5318–5330 *Nucleic Acids Research, 2015, Vol. 43, No. 11* *Published online 01 May 2015*
doi: 10.1093/nar/gkv423

Signal integration by the *CYP1A1* promoter — a quantitative study

Pascal Schulthess[1,2], Alexandra Löffler[3], Silvia Vetter[3], Luisa Kreft[3], Michael Schwarz[3], Albert Braeuning[4,*] and Nils Blüthgen[1,2,*]

[1]Institute for Pathology, Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany, [2]Integrative Research Institute for the Life Sciences and Institute for Theoretical Biology, Humboldt University of Berlin, Philippstr. 13, 10115 Berlin, Germany, [3]Institute for Experimental and Clinical Pharmacology and Toxicology, Department of Toxicology, University of Tübingen, Wilhelmstraße 56, 72074 Tübingen, Germany and [4]Department of Food Safety, Federal Institute for Risk Assessment, Max-Dohrn-Straße 8–10, 10589 Berlin, Germany

Received October 12, 2014; Revised March 11, 2015; Accepted April 17, 2015

ABSTRACT

Genes involved in detoxification of foreign compounds exhibit complex spatiotemporal expression patterns in liver. Cytochrome P450 1A1 (*CYP1A1*), for example, is restricted to the pericentral region of liver lobules in response to the interplay between aryl hydrocarbon receptor (AhR) and Wnt/β-catenin signaling pathways. However, the mechanisms by which the two pathways orchestrate gene expression are still poorly understood. With the help of 29 mutant constructs of the human *CYP1A1* promoter and a mathematical model that combines Wnt/β-catenin and AhR signaling with the statistical mechanics of the promoter, we systematically quantified the regulatory influence of different transcription factor binding sites on gene induction within the promoter. The model unveils how different binding sites cooperate and how they establish the promoter logic; it quantitatively predicts two-dimensional stimulus-response curves. Furthermore, it shows that crosstalk between Wnt/β-catenin and AhR signaling is crucial to understand the complex zonated expression patterns found in liver lobules. This study exemplifies how statistical mechanical modeling together with combinatorial reporter assays has the capacity to disentangle the promoter logic that establishes physiological gene expression patterns.

INTRODUCTION

The liver is the main organ carrying out the metabolism of many drugs and toxins. One important protein involved in phase I of xenobiotic and drug metabolism is cytochrome P450 1A1 (CYP1A1). It is regulated by the aryl hydrocarbon receptor (AhR) which, upon stimulation with the environmental toxin 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), translocates into the nucleus where it binds to its heterodimerization partner Arnt (AhR nuclear translocator) (1,2). The so-formed dimer then binds to dioxin responsive elements (DREs) in the promoter region of target genes (3–11). The AhR pathway induces strong expression of CYP1A and other xenobiotic-metabolizing enzymes upon stimulation by ligands/toxins such as TCDD. This leads to altered toxicokinetics of endogenous metabolites and foreign compounds whose metabolism is dependent on AhR-regulated genes. Other effects of AhR activation include skin toxicity (especially chloracne), tumor promotion and disturbance of developmental processes, at least in laboratory animals (12).

Expression of many genes in the liver is restricted to certain regions of a liver lobule, such as the periportal and the pericentral region (13–21). Many proteins/mRNAs from various members of the CYP family show an expression gradient along the portocentral axis, including the members of the AhR-regulated CYP1A subfamily (22).

The Wnt/β-catenin signaling pathway, which shows high activity in the perivenous region, has been identified as one of the major determinants of hepatic zonation (23–25). In the absence of activating Wnt ligands, the free cytosolic fraction of the key protein in the pathway, β-catenin, is constantly phosphorylated by a large protein complex and thus marked for degradation by the proteasome. Binding of Wnt ligands to Frizzled receptors destabilizes the protein complex which in turn allows β-catenin to escape degradation, translocate into the nucleus and initiate target gene transcription via an interaction with T-cell factor (TCF) transcription factors.

In mice it has been observed that a deregulation of the Wnt/β-catenin signaling pathway affects the expression of *Cyp1a* genes. While mice with a hepatocyte-specific knock-

*To whom correspondence should be addressed. Tel: +493020938924; Fax: +493020938801; Email: nils.bluethgen@charite.de
Correspondence may also be addressed to Albert Braeuning. Tel: +4930184123758; Fax: +4930184123685; Email: albert.braeuning@bfr.bund.de

out of *Ctnnb1,* the gene that encodes for β-catenin, lose the basal expression of *Cyp1a* (26), mouse hepatomas and transgenic hepatocytes with activated β-catenin show an increased *Cyp1a* expression even in the absence of TCDD (27,28). β-Catenin and AhR activation synergize in the induction of *CYP1A1* expression (22,26,29). Five functional transcription factor binding sites related to AhR- or β-catenin-dependent signaling have been identified within the human *CYP1A1* promoter: four AhR-binding DREs (3) and one site that is bound by β-catenin/TCF (22). Thus, both pathways can directly regulate the promoter. Furthermore, a physical interaction between β-catenin and the AhR has been shown (22,30). Still, the molecular details of the interplay between the transcription factors are still unknown.

Previously, there have been various successful attempts to disentangle the interactions of transcription factors with the help of mathematical models, mainly in simpler prokaryotic systems or invertebrates such as *Drosophila melanogaster*. Thermodynamic models that are based on statistical mechanics were first introduced in the early '80s by Ackers *et al.* (31) and Shea and Ackers (32) and further elaborated on by many others (33–49). For example, with the help of thermodynamic modeling, Giorgetti *et al.* (50) showed that NF-κB uses clustered binding sites non-cooperatively. In *D. melanogaster*, thermodynamic models were used to explain the morphogenic gradients (51,52) and the segmentation of the embryo (53).

How does the *CYP1A1* promoter integrate the two signals originating from the AhR and Wnt/β-catenin signaling pathways? How does this integration collude with a possible cooperative interaction between various TFs binding to the promoter to regulate a multitude of expression patterns that give rise to hepatic zonation? To answer these questions we introduced a library of mutant promoter constructs into a hepatoma cell and observed their activity in response to modulation of the two signaling pathways. With the help of thermodynamic models we untangled the cooperative interaction between the TFs and predicted hepatic zonation patterns.

MATERIALS AND METHODS

Plasmids

Generation of a pT81luc-based Firefly luciferase reporter plasmid containing a ∼1200 bp fragment of the human *CYP1A1* promoter has been described previously (54). This plasmid contains four functional AhR-binding DREs, termed C, D, E and F, (3) and one β-catenin/TCF binding site, termed T (22). Different mutant versions of the reporter were generated by site-directed mutagenesis of individual transcription factor binding sites using the QuikChange kit (Stratagene, La Jolla, CA, USA) as recently described (22). An overview of the mutations introduced is given in Supplementary Figure S1. Synthetic promoter constructs were generated containing multiple copies of either C-DRE (sequence 5′-CGCTTCTCACGCGAGCCGG-3′) or D-DRE (sequence 5′-GCCGGCGCACGCAAGCTAG-3′) by cloning synthetic oligonucleotides into the SmaI site of pT81luc. Versions of the 2× C-DRE plasmid with different distances between the two DREs were generated by inserting non-AhR-responsive sectors of different size from the 1.2 kb *CYP1A1* promoter fragment between the two C-DREs of the NaeI-cut pT81-luc/2× C-DRE plasmid, resulting in 2× C-DRE variants with 49, 156 and 292 bp distance between the two AhR binding sites, respectively. The integrity of all plasmids was verified by dideoxy sequencing. In some experiments, a pCMV4-based expression vector for human AhR was used (27); control cells were transfected with empty pCMV4. Activity of the β-catenin pathway was monitored with the SuperTOPFlash vector, which expresses Firefly luciferase under the control of eight β-catenin/TCF binding sites (55). Plasmid pRL-CMV encoding Renilla luciferase under the control of the constitutive cytomegaly virus promoter (Promega, Mannheim, Germany) was co-transfected and used for normalization of Firefly luciferase signals.

Cell culture and transfection

Mouse hepatoma cells from lines 55.1c (22), Hepa1c1c7 and the AhR-deficient sub-clone Hepa12 (56) were cultured in D-MEM/F-12 medium supplemented with 10 % fetal bovine serum and antibiotics (all reagents purchased from Invitrogen, Karlsruhe, Germany) at 37 °C and 5 % CO_2 in a humidified atmosphere. 55.1c cells carry a heterozygous deletion in exon 3 of *Ctnnb1*, encoding a constitutively active version of β-catenin (Figure 1D). Cells were seeded on 24-well plates at a density of 40 000 cells/cm^2 24 h prior to transfection with Lipofectamine 2000 (Invitrogen). Cells were treated with the indicated concentrations of TCDD (Ökometric, Bayreuth, Germany) and/or the β-catenin inhibitors iCRT3 (Merck, Darmstadt, Germany), FH535 (Merck), or PNU74654 (Sigma, Taufkirchen, Germany) for 24 h, starting 24 h after transfection. All compounds were dissolved in dimethylsulfoxide and final concentration of the solvent in culture medium was 0.2 %. Cells were lysed with 1× Passive Lysis Buffer (Promega) for 15 min at room temperature. Dual-luciferase assays for Firefly and Renilla luciferase activities were conducted as recently described (57). Primary hepatocytes were obtained from young male adult mice with hepatocyte-specific knockout of *Ctnnb1* (encoding β-catenin) by standard collagenase perfusion (56) and seeded at a density of 50 000 cells/cm^2 on 6-well plates coated with rat tail collagen in D-MEM/F-12 medium supplemented with 10 % fetal bovine serum and antibiotics. Medium was changed to 1 % serum after 6 h and cells were treated as described above.

Cell viability and growth analysis

All compounds were tested for the absence of cytotoxicity by the neutral red uptake and Alamar blue reduction assays as previously described (58). Cell growth was monitored by the use of the sulforhodamine B assay according to Skehan *et al.* (59). All assays were conducted in octuple determinations on 96-well plates, where cells were seeded at 5 000 cells/cm^2 (sulforhodamine B assay) or 9 000 cells/cm^2 (cytotoxicity assays).

Immunoprecipitation and western blotting

Whole cell lysates were prepared according to Braeuning *et al.* (22). Immunoprecipitation was performed at 4°C

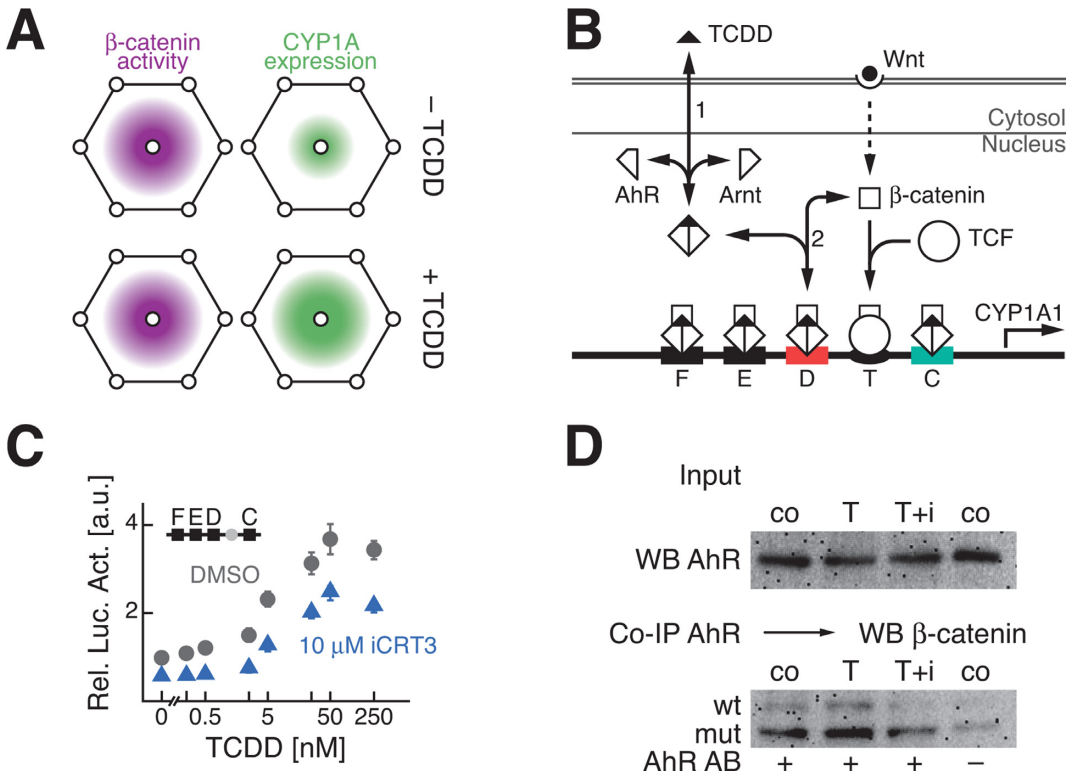

Figure 1. Cross-talk between AhR and Wnt/β-catenin signaling pathways in zonal expression of *CYP1A*. (**A**) Hepatic zonation schematic of β-catenin activity and *CYP1A* expression in response to TCDD treatment. Area of cells displaying high β-catenin activity is thought to remain invariant to presence of TCDD while a larger amount of cells express CYP1A after TCDD exposure. (**B**) Schematic representation of the two signaling pathways that converge on the *CYP1A1* promoter. TCDD can bind to the AhR and Arnt that is then assumed to recruit free β-catenin. β-catenin is activated (dashed arrow) through the binding of Wnt to its surface receptors and forms a TF complex with TCF. The *CYP1A1* promoter possesses five functional TFBS within 1.2 kbp, four for TCDD/AhR/Arnt (rectangular binding sites C, D, E, F) and one for β-catenin/TCF (elliptical binding site T). The double-headed and numbered arrows depict the two reactions used in the signaling model. The C- and D-DRE are color coded as in subsequent figures. (**C**) β-Catenin modulates *CYP1A1* expression independently of the β-catenin/TCF binding site. Relative luciferase activity of a promoter construct in which the β-catenin/TCF binding site was inactivated by point-mutation over a series of TCDD concentrations treated with DMSO (gray circles) or 10 μM of the β-catenin inhibitor iCRT3 (blue triangles). (**D**) The interaction between AhR and β-catenin is increased upon TCDD treatment. Top: Western analysis of AhR expression in lysates from 55.1c cells. Co: solvent control; T: 250 nM TCDD for 1 h; T+i: 250 nM TCDD + 10 μM iCRT3. Bottom: western analysis of β-catenin after immunoprecipitation with an anti-AhR antibody (AhR AB '+'). The two bands show the wild-type (wt) and exon 3-deleted (mut) versions of β-catenin, i.e. the constitutively active β-catenin present in the cells.

over night using ProteinG-agarose beads and an antibody against the AhR (Biomol, Hamburg Germany; 1:200 dilution). Whole cell lysates (50 μg/lane; protein concentration determined by use of the Bradford assay) or immunoprecipitates were separated by SDS-PAGE and transferred to PVDF membranes. Proteins were visualized by using antibodies against the AhR (Biomol; 1:1 000), β-catenin (BD biosciences, Heidelberg, Germany; 1:500), or GAPDH (Merck; 1:1 000) in combination with appropriate alkaline phosphatase-conjugated secondary antibodies (Millipore, Schwalbach, Germany; 1:10 000) and the substrate CDP-Star (Tropix, Darmstadt, Germany). Chemoluminescence was monitored with a CSC camera (Raytest, Straubenhardt, Germany).

Gene expression analysis

Total RNA was isolated by TRIzol (Invitrogen) and reverse transcribed using avian myeloblastosis virus reverse transcriptase (Promega) as described previously (57). Real-time RT-PCRs were performed on a LightCycler system (Roche, Mannheim, Germany) using the FastStart DNA Master^{PLUS} SYBR Green I kit (Roche) according to the manufacturer's instructions and the following primer pairs: 18s rRNA_fwd 5′-CGGCTACCACATCCAAGGAA-3′; 18s rRNA_rev 5′-GCTGGAATTACCGCGGCT-3′; Cyp1a1_fwd 5′-TGTCCTCCGTTACCTGCCTA-3′; Cyp1a1_rev 5′-GTGTCAAACCCAGCTCCAAA-3′; Cyp1a2_fwd 5′-GAGCGCTGTATCTACATAAACCA-3′; Cyp1a2_rev 5′-GGGTGAACATGATAGACACTATTGT-3′. Data were normalized according to the method described by Pfaffl (60) with 18s rRNA as a housekeeping gene.

Animal experiment and immunostaining

Young adult male C3H/HeN mice (n = 5–6 per group) were treated with the AhR inducer 3-methylcholanthrene (Sigma; dissolved in corn oil) by a single i.p. injection of 10, 25 or 50 mg/kg body weight 48 h prior to sacrifice. Mice were killed between 9 and 11 a.m. to avoid circadian variations; livers were excised, transferred to Carnoy's fixative and subsequently embedded in paraffin. Tissue

slices of 5 μm thickness were stained for CYP1A as recently described (61) using a rabbit antiserum at 1:500 dilution (gift of Dr R. Wolf, University of Dundee, UK) and a horseradish peroxidase-conjugated secondary antibody (1:100; Dako, Glostrup, Denmark) with the substrates 3-amino-9-ethylcarbazole/H_2O_2. Mice had access to tap water and standard chow ad libitum. All animals received humane care and protocols complied with institutional guidelines. Width of CYP1A-positive zones was assessed using an AxioImager light microscope and AxioVision software (Zeiss, Oberkochen, Germany).

Data processing

The raw data from the relative luciferase measurements was preprocessed as described in the following to account for measurement inaccuracies. For each promoter configuration the means and standard deviations were calculated from at least three replicates and then normalized to data without TCDD and β-catenin inhibitor. To avoid underestimates of standard deviations, we used 10 % of the mean as the lowest value for the standard deviations, and used a minimum absolute standard deviation of 0.5 for model calibration. A linear correction error model for each promoter construct was used to express standard deviation as a linear function of the mean. The resulting standard deviations were then converted into standard deviations of the mean by dividing by the square root of the number of replicates for each promoter construct. For details, see supporting text and refer to the supporting Spreadsheet S13 for the data sets.

Model implementation

To describe the expression of *CYP1A1* we developed a two-tier mathematical model. Signaling is represented by mass-action kinetics, comprising two reversible reactions (TCDD + AhR + Arnt ↔ TCDD/AhR/Arnt, TCDD/AhR/Arnt + β-catenin ↔ TF, cf. Figure 1B), which we assume to be in thermodynamic equilibrium. Additionally, we assumed that the total concentrations of AhR plus Arnt, TCDD and β-catenin in the system are constant. Since the total concentration of β-catenin was not measured, we multiplied the β-catenin concentration with a parameter describing the activity of β-catenin that can be modulated by an inhibitor, leaving the concentration of β-catenin free for fitting. The statistical mechanics model describes the binding of TFs to binding sites on the promoter. Each TFBS as well as a separate binding site for the RNA polymerase (RNAP) is allowed to be either bound or unbound. For every mutant promoter construct this leads to a collection of possible binding state arrangements. Only those arrangements in which the RNAP is bound to the DNA are assumed to lead to transcription. Interaction of two simultaneously bound TFs is characterized by a binding free energy; all bound TFs interact with the bound RNAP through another binding free energy and with the DNA through an association constant. The formulation of the model follows the formalism developed by Sherman and Cohen (45) that leads to one expression for the binding probability of the RNAP for every promoter construct. For comparison with experimental data each binding probability is normalized to the basal promoter expression, i.e. when only the RNAP is bound. The resulting fold-change in binding probability is assumed to match the observed *CYP1A1* expression. For a detailed deduction of the model, see supporting text.

Model fit

The combination of the equations of the signaling model and the construct specific binding probabilities served as objective functions for a weighted non-linear least square fit (trust-region-reflective algorithm in lsqnonlin, MATLAB R2013a). We utilized Latin hypercube sampling (62) to generate a set of 10 000 starting values for optimization. The ranges of the parameters were limited to be biologically feasible (e.g. association constants greater zero) and it was made sure that the resulting optima did not exceed any artificial boundaries such as an exceeding number of iterations. The identifiability and confidences of the parameter estimates was then assessed by exploiting the profile likelihood (63). A detailed description of the algorithm can be found in the supporting text.

RESULTS AND DISCUSSION

Successful detoxification in the liver requires the induction of genes involved in the metabolism of exogenous substances in a spatially confined area. In this work we use the promoter of the *CYP1A1* gene to study exemplarily how a promoter integrates two stimuli, and how multiple *cis*-regulatory elements shape the response and thus the physiological expression pattern of this mammalian promoter.

Interaction between AhR and Wnt/β-catenin signaling on *CYP1A* expression

Two pathways control the *CYP1A1/2* promoters. The AhR signaling pathway mediates the response to toxins, while the Wnt/β-catenin signaling pathway, that shows a portocentral gradient, restricts *CYP1A* expression to the perivenous region. Without TCDD, *CYP1A* is only expressed in a small perivenous area. After AhR activation by TCDD, the region of expression widens significantly (see schematically in Figure 1A), although the pattern and strength of β-catenin activity remains invariant (22,29).

Stimulation with TCDD leads to the formation of a transcription factor (TF) complex containing TCDD, AhR and Arnt that can bind to the dioxin responsive elements (DREs) within the *CYP1A1* promoter region. Wnt ligands trigger the accumulation of β-catenin, which forms TFs by binding to TCF. Four known functional transcription factor binding sites (TFBS) in a ∼1 200 bp upstream region of the human *CYP1A1* gene are recognized by the TCDD/AhR/Arnt complex (in the following termed C-DRE to F-DRE) and one by β-catenin/TCF (Figure 1B). To test if the β-catenin/TCF binding site is necessary for the Wnt-mediated expression, we cloned a mutant reporter construct that contains the *CYP1A1* promoter with an inactivated β-catenin/TCF TFBS (Supplementary Figure S1). We transfected this reporter into mouse hepatoma cells from line 55.1c, then exposed them to increasing concentrations of TCDD for 24 h and measured promoter activity

(Figure 1C). A 4-fold induction was observed. When we additionally added a β-catenin inhibitor (iCRT3), both basal reporter activity and induction were drastically reduced, although the inhibitor does not effect the expression of the Ah receptor (Supplementary Figure S2). This hints to a β-catenin/TCF binding site-independent mechanism of interaction between the two pathways.

How do the two pathways now interact to control the spatiotemporal expression of *CYP1A1*? First, a direct interaction between the proteins has been reported (22,30). Second, AhR might be a transcriptional target of Wnt/β-catenin signaling. We performed co-immunoprecipitation that indeed confirmed a physical interaction between AhR and β-catenin in our cell system (Figure 1D). In particular, the interaction was strongest when TCDD was applied, and was reduced when β-catenin activity was inhibited. As overexpression of AhR did not alter AhR-dependent reporter activity (Supplementary Figure S3), we concluded that the amount of AhR was not rate limiting for the regulation of *CYP1A1*, and thus modulation of AhR mRNA expression by β-catenin is not relevant for *CYP1A1* regulation under our experimental conditions.

In summary, Wnt/β-catenin and AhR signaling are linked in two ways to regulate *CYP1A1*: firstly, pathway specific transcription factors act on the same promoter and, secondly, some signaling components interact physically. An interesting aspect is that cross-talk exists without a functional β-catenin/TCF binding site. Thus, interactions on the *CYP1A1* promoter are not necessary to establish cross-talk, and it remains unclear which level of cross-talk would be the dominating mechanism by which these two pathways interact in regulating *CYP1A1* expression. To dissect this complex regulation of *CYP1A1* in more detail, the quantitative behavior of the response to AhR signaling as well as the quantitative interaction of both pathways will be studied in a first step.

Cooperativity between DRE TFBS

AhR signaling converges on four DREs in the promoter of *CYP1A1*. Multiple binding sites for TFs can give rise to cooperative behavior and establish complicated promoter logic (38,50,64,65). We therefore decided to first investigate the interaction between these sites by constructing promoter mutants in which we systematically point-mutated the TFBS sequences to disable their function (Supplementary Figure S1), and then measured their promoter activity at increasing concentrations of TCDD (Figure 2A). The first construct in Figure 2A has only one functional, non-mutated AhR binding site (D-DRE) and shows hardly any induction by TCDD. In contrast, a construct that has only the C-DRE binding site available for AhR binding shows a robust, 2-fold increase in luciferase activity at higher amounts of TCDD. Finally, for a construct where both the C-DRE and the D-DRE are present, we observed an almost 4-fold increase in luciferase activity.

Thus, while the D-DRE alone does not change gene induction by TCDD, it cooperates with the C-DRE in inducing expression, demonstrating a more-than-additive cooperativity between the interacting TFs that bind to the DREs.

Synthetic promoter constructs confirm cooperativity between dioxin responsive elements

The cooperative interaction between the TFs that bind to the C- and D-DRE motivated us to investigate the interaction in more detail, in particular how the different binding site sequences and distances between the binding sites influence the cooperative behavior. Although both C- and the D-DRE share an equal core sequence (5'-CACGC-3'), they differ strongly in the flanking regions (Supplementary Figure S1), thus they likely differ in binding affinity. To systematically analyze how these DREs and their interaction change the promoter response in detail, we constructed synthetic promoter mutants with which we generated quantitative data, and then used a mathematical model to analyze these. We generated six constructs with one to six copies of the C-DRE TFBS sequence, three constructs with one to three copies of the D-DRE TFBS sequence. It should be noted that the distance of 19 bp between two binding sites in the C-DRE and the D-DRE constructs is smaller than the mean TFBS distance of 260 bp in the natural promoter. To analyze the effect of distance, we also cloned four constructs with two C-DREs placed at increasing distances from each other (19, 49, 156 and 292 bp, see Figure 2B). The inserted sequences between the two C-DREs were not AhR-responsive. These 13 constructs were then transfected into 55.1c cells, which subsequently were treated with increasing concentrations of TCDD for 24 h. The measured luciferase activity data was then preprocessed as described in the Materials and Methods section and the supporting text. For all synthetic constructs, in which only one DRE was present, no significant induction of reporter gene expression for different TCDD concentrations was observed. In contrast, all remaining constructs showed a dose-dependent induction (dots in Figure 2D and F). For the C-DRE constructs, we observed maximal induction for three present DREs. Thus, adding more binding sites to the constructs does not add to the inducibility of the promoter but conversely seems to decrement its induction. If we furthermore compare the constructs containing different binding sites, i.e. C-DREs versus D-DREs, we observed that the constructs with C-DREs showed higher induction than the corresponding D-DRE constructs. For example, the maximal fold-changes for the double C-DRE construct was 10, whereas it was at most two for the double D-DRE construct.

To quantify the effect of the Wnt/β-catenin signaling pathway on gene expression, we established a system to modulate the Wnt/β-catenin pathway. Again, we used 55.1c cells, which express a mutant, constitutively active form of β-catenin (22). We then used the β-catenin inhibitor iCRT3 for which we confirmed that it does not act as AhR agonist (Supplementary Figure S2), shows no cytotoxicity (Supplementary Figure S4), and can efficiently block β-catenin signaling output (Supplementary Figure S5). When we applied the inhibitor in experiments with the *CYP1A1* promoter reporter, we observed that iCRT3 can indeed inhibit the induction of the reporter by TCDD (Supplementary Figures S3 and S6). To calibrate the concentrations of the iCRT3 inhibitor, we used the 8x β-catenin/TCF-driven SuperTOPFlash reporter and measured luciferase activity as a proxy for β-catenin pathway activity. To ensure

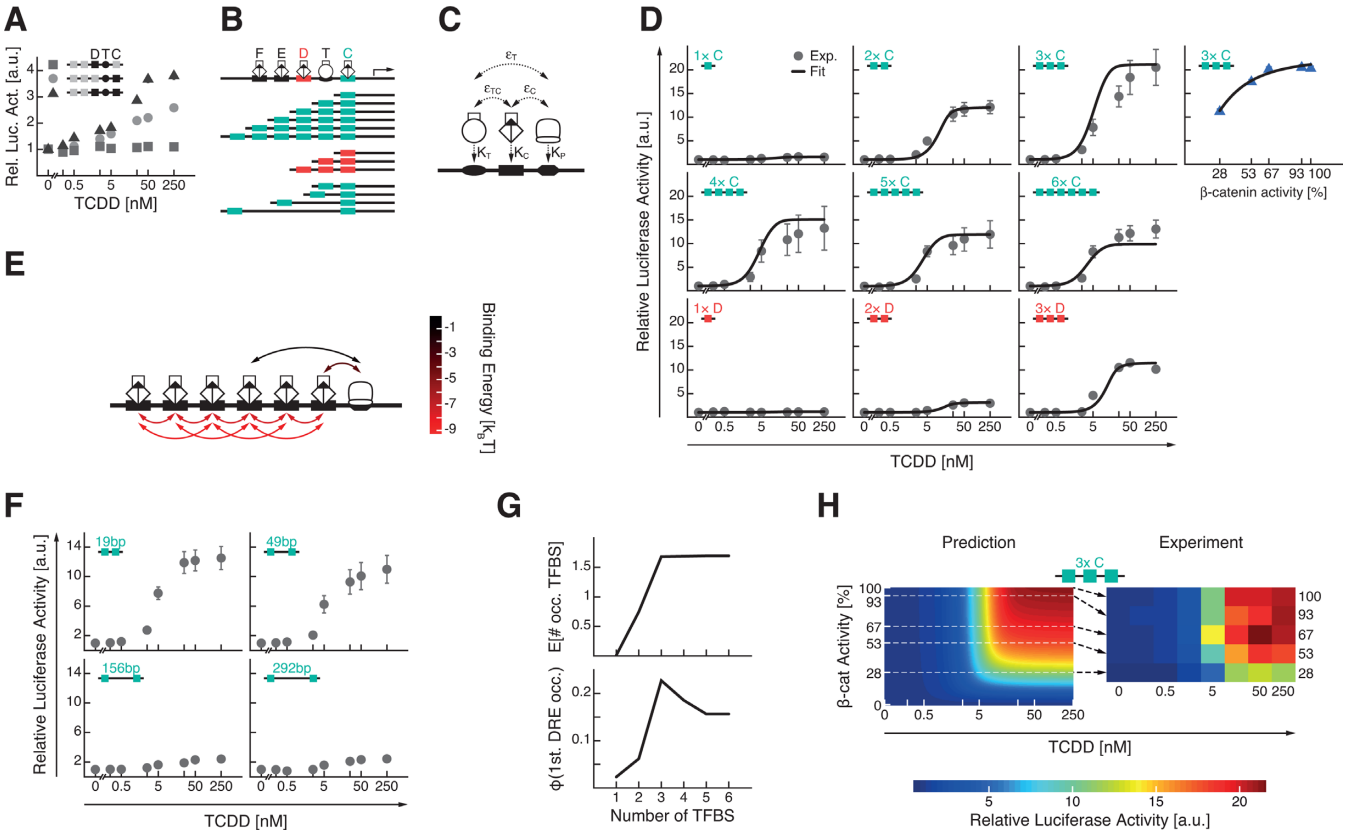

Figure 2. A mathematical model recapitulates cross-talk and cooperativity between binding sites in the control of a synthetic promoter constructed from elements of the *CYP1A1* promoter. (**A**) Cooperativity between TFs binding to DREs on the natural promoter. Relative luciferase activity of three natural promoter constructs over a series of TCDD concentrations was measured. On the three promoter constructs only the following TFBS combinations are present as WT sequence while the others are inactivated through point-mutation: square: β-catenin/TCF and D-DRE; circle: β-catenin/TCF and C-DRE; triangle: β-catenin/TCF, C- and D-DRE. (**B**) Schematic depiction of the synthetic promoter construct library. Six constructs hold one to six copies of the sequence of the C-DRE (turquois), three constructs hold one to three copies of the sequence of the D-DRE (red), and differently long non-TF responsive sequences (19, 49, 156, 292 bp) were inserted between two C-DREs for four constructs. (**C**) Scheme explaining the parameters used in the thermodynamic model. K_i represent the association constants of the TFs to the C and T TFBS as well as between the RNAP and the DNA. ε_C and ε_T are the binding energies between the TFs and the RNAP while ε_{TC} is the interaction term between the TFs. (**D**) TCDD concentration series of the synthetic promoter constructs can be explained with a thermodynamic model. TCDD concentration series of nine synthetic constructs as well as one β-catenin titration series where the relative luciferase activity was measured (gray dots and blue triangles). The upper seven graphs represent the C-DRE constructs while the lower three graphs represent D-DRE constructs. The graph in the top right corner shows a β-catenin titration series measured at 250 nM TCDD concentration. Error bars represent one standard deviation of 6–10 biological replicates. The black curves show the fits of the thermodynamic model. (**E**) Cooperative binding dominates the synthetic promoter constructs. Binding energies resulting from fits are displayed on the promoter construct holding six copies of the C-DRE. Arrows depict significant binding events. Their colors represent binding strength where lower values represent stronger association. All binding events were present in the model. Those equal or close to 0 are not depicted as arrows. (**F**) Stimulus-response curves for constructs with different distances between DREs confirm reduced cooperativity for longer distances. Relative luciferase activity was measured over increasing TCDD concentrations for four promoter constructs with different distances (19, 49, 156, 292 bp) between their C-DREs. (**G**) Model simulations showing why three C-DRE binding sites show maximal induction. Top: The average number of occupied binding sites is plotted over the number of present binding sites in the synthetic constructs. Bottom: The probability that the first DRE, i.e. the one closest to the RNAP is occupied by TFs is shown. Note that these values were calculated for 250 nM of TCDD. (**H**) Thermodynamic model correctly predicts and experimental data and confirms an AND gate relationship between the Wnt/β-catenin and the AhR signaling pathway on the 3× C-DRE construct. Left: prediction of the promoter activity of the synthetic construct holding three C-DREs by the thermodynamic model through variation of the β-catenin activity parameter. Right: corresponding measurements of the relative luciferase activity of the 3× C-DRE construct where the inhibitor iCRT3 modulated the β-catenin activity. The cells were stimulated with increasing TCDD concentrations.

the specificity of the observations, the experiments were repeated using a second, structurally unrelated β-catenin inhibitor, FH535 (Supplementary Figures S2–S6). The results obtained with FH535 closely resemble the findings with iCRT3. We used the iCRT3 inhibitor to generate quantitative data on the construct with three C-DREs, which shows reduction in relative luciferase activity when the β-catenin pathway activity is reduced (Figure 2D, top right).

A thermodynamic model can explain binding site cooperativity. To understand in more detail, how the different binding sites interact, we decided to describe this behavior with the help of a mathematical model. The model has two tiers. The first tier of the model describes the signaling events leading to the formation of the TFs (cf. Figure 1B). This includes the reversible binding of TCDD to AhR and its nuclear translocator Arnt. Furthermore, the association between β-catenin and the TCDD/AhR/Arnt complex is modeled. The second tier of the model is the binding of

the TFs to the binding sites in the promoter, which we decided to model by a thermodynamic approach that describes binding probabilities as function of binding energies (cf. Figure 2C for parameters types used in the model). Our thermodynamic model assumes that the described system operates in thermodynamic equilibrium, similarly to previous work (31,33,35,40,45,46,49,53,66). Since we determine expression on long time-scales, it is likely that the system is at least in steady state, and thus could likely be describe with such an approach. This model includes a binding site for the RNA polymerase (RNAP), and we use the probability of bound polymerase as a proxy for gene expression. A detailed description and deduction of the model expression can be found in the Materials and Methods and the supporting text. The model assumes that AhR concentration is not the limiting step in the pathway, as we found that AhR Co-transfection with different reporter constructs (WT, 3× C-DRE) does not change the measured luciferase activity, even when the cells are stimulated with TCDD (Supplementary Figure S3).

To model the behavior of the synthetic promoter constructs, we fitted the mathematical models for one to six C-DREs and one to three D-DREs at the same time to the data consisting of the TCCD titration series (gray circles in Figure 2D) and to one β-catenin titration series (blue triangles in Figure 2D). Between all these models, those parameters that describe the same biological process were kept equal, limiting the total number of free parameters. For example, the binding energy between the TF that binds to the first DRE and the polymerase was the same in all models because the same TF binds to all the TFBS. Additionally, interaction over the same distances was assumed to be carried out by the same binding energy, i.e. there was one parameter for neighboring TFs, one for binding over one intermediate, etc. Furthermore, the association constants of all constructs using C-DREs were the same, and those constructs using D-DREs had a different parameter. Overall, all models had 17 parameters that were then simultaneously fitted to the corresponding data sets. Using this strategy, we could fit the data set in most quantitative detail (compare solid lines which represent the best model fit with dots that represent the experimental data in Figure 2D). Even though the data of the C-DRE and the D-DRE constructs differed significantly, they could be explained by only one differing parameter that models the binding affinity of the transcription factor to the DRE sequences. In particular, binding affinity of the TFs to the C-DREs is 4× stronger than that to the D-DREs (cf. Supplementary Table S1). The model fit also reproduces the resulting strong sigmoidal response. All constructs show an almost switch-like behavior if the TCDD concentration is increased.

Binding site cooperativity limited to neighboring TFBS. The parameters produced by the model fitting routine now helped us to understand how the TFs interact to modulate various induction behaviors of the synthetic constructs (Figure 2E). To obtain confidence intervals for the parameters, we used the profile likelihood approach (63). Applying a point-wise confidence threshold, i.e. confidence intervals that hold individually for each parameter, we find that all parameters had a clear minimum and thus are identifiable. Ten of the parameters were significantly larger than zero, most with relatively small confidence intervals, while the remaining seven parameters (ε_2, ε_4, ε_{d3}, ε_5, ε_{d4}, ε_6, ε_{d5}) were indistinguishable from zero, i.e. the model can fit the data without these interactions (Supplementary Figure S7 and Table S1). Strongest binding between TFs occurs over short distances whereas binding to the polymerase is only significant for the TFs binding to the first TFBS. Longer distance binding between TFBS and RNAP is either zero or close to zero, as is longer distance cooperative binding between TFs (Supplementary Figure S8). These modeling results could also be confirmed by the promoter constructs with varying distances between the TFBSs (Figure 2F). Here, the maximal fold-change of ∼12-fold can be observed for the smallest distance of 19 bp between the TFBS. With increasing distance the maximal induction decreases to ∼10-fold for a distance of 49 bp and to ∼2-fold for the distances 156 and 292 bp. It is apparent that the inducibility decreases with increasing distance between the TFBS, suggesting that, as predicted by the model parameters, long distance cooperation is limited. The importance of distance between TFBS was also investigated using similar approaches for the Hepcidin promoter, where cooperativity between nearby binding sites for BRE1 and STATBS was essential to reproduce the stimulus-response curves with a mathematical model (64).

Sequestration is responsible for reduced transcriptional induction. A counterintuitive aspect of the data is that adding more DREs to the promoter does not necessarily lead to increased transcriptional output. Adding more than three DREs rather decreases activity (Figure 2D). Interestingly, the mathematical model can accurately describe that behavior. This can be explained as follows: first, only the most proximal DRE significantly interacts with the RNAP, therefore transcription is only initiated if the TCDD/AhR/Arnt/β-catenin TF complex is bound to the first DRE. Second, the DREs only interact over short distances, i.e. with the neighboring DRE and the DRE next to its neighbor. In the case when there are two DREs present, the second DRE can interact with the first and thus increase transcriptional output. Similarly, if three DREs are present both the second and the third DRE can interact with the first DRE and further increase the transcriptional output. However, if the construct holds four DREs, the forth DRE cannot directly interact with the first DRE because the model parameters were zero for long range interactions. Since only pairwise interactions can occur, there are states where the second, third and forth DRE interact with each other but none of them with the first DRE. This reduces transcriptional output and can thus be viewed such that the fourth, fifth and sixth DRE sequester interaction partners for the first DRE. This hypothesis is also supported by the calculation of the average number of occupied binding sites as well as the probability that the first TFBS is bound by a TF (Figure 2G). While no TFs are bound on average if only one C-DRE is present (explaining the lacking induction of the 1× C-DRE construct in Figure 2D), one TFBS is occupied in the 2× C-DRE construct. The maximum number of occupied TFBS is reached in the 3× C-DRE. Adding more DREs thus does not in-

crease the mean number of bound binding sites (Figure 2G, top). Likewise, looking at the probability that the first and thus the strongest cooperator with the RNAP is bound to its binding site we see a peak for the 3× C-DRE construct. This means that the same number of TFs may distribute over four and more TFBS resulting in states where the first TFBS is unoccupied (Figure 2G, bottom) and hence induction is reduced.

In summary, while the number of occupied binding sites remains constant for three and more DREs, the present TFs have more possibilities to bind to the promoter leading to a decreased probability that the first DRE is occupied. And since the TF bound to the first DRE interacts most strongly with the RNAP this leads to decreased induction of the 4× to 6× C-DRE constructs.

The thermodynamic model can predict signal integration by the synthetic promoter. As we have now established a modeling framework that can describe the response of the synthetic constructs to TCDD stimulation well, we next decided to include the impact of the Wnt/β-catenin signaling pathway on the promoter. As the model was trained in cells that have a constitutively active Wnt/β-catenin pathway, we simulated inhibition of β-catenin signaling in our model by reducing the parameter representing β-catenin activity. When plotting the predicted response of the model describing the 3× C-DRE promoter construct as a function of TCDD and β-catenin activity, it becomes apparent that the promoter requires both Wnt/β-catenin and AhR signaling pathways to be active and responds in a sharp ultrasensitive way, resembling the behavior of a logic AND gate (Figure 2H, left).

We then generated a data set by systematically varying the β-catenin activity and the TCDD concentration, and afterwards measured reporter activity for the synthetic construct with triple C-DREs (Figure 2H right). As predicted by the model, a decrease in β-catenin activity through inhibition with iCRT3 results in lower luciferase activity. Similarly, a smaller TCDD stimulus leads to a decrease in *CYP1A1* expression. Even though a full inhibition of β-catenin activity and thus a complete shut-down of *CYP1A1* induction is experimentally not accessible, this data is in quantitative agreement with the model predictions, confirming the AND gate behavior between the Wnt/β-catenin and AhR pathway as predicted by the model in the experimentally accessible range (Figure 2H, left).

We also investigated the behavior of other constructs that included the D-DREs, where we could also see an agreement with the model predictions and the AND-gate behavior, however less pronounced (Supplementary Figure S9).

A model for signal integration of the human *CYP1A1* promoter

After we had established that our mathematical model can indeed describe the behavior of the well-controllable synthetic promoter constructs, we next decided to use this approach to analyze the human *CYP1A1* promoter. Similarly as with the synthetic constructs, we generated a promoter reporter library with 17 different combinations of point mutations that inactivate specific TFBS (Figure 3A and Supplementary Figure S1). Each of these reporters was then transfected into mouse hepatoma cells of line 55.1c and treated with increasing concentrations of TCDD for 24 h. The measured luciferase activity data was preprocessed as described in the Materials and Methods section and the supporting text. From the resulting dataset (gray dots in Figure 3B), it can be seen that the behavior of the natural promoter is less obvious than that of the synthetic promoter constructs. For example, the promoter that contains all binding sites except the F-DRE (EDTC construct) shows the strongest induction upon TCDD treatment, even stronger than the WT construct that harbors all binding sites. Furthermore, it was observed that constructs DTC, FTC, EDT, FET, FDTC, FEDC and FEDT displayed almost the same inducibility although there are different types and numbers of TFBS present. Finally, it was surprising that the constructs TC and ET presented with a 3-fold inducibility while the construct DT is hardly inducible at all.

To disentangle the effects of the different binding sites and analyze if the behavior of the promoter can be understood by the interactions between the binding sites, we again set up a model. The association constants for the C-DRE and D-DRE could be taken from the model fit of the synthetic promoter constructs. Within the synthetic constructs we have seen that cooperative binding is only present between direct or next neighbors, which led us to eliminate long-distance interactions from the model (between C-DRE and E-DRE, C-DRE and F-DRE, and β-catenin/TCF and F-DRE). Furthermore, if a binding site is inactivated through a mutation in the construct we removed the mutant TFBS from the model by setting the corresponding association constant between TF and TFBS to zero.

When fitting the remaining 19 parameters to the entire data including a β-catenin titration series (blue triangles in Figure 3B) we see that the model can recapitulate the data well, except for the very strong induction of the constructs EDTC and FETC (Figure 3B). That may be a result of more complex cooperativity between three factors, which we did not include in the model.

An interesting aspect of the experimental data and the model is that, compared to the synthetic model (Figure 2D), the induction of the natural promoter (Figure 3B) showed a less pronounced sigmoidality, i.e. not such a strong switch-like behavior.

Cooperativity in the wild type promoter. To understand in more detail how the difference in the sigmoidality of the stimulus-response curves can arise, we investigated the parameter values. We used the profile likelihood method to determine confidence intervals and to analyze which parameters can be robustly estimated from the experimental data (63). We found that 18 of the 19 parameters showed a clear minimum and thus were identifiable by the data, when we apply point-wise confidence thresholds. Two of the parameters (q_P and ε_D) were indistinguishable from zero, i.e. irrelevant for the model to describe the data. Only the cooperative binding energy between the D- and the E-DRE was practically non-identifiably (according to the terminology in (63)) and thus could not be reliably determined from the data (Supplementary Figures S10 and S11). When we

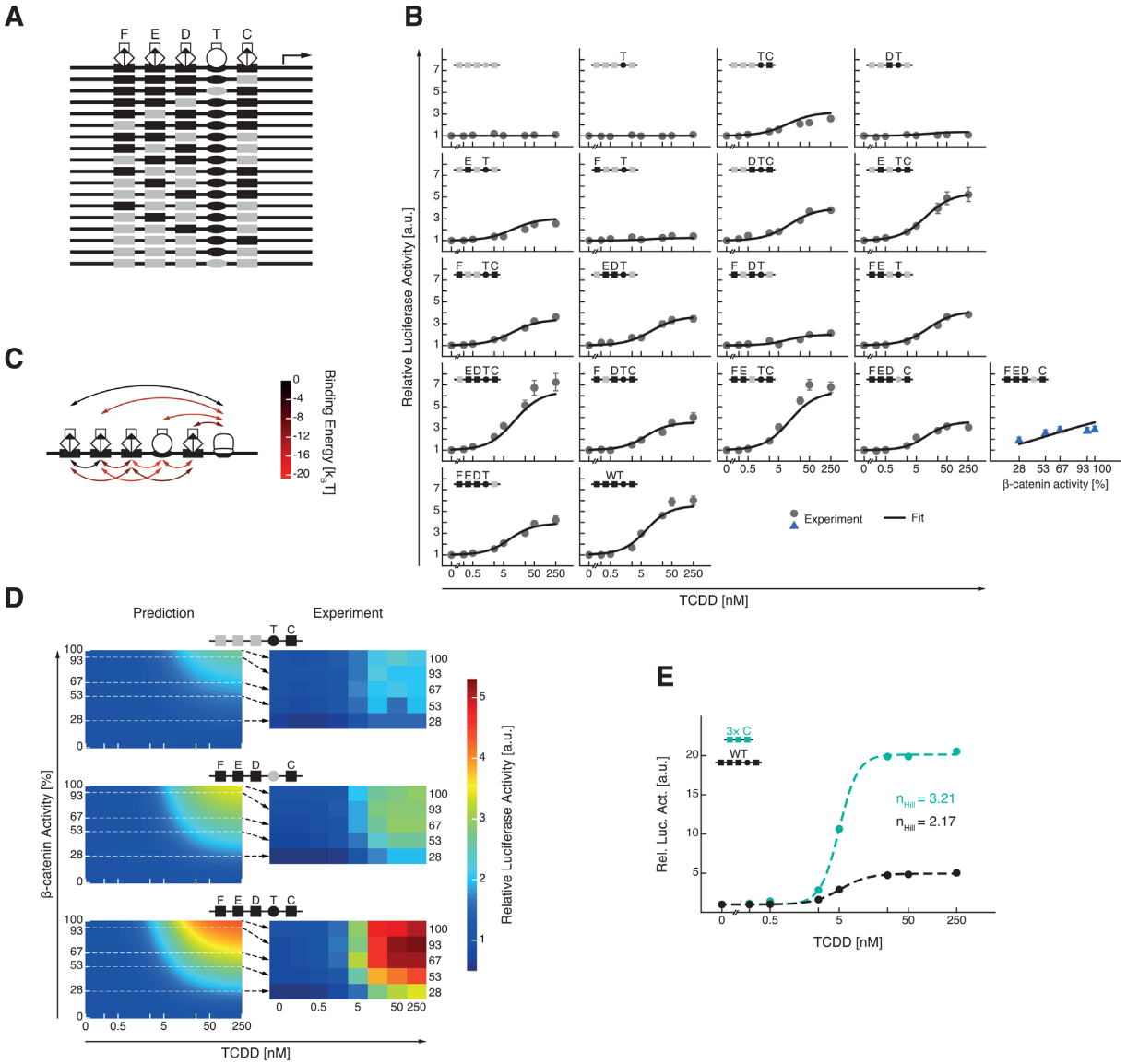

Figure 3. Promoter logic of the human *CYP1A1* promoter dissected by a mathematical model. (**A**) Schematic representation of the reporter library for the natural *CYP1A1* reporter. Rectangular binding sites are DREs while elliptical binding sites are targets of β-catenin/TCF. Gray symbols depict TFBS that were inactivated by point-mutation while the WT binding sequence is represented by black symbols. (**B**) The response of the natural promoter constructs to TCDD concentration series can be explained with a thermodynamic model. Relative luciferase activity was measured over increasing TCDD concentrations and an additional β-catenin titration for 18 natural promoter constructs (gray dots and blue triangles). The β-catenin titration was measured at 250 nM TCDD. Error bars represent one standard deviation of 4–10 biological replicates. The black solid curves show the fits of the mathematical model. (**C**) Binding between the RNAP and the TFs dominates the natural promoter constructs. Binding energies resulting from fits displayed on the natural promoter. Arrows depict existing binding events. Their colors represent binding strength where lower values represent stronger association. Only short range binding events were considered in the model, those equal or close to zero are not depicted as arrows. (**D**) Predicted response to Wnt/β-catenin and AhR signaling pathway for three different natural promoter constructs. Left: prediction of the promoter activity of three natural constructs (C-DRE and β-catenin/TCF TFBS present; β-catenin/TCF TFBS inactivated through point-mutation; WT construct) by the mathematical model through variation of the β-catenin activity parameter. Right: Corresponding measurements of the relative luciferase activity of three constructs where the β-catenin activity was modulated by the iCRT3 inhibitor. The cells were simultaneously stimulated with increasing TCDD concentrations. (**E**) Hill Coefficients confirm more switch-like behavior of synthetic promoter constructs. From the double-stimulated datasets of the synthetic 3× C DRE and the natural WT promoter construct the subsets for 100 % β-catenin activity was fitted to Hill functions. Relative luciferase activity for both constructs is represented by turquoise (synthetic 3× C DRE construct) and black (natural WT construct) dots. The dashed lines depict the corresponding Hill functions. The resulting Hill coefficients for the two constructs are shown.

furthermore compared the interaction energy between the TFs and the RNAP with those from the synthetic promoters, we found that they were generally higher in the natural promoter constructs, and that they were relatively independent of the distance to the RNAP binding site (Figure 3C, Supplementary Figure S10 and Supplementary Table S2). While the binding energy between the D-DRE and the RNAP is zero, its cooperative interaction with the β-catenin/TCF TF is the strongest. On the other hand, the strong binding energies of the C-DRE and the E-DRE with the RNAP resulted also in strong cooperative interactions to the remaining TFBS. But most interesting is the strong influence of the β-catenin/TCF TFBS. It not only has the strongest connection with the RNAP but also communicates most strongly with its direct or next neighbors.

The model is able to predict signal integration by the promoter. After having established a model that is able to quantitatively describe the response to TCCD stimulation, we aimed to predict how TCCD and β-catenin are integrated on the *CYP1A1* promoter (Figure 3D). We first simulated a construct where all TFBS were mutated except the most proximal C-DRE and the β-catenin/TCF binding site. The model predicted a weak induction that requires both β-catenin and TCDD signaling (Figure 3D top left), which is similar to what we observed when we measured the reporter activity for this construct for varying concentrations of TCDD and β-catenin (Figure 3D top right). Next, we investigated the behavior of the promoter where only the β-catenin/TCF site is mutated, allowing us to see if the model correctly predicts the cross-talk of AhR and Wnt/β-catenin signaling that occurs upstream of the promoter. The model predicts a 3- to 4-fold induction if both β-catenin and TCDD are at high doses, and shows rather shallow stimulus-response curves (Figure 3D, middle left). Similarly, the corresponding experimental data set shows a 3-fold induction by TCDD that is modulated by β-catenin, confirming that strong regulatory cross-talk occurs independently of the β-catenin/TCF binding site (Figure 3D, middle right). Finally, when simulating the full model representing the WT promoter, we observe a stronger, ∼5-fold induction that shows a gradual increase with both TCDD and β-catenin, which closely resembles that of the experiment (Figure 3D bottom left and bottom right). When we compared those measurements with the TCDD titration experiments, it was reassuring to see that these independent experiments were in good agreement (Supplementary Figure S12).

When comparing experimental data and mathematical model from the natural promoter with the synthetic promoter, we noticed that the stimulus-response pattern was less switch-like in the natural promoters. To quantify the differences in steepness of the dose-response curves between the synthetic and the natural promoter constructs, we calculated the Hill coefficients by fitting Hill functions to the dose–response curves for TCDD of the synthetic 3× C DRE as well as the natural WT promoter construct (Figure 3E). The resulting Hill coefficients are 3.21 and 2.17 for the synthetic and the natural construct confirming a more switch-like behavior of the synthetic constructs.

In summary, *CYP1A1* expression requires the simultaneous presence of two signals, as a single stimulant — TCDD or β-catenin — is not sufficient to trigger full *CYP1A1* expression. Thereby expression is restricted to the pericentral zone of a liver lobule where β-catenin levels are high. This behavior closely resembles a logical AND gate, which has been widely discussed in recent literature (37,39,40,67,68). It is interesting to see that the apparent AND-gate logic of the *CYP1A1* promoter is established not only by interactions between the *cis*-regulatory binding sites, but in addition also by the interaction between Wnt/β-catenin and AhR signaling pathways.

In another study, modeling the thermodynamics of *cis*-regulatory binding sides of the hepcidin promoter, it was shown that the response to two different pathways (SMAD and STAT) did not follow an AND gate, although some mutants, which are associated with hereditary hemochromatosis, did. This reemphasizes the evolutionary plasticity of promoter logic (64).

Thermodynamic model predicts physiological zonation pattern. We reasoned that less switch-like behavior might be important for zonal gene expression: if no exogenous AhR agonists are present, the area of *CYP1A1* expression around the central vein is small and restricted to very high β-catenin activity, whereas it extends larger into periportal regions when AhR receptors are stimulated (Figure 4). Similarly, in our model the expression of the reporter remains low in cells that show 28 % of β-catenin activity until TCDD concentrations rise above 10–50 nM and increases at higher doses. In contrast, cells with 100 % β-catenin activity show strong and saturating expression already at 5 nM TCDD.

To compare predicted spatial expression patterns with data, we mapped the model predictions for different TCDD concentrations on hexagonal grids that are similar to liver lobules, and assumed a linear gradient for β-catenin from the center to the periphery of each hexagon (Figure 4, top). This type of gradient was deduced from immunohistochemical images showing zonation of active β-catenin in mouse liver (23). For the TCDD concentration, we assumed that the applied TCDD concentration will distribute evenly over the whole lobule. We then performed immunostainings of CYP1A in mouse livers treated with different concentrations of the AhR inducer 3-methylcholanthrene (3-MC) (Figure 4 bottom). When comparing these patterns, we saw that without stimulation, the experimental data showed a basal perivenous expression of *CYP1A* that was not predicted by our model. Most probably, the reason for this is that the AhR retains basal activity in the liver due to the presence of low levels of endogenous AhR agonists (e.g. comparable to the 5 nM TCDD in our stimulations). For 10 mg/kg body weight 3-MC a broadening of the *CYP1A*-expressing region around the central veins could be observed while higher concentrations of 3-MC only lead to an increased expression intensity (represented by darker browns). This effect could be replicated well with the thermodynamic model.

While the thermodynamic model can accurately describe the promoter logic of the promoters in the luciferase assays, it should be noted that the fold-change of the *CYP1A1* gene is higher in the promoter of human cell lines

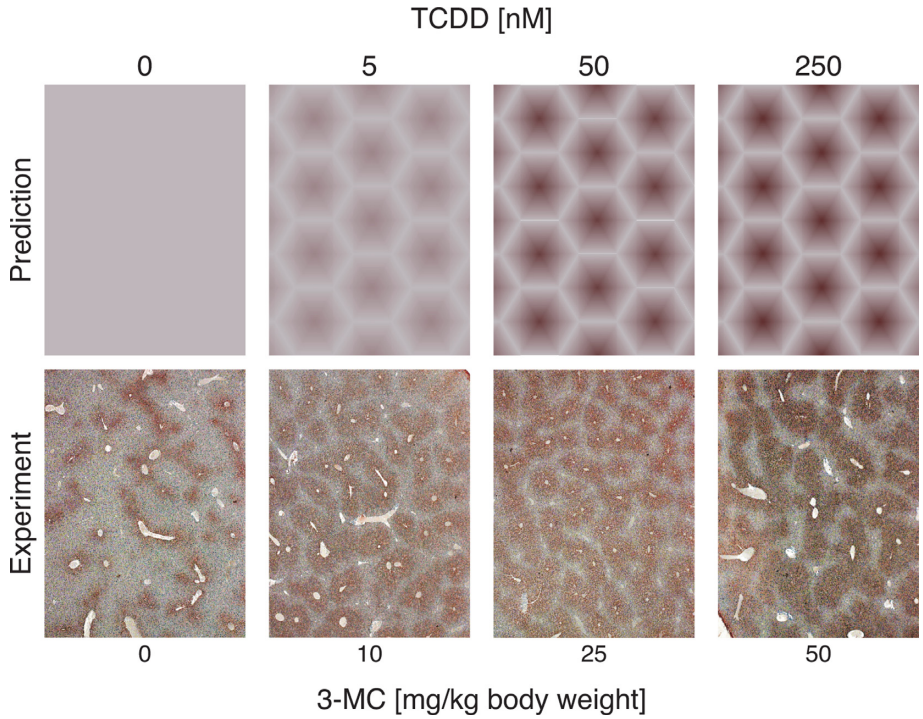

Figure 4. Model predicts spatial expression of *CYP1A1* as a result of its promoter logic. Top: Predictions of the thermodynamic model for the WT construct for selected TCDD concentrations. Each hexagon represents an idealized hepatic lobule that is exposed to a β-catenin activity gradient from central vein (high) to portal vein (low). Bottom: Representative immunostainings of mouse liver for CYP1A for different concentrations of 3-MC.

(HepG2, CaCo-2, HT116) when compared to luciferase assays (29,54). Thus, the model captures the essence of the promoter logic, but additional factors such as other binding sites or chromatin composition may influence the fold-change. It should also be noted that the kinetic details of the interactions in the promoter will determine how large the zone of expression will be and if the zone widens if toxins are present. While the described interaction of Wnt/β-catenin and AhR is sufficient to describe zonation of the *CYP1A* expression, it is likely that the zonation of other genes is controlled by different and/or additional mechanisms. The expression of glutamine synthetase (GS), for example, is restricted to a small pericentral zone which can be explained with an additional regulatory mechanism that silences an 5′-enhancer within the GS gene (69).

Taken together, we reasoned that interactions between Wnt/β-catenin and AhR at the signaling level, together with the complex cooperative interactions between the DREs in the human *CYP1A1* promoter enable the spatiotemporal expression pattern observed *in vivo*.

CONCLUSION

There is currently a huge discrepancy in the level of detail for the quantitative description of signaling systems and gene regulation for mammalian cells. There is for example a large body of literature on kinetic models of kinase signaling (70). In contrast, most models of gene regulation in mammalian cells address the behavior and dynamics of larger networks neglecting the regulatory complexity of a promoter, leaving the question open of how the regulatory logic of a promoter is established. There have been a few notable exemptions recently that either use more complex input-output functions (71), or make use of a thermodynamic description of the promoter (50,64). An interesting aspect of our model is that it can fully describe the behavior of the promoter by a thermodynamic equilibrium, i.e. purely by binding between the proteins of the signaling pathways with each other and the promoter.

During recent years, statistical models have been developed that can be trained on genome-wide data, such as transcription factor binding data, open chromatin, and histone marks to predict gene expression (72,73). These models are phenomenological, and it would be interesting to combine these with a thermodynamic description, as it has been done previously for gene expression in *D. melanogaster* development (53).

Taken together, our work demonstrates that statistical mechanics modeling together with combinatorial reporter libraries can be successfully applied to dissect the molecular details that establish promoter logic. This provides insight into how physiological relevant expression patterns are established.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Technical assistance by Elke Zabinsky and Johanna Mahr is greatly acknowledged. We also thank Dr C. Roland Wolf (Dundee, UK) for the gift of the CYP1A antiserum.

For critically reading the manuscript we thank Franziska Witzel, Bertram Klinger and Mattias Rydenfelt.

FUNDING

Federal Ministry of Education and Research (FORSYS-Partner, OncoPath and SysDT). Funding for open access charge: Institutional Funding at Charite.
Conflict of interest statement. None declared.

REFERENCES

1. Bock,K.W. (1994) Aryl hydrocarbon or dioxin receptor: biologic and toxic responses. *Rev. Physiol. Biochem. Pharmacol.*, **125**, 1–42.
2. Hankinson,O., Hoffman,E.C., Reyes,H., Chu,F.F., Sander,F., Conley,L.H. and Brooks,B.A. (1991) Cloning of a factor required for activity of the Ah (dioxin) receptor. *Science*, **252**, 954–958.
3. Kress,S., Reichert,J. and Schwarz,M. (1998) Functional analysis of the human cytochrome P4501A1 (CYP1A1) gene enhancer. *Eur. J. Biochem.*, **258**, 803–812.
4. Reyes,H., Reisz-Porszasz,S. and Hankinson,O. (1992) Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor. *Science*, **256**, 1193–1195.
5. Fujisawa-Sehara,A., Yamane,M. and Fujii-Kuriyama,Y. (1988) A DNA-binding factor specific for xenobiotic responsive elements of P-450c gene exists as a cryptic form in cytoplasm: its possible translocation to nucleus. *PNAS*, **85**, 5859–5863.
6. Swanson,H.I., Chan,W.K. and Bradfield,C.A. (1995) DNA binding specificities and pairing rules of the Ah receptor, ARNT, and SIM proteins. *J. Biol. Chem.*, **270**, 26292–26302.
7. Yao,E.F. and Denison,M.S. (1992) DNA sequence determinants for binding of transformed Ah receptor to a dioxin-responsive enhancer. *Biochemistry*, **31**, 5060–5067.
8. Denison,M.S., Fisher,J.M. and Whitlock,J.P. (1988) Inducible, receptor-dependent protein-DNA interactions at a dioxin-responsive transcriptional enhancer. *PNAS*, **85**, 2528–2532.
9. Bacsi,S.G., Reisz-Porszasz,S. and Hankinson,O. (1995) Orientation of the heterodimeric aryl hydrocarbon (dioxin) receptor complex on its asymmetric DNA recognition sequence. *Mol. Pharmacol.*, **47**, 432–438.
10. Denison,M.S., Fisher,J.M. and Whitlock,J.P. (1989) Protein-DNA interactions at recognition sites for the dioxin-Ah receptor complex. *J. Biol. Chem.*, **264**, 16478–16482.
11. Kubota,M., Kawajiri,K., Sogawa,K., Kaizu,Y., Sawaya,T., Watanabe,J., Gotoh,O. and Fujino,H. (1991) Xenobiotic responsive element in the 5′-upstream region of the human P-450c gene. *J. Biochem.*, **110**, 232–236.
12. IARC (1997) *IARC Working Group on the Evaluation of Carcinogenic Risks to Humans*. Vol. **69**, pp. 1–631.
13. Oinonen,T. and Lindros,K.O. (1998) Zonation of hepatic cytochrome P-450 expression and regulation. *Biochem. J.*, **329**, 17–35.
14. Lindros,K.O. (1997) Zonation of cytochrome P450 expression, drug metabolism and toxicity in liver. *Gen. Pharmacol.*, **28**, 191–196.
15. Santostefano,M.J., Richardson,V.M., Walker,N.J., Blanton,J., Lindros,K.O., Lucier,G.W., Alcasey,S.K. and Birnbaum,L.S. (1999) Dose-dependent localization of TCDD in isolated centrilobular and periportal hepatocytes. *Toxicol. Sci.*, **52**, 9–19.
16. van Sliedregt,A. and van Bezooijen,C.F. (1990) Effect of different doses of 3-methylcholanthrene on the localization of the 3-methylcholanthrene-inducible isoenzymes of cytochrome P450 within the centrilobular and periportal zones of the rat liver. *Biochem. Pharmacol.*, **39**, 1703–1708.
17. Bonin,M., Braeuning,A., Ittrich,C., Köhle,C., Hailfinger,S., Buchmann,A. and Schwarz,M. (2006) Differential gene expression in periportal and perivenous mouse hepatocytes. *FEBS J.*, **273**, 5051–5061.
18. Wagenaar,G.T., Chamuleau,R.A., de Haan,J.G., Maas,M.A., de Boer,P.A., Marx,F., Moorman,A.F., Frederiks,W.M. and Lamers,W.H. (1993) Experimental evidence that the physiological position of the liver within the circulation is not a major determinant of zonation of gene expression. *Hepatology*, **18**, 1144–1153.
19. Gebhardt,R., Baldysiak-Figiel,A., Krügel,V., Ueberham,E. and Gaunitz,F. (2007) Hepatocellular expression of glutamine synthetase: an indicator of morphogen actions as master regulators of zonation in adult liver. *Prog. Histochem. Cytochem.*, **41**, 201–266.
20. Gebhardt,R. (1992) Metabolic zonation of the liver: regulation and implications for liver function. *Pharmacol. Ther.*, **53**, 275–354.
21. Jungermann,K. and Kietzmann,T. (1996) Zonation of parenchymal and nonparenchymal metabolism in liver. *Annu. Rev. Nutr.*, **16**, 179–203.
22. Braeuning,A., Köhle,C., Buchmann,A. and Schwarz,M. (2011) Coordinate regulation of cytochrome P450 1a1 expression in mouse liver by the aryl hydrocarbon receptor and the beta-catenin pathway. *Toxicol. Sci.*, **122**, 16–25.
23. Benhamouche,S., Decaens,T., Godard,C., Chambrey,R., Rickman,D.S., Moinard,C., Vasseur-Cognet,M., Kuo,C.J., Kahn,A., Perret,C. *et al.* (2006) Apc tumor suppressor gene is the 'zonation-keeper' of mouse liver. *Dev. Cell*, **10**, 759–770.
24. Hailfinger,S., Jaworski,M., Braeuning,A., Buchmann,A. and Schwarz,M. (2006) Zonal gene expression in murine liver: lessons from tumors. *Hepatology*, **43**, 407–414.
25. Torre,C., Perret,C. and Colnot,S. (2010) Molecular determinants of liver zonation. *Prog. Mol. Biol. Transl. Sci.*, **97**, 127–150.
26. Braeuning,A., Sanna,R., Huelsken,J. and Schwarz,M. (2009) Inducibility of drug-metabolizing enzymes by xenobiotics in mice with liver-specific knockout of Ctnnb1. *Drug Metab. Dispos.*, **37**, 1138–1145.
27. Loeppen,S., Koehle,C., Buchmann,A. and Schwarz,M. (2005) A beta-catenin-dependent pathway regulates expression of cytochrome P450 isoforms in mouse liver tumors. *Carcinogenesis*, **26**, 239–248.
28. Schreiber,S., Rignall,B., Braeuning,A., Marx-Stoelting,P., Ott,T., Buchmann,A., Hammad,S., Hengstler,J.G., Schwarz,M. and Köhle,C. (2011) Phenotype of single hepatocytes expressing an activated version of β-catenin in liver of transgenic mice. *J. Mol. Histol.*, **42**, 393–400.
29. Vaas,S., Kreft,L., Schwarz,M. and Braeuning,A. (2014) Cooperation of structurally different aryl hydrocarbon receptor agonists and β-catenin in the regulation of CYP1A expression. *Toxicology*, **325**, 31–41.
30. Kasai,S., Ishigaki,T., Takumi,R., Kamimura,T. and Kikuchi,H. (2013) Beta-catenin signaling induces CYP1A1 expression by disrupting adherens junctions in Caco-2 human colon carcinoma cells. *Biochim. Biophys. Acta*, **1830**, 2509–2516.
31. Ackers,G.K., Johnson,A.D. and Shea,M.A. (1982) Quantitative model for gene regulation by lambda phage repressor. *PNAS*, **79**, 1129–1133.
32. Shea,M.A. and Ackers,G.K. (1985) The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. *J. Mol. Biol.*, **181**, 211–230.
33. Buchler,N.E., Gerland,U. and Hwa,T. (2003) On schemes of combinatorial transcription logic. *PNAS*, **100**, 5136–5141.
34. Bintu,L., Buchler,N.E., Garcia,H.G., Gerland,U., Hwa,T., Kondev,J. and Phillips,R. (2004) Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.*, **15**, 116–124.
35. Bintu,L., Buchler,N.E., Garcia,H.G., Gerland,U., Hwa,T., Kondev,J., Kuhlman,T. and Phillips,R. (2005) Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.*, **15**, 125–135.
36. Veitia,R.A. (2003) A sigmoidal transcriptional response: cooperativity, synergy and dosage effects. *Biol. Rev. Camb. Philos. Soc.*, **78**, 149–170.
37. Frank,T.D., Cheong,A., Okada-Hatakeyama,M. and Kholodenko,B.N. (2012) Catching transcriptional regulation by thermostatistical modeling. *Phys. Biol.*, **9**, 045007.
38. Hermsen,R., Ursem,B. and ten Wolde,P.R. (2010) Combinatorial gene regulation using auto-regulation. *PLoS Comput. Biol.*, **6**, e1000813.
39. Setty,Y., Mayo,A.E., Surette,M.G. and Alon,U. (2003) Detailed map of a cis-regulatory input function. *PNAS*, **100**, 7702–7707.
40. Raveh-Sadka,T., Levo,M. and Segal,E. (2009) Incorporating nucleosomes into thermodynamic models of transcription regulation. *Genome Res.*, **19**, 1480–1496.
41. Coulon,A., Gandrillon,O. and Beslon,G. (2010) On the spontaneous stochastic dynamics of a single gene: complexity of the molecular interplay at the promoter. *BMC Syst. Biol.*, **4**, 2.

42. Kaplan,T., Li,X.-Y., Sabo,P.J., Thomas,S., Stamatoyannopoulos,J.A., Biggin,M.D. and Eisen,M.B. (2010) Quantitative models of the mechanisms that control genome-wide patterns of transcription factor binding during early Drosophila development. *PLoS Genet.*, **7**, e1001290.
43. Chen,Chieh-Chun, Zhu,X.-G. and Zhong,S. (2008) Selection of thermodynamic models for combinatorial control of multiple transcription factors in early differentiation of embryonic stem cells. *BMC Genomics*, **9**, S18.
44. Wray,G.A. (2007) The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.*, **8**, 206–216.
45. Sherman,M.S. and Cohen,B.A. (2012) Thermodynamic state ensemble models of cis-regulation. *PLoS Comput. Biol.*, **8**, e1002407.
46. He,X., Samee,M.A.H., Blatti,C. and Sinha,S. (2010) Thermodynamics-based models of transcriptional regulation by enhancers: the roles of synergistic activation, cooperative binding and short-range repression. *PLoS Comput. Biol.*, **6**.
47. Garcia,H.G., Sanchez,A.A., Kuhlman,T.T., Kondev,J.J. and Phillips,R.R. (2010) Transcription by the numbers redux: experiments and calculations that surprise. *Trends Cell Biol.*, **20**, 11–11.
48. Hermsen,R., Tans,S. and ten Wolde,P.R. (2006) Transcriptional regulation by competing transcription factor modules. *PLoS Comput. Biol.*, **2**, e164.
49. Gertz,J., Siggia,E.D. and Cohen,B.A. (2009) Analysis of combinatorial cis-regulation in synthetic and genomic promoters. *Nature*, **457**, 215–218.
50. Giorgetti,L., Siggers,T., Tiana,G., Caprara,G., Notarbartolo,S., Corona,T., Pasparakis,M., Milani,P., Bulyk,M.L. and Natoli,G. (2010) Noncooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. *Mol. Cell*, **37**, 418–428.
51. White,M.A., Parker,D.S.D., Barolo,S. and Cohen,B.A. (2011) A model of spatially restricted transcription in opposing gradients of activators and repressors. *Mol. Syst. Biol.*, **8**, 614–614.
52. Zinzen,R.P., Senger,K., Levine,M. and Papatsenko,D. (2006) Computational models for neurogenic gene expression in the Drosophila embryo. *Curr. Biol.*, **16**, 1358–1365.
53. Segal,E., Raveh-Sadka,T., Schroeder,M., Unnerstall,U. and Gaul,U. (2008) Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature*, **451**, 535–540.
54. Schreiber,T.D., Köhle,C., Buckler,F., Schmohl,S., Braeuning,A., Schmiechen,A., Schwarz,M. and Münzel,P.A. (2006) Regulation of CYP1A1 gene expression by the antioxidant tert-butylhydroquinone. *Drug Metab. Dispos.*, **34**, 1096–1101.
55. Braeuning,A., Menzel,M., Kleinschnitz,E.-M., Harada,N., Tamai,Y., Köhle,C., Buchmann,A. and Schwarz,M. (2007) Serum components and activated Ha-ras antagonize expression of perivenous marker genes stimulated by beta-catenin signaling in mouse hepatocytes. *FEBS J.*, **274**, 4766–4777.
56. Braeuning,A. and Buchmann,A. (2009) The glycogen synthase kinase inhibitor 3-(2,4-dichlorophenyl)-4-(1-methyl-1H-indol-3-yl)-1H-pyrrole-2,5-dione (SB216763) is a partial agonist of the aryl hydrocarbon receptor. *Drug Metab. Dispos.*, **37**, 1576–1580.
57. Braeuning,A. and Vetter,S. (2012) The nuclear factor κB inhibitor (E)-2-fluoro-4'-methoxystilbene inhibits firefly luciferase. *Biosci. Rep.*, **32**, 531–537.
58. Braeuning,A., Vetter,S., Orsetti,S. and Schwarz,M. (2012) Paradoxical cytotoxicity of tert-butylhydroquinone in vitro: what kills the untreated cells? *Arch. Toxicol.*, **86**, 1481–1487.
59. Skehan,P., Storeng,R., Scudiero,D., Monks,A., McMahon,J., Vistica,D., Warren,J.T., Bokesch,H., Kenney,S. and Boyd,M.R. (1990) New colorimetric cytotoxicity assay for anticancer-drug screening. *J. Natl. Cancer Inst.*, **82**, 1107–1112.
60. Pfaffl,M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, **29**, e45.
61. Braeuning,A., Singh,Y., Rignall,B., Buchmann,A., Hammad,S., Othman,A., von Recklinghausen,I., Godoy,P., Hoehme,S., Drasdo,D. *et al.* (2010) Phenotype and growth behavior of residual β-catenin-positive hepatocytes in livers of β-catenin-deficient mice. *Histochem. Cell Biol.*, **134**, 469–481.
62. McKay,M.D., Beckman,R.J. and Conover,W.J. (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **21**, 239–245.
63. Raue,A., Kreutz,C., Maiwald,T., Bachmann,J., Schilling,M., Klingmuller,U. and Timmer,J. (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, **25**, 1923–1929.
64. Casanovas,G., Banerji,A., d'Alessio,F., Muckenthaler,M.U. and Legewie,S. (2014) A multi-scale model of hepcidin promoter regulation reveals factors controlling systemic iron homeostasis. *PLoS Comput. Biol.*, **10**, e1003421.
65. Korenčič,A., Bordyugov,G., Košir,R., Rozman,D., Goličnik,M. and Herzel,H. (2012) The interplay of cis-regulatory elements rules circadian rhythms in mouse liver. *PLoS ONE*, **7**, e46835.
66. Fakhouri,W.D., Ay,A., Sayal,R., Dresch,J., Dayringer,E. and Arnosti,D.N. (2009) Deciphering a transcriptional regulatory code: modeling short-range repression in the Drosophila embryo. *Mol. Syst. Biol.*, **6**, 341–341.
67. Cox,R.S., Surette,M.G. and Elowitz,M.B. (2007) Programming gene expression with combinatorial promoters. *Mol. Syst. Biol.*, **3**, 145.
68. Mayo,A.E., Setty,Y., Shavit,S., Zaslaver,A. and Alon,U. (2006) Plasticity of the cis-regulatory input function of a gene. *PLoS Biol*, **4**, e45.
69. Gaunitz,F., Deichsel,D., Heise,K., Werth,M., Anderegg,U. and Gebhardt,R. (2005) An intronic silencer element is responsible for specific zonal expression of glutamine synthetase in the rat liver. *Hepatology*, **41**, 1225–1232.
70. Kholodenko,B.N., Hancock,J.F. and Kolch,W. (2010) Signalling ballet in space and time. *Nat. Rev. Mol. Cell. Biol.*, **11**, 414–426.
71. Mrowka,R., Steege,A., Kaps,C., Herzel,H., Thiele,B.J., Persson,P.B. and Blüthgen,N. (2007) Dissecting the action of an evolutionary conserved non-coding region on renin promoter activity. *Nucleic Acids Res.*, **35**, 5120–5129.
72. Karlić,R., Chung,H.-R., Lasserre,J., Vlahovicek,K. and Vingron,M. (2010) Histone modification levels are predictive for gene expression. *PNAS*, **107**, 2926–2931.
73. Natarajan,A., Yardimci,G.G., Sheffield,N.C., Crawford,G.E. and Ohler,U. (2012) Predicting cell-type-specific gene expression from regions of open chromatin. *Genome Res.*, **22**, 1711–1722.