently and thus can support secretion. Both, intramolecular disulfide bridge formation and secretion do not occur in the absence of S165 (Figure 6C). The failure to be secreted can easily be explained by a loss of the N-terminal sequence required for secretion in the corresponding mutants. Obviously, the intramolecular disulfide bridge cannot be formed if the N-terminus is truncated beyond C12 leaving free C108 residues which then leads to the increased amount of disulfide bonded dimers found for HDGF S165A. Recently, an unusual domain swapping mechanism involving the loop between amino acids 31 and 45 has been proposed for the dimerisation of HDGF *hath* region expressed in bacteria [Sue et al., 2007]. In principle, such dimer formation is not in conflict with our model. However, whether such domain swapped dimers exist in fibroblasts and whether they are secreted is unclear. In this context it should be mentioned that we observed that HDGF interacts with HRP-2 and that in particular a HRP-2 splice variant selectively binds to a processed form of HDGF, possibly representing non-phosphorylated at S165 (refer Publication 3).

An interesting and difficult to answer question is how this is structurally related to S165, since the structure of the C-terminal part beyond the *hath* region is unknown. One possible scenario is that phosphorylation of S165 induces a conformational change leading to an interaction of the C-terminal part with the N-terminus of HDGF preventing the loss of the N-terminus and supporting the formation of an intra-molecular disulfide bond between C12 and C108. We noticed that S165 is followed by a proline residue. If phosphorylated such S/T-P motifs are recognized by Pin1 [Landrieu et al., 2001; Sudol et al., 2001], a peptidyl-prolyl isomerase (PPI) catalysing *cis*/*trans* isomerisation of peptide bonds at proline residues, causing a conformational change in the protein substrate [Shaw, 2002]. For example, Pin1 has profound effects on amyloid precursor protein (APP) processing and amyloid-β peptides (Aβ) secretion [Pastorino et al., 2006]. This observation opens the possibility that HDGF secretion might also be regulated by phosphorylation. In the context of the SP motif, S165 is a potential substrate for mitogen-activated protein kinases (MAPK). Interestingly, a recent report provided evidence for a link between HDGF and Erk1/2, a member of the MAPK family [Mao et al., 2008]. Furthermore, in HDGF the sequence ^{8}KEYKCGDLVF^{17} represents a possible MAPK docking motif with 100% identity to the consensus sequence (ELM server to identify functional sites on proteins (*http://elm.eu.org/*)) [Kallunki et al., 1996; Bardwell et al., 2001]. In this context it is relevant to point out that the mutation of S165 leads to a truncated N-terminus resulting in the loss of this potential MAPK docking site. Further on, Pin1 interacts with Cdc2 [Crenshaw et al., 1998] which is another candidate kinase for the phosphorylation of S165 (based on *KinasePhos* prediction server analysis [Huang et al., 2005]). In conclusion, *cis*/*trans*

isomerization of P166 might be a mechanism that could be involved in HDGF processing. However, further studies are necessary to identify the kinase(s) responsible for HDGF phosphorylation and/or the protease(s) involved.

The relevance of interactions between N- and C-terminal parts for maintaining the structure and function have also been described for HMGB1, a protein closely related to HDGF [Thomsen et al., 2004]. The study revealed that the HMGB1 acidic C-terminal domain interacts with the basic N-terminal domain and the intramolecular interaction between the two oppositely charged termini is enhanced when serine residues in the acidic tail of HMGB1 are phosphorylated. Interestingly, HDGF shares 36% protein sequence homology with HMGB1, with particularly high similarity in the C-terminal acidic region, the so-called acidic tail [Nakamura et al., 1994] and has a highly positively charged N-terminal region [Sue et al., 2004], which might lead to the same kind of interaction like observed for HMGB1.

In summary, the N-terminal 10 amino acids of HDGF appear to be responsible and sufficient to mediate non-classical secretion of HDGF and other proteins. Furthermore, a serine residue in the C-terminal part of HDGF is essential for maintaining this N-terminal part, probably by influencing the structure of HDGF.

Materials and Methods

Expression plasmids and recombinant proteins

Plasmids were constructed using standard recombinant cloning techniques and all changes were verified by DNA sequencing. Plasmids coding for C-terminally *Strep*-tagged HDGF wild type (HDGF wt) and untagged HDGF wt were prepared as described before [Thakar et al., 2008]. N-terminally *Strep*-tagged HDGF wt was prepared using the untagged HDGF wt as template, and then subcloned in frame to the N-terminal *Strep*Tag sequence using EcoRI/XhoI restriction sites of the multiple cloning site of the pEXPR-IBA vector (IBA, Göttingen, Germany). The primers used for amplification were
Sense 5'CGGAATTCGATGTCGCGATCCAACCGGCAG and
antisense 5'CGCTCGAGCTACAGGCTCTCTCATGATCT

HDGF mutants were generated by site-directed mutagenesis with the Quickchange Multi-Site Direct Mutagenesis kit (Stratagene, La Jolla, CA, U.S.A.) according to the manufacturer's protocol using the untagged HDGF wt plasmid cloned in pcDNA3 Amp as a template. Positive clones were identified by silent mutations within the following primer pairs
HDGF*untagged* S98A (KpnI silent)
Sense 5´ACAGTCAAGGCCGCTGGGTACCAGTCCTCCCAG
Antisense 5´CTGGGAGGACTGGTACCCAGCGGCCTTGACTGT

HDGF*untagged* S102/103A (without restriction site)
Sense 5´TCTGGCTACCAGGCCGCGCAGAAAAAGAGTTGT
Antisense 5´ACAACTCTTTTTCTGCGCGGCCTGGTAGCCAGA
HDGF*untagged* 128A (PstI silent)
Sense 5´GATAAGAAGGGCGCTGCAGAGGGCAGC
Antisense 5´CTATTCTTCCCGCGACGTCTCCCGTCG
HDGF*untagged* S132A (PstI silent)
Sense 5´AAGGGCTCTGCAGAGGGCGCCAGCGACGAAGAA
Antisense 5´TTCTTCGTCGCTGGCGCCCTCTGCAGCGCCCTT
HDGF*untagged* S133A (PstI silent)
Sense 5´AAGGGCTCTGCAGAGGGCAGCGCCGACGAAGAAGGG
Antisense 5´CCCTTCTTCGTGGCGCTGCCCTCTGCAGCGCCCTT
HDGF*untagged* S165A (XhoI silent)
Sense 5´GATGTGCTCGAGGACGCCCCTAAACGTCCC
Antisense 5´GGGACGTTTAGGGGCGTCCTCGAGCACATC
HDGF*untagged* S202A (PstI silent)
Sense 5´GAGAAGAACAGCACCCCTGCAGAGCCAGACGC
Antisense 5´GCGTCTGGCTCTGCAGGGGTGCTGTTCTTCTC
HDGF*untagged* S206A (NaeI silent)
Sense 5´GAGCCAGACGCCGGCCAGGGACCTCCT
Antisense 5´AGGAGGTCCCTGGCCGGCGTCTGGCTC
HDGF*untagged* C12A (without restriction site)
Sense 5´CAGAAAGAGTACAAGGCTGGAGACCTGGTG
Antisense 5´CACCAGGTCTCCAGCCTTGTACTCTTCTG
HDGF*untagged* C108A (without restriction site)
Sense 5´TCCCAGAAAAAGAGTGCTGCTGCGGCAGAGCCC
Antisense 5´GGGCTCTGCCGCAGCACTCTTTTTCTGGA

For the double mutant HDGF untagged C12A/S165A and C108A/S165A, HDGF S165A untagged was used as a template for mutagenesis.

For construction of HDGF N-terminally truncated mutants HDGF wt in pcDNA3 was used as the template for amplification. HDGF N-terminally truncated fragments were amplified double digested and sub cloned into BamHI/EcoRI pre-cut pcDNA3 vector (Invitrogen, CA, USA) using the antisense primer
5´CGGAATTCCTACAGGCTCTCATGATCTCT and the following sense primers:
HDGF untagged ΔN11
5´CGGGATCCATGTGCGGAGACCTGGTGTTTGCG

HDGF untagged ΔN13
5´CGGGATCCATGGACCTGGTGTTTGCGAAGATG
HDGF untagged ΔN15
5´CGGGATCCATGGTGTTTGCGAAGATGAAAGGA

HDGF wt with N-terminal *Strep*Tag was used as the template to amplify varying lengths of peptides representing the N-terminal amino acids of HDGF. The PCR amplified product was BglII/SalI double digested, cleaned and ligated to BglII/SalI pre-cut pEGFP-N3 vector (Clontech, CA, USA) to obtain EGFP chimeras extended at its N-termini with peptides representing the N-terminal amino acids of HDGF (N-terminal *Strep*Tag HDGF ^{peptide}-EGFP plasmids). Amplification was carried out using the sense primer 5´CGAGATCTATGGCTAGCTGGAGCCACCCG and antisense primer specific for different lengths of the peptide as follows:
HDGF N10
5´ GCGTCGACGTACTCTTTCTGCCGGTTGGA
HDGF N15
5´ GCGTCGACCAGGTCTTCGCACTTGTACTC
HDGF N19
5´ GCGTCGACGTACTCTTTCTGCCGGTTGGA

Further, these constructs were EcoRI/NotI double digested, cleaned and ligated to EcoRI/NotI pre-cut pcDNA3 vector to obtain untagged HDGF ^{peptide}-EGFP plasmids.

For production of HDGF-SNAP-*Strep*Tag in pcDNA3, the coding sequence of the hAGT protein (SNAP) was amplified using the pSNAP-tag® (m) vector (NEB, Ipswich, MA, USA) as a template and subcloned in frame into BamHI/EcoRI digested pcDNA3 Amp *Strep-tag*. The primers used for amplification were sense
5´CGGGATCCCTGGAGGTGCTGTTCCAGGGCCCCATGGACAAAGACTGCGAAATGAAGCG-3´
including the coding sequence for the 3C protease recognition site of the human rhinovirus HRV 3C (LEVLFQGP, coding sequence underlined) and antisense primer
5´CGGAATTCACCCAGCCCAGGCTTGCCCAGA.

HDGF ^{peptide}-EGFP constructs in pcDNA3 vector were BamHI/NotI double digested to remove EGFP. Simultaneously HDGF-SNAP-*Strep*Tag in pcDNA3 was BamHI/NotI double digested to remove SNAP-*Strep*Tag and ligated to previously cut pcDNA3 to obtain HDGF ^{peptide} -SNAP-*Strep*Tag constructs in pcDNA3.

Cell culture and transfections

NIH3T3 (Mouse embryonic fibroblast cell line), COS-7 (kidney fibroblasts from the African green monkey *Cercopithecus aethiops*) and HEK 293 (human embryonic kidney cells) were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% foetal calf serum (FCS), 100 U/mL penicillin and 50µg/mL streptomycin at 37°C in a humidified atmosphere of 5% (vol/vol) CO_2. For transient transfection cells were grown to a density of 80% confluency. Directly before transfection the media was changed from 10% to 2% FCS. Transfection was carried out with the polyethylenimine ExGen 500 according to the manufacturer's protocol (Fermentas, St. Leon-Rot, Germany). For protein expression, COS-7 and HEK 293 cells were transfected with the indicated amount of the different plasmids and were harvested after 24 h.

Preparation of protein extracts and supernatant samples

Cells were washed twice with PBS before treatment with ice cold (1mL/10 cm plate) TNE lysis buffer (20mM Tris/Cl pH 7.4, 150 mM NaCl, 5mM EDTA, 1% NP-40 + protease inhibitors (Roche, Mannheim, Germany)) on ice. After 10 min on ice, cells were scraped from the dish, transferred to a 1.5 mL reaction tube and incubated on ice for additional 10 min. Cell debris was separated by centrifugation, 12000 *g* at 4°C for 20 min, and the supernatant transferred to a new 1.5 mL reaction tube. The cleared lysate was used directly for SDS-PAGE and western blot. Supernatant from cell cultures transiently transfected was collected after 24 h; centrifuged, 12000 *g* at 4°C for 20 min, to clarify cell debris. Four times the volume of acetone was mixed with 250 µL clarified supernatant and incubated at -20°C for 2 h to precipitate proteins. The mixture was then centrifuged, 12000 *g* at 4°C for 15 min. Precipitated protein pellet was resuspended in 80 µL 2x SDS-PAGE sample buffer, incubated at 95°C for 5 min and was used directly for SDS-PAGE and Western blot.

Western Blot analysis

Samples were separated by SDS-PAGE (MiniProtean III; BioRad, München, Germany) and transferred onto poly (vinylidene difluoride) (PVDF; Millepore, Schwalbach, Germany) membranes. The membranes were blocked with 5% non-fat dry milk powder or 5% bovine serum albumin (BSA) in Tris-buffered saline 0.15% Tween20™ for 1h. Immunodetection was performed by incubating the membranes with the different primary antibodies diluted in blocking buffer over-night at 4°C. After four washes with Tris-buffered saline 0.15% Tween 20™, the membranes were incubated with secondary antibody conjugated with horseradish peroxidase for 2 h. After four washes, blots were developed with the enhanced chemiluminescence system (Amersham, Braunschweig, Germany), and membranes were exposed to X-ray film (Amersham, Braunschweig, Germany). The following antibodies were used: polyclonal sheep anti HDGF (1:1000), polyclonal sheep anti *Strep*-tag (1:1000; IBA, Göttingen, Germany),

monoclonal mouse anti EGFP (JL-8; 1:5000; Clontech, Palo Alto, CA, U.S.A.) monoclonal mouse anti Actin (1:2000; Dianova, Hamburg, Germany)

Determination of cell viability

Cell viability was analyzed by determining the activity of lactate dehydrogenase (LDH) released into culture supernatants within 24 h incubation [Dringen and Hamprecht, 1998]. Briefly, 10 µL of the incubation medium were added to 190 µL 80 mM TrisHCl buffer pH 7.2 containing 200 mM NaCl followed by 200 µL reaction mixture (80 mM Tris HCl buffer pH 7.2 containing 200 mM NaCl, 3.2 mM pyruvate and 0.4 mM NADH (final concentration of NADH: 0.2 mM)). The decrease in the absorbance at 340 nm was recorded for up to 10 min in 30 s intervals at room temperature. The total LDH activity present in the cells was determined using Triton X-100 (1% (w/v) for 30 min) extracts of the cells. LDH activities released in supernatants exceeding 15% were considered as indication for increased cell lysis.

Acknowledgments

Our thanks go to Mrs. Nazila Isakovic for her excellent technical support. This work is supported by the Central Research Funding (ZFK, BFK NaWi), University of Bremen, and the Tönjes-Vagt-Stiftung, Project XXI.

References

Abouzied, M.M., El-Tahir, H.M., Prenner, L., Haberlein, H., Gieselmann, V., Franken, S. (2005). Hepatoma-derived growth factor. Significance of amino acid residues 81-100 in cell surface interaction and proliferative activity. J Biol Chem *280*, 10945-10954.

Bardwell, A.J., Flatauer, L.J., Matsukuma, K., Thorner, J., Bardwell, L. (2001). A conserved docking site in MEKs mediates high-affinity binding to MAP kinases and cooperates with a scaffold protein to enhance signal transmission. J Biol Chem *276*, 10374-10386.

Chang, K.C., Tai, M.H., Lin, J.W., Wang, C.C., Huang, C.C., Hung, C.H., Chen, C.H., Lu, S.N., Lee, C.M., Changchien, C.S., Hu, T.H. (2007). Hepatoma-derived growth factor is a novel prognostic factor for gastrointestinal stromal tumors. Int J Cancer *121*, 1059-1065.

Clermont, F., Gonzalez, N.S., Communi, D., Franken, S., Dumont, J.E., Robaye, B. (2008). HDGF is dephosphorylated during the early steps of endothelial cell apoptosis in a caspase-dependent way. J Cell Biochem *104*, 1161-1171.

Crenshaw, D.G., Yang, J., Means, A.R., Kornbluth, S. (1998). The mitotic peptidyl-prolyl isomerase, Pin1, interacts with Cdc25 and Plx1. EMBO J *17*, 1315-1327.

Dietz, F., Franken, S., Yoshida, K., Nakamura, H., Kappler, J., Gieselmann, V. (2002). The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. Biochem J *366*, 491-500.

Dringen, R., Hamprecht, B. (1998). Glutathione restoration as indicator for cellular metabolism of astroglial cells. Dev Neurosci *20*, 401-407.

El-Tahir, H.M., Dietz, F., Dringen, R., Schwabe, K., Strenge, K., Kelm, S., Abouzied, M.M., Gieselmann, V., Franken, S. (2006). Expression of hepatoma-derived growth factor family members in the adult central nervous system. BMC Neurosci *7*, 6.

Enomoto, H., Yoshida, K., Kishima, Y., Kinoshita, T., Yamamoto, M., Everett, A.D., Miyajima, A., Nakamura, H. (2002). Hepatoma-derived growth factor is highly expressed in developing liver and promotes fetal hepatocyte proliferation. Hepatology *36*, 1519-1527.

Everett, A.D., Lobe, D.R., Matsumura, M.E., Nakamura, H., McNamara, C.A. (2000). Hepatoma-derived growth factor stimulates smooth muscle cell growth and is expressed in vascular development. J Clin Invest *105*, 567-575.

Everett, A.D., Stoops, T., McNamara, C.A. (2001). Nuclear targeting is required for hepatoma-derived growth factor-stimulated mitogenesis in vascular smooth muscle cells. J Biol Chem *276*, 37564-37568.

Hu, T.H., Lin, J.W., Chen, H.H., Liu, L.F., Chuah, S.K., Tai, M.H. (2009). The expression and prognostic role of hepatoma-derived growth factor in colorectal stromal tumors. Dis Colon Rectum *52*, 319-326.

Huang, H.D., Lee, T.Y., Tzeng, S.W., Horng, J.T. (2005). KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites. Nucleic Acids Res *33*, W226-229.

Ikegame, K., Yamamoto, M., Kishima, Y., Enomoto, H., Yoshida, K., Suemura, M., Kishimoto, T., Nakamura, H. (1999). A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus. Biochem Biophys Res Commun *266*, 81-87.

Izumoto, Y., Kuroda, T., Harada, H., Kishimoto, T., Nakamura, H. (1997). Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. Biochem Biophys Res Commun *238*, 26-32.

Kallunki, T., Deng, T., Hibi, M., Karin, M. (1996). c-Jun can recruit JNK to phosphorylate dimerization partners via specific docking interactions. Cell *87*, 929-939.

Kishima, Y., Yamamoto, H., Izumoto, Y., Yoshida, K., Enomoto, H., Yamamoto, M., Kuroda, T., Ito, H., Yoshizaki, K., Nakamura, H. (2002). Hepatoma-derived growth factor stimulates cell growth after translocation to the nucleus by nuclear localization signals. J Biol Chem *277*, 10315-10322.

Kuroda, T., Tanaka, H., Nakamura, H., Nishimune, Y., Kishimoto, T. (1999). Hepatoma-derived growth factor-related protein (HRP)-1 gene in spermatogenesis in mice. Biochem Biophys Res Commun *262*, 433-437.

Landrieu, I., Odaert, B., Wieruszeski, J.M., Drobecq, H., Rousselot-Pailley, P., Inze, D., Lippens, G. (2001). p13(SUC1) and the WW domain of PIN1 bind to the same phosphothreonine-proline epitope. J Biol Chem *276*, 1434-1438.

Lukasik, S.M., Cierpicki, T., Borloz, M., Grembecka, J., Everett, A., Bushweller, J.H. (2006). High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. Protein Sci *15*, 314-323.

Mao, J., Xu, Z., Fang, Y., Wang, H., Xu, J., Ye, J., Zheng, S., Zhu, Y. (2008). Hepatoma-derived growth factor involved in the carcinogenesis of gastric epithelial cells through promotion of cell proliferation by Erk1/2 activation. Cancer Sci *99*, 2120-2127.

Mori, M., Morishita, H., Nakamura, H., Matsuoka, H., Yoshida, K., Kishima, Y., Zhou, Z., Kida, H., Funakoshi, T., Goya, S., Yoshida, M., Kumagai, T., Tachibana, I., Yamamoto, Y., Kawase, I., Hayashi, S. (2004). Hepatoma-derived growth factor is involved in lung remodeling by stimulating epithelial growth. Am J Respir Cell Mol Biol *30*, 459-469.

Nakamura, H., Izumoto, Y., Kambe, H., Kuroda, T., Mori, T., Kawamura, K., Yamamoto, H., Kishimoto, T. (1994). Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein. J Biol Chem *269*, 25143-25149.

Nakamura, H., Kambe, H., Egawa, T., Kimura, Y., Ito, H., Hayashi, E., Yamamoto, H., Sato, J., Kishimoto, S. (1989). Partial purification and characterization of human hepatoma-derived growth factor. Clin Chim Acta *183*, 273-284.

Nameki, N., Tochio, N., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Fujikura, Y., Saito, M., Ikari, M., Watanabe, M., Terada, T., Shirouzu, M., Yoshida, M., Hirota, H., Tanaka, A., Hayashizaki, Y., Guntert, P., Kigawa, T., Yokoyama, S. (2005). Solution structure of the PWWP domain of the hepatoma-derived growth factor family. Protein Sci *14*, 756-764.

Oliver, J.A., Al-Awqati, Q. (1998). An endothelial growth factor involved in rat renal development. J Clin Invest *102*, 1208-1219.

Pastorino, L., Sun, A., Lu, P.J., Zhou, X.Z., Balastik, M., Finn, G., Wulf, G., Lim, J., Li, S.H., Li, X., Xia, W., Nicholson, L.K., Lu, K.P. (2006). The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-beta production. Nature *440*, 528-534.

Ren, H., Tang, X., Lee, J.J., Feng, L., Everett, A.D., Hong, W.K., Khuri, F.R., Mao, L. (2004). Expression of hepatoma-derived growth factor is a strong prognostic predictor for patients with early-stage non-small-cell lung cancer. J Clin Oncol *22*, 3230-3237.

Shaw, P.E. (2002). Peptidyl-prolyl isomerases: a new twist to transcription. EMBO Rep *3*, 521-526.

Shu, H., Chen, S., Bi, Q., Mumby, M., Brekken, D.L. (2004). Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line. Mol Cell Proteomics *3*, 279-286.

Singh, D.P., Kimura, A., Chylack, L.T., Jr., Shinohara, T. (2000). Lens epithelium-derived growth factor (LEDGF/p75) and p52 are derived from a single gene by alternative splicing. Gene *242*, 265-273.

Sudol, M., Sliwa, K., Russo, T. (2001). Functions of WW domains in the nucleus. FEBS Lett *490*, 190-195.

Sue, S.C., Chen, J.Y., Lee, S.C., Wu, W.G., Huang, T.H. (2004). Solution structure and heparin interaction of human hepatoma-derived growth factor. J Mol Biol *343*, 1365-1377.

Sue, S.C., Lee, W.T., Tien, S.C., Lee, S.C., Yu, J.G., Wu, W.J., Wu, W.G., Huang, T.H. (2007). PWWP module of human hepatoma-derived growth factor forms a domain-swapped dimer with much higher affinity for heparin. J Mol Biol *367*, 456-472.

Thakar, K., Niedenthal, R., Okaz, E., Franken, S., Jakobs, A., Gupta, S., Kelm, S., Dietz, F. (2008). SUMOylation of the hepatoma-derived growth factor negatively influences its binding to chromatin. FEBS J *275*, 1411-1426.

Thomsen, M.S., Franssen, L., Launholt, D., Fojan, P., Grasser, K.D. (2004). Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein. Biochemistry *43*, 8029-8037.

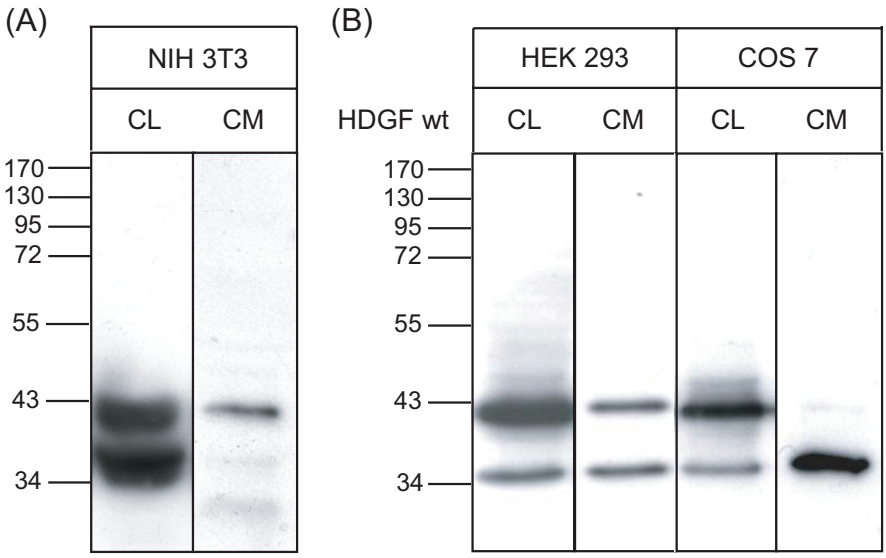
Tsang, T.Y., Tang, W.Y., Tsang, W.P., Co, N.N., Kong, S.K., Kwok, T.T. (2008). Downregulation of hepatoma-derived growth factor activates the Bad-mediated apoptotic pathway in human cancer cells. Apoptosis *13*, 1135-1147.

Tsang, T.Y., Tang, W.Y., Tsang, W.P., Co, N.N., Kong, S.K., Kwok, T.T. (2009). Mechanistic study on growth suppression and apoptosis induction by targeting hepatoma-derived growth factor in human hepatocellular carcinoma HepG2 cells. Cell Physiol Biochem *24*, 253-262.

Yamamoto, S., Tomita, Y., Hoshida, Y., Takiguchi, S., Fujiwara, Y., Yasuda, T., Doki, Y., Yoshida, K., Aozasa, K., Nakamura, H., Monden, M. (2006). Expression of hepatoma-derived growth factor is correlated with lymph node metastasis and prognosis of gastric carcinoma. Clin Cancer Res *12*, 117-122.

Yoshida, K., Nakamura, H., Okuda, Y., Enomoto, H., Kishima, Y., Uyama, H., Ito, H., Hirasawa, T., Inagaki, S., Kawase, I. (2003). Expression of hepatoma-derived growth factor in hepatocarcinogenesis. J Gastroenterol Hepatol *18*, 1293-1301.

Figures



(C)

	LDH level (% max. release)
NIH 3T3	11,7870 ± 2,8086
HEK 293	8,3330 ± 3,6673
COS-7	7,8444 ± 1,5288

Figure 1 HDGF secretion of under endogenous and upon over-expression.

Cell lysates (CL) and media (CM) for (A) NIH3T3 cells and (B) HEK 293 and COS-7 cells transfected with plasmids coding for the HDGF wt were lysed and analysed by SDS-PAGE and Western blotting with HDGF specific antibody (C) LDH assay results for cell media (CM) of NIH 3T3, HEK 293 and COS-7 to assess cell lysis.

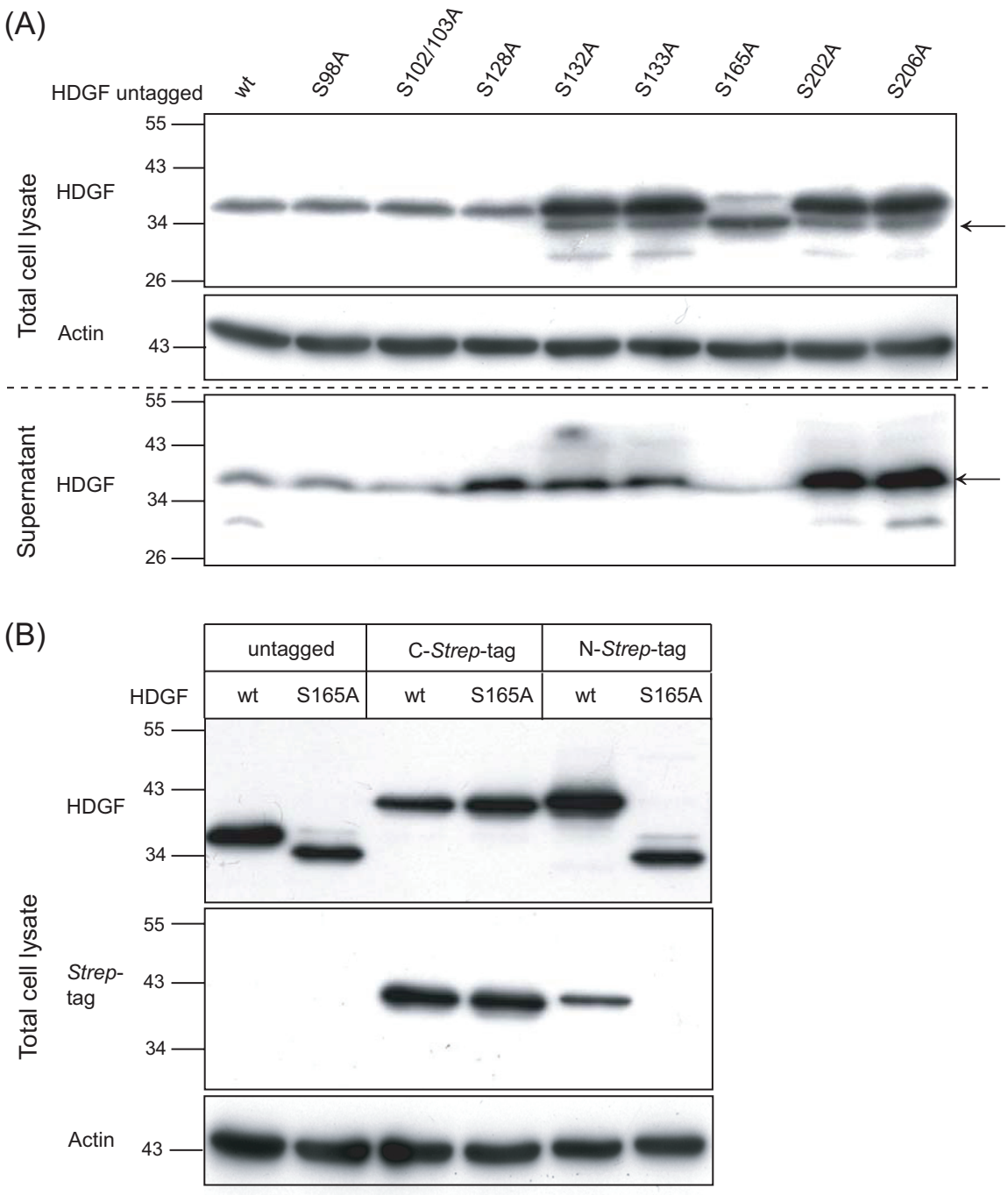


Figure 2 Serine 165 regulates HDGF secretion by processing of the N-terminus.

(A) HEK 293 cells were transfected with plasmids coding for HDGF wt or the indicated serine to alanine mutants. Intracellular (cell lysates) and secreted (supernatant) proteins were analysed by SDS-PAGE and Western blotting with HDGF specific antibody as described under Material and Methods. Arrows point at the lower molecular mass form of HDGF S165A mutant. (B) HEK 293 cells were transfected with plasmids coding for untagged, C-terminally *Strep*-tagged and N-terminally *Strep*-tagged HDGF wt or S165A were lysed and analysed by SDS-PAGE and Western blotting with HDGF and *Strep*-tag specific antibodies.

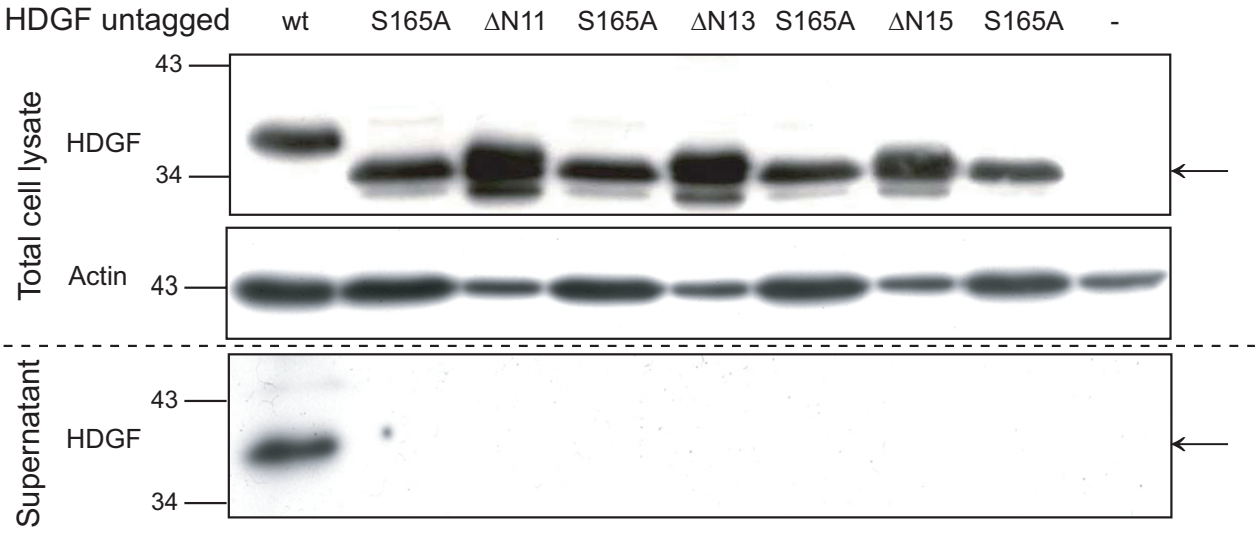
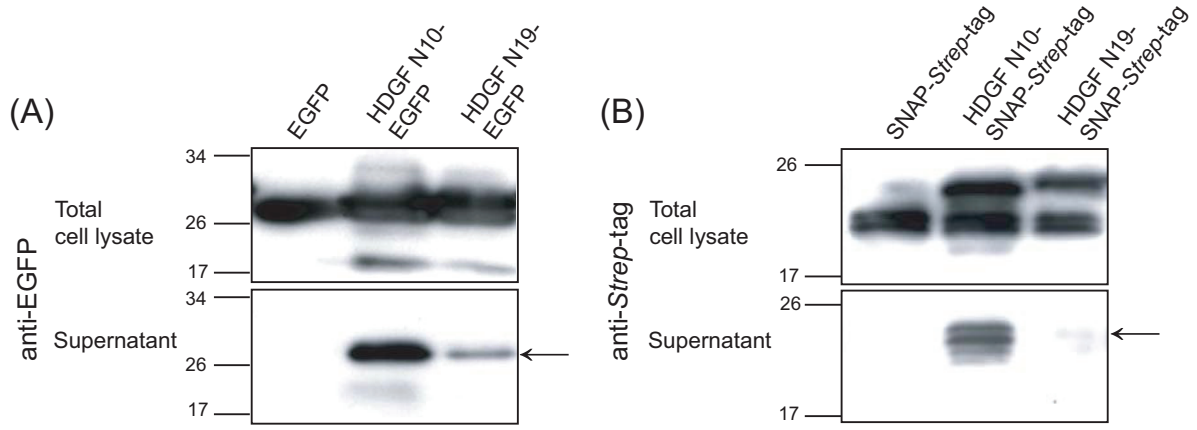


Figure 3 N-terminus truncation of HDGF abolishes secretion.

HEK 293 cells were transfected with plasmids coding for HDGF wt, S165A or N-terminally truncated ΔN11, ΔN13 or ΔN15 HDGF variants. Intracellular (cell lysates) and secreted (supernatant) proteins were analysed by SDS-PAGE and Western blotting with HDGF specific antibody. Arrow points at the lower molecular mass form of HDGF S165A mutant and the truncation variants (upper panel) and secreted HDGF wt protein in the supernatant (lower panel).



(C)

	LDH level (% max. release)
EGFP	16,1870 ± 2.0167
HDGF N10-EGFP	10,4260 ± 2.2252
HDGF N19- EGFP	13,4260 ± 1.2252
SNAP-*Strep*-tag	18,1870 ± 1.3167
HDGF N10-SNAP-*Strep*-tag	13,0260 ± 1.2287
HDGF N19- SNAP-*Strep*-tag	11,4260 ± 1.4452

Figure 4 HDGF N-terminal 10 amino acids of mediate secretion of other proteins.

(A) HEK 293 cells were transfected with plasmids coding EGFP, HDGF N10-EGFP or HDGF N19-EGFP plasmids. Intracellular (cell lysates) and secreted (supernatant) proteins were analysed by SDS-PAGE and Western blotting with EGFP specific antibody. Arrow points out secreted EGFP detected in the supernatant. (B) HEK 293 cells transfected with plasmids coding SNAP-*Strep*Tag, HDGF N10-SNAP-*Strep*Tag or HDGF N19-SNAP-*Strep*Tag plasmids wt were lysed and analysed by SDS-PAGE and Western blotting with *Strep*Tag specific antibody. Arrow points out secreted SNAP-*Strep*Tag detected in the supernatant.

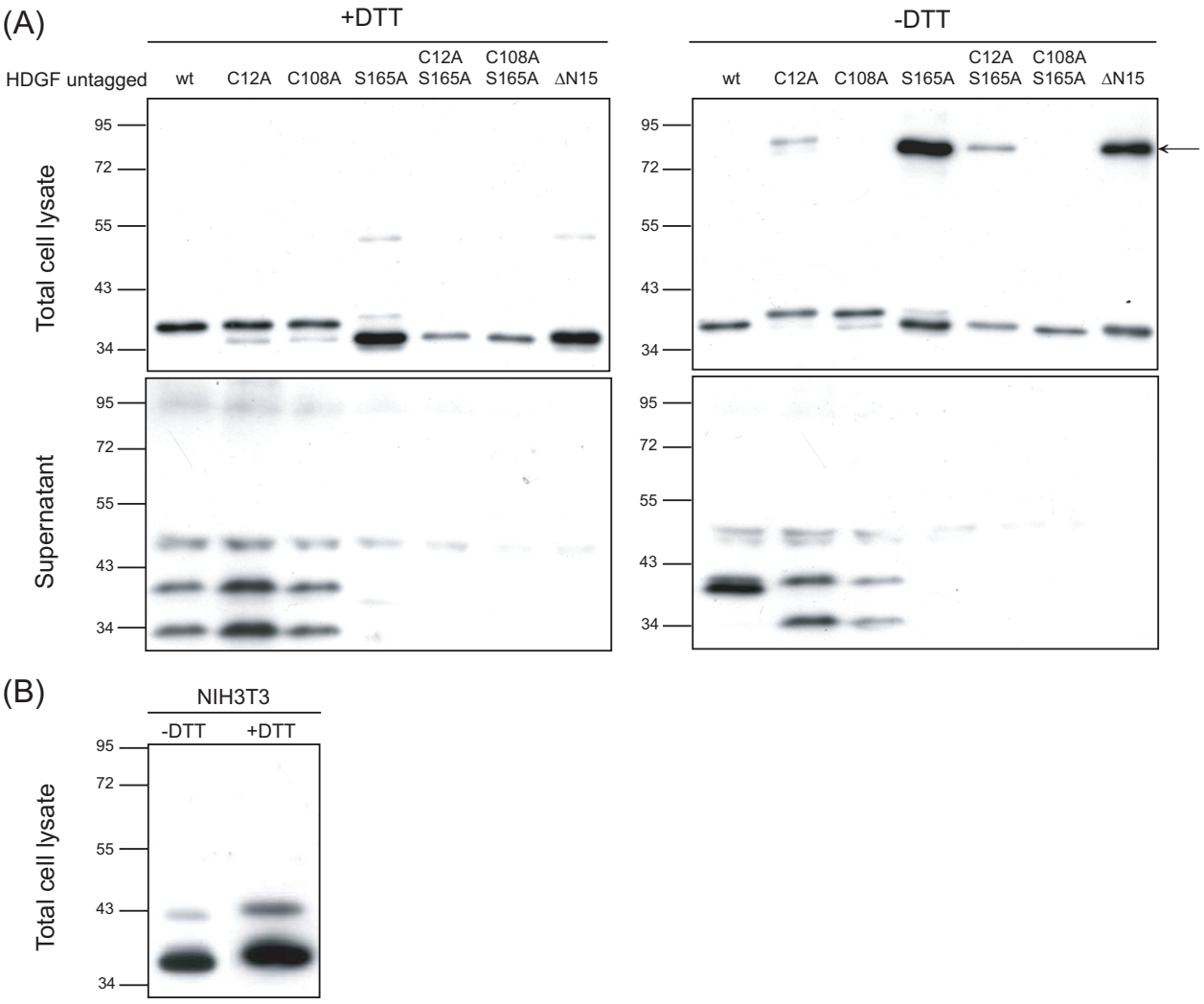


Figure 5 Disulfide bridges in HDGF.

(A) HEK 293 cells transfected with plasmids coding HDGF wt, C12A, C108A, C12A/S165A, C108A/S165A or ΔN15 mutants were lysed; cleared lysate was subjected to immunoprecipitation, as described under Materials and Methods, and analysed (under reducing and non-reducing conditions) along with secreted (supernatant) proteins by SDS-PAGE and Western blotting with HDGF specific antibody. Arrows point at the HDGF dimer band under non-reducing conditions (B) HDGF expression in NIH 3T3 cells was analysed (under reducing and non-reducing conditions) by SDS-PAGE and Western blotting with HDGF specific antibody.

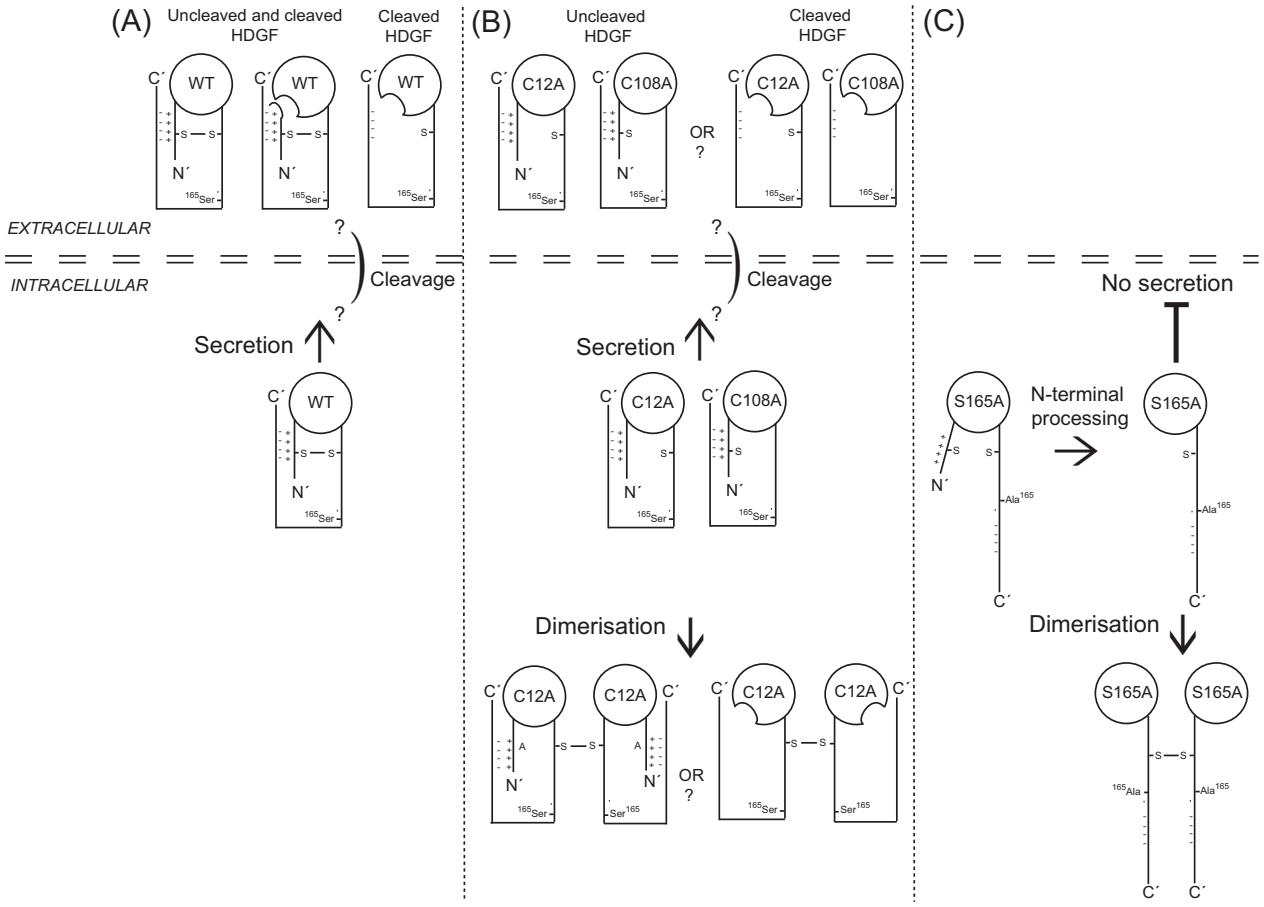


Figure 6 Proposed model for HDGF secretion and dimerisation. (A) HDGF wt; (B) HDGF C12A and C108A mutants; and (C) HDGF S165A mutant.

3.3 Publication 3

Interaction of HRP-2 splice variants with Hepatoma-derived growth factor

K Thakar, D Kelkar, T Shidore, S Gupta, S Kelm, F Dietz

2010

Biological Chemistry
(Submitted; MS no. BIOLCHEM-10-180)

BIOLOGICAL CHEMISTRY

Interaction of HRP-2 splice variants with Hepatoma-derived growth factor

Ketan Thakar , Dipti Kelkar, Teja Shidore, Shivangi Gupta, Sørge Kelm, Frank Dietz [1]

Centre for Biomolecular Interactions Bremen (CBIB), Department of Biochemistry, University of Bremen, Leobener Str. NW2, 28359 Bremen, Germany

[1]To whom correspondence should be addressed:
Centre for Biomolecular Interactions Bremen (CBIB)
Department of Biochemistry, University of Bremen
Leobener Str. im NW2, 28359 Bremen, Germany,
Phone ++4942121863223
FAX ++4942121863221
E-mail: fdietz@uni-bremen.de

12 text pages (including references), 6 Figure pages;
Total word count: 4878; Abstract: 149; Total no. of characters (no spaces): 27,096

Abstract

Hepatoma-derived growth factor related protein 2 (HRP-2) belongs to a family with five additional members: hepatoma-derived growth factor (HDGF); lens epithelium derived growth factor (LEDGF); and the HDGF related proteins 1, 3 and 4. Especially for HRP-2 very little is known regarding its function. Here we demonstrate that HRP-2 interacts with HDGF. Furthermore, we identified a new splice variant of HRP-2 having a 53 amino acid deletion in the N-terminal region of the protein. This deletion outside of the PWWP domain causes a loss of an α-helix within the *hath* region. Moreover, the specific interaction of the new HRP-2 splice variant with a processed HDGF form is attributed to the loss of this α-helix, suggesting a critical role of this element. These observations provide evidence for a previously unknown interplay between different members of the HRP family and open a new perspective for a better understanding of HRP functions.

Key words

Alternative splicing; *hath* (homologous to the amino terminus of HDGF) region; N-terminal processing; phosphorylation; PWWP domain.

Introduction

Hepatoma-derived growth factor related proteins (HRPs) comprise a family of proteins shown to be mitogenic to various cell lines. Besides the prototype of the family, hepatoma-derived growth factor (HDGF) [Nakamura et al., 1989], five HRPs have been identified so far; HRP-1 [Izumoto et al., 1997], HRP-2 [Izumoto et al., 1997], HRP-3 [Ikegame et al., 1999; Abouzied et al., 2004], HRP-4 [Dietz et al., 2002], and LEDGF (lens epithelium-derived growth factor) [Singh et al., 2000a]. All HRPs share the following features: (i) high degree of sequence identity in the first N-terminal 98 amino acid residues, abbreviated as *hath (*homologous to the amino terminus of HDGF) region; (ii) a PWWP domain in the *hath* region; (iii) a canonical bipartite nuclear localization signal in the C-terminal region; and (iv) lack of a hydrophobic signal peptide.

HDGF was originally purified as a heparin binding growth factor stimulating growth of a variety of cell types [Nakamura et al., 1994; Oliver and Al-Awqati, 1998; Everett et al., 2000; Mori et al., 2004; Lepourcelet et al., 2005]. Apart from HDGF, mitogenic activity has been confirmed for HRP-4 in fibroblasts [Dietz et al., 2002]; and LEDGF in lens epithelial cells, keratinocytes, and fibroblasts [Singh et al., 2000b]. Only HRP-2 and HDGF are ubiquitously expressed, whereas HRP-1/4 and HRP-3 are restricted to testis and brain, respectively [Nakamura et al., 1994; Izumoto et al., 1997; Ikegame et al., 1999; Dietz et al., 2002].

HRP-2 is the largest HRP containing 669 amino acids with a high level of basic and acidic residues, a typical feature of nuclear proteins, and a mixed charge cluster found in eukaryotic regulatory proteins, including transcription and replication factors [Izumoto et al., 1997]. Another feature of HRP-2 is the presence of an integrase binding domain (IBD), which has a high degree of similarity to the IBD of LEDGF. For both proteins it was shown that they bind to and enhance the activity of HIV integrase, however unlike LEDGF/p75, HRP-2 lacks the ability to tether HIV-1 IN to chromatin [Cherepanov et al., 2004]. Alternatively spliced gene products (p75 and p52) have been shown for LEDGF [Singh et al., 2000a]. Analysis of the murine genome revealed that the HRP-2 gene encompasses at least 16 exons. Furthermore, mRNA entries for 2 isoforms, one with an insertion of 9 amino acids, suggest that also HRP-2 pre-mRNA is alternatively spliced.

The PWWP domain of HRPs comprises a conserved 70-amino acid sequence found in ~60 eukaryotic proteins. Mainly roles in interactions with other proteins [Stec et al., 2000] or with DNA [Qiu et al., 2002; Chen et al., 2004; Ge et al., 2004] have been attributed to the PWWP domain. Interestingly, the PWWP domain of HDGF binds not only to DNA [Yang and Everett, 2007] but also the heparin-binding activity of HDGF [Nakamura et al., 1989] has been assigned to its PWWP domain [Sue et al., 2007]. A high-resolution structure of the HDGF *hath* region determined by NMR revealed the core structure

consisting of α-helical elements and a β-barrel with five anti-parallel β-strands [Lukasik et al., 2006]. HDGF forms homodimer by an unusual mechanism exchanging β-strands 1 and 2 between two HDGF PWWP domains [Sue et al., 2007]. This dimerisation leads to an increased binding affinity towards heparin.

In this study we addressed the question whether HDGF also interacts with other HRPs like HRP-2. We show that (1) HRP-2 binds to HDGF, that (2) a HRP-2 isoform with a 53 amino acid deletion in the *hath* region preferentially interact with a HDGF variant, and (3) this specificity can be attributed to the loss of an α-helix within the deleted region.

Results

Differential expression of HRP-2 correspond to alternatively spliced isoforms

Previous studies investigating HRP-2 expression in different brain regions have shown distinct bands with a specific anti HRP-2 polyclonal antibody [El-Tahir et al., 2006]. Extending this observation we analyzed the changes in HRP-2 expression during the development of rats (P= post natal days 2–30, adult) in different brain regions (Figure 1). The intensities of the four distinct bands at 100-130 kDa identified by the polyclonal antibody suggested area-specific changes in expression pattern of HRP-2 forms during the development of rat brain.

The heterogeneity of HRP-2 proteins detected by the antibody could be the result of either specific proteolytic processing or alternative splicing of the HRP-2 pre-mRNA. To further explore the hypothesis of alternatively spliced HRP-2 variants, we amplified total HRP-2 mRNA by RT-PCR using mouse adult brain cDNA as a template with sense and antisense primers flanking the coding region. By sequencing the PCR products we discovered a new splice variant of HRP-2, besides the already described HRP-2 coding sequences referred to as HRP-2 *isoform a* in this study (669 aa, accession number BAA22896), and a variant referred to as HRP-2 *isoform b* (678 aa, accession number BC003741). Figure 2 shows a schematic representation of the known intron and exon borders and resulting alternatively spliced HRP-2 gene products. The new splice variant is characterized by an in-frame deletion spanning parts of exons 3 and 4. During this study we refer this new isoform as HRP-2 *isoform c* (616 aa, accession number FN687734).

In HRP-2 *isoform c* the deletion of 53 amino acids (aa 66–118) leaves the predicted β-barrel intact, but skips the remaining part of the conserved *hath* region including a regular α-helix which seems to be a characteristic feature of the PWWP domain of all HRPs according to the solved structure of the *hath* region of HDGF (PDB ID: 1RI0) [Sue et al., 2004] and HRP-3 (PDB ID: 1N27) [Nameki et al., 2005]

All HRP-2 isoforms were expressed in COS-7 cells and analyzed by Western blotting (Figure 3). While the HRP-2 isoforms *a* and/or *b* have an electrophoretic mobility similar to the bands just below 130 kDa found in tissue samples (frontal lobe P5 in Figure 3) *isoform c* migrates similar to the lower bands.

Interaction of HRP-2 with HDGF

HDGF has been shown to form homo- and hetero-dimers [Sue et al., 2007; Yang and Everett, 2007]. Therefore, we have tested the hypothesis that different HRPs interact with each other. To investigate the interaction of HRP-2 isoforms with HDGF, HRP-2 *isoforms a*, *b* or *c* and HDGF wt were co-expressed in COS-7 cells. For affinity precipitation of the interacting HRPs with *Strep*Tactin® beads, either HRP-2 (Figure 4A) or HDGF wt (Figure 4B) contained a *Strep*-tag at its C-terminus. The Western blot results of these precipitates clearly indicate interaction of HDGF with all HRP-2 isoforms. Interestingly, only HRP-2 *isoform c* interacts additionally with a lower molecular mass form of HDGF which is almost not detectable in the total cell lysate fraction.

Immunocytochemical analysis revealed a predominant nuclear localization of HDGF and HRP-2 *isoform a*, *b* and *c*, providing the spatial requirement for the interaction between these proteins in the cell (Figure 5).

In the context of a study addressing the secretion mechanism of HDGF (see accompanying manuscript), we generated eight HDGF variants in which potential serine phosphorylation sites had been mutated to alanine and expressed these in HEK 293 cells. One of these mutants, HDGF S165A, is truncated at its N-terminus leading to a higher electrophoretic mobility similar to the lower molecular mass band co-precipitating with HRP-2 *isoform c*.

The 53 amino acid deletion (aa 66-118) in HRP-2 *isoform c* includes a 11-residue long α-helical element. To test whether this modification is the structural basis for the altered binding specificity, we prepared a deletion mutant of HRP-2 *isoform a* lacking only these amino acids stretch (HRP-2 *isoform* $a^{\Delta helix}$). Co-precipitation studies were performed as described above. The data shown in Figure 6 confirms that all HRP-2 isoforms and the Δhelix mutant interact with both, HDGF wt and HDGF S165A mutant. Furthermore, like HRP-2 *isoform c*, the HRP-2 *isoform* $a^{\Delta helix}$ mutant also co-precipitates the lower molecular mass form of HDGF. Although no precise quantification is possible from these experiments, it appears that HDGF S165A is co-precipitated by HRP-2 *isoform c* and the *isoform* $a^{\Delta helix}$ mutant at a much higher rate than HDGF wt.

Discussion

In this study we demonstrate that HRP-2 and HDGF bind to each other and explored the specificity of this reaction as a prototype for interactions between HRPs. HDGF acts as a transcriptional co-repressor for SMYD1 gene expression by recruiting CtBP [Yang and Everett, 2009]. It is feasible that HRP-2/HDGF hetero-dimerisation modulates DNA binding thus regulating gene expression. One important result of this study has been the observation that HRP-2 *isoform c* selectively binds to a subset of HDGF molecules with a lower molecular mass. This opens the possibility that interactions between HRPs, and consequently their functions, are controlled by differential expression of such isoforms. Here we have shown that the expression pattern of HRP-2 variants is changing during development, extending a previous investigation describing the regional differences of HRPs expression in rat brain [El-Tahir et al., 2006]. Obviously, the observed band pattern can have different origins, such as proteolysis, post-translational modification or the occurrence of different splice variants. All these events can be regulated during developmental processes.

We further provide strong evidence that at least three HRP-2 variants, *isoforms a* and *b*, and a so far unknown *isoform c*, are generated by alternative mRNA splicing. Alternative splicing is not unusual for HRPs. It has been described for LEDGF giving two isoforms p75 and p52 [Singh et al., 2000a] and splice variants of other members like HDGF or HRP-3 can be found in databases. However, the corresponding proteins have not been described.

The new *isoform c* is a 616 aa protein with a deletion in its N-terminal region extending from amino acids 66 to 118. Studies on HDGF have highlighted the importance of this region in regulating or defining HDGF function. For example, amino acids 81-100 of HDGF appear to be relevant for its cell surface interaction and proliferative function [Abouzied et al., 2005] and SUMOylation of K80 has been shown to affect HDGF binding to chromatin [Thakar et al., 2008]. Recombinant HRP-2 *isoforms a* and *b* migrate like proteins with a molecular mass of 120-130 kDa, whereas HRP-2 *isoform c* migrates like a protein of 100-110 kDa correlating with the lower molecular weight bands from the tissue samples. For HRPs it is difficult to directly correlate the predicted molecular mass of the splice variants with bands in SDS-PAGE. As already shown for HDGF and HRP-3 [Abouzied et al., 2004], HRP-2 is migrating at a molecular weight significantly higher than predicted (74 kDa) by its primary amino acid sequence. Since also proteins expressed in bacteria (data not shown) show the same unusual low electrophoretic mobility, it is unlikely to result from post-translational modifications.

The occurrence of various isoforms suggests specialized and specific functions of these proteins related to their structural characteristics. Sequence alignments and structural comparison between different HRPs suggest that the *hath* region is folded similarly in all these proteins, consisting of a five stranded

anti-parallel β-barrel and α-helices [Sue et al., 2004]. The *hath* region of HDGF can form dimers mediated by an unusual exchange of between the two monomers [Sue et al., 2007]. It is tempting to speculate that a similar mechanism mediates the interaction between HRP-2 and HDGF described in this study. In this context it is important to note that the specificity of HRP-2 *isoform c* for interaction with HDGF variants appears to be due to the loss of 53 amino acids including the α-helix named B [Sue et al., 2004] or α_4 [Lukasik et al., 2006]) which is in direct contact with the β-strands 1 and 2 involved in HDGF dimerisation [Sue et al., 2007]. Along this line, an explanation for the binding preferences of HRP-2 *isoform c* could be that the exchange of β-strands 1 and 2 is facilitated, as these are exposed due to the lack of the α-helix covering them in *isoform a* or *b*. This hypothesis is supported by our data, since HRP-2 *isoform a*$^{\Delta helix}$ lacking only this α-helix has the same binding preferences as *isoform c*.

Moreover, the apparently preferred HDGF binding partner of HRP-2 *isoform c* is a variant with an electrophoretic mobility similar to the HDGF S165A mutant which is truncated at the N-terminus. Interestingly, this truncation leads to the loss of the intramolecular disulfide bridge between C12 and C108 (see accompanying manuscript) which may stabilize a more compact conformation of the *hath* region preventing the exchange of β-strands 1 and 2 in HDGF wt.

In summary, this study shows for the first time heteromerisation between different members of the HRP family; HDGF and HRP-2. Additionally, we discovered a new splice variant of HRP-2 with a deletion in its *hath* region which interacts preferentially with a processed form of HDGF probably due to the loss of an α-helix in HRP-2. In conclusion, these observations provide a new perspective for understanding the biological functions of HDGF and related proteins.

Materials and Methods

RNA isolation and cDNA preparation from brain tissue sample

Total RNA was isolated by homogenization of brain tissue samples according to manufacture's protocol (peqGOLD TriFast kit, PEQLAB, Erlangen, Germany). Briefly; 1 mL TriFast for every 50-100 mg of tissue was used. The nucleoprotein complexes were incubated at room temperature for 5 min followed by addition of 0.2 mL chloroform and further incubation for 10 min at room temperature. The samples were then centrifuged at 12000 *g* to separate phases. RNA containing aqueous phase was then separated and precipitated using 500 µL of isopropanol mL^{-1} of TriFast. These samples were further incubated on ice for 15 min and centrifuged for 10 min at 12000 *g*, 4°C. RNA pellet obtained was subsequently washed with 75% ethanol, centrifuged for 10 min at 7500 *g*, 4°C and re-suspended in RNA free water. cDNA synthesis was carried out according to manufactures protocol (RevertAid™ First Strand cDNA Synthesis Kit, Fermentas, St. Leon-Rot, Germany).

Production of recombinant proteins

Plasmids were constructed using standard recombinant cloning techniques and all changes were verified by DNA sequencing. HRP-2 was amplified from mouse brain cDNA using the following nested primers:
Sense 5' CCTCCGCCGACAGCATGCCG and
antisense 5' GGCTGGAGCTCAGCTGTCCT followed by
Sense 5´CGAAGCTTATGCCGCACGCCTTCAAGCCC and
antisense 5´GCGAATTCGCTGTCCTCATTGTCAT primer
containing HindIII and EcoRI restriction sites and cloned in frame to the *Strep*Tag coding sequence into HindIII/EcoRI digested pcDNA3 Amp *Strep*Tag vector as described previously [Thakar et al., 2008].

Untagged HRP-2 isoforms were prepared by amplifying the coding sequence from the respective C-terminally *Strep*-tagged constructs using the same sense primer and antisense primer 5´GCGAATTCTCAGCTGTCCTCATTGTCAT
The PCR amplification product was cloned using the HindIII and EcoRI restriction sites in the pcDNA3 Amp vector.

The cloning strategy to generate C-terminally *Strep*-tagged HRP-2 *isoform* $a^{\Delta helix}$ was as follows: The cDNAs coding for the N-terminal (1 to 74) and the C-terminal region of HRP-2 *isoform a* (87 to 669) were amplified separately using C-terminally *Strep*-tagged HRP-2 *isoform a* as templates. The 3' and 5' ends of the amplification products were designed to allow an in frame ligation using EcoRI and subcloned in frame into HindIII/EcoRI or EcoRI/KpnI digested pcDNA3 Amp. The following primers were used for amplification:

Sense primer for N-terminal region 5' CCCAAGCTTATGCCGCACGCCTTCAAGCCC,
antisense primer for N-terminal region 5' CGGAATTCGCCTTTCCTCTTGTTGGG,
Sense primer for C-terminal region 5' CGGAATTCAACCCCCATGCCAGCTAC,
antisense primer for C-terminal region 5' CGGGTACCGCTGTCCTCATTGTCATC.
The plasmid encoding for HDGF S165A was described elsewhere (see accompanying manuscript).

Cell culture and transfections

For transient transfection COS-7 cells were maintained, transfected using the polyethylenimine ExGen 500 according to the manufacturer's protocol (Fermentas, St. Leon-Rot, Germany) and harvested after 24 h as described previously [Thakar et al., 2008].

Preparation of proteins extracts

Cells were washed twice with PBS before treatment with ice cold (1mL/10 cm plate) TNE lysis buffer containing 20mM Tris/Cl pH 7.4, 150 mM NaCl, 5mM EDTA, 1% NP-40 + protease inhibitors (complete mini, EDTA free from Roche, Mannheim, Germany) on ice. After 10 min on ice, cells were scraped from the dish, transferred to a 1.5 mL reaction tube and incubated on ice for additional 10 min. Cell debris was separated by centrifugation at 12000 *g*, 4ºC for 20 min. The cleared lysate was used directly for SDS-PAGE and Western blot or for the purification of the recombinant *Strep*-tagged HDGF using *Strep*Tactin® (IBA, Göttingen, Germany).

***Strep*Tactin® purification**

Cleared lysate (900 µL from each 10 cm ∅ cell culture dish) were incubated in a batch procedure with 30 µL of *Strep*Tactin® Sepharose beads on a rotor shaker for 6 h at 4°C. After incubation beads were centrifuged at 1200 *g*, 4ºC for 15 s and the supernatant was transferred to a new 1.5 mL reaction tube. Beads were washed 4 times with 1 mL of TNE lysis buffer. After the final washing step, bound proteins were eluted by incubating the beads at 95°C for 10 min in 80 µL 2x SDS-PAGE sample buffer. 20 µL of the eluate was used for SDS-PAGE.

Western Blot analysis

In principle, Western blot analysis was performed as described elsewhere (see accompanying manuscript). The following antibodies were used: rabbit anti-HRP-2 polyclonal (1:250) and sheep anti-HDGF polyclonal (1:1000) [Abouzied et al., 2004], sheep anti-*Strep*-tag polyclonal (1:1000; IBA, Göttingen, Germany), mouse anti-Actin monoclonal (1:2000; Dianova, Hamburg, Germany).

Immunofluorescence

For analysis of sub-cellular localization by immunofluorescence COS-7 cells (3 x 10^4) were plated on cover-slips in a 24 well plate were grown overnight. Post transfection cells were washed twice in phosphate buffered saline (PBS) solution and then fixed in 4% paraformaldehyde in PBS for 10 min. For permeabilization, cells were treated for 2 min with ice cold methanol on ice and afterwards washed twice with PBS/2% BSA before blocking unspecific binding sites with DMEM/10% FCS for 30 min. The cells were immunostained with polyclonal sheep anti-HDGF and rabbit anti-HRP-2 as primary antibodies for 90 min in PBS/2% BSA at room temperature. After washing twice with PBS/2% BSA, cells were treated for 90 min in PBS/2% BSA at room temperature in the dark with donkey anti-rabbit antibody-Cy3 conjugate and donkey anti-sheep antibody-Cy2 conjugate (Dianova, Hamburg, Germany). Cells were

washed in PBS/2% BSA and then treated with the nuclear stain DAPI (Roche, Mannheim, Germany) for 1 min at room temperature. After additional washing cells were embedded in immuno-mount (Thermo Shandon, Pittsburgh, PA, USA) and images were captured on the fluorescence microscope (Olympus, Hamburg, Germany).

Acknowledgments

Our thanks go to Dr. Sebastian Franken from the Department of Physiological Chemistry, University of Bonn, Germany for kindly providing the rabbit anti-HRP-2 polyclonal antibody. We also like to thank Mrs. Nazila Isakovic for excellent technical support. This work is supported by the Central Research Funding (ZFK, BFK NaWi), University of Bremen, and the Tönjes-Vagt-Stiftung, Project XXI.

References

Abouzied, M.M., Baader, S.L., Dietz, F., Kappler, J., Gieselmann, V., Franken, S. (2004). Expression patterns and different subcellular localization of the growth factors HDGF (hepatoma-derived growth factor) and HRP-3 (HDGF-related protein-3) suggest functions in addition to their mitogenic activity. Biochem J *378*, 169-176.

Abouzied, M.M., El-Tahir, H.M., Prenner, L., Haberlein, H., Gieselmann, V., Franken, S. (2005). Hepatoma-derived growth factor. Significance of amino acid residues 81-100 in cell surface interaction and proliferative activity. J Biol Chem *280*, 10945-10954.

Chen, T., Tsujimoto, N., Li, E. (2004). The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. Mol Cell Biol *24*, 9048-9058.

Cherepanov, P., Devroe, E., Silver, P.A., Engelman, A. (2004). Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. J Biol Chem *279*, 48883-48892.

Dietz, F., Franken, S., Yoshida, K., Nakamura, H., Kappler, J., Gieselmann, V. (2002). The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. Biochem J *366*, 491-500.

El-Tahir, H.M., Dietz, F., Dringen, R., Schwabe, K., Strenge, K., Kelm, S., Abouzied, M.M., Gieselmann, V., Franken, S. (2006). Expression of hepatoma-derived growth factor family members in the adult central nervous system. BMC Neurosci *7*, 6.

Everett, A.D., Lobe, D.R., Matsumura, M.E., Nakamura, H., McNamara, C.A. (2000). Hepatoma-derived growth factor stimulates smooth muscle cell growth and is expressed in vascular development. J Clin Invest *105*, 567-575.

Ge, Y.Z., Pu, M.T., Gowher, H., Wu, H.P., Ding, J.P., Jeltsch, A., Xu, G.L. (2004). Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. J Biol Chem *279*, 25447-25454.

Huang, H.D., Lee, T.Y., Tzeng, S.W., Horng, J.T. (2005). KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites. Nucleic Acids Res *33*, W226-229.

Ikegame, K., Yamamoto, M., Kishima, Y., Enomoto, H., Yoshida, K., Suemura, M., Kishimoto, T., Nakamura, H. (1999). A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus. Biochem Biophys Res Commun *266*, 81-87.

Izumoto, Y., Kuroda, T., Harada, H., Kishimoto, T., Nakamura, H. (1997). Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. Biochem Biophys Res Commun *238*, 26-32.

Lepourcelet, M., Tou, L., Cai, L., Sawada, J., Lazar, A.J., Glickman, J.N., Williamson, J.A., Everett, A.D., Redston, M., Fox, E.A., Nakatani, Y., Shivdasani, R.A. (2005). Insights into developmental mechanisms and cancers in the mammalian intestine derived from serial analysis of gene expression and study of the hepatoma-derived growth factor (HDGF). Development *132*, 415-427.

Lukasik, S.M., Cierpicki, T., Borloz, M., Grembecka, J., Everett, A., Bushweller, J.H. (2006). High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. Protein Sci *15*, 314-323.

Mense, M., Vergani, P., White, D.M., Altberg, G., Nairn, A.C., Gadsby, D.C. (2006). In vivo phosphorylation of CFTR promotes formation of a nucleotide-binding domain heterodimer. EMBO J *25*, 4728-4739.

Mori, M., Morishita, H., Nakamura, H., Matsuoka, H., Yoshida, K., Kishima, Y., Zhou, Z., Kida, H., Funakoshi, T., Goya, S., Yoshida, M., Kumagai, T., Tachibana, I., Yamamoto, Y., Kawase, I., Hayashi, S. (2004). Hepatoma-derived growth factor is involved in lung remodeling by stimulating epithelial growth. Am J Respir Cell Mol Biol *30*, 459-469.

Nakamura, H., Izumoto, Y., Kambe, H., Kuroda, T., Mori, T., Kawamura, K., Yamamoto, H., Kishimoto, T. (1994). Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein. J Biol Chem *269*, 25143-25149.

Nakamura, H., Kambe, H., Egawa, T., Kimura, Y., Ito, H., Hayashi, E., Yamamoto, H., Sato, J., Kishimoto, S. (1989). Partial purification and characterization of human hepatoma-derived growth factor. Clin Chim Acta *183*, 273-284.

Nameki, N., Tochio, N., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Fujikura, Y., Saito, M., Ikari, M., Watanabe, M., Terada, T., Shirouzu, M., Yoshida, M., Hirota, H., Tanaka, A., Hayashizaki, Y., Guntert, P., Kigawa, T., Yokoyama, S. (2005). Solution structure of the PWWP domain of the hepatoma-derived growth factor family. Protein Sci *14*, 756-764.

Oliver, J.A., Al-Awqati, Q. (1998). An endothelial growth factor involved in rat renal development. J Clin Invest *102*, 1208-1219.

Qiu, C., Sawada, K., Zhang, X., Cheng, X. (2002). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat Struct Biol *9*, 217-224.

Rubin, S.M., Gall, A.L., Zheng, N., Pavletich, N.P. (2005). Structure of the Rb C-terminal domain bound to E2F1-DP1: a mechanism for phosphorylation-induced E2F release. Cell *123*, 1093-1106.

Singh, D.P., Kimura, A., Chylack, L.T., Jr., Shinohara, T. (2000a). Lens epithelium-derived growth factor (LEDGF/p75) and p52 are derived from a single gene by alternative splicing. Gene *242*, 265-273.

Singh, D.P., Ohguro, N., Kikuchi, T., Sueno, T., Reddy, V.N., Yuge, K., Chylack, L.T., Jr., Shinohara, T. (2000b). Lens epithelium-derived growth factor: effects on growth and survival of lens epithelial cells, keratinocytes, and fibroblasts. Biochem Biophys Res Commun *267*, 373-381.

Stec, I., Nagl, S.B., van Ommen, G.J., den Dunnen, J.T. (2000). The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? FEBS Lett *473*, 1-5.

Sue, S.C., Chen, J.Y., Lee, S.C., Wu, W.G., Huang, T.H. (2004). Solution structure and heparin interaction of human hepatoma-derived growth factor. J Mol Biol *343*, 1365-1377.

Sue, S.C., Lee, W.T., Tien, S.C., Lee, S.C., Yu, J.G., Wu, W.J., Wu, W.G., Huang, T.H. (2007). PWWP module of human hepatoma-derived growth factor forms a domain-swapped dimer with much higher affinity for heparin. J Mol Biol *367*, 456-472.

Thakar, K., Niedenthal, R., Okaz, E., Franken, S., Jakobs, A., Gupta, S., Kelm, S., Dietz, F. (2008). SUMOylation of the hepatoma-derived growth factor negatively influences its binding to chromatin. FEBS J *275*, 1411-1426.

Yang, J., Everett, A.D. (2007). Hepatoma-derived growth factor binds DNA through the N-terminal PWWP domain. BMC Mol Biol *8*, 101.

Yang, J., Everett, A.D. (2009). Hepatoma-derived growth factor represses SET and MYND domain containing 1 gene expression through interaction with C-terminal binding protein. J Mol Biol *386*, 938-950.

Figures

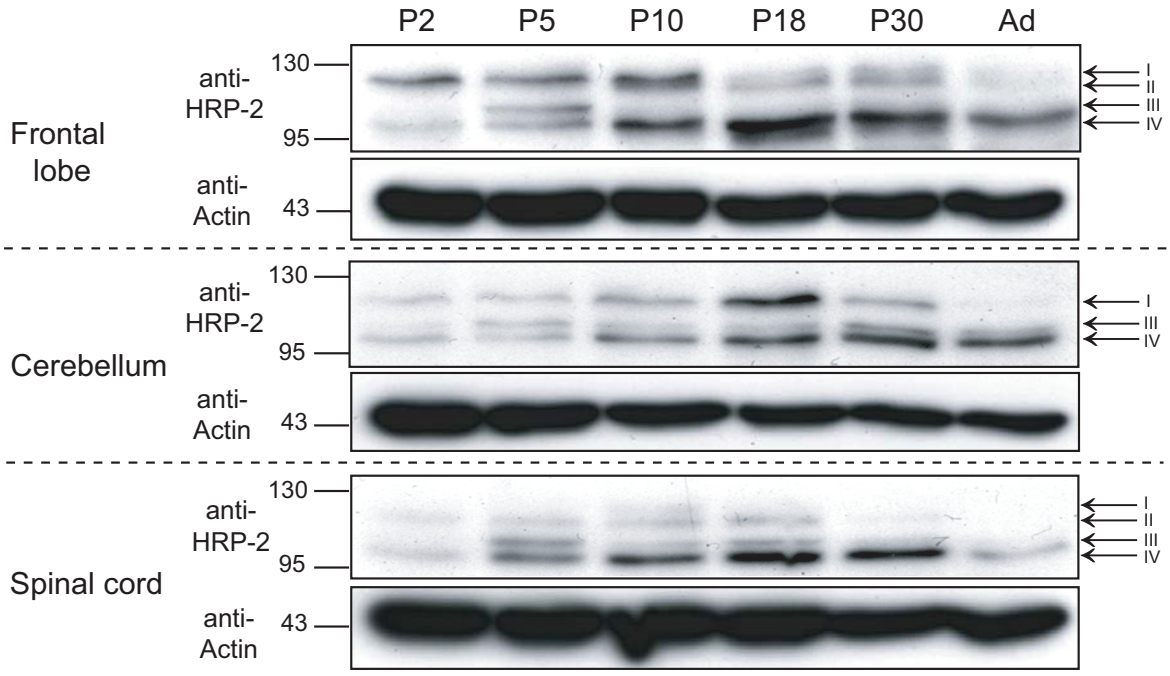


Figure 1 Expression of HRP-2 in different regions during brain development.

Equal amounts (20 µg) of protein extracts from tissue homogenates (P-postnatal, Ad-adult from frontal lobe, cerebellum, and spinal chord) were subjected to Western blot analysis using a rabbit anti-HRP-2 polyclonal antibody. Numbers I, II, III & IV represent different HRP-2 forms. Bars on the left indicate positions of the molecular mass marker (in kDa). Experimental details are described under „Material and Methods“.

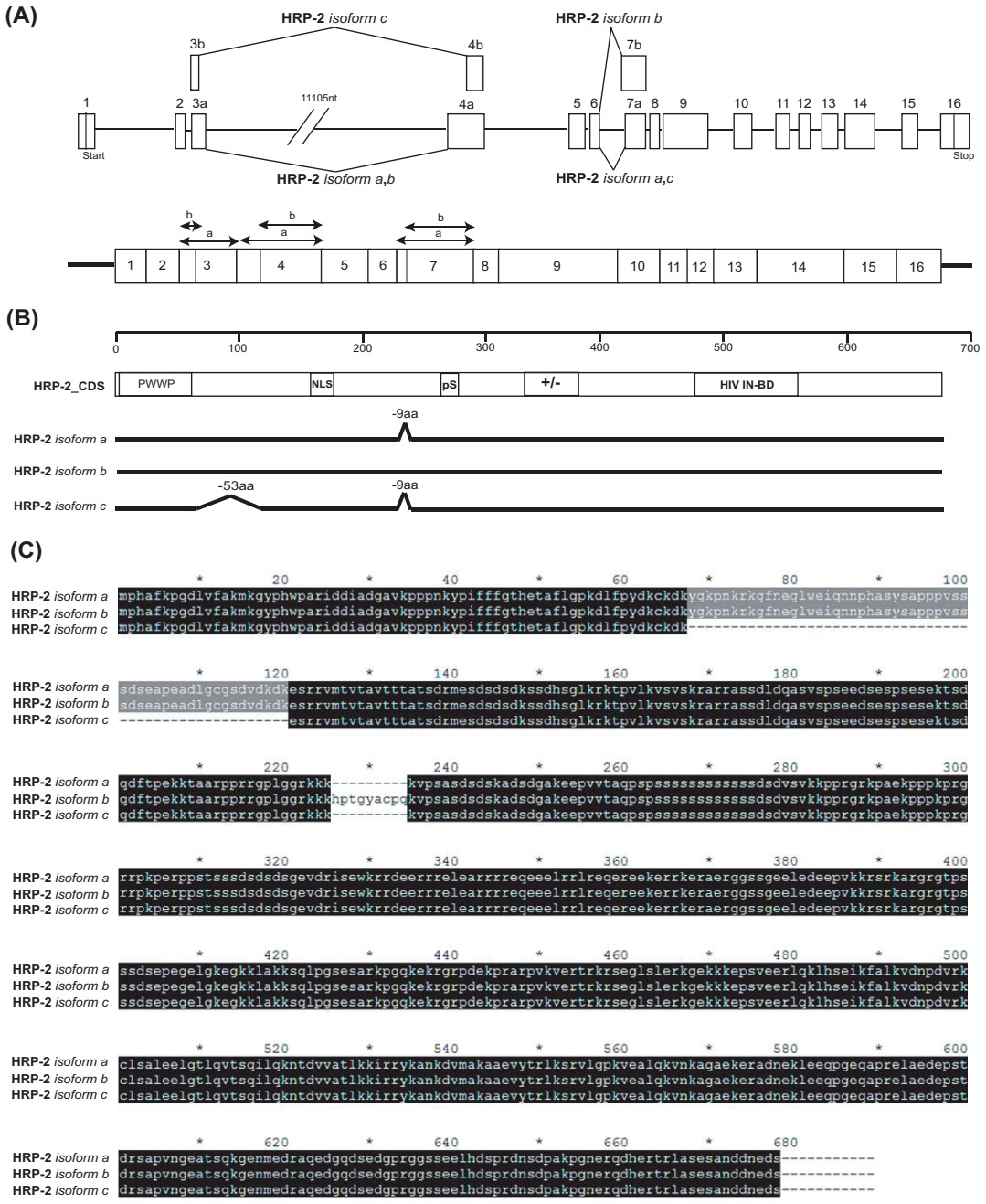


Figure 2 Schematic representations of HRP-2 isoforms.

(A) Intron-exon organization in the HRP-2 gene and generation of alternatively spliced isoforms. (B) Comparison of the HRP-2 isoforms denoting the spliced out regions. The locations of the PWWP domain (PWWP), the nuclear localization signal (NLS), poly-serine stretch (pS), mixed charged cluster (+/-) and HIV integrase binding domain (HIV-IN) are indicated. (C) Amino acid alignment of the HRP-2 isoforms; amino acids identical in all isoforms are on black background, amino acids identical only in HRP-2 *isoform a* and *isoform b* are on grey background, dashes indicate amino acids missing in the HRP-2 *isoform a* and *isoform c* compared with *isoform b*.

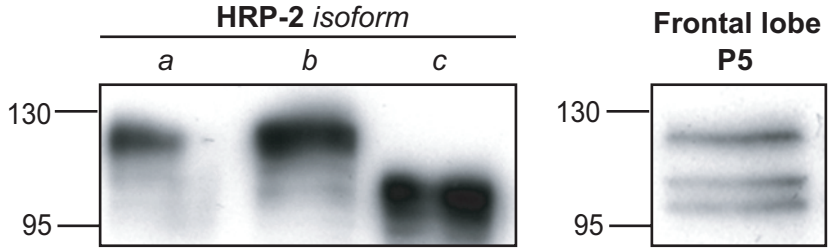


Figure 3 Expression of recombinant HRP-2 isoforms.

COS-7 cells were transfected with plasmids coding for HRP-2 *Strep*-tag *isoform a*, *b* or *c*. Cell lysates were analyzed by SDS-PAGE and Western blotting with anti-HRP-2 antibodies. The band pattern of HRP-2 found in tissue extract from frontal lobe P5 is shown for comparison on the right side.

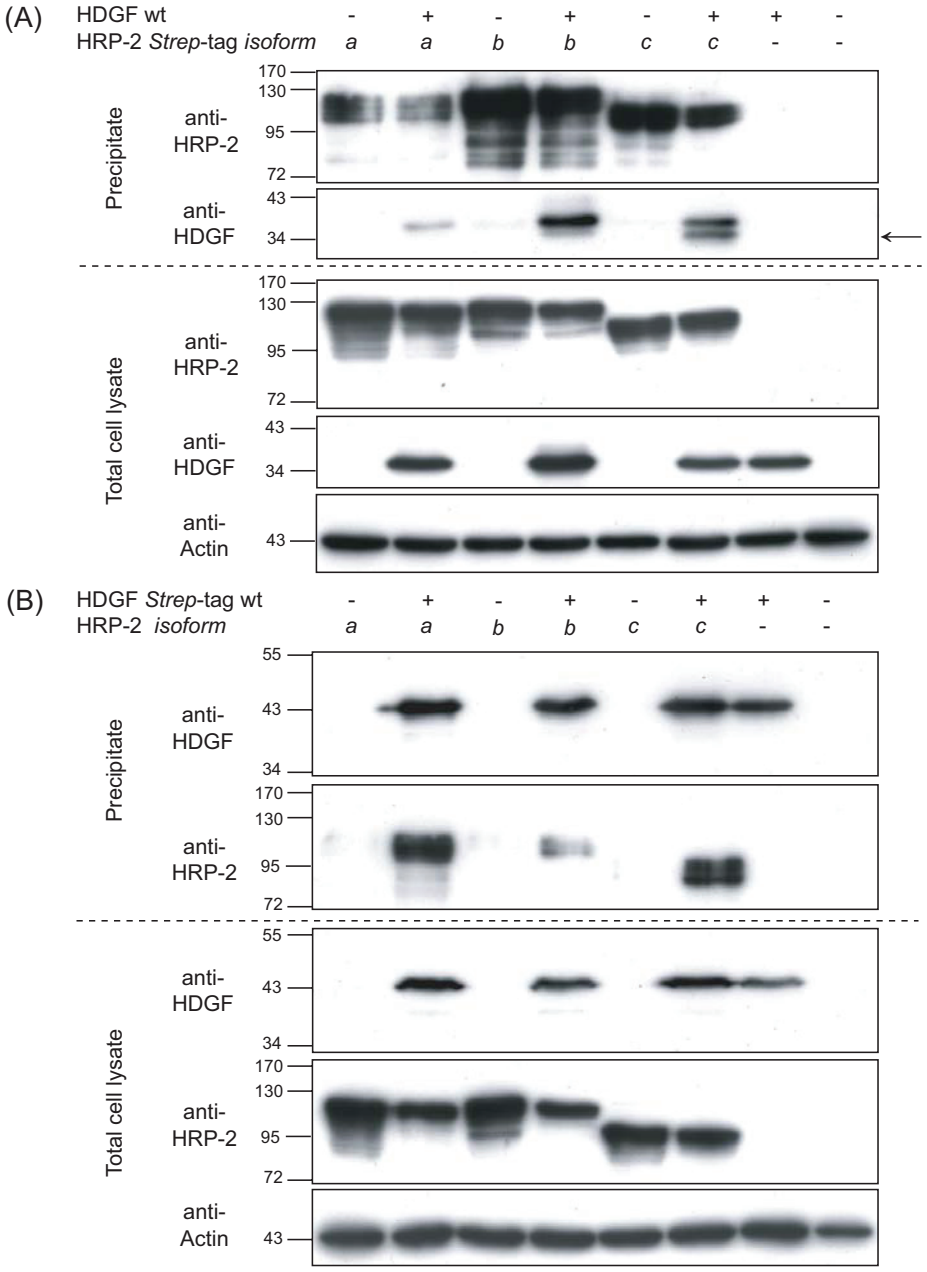


Figure 4 Interaction of HRP-2 with HDGF

Cell lysates of transfected COS-7 cells were prepared and affinity precipitations with *Strep*Tactin® beads were performed. Preciptates and total cell lysates were analyzed by Western blotting detecting HRP-2 *isoform a*, *b* or *c* with anti-HRP-2, HDGF wt with anti-HDGF and actin with anti-actin antibodies as indicated. Experimental details are described under “Material and Methods“.

(A) COS-7 cells were transfected with plasmids coding for HRP-2 *Strep*-tag *isoform a*, *b* or *c*, alone or together with plasmid coding for HDGF wt as indicated. The arrow points to a lower molecular mass form of HDGF precipitated preferentially by HRP-2 *isoform c*. (B) COS-7 cells were transfected with plasmids coding for HRP-2 *isoform a*, *b* or *c*, alone or together with plasmid coding for HDGF wt *Strep*-tag.

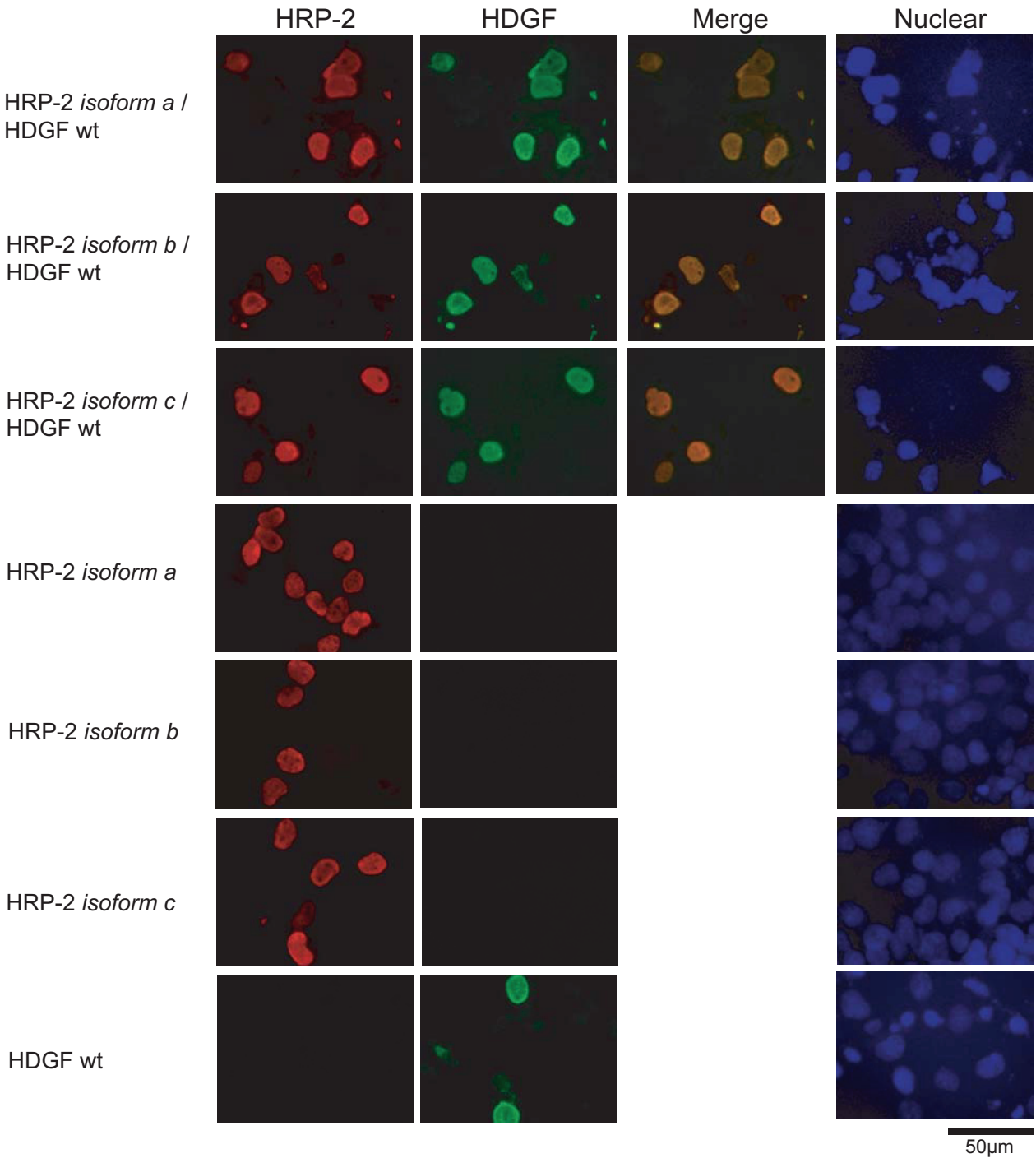


Figure 5 Nuclear localization of HRP-2 and HDGF.

To investigate the subcellular localization of HRP-2 isoforms and HDGF, COS-7 cells were transfected with plasmids coding for HRP-2 *Strep*-tag *isoform a*, *b* or *c* alone or together with untagged HDGF wt as indicated. Paraformaldehyde fixed cells were methanol permeabilized and stained using polyclonal rabbit anti-HRP-2 and sheep anti-HDGF primary antibodies followed by donkey anti-rabbit Cy3-conjugated and donkey anti-sheep Cy2-conjugated secondary antibodies. Immunostaining for HRP-2 and HDGF was analyzed by fluorescence microscopy. DAPI was used for nuclear staining. Scale bar = 50 µm.

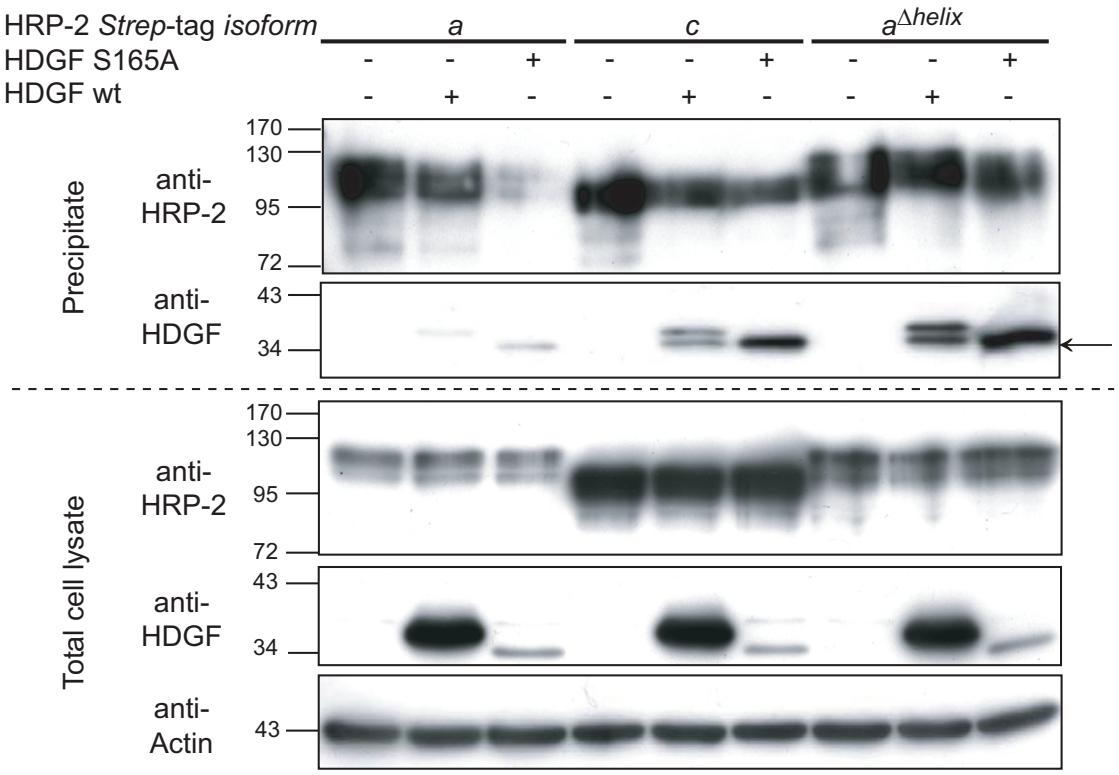


Figure 6 Interaction of HDGF wt and HDGF S165A with HRP-2 *isoform a*, *isoform c* and *isoform a*$^{\Delta helix}$ mutant.

Cell lysates of COS-7 cells transfected with the plasmids indicated were prepared and affinity precipitations with *Strep*Tactin® beads were performed. Precipitates and total cell lysates were analysed by Western blotting detecting HRP-2 *isoforms* with anti-HRP-2, HDGF variants with anti-HDGF and actin with anti-actin antibodies as indicated. Experimental details are described under “Material and Methods“. The arrow points at the lower molecular mass form of HDGF co-precipitated with *Strep*-tagged HRP-2 *isoform c* or *isoform a*$^{\Delta helix}$ mutant.

4

Additional Results

4.1 SUMOylation of HDGF by SUMO isoforms in mammalian cells

We have demonstrated that HDGF is modified by the attachment of SUMO-1 in mammalian cells and the lysine residue on which the modification takes place is not a part of the consensus motif described for SUMOylation [56]. Till date there are four 10-11 kDa SUMO isoforms identified in mammals. SUMO-1 exhibits 47% similarity at the protein level with SUMO-2 and -3, whereas SUMO-2 and -3 are 95% similar [57]. To investigate whether HDGF is also SUMOylated by SUMO-2 and SUMO-3 isoforms and whether the modification takes place on the same acceptor lysine as for SUMO-1 we expressed HDGF wt and K80R mutant in COS-7 cells alone or together with EGFP-SUMO-1, HA-SUMO-2 and HA-SUMO-3. We observed that HDGF wt is SUMOylated by all the three SUMO isoforms and a complete loss of the SUMOylated HDGF signal was observed for the mutant K80R suggesting that all the three SUMO isoforms use the same lysine residue for modification (Figure 4.1).

.

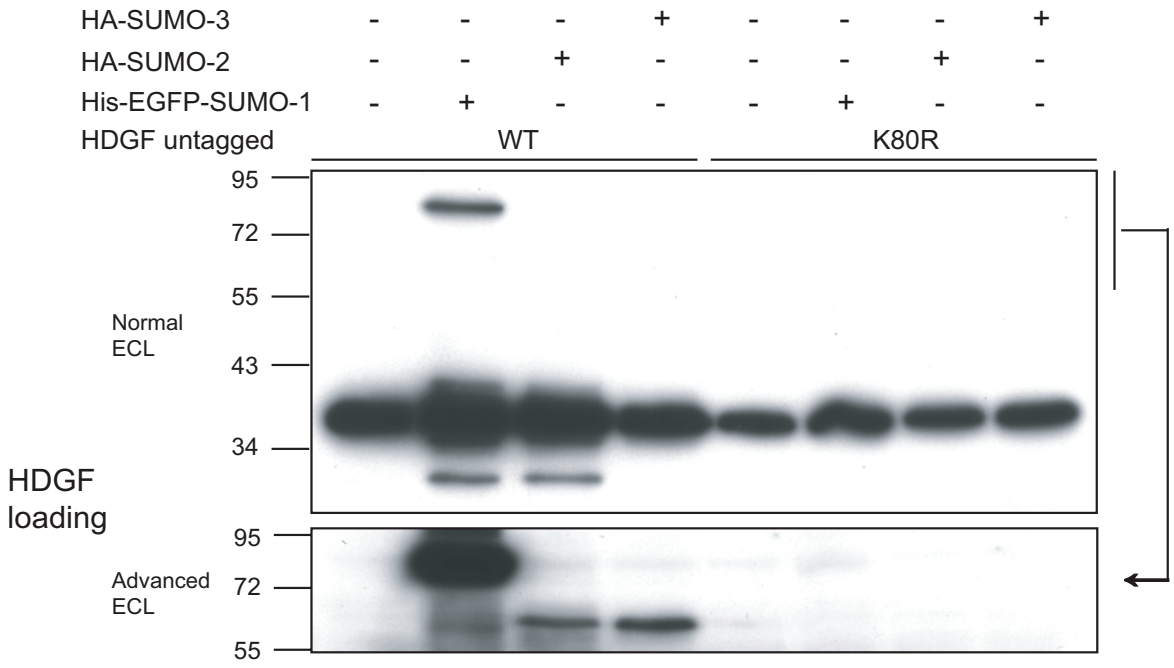


Figure 4.1 HDGF is modified by different SUMO isoforms in mammalian cells. COS-7 cells transfected with plasmids coding for the indicated proteins were lysed and analyzed by SDS-PAGE and Western blotting with HDGF specific antibody. SUMOylated and unSUMOylated HDGF could be detected in the total cell lysate of cells co-expressing HDGF wt with His-EGFP-SUMO-1, HA-SUMO-2 or HA-SUMO-3.

4.2 HDGF is processed C-terminally at a potential caspase cleavage site

In our previous studies with over expression of HDGF wt we especially observed in COS-7 cells a double band pattern in the total cell lysate using a HDGF specific antibody. The two bands observed differ by roughly 4 kDa. We were able to show that the lower molecular weight form is C-terminally processed as it can be also detected using a *Strep*Tag specific antibody using N-terminally *Strep*-tagged HDGF for over-expression. Moreover, the lower molecular weight form is the form which is preferentially secreted in COS-7 cells (Figure 4.2 (a)). These observations lead us to look for potential cleavage sites in the C-terminal region of HDGF. Analysis of the mouse HDGF amino acid sequence in CASVM (1.0) server for caspase cleavage site prediction gave us a predicted site ^{202}SEPD^{205} for HDGF. Interestingly, comparison of mouse and human HDGF amino acid sequences shows that the aspartate residue where the cleavage would take place in mouse HDGF is replaced by glycine in human HDGF (Figure 4.2 (b)).

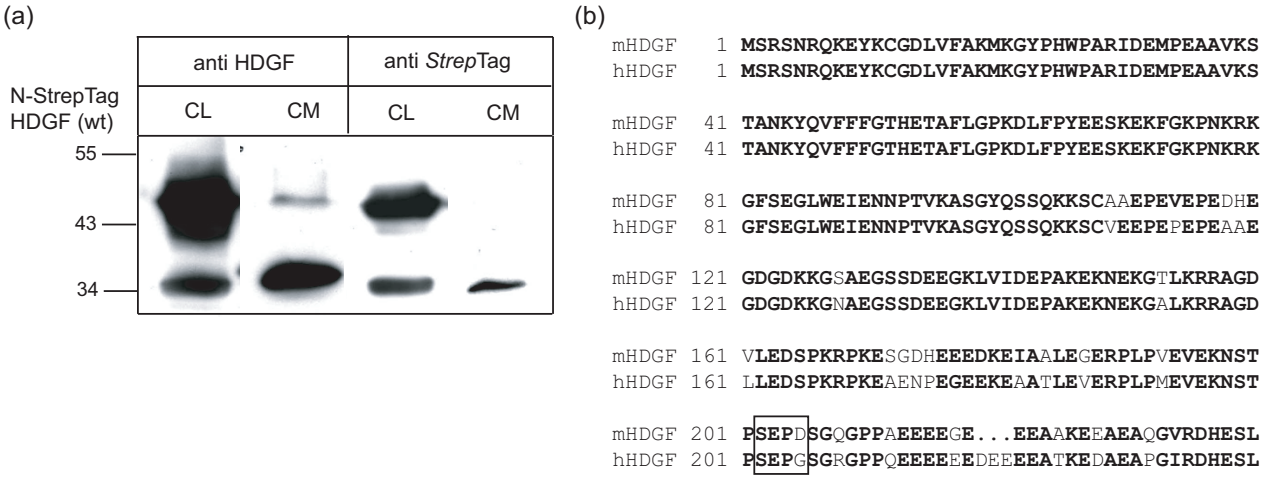


Figure 4.2 (a) Expression and secretion of N-*Strep*Tag HDGF in COS-7 cells. Cell lysate (CL) and media (CM) of COS-7 cells transiently transfected with plasmids coding for N-terminally *Strep*-tagged HDGF wt were detected after western blotting using a polyclonal sheep anti-HDGF antibody and polyclonal rabbit anti-*Strep*Tag antibody. (b) Amino acid alignment of mouse HDGF wt (NCBI Reference Sequence: NP_032257.3) and human HDGF wt (NCBI Reference Sequence: NP_004485.1) proteins. Homologous amino acids between the two proteins are represented in bold. The predicted caspase cleavage site is marked by rectangle

To investigate whether this site plays a role in HDGF C-terminal processing, we transfected plasmids encoding for mHDGF wt, mHDGF D205G, hHDGF wt and hHDGF G205D mutants in COS-7 cells and checked for expression and secretion. We could clearly demonstrate that this site is functional in mouse HDGF and can be mimicked by mutation in human HDGF (Figure 4.3).

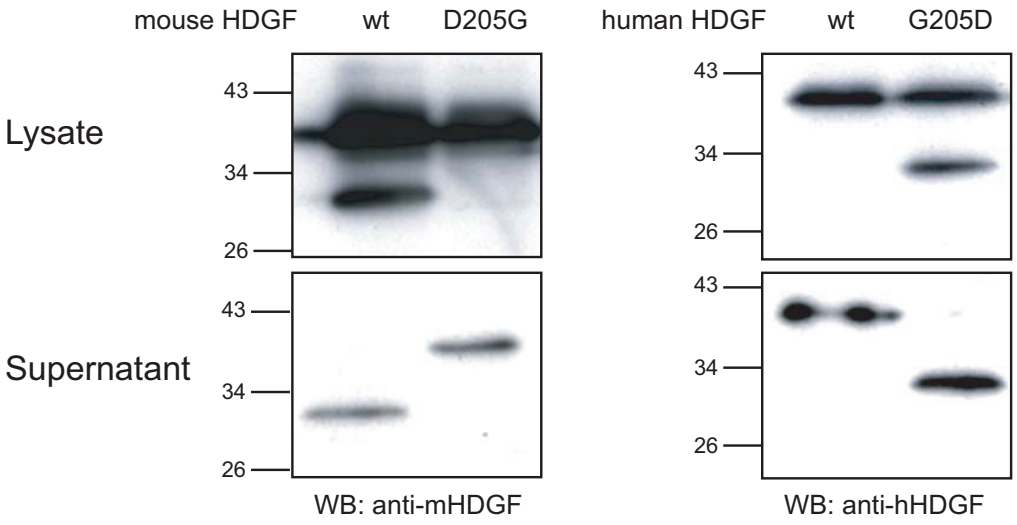


Figure 4.3 HDGF is C-terminally processed at a predicted caspase cleavage site. COS-7 cells transfected with plasmids coding for the indicated proteins. Cells were harvested; supernatants acetone precipitated; and analyzed along with the lysate by SDS-PAGE and Western blotting with mouse and human HDGF specific antibodies. The lower band corresponds to the C-terminally processed HDGF (for both the mouse and human HDGF) which could be detected in lysate as well as the supernatants.

5 Discussion

Post-translational modifications (PTM's) are modifications occurring on proteins, like enzyme-catalyzed covalent additions; structural changes; and proteolytic cleavage of peptide backbones. So far, more than 200 PTM's have been characterized. PTM's modulate the function of most eukaryotic proteins by altering their activity state, localization, turnover, and interactions with other proteins. HDGF is a nuclear-targeted mitogen implicated in fetal development, cancer progression and organ remodeling. HDGF was predicted to be modified by phosphorylation on serine and threonine residues and mass spectroscopy based approach also confirmed the use of serine 132, 133 and 165 [58]. Furthermore, HDGF is rapidly dephosphorylated in response to apoptotic inducers in endothelial cells [53]. However, conclusive evidence for these modifications on HDGF and their effects on its function has not yet been determined. It is therefore important to establish how HDGF activity is regulated by PTM's. The present thesis is based on three major projects described in the introduction and undertaken in publications 1 to 3, addressing the question of PTM's of HDGF.

5.1 SUMOylation of HDGF

SUMO, a ~10kDa protein, is a member of the ubiquitin-like protein super-family and is post-translationally conjugated to various cellular proteins in a process that is mechanistically analogous to ubiquitination. Chemically, SUMOylation involves the formation of an iso-peptide bond between the C-terminal carboxyl group of a mature SUMO and the ε-amino group of a lysine residue in the substrate protein. The target lysine residue is generally embedded within an internal motif with the consensus ψKXE/D (where ψ is any hydrophobic amino acid; K is lysine; X can be any amino acid; E is glutamate and D is aspartate) that is conserved in most, but not all, proteins that are targeted by SUMO-1 [59]. So far, four 10-11 kDa SUMO isoforms in mammals have been described. SUMO-1 exhibits 47% similarity at the protein level with SUMO-2 and -3, whereas SUMO-2 and -3 share 95% similarity [57]. Since the identification of the first SUMO-modified substrate, RanGAP, in 1996 [60], a large number of proteins have been shown to be modified by SUMO and new substrates for SUMO-modification continue to be identified. Many of the known SUMO-modified proteins are located in the nucleus with important roles in regulating transcription,

chromatin structure assembly, and DNA repair. PTM by SUMO is also important for regulation and/or localization of proteins with roles in cell proliferation, differentiation, and apoptosis.

We demonstrated that HDGF is modified by the attachment of SUMO-1 in mammalian cells **(Publication 1)**. Our search for the SUMO-1 acceptor sites in HDGF did not reveal a clear SUMOylation consensus sequence ΨKxE/D. Moreover, mutation of all predicted SUMOylation sites did not lead to the loss of SUMOylation. This was coupled with the problem of a high number of lysine residues in HDGF making the search time consuming. We overcame this problem by expressing HDGF in bacteria together with components of the SUMO-1 conjugating system [61]; and later analyzed the SUMO-conjugated HDGF peptide in a mass spectrometry based analysis, to identify the lysine residue used for the conjugation of SUMO-1 as K80 in the non-consensus motif $RK_{80}GF$. In principle, SUMOylation can have varied functional consequences for HDGF. Previous studies showed that the PWWP domain of HDGF is sufficient to bind DNA *in vitro* [27, 28]. To check for the effect of SUMOylation on HDGF ability to bind DNA, we performed chromatin binding assay and were able to show that unSUMOylated HDGF indeed binds to chromatin or chromatin-associated proteins. In contrast, SUMOylated HDGF is excluded from the chromatin-bound fraction and was only found in the cytoskeleton and nuclear matrix proteins containing fraction

As mentioned previously, there are four SUMO isoforms in mammals. For a few substrates, isoform preferences have been documented. RanGAP1 is modified almost exclusively by SUMO-1 *in vivo* [62], while other substrates, such as the pro-myelocytic leukemia protein (PML), have been reported to be conjugated with both SUMO-1 [63] and SUMO-2/3 [64]. Analysis of functional differences between the SUMO isoforms in mammalian cells has been previously studied by comparing the localization and dynamic behavior of YFP-SUMO fusion proteins [65]. The results showed that YFP-SUMO-1 behaves in a manner that is highly distinct from either YFP-SUMO-2 or YFP-SUMO-3. Although the pattern of localization of the isoforms was overlapping in the nucleoplasm and in PML bodies, YFP-SUMO-1 was also uniquely localized to nucleoli, the nuclear envelope, and cytoplasmic foci. Furthermore, YFP-SUMO-1 showed different dynamics from YFP-SUMO- 2 or -3, with slower rates of both recovery and depletion, and also had substantially different responses to physiological stimuli, like heat stress, a known fact from previous SUMO modified protein studies [62]. These findings make it intriguing to study the complexity of SUMO modification on HDGF. To access whether HDGF can be SUMOylated by different SUMO isoforms, we performed co-expression experiments and demonstrated that SUMO-2 and SUMO-3 are also able to modify HDGF. Moreover, they utilize the same non-consensus motif $RK_{80}GF$ for SUMOylation **(refer Section 4.1; Figure 4.1)**. The functional relevance of this modification however needs to be elucidated. It would be interesting to check if there is competition or preference for the use of the site between the SUMO isoforms and if modification of HDGF by different isoforms leads to different functional consequences.

5.2 Phosphorylation dependent regulation of HDGF secretion and processing

Considering its implicated involvement in fetal development, cancer progression and organ remodeling, it is necessary to understand the mechanism of HDGF secretion and its regulation. In this part of the study we focused on HDGF secretion with respect to PTM. We assessed HDGF wt and S→A mutants for differences in expression and secretion. Surprisingly, we could observe a loss of secretion if a phosphorylation site, serine 165, was mutated. Additionally, we have shown that HDGF secretion is attributed to processing of N-terminal region, which is also influenced by the loss of serine 165 phosphorylation site **(Publication 2)**.

A kinase that might be involved in phosphorylation of HDGF at position S165 can be the mitogen-activated protein kinases (MAPK). MAPKs are serine/threonine kinases that tend to phosphorylate their substrates on S/T-P motif. MAPK signaling pathways are characterized by a cascade of multiple kinases that contribute to regulate a variety of extracellular stimuli, thereby control the cellular environment. The MAPK family includes the ERKs (extracellular signal-regulated kinase), the JNKs (c-Jun N-terminal kinase), the p38 MAPKs and ERK5. The ERKs generally regulate cellular proliferation and differentiation processes in response to growth factors and hormones; while the JNKs and p38 MAPKs are primarily activated in response to extracellular stresses, such as U.V. irradiation, osmotic stress or inflammatory stimuli [66]. As the sequence S/T-P is found in ~ 80% of all proteins, additional factors are required to direct individual kinases towards the correct substrates [67]. MAPKs ensure their substrate specificity by interaction through docking motifs, short amino acid stretches located on MAPK-interacting proteins. The classic MAPK docking site was first identified in c-JUN, a substrate of JNK [68]. It is characterized by a cluster of at least two positively charged amino acids followed by a spacer of 2-to-6 residues from a hydrophobic-X-hydrophobic sequence, where the hydrophobic residues are long chain aliphatics (usually Leu, Ile). Both in the spacer and in the sequence immediately C-terminal to the hydrophobic-X-hydrophobic element, there is a high tendency for the presence of Pro, Asn, and/or Gly, which are residues that are both turn-forming and helix-breaking [69]. Interestingly, such a MAPK docking motif is predicted with 100% identity for amino acids ^{8}KEYKCGDLVF^{17} in HDGF (ELM server to identify functional sites on proteins (*http://elm.eu.org/*)). Moreover, S165 in HDGF is a part of a SP motif which might be phosphorylated by a MAPK. Supporting this speculation, it has already been shown that HDGF can activate Erk1/2, a member of the MAPK family. The study suggested that HDGF is involved in the carcinogenesis of gastric epithelial cells and promotes cell proliferation by Erk1/2 activation [44]. Furthermore, we have demonstrated that HDGF secretion is attributed to a processing of N-terminal region of the protein and the processing of HDGF S165A mutant leads to a loss of 13 or more amino acids (analyzed by N-terminal amino acid sequencing). This would lead to the loss of the potential MAPK docking site.

A mechanism of *cis*/*trans* isomerization occurring at this SP phosphorylation motif can be speculated to play a role in this N-terminal processing of HDGF. *Cis/trans* isomerization of peptidyl-prolyl bonds is carried out by the peptidyl-prolyl isomerase (PPIs), and brings about a conformational change in the protein [70]. Such enzymes have been isolated from yeast and known as Ess1, and its human orthologue is known as Pin1. Pin1 consists of 2 domains, an N-terminal WW domain and a carboxy-terminal catalytic domain joined by a short linker. The WW domain IV (group IV) of Pin1 recognizes serine/ threonine-proline (S/T-P) motifs only under phosphorylated state [71, 72]. Pin1 has profound effects on amyloid precursor protein (APP) processing and amyloid-β peptides (Aβ) production [73]. Pin1 accelerates isomerization of APP by binding a phosphorylated TP motif and Pin1 over-expression reduces Aβ secretion from cell cultures. HDGF S165 being a part of such SP motif might interact with Pin1 causing isomerization from *cis* to *trans* conformation in HDGF wt only when it is phosphorylated at S165, whereas for HDGF S165A mutant, Pin1 cannot interact; isomerization is hindered; and thus leaving the *cis* form prone for processing (Figure 5.1). However, further studies identifying the kinase involved in HDGF phosphorylation and/or the protease performing the processing have to be performed to support these hypotheses.

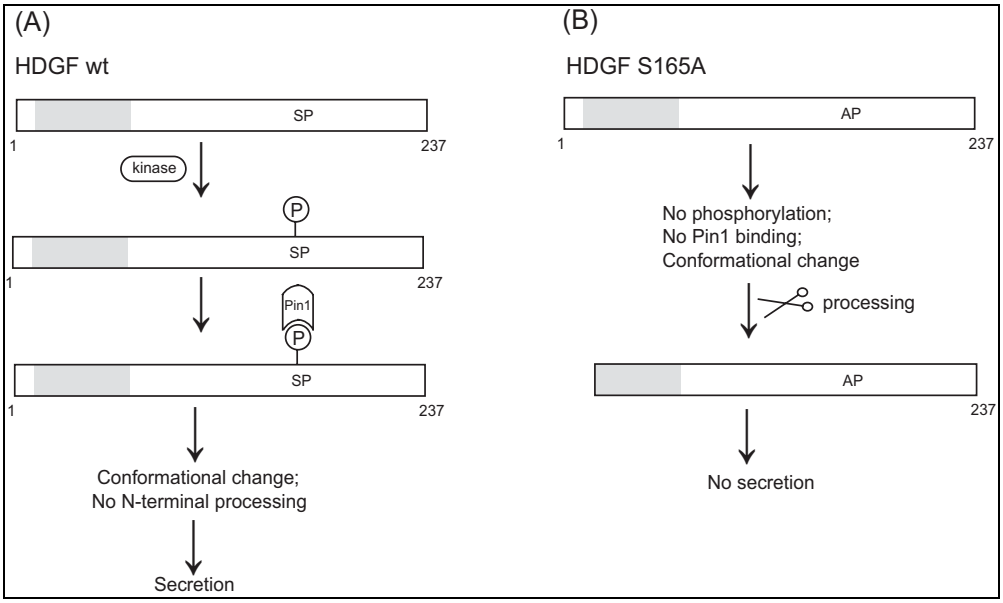


Figure 5.1 Schematic representation of a potential Pin1 mediated *cis*/*trans* isomerization of HDGF regulating its processing and secretion. Part highlighted in grey represents the PWWP domain in HDGF. The potential Pin1 ^{165}SP^{166} binding motif is shown for HDGF wt and corresponding ^{165}AP^{166} mutation for the HDGF S165A.

5.3 HDGF dimerisation

HDGF PWWP module forms dimers by an unusual mechanism of domain swapping which increases its affinity to bind to heparin [29]. The dimer is formed by exchanging the N-terminal β1-β2 strands of the two monomers which are linked by two long L2 loops. The result established that the domain swapped PWWP modules does not occur due to disulfide linkages via Cys_{12}, however this study on the bacterial expressed

PWWP module and excludes the Cys_{108}. Evidence for the existence of a disulfide bond was provided in the original study for HDGF. The mitogenic activity of secreted HDGF was completely lost after treatment by reducing agent [6]. We have demonstrated in our study that HDGF is able to form intra- and inter-molecular disulfide bonds **(Publication 2)**. We provided evidence for HDGF dimerisation is mediated exclusively via Cys_{108}, and N-terminal processing in the S165A mutant or N-terminal truncation of HDGF favors dimer formation. The results suggest that phosphorylation of Ser_{165} in HDGF is required to introduce a structural change in the molecule which might then lead to an interaction of the C-terminal and N-Terminal regions, and further to the formation of an intra-molecular disulfide bond between Cys_{12} and Cys_{108}. In the HDGF S165A mutant the initial structural change cannot occur (possibly by the absence of the motif recognized by a PPI like Pin1) resulting in a loss of the C- and N-terminal interaction and the intra-molecular disulfide bond. Based on these results we propose a model illustrating the possible mechanism **(refer Publication 2; Figure 6)**. N- and C-terminal interactions have been demonstrated HMGB-1, a closely related protein to HDGF, and play a crucial role in maintaining its structure and function. The study revealed that the HMGB-1 acidic C-terminal domain interacts with the basic N-terminal domain and the intramolecular interaction between the two oppositely charged termini is enhanced when serine residues in the acidic tail of HMGB-1 are phosphorylated [74].

In summary, the N-terminal amino acids of HDGF appear to be responsible to mediate non-classical secretion of HDGF. Furthermore, a serine 165 is essential for maintaining N-terminal part, probably by influencing the structure of HDGF.

5.4 HDGF–HRP-2 interaction

5.4.1 Differential expression of HRP-2 correspond to alternatively spliced isoforms

Tissue specific data for expression of HRP-2 was only available at the RNA level which shows ubiquitous distribution [2], while area and cell type specific expression of HRP-2 in the rat brain is also demonstrated at the protein level [14]. In this study we have shown differential expression of HRP-2 protein in brain in a developmental dependent manner **(Publication 3)**. HRP-2 has a wide distribution in cerebrum, heart, muscle, liver, frontal lobe, thymus whereas low levels of expression in lung and spleen tissue (Dietz et al., unpublished data). By screening mouse brain cDNA; along with the two variants already described in the database (referred in this study HRP-2 *isoform a* and *b*); we could also identify a new variant of HRP-2 (referred in this study as HRP-2 *isoform c).* The three different isoforms arise from alternative splicing **(Publication 3)**. The newly discovered isoform is a 616 amino acid protein which shows a deletion in the conserved *hath* region. The 53 amino acid deletion; spanning from aa 66-118; might affect the proper folding of the PWWP domain, which has been suggested as a protein-protein interaction domain. The only other member of the family for which presence of spliced variants has been described in the literature is

LEDGF, occurring as LEDGF/p75 and p52 [75]. All the three HRP-2 isoforms contain two putative helix-turn-helix motifs. The helix-turn-helix motifs described for LEDGF, which are reported to be essential for DNA binding and trans-activation [76], show a high degree of similarity to the putative motifs in HRP-2. The three HRP-2 isoforms displaying the occurrence of the putative helix-turn-helix motifs may well function like LEDGF in modulating gene expression. Also, the presence of a mixed charge cluster, a characteristic of eukaryotic regulatory proteins, makes it plausible that HRP-2 functions as regulatory protein [2]. *In vitro* as well as *in vivo* cells expressing LEDGF exhibit increased resistance to environmental stress and improved survival [76]. Also, LEDGF is produced at higher levels under thermal and oxidative stress [77] and is important in up-regulating stress-associated genes by binding to the stress response element and heat shock elements present in these genes [76]. Considering its high similarity to LEDGF, HRP-2 might also play a role in defense mechanisms of cells against stress. In future experiments should be carried out to test the role of HRP-2 in the stress response of cells. The appearance of various bands in tissue may be attributed to the regulated function of different HRP-2 isoforms and further studies in this area should be undertaken to investigate the precise roles played by the various isoforms in different tissues and stages.

5.4.2 HDGF/HRP-2 interaction studies

As mentioned before, HDGF can form domain swapped dimers derived from the interaction of the structured N-terminal part [29]. Sequence and structural identity between N-terminal regions of all HRPs family proteins raises the possibility for the interplay and heteromerisation between different HRP members. Moreover, a recent study indicates that HDGF acts as a transcriptional co-repressor for SMYD1 gene expression by recruiting CtBP and this interaction is critical for nuclear function of HDGF [78]. It might be that HRP-2/HDGF heteromerisation acts in similar manner modulating DNA binding to specific targets and regulate gene expression. By co-precipitation studies we were able to show that HRP-2 and HDGF interact and the new HRP-2 *isoform c* can additionally interact with a lower molecular mass form of HDGF, which correlates with the N-terminally processed HDGF S165A mutant **(Publication 3)**.

From our results we hypothesize that protein-protein interaction between different members of the HRP family might play a very important role in their functions. This finding opens a new perspective in the field of HRP family research.

5.5 Caspase dependent cleavage of HDGF

Caspases are a family of cysteine proteases expressed in virtually all animal cells and play an essential role in many forms of cell death by apoptosis [79]. A total of 12 cysteine proteases known as caspases have been identified in mammals: caspase-1 to -10, caspase-12 and caspase-14. [80]. There are two main pathways of caspase activation leading to apoptotic cell death. The intrinsic pathway; which involves proteins of the Bcl-2 family, which in response to a cellular damage migrate to the mitochondria to either antagonize anti-apoptotic proteins or activate multi-domain pro-apoptotic proteins Bax and Bak [81]; and the extrinsic pathway, which starts at the plasma membrane by engagement of transmembrane death receptors such as Fas/CD95, tumor necrosis factor receptor 1 (TNF-R1) and TNF-related apoptosis-inducing ligand (TRAIL) receptors. Multi-protein complexes are formed in which caspases are recruited by the adaptor molecule Fas-associated death domain protein (FADD) and activated. In turn, these enzymes either directly activate the caspase cascade or connect the extrinsic and the intrinsic pathway [82]. Previous studies suggests that HDGF is involved in apoptosis and also is regulated by caspase dependent mechanism [53]. Knock-down of HDGF not only induced apoptosis in human cancer cells through the Bad-mediated intrinsic apoptotic pathway, but also the Fas-mediated extrinsic apoptotic pathway suppressing anchorage-independent growth of cancer cells [54, 55]. RNAi-induced silencing of the HDGF prevented the release of pro-apoptotic factors, Smac/Diablo, from the mitochondria upon TNF-α treatment [52]. Triggering endothelial cell apoptosis by TNFα and cycloheximide leads to an early dephosphorylation of HDGF before mitochondrial membrane permeabilization and downstream from an initiator caspase [53]. All these studies suggest involvement and regulation of HDGF by components of the apoptotic pathway.

Based on our previous knowledge for over-expression of mouse HDGF wt we always observed a characteristic double band pattern in the total cell lysate exclusively in COS-7 cells **(refer Section 4.2, Figure 4.2(a))**. This pattern can be only observed for over-expression of the untagged mouse HDGF wt, but interestingly not for the untagged human HDGF wt (data not shown). This lead to the hypothesis that occurrence of such a pattern is not unspecific or not derived from degradation, but involves a specific cleavage product upon over-expression. Our results show that the two forms differ by 4 kDa; the lower molecular mass form is C-terminally processed; and it is preferentially secreted and suggest that C-terminal cleavage of mouse HDGF might have a regulated function under specific conditions. It is important to note that, the lower molecular mass form of HDGF observed here is not same as the processed form of HDGF demonstrated earlier which is observed in regulation of HDGF secretion.

Considering the difference in the molecular mass in HDGF, we searched for cleavage motifs which might be relevant and lead us to search of a potential caspase cleavage site in the C-terminal region of HDGF. Caspases typically recognize tetra-peptide (P_4-P_3-P_2-P_1) motifs in their substrates and cleave between the P_1 amino acid residue and the adjacent C-terminal amino acid in the peptide chain. Analysis of mouse

HDGF amino acid sequence using the CASVM (1.0) server for predicting caspase cleavage sites (*http://www.casbase.org/casvm/*); predicted a ^{202}SEPD^{205} caspase cleavage motif. Such a motif has been shown to be functional for sterol-regulatory-element-binding protein (SREBP) [83]. Cleavage at this predicted motif in HDGF would lead to a loss of C-terminal 32 amino acids. Interestingly, comparison of HDGF amino acid sequences from different species shows that the aspartate residue where the cleavage might occur in mouse HDGF is replaced by glycine in human HDGF (Figure 5.2). Using mutants of HDGF for this cleavage site in mouse and human forms (mHDGF D205G, hHDGF wt and hHDGF G205D mutants); we could confirm the use of this site in mouse HDGF, which can be mimicked for human HDGF. Moreover, we observed that for mouse HDGF wt this cleavage favors secretion and loss of the site in mutant renders the uncleaved full length form that is secreted. Similarly behavior was observed for human HDGF knock-in mutant **(refer section 4.2, Figure 4.2 (b))**.

Mus musculus	^{200}TPSEPDSGQG^{209}
Rattus norvegicus	^{200}TPSEPDSGQG^{209}
Bos taurus	^{200}TPSEPGSGRG^{209}
Pan troglodytes	^{200}TPSEPGSGRG^{209}
Homo sapiens	^{200}TPSEPGSGRG^{209}

Figure 5.2 HDGF amino acid sequence alignment (residues 200 to 209) from various species (*Mus musculus* accession number NM_008231.4 ; *Rattus norvegicus* accession number AF389348.1 ; *Bos taurus* accession number BC123463.1 ; *Pan troglodytes* accession number XM_001167346.1 and *Homo sapiens* accession number BC018991.1). Amino acids highlighted in dark grey background represent the conserved residues from the predicted caspase cleavage motif SEPD in all the species; while the one in light grey background represents the evolutionary changed aspartate to glycine residue.

Our results show that not only the predicted caspase site in HDGF is functional, but also that it might have evolutionary significance for the protein. Further studies dealing with the use of this site in apoptotic conditions and establishing the evolutionary link need to be performed.

5.6 Conclusion and perspectives

There are four major findings from this study; (1) HDGF is modified by SUMOylation and the modification negatively affects its ability to bind to DNA. Additionally, apart from SUMO-1, HDGF can also be modified by SUMO-2 and SUMO-3 at the same site; (2) A phosphorylation dependent mechanism regulates HDGF processing and secretion; (3) HDGF is able to form intra- and inter-molecular disulfide bonds and N-terminal processing favors HDGF dimer formation and; (4) HDGF is able to interact with another member of the HRP family and specificity of this interaction is also affected by phosphorylation dependent processing. Taken together, the data combines various aspects of HDGF PTM's. On the basis of findings discussed in this thesis a number of research areas can be investigated. A few of them are discussed and elaborated below.

For further analysis of HDGF SUMOylation, we can directly check for functional consequences of the modification of SUMO-2 and SUMO-3 as studied for SUMO-1. Also, SUMOylation of other HRP family members can be undertaken. In a preliminary study, it has been demonstrated that HRP-3 is SUMO-1 modified in mammalian cells (Dietz et al., unpublished data). The lysine 80 in HDGF from the non-consensus motif is also conserved in HRP-3 and it also contains five other candidate SUMOylation sites. To determine the exact SUMO-1 acceptor lysine(s) within HRP-3, it would be worthwhile to prepare HRP-3 K→R mutants and test for influence of SUMOylation on HRP-3.

The pivotal role of N-terminal amino acids mediating HDGF secretion was established from our results. To further ascertain and elucidate whether these N-terminal amino acids are relevant only for HDGF or if they can also mediate secretion of other proteins by serving as a non-classical signal peptide, we prepared and analyzed other reporter proteins chimera. We could show that other proteins (EGFP and SNAP-*Strep*Tag) fused with a peptide comprising the first 10 amino acids of HDGF are more efficiently secreted than with the first 19 amino acids. According to the solved HDGF NMR structure the first 10 amino acids are unstructured in solution, while the first structural element of HDGF begins at Leu_{15} with the first β-strand [19]. Possibly, tagging of the first 10 amino acids of HDGF functions as free floating peptide, like in HDGF, whereas with first 19 amino acids lead to formation of a structural element which interferes with the secretion machinery. Although this might be an artificial system, it still functions in mediating secretion of other proteins, which might be used as a tool to study secretion pathways. All eukaryotic cells from yeast to man are characterized by an elaborate secretory machine that recognizes signal peptide-bearing proteins resulting in their translocation across the membrane of the endoplasmic reticulum (ER). Once localized to the lumen of the ER, secretory proteins are packaged into transport vesicles provided they pass ER quality control measures. Following cargo delivery to the Golgi apparatus, post-Golgi transport carriers fuse with the plasma membrane, a process that eventually results in the release of classical secretory proteins into the extracellular space [84]. However, a number of secretory proteins with defined extracellular functions have been shown not to contain functional signal peptides and do not enter the ER membrane translocation machinery. Intriguingly, unconventional secretory proteins comprise a group of molecules of significant biomedical relevance such as the pro-angiogenic growth factor fibroblast growth factor-2 (FGF-2) [85], inflammatory cytokines such as interleukin-1β (IL-1 β) [86] and macrophage migration inhibitory factor (MIF) [87], etc. Few of the non-classical pathways and their inhibitors are listed below in Table 5.1. The mechanisms by which these proteins are secreted provide us the opportunity to search within an array of non-classical pathways to understand the mechanism of HDGF secretion.

Table 5.1 List of few secretion pathways and their specific inhibitors.

No.	Secretion pathway/ mechanism	Example protein	Inhibitor	Reference
1	ER-Golgi	-	Brefeldin A	-
2	Endolysosomal	HMGB-1	Glybenclamide	[88]
3	Exosome mediated	HSP90	Dimethyl amiloride	[89]
4	Membrane blebbing	Trans-glutaminase	Blebbistatin; Y-27632	[90]

The discovery of new HRP-2 splice variants and their differential expression opens a way for better understanding the HRP family functions. Confirming the expression of splice variants using alternative method like real-time PCR, and analyzing the functional behavior of the various isoforms are steps in future direction. Estimating the relative abundance of each variant is also necessary in studying the biological functions of the isoforms. We have shown that there is interaction of HRP-2 isoforms with HDGF. It might be that the interaction within the HRP family acts as central regulator controlling the mode of action of these proteins. The consequences of such an interaction on cellular pathways offer an interesting area of research.

Our results demonstrate that HDGF can also be cleaved at the C-terminus at a potential caspase cleavage site ^{202}SEPD^{205}. However, drawing a conclusive link between the cleavage of a particular protein and a cellular change that takes place has not been straightforward. While the cleavage of certain proteins is probably vital to achieve the controlled and complete collapse of the cell, the cleavage of many other proteins may ultimately be of little consequence to the death process. Such “bystander substrates” may simply contain the appropriate caspase-consensus cleavage site(s) by chance alone [91]. Also, a number of proteins with the ability to repress the cell death program have been discovered. Perhaps the most well known of these is the Bcl-2 protein [92]. A whole family of Bcl-2 related proteins have been discovered, some of which are repressors of apoptosis while others, paradoxically, can promote entry into the death program. Many proteins that block cell death appear to do so by protecting endogenous substrates from the destructive action of the caspases by offering themselves up for proteolysis, often neutralizing the activity of the caspase. In the case of Bcl-2, many binding partners have been discovered in recent years. One study indicates that the apoptosis induced by knock-down of HDGF is mediated by Bad (a pro-apoptotic Bcl-2 family protein), which trigger apoptosis by dimerising with Bcl-2 and Bcl-xL, and neutralize their anti-apoptotic effects. This suggests role of HDGF as a survival factor, although the mechanism to regulate the intracellular signaling pathways is yet to be defined [54]. Investigating the caspase involved in HDGF cleavage, its functional effect on HDGF and checking for HDGF ability to act as cell survival factor by acting in co-operation with other repressor proteins such as Bcl-2 should be a pursued in the future.

Finally, it can be concluded that the advances in the understanding of post-translational modifications of HDGF demonstrated in this thesis, would lead to the realization of importance of these modifications to control and regulate HDGF and HRP family protein functions.

6
References

1. Nakamura H, Izumoto Y, Kambe H, Kuroda T, Mori T, Kawamura K, Yamamoto H & Kishimoto T (1994) Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein. *J Biol Chem* **269**, 25143-25149.

2. Izumoto Y, Kuroda T, Harada H, Kishimoto T & Nakamura H (1997) Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. *Biochem Biophys Res Commun* **238**, 26-32.

3. Ikegame K, Yamamoto M, Kishima Y, Enomoto H, Yoshida K, Suemura M, Kishimoto T & Nakamura H (1999) A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus. *Biochem Biophys Res Commun* **266**, 81-87.

4. Ge H, Si Y & Roeder RG (1998) Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *EMBO J* **17**, 6723-6729.

5. Dietz F, Franken S, Yoshida K, Nakamura H, Kappler J & Gieselmann V (2002) The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. *Biochem J* **366**, 491-500.

6. Nakamura H, Kambe H, Egawa T, Kimura Y, Ito H, Hayashi E, Yamamoto H, Sato J & Kishimoto S (1989) Partial purification and characterization of human hepatoma-derived growth factor. *Clin Chim Acta* **183**, 273-284.

7. Oliver JA & Al-Awqati Q (1998) An endothelial growth factor involved in rat renal development. *J Clin Invest* **102**, 1208-1219.

8. Everett AD, Lobe DR, Matsumura ME, Nakamura H & McNamara CA (2000) Hepatoma-derived growth factor stimulates smooth muscle cell growth and is expressed in vascular development. *J Clin Invest* **105**, 567-575.

9. Kishima Y, Yoshida K, Enomoto H, Yamamoto M, Kuroda T, Okuda Y, Uyama H & Nakamura H (2002) Antisense oligonucleotides of hepatoma-derived growth factor (HDGF) suppress the proliferation of hepatoma cells. *Hepatogastroenterology* **49**, 1639-1644.

10. Everett AD, Stoops T & McNamara CA (2001) Nuclear targeting is required for hepatoma-derived growth factor-stimulated mitogenesis in vascular smooth muscle cells. *J Biol Chem* **276**, 37564-37568.

11. Kishima Y, Yamamoto H, Izumoto Y, Yoshida K, Enomoto H, Yamamoto M, Kuroda T, Ito H, Yoshizaki K & Nakamura H (2002) Hepatoma-derived growth factor stimulates cell growth after translocation to the nucleus by nuclear localization signals. *J Biol Chem* **277**, 10315-10322.

12. Mosevitsky MI, Novitskaya VA, Iogannsen MG & Zabezhinsky MA (1989) Tissue specificity of nucleo-cytoplasmic distribution of HMG1 and HMG2 proteins and their probable functions. *Eur J Biochem* **185**, 303-310.

13. Cherepanov P, Devroe E, Silver PA & Engelman A (2004) Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J Biol Chem* **279**, 48883-48892.

14. El-Tahir HM, Dietz F, Dringen R, Schwabe K, Strenge K, Kelm S, Abouzied MM, Gieselmann V & Franken S (2006) Expression of hepatoma-derived growth factor family members in the adult central nervous system. *BMC Neurosci* **7**, 6.

15. El-Tahir HM, Abouzied MM, Gallitzendoerfer R, Gieselmann V & Franken S (2009) Hepatoma-derived growth factor-related protein-3 interacts with microtubules and promotes neurite outgrowth in mouse cortical neurons. *J Biol Chem* **284**, 11637-11651.

16. Engelman A & Cherepanov P (2008) The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. *PLoS Pathog* **4**, e1000046.

17. Sutherland HG, Newton K, Brownstein DG, Holmes MC, Kress C, Semple CA & Bickmore WA (2006) Disruption of Ledgf/Psip1 results in perinatal mortality and homeotic skeletal transformations. *Mol Cell Biol* **26**, 7201-7210.

18. Fatma N, Singh DP, Shinohara T & Chylack LT, Jr. (2001) Transcriptional regulation of the antioxidant protein 2 gene, a thiol-specific antioxidant, by lens epithelium-derived growth factor to protect cells from oxidative stress. *J Biol Chem* **276**, 48899-48907.

19. Sue SC, Chen JY, Lee SC, Wu WG & Huang TH (2004) Solution structure and heparin interaction of human hepatoma-derived growth factor. *J Mol Biol* **343**, 1365-1377.

20. Abouzied MM, Baader SL, Dietz F, Kappler J, Gieselmann V & Franken S (2004) Expression patterns and different subcellular localization of the growth factors HDGF (hepatoma-derived growth factor) and HRP-3 (HDGF-related protein-3) suggest functions in addition to their mitogenic activity. *Biochem J* **378**, 169-176.

21. Sue SC, Chen JY & Huang TH (2004) Sequence specific 1H, 13C and 15N resonance assignments of the hath-domain of human hepatoma-derived growth factor. *J Biomol NMR* **29**, 95-96.

22. Stec I, Nagl SB, van Ommen GJ & den Dunnen JT (2000) The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett* **473**, 1-5.

23. Qiu C, Sawada K, Zhang X & Cheng X (2002) The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. *Nat Struct Biol* **9**, 217-224.

24. Chen T, Tsujimoto N & Li E (2004) The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol Cell Biol* **24**, 9048-9058.

25. Ge YZ, Pu MT, Gowher H, Wu HP, Ding JP, Jeltsch A & Xu GL (2004) Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *J Biol Chem* **279**, 25447-25454.

26. Slater LM, Allen MD & Bycroft M (2003) Structural variation in PWWP domains. *J Mol Biol* **330**, 571-576.

27. Lukasik SM, Cierpicki T, Borloz M, Grembecka J, Everett A & Bushweller JH (2006) High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. *Protein Sci* **15**, 314-323.

28. Yang J & Everett AD (2007) Hepatoma-derived growth factor binds DNA through the N-terminal PWWP domain. *BMC Mol Biol* **8**, 101.

29. Sue SC, Lee WT, Tien SC, Lee SC, Yu JG, Wu WJ, Wu WG & Huang TH (2007) PWWP module of human hepatoma-derived growth factor forms a domain-swapped dimer with much higher affinity for heparin. *J Mol Biol* **367**, 456-472.

30. Zhou Z, Yamamoto Y, Sugai F, Yoshida K, Kishima Y, Sumi H, Nakamura H & Sakoda S (2004) Hepatoma-derived growth factor is a neurotrophic factor harbored in the nucleus. *J Biol Chem* **279**, 27320-27326.

31. Marubuchi S, Okuda T, Tagawa K, Enokido Y, Horiuchi D, Shimokawa R, Tamura T, Qi ML, Eishi Y, Watabe K, Shibata M, Nakagawa M & Okazawa H (2006) Hepatoma-derived growth factor, a new trophic factor for motor neurons, is up-regulated in the spinal cord of PQBP-1 transgenic mice before onset of degeneration. *J Neurochem* **99**, 70-83.

32. Hu TH, Huang CC, Liu LF, Lin PR, Liu SY, Chang HW, Changchien CS, Lee CM, Chuang JH & Tai MH (2003) Expression of hepatoma-derived growth factor in hepatocellular carcinoma. *Cancer* **98**, 1444-1456.

33. Yoshida K, Nakamura H, Okuda Y, Enomoto H, Kishima Y, Uyama H, Ito H, Hirasawa T, Inagaki S & Kawase I (2003) Expression of hepatoma-derived growth factor in hepatocarcinogenesis. *J Gastroenterol Hepatol* **18**, 1293-1301.

34. Bernard K, Litman E, Fitzpatrick JL, Shellman YG, Argast G, Polvinen K, Everett AD, Fukasawa K, Norris DA, Ahn NG & Resing KA (2003) Functional proteomic analysis of melanoma progression. *Cancer Res* **63**, 6716-6725.

35. Ren H, Tang X, Lee JJ, Feng L, Everett AD, Hong WK, Khuri FR & Mao L (2004) Expression of hepatoma-derived growth factor is a strong prognostic predictor for patients with early-stage non-small-cell lung cancer. *J Clin Oncol* **22**, 3230-3237.

36. Iwasaki T, Nakagawa K, Nakamura H, Takada Y, Matsui K & Kawahara K (2005) Hepatoma-derived growth factor as a prognostic marker in completely resected non-small-cell lung cancer. *Oncol Rep* **13**, 1075-1080.

37. Hayashi E, Kuramitsu Y, Okada F, Fujimoto M, Zhang X, Kobayashi M, Iizuka N, Ueyama Y & Nakamura K (2005) Proteomic profiling for cancer progression: Differential display analysis for the expression of intracellular proteins between regressive and progressive cancer cell lines. *Proteomics* **5**, 1024-1032.

38. Lepourcelet M, Tou L, Cai L, Sawada J, Lazar AJ, Glickman JN, Williamson JA, Everett AD, Redston M, Fox EA, Nakatani Y & Shivdasani RA (2005) Insights into developmental mechanisms and cancers in the mammalian intestine derived from serial analysis of gene expression and study of the hepatoma-derived growth factor (HDGF). *Development* **132**, 415-427.

39. Uyama H, Tomita Y, Nakamura H, Nakamori S, Zhang B, Hoshida Y, Enomoto H, Okuda Y, Sakon M, Aozasa K, Kawase I, Hayashi N & Monden M (2006) Hepatoma-derived growth factor is a novel prognostic factor for patients with pancreatic cancer. *Clin Cancer Res* **12**, 6043-6048.

40. Yamamoto S, Tomita Y, Hoshida Y, Takiguchi S, Fujiwara Y, Yasuda T, Doki Y, Yoshida K, Aozasa K, Nakamura H & Monden M (2006) Expression of hepatoma-derived growth factor is correlated with lymph node metastasis and prognosis of gastric carcinoma. *Clin Cancer Res* **12**, 117-122.

41. Folkman J (1995) Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nat Med* **1**, 27-31.

42. Okuda Y, Nakamura H, Yoshida K, Enomoto H, Uyama H, Hirotani T, Funamoto M, Ito H, Everett AD, Hada T & Kawase I (2003) Hepatoma-derived growth factor induces tumorigenesis in vivo through both direct angiogenic activity and induction of vascular endothelial growth factor. *Cancer Sci* **94**, 1034-1041.

43. Everett AD (2001) Identification, cloning, and developmental expression of hepatoma-derived growth factor in the developing rat heart. *Dev Dyn* **222**, 450-458.

44. Mao J, Xu Z, Fang Y, Wang H, Xu J, Ye J, Zheng S & Zhu Y (2008) Hepatoma-derived growth factor involved in the carcinogenesis of gastric epithelial cells through promotion of cell proliferation by Erk1/2 activation. *Cancer Sci* **99**, 2120-2127.

45. Enomoto H, Yoshida K, Kishima Y, Okuda Y & Nakamura H (2002) Participation of hepatoma-derived growth factor in the regulation of fetal hepatocyte proliferation. *J Gastroenterol* **37 Suppl 14**, 158-161.

46. Raines EW & Ross R (1993) Smooth muscle cells and the pathogenesis of the lesions of atherosclerosis. *Br Heart J* **69**, S30-37.

47. Everett AD & Bushweller J (2003) Hepatoma derived growth factor is a nuclear targeted mitogen. *Curr Drug Targets* **4**, 367-371.

48. Mori M, Morishita H, Nakamura H, Matsuoka H, Yoshida K, Kishima Y, Zhou Z, Kida H, Funakoshi T, Goya S, Yoshida M, Kumagai T, Tachibana I, Yamamoto Y, Kawase I & Hayashi S (2004) Hepatoma-derived growth factor is involved in lung remodeling by stimulating epithelial growth. *Am J Respir Cell Mol Biol* **30**, 459-469.

49. Nakamura H, Yoshida K, Ikegame K, Kishima Y, Uyama H & Enomoto H (2002) Antibodies against hepatoma-derived growth factor and mucosal repair in ulcerative colitis. *J Gastroenterol* **37 Suppl 14**, 8-14.

50. Enomoto H, Nakamura H, Liu W, Yoshida K, Okuda Y, Imanishi H, Saito M, Shimomura S, Hada T & Nishiguchi S (2009) Hepatoma-derived growth factor is induced in liver regeneration. *Hepatol Res* **39**, 988-997.

51. Deng Y, Ren X, Yang L, Lin Y & Wu X (2003) A JNK-dependent pathway is required for TNFalpha-induced apoptosis. *Cell* **115**, 61-70.

52. Machuy N, Thiede B, Rajalingam K, Dimmler C, Thieck O, Meyer TF & Rudel T (2005) A global approach combining proteome analysis and phenotypic screening with RNA interference yields novel apoptosis regulators. *Mol Cell Proteomics* **4**, 44-55.

53. Clermont F, Gonzalez NS, Communi D, Franken S, Dumont JE & Robaye B (2008) HDGF is dephosphorylated during the early steps of endothelial cell apoptosis in a caspase-dependent way. *J Cell Biochem* **104**, 1161-1171.

54. Tsang TY, Tang WY, Tsang WP, Co NN, Kong SK & Kwok TT (2008) Downregulation of hepatoma-derived growth factor activates the Bad-mediated apoptotic pathway in human cancer cells. *Apoptosis* **13**, 1135-1147.

55. Tsang TY, Tang WY, Tsang WP, Co NN, Kong SK & Kwok TT (2009) Mechanistic study on growth suppression and apoptosis induction by targeting hepatoma-derived growth factor in human hepatocellular carcinoma HepG2 cells. *Cell Physiol Biochem* **24**, 253-262.

56. Thakar K, Niedenthal R, Okaz E, Franken S, Jakobs A, Gupta S, Kelm S & Dietz F (2008) SUMOylation of the hepatoma-derived growth factor negatively influences its binding to chromatin. *FEBS J* **275**, 1411-1426.

57. Kim KI, Baek SH & Chung CH (2002) Versatile protein tag, SUMO: its enzymology and biological function. *J Cell Physiol* **191**, 257-268.

58. Shu H, Chen S, Bi Q, Mumby M & Brekken DL (2004) Identification of phosphoproteins and their phosphorylation sites in the WEHI-231 B lymphoma cell line. *Mol Cell Proteomics* **3**, 279-286.

59. Melchior F, Schergaut M & Pichler A (2003) SUMO: ligases, isopeptidases and nuclear pores. *Trends Biochem Sci* **28**, 612-618.

60. Matunis MJ, Coutavas E & Blobel G (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J Cell Biol* **135**, 1457-1470.

61. Uchimura Y, Nakamura M, Sugasawa K, Nakao M & Saitoh H (2004) Overproduction of eukaryotic SUMO-1- and SUMO-2-conjugated proteins in Escherichia coli. *Anal Biochem* **331**, 204-206.

62. Saitoh H & Hinchey J (2000) Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. *J Biol Chem* **275**, 6252-6258.

63. Sternsdorf T, Puccetti E, Jensen K, Hoelzer D, Will H, Ottmann OG & Ruthardt M (1999) PIC-1/SUMO-1-modified PML-retinoic acid receptor alpha mediates arsenic trioxide-induced apoptosis in acute promyelocytic leukemia. *Mol Cell Biol* **19**, 5170-5178.

64. Kamitani T, Nguyen HP, Kito K, Fukuda-Kamitani T & Yeh ET (1998) Covalent modification of PML by the sentrin family of ubiquitin-like proteins. *J Biol Chem* **273**, 3117-3120.

65. Ayaydin F & Dasso M (2004) Distinct in vivo dynamics of vertebrate SUMO paralogues. *Mol Biol Cell* **15**, 5208-5218.

66. Biondi RM & Nebreda AR (2003) Signalling specificity of Ser/Thr protein kinases through docking-site-mediated interactions. *Biochem J* **372**, 1-13.

67. Bardwell L (2006) Mechanisms of MAPK signalling specificity. *Biochem Soc Trans* **34**, 837-841.

68. Kallunki T, Deng T, Hibi M & Karin M (1996) c-Jun can recruit JNK to phosphorylate dimerization partners via specific docking interactions. *Cell* **87**, 929-939.

69. Bardwell AJ, Flatauer LJ, Matsukuma K, Thorner J & Bardwell L (2001) A conserved docking site in MEKs mediates high-affinity binding to MAP kinases and cooperates with a scaffold protein to enhance signal transmission. *J Biol Chem* **276**, 10374-10386.

70. Shaw PE (2002) Peptidyl-prolyl isomerases: a new twist to transcription. *EMBO Rep* **3**, 521-526.

71. Sudol M, Sliwa K & Russo T (2001) Functions of WW domains in the nucleus. *FEBS Lett* **490**, 190-195.

72. Landrieu I, Odaert B, Wieruszeski JM, Drobecq H, Rousselot-Pailley P, Inze D & Lippens G (2001) p13(SUC1) and the WW domain of PIN1 bind to the same phosphothreonine-proline epitope. *J Biol Chem* **276**, 1434-1438.

73. Pastorino L, Sun A, Lu PJ, Zhou XZ, Balastik M, Finn G, Wulf G, Lim J, Li SH, Li X, Xia W, Nicholson LK & Lu KP (2006) The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-beta production. *Nature* **440**, 528-534.

74. Thomsen MS, Franssen L, Launholt D, Fojan P & Grasser KD (2004) Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein. *Biochemistry* **43**, 8029-8037.

75. Singh DP, Kimura A, Chylack LT, Jr. & Shinohara T (2000) Lens epithelium-derived growth factor (LEDGF/p75) and p52 are derived from a single gene by alternative splicing. *Gene* **242**, 265-273.

76. Singh DP, Kubo E, Takamura Y, Shinohara T, Kumar A, Chylack LT, Jr. & Fatma N (2006) DNA binding domains and nuclear localization signal of LEDGF: contribution of two helix-turn-helix (HTH)-like domains and a stretch of 58 amino acids of the N-terminal to the trans-activation potential of LEDGF. *J Mol Biol* **355**, 379-394.

77. Sharma P, Singh DP, Fatma N, Chylack LT, Jr. & Shinohara T (2000) Activation of LEDGF gene by thermal-and oxidative-stresses. *Biochem Biophys Res Commun* **276**, 1320-1324.

78. Yang J & Everett AD (2009) Hepatoma-derived growth factor represses SET and MYND domain containing 1 gene expression through interaction with C-terminal binding protein. *J Mol Biol* **386**, 938-950.

79. Nicholson DW & Thornberry NA (1997) Caspases: killer proteases. *Trends Biochem Sci* **22**, 299-306.

80. Salvesen GS & Abrams JM (2004) Caspase activation - stepping on the gas or releasing the brakes? Lessons from humans and flies. *Oncogene* **23**, 2774-2784.

81. Marsden VS & Strasser A (2003) Control of apoptosis in the immune system: Bcl-2, BH3-only proteins and more. *Annu Rev Immunol* **21**, 71-105.

82. Launay S, Hermine O, Fontenay M, Kroemer G, Solary E & Garrido C (2005) Vital functions for lethal caspases. *Oncogene* **24**, 5137-5148.

83. Wang X, Zelenski NG, Yang J, Sakai J, Brown MS & Goldstein JL (1996) Cleavage of sterol regulatory element binding proteins (SREBPs) by CPP32 during apoptosis. *EMBO J* **15**, 1012-1020.

84. Schatz G & Dobberstein B (1996) Common principles of protein translocation across membranes. *Science* **271**, 1519-1526.

85. Nugent MA & Iozzo RV (2000) Fibroblast growth factor-2. *Int J Biochem Cell Biol* **32**, 115-120.

86. Braddock M & Quinn A (2004) Targeting IL-1 in inflammatory disease: new opportunities for therapeutic intervention. *Nat Rev Drug Discov* **3**, 330-339.

87. Lue H, Kleemann R, Calandra T, Roger T & Bernhagen J (2002) Macrophage migration inhibitory factor (MIF): mechanisms of action and role in disease. *Microbes Infect* **4**, 449-460.

88. Gardella S, Andrei C, Ferrera D, Lotti LV, Torrisi MR, Bianchi ME & Rubartelli A (2002) The nuclear protein HMGB1 is secreted by monocytes via a non-classical, vesicle-mediated secretory pathway. *EMBO Rep* **3**, 995-1001.

89. Yu X, Harris SL & Levine AJ (2006) The regulation of exosome secretion: a novel function of the p53 protein. *Cancer Res* **66**, 4795-4801.

90. Aumuller G, Wilhelm B & Seitz J (1999) Apocrine secretion--fact or artifact? *Ann Anat* **181**, 437-446.

91. Martin SJ & Green DR (1995) Protease activation during apoptosis: death by a thousand cuts? *Cell* **82**, 349-352.

92. Reed JC (1997) Double identity for proteins of the Bcl-2 family. *Nature* **387**, 773-776.

A
Appendix

A.1 Abbreviations

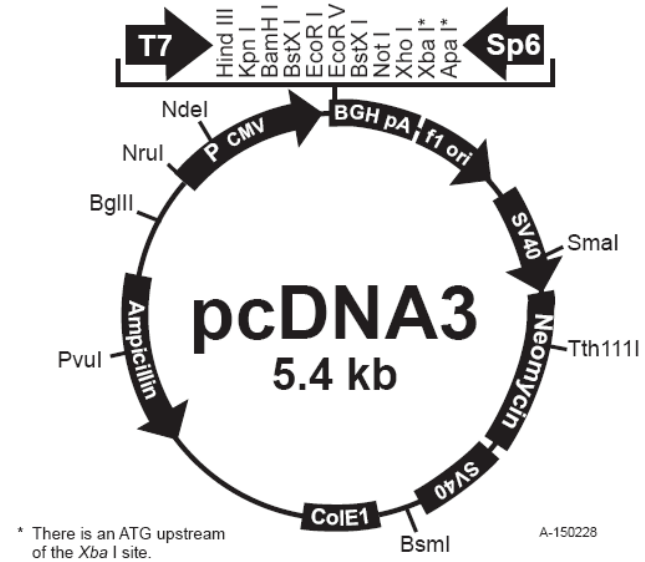
APP	amyloid precursor protein
ATP	adenosine triphosphate
BSA	bovine serum albumin
CAMKII	calmodulin independent protein kinase II
cAMP	cyclic adenosine monophosphate
cDNA	complementary deoxyribonucleic acid
DAPI	4´-6-diamidino-2-phenylindole dihydrochloride
dd	double distilled
dNTP	deoxy nucleotide triphophasphate
ddNTP	dideoxy nucleotide triphophasphate
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
Dnmt	DNA methyltransferase homology domain
dsDNA	double stranded DNA
DTT	dithiothretol
ECL	enhanced chemiluminescence
EDTA	ethylene diamine tetraacetic acid
EGFP	enhanced green fluorescent protein
ERK	extracellular signal-regulated kinase
FCS	fetal calf serum
hath	homologous to the amino terminus of HDGF
HEK	human embryonic kidney
HDGF	hepatoma-derived growth factor
HMG-1	high mobility group protein-1
HRP	HDGF-related protein
HSF-2	heat shock factor-2
IAA	iodoacetic acid
IL-1β	interleukin-1β
JNK	c-Jun N-terminal kinase
KDS	potassium dodecyl sulfate
kDa	kilo Daltons
LEDGF	lens epithelium-derived growth factor
MAPK	mitogen-activated protein kinases
NEM	N-ethylmaleimide
NLS	nuclear localization signal
NMR	nuclear magnetic resonance
NSCLC	non-small cell lung cancer
NTP	nucleoside triphosphate
OD	optical density
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PKA	protein kinase A
PKC	protein kinase C
PML	promyelocytic leukemia protein
PPi	peptidyl-prolyl isomerase
PTM	post-translational modifications
PVDF	polyvinylidene fluoride
RanGAP	RanGTPase-activating protein
RNA	ribonucleic acid
rpm	rounds per minute
RT	room temperature
RT-PCR	reverse transcriptase PCR
SUMO	small ubiquitin-like modifier
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SMC	smooth muscle cell
SPR	surface plasmon resonance
SUMO	small ubiquitin-like modifier
TBS	tris-buffered saline
TGF	transforming growth factor
Ubc-9	ubiquitin conjugating
V	volts
VEGF	vascular endothelium growth factor
WB	western blot
wt	wild type
YFP	yellow fluorescent protein

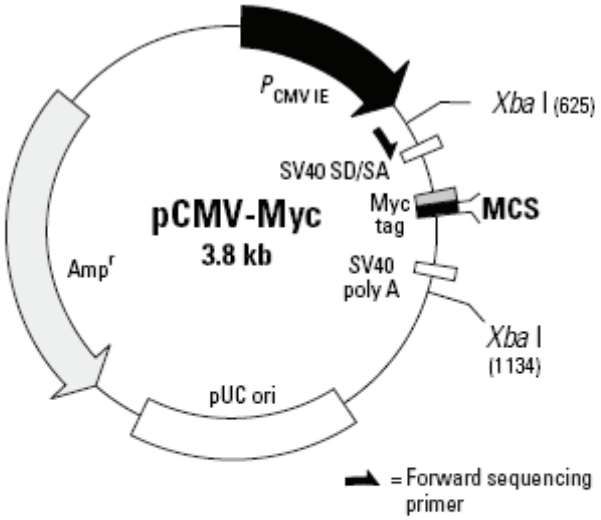
A.2 List of Tables and Figures

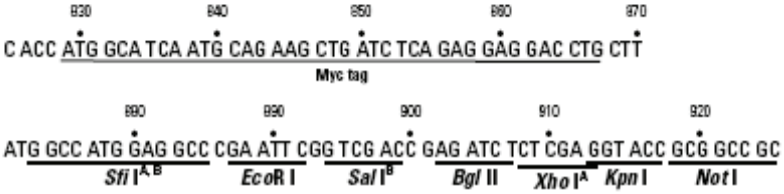
A.2.1 List of Tables

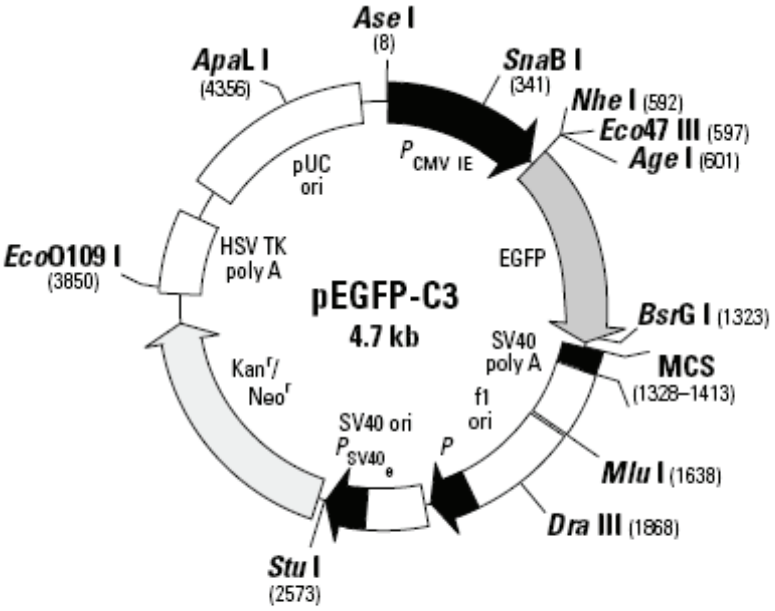
A.2.2 List of Figures

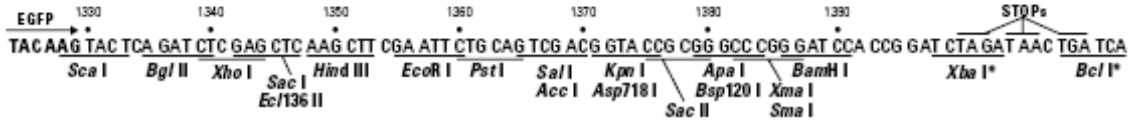
A.3 Vector maps

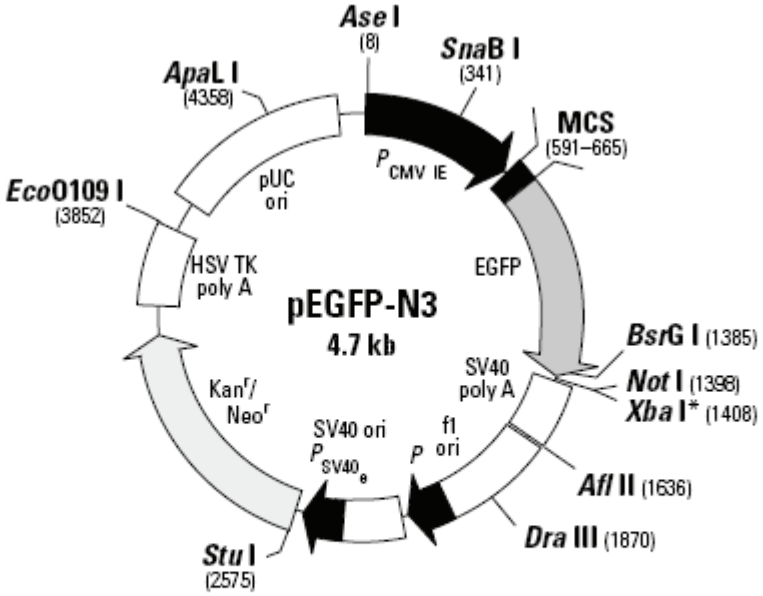


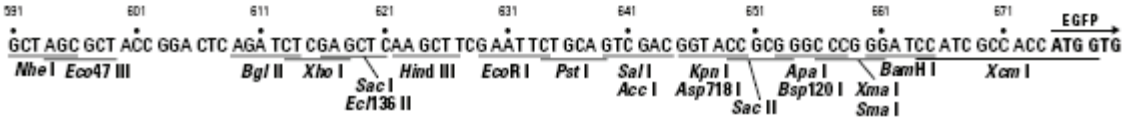





Ase I
(8)
ApaL I
(4356)
SnaB I
(341)
Nhe I (592)
Eco47 III (597)
Age I (601)
P CMV IE
pUC
ori
EGFP
HSV TK
poly A
EcoO109 I
(3850)
pEGFP-C3
4.7 kb
BsrG I (1323)
SV40
poly A
MCS
(1328–1413)
Kanr/
Neor
f1
ori
SV40 ori
P SV40e
P
Mlu I (1638)
Dra III (1868)
Stu I
(2573)


EGFP
1330
1340
1350
1360
1370
1380
1390
STOPs
TAC AAG TAC TCA GAT CTC GAG CTC AAG CTT CGA ATT CTG CAG TCG ACG GTA CCG CGG GCC CGG GAT CCA CCG GAT CTA GAT AAC TGA TCA
Sca I
Bgl II
Xho I
Sac I
Ecl136 II
Hind III
EcoR I
Pst I
Sal I
Acc I
Kpn I
Asp718 I
Sac II
Apa I
Bsp120 I
Xma I
Sma I
BamH I
Xba I*
Bcl I*


Ase I
(8)
SnaB I
(341)
ApaL I
(4358)
MCS
(591–665)
P CMV IE
pUC
ori
EcoO109 I
(3852)
EGFP
HSV TK
poly A
pEGFP-N3
4.7 kb
BsrG I (1385)
Not I (1398)
Xba I* (1408)
SV40
poly A
Kanr/
Neor
SV40 ori
P SV40e
f1
ori
P
AflII (1636)
Dra III (1870)
Stu I
(2575)


591
601
611
621
631
641
651
661
671
EGFP
GCT AGC GCT ACC GGA CTC AGA TCT CGA GCT CAA GCT TCG AAT TCT GCA GTC GAC GGT ACC GCG GGC CCG GGA TCC ATC GCC ACC ATG GTG
Nhe I
Eco47 III
Bgl II
Xho I
Sac I
Ecl136 II
Hind III
EcoR I
Pst I
Sal I
Acc I
Kpn I
Asp718 I
Sac II
Apa I
Bsp120 I
Xma I
Sma I
BamH I
Xcm I

A.4 List of Manufacturers

A.4.1 Chemicals and consumables

Item	Provider
Acrylamide, 30 % (Acrylamid / Bisacrylamid 37, 5:1)	Bio-Rad
Agarose (SeaKem)	Cambrex
Ammoniumpersulfate (APS)	Serva
Ampicillin	Roth
Amersham Hyperfilm ECL	GE healthcare
Blotting paper	Whatman
Bovine Serum Albumin (BSA fraction V)	PAA laboratories
Bromophenol blue	Sigma
Cover-slips	Starstedt
DAPI (4´-6-Diamidino-2-phenylindole dihydrochloride)	Fluka
Disposable pipette tips	Sarstedt
Double distilled water (dd H_2O)	UltraClear UV+ system
DMEM (Dulbecco's modified Eagle's medium)	Invitrogen
DpnI endonuclease	Fermentas
DTT (Dithiothreitol)	Fermentas
Embedding media	Fluka
Enhanced ChemiLuminescence (ECL)	Amersham Biosciences
Epi-cups (1.5 mL /2 mL)	Eppendorf
Ethanol, 99.8 %	Roth
Ethylendiamintetraacetatic acid (EDTA)	Sigma
Ethidium Bromide	Sigma
ExGen 500	Fermentas
Falcon tubes (15 and 50 mL)	Sarstedt
Fetal calf serum (FCS)	Invitrogen
Gentamycin, 50 mg/mL	PAA laboratories
Glycerol, 86-88 %	Riedel-deHaen
Glycine	Riedel-deHaen
Hydrochloric acid (HCl)	Fischer Scientific
Iodoacetamide	Sigma
Isopropanol (99.7 %)	Roth
Kanamycin (10,000 mcg/mL)	Gibco
Magnesium dichloride (Hexahydrate, 99 %)	Acros Organics
Mass Ruler DNA Ladder Mix (80-10000 base pairs)	Fermentas
Methanol (99.6 %)	Fluka
Microcentrifuge tubes, 1.5 mL	Sarstedt
N-Ethylmaleimide	Sigma
NP-40	Sigma
Parafilm	American National Can
Pasture pipette (Glass)	Brand
Pfu DNA polymerase	Fermentas
Piperazine-1, 4-bis (2-ethanesulfonic acid) (PIPES)	Roth
Phenylmethylsulfonyl fluoride (PMSF)	Sigma
Potassium chloride	Merck
Prestained protein molecular weight marker	MBI Fermentas
Protease inhibitor Cocktails tabs EDTA free	Roche
PVDF membrane Hybond P	Roth
Restriction endonucleases (used in this project)	Fermentas
Serological pipettes (5, 10 and 25 mL)	Sarstedt
Skimmed milk	Heirler
Sodium chloride	Fluka
Sodium deoxycholate	Merck
Sodium dodecyl sulfate (SDS)	Biomol
Sodium hydrogen carbonate	Riedel-deHaen
Sodium hydroxide	Fluka
Sucrose	Acros Organics
T4 DNA ligase	Fermentas
Taq DNA polymerase	Fermentas
TEMED (Tetramethylendiamine)	Roth
Tisuue cell culture plates (10 cm)	Sarstedt
Tisuue cell culture plates (6, 24 and 96 well)	Greiner
Tris-(hydoxymethyl-) aminomethane (Tris)	Sigma
Triton X-100	USB

Trypsin/EDTA	PAA Laboratories
Tryptone	Fluka
Tween 20 (Polyethylene-sorbitan monolaurate)	Sigma
UV-cuvette	Brand
Yeast extract	Fluka
X-Ray Hyperfilm ECL	Amersham
Tissue culture plates	Sarstedt

A.4.2 Devices

Item	Manufacturer
Centrifuge and rotor (Benchtop minispin; 5804R & 5810 R)	Eppendorf
Centrifuge rotor (Sorvall evolution RC)	Sorvall
Electrical power supply (SDS-PAGE)	Bio-Rad
Electrical power supply (Western blotting)	Bio-Rad
Electrophoresis system (DNA)	Bio-Rad
Electrophoresis system (Protein)	Bio-Rad
Heating block	Eppendorf
Ice-machine	Scotsman
Incubator (Cell culture)	Heraeus
Incubator (Bacteria)	Binder
Laminar air flow system	Holten
Magnetic stirrer	IKA
Microtiter plate photometer (Sunrise-Basic)	Tecan
Microtiter plate photometer (Multiscan Ascent)	Thermo Labsystems
Microwave	Bosch
Neubauer chamber	Marienfeld
pH meter	Inolab
Photometer	Eppendorf
Pipettes	Eppendorf
Pipette boy	Eppendorf
Pipette (multi-channel)	Eppendorf
Refrigerator (4ºC)	Liebherr
Refrigerator (-20ºC)	Vestfrost
Refrigerator (-80ºC)	Heraeus
Shaker (Vibramax 100)	Heidolph
Thermocycler (Master cycler)	Eppendorf
UV-camera system	Herolab
Vortex	IKA
Water-bath	GFL
Weighing Balance	Sartorius
Weighing Balance (fine measurement)	Sartorius
Western blotting device (tank)	Bio-Rad
X-ray film developing machine (Curix image processor)	Agfa
X-ray film cassette	Amersham

Acknowledgements

Dr. Frank Dietz, thank you for giving me the opportunity to work on this project and also for your expert supervision over the entire period of the thesis. I thank you for friendly working environment and thought stimulating discussions. You have been a great mentor, your work style and dedication will keep on inspiring me throughout my life. I specially would like to show appreciation for your excellent, quick and prompt reviewing during the writing of the manuscripts and the thesis. I owe a lot to your perseverance regarding the project and keeping faith in my abilities.

Prof. Dr. Sørge Kelm, thank you for reviewing my thesis work and giving me the opportunity to work in your working group. I thank you for excellent reviewing of the manuscripts. I am also grateful for all the guidance and support you have showed during entire period of the thesis. I highly appreciate the fact that you were always eager to expand the boundaries of work in different quadrants, encouraging the use of various methods and tools, which in turn benefited the project and also my personal development. **Prof. Dr. Reimer Stick**, thank you for showing interest in the project and reviewing my thesis work. **Dr. Kathrin Mädler**, thank you for being willing to be an examiner for my thesis.

I thank **Dr. Sebastian Franken**, Bonn, Germany and **Dr. Rainer Niedenthal**, Hannover, Germany for their collaboration and **Dr. Jens Harder**, Max Planck Institute for Marine Microbiology, Bremen, Germany for providing access to the DNA sequencing facility. Also, I thank **Prof. Dr. Ralf Dringen**, Bremen, Germany for allowing use of experimental facilities in his laboratory. I personally would like to mention gratitude of the International office, University Bremen, Germany for providing additional financial assistance.

Mrs. Nazilla Isakovic and **Mrs. Petra Berger**, thank you for your excellent technical assistance. **Mr. Thaddeus T. Gbem**, thank you for all the discussions and also for your encouragement and friendship. I like to mention gratitude towards **Mrs. Sabine Limberg** for always helping with administrative work. I thank all the members of AG Kelm for providing friendly working environment. I also thank members of AG Mädler, with special mention to Dr. Federico Paroni and Mr. Amin Ardestani, for friendly atmosphere and interactions. I thank Mr. Shaunak Khandekar and Mr. Thaddeus T. Gbem for proof-reading my thesis.

I am grateful and indebted for the support and motivation received from my parents, parents-in-law, and my family throughout my studies. Last; but certainly not the least; I thank my wonderful wife, Teja, for all her support, encouragement, perseverance and love.

Erklärung

Gemäß § 6 der Promotionsordnung der Universität Bremen für die mathematischen, natur- und ingenieurswissenschaftlichen Fachbereiche vom 14 März 2007 versichere ich, dass

1. die Arbeit ohne unerlaubte fremde Hilfe angefertigt wurde

2. keine anderen als die angegebenen Quellen und Hilfsmittel benutzt werden

3. die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht wurden.

Bremen, 27 April 2010

Ketan Thakar