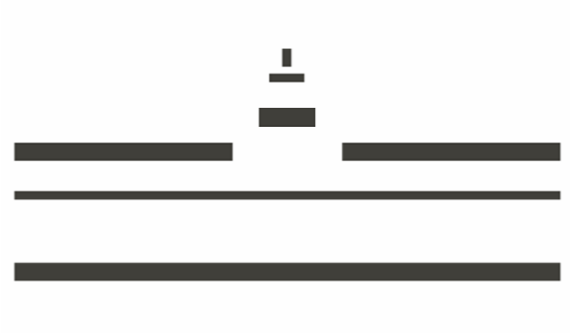


Molecular functional characterization of the human biglycan gene 5'-flanking region

Dipl. Biol. Boris Schmitz

-2010-

Molecular functional characterization of the human biglycan gene 5'-flanking region

Molekular funktionelle Charakterisierung der 5'-flankierenden Region des humanen Biglykan Gen-Promotors

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften im Fachbereich Biologie

der Mathematisch-Naturwissenschaftlichen Fakultät

der Westfälischen Wilhelms-Universität Münster

vorgelegt von

Dip. Biol. Boris Schmitz

aus Remscheid

-2010-

Dekan:	Univ.-Prof. Dr. Christian Klämbt
Erster Gutachter:	Univ.-Prof. Dr. Dirk Prüfer
Zweiter Gutachter:	Univ.-Prof. Dr. Christian Klämbt
Datum der mündlichen Prüfung:	13.12.2010
Datum der Promotion:	17.12.2010

The path is hard,

but will be amply rewarding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABREVIATIONS

Acc#	Accession number
AP-1	Activator Protein-1
AS	Antisense Strand
BGN	Biglycan
CBP	CREB-Binding Protein
ChIP	Chromatin Immunoprecipitation
cAMP	Cyclic Adenosine Monophosphate
cDNA	complementary Desoxyribo Nucleic Acid
COS-7	African green monkey kidney fibroblast-like cell line
CRE	cAMP-Response Element
CREB	cAMP Response Element-Binding protein
CAD	Coronary Artery Disease
CVD	Cardiovascular Disease
DBTSS	Data Base of Transcription Start Sites
DPE	Downstream Promoter Element
Poly (dI●dC)	Polydeoxy (Inosinate-Cytidylate) Acid
EA.hy926	Human vascular endothelial cell line
ECM	Extracellular Matrix
ECTIM	Etude Cas-Témoins de I'infarctus du Myocarde
EMSA	Electrophoretic Mobility Shift Assay
ETS	E26 Transformation-specific Sequence
GATA1	GATA-motif binding protein 1
GWA	Genome Wide Association
HapMap	Haplotype Map
HEK293T	Human Embryonic Kidney cells expressing the large T-antigen of SV40
ICAM-1	Intercellular Adhesion Molecule-1
IL1-β	Interleukin 1-beta
Inr	Initiator Element

LDL	Low-Density Lipoprotein
MI	Myocardial Infarction
MolHap	Molecular Haplotype
MolProMD	Münster Molecular Functional Profiling for Mechanism Detection
PAGE	Polyacrylamide Gel Electrophoresis
PDGF	Platelet-Derived Growth Factor
PKA	Protein Kinase A
PKC	Protein Kinase C
PMA	Phorbol 12-myristate 13-acetate
PU.1	Hematopoietic transcription factor
n.s.	not significant
oxLDL	oxidized LDL
RACE	Rapid Amplification of cDNA Ends
RLU	Relative Light Unit
RP27	Ribosomal Protein 27
SDS	Sodium Dodecyl Sulphate
SMC	Smooth Muscle Cell
SNP	Single Nucleotide Polymorphism
SP1	Specificity Protein 1
SS	Sense Strand
SSCP	Single-Strand Conformation Polymorphism
TBP	TATA-box Binding Protein
TF	Transcription Factor
TFBS	Transcription Factor Binding Site
TGF-β1	Transforming Growth Factor-beta1
TGFBR	Transforming Growth Factor-beta1 Receptor
THP-1	Human acute monocytic leukemia cell line-1
TSS	Transcription Start Site
UTR	Untranslated Region
VCAM-1	Vascular Cell Adhesion Molecule-1

VSMC	Vacular Smooth Muscle Cell
wt	wild type

ABSTRACT

Atherosclerosis is a multifactorial and polygenic disease. Origin and progression of atherosclerosis involve multiple biological systems. The extracellular matrix protein biglycan (BGN) is involved in the pathogenesis of atherosclerosis at different stages. BGN affects the constitution and stability of the fibrous cap of the atherosclerotic plaque, mediates binding of low-density lipoproteins to the artery wall and exerts proinflammatory actions. We analyzed the transcriptional regulation of the human *BGN* gene, also with respect to individual promoter allelic constellations. In transient transfection experiments of the *BGN* 5'-flanking region and 5'-UTR, we identified an 1025 bp portion with sufficient transcriptional activity in endothelial (EA.hy926) and monocytic (THP-1) cell lines. Screening of 1198 bp of the promoter region in 57 individuals with cardiovascular disease (CVD) led to the validation of three genetic variants: G-578A (rs11796997), G-151A (Brand, unpublished data) and G+94T (rs5945197). Subcloning and resequencing revealed three common MolHaps: BGN-MolHap1 [G^{-578}-G^{-151}-G^{+94}], BGN-MolHap2 [G^{-578}-A^{-151}-T^{+94}] and BGN-MolHap3 [A^{-578}-G^{-151}-G^{+94}]. Introduction of either MolHap2 or MolHap3 in a 1025 bp construct resulted in a significant decrease (all P-values <0.05) of transcriptional activity in both EA.hy926 and THP-1 cell lines. To identify the positions of *cis*-regulatory elements in the 5'-flanking region, we generated serial deletion constructs. Stimulation of THP-1 cells with transforming growth factor-beta1 (TGF-β1) increased transcriptional activity of wild type *BGN* promoter fragments up to 3-fold (all P-values <0.01). To identify *trans*-acting transcription factors, we performed co-expression, chromatin immunoprecipitation (ChIP), and EMSA experiments. Transcription factor SP1 was shown to activate transcriptional activity of promoter fragments up to 4-fold compared to mock control (all P-values <0.001) and physical SP1 interaction was confirmed in ChIP experiments for alleles G-151 and G+94 in EA.hy926 cells. EMSA experiments revealed binding of c-FOS to the 5'-UTR position G+94T. TGF-β1 stimulation of THP-1 nuclear extracts enhanced SP1 interaction with position G-578A. In THP-1 cells, ETS family member PU.1 bound the major allele G-578 with higher affinity (4-fold) compared to the minor allele -578A. We conclude that *BGN* gene expression is under the control of activating transcription factor SP1. TGF-β1 reinforces SP1 binding and enhances transcriptional activity of the *BGN* promoter. The polymorphic position G+94T reside within a *cis*-active promoter element where AP-1 complex formation was observed. Monocyte-specific regulation of *BGN* expression is also controlled by ETS TF PU.1.

1 INTRODUCTION

1.1 Cardiovascular Disease (CVD)

CVD include hypertension, coronary heart disease (CHD), cerebrovascular disease, peripheral artery disease, rheumatic heart disease, congenital heart disease, and heart failure. Globally, CVD is the leading cause of death for women and men. In 2005 an estimated 17.5 million people died from CVD, representing 30% of all global events (WHO). Sixty-one percent of cardiovascular deaths occur due to specified risk factors, which include high blood pressure (~54% of deaths in European countries), high body mass index, high cholesterol, high blood glucose, unhealthy diet and physical inactivity. Above all, family history of CVD strongly and independently determines future CVD risk in individuals (Myers et al., 1990), underlining a major impact of genetic predisposition.

1.2 Arteriosclerosis

Arteriosclerosis represents the pathophysiological basis of CVD. The term literally refers to 'hardening of the arteries', which resembles early reports on its observed pathophysiology. Terminology and classification of arteriopathic subcategories are still part of ongoing debates (Fishbein & Fishbein, 2009). Commonly accepted, arteriosclerosis is a systemic disease with particularly noticeable manifestations in the medium and large arteries and a strong inflammatory background. Atherosclerosis represents the most important form of arteriosclerosis, describing loss of elasticity of affected arteries due to atheromatous plaque formation (fig. 1).

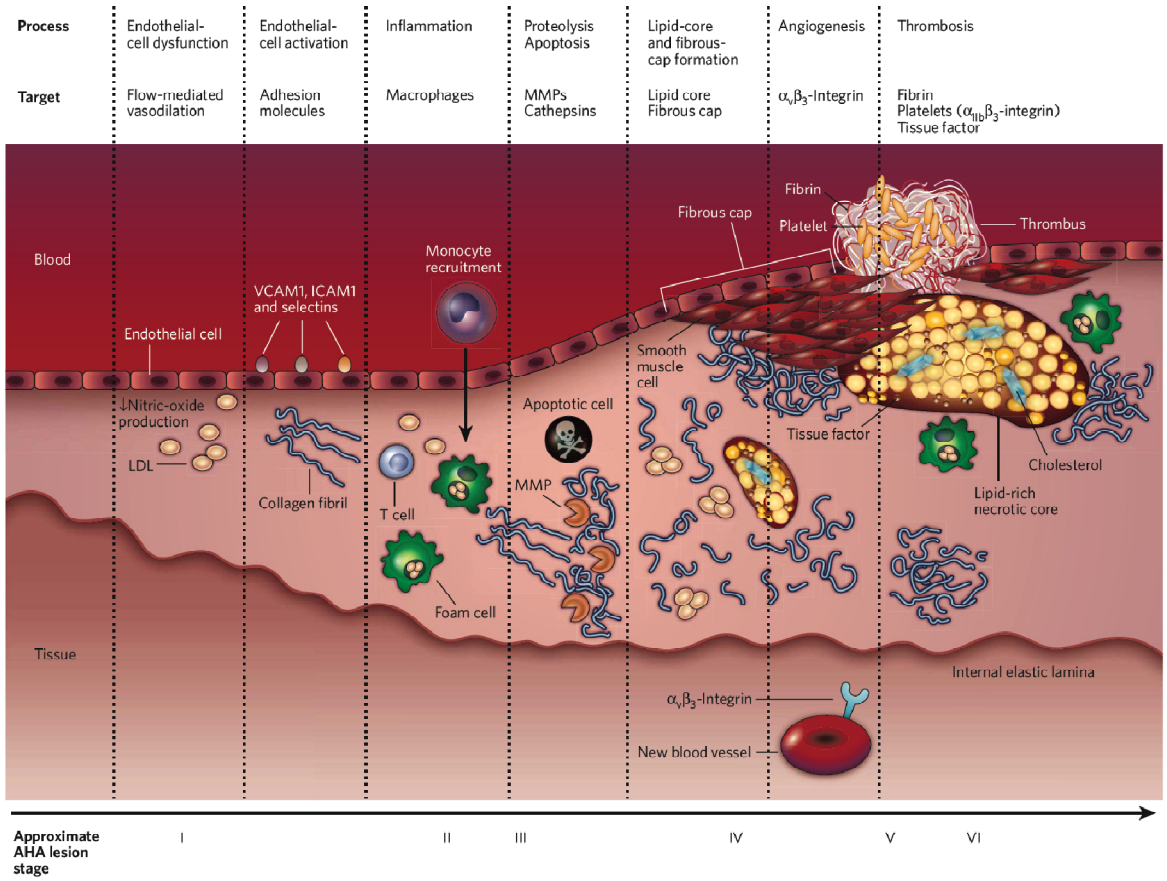


Figure 1: Development of atherosclerotic lesions.

The development of an atherosclerotic lesion is shown from left to right, from a normal blood vessel to a vessel with an atherosclerotic plaque and superimposed thrombus. Stages I to VI (below) are depicted according to the American Heart Association. The process representing specific stages of the lesion is shown at the top of the figure. (I) Lipoprotein particles, especially LDL, enter the arterial wall and undergo modification, including oxidation. The endothelial cells dysfunction is marked by inhibition of nitric oxide production. Oxidised LDL particles lead to endothelial cell activation with subsequent expression of adhesion molecules (ICAM-1, VCAM-1 and selectines) and inflammatory cytokines. (II) Monocytes are recruited into the subendothelial space, guided by chemokines. Differentiated into macrophages, they take up oxLDL and transform into lipid-laden foam cells. Infiltrating T-cells produce adhesion molecules, cytokines and chemokines leading to an ongoing immune activation. (III-IV) Apoptosis of foam cells leads to necrotic core formation. VSMC migrate into the developing plaque, forming a fibrous cap. (V-VI) Formation of stable or unstable plaques with potential rupture and thrombus formation. Figure adapted with permission from Z. Fayad (Sanz & Fayad, 2008).

1.2.1 Pathophysiology of atherosclerosis

Atherosclerosis is a progressive, multifactorial, and proinflammatory disease characterized by the accumulation of lipids in distinct regions of the arterial wall designated as atherosclerotic plaques. Changes of the arterial wall due to atherosclerotic processes begin in early childhood and is a slow and silent process.

Clinical manifestation with noticeable symptoms is not commonly detectable before 30 years of age, depending on individual predisposing factors. Despite other potential inducers of atherosclerosis pathophysiology, infiltration and retention of lipids from blood into the intima of the arterial wall is a pivotal process in disease development and progression (fig. 1). Elevated low-density lipoprotein (LDL) plasma levels result in binding of LDL to proteoglycans in the subendothelial matrix, where modification through oxidative and enzymatic (i.e. myeloperoxidase, lipoxygenases and others) processes takes place yielding oxLDL (Skålén et al., 2002). LDL thereby passively diffuses through endothelial cell junctions, with preference to regions of arterial branching or curvature, were no particular orientation of the cells in blood flow direction is observed. Activation of endothelial cells by phospholipids released from LDL may result in expression of several adhesion molecules such as VCAM-1, ICAM-1 and selectins on the endothelial surface (Steinberg, 2009). Upon interaction of adhesion molecules with carbohydrate ligands on the surface of blood monocytes, rolling, adhesion, and subsequently transmigrate of the monocyte into the subendothelial space takes place in response to different chemokines. Macrophage colony-stimulating factor (M-CSF) induces proliferation and differentiation into macrophages (Yan & Hansson, 2007). This process is associated with upregulation of pattern-recognition receptors including scavenger receptors (Hofnagel et al., 2007) which mediate the uptake of oxLDL leading to foam cell formation and Toll-like receptors initiating inflammatory activation (Hansson, 2009). Native LDL is not taken up by macrophages very rapidly and LDL modification is crucial for early lesion formation. Anti-atherogenic processes are thereby interfered by oxLDL since it inhibits the production of nitric oxide, a strong mediator of vasorelaxation. This early stage in atherosclerosis pathophysiology, characterized by the subendothelial accumulation of lipid-laden macrophages, is termed the fatty streak. During the formation of the atherosclerotic plaque, the shoulder regions determine its future stability to a great extend. Here, T-cells of the lesion, with domination of CD4+ cells over CD8+ cells, are concentrated in clusters (Robertson & Hansson, 2006). On ongoing immune activation in these regions is marked by MHC class II-expressing macrophages and dendritic cells

(Jonasson et al., 1985). Infiltrating T-cells are activated by local antigens such as oxLDL components or heat shock protein 60/65 (Robertson & Hansson, 2006), resulting in the expression of various cytokines and differentiation into Th1 effector cells. The prototypic Th1 cytokine interferon-γ (IFN-γ) promotes macrophage and endothelial activation with production of adhesion molecules, cytokines and chemokines and secondary increases the expression of tumor necrosis factor-α (TNF-α) and IL-1 (Hansson, 2001). Targeted deletion of IFN-γ or its receptor reduces disease progression, while recombinant IFN-γ accelerates lesion formation. The resulting local concentration of synergistic cytokines induce the proliferation and migration of VSMC, which in turn concentrate ECM components within the advanced atherosclerotic lesion. The advanced lesion is mainly composed of a necrotic core with accumulation of lipid-laden macrophages and extracellular lipid droplets derived from dispersed foam cells. A cap of VSMC and a collagen rich matrix is formed, characterizing the composition of the fibrous cap and its mechanical strength.

Acute coronary events depend principally on the composition and vulnerability of the plaque and the status of a plaque from stable to unstable can change within a very short period of time under certain pathophysiological conditions. Underlying mechanisms are still under investigation, with a focus on endothelial (Hirschi et al., 2008) and smooth muscle (Sata et al., 2002) progenitor cells trafficking in the vascular wall (for a review see Hristov & Weber, 2008). Increasing numbers of inflammatory cells may contribute to unstable plaque formation by weakening of the plaque shoulder regions, making the cap more vulnerable to shear stress, leading incidentally to plaque rupture with subsequent thrombus formation by adhering platelets and fibrin crosslinks. Plaque erosion is another cause of acute coronary thrombosis. Eroded plaques are rich in SMC and proteoglycans, with relatively few inflammatory cells. The main cap proteoglycans, versican, BGN, and decorin (DCN), in combination with hyaluronan accumulate in topographically distinct patterns. Not only do these molecules contribute to plaque burden, but influence the biomechanical properties of vascular lesions and the ability of plaques to resist rupture. Stable plaques are thereby marked by increased accumulation of versican and BGN whereas plaque erosion is characterized by aggregation of hyaluronan and absence of BGN (Kolodgi et al., 2002)

1.3 BGN as a candidate for CVD pathophysiology

BGN, or proteoglycan-I (PG-I), has been first described in 1983 by Fisher et al. (Fisher et al., 1983). The human *BGN* represents a single copy gene and was mapped to the long arm of the X-chromosome (Xq28). *BGN* consists of eight exons, including exon one encoding the 5'-untranslated region of the mRNA, seven introns and spans over 9.5 kbp of DNA. It is transcribed into nine transcripts by the alternative use of TSS and multiple splicing events. *BGN* is broadly expressed in the human body with high mRNA levels in lung, placenta, liver, retina and the heart. Altered *BGN* expression has been reported under several pathological conditions including cancer (Wong et al., 2009; Valladares et al., 2006), hypertension (Sardo et al., 2009) and atherosclerotic plaque formation (Riessen et al., 1994; Kolodgie et al., 2002; Adiguzel et al., 2009). *BGN* expression is influenced by a number of cytokines. TGF-β1 has been identified as a positive regulator of *BGN* expression (Lijnen et al. 2000; Ungefroren et al., 2003; Burch et al., 2010) and a negative feedback loop for BGN regulating TGF-β1 activity has been suggested (Ruoslahti & Yamaguchi, 1991). TNF-α decreased *BGN* steady-state mRNA levels (-62%) and the *BGN* core protein gene transcription rate (-18%) in human chondrocytes (Dodge et al., 1998). The combination of IFN-γ and TNF-α resulted in a potentiation of the observed inhibitory effect. The important role of BGN in circulating monocytes has been emphasized repeatedly and a decrease in BGN expression has been reported after incubation with glucocorticoids (Kimoto et al., 1994). Treatment with angiotensin II has been shown to increase BGN mRNA expression in patients' blood monocytes (Sardo et al., 2009).

BGN is one of at least nine small leucine-rich proteoglycans (SLRP) which are secreted extracellular matrix components with different localization, expression and function. Grouped into three different classes (Iozzo, 1999), all of them can be described as either glycoproteins containing N-linked oligosaccharides or as proteoglycans containing chondroitin/dermatan sulfate or keratin sulfate chains (fig. 2). Class I comprises DCN and BGN which show a high homology (57%) and have been proposed to have evolved from an ancestral gene by gene duplication (Krusius et al., 1986). BGN and DCN are the only SLRP members that contain a highly conserved pro-peptide which might serve as a recognition signal for the xylosyltransferase. Both contain ten leucine-rich repeats (LRR) flanked by cystein-rich regions (Krishnan et al., 1999) and either one (DCN) or two (BGN) chondroitin/dermatan

sulfate chains at their N-terminal domain. Other members of the SLRP family are fibromodulin (FMOD), lumican (LUM), keratocan (KERA), PRELP and osteomodulin (OMD) in class II as well as epiphycan (DSPG3) and osteoglycin (OGN).

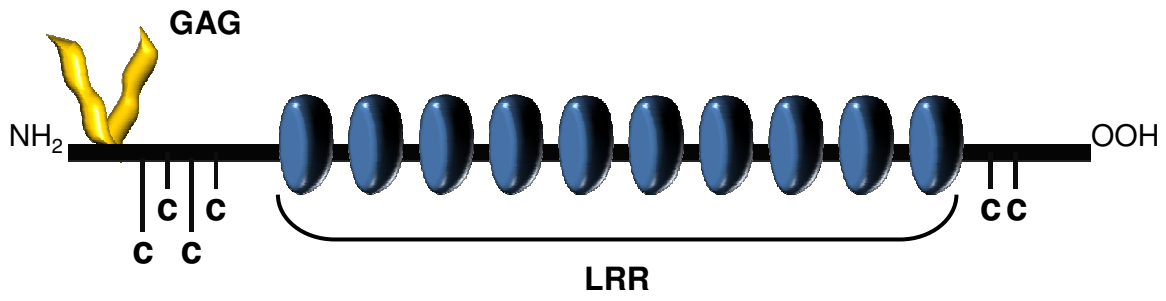


Figure 2: Schematic representation of the BGN protein, indicating major domains.

N-terminal: two glycosaminoglycan (GAG) chains (yellow), C: cystein residues, 10 leucine-rich repeats (LRR, blue). BGN self-aggregates into dimers and hexamers under physiological conditions. With modifications from Krishnan et al., 1999.

The versatile role of BGN in pathophysiological processes has been extensively studied in animal models. Heegaard et al. (Heegaard et al., 2007) reported a vascular phenotype in *BGN*-knock out (ko) mice. BGN-deficiency caused aortic rapture across intima and media and dissection of the aortic wall. This observation was predominantly male-specific, since 50% of male ko mice died within three month of age but female mice showed similar phenotypes as wt. Corsi and colleagues (Corsi et al., 2002) highlighted the impact of *BGN*-deficiency on collagen fibrillogenesis. The group observed structural abnormalities in collagen fibril deposition in bone, dermis, and tendon of male *BGN*-ko mice. A very comprehensive study was presented by Westermann et al. (Westermann et al., 2008), investigating the role of BGN during cardiac ECM remodeling and cardiac hemodynamics after MI. They observed increased mortality, high frequency of cardiac ruptures, and aggravated cardiac failure in male ko animals after experimental MI due to perturbed collagen remodeling in the infarct scar tissue. In wt mice, *BGN* expression was strongly upregulated after MI, peaking in parallel to collagen content and was found in the infarct scar and border zones using

immunohistochemistry. The dual nature of BGN as a signalling molecule and a crucial proinflammatory factor was highlighted by Schaefer and colleagues (Schaefer et al., 2005). *BGN*-ko mice had a survival benefit compared with wt animals in induced sepsis due to decreased inflammatory response resulting in less organ damage. In addition, they provide evidence that macrophages are capable of producing BGN upon stimulation by lipopolysaccharide-induced proinflammatory factors. In a previous report the group observed concurrence of BGN overexpression and increased numbers of infiltrating cells in a model of renal inflammation (Schaefer et al., 2002).

1.4 Polymorphic structure of the *BGN* gene

In initial studies (Brand, unpublished data), the 5'-flanking region, exons, introns, and the 3'-flanking region of the *BGN* gene were scanned for alterations in genomic DNA from 95 patients of the ECTIM study. ECTIM (Etude Cas-Témoins de l'Infarctus du Myocarde) is a study of patients with MI (Cambien et al., 1992) from regions covered by the World Health Organization's Monitoring Trends and Determinants in Cardiovascular Disease (MONICA) registers and of control subjects (cases [n=988, mean±SD age 55.8±8.1 years, 26.2% women], controls [n=949, mean±SD age 56.6±8.3, 27.0% women]).

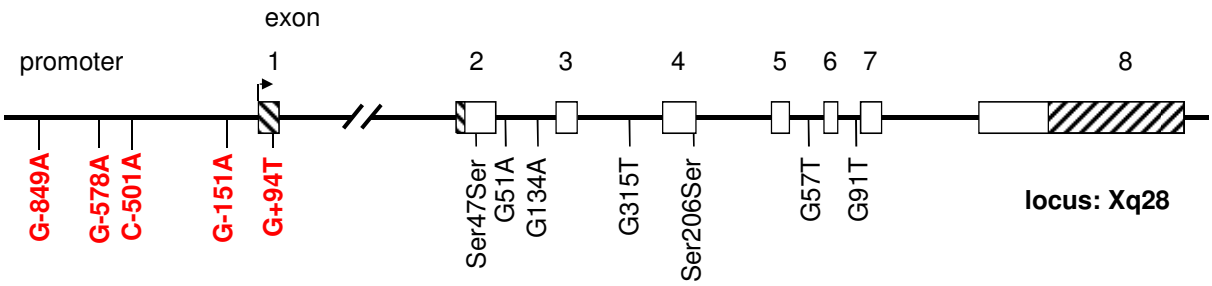


Figure 3: Schematic representation of the *BGN* gene.

The *BGN* gene consists of eight exons, shown in blank bars and seven introns. The single copy gene is located on the long arm of the X-chromosome at the proximal region Xq28. The untranslated exonic regions (5' and 3') are indicated by shaded bars. Identified SNPs with relative position are depicted below. Promoter SNPs are indicated in red. The major TSS is marked by an arrow (adopted from Brand, unpublished and Rüssmann, MD thesis).

Twelve SNPs were identified using SSCP analysis (fig. 3), five of which were located in intronic regions (G51A, G134A, G315T, G57T and G91T; designation according to end of upstream exon), two synonymous polymorphisms were found in exons two (Ser47Ser) and four (Ser206Ser). In the 5'-region, five SNPs were detected, one residing in the 5'-UTR of the gene. SNPs G-849A (rs56134709), G-578A (rs11796997) and G+94T (rs5945197) were already listed at NCBI (National Center for Biotechnology Information), SNP G-151A and C-501A had not been reported before. SNPs G-849A was detected one and C-501A twice within the 95 ECTIM DNA samples.

In a further analysis (Rüssmann, MD thesis), three MolHaps have been suggested to be formed by variants G-578A, G-151A and G+94T, based on SSCP results only (sensitivity of variant detection by SSCP <80%; fig. 4).

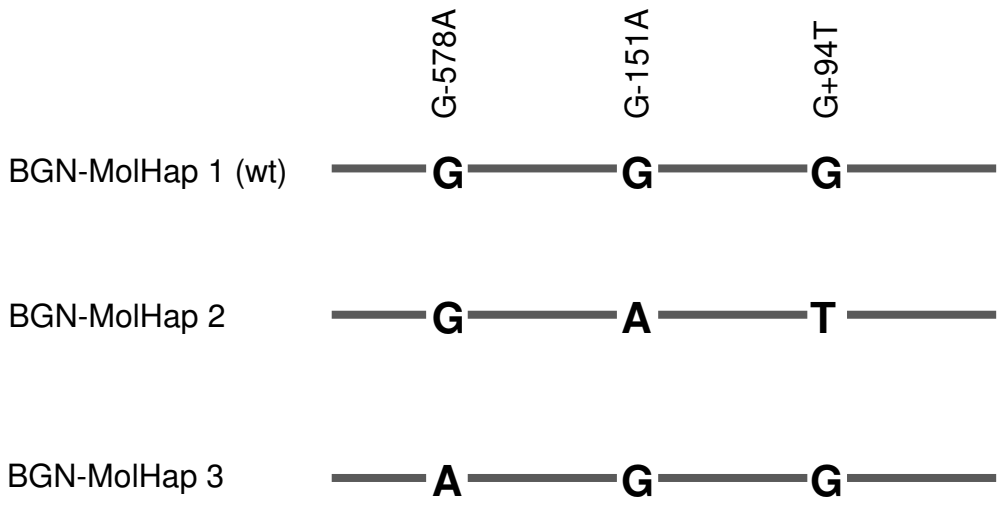


Figure 4: Postulated MolHaps of the human *BGN* 5'-flanking region.

Estimated frequencies among ~72% of analysed and evaluable DNA samples: BGN-MolHap1: ~19%, BGN-MolHap2: ~46% and BGN-MolHap3: ~36% (adopted from Rüssmann, MD thesis).

1.5 TGF-β1 cytokine signalling pathway

TGF-β1 is a multifunctional peptide that controls proliferation, differentiation, and other functions in different cell types. TGF-β1 binds to membrane receptor serine/threonine kinases which form a bi-dimeric receptor complex (fig. 5), consisting of two pairs of subunits known as receptor type I (TGFBR1) and type II (TGFBR2). A membrane-anchored proteoglycan, designated type III receptor (TGFBR3), contributes to this process by capturing and presentation of TGF-β1 to the signalling receptors I and II. Upon TGF-β1 binding, TGFBR1 phosphorylates receptor-bound SMAD2/3 TF, releasing them into the cytoplasm with subsequent translocation into the nucleus. As a common partner, SMAD4 serves as a mediator of transcriptional activation (for a comprehensive review see Massagué & Gomis, 2006). Animal models provided evidence that TGF-β1 signalling pathway plays a pivotal role for ECM formation. In rats, administration of TGF-β1 neutralizing antibodies suppressed the expression of genes ECM components (Sharma et al., 1996). Transgenic mice overexpressing TGF-β1 predominantly in odontoblasts showed increase in dentin ECM components and abnormal deposition in the dental pulp (Thyagarajan et al., 2001). SMAD2, SMAD3, and SMAD4 are highly expressed in macrophage-derived foam cells of fatty streaks (Kalinina et al., 2004). SMC in these lesions do not express SMAD and TGF-β1 signalling is impaired with subsequent suppression of ECM components. In contrast, SMC in stable fibrous plaques express high levels of SMAD and ECM proteins. TGF-β1 is therefore a potential therapeutic target and it has recently been suggested to block TGF-β signal transduction and downstream proteoglycan synthesis in vascular SMC (VSMC) for the prevention of CVD using p38 MAP kinase inhibitors (Osman et al., 2008). This approach is problematic to some extend since plaque formation is an ongoing process with early onset in the development of atherosclerosis and TGF-β signal inhibition could support formation of unstable plaques. Selective local restoration of TGF-β responsiveness, leading to stabilization of atherosclerotic lesions is suggested to be a promising concept. On the other hand, findings of Suthanthiran et al. (Suthanthiran et al., 2000) indicate that TGF-β1 overexpression is a risk factor for hypertension and hypertensive complications.

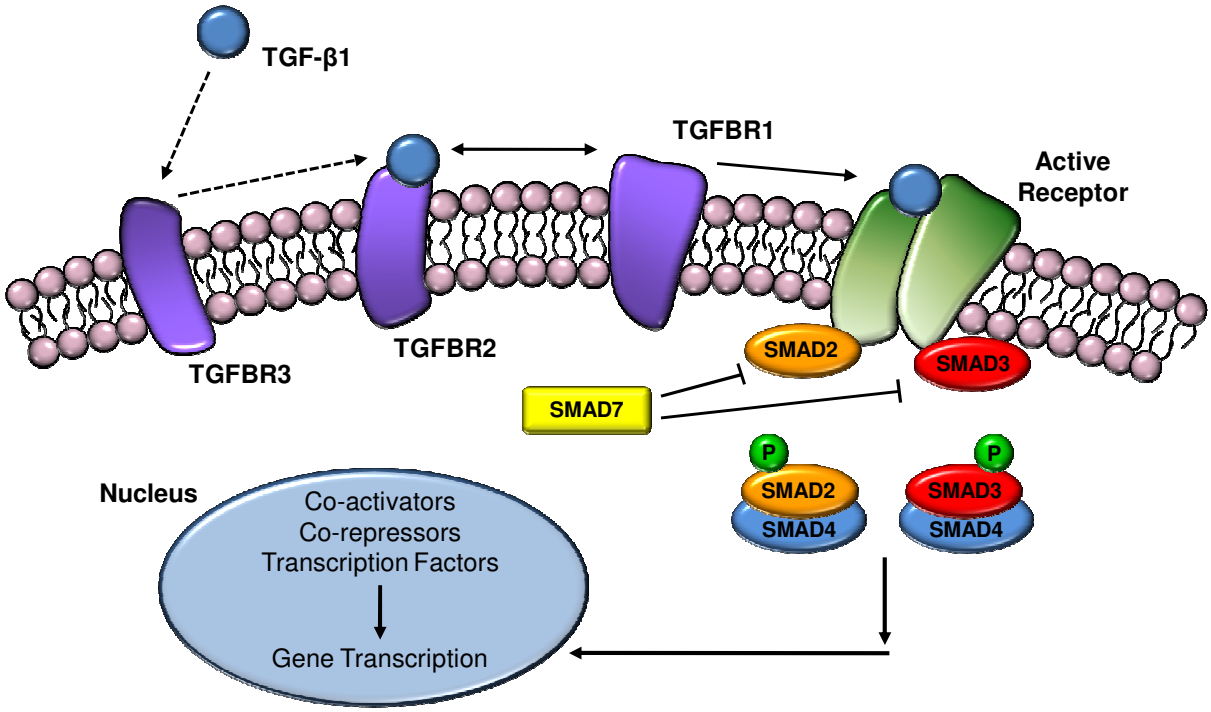


Figure 5: The TGF-β1 signalling pathway.

TGF-β1 binds to the receptor TGFBR2 upon capturing and presentation by TGFBR3 (dotted lines). TGFBR2 recruits and phosphorylates TGFBR1 with subsequent phosphorylation (P) of SMAD2 and SMAD3. This process is inhibited by SMAD7. Activated SMAD2/3 form heterodimers with SMAD4 and translocate to the nucleus. Regulation of gene expression is coordinated by co-activators, co-repressors and other transcription factors. With modifications from Pinzani & Marra, 2001.

1.6 Gene expression control

As the Human Genome Project has been almost completed, the complexity of gene expression control has become exposed. Only 2–3% of the human genome encode for proteins, far less than originally apparent. The observed phenotype plasticity distinguishing the various types of cells throughout the human body are rather the result of differences in gene expression than a great variety of different coding genes. Orchestrated cellular proliferation and differentiation processes are controlled by a diversity of programs, regulating gene expression. These processes are crucial steps in maintenance of cellular homeostasis as well as in specific responses after signal transduction stimulation under pathophysiological conditions. Gene

expression can be regulated at different control levels. The most important being chromatin packaging, transcription, transcript processing, nuclear export, transcript stability, translation and protein stability/activity. Despite the diversity of potential levels of regulation, the majority of regulatory events are under transcriptional control (Villard, 2004).

1.6.1 Transcriptional control

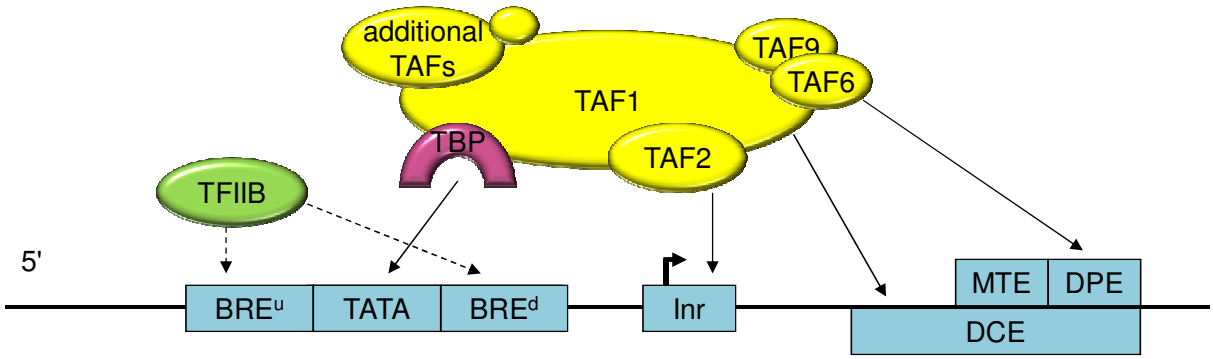
Coding genes are under the control of at least one but more likely several, alternative promoters. High-resolution maps of active promoters in the human genome have suggested an average of about 3.1 promoters per gene (Kim et al., 2005; Cheong et al., 2006), underlining the complexity of transcriptional control. In general, one promoter is located within the 5'-flanking region of a gene and alternative promoters can reside within intronic or exonic regions of the same or an adjacent gene. To date, the exact localization of promoter regions for most human genes is still unknown (Barrera & Ren, 2006). In eukaryotes transcription is controlled at different levels. Chromatin condensed by nucleosomal packaging is inaccessible for RNA Polymerase II (Pol II) and genes residing in packed regions are transcriptionally silent (Segal & Widom, 2009). Histone acetylation and gene methylation are the two main mechanisms involved in reactivation of silenced genes. Eukaryotic Pol II requires the assembly of the general transcription machinery at the promoter for initiation of gene transcription. In a next step, binding of regulatory *trans*-acting factors at *cis*-active elements occurs, which interact with the general transcription machinery. Co-regulatory factors (bridging-proteins such as p300), which do not bind DNA but interact with proteins of the general machinery and TF are involved in mediation of distal regulatory regions with the core promoter. After initiation of RNA synthesis, the transition to productive elongation is controlled as well. Pol II has been observed to stop after elongation of 20–50 nucleotides into the gene, designated as Pol II stalling within the proximal promoter regions (Muse et al., 2007). This concept has been suggested for transcripts involved in rapid signal cascades, where Poly II is pre-engaged at the promoter without active transcription awaiting signals for transcript elongation (Buratowski, 2008). Activated genes display another characteristic feature, which is the ability to switch their conformation. Interaction of promoter elements and 3'-associated factors may initiate a loop structure, retaining Pol II once a full transcription cycle has completed with no need for reassembly of the transcription machinery (Moore & Proudfoot, 2009). In addition, a connection between promoter usage and alternative splicing

has been discussed, involving the recruitment of factors with dual functions in transcription and splicing, linking the two processes (Kornblihtt, 2005). Generated and stable mRNA transcripts are targets of naturally occurring knock-down events by RNA interference (RNAi) (Fire et al., 1998).

1.6.2 The general transcription machinery

The core promoter in eukaryotes serves as a platform for the assembly of the transcription preinitiation complex (PIC). Formation of the PIC is necessary since Pol II itself lacks sequence-specific recognition ability. For promoter function, proper assembly and formation of the PIC at characteristic DNA sequences is required. Seven eukaryotic core promoter elements have been identified so far (fig. 6; Thomas & Chiang, 2006), the most prominent being the TATA box, represented by an A/T-rich sequence located 25 to 30 nucleotides upstream of the TSS. PIC formation at this motive depends on recognition by the TATA-binding protein (TBP) or the TBP subunit of the TFIID complex. A TATA box (Smale & Kadonaga, 2003) is present in only <20% of human promoters (Kim et al., 2005) and other core promoter elements, either alone or in combinations are used. Data from functional studies suggest that human housekeeping genes, growth factors and transcription factors often lack a TATA element (Zhou & Chiang, 2001). In close vicinity of the TSS (-2 to +5 bp), the initiator (Inr) is located, which is recognized by the TAF1/TAF2 components of TFIID. It is capable of directing transcription initiation alone or in conjunction with the TATA box or other core promoter elements. The downstream promoter element (DPE) is often found in human TATA-less promoters and recognized by TAF6 and TAF9 dimeric complex (Shao et al., 2005). In addition to the DPE, two other downstream core promoter elements have been identified. The motif ten element (MTE) usually interacts with the Inr to enhance Pol II transcription but is capable of substituting the loss of a TATA box or DPE and can act synergistically with both elements in an Inr-dependent manner (Lim et al., 2004). In contrast to all other core promoter elements, the downstream core element (DCE) contains three discontinuous subelements (CTTC, CTGT, AGC), spanning +6 to +34 adjacent to the TSS (Lee et al., 2005b). Although originally identified in TATA box promoters, the two TFIIB-recognition elements (BRE) are found in many TATA-less promoters. The upstream BRE (BRE^u) is also capable of binding TFIIB independently of TBP, while downstream BRE (BRE^d) appears to be TBP-dependent (Deng & Roberts, 2005). The variability of these seven

core promoter elements with their diverse consensus sequences and potential of multivalent combination underline the complexity of eukaryotic promoters. This complexity as well accounts for so far still inaccurate *in silico* prediction of core promoter elements, despite new approaches with improved accuracy (Zhang, 2007; Zeng et al., 2009).



Core Promoter Element	Position	Consensus Sequence (5'-3')	Bound Protein
BRE^u	-38 to -32	(G/C)(G/C)(G/A)CGCC	TFIIB
TATA	-31 to -24	TATA(A/T)A(A/T)(A/G)	TBP
BRE^d	-23 to -17	(G/A)T(T/G/A)(T/G)(G/T)(T/G)(T/G)	TFIIB
Inr	-2 to +5	PyPyAN(T/A)PyPy	TAF1/TAF2
MTE	+18 to +29	C(G/C)A(A/G)C(G/C)(G/C)AACG(G/C)	n.a.
DPE	+28 to +34	(A/G)G(A/T)CGTG	TAF6/TAF9
DCE	3 subelements +6 to +11 +16 to +21 +30 to +34	core sequence: CTTC CTGT AGC	TAF1

Figure 6: Core promoter elements recognized by TFIIB or TFIID.

TFIIB (green) and TFIID (yellow) recognize different core promoter elements. The TATA box is recognized by the TBP subunit of the TFIID complex. The table below lists the consensus sequences and respective positions for each core promoter element. n.a.: not available. With modifications from Thomas & Chiang, 2006.

1.6.2.1 Different pathways for PIC assembly

Understanding the assembly pathways of transcription complexes at distinct core promoter regions is crucial for the understanding of the gene transcription mechanisms. Two pathways have so far been identified to lead to functional PIC assembly at the site of core promoter elements. Transcription factors required for Pol II-mediated transcription and identified accessory factors had been designated earlier by their presence in different isolated fractions of nuclear extracts TFIIA to TFIIH (Thomas & Chiang, 2006).

1.6.2.2 The sequential assembly pathway

This pathway has been identified after intensive study of *in vitro* transcription experiments, describing multiple essential steps and a hierarchical nature for the assembly of an initiation-competent PIC. Buratowski et al. (Buratowski et al., 1989) specified the order of entry and relative position of each general transcription factor, which was refined in later experiments. Initially, TFIID recognizes the TATA box, followed by separate addition of TFIIA and TFIIB stabilizing promoter-bound TFIID, followed by recruitment of Pol II/TFIIF. This stable complex is completed by addition of accessory factors TFIIE and TFIIH.

1.6.2.3 The Pol II holoenzyme pathway

This alternative pathway of PIC assembly was first described after the purification of a preassembled complex of Pol II with or without a subset of general TF (Ossipow et al., 1995) and proteins involved in chromatin remodeling (Cairns et al., 1996), DNA repair (Maldonado et al., 1996) and mRNA processing (McCracken et al., 1997). Specifically, PolII and TFIIB, TFIIE, TFIIF and TFIIH have been identified in Pol II holoenzyme formation in the absence of TFIID. In this pathway, the core promoter-binding factor TFIID is suggested to facilitate the entry of Pol II holoenzyme to the promoter region in a process parallel to prokaryotic RNA polymerase recruited by the dissociable σ factor.

Both, the sequential assembly pathway and the Pol II holoenzyme pathway may exist in vivo, depending on specific needs for individual gene transcription or transcription efficiency (Lemon & Tjian, 2000), as in the suggested process of loop structure formation (cf. 1.5.1).

1.6.2.4 The TFIID complex

In either pathway, the multiprotein complex TFIID is one of the first general TF that binds the core promoter and facilitates PIC assembly. Historically, TFIID has been described as a TATA-binding factor through *in vitro* binding assays, but human TFIID seems to act primarily through TATA-less promoters *in vivo* as shown by ChIP-on-chip experiments (Kim et al., 2005). Together with the TBP, 13 TBP-associated factors (TAF) have been identified. The TBP subunit of TFIID thereby is specified in contacting TATA boxes, allowing TFIID to recognize TATA-containing promoters. TATA-less promoters displaying either one of the core promoter elements (fig. 6) are thereby recognized by the diversity of TAF subunits of TFIID (Pugh & Tjian, 1991). In depths characterization of TAF revealed that TFIID exerts potential as a co-activator in transcription, modulating activator function through multiple domain interactions. This has been explicitly demonstrated for the transcriptional activator SP1. The DNA-binding domain of SP1 specifically interacts with TAF7 of the TFIID complex (Chiang & Roeder, 1995), while the SP1 activation domain contacts TAF7 (Hoey et al., 1993). Human TAF7 is able to interact with different activators of transcription in addition to SP1, such as YY1 and USF, through its N-terminal domain and TAF1 in its central region, documenting its co-activator potential by bridging the bound activator to the general transcription machinery with simultaneous bending of interlaced DNA regions. These TAF-TAF interactions suggest transcriptional activators to function by recruiting TFIID to the promoter with subsequent PIC formation (Albright & Tjian, 2000), leading to the concept of TAF-dependent activation (fig. 7a). However, some genes do not rely on TAF, and TBP serves as a TATA-binding factor without mediation of TFIID in a TAF-independent activation (fig. 7b). In the absence of TAF, TBP can function with general cofactors such as TFIIA and TFIIH, replacing TAF by submitting signals between activators and the general transcription machinery.

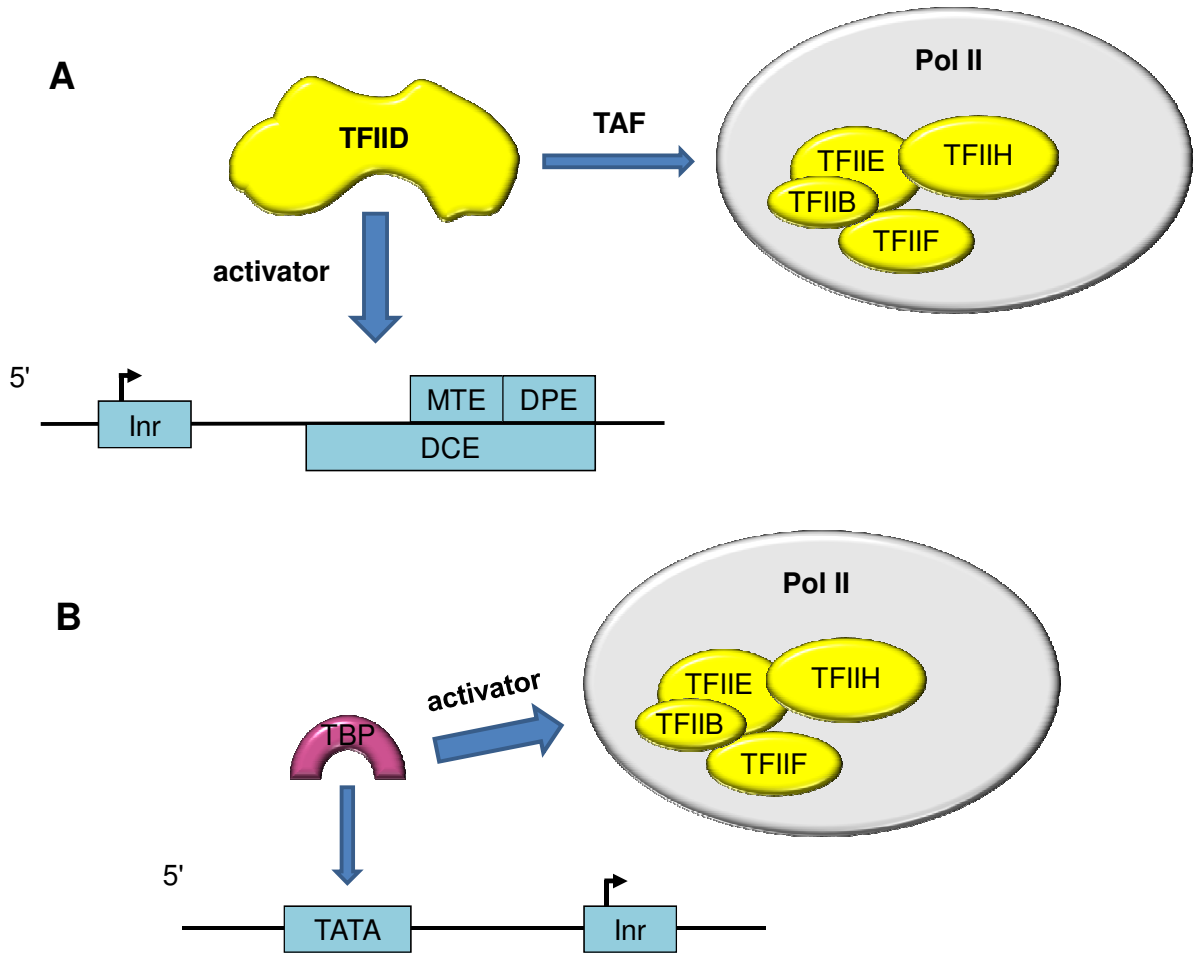


Figure 7: TAF-dependent and TAF-independent transcriptional activation.

The schematic representation shows promoter recognition in case of (A) TATA-less promoters and (B) TATA promoters. (A) TFIID-mediated promoter recognition is the rate-limiting step (wide arrow) facilitated by transcriptional activators. (B) When TBP is used as TATA-recognition factor, Pol II recruitment becomes the rate-limiting process. With modifications from Thomas & Chiang, 2006.

1.6.3 CpG islands promoters

The CpG dinucleotide, which resembles a substrate of the DNA methyltransferase, is underrepresented in the human genome because of ongoing deamination and thymine formation without correction by DNA repair enzymes. Although these mutations seem to take place accidentally and at random, extended stretches of DNA (0.5 – 2 kbp) exist that possess a relatively high density of CpG dinucleotides, contradicting this principle (Smale & Kadonaga, 2003). CpG islands remain mostly unmethylated in different tissues and stages of development compared to other DNA sequences and are associated with an accessible chromatin structure (Antequera & Bird, 1993). This characteristic feature has been suggested to be the cause of common CpG promoter bidirectionality after parasitic DNA insertion in close proximity (Kalitsis & Safferi, 2009). In humans, the percentage of CpG-associated promoters has been reported to be 88% (Kim et al. 2005), suggesting that CpG islands may play a more general role in gene expression than previously anticipated. Despite the high prevalence of CpG islands promoters, the elements that are essential for their core promoter function remain poorly defined. They are often characterized by the presence of multiple TSS that span a region of 100 bp or more. The TSS can co-localize with sequences exhibiting homology to the Inr consensus sequence but have also been reported to function independently of this core promoter element. One common feature of CpG islands is the presence of multiple binding sites for SP1. TSS are often located 40–80 bp downstream of the SP1 consensus sites and SP1 is suggested to direct the basal machinery to form a PIC within a loosely defined sequence portion (cf. 5.3). As discussed above, TFIID subunits are capable of core promoter recognition interact with the sequences most compatible with their DNA recognition motifs within that region. Binding of basal factors is strongly dependent on recruitment by activator proteins (fig. 6) bound to distal promoter elements. Activators such as SP1, may recruit TFIID to the core promoter, via a mechanism that is less dependent on the affinity of TFIID for the core promoter. Furthermore, active binding of SP1 in addition to promoter activation plays a key role in protecting CpG islands from de novo methylation (Brandeis, 1994).

1.7 Linkage and association studies

1.7.1 Family-based linkage analyses

The prevalence of complex genetic diseases results from interactions between various environmental factors and genetic variations. Identification of alleles affecting disease risk will help to better understand disease aetiology and pathophysiology. Genome-wide linkage analysis using polymorphic loci or markers such as SNPs spread across the genome is one common approach to map disease-associated genes. It is the simplest form of genetic mapping, measuring correlated segregation of Mendelian inherited markers for a trait of interest in families. Given meiotic recombination, linked markers are supposed to be located in close proximity to a disease-influencing gene. In this setting, disease-associated genomic regions for a continuous phenotype are identified that more likely harbour a causal genetic variant contributing to the trait. While linkage analysis has been proven to be successful in mapping genes underlying mono- or oligogenic diseases, the approach has its limitations in detecting genes underlying multifactorial complex traits. Another limitation of linkage mapping in humans is the recruitment of a sufficient number of affected families.

1.7.2 Population-based association analyses

One widely used method for identification of genes involved in complex disease pathophysiology is to compare allele or genotype frequencies of a given variant, mainly SNPs, between two groups (Couzin-Frankel, 2010). If a particular allele in either group is observed with higher frequency than expected by chance, it suggests an association with the disease or trait of interest. Most often, the candidate allele serves as a marker rather than representing the functional (causal) variant influencing the disease phenotype. The correlation between genetic variants and trait differences is most often assessed in case/control samples on a population scale. Until now, candidate gene association studies were performed to dissect common variants within genes of previously identified linkage regions or putative diseased pathways. However, association studies based on one or few candidate genes examine only a small fraction, disregarding the multifactorial nature of complex traits. Moreover, the hypothesis-driven approach leads to potential bias in selecting candidate genes. The completion of the human genome project and the development of cost-efficient high-throughput SNP genotyping platforms have set the stage for GWA studies. They represent an

important step beyond candidate gene studies, allowing SNP marker queries of the entire genome at levels of high resolution, unaffected by prior hypotheses. GWAs are based on the common disease/common variant hypotheses suggesting that a distinct number of variants with allele frequencies of more than 1% in the observed population contribute to the more common forms of human diseases. The number of GWAs is exponentially growing, addressing several disease categories including cancer, cardiovascular, autoimmune, metabolic, and neurodegenerative diseases, etc. Due to insufficient sample size or random errors, associated SNP markers found in GWAs often exhibit a lack of reproducibility and suffer from high false positive rates (Manolio et al., 2009). However, newer studies rely on very large sample sizes and appropriate statistical features (P-values corrected for multiple testing in GWAs) that allow reproducibility of allelic associations (Ehret, 2010).

1.8 Aim and design of the study

BGN was chosen as a candidate gene for CVD with regard to its emerging role in atherosclerosis pathophysiology, which has already been proven in appropriate animal models. As a multifunctional and proinflammatory component of the ECM in the heart but also in the arterial system, BGN is involved in the different steps of atherosclerosis, heart disease progression and outcome, etc.

Individual functional genetic variants and molecular haplotype patterns may account for the genetic predisposition to disease traits. Genetic variants that possibly affect the transcriptional regulation of *BGN* may thereby influence local *BGN* expression and stoichiometric availability in a given regulatory context, altering downstream cellular processes.

To identify genetic variants of the *BGN* 5'-flanking region within 57 CVD patients of the MolProMD study, we performed a systematic screening of patients' genomic DNA. Rare variants with a distribution of ~1% were calculated to be detected with ~80% likelihood, analysing 120 chromosomes. Molecular haplotype constellations were identified by individual subcloning and resequencing of patients' DNA. For subsequent *in vitro* functional analyses of variants altering *BGN* gene regulation, a sufficiently active promoter portion had to be determined, in a cell system endogenously expressing *BGN*, by the use of reporter gene assays. Reporter gene assays allow the determination and quantification of transcriptional activity of a defined promoter portion cloned into a promoter-less reporter gene vector.

Observed reporter gene expression is thereby explicitly driven from the inserted promoter fragment. In addition, the selected cell system required the transposition of *in situ* cellular processes such as ECM generation and processes of inflammation. Alternative TSS in the employed cell lines were assessed by 5'RACE. Using serial 5'-deletion of a sufficiently active promoter portion, cell specific *cis*-regulatory elements were identified. Possible alterations in DNA/protein interactions of *trans*-acting factors at *cis*-regulatory positions by introduction of detected genetic variants were analysed on a first level using computational models for *in silico* TFBS prediction. Co-expression experiments were used to control *in vivo* potential of predicted interacting transcriptional regulators in cell lines. *In vitro* binding of promoter elements by TF, with special respect to allele-specific differences, were assessed by gel shift experiments. EMSA competition assays were scheduled to identify interacting proteins. Where applicable novel approaches for promoter analyses and identification of TF were implemented in the work.

2 MATERIAL

2.1 Chemicals

Laboratory chemicals (highest quality) were generally purchased from Merck (Darmstadt), Sigma (Steinheim) or Roth (Karlsruhe). Chemicals from different suppliers are listed as follows.

Agarose	Biozym, Hess. Oldendorf
Bacto Agar	Becton Dickinson, Heidelberg
Bacto Tryptone	Becton Dickinson, Heidelberg
Bacto Yeast Extract	Becton Dickinson, Heidelberg
Blocking reagent	Roche Diagnostics, Mannheim
Biotin dd-UTP	Roche Diagnostics, Mannheim
Chloroform	Fluka Riedel-de Haën, Seelze
dNTPs (dATP, dCTP, dGTP, dTTP)	Fermentas, St. Leon-Rot
Ficoll	Fluka Riedel-de Haën, Seelze
Glutamin	Gibco, Karlsruhe
Nonidet P-40 (IGEPAL)	Sigma Aldrich, Schnelldorf
Poly(dI•dC)	USB, Staufen
Protease Inhibitor cocktail with EDTA (Complete)	Roche Diagnostics, Mannheim
Spermidine	Fluka Riedel-deHaën, Seelze
TGF-β1, human, recombinant	Calbiochem, San Diego, USA
IL-1β, human, recombinant	Calbiochem, San Diego, USA
PDGF, human, from platelets	Calbiochem, San Diego, USA

2.2 Other solutions and reagents

2.2.1 Sera and media

Dulbecco's modified eagle's medium (DMEM)	Sigma, Steinheim
Dulbecco's phosphate buffered saline (PBS)	Sigma, Steinheim
Fetal bovine serum (conditioned)	PAA, Pasching
Fetal bovine serum, iron supplemented	Cell Concepts, Umkirch
Roswell Park Memorial Institute 1640 medium (RPMI)	Sigma, Steinheim

2.2.2 DNA ladder and protein marker

Precision Plus Protein Dual Color Standard Plus	BioRad, Munich
Precision Plus Protein Western C	BioRad, Munich
GeneRuler 100 bp DNA ladder	Fermentas, St. Leon-Rot
GeneRuler 1 kb DNA ladder	Fermentas, St. Leon-Rot

2.2.3 Enzymes and antibiotics

Trypsine-EDTA (0.05%)	Gibco, Karlsruhe
Ampicillin	Roth, Karlsruhe
Penicillin/Streptomycine solution	Sigma Chemie, Steinheim
Spectinomycine	Sigma, Steinheim

2.2.4 Consumables and kits

BCA Protein Assay Kit	Thermo Fischer, Bonn
CL-XPosure Film	Thermo Fischer, Bonn
High Pure PCR Product Purification Kit	Roche Diagnostics, Mannheim
Immobilon-P Transfer Membrane (PVDF)	Millipore, Bedford, USA
LightShift Chemiluminescent EMSA Kit	Thermo Fischer, Bonn

Lipofectamine 2000	Invitrogen, Karlsruhe
Luciferase Assay System	Promega, Mannheim
Magnetic Protein-G beads	Invitrogen, Karlsruhe
Nanofectin	PAA, Pasching
Passive Lysis Buffer (5 x)	Promega, Mannheim
PureLink HiPure Plasmid DNA Purification Kit	Invitrogen, Karlsruhe
QIAamp DNA Blood Mini Kit	Qiagen, Hilden
NucleoSpin Plasmid	Macherey-Nagel, Düren
NucleoSpin RNA II	Macherey-Nagel, Düren
QIAquick Gel Extraction Kit	Qiagen, Hilden
QuikChange Multi Site-directed Mutagenesis Kit	Stratagene, Amsterdam, The Netherlands
SuperSignal West Pico Chemiluminescent Substrate	Thermo Fischer, Bonn
Whatman Paper 3MM Chr.	Biometra, Göttingen
Pipette tips 0.1 µl - 1000 µl	Sarstedt, Nürnbrecht
Reaction tubes 0.2 ml - 2 ml	Eppendorf, Hamburg Biozym, Hess. Oldendorf
15 ml/50 ml tubes	Greiner, Kremsmünster Nunc, Wiesbaden
Petri dishes	Sarstedt, Nürnbrecht
Plastics for cell culture	Greiner, Kremsmünster
PCR plates, microtiter plates	Abgene, Hamburg

2.2.5 DNA-modifying enzymes

BigDye3.1	Applied Biosystems, Foster City, USA
DNase I	Fermentas, St. Leon-Rot
Exonuclease I	Fermentas, St. Leon-Rot
GoTaq DNA Polymerase	Promega, Mannheim

Herculase II DNA Polymerase	Stratagene, Amsterdam, The Netherlands
High Fidelity PCR Enzyme Mix	Fermentas, St. Leon-Rot
Restriction endonucleases	Fermentas, St. Leon-Rot
RNase H	Fermentas, St. Leon-Rot
RNase A	Applichem, Darmstadt
Shrimp Alkaline Phosphatase	Fermentas, St. Leon-Rot
Superscript III	Invitrogen, Karlsruhe
RNaseOUT RNase Inhibitor	Invitrogen, Karlsruhe
T4 DNA Ligase	Fermentas, St. Leon-Rot
TdT terminal transferase	Roche, Mannheim

2.2.6 Antibodies

Table 1: Antibodies

Antibody	Host	Dilution	Manufacturer
c-FOS	rabbit	1:1000	Santa Cruz, Heidelberg
CREB/ATF1	rabbit	1:1000	Nanotools, Teningen
GATA1	rat	1:1000	Santa Cruz, Heidelberg
PU.1	rabbit	1:1000	Santa Cruz, Heidelberg
SP1	rabbit	1:1000	Upstate, Schwalbach
SRF	mouse	1:1000	Millipore, Schwalbach

2.2.7 Plasmids and vectors

pCRII TOPO cloning vector	Invitrogen, Karlsruhe
pCR8/GW/TOPO cloning vector	Invitrogen, Karlsruhe
pGL3-Basic reporter gene vector	Promega, Mannheim
pGL3-Control reporter gene vector	Promega, Mannheim
pGL3-Promoter reporter gene vector	Promega, Mannheim
pGFP reporter gene vector	Amaxa, Cologne
pRc/CMV expression vector	Dr. Dimitris Kardassis, Heraklion, Greece
pSP1/CMV expression vector	Dr. Dimitris Kardassis, Heraklion, Greece

2.2.8 Bacteria (*E. coli*)

Strain	Genotype	Manufacturer
DH5α (K12)	F- φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17(rk-, mk+) phoA supE44 thi-1 gyrA96 relA1 λ-	Invitrogen, Karlsruhe
Mach1	derivatives of *E. coli* W strains ΔrecA1398 endA1 tonA Φ80ΔlacM15 ΔlacX74 hsdR(rK- mK+)	Invitrogen, Karlsruhe

2.2.9 Eukaryotic cells

Line	Origin	Reference
COS-7	African green monkey kidney, fibroblast-like	ATCC no.: CRL-1651
EA.hy926	Human vascular endothelium	Edgell et al., 1983
HEK293T	Human embryonic kidney	ATCC no.: CRL-11268
THP-1	Human monocytes	ATCC no.: TIB-202

2.2.10 Laboratory equipment

Instrument	Specification	Manufacturer
Cell counter	Casy Model TT	Innovatis, Bielefeld
Centrifuge	Multifuge 3SR	Heraeus, Hanau
Centrifuge	5415C	Eppendorf, Hamburg
Centrifuge	5417R	Eppendorf, Hamburg
Centrifuge	5810R	Eppendorf, Hamburg
Centrifuge	J2-21M/E	Beckman Coulter, Krefeld
CO_2-Incubator (eukaryotic cells)	MCO-18AIC	Sanyo, Munich
Coffee maker	TKA1410V	Bosch, Munich
Developing machine	Optimax	Protec, Oberstenfeld
Gel electrophoresis chamber	Mini PROTEAN	BioRad, Munich
Gel electrophoresis chamber	StarPhoresis	Starlab, Ahrensburg
Gel imaging	AlphaImagerEC	Alpha Innotech Corp, San Leandro, USA
Incubator shaker	Series 25	New Brunswick Scientific, Nürtingen
Luminometer	Sirius V12	Berthold Detection Systems, Pforzheim
Microbiological incubator	B 6120	Heraeus, Hanau
Microscope	Axiovert 40 CFL	Zeiss, Jena
pH-Meter	Calimatic 766	Knick, Dülmen
Spectrophotometer	NanoPhotometer	Implen, Munich
Power supply	PowerPackBasic	BioRad, Munich
Sequence detection system	7500 ABIprism	Applied Biosystems, Foster City, USA

Sonicator	Bioruptor UCD-200	Diagenode, Liège, Belgium
Sterile hood (bacteria)	Class II type EF	Clean air Techniek B.V., Woerden, The Netherlands
Sterile hood (eukaryotic cells)	HS 12	Heraeus, Hanau
Tank Blot chamber	Mini Trans-Blot Cell	BioRad, Munich
Thermocycler	PTC-225, PTC-240 DNA Engine Tetrad (2)	MJ Research, Miami, USA
UV-table	Transilluminator	Intas, Göttingen

3 METHODS

3.1 Molecular biological methods

General molecular methods were performed as described in "Molecular Cloning" (Sambrook & Russel, 2001). Variations in protocols are indicated where appropriate.

3.1.1 Preparation of genomic DNA

Genomic DNA was extracted from mononuclear cells or cell lines using the QIAamp DNA Blood kit (Qiagen) according to manufacturers' instructions. Briefly, 200 µl of human whole blood (EDTA-treated) or 5 x 10^6 cultured cells were mixed with 20 µl Proteinase K and 200 µl binding buffer, incubated at 56°C for 10 min and loaded to the spin column. After two washing steps, DNA was eluted with dH_2O (pH 7 – 8.5) or TE. Genomic DNA was held at 4°C and stored at -20°C.

3.1.2 Preparation of total RNA

RNA was extracted from cultured cells using the NucleoSpin RNA II kit (Macherey-Nagel) following the manufacturers' instructions. Cells (5 x 10^6) were lysed by adding 350 µl lysis buffer/1% β-mercaptoethanol. The lysate was cleared by filtration through filter columns and 350 µl ethanol was added to adjust binding conditions. The lysate was loaded onto RNA binding column followed by desalting of the membrane. DNA was removed from the column by incubation with DNase for 15 min. Washed thrice, total RNA was eluted in RNase-free water and stored at -80°C.

3.1.3 Preparation of plasmid DNA

Plasmid DNA from *E. coli* cultures (1.5 ml) for cloning was obtained by use of the NucleoSpin Plasmid kit (Macherey-Nagel) following manufacturers' instructions. Pelleted cells were resuspended and lysed for 5 min at RT. After addition of neutralization buffer, lysate was cleared by centrifugation. DNA was bound to silica membrane and washed twice. Ethanol was qualitatively removed by centrifugation before elution of plasmid DNA.

Alternatively and in case of bacterial artificial chromosome (BAC) preparation was performed by alkaline lysis with SDS (Birnboim & Doly, 1979).

Transfection grade plasmid DNA of highest quality without endotoxins was preparated from 100 ml *E. coli* cultures using the PureLink HiPure Plasmid DNA Purification kit (Invitrogen) according manufacturers' instructions. Cells were spun down, resuspended and lysed with simultaneous RNase A treatment (5 min). Lysate was cleared by precipitation and centrifugation (12000 x g, 10 min, RT). Supernatant was routinely cleared from bacterial endotoxins using a separate endotoxin removal protocol (Buffer A and B). The cleared lysate was loaded onto an anion exchange column, washed and eluted by gravity flow. Eluate was precipitated by addition of isopropanol (70% v/v) and centrifugation (15000 x g, 30 min, 4°C). DNA was washed in 70% ethanol, centrifuged, air-dried and resuspended in TE buffer. Plasmid DNA for transfection was held at 4°C and stored at -20°C.

Endotoxin Removal Buffer A		Endotoxin Removal Buffer B	
50 mM	MOPS, pH 7.0	100 mM	sodium acetate, pH 5.0
750 mM	sodium chloride	750 mM	sodium chloride
10% (w/v)	Triton X-100	1% (w/v)	Triton X-100
10% (v/v)	isopropyl alcohol		

3.1.3.1 Quality and quantity control of nucleic acids

Concentration and purity of nucleic acids was measured photometrically using a nanophotometer (Implen). Identity and potential contamination of plasmid DNA was routinely tested by hydrolysation with selected restriction enzymes and automated sequencing.

3.1.4 Polymerase Chain Reaction (PCR)

PCR was performed with variations according to Mullis et al. (Mullis et al., 1990). For standard applications, GoTaq DNA polymerase (Promega) was used. To guaranty amplicon sequence identity with template DNA for cloning of transfection vectors, a *Taq/Pfu* proofreading enzyme mix was used (High Fidelity Mix, Fermentas; Herculase II, Stratagene).

Standard PCR reaction contained:

5 ng	of genomic DNA
10 µM	of each primer
200 µM	of each dNTP
1 x	polymerase buffer
2 mM	$MgCl_2$
0.6 U	DNA polymerase

add nuclease free water to 25 µl

Enhancing additives were used where appropriate to reduce the formation of secondary structures (1 M Betain, 2 – 10% DMSO, 1/3 deazaGTP/dGTP) or to stabilize DNA polymerase (0.1 - 1% Triton X-100, Tween 20, NP - 40).

Standard PCR program:

A:	Initial denaturation	95°C, 5 min	
B1:	Denaturation	95°C, 1 min	25 – 35 cycles
B2:	Annealing	x°C, 45 sec	
B3:	Elongation	72°C, 1 min/kb	
C:	Final elongation	72°C, 10 min	

Modifications of the program were applied where appropriate.

Calculation of oligo T_m was performed using the algorithm of Chester & Marshak (Chester & Marshak, 1993):

$$(69.3 + \%GC \sim \text{content x } 0.41) - \frac{650}{\text{lenght of oligonucleotide}}$$

Touch down PCR:

To enrich specific PCR products, annealing temperatures was gradually decreased for maximum annealing stringency. The annealing temperature was set 5 - 10°C over primer annealing temperature and reduced by 2°C every second cycle until calculated annealing temperature was reached, followed by 25 cycles at final annealing temperature.

Nested PCR:

To amplify weak PCR signals and to generate specific fragments from a mixture of unspecific amplicons, a second set of inner or nested primers was used. Amplified PCR product was either used directly as template (1 µl) or after extraction from agarose gels (3.1.11).

3.1.5 cDNA synthesis

Generation of cDNA was performed using Superscript III (Invitrogen) according to the manufacturer. Total RNA (500 ng – 5 µg, extracted as described in 3.1.2) were reversely transcribed using either 1 µl of oligo(dT) or 2 pmol of gene-specific primer in presence of a RNase inhibitor (RNaseOUT, Invitrogen) for 60 min at 50°C. Reaction temperature was increased to 55°C for gene-specific primers. Reaction was inactivated by heating at 70°C for 15 min. Prior to amplification of longer PCR targets (>1 kb) and in case of 5'RACE experiments, complementary RNA was removed by incubation with RNase H (2 units, 20 min, 37°C). Integrity of cDNA was routinely controlled by diagnostic PCR for hRP27.

3.1.6 5'RACE

RACE PCR was performed as described previously (Frohman, 1993). All experimental steps were performed within the same day, without freezing of the generated cDNA to achieve maximum accuracy at TSS identification. Briefly, total RNA was extracted (3.1.2) and first strand cDNA was generated (3.1.5) using a *BGN* specific antisense primer, followed by poly(A)tailing of the 3'-end using terminal desoxynucleotidyl transferase (TdT). A homopolymeric d(T)-anchor primer, complementary to the 3'-poly(A) tail, and an antisense primer located within the gene were used to amplify the first PCR product. Nested primer were used to increase the amount of PCR product for subsequent sequencing (tab.2).

Table 2: Oligonucleotide sequences for 5'RACE

Description	Sequence 5' to 3'	Position	Ref. Accession #
BGN RT	AGCTTCCTGGCTCTGCCTCC	Exon 2	NT_011726.13
dT-anchor primer	GACACGCGTATCGATGTCGACTTTTTTTTTTT TTTTT	-	-
Adaptor primer	GACCACGCGTATCGATGTCGAC	-	-

First strand cDNA synthesis reaction using SuperScript III (Invitrogen)

2 µg total RNA

2 pmol gene-specific primer

500 µM of each dNTP

ddH_2O to 13 µl

Mixture was heated to 65°C for 5 min and incubated on ice for at least 1 min.

4 µl synthesis buffer

5 mM DTT

1 µl RNaseOUT RNase inhibitor (Invitrogen)

1 µl SuperScript III reverse transcriptase

Mixture was incubated for 60 min at 55°C.

RNA complementary to the cDNA was removed by addition of RNaseH (2 U) and incubation at 37°C for 20 min.

Poly(A) tailing

16.5 µl purified cDNA

5 µl reaction buffer

2.5 µl dATP (2 mM)

Mixture was incubated at 94°C for 3 min and chilled on ice.

1 µl TdT (80 U/µl)

Incubation at 37°C for 30 min followed by heat inactivation for 10 min at 70°C.

PCR reaction was performed as described (3.1.4) using gene-specific primer and d(T)-anchor primer at 55°C for 35 cycles.

3.1.7 DNA/RNA-modifying reactions

3.1.7.1 Hydrolysation with bacterial endonucleases

Hydrolysation with endonucleases was conducted using 100 - 500 ng DNA and 1 U restriction enzyme. H_2O and 10 x reaction buffer were added to a total volume of 20 µl, incubated for 1 hr at 37°C and heat-inactivated at 70°C for 10 min.

3.1.7.2 Dephosphorylation of DNA

Shrimp Alkaline Phosphatase (SAP) was used for dephosphorylation of 5'-ends of linearized plasmid DNA to disable re-ligation. One unit of SAP, dH_2O and 10 x reaction buffer were added to the digestion reaction (cf. 3.1.7.1) to a total volume of 25 µl. Reaction mixtures were incubated at 37°C for 30 min and heat-inactivated for 10 min at 65°C.

3.1.7.3 Biotinylation of oligonucleotides for EMSA experiments

Single-stranded oligonucleotides (~30 bp; IBA, Göttingen) for EMSA experiments were synthesized at a minimum coupling efficiency of > 98.5% and purified twice by high pressure liquid chromatography. Oligonucleotides were 3'-biotinylated with biotin-16-ddUTP (Roche) using TdT. 5 pmol of each oligo were labelled in a reaction mix containing 2 mM $CoCl_2$, 500 pmol biotin-16-ddUTP, and 60 U TdT at 37°C for 30 min. Labelled probes were chloroform extracted and centrifuged twice (2 min, 14000 x g, RT). Specific annealing of oligonucleotides (20 fmol) was achieved by denaturation at 95°C for 10 min in 100 mM NaCl, followed by slow over night cool down to RT. Double-stranded competitors were generated using a concentration of 2 pmol of each unlabeled primer. Annealing quality and signal intensity of identical probes holding single base pair mutations were controlled routinely by gel electrophoresis.

3.1.8 Agarose gel electrophoresis

Agarose concentrations of 0.5% to 3% were applied in 1x TAE buffer, depending on fragment sizes.

50 x TAE		6 x Loading buffer	
2 M	Tris base	0.02% (w/v)	bromphenol blue
50 mM	EDTA	0.02% (w/v)	xylene cyanol
adjusted to pH 8.0 with acetic acid		30% (v/v)	glycerol
		20 mM	Tris-HCl, pH 7.6
		2 mM	EDTA

For visualizing DNA double-strands, 0.05 µg/ml ethidium bromide (EtBr) was added to the gel solution. Longer fragments (>5 kbp) generated for cloning after gel extraction were stained using a Crystal Violet (Sigma) solution (0.01% w/v in 1 x TAE) and were cut from gels without the use of UV light to prevent DNA strand breaking.

3.1.9 Site-directed mutagenesis

Site-directed mutagenesis was used for selective introduction of genetic variants into the respective reporter gene vectors using a highly efficient commercial system (QuikChange Multi Site-directed Mutagenesis Kit, Stratagene). Allelic variants -578A, -151A, +94T were introduced into *BGN* promoter fragments as follows: denaturation of input vector wt DNA was followed by annealing and extension of the mutagenic primers (tab. 3), and ligation of nicks. Methylated and hemimethylated DNA was digested with endonuclease *Dpn*I (60 min, 37°C) followed by bacterial transformation. When indicated, up to three mutations at different sites were performed in a single reaction.

Standard reaction		Standard mutagenesis PCR	
2.5 µl	QuikChange Multi reaction buffer	Initial denaturation	95°C, 5 min
0.5 µl	Quik Solution	Denaturation	95°C, 1 min
100 ng	template DNA	Annealing	55°C, 1 min
100 ng	of each mutagenic primer	Elongation	65°C,
1 µl	dNTP mix		(2 min/1 kb)
1 µl	QuikChange Multi enzyme blend	30 cycles	

Table 3: Oligonucleotide sequences for site-directed mutagenesis

Description	Sequence 5' to 3'	Position	Ref. Accession #
BGN MUT -578A	AAGGGAAGAAAGTCTAGAGTGGAAGGGAGGG	-593	NT_011726.13
BGN MUT -151A	GAAGCTGCCAGGGGGACCGGGAAGCCTGCCC	-136	NT_011726.13
BGN MUT +94T	CACCACCCCAGCCCTCCAACTAGTCAGCCT	+79	NT_011726.13

3.1.10 Construction of reporter gene plasmids

Promoter wt fragments were generated using MolProMD (not ECTIM) patients genomic DNA. For the analysis of the three *BGN* MolHaps, a 1025 bp fragment was generated using primer at position -893 (upstream TSS1) and +132. Genetic variants were introduced by site-directed mutagenesis (cf. 3.1.9) yielding constructs pBGN-MoLHap1/luc, pBGN-MolHap2/luc, pBGN-MolHap3/luc. Deletion constructs of the *BGN* 5'-flanking region were amplified using one antisense primer at position +162 bp and sense primers generating constructs shown in fig. 8. For all deletion constructs, genomic DNA representing the wt sequence was used as template.

For transient transfection assays, the generated PCR fragments were introduced in 5'-3'-orientation into the promoter-less luciferase reporter gene vector pGL3-Basic (Promega) using the Gateway cloning system (Invitrogen). This highly efficient cloning technique is based on site-specific recombination properties of bacteriophage λ (Landy et al., 1989). Recombination events occur at specific attachment sequences on phage DNA (*attP*) and bacteria DNA (*attB*). Gel extracted PCR fragments were primarily cloned into entry vector

pCR8/GW/TOPO where the fragment is flanked by *attL* sequences, and transformed into competent Mach1 bacterial cells (cf. 3.3.1.2). After plasmid preparation and purification (cf. 3.1.3), the entry clone was incubated with the modified pGL3-Basic destination vector carrying artificial *attR* sites. Addition of the LR Clonase enzyme results in the exchange of the Gateway cassette in combination with the insert of interest. Generated plasmids were double digested with selected, sequence-specific restriction endonucleases (3.1.7.1) followed by agarose gel electrophoresis to verify 5'-3'-orientation and correct insert size. All generated vectors were directly sequenced (cf. 3.1.12) to ensure sequence accuracy and identity.

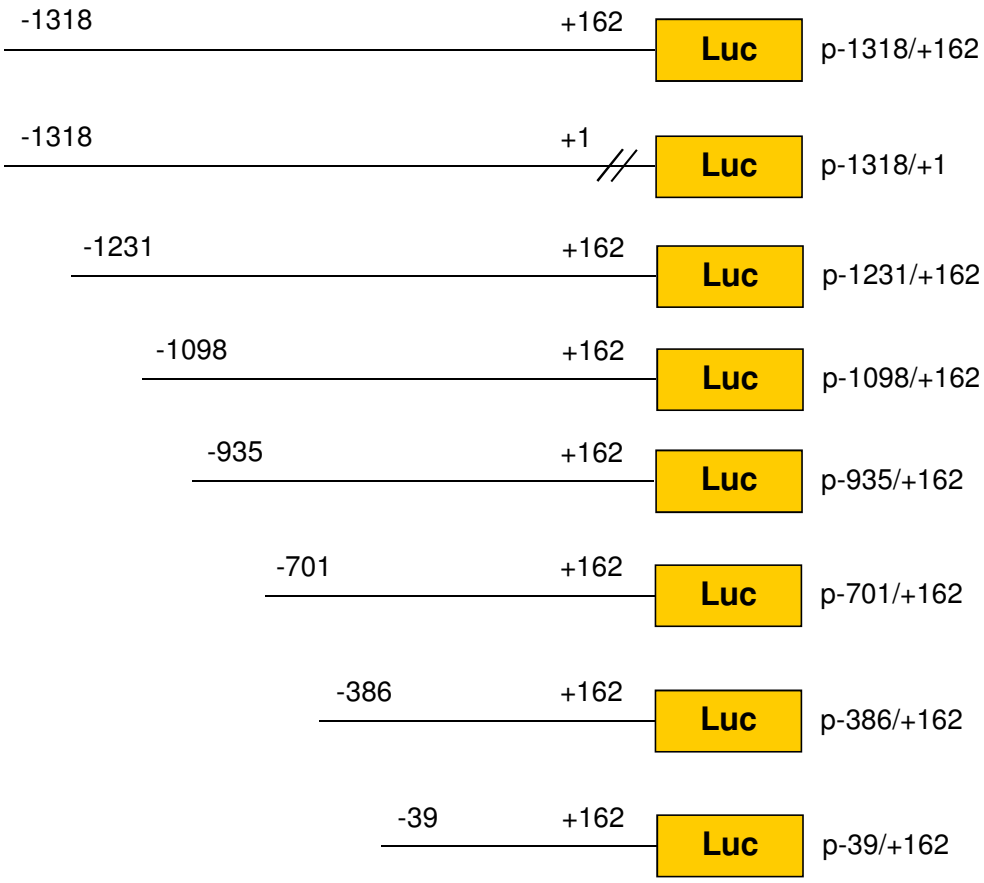


Figure 8: Schematic representation of deletion constructs of the human *BGN* promoter.

Deletion constructs were cloned into the pGL3-Basic vector system (Promega). This vector does not possess promoter activity of its own but features a luciferase cassette adjacent to the multiple cloning site. Transcriptionally active promoter fragments result in the expression of the luciferase protein, permitting measurement of promoter activity in relative light units. Sequence positions are shown according to TSS1. Luc: luciferase.

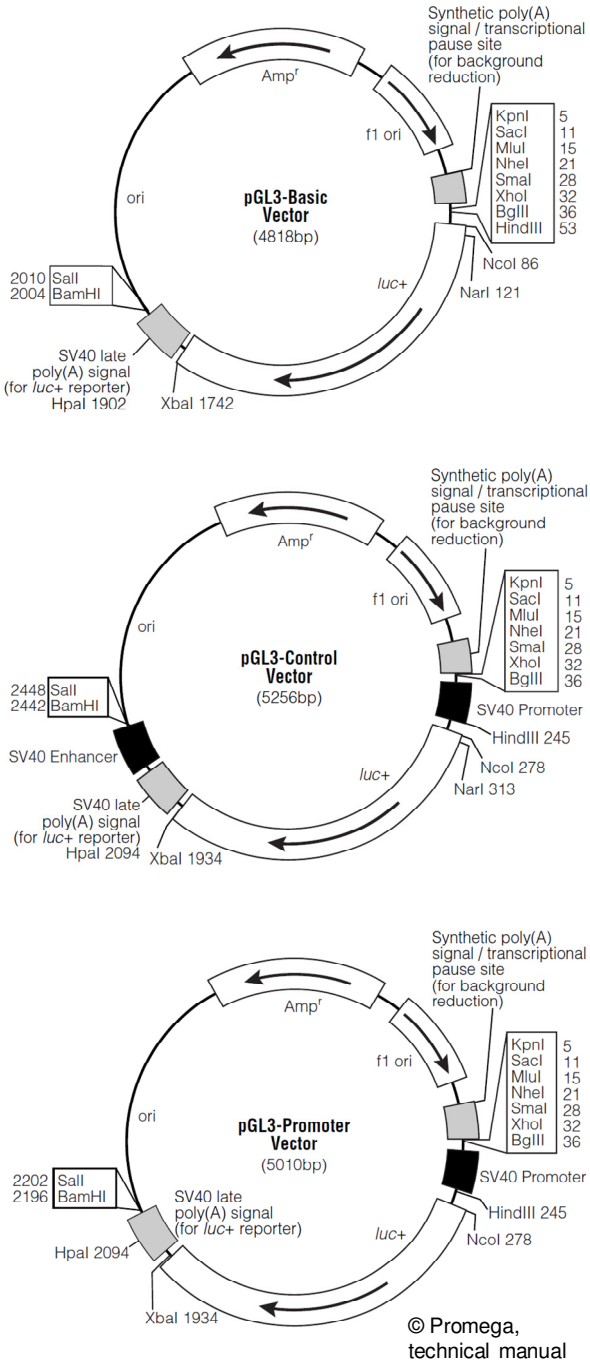


Figure 9: pGL3-System vector circle maps.

The pGL3-Basic vector lacks eukaryotic promoter and enhancer sequences, pGL3-Control vector contains SV40 promoter and enhancer sequences, the pGL3-Promoter vector contains an SV40 promoter upstream of the luciferase gene. Amp^r: gene conferring ampicillin resistance in *E. coli; luc*$^+$: cDNA encoding the modified firefly luciferase; f1 ori: origin of replication derived from filamentous phage; ori: origin of replication in *E. coli.* Arrows within *luc*$^+$ and the Amp^r gene indicate the direction of transcription; the arrow in the f1 ori indicates the direction of ssDNA strand synthesis.

Standard pCR8/GW/TOPO cloning reaction	LR Clonase reaction
1 µl salt solution (1.2 M NaCl, 0.06 M $MgCl_2$)	100 ng entry vector
1 µl pCR8/GW/TOPO cloning vector (10 ng/µl)	150 ng destination vector
4 µl purified insert	2 µl LR Clonase
incubation for 5 min at RT,	add 8 µl TE-buffer
transformation in competent	incubation for 1 hr at 25°C
Mach1 bacterial cells	add 1 µl Proteinase K
	incubation for 10 min at 37°C

Table 4: Oligonucleotide sequences for promoter constructs

Description	Sequence 5' to 3'	Position	Ref. Accession #
BGN MolHap SS	GCTAGCCAGTGTCGTACTAAGGACCT	-893	NT_011726.13
BGN MolHap AS	CCCGGGTGGAGAGGGGAGGCGCCA	+132	NT_011726.13
BGN -1318/+162 SS	ACAGTTTGGTCAAGGTGCCA	-1318	NT_011726.13
BGN -1231/+162 SS	CATGCCCCCAGGTAGAGC	-1231	NT_011726.13
BGN -1098/+162 SS	GATGCCGCCTCTCTCTAGC	-1098	NT_011726.13
BGN -935/+162 SS	GCTCTTGCTGAGGGAAGAAG	-935	NT_011726.13
BGN -701/+162 SS	GAAGACTGACTCGCTGGAT	-701	NT_011726.13
BGN -386/+162 SS	GTTTGAGTGAGTGCGAGTGT	-386	NT_011726.13
BGN -39/+162 SS	TGCCCAGCCTTTAGCCTC	-39	NT_011726.13
BGN +162 AS	CTGAGGAGGCAGCTTGAAG	+162	NT_011726.13

3.1.11 Purification of PCR products

Purification of PCR products for sequencing, cloning or generation of probes was performed by extraction from agarose gels or directly from PCR reactions.

Gel extraction was performed using QIAquick Gel Extraction Kit (Qiagen). DNA fragments (between 70 bp and 10 kbp) were excised from low agarose gels, mixed with solubilization buffer QG (pH 7.5) and heated for 10 min at 50°C. If changes in colour of buffer QG occurred, 10 µl of 3 M sodium acetate, pH 5.0 were added to restore optimal binding pH. Samples were mixed with one gel volume of isopropanol (100%) loaded onto column silica membranes (≤400 mg agarose) and washed twice. DNA was eluted in buffer EB (10 mM

Tris-HCl, pH 8.5). PCR product purification was performed using the High Pure PCR Product Purification Kit (exclusion limit 100 bp; Roche). Reactions (up to 100 µl) were mixed with binding buffer and loaded onto a silica membrane column. After two washing steps, the DNA was eluted in 10 mM Tris-HCl, pH 8.5.

For sequencing reactions, a rapid one-step PCR clean-up was performed using the ExoSAP protocol. A mixture of Exonuclease I and Shrimp Alkaline Phosphatase (SAP) (both Fermentas) was used to digest small single-stranded fragments (e.g. primers) and to remove dNTPs.

ExoSAP mixture

20 U	Exonuclease I (*E. coli*)
10 U	Shrimp Alkaline Phosphatase (SAP)

add dH_2O to 100 µl

PCR products (5 µl) were prepared with 1 µl of ExoSAP mix and incubated at 37°C for 30 min followed by heat-inactivation for 15 min at 80°C.

3.1.12 Sequencing

Samples were sequenced for detection and localization of genetic variants in the MolProMD study (not ECTIM), and to control for sequence identity of DNA fragments and plasmid constructs using an automated ABI 3730 fluorescence sequencer with BigDye terminator chemistry (PE Applied Biosystems).

Standard sequencing reaction

5 - 50 ng	purified PCR fragment or 120 ng plasmid DNA
1.6 µM	sequencing primer
2 µl	5 x BigDye buffer
1 µl	BigDye3.1 (Applied Biosystems)

add dH_2O to 10 µl

3.1.13 EMSA

To reveal *in vitro* DNA/protein interactions, EMSA experiments were performed. Oligonucleotides (tab. 4) representing the DNA sequence of interest were purchased from IBA (Göttingen) with a maximum coupling efficiency of 98.5%. Oligonucleotides were 3'-biotinylated with biotin-16-ddUTP (Roche) using TdT (cf. 3.1.7.3) for detection with anti-biotin antibodies. Protein binding to the applied probe is visualized by a shifted band in lanes containing protein extract in comparison to the control lane representing the unbound DNA probe. Per standard reaction, 5 µg nuclear protein extract was incubated with 500 ng presheared poly (dI●dC) as non-specific competitor, 250 mM Betaine, and a 200-fold molar excess of unlabeled oligonucleotide as specific competitor for 5 min at RT. After addition of the labelled probe, reactions were incubated for 15 min at RT. Probes and bound complexes were separated on a 6% native PAGE (0.5x TBE; 100 V), and blotted afterwards onto PVDF membranes (cf. 3.2.6) (0.5x TBE; 100 V; 60 min). DNA probes were cross-linked to the membrane by UV-light (312 nm) for 10 min.

4 x binding buffer	
20 mM	$MgCl_2$
240 mM	KCl
40 mM	HEPES/KOH, pH 7.9
5 mM	spermidine
16% (w/v)	Ficoll

6% PAGE	
2 ml	Acrylamid-Bis 30%
1 ml	5 x TBE
83.7 µl	APS, 10%
3.7 µl	TEMED

5 x TBE	
45 mM	Tris base
45 mM	boridic acid
10 mM	EDTA

Visualization of probes was achieved using the Chemiluminescent Nucleic Acid Detection Module (Thermo Fischer). Membranes were blocked in blocking buffer containing streptavidin-horseradish peroxidase conjugate. After washing and equilibration, membranes

were incubated in substrate working solution (luminol/enhancer and peroxidase solution). Membranes were than exposed to CL-X Posure Film (Thermo Fischer).

Table 5: Oligonucleotide sequences of EMSA probes

Description	Sequence 5' to 3'	Position	Ref. Accession #
BGN-578A SS	AGGGAAGAAAGTCTG**A**AGTGGAAGGGAGGGCACAGG	-593	NT_011726.13
BGN-578A AS	TCCCTTCTTTCAGAC**T**TCACCTTCCCTCCCGTGTCC	-593	NT_011726.13
BGN-578G SS	AGGGAAGAAAGTCTG**G**AGTGGAAGGGAGGGCACAGG	-593	NT_011726.13
BGN-578G AS	TCCCTTCTTTCAGAC**C**TCACCTTCCCTCCCGTGTCC	-593	NT_011726.13
BGN-151G SS	GCTGCCAGGGGGG**G**CCGGGAAGCCTGCCCCCT	-163	NT_011726.13
BGN-151G AS	CGACGGTCCCCC**C**GGCCCTTCGGACGGGGGA	-163	NT_011726.13
BGN-151A SS	GCTGCCAGGGGGG**A**CCGGGAAGCCTGCCCCCT	-163	NT_011726.13
BGN-151A AS	CGACGGTCCCCC**T**GGCCCTTCGGACGGGGGA	-163	NT_011726.13
BGN+94G SS	GCCCACCACCCCAGCCC**G**CCAACTAGTCAGCCT	+77	NT_011726.13
BGN+94G AS	CGGGTGGTGGGGTCGGG**C**GGTTGATCAGTCGGA	+77	NT_011726.13
BGN+94T SS	GCCCACCACCCCAGCCC**T**CCAACTAGTCAGCCT	+77	NT_011726.13
BGN+94T AS	CGGGTGGTGGGGTCGGG**A**GGTTGATCAGTCGGA	+77	NT_011726.13
BGN-151MUT SS	GCTGCCATGGTG**G**CCGGGAAGCCAGCACCCT	-163	NT_011726.13
BGN-151MUT AS	CGACGGTACCAC**C**GGCCCTTCGGTCGTGGGA	-163	NT_011726.13
AP-1 SS consensus	CGCTTGATGACTCAGCCGGAA	-	Lee et al., 1987a
AP-1 AS consensus	GCG AACTACTGAGTCGGCCTT	-	Lee et al., 1987a
SP1 SS consensus	ATTCGATCGGGGCGGGGCGAGC	-	Kodonaga et al., 1987
SP1 AS consensus	TAAGCTAGCCCCGCCCCGCTCG	-	Kodonaga et al., 1987

3.1.14 ChIP

To access the binding of transcription factors to DNA sequences of interest *in vivo*, ChIP assays were performed using a modified protocol (Boyd et al., 1998; Liu et al., 2000). Proteins interacting with the DNA were crosslinked and bound chromatin was precipitated using selected specific antibodies. PCR amplification was used to identify the segment of the genome associated with the protein. A total of about 10^7 cells were fixed by adding formaldehyde to a final concentration of 1% (v/v) and incubated for 30 min at RT. Cells were washed twice with cold PBS (Sigma) and lysed. Cellular debris was removed by centrifugation. Isolated nuclei were lysed followed by sonication using a Bioruptor (Diagenode; intensity: high, interval: 0.5, 30 min) with resulting chromatin length of ~300 bp. Fragment size was routinely controlled on agarose gels. After centrifugation the supernatant was incubated for 30 min (4°C) with rabbit pre-immune serum followed by incubation (30 min, 4°C) with freshly prepared (blocked with BSA and tRNA, 1 hr, 4°C) magnetic Protein-G beads (Invitrogen). After centrifugation, supernatant was transferred to low-binding tubes and 3 µg of specific antibody was added, followed by overnight incubation. Magnetic Protein-G beads were added the next day (10 µl, 4°C, for 3 hrs). After extensive washing the antibody/protein/DNA complex was eluted from the beads, followed by removal of formaldehyde crosslinks and protein digestion with proteinase K (67°C over night). DNA was prepared by phenol/chloroform/isoamyl alcohol extraction and used for PCR analysis.

Cellular lysis buffer	Nuclear lysis buffer	Dilution buffer
10 mM Tris pH 8.0	50 mM Tris pH 8.0	20 mM Tris pH 8.0
10 mM NaCl	10 mM EDTA	2 mM EDTA
0.2% (v/v)NP-40	1% (w/v) SDS	150 mM NaCl
Roche Complete	Roche Complete	1% (w/v) Triton X-100
proteinase inhibitor	proteinase inhibitor	Roche Complete
		proteinase inhibitor

Wash buffer I	Wash buffer II	Wash buffer III
20 mM Tris pH 8.0	10 mM Tris pH 8.0	20 mM Tris pH 7.6
2 mM EDTA	1 mM EDTA	50 mM NaCl
50 mM NaCl	0.25 M LiCl	
1% (w/v) Triton X-100	1% (v/v) NP-40	Elution buffer
0.1% (w/v) SDS	1% (w/v) Ceoxycholic	100 mM $NaHCO_3$
Roche Complete	acid	1% (w/v) SDS
proteinase inhibitor		

Table 6: Oligonucleotide sequences for ChIP

Description	Sequence 5' to 3'	Position	Ref. Accession #
BGN ChIP -28_+78 SS	AGCCTCCCGCCCGCCGCCT	-28	NT_011726.13
BGN ChIP -28_+78 AS	CGCACGTCTATCTGTCCGGTG	+78	NT_011726.13
BGN ChIP -918_-806 SS	AGGGGACACTACGGGACAG	-918	NT_011726.13
BGN ChIP -918_-806 AS	TTACCCCACCAAGACTC GC	-806	NT_011726.13
BGN ChIP -1164_-1066 SS	CGGTTGAGTGATGGCACTG	-1164	NT_011726.13
BGN ChIP -1164_-1066 AS	GGAGAGAGGGGTGGCTAGAG	-1066	NT_011726.13
BGN ChIP G-578A SS	GATCGGGGCCTCTTTTTAAG	-610	NT_011726.13
BGN ChIP G-578A AS	TAGCTGTTGTGGATTTCTGG	-511	NT_011726.13
BGN ChIP G-151A SS	CGTCTACAAGAAAATTGCTC	-191	NT_011726.13
BGN ChIP G-151A AS	GGGGAGGGAGGAAAGGAG	-77	NT_011726.13
BGN ChIP G+94T SS	TGCCCAGGAGTGAGTAGCTG	+14	NT_011726.13
BGN ChIP G+94T AS	AGGCGCAGGCTGACTAGTTG	+115	NT_011726.13

3.2 Protein biochemical methods

3.2.1 Preparation of proteins

Crude protein extracts were prepared from cell culture by scraping cells in presence of lysis buffer. Cellular debris was removed by centrifugation (12000 x g, 5 min, 4°C). Pre-heated SDS-PAGE sample buffer was added to supernatants and heated to 95°C for 5 min. Samples were chilled on ice and loaded directly onto a gel or kept frozen at -20°C until used.

Lysis buffer	
20 mM	imidazole, pH 6.8
100 mM	KCl
1 mM	$MgCl_2$
10 mM	EGTA
0.2% (v/v)	Triton X-100
10 mM	NaF
1 mM	sodium vanadate
1 mM	sodium molybdate
10 mM	sodium pyrophosphate
25 mM	β-glycerophosphate

4 x SDS sample buffer	
200 mM	Tris-HCl, pH 6.8
8% (w/v)	SDS
0.4% (w/v)	bromphenol blue
40% (v/v)	Glycerol

3.2.2 Isolation of nuclear proteins

Nuclear protein extracts were obtained using a modified protocol by Schreiber et al. (Schreiber et al., 1989). Briefly, 10^7 cells were washed twice and harvested by scraping in ice cold PBS. Cells were spun down (5000 x g, 2 min, 4°C) and pellets were resuspended in an appropriate volume of "low salt" buffer (400–800 µl) and allowed to swell on ice for 15 min. After addition of detergent NP-40 (25–75 µl of a 10% solution, depending on cell line), cells were vortexed and incubated at room temperature for up to 5 min until cells were lysed. Efficiency of this step was controlled by microscopy and optimized for each cell line. The

supernatant containing the cytosolic protein fraction was removed by centrifugation (5000 x g, 5 min, 4°C). The pellet was resuspended in an appropriate volume of "high salt" buffer (50–150 µl, adjusted to pellet size). After 2 hrs of rotation on a wheel at 4°C, cellular debris was spun down (24000 x g, 1 hr, 4°C) twice. Supernatant was moved to a fresh tube after the first centrifugation step. The nuclear protein fraction was aliquoted on ice, snap frozen in liquid nitrogen and kept at –70°C. Nuclear protein fractions were routinely tested for protein concentration using BCA assay. Amount of DNA contamination was tested using spectroscopy. Integrity of proteins was tested using Coomassie SDS gel staining.

"low salt" buffer	
10 mM	HEPES, pH 7.9
10 mM	KCl
1 mM	DTT
1.5 mM	$MgCl_2$
Roche Complete	
proteinase inhibitor	

"high salt" buffer	
20 mM	HEPES, pH 7.9
0.2 mM	EDTA
1 mM	DTT
420 mM	NaCl
1.5 mM	$MgCl_2$
0.5 mM	PMSF
25% (v/v)	Glycerol
Roche Complete	
proteinase inhibitor	

3.2.3 Protein quantification

Nuclear extracts were analysed spectrometrically using a sensitive BCA Protein Assay Kit (Thermo Fischer) with a high tolerance rate for detergents and salts. Standard curve was prepared from bovine serum albumin (BSA).

3.2.4 SDS-Polyacrylamide Gel Electrophoresis (PAGE)

Protein samples were separated on a 10% SDS gel as described by Rittenhouse & Marcus (1984). Proteins were denatured in SDS sample buffer (95°C, 5 min), spun down and chilled on ice (5 min). Samples ran on stacking gel (4% polyacrylamide) at 80 V and were separated afterwards at 140 V. Running time depended on protein size. A pre-stained marker was used to control running of the gel.

Stacking gel (4%)

560 µl	Acrylamide-bis 30%
675 µl	0.5 M Tris-HCl, pH 6.8
675 µl	0.5 M imidazole, pH 6.8
75 µl	SDS, 10%
5 µl	TEMED
40 µl	APS, 10%

add dH_2O to 4.2 ml

Running gel (10%)

2.5 ml	Acrylamide-bis 30%
1.9 ml	1.5 M Tris-HCl, pH 8.8
75 µl	SDS, 10%
5 µl	TEMED
25 µl	APS, 10%

add dH_2O to 7.5 ml

1 x SDS running buffer

25 mM	Tris base
192 mM	glycine
1% (w/v)	SDS

3.2.5 Coomassie blue staining

Protein bands were visualized by incubation of the gel in Coomassie blue staining solution (1 hr), followed by destaining until background was minimized. Gels were kept in storage solution and dried when required.

Coomassie staining solution

0.25% (w/v)	Coomassie Brilliant Blue R-250
45% (v/v)	methanol
10% (v/v)	acetic acid
add dH_2O	

Destaining solution

45% (v/v)	methanol
10% (v/v)	acetic acid
add dH_2O	

Storage solution

5% (v/v)	methanol
5% (v/v)	acetic acid
add dH_2O	

3.2.6 Western Blot (tank blot)

After separation on SDS gels, proteins were blotted onto PVDF membranes (Immobilon-P, 0.45 µm, Millipore) following the Towbin tank blot protocol (Towbin et al., 1979). PVDF membranes were activated in methanol and equilibrated in blotting buffer. Gel and membrane were placed with two layers of Whatman paper and fibre pads inside a gel holder cassette, avoiding any air bubbles and placed into the pre-cooled blotting buffer tank. Blots were run for 1 hr at 100 V using cooling units.

1 x Blotting buffer

25 mM	Tris base
192 mM	glycine
10% (v/v)	methanol
adjusted to pH 7.6	

Membranes were saturated in blocking solution over night (4°C). Proteins of interest were detected by immunodetection using specific primary antibodies (cf. 2.2.6). Membranes were incubated with primary antibodies in blocking solution for 1 hr and with HRP-coupled secondary antibodies for 45 min at RT with gentle agitation. After extensive washing, membranes were incubated for 5 min in SuperSignal West chemiluminescent substrate (Pico or Femto, Thermo Scientific) and exposed to CL-X Posure Film (Thermo Fischer).

Blocking solution	Wash solution (1 x TBS-T)
0.5% (w/v) casein	100 mM Tris base
in 1 x TBS-T, pH 7.6	1.5 mM NaCl
filtered through 320 mm filter	0.03% (v/v) Tween-20
(Schleicher & Schuell, Dassel)	adjusted to pH 7.6

3.3 Cell biological and microbiological methods

3.3.1 Prokaryotic cells

Bacteria were used to generate and amplify plasmid DNA and cultured at 37°C either in LB Medium or on LB Agar plates. Antibiotics were added for specific selection of transformed bacteria. Glycerol cultures (15% (v/v) glycerol) were snap frozen in liquid nitrogen and stored at –80°C without antibiotics.

LB Medium	LB Agar
10 g Bactotryptone	15 g Bacto agar in 1000 ml LB Medium
10 g Sodium chloride	add appropriate antibiotics after cool down to
5 g Yeast extract	56°C
add dH_2O to 1000 ml, pH 7.0	
autoclave at 121°C for 20 min	

3.3.1.1 Generation of chemically competent cells

Competent bacterial cells were generated using a modified protocol by Hanahan for preparation and transformation of *E. coli* cells (Hanahan, 1983).

200 ml of LB Medium were inoculated with bacterial cells and grown to an OD_{600} of about 0.4 at 37°C. After 20 min of incubation in an iced water bath with gentle agitation, cells were harvested at (4000 x g, 15 min, 4°C,). The pellet was resuspended in 10 ml iced $MnCl_2$-transformation buffer, and kept on ice for 10 min, followed by another centrifugation step (3000 x g, 10 min, 4°C). The pellet was resuspended in 7.4 ml $MnCl_2$-transformation buffer and mixed gently, followed by dropwise addition of 560 µl DMSO. Aliquots of 100 µl were prepared on ice and snap frozen in liquid nitrogen. Competent cells were stored at –80°C. Transformation efficiency of generated competent cells was routinely tested using pUC19 as control vector.

$MnCl_2$-transformation buffer

10 mM	HEPES, pH 6.8
15 mM	$CaCl_2$
20 mM	KCl
55 mM	$MnCl_2$

3.3.1.2 Transformation of competent cells

An aliquot of competent cells (100 µl) was incubated in an iced water bath for 3 min, mixed with 50 ng of transforming DNA (volume ≤5%) and kept on ice for 30 min. Cells were heat-shocked at 42°C for 30 sec, and briefly cooled on ice for 2 min. Pre-warmed LB-Medium (250 µl) was added and cells were shook for 45 min at 37°C (225 rpm) for recovery and expression of the antibiotic resistance marker encoded by the plasmid. Thereof, 50 and 150 µl were plated onto antibiotic agar plates using a freshly prepared bent glass rod and incubated over night at 37°C.

3.3.2 Eukaryotic cells

Eukaryotic cell lines were used to analyse generated reporter gene plasmid containing promoter fragments towards their transcriptional activity using transient transfection (cf. 3.3.2.3). Different stimulatory regimes were tested in cultured cells to detect regulatory effects on gene expression. Cells were also used for ChIP experiments (cf. 3.1.14) to identify *in vivo* interactions of proteins with selected chromatin regions. Nuclear proteins extracted from eukaryotic cells were used in EMSA experiments (cf. 3.1.13) to analyse DNA/protein interactions *in vitro*.

3.3.2.1 Eukaryotic cell culture

The human embryonic kidney cell line HEK293T, the vascular endothelial cell line EA.hy926 and the African green monkey kidney cell line COS-7 were maintained in DMEM containing 10% (v/v) fetal calf serum (PAA), 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM/ml L-Glutamine (all GIBCO). For cultivation of HEK293T iron-supplemented fetal calf serum was used (Cell Concepts). The monocytic cell line THP-1 was maintained in RPMI 1640 medium (Roswell Park Memorial Institute) containing 10% (v/v) fetal calf serum (PAA), 100 U/ml penicillin, 100 µg/ml streptomycin, 2 mM/ml L-Glutamine and 1 x modified Eagle's medium amino acid solution (Sigma). THP-1 cells were kept at a concentration of 0.5 to 1 x 10^6/ml. Differentiation of THP-1 cells was performed by stimulation with 10^{-8} m PMA for 72 hrs. For stimulation of cells 0.5 mM 8-bromo-cAMP (Biolog), 10^{-8} PMA (Sigma) or 5–20 ng/ml recombinant cytokines TGF-β1, IL-1β and PDGF (Calbiochem) were used. When state of confluence was reached, cells were detached from surface by trypsination and splitted at appropriate ratios for further cultivation. The number of passages did not exceed 40 in any case.

3.3.2.2 Storage

For long term storage, cells were washed twice with PBS, trypsinated, and transferred to fresh medium. After centrifugation, cells were placed on ice and resuspended in fetal calf serum containing 10% (v/v) DMSO. Cryo tubes were cooled slowly to –80°C and transferred to liquid nitrogen the next day. Thawing of cells was performed in a water bath at 37°C. To

remove DMSO from the freezing medium, cells were washed with PBS and transferred into pre-warmed medium after centrifugation.

3.3.2.3 Transient transfection

Cell lines EA.hy926 and THP-1 were transfected using Nanofectin transfection reagent (PAA). Nanofectin consists of two components, a positively charged polymer with DNA-binding capacity which is embedded into a porous nanoparticle. In case of EA.hy926 cells, 10^5 cells/well were plated in a 24-well plate format and transfected the next day after a change of medium 2 hrs prior addition of transfection complexes. For both, EA.hy926 and THP-1 cells, 1 µg DNA and 3.2 µl Nanofectin solution were diluted in 50 µl NaCl solution (150 mM) and incubated for 10 min at RT. Both preparations were mixed and incubated for 30 min at RT. Transfection complex was added dropwise to cell medium. In case of EA.hy926 cells, transfection reagent was removed by change of medium after 4 hrs. Cells were harvested 24 hrs post transfection with 100 µl Passive Lysis buffer (Promega). Luciferase activity was determined using a Sirius single-tube luminometer (Berthold detection systems). The cell lysate/luciferase substrate ratio was routinely 20 µl/75 µl.

Transfection of the promoter constructs was performed in equimolar amounts of reporter vectors using the inert vector p0GH (Nichols Institute) for adjustment of DNA content. The pGL3-Control vector in which transcription is driven by the competent SV40 viral promoter and additional enhancer served as positive control. Promotorless pGL3-Basic vector served as empty vector control. Transfection experiments were repeated at least three times, in triplicates for each plasmid.

3.3.2.4 Cotransfection

Transient overexpression of proteins was used to analyze possible effects on transcription of the cotransfected reporter gene vectors containing promoter fragments. Expression vector to reporter gene vector ratio was routinely 3:1. In this study, expression vectors for members of the C/EBP family, CBP and SP1 were used.

3.4 Study populations

The current investigation was based on the Münster Molecular Functional Profiling for Mechanism Detection (MolProMD) study. Genomic DNA from patients of this study was mainly used for the detection of existing genetic variants by sequencing as well as for subcloning, generating gene promoter reporter vectors. The Münster MolProMD Study is a prospective study of patients with CVD (MI, essential hypertension) and families, aimed at studying molecular genetic mechanism of CVD. The study was approved by the ethics committee of the Medical Faculty, Westphalian Wilhelms-University of Münster and written informed consent was obtained from all study subjects.

3.5 Computational sequence analyses

Web based algorithms were used for *in silico* DNA sequence analyses. For phylogenetic and gene family footprinting, sequences were aligned using the ClustalW program at EMBL (http://www.ebi.ac.uk/Tools/clustalw/index.html). For prediction of TF binding sites, DNA sequences were compared to the TRANSFAC 7.0 database by using AliBaba2.1 (http://www.gene-regulation.com/pub/programs/alibaba2/index.html) and PROMO (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3). PROMO (accessing TRANSFAC 8.3) was also used to analyse regulatory sequences of different genes expressed in the same tissue using the MultiSearch site function and for phylogenetic footprinting with comparison of TF binding sites. Furthermore, PROMO was used for consensus sequence identification and matrix visualization to design oligos with defined TF binding qualities serving as controls in EMSA experiments. Settings of the used algorithms are available upon request (Chenna et al., 2003; Grabe, 2002; Messeguer et al., 2002).

4 RESULTS

4.1 Phylogenetic footprinting

To reveal nucleotides under evolutionary selection in the *BGN* 5'-flanking region and 5'-UTR, phylogenetic footprinting was performed by comparison of human (Acc#: NT_167198.1) and mouse (Acc#: NT_039706.7) sequences using ClustalW at http://www.ebi.ac.uk. Evolutionary conservation of sequences over longer time periods is assumed to be the result of negative selection, implying identical function and expression of the compared orthologous genes (Tagle et al., 1988). Differential phylogenetic footprinting will potentially be observed if a change in expression patterns occurred during the evolutionary process (Fischer & Backendorf, 2007). Results obtained by this method depend considerably on the species used for comparison. Sequences from species with small evolutionary distances may well not reveal enough mutations outside of *cis*-elements whereas species which are evolutionary more distant will not reveal any residual homology. Phylogenetic footprinting may furthermore not be used unrestrictedly to identify regions containing *cis*-active elements. Since the structure of the genome is highly variable between organisms and control regions may well be localized in intronic regions, downstream, or in considerable distance of a gene (Trinklein et al., 2007; Kleinjan et al., 2005, Lower et al., 2009), control regions are not annotated in databases and may be tracked only by their flanking gens. Different arrangements of genes (i.e. copy number, orientation) in different organisms introduce considerable uncertainty to this procedure. For *BGN*, sequence homology of ~70% was observed for the 5'-UTR between human and mouse, whereas an overall score of ~65% was reported for adjacent upstream 1150 bp (fig. 10). A notable sequence homology between the two species of ~80% was detected between the proximal position -250 and +25 of the human *BGN* sequence. Sequence analyses of the 5'-flanking region using different web-based algorithms (cf. references for details) did not reveal a TATA or CAAT box, but multiple GC elements.

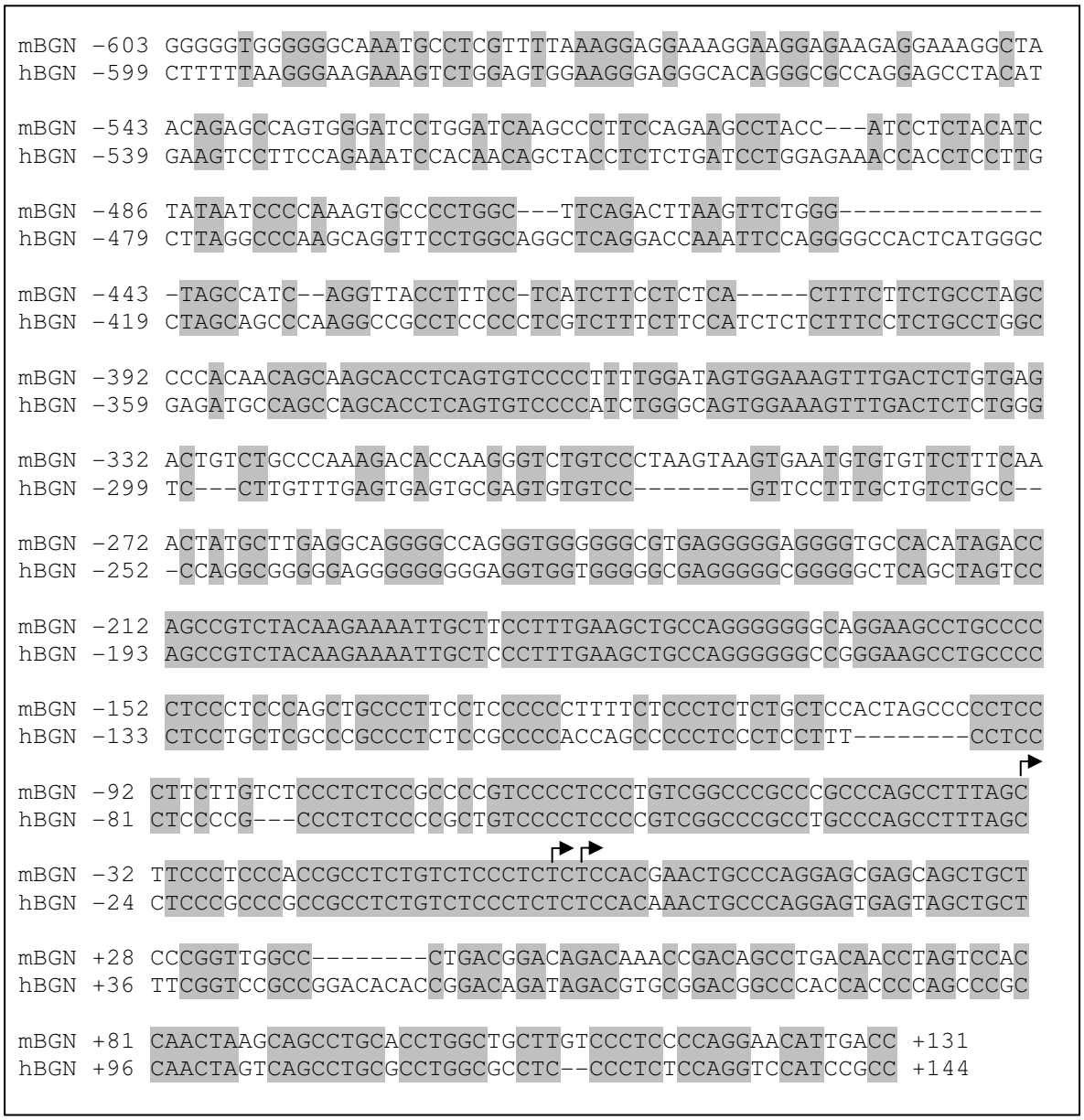


Figure 10: Comparison of the mouse and human *BGN* 5'-flanking region and 5'-UTR.

Conserved domains are marked in gray. Main human TSS according to DBTSS are indicated by arrows. The distal part of the human sequence, position -599 to -230, is not well conserved (conservation score ~20%) but holds a highly conserved region of 50 bp at position -352 to -302. The proximal part of the 5'-flanking region and the 5'-UTR have an overall conservation of ~80%. Alignment was conducted using ClustalW.

4.2 Identification of cell lines for *BGN* promoter studies

For *in vitro* functional analyses of gene promoters, a cell system expressing the gene of interest is mandatory since a defined set of TF accounts for sequence-specific binding of accessible *cis*-active promoter elements and coordinated gene expression control (cf. 1.5). Cell lines differ in their TF constitution and selection of a cell system missing on endogenous expression can lead to artificial results. Concerning the biological system or pathophysiology the gene product is hypothetically involved in, cell lines known to play pivotal roles in these processes should be screened routinely.

4.2.1 Endogenous *BGN* expression analysis

BGN is known to be expressed ubiquitously throughout the human body but does posses different functional properties. It has been shown to be a major component of the ECM and its potential as a proinflammatory factor has been reported recently (for details cf. 1.3). We therefore analyzed the endogenous expression of *BGN* in the embryonic kidney cell line HEK293T, an African green monkey kidney cell line COS-7, a differentiated vascular endothelial cell line EA.hy926 and the monocytic cell line THP-1. Standard transcript detection was performed using a SS primer in exon two (5'-TGACACCTCGGGCGTCCTGG-3') and an AS primer in exon four (5'-GAGCTGGGTAGGTTGGGCGG-3') at 28 cycles in a touchdown PCR with cDNA template (amplicon length 348 bp). All cell lines tested positive for *BGN* endogenous expression (fig. 11), in good consent with its ubiquitous expression *in vivo*. To address existing stimulatory aspects in *BGN* gene expression, cells were incubated with cAMP and the phorbol esther PMA. Both compounds elicit broad inflammatory signal transduction pathways. cAMP induces the PKA pathway while PMA leads to activation of the PKC signal cascade. Stimulation of THP-1 monocytes with PMA (78 hrs) additionally leads to differentiation into macrophages. Both stimulatory regimes did not alter the *BGN* transcript level in our experiments (fig. 11).

4.2.2 Identification of TSS

Current databases such as DBTSS (Wakaguri et al., 2008; Ota et al., 2004) hold precise positional information on TSS of eukaryotic mRNA, since evolution of high-throughput technologies allowed massive sequence analysis of full-length cDNA in the recent past (Yamashita et al., 2010). Lately, a continuative refinement has been integrated to take differential usages of the TSS, depending on the cellular circumstances, into account. Tags were collected from a total of 33 different cell types or culturing conditions. For *BGN* (Entrez gene ID 633) DBTSS reports three main TSS in normal adult tissue (fig. 10). To identify the cell type-specific usage of TSS in the human *BGN* gene promoter, 5'RACE experiments and diacritic PCR were used. We were able to show the alternative usage of two TSS (fig. 11) in correlation with DBTSS data and termed the identified proximal position TSS1 and the more distal position TSS2. The flexible structure of transcription initiation was as well addressed with respect to different stimulatory regimes. TSS1 was shown to be actively involved in transcription in all tested cell lines independently of the applied stimulation, while parallel usage of TSS1 and TSS2 was exclusively detected in EA.hy926 cells. Stimulation of EA.hy926 cells with phorbol ester PMA lead to concentration of transcriptional initiation upon TSS1 and abrogation of usage of TSS2 (fig. 10).

We identified four cell lines expressing the human *BGN*. One of them, the vascular endothelial cell line EA.hy926, featuring an alternative TSS (TSS2, fig. 10). Considering the dual nature of BGN as a structural protein as well as an inflammatory factor, we decided to focus on the EA.hy926 and THP-1 cell lines for further investigations.

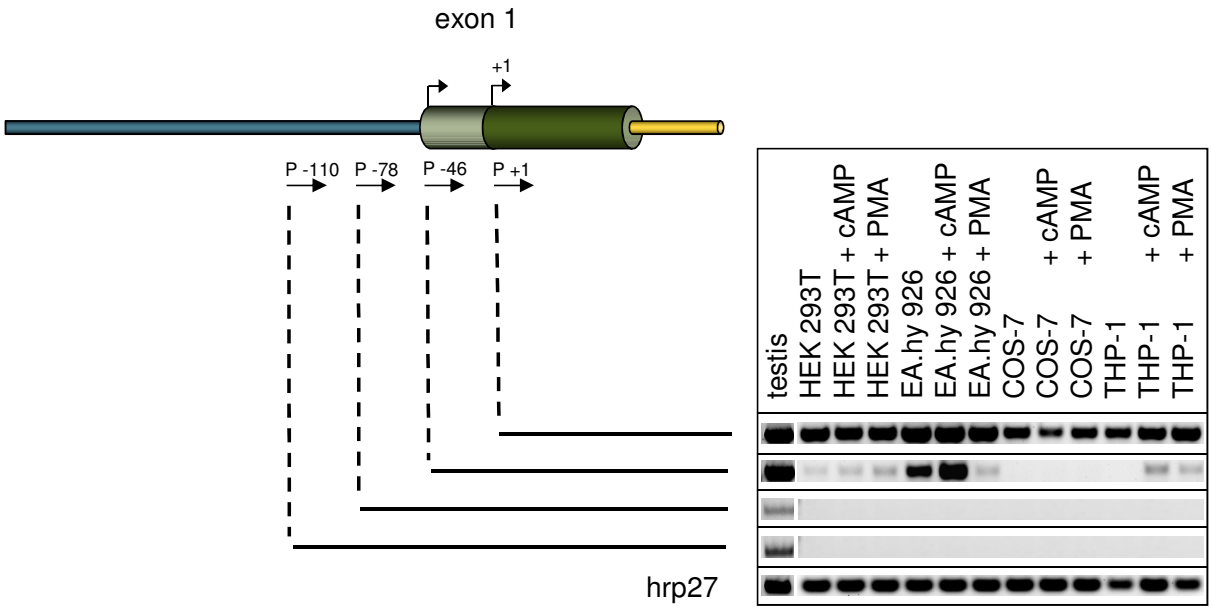


Figure 11: Analysis of *BGN* TSS in different cell lines.

The schematic illustration shows exon one and the upstream 5'-flanking region. Notation of primers below is based on the main identified transcription start site, TSS1. Total RNA was used for cDNA synthesis followed by semi-quantitative PCR with the indicated 5'-primers and 3'-primer P+623. Integrity of cDNA was tested by amplification of hRP27. RNA from human testis tissue served as positive control. For stimulation, cells were grown in medium with either 8-bromo-cAMP (0.5 mM) or PMA (10^{-8} M) for 24 hrs.

4.3 Characterization of *BGN* promoter transcriptional activity

For *BGN* promoter characterization, its transcriptional potential and the identification of regulatory sequence regions within, deletion constructs were generated. Since the 5'-flanking region of the human *BGN* gene does not feature a TATA box or other comparable rather focused promoter elements to drive its expression, a CpG island sequence analysis (Takai & Jones, 2002) was performed to identify regions with potential promoter properties *in silico* (fig. 12). Since the sequence portion of 336 bp predicted to be a potential CpG island spans almost completely the 5'-UTR and adjacent ~200 bp of 5'-flanking region, the untranslated region of exon one was included in the generated promoter constructs. Serial truncation of deletion constructs was planned taking further *in silico* predictions for TFBS into consideration (for details cf. 3.5). Sequence positions with highly conserved TF binding clusters were used as markers for each subsequent shortening step of fragments (for schematic representation of all deletion constructs and nomenclature see fig. 8). Finally, a fragment representing ~1300 bp of the 5'-flanking region with inclusion of the 5'-UTR was selected to be analysed in reporter gene assay analyses.

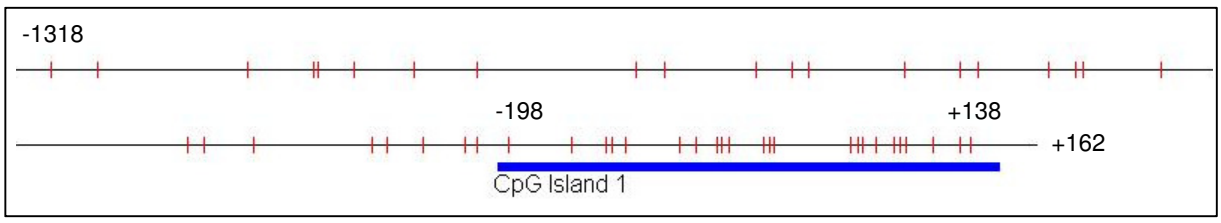


Figure 12: CpG island analysis of the human *BGN* 5'-flanking region and 5'-UTR.

We identified one CpG-rich region spanning position -198 to +138 (336 bp), underlined in blue. Vertical lines in red mark cytosine/guanine phosphodiesters. Sequence positions are based on TSS1. Position -1318 and +162 mark the longest generated promoter deletion construct. The web based application CpG Island Searcher (http://www.cpgislands.com) was used with the following settings: selected lower limits: %GC=60, ObsCpG/ExpGpG=0.65, length=100, distance=100.

4.3.1 Reporter gene assays

Transient transfection of generated *BGN* promoter deletion constructs in EA.hy926 cells revealed a region with high transcriptional activity in the 5'-flanking region of the gene (fig. 13). The region between -1231 and -935, covered by three promoter constructs, holds strong transcriptional activity reaching the level of pGL3-Control in our experiments. Truncation of a 234 bp fragment resulted in a significant ($p<0.001$) decrease of transcriptional activity. The constructs -701/+162 and -386/+162 covering the predicted CpG region of 336 bp (fig. 12) hold only limited potential to drive gene expression. Further truncation to a promoter fragment representing 39 bp of the 5'-region (-39/+162) but including the 162 bp of the 5'-UTR, significantly ($p<0.05$) regenerated promoter transcriptional activity. Notably, the promoter construct -1318/+162 did not display any promoter activity at all (RLU beyond baseline level). In part restoration of transcriptional activity was achieved by truncation of the 162 bp 5'-UTR generating the fragment -1318/+1.

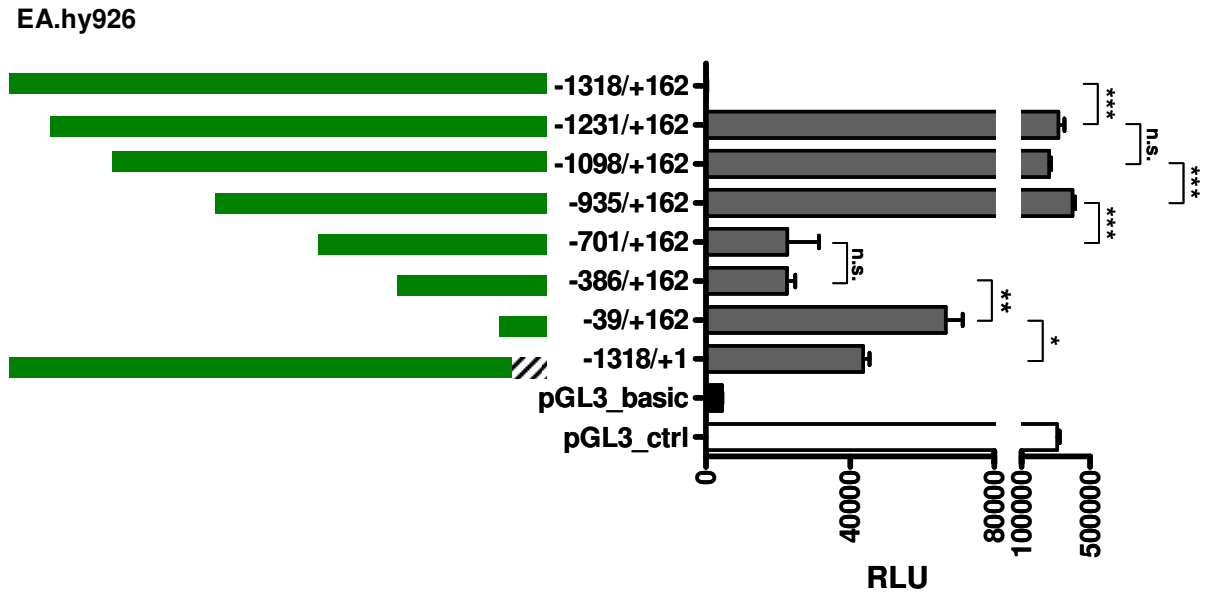


Figure 13: Transcriptional activity of selected *BGN* promoter fragments.

BGN promoter fragments expressing luciferase gene were transiently transfected into EA.hy926 cells to identify transcriptionally active regions within the 5'-flanking region. Length of the fragments is schematically indicated by filled boxes (left), the 5'-UTR is marked by a shaded box. The positive control pGL3-Control (white bar) contains a strong viral promoter and enhancer, the negative control pGL3-Basic (black bar) lacks any active promoter. The region between -1231 and -935 holds a strong transcriptional activity reaching activity of pGL3-Control, whereas the proximal region starting at position -701 exhibits declined transcriptional activity. The fragment spanning -1318 to +162 including the 5'-UTR (top) shows no transcriptional activity. Excision of the 5'-UTR (bottom) leads to partly reconstitution of activity. Levels of significance are shown by asterisk on the right (n.s.: not significant, *P<0.05, **P<0.01, ***P<0.001). RLU: Relative Light Units.

4.4 Verification of *BGN* gene variants and MolHaps in the MolProMD study

To reveal the composition of identified promoter polymorphisms on single DNA strands and to define MolHaps, genomic DNA from 57 patients of the MolProMD study were screened. One-thousand twenty-five bp of the promoter region harbouring G-578A, G-151A, and G+94T were amplified, subcloned, and sequenced twice (both DNA strands with sense and antisense primers) using an automated sequencing device. Three MolHaps were verified according to previous assumptions: BGN-MolHap1 [G^{-578}-G^{-151}-G^{+94} (wt)] BGN-MolHap2 [G^{-578}-A^{-151}-T^{+94}] and BGN-MolHap3 [A^{-578}-G^{-151}-G^{+94}]. BGN-MolHap1 has been designated wt since ancestral allele information was only available for position G-578A, with G mutated to A. The variants G-849A and C-501A were not detected in our 57 samples from the MolProMD study.

4.4.1 *BGN* MolHap promoter fragments in reporter gene assays

Molecular promoter haplotypes have been shown to significantly influence transcriptional activity of different genes in reporter gene assays (Hagedorn et al., 2009; Dördelmann et al., 2008; Hasenkamp et al., 2008). For the analyses of *BGN* MolHaps, a 1025 bp fragment was cloned into pGL3-Basic vector, harbouring either of the three indentified SNP combinations (cf. 4.4) and transiently transfected into EA.hy926 and THP-1 cells. Previous analyses in HEK293T cells pointed to a higher transcriptional activity of MolHap3 compared to MolHap2 (Rüssmann, MD thesis) but subsequent EMSA experiments were not sufficiently conclusive to reveal the underlying nature of protein/DNA interaction. The analysed promoter fragment, spanning position -893 to +132, exerted significant transcriptional activity (~250000 RLU in Ea.hy926) comparable to identified fragments with highest transcriptional activity in EA.hy926 promoter deletion studies (4.3.1), when wt alleles were present. Introduction of MolHap2 and MolHap3 led to a significant reduction of transcriptional activity in EA.hy926 cells ($P<0.001$), with total abrogation beyond the level of pGL3-Basic for MolHap3 ($P<0.001$, fig. 14). A consistent result was observed upon transfection in the monocytic cell line THP-1. Fragments harbouring variants corresponding to MolHap2 and MolHap3 lead to a significant reduction of transcriptional activity (P-values<0.01). The same effects observed under basic conditions were also confirmed under the stimulatory regime of cAMP and PMA (not shown).

Both compounds did not have any effect on the transcriptional activity of the wt fragment in neither EA.hy926 nor THP-1 cells.

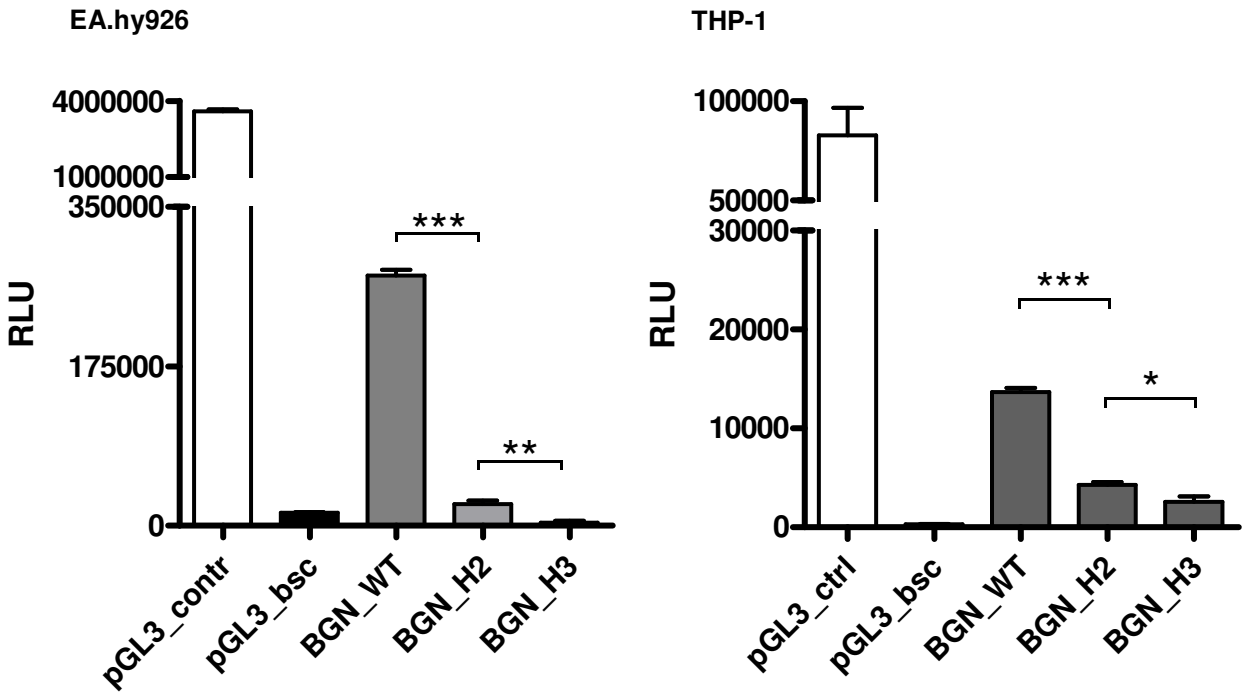


Figure 14: Transcriptional activity of *BGN* MolHaps.

BGN MolHaps expressing luciferase gene were transiently transfected into EA.hy926 (left) and THP-1 cells (right) to identify the influence of SNP combinations on transcriptional activity. The wt fragment, representing BGN-MolHap1 [G^{-578}-G^{-151}-G^{+94} (wt)], exhibited significant transcriptional activity in both cell lines. Transcriptional activity of BGN-MolHap2 [G^{-578}-A^{-151}-T^{+94} (H2)] was significantly reduced, whereas introduction of BGN-MolHap3 [A^{-578}-G^{-151}-G^{+94} (H3)] led to total abrogation of transcriptional activity in EA.hy926 cells and further reduction in THP-1 cells. The positive control pGL3-Control (white bar) contains a strong viral promoter and enhancer, the shuttle control pGL3-Basic (black bar) lacks any active promoter. Levels of significance are shown by asterisk (*P<0.05, **P<0.01, ***P<0.001). RLU: Relative Light Units.

4.4.2 Analysis of *BGN* promoter SNPs in transfection experiments

For in-depth characterization of each genetic variant in the context of its narrowed promoter region, fragments containing each variant with minimal surrounding sequence portions cloned into the pGL3-Promoter vector (kind gift of Christina Rüssmann) were used. This vector of the pGL3-System, which contains an SV40 promoter, was initially generated to analyse isolated promoter portions towards their potential as individual enhancer elements. The six vectors contained either allele at position G-578A in a 187 bp fragment, the alleles at position G-151A in a 188 bp fragment and alleles at position G+94T in a 158 bp fragment. The generated vectors were transiently transfected into EA.hy926 cells (fig. 15). Position G-578A was able to slightly enhance activity of the pGL3-Promoter vector. A mild enhancing effect was also observed for -151A, whereas the wt allele -151G reduced transcriptional activity below empty pGL3-Promoter vector activity. Neither allele present at position G+94T did exert any enhancing activity in the present transfection experiments (all P-values <0.01).

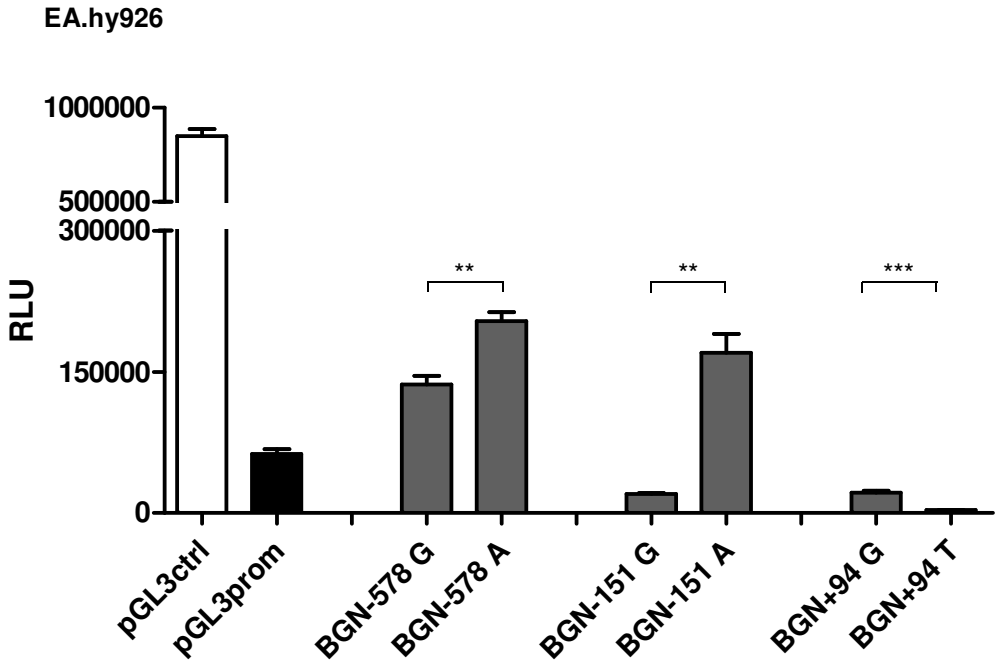


Figure 15: Transient transfection of *BGN* promoter fragments harbouring insulated variants.

Isolated variants were transiently transfected into EA.hy926 cells. Both alleles at position G+94T reduced basic transcriptional activity of the pGL3-Promoter vector. Introduction of the A allele at position G-151A and both alleles at position G-578A enhanced transcriptional activity of pGL3-Promoter. The positive control pGL3-Control (white bar) contains a strong viral promoter and enhancer, the shuttle control pGL3-Promoter (black bar) contains an SV40 promoter without enhancer. Levels of significance are shown by asterisk (**P<0.01, ***P<0.001). RLU: Relative Light Units.

4.4.3 *In silico* analysis of *BGN* promoter regions

An extensive *in silico* analysis was performed for the characterized *BGN* promoter region (4.3.1) with special respect to alterations in TFBS prediction upon introduction of alleles corresponding to BGN-MolHap1-3. A preliminary analysis performed by Rüssmann (MD thesis) using AliBaba2.1/TRANSFAC 7.0 revealed, in addition to a few other TF, SP1 TFBS for all polymorphic positions. With SP1 being the most frequent TF overall predicted for the entire GC-rich *BGN* promoter, its predicted binding to the polymorphic regions seemed mandatory. To correct for random prediction of TF due to sequence characteristics, the web based tool PROMO accessing TRANSFAC 8.3 was used. The tool allows to measure the reliability of predictions by using a random sequence of 1000 nucleotides, considering both, a model with equiprobability of the four nucleotides and a model with the same nucleotide frequency as the query sequence, represented by Query Expectation Values (QEV). The results shown in table 7 were obtained using Matrix Dissimilarity Rates (MDR) ≤6% (i.e. 94% identity to TF consensus sequence). Position G-578A showed binding for the stress-related protein p53, the B-cell linage-specific activator PAX-5 and a member of the ETS TF family (c-Ets1). TFBS prediction for the proximal position G-151A showed condensed binding for SP1/SP3 and the B-/T-cell-specific factor LEF-1, with an additional ETS binding site. The variant G+94T positioned in the 5'-UTR was predicted to generate a binding site for AP-1 TF complex by detection of c-FOS/c-JUN consensus motifs. With AP-1 being not a single protein, it constitutes a group of related dimeric basic region-leucine zipper proteins that belong to the JUN, FOS, MAF and ATF subfamilies (Shaulian & Karin, 2002).

Table 7: *In silico* prediction of TFBS. Analyses were performed with sequence input of 60 bp contain the identified genetic variants using PROMO accessing database TRANSFAC 8.3.

-578G		-578A		-151G		-151A		+94G		+94T	
Name	**QEV**	**Name**	**QEV**	**Name**	**QEV**	**Name**	**QEV**	**Name**	**QEV**	**Name**	**QEV**
TFIID	0.011	TFIID	0.015	LEF-1	0.000	LEF-1	0.000	HIF-1	0.003	HIF-1	0.003
c-Ets1	0.020	P53	0.021	c-Ets1	0.002	c-Ets1	0.002	E2F-1	0.003	c-FOS	0.004
P53	0.022	c-Ets1	0.021	SP1/SP3	0.002	SP1/SP3	0.002	c-FOS	0.004	c-JUN	0.005
P53	0.031	P53	0.029	NF-E4	0.002	NF-E4	0.002	c-JUN	0.004	CP2	0.010
Pax-5	0.040	Pax-5	0.037	Pax-4a	0.005	C/EBP-β	0.006	Yi	0.004	NF-1	0.010
MDR: 6% species: human		MDR: 6% species: human		MDR: 5% species: all		MDR: 5% species: all		MDR: 5% species: all		MDR: 5% species: all	

4.5 Overexpression of TF SP1

In silico TFBS prediction for the *BGN* promoter in general revealed a number of consensus sequences for TF SP1. The TF SP1 is ubiquitously expressed, possesses three C_2H_2-type zinc fingers as DNA-binding domain (DBD) and has two glutamine-rich transactivation domains (TAD; Kadonaga et al., 1987; Kadonaga et al. 1988; Courey et al., 1989). Binding was predominantly predicted in conserved regions (fig. 10), spanning the proximal positions -230 to -200 and the region -25 to +5 flanking the main TSS (all matrix dissimilarity rates =0, QEV ≤0.2). In addition, SP1 binding was predicted for the three identified genetic variants composing *BGN* MolHaps. Based on these findings, we performed co-expression experiments in EA.hy926 cells using an expression vector for SP1 (pSP1/CMV). This mammalian expression vector is under the control of a promoter-regulatory region of the human cytomegalovirus for optimal performance. Effector (pSP1/CMV) to reporter gene vector (*BGN* promoter constructs) ratio was 1:3 in EA.hy926 transient transfections. The empty vector pCMV served as shuttle control. Strong induction of *BGN* promoter transcriptional activity in the presence of overexpressed SP1 was observed for all deletion constructs (fig. 17, all FI values ≥2.3). Explicitly high FI values over empty shuttle control were observed in promoter regions predicted to harbour conserved SP1 TFBS. Fragment -39/+162, including the main TSS flanking region and 5'-UTR, featured a 3.5-fold increase in transcriptional activity, while fragment -1318/+1 covering the distal part of the promoter and lacking the 5'-UTR presented a 4.1-fold increase in transcriptional activity. The section between -1231 and -935, covered by three promoter constructs with highest absolute transcriptional activity, displayed a mean induction of transcriptional activity upon SP1 co-transfection by 2.3-fold.

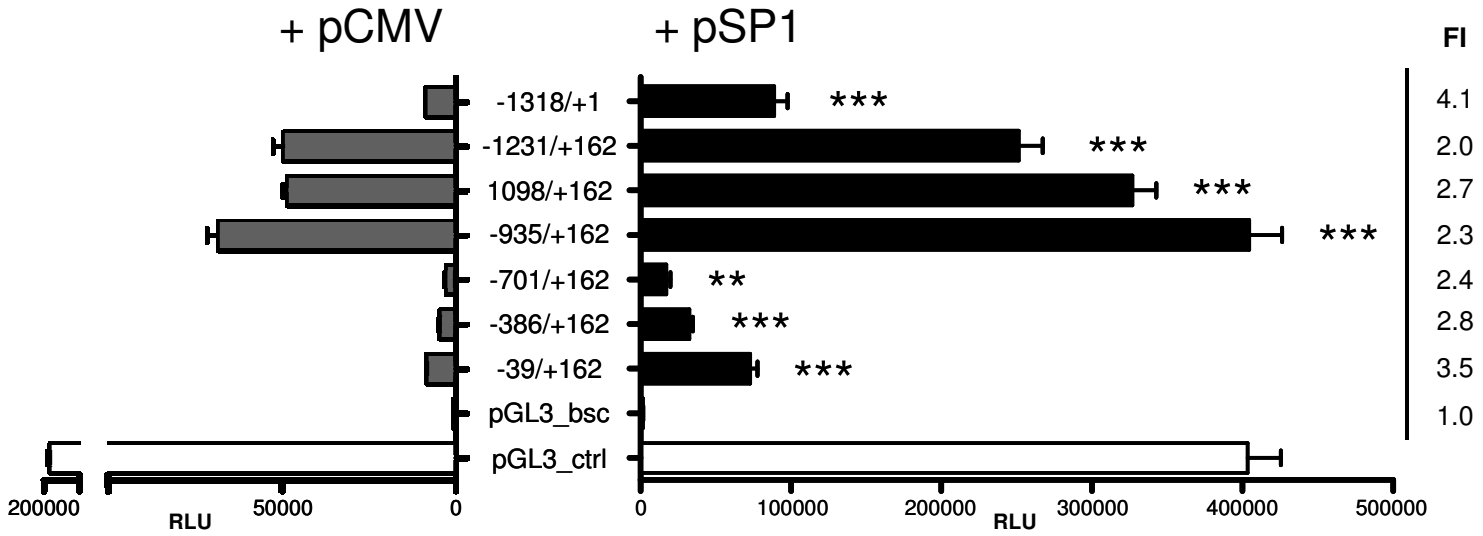


Figure 17: Co-expression of TF SP1 and *BGN* promoter constructs in EA.hy926 cells.

Significant induction of *BGN* promoter deletion constructs transcriptional activity was observed upon overexpression of TF SP1. The empty vector pCMV served as shuttle control (gray bars, left), corresponding transcriptional activity in presence of overexpressed SP1 is represented by black bars (right). All results reached high levels of significance ($P \leq 0.01$). Each constructs' relative activity over the empty shuttle vector pGL3-Basic was calculated and expressed as fold induction (FI) on the far right. High levels of transcriptional activity induction were observed for fragments -39/+162 (FI by 3.5) and-386/+162 (FI by 2.8) in correspondence with SP1 TFBS predictions in these regions. The promoter portion covering positions -1318/+1 (no 5'-UTR) reached the highest level of FI (by 4.1). Levels of significance are shown by asterisk (**P<0.01, ***P<0.001). RLU: Relative Light Units. pGL3-Control is represented by white bars.

4.6 ChIP experiments

Since overexpression of TF SP1 resulted in strong induction of *BGN* promoter transcriptional activity, ChIP experiments were performed for *in vivo* verification of these results. Stimulation of *BGN* transcriptional activity might not originate from direct binding of SP1 to the promoter itself, but could exert indirect regulatory effects. Demonstration of SP1 binding to the *BGN* 5'-flanking region in ChIP experiments would verify direct physical interaction of SP1 with the *BGN* promoter.

Proteins interacting with chromatin were chemically cross-linked using formaldehyde followed by sonification of DNA. Precipitation of bound promoter segments was achieved by the use of selected specific antibodies against SP1, GATA1 and phosphorylated CREB. Three distinct regions of the *BGN* promoter were amplified in PCR reactions following ChIP, with mean amplicon length of 100 bp. Neither binding of GATA1 nor binding of CREB was detected at any of the analysed sections, while binding of SP1 was demonstrated for region -28 to +71 and -918 to -806 (fig. 18). ChIP results were in correspondence with co-expression experiments since the region flanking the main TSS and the more distal parts showed strong induction of transcriptional activity upon SP1 overexpression (fig. 17). The promoter portion spanning -1164 to -1066 did not show binding of TF SP1 in ChIP experiments.

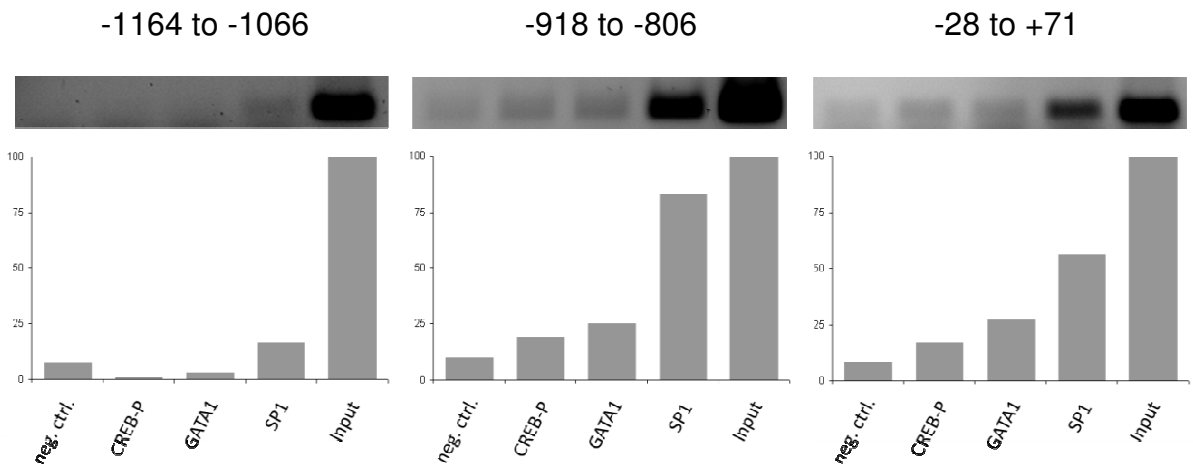


Figure 18: TF SP1 interacts selectively with *BGN* promoter portions in EA.hy926 cells.

ChIP was used to analyse DNA/protein interactions *in vivo*. Portions of 100 bp were amplified using specific oligonucleotides located in the *BGN* promoter region. Sonicated chromatin was used as positive input control and determined 100% for densitometric PCR signal intensity evaluation (bars below). Amplified promoter regions are depicted above with corresponding positions. Interaction of SP1 with the *BGN* promoter was detected in the region -28 to +71 flanking the main TSS as well as the distal part spanning -918 to -806. Position -1164 to -1066 did not interact with SP1 in EA.hy926 cells. TF GATA1 and phosphorylated CREB did not bind to any of the analysed regions.

4.6.1 ChIP analysis of *BGN* MolHaps

In vivo binding of TF SP1 was also tested for polymorphic regions of the *BGN* promoter (fig. 19) since *BGN* MolHaps2 and 3 showed significant influence on transcriptional activity which could be due to differential binding of SP1. EA.hy926 cells exhibit *BGN* MolHap1 [G^{-578}-G^{-151}-G^{+94} (wt)] and polymorphic regions were amplified in PCR reactions following ChIP, with mean amplicon length of 100 bp. Interaction of SP1 was shown in ChIP experiments for the proximal position G-151 and the 5'-position G+94 as predicted by computational TFBS analysis. ChIP experiments in EA.hy926 cells did however not reveal *in vivo* binding to position G-578. Neither binding of GATA1 nor binding of phosphorylated CREB was detected at any of the analysed polymorphic positions.

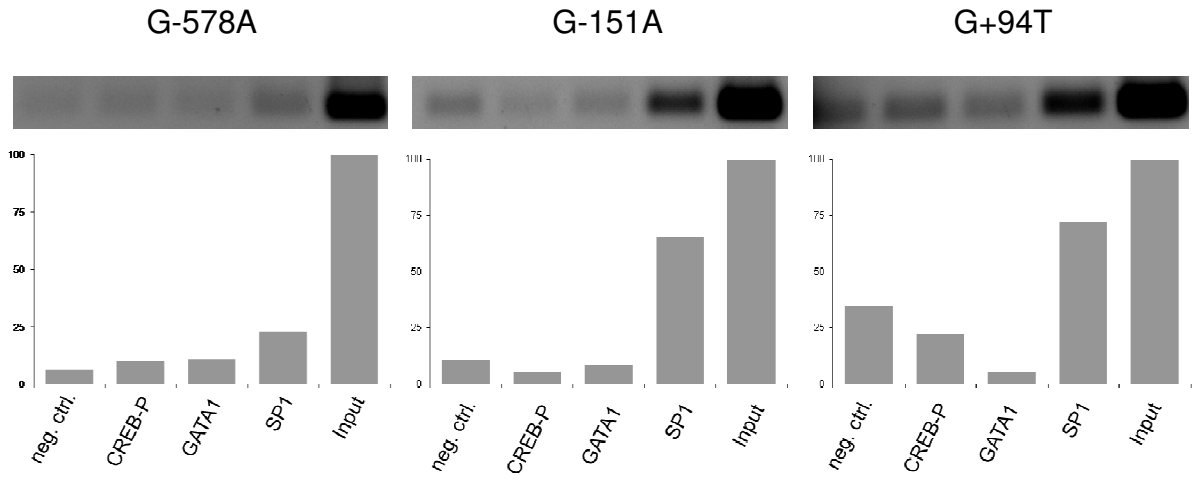


Figure 19: TF SP1 interacts with polymorphic *BGN* promoter regions in EA.hy926 cells.

Interaction of SP1 was detected at the proximal position G-151A and the 5'-UTR position G+94T, but not for the more distal position G-578A. Portions of 100 bp were amplified using specific oligonucleotides flanking polymorphic regions. Sonicated chromatin was used as positive input control and determined 100% for densitometric PCR signal intensity evaluation (bars below). TF GATA1 and phosphorylated CREB did not bind to any of the analysed regions.

4.7 TGF-β1 signal transduction effects on *BGN* gene expression

The multifunctional TGF-β family has been described to involve key regulators of ECM assembly and remodeling. At the cellular level, TGF-β cytokines inhibit proliferation of epithelial, endothelial, and haematopoietic cells and regulates their differentiation. Among this large family with at least 42 members encoded in the human genome, TGF-β1 has been reported to regulate BGN on transcriptional and posttranscriptional levels (Schmidt et al., 2006; Tiede et al., 2010) and thereby influence its availability and protein characteristics (Little et al., 2008). Most of the currently available data has however come from studies concentrating on the role of proteoglycans in fibrotic processes, using fibroblast cell lines. Since cytokines from the TGF-β subfamily are known for their remarkable regulatory versatility, a detailed analysis of TGF-β1 effects on our cell system was mandatory.

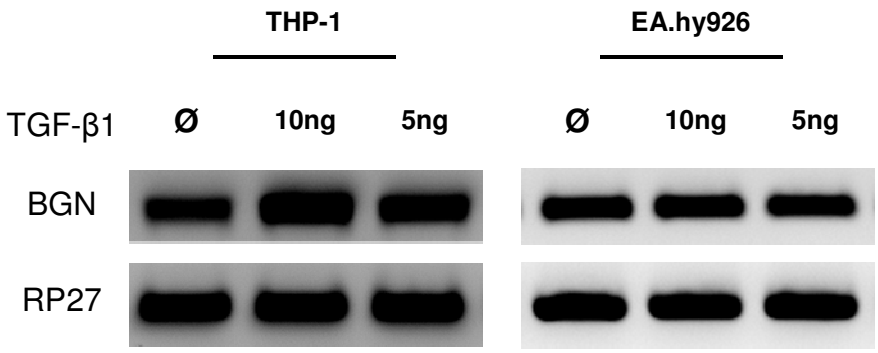


Figure 20: TGF-β1 stimulates *BGN* mRNA expression in THP-1 monocytes.

EA.hy926 cells and THP-1 monocytes were stimulated with TGF-β1 for 24 hrs (5 and 10 ng/ml). No effect upon stimulation was observed in EA.hy926 cells while THP-1 monocyte *BGN* mRNA expression was upregulated depending on cytokine concentration. Integrity of cDNA was tested by amplification of hRP27. Ø: mock control.

Semi-quantitative PCR analysis of BGN mRNA levels upon TGF-β1 stimulation revealed an activating effect in THP-1 monocytes, while EA.hy926 cells did not respond to the cytokine (fig. 20). To execute its intracellular signal transduction capacity, TGF-β1 needs to activate either one of two transmembrane serine/threonine receptor kinases. The type II receptor (TGFBR2) is required for the antiproliferative activity of TGF-β, whereas the type I receptor (TGFBR1) mediates the induction of several genes involved in cell-matrix interactions (Ebner et al., 1993). Expression of TGFBR1 is therefore essential for downstream regulatory effects of TGF-β1 on *BGN* gene expression. We tested both cell lines, THP-1 and E.hy926, using semi-quantitative PCR and verified the presence of TGFBR1 mRNA (not shown). We further analysed the downstream effects of TGF-β1 on the level of transcriptional promoter activation, since variations in BGN mRNA levels could depend on post-transcriptional RNA modification. Transient transfection of *BGN* promoter deletions constructs under the stimulatory regime of TGF-β1 (fig. 21) revealed inducibility of transcriptional activity in THP-1 cells in correspondence with diagnostic PCR results (fig. 20). Significant changes in transcriptional activity over pGL3-Basic compared to basic conditions were observed in fragments representing region -1231 to -935, displaying highest transcriptional activity in THP-1 cells. Overall FI observed was 2.6. In analogy to BGN mRNA levels, no stimulatory effect of TGF-β1 on transcriptional activity of any promoter fragment was observed in EA.hy926 cells (fig. 22).

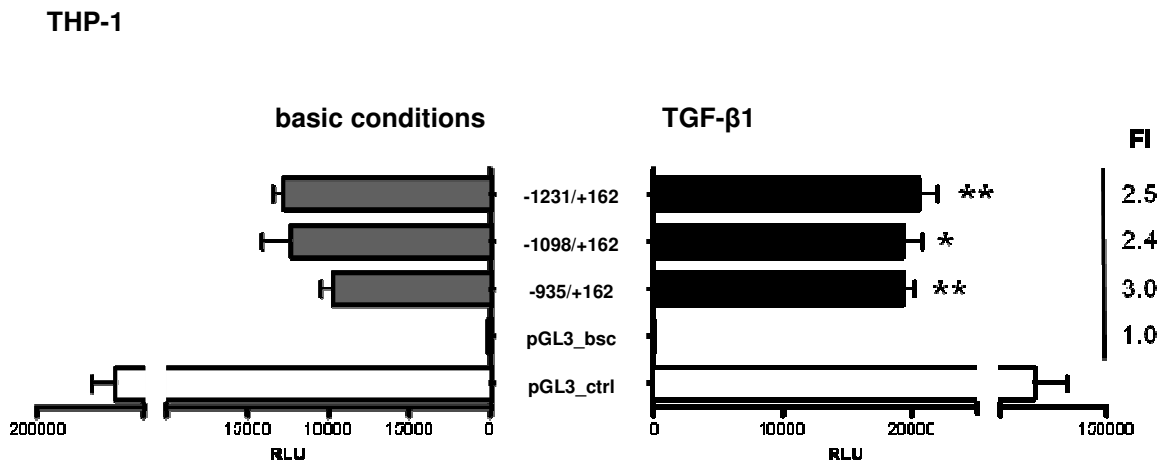


Figure 21: TGF-β1 induces transcriptional activity of the *BGN* promoter in THP-1 cells.

Transcriptional activity of *BGN* promoter fragments was induced up to 3-fold upon stimulation with TGF-β1 (10 ng/ml, 24 hrs). The effect was most prominent for fragments spanning the region -1231 to -935. Basic conditions are shown on the left (gray bars), corresponding transcriptional activity in presence of TGF-β1 is represented by black bars (right). Each constructs' relative activity over the empty shuttle vector pGL3-Basic was calculated and expressed as fold induction (FI) on the far right. Levels of significance are shown by asterisk (*P<0.05, **P<0.01). RLU: Relative Light Units. pGL3-Control is represented by white bars.

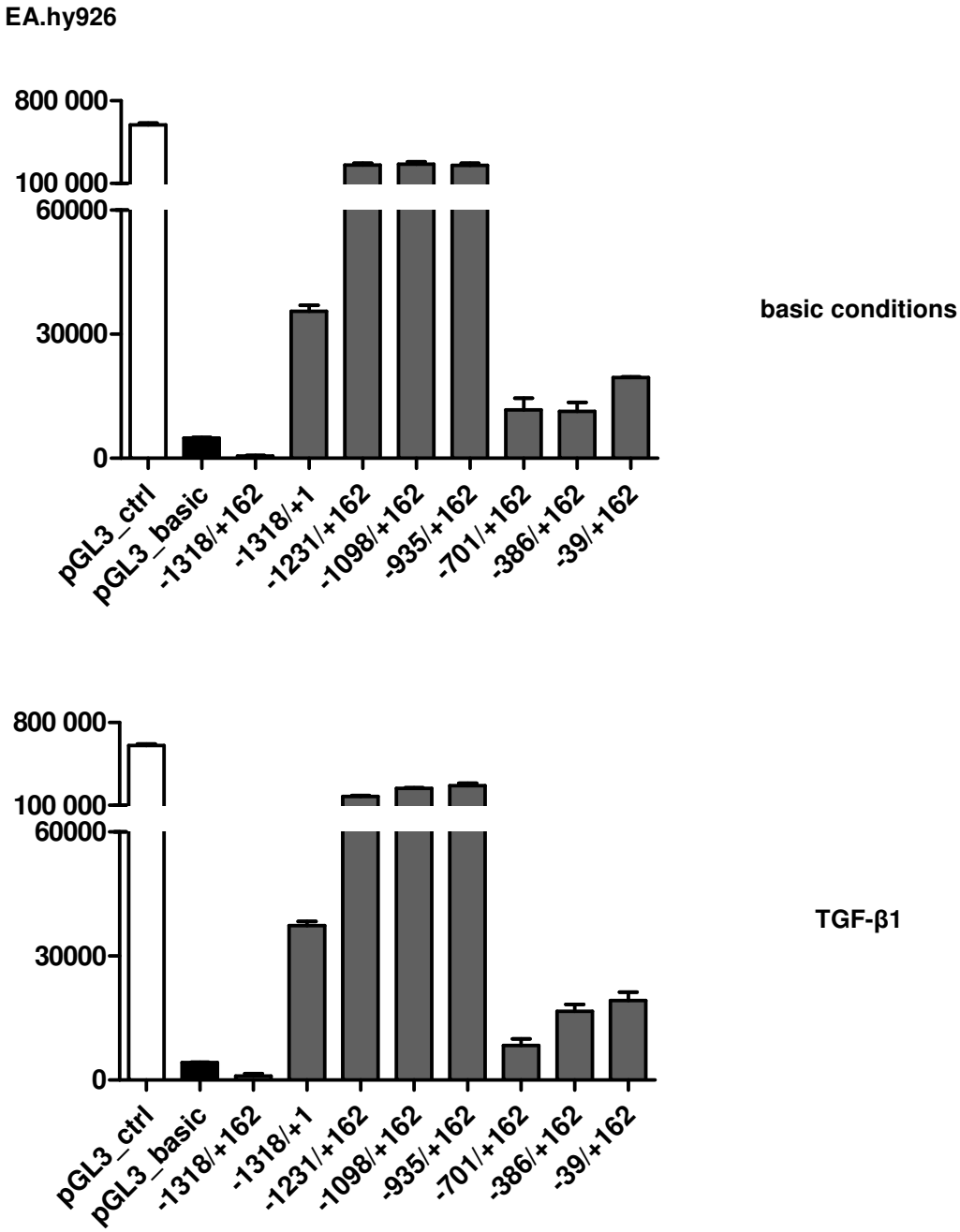


Figure 22: TGF-β1 does not affect *BGN* transcriptional activity in EA.hy926 cells.

Transcriptional activity of *BGN* promoter fragments was not altered upon stimulation with TGF-β1 (10 ng/ml, 24 hrs). Basic conditions are shown on top, corresponding transcriptional activity in presence of TGF-β1 below. RLU: Relative Light Units. pGL3-Control is represented by white bars.

4.7.1 TGF-β1 effect on nuclear SP1 concentration

Since we observed strong effects upon overexpression of TF SP1 in EA.hy926 cells but diverse effects upon stimulation with TGF-β1, a regulatory impact of the cytokine on the nuclear availability of SP1 was assumed. We performed western blot analysis using nuclear extracts derived from both, THP-1 and EA.hy926 cells, comparing unstimulated and TGF-β1 treated protein fractions.

Detected levels of SP1 in nuclear fractions derived from TGF-β1 stimulated EA.hy926 cells showed only marginal differences in protein concentrations compared to untreated cells (fig. 23). In comparison to untreated THP-1 cells, we observed a higher absolute concentration of TF SP1 in the endothelial cell line. THP-1 cells treated with TGF-β1 displayed a reduction in SP1 protein concentration compared to mock control lacking the cytokine.

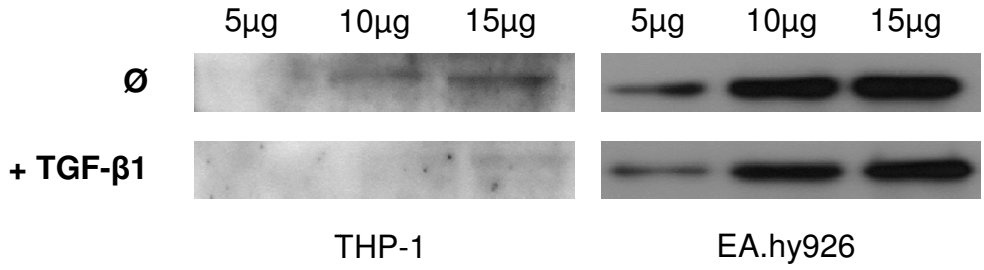


Figure 23: TF SP1 is decreased by TGF-β1 stimulation in THP-1 cells.

SP1 protein concentrations in nuclear extracts were analysed using western blot. THP-1 cells showed a reduced overall concentration of TF SP1 compared to EA.hy926 cells. Stimulation of THP-1 cells with TGF-β1 for 24 hrs lead to a considerable reduction of SP1, while EA.hy926 cells showed only marginal effects. A gradient of total nuclear protein applied to the gel served as loading control. Ø: mock control.

4.8 Band shift experiments

Gel shift experiments were used to access binding of TF with defined *BGN* promoter regions *in vitro*. Biotinylated probes generated resembled the flanking regions of the three genetic variants which constitute *BGN* MolHaps. All probes were between 31 and 36 bp in size, limiting potential binding partners to the exact polymorphic region. Since *BGN* MolHaps span a region of 672 bp, combined analyses in EMSA experiments were inexecutable. To visualize sequence-specific binding, unlabelled probes in a 200-fold excess were used in any gel shift for signal competition. For identification of proteins binding at positions of observed gel shifts, specific antibodies were used according to *in silico* TF binding predictions.

4.8.1 EMSA at position G+94T

Position G+94T has been shown to influence transcriptional activity of the *BGN* promoter significantly in transient transfection experiments (4.4.1). Introduction of the T resulted in strong abrogation of reporter gene plasmid transcriptional activity. Using biotinylated probes resembling the flanking sequence of position G+94T with EA.hy926 and THP-1 nuclear extracts, we observed considerable differences in DNA/protein binding patterns (fig. 24). For both cell lines, a prominent and sequence-specific band shift (black arrow) was detected with significant higher binding affinity to the probe bearing the +94T allele. Of lower intensity, a second specific shift (open arrow) emerged.

Using *in silico* methods (tab. 7), AP-1 was predicted to be a potential binding partner for position G+94T. AP-1 is known to directly affect gene transcription and is individually composed of multimeric TF, often including a heterodimer of the basic leucine zipper proteins c-FOS and c-JUN (Shaulian & Karin, 2002). Regulatory properties of AP-1 are versatile. Of all components, the most potent transcriptional activator is c-JUN, whose transcriptional activity can be antagonized by JUNB. FOS proteins enhance JUN DNA-binding activity upon formation of stable heterodimers. Gel shifts with unlabelled probes harbouring the +94 G or T allele and EA.hy926 nuclear extracts were detected using a c-FOS specific antibody to demonstrate AP-1 complex formation at this position (fig. 25). An AP-1 consensus sequence served as positive control (Kodonaga et al., 1987). We detected a strong c-FOS signal when the AP-1 oligonucleotide was used, demonstrating potential complex formation under applied EMSA conditions. By use of this positive control, binding of c-FOS to both alleles, G and T

at position +94 was demonstrated (open arrow). The band shift (fig. 25, open arrow) identified to involve c-FOS does not exhibit differences in signal intensity in consent with the results obtained using c-FOS antibody.

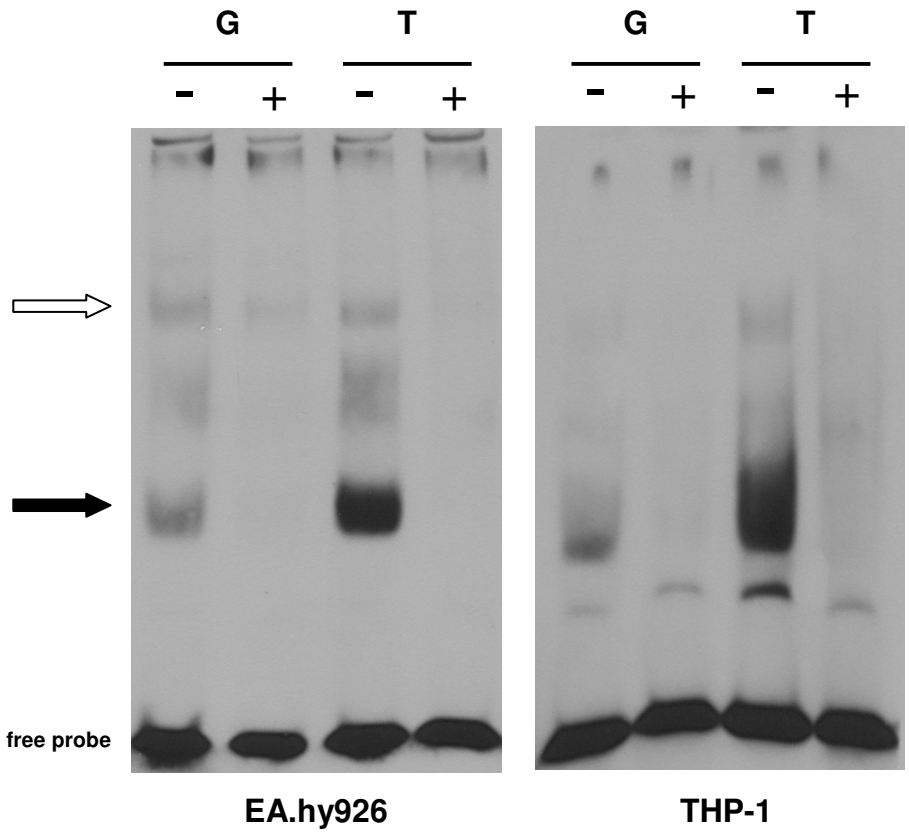


Figure 24: Allele-specific interaction of nuclear extracts with position G+94T.

Nuclear extracts from EA.hy926 and THP-1 cells bind explicitly to the +94T allele, while only minor interaction with the +94G allele was observed (black arrow). A less prominent but specific shift is marked by an open arrow. **-**: 3'-biotinylated probe; **+**: 200-fold excess (8 pmol) of sequence-specific competitor. Unbound oligonucleotides are visible at the bottom (free probe).

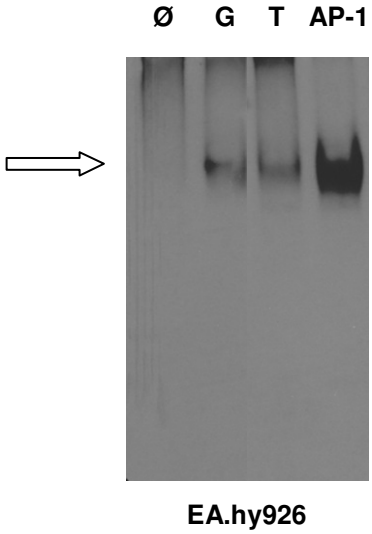


Figure 25: TF c-FOS binds to position G+94T in EA.hy926 cells.

Using a c-FOS-specific antibody, one of the observed shifts (cf. fig. 24, open arrow) was identified to involve binding of TF c-FOS. A consensus site for TF complex AP-1 served as positive control. Ø: lanes without any oligonucleotide served as control for unbound protein separation in native gels. All oligonucleotides were used at a concentration of 8 pmol.

4.8.2 EMSA at position G-151A

Gel shift experiments with biotinylated probes resembling position G-151A and EA.hy926 nuclear extracts revealed two specific bands for both alleles (fig. 26, A). A different separation of band shifts was observed using nuclear extracts from THP-1 cells, while bands were also specific but no differences between alleles was observed (fig. 26). ChIP analysis of this position (cf. fig. 19) revealed binding of TF SP1 in EA.hy926 cells. Genotype of EA.hy926 cells was determined by sequencing to represent *BGN* MolHap1, harbouring the G allele at position -151. EMSA experiments were used to verify these results and to further access binding of SP1 to the A allele. Taking *in silico* predictions (tab. 7) for position G-151A and consensus site weight matrix information into account, a mutated probe was designed by altering either two nucleotides at both ends of the probe basic for TF SP1

binding. Application of the altered oligonucleotide as specific competitor did not resulted in competition and signal intensity of both probes representing either G and A allele did not changed (fig. 26, B). Use of the mutated sequence as biotinylated probe itself featured a completely different binding pattern. The most considerable and strongest signal (black arrow) was undetectable in lanes using the mutated probe. To directly identify the presence of SP1 in this DNA/protein complex, a SP1-specific antibody was used for signal detection. This approach revealed binding of TF SP1 to both alleles at position G-151A with identical signal intensity. As designed, the oligonucleotide missing perfect SP1 consensus sites did not bind to the TF.

We further investigated whether SP1 binding was of stoichiometric nature, or binding was linear to probe concentration. Application of the unlabeled -151G probe in different concentrations (fig. 27) resulted in consequent increase of signal intensity linear to amount of available oligonucleotide. Level of saturation was reached at a concentration of 4.0 pmol.

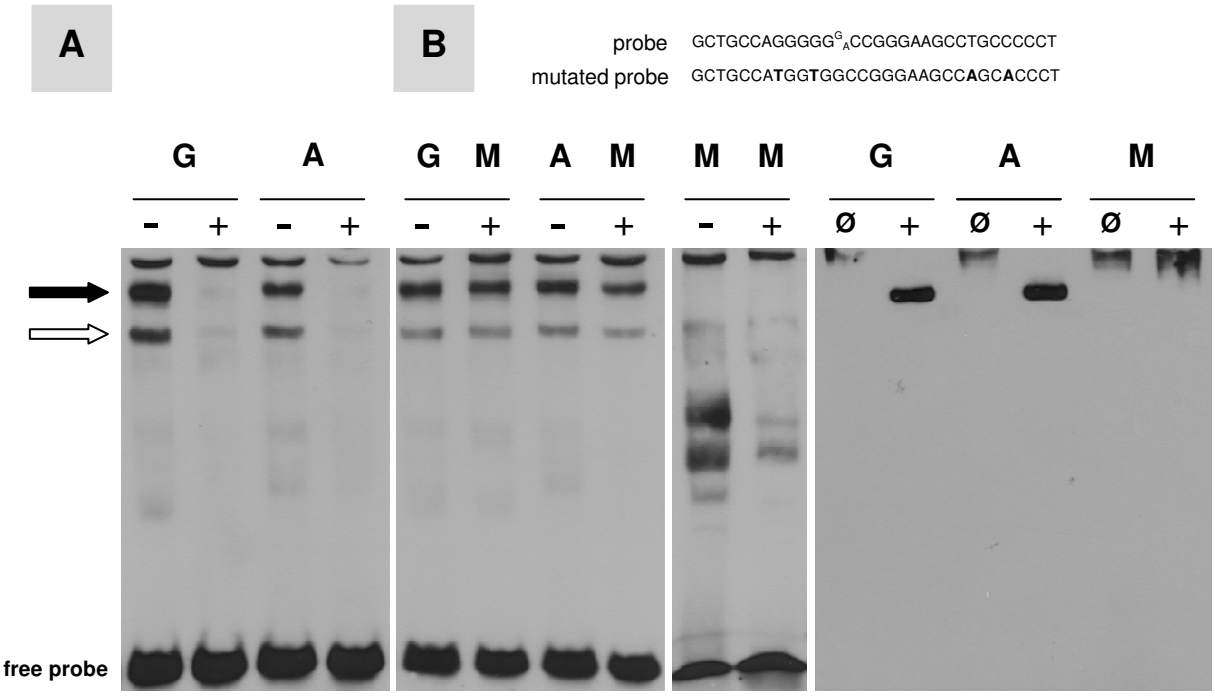


Figure 26: Sequence-specific binding of SP1 at position G-151T.

A: Nuclear extracts from EA.hy926 cells show explicit binding to the -151 G and T allele (arrows). **B**: The mutated sequence M (depicted above, mutated positions in bold letters) is not able to compete binding of the unaltered sequence. Using the mutated sequence as biotinylated probe results in a different binding pattern. The band indicated by a black arrow was identified to bind SP1 by using a SP1 antibody (far right). The mutated probe does not show any residual binding of SP1. **-**: 3'-biotinylated probe; **+**: 200-fold excess (8 pmol) of sequence-specific competitor. Ø: lanes without any oligonucleotide served as control for unbound protein separation in native gels. Unbound oligonucleotides are visible at the bottom (free probe).

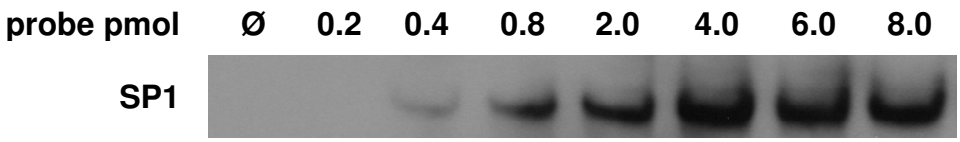


Figure 27: Binding of SP1 in -151G probe serial dilution.

TF SP1 binds the unlabelled -151G probe in relation to oligonucleotide quantity. Saturation of binding was reached at a probe concentration of 4.0 pmol using EA.hy926 nuclear extracts. Ø: lane without any oligonucleotide served as control for unbound protein separation in native gels.

4.8.2.1 THP-1 band shift experiments

Gel shift experiments using THP-1 nuclear extracts featured a different binding pattern compared to nuclear extracts derived from EA.hy926, missing the most prominent signal (fig. 26, black arrow) representing SP1 interaction. Two specific interactions were observed (fig. 28, arrows) but gel shifts did not differ in dependency of the present allele. Using a SP1 consensus site as sequence-specific competitor did not result in signal intensity decrease and application of the mutated probe (cf. fig. 26) showed identical binding compared with unaltered probe sequences.

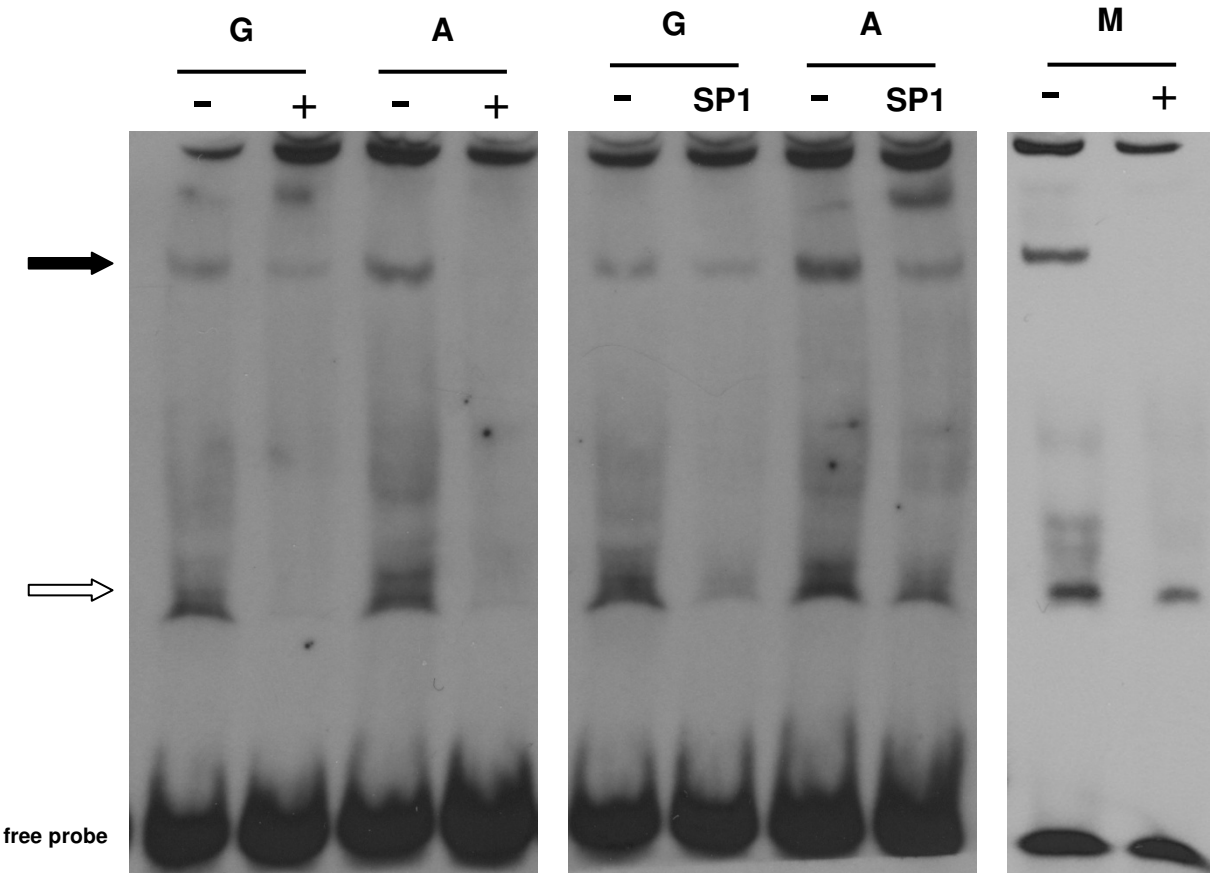


Figure 28: Sequence-specific interaction of THP-1 nuclear extract with G-151A is independent of TF SP1.

Left: THP-1 nuclear extracts specifically bound to position G-151A (arrows). Centre: Incomplete competition of bands was observed using SP1 consensus sequences. Right: The mutated probe M (fig. 26) shows identical specific binding pattern. -: 3'-biotinylated probe; +: 200-fold excess (8 pmol) of sequence-specific competitor. SP1: A SP1 consensus site served as specific competitor. Unbound oligonucleotides are visible at the bottom (free probe).

4.8.3 EMSA at position G-578A

Band shift analysis of position G-578A using EA.hy926 nuclear extracts revealed a prominent allele-specific effect for probes harbouring the G allele (fig. 29 left, black arrow), while no specific effect was observed for position -578A. Application of nuclear extracts obtained from TGF-β1 stimulated cells changed the observed DNA/protein bands, resulting in amplification of a second shift (open arrow) with higher signal intensity for the A allele and abrogation of the interaction with the G allele (black arrow). Signal detection using SP1-specific antibody did not reveal any interaction of unstimulated nuclear extracts with SP1 (fig. 29, right), while a positive result was obtained for protein extracts from TGF-β1 treated cells. Binding of TF SP1 is observed in parallel to gel shift intensification (open arrow) upon TGF-β1 stimulation.

Nuclear extracts from THP-1 cells presented a different binding pattern, with one considerable specific band (fig. 30, black arrow). Upon application of TGF-β1 stimulated nuclear extracts, an additional shift occurred for THP-1 cells (fig. 30, angled arrows). In consistence with results observed using stimulated EA.hy926 extracts, the emerging DNA/protein interaction was identified to depend on SP1 binding.

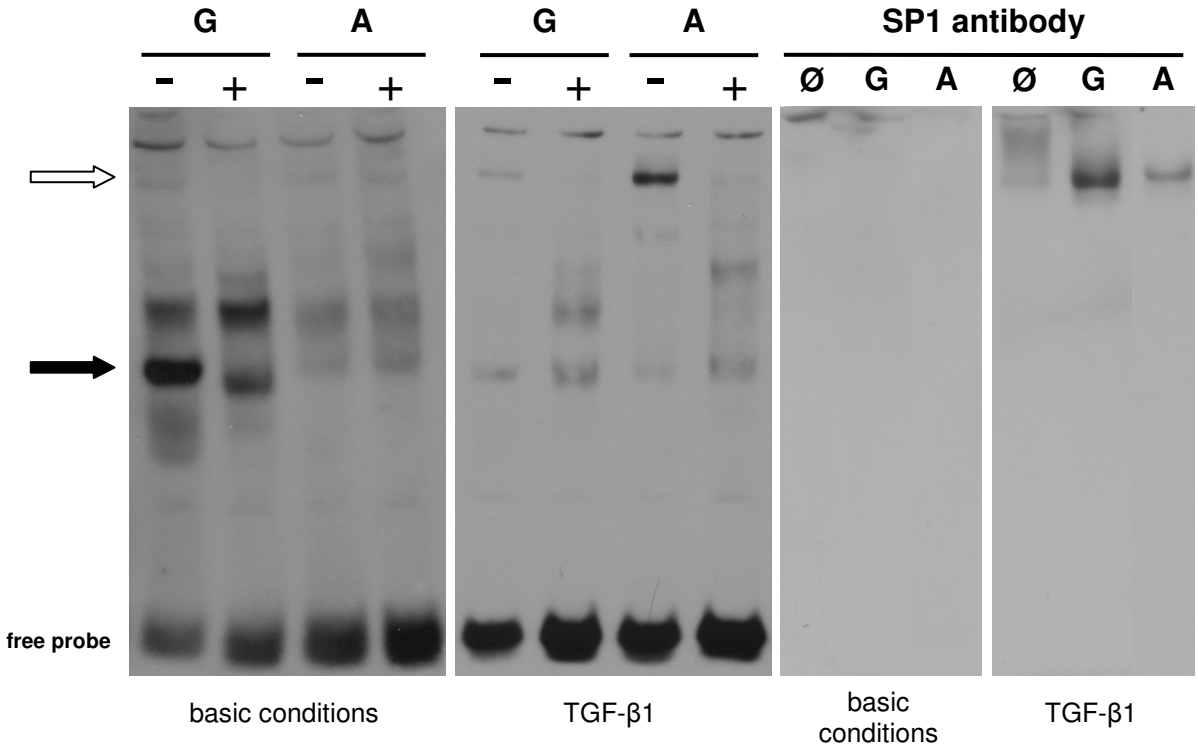


Figure 29: TGF-β1 stimulation alters interaction of EA.hy926 nuclear extracts with position G-578A.

An allele-specific gel shift (black arrow) was detected for the -578G allele using nuclear extracts from cells held under basic conditions (left). TGF-β1 stimulated extracts show differential binding pattern with an emerging band at the top of the gel (open arrow). The emerging band was identified to depend on SP1 binding to the probes (far right). -: 3'-biotinylated probe; +: 200-fold excess (8 pmol) of sequence-specific competitor. Unbound oligonucleotides are visible at the bottom (free probe). Ø: lane without any oligonucleotide served as control for unbound protein separation in native gels.

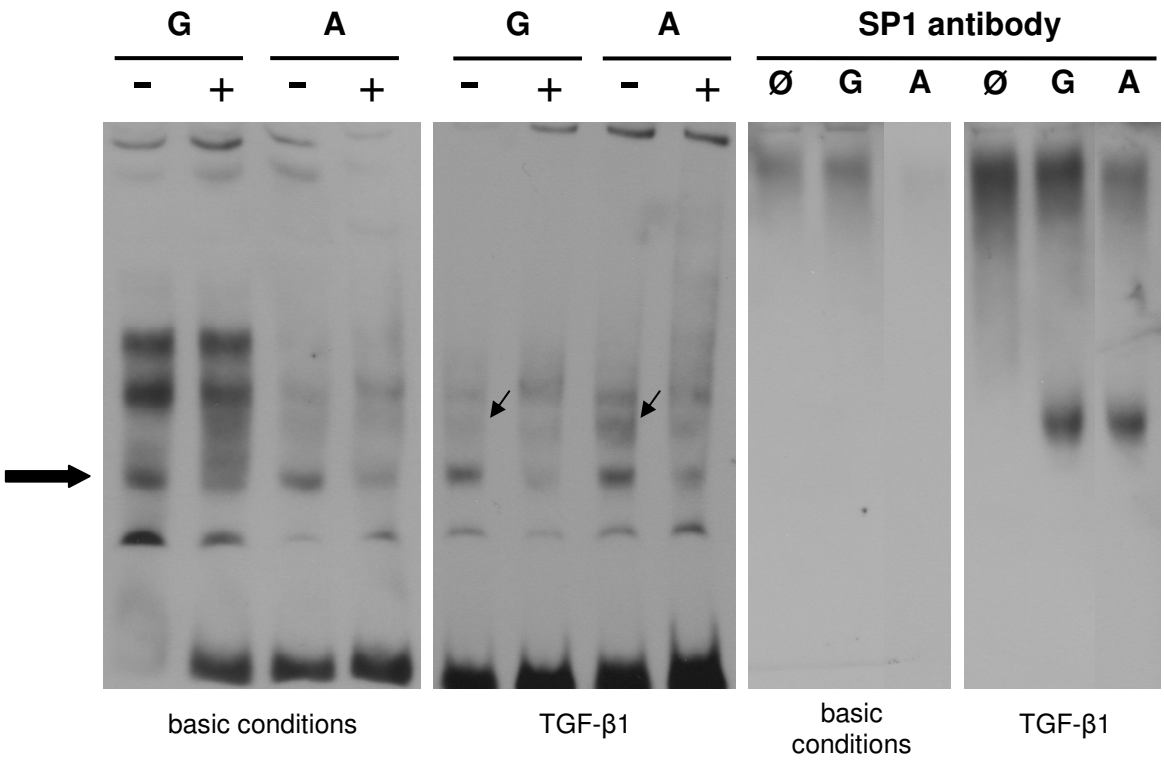


Figure 30: TGF-β1 stimulation enhances SP1 binding to position G-578A in THP-1 cells.

A specific gel shift (black arrow) was detected for the G-578A position allele using nuclear extracts from cells held under basic conditions (left). TGF-β1 stimulated extracts show differential binding pattern with an emerging band (angled arrows). The emerging band was identified to depend on SP1 binding to the probes (far right). **-**: 3'-biotinylated probe; **+**: 200-fold excess (8 pmol) of sequence-specific competitor. Unbound oligonucleotides are visible at the bottom (free probe). Ø: lane without any oligonucleotide served as control for unbound protein separation in native gels.

We further investigated *in silico* binding predictions for both alleles with special respect towards cell-specific regulatory factors. Computational analysis revealed an altered PU.1 consensus site upon introduction of the A allele at position -578. The hematopoietic ETS-domain transcription factor PU.1 has been identified to regulate human monocyte/macrophage differentiation (Rosa et al., 2007) and could account for cell-specific *BGN* gene expression. The sequence-specific gel shifts observed using nuclear extracts from untreated and TGF-β1 stimulated THP-1 cells (fig. 30, black arrow) were detected using specific PU.1 antibody (fig. 31). Binding of PU.1 was identified to probes harbouring either one allele with increased binding of PU.1 when stimulated protein extracts were applied. Potential differences in binding affinity were assessed using probe serial dilution (fig. 32). A 4-fold higher binding affinity of PU.1 towards the probe harbouring the -578G allele compared with the probe presenting the A allele was observed in these experiments, in agreement with previous *in silico* analyses.

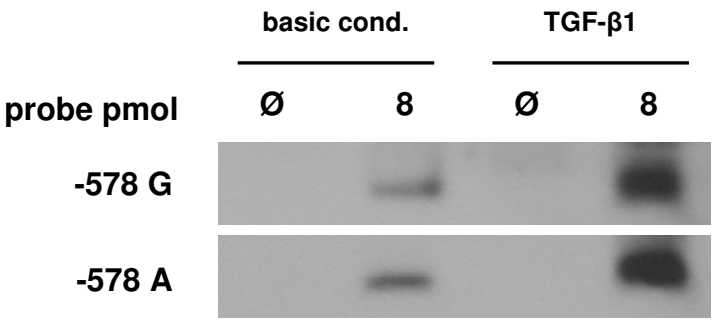


Figure 31: TF PU.1 binds position G-578A in THP-1 cells.

TF PU.1 binds unlabelled -578 probes independent of the present allele. Binding is increased in nuclear extracts from TGF-β1 stimulated cells. Ø: lane without any oligonucleotide served as control for unbound protein separation in native gels.

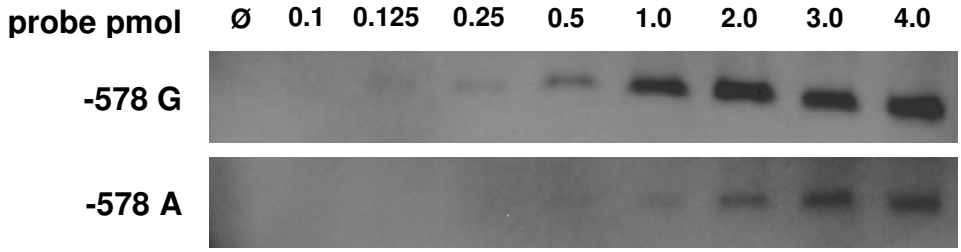


Figure 32: TF PU.1 binds the G allele with higher affinity.

TF PU.1 binds unlabelled -578 probes in relation to oligonucleotide quantity. Signal intensity of the probe bearing the G allele detected by specific PU.1 antibody was increased 4-fold compared to probe -578 harbouring the A allele. Ø: lane without any oligonucleotide served as control for unbound protein separation in native gels.

5 DISCUSSION

In the present work we analysed the transcriptional regulation of the human *BGN* gene expression. We have put our emphasis on the impact of genetic variants G-578A (rs11796997), G-151A (Brand et al., 2002) and G+94T (rs5945197) within potential *cis*-regulatory regions, i.e. potential TF binding sites. To our knowledge only a view publications (Ungefroren & Krull, 1996; Heegaard et al. 2004) aimed at characterizing the molecular basis of *BGN* transcriptional regulation and little is known about its polymorphic promoter structure and consequences for its regulation. We identified a promoter portion of 1025 bp exerting sufficient transcriptional activity in EA.hy926 and THP-1 cells. In the context of this 1025 bp construct, we observed a prominent effect of SNPs (G-578A, G-151A and G+94T) reducing *BGN* promoter transcriptional activity. By generation of truncated promoter constructs, co-expression experiments and detailed band shift analyses, we identified c-FOS and SP1 to up-regulate *BGN* gene expression. Moreover, haematopoietic TF PU.1 significantly up-regulated monocyte-specific *BGN* promoter activity.

5.1 Variable TSS of the *BGN* gene

Transcription of the *BGN* gene is documented to initiate from multiple TSS (Ungefroren & Krull, 1996; DBTSS) and results in varying lengths of transcripts' 5'-UTR. In this study, we used the 5'RACE approach to analyse the transcriptional organization of the *BGN* promoter. This led to the identification of a TSS (TSS1) three bp upstream of a major TSS (AK092954) listed in DBTSS, which was detected by oligo-capped cDNA analysis. In the endothelial cell line EA.hy926, we further identified a TSS located 46 bp upstream of TSS1. This second TSS is used in parallel to TSS1, resulting in two diverse *BGN* transcripts, simultaneously expressed under basic conditions as well as under protein kinase A pathway stimulation with cAMP. Using the phorbol ester PMA, initiation of transcription was preferred from TSS1. In contrast to other reports (Ungefroren & Krull, 1996), no further upstream TSS existed in our tested cell lines. Detection of varying TSS in close vicinity, as observed in our experiments for TSS1, has been discussed with regard to technical inaccuracy of applied methods (Hagedorn et al., 2009). This has been particularly proven for primer extension of cDNA

products using radiolabelled primers, where excessive radiation may result in degradation of 5'-ends and mis-determination of TSS. 5'RACE methodology, which includes primer extension enhanced by PCR amplification and allows determination of the 5'-end of lower abundant transcripts produced from weaker promoters (Scotto-Lavino et al., 2006), has been proven to identify TSS more accurately.

The distribution of the TSS could reflect the dynamic nature of the transcriptional initiation events, labelling them as regions with TSS in close proximity rather than static positions. This flexible structure of transcript initiation has explicitly been reported for CpG island promoters (Smale & Kadonaga, 2003). Although multiple TSS are present, broad-type promoters still display a preference for specific initiation sites (Carninci et al., 2006). Our results as well as findings from other groups (Wegrowski et al., 1995; Ungefroren & Krull, 1996) suggest a broad TSS distribution over the *BGN* promoter. Using multiple TSS over an extended region thereby requires exclusion of ATG start codons from sequences downstream of initiation sites since translation generally starts from the first ATG in the nascent mRNA (Lee et al., 2005a). Interestingly, general ATG distribution within the human *BGN* 5'-flanking region is adjourned 362 bp upstream of TSS1, generating a so called ATG desert. The intention of multiple promoter TSS usage, specifically for the major group of CpG island promoters in mammals, is still under investigation. Of importance, these findings imply that we cannot consider the 5'-end of the longest cDNA of a gene as the true full-length transcript. There is evidence that TNFIID complex formation actually occurs relatively non-specifically and scans along the DNA for a TSS (Sandelin et al., 2007). This will allow for parallel assembly of PIC in presence of multiple TSS and could regulate transcript initiation and gene expression. In our work, the observed concentration of transcript initiation on TSS1 after stimulation with PMA did nevertheless not translate in reduction of total *BGN* transcript level. Altered mRNA 5'-regions could also result in modified mRNA half-life but mRNA is stabilized by intron size rather than destabilized by the lack of exonic regions (Ross, 1995). Posttranscriptional effects of altered mRNA 5'-cap thus may play a role for *BGN* translational processes. After release from the nuclear pore, mRNAs are often directed to particular sites in the cytoplasm, driving subcellular protein localization patterns (Lécuyer et al., 2007).

5.2 *BGN* promoter capacity is altered by genetic variants

The impact of genetic variants residing within or near *cis*-active elements on promoter activity has been reported repeatedly by our group and others (Dördelmann et al., 2008; Hagedorn et al. 2009; Telgmann et al., 2008; Funke-Kaiser et al., 2003; Frisdahl et al., 2005; Schultz et al., 2009). Transient transfection assays in cell lines expressing the gene of interest are state of the art for *in vitro* determination of effects on promoter transcriptional activity, but certain conditions are necessary for their quantification. The analysed deletion construct, resembling the promoter portion carrying a certain allele, needs to display sufficient transcriptional activity. Therefore, assembly of the basic transcription machinery and formation of PIC on the analysed fragment is mandatory since specific transcriptional activity is defined by differential interplay of *cis*-active elements with the core promoter (Lemon & Tjian, 2000). To translate into measurable reporter gene expression, a downstream TSS is necessary. Reliable results also depend on transfection efficiency and technical accuracy indicated by low standard deviations across experiments. Genetic variants in a given promoter fragment may contribute differentially to overall transcriptional activity of the analysed fragment since sequence alterations can influence core promoter motives such as TATA boxes (as shown for RhoA by Schröer, Ph.D. thesis) and recognition of consensus sequences by TF at *cis*-elements. Combinations of alleles on the same DNA strand, represented by MolHaps are thereby often observed to result in overall differential effects on transcriptional activity, acting synergistically or competitively (Dördelmann et al., 2008; Hagedorn et al., 2009). A preliminary analysis of ~200 bp isolated promoter portions harbouring *BGN* variants in the context of the pGL3-Promoter vector (Rüssmann, MD thesis) suggested limited enhancer capacity of the single fragments upon transfection into HEK293T cells. The results suggested the necessity to further characterize the 'entire' *BGN* promoter and to analyse the genetic variants in their original MolHap composition. Despite the endogenous *BGN* expression of HEK293T cells, the composition of HEK293T nuclear proteins (i.e. DNA binding proteins) seemed to interact indifferently with *BGN* MolHaps, lacking the potential do drive *BGN* transcriptional activity with respect to the individual *BGN* promoter composition. This was also supported by the observations obtained from EMSA experiments with HEK293T nuclear extracts for positions G-578A and G-151A. Band shift experiments of

multiple compositions repeatedly failed to exhibit specific DNA/protein interactions. EA.hy926 and THP-1 cells were therefore chosen for the current *BGN* promoter analysis, in addition to their close connection with atherosclerosis pathophysiology.

In the present characterization of the *BGN* promoter, we report sufficient transcriptional activity of deletion constructs presenting the region between bp -1231 to -935. The promoter was sufficiently transcriptionally active under basal conditions in the vascular endothelial cell line EA.hy926. The vascular endothelium lines all blood vessels in a luminal monolayer and forms, together with the basal membrane, the tunica intima. The basal membrane thereby consists of a specialized ECM (Wiradjaja et al., 2010), its composition being influenced by the endothelial cell phenotype, connecting the endothelium to vascular smooth muscle cells. During development of collagen matrices initial fibrils are cross-linked by SLRP such as BGN, binding to different sites of a collagen monomer (for comprehensive review see Kalamajski & Oldberg, 2010). Hence, sufficient transcriptional activity of the *BGN* gene under basic conditions seems conceivable since assembly of the ECM is a permanent process. Potential effects of the identified MolHaps were analyzed within the context of a 1025 bp portion of the *BGN* promoter (-893/+132) after in-depth analysis of 1300 bp 5'-flanking region and 5'-UTR of the *BGN* gene. We are aware that the analysed portion may not be considered as the entire promoter since recent publications have shown the profound effect of *cis*-acting elements located at considerable (up to 1000 kb) distances (Lower et al., 2009, Higgs et al., 2008). On the other hand, distal promoter elements acting independently of their location have been described to consist of enhancers and insulators influencing the activation of high-level transcription rather than basic promoter activity (West & Fraser, 2005). *BGN* MolHap2 and MolHap3 showed a significant decrease of promoter performance in both, EA.hy926 and THP-1 cells under basic conditions compared to wt. To verify our hypothesis of concerted interaction of TF binding to polymorphic regions in haplotype constellations with their core promoter, we tested each variable region with potential enhancer activity in the context of the pGL3-Promoter vector. The vector contains a SV40 promoter and introduced DNA fragments potentiate luciferase gene activity in case they exert enhancer capacity. *BGN* promoter fragments proximal to TSS1 at positions G-151A and G+94T did not enhance promoter activity when the major allele was present. For the minor +94T allele carrying construct, we observed total inhibition of promoter activity, while introduction of the minor -151A allele in the context of 188 bp flanking region activated transcriptional activity to some

extent. The more distal promoter portion harbouring G-578A exerted enhancer capacity in interaction with the SV40 promoter. These results implicate an overall compensatory effect of the analysed polymorphic positions in their allelic MolHap constellations, demonstrated explicitly by positions G-578A (MolHap1 and 3) and G-151A (MolHap1 and 2). Presence of the minor alleles in the MolHap situation did not enhance transcriptional activity but resulted in strong abrogation of activity. However, single variant analysis in context of the pGL3-Promoter revealed differences between the introduced alleles and their effect on transcriptional activity, suggesting alterations in consensus motives and TF binding.

5.3 Identification of DNA/protein interactions

At the molecular level, gene expression is orchestrated by the recruitment of *trans*-acting proteins (TF) to *cis*-acting sequences (binding sites), creating a series of *cis*-regulatory modules (CRM) along a given sequence. CRM act as modular units, integrating the input from multiple TF to a specific spatio-temporal output of gene expression. Binding of TF thereby depends on electrostatic interaction with the DNA molecule at TFBS, as well as the stoichiometric availability of the TF participating in module assembly. Overlapping or superimposed binding sites for multiple factors may thereby result in binding competition of diverse regulatory proteins (Fry & Farnham, 1999), some of them contributing differently to overall transcriptional activity. TFBS exhibit specific characteristics with each nucleotide accounting differentially to the molecular function of the entire site. The relevance of each nucleotide is commonly illustrated by a position weight matrix (Piipari et al., 2010), also used for computational prediction of TFBS.

Modification of TFBS by genetic variants may result in a complete loss of TF binding but alteration of binding affinity is more likely to be observed since consensus sequences are multivalent. Computational TF binding prediction, although integrating position weight matrixes, are limited to DNA sequence recognition but cannot account for interactions of heteromeric CRM. Comprehensive prediction of protein binding in MolHap constellations therefore has not been achieved yet and is in addition suffering from overall missing information on the physiological state of a given cell. To account for the stated limitations of *in silico* TFBS prediction, we started our analyses on possible transcriptional regulators of the

BGN promoter with co-expression experiments rather than with single position analyses. Thereby, TF SP1 was one of the most likely candidates, proposed by different computational predictions and the overall GC-rich promoter sequence (Briggs et al., 1986). Additional indication were derived from the report of Verrecchia et al. (Verrecchia et al., 2001), demonstrating downregulation of *BGN* and other ECM components as an effect of antisense *SP1* expression. Here we present substantial evidence for SP1 activation of the *BGN* promoter upon overexpression, enhancing transcriptional activity up to 4-fold. Co-transfection with other predicted binding factors such as CEBP-α and δ or CBP did not result in any alteration of promoter activity. ChIP analyses of different promoter regions demonstrated direct and parallel physical association of SP1 with proximal and distal regions under basic conditions. Both results indicated participation of SP1 in the regulation of basal *BGN* gene expression as well as the enhancing nature of SP1, rather than prompting towards indirect downstream SP1 effects. It has been shown that transcriptional activation by enhancer-binding factors such as SP1 requires TFIID (Hoey et al., 1993). SP1 glutamine-rich activation domains thereby interact with the TAF_{II} subunits of TFIID and the essential cofactor complex CRSP (Cofactor Required for SP1 transcriptional activation; Ryu et al., 1999). Notably, Geerkens and colleagues (Geerkens et al., 1995) found that *BGN* expression levels are increased in patients with additional sex chromosomes and suggested the observed 'pseudo-pseudoautosomal expression' of *BGN* to result from a gene that escapes X-inactivation and regulates *BGN* transcriptional activity. The gene encoding one domain of the CRSP cofactor complex, $CRSP_{150}$, has been mapped to the X-chromosome. $CRSP_{150}$ displays pseudoautosomal characteristics, escapes inactivation (Wutz & Gribnau, 2007; Carrel et al., 1999) and could contribute to this phenomenon.

5.4 *BGN* MolHaps

In contrast to MolHap3 [A^{-578}-G^{-151}-G^{+94}], *BGN* MolHap2 [G^{-578}-A^{-151}-T^{+94}] comprises two minor alleles, one at the proximal position G-151A and one at the 5'-UTR position G+94T. The observed reduction of promoter transcriptional activity upon introduction of MolHap2 could thus originate from altered TF binding at either one of the two positions but is more likely to be explained by a deranged interplay of factors interacting with both polymorphic promoter segments.

We identified binding of the TF complex AP-1 at position G+94T in EMSA experiments using a specific c-FOS antibody. AP-1 is known to regulate a wide range of cellular processes, including cell proliferation, cell death and differentiation. AP-1 is not a single factor, but a combination of dimeric basic leucine zipper proteins that belong to the JUN, FOS, MAF and ATF sub-families, which recognize with highest affinity 12-*O*-tetradecanoylphorbol-13-acetate (TPA) response elements (5'-TGAG/CTCA-3') and slightly lower affinity cAMP response elements (5'-TGACGTCA-3'; Chinenov & Kerppola, 2001). AP-1 binding may thereby be positioned in the 5'-UTR of a promoter, as reported for the sodium-dependent bile acid transporter by Duane and colleagues (Duane et al., 2007).

The first report on AP-1 described its activation potential on the $MTII_A$ promoter (Lee et al. 1987a), establishing its name and implicit function as activator of transcriptional activity. In recent years, further analyses offered increasing evidence that some effects of AP-1 are mediated by gene repression (Murphy et al., 1996). These effects may depend on the interactions of AP-1 family members with transcriptional corepressors (Pessah et al., 2000) and the nature of the AP-1 consensus site (Hsu et al., 1993). Other groups (Kouzarides & Ziff, 1988; Gentz et al., 1989; Turner & Tjian, 1989) demonstrated that the leucine zipper domain of both c-FOS and c-JUN are necessary for heterodimer formation and transcriptional transactivation. The mere presence of a leucine zipper is, however, not sufficient for dimer formation. c-FOS does not form homodimers, whereas such a formation has been reported for c-JUN complexes with weak AP-1 binding capacity (Dwarki et al., 1990). It therefore seems conceivable that detected binding of c-FOS at either allele at position G+94T marks a differently composed AP-1 complex with activating capacity in case of the G allele and repressor properties in presence of the T allele. Shah et al. (Shah et al., 2006) demonstrated transcriptional suppression by AP-1 binding to a polymorphic promoter position (-1347C/T) of *TGFB1*. Cellular studies showed that an AP-1 complex containing JUND and c-FOS was recruited to the *TGFB1* promoter *in vivo* only when the -1347C allele was present. The -1347T allele was associated with increased TGF-β1 levels because of impaired negative regulation by AP-1. Schreiber and colleagues (Schreiber et al., 1999) reported that the *p53* promoter contains a motif that differs from the consensus AP-1-binding site in only one single position, creating a negative regulatory element which is bound by c-JUN, resulting in a transcriptional repression of the *p53* promoter.

The transactivating potential of AP-1 has been reported frequently to depend on co-activators such as SP1 (Wu et al., 2003; Kardassis et al., 1999). In addition, SP1/SP3 and AP-1 are often found as components in large multimeric transcription factor complexes. AP-1 has also been reported to fail in transcriptional activity activation upon SP1 inhibition on promoters where interaction with SP1 has been observed, suggesting the interaction of AP-1 and SP1 to be required occasionally to transmit AP-1 effects (Lee et al., 1987b). The interplay of position G+94T and G-151A may therefore be of importance for an overall activating potential on transcriptional activity.

The observed effects at position G-151A have two possible explanations. As a common base to both of them, we hypothesize competition of TF with the activating factor SP1 at its consensus motif. As explained above, DNA/protein interaction is thought to be a rather flexible than static process, with TF tendency towards association or dissociation depending to a great extent on the physiological state of the cell. Alteration of cognate sites by genetic variation can thereby result in reallocation of balanced conditions. *In silico* analyses for position G-151A predicted a superimposed EGR family consensus motif to the identified SP1 binding sites. Competition of EGR family members with TF SP1 has been reported frequently, in particular for EGR1 (Hagedorn et al., 2009; Hsu et al., 2009; Khachigian et al., 1995). Both TF share very similar consensus motifs, recognized in either case by three zinc-finger DNA-binding domains of the protein. In contrast to SP1, which is an activating TF, EGR family members are able to generate repressor characteristics. Except for EGR4, all EGR family members exhibit a distinct repression domain (R1), through which binding to NAB (NGFI-A binding) proteins can be mediated and the activating potential of EGR is suppressed (Russo et al., 1995). This has been shown conclusively by Tan et al. (Tan et al., 2003) for inhibition of the collagen promoter (COL2A1) by IL1β. In case of *BGN*, EMSA experiments at position G-151A using EGR1 and EGR2 antibodies were inconclusive and did not support specific interaction of EGR family members with this distinct promoter segment.

Yet another explanation for altered transcriptional activity upon TF competition at *cis*-active elements can be found in the SP TF family itself. The ubiquitous TF SP1 and SP3 exhibit very similar DNA-binding specificities and compete for binding to the same GC-boxes (Kingsley & Winoto, 1992). Therefore, the SP1/SP3 ratio plays an important role in gene regulation (Yu et al., 2003). On some promoters, SP3 cooperates with SP1 (Gartel et al., 2000) whereas on others, SP3 antagonizes the SP1-mediated activation (Kumar & Butler, 1997). This effect is explained as follows. In general, SP3 represses the SP1-mediated transactivation of promoters

with two or more SP1 sites but does not affect the SP1-mediated transactivation of promoters with only one SP1 site (Ritchie et al., 2000). Moreover, SP3 lacks the ability of SP1 to transactivate synergistically via two or more SP1 sites (Majello et al, 1997). Consequently, if SP3 displaces SP1 from promoters with at least two SP1 sites, a net repression of the initial SP1-mediated transactivation is observed. We have shown binding of SP1 to the *BGN* promoter on at least two distinct positions under basic conditions. Binding of SP3 to the promoter, triggered by the G-151A site, would most likely result in the loss of SP1-mediated transactivation. Introduction of the minor A allele resembling MolHap2 creates a CT-box motif of the 5'-GGGGAGGGGC-3' type. Binding of SP1 to CT-boxes has been reported to be sixfold weaker compared to GC-boxes (Letovsky & Dynan, 1989), whereas SP3 binds CT-boxes with identical affinity (Hagen et al., 1992).

In transient transfection assays, we have shown the impact of MolHap3 on *BGN* transcriptional activity, created by introduction of the G-578A site. This observation strongly supports the concept of potent alteration of promoter activity by single nucleotide variations. In contrast to MolHap2, where interaction of SP1 has been detected at position G-151A, no binding of SP1 to position G-578A under basic conditions, neither in EMSA experiments nor in ChIP analysis, was observed despite *in silico* TFBS prediction. Taken the identification of SP1 at the nearby position bp -918 to -806 using ChIP technology into account, this result is not likely to be explained by methodical inaccuracy but rather provides evidence for ChIP assay stringency.

Using nuclear extracts from THP-1 monocytes we identified the ETS TF PU.1 to bind position G-578A with significantly higher affinity towards the G allele. Very interestingly, we were recently able to demonstrate an association of G-578A with cardiac resynchronisation therapy phenotype, in that the G allele was associated with response to cardiac resynchronisation therapy in patients with heart failure after MI (Schmitz, for FP7 VPH2). Computational analysis of the proximal region of the exact position identified a repeat of PU boxes of the 5'-GGA(A/T)-3' type, which is partly corrupted by introduction of the A allele. Multiple studies have demonstrated that PU.1 is essential for the development of hematopoietic cells of myeloid, B-cell, and T-cell lineages (Scott et al, 1994) and is further actively involved in regulation of genes in monocyte-derived macrophages (Nacu et al., 2008). Despite its predominant expression and pivotal function in hematopoietic lineages, PU.1 is in addition expressed in fibroblasts, which are important structural elements of tissue integrity synthesizing ECM components, and has the potential to convert fibroblasts into

macrophage-like cells (Feng et al., 2007). PU.1 can directly bind to TFIID, mediated by the activation domain of PU.1, without the need for an adaptor protein as shown necessary for SP1 (Hagemeier et al., 1993). Since PU.1 is not expressed in endothelial cells such as the EA.hy926 linage, other members of the ETS family could potentially bind to the presented PU boxes.

In fact, ETS proteins are described to play a major role in endothelial cells and all characterized endothelial enhancers contain multiple ETS binding sites (Bernat et al., 2006; De Val & Black, 2008). At least 19 different ETS proteins are expressed in human endothelial cells (Hollenhorst et al., 2004; Liu & Patient, 2008). ETS factors are critical for ECM remodeling (Trojanowska, 2000) and deletion of ETS genes leads to impairment of this process (Pham et al., 2007; Oettgen, 2006). Great redundancy is observed among the majority of ETS factors, which probably reflects their binding to identical *cis*-acting elements, the consensus motif highly conserved among all members of the family (De Val & Black, 2009). Notably, the exact consensus motif (the PU box) was initially identified to substitute for the deleted wt enhancer element of the SV40 promoter (Klemsz et al., 1990). In our study, we observed a significant enhancer activity of the G-578A region in the pGL3-Promoter context, too. Controversially and despite the observation that both alleles were able to enhance SV40 promoter activity, total abrogation of promoter transcriptional activity was observed upon introduction of the A allele in the MolHap context cloned in pGL3-Basic vector. Indeed, it has been estimated that ETS proteins contribute to endothelial-specific gene expression, despite their overall ubiquitous expression in adult tissues and ETS binding sites being not specific to endothelial-expressed gene loci, by functioning in combination with other transcription factors. An interesting effect has been described concerning the interplay of ETS proteins and SP1. Trojanowska (Trojanowska, 2000) reported that two distinct ETS protein, FLI1 and ETS1, contribute to the regulation of the *COL2A1* gene in fibroblasts. A functional PU box was identified in the *COL2A1* promoter in close proximity to the SP1 sites. ETS1 and FLI1 had opposite effects on the activity of this promoter. While ETS1 stimulated transcriptional activity, FLI1 inhibited the promoter activity and SP1 binding was essential for this inhibitory effect of FLI1. It is conceivable that different ETS factors, like ETS1 and FLI1, are in active competition at position G-578A in EA.hy926 cells, as we have shown binding of the ETS family member PU.1 at this position in THP-1 cells. Binding affinity for PU.1 is altered in this cell line upon introduction of the A allele and a similar effect, leading to

augmented binding of an excessive TF, could contribute to the observed abrogation of transcriptional activity in EA.hy926 cells.

5.5 Effect of TGF-β1 on *BGN* gene expression and TF binding

The multifunctional cytokine TGF-β1 has been described to be a key regulator of ECM assembly and remodeling. TGF-β1 is well known to modify BGN at the posttranslational level (Little et al., 2008) by hyperelongation of chondroitin sulfate chains, leading to enhanced BGN binding affinity to LDL. However, inconsistent TGF-β1 effects on *BGN* gene expression have been reported. Ungefroren & Krull (Ungefroren & Krull, 1996) did not observe any effect on transcriptional activity of *BGN* promoter fragments, while Heegaard and colleagues (Heegaard et al., 2004) reported increased *BGN* mRNA levels as well as increased reporter gene activity. Here we report a significant enhancing effect of TGF-β1 on both, mRNA levels and promoter transcriptional activity. This effect was however limited to THP-1 monocytes, whereas no such effect was observed in EA.hy926 cells. Controverse findings from other groups could thus as well originate from different cell lines analysed. To reveal the mechanistic process of TGF-β1 signal transduction towards the *BGN* promoter, we used nuclear extracts derived from TGF-β1 treated cells in EMSA experiments and observed an altered DNA binding pattern at positions G-578A. Using SP1-specific antibodies we identified selective SP1 binding at this distal position in EA.hy926 and THP-1 cells after TGF-β1 stimulation, while no binding of SP1 was detected under basic conditions.

Previously published reports have established a role for SP1 as the primary TF to be essential for the response to the TGF-β1 signal (Datto et al., 1995) and consequently, TGF-β1 induction mediated by SP1 has been shown for several genes (Inagaki et al., 1994; Li et al., 1998; Botella et al., 2001; Greenwel et al., 1997). Considering a simple linear effect of promoter activation by an increase in SP1 protein level, we assessed nuclear SP1 availability after TGF-β1 stimulation in comparison to basic conditions. We did not observe alterations in SP1 protein levels in neither cell line, but rather a slight tendency towards SP1 reduction in THP-1 cells. The observed effect might more likely be explained by enhanced SP1 binding to its consensus motifs. Recent reports have proposed a TGF-β1-activated SMAD2/3 phosphorylation pathway, along with p38MAPkinase signalling, to be responsible for the

TGF-β1-mediated elongation of GAG chains on proteoglycans. Subsequently, inhibiting SMAD2/3 phosphorylation antagonized the GAG elongation effect (Dadlani et al., 2008). Overexpression of both SMAD2 and SMAD3 potentiated the TGF-β1 effect on *BGN* expression, whereas overexpression of SMAD7 impaired the effect, suggesting the participation of the entire SMAD signalling cascade in TGF-β1 regulation of *BGN* gene expression (Chen et al., 2002; Ungefroren et al., 2005). This mechanism relies on the physical interaction of SMAD with SP1 and results in enhanced binding of the TF to its consensus motifs. High expression levels of TF SP1 in EA.hy926 cells under basic conditions could provide an explanation for the continuously high *BGN* promoter activity and the absence of any TGF-β1 effect in the endothelial cell line. Selective phosphorylation and nucleocytoplasmic shuttling of SMAD in case of THP-1 cells expressing SP1 at low levels could account for an orchestrated response to TGF-β1 in monocytes.

Characterizations of knock-out mice for TGF-β signalling components have furthermore demonstrated the critical role of SMAD proteins in vascular development and disease (ten Dijke & Arthur, 2007), with an essential function of SMAD3 in angiotensin II‑induced vascular fibrosis, characterized by arterial wall thickening through excessive deposition of ECM (Wang et al., 2006). The authors explicitly state that activation of SMAD3 but not SMAD2 is a key mechanism by which angiotensin II mediates arteriosclerosis. These observations might contribute to the positive association of SMAD3 with CAD (Samani et al., 2007) in the Welcome Trust Case Control Consortium study (which involved 1926 case subjects with CAD and 2938 controls) and combined analysis in the German MI Family Study (which involved 875 case subjects with MI and 1644 controls).

5.6 Conclusion

The impact of genetic variants on molecular biologic function has been repeatedly demonstrated. A *pars pro toto* example is given by the listed co-authorships including *in vitro* functional analyses on human ICAM-1, VCAM-1 and osteoprotegerin. Genetic variants located in regulatory regions do not directly alter amino acid sequences or protein structure, but can initiate profound transcriptional (i.e. regulatory) consequences. Individual genetic predisposition strongly determines the clinical progression of arteriosclerosis and CV outcome alongside with environmental and lifestyle factors as shown originally by twin

studies (Marenberg et al., 1994). Most recent whole-genome approaches with large study cohorts using DNA micro array technologies have provided new insight into aetiology of CVD and related diseases, generating a heterogeneous set of candidate genes and associated loci (Vasan et al., 2009; Samani et al., 2007; MIGC, 2009; Ikram et al., 2009; Newton-Cheh et al., 2009). To some extend surprising in the first place, common base to most of the identified loci was their position in intergenic regions rather than within or near protein coding regions. Simultaneously, advances in deciphering of gene regulatory mechanisms have underlined the pivotal role of non-coding DNA regions, executing their function despite their distal position to the multiple genes they control (Heintzman & Ren, 2009). Additional aspects of association with non coding DNA regions may exist in small-RNA mediated pathways that have central roles in the silencing of gene expression in eukaryotic cells, with at least 30% of human genes thought to be regulated by microRNAs (Jinek & Doudna, 2009). With the newly identified genetic variants available, profiling the functional basis of disease association and generation of mechanistic insights becomes ever more important (Brand-Herrmann, 2008). This includes not only analysis of single polymorphic positions towards their functionality but investigation of the entire structure of candidate gene promoters and interacting molecules with respect to pathophysiological conditions.

As shown in the current work, the analysis of naturally occurring MolHap constellations within a defined cellular and genomic context is needed to identify the effects of single genetic variants on TFBS and regulation of gene expression. We provide substantial evidence that MolHaps affect the transcriptional regulation of the human *BGN* gene to a great extent. This is caused by alterations in TFBS at *cis*-regulatory elements and a subsequent change in TF binding affinities (as shown explicitly for PU.1 and the -578A allele). We further propose the regulation of *BGN* by TGF-β1 signal transduction and SMAD proteins to act, at least in part, through increased binding of TF SP1 to the polymorphic positions. Due to the dual role of BGN in arteriosclerosis pathophysiology, estimation of potential effects in patients carrying either *BGN* MolHap seems rather conjectural. However, the following arguments will have to be considered. Local BGN availability is essential for collagen deposition and formation of stable fibrous caps in aortic plaques as well as for remodeling processes after MI. Reduced *BGN* gene expression due to the presence of MolHap2 and 3 could violate these processes, resulting in more severe outcomes. Looking at the other side of the medal, with systemic and *in situ* inflammatory processes influencing arteriosclerosis to a great extend, proinflammatory proteins such as BGN may execute their negative effects on disease

progression or early onset. *BGN* MolHap2 and 3 could therefore also account for protective effects, reducing *BGN* gene expression and the inflammatory burden.

6 PERSPECTIVE

The genetic predisposition for multifactorial and complex diseases such as CAD has become common knowledge. Large twin studies of the past 25 years provided comprehensive data numbering the genetic impact on disease susceptibility to at least 50%. Despite a rising number of proposed and analysed candidate genes, the multilayer and interlaced pathophysiology combined with different stages of the disease has circumvented explicit explanations of its development, progression and severity. GWA studies using large prospective and retrospective cohorts have provided large amounts of data widening the insight into the disease complexity even more, identifying significant associations of genetic variants with progression and severity of CAD. However, the molecular mechanisms of the identified variants are not well understood, impeding their use for therapeutic target identification or even risk prediction. It is obvious that understanding the functionality of human gene regulation and the impact of the interpersonal variable genome structure, especially in promoter regions, is mandatory. Since we have only started to understand the variable structure of the genome and its alterations throughout an individual's life span by epigenetic processes, this task seems unobtainable with currently available resources.

Molecular functional characterization of genetic variants, as presented exemplarily in this work, may contribute to our understanding of gene regulation by *cis*-active elements and *trans*-acting factors. The concept of individual MolHaps thereby is a next step accounting for an individual genome. Future work will have to account for the interaction within relevant pathophysiological pathways, considering the impact of genetic variants on cell-specific gene regulation. Identified functional variants within candidate genes are meanwhile integrated in latest approaches towards disease modelling (e.g. the VPH2 project), using data mining modules to identify combinations of genetic variants for risk prediction and identification of yet unknown underlying pathways and mechanisms.

The field of structural biology has so far provided three-dimensional (3D) structures of more than 45000 proteins. High-throughput methods for solving protein 3D structures are currently under development. Protein sequences can be predicted from their coding DNA and structural information can be used to predict the preferred substrate of a protein, identifying potentially recognized DNA sequences. Variations in the coding region of a gene can thus be translated into altered amino acid sequences and altered DNA binding capacity and kinetics. Most

promising technological advances thereby come from the field of NMR spectroscopy, decreasing size limitations and providing greater accuracy. As a result, information has become available on the way in which proteins interact with other biomolecules such as DNA. Coupling these advances with the exponential increase in data processing and storage could open the door for life imaging of gene regulation processes.

7 REFERENCES

Adiguzel E, Ahmad PJ, Franco C, Bendeck MP. Collagens in the progression and complications of atherosclerosis. Vasc Med. 2009 Feb;14(1):73-89.

Albright SR, Tjian R. TAFs revisited: more data reveal new twists and confirm old ideas. Gene. 2000 Jan 25;242(1-2):1-13.

Antequera F, Bird A. CpG islands. EXS. 1993;64:169-85.

Barrera LO, Ren B. The transcriptional regulatory code of eukaryotic cells--insights from genome-wide analysis of chromatin organization and transcription factor binding. Curr Opin Cell Biol. 2006 Jun;18(3):291-8.

Benyamin B, Visscher PM, McRae AF. Family-based genome-wide association studies. Pharmacogenomics. 2009 Feb;10(2):181-90.

Bernat JA, Crawford GE, Ogurtsov AY, Collins FS, Ginsburg D, Kondrashov AS. Distant conserved sequences flanking endothelial-specific promoters contain tissue-specific DNase-hypersensitive sites and over-represented motifs. Hum Mol Genet. 2006 Jul 1;15(13):2098-105.

Birnboim HC, Doly J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 1979 Nov 24;7(6):1513-23.

Botella LM, Sánchez-Elsner T, Rius C, Corbí A, Bernabéu C. Identification of a critical Sp1 site within the endoglin promoter and its involvement in the transforming growth factor-beta stimulation. J Biol Chem. 2001 Sep 14;276(37):34486-94.

Boyd KE, Wells J, Gutman J, Bartley SM, Farnham PJ. c-Myc target gene specificity is determined by a post-DNAbinding mechanism. Proc Natl Acad Sci USA. 1998 Nov 10;95(23):13887-92.

Brand-Herrmann SM. Where do we go for atherothrombotic disease genetics? Stroke. 2008 Apr;39(4):1070-5.

Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H. Sp1 elements protect a CpG island from de novo methylation. Nature. 1994 Sep 29;371(6496):435-8.

Briggs MR, Kadonaga JT, Bell SP, Tjian R. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science. 1986 Oct 3;234(4772):47-52.

Buratowski S, Hahn S, Guarente L, Sharp PA. Five intermediate complexes in transcription initiation by RNA polymerase II. Cell. 1989 Feb 24;56(4):549-61.

Buratowski S. Transcription. Gene expression--where to start? Science. 2008 Dec 19;322(5909):1804-5.

Burch ML, Yang SN, Ballinger ML, Getachew R, Osman N, Little PJ. TGF-beta stimulates biglycan synthesis via p38 and ERK phosphorylation of the linker region of Smad2. Cell Mol Life Sci. 2010 Jun;67(12):2077-90.

Cairns BR, Lorch Y, Li Y, Zhang M, Lacomis L, Erdjument-Bromage H, Tempst P, Du J, Laurent B, Kornberg RD. RSC, an essential, abundant chromatin-remodeling complex. Cell. 1996 Dec 27;87(7):1249-60.

Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engström PG, Frith MC, Forrest AR, Alkema WB, Tan SL, Plessy C, Kodzius R, Ravasi T, Kasukawa T, Fukuda S, Kanamori-Katayama M, Kitazume Y, Kawaji H, Kai C, Nakamura M, Konno H, Nakano K, Mottagui-Tabar S, Arner P, Chesi A, Gustincich S, Persichetti F, Suzuki H, Grimmond SM, Wells CA, Orlando V, Wahlestedt C, Liu ET, Harbers M, Kawai J, Bajic VB, Hume DA, Hayashizaki Y. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet. 2006 Jun;38(6):626-35.

Carrel L, Cottle AA, Goglin KC, Willard HF. A first-generation X-inactivation profile of the human X chromosome. Proc Natl Acad Sci USA. 1999 Dec 7;96(25):14440-4.

Chen WB, Lenschow W, Tiede K, Fischer JW, Kalthoff H, Ungefroren H. Smad4/DPC4-dependent regulation of biglycan gene expression by transforming growth factor-beta in pancreatic tumor cells. J Biol Chem. 2002 Sep 27;277(39):36118-28.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 2003 Jul 1;31(13):3497-500.

Cheong J, Yamada Y, Yamashita R, Irie T, Kanai A, Wakaguri H, Nakai K, Ito T, Saito I, Sugano S, Suzuki Y. Diverse DNA methylation statuses at alternative promoters of human genes in various tissues. DNA Res. 2006 Aug 31;13(4):155-67.

Chester N, Marshak DR. Dimethyl sulfoxide-mediated primer Tm reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. Anal Biochem. 1993 Mar;209(2):284-90.

Chiang CM, Roeder RG. Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. Science. 1995 Jan 27;267(5197):531-6.

Chinenov Y, Kerppola TK. Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. Oncogene. 2001 Apr 30;20(19):2438-52.

Corsi A, Xu T, Chen XD, Boyde A, Liang J, Mankani M, Sommer B, Iozzo RV, Eichstetter I, Robey PG, Bianco P, Young MF. Phenotypic effects of biglycan deficiency are linked to collagen fibril abnormalities, are synergized by decorin deficiency, and mimic Ehlers-Danlos-like changes in bone and other connective tissues. J Bone Miner Res. 2002 Jul;17(7):1180-9.

Courey AJ, Holtzman DA, Jackson SP, Tjian R. Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell. 1989 Dec 1;59(5):827-36.

Couzin-Frankel J. Major heart disease genes prove elusive. Science. 2010 Jun 4;328(5983):1220-1.

Dadlani H, Ballinger ML, Osman N, Getachew R, Little PJ. Smad and p38 MAP kinase-mediated signaling of proteoglycan synthesis in vascular smooth muscle. J Biol Chem. 2008 Mar 21;283(12):7844-52.

Datto MB, Yu Y, Wang XF. Functional analysis of the transforming growth factor beta responsive elements in the WAF1/Cip1/p21 promoter. J Biol Chem. 1995 Dec 1;270(48):28623-8.

De Val S, Black BL. Transcriptional control of endothelial cell development. Dev Cell. 2009 Feb;16(2):180-95.

Deng W, Roberts SG. A core promoter element downstream of the TATA box that is recognized by TFIIB. Genes Dev. 2005 Oct 15;19(20):2418-23.

Dördelmann C, Telgmann R, Brand E, Hagedorn C, Schröer B, Hasenkamp S, Baumgart P, Kleine-Katthöfer P, Paul M, Brand-Herrmann SM. Functional and structural profiling of the human thrombopoietin gene promoter. J Biol Chem. 2008 Sep 5;283(36):24382-91.

ten Dijke P, Arthur HM. Extracellular control of TGFbeta signalling in vascular development and disease. Nat Rev Mol Cell Biol. 2007 Nov;8(11):857-69.

Dodge GR, Diaz A, Sanz-Rodriguez C, Reginato AM, Jimenez SA. Effects of interferon-gamma and tumor necrosis factor alpha on the expression of the genes encoding aggrecan, biglycan, and decorin core proteins in cultured human chondrocytes. Arthritis Rheum. 1998 Feb;41(2):274-83.

Duane WC, Xiong W, Wolvers J. Effects of bile acids on expression of the human apical sodium dependent bile acid transporter gene. Biochim Biophys Acta. 2007 Nov;1771(11):1380-8.

Dwarki VJ, Montminy M, Verma IM. Both the basic region and the 'leucine zipper' domain of the cyclic AMP response element binding (CREB) protein are essential for transcriptional activation. EMBO J. 1990 Jan;9(1):225-32.

Ebner R, Chen RH, Shum L, Lawler S, Zioncheck TF, Lee A, Lopez AR, Derynck R. Cloning of a type I TGF-beta receptor and its effect on TGF-beta binding to the type II receptor. Science. 1993 May 28;260(5112):1344-8.

Edgell CJ, McDonald CC, Graham JB. Permanent cell line expressing human factor VIII-related antigen established by hybridization. Proc Natl Acad Sci USA. 1983 Jun;80(12):3734-7.

Ehret GB. Genome-wide association studies: contribution of genomics to understanding blood pressure and essential hypertension. Curr Hypertens Rep. 2010 Feb;12(1):17-25.

Feng R, Desbordes SC, Xie H, Tillo ES, Pixley F, Stanley ER, Graf T. PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. Proc Natl Acad Sci USA. 2008 Apr 22;105(16):6057-62.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998 Feb 19;391(6669):806-11.

Fischer DF, Backendorf C. Identification of regulatory elements by gene family footprinting and in vivo analysis. Adv Biochem Eng Biotechnol. 2007;104:37-64.

Fishbein GA, Fishbein MC. Arteriosclerosis: rethinking the current classification. Arch Pathol Lab Med. 2009 Aug;133(8):1309-16.

Fisher LW, Termine JD, Dejter SW Jr, Whitson SW, Yanagishita M, Kimura JH, Hascall VC, Kleinman HK, Hassell JR, Nilsson B. Proteoglycans of developing bone. J Biol Chem. 1983 May 25;258(10):6588-94.

Frisdal E, Klerkx AH, Le Goff W, Tanck MW, Lagarde JP, Jukema JW, Kastelein JJ, Chapman MJ, Guerin M. Functional interaction between -629C/A, -971G/A and -1337C/T polymorphisms in the CETP gene is a major determinant of promoter activity and plasma CETP concentration in the REGRESS Study. Hum Mol Genet. 2005 Sep 15;14(18):2607-18.

Frohman MA. Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. Methods Enzymol. 1993;218:340-56.

Fry CJ, Farnham PJ. Context-dependent transcriptional regulation. J Biol Chem. 1999 Oct 15;274(42):29583-6.

Funke-Kaiser H, Reichenberger F, Köpke K, Herrmann SM, Pfeifer J, Orzechowski HD, Zidek W, Paul M, Brand E. Differential binding of transcription factor E2F-2 to the endothelin-converting enzyme-1b promoter affects blood pressure regulation. Hum Mol Genet. 2003 Feb 15;12(4):423-33.

Gartel AL, Goufman E, Najmabadi F, Tyner AL. Sp1 and Sp3 activate p21 (WAF1/CIP1) gene transcription in the Caco-2 colon adenocarcinoma cell line. Oncogene. 2000 Oct 26;19(45):5182-8.

Geerkens C, Vetter U, Just W, Fedarko NS, Fisher LW, Young MF, Termine JD, Robey PG, Wöhrle D, Vogel W. The X-chromosomal human biglycan gene BGN is subject to X inactivation but is transcribed like an X-Y homologous gene. Hum Genet. 1995 Jul;96(1): 44-52.

Gentz R, Rauscher FJ 3rd, Abate C, Curran T. Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains. Science. 1989 Mar 31;243(4899):1695-9.

Grabe N. AliBaba2: context specific identification of transcription factor binding sites. In Silico Biol. 2002;2(1):S1-15.

Greenwel P, Inagaki Y, Hu W, Walsh M, Ramirez F. Sp1 is required for the early response of alpha2(I) collagen to transforming growth factor-beta1. J Biol Chem. 1997 Aug 8;272(32):19738-45.

Hagedorn C, Telgmann R, Dördelmann C, Schmitz B, Hasenkamp S, Cambien F, Paul M, Brand E, Brand-Herrmann SM. Identification and functional analyses of molecular haplotypes of the human osteoprotegerin gene promoter. Arterioscler Thromb Vasc Biol. 2009 Oct;29(10):1638-43.

Hagen G, Müller S, Beato M, Suske G. Cloning by recognition site screening of two novel GT box binding proteins: afamily of Sp1 related genes. Nucleic Acids Res. 1992 Nov 11;20(21):5519-25.

Hagemeier C, Bannister AJ, Cook A, Kouzarides T. The activation domain of transcription factor PU.1 binds the retinoblastoma (RB) protein and the transcription factor TFIID in vitro: RB shows sequence similarity to TFIID and TFIIB. Proc Natl Acad Sci USA. 1993 Feb 15;90(4):1580-4.

Hanahan D. Studies on transformation of Escherichia coli with plasmids. J Mol Biol. 1983 Jun 5;166(4):557-80.

Hansson GK. Regulation of immune mechanisms in atherosclerosis. Ann N Y Acad Sci. 2001 Dec;947:157-65.

Hansson GK. Inflammatory mechanisms in atherosclerosis. J Thromb Haemost. 2009 Jul;7 Suppl 1:328-31.

Hasenkamp S, Telgmann R, Staessen JA, Hagedorn C, Dördelmann C, Bek M, Brand-Herrmann SM, Brand E. Characterization and functional analyses of the human G protein-coupled receptor kinase 4 gene promoter. Hypertension. 2008 Oct;52(4):737-46.

Heegaard AM, Gehron Robey P, Vogel W, Just W, Widom RL, Schøller J, Fisher LW, Young MF. Functional characterization of the human biglycan 5'-flanking DNA and binding of the transcription factor c-Krox. J Bone Miner Res. 1997 Dec;12(12):2050-60.

Heegaard AM, Xie Z, Young MF, Nielsen KL. Transforming growth factor beta stimulation of biglycan gene expression is potentially mediated by sp1 binding factors. J Cell Biochem. 2004 Oct 15;93(3):463-75.

Heegaard AM, Corsi A, Danielsen CC, Nielsen KL, Jorgensen HL, Riminucci M, Young MF, Bianco P. Biglycan deficiency causes spontaneous aortic dissection and rupture in mice. Circulation. 2007 May 29;115(21):2731-8.

Heintzman ND, Ren B. Finding distal regulatory elements in the human genome. Curr Opin Genet Dev. 2009 Dec;19(6):541-9.

Higgs DR, Vernimmen D, Wood B. Long-range regulation of alpha-globin gene expression. Adv Genet. 2008;61:143-73.

Hirschi KK, Ingram DA, Yoder MC. Assessing identity, phenotype, and fate of endothelial progenitor cells. Arterioscler Thromb Vasc Biol. 2008 Sep;28(9):1584-95.

Hochheimer A, Tjian R. Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. 2003 Jun 1;17(11):1309-20.

Hoey T, Dynlacht BD, Peterson MG, Pugh BF, Tjian R. Isolation and characterization of the Drosophila gene encoding the TATA box binding protein, TFIID. Cell. 1990 Jun 29;61(7):1179-86.

Hoey T, Weinzierl RO, Gill G, Chen JL, Dynlacht BD, Tjian R. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell. 1993 Jan 29;72(2):247-60.

Hofnagel O, Luechtenborg B, Weissen-Plenz G, Robenek H. Statins and foam cell formation: impact on LDL oxidation and uptake of oxidized lipoproteins via scavenger receptors. Biochim Biophys Acta. 2007 Sep;1771(9):1117-24.

Hollenhorst PC, Jones DA, Graves BJ. Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. Nucleic Acids Res. 2004 Oct 21;32(18):5693-702.

Hristov M, Weber C. Ambivalence of progenitor cells in vascular repair and plaque stability. Curr Opin Lipidol. 2008 Oct;19(5):491-7.

Hsu HH, Duning K, Meyer HH, Stölting M, Weide T, Kreusser S, van Le T, Gerard C, Telgmann R, Brand-Herrmann SM, Pavenstädt H, Bek MJ. Hypertension in mice lacking the CXCR3 chemokine receptor. Am J Physiol Renal Physiol. 2009 Apr;296(4):F780-9.

Hsu JC, Cressman DE, Taub R. Promoter-specific trans-activation and inhibition mediated by JunB. Cancer Res. 1993 Aug 15;53(16):3789-94.

Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, Debette S, Lumley T, Folsom AR, van den Herik EG, Bos MJ, Beiser A, Cushman M, Launer LJ, Shahar E, Struchalin M, Du Y, Glazer NL, Rosamond WD, Rivadeneira F, Kelly-Hayes M, Lopez OL, Coresh J, Hofman A, DeCarli C, Heckbert SR, Koudstaal PJ, Yang Q, Smith NL, Kase CS, Rice K, Haritunians T, Roks G, de Kort PL, Taylor KD, de Lau LM, Oostra BA, Uitterlinden AG, Rotter JI, Boerwinkle E, Psaty BM, Mosley TH, van Duijn CM, Breteler MM, Longstreth WT Jr, Wolf PA. Genomewide association studies of stroke. N Engl J Med. 2009 Apr 23;360(17):1718-28.

Inagaki Y, Truter S, Ramirez F. Transforming growth factor-beta stimulates alpha 2(I) collagen gene expression through a cis-acting element that contains an Sp1-binding site. J Biol Chem. 1994 May 20;269(20):14828-34.

Iozzo RV. The biology of the small leucine-rich proteoglycans. Functional network of interactive proteins. J Biol Chem. 1999 Jul 2;274(27):18843-6.

Jinek M, Doudna JA. A three-dimensional view of the molecular machinery of RNA interference. Nature. 2009 Jan 22;457(7228):405-12.

Jonasson L, Holm J, Skalli O, Gabbiani G, Hansson GK. Expression of class II transplantation antigen on vascular smooth muscle cells in human atherosclerosis. J Clin Invest. 1985 Jul;76(1):125-31.

Kadonaga JT, Carner KR, Masiarz FR, Tjian R.Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell. 1987 Dec 24;51(6): 1079-90.

Kadonaga JT, Courey AJ, Ladika J, Tjian R. Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science. 1988 Dec 16;242(4885):1566-70.

Kalamajski S, Oldberg A. The role of small leucine-rich proteoglycans in collagen fibrillogenesis. Matrix Biol. 2010 May;29(4):248-53.

Kalinina N, Agrotis A, Antropova Y, Ilyinskaya O, Smirnov V, Tararak E, Bobik A. Smad expression in human atherosclerotic lesions: evidence for impaired TGF-beta/Smad signaling in smooth muscle cells of fibrofatty lesions. Arterioscler Thromb Vasc Biol. 2004 Aug;24(8):1391-6.

Kalitsis P, Saffery R. Inherent promoter bidirectionality facilitates maintenance of sequence integrity and transcription of parasitic DNA in mammalian genomes. BMC Genomics. 2009 Oct 27;10:498.

Kardassis D, Papakosta P, Pardali K, Moustakas A. c-Jun transactivates the promoter of the human p21(WAF1/Cip1) gene by acting as a superactivator of the ubiquitous transcription factor Sp1. J Biol Chem. 1999 Oct 8;274(41):29572-81.

Khachigian LM, Williams AJ, Collins T. Interplay of Sp1 and Egr-1 in the proximal platelet-derived growth factor A-chain promoter in cultured vascular endothelial cells. J Biol Chem. 1995 Nov 17;270(46):27679-86.

Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B. A high-resolution map of active promoters in the human genome. Nature. 2005 Aug 11;436(7052):876-80.

Kimoto S, Cheng SL, Zhang SF, Avioli LV. The effect of glucocorticoid on the synthesis of biglycan and decorin in human osteoblasts and bone marrow stromal cells. Endocrinology. 1994 Dec;135(6):2423-31.

Kingsley C, Winoto A. Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. Mol Cell Biol. 1992 Oct;12(10):4251-61.

Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet. 2005 Jan;76(1):8-32.

Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. Cell. 1990 Apr 6;61(1):113-24.

Kolodgie FD, Burke AP, Farb A, Weber DK, Kutys R, Wight TN, Virmani R. Differential accumulation of proteoglycans and hyaluronan in culprit lesions: insights into plaque erosion. Arterioscler Thromb Vasc Biol. 2002 Oct 1;22(10):1642-8.

Kornblihtt AR. Promoter usage and alternative splicing. Curr Opin Cell Biol. 2005 Jun;17(3):262-8.

Kouzarides T, Ziff E. The role of the leucine zipper in the fos-jun interaction. Nature. 1988 Dec 15;336(6200):646-51.

Krishnan P, Hocking AM, Scholtz JM, Pace CN, Holik KK, McQuillan DJ. Distinct secondary structures of the leucine-rich repeat proteoglycans decorin and biglycan. Glycosylation-dependent conformational stability. J Biol Chem. 1999 Apr 16;274(16):10945-50.

Krusius T, Ruoslahti E. Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA. Proc Natl Acad Sci USA. 1986 Oct;83(20):7683-7.

Kumar AP, Butler AP. Transcription factor Sp3 antagonizes activation of the ornithine decarboxylase promoter by Sp1. Nucleic Acids Res. 1997 May 15;25(10):2012-9.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinformatics. 2007 Nov 1;23(21):2947-8.

Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell. 2007 Oct 5;131(1):174-87.

Lee DH, Gershenzon N, Gupta M, Ioshikhes IP, Reinberg D, Lewis BA. Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. Mol Cell Biol. 2005b Nov;25(21):9674-86.

Lee MP, Howcroft K, Kotekar A, Yang HH, Buetow KH, Singer DS. ATG deserts define a novel core promoter subclass. Genome Res. 2005a Sep;15(9):1189-97.

Lee W, Mitchell P, Tjian R. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell. 1987a Jun 19;49(6):741-52.

Lee W, Haslinger A, Karin M, Tjian R. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature. 1987b Jan 22-28;325(6102):368-72.

Lemon B, Tjian R. Orchestrated response: a symphony of transcription factors for gene control. Genes Dev. 2000 Oct 15;14(20):2551-69.

Letovsky J, Dynan WS. Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence. Nucleic Acids Res. 1989 Apr 11;17(7):2639-53.

Li JM, Datto MB, Shen X, Hu PP, Yu Y, Wang XF. Sp1, but not Sp3, functions to mediate promoter activation by TGF-beta through canonical Sp1 binding sites. Nucleic Acids Res. 1998 May 15;26(10):2449-56.

Lijnen PJ, Petrov VV, Fagard RH. Induction of cardiac fibrosis by transforming growth factor-beta(1). Mol Genet Metab. 2000 Sep-Oct;71(1-2):418-35.

Lim CY, Santoso B, Boulay T, Dong E, Ohler U, Kadonaga JT. The MTE, a new core promoter element for transcription by RNA polymerase II. Genes Dev. 2004 Jul 1;18(13):1606-17.

Liu F, Patient R. Genome-wide analysis of the zebrafish ETS family identifies three genes required for hemangioblast differentiation or angiogenesis. Circ Res. 2008 Nov 7;103(10):1147-54.

Liu S, Spinner DS, Schmidt MM, Danielsson JA, Wang S, Schmidt J. Interaction of MyoD family proteins with enhancers of acetylcholine receptor subunit genes in vivo. J Biol Chem. 2000 Dec 29;275(52):41364-8.

Little PJ, Ballinger ML, Burch ML, Osman N. Biosynthesis of natural and hyperelongated chondroitin sulfate glycosaminoglycans: new insights into an elusive process. Open Biochem J. 2008;2:135-42.

Lower KM, Hughes JR, De Gobbi M, Henderson S, Viprakasit V, Fisher C, Goriely A, Ayyub H, Sloane-Stanley J, Vernimmen D, Langford C, Garrick D, Gibbons RJ, Higgs DR. Adventitious changes in long-range gene expression caused by polymorphic structural variation and promoter competition. Proc Natl Acad Sci USA. 2009 Dec 22;106(51):21771-6.

Majello B, De Luca P, Lania L. Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. J Biol Chem. 1997 Feb 14;272(7):4021-6.

Maldonado E, Shiekhattar R, Sheldon M, Cho H, Drapkin R, Rickert P, Lees E, Anderson CW, Linn S, Reinberg D. A human RNA polymerase II complex associated with SRB and DNA-repair proteins. Nature. 1996 May 2;381(6577):86-9.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature. 2009 Oct 8;461(7265):747-53.

Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic susceptibility to death from coronary heart disease in a study of twins. N Engl J Med. 1994 Apr 14;330(15):1041-6.

Massagué J, Gomis RR. The logic of TGFbeta signaling. FEBS Lett. 2006 May 22;580(12):2811-20.

McCracken S, Fong N, Yankulov K, Ballantyne S, Pan G, Greenblatt J, Patterson SD, Wickens M, Bentley DL. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature. 1997 Jan 23;385(6614):357-61.

Messeguer X, Escudero R, Farré D, Núñez O, Martínez J, Albà MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics. 2002 Feb;18(2):333-4.

Moore MJ, Proudfoot NJ. Pre-mRNA processing reaches back to transcription and ahead to translation. Cell. 2009 Feb 20;136(4):688-700.

Mullis KB. Target amplification for DNA analysis by the polymerase chain reaction. Ann Biol Clin (Paris). 1990;48(8):579-82.

Murphy C, Nikodem D, Howcroft K, Weissman JD, Singer DS. Active repression of major histocompatibility complex class I genes in a human neuroblastoma cell line. J Biol Chem. 1996 Nov 29;271(48):30992-9.

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. RNA polymerase is poised for activation across the genome. Nat Genet. 2007 Dec;39(12):1507-11.

Myers RH, Kiely DK, Cupples LA, Kannel WB. Parental history is an independent risk factor for coronary artery disease: the Framingham Study. Am Heart J. 1990 Oct;120(4):963-9.

Myocardial Infarction Genetics Consortium. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009 Mar;41(3):334-41.

Nacu N, Luzina IG, Highsmith K, Lockatell V, Pochetuhen K, Cooper ZA, Gillmeister MP, Todd NW, Atamas SP. Macrophages produce TGF-beta-induced (beta-ig-h3) following ingestion of apoptotic cells and regulate MMP14 levels and collagen turnover in fibroblasts. J Immunol. 2008 Apr 1;180(7):5036-44.

Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S, Papadakis K, Voight BF, Scott LJ, Zhang F, Farrall M, Tanaka T, Wallace C, Chambers JC, Khaw KT, Nilsson P, van der Harst P, Polidoro S, Grobbee DE, Onland-Moret NC, Bots ML, Wain LV, Elliott KS, Teumer A, Luan J, Lucas G, Kuusisto J, Burton PR, Hadley D, McArdle WL; Wellcome Trust Case Control Consortium, Brown M, Dominiczak A, Newhouse SJ, Samani NJ, Webster J, Zeggini E, Beckmann JS, Bergmann S, Lim N, Song K, Vollenweider P, Waeber G, Waterworth DM, Yuan X, Groop L, Orho-Melander M, Allione A, Di Gregorio A, Guarrera S, Panico S, Ricceri F, Romanazzi V, Sacerdote C, Vineis P, Barroso I, Sandhu MS, Luben RN, Crawford GJ, Jousilahti P, Perola M, Boehnke M, Bonnycastle LL, Collins FS, Jackson AU, Mohlke KL, Stringham HM, Valle TT, Willer CJ, Bergman RN, Morken MA, Döring A, Gieger C, Illig T, Meitinger T, Org E, Pfeufer A, Wichmann HE, Kathiresan S, Marrugat J, O'Donnell CJ, Schwartz SM, Siscovick DS, Subirana I, Freimer NB, Hartikainen AL, McCarthy MI, O'Reilly PF, Peltonen L, Pouta A, de Jong PE, Snieder H, van Gilst WH, Clarke R, Goel A, Hamsten A, Peden JF, Seedorf U, Syvänen AC, Tognoni G, Lakatta EG, Sanna S, Scheet P, Schlessinger D, Scuteri A, Dörr M, Ernst F, Felix SB, Homuth G, Lorbeer R, Reffelmann T, Rettig R, Völker U, Galan P, Gut IG, Hercberg S, Lathrop GM, Zelenika D, Deloukas P, Soranzo N, Williams FM, Zhai G, Salomaa V, Laakso M, Elosua R, Forouhi NG, Völzke H, Uiterwaal CS, van der Schouw YT, Numans ME, Matullo G, Navis G, Berglund G, Bingham SA, Kooner JS, Connell JM, Bandinelli S, Ferrucci L, Watkins H, Spector TD, Tuomilehto J, Altshuler D, Strachan DP, Laan M, Meneton P, Wareham NJ, Uda M, Jarvelin MR, Mooser V, Melander O, Loos RJ, Elliott P, Abecasis GR, Caulfield M, Munroe PB. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet. 2009 May 10;41(6):666-76.

Oettgen P. Regulation of vascular inflammation and remodeling by ETS factors. Circ Res. 2006 Nov 24;99(11):1159-66.

Osman N, Ballinger ML, Dadlani HM, Getachew R, Burch ML, Little PJ. p38 MAP kinase mediated proteoglycan synthesis as a target for the prevention of atherosclerosis. Cardiovasc Hematol Disord Drug Targets. 2008 Dec;8(4):287-92.

Ossipow V, Tassan JP, Nigg EA, Schibler U. A mammalian RNA polymerase II holoenzyme containing all components required for promoter-specific transcription initiation. Cell. 1995 Oct 6;83(1):137-46.

Ota T, Suzuki Y, Nishikawa T, Otsuki T, Sugiyama T, Irie R, Wakamatsu A, Hayashi K, Sato H, Nagai K, Kimura K, Makita H, Sekine M, Obayashi M, Nishi T, Shibahara T, Tanaka T, Ishii S, Yamamoto J, Saito K, Kawai Y, Isono Y, Nakamura Y, Nagahari K, Murakami K, Yasuda T, Iwayanagi T, Wagatsuma M, Shiratori A, Sudo H, Hosoiri T, Kaku Y, Kodaira H, Kondo H, Sugawara M, Takahashi M, Kanda K, Yokoi T, Furuya T, Kikkawa E, Omura Y, Abe K, Kamihara K, Katsuta N, Sato K, Tanikawa M, Yamazaki M, Ninomiya K, Ishibashi T, Yamashita H, Murakawa K, Fujimori K, Tanai H, Kimata M, Watanabe M, Hiraoka S, Chiba Y, Ishida S, Ono Y, Takiguchi S, Watanabe S, Yosida M, Hotuta T, Kusano J, Kanehori K, Takahashi-Fujii A, Hara H, Tanase TO,Nomura Y, Togiya S, Komai F, Hara R, Takeuchi K, Arita M, Imose N, Musashino K, Yuuki H, Oshima A, Sasaki N, Aotsuka S, Yoshikawa Y, Matsunawa H, Ichihara T, Shiohata N, Sano S, Moriya S, Momiyama H, Satoh N, Takami S, Terashima Y, Suzuki O, Nakagawa S, Senoh A, Mizoguchi H, Goto Y, Shimizu F, Wakebe H, Hishigaki H, Watanabe T, Sugiyama A, Takemoto M, Kawakami B, Yamazaki M, Watanabe K, Kumagai A, Itakura S, Fukuzumi Y, Fujimori Y, Komiyama M, Tashiro H, Tanigami A, Fujiwara T, Ono T, Yamada K, Fujii Y, Ozaki K, Hirao M, Ohmori Y, Kawabata A, Hikiji T, Kobatake N, Inagaki H, Ikema Y, Okamoto S, Okitani R, Kawakami T, Noguchi S, Itoh T, Shigeta K, Senba T, Matsumura K, Nakajima Y, Mizuno T, Morinaga M, Sasaki M, Togashi T, Oyama M, Hata H, Watanabe M, Komatsu T, Mizushima-Sugano J, Satoh T, Shirai Y, Takahashi Y, Nakagawa K, Okumura K, Nagase T, Nomura N, Kikuchi H, Masuho Y, Yamashita R, Nakai K, Yada T, Nakamura Y, Ohara O, Isogai T, Sugano S. Complete sequencing and characterization of 21,243 full-length human cDNAs. Nat Genet. 2004 Jan;36(1):40-5.

Pessah M, Prunier C, Marais J, Ferrand N, Mazars A, Lallemand F, Gauthier JM, Atfi A. c-Jun interacts with the corepressor TG-interacting factor (TGIF) to suppress Smad2 transcriptional activity. Proc Natl Acad Sci USA. 2001 May 22;98(11):6198-203.

Pham VN, Lawson ND, Mugford JW, Dye L, Castranova D, Lo B, Weinstein BM. Combinatorial function of ETS transcription factors in the developing vasculature. Dev Biol. 2007 Mar 15;303(2):772-83.

Piipari M, Down TA, Hubbard TJ. Metamotifs--a generative model for building families of nucleotide position weight matrices. BMC Bioinformatics. 2010 Jun 25;11:348.

Pinzani M, Marra F. Cytokine receptors and signaling in hepatic stellate cells. Semin Liver Dis. 2001 Aug;21(3):397-416.

Pugh BF, Tjian R. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. Genes Dev. 1991 Nov;5(11):1935-45.

Riessen R, Isner JM, Blessing E, Loushin C, Nikol S, Wight TN. Regional differences in the distribution of the proteoglycans biglycan and decorin in the extracellular matrix of atherosclerotic and restenotic human coronary arteries. Am J Pathol. 1994 May;144(5): 962-74.

Ritchie S, Boyd FM, Wong J, Bonham K. Transcription of the human c-Src promoter is dependent on Sp1, a novel pyrimidine binding factor SPy, and can be inhibited by triplex-forming oligonucleotides. Biol Chem. 2000 Jan 14;275(2):847-54.

Rittenhouse J, Marcus F. Peptide mapping by polyacrylamide gel electrophoresis after cleavage at aspartyl-prolyl peptide bonds in sodium dodecyl sulfate-containing buffers. Anal Biochem. 1984 May 1;138(2):442-8.
Robertson AK, Hansson GK. T cells in atherogenesis: for better or for worse? Arterioscler Thromb Vasc Biol. 2006 Nov;26(11):2421-32.

Rosa A, Ballarino M, Sorrentino A, Sthandier O, De Angelis FG, Marchioni M, Masella B, Guarini A, Fatica A, Peschle C, Bozzoni I. The interplay between the master transcription factor PU.1 and miR-424 regulates human monocyte/macrophage differentiation. Proc Natl Acad Sci USA. 2007 Dec 11;104(50):19849-54.

Ross J. mRNA stability in mammalian cells. Microbiol Rev. 1995 Sep;59(3):423-50.

Ruoslahti E, Yamaguchi Y. Proteoglycans as modulators of growth factor activities. Cell. 1991 Mar 8;64(5):867-9.

Russo MW, Sevetson BR, Milbrandt J. Identification of NAB1, a repressor of NGFI-A- and Krox20-mediated transcription. Proc Natl Acad Sci USA. 1995 Jul 18;92(15):6873-7.

Ryu S, Zhou S, Ladurner AG, Tjian R. The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1. Nature. 1999 Feb 4;397(6718):446-50.

Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, Dixon RJ, Meitinger T, Braund P, Wichmann HE, Barrett JH, König IR, Stevens SE, Szymczak S, Tregouet DA, Iles MM, Pahlke F, Pollard H, Lieb W, Cambien F, Fischer M, Ouwehand W, Blankenberg S, Balmforth AJ, Baessler A, Ball SG, Strom TM, Braenne I, Gieger C, Deloukas P, Tobin MD, Ziegler A, Thompson JR, Schunkert H; WTCCC and the Cardiogenics Consortium. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007 Aug 2;357(5):443-53.

Sambrook J, Russell DW. Molecular cloning: a laboratory manual. CSHL Press, Cold Spring Harbor, New York. 2001.

Sandelin A, Carninci P, Lenhard B, Ponjavic J, Hayashizaki Y, Hume DA. Mammalian RNA polymerase II core promoters: insights from genome-wide studies. Nat Rev Genet. 2007 Jun;8(6):424-36.

Sanz J, Fayad ZA. Imaging of atherosclerotic cardiovascular disease. Nature. 2008 Feb 21;451(7181):953-7.

Sardo MA, Mandraffino G, Campo S, Saitta C, Bitto A, Alibrandi A, Riggio S, Imbalzano E, Saitta A. Biglycan expression in hypertensive subjects with normal or increased carotid intima-media wall thickness. Clin Chim Acta. 2009 Aug;406(1-2):89-93.

Sata M, Saiura A, Kunisato A, Tojo A, Okada S, Tokuhisa T, Hirai H, Makuuchi M, Hirata Y, Nagai R. Hematopoietic stem cells differentiate into vascular cells that participate in the pathogenesis of atherosclerosis. Nat Med. 2002 Apr;8(4):403-9.

Schaefer L, Babelova A, Kiss E, Hausser HJ, Baliova M, Krzyzankova M, Marsche G, Young MF, Mihalik D, Götte M, Malle E, Schaefer RM, Gröne HJ. The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages. J Clin Invest. 2005 Aug;115(8):2223-33.
Schaefer L, Macakova K, Raslik I, Micegova M, Gröne HJ, Schönherr E, Robenek H, Echtermeyer FG, Grässel S, Bruckner P, Schaefer RM, Iozzo RV, Kresse H. Absence of decorin adversely influences tubulointerstitial fibrosis of the obstructed kidney by enhanced apoptosis and increased inflammatory reaction. Am J Pathol. 2002 Mar;160(3):1181-91.

Schmidt A, Lorkowski S, Seidler D, Breithardt G, Buddecke E. TGF-beta1 generates a specific multicomponent extracellular matrix in human coronary SMC. Eur J Clin Invest. 2006 Jul;36(7):473-82.

Schreiber E, Matthias P, Müller MM, Schaffner W. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. Nucleic Acids Res. 1989 Aug 11;17(15):6419.

Schreiber M, Kolbus A, Piu F, Szabowski A, Möhle-Steinlein U, Tian J, Karin M, Angel P, Wagner EF. Control of cell cycle progression by c-Jun is p53 dependent. Genes Dev. 1999 Mar 1;13(5):607-19.

Schultz J, Lorenz P, Ibrahim SM, Kundt G, Gross G, Kunz M. The functional -443T/C osteopontin promoter polymorphism influences osteopontin gene expression in melanoma cells via binding of c-Myb transcription factor. Mol Carcinog. 2009 Jan;48(1):14-23.

Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science. 1994 Sep 9;265(5178):1573-7.

Scotto-Lavino E, Du G, Frohman MA. 5' end cDNA amplification using classic RACE. Nat Protoc. 2006;1(6):2555-62.

Segal E, Widom J. What controls nucleosome positions? Trends Genet. 2009 Aug;25(8): 335-43.

Shah R, Hurley CK, Posch PE. A molecular mechanism for the differential regulation of TGF-beta1 expression due to the common SNP -509C-T (c. -1347C > T). Hum Genet. 2006 Nov;120(4):461-9.

Sharma K, Jin Y, Guo J, Ziyadeh FN. Neutralization of TGF-beta by anti-TGF-beta antibody attenuates kidney hypertrophy and the enhanced extracellular matrix gene expression in STZ-induced diabetic mice. Diabetes. 1996 Apr;45(4):522-30.

Shao H, Revach M, Moshonov S, Tzuman Y, Gazit K, Albeck S, Unger T, Dikstein R. Core promoter binding by histone-like TAF complexes. Mol Cell Biol. 2005 Jan;25(1):206-19.

Shaulian E, Karin M. AP-1 as a regulator of cell life and death. Nat Cell Biol. 2002 May;4(5):E131-6.

Skålén K, Gustafsson M, Rydberg EK, Hultén LM, Wiklund O, Innerarity TL, Borén J. Subendothelial retention of atherogenic lipoproteins in early atherosclerosis. Nature. 2002 Jun 13;417(6890):750-4.

Smale ST, Kadonaga JT. The RNA polymerase II core promoter. Annu Rev Biochem. 2003;72:449-79.

Steinberg D. The LDL modification hypothesis of atherogenesis: an update. J Lipid Res. 2009 Apr;50 Suppl:S376-81.

Suthanthiran M, Li B, Song JO, Ding R, Sharma VK, Schwartz JE, August P. Transforming growth factor-beta 1 hyperexpression in African-American hypertensives: A novel mediator of hypertension and/or target organ damage. Proc Natl Acad Sci USA. 2000 Mar 28;97(7):3479-84.

Tagle DA, Koop BF, Goodman M, Slightom JL, Hess DL, Jones RT. Embryonic epsilon and gamma globin genes of a prosimian primate (Galago crassicaudatus). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. J Mol Biol. 1988 Sep 20;203(2):439-55.

Takai D, Jones PA.Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci USA. 2002 Mar 19;99(6):3740-5.

Tan L, Peng H, Osaki M, Choy BK, Auron PE, Sandell LJ, Goldring MB. Egr-1 mediates transcriptional repression of COL2A1 promoter activity by interleukin-1beta. J Biol Chem. 2003 May 16;278(20):17688-700.

Telgmann R, Dördelmann C, Brand E, Nicaud V, Hagedorn C, Pavenstädt H, Cambien F, Tiret L, Paul M, Brand-Herrmann SM. Molecular genetic analysis of a human insulin-like growth factor 1 promoter P1 variation. FASEB J. 2009 May;23(5):1303-13.

Thomas MC, Chiang CM. The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol. 2006 May-Jun;41(3):105-78.

Thyagarajan T, Sreenath T, Cho A, Wright JT, Kulkarni AB. Reduced expression of dentin sialophosphoprotein is associated with dysplastic dentin in mice overexpressing transforming growth factor-beta 1 in teeth. J Biol Chem. 2001 Apr 6;276(14):11016-20.

Tiede K, Melchior-Becker A, Fischer JW. Transcriptional and posttranscriptional regulators of biglycan in cardiac fibroblasts. Basic Res Cardiol. 2010 Jan;105(1):99-108.

Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA. 1979 Sep;76(9):4350-4.

Trinklein ND, Karaöz U, Wu J, Halees A, Force Aldred S, Collins PJ, Zheng D, Zhang ZD, Gerstein MB, Snyder M, Myers RM, Weng Z. Integrated analysis of experimental data sets reveals many novel promoters in 1% of the human genome. Genome Res. 2007 Jun;17(6):720-31.

Trojanowska M. Ets factors and regulation of the extracellular matrix. Oncogene. 2000 Dec 18;19(55):6464-71.

Turner R, Tjian R. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional cFos-cJun heterodimers. Science. 1989 Mar 31;243(4899):1689-94.

Ungefroren H, Krull NB. Transcriptional regulation of the human biglycan gene. J Biol Chem. 1996 Jun 28;271(26):15787-95.

Ungefroren H, Lenschow W, Chen WB, Faendrich F, Kalthoff H.Regulation of biglycan gene expression by transforming growth factor-beta requires MKK6-p38 mitogen-activated protein Kinase signaling downstream of Smad signaling. J Biol Chem. 2003 Mar 28;278(13): 11041-9.

Ungefroren H, Groth S, Ruhnke M, Kalthoff H, Fändrich F. Transforming growth factor-beta (TGF-beta) type I receptor/ALK5-dependent activation of the GADD45beta gene mediates the induction of biglycan expression by TGF-beta. J Biol Chem. 2005 Jan 28;280(4):2644-52.

Valladares A, Hernández NG, Gómez FS, Curiel-Quezada E, Madrigal-Bujaidar E, Vergara MD, Martínez MS, Arenas Aranda DJ. Genetic expression profiles and chromosomal alterations in sporadic breast cancer in Mexican women. Cancer Genet Cytogenet. 2006 Oct 15;170(2):147-51.

Vasan RS, Glazer NL, Felix JF, Lieb W, Wild PS, Felix SB, Watzinger N, Larson MG, Smith NL, Dehghan A, Grosshennig A, Schillert A, Teumer A, Schmidt R, Kathiresan S, Lumley T, Aulchenko YS, König IR, Zeller T, Homuth G, Struchalin M, Aragam J, Bis JC, Rivadeneira F, Erdmann J, Schnabel RB, Dörr M, Zweiker R, Lind L,Rodeheffer RJ, Greiser KH, Levy D, Haritunians T, Deckers JW, Stritzke J, Lackner KJ, Völker U, Ingelsson E, Kullo I, Haerting J, O'Donnell CJ, Heckbert SR, Stricker BH, Ziegler A, Reffelmann T, Redfield MM, Werdan K, Mitchell GF, Rice K, Arnett DK, Hofman A, Gottdiener JS, Uitterlinden AG, Meitinger T, Blettner M, Friedrich N, Wang TJ, Psaty BM, van Duijn CM, Wichmann HE, Munzel TF, Kroemer HK, Benjamin EJ, Rotter JI, Witteman JC, Schunkert H, Schmidt H, Völzke H, Blankenberg S. Genetic variants associated with cardiac structure and function: a meta-analysis and replication of genome-wide association data. JAMA. 2009 Jul 8;302(2):168-78.

Verrecchia F, Rossert J, Mauviel A. Blocking sp1 transcription factor broadly inhibits extracellular matrix gene expression in vitro and in vivo: implications for the treatment of tissue fibrosis. J Invest Dermatol. 2001 May;116(5):755-63.

Villard J. Transcription regulation and human diseases. Swiss Med Wkly. 2004 Oct 2;134 (39-40):571-9.

Wakaguri H, Yamashita R, Suzuki Y, Sugano S, Nakai K. DBTSS: database of transcription start sites, progress report 2008. Nucleic Acids Res. 2008 Jan;36:D97-101.

Wang W, Huang XR, Canlas E, Oka K, Truong LD, Deng C, Bhowmick NA, Ju W, Bottinger EP, Lan HY. Essential role of Smad3 in angiotensin II-induced vascular fibrosis. Circ Res. 2006 Apr 28;98(8):1032-9.

Wegrowski Y, Pillarisetti J, Danielson KG, Suzuki S, Iozzo RV. The murine biglycan: complete cDNA cloning, genomic organization, promoter function, and expression. Genomics. 1995 Nov 1;30(1):8-17.

West AG, Fraser P. Remote control of gene transcription. Hum Mol Genet. 2005 Apr 15;14 Spec No 1:R101-11.

Westermann D, Mersmann J, Melchior A, Freudenberger T, Petrik C, Schaefer L, Lüllmann-Rauch R, Lettau O, Jacoby C, Schrader J, Brand-Herrmann SM, Young MF, Schultheiss HP, Levkau B, Baba HA, Unger T, Zacharowski K, Tschöpe C, Fischer JW. Biglycan is required for adaptive remodeling after myocardial infarction. Circulation. 2008 Mar 11;117(10):1269-76.

Wiradjaja F, DiTommaso T, Smyth I. Basement membranes in development and disease. Birth Defects Res C Embryo Today. 2010 Mar;90(1):8-31.

Wong FH, Huang CY, Su LJ, Wu YC, Lin YS, Hsia JY, Tsai HT, Lee SA, Lin CH, Tzeng CH, Chen PM, Chen YJ, Liang SC, Lai JM, Yen CC. Combination of microarray profiling and protein-protein interaction databases delineates the minimal discriminators as a metastasis network for esophageal squamous cell carcinoma. Int J Oncol. 2009 Jan;34(1):117-28.

World Health Organization, 2009 Global health risks: mortality and burden of disease attributable to selected major risks.

Wu Y, Zhang X, Zehner ZE. c-Jun and the dominant-negative mutant, TAM67, induce vimentin gene expression by interacting with the activator Sp1. Oncogene. 2003 Dec 4;22(55):8891-901.

Wutz A, Gribnau J. X inactivation Xplained. Curr Opin Genet Dev. 2007 Oct;17(5):387-93.

Yamashita R, Wakaguri H, Sugano S, Suzuki Y, Nakai K. DBTSS provides a tissue specific dynamic view of Transcription Start Sites. Nucleic Acids Res. 2010 Jan;38:D98-104.

Yan ZQ, Hansson GK. Innate immunity, macrophage activation, and atherosclerosis. Immunol Rev. 2007 Oct;219:187-203.

Yu B, Datta PK, Bagchi S. Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites. Nucleic Acids Res. 2003 Sep 15;31(18):5368-76.

Zeng J, Zhu S, Yan H. Towards accurate human promoter recognition: a review of currently used sequence features and classification methods. Brief Bioinform. 2009 Sep;10(5):498-508.

Zhang MQ. Computational analyses of eukaryotic promoters. BMC Bioinformatics. 2007 Sep 27;8 Suppl 6:S3.

Zhou T, Chiang CM. The intronless and TATA-less human TAF(II)55 gene contains a functional initiator and a downstream promoter element. J Biol Chem. 2001 Jul 6;276(27):25503-11.

Electronic resources:

ALGGEN – PROMO (release 3.0.2):
http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3
(last access: 26.08.2010)

AliBaba (release 2.0):
http://www.gene-regulation.com/pub/programs/alibaba2/index.html?
(last access: 26.08.2010)

CpG Island Searcher (release 10/29/04):
http://www.cpgislands.com/
(last access: 28.07.2010)

Database of Transcriptional Start Sites (DBTSS, release 7.0):
http://dbtss.hgc.jp/
(last access: 03.09.2010)

European Molecular Biology Laboratory/European Bioinformatics Institute (EMBL-EBI):
http://www.ebi.ac.uk/
(last access: 12.08.2010)

International HapMap Project (browser release #28):
http://hapmap.ncbi.nlm.nih.gov/ (last access: 06.08.2010)
(last access: 06.08.2010)

National Center for Biotechnology Information (NCBI):
http://www.ncbi.nlm.nih.gov/guide/
(last access: 10.09.2010)

TRANSFAC (release 7.0):
http://www.gene-regulation.com/cgi-bin/pub/databases/transfac/search.cgi
(last access: 26.08.2010)

VPH2:
http://www.vph2.eu/
(last access: 4.09.2010)

8 CONFERENCES

Boris Schmitz, Bianca Schröer, Christina Rüßmann, Ralph Telgmann, Jens Fischer, François Cambien, Eva Brand, Martin Paul, Stefan-Martin Brand-Herrmann. Transcriptional Activity of the human biglycan promoter is altered by single nucleotide polymorphisms. 31. Wissenschaftlicher Kongress der Deutschen Liga zur Bekämpfung des Hochdrucks/Deutsche Hypertoniegesellschaft, Bochum 22.-24. November 2007.

Boris Schmitz, Christina Rüßmann, Ralph Telgmann, Jens Fischer, François Cambien, Eva Brand, Martin Paul, Stefan-Martin Brand-Herrmann. Transcriptional Activity of the human biglycan promoter is altered by single nucleotide polymorphisms. 49. Frühjahrstagung der DGPT, Mainz 11.-13. März 2008, Naunyn-Schmiedebergs Archive of Pharmacology 2008; 377, 482. Oral presentation.

Boris Schmitz, Christina Rüßmann, Ralph Telgmann, Jens Fischer, François Cambien, Eva Brand, Martin Paul, Stefan-Martin Brand-Herrmann. Transcriptional Activity of the human biglycan promoter is altered by single nucleotide polymorphisms. 74. Jahrestagung der Deutschen Gesellschaft für Kardiologie, Mannheim 27.-29. März 2008.

Boris Schmitz, Christina Rüßmann, Ralph Telgmann, Jens Fischer, François Cambien, Eva Brand, Martin Paul, Stefan-Martin Brand-Herrmann. Der Einfluss genetischer Varianten auf die Transkriptionsaktivität des humanen Biglycan-Promotors. 39. Kongress der Gesellschaft für Nephrologie (GfN)/41. Jahrestagung der Deutschen Arbeitsgemeinschaft für Klinische Nephrologie (DAGKN), Tübingen 27. - 30. September 2008.

Boris Schmitz, Ralph Telgmann, Eva Brand, Christina Rüßmann, Jens Fischer, Martin Paul, Stefan-Martin Brand-Herrmann. Human Biglycan gene promoter: Examination of Molecular Haplotypes. 50. Jahrestagung der Deutschen Gesellschaft für Experimentelle und Klinische Pharmakologie und Toxikologie (DGPT), Mainz 10. - 12. März 2009. Oral presentation.

Boris Schmitz, Ralph Telgmann, Eva Brand, Christina Rüßmann, Jens Fischer, Martin Paul, Stefan-Martin Brand-Herrmann. Profiling of Biglycan Molekular Promoter Haplotypes. 1. Jahrestagung der Deutschen Gesellschaft für Nephrologie, Göttingen 26. - 29 September 2009.

Boris Schmitz, Ralph Telgmann, Eva Brand, Christina Rüßmann, Jens Fischer, Martin Paul, Stefan-Martin Brand-Herrmann. Profiling of Biglycan Molecular Promoter Haplotypes. 51. Jahrestagung der Deutschen Gesellschaft für Experimentelle und Klinische Pharmakologie und Toxikologie (DGPT), Mainz 23. - 25. März 2010. Oral presentation.

Boris Schmitz, Ralph Telgmann, Eva Brand, Christina Rüßmann, Jens Fischer, Martin Paul, Stefan-Martin Brand-Herrmann. Profiling of Biglycan Molecular Promoter Haplotypes. 78th European artherosclerosis society congress, Hamburg 20. - 23. Juni 2010.

Boris Schmitz, Klaus Schmidt-Petersen, Eva Brand, Ralph Telgmann, Vivienne Nicaud, Claudia Hagedorn, J. Labreuche, Corinna Dördelmann, A. Elbaz, M. Gautier-Bertrand, Jens Fischer, A. Evans, C. Morrison, D. Arveiler, Monika Stoll, Pierre Amarenco, Francois Cambien, Martin Paul, Stefan-Martin Brand-Herrmann. Osteopontin gene variation and cardio/cerebrovascular disease phenotypes. 78th European Artherosclerosis Society Congress, Hamburg 20. - 23. Juni 2010.

Boris Schmitz, Ralph Telgmann, Eva Brand, Alois Rötrige, Jens Fischer, Martin Paul, Stefan-Martin Brand. Functional analyses of Biglycan molecular promoter haplotypes.15th annual meeting of the ECCR, Nice 8. - 10. Oktober 2010.

9 PUBLICATIONS

Hagedorn C, Telgmann R, Dördelmann C, Schmitz B, Hasenkamp S, Cambien F, Paul M, Brand E, Brand-Herrmann SM. Identification and Functional Analyses of Molecular Haplotypes of the Human Osteoprotegerin Gene Promoter. Arterioscler Thromb Vasc Biol. 2009 Oct;29(10):1638-43.

Vischer P, Telgmann R, Schmitz B, Hasenkamp S, Schmidt-Petersen K, Beining K, Huge A, Paul M, Amarenco P, Cambien F, Brand E, Brand-Herrmann SM. Molecular investigation of the functional relevance of missense variants of ICAM-1. Pharmacogenet Genomics. 2008 Nov;18(11):1017-9

Vischer P, Schmitz B, Brand E, Schmidt-Petersen K, Fabritius C, Büchele B, Paul M, Cambien F, Simmet T, Brand SM. Investigation of the functional relevance of VCAM-1 missense variants. J. Hypertens. 2010, in revision.

Dimitris Gatsios, John Garofalakis, Theodora Chrysanthakopoulou, Evanthia Tripoliti, Renata De Maria, Maria Grazia Franzosi, Boris Schmitz, Stefan-Martin Brand and Oberdan Parodi. Knowledge extraction in a population suffering from heart failure. ITAB 2010, in revision.

Danksagung

Mein besonderer Dank gilt an dieser Stelle Herrn Prof. Dr. Stefan-Martin Brand für die Bereitstellung des Themas und die Möglichkeit meine Arbeiten auf diversen nationalen und internationalen Kongressen vertreten zu dürfen. Sein Vertrauen in meine Arbeit hat mich stets motiviert. Auch möchte ich meinem Betreuer an der biologischen Fakultät der WWU, Prof. Dr. Dirk Prüfer danken. Er hat diese Arbeit mit intensiven Gesprächen unterstützt.

Ein ausdrücklicher Dank geht an alle, die im Labor der AG Prof. Brand direkt oder indirekt an dieser Arbeit beteiligt waren. Danke an Corinna Dördelmann, Claudia Hagedorn und Sandra Hasenkamp, die mir in der Anfangszeit zur Seite gestanden haben, sowie Bianca Schröer und Friedericke Bruns. Vielen Dank auch an die anderen Doktoranden Andrea Salomon, Katrin Guske und Mareike Herrmann, die mich immer auf Trab gehalten haben. Spezieller Dank gilt auch dem technischen Personal: Christine Fabritius, Karin Tegelkamp und Margit Käse. Besonders erwähnen möchte ich Alois Rötrige, der mir eine besondere Hilfe war. Danke Alois! Vielen Dank an Euch alle, für anregende Gespräche, praktische Hilfe und interessant Stunden außerhalb der Arbeitszeit. Nicht zuletzt möchte ich Dr. Ralph Telgmann danken. Für die Hilfe bei dieser Arbeit und auch persönlichen Rat.
Ein Dank geht auch an alle anderen Arbeitsgruppen im Haus, die stets mit Rat und Tat zur Stelle waren.

An dieser Arbeit waren auch meine Freunde und Familie maßgeblich beteiligt. Danke an alle die immer an mich geglaubt und mich unterstützt haben. Besonders danke ich meiner Schwester Helena und meiner Mutter. Danke für Eure Stärke.
Zuletzt: Papa, das ist für Dich!