Generating Cell-Specific CRISPR Knockouts to Investigate the Role of Neuropeptide Receptors on *Drosophila* Behavior

Senior Thesis

Presented to
The Faculty of Arts and Sciences
Brandeis University

Undergraduate Program in the Department of Biology
Dr. Michael Rosbash, Advisor

In partial fulfilment of the requirements for the degree of Bachelor of Science

By: Jason Xin
April 19, 2018

Committee Members:
Dr. Michael Rosbash
Dr. Michael Marr
Dr. Amy Lee
Abstract

Neuropeptides and their receptors have diverse and important roles in regulating physiology, including appetite, stress, inflammation and pain, and they function in Drosophila as well as mammals. My longer-term goal is to create a library of CRISPR-mediated knock-outs of all neuropeptide receptor genes in Drosophila, and then develop/use novel methods to create and assay adult cell-specific knock-outs of these genes. I have begun with the adult brain circadian clock network because it is known to be highly regulated by neuropeptides, in part to anticipate environmental changes during the 24 hour cycle. Knocking-out neuropeptide receptors in a neuron-specific manner may therefore provide important insights into communication within the circadian network. To assess our methods and strategies, I began by generating cell-specific knockouts of the core clock proteins Period (PER) and Timeless (TIM) to show that the tRNA-sgRNA CRISPR-Cas9 system works efficiently to disrupt these genes within the clock neural network. The results also have some surprising features. I have just begun to assay the application of this technique to neuropeptide receptors but have already confirmed that the short neuropeptide F receptor (sNPFR) positively and strongly regulates sleep. Further investigation promises to provide a deeper understanding of how neuropeptides function in neuronal communication and how they influence health and disease.
Introduction

The rotation of the earth results in different conditions during day and night. To cope with these ever changing conditions, life on earth has developed a time-measuring mechanisms that allows the anticipation of changes in environment during the 24 hour cycle, and this is called the circadian clock. Other than governing sleep-activity cycles, the circadian clock influences a variety of important processes in a rhythmic manner, including digestion, body temperature, blood pressure, hormone release, and metabolism [1]. There are a few fundamental features of circadian clocks. Firstly, the circadian clock persists in the absence of environmental cues, with a period slightly different from 24 hours. In order to reset the circadian clock such that it matches the local time, it is entrained, or adjusted by signals called Zeitgebers, meaning time-givers. These signals include light, temperature, activity, and availability of nutrients. Thirdly, the circadian clock is temperature compensated, meaning that the period of the clock persists over a wide range of temperature [1]. This is interesting as the rates of almost all biochemical reactions are very temperature sensitive.

The mechanism of maintaining these circadian rhythms involves interlocked transcriptional-translational feedback loops. These drive the rhythmic transcription of “clock” genes, encoding feedback loop components, as well as “output genes”, which regulate physiological and behavioral rhythms, as mentioned previously [2]. In mammals, it is the suprachiasmatic nucleus (SCN) of the hypothalamus that acts as the central clock [1]. The SCN receives input from the alternating light and dark cycles through photoreceptors in the eye, and relays the information to the rest of the body and brain, thereby coordinating the clocks in cells throughout the body and leading to alterations in gene expression and behavior [1]. This highlights the three main components of circadian systems in all organisms: the core clock, the input pathways, and the output pathways [3].
Although not all clock components are conserved between the different animal kingdoms, the transcription feedback loop mechanism by which the central clock operates is conserved [4,5]. The fruit fly, *Drosophila melanogaster* has been a key model organism in the understanding of how the circadian clock operates on a molecular level [6]. *Drosophila melanogaster* is such a useful and attractive model organism for various reasons- its genome has been completely sequenced, it has a rapid life cycle, and has a simpler neuronal circuitry compared to mammals, to name a few [7].

The central clock in *Drosophila melanogaster* resides in the head, composed of approximately 150 clock neurons that rhythmically express the core clock proteins [6]. These neurons are divided based upon both their anatomical position and size. There are the dorsal neurons which are further divided into 3 clusters, the DN1s, DN2s, and DN3s. There are also lateral neurons, which are further divided into 4 clusters, the small and large ventral lateral neurons (sLNvs and lLNvs), the dorsal lateral neurons (LNds), and the posterior lateral neurons (LPNs) [3]. In the past it was believed that these lateral neurons were the key pacemaker clock neurons, while the dorsal neurons had lesser more subtle roles in maintaining rhythmicity [8]. However, it is now known that there is a more complex relationship between the neuronal clusters for coordinating the daily activity pattern of the fly [8].

For *Drosophila melanogaster*, there is highest locomotor activity when the lights first come on, and when the lights go off, in a 12 hour light-dark cycle. These levels of maximum activity are referred to as the morning peak and the evening peak, respectively [9]. Between these two time periods there is a mid-day siesta, and after the evening peak is when there is consolidated sleep [10]. Before the bouts of morning and evening activity, there is increasing activity before lights on and lights off, referred to as anticipation. Through selectively ablating clock neurons, it was found that the LNvs expressing the neuropeptide pigment dispersing factor (PDF) are important for morning anticipation and are therefore the M oscillator, while LNds are important for evening anticipation, and are therefore the E oscillators [11]. These results suggest that the 2 oscillators reside in distinct clock neurons, and that there is
communication between the different neuronal groups to coordinate the behaviour in changing environmental conditions. However, these interactions are still not yet well understood.

Since the sequencing and annotation of the *Drosophila* genome [12], different neuropeptides have been identified in a variety of neuron types [13]. Neuropeptides are small proteinaceous substances produced and released by neurons which act on other neural substrates [14]. Neuropeptides and their respective receptors are known to play roles in various physiological responses such as appetite, mood, and inflammatory responses [15]. There are a few key differences between neuropeptides and classical neurotransmitters. The major difference between the two is in their biosynthesis. Neuropeptides are derived from larger, inactive precursors that are over 90 amino acid residues in length, and are synthesized at the cell soma. They are then transported down the axon before being stored in peptide-containing, large dense core vesicles (LDCVs) for exocytosis [16]. In contrast, neurotransmitters are synthesized in the cytoplasm of the presynaptic neuron and stored in small clear vesicles [16]. Neuropeptides are also present in tissues at lower concentrations, and are active at their receptors at lower concentrations [16]. Most neuropeptides act through binding to G-protein coupled receptors, leading to the activation of intracellular G-proteins [17]. The signal is then further amplified by inducing other intracellular pathways involving adenylyl cyclase, cAMP, MAPK/ERK, PKA, and phosphorylation of certain target proteins [17]. These signalling cascades can result in secretion of substances from the cell, opening of ion channels, and also impact transcription rates and therefore gene-expression levels [18]. However, GPCRs do not only bind to neuropeptides, but also bind an extremely diverse set of ligands, including proteins, hormone, drugs, photons, and small molecules [19]. In *Drosophila melanogaster*, a significant number of GPCRs are orphans [19], meaning that the endogenous ligand has not yet been identified, and is therefore open to new discoveries.

The most potent circadian communication signal in *Drosophila melanogaster* is the neuropeptide pigment dispersing factor (PDF) [6]. PDF has an important role as the pacemaker
coordinator [20], and is also important for promoting wakefulness and arousal [21]. Therefore, PDF is important for both internal signalling in the circadian network and also output signalling to downstream effectors [22]. The multiple functions of PDF can also be seen by observing the distribution of its receptor, PDFR. In the circadian network, all five of the sLNvs, 2 lLNvs, 3 LNds, both DN2, DN1a, and several DN1ps and DN3s express PDFR, highlighting its importance in internal signalling. Outside of the circadian network, there is PDFR expression at the ellipsoid body [21], a locomotor centre, as well as at the boundary of the lamina and retina, possible modulating the light-input pathway [23]. When *Drosophila* bears the loss-of-function *pdf*01 mutation, these flies characterized by a loss of the morning activity anticipation and an advanced evening peak in light/dark conditions, while showing high levels of arrhythmicity in constant darkness [24]. However, *Drosophila* bearing the *pdfr* mutation do show reduced morning anticipations, but oscillations of the core clock protein Period (PER) are similar to what is seen in wild-type flies under light/dark conditions [25]. This suggests that PDF may also be acting on other receptors in the circadian network.

Another neuropeptide that plays a significant role in the control of behavioral rhythms is the ion transport peptide (ITP). ITP is expressed in the 5th sLNv, 1 LNd, and in a few non-clock cells present in the brain [26]. Through the use of RNA interference to knockdown ITP in two clock cells, *Drosophila* showed reduced evening activity, as well as increased nocturnal activity. When ITP was overexpressed in the clock neurons, there was complete disruption of *Drosophila* behavioral rhythm [26]. Surprisingly, the overexpression of ITP only slightly dampened PER cycling in the important pacemaker neuron clusters [26], suggesting that ITP is important in the output pathways and not for communication within the circadian network. Interestingly, when there is simultaneous knockdown of both ITP and PDFR, the flies displayed hyperactive phenotypes and were almost completely arrhythmic under constant darkness [26].
The short neuropeptide F (sNPF) is a neuropeptide that is expressed in a large population of diverse neurons in the brain, with most of the neurons being interneurons of mushroom bodies, central nervous system structures that are implicated in memory and sleep [27]. When sNPF is knocked down in *Drosophila*, the flies exhibit decreased night-time sleep, and the same is observed when there is loss of function of the receptor sNPFR [28]. The neuropeptide sNPF also plays a role in the regulation of fly behaviour after sleep deprivation, as sNPF knockdown flies show much less rebound or recovery sleep after deprivation compared to wild type flies [28].

Recently, three different key subgroups of clock neurons have been purified, and the expressed genes at the different times of the day have been identified through RNA-sequencing [29]. This revealed transcripts encoding neuropeptides known to be expressed in the circadian network, several neuropeptides that are not known to be expressed, and also some receptors for neuropeptides [29]. Interestingly, the expression of the neuropeptides in the three clock neuron clusters were largely distinct, with little shared expression, indicating complex interaction and cell-specific circadian function [29]. Despite these signalling peptides being implicated in functions such as heartbeat and gland secretion [30], the exact physiological function is yet to be determined. The goal of the project is to use a novel CRISPR technique to create cell specific knockouts of various identified neuropeptide receptors and observe their impact on behaviour and circadian rhythm.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) are the hallmark of a bacterial defence mechanism [15]. They form the basis of the CRISPR/Cas9 system, which has revolutionized genome editing. The genome-editing technique requires a guide RNA composed of the tracrRNA (trans-activating CRISPR RNA) that binds to the Cas nuclease, and the crRNA (CRISPR-RNA) that binds to the target DNA sequence to provide specificity for cutting. However, there are still improvements that can be made to the efficiency of this technique, which has led to the development of a novel CRISPR technique incorporating transfer RNA-flanked sgRNAs (single guide RNA). This
development results in the simultaneous expression of multiple sgRNAs, an effective method for targeted cutting of multiple sites at the gene of interest [31]. Transgenic Drosophila containing a vector construct containing four sgRNAs targeting the genes sepia (se), ebony (e), curled (cu), and forked (f) crossed to flies expressing the actin-Cas9 promoter resulted in offspring that had a high penetrance of the visible phenotype, as well as biallelic disruption of se in 99% of eye tissue [31]. This novel CRISPR technique can therefore be used to circumvent the potential problems of functional in-frame mutations as well as single guide RNAs with low activity.

The ability to express the created vector in a spatially restricted manner is very important, and requires the GAL4/UAS (upstream activator sequence) binary system. This system requires transgenic flies that express the yeast transcription factor GAL4 under a Drosophila promoter, and a second transgenic fly line which has the designed vector containing UAS, the GAL4 binding site [32]. The progeny of these two fly lines will carry both transgenes, leading to expression of the vector in a spatial manner that is dependent on the promoter controlling GAL4 expression. Thus, fly lines with different promoters can be used in order to express GAL4 in specific clock neurons. Currently, there are roughly 3850 different lines in which the promoter drives the expression of GAL4 in a range of 20 to 5000 central-brain neurons [33].

In this project, I will be using a variety of GAL4 drivers that express the designed construct containing 3 tRNA flanked sgRNA in various different cell types. This technique can therefore be used to create cell-specific knockouts of the neuropeptide receptors of interest and investigate if the neuropeptides have cell-specific functions, or if they function in the whole clock network to coordinate the circadian rhythm. I will also be aiming to create a library of fly lines expressing the sgRNA constructs targeting the different neuropeptide receptors, which can then be used in future studies.
Materials and Methods

Cloning and Generating Transgenic Fly Lines

All oligonucleotides used in this study are listed in Supplementary Table 1.

The pCFD6 plasmid [31] was digested with the Bbs1 restriction enzyme and incubated at 37°C overnight. The digest was then run on a 1% agarose gel, and the band was cut out and extracted using a QIAquick Gel Extraction kit. A PCR reaction was setup with the oligos ordered in Table 1 and using the Q5 polymerase (NEB). The annealing temperature was set to 60°C with an elongation time of 20 seconds. The PCR products were run on a 1% agarose gel, and the bands were cut and extracted using the QIAquick Gel Extraction kit. A Gibson Assembly (NEB) reaction was then performed using ~50ng of digested backbone and a two-fold molar excess of each insert. The reaction mixture was then diluted with 45 uL of water, and 1uL of the Gibson Assembly mix was then used to transform chemically competent DH5α cells using standard procedures. The cells were plated on LB plates containing 100ug/ml ampicillin and incubated at 37°C overnight.

Single colonies were picked using a pipette tip and incubated in 40uL of mH2O. A colony PCR reaction was then setup using REDTaq PCR Reaction Mix (Sigma-Aldrich), with an annealing temperature of 60°C, elongation time of 1 minute, and using the following primers:

fwd: 5’ – GTAGACATCAAGCATCGGTGG – 3’

rev: 5’ – TTAGAGCTTTAAATCTCTGTAGGTAG – 3’

The PCR products were then run on an 1% agarose gel, and colonies giving the correct size of ~700 bp were cultivated in 3mL of LB + AMP media overnight and plasmid DNA was isolated using the QIAprep Spin Miniprep Kit. The DNA was then sequenced with the primer:

pCFD6seqfwd: 5’ – GTAGACATCAAGCATCGGTGG – 3’.
Colonies with the correct sequence were grown in 50 mL of LB + AMP media overnight so that a midiprep could be performed using the QIAGEN plasmid midiprep kit.

Transgenic fly lines were then generated by Rainbow gene injection services. The transgenes used in the study were integrated at the genomic landing site $\gamma[1]w[67c23];P[\gamma[+t7.7]=CaryP]attP1$ on the second chromosome using the PhiC31/attp/attB system. The injected flies were screened using the mini-white marker by crossing single flies to $w^{1118}$ flies (Stock 3605, Bloomington Stock Center), followed by screening for orange eyes. Stable lines were then generated by crossing the flies positive for mini-white to balancer flies with the genotype Cyo/Sco; TM6B/MKRS (Stock 2551, Bloomington Stock Center).

For this study the following genes were targeted: *period (per)*, *timeless (tim)*, short neuropeptide F receptor (*sNPFR*), Myosuppressin receptor (*DMSR*), Allostatin A receptor 1 (*AstA-R1*), Allostatin C receptor 2 (*AstC-R2*), CG13995, and CG13229.

Fly Brain Immunostaining

Flies were entrained for 3-4 days in 12 hours light and 12 hours dark conditions, and were fixed at time-point ZT23. Whole flies were fixed in PBST (0.5% TritonX) with 4% paraformaldehyde for 2 hrs and 45 minutes on a rotating wheel, before washing 5 times for 10 minutes with PBST (0.5% TritonX). The heads were then removed and the brains were dissected in PBST. The brains were blocked in 5% NGS in PBST for 3 hrs at room temperature. The primary antibody (anti-PDF 1:1000 (Developmental Studies Hybridoma Bank), rabbit anti-PER 1:1000 (provided by Ralf Stanewsky [34]), rat anti-TIM 1:200 (provided by Francois Rouyer [35]), 0.02% NaN$_3$, 5% NGS, and PBST) was added and the brains were incubated at 4°C overnight. The brains were then washed 5 times for 10 minutes each time with PBST. The secondary antibody (1:200 dilution of Alexa Fluor 488 conjugated anti-rabbit, Alexa Fluor 635 conjugated anti-mouse, Alexa Fluor 546 conjugated anti-rat, 5% NGS and PBST) was added to the brains
and incubated at room temperature for 3 hours. The brains were then washed 5 times for 10 minutes each time with PBST. The brains were mounted in Vectashield Mounting Medium and scanned in 1.5 μM sections on a Leica SP5 confocal microscope. The scanning parameters were kept constant for experimental lines and the according controls.

**Fly Strains**

All fly stocks were maintained on standard cornmeal food. Crosses were performed at 25°C using virgin females transgenic for X-GAL4 UAS-Cas9 and males containing the transgenic sgRNAs. Different GAL4 lines were used, and the expression of GAL4 for each line was controlled by different tissue specific promoters, thus expressing the sgRNAs in specific cells. The following GAL4 lines were used: Mai-GAL4 [11], Clk856-GAL4 [36], and Tim-GAL4 [37]. All the GAL4 fly lines were crossed to the UAS-Cas9 [31] fly line. For each experimental strain there were 2 control strains. The first is the GAL4 control where flies transgenic for GAL4 UAS-Cas9 are crossed to w¹¹¹⁸ (Stock 3605, Bloomington Stock Center) flies. The second is the guide control, where flies containing the transgenic sgRNAs are crossed to w¹¹¹⁸ flies (Stock 3605, Bloomington Stock Center).

**Behavior Recording and Analysis**

The *Drosophila* Activity Monitoring system (Trikinetics) was used to record behavior by measuring the number of times the fly crossed the infrared beam in 1-minute intervals. Individual male flies of 5-7 days in age were loaded into glass tubes containing agar sucrose media. The flies were first entrained in standard 12 hour light phase and 12 hour dark phase (LD 12:12) conditions for 7 days, before releasing them in constant darkness for another 7 days.
The raw data obtained from the DAM system was then plotted as actograms using the plugin ActogramJ on Fiji. In order to factor in the time for initial entrainment, the raw data before lights-on on the second day was cut-off. To create the average activity profiles, the total activity every hour for each fly was averaged over days 3-7 and plotted. To determine the morning anticipation index, the sum of the total activity 3 hours before lights-on is divided by the sum of the total activity 6 hours before lights-on [38]. To determine the evening anticipation index, the sum of the total activity 3 hours before lights-off is divided by the sum of the total activity 6 hours before lights-off [38]. Sleep is defined as inactivity for 5 minutes. To calculate the average sleep profiles, the sleep amount per every 30 minutes for each fly was averaged over days 3-7 and plotted. The total sleep is a measure of the amount of sleep in 24 hours, and is then divided into day-time and night-time sleep. Groups were compared using a one-way ANOVA followed by a Tukey’s multiple comparison test, and results were deemed non-significant (ns) or significant at $P \leq 0.05$ (*), $P \leq 0.01$ (**), or $P \leq 0.001$ (***).
Results

The tRNA-sgRNA CRISPR system involves inducing multiple DNA double stranded breaks in the gene of interest through the Cas9 endonuclease. The initial cut is then repaired by the non-homologous end joining (NHEJ) pathway, which is prone to generating indel errors and causing a frameshift mutation of the gene [39]. In order to first test the efficiency of the tRNA-sgRNA CRISPR system, I created transgenic flies containing 3 different sgRNAs targeting the genes per and tim. These genes were chosen due to the well-known behavioral phenotypes of both per-null (per0) and tim-null (tim0) mutant flies, and the availability of antibodies that can be used for PER and TIM immunostaining.

Two different approaches were implemented when cloning the sgRNAs. With the first approach, the 3 sgRNAs were targeted within 60 base pairs on either side of the ATG transcriptional start site (Fig. 1). This method was originally chosen as the guides can not only be used to cause mutations in the gene of interest, but they can also be used to induce overexpression through Cas9 activators [40]. This technique involves the fusion of cutting-deficient Cas9 mutants with a transcriptional activator domain [40]. However, due to the obtained behavioral phenotypes and immunostaining results shown later, a second cloning approach was implemented. For the second approach, the 3 sgRNAs were targeted to the protein coding sequence of the gene (Fig. 1). This was chosen as a previous study employing the same tRNA-sgRNA CRISPR system showed disruption of sepia (se) in 99% of eye tissue when four sgRNAs were targeting the coding sequence of the gene [31].
Figure 1. The target sites of guides 1 and guides 2 for *per* and *tim* (a) Target sites of the sgRNAs using cloning method 1 (black arrow), and cloning method 2 (blue arrows) for the *per* gene. (b) Target sites of the sgRNAs using cloning method 1 (black arrow), and cloning method 2 (blue arrows) for the *tim* gene.
After completing the cloning of both sets of *per*-guides (*per*-guides 1 and *per*-guides 2), the transgenic flies containing the desired construct were then crossed to flies expressing GAL4 under the control of a *clk856* driver, thus expressing the sgRNAs and the Cas9 endonuclease in the majority of the clock neurons, excluding a subset of DN1s and a subset of DN3s.

Flies were entrained for 7 days in LD 12:12 conditions, and the average locomotor activity profile for the flies were calculated (Fig. 2a). As expected, both of the parental controls showed the characteristic bimodal activity pattern, consisting of a morning activity peak around lights-on and an evening activity peak around lights-off, which are clearly separated by a siesta during mid-day. Also, as expected, the control groups showed almost no activity during the night. Flies expressing *per*-guides 1 that target the transcriptional start site retained the bimodal activity pattern that is similar to that of the control flies. Flies expressing *per*-guides 1 also showed no significant change in anticipatory behavior for both the morning and evening activity peaks (Fig. 2b), suggesting little or no effect of mutating the *per* transcription start site on behavior in LD 12:12.

However, flies expressing *per*-guides 2 that target the protein coding region show a very different activity profile. They show no true morning and evening activity peaks, and the increase in activity at lights-on and lights-off is merely a masking effect of the flies reacting to light transitions. In addition, the flies show an increase in mid-day activity. The expression of *per*-guides 2 also resulted in a loss of anticipatory behavior to morning and evening lights transitions, as highlighted by the anticipation index values indistinguishable from 0.5 (Fig. 2b). Previous research using the same quantification methods of the anticipation index has shown that wild-type flies have an anticipation index (AI) value of between 0.7-0.8 for both lights-on and lights-off, while flies that do not show any anticipation have lower values closer to 0.5 [38]. Although the activity levels during the 6 hours before lights-on is greater than what is observed in both controls, there is no increase in locomotor activity right before lights-on, or before lights-off. Taken together this data shows high parallels to *per*\(^0\) mutant behavior [41].
When released in constant darkness (DD), the flies expressing per-guides 1 showed a significantly increased free-running period compared to the controls, which maintained the roughly 24 hour period (Fig. 2c). For the flies expressing per-guides 2, there was only one fly that still retained rhythmicity under DD conditions, and had an extremely long period of 28.8 hours. However, due to being only a single fly, this long period value may not be representative. This large reduction in rhythmicity (Fig. 2d) is also observed in per\(^0\) mutants [41], suggesting a high efficiency of per-guides 2. Flies expressing per-guides 1 show a smaller loss of rhythmicity, with 80% of the flies still retaining rhythmic behavior in constant darkness.
Figure 2. Expression of *per*-guides 2 results in decreased anticipatory activity in LD 12:12, and decreased rhythmicity in constant darkness DD (a) The average activity profiles of flies expressing *per*-guides 1 and *per*-guides 2 using a *clk*856 driver in LD 12:12 conditions. (b) Quantification of M and E anticipation. Flies expressing *per*-guides 2 showed a significant reduction in M and E anticipation, while expression of *per*-guides 1 show no significant change. (c) Free-running period of flies released in constant darkness. Flies expressing *per*-guides 1 showed a significantly longer period of the circadian clock when released in DD. (d) Percentage of rhythmic flies in DD. Flies expressing *per*-guides 1 showed a slight decrease in the percentage of flies maintaining rhythmicity in DD, while those expressing *Per*-guides 2 showed a greater decrease. ns= not significant (results compared with controls), P≤0.05 (*), P≤0.01 (**), P≤0.001 (***)
To address the question of different cutting efficiencies for \textit{per}-guides 1 and \textit{per}-guides 2, I performed immunostaining against PER for both experimental groups and the respective controls (Fig. 3). The staining of a subset of the lateral neurons are shown, including the sLNvs, the lLNvs, and the LNds. These neurons are chosen as representative examples, as they are included by the \textit{clk}856-GAL4 driver. The PDF antibody was used in order to differentiate the PDF cells (sLNvs and lLNvs) from the LNds, which in some fly brains can be in very close proximity to one another, and are therefore difficult to separate.

As expected from the behavioral phenotype, the expression of \textit{per}-guides 1 resulted in no observable decrease in intensity of PER staining compared to controls (Fig. 3a), thus suggesting low efficiency of the guides targeting the transcriptional start site of \textit{per}. In contrast, the expression of \textit{per}-guides 2 resulted in a very large decrease in PER staining intensity in comparison to the controls (Fig. 3b), in line with the decreased rhythmicity and loss of anticipatory behavior observed. This shows that the cutting efficiency of \textit{per}-guides 2 is very high. However, approximately 2 neurons per brain may still show PER expression when the guides are expressed using \textit{clk}856-GAL4, as highlighted by the 2 stained LNds (Fig. 3b, left).
Figure 3. Expression of *per*-guides 1 results in no observable decrease in PER staining intensity, while expression of *per*-guides 2 show significant reduction in staining intensity. (a) The immunostaining of fly brains when *per*-guides 1 were expressed. The PDF antibody was used in order to differentiate the sLNvs and iLNvs from the LNds. There is no observable decrease in staining intensity. (b) The immunostaining of fly brains when *per*-guides 2 were expressed, this time showing a large decrease in staining intensity. There are 2 stained LNds and thus still show some PER expression. All flies were fixed at time-point ZT 23.
After looking at the effect of the guides targeting the *per* gene, I then created transgenic flies expressing sgRNAs targeting the *tim* gene. When entrained in LD 12:12 conditions, the flies that are expressing *tim*-guides 1 (targeting the transcriptional start site) using the *clk856* driver showed a characteristic bimodal activity pattern similar to that of the controls (Fig. 4a). There was, however, a tendency to reduce morning anticipation, as well as a significant reduction in evening anticipatory behavior (Fig. 4b), suggesting that there is partial functionality of the guides. Flies expressing *tim*-guides 2 (targeting the coding sequence) reproduce the phenotype described previously for *per*-guides 2; they show a reduction in activity levels during morning and evening light transitions, and display increased activity during mid-day. They also showed a significant reduction in both morning and evening anticipation compared to not only the control flies, but to the flies expressing *tim*-guides 1 as well. This suggests a more completely knockout and thus a higher efficiency of *tim*-guides 2. This lack of anticipatory behavior to transitions of lights-on and lights-off is also observed in *tim*⁰ mutants [42].

When released into DD, expression of both sets of guides resulted in significant changes in behavior. For flies expressing *tim*-guides 1, the period is significantly longer in comparison to both sets of controls (Fig. 4c). Surprisingly, flies expressing *tim*-guides 2 display a period that is shorter compared to those expressing *tim*-guides 1, but still significantly longer compared to both controls. In addition to the increased period, the flies also showed a loss of rhythmicity. When expressing *tim*-guides 1, the flies showed a 40% loss of rhythmicity, while flies expressing *tim*-guides 2 displayed a larger 64% loss of rhythmicity (Fig. 4d).
Figure 4. Expression of sgRNAs targeting *tim* results in decreased anticipatory activity in LD 12:12 and decreased rhythmicity in DD (a) The average activity profiles of flies expressing *tim*-guides 1 and *tim*-guides 2 using a clk856 driver in LD 12:12 conditions. (b) Quantification of M and E anticipation. Flies expressing *tim*-guides 2 showed significant reduction in M and E anticipation, while flies expressing *tim*-guides 1 showed reduced E anticipation. (c) Free-running period of flies released in DD. Both flies expressing *tim*-guides 1 and 2 show significant increases in circadian period when released in DD. (d) Percentage of rhythmic flies in DD. Flies expressing *tim*-guides 1 and 2 show a loss of flies maintaining rhythmicity in constant darkness. ns = not significant (results compared with controls), P≤0.05 (*), P≤0.01 (**), P≤0.001 (***)
To compare the efficiency of both sets of *tim*-guides, I performed immunostaining against TIM when both guides were expressed using a *clk856* driver, as well as for the corresponding controls (Fig. 5). As with the immunostaining against PER for the *per*-guides, the sLNvs, ILNvs, and LNds are the representative cells shown here. Similarly, the PDF antibody was used in order to differentiate the PDF cells (sLNvs and ILNvs) from the LNds.

The expression of *tim*-guides 1 resulted in a moderate decrease in staining intensity when compared to controls. Although all the lateral neuron cells in this area are stained, it is a lot fainter, indicating that TIM expression is reduced. This is in line with the observed behavioral phenotype, where a loss of anticipatory behavior before the evening peak but not the morning peak is observed. This again suggests that *tim*-guides 1 do show partial functionality. The expression of *tim*-guides 2 resulted in an even greater decrease in TIM staining intensity. None of the cells are labelled, indicating that there is no or extremely low amounts of TIM expression in these lateral neurons, and suggests very high cutting efficiency of *tim*-guides 2.
Figure 5. Expression of *tim*-guides 2 results in a greater decrease in staining intensity compared to *tim*-guides 1, but both show observable reduction compared to controls. (a) The immunostaining of the brains when *tim*-guides 1 were expressed. There is an observable reduction in staining intensity, although most cells show low amounts of TIM expression. (b) The immunostaining of fly brains when *tim*-guides 2 were expressed, showing a larger reduction in staining intensity compared to both *tim*-guides 1 and controls. All flies were fixed at time-point ZT 23.
After testing the efficiency of the sgRNAs targeting the core clock components *tim* and *per*, I then created transgenic flies expressing sgRNAs targeting the transcriptional start site of different neuropeptide receptors. Some of the guides targeting the neuropeptide receptors showed no observable change in circadian behavior when expressed with drivers targeting the clock neuron network, including *tim*-GAL4, *clk856*-GAL4, and *mai*-GAL4. One representative neuropeptide receptor is the myosuppressin receptor (DmsR), a G-protein coupled receptor implicated in regulating locomotor activity [43].

When entrained in LD 12:12 conditions, the flies expressing the sgRNAs targeting DmsR under the control of a *clk856*-GAL4 driver showed a characteristic bimodal activity pattern (Fig. 6). As with the control groups, the flies show increased locomotor activity right before lights-on and lights-off, indicating that both morning and evening anticipation are still present. The same activity profile is observed when the DmsR-guides are expressed under *tim*-GAL4 and *mai*-GAL4 drivers. As *tim*-GAL4 drives expression in all of the clock neurons, this suggests that DMS may be acting on cells outside of the clock network to produce the published change in locomotor activity.
Figure 6. The average activity profiles of *D. melanogaster* expressing DmsR-guides using drivers *clk856-GAL4*, *tim-GAL4*, and *mai-GAL4*, in LD 12:12 conditions. Expression of sgRNAs targeting DmsR show no changes in behavioral phenotypes when expressed under the control of *clk856-GAL4*, *tim-GAL4*, or *mai-GAL4* drivers.
One of the neuropeptide receptors that showed major phenotypic changes when targeted was the short neuropeptide F receptor (sNPFR), which has been implicated in sleep promotion [28]. When entrained in LD 12:12 conditions, the flies that are expressing the sNPFR-guides under the control of a *tim*-GAL4 driver showed a characteristic bimodal activity pattern similar to that of both controls groups (Fig. 7a). However, an increased amount of activity during night-time is observed, while activity during mid-day does not seem to be impacted, suggesting that sNPF may have a role in modulating night time activity and sleep.

To test the impact of the sNPFR-guides on sleep in LD 12:12 conditions, a sleep analysis was performed. The state of sleep is defined as inactivity for 5 consecutive minutes. The total sleep in 24 hours was determined, which was then further divided into both day-time and night-time sleep (Fig. 7b). When the sNPFR-guides were expressed using a *tim*-GAL4 driver, a significant reduction in total sleep was observed. Interestingly, day-time sleep was not impacted, while night-time sleep was significantly reduced. This suggests that sNPF may play a role in promoting normal night-time sleep.

Surprisingly, when the sNPFR-guides are expressed using a *clk856*-GAL4 driver, there was no significant reduction in the amount of total sleep (Fig. 7c, left). When the sNPFR-guides are expressed using the *mai*-GAL4 driver, there was a slight reduction in total and night-time sleep, but the decrease was not significant (Fig. 7c, right). Given the broader expression pattern of *tim*-GAL4 (all clock cells, compound eyes and glia cells) compared to both other GAL4 drivers used, this data suggests that sNPFR might act as a sleep promoting factor in glia cells and/or DN1/DN3 clock neurons.
Figure 7. Expression of sgRNAs targeting sNPFR under the control of a tim-GAL4 driver results in reduced night time sleep, while day time sleep is unaffected  (a) The average activity profiles of flies expressing sNPFR using a tim-GAL4 driver under 12:12 LD conditions. A large increase in night-time activity is observed.  (b) Sleep profile and quantification of total, day-time, and night-time sleep. Flies expressing sNPFR-guides using a tim-GAL4 driver show decreased total sleep. This reduction was due to significantly less night-time sleep, while the amount of day-time sleep was not impacted.  (c) Quantification of total, day-time, and night-time sleep. Flies expressing sNPFR-guides using both the clk856-GAL4 and mai-GAL4 driver did not result in any significant reduction of total sleep, nor day-time or night-time sleep.
Discussion

One aim of this study was to implement and test a cell specific CRISPR/Cas9 based system to manipulate the clock neuron network in a cell-specific manner. To do so, the efficiency of sgRNAs targeting the transition start site and sgRNAs targeting the coding sequence of per were compared in this study. I show that flies expressing per-guides 1 in the clock neurons behave like wild type flies. Together with the lack of any observable decrease in PER staining intensity, this shows that per-guides 1 have low functionality in clock neurons. However, when per-guides 1 are expressed under the control of GMR-GAL4, and thus driving expression in the eye, a large decrease in staining intensity of the retina was observed (data not shown). This suggests that per-guides 1 have high functionality in the eye, and therefore per may potentially have alternate transcriptional start sites in clock neurons. This can be tested in the future using RNA sequencing and making comparisons between results from clock neurons and photoreceptor cells in flies expressing per-guides 1 and wild-type flies.

In contrast to per-guides 1, flies expressing per-guides 2 showed a behavioral phenotype that is closely replicated by per<sup>0</sup> flies, lacking bimodal organisation of activity in 12:12 LD, and a major loss of rhythmicity (~80%) under DD conditions [44]. The PER immunostaining also showed a large decrease in staining intensity, indicating that per-guides 2 have high cutting efficiency in clock neurons. However, certain differences are noted between per<sup>0</sup> mutants and flies expressing per-guides 2. per<sup>0</sup> mutants only show very small mean activity peaks provoked by lights-on and lights-off [39], while flies expressing per-guides 2 show more prominent activity peaks as a result of reacting to light transitions. However, this could be a result of different recording conditions, as the compared per<sup>0</sup> mutants were monitored at 20°C. When per<sup>0</sup> flies were monitored at 25°C, they showed a strong activity peak when reacting to lights-on and lights-off transitions in LD 12:12 conditions [35], indicating that temperature has an effect on the general onset of activity in flies [55].
In addition to evaluating the efficiency of the guides against *period*, the efficiency of sgRNAs targeting the transcriptional start site and sgRNAs targeting the coding sequence of *timeless* were compared as well. When *tim*-guides 1 are expressed under the control of *clk856-GAL4*, the flies retain bimodal activity in LD 12:12 conditions, albeit with a lower morning and evening activity peak, as well as significant reduction in evening anticipatory behavior. There is also a moderate 40% loss of rhythmicity when released in DD conditions, showing that *tim*-guides 1 are partially functional. This is also confirmed by the TIM immunostaining, where a decreased staining intensity is observed, but the lateral neurons still show some TIM expression. This shows that reduced amounts of TIM are still sufficient to maintain the bimodal activity pattern in LD 12:12, and a rather high level of rhythmicity in DD.

The observable decrease in immunostaining intensities also indicates that *tim*-guides 1 have higher cutting efficiency in comparison to *per*-guides 1 in clock neurons. This can explain the difference in percentage of rhythmic flies in DD conditions, where a 40% loss of rhythmicity was observed with *tim*-guides 1, and only a 20% loss was observed with *per*-guides 1.

Expression of *tim*-guides 2 closely replicates the behavioral phenotype of both *per*-guides 2 and *tim*⁰ mutants, showing a lack of morning and evening anticipatory behavior in LD 12:12 conditions, as well as a major loss of rhythmicity in DD conditions. There were also no labelled lateral neurons when the TIM immunostaining was performed, thus showing that by simultaneously targeting several sites within the protein coding sequence, this can largely circumvent problems of functional in-frame mutations and sgRNAs that have low levels of activity. The high efficiency of this technique provides a new option to test the M and E oscillator model, where distinct groups of circadian neurons modulate the timing of morning and evening activity [11]. Due to the sgRNAs targeting the coding sequence having a much stronger effect than sgRNAs targeting the transcription start site, I went on to generate a library for guides targeting the coding sequence of neuropeptide receptors, and 70 clonings have been completed so far.
Given the positive results from manipulating PER and TIM, I also tested sgRNAs that are targeting neuropeptide receptors. Some have shown no phenotypic changes when expressed using the drivers tested so far, including clk856-GAL4, tim-GAL4, and mai-GAL4. One example of this is the receptor for myosuppressin (DmsR), a neuropeptide that has been shown to play a role in locomotor activity [43]. Previous research has shown that DmsR-RNAi knockdown using an elav-GAL4 diver resulted in reduced locomotor activity, in particular the mean velocity of movement [43]. However, when the DmsR-guides are expressed under the control of the aforementioned drivers, the flies displayed average activity curves indistinguishable to wild-type flies, with no significant changes in locomotor activity. This discrepancy in results could be due to the previous study’s use of the pan-neuronal elav-GAL4 driver, which has a broader expression pattern compared to the drivers that I have tested so far. In the future, I could first try to replicate the published results using elav-GAL4, before using more specific drivers that drive expression in neurons outside of the clock network to investigate how DMS promotes normal locomotor activity levels.

Interestingly, the same study found that the ubiquitous knockdown of DMSR resulted in a lethal phenotype, while this is not observed when using elav-GAL4, which induces expression in all neurons [43]. This suggests that the cause of lethality is outside of the CNS. FlyATLAS has shown that there is a high expression of DmsR in the crop and Malpighian tubules [43], which play roles in storing liquid food and excretion respectively [46,47]. Therefore, using drivers that express specifically in these cells, such as DJ717-GAL4 [48] and Tdc1-GAL4 [49] respectively, I could potentially investigate the role of DMS outside of the CNS and why ubiquitous knockdown of DMS results in a lethal phenotype.

In contrast to the sgRNAs targeting DmsR, expression of sNPFR-guides using tim-GAL4 showed major phenotypic changes, displaying a significant decrease in total sleep. This decrease was due to a significant reduction in night-time sleep, while day-time sleep was unaffected. This is also observed in previous research where RNAi knockdown of sNPF using pdf-GAL4 resulted in small but significant
decreases in night-time sleep, without affecting daytime sleep [28]. Surprisingly, sNPFR-guides expressed using clk856-GAL4 and mai-GAL4 showed no significant reduction in total or night-time sleep. This could indicate that while sNPF within sLNvs and lLNvs promotes normal night-time sleep, it may be acting on its corresponding receptor in the DN1s or DN3s, or potentially the glia cells that are included in tim-GAL4. To test this, I could use drivers that induce expression in the DN1s and DN3s but not the other clock neurons, such as PDFREY11851-GAL4 [50]. Outside the clock network, a previous study has shown that sNPFR expression is also enriched in the mushroom body and fan-shaped body [51], both of which are implicated in sleep and arousal [52]. Thus, I can use drivers such as C765-GAL4 [53] to understand if sNPF communicates and promotes normal night-time sleep via these cells.

In this study, I have shown that the tRNA-sgRNA CRISPR system does have high efficiency, especially when targeting the protein coding sequence. The ability to therefore create cell specific knockouts has many potential applications. A novel split-cas9 system can be used to investigate adult-specific effects of knocking out specific neuropeptide receptors. This system involves the splitting of Cas9 into two separate fragments that are chemically inducible by rapamycin for controlled reassembly to mediate gene editing [54]. Many neuropeptides are involved in development and therefore this technique can circumvent many developmental issues, and allows one to study functions of receptors in an adult specific manner. In addition to testing novel methods, I also aim to create a complete library of CRISPR-mediated knock-outs for all 115 neuropeptide receptor genes in the Drosophila. This can be used to reveal more about the communication within neuronal clusters in the circadian network, and how they coordinate many physiological responses, such as appetite and stress, in Drosophila and other organisms.
Acknowledgements

I want to express my sincere gratitude to Dr. Michael Rosbash for giving me the opportunity to work in the Rosbash lab, and participate in a project I truly enjoy. I am very grateful for the advice regarding both lab-work and outside of it, and I look forward to continue working in the Rosbash lab after graduation. I am also extremely grateful to Matthias Schlichting for his mentorship during my time here. He has truly taught me a lot regarding experimental procedures, writing, preparing for talks, to name a few. This thesis would not be possible without his guidance and feedback. I also want to thank all the members of the Rosbash Lab, who have created a great environment for undergraduates to do research. I want to thank Madelen Diaz especially for her help in creating the figures, as well as for her writing advice. Finally, I want to thank my committee members Dr. Michael Marr and Dr. Amy Lee, and of course Dr. Michael Rosbash, for taking time to evaluate my work.
References


Supplementary Table 1.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer Number</th>
<th>Primer Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>per</em> (guides 1)</td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>per</em> (guides 1)</td>
<td>Primer 2</td>
<td>GAATTGTCCATTGCTTGGACCCACCCCGGAATCGAACC</td>
</tr>
<tr>
<td><em>per</em> (guides 1)</td>
<td>Primer 3</td>
<td>ACAGCGACTCGATCATGGCTTGTCGACAGCTATGCTGGAAC</td>
</tr>
<tr>
<td><em>per</em> (guides 1)</td>
<td>Primer 4</td>
<td>TTCATTTGCATTCTACGTAAAAACCCGCAAGTCCAGCAGGGAATCGAACC</td>
</tr>
<tr>
<td><em>per</em> (guides 2)</td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>per</em> (guides 2)</td>
<td>Primer 2</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>per</em> (guides 2)</td>
<td>Primer 3</td>
<td>GAATTGTCCATTGCTTGGACCCACCCCGGAATCGAACC</td>
</tr>
<tr>
<td><em>per</em> (guides 2)</td>
<td>Primer 4</td>
<td>GACCAAACAGTCCAGCAGGGAATCGAACC</td>
</tr>
<tr>
<td><em>tim</em> (guides 1)</td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>tim</em> (guides 1)</td>
<td>Primer 2</td>
<td>GAATTGTCCATTGCTTGGACCCACCCCGGAATCGAACC</td>
</tr>
<tr>
<td><em>tim</em> (guides 1)</td>
<td>Primer 3</td>
<td>GAATTGTCCATTGCTTGGACCCACCCCGGAATCGAACC</td>
</tr>
<tr>
<td><em>tim</em> (guides 1)</td>
<td>Primer 4</td>
<td>GACCAAACAGTCCAGCAGGGAATCGAACC</td>
</tr>
<tr>
<td><em>tim</em> (guides 2)</td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>tim</em> (guides 2)</td>
<td>Primer 2</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>tim</em> (guides 2)</td>
<td>Primer 3</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>tim</em> (guides 2)</td>
<td>Primer 4</td>
<td>GAATTGTCCATTGCTTGGACCCACCCCGGAATCGAACC</td>
</tr>
<tr>
<td><em>sNPFR</em></td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>sNPFR</em></td>
<td>Primer 2</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>sNPFR</em></td>
<td>Primer 3</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>sNPFR</em></td>
<td>Primer 4</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>DmsR</em></td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>DmsR</em></td>
<td>Primer 2</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>DmsR</em></td>
<td>Primer 3</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>DmsR</em></td>
<td>Primer 4</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>AstA</em></td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>AstA</em></td>
<td>Primer 2</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>AstC</em></td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>AstC</em></td>
<td>Primer 2</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>AstC</em></td>
<td>Primer 3</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>AstC</em></td>
<td>Primer 4</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>CG13995</em></td>
<td>Primer 1</td>
<td>GGCCCGGGTTCGATTCCCGGCCGATGCA</td>
</tr>
<tr>
<td><em>CG13995</em></td>
<td>Primer 2</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>CG13995</em></td>
<td>Primer 3</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td><em>CG13995</em></td>
<td>Primer 4</td>
<td>CAAGTTCGCCCATCTTGCAGGCAGGGCAATCGAACC</td>
</tr>
<tr>
<td>CG13229</td>
<td>Primer 3</td>
<td>TCTACAATGATGCAAGAGACGTTCAGCTATGCTGGAAAC</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>CG13229</td>
<td>Primer 4</td>
<td>ATTTTAACCTTCTATTTTCTAGCTCTAAAAACTGTTGAAAGGAAACCCGCTGCACCAGCGAATCGAACC</td>
</tr>
</tbody>
</table>

Table 1: List of oligonucleotides used to clone 3 gRNAs from 2 overlapping PCR products for *per* and *tim*, and each neuropeptide receptor.