

Deciphering Regulatory Patterns of Inflammatory Gene Expression From Interleukin-1–Stimulated Human Endothelial Cells

Herbert Mayer, Martin Bilban, Vladislav Kurtev, Florian Gruber, Oswald Wagner, Bernd R. Binder, Rainer de Martin

Objective—Endothelial cells comprise a key component of the inflammatory response. We set out to obtain a comprehensive overview of the immediate-early to early gene expression program of interleukin-1 (IL-1)–stimulated endothelial cells and to identify novel transcription factors and regulatory elements.

Methods and Results—Human umbilical vein endothelial cells (HUVECs) were stimulated with IL-1 for 0, 0.5, 1, 2.5, and 6 hours and analyzed using Affymetrix U133 microarrays. A total of 137 genes were found to be regulated >4-fold, including 18 transcription factors. The expression of selected genes was confirmed by real-time polymerase chain reaction. Cluster analysis was performed in order to group genes according to their expression profiles. To identify novel transcription factor–binding sites, the corresponding promoters were extracted from databases and analyzed for regulatory elements that were over-represented in specific clusters. Several potentially novel DNA binding sites were identified, and one was shown to specifically bind an IL-1–inducible protein from HUVEC.

Conclusions—These results demonstrate that in the early phase after stimulation, IL-1 evokes a complex gene expression program that includes positive but also negative (feedback) regulators of diverse endothelial cell functions. Furthermore, the identification of a new promoter regulatory element demonstrates the feasibility of the bioinformatics-driven approach to discover novel regulatory mechanisms. (*Arterioscler Thromb Vasc Biol.* 2004;24:1192-1198.)

Key Words: endothelial cells ■ inflammation ■ interleukin-1 ■ microarrays ■ transcription factors

During the inflammatory response, the endothelium undergoes a variety of changes, including increase of permeability, adhesiveness for leukocytes, and a switch from anticoagulant to procoagulant properties. This is accomplished predominantly by expression of genes encoding interleukins and chemoattractants (interleukin-1 [IL-1], IL-6, IL-8), cell adhesion molecules (E-selectin, intercellular adhesion molecule-1 [ICAM-1], vascular cell adhesion molecule-1), procoagulant molecules (tissue factor, plasminogen activator inhibitor-1), and several others (inducible nitric oxide synthase, cyclooxygenase), as well as the downregulation of certain genes.^{1–4} Together, they orchestrate the response of the endothelium to inflammatory noxae such as IL-1, tumor necrosis factor- α (TNF- α), or lipopolysaccharide.

This gene expression program requires the action of tightly controlled regulatory mechanisms. Transcriptional regulation appears to be of major importance, and several of the transcription factors (TFs) that are operative in the inflammatory context have been the subject of detailed

investigation (eg, nuclear factor κ B [NF- κ B], early growth response 1 [EGR1], and NAK-1).^{2,5,6} Their tight regulation is of outstanding importance because an exaggerated or prolonged activation may lead to chronic inflammation and cause detrimental effect, as this occurs in a variety of diseases. Appropriate mechanisms to restrict the inflammatory response in time and magnitude can be expected to exist, and some have been described.^{7–9}

With the aim of acquiring a comprehensive survey over the kinetics of the early gene expression program and of regulatory mechanisms of the endothelium during inflammation, we performed gene expression profiling of human umbilical vein endothelial cells (HUVECs) stimulated with IL-1. With the main focus on the identification of regulatory molecules, we present here several TFs that are novel in the context of endothelial cell (EC) biology. Furthermore, we used a new bioinformatics strategy that is based on the promoter analysis of clustered genes sharing a common regulatory behavior to delineate novel regulatory elements.

Received February 13, 2004; accepted March 10, 2004.

From the Department of Vascular Biology and Thrombosis Research (F.G., B.R.B., R.d.M.), the Clinical Institute of Medical and Chemical Laboratory Diagnostics, and Ludwig Boltzmann Institute of Clinical Experimental Oncology (M.B., O.W.), University of Vienna, Austria; and the Competence Center Biomolecular Therapeutics GmbH (H.M., V.K.), Vienna, Austria.

Correspondence to Rainer de Martin, PhD, Department of Vascular Biology and Thrombosis Research, University of Vienna, Brunnerstr 59, A-1235 Vienna, Austria. E-mail rainer.de.martin@univie.ac.at

© 2004 American Heart Association, Inc.

Arterioscler Thromb Vasc Biol. is available at <http://www.atvbaha.org>

DOI: 10.1161/01.ATV.0000131263.06296.77

Methods

Cell Culture

HUVECs were isolated and cultured on gelatin-coated cell culture dishes in M199 medium supplemented with 20% FCS, antibiotics, EC growth supplement and heparin as described.¹⁰

RNA Preparation

HUVECs were stimulated with 100 U/mL of human IL-1 (Bio-source) for various periods of time and total RNA isolated using the RNeasy kit (Qiagen) according to the instructions of the manufacturer.

Microarray Hybridization and Primary Data Analysis

Preparation of cRNA, hybridization, and scanning of the arrays were performed according to protocols of the manufacturer. A detailed description is available as an online supplement (online Methods, available at <http://www.ahajournals.org>). Hybridization to one set of human U133A GeneChips (Affymetrix) and scanning of the arrays were performed according to protocols of the manufacturer.¹¹ The arrays were scanned using the GeneArray scanner (Affymetrix). Image analysis was performed with GeneChip software (MAS 5.0; Affymetrix). Normalization was performed by global scaling, with the arrays scaled to an average intensity of 500. To be microarray experiment compliant,¹² the microarray data were submitted to Gene Expression Omnibus (GEO),¹³ which is the public microarray repository of the National Center for Biotechnology Information. A series record (GSE973) was created, describing the experiment in detail and providing a focal point for the 5 individual sample records (GSM15389 to GSM15393), representing the 5 time points of the experiment.

Clustering Analysis

After filtering of genes (ProbeSets) showing absolute calls of “absence” and difference calls of “no change” for all data points, cluster analysis was performed on the remaining set, with a limit of 4-fold regulation (137 genes) in at least one of the time points using the Euclidean distance and average linkage (EPCLUST) program. The K-means clustering algorithm was applied, and the number of clusters was set to 10.¹⁴ Correlation measure-based distance was rated superior to Euclidean distance, thereby emphasizing the general shape of the profile and not the absolute magnitude of the expression response. In an alternative approach, we also generated a functional classification system for the set of genes using the Gene Ontology mining tool, which is part of the Affymetrix NetAffx portal,¹⁵ as a basis; however, manual literature search had to be added to achieve a reliable annotation of the full data set. The individual functional clusters were then subjected to hierarchical clustering¹⁶ using Euclidean distance and average linkage. The

clustering procedure and the visualization process were performed using EPCLUST.

Real-Time Polymerase Chain Reaction Analysis

The reliability of the microarray data were confirmed by quantitative real-time-polymerase chain reaction (PCR) using a LightCycler (Roche) together with the 2-step real-time-PCR kit and SybrGreen detection (Roche), according to the instructions of the manufacturer. Briefly, 1 μ g of total RNA was reverse transcribed in a total volume of 20 μ L of RNA PCR Master Mix. The thermocycler was programmed to 1 cycle of 15 minutes reverse transcription at 42°C, 5 minutes denaturation at 96°C, and a final cooling step. The first-strand cDNA was then diluted 1:3 with water, and 1.5 μ L (representing 25 ng of total RNA) was used for real-time PCR. Primers were designed using the program Primer3.¹⁷ Primer sequences are listed in the online supplement (Table I). To choose sense and antisense primers that belong to different exons, the genomic organization of the genes was extracted from the University of California, Santa Cruz Human Genome Browser, data freeze July 2003.¹⁸ In all cases, the primer efficiency was determined using cDNA dilutions.

Comparative Multiple Promoter Analysis

The Java program TOUCAN was the primary software used for comparative promoter analyses of the coclustering genes.¹⁹ First, proximal promoter sequences (1 kb upstream and 0.2 kb downstream of the transcriptional start sites) were extracted from genomic databases. As a quality control, all sequences were checked against the UCSC Human Genome Browser,¹⁸ demonstrating an accuracy of \approx 95%. The remaining sequences were added manually. The TOUCAN tool MotifScanner, which searches the TRANSFAC database,²⁰ was used to detect TF binding sites (TFBSs) in the sets of sequences. The prior (stringency level) was set to a value of 0.1, and the human promoter set of the Eukaryotic Promoter Database (EPD) was chosen as third-order background model. The statistics tool of TOUCAN was applied to the data produced by MotifScanner in combination with the appropriate expected frequencies file (human EPD), thereby detecting over-represented features (showing positive significance values) in the sets of coclustering genes. The tool ModuleSearcher was used to scan the sequences for high-scoring combinations of TFBSs. Default parameters were applied, except that the allowed distance was raised to 1 kb. Additionally, the clusters were scanned by the tool MotifSampler to detect over-represented patterns in the sequence sets using a prior level of 0.2. Finally, the extracted patterns were compared with the TRANSFAC site table to determine whether they correspond to known TFBSs. Selected motifs were subjected to a sequence conservation analysis using the tool WebLogo.²¹ For this purpose, all sequences defined by a motif were analyzed and a graphic representation created, which displays the frequencies of bases at each position as the relative heights of letters.

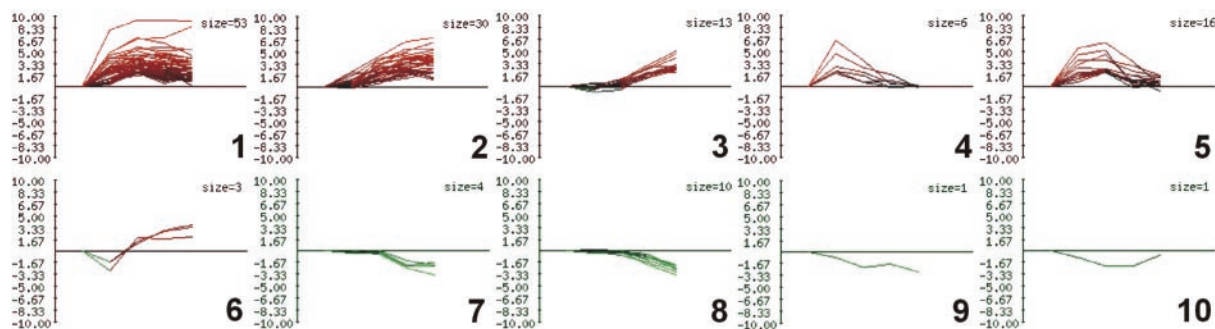


Figure 1. Cluster analysis of the 137 most regulated genes, showing fold changes higher than 4, as performed by the program EPCLUST. The K-means clustering algorithm was applied to the data set, and the number of clusters was set to a value of 10. Correlation measure-based distance was used for distance measure. The profile displays the time points of IL-1 treatment (0, 0.5, 1, 2.5, and 6 hours) at the x axis vs the corresponding signal log ratios at the y axis. Gene identifiers for each cluster are listed in the online supplement (Table III).

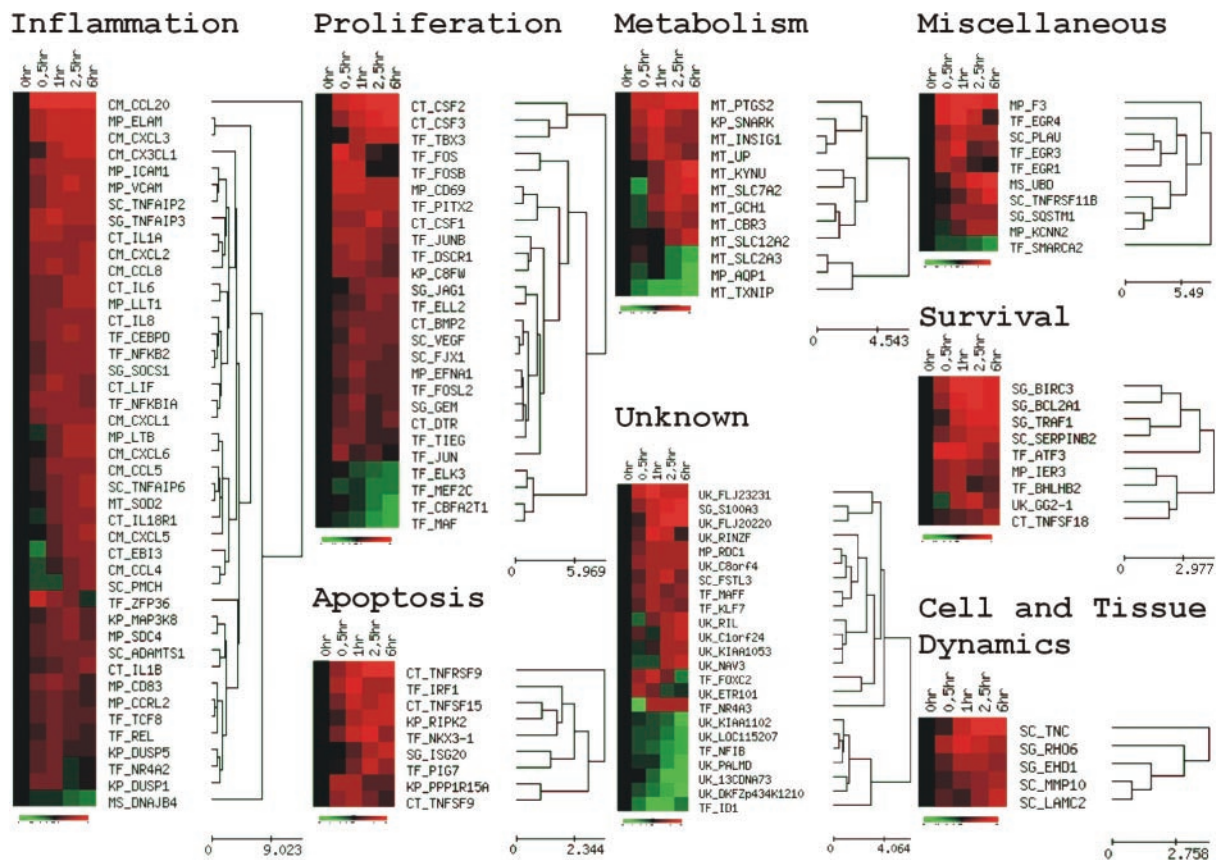


Figure 2. Functional classification and hierarchical clustering. The 137 genes induced by more than 4-fold or repressed to less than 25% of their uninduced state, were classified according to biological processes and subjected to hierarchical tree clustering using EPCLUST. Induction of expression is indicated by shades of red, whereas repression is represented by shades of green. CM indicates chemokines and chemokine receptors; CT, cytokines and cytokine receptors; KP, kinases and phosphatases; MP, membrane proteins; MS, miscellaneous; MT, metabolic proteins; SC, secreted proteins; SG, signaling proteins; TF, transcription factors; UK, unknown.

Nuclear Extracts and Electrophoretic Mobility Shift Assay

Nuclear proteins were extracted from IL-1-treated and control HUVECs as described.²² Double-stranded oligonucleotide probes for the motif over-represented in cluster 1 (C1: 5'-ggatccCGG/CG/CAG/ACGA/CCCgaattc-3') and FREAC7 (5'-cttaaACATAAACAagcatg-3';²³) were labeled with ³²P-γ-ATP using T4 polynucleotide kinase, and 100 000 cpm per binding reaction were used in electrophoretic mobility shift assay (EMSA) as described.²⁴ A 50-fold excess of unlabeled wild-type or mutant oligonucleotides was used for competition experiments (mutC1: 5'-ggatccCGG/CG/CAG/ACGA/CCCgaattc-3'; mut-FREAC7: 5'-cttaaACATCCACAgcatg-3'). Gels were dried and exposed on a PhosphorImager (Molecular Dynamics) and signals quantified using ImageQuant 5.0 software.

Results

Reprogramming of the Early EC Gene Expression Pattern Evoked by IL-1

To characterize gene expression and regulatory mechanisms in the early phase of the response of EC to IL-1, we performed microarray analysis of HUVECs. Cells were stimulated with IL-1 for various periods of time (0.5, 1, 2.5, and 6 hours), and the expression of 2 known markers of activation, IL-8 and E-selectin, was confirmed by real-time PCR (data not shown). The pattern of gene expression at each time point was then determined using Affymetrix U133A microarrays. Figure I (available online at <http://atvb.ahajournals.org>)

shows scattergrams comparing the signal intensities from IL-1-treated cells with those derived from untreated cells. Each time point thus represents the ratio of the signal intensities of mRNAs at the 2 time points. It becomes evident that at early time points (especially 0.5 hours), only a few genes are upregulated (“immediate early response”), and almost none are repressed. Although the number of induced genes stays about the same from 1 to 6 hours, downregulated genes markedly increase at later time points, particularly at 6 hours. The microarray data were submitted to the GEO¹³ microarray repository under the series record (GSE973) and the 5 individual sample records (GSM15389 to GSM15393).

On the basis of these data, we generated a data set of high reliability (see Methods), both in terms of absolute signal intensities and of significant changes of expression levels, using the criterion of more than 4-fold induction or repression in at least 1 of the time points. This data set consisted of 118 upregulated and 16 downregulated genes, with an additional 3 genes showing both effects. An extended description of these 137 genes, containing all values of fold changes sorted separately for each time point, is available in an online supplement (Table II, available online at <http://atvb.ahajournals.org>). To subgroup the data set according to the different profiles, the K-means clustering algorithm was applied,¹⁴ and the number of clusters, K, was set to 10 (Figure 1; Table III,

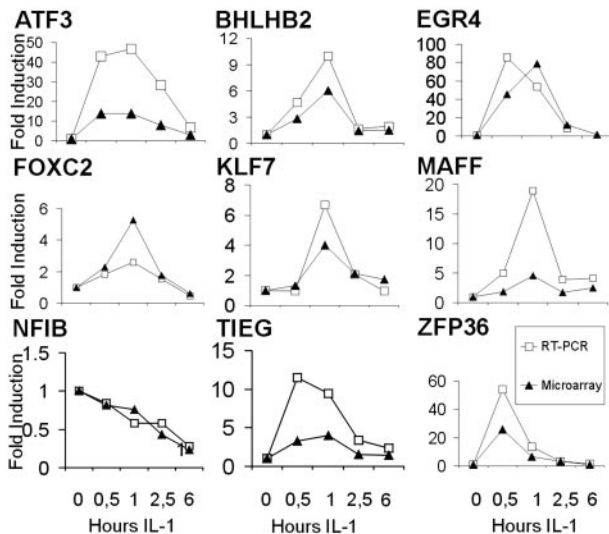


Figure 3. Comparison of microarray (closed symbols) and real-time-PCR (open symbols) analyses of selected genes. For real-time-PCR, primer efficiencies were determined using different dilutions of cDNA and values normalized to $\beta 2$ -microglobulin. The gene expression values are plotted as fold inductions (y axis) vs time points of IL-1 stimulation (x axis). For descriptions of the genes, please refer to the text.

available online at <http://atvb/ahajournals.org>). In an alternative approach, the set of 137 genes was functionally classified (Figure 2; Table IV, available online at <http://atvb/ahajournals.org>). For this purpose, genes are categorized according to “biological processes” and “molecular functions” and subjected to hierarchical clustering.¹⁶ The largest group is built of 43 genes (31%) that are known to be involved in processes of inflammation or immune modulation. A total of 26 genes (19%) have functions in proliferation or differentiation, and 12 genes (9%) are part of metabolic pathways. A total of 10 genes (7%) are described to fulfill different functions (“miscellaneous”). The categories apoptosis and survival are composed of 9 genes (7%) each. Only 5 genes (4%) are known to play a role in processes of cell and tissue dynamics. Finally, a considerable list of 23 genes (17%) is unknown in the context of a biological process, most of which are also in the context of molecular function. Interestingly, the last category shows the highest internal fraction of downregulated genes (30%), followed by metabolism (25%) and proliferation (15%).

Expression Profiles of Selected TFs Are Confirmed by Real-Time PCR

To substantiate the data obtained by microarray analysis, we performed real-time PCR on some selected genes. We chose TFs and other potentially regulatory molecules that were of special interest in the context of this study. Primers were chosen to avoid amplification of genomic DNA, and primer efficiencies determined using a standard curve. Although the magnitude of gene expression differed in some cases between the microarray and real-time PCR experiments, the kinetics of expression were usually very similar (Figure 3).

Comparative Regulatory Sequence Analysis of Coclustering Genes

We hypothesized that genes showing a common regulatory behavior may also share common regulatory mechanisms such as TFBSs in their respective promoter regions. To identify these possible common regulatory patterns that should be over-represented in certain clusters, we took advantage of the newly emerging TOUCAN bioinformatics tool.¹⁹ In a first step, promoters were extracted from genomic databases (see Methods). Second, a procedure of regulatory sequence analysis was performed in 3 stages, namely the extraction of cognate TFBSs, of combinations (“modules”) of TFBSs, and of sequence patterns. From each of these classes of sequence elements, several examples were found over-represented in certain clusters. The results are listed in parts A, B, and C, respectively, of the Table. Strikingly, the DNA binding site for NF- κ B, a well-characterized TF involved in the inflammatory response, was identified, thus serving as a positive control for this novel procedure. Please note that the total number of genes in clusters 1, 2, 3, and 8 differs between the Table and Figure 1 because of the unavailability of the promoter sequences for the genes INSIG1, FLJ20220, PMCH, KIAA1053, and DKFZp434K1210.

Identification of a Novel TFBS

To further confirm these in-silico predictions experimentally, we performed electrophoretic mobility shift analysis with the most prominent unknown DNA sequence, the CsGAGCGC element (“C1”) over-represented in clusters 1 and 5, and with a FREAC7 binding site. An NF- κ B binding site served as positive control. Using HUVEC nuclear extracts, a band was detected that bound to the double-stranded C1 oligonucleotide, showing a moderate but significant (1.8-fold) induction on IL-1 stimulation (Figure 4A); no binding was detected with the FREAC7 site. To test for specificity, we generated a mutant C1 oligonucleotide by substituting those residues that were highly conserved in the multiple sequence alignment of the C1 element derived from 29 different promoters (Figure 4B). Binding of the protein to the C1 element was inhibited by an excess of wild-type but not of mutant oligonucleotide, demonstrating the specificity of binding (Figure 4A).

Discussion

A large fraction of the 137 regulated genes, namely 53, coclustered in K-means cluster 1, exhibiting a peak of induction at 1 hour, and more or less staying at a constant level throughout the rest of the time course. Of these 53 genes, 21 are known to be involved in processes of inflammation or immune modulation, including many chemokines and chemokine receptors (CCL8, CCRL2, and CXCL1, 2, 3), cytokines (IL1 α , IL8, and LIF), phosphatases, and other signaling proteins (eg, DUSP1, DUSP5, SOCS1, and TNFAIP3), along with two well-known adhesion proteins (E-selectin and ICAM-1). These molecules also serve as additional positive controls for the microarray analysis.

In this study, we chose a narrow time course of IL-1 stimulation for up to 6 hours. In contrast to previous experiments,^{25,26} in which either later time points of stimulation or smaller numbers of genes have been investigated, this ap-

Regulatory Sequence Analysis of Coclustering Genes

A TFBS	Cluster No./Site in x of y =% of Promoters									
	1	2	3	4	5	6	7	8	9	10
CREB/ATF	12/52=23	3/29=10	1/11=9	5/6=83	9/16=56	1/3=33	1/4=25	2/9=22	1/1=100	1/1=100
GATA1	14/52=27	11/29=38	4/11=36	3/6=50	6/16=38	2/3=67	1/4=25	2/9=22	0/1=0	1/1=100
GATA2	8/52=15	3/29=10	1/11=9	1/6=17	3/16=19	0/3=0	0/4=0	4/9=44	0/1=0	0/1=0
GATA3	10/52=19	8/29=28	2/11=18	1/6=17	3/16=19	1/3=33	0/4=0	3/9=67	0/1=0	0/1=0
HSF1	4/52=8	2/29=7	0/11=0	1/6=17	1/16=6	1/3=33	0/4=0	4/9=44	0/1=0	1/1=100
NF-κB	28/52=54	6/29=21	4/11=36	1/6=17	8/16=50	2/3=67	0/4=0	1/9=11	0/1=0	0/1=0
OCT1	22/52=42	18/29=62	5/11=45	3/6=50	7/16=44	1/3=33	3/4=75	3/9=67	1/1=100	0/1=0
PBX1	12/52=23	9/29=31	3/11=27	1/6=17	3/16=19	2/3=67	3/4=75	0/9=0	0/1=0	0/1=0
RREB1	13/52=25	1/29=3	0/11=0	1/6=17	3/16=19	1/3=33	2/4=50	5/9=56	0/1=0	0/1=0
SP1	27/52=52	8/29=28	2/11=18	5/6=83	14/16=88	1/3=33	1/4=25	7/9=78	0/1=0	1/1=100
SREBP1	8/52=15	1/29=3	1/11=9	2/6=33	1/16=6	0/3=0	1/4=25	0/9=0	0/1=0	0/1=0
SRF	4/52=8	1/29=3	0/11=0	2/6=33	4/16=25	0/3=0	0/4=0	0/9=0	0/1=0	0/1=0
XBP1	5/52=10	0/29=0	0/11=0	0/6=0	2/16=13	2/3=67	0/4=0	1/9=11	0/1=0	0/1=0
Ap2	18/52=35	6/29=21	3/11=27	4/6=67	8/16=50	2/3=67	0/4=0	3/9=67	0/1=0	0/1=0
FOXD3	13/52=25	12/29=41	5/11=45	2/6=33	2/16=13	2/3=67	2/4=50	4/9=44	0/1=0	0/1=0
FREAC7	12/52=23	9/29=31	3/11=27	0/6=0	2/16=13	0/3=0	2/4=50	0/9=0	0/1=0	0/1=0
HFH3	13/52=25	9/29=31	2/11=18	3/6=50	2/16=13	2/3=67	3/4=75	1/9=11	0/1=0	1/1=100
PAX4	22/52=42	11/29=38	2/11=18	4/6=67	8/16=50	1/3=33	1/4=25	3/9=67	0/1=0	0/1=0
B Modules	1	2	3	4	5	6	7	8	9	10
NFκB+FREAC7	7/52=13	4/29=14	0/11=0	0/6=0	1/16=6	0/3=0	0/4=0	0/9=0	0/1=0	0/1=0
AP2+SP1	11/52=21	5/29=17	2/11=18	3/6=50	5/16=31	1/3=33	0/4=0	3/9=67	0/1=0	0/1=0
OCT1+FOXD3	9/52=17	9/29=31	3/11=27	2/6=33	1/16=6	1/3=33	3/4=75	1/9=11	0/1=0	0/1=0
ATF+HFH3+PAX4	0/52=0	0/29=0	0/11=0	3/6=50	1/16=6	0/3=0	0/4=0	0/9=0	0/1=0	0/1=0
AP2+CREB	2/52=4	1/29=3	0/11=0	4/6=67	6/16=38	1/3=33	0/4=0	1/9=11	0/1=0	0/1=0
CREB+SP1	6/52=12	1/29=3	0/11=0	3/6=50	6/16=38	0/3=0	0/4=0	1/9=11	0/1=0	1/1=100
C Motifs	1	2	3	4	5	6	7	8	9	10
CsGAGCGC	16/52=31	3/29=10	1/11=9	1/6=17	4/16=25	0/3=0	0/4=0	2/10=20	0/1=0	0/1=0
CTGGGATy	13/52=25	9/29=31	2/11=18	2/6=33	1/16=6	0/3=0	2/4=50	2/10=20	0/1=0	0/1=0
TTCTAGCT	0/52=0	2/29=7	4/11=36	0/6=0	0/16=0	0/3=0	2/4=50	0/9=0	0/1=0	0/1=0
sGGATsCG	5/52=10	0/29=0	0/11=0	4/6=67	3/16=19	0/3=0	0/4=0	0/9=0	1/1=100	0/1=0

s=G/C; y=T/C

Regulatory sequence analysis of coclustering genes. The proximal 1-kb promoter regions of the genes from different clusters were analyzed using the TOUCAN program. Sequences or DNA binding sites that are over-represented in individual clusters are indicated in bold. B, Modules defining combinations of TFBS were detected by the TOUCAN tool ModuleSearcher using a moderate stringency. C, Motifs are over-represented sequence patterns and were defined by the TOUCAN tool MotifSampler. All values represent 'perfect matches' of the given motifs within the respective promoter sequences.

proach has allowed us to study immediate-early to early regulatory mechanisms of EC activation. Indeed, within this time period, 36 TFs (26% of all regulated genes) that are either cognate or potentially novel in the context of inflammation were identified. Well-characterized TFs that mediate the immediate-early response are found in cluster 4 (JUN, FOS, EGR1). Similarly, cluster 5 shows a strong predominance of TFs (10 of the 16 genes), suggesting that these sets of genes peaking at 1 hour after stimulation govern the early response. The expression profiles of 8 TFs were confirmed by real-time-PCR (Figure 3). Of these, BHLHB2, EGR4, TIEG, MAFF, and NFIB (a downregulated TF) have not been described in ECs. These TFs are expected to trigger a second wave of gene expression to regulate specific aspects of the

inflammatory response at later time points; the identification of their target genes will be the subject of further studies. Notably, 5 downregulated TFs (CBFA2T1, ELK3, MAF, MEF2C, NFIB) are known to be involved in proliferation or differentiation, suggesting a concerted repression of these processes by IL-1. Consistent with these data are also the temporary downregulation of ID1 (inhibitor of DNA binding 1). ID1 is a potent repressor of thrombospondin-1 transcription,²⁷ a major inhibitor of neovascularization.

From a therapeutic point of view, induced genes with inhibitory function are of special interest because they may represent endogenous feedback mechanisms to shut down the inflammatory response at later times. In our experiments, induction of IκBα, an inhibitor of the TF NF-κB, is a

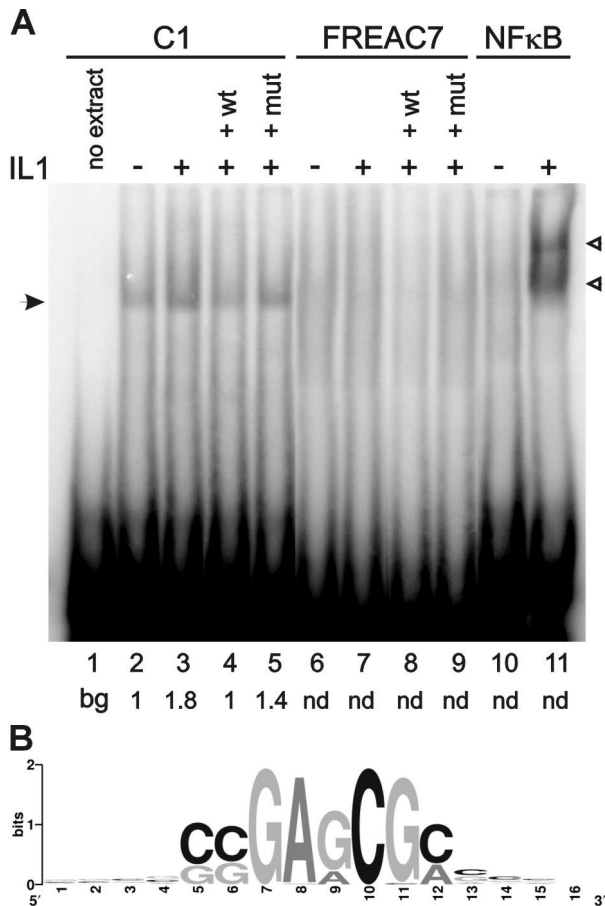


Figure 4. A, Electrophoretic mobility shift assay using the binding sites C1 (lanes 1 to 5) and FREAC7 (lanes 6 to 9). Extracts from IL-1-stimulated and control HUVECs, as well as competition with wild-type and mutant oligonucleotides, are indicated. Lanes 10 and 11 show binding of NF- κ B used as positive control. Lane 1, No extract. Arrowheads indicate specific bands. A representative experiment (of a total of 3) is shown. Fold induction (shown below lane numbers) was determined by densitometric quantification of the specific bands. Signal intensity in lane 2 was arbitrarily set to a value of 1. bg indicates background correction; nd, not determined. B, Sequence logo of motif CsGAGCGC (C1). This motif was found to be over-represented in the K-means cluster 1 by the tool MotifSampler. All sequences defined by this motif were extracted from cluster 1 and analyzed by the tool WebLogo. A graphic representation ("logo") was produced, which displays the degree of sequence conservation at each position (see Methods).

well-described example.²⁸ Another is DSCR1, a calcineurin inhibitor peaking at 1 hour, that has been described recently to shut down NF-AT activity.⁹ It will be interesting to investigate whether the latter mechanism is also operative in ECs. Last but not least, the genes encoding ZFP36 (Figure 3) and C8FW have been associated with negative regulation on several levels.^{29,30} In particular, ZFP36 has been shown to bind to and accelerate mRNA degradation,²⁹ as already demonstrated for TNF- α and granulocyte-macrophage colony-stimulating factor mRNAs. C8FW is highly homologous to a protein (SINK) that was characterized recently as a negative regulator of NF- κ B-dependent transcription via specific interaction with the p65/RelA transactivation domain.³¹ Our findings on the regulated expression of ZFP36

and C8FW in ECs will serve as a starting point for the investigation of multiple negative regulatory mechanisms during inflammation.

To investigate transcriptional regulation, researchers have traditionally dissected individual promoters to identify novel TFBSs. We reasoned that the large number of regulated genes derived from microarray experiments would allow a different strategy, namely a comparative promoter analysis. This is based on the hypothesis that coregulated genes, because they are grouped in individual clusters, might share common regulatory mechanisms that govern their expression. Promoters of genes grouped in certain clusters would thereby contain a different setup of regulatory elements compared with those in other clusters.

Recently, bioinformatics solutions for this purpose have emerged, including the Genomatix and TOUCAN program packages.¹⁹ We used the TOUCAN program, which is freely available, first to extract the promoter regions of all genes, and second, to analyze them with regard to TFBSs, combinations thereof, and sequence motifs. As listed in the Table, a number of over-represented TFBSs (compared with the average in the EPD database) and sequence motifs were identified. In cluster 1 (and also 5, which both peak at 1 hour), the binding site for NF- κ B was prominent, serving as an excellent positive control for the procedure. In contrast, NF- κ B sites play only a minor role in promoters of cluster 4 (immediate-early response genes peaking at 0.5 hours), and a moderate role in "late responders" (clusters 2 and 3). This is consistent with the expression kinetics of many known NF- κ B-dependent genes. In contrast, the binding sites for so-called "general" TFs such as SP-1 are more evenly distributed over the set of clusters, as expected. Another striking observation is the prominent occurrence of binding sites for CREB/ATF clusters of immediate-early (cluster 4) and early (cluster 5) induced genes compared with late responders (clusters 2 and 3) and downregulated genes.

The second part was the search for significant combinations of 2 or more TFBSs using the TOUCAN tool Module-Searcher (see the Table, B). This approach is based on the suggestion that rather the combinations of TFs are relevant for gene expression. Although parts of this hypothesis, especially the contribution of the distance between binding sites, remain a matter of debate, this approach has yielded several combinations of known TFBSs, involving AP2, FOXD3, FREAC7, HFH3, and PAX4. However, we were not able to confirm a role for FREAC7,²³ in EMSA (see Results). Therefore, the significance of this second approach remains to be established.

The third stage of the analysis concentrated on the search for over-represented motifs (patterns) in promoters of co-clustering genes using the TOUCAN tool MotifSampler. A motif length of 8 bases was chosen, and in the first run, hits allowing a certain degree of mismatches to the "perfect motif" were also produced. Finally, only "perfect matches" are listed in the Table (C), allowing for comparison between the different clusters. We decided to investigate further the motif C1 (CsGAGCGC, where s is C or G), which occurs in clusters 1 and 5. Searching this motif in the TRANSFAC site table did not yield any known TFBSs. Using EMSA, we

could demonstrate that a protein, possibly representing a novel TF, binds to this sequence. This protein also showed IL-1 inducibility, further supporting its involvement in regulated inflammatory gene expression. Its cloning and additional characterization will be the subject of future studies. Moreover, our results demonstrate the validity of the comparative promoter analysis approach that can be applied as a general follow-up study for gene expression profiling.

Acknowledgments

This work was supported by grants from the Austrian Science Foundation (SFB 005-11, SFB 005-12; to B.R.B., R.d.M., respectively).

References

- Collins T. Endothelial nuclear factor-kappa B and the initiation of the atherosclerotic lesion. *Lab Invest.* 1993;68:499–508.
- de Martin R, Hoeth M, Hofer-Warbinek R, Schmid JA. The transcription factor NF-kappa B and the regulation of vascular cell function. *Arterioscler Thromb Vasc Biol.* 2000;20:E83–E88.
- Mantovani A, Sozzani S, Vecchi A, Introna M, Allavena P. Cytokine activation of endothelial cells: new molecules for an old paradigm. *Thromb Haemost.* 1997;78:406–414.
- Pober JS. Activation and injury of endothelial cells by cytokines. *Pathol Biol (Paris).* 1998;46:159–163.
- Mechtcheriakova D, Schabbauer G, Lucerna M, Clauss M, De Martin R, Binder BR, Hofer E. Specificity, diversity, and convergence in VEGF and TNF- α signaling events leading to tissue factor up-regulation via EGR-1 in endothelial cells. *FASEB J.* 2001;15:230–242.
- Gruber F, Hufnagl P, Hofer-Warbinek R, Schmid JA, Breuss JM, Huber-Beckmann R, Lucerna M, Papac N, Harant H, Lindley I, de Martin R, Binder BR. Direct binding of Nur77/NAK-1 to the plasminogen activator inhibitor 1 (PAI-1) promoter regulates TNF α -induced PAI-1 expression. *Blood.* 2003;101:3042–3048.
- Lucerna M, Mechtcheriakova D, Kadl A, Schabbauer G, Schafer R, Gruber F, Koshelnick Y, Muller HD, Issbrucker K, Clauss M, Binder BR, Hofer E. NAB2, a corepressor of EGR-1, inhibits vascular endothelial growth factor-mediated gene induction and angiogenic responses of endothelial cells. *J Biol Chem.* 2003;278:11433–11440.
- Sun SC, Ganchi PA, Beraud C, Ballard DW, Greene WC. Autoregulation of the NF-kappa B transactivator RelA (p65) by multiple cytoplasmic inhibitors containing ankyrin motifs. *Proc Natl Acad Sci U S A.* 1994;91:1346–1350.
- Ryeom S, Greenwald RJ, Sharpe AH, McKeon F. The threshold pattern of calcineurin-dependent gene expression is altered by loss of the endogenous inhibitor calcipressin. *Nat Immun.* 2003;4:874–881.
- Zhang JC, Fabry A, Paucz L, Wojta J, Binder BR. Human fibroblasts downregulate plasminogen activator inhibitor type-1 in cultured human macrovascular and microvascular endothelial cells. *Blood.* 1996;88:3880–3886.
- Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, Patapoutian A, Hampton GM, Schultz PG, Hogenesch JB. Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A.* 2002;99:4465–4470.
- Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet.* 2001;29:365–371.
- Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002;30:207–210.
- Tavazoie S, Hughes JD, Campbell MJ, Cho RJ, Church GM. Systematic determination of genetic network architecture. *Nat Genet.* 1999;22:281–285.
- Liu G, Loraine AE, Shigeta R, Cline M, Cheng J, Valmeekam V, Sun S, Kulp D, Siani-Rose MA. NetAffx: affymetrix probesets and annotations. *Nucleic Acids Res.* 2003;31:82–86.
- Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A.* 1998;95:14863–14868.
- Rozen R, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S eds. *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Totowa, NJ: Humana Press; 2000:365–386.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res.* 2002;12:996–1006.
- Aerts S, Thijs G, Coessens B, Staes M, Moreau Y, De Moor B. Toucan: deciphering the cis-regulatory logic of coregulated genes. *Nucleic Acids Res.* 2003;31:1753–1764.
- Wingender E, Chen X, Fricke E, Geffers R, Hehl R, Liebich I, Krull M, Matys V, Michael H, Ohnhauser R, Pruss M, Schacherer F, Thiele S, Urbach S. The TRANSFAC system on gene expression regulation. *Nucleic Acids Res.* 2001;29:281–283.
- Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 1990;18:6097–6100.
- Dignam JD, Lebovitz RM, Roeder RG. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 1983;11:1475–1489.
- Pierrou S, Hellqvist M, Samuelsson L, Enerback S, Carlsson P. Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending. *EMBO J.* 1994;13:5002–5012.
- Cheng Q, Cant CA, Moll T, Hofer-Warbinek R, Wagner E, Birnstiel ML, Bach FH, de Martin R. NK-kappa B subunit-specific regulation of the I kappa B α promoter. *J Biol Chem.* 1994;269:13551–13557.
- Bandman O, Coleman RT, Loring JF, Seilhamer JJ, Cocks BG. Complexity of inflammatory responses in endothelial cells and vascular smooth muscle cells determined by microarray analysis. *Ann N Y Acad Sci.* 2002;975:77–90.
- Zhao B, Stavchansky SA, Bowden RA, Bowman PD. Effect of interleukin-1 β and tumor necrosis factor- α on gene expression in human endothelial cells. *Am J Physiol Cell Physiol.* 2003;284:C1577–C1583.
- Volpert OV, Pili R, Sikder HA, Nelius T, Zaichuk T, Morris C, Shiflett CB, Devlin MK, Conant K, Alani RM. Id1 regulates angiogenesis through transcriptional repression of thrombospondin-1. *Cancer Cell.* 2002;2:473–483.
- de Martin R, Vanhove B, Cheng Q, Hofer E, Csizmadia V, Winkler H, Bach FH. Cytokine-inducible expression in endothelial cells of an I kappa B α -like gene is regulated by NF kappa B. *EMBO J.* 1993;12:2773–2779.
- Blackshear PJ. Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of mRNA turnover. *Biochem Soc Trans.* 2002;30:945–952.
- Wilkin F, Suarez-Huerta N, Robaye B, Peetermans J, Libert F, Dumont JE, Maenhaut C. Characterization of a phosphoprotein whose mRNA is regulated by the mitogenic pathways in dog thyroid cells. *Eur J Biochem.* 1997;248:660–668.
- Wu M, Xu LG, Zhai Z, Shu HB. SINK is a p65-interacting negative regulator of NF-kappaB-dependent transcription. *J Biol Chem.* 2003;278:27072–27079.