DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Binding of Circadian Transcription Regulators and Rhythmic Gene Expression in Human Cells

presented by
Diplom-Biochemist Julia Stefanski
born in Poznań, Poland

DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Biochemist Julia Stefanski
born in Poznań, Poland

Date of oral examination: 30.04.2013

Binding of Circadian Transcription Regulators and Rhythmic Gene Expression in Human Cells

Referees: Prof. Dr. Michael Brunner
Prof. Dr. Felix Wieland

Für meinen Vater

Summary

The circadian clock is a tightly regulated mechanism that has evolved in most organisms to enable them to anticipate daily reoccurring changes in the environment. Expression analysis in mouse suggests that about 10% of the mammalian transcriptome may be under circadian control, however the overlap between different tissues is very small [1]. Mechanisms determining this cell type specificity of clock controlled genes are not yet understood. The human osteosarcoma cell line U2OS expresses a functional circadian clock and has been used in many circadian studies [2, 3, 4], yet, except for components of the core clock machinery, no genes have been found to be rhythmically expressed in this cell line in a microarray analysis [3] whereas about 3000 rhythmic transcripts have been identified in mouse liver [3, 5, 6].

This work shows ChIP-sequencing data of BMAL1, CLOCK, and CRY1 that indicates that the circadian transcription regulators bind in sequence-specific manner to several thousand sites in the genome of U2OS, comparable to the number found in mouse liver [5]. However, time course microarray experiments carried out to verify functionality of these sites, found only 58 rhythmic genes harboring a binding site, suggesting a high amount of opportunistic binding that cannot be linked to transcription. Even most genes with high-scoring binding sites in their promoter were apparently arrhythmic. U2OS cell lines stably expressing luciferase reporter constructs under control of representative promoters were generated and revealed that BMAL1 and CLOCK functionally act on these, but rhythmic expression appears to be superimposed by a high basal transcription. The data indicates that high amplitude transcription rhythms are masked by constitutive transcription that overwrites the rhythmic contribution of the circadian clock, resulting in apparent arrhythmicity due to low amplitude oscillations. This work suggests that the cell-type specific circadian transcriptome is not determined on the level of BMAL1-CLOCK binding and rhythmic activation of transcription by these circadian regulators, but rather by the cell-type specific ratio of circadian versus gene-specific or general transcription regulators.

Zusammenfassung

Die zirkadiane Uhr ist ein streng regulierter Mechanismus, der sich in den meisten Organismen entwickelt hat um täglich wiederkehrende Veränderungen in der Umgebung zu antizipieren. Analysen des Maus Transkriptoms deuten darauf hin, dass ungefähr 10% der in Säugetieren exprimierten Gene unter der Kontrolle der zirkadianen Uhr stehen, jedoch scheint die Überlappung dieser in verschiedenen Geweben sehr klein zu sein [1]. Mechanismen, wie es zu dieser Zelltyp-Spezifität von durch die Uhr regulierten Genen kommt, sind bislang unverstanden. Die humane Osteosarkoma Zelllinie U2OS besitzt eine funktionelle zirkadiane Uhr und wurde in vielen chronobiologischen Studien verwendet [2, 3, 4], dennoch wurden in einer Microarray-Analyse, außer den Hauptkomponenten der zirkadianen Uhr, keine rhythmischen Gene in dieser Zelllinie gefunden [3], während in der Leber von Mäusen ungefähr 3000 rhythmische Transkripte identifiziert wurden [3, 6, 5].

Diese Arbeit zeigt ChIP-sequencing Daten von BMAL1, CLOCK und CRY1, die darauf hindeuten, dass die zirkadianen Transkriptionsregulatoren sequenz-spezifisch an mehrere tausend Stellen im Genom von U2OS binden, einer Zahl vergleichbar mit der in Maus Leber [5]. Zwei Zeitreihen wurden auf Microarrays getestet um die Funktionalität der Bindungsstellen zu verifizieren, jedoch wurden nur 58 rhythmische Gene mit Bindungsstellen gefunden, was darauf schließen lässt, dass der Großteil der Bindungsstellen opportunistisch ist und nicht mit rhythmischer Transkription in Verbindung gebracht werden kann. Sogar Gene mit besonders guten Bindungsstellen in ihrem Promotor erscheinen arrhythmisch. U2OS Zelllinien wurden generiert, bei denen Luziferase-Reporter Konstrukte unter Kontrolle von repräsentativen Promotoren stabil integriert wurden. Diese Zellen zeigen, dass BMAL1 und CLOCK zwar auf diese Promotoren wirken, eine rhythmische Expression aber gleichzeitig überlagert wird durch eine hohe Basal-Transkription. Die Daten deuten darauf hin, dass Transkriptionsrhythmen mit hoher Amplitude durch konstitutive Expression maskiert werden, die die rhythmische Komponente der zirkadianen Uhr überschreibt, so dass eine scheinbare Arrhythmizität, bedingt durch Oszillationen mit sehr niedriger Amplitude, entsteht. Diese Arbeit schlägt vor, dass das Zelltyp-spezifische zirkadiane Transkriptom nicht auf der Ebene von BMAL1-CLOCK Bindung und rhythmischer Aktivierung der Transkription bestimmt wird, sondern durch ein Zelltyp-spezifisches Verhältnis von zirkadianen zu genspezifischen oder allgemeinen Transkriptionsregulatoren.

Contents

1. Introduction

1.1. Circadian clocks - an overview

Circadian clocks (Latin "circa diem" = about a day) are endogenous devices that have evolved in all kingdoms of life in order to anticipate the alternating challenges of day and night imposed on all living beings by the rotation of the earth. Self sustained molecular clocks are found on the level of single cells not only in unicellular organisms but also in multicellular organisms [7]. The circadian clock is not essential for survival but it conveys an evolutionary advantage to the organism. In Cyanobacteria and lower eukaryotes like the filamentous fungus *Neurospora crassa*, the clock organizes metabolism and growth in a temporal manner. In higher eukaryotes the clock regulates rhythmic behavior in addition to physiological and metabolic rhythms in the course of the day.

There are three main properties of all circadian clocks: First, they are endogenous mechanisms, which means that the core of the clock lies in an internal oscillator encoded by the genome. Second, circadian clocks are self-sustained, therefore, even in the absence of environmental stimuli, the clock remains functional and influences the expression of genes in a rhythmic fashion. Third, they are temperature-compensated. Although the molecular oscillator depends on a series of chemical reactions, these are balanced in such a way, that the period retains approximately the same length even when the temperature changes. This is most important in poikilotherm organisms whose internal temperature varies with the external temperature.

The circadian clock is organized in three parts: An input pathway, a feedback loop, and an output pathway (Fig. 1.1). At the core of the molecular circadian clock in all eukaryotes there is a set of interconnected transcriptional and translational feedback loops that establish a rhythmic response with a period of about 24 hours. Briefly, a positive regulator activates transcription of a gene that negatively feedbacks on its own expression. Over time the negative regulator accumulates, resulting in repression of its own transcription. Protein turnover in turn decreases levels of the negative regulator, thus activating its expression. The kinetics of this feedback loop are delayed via sev-

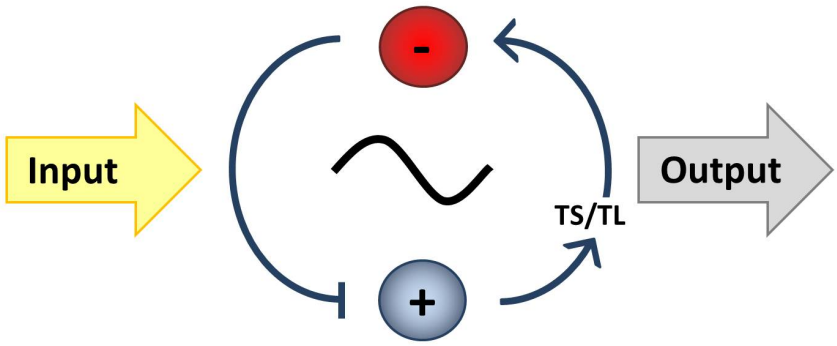

Figure 1.1.: Scheme of the general mechanism of the circadian clock. A Zeitgeber (Input), such as light or temperature, conveys timing information to the endogenous molecular clock. At the core of the feedback loop a transcriptional activator (+) leads to expression of its own negative regulator (-). Transcription and translation (TS/TL) delay this process such that the feedback loop takes about 24 hours. Other genes that are under control of the activator influence diverse functions in the cell (Output).

eral mechanisms to generate a period that approximately reflects the external day of 24 hours. This is achieved in part on a post-transcriptional level by influencing RNA stability and splicing, as well as on posttranslational levels via acetylation, phosphorylation, ubiquitination, and sumoylation [8, 9, 10].

Several input pathways convey timing information to the core oscillator and synchronize it with the environment. Entrainment, i.e. adjustment of the clock to a specific rhythm, for example of light, temperature or nutrients, can reset the clock. Different mechanisms have evolved to sense light, which is one of the strongest cues. Unicellular organisms have developed light-sensitive elements in one or more components of the molecular clock. Some multicellular organisms like zebrafish and *Drosophila* can detect illumination via specific light receptors in each cell of the body. Most vertebrates have one or a small number of specialized cells, which sense light, and relay this timing information to another specialized structure in the brain, the so called master-clock, which then entrains all other clocks of the body, the peripheral clocks, to the exogenous timing regime. When traveling to a different timezone, for instance, the light of the sun phase-shifts our rhythm from the internal, endogenous time, to the exogenous, local time, which is why we experience the so called jet lag

The core oscillator activates clock-controlled genes that generate the output of the circadian clock by rhythmically modulating many cellular functions, for example via control of key enzymes in metabolic pathways. The level of clock-controlled genes oscillate throughout the day resulting in abundance rhythms of the encoded protein.

In some cases these genes are found to feedback on the clock itself, generating a further loop in the circadian system. This feedback can transmit information about the metabolical state to the core oscillator and lead to more robustness. A major goal in the field of chronobiology is to elucidate which genes are controlled by the circadian clock and how their expression influences other pathways and the clock itself. In higher eukaryotes this is a challenging task, because studies have shown, that there is only little overlap between clock-controlled genes in different cell types and tissues [1] and the mechanisms that convey this tissue-specificity are yet unknown.

1.2. The circadian clock in different organisms

The circadian clock can be found in almost all branches of life. In this chapter a general overview shall be given over input and output of the circadian clock, as well as the key players of the core-loop. Moreover, common methods to analyze the rhythm in these organisms will be described briefly. A detailed description of the circadian clock in mammals, will be given in the following chapter.

1.2.1. Invertebrates

The existence of a circadian clock has first been described in plants. Later, the first successful studies in elucidating the molecular basics of the circadian clock have been performed in invertebrates, namely cyanobacteria, the filamentous fungus *Neurospora crassa* and the fruit fly *Drosophila melanogaster*. Nowadays improved genetic tools enable us to study the molecular basis of the clock in vertebrates as well. Nevertheless, the machinery underlying the circadian clock in invertebrates remains the best understood, and therefore gives interesting insights into the even more complex mechanisms in higher organisms.

1.2.1.1. Cyanobacteria *(Synechococcus elongatus)*

Cyanobacteria are the simplest and the most ancient forms of life that have been found to possess a circadian clock. As phototrophs, the biochemistry and physiology of cyanobacteria must be attuned to the availability of energy from light. To date, cyanobacteria remain the only prokaryotes shown to possess a true circadian clock. *Synechococcus elongatus*, a freshwater unicellular cyanobacterium, has been used extensively for circadian studies because of the wide array of genetic tools available for

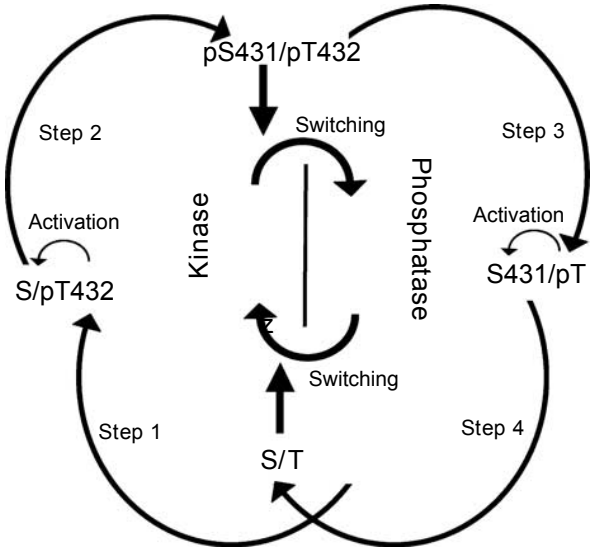

Figure 1.2.: Schematic overview of KaiC phosphorylation/dephosphorylation. The KaiC phosphorylation cycle consists of four sequential reactions. During steps 1 and 2, KaiC exhibits autokinase activity and during steps 3 and 4, KaiC serves as an autophosphatase. Phosphorylation of S431 creates the double-phosphorylated form (pS/pT) and switches KaiC from an autokinase to an autophosphatase, whereas dephosphorylation of S431 leads to the completely dephosphorylated form (S/T), switching KaiC from a phosphatase to a kinase. Phosphorylation of T432 is stimulated by KaiA (step 1) and activates phosphorylation of S431 (step 2). Likewise, dephosphorylation of T432 (step 3) activates dephosphorylation of S431 (step 4) upon which KaiB binds to KaiC, inactivating KaiA. Figure and Legend adapted from Nishiwaki et al. [11].

this organism. Its core oscillator is made up of three proteins: KaiA, KaiB, and KaiC [12] (Kai means "cycle" in Japanese). Both, the abundance and the phosphorylation states, of KaiC cycle *in vivo*. *In vitro*, the three core proteins and ATP are sufficient to produce circadian oscillation in KaiC phosphorylation with a period of about a day [13]. As opposed to other clocks, where the negative regulator is degraded, this system does not require protein synthesis for the subsequent circadian cycle. Rhythmic phosphorylation occurs, because KaiC switches from autokinase to autophosphatase activity upon double phosphorylation. Stimulation of KaiC by KaiA results in autophosphorylation of threonine 432 of KaiC. KaiC subsequently autophosphorylates serine 431. KaiC now has increased dephosphorylation activity and first removes a phosphate at T432 upon which KaiB binds to KaiC. This traps and inactivates KaiA, thus allowing further KaiC dephosphorylation [11] (Fig. 1.2).

Various studies suggest that there is no specific interaction of clock proteins with

the kaiBC promoter in prokaryotes, but rather that KaiC affects all gene promoters in a common genome-wide manner by rhythmically modifying chromosome topology [14, 15, 16]. Furthermore the *S.elongatus* circadian clock has been found to gate cell division, whereas the cell cycle has no effect on the circadian clock [17].

Cyanobacteria sense light via several pathways that inform the cell about its redox state, which correlates with light in a photosynthetic organism, and thereby keep the circadian clock synchronized with the environment [18, 19]. Recent experiments even suggest that KaiA itself can bind the oxidized form of quinone and that its aggregation upon such binding negatively affects KaiC phosphorylation [20].

1.2.1.2. Plants

The possibility for an intrinsic circadian clock has first been discussed in plants. In 1729, the scientist Jean-Jacques de Marain realized that rhythmic petal movement of *Mimosa pudica* was not simply a response to the sun, but persisted even in the absence of sunlight [21]. In the following centuries the intrinsic circadian clock remained a matter of debate, but is now accepted as a fact by scientists.

In *Arabidopsis thaliana* the circadian clock regulates numerous biological processes, such as rhythmic leaf movement and petal opening [22], hormone biosynthesis and response [23], water uptake [24], and even pathogen response [25]. Plants with a clock period that matches the environment contain more chlorophyll, fix more carbon, grow faster, and survive better than plants with circadian periods differing from that of their environment, which shows that the circadian clock confers a photosynthetic advantage to them [26]. Furthermore, plants are able to measure the environmental photoperiod, in order to permit flowering to occur only during the correct season [27]. Circadian regulation of transcription affects roughly one third of the genes in Arabidopsis [28, 23, 29]. The circadian clock in plants is reset by light, which is sensed through phytochromes (red and far-red photoreceptors), as well as cryptochromes [30, 31] and the proteins of the ZEITLUPE (ZTL) family (blue-light receptors) [32]. By controlling their expression the circadian clock is able to modulate susceptibility for light entrainment cues [33].

The plant clock is very complex and made up of several interconnected feedback loops with various key players for each time of the day. The main components will be briefly explained here. In the so called three-loop model (Fig. 1.3) there are a morning loop and an evening loop which interlock with the central core loop [34]. The main players of the central loop are LHY (LATE AND ELONGATED HYPOCOTYL) and CCA1 (CIRCADIAN CLOCK ASSOCIATES 1) which are expressed in the morning and

repress transcription of *TOC1* (*TIME OF CAB EXPRESSION 1*), confining its expression to dusk. Conversely, TOC1, promotes the expression of *CCA1* and *LHY*. In the morning loop, PRR (PSEUDO RESPONSE REGULATOR) 9, PRR7, and PRR5 associate sequentially with the promoter of *LHY* and *CCA1* to repress their transcription and counteract the activating function of TOC1. An evening loop is proposed between an unknown light-inducible component Y and TOC1. Partial Y function was assigned to the evening gene *GI* (*GIGANTEA*) and is mediated by its physical interaction with ZTL, an F-box protein that facilitates ubiquitination and subsequent degradation of both TOC1 and PRR5 [35]. Additional loops connect ELF3 (EARLY FLOWERING 3)-ELF4 and LUX with LHY-CCA1 and PPR9. Recently, rhythmic histone-acetylation changes have been found to be important for clock transcriptional regulation as well [36, 37]

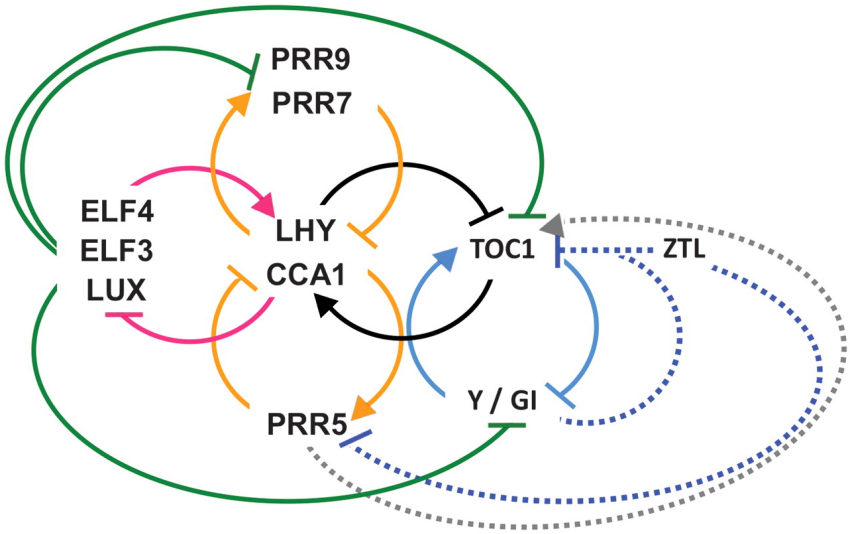

Figure 1.3.: Circadian feedback loops in plants. Circadian clock components are connected by several feedback loops. Central loop (black), morning loop (orange), and evening loop (light blue) conform to the three-loop model. Additional loops connect ELF3-ELF4 and LUX with LHY and CCA1 (pink) and with PRR9 (green). Protein-protein interactions between the evening loop components, ZTL and PRR5, are indicated with dashed lines. The model is further explained in the text. Figure and Legend from Herrero and Davis [35].

Apart from its daily role, the circadian clock also fulfills roles in overwintering of perennial plants that live for more than two years through initiation of growth cessation and dormancy [38] and contribution to cold hardiness [39, 40]. In trees, as in Arabidopsis, cold tolerance is correlated with CCA1 and LHY expression [40]. Current studies are aiming at establishing more precisely how cold regulates CCA1 and LHY expression to provide new tools for breeding of plants that are better adapted to local

climates and for broadening the latitudinal range of crops.

1.2.1.3. Fungi (*Neurospora crassa*)

Neurospora crassa is a filamentous fungus that is characteristically found as first settlers after forest fires on dead organic material. It commonly grows as a mycelium of haploid hyphae that contain multiple nuclei. Vegetative growth is interspersed by formation of bright orange aerial hyphae that segment into the asexual spores called conidia. This process is controlled by the circadian clock, making rhythmic conidiation of *Neurospora* an easy way for following the circadian rhythm of the organism. If applied to one end of a glass tube with an agar-bottom (so called race tubes) *Neurospora* band-strains grow in alternating patterns of vegetative hyphae and conidia along the tube. This enables scientists to study the effect of mutations on period and phase of the circadian clock. This simple way of analyzing the clock contributed significantly to the current understanding of the circadian clock.

The key players of the circadian clock in *Neurospora* are the transcription factors WC-1 (White-Collar 1) and WC-2 (White-Collar 2) and their negative regulator FRQ (Frequency). WC-1 and WC-1 form a dimer, WCC, the White-Collar-Complex, which drives directly and indirectly rhythmic expression of clock-controlled genes [41]. In the morning the WCC is located in the nucleus where it activates transcription of *frq* and other target genes. During the course of the day FRQ protein is produced, dimerizes and assembles with FRH (FRQ-interacting helicase) and CK1a (Casein kinase 1a) into a high molecular mass complex, which is transported into the nucleus. Progressive phosphorylation on up to 113 phosphorylation sites of FRQ changes its three-dimensional conformation, exposes a PEST (proline, glutamic acid, serine, threonine) sequence and results in its degradation [42]. In the nucleus FRQ interacts with the WCC and mediates its phosphorylation via CK1a [43, 44]. The transcriptionally active unphosphorylated WCC is unstable [45], whereas phosphorylated WCC is inactive but stable and accumulates in the cytosol, where it is dephosphorylated by Protein-Phosphatase 2A, which allows a new cycle to begin [45].

Extensive studies on light response and -adaptation have been carried out in *Neurospora crassa*. The filamentous fungus has developed a system that does not react to disturbing illumination cues during the night, such as the moon, and at the same time remains able to sense very low levels of light if necessary. The response to light is mediated by the WCC and the clock-controlled protein VVD (Vivid). Both, VVD and WCC, contain a LOV (Light, Oxygen, Voltage) domain which enables dimerization

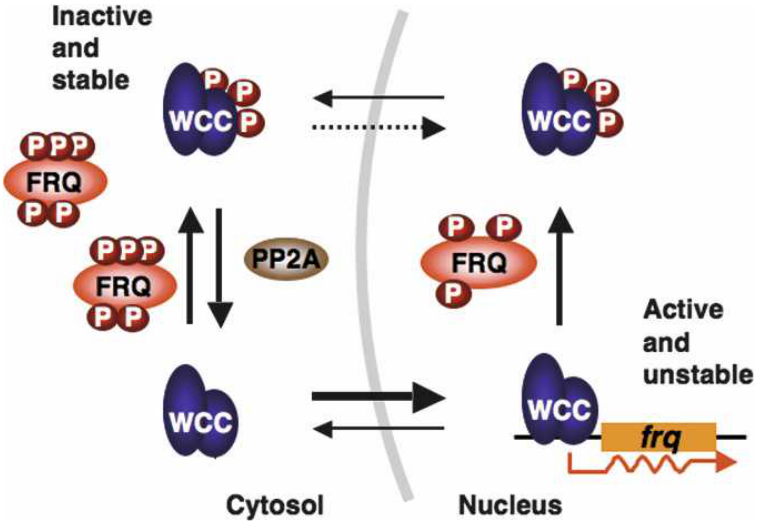

Figure 1.4.: Molecular mechanism of the circadian clock in *Neurospora crassa* as described in the text. Figure from Schafmeier et al. [45].

upon illumination. Light-induced homodimerization of the WCC activates its function as a transcription factor, but when VVD accumulates during the day it inactivates the WCC by competitive formation of WCC-VVD heterodimers. During the night, remaining protein levels of VVD suppress responses to low light cues [46].

ChIP-sequencing of the targets of WCC in *Neurospora* lead to discovery of many new clock-controlled genes, a substantial portion of which are transcription factors themselves. Among these, CSP1 (Conidial Separation 1), a transcriptional repressor, has been found to regulate activation of its target genes in antiphase to the rhythm produced by the WCC [47]. It is a very good example of how a small set of clock-controlled genes can control a larger set of genes in a rhythmic fashion on a genome-wide scale. CSP1 modulates expression of the fatty acid desaturases, and thereby regulates lipid homeostasis of *Neurospora* in the course of the day. During the night, when CSP1 repression is lifted, fatty acid desaturases are produced, leading to a higher proportion of desaturated acids in the membranes, which makes the membranes more fluid - a reaction that counteracts the increased rigidity of membranes in the cold [47].

1.2.1.4. Fruit fly (*Drosophila melanogaster*)

The fruit fly *Drosophila melanogaster* provides an excellent model system for studying the molecular and neurogenetic basis of circadian behavioral rhythms in higher eukaryotes [48]. In 1971 the first clock gene, *period*, was found in *Drosophila* [49]. Many more clock genes followed in the next decades and it turned out that *period* was conserved

in mouse, zebrafish and humans [50, 51, 52, 53]. The simplicity of *Drosophila*'s neuronal network organization in the brain makes it of special use in neurological studies regarding the clock. A so called master clock" or "pacemaker" is located in the brain, that controls behavioral rhythms and can be reset by light. About 150 clock cells per brain hemisphere permit dissection of the function of single cells or cell clusters in the brain [54].

The current model (see Figure 1.5) [54] predicts that the helix-loop-helix transcription factors Clk (Clock) and Cyc (Cycle) bind as heterodimers to E-box sequences (CACGTG) in the promoter regions of their target genes at midday, thereby activating transcription of the genes *per* (*period*) and *tim* (*timeless*). Per appears to be the main player in the negative feedback loop, whereas the function of Tim is mainly to protect Per from phosphorylation by Dbt (Double-Time) Kinase and subsequent ubiquitination by Slimb leading to proteasomal degradation. This can be counteracted by PP2A (Protein phosphatase 2 A) as well. Mediated by CK2 (Casein Kinase 2) the Per-Dbt-Tim complex translocates into the nucleus, where it binds to Clk/Cyc dimers causing hyperphosphorylation of Clk, which prevents the Clk/Cyc dimer from binding to the DNA. As in the clock of *Neurospora crassa*, the negative regulators cause phosphorylation of the transcriptional activators leading to their inactivation. Recently, another clock controlled gene encoding for the basic helix-loop-helix orange-domain putative transcription factor Cwo (clockwork orange) has also been found to repress transcription of Clk target genes [55]. A second feedback loop interlocks with the first by regulating the transcription of *clk*, which is transcribed in a reciprocal way, called "in antiphase", to *tim/per*. Clc/Cyc also activate transcription of *vrille* and *pdp1*. Both proteins bind to so called V/P boxes in the promoter region of *clk*. While Vrille is inhibiting expression of *clk*, Pdp1 activates *clk* transcription.

The 150 neurons per hemisphere of the *Drosophila* brain that make up the master clock, are divided into seven major groups, named after their anatomical position. There are three dorsal neuronal groups (DN1-3), 4 lateral neuronal groups (LNd, l-LNv, LPN, s-LNv), as well as an additional set of a few hundred glia cells [56]. DNs and LPNs seem to be primarily temperature-entrainable, whereas the LNs seem to be preferentially entrainable by light [57].

Drosophila has three different photoreceptive organs, the complex eyes, the ocelli, and the Hofbauer-Buchner eyelet, which all contain different opsin-based photopigments. Two or more additional photopigments exist, that are not restricted to these photoreceptive organs. At least three molecular pathways in *Drosophila* respond to illumination and only when all three input routes are absent, is the circadian system

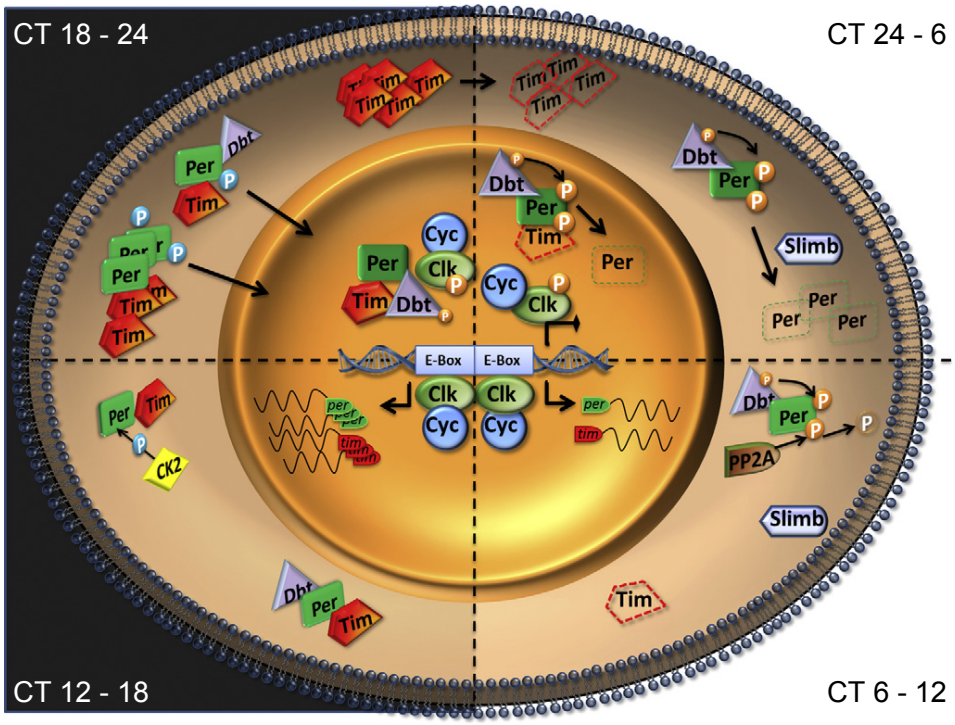

Figure 1.5.: Molecular mechanisms of the *Drosophila* circadian clock. For details see main text. The figure displays a neuronal clock cell at different times of the day. The neuron is compartmented into nucleus and cytoplasm. Each quarter of the cell roughly represents six hours of the circadian day, while the right side (0 - 12) is in light and the left side (12 - 24) in darkness. P indicates phosphate/phosphorylation, the accompanying arrow indicates transfer of phosphate. Dashed arrows or dashed proteins symbols indicate proteasomal degradation. Angular arrows display transcription, sinuous lines mark mRNAs. Normal arrows indicate movement from the cytoplasm to the nucleus. Figure and Legend adapted from Peschel & Förster [54].

unable to react on light [58]. One pathway for synchronization to light occurs via the blue-light photopigment Cry that is expressed in some specialized DN1 cells and in the compound eyes of the fly [59]. Illumination probably induces a conformational change in Cry, enabling it to bind Tim and thus triggering its degradation. As a consequence of the lack of Tim, Period becomes vulnerable to proteasomal degradation as well [54].

Drosophila was found to be a crepuscular insect, an insect that is mainly active at dusk and at dawn. This activity pattern can be measured by putting individual animals into small tubes and measuring the locomotor activity of the fly by a laser or a detector at one end of the tube. The fact that *Drosophila* shows an activity peak in the morning and another one in the evening is difficult to combine with the theory of one circadian clock. Therefore it was proposed that *Drosophila* contains a set of two coupled oscil-

lators, a morning oscillator that is accelerated by light, and an evening oscillator that is slowed down by light [54]. This could enable *Drosophila* to adapt to different photoperiods. The model of the two oscillators is still under debate in the *Drosophila*-field. A recent Letter to Nature [60] is challenging many key laboratory-based assumptions by studies that were carried out under natural conditions. Among other things, they found that under high daytime temperatures *Drosophila* shows a third - afternoon - peak and become rather diurnal than crepuscular.

1.2.2. Vertebrates

The molecular mechanism of the core clock among the various species of vertebrates is relatively conserved and therefore the basics of this mechanism with some minor differences apply to zebrafish, birds, and mammals alike and will be discussed together with the mammalian clock in more detail in the next chapter. For this reason special emphasis will be put onto the differences between the organisms in the following section.

1.2.2.1. Zebrafish (*Danio rerio*)

The zebrafish represents a very useful model for studying key aspects of the vertebrate circadian clock, such as the establishment of the molecular clock during embryogenesis and the light input pathway.

One of the characteristic features of the vertebrate in comparison to the invertebrate clock is the presence of additional copies of key clock genes. In the zebrafish a genome duplication event during the evolution of the teleost lineage has lead to even more additional copies [61]. Some of these have been lost during evolution, others persist. The latter ones either have a redundant function or the multiple functions of one original protein are now performed by several more specialized ones [62]. The most impressive example for this can be found in the *cryptochrome* gene, which exists once in *Drosophila*, twice in mammals, but with six copies in zebrafish, all of these with diverse functions and expression patterns [63].

Remarkably, the zebrafish appears to be the only vertebrate, in which peripheral clocks can be entrained by light, even in cell culture [64]. However, the nature of the peripheral photoreceptor remains elusive to date. Three possible candidates have been proposed. Most likely one of the Cry homologs, Cry4, which shares closest sequence similarity to *Drosophila* Cry, has the ability to act as a light sensor [63]. In *Drosophila* Cry is known to play an important role in photoreception in the brain. Experimental

data also suggests that zebrafish senses the redox state through hydrogen peroxide production by a phototransducing flavin-containing oxidase [65], which would resemble the light input sensing pathway of cyanobacteria. Moreover, extra retinal opsin-like genes have been found in zebrafish, but so far no functional data has confirmed that they indeed serve as photoreceptors [66].

A light responsive module, consisting of closely spaced E- and D-box elements has been found in proximity to the transcription start site of *per2* and is critical for its light-driven rhythmic expression. The D box binding transcription factor, Tef (thyrotrope embryonic factor), was shown to be a key player for mediating this in zebrafish [67]. Interestingly, the *per2* promoter region is highly conserved and Tef is also expressed among vertebrates, although other species like mouse and human lack directly light-entrainable peripheral cells. It is proposed that during the course of evolution ancestral light-responsive mechanisms may have been subverted to signals other than light [62].

There are two main methods how effects on the circadian clock are monitored in zebrafish. Individual larvae can be placed in 24 well microtiter plates and then an automated video analysis system documents swimming activity, which is diurnal in zebrafish. Locomotor activity in adult fish, however, is less robust than in larvae [68, 69]. Alternatively, the rhythmic release of melatonin from the pineal gland can be monitored. Melatonin release is first detected around 24 hours post fertilization and therefore one of the first measurable clock outputs during development [70, 71]. In addition, adult pineal glands continue to rhythmically synthesize melatonin even in culture in a light and clock-dependent manner [72]. Nowadays, luciferase-reporter gene activity under control of a clock-regulated promoter is also used for studying the clock output [73]. This method is quite convenient in zebrafish, because embryos and larvae can be screened alive in a 96 well microtiter plate, since they are completely transparent.

1.2.2.2. Birds

The circadian pacemaking system of birds comprises three major components: the pineal gland, which rhythmically synthesizes and secretes melatonin, a specific hypothalamic region, and the retinae of the eyes. These components work together to produce the circadian output, but their relative contribution to overt rhythmicity appears to differ between species. What is more, the system may change its properties even within an individual, depending, for instance, on its state in the annual cycle or its photic environment [74].

There are some important aspects that appear to function differently in birds. Most

notably, birds possess extraoccular photoreceptors in the pineal gland and brain, primarily within circumventricular structures that can be entrained directly to light-dark cycles. Moreover, several specialized photopigments have been associated with clock function, including pinopsin and melanopsin. Structural analyses of chick CRY2 predicts its photoreceptive capabilities [75]. Furthermore, there are several reports stating that *CRY1* and *CRY2* levels respond to lighting conditions [76].

A very special characteristic of the circadian clock in birds is that it is necessary for navigation and migration. Many experiments have shown that migrating birds use their internal clock in conjunction with the position of the sun to determine their longitudinal localization [77]. Retinal cryptochromes might further be involved in magnetoreception [78]. Apart from navigation, recent work on the night-migratory blackheaded bunding showed that long days induce the migratory phenotype, including body fattening and *Zugunruhe*, as well as testis maturation, suggesting an effect of the circadian clock on seasonal breeding [79] as well.

1.3. Circadian clocks in mammals

1.3.1. Overview

The circadian clock among mammals is quite conserved. The master clock is located in a substructure of the hypothalamus just above the optic chiasm and is hence called the suprachiasmatic nucleus (SCN). The SCN is composed of bilateral nuclei containing approximately 10,000 neurons each. It receives photic input via a novel class of intrinsically light-sensitive retinal ganglion cells that express the photopigment melanopsin and project the information to the SCN via the retinohypothalamic tract [80, 81]. The SCN then relays the information to other regions in the brain and regulates all other rhythmic cells, so called peripheral oscillators, in the body [82] through both, neuronal and humoral signals, which are responsible for proper timing of hormone release, feeding behavior and body-temperature fluctuations [10, 83] as well as other rhythmic functions. Of interest, temporal feeding restriction under light-dark or dark-dark conditions have been able to change the phase of circadian gene expression in peripheral cell types by up to 12 hours, while leaving the phase of cyclic gene expression in the SCN unaffected. Apparently, they thereby uncouple peripheral oscillators from the central pacemaker [84]. Recent findings support a role of PARP1 (poly ADP-ribose polymerase 1) in connecting feeding with the mammalian system [85].

Because of the large set of genetic tools available, the majority of research in mam-

mals focuses on the mouse as a model organism to study the molecular basis of the mammalian clock. One method of choice to monitor the output of the circadian clock is by recording the wheel-running behavior of mice. Moreover, knock-in mice have been engineered to express a luciferase-reporter under control of circadian promoters such as the *Per2*-luc mice. Tissue explants of these have demonstrated that there is a functional molecular clock on the level of almost every type of cell in the body [86]. Depending on the tissue, however, consistent luciferase rhythms can be found for variable amounts of time. While the SCN contains a very robust circadian clock displaying rhythmic luciferase expression for several weeks after explantation, the expression of luciferase in peripheral oscillators dampens within a couple of days. Moreover, peripheral oscillators can be entrained to temperature rhythms, whereas the SCN itself cannot [87].

Circadian studies on humans are mostly carried out by analyzing questionnaires, comparing behavioral rhythms, and by identifying hormone and protein levels from the blood. Nowadays the genetic basis of the human clock is directly explored to an increasing extent by means of cell lines derived from human cells. One of the key cell lines used to study the circadian clock in human cell culture is the U2OS cell line from an osteosarcoma patient.

The core feedback loop in mammals will be discussed in detail in the next section.

1.3.2. Molecular mechanism

1.3.2.1. Main feedback loop

The core circadian clock in mammals consists of several interconnected transcription translation feedback-loops which have been the subject of extensive research in the past decade. The key players of the main regulatory loop are the transcriptional activators CLOCK (circadian locomotor output cycles kaput) [88, 89] or its paralogue in some regions of the brain, NPAS2 (neuronal PAS domain protein 2)[90], as well as BMAL1 and -2 (Brain and muscle arnt like 1) [91, 92, 93] and their negative regulators PER1, -2, and -3 (Period homologue 1, -2, -3) [94] and CRY 1 and -2 (Cryptochrome 1, -2) [95, 96].

The two basic helix-loop-helix transcription factors CLOCK and BMAL1 form a heterodimer, which binds to consensus-E-box-sequences (CACGTG) in the promoter region of their target genes, among these the *PER* and *CRY* genes, thereby activating their transcription [89, 91]. PER and CRY form a dimer in the cytosol where they undergo further posttranslational modifications which regulate their activity and stability.

Eventually, the PER-CRY-dimer binds to and becomes phosphorylated by CK1 (Casein Kinase 1) and is translocated into the nucleus, where it represses its own transcription through interaction with the CLOCK-BMAL1-heterodimer [97, 98]. As a consequence, levels of newly synthesized PER and CRY decrease, while the existing PER-CRY repressor complex is degraded, thus releasing repression and leading to new transcriptional activation by BMAL1 and CLOCK (Figure 1.6).

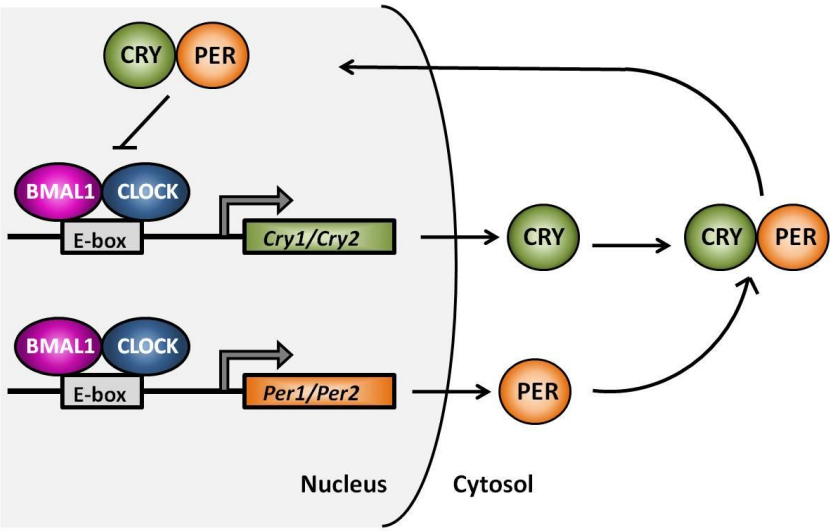

Figure 1.6.: Central negative feedback loop of the circadian clock in mammals. The bHLH-transcription factors BMAL1 and CLOCK bind to E-boxes in their target genes and thereby activate their transcription. Among the target genes are the CRY and PER proteins, which feedback on their own expression by negatively interacting with the BMAL1-CLOCK-heterodimer.

Post-translational modification and degradation are crucial steps for maintaining the period of about 24 hours of the molecular clock. The basic mechanism of transcriptional and translational feedback is maintained, adapted and supported by several other regulatory pathways and interconnected feedback loops, which in many cases are not clearly understood yet. Even the exact mode of activation by BMAL1 and CLOCK and repression by PER and CRY in mammals is still a matter of ongoing research. Experimental data and recent crystal structure analysis of the BMAL1-CLOCK heterodimer suggests that CRY1 binds directly to CLOCK [99, 100] and that PER translocates the inhibitor complex into the nucleus [101]. BMAL1 and CLOCK appear to remain bound to DNA throughout the circadian cycle, where they are rhythmically inhibited, probably via phosphorylation by various kinases and possible subsequent degradation of the BMAL1-CLOCK, which could be replaced by the newer unphosphorylated heterodimer [101, 102, 103]. More recent experiments have been able to show rhythmic binding of

BMAL1 and CLOCK in mouse liver [5, 6].

PER and CRY also undergo several posttranslational modifications. Both, the PER and the CRY proteins are progressively phosphorylated in the cytosol, which, among other things, targets them for polyubiquitination and degradation by the 26S proteasomal pathway. The turnover of PER and CRY is specifically regulated by two different SCF (Skp1-Cullin-F-box) E3 ubiquitin ligase complexes, the β-TrCP1 (beta-transducin repeat containing) [104] and FBXL3 (F-box and leucine-rich repeat protein 3) E3 ubiquitin ligase complexes [105, 106], respectively. PER is phosphorylated by CK1ε (Casein kinase 1ε) and CK1δ (Casein kinase 1δ) [107].

The three PER proteins and the two CRY proteins form a large nuclear complex, the PER complex, which interacts with the BMAL1-CLOCK heterodimer to repress transcription. Other constituents of this complex are CK1ε and CK1δ, as well as six RNA binding proteins. The PER complex RNA-binding proteins PSF (PTB-associated splicing factor), NONO (non-POU domain containing, octamer binding) and WDR5 (WD repeat domain 5) act in transcriptional repression, that is probably achieved in part via recruitment of the SIN3-HDAC (SIN3A-Histone Deacetylase) complex to clock gene promoters, which promotes deacetylation at histones 3 and 4 [108]. The PER complex components DDX5 (DEAD box helicase 5) and DHX9 (DEAH box polypeptide 9) on the other hand, appear to act on coupling between initiation and termination of transcription. The latter one is likely achieved at the 3'-termination site by antagonizing the action of SETX (Senataxin), an RNA helicase that promotes transcriptional termination. This impedes RNA Pol II release and thereby represses transcriptional re-initiation [109].

1.3.2.2. Positive feedback via REVERB and ROR

When the CLOCK-BMAL1 dimer activates transcription of *PER* and *CRY* it also activates expression of the nuclear receptors *NR1D1* and *NR1D2*, encoding for the proteins REVERBα and REVERBβ, respectively. These functionally redundant proteins strongly repress *BMAL1*-transcription and are therefore required for rhythmic *BMAL1* expression [110, 111, 112, 113]. The more PER and CRY protein is produced, the more they inhibit the CLOCK-BMAL1-dimer, resulting in less transcription of *NR1D1* and *NR1D2*. *BMAL1*-transcription is thus de-repressed by lower REVERB protein levels. REVERBα mediates transcriptional repression via NCoR1 (nuclear receptor corepressor 1), which activates the histone deacetylase HDAC3 [114, 115] (Figure 1.7).

Antagonistic to the function of the REVERB proteins is the function of the ROR (RAR-

related orphan receptors) proteins [112, 116]. They contribute to *BMAL1* amplitude but are dispensable for *BMAL1* rhythm [111]. REVERBs and RORs compete for binding to specific ROREs (retinoic acid-related orphan receptor response elements) and repress or activate transcription, respectively [112]. Binding to ROREs was found in the promoter of the *BMAL1* gene as well as in the promoter of *NPAS2* [117] and the first intron of *CLOCK* [118]. This secondary loop is not essential, but it is thought to add robustness to the molecular clock. PER2 has been found to rhythmically interact with promoters of nuclear receptors, such as *NR1D1* and thereby modulate their expression [119]. Cryptochromes, on the other hand, appear to bind to glucocorticoid receptors and repress glucocorticoid-receptor-dependent transcription [120].

1.3.2.3. Chromatin Remodeling

CLOCK has been shown to act as a histone acetyltransferase on histone H3 Lys-14 and Lys-9 [8] at the promoter of its target genes, thereby inducing a transcription-permissive state. Furthermore, it can also specifically acetylate its binding partner BMAL1 on Lys-537. This acetylation facilitates recruitment of CRY1 to CLOCK-BMAL1, thereby promoting transcriptional repression [65]. The NAD+ dependent histone deacetylase SIRT1 (Sirtuin 1) interacts with CLOCK and rhythmically counteracts its acetylation [121]. SIRT1 governs critical metabolic and physiological processes through deacetylation of several regulatory proteins involved in the control of metabolism, including members of the FOXO (Forkhead Box, subclass O) protein family [122], PGC-1α (Peroxisome Proliferator-Activated Receptor Gamma (PPARγ) Coactivator 1A), and the nuclear receptor LXR (Liver X Nuclear Receptor) [123]. CLOCK and SIRT1 contribute to circadian chromatin remodeling at the promoter of *NAMPT* (Nicotinamide Phosphoribosyltransferase) promoter [124] which is a key enzyme in the NAD+ salvage pathway. Thus, SIRT1 contributes to the circadian synthesis of its own coenzyme, NAD+, which tightly couples cellular metabolism and circadian physiology (Figure 1.7). Interestingly, histone acetylation and BMAL1 acetylation occur nearly in antiphase, pointing to a dual role, first in activation of target genes and later in their repression. Other histone-modifying enzymes, such as JARID1a (Jumonji, AT Rich Interactive Domain 1a), also seem to influence circadian transcription. JARID1a forms a complex with CLOCK-BMAL1, which is recruited to the *Per2* promoter. It increases histone acetylation by inhibiting histone deacetylase 1 function and enhances transcription by CLOCK-BMAL1 in a demethylase-independent manner [125].

A recent profile of the epigenetic landscape in mouse liver [6] showed that on a

genome-wide scale there are rhythms in RNA polymerase II recruitment and initiation, as well as in histone H3 trimethylation (H3K4me3) and acetylation (H3K9ac, H3K27ac), which occur at thousands of expressed genes whether or not RNA cycling was detectable. It was suggested that circadian transcriptional regulators appear to be involved in the initial stages of RNA polymerase II recruitment and initiation and that the histone modifications associated with these events set the stage for gene expression on a global scale, but that additional control steps must then determine the ultimate transcriptional outputs from these sites.

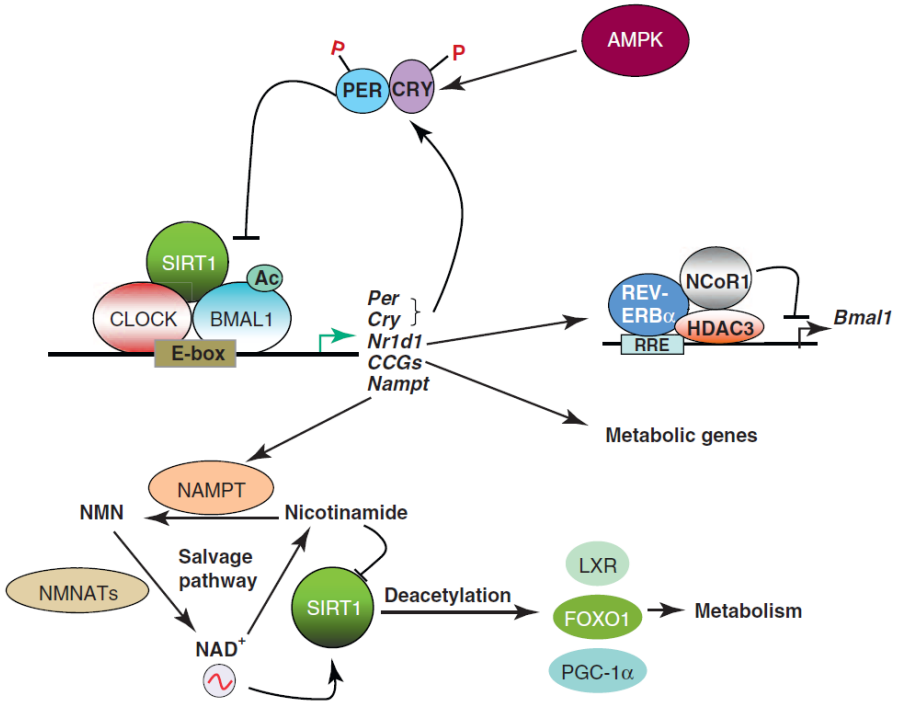

Figure 1.7.: Interplay between regulators of circadian clock and metabolism. Some clock controlled genes are regulators of metabolic pathways, and others can influence metabolite levels. Nutrient sensors AMPK and SIRT1 are regulated in a circadian manner, and they can modulate circadian rhythms and metabolism by post-translational modifications of key regulators. Figure and Legend from Sahar and Sassone-Corsi [126].

1.3.2.4. Link to metabolism via AMPK and the PPAR proteins

Both, CRY1 and Casein kinase 1ε, have recently been found to be phosphorylated by AMPK (AMP-activated protein kinase), which leads to increased degradation of CRY1 [127] and PER2. This mechanism allows nutrient signals to feed back on the core circadian clock because AMPK is activated when the ratio of AMP/ATP is high, signaling

that the overall cellular energy status is low (Figure 1.7). Moreover, the activity of AMPK was found to be rhythmic in mouse liver, hypothalamus, and fibroblasts [127, 128].

Another important link between circadian clocks and metabolism has been established in recent years based on the nuclear receptors from the family of the peroxisome proliferator-activated receptors (PPARs). Three subtypes of PPARs are known: PPARα, PPARδ, and PPARγ. The expression of the different PPARs exhibit characteristic but distinct tissue-specific rhythms in mouse. *PPARA* transcripts cycle in both, white and brown adipose tissue and in liver but not in muscle. It is known for its role in promoting hepatic fatty acid oxidation and ketogenesis in response to fasting. *PPARG* selectively cycles in white adipose tissue and liver and is most abundant in adipose tissue, where it activates transcriptional programs for lipid storage and lipogenesis, whereas *PPARD* transcripts oscillate only in brown adipose tissue and liver [129].

Not only are the PPARs expressed in a rhythmic manner, they also provide a feedback on expression of important core clock genes themselves. PPARγ directly induces expression of *NR1D1* in adipose tissue of mice by binding to its promoter [130]. On the other hand PER2 was shown to directly regulate PPARγ in white-adipose tissue by blocking PPARγ recruitment to target promoters and thereby inhibiting transcriptional activation. This mechanism is conceptually different from the repression that PER2-CRY exerts on CLOCK-BMAL1 and does not require CRY [131]. The coactivator of PPARα, PGC-1α, provide further feedback to the circadian network. *PGC1A* was found to be rhythmically expressed in liver, brown adipose tissue, and skeletal muscle of mice [129, 132]. It stimulates the expression of clock genes, notably *BMAL1* and *NR1D1*, through coactivation of the ROR family of orphan nuclear receptors [132].

1.3.2.5. Fine-tuning via the DEC proteins

The clock-controlled proteins DEC1 (Differentiated embryo chondrocyte 1) and DEC2 play a role in the fine-tuning of the circadian clock. Both proteins are bHLH transcription factors like BMAL1 and CLOCK and suppress BMAL1-CLOCK-enhanced promoter activity through competitive binding to canonical E-boxes (CACGTG) [133], thus influencing the phase of expression of genes that contain this motif in their promoter region, such as *DEC1* and *-2* themselves, as well as *PER1*, *DBP*, and *NR1D1*. But unlike BMAL1 and CLOCK they do not bind to E'-boxes (CACGTT) and therefore do not affect the phase of *PER2* and *CRY1* [134].

1.3.2.6. D-box regulation via the bZIP proteins

There are four proteins of the bZIP (basic leucine zipper) family that are regulated in a circadian fashion and affect circadian output target genes by interacting with D-box elements in the promoter region. The PAR (proline and acidic amino acid-rich) bZIP proteins DBP (albumin D-site-binding protein), HLF (Hepatic leukemia factor), and TEF (Thyrotroph embryonic factor) are highly rhythmic transcriptional activators in several tissues, including liver, SCN and kidney [1, 135, 136]. The rhythmic bZIP transcription factor *E4BP4* (E4 promoter-binding protein) is expressed in antiphase to *DBP, HLF* and *TEF* and functions as a negative regulator of D-boxes. Whereas the positive activators are able to dimerize, E4BP4 is not, but it can compete for the same binding sites with them [137]. The PAR bZIP transcription factors modulate the expression of many proteins involved in detoxification and drug metabolism [138]. Cellular DBP and E4BP4 proteins are critical for determining the circadian period length and have reciprocal effects on it [139].

1.3.3. Physiological rhythms and pathophysiology

The circadian clock influences important physiological processes, such as the proper timing of hormone release, feeding behavior, cardiovascular- and renal activity, the hepatic metabolism, and body-temperature fluctuations [10, 83, 140]. The SCN is the master pacemaker that senses light and relays timing information to peripheral clocks in the body via different pathways. Diurnal body-temperature fluctuations have been found to entrain peripheral oscillators, but not the SCN [87]. Recent analysis of cycling transcripts in mice with and without a functional liver clock showed, that rhythmic transcription of most genes in this organ depended on a functional hepatocyte clock, whereas a small set of other genes were circadianly expressed even in the absence of a liver clock, among these *Per2*. Rhythms in these genes were probably generated by systemic signals [141] and could provide another mechanism by which peripheral oscillators are entrained by the SCN. Moreover, several pathways have developed to also allow a feedback of information about the metabolic state of a cell to its local circadian system. An example for this is the finding that peripheral oscillators can be uncoupled from the central pacemaker by restricted feeding [84]. In order to establish communication between circadian clocks and metabolism, sensors affecting both systems exist. Among these are redox sensors, AMP/ATP ratio sensors, glucose sensors, and fatty acid sensors [83].

The entire circadian system has evolved to optimize an organism's performance over

the course of the day and to maintain homeostasis by keeping various physiological processes in balance. When this equilibrium is lost, for example by rotating shift work, it can increase the risk of developing severe diseases. Shift-work and sleep deprivation have been found to dampen rhythms in growth hormone and melatonin secretion, reduce insulin sensitivity, and elevate circulating cortisol levels. These changes favor weight gain, obesity, and development of metabolic syndrome [142]. Along the same line, loss of adequate blood pressure decline during the sleep period has been found to be a significant risk factor for cardiovascular organ damage [143].

Hamsters with a point mutation in casein kinase-1ε, the so called *tau* mutation, have a 22-hour rather than a 24-hour circadian cycle. When kept under a 24-hour cycle (light/dark, 12 h/12 h) these hamsters developed cardiomyopathy and severe renal disease which was reversed when they were kept under a 22-hour cycle, the length of their intrinsic rhythm, or when the SCN was ablated. This impressively shows, that not the disruption of the circadian clock, but the desynchrony between the internal and external clock were responsible for the development of cardiovascular disease [144]. Similar effects were observed in mice [145].

Probably the most obvious output of the circadian clock in humans is the rhythmic pattern of sleep and wakefulness, the timing of which is regulated by two processes: a sleep homeostatic process that underlies the rise of need for sleep during wakefulness and a circadian process that determines the threshold for switching between sleep and wake. Several sleep disorders have been clearly linked to circadian alterations in the timing of sleep such as the advanced sleep phase syndrome and the delayed sleep phase syndrome. FASPS (Familial advanced sleep-phase syndrome) was the first human sleep disorder, where a direct link to a known core clock gene was found. FASPS is inherited in an autosomal dominant fashion and is characterized by persistent 3-4 hours advanced sleep onset- and waking times. This causes affected individuals to not be able to function on the conventional time scale in society and work-life, often leading to depression. FASPS results from the missense mutation S662G in PER2, which coincides with a phosphorylation site within its CK1 binding domain. The missing phosphorylation site appears to cause a faster turnover of nuclear PER2, either by enhanced degradation or by reduced nuclear retention of PER2 [146, 147].

1.3.4. Relation between clocks, cell cycle, and cancer

In the last decade many connections of various core clock proteins to cell cycle control and the development of cancer have been found. The finding that disruption of

circadian rhythms led to increased mammary tumor development was first reported in the 1960s [148]. Since then, several epidemiological studies have pointed to an effect of a de-regulated circadian clock on the development of cancer. Among these are studies that show that female shift-workers have a higher predisposition for developing cancer, which increases with the number of years and hours per week that individuals spent working at night [149, 150]. Moreover, disruption of circadian rhythms appears to accelerate tumor progression, which was studied in SCN-ablated mice [151]. There is direct evidence that demonstrates an association between loss of human PER1, PER2, and PER3 function and breast- and endometrial cancer. One study showed that 95% of breast tumors display no or deregulated level of PER1, PER2, or PER3 proteins compared to adjacent normal tissues [152].

On the molecular level the BMAL1-CLOCK heterodimer appears to control the expression of many genes that are involved in the cell cycle. Several reports show rhythmic expression of major proto-oncogenes like *c-MYC* (*v-myc myelocytomatosis viral oncogene homolog (avian)*) and *MDM2* (*Mouse Double-Minute*), as well as tumor-suppressor genes like *P53*, and *GADD45*α (*Growth Arrest And DNA-Damage-Inducible, alpha*) [1, 153, 154]. The *c-Myc* promoter appears to be suppressed by BMAL1-CLOCK, whereas CRY1 relieves repression [153]. WEE1 kinase, which regulates the timing of G2/M transition [155] and *METAP2* (*Methionyl Aminopeptidase*) which is important in cell cycle initiation in eukaryotes [156, 157] are also directly controlled by BMAL1 and CLOCK. PER1 was found to interact with both, ATM (Ataxia-Telangiectasia Mutated) kinase and CHK2 (Checkpoint Kinase 2), which are involved in the cellular response to ionizing radiation and DNA double-strand-break-inducing events [158]. REVERBα, in turn, has been found to regulate expression of the cell cycle inhibitor *p21*$^{WAF1/CIP1}$ [159]. The protein NONO that is bound in the PER inhibitor complex acts together with PER to rhythmically activate the promoter of the cell cycle checkpoint gene *p16-INK4A*, which regulates G1/S transition [160].

Although CRY1 and CRY2 have structural similarity to photolyases, proteins involved in the repair of UV-damaged DNA, biochemical evidence of participation in DNA-repair is still missing [161]. A study on *Cry1*-/- and *Cry2*-/- mice even reports no effect of either *cry* knockout on spontaneous and ionizing radiation-induced cancers [162].

All these observations lead to the assumption, that there is a common pathway shared by regulators of the circadian clock and cell proliferation. This pathway need not to be shared at all the times, but may come into play in stress-situations only, like after irradiation-induced DNA-damage, and therefore be opportunistic and episodic rather than operating at all times [161].

1.4. Binding motif of BMAL1 and CLOCK

The core transcriptional activator in mammals is a heterodimer of BMAL1 and CLOCK, which both are basic helix-loop-helix (bHLH) PAS domain proteins. There are more than 125 different bHLH proteins [163], which are classified by sequence in A - F. BHLH proteins can form homo- and heterodimers mediated by their HLH domain [164] and, in case of class C, as for BMAL1 and CLOCK, also by their PAS-domain [100]. At the amino-terminal end of the bHLH region is the basic domain, which binds the transcription factor to DNA at a consensus hexanucleotide sequence known as the E-box. Different families of bHLH proteins recognize different E-box consensus sequences and dimerization with different partners can shift affinities [163]. BMAL1 and CLOCK were found to bind to E-box motifs with the consensus sequence CACGTG (allowing up to one mismatch) or double E-box motifs containing two E-box sequences (with up to 3 mismatches total) in tandem with a short interval of six nucleotides between them [5, 6, 165]. Very recently, it was suggested that BMAL1 and CLOCK bind to a 7-bp sequence rather than a 6-bp motif, and that two noncanonical E-boxes were bound with high affinities, consisting of the sequences AACGTGA and CATGTGA [166]. The former motif has been observed as a 6-bp E'-box CACGTT to account for most of the transcriptional drive in the *mPer2* promoter [167].

1.5. Project Aim: Analysis of binding of the circadian transcription regulators in human cells

Many cellular pathways are influenced by the circadian clock and an important aim of the field of chronobiology is to find out which genes or proteins the circadian clock directly acts on. Mouse experiments suggest that around 10% of all actively transcribed genes in mammals could be clock-controlled and that they are highly tissue specific [1]. The molecular basis of tissue specificity of circadian gene expression, however, is not yet understood. Many published works have therefore focused on identifying oscillating mRNA expression in various organs through microarray experiments using mouse tissue extracts [1, 3, 168, 169, 170]. In recent years, our group, as well as some other groups, have used a new approach for finding direct targets of the circadian clock on a genome wide scale in various organisms [5, 6, 171, 172]. This approach is based on the combination of Chromatin-Immunoprecipitation (ChIP) of the core transcriptional activator(s) of the circadian clock, followed by deep sequencing, so

called ChIP-sequencing or ChIP-seq. Using this method rhythmic binding of BMAL1 has been observed in 2047 target genes in mouse liver [5]. Most studies so far have been performed in mice, because tissue extracts of wild type and knockout animals under controlled conditions are readily available. Lately, advances in studying the human clockwork have been made due to the fact that some immortal cell lines contain a functional circadian clock which can be reset by stimuli like dexamethasone or entrained to temperature rhythms. One cell line that has been used in various circadian studies is the U2OS cell line derived from human osteosarcoma cells [2, 3, 4] with which all experiments in this work have been performed.

A stringent genome-wide transcriptome analysis in U2OS did not find any rhythmic transcripts besides seven core clock genes [3]. In this work we want to analyze whether the core circadian transcription factors bind to and regulate fewer genes in U2OS than in mouse liver. Therefore we expanded the ChIP-seq approach to find genes that are bound by three key players of the mammalian circadian clock, BMAL1, CLOCK, and CRY1, and combined ChIP-seq binding site analysis and motif search in U2OS cells with microarray data of a two-day RNA time course in the same cell type. With this data at hand, further analysis was carried out to gather insights into the functionality of the obtained binding sites by producing U2OS cells that stably expressed promoter-luciferase constructs of promoters that contained or lacked binding sites for BMAL1 and CLOCK. The combination of these experiments increases the insight into the circadian clock of U2OS cells and poses a hypothesis for tissue specificity.

2. Material and Methods

2.1. Material

2.1.1. Organisms

2.1.1.1. Cell lines

Most experiments were performed with cell lines derived from the human osteosarcoma cell line U2OS. These cells express a functional circadian clock and are therefore used as a model system for clocks in humans. J. Ponten and E. Saksela derived this line (originally called "2T") in 1964 from a moderately differentiated sarcoma of the tibia of a 15-year-old girl. The cultures were started from a non-necrotic intraosseous part of the tumor obtained at the time of amputation of the left leg. This cell line was first found to be hypodiploid [173]. The culture collection ATCC (American Type Culture Collection) received the U2OS cell line in 1982 from the Human Tumor Cell Bank at passage number 28. ATCC found that U2OS cells are not hypodiploid but contain on average 76 chromosomes per cell and that normal chromosome representation was low. Most of the chromosomes were monosomic and 22 stable marker chromosomes had formed by rearrangement of several of the missing chromosomes, particularly of chromosomes 1, 7, 9, and 11. The U2OS cells used in this work were received at passage number 28. It is uncertain whether the described rearrangements can be found in these U2OS, because changes appear to be quite frequent.

In a study comparing 6 different osteosarcoma cell lines, U2OS had the lowest level of chromosomal numerical variations and showed the most normal pheno- and genotype concerning the genes *P53*, *RB1*, and *MYC* [174]. Although the *MYC* gene was 5 times amplified, induction upon serum-stimulation appeared to function normally. U2OS were the only osteosarcoma cells with a non-mutated *P53* [175]. Therefore, even though it is a tumor cell line, it seems to be a comparably good system for studying circadian rhythms in humans and it allows researchers to use genetic tools in a human cell line. So far several studies have been carried out, including knock-down

studies [2, 4, 176], experiments with rhythmic *BMAL1*-promoter-luciferase constructs [4], as well as studying rhythmic expression of genes on a genome-wide scale [3].

While tissues in the body are constantly synchronized by the SCN, this synchrony is lost over time in all explanted tissues, a phenomenon called dampening. When working with cell lines it is therefore necessary to be able to reset the clock. This is often performed by adding high amounts of serum for a short amount of time [7] or by using the glucocorticoid dexamethasone [3, 177]. Recent studies have shown that some cell lines, such as NiH-3T3 cells [146] and U2OS cells (unpublished data) can be entrained to temperature rhythms as well, providing a convenient method to obtain a consistent rhythm without upsetting the system by addition of metabolites.

In this work single cell clones of U2OS cells stably expressing promoter-luciferase constructs were generated using lentiviral transduction or transient transfection and selection. For lentiviral transduction a modified U2OS cell line, $U2OS_{mcat}$ (generated by Dr. Claudia Seelenmeyer), was used as target cell line. This cell line was selected with the antibiotic zeocin to stably express the murine MCAT (Murine Cation Transporter) receptor on its surface [178], thus allowing virus entry into a human cell line via a mouse-specific protein from the pV Pack Eco System. The stable cell lines that have been used for this work are listed in Table 2.1. U2OS *BMAL1-luc* was described previously [4].

Table 2.1.: U2OS single cell clones that stably express promoter-luciferase constructs. From each cell line several single cell clones were grown and used. The single cell clones used to obtain the data shown in the results-section are named first. In parenthesis are other clones that produced good rhythms. *GAPDH*-luc cell lines were not rhythmic.

Cell line	plasmid	luciferase	promoter	Single cell clone
$U2OS_{mcat}$	pSIN_MCS	*luc2P*	*ATG3*	B11 (E6, E8)
$U2OS_{mcat}$	pSIN_MCS	*luc2P*	*EIF5A2*	E8 (E10)
$U2OS_{mcat}$	pSIN_MCS	*luc2P*	*GAPDH*	G3 (C7)
$U2OS_{mcat}$	pSIN_MCS	*luc2P*	*SCN5A*	E8 (G7)
U2OS	pGL4.20	*luc2*	*CRY1*	E2 (E3, G7)
U2OS	pGL4.20	*luc2*	*NR1D1*	D5 (C7, C9)
U2OS	pGL4.20	*luc2*	*PER2*	E4 (B8, B7)
$U2OS_{mcat}$	pSIN_MCS	*luc2*	*PER2*	B9 (D7)

HEK (Human Embryonic Kidney) 293T cells are the highly transfectable derivative of the 293 cell line into which the temperature sensitive gene for SV40 (Simian Virus 40) T-antigen was inserted. These cells have been developed specifically for the purpose

of high transfectability and their capacity of producing high titer of infectious retrovirus. In this work they were used to produce the virus containing the promoter-luciferase constructs of interest using the pV Pack Eco System.

2.1.1.2. Escherichia Coli

Escherichia coli (*E.coli*) is a Gram-negative, rod-shaped bacterium from the family of Enterobacteriaceae of the Gammaproteobacteria. It is commonly found in the lower intestine of endothermic organisms and forms part of the gut flora. It is about 2 µm long and 0.5 µm in diameter. *E.coli* was one of the first organisms to have its genome sequenced [179]. Due to its rapid division rate of 20-30 minutes in culture and its well known and simple genetics it has become a widely used model organism. Apart from this function, it is used ubiquitously as a host and propagator for DNA plasmids as well as for protein overexpression. In this work the competent *E.coli* strains DH5α (Invitrogen) and XL-1 blue (Agilent Technologies) were transformed for propagation of plasmids.

2.1.1.3. Murine leukemia virus

Murine leukemia viruses (MLV) belong to the γ-retroviral genus of the Retroviridae family. Retroviruses are RNA viruses that replicate through an integrated DNA intermediate. Retroviral particles encapsidate two copies of the full-length viral RNA, each copy containing the complete genetic information needed for virus replication. Retroviruses possess a lipid envelope and use interactions between the virally encoded envelope protein that is embedded in its membrane and a cellular receptor to enter a host cell. After entering a host cell the virally encoded enzyme Reverse Transcriptase produces a DNA-copy of the viral RNA using reverse transcription. Viral DNA is permanently integrated into the host genome by the virally encoded integrase and is referred to as provirus. The transcription and translation machinery from the host cell carries out expression of the viral genes. Newly synthesized viral proteins and full-length RNAs are assembled together to form new viruses that bud out of the host cells (Figure 2.1) [180].

MLV encodes the retroviral genes *gag* (*group-specific antigen*), *pro* (*protease*), *pol* (*polymerase*), and *env* (*envelope*). The *gag* gene codes for the Gag polyproteins that make up the viral capsid. After assembly of the virus particle, this polyprotein is proteolytically cleaved into several proteins. The *pol* gene encodes Reverse Transcriptase as well as Integrase. The *env* gene codes for the Envelope polyprotein, which is also

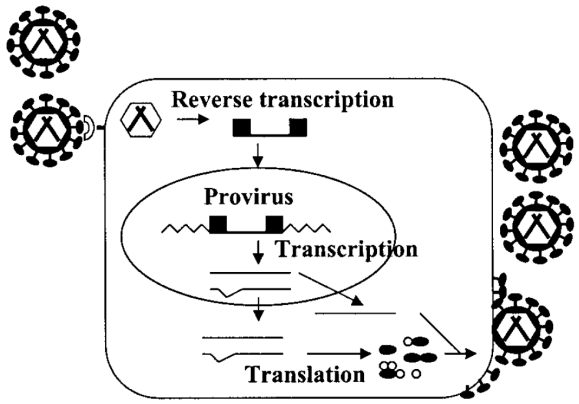

Figure 2.1.: Life-cycle of a retrovirus, as described in the text. Figure from Hu and Pathak [180].

cleaved into two parts. Viral DNA contains large redundant sequences at both ends of the genome, designated long terminal repeats (LTR). These can be further divided into U3 (unique 3'), R (repeat), and U5 (unique 5') regions. Viral promoters and enhancers are located in the U3 region. The R region is essential for reverse transcription and replication of all retroviruses. In addition MLV also contains a polyadenylation signal. The U5 region contains sequences that facilitate the initiation of reverse transcription [180]. The viral packing signal (Ψ) is required for the viral RNA to be packaged into virion particles. Sequences extending into the *gag* open reading frame can enhance the efficiency of packaging and are called $\Psi+$ [181].

2.1.2. Plasmids

2.1.2.1. pGL4.20

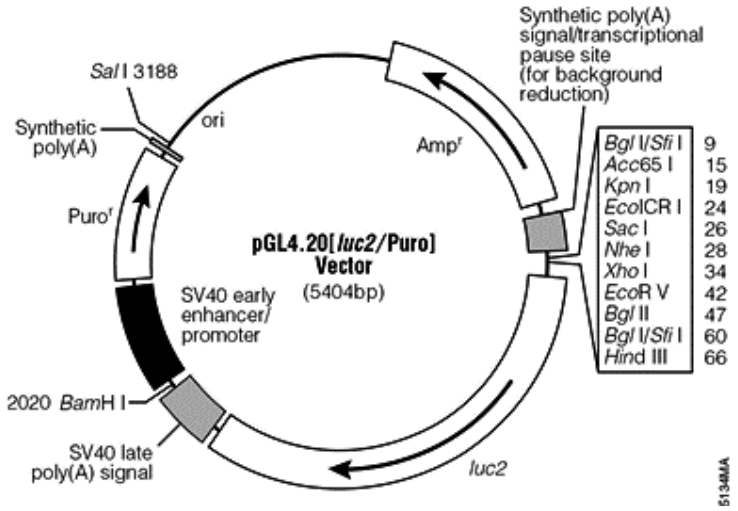

Figure 2.2.: Vectorcard of pGL4.20 as provided by the manufacturer (Promega). The vector encodes for *luc2* and contains a puromycin resistance for selection of stable cell lines. Promoter constructs were cloned into the vector according to Table 2.2.

The plasmid pGL4.20 (Figure 2.2) was obtained from Promega. It contains a synthetic firefly (*Photinus pyralis*) luciferase2 (*luc2*), which has been codon-optimized for more efficient expression in mammalian cells. Moreover pGL4.20 contains a puromycin resistance for selection of stable cell lines. The vector neither contains promoter nor enhancer elements. The promoter regions of human *CRY1*, *NR1D1* and *PER2* were amplified from U2OS gDNA and cloned into the multiple cloning site (MCS) of this vector. Information on the promoter and the restriction enzymes that were used are listed in Table 2.2.

Table 2.2.: Promoter-luciferase constructs were generated with the promoters of the known clock-controlled genes *CRY1*, *NR1D1*, and *PER2*. Information on promoter-regions and flanking restriction sites are shown.

Gene	Region (bp from TSS)	Length (bp)	Restr.Enzymes
CRY1	-1107 to +799	1906	NheI, HindIII
NR1D1	-858 to +625	1483	NheI, BglII
PER2	-848 to +204	1053	SacI, HindIII

2.1.2.2. pSGG_prom

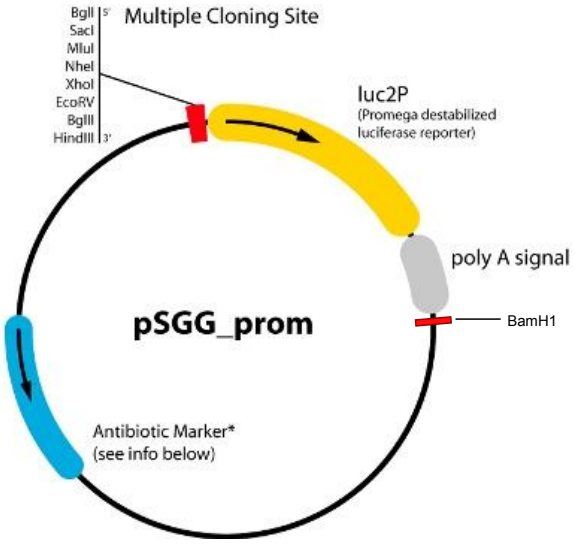

Figure 2.3.: The plasmid pSGG_prom as provided by the manufacturer (SwitchGear Genomics). Constructs were cut using MluI and BamHI. The antibiotic marker was ampicillin.

The promoter-luciferase constructs for *ATG3-luc2P*, *EIF5A2-luc2P*, *GAPDH-luc2P*, and *SCN5A-luc2P* were purchased from SwitchGear Genomics and delivered in the plasmid pSGG_prom with the information provided in Figure 2.3 and Table 2.3. The vector is derived from the pGL4 family of vectors (Promega) and contains a multiple cloning site, *luciferase2P* (*luc2P*), and an ampicillin resistance for propagation in E.coli cells. Luc2P is a modified version of luc2 harboring a destabilizing c-terminal PEST sequence, reducing its half-life from 3 h to 1 h. Destabilized reporter proteins are expected to be more responsive and better suited to monitor rapid processes than those with slower degradation rates. In circadian biology their advantage is that they provide a better signal-to-noise ratio. The entire promoter-luciferase2P region was cloned into the vector pSIN_MCS after digestion with the enzymes MluI and BamHI.

Table 2.3.: Information on promoter-*luc2P* constructs as provided by Switch Gear Genomics.

productID	gene	chr	start	end	strand	length (bp)
108396	*ATG3*	chr3	113763005	113763997	-	992
106995	*EIF5A2*	chr3	172109014	172109967	-	953
121624	*GAPDH*	chr12	6513163	6514226	+	1063
109730	*SCN5A*	chr3	38665940	38666935	-	995

2.1.2.3. pSIN_MCS

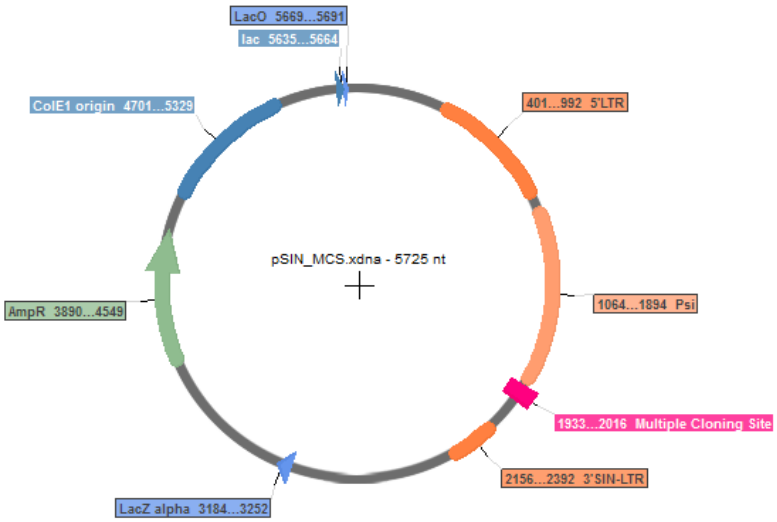

Figure 2.4.: Vectorcard of pSIN_MCS. The vector contained the retroviral elements 5'-LTR and 3'SIN-LTR and the viral packaging signal $\Psi+$ (orange). For cloning, two restriction sites in the MCS (pink), MluI and BamHI (light blue), were used. The plasmid contains an ampicillin-resistance (green).

This vector is an adapted version of the pSIN.Bi-F3-lg-F plasmid which was developed by Dr. Ina Weidenfeld [182] and modified by Christian Nahstoll. The pSIN.Bi-F3-lg-F plasmid is a derivative of the self-inactivating (SIN) Moloney Murine Leukemia Virus-based retroviral vector S2f-IMCg [183] and can be used for retrovirus production. SIN vectors distinguish themselves from conventional γ-retroviral vectors by deletions within the U3 sequence of the 3'-LTR, which covers the LTR promoter and enhancer region. Upon reverse transcription and genomic integration, the γ-retroviral vector will carry the SIN-LTR at both, the 3'- and the 5'-end [184]. Thus, transcription of a gene of interest inserted between both LTRs is not regulated by the strong viral promoters but by a promoter which must be specifically inserted along with the target gene between the LTR regions. In the pSIN_MCS a 3.5 kb large fragment of the pSIN.Bi-F3-lg-F, containing a Tet promoter, *luciferase*, and *GFP* was substituted with a multiple cloning site which was used for cloning promoter-luciferase constructs from the pSGG_prom into the pSIN_MCS vector using the restriction enzymes BamHI and MluI.

The retroviral elements in this plasmid include the extended version of the viral packaging signal $\Psi+$, the 5'-LTR, and the 3'SIN-LTR [182] (Figure 2.4).

2.1.3. Antibodies

2.1.3.1. Peptide antibodies

For this work four peptide-based antibodies against the proteins BMAL1, CLOCK, CRY1, and CRY2 (Table 2.4) were generated by Pineda Antikörper-Service. Three-month-old rabbits were repeatedly immunized with specific peptides and blood was collected every 30 days. The sera were tested for specificity using western blot, beginning 60 days after the first immunization. The peptides contained a cystein at the n-terminus, which allowed purification of the antisera through peptide-bound columns via disulfide-bridges. Antibodies were tested for specificity against the peptide on western blots with U2OS total protein extracts (Figure 2.5).

Table 2.4.: Four peptide-antibodies were generated against the core clock proteins BMAL1, CLOCK, CRY1, and CRY2. All, except for CRY2 were used for ChIP-sequencing. For each antibody two rabbits were immunized (R1, R2). The antibody with the highest affinity was purified from serum at the indicated number of days after first immunization (d150, d230).

Target	Designation	Animal and day	Peptide
BMAL1	BMAL1 c-term	R2 d230	CLEADAGLGGPVDFSDLPWPL
CLOCK	CLNT	R2 d150	CIFDGLVEEDDKDKAKRVS
CRY1	CRY1 c-term	R1 d150	CQEEDTQSIGPKVQRQSTN
CRY2	CRY2 c-prox	R2 d150	CLEAAEEPPGEELSKRARVAE

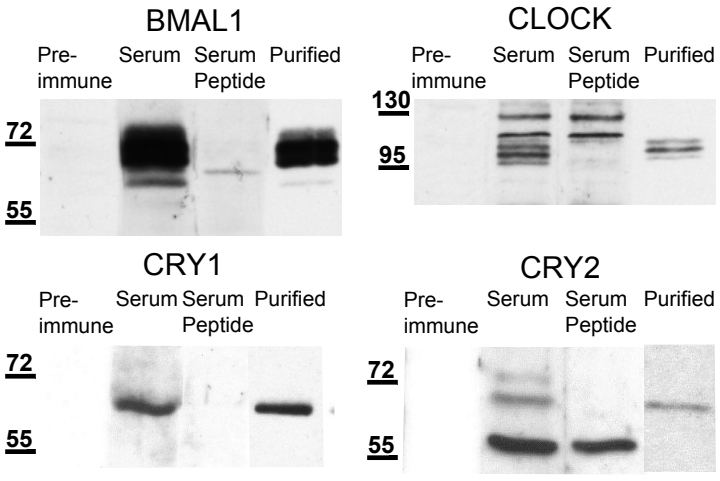

Figure 2.5.: Western blots showing specificity of the generated peptide-antibodies. Antibody-specificity was analyzed with U2OS total extracts, comparing serum (1:1000), peptide-blocked-serum (serum: 1:1000, peptide: 10 µg/ml), and affinity-purified antibody (BMAL1 - 1:750, all others: 1:300).

2.1.3.2. Commercial antibodies

The secondary goat-anti-rabbit antibody, conjugated to horseradish-peroxidases is commercially available from BIORAD.

2.2. Methods

2.2.1. Collaborations

Bioinformatical analysis of the data from ChIP-seq and microarray experiments were carried out in collaboration with Prof. Felix Naef and Laura Symul at the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Descriptions of the analyses were adapted from Laura Symul.

Microarray design and analysis was done in collaboration with Dr. Tamás Fischer, Biochemistry Center of the Heidelberg University, Germany.

FACS sorting was carried out by Dr. Monika Langlotz, Center for Molecular Biology of the Heidelberg University (ZMBH), Germany.

HEK 293T, U2OS, and U2OS *BMAL1-luc* were a kind gift of Prof. Achim Kramer, Charité Berlin, Germany.

2.2.2. Cultivation of mammalian cells

2.2.2.1. Growing and passaging of cells

For all experiments which did not require temperature entrainment, U2OS cells were grown at 37°C with 5% CO_2. Cells were cultured in DMEM (Dulbecco's modified Eagles medium) with 4.5 g/L glucose, supplemented with 10% (v/v) FCS (Fetal Calf Serum) and Penicillin/Streptomycin (Final conc.1x). HEK cells and U2OS cells, as well as its derivatives, grow adherent in monolayers on the surface of non-coated cell culture dishes.

Protocol: Cells were washed once with PBS and then incubated with 0.9 ml Trypsin-EDTA for 8-10 min at 37°C. The serine protease trypsin interferes in cadherin-mediated cell-cell adhesion. EDTA (Ethylenediaminetetraacetic acid) sequesters Ca2+ ions, making the cadherins susceptible to trypsin. Cells were taken up in 10 ml growth medium and resuspended by pipetting. For U2OS cells 1 ml was kept and the rest was

discarded, whereas only 0.5 ml were kept from the fast-growing HEK cells. This way cells were passaged twice a week. All chemicals were obtained from PAA.

2.2.2.2. Freezing and thawing cells

In order to use cells with comparable passage numbers and to prevent mycoplasma contamination, a new stock of cells was thawed every 6-8 weeks. Moreover, each thawed aliquot was first treated for a week with MRA (Mycoplasma Removal Agent from MP) and handled in a separate incubator before cells were released for experiments.

Protocol Freezing: A 2x cryomedium containing growth medium (40%), FCS (40%), and DMSO (Dimethyl sulfoxide, 20%) was freshly prepared and sterile filtered. DMSO is a cryoprotectant, which reduces ice formation and thereby prevents cell death during the freezing process [185]. One 100 mm dish was used for freezing two 1 ml-aliquots of cells. The cells were washed with PBS and trypsinized as described above. Cells were sedimented by centrifugation at 200 g for 5 minutes. Pellets were dislodged by flicking and taken up in 1 ml growth medium. Directly before aliquoting 1 ml cryomedium was added. Cells were frozen in a special container with isopropanol at -80°C. The isopropanol enables cells to freeze at the rate of about 1 K per hour, thereby preventing cell damage [185]. After 3 days cells were put in liquid nitrogen for long-term storage.

Protocol Thawing: Frozen cells were thawed quickly in a water bath and were then directly diluted in 20 ml growth medium and centrifuged at 200 g for 5 minutes. The pellet was resuspend in 10 ml growth medium and 1 ml was put into a 35 mm dish and filled up to 1.5 ml with growth medium and MRA (at a final dilution of 1:100). The rest was kept at 37°C without MRA as a backup or for freezing new cell stocks, which were then marked as "px+1", referring to them being passaged one more time than the original stock. During the incubation with MRA cells were split once to provide fresh medium and MRA. After treatment cells were passaged into 100 mm dishes. When these dishes were confluent, the cells were ready to be used for experiments.

2.2.3. Generation of stable cell lines

U2OS cell lines were generated which stably expressed the *luciferase* gene under control of the human promoters of the three core clock genes *CRY1*, *NR1D1* and *PER2*,

as well as promoters of *ATG3*, *EIF5A2*, and *SCN5A*, which contained a ChIP-seq binding site. For control, a cell line expressing *luciferase* under the *GAPDH*-promoter was generated. Core clock promoter-*luc2* constructs were generated by transfection and selection, while the other promoter-*luc2P* cell lines were produced using lentiviral transduction. U2OS *PER2-luc2* was generated in both ways.

2.2.3.1. Transfection

Transient transfection was performed using Lipofectamine2000 (Invitrogen) according to the manufacturer's protocol for the 24-well format. U2OS cells are very sensitive to transient transfection and transfection efficiency usually ranges between 10-50 %. HEK cells show consistently high transfection efficiencies.

Protocol: One day before transfection 150,000 cells were plated in 500 µl DMEM supplemented with 10% FCS. For transfection 0.8 µg plasmid and 3 µl Lipofectamine were diluted in 50 µl Optimem each. After 5 min incubation the Optimem-Lipofectamine mix was added to the Optimem-plasmid mix and incubated for 20 minutes. 100 µl of this mixture were added dropwise to the cells, which were then incubated for 18-22 hours at 37℃. After incubation the medium was exchanged to DMEM containing FCS and Penicillin/Streptomycin.

For selection, 24 hours after transfection cells were trypsinized and split onto two 35 mm dishes, which corresponded to a dilution of 1:12, needed for selection using antibiotics. When cells were diluted less than this, more negative cells survived selection. 24 hours later the antibiotic for selection was added to the medium. It took 2-3 weeks until resistant cells filled one 35 mm dish under selection. During this time medium was changed twice a week and cells were trypsinized regularly to support growth. Selection medium was applied 24 hours after trypsinization. The smallest effective dosage of puromycin was 1.3 µg/ml, which was used for all experiments, but five times higher doses were tolerated as well. When cells reached confluence in a 100 mm dish, single cells were sorted into 3 - 4 white 96-well clear-bottom plates using FACS according to the standard protocol. Two aliquots of unsorted cells were frozen before sorting.

2.2.3.2. Viral transduction

MLV-based vectors are very commonly used for gene therapy studies and for producing stable cell lines, because they efficiently stably integrate the gene of interest into the genome of a target cell line. In order to provide more safety all the cis and trans

elements required for producing and packaging infectious viruses with the desired gene of interest are separated onto three plasmids, with minimal or no sequence overlap between them, thus reducing the possibility of homologous recombination between the vectors. A host cell line, 293T HEK, is used to produce the virus for infection. Once the viral RNA is encapsidated, virus particles bud off and are released into the cell supernatant. The supernatant is then used for infecting target cells. The produced virus itself, although capable of infecting cells, is replication-deficient. The Eco-plasmid containing the envelope-gene allows to produce ecotropic MLVs which are capable of infecting cells expressing the murine MCAT receptor.

Viral transduction was performed using the pVPack Eco-System (Agilent) according to the manufacturer's protocol with the exception of upscaling the virus-production to 100 mm dishes.

Protocol: Four million 293T HEK cells were plated on a collagen-coated 100 mm dish. For collagen-coating 1 ml of a collagen-solution was dispersed on a 100 mm dish and allowed to dry under the hood. 9 µg each of the plasmids pSIN_MCS which contains the gene of interest, the pV Pack-GP, that codes for gag and pol, and pVPack-Eco which encodes the envelope protein eco, were combined in one tube. Then 1 ml of 100% (v/v) ethanol and 0.1 volume 3 M sodium acetate were added to the plasmids. After mixing by inverting the tube, the DNA was precipitated at -80 °C for 30 minutes and collected by centrifugation at 12,000 g for 10 minutes at 4 °C. 1 ml 70% ethanol was added to the pellet, vortexed briefly, and centrifuged again at 12,000 g for 5 minutes at 4 °C. The wet pellets were stored at 4 °C overnight. On the second day HEK cells, which should be about 80% confluent, were transiently transected with Lipofectamine2000 according to the manufacturer's protocol for 100 mm dishes, using 27 µg plasmid and 67,5 µl Lipofectamine. After 16-20 hours incubation medium was exchanged to 6 ml fresh growth medium, which was incubated for another 48 hours, allowing viruses to be released into the supernatant. This supernatant was collected and filtered through a 0.45 µm filter. Part of it was added freshly to the target cells, the rest was aliquoted and snap frozen at -80 °C. Of note, the virus titer after one freeze-thawing cycle decreases by about 50%, therefore aliquots were only frozen and thawed once.

One day before viral transduction, the $U2OS_{mcat}$ target cells were plated to a confluency of about 30% in a 35 mm-dish. Before transduction target cells were washed once with PBS. Then 2 ml growth medium containing the viral supernatant and DEAE (Diethylethanolamine)-dextran solution at a final concentration of 10 µg/ml was added. After 3 hour incubation at 37 °C 1 ml growth medium was added to the cells. Importantly,

U2OS cells were most efficiently transduced with small amounts of viral supernatant (200 µl fresh undiluted supernatant or 500 µl frozen supernatant) rather than with large amounts. 24 hours after transduction the medium was exchanged and another 24 hours later single cells were sorted by FACS according to the standard protocol. Two aliquots of unsorted cells were frozen to allow new sorting at a later time point.

2.2.3.3. FACS

Fluorescence assisted cell sorting (FACS) is used for high throughput examining, counting, and sorting of single cells and is a convenient method to generate single cell clones of a stable cell line. The flow cytometer detects and analyses physical and chemical characteristics of single cells in a stream of fluid. Two liquids of sufficient different velocity or density to not mix, flow through a tube, where they form a two-layer stable flow, with the sheath fluid enveloping the sample fluid. This produces a very small sample flow cross-section, where a single cell can pass through the laser beam. Several detectors are aimed at this point, one in line with the light beam (Forward Scatter; FSC), and several perpendicular to it (Side Scatter; SSC, and one or more fluorescent detectors). Each suspended particle scatters the light and fluorescent chemicals that are bound to or were produced within a cell are excited into emitting light with the appropriate laser. The FSC correlates with the cell volume, whereas the SSC provides information about the inner complexity of the cell, i.e. the shape of its nucleus, amount and type of cytoplasmic granules, or the membrane roughness.

Promoter-luciferase-construct expressing cells were sorted for their healthiness, meaning that their shape and size were within the range of the normal population and that cells were propidium iodide negative. The fluorescent molecule propidium iodide, which was added to the sorting medium, binds with DNA but is membrane impermeant in viable cells. Therefore it is often used to differentiate dying from viable cells.

Protocol: The sorting medium contained cell dissociation buffer (5%), FCS (0.2%), propidium iodide (1 µg/ml) and penicillin/streptomycin (1x) and was freshly prepared and sterile filtered for each experiment. For sorting, cells were trypsinized and centrifuged in 5 ml DMEM growth medium at 200 g and 4 ℃ for 5 minutes. The supernatant was carefully aspired and sorting medium was added to the cells (500 µl for a 35 mm dish, 1 ml for a 100 mm dish). After careful resuspension by pipetting three times up and down, cells were separated by pipetting them through a single-cell filter and sorted into 96-well plates containing conditioned medium. This medium has been incubated

with growing cells for a couple of days, was then sterile filtered with a 0.22 µm filter, and complemented with the same amount of fresh medium. Conditioned medium contains all grow factors secreted by the cells and preparing and incubating it at 37℃ and 5% CO_2 a day before sorting increased survival rates of the cells, achieving viability rates of about 40%. Moreover, it was important that U2OS cells were not sorted directly after thawing, but only after about 2 weeks to increase survival rate.

2.2.4. Circadian experiments

2.2.4.1. Real-time bioluminescence monitoring

The firefly luciferase generates light from its substrate D-luciferin in a chemical reaction:

$$luciferin + ATP \rightarrow luciferyl\ adenylate + PP_i$$

$$luciferyl\ adenylate + O_2 \rightarrow oxyluciferin + AMP + light$$

Luciferase is inhibited by its product oxyluciferin and activated allosterically by ATP. Moreover, in a side reaction with luciferyl adenylate oxygen is reduced to hydrogen peroxide.

In cell lines stably expressing promoter luciferase constructs, the expression of luciferase is under control of the promoter of interest. Circadian activation of the promoter can be followed through detection of the resulting rhythmic luminescence, as shown for U2OS *BMAL1-luc* [4]. Cells in culture desynchronize over time. A medium change on the day of the measurement was however sufficient to yield consistent rhythms of luciferase expression in *BMAL1*-luc U2OS cells. This is probably due to the fact, that fresh medium contains a higher glucose concentration, as well as a higher concentration of other growth factors, resulting in a resetting of the clock through a small serum shock, a known method to reset cells [7].

Protocol: One day before the experiment 30,000 cells/well were split into a white 96-well plate. For optimal results cells needed to be 100% confluent at the time of starting the measurement. Cells were washed once with PBS and then medium was exchanged to DMEM (without phenolred) supplemented with FCS (10%), Hepes (25 mM), Luciferin (0.125 µM; P.J.K), and Penicillin/Streptomycin (1x). Medium was supplemented with Hepes to buffer the pH at physiological CO_2-concentrations. Cells were then incubated for 1-2 hours at 37℃ and 5% CO_2 before sealing them. Measurement of luminescence occurred every 20-30 min in the EnVision Excite (PerkinElmer) for

about a week. During this time cells were kept at 37℃, without additional CO_2. For each cell line 8 wells were measured at the same time and constructs were considered as rhythmic when more than one cell line with repeatable rhythmic expression of luciferase was produced. Thus obtained data was analyzed using the Multicycle software (Actimetrics). Raw data was detrended over the running average of the past 24 hours and smoothened by the software.

2.2.4.2. Time courses

In order to monitor rhythmic changes in the rate of transcription time course experiments were conducted. Expression of target genes was tested using qPCR (time courses Z1 and Z3) or a combination of qPCR and microarray studies (time courses Z4 and Z5). Time courses were performed by harvesting cell extracts for two days every 3-4 hours, yielding 16 or 12 time points, respectively.

Because cells desynchronize over time in cell culture and in order to be able to compare data from different measurements a precise starting point needed to be established, which means that cells needed to be reset by or entrained to a specific Zeitgeber. Two published methods were used to reset U2OS, one using the glucocorticoid analog dexamethasone (time courses Z1 and Z3) [177, 3] and the other through entrainment to rhythmic temperature changes (time courses Z4 and Z5) [87, 186]. Glucocorticoids are known as timing cues for peripheral oscillators *in vivo* and dexamethasone was found to be able to entrain circadian rhythms in rat-1 fibroblasts and in the periphery but not in the SCN. It is therefore believed, that glucocorticoids might participate in the resetting of peripheral clocks induced by feeding [177, 187]. Environmental temperature cycles are a universal entrainment cue for all circadian systems at the organismal level. In homeothermic vertebrates, it was recently shown that although the SCN can resist temperature entrainment, the peripheral oscillators cannot. It is therefore assumed that the SCN drives circadian rhythms in body temperature, which then act as a universal cue for entrainment of peripheral oscillators [87, 186]. Temperature-cycles were found to be suitable in U2OS cells (unpublished data). In this work it was preferably used over dexamethasone, because it yields better rhythms with a slightly higher amplitude and does not require treatment with a chemical.

Protocol: For resetting cells with dexamethasone cells were grown to confluence. The growth medium was exchanged to fresh medium containing dexamethasone at a final concentration of 1 µM. Cells were incubated with this medium for 20 minutes

at 37℃ degrees and then washed with PBS before adding freshly prepared growth medium.

For temperature entrainment U2OS cells were incubated at alternating 12 h, 33℃ and 12 h, 37℃ cycles. Medium changes were carried out at the transition from 33℃ to 37℃. For time course Z5, cells were pre-incubated under temperature entrainment for 4 days. On day 5, 160,000 cells/well were seeded in 32 x 35 mm dishes, which corresponds to a confluency of about 16% on the first day. Cells were grown in the same incubator for 5 more days under temperature entrainment until they reached confluence. Growth medium was exchanged directly before the cells were released to constant 37℃. 24 h and 48 h after medium change cells were harvested at 3 h intervals for 24 h. The time course Z4 was carried out similarly, with the exception that cells were incubated in two incubators with opposite temperature rhythms and harvested for 12 h on two days. Cells were harvested in RLT lysis buffer and snap-frozen for later RNA isolation using the RNeasy kit (Qiagen).

2.2.5. Biochemical methods

2.2.5.1. Protein extracts

Total protein extracts of U2OS cells were prepared for western blot analysis.

Protocol: Cells were trypsinized and cell pellets were either used directly or frozen at -80℃ and thawed for later protein extraction. 100 µl per 35 mm dish of cells were incubated with RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% Triton-X, 0.25% Sodium-Deoxycholate, 3 mM EDTA) that was freshly supplemented with protease inhibitors (cOmplete, Roche) and phosphatase inhibitors (PhosSTOP, Roche). After vortexing for 20 seconds, cells were incubated in an ultrasound bath at 4℃ for 10 minutes. Lysates were centrifuged at maximal velocity for 5 minutes at 4℃. The protein concentration of the supernatant was measured using the NanoDrop and then boiled with the appropriate amount of 4x Laemmli buffer (250 mM Tris HCl pH 6.8, 6% (w/v) SDS, 40% (w/v) Glycerol, 0.04% (w/v) Bromphenolblue, 20% β-Mercaptoethanol) for 3 minutes at 95℃.

2.2.5.2. SDS-PAGE and western blot

Protein extracts are commonly separated and analyzed using SDS-PAGE (Sodium-Dodecyl-Sulfate Polyacrylamide-Gelelectrophoresis) [188]. The combination of SDS-

PAGE with western blotting allows the detection of specific proteins on the gel [189].

In a SDS-PAGE proteins are separated according to their electrophoretic mobility, which is a function of both the length of the polypeptide and its charge. Laemmli-buffer is added to protein lysates. It contains the anionic detergent SDS that adds negative charges in proportion to the length of the polypeptide. Thereby, proteins on a denaturating acrylamide gel are separated mainly according to their size. Electrophoresis on polyacrylamide gels takes place vertically and the proteins run through two different types of gel. The upper portion, the stacking gel, contains pockets and collects ("stacks") all proteins at the border to the lower portion of the gel, the running gel. This portion can contain variable amounts of acrylamide/bisacrylamide, thereby changing the pore-size of the gel. In a western blot the separated proteins are subsequently transferred from the gel to a nitrocellulose membrane using the semi-dry blotting technique during which an electric current transfers the negatively charged proteins to the membrane. Proteins bind to the membrane by hydrophobic and charged interactions. They are detected using antibodies specifically for the proteins of interest. A second antibody is directed against the species from which the first antibody originated and coupled to a peroxidase. Proteins are detected by a chemiluminescence reaction carried out by this conjugated peroxidase that oxidizes luminol to emit visible light. Several secondary antibodies can bind to one primary antibody, thus enhancing the signal. The incubation with a light-sensitive film on the thus-treated blots leads to the visualization of the bands.

Table 2.5.: Protocol for the used polyacrylamide gels. Temed and APS were added at the end because they start the polymerization reaction.

Chemicals	Running Gel (10%)	Stacking Gel (5%)
Acrylamide (30%)/Bisacrylamide (0.5%)	5.7 ml	0.85 ml
Tris-HCl, pH 8.8 (1 M)	6.5 ml	0.3 ml
H_2O	4.5 ml	3.8 ml
SDS (10%)	0.167 ml	0.05 ml
Temed	20 µl	5 µl
APS (10%)	100 µl	25 µl

Protocol: For the experiments 10% gels were made, which effectively separate the proteins of the desired size of about 60-90 kDa. The protocol for both the separating and the stacking gels can be found in Table 2.5. Gels were run in SDS-Buffer (60.55 g TRIS, 288.3 g Glycine,100 ml SDS-solution [10% w/v], H_2O ad 10 L, pH 6.8) over

night at 5 mA per gel, and transferred to an Optitran nitrocellulose membrane (BA-S 85, Whatman) using the semi-dry blotting technique. Whatman-paper (3 mm) and membrane were soaked in Blot-buffer (24.2 g TRIS, 112.6 g Glycine, 2 L methanol, 80 ml SDS-solution (10%[w/v]), H_2O ad 10 L). Three layers of whatman-paper were placed on the bottom of the blotting chamber, followed by a nitrocellulose membrane, the gel and another layer of three whatman-paper. Gels were blotted for 2 hours at 300 mA/gel, after which the bands were visualized by Ponceau S solution (0.1% Ponceau S in 5% acetic acid), a diazo dye that reversibly and unspecifically stains all proteins on the blot. Staining was used to cut out individual lanes of protein-extracts, which were later incubated with different antibodies. Blots were first blocked for 45 min at room temperature with 5% milk powder in TBS (TRIS Buffered Saline) in order to prevent unspecific binding of the antibody. They were incubated with the first antibody diluted as indicated in table 2.6 for 2 hours at room temperature. In order to test the affinity-purified antibodies total extracts were incubated with the pre-immune serum (1:1000), serum before affinity-purification (1:1000), and the same serum (1:1000) blocked with the specific peptide (10 µg/ml). After a 4 x 2-minute wash with TBS the secondary antibody was applied over night at 4℃. The blot was washed 3 x 10 minutes with TBS. Chemiluminescence detection with a solution containing luminol, p-coumaric acid and H_2O_2 was carried out on a light-sensitive film (Fuijifilm SuperRX) .

Table 2.6.: Western blot conditions for peptide-antibodies. The antibodies were tested for specificity in western blots. Affinity-purified antibodies were used in the indicated concentrations. The first antibody was incubated for 2 hours at room temperature or at 4℃ over night.

Target protein	Dilution	2nd Antibody	Signal height
BMAL1	1:750	1:5000	75 kDa
CLOCK	1:300	1:5000	9 kDa
CRY1	1:300	1:5000	60 kDa
CRY2	1:300	1:5000	60 kDa

2.2.6. Molecular biological methods

2.2.6.1. Quantification of RNA and DNA

The concentration of RNA and DNA was determined using the NanoDrop.

For RNA the absorbance ratio at 260 nm and 230 nm was determined. The 260/230 values should be in the range of 2.0 - 2.2. Lower ratios indicate the presence of con-

taminants which absorb at 230 nm. If the ratio was lower RNA was purified again using the Qiagen RNeasy Kit. Moreover, the ratio of absorbance at 260 nm and 280 nm was measured. A ratio of 2 was considered as "pure" for RNA, whereas a ratio of 1.8 was considered as "pure" for DNA. A lower ratio might indicate the presence of proteins or other contaminants that absorb strongly around 280 nm.

For the microarray RNA samples were labeled with Cy3 and Cy5 (see Figure 2.8). The NanoDrop software offers a test type called "MicroArray" which estimates the amount of both dyes in the sample along with the concentration of the RNA. Value and purity of the specific dye were assessed and a ratio of pmol dye per µg cRNA was determined, which was used for calculating the amount of cRNA to be used on the array.

2.2.6.2. PCR

Table 2.7.: Standard protocol used for all PCR reactions using Finnzymes Phusion polymerase.

PCR protocol		
Repeat	**Temp (℃)**	**Time**
1x	98	30s
28x	98 72	10s 30s + 30s/kb
1x	72	180s
hold	4	∞

Table 2.8.: Standard 1x Mastermix for PCR using Phusion. For all PCRs the High Fidelity (HF) Master Mix from Finnzymes was used. It contains $MgCl_2$.

1x PCR Mastermix	
Volume (µl)	**Reagent**
10 µl	HF buffer
1 µl	dNTPs (10 µM each)
100 ng	gDNA
0,5 µl	Phusion Pol
3%	DMSO
ad 50 µl	water

The Polymerase Chain Reaction (PCR) was first presented in 1985 by Kary B. Mullis [190, 191] and contributed so profoundly to today's scientific work that already in 1993 he received the Nobel Prize for Chemistry for its invention.

PCR is based on a simple 3-step process during which a specific region of DNA is amplified. First, the DNA template is denaturated by heat. Next, specific primers, short DNA templates complementary to the beginning and the end of the amplified region, anneal to the template. The primers are then extended by a thermostable polymerase. Each cycle leads to a duplication of the target sequence. After repeating this procedure multiple times, a high amplification of the region of interest is obtained, which can be visualized, for example on an agarose gel. Preceded by reverse transcription this method can even be applied for amplification of mRNA. The basic PCR method has proven to be useful in order to detect small quantities of DNA in a sample as well as for amplification of a specific gene.

In this work PCR was used for amplification of the promoter regions for *CRY1, NR1D1*, and *PER2* from genomic DNA obtained from U2OS cells. For all experiments Phusion Polymerase (Finnzymes) was used. The basic PCR protocol can be found in tables 2.7 and 2.8. Primers and PCR-conditions for the constructs are in Tables 2.9 and 2.10. The constructs were first cloned into the pGL3-vector (Promega). This vector was then used as template for amplification and cloning into the pGL4.20 vector. The *CRY1*-promoter was amplified using the same primers in both cases. Because of already existing restriction sites in the pGL4 vector, primers with new restriction sites had to be used for *NR1D1* and *PER2*. This did not affect the sequence of the construct.

Table 2.9.: List of primers for cloning. The added restriction site is written in minuscule and is preceded by an overhang of bases to allow processing by the enzyme.

Name	Sequence
Cry1_NheI_FWD	TTTgctagCCTTGAGAAATCACACTTCAGAGGG
Cry1_HindIII_REV	TTTaagcttGCCGACACCTTCGCTTCC
NR1D1_P1_BglII_FWD	CCCagatctGGTCGGTCACTTCTCCTAAGCTTCC
NR1D1_P1_NcoI_REV	CCccatgGTCTTCACCAGCTGAGAGCGGTC
NR1D1_P1_NheI_FWD	CCCgctagcGGTCGGTCACTTCTCCTAAGCTTCC
NR1D1_P1_BglII_REV	TTTagatctGTCTTCACCAGCTGAGAGCGGTC
Per2_2_MluI_FWD	CCacgcgtCGACCACAGGCACCTCAGAG
Per2_2_NcoI_REV	CCccatggGCTCCTGAGCTGCGAAGCG
Per2_SacI_FWD	CCgagctcCGACCACAGGCACCTCAGAG
Per2_HindIII_REV	GGaagcttGCTCCTGAGCTGCGAAGCG

Table 2.10.: Sequence and PCR-conditions used for amplification of promoter-regions. The table shows into which plasmids the inserts were cloned, which part of the promoter was amplified (bp rel to transcription start site), the length of the fragment (bp) and which PCR-conditions (temperature and DMSO amount) were used.

Name	from ... to	Position	Length	PCR
Cry1_NheI_FWD Cry1_HindIII_REV	gDNA to pGL3 pGL3 to pGL4.20	-1107 to + 799	1906	69℃ 3% DMSO
NR1D1_P1_BglII_FWD NR1D1_P1_NcoI_REV NR1D1_P1_NheI_FWD NR1D1_P1_BglII_REV	gDNA to pGL3 pGL3 to pGL4.20	-858 to + 625	1483	72℃ 3% DMSO 72℃ 3% DMSO
Per2_2_MluI_FWD Per2_2_NcoI_REV Per2_SacI_FWD Per2_HindIII_REV	gDNA to pGL3 pGL3 to pGL4.20	-848 to +204	1053	72℃ 10% DMSO 72℃ 10% DMSO

2.2.6.3. Quantitative PCR

Using real-time PCR technology the basic PCR method was expanded for quantitative measurements as well, called quantitative PCR (qPCR) or real-time PCR. For quantitative PCR a fluorescent probe is added to the PCR Mastermix, which will generate a fluorescent signal relative to its incorporation into the PCR-amplicon, that can be quantified in real-time during the PCR-run. All probes in this work are fluorescent hydrolysis probes, which either stem from the Universal probe library (UPL, Roche) or have been designed specifically and are tagged with the fluorophor 6-FAM (Carboxyfluorescein) at the 5'end and with the quencher TAMRA (Rhodamine) at the 3'end. Hydrolysis probes bind within the amplicon generated by forward and reverse primer and only emit a fluorescent signal once the quencher is cleaved by the polymerase during the elongation of the PCR-fragment. The generation of the amplified product is monitored through an increase in fluorescence after each cycle of amplification. Typically the curve has a sigmoidal shape, with a flat line at the beginning and the end and a steep rise in between. Quantification of gene targets on the Roche Light Cycler is based on crossing point cycle (Cp) values. The crossing point cycle is determined by taking the second derivative maximum of the generated signal, which corresponds to a point very early in the amplification process, where amplification occurs in a linear fashion.

Quantitative PCR was performed on the Roche Light Cycler 480 using the TaqMan Gene Expression Master Mix from Applied Biosystems. For normal quantifica-

tion, cDNA was diluted 1:20. ChIP-DNA was eluted in 200 µl after clean-up for direct use. The enrichment of distinct regions on chromosome 7 was measured with 5 ng and 50 ng genomic DNA. For preRNA measurements the Fast Advanced Master Mix from Applied Biosystems was used which provided Cp values 2-4 cycles earlier than the Gene Expression Master Mix. PCR protocol and mastermix were identical in both experiments and are listed in Tables 2.11 and 2.12.

Table 2.11.: Standard protocol for the qPCR master mix of one 20 µl reaction. 15 µl mastermix were prepared and pipetted to 5 µl of sample.

qPCR Mastermix	
Volume (µl)	**Reagent**
10 µl	2x Master Mix
0.2 µl	Forward primer (20 µM)
0.2 µl	Reverse primer (20 µM)
0.2 µl	probe (10 µM)
4.4 µl	water

Table 2.12.: Standard protocol used for qPCR-reactions at the Light Cycler 480.

qPCR protocol			
Step	**Repeat**	**Temp (°C)**	**Time**
Initial Denaturation	1x	50 95	2 min 10 min
Amplification	40x	95 60	15 s 1 min
Cooling	1x	37	30s

2.2.6.4. PCR clean-up

The Wizard (R) SV Gel and PCR Clean-Up System from Promega is a membrane-based system, designed to extract and purify DNA fragments of 100 bp to 10 kb. DNA binds to silica membranes in the presence of chaotropic salts, allowing other substances to be washed through the column. DNA was purified from agarose gels after PCR amplification or restriction digestion as well as directly from solution after ChIP.

Protocol: The protocol for all purifications is similar, but for the first step. When purifying a DNA fragment from an agarose gel, the gel was run according to standard procedures. Subsequently the band was excised from the gel and dissolved in the presence of guanidine isothiocyanate (1 ml membrane binding solution per g of gel)

Table 2.13.: Primers used for detection of mRNA levels using quantitative PCR.

Primer Name	Sequence 5' - 3'	Probe No
mRNA - probe: UPL		
ANKRD12_FWD	cgtccagtggatgtagcaga	82
ANKRD12_REV	tgagcctcttcggaatctgt	
ATG3_FWD	aacagtgaccattgaaaatcacc	87
ATG3_REV	gattttcttcatcacctcagcat	
EIF5A2_FWD	ttagcctcggcaaaccaa	60
EIF5A2_REV	cttgtgctttagaaatttcctttgt	
SCN5A_FWD	gagcaacttgtcggtgctg	12
SCN5A_REV	gatttggccagcttgaagac	
Luc2P_FWD	cgcagtaggcaaggtggt	77
Luc2P_REV	cagtgtcttaccggtgtcca	
18s_rRNA_FWD	ttgactcaacacgggaaacc	77
18s_rRNA_REV	cgctccaccaactaagaacg	
mRNA - probe: 5' 6-Fam-3'Tamra		
HPRT_FWD	ctggcgtcgtgattagtgat	
HPRT_REV	ctcgagcaagacgttcagtc	
HPRT_probe	cacccttccaaatcctcagcataatg	X
GAPDH_FWD	catcaatggaaatcccatca	
GAPDH_REV	gactccacgacgtactcagc	
GAPDH_probe	tccaggagcgagatccctcca	X
BMAL1_FWD	ttggacgactgcattctcat	
BMAL1_REV	aatagctgttgccctctggt	
BMAL1_probe	tccatctatcgcgtgccgaga	X
CRY1_FWD	cttgatgcagattggagcat	
CRY1_REV	ccattgggatctgttctcct	
CRY1_probe	acagccacatccaacttccagca	X

Table 2.14.: Primers used for detecting introns and ChIP-enrichment using quantitative PCR.

Primer Name	Sequence 5' - 3'	Probe No
pre-mRNA - UPL/5' 6-FAM-3'Tamra		
BMAL1_intron13_FWD	ctaggcagtacaagaaccaaaagac	25
BMAL1_intron13_REV	gatgaatgtagcttttgggtgac	
ANKRD12_intron10_FWD	gagcgagtcttctggtttgac	31
ANKRD12_intron10_REV	ttctccagtcttccacagttca	
ATG3_intron10_FWD	ccctctcagttttggcttattct	26
ATG3_intron10_REV	cacatggtatcagttggacaaaa	
GAPDH_intron4_FWD	gtgtccctcaatatggtcctg	X
GAPDH_intron4_REV	ttctccatggtggtgaagac	
GAPDH_intron4_probe	tactcagcgccagcatcgcc	
qPCR ChIP peak - 5' 6-FAM-3'Tamra		
GAPDH_gDNA_FWD	catcaatggaaatcccatca	
GAPDH_gDNA REV	ttctccatggtggtgaagac	
GAPDH_gDNA_probe	tactcagcgccagcatcgcc	X
PER1_Ebox1,2_FWD	agacctctcagcctatgagaaagc	
PER1_Ebox1,2_REV	cccgacctgccaagattg	
PER1_Ebox1,2_probe	tggagagcgggactggcatttacg	X
PER1_Ebox3_FWD	cacagagccctcccttca	
PER1_Ebox3_REV	ggcactcccctcactgtca	
PER1_Ebox3_probe	ccttccctcattgttcaggaaagcttaggc	X
PER1_Ebox4_FWD	catccttcctccacccactc	
PER1_Ebox4_REV	caaatcctgggagaggacagtta	
PER1_Ebox4_probe	cccaccggtcacacgtggacc	X

Table 2.15.: List of primers used to verify amplification of certain regions on chr 7 using quantitative PCR.

Name	Sequence 5' - 3'	Probe No	Position on chr 7
PDGFA_intron2_FWD	agtgctggtcttggctctg	43	0.5 Mb
PDGFA_intron2_REV	ctgcgagccgtgtttacag		
LFNG_peak_FWD	ccaggccactgttggaaa	64	2.6 Mb
LFNG_peak_REV	gatatggctggtgcctcac		
FOXK1_up_chr7_FWD	ctccttgatcaggtcccaac	20	4.6 Mb
FOXK1_up_chr7_REV	tgtgtgtgttgtgcagtgtgt		
EGFR_down_chr7_FWD	cctggttcaacacacagctc	37	55 Mb
EGFR_down_chr7_REV	ttgtggttctaaaggcgaaac		
Serpine1_prom_FWD	taacccctggtcccgttc	77	100 Mb
Serpine1_prom_REV	caacaaccttgtctggctga		

in a thermocycler at 65℃. The solution was then incubated on the columns. DNA-fragments that were purified from solution (i.e. after ChIP) were mixed with one volume of membrane binding solution and incubated directly on the silica columns. After centrifugation at 16,000 g the flow-through was discarded and columns were washed first with 700 µl wash buffer followed by 1 minute centrifugation, next with 500 µl wash buffer and 5 minute centrifugation. The empty columns were dried at 1 min maximum velocity in the centrifuge. DNA was eluted in 30-50 µl prewarmed water.

2.2.6.5. Agarose gel electrophoresis

DNA is a highly negatively charged molecule and can be separated according to its size on agarose gels. It is then visualized by adding ethidiumbromide to the gel, which intercalates into DNA and fluoresces when exposed to ultraviolet (UV) light. Agarose gels are easily cast and electrophoresis takes place horizontally in TAE (TRIS Acetic Acid EDTA) buffer (4.84 g TRIS, 1.14 g Acetic Acid, 0.744 g Na_2EDTA per L). Loading dye (30% glycerol, Bromphenol Blue, and Xylene cyanol) is added to the samples which increases viscosity and thereby allows pipetting of the sample into the gel pockets. For the DNA-products in this work 1% [w/v] (in TAE) agarose gels were used. To prepare these gels the appropriate amount of agarose is boiled in TAE buffer and then stored at 60℃ until used. Gels are cast and 1 drop ethidiumbromide is added and mixed. When the gel cools down it becomes solid. Gels were run in a specialized chamber filled with TAE buffer at 100 V for 30 minutes or until desired separation has taken place. For cloning exposure to UV-light must be minimalized in order to prevent UV-light induced DNA-damage.

2.2.6.6. RNA purification

The RNeasy Kit (Qiagen) was used for extraction of RNA from cell lysates and for cleaning-up RNA after labeling for the microarray experiments.

Protocol: Cell lysates were prepared by washing cells once with PBS, and then adding 350 µl RLT buffer to the dish. Cells were then detached using a cell scraper and pipetted into a tube. This tube was vortexed for 20 seconds, after which the lysate was sheared by pipetting 10x up and down using a syringe with a 20G-needle. The resulting lysate was snap-frozen and stored at -80℃ until further preparation.

Frozen aliquots were quickly thawed at 37℃ and mixed by pipetting after addition of 1 volume ethanol (70%). The mixture was transferred to an RNeasy column and

centrifuged for 15 seconds at 10,000 g. The columns were washed first with 700 µl RW1 and then with 500 µl RPE. A final washing step was carried out with another 500 µl RPE, followed by 2 minutes centrifugation at 10,000 g. The column was then placed into a new collection tube and dried by 1 minute centrifugation at 16,000 g. The RNA was eluted with 40 µl RNase free water.

Labeled samples were purified similarly. First, the volume of the sample was adjusted to 100 µl with water. Next, 350 µl RLT were added, followed by 250 µl of ethanol (100%). Then the sample was transferred to the RNeasy column. The same protocol for purification was used as before, with the exception of starting directly with wash buffer RPE1.

2.2.6.7. gDNA extraction

The DNeasy Kit (Qiagen) was used for extraction of genomic DNA from cell lysates.

Protocol: First, cells were trypsinized as described before. Cells were taken up in PBS (2 ml per 35 mm-dish, 10 ml per 100 mm-dish) and centrifuged 5 minutes at 300 g. The pellet was resuspended in 200 µl PBS. Next, the lysate was incubated at room temperature for 2 minutes with 20 µl Proteinase K (provided) and 10 µl RNase (Stock conc. 10 mg/ml). Then, 200 µl buffer AL was added to the sample and vortexed before incubation at 56℃ for 10 minutes. Next, 200 µl ethanol (100%) was added and mixed by vortexing. This lysate was then transferred to a minicolumn and centrifuged 1 minute at 8000 g. The columns were washed first with 500 µl AW1 followed by 1 minute centrifugation at 8000 g and next with 500 µl AW2 followed by 3 minutes centrifugation at 20,000 g. The gDNA was eluted after 1 minute incubation with 200 µl buffer AE.

2.2.6.8. cDNA synthesis

For cDNA synthesis the QuantiTect Reverse Transcription Kit (Qiagen) was used. It comprises two main steps: elimination of genomic DNA and reverse transcription. Reverse transcriptase requires primers to initiate the transcription process. The primers in the supplied RT Primer Mix allow cDNA synthesis from all regions of the RNA transcripts. Quantiscript Reverse Transcriptase contains an RNA-dependent DNA-polymerase activity which transcribes cDNA from an RNA template. Moreover, an RNase H activity of the enzyme specifically degrades the RNA from the RNA:cDNA hybrids, but has no effect on pure RNA.

Protocol: 1 µg of RNA was used for preparation of 20 µl of cDNA. The volume was adjusted to 12 µl and incubated with 2 µl gDNA wipe-out buffer at 42℃ for 2 minutes and then placed on ice. The reverse-transcription master mix was prepared by combining 1 µl Quantiscript Reverse Transcriptase, 4 µl Quantiscript RT buffer (5x) and 1 µl RT Primer Mix. Template RNA was incubated with the reverse-transcription master mix for 30 minutes at 42℃. An incubation step of 3 minutes at 95℃ inactivated the reverse transcriptase. The cDNA was diluted 20-fold and stored at -20℃.

2.2.6.9. Cloning

The *CRY1*-, *NR1D1*- and *PER2*-promoters were amplified by PCR from human U2OS genomic DNA. They were first cloned into the pGL3 vector (Promega) which did not allow for selection due to lack of an antibiotic-resistance marker. This plasmid was used as template for PCR amplification of the constructs for cloning them into the pGL4.20 which harbored a puromycin-resistance. The pGL3-*PER2-luc* construct was also digested with the enzymes MluI and BamH1 to yield a promoter-luc fragment that could be cloned into the pSIN_MCS. The promoter-luciferase constructs for *ATG3, EIF5A2, GAPDH*, and *SCN5A* were cloned after restriction digestion from the pSGG plasmid into the pSIN_MCS-plasmid.

Protocol: The promoters were amplified with PCR and purified using the PCR clean-up kit as described in the PCR-section. All of the purified PCR-constructs and 2-4 µg vector were then digested at 37℃ for 2-3 hours with the restriction enzymes (NEB). Amounts of enzymes were calculated based upon their activity and the data given in the NEB catalog. In order to prevent religation of the vector the 5'-phosphate was removed by 30 minutes incubation at 37℃ with 1 µl CIP (Calf-Intestinal Alkaline Phosphatase). After digestion the constructs were purified directly by a PCR clean-up column, while the vector was run on a 1% agarose gel before the appropriate band was dissected from the gel and purified using the PCR clean-up kit. Ligation reactions were prepared using 50 ng vector and 3x and 9x the amount of insert, as estimated by the DNA-concentration and taking into consideration the varying molecular weight. Per 20µl reaction 2 µl 10x Ligation buffer (Fermentas), 1µl vector and 1µl ligase (Fermentas) were combined. Ligation occurred for 1 hour at 22℃. Half of the reaction was then transformed into competent *E.coli* DH5α or *E.coli* XL1-blue cells according to the protocol and plated out on LB-ampicillin (LB-Amp) agar plates (final conc. Amp 100 µg/ml) that were incubated at 37℃ over night. From each plate 3-6 colonies were picked and

incubated over night in a 3 ml LB-Amp (final conc. Amp 100 µg/ml) liquid culture. From these cells plasmid preparations were done according to the protocol. These plasmid preparations were test-digested in a 15 µl reaction containing 4 µl plasmid and 0.25 µl of each restriction enzyme. Half of the reaction was loaded on a 1% agarose gel. When the digested fragment matched the expected height the fragment was purified using the PCR clean-up kit and the entire insert was sequenced (GATC, LGC) before generation of stable cell lines.

2.2.6.10. Transformation of competent cells and retransformation

Transformation is a process by which bacteria take up extracellular DNA, that will subsequently be replicated in each cell division. For transformation to happen, bacteria must be in a state of competence, for example as a response to environmental conditions. In the laboratory, competent *E.coli* strains are prepared by incubation in a solution containing divalent cations, which makes them take-up plasmids after a heat-shock to 42℃.

Protocol: Competent *E.coli* cells were stored in 50 µl aliquots at -80 ℃. For transformation, cells were thawed on ice for 5-10 minutes. Then, about 25 ng plasmid from ligation or 200 ng purified plasmid for retransformation, were added to the cell suspension and incubated on ice for 30 minutes. Next, cells were driven to take up the plasmids by heat-shock through a 75-second incubation at 42℃, with a subsequent incubation at 4℃ for at least 3 minutes. After re-transformation of purified plasmids, the *E.coli* cells were directly transferred to a liquid culture of LB-Ampicillin. However, after ligation and transformation cells were first incubated at 37℃ for at least 3 rounds of replication (1-2 hours) in 950 µl SOCS medium without antibiotics. Cells were then centrifuged for 30 seconds at 13,000 rpm and most of the supernatant was discarded. About 50-100 µl were retained to resuspend cells and plate them on LB-Amp agar plates, which allows for picking single colonies that contain identical copies of the same plasmid.

2.2.6.11. Purification of plasmids

Protocol Mini Prep precipitation: For small-scale plasmid purification after cloning a 2 ml overnight-culture of *E.coli* cells was centrifuged for 2.5 min at 13,000 rpm. The pellet was resuspended in 300 µl resuspension buffer S1 (50 mM Tris-HCl, 10 mM EDTA, 100 µg/ml RNaseA; pH 8). After addition of 300 µl lysis buffer S2 (200 mM NaOH, 1%SDS) the tubes were inverted 5x and incubated for 3 minutes. Then, 300

µl neutralization buffer S3 (2.8 M KAc, pH 5.1) was added, the tubes were inverted 5x and incubated for another 3 minutes. The lysed cells were centrifuged for 10 minutes at 13,000 rpm. 800 µl of the supernatant was added to 600 µl isopropanol and mixed well. After subsequent 15 minutes centrifugation at 13,000 rpm the pellet was washed with 500 µl ethanol (70%) and centrifuged for another 5 minutes at 13,000 rpm. The supernatant was discarded and the pellets were dried and resuspended in 50 µl water. The plasmids were tested for integration of the construct using test digestion. When the digestion looked as expected the plasmids were purified using the PCR clean-up kit and subsequently sequenced.

Protocol Midi Prep columns: Plasmids where the insert was verified by sequencing were prepared in a larger scale and purified by Midi Prep columns (NucleoBond) using the buffers as before. Briefly, 50 ml overnight culture were pelleted by centrifugation at 4500 g for 30 minutes. The pellet was resuspended in 4 ml buffer S1, lysed in 4 ml buffer S2 and neutralized in 4 ml buffer S3, as described above. In the meantime a column was equilibrated with 2.5 ml buffer N2 (100 mM Tris-HCl, 900 mM KCl, 0.15% Triton X-100, 15% EtOH, pH 8.5). The lysate was clarified by filtration and then loaded onto the NuceloBond column. The column was washed with 10 ml buffer N3 (100 mM Tris-HCl, 1.15 M KCl, 15% EtOH, pH 6.3). DNA was eluted with 5 ml buffer N5 (100 mM Tris-HCl, 1 M KCl, 15%EtOH, pH 8.5). 3.5 ml room-temperature isopropanol was added and mixed carefully to precipitate the eluted DNA. The mix was centrifuged at 15,000 g for 30 min at 4℃. The pellet was washed with 2 ml room-temperature ethanol (70%) and dried for 15 minutes. It was redissolved in 200 µl water.

2.2.7. ChIP-sequencing

Chromatin-Immunoprecipitation (ChIP) is a powerful technique for evaluating interactions of transcription factors and DNA-associated proteins with the DNA. Figure 2.6 displays the important initial steps of ChIP. First, all proteins are crosslinked to the DNA through a chemical agent like formaldehyde. Next, lysates are prepared and the chromatin is sheared to small fragments of about 200-500 bp length using ultrasound sonication. DNA-fragments bound by the protein of interest are then immunoprecipitated through beads that are covered with an antibody specific against the protein of interest. These DNA fragments are then eluted and purified. Enrichment of specific sequences in comparison to the input is tested using qPCR.

Many advances using this method have been made after the invention of high-

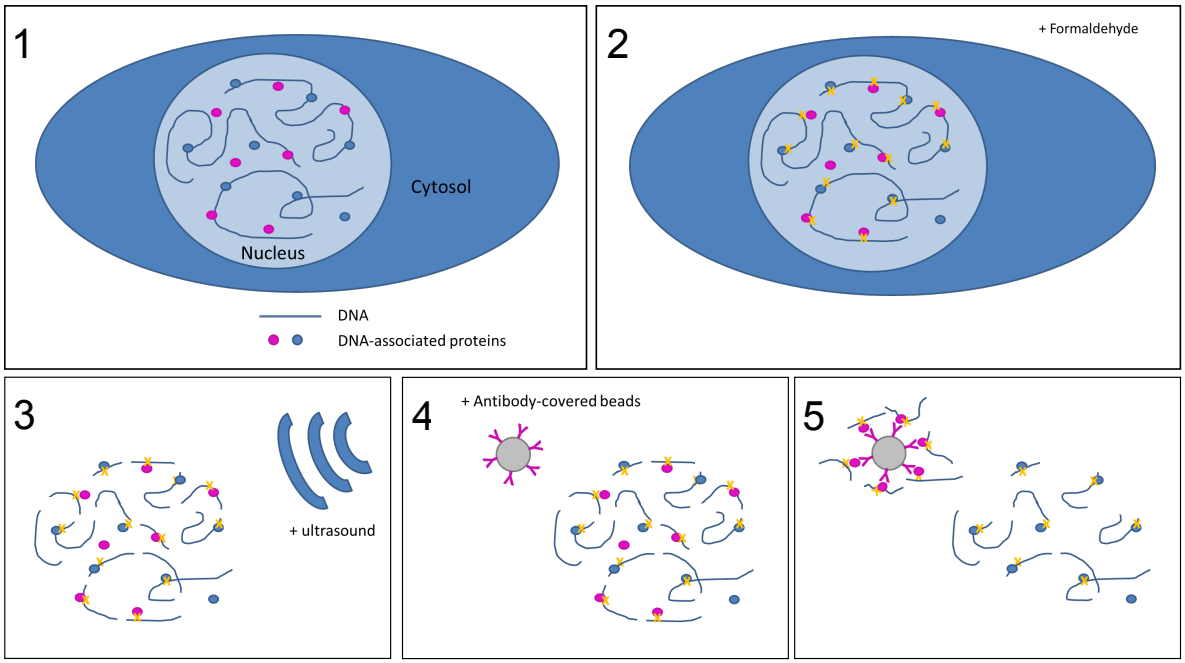

Figure 2.6.: Overview over chromatin-immunoprecipitation (ChIP). Schematic representation of a cell (1). For ChIP proteins were crosslinked to DNA using formaldehyde (2). Next, cell lysates were sonicated to shear DNA into small fragments of about 200-500 bp (3). Beads covered with an antibody specific against the protein of interest were incubated with the lysate (4) in order to enrich and precipitate specifically all DNA-fragments bound by this protein (5). The DNA-fragments are eluted from the beads and purified (not shown).

throughput genome-wide sequencing, so-called "deep sequencing". When ChIP-samples are sequenced on a high-throughput scale enriched DNA-regions, and thus, binding sites for the protein of interest, can be assigned on a genome-wide scale.

Several techniques for deep-sequencing have evolved in the past. In this work deep sequencing was performed using the Illumina sequencing-by-synthesis (SBS) technology (Figure 2.7), the bases of which have been developed by S. Balasubramanian and D. Klenerman in the mid-1990s [192]. The basic concept of SBS is that clusters of DNA fragments are sequenced during amplification of the complimentary strand through addition of one nucleotide at a time and subsequent monitoring. First, individual short DNA fragments are bound via adaptors to a solid phase, where they are clonally amplified to produce several million dense clusters of double-stranded DNA. Then, the complimentary strands to these clusters are synthesized. For this, reversible fluorescent terminator nucleotides in four colors (one per nucleotide) are added during amplification. The terminator sequence allows addition of only one nucleotide to

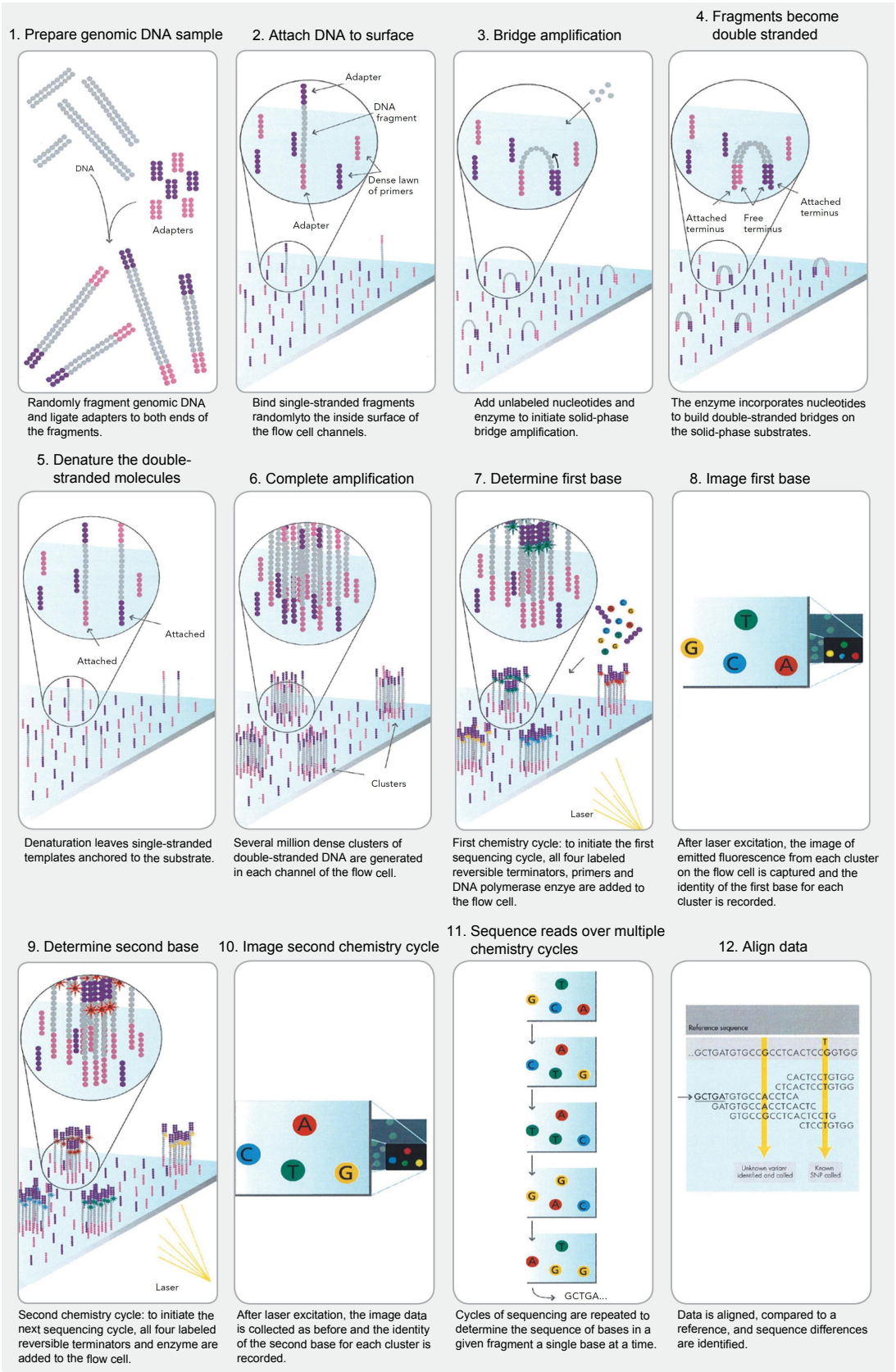

Figure 2.7.: Overview over the Illumina sequencing process. Explanations as given in the figure. Figure adapted from the Illumina-Systems and Software Manual.

a DNA-strand. After addition of the nucleotide the fluorescent signal emitted by each cluster of amplified DNA fragment is measured. Then, the next nucleotide is added to the DNA strands, followed by a new recording of the emitted signal. Sequence reads over multiple cycles of nucleotide-additions allow determination of the sequence of the given fragments. On the Illumina Genome Analyzer IIx about 38 bases from each strand from about 16 million clusters can thus be sequenced, resulting in the sequencing about about 1.2 Gigabases (roughly one third of the human genome) which are then aligned to the genome to identify the sequenced region. Through refinement and optimization the 2011 generation of Illumina SBS technology-based instruments generates up to 600 Gigabases per run, sequencing up to 55 Gb per day.

2.2.7.1. ChIP-protocol

ChIP was performed with antibodies against BMAL1, CLOCK, and CRY1 in confluent desynchronized U2OS cells. Cells were grown for at least 5 days in 100 mm dishes until they reached confluence in order to avoid any bias due to rhythmicity or the cell cycle on the cells. The protocol was adjusted and optimized from a protocol published on the web page of Abcam [193] and was then first published in mouse ChIP-seq experiments with our collaboration partners [5].

Protocol: 10 x 100 mm-dishes were sufficient for 4 ChIPs with enough material for one sequencing run. First, proteins were cross-linked to DNA by incubating them for 10 min at room temperature with formaldehyde (final conc. 0.4%). Cross-linking was stopped by a 5 minute incubation with glycine (final conc. 125 mM). Then cells were rinsed twice with cold PBS and lysed in 500 µl FA lysis buffer (50 mM Hepes-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS; freshly supplemented with 10% cOmplete) per dish. Next, chromatin was sheared using a conventional sonicator with a MS72 sonotrode which was introduced into the sample for a total of 6 minutes at 50% output. The ultrasound was pulsed 12 times for one minute at 1 second on, 1 second off. After each minute a cooling time of 30 seconds was allowed. The entire sonication process was performed in an ice bath. 1 ml was sheared at a time.

The sonicated lysates were centrifuged for 10 min at 8000 g at 4℃. After this, samples were snapshot-frozen and stored at -80℃. One freeze-thaw cycle did not decrease ChIP-efficiency. Protein A Sepharose-Beads CL-4B were used for immunoprecipitation. Beads were washed 3x with RIPA buffer and then taken up in the corre-

sponding volume of RIPA buffer containing 10% 20x Protease Inhibitor (cOmplete). In order to rid the lysate from everything that would unspecifically bind to the beads, a pre-clearing step was carried out. 60 µl beads were incubated with the lysate for 1.5-2 hours at 4℃. After incubation the beads were centrifuged for 2 minutes at 1500 g. The protein concentration of the sheared supernatant was estimated using the NanoDrop. For the ChIP 25 mg protein lysate were used and 2 mg of protein were stored as input control. Next, the antibody was bound to the proteins in the supernatant during 3-5 hours incubation on the rotating wheel at 4℃. For the BMAL1 ChIP 4 µl affinity purified antibody were used, and for the CLOCK and the CRY1 ChIP 16 µl . After overnight incubation on a rotating wheel with 60 µl beads suspension the protein-antibody-beads conjugates were precipitated at 1500 g for 2 minutes. The pellet was washed 3 times with 1 ml ChIP wash buffer. After the first wash the beads were transferred to a new tube in 300 µl ChIP wash buffer in order to rid the sample of DNA unspecifically bound to the tube wall. The remaining beads were transferred with another 700 µl wash buffer. The fourth wash was performed in 1 ml ChIP final wash buffer which provided more stringent wash conditions through higher salt concentration. After this the supernatant was carefully taken off with a hamilton-pipette and discarded. DNA was eluted by incubation with 120 µl ChIP elution buffer for 15 min at 30℃, shaking at 900 rpm. Beads were centrifuged for 2 minutes at 2000 g and the supernatant was transferred to a new tube. Decrosslinking was performed with the ChIP-eluate and the input in the same fashion. The SDS concentration of the input was adjusted to 1%. Cells were first incubated for 1.5-2 hours at 37℃ with RNase (4 µg per 120 µl volume) and then at 65℃ with Proteinase K (100 µg per 120 µl volume) shaking for at least 5 hours.

The de-crosslinked DNA was purified using the PCR clean-up kit according to the manufacturer's protocol. First, the corresponding volume of membrane binding solution (120 µl for the ChIPs) was added to the eluate and then incubated on a silica column for 1 minute and centrifuged down for 1 minute at 16,000 g. Washes were carried out with 700 µl and 500 µl membrane wash solution followed by 1 minute and 5 minutes centrifugation, respectively. DNA was eluted after 1 minute incubation with 100 µl prewarmed RNAse-free water.

Before sequencing ChIPs were tested for enrichment using quantitative PCR. For this the DNA solution was diluted with 100 µl RNAse-free water and 5 µl per reaction were used as a template.The enrichment of ChIP samples was compared to the input using a primer/probe set for *GAPDH* genomic DNA and a region containing two E-boxes in the *PER1* promoter. Primers used for testing ChIP-enrichment in qPCR are listed in Table 2.14.

2.2.7.2. Deep-sequencing

For each ChIP several samples were prepared and only those with the best enrichment, as tested at the *PER1* promoter, were used for sequencing. Immunoprecipitated chromatin (10 ng DNA of 8 independent BMAL1 ChIPs with enrichment levels > 80-fold, 9 ng DNA of a CRY1 ChIP with 10-fold enrichment, and 8 ng DNA of a CLOCK ChIP with 40-fold enrichment) was used for deep sequencing at the Lausanne Genomic Technologies Facility, Switzerland. Sequencing libraries were prepared using the ChIP-seq Sample Preparation Kit (Illumina) according to the protocol supplied with the reagents. Products were quantified using the Qubit fluorometer (Invitrogen). Three lanes of the BMAL1 library, and one lane each of the CRY1 and CLOCK library were sequenced on the Illumina Genome Analyzer IIx using either Single-Read Cluster Generation Kit and 36 Cycle Sequencing Kit v2. Data were processed using the Illumina Pipeline Software v1.40. As a standard control lane a genomic DNA library made with PhiX DNA was included, which was used to set the base calling parameters used by the Illumina software for all of the lanes of the flow cell, regardless of the library type.

2.2.7.3. ChIP-seq data analysis

Sequenced DNA reads (Illumina, 38 bases long reads) were mapped to the human genome (Homo Sapiens Genome Reference Consortium Human Build 37 [GRCh37/hg19; Feb. 2009]) using Bowtie [194] with maximally two mismatches and only one hit allowed on the genome. If several reads from the same library mapped to the very same genomic position and on the same strand, they were considered redundant tags because they might have resulted not by ChIP, but rather by the amplification process during library preparation. Thus, from redundant tags only one read was kept for the rest of the analysis, resulting in 52 million tags for the BMAL1-ChIP, as well as 2.4 million tags for the CLOCK-ChIP and 5.3 million tags for the CRY1-ChIP. Each lane of the BMAL1 ChIP-Seq yielded 1.6×10^7 reads, the CLOCK ChIP-Seq 2.1×10^7, and the CRY1 ChIP-Seq 2.5×10^7.

For each protein separately, BMAL1 or CLOCK or CRY1-bound regions were detected by MACS [195] with the following parameters: shift = 70 bp (BMAL1), 90 bp (CLOCK) and 100 bp (CRY1), bandwidth = 2 x shift, genome size = 2.16 Gb, and an input chromatin sample from human 293T HEK cells as control data. This yielded a total of 9459 regions bound by at least one protein. A refined estimate of the binding site location in each of these regions was then obtained by a deconvolution algorithm that models the expected distribution of tags on the positive and negative strands [5]. This

was done separately for each protein in order to accurately detect the genomic location of the bound protein. For the rest of the analysis, local maxima in the deconvoluted signal were used as the center positions of the binding sites. The deconvolution methods also provides a goodness of fit score that quantifies whether the plus and minus tag distributions fit the expected patterns [5]. This score was used to reject spurious sites. The binding sites were then attributed a score that combines the enrichment in tag counts in window of $\pm$ 100 bp around their center compared to input chromatin, as well as the goodness of fit criterion. Of note, sites that overlapped within a 200 bp region, were extended to cover a larger region, thus resulting in some sites that are longer than 200 bp. The overlaps between the binding sites of the three proteins were computed using the command "intersectBed" from the BedTools suite [196]. Each binding site was annotated with the RefSeq gene having the closest TSS. Visualization of the ChIP-seq and of the detected binding sites for each protein can be found on the UCSC genome browser:

DNA sequences in the binding sites were retrieved from the UCSC Genome Browser database. In order to analyze the E-box composition of binding sites, the sequences for all binding sites were scanned with a weight matrix describing a perfect, highly polarized E-box (probability of 0.97 for each base of the motif and probabilities of 0.01 for the 3 other bases). To compute the spatial correlation in the positions of E-boxes, the likelihood scores returned by the E-box scanning were used and converted into occupancies using a sigmoid transformation with threshold corresponding to two mismatches. The correlation signal was then computed on the occupancies. The enrichment for tandem E-boxes was tested by re-scanning the binding sites sequences using a weight matrix describing 2 perfect, highly polarized E-boxes joined by 6 unspecific (background frequencies) bases.

2.2.8. Microarray

In order to test the obtained binding sites for functionality, gene expression was assayed in large scale on a microarray.

2.2.8.1. Array design

A 60k customized microarray was designed for 6356 genes, which corresponds to about one fourth of the human transcriptome. For each gene 10 different probes were used to increase reliability of the data. The median signal of all significant probes was employed for data analysis. The probes were specific for 1373 genes found closest

to a ChIP-seq binding site. For control two sets of genes without binding sites were tested. One set contained 3480 completely random genes, whereas the other set consisted of 1503 specifically selected genes. This latter category was made up of a large set of genes that seemed potentially to be rhythmic, i.e. 443 genes from ChIP-sequencing that were not the closest gene to a binding site but within $\pm$ 20 kb of it, 78 genes that included *BMAL1*, *NPAS2*, casein kinases and nuclear receptors, 449 genes that were most rhythmic in a previously published microarray analysis in U2OS cells [3] (modified supplementary data), 93 genes that were differentially regulated through RevErbα knockout, 83 genes involved in the cell cycle, 90 genes related to epigenetic regulation, and 83 and 79 genes related to glucose metabolism and oxidative stress, respectively.

2.2.8.2. Microarray sample preparation

Two 48-hour time courses (Z4 and Z5) were performed as described and cell pellets were harvested every 3 hours. RNA was purified and samples were labeled with the Two-Color Low Input Quick Labeling kit (Agilent) according to the provided protocol. Labeling using two colors allowed a comparison of RNA amounts from each of the 16 time points (labeled with Cyanine-5 (Cy5)) to the average of all 16 time points (labeled together with Cy3). The signal for Cy5 was set in relation to the signal from Cy3 (sample relative to average). An internal fluorescent control was provided by adding Spike A and Spike B to the samples. A schematic overview over the procedure is given in Figure 2.8.

Protocol: After RNA-extraction RNA dilutions of 66.6 ng/µl were prepared. In the first step 2 µl of spike A or spike B mix were added in a 1:25600 dilution to 100 ng (1.5 µl) of each RNA sample. These dilutions were prepared in a 4-step serial dilution (1:20, 1:40, 1:16, 1:2) as described in the protocol. In the second step, the labeling reaction was prepared and cDNA was transcribed from the RNA. First, 1.8 µl of T7 Promoter Primer Mix was added to the RNA-samples. The T7 promoter primer is an Oligo dT-primer that binds to poly A tails of RNA and tags them with an additional T7-promoter. This mix was denaturated by incubating it at 65 ℃ for 10 minutes, followed by a 5-minute incubation at 4 ℃. 4.7 µl of the cDNA Master Mix, containing AffinityScript RNase Block mix were added to each sample, which was then incubated at 40 ℃ for 2 hours. During this step a complimentary DNA-strand (cDNA) was produced. An incubation at 70 ℃ for 15 minutes inactivated the Affinity script enzyme. The entire

sample was placed on ice for 5 minutes.

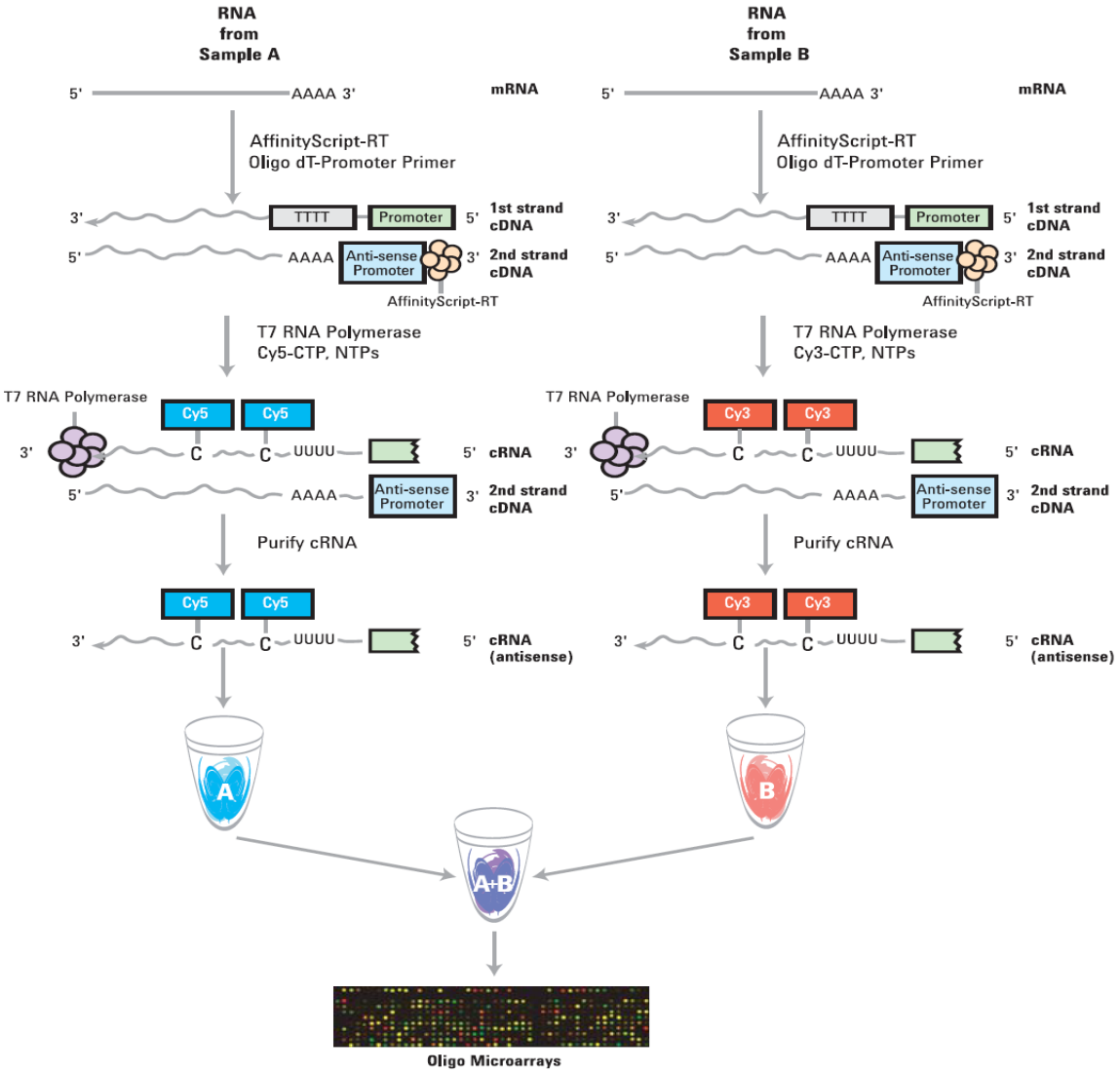

Figure 2.8.: Schematic overview of cRNA-labeling for the microarray as provided by the manufacturer (Agilent). The procedure is described in the text.

Next, 6 µl of the T7 RNA polymerase Transcription Master Mix were added to the samples, which were then incubated for 2 hours at 4 ℃ during which T7 RNA Polymerase replaced the complimentary DNA strand with a complimentary RNA strand (cRNA) at the same time incorporating Cy3- or Cy5-labeled CTP. In the third step the labeled RNA was purified using the Qiagen RNeasy mini spin columns as described. The RNAs were eluted in 30 µl RNase-free water and quantified using the Microarray Measurement tab in the NanoDrop software. The concentrations of cRNA (ng/µl) and Cy3 or Cy5 (pmol/µl) were used to determine the specific activity of either Cy3 or Cy5

per µg cRNA. For hybridization to the microarray 300 ng Cy3-labeled cRNA and 300 ng Cy5-labeled cRNA time point-average were combined in the fragmentation mix and incubated at 60℃ for exactly 30 minutes to fragment the RNA. After cooling on ice 25 µl cRNA from the fragmentation mix were complemented with the same volume of 2x GE Hybridization buffer HI-RPM. After spinning them down the samples were immediately loaded onto the microarray and hybridized for 17 hours at 65℃ rotating at 10 rpm. After hybridization the array was washed twice for 5 minutes in acetonitrile, several times shortly with Milli-Q water, 1 minute in GE Wash Buffer 1, and 1 min with GE Wash Buffer 2 which had been prewarmed to 37℃. Then the slides were put into a slide holder and scanned immediately.

2.2.8.3. Microarray data analysis

After rejecting probes with too low absolute hybridization signals (less than 20 fold of the median signal of negative control probes), the raw ratio data of the two independent 48 h time courses was analyzed. Genes with less than two valid probes were rejected from the analysis. For each gene, the median ratio of all remaining probes was then used as relative expression measure of one probeset at each time point. For the identification of rhythmic probesets, the Fisher test was used as described [5, 197]. To be stringent the following combined filters were used: the Fisher test p-value needed to pass a threshold of a FDR (False Discovery Rate) [198] of 20% in at least one of the two arrays, and of 50% in both arrays; the amplitude needed to be higher than 0.25 in log2 scale on both arrays, which corresponds to a 1.2-fold change from peak to trough; given that the signal for one gene consisted of the median of at least 3 probes and at most 10 probes, sufficient correlation between these probes was ensured by requiring that the temporal average of the standard deviation between the probes does not exceed 33% of the amplitude of the median signal; the phase of a given gene needed to be comparable in both arrays, allowing for a deviation of $\pm$ 3 h between both arrays. Together, this analysis yielded a set of 118 robustly rhythmic transcripts.

3. Results

3.1. Circadian regulator binding site analysis

3.1.1. Sequencing

Three lanes of the BMAL1 ChIP and one lane each of the CLOCK- and the CRY1 ChIP were sequenced. In each lane of the BMAL1 ChIP about 16 million clusters passed the filter, which corresponds to the number of sequence tags that were obtained. From the CLOCK- and the CRY1 ChIP 21 and 25 million clusters, respectively, passed the filter (Table 3.1).

Table 3.1.: Information on sequenced reads. Of note: BMAL1-ChIP was sequenced on three lanes, whereas the CLOCK-ChIP and the CRY1 ChIP were sequenced on one lane each.

ChIP	Read Length (bp)	Loading conc. (pM)	Pass Filter Clusters (million)
BMAL1	38	12	16
	38	12	16
	38	12	16
CLOCK	38	10	21
CRY1	38	10	25

3.1.2. Binding site annotation

Circadian regulator binding sites (CRBSs) were annotated as described in detail in the methods section. During the sequencing process 38 bp were sequenced from the 5'-end of both, the leading and the lagging strand of the DNA-fragment. This should yield sequencing profiles for both strands that are similar but slightly offset to one another. Reads from each strand were mapped separately to the human genome and binding sites were detected for each ChIP-seq individually using several algorithms taking into

consideration the shape of the peaks on each DNA-strand. Finally, the binding sites were extended to a window of $\pm$100 bp and attributed a score which correlated with the enrichment of reads and with the agreement to the expected distribution on both strands. This analysis yielded a total of 3040 CRBSs, 2048 of which were bound by BMAL1, 787 by CLOCK, and 1063 by CRY1. All binding sites were mapped on the UCSC genome browser and annotated to the closest gene, resulting in a total of 1454 genes with CRBSs. Among these, 956 genes had only one binding site, whereas the remaining 498 genes harbored 2084 sites. Most known clock-controlled genes, such as *PER1, PER2, PER3, CRY2, NR1D1, DEC1, DEC2, DBP*, and *TEF* possessed 2-4 binding sites, whereas *CRY1* and *NR1D2* had only one. All these known circadian genes had at least one common site for BMAL1 and CLOCK binding, in most cases even for BMAL1, CLOCK, and CRY1. The ability to detect binding at the known clock-controlled genes shows that the data analysis was reliable.

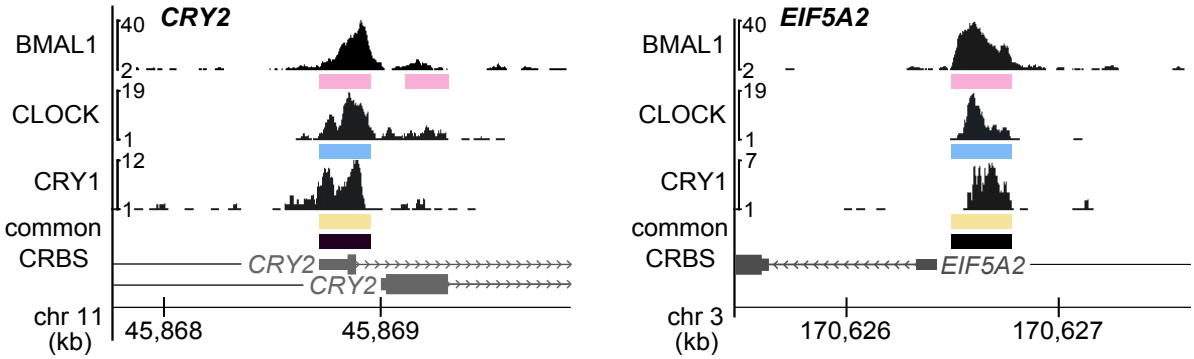

Figure 3.1.: UCSC browser views of BMAL1, CLOCK, and CRY1 occupancy at the promoters of *CRY2* (left) and *EIF5A2* (right). Each black dot represents one sequenced read. Data was scaled as indicated on the left of each ChIP-representation. Regions detected as binding sites of the individual transcription regulators are marked by colored bars. Black bars show common CRBSs of the three regulators. In both examples the peaks are close to the transcription start site.

Two examples of ChIP-seq data with binding sites for all three transcription factors, as presented on the UCSC genome browser, are shown in Figure 3.1. The left panel displays the alignment of sequenced tags around the Transcription Start Site (TSS) of the known clock-controlled gene *CRY2*, while the right panel shows the alignment around the TSS of *EIF5A2*, a gene that was so far not known to be regulated by the circadian clock. *CRY2* contains a second BMAL1 peak downstream of the gene (pink bar). Although it looks like there could be some enrichment from the CLOCK- and the

CRY1-ChIP as well, this example demonstrates, that the parameters for detecting a peak were relatively stringent in order to reduce the amount of false positives and that in this case either the enrichment or the shape of the second peak for CLOCK and CRY1 must have been below the set parameters for peak detection.

3.1.3. Distribution of CRBSs

The binding sites for the individual clock proteins showed highly significant though imperfect overlap (Figure 3.2). The overlap between BMAL1 and CLOCK (214 CRBSs) as well as between all three ChIPs (179 CRBSs), altogether accounted for 13% of the binding sites. Of note, about 50% of the CLOCK binding sites overlapped with BMAL1, whereas only 19% of BMAL1 binding overlapped with CLOCK. About 44% of the CRY1 CRBSs overlapped with sites for BMAL1 and/or CLOCK.

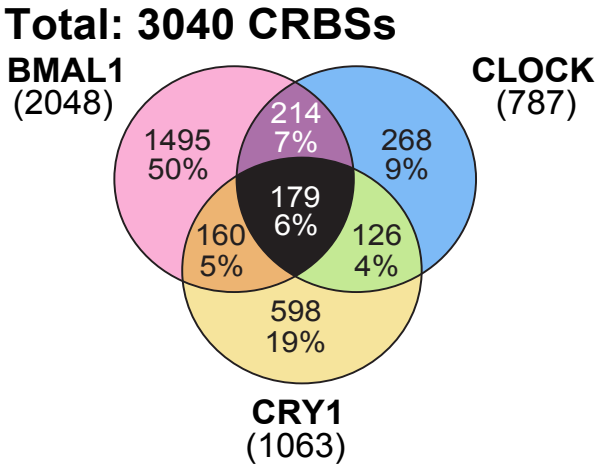

Figure 3.2.: Venn diagram showing numbers and percentage of individual and overlapping binding sites of BMAL1, CLOCK, and CRY1. The total number of binding sites is given in parenthesis under the name of each ChIP-seq, percentages represent the fraction of the CRBSs over the total number of sites for the three proteins.

3.1.4. Localization of binding sites

Binding sites showed a pronounced enrichment at transcriptional start sites, with 14.6% within 1 kb of a TSS, suggesting a role in gene expression. In contrast, a simulated random set of binding sites provided only 0.1% within 1 kb of a TSS (Figure 3.3 top panels). Individual analysis of the three ChIP-seq data sets (Figure 3.3 bottom panels) showed a strong accumulation for BMAL1 and CLOCK at the TSS, while CRY1

was only weakly enriched there. Notably, the strongest enrichment around the TSS occurred at common sites between two or all three ChIPs. Binding sites detected for BMAL1 only were also highly enriched at the TSS, whereas CLOCK-only binding sites displayed a broader distribution. CRY1 sites that coincided with binding for BMAL1 and/or CLOCK showed a slight enrichment around the TSS, whereas CRY1-only binding sites were randomly distributed.

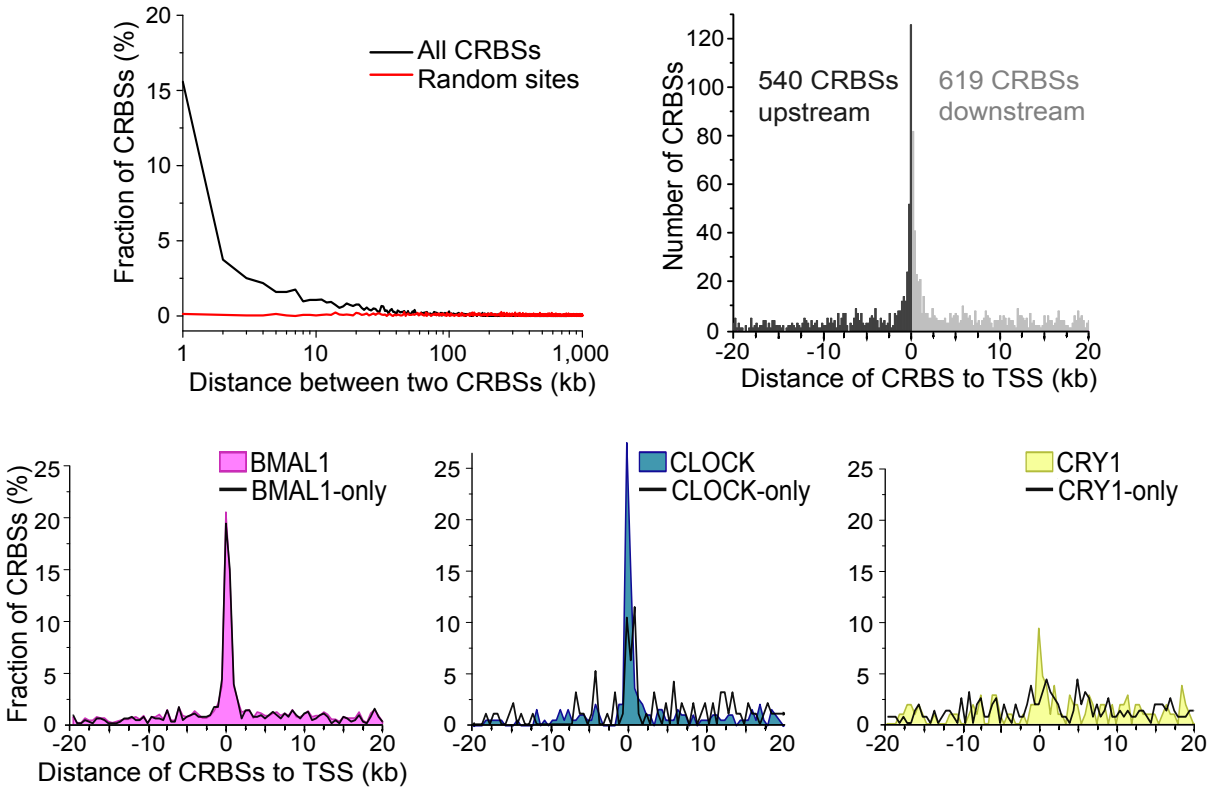

Figure 3.3.: Binding sites for BMAL1 and CLOCK, but not for CRY1, were highly enriched around the transcription start site (TSS). Upper panel left: Histogram of the distance between two consecutive CRBSs (black) in comparison to a set of 3040 random sites (red). Upper panel right: CRBSs cluster near TSSs. Histograms of positions of all CRBSs are shown for a window of $\pm$ 20 kb around TSSs with a bin size of 200 bp. The numbers of CRBSs upstream and downstream of the TSS are indicated. Bottom panels: Histograms of individual regulators (bin size 1 kb) show that CRBSs for BMAL1 and CLOCK are enriched at the TSS. The black line in each plot marks the distribution of binding sites found only for the respective protein. Common CRBSs for two or all three regulators are most strongly enriched around the TSS suggesting a relation to transcription. This enrichment is less for binding sites found for CLOCK only and essentially absent in CRY1-only binding sites.

The vast majority of CRBSs was located in intergenic regions (49.3%) and introns (28.4%), 7.5% of the CRBSs were in promoters, 10.0% in 5'-UTRs, 2.3% in exons and 2.4% in 3'-UTRs (Figure 3.4). However, when taking into consideration that intergenic regions and introns make up 64.1% and 27.1% of the genome, respectively, while the contribution to the genome of promoters (0.6%), 5'-UTRs (4.8%), exons (1.8%) and 3'-UTRs (1.9%) is considerably small, CRBSs were slightly under-represented in intergenic regions (0.8x) and highly enriched in promoters (12.1x) (Figure 3.4, right). They were slightly enriched in 5'-UTRs (2.1x), but occurred approximately at the expected rate in exons (1.3x), introns (1.1x) and 3'-UTRs (1.3x). The enrichment in 5'-UTRs reflects the general tendency of CRBSs to accumulate in a narrow region of about $\pm$ 1 kb around TSSs.

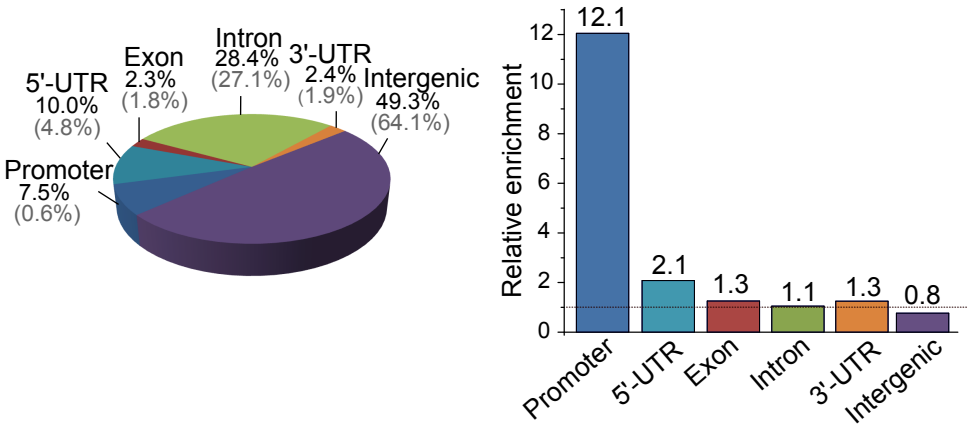

Figure 3.4.: Location of binding sites. The pie chart (left) shows the percentage of binding sites in a genomic region (black) and the contribution of the region (%) to the genome (grey). The bar chart (right) displays the relative enrichment of CRBSs in a genomic region. CRBSs are highly enriched in promoters. The center of a binding site was assigned to its localization in the gene: promoter (-1 kb to TSS), 5'-UTR, exon, intron, 3'-UTR, intergenic.

3.1.5. Binding motif analysis

It was published, that in the mouse BMAL1 and CLOCK bind as a heterodimer to E-boxes or tandem E-boxes with a spacer of 6 or 7 nucleotides [5]. Binding motif analysis on the dataset from this work supports these findings in humans as well. The E-box motif with the consensus sequence CACGTG, was highly enriched among the binding sites (9.8-fold). Interestingly though, the sequence logo for the consensus sequence (Figure 3.5 upper panel) suggested, that among the E-box motifs in the CRBSs the

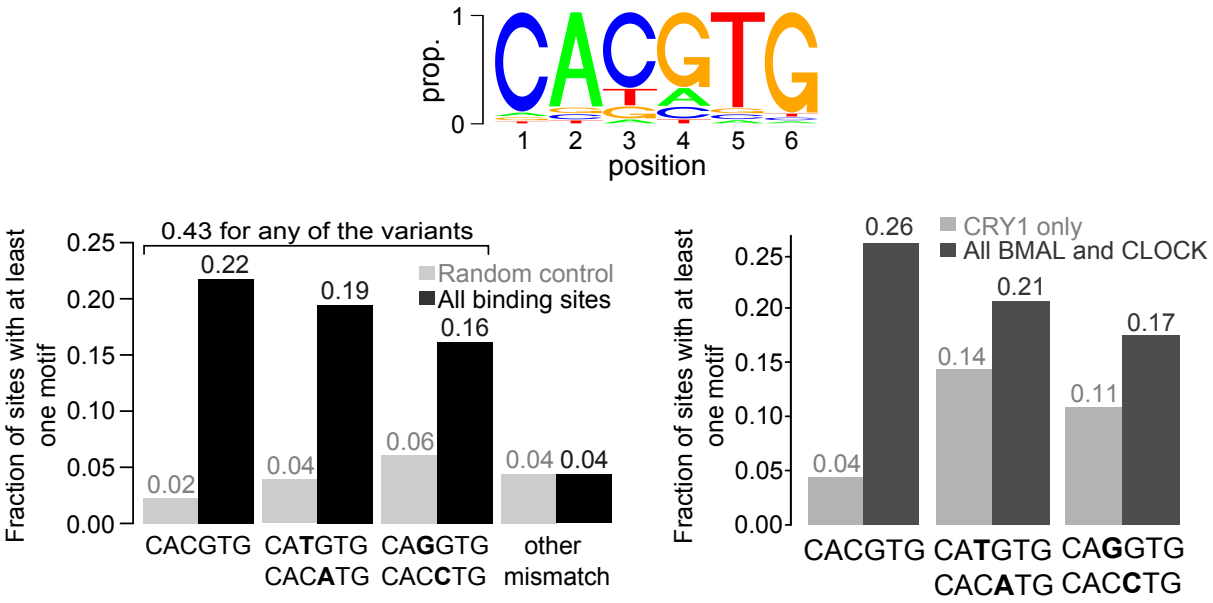

Figure 3.5.: The E-box motif was enriched in BMAL1- and CLOCK-, but not in CRY1-binding sites. Top panel: CRBSs with imperfect E-boxes (1 mismatch) display preferential mismatches at the central positions. The logo shows average proportions of bases. Bottom panel left: Fraction of CRBSs containing the indicated E-box motifs (black). To control for bias due to sequence composition, the sequences were randomly shuffled and analyzed similarly (grey bars). The three main variants cover 43% of all CRBSs. Bottom panel right: Fractions of CRY1-only sites (light grey bars, n = 598) and BMAL1 and/or CLOCK sites (dark grey bars, n = 2442) containing the indicated E-box motifs.

central positions were less conserved. Therefore the fraction of sites containing E-boxes with mismatches in these positions was calculated (Figure 3.5 bottom panel left). In order to not produce a strand-related bias because the canonical E-box is palindromic, whereas E-boxes with one mismatch are not, the (+)-strand was examined for the binding motif as well as for its reverse complement. Interestingly, a similar amount of CRBSs harbored an E-box (22%) as the motifs CA**T**GTG and CAC**A**TG (19%), an E-box with a pyrimidine-to-pyrimidine- or a purine-to-purine-exchange in position 3 and 4, respectively (hereafter referred to as E*-boxes). 16% of the CRBS contained the motifs CA**G**GTG and CAC**C**TG, with a double-guanine or double-cytosine in positions 3 and 4 (hereafter referred to as E-GG-boxes). There was no significant fraction of binding sites with a mismatch in any other position. Of note, only 43% of the CRBSs featured any of the three E-box-motifs at all, but some sites possessed several motifs.

Interestingly, these three motifs were differently distributed among BMAL1, CLOCK, and CRY1 binding sites. Table 3.2 shows that 28% of each, BMAL1 and CLOCK

Table 3.2.: Fraction of binding sites (in %) that harbor an E-box, E*-box or E-GG-box.

E-box type	BMAL1	CLOCK	CRY1
E-box (CACGTG)	28.5%	28.1%	8.3%
E*-box (CA**T**GTG/CAC**A**TG)	20.7%	25.7%	20.7%
E-GG-box (CA**G**GTG/CAC**C**TG)	17.9%	17.3%	11.6%

CRBSs harbored an E-box, whereas 21% and 26%, respectively, contained an E*-box. About 18% of both possessed the E-GG-box. At the same time only 8% of the CRY1 binding sites harbored a perfect E-box, whereas 21% contained an E*-box and 12% an E-GG-box. The relative amount of E*-boxes in CRY1 sites was thus the same as in BMAL1 and CLOCK sites, whereas E-GG-boxes were present slightly less.

Figure 3.5 (bottom panel right) shows, that CRY1-only CRBSs harbored less E*-boxes (14%) than CRY1 sites overlapping with BMAL1 and/or CLOCK, but that the canonical E-box was virtually absent (4%). E-GG-boxes, however, were equally present in all CRY1-sites (11%).

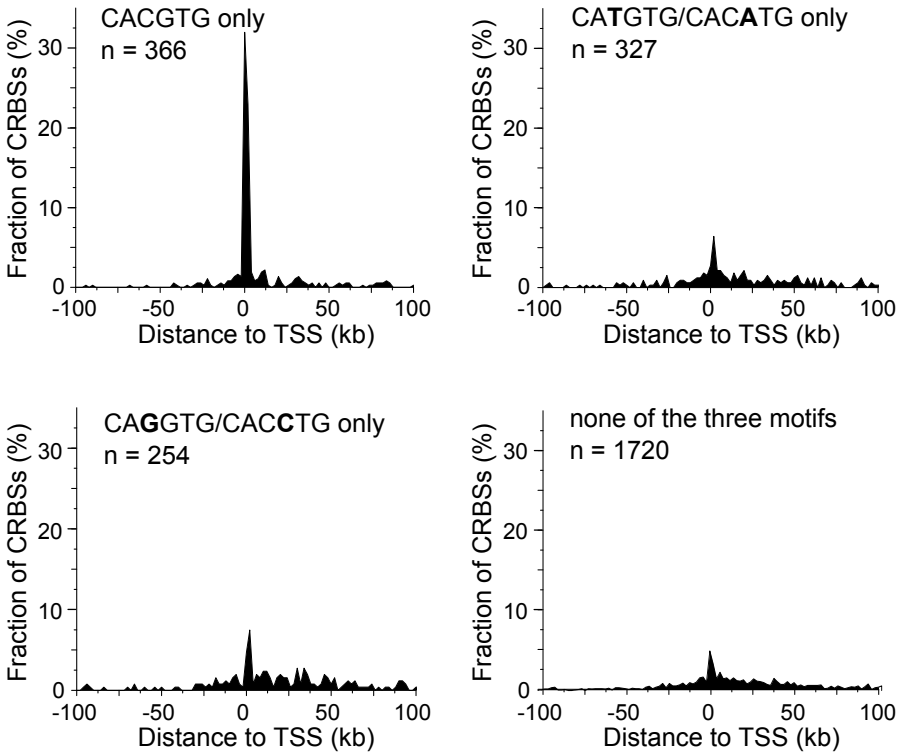

Figure 3.6.: Binding sites harboring a perfect E-box were strongly enriched around the TSS. Histogram of CRBSs containing only the canonical E-box motif (top left), the E*-box (top right), or the E-GG-box motif (bottom left) or none of these (bottom right). The distance was calculated in 2 kb-bins.

Of interest, CRBSs that only contained the perfect E-box motif were highly enriched around the TSS of genes, whereas binding sites that either only harbored the E*-box or the E-GG-box motif, were not enriched around the TSS and were distributed like sites not containing any of these motifs at all (Figure 3.6).

Similar to the published data in mouse [5], auto correlation analysis on the CRBSs from this work found the highest probability for a repeating tandem E-box sequence at a spacer of 6 nucleotides (Figure 3.7, left-hand side). The average sequence logo for tandem E-boxes with zero to three mismatches was computed and the total number of tandem motifs was plotted (Figure 3.7, right-hand side).

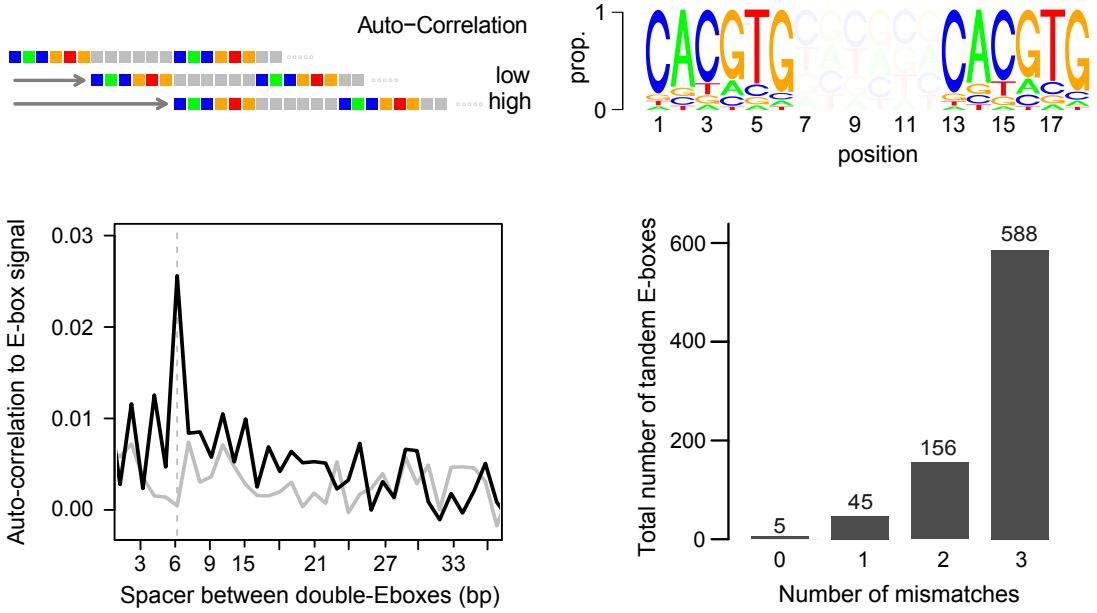

Figure 3.7.: Occurrence of tandem E-box motifs in CRBSs. Auto-correlation showed an enrichment of tandem E-boxes with a spacer of 6 bp. The optimal spacer length was calculated by shifting the entire sequence step by step by one nucleotide (upper left panel). The fit between the original E-box (black) and the next E-box was calculated by auto-correlation. To control for sequence composition bias the sequences were randomly shuffled and analyzed again (grey). The sequence logo for the best-fitting E-box motif with $\leq$ 3 mismatches is shown in the upper right panel. The lower right panel shows the total number of tandem E-boxes found in the CRBSs.

There were only few tandem motifs with no (5) or one (45) mismatch, but there were 156 tandem motifs with two mismatches and 588 with three mismatches. Some CRBSs featured multiple tandem E-box motifs, overall yielding 555 (18.2%) CRBSs with at least one tandem motif with up to three mismatches. This is >5-fold above what would be expected in a random data set. The core clock genes *DBP, DEC1, DEC2,*

PER3, and *TEF* had one binding site each which harbored a tandem E-box with three mismatches, the other core clock genes did not contain a tandem E-box in their binding sites at all.

CRY-genes have been reported to bind to glucocorticoid receptors and to repress transcription at glucocorticoid response elements (GREs) in target genes [120]. This finding cannot be supported by the data at-hand. Analysis of the consensus GRE-motif in the CRBSs showed that in fact only a small fraction of CRY1-sites contained a GRE (Figure 3.8) and that there was only a very slight enrichment in comparison to the randomly shuffled control. A higher fraction of BMAL1 and CLOCK binding sites harbored a GRE, however, these seem to be sequence biased for the motif, because in comparison to the randomly shuffled control, again, there is only a slight enrichment in GREs among the CRBSs.

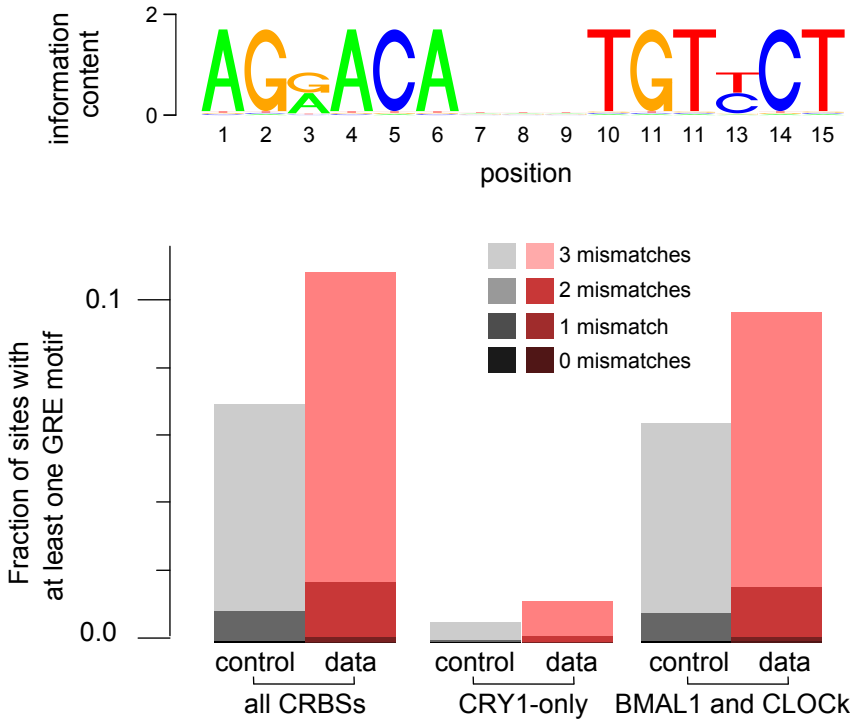

Figure 3.8.: There is no enrichment of the GRE-motif in the binding sites. Upper panel: GRE sequence logo. Lower panel: Fraction of sites harboring a GRE-motif with 0-3 mismatches (as indicated by the color chart) for all CRBSs (left), sites found for CRY1-only (center), all BMAL1 and CLOCK sites (right). To control for sequence bias the sequences of the binding sites were randomly shuffled and analyzed again (grey bars).

3.1.6. Binding site clusters

Finally, we observed that the CRBSs were, unlike data obtained from mouse liver, not evenly distributed over the DNA. Some chromosomes had a substantially higher amount of binding sites, especially towards the outer regions. Histograms of the CRBSs show that especially on chromosomes 5, 7, 8, 13, and 14 there appear to be regions of about 10 Mb with abnormally high densities of binding (Figure 3.9). Since the work was performed in a tumor cell line, it may be that these regions with clustered binding sites could be a result of amplification of certain areas. Amplified regions should yield a higher background in ChIP-sequencing because unspecific binding can occur proportionally more often. Regions with high density of CRBSs were compared to the remaining portion of the chromosome (Table 3.3) and the difference in background tags was computed bioinformatically. Only the data from BMAL1 ChIP-seq was used, because sequencing was deepest there, yielding a higher coverage of the background. BMAL1-tags were compared to a sequenced sample of 293T HEK DNA, where no chromosome alterations have been reported. Interestingly, for all tested regions, the enrichment was only 1.5 to maximally 3-fold, which by itself cannot account for up to 100x more binding sites in these regions. Moreover, a comparison of pure genomic DNA from U2OS and HEK cells using qPCR on several regions on the chromosome with the highest enrichment, chromosome 7, yielded a similar enrichment of up to 3-fold (data not shown).

Table 3.3.: Regions with extraordinary high densities of binding sites were found in U2OS cells. Most of these sites were localized at the outer regions of chromosomes.

Chromosome	Region	Localization	Enrichment
chr 5	0 - 10 Mb	beginning	1.5x
chr 7	0 - 10 Mb	beginning	3x
chr 8	140 - 150 Mb	end	1.6x
chr 13	110 - 115 Mb	end	1.5x
chr 14	60 - 80 Mb	centered	1.7x

Histograms of E-boxes in chromosomes 5, 7, 8, 13, and 14 showed, that in many cases the regions with extremely high amounts of binding sites also contained a higher-than-average number of E-boxes. It appears possible that the occurrence of slightly amplified chromosomal regions, in combination with a high amount of E-boxes resulted in the observed strong clustering of binding. Figure 3.9 displays the histograms of CRBSs and E-boxes in the chromosomes with extraordinary high binding site densities.

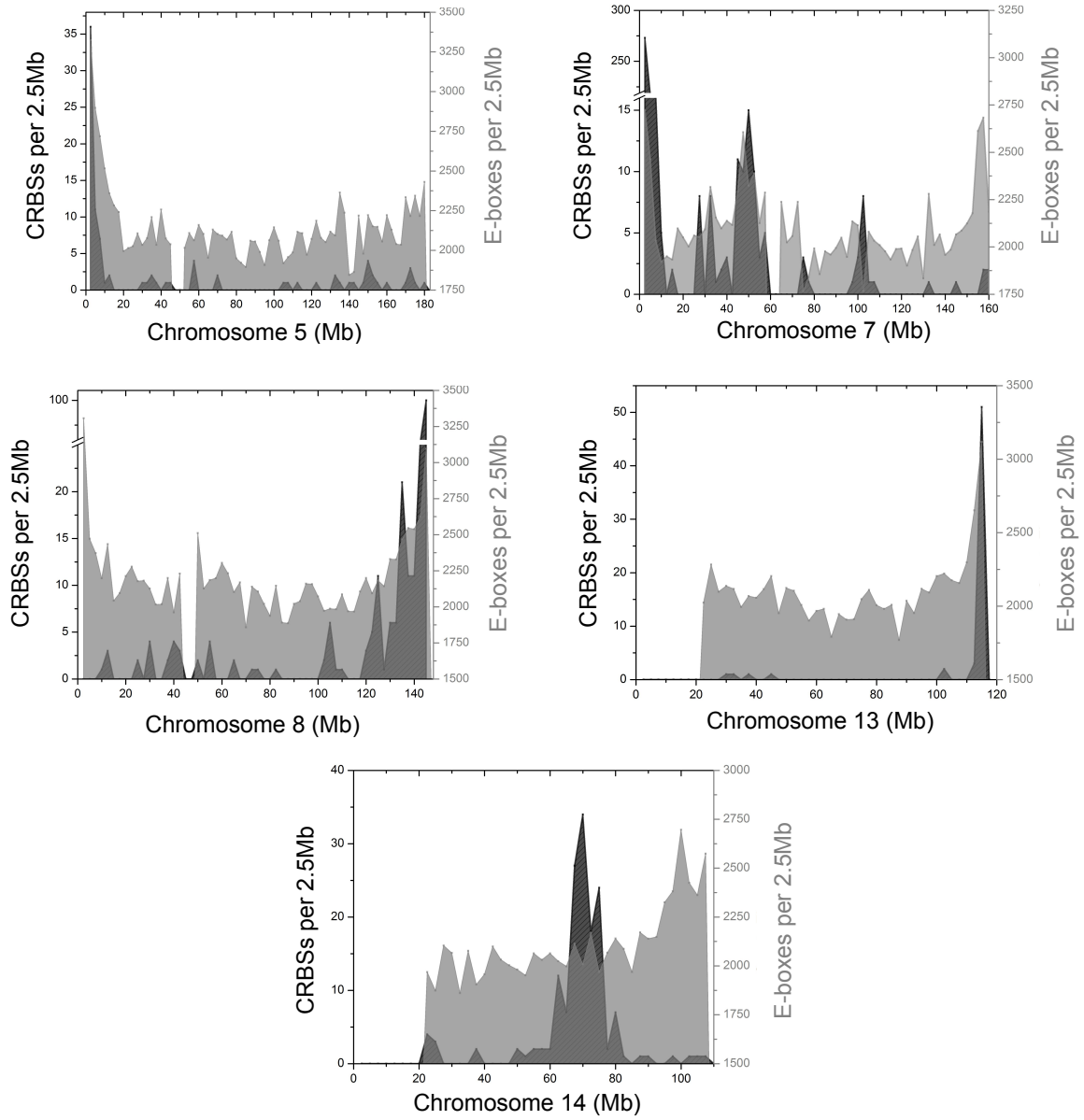

Figure 3.9.: Clustered binding sites can often be found in regions with high E-box densities. Histograms for the number of CRBSs (dark grey, left y-axis) and the number of E-boxes (light grey, right y-axis) were made with a bin of 2.5 Mb for the chromosomes 5, 7, 8, 13, and 14, where regions with exceptionally high binding site densities were found. Note that for some regions no sequence data is available, which is shown by the lack of a count of E-boxes.

3.2. Expression analysis revealed about 1% rhythmic genes in U2OS

Each CRBS identified by ChIP-seq was attributed to the gene with the closest TSS. By this means 1454 genes with binding sites were identified. Next, we analyzed whether genes with a binding site show circadian mRNA expression rhythms. Two microarray-experiments were performed with RNA from the time courses Z4 and Z5 in temperature-entrained U2OS cells that were harvested every three hours over two consecutive days.

A custom-made 60k-microarray was designed (see Methods). The probes were specific for 1373 genes found closest to a ChIP-seq binding site. For control, two sets of genes without binding sites were tested. One set contained 3480 completely random genes, whereas the other set consisted of 1503 specifically selected potentially rhythmic genes. Among these were 443 genes from ChIP-sequencing that were not the closest gene to a binding site but nevertheless harbored a CRBS within $\pm$ 20 kb of their TSS. 1285 genes with CRBSs (93.6%), 3059 random genes (87.9%) and 1364 (90.8 %) of the group of selected genes without CRBSs were expressed.

For data analysis three main criteria from chronobiology were used to define rhythmic genes: a 24h-oscillation, amplitude and phase (Figure 3.10).

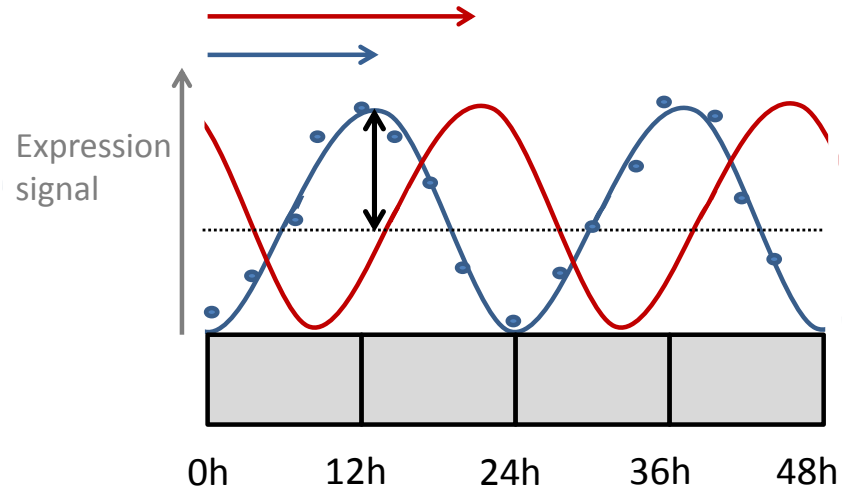

Figure 3.10.: Simplified overview over the main criteria used to define rhythmicity. The rhythm of expression can be approximated by fitting a cosine-curve with a 24h-period to the data points. The amplitude of gene expression is depicted by the black arrow and marks the influence of the circadian clock on expression of that particular gene. The phase is determined by the first maximum of the fitted cosine-curve (blue arrow represents phase ZT 12, red arrow represents phase ZT 23).

The first criterion is a prerequisite for rhythmicity, because a clock-controlled gene should oscillate with a period of about 24 hours. For this, the Fisher test was used as described [5, 197]. Roughly, this mathematical method describes how well the data fits a cosine curve. The curve was defined to contain a 24-hour component for the 24-hour rhythmicity, as well as a 48-hour component, which would allow to fit a dampening curve. The other two criteria define the rhythmicity more precisely. The amplitude provides information on how strongly a gene is regulated by the circadian clock. In this work it is defined as one half of the value between the peak and the trough of the fitted cosine curve, if not specified otherwise. The phase is set in relation to a *Zeitgeber* and enables a comparison of the time point of expression between two or more genes. In this work, the phase is defined as the first peak in expression 24h after release to 37°C.

Rhythmic genes were defined applying a set of stringent criteria, which reflect these key factors and their comparability between two independent data-sets (see methods for more detail). Thus, a set of 118 oscillating genes was identified, 58 with a CRBS, and 60 genes without (Figure 3.11).

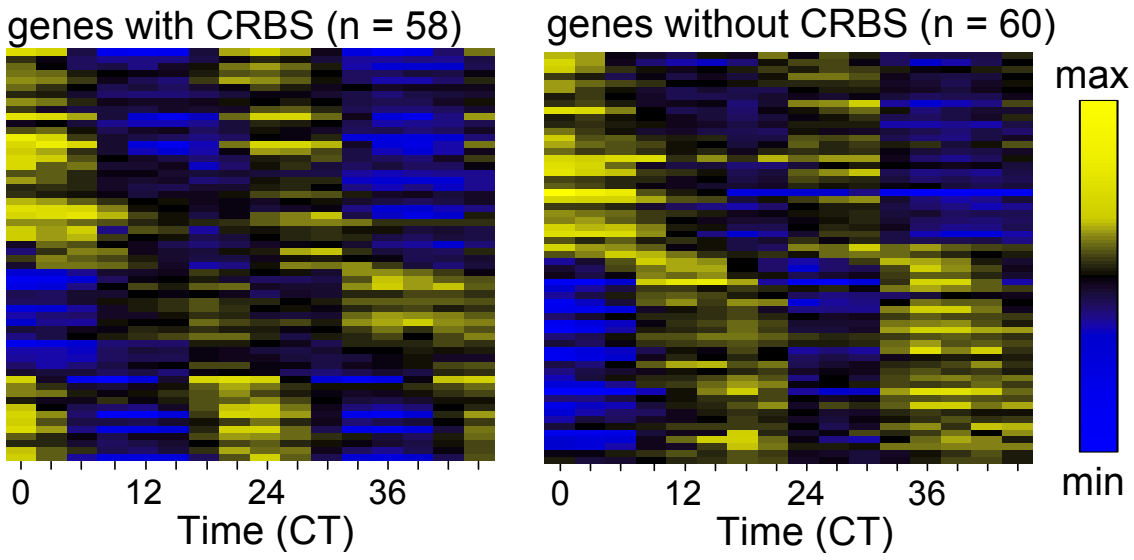

Figure 3.11.: Heatmap view of the 24h cycling genes from the microarray ordered by phase of expression. The median expression level of each time point is shown in relation to the mean of all time points (upregulation: yellow, downregulation: blue). The left panel represents rhythmic genes with an assigned binding site, the left panel circadian genes without a CRBS.

Among the rhythmic genes with a CRBS were the core clock genes *CRY1, CRY2, DBP, DEC1, DEC2, PER1, PER2, PER3, REVERB*α*, REVERB*β and *TEF*, which are known to be regulated by BMAL1/CLOCK. In the group of oscillating genes without binding sites were, as expected, *BMAL1, E4BP4* and *NPAS2*. Thus, virtually all core clock genes were identified, indicating that the expression analysis was suitable for the detection of circadian genes.

A plot of the amplitude versus the circadian phase color-coded by the degree of rhythmicity revealed a biphasic profile of expression with two peaks around CT 15 and CT 3 (Figure 3.12 top panel) for highly rhythmic genes. The expression of circadian genes without CRBSs was enriched in two sharp peaks at CT 15 and CT 3, whereas the phases of rhythmic genes with CRBSs peaked at CT15 and to a larger extent at CT21 - CT6. The latter peak contained all core clock genes that are known to be controlled by BMAL1/CLOCK (Figure 3.12 middle and bottom panel). Examples for rhythmic and non-rhythmic genes are shown in Figure 3.13.

The microarray data permits a comparison between genes with and without a CRBS. 4.2% (58 of 1373) of the genes with a binding site were rhythmic. In contrast, only 0.9% (33 of 3480) of the set of randomly chosen genes were expressed in a circadian fashion. Considering this fraction as the rate by which genes become rhythmic in an indirect manner, it can be estimated that about 12 of the 58 genes with a CRBS oscillated for reasons other than their binding site. The remaining 27 rhythmic genes are from the group of specifically selected genes without a BMAL1-binding site like *BMAL1* and *E4BP4*, and other potentially rhythmic genes. Of note, from the 443 genes that did have a CRBS within $\pm$ 20 kb of their TSS, but were not the closest to a binding sites, only 4 genes (0.9%) were rhythmic. Therefore these genes were not more often rhythmic than the random genes. If a gene contained at least one binding site close to the TSS (within $\pm$ 1 kb) there was an 8-fold higher probability for the gene to be rhythmically expressed as compared to the random genes, whereas genes with only distant binding sites (> 1 kb of the TSS) had a 3-fold higher probability (Figure 3.14).

E-boxes and E*-boxes in the CRBS were, although enriched in general, similarly present in rhythmic and non-rhythmic genes. Out of the 1373 genes with a CRBS 791 (58%) possessed at least one binding site with an E- or E*-box. From the 58 rhythmically expressed genes with CRBSs, a similar fraction, 36 genes (62%), featured an E- or E*-box in at least one site. About 5% of the genes with an E-box (30 of 571) or E'-box (21 of 431) containing CRBS were expressed in a circadian fashion.

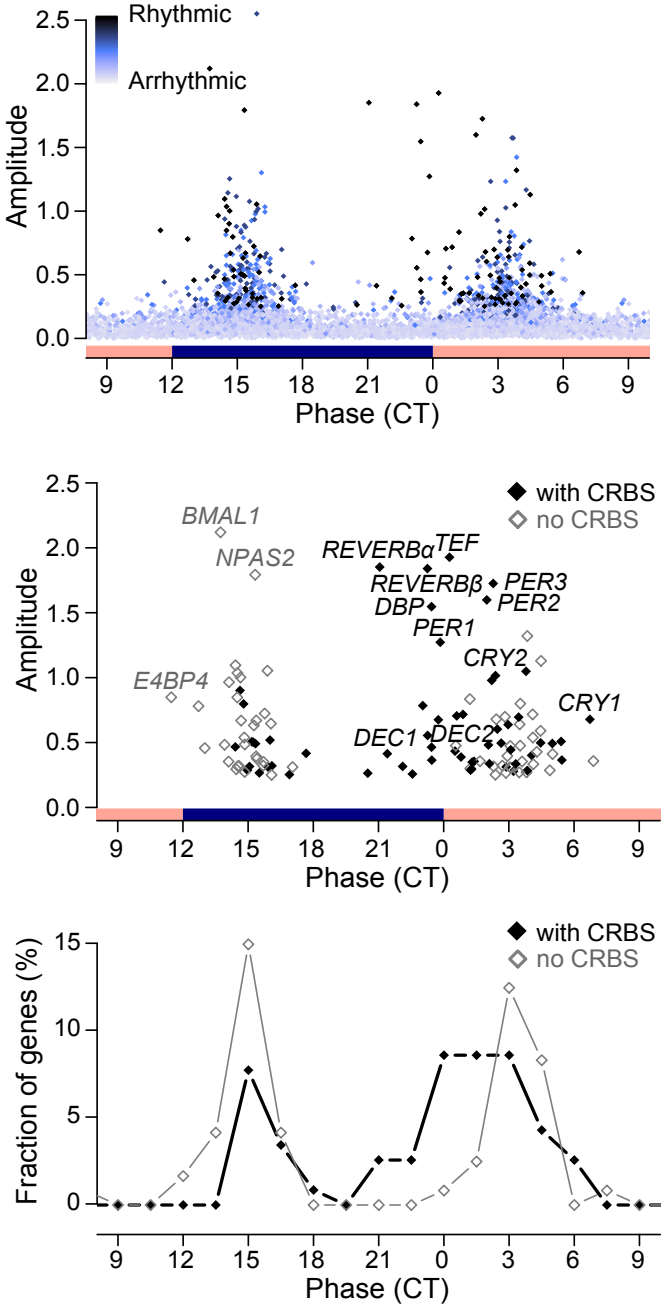

Figure 3.12.: Rhythmic genes show a tendency to cluster in two phases. Analysis of temporal expression profiles of 5708 expressed genes in two microarray replicates identified 118 common genes with diurnal expression rhythms. Data from time course Z5 is shown. Top: The amplitudes of expressed genes were plotted versus the circadian phase. The 118 rhythmic genes are indicated by black symbols. The shade of blue corresponds to the rhythmicity of a gene (light blue = low 24 h-rhythmicity and/or phase undefined; dark blue = high 24 h-rhythmicity and reliable phase). Middle: The amplitudes of the 118 rhythmic genes are plotted against the phase. Genes with CRBSs are shown with black diamonds, the other rhythmic genes with gray diamonds. Core circadian clock genes are indicated. Bottom: Phase distribution of rhythmic genes with and without CRBSs.

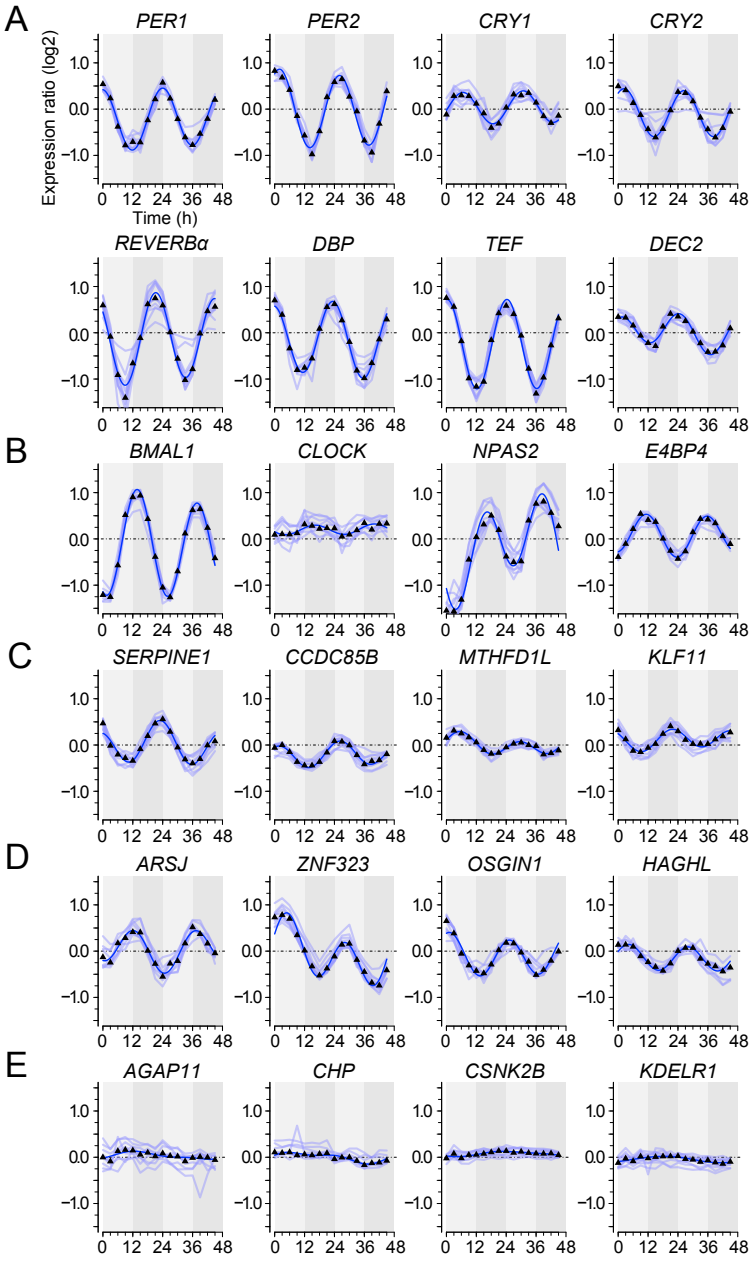

Figure 3.13.: Examples for microarray time course data from temperature-entrained U2OS cells. RNA levels of U2OS cells were analyzed in 3 h intervals over a period of two days. Expression ratios of each time point (0 - 45 h) relative to the average were calculated and plotted on a log2 scale versus the time (h). For each gene up to 10 probes were spotted on the arrays, represented by light blue lines. The black triangles and the fitted dark blue sine curve correspond to the median of the data. Light and dark areas in the background indicate subjective day and night, respectively. (A) Clock genes with CRBSs. (B) Clock-genes without CRBSs. (C) Rhythmic genes with CRBSs. (D) Rhythmic genes without CRBSs. (E) Genes not expressed in a rhythmic fashion. AGAP11 harbors a highly enriched binding site for BMAL1, CLOCK, and CRY1, CHP has a CRY1 binding site. In both genes the CRBSs were close to the TSS. CSNK2B and KDELR1 do not have a CRBS in a window of $\pm$ 20 kb.

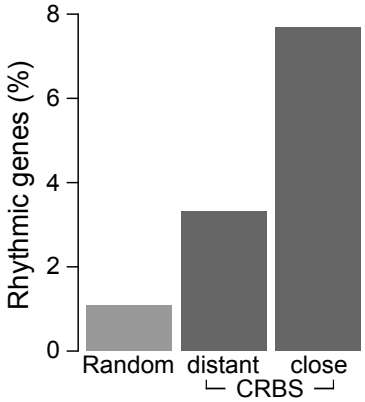

Figure 3.14.: Impact of the presence of a CRBS on rhythmicity. Fractions of rhythmic genes without CRBSs (light grey) and genes with a CRBS (dark grey) located distant (n = 1021) and close (n = 352) to the TSS (within $\pm$1 kb).

3.3. Genes with a binding site have the potential to be rhythmic

Among the oscillating genes, core clock genes were generally expressed with considerably higher amplitude rhythms than other clock-controlled genes. The majority of genes with CRBSs was not expressed in a circadian fashion, not even genes with high-scoring CRBSs in their promoters, such as *ATG3*, *EIF5A2*, or *SCN5A*. This is in contrast to mouse liver, where nearly all strongest binding sites were linked to rhythmic transcript expression [5]. Closer inspection of the microarray data revealed, that some genes might have been oscillating with a too low amplitude to be detected by the applied criteria, excluding possibly rhythmic genes like *CLOCK* (Figure 3.13 B), *EIF5A2*, and *SCN5A* (Figure 3.17 A). The gene *ATG3* did not appear to be expressed in a rhythmic fashion at all (Figure 3.17 A). If transcription is controlled by the circadian clock, but for example the half-life of an mRNA is very long, it is possible, that the preRNA of a gene is rhythmic, whereas the processed mRNA is not. The probes on the microarray were designed to detect mature mRNA only, therefore the expression of the pre-RNA of some of the arrhythmic genes from the array was tested using qPCR on the same time course that were used for the arrays. However, although the preRNA of *BMAL1* was circadian, the preRNAs of the tested genes were not more rhythmic than before (Figure 3.15).

In order to gain more insights into this phenomenon U2OS cell lines were generated, that expressed the destabilized *luc2P* under control of the promoters (about -1

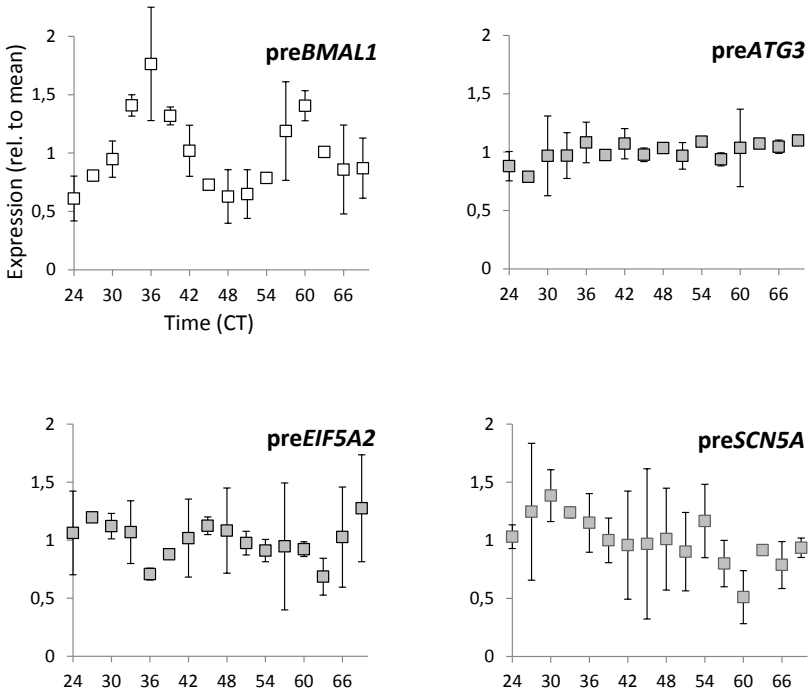

Figure 3.15.: The preRNA levels of *ATG3*, *EIF5A2*, and *SCN5A* are not rhythmic. PreRNA-expression *BMAL1*, *ATG3*, *EIF5A2*, and *SCN5A* was tested using qPCR on the two time courses (Z4 and Z5) that were assayed on the microarray. The data of each time course was normalized to the mean.

kb to the TSS) of these genes. Of each cell line, more than one single cell clone was obtained and expression of luciferase was monitored over several days and compared to similarly obtained controls with known clock-controlled promoters, i.e. *CRY1-luc2*, *PER2-luc2*, and *NR1D1-luc2* and the previously published cell line U2OS *BMAL1-luc* [4], as well as a negative control cell line with *luc2P* under control of the *GAPDH*-promoter. The data was detrended using the MultiCycle software (Actimetrics).

Interestingly, the tested promoter-luciferase constructs *ATG3-luc2P*, *EIF5A2-luc2P*, and *SCN5A-luc2P* yielded a solid, reproducible circadian rhythm, even though the genes themselves were not expressed in a diurnal fashion on the microarray. The expression of luciferase was in a different phase than expression of luciferase regulated by the *BMAL1*-promoter. Control *luc2*-expressing cell lines with the promoters of the known core-clock genes *CRY1*, *PER2*, and *NR1D1*, showed a robust luciferase-oscillation in a similar phase to the tested constructs. Luc2P under the control of the *GAPDH*-promoter, did not exhibit a circadian rhythm (Figure 3.16).

The raw data from the promoter-luciferase constructs (Figure 3.17 B bottom vs. top) shows that despite the use of the unstable luc2P reporter, the amplitudes (peak

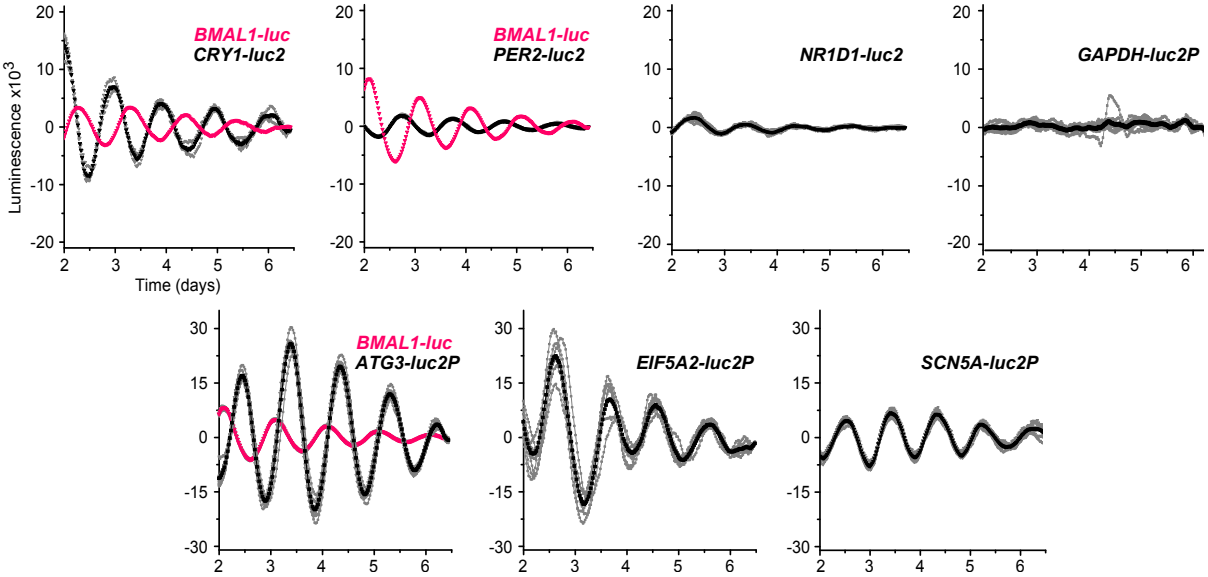

Figure 3.16.: *Promoter-luc2P* constructs of genes with a CRBS are rhythmic. U2OS promoter-luciferase cell lines were generated that stably expressed constructs, in which the destabilized *luc2P* was under control of promoters of the three non-rhythmic genes *ATG3*, *EIF5A2*, and *SCN5A*, that harbor a CRBS in this region. *Luc2* regulated by the promoter of three core-clock genes was used as positive control, while *GAPDH-luc2P* served as a negative control (upper panel). The *NR1D1-luc2* data is from a pool of cells after selection. For all other constructs several single cell clones were selected and measured at 37℃ for 6 days at 30 min intervals. Data was detrended using the Multicycle software. The data from several wells for one single cell clone is shown (grey), as well as the average of all wells (black). The pink line marks the expression of *BMAL1*-luc when measured in the same experiment. Note that the scale is different for the top row and the bottom row.

- trough) of the *ATG3*, *EIF5A2*, and *SCN5A*-promoter controlled rhythms were low in comparison to their Luc2P mean expression levels. Baseline expression of these genes is probably high, masking a circadian rhythm *in vivo*. For *BMAL1-luc* and *PER2-luc2* both, raw and detrended data looked very much alike because the amplitude was high in comparison to the basal expression of the luciferase (Figure 3.17 bottom). Baseline expression of these genes is probably mainly controlled by the circadian activation of transcription. *In vivo* this discrepancy between *BMAL1* and *PER2* expression on one side and *ATG3*, *EIF5A2*, and *SCN5A* expression must even be stronger, since a more stable form of luciferase was used in *BMAL1-luc* and *PER2-luc2*-cells.

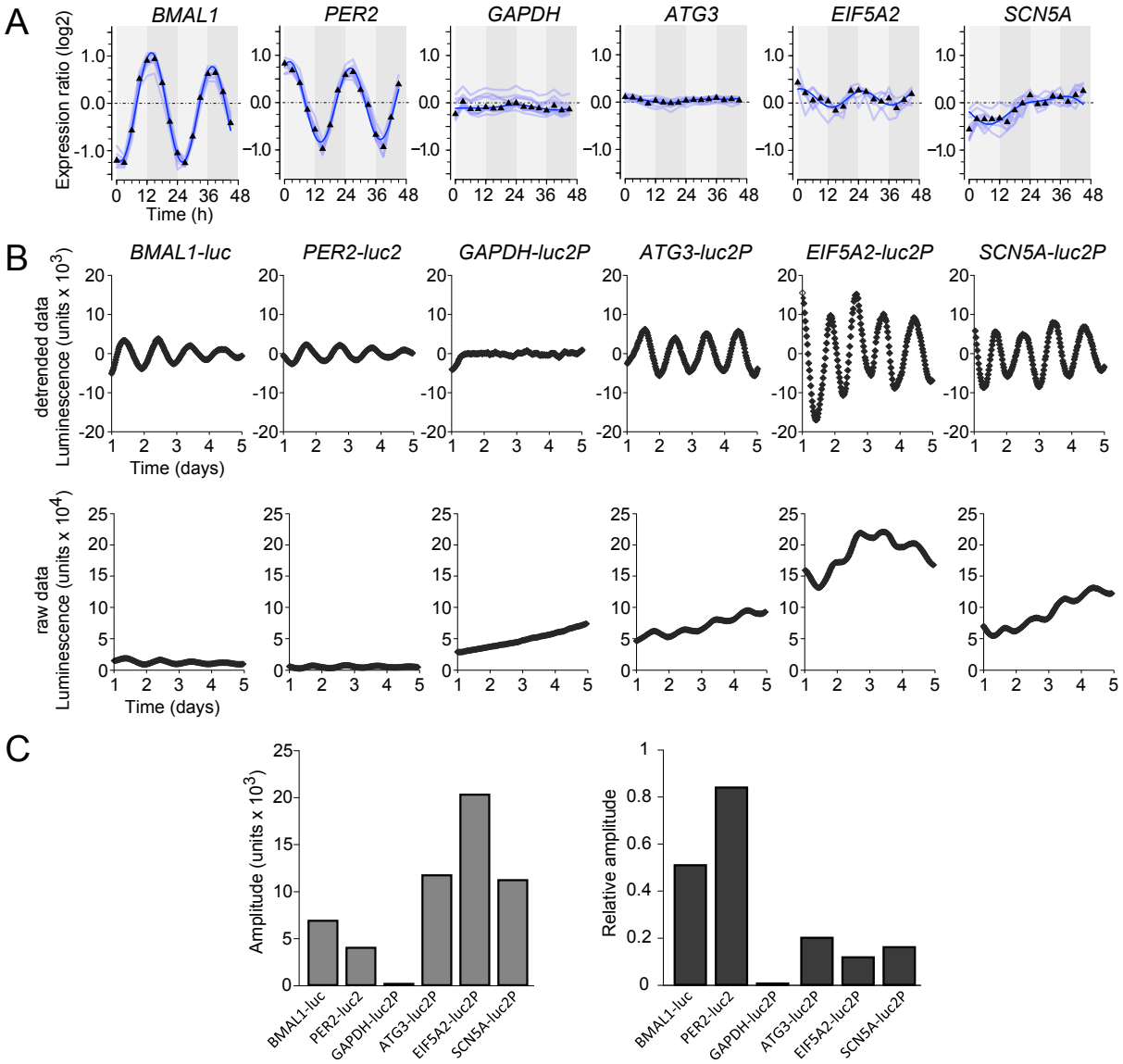

Figure 3.17.: Rhythmic transcription of genes with CRBSs in their promoter is masked by a high baseline expression. (A) Temporal expression profiles of *ATG3*, *EIF5A2*, and *SCN5A*. The rhythmic genes *BMAL1* and *PER2* and the housekeeping gene *GAPDH* are shown for control. (B) Low amplitude transcription rhythms are revealed by long-term luminescence measurements of stable U2OS cell lines expressing the destabilized luciferase2P (luc2P) under control of the promoters of *ATG3*, *EIF5A2*, and *SCN5A*. *BMAL1-luc*, *PER2-luc2*, and *GAPDH-luc2P* cell lines are shown for control. Luminescence of synchronized cultures was measured at 30 min intervals. Examples for detrended data (top) and raw data (bottom) for day 1 to day 5 after synchronization are shown. Several single cell clones were obtained showing the same behavior. (C) Contribution of circadian oscillations to the expression levels of these genes. The grey columns correspond to amplitude (peak - trough) of the oscillation of the first day. The black columns set the amplitude in relation to the mean expression levels of day 1 through day 3.

4. Discussion

4.1. BMAL1 and CLOCK CRBSs harbor binding motifs distinct from CRY1-sites

Most genome-wide studies in mammals to date have been performed in mice. This work provides insights into the circadian network in humans by means of a rhythmic human osteosarcoma cell line, U2OS. Recent publications visualized rhythmic binding of BMAL1 and CLOCK to DNA by ChIP-seq using several time points around the clock in mouse liver explants [6, 5]. This strongly rhythmic organ is known to exhibit high amplitude mRNA oscillations, whereas a former expression array study with U2OS cells yielded only low amplitude mRNA rhythms [3]. Therefore, in this work ChIP-Seq was performed in unsynchronized U2OS cells. Three individual ChIPs with antibodies directed against BMAL1, CLOCK, and CRY1, were analyzed by deep sequencing, and 3040 circadian regulator binding sites for all three proteins, comparable to the number found in highly rhythmic tissues like liver [5], were detected. The known clock-controlled genes *CRY1, CRY2, DBP, DEC1, DEC2, PER1, PER2, PER3, NR1D1, NR1D2*, and *TEF* possessed 1-4 binding sites each and all of them had at least one common site for BMAL1 and CLOCK binding, in most cases even for CRY1 as well.

Out of the set of 3040 CRBSs only about 13% overlapped between BMAL1 and CLOCK, and/or CRY1. This overlap is likely somewhat higher, because some weak peaks might have been below the identification threshold used in the bioinformatical analysis. Half of the CLOCK binding sites overlapped with BMAL1 binding sites, whereas only one fifth of BMAL1 binding sites overlapped with CLOCK. This is surprising, taking into consideration that BMAL1 and CLOCK form a heterodimer that binds to E-boxes. An explanation for this could be that the BMAL1-ChIP yielded data with a better signal-to-background ratio, allowing more peaks to be detected. This could be in part due to technical reasons, such as the sensitivity of the antibody and the fact that three lanes of the BMAL1-ChIP were sequenced and compared to one lane of the CLOCK-ChIP. Another possibility could be that NPAS2, which is highly related in amino

acid sequence to CLOCK and is known to be able to substitute CLOCK function in the mammalian forebrain [90], binds to BMAL1 in U2OS cells, thereby reducing the total signal that can be obtained by the CLOCK antibody. The microarray data from this work indicates that *NPAS2*-mRNA is expressed in similar levels as *CLOCK*.

Binding sites for both, BMAL1 and CLOCK, were enriched around transcription start sites, as would be expected for a transcription factor. However, binding sites for CRY1 were not enriched around the TSS. CRY1 exerts its function in the negative feedback loop of the clock by binding to BMAL1 and CLOCK [99, 100] and therefore might only interact indirectly with DNA. Depending on the state of the chromatin, it could be possible, that during the cross-linking process for the ChIP CRY1 that bound to BMAL1 and CLOCK was crosslinked to other sterically close DNA regions as well, which could account for a portion of the binding sites that was not enriched around the transcription start site. This is supported by the fact that a perfect E-box, CACGTG, was found only in 8% of the CRY1 binding sites, whereas it was found in 28% of each, BMAL1 and CLOCK binding sites. Interestingly though, a second E-box motif with a mismatch in the middle positions, CA**T**GTG and CAC**A**TG, which was called E*-box motif in this work, was found to be similarly present in all three ChIPs with 21% occurrence in both, BMAL1 and CRY1 binding sites, and 26% in CLOCK binding sites. Another motif, with a -GG- or a -CC- in the middle-positions, called E-GG-box here, was found to be enriched in all binding sites as well. This motif was more present in BMAL1 and CLOCK binding sites (18% each) than in CRY1 binding sites (12%), but the difference was not as pronounced as for perfect E-boxes. The E*-box motif was described recently as part of a possible 7-bp noncanonical binding motif for BMAL1 and CLOCK [166], supporting the notion of its special importance. However, no preference for the base following the E*-box motif was found in this work. Moreover, there was no enrichment for the E'-box motif ((**T**)CACGT**T**), which was described to be important for *mPer2*-transcription [166, 167].

Our data shows that binding sites containing the E*-box or the E-GG-box motif were only weakly enriched around the TSS and displayed a behavior similar to CRY1-binding sites. This suggests that there might exist at least two different classes of binding sites. One class harbors perfect E-boxes and is found preferentially in the vicinity of transcription start sites, whereas the other class contains E*-boxes or E-GG-boxes and is distributed randomly. Perhaps the binding to a perfect E-box results in a conformation of the BMAL1-CLOCK heterodimer that reduces the ability of CRY1 to interact with BMAL1-CLOCK. Alternatively, it may not be the CRY1-binding that is affected, but only the ability to immunoprecipitate CRY1 at BMAL1-CLOCK sites with perfect E-boxes. If,

for instance, CRY1 were able to release BMAL1 and CLOCK from the DNA efficiently at sites with perfect E-boxes, but only slowly at noncanonincal E-boxes, then there would be a higher probability to immunoprecipitate CRY1 at these sites. In any case CRBSs with E*-boxes appear to be functionally distinct from sites harboring perfect E-boxes.

The relatively high occurrence of the E*-box motif in CRY1 binding sites that do not overlap with BMAL1 and CLOCK suggests that there also might be another binding partner for CRY1, possibly a different bHLH-transcription factor. Crystal structure analysis of the BMAL1-CLOCK heterodimer indicates that CRY1 binds the complex via CLOCK [100]. It therefore seems likely that BMAL1 could be substituted by another bHLH-protein. Most bHLH transcription factors bind to E-box motifs, but they show different specificities as to the exact sequence of the motif [163]. CLOCK can form functional dimers with the BMAL1 paralog, BMAL2 [93, 199], which is not recognized by the BMAL1-antibody used in this work [5]. Not enough is known about BMAL2-CLOCK binding sites to tell, whether they possess a different sequence specificity. Another protein, which was found to preferably bind to the E*-box motif is the MYCN protein [200]. Although nothing is published about its possible binding to CLOCK or CRY1, this sequence specificity for E*-boxes makes it an interesting candidate for interactions with CRY1, even more so since MYC proteins are often overexpressed in tumor cells. It has previously been shown, that CRY1 can bind to glucocorticoid-receptors, which then bind to GREs on the DNA [120]. However this work cannot support this finding because no considerable enrichment for GREs was found among the binding sites.

It is to be noted, that the E-box motif, even when allowing one mismatch, was found in only 65% of the binding sites, which is along the lines of what has been found in other studies as well [5, 6]. Two thirds of these sites contained a perfect E-box, an E*-box, or the E-GG-box motif. Although there was little or no enrichment of other E-boxes with one mismatch among the binding sites, they still occurred with a certain frequency, making up the other third of sites with E-box motifs. The occurrence of E-boxes alone does not appear to be able to explain the binding of BMAL1 and CLOCK. In fact, even the known clock-controlled gene *CRY1*, did not contain an E-box or E*-box sequence in or near the CRBS that was found. It did harbor the already mentioned CACGT**T** motif, which was not enriched among the binding sites.

Tandem E-boxes with a spacer of 6 nucleotides have been found to be required for rhythmic expression of genes [165]. A tandem E-box motif with up to three mismatches and a spacer of 6 nucleotides was present in 18% (555 of 3040) of the binding sites, corresponding to 385 of 1454 (26%) genes in the dataset at-hand. The notion, that tandem E-boxes are required for rhythmic expression can, however, not be fully sup-

ported through this work. Only the clock genes *DBP, DEC1, DEC2, PER3*, and *TEF* possessed binding sites with a tandem E-box with three mismatches, the other core clock genes didn't contain a tandem E-box at all in their binding sites. Overall, a tandem E-box (with 2 or 3 mismatches) was present in 31% (18 of 58) of rhythmically transcribed genes with a CRBS, marking only a slight increase of this motif among the oscillating genes.

In the data set a clustering of binding sites was detected mostly at the outer 10 Mb of chromosomes 5, 7, 8, 13, and 14. These five regions, which amount to an overall size of about 50 Mb, account for about 1/3 of the detected CRBSs. This work was performed in a tumor cell line with an average of 76 chromosomes. Therefore, it seems likely that a special amplification of these chromosomal regions could have lead to a higher background signal, yielding an increased false discovery rate of binding sites. Further analysis though showed that these regions were not present in significantly higher copy numbers than the rest of the chromosome. In most cases a 1.5-fold enrichment of these regions was detected, which could result from a single crossing-over event at the outer regions, eventually leading to one more copy of this region in U2OS cells compared to the average amount. This by itself however, does not seem to be able to account for the massive overrepresentation of binding sites from these regions in the data set. Finally, it was also observed, that these regions often contained a high amount of E-boxes, which probably contributed partially to the large number of detected CRBSs. Therefore these clustered regions were not regarded separately in the entire analysis. Moreover, these binding sites contributed only to a relatively small number of genes, some of which were found to be expressed in a circadian manner.

4.2. A small fraction of genes with a CRBS is rhythmic

Rhythmicity of genes in U2OS cells was analyzed by applying an RNA-time course of temperature-entrained U2OS cells on a custom-made microarray. Using stringent criteria, a subset of 118 genes was defined as rhythmic and 58 of these harbored a CRBS. Among these were most of the known clock-controlled genes. The fraction of random oscillating genes was 0.9%. The set of genes that were within $\pm$ 20 kb of a TSS, but not the closest gene to a binding site, was equally often rhythmic as the random genes. In general, genes with a CRBS close to the TSS were considerably more often circadianly expressed. By extrapolation, it can be estimated that about 200-300 of all

transcribed genes are rhythmically expressed in U2OS cells, but probably only 46-58 genes are controlled directly by BMAL1 and CLOCK. This is higher than the 7 rhythmic genes reported before for U2OS cells [3], but nonetheless it is about 10-fold lower than the number found in mouse [1, 3, 5, 6]. Genes with BMAL1-only or CLOCK-only sites were not more often rhythmic than genes with a CRY1-only binding site. About 3-3.5% of the genes with a single CRBS were rhythmically expressed on the microarray. The values were similar for common BMAL1-CRY1 binding sites, whereas no genes with sites for CLOCK-CRY1 binding became rhythmic, but it must be noted that the numbers for these common CRBSs were very low. A significantly higher fraction of genes was rhythmically expressed, when a BMAL1-CLOCK (9%) or a BMAL1-CLOCK-CRY1 (10%) binding site was found. This shows, that common BMAL1-CLOCK CRBSs, especially within $\pm$ 1 kb of a TSS, had a certain predictive power, whereas single binding sites did not. This must be noted with caution, though, due to the overall small number of circadian genes with CRBSs.

The expression of rhythmic genes with a binding site clustered in two phases which were about 12 hours apart. 43 genes were expressed in phase with known BMAL1-CLOCK controlled genes of the core-clock like the *PER, CRY* and *REVERB*-genes, whereas 15 genes were expressed in phase with REVERBα regulated genes like *BMAL1* and *NPAS2*. The latter genes harbored only single binding sites with a low global score, that reflects the quality of the binding site and it seems likely that a substantial portion of these were rhythmic due to other reasons than their binding sites, which would be in accordance with the false positive rate of 0.9%.

All known core-clock genes with the exception of the *ROR* genes were rhythmic and expressed as reported before: The *CLOCK* gene showed only a weak rhythmicity and was not among the selected 118 rhythmic genes, which is in accordance with reports in the literature, attributing varying degrees of rhythmicity to *CLOCK* expression in different tissues of mice [1]. The gene for the bZIP-protein E4BP4, a transcription factor that counteracts the function of DBP and TEF, was expressed in antiphase to their mRNA [137]. *CRY1* has been reported to be expressed extremely late [201] and with a phase-delay relative to *NR1D1* [132]. This is reflected in the microarray data where the *REVERB*-genes are the earliest core-clock gene to be expressed, whereas *CRY1* is the latest core-clock gene in the defined phase-window. These correlations to the literature demonstrate the high quality of the time course data and increase reliability of the so far unknown data.

4.3. Functional classes of rhythmic genes in U2OS

The proteins encoded by most of the newly identified directly-controlled rhythmic genes have important roles in cellular function, such as glycogen synthesis (PPP1R3B, GYG1), siRNA-induced gene silencing (EIF2C1), cell-cell-interaction (COL22A1, COL17A1), transport of molecules or ions (ABCG1, ANO2), and nucleotide-synthesis (MTHFD1L). Among the rhythmic genes there are also many regulatory genes with the potential to control another subset of genes on a second level of hierarchy, such as genes involved in stress response (*GADD45B*), signal transduction (*PRKCH*, *PREX1*, *RALGDS*, *SULF2*) and regulation of transcription, which is the largest class of oscillating genes. Rhythmically expressed transcription factors such as the genes for AFF3, CCDC85B, FOXJ2, KLF11, and TFEB each have different groups of target genes, which thus may be indirectly regulated by the circadian clock.

Among the circadianly expressed genes without a CRBS transcriptional regulators apart from the core-clock genes *BMAL1*, *NPAS2*, and *E4BP4*, were less prominent, but for a couple of examples (*HMGA2, TSC22D3, ZNF323*). The protein products encoded by many of the oscillating genes without CRBSs are involved in important processes in the cell, such as glucose and fatty acid metabolism (ACAD, H6PD, PYGL) and signal transduction (CAMK2D, FGFR1, NRG1, PDE4B, PPP3CA, TNFSF15, WNT5B). Moreover, some proteins are important in the cytoskeleton and cell adhesion (LIMS2, NRG1, PALLD), in the extracellular matrix (MMP3, MMP24, SUN3), the oxidative stress response (AOX1, OSGIN1), RNA-binding (HNRPDL), and transport and channels (ATP1B1, CLTCL1, SLC2A1).

Interestingly, there are several examples for rhythmic genes that are involved in the same pathway. The *SERPINE1* (Serpin Peptidase Inhibitor) gene, which harbored three BMAL1-binding sites close to its transcription start site, and the genes *PLAT* (Tissue Plasminogen Activator) and *PLAU* (Plasminogen Activator), the proteins of which are also known as tPA and Urokinase, respectively, which did not have any binding sites, are all classically involved in the fibrinolysis pathway, regulating the turn-over of blood-clots. SERPINE1 is the principal inhibitor of tPA and Urokinase, which both can cleave Plasminogen in order to give rise to the active fibrinolytic protease Plasmin. In mouse SCN, there are several reports showing that *Serpine1* is rhythmically expressed [1, 202] and diverse reports about the expression of the gene *Plat*, which was rhythmic and in the same phase as *Serpine1* in one report [1], but constantly expressed in another [202]. Moreover, comparison of SCN of wild-type, Clock-mutant,

and *Cry1/Cry2*-deficient mice has shown that expression of mouse *Serpine* appears to be regulated directly by the circadian clock. This work can substantiate this notion in a human cell line through the finding of BMAL1 binding sites in the promoter of *SERPINE1*. Moreover, in U2OS cells, not only mRNA expression of *SERPINE1*, but also of *PLAT* and *PLAU* are circadian. *SERPINE1* expression peaked at CT 23 together with the early core-clock genes like *NR1D2*, which was in a different phase to expression of *PLAT* and *PLAU*, that both peaked at CT 15. The latter ones appear to be regulated on a second-level of hierarchy. A tempting speculation would be that there is some regulation, that purposefully enables expression of these genes in different phases in order to separate expression of the proteases from their inhibitor, thus allowing a temporal window in which the proteases can function. It would be interesting to know what the reason for rhythmic expression of these proteins in U2OS cells is. One publication suggests a function of plasmin which is not related to fibrinolysis, but to regulation of the protein BDNF (Brain-Derived Neurotrophic Factor) which has been shown to be able to modulate light- and glutamate-induced resetting of the clock in the SCN [203]. This hints to a new function of the plasmin-system in the circadian clock at least in the brain. On another line of investigation, the expression of tPA and Urokinase was reported in mouse osteoblasts and osteogenic sarcoma cells of rat origin, where they were assumed to regulate bone metabolism [204, 205, 206]. It thus appears likely, that the various processes of bone turn-over could be regulated in a timely manner by the circadian clock, which is reflected in the bone-derived U2OS cell line.

Another example for such antiphasic behavior of two related proteins is found for the genes *ANO1* and *ANO2*, which are close members of the Anoctamin family and are probably Ca^{2+}-activated Cl^- channels. *ANO2* harbored a CRY1-binding site and was rhythmic with a peak at CT2, while *ANO1*, which did not contain a binding site, was rhythmic in antiphase to it with a peak at CT18. These two ANO proteins might have different kinetics through variable gating characteristics. Anoctamins are highly overexpressed in some cancers. It is assumed that they participate in cell proliferation or tumor progression, since the activity of several anion channels correlates with the cell cycle [207]. As to now, there are little studies involving the regulation of channels by the molecular clock, although many channels and transporters have been found to be regulated in a circadian fashion in this work.

Two key enzymes of metabolic pathways were among the rhythmic genes without CRBS: *ACADL* (Acyl-Coa Dehydrogenase) and *H6PD* (Hexose-6-phosphate Dehydrogenase). ACADL is one of the four enzymes that catalyze the initial step of mitochondrial beta-oxidation of straight-chain fatty acids and therefore is important for en-

ergy generation. H6PD is the rate-limiting enzyme in the pentose phosphate pathway, a metabolic pathway that increases cellular levels of the co-enzyme NADPH (nicotinamide adenine dinucleotide phosphate). Fatty acid biosynthesis consumes NADPH and thereby increases the cellular level of NADP+, thus stimulating H6PD to produce more NADPH. In this respect these two enzymes function in opposing pathways, ACADL in fatty acid oxidation and G6PD in fatty acid biosynthesis. This may be reflected in their circadian expression pattern that is also in antiphase to one another, with *ACADL* being expressed at CT 5 and *H6PD* peak expression at CT 17.

An important link between circadian clocks and metabolism has been established in recent years based on the family of the peroxisome proliferator-activated receptors (PPAR). Although ChIP-seq identified a BMAL1-CRY1 binding site in an intron of *PPARA*, it was not among the 118 rhythmic genes from the microarray. Nevertheless, two other genes from the PPAR-family were rhythmic in U2OS-cells even with no CRBS, *PPARG* and the coactivator of PPARα, *PGC1A*. *PPARG* peaks at CT4, whereas *PGC1A* peaks at CT16. SIRT1 was found to regulate transcription of *PGC1A*, which then stimulates the expression of clock genes, especially *BMAL1* and *NR1D1* through coactivation of the ROR family of orphan nuclear receptors [132]. PPARγ induces expression of *Nr1d1* in adipose tissue of mice [130].

These examples illustrate that there are multiple ways how the relatively small set of rhythmic genes could influence the cell in a global way. Not only transcription factors, but also ion channels and key enzymes in signal transduction and metabolic pathways were found to be rhythmic. Circadian expression of genes without CRBSs could be regulated on a second level of hierarchy by the directly clock controlled genes, for example via rhythmically expressed transcription factors, or through a feedback of a rhythmic cellular state, e.g. the metabolism, on gene expression.

4.4. Rhythmic transcription of genes with CRBSs can be masked by a high baseline expression.

96% of the genes with a CRBS were not found to be rhythmic. A certain portion, but not all of them, might have been oscillating with a very low amplitude that could not be detected by the applied techniques and criteria. Therefore specific candidate genes were chosen that did have a binding site in their promoter region, but were not oscillating on the array. The gene *EIF5A2* harbored a common CRBS for all three

proteins, BMAL1, CLOCK, and CRY1 in its promoter region. Its expression profile on the microarray displayed a possible low-amplitude rhythm. The gene *ATG3*, on the other hand, appeared constitutively expressed on the microarray. It possessed a common binding site for BMAL1 and CLOCK and was located in a chromosomic region with high gene density. Although the binding site was only 255 bp upstream of the TSS of *ATG3*, it was closer to the TSS of another gene, transcribed on the other strand, *SLC35A5*, which was not rhythmic either. Expression levels of *ATG3* were high and similar to those of the highest expressed rhythmic genes. *SCN5A* harbored two binding sites for BMAL1 and one for CLOCK. Although none of these genes were rhythmic on the microarray, all three promoters featured perfect E-boxes within their CRBSs. The CLOCK site of *SCN5A* even contained a double E-box motif with a perfect E-box in tandem with an E-box harboring two mismatches.

Recent work suggested that about 15% of the oscillating genes in mouse liver contained rhythmic introns, which reflected the preRNA levels, whereas their exons did not cycle detectably [6, 208]. This behavior could result from post-transcriptional regulation, such as mRNA turnover. A long mRNA half-life leads to an accumulation of mRNAs that would mask a circadian rhythmicity in gene expression. qPCR analysis of preRNA in U2OS showed, that although *BMAL1* preRNA was rhythmic, the preRNAs of *ATG3*, *EIF5A2*, and *SCN5A* were not.

U2OS cells stably expressing promoter-*luc2P* constructs of these genes were rhythmic for many days, displaying a 24h-rhythm in similar phase to constructs of known clock-controlled genes such as *CRY1-luc2* and *PER2-luc2*. Importantly, the raw-data showed a 24h-oscillation, overlain by a constitutive high expression. The amplitudes of the transcription rhythms were rather low in comparison to the mean transcription activity of these promoters. This behavior was not found for the *BMAL1-luc* or *PER2-luc2* cell lines, although Luc and Luc2 are more stable than Luc2P. It appears, that there are two or more components regulating expression of the tested target genes: a rhythmic 24h-component and another component constitutively driving transcription. The combined action of both yields the observed behavior for the tested promoter-luciferase constructs. Detrending of the data uncoupled the circadian rhythm from other components and therefore displayed a clear rhythm. *In vivo* such a rhythm is probably too weak to be detected with a 3 h-resolution on a microarray, but can be discerned by promoter-luciferase constructs where data was collected every 30 minutes.

The examples at hand probably show the superimposed action of at least one other activating transcription factor, which acts along with BMAL1 and CLOCK, resulting in a non-rhythmic or weakly rhythmic expression profile. In a cell, there are always several

transcription factors and modifiers competing for binding sites and acting on transcription. Some transcription factors inhibit, others activate gene expression. In the end a balance is obtained between weak and strong, activating and inhibiting action, yielding a specific expression profile of a gene. Since these factors can be expressed in a tissue-specific or time-specific manner the expression of a gene can vary greatly between different tissues. In a different cell type the circadian clock could dominate over other transcription factors, thus resulting in a more strongly rhythmic expression. Alternatively a transcriptional repressor could dominate over the circadian clock, switching transcription completely off.

SCN5A is a subunit for a sodium-channel, which is found primarily in cardiac muscle. This channel is responsible for the initial upstroke of the action potential in an electrocardiogram. Interestingly, although its mRNA is not rhythmic in U2OS cells, there is a (yet unpublished) report from another group, that found *Scn5a* to be rhythmic in the heart muscle of mice [209]. This gene might be a good example for tissue-specific regulation. Its potential rhythmicity was acknowledged by its multiple CRBSs and the promoter-luciferase rhythms, but could only be shown in a tissue where it contains an important function.

A number of key proteins are rhythmically expressed in a tissue specific manner, allowing the circadian clock to optimize protein function at specific times of the day, e.g. metabolic enzymes in liver, bone turnover regulation in osteoblasts, and a sodium-channel for heart-function in the heart muscle. In most other tissues, however, these genes might not be expressed in a rhythmic fashion. Mechanisms determining the cell type specific complexity of circadian transcriptomes are not well understood. This work shows, that although a similar number of binding sites was found in ChIP-seq experiments, in U2OS cells only about 200-300 genes are probably rhythmically expressed, while in mouse liver, up to 3700 transcripts were expressed in a diurnal fashion [5, 6, 3, 210]. How can this finding be explained? A substantial portion of the 3040 CRBSs in U2OS is E-box specific. There could be a functional difference between CRBSs with canonical E-boxes that are mainly found in the vicinity of TSSs and CRBSs with E*-boxes that are rather randomly distributed, however no difference in rhythmicity was observed in this work. Amongst the oscillating genes in U2OS were 14 genes that are part of the core circadian clock machinery. These clock genes displayed relatively high amplitude expression rhythms (peak to trough ratios of 2 - 4), which are, however, substantially lower than the expression amplitudes of the corresponding genes in mouse liver [210]. The amplitudes of the remaining clock-controlled genes

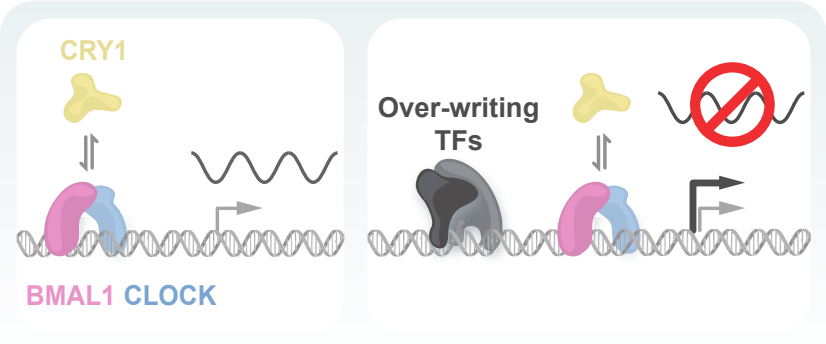

Figure 4.1.: Model of the effect of overwriting transcription factors.

with and without CRBSs were considerably lower (<1.5-fold). The promoter-luciferase constructs indicated, that many more genes with a binding site have the potential to be rhythmically expressed, but that the rhythmic component could be masked by binding of another transcription factor. We therefore hypothesize that BMAL1 and CLOCK bind opportunistically to accessible E-boxes in the DNA, but depending on the amount of other specific transcription factors in the cell, this binding will or will not be translated into a detectable rhythm of gene expression (Figure 4.1). Tissue-specific transcriptional activators or repressors could bind in the proximity of BMAL1 and CLOCK or modify the epigenetic landscape around the binding site, thus overriding a circadian effect or reducing it. Other scenarios are possible, where integration of various signals would lead to modification of the expression pattern, depending on the state of the cell. Tissue-specific modifiers would allow for genes to be rhythmic in one cell type, but not in another, although the same binding site is present in all cells, thus explaining the finding that the overlap of rhythmic genes between different tissues is very small [1].

This work therefore provides a list of genes that are potentially regulated by the circadian clock in one or another cell type. It can be used as a starting point for further work on more specialized pathways or for specific analysis of gene expression in U2OS cells. Moreover, it provides a model on how tissue-specificity can be obtained with the same binding site repertoire in different tissues, resulting in flat expression of a gene in one cell type and rhythmic expression in another. This hypothesis needs to be substantiated with more in-depth experiments in various tissues in the future.

A. Bibliography

[1] Satchidananda Panda, Marina P. Antoch, Brooke H. Miller, Andrew I. Su, Andrew B. Schook, Marty Straume, Peter G. Schultz, Steve A. Kay, Joseph S. Takahashi, and John B. Hogenesch. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell*, 109(3):307–20, May 2002.

[2] Christopher Vollmers, Satchidananda Panda, and Luciano DiTacchio. A high-throughput assay for siRNA-based circadian screens in human U2OS cells. *PloS ONE*, 3(10):e3457, January 2008.

[3] Michael E. Hughes, Luciano DiTacchio, Kevin R. Hayes, Christopher Vollmers, S. Pulivarthy, Julie E. Baggs, Satchidananda Panda, and John B. Hogenesch. Harmonics of circadian gene transcription in mammals. *PLoS Genetics*, 5(4):e1000442, April 2009.

[4] Bert Maier, Sabrina Wendt, Jens T. Vanselow, Thomas Wallach, Silke Reischl, Stefanie Oehmke, Andreas Schlosser, and Achim Kramer. A large-scale functional RNAi screen reveals a role for CK2 in the mammalian circadian clock. *Genes & Development*, 23(6):708–18, March 2009.

[5] Guillaume Rey, François Cesbron, Jacques Rougemont, Hans Reinke, Michael Brunner, and Felix Naef. Genome-Wide and Phase-Specific DNA-Binding Rhythms of BMAL1 Control Circadian Output Functions in Mouse Liver. *PLoS Biology*, 9(2):e1000595, February 2011.

[6] Nobuya Koike, Seung-Hee Yoo, Hung-Chung Huang, Vivek Kumar, Choogon Lee, Tae-Kyung Kim, and Joseph S. Takahashi. Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals. *Science*, 338(6105):349–54, August 2012.

[7] Aurélio Balsalobre, Francesca Damiola, and Ueli Schibler. A serum shock induces circadian gene expression in mammalian tissue culture cells. *Cell*, 93(6):929–37, June 1998.

[8] Masao Doi, Jun Hirayama, and Paolo Sassone-Corsi. Circadian regulator CLOCK is a histone acetyltransferase. *Cell*, 125(3):497–508, May 2006.

[9] Jiwon Lee, Yool Lee, Min Joo Lee, Eonyoung Park, Sung Hwan Kang, Chin Ha Chung, Kun Ho Lee, and Kyungjin Kim. Dual modification of BMAL1 by SUMO2/3 and ubiquitin promotes circadian activation of the CLOCK/BMAL1 complex. *Molecular and Cellular Biology*, 28(19):6056–65, 2008.

[10] Steven M. Reppert and David R. Weaver. Coordination of circadian timing in mammals. *Nature*, 418(6901):935–41, August 2002.

[11] Taeko Nishiwaki, Yoshinori Satomi, Yohko Kitayama, Kazuki Terauchi, Reiko Kiyohara, Toshifumi Takao, and Takao Kondo. A sequential program of dual phosphorylation of KaiC as a basis for circadian rhythm in cyanobacteria. *EMBO Journal*, 26(17):4029–37, September 2007.

[12] Masahiro Ishiura. Expression of a Gene Cluster kaiABC as a Circadian Feedback Process in Cyanobacteria. *Science*, 281(5382):1519–23, September 1998.

[13] Masato Nakajima, Keiko Imai, Hiroshi Ito, Taeko Nishiwaki, Yoriko Murayama, Hideo Iwasaki, Tokitaka Oyama, and Takao Kondo. Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science*, 308(5720):414–5, April 2005.

[14] Tetsuya Mori, Sergei V. Saveliev, Yao Xu, Walter F. Stafford, Michael M. Cox, Ross B. Inman, and Carl H. Johnson. Circadian clock protein KaiC forms ATP-dependent hexameric rings and binds DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 99(26):17203–8, December 2002.

[15] Yoichi Nakahira, Mitsunori Katayama, Hiroshi Miyashita, Shinsuke Kutsuna, Hideo Iwasaki, Tokitaka Oyama, and Takao Kondo. Global gene repression by KaiC as a master process of prokaryotic circadian system. *Proceedings of the National Academy of Sciences of the United States of America*, 101(3):881–5, January 2004.

[16] Vikram Vijayan, Rick Zuzow, and Erin K. O'Shea. Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 106(52):22564–8, December 2009.

[17] Tetsuya Mori and Carl Hirschie Johnson. Independence of Circadian Timing from Cell Division in Cyanobacteria Independence of Circadian Timing from Cell Division in Cyanobacteria. *Journal of Bacteriology*, 183(8):2439 – 44, 2001.

[18] Natalia B. Ivleva, Matthew R. Bramlett, Paul A. Lindahl, and Susan S. Golden. LdpA: a component of the circadian clock senses redox state of the cell. *EMBO Journal*, 24(6):1202–10, March 2005.

[19] Guogang Dong and Susan S. Golden. How a cyanobacterium tells time. *Current Opinion in Microbiology*, 11(6):541–6, December 2008.

[20] Thammajun L. Wood, Jennifer Bridwell-Rabb, Yong-Ick Kim, Tiyu Gao, Yong-Gang Chang, Andy LiWang, David P. Barondeau, and Susan S. Golden. The KaiA protein of the cyanobacterial circadian oscillator is modulated by a redox-active cofactor. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13):5804–9, March 2010.

[21] Jean-Jacques de Marain. Observation Botanique. *Histoire de l'Acadamie Royale des Sciences*, page 35, 1729.

[22] Andrew J. Millar, Isabelle A. Carré, Carl A. Strayer, Nam-Hai Chua, and Steve A. Kay. Circadian clock mutants in Arabidopsis identified by luciferase imaging. *Science*, 267(5201):1161–3, February 1995.

[23] Michael F. Covington, Julin N. Maloof, Marty Straume, Steve A. Kay, and Stacey L. Harmer. Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome Biology*, 9(8):r–130, 2008.

[24] Tomoyuki Takase, Haruki Ishikawa, Haruko Murakami, Jun Kikuchi, Kumi Sato-Nara, and Hitoshi Suzuki. The circadian clock modulates water dynamics and aquaporin expression in Arabidopsis roots. *Plant & Cell Physiology*, 52(2):373–83, February 2011.

[25] Wei Wang, Jinyoung Yang Barnaby, Yasuomi Tada, Hairi Li, Mahmut Tör, Daniela Caldelari, Daeun Lee, Xiang-Dong Fu, and Xinnian Dong. Timing of plant immune responses by a central circadian regulator. *Nature*, 470(7332):110–4, February 2011.

[26] Antony N. Dodd, Neeraj Salathia, Anthony Hall, Eva Kévei, Réka Tóth, Ferenc Nagy, Julian M. Hibberd, Andrew J. Millar, and Alex A. R. Webb. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science*, 309(5734):630–3, July 2005.

[27] W. W. Garner and H. A. Allard. Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *Monthly Weather Review*, page 415, 1920.

[28] Stacey L. Harmer, John B. Hogenesch, Marty Straume, Hur-Song Chang, Bin Han, Tong Zhu, Xun Wang, Joel A. Kreps, and Steve A. Kay. Orchestrated Transcription of Key Pathways in Arabidopsis by the Circadian Clock. *Science*, 290(5499):2110–3, December 2000.

[29] Katharine E. Hubbard, Fiona C. Robertson, Neil Dalchauz, Alex A. R. Webb, and Katharine Hubbard. Molecular BioSystems. *Molecular BioSystems*, 5:1502–11, 2009.

[30] David E. Somers, Paul F. Devlin, and Steve A. Kay. Phytochromes and Cryptochromes in the Entrainment of the Arabidopsis Circadian Clock. *Science*, 282(5393):1488–90, November 1998.

[31] Paul F. Devlin and Steve A. Kay. Cryptochromes are required for phytochrome signaling to the circadian clock but not for rhythmicity. *The Plant Cell*, 12:2499–510, December 2000.

[32] David E. Somers, Thomas F. Schultz, Maureen Milnamow, and Steve A. Kay. ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. *Cell*, 101(3):319–29, April 2000.

[33] Réka Tóth, Éva Kevei, Anthony Hall, Andrew J. Millar, and Ferenc Nagy. Circadian Clock-Regulated Expression of Phytochrome and Cryptochrome Genes in Arabidopsis 1. *Plant Physiology*, 127(December):1607–16, 2001.

[34] James C. W. Locke, László Kozma-Bognár, Peter D. Gould, Balázs Fehér, Eva Kevei, Ferenc Nagy, Matthew S. Turner, Anthony Hall, and Andrew J. Millar. Experimental validation of a predicted feedback loop in the multi-oscillator clock of Arabidopsis thaliana. *Molecular Systems Biology*, 2:59, January 2006.

[35] Eva Herrero and Seth J. Davis. Time for a Nuclear Meeting: Protein Trafficking and Chromatin Dynamics Intersect in the Plant Circadian System. *Molecular Plant*, 5(3):1–12, February 2012.

[36] Mariano Perales and Paloma Más. A functional link between rhythmic changes in chromatin structure and the Arabidopsis biological clock. *The Plant Cell*, 19(7):2111–23, July 2007.

[37] Benoit Farinas and Paloma Mas. Functional implication of the MYB transcription factor RVE8/LCL5 in the circadian control of histone acetylation. *The Plant Journal*, 66(2):318–29, April 2011.

[38] Antje Rohde and Rishikesh P. Bhalerao. Plant dormancy in the perennial context. *Trends in Plant Science*, 12(5):217–23, May 2007.

[39] C. J. Weiser. Cold Resistance and Injury in Woody Plants. *Science*, 169(3952):1269–78, 1970.

[40] Annikki Welling and E. Tapio Palva. Molecular control of cold acclimation in trees. *Physiologia Plantarum*, 127(2):167–81, June 2006.

[41] Rongmin Chen, Aaron Schirmer, Yongjin Lee, Hyeongmin Lee, Vivek Kumar, Seung-Hee Yoo, Joseph S. Takahashi, and Choogon Lee. Rhythmic PER abundance defines a critical nodal point for negative feedback within the circadian clock mechanism. *Molecular Cell*, 36(3):417–30, November 2009.

[42] Christina Querfurth, Axel C. R. Diernfellner, Elan Gin, Erik Malzahn, Thomas Höfer, and Michael Brunner. Circadian conformational change of the Neurospora clock protein FREQUENCY triggered by clustered hyperphosphorylation of a basic domain. *Molecular Cell*, 43(5):713–22, September 2011.

[43] Tobias Schafmeier, Andrea Haase, Krisztina Káldi, Johanna Scholz, Marc Fuchs, and Michael Brunner. Transcriptional feedback of Neurospora circadian clock gene by phosphorylation-dependent inactivation of its transcription factor. *Cell*, 122(2):235–46, July 2005.

[44] Qun He, Joonseok Cha, Qiyang He, Heng-Chi Lee, Yuhong Yang, and Yi Liu. CKI and CKII mediate the FREQUENCY-dependent phosphorylation of the WHITE COLLAR complex to close the Neurospora circadian negative feedback loop. *Genes & Development*, 20(18):2552–65, September 2006.

[45] Tobias Schafmeier, Axel Diernfellner, Astrid Schäfer, Orfeas Dintsis, Andrea Neiss, and Michael Brunner. Circadian activity and abundance rhythms of the Neurospora clock transcription factor WCC associated with rapid nucleo-cytoplasmic shuttling. *Genes & Development*, 22:3397–402, 2008.

[46] Erik Malzahn, Stilianos Ciprianidis, Krisztina Káldi, Tobias Schafmeier, and Michael Brunner. Photoadaptation in Neurospora by Competitive Interaction of Activating and Inhibitory LOV Domains. *Cell*, 142(5):762–72, September 2010.

[47] Gencer Sancar, Cigdem Sancar, Britta Brügger, Nati Ha, Timo Sachsenheimer, Elan Gin, Simon Wdowik, Ingrid Lohmann, Felix Wieland, Thomas Höfer, Axel Diernfellner, and Michael Brunner. A global circadian repressor controls antiphasic expression of metabolic genes in Neurospora. *Molecular Cell*, 44(5):687–97, December 2011.

[48] Paul E. Hardin. The Circadian Timekeeping System of Drosophila. *Current Biology*, 15:714–22, 2005.

[49] Ronald J. Konopka and Seymour Benzer. Clock Mutants of Drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 68(9):2112–6, 1971.

[50] Zhong Sheng Sun, Urs Albrecht, Olga Zhuchenko, Jennifer Bailey, Gregor Eichele, and Cheng Chi Lee. RIGUI , a Putative Mammalian Ortholog of the Drosophila period Gene. *Cell*, 90:1003–11, 1997.

[51] Hajime Tei, Hitoshi Okamura, Yasufumi Shigeyoshi, Chiaki Fukuhara, Ritsuko Ozawa, Matsumi Hirose, and Yoshiyuki Sakaki. Circadian oscillation of a mammalian homologue of the Drosophila period gene. *Nature*, 389:512–16, 1997.

[52] Urs Albrecht, Zhong Sheng Sun, Gregor Eichele, and Cheng Chi Lee. A Differential Response of Two Putative Mammalian Circadian Regulators , mper1 and mper2 , to Light. *Cell*, 91:1055–64, 1997.

[53] Franck Delaunay, Christine Thisse, and Oriane Marchand. An Inherited Functional Circadian Clock in Zebrafish Embryos. *Science*, 289:297 – 300, 2000.

[54] Nicolai Peschel and Charlotte Helfrich-Förster. Setting the clock–by nature: circadian rhythm in the fruitfly Drosophila melanogaster. *FEBS Letters*, 585(10):1435–42, May 2011.

[55] Sebastian Kadener, Dan Stoleru, Michael McDonald, Pipat Nawathean, and Michael Rosbash. Clockwork Orange is a transcriptional repressor and a new Drosophila circadian pacemaker component. *Genes & Development*, 21(13):1675–86, July 2007.

[56] Charlotte Helfrich-Förster. The neuroarchitecture of the circadian clock in the brain of Drosophila melanogaster. *Microscopy Research and Technique*, 62:94–102, October 2003.

[57] Yoko Miyasako, Yujiro Umezaki, and Kenji Tomioka. Separate sets of cerebral clock neurons are responsible for light and temperature entrainment of Drosophila circadian locomotor rhythms. *Journal of Biological Rhythms*, 22(2):115–26, April 2007.

[58] Charlotte Helfrich-Förster, Christine Winter, Alois Hofbauer, Jeffrey C. Hall, and Ralf Stanewsky. The circadian clock of fruit flies is blind after elimination of all known photoreceptors. *Neuron*, 30:249–61, April 2001.

[59] Taishi Yoshii, Takeshi Todo, Corinna Wülbeck, Ralf Stanewsky, and Charlotte Helfrich-Förster. Cryptochrome is present in the compound eyes and a subset of Drosophila's clock neurons. *The Journal of Comparative Neurology*, 508(6):952–66, June 2008.

[60] Stefano Vanin, Supriya Bhutani, Stefano Montelli, Pamela Menegazzi, Edward W. Green, Mirko Pegoraro, Federica Sandrelli, Rodolfo Costa, and Charalambos P. Kyriacou. Unexpected features of Drosophila circadian behavioural rhythms under natural conditions. *Nature*, 000:1–6, April 2012.

[61] John H. Postlethwait, Yanm Yi-Lin, Michael A. Gates, Sally Horne, Angel Amores, Alison Brownlie, Adriana Donovan, Elizabeth S. Egan, Allan Force, Zhiyuan Gong, Carole Goutel, Andreas Fritz, Robert Kelsh, Ela Knapik, Eric Liao, Barry Paw, David Ransom, Amy Singer, Margaret Thomson, Tariq S. Abduljabbar, Pam Yelick, Dave Beier, J.-S. Joly, Dan Larhammar, Frederiic Rosa, Monte Westerfield, Leonard I. Zon, Steve L. Johnson, and Willam S. Talbot. Vertebrate genome evolution and the zebrafish gene map. *Nature Genetics*, 18:345–9, 1998.

[62] Gad Vatine, Daniela Vallone, Yoav Gothilf, and Nicholas S. Foulkes. It's time to swim! Zebrafish and the circadian clock. *FEBS letters*, 585(10):1485–94, May 2011.

[63] Yuri Kobayashi, Tomoko Ishikawa, Jun Hirayama, Hiromi Daiyasu, Satoru Kanai, Hiroyuki Toh, Itsuki Fukuda, Tohru Tsujimura, Nobuyuki Terada, Yasuhiro Kamei, Shunsuke Yuba, Shigenori Iwai, and Takeshi Todo. Molecular analysis of zebrafish photolyase / cryptochrome family : two types of cryptochromes present in zebrafish. *Genes to Cells*, 5:725–38, 2000.

[64] David Whitmore, Nicholas S. Foulkes, and Paolo Sassone-Corsi. Light acts directly on organs and cells in culture to set the vertebrate circadian clock. *Nature*, 404(6773):87–91, March 2000.

[65] Jun Hirayama, Saurabh Sahar, Benedetto Grimaldi, Teruya Tamaru, Ken Takamatsu, Yasukazu Nakahata, and Paolo Sassone-Corsi. CLOCK-mediated acetylation of BMAL1 controls circadian function. *Nature*, 450(7172):1086–90, December 2007.

[66] Paraskevi Moutsaki, David Whitmore, James Bellingham, Katsuhiko Sakamoto, Zoë K. David-Gray, and Russell G. Foster. Teleost multiple tissue (tmt) opsin: a candidate photopigment regulating the peripheral clocks of zebrafish? *Brain research. Molecular brain research*, 112:135–45, April 2003.

[67] Gad Vatine, Daniela Vallone, Lior Appelbaum, Philipp Mracek, Zohar Ben-Moshe, Kajori Lahiri, Yoav Gothilf, and Nicholas S. Foulkes. Light directs zebrafish period2 expression via conserved D and E boxes. *PLoS Biology*, 7(10):e1000223, October 2009.

[68] Mark W. Hurd, Jason Debruyne, Martin Straume, and Gregory M. Cahill. Circadian rhythms of locomotor activity in zebrafish. *Physiology & behavior*, 65(3):465–72, December 1998.

[69] Gregory M. Cahill, Mark W. Hurd, and Matthew M. Batchelor. Circadian rhythmicity in the locomotor activity of larval zebrafish. *Neuroreport*, 9(15):3445–9, October 1998.

[70] Yoav Gothilf, Steven L. Coon, Reiko Toyama, Ajay Chitnis, M. A. A. Namboodiri, and David C. Klein. Zebrafish serotonin N-acetyltransferase-2: marker for development of pineal photoreceptors and circadian clock function. *Endocrinology*, 140(10):4895–903, October 1999.

[71] Nasser Kazimi and Gregory M. Cahill. Development of a circadian melatonin rhythm in embryonic zebrafish. *Brain research. Developmental brain research*, 117(1):47–52, October 1999.

[72] Gregory M. Cahill. Circadian regulation of melatonin production in cultured zebrafish pineal and retina. *Brain Research*, 708:177–81, 1996.

[73] Maki Kaneko and Gregory M. Cahill. Light-dependent development of circadian gene expression in transgenic zebrafish. *PLoS Biology*, 3(2):e34, March 2005.

[74] Eberhard Gwinner and Roland Brandstätter. Complex bird clocks. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 356(1415):1801–10, November 2001.

[75] Michael J. Bailey, Phillip D. Beremand, Rick Hammer, Elizabeth Reidel, Terry L. Thomas, and Vincent M. Cassone. Transcriptional profiling of circadian patterns of mRNA expression in the chick retina. *The Journal of biological chemistry*, 279(50):52247–54, December 2004.

[76] Vincent M. Cassone, Jiffin K. Paulose, Melissa G. Whitfield-Rucker, and Jennifer L. Peters. Time's arrow flies like a bird: Two paradoxes for avian circadian biology. *Gen Comp Endocrinol.*, 163(1-2):109–16, 2009.

[77] Kenneth P. Able and Jeffrey D. Cherry. Mechanisms of dusk orientation in white-throated sparrows (Zonotrichia albicollis): Clock-shift experiments. *Journal of Comparative Physiology A*, 159(1):107–13, 1986.

[78] Andrea Möller, Sven Sagasser, Wolfgang Wiltschko, and Bernd Schierwater. Retinal cryptochrome in a migratory passerine bird: a possible transducer for the avian magnetic compass. *Die Naturwissenschaften*, 91(12):585–8, December 2004.

[79] Ashutosh Rastogi, Yatinesh Kumari, Sangeeta Rani, and Vinod Kumar. Phase inversion of neural activity in the olfactory and visual systems of a night-migratory bird during migration. *European Journal of Neuroscience*, 34(1):99–109, July 2011.

[80] David M. Berson. Strange vision: ganglion cells as circadian photoreceptors. *Trends in Neurosciences*, 26(6):314–20, June 2003.

[81] Robert J. Lucas, Melanie S. Freedman, Daniela Lupi, Marta Munoz, Zoe K. David-Gray, and Russell G. Foster. Identifying the photoreceptive inputs to the mammalian circadian system using transgenic and retinally degenerate mice. *Behavioural Brain Research*, 125(1-2):97–102, November 2001.

[82] Shin Yamazaki. Resetting Central and Peripheral Circadian Oscillators in Transgenic Rats. *Science*, 288(5466):682–5, April 2000.

[83] Urs Albrecht. Timing to Perfection: The Biology of Central and Peripheral Circadian Clocks. *Neuron*, 74(2):246–60, 2012.

[84] Francesca Damiola. Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes & Development*, 14(23):2950–61, December 2000.

[85] Gad Asher, Hans Reinke, Matthias Altmeyer, Maria Gutierrez-Arcelus, Michael O. Hottiger, and Ueli Schibler. Poly(ADP-ribose) polymerase 1 participates in the phase entrainment of circadian clocks to feeding. *Cell*, 142(6):943–53, September 2010.

[86] Seung-Hee Yoo, Shin Yamazaki, Phillip L. Lowrey, Kazuhiro Shimomura, Caroline H. Ko, Ethan D. Buhr, Sandra M. Siepka, Hee-Kyung Hong, Won Jun Oh, Ook Joon Yoo, Michael Menaker, and Joseph S. Takahashi. PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proceedings of the National Academy of Sciences of the United States of America*, 101(15):5339–46, April 2004.

[87] Ethan D. Buhr, Seung-Hee Yoo, and Joseph S. Takahashi. Temperature as a universal resetting cue for mammalian circadian oscillators. *Science*, 330:379–85, October 2010.

[88] David P. King, Yaliang Zhao, Ashvin M. Sangoram, Lisa D. Wilsbacher, Minoru Tanaka, Marina P. Antoch, Thomas D. L. Steeves, Martha Hotz Vitaterna, Jon M. Kornhauser, Phillip L. Lowrey, Fred W. Turek, and Joseph S. Takahashi. Positional cloning of the mouse circadian clock gene. *Cell*, 89(4):641–53, May 1997.

[89] Nicholas Gekakis, David Staknis, Hubert B. Nguyen, Fred C. Davis, Lisa D. Wilsbacher, David P. King, Joseph S. Takahashi, and Charles J. Weitz. Role of the CLOCK Protein in the Mammalian Circadian Mechanism. *Science*, 280(5369):1564–9, June 1998.

[90] Martin Reick, Joseph A. Garcia, Carol Dudley, and Steven L. McKnight. NPAS2: an analog of clock operative in the mammalian forebrain. *Science*, 293(5529):506–9, July 2001.

[91] John B. Hogenesch, Yi-Zhong Gu, Sanjay Jain, and Christopher A. Bradfield. The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. *Proceedings of the National Academy of Sciences of the United States of America*, 95(10):5474–9, May 1998.

[92] Maureen K. Bunger, Lisa D. Wilsbacher, Susan M. Moran, Cynthia Clendenin, Laurel A. Radcliffe, John B. Hogenesch, M. Celeste Simon, Joseph S. Takahashi, and Christopher A. Bradfield. Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell*, 103(7):1009–17, December 2000.

[93] Shuqun Shi, Akiko Hida, Owen P. McGuinness, David H. Wasserman, Shin Yamazaki, and Carl Hirschie Johnson. Circadian clock gene Bmal1 is not essential; functional replacement with its paralog, Bmal2. *Current Biology*, 20(4):316–21, February 2010.

[94] Binhai Zheng, David W. Larkin, Urs Albrecht, Zhong Sheng Sun, Marijke Sage, Gregor Eichele, Cheng Chi Lee, and Allan Bradley. The mPer2 gene encodes a functional component of the mammalian circadian clock. *Nature*, 400(6740):169–73, July 1999.

[95] Gijsbertus T. J. van der Horst, Manja Muijtjens, Kumiko Kobayashi, Riya Takano, Shin-Ichiro Kanno, Masashi Takao, Jan de Wit, Anton Verkerk, Andre P. M. Eker, Dik van Leenen, Ruud Buijs, Dirk Bootsma, Jan H. J. Hoeijmakers, and Akira Yasui. Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature*, 398(6728):627–30, April 1999.

[96] Martha Hotz Vitaterna, Christopher P. Selby, Takeshi Todo, Hitoshi Niwa, Carol Thompson, Ethan M. Fruechte, Kenichi Hitomi, Randy J. Thresher, Tomoko Ishikawa, Junichi Miyazaki, Joseph S. Takahashi, and Aziz Sancar. Differential regulation of mammalian period genes and circadian rhythmicity by cryptochromes 1 and 2. *Proceedings of the National Academy of Sciences of the United States of America*, 96(21):12114–9, October 1999.

[97] Kazuhiko Kume, Mark J. Zylka, Sathyanarayanan Sriram, Lauren P. Shearman, David R. Weaver, Xiaowei Jin, Elizabeth S. Maywood, Michael H. Hastings, and Steven M. Reppert. mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell*, 98(2):193–205, July 1999.

[98] Ashvin M. Sangoram, Lino Saez, Marina P. Antoch, Nicholas Gekakis, David Staknis, Andrew Whiteley, Ethan M. Fruechte, Martha Hotz Vitaterna, Kazuhiro Shimomura, David P. King, M ichael W. Young, Charles J. Weitz, and Joseph S. Takahashi. Mammalian circadian autoregulatory loop: a timeless ortholog and mPer1 interact and negatively regulate CLOCK-BMAL1-induced transcription. *Neuron*, 21(5):1101–13, November 1998.

[99] Edmund A. Griffin Jr., David Staknis, and Charles J. Weitz. Light-Independent Role of CRY1 and CRY2 in the Mammalian Circadian Clock. *Science*, 286(768):768–71, October 1999.

[100] Nian Huang, Yogarany Chelliah, Yongli Shan, Clinton A. Taylor, Seung-Hee Yoo, Carrie Partch, Carla B. Green, Hong Zhang, and Joseph S. Takahashi. Crystal Structure of the Heterodimeric CLOCK:BMAL1 Transcriptional Activator Complex. *Science*, 337:189–94, May 2012.

[101] Choogon Lee, Jean-Pierre Etchegaray, Felino R. A. Cagampang, Andrew S. I. Loudon, and Steven M. Reppert. Posttranslational mechanisms regulate the mammalian circadian clock. *Cell*, 107:855–67, December 2001.

[102] Kamon Sanada, Toshiyuki Okano, and Yoshitaka Fukada. Mitogen-activated protein kinase phosphorylates and negatively regulates basic helix-loop-helix-PAS transcription factor BMAL1. *The Journal of Biological Chemistry*, 277(1):267–71, January 2002.

[103] Hikari Yoshitane, Toshifumi Takao, Yoshinori Satomi, Ngoc-hien Du, Toshiyuki Okano, and Yoshitaka Fukada. Roles of CLOCK Phosphorylation in Suppression of E-box-dependent Transcription. *Molecular and Cellular Biology*, 29(13):3675–86, 2009.

[104] Silke Reischl, Katja Vanselow, Pål O. Westermark, Nadine Thierfelder, Bert Maier, Hanspeter Herzel, and Achim Kramer. Beta-TrCP1-mediated degradation of PERIOD2 is essential for circadian dynamics. *Journal of Biological Rhythms*, 22(5):375–86, October 2007.

[105] Luca Busino, Florian Bassermann, Alessio Maiolica, Choogon Lee, Patrick M. Nolan, Sofia I. H. Godinho, Giulio F. Draetta, and Michele Pagano. SCFFbxl3 controls the oscillation of the circadian clock by directing the degradation of cryptochrome proteins. *Science*, 316(5826):900–4, May 2007.

[106] Sofia I. H. Godinho, Elizabeth S. Maywood, Linda Shaw, Valter Tucci, Alun R. Barnard, Luca Busino, Michele Pagano, Rachel Kendall, Mohamed M. Quwailid, M. Rosario Romero, John O'neill, Johanna E. Chesham, Debra Brooker, Zuzanna Lalanne, Michael H. Hastings, and Patrick M. Nolan. The after-hours mutant reveals a role for Fbxl3 in determining mammalian circadian period. *Science*, 316(5826):897–900, May 2007.

[107] Makoto Akashi, Yoshiki Tsuchiya, Takao Yoshino, and Eisuke Nishida. Control of Intracellular Dynamics of Mammalian Period Proteins by Casein Kinase I epsilon (CKIepsilon) and CKI delta in Cultured Cells. *Molecular and Cellular Biology*, 22(6):1693–703, 2002.

[108] Hao A. Duong, Maria S. Robles, Darko Knutti, and Charles J. Weitz. A molecular mechanism for circadian clock negative feedback. *Science*, 332(6036):1436–9, June 2011.

[109] Kiran Padmanabhan, Maria S. Robles, Thomas Westerling, and Charles J. Weitz. Feedback Regulation of Transcriptional Termination by the Mammalian Circadian Clock PERIOD Complex. *Science*, 337(6094):599–602, March 2012.

[110] Han Cho, Xuan Zhao, Megumi Hatori, Ruth T. Yu, Grant D. Barish, Michael T. Lam, Ling-Wa Chong, Luciano DiTacchio, Annette R. Atkins, Christopher K. Glass, Christopher Liddle, Johan Auwerx, Michael Downes, Satchidananda Panda, and Ronald M. Evans. Regulation of circadian behaviour and metabolism by REV-ERB-α and REV-ERB-β. *Nature*, 485(7396):123–7, May 2012.

[111] Andrew C. Liu, Hien G. Tran, Eric E. Zhang, Aaron A. Priest, David K. Welsh, and Steve A. Kay. Redundant function of REV-ERBalpha and beta and non-essential role for Bmal1 cycling in transcriptional regulation of intracellular circadian rhythms. *PLoS Genetics*, 4(2):e1000023, February 2008.

[112] Fabienne Guillaumond, Hugues Dardente, Vincent Giguère, and Nicolas Cermakian. Differential control of Bmal1 circadian transcription by REV-ERB and ROR nuclear receptors. *Journal of Biological Rhythms*, 20(5):391–403, October 2005.

[113] Nicolas Preitner, Francesca Damiola, Joszef Zakany, Denis Duboule, Urs Albrecht, and Ueli Schibler. The Orphan Nuclear Receptor REV-ERB a Controls Circadian Transcription within the Positive Limb of the Mammalian Circadian Oscillator. *Cell*, 110:251–260, 2002.

[114] Lei Yin and Mitchell A. Lazar. The orphan nuclear receptor Rev-erbalpha recruits the N-CoR/histone deacetylase 3 corepressor to regulate the circadian Bmal1 gene. *Molecular Endocrinology*, 19(6):1452–9, June 2005.

[115] Theresa Alenghat, Katherine Meyers, Shannon E. Mullican, Kirstin Leitner, Adetoun Adeniji-Adele, Jacqueline Avila, Maja Bućan, Rexford S. Ahima, Klaus H. Kaestner, and Mitchell A. Lazar. Nuclear receptor corepressor and histone deacetylase 3 govern circadian metabolic physiology. *Nature*, 456(7224):997–1000, December 2008.

[116] Trey K. Sato, Satchidananda Panda, Loren J. Miraglia, Teresa M. Reyes, Radu D. Rudic, Peter McNamara, Kinnery A. Naik, Garret A. FitzGerald, Steve A. Kay, and John B. Hogenesch. A functional genomics strategy reveals Rora as a component of the mammalian circadian clock. *Neuron*, 43(4):527–37, August 2004.

[117] Christine Crumbley, Yongjun Wang, Douglas J. Kojetin, and Thomas P. Burris. Characterization of the core mammalian clock component, NPAS2, as a REV-ERBalpha/RORalpha target gene. *The Journal of Biological Chemistry*, 285(46):35386–92, November 2010.

[118] Christine Crumbley and Thomas P. Burris. Direct regulation of CLOCK expression by REV-ERB. *PloS ONE*, 6(3):e17290, January 2011.

[119] Isabelle Schmutz, Jürgen A. Ripperger, Stéphanie Baeriswyl-Aebischer, and Urs Albrecht. The mammalian clock component PERIOD2 coordinates circadian output by interaction with nuclear receptors. *Genes & Development*, 24(4):345–57, February 2010.

[120] Katja A. Lamia, Stephanie J. Papp, Ruth T. Yu, Grant D. Barish, N. Henriette Uhlenhaut, Johan W. Jonker, Michael Downes, and Ronald M. Evans. Cryptochromes mediate rhythmic repression of the glucocorticoid receptor. *Nature*, 480(7378):552–6, December 2011.

[121] Yasukazu Nakahata, Milota Kaluzova, Benedetto Grimaldi, Saurabh Sahar, Jun Hirayama, Danica Chen, Leonard P Guarente, and Paolo Sassone-Corsi. The NAD+-dependent deacetylase SIRT1 modulates CLOCK-mediated chromatin remodeling and circadian control. *Cell*, 134(2):329–40, July 2008.

[122] Anne Brunet, Lora B. Sweeney, J. Fitzhugh Sturgill, Katrin F. Chua, Paul L. Greer, Yingxi Lin, Hien Tran, Sarah E. Ross, Raul Mostoslavsky, Haim Y. Cohen, Linda S. Hu, Hwei-Ling Cheng, Mark P. Jedrychowski, Steven P. Gygi, David A. Sinclair, Frederick W. Alt, and Michael E. Greenberg. Stress-dependent regulation of FOXO transcription factors by the SIRT1 deacetylase. *Science*, 303(5666):2011–5, March 2004.

[123] Qing-Hua Li and Hong-Quan Yang. Cryptochrome signaling in plants. *Photochemistry and Photobiology*, 83(1):94–101, 2007.

[124] Yasukazu Nakahata, Saurabh Sahar, Giuseppe Astarita, Milota Kaluzova, and Paolo Sassone-Corsi. Circadian control of the NAD+ salvage pathway by CLOCK-SIRT1. *Science*, 324(5927):654–7, May 2009.

[125] Luciano DiTacchio, Hiep D. Le, Christopher Vollmers, Megumi Hatori, Michael Witcher, Julie Secombe, and Satchidananda Panda. Histone Lysine Demethylase JARID1a Activates CLOCK-BMAL1 and Influences the Circadian Clock. *Science*, 333(6051):1881–5, September 2011.

[126] Saurabh Sahar and Paolo Sassone-Corsi. Regulation of metabolism: the circadian clock dictates the time. *Trends in Endocrinology and Metabolism*, 23(1):1–8, January 2012.

[127] Katja A. Lamia, Uma M. Sachdeva, Luciano DiTacchio, Elliot C. Williams, Jacqueline G. Alvarez, Daniel F. Egan, Debbie S. Vasquez, Henry Juguilon, Satchidananda Panda, Reuben J. Shaw, Craig B. Thompson, and Ronald M. Evans. AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation. *Science*, 326(5951):437–40, October 2009.

[128] Jee Hyun Um, Shutong Yang, Shin Yamazaki, and Hyeog Kang. Activation of 5-AMP-activated Kinase with Diabetes Drug Metformin Induces Casein Kinase I epsilon (CKI epsilon) -dependent Degradation of Clock Protein mPer2. *Journal of Biological Chemistry*, 282(29):20794–8, 2007.

[129] Xiaoyong Yang, Michael Downes, Ruth T. Yu, Angie L. Bookout, Weimin He, Marty Straume, David J. Mangelsdorf, and Ronald M. Evans. Nuclear receptor expression links the circadian clock to metabolism. *Cell*, 126(4):801–10, August 2006.

[130] Coralie Fontaine, Guillaume Dubois, Yannick Duguay, Torben Helledie, Ngoc Vu-Dac, Philippe Gervois, Fabrice Soncin, Susanne Mandrup, Jean-Charles Fruchart, Jamila Fruchart-Najib, and Bart Staels. The orphan nuclear receptor Rev-Erbalpha is a peroxisome proliferator-activated receptor (PPAR) gamma target gene and promotes PPARgamma-induced adipocyte differentiation. *The Journal of Biological Chemistry*, 278(39):37672–80, September 2003.

[131] Benedetto Grimaldi, Marina Maria Bellet, Sayako Katada, Giuseppe Astarita, Jun Hirayama, Rajesh H. Amin, James G. Granneman, Daniele Piomelli, Todd Leff, and Paolo Sassone-Corsi. PER2 controls lipid metabolism by direct regulation of PPARγ. *Cell Metabolism*, 12(5):509–20, November 2010.

[132] Chang Liu, Siming Li, Tiecheng Liu, Jimo Borjigin, and Jiandie D Lin. Transcriptional coactivator PGC-1alpha integrates the mammalian clock and energy metabolism. *Nature*, 447:477–81, May 2007.

[133] Sato Honma, Takeshi Kawamoto, Yumiko Takagi, and Katsumi Fujimoto. Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature*, 419:841–4, 2002.

[134] Ayumu Nakashima, Takeshi Kawamoto, Kiyomasa K. Honda, Taichi Ueshima, Mitsuhide Noshiro, Tomoyuki Iwata, Katsumi Fujimoto, Hiroshi Kubo, Sato Honma, Noriaki Yorioka, Nobuoki Kohno, and Yukio Kato. DEC1 modulates the circadian phase of clock gene expression. *Molecular and Cellular Biology*, 28(12):4080–92, 2008.

[135] E. Falvey, F. Fleury-Olela, and U. Schibler. The rat hepatic leukemia factor (HLF) gene encodes two transcriptional activators with distinct circadian rhythms, tissue distributions and target preferences. *The EMBO journal*, 14(17):4307–17, September 1995.

[136] Philippe Fonjallaz, Vincent Ossipow, Gerhard Wanner, and Ueli Schibler. The two PAR leucine zipper proteins, TEF and DBP, display similar circadian and tissue-specific expression, but have different target promoter preferences. *The EMBO journal*, 15(2):351–62, January 1996.

[137] Shigeru Mitsui, Shun Yamaguchi, Takuya Matsuo, Yoshiki Ishida, and Hitoshi Okamura. Antagonistic role of E4BP4 and PAR proteins in the circadian oscillatory mechanism. *Genes & Development*, 15:995–1006, 2001.

[138] Frédéric Gachon, Fabienne Fleury Olela, Olivier Schaad, Patrick Descombes, and Ueli Schibler. The circadian PAR-domain basic leucine zipper transcription factors DBP, TEF, and HLF modulate basal and inducible xenobiotic detoxification. *Cell Metabolism*, 4(1):25–36, July 2006.

[139] Daisuke Yamajuku, Yasutaka Shibata, Masashi Kitazawa, Toshie Katakura, Hiromi Urata, Tomoko Kojima, Satoko Takayasu, Osamu Nakata, and Seiichi Hashimoto. Cellular DBP and E4BP4 proteins are critical for determining the period length of the circadian oscillator. *FEBS Letters*, 585(14):2217–22, July 2011.

[140] Frédéric Gachon, Emi Nagoshi, Steven A. Brown, Juergen Ripperger, and Ueli Schibler. The mammalian circadian timing system: from gene expression to physiology. *Chromosoma*, 113(3):103–12, September 2004.

[141] Benoît Kornmann, Olivier Schaad, Hermann Bujard, Joseph S. Takahashi, and Ueli Schibler. System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. *PLoS Biology*, 5(2):e34, February 2007.

[142] Karine Spiegel, Esra Tasali, Rachel Leproult, and Eve Van Cauter. Effects of poor and short sleep on glucose metabolism and obesity risk. *Nature Reviews Endocrinology*, 5(5):253–61, May 2009.

[143] Norihiko Takeda and Koji Maemura. Circadian clock and vascular disease. *Hypertens Research*, 33(7):645–51, July 2010.

[144] Tami A. Martino, Gavin Y. Oudit, Andrew M. Herzenberg, Nazneen Tata, Margaret M. Koletar, Golam M. Kabir, Denise D. Belsham, Peter H. Backx, Martin R. Ralph, and Michael J. Sole. Circadian rhythm disorganization produces profound cardiovascular and renal disease in hamsters. *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, 294(5):R1675–83, May 2008.

[145] Tami A. Martino, Nazneen Tata, Denise D. Belsham, Jennifer Chalmers, Marty Straume, Paul Lee, Horia Pribiag, Neelam Khaper, Peter P. Liu, Fayez Dawood, Peter H. Backx, Martin R. Ralph, and Michael J. Sole. Disturbed diurnal rhythm alters gene expression and exacerbates cardiovascular disease with rescue by resynchronization. *Hypertension*, 49(5):1104–13, May 2007.

[146] Katja Vanselow, Jens T. Vanselow, Pål O. Westermark, Silke Reischl, Bert Maier, Thomas Korte, Andreas Herrmann, Hanspeter Herzel, Andreas Schlosser, and Achim Kramer. Differential effects of PER2 phosphorylation: molecular basis for the human familial advanced sleep phase syndrome (FASPS). *Genes & Development*, 20(19):2660–72, October 2006.

[147] Y. Xu, K. L. Toh, C. R. Jones, J.-Y. Shin, Y.-H. Fu, and L. J. Ptácek. Modeling of a human circadian mutation yields insights into clock regulation by PER2. *Cell*, 128(1):59–70, January 2007.

[148] T. Hamilton. Influence of environmental light and melatonin upon mammary tumour induction. *Br J Surg*, 56(10):764–7, 1969.

[149] Eva S. Schernhammer, Francine Laden, Frank E. Speizer, Walter C. Willett, David J. Hunter, Ichiro Kawachi, and Graham A. Colditz. Rotating night shifts and risk of breast cancer in women participating in the nurses' health study. *Journal of the National Cancer Institute*, 93(20):1563–8, October 2001.

[150] Johnni Hansen. Increased breast cancer risk among women who work predominantly at night. *Epidemiology*, 12(1):74–7, January 2001.

[151] Elisabeth Filipski, Verdun M. King, XiaoMei Li, Teresa G. Granda, Marie-Christine Mormont, XuHui Liu, Bruno Claustrat, Michael H. Hastings, and Francis Lévi. Host circadian clock as a control point in tumor progression. *Journal of the National Cancer Institute*, 94(9):690–7, May 2002.

[152] Shou-Tung Chen, Kong-Bung Choo, Ming-Feng Hou, Kun-Tu Yeh, Shou-Jen Kuo, and Jan-Gowth Chang. Deregulated expression of the PER1, PER2 and PER3 genes in breast cancers. *Carcinogenesis*, 26(7):1241–6, July 2005.

[153] Loning Fu, Helene Pelicano, Jinsong Liu, Peng Huang, and Cheng Lee. The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. *Cell*, 111:41–50, October 2002.

[154] Benoît Kornmann, Nicolas Preitner, Danièle Rifat, Fabienne Fleury-Olela, and Ueli Schibler. Analysis of circadian liver gene expression by ADDER, a highly sensitive method for the display of differentially expressed mRNAs. *Nucleic Acids Research*, 29(11):E51, June 2001.

[155] Takuya Matsuo, Shun Yamaguchi, Shigeru Mitsui, Aki Emi, Fukuko Shimoda, and Hitoshi Okamura. Control mechanism of the circadian clock for timing of cell division in vivo. *Science*, 302:255–9, October 2003.

[156] Hiroo Nakagawa, Satoru Koyanagi, Takako Takiguchi, Yukako Kuramoto, Shinji Soeda, Hiroshi Shimeno, Shun Higuchi, and Shigehiro Ohdo. 24-Hour Oscillation of Mouse Methionine Aminopeptidase2, a Regulator of Tumor Progression, Is Regulated By Clock Gene Proteins. *Cancer Research*, 64(22):8328–33, November 2004.

[157] W. Todd Lowther and Brian W. Matthews. Structure and function of the methionine aminopeptidases. *Biochimica et Biophysica Acta*, 1477:157–67, March 2000.

[158] Sigal Gery, Naoki Komatsu, Lilit Baldjyan, Andrew Yu, Danielle Koo, and H Phillip Koeffler. The circadian gene per1 plays an important role in cell growth and DNA damage control in human cancer cells. *Molecular Cell*, 22(3):375–82, May 2006.

[159] Aline Gréchez-Cassiau, Béatrice Rayet, Fabienne Guillaumond, Michèle Teboul, and Franck Delaunay. The circadian clock component BMAL1 is a critical regulator of p21WAF1/CIP1 expression and hepatocyte proliferation. *The Journal of Biological Chemistry*, 283(8):4535–42, February 2008.

[160] Elzbieta Kowalska, Juergen A. Ripperger, Dominik C. Hoegger, Pascal Bruegger, Thorsten Buch, Thomas Birchler, Anke Mueller, Urs Albrecht, Claudio Contaldo, and Steven A. Brown. Feature Article: NONO couples the circadian clock to the cell cycle. *Proceedings of the National Academy of Sciences of the United States of America*, 110(5):1592–9, December 2012.

[161] Tim Hunt and Paolo Sassone-Corsi. Riding tandem: circadian clocks and the cell cycle. *Cell*, 129(3):461–4, May 2007.

[162] Michele A. Gauger and Aziz Sancar. Cryptochrome, Circadian Cycle, Cell Cycle Checkpoints, and Cancer. *Cancer research*, 65(15):6828–34, August 2005.

[163] Susan Jones. An overview of the basic helix-loop-helix proteins. *Genome Biology*, 5(6):226, 2004.

[164] Cornelis Murre, Patrick Schonleber McCaw, H. Vaessin, M. Caudy, L. Y. Jan, Y. N. Jan, Carlos V. Cabrera, Jean N. Buskin, Stephen D. Hauschka, Andrew B. Lassar, Harold Weintraub, and David Baltimore. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*, 58:537–44, August 1989.

[165] Yasukazu Nakahata, Mayumi Yoshida, Atsuko Takano, Haruhiko Soma, Takuro Yamamoto, Akio Yasuda, Toru Nakatsu, and Toru Takumi. A direct repeat of E-box-like elements is required for cell-autonomous circadian rhythm of clock genes. *BMC Molecular Biology*, 9:1–11, January 2008.

[166] Zixi Wang, Yaling Wu, Lanfen Li, and Xiao-Dong Su. Intermolecular recognition revealed by the complex structure of human CLOCK-BMAL1 basic helix-loop-helix domains with E-box DNA. *Cell Research*, pages 1–12, December 2012.

[167] Seung-Hee Yoo, Caroline H. Ko, Phillip L. Lowrey, Ethan D. Buhr, Eun-joo Song, Suhwan Chang, Ook Joon Yoo, Shin Yamazaki, Choogon Lee, and Joseph S. Takahashi. A noncanonical E-box enhancer drives mouse Period2 circadian oscillations in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7):2608–13, February 2005.

[168] Hiroki R. Ueda, Satoko Hayashi, Wenbin Chen, Motoaki Sano, Masayuki Machida, Yasufumi Shigeyoshi, Masamitsu Iino, and Seiichi Hashimoto. System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nature Genetics*, 37(2):187–92, March 2005.

[169] Ruth A. Akhtar, Akhilesh B. Reddy, Elizabeth S. Maywood, Jonathan D. Clayton, Verdun M. King, Andrew G. Smith, Timothy W. Gant, Michael H. Hastings, and Charalambos P. Kyriacou. Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. *Current Biology*, 12(7):540–50, April 2002.

[170] Kai-Florian Storch, Ovidiu Lipan, Igor Leykin, N. Viswanathan, Fred C. Davis, Wing H. Wong, and Charles J. Weitz. Extensive and divergent circadian gene expression in liver and heart. *Nature*, 417(6884):78–83, May 2002.

[171] Kristina M. Smith, Gencer Sancar, Rigzin Dekhang, Christopher M. Sullivan, Shaojie Li, Andrew G. Tag, Cigdem Sancar, Erin L. Bredeweg, Henry D. Priest, Ryan F. McCormick, Terry L. Thomas, James C. Carrington, Jason E. Stajich, Deborah Bell-Pedersen, Michael Brunner, and Michael Freitag. Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for neurospora white collar complex. *Eukaryotic Cell*, 9(10):1549–56, October 2010.

[172] Fumiyuki Hatanaka, Chiaki Matsubara, Jihwan Myung, Takashi Yoritaka, Naoko Kamimura, Shuichi Tsutsumi, Akinori Kanai, Yutaka Suzuki, Paolo Sassone-Corsi, Hiroyuki Aburatani, Sumio Sugano, and Toru Takumi. Genome-wide profiling of the core clock protein BMAL1 targets reveals a strict relationship with metabolism. *Molecular and Cellular Biology*, 30(24):5636–48, 2010.

[173] Jan Pontén and Eero Saksela. Two established in vitro cell lines from human mesenchymal tumours. *International Jurnal of Cancer*, 2(5):434–47, September 1967.

[174] Katerina N. Niforou, Athanasios K. Anagnostopoulos, Konstantinos Vougas, Christos Kittas, Vassilis G. Gorgoulis, and George T. Tsangaris. The proteome profile of the human osteosarcoma U2OS cell line. *Cancer Genomics & Proteomics*, 5:63–78, 2008.

[175] Robert J. Isfort, David B. Cody, Glenda Lovell, and Claus-Jens Doersen. Analysis of oncogenes, tumor suppressor genes, autocrine growth-factor production, and differentiation state of human osteosarcoma cell lines. *Molecular Carcinogenesis*, 14(3):170–8, November 1995.

[176] Eric E. Zhang, Andrew C. Liu, Tsuyoshi Hirota, Loren J. Miraglia, Genevieve Welch, Pagkapol Y. Pongsawakul, Xianzhong Liu, Ann Atwood, Jon W. Huss, Jeff Janes, Andrew I. Su, John B. Hogenesch, and Steve A. Kay. A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell*, 139(1):199–210, October 2009.

[177] Aurélio Balsalobre. Resetting of Circadian Time in Peripheral Tissues by Glucocorticoid Signaling. *Science*, 289(5488):2344–7, September 2000.

[178] André Engling, Rafael Backhaus, Carolin Stegmayer, Christoph Zehe, Claudia Seelenmeyer, Angelika Kehlenbach, Blanche Schwappach, Sabine Wegehingel, and Walter Nickel. Biosynthetic FGF-2 is targeted to non-lipid raft microdomains following translocation to the extracellular surface of CHO cells. *Journal of Cell Science*, 115(18):3619–31, September 2002.

[179] Frederick R. Blattner. The Complete Genome Sequence of Escherichia coli K-12. *Science*, 277(5331):1453–62, September 1997.

[180] Wei-Shau Hu and Vinay K. Pathak. Design of retroviral vectors and helper cells for gene therapy. *Pharmacological Reviews*, 52(4):493–511, December 2000.

[181] Michael A. Bender, Theo D. Palmer, Richard E. Gelinas, and A. Dusty Miller. Evidence that the packaging signal of Moloney murine leukemia virus extends into the gag region. *Journal of Virology*, 61(5):1639–46, May 1987.

[182] Ina Weidenfeld. *A conditional mammalian expression system based on a retargetable genomic locus (Doctoral Thesis)*. PhD thesis, Ruperto-Carola University of Heidelberg, Germany, 2009.

[183] Rainer Loew, Elisa Vigna, Dirk Lindemann, Luigi Naldini, and Herman Bujard. Retroviral vectors containing Tet-controlled bidirectional transcription units for simultaneous regulation of two gene activities. *Journal of Molecular and Genetic Medicine*, 2(1):107–18, January 2006.

[184] Christian J. Buchholz and Klaus Cichutek. Is it going to be SIN ? *The Journal of Gene Medicine*, 8:1274–6, 2006.

[185] David E. Pegg. Principles of cryopreservation. *Methods in Molecular Bology*, 368(2):39–57, January 2007.

[186] R. Keith Barrett and Joseph S. Takahashi. Temperature Compensation and Temperature Entrainment of the Chick Pineal Cell Circadian Clock. *Journal of Neuroscience*, 15(8):5681–92, 1995.

[187] Aurélio Balsalobre, Lysiane Marcacci, and Ueli Schibler. Multiple signaling pathways elicit circadian gene expression in cultured Rat-1 fibroblasts. *Current Biology*, 10(20):1291–4, October 2000.

[188] U.K. Laemmli. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227:680–5, 1970.

[189] Harry Towbin, Theophil Staehelin, and Julian Gordon. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America*, 76(9):4350–3, January 1979.

[190] Randall K. Saiki, Stephen Scharf, Fred Faloona, Kary B. Mullis, Glenn T. Horn, Henry A. Erlich, and Norman Arnheim Arnheim. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230:1350–4, January 1985.

[191] Kary B. Mullis, Fred Faloona, Stephen Scharf, Randall K. Saiki, Glenn T. Horn, and Henry A. Erlich. No Title. *Cold Spring Harb Symp Quant Biol.*, 51:263–73, 1986.

[192] W. Scott Furey, Catherine M. Joyce, Mark A. Osborne, David Klenerman, James A. Peliska, and Shankar Balasubramanian. Use of fluorescence resonance energy transfer to investigate the conformation of DNA substrates bound to the Klenow fragment. *Biochemistry*, 37(9):2979–90, March 1998.

[193] Abcam. CROSS-LINKING CHROMATIN IMMUNOPRECIPITATION. 2008.

[194] Ben Langmead, Cole Trapnell, Mihai Pop, and Steven L. Salzberg. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3):R25, January 2009.

[195] Yong Zhang, Tao Liu, Clifford A. Meyer, Jérôme Eeckhoute, David S. Johnson, Bradley E. Bernstein, Chad Nusbaum, Richard M. Myers, Myles Brown, Wei Li, and X. Shirley Liu. Model-based analysis of ChIP-Seq (MACS). *Genome Biology*, 9(9):R137, January 2008.

[196] Aaron R. Quinlan and Ira M. Hall. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6):841–2, March 2010.

[197] Gwendal Le Martelot, Donatella Canella, Laura Symul, Eugenia Migliavacca, Federica Gilardi, Robin Liechti, Olivier Martin, Keith Harshman, Mauro Delorenzi, Béatrice Desvergne, Winship Herr, Bart Deplancke, Ueli Schibler, Jacques Rougemont, Nicolas Guex, Nouria Hernandez, and Felix Naef. Genome-Wide RNA Polymerase II Profiles and RNA Accumulation Reveal Kinetics of Transcription and Associated Epigenetic Changes During Diurnal Cycles. *PLoS Biology*, 10(11):e1001442, November 2012.

[198] Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, 57(1):289–300, 1995.

[199] Momoko Sasaki, Hikari Yoshitane, Ngoc-Hien Du, Toshiyuki Okano, and Yoshitaka Fukada. Preferential inhibition of BMAL2-CLOCK activity by PER2 reemphasizes its negative role and a positive role of BMAL2 in the circadian transcription. *The Journal of Biological Chemistry*, 284(37):25149–59, September 2009.

[200] Derek M. Murphy, Patrick G. Buckley, Kenneth Bryan, Sudipto Das, Leah Alcock, Niamh H. Foley, Suzanne Prenter, Isabella Bray, Karen M. Watters, Desmond Higgins, and Raymond L. Stallings. Global MYCN Transcription Factor Binding Analysis in Neuroblastoma Reveals Association with Distinct E-Box Motifs and Regions of DNA Hypermethylation. *PloS ONE*, 4(12):e8154, 2009.

[201] Maki Ukai-Tadenuma, Rikuhiro G. Yamada, Haiyan Xu, Jürgen A. Ripperger, Andrew C. Liu, and Hiroki R. Ueda. Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell*, 144:266–81, 2011.

[202] N. Ohkura, K. Oishi, N. Fukushima, M. Kasamatsu, G-I. Atsumi, N. Ishida, S. Horie, and J. Matsuda. Circadian clock molecules CLOCK and CRYs modulate fibrinolytic activity by regulating the PAI-1 gene expression. *Journal of Thrombosis and Haemostatsis*, 4:2478–85, 2006.

[203] Xiang Mou, Cynthia B. Peterson, and Rebecca A. Prosser. Tissue-type plasminogen activator-plasmin-BDNF modulate glutamate-induced phase-shifts of the mouse suprachiasmatic circadian clock in vitro. *European Journal of Neuroscience*, 30(8):1451–60, October 2009.

[204] J. A. Hamilton, S. Lingelbach, N. C. Partridge, and T. J. Martin. Regulation of plasminogen activator production by bone-resorbing hormones in normal and malignant osteoblasts. *Endocrinology*, 116(6):2186–91, June 1985.

[205] Bayard D. Catherwood, Louisa Titus, Scott D. Boden, and Mark S. Nanes. Increased Expression of Tissue Plasminogen Activator Messenger Ribonucleic Acid Is an Immediate Response to Parathyroid Hormone in Neonatal Rat Osteoblasts. *Endocrinology*, 134(3):1429 – 36, 1994.

[206] Yosuke Kanno, Akira Ishisaki, Eri Kawashita, Naoyuki Chosa, Keiichi Nakajima, Tatsuji Nishihara, Kuniaki Toyoshima, Kiyotaka Okada, Shigeru Ueshima, Kenji Matsushita, Osamu Matsuo, and Hiroyuki Matsuno. Plasminogen/plasmin modulates bone metabolism by regulating the osteoblast and osteoclast function. *The Journal of Biological Chemistry*, 286(11):8952–60, March 2011.

[207] H .Criss Hartzell, Kuai Yu, Qinhuan Xiao, Li-Ting Chien, and Zhiqiang Qu. Anoctamin/TMEM16 family members are Ca2+-activated Cl- channels. *Journal of Physiology*, 587(Pt 10):2127–39, May 2009.

[208] Shihoko Kojima, Elaine L. Sher-Chen, and Carla B. Green. Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. *Genes & Development*, 26(24):2724–36, December 2012.

[209] Elizabeth A. Schroder, Karyn A. Esser, Jonathan C. Makielski, and Brian P. Delisle. Circadian Variation in SCN5A Driven by the Molecular Clock (Poster-Abstract). In *Conference Poster (unpublished)*, 2011.

[210] Christopher Vollmers, Robert J. Schmitz, Jason Nathanson, Gene Yeo, Joseph R. Ecker, and Satchidananda Panda. Circadian Oscillations of Protein-Coding and Regulatory RNAs in a Highly Dynamic Mammalian Liver Epigenome. *Cell Metabolism*, 16(6):833–45, December 2012.

B. List of Abbreviations

Table B.1.: List of Abbreviations A-C

A-C	
ABCG1	ATP-Binding Cassette, Sub-Family G (WHITE), Member 1
ACAD	Acyl-CoA Dehydrogenase
ACADL	Acyl-CoA Dehydrogenase, Long Chain
AFF3	AF4/FMR2 Family, Member 3
AMP	Adenosine Monophosphate
AMPK	AMP-Activated Protein Kinase
ANO1/2	Anoctamin 1/2
AOX1	Aldehyde Oxidase 1
ATCC	American Type Culture Collection
ATM	Ataxia-Teleangiectasia Mutated
ATP	Adenosine-5\strich -Triphosphate
ATP1B1	ATPase, Na+/K+ Transporting, Beta 1 Polypeptide
betaTrCP1	Beta-Transducin Repeat Containing
BDNF	Brain-Derived Neurotrophic Factor
bHLH	Basic Helix-Loop-Helix
BMAL1	Brain And Muscle Arnt Like 1
bZIP	Basic Leucine Zipper
CAMK2D	Calcium/Calmodulin-Dependent Protein Kinase II Delta
CCA1	Circadian Clock Associates 1
CCDC85B	Coiled-Coil Domain Containing 85B
ChIP	Chromatin Immunoprecipitation
CHK2	Checkpoint Kinase 2
CK1/2	Casein Kinase 1/2
CIP	Calf-Intenstinal Alkaline Phosphatase
Clk	Clock
CLOCK	Circadian Locomotor Output Cycles Kaput
CLTCL1	Clathrin Heavy Chain-Like 1
COL17A1	Collagen, Type XVII, Alpha 1
COL22A1	Collagen, Type XXII, Alpha 1
CPEB1	Cytoplasmic Polyadenylation Element Binding protein
CRY	Cryptochrome
CRBS	Circadian Regulator Binding Site
CSP1	Conidial Separation 1
CTP	Cytidine-5\strich -Triphosphate
Cy3/5	Cyanine 3/5
Cyc	Cycle
c-MYC	v-myc Myelocytomatosis Viral Oncogene Homolog (avian)

Table B.2.: List of Abbreviations D-G

D-G	
DBP	Albumin D-Site-Binding Protein
Dbt	Double-Time
DEAE	Diethylethanolamine
DEC1/2	Differentially Expressed in Chondrocytes 1
DDX5	DEAD Box Helicase 5
DHX9	DEAH Box Polypeptide 9
DMEM	Dulbecco's Modified Eagles Medium
DMSO	Dimethyl Sulfoxide
DN1-3	Dorsal Neuronal
E4BP4	E4 Promoter-Binding Protein
EDTA	Ethylenediaminetetraacetic Acid
EIF2C1	Eukaryotic Translation Initiation Factor 2C, 1
EIF5A2	Eukaryotic Translation Initiation Factor 5A2
ELF3/4	Early Flowering 3/4
env	Envelope
FACS	Fluorescence Activated Cell Sorting
6-FAM	Carboxyfluorescein
FASPS	Familial Advanced Sleep-Phase Syndrom
FBXL3	F-box And Leucine-Rich Repeat Protein 3
FCS	Fetal Calf Serum
FDR	False Discovery Rate
FGFR1	Fibroblast Growth Factor Receptor 1
FOXJ2	Forkhead Box J2
FOXO	Forkhead Box, Subclass O
FRH	Frequency Interacting Helicase
FRQ	Frequency
FSC	Forward Scatter
GADD45A	Growth Arrest And DNA-Damage-Inducible, Alpha
gag	Group-Specific Antigen
GI	Gigantea
GRE	Glucocorticoid-Response-Element
GYG1	Glycogenin 1

Table B.3.: List of Abbreviations H-O

H-O	
H6PD	Hexose-6-Phosphate Dehydrogenase (Glucose 1-Dehydrogenase)
HDAC	Histone Deacetylase
HEK	Human Embryonic Kidney Cells
HLF	Hepatic Leukemia Factor
HMGA2	High Mobility Group AT-hook 2
HNRPDL	Heterogeneous Nuclear Ribonucleoprotein D-Like
JARID1a	Jumonji, AT Rich Interactive Domain 1a
KLF11	Kruppel-Like Factor 11
LB-Amp	LB-Ampicilin
LHY	Late And Elongated Hypocotyl
LN	Lateral Neuronal
LOV	Light Oxygen Voltage
LTR	Long Terminal Repeat
LXR	Liver X Nuclear Receptor
MCAT	Murine Cation Transporter
MCS	Multiple Cloning Site
MDM2	Mouse Double-Minute
METAP2	Methionyl Aminopeptidase
MLV	Murine Leukemia Virus
MCAT	Murine Cation Transporter
MLV	Murine Leukemia Virus
MRA	Mycoplasma Removal Agent
MTHFD1L	Methylenetetrahydrofolate Dehydrogenase (NADP+ dependent) 1-Like
NAMPT	Nicotinamide Phosphoribosyltransferase
NCoR1	Nuclear Receptor Corepressor 1
NONO	Non-POU Domain Containing, Octamer Binding
NPAS2	Neuronal PAS Domain Protein 2
NR1D1/2	Nuclear Receptor Subfamily 1, Group D, Member 1/2
NRG1	Neuregulin 1
OSGIN1	Oxidative Stress Induced Growth Inhibitor 1

Table B.4.: List of Abbreviations P-R

P-R	
PAR-domain	Proline And Acidic Amino Acid-Rich-Domain
PAS	Per Arnt Sim
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDE4B	Phosphodiesterase 4B, cAMP-Specific
Pdp1	PAR-Domain Protein 1
PER	Period
PEST	Proline (P), Glutamic Acid (E), Serine (S), Threonine (T)
PGC1A	Peroxisome Proliferator-Activated Receptor Gamma Coactivator 1A
PLAT	Tissue Plasminogen Activator (also tPA)
PLAU	Plasminogen Activator (also Urokinase)
pol	Polymerase
PP2A	Protein Phosphatase A
PPAR A/D/G	Peroxisome Proliferator-Activated Receptor Alpha/Delta/Gamma
PPP1R3B	Protein Phosphatase 1 Subunit 3B
PPP3CA	Protein Phosphatase 3, Catalytic Subunit, Alpha Isozyme
PREX1	Phosphatidylinositol-3,4,5-trisphosphate-dependent Rac Exchange Factor 1
PRKCH	Protein Kinase C
pro	Protease
PRR5/7/9	Pseudo Response Regulator
PSF	PTB-Associated Splicing Factor
PYGL	Phosphorylase, Glycogen, Liver
qPCR	quantitative PCR
R	Repeat
RALGDS	Ral Guanine Nucleotide Dissociation Stimulator
ROR	RAR-Related Orphan Receptors
ROREs	Retinoic Acid-Related Orphan Receptor Response Element

Table B.5.: List of Abbreviations S-Z

S-Z	
SBS	Sequencing-By-Synthesis
SCF	Skp1-Cullin-F-box
SCN	Suprachiasmatic Nucleus
SDS-PAGE	Sodium-Dodecyl-Sulfate Polyacrylamide-Gelelectrophoresis
SETX	Senataxin
SIN	Self-Inactivating
SIN3	SIN3 Transcription Regulator Homolog A (yeast)
SIRT1	Sirtuin
SLC2A1	Solute Carrier Family 2 (Facilitated Glucose Transporter), Member 1
SSC	Side Scatter
SV40	Simian Virus 40
SULF2	Sulfatase 2
TAE	TRIS Acetic Acid EDTA
TBS	TRIS Buffered Saline
TEF	Thyrotroph Embryonic Factor
TFEB	Transcription Factor EB
Tim	Timeless
TNFSF15	Tumor Necrosis Factor (ligand) Superfamily, Member 15
TOC1	Time Of Cab Expression 1
tPA	Tissue Plasminogen Activator (also PLAT)
TSS	Transcription Start Site
TSC22D3	TSC22 domain family, Member 3
U3/5	Unique 3\strich/5\strich
UPL	Universal Probe Library
UV	Ultraviolet
VVD	Vivid
WC-1/2	White Collar 1/2
WCC	White Collar Complex
WDR5	WD Repeat Domain 5
WNT5B	Wingless-Type MMTV Integration Site Family, Member 5B
ZNF323	Zinc Finger Protein 323
ZTL	Zeitlupe

C. List of Figures

D. List of Tables

Acknowledgments

First of all, I would like to thank everybody who contributed directly to the work in my thesis, especially my supervisor Prof. Dr. Michael Brunner, who decided to extend his research from Neurospora Crassa to mammals and enabled me to do the first PhD thesis on mammalian cell culture in his lab. I am very grateful for the time he invested on the many, many discussions we had on the project. I also want to thank our collaborators in Lausanne, Prof. Dr. Felix Naef and his PhD student Laura Symul. Laura did a great job on the bioinformatical analysis of my experimental data and helped me carry out my own analyses using informatics. I don't know what I would have done without her. Thanks also to Dr. Tamás Fischer, who provided the know-how for the microarray-experiments and as such helped me go beyond a difficult point in my thesis. I thank everybody who helped me improve my dissertation by critical comments and Prof. Dr. Felix Wieland for his function as second referee of my thesis.

Then of course I would like to thank our entire working group, but especially the people who during the last four years were involved in the work carried out in the cell culture and therefore contributed at one point or the other to this work: Christian Maurer, Claudia Seelenmeyer, Axel Diernfellner, Christian Nahstoll, François Cesbron, and those that are now following up on the ideas that were started by my work: Réka Haraszti, Bianca Baumert, and Anton Shostak.

Last but not least I want to thank all the people that supported me throughout my thesis, my friends from HBIGS and the BZH, and my "lunch-dates". My most special thanks go to my parents, who always supported my ideas, even if they involved moving far away, and to my friend Monika Gaik and my love Max Hoffmann, who would sit down in a car early on a Saturday morning and bring me back to my parents if needed. The last couple of years were difficult in many ways, but you helped me live through that. Thank you so much for this and everything else.

Hiermit erkläre ich, dass ich die vorliegende Dissertation eigenständig und ohne Verwendung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäss aus veröffentlichten und nicht veröffentlichten Schriften entnommen sind, sind als solche gekennzeichnet. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Fakultät oder Prüfungsbehörde vorgelegen.

Heidelberg, den 12.03.2013 Julia Stefanski