Functional analysis of the *period* and *timeless* circadian molecular mechanism within the diptera.

Thesis submitted for degree of Doctor of Philosophy at The University of Leicester

By

Seth Racey Bsc (St Andrews), Department of Genetics, The University of Leicester

October 2001

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Functional analysis of the *period* and *timeless* circadian molecular mechanism within the diptera.

Seth Racey

Abstract

In *D. melanogaster* the *period* (*per)* and *timeless* (*tim*) genes form a feedback loop that rhythmically inhibits their own transcription. This mechanism creates a twenty four hour clock that regulates many of the flies behavioural and molecular circadian cycles. The conservation of at least *per* as part of this molecular mechanism in mammals suggests that clock gene functions are well conserved across the animal kingdom. However analysis of *per* and *tim* functions in *Antherea pernyi* and *Musca domestica* illustrated that the function of the genes did not comply with the *D. melanogaster* feedback model, and that the genes functions may have diverged significantly within the insects.

To further test this divergence of function , analysis of the *per* and *tim* function was performed on *D. virlis* and *D. pseudoobscura* and extended in *Musca*. The functional analysis confirmed that *Musca* PER does not obey the *D. melanogaster* circadian clock model. In particular it does not cycle in abundance or subcellular localisation. Furthermore the timing cycles of *per* and *tim* transcripts are significantly different between *D. virilis* and *D .pseudoobscura* and *D. melanogaster*. Furthermore unusual PER and TIM cycles in *D. virilis* suggest that it may have a quite different method of generating a twenty four hour cycle. All of which suggests that the genetic mechanisms behind the molecular clocks of the insects have undergone widespread changes over evolutionary history, and may mean that they are a target for change during speciation events.

Abbreviations:

A	Adenine
AP	Alkaline Phosphatase
ATP	Adenosine Triphosphate
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
A	Alanine
CD	Circular Dichroism
CPM	counts per minute
Cys	Cysteine
C	Cytosine
C	Cysteine
DAB	diaminobenzidine
DMF	Dimethylformamide
DMSO	dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
Da	Dalton
D	Aspartic Acid
EDTA	ethylenediaminetetraacetic acid
E	Glutamic Acid
E.coli	Escherichia coli
EMBL	European Molecular Biology Laboratory
F	Phenylalanine
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
G	Glycine
GST	Glutathione-S-Transferase
G	Guanine
HLH	Helix-Loop-Helix
HA	Hemagglutinin
Hepes	N-2-hydroxyethylpiperazine-N'-2-ethane sulfonic acid
His	Histidine
H	Histidine
Ile	Isoleucine
I	Isoleucine
IPTG	Isopropyl-beta-D-thiogalactoside
kDa	Kilodalton
K	Lysine
LSC	Liquid Scintillation Counter
Leu	Leucine
Lys	Lysine
L	Leucine
MALDI	matrix-assisted laser desorption ionization
Met	Methionine
M	Methionine
mRNA	messenger-Ribonucleic Acid
NMR	Nuclear Magnetic Resonance
OD	Optical Density
PAGE	polyacrylamide gel electrophoresis
Phe	Phenylalanine
Pro	Proline

P	Proline
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
Q	Glutamine
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
R	Arginine
SDS	sodium dodecyl sulfate
Ser	Serine
S	Serine
TEMED	N,N,N',N-Tetramethylethylenediamine
Thr	Threonine
Tm	Tropomyosin
Trp	Tryptophan
Tyr	Tyrosine
T	Thymine
T	Threonine
tRNA	transfer-Ribonucleic Acid
Val	Valine
V	Valine
W	Tryptophan
Y	Tyrosine
U	Uracil
WT	Wild Type
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactoside

Abstract

Introduction

Chapter 2:
Materials and Methods

Chapter 3:
PER expression in *Musca domestica*

Chapter 4:
Raising anti-sera against *Musca* PER

Chapter 5
Purification of protein fragments for antibody production and structural studies.

Introduction

It should come as little surprise that organisms have felt the need to time the passing of a day. The change from dark to dawn means many physical variables change; sunlight, warmth and wind can mean food for predators, predation for the hunted and energy for carbon fixers. In this context the ability to keep time and predict these change is of obvious value. Recently the molecular components of the 24 hour or circadian clock (*circa*-around; *dien*- day) have been described in bacteria, fungi, diptera and mammals. Of the organisms studied two models, *Drosophila melanogaster* and *Neurospora crassa,* have provided many of the key insights into circadian clocks including the unifying concept of negative feedback. To understand how negative feedback works each model is best described separately and in detail.

***Drosophila melanogaster* circadian clock.**

The detailed analysis of the circadian molecular clock started with the identification of the *period* (*per*) gene by forward mutagenesis (Konopka and Benzer, 1971). The *per* gene was then cloned (Bargiello and Young, 1984) and three mutants were identified. The per^0 gave arrhythmic behavioural rhythms. Whereas per^S and per^L, slow and fast mutants respectively, gave dramatic period shifts in free-running conditions; indicating that the gene was essential for circadian clock function. Analysis of mRNA cycles in these mutants suggested that *per*-encoded protein fed back on the levels of its own transcript (Hardin et al., 1990). *per* mRNA cycled in abundance in a daily manner, with a 5- 10 fold amplitude. In fly heads the mRNA cycle appeared shifted in the mutants, per^S peaked earlier and per^L later in the day, and the cycle persisted in experiments where the flies were moved from

their normal entrainment of LD into DD. Under DD conditions wild-type flies showed a ~24h mRNA cycle, the per^S mutant gave a faster cycle of 20h. As the per^S is a single amino acid substitution this meant that the protein must feed back on its own mRNA cycle (Baylies et al., 1992; Yu et al., 1987).

The second component, the *timeless (tim)* gene was identified by forward mutagenesis (Myers et al., 1995) and by using a biochemical approach (Gekakis et al., 1995). Molecular analysis showed that the tim^0 mutant obliterates the circadian *per* mRNA cycle, *tim* mRNA also cycled in a manner indistinguishable from *per* mRNA with the *tim* mutant also blocking the rhythmic expression of its own transcript (Price et al., 1995; Sehgal et al., 1995). At the same time yeast-two-hybrid analysis showed that PER and TIM fragments dimerized and that the per^L mutation weakened the binding of the PER fragment to TIM (Gekakis et al., 1995). This was a partnership between two genes *per* and *tim* that share the following features, both have rhythmic mRNA expression, both feed back on their own mRNA transcripts, both physically associate and mutations in both disrupt clock phenotypes. All of these facts suggested that *per* and *tim* were two central genes in the circadian mechanism. Subsequent detailed analysis has revealed how this feedback occurs and also how a series of biochemical events have created a negative feedback system that has a 24h rhythm.

PER protein localisation was first analysed by immunhistochemistry and with *per-lacZ* fusions (Liu et al., 1988; Siwicki et al., 1988). The PER protein was found in many areas including the optic lobes, gut, malphigian tubules and central nervous system (CNS) (Liu et al., 1988; Siwicki et al., 1988). Furthermore the PER protein cycled in the CNS, peaking late at night and falling throughout the day (Siwicki et al., 1988; Zerr et al., 1990). This cycle persisted even in DD suggesting

that the protein cycle was not simply responding to light cues. Further analysis showed that a small subset of PER containing cells were important to the generation of rhythmic behaviour (Ewer et al., 1992) supporting extensive earlier work in which it was demonstrated that the central nervous system of the fly was an important centre for controlling behavioural rhythms (Handler and Konopka, 1976; Konopka et al., 1983).

Three further studies dissected these key brain areas responsible for rhythmic behaviour. The first over-expressed PER protein to accumulate to a constantly high level only in the photoreceptors, a tissue that normally expresses PER. This however had no effect on the rhythmic behaviour of these flies; therefore the master regulator does not reside in the photoreceptors (Zeng et al., 1994). The second study drove PER expression off a *glass* promoter in per^0 flies permitting expression only in the photoreceptors and a small subset of cells in the ocelli and CNS (Vosshall and Young, 1995) This resulted in rescue of rhythmic behaviour in the normally arrhythmic mutant flies. Additionally, when the photoreceptors were genetically removed, rhythmic behaviour persisted (Vosshall and Young, 1995). Both studies suggested that the CNS was the site for master control of rhythmic behaviour. The third study generated wild type/ per^0 mosaics and rhythmic behaviour was mapped to the CNS in a subgroup of glia (Ewer et al., 1992). These studies demonstrated conclusively that the observed staining in the optic lobes and photoreceptors did not drive rhythmic behaviour in *D. melanogaster* . This contrasted to previous work that had shown that the optic lobes governed behavioural rhythmicity in cockroaches (Jackson et al., 2001). Ewer *et al.* concluded that although the PER expression in glial provided weakly rhythmic behaviour it was probable that the nearby set of neural cells, the lateral neurones

(LNs), could be the site responsible for robust rhythmic behaviour. In addition a subset of LNs were shown to colocalise pigment dispersing hormone (PDH) and antibodies to PDH were able to reveal that these LNs had extensive neural projections (Helfrichforster, 1995). These cells therefore provided candidates for a master clock as they could communicate with the rest of the brain, something that the smaller glial cells, whose role is thought to be in supporting neuronal cells, could not. The colocalisation of PDH also provided a convenient way of unambiguously identifying a subset of the LNs (HelfrichForster, 1996). In addition, cell death genes *rpr* and *hid* were used to selectively ablate LNs that colocalised PDH, a *pdh-gal*4 transgene was used to drive *UAS-rpr* or *UAS-hid* expression, restricting their expression to neurones that would normally express PDH. The transgenic animals were poorly rhythmic (Renn et al., 1999) conclusively demonstrating the LNs importance to behavioural rhythm generation.

More detailed ICC revealed that nuclear translocation of PER was gated (Curtin et al., 1995; Ewer et al., 1992; Liu et al., 1992) with delayed entry in per^{L} mutants (Curtin et al., 1995). The nuclear localisation of PER supported the idea that it could directly feed back on its own transcription, and this was confirmed by PER over-expression studies which revealed that the endogenous *per* mRNA levels were suppressed to low levels (Zeng et al., 1994). TIM was also shown to enter the nucleus in a gated manner and this could be blocked in per^{0} mutants (Sehgal et al., 1995) as could, conversely, the entry of PER in tim^{0} mutants (Vosshall et al., 1994). Therefore TIM required PER to be present for nuclear entry and *vice-verse* (Myers et al., 1996). This gated nuclear entry provided a mechanism by which the transcription of *per* and *tim* mRNA could occur sometime after the rise of PER and TIM protein. Therefore a necessary lag between protein accumulation and inhibition

of transcription was built into the cycling of PER and TIM. Mathematical modelling of a transcriptional/translational negative feedback loop suggests that without such a lag there is a dampening of the cycle (Friesen et al., 1993).

PER TIM complexes

Yeast-two-hybrid analysis has shown that PER and TIM can dimerise (Gekakis et al., 1995) Although the protein can interact in yeast this does not offer proof that this interaction is biologically important in *Drosophila*. However immunoprecipitation and protein size fractionation experiments confirmed that the PER-TIM complex is formed in fly heads (Zeng et al., 1996). The apparent molecular weight of PER was determined from its migration during gel filtration or in sucrose gradients which had been calibrated for protein size migration. Head extracts from night or day time points were fractionated by sucrose gradient or gel filtration and then subject to western analysis. In night time head extracts PER migrated in a manner which suggested its molecular weight was considerably greater than its amino acid sequence predicted, suggesting PER was in a complex. However, during the day PER migrated much closer to its predicted size. Therefore PER was in a complex during the night but was monomeric during daylight. In tim^0 flies the night time complex was removed. When two differently tagged PER constructs (*per-ha* and *per-myc*) were transformed into flies, HA tagged PER pulled down TIM during immunoprecipitation (IP) of the HA tag. However the second tagged PER (PER-MYC) was not immunoprecipitated (IP), suggesting that PER-PER complexes were not a significant fraction of the complex as had been previously predicted by (Huang et al., 1995)

TIM is responsible for light entrainment.

Not only has TIM been found to be involved in complex formation but its role in mediating light entrainment of the clock has been demonstrated by multiple authors (Lee et al., 1996; Myers et al., 1996; Zeng et al., 1996). Myers *et al* demonstrated that TIM is rapidly degraded by light, through entraining per^0 flies to DD, a regime that leaves TIM protein at constantly high levels, and then pulsing them with 15 mins of light, TIM levels were immediately reduced, demonstrating that the TIM response to light was independent of PER and was also extremely rapid (Myers et al., 1996). TIM in wild type flies was also destabilised on light stimulation (Myers et al., 1996). This, combined with the observation that PER is still present at ZT4 (Edery et al., 1994), indicates that TIM is the partner that is more sensitive to light. Further studies have shown that this light induced reduction in TIM reduces amounts of the PER-TIM complex not just free TIM (Lee et al., 1996; Zeng et al., 1996). Although TIM is monomeric at the start of its accumulation it soon forms complexes with PER. However at the start of it decline, between ZT 22 and ZT24, the amount of TIM involved in complexes was reduced by 88% (Zeng et al., 1996). In contrast amounts of PER decline much more slowly (between ZT 22 and ZT6), with PER occurring as a monomer throughout its decline over the day. Nevertheless PERs accumulation is dependent upon TIM as PER is unstable in a tim^0 background but TIM is stable in a per^0 background, suggesting that PER is stabilised by TIM. This indicates that when TIM levels fall at the end of the night (ZT 22-24), PER is no longer shielded and is degraded.

Phase Response Curves (PRC)

During the accumulation phase, from ZT 12 to ZT 16, PER and TIM levels rise but PER and TIM are present as monomers blocking transit into the nucleus. A pulse of light during this phase reduces TIM levels thus preventing the formation of the PER-TIM dimer, but a pool of *tim* mRNA allows reaccaccumulation of TIM (So and Rosbash, 1997). After a delay enough TIM accumulates to form the PER-TIM dimer which then moves into the nucleus and inhibits *per* and *tim* transcription, resulting in a phase delay as the onset of repression of transcription is delayed (Zeng *et al.*, 1996). This phase delay in molecular rhythms was maintained because the flies were kept in DD conditions with no light input to reset the clock. Once nuclear entry had occurred around ZT 17-18 the effect of light pulses was opposite, the pools of *per* and *tim* mRNA were much reduced (So and Rosbash, 1997), and re-synthesis of TIM could not occur. Without TIM to protect it PER degraded more quickly and repression was released sooner, thus advancing the next phase of accumulation (Zeng *et al.*, 1996).

Positive Factors

With the mounting body of evidence that negative elements of the feedback loop had been found the search for the positive factors began. The first breakthrough came with the positional cloning of a Clock (Clk); a gene initially identified by mutagenesis, in which a circadian behavioural variant had been produced. *mClk* encodes a βHLH DNA binding protein with a PAS domain (King et al., 1997). *dClock (dClk)* was then cloned (it was initially called *jerk* (*jrk*) because of an arrhythmic mutant associated with this locus) and was homologous to *mClk*. The *jerk* mutation removed most of the *dClk* activation domain, and gave low levels

of PER and TIM (Allada et al., 1998). A second transcription factor essential for clock function in *Drosophila* was then cloned (Rutila et al., 1998). The *cycle* (*cyc*) gene was another bHLH-PAS protein, with homology to human *BMAL1* or *MOP3*. Mutants of *cyc* were arrhythmic. Rutila *et al* (1998) proposed that dCYC and dCLK dimerized and bonded to the E-boxes (a CACGTG tetramer) of the *per* and *tim* promoters (Lee et al., 1999) where they activate transcription from the *per* and *tim* promoters (Darlington et al., 1998; Lee et al., 1999). In addition the PER-TIM complex was shown to interact with the dCYC-dCLK complex during the night (Lee et al., 1998) but did not disrupt it (Bae et al., 2000; Lee et al., 1999), leading to the possibility of PER-TIM-CYC-CLOCK tetramers (Bae et al., 2000). Furthermore *dClk* transcription appears to be upregulated by PER and TIM because dCLK is held at low levels in per^0 or tim^0 mutants. This interconnected feedback loop may promote robustness of the circadian cycle (Lee et al., 1998), but how PER and TIM act as positive factors in this cycle remains unknown.

Cryptochromes are the light sensors in the *Drosophila* clock.

A mutation in the *Drosophila* cryptochrome gene (cyr^b) eliminated the *per-luc* reporter cycle (Stanewsky et al., 1998). The mutant cry^b may be a null form of the *cryptochrome* gene and has poor synchronization to light cycles (Stanewsky et al., 1998). Furthermore PER-TIM levels did not cycle in cry^b mutants and TIM, in particular, appeared to be resistant to light induced degradation in the cry^b mutants (Stanewsky et al., 1998). Therefore in wildtype flies TIM degradation was probably a consequence of a CRY mediated signal (Stanewsky et al., 1998) Additionally over expression of CRY resulted in an increased photosensitivity of the circadian clock (Emery et al., 1998). The *cry* gene family had been characterised in plants in which

all its members were blue light receptors (Stanewsky et al., 1998). Transfection of a *tim*-luc reporter in S2 cells in conjunction with *per, tim, Clk*, *cyc* and *cry* illustrated that the PER-TIM inhibition of the reporter was removed only when CRY was expressed in conjunction with light stimulation (Stanewsky et al., 1998). In the S2 cell line used, however, TIM was not degraded by light and the removal of PER-TIM inhibition was uncoupled from TIM degradation. This lead to the conclusion that TIM degradation, as occurs in adult brains, was probably a downstream consequence of the CRY interaction with PER –TIM, but was not necessary for removal of PER-TIM inhibition (Ceriani et al., 1999).

*Cry*s effect on PER-TIM complex and its subsequent effect on the circadian clock is dependent on its expression in the LNs. CRY expression was driven in *Drosophila* LNs using *tim-GAL4*, *UAS-cry* line (Emery *et al.* 2000). The *tim-GAL4* driven CRY expression resulted in rescue of cry^b behavioural arrhythmicity. However *CRY* expression in the photoreceptors, driven off a *rodopsin1* promotor, did not rescue cry^b arrhythmicity (Emery et al., 2000).

What signals TIMs degradation is not clear, although CRY did not signal degradation in S2 cell lines (Stanewsky et al., 1998) a CRY mediated TIM degradation signal cannot be ruled out because the experimental analysis involved expression constructs in cultured cells, and may not reflect the true protein dynamics in fly heads. However, once TIM has been signalled for degradation by ubiquitination it is quickly degraded by the proteasome, presumably through TIMs seven PEST sequences. (Naidoo et al., 1999).

Other factors involved in the *Drosophila* clock.

The positive and negative elements of the *Drosophila* circadian clock were therefore in place. However a series of new mutants with clock phenotypes added an important new dimension to the feedback loop.

The *Double-time (dbt)* mutants have short (dbt^S) and long (dbt^L) circadian behavioural phenotypes. The dbt^P allele, was pupal lethal, however, in the mutant larvae PER was hypophosphorylated and accumulates without the presence of TIM (Price et al., 1998). Therefore *dbt* appears to be implicated in the degradation of PER. The *dbt* gene encodes a protein with homology to human casein kinase1ε and DBT binds PER, phosphorylating it and causing it to become unstable (Price et al., 1998). Another two *dbt* mutants (dbt^g dbt^H) removed the lag between *per* mRNA and protein cycles causing a lengthening in the behavioural rhythms (Suri et al., 2000). Therefore *dbt*'s role in the circadian clock appears to be the generation of the lag between *per* mRNA and protein accumulation.

The *Vrille* (*vri*) gene encodes a transcription factor and was known as a maternal enhancer of dorsoventral patterning (George and Terracol, 1997) but was identified as a cycling transcript by Blau and Young (1999). The *vri* mRNA cycling is directly regulated by *d*CLK and CYC and removal of the *vri* cycle leads to repression of *per* and *tim* cycles and associated long or arrhythmic phenotypes (Blau and Young, 1999). Furthermore VRI and CLK independently regulate PDF which is involved in overt rhythmic behaviour (Blau and Young, 1999). Placing *vri* as an input, output or central part of the oscillator is difficult and shows the complexity inherent in the circadian clock.

The *Lark* mutants were identified in a screen for altered eclosion timing mutants that did not disrupt activity rhythms (Newby and Jackson, 1993). LARK

has RNA binding characteristics (Newby and Jackson, 1996) and its mRNA cycles are disrupted by per^0 and tim^0 (McNeil et al., 1998). Its role in controlling temporal adult eclosion rhythmes puts it as an output of the clock.

dCREB2 (cAMP response element binding proteins) mutations also affect the clock, causing shortened periods and dampening of *per* mRNA cycles. The per^0 mutant removes circadian cycles in CREB protein accumulation. Given that the two genes have affects on each others cycling, this suggests that they are involved in the same regulatory loop and CREB appears to be another input to the clock (Belvin et al., 1999).

Finally, the *pigment dispersing factor* (*PDF*) gene is implicated in circadian clocks. Mutations in this gene cause severe abnormalities in circadian rhythms and implicate PDF as a principal output transmitter in the behavioural pathway (Renn et al., 1999).

A Drosophila model

With the identification of a large number of the molecular components involved in circadian clocks a general model of clock function can be made (figure 1.1)

The cycle begins with CLOCK and CYCLE binding together on the *per/tim* e-box promoter and initiate transcription of *per/tim* mRNA. PER and TIM then slowly accumulate over time. When enough PER and TIM has been made they dimerise and enter the nucleus and bind to the CLOCK-CYCLE complex and inactive it, repressing *per/tim* transcription in the process. Therefore the levels of

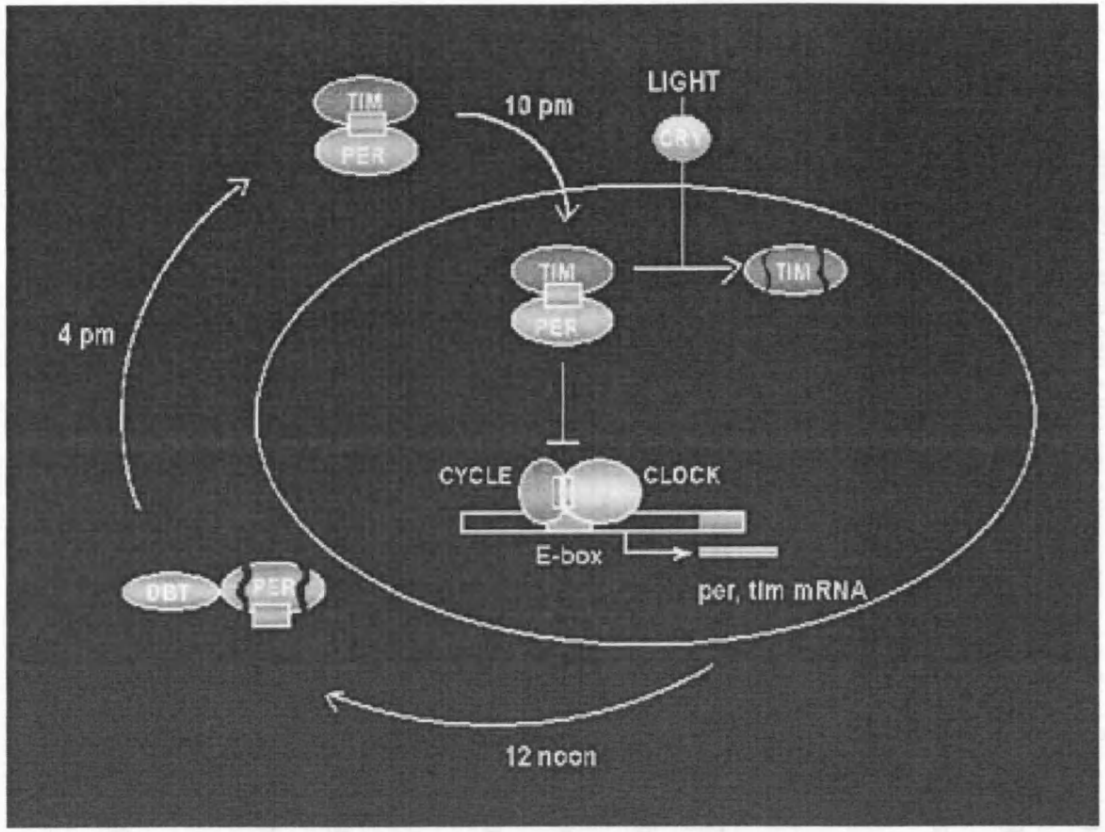


Figure 1.1 General model of circadian molecular mechanism

their mRNA falls and subsequently the protein levels fall until the CLOCK-CYCLE complex is nolonger inactivated and a new cyle of the clock begins.

Two other proteins play critical roles in this cycle; DBT and CRY. DBT ensures that the cycle takes 24 hours to complete by regulating the accumulation of the PER protein through phosporylation modification. CRY on the other hand ensures that the components cycle matches the outside world by rapidly degrading TIM in the presence of light.

Fly behavioural rhythms

The circadian clock mechanism described above then controls a great many aspects of the flies rhythmic behaviour. Examples of such behaviour are locomotor activity and egg eclosion. Drosophila flies have a distinct circadian profile of locomotor activity. They are typically inactive at night then wake up during the morning and have a bout of prolonged activity, at mid day they then siesta, with decreased activity for a few hours and then have another extended bout of activity until dusk (Piccin et al, 2000). Egg eclosion also has a distinct circadian pattern with the majority of flies eclosing in the hours before dawn, Flies that miss this window of opportunity then wait 24 hours until the next window of opportunity (Sauman et al , 1998)

Neurospora crassa

Forward genetics identified the *frequency* (*frq*) gene locus by isolating short, long and null mutants. The *frq* transcript oscillated with a period of 23h in DD, the long mutant *frq7* lengthened the cycle and the null mutant *frq9* abolished it (Aronson et al., 1994). Aronson *et al* hypothesised that FRQ was a negative

repressor of its own transcript, like PER and TIM. When a *frq* constitutive over-expression construct was induced in a rhythmic line it became arrhythmic. Furthermore this construct was not able to rescue null mutants. However a inducible, weak expression construct was able to reset rhythmic lines. FRQ was, therefore, a negative repressor of its own transcript. Given the mutant phenotypes, *frq* was described as a "central component in a circadian oscillator" (Aronson et al., 1994).

Analysis of two mutant loci *white collar-1* (*wc-1*) and *white collar*-2 (*wc-2*) showed that they encoded zinc finger transcripition factors (Ballario et al., 1996) that associated through PAS domains (Ballario et al., 1998). Both are necessary for *frq* transcription, as *wc* null mutants have low levels of *frq* mRNA and protein. The *frq* transcript oscillation cannot be driven in the absence of *wc-2,* and *wc-1* is responsible for the light mediated enhancing of *frq* transcription (Crosthwaite et al., 1997). Therefore a fungal system had been described that has striking similarity to the core oscillator of *Drosophila*. The negative elements are *per* and *tim* or *frq* whose mRNAs accumulate rhythmically. They then enter the nucleus after a lag, and sequester positive regulators for their own transcription. The positive factors consist of two protein partners; *clock* and *cycle* or *wc-1* and *wc-2*, that associate through PAS-PAS interactions. The negative regulators PER,TIM or FRQ decay and negative regulation is removed leading to a new cycle.

Furthermore there is lengthy lag in translation of *frq* (Garceau et al., 1997) similar to that found for *per* and *tim* (So and Rosbash, 1997). Nuclear localisation is required in both systems (Curtin, 1995 et al; Liu et al., 1992; Luo et al., 1998) and is mediated through an NLS. However unlike *Drosophila,* in the fungus there is no delayed entry into the nucleus and this does not contribute to the temporal lag in

negative repression (Luo et al., 1998). Alternative translational starts are also used; TIM is polymorphic for two such isoforms (Rosato et al., 1997) whereas FRQ has two forms (Garceau et al., 1997) which are regulated by temperature and allow rhythmicity over a wider range of temperatures (Liu et al., 1997). In addition phosphorylation is also used as a signal for degradation of FRQ much like PER. Removal of phosphorylation leads to a reduction in degradation and a longer period (Liu et al., 2000). Furthermore FRQ positively regulates levels of WC-1 providing an interconnected circadian loop which is proposed to enhance the robustness of the clock (Lee et al., 2000). This mechanism of interconnected loops for input, output and central components is used in *Drosophila*, *Neurospora* and mammals (Glossop et al., 1999; Lee et al., 2000; Shearman et al., 2000b).

Despite its obvious similarities to *Drosophila* the roles of *frq*, *wc-1* and *wc-2* as central components in a circadian oscillator has now been questioned. The first evidence that doubted the position of these genes at the centre of the clock relied on elaborate manipulations of entrainment protocols. Merrow *et al* (1999) showed that in constant light, circadian rhythms could be driven by temperature changes. However constant light should keep *frq* mRNA levels high, a condition that should stop the clock if *frq* is the central regulator. From this they conclude that *frq* is a circadianly regulated part of the light input pathway (Merrow et al., 1999). The *chol-1* gene, whose mutants affect the clock, can be combined with *wc-1*, *wc-2* and *frq* null mutants and still give rhythmicity in circadian growth cycles, providing unequivocal evidence that *frequency*, *whitecollar-1* and *whitecollar-2* were not the central circadian regulators. However, the strong clock phenotypes of the mutant *frq*, *wc-1*, *wc-2* transcriptional loop illustrates that many clock functions can be

controlled by what now appears to be the light and temperature input into the clock which has its own rhythmicity.

As a note of caution a recent review of fungal circadian clocks points out that a *Frequency* less oscillator (FLO) has been known for some time (Lakin-Thomas, 2000). The oscillator can create weak 12-13h rhythms in fungal growth patterns in *frq* null mutants, and the authors point out that the *chol-1* mutants, which produce variable growth rates may be creating an apparent circadian rhythm in the FLO (Lakin-Thomas, 2000). Until the effect of *chol-1* mutants on the circadian clock has been assessed using a variety of circadian phenotypes then the debate as to *frq, wc-1* and *wc-2*'s proper role assignment in the clock remains open.

Mammalian clocks

The molecular characterisation of mammalian clocks began with the cloning of the first *per* gene homologue in humans and mice, h*Per1* and m*Per1* (Sun et al., 1997). The m*Per1* transcript cycled in the supra chiasmatic nucleus (SCN), the mammalian site for the master clock. A second *per* homologue, m*Per2* (Albrecht et al., 1997; Shearman et al., 1997) and a third homologue, m*Per3*, were quickly added (Takumi et al., 1998; Zylka et al., 1998). The gene encoding the positive transcription factor m*Clk* was isolated by positional cloning (King et al., 1997) and BAC rescue (Antoch et al., 1997). A *cyc* homologue (*Bmal1* or *Mop3*) had been previously cloned (Ikeda and Nomura, 1997) and was subsequently identified as part of the clock (Abe et al., 1998) (Bunger et al., 2000). Two *cryptochromes* (m*Cry*1 and m*Cry*2) were also identified as clock components (Kume et al., 1999; Thresher et al., 1998) and m*Tim* was shown to physically interact with clock components (Sangoram et al., 1998; Takumi et al., 1999). The homologue of fly *dbt*, *casein*

*kinase 1*ε was shown to phosphorylate the mPER proteins (Keesler et al., 2000; Lowrey et al., 2000). Therefore within three years (1997-2000) homologues of all the main components of the *Drosophila* clock had been identified in mammals (Table 1.1), and all but mTIM had been shown to be involved in normal clock function. Nonetheles, despite the conservation of genes involved there were changes in their function.

mPer 1- 3

The *mPer1* and *mPer2* genes appear to play a role both in responses to light input and in resetting the central components of the oscillator, a function associated with *tim* in *Drosophila*. *mPer1* and *mPer2* transcription is induced by a light pulse in the night (Albrecht et al., 1997; Shearman et al., 1997; Shigeyoshi et al., 1997; Zylka et al., 1998). There is controversy over the role of *mPer2* with Albrecht *et al* (1997) disagreeing with the consensus view of light responsiveness claimed in two later papers (Shearman et al., 1997; Zylka et al., 1998). Additionally, disruption of *mPer2* leads to arrhythmicity (Zheng et al., 1999). To examine the molecular characteristics of light exposure mRNA and protein cycles for the *mPer*s, *mTim* and *mCry*s were studied. Only *m*PER1 was altered by exposure to light in the night, with no alteration in *m*CRY or *m*TIM levels (Field et al., 2000). This provides *m*PER1 with the main function of keeping the central clock entrained to light.

mPer3 cycles in the SCN but is not responsive to light (Takumi et al., 1998; Zylka et al., 1998), so is not an acute light responsive element. However the role that m*Per3* plays is highly contentious with two studies giving it strikingly different roles. The first shows that disruption of the *mPer3* gene leaves *m*PER3 undetectable and leads to no observable change in cycles of *mPer*1, *mPer*2, *Bmal1* and *mCry1*

and a subtle, 0.5 hrs, shortening in behavioural rhythms. Therefore *mPer3* is not needed for rhythmicity in mice (Shearman et al., 2000a) and appears functionally redundant. However in a cultured cell model *m*PER3 translocates *m*PER1 and *m*PER2 into the nucleus (Yagita et al., 2000). In addition, *m*PER1 and *m*PER2 interact physically with mPER3 and, furthermore, mPER3 can perform this function despite the absence of *m*CRY1 and *m*CRY2; a condition that stops the clock functioning (Yagita et al., 2000). Yagita *et al* also identify a cytoplasmic localisation domain (CLD) and nuclear localisation signal (NLS) motif in mPER3 and suggest that interplay between the two motifs regulate the movement into the nucleus as in *Drosophila* TIM (Saez and Young, 1996). This puts mPER3 function as a nuclear localisation factor and although this study does not directly contradict any evidence presented by Shearman *et al*, the nuclear localisation function disagrees strongly with their contention that mPER3 is not important for clocks. A note of caution is that nuclear localisation is a function postulated for CRYs and TIM in mammals (see separate *cry* and *tim* discussions) and resolving whether mPER3, CRY or TIM is responsible for nuclear localisation is paramount. In conclusion the weak phenotypic effects of m*Per3* when knocked-out means it is unlikely to be part of the central clock. Nevertheless, given that mPER3 can interact with many of the clock components and import mPER1 and mPER2 into the nucleus, it probably has a clock function and may be part of an

output pathway. It is intriguing, however, that duplication of m*Per* genes in mammals may have lead to different m*Per*s performing the functions of TIM in *Drosophila*, namely those of nuclear localisation of PER and light input into the central clock.

Cryptochromes

Mammalian cryptochrome function was first described as an input mechanism for light into the clock with the observation that a null mutant of *mCry2* leads to less sensitive light induction of *mPer* (Thresher et al., 1998). In addition, when homozygous double null mutants for *mCry1* and *mCry2* are produced the mice are arrhythmic in darkness (Vitaterna et al., 1999). This was entirely consistent with *mCry*s role as a light sensor and input into the central clock. However three studies then showed that *mCry1* and *2* could act as negative repressors of *mPer*s. *mCry1* (-/-) *mCry2* (-/-) double mutants abolished *mPer1* and *mPer2* cycles, and left both *mPer1* and *mPer2* levels high (Okamura et al., 1999). An additional study by Vitaterna *et al* shows similar responses for these double null mutants, however, they report that only the *mPer1* cycle is removed whilst *mPer2* cycling continues. (Vitaterna et al., 1999). Both studies show that *mCry1 mCry2* (-/-) remove the ability of *mPer1* and *mPer2* to respond to light. Furthermore, the double mutant is arrhythmic in DD. The *mCry* genes are, therefore, important parts of the clock and act as negative repressors of at least m*Per1* and m*Per2* (Shearman *et al.* 2000) . Both *m*CRYs have been shown to form multimeric complexes with the *m*PERs and move into the nucleus at least in cell lines (Kume et al., 1999). Once inside the nucleus CRY acts on CLOCK-BMAL-1 transcription factor complex and inhibits *mPer* transcription (Shearman et al., 2000b) (see below for discussion of *BMAL1* and *mClock* function). However CRY protein levels do not respond to light pulse as TIM does in *Drosophila* (Field et al., 2000). How CRY mediates its function as part of the light input is unknown.

***Clock* and *BMAL1* (m*Cyc*)**

As with Drosophila CLK-BMAL-1 complexes bind to E-box sequences and activate transcription (Gekakis et al., 1998). Mutant mCLK forms a complex, binds E-boxes but fails to activate transcription (Gekakis et al., 1998). E-box elements have been shown to be present in at least mPER1 with five E-box sequences occurring upstream of the coding sequence (Hida et al., 2000). The mCLK-mBMAL1 complex is able to bind the E-boxes of mPER1 and *Drosophila per* and drive the expression of reporter genes (Gekakis et al., 1998; Hida et al., 2000). This complex has been shown to drive the rhythmic expression of m*Per1*; however it is likely that it drives the expression of some or all the negative factors m*Per2*, m*Per3*, m*Cry1* and m*Cry2*. There is also an interesting twist in the positive regulation of m*Per*1. In addition to the E-boxes, upstream of m*Per1* coding sequence there are four cAMP response elements (Hida et al., 2000); these elements may be responsible for the very rapid induction of m*Per1* transcription in response to light, presumably by CREB proteins.

m*Tim*

The human and mouse *Tim* genes present something of a conundrum. They are fully functional proteins which can translocate *Drosophila* PER into the nucleus of *Drosophila* S2 cells and mTIM interacts with mPER1 and inhibits CLOCK-BMAL-1 transcription (Sangoram et al., 1998; Takumi et al., 1999). However m*Tim* RNA levels fail to oscillate in the SCN and TIM is an entirely nuclear protein whose profile is not altered by light pulses. It is therefore unlikely to fufill a similar function to dTIM. Furthermore studies using immmunoprecipitation show that mTIM does not associate with any of the mPERs (Field et al., 2000), and mTIM levels are unresponsive to light although it does associate with CRY. A recent paper,

however, cast doubt on whether the relevant clock gene homologue has been cloned. A second *Drosophila tim* gene, d*tim2*, has been identified in *Drosophila*. Phylogenetic studies show that m*Tim* clusters much closer to d*tim2* than d*tim* (Benna et al., 2000). Therefore it is more likely that m*Tim* is a functional homologue of d*tim2*, a gene with no known clock phenotype. However a search of the now complete human genome failed to find further mammalian h*Tim* homologues (Clayton et al. 2001)

Caesin kinase 1ε and 1δ functions.

The Syrian hamster has provided a clock mutant named *tau* (Loudon et al., 1994). Until recently cloning efforts failed to identify the gene at the molecular level, but *tau* has now been cloned and is found to be the homologue of murine CK1ε (Lowrey et al., 2000). CK1ε is able to bind to mPER1 and mPER2 and the *tua* mutant reduces the rate at which they are phosphorylated (Lowrey et al., 2000). Furthermore there is another CK1 termed CK1δ that can bind and phosphorylate mPER1 *in vitro* (Vielhaber et al., 2000). The phosphorylation apparently masks the mPER1 nuclear localisation signal and retards mPER1 entry into the nucleus (Vielhaber et al., 2000). Therefore its function is similar to its *Drosophila* homologue, DBT, in that it contributes to the length of the cycle by disrupting PER entry into the nucleus, thereby tuning the rhythm to 24 hours. A recent analysis of people suffering Familial Advanced Sleep Phase Syndrome (FASPS) has shown that the disorder maps to m*Per*2, resulting in a serine to glycine mutation in the CK1 binding region (Toh et al., 2001). The missense mutation results in a 20 hr body clock and "morning lark" behaviour. The human FASPS has a strong parallel with

the per^s mutation in *Drosophila* which also produces a shorter rhythm and results from a serine to araginine missense mutation.

***Antheraea pernyi*; new lessons from a silkmoth.**

Analysis of the circadian clock in *Antheraea pernyi* began with the cloning of the *Apper* gene (Reppert et al., 1994) Although there was only a 39% conservation of sequence at the protein level. *In situ* analysis demonstrated that *Apper* mRNA cycled in eight neurons in the central brain and that the protein cycled in the photoreceptors (Reppert et al., 1994). Furthermore, transgenic P-element rescue demonstrated that the gene was capable of rescuing *Drosophila* rhythms although poorly (Levine et al., 1995). In *A. peryni,* PER immunoreactivity was restricted to the photoreceptors, midgut epithial cells and eight large cells in the CNS (Sauman et al., 1996). Transplantation experiments were carried out in which donor larvae were entrained in an LD cycle advanced by eight hours with respect to recipient larvae. The transplantation of the midgut epithium had no effect on the egg hatching of the recipient larvae. However, transplantation of the donor CNS advanced the gated egg hatching behaviour of the recipient by eight hours (Sauman and Reppert, 1998). It is, therefore thought that the *A. Pernyi* body clock resides in the CNS, exerting its control through humoral agent (Sauman and Reppert, 1998).

The natural conclusion from this body of work was that the *A. Pernyi* clock appeared to operate in a conventional *Drosophila* like way. However analysis of the staining pattern of the eight central clock cells in adult revealed some startling surprises. Although these neurons co-express PER and TIM and *per* mRNA in a cyclical manner and these cycles are repressed by light the PER proteins never enter the nucleus (Sauman and Reppert, 1996). This is an absolute requirement for PER to

directly interact with its positive factors and negatively regulate its transcription. Furthermore a cycling antisense *per* mRNA was identified. Duplex formation between the *Apper* sense and the cycling antisense message would prevent translation of the *per* sense message and PER protein levels would fall. A PER cycle could, therefore, be set up without the requirement for PER inhibition of its own transcription through nuclear entry. The cycling antisense mRNA provided a plausible way that *Apper* could set up a negative feedback loop in a different way to *Drosophila*. The antisense transcript could hybridise with the *per* transcript to form double stranded RNA molecules which would prevent translation. However it was later shown that the antisense transcript was produced from an additional copy of *per* on the heterogametic W chromosome of the *A. Pernyi* females, and could therefore not be responsible for rhythmicity in ZZ homogametic males (Gotter et al., 1999). Nevertheless the question of how PER regulated rhythms apparently without nuclear entry was still open.

Bacterial Clocks

Bacterial clocks truly demonstrate the ubiquitous nature of negative feedback in clocks. Initially it was thought that bacteria did not contain circadian clocks, but it was found that nitrogen fixing cyanobacteria had several circadian phenotypes ((Kondo and Ishiura, 2000). A luciferase reporter driven off a *psbA* promoter, a major component of photosystem II, cycled with a 24 h period for 11 days in constant light free-running conditions. This rhythm could be phase shifted by dark pulses in a phase dependent manner, and the rhythm was insensitive to temperature changes (Kondo et al., 1993). A promotor-trapping study then sought to identify how many genes were under circadian regulation. A promotorless luciferase

gene was inserted into the genome by recombination. Surprisingly of the 800 genes that gave enough luminescence to be studied almost all had a circadian rhythm (Liu et al., 1995). Most clones cycled in the same phase as *psbA* but 20% had different phases or profiles. A screen for altered *psbA* luciferase rhythms identified a huge number of mutants (>100), 19 of which mapped to three ORFs, KAIABC, in a single cluster (Ishiura et al., 1998). The genes had no similarity to any of the previous clock genes found. The cluster is transcribed by two promoters, *PkaiA* and *PkaiBC,* with both promotors mediating rhythmicity. *kaiC* appears to act as a negative repressor with overexpression resulting in immediate and complete repression of *PkaiBC.* Whilst kaiA appears to act as a positive factor overexpression dramatically increasing *PkaiBC* mRNA (Ishiura et al., 1998).

Phase shift experiments revealed that KaiC was the negative regulator. In the *PkaiBC-luc* strain *kaiC* was expressed temporarily from a inducible promotor, it caused advances in bioluminescence cycles when *kiaC* was normally on the rise and delays when *kaiC* was normally on the wane (Ishiura et al., 1998).

Furthermore the proteins interact *in vitro* and *in vivo*, KaiC enhancing a KaiA-KaiB interaction suggesting they form a multimer. In addition, the long period KaiA mutant enhances the KaiA-KaiB interaction (Iwasaki et al., 1999; Xu et al., 1999), suggesting that this association may be important in keeping the 24 hour timing of the clock. Further to this a histidine kinase encoded by the *sasA* gene, has been shown to bind KaiC and is essential for the clock to function (Iwasaki et al., 2000). It has also been shown that KaiC can autophosphorylate (Nishiwaki et al., 2000), and KaiC has DNA binding activity which is essential for its clock function (Nishiwaki et al., 2000). The light input has also been characterised in the gene

cikA, a phytochrome responsible for the dark resetting of the clock (Schmitz et al., 2000)

Initially, the cyanobacterial circadian clock appears quite different from that proposed for *Drosophila* and mammals. KaiABC have no similarity to other clock protein and do not contain the PAS domains found in all eukaryotic clocks. In addition the negative factor itself has DNA binding activity and may not just mediate removal of positive regulation. However, the model of negative feedback providing clock function still applies and kinases regulate the clock. The positive factor dimerises to another clock factor to function, and light (or the absence of light) as sensed by a phytochrome, is the major resetting stimulus for the clock.

Plant Clocks.

So far the study of circadian clocks in plants has failed to clearly identify the central components of the clock (McClung, 2000). However a large number of genes have been identified as inputs or outputs of the circadian clock. A genome wide micro-array screen of genes that cycled in free running conditions of constant light showed that 5.5% of studied genes cycled reproducibly ((Harmer et al., 2000) 8200 genes analysed). A large number of these loci encode genes for phenolic sunscreens, photosynthetic components, carbon metabolism, nitrogen and sulphur assimilation, cell elongation, flowering and genes that encode for information sources such as *phyB* and the *cryptochromes*, demonstrating the far-reaching importance of the circadian clock in regulating gene expression.

The light input mechanism of the plant circadian clock is the best characterised component of the circadian mechanism. *Phy A* and *B* have been identified as additive red light inputs and *cry 1* and *2* are redundant blue light inputs.

Furthermore *phy D* and *E* also input into the clock (Devlin and Kay, 2000). *Phy A* requires *cry-1* illustrating a genetic interaction between the two photosensitive light input systems. Additionally, *cry-1* and *2* do not form central components of the circadian clock as they do in mammals because the double *cry-1,2* mutant does not disrupt the clock (Devlin and Kay, 2000). Indeed the quadruple *phyA,B cry-1,2* mutant still has a light sensitive clock so there may be a high degree of redundancy in the light input mechanisms (Yanovsky et al., 2000). Further input mutants have been identified in the *gigantea* and *zeitlupe* genes: *gigantea* encodes a transmembrane protein with mutants altering the phenotype of light sensitivity of flowering and circadian rhythms (Park et al., 1999): *zeitlupe* encodes a protein with a PAS domain with highest similarity to those of the light sensors *wc-1* and *nph-1*(an *Arabidopsis* blue light receptor) and mutants that suggest a role in light input (Somers et al., 2000).

Cycling input and output factors are a feature that plant clocks share with other circadian models. The *phy A* part of the light input pathway, undergoes a daily cycle controlled by the circadian clock (Bognar et al., 1999). *Toc-1*, which appears to be an output transcription factor, is circadianly regulated and feeds back on its own transcription (Strayer et al., 2000). However a gene, or gene network, with negative feedback and a clock phenotype is not necessarily part of the central clock as appears to have been demonstrated with the *wc-1*,*wc-2* and *frq* network in *Neurospora* (Lakin-Thomas and Brody, 2000). The challenge for plant circadian biologist is to try and dissect the central components out of the numerous genes with clock phenotypes. As a note of caution, however, plants appear not to have a central tissue for circadian regulation; parts of the same plants can be entrained to different phases, and the phases are retained after removal of the entraining cue (Thain et al.,

2000). Indeed it will be interesting to see if there is one central molecular mechanism in plant cells or whether circadian regulation is run by a variety of redundant oscillators, some which have already been characterised.

Project aims:

The *per* gene of *Musca domestica* was previously cloned in this laboratory and *Musca per* (*Mdper*) transgenes transformants rescue rhythmicity in *Drosophila* per^0 (Piccin et al 2000). However the *Md*PER protein did not cycle in *Musca* and showed unusual cycling characteristics in the *Drosophila* per^0 transformants. As PER cycles had previously been assigned a role in the canonical clock loop (Ewer et al., 1988) the lack of PER cycle in *Musca* is problematic for the negative feedback loop. Nevertheless, it is obvious from the behavioural rhythmicity of *Musca* and transformants that the clock is perfectly functional in both backgrounds (Piccin et al. 2000). The model would still hold, however, if the MdPER still entered the nucleus in a gated manner in the central clock cells, the true canonical feature of the circadian clock (Lee et al., 1996) and a obligatory part of the molecular mechanism of clock timing. My initial aim, therefore, was to observe *Musca* PER protein by ICC and determine its subcellular localisation during a circadian time course. It became apparent that this would require the manufacture of new *Musca* PER protein specific antibodies requiring the production and purification of *Md*PER fragments. This also offered the opportunity of using the purified *Musca* fragments to perform structural studies. As yet there has been not crystallographic or NMR data on any PER proteins, with the exception of studies on the conformation of the Thr-Gly repeat (Castiglionemorelli et al., 1995). Indeed the PAS domain has not been structurally characterised in any of the clock proteins or indeed any eukaryotic

transcription factors. Due to the surprising result from *Musca*, I decided to extend the use of the comparative approach to observe mRNA and protein cycles within the diptera, including the *per* and *tim* genes of *D. pseudoobscura* and *D. virilis*. It has been taken for granted that the *D. melanogaster* model applies to all Drosophilids. I will test whether this is indeed the case with these other fly species.

Chapter 2:

Materials and Methods

Fly maintenance

Fly stocks of D. *melanogaster D. virilis* and *D. pseudoobscura* were maintained on sugar/agar media (4.63g of sucrose, 4.63g of dried brewers yeast, 0.71 of agar and 0.2g of Nipagin in 100mls of water), in either glass vials (10 X 2.2cm) or one-third pint milk bottles. The stocks were kept in temperature-controlled rooms at 18^0C or 25^0C and in light-dark cycles (LD 12:12).

Musca domestica stocks were quarantined, as they are susceptible to mite infestations. Adults were maintained in Perspex hamster cages sealed with thick ladies' tights. The flies had access to sugar, marvel skimmed milk powder and water, which were distributed through damp tissues kept in 500ml bottles of water. Laying media were made up in yellow gilson tip boxes; tip boxes were filled two thirds full with Jordans bran with the addition of a teaspoon of dried yeast powder, 180mls of skimmed milk and 1.5mls of 20% Nipagin mixed in. The laying media was placed in the fly cage for three to five days, stirring and watering each day to keep moist and aerated. The media were moved to a clean cage and once large maggots could be observed the laying media were allowed to dry to aid pupation. All *Musca* stock were maintained at 25^0C . In order to collect the *Musca,* the flies were entrained as detailed below and then moved into, and maintained on sugar food vials the day prior to the start of the experiment.

Entraining flies to light-dark Cycles (LD)

Flies were placed in temperature controlled incubators (LMS brand) maintained at 18^0C or 25^0C, on a constant cycle of 12 hours light: 12 hours dark (LD 12:12) for at least three complete light cycles. Once the experiment had begun flies were then collected by freezing in liquid nitrogen. Flies collected from the dark were prepared prior to collection by wrapping the vials in foil just as they were entering the dark phase of their cycle. The samples were then collected in a dark room with only a photographic red light source for illumination. This ensured that the flies' circadian rhythms were not perturbed by stray light contamination during processing.

Entraining flies to dark-dark conditions (DD)

Flies were entrained as above, and then moved to constant darkness at the start of their next dark cycle when the vials of flies were wrapped in foil and harvested over three days, as described above.

Preparation of fly head samples

D. melanogaster, D. virilis and *D. pseudoobscura* head samples were all prepared in the same way. First the flies were snap frozen in liquid nitrogen and stored at -80^0C until decapitation. The fly heads were then removed from their bodies by briefly vortexing the frozen flies in a red cap 15ml tube. The body parts were separated from the heads by sieving the samples through metal tea strainers that had been pre-chilled in liquid nitrogen. The separated fly heads were tipped onto a metal plate (chilled on dry ice) so contaminating body parts could be removed with a small brush. The sample was decanted into an eppendorf tube and

stored at -80^0C. *M. domestica* heads were prepared in the same way except they were not sieved but were separated from the body parts on the metal plate using a cooled scalpel.

RNA isolation

RNA was isolated from fly head samples using RNAZOL^{TM}B with a modified protocol. Lysis of the tissue was performed in 1.5ml eppendorf tubes using a mini-pestle as follows. 20mg of frozen tissue was maintained frozen on dry ice, 100μl of RNAZOL was added and the mixture thawed by lysing with the mini-pestle attached to a Black and Decker drill, on full power for 30s. 900ul of ice cold RNAZOL was immediately added and the extraction was performed in accordance with the manufacturer's instructions. This modified protocol achieved 5-10 fold greater yield of pure RNA than a hand lysis method.

RNA quantification

RNA was quantified using a Qiagen RNA club newsletter Number 6 protocol (Qiagen RNA Club, 1998). Sample RNAs are mixed with ethidium bromide solution and compared to similar preparations of yeast tRNA of known concentration on a transilluminator. This provides accurate quantification of small amounts of RNA.

Spectrophotometry was noted to give inaccurate concentrations due to eye pigments being carried over from the head extractions. These pigments then interfere with spectrophotometry. Subsequent phenol chloroform extractions and ethanol precipitations failed to remove the pigments.

First strand synthesis

Fly head mRNA was isolated directly from the tissue using the Qiagen Oligotex method as directed by the manufacturer. First strand synthesis was then performed with AMV-RT and Oligo dT (both Gibco) as recommended by the manufacturer.

PCR

PCR was performed using the Expand Kit (Roche). This consisted of a mixture of *Taq* and *Pfu* polymerase and provided robust PCR along with low error rates whilst still allowing T-A cloning. The manufacturer's recommended reaction conditions were adhered to and reactions were performed on a Hybaid tri block thermal cycler.

Primers were manufactured by PNACL on a 0.2μm scale and lyophilised. They were then resuspended in 0.5mls DDH2O and concentration determined by spectrophotometry. The primer concentrations were then adjusted to a ten times stock of 0.3μM and used at a final concentration of 30ηM, unless otherwise stated. Annealing temperature for each reaction was optimised empirically. The primers used are detailed in the experimental chapters.

Digestions

Digestions were performed using Gibco or NewEngland biolabs enzymes which were used in accordance with the manufacturer's instructions. The enzymes used are detailed in the experimental chapters.

Ligations and transformation of bacterial cells

Ligations and transformations were performed as published (Maniatis et al., 1988) and transformed into either XL1-blue or BL21 *E.coli* strains

Plasmid purification.

Plasmid DNA was purified using Qiagen Mini and Maxi Kits in accordance with the manufacturer's instructions.

Recombinant proteins

Recombinant proteins were produced by expressing in frame PCR products from the pET14b plasmid in the BL21 *E. coli* strain. PCR fragments were engineered to contain 5' nde1 and 3' xho1 sites that allowed directional cloning into the pET14b vector. Expression analysis was performed as detailed in the Novagen pET vector booklet (version 7). All recombinant proteins were tagged with poly histidine tags (His_6) and were purified from cell lysates using the Qiagen NTA-Ni resin.

Protein isolation from flies

Fly heads were collected as detailed above. The tissues were homogenised in approximately three volumes (typically 100μl) of extraction buffer using a micro-pestle and Black and Decker drill, essentially as described in the RNA extraction method. All homogenised tissues were kept on ice during quantification and the

concentration was determined by Bradford assay as detailed bellow, the tissue concentrations were equalised to 2μg/μl by adding extraction buffer.

The extraction buffer consisted of 20mM HEPES, pH 7.5, 100mM NaCl (note this is a modification from the original Edery recipe which uses KCl to precipitate SDS), 0.2mM EDTA, 1mM MgCl, 5% Glycerol, 1mM DTT, 0.1% SDS. Just prior to use, protease inhibitors were added to the following concentrations; 10μg/ml Aprotonin, 5 μg/ml Leupeptin, 1μg/ml Pepstsatin A and 50mM PMSF.

Western and Coomasie analysis

Western analysis was performed on 12% SDS-PAGE gels, essentially as described by Maniatis (Maniatis et al., 1988). Proteins were blotted onto Schlicer and Schull Protean^{TM} nitrocellulose paper, 5% Marvel milk powder in TBST was used as the blocking agent. Blocking was typically performed for one hour at room temperature. The primary antibody was incubated with the filter overnight (16 hours) at 4^{0}C. HRP labelled secondary antibodies were incubated for 1hr and the blot was developed using ECL (Amersham-Pharmacia biotech ECL kit).

Bradford assay in microtitre plates

Protein concentrations were determined using the (Bradford assay). 50μl of BIO-RAD protein reagent was mixed with 150 μl of elgastat (protein free) water in a microtitre plate. The sample was added and made up to 250ul with water. The reaction was left for at least 5 minutes and then the absorbance was measured at 595nm in a biorad model 3550-microplate reader. This was compared to a standard curve of Bovine Serum Albumin. The samples were then diluted until they could be compared to the linear range of the assay. This was determined empirically for every

experiment and the samples were blanked on with appropriate control buffers without tissue or bacterial extracts.

Hydrophobic interaction chromatography

Hydrophobic interaction chromatography was performed on Phenyl, Octyl or Butyl Sepharose^{TM} (Pharmacia Biotech). Yeast lysate was bound to the column in high salt conditions (Tris 20mM, pH 8, 1.7M $(NH_4)SO_4$) and eluted of in a gradient of Tris 20mM, pH 8, 1.7M $(NH_4)SO_4$: Tris 20mM pH 8. All elutions were performed on Biorad Biologic LP chromatography system.

Entrainment and collections for immunocytochemistry

Flies were entrained and collected as detailed above. Flies were anaesthetised with CO_2 and *Drosophila* imbedded directly in Oct^{TM} (Tissue Tech) as detailed below. *M. domestica* were killed in an ethanol chamber and their proboscis removed. The flies were then moved to ice chilled 4% paraformaldhyde PBS fixative. The cuticle surrounding the brain was removed by dissection under a light microscope and the flies bathed in chilled fixative. Dissection typically took two minutes and was performed at 4^0C. Following dissection the flies were fixed overnight at 4^0C in foil wrapped tubes, and were then either imbedded immediately or stored at 4^0C in PBS for up to a month.

Imbedding

A small amount of tissue imbedding medium (OctTM, Tissue Tek) was used to coat a pre-cooled (-20^0C) cryostat chuck. The fly was placed dorsal side down in imbedding medium (Oct^{TM}, Tissue Tek) and arranged so that the wings and legs lay

swept back along the body. Air bubbles were removed using forceps thus preventing interference with head sections. More imbedding medium was added to cover the fly. The chuck and fly were placed in a beaker of 2-methybutane that had been pre-cooled in liquid nitrogen. The beaker was returned to liquid nitrogen to ensure snap freezing. The small ice crystals generated by such rapid snap freezing helped maintain tissue structure during sectioning. Flies prepared with methods that resulted in slower freezing resulted in poor quality sections. Imbedded flies could be stored at -20^0C for 24h.

Cryosectioning

Imbedded flies were placed in a Jung Frigocut 2000E cryostat for 20mins to equilibrate. 10 μm horizontal sections were produced at -14^0C using a Jung stainless steel blade (disposable blades resulted in poor quality sections).

Antibodies

Primary antibodies used are detailed in the experimental chapters. Secondary antibodies used were goat anti mouse CY-3, horse anti rabbit CY-3 and goat anti rat CY-3 (all provided by Jackson Immunochemicals).

Confocal microscopy

A Biorad MRC 600 confocal microscope was used for ICC using an argon ion laser exciting with 488ηm and 514 ηm beam lines. Images were captured through optical filters that only allowed the relevant wave lengths of light form the

excited flourchromes through (CY3 500-530, Propidium Iodide 620-650). The resultant light was then captured on a CCD providing a black and white image that reflects the intensity of the signal produced. False colour was then added later using the biorad-MRC image analysis software (false colour of CY3 was given from the FITC look up table and propidium iodide from the RITC look up table).

RNA protection

RNA protection analysis was carried out according to the published protocol (Ausebel, 1990). Probe incorporation was determined by Crenkov counting in a Packard scintillation counter. All probes were hybridised at 45^0C overnight.

For RNA protection an antisense RNA probe was run off a plasmid containing a RNA polymerase promotor such as T7, T3 or SP6. This was done in the presence of ^{32}P-CTP that was incorporated to produce a radioactively labelled probe. The probe was then hybridised to a large pool of the sample RNA and the nonhybridised RNA digested by RNase A and RNase T1. These enzymes degrade single stranded RNA but cannot degrade double stranded RNA, thus the sense-target/antisense-probe duplex is not digested. The amount of protected probe provided quantification of the amount of transcript present. However duplex formation must be perfect to prevent cleavage of the probe at sites where the duplex "breathes". Therefore the inserts used to make probes must match the transcript perfectly. Polymorphisms or PCR induced errors would result in degradation at the site of mismatch and would make interpretation difficult. The cloning and sequencing of the *per, tim* and *RP49* directly from the experimental stocks prevented cloning of inserts that had mismatches to the stock transcripts. Most stocks are inbred and therefore polymorphisms were unlikely to occur within the

stock, however the stock may be polymorphic with respect to the original sequenced cDNA. In addition three clones were obtained and sequenced for each probe so that the consensus sequence could be determined and consensus clones used, negating PCR induced mismatch errors.

Cloning strategy

Gene specific primers were designed to the *per, tim* and *rp49* genes of all three species using the Primer 3 primer picking programme (primers used are listed in table 2.1). mRNA was isolated from fly heads using the Qiagen Oligotex kit^{TM}, and was used as a template for cDNA synthesis as detailed in the methods section. The resultant first strand synthesis was used as a template for PCR. Fly head samples were collected every 4 hours throughout the night and then pooled, increasing the likelihood of *per* and *tim* transcript being present. The Roche expand kit was used for PCR, which insured low error rate whilst still allowing the use of A-T cloning. PCR fragments were ligated into Nco1 linearised pDk101 (a derivative of pGEM) and blue white selection was used to select positive clones.

Primer	Sequence 5'-3'
D.virilis per F	CGAGAAGGTCGTGAAGAAGG
D.virilis per F	CTGCTCATCCTCCGGTATGT
D.virilis tim F	CGCAACATTTTGCACATACC
D.virilis tim R	GGGTGGCGTTGTGTTACTCT
D.pseudoobscura per F	AGGAGCACTTTCTGCGTGAT

D.pseudoobscura per R	CCTCGGGAACTCTGTAGCTG
D.pseudoobscura tim R	ATGAACGTCAGCATGAGCAG
D.pseudoobscura tim F	GATAGGCTCGTGGTTGTGGT
D.pseudoobscura RP49 F	GTCGGATCGTTATGCCAAGT
D.pseudoobscura RP49 R	CAAGTATTGGCCCTTGAAGC
Musca tim F	CCAGCACATACCCACAGATG
Musca tim R	TAAGGATTCGGATCACTGCC

Table 2.1, PCR primer sequence for RNase protection probes.

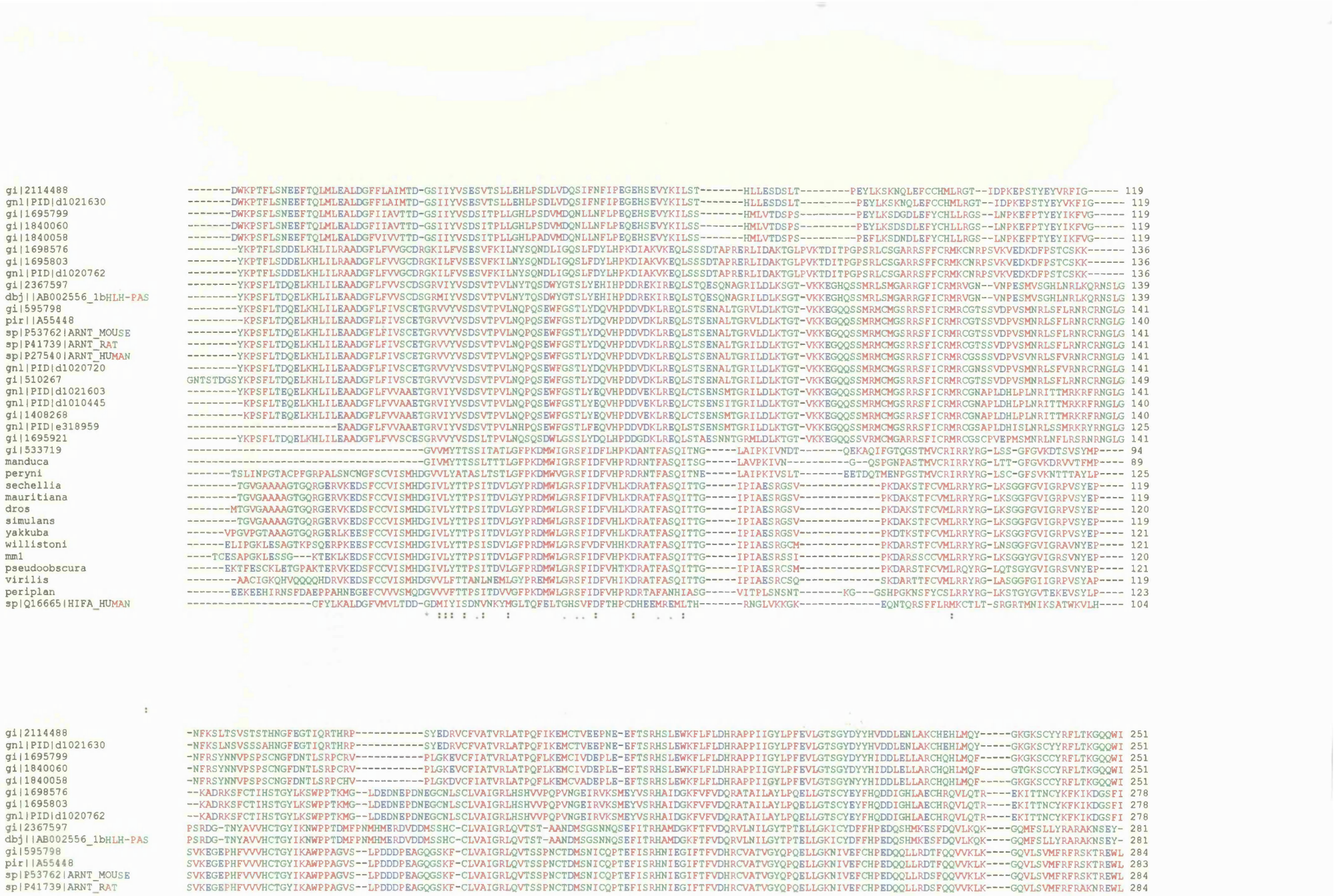
Degenerate PCR

There was no available sequence for *D. virilis RP49* in the databases, therefore degenerate primers were designed to clone a small fragment. All the sequenced dipteran *RP49* genes were identified using BLAST searches and aligned using CLUSTAL-W multiple alignment tool (figure 2.1), and two primer pairs were designed to consensus sequence. Primers were designed by hand using four criteria;

- Primers were designed to consensus sequence on the protein alignments (Figure 2.1).
- Degeneracy was minimised.
- A GC clamp on 3' end of the primer for specificity.
- No degeneracy was allowed at the 3' end of primer.

Primer sites are shown in figure 6.1 and primers listed in table 6.2.

Primer pair 1	Sequence 5'-3'	Degeneracy


gi|2114488
gnl|PID|d1021630
gi|1695799
gi|1840060
gi|1840058
gi|1698576
gi|1695803
gnl|PID|d1020762
gi|2367597
dbj||AB002556_1bHLH-PAS
gi|595798
pir||A55448
sp|P53762|ARNT_MOUSE
sp|P41739|ARNT_RAT
sp|P27540|ARNT_HUMAN
gnl|PID|d1020720
gi|510267
gnl|PID|d1021603
gnl|PID|d1010445
gi|1408268
gnl|PID|e318959
gi|1695921
gi|533719
manduca
peryni
sechellia
mauritiana
dros
simulans
yakkuba
willistoni
mml
pseudoobscura
virilis
periplan
sp|Q16665|HIFA_HUMAN


sp|P27540|ARNT_HUMAN SVKDGEPHFVVVHCTGYIKAWPPAGVS--LPDDDPEAGQGSKF-CLVAIGRLQVTSSPNCTDMSNVCQPTEFISRHNIEGIFTFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQQVVKLK----GQVLSVMFRFRSKNQEWL 284
gnl|PID|d1020720 SVKDGEPHFVVVHCTGYIKAWPPAGVS--LPDDDPEAGQGSKF-CLVAIGRLQVTSSPNCTDMSNVCQPTEFISRHNIEGIFTFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQQVVKLK----GQVLSVMFRFRSKNREWL 284
gi|510267 SVKEGEPHFVVVHCTGYIKAWPPAGVS--LPDDDPEAGQGSKF-CLVAIGRLQVTSSPNCTDMSNICQPTEFISRHNIEGIFTFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQQVVKLK----GQVLSVMFRFRAKNREWL 292
gnl|PID|d1021603 PVKEGEAQYAVVHCTGYIKAWPPAGMT--IPEEDADVGQGSKY-CLVAIGRLQVTSSPVCMDMNGMSVPTEFLSRHNSDGIITFVDPRCISVIGYQPQDLLGKDILEFCHPEDQSHLRESFQQVVKLK----GQVLSVMYRFRTKNREWM 284
gnl|PID|d1010445 PVKEGEAQYAVVHCTGYIKAWPPAGMT--IPEEDADVGQGSKY-CLVAIGRLQVTSSPVCMDMSGMSVPTEFLSRHNSDGIITFVDPRCISVIGYQPQDLLGKDILEFCHPEDQSHLRESFQQVVKLK----GQVLSVMYRFRTKNREW- 282
gi|1408268 PVKEGEAQYAVVHCTGYIKAWPPAGMS--IPEEDADVGQGSKY-CLVAIGRLQVTSSPVCMDMSGMSVPTEFLSRHNSDGIITFVDPRCISVIGYQPQDLLGKDILEFCHPEDQSHLRESFQQVVKLK----GQVLSVMYRFRTKNREW- 282
gnl|PID|e318959 PSKEGEAQYSVVHCTGYIKAWPPAGMT--IPDEDTEAGQTSKY-CLVAIGRLQVTSSPVSMDMNGLSVLTEFLSRHNSDGIITFVDPRCINVIGYQPQDLLGKDILEFCHPEDQQ---------- 237
gi|1695921 PPKDGEPQYVVVHCTGYIKSWPPTGVN--LTDEEADNILGSRY-CLVAIGRLQVTSCPSDTDMNSISVPVEFISRHNCQGLFTFVDHRCMATVGYQPQELLGKNILELAHPEDQELLRDSFQQVVKLK----GQVLSVMFRFLSKTRDWL 284
gi|533719 --FLLKFRFRNISDD---K--GL----------VVYLVIQTVPFFSAYKTPNEILTQEVS-FIMRHSANGNLEYIDPDCVPYLGYIPQDITNRNALVLYHPGDLPFLQEVYQAIVKE----GSVTRSKSYRMVTQNGHFI 212
manduca --FLLKFTFKNVSDE---E--GK----------VIYLVIQATQFFSAFRIPSEVVSKAVP-FVMRHAANGNLEYIDPESVPYLGYLPQDVTDKDALQLYHPEDLDYLQQVYETIVKE----GGVPRTKAYRMMAQNGDYL 207
peryni --FLLKFKFKNVNED---K--GN----------VIYLVIQAVPFFSAFKTSNEVLAKTVS-FVIRHSADGNLEYIDAESVPYLGYLPQDITNRDALLLYHPGDLGYLQEIYGSLVKE----GNVTRSKTYRMMTQNGHYM 243
sechellia --FRLGLTFREAPEEARPDNYMVSNG----------TNMLLVICATPIKSSYKVPDEILSQKSPKFAIRHTATGIISHVDSAAVSALGYLPQDLIGRSIMDFYHHEDLSVMKETYETVMKKGQTAGASFCSKPYRFLIQNGCYV 251
mauritiana --FRLGLTFREAPEEARPDNYMVSNG----------TNMLLVICATPIKSSYKVPDEILSQKSPKFAIRHTATGIISHVDSAAVSALGYLPQDLIGRSIMDFYHHEDLSVMKETYETVMKKGQTAGASFCSKPYRFLIQNGCYV 251
dros --FRLGLTFREAPEEARPDNYMVSNG----------TNMLLVICATPIKSSYKVPDEILSQKSPKFAIRHTATGIISHVDSAAVSALGYLPQDLIGRSIMDFYHHEDLSVMKETYETVMKKGQTAGASFCSKPYRFLIQNGCYV 252
simulans --FRLGLTFREAPEEARPDNYMVSNG----------TNMLLVICATPIKSSYKVPDEILSQKSPKFAIRHTATGIISHVDSAAVSALGYLPQDLIGRSIMDFYHHEDLSVMKETYETVMKKGQTAGASFCSKPYRFLIQNGCYV 251
yakkuba --FRLGLTFREAPEEARPDNYMVSNG----------TNMLLVICATPIKSSYKIPDEILSQKSPKFAIRHTATGIISHVDSAAVSALGYLPQDLIGRSIMDFYHQEDLSVMKETYEMVMKKGQTAGASFCSKPYRFLIQNGCYV 253
willistoni --FRLGLTFREAPEEARPDNYMVSNG----------TNMLLVICATPIKSSYKVPDEILSQKSPKFAIRHTATGIISHVDSAAVSALGYLPQDLIGRSIMDFYHHEDLSVMKDTYETVMKKGQTAGASFCSKPYRFLIQNGCFV 253
mml --FRLGLTFREAPEEARSDNSLPS-G----------TNMLLVICATPIKSAYKVCDELLSRKTPKFAIRHTKTGIISTVDSGAVSALGYLPQDLIGRSILDFYHHEDLTVLKEIYETVMKKGQTAGASFCSKPYRFLVQNGCYV 251
pseudoobscura --FRLGMSFREAPEEERSDNYMVANS----------SNMLLVICATPIKSSYRVPEEIHSQRSPKFAIRHTAAGIISHVDSAAVSALGYLPQDLMGRSIMDLYHHDDLPVIKEIYESVMKKGQTAGASFCSKPYRFLIQNGCYI 253
virilis --FRLGLTFREAPEEVQPDGCTLSNA----------TSMLLVISATPIKSCYKEPDEFLSPKGPKFAIQHTAAGIISHVDTAAVSALGYLPQDLIGRSILDFYHHEDLSDIKDIYEKVVKKGQTVGATFCSKPFRFLIQNGCYI 251
periplan --FQLNMTFRELLPHSNPLELEGNTSP-------ESV-PGGCNNMFLVITAKLICPAYKHAGETCA--SPKFVTRHLATCKLNYVDPECMPYLGYLPHEMLGNSVLDFYHPEDLPFLKEVYQIVMQEN---GAPFRSKPYRFRSHNGGYI 258
sp|Q16665|HIFA_HUMAN --CTGHIHVYDTNSNQPQCGYKKPPMT----------CLVLICEPIPHPS----NIEIPLDSKTFLSRHSLDMKFSYCDERITELMGYEPEELLGRSIYEYYHALDSDHLTKTHHDMFTK-----GQVTTGQYRMLAKRGGYV 226
WLQTHYYITYHQWN 265
WLQTHYYITYHQWN 265
WLQTHYYITYHQW 264
WL 253
WLQTHYYITYHQW 264
TLRSRWFSFM 288
TLRSRWFSFM 288
TLRSRWFSFM 288
WMRTSSFTFQNPYS 298
WMRTSSFTFQNPYS 297
WMRTSSFTFQNPYS 298
WMRTSSFTFQNPYS 298
WMRTSSFTFQNPYS 298
WTRTSSFTFQNPYS 298
WMRT 296
WIRTSSFTFQNPFS 298
KVETEWSAFINPWSRKLEFVNGKYYIIEGPANPDVFESPDPEKTPKLTEE 262
KLETEWSS 215
KVETEWSAFIN 254
LLETEWTSFV 261
LLETEWTSFVNPLETEWTSFVNP 274
LLETEWTSFV 262
LLETEWTSFVNPWSR 266
LLETEWTSFVNPWSRKLE 271
LLETEWTSFVNPWSRKLEF 272
LLDTEWTSFVNP 263
LLETEWSSFVNPWSRKLEF 272
LLETEWTSFVNPWS 265
LLETEWSSFVNPWSKKL 275
WVETQATVIYNTKN 240

Forward	AARTGGMGIAARCCIAARGG	256
Reverse	CARTAIACICKRTTYTGC	256
Primer pair 2		
Forward	CAYAARTGGMGIAARCCIAARGG	256
Reverse	TGIARRAAYTTYTTRAAICC	256

Table 2.2 Degenerate primer design

PCR conditions were optimised and primer pair 1 and 2 produced a single bands of the right size with two step PCR (initial anneal at 30^0C for five cycles then a step up to 50^0C for remaining cycles, primer concentration 6μm (20-fold normal)). Primer pair 1 was first to produce three positive clones, which were then sequenced. The all the sequences showed a high degree of homology to *D. melanogaster RP49* during a BLAST search, indicating that the *D. virilis RP49* gene had been cloned.

Clones for *Musca RP49* were kindly donated by Alberto Piccin. RNA probes were manufactured so that they ran in the antisense direction; 5' polymerise site-stop codon- ATG start, with respect to full gene cDNA sequence.

Probes were produced from linearised plasmids (see table 2.3 for details) and, where necessary, the fragments were blunted using T4 DNA polymerase (Gibco) in accordance with manufacturer's guidelines. RNA probes were produced using Newengland biolabs SP6, T3 or T7 polymerase in accordance with the manufacturer's instructions, and the probes were purified by gel filtration. They were then counted for activity and 10^5 cpm was hybridised to 10μg of total RNA for *D. virilis* and *Musca* or 20μg for *D. pseudoobscura*. The RNA was then degraded by

RNase A and T1 under optimised conditions and the duplexed RNA was denatured and resolved on a sequencing gel.

Probe	Linearised with	T4 polymerase blunted	RNA Polymerase.	Probe length	Protected fragment length
Virilis per	EcoRV	-	T7	343	241
Virilis tim	EcoRV	-	T7	420	309
Virilis RP49	Sph-1	Yes	SP6	237	136
Pseudoobscura per	Sph-1	Yes	SP6	344	233
Pseudoobscura tim	Sph-1	Yes	SP6	168	347
Pseudoobscura RP49	EcoRV	-	T7	157	91
Musca tim	EcoRV	-	T7	770	550
Muscas RP49	Sa1-1	-	T7	150	91

Table 3.3 details of procedures used for probe runoff.

Analysis of message abundance

The protected product was resolved on a sequenceing gel (biorad system) along with radiolabelled 25 bp markers (Gibco BRL, see manufacturers notes for ATP exchange radiolabelling). The gel was dried and exposed to a phosphor imager

screen for 12-168 hrs and the resultant image was captured on a Molecular Dynamics phospor imager. Bands were manually chosen and the intensity and area of the protected band used to produce a band "volume". The protections and anlaysis were repeated for two independent replicas of each experiment and the band "volumes" graphed.

Protein cycle analysis

The abundance of the PER and TIM in *D. pseudoobscura, D. virilis* and *Musca* were observed by western analysis. Blots of proteins, extracted over a time course, were probed with a battery of antibodies raised against PER or TIM protein fragments. The antibodies were gifts from a variety of laboratories. The specificity of the patterns achieved was verified by comparison with the pattern achieved with other antibodies. A similar band profile, on replica blots from two independently produced antibodies, provides evidence of a specific interaction. Also cycling bands that occurred close to the predicted size of PER or TIM were assumed to be due to a specific antibody staining.

Method

Soluble protein was extracted from head tissues that had been collected over a circadian time course. Briefly, flies were collected and snap frozen, the heads were then collected, without defrosting, and lysed by homogenisation in ice cold protein extraction buffer. The protein extraction buffer contained protease inhibitors and the detergents SDS and Triton-X-100, which aid tissue break up but were unlikely to be in sufficient concentrations to solublise insoluble proteins. Protein concentration was then determined by Bradford assay (see methods section) and equalised to

2mg/ml. The 20μl of each of the protein extracts was loaded in each lane and was then subject to SDS-PAGE and blotted.

Image capture and Image analysis

All western images were analysed using the Amersham Pharmacia Biotech 1D Image analysis software. The computer program chose the bands (which were confirmed by eye) and band "volumes" were calculated and plotted. Unless otherwise stated at least three replica were performed for each experiment.

Chapter 3:

PER expression in *Musca domestica*.

Introduction

Previous working in this laboratory, primarily by Alberto Piccin, had suggested that *Musca* provides an interesting comparative model for circadian clocks (Piccin, 1998). *Musca* showed no rhythmic PER protein expression (**Fig. 3.1**), but the transcript appeared to cycle with a peak early at night (**Fig. 3.2**) When a construct containing the coding sequence of *Mdper* was used to rescue arrhythmic *D. melanogaster per* null mutants it provided robust behavioural rescue (Piccin *et al.*, 2000). In transformants the *Mdper* transgene gave normal mRNA cycles (**Fig. 3.3**), however, even though the protein levels cycle (**Fig. 3.4**) there was no lag between the mRNA and protein cycles, both peak late in the light phase of LD (ZT10-12) and decline with similar profiles.

Although there were rhythmic mRNA and protein cycles in the transgenic transformants the lack of a PER cycle in *Musca* is problematic for the negative feedback loop. If one of the so-called "state variables" is held constant then the clock should stop (Aronson, et al 1994). Therefore in *Musca* either PER is not a state variable or the model of negative feedback based on PER does not apply. It is difficult to see how a protein that is a canonical part of the clock could rhythmically feedback on its own transcription when it is not rhythmically produced. A plausible explanation is that although PER is not rhythmically expressed it does still move into the nucleus in a rhythmic manner. In the traditional feedback model the rhythmic build up of both PER and TIM and their interactions provide the information necessary for the lag in entry into the nucleus, and subsequent

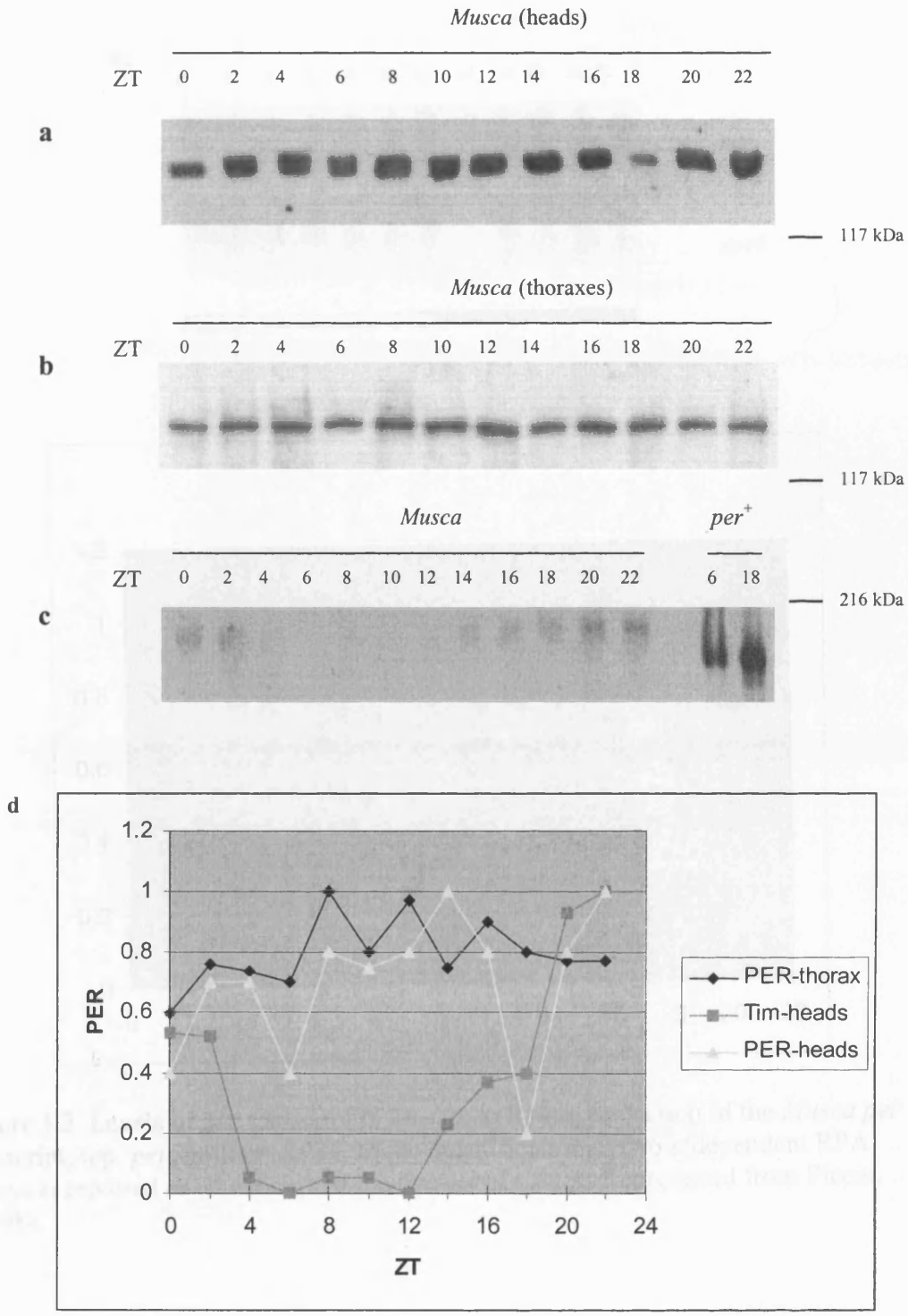


Figure 3.1. **a** and **b**: immunoblots of *Musca* total head and thorax extract, respectively, collected at various Zeitgeber times and analysed with αPER-978. **c**: immunoblot of head extracts probed with anti-Tim. The quantification of these data is shown in **d** (from Piccin,1998)

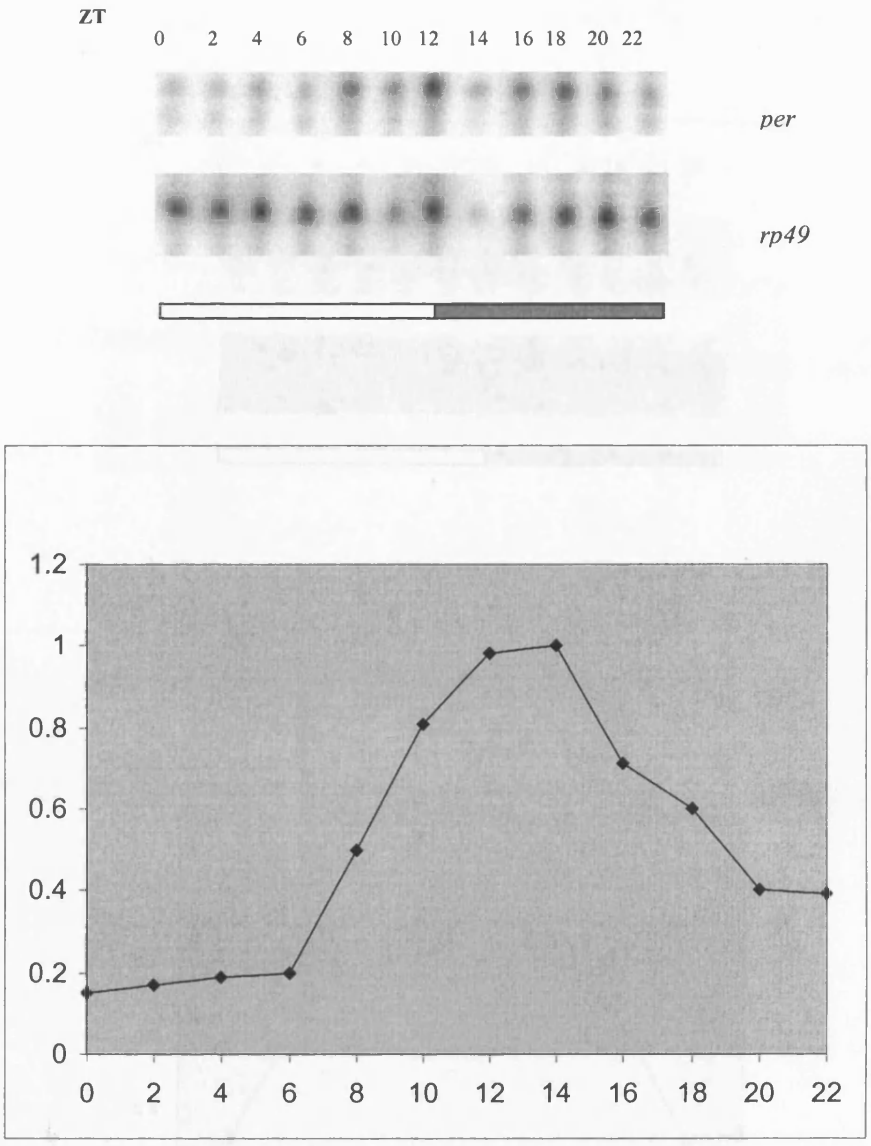
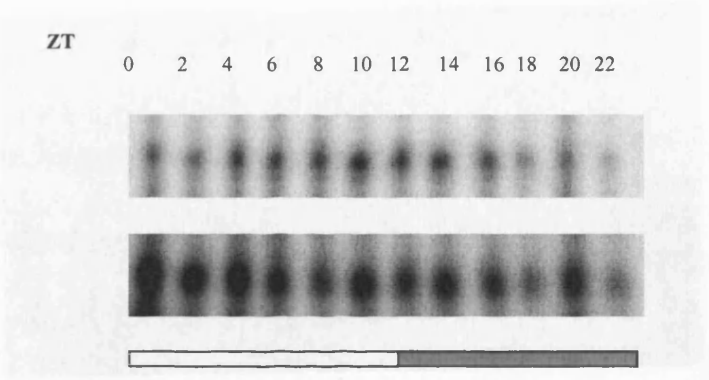


Figure 3.2. Levels of *per* transcript in *Musca*. **a**: RNase protection of the *Musca per* transcript; top: *per*; bottom: *RP49*. **b**: the quantification of two independent RPA assays is reported as an average of the two set of values; Reproduced from Piccin, (1998).

a



b

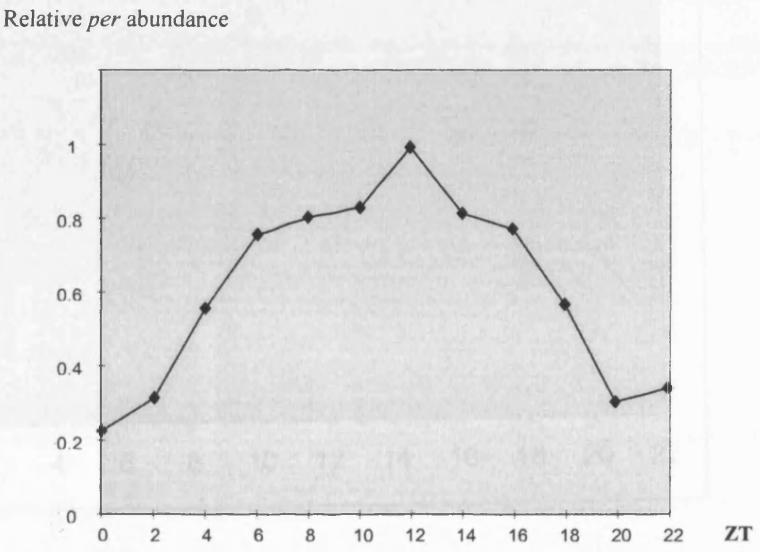
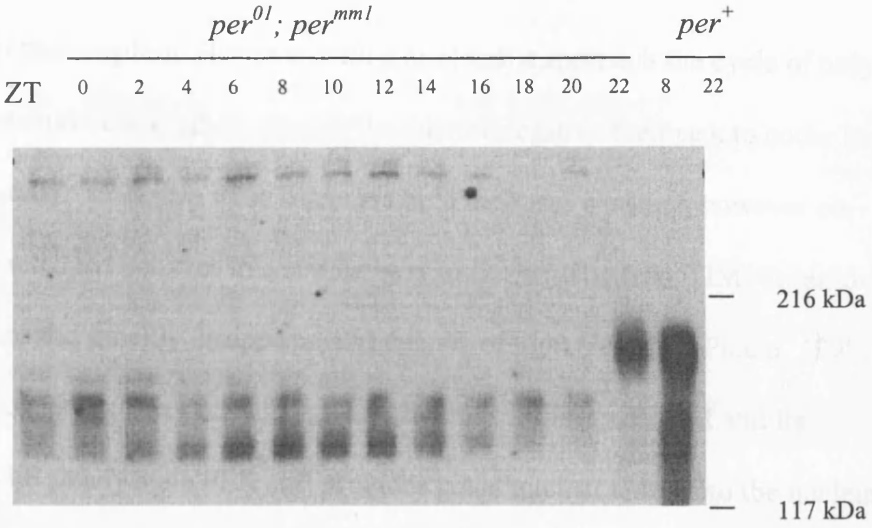


Figure 3.3, *Musca per* transcript in *per*^{0} *D. melanogaster* transformant. **a**: RNase protection of *Mdper*, top RP49, bottom. **b:** Quantification of two independent RPA assays. Reproduced from Piccin (1998).

a:



b:

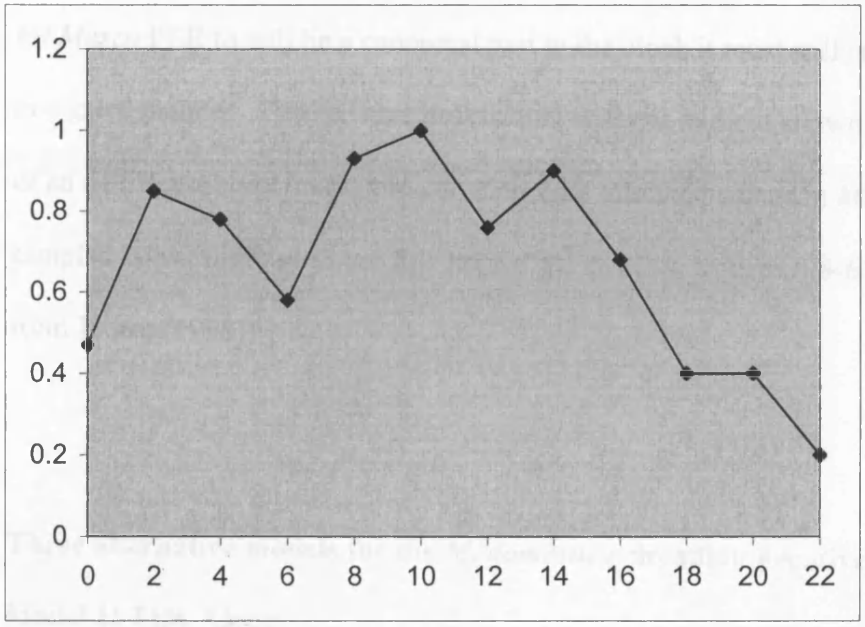


Figure 3.4 Oscillation in the levels of PER *D. melanogaster* per^0 *Mdper* transformant carrying two doses of the *Mdper* transgene. **a**: immunoblot probed with αPer-978 antibody. **b**: graphical representation of the amount of PER calculated as the sum of the 130 and 150 kDa bands (from Piccin 1998).

repression of transcription. However with a minimalist approach the cycle of only one protein would be enough to provide the cue for negative feedback to occur in a rhythmic manner. In *Musca* PER levels are held high and constant, however on westerns *Musca* TIM behaves in a similar way to *D. melanogaster* TIM. It builds up in the night and quickly disappears at the onset of light (*Md*TIM, Piccin, 1998; *d*TIM, So *et al* 1999). It is possible that the rhythmic build up of TIM and its interaction with steady state PER still provides gated nuclear entry into the nucleus with an appropriate delay after transcription. Thereafter the negative repression could take place much in the way it is shown to occur in *D. melanogaster*. However for *Musca* PER to still be a canonical part to the clock it must still enter the nucleus in a gated manner. Furthermore provisional analysis had not shown any movement of an αPER antibody immunoreactive product into the nucleus in *Musca* heads sampled late at night or in the first half of the day (see **figures 6.6-6.12** reproduced from Piccin (1998))

Three alternative models for the *M. domestica* circadian negative feedback loop

Model 1: TIM Alone.

TIM alone is responsible for the lag in negative feedback. *Musca* TIM is translated some hours after it is first transcribed giving a sufficient lag for the negative feedback loop to create a 24h cycle. Alternatively the dynamics of the PER-TIM interaction is such that it does not dimerise efficiently, creating the necessary lag through a delay in entrance into the nucleus. Once TIM accumulates it interacts with the pooled PER to enter into the nucleus, where the dimer represses

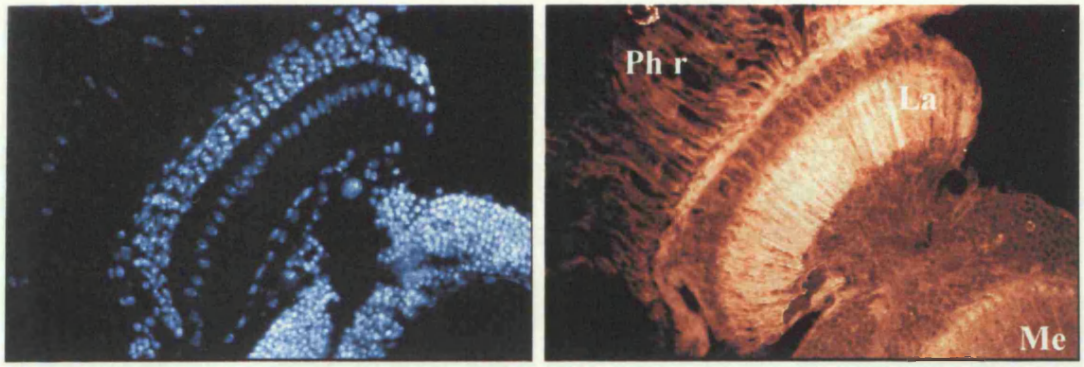


Figure 6.7. Staining observed in photoreceptors and lamina of a housefly collected at ZT12. Strong staining is detected in the lamina but no anti-Per signal is found in the photoreceptor cells. On the left is the DAPI staining which marks the photoreceptor nuclei. On the right, the same frame is seen with a different filter in order to visualise the antibody. **Ph r**: photoreceptors; **La**: lamina; **Me**: medulla. (400X).
Control slides processed with the secondary antibody only showed staining associated with the photoreceptors and none at the other structures detected with the monoclonal mix (data not shown). No attempt was made to preadsorb the primary antibody against the Per protein.

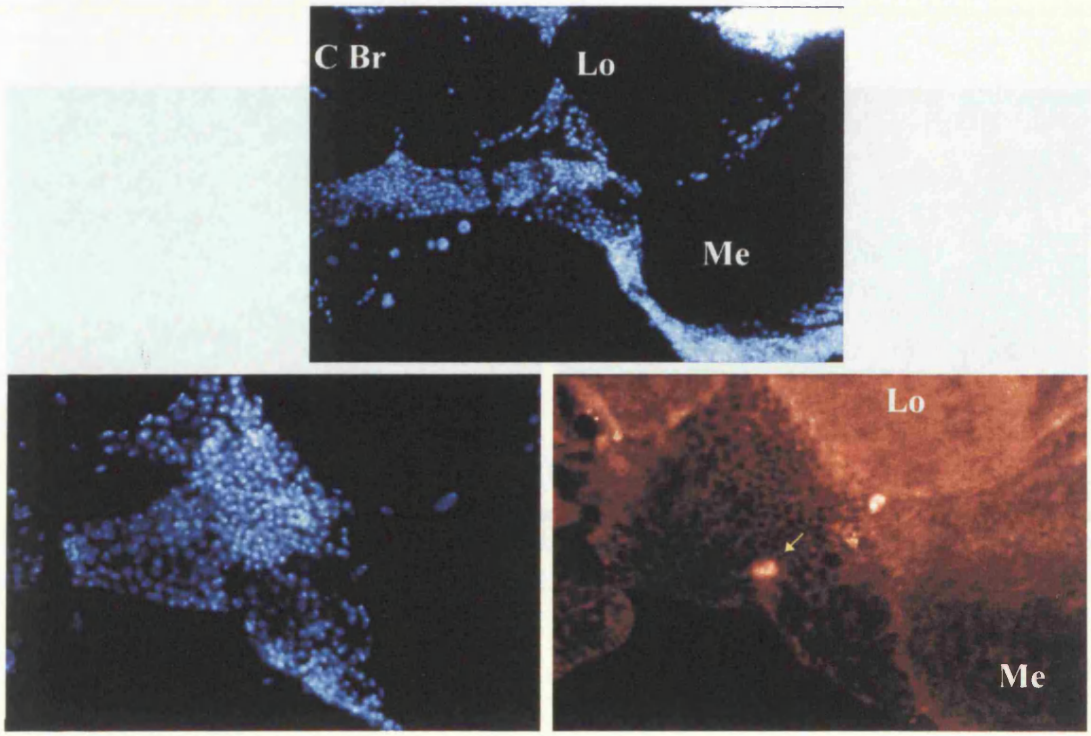


Figure 6.8. Section showing lateral neuron positive staining at ZT12 (yellow arrow). Notice that the Cy3 fluorescence is emitted in a region devoid of nuclei (no DAPI staining).**C Br**: central brain: **Lo**: lobula; **Me**: medulla. The lower magnification shows the location of the lateral neurons in respect to other major structures of the brain (bottom images: 400X; top: 200X).

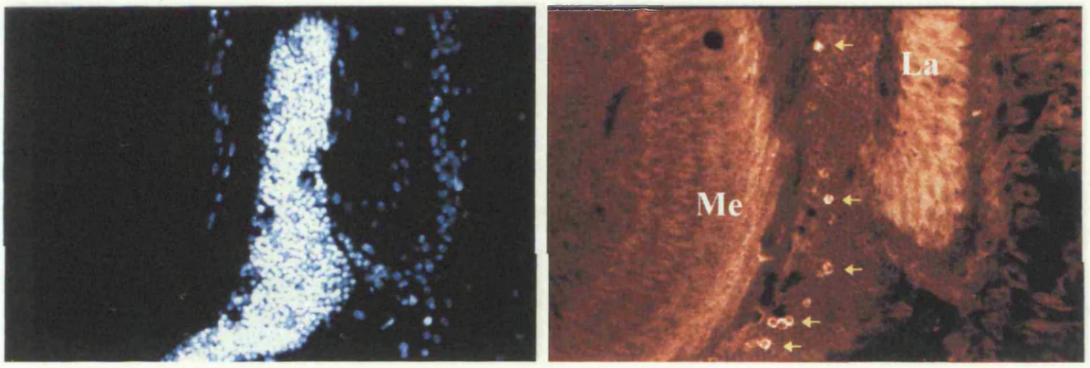


Figure 6.9. Glial staining at ZT0 (yellow arrows). The dense concentration of nuclei visible in DAPI correspond to glial cells between the lamina and the medulla. Positive antibody staining is detected in several glial cells, but in none is the staining nuclear (note the characteristic "doughnut" shape of the Cy3 signal). **La**: lamina; **Me**: medulla. Left: DAPI staining; right: Cy3 staining (400X).

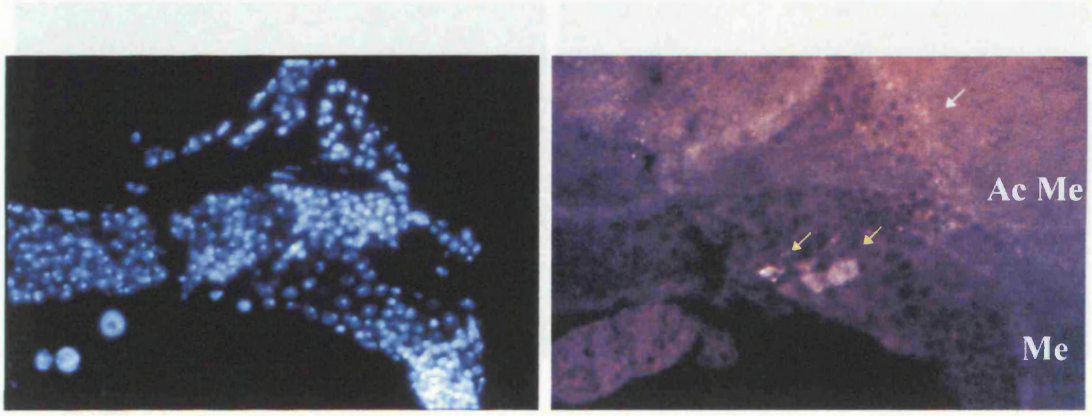


Figure 6.10. Positive signal detected in the neuropil of the accessory medulla (white arrow), a region of the fly brain where the LN extend their projections. Also visible Lateral Neuron staining (yellow arrows). **Me**: medulla; **Ac Me**: accessory medulla. Sample collected at ZT12 (400X).

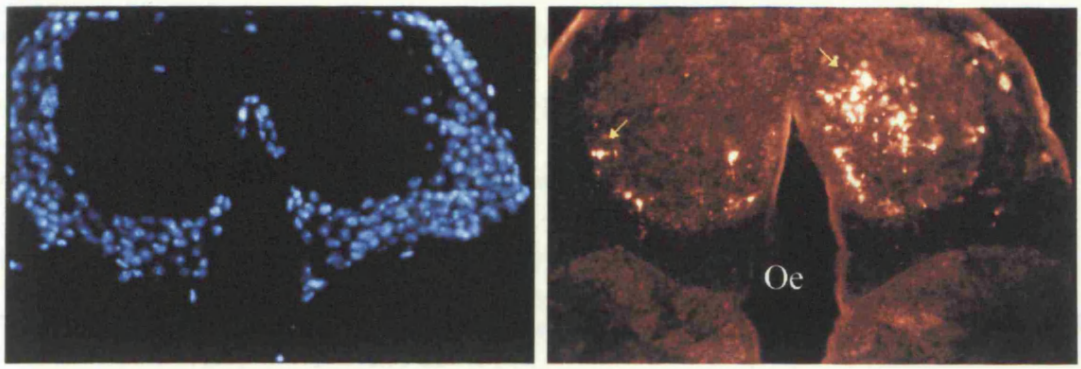


Figure 6.11. Synaptic Per expression in the central brain neuropil at ZT20. **Oe**: oesophagus. Left: DAPI staining; right: Cy3 staining (400X).

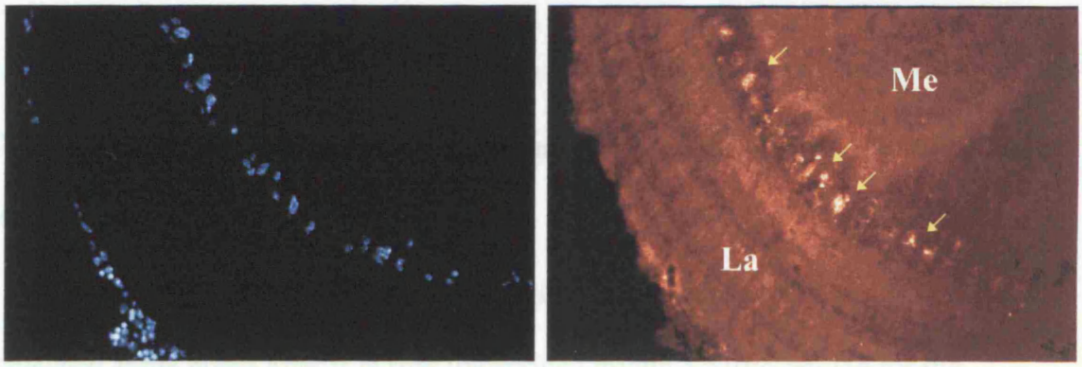


Figure 6.12. Glial staining at ZT4. Nuclear staining is visible in glial cells between lamina and medulla (confront position of positive Per staining, indicated by yellow arrows, and that of glial nuclei evidenced by DAPI). **La**: lamina; **Me**: medulla. Left: DAPI staining; right: Cy3 staining (400X).

per and *tim* transcription. TIM levels then fall with the onset of light and repression is removed.

Model 2: The PER Cycle is swamped by non-cycling of other tissues.

The observed lack of cycling is due to differential expression of PER and TIM throughout the head of the fly. If a large population of cells have PER expression without DBT or other degradation factors then PER would accumulate in these cells. However the equivalents of the LNs could still have rhythmic expression of all the necessary clock genes. *Musca* PER could be degraded through the DBT mediated pathway and the clock would function as in *D. melanogaster.* This would remain undetected in westerns as the whole head was assayed and a PER cycle would be swamped by non-rhythmic expression elsewhere. This is not unprecedented as *D. melanogaster* PER expression is not rhythmic in the ovaries, and if the whole abdomen is studied no cycle is observed, even though other tissue have PER cycles (Hardin *et al.*, 1994).

Model 3: Rhythmic translation of functional PER.

Only a small proportion of PER manufactured each day is functional, in that it dimerises with TIM and moves into the nucleus. It is this small fraction of the daily PER manufactured that is responsible for generating rhythmicity. This functionality could be conferred by phosphorylation through DBT kinase. The inactive form of PER is not modified by any kinase and is unable to bind TIM, so it has no effect on the clock and would presumably build up in the cytoplasm. The same kinase would target PER for degradation so the inactive form would pool while the active form goes through the daily cycle of production and degradation.

The casein kinase family, to which DBT belongs, typically require the target protein to be prephosphorylated . The kinase involved in this prephosporylation may be less efficient in *Musca* and lead to a significant proportion of PER being completely unphosphorylated. This model assumes that phosphorylation is required for PER-TIM binding *Musca*. There is little evidence for or against phosphorylation being required for dimerisation, although fragments of PER and TIM can bind in the yeast two hybrid system without the presence of any *D. melanogaster* kinases (Gekakis et al., 1995).

The production of a small proportion of active PER would be difficult to observe. It need not have any observable change in mobility when compared to the presumed non-functional PER. One phosporylational event could confer the functionality through changing the protein conformation (Karin and Hunter, 1995). Indeed there is a precedent in clocks, with the single phosphorylation of Ser-513 in FRQ being responsible for its rapid degradation (Liu et al., 2000). This kind of event would be difficult to detect on westerns as a single phosphorylational event is unlikely to profoundly alter the proteins migration. In this model the apparent non-cycling of PER would be mediated through changes in *Md*PER amino acid sequence, which results in a protein that is less efficiently phosphorylated by a kinase. In addition changes in *Musca* DBT kinase activity could be responsible.

All of these models assume that the genes involved in the circadian clock are conserved between *Musca* and *D. melanogaster*. This appears a reasonable assumption considering the level of conservation of genes between *D. melanogaster* and mammals(Reppert, 2000).

Given these three alternative models it is important to determine whether, in *Musca,* PER enters the nucleus in the putative pacemaker cells and to observe its

general expression pattern. If it is shown not to enter the nucleus then it would raise serious doubts about the role of PER in the circadian timing of *Musca*. The circadian clock would function in a similar manner to the enigmatic *Antheraea peryni* in which PER does not move in and out of the nucleus (Sauman and Reppert, 1996).

In this chapter a collection of antibodies raised against PER were used to perform immunocytochemistry (ICC). Immunoflorescence based detection was used in conjunction with confocal microscopy, allowing counterstaining of the nucleus along with florescent staining of the PER protein. This provided unambiguous determination of cytoplasmic or nuclear localisation of PER. Previous studies had used enzyme based techniques for protein localisation (Siwicki et al., 1988; Zerr et al., 1990). This has the disadvantage that no counter stain could be used. Instead the pattern of staining was used to assign cellular location, "doughnut" patterns being assigned as cytoplasmic and "tight ball" staining as nuclear. This however introduced the problems of interpretation into the analysis and required confirmation by electron microscopy (Liu et al., 1992). By contrast, confocal microscopy not only allows for counterstaining but the cells can be optically sectioned on the microscope. This acts to reduce background florescence that can occur through staining outside the plane of the optical section and permits exact determination of PER location in each cell. Whole brains of *Musca* were sectioned to 10μm in a cryostat (see methods section.). Cryogenic sectioning was chosen because it was most likely to preserve the antigens. To provide a positive control ICC was performed on the mutant *D. melanogaster* strain *yellow, white* (*yw*). The mutant *Musca* strain *white* (*w*) was also used for ICC as white eyed flies avoid the autoflorescence of eye pigments.

Methods

Flies were entrained for at least three days in LD at 25^0C and were collected every two hours and immediately prepared for cryosectioning, as detailed in the methods section. 10μm sections were prepared on a cryostat, fixed and then used for ICC analysis. The antibodies used are detailed in the results sections and table 3.5. Immunoflorescence was visualised using a confocal microscope and captured using a CCD camera. Enzyme based staining was visualised on a compound microscope and captured on a digital camera.

Results

D. melanogaster positive controls:

D. melanogaster PER expression provides a convenient positive control. Most of the antibodies used have been raised and titred against *D. melanogaster* PER. In addition, the *D. melanogaster* staining pattern provides a system in which the observation of nuclear or cytoplasmic location can be developed. Initial studies using enzyme based labels did not report cellular location (Siwicki et al., 1988) and only through careful examination could cellular localisation be observed (Zerr et al., 1990). However the use of florescence labelling in conjunction with confocal microscopy allows more detailed and less ambiguous assignment of cellular and subcellular location (Ewer et al., 1992; Saez and Young, 1996). Therefore fluorescent based labels were used with the αPER-978 antibody (a gift from J Hall) and the pattern was visualised on a confocal microscope. αPER-978 generated a strong staining pattern throughout the CNS. The results confirmed those of previous studies (Ewer et al., 1992; Liu et al., 1988; Siwicki et al., 1988; Zerr et al., 1990). **Figure 3.5** shows the patterns obtained from *yw D. melanogaster* sectioned

Antibody Indentity	**Laboratory source**	**Immunogen**	**Type of antibody produced**	**Pattern achieved**	**Figures**	**Reason for rejection**
αPER-KS1 (10C3C9)	Kathy Siwicki	*Manduca/Drosophila* PER chimeric protein	Mouse monoclonal	Strong pattern of many doughnut shaped cells throughout the CNS.	3.11	Not repeatable, Probably due change in activity of a cell line.
αPER-KS2 (5F7F6)	Kathy Siwicki	*Manduca/Drosophila* PER chimeric protein	Mouse monoclonal	Strong pattern of many doughnut shaped cells throughout the CNS.	3.11	Not repeatable, Probably due change in activity of a cell line.
αPER-978	Jeff Hall	*Drosophila* PER	Rabbit polyclonal	Very weak pattern requiring optical enhancement.	3.7-3.10	Same pattern obtained in negative control.
αPER-473	Alberto Piccin/ Kyriacou	*Musca* PER Fragment	Rabbit polyclonal	Very strong stain of the entire optic lobes in *per^0 musca* transgene, no apparent strong *Musca* signal.	3.13	Non specific staining of optic lobes with no observable specific stain in the central brain of *per^0 musca* transgene. No *Musca* pattern.
αPER-474	Alberto Piccin/ Kyriacou	*Musca* PER Fragment	Rabbit polyclonal	Very strong stain of the entire optic lobes in *per^0 musca*	3.13	Non specific staining of optic lobes with no observable specific stain in the

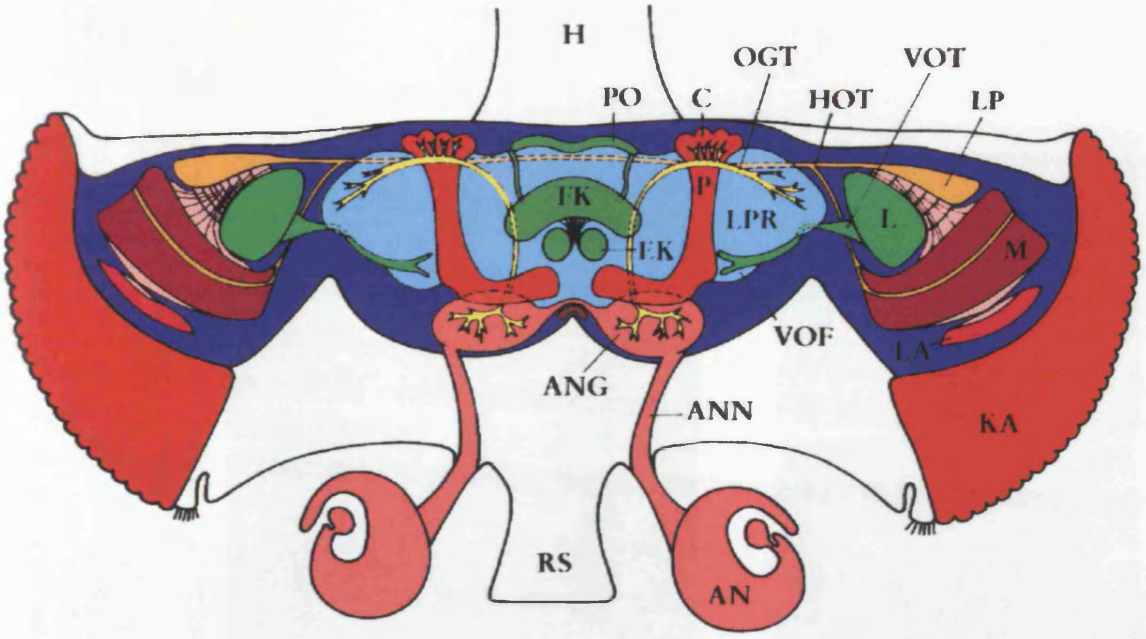
				transgene, no apparent strong *Musca* signal		central brain of *per^0 musca* transgene. No *Musca* pattern.
αPER-Full	Micheal Rosbash	*Drosophila* PER	Rabbit polyclonal	No specific staining	3.13	Used at maximum practically achievable concentration. Failure to produce any staining pattern mitigated against further trials.
αPER-NTer	Micheal Rosbash	*Drosophila* PER, purified against N-terminal fragment (1-640aa)	Affinity purified Rabbit polyclonal	No specific staining	3.13	Used at maximum practically achievable concentration. Failure to produce any staining pattern mitigated against further trials.
αPER-RAT4	Micheal Rosbash	*Drosophila* PER, purified against N-terminal fragment (250-450aa)	Affinity purified Rat polyclonal	No specific staining	3.13	Used at maximum practically achievable concentration. Failure to produce any staining pattern mitigated against further trials.
αTIM-316	Justin Blua/ Micheal Young	*Drosophila* TIM	Rabbit polyclonal	Stained all the neurons in *Musca* head sections.	3.12	Presumed as non-specific.

Table 3.5 Antibodies tested for specific *Musca* PER and TIM staining.

at ZT 17-18, which consists of strong labelling at the tips of the photoreceptors with a few cells at their base also staining. In addition there is also a band of glial cells on the periphery of the lamina. Further glial staining occurs around the medulla, lobulla and lobulla Plate, with characteristic large glial cells lighting up in an area surrounded by the three structures. The large glial, located in the neuropil between the lobula and the central brain, are very distinctive because of their size and position. Assignment of cell type was based on Ewer *et al* (1992) in which glial staining was differentiated from neuronal staining by the use of αELAV neuronal counter stain. Careful comparison of the pattern achieved to that in the literature allows accurate assignment of cell type, cell location was identified by comparison to anatomical maps od the drosophila brain, shown in **figure A**. Furthermore the excellent flybrain database tutorials and resource of horizontal sections was used to firmly identify the anatomical structures of the fly brain.

Figure 3.6, shows the cytoplasmic staining in the lateral neurons at ZT 17. The propidium iodide counter stain allows a more confident assignment of subcellular distribution. In addition to the pattern mentioned in the literature it is worth noting that many neuronal cells surrounding the LNs exhibit weak staining patterns. **Figure 3.7** and **3.8** illustrate the movement of PER into the nucleus in the photoreceptors. In **Plate 3.7** at ZT 17, when PER is cytoplasmic in the LNs, there is a dispersed cytoplasmic staining pattern for PER, and doughnut shapes are much more difficult to observe. In **Figure 3.8** the staining pattern becomes much more punctate and shows nuclear PER distribution in the photoreceptors without the need for a nuclear counter stain.

In conclusion the LNs stained strongly along with small glial cells throughout the optic lobes and CNS and the photoreceptors also stained strongly



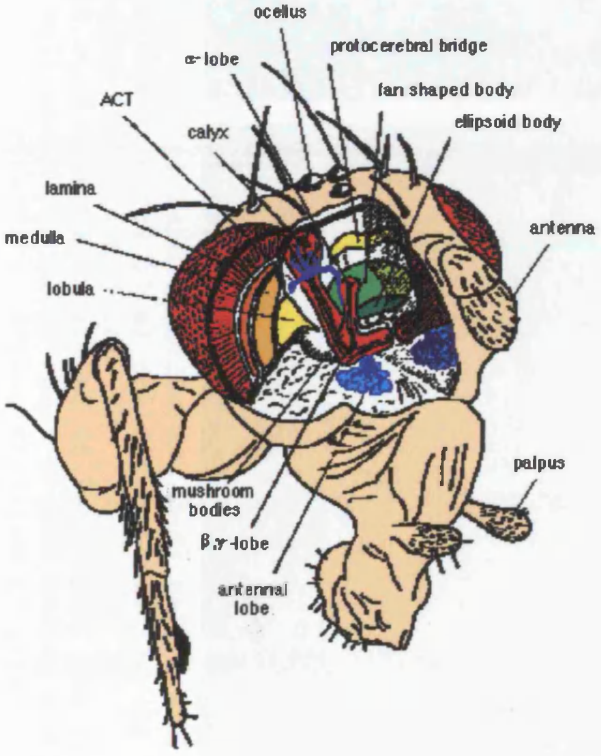


Figure A, Horizontal view through a Drosophila head, and whole head representation.

The main areas of activity for circadian cell lies at the photoreceptors (KA) and the lamina (LA), Medulla (M), Lobula plate (LA) and the CNS shown in light blue.

Figures reproduced from flybrain database.

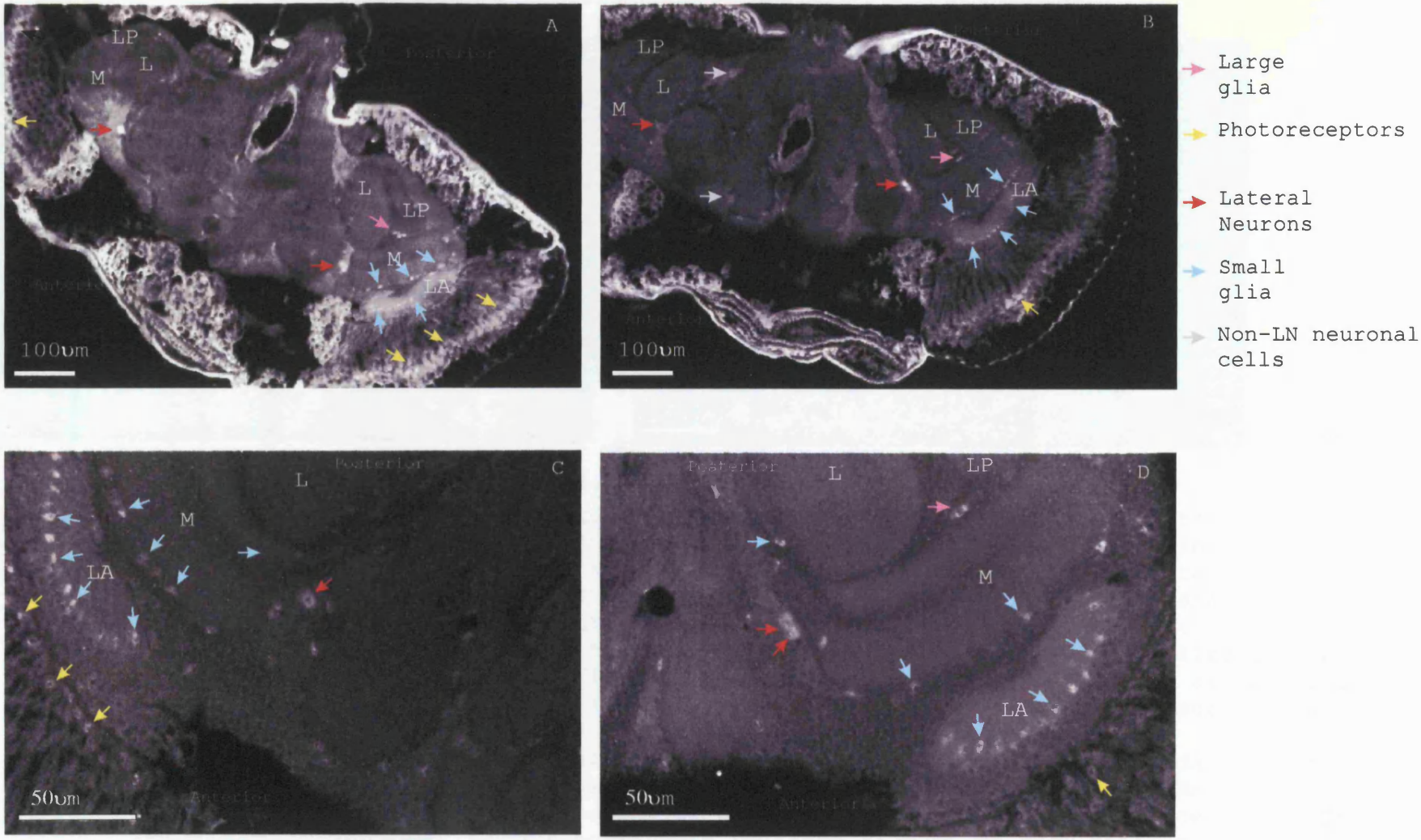
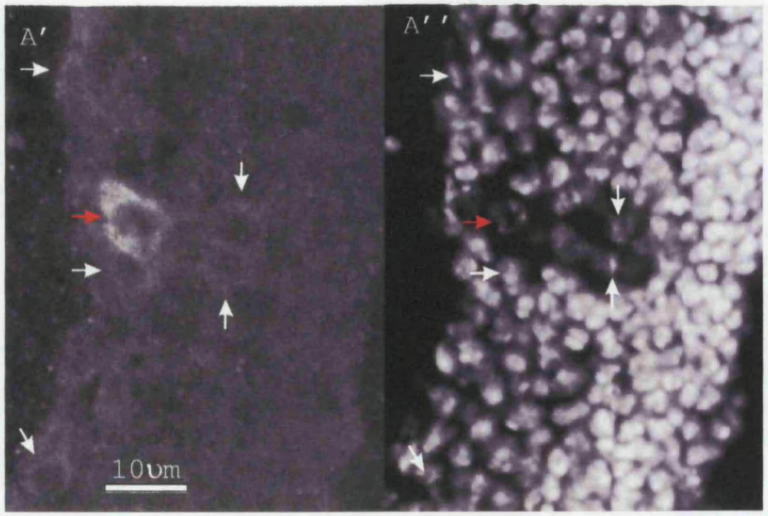
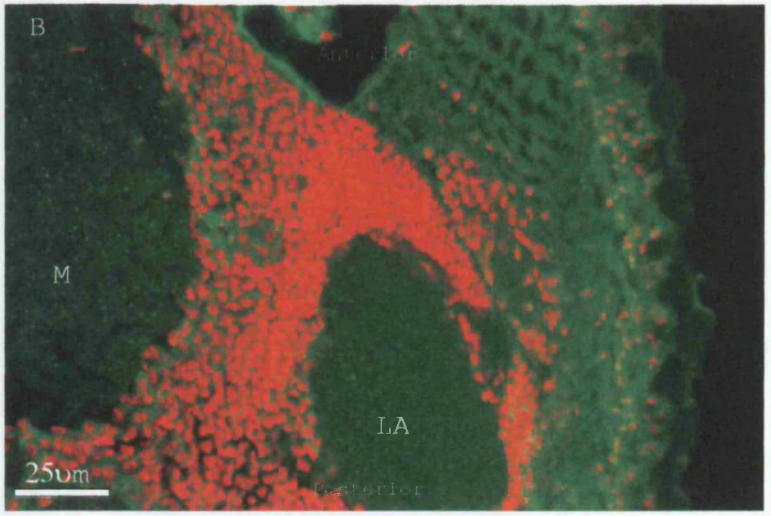


Figure 3.5 Whole brain *yw Drosophila probed with rabbit anti-dPER (Jeff Hall bleed 10-1)*
A: ZT 18, Lateral Neurons in ventral position. Large glial cellbetween lobular (L), lobular plate (LP) and medulla(M) . Glial cells around lamina (LA) and medula.
Band at tips of photoreceptors **B:** ZT 18 weak neuronal staining on left side, not from LNs, with strong ventral LN on right side ,large glial cell and small glial cells around lamina and medulla. **C: ZT 17** Weak signal in cells at the base of the photoreceptors. Strong staining in glial cells around lamina, medulla and large glial cell, also a single LN close to the lobula. **D: ZT 17** similar pattern with pair of ventral Lns.* NOTE doughnut shape common at **ZT 17** but less apparent in **ZT 18**





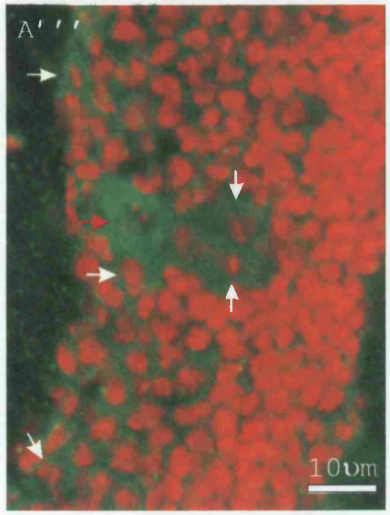


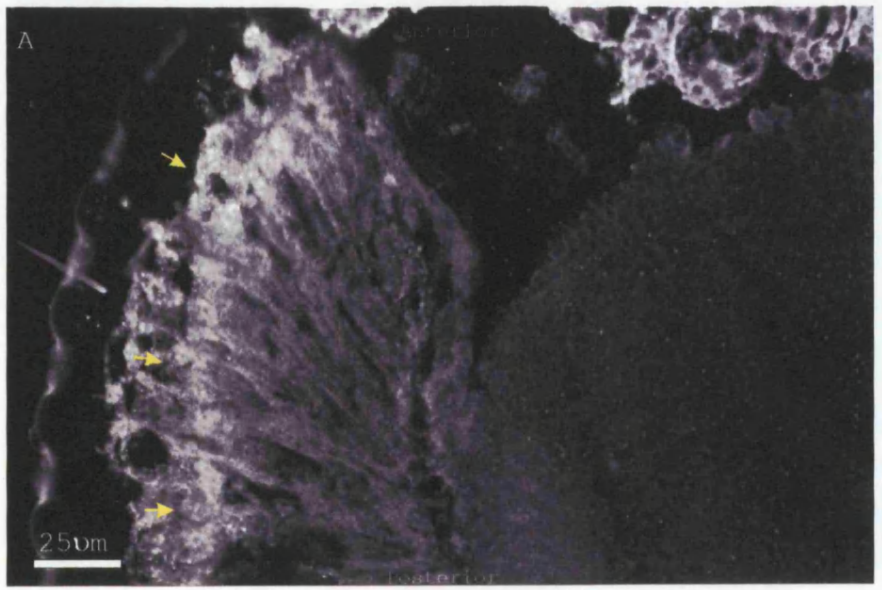
Figure 3.6 ZT 17 *yw Drosophila*

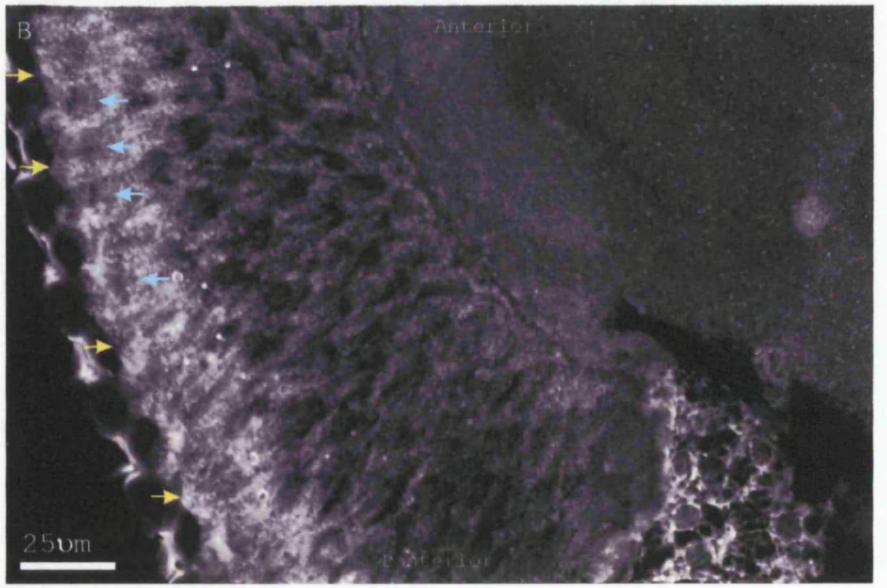
A′: Image from 1st filter, αPER-978 visualised with Cy-3. Strong staining of one of the LNs with weaker staining. The red arrows mark the strong LN which is characteristically large. The white arrows mark areas of ghost like staining close to the LN. The large size of these neurons suggest that they belong to the LN class, however a few are of smaller size and do not fit the characterisitics of the LNs. These are probabely other neuron types. All the staining occurs in doughnut patterns at ZT 17.

A′′: Image from the second filter, Propidium Iodide nucleic acid counterstain.Illustrates the Large size of the LNs with a less dense packing of neurons than in the rest of the neuropil. The arrows correspond to those from A′ to show that staining surrounds the nuclei.

A′′′: A combined image of A′ and A′′ with the same cells marked. The PER stain surrounds the nuclei. Nuclear staining would result in a yellow colour (red +green).

B: lower magnification image of the same area. The LN lies between the medulla (M) of and lamina (LA).





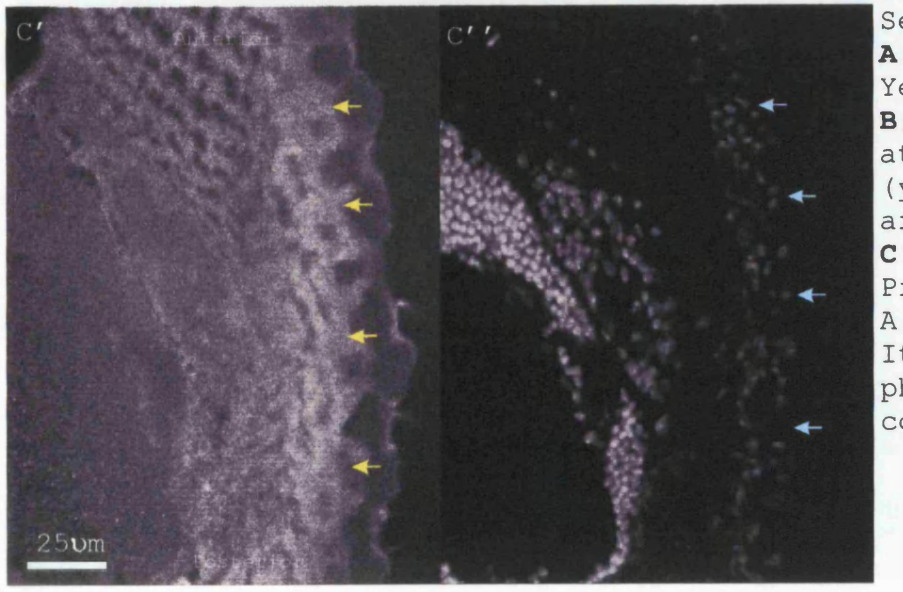
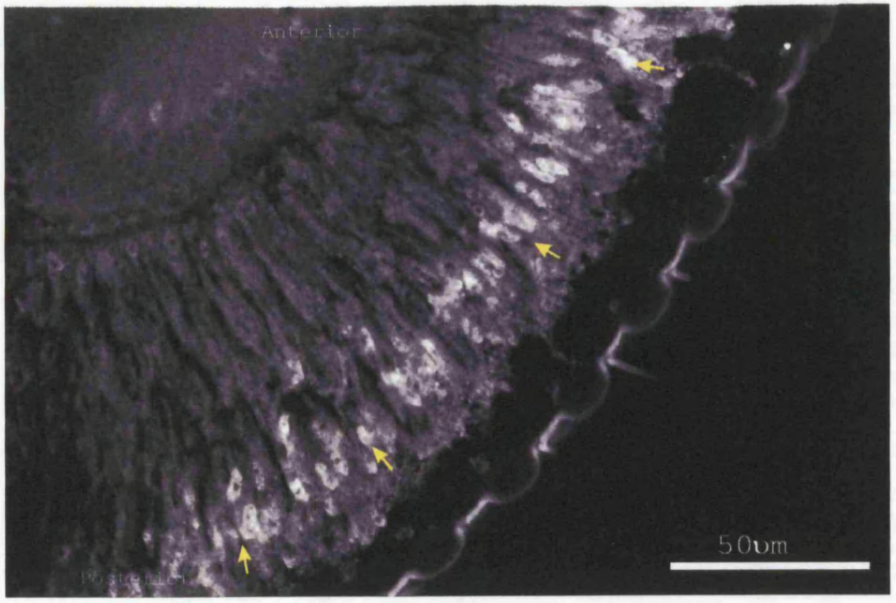
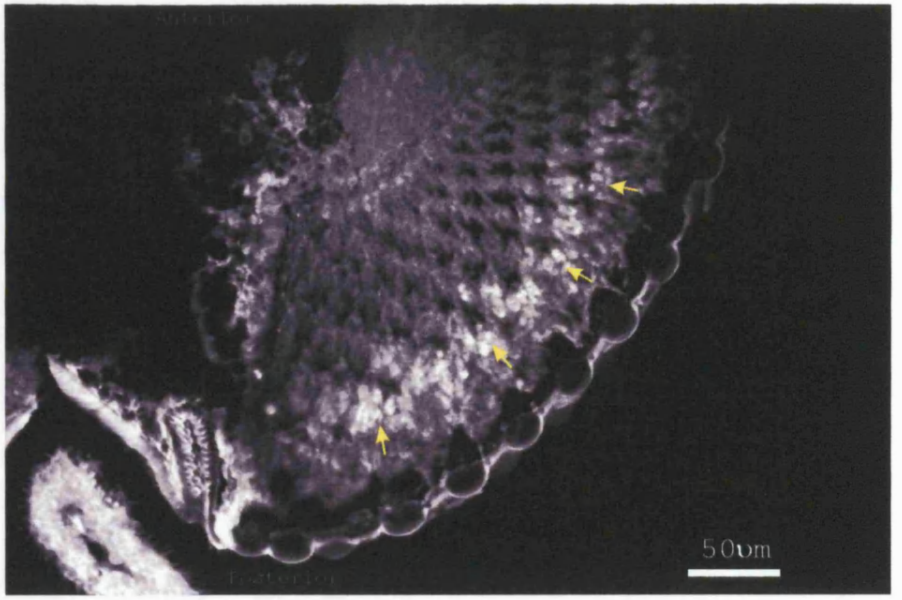


Figure 3.7 ZT 17 *yw Drosophila*. αPER-978, Cy-3. Selection of images from the Photoreceptors.
A: Photoreceptors. A diffuse band of PER marked with Yellow arrows.
B: Photoreceptors. Unusually well preserved structure at tip of the photoreceptors, shows a band of staining (yellow arrows)centering around darker circles (blue arrows) which are likely to be photoreceptor nuclei.
C: Dual image form both filters.**C'** Anti-PER, **C''** Propidium Iodide counterstain of nuclie (blue arrows) A diffuse band of staining is marked (yellow arrows). It occurs in same region as nuclei at the tip of the photoreceptors illustrating that the two patterns co-localise.





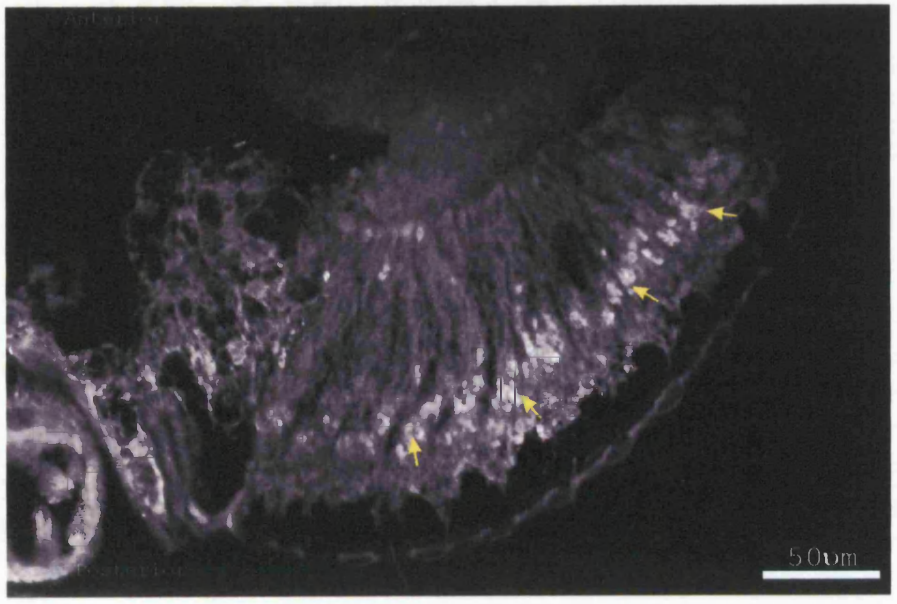


Figure 3.8 *yw Drosophila* **ZT 20** αPER Cy-3.
Selection of photoreceptor images.
Yellow arrows mark strong punctate staining at the tips of the photoreceptors. Although there is no nuclear counterstain the pattern clearly shows nuclear location of PER in the photoreceptors at ZT 20.

(**Fig. 3.5**). Cytoplasmic locations where observed for PER during the early part of the night (ZT 16-17) with photoreceptor staining showing a diffuse cytoplasmic pattern (**Fig. 3.6**). The LNs also showed strong doughnut staining (**Fig. 3.7**), which was confirmed as cytoplasmic by propidium iodide counterstaining of the nucleus. In the late night (ZT 18-24) staining of the photoreceptors showed a punctate nuclear pattern without the need for co-localisation with propidium iodide (**Fig. 3.8**). The subcellular localisation of the PER is best described in the LNs by propidium iodide counter stain, which under the confocal microscope allows a confident assignment of subcellular location.

***Musca* Labelling:**

The *Musca domestica* brain is much larger than the *D. melanogaster* head and had to be treated differently. It was dissected out of the head cuticle in ice chilled fixative. The head was then treated with a cryopreservative and sectioned in a cryostat and subject to ICC. **Figure 3.9**, illustrates the staining pattern achieved with αPER-978, the antibody used on *D. melanogaster* sections and *Musca* westerns. The cell locations were identified by comparison to the *Drosophila* brain maps (**figure A**) which compare well to maps of the *Musca* brain in the Atlas of an insect brain by Nicholas Strausfeld (1942). The staining pattern from *Musca* sections was very weak in comparison with *D. melanogaster*. Indeed background florescence in *D. melanogaster* gives a signal on which specific staining is superimposed. However in *Musca* there is only very slight background, which appears to be an intrinsic property of the two species tissues as the background levels can still be observed when no antibody has been applied. *Musca* staining had to be observed with image enhancement and the most noticeable feature of *Musca*

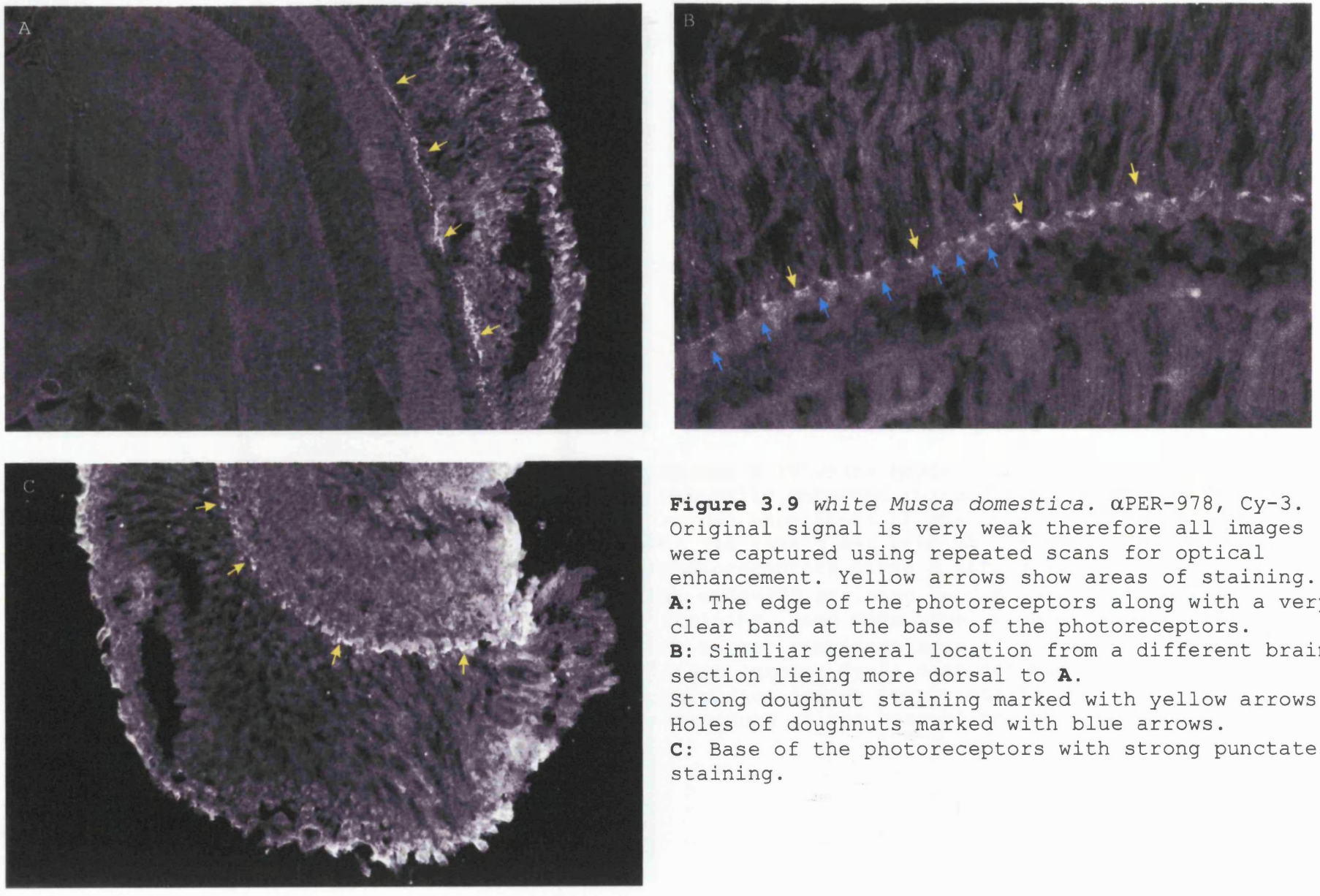
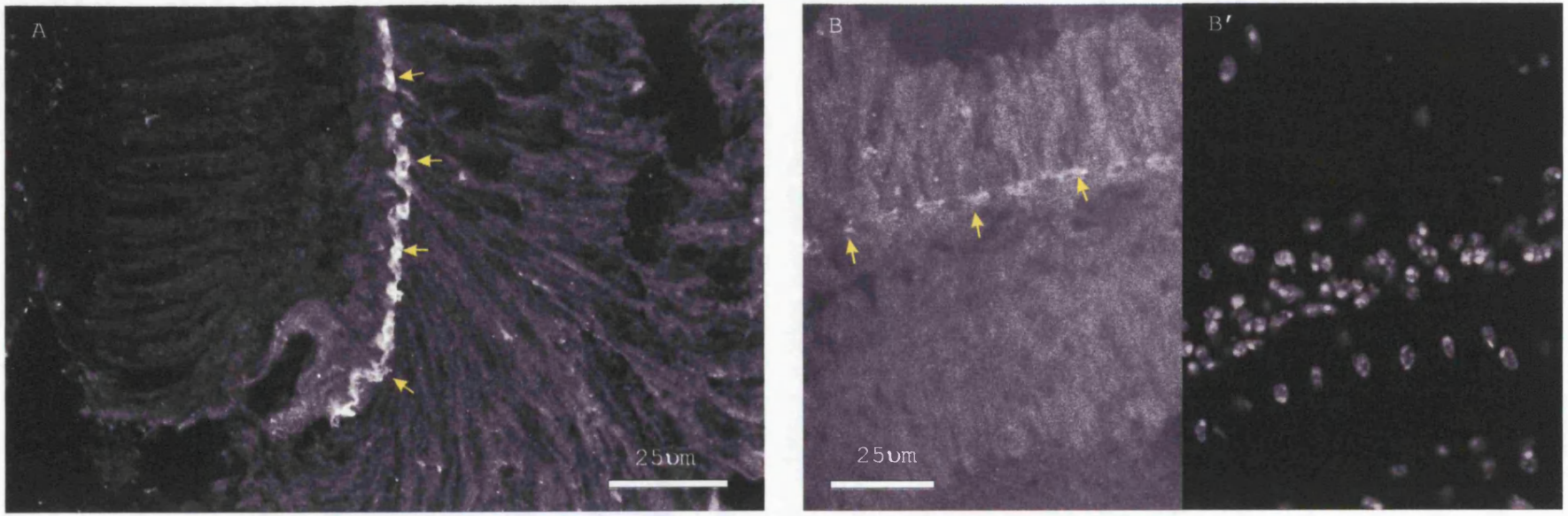


Figure 3.9 *white Musca domestica.* αPER-978, Cy-3. Original signal is very weak therefore all images were captured using repeated scans for optical enhancement. Yellow arrows show areas of staining.
A: The edge of the photoreceptors along with a very clear band at the base of the photoreceptors.
B: Similiar general location from a different brain section lieing more dorsal to **A**.
Strong doughnut staining marked with yellow arrows. Holes of doughnuts marked with blue arrows.
C: Base of the photoreceptors with strong punctate staining.



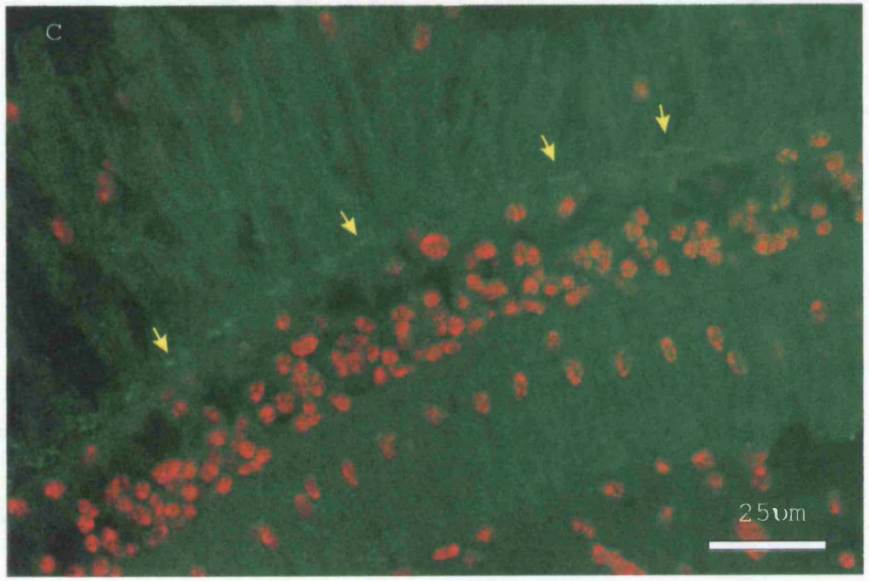


Figure 3.10 *white Musca domestica*. αPER-978, Cy-3. Optically enhanced through repeated scans. Yellow arrows show areas of staining.
A: The same clear band at the base of the photoreceptors gives a strong signal.
B: αPER-978 staining band **B'** propidium iodide staining, band does label around nuclei.
C: Larger combined image of B. The band of αPER-978 stain (green, Cy-3) does not map onto nuclie (red, propidium iodide counterstain) but maps onto a structure at the base of the photoreceptors.

was a band of staining at the base of the photoreceptors and more occasionally at the tips of the photoreceptors when they had not been removed during the dissection. The pattern appears to be cytoplasmic with doughnut like staining. **Figure 3.10**, Illustrates that this candidate area of staining does not focus on the nuclei, but instead probably occurs at the base of the photoreceptor cells. The propidium iodide counter stain showed that the cellular nuclei are located away from the staining pattern.

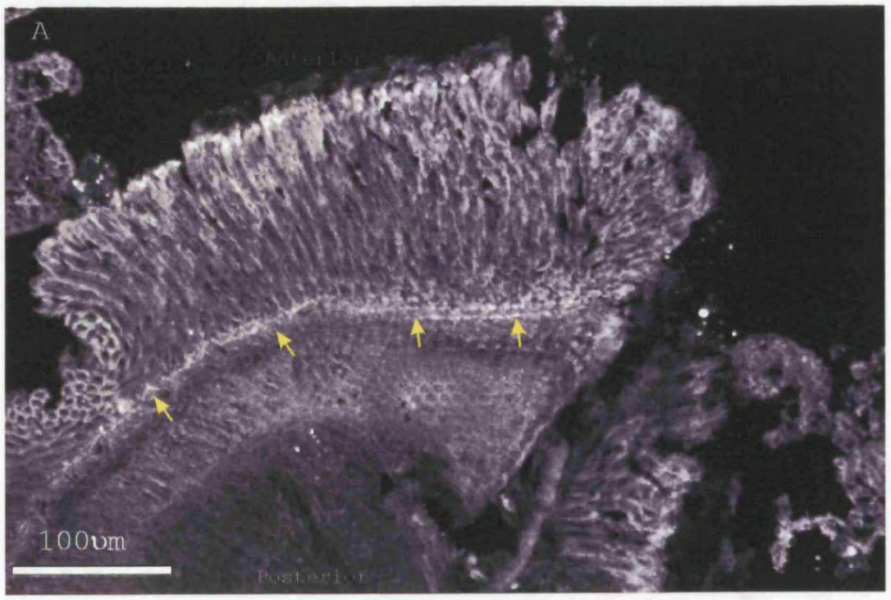
Alberto Piccin and Kathy Siwicki had shown that two monoclonal antibodies (two anti *Munduca/Drosophila*PER chimeras termed αPER-KS1 and αPER-KS2) gave robust labelling on *Musca* sections. Although I repeated this result during a visit to Kathy Siwicki's laboratory I attempted, unsuccessfully, to repeat this result with different batches of the same antibody in Leicester. There is no obvious explanation of why this should be so as detection worked well with the positive control (αPER-978 on *D. melanogaster*) and more than one anti-mouse secondary was used for detection of the monoclonal lines. In order to try to reproduce the initial result the antibodies were concentrated by a factor of ten. Microcon™ centrifuge columns (10000 MW cut off) were used (see methods section). The columns were blocked with BSA, to avoid antibody binding to the filter, and spun. The monoclonal antibody was then applied directly to sections as detailed (see methods section) but the staining pattern achieved was much weaker than that achieved previously in the USA. Monoclonal lines should express the same antibody throughout their lifetime, as they are cell lines produced from one clonal cell and should therefore produce one antibody only. However given that the monoclonal antibodies from Kathy Siwicki's laboratory had been shown to change activity over time (Jocelyne Noveral, pers. comm.) the lack of reproducibility was

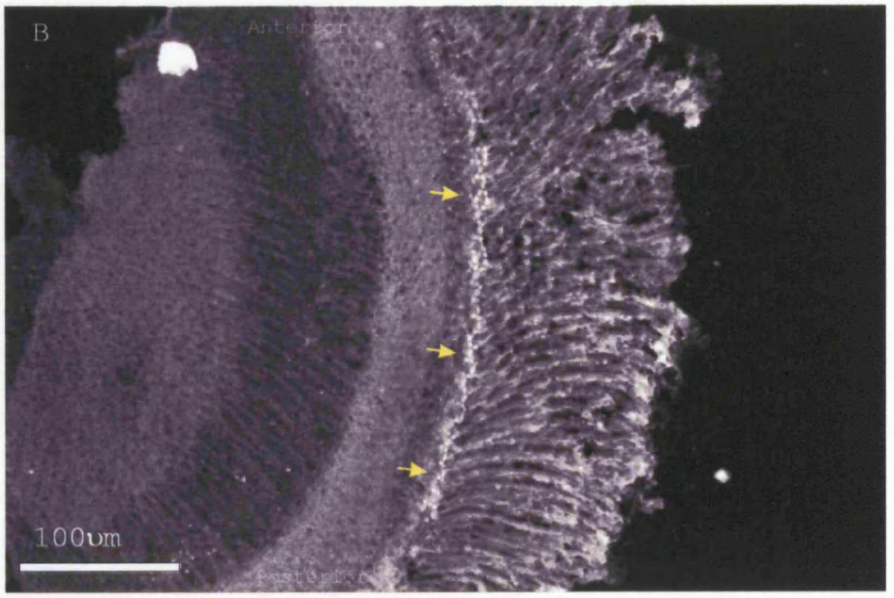
probably due to poor maintenance of the cell lines. Unfortunately although these monoclonal cell lines gave promising results they had to be discarded, as the initial results could not be repeated in Leicester by myself or in the USA by Kathy Siwicki (pers. comm.., K Siwicki).

Figure 3.11, shows the results from the concentrated monoclonal antibodies in *Musca*. They produced the same strong band of staining as achieved in **Figure 3.9** and **3.10**. However the pattern is much more punctate (as demonstrated by the solid balls of staining marked by yellow arrows in **Figure 3.11A** and **B)** suggesting a nuclear localisation. Again there is strong staining at the tips of the photoreceptors.

Figure 3.12 shows selected images from sections of a *Musca* whole brain. The immunocytochemistry procedure used was repeated exactly as was previously reported except no primary or secondary antibodies were applied, the images were then captured as detailed in the methods. This therefore provides a control that demonstrates the autoflourescent signal from the sample.

There are identical patterns to those achieved with the polyclonal αPER-978 and monoclonal KS1 and KS2 antibodies. Therefore these patterns can be discounted as representing specific PER staining. This pattern was not noticed in previous negative controls where individual sections from a series were used. Only through sectioning a whole head and analysing all the areas of the brain can a proper negative control be achieved. The only specific pattern of *Musca* staining was observed in the cytoplasmic staining of *Musca* neurons near the central brain (**Figure 3.11C**). However the pattern was very weak and was not easily repeated. Given that this was from a concentrated antibody, that was in short supply, it did not provide a useful reagent for subsequent use.





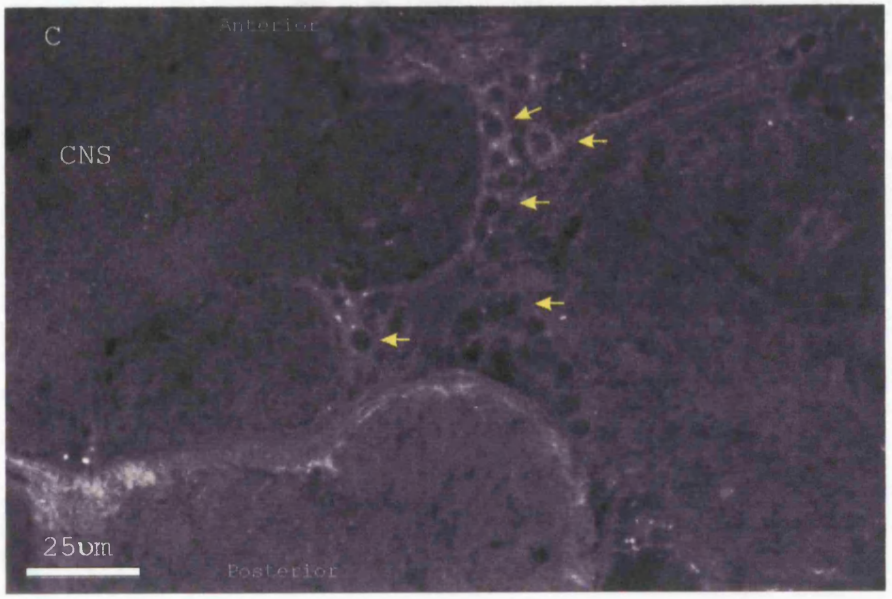


Figure 3.11 *white Musca domestica*. αPER-KS1. Optically enhanced through repeated scans. Yellows arrows show areas of staining
A: Strong band at base of photoreceptors much like αPER-978.
B: Photoreceptor band in sections from a different head.
C: αPER-KS2 concentrated by microcon filtration. Staining between the optic lobe and central brain

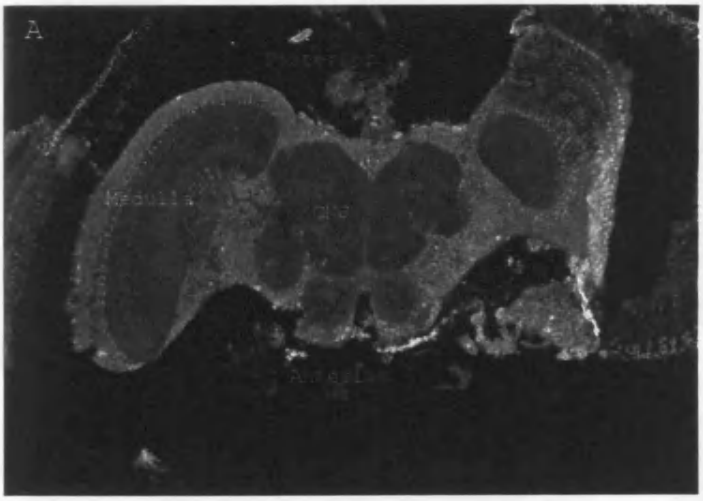
Further trials were performed using more anti-PER and TIM antibodies but none produced a convincing pattern, table 3.1 and **figure 3.13** and **3.14**.

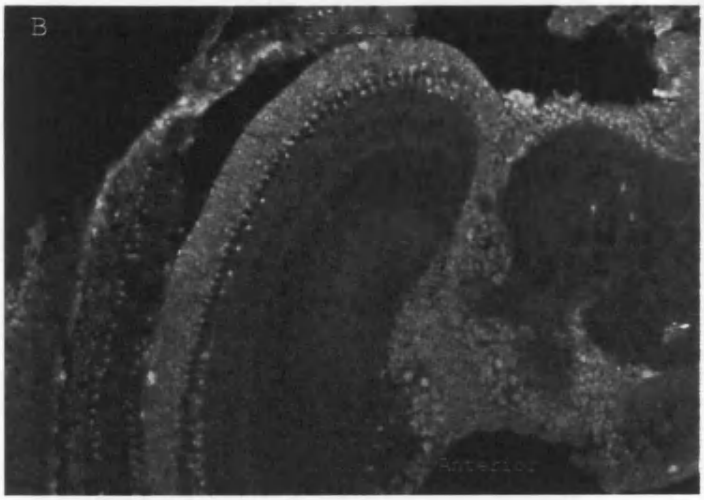
Discussion.

The unusual behaviour of *Musca* PER during western analysis required that cellular analysis of *Md*PER be performed. However no definite anti-PER staining could be reproduced in *Musca*. The αPER-978 antibody failed to produce any unambiguous staining in *Musca*, although a pattern of labelling was produced with some interesting features that could be interpreted as cytoplasmic. However this pattern was repeated with the negative control. **Figure 3.12** demonstrates the pitfalls of using staining patterns alone as a basis for interpretation of the results. Only by careful examination of anatomical features and nuclear counter staining can the validity of a pattern be assessed. It is frustrating and perplexing as to why αPER-978 antibody did not work on *Musca* sections when it worked on westerns. The only plausible explanation is that the polyclonal antibody recognises an epitope that is masked when the protein is folded, as in sections, but is exposed when unfolded as in SDS-PAGE westerns. A similar phenomenon was noted to occur with the first antibodies raised against *d*PER. The Kathy-Siwicki-monoclonal raised against peptides centered on the per^{s} site recognised PER on sections but had no specific reaction on westerns.

Given the lack of an antibody that recognises *Md*PER on sections it was therefore necessary to produce an antibody that would work on sections. The manufacture of *Musca* antibodies is discussed in chapter 4.

Even assuming that *Md*PER enters the nucleus to fulfil its normal functions the *Musca* clock may operate in quite a different fashion to the *D. melanogaster*





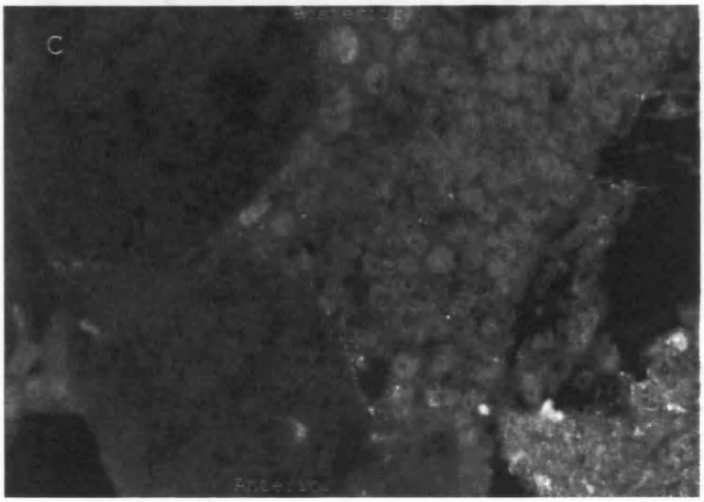


Figure 3.13 αTIM-316, ZT20.
A Low power magnification.
B medium power magnification.
C high power image magnification.
The antibody labels all neuronal cell bodies. Labelling appears to have a cytoplasmic localisation under high power magnification.

clock as *Md*PER is not subject to the daily degradation cycle. In *D. melanogaster* phosphorylation of PER by DBT provided the cue for degradation and *dbt* mutants show fast or slow circadian phenotypes (Price et al., 1998). Therefore the phosphorylation of PER is the key to regulating its degradation. When no phosphorylation of PER occurs, as in the *dbt* null mutant, dbt^{P}, then PER accumulation is no longer dependent on TIM (Price et al., 1998), PER is no longer degraded and can accumulate to high levels. A series of mathematical models point to degradation being of paramount importance in adjusting the length of cycle. Using the Goodwin oscillator to model the molecular mechanisms of the clock it was shown that mRNA and protein turnover was important in adjusting the cycle length (Ruoff et al., 1997; Ruoff et al., 1999). Further to this two more models elaborate and included phosphorylation targeting PER for degradation. It is concluded that the phosphorylation of PER and its subsequent degradation is partly responsible for the time lag between translation and negative feedback (Goldbeter, 1995; Leloup and Goldbeter, 1998), without which the clock would stop oscillating and return to a steady situation.

The bare essentials of the clock is the time delay between transcription and negative feedback (Lema et al., 2000). It is clear from both the models and empirical evidence from the *dbt* mutants that phosphorylation and degradation are also key components in *D. melanogaster*. In addition phosphorylation has been shown to be involved in circadian clocks of mammals (Keesler et al., 2000 ; Lowrey et al., 2000), *Neurospora* (Liu et al., 2000) and bacteria (Iwasaki, et al 2000). However phosphorylation changes in mobility of *Musca* PER has not been observed and it appears hypophosphorylated; treatment to remove phosphorylation has little effect on the mobility of the protein extracts (Piccin, 1998). In contrast *D.*

melanogaster PER undergoes an easily observed change in mobility during the day due to phosphorylation (Edery et al., 1994). Therefore how the *Musca* clock maintains time is intriguing as there appears to be little or no phosphorylation of *Md*PER (Piccin, 1998). *Musca* therefore appears to remove one of the important mechanisms for achieving the lag between transcription and translation of PER. This property appears not to be an intrinsic property of *Musca* PER; *D. melanogaster* per^0 mutants rescued by a *Musca per* transgene were rhythmic. Furthermore westerns analysis shows that the protein cycles in these transgenic flies and there is some evidence of phosphorylation modification One would predict that if the *Md*PER protein was instead incapable of binding DBT in the *D. melanogaster* per^0 transformant then it should have an arrhythmic phenotype and PER should have a similar profile as the dbt^P mutant. Rescue of per^0 with chimeric *per* transgenes with the *D. melonagaster* N-terminus and the *Musca* C-terminus provided behavioural rescue with protein cycles that lagged the RNA in a similar manner to wildtype *D. melanogaster (Piccin, 1998)*. Therefore it is likely that DBT has normal function on *Md*PER in these transformants, so the N-terminus sequences of Musca affect PER dynamics. It is expected that *MdP*ER must be capable of binding TIM and moving into the nucleus for negative repression of *per* and *tim* transcription in transformants.

Chapter 4:

Raising anti-sera against *Musca* PER

Introduction:

To increase the likelihood of raising an antibody that works on sections it is necessary to consider why the battery of antibodies tested fail to recognise *Musca* PER. Antibody response to a inoculated protein has the potential to produce individual antibodies that recognise protein sequences across the whole length of the presented fragment. However in reality this rarely happens. Instead the ability to elicit an antibody immune response varies along the length of the protein and between proteins. Some proteins provide excellent antigens that elicit a strong response whereas others produce little or no response at all. It is unfortunate that PER appears to fall into the latter category. The first attempts to produce antibodies against *Drosophila* PER failed to produce usable sera (Kathy Siwicki, pers. comm.).

The antibodies used against *Musca* have several failings. Jeff Hall's and Kathy Siwicki's antibodies were raised against PER fragments from other species and strongly label PER from these species. However antibodies recognise short sequences and it is possible that these sequences were not conserved in *Musca*. Albertos Piccin's antibodies were raised against a short *Musca* PER fragment. This however failed to produce a strong serum. This may be due to only inoculating a small part of the protein; parts of the protein that may have provide good epitopes may not have been present in the small fragment. In addition the inoculation protocol did not include the use of an adjuvant. Adjuvants act to boost the general immune system and increase the response to presented antigens. Given that PER has a history of eliciting a weak immune response this was probably a major draw back

Aim:

To produce a strong, specific sera to *Musca* PER that work on tissue sections and westerns.

Approach:

The approach used to generate antibodies against *Musca* relied upon inoculating with full length *Musca* PER in conjunction with a strong adjuvant. This approach presented all of the epitopes of *Musca* PER along with a strong boost to the immune response through the adjuvant. Rabbit was chosen as the host because they provide large amounts of polyclonal sera. Polyclonal sera are most likely to be successful because they potentially contain antibodies that are specific to many sequence stretches of the presented fragment. One reason for failure can be that the epitopes that a particular antibody recognises can be masked. Epitopes that lay in the protein core can be inaccessible because the outer parts of the protein stericaly interfere with antibody binding. This is a particular problem when working with sections because no denaturant is used and so the protein is presented in its native state. Polyclonals with their avidity to many epitopes are therefore more likely to be successful. Indeed it is the case for Jeff Hall's αPER antibody that it must recognise some stretches of sequences for *Musca* PER for there to be a specific affinity for *Musca* PER on western blots. However, the failure of this antibody on sections is likely to result form the masking of these specific sequences which happens when the protein is in its native conformation in tissue sections. When possible, soluble proteins were used to inoculate; soluble proteins are likely to be correctly folded into their native states. Although the antigen presenting cells

degrade the inoculated proteins and present small fragments on the cell surface, the resultant antibodies can still be specific for native structures in preference to denatured proteins. A phenomenon that is well documented (Harlow and Lane 1988).

Production of *Musca* PER:

A bacterial expression system was chosen because it was cheap, quick and easy to use. This required the cloning of cDNAs into the expression plasmid. There were no full length cDNA clones available;, the sequencing of the *Musca* PER and manufacture of the mm1 constructs thus relying on genomic DNA clones. Therefore cDNA clones had to be produced. Polyadenalated RNA was isolated as described in methods section. It was then reverse transcribed to produce single strand of DNA which was then used as a template for PCR. Primer design and optimisation is detailed in chapter 5 along with design of the cloning regime, expression system and protein purification.

The resultant six protein fragments covered the whole length of *Musca* PER, as shown in figure 4.1, with each protein having a N-terminal polyhistidine tag, detailed in figure 5.5 (chapter 5), which aided purification. This tag was known to be a poor epiptope and was not removed before inoculation of the proteins.

Inoculation procedure:

Two male Newzealand White rabbits. 1.8-2kg, where inoculated every 4 weeks with 0.8ml of 1:1, titremax adjuvant: protein in PBS. A total of 125 μg of protein was made up to 0.5mls in PBS and mixed 1:1 with Tiremax Gold^{TM} to form a complete emulsion. 0.8mls were injected per rabbit, per boost. The concentration

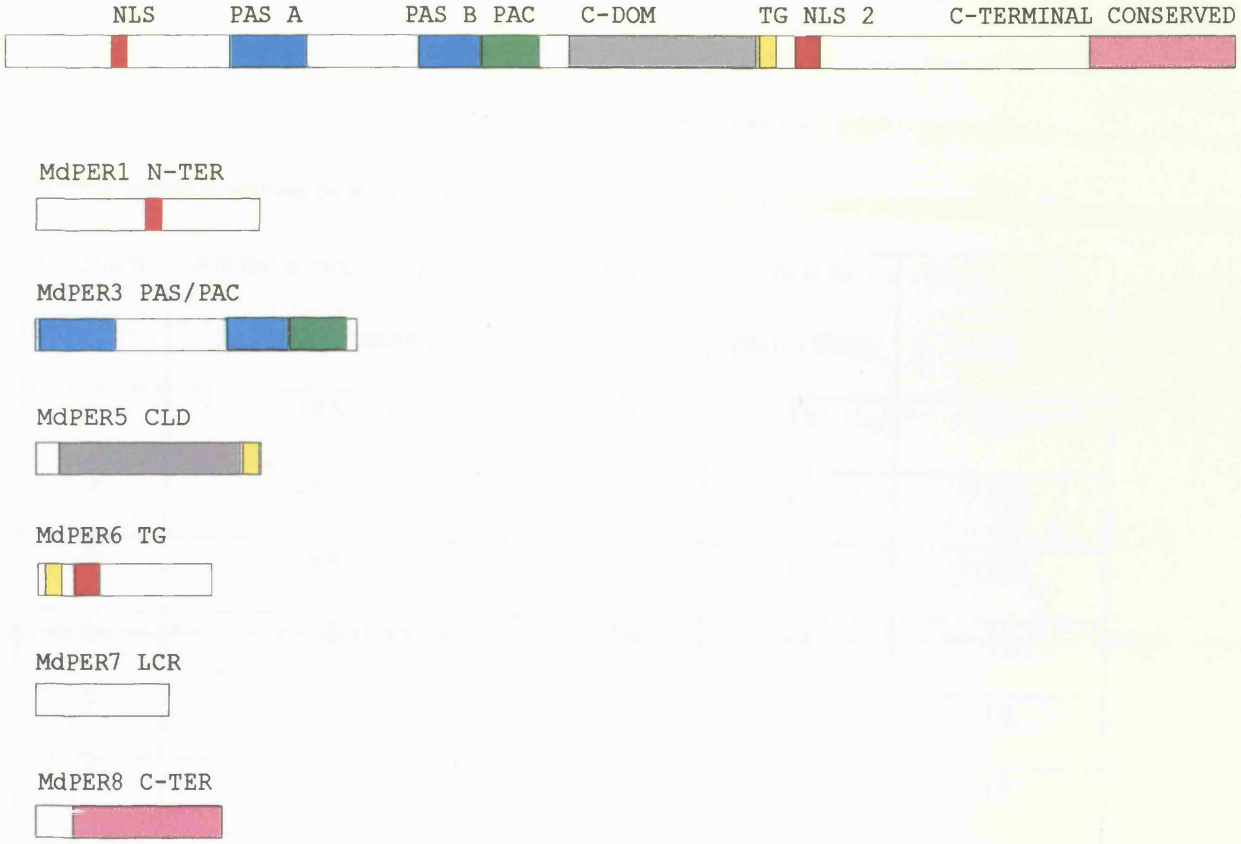


Figure 4.1, Schematic representation of Musca PER with functional/comserved domains marked and the fragments used during inoculation.

of each fragment was determined with the Bradford assay (see methods section). All the fragments were then mixed, keeping the molar ratio of each fragment equal (see table 4.1) and the volume adjusted to 0.5mls in PBS. Molar ratio was calculated by dividing amino acids in the fragment by total for the *Musca* PER; giving a usable approximation of molar ratio.

Fragment	Amino acids in fragment / total in protein.	ratio	Total in each 100ug	Total in 125 ug
1	195/1080	.18	18	23
3	283/1080	.26	26	32
5	198/1080	.18	18	23
6	152/1080	.14	14	18
7	118/1080	.11	11	14
8	164/1080	.15	15	19

Table 4.1

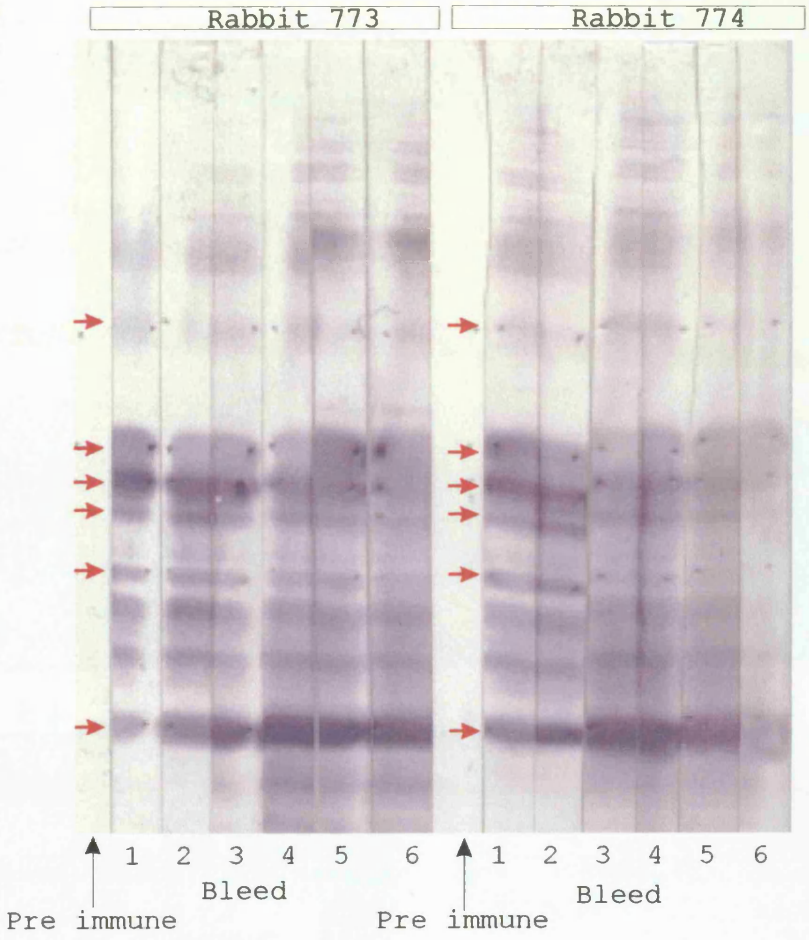
Bleeding procedure:

Pre immune bleeds were taken prior to the first inoculation. The rabbits where then serially inoculated, once every four weeks, with the same Titremax Gold^{TM}, protein mix. The rabbits were bled two weeks after every inoculation and sacrificed, and exsanguinated, two weeks after the fifth inoculation. The blood was spun to separate serum, aliquoted and stored at -20^{O}C prior to testing.

Testing the antibody.

The antibody was tested for avidity before application to sections. Purified fragments were mixed and applied to a SDS-PAGE gel, and then subject to western analysis. The blot was stained with Ponceaus S protein stain and the positions of each fragment marked. The blot was cut into strips and each strip used in western analysis with a different bleed of the primary antibodies and then developed with an enzymatic colour reaction.

Figure 4.2 shows that the inoculation produced a strong reaction to all of the protein fragments that were visible by Ponceaus S stain. Comparison of the pre-immune sera to test bleed shows that all of the visualised bands were due to the inoculation. Bands that developed in areas that did not co-localise with the Ponceaus S stain were either due to contaminating bacterial protein or low concentration *Musca* PER fragments. Degradation products or multimers of the fragments could produce the multitude of bands visualised. The combination of fragments covering the full length of the protein and Titremax^{TM} adjuvant had illicted a strong, and specific, immune response to *Musca* PER bacterial fragments. Although the inoculation had produced a strong reaction to *Musca* PER this still does not guarantee a useful antibody. Therefore the *sera* still have to be assessed for avidity to *Musca* PER from tissue extracts. Factors such as concentration, protein folding and post-translational modification can prevent antibody binding. Tissue extracts from 2 hourly collections (from ZT 0-22) were equalised to 2μg/ul, mixed and loaded onto a SDS-PAGE gel and blotted. The blot was cut into strips and these strips were subjected to western analysis, with varying concentrations of primary antibody, and then developed with ECL (Figure 4.3). Lane 1 consists of rabbit anti PER from Rosbash's lab. The antibody produced a high background, however



Figuire 4.2 antibody testing of polyclonal 773 and 774 against the bacterial expressed and purified fragments. Fragments were mixed, loaded on to a SDS-PAGE gel then blotted. The blot was developed with Ponceaus S total protein stain. The bands were then marked on the blot with pencil and the blot cut into strips. Each strip was subject to western anlaysis with either pre immune sera or post innoculation bleeds (numbered 1-6). Red arrows mark the Ponceaus S bands.

The Ponceaus S bands colocalisation with the bands produced by western analysis. Illustrating that the antibodies reacted against the inoculated proteins in both rabbits. Also the antibodies react against all of the bands visible by Ponceaus S staining illustrateing a broad spectrum reaction to many *Md*PER epitopes.

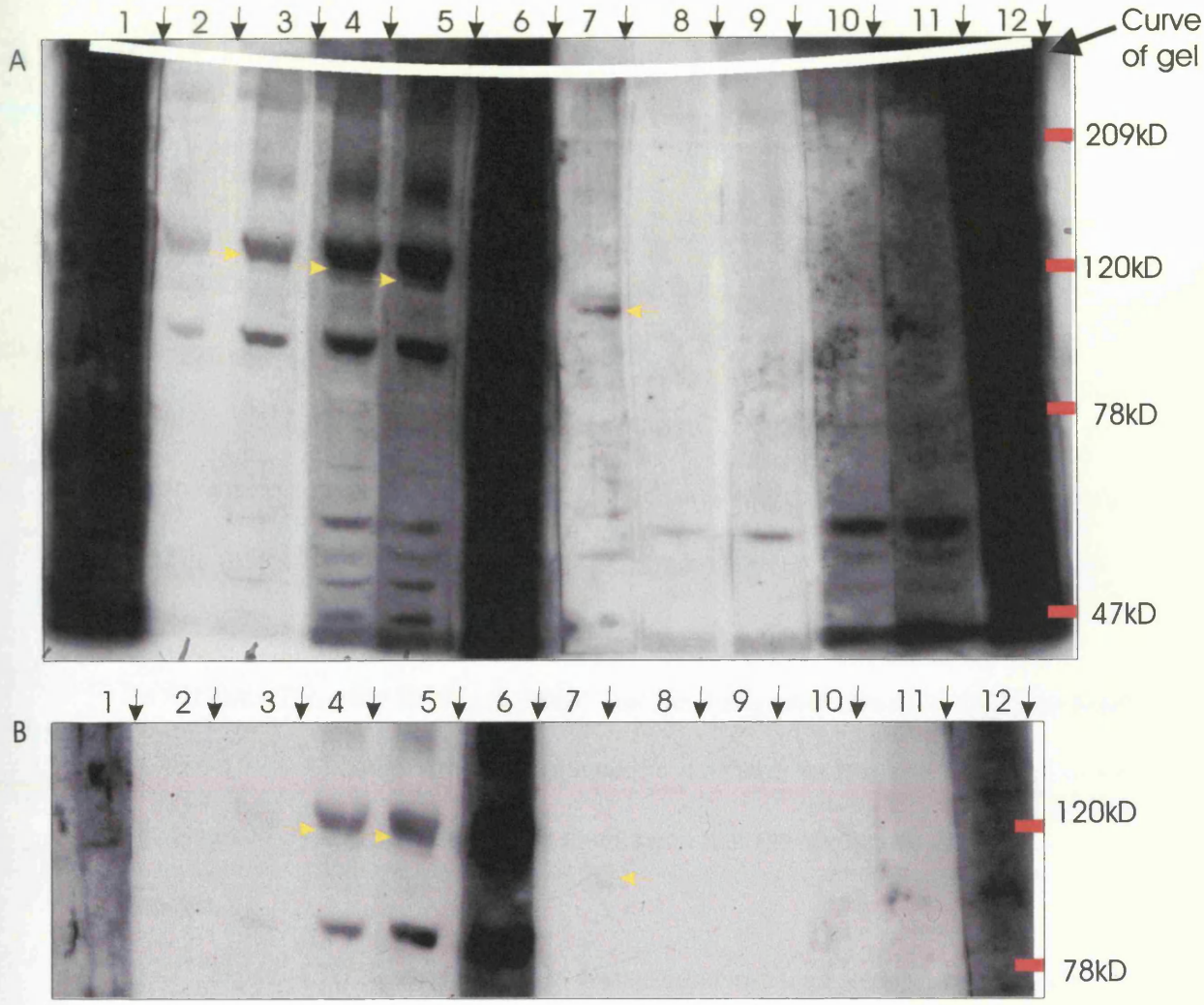


Figure 4.3 Antibody testing.

Musca extracts, 2 hour samples, Zt 0-22, were combined and loaded on SDS PAGE. Each lane, which contained an identical loading, was visualised with Ponceus S stain and cut into strips, each containing one lane. The strips were then exposed to different antibody regimes as detialed bellow. Size markers shown in red.

A: 1, Rosbash Rabbit PER, Lanes 2-6 Test bleeds, Rabbit 774: 1/10000, 1/5000, 1/1000, 1/500, 1/100 respectively. 7, Jeff Hall Rabbit, bleed 12-1, 3rd use.. Lanes 8-12,Test bleeds, Rabbit 773: 1/10000, 1/5000,1/1000, 1/500, 1/100 respectively..

B: weaker exposure of A.

subsequent weaker exposures did not demonstrate any specific bands that occur above background. Lanes 2-5 consist of increasing dilutions of rabbit 774 serum. The main band that developed was marked with yellow arrows. This band migrated at the same relative position as reported for *Musca* PER (Piccin 1998), and provided a strong candidate for a specific PER reaction under western analysis. Dilutions of 1/10000-1/500 appear effective. The background of this antibody was very low with a strong ratio between the main reacting band and non-specific background. There are two minor but strong bands flanking the main band, which could also provide candidates for PER specific bands. However Lane 7, developed with Jeff Hall's Rabbit anti-PER, only labels one band. This band migrated closely to the main band of Rabbit 774. Although the band appeared to run faster on the blot, the curve of the blot (marked by a white band) gives the illusion that the two bands do not co-migrate.

Figure 4.3B illustrates the low background and high strength of antibody 774. The bands in 774 develop with little background, making quantification by densitometry possible. Jeff Hall's antibody failed to produce a strong band under the same exposure (band position marked by yellow arrow, Lane 7). Figure 4.4 shows a full time course western analysis of *Musca* PER with Jeff Hall's Rabbit anti-PER antibody, bleed 12-1. The PER band, marked by yellow arrows, only labelled weakly. Although there is no apparent strong cycle it would be difficult to quantify the band due to its relatively weak ratio to background. Previous studies had used Jeff Hall's Rabbit anti-PER antibody, bleed 9-1, which had produced a strong band. Unfortunaelty this bleed resource had run out in our lab. Bleed 12-1 had much less specific avidity to *Musca* PER, probably due to changes in the antibody producing cell profile in the rabbit over time. Figure 4.4 represents

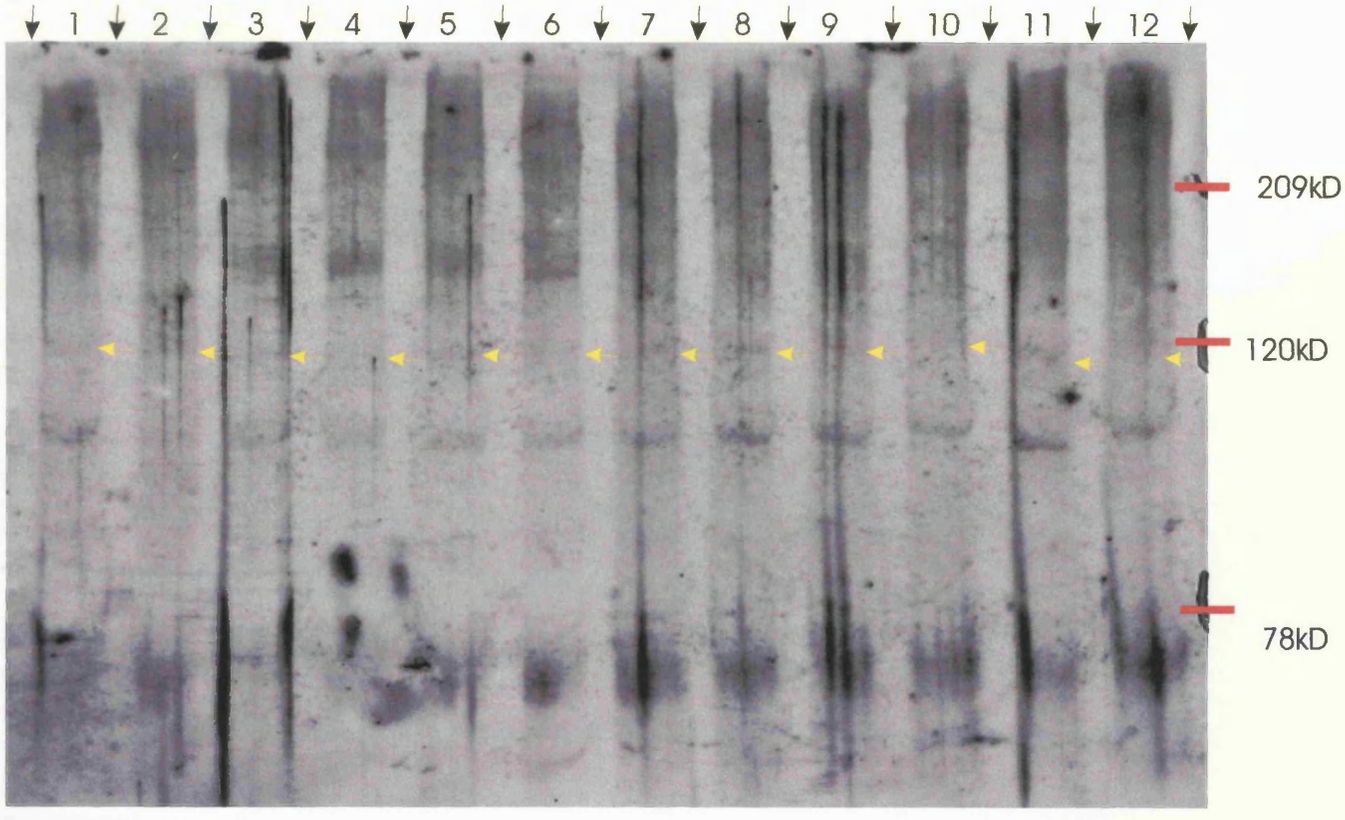


Figure 4.4. Musca, 2 hour collections, ZT 0-22. Lanes 1-12 ZT 22-0 respectively. Visualised with Jeff Hall Rabbit anti-PER, bleed 12-1, 3rd use. Yellow arrows mark PER band..

optimised exposure of Jeff Hall's Rabbit anti-PER antibody, bleed 12-1 to *Musca* PER. The antibody had been used twice previously on blots. The background was much higher during these exposures due to a higher ratio of non-specific to specific antibody binding. Repeat re-use reduces the non-specific binding. Further re-use resulted in a fainter *Musca* PER band due to absorption of PER binding antibody.

Piccin (1998) showed that on 6% SDS-PAGE gel *Md*PER migrated at around 130kDa, which is larger than 120kDa that the amino acid predicts. On 12% SDS-PAGE gels used here the same bands runs at120kDa. Therefore the αPER-774 antibody was used during western analysis on a 6% SDS-PAGE gel loaded with *Musca* head tissue extracts (figure 4.6). On 6% SDS-PAGE the band migrates at closer to 130 kDa. Therefore it would appear to be the same band found by Piccin (1998) and identified as *Md*PER by immunoprecipitation anlaysis. The *Md*PER protein therefore appears to have slightly different mobility on different strength SDS-PAGE gels.

Other control for antibody specificity that could have been included. These include the use of the *Md*PER per^0 , *D. melanogaster* transformant that contained the entire coding sequence of *Md per*. Demonstration of a specific antibody pattern that did not occur in non transformed per^0 flies would have been strong evidence of a specific *Md*PER reaction. Immunoprecipitation using the α*Md*PER antibody and subsequent coprecipitation of a αTIM reacting band would have also provided strong evidence for specific *Md*PER reaction.

Results from antibody on sections:

The αPER-774 antibody was then used to determine the subcellular localisation of *Md*PER in wildtype flies. Immunocytochemistry was performed on

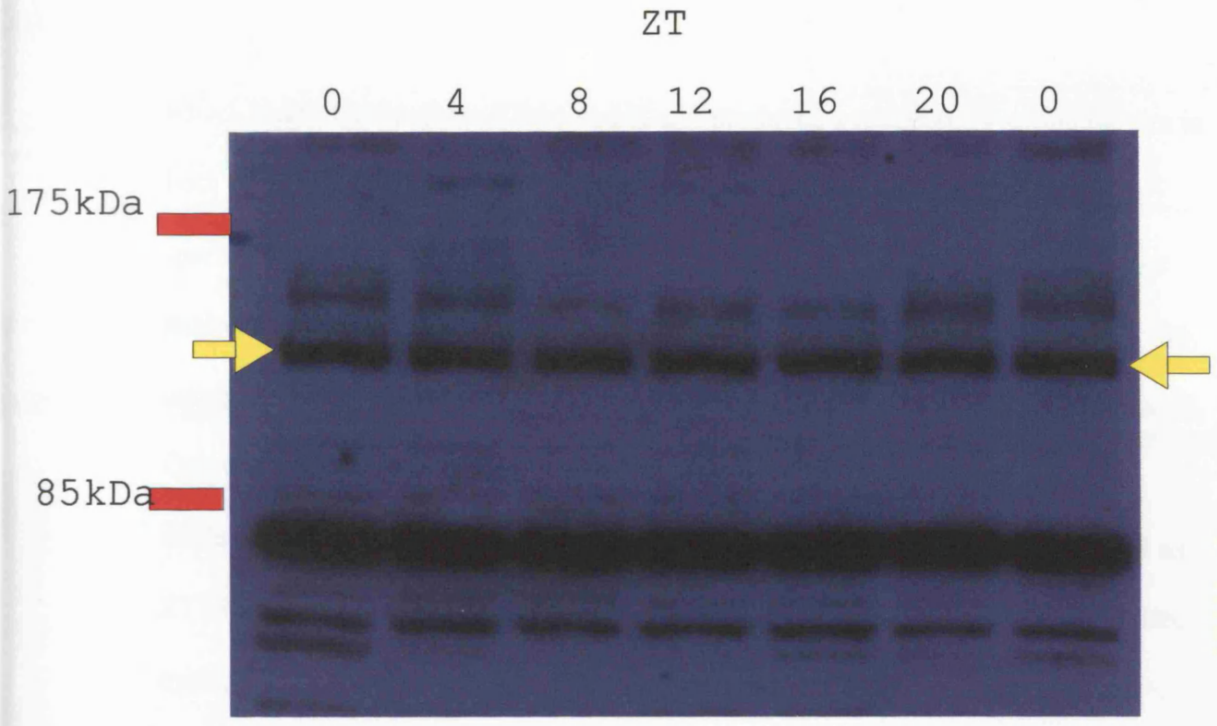


Figure 4.6. 6% SDS-PAGE analysis of αPER-774 on MdPER tissue extracts. The *Md*PER specific band is marked with a yellow arrow and molecular weight markers shown in red.

whole head sections form ZT 16 and 8, and produced a moderately strong pattern in both time points (figure 4.7). The only area of the *Musca* brain that showed a specific signal was located between the optic lobe and CNS, which is in a similar anatomical position to *D. melanogasters* LNs. Furthermore a doughnut pattern was obtained in both time points, suggesting that the majority of *Md*PER is cytoplasmic. One would expect that if there was a rhythmic movement of MdPER into the nucleus then it should have been detected in the two time points analysed (ZT8 and ZT16). The time points were chosen because they come from opposing parts of the cycle, and in *D.melanogatser* there was a strikeing difference in abundance and sub cellular localisation of PER More time points could not be analysed due to the laborious and time consuming nature of immunocytochemistry. However these results certainly suggests that the only *Md*PER expressing cells in *Musca* heads are located in positions that make them candidates for LNs. Further more there was no gross cycle in *Md*PERs abundance or its cellular or subcellular localisation. Therefore it seem likely that western analysis of *Musca* that shows a non cycleing *Md*PER) product (Piccin 1998) truly represents *Md*PERs abundance in the clock relevant *Musca* cells.

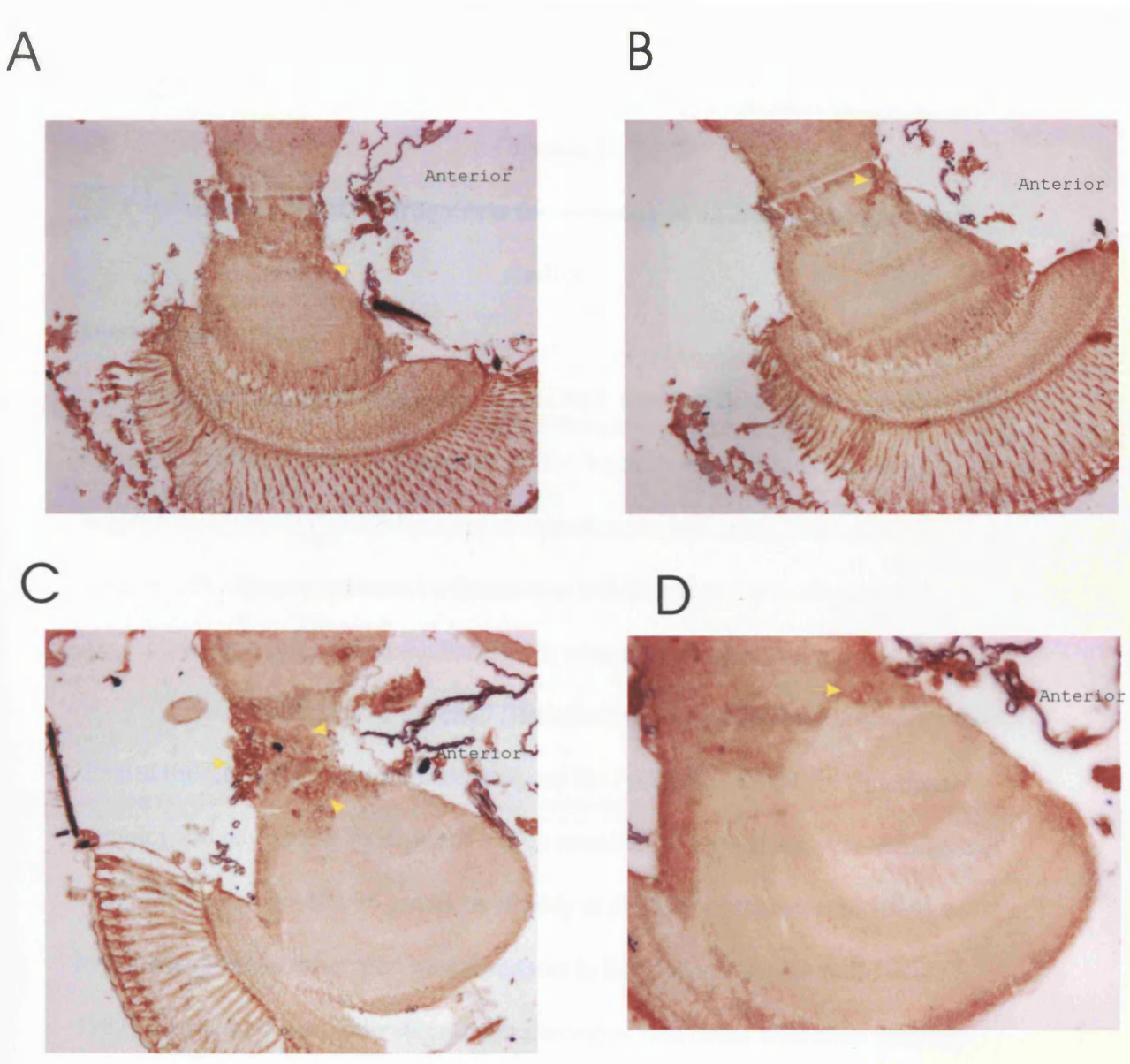


Figure 4.7 enzyme labelled immunocytochemistry of *Musca* brain sections with antibody αPER-774.
A and **B** ZT 16. **C** and **D** ZT8.

Chapter 5

Purification of protein fragments for antibody production and structural studies.

Introduction

Although the amino-acid sequence of PER was first determined in 1987 (Citri et al., 1987) it alone does not describe PERs' function. The problem of using sequence analysis to deduce function was nicely summed up with the quote,

> "...these sequences by themselves tell little more about the biological system than the New York telephone directory tells about the function of the city" (Branden and Tooze, 1991).

Indeed the first publications which analysed the sequence of PER for functional homology to other proteins declared PER a member of the proteoglycan family, and as such involved in cell-cell junctions (Reddy et al., 1986: Jackson et al 1986) a hypothesis given direct experimental support in Bargiello *et al* (Bargiello et al., 1987). However these publications gave a spurious functional allocation for PER (Saez and Young et al., 1992) which has since been retracted (Bargiello, 1992). Subsequent allocations of function from sequence analysis have been much more cautious.

In order to correctly assign function to a protein or protein fragment a rigorous analysis is required. One approach includes three main lines of evidence:

- Homology searches by sequence analysis provide cues on which to base functional hypotheses.
- Experimental tests of hypothesised function, confirming or rejecting putative functions.

- Structural determination, allowing comparison to structures of proteins of known function.

Homology searches

The homology search approach is based on using the amino acid sequence of the protein to query a database of all known proteins. An algorithm searches for other proteins that have stretches of amino acid sequence similar to the query sequence. Similar sequences are likely to have arisen through common ancestry and are likely to have a similar function .

The similarity is scored on its strength, and the statistical significance of the similarity is used to determine which hits are likely to represent true common ancestry. Database searching with the PER sequence revealed two PAS domain repeats followed by a PAC domain. The PAS domains were first identified by the sequence similarity of *period*, *Arnt* and *Singleminded* (hence PAS) (Crews et al., 1988). ARNT and SIM were both transcription factors with bHLH DNA binding motifs N-terminal of the PAS. However PER lacked the bHLH (Huang et al., 1993), and was unlikely to function as a DNA binding transcription factor. The PAC domain was identified through sequence similarity immediately C-terminal of the second PAS repeat, PAS-B, and is thought to be involved in stabilising the structure of some PAS domains (Ponting and Aravind, 1997).

The PAS repeats and PAC account for a stretch of sequence from residues 240-510 (Taylor and Zhulin, 1999). However, there are 1224 amino acids in the *D. melanogaster* PER sequence (SWP:P07663). Therefore large stretches of sequence have had no published functional allocation through sequence analysis. However there was significant conservation of sequence in insects' PER outside of the

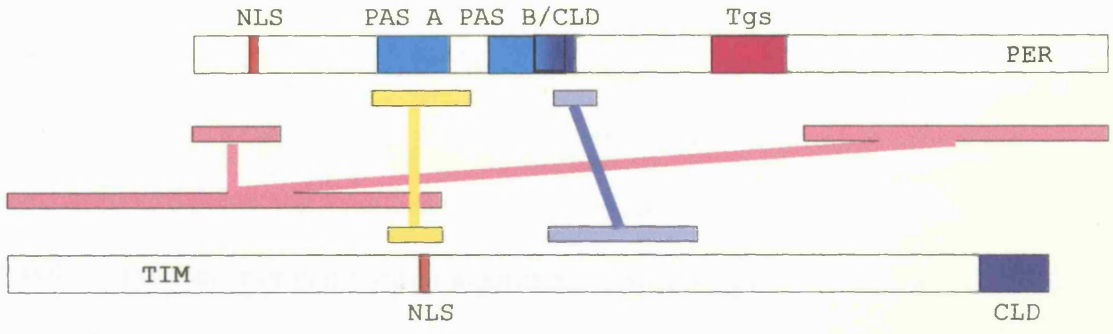


Figure 5.1 Functional domains of PER and interaction with TIM.

- PER: NLS, 66-80aa (lysine rich sequence in 1-95 aa identified as NLS (Saez and Young 1996))
- PAS-A 239-334(Taylor and Zhulin 1999)
- PAS-B 391-496 aa (Taylor and Zhulin 1999)
- CLD, 453-511 (Saez and Young 1996)
- TG (Peixoto, Costa et al. 1992).

- TIM: NLS 541-553 (KELRRKKLVKR) putataive NLS sequence in PIR A57655 from 446-577aa identified as NLS (Saez and Young 1996)
- CLD 1238-1389 aa (Saez and Young 1996)

Interacting regions:

- PER 223-512aa - TIM 505-578aa (Saez and Young 1996)
- PER 448-512aa - TIM 715-914aa (Saez and Young 1996)
- PER 1-118 and 819-1224aa - TIM 1 -580aa are reported to bind however no functional allocation was given (Sangoram, Saez et al. 1998) report as unpublished data from (Saez and Young 1996)]. However it suggests that PER may be a globular protein which doubles back on itself to allow both ends of PER to bind one end of TIM.

Note: Although Saez and Young put 453-511aa outside PAS-B a comprehensive study of what constitutes a PAS domain shows there is considerable overlap between the CLD of PER and the PAS-B domain (Taylor and Zhulin 1999) Saez and Young only included the mostconserved sequences in their annotation of PAS domains. However sequence homology outside the core strongly suggested that the PAS structure extends out past the conserved core.A model for the interaction is that PAS-B/ CLD is masked by TIM 715-914. PER can then go nuclear because its cytoplasmic retention through the CLD is prevented by the masking. Masking and unmasking of NLS and CLD are commonfeatures of nuclear import/export (Nigg 1997) . However as yet no sequence fragments have been shown to directely interact with the CLD of TIM. However it is not necessary that there to be a direct interaction for masking. Interactions elsewhere can lead to confimation changes stearic interference that could mask the CLD of TIM.

Reference sequence TIM: 1389 aa (PIR A57655), PER: 1224 aa (SWP: P07663).

PAS/PAC domains (figure 5.2). It is likely that this conservation of sequence represents selection acting to maintain functional domains within the protein. Given that protein evolution is dominated by gain, loss or duplication of protein modules (Fuchs and Buta, 1997; Riley and Labedan, 1997) it is unlikely that uncharacterised domains have arisen independently. A lack of homology is more likely to reflect the domains sequence diverging from a distant common ancestor lying out with the sensitivity of the sequence comparison algorithm. However lack of sequence homology does not imply lack of functional homology as sequences with low or no sequence identity have been shown to be structurally and functionally homologous. Therefore genetic and structural analysis could provide further information on these domains.

Genetic analysis

Further information on the function of PER has been gained by genetic analysis of *per* and its partners. The PAS domain was identified as a protein-protein interaction domain in the yeast-two-hybrid system (Huang et al., 1993). The PAS of PER dimerised with other PAS domains and the efficiency of these interactions was reduced by the mutant per^{L} (Huang et al., 1993). Furthermore TIM was identified during a yeast-two-hybrid screen for protein that interact with PER (Gekakis et al., 1995). PER was shown to dimerise with TIM through its PAS domains, the per^{L} mutant again reducing the efficiency of the interaction (Gekakis et al., 1995). The TIM sequence does not share homology with any PAS domains and was the first demonstration of a non PAS-PAS interaction. Amino acids 233-365 of PER (corresponding to the first PAS repeat, PAS A) were sufficient for this interaction (Gekakis et al., 1995). However, subsequent analysis has shown that a second

ORIENTATION OF PAGE IS AS PER THE ORIGINAL IN THE BOOK.

```
        -------->                                                                    NLS XXXXXX NLS
               bbbbbbbbbbbbbbbbbb   bbbbbbbbbbbbbb bbbbbbb     aaaaaaaaaaaaaaaaaaaaaaa         bbbbbbbb                  b
D_SIMU -------------KVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASSSD------MIIKRNKDKSRKKKKNKGTGQGAGQA-QTLISASTSLEGRDEEKPRPSGTGC 100
D_MAUR -------------KVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASSSD------MIIKRNKDKSRKKKKNKGAGQGAGQA-QTLISASTSLEGRDEEKPRPSGTGC 100
D_SECH -------------KVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASSSD------MIIKRNKDKSRKKKKNKGAGQGAGQA-QTLISASTSLEGRDEEKPRPSGTGC 100
D_YAKK MEGGESAESTHNTKVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASSSD------MIIKRNKDKSRKKKKNKGAGQGAGQAGQSLISASTSLEGGAEEKPRPSGSGC 114
D_MELA MEGGESAESTHNTKVSDSTYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASSSD------MIIKRNKDKSRKKKKNKGAGQGAGQA-QTLISASTSLEGRDEEKPRPSGTGC 51
D_WILL --------------------------------------------------------------------------NKDKSRKKKKPKCIALATATA--------VSLEG-TRESPLPASGSC 38
D_PSEU -MEGESTESTQNTKVSDSAYSNSCSNSQSQRSGSSKSMLSGSHSSGSSGYGGKPSIQTSSSD------MAIKRNKEKSRKKKKAKCTQAQATIS--------SSLEG-AEEQPHSSGTTC 104
M_DOME -MEGESTESTHNTKVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQTSR-----------LKLKSPFPYYSETNCNFFSINTQSEIHITKRGKDKGRKKKKQKCSSTVT 108
D_VIRI -MEGESTESTHNTKVSDSAYSNSCSNSQSQRSGSSKSRLSGSHSSGSSGYGGKPSTQASSSD------MAVKRNKDKSRKKKKAKSPAQATAAT----TTTIKSLEQTEEPLLVKPNNGS 109
M_SEXT ------------------------------------------------------------------------------------------------------------------------
A_PERN --MNNMDGSENNAKVSDSAYSNSCSNSQSRRSHSSKSTHSGSNSSGSSGYGGQPSTSSSND-------LSDQKKEKELKKKKQVETLMPDTQIEV------------------------- 87
P_AMER ----MEETATHNTKISDSAYSNSSSNSQSQRSSGSSKSRHS--NSSGSSGYCGHGSSIQGSSNEPFPQPSVAKRNKDKEHKKKKLKSSVTTAATVTVTSVVTTVSEYTEHENGTSHMSLGV 114

                                                                                                         <---------
                                                                                                              ---------
       bbbbbb                  aaaaaaaaaaaaaaa     aaaaaaaaaaaaaa           bbbb         bbbbb  aaaaaaaaaaaa        aaaabbbbbbbb
D_SIMU -VEQQICRELQ-DQQHGEDHSEPQATEQLQQEEE-DQSGSESEADRVEGVAKSEAAQSFPIPSPLSVTIVPPSMGGCGGVGHAAGLDSGLAKFDKTWEAG--PGKLESMTGV-GAAAAGT 214
D_MAUR -VEQQICRELQ-DQQHGEDHSEPQATEQLQQEEE-DQSGSESEADRVEGVAKSEAAQSFPIPSPLSVTIVPPSMGGCGGVGHAAGLDSGLAKFDKTWEAG--PGKLESMTGV-GAAAAGT 214
D_SECH -VEQQICRELQ-DQQHGEDHSEPQAAEQLQQEE--DQSGSESEADRVEGVAKSEAVQSFPIPSPLSVTIVPPSMGGCGGVGHAAGLDSGLAKFDKTWEAG--PGKLESMTGV-GAAAAGT 213
D_YAKK GVEQQSCRELLQDQQHGEDHSEPKATEQLQQEEG-DRSGSESEAERVENAAKSEAAQSFPIPSPLSVTIVPPSMGGCAGVGHAASLDSGLAKLDKTWEAGG-PGKVEPVPGVPGTAAAGT 232
D_MELA -VEQQICRELQ-DQQHGEDHSEPQAIEQLQQEEEEDQSGSESEADRVEGVAKSEAAQSFPIPSPLSVTIVPPSMGGCGGVGHAAGLDSGLAKFDKTWEAG--PGKLESMTGV-GAAAAGT 166
D_WILL ---EKVLQELQDTQQLG----EPLVVTETQLSEQ------LLETEQNEDQNKSEQLAQFPLPTPIVTTLSPGIGPGHDCVGGASG--GAVAGGCLVVGAGT-DKTSELIPGK--LESAGT 140
D_PSEU --DQKILHVLATTQQLGD---QPSSLDHKLGEQ--------LEARHNCGVGKAEQSQSFSLPCPLSVSTLMPGIGVCHG-GNAPG-----GKWEKTFESC----KLETGPAK-------- 193
M_DOME NALDSQEDVAKTNETKADEEIGLNDITMTQGNRQ---KENKENLEKENNEPKIEKSLQSPIPSPLSATTNQG------------------VKSEKTCESA--PGKLESSGKT-------- 197
D_VIRI ---CEQQLELQDAQQLGAP--TPSDAHDAHGDKP------QLDVDEQQDDPQAEQIQQLETATAATISPDTMSASVTVTIDGCTS-------MEKTCEWTDRPGRLEAHAAC-IGKQHVQ 210
M_SEXT ------------------------------------------------------------------------------------------------------------------------
A_PERN --------------------ECRPEEDVIN---------------IPSEEGGAADDVLVPSPKQTLQTDNDIADIEVAIPDTN-------NDKEEAIVYNTSLINPGTAC------PF 157
P_AMER ---SGTVVPLSGTVAEETEITEAGSEGSVISSHAGVALGAAGVVPATTPGPENEAHQMASLTQTLNSIKKMKKMKDLSTDIPEET-------EGHSFSLPMVAEEKEEHIRN------SF 218

           -->PAS A XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX PAS A
                   bbbb   bbbbbbb    bbbbbbb    bbbb  aaaaaaaaaaaaaaaaaabbbbbbbb                bbbbbbb        bbbbb          aaaa
D_SIMU GQRGERVKEDSFCCVISMHDGIVLYTTPSITDVLGYPRDMWLGRSFIDFVHLKDRATFASQITTGIPIAESRG-------SVPKDAKSTFCVMLRRYRGLKSGGFGVIGRPVSYEPFRLG 327
D_MAUR GQRGERVKEDSFCCVISMHDGIVLYTTPSITDVLGYPRDMWLGRSFIDFVHLKDRATFASQITTGIPIAESRG-------SVPKDAKSTFCVMLRRYRGLKSGGFGVIGRPVSYEPFRLG 327
D_SECH GQRGERVKEDSFCCVISMHDGIVLYTTPSITDVLGYPRDMWLGRSFIDFVHLKDRATFASQITTGIPIAESRG-------SVPKDAKSTFCVMLRRYRGLKSGGFGVIGRPVSYEPFRLG 326
D_YAKK GQRGERLKEESFCCVISMHDGIVLYTTPSITDVLGYPRDMWLGRSFIDFVHLKDRATFASQITTGIPIAESRG-------SVPKDTKSTFCVMLRRYRGLKSGGFGVIGRPVSYEPFRLG 345
D_MELA GQRGERVKEDSFCCVISMHDGIVLYTTPSITDVLGYPRDMWLGRSFIDFVHLKDRATFASQITTGIPIAESRG-------SVPKDAKSTFCVMLRRYRGLKSGGFGVIGRPVSYEPFRLG 279
D_WILL KPSQERPKEESFCCVISMHDGIVLYTTPSISDVLGFPRDMWLGRSFVDFVHHKDRATFASQITTGIPIAESRG-------CMPKDARSTFCVMLRRYRGLNSGGFGVIGRAVNYEPFRLG 253
D_PSEU ---TERVKEDSFCCVISMHDGIVLYTTPSITDVLGFPRDMWLGRSFIDFVHTKDRATFASQITTGIPIAESRC-------SMPKDARSTFCVMLRQYRGLQTSGYGVIGRSVNYEPFRLG 303
M_DOME ----EKLKEDSFCCVISMHDGIVLYTTPSITDVLGFPRDMWLGRSFIDFVHPKDRATFASQITTGIPIAESRS-------SIPKDARSSCCVMLRRYRGLKSGGYGVIGRSVNYEPFRLG 306
D_VIRI QQQHDRVKEDSFCCVISMHDGVVLFTTANLNEMLGYPREMWLGRSFIDFVHIKDRATFASQITTGIPIAESRC-------SQSKDARTTFCVMLRRYRGLASGGFGIIGRPVSYAPFRLG 323
M_SEXT -------------------GIVMYTTSSLTTTLGFPKDMWIGRSFIDFVHPRDRNTFASQITSGLAVPKIVN-----GQSPGNPA-STMVCRIRRYRGLT-TGFGVKDRVVTFMPFLLK 93
A_PERN GRPALSNCNG-FSCVISMHDGVVLYATASLTSTLGFPKDMWVGRSFIDFVHPRDRNTFASQITNELAIPKIVSLTEETDQTMENPG-STMVCRIRRYRGLS-CGFSVKNTTTAYLPFLLK 274
P_AMER DAEPPAHNEGEFCVVVSMQDGVVVFTTPSITDVVGFPKDMWLGRSFIDFVHPRDRTAFANHIASGVITPLSNS----NTKGGSHPGKNSFYCSLRRYRGLKSTGYGVTEKEVSYLPFQLN 334
                          *:*:::*..:.  :*:*::**:****:****  :**  :**.:*:. :  .                 :   .:    :*:****    *:.:     . : ** *
```

Figure 5.2 continued over leaf

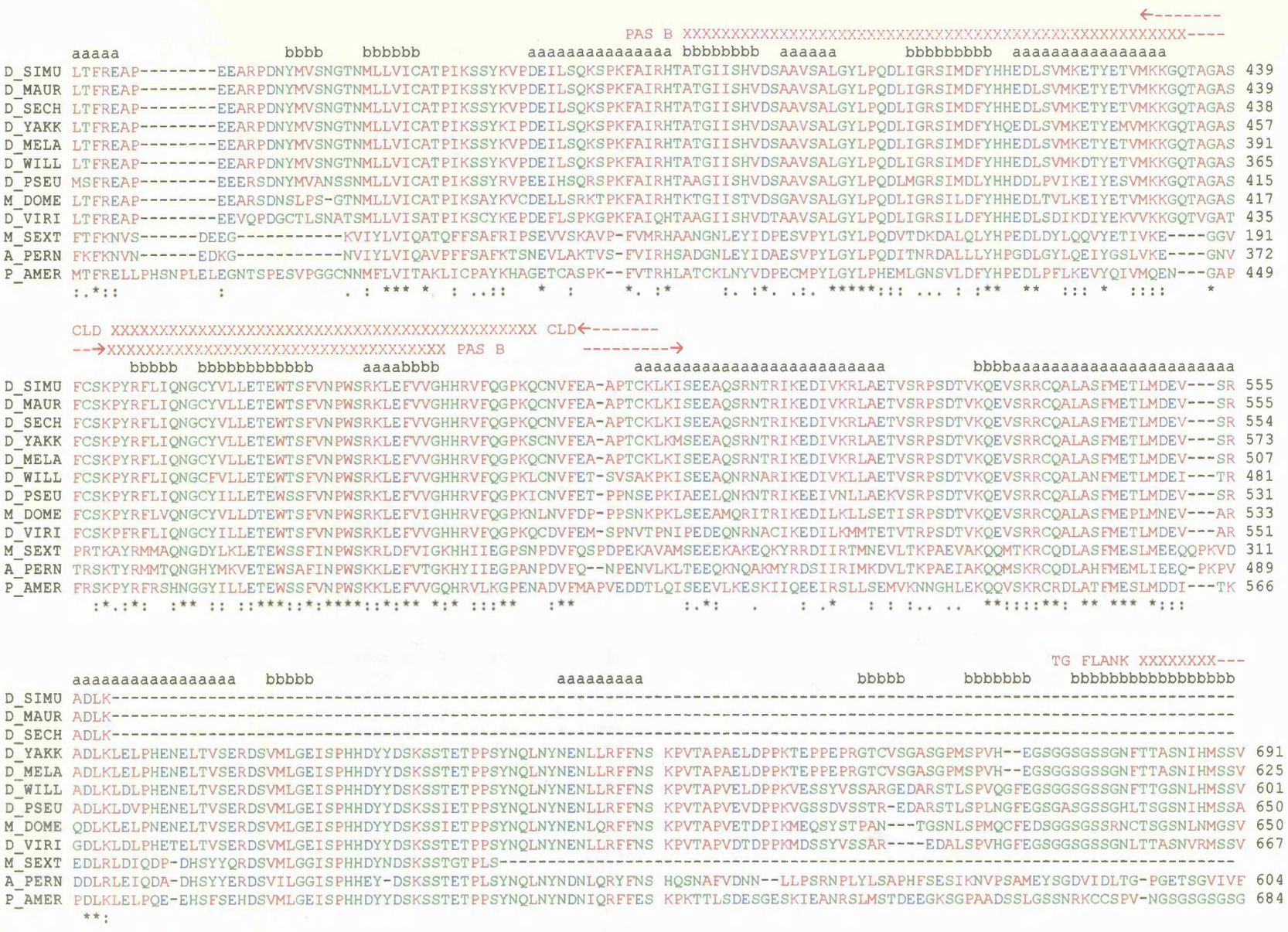


Figure 5.2 continued over leaf

Figure 5.2 continued over leaf

Figure 5.2 Clustal W alignment of insect PERs

-----→ primer site
XXXXXXX Functional domains
aaaaaaa α-Helix
bbbbbbb β-sheet

region of PER (amino acids 453-511) was involved (Saez and Young, 1996). Two further functional domains were identified, the Cytoplasmic Localisation Domain (CLD) and the Nuclear Localisation Signal (NLS). The CLD was first suggested through fusion gene analysis which showed that amino acids 95-529 inhibit the nuclear localisation of fusion proteins. Constructs without this sequence stretch could enter the nucleus without dimerising with TIM (Vosshall et al., 1994), nuclear localisation of PER normally requiring TIM. Further mapping showed that amino acids 453-511 were responsible for PERs cytoplasmic location, these amino acids were termed the CLD (Saez and Young, 1996). In addition a NLS was mapped to the first 95 amino acids of PER (Saez and Young, 1996) and is likely to consist of a Lysine rich stretch from amino acids 66-80 which are common signals for nuclear import factors (Nigg, 1997). It has also been shown that PER binds TIM across two stretches of sequence, PER 233-365 binding TIM 505-578 (Gekakis et al., 1995), PER 448-512 binding TIM 715-914 (Saez and Young, 1996). In addition PER 1-118 and PER 819-1224 aa are reported to bind TIM 1 –580aa. However no functional allocation was given and this interaction was reported as unpublished data (Sangoram et al., 1998). The sequence homologies and functional analysis allows a model of PER functional domains and its interacting regions with TIM to be developed, (**Fig. 5.1**).

With a number of functional sequence stretches identified, the determination of the structures of PER and TIM should provide valuable information on how they function. In particular, large proteins like PER and TIM can often be viewed as consisting of several functional domains. Domains are globular units linked together by short lengths of polypeptide chain, which can be independent units or interact structurally (Branden and Tooze, 1991). The concept of domain is not clear

cut, with a variety of definitions. However it is convenient to refer to a domain as an autonomous folding unit, because they usually fold independently of each other (Branden and Tooze, 1991). Only through structural determination can the domains of PER and TIM be fully determined.

Structural analysis

Structural analysis can reveal several aspects of protein structure.

- Primary structure- describes amino acid sequence.
- Secondary structure- describes regular local confirmations, especially α-helices and β-sheets.
- Tertiary structure describes assemblies of elements of secondary structure into domains.
- Quaternary structure describes the three-dimensional arrangements of different domains within the whole native protein structure.

The primary structure of PER and TIM has already been determined through sequencing the coding sequence of the genes (Citri et al., 1987, Myers et al., 1995). Secondary structure is only truly determined through structural analysis such as NMR or X-ray diffraction. CD provides gross analysis of overall secondary structure but does not provide information on which amino-acids are involved in α-helices, β-sheets or loops. Tertiary structure can be determined by NMR and X-ray diffraction. Quaternary structure can be determined if the whole protein has been crystallised. Quaternary structure can often also be pieced together by analysis of tertiary structure of fragments of the whole protein.

A structure on its own does not allow determination of function. Even strong structural similarity can be misleading, for instance CASK a membrane associated protein has strong homology to a guanylate kinase. However the key catalytic serine residue that would allow it to act as a kinase is not present (Hat et al., 1996). Instead the domains appear to be involved in protein-protein interactions (Hat et al., 1996). It is likely that the stable backbone of the kinase has co-opted into another function during the evolution of the protein. In such a case structural homology alone could misassign the function of CASK into a kinase function.

Only with functional data, obtained through genetic or biochemical analysis, does structure come into its own. Once a functional site such as a protein-protein interaction domain has been identified through genetic or biochemical assays then sequences involved can be mapped onto the structure. The structure can be used to suggest areas physically adjacent to the relevant sequence which may also be contribute to function. Furthermore, residues which are on the protein surface and are more likely to mediate an interaction can be identified, along with amino acids that lie deep within the structure and contribute to structural stability through packing interactions. The structure can then be used to suggest sites for mutagensis experiments to map the extent of the interacting residues. With genetic and biochemical information on function, structural analysis becomes an invaluable tool for refining the identification of residues which are important in the function of the protein.

PER and TIM are two proteins with an expanding wealth of genetic and biochemical information on function. Structural analysis of both proteins would help refine our models of PER and TIM functions. PER in particular has a large

dataset of interacting regions, functional domains and mutations. A structural model would help fully define the functional regions.

Aim

The aims of this chapter were to purify fragments of *MdPER* for structural studies, and to provide enough pure *Md*PER to inoculate rabbits for antibody production as detailed in Chapter 4.

Approach

Antibody production required the purification of protein fragments spanning the entire length of *MdPER*. With an additional effort the fragments could be purified for structural analysis. By breaking the protein into small fragments, an approach designed for ease of PCR and cloning, all of the putative domains of *MdPER* could be expressed and purified. This approach lent itself very well to the production of protein fragments for structural analysis, all the putative domains of *MdPER* could be covered. The separate *MdPER* fragments were designed so they spanned domains and should therefore be able to fold independently and provide soluble, folded candidates for structural analysis. The approach used to identify the domains of *MdPER* is listed bellow, and then described fully:

- Align PER homologues.
- Overlay predicted secondary structure on alignments.
- Assign domain boundaries. Using the alignments, predicted secondary structure and information on domains and functional sequence stretches to break proteins up into independently folding domains.
- Create primers and PCR.

- Clone.
- Express.
- Purify.
- Subject to structural analysis.

Alignments

The alignments were performed during April 1998 and only included information current to the databases at this time.

The sequence of insect PERs were obtained through database searches. The *D. melanogaster* PER sequence was used to query the gene sequence databases (using the BLAST algorithm) and return homologues from other species. Unpublished PER sequences were also used; *Musca domestica*, provided by Alberto Piccin (pers. comm.) and *Periplaneta americana*, provide by Steve Reppert (pers. comm.). The sequences were then aligned with the Clustal-W algorithm, to produce a global alignment of all the sequences (Thompson et al., 1997).

Secondary structure prediction

The secondary structure prediction for *D. melanogaster* (Tamburro. Pers. commun.) was then marked onto the alignment. Secondary structural prediction is greater than 70% accurate (Rost and Sander, 1993) and therefore provided valuable information on domain boundaries.

Domain boundaries

Domains of *MdPER* were assigned using three approaches. Firstly, *Md*PER was aligned to the functional domains of *d*PER published in the literature (see **Fig.**

5.1), primers were placed between functional domains. Second, conserved regions with no functional data were identified in the aligned insect PER sequences. Stretches of conserved sequence were maintained and the sequence broken at areas of low complexity or conservation. Thirdly, primers were placed between regions of predicted secondary structure in order not to break structural units.

The resultant alignment along with the secondary structural information and functional information was used to break the protein into candidates that should consist of separate functional domains. Figure 5.2 shows the alignment of all the insect PERs with secondary structure and primer sites marked on the alignment. This method for identifying domains resulted in the production of eight fragments (**Fig. 5.3**).

Primer design, PCR and cloning strategy

Primers were designed to produce in frame fragments of the identified domains. The forward primers were designed with a 5' Nde1 site, which it adds to the PCR product, immediately after this a methionine is also introduced followed by the in frame *MdPER* cDNA sequence. The reverse primers introduced Xho1 site at the extreme 3' end of the cDNA fragment, immediately 5' to the Xho1 site a stop codon was also introduced, see figure 5.4 below.

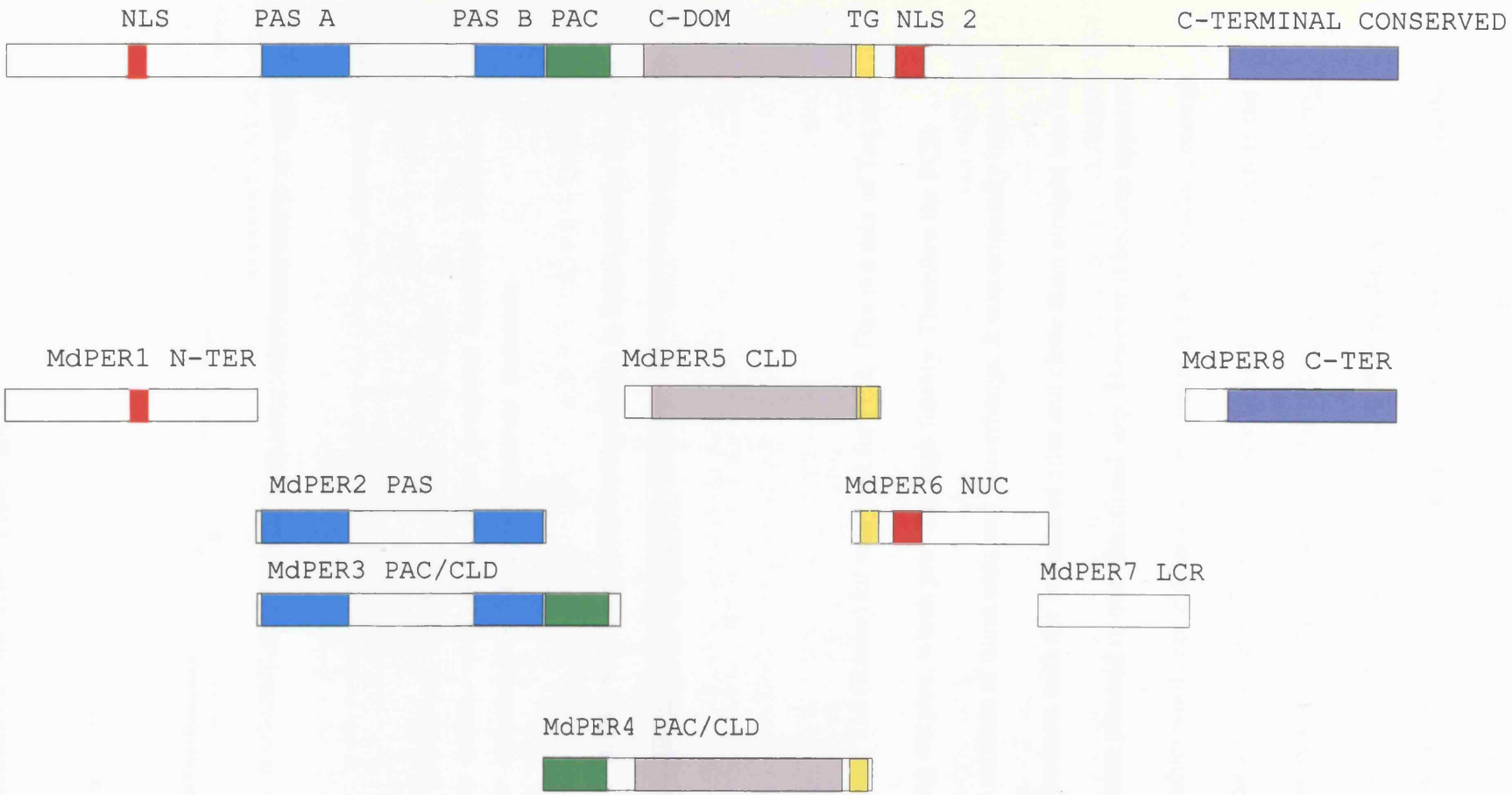
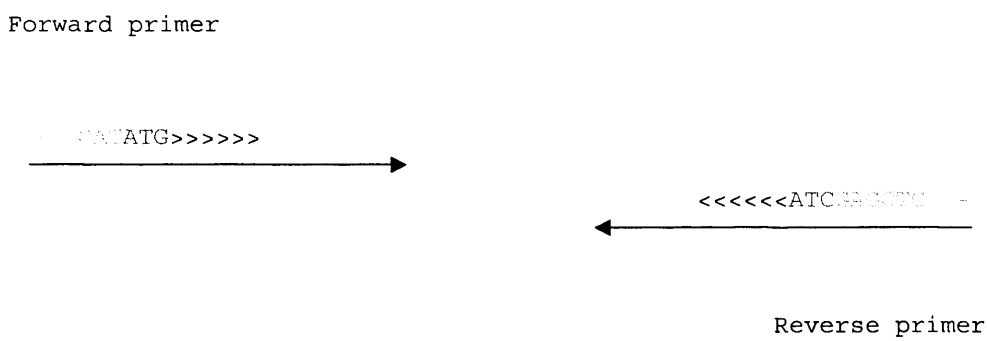


Figure 5.3 *Md*PER sequence broken up into eight fragments for bacterial expression.

Figure 5.4 Protein expression primer design.



ATG = Nde1 site and ATG, = Xho1 site,
ATC = Stop codon, >>>>>>= in frame sequence specific primer,
= 5' Overhanging non-complementry sequence

A table of primer sequences with corresponding in frame protein sequence and protein fragments is given in table 5.1.

Methods

The Expand TM (Roche) kit was used for PCR. This is a mix of Taq and Pfu, proof reading enzyme, which provides high fidelity. Therefore the PCR fragments had a mixture of blunt ends and T overhangs. It was originally intended to cut the PCR products with the introduced sites and clone them straight into the pET-14b expression plasmid in one directional step. However it became apparent that the PCR product would not cut because the primers did not introduce enough sequence 5' of the sites to allow the endonuclease to bind and cut. Therefore the product were first cloned into Nco1 linearised pDk101 through T-A cloning. This step was inefficient because only a proportion of the PCR products had a T overhang. Once the PCR fragments had been cloned they were then cut out of the pDk101 with the introduced sites and ligated directionally into pET14b. The resultant insert containing pET14b was transformed into a bacterial host and used to

produce large amounts of the protein fragments encoded by the PCR product. The pET14b plasmid introduces a N-terminal His-tag 5' to the Nde cloning site which aided purification (See figure5.5).

Figure 5.5 pET14b His Tag.

T7 promotor>>>>ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGC[illegible]ATG[illegible]

MetGlySerSerHisHisHisHisHisHisSerSerGlyLeuValProArgGlySerHisMetLeuGlu

ATG = Start Methionine, [illegible]ATG = Nde1, [illegible] = Xho1.

Expressions and purifications

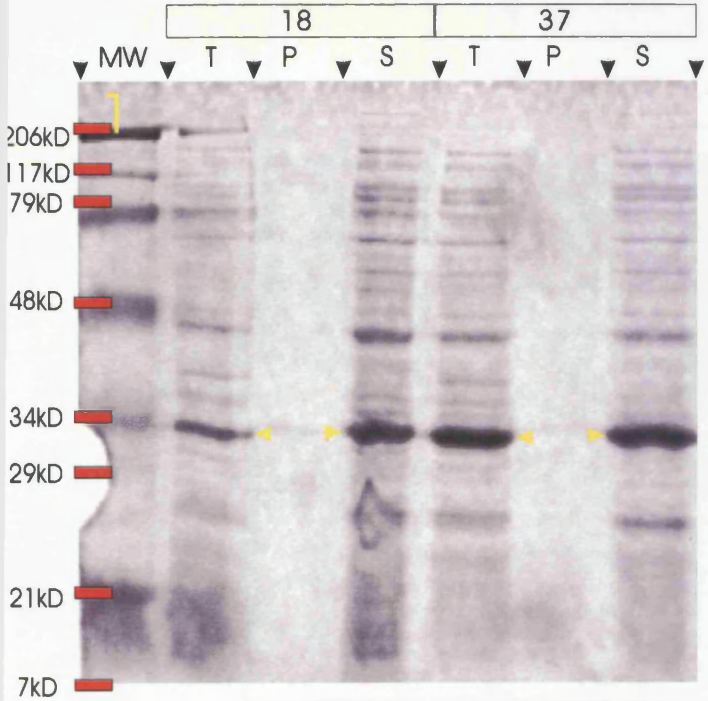
The goal of protein production and purification was to produce a relatively large amount of very pure protein in a soluble form. Typically NMR and X-ray crystallography require 1ml of 10mg/ml of soluble pure protein. The only practicable way of achieving this was to overproduce the proteins in a transgenic cell line. There are many commercially available technologies for protein production: bacteria, yeast, insect or mammalian cell lines can all be used. Each one has advantages and disadvantages, bacteria are cheap and easy to use, and offer high levels of protein expression. However they do not provide any posttranslantional modification and have a different cellular environment to eukaryotic cells, resulting in fewer disulphide bridges, which are important in stabilising some protein structures (Darby and Creighton, 1990; Derman et al., 1993). Yeast and insect cells offer some of the advantages of eukaryotic posttranslational modification but can often result in poorer expression and are less easy to handle. Mammalian cell lines can provide complex posttranslational modification but require dedicated facilities, are expensive and often result in lower protein expression.

Bacterial expression was chosen for its speed and ease of use. In order to purify the over-expressed gene away from the bacterial proteins a polyhistidine tag was used. The tag allowed binding of metal ions; nickel in particular. Polyhistidine stretches are rare in proteins and therefore afforded a reasonable level of purification of expressed proteins. In addition the tag itself is short and flexible and is unlikely to interfere with native structure of the protein; it is often unnecessary to remove the tag prior to structural analysis.

All bacterial expression was performed in the BL21 *E. coli* strain and expressions, lysis and purifications were performed as detailed in the methods section.

Expression

Initial expression in bacteria were performed to determine the apparent size, solubility and yield of the eight expression constructs. All Cultures were grown to 0.6-1.0 OD at 37^0C: cultures which were expressed at 37^0C were then induced for three hours and harvested (see methods section). For 18^0C expression, cultures were moved to 18^0C for one hour to equilibrate and then induced overnight (14-16 hours) and harvested. Different expression temperatures were used because lowering the temperature of expression can greatly increase the solubility of recombinant proteins (Schein and Noteborn, 1988). The bacterial were harvested by centrifugation, lysed, fractionated and loaded on SDS-PAGE for analysis (**Fig. 5.6** and **5.7**) (see methods section for SDS-PAGE analysis). Fragments *Md*PER1, 5, 8 appeared soluble and had high levels of expression. Fragments *Md*PER2, 3, 4 and 7 were insoluble and well expressed at all temperatures. Fragment *Md*PER6 appeared to be toxic or degraded, as it could not be easily visualised through SDS-PAGE and



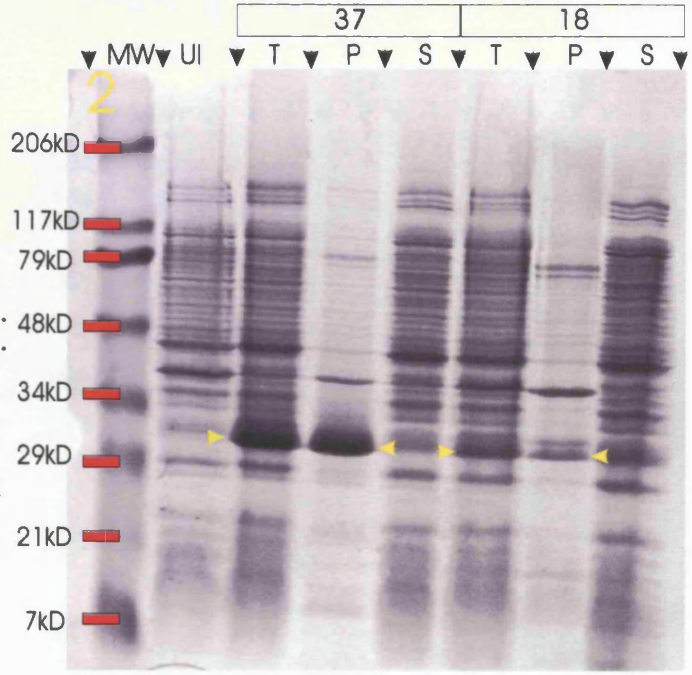
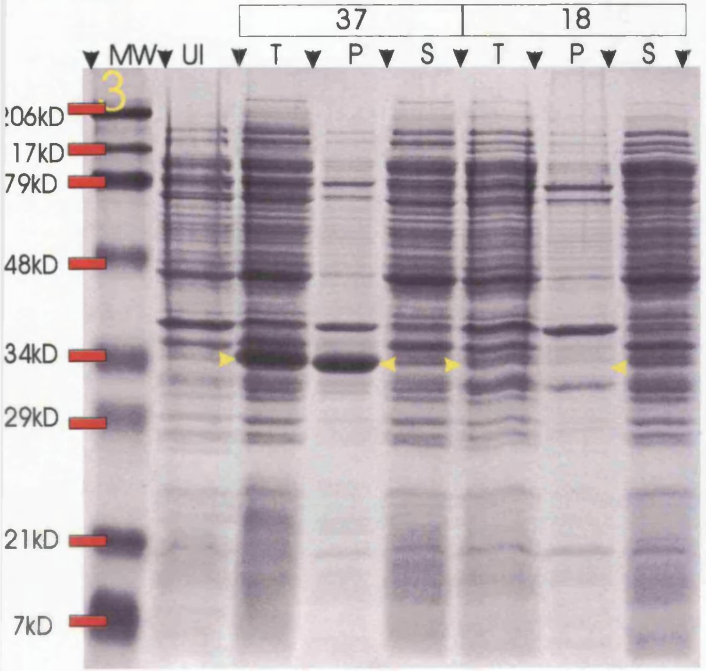
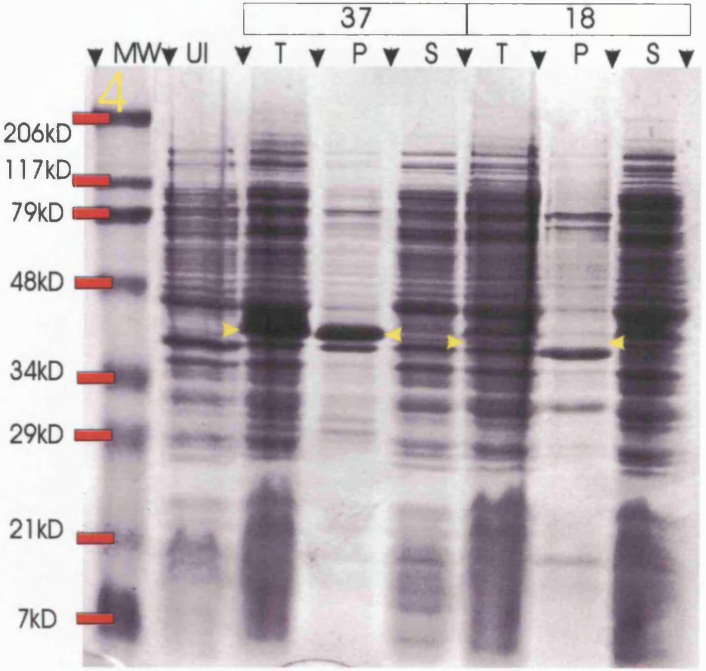
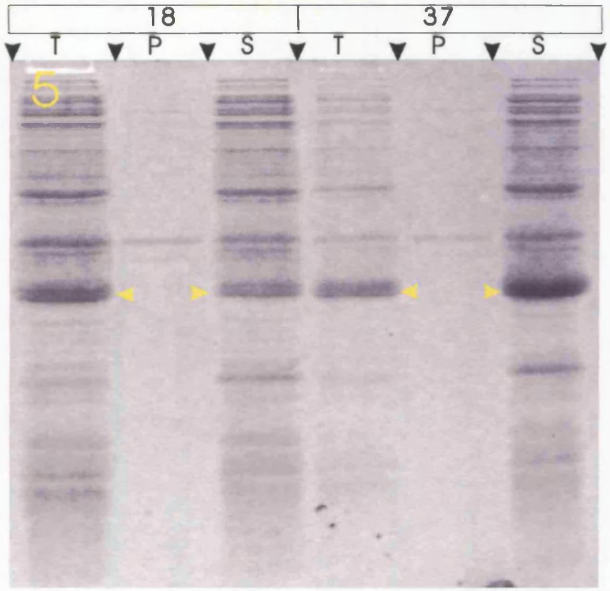


Figure 5.6 SDS-PAGE analysis of fractionated bacterial expressions. MW= Molecular weight broad range markers from Biorad. UI= Uninduced protein sample. T= Total protein. P= Pelleted protein or insoluble fraction. S= Soluble proteins. 18=18^{0}C incubation and induction. 37=37^{0}C incubation and induction. By comparing uninduced protein to induced protein novel bands resulting from overexpression of the protein can be identified. Yellow arrows mark novel induced protein bands. *Md*PER fragment number in yellow at the top left hand corner of gel.







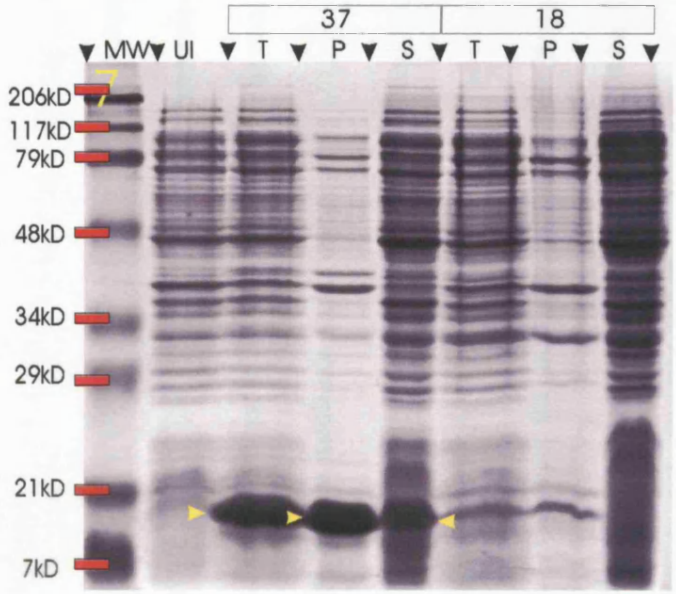
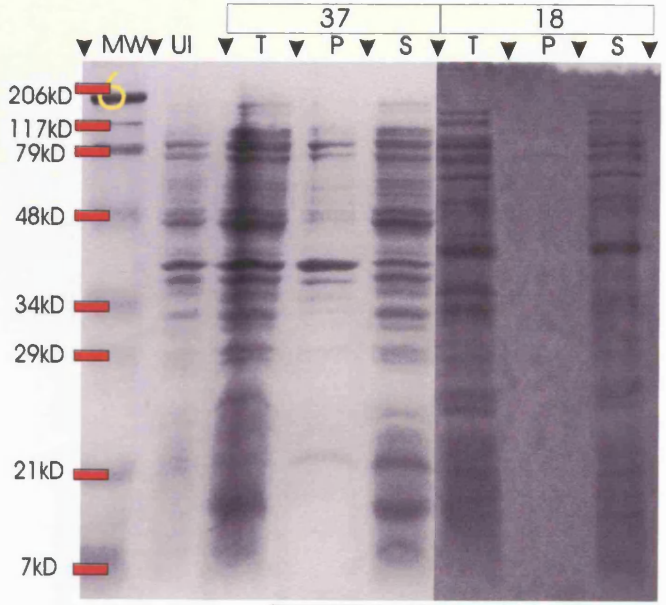
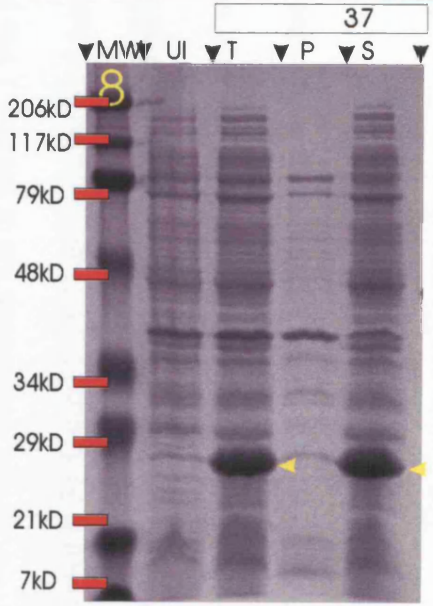


Figure 5.7 SDS-PAGE analysis of fractionated bacterial expressions. MW= Molecular weight broad range markers from Biorad. UI= Uninduced protein sample. T= Total protein. P= Pelleted protein or insoluble fraction. S= Soluble proteins. 18=18^{0}C incubation and induction. 37=37^{0}C incubation and induction. By comparing uninduced protein to induced protein, novel bands resulting from overexpression of the protein could be identified. Yellow arrows mark novel induced protein bands. *Md*PER fragment number in yellow at the top left hand corner of gel.Fragment 8 was only expressed at 37^{0}C. Initial clones were misidentified and did not express, subsequent clones were used and expressed at 37^{0}C only.





Coommasie staining. The results for all the expressed fragments are summarised in table 5.2. All the proteins had highest expression at 37^0C, so all further expressions were carried out at 37^0C. At this point fragment *Md*PER 2 and 4 were discarded. Both were insoluble and covered parts of the protein that had already been covered by the other fragments. The purpose of producing fragments 2 and 4 was to provide soluble proteins that contained some of the interacting motifs identified in figure 5.1. However the lack of solubility of both fragments and the redundant nature of there amino acid coverage allowed them to be discarded.

Purification of protein fragments

With the expression levels of the fragments characterised, strategies for purification had to be developed and are detailed below.

Soluble proteins

***Md*PER1**

A trial purification run of MdPER 1 was performed. A 50ml culture was grown at 37^0C until 0.8 OD and induced for 3 hours. Following induction the cells were harvested and stored. Lysis was performed and the sample bound to 0.25mls prepared Ni-NTA matrix. The matrix was then washed and eluted on a biologic LP chromatography system (details on growth, induction, harvesting, storage, lysis and elution is given in methods section).

The elution profile was shown in Figure 5.8A. 60mM imidazole was used as a primary elution to remove weakly bound proteins from the column, shown in lanes 1-4. The expressed protein fragment was then eluted in 200mM imidazole, lanes 5-7. A full length MdPER was eluted (yellow arrows). In addition, two minor

Fragment	Base pairs	Sequence ID	Length amino acids	Features	Calculated kDa	Aparent kDa	37 solubility	15 solubility
MdPER1	585	1-195	195	N-terminal with Nuclear Signal 1	23	29	Y	Y
MdPER2	678	193-417	224	PAS A-B	27	31	N	N
MdPER3	849	193-474	281	PAS A-B + PAC	34	34	N	N
MdPER4	675	415-669	254	PAC+CLD	31	31	N	N
MdPER5	594	472-669	197	CLD	24	30-34	Y	Y
MdPER6	456	648-800	152	Nuclear Signal 2	18	27	Low level	Low level
MdPER7	654	796-913	117	Low Complexity Region	14	18	N	N
MdPER8	492	909-1072	163	C-Terminal conserved region in diptera	20	20	Y	Y

Table 5.2 Summary of trial inductions and lysis of the MdPER fragments and additional information for each protein

bands were eluted (blue arrows), these may have been due to impurities or degradation products of the major band. However the protein produced was remarkably pure.

The experiment was then scaled up to produce sufficient protein for structural studies. Half litre of culture was lysed and bound to 2.5mls Ni-NTA, representing a 10 fold increase in scale compared to the previous purification. The pattern of eluted proteins achieved during the purification was quite different to that of the smaller scale purification, although the only difference in treatment was the increase in scale of the experiment (**Fig. 5.8A, 5.8B**: SDS-PAGE analysis of the smaller scale purification and larger scale purification respectively). The elution peak was spread over many fractions, as would be expected from an increase in scale. However there was a large increase in the number of contaminating bands (marked with blue arrows). This was due to increased degradation of the full length fragments. Comparing the relative band intensities of the total protein resultant from the two lyses (**Fig. 5.8A** lane T and 5.8B lane T) there was one clear expressed band in the smaller scale lysis, with much less intensely stained degradation products (blue arrows). However the larger scale lysis produced more intensely stained degradation products which were then eluted from the Ni-NTA along with the full length protein (**Fig. 5.8B** lane 6).

The increased intensity of these degradation products was likely to be due to the slower lysis of the larger cell pellets during the increased scale purification. The larger frozen bacterial cell pellets took longer to thaw and resuspend. In addition the pellets required extended sonication to break open the cells and reduce the viscosity resultant from genomic DNA. Although a suite of protease inhibitors was used in the lysis buffer these are only effective when the proteins are completely

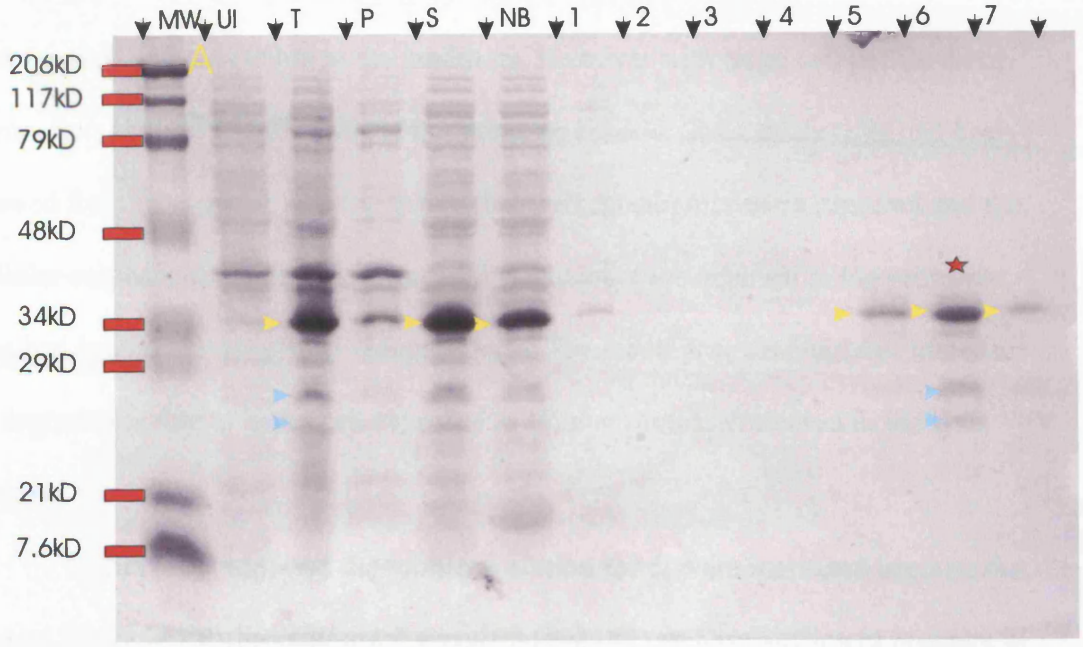


Figure 5.8 A, small scale purification run for *Md*PER1. MW= Molecular weight markers, UI= Uninduced cell pellet, T= Total induced cell protein, P= Insoluble cell protein, S= Soluble protein, NB= Not bound protein, 1-7 = Eluted fractions. Yellow arrows mark full length *Md*PER1 and blue arrow mark *Md*PER1 degradation products. The red star marks the fraction used for inoculation.

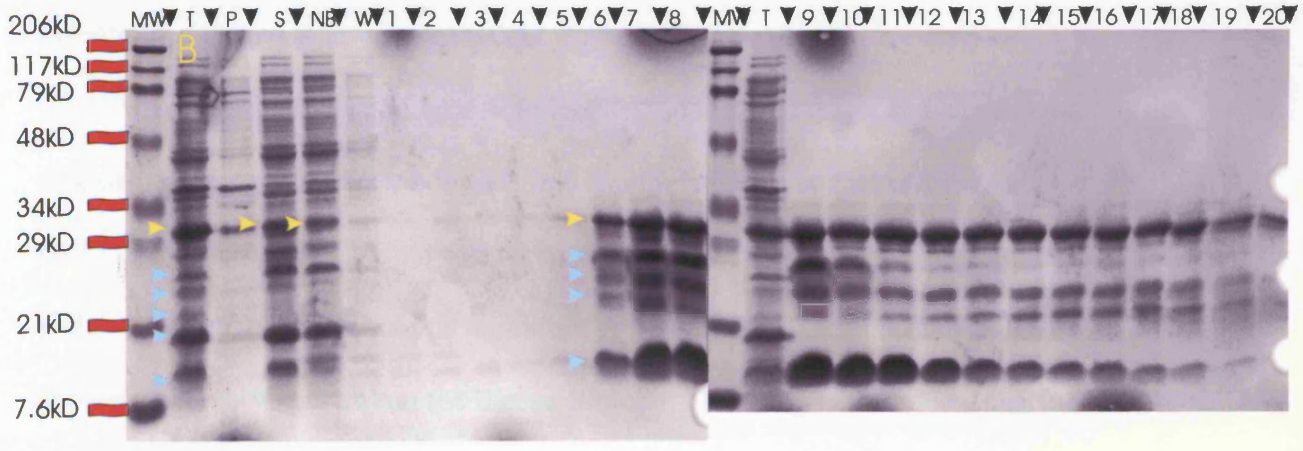


Figure 5.8 B, large scale purification run for *Md*PER1. MW= Molecular weight markers, UI= Uninduced cell pellet, T= Total induced cell protein. P= Insoluble cell protein, S= Soluble protein, NB= Not bound protein, W= Batch wash, 1-20 = Eluted fractions. Yellow arrows mark full length *Md*PER1 and blue arrow mark degradation products.

resuspended and accessible to the inhibitors. However with large cell pellets there were often clumps of cells that did not break up readily. Since these cells had been thawed from freezing, it was likely that their cell membranes were ruptured and the cellular contents, including the expressed fragment, were exposed to the proteases that had leaked out of cellular compartments. The result is an unavoidable increase in degradation due to increased exposure to cellular protease released in the lysis process.

To further compound the problem, elution times were increased because the Biorad Biologic LP chromatography system used did not have sufficient pressure to quickly force elutant through the increased bed volume of matrix. Inhibitors were not used throughout the elution because of the prohibitive cost of adding inhibitors to the large volume of washes and elutants used. Therefore the bound proteins were potentially exposed to any proteases not washed off the column.

In order achieve good lysis with less degradation the larger scale lysis was repeated, except pellets where divided into smaller aliquots so that resuspension could occur more rapidly. Care was taken to insure that the lysis, binding, elution and analysis were completed in a minimum of time. A few fractions that showed protein peaks (UV spectra) on the Biorad Biologic LP during elution were analysed by SDS-PAGE (**Fig. 5.9A**). The fractions were larger, due to the increased flow rate during the elution and so had a larger volume of eluted protein. The purification resulted in slightly fewer degradation products, however there was still a large number of low molecular weight bands (marked in blue). The fractions were stored at 4^0C over a weekend with the addition of protease inhibitors (methods section for a full list). They were then resampled and subject to SDS-PAGE analysis (**Fig. 5.9B**). The number of bands had increased in each fraction; fraction 24 and 38 were

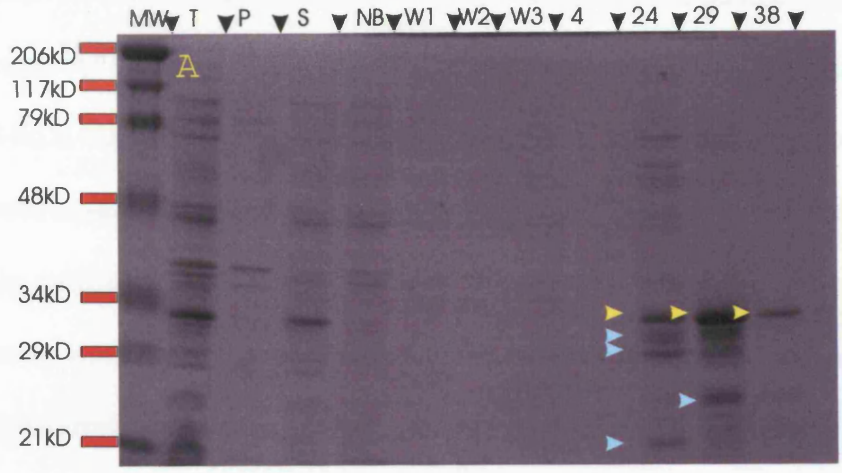
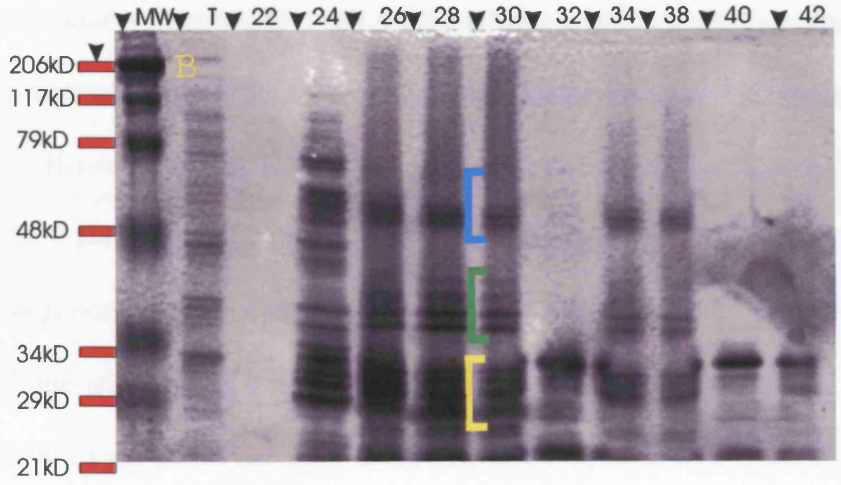
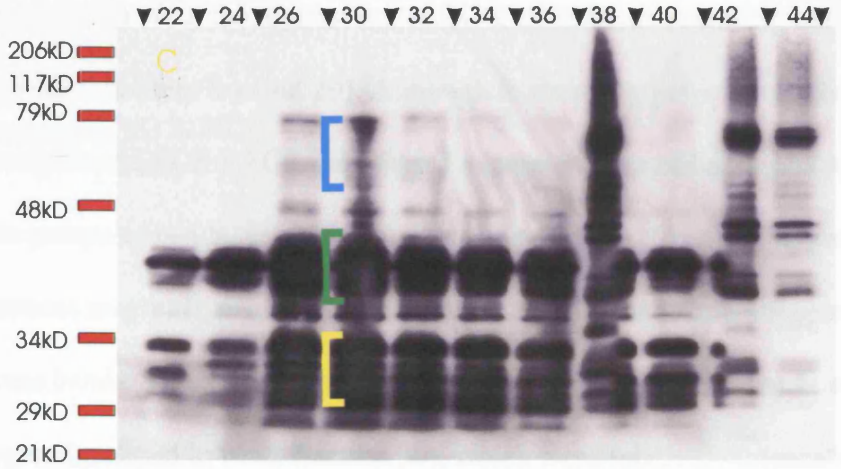


Figure 5.9A. A quick ½l purification run of *Md*PER1.
MW= Molecular weight markers, T= Total induced cell protein, P= Insoluble cell protein, S= Soluble protein, NB= Not bound protein, W1-3= Batch washes, 4-38 = Eluted fractions.
Yellow arrows mark full length *Md*PER1 and blue arrows mark degradation products.



B: SDS-PAGE, Coomassie stain on stored fractions.
MW= Molecular weight markers, T= Total induced cell protein.
22-42 = Eluted fractions.



C: Western blot with anti-his-tag antibody on stored eluted fractions from A.
22-44 = Eluted fractions. The yellow arrows mark full length *Md*PER1 and blue arrow mark degradation products. The blue,green and yellow brackets mark different size classes of multimers common to both SDS-PAGE and western analysis.

common to both gels and show a multiplication of band number. This would be an expected consequence of degradation. However most of the bands occur above the 34kDa weight of full length expressed *MdPER*1 and degradation could not produce such a pattern. I hypothesise that the extra bands occurred through the binding of the purified full length and truncated fragments to each other. This would have to occur with a very strong affinity as SDS-PAGE subjected the proteins to harsh denaturing conditions. Another plausible hypothesis was that bacterial growth has not been inhibited over the weekend by the addition of sodium azide and the high molecular weight bands result from bacterial growth. These hypotheses were tested by resampling the stored fractions and subjecting them to western analysis with an anti-His-tag antibody (**Fig. 5.9C**). Although the pattern had changed from the SDS-PAGE gel (**Fig. 5.9B**) the staining pattern between the SDS-PAGE and western blot was generally comparable. Bands with similar molecular weight on the SDS-PAGE gel and western blot have been grouped for comparison. There was a strong group of 5-6 strong bands between 34-29 kDa (grouped in yellow, fraction 30) and another group of strong bands between the 34 and 48 kDa markers (grouped in green, fraction 30), with the highest molecular weight bands occurring between 48-79 kDa (grouped in blue, fraction 30). Although the banding pattern was not identical, the westerns and SDS-PAGE gels shared a large number of bands and it was likely that the groups of bands resulted from the same proteins. Changes between banding patterns may have resulted from an extra day of storage or from epitope masking in some bands. This was likely to occur with bands found between in 48-79 kDa region (grouped in blue, fraction 30), which were relatively intensely Coomassie stained but much were weakly labelled on the blot. The variety of bands found could result from the binding of three or four truncated degradation products

sticking in a number of combinations, each combination resulting in a protein complex with a unique molecular weight. The comparable patterns rule out bacterial contamination as most of the bands contain his-tags and must result from truncated products and multimers of the products. The anti-his tag antibody was a commercial antibody that was reputed to have an avidity for only his-tagged proteins. Indeed it did not react with any bacterial samples that were not induced to produce his-tagged proteins, therefore cross-reaction with invading bacteria could be ruled out.

The red star in figure 5.8A marks the fraction chosen for use in the inoculation for antibody production. It is the purest fraction achieved. Concentration was achieved using amicon centiprep 10kDa column. The concentration was determined using the Bradford assay calibrated against a BSA standard curve (methods section). The preparation of the inoculum was detailed in chapter 4.

*Md*PER 5

A trial purification was performed for *MdPER5*, 50mls of the culture was lysed, bound and eluted as detailed for *MdPER1*. SDS-PAGE analysis showed poor lysis had been achieved. The pellet and total fractions have similar protein profiles, whereas the soluble fraction showed a much weaker banding pattern (Fig. 5.10A). Therefore most bacteria had failed to lyse and were spun down into the pellet fraction leading to less protein in the soluble fraction available for matrix binding. The 60mM wash eluted no proteins (lanes 1-3). The 200mM elution produced one major band at approximately 34kDa (yellow arrows lanes 5-7), close to the size at which *Md*PER5 migrated during the initial expressions (Fig. 5.6) and above the size predicted full length *Md*PER5 (table 5.2). A few minor degradation products were also copurified (blue arrows, lanes 5-7).

The scale was increased, a ½ litre culture was lysed, bound and eluted in the same manner as *MdPER*1. The 60mM imadoazole elution appeared to produce a large number of bands (**Fig. 5.10B**) and no such pattern had been achieved during the small scale purification (**Fig. 5.10A**). The 200mM elution gave a pattern similar to the 60mM elute (fractions 14-33), however two main bands were predominant (marked with a blue and yellow arrows fraction 13). Towards then end of the 200mM elution these two bands were eluted exclusively. The lysis and elution appeared to produce one major degradation product of the ~34kDa full length product, a major band at 27kDa. This product was not present in the small scale lysis, (**Fig. 5.10A**). Again the increased time it took to lyse a ½ litre pellet appears to introduce the problems of degradation. However from the intensity of the bands it is clear that MdPER 5 expressed well and bound the matrix efficiently. Therefore large quantities of the protein could be purified away from the majority of the bacterial proteins. With further purification sufficient protein for structural analysis could be achieved.

Another half litre culture of *MdPER*5 was lysed, with the lysis optimised so that it was achieved as rapidly as possible (see *MdPER*1). Fractions that contain protein peaks (UV absorbance) during the Biologic LP purification were subject to SDS-PAGE analysis (**Fig. 5.11A**). Washes were performed prior to Biologic LP elution in order to remove unbound proteins as quickly as possible (**Fig. 5.11A, w1-w3**). The 60mM elute (fraction 9) did not produce the same multiple banding pattern as 60mM elute in the earlier lysis (**Fig. 5.10B, 10-13**). Therefore it was likely that in 5.10B poor lysis resulted in a large number of degradation products that formed multimers, which were eluted earlier from the column. The quicker lysis (**Fig. 5.11A**) resulted in fewer degradation products in 200mM elutant as well

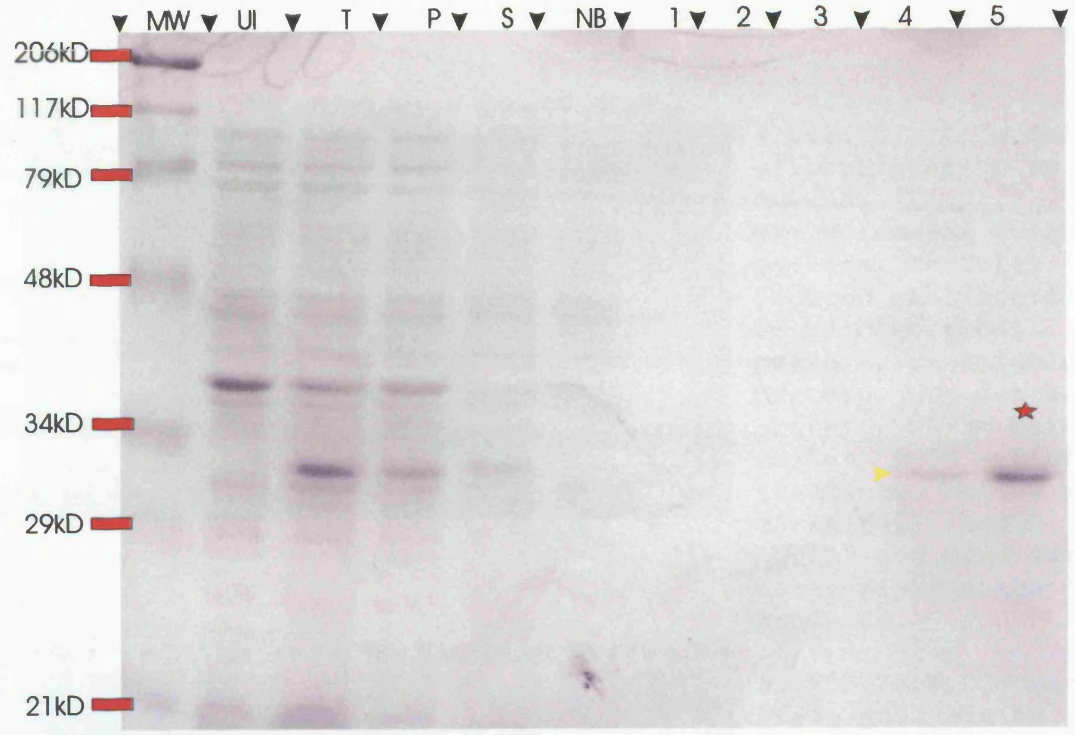


Figure 5.10 A, small scale purification run for *Md*PER5. MW= Molecular weight markers, UI= Uninduced cell pellet, T= Total induced cell protein, P= Insoluble cell protein, S= Soluble protein, NB= Not bound protein, 1-5 = Eluted fractions. Yellow arrows mark full length MdPER5 and the red star marks fraction used for inoculation.

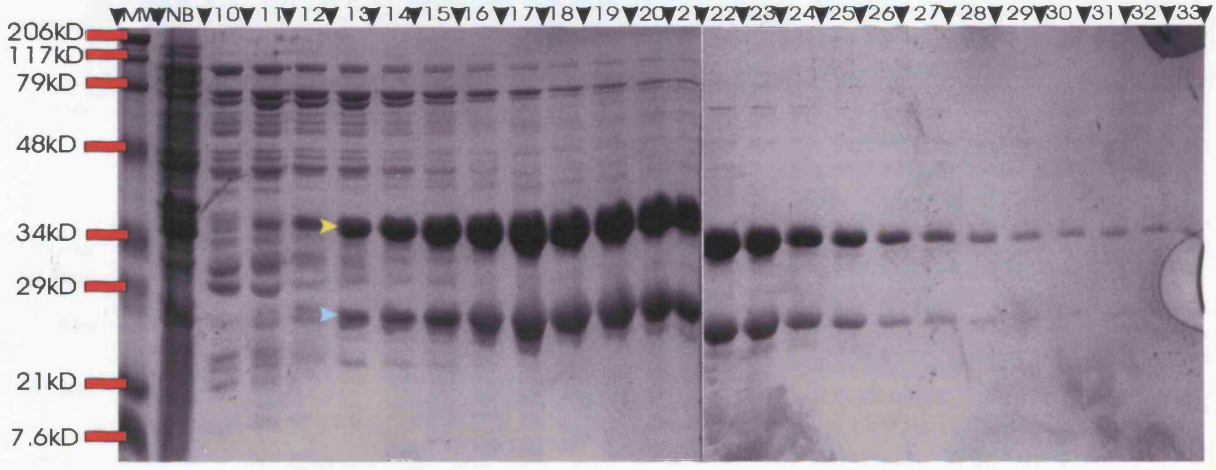


Figure 5.10 B, large scale purification run for *Md*PER5. MW= Molecular weight markers, NB= Not bound protein, 10-13 60mM elute, 14-33 = 200mM eluted fractions with a yellow arrow marking full length *Md*PER5 and a blue arrow marking a degradation product.

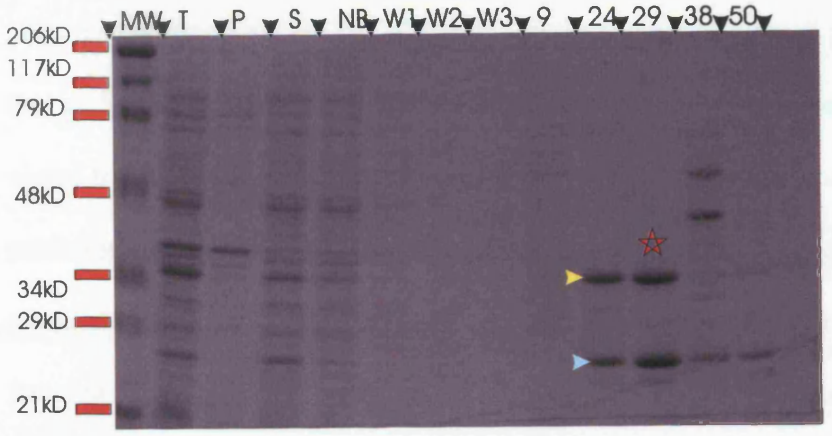
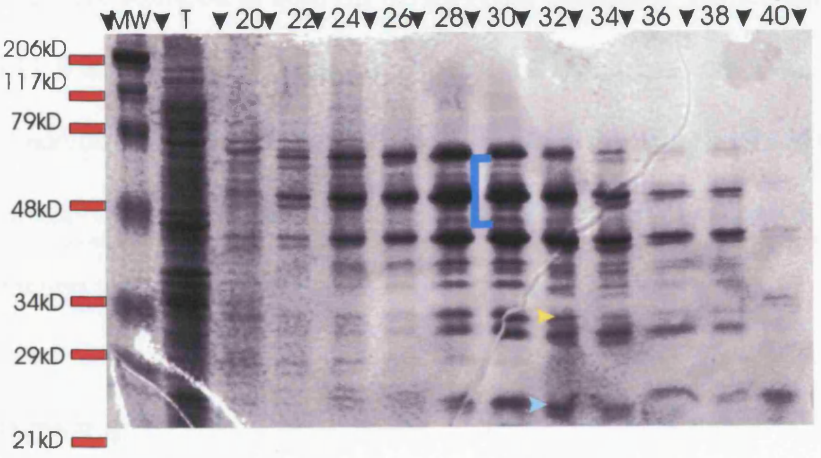
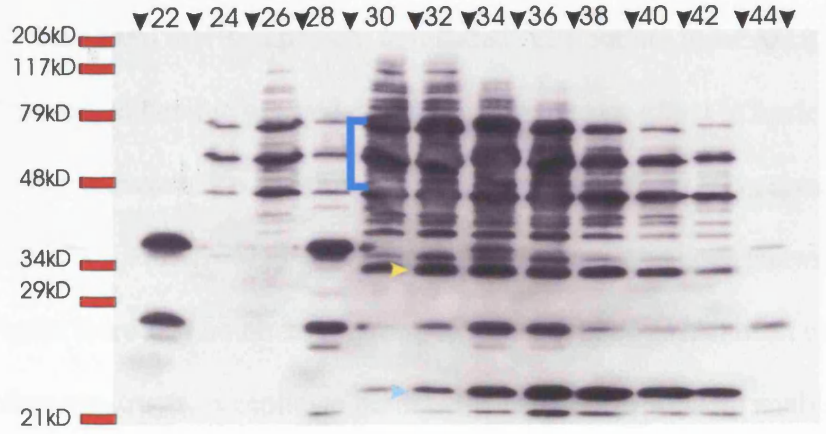


Figure 5.11A, a quick ½l purification run of *Md*PER5. MW= Molecular weight markers, T= Total induced cell protein. P= Insoluble cell protein, S= Soluble protein, NB= Not bound protein, W1-3= Batch washes, 9-50 = Eluted fractions. Yellow arrow marks full length *Md*PER5 and blue arrow marks degradation products.



B: SDS-PAGE, Coomassie stain analysis on stored fractions. MW= Molecular weight markers, T= Total induced cell protein, 20-40 = Eluted fractions, yellow arrow marks full length *Md*PER5 and blue arrow marks degradation products.



C: Western blot with anti-his-tag antibody on stored fractions. 22-44 = Eluted fractions. Yellow arrow marks full length *Md*PER5 and blue arrow marks degradation products. Blue bracket marks multimers common to SDS-PAGE and western analysis.

(**Fig. 5.11A**, fractions 24-29), although, oddly, there was a small peak of multimers that eluted later (fraction 38). Protease and bacterial inhibitors were added to the eluted fractions and were stored overnight at 4^0C. SDS-PAGE analysis was then performed on the fractions (**Fig. 5.11B**). Although quick lysis had reduced the number of contaminating multimers overnight storage resulted in larger multimers forming. The multimers were his-tag containing as they cross reacted with the anti-his-tag antibody (**Fig. 5.11C**). A group of three strong bands between 49 and 79 kDa were common to both the SDS-PAGE Coomassie stained gel and blots (**figure 5.11 B** and **Fig. 5.11C** respectively, marked with a blue bracket). The large number of additional bands that occur in both illustrate the complexity of the multimer formation that must have been occurring. The red star in figure 5.11 marked the fraction used for antibody production (detailed in chapter 4).

***Md*PER 6.**

Initial analysis showed that *Md*PER 6 did not express well as there was no obvious band that was present in induced cells but not induced cultures (**Fig. 5.7**). This was either due to rapid degradation, or a toxic effect in bacteria, that prevents good expression. To test this, the induction of *Md*PER6 was repeated at both 18^0C and 37^0C, with the SDS-PAGE analysis of these inductions shown in figure 5.12A. Again there was no obvious strong expressing band as had been produced for the other constructs. A replicate gel was then subject to western analysis with the anti-his-tag antibody and developed with an alkaline phosphatase colour reaction, (**Fig. 5.12B**). The anti-his-tag antibody produced weak bands at around 21kDa, close to the 18 kDa predicted for *Md*PER6. In addition the western blot shows that *Md*PER 6 (marked with yellow arrows) expresses best at 18^0C and is mainly found in the

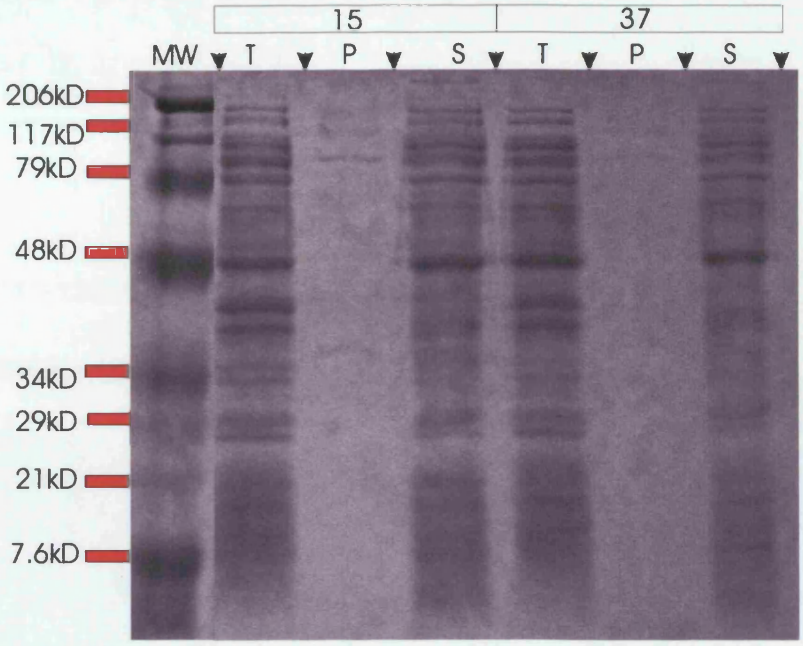


Figure 5.12A, SDS-PAGE Coomassie analysis of *Md*PER6 inductions.
MW= Molecular weight markers, T= Total protein, P= Pellet protein, S= Soluble protein, 15= 15^{0}C induction, 37= 37^{0}C induction.

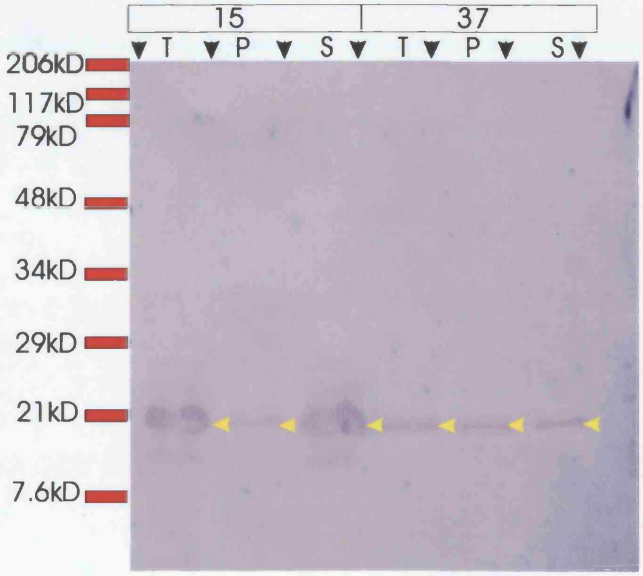
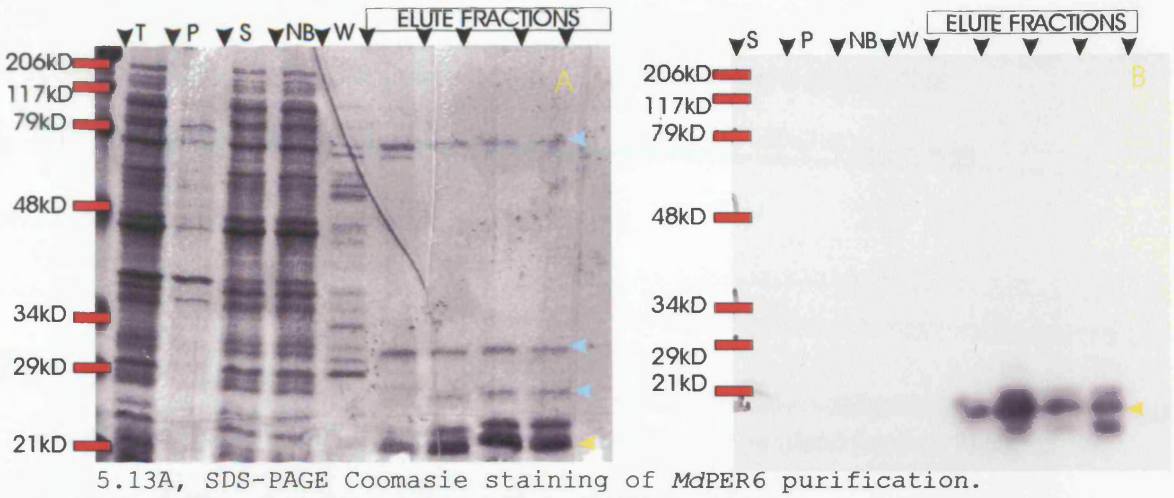


Figure 5.12B western analysis of a replica gel using a anti-his-tag antibody and alkaline phosphotase colour reaction for visualisation. Yellow arrows mark full length *Md*PER6.
T= Total protein, P= Pellet protein, S= Soluble protein, 15= 15^{0}C induction, 37= 37^{0}C induction.

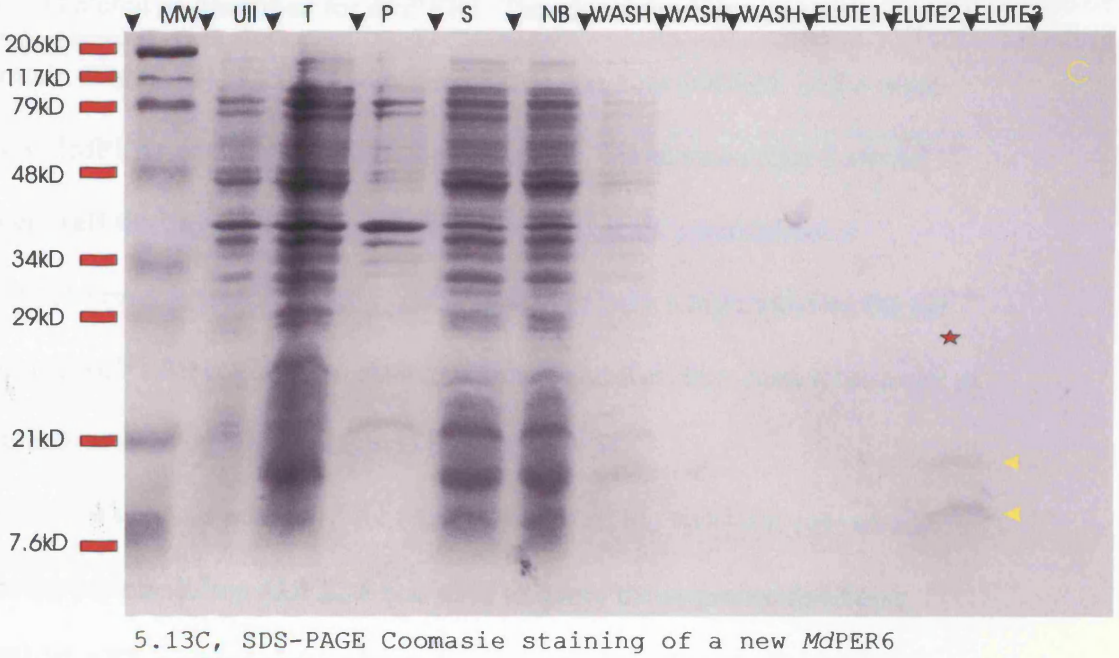
soluble fraction. There were also a few minor degradation bands surrounding the main band.

Although MdPER 6 expressed at low levels this did not preclude purification of the protein for antibody production. Inoculation only requires small quantities of pure protein. However the poor expression levels would prevent structural analysis as large volumes of expressing cells would be required. Purification was therefore attempted with a view to producing protein for inoculation. 1.5 litres of cultures were grown to 0.8 OD and moved to 18^0C for one hour prior to induction. The cells were induced to express overnight at 18^0C and harvested. Once lysed the cells were then bound to 2 mls of Ni-NTA and eluted with 200mM imidazole. Figure 5.13A shows SDS-PAGE Coomasie stained protein analysis of the purification run. The eluted fractions contained the strongly stained bands at around 21kDa, corresponding to full length *Md*PER 6 (marked with yellow arrows). However there were contaminating bands at high molecular weights (blue arrows). Western analysis of a replica gel with anti-his-tag antibody confirmed that the major band at 21kDa was indeed due to *Md*PER6 (**Fig. 5.13B**). However the high molecular weight bands did not cross react with the antibody and were due to bacterial protein impurities and were unacceptable for inoculation. The impurities were likely to have occurred through the lack of his-tagged proteins competing for the immobilised Ni ions, therefore weakly binding bacterial protein could occupy vacant Ni ions. In order to circumvent this problem the ratio of Ni-NTA to lysed culture was reduced. A bacterial pellet from 1.5 litres of culture was lysed and bound to 0.5 mls of Ni-NTA, representing a four fold increase in the culture volume to Ni-NTA ratio. The bound proteins were then eluted and subject to SDS-PAGE (**Fig. 5.13C**). There were no contaminating bands at high molecular weights. Two



5.13A, SDS-PAGE Coomasie staining of *Md*PER6 purification. Yellow arrow marks full length *Md*PER6 and blue arrows mark contaminating proteins.
T= Total protein, P= Pelleted protein, S= Soluble protein., NB= Not Bound, W= Batch wash

B: western analysis of replica gel with anti-his antibody and ECL detection.
T= Total protein, P= Pelleted protein, S= Soluble protein, NB= Not Bound, W= Batch wash. Yellow arrow marks full length *Md*PER6.



5.13C, SDS-PAGE Coomasie staining of a new *Md*PER6 purification.
UI= Uninduced protein T= Total protein, P= Pellet protein, S= Soluble protein., NB= Not Bound, WASH= Batch washes. The red star marks faction used for inoculation.

bands between 21kDa and 7.6kDa (marked by yellow arrows) were eluted, The bands represent full length MdPER6 and a degradation product. This fraction was chosen for antibody production as detailed in chapter 4.

***Md*PER 8**

Initial studies of *Md*PER8 had to be rejected when a clone identification mix up was discovered. Therefore there was insufficient time to perform structural scale purifiation runs and *Md*PER8 was only expressed and purified for antibody production.

50mls of induced *Md*PER8 culture were lysed and bound to Ni-NTA and washed and eluted as described for *Md*PER1. The eluted fractions showed a highly pure protein (**Fig. 5.14**) at 25kDa and this fraction was concentrated in the same manner as *Md*PER1 and 5 and used in the inoculum. Given more time it would have been well worth investing the effort in structural scale purification of *Md*PER8. The protein purified easily and appeared to have a high yield on the gel. Furthermore *Md*PER8 contains no cystein residues and therefore cannot have any of the problems associated with disulphide bond formation.

*Md*PER8 had no published functional allocation but was well conserved amongst the insects. When *Md*PER8 was used to query the sequence databases using PSI-BLAST, a very sensitive homology search algorithm, the homology matches showed the sequence was well conserved within insect proteins but was not found in any other proteins. Therefore it would be difficult to convince a structural biologist of the value of expending effort on *Md*PER8.

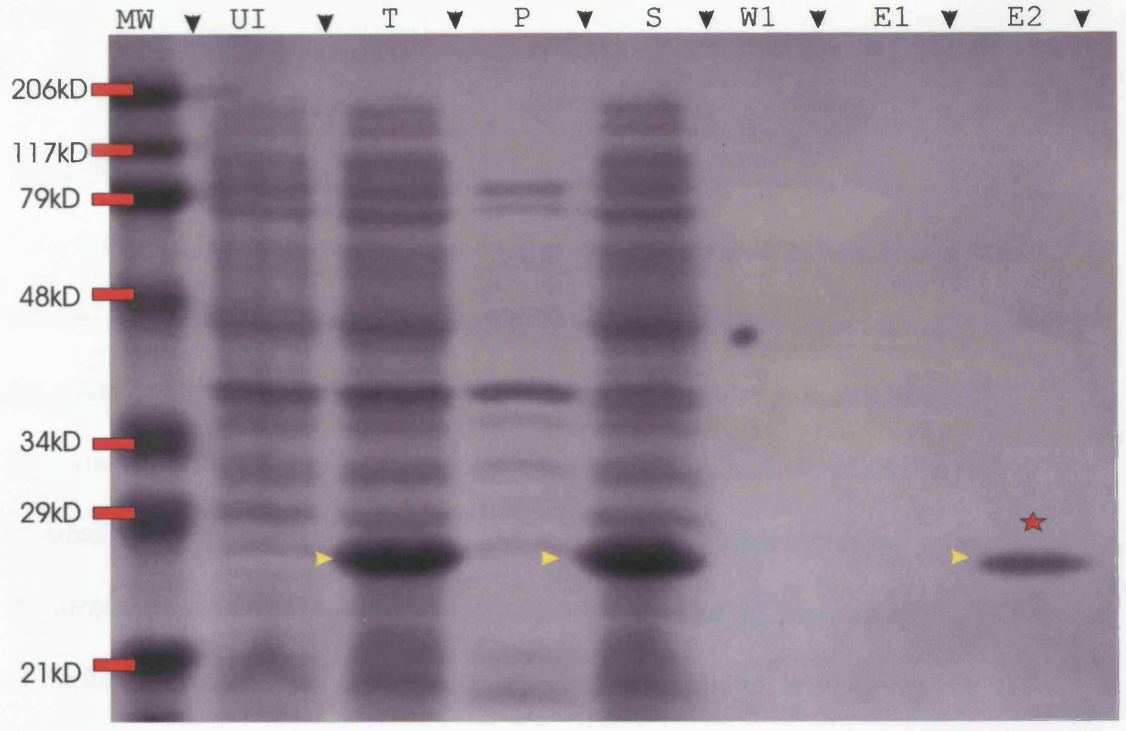


Figure 5.14 SDS-PAGE Coomasie stain of small scale *Md*PER8 purification. UI = uninduced cell pellet, T = Total induced protein, P = pelleted proteins, S = soluble proteins, W1 = batch wash, E1 = 60 mM Elute, E2 = 200mM Elute, Yellow arrows mark the full length *Md*PER8 bands and the red star marks the fraction used for the inoculation.

Insoluble proteins

One of the problems with producing eukaryotic proteins in *E.coli* is that the product frequently accumulates as insoluble aggregates. These insoluble aggregates are packaged into inclusion bodies, which occur as granular particles within the bacterial cell. Once the cell has been broken open the inclusion bodies can be isolated as a pellet by centrifugation. The main problem with inclusion bodies is that the proteins form inappropriate folded states that remove their native activity. One of the key elements in the formation of non-native structures is the formation of inappropriate disulfide bonds between cystein residues. Disulfide bonds often stabilise the native structure of eukaryotic proteins. However bacterial cells have a highly reducing environment and this often leads to inappropriate disulfide bond formation and non-native structure (Fischer et al., 1992). Despite the obvious disadvantage of non-native structure formation, inclusion bodies offer the advantage that they can be purified away from cellular proteins very effectively and efficiently, and the by the use of simple lysis and centrifugation can achieve greater than 60% purity common (Guise et al., 1996). Furthermore, because the proteins are packaged away from the rest of the cell contents they have no toxic effect and accumulate to extremely high levels, often constituting more than 50% total cell protein (Guise et al., 1996). Therefore inclusion bodies provide a protein purification step by which highly concentrated and relatively pure protein can be achieved (Guise et al., 1996; Lin and Cheng, 1991).

Protein purification has two goals, firstly to achieve high concentrations of pure protein, secondly to retain the native activity of the recombinant protein. Inclusion bodies are very useful in achieving the first goal, to achieve the second

goal requires protein refolding into the native state. Understanding protein folding is one of the major challenges facing biology, and represents a burgeoning field that that is only beginning to understand how a simple polypeptide chain can fold into a complex structure. Refolding is a difficult task and has been likened to "the unboiling of an egg" (Guise et al., 1996). Reviews of refolding recombinant proteins from inclusion bodies identify common approaches to refolding proteins: isolation by sonication/pressure lysis, often accompanied by lysozyme to aid the lysis and DNase treatment to remove contaminants, renaturisation by urea/GnHCL chaotrophic reagents, removal of denaturants by dilution or dialysis which allows folding to commence and breaking and reformation of disulphide bonds by oxidising reagents(Fischer et al., 1993; Guise et al., 1996). There are however numerous approaches to achieving soluble folded proteins, a few of these approaches have been reviewed and listed in the table 5.4. In order to maximise the likelihood of refolding the two insoluble proteins (*Md*PER3 and 7) as wide a range of refolding techniques as possible were trialed. Methods were selected on two criteria; their ease of application and whether they where significantly different from the other methods selected. The rationale being that a wide range of methods were more likely to be successful.

*Md*PER3 contains two PAS domains along with the PAC domain. A protein module of two PAS domains followed by a PAC domain is used in at least one of the circadian clock genes of all eukaryotic clocks studied. Therefore most effort was expended on refolding *Md*PER3. In contrast MdPER7 does not appear to have any functional homologues. Furthermore when the sequence is used to query database with the BLAST algorithm most of the sequence is removed by the low complexity

filter. Therefore the sequence is unlikely to have an ordered structure and less effort was expended on refolding *Md*PER7.

Refolding methods

The technique of Lin *et al* (Lin and Cheng, 1991) was attempted first. The technique was carried out exactly as published. During the shock dilution of denatured inclusion bodies into the renaturation buffer aggregation was observed. The sample was then centrifuged and the supernatant analysed for soluble protein by Bradford assay (see methods section). No soluble protein was detected and the technique was rejected. However prior to unfolding and renaturisation steps, the protocol introduces techniques for removing impurities form inclusion bodies. Fresh, unfrozen bacterial cell pellets were resuspended in $1/20^{th}$ culture volume of extraction buffer (20mM Tris-HCl/pH7.5, 20% sucrose, 1mM EDTA) and incubated on ice for 10min. Cells were pelleted by centrifugation and resuspended in ice cold distilled water, then recentrifuged at 800g to remove cell walls from the spheroplasts. The pelleted spheroplasts were resuspended in (PBS without Mg^{2+} and Ca^{2+}, 5mM EDTA, + protease inhibitors, without β-mercaptanethanol). Lysis was carried out as usual. The inclusion bodies were then washed three times by resuspension in (PBS without Mg^{2+} and Ca^{2+}, 5mM EDTA, 25% sucrose, 1% Triton-X-100), incubation on ice and recentrifugation. These steps allowed a high degree of purification as demonstrated in figure 5.15. Small quantities of bacteria could be lysed efficiently to produce relatively pure inclusion bodies. Although this could produce enough protein for antibody production, larger scale cultures were required for structure scale protein purification. Larger scale lysis was less efficient, which resulted in inclusion bodies contained with unlysed bacterial cell proteins

(**Fig. 5.15A**, lane P, blue bracket marks contaminants). However these contaminants were removed by the purification protocol stated above (**Fig. 5.15A**, lane X note removal of contaminants when compared to lane P), and this technique was used to obtain enough pure protein with which to commence structural scale refolding experiments.

Denaturing Ni-NTA affinity chromatography and refolding.

Ni-NTA binding of denatured inclusion bodies containing the his- tagged expressed fragment was trialed with both *Md*PER 3 and *Md*PER7. The methods detailed in Rogl *et al* (Rogl et al., 1998) were followed faithfully. However in addition to urea denaturation, GnHCl (6M) was also used to replace urea as the denaturant in the solution as there is some evidence that GnHCl is a better denaturant which results in more active protein (Fischer et al., 1992). 2.5mg of inclusion bodies from both constructs were denatured and bound to Ni-NTA matrix. The bound matrix was washed to remove the denaturant and the eluted with 200mM imidazole. The yield of soluble product was determined by Bradford assay and the results presented in table 5.3. Although 2.5mg of protein was used in each experiment it was obvious that this is a miscalculation as significantly more protein was found in the combined, not bound and elute fractions. However inclusion bodies were difficult to quantify, as they did not easily form a uniform suspension. In addition it was possible that some of the protein is shielded from the Bradford assay when inside the inclusion bodies. However measuring inclusion body protein content with the Bradford assay still allowed an approximate equalisation of the quantity used. Figure 5.16A and B show the refolding of MdPER 3 and MdPER 7

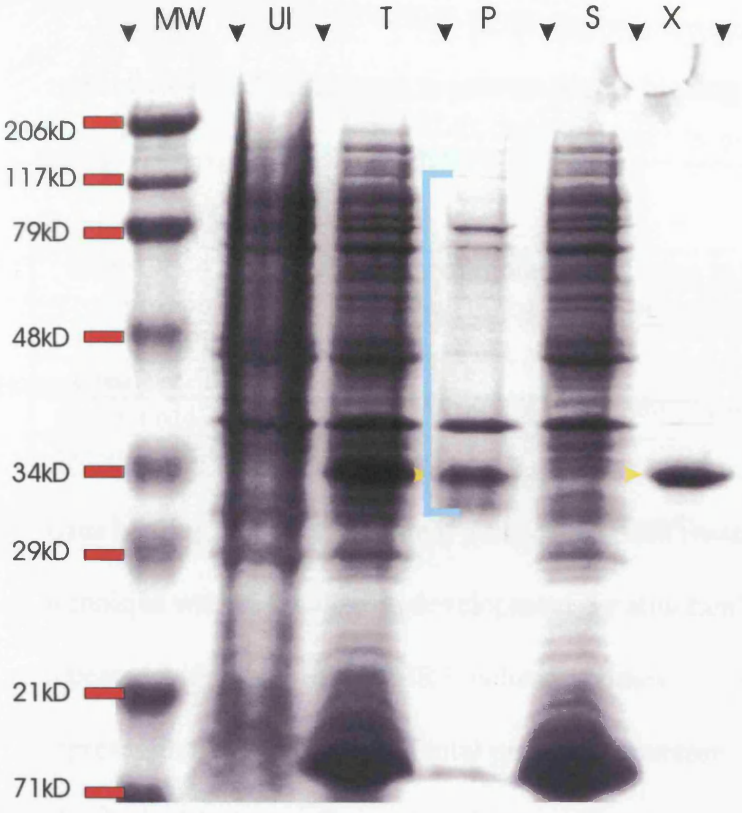


Figure 5.15A, SDS-PAGE Coomassie stained large scale *Md*PER3 purification. MW= Molecular weight markers, UI= Uninduced protein, T= Total protein, P= pellet, S= soluble, X= purified inclusion bodies. Full length *Md*PER3 is marked by a yellow arrow and contaminating proteins are marked by a blue bracket.

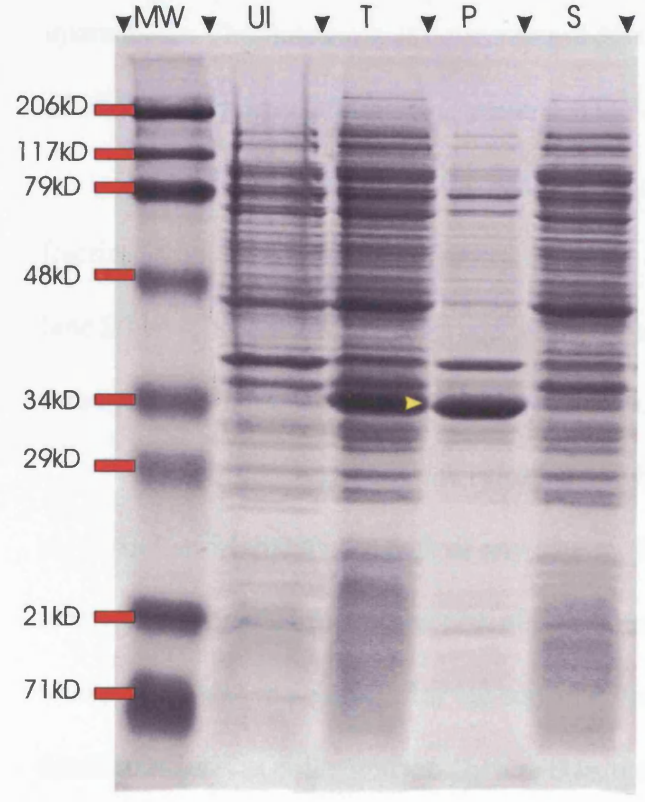


Figure 5.15B, SDS-PAGE Coomassie stain of trial *Md*PER3 lysis. MW= Molecular weight markers, UI= Uninduced protein, T= Total protein, P= pellet, S= soluble. Full length MdPER 3 is marked by a yellow arrow.

respectively. GnHCl appears to prevent proper binding to the Ni-NTA matrix as the protein is removed in the washes.

Table 5.3

Denaturant and protein	Quantity of inclusion bodies in mg	mg in not bound fraction	mg eluted in soluble form	yield
Urea MdPER 3	2.5	3.4	1	23%
Urea MdPER 7	2.5	3.4	2	46%
GnHCl MdPER 3	2.5	3.7	0	0%
GnHCl MdPER 7	2.5	4.3	.6	12%

Urea binding achieved excellent purification with reasonable yields, therefore this technique was well suited to development for structural trials. The technique was repeated with 12mg of *Md*PER3 inclusion bodies. 1.4 mgs was eluted in 15 mls representing a 12% elution of total protein, a decrease in yield from the 23% obtained with the smaller scale refolding. The eluted fraction was concentrated by a Amicon centricon 10 kDa cut off concentrator, used according to the manufacturer's instructions. The fraction was concentrated down to 0.8mls; a concentration factor of 18, however only 0.34mg was present in the concentrated fraction.

SDS-PAGE analysis of the Amicon concentration run shows that the eluted fraction is not significantly concentrated by the column (**Fig. 5.16C**, comparison of lane E to MC), *Md*PER3 is absorbed by the column as it is not present in the flow through but is not concentrated in the retained sample.

It is well documented that refolded yields decrease on scale up (Shimamoto et al., 1998). Unfortunately the lack of any appreciable concentration by the Amicon centricons prevented the production of concentrated pure protein; only 3% of the total protein was recovered after the concentration step, and the level of concentration was still significantly less than the 10mg/ml required for structural studies. Consequently, only when a technique for efficiently concentrating the proteins becomes available will structural studies become possible. It is worth

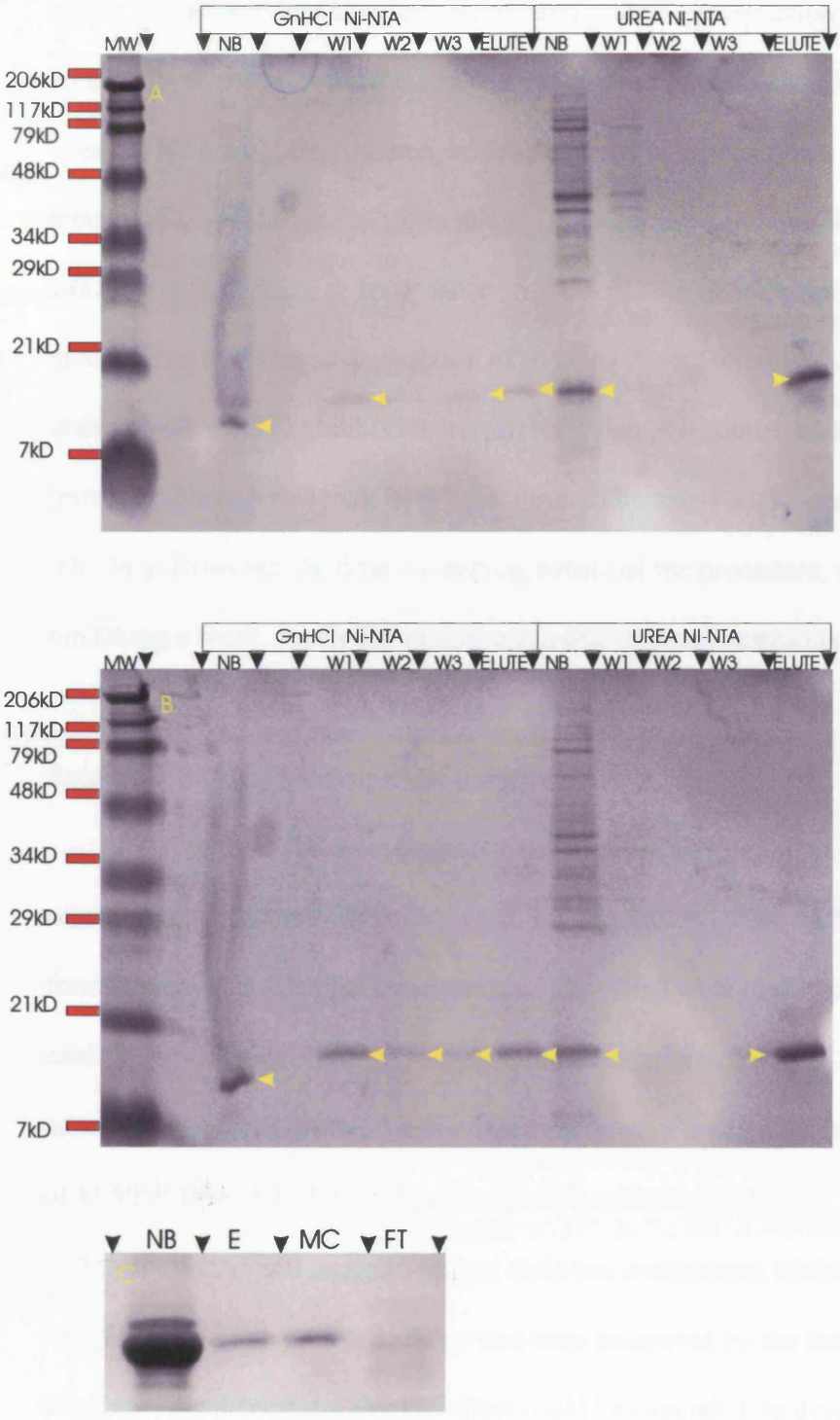


Figure 5.16A and B, SDS-PAGE Coomassie stained GnHCL and Urea Ni-NTA purification and refolding of *Md*PER3 and *Md*PER7 respectively.

MW= Molecular weight markers, NB= Not bound, W1-3= Washes, Elute= 200mM imidazole. *Md*PER3 and *MdP*ER7 are marked with yellow arrows. Note the high degree of purity of eluted bands. Also note GnHCL distorts the running of the NB sample through the cell matrix resulting in smearing.

C SDS-PAGE Coomassie stained microcon concentration of eluted *Md*PER3 fraction. MC= Microcon concentrate, FT= Microcon flowthrough.

noting that ammonium sulphate and ethanol protein precipitation methods for concentrating proteins were also trailed without any success. In addition to the Urea/Ni-NTA refolding method, which achieved most success, a glycerol based method (Shimamoto et al., 1998) and the urea dialysis method (Maniatis et al., 1981) were performed in accordance with the published techniques. The glycerol method lead to obvious aggregation of proteins during renaturation and a undetectable soluble yield. The urea dialysis also procedure yielded no soluble proteins, although this may have been through bacterial degradation during refolding. However the time consuming nature of the procedure, with each refolding run taking a week, precluded optimisation due to time constraints.

Summary of MdPER fragment purification

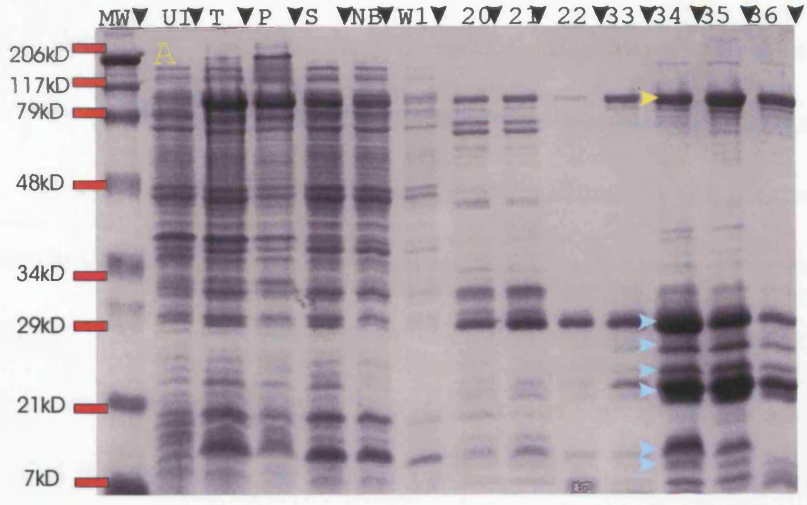
It was clear that each fragment still requires development of a purification scheme in order to provide pure protein for structural studies. *Md*PER1, 3 and 5 all form strong candidates for structural analysis. They were highly expressed and were easily purified away from the majority of contaminating bacterial proteins. In addition these three protein fragments cover most of the known functional domains of *Md*PER (**Fig. 5.2**). However, after initial purification, *Md*PER 1 and 5 formed multimers that would mitigate against their use in structural trials. *Md*PER 3 could be refolded but attempts to scale up had been hampered by the lack of a suitable concentrating procedure. Further effort could be expended on developing their purification procedures, however developing the purification of all three would be a time consuming process.

As a solution a new fragment that covers all three was trailed for purification. The fragment, termed *Dm*PERX, spans the start methionine to the start

of the TG regions *of D. melanogaster* PER and had an N-terminal His-tag introduced. Initial expression studies, as part of a parallel project undertaken in this laboratory, had shown that the fragment was soluble and well expressed (Jon Clayton, pers. comm.). Jon Clayton then donated the fragment for further Ni-NTA purification trials.

***Dm*PERX purification.**

Two litres of induced culture were lysed, bound and eluted as described in the methods section, the gel analysis of the purification was presented in figure 5.17. Lysis was not complete as there were many protein bands in the pellet (**Fig. 5.17A**, lane P). This was a common occurrence amongst the large cultures where complete lysis was difficult to achieve. However full length *Dm*PERX was purified with contaminating degradation products (lanes 33-36). There was still a significant amount of *Dm*PERX present in the Not Bound fraction. Therefore this fraction was rebound and eluted in a second purification run (**Fig. 5.17B**). The eluted fractions were relatively free from degradation products, which had obviously been depleted by the previous purification run. Maximum protein recovery was required for structural analysis, therefore gel filtration was performed to separate the full length protein from the degradation products in the eluted fractions from the first purification. To achieve good resolution during gel filtration the protein was concentrated into a smaller volume. Concentration was attempted with fraction 35 (**Fig. 5.17C**). There was aggregation of protein during overnight storage as shown by the reduction in staining from 35 to 35'. The sample was then concentrated in an Amicon microcon 10kDA as detailed previously. However this resulted in a significant loss of protein (**Fig. 5.17C** comparison of 35' to lane C), as there were no proteins in the flow through (lane F) and so loss must be due to binding of the



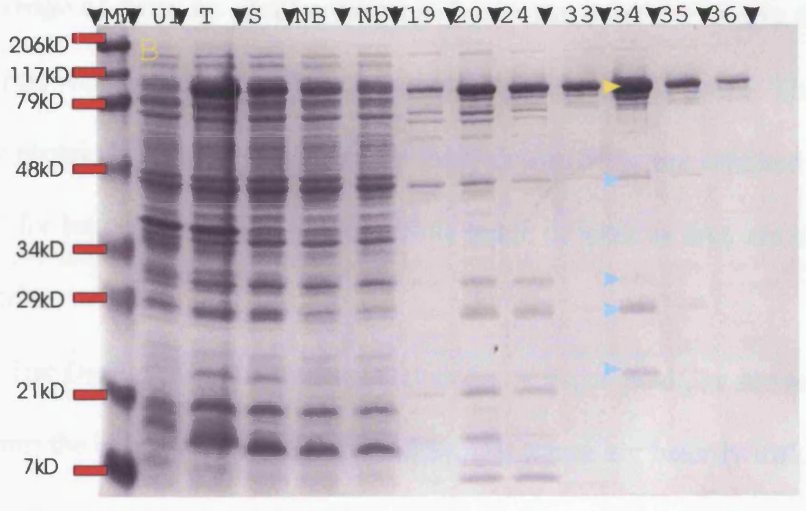
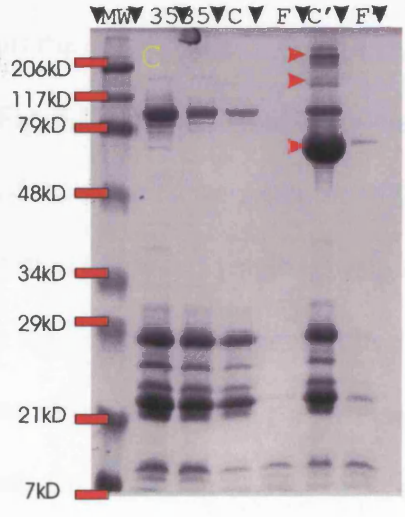


Figure 5.17A SDS-PAGE Coomasie analysis of a 21 culture of *Dm*PERX lysed bound and eluted from a Ni-NTA column. Mw= molecular weight markers, UI=uninduced protein, T= Total protein, P=Pelleted protein, S= Soluble protein, NB= Not Bound protein, W1= Batch Wash 1, 20-21= 60mM elute, 22-36= 200mM elute. Full length *Dm*PERX was marked with a yellow arrow and degradation products marked with blue arrows.



B repeat purification where the Not Bound fraction from the 1st run was bound to Ni-NTA to further capture *Dm*PERX. NB'= second Not Bound fraction. 19-20= 60mM Elute, 24-36= 200mM Elute .

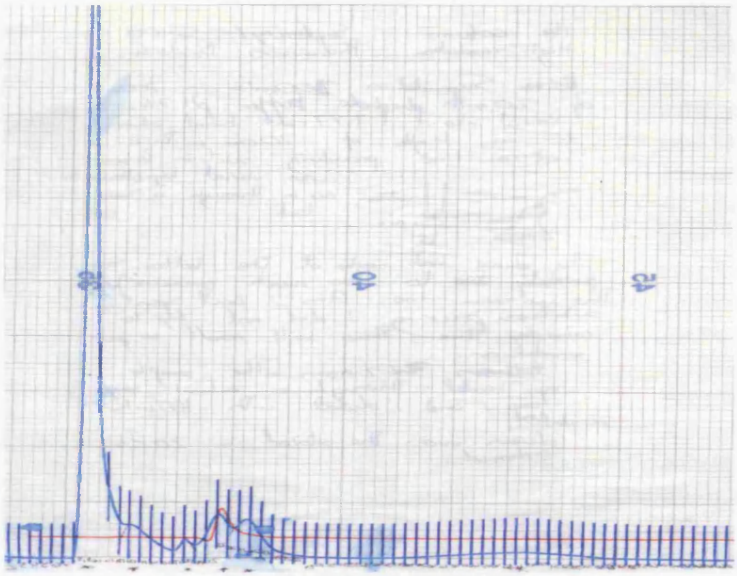
C, fraction 35 from the 1st purification run was concentrated prior to gel filtration. Mw= Molecular weight marker, 35 = fraction 35 before overnight storage, 35'= after over night storage. C= microcon concentrate. F= flowthrough. C'= microcon concentrate with BSA blocking. F'= flow through following blocking.

protein to the filtration membrane. The Amicon microcon was then blocked with BSA and the procedure repeated. This resulted in less protein loss (comparison of C to C'), however as the column was not washed sufficiently the concentrated fraction was contaminated by BSA (red arrows). Fractions 33-37 were then concentrated with the blocked Amicon microcon and applied to the gel filtration column to separate the full-length fraction from the smaller degradation products.

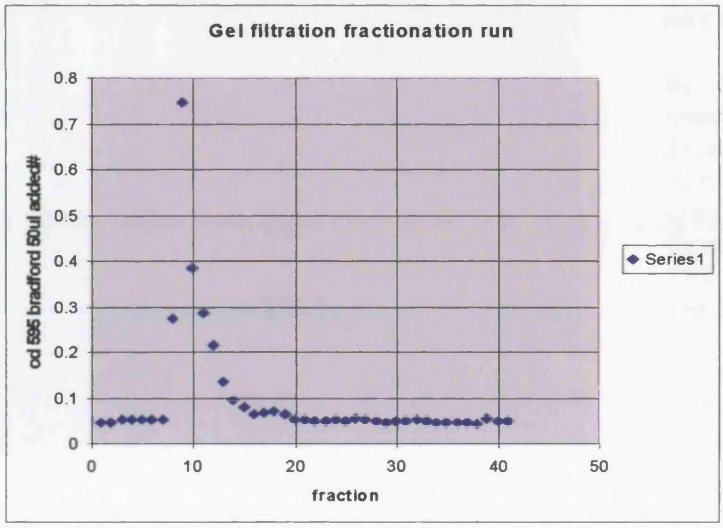
Gel filtration relies on a packed column of matrix; the matrix contains pores with a range of sizes, so small proteins and salts can enter the matrix through all pore sizes, whereas larger proteins enter only through larger pores. Therefore the smaller proteins and salts that enter the matrix more often are retained on the column for longer; the larger proteins elute much quicker as they are not retained by the matrix.

The *Dm*PERX gel filtration produced one major peak, as shown in the UV trace from the column elution (**Fig. 5.18A**). UV traces are heavily influenced by aromatic residue content of the eluted proteins and are also influenced by DNA eluted off the column. Therefore the eluted fractions were checked by Bradford assay (**Fig. 5.18B**) which confirmed that protein had been eluted off in the major peak, and the samples were subject to SDS-PAGE analysis (**Fig. 5.19C**). There was no separation of the full length protein from its contaminants, due either to binding between the full length protein and its degradation products or because the resolution of the matrix was not sufficient to separate the proteins. To test the last hypothesis Gel Filtration was repeated with greater resolution and a fresh culture was lysed. Purification by Ni-NTA again resulted in a mixture of full length *Dm*PERX and degradation products (SDS-PAGE analysis, **Fig. 5.19A**).

A



B



C

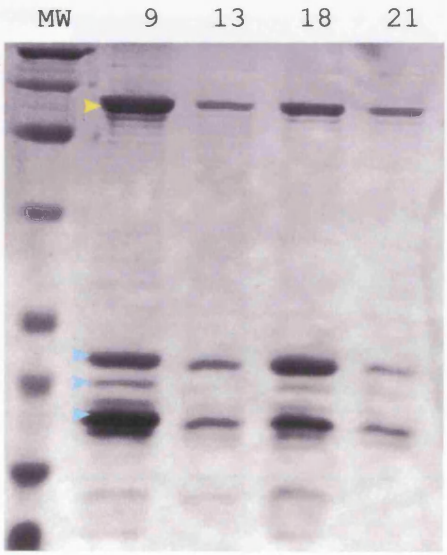
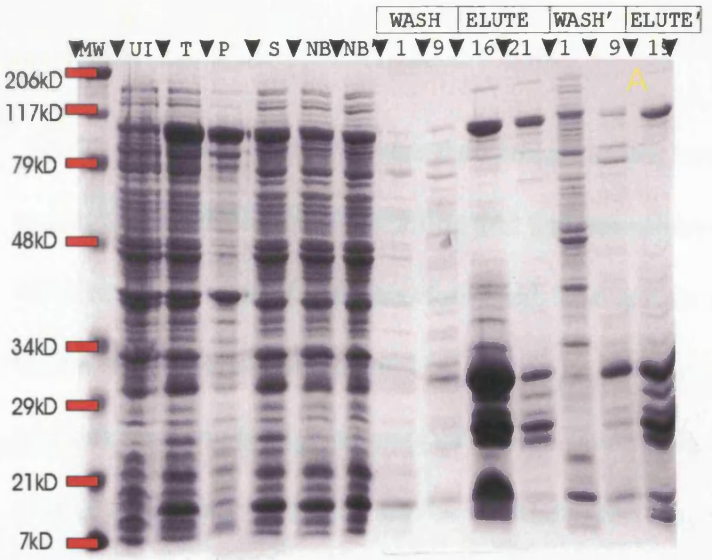
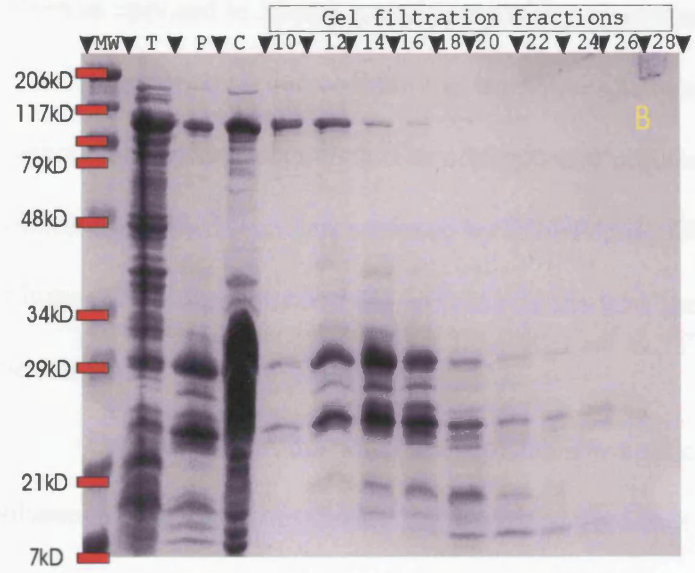


Figure 5.18 Gel filtration using a 1.75cm^2 diameter, 20cm long Biorad LP column packed with Sephacryl s-100 HR matrix.
A, UV trace of eluted fractions with one major peak visible at around fraction nine and with four minor peaks in the next ten fractions.
B biorad protein determination of eluted peaks confirms UV peaks.
C, SDS-PAGE coomasie stained analysis of gelfiltration fractions. Full length *Dm*PERX marked with yellow arrows and degradation products marked with blue arrows.





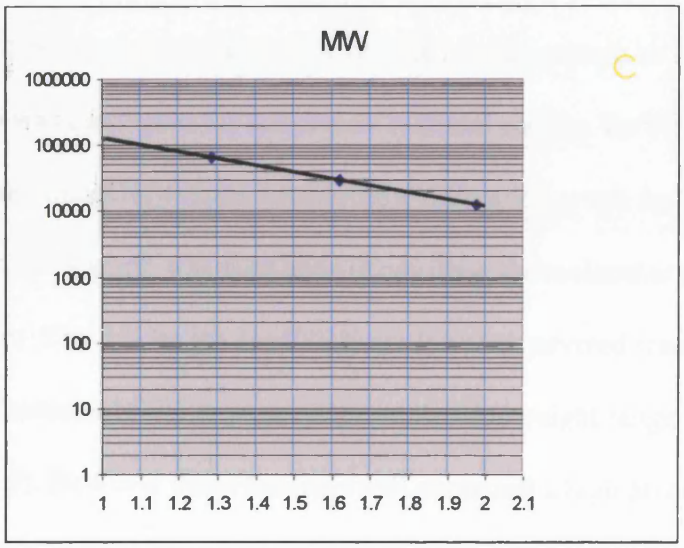


Figure 5.19A, a 21 culture of *Dm*PERX was lysed bound and eluted from an Ni-NTA column. The not bound fraction was rebound and eluted again.
MW= molecular weight markers, UI=uninduced protein, T= Total protein, P=Pelleted protein, S= Soluble protein, NB= Not Bound protein, NB'= Not Bound second Elute., WASH= 60mM elute, ELUTE= 200mM elute, WASH'= second binding 60mM elute, ELUTE'= 200mM elute.
Full length *Dm*PERX was marked with yellow arrows and degradation products marked with blue arrows.

B, Gel filtration of the combined 200mM eluted fractions from purification runs shown in A.
Concentration of the fractions was achieved using PEG dialysis and is shown in lane C.
Gel filtration was performed with sepahacryl s-100 HR matrix packed in a 50cm long, 1.5inch internal diameter low pressure column. The collected fractions then analysed by SDS-PAGE Coomassie staining.

C= Graph of calibration of the gel filtration column.

The eluted fractions were then concentrated, however the Amicon centricon method had proven time-consuming and had not achieved high concentrating effects. A new method was therefore tried. The sample was placed in dialysis tubing with a low molecular weight cut off and dialysed against solid PEG (see methods section) resulting in high degrees of concentration (**Fig. 5.19B**, lane C). Therefore a small concentrated sample could be applied to the longer gel filtration column (50cm as opposed to 20cm), greatly improving resolution.

The Gel filtration column had been previously calibrated using protein size markers (see methods section). The concentrated protein was loaded on to the column and eluted fractions analysed by SDS-PAGE (**Fig. 5.19A**). The eluted volumes of the fractions and the protein size markers are presented in table 5.5, with the calibration graph presented in 5.19C.

The eluted proteins' apparent molecular weight can be calculated by the volume in which it elutes during gel filtration. Vo describes the volume at which a protein that is not able to enter the matrix is eluted off; given by blue dextran elution volume, Ve describes the elution volume of each protein or eluted fraction. The Ve/Vo was calculated for the protein markers and the Ve/Vo plotted against the markers molecular weight, producing a calibration graph for the column (**Fig. 5.19A**). This graph was then used to calculate the molecular weight of the eluted proteins. The full length *Dm*PERX elution peak covered fractions 10-12, and tailed off in fractions 14-16, representing a molecular weight range of 90-60,000 kDA (see table 5.5). However these fractions still contained a high proportion of degradation products. The degradation products migrated at less than 30,000 kDa during SDS-PAGE and should elute at a Ve/Vo greater than 1.62 (more than 22mls). Therefore *Dm*PERX should mainly elute between fractions 10-12 and the degradation product

should elute after fraction 14. However degradation products occurred in fraction 10 and were prevalent in fraction 12. Therefore during gel filtration the degradation products behaved as 90-60, kDa proteins, not the 29-25kDa that they migrated on during SDS-PAGE.

I propose that the degradation products bind to form multimers increasing their size during gel filtration, and moving the elution profile into earlier fractions. Therefore instead of having two non-overlapping profiles of *Dm*PERX and degradation products as should occur, the two elution profiles overlap considerably. However the binding was not strong enough to be maintained under SDS-PAGE denaturing conditions as there was no evidence of multimer formation. Therefore the Gel Filtration was repeated with the addition of 1M Urea in the running buffer. 1M urea was chosen as it should not completely unfold the proteins but should disrupt protein binding interactions. However the elution profiles of the degradation products and full length *Dm*PERX still overlapped. Therefore gel filtration was unable to provide high recovery of pure full length *Dm*PERX.

Structural analysis through Circular Diochroism

The main goal of the purification of the fragments was to provide enough pure protein for structural studies. However, either insufficient protein was obtained or it could not be purified away from contaminants. Achieving pure protein would obviously require a large commitment of dedicated effort. Therefore it was decided to assess whether the proteins that had been expressed were sufficiently structured for crystallisation and structural trials. As a quick means of assessing structure the fragments were purified by denaturing Ni-NTA chromatography.

Introduction

Absorption spectra were used to provide information on the purity and concentration of the purified MdPER fragments and Circular Diochroism (CD) was used to assay the folded state of the fragments.

Absorption Spectra

Although it is common for the 260/280nm ratio to be used to determine whether protein or DNA is responsible for absorption it is much more informative to consider the whole absorption spectra.

Phenylalanine absorbance is weak and gives a spiky profile around 250-260nm and virtually zero absorbance at 270nm; tyrosine and tryptophan contribute much more around 280nm. The peak absorbance of DNA is at 260nm with a sloping base line from 310-400nm indicating aggregates. Therefore spectra that peak at 260nm indicate DNA contamination although spiky profiles, around 260nm, may be due to phenyalanine. Strong peaks at ~280nm, and a low 260/280nm ratio indicates pure protein, and the contribution of aggregates can be determined by extrapolating the sloping 310-400nm baseline back across 250-300nm and removing it from the signal (Creighton, 1997).

Fragments *Md*PER1 and 5 showed slightly sloping baselines, indicating slight aggregation. The slopes were extrapolated and the aggregates' contribution to the spectra removed for the purpose of calculation of purity and concentration. Initial analysis of the spectra showed that most of the fragments had high 260/280 ratios indicating DNA contamination. However all of the fragments used did not contain tryptophan residues and had relatively high proportions of phenylalanine

(see table 5.6). Therefore phenyalanine was likely to be responsible for the high 260/280 ratio as DNase was added during the lysis and should have acted to prevent DNA contamination of the purified proteins. The 280nm absorption was used in conjunction with the molar extinction co-efficient (table 5.6) to calculate the protein concentration and the proteins were adjusted to 3μM for CD analysis.

Circular Diochroism

Proteins absorb light through peptide groups and aromatic side chains. When a chromophore is part of an asymmetric structure then it absorbs left or right handed polarised light to different extents, this is called Circular Diochroism or CD

CD can be carried out, in solution, on minute amounts of protein and can be used to evaluate the secondary structure of a protein; giving information on secondary and tertiary structures. CD measurements can occur in two regions, the far-UV (170nm-250nm), also known as the amide region, where absorbance is dominated by peptide bonds, and the near-UV (250nm-300nm), which originates from immobilised aromatic side chains. Far-UV CD provides information about the structure of the backbone of the protein and is particularly sensitive to α-helices. In contrast, CD from the near-UV is informative about the aromatic side chains immobilised by secondary structure and is sensitive to precise orientation in a protein. It therefore provides a fingerprint of correctly folded proteins.

In order to determine whether a protein is folded or unstructured the far-UV spectra of the proteins were analysed. Unstructured proteins should give low signals for α-helices, β-turns and β-sheets with the major contribution to the CD being from unstructured backbone. However proteins that have been correctly folded should give signals for α-helices and β-turns or β-sheets from their

structured backbone. Therefore CD provided a convenient assay for the folded state of the proteins. However it assigns proportions of α-helices and β-turns or β-sheets and gives no information on their arrangement. Nevertheless it fulfilled the important task of determining whether the proteins were structured.

CD Method

The absorption differences between left and right polarised light is tiny, therefore the measurement of CD has to be very sensitive. Less than 0.1% of absorption is due to CD so it was important that the measurement was performed very accurately. Several steps were taken to improve the signal to noise ration; Firstly the protein concentration was varied until strong spectra were achieved. Also phosphate buffer was used because it is virtually non-absorbant in the UV range used. In addition the buffer was degassed because O_2 absorbs UV, and N_2 was used to flush the CD to prevent ozone build up that is damaging to the machine and protein. The proteins were centrifuged to remove dust and aggregates that can distort the spectra, and the CD scans were repeated 10 times for each protein and the spectra were averaged. The resultant spectra were then analysed using a neural network program (CDNN) that had been trained on CD of proteins of known structure.

Results

The absorption spectra of *Md*PER5 and 7 are shown in figure 5.22 and their CDNN assignment of structure (given in Appendix I) was compared to that predicted by the PHD structural prediction algorithm. However the CDNN assignment of structure provides information estimates of β-turn content as well as α-helix and β-sheets, whereas the PHD algorithm only provides a prediction for α-

helix and β-sheet content. Therefore the CDNN assignment of random coil and β-turn were summed for the purpose of comparison to the PHD assignment. Typically β-turns are short units of structure found in unstructured loops and are therefore more likely to be assigned as random coil by the PHD algorithm. The comparison of the predicted structure and observed structure is given in table 5.7. The observed percentage of random coiled backbone was less than that predicted by the PHD structural prediction algorithm. Indeed the relative proportions of β-sheet, α-helix and random coil, predicted and observed, were broadly similar. If the protein fragments had demonstrated a much lower proportion of structure than predicted this would have indicated an unfolded state, implicating inappropriate or poor folding during purification. However the parity between the PHD predicted levels of structure and those achieved, suggests that bacterial production and purification resulted in appropriately folded proteins. Therefore structural scale protein purification would be likely to result in structured proteins that crystallise and diffract. Indeed the use of a denaturant during purification did not prejudice against achieving structured proteins, which was one of the main concerns with using denaturants.

Conclusions

*Md*PER1, 5 and *Dm*PERX were all highly expressed soluble proteins that could be purified away from the majority of bacterial proteins. However each fragment was co-purified with shorter degradation products. In *Md*PER1 and 5 these formed multimers with the full length proteins during storage. The binding of these multimers was preserved under SDS-PAGE and therefore must constitute a strong binding. *Dm*PERX and its degradation products did not form the same strong

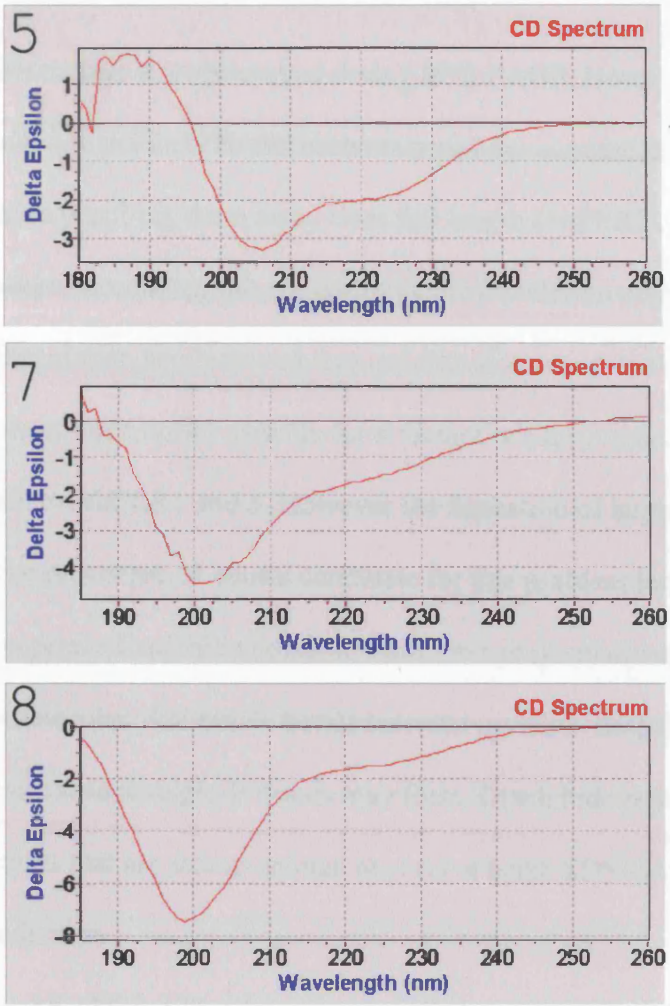


Figure 5.23 CD-spectra of Md PER 5, 7, 8. Proteins identified by number in top left hand corner.

	predicted %	observed %
	Md **PER 5**	
α-helix	24	15
β-sheet	13	25
Random Coil	62	60
	Md **PER 7**	
α-helix	10	10
β-sheet	15	22
Random Coil	75	68
	Md **PER 8**	
α-helix	0	8
β-sheet	8	18
Random Coil	92	80

Table 5.7 Comparison of PHD structural prediction and CD analysis of Md PER 5, 7 and 8.

multimers that were preserved during SDS-PAGE. However the *Dm*PERX degradation products bound each other and the increase in size prevented gel filtration purifying them away from full length *Dm*PERX. Prior to gel filtration, ion exchange chromatography or hydrophobic interaction chromatography may remove the degradation products and then gel filtration could "polish" the full length protein and elute it in a buffer suitable for structural studies. This approach could also be applied to *Md*PER1 and 5. However the formation of large multimeric complexes must be prevented. A causal candidate for this problem is the formation of inappropriate disulphide bonds to form inter-molecular, rather than appropriate intra-molecular, disulphide bonds between cysteins. Both fragments have more than one cystein so disulphide bonds may form. Disulphide bond binding may form multimers that are strong enough to survive harsh SDS-PAGE analysis, a feature of the multimers.

An additive (β–mercaptoethanol) that prevents oxidation of the proteins was added to all the storage buffers. However it may have contributed to the disulphide problem. The anti-oxidisers, β–mercaptoethanol and ditiothreitol (DTT), are common additives to inclusion body refolding experiments, where they function to promote exchange of disuphide bonds with the goal of forming appropriate disuphide bonds (Guise et al., 1996). Removal of these agents may therefore prevent the formation of the multimers. Alternatively approaches for disuphide exchange such as, air oxidation, glutathione/cystein addition or DTT/β–mercaptoethanol addition may help appropriate disuphide bonds form (Guise et al., 1996). However the use of more advanced equipment would help increase the yield and speed of purification. The large cultures required for structural studies proved very difficult to lyse with the lysozyme/ sonication

protocol used, and typically took an hour to achieve poor levels of lysis. Therefore there was opportunity for protease to create the degradation products due to the proteins being exposed to bacterial protease that have spilled from cellular compartments during thawing. The proteases inhibitors can prevent degradation only once the bacteria cell has been lysed. The use of a French press, which lyses large volumes of cells quickly and efficiently, would prevent much of the degradation. In addition, each purification procedure typically took a day on the low pressure system used because of the slow flow rates; a three step protocol would therefore take three days to complete. With a high pressure chromotography system a three step protocol could be carried out in a matter of hours preventing aggregation and multimer build up. With the current equipment *Md*PER 3 represents the best chance of producing pure, structural scale, soluble protein due to the ease of inclusion body production. The inclusion body purification scheme used, combined with the Urea/Ni-NTA refolding, produced very pure soluble proteins. Problems of decreased yield with increased scale should be easily solved, probably by using many simultaneous small batches and pooling the results. The major problem encountered was inability to concentrate efficiently with the Amicon microcon 10kDa spin columns. However dialysis against PEG, used in the *Dm*PERX gel filtration runs, proved a efficient and simple way of concentrating samples.

*Md*PER 3 contains both PAS domains and the PAC. There are no crystallographic data for an entire eukaryotic double PAS domain in a transcription factor. As this is an ubiquitous motif in all eukaryotic clocks then crystallographic data would greatly advance our understanding of PAS function. *Md*PER 8 eventually proved to be a highly expressed soluble protein. However time

constraints did not allow for anything other then purification for antibody production. *Md*PER 8 does not have any well characterised domains, however it was well conserved with in the Dipteran period genes but had no homologous domains outside of the Diptera. Therefore solving its structure would be unlikely to be generally informative about domains in other proteins. Nevertheless *Md*PER 8 contained no cysteins and would be incapable of forming disulphide bonds and was therefore unlikely to be subject to the same multimer problem as *Md*PER 1 and 5. *Md*PER 7 proved easy to purify and refold but the low complexity of its sequence probably mitigates against expending effort on determining its structure.

CD analysis on crudely purified fragments illustrated that they can form spontaneous structured domains even when denatured. Therefore effort expended on further purification is unlikely to be wasted on the production of unstructured, poorly folded proteins. However aggregation was observed in all the fragments. Prevention of aggregates is paramount for structural studies. Changing the buffer conditions can help prevent hydrophobic interactions which can be instrumental in the formation of aggregates, additionally detergents may also prevent hydrophobic interactions and aggregation. Indeed there are some reports of detergents preventing aggregations even during crystallisation (Longenecker et al., 1998).

Chapter 6:

Comparison of per and tim RNA and Protein cycling within the Diptera.

Introduction

Comparative approach

Comparisons of the per and tim circadian molecular mechanism across the insects have shown that their mode of action was not conserved (Piccin. 1998; Reppert et al., 199). As detailed previously (in chapter 1) Antheraea pernyi and M. domestica provided two model circadian clocks that appear to function in different ways from the D. melanogaster circadian clock.

To recap, in A. pernyi, the per and tim transcripts and proteins cycled in heads (Reppert et al.,1994); however PER and TIM did not enter the nucleus in the eight neurones which expressed the proteins in the CNS (Reppert et al.,1994) and therefore could not directly inhibit their transcription. Although PER and TIM could be observed to enter the nucleus of the photoreceptors and ovaries (Reppert et al.,1994) ligation and transplant experiments demonstrated that the CNS drove this cyclic movement (Sauman and Reppert, 1998). An antisense transcript of per had been identified as a possible mechanism for generating a negative feedback loop (Reppert et al.,1994), but this was found to be a sex specific phenomenon and, as such, was not involved in circadian rhythm generation (Gotter et al.,1999). It is yet to be shown how the enigmatic A. pernyi circadian clock generates a negative feedback loop in these cells.

In *Musca,* although the TIM protein and *per* transcript cycled the PER protein did not cycle in head tissue extracts (Piccin, 1998). Nevertheless the lack of a PER protein cycle did not prevent rhythmic behaviour in wildtype *Musca* (Piccin, 1998). Furthermore a *D. melanogaster* promotor - *Musca* coding sequence fusion gene rescued *D. melanogaster* per^0 flies behavioural rhythms, and the PER protein levels cycled in these flies. Intriguingly this protein cycle was in the same phase as that of the transcript (Piccin, 1998), in contrast to the lag observed in *D. melanogaster* (So and Rosbash, 1997).

Given that the comparison of two species *Musca* and *A. pernyi* had illuminated the divergence of clock function within the insects it would be informative to observe the molecular mechanisms behind clock function of other insect species. Comparative studies were already underway on more distantly related insect species such as the cockroach, *Periplaneta americana* (J. Clayton and S. Reppert, pers. comm.) and the honey bee, *Apis melifera* (Toma et al., 2000). However analysis of *Musca* had shown that there was functional divergence within the diptera (Piccin, 1998; Piccin et al., 2000). Therefore it was decided to analyse the molecular mechanisms of two *drosophilid* species, *D. pseudoobscura* and *D. virilis.* Both species had easily accessible stocks and the *per* and *tim* genes had been cloned in both (Petersen et al., 1988) making molecular analysis simpler. *D. pseudoobscura per* had already been used to rescue *D. melanogaster* per^0 flies and had provided poor rescue of behavioural rhythms (Petersen et al., 1988; Piccin, 1998), indicating that there was sufficient functional divergence in the *D. pseudoobscura per* gene to prevent it functioning normally in the *D. melanogaster* transformants. Hence the function of the *per* gene in the circadian clock may have altered over evolutionary history. *D. virilis* is a more distant relative of *D.*

melanogaster than *D. pseudoobscura,* which is reflected by decreased *per* sequence homology (Piccin, 1998). Interestingly the PAS domain of *per* in *D pseudoobscura* and *D. virilis* were more diverged from *D. melanogaster* than *Musca (Piccin et al., 2000).* The PAS domain, a major dimerisation motif, associates PER with TIM (Gekakis et al., 1995; Huang et al., 1995; Huang et al., 1993) and divergence of sequence suggests change in dimerisation between the two proteins. No functional analysis has yet been performed on *D. virilis per*, but sequences of *per* were used to infer gene flow and genetic divergence between *D.virilis* populations (Ousley et al., 1998).

There was surprising little information on the molecular function of the clock components in the *Drosophila* group. Although many of the *Drosophila per* genes had been cloned they had generally been used to identify conservation in *per* sequences and therefore the likely functional domains. Indeed there was good coverage of insect *per* sequences in the databases as many *per* genes had been cloned for sequence analysis (Regier et al., 1998). There was, however there was a relative paucity of functional information for the *per* genes.

The observation of transcriptional cycling of *per* and *tim* across the *dipteran* species *Musca, D. pseudoobscura* and *D. virilis* would provide a broad functional analysis of *dipteran* circadian clocks. Furthermore the observation of *per* and *tim* transcript cycles could be used to infer functional differences in transcriptional or posttranscriptional regulation within the *diptera*. In addition protein cycles were analysed where possible and comparisons of protein to RNA cycle were used to indicate whether there was a lag in the negative feedback system in these *dipteran* species.

Transcriptional cycling and negative feedback

In order to create a negative feedback loop PER and TIM must negatively feed back on their own transcription in a circadian manner. This feed back can be observed by determining the *per/tim* transcript cycles. Failure to feedback would lead to high transcript levels and no cycle, as has been shown to occur in per^0 flies where the per^0 transcript levels over accumulate (Hardin et al., 1990). Hardin *et al.* (1990) were the first to observe the cycling of the *per* transcript, which had a 9-fold amplitude cycle; transcript levels were lowest at the end of the night (ZT 0), started to rise in the first half of the day (ZT4) and peaked towards the end of the day (ZT 10) then remaining at relatively high levels until the middle of the night (ZT 18) after which *per* transcript levels fell rapidly (Hardin et al., 1990). The mutants per^S and per^L caused advances and delays to peak levels respectively (Hardin et al., 1990). The *per* transcript cycles, and the mutants affects on these cycles strongly suggested that transcription of *per* was under circadian control and that *PER* itself was involved in its own transcriptional regulation.

To fully test the hypothesis of transcriptional regulation fragments upstream of the *per* translation start were fused to the *chloramphenicol acetyltransferase (Cat)* reporter gene, whose abundance was measured by RNAse protections (Hardin et al., 1992). The rationale being that mRNA cycles may arise without transcriptional regulation if, for instance, there were significant posttranscriptional effects. A large fragment, from –4000 to +25bp with respect to the *per* translation start, caused robust cycling in the *Cat* transcript. Therefore the promotor activity was under circadian control (Hardin et al., 1992). Furthermore RNAse protections were performed to observe whether posttranscriptional regulation was a feature of *per* cycles; pre-mRNA was observed by protection with an antisense *per* fragment

which included an intron, mature mRNA would be observed by the removal of this intron to create a shorter protected fragment. The pre-mRNA and mRNA cycled in the same phase and with similar relative amplitudes. This lead to the conclusion that the virtual identity in phase and amplitude of the mature and precursor transcripts indicated that there is little difference in synthesis or turnover and that there is no detectable lag due to posttranscriptional processing (Hardin et al., 1992).

The observation that promotorless *per* transgenes could cause behavioural cycles suggested that there must be some regulatory elements downstream of the *per* promotor (Hamblen et al., 1986). In contrast to Hardin *et al.* (1992) posttranscriptional processing was observed through an experiment that sought to directly determine transcriptional activity of the *per* and *tim* loci. The nuclear run off assay was used to observe transcriptional activity of both loci. The *per* and *tim* transcriptional activity both cycled with a 9-fold and 30-fold amplitude respectively. However mathematical analysis of *per* transcriptional activity and mature mRNA cycles showed that the stability of *per*-mRNA changed over time, becoming less stable in its declining phase, leading to the conclusion that this was likely to be due to posttranscriptional regulation (So and Rosbash, 1997). Furthermore the promotorless *per* transgene, which had previously been shown to rescue per^0 behaviour (Hamblen et al., 1986), was crossed into a *per* deletion bacground (per^-) but the *per* transcript cycled, clearly indicateing that there was posttranscriptional regulation (So and Rosbash, 1997), contradicting earlier results (Hardin et al., 1992). In addition alternative splicing events produced two *per* transcripts (type A and B) that differed in their 3' UTR. Transgenic flies that only produced type A transcripts showed slower protein accumulation cycles (Cheng et al., 1998) and the splicing of the 3' UTR was found to be important in maintaining a 24 hour clock

during temperature fluctuations (Majercak et al., 1999). Splicing was enhanced in cold conditions, causing an advance in the *per* transcript and protein cycles (Majercak et al., 1999), compensating the clock to colder temperature. The circadian clock requires temperature compensation to maintain its 24hr rhythm, as lower temperatures should lead to slower reaction kinetics and a slower molecular rhythm.

Approach to observing transcriptional activity

RNA protection analysis was chosen to determine transcriptional cycling of *per* and *tim* in the three species. All published analysis of *per* and *tim* transcriptional cycles had been achieved using RNA protection analysis (Hardin et al., 1992; Piccin, 1998; So and Rosbash, 1997). RNA protection offered enhanced sensitivity when compared to other techniques; large amounts of RNA could be hybridised to a probe and the non-specific RNA digested away to leave only the specific signal. This acts to reduce the background and would allow the quantification of low-level transcripts from large RNA pools. Another advantage was that more than one probe could be used simultaneously therefore *per* and *tim* cycles could observed in the same RNA sample. In addition a housekeeping gene such as *Ribsomal Protein-49kDA* (*RP49*) could be used to control for RNA loading, providing a reliable quantification of the *per* and *tim* transcripts. Although RNA protections were more laborious and demanding the increase in sensitivity over northern analysis meant it was the technique of choice.

Results: RNA protections and transcriptional cycling

Comparison of the probe length to protected fragment length, and the 25bp ladder (GibcoBRL) revealed that all probes had been digested to the correct size for full length protection. RP49 often produced a multiple banding pattern just below the protected fragment size. However this did not prevent quantification and was a feature of the published RNase protection analysis (Hardin et al., 1992). Each protection analysis was duplicated twice using flies from different entraining runs and the results were then analysed and plotted with standard error bars. *D.pseudoobscura* could not be repeated twice because more RNA, and therefore more flies, were required to achieve a signal. Furthermore the *D. pseudoobscura* stock used had slow growth rates making tissue collection difficult.

Very variable RNA loadings were observed (as shown by the large differences in the protected RP49). However attempts were made to equalise the loading of the RNA samples used for each protection. RNA was extracted with RNAzol and then cleaned up with phenol extraction and isopropanol precipitation. It was then quantified with either the qiagen ethidium bromide/ transiluminator method (see methods section) or by spectrophotometer. However although both methods produced reliable standard curves with yeast tRNA, the calculated concentrations were not comparable between the two methods when fly head RNA isolates was used. It was noticed that the RNA pellets were red in colour, presumably due to the red eye pigment in the WT flies (white eyed flies did not have red colouration to the RNA pellets). Presumeabley this pigmentation interfered with the spectrophotometer readings and gave unreliable results. Another confounding factor may have been that brain tissue is notorious for having a high proteoglycan content which is often carried over in RNA preparations. A protocol such as the

Trizol method (Gibco) which includes a proteoglycan removal step may have improved the reliability of RNA quanification..

D. pseudoobscura

In *D. pseudoobscura* the *per* transcript peaked at ZT 14 and reached its lowest level at ZT 22 (**Fig. 6.2**), with the demise of *per* taking eight hours to complete. Re-accumulation of the transcript began sometime between ZT 4 and 10.

The peak in *tim* transcript occurred much earlier than the *per* transcript, peaking at ZT 8 (**Fig. 6.3**) and reached its lowest level ZT 2, with the demise taking eighteen hours to complete. Quantification of the *tim* transcript appeared to give a noisy profile, with levels fluctuating from 50% and 80% of maximum between ZT 10-18, until a consistent declining phase occurred from ZT 20 until ZT 2.
In comparison to *per*, the *tim* transcript obviously peaked first, some 4 to 8 hours earlier. These data could be interpreted in two ways, firstly both genes were under different transcriptional control with *tim* transcription being activated earlier than *per*, secondly transcriptional control of *per* and *tim* were similar and the difference was due to posttranscriptional regulation of the gene transcript. Posttranscriptional control was most likely to occur on the *per* transcripts as *Drosophila per* undergoes posttranscriptional regulation which affects its accumulation, whereas *tim* does not show such a striking posttranscriptional effects (So and Rosbash, 1997). Whichever transcript is subject to posttranscriptional regulation, binding of the non-matured transcript to a RNA binding protein and subsequent processing could create a lag in accumulation.

D.virilis

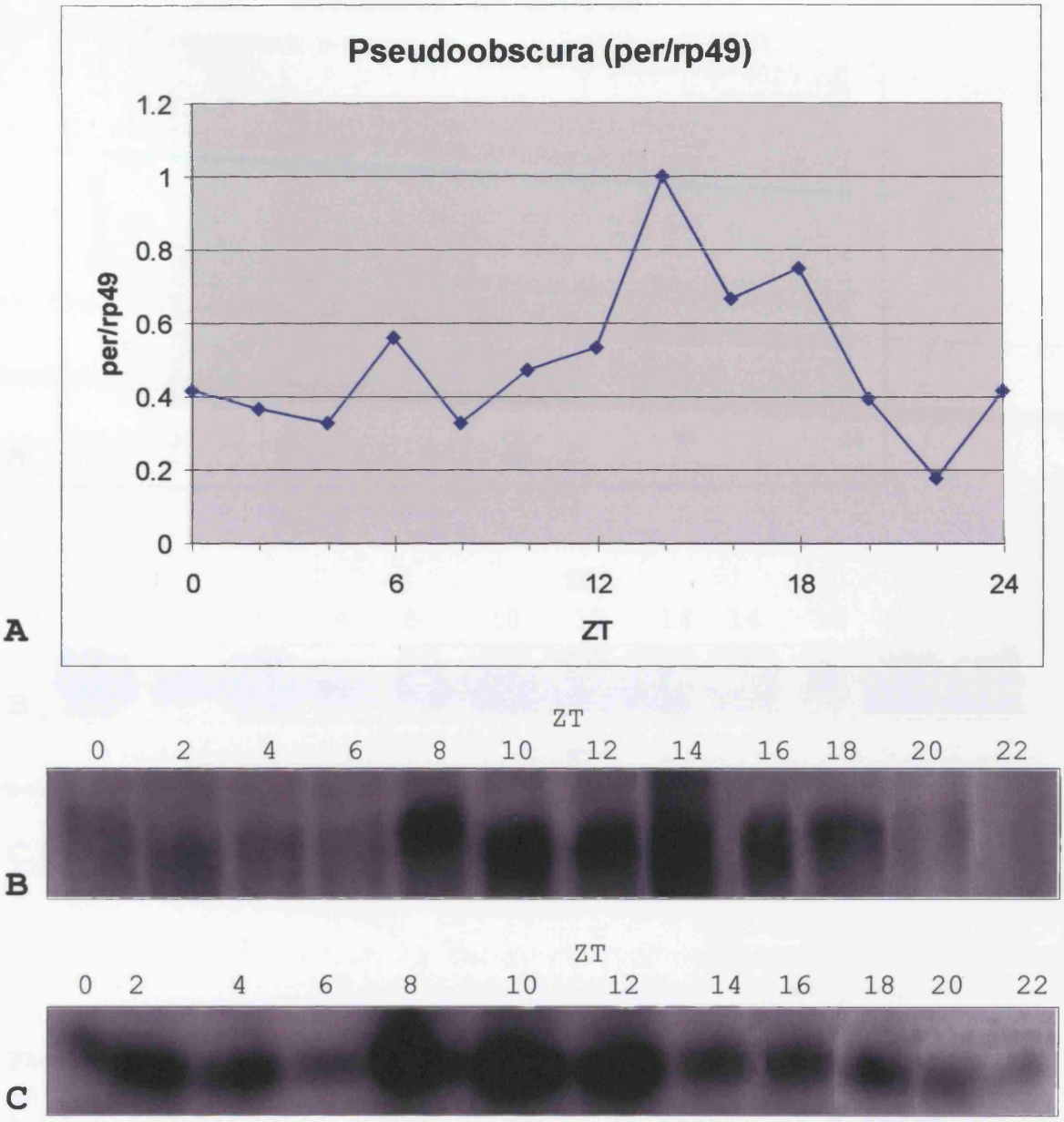


Figure 6.2, RNA protection analysis of *D. pseudoobscura per* and *rp49*.
A Quantification of *per:rp49* ratio expressed as a fraction of the highest value.
B *per* protected band.
C *rp49* protected band.

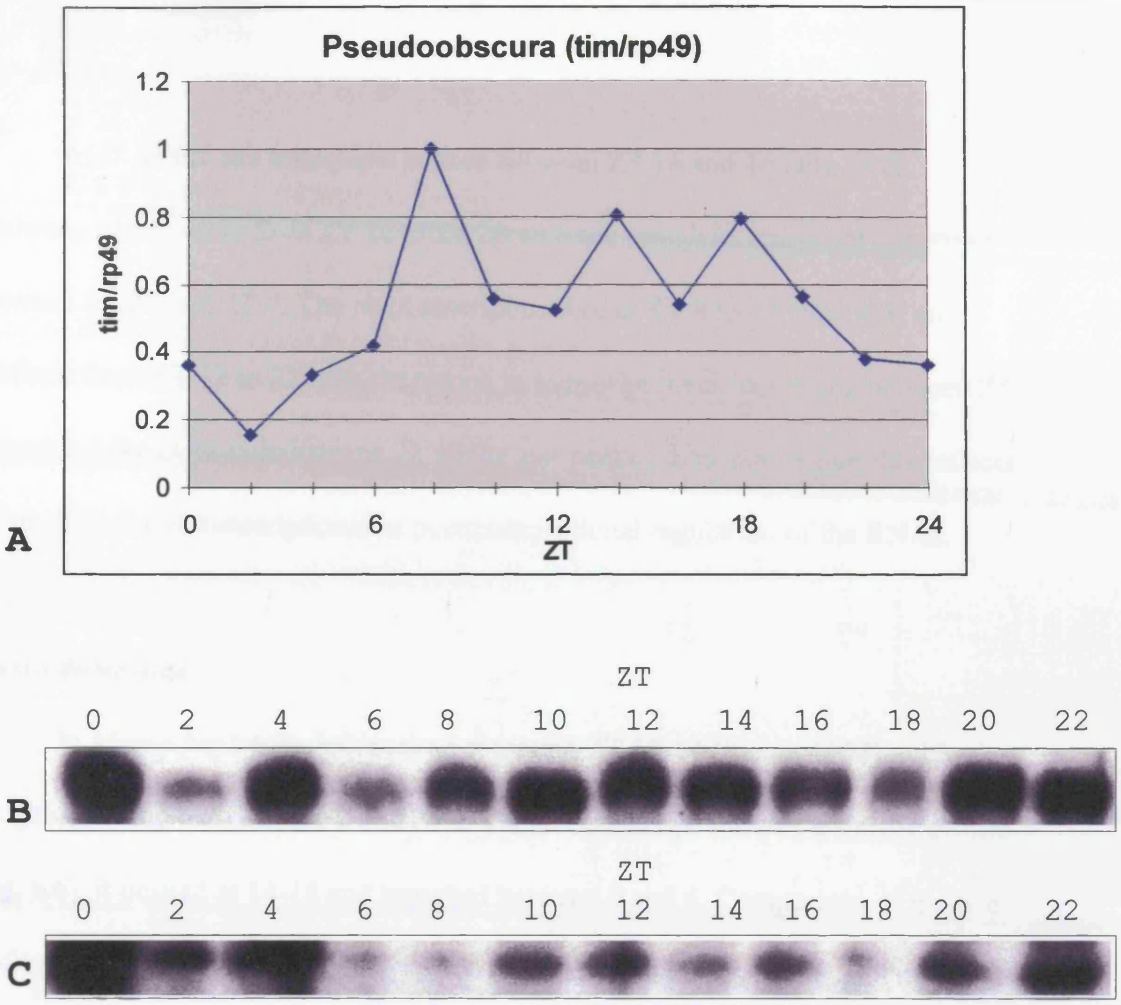


Figure 6.3, RNA protection analysis of *D. pseudoobscura tim* and *rp49*.
A Quantification of *tim:rp49* ratio expressed as a fraction of the highest value.
B *tim* protected band.
C *rp49* protected band.

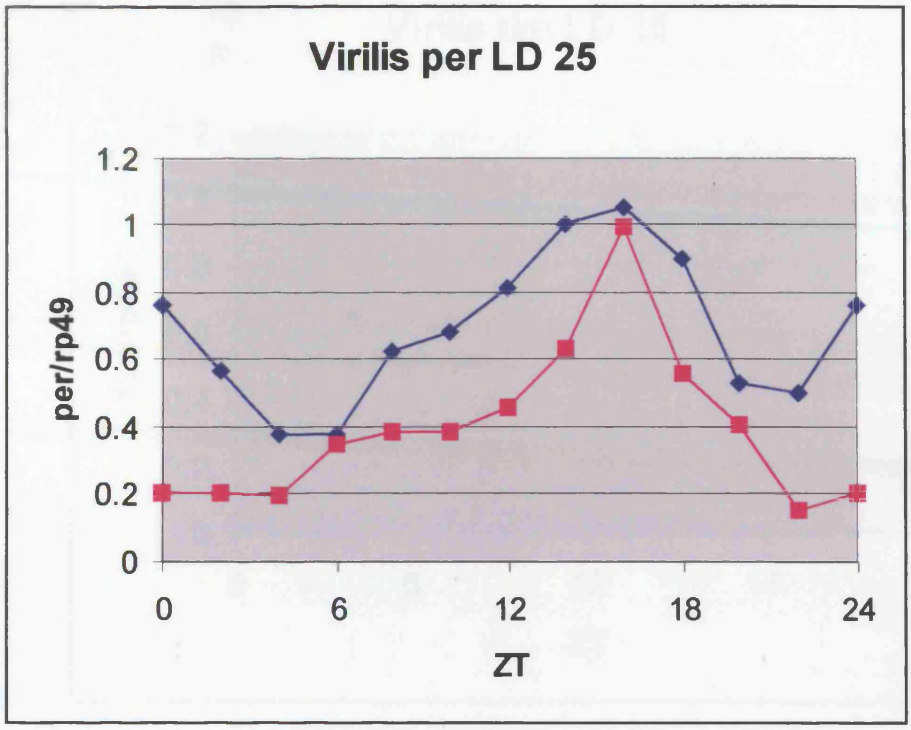
In *D. virilis per* transcripts peaked between ZT 14 and 16 (**Fig. 6.4**), declining consistently from ZT 16 to ZT 20 with the trough in transcript occurring between ZT 20 and ZT 4. The *tim* transcript peaked at ZT 8 to 10 (**Fig. 6.5**) and declined from ZT 12 to 22 with the trough in transcript levels occurring between ZT 20 and 2. Like *D. pseudoobscura, D. virilis per* peaked after *tim*. Again this reflects either differential transcriptional or posttranscriptional regulation of the RNAs.

Musca domestica

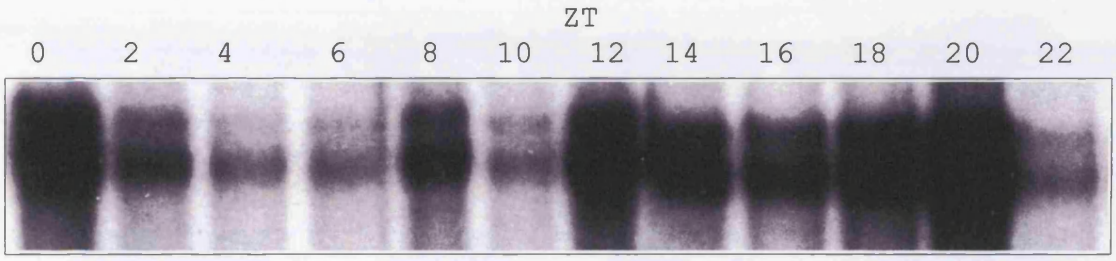
In *Musca per* transcripts peaked at around ZT 12-14 (Piccin 1998), with a trough occurring between ZT 0 and ZT4. The *tim* transcript showed a similar profile (**Fig. 6.6**), it peaked at 14-18 and troughed between 0 and 6. Comparison of the two profiles shows that, generally, they cycled in a similar manner peaking and declining at roughly the same time. If anything the *tim* profile is delayed with respect to the *per* profile, therefore there were unlikely to be any differences in the transcriptional or posttranscriptional regulation of the genes. This was in sharp contrast to the *D.virilis* and *D. pseudoobscura* profile where *per* transcripts peaked noticeably later than *tim*.

Results of Protein abundance analysis

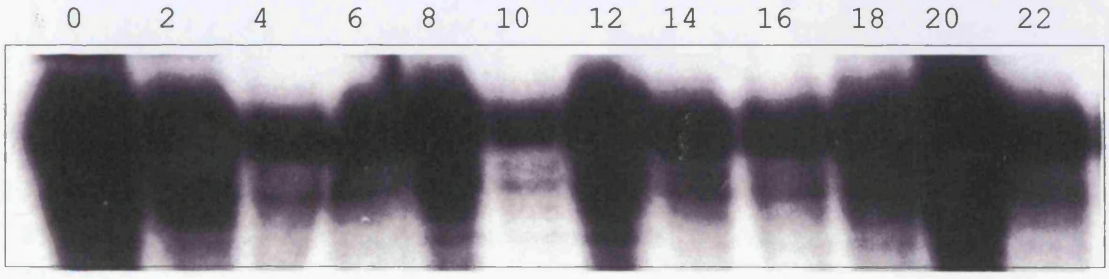
Attempts were made to observe the daily abundance cycles of PER and TIM in *Musca*, *D. virilis* and *D. pseudoobscura*. Only *Musca* PER and TIM cycles had been previously observed (Piccin, 1998), nevertheless a battery of antibodies raised against PER and TIM were available. However the antibodies had not been characterised for specificity to *D. virilis* and *D.pseudobscura* protein blots. Therefore the first priority was to characterise the staining pattern obtained with



A



B



C

Figure 6.4, RNA protection analysis of *D. virilis per* and *rp49*.
A Quantification of *per:rp49* ratio expressed as a fraction of the highest value.
B *per* protected band.
C *rp49* protected band.

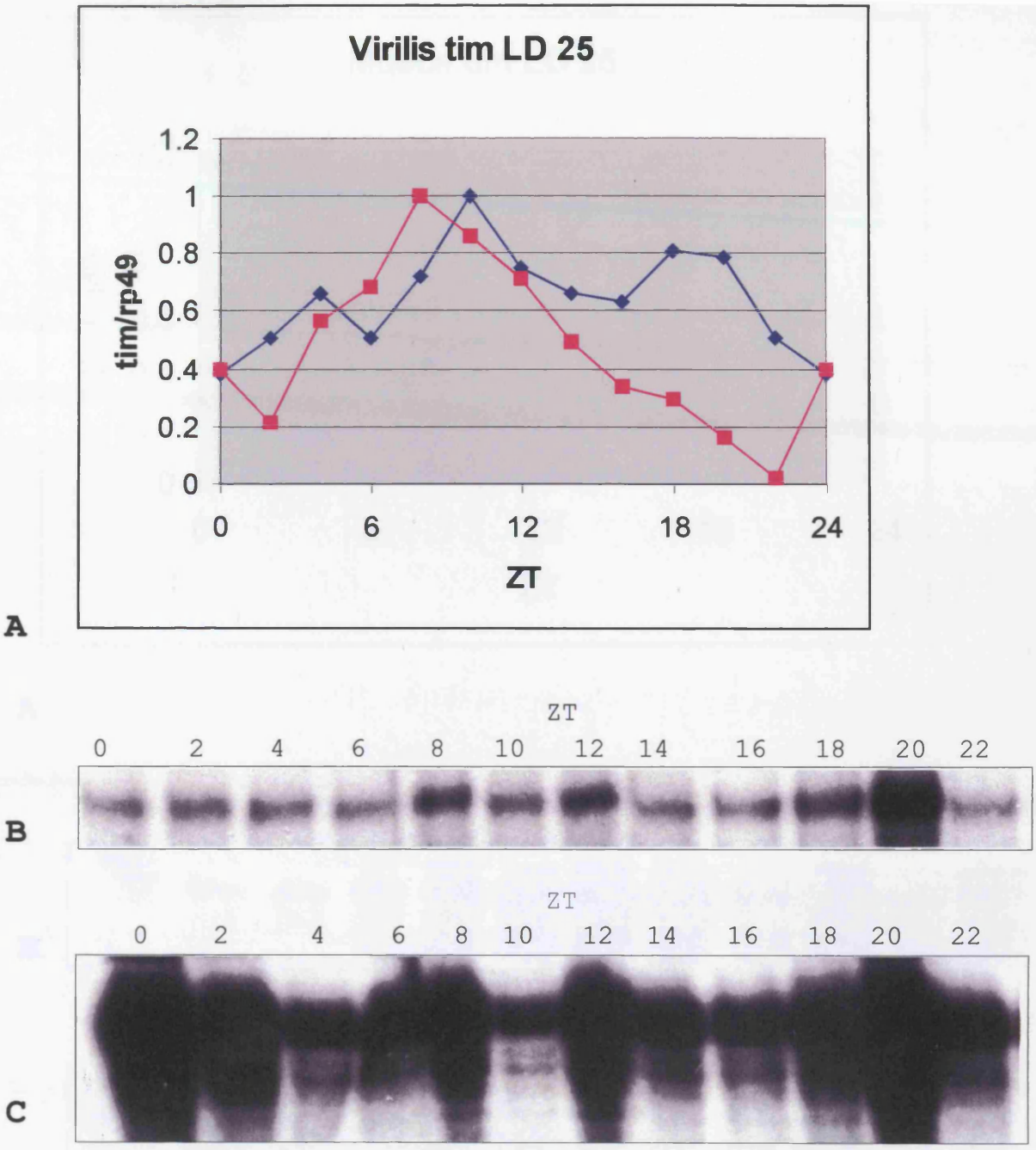
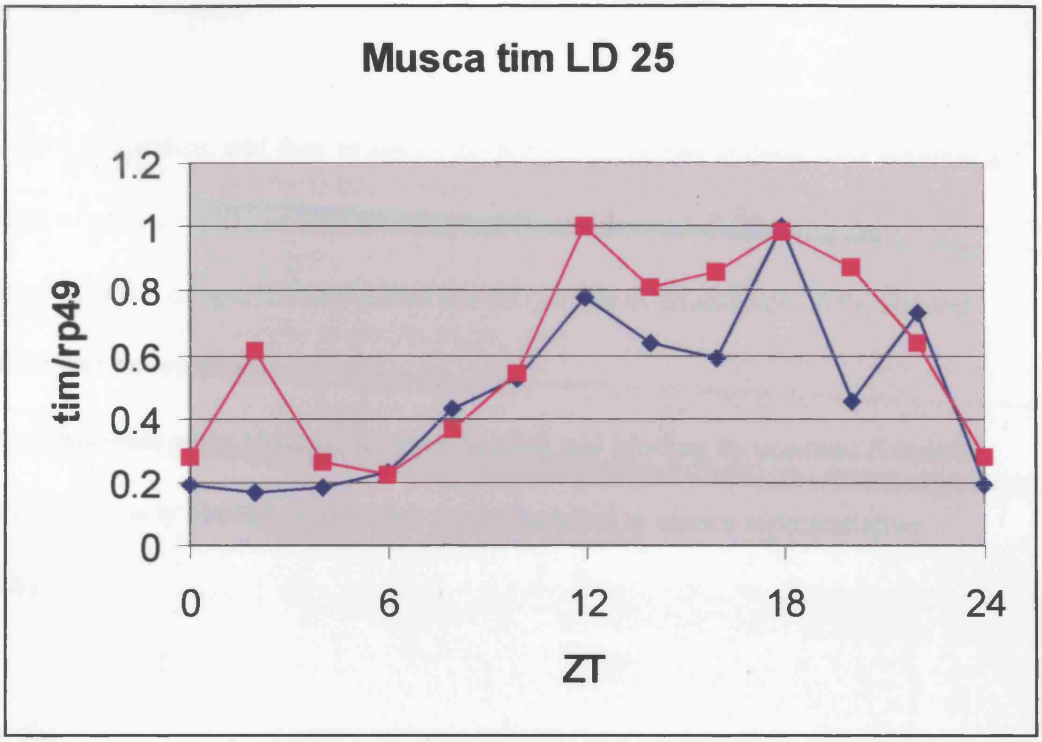
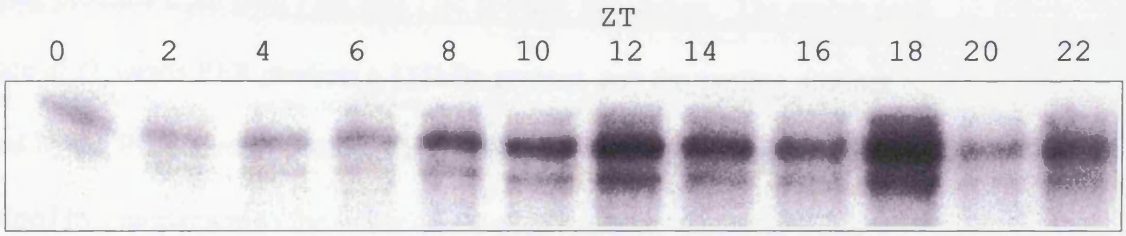


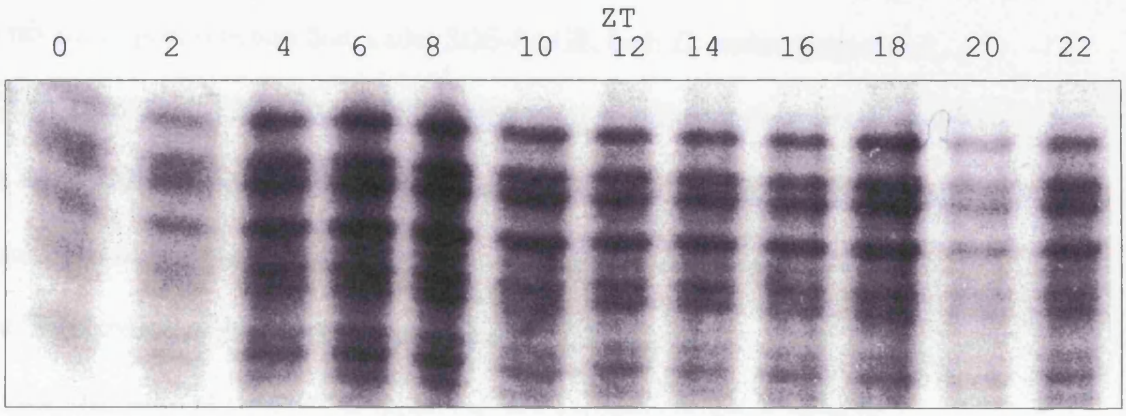
Figure 6.5, RNA protection analysis of *D. virilis tim* and *rp49.*
A Quantification of *tim:rp49* ratio expressed as a fraction of the highest value.
B *tim* protected band.
C *rp49* protected band.



A



B



C

Figure 6.6, RNA protection analysis of *Musca tim* and *rp49*.
A Quantification of *tim:rp49* ratio expressed as a fraction of the highest value.
B *tim* protected band.
C *rp49* protected band.

different antibodies, and then to assess the pattern produced to determine whether a pattern of specific PER or TIM staining had been achieved. Following the characterisation of specific antibodies the daily cycle of abundance of the stained protein was determined.

All blots were assessed for even loading and blotting by ponceau S staining. Uneven or poorly blotted membranes were discarded to ensure representative results.

D.virilis

The characterisation of antibodies on *D. virilis* blots lead to the observation of cycling proteins from both PER and TIM specific antibodies. The amino acid sequence of *D. virilis* PER predicts a 135kDa protein, and the cycling product obtained by the PER specific antibody migrates at around 160-170kDa (**Fig. 6.7**), as determined by comparison to the molecular weight markers. However it is a commonly noted phenomenon that, under SDS-PAGE, both *D. melanogaster* and *Musca* PER (Lee et al., 1996; Piccin, 1998) migrate at significantly higher molecular weights than predicted from their amino acid sequence. The identification of a band with a strong circadian cycle, in abundance, at a molecular weight slightly higher than predicted indicates that the staining pattern is PER specific.

Quantification of the *Dv*PER protein shows a robust circadian cycle in abundance (there is a >7 fold change in abundance) (**Fig. 6.7B**). The levels of PER protein were lowest at ZT 6 and remained low until ZT 12, which marked the start of the accumulation phase. The PER protein accumulation phase then peaked at ZT 20 and remained at steady state levels until ZT 2 after which there was rapid decline in abundance to a minimum at ZT 6.

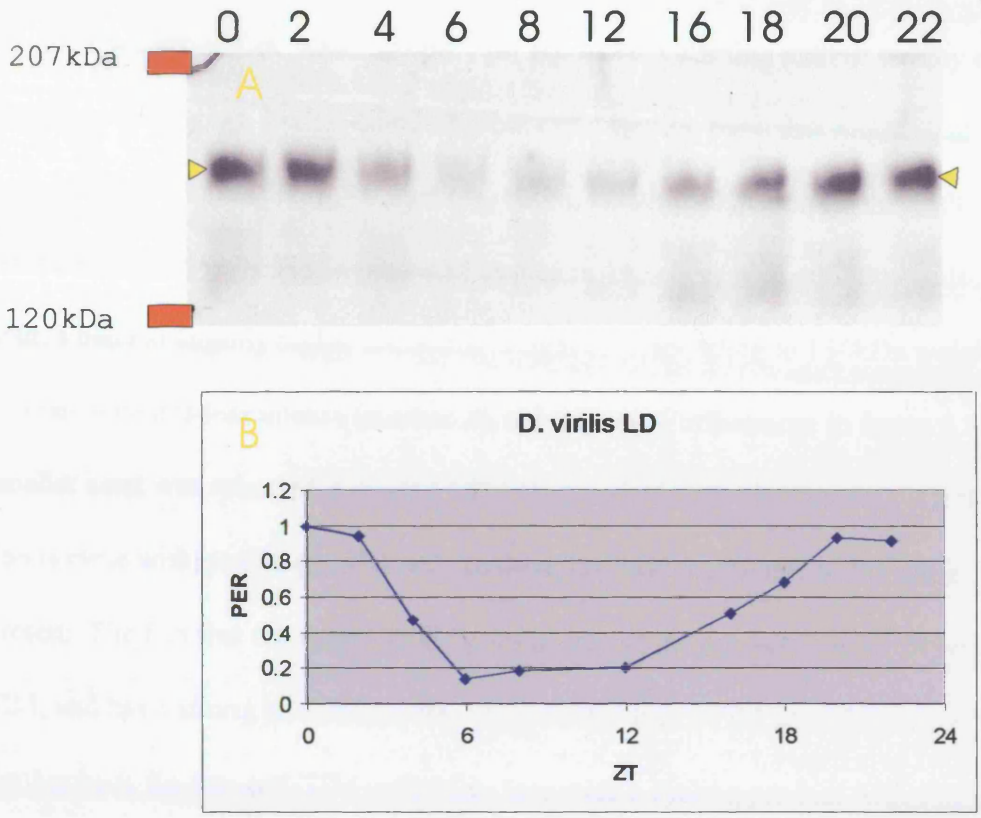


Figure 6.7 PER protein quantification for *D. virilis* LD cycle (25^{0}C), antiPER-rat4 primary (1:10000 dilution), Secondary anti-ratHRP conjugate.
A PER specific band, marked with yellow arrows. Molecular weight marker illustrated with red boxes, ZT timepoints numbered across the top
B Quantification of PER expressed as a fraction of the highest value.

Two antibodies raised against TIM produced a staining pattern worthy of further investigation. Both antibodies produced a cycling band that migrates at around 130-140kDa (**Fig. 6.8A and Fig. 6.8B**), which was just under the 150kDa predicted by *D. virilis* TIM amino acid sequence. However both antibodies also stain a band at slightly higher molecular weight corresponding to 150kDa weight protein, which is less intense (marked by red arrows). Furthermore in figure 6.8B, a smaller band was revealed at around 117 kDa (marked with blue arrows). All the bands cycle with similar profiles and are therefore likely to be due to the same protein. The fact that the major band cycles at around the correct size for *D. virilis* TIM, and has a strong circadian cycle, suggests that it is a TIM specific pattern. Furthermore the two anti-TIM antibodies recognise a cycling product with similar sizes on an SDS-PAGE gels, providing strong evidence of a specific TIM staining pattern. The presence of multiple bands above the major band indicates phosphorylation maybe occurring, which would increase the relative size of TIM (the relevance of this apparent phosphorylation is discussed later).

The cycle in abundance of *D. virilis* PER compares well to that published for *D. melanogaster* PER (So and Rosbash, 1997). In *D. virilis* PER peaks slightly later at ZT 20 (compared to ZT 19). However, of much more significance is the decline phase of *D.virilis* PER which occurs much more quickly than in *D. melanogaster*. *D. virilis* PER declines from near maximal levels at ZT 2 to an abundance minimum at ZT 6. In contrast the decline of *D. melanogaster* PER takes much longer, with a progressive decline occurring between ZT 23 and ZT 9. Therefore the cycle of abundance of *D. virilis* PER appeared to produce relatively compressed peak, mainly due to PERs shortened phase of decline. However it is worth noting that *D. virilis* PER did undergo an observable fluctuation in mobility

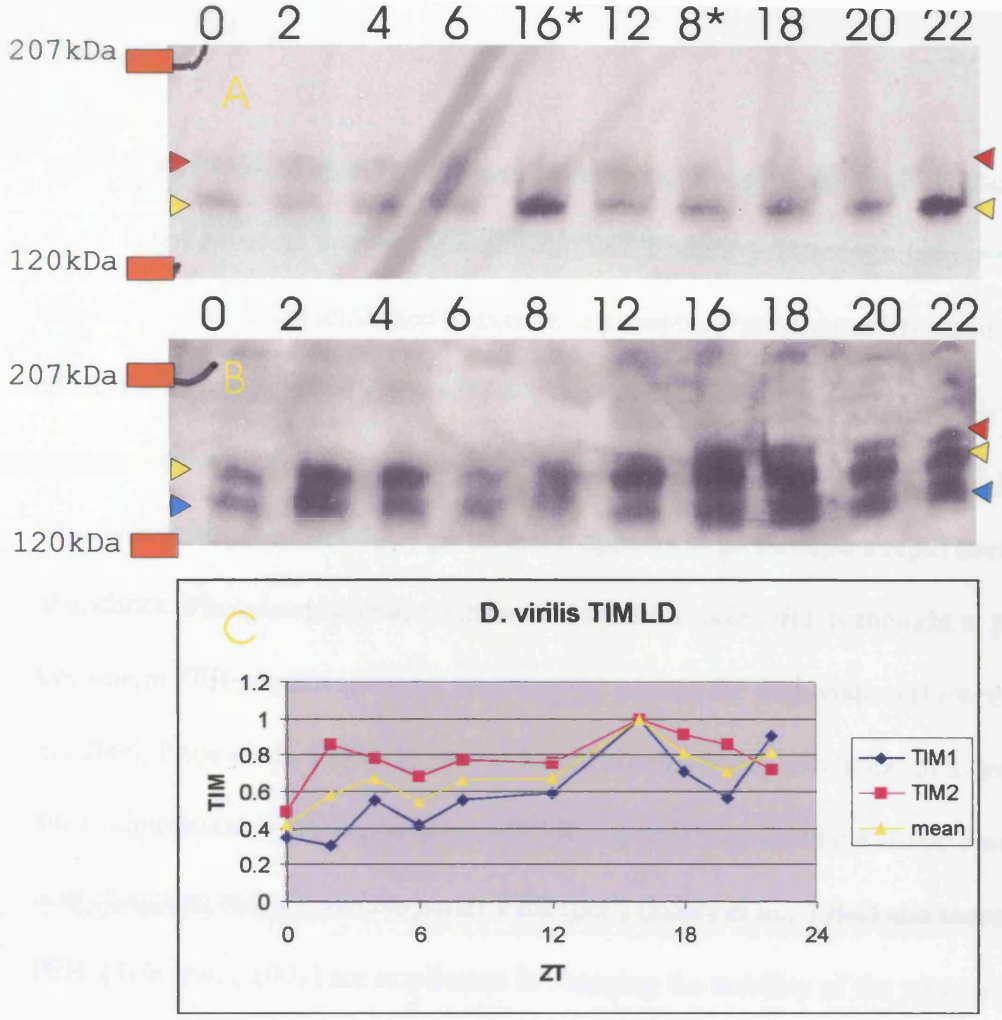


Figure 6.8 TIM protein quantification for *D. virilis* LD cycle (25^{0}C). Blot **A** antiTIM-UPR8 primary (1:1000 dilution) and secondary anti-ratHRP conjugate. Blot **B** anti-TIM306 , and secondary anti-rabbitHRP conjugate.
A and **B**, TIM specific band, marked with yellow arrows. Molecular weight marker illustrated with red boxes, ZT timepoints numbered across the top. Red arrows mark possible phosphorylated products, blue arrows mark degradation products.
C Quantification of TIM expressed as a fraction of the highest value. Results from both antibodies plotted and then given as a mean value

under SDS-PAGE (**Fig. 6.7**). However in contrast, *D. melanogaster* PER showed a much larger apparent increase in molecular weight during its accumulation (Edery et al., 1994), which was attributed to extensive phosphorylational modification (Edery et al., 1994; Price et al., 1998).

The evidence of less extensive phosphorylational modification in *D. virilis* PER is something of an oddity given that it appears to go through a rapid decline in abundance. Phosphorylational modification by the kinase DBT is thought to play a key role in PER protein stability, targeting the protein for degradation (Lowrey et al., 2000; Price et al., 1998). However it is likely that phosphorylation of a few key sites is important to the degradation of PER, especially given that a single amino acid change in both *D. melanogaster* PER (per^{s}) (Edery et al., 1994) and human PER (Toh *et al.*, 2001) are implicated in changing the stability of the protein through changes in phosporylational modification. Therefore the relatively modest change in phosporylational modification of *D. virilis* PER may still have a large effect on its stability.

The *D. virilis* TIM profile showed some striking differences to the *D. melanogaster* profile. Firstly, in *D. melanogaster*, TIM accumulated during the night (ZT 12-ZT 22) but was very quickly degraded by light stimulation during the day with very low or undetectable levels of TIM from ZT 1 to ZT 10 (Lee et al., 1996; So and Rosbash, 1997; Suri et al., 1998) and this lead to a large amplitude cycle in abundance. However staining from both antibodies (**Fig. 6.8A and 6.8B**) showed that *D. virilis* TIM was clearly detectable during the day at between half and a quarter of it peak levels. Secondly in *D. virilis* the change in mobility of TIM on both blots provided some evidence for phosphorylation. Although phosphorylation of TIM has been observed in *D. melanogaster,* it only results in subtle changes in

mobilty (Lee et al., 1996; Price et al., 1995; Sehgal et al., 1995; Zeng et al., 1996). In comparison if the changes in *D. virilis* TIM mobility are due to phosphorylation then TIM is either hyperphosphorylated or phosporylation causes confirmational changes that are preserved under SDS-PAGE. Therefore it would be worthwhile investigating *D. virilis* TIM and its kinase, *shaggy* (Martinek *et al.*, 2001) to see if the interactions have been altered.

Musca

The profiles of *Md*PER and *Md*TIM had already been observed (Piccin, 1998); *Md*TIM cycled in a similar manner to *D. melanogaster* TIM. However the *Md*PER protein was shown to have no circadian cycle in abundance and was constitutively expressed at high levels throughout the circadian cycle. This constitutive expression should stop the PER/TIM circadian clock mechanism from functioning if PER is a state variable. However *Musca* flies have a strongly rhythmic behaviour despite the lack of PER cycle (Piccin, 1998; Piccin et al., 2000). Therefore, either the model of PER/TIMs negative feedback does not apply in *Musca,* or negative feedback occurs but there is a novel mechanism which allows PER to accumulate without affecting subsequent behavioural rhythmicity. Several models of such mechanisms were detailed in chapter 3, however all rely on the negative feedback mechanism still applying but with different PER protein dynamics. If negative feedback still occurs then *Md*PER transcription and translation must still occur in a rhythmic manner. It had already been demonstrated that *Mdper* was rhythmically transcribed (Piccin, 1998). Therefore if *Md*PER could be observed before it over accumulates then a rhythmic build up of *Md*PER protein

may been observed, such an observation would then strengthen the role of negative feedback at the *per* locus in *Musca* .

An experiment was performed to observe *Md*PER protein abundance in pre-adult stages. Previous experiments had failed to observe *Md*PER accumulation in larvae (data not shown), therefore *Md*PER accumulation was observed in pupae and early adults. Adult *Musca* were entrained to at least three days of LD cycles and then allowed to lay on laying media for a 12hr period. The adults were then removed and the entrainment continued. Pupation was observed and pupae sorted to provide a cohort of pupae that had undergone pupation during the same 12hr period. Pupae were harvested at ZT 2 and ZT 14 until the 1st day after eclosion. Protein was extracted and subject to western analysis with anti-*Md*PER 774 polyclonal rabbit antibody. This antibody had already shown a specific *Md*PER staining profile (Chapter 4) where the *Md*PER specific band had been shown to migrate at 120kDa on 12% SDS-PAGE gels. In the pupal extract a 120kDa band provided weak staining on day one of pupation (**Fig. 6.9A**), it then disappeared from day two to day four of pupation, reappearing in the adult stage the 1st day after eclosion. However there appears to be no evidence of a stepwise increase in *Md*PER. There was no 120kDa *Md*PER band at ZT14 on day four of pupation, eight hours later, following eclosion, there was a strong band at ZT 2. This band then appeared to be at steady state levels as there was little change in its intensity 12 hrs later at ZT 14. Therefore rhythmic accumulation of *Md*PER was not found. The sudden presence of the 120kDa band at ZT 2 directly following eclosion is puzzling. There did not appear to be any gradual accumulation of *Md*PER. However below the 120kDa band there is a 100-110kDa band and, intriguingly there is some evidence that this band cycles. On the 2nd and 4th days of pupation this band is most abundant at night; following

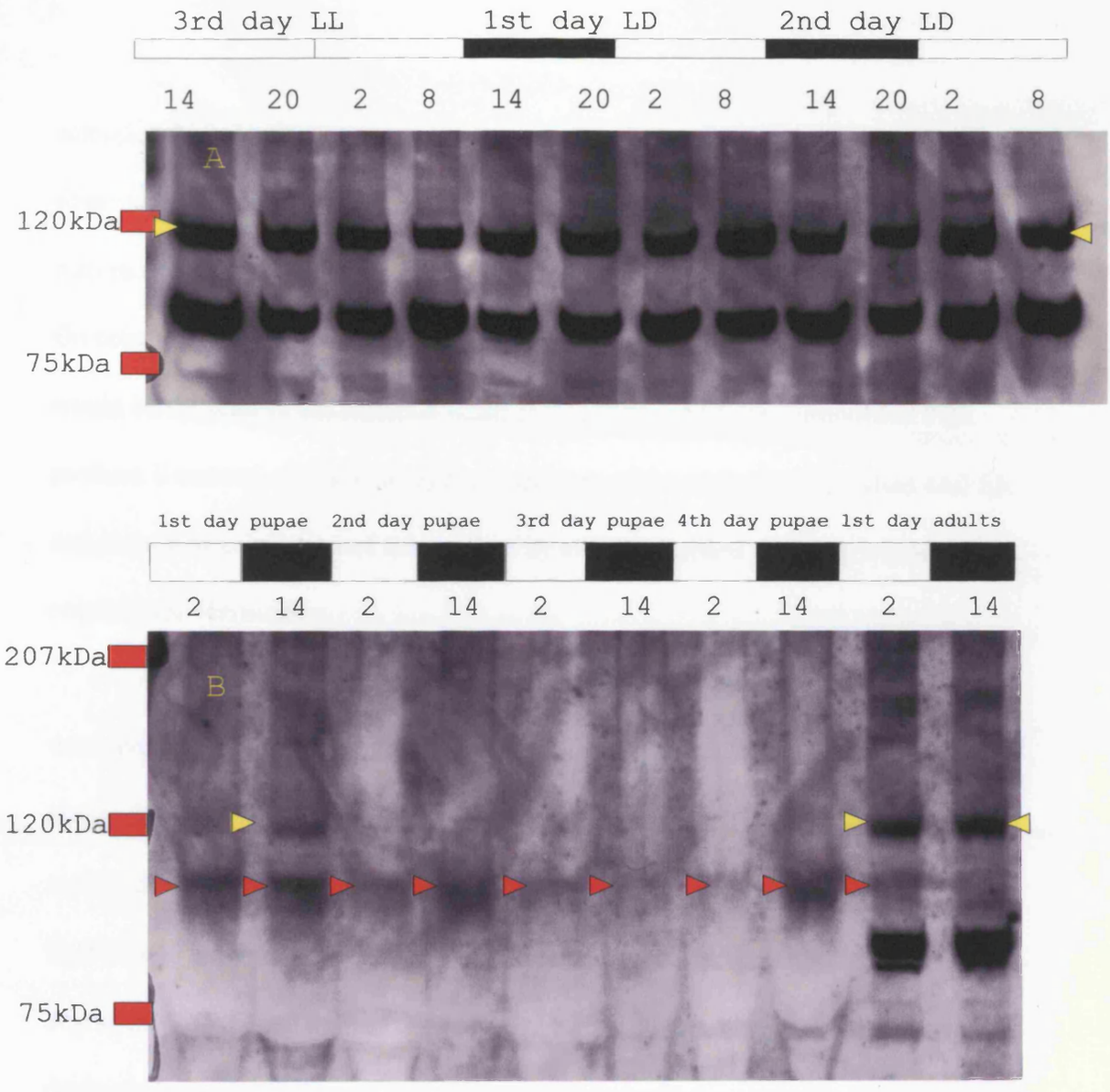


Figure 6.9 PER protein quantification for *Musca* PER labelled with anti*Md*PER-774
A LL experiment. Flies were entrained to three days of LL conditions and then moved to LD cycles for two days. ZT was numbered across the top, with corresponding days labelled above the time points. The PER specific band was marked with yellow arrows and the molecular weight marker was illustrated with red boxes.
B. Early developmental time course experiment. A cohort of similarly aged larvae were entrained to LD cycles and pupae and adults were then collected from the cohort, over time points covering the four days as pupae and the first day after eclosion. The PER specific band was marked with yellow arrows and the molecular weight marker was illustrated with red boxes. A possible smaller PER specific band was labelled with red arrows.

eclosion, the intensity of the band reduces and in blots of mature adult it is not observed (see Chapter 4). It is possible that this represents a PER specific staining pattern and the reduction in size was due to a lack of posttranslational modification. On eclosion this band may be modified to produce a 120kDa band, however it would be difficult to demonstrate whether this band is truly an unmodified PER product. Removal of this band by preabsorption of the antibody to purified *Md*PER and immunoprecipitation of this product by affinity purified antibody would be required for verification.

A further experiment was performed inorder to determine whether *Md*PER was rhythmically produced. In *D. melanogaster* PER is stabilised by TIM and is absent from tim^0 flies (Price et al., 1995). Furthermore the TIM protein is degraded rapidly during light exposure (Myers et al., 1996). Therefore conditions of constant light should leave TIM levels low and destabilise PER, and result in PER degradation. Once the light stimulus was removed PER and TIM should reaccumulate; observation of this reaccumulation could then provide evidence of rhythmic *Md*PER production.

In an experiment designed to observe this reaccumulation adult *Musca* were subject to three days of constant light conditions (LL) after which they were moved back into LD conditions. The flies were then collected over a time course and subject to western analysis with anti-PER 774. Quantification of the *Md*PER band at 120kDa (**Fig. 6.9A**) shows that, even after three days of LL entrainment, *Md*PER levels were high. Following release from LL there is no discernible cycle in *Md*PER abundance. The fact that *Md*PER is still present following LL would prevent the observation of rhythmic *Md*PER production. In retrospect it was not surprising that MdPER did not respond to the LL conditions as expected. Under normal LD

conditions *Md*TIM has been shown to be degraded during light exposure (Piccin, 1998), however this removal of *Md*TIM dos not destabilise *Md*PER sufficiently for it levels to decline over the light phase of the LD entrainment (Piccin, 1998). Removal of *Md*TIM by LL should result in the removal of negative inhibition by the PER/TIM complex and result in high levels of *Mdper* transcription. Therefore there should be constant translation of the *Md*PER product and this would lead to overaccumulation. This experiment highlights the insensitivity of *Md*PER to light stimuli and MdTIM removal, but fails to determine whether *Md*PER is rhythmically produced.

D. Pseudoobscura

No obvious candidate band for *D. pseudoobscura* PER and TIM were identified. Only one antibody provided a pattern (anti-PER rosbash full length); two strong bands were observed at 207kDa (**Fig. 6.10A**). Under 6M denaturing SDS-PAGE the doublet runs at 150kDa (**Fig. 6.10B**), which is sufficiently close to the 127kDA predicted for *Dp*PER for the band to be a candidate for specific staining. Comparison of the protein to that of a *Dp*PER transformed *D. melanogaster* per^0 fly would help identify this band as PER specific. A *D. melanogaster* per^0 line carrying a *Dpper* construct had already been produced. However no *Dp*PER specific band was found when the construct pattern was compared to per^0 (data not shown). However it had been noted that the construct did not provide good behavioural rescue (Piccin et al., 2000) and its mRNA did not appear to be well expressed (Mike Hennessey, pers comm.). Therefore the doublet band at 150kDa cannot be confirmed as PER. It is intriguing that this candidate band fails to cycle, just as *Md*PER fails to cycle.

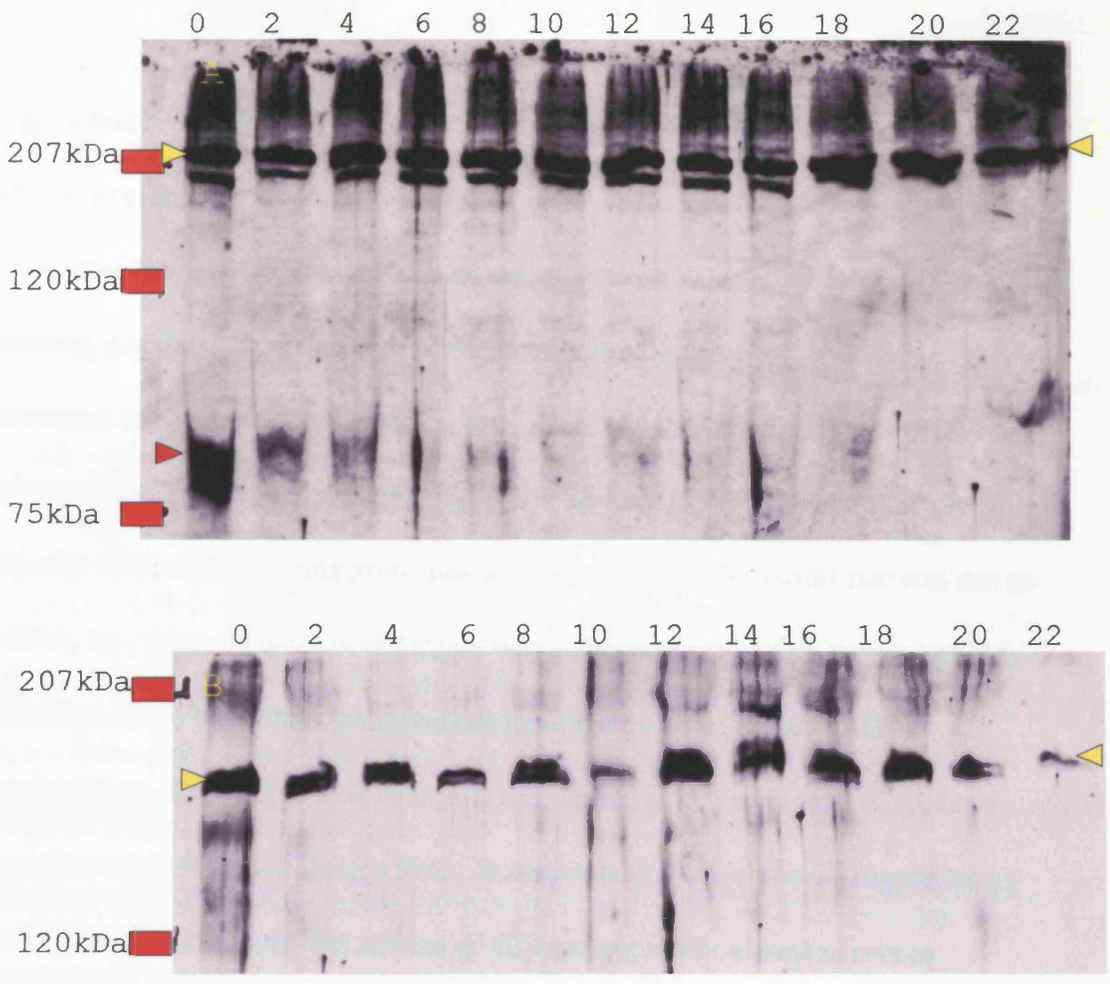


Figure 6.10 Blots of a *D. pseudoobscura* LD time course labelled with antiPER-771. Candidate band marked with yellow arrows and the molecular weight marker was illustrated with red boxes.
A Normal SDS-PAGE gel.
B 6M UREA SDS-PAGE gel.

Conclusions

RNAse protection

It was outwith the experimental design to discriminate between differences occuring due to transcriptional or posttranscriptional regulation. Indeed the protection analysis was performed on total RNA and therefore both pre-mRNA and mature mRNA would be protected. However So *et al* (1997) showed that the majority of the signal during protection analysis of *D. melanogaster per* was due to mRNA, less than one tenth of the signal was achieved with probes for pre-mRNA. Therefore it was likely that the signals observed were mainly due to mRNA accumulation. This tallies well with polyadenylation acting to protect RNA from degradation (reviewed in Ross, 1996). Differences in transcriptional regulation of *per* and *tim* would require the actions of different promotor elements and/or transcriptional activators. This was unlikely considering that the E-box and CREM-responsive promotor elements of *per* were conserved between mice and *Drosophila* (Belvin et al., 1999; Takahata et al., 2000). However this does not preclude differences in the enhancer elements of *per* and *tim* accounting for the differences in transcriptional profile, however the most compelling explanation for the differences in transcriptional profiles is posttranscriptional regulation. Posttranscriptional regulation had been demonstrated to occur on *D. melanogaster per* transcripts delaying its cycle (So and Rosbash, 1997), which is in accordance with the results presented here for *D. virilis* and *D. pseudoobscura*. Alternative splicing of *per* mRNA had been shown to be involved in maintaining rhythmic period across temperature ranges (Majercak et al., 1999). Therefore it seems likely that posttranscriptional regulation of the *per* transcript accumulation has been

maintained across the *drosophilids* but not the *diptera* as a whole, as there is no delay between *Musca per* and *Musca tim* transcript cycles. Whether alternative splicing has been conserved could not be answered by this study but the likely conservation of posttranscriptional regulation in the drosophilds suggests it would certainly be worth investigating

Protein cycles: features and comparison to transcript cycles

The comparison of both protein and RNA cycles of *per* and *tim* allow inferences to be made about how the clock maintains it rhythm. In particular there is a strong theoretical framework, which predicts that the time lag between RNA transcription and protein translation is an important factor in maintaining the circadian feedback cycle (Gonze et al., 2000; Lewis et al., 1997). The lag caused by posttranscriptional processing can be indirectly observed through comparison of protein and RNA cycles. Furthermore posttranslational processing, such as phosphorylation events, could also be observed. Phosphorylation by the kinase DBT has been shown to introduce a lag between translation and nuclear entry in *D. melanogaster* (Kloss et al., 2001; Kloss et al., 1998; Suri et al.,2000).

In *D. virilis* the *per* transcript peaks at ZT 16 (**Fig. 6.4**) and the protein peaks at ZT 20 (**Fig. 6.7**), therefore the PER protein undergoes a four hour lag in accumulation, which is similar to the lag that occurs in *D. melanogaster*. However *D.virilis* PER protein does not undergo an easily observable change in mobility over the day which is unlike *D. melanogaster* PER (Edery et al., 1994). The phosphorylational of PER alters the period length of the clock, with mutants in the PER kinase, DBT a causing short and long behavioural phenotypes (Price et al., 1998).

An unusual aspect of the *D. virilis* circadian mechanism is that there is a long delay between the *tim* transcript and TIM protein. The transcript peaks at ZT 10 (**Fig. 6.5**) and the protein peaks six hours later at ZT 16 (**Fig. 6.8**). How this delay is achieved cannot be determined by this set of experiments, h`owever it is likely that the delay is mediated by post-transcriptional processes. Furthermore the unusual, apparent phosphorylation (**Fig. 6.8**) of *D. virilis* TIM may contribute to this delay. The TIM kinase, SHAGGY (Martinek et al., 2001) may play a similar role to DBT, in that it contributes to the protein stability and therefore affects the timing of the circadian cycle (Kloss et al., 2001; Kloss et al., 1998; Suri et al.,2000). It may be possible that *tim* gene has a larger contribution to maintaining cycle length in *D. virilis* than in *D. melanogaster*, where it appears to mainly function as the light sensitive input into the negative feedback.

Perhaps one of the most intriguing aspects of the *D. virilis* molecular mechanism is that the *per* transcript peaks at ZT 16 (**Fig. 6.4**) some six hours later than the *tim* transcript, which peaks at ZT 10 (**Fig. 6.5**). Therefore there must be a difference in either transcriptional control or post transcriptional control of the *D. virilis per* and *tim* genes, in order to generate the large lag between the *per* and *tim* transcript cycles.

In contrast the TIM cycle in *Musca*, which has been previously published (Piccin, 1998), shows little delay from the *tim* transcript cycle observed here (**Fig.** 6.6). Therefore the *tim* gene does not seem to be responsible for generating a lag in the circadian molecular mechanism in *Musca*. Additionally, an experiment that sought to observe the profile of PER translation failed to observe any strong evidence of a circadian cycle in translation of PER in *Musca*. Instead these

experiments illustrated the insensitivity of the *Musca* PER protein to degradation during the light phase of the circadian cycle.

It was unfortunate that no obvious candidates for *D. pseudoobscura* PER and TIM were identified by the antibodies used as this does not allow a comparison of transcript to protein cycle. However the *D. pseudoobscura tim* transcript peaked at ZT8 (**Fig. 6.3**) whereas the *per* transcript peaked six hours later, at ZT 14 (**Fig. 6.2**). It would, therefore, appear that once again the *per* transcript cycle is delayed, and this is likely to be through post-transcriptional regulation of the *per* transcript. However as a note of caution, the cycles of the transcript have only been observed once in *D. pseudoobscura*. The low levels of both transcripts meant that a large amount of RNA (20 μg) had to hybridised to a strongly labelled probe in order to observe a protected product. Attempts to repeat the experiment were hampered by the requirement of fresh probe and large amounts of head tissues required from this slow growing strain.

Chapter 7: Discussion

One of the main aims of this thesis was to observe *Md*PER accumulation in Musca and, in particular, its subcellular localisation throughout the day. It had already been shown that this protein failed to cycle in abundance (Piccin, 1998) and appeared to be constitutively expressed. Under the *D. melanogaster* PER/TIM negative feedback model this state should stop the clock. Constitutive PER expression should result in constant negative feedback of its own transcription and should prevent rhythmic accumulation of PER, TIM or any of the downstream targets. However the *Musca* flies have strong behavioural rhythmicity despite the apparent lack of a PER protein cycle (Piccin, 1998; Piccin et al., 2000).

Attempts to analyse the daily accumulation of *Md*PER in *Musca* were hampered by the lack of an antibody with strong avidity for *Md*PER in sections. A comprehensive trial of all of the available anti-PER antibodies was performed and is detailed in Chapter 3; it failed to reveal an appropriate antibody

All of the antibodies used in Chapter 3 were donations from circadian clock laboratories and the small volumes of antibodies received hampered efforts to properly characterise them. Only a limited dilution series could be performed, as there was not enough of each of the antibodies to use them at high concentrations. Antibodies recognise short sequence lengths in the target amino acid (Harlow *et al*, 1988). As there was a significant sequence divergence between *Md*PER and *d*PER (Piccin et al., 2000) then only antibodies that recognise conserved sequences in *Musca* are likely to be useful. Given that most of the antibodies were raised to *d*PER then it would be reasonable to expect that any avidity to

*Md*PER is going to be limited to a small minority of antibodies within the polyclonal antibody population. A further compounding problem is that the proteins are not denatured during immunocytochemistry. Therefore a successful *d*PER antibody must also recognise an epitope sequence that is not buried or inaccessible within the protein structure. A *d*PER antibody with affinity to denatured *Md*PER had previously been identified by western analysis (Piccin, 1998), however this antibody did not show any *Md*PER avidity during immunocytochemical analysis in Chapter 3 and this was likely to be due to the epitope being buried. Furthermore the antibodies used by (Piccin 1998) were no longer available. Although different bleeds of the polyclonal anti-PER (αPER-978) antibody were available, western analysis (chapter 4) had shown that the new bleed did not have the same anti-*Md*PER activity. This could be attributed to changes in the antibody population in response to repeated innoculations. The practice of repeat bleeding used to produce the polyclonal antibody αPER-978 is unusual; normally a rabbit is exsanguinated once a sufficient titre has been achieved, which produces a pool of antibody with a known activity. The practice of taking many bleeds to produce αPER-978 allowed the antibody population to vary over time and therefore required the antibody be recharacterised with each bleed.

Attempts were made to produce an antibody that would allow the characterisation of *Md*PER on sections. The approach used relied upon inoculating fragments of *Md*PER that covered the entire amino acid sequence, so all of the potential epitopes were therefore presented to the rabbit's immune system. Furthermore, where possible, soluble protein was inoculated, as some antibody reactions can be sensitive to secondary structure (Harlow and Lane, 1988), and these antibodies were mostly likely to provide a

strong avidity during non-denaturing immunocytochemistry. Characterisation of this antibody (α*Md*PER-774) by assessing its reaction strength against the inoculated *Md*PER fragments by western analysis (Chapter 4), showed that it reacted strongly to the inoculated proteins. Indeed it appeared that most, if not all, of the inoculated fractions had produced an immune reaction. This high coverage of the *Md*PER protein meant it was likely that an accessible epitope would be recognised during immunocytochemistry.

Furthermore western analysis of *Musca* tissue extracts showed that, when run under the same conditions to those used by Piccin (1998), the αPER-774 produced an identically sized band as the 150kDa achieved with the αPER-978 antibody.(Piccin 1998) had demonstrated that this band represented a specific α*Md*PER product. Affinity purification of the antibody serum with the α *Md*PER fragment removed the band and elution of the affinity purified antibody fraction then reinstated the band. (Piccin 1998) used 6.5% SDS-PAGE gels during western analysis, however during most of the western analysis in this thesis 12% SDS-PAGE gels were used. Under these conditions the 150kDa product found by Piccin (1998) migrates at closer to 120kDa, which incidentally is the size predicted for *Md*PER from its amino acid sequence. The apparent size discrepancies of the bands produced by western analysis with the αPER-774 antibody on 6.5% and 12% represent common phenomena in western analysis. Although the SDS in SDS-PAGE is supposed to be a powerful protein denaturant it does not fully denature the proteins into a linear amino acid sequence (Maniatis et al., 1998). Instead, a lot of the globularity of the protein remains. The behaviour of this non-linear structure as it tumbles through the SDS-PAGE gel can result in quite different migration characteristics when the strength of the gel is changed. Indeed at least one proprietary brand of protein

molecular weight marker publishes differing migration sizes for each protein size standards depending on the gel strength (biorad technical bulletin).

Immunocytochemistry using this αPER-774 antibody was performed on *Musca* head sections. The staining pattern produced was very weak, however it did show that some cells between the optic lobe and the CNS stained with αPER-774. The staining analysis was performed with sera that had been preabsorbed against per^0 flies, which acted to reduce the background. This staining was therefore likely to represent a true PER specific pattern. Unfortunately, due to the laborious nature, only two time points could be analysed for a staining pattern. One time point was taken from late at night and the other from the middle of the day. Both time points produced a similar pattern with weak staining occurring in neurones between the optic lobes and CNS. In both time points the cells staining appeared to be cytoplasmic as a doughnut pattern was produced. However the pattern could not be confirmed as cytoplasmic because the enzyme based labelling system did not allow the use of a nuclear counterstain. However the enzyme based system does allow a degree of signal amplification so the use of fluorescent based labelling may have prevented the observation of the weak signal. One of the problems that arose in the fluorescence based staining (chapter 3) was that of artifactual patterns being produced when tissues autofloresce. However the pattern of artefacts was not so pronounced with enzyme based methods as few tissues have peroxidase activity, and precautions such as preincubation with H_2O_2 removes any residual staining.

The location of the stained cells is intriguing because it is in an area of the *Musca* brain that corresponds to where the LNs were found in *D. melanogaster*. The *Musca* staining is located in a dorsal area of the neuropil, between the optic lobe and CNS,

which corresponds well to the dorsal LNs of *D. melanogaster* (Ewer et al., 1992). It would be extremely difficult to assess whether these cells were LNs in *Musca* as this cell type has not been identified previously in this species. PDH co-staining with PER could be used to confirm nature of these cells (Helfrichforster, 1995), however this would require an antibody that has avidity to *Musca* PDH. Furthermore only a minority of LNs in *D. melanogaster* co-localise with PDH, so the lack of PDH staining could not be used to dismiss the pattern as non-LN staining; and co-localisation with PDH could only act to confirm the LN type. In fact, studies on the beetle *Pachymorpha sexguttata* showed that strong candidates for PER expressing cells did not counterstain with PDH (Frisch et al., 1996).

No other cell locations showed a specific pattern with the αPER-774 antibody. The candidate areas that gave strong patterns in the published literature with *D. melanogaster,* such as the photoreceptors and glia around the medulla, lobular and lobular plate, were rigorously assessed in *Musca* for a staining pattern and none was found. However the weak staining in the LN candidate cells between the optic lobe and CNS is likely to be due to poor antibody avidity rather than poor *Md*PER expression as western analysis shows a high level of constitutively produced protein. The staining pattern found in the cells at the two time points analysed, where the intensity did not change between night-time conditions and daylight conditions, tallies well the constitutive expression seen in westerns. It is possible that this cell type is entirely responsible for the pattern seen on westerns of head extracts. However it is possible that other cells with a lower expression profile are responsible for the western signal; *per-luciferase* fusion gene analysis of *D. melanogaster* showed that the majority of cell types

actually have rhythmic PER expression (Brandes et al., 1996; Plautz et al., 1997) and previous antibody analysis had failed to identify them because they could not detect the low level expression.

The two time points used for the immunocytochemical analysis were chosen because, under the PER/TIM negative feedback model, they should show two different states. The time point from the dark phase of the cycle was chosen because TIM was high (Piccin, 1998). The constitutively expressed *Md*PER should therefore be in a dimer with TIM; a state that permits it to enter the nucleus. Observation of this nuclear movement would confirm that the model of PER/TIM negative feedback could still apply. The time point from the light phase was chosen because TIM levels were low and had been low for sometime. Therefore PER should have been degraded and thus be absent from the nucleus, releasing negative feedback and allowing transcription and a new feedback cycle to begin

Two experiments were performed that sought to observe the accumulation phase of *Md*PER. The first relies on the suggestion that, perhaps, rhythmic accumulation cannot be observed in adults because there is already an over accumulation of the PER product so 'old PER' masks rhythmic translation of 'new' *Md*PER Rhythmic translation of *Md*PER may allow *Md*PER to function in the negative feedback loop despite its apparent constitutive expression. One model of how the *Musca* clock may function detailed in (chapter 3) is that the pool of *Md*PER that accumulates does not degrade because it has been improperly phosphoryated by the kinase double time. This prevents it binding to TIM and therefore this form of *Md*PER does not have a clock function. However a small proportion of the *Md*PER that is translated each day may be properly phosphorylated and

binds TIM to have a clock function. This model requires rhythmic translation of the *Md*PER, for which there is no experimental evidence other than rhythmic transcription of *Mdper*.

The two experiments relied on different approaches. One sought to observe the *Md*PER protein before it over accumulates. A time course was collected from a cohort of larvae as they underwent pupation and eclosion, which should allow observation of rhythmic accumulation as *Md*PER is first translated. The other experiment sought to drive down the *Md*PER levels in adults. Constant light was used because in *D. melanogaster* light stimuli results in PER degradation. However no cyclic accumulation of *Md*PER was observed in either experiment and the only thing they illustrated was *Md*PERs lack of sensitivity to light.

Setting up a cycle without PER nuclear entry.

The experimental evidence accumulated in this thesis fails to demonstrate either that there is rhythmic translation of MdPER or that it enters the nucleus in a gated manner. In fact the only *Md*PER positive cells identified showed that the protein was cyctoplasmic during both the light and dark phases of entrainment. This is in agreement with preliminary data from Piccin (1998) where *Md*PER positive cells had an entirely cytoplasmic localisation. All the experimental evidence suggests that MdPER is an entirely cytoplasmic protein in *Musca* heads. Therefore it is worth considering how a circadian molecular mechanism, which is largely conserved between *D. melanogaster* and mammals, could function without nuclear entry of PER.

One plausible explanation is that the *Md*PER protein no longer participates in the negative arm of the clock. Functional redundancy in the clock mechanism would allow the roles of PER to be carried out by other genes in the circadian molecular mechanism. Taking the most parsimonious view, where the majority of the clock genes in *D. melanogaster* have a conserved function in *Musca*, it might be possible for one or a few of the genes to perform PERs function.

The *D. melanogaster* clock mechanisms involve the formation of multi protein complexes. Published data illustrate that PER, TIM and DBT form a complex in the nucleus (Kloss et al. 2001), as do CRY, PER and TIM (Ceriani et al., 1999), and *d*CLOCK, CYC, PER and TIM (Lee et al., 1999). Therefore it is likely that there is a multi protein complex involving all or most of these genes in *D. melanogaster.*

The main function of the PER/TIM complex in the nucleus is to prevent CYC/*d*CLOCK activation of transcription, which is achieved without disrupting the CYC/dCLOCK complex (Lee et al., 1999). It is plausible that, in *Musca,* CRY or TIM could block transcription. CRY has already been shown to have different roles in *Drosophila* peripheral oscillators and is not just involved in photic input (Krishnan et al. 2001). Furthermore CRY forms part of the negative arm of the clock in mammals (Griffin et al., 1999). CRY may therefore perform the negative inhibition of transcription in *Musca*. It is less plausible that TIM could perform this function in *Musca* as sequence comparison of *d*TIM and a partial *Md*TIM clone did not find any notable sequence divergence (Jon Clayton, pers. comm.); additionally TIM has not been shown to have additional functions in any characterised clock systems.

Musca Doubletime

One problem created by the removal of PER from the negative arm of the circadian clock is that it is mainly responsible for introducing lags in the molecular mechanisms that ensure the clock has a 24hr rhythm (Price et al., 1998); removal of PER would presumably mean removing the contribution of *doubletime* to the timing of the circadian molecular mechanism. The phosphorylation of PER by DBT in *Drosophila* results in the degradation of PER, particularly when it is in a monomeric form during the light phase of entrainment (Kloss et al. 2001). This degradation results in the removal of PER from the nucleus and controls the timing of the start of transcription and the beginning of the next cycle; mutations in DBT delay the transcriptional profile of PER and result in the removal of the lag between transcription of *per* and translation (Price et al., 1998).

Although all of the experimental evidence suggests that *Md*PER does not enter the nucleus it cannot be said to unequivocally demonstrate that nuclear entry does not occur. Potentially the nuclear entry of *Md*PER in a very small subset of cells could provide clock function; the expression of PER in one LN of the *Drosophila disconnected* mutant has been shown to provide behavioural rhythmicity (HelfrichForster, 1996). Furthermore it may be possible that only a minority of PER enters the nucleus to inhibit transcription. This movement of a minority of MdPER would be difficult to observe, as it would still appear to have a predominantly cytoplasmic location.

Even assuming that *Md*PER enters the nucleus to fulfil its normal functions, the *Musca* clock may operate in quite a different fashion to the *D. melanogaster* clock as *Md*PER is not subject to the daily degradation cycle. In *D. melanogaster* phosphorylation

of PER by DBT provided the cue for degradation, and *dbt* mutants show fast or slow circadian phenotypes (Price et al., 1998). The phosphorylation of PER, therefore, is the key to regulating its degradation. When no phosphorylation of PER occurs, as in the *dbt* null mutant dbt^{P}, then PER accumulation is no longer dependent on TIM (Price et al., 1998), PER is no longer degraded and can accumulate to high levels. A series of mathematical models point to degradation being of paramount importance in adjusting the length of cycle. Using the Goodwin oscillator to model the molecular mechanisms of the clock, it was shown that mRNA and protein turnover was important in adjusting the cycle length (Ruoff et al., 1997; Ruoff et al., 1999). Further to this, two more models elaborate, and included phosphorylation targeting PER for degradation. It is concluded that the phosphorylation of PER and its subsequent degradation is partly responsible for the time lag between translation and negative feedback (Goldbeter, 1995; Leloup and Goldbeter, 1998), without which the clock would stop oscillating and return to a steady state.

The bare essentials of the clock is the time delay between transcription and negative feedback (Lema et al., 2000). It is clear from both the models and empirical evidence from the *dbt* mutants that phosphorylation and degradation are also key components in *D. melanogaster*. Additionally phosphorylation has been shown to be involved in circadian clocks of mammals (Keesler et al., 2000; Lowrey et al., 2000), *Neurospora* (Liu et al., 2000) and bacteria (Iwasaki et al., 2000). However phosphorylation changes in mobility of *Md*PER have not been observed and it appears hypophosphorylated; treatment to remove phosphorylation has little effect on the mobility of the protein extracts (Piccin, 1998). In contrast *D. melanogaster* PER

undergoes an easily observed change in mobility during the day due to phosphorylation (Edery et al., 1994). Therefore how the *Musca* clock maintains time is intriguing as there appears to be little or no phosphorylation of *Md*PER (Piccin, 1998). *Musca* therefore appears to be missing one of the important mechanisms for achieving the lag between transcription and translation of PER. This property appears not to be an intrinsic property of *Md* PER. *D. melanogaster* per^0 mutants rescued by a *Musca per* transgene were rhythmic. Furthermore western analysis shows that the protein cycles in these transgenic flies. However the protein appears hypophosphorylated when compared to wildtype *D. melanogaster* PER but there is some evidence of phosphorylation modification (Piccin, 1998). This apparent hypophosphorylation may be due in part to the truncation of the repetitive domain that is termed the Thr-Gly domain in *D. melanogaster,* which contains stretches of residues theoretically capable of phosphorylational modification. One would predict that if the *Md*PER protein was instead incapable of binding DBT in the *D. melanogaster* per^0 transformant then it should have an arrhythmic phenotype and PER should have a similar profile as the dbt^P mutant. Rescue of per^0, with *per* transgenes chimeric for *D. melanogaster per* sequence 5' of the TGs and *Musca per* 3' of the Thr-Gly, provided behavioural rescue with protein cycles that lagged the RNA in a similar manner to wildtype *D. melanogaster (Piccin, 1998)*. Therefore it is likely that DBT has normal function on *Md*PER in the transformants, and 5' sequences upstream of the Thr-Gly region in PER affect its dynamics

***D. virilis* and *D. pseudoobscura* molecular mechanisms.**

Superficially, the *D. virilis* circadian molecular mechanism, as described in chapter 6, functions in a similar manner to *D. melanogaster*. The *per* and *tim* transcripts cycle as do the PER and TIM proteins. This is in agreement with the *D. melanogaster* negative feedback model, but there are several aspects in which the *D. virilis* molecular mechanism differs. The *per* transcript peaks later than in *D. melanogaster (*ZT 14-16 compared to ZT 13) and there is a profound increase in the delay between the *per* and *tim* transcripts, with a lengthy delay of six hours between the *tim* and *per* transcript peaks, whereas *D. melanogaster per* peaks two hours after *tim* (ZT 13 and ZT 11 respectively) (So and Rosbash, 1997). Furthermore both the PER and TIM protein dynamics are significantly different, the *Dv*PER protein peaks at much the same time as *D. melanogaster(* ZT 20 and ZT 19 respectively) but does not show any easily observed change in apparent molecular weight as measured by migration under SDS-PAGE, during the circadian time course. This is in marked contrast to *D. melanogaster* PER, where phosphorylation profoundly alters its molecular weight (Edery et al., 1994). In addition the *Dv*PER accumulation and decline profile is shorter than *D. melanogatser* Furthermore *Dv*TIM appears to be less sensitive to light, has a modest cycle in abundance and shows an increase in molecular weight during the circadian time course, of which all three properties profoundly differ from *D. melanogaster* TIM, which is highly sensitive to light, has a large magnitude cycle in abundance and shows no significant change in molecular weight (HunterEnsor et al., 1996).

In conclusion, *D. virilis* may have a significantly different mechanism for generating the necessary lags between transcription and negative feedback. This lag has mainly been attributed to the interactions between DBT and PER in *D. melanogaster* (Kloss et al 2001; Suri et al 2000), however it appears improbable that this mechanism applies in *D. virilis*. The hypophosphorylation of *Dv*PER suggests a either a lower kinase activity for DBT or a PER sequence that is less prone to modification by the kinase. If the DBT kinase was indeed weaker you would predict it should provide a phenotype similar to the long *D. melanogaster dbt* mutants, dbt^{G}, dbt^{h}, and dbt^{L} (Suri et al. 2000; Price et al.1998). However *Dv*PER has a shorter accumulation and decline phase and is quickly degraded following the transition to light at ZT0. This contradicts the hypothesis that *Dv*DBT has a lower kinase activity; in *D. melanogaster* DBT acts to degrade PER from the nucleus, and the mutants $dbt^{G}$and dbt^{h} result in a slower PER decline (Kloss et al 2001; Suri et al 2000). It seems much more likely that the *Dv*PER provides a poor substrate for the kinase. The shortened accumulation and decline phase is somewhat similar to that of the per^{s} mutant which (Edery et al., 1994) which is likely to be a mutation in a site which is a target for phosphorylation (Toh et al., 2001). This hypothesis could be tested by using a *Dv*PER p-element construct to rescue per^{0} *D. melanogaster* flies. If the *Dv*PER sequence was responsible for the shortened profile then you would expect that it would produce a phenotype similar to per^{s} mutants

The accumulation profile of TIM suggests it may have a role in generating a lag between transcription and negative feedback. The molecular weight change observed for *Dv*TIM suggest that it is extensively modified by a kinase. Furthermore *Dv*TIMs apparent insensitivity to light and its slow degradation may mean that it is able to generate a lag by

inhibiting transcription by remaining in the nucleus for longer. This would be in sharp contrast to *D. melanogaster* TIM which has not been implicated in lag generation, other than partnering PER in the nucleus (Vosshall et al., 1994). It may be that the TIM kinase *shaggy* is important in the generation of the lag in *D. virilis* instead of the PER kinase DBT (Martinek et al., 2001).

Whatever mechanism is responsible for generating the lag between transcription and negative feedback in *D. virilis,* there must be coevolution within the molecular components that allows PER's shorter accumulation and decline to be compensated by the other molecules. It may be that the molecular mechanisms are in fact quite plastic and there is redundancy of function that would allow changes in the clock function of the genes to be tolerated. Indeed several studies have demonstrated that the molecular mechanism does appear to have a certain redundancy. Firstly, the removal of transcription activation of *per* did not prevent its mRNA from cycling (Frisch et al. 1994). Furthermore when the slow dbt^{G}, dbt^{h} were allowed to free run in constant darkness the protein profiles shifted and lengthened, but the mRNA profiles did not (Suri et al. 2000). In addition *D. melanogaster* photoreceptors have a circadian oscillator that can function without *per* mRNA cycling (Krishnan, et al 2001).

It is unfortunate that *D. pseudoobscura* PER and TIM could not be unambiguously identified during western analysis. Nevertheless the observation of a non-cycling PER candidate band suggest that functional redundancy of clock component has allowed the molecular mechanism of insect clocks to evolve different timing mechanisms.

The evidence that *per* mRNA did not respond to the PER protein accumulation shifts (Suri et al. 2000) and the fact that the promotorless *per* gene produced a cycling mRNA (Frisch et al. 1994) suggests that the *per* mRNA may be regulated following transcription. This certainly appears to be the case in both *D. virilis* and *D. pseudoobscura*. Both have lengthy delays between the *tim* and the *per* transcript accumulation. However under the feedback model their transcription should be activated and repressed at the same time. Therefore it is likely that the *per* mRNA accumulation is delayed. This may occur through mRNA binding factors which would process the *per* transcript and delay its accumulation. Regulation of the *per* mRNA had been shown to occur in *D. melanogaster* with evidence that there was a delay its accumulation (So and Rosbash, 1997).

It is unfortunate that *D. pseudoobscura* PER and TIM could not be unambiguously identified during western analysis. Nevertheless the observation of a non-cycling PER candidate band suggests that it may have a similar PER profile to *Musca.* Which would strengthen the hypothesis that functional redundancy of clock components has allowed the molecular mechanism of insect clocks to evolve different timing mechanisms.

The overall impression from comparative studies of the molecular clocks of insects *Antheraea pernyi, M. domestica D virilis* and *D. pseudoobscura* is that, although the genes are conserved, their mode of function may not be. Therefore the circadian clock may have been under selective pressure during speciation, which has lead to

different molecular functions for the genes. Circadian timing provides a compelling way of providing pre-mating isolation during speciation, flies who undergo foraging or courtship behaviours at different times of day are likely to become sexually isolated a process known to speed up speciation events. Indeed in *Drosophila* love song studies the interpulse interval (a component of the love song) has been shown to be altered in *period* mutants (Kyriacou, 1988), and this change may be a causative factor of pre-mating isolation. In addition to this the significant changes in the timing of *per/tim* molecular rhythms shown in this study may have been the result of selective pressure favouring speciation. Which may have acted to alter the molecular cycles of *per* and *tim* to create a pre-mating barrier and aid speciation.

Alternatively functional redundancy in the molecular component may have allowed the functions of the molecular mechanisms drift over time. The interconnected loops that reinforce the robustness of the clock cycle (Lee et al., 2000) may be responsible in part for this functional redundancy. This robust nature of the clock cycle may allow it components to undergo functional divergence as changes in their function are buffered by the robust ability of the clock to keep time.

Bibliography

CBS Phosphobase v 2.0. In A database of phosphorylation sites.

Abe, H., Honma, S., Namihira, M., Tanahashi, Y., Ikeda, M., and Honma, K. (1998). Circadian rhythm and light responsiveness of BMAL1 expression, a partner of mammalian clock gene Clock, in the suprachiasmatic nucleus of rats. Neuroscience Letters *258*, 93-96.

Albrecht, U., Sun, Z. S., Eichele, G., and Lee, C. C. (1997). A differential response of two putative mammalian circadian regulators, mper1 and mper2, to light. Cell *91*, 1055-1064.

Allada, R., White, N. E., So, W. V., Hall, J. C., and Rosbash, M. (1998). A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless. Cell *93*, 791-804.

Antoch, M. P., Song, E. J., Chang, A. M., Vitaterna, M. H., Zhao, Y. L., Wilsbacher, L. D., Sangoram, A. M., King, D. P., Pinto, L. H., and Takahashi, J. S. (1997). Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. Cell *89*, 655-667.

Aronson, B. D., Johnson, K. A., and Dunlap, J. C. (1994). Circadian Clock Locus Frequency - Protein Encoded By a Single Open Reading Frame Defines Period

Length and Temperature Compensation. Proceedings of the National Academy of Sciences of the United States of America *91*, 7683-7687.

Ausebel, F., Brent, R., Kingston, R., Moore, D., Seidman, J., Smith, J., and Struhl, K. (1990). Current protocols: Wiley).

Bae, K., Lee, C., Hardin, P. E., and Edery, I. (2000). dCLOCK is present in limiting amounts and likely mediates daily interactions between the dCLOCK-CYC transcription factor and the PER-TIM complex. Journal of Neuroscience *20*, 1746-1753.

Ballario, P., Talora, C., Galli, D., Linden, H., and Macino, G. (1998). Roles in dimerization and blue light photoresponse of the PAS and LOV domains of Neurospora crassa white collar proteins. Molecular Microbiology *29*, 719-729.

Ballario, P., Vittorioso, P., Magrelli, A., Talora, C., Cabibbo, A., and Macino, G. (1996). White collar-1, a central regulator of blue light responses in Neurospora, is a zinc finger protein. Embo Journal *15*, 1650-1657.

Bargiello, T. A. (1992). Correction. Nature *360*, 610-610.

Bargiello, T. A., Saez, L., Baylies, M. K., Gasic, G., Young, M. W., and Spray, D. C. (1987). The Drosophila Clock Gene Per Affects Intercellular Junctional Communication. Nature *328*, 686-691.

Bargiello, T. A., and Young, M. W. (1984). Molecular gentics of a biological clock in Drosophila. Proc. Natl. Acad. Sci. USA *81*, 2142-2146.

Baylies, M. K., Vosshall, L. B., Sehgal, A., and Young, M. W. (1992). New Short-Period Mutations of the Drosophila Clock Gene Per. Neuron *9*, 575-581.

Belvin, M. P., Zhou, H., and Yin, J. C. P. (1999). The Drosophila dCREB2 gene affects the circadian clock. Neuron *22*, 777-787.

Benna, C., Scannapieco, P., Piccin, A., Sandrelli, F., Zordan, M., Rosato, E., Kyriacou, C. P., Valle, G., and Costa, R. (2000). A second timeless gene in Drosophila shares greater sequence similarity with mammalian tim. Current Biology *10*, R512-R513.

Blau, J., and Young, M. W. (1999). Cycling vrille expression is required for a functional Drosophila clock. Cell *99*, 661-671.

Branden, C., and Tooze, J. (1991). Introduction to Protein Structure (New York and London: Gerland Publishing Inc).

Brandes, C., Plautz, J. D., Stanewsky, R., Jamison, C. F., Straume, M., Wood, K. V., Kay, S. A., and Hall, J. C. (1996). Novel features of Drosophila period transcription revealed by real-time luciferase reporting. Neuron *16*, 687-692.

Bunger, M. K., Wilsbacher, L. D., Moran, S. M., Clendenin, C., Radcliffe, L. A., Hogenesch, J. B., Simon, M. C., Takahashi, J. S., and Bradfield, C. A. (2000). Mop3 is an essential component of the master circadian pacemaker in mammals. Cell *103*, 1009-1017.

Castiglionemorelli, M. A., Guantieri, V., Villani, V., Kyriacou, C. P., Costa, R., and Tamburro, A. M. (1995). Conformational Study of the Thr-Gly Repeat in the Drosophila Clock Protein, Period. Proceedings of the Royal Society of London Series B-Biological Sciences *260*, 155-163.

Ceriani, M. F., Darlington, T. K., Staknis, D., Mas, P., Petti, A. A., Weitz, C. J., and Kay, S. A. (1999). Light-dependent sequestration of TIMELESS by CRYPTOCHROME. Science *285*, 553-556.

Cheng, Y. Z., Gvakharia, B., and Hardin, P. E. (1998). Two alternatively spliced transcripts from the Drosophila period gene rescue rhythms having different molecular and behavioral characteristics. Molecular and Cellular Biology *18*, 6505-6514.

Citri, Y., Colot, H. V., Jacquier, A. C., Qiang, Y., Hall, J. C., Baltimore, D., and Rosbash, M. (1987). A Family of Unusually Spliced Biologically-Active Transcripts Encoded By a Drosophila Clock Gene. Nature *326*, 42-47.

Creighton, T. E. (1997). Protein Structure: A practical appraoch, 2nd Edition, B. D. Hames, ed. (Oxford New York Tokyo: IRL press at Oxford University Press).

Crews, S. T., Thomas, J. B., and Goodman, C. S. (1988). The Drosophila Single-Minded Gene Encodes a Nuclear-Protein With Sequence Similarity to the Per Gene-Product. Cell *52*, 143-151.

Crosthwaite, S. K., Dunlap, J. C., and Loros, J. J. (1997). Neurospora wc-1 and wc-2: Transcription, photoresponses, and the origins of circadian rhythmicity. Science *276*, 763-769.

Curtin, K. D., Huang, Z. J., and Rosbash, M. (1995). Temporally Regulated Nuclear Entry of the Drosophila Period Protein Contributes to the Circadian Clock. Neuron *14*, 365-372.

Darby, N. J., and Creighton, T. E. (1990). Folding Proteins. Nature *344*, 715-716.

Darlington, T. K., Wager-Smith, K., Ceriani, M. F., Staknis, D., Gekakis, N., Steeves, T. D. L., Weitz, C. J., Takahashi, J. S., and Kay, S. A. (1998). Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. Science *280*, 1599-1603.

Derman, A. I., Prinz, W. A., Belin, D., and Beckwith, J. (1993). Mutations That Allow Disulfide Bond Formation in the Cytoplasm of Escherichia-Coli. Science *262*, 1744-1747.

Edery, I., Rutila, J. E., and Rosbash, M. (1994). Phase-Shifting of the Circadian Clock By Induction of the Drosophila Period Protein. Science *263*, 237-240.

Edery, I., Zwiebel, L. J., Dembinska, M. E., and Rosbash, M. (1994). Temporal Phosphorylation of the Drosophila Period Protein. Proceedings of the National Academy of Sciences of the United States of America *91*, 2260-2264.

Emery, P., So, W. V., Kaneko, M., Hall, J. C., and Rosbash, M. (1998). CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity. Cell *95*, 669-679.

Emery, P., Stanewsky, R., Hall, J. C., and Rosbash, M. (2000). Drosophila cryptochromes - A unique circadian-rhythm photoreceptor. Nature *404*, 456-457.

Emery, P., Stanewsky, R., Helfrich-Forster, C., Emery-Le, M., Hall, J. C., and Rosbash, M. (2000). Drosophila CRY is a deep brain circadian photoreceptor. Neuron *26*, 493-504.

Ewer, J., Frisch, B., Hamblencoyle, M. J., Rosbash, M., and Hall, J. C. (1992). Expression of the Period Clock Gene Within Different Cell-Types in the Brain of Drosophila Adults and Mosaic Analysis of These Cells Influence On Circadian Behavioral Rhythms. Journal of Neuroscience *12*, 3321-3349.

Ewer, J., Rosbash, M., and Hall, J. C. (1988). An Inducible Promoter Fused to the Period Gene in Drosophila Conditionally Rescues Adult Per-Mutant Arrhythmicity. Nature *333*, 82-84.

Field, M. D., Maywood, E. S., O'Brien, J. A., Weaver, D. R., Reppert, S. M., and Hastings, M. H. (2000). Analysis of clock proteins in mouse SCN demonstrates phylogenetic divergence of the circadian clockwork and resetting mechanisms. Neuron *25*, 437-447.

Fischer, B., Sumner, I., and Goodenough, P. (1992). Isolation and Renaturation of Bioactive Proteins Expressed in Escherichia-Coli As Inclusion-Bodies. Arzneimittel-Forschung/Drug Research *42-2*, 1512-1515.

Fischer, B., Sumner, I., and Goodenough, P. (1993). Isolation, Renaturation, and Formation of Disulfide Bonds of Eukaryotic Proteins Expressed in Escherichia-Coli As Inclusion- Bodies. Biotechnology and Bioengineering *41*, 3-13.

Frisch, B., Fleissner, G., Brandes, C., and Hall, J. C. (1996). Staining in the brain of Pachymorpha sexguttata mediated by an antibody against a Drosophila clock-gene product: Labeling of cells with possible importance for the beetle's circadian rhythms. Cell and Tissue Research *286*, 411-429.

Frisch, B., Hardin, P. E., Hamblencoyle, M. J., Rosbash, M., and Hall, J. C. (1994). A Promoterless Period Gene Mediates Behavioral Rhythmicity and Cyclical Per

Expression in a Restricted Subset of the Drosophila Nervous-System. Neuron *12*, 555-570.

Garceau, N. Y., Liu, Y., Loros, J. J., and Dunlap, J. C. (1997). Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. Cell *89*, 469-476.

Gekakis, N., Saez, L., Delahayebrown, A. M., Myers, M. P., Sehgal, A., Young, M. W., and Weitz, C. J. (1995). Isolation of Timeless By Per Protein-Interaction - Defective Interaction Between Timeless Protein and Long-Period Mutant Per(L). Science *270*, 811-815.

Gekakis, N., Staknis, D., Nguyen, H. B., Davis, F. C., Wilsbacher, L. D., King, D. P., Takahashi, J. S., and Weitz, C. J. (1998). Role of the CLOCK protein in the mammalian circadian mechanism. Science *280*, 1564-1569.

George, H., and Terracol, R. (1997). The vrille gene of Drosophila is a maternal enhancer of decapentaplegic and encodes a new member of the bZIP family of transcription factors. Genetics *146*, 1345-1363.

Glossop, N. R. J., Lyons, L. C., and Hardin, P. E. (1999). Interlocked feedback loops within the Drosophila circadian oscillator. Science *286*, 766-768.

Goldbeter, A. (1995). A Model For Circadian Oscillations in the Drosophila Period Protein (Per). Proceedings of the Royal Society of London Series B-Biological Sciences *261*, 319-324.

Gonze, D., Leloup, J. C., and Goldbeter, A. (2000). Theoretical models for circadian rhythms in Neurospora and Drosophila. Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences *323*, 57-67.

Gotter, A. L., Levine, J. D., and Reppert, S. M. (1999). Sex-linked period genes in the silkmoth, Antheraea pernyi: Implications for circadian clock regulation and the evolution of sex chromosomes. Neuron *24*, 953-965.

Griffin, E. A., Staknis, D., and Weitz, C. J. (1999). Light-independent role of CRY1 and CRY2 in the mammalian circadian clock. Science *286*, 768-771.

Guise, A. D., West, S. M., and Chaudhuri, J. B. (1996). Protein folding in vivo and renaturation of recombinant proteins from inclusion bodies. Molecular Biotechnology *6*, 53-64.

Hamblen, M., Zehring, W. A., Kyriacou, C. P., Reddy, P., Yu, Q., Wheeler, D. A., Zwiebel, L. J., Konopka, R. J., Rosbash, M., and Hall, J. C. (1986). Germ-Line Transformation Involving Dna From the Period Locus in Drosophila-Melanogaster - Overlapping Genomic Fragments That Restore Circadian and Ultradian Rhythmicity to Per0 and Per- Mutants. Journal of Neurogenetics *3*, 249-291.

Handler, A. M., and Konopka, R. J. (1976). Transplanation of a circadian pacemaker in Drosophila. Nature *279*, 236-238.

Hardin, P. E. (1994). Analysis of Period Messenger-Rna Cycling in Drosophila Head and Body-Tissues Indicates That Body Oscillators Behave Differently From Head Oscillators. Molecular and Cellular Biology *14*, 7211-7218.

Hardin, P. E., Hall, J. C., and Rosbash, M. (1992). Behavioral and Molecular Analyses Suggest That Circadian Output Is Disrupted By Disconnected Mutants in Drosophila-Melanogaster. Embo Journal *11*, 1-6.

Hardin, P. E., Hall, J. C., and Rosbash, M. (1992). Circadian Oscillations in Period Gene Messenger-Rna Levels Are Transcriptionally Regulated. Proceedings of the National Academy of Sciences of the United States of America *89*, 11711-11715.

Hardin, P. E., Hall, J. C., and Rosbash, M. (1990). Feedback of the Drosophila Period Gene-Product On Circadian Cycling of Its Messenger-Rna Levels. Nature *343*, 536-540.

Harlow, E., and Lane, D. (1988). Antibodies: A Laboratory Manual: Cold Spring Harbour Laboratories).

HelfrichForster, C. (1996). Drosophila rhythms: From brain to behavior. Seminars in Cell & Developmental Biology *7*, 791-802.

Helfrichforster, C. (1995). The Period Clock Gene Is Expressed in Central-Nervous-System Neurons Which Also Produce a Neuropeptide That Reveals the Projections of Circadian Pacemaker Cells Within the Brain of Drosophila-Melanogaster. Proceedings of the National Academy of Sciences of the United States of America *92*, 612-616.

Hida, A., Koike, N., Hirose, M., Hattori, M., Sakaki, Y., and Tei, H. (2000). The human and mouse Period1 genes: Five well-conserved E-boxes additively contribute to the enhancement of mPer1 transcription. Genomics *65*, 224-233.

Huang, Z. J., Curtin, K. D., and Rosbash, M. (1995). Per Protein Interactions and Temperature Compensation of a Circadian Clock in Drosophila. Science *267*, 1169-1172.

Huang, Z. J., Edery, I., and Rosbash, M. (1993). Pas Is a Dimerization Domain Common to Drosophila Period and Several Transcription Factors. Nature *364*, 259-262.

HunterEnsor, M., Ousley, A., and Sehgal, A. (1996). Regulation of the Drosophila protein timeless suggests a mechanism for resetting the circadian clock by light. Cell *84*, 677-685.

Ikeda, M., and Nomura, M. (1997). cDNA cloning and tissue-specific expression of a novel basic helix-loop-helix/PAS protein (BMAL1) and identification of alternatively

spliced variants with alternative translation initiation site usage. Biochemical and Biophysical Research Communications *233*, 258-264.

Iwasaki, H., and Dunlap, J. C. (2000). Microbial circadian oscillatory systems in Neurospora and Synechococcus: models for cellular clocks. Current Opinion in Microbiology *3*, 189-196.

Iwasaki, H., and Kondo, T. (2000). The current state and problems of circadian clock studies in cyanobacteria. Plant and Cell Physiology *41*, 1013-1020.

Iwasaki, H., Williams, S. B., Kitayama, Y., Ishiura, M., Golden, S. S., and Kondo, T. (2000). A KaiC-interacting sensory histidine kinase, SasA, necessary to sustain robust circadian oscillation in cyanobacteria. Cell *101*, 223-233.

Jackson, F. R., Schroeder, A. J., Roberts, M. A., McNeil, G. P., Kume, K., and Akten, B. (2001). Cellular and molecular mechanisms of circadian control in insects. Journal of Insect Physiology *47*, 833-842.

Karin, M., and Hunter, T. (1995). Transcriptional Control By Protein-Phosphorylation - Signal Transmission From the Cell-Surface to the Nucleus. Current Biology *5*, 747-757.

Keesler, G. A., Camacho, F., Guo, Y., Virshup, D., Mondadori, C., and Yao, Z. B. (2000). Phosphorylation and destabilization of human period I clock protein by human casein kinase I epsilon. Neuroreport *11*, 951-955.

King, D. P., Zhao, Y. L., Sangoram, A. M., Wilsbacher, L. D., Tanaka, M., Antoch, M. P., Steeves, T. D. L., Vitaterna, M. H., Kornhauser, J. M., Lowrey, P. L., Turek, F. W., and Takahashi, J. S. (1997). Positional cloning of the mouse circadian Clock gene. Cell *89*, 641-653.

Kloss, B., Price, J. L., Saez, L., Blau, J., Rothenfluh, A., Wesley, C. S., and Young, M. W. (1998). The Drosophila clock gene double-time encodes a protein closely related to human casein kinase I epsilon. Cell *94*, 97-107.

Kloss, B., Rothenfluh, A., Young, M. W., and Saez, L. (2001). Phosphorylation of PERIOD is influenced by cycling physical associations of DOUBLE-TIME, PERIOD, and TIMELESS in the Drosophila clock. Neuron *30*, 699-706.

Konopka, R., Wells, S., and Lee, T. (1983). Mosiac Analysis of a Drosophila Clock Mutant. Mol Gen Genet *190*, 284-288.

Krishnan, B., Levine, J. D., Lynch, M. K. S., Dowse, H. B., Funes, P., Hall, J. C., Hardin, P. E., and Dryer, S. E. (2001). A new role for cryptochrome in a Drosophila circadian oscillator. Nature *411*, 313-317.

Kume, K., Zylka, M. J., Sriram, S., Shearman, L. P., Weaver, D. R., Jin, X. W., Maywood, E. S., Hastings, M. H., and Reppert, S. M. (1999). mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. Cell *98*, 193-205.

Kyriacou, C. P. (1988). Phenotypic Effects of the Period Gene in Drosophila. Journal of Interdisciplinary Cycle Research *19*, 189-189.

Lakin-Thomas, P. L. (2000). Circadian rhythms - new functions for old clock genes? Trends in Genetics *16*, 135-142.

Lee, C., Bae, K., and Edery, I. (1999). PER and TIM inhibit the DNA binding activity of a Drosophila CLOCK-CYC/dBMAL1 heterodimer without disrupting formation of the heterodimer: a basis for circadian transcription. Molecular and Cellular Biology *19*, 5316-5325.

Lee, C. G., Bae, K. H., and Edery, I. (1998). The Drosophila CLOCK protein undergoes daily rhythms in abundance, phosphorylation, and interactions with the PER-TIM complex. Neuron *21*, 857-867.

Lee, C. G., Parikh, V., Itsukaichi, T., Bae, K., and Edery, I. (1996). Resetting the Drosophila clock by photic regulation of PER and a PER-TIM complex. Science *271*, 1740-1744.

Lee, K., Loros, J. J., and Dunlap, J. C. (2000). Interconnected feedback loops in the Neurospora circadian system. Science *289*, 107-110.

Leloup, J. C., and Goldbeter, A. (1998). A model for circadian rhythms in Drosophila incorporating the formation of a complex between the PER and TIM proteins. Journal of Biological Rhythms *13*, 70-87.

Lema, M. A., Golombek, D. A., and Echave, J. (2000). Delay model of the circadian pacemaker. Journal of Theoretical Biology *204*, 565-573.

Levine, J. D., Sauman, I., Imbalzano, M., Reppert, S. M., and Jackson, F. R. (1995). Period Protein From the Giant Silkmoth Antheraea-Pernyi Functions As a Circadian Clock Element in Drosophila- Melanogaster. Neuron *15*, 147-157.

Lewis, R. D., Warman, G. R., and Saunders, D. S. (1997). Simulations of free-running rhythms, light entrainment and the light-pulse phase response curves for the locomotor activity rhythm in period mutant of Drosophila melanogaster. Journal of Theoretical Biology *185*, 503-510.

Lin, K. H., and Cheng, S. Y. (1991). An Efficient Method to Purify Active Eukaryotic Proteins From the Inclusion-Bodies in Escherichia-Coli. Biotechniques *11*, 748-&.

Liu, X., Lorenz, L., Yu, Q. N., Hall, J. C., and Rosbash, M. (1988). Spatial and Temporal Expression of the Period Gene in Drosophila-Melanogaster. Genes & Development *2*, 228-238.

Liu, X., Zwiebel, L. J., Hinton, D., Benzer, S., Hall, J. C., and Rosbash, M. (1992). The Period Gene Encodes a Predominantly Nuclear-Protein in Adult Drosophila. Journal of Neuroscience *12*, 2735-2744.

Liu, Y., Garceau, N. Y., Loros, J. J., and Dunlap, J. C. (1997). Thermally regulated translational control of FRQ mediates aspects of temperature responses in the Neurospora circadian clock. Cell *89*, 477-486.

Liu, Y., Loros, J., and Dunlap, J. C. (2000). Phosphorylation of the Neurospora clock protein FREQUENCY determines its degradation rate and strongly influences the period length of the circadian clock. Proceedings of the National Academy of Sciences of the United States of America *97*, 234-239.

Lloyd, A. L., and Lloyd, D. (1993). Hypothesis - the Central Oscillator of the Circadian Clock Is a Controlled Chaotic Attractor. Biosystems *29*, 77-85.

Longenecker, K. L., Roach, P. J., and Hurley, T. D. (1998). Crystallographic studies of casein kinase I delta: toward a structural understanding of auto-inhibition. Acta Crystallographica Section D-Biological Crystallography *54*, 473-475.

Lorenz, L. J., Hall, J. C., and Rosbash, M. (1989). Expression of a Drosophila Messenger-Rna Is Under Circadian Clock Control During Pupation. Development *107*, 869-880.

Loudon, A. S. I., Wayne, N. L., Krieg, R., Iranmanesh, A., Veldhuis, J. D., and Menaker, M. (1994). Ultradian Endocrine Rhythms Are Altered by a Circadian Mutation in the Syrian-Hamster. Endocrinology *135*, 712-718.

Lowrey, P. L., Shimomura, K., Antoch, M. P., Yamazaki, S., Zemenides, P. D., Ralph, M. R., Menaker, M., and Takahashi, J. S. (2000). Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. Science *288*, 483-491.

Luo, C. H., Loros, J. J., and Dunlap, J. C. (1998). Nuclear localization is required for function of the essential clock protein FRQ. Embo Journal *17*, 1228-1235.

Majercak, J., Sidote, D., Hardin, P. E., and Edery, I. (1999). How a circadian clock adapts to seasonal decreases in temperature and day length. Neuron *24*, 219-230.

Maniatis, T., Fritsch, E., and Sambrook, J. (1981). Molecular Cloning: a laboratory manual: Cold Spring Harbour Laboratory).

Martinek, S., Inonog, S., Manoukian, A. S., and Young, M. W. (2001). A role for the segment polarity gene shaggy/GSK-3 in the Drosophila circadian clock. Cell *105*, 769-779.

McClung, C. R., Fox, B. A., and Dunlap, J. C. (1989). The Neurospora Clock Gene-Frequency Shares a Sequence Element With the Drosophila Clock Gene Period. Nature *339*, 558-562.

McNeil, G. P., Zhang, X. L., Genova, G., and Jackson, F. R. (1998). A molecular rhythm mediating circadian clock output in Drosophila. Neuron *20*, 297-303.

Merrow, M. W., Garceau, N. Y., and Dunlap, J. C. (1997). Dissection of a circadian oscillation into discrete domains. Proceedings of the National Academy of Sciences of the United States of America *94*, 3877-3882.

Myers, M. P., WagerSmith, K., RothenfluhHilfiker, A., and Young, M. W. (1996). Light-induced degradation of TIMELESS and entrainment of the Drosophila circadian clock. Science *271*, 1736-1740.

Myers, M. P., Wagersmith, K., Wesley, C. S., Young, M. W., and Sehgal, A. (1995). Positional Cloning and Sequence-Analysis of the Drosophila Clock Gene, Timeless. Science *270*, 805-808.

Newby, L. M., and Jackson, F. R. (1993). A New Biological Rhythm Mutant of Drosophila-Melanogaster That Identifies a Gene With an Essential Embryonic Function. Genetics *135*, 1077-1090.

Newby, L. M., and Jackson, F. R. (1996). Regulation of a specific circadian clock output pathway by lark, a putative RNA-binding protein with repressor activity. Journal of Neurobiology *31*, 117-128.

Nigg, E. A. (1997). Nucleocytoplasmic transport: Signals, mechanisms and regulation. Nature *386*, 779-787.

Nishiwaki, T., Iwasaki, H., Ishiura, M., and Kondo, T. (2000). Nucleotide binding and autophosphorylation of the clock protein KaiC as a circadian timing process of cyanobacteria. Proceedings of the National Academy of Sciences of the United States of America *97*, 495-499.

Okamura, H., Miyake, S., Sumi, Y., Yamaguchi, S., Yasui, A., Muijtjens, M., Hoeijmakers, J. H. J., and van der Horst, G. T. J. (1999). Photic induction of mPer1 and mPer2 in Cry-deficient mice lacking a biological clock. Science *286*, 2531-2534.

Ousley, A., Zafarullah, K., Chen, Y. F., Emerson, M., Hickman, L., and Sehgal, A. (1998). Conserved regions of the timeless (tim) clock gene in Drosophila analyzed through phylogenetic and functional studies. Genetics *148*, 815-825.

Park, J. H., and Hall, J. C. (1998). Isolation and chronobiological analysis of a neuropeptide pigment-dispersing factor gene in Drosophila melanogaster. Journal of Biological Rhythms *13*, 219-228.

Park, J. H., Helfrich-Forster, C., Lee, G., Liu, L., Rosbash, M., and Hall, J. C. (2000). Differential regulation of circadian pacemaker output by separate clock genes in Drosophila. Proceedings of the National Academy of Sciences of the United States of America *97*, 3608-3613.

Peixoto, A. A., Costa, R., Wheeler, D. A., Hall, J. C., and Kyriacou, C. P. (1992). Evolution of the Threonine-Glycine Repeat Region of the Period Gene in the Melanogaster Species Subgroup of Drosophila. Journal of Molecular Evolution *35*, 411-419.

Petersen, G., Hall, J. C., and Rosbash, M. (1988). The Period Gene of Drosophila Carries Species-Specific Behavioral Instructions. Embo Journal *7*, 3939-3947.

Piccin. (1998). The clock gene period int the Housefly, Musca domestica: a molecular analysis. In Genetics (Leicester: University of Leicester).

Piccin, A., Couchman, M., Clayton, J. D., Chalmers, D., Costa, R., and Kyriacou, C. P. (2000). The clock gene period of the housefly, Musca domestica, rescues behavioral rhythmicity in Drosophila melanogaster: Evidence for intermolecular coevolution? Genetics *154*, 747-758.

Plautz, J. D., Kaneko, M., Hall, J. C., and Kay, S. A. (1997). Independent photoreceptive circadian clocks throughout Drosophila. Science *278*, 1632-1635.

Plautz, J. D., Straume, M., Stanewsky, R., Jamison, C. F., Brandes, C., Dowse, H. B., Hall, J. C., and Kay, S. A. (1997). Quantitative analysis of Drosophila period gene transcription in living animals. Journal of Biological Rhythms *12*, 204-217.

Ponting, C. P., and Aravind, L. (1997). PAS: a multifunctional domain family comes to light. Current Biology *7*, R674-R677.

Price, J. L. (1995). Are Competing Intermolecular and Intramolecular Interactions of Period Protein Important For the Regulation of Circadian- Rhythms in Drosophila. Bioessays *17*, 583-586.

Price, J. L. (1997). Insights into the molecular mechanisms of temperature compensation from the Drosophila period and timeless mutants. Chronobiology International *14*, 455-468.

Price, J. L., Blau, J., Rothenfluh, A., Abodeely, M., Kloss, B., and Young, M. W. (1998). Double-time is a novel Drosophila clock gene that regulates PERIOD protein accumulation. Cell *94*, 83-95.

Price, J. L., Dembinska, M. E., Young, M. W., and Rosbash, M. (1995). Suppression of Period Protein Abundance and Circadian Cycling By the Drosophila Clock Mutation Timeless. Embo Journal *14*, 4044-4049.

Reddy, P., Jacquier, A. C., Abovich, N., Petersen, G., and Rosbash, M. (1986). The Period Clock Locus of D Melanogaster Codes For a Proteoglycan. Cell *46*, 53-61.

Regier, J. C., Fang, Q. Q., Mitter, C., Peigler, R. S., Friedlander, T. P., and Solis, M. A. (1998). Evolution and phylogenetic utility of the period gene in Lepidoptera. Molecular Biology and Evolution *15*, 1172-1182.

Renn, S. C. P., Park, J. H., Rosbash, M., Hall, J. C., and Taghert, P. H. (1999). A pdf neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in Drosophila. Cell *99*, 791-802.

Reppert, S. M. (2000). Cellular and molecular basis of circadian timing in mammals. Seminars in Perinatology *24*, 243-246.

Reppert, S. M., and Sauman, I. (1995). Period and Timeless Tango - a Dance of 2 Clock Genes. Neuron *15*, 983-986.

Reppert, S. M., Tsai, T., Roca, A. L., and Sauman, I. (1994). Cloning of a Structural and Functional Homolog of the Circadian Clock Gene Period From the Giant Silkmoth Antheraea-Pernyi. Neuron *13*, 1167-1176.

Reppert, S. M., and Weaver, D. R. (2000). Comparing clockworks: Mouse versus fly. Journal of Biological Rhythms *15*, 357-364.

Riley, M., and Labedan, B. (1997). Protein evolution viewed through Escherichia coli protein sequences: Introducing the notion of a structural segment of homology, the module. Journal of Molecular Biology *268*, 857-868.

Rogl, H., Kosemund, K., Kuhlbrandt, W., and Collinson, I. (1998). Refolding of Escherichia coli produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. Febs Letters *432*, 21-26.

Rost, B., and Sander, C. (1993). Improved prediction of protein secondary structure by use of sequence profiles and neural networks. Proc. Natl. Acad. Sci. U.S.A. *90*, 7558-7562.

Ruoff, P., Rensing, L., Kommedal, R., and Mohsenzadeh, S. (1997). Modeling temperature compensation in chemical and biological oscillators. Chronobiology International *14*, 499-510.

Ruoff, P., Vinsjevik, M., Monnerjahn, C., and Rensing, L. (1999). The Goodwin oscillator: On the importance of degradation reactions in the circadian clock. Journal of Biological Rhythms *14*, 469-479.

Rutila, J. E., Maltseva, O., and Rosbash, M. (1998). The tim(SL) mutant affects a restricted portion of the Drosophila melanogaster circadian cycle. Journal of Biological Rhythms *13*, 380-392.

Rutila, J. E., Suri, V., Le, M., So, W. V., Rosbash, M., and Hall, J. C. (1998). CYCLE is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of Drosophila period and timeless. Cell *93*, 805-814.

Saez, L., and Young, M. W. (1996). Regulation of nuclear entry of the Drosophila clock proteins period and timeless. Neuron *17*, 911-920.

Sangoram, A. M., Saez, L., Antoch, M. P., Gekakis, N., Staknis, D., Whiteley, A., Fruechte, E. M., Vitaterna, M. H., Shimomura, K., King, D. P., Young, M. W., Weitz,

C. J., and Takahashi, J. S. (1998). Mammalian circadian autoregulatory loop: A timeless ortholog and mPer1 interact and negatively regulate CLOCK-BMAL1-induced transcription. Neuron *21*, 1101-1113.

Sauman, I., and Reppert, S. M. (1998). Brain control of embryonic circadian rhythms in the silkmoth Antheraea pernyi. Neuron *20*, 741-748.

Sauman, I., and Reppert, S. M. (1996). Circadian clock neurons in the silkmoth Antheraea pernyi: Novel mechanisms of period protein regulation. Neuron *17*, 889-900.

Sauman, I., and Reppert, S. M. (1996). Molecular characterization of prothoracicotropic hormone (PTTH) from the giant silkmoth Antheraea pernyi: Developmental appearance of PTTH-Expressing cells and relationship to circadian clock cells in central brain. Developmental Biology *178*, 418-429.

Sauman, I., Tsai, T., Roca, A. L., and Reppert, S. M. (1996). Period protein is necessary for circadian control of egg hatching behavior in the silkmoth Antheraea pernyi. Neuron *17*, 901-909.

Sauman, I., Tsai, T., Roca, A. L., and Reppert, S. M. (2000). Period protein is necessary for circadian control of egg hatching behavior in the silkmoth Antheraea pernyi (vol 17, pg 901, 1996). Neuron *27*, U12-U12.

Schein, C. H., and Noteborn, M. H. M. (1988). Formation of Soluble Recombinant Proteins in Escherichia-Coli Is Favored By Lower Growth Temperature. Bio-Technology *6*, 291-294.

Schmitz, O., Katayama, M., Williams, S. B., Kondo, T., and Golden, S. S. (2000). CikA, a bacteriophytochrome that resets the cyanobacterial circadian clock. Science *289*, 765-768.

Sehgal, A. (1995). Genetic Dissection of the Circadian Clock - a Timeless Story. Seminars in the Neurosciences *7*, 27-35.

Sehgal, A., Price, J. L., Man, B., and Young, M. W. (1994). Loss of Circadian Behavioral Rhythms and Per Rna Oscillations in the Drosophila Mutant Timeless. Science *263*, 1603-1606.

Sehgal, A., Rothenfluhhilfiker, A., Hunterensor, M., Chen, Y. F., Myers, M. P., and Young, M. W. (1995). Rhythmic Expression of Timeless - a Basis For Promoting Circadian Cycles in Period Gene Autoregulation. Science *270*, 808-810.

Shearman, L. P., Jin, X. W., Lee, C. G., Reppert, S. M., and Weaver, D. R. (2000). Targeted disruption of the mPer3 gene: Subtle effects on circadian clock function. Molecular and Cellular Biology *20*, 6269-6275.

Shearman, L. P., Sriram, S., Weaver, D. R., Maywood, E. S., Chaves, I., Zheng, B. H., Kume, K., Lee, C. C., van der Horst, G. T. J., Hastings, M. H., and Reppert, S. M.

(2000). Interacting molecular loops in the mammalian circadian clock. Science *288*, 1013-1019.

Shearman, L. P., Zylka, M. J., Weaver, D. R., Kolakowski, L. F., and Reppert, S. M. (1997). Two period homologs: Circadian expression and photic regulation in the suprachiasmatic nuclei. Neuron *19*, 1261-1269.

Shigeyoshi, Y., Taguchi, K., Yamamoto, S., Takekida, S., Yan, L., Tei, H., Moriya, T., Shibata, S., Loros, J. J., Dunlap, J. C., and Okamura, H. (1997). Light-induced resetting of a mammalian circadian clock is associated with rapid induction of the mPer1 transcript. Cell *91*, 1043-1053.

Shimamoto, N., Kasciukovich, T., Nagia, H., and Hayward, R. (1998). Effiecent solubalization of prtoeins overproduced as incluion bodies by use of an extreme concentration of glycerol. Technical Tips Online.

Siwicki, K. K., Eastman, C., Petersen, G., Rosbash, M., and Hall, J. C. (1988). Antibodies to the Period Gene-Product of Drosophila Reveal Diverse Tissue Distribution and Rhythmic Changes in the Visual- System. Neuron *1*, 141-150.

So, W. V., and Rosbash, M. (1997). Post-transcriptional regulation contributes to Drosophila clock gene mRNA cycling. Embo Journal *16*, 7146-7155.

So, W. V., Sarov-Blat, L., Kotarski, C. K., McDonald, M. J., Allada, R., and Rosbash, M. (2000). Takeout, a novel Drosophila gene under Circadian Clock transcriptional regulation. Molecular and Cellular Biology *20*, 6935-6944.

Strayer, C., Oyama, T., Schultz, T. F., Raman, R., Somers, D. E., Mas, P., Panda, S., Kreps, J. A., and Kay, S. A. (2000). Cloning of the Arabidopsis clock cone TOC1, an autoregulatory response regulator homolog. Science *289*, 768-771.

Sun, Z. S., Albrecht, U., Zhuchenko, O., Bailey, J., Eichele, G., and Lee, C. C. (1997). RIGUI, a putative mammalian ortholog of the Drosophila period gene. Cell *90*, 1003-1011.

Suri, V., Hall, J. C., and Rosbash, M. (2000). Two novel doubletime mutants alter circadian properties and eliminate the delay between RNA and protein in Drosophila. Journal of Neuroscience *20*, 7547-7555.

Suri, V., Lanjuin, A., and Rosbash, M. (1999). TIMELESS-dependent positive and negative autoregulation in the Drosophila circadian clock. Embo Journal *18*, 675-686.

Suri, V. P., Qian, Z. W., Hall, J. C., and Rosbash, M. (1998). Evidence that the TIM light response is relevant to light- induced phase shifts in Drosophila melanogaster. Neuron *21*, 225-234.

Takahata, S., Ozaki, T., Mimura, J., Kikuchi, Y., Sogawa, K., and Fujii-Kuriyama, Y. (2000). Transactivation mechanisms of mouse clock transcription factors, mClock and mArnt3. Genes to Cells *5*, 739-747.

Takumi, T., Matsubara, C., Shigeyoshi, Y., Taguchi, K., Yagita, K., Maebayashi, Y., Sakakida, Y., Okumura, K., Takashima, N., and Okamura, H. (1998). New mammalian period gene predominantly expressed in the suprachiasmatic nucleus. Genes to Cells *3*, 167-176.

Takumi, T., Nagamine, Y., Miyake, S., Matsubara, C., Taguchi, K., Takekida, S., Sakakida, Y., Nishikawa, K., Kishimoto, T., Niwa, S., Okumura, K., and Okamura, H. (1999). A mammalian ortholog of Drosophila timeless, highly expressed in SCN and retina, forms a complex with mPER1. Genes to Cells *4*, 67-75.

Takumi, T., Taguchi, K., Miyake, S., Sakakida, Y., Takashima, N., Matsubara, C., Maebayashi, Y., Okumura, K., Takekida, S., Yamamoto, S., Yagita, K., Yan, L., Young, M. W., and Okamura, H. (1998). A light-independent oscillatory gene mPer3 in mouse SCN and OVLT. Embo Journal *17*, 4753-4759.

Taylor, B. L., and Zhulin, I. B. (1999). PAS domains: Internal sensors of oxygen, redox potential, and light. Microbiology and Molecular Biology Reviews *63*, 479-+.

Thresher, R. J., Vitaterna, M. H., Miyamoto, Y., Kazantsev, A., Hsu, D. S., Petit, C., Selby, C. P., Dawut, L., Smithies, O., Takahashi, J. S., and Sancar, A. (1998). Role of

mouse cryptochrome blue-light photoreceptor in circadian photoresponses. Science *282*, 1490-1494.

Toh, K. L., Jones, C. R., He, Y., Eide, E. J., Hinz, W. A., Virshup, D. M., Ptacek, L. J., and Fu, Y. H. (2001). An hPer2 phosphorylation site mutation in familiar advanced sleep phase syndrome. Science *291*, 1040-1043.

Vielhaber, E., Eide, E., Rivers, A., Gao, Z. H., and Virshup, D. M. (2000). Nuclear entry of the circadian regulator mPER1 is controlled by mammalian casein kinase I epsilon. Molecular and Cellular Biology *20*, 4888-4899.

Vitaterna, M. H., Selby, C. P., Todo, T., Niwa, H., Thompson, C., Fruechte, E. M., Hitomi, K., Thresher, R. J., Ishikawa, T., Miyazaki, U., Takahashi, J. S., and Sancar, A. (1999). 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*, 12114-12119.

Vosshall, L. B., Price, J. L., Sehgal, A., Saez, L., and Young, M. W. (1994). Block in Nuclear-Localization of Period Protein By a 2nd Clock Mutation, Timeless. Science *263*, 1606-1609.

Vosshall, L. B., and Young, M. W. (1995). Circadian-Rhythms in Drosophila Can Be Driven By Period Expression in a Restricted Group of Central Brain-Cells. Neuron *15*, 345-360.

Yagita, K., Yamaguchi, S., Tamanini, F., van der Horst, G. T. J., Hoeijmakers, J. H. J., Yasui, A., Loros, J. J., Dunlap, J. C., and Okamura, H. (2000). Dimerization and nuclear entry of mPER proteins in mammalian cells. Genes & Development *14*, 1353-1363.

Zeng, H. K., Hardin, P. E., and Rosbash, M. (1994). Constitutive Overexpression of the Drosophila Period Protein Inhibits Period Messenger-Rna Cycling. Embo Journal *13*, 3590-3598.

Zeng, H. K., Qian, Z. W., Myers, M. P., and Rosbash, M. (1996). A light-entrainment mechanism for the Drosophila circadian clock. Nature *380*, 129-135.

Zerr, D. M., Hall, J. C., Rosbash, M., and Siwicki, K. K. (1990). Circadian Fluctuations of Period Protein Immunoreactivity in the Cns and the Visual-System of Drosophila. Journal of Neuroscience *10*, 2749-2762.

Zylka, M. J., Shearman, L. P., Weaver, D. R., and Reppert, S. M. (1998). Three period homologs in mammals: Differential light responses in the suprachiasmatic circadian clock and oscillating transcripts outside of brain. Neuron *20*, 1103-1110.

CBS Phosphobase v 2.0. In A database of phosphorylation sites.

Qiagen RNA club news letter No 6, www.qiagen.com