Expression, function and regulation of the *Him* gene during *Drosophila* heart development

Karen Wessel

PhD Thesis

School of Biosciences
Cardiff University

2013

Acknowledgements

My first and biggest thank you goes to Mike Taylor, for his trust and patience in allowing me to do this work in his lab. I greatly appreciated all the help and advice he has given me through the years. Without him this thesis would not have been possible.

I am also grateful to the British Heart Foundation for funding me and for supporting *Drosophila* heart research.

I also wish to thank my parents for their generosity, love and unwavering support through this time. They are truly wonderful and I am very lucky to have them.

Major thanks also go to current and previous lab members for help, advice and all the attempts to keep me smiling. Thank you Jun, Stuart, David, Dan and Alina. And all the other people here in the office: Simona, Jodie, Stuart, Clare and Leanne, thank you. It would have been very boring without them.

I would also like to acknowledge and thank Achim Paululat, Maik Drechsler and Stefanie Albrecht for their invaluable help with the FISH protocol. I learned a lot in those two weeks.

And another thank you goes to my friends who have kept me laughing and distracted in the past years, whether I wanted to or not. You have always had time for me and I really appreciate that. So, thank you Tanja, Jeni, Alice, Mary, Birgit, Kerstin, and Doro.

Summary

I have analysed the regulation and function of the *Him* gene to gain new insights into *Drosophila* heart development and its controlling factors. My results show that *Him* is important during the early specification of pericardial cells and cardioblasts. Loss of *Him* leads to a reduced number in both of these cell types by the end of embryogenesis. Over-expression of *Him* throughout the heart results in supernumerary pericardial cells.

Him is expressed in all embryonic pericardial cells from embryonic stage 12 to approximately stage 15. I have identified an enhancer fragment that reproduces this expression pattern. Phylogenetic footprinting revealed three highly conserved regions within this sequence. I undertook an extensive mutational analysis of this enhancer to identify regulatory elements within it. I identified Tinman as a direct activator of *Him* expression. My data indicate that *Him* is activated in a widespread area of the dorsal mesoderm and the amnioserosa and is actively limited to the pericardial cells. A 5 bp mutation within the enhancer sequence allows for expression within the cardioblasts.

Both heart cell types develop from the dorsal mesoderm and some share immediate progenitors. By stage 13, *Him* is pericardial cell specific and *Mef2* is cardioblast specific. This is essential for normal heart development. If *Him* is not excluded from the cardioblasts, expression of the muscle-cell specific differentiation gene *myosin* is disrupted, similar to what has been described for *Mef2* null mutants. If *Mef2* is expressed in pericardial cells, the larval development of the pericardial cells is severely disturbed.

A possible explanation for these data is that *Him* is part of a genetic program that prevents the premature differentiation of heart cells and its down-regulation permits the pericardial cells to undergo their correct development.

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, not is being submitted concurrently in candidature for any degree or other award.

Signed _____ Date _____

Karen Wessel

Statement 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed _____ Date _____

Karen Wessel

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed _____ Date _____

Karen Wessel

Statement 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for interlibrary loan, and for the title and summary to be made available to outside organisations.

Signed

Date

_____ _____

Karen Wessel

Table of Contents

Chapter 3 – *Him* expression and phenotype

List of Figures

List of Tables

List of DVD contents

Electronic version of thesis

- Table of contents
- Chapter 1 Introduction
- Chapter 2 Materials and Methods
- Chapter 3 *Him* expression and phenotype
- Chapter 4 The Pericardial Cell Enhancer (PCE)
- Chapter 5 Tinman regulates *Him* expression
- Chapter 6 Further regulatory input in *Him* expression
- Appendix
- Bibliography

Moveis Chapter 3 *Mef2* over-expression phenotype

- Folder:
- control handCGFP2.3 3rd instar larvae
  - contains two movies of one control larvae at different magnifications
- HimGal4 x UAS-Mef2high 3rd instar larvae
  - contains six movies of three different larvae at different mangnifications

1. Introduction

The fruit fly, *Drosophila melanogaster,* has become a widely studied model organism in many scientific areas. *Drosophila* offers many advantages such as small size, easy cultivation, short generation time and this model system is suitable to both well-developed classic and modern genetic approaches. In *Drosophila*, a functional heart is only necessary after completed embryogenesis.

The early patterning mechanisms that establish the basic body plan and its further subdivision are well studied. It is now known that many of these steps and processes are well conserved throughout the animal kingdom. This amazing conservation of patterning and the corresponding transcription networks has also been shown to be true especially for early heart development (Bodmer, 1995; Harvey, 1996; Olson and Srivastava, 1996, Zaffran and Frasch, 2002).

In *Drosophila*, a functional heart is only necessary after completed embryogenesis. The *Drosophila* heart derives, as all hearts do, from the mesodermal germ layer. After the specification of the mesoderm this germ layer is subdivided into the fat-body and the somatic, visceral and heart mesoderm, each of which then differentiates further (Azpiazu and Frasch, 1993; Riechman et al., 1997).

1.1 *Drosophila* development

1.1.1 Life cycle

The *Drosophila* life cycle takes about 11 days (at 25°C). The female flies deposit a fertilized egg on the surface of the food. This egg develops externally and is thus easily accessible for experiments. The embryonic development is completed within 24 hours and the first instar larva hatches by this time. The first larval instar lasts 24 hours. It takes another 48 hours (i.e. 96 hours AEL = after egg laying) for the third instar larva to develop from the second larval instar larva. This larva then continues to eat and grow to eventually pupate to begin metamorphosis into the adult fly (ca. 120 hours AEL). During pupation most of the larval tissues are histolysed and the adult fly develops. It

ecloses approximately ten to eleven days after egg laying. One of the few tissues that is not completely histolysed during metamorphosis is the heart.

1.1.2 Embryonic development

The complete process of embryonic development has been divided into 17 distinct stages (Campos-Ortega and Hartenstein, 1985) to aid in its study. After the first nuclear divisions the embryo cellularizes (between two and three hours AEL, embryonic stage 5) which is followed by gastrulation. During gastrulation the three germlayers (endoderm, ectoderm and mesoderm) are formed. The endoderm will develop into the fore- and hindgut. The ectoderm will develop into the nervous system and epidermis, which secretes the cuticle of the developing fly. The mesoderm differentiates into the visceral muscle, muscles and heart and as such is of particular importance to this thesis, and will be explored further.

1.1.3 Formation of the mesodermal germlayer and its subdivision

The mesodermal germlayer is established on the ventral side of the *Drosophila* embryo. The mesodermal cells are internalized through the process of gastrulation. The maternally deposited factor Dorsal becomes active in the nuclei of the ventral most cells (Chasen and Anderson, 1993) and thus activates the zygotic genes *twist* (*twi*) and *snail* (*sna*) (Ip et al., 1992; Jiang et al., 1991; Pan et al., 1991; Thisse et al., 1988; Boulay et al., 1987). Both Twist and Snail are transcription factors necessary for the development of the mesoderm. Twist activates later mesodermal genes and Snail represses non-mesodermal genes in the developing germlayer. The Twist protein is a helix-loop-helix protein and is necessary to activate the homeobox gene *tinman* (*tin*) and the MADS domain transcription factor *Mef2* (Myocyte enhancer factor 2) (Taylor et al., 1995; Yin et al., 1997). The early mesoderm is characterised by the uniform expression of Twist, Tinman and Mef2.

During gastrulation, the mesodermal cells divide twice and migrate along the inner surface of the ectoderm towards the dorsal side of the embryo subsequently forming a monolayer of cells underneath the ectoderm (this corresponds to embryonic

stage 8; Leptin and Grunewald, 1990; Dunin Borkowski et al., 1995). The outer ectodermal layer now influences the subdivision of the mesoderm through signals such as Decapentaplegic (Dpp) along the dorso-ventral axis and Wingless (Wg) and Hedgehog (Hh) along the anterior-posterior axis (during embryonic stage 9; Staehling-Hampton et al., 1994; Dunin Borkowski et al, 1995; Frasch, 1995; Baylies et al., 1998).

The mesoderm is also subdivided along the anterior-posterior axis by the expression domains of the pair-rule genes *even-skipped* (*eve*) and *sloppy-paired* (*slp*). While mesodermal areas under the influence of Even-skipped will form the visceral musculature and the fat body, the areas under the influence of Sloppy-paired will form the somatic musculature and the heart (Riechmann et al., 1997).

The developing dorsal mesoderm in its early stages is characterised through uniform *tinman* expression, which becomes restricted to this part of the mesoderm at this stage, and will give rise to the visceral musculature and the heart. *tinman* mutants do not form either heart or visceral mesoderm (Bodmer, 1993). The expression of Sloppy-paired and Wingless (in the ectoderm) induces the differentiation of the heart mesoderm in the mesodermal cells that are in close contact with the ectoderm (Azpiazu and Frasch, 1993; Riechmann et al., 1997). In the cells further away from the ectoderm, Tinman activates *bagpipe* (*bap*) and thus promotes the formation of the visceral mesoderm (Azpiazu and Frasch, 1993). The generation of the visceral and the heart domain within the mesoderm are mutually exclusive since the expression of the Sloppy-paired protein represses *bagpipe* expression.

During these specification processes, the mesodermal cells divide a third time and form two layers of cells. As stated above, the inner layer gives rise to the fat body and the visceral muscles, while the outer mesodermal layer remains under the control of high levels of Twist and also remains in physical contact with the overlaying ectoderm. This outer layer will form the somatic muscles and the heart (Dunin Borkowski et al., 1995; Riechmann et al., 1997).

1.2 The dorsal vessel

By late embryonic stage 12, the dorsal mesoderm extends as a continuous band along the sides of the embryo. Following germband retraction and dorsal closure of the

embryo these cells migrate towards the dorsal midline of the embryo and will form the dorsal vessel or heart along the dorsal midline shortly after dorsal closure (Frasch, 1999). During the process of dorsal migration and closure the presumptive heart cells are patterned further under the control of a network of various genes (Cripps and Olson, 2002).

As a result of the inwards movement of the progenitors of the fat body and the visceral muscles, the mesoderm forms a "crest and trough" pattern along the anterior-posterior axis by early embryonic stage 10; the peak of each crest forms under the part of the ectodermal stripe with highest levels of Wingless (Dunin Borkowski et al., 1995). The heart progenitors start to be specified at this stage as well. During embryonic stage 10 a transient cluster of cells expressing Lethal of scute (L'sc) appears in the peak area of the dorsal mesoderm that was exposed to the highest levels of Decapentaplegic (Dpp) expression. The process of lateral inhibition singles out precursors for the heart in this cluster of cells and the first specific heart progenitors that are detectable are one or two Eve-positive cells within this peak area in late stage 10/early stage 11 embryos (Dunin Borkowski et al., 1995; Park et al., 1998; Carmena et al., 1998). The crest cells that contain the heart precursors begin to spread out to the sides during embryonic stage 11 and by the end of this stage the division of the progenitor into the two Eve-positive pericardial cells has occurred (Park et al., 1998; Carmena et al., 1998). Two different models for the generation of the Eve-positive pericardial cells have been proposed; they disagree in the exact lineage of the Eve-positive pericardial cells (Park et al., 1998; Carmena et al., 1998). The analysis of Park et al. (1998), based on the analysis of induced clones and the expression pattern of marker genes, argues that two precursors asymmetrically divide to each yield one Eve-positive pericardial cell and a cell that contributes to the DA1 muscle of the somatic mesoderm. Contrary to this, the study of Carmena et al. (1998) used immuno-histochemical and over-expression techniques together with rescue experiments to argue that the asymmetric division occurs one step earlier and that the immediate progenitor of the Eve-positive pericardial cells divides symmetrically to produce these two daughter cells. This latter model is the model depicted in Figure 1.2.1 F and is the more widely referenced model.

During early stage 12, together with the beginning of the germband retraction, the first precursors for the Odd-positive cells appear underneath the ectodermal Odd-skipped stripe (Ward and Skeath, 2000). The four Odd-positive pericardial cells of the mature embryonic heart are generated from three different precursors. Two of these

precursors divide asymmetrically during the early part of stage 12 and generate one Odd- and Svp-positive pericardial cell and one Svp-positive cardioblast each; the third precursor appears slightly later during late stage 12/early stage 13 and divides symmetrically into two Odd-positive pericardial cells (Ward and Skeath, 2000). By stage 14, four Odd-positive pericardial cells are arranged in a row along the anterior-posterior axis of each hemisegment (see Figure 1.2.1 B for a schematic of the Odd-positive pericardial cell lineage).

While the lineage of the remaining heart cells is also known, the timing of their division is less well studied. It is, however, generally accepted that these cells are also generated during late stage 11/early stage 12 (Ward and Skeath, 2000). In their study from 2000, Ward and Skeath used immuno-histochemical techniques in combination with clonal analysis and assays in *numb* and *sanpodo* mutants to show that the four Tinman-positive cardioblasts share one common superprogenitor (TSP in Figure 1.2.1 C), which divides into two progenitors, one of which gains *Ladybird early* expression and divides into the two Tinman- and Ladybird-expressing cardioblasts, while the other progenitor generates the two Tinman-expressing cardioblasts (Ward and Skeath, 2000). The remaining two cardioblasts of each hemisegment are generated from a Seven-up-superprogenitor (SSP in Figure 1.2.1 B). This superprogenitor divides into two equivalent progenitor cells, which then continue to divide asymmetrically each into a Seven-up-positive cardioblast and pericardial cell (see above; Ward and Skeath, 2000). The Seven-up pericardial cells are also two of the four Odd-positive pericardial cells.

The two Tinman- and Ladybird-positive pericardial cells and the two Tinman-positive pericardial cells are all generated through the symmetric division of a precursor cells (Ward and Skeath, 2000); these make up the four Tinman-positive pericardial cells. It is currently unknown if these two progenitors for the four Tinman-positive pericardial cells share a superprogenitor as do the Tinman-positive cardioblasts.

1.2.1 The embryonic heart

The segmentally organised *Drosophila* heart stretches from the third thoracic segment to the eighth abdominal segment. It pumps haemolymph towards the anterior of the animal. As is typical of insects, all other parts of the circulatory system are open;

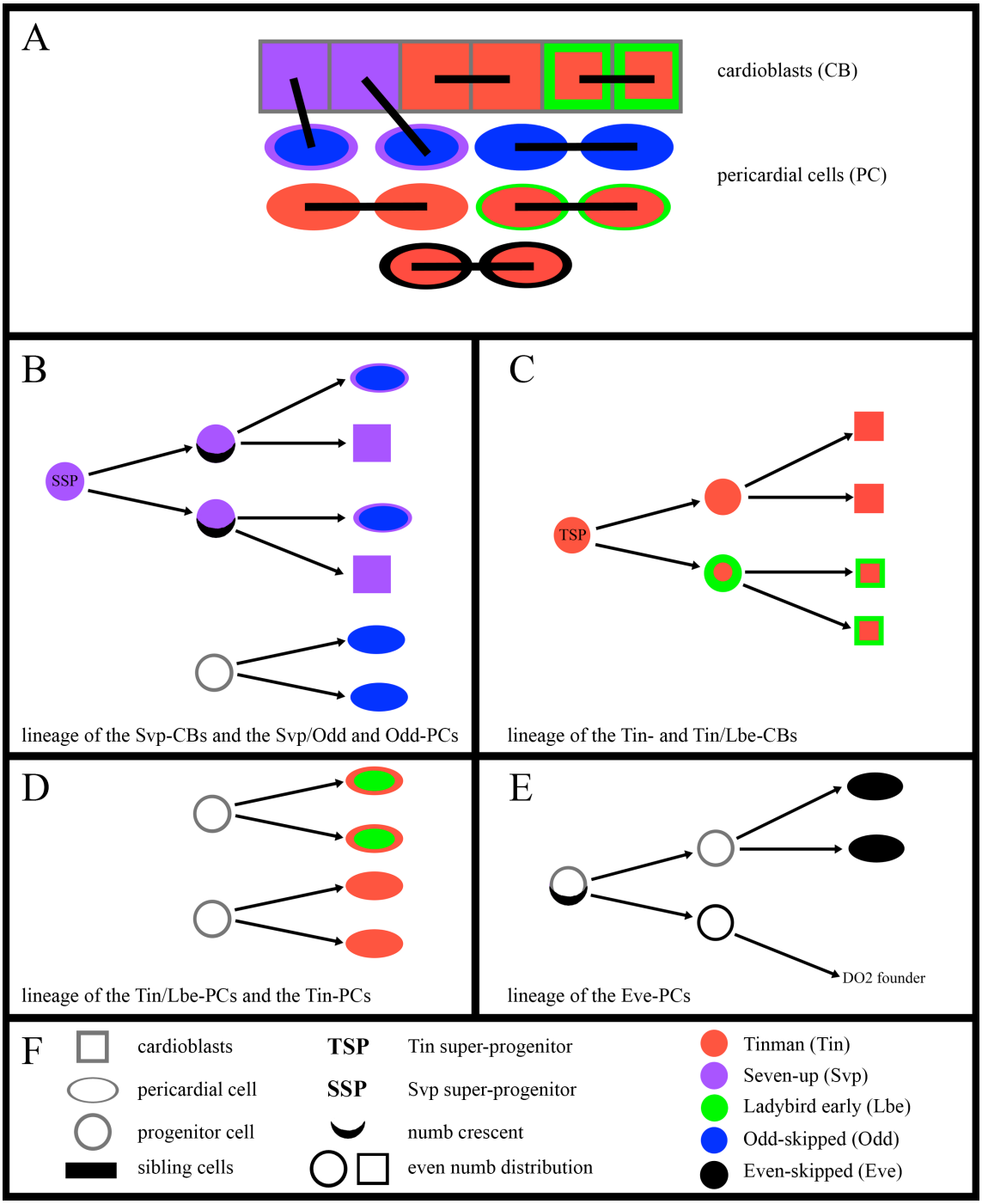


Figure 1.2.1: The cells of the *Drosophila* heart and their lineage (adapted after Ward & Skeath, 2000; Alvarez et al., 2003; Han & Bodmer, 2003).
A Schematic representation of the organisation of a hemisegment of the embryonic heart. **B** lineage of the Svp-positive cardioblasts and the Svp/Odd-positive and Odd-positive pericardial cells. **C** lineage of the Tin- and Tin/Lbe-positive cardioblasts. **D** lineage of the Tin- and Tin/Lbe-positive pericardial cells. **E** lineage of the Eve-positive pericardial cells. **F** key to colours, symbols and abbreviations used throughout the figure.

the haemolymph re-enters the dorsal vessel through openings, termed ostia, in the posterior part of the heart.

The embryonic *Drosophila* heart is subdivided into two morphologically different parts along the anterior-posterior axis of the animal: the heart proper in the posterior four segments and the aorta in the anterior segments. A pair of cardiovascular valves separates these two regions. The aorta appears thinner than the heart proper and also has a smaller lumen. Situated around the anterior-most region of the aorta are the lymph-gland and the endocrine ring-gland. The lymph-gland is responsible for the generation of the haemocytes that start circulating through the animal with the start of metamorphosis (Holz et al., 2003; for a review of the hematopoietic function of the lymph gland see Evans et al., 2003). The ring gland is part of the endocrine system of the fruit fly and is responsible for the production of ecdysone (Harvie et al., 1989; DeValesco et al., 2004).

The dorsal vessel consists of a closed, contractile tube of myocardial cells or cardioblasts, and an outer layer of loosely associated pericardial cells (Rizki, 1978). The pericardial cells are non-contractile and have a function in the filtration of the haemolymph (Rizki, 1987; Das et al. 2008; Weavers et al., 2009, Zhuang et al., 2009). The pericardial cells are one of two types of nephrocytes found in *Drosophila* and together with the Malphighian tubules make up the excretory system of the fruit fly. The second types of nephrocytes found in *Drosophila* are the garland cells, which are located around the oesophagus of the animal. Both types of nephrocytes take up waste particles smaller than 10-12nm through filtration and endocytosis (Crossley, 1972; Weavers et al., 2009) and are able to metabolize, store and exocytose these waste materials (for a recent review see: Denholm and Skaer, 2009).

As previously described the heart is organised in a segmentally repeating pattern. Exceptions to this are the anterior-most and the posterior-most segments, T3 and A8, respectively. In these two segments cells express different genes and are also present in different numbers. The lymph gland lies posterior to the ring gland and encompasses the segments T3 and A1. The remaining heart containing segments (A2-A7) all contain the same number of cells; each of these segments contains twelve cardioblasts and twenty pericardial cells (Ward and Skeath, 2000; Han et al., 2002). Unless otherwise stated, I will further describe the arrangement of the embryonic heart per hemisegment, i.e. a segment split in half along midline and the anterior-posterior

axis. Each hemisegment has an identical arrangement of sub-groups of heart cells. These sub-groups of heart cells are defined by the different expression of known transcription factors and marker genes (see Figure 1.2.1 for reference).

The cardioblasts can be distinguished from the pericardial cells by the expression of *Mef2* (Taylor et al., 1995; Yin et al., 1997; Bour et al., 1995; Lilly et al., 1995). Four of the six cardioblasts of each hemisegment express *tinman* (Bodmer, 1993; Azpiazu and Frasch,1993). The remaining two cardioblasts do not express *tinman,* but instead express *seven-up* (*svp*; Bodmer and Frasch, 1999) and *dorsocross* (*doc*; Lo and Frasch, 2001). Two of the four Tinman-positive cardioblasts express the *ladybird* genes, *ladybird-early (lbe)* and *ladybird-late (lbl)* (Jagla et al., 1997). These four cardioblasts also express the T-box genes *midline (mid)* and *H15* (Miskolczi-McCallum et al., 2005; Qian et al., 2005, Reim et al., 2005).

The pericardial cells are characterised by the uniform expression of *zfh-1* (*zinc finger homeodomain 1*) and *pericardin* (*prc*; Lai et al., 1991; Zaffran et al., 1995; Chartier et al., 2002). As with the cardioblasts there are also sub-groups within the pericardial cells. *tinman* is also expressed in the pericardial cells; the four pericardial cells ventral to the cardioblasts are called Tinman-positive pericardial cells (see Figure 1.2.1 A). The posterior two of these four Tinman-expressing cells also express *ladybird.* The group of four pericardial cells positioned lateral and slightly ventral to the cardioblasts are named Odd-pericardial cells after their expression of *odd-skipped* (*odd*; Ward and Coulter, 2000; Ward and Skeath, 2000). The anterior two of these four cells express *seven-up* as well as *odd-skipped* and are termed Odd/svp-lacZ pericardial cells. In the late embryonic heart the Eve-pericardial cells, which also express *tinman*, are located dorsally to the rest of the heart (Frasch and Levine, 1987).

The early heart mesoderm is characterised by the uniform expression of *tinman*. Within the cluster of *tinman* expressing cells there are two mutually exclusive clusters of cells expressing *ladybird* and *even-skipped*. These two genes repress each other and form two distinct groups of precursors (Han et al., 2002). The T-box gene cluster *dorsocross1, dorsocross2* and *dorsocross3* (all three are usually grouped together and referred by one single abbreviation: *doc*) is also expressed in these early heart progenitors (Reim and Frasch, 2005). The *doc* genes are activated by Dpp and Wingless and are independent of Tinman in the early stages of heart development. The *doc* genes are necessary for the correct specification of the heart cells.

The GATA-factor Pannier starts to be expressed around stage 11 and its expression is dependent on both Tinman and Doc, which are, in turn, also positively regulated by Pannier (Gajewski et al., 2001; Reim and Frasch, 2005). Doc expression is initiated in all heart precursors, but is then limited to the two anterior-most cardioblasts per hemisegment expressing *seven-up* during stage 12 to 13. Pannier and Doc activate two further T-box genes, *midline/neuromancer2* (*mid* or *nmr2*) and *H15/neuromancer1* (*H15* or *nmr1*). These two genes are expressed in the cardioblasts only; their expression has not been detected in the pericardial cells. *midline* expression starts in early stage 12, while the expression of *H15* starts slightly later during mid stage 12. Midline has been shown to be necessary for *tinman* expression from late stage 12 to stage 14 and to negatively influence *doc* expression in the cardioblasts. However, both of these interactions might be indirect (Miskolczi-McCallum et al., 2005; Qian et al., 2005, Reim et al., 2005).

Another direct Tinman target, which is expressed in all heart cells, is the bHLH factor Hand (Heart, Autonomic Nervous system, neural crest-Derived cell types). Both Tinman and Pannier are necessary to activate expression of *hand* during stage mid 12 (Han and Olsen, 2005). So far, the function of the Hand protein in the embryonic *Drosophila* heart development is unknown, but it has been shown to be necessary for the formation of the adult heart from the larval heart during metamorphosis (Lo et al., 2007) and to be a direct target of the transcription factor Biniou in the visceral mesoderm (Popichenko et al., 2007).

Further direct Tinman targets are the *ß3-tubulin* and *Mef2* genes (Gajewski et al., 1998; Gajewski et al., 1999; Kremser et al., 1999). *Mef2* is initially expressed in all of the mesoderm, but becomes restricted to the cardioblasts within the heart mesoderm (Gajewski et al., 1998; Gajewski et al., 1999). Expression of *ß3-tubulin* in the heart is restricted to the four Tinman expressing cardioblasts and not observed in the two anterior Seven-up-positive cardioblasts.

The gene *zinc finger homeodomain 1* (*zfh-1*) is initially expressed in the developing mesoderm and is then restricted to the pericardial cells with the *Drosophila* heart (Lai et al., 1991). Zfh-1 is the most commonly used marker for the pericardial cells. Towards the end of embryonic development, *zfh-1* expression becomes further restricted to a sub-set of pericardial cells as the Eve-positive pericardial cells lose *zfh-1* expression (Su et al., 1999; Johnson et al., 2003).

1.2.2 The larval heart

The development of the *Drosophila* heart continues during the larval life of the animal. By the end of embryogenesis the heart is capable of pumping the haemolymph through the body. During the larval development, the heart grows in size in parallel with the whole animal. Generally there is still little that is known about the expression of genes in the heart during larval development. This is in part due to the fact that the larva secretes a cuticle, which makes immuno-histochemical methods more difficult as antibodies are unable to diffuse through this layer.

The transcription factor *hand* remains expressed in all heart cells (cardioblast and pericardial cell) throughout larval life and into the adult fly (Sellin et al., 2006). *tinman* expression has also been described in the heart tube (Qian et al., 2008) and *Mef2* expression has been described for the cardioblasts of third instar larvae (Molina and Cripps, 2001). Sellin et al. (2006) have described the organisation of the larval heart using a hand-GFP reporter construct. Their study showed that the number of hand-GFP-positive pericardial cells is reduced dramatically from the first instar larvae to the second instar larvae. There is evidence that the majority of pericardial cells are lost from the anterior part of the larval dorsal vessel and that these cells most likely degenerate (Sellin et al., 2006; Das et al., 2008). However, Sellin et al. (2006) state that they could not obtain reproducible results with propidium iodide and acridine orange, two commonly used stains for apoptosis; leaving the fate of this set of pericardial cells unclear. It is currently also not known if this loss of pericardial cells affects a specific sub-set of the pericardial cells. Das et al. (2008) have shown that in third instar larvae, all larval pericardial cells express *odd-skipped* and *even-skipped* but not *tinman* or *seven-up* and that the pericardial cells have the same identity with regard to the markers they tested for. They also observe *pericardin* expression throughout larval development (Das et al., 2008). Serpent, a GATA factor, is also expressed in third instar larvae (Brodu et al., 1999). In a recent paper, Lehmacher et al. (2012) mention that the embryonic Odd-positive pericardial cells are maintained into larval life.

During the first larval instar the cardioblasts begin to grow and extend themselves along the anterior-posterior axis and the pericardial cells lose their close association with the cardiac tube (Sellin et al., 2006). Towards the end of the first instar the remaining pericardial cells also begin to grow in diameter and their nucleus enlarges

considerably as well. This nucleus enlargement is most likely due to the polytenization of the chromosomes (Kambysellis and Wheeler, 1972; Rizki, 1978).

The continuous arrangement of the pericardial cells along the heart tube as seen in the embryo is interrupted in the early second instar larvae. The remaining pericardial cells are all arranged in one plane, but they do not resemble "pearls-on-a-string" anymore; especially in the anterior region there are large gaps between the pairs of pericardial cells. So far, no connective structures between the pericardial cells have been described. Along the heart proper the spacing of the pericardial cells is, however, quite close.

1.2.3 The adult heart

The heart is one of the organs of the fly that is not completely histolysed and re-built during the metamorphosis of the larva to the adult fly. The adult heart remains at the dorsal midline of the animal. It does however undergo several morphologic changes. The transcription factor Hand appears to have a crucial role for the correct formation of the adult heart as its loss leads to a number of defects associated with the development of the adult heart, i.e. the myofibrillar structure is disorganized and the systolic and diastolic diameter is reduced leading to a premature death of the adult flies (Lo et al., 2007). The adult heart is shortened to what was the embryonic "aorta" when compared to the larval and embryonic heart, it now only extends from the brain in the dorsal anterior region as far as the segment A5 and also gains a curved structure following the general shape of the abdomen (Ritzki, 1978; Curtis et al., 1999; Molina and Cripps, 2001; Sellin et al., 2006). The cardiac cells of the posterior most segments undergo programmed cell death (Jensen, 1973; Molina and Cripps, 2001; Zeitouni et al., 2007).

This shorter adult heart is formed in response to the last peak of ecdysone expression during morphogenesis, which in turn controls the expression of the Hox genes *ultrabithorax (Ubx*) and *abdominalA (abdA*; Monier et al., 2005). The anterior segments A1 to A4 express *Ubx* and the posterior three segments A5 to A7 express *abdA*. The cardiac cells of the last segment (A5) "transdifferentiate" from regular myocardial cells in the middle of the larval heart into the specialized cells of the terminal chamber of the heart; this chamber is innervated but lacks automatic contractile

activity (Zeitouni et al., 2007). This change is dependent on the expression of *abdA* and the inhibition of the Wnt signalling cascade (Zeitouni et al., 2007). The Wnt signalling cascade is also necessary for the *de novo* formation of the inflow tract in the abdominal segments A2 and A3 and the differentiation of the cardiac cells (Curtis et al., 1999; Sellin et al., 2006; Zeitouni et al., 2007). During the remodelling of the larval to adult heart no new cardiomyocytes are formed, instead the already present and functional cells differentiate and grow in size and in the number of myofibrils they contain (Monier et al., 2005; Zeitouni et al., 2007). The adult heart possesses five ostia/inflow tracts that develop from the Seven-up positive cardioblasts of the embryo (Curtis et al., 1999; Wasserthal, 2007; Shah et al., 2011). The PDGF-VEGF pathway is necessary for this valve formation and is, similarly to the already mentioned pathways, triggered by the ecdysone cascade during pupation (Zeitouni et al., 2007).

Recently it has been shown that two separate populations of pericardial cells contribute to tissues outside the heart proper. During pupation a ventral layer of longitudinal muscle develops beneath the adult heart (Curtis et al., 1999; Zeitouni et al., 2007; Shah et al., 2011). This layer of ventral longitudinal fibres is thought to originate from the secondary lymph gland cells, which are located posterior to the actual lymph glands of the animal (Shah et al., 2011). This muscle layer is syncytial and its formation is dependent on the activation of the FGF pathway (Zeitouni et al., 2007). The same study also shows the necessity of Notch signalling for the correct differentiation of the ventral heart muscle.

Tögel et al. (2008) described the *Drosophila* wing hearts (pulsatile organs under the scutellum of the adult flies) as organs necessary to remove the epithelial cells from the wings upon eclosion of the adult fly. They have identified the origins of these organs as the anterior-most Eve-positive pericardial cells (Tögel et al., 2008). These anterior-most Eve-positive pericardial cells are located anterior to the lymph gland and are a sub-set of the Eve-positive pericardial cells described earlier. However, contrary to the already described Eve-positive pericardial cells these cells exclude *tinman* expression by embryonic stage 13 (Tögel et al., 2008). The origin of the wing hearts adds a further example that pericardial cells can indeed contribute to myogenic tissue in the adult fly.

1.3 Bioinformatic analysis of non-coding sequences

One of the big advantages of *Drosophila* as a model system is the immense effort that has been put into sequencing other *Drosophila* species. The genome of twelve *Drosophila* species has been sequenced, assembled and has been made publicly available (*Drosophila* 12 Genomes Consortium, 2007). This project has been unique in its scope and in the number of different species of one genus that have been sequenced. With the help of this project it is possible to analyse the degree of conservation of non-coding and coding sequences between the different species. The lessons learnt from these analyses will also facilitate bioinformatic approaches for other species enabling predictions of protein function, potential regulators and transcription factor binding sites to become more accurate.

The data of the sequenced genomes for *D. simulans, D. sechellia, D. yakuba, D. erecta, D. ananassae, D. persimilis, D. willisoni, D. virilis, D. mojavensis* and *D. grimshawi* are publicly available (*Drosophila* 12 genome consortium, 2007). Together with the previously released sequence data for *D. melanogaster* and *D. pseudoobscura* these species cover different evolutionary distances from *D. melanogaster*. Their relation to each other is shown in Figure 1.3.1 (Russo et al., 1995; Powell, 1997). *D. virilis*, *D. mojavensis* and *D. grimshawi* are the species that are farthest removed from *D. melanogaster*; the evolutionary distance between *D. melanogaster* and *D. virilis* is equivalent to the distance between mice and humans (Hartl and Lozovskaya, 1994). Approximately 22 to 26 % of non-coding sequence between these species is conserved (Bergman et al., 2001). All other *Drosophila* species, for which sequence data are available, are within the *Sophophora* sub-group of the *Drosophila* genus and are thus more closely related. The two species closest related to *D. melanogaster* are *D. simulans* and *D. sechellia*. *D. simulans* was separated approximately ten millions years ago from *D. melanogaster*, while *D. yakuba* was separated from *D. melanogaster* about eight million years ago. The *Melanogaster* subgroup of the *Sophophora* also includes *D. erecta* and the complete *Melanogaster* group also includes *D. ananassae*. This group diverged around ten million years ago and split from the *Obscura* group about 22 million years ago. Two sequenced members of the *Obscura* group, *D. pseudoobscura* and *D. persimilis*, have been included in the sequencing project to allow for the study

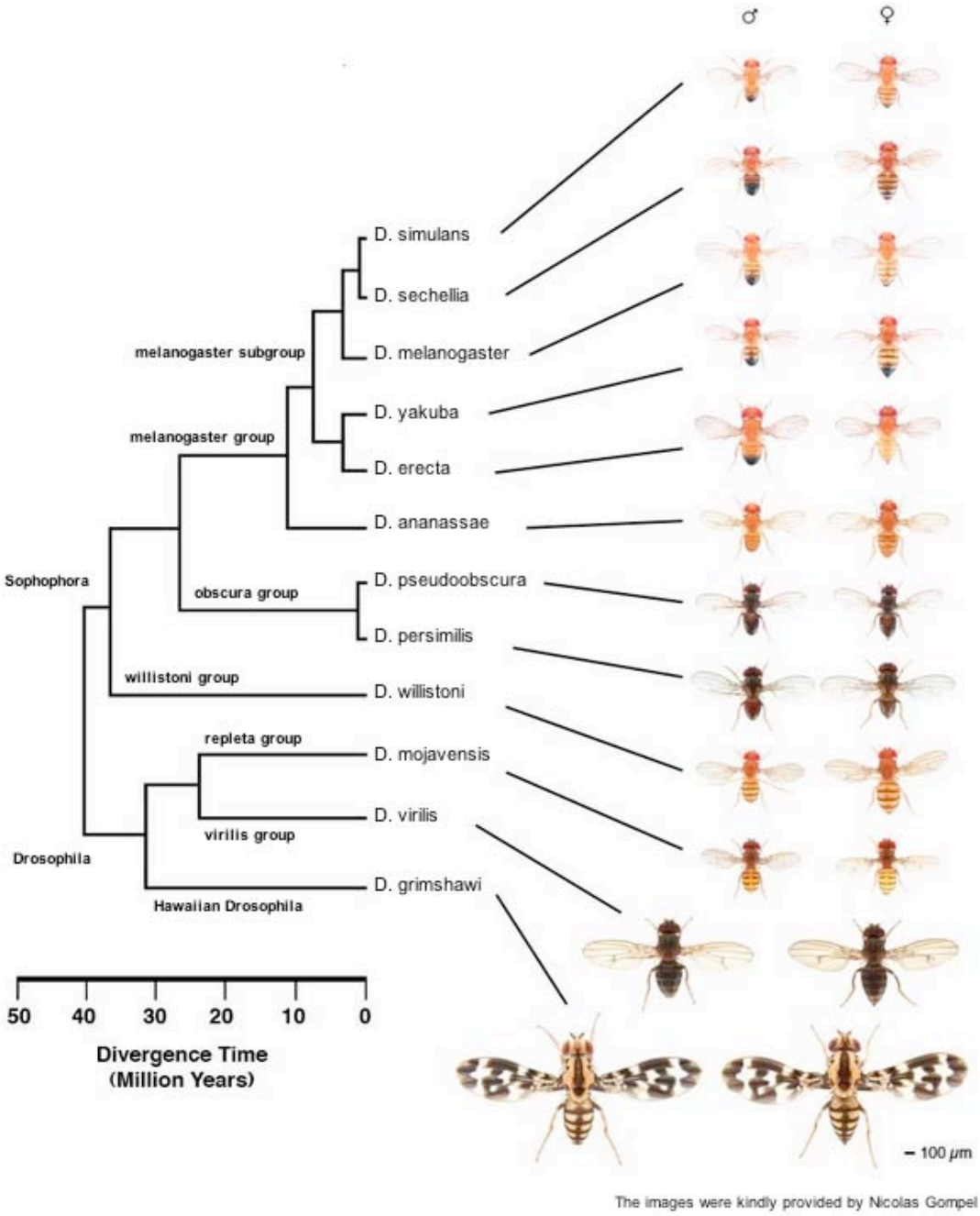


Figure 1.3.1: Phylogenetic relationships between the twelve *Drosophila* species used in this study (image taken from Flybase (http://flybase.org/static_pages/species/sequenced_species.html)).
This diagram only shows the relevant branching points of the different species and is simplified (this is especially true for the *virilis-repleta* radiation).

of an evolutionary outgroup to the *Melanogaster* subgroup. There are differences in the estimates for the evolutionary distance between *D. pseudoobscura* and *D. melanogaster.* They range from 40 to 50 millions years (Russo et al., 1995; Powell, 1997; Tamura et al., 2004; *Drosophila* 12 Genomes Consortium et al., 2007; Obbard et al., 2012).

For my project the sequence data have helped me to define the regions that are most likely to be involved in the transcriptional regulation of *Him*. This so-called "Phylogenetic footprinting" of the non-coding sequence, i.e. the alignment of sequence data for the same region of different species, reveals those regions that have been conserved through the course of time. As for coding sequence, the degree of conservation of non-coding sequence is seen as an indicator of essential functions e.g. transcription factor binding sites. It has also been shown that the *cis*-regulatory sequences of *Drosophila* are highly constraint (Bird et al., 2006; Wittkopp, 2006; *Drosophila* 12 Genome Consortium, 2007; Ozdemir et al., 2011). This approach will facilitate the search for transcriptional regulators of *Him*.

Based on the availability of the *Drosophila* genome sequence and on the development of new high-throughput techniques established in the ENCODE (Encyclopedia of DNA Elements) project, the modENCODE (model organism ENCODE) project for *Drosophila melanogaster* and *Caenorhabditis elegans* (worm) was established (Celniker et al., 2009; The modENCODE Consortium et al., 2010). The aim of the modENCODE project is to generate and link the various different but comparable data sets about gene expression and regulation, known target genes and chromatin structure etc. Thus, the modENCODE project will facilitate and help integrative studies of the function of particular genes in relation to its position in the genome and the role it has during the animals' life cycle. The modENCODE project has also vastly improved the annotation of the genomes. One area of interest for the project has been to build and establish tools to reliably predict transcription factor binding sites and their binding factors (Flynet; Tian et al., 2009). This in turn, has allowed for the development of the first *cis*-regulatory maps of *Drosophila*, which has identified many candidate regulatory elements in combination with chromatin activity and accessibility and new transcription factor co-binding relationships. The transcriptional network described in this map currently includes about 800 different regulatory relationships (Nègre et al., 2011). Ozdemir et al., 2011 show some of the possibilities and power this approach has by analysing the binding characteristics, function and evolutionary development of the transcription factor Twist at various stages during the embryonic

development of *Drosophila*. This genome-wide, computational approach generates a lot of data that looks promising. It is however, essential to prove and connect these data to its functional relevance as transcription factor binding sites can be assigned in error as pointed out by Halfon et al., 2011.

1.4 Conservation of heart development between vertebrates and *Drosophila*

Even though the adult *Drosophila* and vertebrate heart appear very different, the early development of both heart types has been shown to be conserved and comparable (Bodmer, 1995; Harvey, 1996; Olson and Srivastava, 1996, Zaffran and Frasch, 2002). In both vertebrates and *Drosophila*, the presumptive heart cells are generated from the area of the mesoderm that moved furthest from the invagination point during gastrulation. These cells form two symmetric cell bands, which then move towards each other in order to form a linear heart "tube". The embryonic *Drosophila* heart remains in this tube-like shape, while the vertebrate heart undergoes many morphological and structural changes, i.e. the looping and formation of compartments within the heart (Christoffels et al., 2000; Harvey, 1998). According to Hartenstein and Mandal (2006), all parts of the *Drosophila* cardiovascular system (contractile cardioblasts, the pericardial cells (nephrocytes) and the lymph gland that produces the haemolymph) are equally related to all the cell types of the vertebrate vascular system. Both cardiovascular systems are of mesothelial origin and contain cells and tissues that are responsible for the formation of blood cells, the excretory system, the circulatory system and the contractile heart (Mandal et al., 2004; Hartenstein and Mandal, 2006; Medioni et al., 2009). Evidence for this can also be found on a genetic level as the *Drosophila rudhira (rudh)* gene is exclusively expressed in the pericardial and garland cells from the third larval instar onwards and into adult life. The Rudhira protein belongs to the WD40 class and regulates macropinocytic uptake in the post-embryonic pericardial cells (Das et al., 2008b). Its ortholog in mouse, Rhudhira (Rudh), is expressed in endothelial precursors during vasculogenesis and angiogenesis (Siva and Inamdar, 2006; Das et al., 2008).

As indicated by this example, not only the morphological and structural steps of the early heart development are conserved between the *Drosophila* and vertebrate heart, but also many transcription factors have been conserved in their role and function (Bodmer and Venkatesh, 1998; Zaffran and Frasch, 2002). For example, the vertebrate orthologs for the *Drosophila tinman* gene (Nkx2.5 in mouse) are also the earliest known genes expressed in the presumptive heart cells (Lints et al., 1993; Harvey, 1996; Tanaka et al., 1998). Both *Drosophila* Tinman and murine Nkx2.5 have been shown to interact with GATA factors and regulate the transcription of downstream genes (Durocher et al, 1997; Gajewski et al., 1998; Gajewski et al, 1999; Bruneau et al., 2001; Gajewski et al, 2001). In a ChIP-on-chip study, Liu et al. (2009) showed that *Drosophila* Tinman activates *eyes absent*, a member of the Eya-Six family of transcription factors that are involved in muscle development in both flies and vertebrates, and that *Drosophila* Tinman regulates *stat92E*, a component of the JAK/STAT signalling pathway, which is also necessary for the somatic muscle development in *Drosophila* and myotome morphogenesis in zebrafish; showing that Tinman also influences the development of at least two further mesoderm derivatives (Liu et al., 2009). *Drosophila* Tinman is essential for heart development; *tinman* null mutants do not form a heart (Azpiazu and Frasch, 1993; Bodmer, 1993). In mice Nkx2.5 is not essential for the formation of the heart; the null mutant mice begin the process of heart formation but then the "heart" fails to continue to develop past the linear stage (Lyons et al., 1995; Harvey, 1996). In vertebrates the function of Nkx2-5 is not to initiate heart development but to ensure the activation of down-stream genes and the proper continuous development of the heart past the first few initial steps (Bartlett et al., 2010; Behrens et al., 2013; Guner-Ataman et al., 2013). In humans Nkx2-5 is implicated in congenital heart disease (for a recent review see Reamon-Buettner and Borlak, 2010).

The transcription factors of the *MEF-2 (myocyte enhancer factor 2) MADS*-Box family have important roles during the differentiation of the mesoderm in vertebrates and *Drosophila* (Bour et al., 1995; Lilly et al., 1995; Olson et al., 1995; Lin et al., 1997). Vertebrates have four members of the Mef2 family of transcription factors, Mef2-A, -B, -C and –D. The heart of mice lacking Mef2C fail to undergo heart looping and also lack the right ventricle chamber due to the down-regulation of the hand2 transcript (Lin et al., 1997). In both, vertebrates and *Drosophila,* Mef2 is known to be a central node for the regulation of gene expression in conjunction with various co-factors (for a review see: Potthoff and Olson, 2007). *Drosophila* has only one *Mef2* gene (Bour

et al., 1995; Lilly et al., 1995; Taylor et al., 1995) and its loss leads to a nearly complete loss of Myosin-expressing cells within the somatic mesoderm and a complete loss of *myosin* expression in the heart while the cardioblasts are still present (Bour et al., 1995; Lilly et al., 1995; Ranganayakulu et al., 1995; Gunthorpe et al., 1999; D. Liotta, PhD thesis).

Another example of conservation of the molecular mechanisms is the group of *GATA* transcription factors. The *Drosophila* GATA factor Pannier is required for the specification of both myocardial and pericardial cells (Gajewski et al., 2001; Alvarez et al., 2003; Klinedinst and Bodmer; 2003). The corresponding vertebrate factors *GATA-4, GATA-5* and *GATA-6* are also expressed in the developing heart and regulate the expression of cardiac specific genes (Molkentin, 2000; Latinkic et al., 2003; Sorrentino et al., 2005). GATA-4 is also necessary for the migration of the heart primordial towards the midline (Molkentin et al., 1997; Kuo et al., 1997). The murine GATA-4 protein is able to function as a cardiogenic factor in *Drosophila* (Gajewski et al., 1999).

T-box genes are also conserved between *Drosophila* and vertebrates. Similarly to *tinman*, these genes are also implicated in congenital heart disease. Tbx20 is expressed in the cardiac crescent in mouse development and its *Drosophila* orthologs, *H15/Neuromancer2/midline* are expressed in the cardioblasts (Ryan and Chin, 2003; Miskolczi-McCallum et al, 2005; Qian et al., 2005; Reim et al., 2005). The three *dorsocross* genes (*doc1, doc2* and *doc3*) are expressed in the Seven-up-positive cardioblasts. The *dorsocross* genes are most closely related to the vertebrate Tbx6 genes that are expressed in the paraxial mesoderm of vertebrate embryos (Reim et al., 2003; Reim and Frasch, 2005). Both *Drosophila* and vertebrate T-box genes are involved in the development, specification and differentiation of the heart (Hoogaars et al., 2007; Reim et al., 2003; Reim and Frasch, 2005; Reim et al., 2005).

This remarkable level of evolutionary conservation stretches as far as to the physiology of the heart and thus the *Drosophila* heart is becoming a well-established model system for the study of heart disease (for recent reviews see: Piazza and Wessells, 2011; Medioni et al., 2009; Cammarato et al., 2011; Nishimura et al., 2011; Maruyama and Andrew, 2012; Qian and Bodmer, 2012).

1.5 *Him (Holes in muscle)*

Him was discovered in a subtractive hybridisation screen, which was aimed at identifying genes specifically expressed in the progenitors of developing mesoderm (Taylor, 2000). Within the heart, *Him* is only expressed in the pericardial cells but prior to the work contained in this thesis, this had not been confirmed by evaluating the co-localisation of *Him* with known pericardial cell markers. This localisation would make it one of the few known genes that are expressed in all pericardial cells, but not expressed in the cardioblasts. The two other examples of proteins expressed in the pericardial cells but not in the cardioblasts include Zfh-1, which is expressed in all pericardial cells up to embryonic stage 15 and Pericardin. Zfh-1 expression is lost from the two Eve-positive pericardial cells per hemisegment by stage 15 (Su et al., 1999; Johnson et al., 2003) and Pericardin is secreted from the cells into the intercellular space and becomes a component of the extracellular matrix, making it unsuitable to detailed studies of cell numbers. Thus, Him and Zfh-1 are the only two known embryonic markers for the complete pericardial cell population.

Him is a novel inhibitor of myogenesis and interacts with Groucho (Liotta et al., 2007). In the somatic mesoderm of the embryo, Him is necessary to prevent premature activity of Mef2 and for the specification of the adult muscle precursors (Liotta et al., 2007). Our lab has further shown that Him plays a similar role in the development of the dorso-longitudinal muscles of the indirect flight musculature of the adult fly (Soler and Taylor, 2009).

Him is a relatively small gene located at 17A2 on the X chromosome of *Drosophila melanogaster*. It consists of only two exons and its cDNA has a length of 1324 nucleotides. The Him protein is located in the nucleus and contains a WRPW-domain in its C-terminus. This WRPW-domain is necessary for its interaction with Groucho (Liotta et al., 2007; D. Liotta, PhD thesis).

The expression pattern of *Him* is similar to that of *zfh-1* (*zn finger homeodomain-1*). The transcription factor *zfh-1* is expressed in all pericardial cells and has been shown to be necessary for the formation of the Eve-expressing pericardial cells (Lai et al., 1991; Lai et al., 1993; Su et al., 1999). Prior to my work, the function of Him during heart development was not known. It is also unknown if and from which developmental

stage onwards the expression patterns of *Him* and *Mef2* in the heart are separate or even mutually exclusive. We assume their expression patterns to be mutually exclusive from an early stage onwards and have shown that at least from embryonic stage 13 onwards this is indeed the case (see chapter 3). Thus, it is possible that Him has a similar function in the *Drosophila* heart to its function described in the embryonic somatic musculature and the indirect flight muscles of the adult fly. We speculate that during the later stages of the development of the embryonic somatic musculature, Him is necessary to keep the adult muscle progenitors (AMPs) in a specified but undifferentiated state by preventing premature Mef2 activity in these cells. The adult muscle progenitors are some of the few mesodermal cells that maintain *Him* expression until the end of embryogenesis. It is currently not known how the adult muscle progenitors and the heart cells react to a loss of *Him*. While the expression pattern of *Him* in the embryonic somatic musculature has been described (Liotta et al., 2007), its expression pattern in the embryonic heart has not been studied in detail and we know little about *Him* expression in the developing larvae, generating further questions that need answering.

The best-studied example of transcriptional regulation during heart development is the *even-skipped* enhancer and it has been shown that this enhancer is an important integrating point for various regulatory pathways activated in pericardial cells (Knirr and Frasch, 2001; Han, Z. et al., 2002; Jagla, 2002). The Even-skipped protein, however, is only expressed in two of the ten pericardial cells per hemisegment. This sub-population of pericardial cells appears to separate itself from the rest of the developing pericardial cells at an early stage during heart development. The Eve-positive pericardial cells are also the only pericardial cells that share a close lineage with dorsal muscle cells as described previously (see section 1.2.1). Very little is known about the regulation of any other genes in the pericardial cells. While nothing is known about the regulation of Odd in the pericardial cells, it has been shown that the expression of Tin in the pericardial cells requires a separate enhancer fragment than the expression in the cardioblasts (Venkatesh et al., 2000). Furthermore it is only known that Dpp represses *zfh-1* expression in a subset of pericardial cells (Johnson et al., 2003). Thus, the analysis of the transcriptional regulation of *Him*, a gene that is expressed in all pericardial cells, will greatly contribute to the understanding of cell-type specific gene expression within the complete set of pericardial cells and not just a sub-set like the Eve-positive pericardial cells. This in turn, will enable us to further

understand the basic underlying principles of the development of the main two cell types in the *Drosophila* heart and will possibly allow us to make predictions about similar processes in vertebrates.

1.6 Aims of the project

The first aims of my research were to establish and describe the expression pattern of the *Him* transcript in the development of the *Drosophila* heart and to determine whether Him was a valuable marker for pericardial cells. Subsequent aims were to analyse the cardiac phenotype caused by the loss or gain of *Him* and to analyse the transcriptional regulation of *Him.* Together this would be the first study of role of Him in the pericardial cells and would give us an insight into the underlying mechanisms that distinguish the myogenic and non-myogenic cardiac cell types and their close developmental relationship. Knowledge and understanding of these mechanisms is of importance for the development of treatments for congenital and acquired heart disease as changes in the regulation of heart development lead to congenital heart disease which affect nearly 1 % of live births and is considered the number one cause of neonatal death (Bruneau, 2008; Narlikar et al., 2010).

2. Materials and Methods

2.1 *Drosophila melanogaster*

2.1.1 Maintenance of Fly Stocks

The fly stocks were maintained in plastic tubes or bottles on standard culture medium (for recipe see Appendix) at 18 °C or 25 °C. The stocks were changed approximately every two weeks (if kept at 25 °C) or every four weeks (if kept at 18 °C). Tables 2.1.1, 2.1.2 and 2.1.3 list all fly stocks that were used or newly generated for this work.

2.1.2 Embryo collection and fixation for non-fluorescent immuno-histochemical analysis

The embryos were collected on apple juice agar plates (for recipe see Appendix) for the times and temperature calculated to produce embryos of the desired stage; most experiments were done on over-night embryo collections (25 °C) to allow for a range of developmental stages. The embryos were collected into a wire-mesh basket with a paintbrush and washed with water. In order to remove the chorion, the embryos were placed in a water-bleach (1:1) mix for approximately two minutes (Nationwide; bleach purchased from New Hall Janitorial). This process was monitored under a dissecting microscope. De-chorionated embryos become shiny and float towards the surface. The embryos were then rinsed with water, transferred to a 2 ml tube containing a 1:1 mix (1 ml each) of heptane and 4 % paraformaldehyde (in PBS) and fixed for 20 minutes on a Lukeham shaker (setting 6-7). The paraformaldehyde was removed and 1 ml of methanol added. Vigorous shaking of the embryos in this mixture for 30 seconds removed the vitelline membrane of the embryos. The embryos were then rinsed in 1 ml fresh 100 % methanol several times and stored in methanol at -20 °C until required.

Table 2.1.1: List of fly stocks used in this work. The Gal4 and UAS-lines used for this work are listed separately in tables 2.1.2 and 2.1.3, respectively.

Fly stock	**Origin**
w^{1118}; DF(2L)Exel7046/cyo wg lacZ	Bloomington stock centre (7819)
handCGFP2.3	Sellin et al., 2006
handCGFP3.1	Sellin et al., 2006
Him 195	Z. Han; R. Bodmer
Him 195;; zfh-1^2/TM3 ftz lacZ	D. Hancock; M. Taylor
Him 52	D. Hancock; M. Taylor
Him 52;; zfh-1^2/TM3 ftz lacZ	D. Hancock; M. Taylor
Him 74	Z. Han; R. Bodmer
Him 74;; zfh-1^2/TM3 ftz lacZ	D. Hancock; M. Taylor
$mef2^{22\text{-}21}$/cyo wg lacZ	Bour et al., 1995
p43 (Eco/Xho)	M. Taylor
pnr^{VX6}	Heitzler et al., 1996
Posakony 2.0kb	Rebeiz et al., 2002
Posakony 4.0kb	Rebeiz et al., 2002
tinABD-1B2; tin346/TM3 eve lacZ	Manfred Frasch; Ingolf Reim
Tin^{EC40}/TM3 AbdA lacZ	Bodmer, 1993
ush^2	Bloomington stock centre
ush^{rev18}	Cubadda et al., 1997
zfh-1^2/Tm3 ftz lacZ, Sb	Lai et al., 1993

Table 2.1.2: List of the various Gal4–driver lines used in this study.

Line	**Origin**	**Expression pattern**
(CA) hand-Gal4	A. Pauluat	Visceral mesoderm and all heart cells from embryonic stage 12 onwards
Da-Gal4	Wodartz et al., 1995	Ubiquitous
En-Gal4 (II)	Gaumer et al., 2000	Ectodermal engrailed stripes
Him-Gal4 L3-5	D. Liotta, M. Taylor	See appendix
Him-Gal4 L4-1	D. Hancock, M. Taylor	See appendix
Mef2-Gal4 (III)	Bour et al., 1995	Somatic mesoderm, later somatic muscles and cardioblasts
TinCΔ4-Gal4 (III)	Lo and Frasch, 2001	In the four svp-negative cardioblasts.
Him-Gal4 L3-5; handCGFP3.1	Generated for this study	Him-Gal4 L3-5 expression pattern
handCGFP2.3; (CA) hand-Gal4	Generated for this study	(CA) hand-Gal4 expression pattern
PCE 6; TinCΔ4-Gal4	Generated for this study	TinCΔ4-Gal4 expression pattern
PCE 6; Da-Gal4	Generated for this study	Da-Gal4 expression pattern
PCE 6; (CA) hand-Gal4	Generated for this study	(CA) hand-Gal4 expression pattern

Table 2.1.3: List of the various UAS expression lines used in this study.

Line	**Origin**
UAS-DN-Tin; UAS-DN-Tin	Han et al., 2002
UAS-Him (J7)	D. Liotta, M. Taylor
UAS-Mef2 (high)	Bour et al., 1995
UAS-Mef2 (low)	Gunthrope et al., 1999
UAS-pnr^{D4} (II)	Ramain et al., 1993
UAS-tin1 (III)	Yin and Frasch, 1989
UAS-tin5 (II)	Yin and Frasch, 1989
UAS-zfh-1	A. Postigo (Bloomington, stock 6879)
UAS-hand A9.0 (II)	A. Pauluat
UAS-pnrEnr	Klinedinst et al., 2003
UAS-pnr/TM3 (AS 33)	Haenlin et al., 1997
UAS-ush/cyo (U10)	Cubadda et al., 1997
UAS-lacZ	M. Baylies (Bloomington, stock 1776)
UAS-GFP/cyoftzlac; UAS-GFP/TM6B	Dutta et al., 2002
UAS-tin5; PCE*S5 L4-1	Generated for this study

2.1.3 Embryo collection and fixation for FISH (Fluorescent *In Situ* Hybridisation)

The embryos were collected as in 2.1.2 with the following changes: the embryos were collected into a wire-mesh basket submersed in TXN (0.7% NaCl, 0.04% TritonX-100) and rinsed in TXN and then in water. The chorion was removed as described in 2.1.2. The embryos were then again washed in TXN and water, transferred to a 2 ml tube containing a 1:1 mix (1 ml of each solution) of heptane and "ribofix solution" (containing 4 % formaldehyde; for 1ml: 25µl 20xPBS, 50µl 0.5M EGTA, pH 8.0, 896µl 10% methanol-free Formaldehyde (Polysciences)) and fixed for 25 minutes on a Lukeham shaker (setting 6-7). The ribofix solution was then removed and replaced with the same volume (1 ml) of methanol. Vigorous shaking of the embryos in this mixture for 30 seconds removed the vitelline membrane of the embryos. The embryos were then rinsed in 1 ml 100 % methanol, which was then removed and replaced by 1 ml 100% ethanol. The embryos were washed several times in 100% ethanol and stored at -20 °C in ethanol until required.

2.1.4 Counting pericardial cells in larvae

The number of pericardial cells in larvae was counted using a handCGFP reporter fragment which is expressed in all heart cells at all stages of the development of the fly (Sellin et al., 2006). The size and shape differences between the pericardial cells and the cardioblasts allow for an easy distinction of the two cell types in the larvae. In order to ensure that only larvae of the same developmental stage were compared, I collected embryos for each experiment on apple juice agar plates for a two hour window. These plates were then incubated at 25°C until the larvae had reached the desired point in their development. The pericardial cells were assayed in young first instar larvae (28-30 h AEL = after egg laying), young second instar larvae (52-54 h AEL) and mid-stage third instar larvae (100-102 h AEL).

At the correct time-point the larvae were picked off the apple juice plates, rinsed in distilled water, dried and mounted on a coverslide in a drop of immersion oil (Zeiss Immersol 518 N) and covered with a coverslip. As spacers I used the following: 2 layers of Scotch® Magic tape (3M, 19 x 25 mm) for the first instar larvae and a window cut into one layer of white insulation tape for second instar larvae and a window cut into

two layers of white insulation tape for third instar larvae. Slides with second and third instar larvae were then placed on an ice block that had been stored at -20°C to further slow down movement of the larvae. First instar larvae are not robust enough for this treatment, but their movement slowed down after approximately 3 to 5 min of exposure to the immersion oil. Despite this treatment the larvae remained alive for at least the 2 hour time window used to score them.

The pericardial cells of live larvae were then imaged under fluorescence on a Zeiss Axioscope (using the Zeiss 15 GreenH 546 filterset) using 10x, 20x and 40x magnification. Images were taken with an AxioCam HRc camera (Zeiss) and the 4.0 version of the Axiovision software (Zeiss) and then processed minimally in Photoshop CS2 or CS6 (Adobe).

2.1.5 Statistical analysis

For each cell counting experiment I aimed to count 20 different animals, for a few experiments I was not able to count the full 20 animals for a particular genotype and developmental stage. When this was the case this information is supplied in the appropriate graph. All graphs show the mean cell number and standard error for each experiment as calculated in Excel (Microsoft). All further statistical analysis was performed in Minitab (Minitab, Ltd.).

For the embryonic experiments I used three different controls (OR, handCGFP2.3 and handCGFP3.1; insertions of the handCGFP reporter fragment on the second and third chromosome, respectively). The results obtained for these controls in the embryonic experiments passed the three assumptions of ANOVA (Analysis of Variance; they are on an interval scale, normally distributed and have a similar degree of variance). Thus I proceeded using a 1-way ANOVA to verify my hypothesis that there is no significant difference between the three controls. This was indeed the case, thus all three controls have similar pericardial cell numbers during the embryonic stages.

As my analysis of the larval pericardial cells is dependent on the presence of the handCGFP reporter construct, the two different handCGFP lines served as controls for these stages. As with the embryonic analysis, the results obtained for these tests were suitable for an ANOVA at each time point. For the second and third instar larvae my

hypothesis that there is no significant difference between the controls was correct and I proceeded as with the embryonic cell counts. However, the controls at the first instar time-point did show a significant difference in the pericardial cell number obtained for the two controls. This most likely highlights a problem in the counting of the pericardial cells as they are still very similar in size to the cardioblasts and arranged partly on top of each other and in a three dimensional structure around the cardioblasts. In consequence of these results I have discounted the result for my cell counts in the first instar larvae as not reliable.

All experimental (non-control data) were grouped together (e.g. all *Him* null mutants and appropriate controls for one developmental point and type of cells). A large proportion of the data did not meet one assumption of ANOVA, they were not normally distributed, and so all data were thus analysed with non-parametric test methods. Kruskal-Wallis tests were used to detect significant differences between data sets, followed by pairwise Mann-Whitney tests to determine which experiments differed significantly when compared directly with the appropriate control. The p-value for each tested pairwise combination was noted. As the chance of a significant result becomes more likely with the large number of pairwise combinations tested, I applied the Bonferroni method (post-hoc test) to the results of the Mann-Whitney test. This test takes the number of pairwise comparisons that were tested into consideration and as a result lowers the p-value required for a statistically significant result correspondingly. All instances where a cell count is given as significantly different to the control in the results section have thus passed the Mann-Whitney test and the post-hoc test, which requires p-values of at least under 0.00625 for the chosen interval of 95 %.

2.1.6 Wing analysis

The analysis of the degree of wing translucence was performed as described in Tögel et al., 2008 and as described in personal communications with M. Tögel and A. Paululat.

In order to test for the correct “clearing” of the wings after eclosing from the pupae, the wings were removed from 6 hour old adult flies. The wings were then mounted on a dry slide and covered with a cover-slip (no mounting medium was used). The cover-slip was then sealed to the microscope slide with nail varnish. The

translucency of the wings was then evaluated under a Zeiss Axioscope using 5x, 10x and 20x magnification. Images were taken with a AxioCam HRc camera (Zeiss) and the 4.0 version of the Axiovision software (Zeiss) and then minimally processed in Photoshop CS2 or CS6 (Adobe).

2.1.7 Generation of transgenic lines of *Drosophila melanogaster*

New transgenic lines of *Drosophila* were generated by microinjection of the construct into the posterior of a pre-cellularised *yw* embryo, the area where the pole cells and thus the future gonads of the animal are formed. Microinjections were performed by Jun Han. The construct integrates into the fly genome from the plasmid with the help of a separately supplied transposase; insertion of the construct into the germline established a stable transgenic line. Each construct also carried the *mini-white* selection marker, which, if integrated into the genome, restores the red eye colour to the previously w^- adult flies. This allowed for an easy selection of transgenic flies.

The DNA mix for microinjection consisted of a final concentration of 1µg/µl construct DNA and 0.25µg/µl helper DNA (plasmid pπ25.1; Rubin and Spradling, 1982; source of the Δ3-4 transposase) in 1x Spradling buffer (1mM Sodium Phosphate buffer pH 6.8, 0.5mM KCl). Before injection the mix was centrifuged at maximum speed for 10 minutes to remove any salt percipiate and 0.5 to 1µl of the mix were loaded into the injection needles.

For microinjection, embryos of *yw* flies were collected in intervals of 30 minutes and then de-chorionated with a 1:1 mix of water:bleach (sodium hydrochloride, Sigma). The de-chorionated embryos were aligned in a row and transferred to a cover-slip edge that had previously been treated with glue (the glue solution is obtained by placing double-sided Scotch® tape (3M) in 50 to 100ml of heptane). The posterior of each embryo was oriented just at the edge of the cover-slip. The aligned embryos were then desiccated for approximately 10 minutes in a sealed desiccator chamber (Simax GL36) with Silica-gel 6-16 (mesh self-indicating, Fisher) and then covered in a thin layer of halocarbon oil (Voltalef, H10S) to prevent further drying out of the embryos. The microinjections needles used were purchased from Eppendorf (Femtotips II). A micromanipulator connected to a pump was then used to inject the construct and helper DNA into the embryo under a microscope (Nikon).

The injected embryos on the cover-slip were then placed on an apple-juice plate, which contained some yeast paste (dried bakers yeast mixed in dH_2O) and a wet filter paper (Whatman) to provide food and prevent drying out of the larvae, respectively. Hatched first instar larvae were transferred into a vial containing fly food and kept at 25°C until the adult (termed F_0) eclosed.

The eclosed F_0 flies were individually crossed with *yw* virgins. The offspring of these flies (F_1) had red eyes if the injected construct integrated into the fly genome. These were selected and by successive crossing of these flies to balancer stock (flies that carry a marker on each chromosome) the chromosome carrying each insertion was established. Where possible the transgenic lines were brought to homozygosity to establish a stock. A list of all generated transgenic fly strains and the lines tested for each strain can be found in Table 2.1.4 and Figure 3.5.1.

2.1.11 The Gal4/UAS system

The Gal4/UAS system allows over-expression of genes in specific tissues and subsets of cells (Brand and Perrimon, 1993). One line carries the coding sequence of the Gal4 protein under the control of a characterised promoter or enhancer element and a second fly line carries the gene, which is to be over-expressed, under the control of the UAS element. If females of the first line are crossed with males of the second line, the offspring expresses the gene under the control of the characterised promoter or enhancer element used to drive Gal4 expression.

Example of the fly crosses:

$$\frac{\text{Da-Gal4}}{\text{Da-Gal4}} \times \frac{\text{UAS-tinman}}{\text{UAS-tinman}}$$

$$\frac{\text{Da-Gal4}}{\text{UAS-tinman}}$$

Table 2.1.4: List of established transgenic lines for this study, their expression pattern and primers used to clone each enhancer fragment.
Lines: tested independent lines; PC: expression in pericardial cells; Cb: expression in cardioblasts; Origin: all constructs and lines were cloned and established in this lab, none are currently published; Fwd: forward primer; Rev: reverse primer

Name	Lines	PC	Cb	Origin	Fwd	Rev
PCE	6; 34; 41	✓	✗	S. McConnell	1C6SET3 FOR	1C6SET3 REV
PCE S4	L1-1; L1-2	✓	✗	S. McConnell	Set3_S4 _fwd	1C6SET3 REV
PCE Δextreg1	L1-1; L7-1; L9-1	✓	✓	this study	Setextreg1 _Fd-EcoRI	1C6SET3 REV
PCE Δreg1	L3-1; L5-2; L7-1; L13-1	✓	✓	this study	1C6_reg1 _Fwd-EcoRI	1C6SET3 REV
PCE LSA	L5-1; L7-2; L11-1	✓	✓	this study	Set3_LSA _reg1_Fwd	Set3_LSA _reg1_Rev
PCE LSA1-5	L1-2; L6-2; L6-3; L8-1	✓	✓	this study	Set3_LSA _1-5_Fwd	Set3_LSA _1-5_Rev
PCE Oct1	L2-1; L3-2	✓	✗	this study	Set3*oct -1.1_Fwd	Set3*oct -1.1_Rev
PCE* S5/S6	L1-1; L2-2; L3-1; L4-1	✓	✗	S. McConnell	1C6Set3 S5-FOR + 1C6Set3 S6-FOR	1C6Set3 S5-REV + 1C6Set3 S6-REV
PCE* S6	L1-1; L1-3; L3-1; L4-1 L4-3; L8-3)	✓	✗	S. McConnell	1C6Set3 S6-FOR	1C6Set3 S6-REV
PCE* S6new	L2-2; L3-2; L5-1; L5-2; L6-2	✓	✗	this study	Set3*S6 new_fwd	Set3*S6 new_rev
PCE Δtin1	L7-1; L9-1; L10-1	✗	✗	this study	1C6_reg1 /-tin_Fwd -EcoRI	1C6SET3 REV
PCE ΔMed	L1-1; L2-1; L6-1; L7-1; L8-1	✓	✗	this study	Set3med_ EcoRI_Fwd	1C6SET3 REV
PCE ΔEbox	L3-2; L4-1; L5-1; L6-2	✓	✗	this study	Set3EBox _EcoRI_Fwd	1C6SET3 REV
PCE S1	L6-6	✗	✗	S. McConnell	Set3_S1 _fwd	1C6SET3 REV
PCE Δtin2	L10-2	✗	✗	this study	1C6SET3 FOR	1C6-reg3/4 _Rev-XhoI
PCE S2	L10-1; “no line”	✓	✗	S. McConnell	1C6SET3 FOR	Set3_S2 _rev
PCE Δreg1 Δtin2	L2-2; L4-3	✗	✗	this study	1C6_reg1 _Fwd-EcoRI	1C6-reg3/4 _Rev-XhoI
PCE* S5	L1-1; L4-1	✗	✗	S. McConnell	1C6Set3S5 -FOR	1C6Set3S5 -REV
PCE* S5C-T	L2; L8	✓	✗	this study	Set3*S5 C-T_fwd	Set3*S5 C-T_rev
PCE* S5T-A	L8; L12	✓	✗	this study	Set3*S5 T-A_fwd	Set3*S5 T-A_rev
PCE LSD4-9	L2-1; L3-1; L5-2; L6-2	✓	✗	this study	Set3_LSD _4-9_Fwd	Set3_LSD _4-9_Rev
PCE LSE	L2-1; L5-1; L5-3	✓	✗	this study study	Set3_LSE _reg2_Fwd	Set3_LSE _reg2_Rev
PCE LSF	L1-2; L2-4; L3-2; L4-4s	✓	✗	this study	Set3_LSF _reg2_Fwd	Set3_LSF _reg3_Rev

2.2 Cell biology

2.2.1 Immunohistochemistry (detection of individual proteins in *Drosophila* embryos)

The antibody stains were performed essentially as described in Rushton et al. (1995). The fixed embryos were blocked for 3 x 10 minutes on a Lukeham shaker (setting 2) at room temperature in 1 ml PBS-Triton (0.1 % (v/v) Triton in 1 x PBS (140 mM NaCl, 6.5 mM KCl, 2.5 mM Na_2HPO_4, 1.5 mM KH_2PO_4; pH 7.5)) with 0.5 % (w/v) bovine serum albumin (PBS-Triton-BSA; BSA, Sigma), rinsed three times in 1 ml PBS and then incubated at 4 °C overnight on an A600 rocker (Denley) with the appropriate primary antibody (see table 2.2.1 for all antibodies used in this study). The embryos were rinsed in 1 ml PBS-Triton and washed 3 x 10 minutes on a Lukeham shaker (setting 2) in 1 ml PBS-Triton and incubated with biotinylated secondary antibody (1:200 dilution in 1 x PBS) for one hour at room temperature without any shaking. After rinsing for 3 times in 1 ml PBS-Triton and then 3 washes for 10 min at room temperature in 1 ml PBS-Triton, the embryos were incubated with the AB-complex of the Vectastain Elite ABC kit (Vector Laboratories) to amplify the signal. For each antibody reaction 2 µl of solution A and 2 µl of solution B were gently mixed in 200 µl 1 x PBS and the AB (Avidin-Biotin) complex was allowed to form by incubating the AB mixture for 30 minutes without shaking at room temperature. The pre-incubated AB-solution was then added to the embryos and incubated for a further 30 min at room temperature without shaking to allow the AB-complex to bind to the Biotin molecules on the secondary antibody. This was followed by three rinses in 1 ml PBS-Triton and three 10 minutes washes in 1 ml PBS-Triton on a Lukeham shaker (setting 2). The colour of the stain was then developed with 0.5 mg/ml 3,3'-diaminobenzadine-tetrahydrochloride (DAB; Sigma D9015) and 0.06 % (v/v) hydrogen peroxide in dH_20 (a 1/100 dilution of a commercially available 6% Hydrogen Peroxide Solution, 20 vols (Care +, L.C.M. Ltd). The staining reaction was stopped by rinsing the embryos three times in 1 ml of PBS-Triton, followed by three 10 minutes washes in 1 ml PBS-Triton.

The embryos were mounted in 80 % glycerol and analysed with a Zeiss Axioscope using 10x, 20x and 40x magnification. Images were taken with an AxioCam

Table 2.2.1: List of antibodies used in this study

Antibody	Final dilution	Animal raised in	origin
Anti-DIG	1:1000	Sheep	Roche
Anti-DIG-AP Fab Fragments	1:2000	Sheep	Roche
Anti-Even-skipped	1:3000	Rabbit	Manfred Frasch
Anti-GFP	1:2000	Mouse	Sigma
Anti-Mef2	1:1000	Rabbit	Bruce Patterson
Anti-mouse 488 IgG (Alexa)	1:200	Goat	Molecular Probes, Invitrogen
Anti-mouse 546 IgG (Alexa)	1:200	Goat	Molecular Probes, Invitrogen
Anti-Muscle Myosin	1:500	Rabbit	Daniel Kiehart
Anti-Odd	1:400	Rabbit	James Skeath
Anti-rabbit 488 IgG (Alexa)	1:200	Goat	Molecular Probes, Invitrogen
Anti-rabbit 546 IgG (Alexa)	1:200	Goat	Molecular Probes, Invitrogen
Anti-Stumps (Dof/Hbr)	1:200	Rabbit	Maria Leptin
Anti-tinman	1:800	Rabbit	Manfred Frasch
Anti-Zfh-1	1:1000	Mouse	Zhe-Chun Lai
Anti-β3-tubulin	1:1000	Rabbit	R. Renkawitz-Pohl
Anti-βgalactosidase	1:5000	Mouse	Promega
Anti-βgalactosidase	1:5000	Rabbit	Cappel (MP Biotech)
Biotinylated anti-mouse	1:200	Goat	Vector Laboratories
Biotinylated anti-rabbit IgG	1:200	Goat	Vector Laboratories
Biotinylated anti-sheep	1:200	Goat	Vector Laboratories

HRc camera (Zeiss) and the 4.0 version of the Axiovision software (Zeiss) and then processed minimally in Photoshop CS2 or CS6 (Adobe).

2.2.2 Immunohistochemistry (detection of multiple proteins in *Drosophila* embryos)

Multiple labelling of proteins in whole-mount *Drosophila* embryos essentially followed the same protocol described in 2.2.1. If the two primary antibodies were raised in different species, they were applied at the same time and then developed successively. For this the embryos were treated as described in section 2.2.1 but only one of the biotinylated secondary antibodies was added. After developing the first antibody in the presence of $NiCl_2$ to give a black precipitate, the embryos were rinsed three times in 1 ml PBS-Triton and then washed three times for 10 minutes on a Lukeham shaker (setting 2) in 1 ml PBS-Triton-BSA. Then the second biotinylated secondary antibody was added and developed as described in section 2.2.1. The less widely expressed protein and/or weaker antibody was usually developed first to ensure the stain is clearly visible. The first antibody was developed in the presence of $NiCl_2$ and thus the formed precipitate has a black colour. The second antibody is developed without any salts and the precipitate is brown allowing for a distinction between the two antibodies.

If both antibodies were raised in the same animal, the protocol described in 2.2.1 was followed for the first antibody. After developing this antibody (again in the presence of $NiCl_2$ to give a black colour) the embryos were rinsed three times in 1 ml of PBS-Triton and then washed three times for 10 minutes on a Lukeham shaker (setting 2) in 1 ml PBS-Triton each. The embryos were then blocked in 1 ml PBS-Triton-BSA on a Lukeham shaker (setting 2) for 30 minutes. Then, the second primary antibody was added to the embryos and the embryos were again incubated on an A600 rocker at 4°C over-night. The remainder of the stain was carried out as described in section 2.2.1 and the second colour reaction was again performed in the absence of salts to give a brown precipitate.

2.2.3 Single fluorescent labelling of proteins in whole-mount *Drosophila* embryos

The fluorescent antibody stains followed the same protocol as the immunohistochemical procedure with the following alterations I made after optimizing the protocol described in 2.2.1 for fluorescence. The fixed embryos were washed three times for 10 minutes in 1 ml PBS-Triton and 2 % (v/v) normal goat serum (Sigma) before the primary antibody was added. After the incubation with the primary antibody the embryos were rinsed three times and then washed four times for 10 min in PBS-Triton and 2 % normal goat serum. If the used antibody was known to generate high levels of background, before the last wash the embryos were incubated in Vectastain SignalEnhancer (Molecular Probes, Invitrogen) for 30 min on a Lukeham shaker (setting 2) to reduce background as much as possible. The secondary antibodies (see Table 2.2.1, Molecular Probes, Invitrogen) were added in a 1/200 dilution in PBS-Triton with 2 % (v/v) normal goat serum (total volume 200 µl per reaction) and incubated for approximately two hours at room temperature without shaking. The embryos were then rinsed three times in 1 ml PBS-Triton and washed six times for 10 min in 1 ml PBS-Triton on a Lukeham shaker (setting 2) before mounting in "Fluorescent Mounting Medium" (90% glycerol, pH 8.6, 25mg/ml DABCO (Sigma; Wood, W. pers. communication). The embryos were imaged with a Zeiss Axioscope using 10x, 20x and 40x magnification. Images were taken with an AxioCam HRc camera (Zeiss; using the Zeiss 15 GreenH 546 and the Zeiss 10 Blue 450-490 SB filtersets) and the 4.0 version of the Axiovision software (Zeiss) and processed minimally in Photoshop CS2 or CS6 (Adobe). Alternatively, the embryos were analysed with a Leica DM6000 microscope and the Leica TCS SP2 AOBS spectral confocal system. Single image slices were selected or merges of image stacks created and further image processing was then done in the accompanying Leica Confocal Software.

2.2.4 Multiple fluorescent labelling of proteins in whole-mount *Drosophila* embryos

Multiple labelling of proteins in whole-mount *Drosophila* embryos with fluorescent secondary antibodies essentially followed the same protocol described in 2.2.2 and 2.2.3. If the two primary antibodies were raised in different species, they were

applied at the same time and the secondary fluorescent-labelled antibodies were also applied in one step during the secondary antibody incubation.

If both primary antibodies were raised in the same animal, double-labelling was achieved by completing the protocol for single fluorescent labelling for the first antibody (see section 2.2.3). Then the embryos were rinsed three times in 1 ml PBS-Triton and washed four times for 10 minutes in PBS-Triton + 2 % (v/v) normal goat serum to make sure to block the embryos well between the incubation between the first secondary and the second primary antibodies. The second first antibody was then added to the embryos and the staining was finished as described in section 2.2.3.

2.2.5 RNA *In situ* Hybridisation

The distribution of a specific transcript in whole-mount embryos is visualized using the method described by Taylor (2000) and uses Digoxigenin-labelled RNA probes (for the generation and quantification of the DIG labelled RNA probes please see 2.3.11 and 2.3.12, respectively). The fixed embryos were transferred to a 1.5 ml tube and slowly re-hydrated in 3 steps (rinse in 1.5 ml 100 % methanol, then in 1.5 ml of 3:2 MeOH/4 % paraformaldehyde in PBS solution, then let stand for 5 min in 1:2 MeOH/4 % paraformaldehyde in PBS solution). Following this, a 10 minutes post-fixation step in 1.5 ml of 4 % paraformaldehyde in PBS on the Lukeham shaker (setting 1.5) was performed. The embryos were rinsed and washed in 1.5 ml PBT (1x PBS + 0.1% Tween20) and pre-hybridised for 1 hour at 55°C in hybridsation buffer (50% formamide (Fluka), 4 x SSC (Sigma), 1 x Denhardt's solution (Sigma), 250 µg/ml yeast tRNA (BRL), 250 µg/ml salmon testis DNA (Sigma), 50 µg/ml heparin (Sigma) and 0.1 % Tween-20). The embryos were then incubated on the rocking platform of the hybridisation oven (Techne Hybridizer HB-1D) for a minimum of 14 hours at 55°C in the presence of a final concentration of 0.125µg of DIG-labelled RNA probe in a total of 500 µl hybridisation buffer in a 1.5 ml tube.

After the hybrdisiation the embryos were washed in 1 ml washing solution (50 % formamide, 2 x SSC, 0.1% Tween-20) for 4 times spaced evenly throughout the day at the hybridisation temperature (e.g. 10 am, 12.30 pm, 2.30 pm and 5 pm), with the last wash being overnight. The embryos were then rinsed in 1 ml PBT and washed in 1 ml PBT for 30 minutes at room temperature and then incubated for 90 minutes in anti-DIG-

AP antibody (Roche, 1:2000 dilution) in 200 µl total volume of PBT + 5 % (v/v) normal goat serum. The embryos were rinsed again in 1 ml PBT and washed 4 x 20 minutes in 1 ml PBT and then rinsed twice and washed once for 5 minutes in AP buffer (100 mM Tris pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$, 0.1% Tween). The colour reaction was then allowed to proceed in the dark by the addition of 0.3 ml of AP buffer and 2.7 µl NBT (4-NitroBlue Tetrazolium Chloride Solution; Roche) and 2.1 µl BCIP (5-Bromo-4-chloro-3-indolyl-phosphate, 4-toluidine salt; Roche) to the embryos and the progress monitored under a dissecting microscope. The reaction was stopped by rinsing the embryos in 1 ml PBT three times and washing the embryos in 1 ml PBT for 10 minutes for another three times. The embryos were then left to equilibrate in 80 % glycerol overnight before being analysed using a Zeiss Axioscope using 10 x, 20 x and 40 x magnification. Images were taken with a AxioCam HRc camera (Zeiss) and the 4.0 version of the Axiovision software (Zeiss) and then processed minimally in Photoshop CS2 or CS6 (Adobe).

2.2.6 RNA *In Situ* and immunohistochemical double labelling of RNA transcripts and proteins

The protocol for the first two days is the same as the one described in 2.2.4. The protocols differ from the antibody incubation onwards. Here, the primary antibody at the appropriate dilution and the anti-DIG-AP antibody (Roche, 1:2000) were added together in a total volume of 200 µl of PBT and incubated for 2 hours at room temperature without shaking, the tubes were placed on their sides. The embryos were then washed 4 x 15 minutes in 1 ml PBT and incubated in the biotinylated secondary antibody (1:200) and the anti-DIG-AP antibody (1:2000) in a total volume of 200 µl of PBT for 1 hour without any shaking. Following this, the embryos were again washed 4 x 15 minutes in 1 ml PBT and then incubated in the Vectastain Elite ABC (Molecular Probes) reagent for 30 minutes (leave the Avidin-Biotin complex to form for 30 minutes before applying to the embryos, see section 2.2.1 for details). The embryos were washed three times 10 minutes in 1 ml PBT. The antibody stain is developed first (according to the protocol described in 2.2.1, 600 µl PBT + 30 µl DAB (0.5 mg/ml), catalyst are 20 µl 0.06 % (v/v) H_2O_2 in dH_2O, no salts). This reaction mix was removed and the embryos washed three times in 1 ml PBT for 15 minutes before being transferred into AP buffer

(100 mM Tris pH9.5, 100 mM NaCl, 50 mM $MgCl_2$, 0.1% Tween-20) and washed 3 x 5 minutes in 1 ml of this buffer. The labelled RNA probe was visualised according to the protocol described in 2.2.4 (600 µl AP buffer + 2.7 µl NBT + 2.1 µl BCIP). The embryos were then washed in PBT and were, before analysis under the Zeiss Axioskop (see above), left to equilibrate in 80 % glycerol overnight.

2.2.7 Fluorescent *In situ* Hybridisation (FISH)

M. Drechsler and A. Paululat (University of Osnabrück, Germany) kindly gave the FISH protocol to us. The embryos were washed on a Lukeham shaker (setting 2) for five times 10 minutes in 1 ml PBT (1 x PBS, 0.1 % Tween-20) and then post-fixed in 1 ml 10 % methanol-free formaldehyde (Polysciences, Germany) for 20 min and washed again five times 10 minutes in 1 ml PBT on a Lukeham shaker (setting 2). This was followed by a wash in 1:1 PBT:hybridisation buffer (for 100 ml: 50 ml Formamide, 25 ml 20 x SSC, 0.1 ml 50 mg/ml heparin, 1 ml 10% TritonX-100) and then two 10 minute washes in 1 ml hybridisation buffer on the Lukeham shaker (setting 2). The embryos were then incubated in 1 ml of hybridisation buffer on the rocker platform of the hybridisation oven (Techne Hybridizer HB-1D) for at least an hour at 55°C.

The DIG-labelled RNA probe was diluted to 1:500 (for both *Him* and *lacZ*) in hybridisation buffer (500 µl total volume per reaction) and denatured at 80°C for a minimum of 15 minutes and then stored on ice for 1 minute immediately before use. The probe was transferred onto the embryos and incubated for at least 40 hours at the hybridisation temperature (55°C) on the rocker platform of the hybridisation oven.

After hybridisation the embryos were washed three times 60 minutes in 1 ml of hybridization buffer on the rocker platform of the hybridisation oven and then twice for 30 minutes in a 1:1 mixture of PBT:hybridisation buffer (1 ml final volume). All these wash steps were carried out at the hybridisation temperature of 55°C. After the last wash in PBT:hybridisation buffer the embryos were washed 5 x 20 min at room temperature in 1 ml PBT and incubated with pre-absorbed (to reduce unspecific binding of the antibody) anti-DIG antibody (1:1000, made in sheep) at 4°C overnight. On the next day the embryos were washed four times 30 minutes in 1 ml PBT at room temperature on the Lukeham shaker (setting 2) and incubated in biotinylated anti-sheep antibody (1:200) in a total volum of 200 µl PBT at 4°C overnight on a rocker.

The next day the embryos were washed four times 30 minutes in 1 ml PBT and incubated in the Vectastain AB Complex (Molecular Probes, 200 µl final volume per reaction, see section 2.2.1 for details) for 1 hour at room temperature (the Avidin-Biotin complex was left to form for 30 minutes before incubation). The embryos were washed five times 10 minutes in 1 ml PBT and incubated 45 to 60 minutes in TSA Cy3 or TSA Cy5 (diluted 1:50 in amplifier solution which was supplied with the Vectastain TSA Cy3 or Cy5 Kits, Perkin Elmer). From this point onwards the embryos need to be kept in the dark under aluminium foil. The embryos were then washed again for five times 10 minutes in 1 ml PBT, followed by heating the embryos to 70°C for 15 minutes while suspended in PBT, followed by 3 more 10 minute washes in 1 ml PBT.

If no further antibody stain was necessary, the embryos were then transferred into approximately 500 µl Fluorescence Mounting Medium (90% glycerol in PBS, pH 8.6, 25mg/ml DABCO; W. Wood, pers. communication) and left to equilibrate overnight at 4°C before analysis. The embryos were imaged with a Zeiss Axioscope using 10x, 20x and 40x magnification. Images were taken with an AxioCam HRc camera (Zeiss; using the Zeiss 15 GreenH 546 and the Zeiss 10 Blue 450-490 SB filtersets) and the 4.0 version of the Axiovision software (Zeiss) and then processed minimally in Photoshop CS2 or CS6 (Adobe). Alternatively, the embryos were analysed with a Leica DM6000 microscope and the Leica TCS SP2 AOBS spectral confocal system. Single image slices were selected or merges of image stacks created and further image processing was performed in the accompanying Leica Confocal Software.

2.2.8 Antibody stain after FISH

This slightly altered protocol (M. Drechsler, S. Albrecht, A. Paululat, pers. communication) for fluorescent antibody staining was followed after the completion of the FISH protocol. It is based on the protocol described in sections 2.2.3 and 2.2.4.

Following the last wash of the FISH protocol the embryos were incubated for 30 minutes in 1 ml freshly made blocking solution (PBT + 2% normal goat serum) at room temperature on the Lukeham shaker (setting 2) and then the appropriate primary antibody (diluted in blocking solution, 200 µl final volume, see table 2.2.1) was added and the embryos were incubated on a slow rocker at 4°C overnight.

The next day the primary antibody was removed and the embryos rinsed three times in 1 ml blocking solution, washed four times 10 minutes in 1 ml blocking solution on the Lukeham shaker (setting 2) and blocked for 1 hour at room temperature in 1 ml blocking solution on the Lukeham shaker (setting 2). During this time the fluorescent secondary antibody was pre-absorbed to reduce background fluorescence. For the pre-absorption the secondary antibody was incubated (1:200 dilution) without shaking on embryos not expressing the antigen for approximately 1 hour at room temperature. These embryos were treated as described in section 2.2.1 until the first antibody incubation. The diluted fluorescent secondary antibody was then removed from the blank embryos and transferred to the tube with the experimental embryos and incubated for 2 hours at room temperature on a Lukeham shaker (setting 1). The embryos were then rinsed three times in 1 ml PBT and washed six times for 10 minutes in 1 ml PBT. After the final wash the embryos were transferred into Fluorescent Mounting medium (90% glycerol in PBS, pH 8.6, 25mg/ml DABCO; W. Wood, pers. communication) and left to equilibrate overnight at 4°C before analysis. The imaging was done as described in section 2.2.6.

2.3 Molecular Biology

2.3.1 Bacteria

The bacterial strain used was *Escherichia coli DH5α.* The *E. coli* bacteria were grown in LB-medium in a shaking incubator (250 rpm) or on LB-Agar plates (containing ampicillin at a concentration of 100 μg/ml) at 37°C.

2.3.2 Preparation of Plasmid DNA

a) Boiling Mini Prep

This preparation was used to analyse 1.5 to 5 ml of culture. The bacteria were grown overnight and spun down for 15 min at maximum speed in a benchtop centrifuge (13 000 rpm; Biofuge pico, Heraeus Instruments). The supernatant was decanted and the bacterial pellet was air-dried for approximately 5 minutes. The pellet was then

resuspended in 150 µl STET-buffer (10 mM Tris (pH 8.0), 25 mM EDTA, 100 mM NaCL, 0.1 % Triton X-100, 8 % sucrose) with Lysozyme (1 mg/ml final concentration; Sigma) and boiled for exactly 45 seconds. This step lysed the cells. The genomic DNA, proteins and cell debris were spun down for 15 min at maximum speed (13 000 rpm) in a benchtop centrifuge. The supernatant was transferred to new 1.5 ml tube and the same volume of isopropanol (150 µl) at room temperature was added to precipitate the DNA. The samples were carefully mixed by inverting the tubes and the DNA was pelleted in 1.5 ml tubes by centrifugation for 15 min at maximum speed. The supernatant was decanted and the pellets washed with 1 ml 70 % ethanol and then dried. The dried pellet was resuspended in 30 µl dH_2O and further analysed as described below.

b) Midi Preps

For the preparation of plasmid DNA in larger quantities the Qiaquick Filter Plasmid Midi Kit (Qiagen) was used. The preparation was done according to the manufacturer's manual; the DNA was then resuspended in 100 µl dH_2O.

2.3.3 Analysis of DNA with restriction enzymes

All restriction enzymes were supplied by NEB and used according to their guidelines. Digests were in a total volume of 50 µl, in the presence of BSA (final concentration of 100 µg/ml) using 5 units of enzyme. Double digests were done in the buffer specified by the NEB buffer compatibility table.

2.3.4 Phenol-Chloroform extraction

Phenol-Chloroform extraction was used to produce clean DNA preparations for cloning after restriction digests. The volume of the sample to be extracted was adjusted with ddH_20, e.g. to 100 µl if necessary. Then the same volume of phenol-chloroform-isoamylalcohol (25:24:1) was added and the sample was mixed well by inversion of the tube. Following centrifugation at maximum speed for 5 min in a centrifuge (13 000 rpm, Biofuge pico, Heraeus Instruments), the upper phase was removed and the supernatant was ethanol precipitated (Section 2.3.5) before further analysis.

2.3.5 Ethanol precipitation

The volume of the sample to be precipitated was brought up to a volume that is easy to handle, e.g. 100 µl. A tenth of the sample volume (e.g. 10 µl) of NaOAc (3 M, pH 5.2; Sigma) was added and the sample was mixed well by inversion. Then twice the sample volume (e.g. 220 µl) of ice-cold 100 % ethanol were added and mixed again by inverting the sample tube. The sample was kept on ice or in the freezer (-20°C) for at least 30 min. After this, the sample was centrifuged at maximum speed for 20 min. The pellet was washed in 70 % ethanol, air-dried and resuspended in the desired volume dH_20 (e.g. 100 µl).

2.3.6 Dephosphorylation of DNA

In order to reduce the self-ligation of vectors during ligation, digested vectors were dephosphorylated. The digested vector was extracted with phenol-chloroform (Section 2.3.4) and ethanol precipitated (Section 2.3.5). The precipitated vector was resuspended in 27 µl ddH_2O. Then 3 µl of the 10 x CIP-buffer (NEB) and 2 µl of CIP (calf intestinal phosphatase, NEB) were added. The mix was incubated at 37°C for 20 min. After this time, 1 µl of CIP was added and the mix was incubated at 37°C for a further 10 min. After the dephosphorylation of the DNA the alkaline phosphatase was inactivated by incubating the mix for 10 min at 65°C.

2.3.7 Ligation of DNA fragments

Ligations were done in a total volume of 10 or 20 µl using 5 or 10 units of T4 Ligase (Roche), respectively. The 2 x T4 Rapid Ligation Buffer was used according to the manufacturer's specifications. The amounts of DNA used for the ligations was calculated according to the following formula:

ng of vector x kb size of insert/kb size of vector x molar ratio of insert/vector = ng of insert

(obtained from the Promega manual for pGEM-T, the molar ratio used was 1:1). Since all ligations for this study involved an *EcoRI* restriction site, the ligations were done at 12°C overnight. Table 2.3.1 lists all plasmids used in this study.

2.3.8 pGEM-T cloning

DNA fragments generated by PCR were immediately sub-cloned into the pGEM-T vector (Promega). The pGEM-T Kit was used according to the manufacturers (Promega) instructions. The amount of insert needed was calculated according to the formula given in Section 2.3.7, the pGEM-T vector was supplied at 50 ng/µl and has an approximate size of 3 kb.

2.3.9 Generation of competent *E. coli* DH5α cells

DH5α *E. coli* cells were grown overnight at 37°C at approximately 250 rpm in 5 ml of SOB medium (for 1l SOB medium: 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, dissolve in 950 ml dH_2O and add 10 ml 250 mM KCl, adjust to pH 7.0 and to a volume of 1 L, autoclave. Before use add 5 ml of a sterile 2 M $MgCl_2$ solution.) in a 14 ml round-bottomed sterile plastic tube. This culture was added to 500 ml of SOB medium in a sterile 2 L culture flask. The bacteria were then grown in the described conditions until they reach an OD_{650} of 0.5. It is important to not overgrow the bacteria. Once this OD has been reached the cells were kept on ice as much as possible. The culture was transferred into sterile 250 ml centrifuge containers and spun for 5 minutes at 4 000 rpm in a pre-cooled rotor (Sorval centrifuge, GSA rotor). The supernatant was discarded while keeping the cells on ice and the cell pellet was gently resuspended in 150 ml cold (4°C) TFB1 (100 mM $RbCl_2$, 50 mM $MnCl_2 \cdot 4H_2O$, 30 mM KAc, 10 mM $CaCl_2 \cdot 2H_2O$, 15 % (v/v) glycerol, pH adjusted to 5.8 with 0.2 M glacial acetic acid, filter sterilized and store at 4°C in a sterile bottle). The cells were pelleted by

Table 2.3.1: List of all plasmids used in this study.

Name	Notes	Origin
Him cDNA	In *pBluescript* Used to generate *Him RNA* probe	Our lab
lacZ construct	In *pBluescript* Used to generate *LacZ RNA* probe	Richard Cripps
Sd construct	Generated from cDNA clone RE23308	DGRC
Hand construct	Used to generate *Hand RNA* probe	Achim Paululat
Tin cDNA	Suitable for in vitro transcription/translation	Manfred Frasch
pGEM-T®	Cloning vector	Promega
pBluescript II KS+	Cloning vector	Stratagene
pCaSpeR hs43 βGal	Vector containing the hs43 promoter to test for enhancer activity, Fly injection vector	Carl Thummel, Hanh Nughyen

centrifuging the suspension at 4 000 rpm for 5 minutes. The supernatant was removed again and the cell pellet resuspended in 20 ml TFB2 (10 mM MOPS, 10 mM $RbCl_2$, 75 mM $CaCl_2 \cdot 2H_20$, 15 % (v/v) glycerol, pH adjusted to between 6.8 and 7.0, filter sterilize and stored at 4°C in an autoclaved bottle). The cell suspension was aliquoted into 200 µl amounts in 1.5 ml tubes on ice and at a room temperature of 4°C. The aliquots were then snap frozen in liquid nitrogen, and the competent cells stored at -70°C.

2.3.10 Transformation of competent *E. coli* cells

The *E. coli* strain DH5α was used for transformations. The competent cells were generated as described in Section 2.3.9. The cells were removed from storage at -70°C and thawed on ice. 3 µl of the ligation mix was pre-chilled in a round-bottom Falcon tube and then 50 µl of the thawed competent cells were added to the ligation mix. The cells and the ligation mix were incubated on ice for 30 min and then heat-shocked at 42°C for 30 seconds. After the heat-shock, the cells were immediately returned to ice for 2 min. Then 450 µl of pre-warmed SOC medium (for 1l SOC medium: 20 g tryptone, 5 g yeast extract, 0.5 g NaCl, dissolved in 950 ml dH_2O and add 10 ml 250 mM KCl, adjust to pH 7.0 and to a volume of 1 l, autoclave. Before use add 5 ml of a sterile 2 M $MgCl_2$ solution and 20 ml sterile 1 M glucose solution) were added to the cells, and the cells incubated for 30 min at 37°C to enable establishment of Amp-resistance. After this incubation, the cells were plated on pre-warmed LB-Amp-Agar plates and incubated overnight at 37°C.

2.3.11 PCR (Polymerase-Chain-Reaction)

The primers for the enhancer fragments created in this study were ordered from MWG and resuspended in Molecular Biology Reagent grade water (Sigma) to a stock concentration of 100 pmol/µl and stored at -20 °C. This stock was then further diluted to 25 pmol/µl which was used for the reactions. The primers were used to introduce an EcoRI restriction site on the 5' end and an XhoI restriction site at the 3' end of the amplified DNA fragment. These restriction sites were then used to check for the

correctly sized fragment after cloning into the pGEM-T vector and then to directionally sub-clone the fragments into the pCaSpeR injection vector.

The PCR reactions were done in a total volume of 50 µl using 2 units of polymerase, approximately 10 ng of template DNA (the starting point for the constructs used for this study was the 1C6 Set3 pGEM clone 3;2 which was generated by S. McConnell) and 25 pmol/µl of each primer per reaction (see Table 2.3.2 for all primers used in this study). All reactions were initially optimised with Taq-Polymerase (Bioline) and then performed with Dynazyme (Finnzyme), which is mix of a proof-reading and non-proof-reading polymerase. Thus, this amplified fragment was suitable for the A/T-cloning. The fragment generated in the reaction with the proof-reading polymerase was then subsequently cloned into pGEM-T and the sequence verified (see attached DVD for all sequence data). A list of all primers and the constructs generated used in this study can be found in Table 2.3.2.

2.3.12 RNA probe synthesis

The template plasmid (a minimum of 2 µg template DNA) for the *RNA* probe (*Him* cDNA sequence cloned into pBluescript, D. Liotta, pers. communication) was linearised at the *5'* end of the cDNA insert using the appropriate restriction enzyme (for *Him*: Hind III, NEB) in a total volume of 10µl. Of this reaction, 5 µl (for probe used in regular RNA *in situ;* for probe used in FISH, the whole reaction (10 µl) was used) were added to the polymerase-mix (total final volume 20µl) that contains 2 µl 10 x DIG labelling mix (Boehringer), 1 µl RNase Inhibitor (Boehringer) and 1 µl (20 units) of the appropriate polymerase (Boehringer) for the plasmid (*Him*: T7 polymerase). This was incubated for 2-3 hours at 37°C. Then a DNase mix (1 µl 10 x DNase1 buffer, 3 µl (30 units) DNase1, 6 µl dH_2O; Boehringer) was added to destroy the plasmid template and incubated for 15 minutes at 37°C. This reaction was stopped by the addition of 80 µl 125 mM sodium carbonate buffer (pH 10.2) and incubated at 60°C for 15 minutes. The reaction was then immediately stored on ice and 50µl of 7.5M ammonium acetate were added and the solution was mixed. Then 375µl of ethanol were added on ice and the reaction was left to precipitate on ice for a minimum of 10 min and centrifuged at maximum speed for 15 minutes. The supernatant was then removed and the pellet was

Table 2.3.2: Primers used to generate the enhancer constructs used in this study. Forward primers introduced an EcoRI restriction site, reverse primers introduced an XhoI restriction site into the PCR product.

Name	Notes	sequence
1C6_reg1_Fwd-EcoRI	Introduced an EcoRI site; forward primer for the PCEΔreg1 construct	5'GGGAATTCCGAGATATTCCCAAGCGGC 3'
1C6_reg1/-tin_Fwd-EcoRI	Introduced an EcoRI site; forward primer for the PCEΔtin1 construct	5'GGGAATTCCACTCTGAAACATCTTGAAGTC 3'
1C6-reg3/4_Rev-XhoI	Introduced an XhoI site; reverse primer for the PCEΔtin2 construct	5'GGCTCGAGTGTTCTGTGTGCGCAATG 3'
1C6SET3FOR	Introduced an EcoRI site; designed by S. McConnell; forward primer for the standard PCE construct	5'GGGAATTCGGGTTTGTCGGCAATGATTTAC 3'
1C6SET3REV	Introduced an XhoI site; designed by S. McConnell, reverse primer for the standard PCE construct	5'GGCTCGAGCATCATCGTGTCGAGCGTCATAATC 3'
Set3_S1_fwd	Introduced an EcoRI site; designed by S. McConnell; forward primer for the PCE S1 construct	5'GGGAATTCGTATATATGTATGTATGTA 3'
Set3_S2_rev	Introduced an XhoI site; designed by S. McConnell	5'GGCTCGAGCATATTATAAATAGCGCCTG 3'
Set3_S4_fwd	Introduced an EcoRI site; designed by S. McConnell; forward primer for the PCE S4 construct	5'GGGAATTCCTCTTCCCCCCTCCAAA 3'
Set3E-Box_EcoRI_Fwd	Introduced an EcoRI site; forward primer for the PCEΔEbox construct	5'GGGAATTCCATACACATGTATATATGTATGTATGTAC 3'
Set3med_EcoRI_Fwd	Introduced an EcoRI site; forward primer for the PCEΔMed construct	5'GGGAATTCCTAGCAGAGAGACCCACAATA 3'
Set3_extreg1_Fwd	Introduced an EcoRI site; forward primer for the PCEΔextreg1 construct	5'GGGAATTCATTCCCCGAGA<TATTCCCAAG 3'

air-dried and resuspended in 30µl TE:formamide (1:1) for regular in situ use or in 150µl FISH hybridisation buffer for use in FISH.

2.3.13 Quantification of RNA probes

For use in the non-fluorescent *in situ* protocol, the generated RNA probe was diluted to a concentration of 5 ng/µl. In order to achieve this, specific dilutions of the synthesized probe were compared to standard concentration of a control RNA (Boehringer). 1µl of each dilution was spotted onto a positively charged nylon membrane (Roche, product number 1 209 272) and crosslinked using the "auto-crosslink mode" to the membrane (UV Stratalinker). The membrane was then washed 2 x 5 minutes in blocking solution (10 % (w/v) blocking reagent (Roche catalogue number 11096176001) in maleic acid buffer (4 ml 1M maleic acid, 6 ml of 1M NaCl for 40 ml, filter before adding blocking reagent and adjusting pH to 7.5)). The membrane was blocked for 30 minutes in blocking solution at room temperature and then incubated in a 1:2000 dilution of anti-DIG-AP Fab fragment in blocking solution for 30 minutes. The antibody was removed and the membrane was washed 2 x 15 minutes in blocking solution, rinsed once and washed twice for 10 minutes in washing solution (for 20ml: 2ml 1M NaCl, 100µl 1M $MgCl_2$, 2ml 1M Tris pH9.5, 200µl 10 % Tween). The colour reaction was left to develop in the developing solution in the dark (washing solution + 4.5µl/ml NBT + 3.5µl/ml BCIP). This reaction was stopped by repeated washes in PBT and the intensity of the spotted probe dilutions was compared to that of the RNA standard.

2.3.14 Generation of the deletion constructs

All enhancer fragments were amplified with specifically designed primers and a high fidelity polymerase and then cloned into the pGEM-T vector (pGEM-T Kit; Invitrogen). Once the PCR product was cloned into the pGEM-T vector, a medium scale DNA preparation was produced (Qiaquick Filter Plasmid Midi Kit, see section 2.3.2 b). This DNA was sequenced by Lark Technologies, Inc. (to their Silver Standard) to verify the correct sequence (see attached DVD for all sequencing data). The DNA preparation

was treated with restriction enzymes and the insert was sub-cloned into a *lacZ* reporter vector (pCaSpeR, Thummel et al., 1988). The reporter vector was injected into *Drosophila* embryos in the blastoderm stage. Eclosing flies were screened for transgenic animals (see Section 2.1.7). A list of the generated deletion transgenic flies is available in Table 2.1.4; Table 2.3.2 contains a list of all primers used to generate the analysed deletion constructs and Figure 3.5.1 is an overview about all enhancer constructs generated in this study.

2.3.15 Site-directed mutagenesis

The site-specific mutations of transcription factor binding sites were introduced using the QuickChange Site-Directed Mutagenesis Kit (Stratagene). This kit uses the ability of certain DNA polymerases to allow sequence mis-matches in the middle of a primer sequence. Combined with the use of the *DpnI* endonuclease, which specifically recognizes methylated DNA (i.e. only the template DNA produced and methylated in bacteria) and not the DNA produced by the *in vitro* polymerase reaction, this kit allows for the efficient introduction of mutations into known sequences.

The *Set3 *S5/S6* and *Set3 *S5* constructs and the primers for the *Set3 *S6* construct had been created previously (S. McConnell & M. Taylor, unpublished data). All other primers were designed with "QuickChange primer" program available from the Stratagene homepage (http://labtools.stratagene.com/QC). The polymerase reaction to introduce the mutated primer into the sequence was done according to the manufacturer's instructions. All potential mutated clones were sent off for sequencing to Lark Technologies, Inc. Once the correct sequence was verified (see attached DVD for sequencing data), the sequences were cloned into the pCaSpeR vector as described in Section 2.3.14. Table 2.3.3 lists all primers used to introduce mutations into the PCE sequence and Table 2.1.4 lists all fly lines generated in this study.

Table 2.3.3: Primers used to introduce point-mutations into the PCE sequence.

Name	Sequence
Set3_LSA_1-5_Fwd	5'ACAGATATGTATATTCCCCGAGAGCGGACCAAGCGGCCAAAAATAGACGC 3'
Set3_LSA_1-5_Rev	5'GCGTCTATTTTTGGCCGCTTGGTCCGCTCTCGGGGAATATACATATCTGT 3'
Set3_LSA_reg1_Fwd	5'CAGATATGTATATTCCCCGAGAGCTCTAGAGCAGGCCAAAAATAGACGC 3'
Set3_LSA_reg1_Rev	5'GCGTCTATTTTTGGCCTGCTCTAGAGCTCTCGGGGAATATACATATCTG 3'
Set3_LSE_reg2_Fwd	5'GGGTTGCATTGCGCACACACAGAATGCTCTAGACGAGAGTTCAAGTGCATGCCG 3'
Set3_LSE_reg2_Rev	5'CGGCATGCACTTGAACTCTCGTCTAGAGCATTCTGTGTGTGCGCAATGCAACCC 3'
Set3_LSD_4-9_Fwd	5'GTAACGCACTTGAAGTGCACTCTGAACACGATTGAAGTCCAAATAAAATAGCAGAGAGAC 3'
Set3_LSD_4-9_Rev	5'GTCTCTCTGCTATTTTATTTGGACTTCAATCGTGTTCAGAGTGCACTTCAAGTGCGTTAC 3'
Set3*oct-1.1_Fwd	5'CCCAAGCGGCCAAAAATAGACATGGCTTGTAACGCACTTGAAGTGCACTC 3'
Set3*oct-1.1_Rev	5'GAGTGCACTTCAAGTGCGTTACAAGCCATGTCTATTTTTGGCCGCTTGGG 3'
Set3*S5C-T_fwd	5'CAGAACATGTGAATGCAGAGTTTAAGTGCATGCCG 3'
Set3*S5C-T_rev	5'CGGCATGCACTTAAACTCTGCATTCACATGTTCTG 3'
Set3*S5T-A_fwd	5'ACATGTGAATGCAGAGTACAAGTGCATGCCGTGAC 3'
Set3*S5T-A_rev	5'GTCACGGCATGCACTTGTACTCTGACATTCACATGT 3'
Set3*S6new_fwd	5'GCCAAAAATAGACGCAAATTGTAACGGTCTTTCAGTGCACTCTGAAACATCTTGAAGTC 3'
Set3*S6new_rev	5'GACTTCAAGATGTTTCAGAGTGCACTGAAAGACCGTTACAATTTGCGTCTATTTTTGGC 3'
Set3_LSF_reg2_Fwd	5'GCAGAGTTCAAGTGCATGCCGATCACACTCAACGCACACACACACACGC 3'
Set3_LSF_reg3_Rev	5'GCGTGTGTGTGTGTGCGTTGAGTGTGATCGGCATGCACTTGAACTCTGC 3'
1C6Set3S5-FOR	5'GCAAATTGTAACGGTCTACAAGTGCACTCTG 3'
1C6Set3S5-REV	5'CAGAGTGCACTTGTAGACCGTTACAATTTGC 3'
1C6Set3S6-FOR	5'GGAATGCAGAGTTGTAGACCATGCCGTGAC 3'
1C6Set3S6-REV	5'CTCACGGCATGGTCTACAACTCTGCATTCC 3'

2.3.16 Purification of DNA fragments from agarose gels

The correct size of the DNA fragments produced in the PCR reactions was verified using gel electrophoresis. Depending on their size, 1 to 1.5 % 1 x TBE agarose gels were used to separate the DNA fragments (for 500 ml TBE: 54 g Tris base, 27.5 g oric acid, 20 ml 0.5 M EDTA (pH 8.0), autoclave; pH 8.3). The bands were cut out under long wavelength UV light (bulbs used were Sylvania 8W F8T5 Black Light; too short wavelength can interfere with the cloning of the DNA fragment) and transferred into pre-weighed 1.5 ml tubes. The weight of the cut out agarose piece was determined and the DNA fragment was then purified using the Qiaquick Gel Extraction Kit (Qiagen) according to the manufacturer's instructions. The final elution step was done with 30 µl of dH_2O, which yielded approximately 100 ng/µl DNA.

2.3.17 *In vitro* translation of the Tinman protein

To generate in vitro translated Tinman protein, the TNT Coupled Reticulocyte Lysate System (Promega) was used. The reactions were carried out according to the instructions provided with the TNT Kit. Based on these instructions, the reaction for the translation of the Tinman protein was set up as follows: 25 µl reticulocyte lysate, 2 µl reaction buffer, 1 µl SP6 polymerase, 0.5 µl aminoacid mix –Meth, 0.5 µl aminoacid mix –Leu (these two aminoacid mixes need to both be added to ensure all aminoacids are present; if supplied separately these aminoacids can be used to introduce a radioactive label), 1 µl RNasin, 1 µl linearized *tin* cDNA (1 µg/µl; kindly provided by M. Frasch) and 19 µl ddH_2O (DNase and RNase free, Sigma). This reaction was incubated at 30ºC for 90 min. For storage, equal amounts of glycerol were added and the reaction was kept at -20 ºC.

2.3.18 Electrophoretic Mobility Shift Assay (EMSA)

The oligos used for the EMSAs were 24-mers containing the putative Tinman binding sites and the surrounding sequence. Oligos were ordered from MWG Biotech. A list of all oligos used for the EMSAs can be found in Table 2.3.4.

Table 2.3.4: Oligos and their sequence used in the Electrophoretic Mobility Shift Assay.

Name	Sequence	Purification	Supplier
Set3_tin1_Wt(+)	TTGTAACGCACTTGAAGTGCACTCTGA	HPLC	MWG
Set3_tin1_Wt(-)	TCAGAGTGCACTTCAAGTGCGTTACAA	HPLC	MWG
Set3_tin1_Mut(+)	TTGTAACGGTCTACAAGTGCACTCTGA	HPLC	MWG
Set3_tin1_Mut(-)	TCAGAGTGCACTTGTAGACCGTTACAA	HPLC	MWG
Set3_tin2_Wt(+)	TGCAGAGTTCAAGTGCATGCCGTG	HPLC	MWG
Set3_tin2_Wt(-)	CACGGCATGCACTTGAACTCTGCA	HPLC	MWG
Set3_tin2_Mut(+)	GATCTGAGTACAGGTAAGGGCTTC	HPLC	MWG
Set3_tin2_Mut(-)	GAAGCCCTTACCTGTACTCAGATC	HPLC	MWG

500 fmol of the plus strand of each oligo were end-labelled with gamma^{32}P-ATP in a 10 µl reaction using 2 µl T4 polynucleotide kinase (20 units; NEB) and 1 µl 10 x Polynucleotide Kinase Reaction Buffer (supplied with the enzyme; NEB). This reaction was incubated at 37ºC for 30 min. The plus and the minus strand oligos were annealed in a total volume of 100 µl TNE buffer (10 mM Tris (pH 8.0), 100 mM NaCl, 1 mM EDTA) to yield a final concentration of 5 fmol/µl for the annealed oligo. To purify the labelled oligo after annealing approximately 700 µl of G25 matrix (Sephadex G25, Sigma) resuspended in TNE buffer were added to a Spin-X column. This was spun for 30 s at 14000 rpm and then the annealed and labelled probe was added to the column and spun again for 30 s. The unincorporated ATP will be retained within the matrix while the purified probe will be in the flow-through. The purified probe was then stored at -20 ºC. The competitor oligos were not labelled, but were annealed following the same protocol. However, the final concentration for the competitors was 500 fmol/µl.

The EMSAs were run on native Acrylamide gels (5 ml Accu-Gel (29:1, Arcylamide:Bis-Acrylamide; National Diagnostics), 1.25 ml 10x TBE, 43.25 ml dH_2O, 50 µl TEMED, 400 µl 10% APS (Ammonium persulfate, Sigma)).

The binding reaction was carried out in 10 µl final volume. The reaction contained 6 µl TNT lysate (protein, this includes 50 % glycerol from storage), 1 µl labelled probe, 2 µl 5x buffer (described in Kremser et al., 1999), 0.15 µl dI-dC (2 µg/µl) and 0.9 µl DTT (50mM). The 5 x binding reaction buffer (Kremser et al., 1999) contains 20 mM Tris/Cl (pH 7.5), 200 mM NaCl, 0.1 % Trition X-100, 5 mM DTE, 0.02 % BSA (w/v) and 5 % glycerol (w/v). The reaction was incubated for 30 min at room temperature.

The gel was pre-run for 45 min at 150 V, then the samples were mixed in a 1:1 ratio with 2 x SDS Gel-loading buffer (100 mM Tris-Cl (pH 6.8), 4 % (w/v) SDS (sodium dodecyl sulfate), 0.2 % (w/v) bromophenol blue, 20 % (v/v) glycerol, 200 mM ß-mercaptoethanol) and were loaded onto the gel. The gel was run at 150 V for approximately 60 min. The gel was dried and the X-Ray film (FujiFilm) was exposed overnight at -80ºC.

2.4 Bioinformatics

2.4.1 Alignment programs

The *D. melanogaster* and *D. pseuodoobscura* sequences for the PCE (=Set3) enhancer fragment were obtained from the BDGP project website (http://www.hgsc.bcm.tmc.edu/projects/drosophila/). The sequence data of the other *Drosophila* species were obtained from the NCBI trace archive, which contains all the raw sequence project sequences and which was updated every 24 hours (http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?). In all cases the *D. melanogaster* sequence was used to search the trace archive using Blast (discontiguous Basic Local Alignment Search Tool). The retrieved sequences were then edited, i.e. stretches of sequence that did not show up in the BLAST results and did not align with the *D. melanogaster* PCE sequence were deleted. The edited sequences were then analysed with various alignment programs. Initially this analysis was done before the release of the compiled and annotated sequences of the ten further *Drosophila* species. However, repeats at later time-points have shown similar results.

Initially only pairwise alignments were generated between *D. melanogaster* and *D. pseudoobscura* using VISTA (http://pipeline.lbl.gov/cgi-bin/gateway2; Dubchak et al., 2000; Mayor et al., 2000; Bray et al., 2003), lalign (http://www.ch.embnet.org/software/LALIGN_form.thml), and ClustalW (http://www.ebi.ac.uk/clustalw/#; Chenna et al., 2003). These alignment programs search for homologous regions within the submitted sequence. All these alignment programs fall into one of two categories: global or local. The aim of global alignment programs is to find the best possible fit over the complete submitted sequence, while local alignment programs aim to find discrete regions within the submitted sequence data that feature a defined level of conservation. Since local alignment tools will show multiple alignment possibilities for the same region, they are more flexible than global alignment tools. The programs used for the alignments in this study cover both categories: Vista and ClustalW produce global alignments, while lalign produces a local alignment.

With the availability of sequence data of further *Drosophila* species, it became necessary to align multiple sequences with each other. Of the programs used for the

previous alignments only ClustalW is capable of calculating multiple alignments. In order to be able to compare the result of a multiple local alignment program with the global ClustalW, alignments were also generated with PipMaker (Percent Identity Plot; Schwartz, S. et al., 2000), which is based on the BLAST Z program, a local alignment program.

2.4.2 Search for transcription factor binding sites

The search for transcription factor binding sites was superimposed onto the generated alignments of the *Drosophila* species. The complete PCE sequence was screened for potential transcription factor binding sites, but special attention was paid to the conserved regions identified by the alignment programs.

For the search for transcription factor binding sites two methods were used. The first method was based on the TRANSFAC database (www.gene-regulation.com/pub/databases.html#transfac; Wingender, E. et al., 2001). This database contains information of binding sites of eukaryotic transcription factors in the form of position weight matrices. These matrices contain the probability of a particular base at a given position within the binding site and are either based on *in vitro* experimental evidence such as Selex or were generated from compiled sites of validated binding sites. This analysis included only those sites for which there was experimental evidence.

Since the TRANSFAC database contains mostly vertebrate transcription factor binding sites with very little *Drosophila* or even insect data, these data have to be treated with care as not every factor has a conserved function and/or binding site in *Drosophila*. To ensure the inclusion of the available *Drosophila* transcription factor data, a list of potential *Drosophila* factors was compiled through a literature search (S. Elgar, pers. communication). Potential transcription factors for this list were chosen according to their spatial and temporal expression pattern during the development of the embryo. As for the TRANSFAC analysis, factors were only included into the list if there was experimental evidence for their functionality.

If there was more than one reported binding site for one transcription factor, all available data for this transcription factor were compiled and a “consensus” binding site was generated. To allow for different levels of degeneration within the consensus sequence, three different consensuses were generated; these were classified as high,

moderate and low stringency. This list of potential transcription factor binding sites was then converted into a DS Gene 'analyse subsequence' file and the *D. melanogaster* PCE sequence was analysed using the "nucleic acid subsequence analysis" function of DS Gene (Accelrys).

2.5 Software

Table 2.5.1: Software used during this study

Program	Manufacturer	Used for
Axiovision	Zeiss	Imaging software
Clustal W	Chenna et al., 2003	Algorithm for multiple sequence alignment
DS Gene Software Suite	Accelrys	Prediction of restriction and transcription factor binding sites, primer design
Excel 2008	Microsoft	Basic statistics and data storage
ImageJ	W. S. Rasband (U.S. National Institutes of Health)	Image/Video processing
lalign	William Pearson (www.ch.EMBnet.org)	Algorithm for pairwise sequence alignment
Leica Confocal Software	Leica	Imaging software for the confocal microscope
MatInspector	Genomatix	Search for transcription factor binding sites
Minitab	Minitab, Inc.	Statistical analysis
MultiPipmaker	Schwartz et al., 2000	Algorithm for multiple sequence alignment
Photoshop CS2/CS 6	Adobe	Image processing and creating figures
Primer3		Primer design
QuickChange primer	Stratagene	Primer design for site-directed mutagenesis
Vista	Mayor et al., 2000	Algorithm for pairwise sequence alignment
VLC	VideoLan Organisation	Video displaying
Word 2008	Microsoft	Word processing
Vista Browser 2.0	Dubchak et al., 2009	Whole-genome alignment

2.6 Commercial Kits used in this study

Table 2.6.1: Commercially available "kits" used in this study.

Name	Producer	notes
pGEM-T®	Promega	Cloning
Qiaquick® Filter Plasmid Midi Kit	Qiagen	Preparation of plasmid DNA for PCR and cloning
Qiaquick® Gel Extraction Kit	Qiagen	Purification of DNA from agarose gels
Quick Change® Site-directed Mutagenesis Kit	Stratagene	Used to introduce site-directed mutagenesis into the PCE construct
TnT Coupled Reticulocyte Lysate Systems®	Promega	Used to translate Tinman protein for the EMSA
TSA Cy3® and Cy5® kits	Perkin Elmer	Fluorescent labelling of *RNA in situs*
TSA® Biotin Kit	Perkin Elmer	Signal amplification in immunostains
Vectastain® ABC-Kit Vector	Vector Laboratories	Signal amplification in immunostains

Chapter 3

Him expression and phenotype

3.1 Introduction

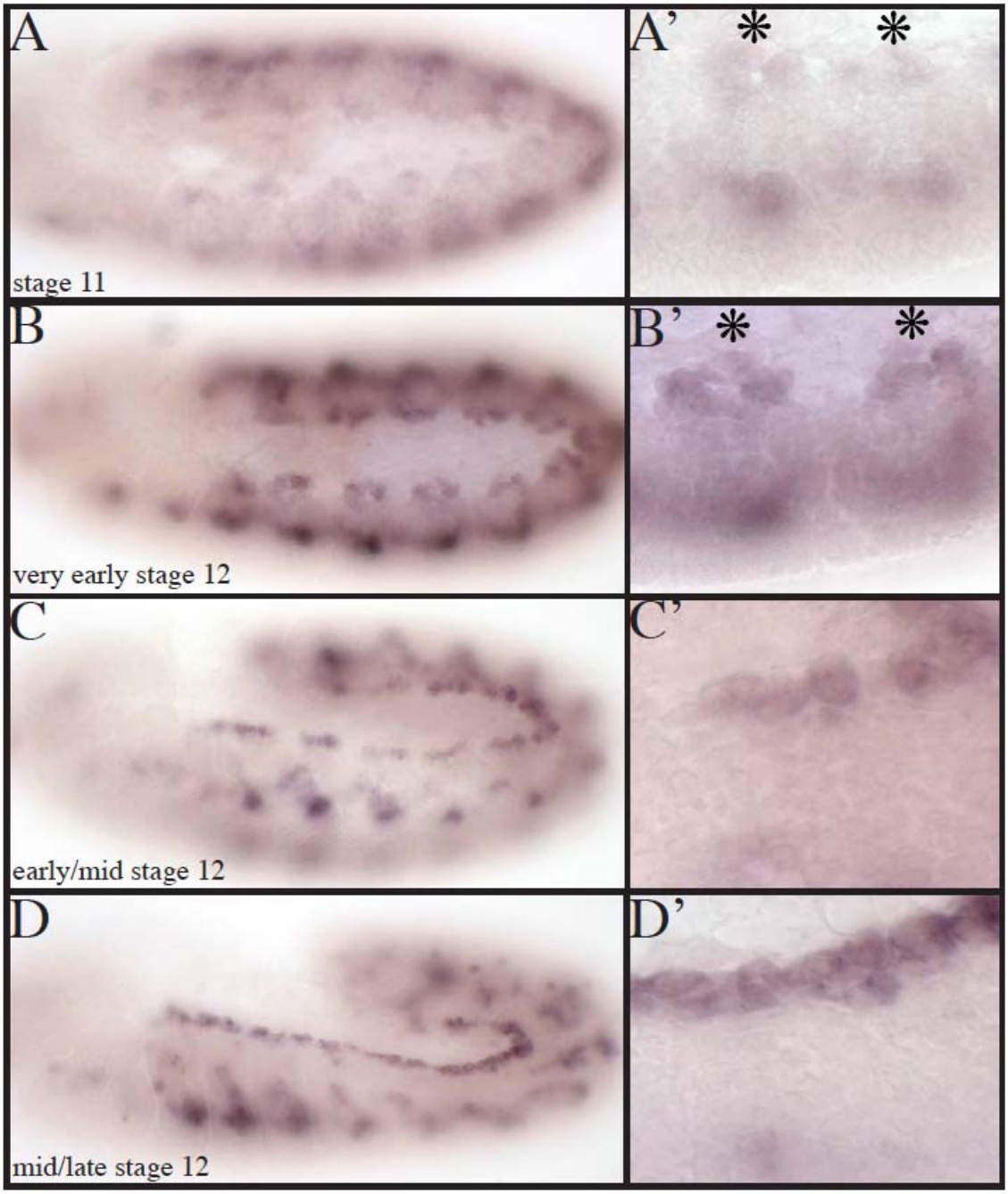
Knowledge of the particular expression pattern of a gene is a very useful tool when analysing its regulation and function. While adjacent cells might provide key regulatory signals, only genes that are expressed in the same tissue at the same time point during development are potential direct positive or negative transcriptional regulators. Knowing when and where a gene is expressed also allows one to develop ideas as to its function.

The technique of *In-Situ*-Hybridisation enables one to visualize the distribution of an RNA transcript of a gene of interest. This establishes where in the embryo and at which stages a gene is transcribed. Where available, the distribution of the protein encoded by the gene can also be analysed using an antibody specific to this gene/protein.

3.2 *Him* expression in the heart

3.2.1 *Him* expression in the heart of the developing embryo

The exact expression pattern of *Him* in the dorsal vessel during development has not yet been described. The *Him* RNA transcript is initially expressed in the whole mesoderm from stage 9 onwards (Liotta et al., 2007). Thus, it is expressed in the cells that become the heart precursors during stage 11 (Bodmer et al., 1990; Wu et al., 1995; see Figure 3.2.1). Expression in all heart precursors is maintained until stage 12. By stage 13, *Him* expression is limited to pericardial cells of the heart (Figure 3.2.1 C and C', Figure 3.2.2 G and H).


A
A'
stage 11
B
B'
very early stage 12
C
C'
early/mid stage 12
D
D'
mid/late stage 12

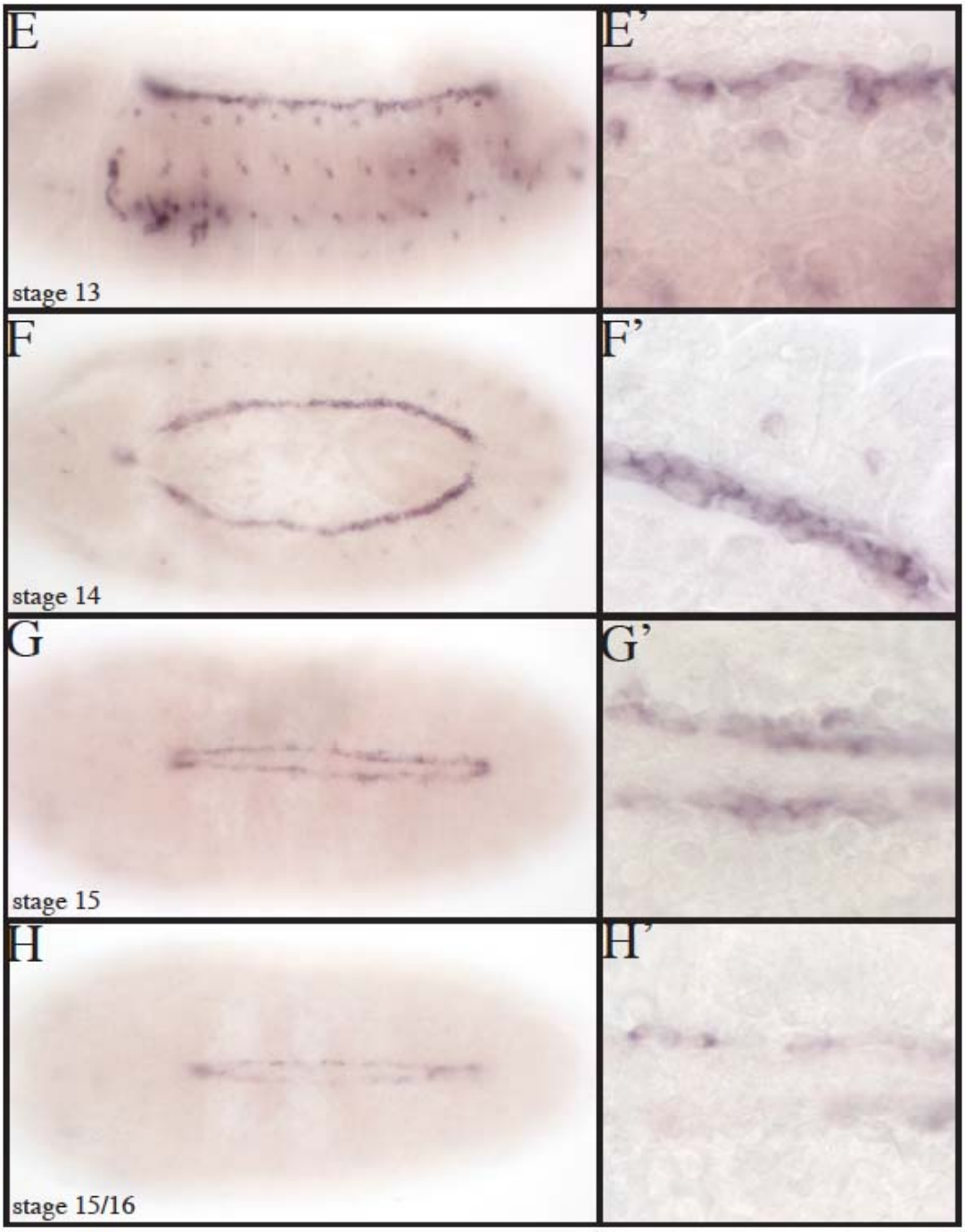


Figure 3.2.1: Expression of the *Him* transcript during *Drosophila* heart development.
Him RNA *in situ* hybridisation on *OR* embryos. Images A-E show a lateral view of the embryo, images F to H show a dorsal view. Images A' - H' show the same embryo as in A - H at a higher magnification of the heart. The stars indicate the top of the mesodermal crescent in two hemisegments.

Initially *Him* expression in the dorsal mesoderm is weaker than its expression in the somatic mesoderm, however, the expression in the dorsal mesoderm increases during stage 11 and early stage 12. At the same time, *Him* expression decreases in the somatic mesoderm. The progenitors of the future heart cells are specified during late stage 10 and early 11 at the very peak of each mesodermal crescent. The precursors for the pericardial cells are located at the peak of each crescent (stars in Figure 3.2.1 A' and B'). As can be seen in the images 3.2.1 A' and B' the cells at the peak of the crescent initially express less *Him* than the ones slightly more lateral. The peaks of the developing mesoderm correspond to the areas of *wingless* expression in the overlaying ectoderm. These observations indicate that the precursors of the Eve-positive heart cells are likely to be the last heart cells to start expressing *Him*. If this is due to an influence of *wingless* is currently still an open question that needs answering.

From mid-stage 12 onwards *Him* expression in the dorsal mesoderm is limited to the pericardial cells of the heart. *Him* expression is then maintained at the same level until stage 15 (see Figure 3.2.1 G and G'). At this point the expression of the *Him* transcript in the heart becomes weaker. During the following last two stages of embryonic development *Him* expression is lost from all pericardial cells. By stage 16, *Drosophila* embryos show a very weak and patchy *Him* transcript expression pattern in the dorsal vessel and no *Him* expression is visible by stage 17.

I used co-localisation with the established cell marker Zfh-1 (Lai et al., 1991) to determine whether *Him* is expressed in all known pericardial cells (arrows in Figure 3.2.2). Zfh-1 has been established as a marker for pericardial cells; it is initially expressed in all known pericardial cells, the two Eve-positive pericardial per hemi-segment cells lose *zfh-1* expression during the later stages of embryonic development (Su et al., 1999; Johnson et al., 2003). Zfh-1 is labelled in red in Figure 3.2.2. The *Him* transcript (labelled in blue in Figure 3.2.2) is also expressed in the Zhf-1-expressing cells. The *Him* transcript is mainly located in the cytoplasm of the cells and around the nuclear localisation of the Zhf-1 protein (Figure 3.2.2 H), however the purple colour of the overlay indicates co-localisation of *Him* and Zfh-1 even within individual cells. The so-marked cell populations are identical, thus, one can conclude that *Him* is expressed in all pericardial cells that also express Zfh-1.

Two of the ten pericardial cells per hemisegment lose Zfh-1 expression in late embryogenesis, these two cells also express the cardiac marker Tinman (Frasch and

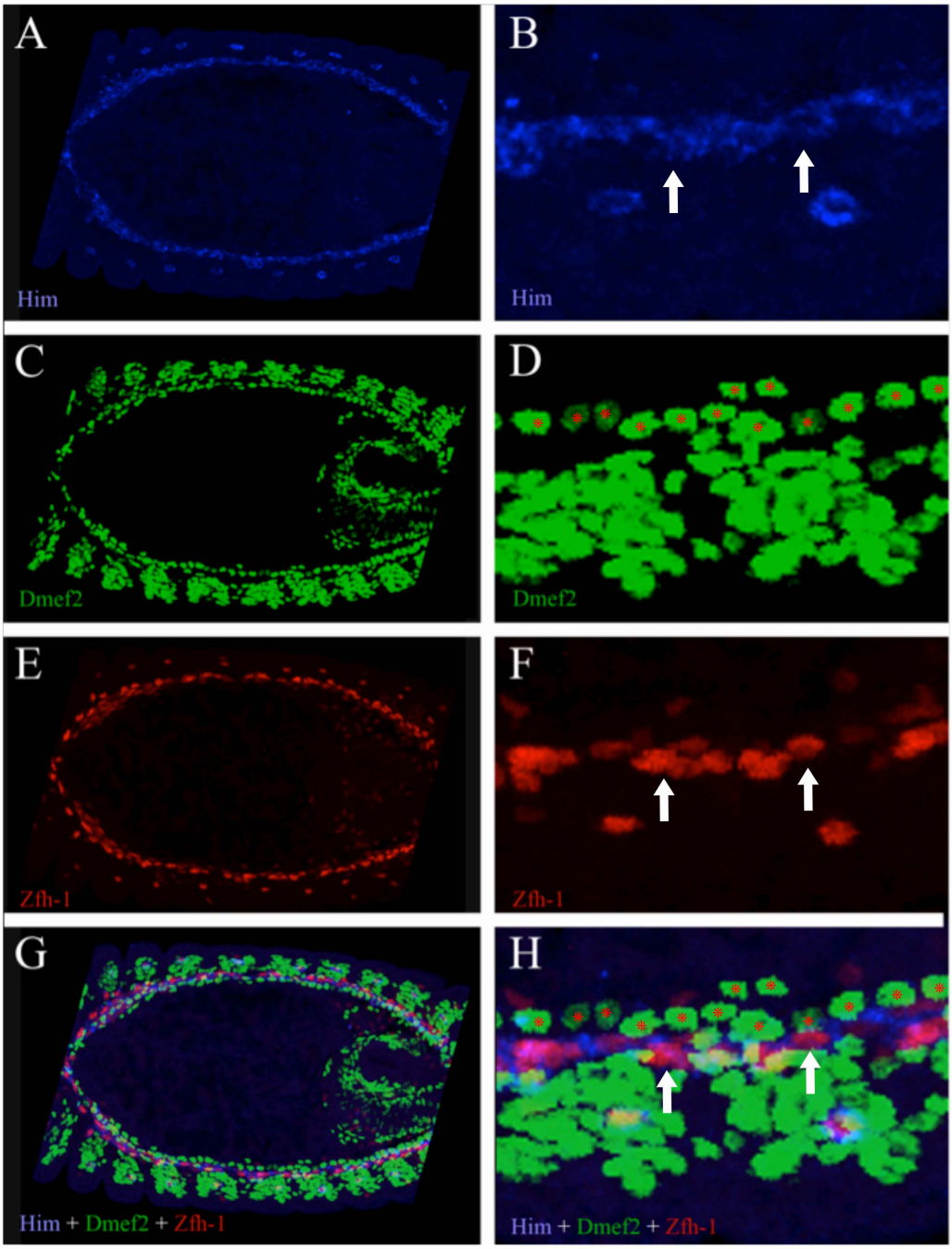


Figure 3.2.2: *Him* is expressed in all pericardial cells and not in the cardioblasts.
Him RNA fluorescent in situ hybridisation on OR embryos. Images A, C, E, G show a dorsal view of a stage 13 embryo. Images B, D, F, H show the same embryo at a higher magnification. The arrows indicate areas where the *Him* transcript and Zfh-1 are coexpressed and the stars mark the Mef2-positive cardioblasts.
A, B *Him RNA in situ* (blue); **C, D** *Dmef2* antibody (green); **E, F** *Zfh-1* antibody (red); **G, H** overlay of all images

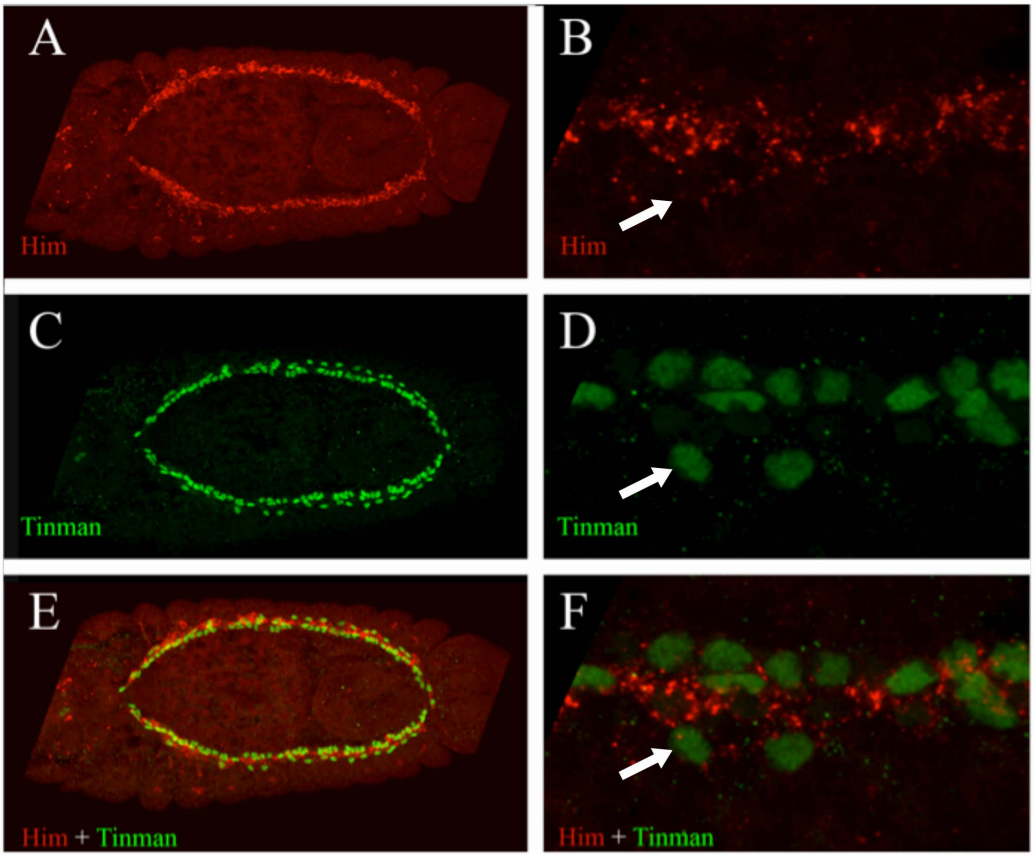


Figure 3.2.3: Him expression in the *Tin-positive* pericardial cells.
Him RNA in situ hybridisation and Tinman antibody stain on OR Embryos. Images A, C, E show a dorsal view of a stage 13 embryo. Images B, D, F show the same embryo at a higher magnification. The arrows point towards the Eve-positive pericardial cells.
A, B *Him RNA in situ* (red); **C, D** Tinman antibody (green); **E, F** overlay of all images.

Levine, 1987). Figure 3.2.3 shows the co-localisation of *Him* transcript and Tinman, which is expressed in six of the ten pericardial cells, including the Eve-positive pericardial cells (arrows in Figure 3.2.3).

Within the cardiac mesoderm of *Drosophila* embryos *Him* expression is limited to the pericardial cells as shown in Figure 3.2.2 G and H. Mef2 is a protein expressed in contractile muscle cells and not in the pericardial cells of the heart (Taylor et al., 1995; Yin et al., 1997; Bour et al., 1995 and Lilly et al., 1995) and *Mef2* and *Him* expression do not overlap in wild-type embryos past embryonic stage 12/13 (the stars in Figure 3.2.2 H indicate the Mef2-positive cardioblasts). I was unable to take an image of an earlier embryo that would allow for the study of *Him* and Mef2 co-localisation in the early and in some cases shared progenitors of the pericardial cells and cardioblasts.

3.2.2 *Him* expression in the heart during the three larval stages

As described previously in the introduction, the pericardial cells continue to undergo differentiation during larval development. Especially noteworthy is the decrease in the number of pericardial cells between the first and third instar from around 120 Zfh-1-positive pericardial cells in a stage 17 embryo (Alvarez et al., 2003) to about 40-50 pericardial cells by the third instar (Rizki et al., 1978, Sellin et al., 2006). The loss of pericardial cells occurs mainly between 36 and 40 hours AEL (after egg laying). This loss of some pericardial cells is accompanied by the increase in the size of some hand-GFP-positive nuclei through endoreduplication in the remaining hand-GFP-positive pericardial cells (Kambysellis & Wheeler, 1972; Sellin et al., 2006, see chapter 1). It is possible that the removal of Him from the pericardial cells prior to these steps is necessary to "allow" the further differentiation of the pericardial cells. Prior to my work, nothing was known about the expression of *Him* in the heart during larval life. I have used several reporter and Gal4-lines to identify if and where *Him* is expressed in the heart after embryonic development. For this analysis, I immobilized larvae at different time points (first instar (26-28h AEL), second instar (52-54h AEL) and third instar (100-102h AEL)) and evaluated the live animals for the fluorescence of the GFP reporter protein under the Zeiss microscope.

The *p43 (Eco/Xho)* construct was made in this lab previous to my work and contains the sequence between the *EcoRI* and *XhoI* restriction sites upstream of the *Him* gene linked to a minimal promoter and GFP reporter (S. McConnell & M. Taylor, unpublished data). The EcoRI/Xho fragment is about 2.8 kb in length and does not contain the 1 kb of sequence immediately upstream of the transcriptional start site of *Him*. Its larval expression pattern is shown in Figure 3.2.4. During the first instar, the GFP expression observed during embryogenesis is still visible in all of these cells (arrow in Figure 3.2.4), but I do not observe any GFP expression in the heart during the second or third instar. Most likely the GFP expression I observed for the first instar is due to the persistence of the GFP protein expressed during embryogenesis. During the second and third instar, I observe a group of cells that do express GFP. These cells are located in close proximity to the heart. The GFP expression in these cells is initially weak in the second instar, but increases in strength by the third instar (see arrows in Figure 3.2.4 C and D). The cells are located towards the anterior of the animal but posterior to the lymph gland and the expression pattern is not similar to that described for the hand-GFP reporter (Sellin et al., 2006). The overall level of GFP expression in these cells is not very strong, which makes identifying them difficult. A step towards this would be to create larvae that carry both the p43 and the *hand* reporter constructs expressing GFP and RFP, respectively, to clearly identify any overlaps in expression and to verify my assumption that these cells are indeed separate from the pericardial cells expressing the *hand* reporter construct.

In addition to this construct, our lab has generated two different *Him-Gal4* driver lines, *Him-Gal4 L3-5* and *Him-Gal4 L4-1*, which I also tested for larval expression using a *UAS-GFP* construct (Dutta et al., 2002). The construct I used throughout the remainder of my thesis, *Him-Gal4 L3-5*, is known to reproduce the embryonic expression pattern of *Him* (D. Liotta, PhD thesis) but so far any expression in larval stages has not been determined. This Gal4-construct contains the same sequence as the *p43 (Eco/Xho)* construct. My tests revealed that it does not drive any GFP expression during the later larval stages. Figure 3.2.5 D and F show the lack of GFP expression during the second and third instar in animals carrying the *Him-Gal4 L3-5* construct. There is only weak expression of GFP during the first instar (Figure 3.2.5 B). The observed levels of GFP expression for the *p43 (Eco/Xho)* construct were stronger (Figure 3.2.4.B). As both constructs contain the same sequence, I conclude that the reduction of the larval GFP expression in the *Him-Gal4 L3-5* construct is due to the

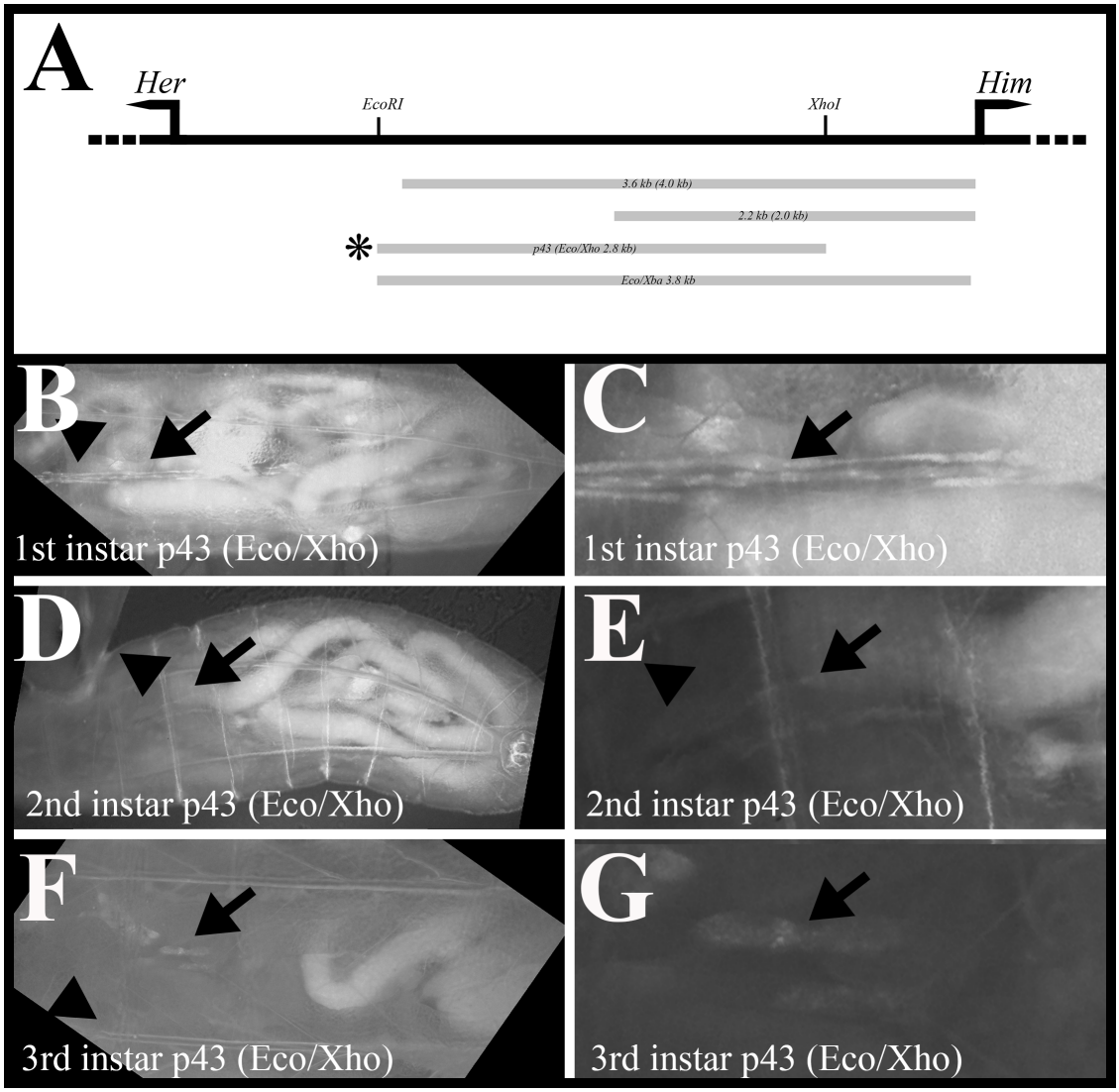


Figure 3.2.4 Larval expression pattern of the *p43 (Eco/Xho 2.8kb)* construct.
Expression pattern of the GFP reporter gene used in the *p43 (Eco/Xho 2.8 kb)* enhancer construct during all three larval instars. The arrows point to the area of observed GFP expression and the arrowhead indicates the position of the lymph gland. C, E and G are magnifications of the images to the left.
A overview of the available constructs to study *Him* expression past embryogenesis. The star indicates the *p43 (Eco/Xho 2.8 kb)* construct shown in B to G; **B, C** 1st instar larvae transgenic for the *p43 (Eco/Xho 2.8kb)* construct; **D, E** 2nd instar larvae carrying the *p43 (Eco/Xho 2.8kb)* construct; **F,G** 3rd instar larvae carrying the same GFP reporter construct.

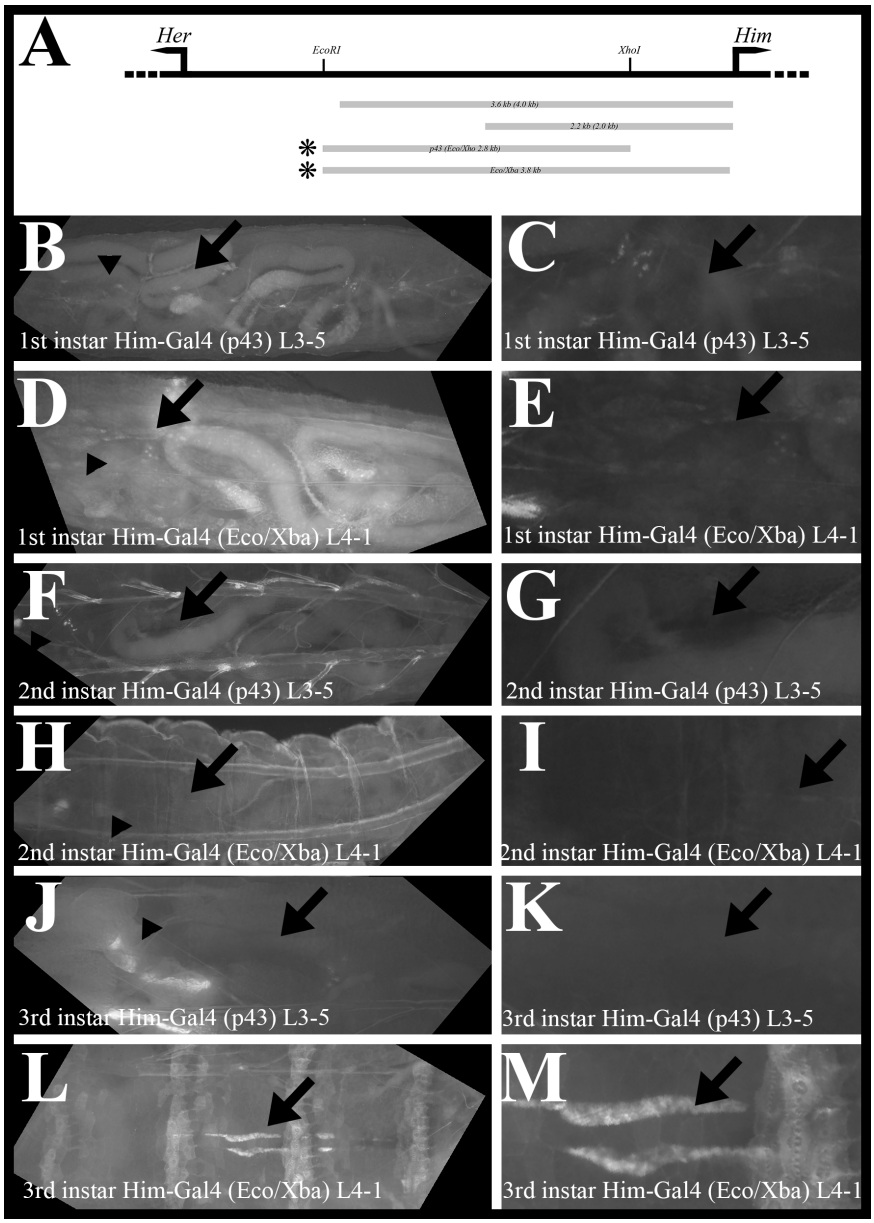


Figure 3.2.5 Expression pattern of the available *Him-Gal4* driver lines.
Two different *Him-Gal4* lines (*L3-5* and *L4-1*) were used to drive GFP expression to verify their expression pattern during larval development. The arrows point to the areas where GFP expression observed in the *Him-Gal4 L4-1* larvae and the arrowhead indicates the position of the lymph gland.. C, E, G, I, K and M show magnifications of the areas the arrow in B, D, F, H, J and L point to.
A overview of all available constructs to study the larval expression pattern of *Him* (the L3-5 line contains the same sequence as the *p43* construct). The star indicates the constructs used in this figure; **B, C** 1st instar larvae shows only very weak GFP expression in the heart if the *Him-Gal4 L3-5* line is used; **D, E** the *Him-Gal4 L4-1* line also only shows weak GFP expression in the heart during the 1st instar; **F, G** no GFP expression in the 2nd instar larvae with the *Him-Gal4 L3-5* construct; **H, I** very weak GFP expression in cells along the anterior part of the aorta if the *Him-Gal4 L4-1* line is used; **J, K** I observed no GFP expression during the 3rd instar if the *Him-Gal4 L3-5* line drives the *UAS*-construct; **L, M** strong GFP expression in cells along the anterior part of the heart if the *Him-Gal4 L4-1* construct is used.

variability of the UAS/Gal4 system and the different insertion points of the constructs into the genome. It is also possible that the reporter construct integrated into a part of the genome that is heavily repressed during larval development. This could explain the difference in GFP expression levels between the *p43 (Eco/Xho)* and the *Him-Gal4 L3-5* constructs. However, I only tested one transgenic line of the *Him-Gal4* construct (L3-5; chosen for its strength of expression in the embryo) for larval expression and it is possible that other lines have a stronger/different effect. In addition to this, the Gal4/UAS-system is temperature dependent and if these experiments were repeated at a higher temperature (these experiments were all done at 25°C), one might be able to observe larval GFP expression in the *Him-Gal4 L3-5* construct.

More recently, a second reporter construct was made in the lab, *Him-Gal4 L4-1*, which contains the ca. 1 kb part of the *5'* upstream sequence of *Him* that is missing in the *Him-Gal4 L3-5* construct in addition to the p43 (Eco/Xho) sequence (D. Hancock, PhD thesis). The larval expression pattern of the *Him-Gal4 L4-1* construct is shown in images 3.2.5 D, E, H, I, L and M. This construct drives expression similar (in strength) to that observed for the *p43 (Eco/Xho)* construct in the second instar larvae. In the third instar, GFP expression is strong in a group of cells close to the heart but there is no expression within the actual heart cells (see arrow in Figure 3.2.5 M). The GFP-expressing cells are in the same position as the GFP-positive cells that I saw with the *p43 (Eco/Xho)* reporter construct. Under these conditions the cells are easier to observe than before but I am still not able to identify them by their position alone. However, these cells appear to have many, relatively small nuclei in close proximity to each other and based on my observation I hypothesize that these GFP-positive nuclei are part of a multi-nucleated cell or tissue.

In addition to the constructs generated in our lab, there are two further reporter constructs for the 5' region of *Him* available. These *Him* reporter constructs were published by Rebeiz et al. and are a 2.2 kb and a 3.6 kb enhancer fragment (in their paper these constructs are called a 2.0kb and a 4.0kb enhancer/reporter construct, respectively), located upstream of the ORF (open reading frame) of *Him*, which have been cloned in front of a GFP reporter (Rebeiz et al., 2002).

The results for this analysis of the 2.0 kb construct is shown in Figure 3.2.6. During the first instar larvae (see Figure 3.2.6 B) the GFP expression pattern is identical to that I observed in stage 16/17 embryos. This indicates that the 2.0 kb construct is at

least active until late in embryogenesis as the GFP protein will persist for a while after its expression ceases. During later larval stages (2nd and 3rd instar; Figure 3.2.6 C and D) this expression pattern is not visible anymore. This result corresponds to the expression pattern I observed with the constructs generated in our lab. Similar to the reporter constructs established in our lab, I have also detected GFP expression in a group of cells very close to the heart tube during these stages (see arrows in Figure 3.2.6 C and D). This expression is initially relatively weak during the second instar larvae but then quite pronounced in the third instar animals. These GFP expressing cells are in the same location as those previously described for the other Him-GFP reporter constructs. The analysis of the expression pattern of the endogenous Him transcript is hampered by the larval cuticle and will require dissections of the animal to allow the probe to penetrate into the tissues.

Based on these results, *Him* expression in the heart is limited to the embryonic stages. The GFP expression in the heart that I observed during the first instar is most likely due to the stability of the GFP protein. The identity of the GFP-positive cells observed in the majority of constructs in the anterior to middle part of the larval body during the second and third instar is currently not known, however, my results point towards the importance of the 1 kb region immediately upstream of the ORF for *Him* expression in these unidentified cells, as I have only observed GFP expression in these cells when the constructs included these 1kb of DNA.

3.3 *Him* null phenotype in the heart

As shown above, *Him* is expressed in all pericardial cells of the embryonic heart. The *Him* transcript can be visualized until the late stages of embryogenesis before possibly being re-expressed in a set of un-identified cells in the second and third instar larvae (see above). The pericardial cells of the *Drosophila* heart are non-myogenic cells thought to be involved in ultra-filtration of the haemolymph (Rizki, 1978; Rugendorff et al., 1994; Das et al., 2008; Weavers et al., 2009). Previously this lab has shown, that, in the somatic mesoderm, Him has a role in down-regulating Mef2 activity and that Him may contribute to keeping the Adult Muscle Precursors (AMPs) in a committed but undifferentiated state during embryogenesis (Liotta et al., 2007). Only recently, it has

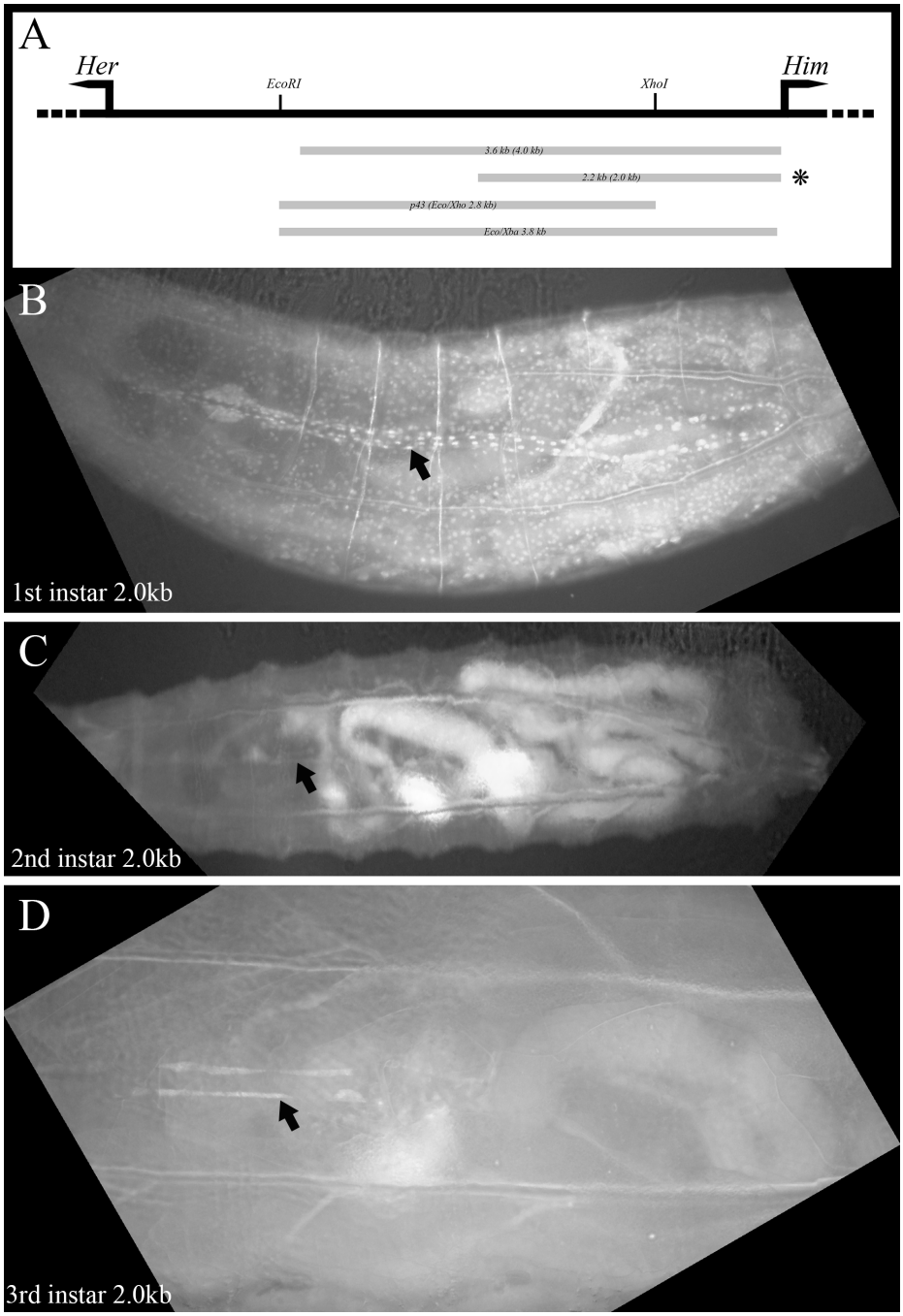


Figure 3.2.6: Larval expression pattern of the *2.0 kb* construct of Rebeiz et al., 2002.

Expression pattern of the GFP reporter gene used in the *2.0 kb* enhancer construct (Rebeiz et al., 2002) during all three larval instars. The arrows point to the areas of observed GFP expression.

A overview of the available constructs to study *Him* expression in larvae. The star indicates the construct used in this figure; **B** 1st instar larvae transgenic for the *2.0 kb* construct shows GFP expression in all heart cells; **C** 2nd instar larvae carrying the *2.0 kb* construct only show GFP expression in unidentified cells close to the heart; **D** 3rd instar larvae carrying the same GFP reporter construct also show GFP expression in the unidentified cells.

been shown that two sub-populations of the pericardial cells contribute to the muscle tissues of the wing hearts and ventral heart muscle during the larval and pupal development (Tögel et al., 2008; Shah et al., 2011). This poses the question if *Him* plays a similar role in the pericardial cells as it does in the embryonic somatic mesoderm. If *Him* is needed to establish a non-myogenic identity (potentially by regulating Mef2 activity) within the pericardial cells during their embryonic development, it is likely that the mis-expression of *Him* in closely associated cells like the cardioblasts or loss of *Him* in the pericardial cells will have an effect on their number and possibly on their cell fate.

There were three lines of *Him* null mutants available to me. One was made in this lab and the other two null lines were created by Zhe Han (D. Hancock, PhD thesis, Han, Z. pers. communication). The *Him 52* null mutant was established in our lab by Dan Hancock. He targeted an area of 98 kb containing six genes in total (*Frq1, andorra, Frq2, Her, Him* and *CG33639*) using P-element mediated recombination (D. Hancock, PhD thesis). The other two *Him* null mutants, *Him 74* and *Him 195*, were created by ends-out homologous recombination by Dr. Zhe Han (in the lab of E. Olson). All three *Him* null mutant lines are negative for *Him* transcript expression and homozygous viable, indicating that *Him* is not essential for survival of the fly. This fact allows me to analyse both *Drosophila* embryos and larvae for a potential *Him* phenotype in the heart during development.

Over-expression of *Him* in the developing dorsal vessel was accomplished using the Gal4/UAS-system (Brand & Perrimon, 1993), which allows for the targeted expression of a gene of interest at a specific time and place. The Gal4-drivers I used for these experiments are the *hand-Gal4*, *Mef2-Gal4* and the *TinCΔ4-Gal4* drivers (Bour et al., 1995; Lo & Frasch, 2001; Sellin et al., 2006) as these allow mis-expression in the whole heart (*hand-Gal4*) or specifically within the cardioblasts (*Mef2*- and *TinCΔ4-Gal4*).

3.3.1 *Him* null phenotype in the heart of the embryo

In order to analyse any potential phenotype in the developing embryonic hearts of the *Him* null mutant lines, I used the expression of several known markers for

pericardial cells and cardioblasts. Zfh-1 is initially expressed in all pericardial cells and subsequently lost in the two Eve-positive pericardial cells per hemisegment during stage 15 (Lai et al., 1991; Su et al., 1999; Johnson et al., 2003), while Odd-skipped (Odd) is expressed in a row of 4 pericardial cells per hemisegment from stage 13 onwards (Ward and Coulter, 2000; Ward and Skeath, 2000) and Even-skipped (Eve) is expressed in two dorsally positioned pericardial cells per hemisegment (Frasch and Levine, 1987).

To visualize the cardioblast cell population, I have used ß3-tubulin. It is expressed in four of the six cardioblasts per hemisegments (Gasch et al., 1988). I have chosen β3-tubulin as my standard cardioblast marker as it is known to be unaffected in *Mef2* mutants and should thus be expressed in any present cardioblasts. According to my hypothesis Mef2 function is likely affected by altering the *Him* expression pattern. As Myosin expression is reduced in *Mef2* mutant embryos (Ranganayakulu et al., 1995), I have included a Myosin stain in the section discussing the *Him* over-expression.

As explained in the introduction (Chapter 1) several of the pericardial cells are generated through an asymmetric division, which generates a pericardial cell as one daughter cell and a myogenic cell (cardioblast or somatic muscle cell) as the second daughter cell (Carmena et al., 1998; Park et al., 1998, Ward and Skeath, 2000). Thus the relationship between the non-myogenic pericardial cells and the contractile muscle cells of the heart and the dorsal somatic mesoderm is very close. It has been shown that loss or mis-expression of a key regulator is capable of increasing one cell population at the cost of another. For example, Hedgehog (Hh) is necessary for the specification of the correct number of Eve- and Ladybird- (Lbe) positive cardiac cells (Liu et al, 2006), while Pannier (Pnr) generally promotes all heart cells and Pointed (Pnt) promotes the specification of pericardial cells at the expense of the cardioblasts (Alvarez et al., 2003).

The influence *Him* has on the specification and differentiation of the heart has not been previously analysed. In order to investigate this I have assayed the expression of the above mentioned marker genes in the available *Him* deficiency lines. The lines I used were *Him 52; If/cyo; handCGFP3.1*, *Him 74; handCGFP2.3, TM2/MKRS* and *Him 195; +/+; +/+*. The hand-GFP construct was introduced in the *Him* lines to allow analysis of the same line in both embryos and larvae. However, I was not able to introduce the hand-GFP reporter fragment into the *Him 195* line.

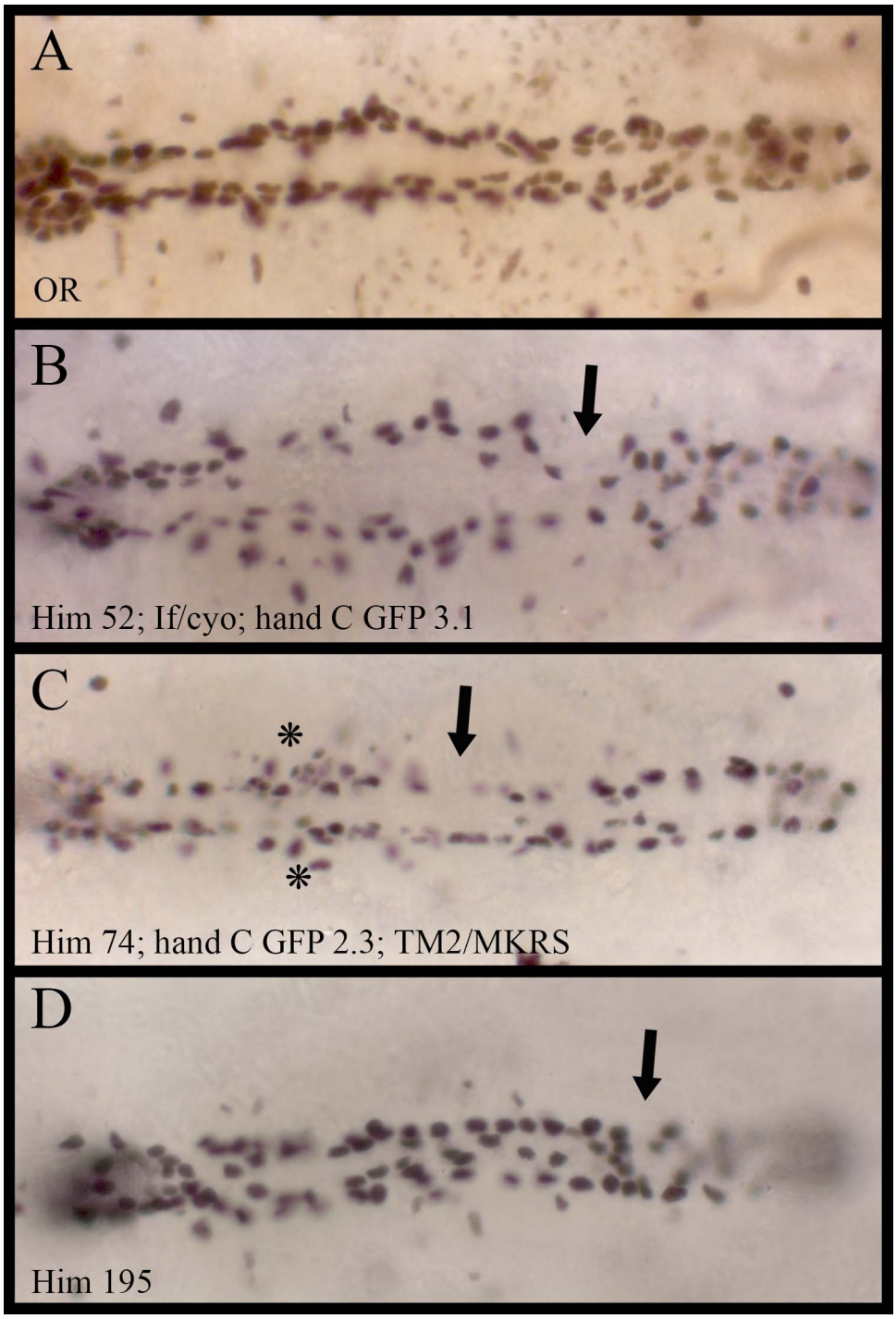


Figure 3.3.1: Zfh-1 expression in *Him* null mutants.
Stage 16/17 embryos were labelled for the distribution of the Zfh-1 protein. The arrows indicate the gaps within the *zfh-1* expression pattern and the stars indicate areas of clustered Zfh-1 expressing cells.
A *OR* wild-type embryo; **B** Zfh-1 expression in the *Him 52* null mutant line; **C** Zfh-1 expression in the *Him 74* null mutant line; **D** Zfh-1 expression in the *Him 195* null mutant line.

The first marker whose expression I tested was Zfh-1. It is expressed in all pericardial cells up until stage fifteen and after that in eight out of the ten pericardial cells per hemisegment (Lai et al., 1991; Su et al., 1999). As can be seen in Figure 3.3.1 Zfh-1 is expressed in the pericardial cells of all three *Him* deficient fly lines. However, when compared to the wild-type expression pattern it can be seen that the *Him* null mutant embryos have fewer Zfh-1-positive pericardial cells. As this loss of Zfh-1-positive pericardial cells is spread evenly along the anterior-posterior axis, it becomes only visible upon close inspection of multiple embryos (for all these assays n = 20). This even loss of cells along the anterior-posterior axis also indicates that the same cells of each hemisegment are affected. However, just based on the Zfh-1 expression pattern I am not able to identify which sub-sets of pericardial cells are affected. This loss of pericardial cells could either be caused by a lack of the Zfh-1-positive pericardial cells or by the loss of Zfh-1 expression in a subset of the pericardial cells. In about half of all three *Him* null embryos I have observed small gaps that are never larger than one hemisegment (arrows in Figure 3.3.1) and most embryos do not show more than one gap at a time. The position of this gap varies along nearly the complete A/P axis of the heart with the exception of the segments containing the lymph gland and the immediately adjacent segment. Approximately 50 % of the *Him 74* embryos shows an accumulation of the pericardial cells around the attachment sites of the alary muscles, which can be identified in this line (see stars Figure 3.3.1). Neither the accumulation of pericardial cells nor the alary muscles were visible in the other two alleles. Counting the exact number of Zfh-1-positive pericardial cells has proven unreliable. The three-dimensional arrangement of the pericardial cells around the inner heart tube causes cells to lie "on top" of each other when viewed dorsally, it thus becomes difficult to count every cell present.

In contrast, numerical analysis of the pericardial cell markers Odd and Eve is reliable, as they are only expressed in a subset of pericardial cells. Together Odd and Eve mark six of the ten pericardial cells per hemisegment. These two groups of pericardial cells also contain those cells that share precursors with either the cardioblasts or the dorsal somatic muscles. The four remaining cells express Tinman. Tinman is however also expressed by the Eve-positive pericardial cells and the cardioblasts which are positioned immediately dorsal to the Tin-positive pericardial cells, thus this analysis would suffer from the same problems as that of the Zfh-1 expression in that counting only the pericardial cells would be extremely difficult.

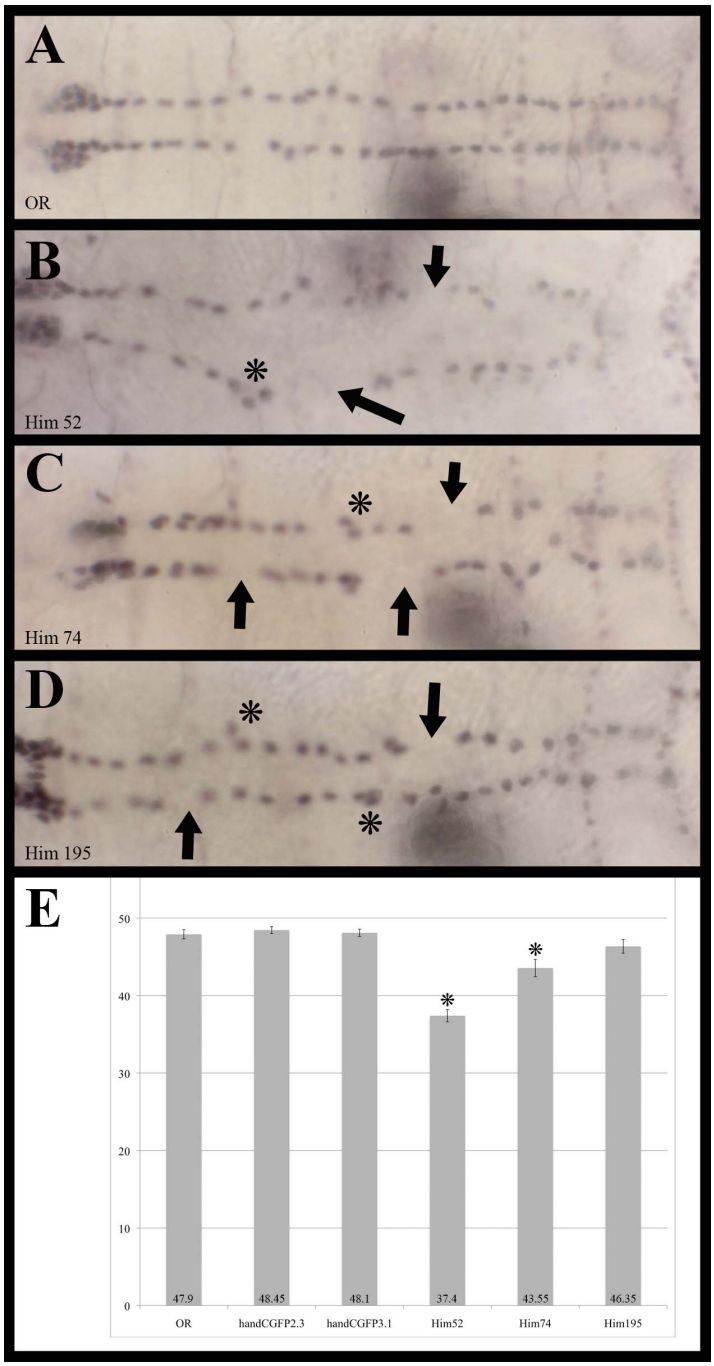


Figure 3.3.2: Expression pattern of Odd in *Him* null mutant embryos.
Stage 16/17 embryos were labelled for the Odd protein and the number of Odd-positive pericardial cells was counted. The arrows indicate the gaps within the *odd* expression pattern and the stars indicate areas of clustered Odd expressing cells.
A Wild-type embryo; **B** Odd expression in a *Him 52* null mutant embryo; **C** the Odd expression pattern in a *Him 74* null mutant embryo; **D** Odd expression pattern in a *Him 195* embryo; **E** Histogram showing the obtained numbers for the Odd-positive pericardial cells in all three controls (*OR, handCGFP2.3, handCGFP3.1*) and the *Him* null lines. The error bars indicate the standard error and stars indicate a statistically significant difference in regard to the wild-type.

Odd is expressed in four pericardial cells per hemisegment and these four cells form a line nearly parallel to the inner heart tube, reminiscent of a "pearls on a string" arrangement (Ward and Coulter, 2000; Ward and Skeath, 2000; see Figure 3.3.2 A). Figure 3.3.2 A-D shows the expression pattern of Odd and Figure 3.3.2 E shows the numbers of Odd-positive pericardial cells in all three lines lacking *Him* and wild-type embryos. All three lines null-mutant for *Him* show a trend towards a loss of Odd-positive pericardial cells. However, only two of the three lines (Him 52 and Him 74) have statistically significant fewer Odd-positive pericardial cells. The observed loss of Odd-positive pericardial cells in the Him 195 null embryos is not statistically significant at the chosen confidence interval of 0.05. All three lines show gaps in some areas of the arrangement of Odd-positive pericardial cells (see arrows in Figure 3.3.2) while in other areas the cells aggregate closer together than they do in wild-type embryos (see stars in Figure 3.3.2). The size of the gaps within the row of Odd-positive pericardial cells does not vary much between the three different lines. Together these observations indicate that loss of *Him* function results in a disruption of the orderly arrangement of the pericardial cells. Whether this organisational phenotype is due to an alteration in cell fate or due to a defect in the even distribution of the Odd-positive pericardial cells is a currently unanswered question.

Eve is expressed in the two pericardial cells per hemisegment that are located dorsally to the contractile heart tube. Figure 3.3.3 A-D shows the expression of Eve in the pericardial cells of wild-type embryos and embryos null mutant for *Him*. Figure 3.3.3 E is a graph of the number of Eve-positive pericardial cells counted in each line. Similar to the behaviour of the Odd-positive pericardial cells, there is a decrease in the number of these cells in the embryos lacking *Him*; all three *Him* null mutant lines lose a statistically significant number of Eve-positive pericardial cells. For all three lines, I have observed one or both of the cells per hemisegment missing and the expression of Eve in the dorsal muscle appears reduced in all three lines if judged by the intensity of the staining. This is interesting as the dorsal muscles DA1 and DO2 (the latter of which also expresses Eve) and the Eve-positive pericardial cells share a common lineage (Carmena et al., 1998; Park et al., 1998). Because of this, it is possible that these two dorsal muscles are more affected by the loss of *Him* than any other somatic muscles.

Indeed, previous analysis for the presence of the somatic muscles in segments A2 to A4 in stage 17 embryos by Dan Hancock has revealed that the DA1 muscle is one of the most frequently (27.3 % in the *Him 52* line) of all muscles affected by duplication

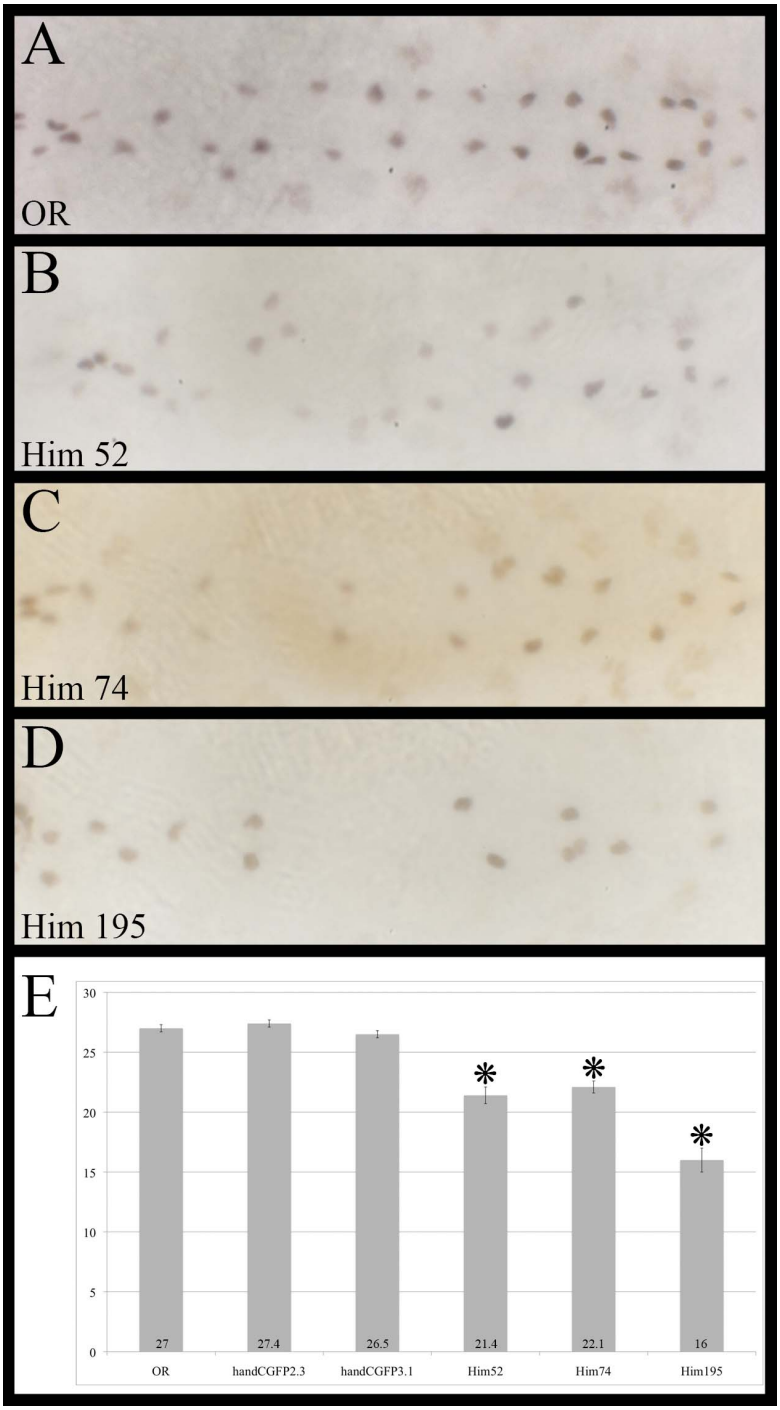


Figure 3.3.3: Expression pattern of Eve in *Him* null mutant embryos.
Stage 16/17 embryos were labelled for the Eve protein and the number of Eve-positive pericardial cells was counted.
A Wild-type embryo; **B** Eve expression in a *Him 52* null mutant embryo; **C** the Eve expression pattern in a *Him 74* null mutant embryo; **D** Eve expression pattern in a *Him 195* embryo; **E** Histogram showing the obtained numbers for the Eve-positive pericardial cells in all three controls (*OR, handCGFP2.3, handCGFP3.1*) and the *Him* null lines. The error bars indicate the standard error and stars indicate a statistically significant difference in regard to the wild-type.

or loss in all three *Him* null mutant lines (D. Hancock, PhD thesis). According to his results there is no compensation through the loss of another somatic muscle for the duplication or *vice versa*. He also measured the thickness for the DO2 muscle and compared his results to those he obtained for wild-type embryos. In his PhD thesis he describes a "relative decrease in average thickness" for the DO2 muscle (D. Hancock, PhD thesis). This result indicates that there might be fewer cells that contribute to each of these muscles. A similar effect on the DO1 muscle has been linked to an increase in Tinman-positive pericardial cells in embryos mutant for *dpp* (Johnson et al., 2007). As I have also observed a loss in the Eve-positive pericardial cells in the *Him* null mutant embryos, this would indicate that *Him* function is necessary early on, during the specification of the initially shared precursors of the Eve-pericardial cell and the DO2 muscle. *Him* is most likely to play a role in the specification or possibly maintenance of these specific precursors. In his thesis, D. Hancock also evaluated the number of Krüppel-positive founder cells in the *Him 52* null mutant line. The Krüppel-positive founder cells are specific for the DA1 and the DO2 muscles and an increase or decrease in their number would indicate an effect of *Him* on their specification. Their number is indeed increased when compared to wild-type embryos (D. Hancock, PhD thesis). It is possible that these founder cells are formed at the expense of the precursors needed for the Eve cell lineage that contributes to the heart and thus accounts for the loss of heart cells. This again points towards a role for *Him* during the specification of the precursors of the Krüppel-positive founder cells.

As explained above, some of the cardioblasts and pericardial cells share a common lineage during development and the number of one cell type can be increased or decreased at the cost of the second cell type in the heart when specific factors are missing or over-expressed (Gajewski et al., 1999; Alvarez et al., 2003). As my data indicate a loss of Eve- and Odd-positive pericardial cells in embryos lacking *Him*, I was interested to see if this is also the case here.

I have used ß3-tubulin as a marker for the cardioblasts. *ß3-tubulin* is expressed in four of the six cardioblasts per hemisegment of the embryo (Leiss et al., 1988). I scored these embryos under DIC optics, which enables me to also visualize the two unlabelled cardioblasts per hemisegment. The number of ß3-tubulin-positive cardioblasts is reduced at a statistically significant level in all three *Him* null mutant lines as well as the total number of cardioblasts I observed (see Figure 3.3.4). The

expression of this protein shows that the regular pattern of two ß3-tubulin-negative cardioblasts followed by four ß3-tubulin-positive cardioblasts per hemisegment is disturbed. Within the individual hemisegments, there is a large variation between a loss and gain of ß3-tubulin-positive cardioblasts. Contrary to my expectations, these data indicate that the loss of pericardial cells in the *Him* null mutants is not accompanied by a corresponding gain of cardioblasts but that *Him* might also necessary for the specification of the correct number of myogenic cardioblasts.

In summary, the loss of *Him* leads to a reduction in all heart cells by embryonic stage 16/17 (see section 3.3.1). While initially surprising, these results do agree with the results of Dan Hancock's analysis of the dorsal somatic muscles, that are often duplicated or missing without compensation from other somatic muscles. All these results point towards a positive role for *Him* during the early specification of heart cells. It is possible that the fate of these early progenitors is changed to that of other mesodermal tissues (fat body and/or visceral mesoderm). A second hypothesis is that *Him* is necessary to achieve the correct "number" of cell divisions necessary for the generation of all heart precursors. For both scenarios, the lack of *Him* would result in a loss of embryonic heart cells.

It must also be mentioned that the number of heart cells scored for each marker vary between the three *Him* null mutant lines. However, all scored cell populations do behave in a similar fashion with regard to the loss of cells when compared to wild-type cell populations. The differences in numbers can be due to several reasons. All embryos of one line where counted at the same time and a continuous counting error cannot be ruled out. The three used *Him* null alleles are also all slightly different in their genetic break points due to the different methods used to generate them. Analyses from our lab indicate that the *Him 74* and *Him 195* alleles are possibly equivalent on a molecular level but that the *Him 52* allele is not. Using PCR, our lab has been able to demonstrate that the break points for the *Him 52* null allele (the one established in our lab) are in agreement with the positions of the P-elements that were used and that five further genes are removed by this deletion (J. Han and M. Taylor, pers. communication). The equivalent analysis for the other two *Him* null alleles (74 and 195) has established that the 5' excision point is in between the position 18101901 and 18102546 of the *X* chromosome (the *Him* coding regions starts at position 18103460 and ends at position 18104774) according to the annotated genome and that the 3' excision point is

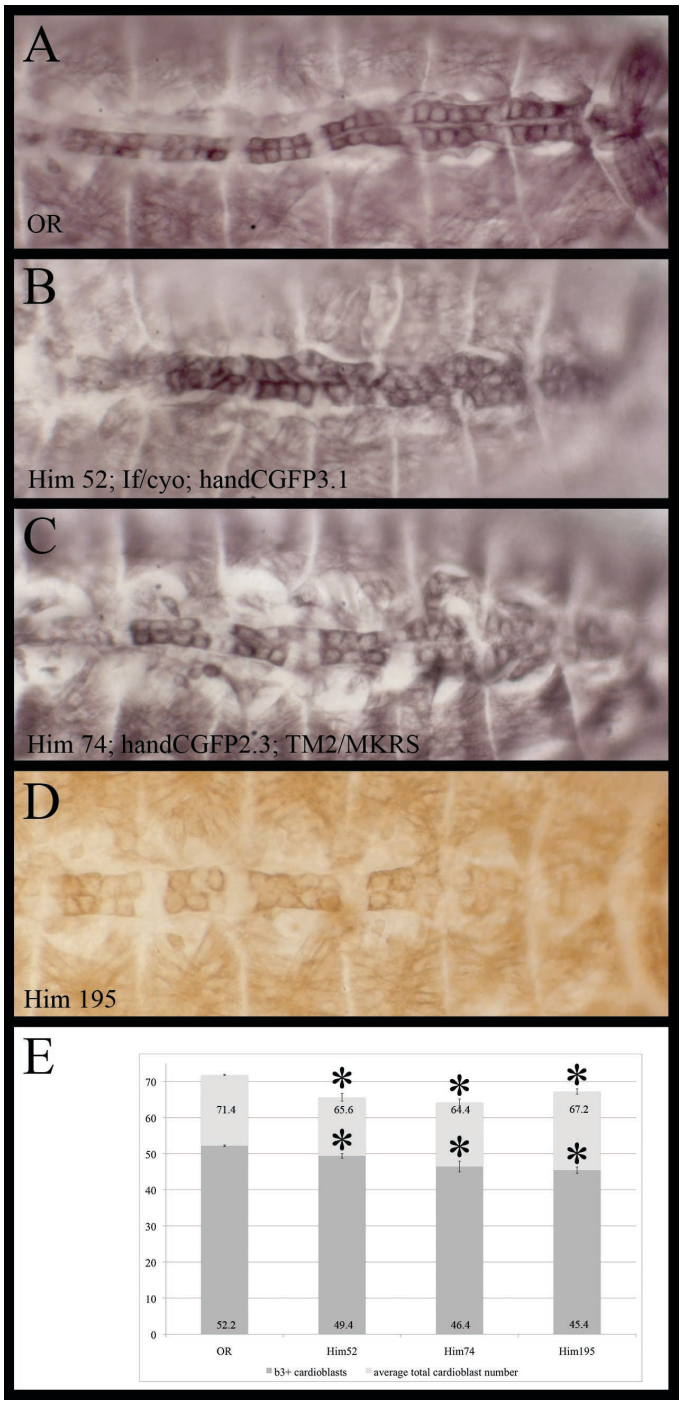


Figure 3.3.4: Expression pattern of ß3-tubulin in *Him* null mutant embryos.
Stage 16/17 embryos were labelled for the ß3-tubulin protein and the number of ß3-tubulin-positive cardioblasts was counted.
A Wild-type embryo; **B** ß3-tubulin expression in a *Him 52* null mutant embryo; **C** ß3-tubulin expression pattern in a *Him 74* null embryo; **D** the ß3-tubulin expression pattern in a *Him 195* null mutant embryo; **E** Histogram showing the numbers for the ß3-tubulin-positive cardioblasts and the total number of cardioblasts in the controls (*OR*) and the *Him* null lines. The error bars indicate the standard error and stars indicate a statistically significant difference compared to the wild-type.

contained within close proximity of the 3' end of the *Him* gene (J. Han, M. Taylor, pers. communication).

3.3.2 *Him* null phenotype in the heart of the larvae

When compared to the embryo, we know relatively little about the development of the pericardial cells post the embryonic stages. Sellin et al. (2006) describe the hand-GFP reporter construct, which drives GFP expression in the heart from stage 12 onwards throughout embryonic, larval and adult heart development. This allows me to visualize the pericardial cells and cardiac cells in live larvae and analyse the behaviour of the marked pericardial cells in various genetic backgrounds. This is important as the complete heart undergoes vast morphological changes during larval and pupal development and it is known that the number of pericardial cells drops sharply from first to second instar larvae (Sellin et al., 2006). This indicates that the pericardial cells undergo further development at this time.

Figure 3.3.5 shows the expression of the two independent hand-GFP reporter lines used as a wild-type control for these experiments (Sellin et al., 2006). I assayed the number, appearance and organisation of the larval pericardial cells at three time points, young first instar larvae (28 to 30 hrs AEL), young second instar larvae (52 to 54 hrs AEL) and mid-stage third instar larvae (100-102 hrs AEL). The heart cells of the first instar larvae (Figure 3.3.5 A, B) are still very similar in number, shape and size to those of a late stage embryo. As the pericardial cells are arranged three-dimensionally around the heart tube, differentiating between the cardioblasts and pericardial cells (both are equally marked by GFP expression) is very difficult at this time point.

By early second instar (Figure 3.3.5 C, D) the number of pericardial heart cells has decreased as described in Sellin et al (2006), especially in the anterior regions of the aorta. The cell body and nuclei of the remaining pericardial cells have increased in size by this stage. The difference in cell shape and size make for an easily scorable difference between the two cell types. The pericardial cells (arrows in Figure 3.3.5) have adopted their typical large round, globular shape with enlarged nuclei while the cardioblasts (see stars in Figure 3.3.5) have retained their original size and the structure

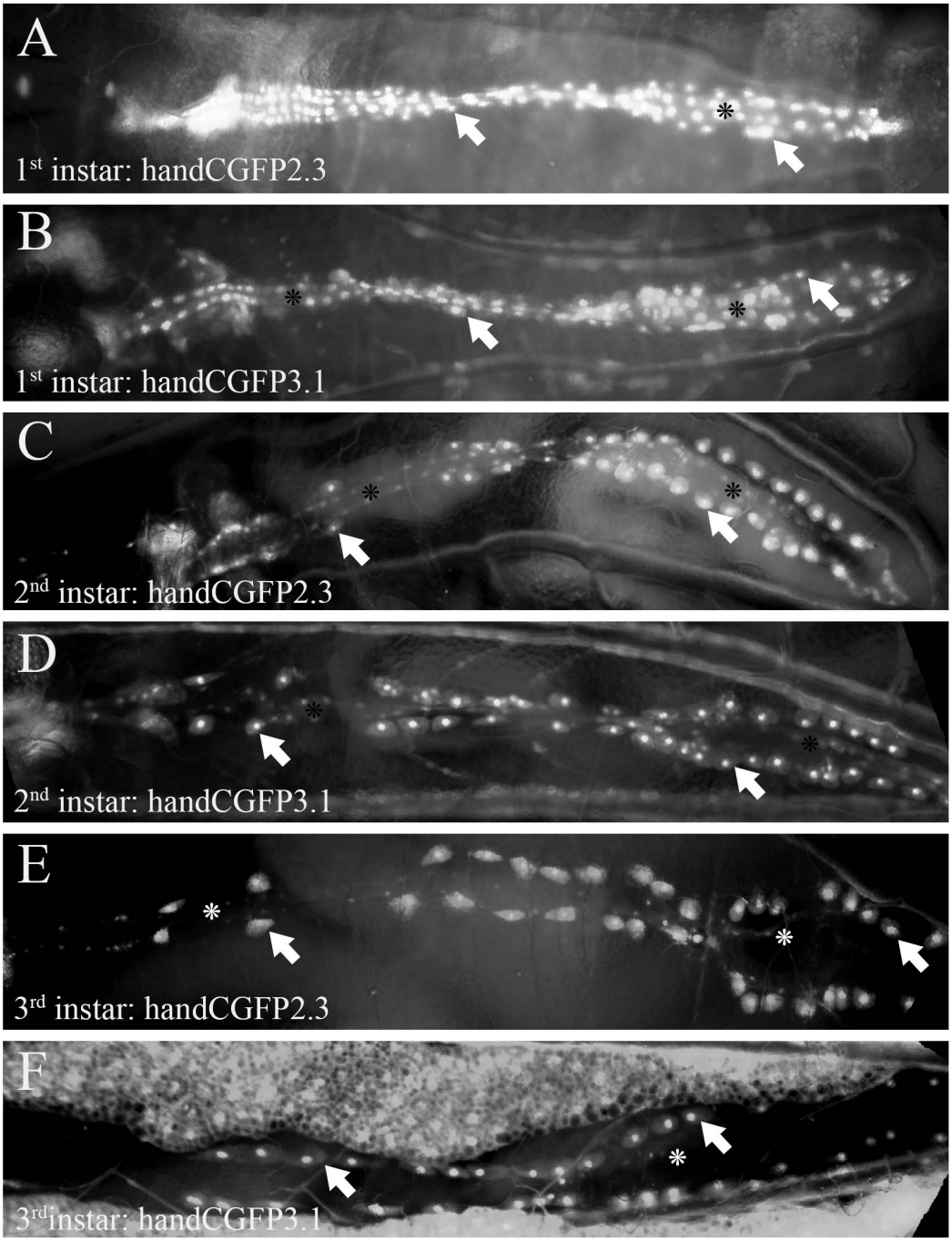


Figure 3.3.5: Expression pattern of the *hand-GFP* reporter during larval life.
Live larvae were photographed under the microscope to visualize the GFP expression pattern. Stars indicate the location of the cardioblasts, arrows point towards the pericardial cells.
A Expression pattern of the *handCGFP2.3* reporter construct in the first larval instar; **B** Expression pattern of the *handCGFP3.1* reporter construct in the first larval instar; **C** Expression of the *handCGFP2.3* reporter construct in the second larval instar; **D** Expression pattern of the *handCGFP3.1* reporter construct in the second larval instar; **E** Expression of the *handCGFP2.3* reporter construct in the third larval instar; **F** Expression pattern of the *handCGFP3.1* reporter construct in the third larval instar.

of the muscular cardiac tube can now also be identified. The size difference between the pericardial cells and the cardioblasts results in a noticeable difference in hand-GFP expression between these two cell types.

The difference between second and third instar larvae lies mainly in the overall size of the larvae (compare Figure 3.3.5 C with E and D with F). The shape and positioning of the pericardial cells in third instar larvae is more or less unchanged from that observed in a second instar larvae, with the exception of a general increase in size of all the heart cells. In all of the experiments in larva, I have relied on the hand-GFP construct to mark the heart cells. The pericardial cells were counted based on being hand-GFP positive and not included within the heart tube.

Him expression is lost from the heart at the end of embryogenesis, thus it is possible that this loss of *Him* expression is necessary for the correct development and/or maintenance of the pericardial cells. As described above, in the embryo the loss of *Him* results in a decrease in pericardial cell number. If this were the case, I would expect to see a similar reduced number of pericardial cells in the first instar and a change in cell number and/or shape during the first two larval instars.

For the two of the *Him* null mutant lines I was able to generate a stable, doubly homozygous line carrying both the *Him* deletion and the hand-GFP reporter construct. I was unable to create a viable fly line both carrying the *Him 195* null mutation and either hand-GFP reporter construct. Crossing the two homozygous fly lines together (*Him 195* and handCGFP2.3 construct) and selecting for the GFP-positive male larvae generated the hand-GFP positive larvae for this line.

Figure 3.3.6 shows the appearance of the pericardial cells as visualized by the hand-GFP reporter construct in first instar wild-type (i.e. both hand-GFP-reporter constructs on their own) and *Him* null mutant larvae. The first instar larvae deficient for *Him* do not appear greatly different to the wild-type larvae, all three lines still express the hand-GFP reporter construct, indicating that the identity of the pericardial cells is not affected when it comes to the factors necessary for the expression of *hand* through the *hand*-enhancer fragment. While I initially counted the number of *hand*-GFP-positive pericardial cells at this time point, my analysis pointed to a statistically significant difference between the numbers for both wild-type strains used. This is most likely due to the similarities in cell shape of pericardial cells and cardioblasts and their position around the heart. This has led me to disregard my counts of the pericardial cell numbers of this stage, limiting my analysis for this time point to observations only.

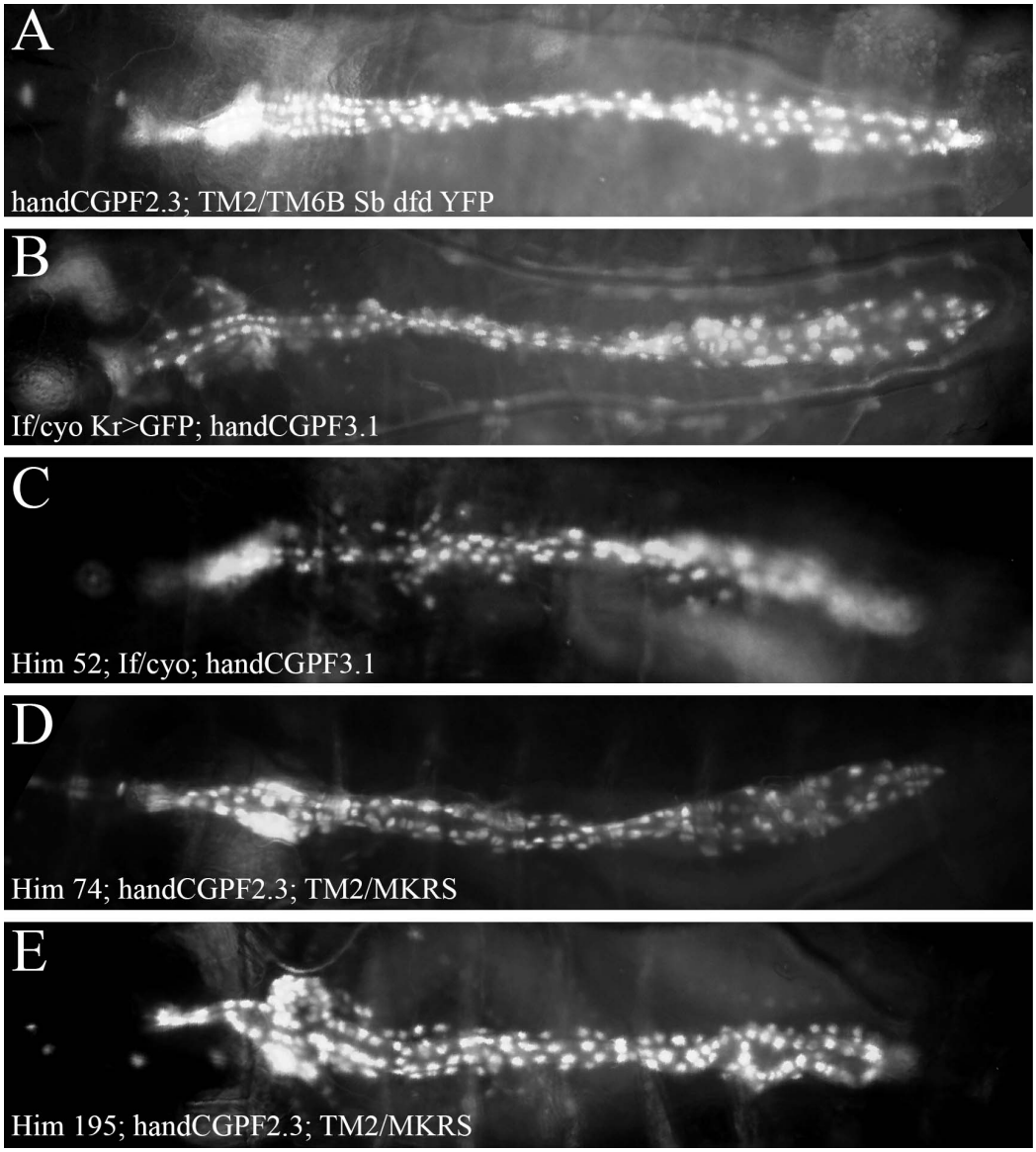


Figure 3.3.6: Expression pattern of the *hand-GFP* reporter construct in first instar larvae carrying *Him* deletions.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the *GFP* expression pattern.
A *handCGFP2.3* reporter construct, wild-type control; **B** *handCGFP3.1* reporter construct, wild-type control; **C** GFP expression pattern in the *Him 52* null mutant first instar larvae; **D** GFP expression pattern in the *Him 74* null mutant first instar larvae; **E** GFP expression pattern of the *Him 195* null mutant first instar larvae.

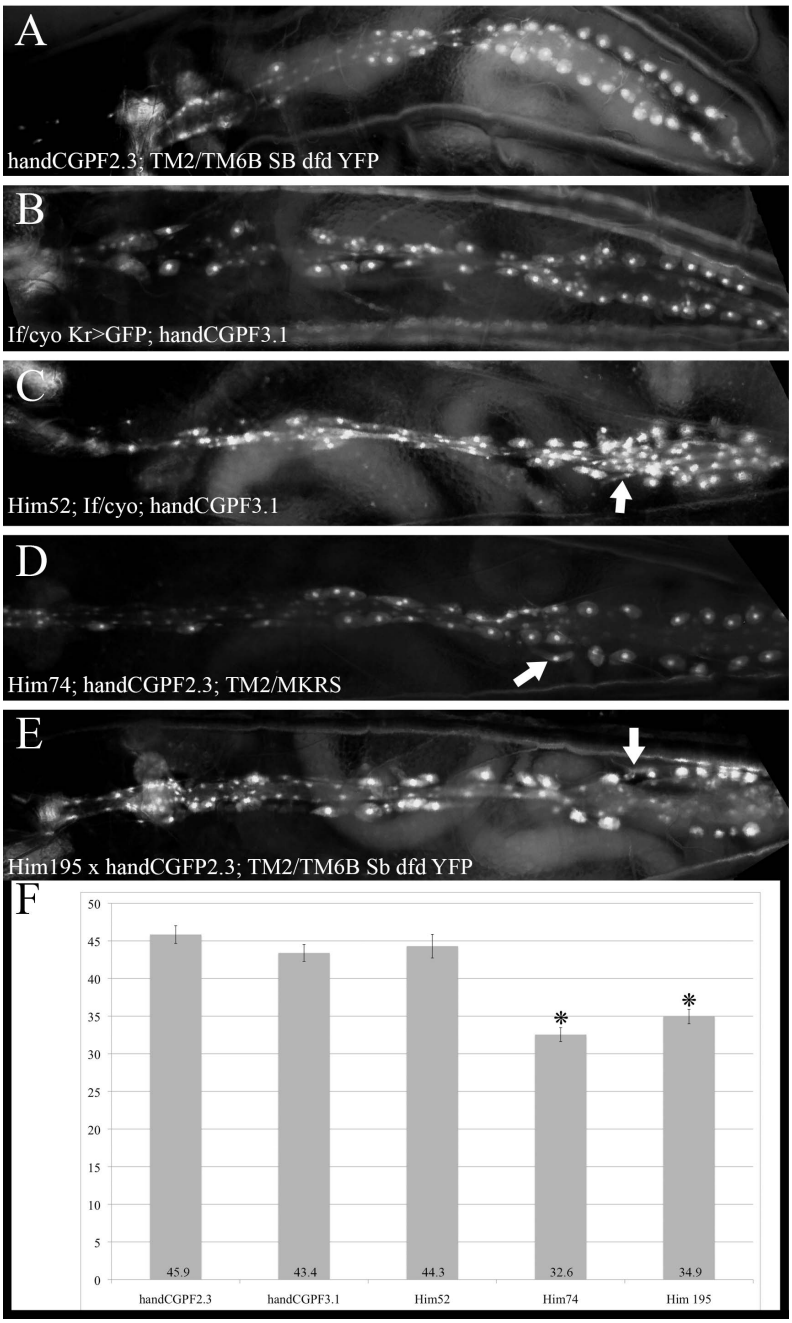


Figure 3.3.7: Expression pattern of the *hand-GFP* reporter construct in second instar larvae carrying *Him* deletions.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the *GFP* expression pattern and the number of GFP positive pericardial cells was counted.
A *handCGFP2.3* reporter construct, wild-type control; **B** *handCGFP3.1* reporter construct, wild-type control; **C** GFP expression pattern in the *Him 52* null mutant second instar larvae; **D** GFP expression pattern in the *Him 74* null mutant second instar larvae; **E** GFP expression pattern of the *Him 195* null mutant second instar larvae; **F** Histogram showing the obtained numbers for the handGFP-positive pericardial cells in the two controls (*handCGFP2.3, handCGFP3.1*) and the second instar *Him* null larvae. The error bars indicate the standard error and stars indicate a statistically significant difference compared to the wild-type.

Figure 3.3.7 shows representative images of all three *Him* null lines and the control at the second instar time-point. The *Him 74* and *Him 195* lines show a statistically significant loss of pericardial cells, while the number of pericardial cells of the *Him 52* line is lower than that of the control animals but not significantly so. In the larvae of the *Him 74* and *Him 195* lines the pericardial cells are spaced further apart along the anterior-posterior axis than in the wild-type larvae, so that the loss is distributed relatively evenly across the whole larval heart and less of a difference between the anterior aorta and heart proper is visible. The loss of pericardial cells I observed at the end of embryogenesis is maintained in these two lines until the beginning of the second instar. The size of the pericardial cells appears not to be affected by the loss of *Him*. However, some cells do have a less round shape than those of the control larvae at this time point (compare arrows in Figure 3.3.7). Without the help of further markers for sub-sets of pericardial cells I cannot identify if there is a specific sub-set of pericardial cells that is more affected by the loss of *Him* than other pericardial cells.

Figure 3.3.8 depicts all three *Him* null mutant lines and the control at the third instar time-point. At this time-point, only the *Him 52* line shows a significant loss of pericardial cells. The pericardial cell numbers for the other two lines, *Him 74* and *Him 195*, are, while still lower than those of the control animals, not statistically different to the control animals at the chosen confidence interval of 0.05. It appears that the loss of pericardial cells that has been described in wild-type larvae between the first and third instar larvae does not occur in larvae lacking *Him*. It is possible that the pericardial cells usually lost at this time are the ones that are missing in the embryos already. This would explain why the number of *hand*-GFP-positive pericardial cells approaches the wild-type level again. Similarly any changes in cell shape are not observable anymore and it thus not possible to assign any importance to that observation. It is possible that other genes can compensate for the loss of *Him* function, at least partially. Another possibility is that the loss of pericardial cells that occurs between the first and third instar wild-type larva, occurs at an earlier stage in *Him* mutants.

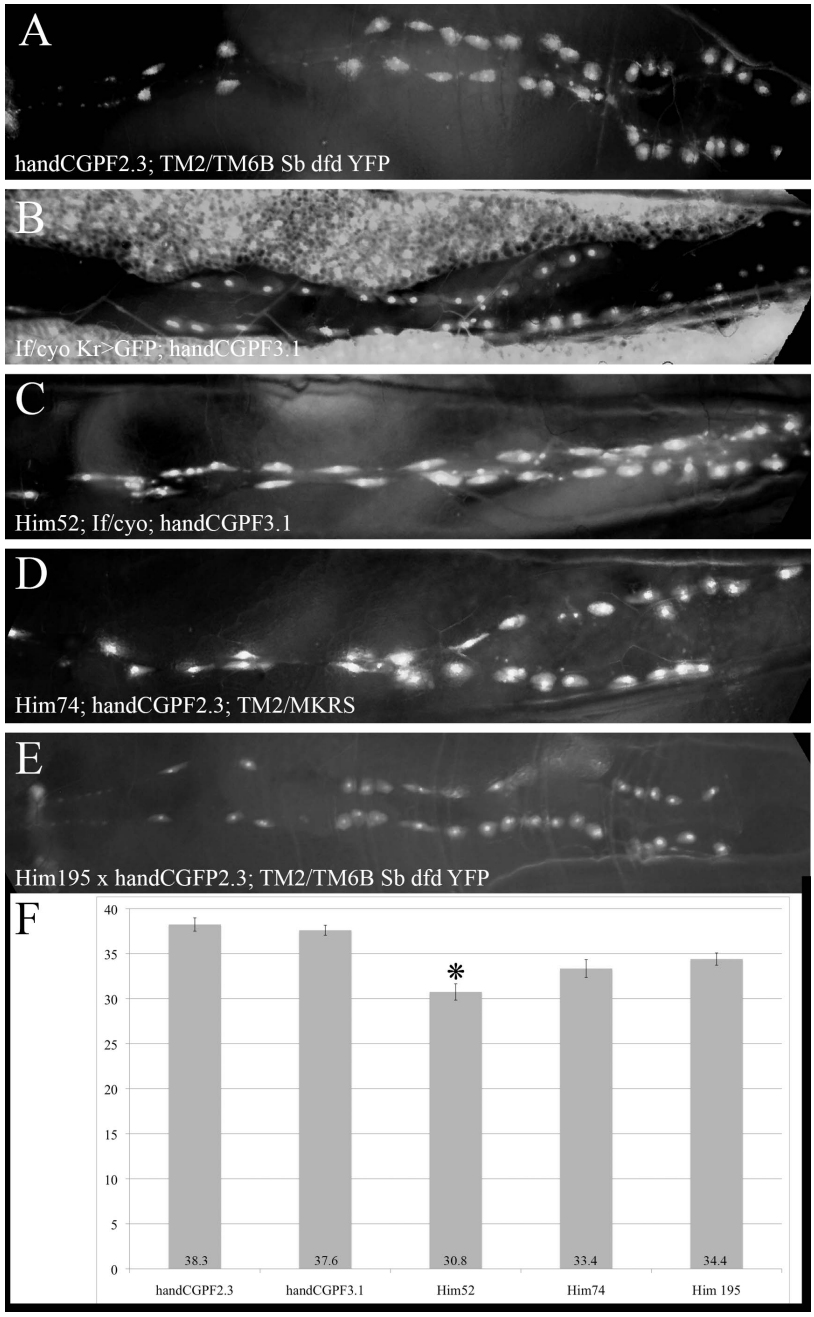


Figure 3.3.8: Expression pattern of the *hand-GFP* reporter construct in third instar larvae carrying *Him* deletions.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the *GFP* expression pattern and the number of GFP positive pericardial cells was counted.
A *handCGFP2.3* reporter construct, wild-type control; **B** *handCGFP3.1* reporter construct, wild-type control; **C** GFP expression pattern in the *Him 52* null mutant third instar larvae; **D** GFP expression pattern in the *Him 74* null mutant third instar larvae; **E** GFP expression pattern of the *Him 195* null mutant third instar larvae; **F** Histogram showing the obtained numbers for the handGFP-positive pericardial cells in the two controls (*handCGFP2.3, handCGFP3.1*) and the third instar *Him* null larvae. The error bars indicate the standard error and stars indicate a statistically significant difference compared to the wild-type.

3.3.3 *Him* is not necessary for the development of functional wing hearts

The wing hearts are pulsatile organs located under the scutellum of the adult fly (Tögel et al., 2008; Lehmacher et al., 2009). Recent publications have shown that these organs are necessary for the correct formation of the wings and that the lack of them results in a non-clearing of the haemolymph from between the wing layers which is visible if examined under the microscope as an "opaque" sheen of the wing (Tögel et al., 2008; M. Tögel, pers. communication). It has also been demonstrated that these organs originate from the anterior most Eve-positive pericardial cells of the larva (Tögel et al., 2008). It is thus possible that the effect of the loss of *Him* extends to these organs as *Him* is expressed in these Eve-positive pericardial cells during embryogenesis.

Figure 3.3.9 shows the wings of the *Him* null mutant lines and those of wild-type animals under the microscope. All wings, including those of the *Him* null flies appear equally clear and translucent. From this I conclude that the loss of *Him* does not affect the function of the wing hearts in clearing out the haemolymph from the wings after the fly ecloses. I however cannot exclude the possibility that the function of the wing hearts might be impaired as I assayed the wings of 6-hour-old flies. If the wing hearts of the *Him* null mutants function less efficiently than that of wild-type flies this stage might be too late to detect any slower clearing of the wings. It is also possible, that part of the development of the wing hearts is affected due to the loss of *Him* but that the wing hearts still remain mainly functional. This will require further investigation.

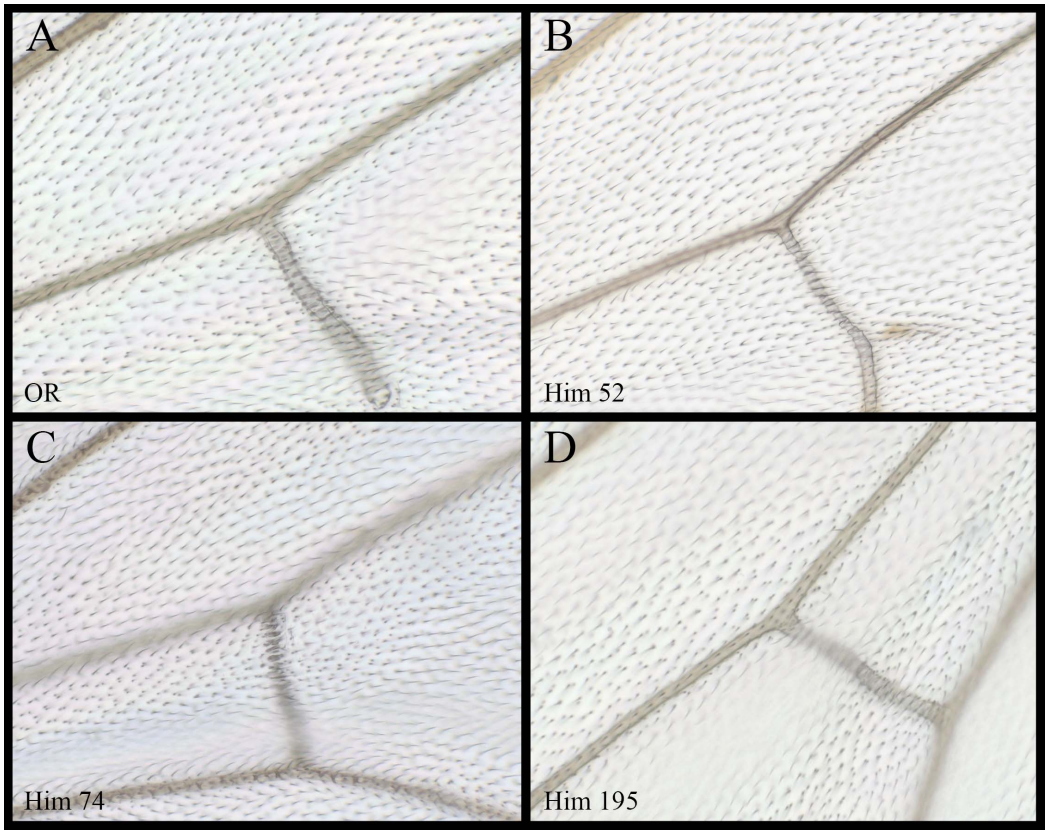


Figure 3.3.9: The function of the wing hearts is not affected in *Him* null mutants.
The wings of young flies were removed and analysed for their clarity under the microscope (Zeiss Axiovision).
A wing of the *OR* wild-type control; **B** wing of a fly homozygous for the *Him 52* null deletion; **C** wing of a fly homozygous for the *Him 74* null deletion; **D** wing of a fly homozygous for the *Him 195* null deletion.

3.4 *Him* over-expression phenotype in the heart

3.4.1 *Him* over-expression phenotype in the embryonic heart

It is possible to target the ectopic expression of a gene of interest in spatial and temporal pattern using the Gal4/UAS system; the components of which were originally isolated from yeast (Brand and Perrimon, 1993). This system uses an enhancer fragment to "drive" expression of the Gal4 protein in a known pattern. A fly line containing this is crossed to a fly line containing the gene of interest under the control of the UAS (Upstream Activating Sequence). The offspring of this cross has copies of both transgenes, enabling the Gal4 protein to bind to the UAS and initiate transcription of the target gene. This method is a very useful tool when analysing the phenotype of genes as this allows an investigation of what happens if the gene is expressed outside its "normal" parameters. The Gal4/UAS system has successfully contributed to help define the function of genes such as *Mef2*, *twist*, *tin* etc.

We hypothesise that in the somatic mesoderm of the embryo the Him protein is necessary to set apart the AMPs from the remaining somatic musculature by regulating Mef2 activity (Liotta et al., 2007). In the established heart, *Mef2* is only expressed in the contractile cardioblasts, while *Him* is only expressed in the non-contractile pericardial cells (see Figure 3.2.2). The *Mef2* transcript is thought to be expressed in the shared heart precursors. This has been shown through single *in situ* hybridisation against the *Mef2* transcript (Lilly et al., 1994; Nguyen et al., 1994). I have not been able to grab an image of an earlier (stage 10 and/or 11) embryo double-labelled for both the *Him* transcript and Mef2 of sufficient quality to analyse their possible co-location in these stages. It is likely that both *Him* and *Mef2* are co-expressed in the very early developmental stages of the heart progenitors but this hypothesis has yet to be proven. Both, cardioblasts and pericardial cells, originate from the dorsal mesoderm and in the case of the Seven-up-positive cardioblasts and two of the Odd-positive pericardial cells, a cardioblast and a pericardial cell each share a common mother cell (Ward and Skeath, 2000; see Introduction). It is thus possible that Him has a similar function in the developing cardiac tissue as it does in the somatic mesoderm, i.e. delineate contractile muscle cells from non-contractile mesodermal cells by regulating Mef2 activity. This has led me to the hypothesis that the exclusion of *Him* from the cardioblasts and the

prevention of Mef2 function in the pericardial cells are equally necessary for normal heart development. To test this hypothesis, I have used the UAS/Gal4 system to ectopically express *Him* in the cardioblasts and analysed the effects this has on the development of the heart. The reciprocal experiments of removing and over-expressing *Mef2* will be described in the next section.

Figure 3.4.1 shows the effect of driving *Him* over-expression in the heart using the *hand-Gal4* driver and the *UAS-Him J7* construct previously generated in this lab (Liotta et al., 2007). In the heart, the *hand-Gal4* driver is expressed from embryonic stage 12 onwards and throughout larval and adult life in all cardioblasts and pericardial cells (Sellin et al., 2006, see Appendix). The pericardial cells are labelled for the Zfh-1 protein. *Him* over-expression in the heart causes an increase in pericardial cells when compared to wild-type. The supernumerary pericardial cells are spread evenly across the length of the heart in additional "rows" when compared to the wild-type organisation of the Zfh-1-positive pericardial cells. As with the evenly distributed loss of pericardial cells in the *Him* null mutants, this indicates that multiple hemisegments are affected in the same way.

As previously mentioned, an exact numerical analysis of the Zfh-1-positive pericardial cells by counting has proven to be unreliable, so instead I have concentrated on scoring the number of Eve- and Odd-positive pericardial cells as I did in the *Him* null mutants. Figure 3.4.2 shows the expression pattern of Eve in wild-type embryos and those over-expressing *Him*. There is no detectable visual difference between the embryos. Counting of the Eve-positive pericardial cells in both wild-type and *Him* over-expressing late embryonic embryos (stage 16/17) confirms this observation (Figure 3.4.2 C). Thus, over-expression of *Him* in the heart cells does not affect the expression of *eve* in the heart or alter Eve-positive pericardial cells number. The loss of *Him* leads to a reduction in the number of Eve-positive pericardial cells. This indicates that while *Him* is necessary for the development of the Eve-pericardial cells, over-expression in the heart only is not capable of inducing hyperplasia of these cells. Interestingly, neither of the two Eve-positive pericardial cells per hemisegment share a common mother cell with a cardioblast but have a common lineage with the founder of the DA1 somatic muscle (Ward and Skeath, 2000) in which *Him* has not been over-expressed in this experimental set-up. It would be interesting to see how the Eve-positive pericardial cells behave if a different driver expressed in the dorsal somatic muscles or at an earlier stage than the *hand-Gal4* driver, is used.

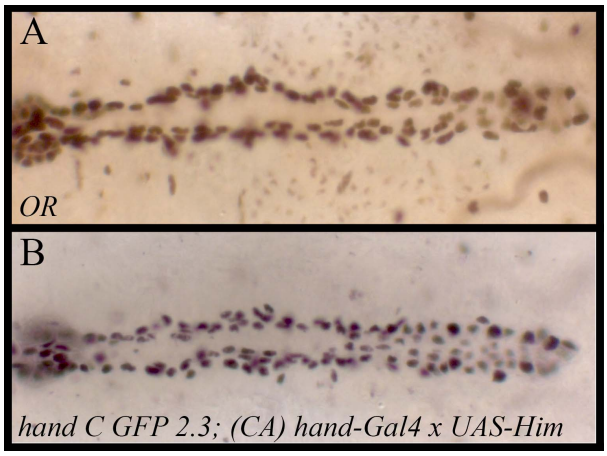


Figure 3.4.1: Zfh-1 expression in embryos over-expressing *Him* in the heart.
Stage 16/17 embryos were labelled for the distribution of the Zfh-1 protein.
A *OR* wild-type embryo; **B** Zfh-1 expression in an embryo ectopically over-expressing *Him* under the control of the *hand-Gal4* driver.

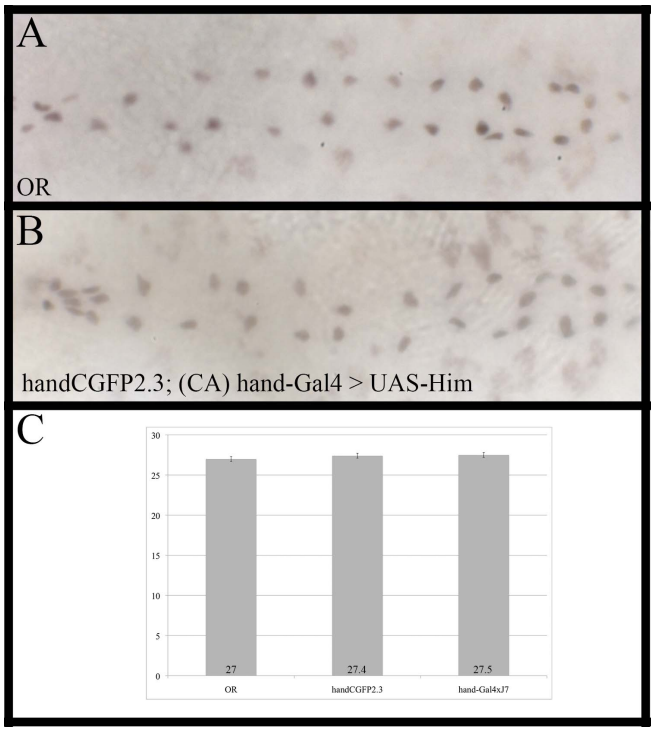


Figure 3.4.2: Expression pattern of Eve in embryos expressing ectopic *Him* in the heart.
Stage 16/17 embryos were labelled for the Eve protein and the number of Odd-positive pericardial cells was counted.
A Wild-type embryo; **B** Odd expression in an embryo ectopically over-expressing *Him* under the control of the *hand-Gal4* driver; **C** Histogram showing the numbers for the Odd-positive pericardial cells in the two controls (*OR, handCGFP2.3*) and the over-expression of *Him*. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

Figure 3.4.3 shows the effect of *Him* over-expression on the Odd-positive pericardial cells. When image 3.4.3 B is compared to the wild-type embryo in 3.4.3 A, it can be seen that over-expression of *Him* driven by *hand-Gal4* causes an increase in Odd-positive pericardial cells. While the increase of pericardial cells is generally distributed equally along the heart, I have observed clustering of two to five Odd-positive pericardial cells in a small bunch in about two thirds of the embryos (see stars in Figure 3.4.3 B). As can be seen in the graph depicted in Figure 3.4.3 C, counting of the Odd-positive cells revealed a statistically significant increase of over ten pericardial cells per animal when *Him* is over-expressed. Thus, at least a part of the hyperplasia of the Zfh-1-positive pericardial cells observed but not scored is due to a significant increase in Odd-positive pericardial cells. I believe that the Tin-positive pericardial cells, which were not assayed in this study, are likely to also have increased in number under these circumstances as the hyperplasia observed when labelling for the Zfh-1 protein appears to be more widespread than that of the Odd-positive pericardial cells alone. A sub-set of the Odd-positive pericardial cells shares an immediate common progenitor with the Seven-up-positive cardioblasts, while the Eve-positive pericardial cells share progenitors with specific dorsal muscles. The genetic regulatory network is different in these cells thus explaining the difference in the reaction to an increased amount of *Him* present in these cells.

In a similar experiment described in David Liotta's PhD thesis, he over-expressed *Him* using the *Him-Gal4 L3-5* driver, which drives expression slightly earlier (from late stage 11/ early stage 12) than the *hand-Gal4* driver I used but its expression is restricted to the pericardial cells only; he observed a similar phenotype (D. Liotta, PhD thesis). He describes a slight but significant increase in the Odd-positive pericardial cells and no change in the number of Eve-positive pericardial cells. This further substantiates my results.

The origin of these supernumerary pericardial cells is unclear so far. My results point towards the cardioblasts as a possible source. In order to investigate this, I analysed the organisation and number of the ß3-tubulin-positive cardioblasts in embryos that over-express *Him* in all embryonic heart cells under the control of the *hand-Gal4* driver. ß3-tubulin is expressed in four of the six cardioblasts per hemisegment. Figure 3.4.4 shows the results of this experiment. Analysing the embryos under DIC optics, I did not observe any organisational differences between the wild-type embryo and the embryos over-expressing *Him* (compare Figure 3.4.4 A and B). Furthermore, scoring of

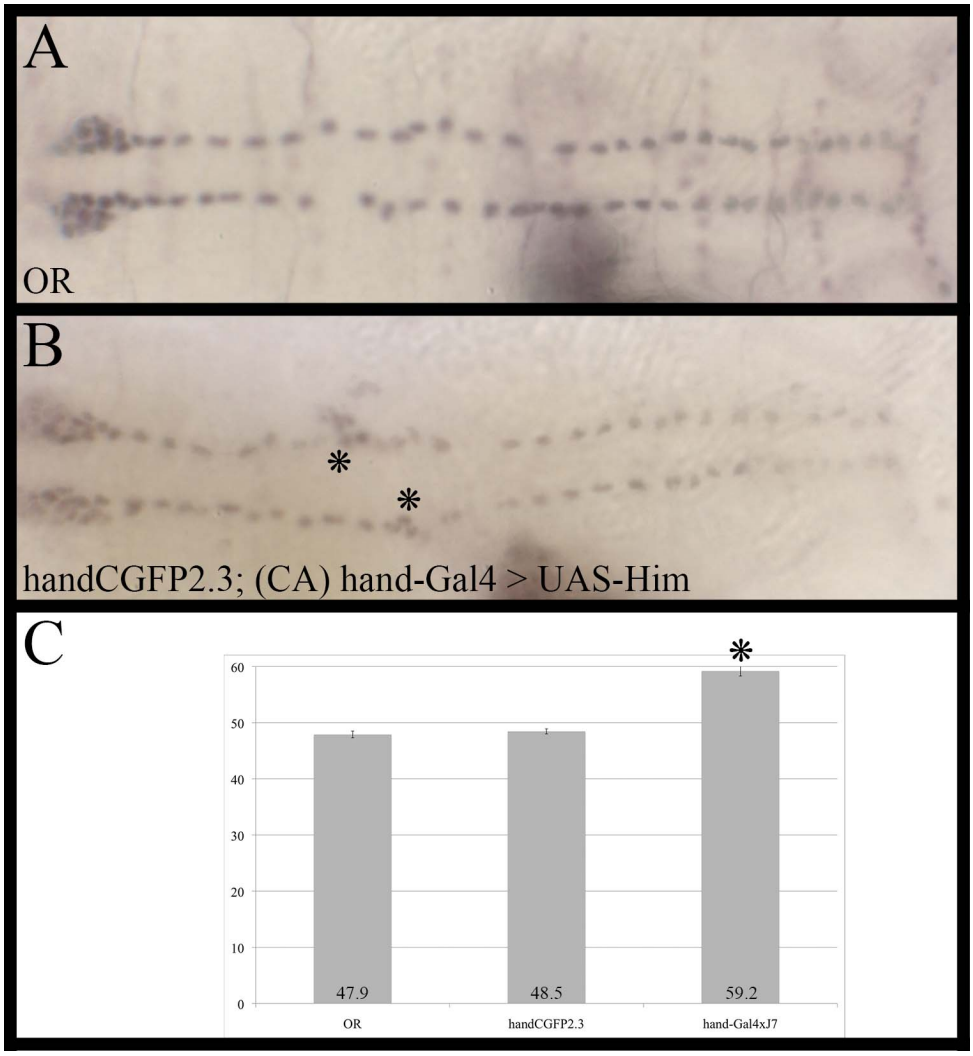


Figure 3.4.3: Expression pattern of Odd in embryos that over-express *Him* in the heart.
Stage 16/17 embryos were labelled for the Odd protein and the number of Odd-positive pericardial cells was counted.
A Wild-type embryo; **B** Odd expression in an embryo ectopically over-expressing *Him* under the control of the *hand-Gal4* driver, stars indicate clustering of cells; **C** Histogram showing the numbers of Odd-positive pericardial cells in the two controls (*OR, handCGFP2.3*) and the over-expression of *Him*. The error bars indicate the standard error and stars indicate a statistically significant difference compared to wild-type. The *UAS-Him* construct is called "J7" in the graph.

the total cardioblasts and those positive for ß3-tubulin confirms this, as there was no significant change in cell number between the two types of embryos.

In order to have a more specific look at the behaviour of the cardioblasts under these conditions, I exclusively over-expressed *Him* in the cardioblasts within the heart with the *TinCΔ4-* and *Mef2-Gal4* drivers. These experiments should also establish if the exclusion of *Him* expression from the cardioblasts (see chapter 6) is important for the correct development of these cells. I assayed the expression pattern of two different marker genes in these embryos, ß3-tubulin (Figure 3.4.4 C and D) and Myosin (Figure 3.4.5). Myosin is a marker for all cardioblasts; it also marks the two ß3-tubulin-negative cardioblasts that arise by asymmetric division. Furthermore, while ß3-tubulin expression has been shown to be unaffected in a *Mef2* mutant (Bour et al., 1995), we know that the expression pattern of Myosin is greatly disturbed in *Mef2* mutants (Lilly et al., 1995; Ranganayakulu et al., 1995; D. Liotta, PhD thesis). If Mef2 and Him have opposing roles during heart development as they do in the development of the somatic muscles, it is possible that the *Mef2* null mutant phenotype and the *Him* over-expression phenotype resemble each other. As can be seen in Figure 3.4.5 A-C the expression of Myosin is affected in the cardioblasts if *Him* is over-expressed in these. The strength of the Myosin expression is very much reduced if *Him* is over-expressed using the *Mef2-Gal4* driver (compare Figure 3.4.5 A to C). As *myosin* is a target gene of Mef2 in the cardioblasts, this points into the direction of Him interfering with Mef2 function in a way similar to what has been described for the somatic mesoderm (Liotta et al., 2007). As the expression of *myosin* is affected in these embryos, scoring them was not possible. Figures 3.4.4 C and D show the same result for the ß3-tubulin-positive cardioblasts if these drivers are used. The arrangement, shape and number of the ß3-tubulin-positive cardioblasts is indistinguishable from that of the control embryos. The results are further substantiated by the total number of cardioblasts present as counted in the ß3-tubulin stain, as these remain very close to that counted for the wild-type. The ectopic expression of *Him* in the cardioblasts leaves at least the Odd-positive pericardial cells unaffected as can be seen in Figure 3.4.6. Cell numbers and arrangement of the Odd-positive pericardial cells are like that in wild-type embryos when the *Mef2-Gal4* driver controls the *UAS-Him* construct. It is very likely that this would also be the case for the *TinCΔ4-Gal4* driver as *Mef2-Gal4* drives mesodermal expression from stage 7 onwards and thus is expressed in the shared precursors of cardioblasts and pericardial

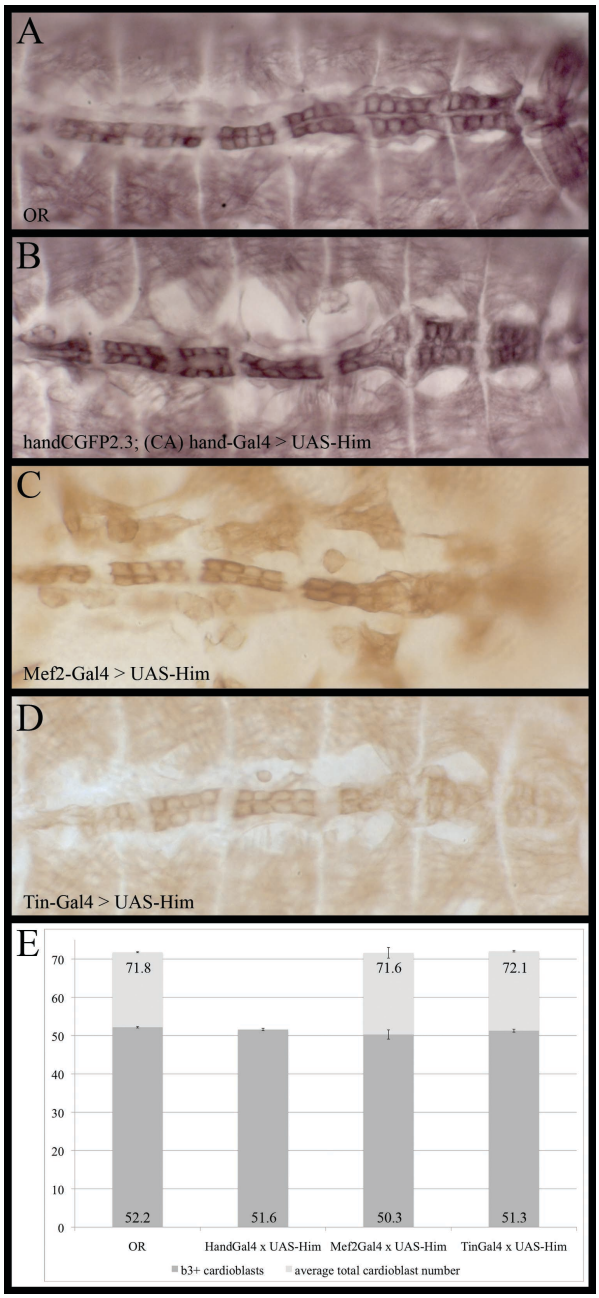


Figure 3.4.4: Expression pattern of ß3-tubulin in embryos that over-express *Him* in the heart.
Stage 16/17 embryos were labelled for the ß3-tubulin protein and the number of ß3-tubulin-positive cardioblasts and the total number of cardioblasts were counted.
A Wild-type embryo; **B** ß3-tubulin expression in an embryo ectopically over-expressing *Him* under the control of the *hand-Gal4* driver; **C** ß3-tubulin expression in an embryo ectopically over-expressing *Him* in the cardioblasts under the control of the *Mef2-Gal4* driver **D** ß3-tubulin expression in an embryo ectopically over-expressing *Him* in four of the six cardioblasts under the control of the *TinCΔ4-Gal4* driver; **E** Histogram showing the numbers of ß3-tubulin-positive cardioblasts and the total number of cardioblasts in the control (OR) and the different over-expressions of *Him*. For the *hand-Gal4* driver the total number of cardioblasts was not determined. The error bars indicate the standard error, there is no statistically significant difference compared to wild-type.

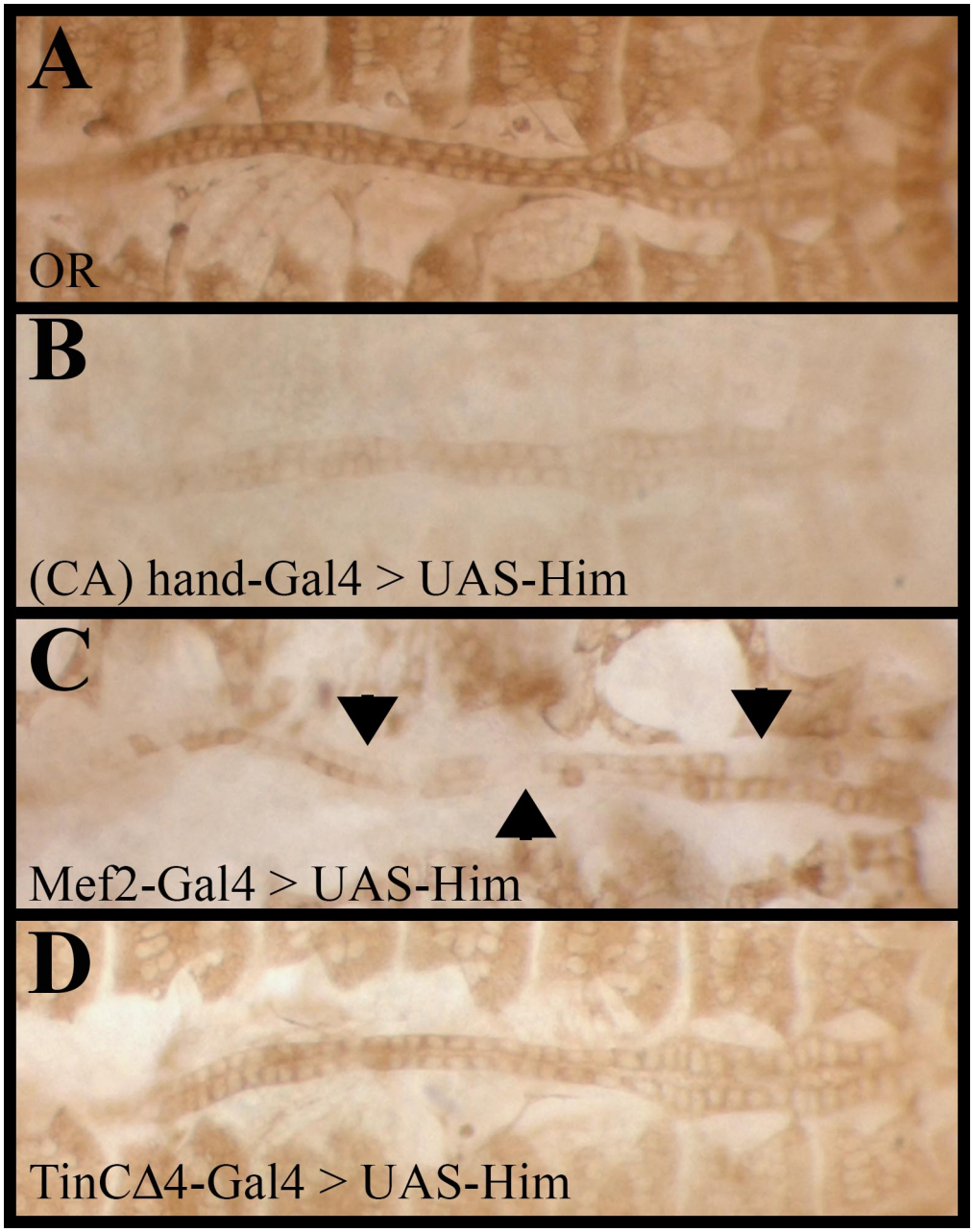


Figure 3.4.5: *Him* over-expression in the cardioblasts can disrupt Myosin expression.
Myosin immuno-stain on stage 17 embryos. Depending on the Gal4 driver used, ectopic *Him* expression in the cardioblasts has a disruptive effect on the expression of Myosin.
A Myosin immuno-stain on wild-type embryos (*OR*). **B** Myosin immuno-stain on embryos over-expressing *Him* under the control of the *hand-Gal4* driver. This leads to a general reduction in Myosin expression. **C**: Myosin expression is reduced and/or completely lost from individual cells (see arrowheads) when *Mef2-Gal4* drives *Him* over-expression. **D** Myosin expression is not affected if *Him* expression is under the control of the *TinCΔ4-Gal4* driver.

cells. Thus, the expression of *Him* in the cardioblasts has no effect on the Odd-positive pericardial cells.

These experiments rule out that the supernumerary pericardial cells observed in embryos over-expressing *Him* in all heart cells developed at the expense of the ß3-tubulin-positive cardioblasts. These experiments however do not directly account for the two non-ß3-tubulin cardioblasts per hemisegment with a positive stain as a possible source for the extra pericardial cells if the *hand-Gal4* driver is used.

The expression pattern of myosin in the embryos over-expression *Him* indicates that these cells are formed normally but that the expression of at least one Mef2 target gene (*myosin*) is greatly reduced. It would be interesting to see if the number of Svp- and/or Mef2-positive cardioblasts, which includes the two ß3-tubulin-negative cells, changes as counting the Myosin-positive cardioblasts was not possible. Svp is a marker for the two cardioblasts that do not express ß3-tubulin and Mef2 is expressed in all six cardioblasts per hemisegment. These two markers should be the next markers to be analysed.

Another possibility for the origin of the pericardial hyperplasia is that the ectopic expression of *Him* in all of the pericardial cells interferes with the number of cell divisions specific for the pericardial cells that would normally occur in a wild-type environment. This hypothesis would explain the stable number of ß3-tubulin-positive cardioblasts. The results presented in David Liotta's thesis also point in this direction. He showed that just over-expressing *Him* in the pericardial cells it is intrinsically expressed in, is sufficient to generate a slight increase in pericardial cell number. These experiments used the *Him-Gal4 L3-5* driver, which drives expression slightly earlier than the *hand-Gal4* driver I used in my experiments. It is possible that by over-expressing *Him* at an earlier time during embryonic development, progenitors were targeted that usually contribute to a different lineage, likely the somatic mesoderm, which loses *Him* expression much earlier than the pericardial cells and forced them into a pericardial cell fate. These embryos show a high degree of disruption of the somatic myogenesis at the stage of differentiation (D. Liotta, PhD thesis).

The result of the over-expression of *Him* in the cardioblasts also shows that the simple ectopic expression of *Him* in the cardioblasts can, at least partially, mimic the *Mef2* mutant phenotype of the cardioblasts as the levels of *myosin* expression are reduced in these embryos. This indicates that, in the early heart progenitors, there is a

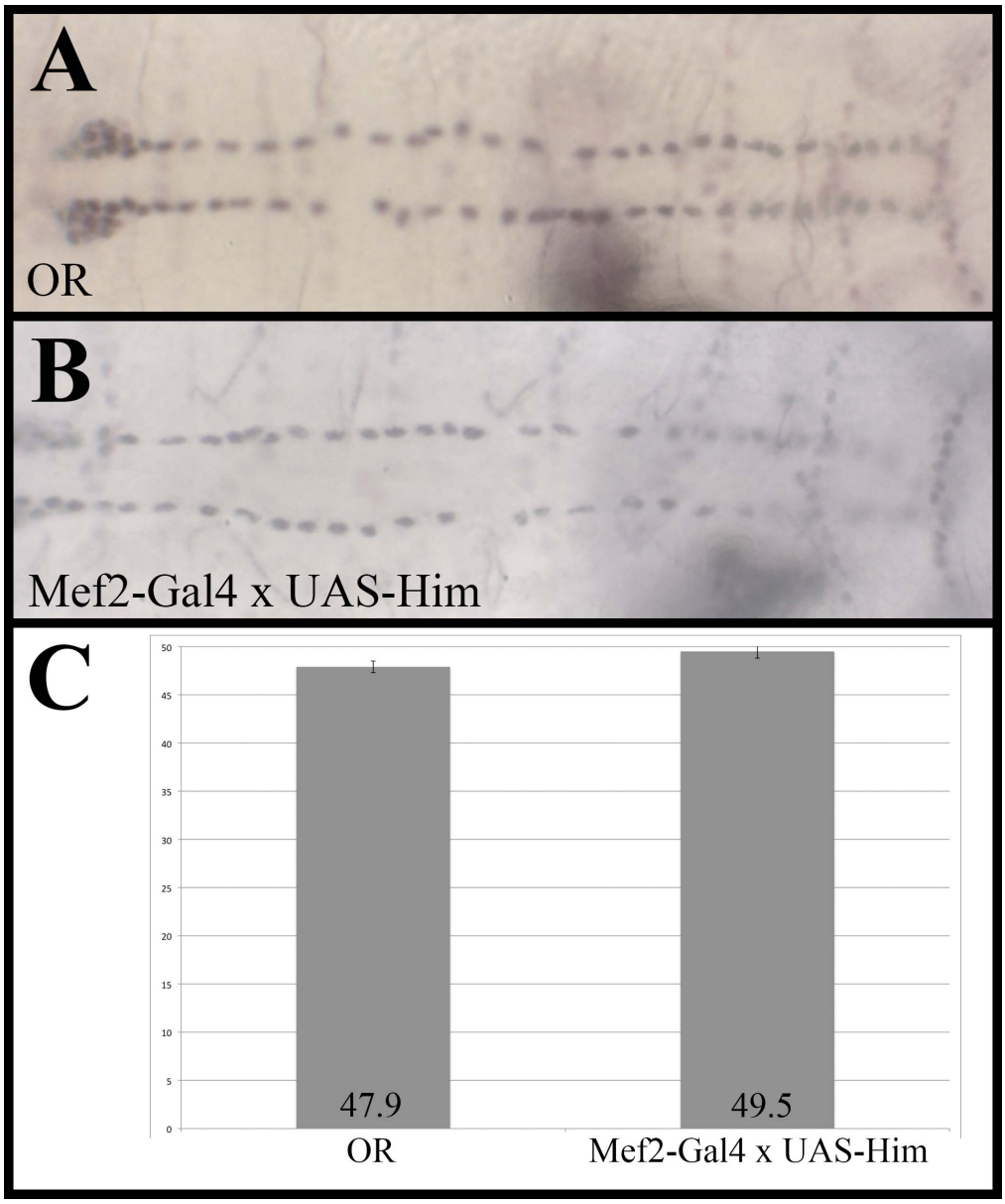


Figure 3.4.6: Expression pattern of Odd in embryos over-expressing *Him* in only the cardioblasts.
A Wild-type embryo; **B** Odd expression pattern in an embryo over-expressing *Him* under the control of the *Mef2-Gal4* driver; **C** Histogram showing the number of Odd-positive pericardial cells in embryos over-expressing *Him* in the cardioblasts. The error bars indicate the standard error, there is no statistically significant difference between wild-type and *Him* over-expressing embryos.

good chance of more than one factor *Him* and Mef2 interact with. This interaction seems to be similar in fashion to what has been described in the somatic mesoderm and that the introduction of *Him* into the cardioblasts alone is sufficient to interfere with the development of the myogenic cardioblasts.

3.4.2 *Him* over-expression phenotype in the larval heart

As for the *Him* null mutant animals, I followed the development of the pericardial cells in animals that over-express *Him* during the larval stages using the hand-GFP construct described in Sellin et al., 2006. I wanted to find out if the increase in pericardial cells I observed in the embryos over-expressing *Him* in the heart is maintained in the larvae or if this phenotype can be compensated for during larval development.

Figure 3.4.7 shows the behaviour of the pericardial cells in larvae over-expressing *Him* in all heart cells by the *hand-Gal4* driver, as marked by hand-GFP during the first larval instar. The pericardial cells of the first instar larvae over-expressing *Him* appear more crowded and closer together than in the wild-type. There also appears to be an increase in the number of pericardial cells but as described previously counting their number at this stage is not reliable. The cells themselves have the shape and appearance of wild-type pericardial cells and the supernumerary pericardial cells are distributed evenly along the length of the heart.

The same is true for the second instar larvae (see Figure 3.4.8). The increase in pericardial cells is obvious as well, but by this stage the extra cells seem to be located more towards the posterior part of the aorta and the heart proper. The pericardial cells around the anterior section of the aorta show a normal arrangement with similar large gaps between them as observed in wild-type larvae. The shape and size of all the pericardial cells are also very similar to the cells of the wild-type larvae. Pericardial cells counts reveal a statistically significant increase when over-expressing *Him*, mainly caused by the accumulation of supernumerary pericardial cells in the posterior sections of the heart.

Figure 3.4.9 shows the expression pattern of the hand-GFP reporter construct in third instar larvae over-expressing *Him*. The increase in pericardial cell number is

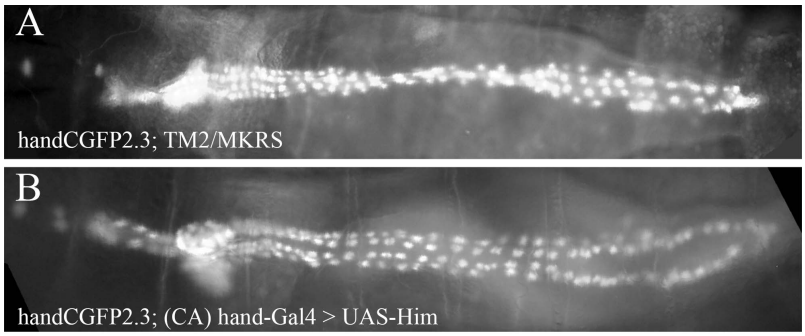


Figure 3.4.7: Expression pattern of the *hand-GFP* reporter construct in first instar larvae that over-express *Him* in the heart.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the GFP expression pattern.
A *handCGFP2.3* reporter construct, wild-type control; **B** hand-GFP expression pattern in a first instar larvae over-expressing *Him* under the control of the *hand-Gal4* driver.

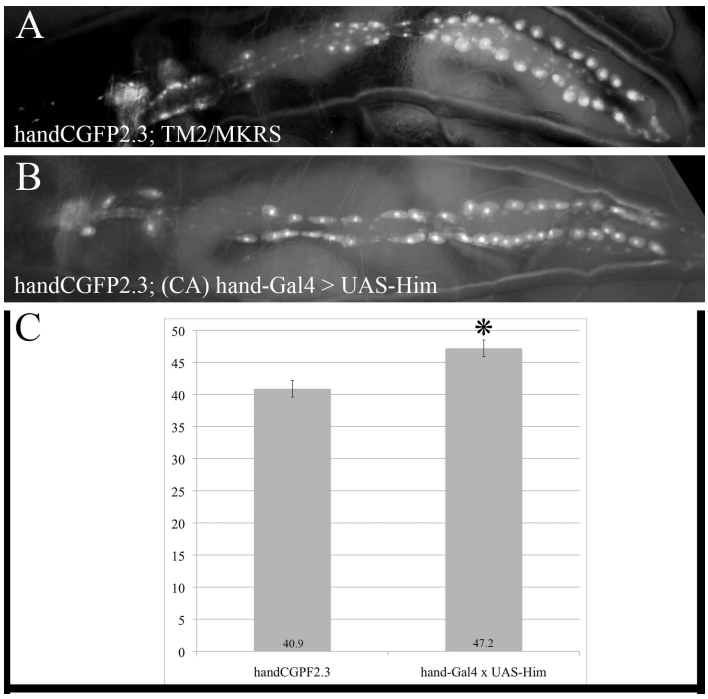


Figure 3.4.8: *Him* over-expression in the heart causes an increase in pericardial cells in second instar larvae.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the hand-GFP expression pattern and the number of GFP positive pericardial cells was counted.
A *handCGFP2.3* reporter construct, wild-type control; **B** *GFP* expression pattern in a second instar larvae over-expressing *Him* under the control of the *hand-Gal4* driver; **C** Histogram showing the numbers for the *handGFP-positive* pericardial cells in the controls (*handCGFP2.3*) and the second instar larvae that expresses *Him* ectopically in the heart. The error bars indicate the standard error and stars indicate a statistically significant increase compared to the wild-type.

maintained. However, the organisation of the pericardial cells is affected by the increase in number. The pericardial cells in the larvae that over-express *Him* are closer to each other, especially in the heart proper region. In an estimated 80 % of all larvae, the pericardial cells in this heart proper region are in two rows or individual hand-GFP positive pericardial cells are displaced laterally, disturbing the "pearls-on-a-string" arrangement that these cells usually have at this stage (see arrows in Figure 3.4.9 C). In the anterior part of the heart, the aorta, the gaps that appear between the pericardial cells from the second instar onwards appear shorter than in the wild-type larvae or are completely missing (see stars in Figure 3.4.9 C). This shows that the effect of over-expressing *Him* is not only restricted to the posterior part of the heart. Counting of the cells shows a statistically significant increase of more than ten pericardial cells per animal.

My observations show that if *Him* expression is maintained throughout larval development, the proper development of the pericardial cells is disturbed. It is currently unclear if this increase in pericardial cell number is due to the failure of the process that reduces pericardial cell numbers during the larval life or if this process does indeed occur but further pericardial cells are generated during this time as well. It is important for the larvae to generate the correct number of pericardial cells during embryogenesis. It would be interesting to see if the change in pericardial cell number has any effects on the physiology of the heart and/or the life expectancy of the flies as it has been shown that the pericardial hyperplasia in *dpp* mutant embryos leads to a decrease in the volume of pumped haemolymph (Johnson et al., 2007). Analysing the pattern of the heartbeat and its response to physiological stress is however beyond the scope of this work.

3.4.3 Conclusions for the *Him* phenotype

These data presented in this section show that *Him* is important for ensuring the specification of the correct number of all heart cells and especially the pericardial cells in the embryo. Table 3.4.1 includes all pericardial cell numbers mentioned in the previous two sections. My analysis has also shown that while *Him* expression is

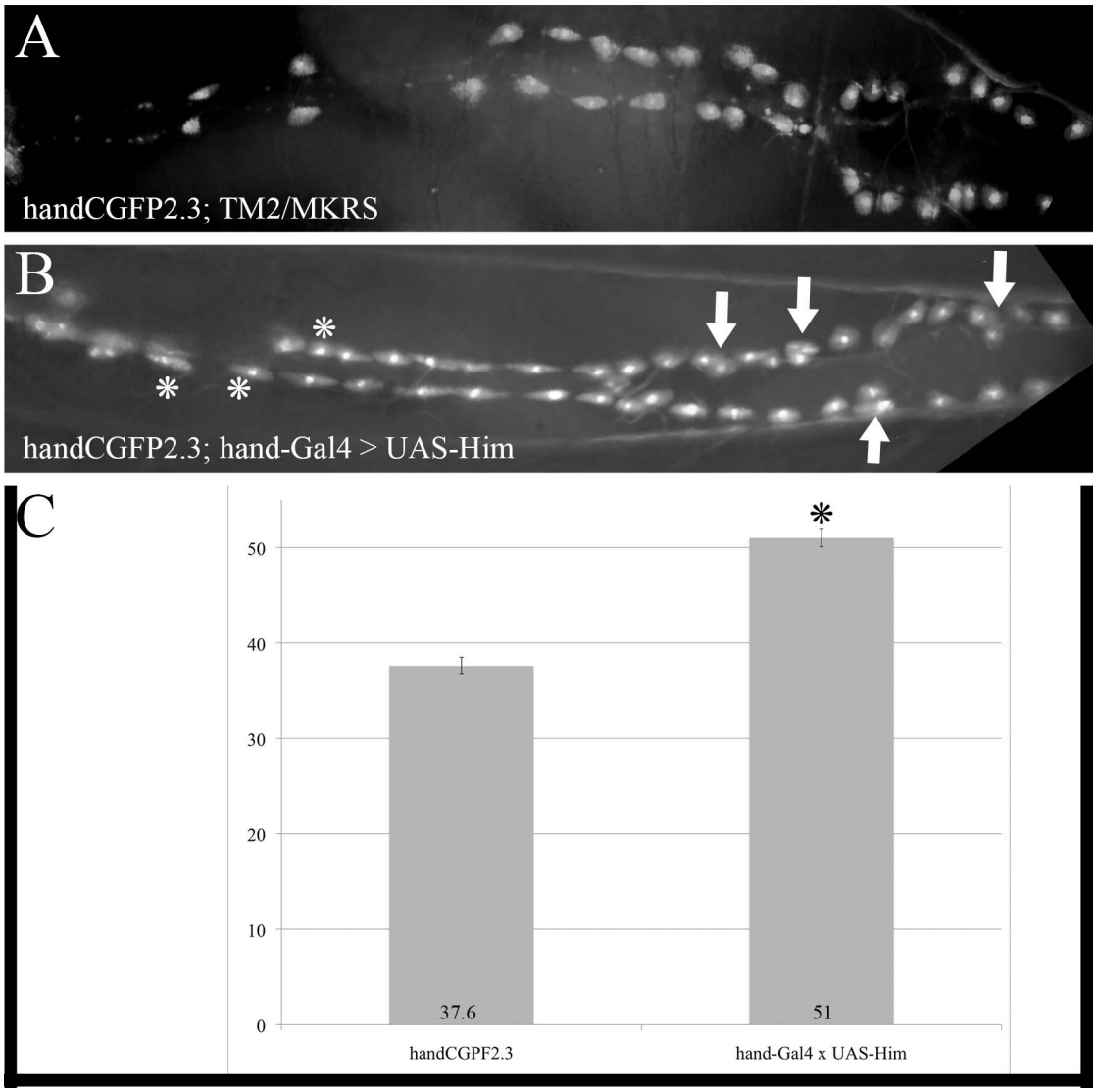


Figure 3.4.9: *Him* over-expression in the heart causes an increase in pericardial cells in third instar larvae.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the hand-*GFP* expression pattern and the number of GFP positive pericardial cells was counted.
A *handCGFP2.3* reporter construct, wild-type control; **B** Over-expression of *Him* using the *hand-Gal4* driver causes an increase in pericardial cells of third instar larvae and also disturbs their organisation. Arrows point to pericardial cells that are laterally displaced and stars indicate areas that should have gaps in the anterior aorta. **C** Histogram showing the numbers of handGFP-positive pericardial cells in the controls (*handCGFP2.3*) and the third instar larvae that express *Him* ectopically in the heart. The error bars indicate the standard error and the star indicate a statistically significant increase compared to the wild-type.

Table 3.4.1: Heart cell numbers obtained from all *Him* loss-of-function and gain-of-function experiments described in this section. OPCs: Odd-positive pericardial cells; EPCs: Eve-positive pericardial cells, CB: total cardioblast number, β3: β3tubulin-positive cardioblast, 2nd instar: number of hand-GFP positive pericardial cells in the second instar, 3rd instar: number of hand-GFP positive pericardial cells in the third instar, n/a: data not available, a grey underlay indicates a significant change in cell number when compared to wild type.

	pericardial cells		cardioblasts		larval stages	
	OPCs	EPCs	β3	CB	2nd instar	3rd instar
OR	47.9	27	52.2	71.8	n/a	n/a
handCGFP 2.3	48.5	27.4	n/a	n/a	45.9	38.3
handCGFP 3.1	48.1	26.5	n/a	n/a	43.4	37.6
Him 52	37.4	21.4	49.4	65.6	44.3	30.8
Him 74	43.6	22.1	46.4	65.4	32.6	33.4
Him 195	46.4	16	45.4	67.2	34.9	34.4
hand-Gal4 x UAS-Him	59.2	27.5	51.6	n/a	47.2	51
Mef2-Gal4 x UAS-Him	49.5	n/a	50.3	71.6	n/a	n/a
TinCΔ4-Gal4 x UAS-Him	n/a	n/a	51.3	72.1	n/a	n/a

necessary for the correct development of the pericardial cells, its exclusion from the cardioblasts is of importance for the correct development of these cells.

Too much *Him* expression in the heart leads to an increase in pericardial cells while the loss of *Him* leads to a loss of pericardial cells and a mild loss of cardioblasts. The loss of pericardial cells during embryogenesis is compensated for during larval life by the third instar. However, the gain of pericardial cells caused by over-expression of *Him* cannot be compensated for during larval life. These results point to *Him* being directly or indirectly involved in the regulation of the decrease of pericardial cells during larval development. It is possible that, if *Him* is expressed during larval life, this loss of pericardial cells cannot happen to the degree normal in wild-type larvae. This leads to the "normalized" appearance of pericardial cells number by the third instar. Johnson et al. (2007) have shown that an excess of pericardial cells is detrimental to the life expectancy and heart function in adult flies. As the pericardial cell number in the *Him* null mutants is within the wild-type range, this effect might be a partial explanation for the viability of the homozygous *Him* null mutant lines. However, continuous over-expression of *Him* in the pericardial cells also changes the pericardial cell number, resulting in third instar larvae that still show an excess of pericardial cells. Accordingly, if these flies were followed into adulthood, they should show a diminished heart function and life span.

These results point to two distinctive roles for *Him* during heart development. During embryogenesis, *Him* is necessary for the specification of the correct number of all heart cells, pericardial cells and cardioblasts. The phenotype described for the misexpression of *Him* in the heart in this thesis and in David Liotta's PhD thesis is reminiscent to that observed in embryos mutant for *decapentaplegic* (*dpp*, Johnson et al., 2007, Liotta, D. PhD thesis). In both cases the Eve-positive pericardial cells are not affected while the Odd-positive pericardial cells increase in number. In their paper Johnson et al., show that this increase of Odd-positive pericardial cells correlates with the expansion of Zfh-1 expression, which is under the control of Dpp. They also describe a reduction in size of some DO1 muscles and an increase in the space between the DO1 and DO2 muscles (Johnson et al., 2007). *Him* null mutants display a "relative decrease in average thickness for the DO2 muscle" and also a larger space between the DO1 and DO2 muscles (D. Hancock, PhD thesis). These phenotypes resemble each other and point towards the possibility that *Him* might also be involved in mediating the

response of *dpp*, similarly to Zfh-1. The possibility that *Him* is a target of Dpp signalling is discussed in chapter 6.

Him clearly also has an effect on the cardioblasts. It has been demonstrated that the hyperplasia of the Tin- and Odd-positive pericardial cells observed in *dpp* mutants is due to a loss of cells of the dorsal muscles (Johnson et al., 2007). My observations of no gain, but rather a loss, of cardioblasts in *Him* null mutant embryos show that the loss of *Him* affects both cardioblasts and pericardial cells in a very similar fashion and not that one cell type gains cells at the cost of the other. Both, the results from previous studies in this lab and the results of Johnson et al., point towards the dorsal muscle cells as the possibly affected tissue. This should be investigated with priority. Furthermore, it is also possible that *Him* expression is necessary to delineate the heart precursors from those for the visceral mesoderm as both develop from the dorsal mesoderm. The presumptive heart cells form under the direct influence of wingless (Riechman et al., 1997). The co-repressor Groucho has been shown to directly interact with dTcf of the Wingless/Wnt signalling pathway and with Brinker, a target of the Dpp signalling pathway (Cavallo et al., 1998; Lawrence et al., 2000; Brantjes et al., 2001 and Hasson et al., 2001). As a previous study in this lab has shown that Him and Groucho also interact physically and work together in the regulation of the activity of Mef2 in the somatic mesoderm this might be a possible link to these pathways (Liotta et al., 2007; Liotta, D. PhD thesis). These possibilities of the mechanistic function of *Him* still require further investigation.

At later stages in development the temporal regulation of the expression of *Him* is also very important. While Him needs to be present during early and mid embryogenesis as described above, it is equally important for the Him protein to not be present by the beginning of larval development to allow for the proper development of the pericardial cells. Disturbances in the expression of *Him* lead to a failure in the specification of the correct pericardial cell number during larval development. While the significance of this reduction is still mainly unclear, it has been shown that pericardial cell hyperplasia has a negative effect on the amount of haemolymph transported with each heart beat (Johnson et al., 2007). Further investigations need to be made into the physiology and function of the pericardial cells of the *Him* null mutant embryos and those over-expressing *Him* during larval life. In my opinion it is highly likely that one will be able to detect aberrant behaviour of these cells and in the over-all functionality of the heart in late larval stages and in the adult. We are currently also

investigating if the failure to reduce the number of larval pericardial cells happens at the expense of the ventral heart muscle that is generated during late larval and pupal life.

The data collected for the embryo and larvae pericardial cells numbers show that all available *Him* null mutants generally behave in the same way (Table 3.4.1). While the genetic mapping of the null mutants is not yet complete with regard to the their exact breakpoints, previous tests in our lab have clearly established all three lines as deficient for *Him* (D. Hancock, PhD thesis, J. Han and M. Taylor, unpublished data). Any differences in the scored exact cell numbers have to be attributed to human error, biological variation or the context of the *Him* deletions. The *Him 52* deletion line lacks a further four genes and there is a possibility that the other two lines have retained some functional regulatory elements and a small part of the 3'-UTR of the *Him* gene. These sequences might still be capable of producing gene products like microRNA or other (partial) RNA fragments that can affect neighbouring genes.

Him is thus implicated in several different functions within the development of the *Drosophila* heart. *Him* is necessary for the correct specification of all heart precursors and the further specification and differentiation of pericardial cells and cardioblasts during embryonic development. In the pericardial cells it also is necessary to lose *Him* expression at the end of embryogenesis to allow the reduction of the number of pericardial cells, the failure of which has been linked to physiological functions of the heart.

3.5 The role of *Mef2* in *Drosophila* heart development

Previous studies in this lab have shown that *Him* can regulate Mef2 activity in the somatic mesoderm and that different levels of Mef2 activity are important during development (Liotta et al., 2007; Elgar et al. 2008). Lowering Him levels using RNAi helps restore the phenotype of *Mef2* hypomorph embryos towards that of a wild-type embryo (Liotta et al., 2007). Similarly, the phenotype observed in the somatic mesoderm of *Drosophila* embryos over-expressing *Him* can be rescued toward a wild-type phenotype by co-over-expressing *Mef2* in the same animal. If this relationship between *Him* and Mef2 function is also of importance in the heart, the loss or gain of

Mef2 expression in the heart should result in a phenotype reminiscent of the phenotypes I have described for the gain or loss of *Him* expression.

The analysis of *Mef2* null mutants has shown that the cardioblasts and pericardial cells of the heart are formed but lack the expression of muscle structural genes likes *myosin heavy chain* (Bour et al., 1995; Lilly et al., 1995; Ranganayakulu et al., 1995; Gunthorpe et al., 1999). The cardioblasts in these animals do express ß3-tubulin in a regular pattern (Gunthorpe et al., 1999, D. Liotta, PhD thesis), the general pericardial cell population appears intact as assayed by the expression of the extra-cellular marker Pericardin and the Eve-positive pericardial cells are not affected by the loss of *Mef2* (Bour et al., 1995, Gunthorpe et al., 1999). There have not been any further studies of the effect the loss of *Mef2* has on the heart of the embryo or larvae.

To my knowledge, so far *Mef2* has not been over-expressed in the *Drosophila* heart with cardiac-specific drivers. Lin et al., 1997 used a selection of epidermal, pan-mesodermal, neuronal and heat-shock induced drivers (*69B*, *Gal4-1407* and *24B*) and describe "no noticeable effect on the viscera and the dorsal vessel". In a later study Gunthorpe et al. (1999) show an effect on the visceral muscles and the heart. In their study, the embryos were labelled for *myosin* expression and they observe a shorter heart and an accumulation of cardioblasts in the posterior part of the heart with an over-all reduction of the cardioblast number in the embryos over-expressing *Mef2*. There are no reports of the effect of over-expression of *Mef2* on the pericardial cells.

3.5.1 The heart cells in *Mef2* null mutant embryos

As none of the previously performed studies scored the heart cells of *Mef2* null mutant embryos and a closer analysis of the images in the studies done by Bour et al. and Gunthorpe et al. raised the possibility that the expression of *ß3-tubulin* might not be completely identical to that observed in wild-type embryos, I have repeated some of the already published experiments. In addition to this I also scored the Odd-positive pericardial cells.

As I did for the *Him* null mutant embryos, I scored the *Mef2* null mutant embryos for the number of ß3-tubulin-positive cardioblasts and the total number of cardioblasts using DIC optics (Figure 3.5.1). While the number of cardioblasts counted

using the DIC optics was similar to that of the wild-type embryos, I observed a marked drop in the number of ß3-tubulin-positive cardioblasts; this loss of ß3-tubulin-positive cardioblasts is statistically significant (see chart in Figure 3.5.1). The arrangement of the cardioblasts is also not quite as regular and orderly as they are in wild-type embryos. This reduction in the number of ß3-tubulin-positive cardioblasts has not been described before. As the total number of cardioblasts is unaffected, I interpret this result as follows: while Mef2 is not necessary for the specification of the cardioblasts, it is involved in the regulation of genes that contribute to the identity of the cells it is expressed in (Bour et al., 1995; Lilly et al. 1995; Ranganayakulu et al, 1995). However, according to my results this effect is not exclusively limited to the late expression of muscle structural genes but also affects the expression of ß3-tubulin.

Scoring the number of Odd-positive pericardial cells (see Figures 3.5.2) showed no change in Odd-positive pericardial cell number. There was also no change in the arrangement or size of the Odd-positive pericardial cells. The previously published experiments (Bour et al., 1995), describe a very similar pericardial phenotype based on the expression of Even-skipped and Pericardin. Thus, I have confirmed the previously described phenotype and extend the analysis to the Odd-positive pericardial cells and to exact cell numbers for all markers I tested. It appears that Mef2 has only a very minor role or even no importance for the correct specification and early differentiation of the pericardial cells of the embryonic heart, but that it is important for the regulation of genes involved in the late differentiation of the cardioblasts.

3.5.2 The effect of *Mef2* over-expression in the pericardial cells of the embryo

In order to achieve *Mef2* over-expression in the pericardial cells I have used two different Gal4 drivers: *Him-Gal4 L3-5* and *hand-Gal4*. *Hand-Gal4* is expressed in all known heart cells from embryonic late stage 12 onwards throughout the animals life while the *Him-Gal4* only drives expression in the pericardial cells of the embryo from stage 11/12 onwards to the end of embryogenesis but is not active during larval life. I used two different *UAS-Mef2* lines, one more strongly expressing line (*UAS-Mef2high*, Bour et al., 1995) and one less strongly expressing line (*10T4A*, Gunthorpe et al., 1999)

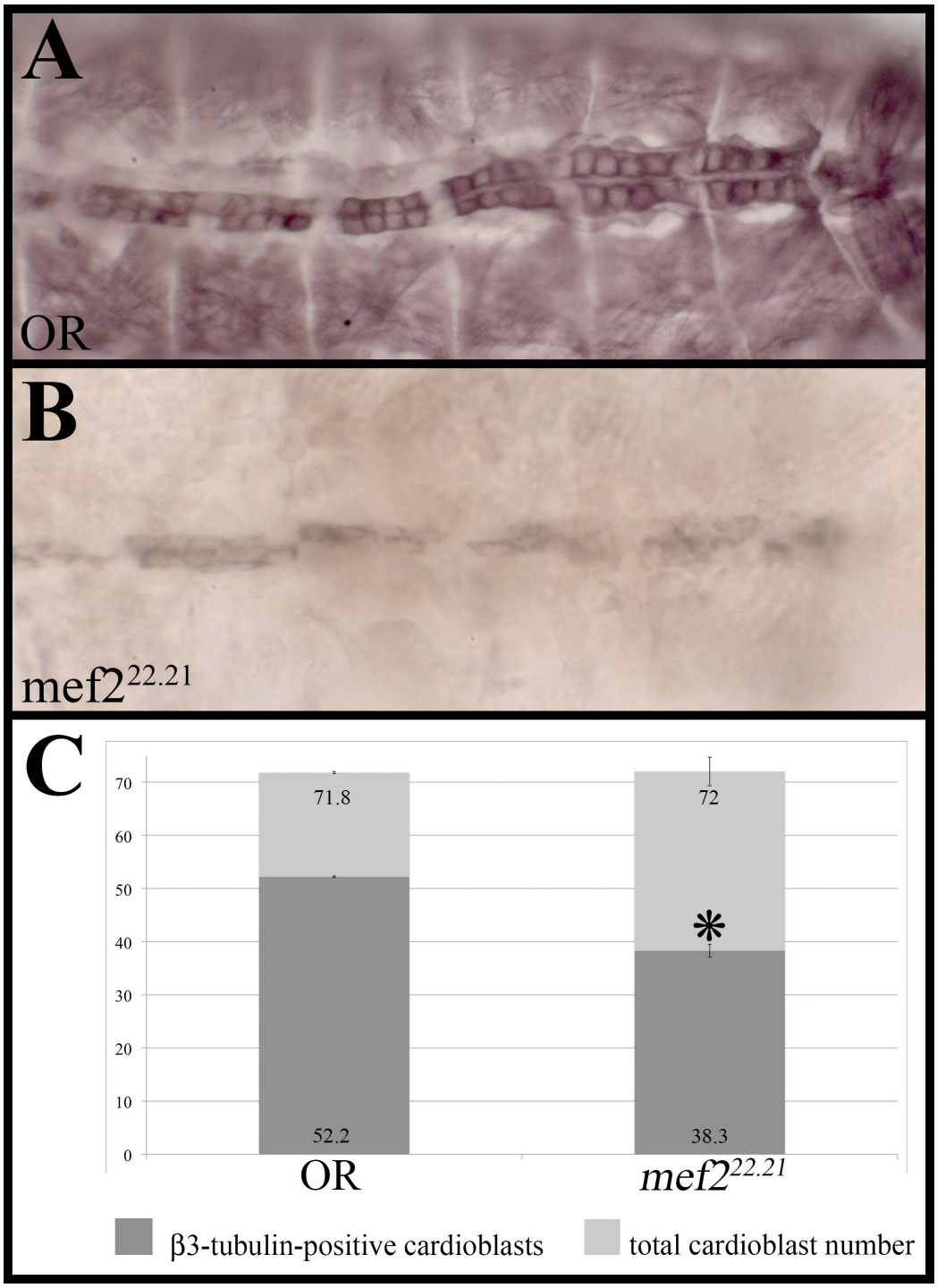


Figure 3.5.1: Expression pattern of ß3-tubulin in *mef2*$^{22.21}$ null mutant embryos.
Stage 16/17 embryos were labelled for the ß3-tubulin protein and the number of ß3-tubulin-positive cardioblasts was counted. Homozygous mutant embryos were chosen by absence of the *lacZ* balancer.
A Wild-type embryo; **B** ß3-tubulin expression in a homozygous *mef2*$^{22.21}$ mutant embryo; **C** Histogram showing the numbers of ß3-tubulin-positive cardioblasts and the total number of cardioblasts in the controls (*OR*) and in the *mef2*$^{22.21}$ embryos. The error bars indicate the standard error and stars indicate a statistically significant decrease compared to the wild-type.

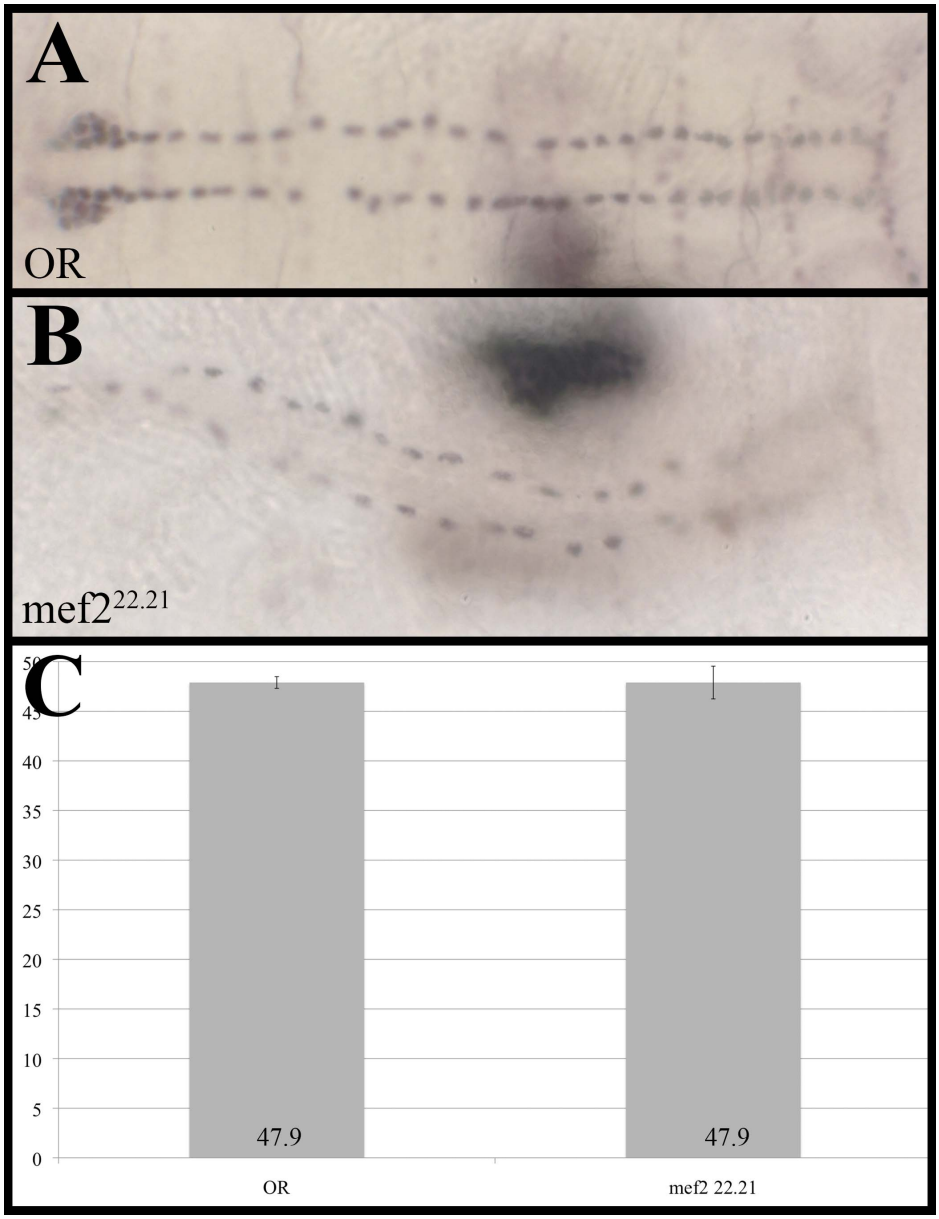


Figure 3.5.2: Expression pattern of Odd in *mef2*$^{22.21}$ null mutant embryos.
Stage 16/17 embryos were labelled for Odd protein and the number of odd-positive pericardial cells was counted. Homozygous mutant *mef2*$^{22.21}$ embryos were chosen through the absence of the *lacZ* balancer.
A Wild-type embryo; **B** Odd expression in a homozygous *mef2*$^{22.21}$ mutant embryo; **C** Histogram showing the numbers of Odd-positive pericardial cells in the controls (*OR*) and in the *mef2*$^{22.21}$ embryos. The error bars indicate the standard error; there is no statistically significant difference compared to the wild-type.

and tested all four possible combinations. As before I used Zfh-1 to gain a first insight into the behaviour of the pericardial cells when *Mef2* is over-expressed in these cells. As can be seen in Figure 3.5.3, distribution and organisation of the Zfh-1-positive pericardial cells is indistinguishable from the pattern observed in a wild-type embryo.

For the analysis in the embryos I again assayed the cell number for the same markers that I used in the *Him* null mutant and *Him* gain-of-function experiments: Odd, Eve and ß3-tubulin. Figure 3.5.4 shows the results I obtained for Eve. Again, it is not possible to distinguish any of the embryos over-expressing *Mef2* from the wild-type embryo based upon the expression pattern of Eve. Counting of the Eve-positive pericardial cells revealed furthermore that all four of the heart *Gal4 > UAS-Mef2* combinations have cell numbers very close to the number of Eve-positive pericardial cells I have counted in wild-type embryos and that there are no significant differences. Thus, ectopic expression of *Mef2* in the pericardial cells does not affect the number and organisation of the Eve-positive pericardial cells in the embryo.

As can be seen in Figure 3.5.5, the expression pattern of Odd protein in these *Mef2* over-expressing embryos it also indistinguishable from the wild-type. The organisation of the Odd-pericardial cells is also the same in both, wild-type and *Mef2* over-expressing embryos. The graph in 3.5.5 F shows that the number of Odd-positive pericardial cells is close to that observed in wild-type embryos in all *Mef2* over-expressing combinations and that there are no statistically significant differences between the embryos over-expressing *Mef2* and the wild-type embryos. Thus, the ectopic expression of *Mef2* also has no influence on the Odd-positive pericardial cells.
I again used ß3-tubulin to visualize the cardioblasts. Figure 3.5.6 shows these results. Here it is important to note that the expression of ß3-tubulin is not affected in *Mef2* loss-of-function mutants according to previous studies, which did not rely on accurate cell counts while I have observed a reduction in the number of ß3-tubulin-positive cardioblasts (figure 3.5.1; Bour et al., 1995; Lilly et al., 1995; Gunthorpe et al., 1999; PhD thesis, D. Liotta). If *Mef2* is over-expressed within the heart, I see a mild organisational phenotype if the more strongly expressing form of the *UAS-Mef2* construct is over-expressed under the control of the *hand-Gal4* driver. The cardioblasts do not line up as neatly as in the wild-type and occasionally bunching up of cardioblasts occurs (see arrows in Figure 3.5.6). Counting of the ß3-tubulin positive cardioblasts showed that all four different *Mef2* over-expression combinations are very close to the

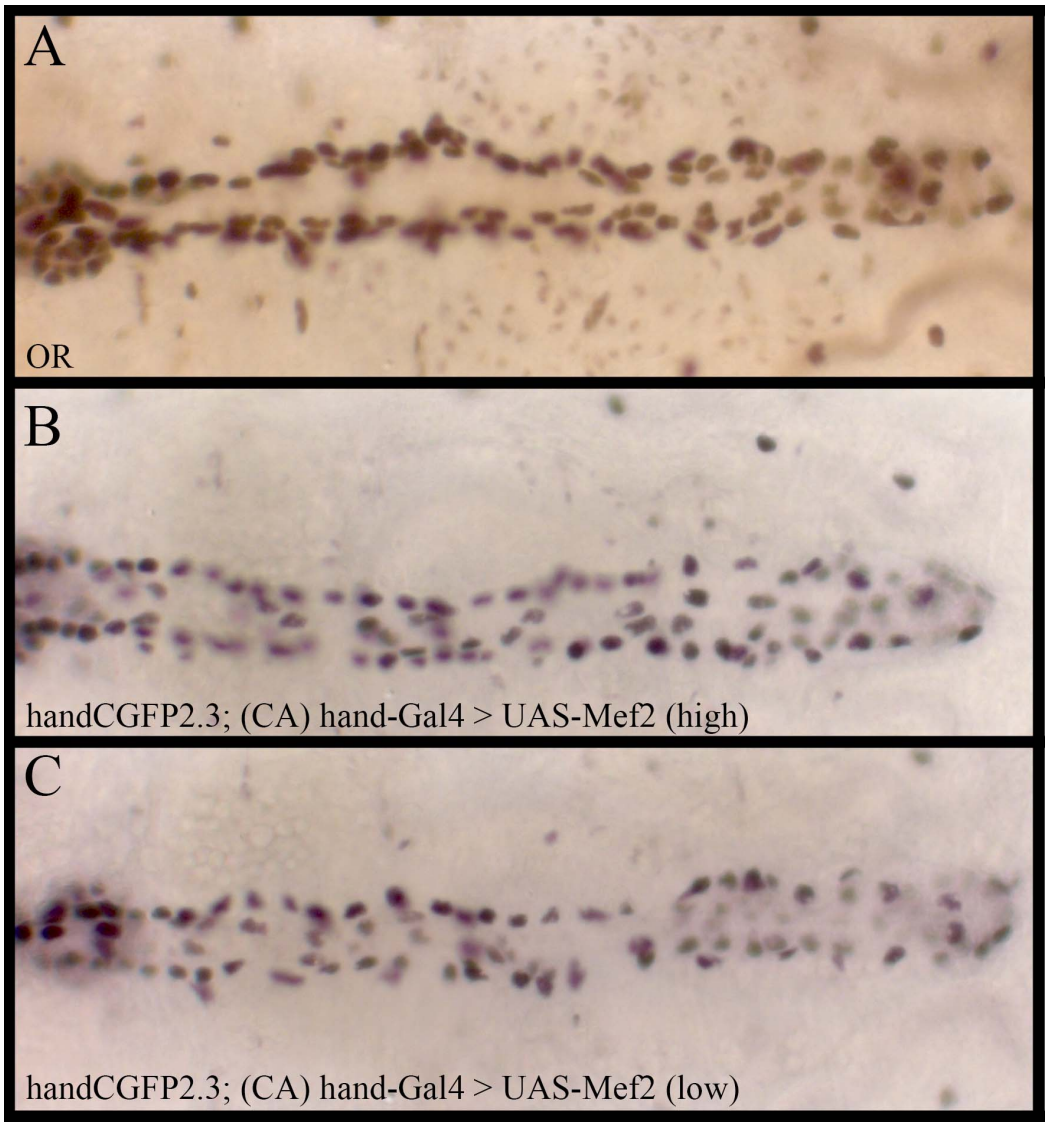


Figure 3.5.3: Zfh-1 expression in embryos over-expressing *Mef2* in the pericardial cells.
Stage 16/17 embryos were labelled for the distribution of the Zfh-1 protein.
A *OR* wild-type embryo; **B** Zfh-1 expression in an embryo if the *hand-Gal4* driver controls the stronger expressing form of the *UAS-Mef2* construct; **C** Zfh-1 expression pattern in an embryo if the *hand-Gal4* driver controls the less strong expressing *UAS-Mef2* construct.

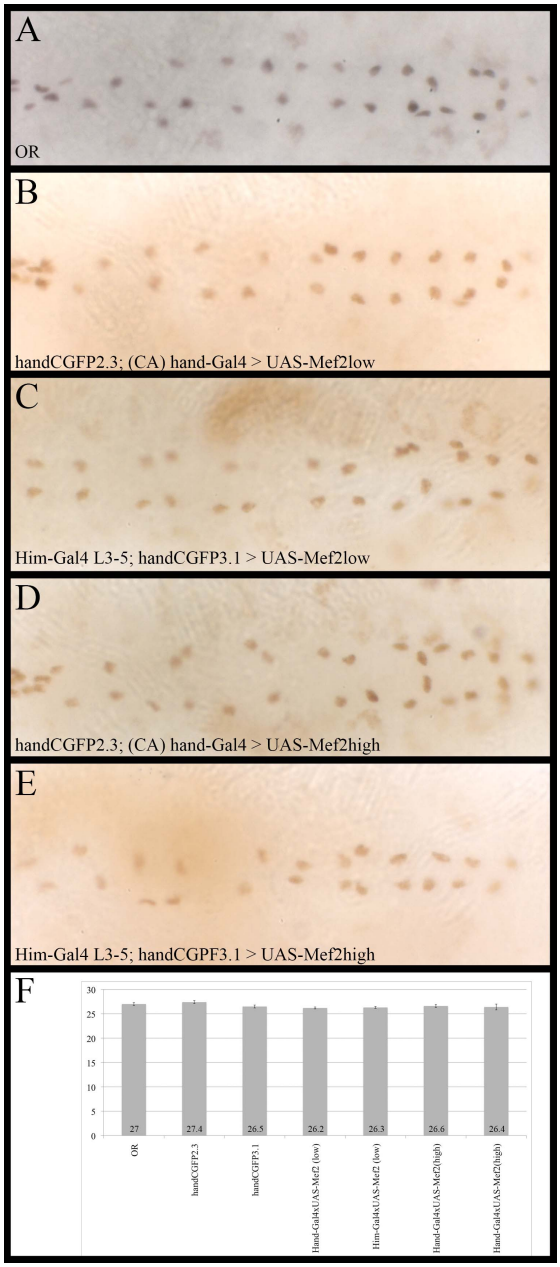


Figure 3.5.4: Expression pattern of Eve in embryos that over-express *Mef2* in the pericardial cells.
Stage 16/17 embryos were labelled for the Eve protein and the number of Eve-positive pericardial cells was counted.
A Wild-type embryo; **B** Eve expression in an embryo if the *hand-Gal4* driver controls the less strong expressing form of the *UAS-Mef2* construct; **C** Eve expression in an embryo if the *Him-Gal4* driver controls the less strong expressing form of the *UAS-Mef2* construct; **D** Eve expression in an embryo if the *hand-Gal4* driver controls the stronger expressing form of the *UAS-Mef2* construct; **E** Eve expression in an embryo if the *Him-Gal4* driver controls the stronger expressing form of the *UAS-Mef2* construct; **F** Histogram showing the numbers of Eve-positive pericardial cells in all three controls (*OR, handCGFP2.3, handCGFP3.1*) and the different conditions used to over-express *Mef2*. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

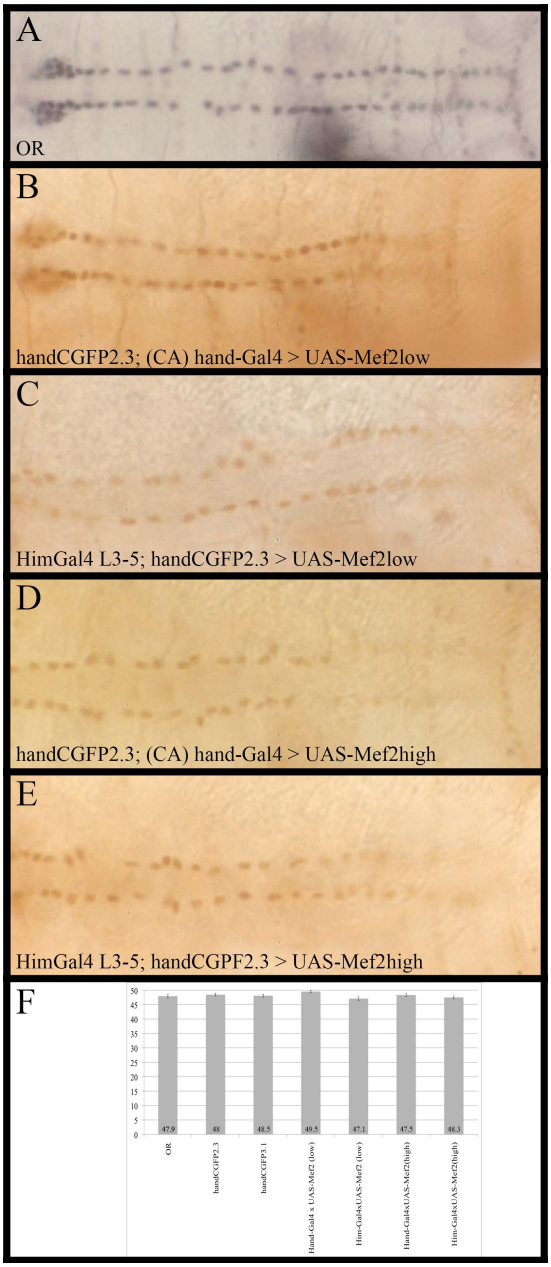


Figure 3.5.5: Expression pattern of Odd in embryos expressing ectopic *Mef2* in the pericardial cells.
Stage 16/17 embryos were labelled for the Odd protein and the number of Odd-positive pericardial cells was counted.
A Wild-type embryo; **B** Odd expression in an embryo if the *hand-Gal4* driver controls the less strong expressing form of the *UAS-Mef2* construct; **C** Odd expression in an embryo if the *Him-Gal4* driver controls the less strong expressing form of the *UAS-Mef2* construct; **D** Odd expression in an embryo if the *hand-Gal4* driver controls the stronger expressing form of the *UAS-Mef2* construct; **E** Odd expression in an embryo if the *Him-Gal4* driver controls the stronger expressing form of the *UAS-Mef2* construct; **F** Histogram showing the numbers for the Odd-positive pericardial cells in all three controls (*OR, handCGFP2.3, handCGFP3.1*) and the different conditions used to over-express *Mef2*. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

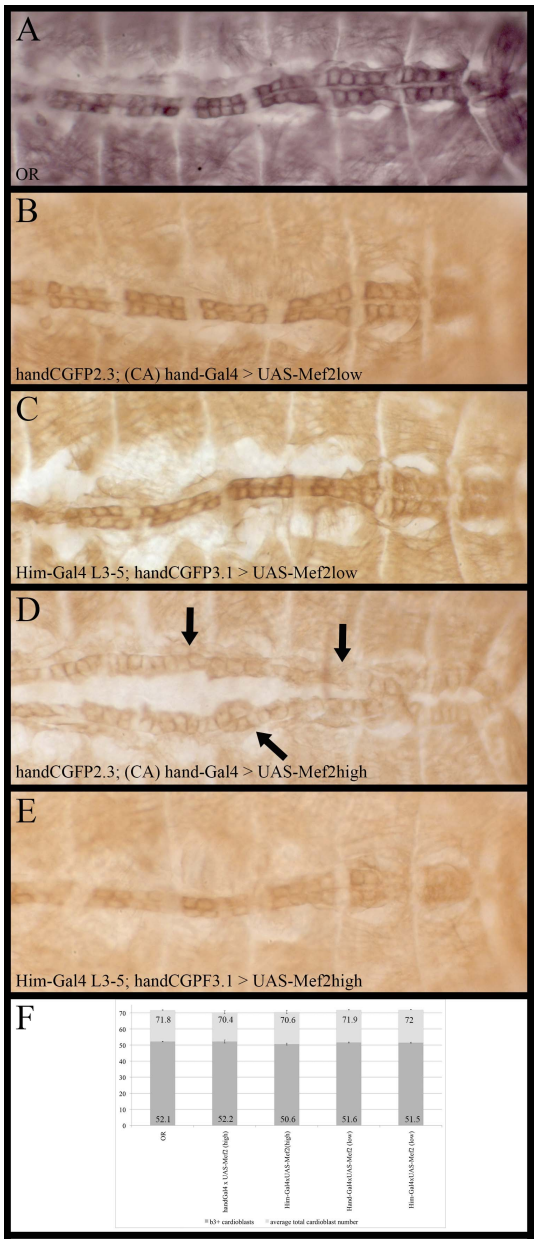


Figure 3.5.6: Expression pattern of ß3-tubulin in embryos that over-express *Mef2* in the pericardial cells.

Stage 16/17 embryos were labelled for the ß3-tubulin protein and the number of ß3-tubulin-positive cardioblasts was counted.

A Wild-type embryo; **B** ß3-tubulin expression in an embryo if the *hand-Gal4* driver controls the less strong expressing form of the *UAS-Mef2* construct; **C** ß3-tubulin expression in an embryo if the *Him-Gal4* driver controls the less strong expressing form of the *UAS-Mef*2 construct; **D** ß3-tubulin expression in an embryo if the *hand-Gal4* driver controls the stronger expressing form of the *UAS-Mef2* construct; **E** ß3-tubulin expression in an embryo if the *Him-Gal4* driver controls the stronger expressing form of the *UAS-Mef2* construct; **F** Histogram showing the numbers of ß3-tubulin-positive and total number of cardioblasts in the controls and the different conditions used to over-express *Mef2*. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

ß3-tubulin-positive cardioblast number observed in wild-type embryos and show no significant differences to the number of ß3-tubulin-positive cardioblasts in wild-type embryos. The total number of cardioblasts is also the same as in wild-type embryos (see Figure 3.5.6 F).

It must be noted that ß3-tubulin is only expressed in four of the six cardioblasts per hemisegment, thus the above analysis excludes a direct analysis the two Svp-positive cardioblasts of each hemisegment. I have, however counted all cardioblasts using DIC optics and not observed a change in this number either. In order to answer this question completely, an antibody stain for Svp, Doc, Mef2 or Myosin would help identify the fate of the two remaining cardioblasts. As the Zfh-1 expression pattern of the embryos over-expressing *Mef2* under the control of the *hand-Gal4* driver is very close to that of wild-type embryos (Figure 3.5.3), I believe that it is relatively unlikely that the Tin-positive pericardial cells are affected by the over-expression of *Mef2*.

3.5.3 The effect of *Mef2* over-expression in the larval heart.

I continued the analysis of the pericardial cells in animals over-expressing *Mef2* into the larval stages utilizing the same hand-GFP reporter construct as before. For an overview of all cell numbers of this section please refer to Table 3.5.1. While I did not observe any phenotype except for a mild organisational phenotype during the embryonic development of the heart, my analysis of the function of *Him* indicates that *Him* has an importance for the correct development of the larval pericardial cells as well. I was interested to see if this could be disturbed by continued *Mef2* expression in the pericardial cells once *Him* is not expressed in these cells anymore. Such a result would further point towards a connection between *Him* expression and Mef2 function and the importance of *Him* expression to prevent Mef2 function in the pericardial cells during embryogenesis. As I have shown previously and as is published, there is a marked drop in the number of pericardial cells between the first and third instar of wild-type animals and this step is disturbed when *Him* is missing or over-expressed. For these experiments of over-expressing *Mef2* in the larvae, I have used the same drivers as before*: Him-Gal4 L3-5* and *hand-Gal4*.

Figure 3.5.7 shows the behaviour of the pericardial cells in the early first instar larvae when *Mef2* is over-expressed under the control of *hand-* or *Him-Gal4*. The organisation and cell number of the pericardial cells are indistinguishable from that of wild-type larvae at this stage. Their shape and size is also virtually identical to that of the wild-type animals. As I was unable to reliably analyse the exact cell numbers for this developmental stage, I have no cell number counts for this and have to rely on these observations only.

During the second larval instar some of the cells begin to change their shape (see arrows in Figure 3.5.8 and Table 3.5.2). Some of the pericardial cells lose their typical round and globular shape. The percentage of these aberrant pericardial cells in relation to the total number of hand-GFP positive pericardial cell depends on the specific combination of Gal4-driver and UAS-construct used. I have observed cell changes in 85 to 100 % of all scored larvae (see Table 5.3.2 and stars in Figure 3.5.8). There are two different changes in cell shape, one is an accumulation of several small nuclei in clusters and other individual cells remain large but lose their round, globular shape.

The small, clustered nuclei have the same hand-GFP intensity as the "normal" pericardial cells but can be as small as half the size of the nuclei of the wild-type cells. There are usually at least two or three nuclei in a cluster and these clusters occur more readily towards the anterior end of the heart. Other hand-GFP positive pericardial cells begin to elongate and form cytoplasmic protrusions (see arrows in Figure 3.5.8 D and E), losing the typical globular shape.

Counting of all GFP-positive pericardial cells in these experiments revealed that for three of the four *Mef2* over-expression experiments the number of hand-GFP-positive pericardial cells is not significantly different to the number I counted in wild-type larvae; only if the stronger expressing form of the *UAS-Mef2* construct is expressed under the control of the *Him-Gal4* driver is there a statistically significant increase in hand-GFP positive pericardial cells (see figure 3.5.8 F). For this developmental stage, these numbers include the aberrantly shaped pericardial cells. This is because cells were often not easily identifiable as either wild-type or spindly-shaped cells, thus not allowing for a reproducible count of how exactly how many non-globular cells are present in one larva. This change of pericardial cell type, arrangement and number is especially interesting considering that so far no phenotype has been described for the embryonic pericardial cells if *Mef2* is over-expressed and that the expression of the

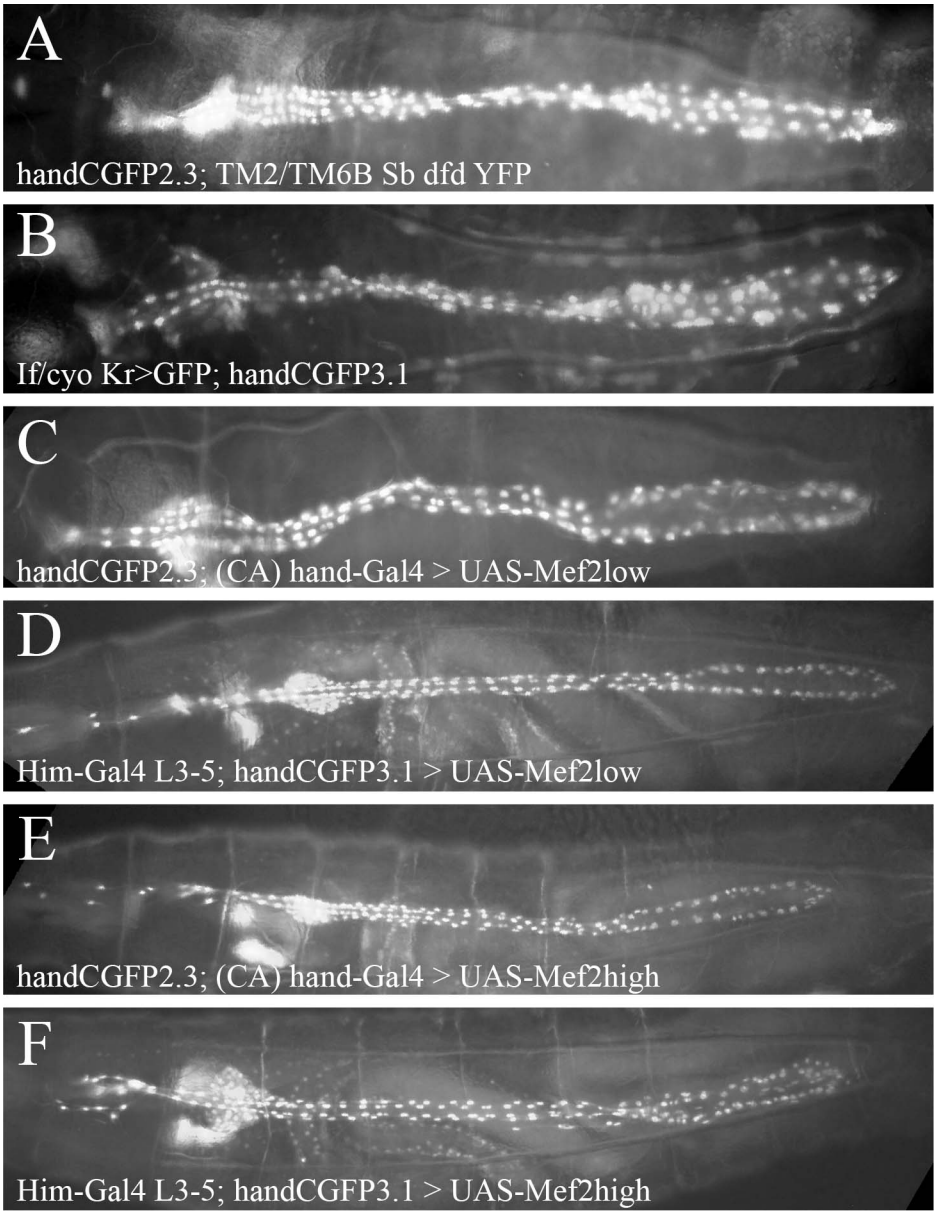


Figure 3.5.7: Expression pattern of the hand-GFP reporter construct in first instar larvae over-expressing *Mef2* in the pericardial cells.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the GFP expression pattern.
A *handCGFP2.3* reporter construct, wild-type control; **B** *handCGFP3.1* reporter construct, wild-type control; **C** GFP expression pattern in a first instar larvae expressing the weaker *UAS-Mef2* construct under the control of the *hand-Gal4* driver; **D** GFP expression pattern in a first instar larvae expressing the weaker *UAS-Mef2* construct under the control of the *Him-Gal4* driver; **E** GFP expression pattern in a first instar larvae expressing the stronger *UAS-Mef2* construct under the control of the *hand-Gal4* driver; **F** GFP expression pattern in a first instar larvae expressing the stronger *UAS-Mef2* construct under the control of the *Him-Gal4* driver.

Table 3.5.1: Heart cell numbers obtained from all *Mef2* loss-of-function and gain-of-function experiments described in this section. OPCs: Odd-positive pericardial cells; EPCs: Eve-positive pericardial cells, CB: total cardioblast number, β3: β3tubulin-positive cardioblast, n/a: data not available, a grey underlay indicates a significant change in cell number when compared to wild type.

	pericardial cells		cardioblasts		larval stages	
	OPCs	EPCs	β3	CB	2nd instar	3rd instar
OR	47.9	27	52.2	71.8	n/a	n/a
handCGFP 2.3	48.5	27.4	n/a	n/a	45.9	38.3
handCGFP 3.1	48.1	26.5	n/a	n/a	43.4	37.6
mef2$^{22.21}$	47.9	n/a	38.3	72	n/a	n/a
hand-Gal4 x UAS-Mef2low	49.5	26.2	52.2	70.4	43.5	42.2
Him-Gal4 x UAS-Mef2low	47.1	26.3	50.6	70.6	47	51.7
hand-Gal4 x UAS-Mef2high	47.5	26.6	51.6	71.9	44.8	59.9
Him-Gal4 x UAS-Mef2high	48.3	26.4	51.5	72	52.2	44.8

Table 3.5.2: Percentage of scored *Mef2* over-expressing larva that contain aberrantly shaped pericardial cells during larval life.

	handCGFP 2.3	*hand-Gal4 x UAS-Mef2low*	*Him-Gal4 x UAS-Mef2low*	*hand-Gal4 x UAS-Mef2high*	*Him-Gal4 x UAS-Mef2high*
1st instar	0 %	0 %	0 %	0 %	0 %
2nd instar	5 %	80 %	85 %	90 %	100 %
3rd instar	35 %	85 %	100 %	100 %	100 %

Table 3.5.3: Average numbers of wild-type and aberrantly shaped pericardial cells per animal (third instar). PCs: pericardial cells

	handCGFP 2.3	*hand-Gal4 x UAS-Mef2low*	*Him-Gal4 x UAS-Mef2low*	*hand-Gal4 x UAS-Mef2high*	*Him-Gal4 x UAS-Mef2high*
globular PCs	37	32.4	44.6	44.5	26.4
non-globular PCs	0.6	9.8	7.1	15.4	18.4
% of non-globular PCs per animal	1.6	23.2	13.7	25.7	41.1

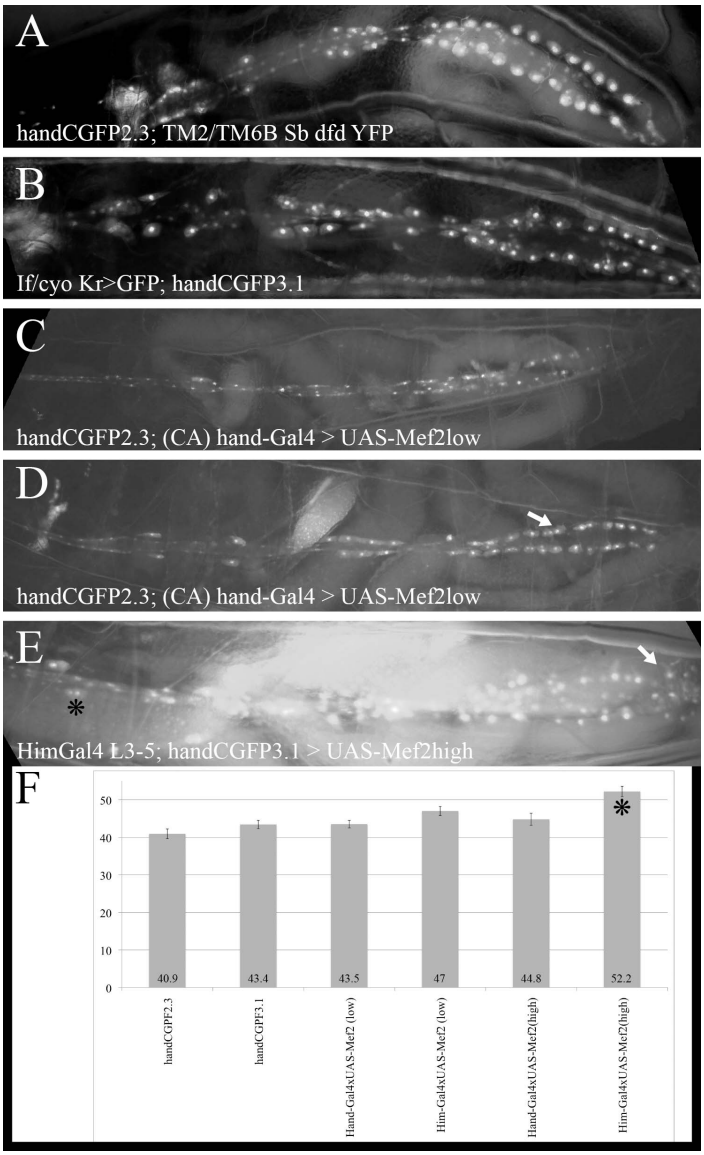


Figure 3.5.8: Expression pattern of the *hand-GFP* reporter construct in second instar larvae expressing ectopic Mef2 in the pericardial cells.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the GFP expression pattern and the number of GFP-positive pericardial cells was counted. Arrows point to elongated non-globular pericardial cells and stars indicate closely associated nuclei.
A *handCGFP2.3* reporter construct, wild-type control; **B** *handCGFP3.1* reporter construct, wild-type control; **C** GFP expression pattern in a second instar larvae expressing the weaker *UAS-Mef2* construct under the control of the *hand-Gal4* driver; **D** GFP expression pattern in a second instar larvae expressing the weaker *UAS-Mef2* construct under the control of the *Him-Gal4* driver; **E** GFP expression pattern in a second instar larvae expressing the stronger *UAS-Mef2* construct under the control of the *Him-Gal4* driver; **F** Histogram showing the numbers of hand-GFP-positive pericardial cells in the two controls (*handCGFP2.3, handCGFP3.1*) and the different conditions used to over-express *Mef2* in the second instar. The error bars indicate the standard error and stars indicate a statistically significant difference compared to wild-type.

Him-Gal4 L3-5 driver is limited to the embryonic stages only (see section 3.2.2 and Appendix).

Third instar larvae show an even more pronounced change in cell shape (Table 5.3.2 and 5.3.3). Figure 3.5.9 shows the expression pattern of the hand-GFP construct in third instar larvae over-expressing *Mef2*. Counts of the number of hand-GFP-positive pericardial cells demonstrate a significant increase in these numbers for three of the four experiments; only in the animals that drive the strong version of the *UAS-Mef2* construct under the control of the *Him-Gal4* driver there is no significant change in pericardial cell number (Figure 3.5.9 G) These animals showed an increase in pericardial cell number in the second instar (Figure 3.5.8 G).

Table 3.5.3 shows the average number of globular and non-globular pericardial cells in the third instar larvae of these experiments. As can be seen in the images in Figure 3.5.9 C', D', E' and F', a good part of the pericardial cells in each animal change their cell shape from the wild-type globular pericardial cell to a more elongated cell with more than one or two cytoplasmic protrusions (see arrows). There are also a considerable number of animals where I detect clusters of nuclei as I already described for the second instar (see stars in Figure 3.5.9) This "clumping" together of several nuclei has so far not been described in the literature and I can only detect it in approximately 35 % of the third instar wild-type larvae I analysed (see Table 3.5.2) and a much lower percentage of pericardial cells per animals is affected by this change in cell shape in these wild-type larvae than in Mef2-overexpressing animals (Table 3.5.3 and figure 3.5.9).

The amount of pericardial cells that change their cell shape per animal is tied in with the strength of the *UAS-Mef2* construct and driver used (Table 3.5.3). If the strong version of the *UAS-Mef2* is expressed, all of the larvae assayed show altered cell shapes but the used driver determines the percentage of aberrantly shaped cells per larva. If the *hand-Gal4* driver is used, on average 25.7 % of the hand-GFP positive pericardial cells of an animal have a non-globular shape. This percentage rises to 41.1 % if the *Him-Gal4* driver is used.

The different expression strength of the two *UAS-Mef2* constructs is illustrated by the decrease in phenotype frequency. If the same Gal4 driver but the less strong expressing *UAS-Mef2* construct is used, 85 % (for the *hand-Gal4* driver) or 100 % (for

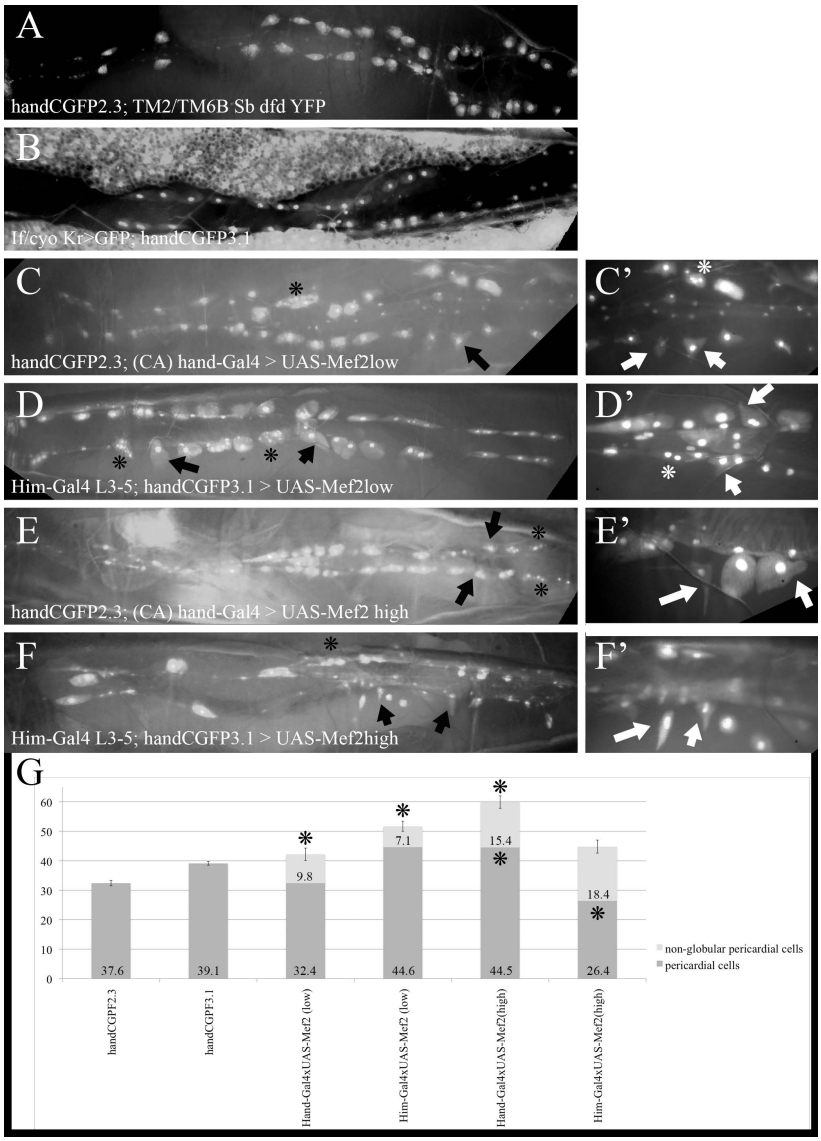


Figure 3.5.9: Expression pattern of the *hand-GFP* reporter construct in third instar larvae over-expressing *Mef2* in the pericardial cells.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the GFP expression pattern and the number of GFP-positive pericardial cells was counted. C' to F' show a section of the same larva shown in C to F at a higher magnification. Arrows point to elongated non-globular pericardial cells and stars indicate closely associated nuclei.
A *handCGFP2.3* reporter construct, wild-type control; **B** *handCGFP3.1* reporter construct, wild-type control; **C** GFP expression pattern in a third instar larvae expressing the weaker *UAS-Mef2* construct under the control of the *hand-Gal4* driver; **D** GFP expression pattern in a third instar larvae expressing the weaker *UAS-Mef2* construct under the control of the *Him-Gal4* driver; **E** GFP expression pattern in a third instar larvae expressing the stronger *UAS-Mef2* construct under the control of the *hand-Gal4* driver; **F** GFP expression pattern in a third instar larvae expressing the stronger *UAS-Mef2* construct under the control of the *Him-Gal4* driver; **G** Histogram showing the numbers of hand-GFP-positive pericardial cells in the two controls (*handCGFP2.3, handCGFP3.1*) and the different conditions used to over-express *Mef2* in the third instar. The error bars indicate the standard error and the star indicates a statistically significant difference compared to wild-type.

the *Him-Gal4* driver) of animals contain differently shaped cells (Table 3.5.2) and the percentage of non-globular pericardial cells per animal is also lower (23.2 and 13.7 %, respectively). These data illustrate that both the time and location of expression and the strength of the used UAS construct have an influence on the observed phenotype. However, all third instar larvae show the same changes in cell shape, proving that this is indeed a phenotype caused by the introduction of *Mef2* into the embryonic pericardial cells.

I have not been able to determine if the clusters of nuclei are indeed one multi-nucleated cell or a close aggregation of several smaller cells, as all these assays were done in whole-mount live larvae and the cuticle and trachea of the larvae interfere with the resolution of the image at high power magnification. For this, it would be advisable to dissect larvae and then visualize the cell membranes in the dissected animals together with the hand-GFP reporter construct.

As the increase in pericardial cell number and the change in cell shape happens in the later larval stages and I did not observe any embryonic phenotype, it is possible that these cells are generated from the pericardial cells that are lost in these stages in a wild-type situation. Indeed, if one subtracts the number of misshapen cells from the total pericardial cell numbers, the resulting numbers for normal, globular shaped pericardial cells are not significantly different to the numbers of these cells in the wild-type if the animals carry the less strongly expressing *UAS-Mef2* construct (Figure 3.5.9 G). The stronger expressed UAS-Mef2 construct has a different effect on the number of pericardial cells. If this construct is driven by *hand-Gal4*, even after subtracting the misshapen cells from the total pericardial cell number indicates an increase in the wild-type shaped cells. This indicates that *Mef2*, if expressed strongly in both the cardioblasts and pericardial cells in the embryo and larva leads not only to the observed changes in cell shape but also to an increase in wild-type like pericardial cells by the third instar. In order to determine if this observation is due to either an increase in pericardial cells or a less severe decrease of these cells during larval development, exact pericardial cell numbers for first instar larvae are necessary. As mentioned before, due to the expression pattern of the hand-GFP reporter construct I used, counts of this stage have proven to be unreliable.

If the strong version of the *UAS-Mef2* construct is expressed by *Him-Gal4* the observed phenotype is altered again. The total number of counted hand-GFP-positive pericardial cells (including the aberrantly shaped cells) is not statistically significant

different to that of the control larvae (Figure 3.5.9 G). Consequently, subtracting the number of misshapen hand-GFP-positive pericardial cells leads to a decrease in number of wild-type shaped hand-GFP-positive pericardial cells. This indicates that the levels at which *Mef2* is over-expressed in the *Him* expression pattern are of importance. Introduction of *Mef2* into the embryonic pericardial cells leads to a change in cell shape and behaviour (see later) in a proportion of the pericardial cells. However, if *Mef2* is over-expressed at strong levels and in the expression pattern of *Him*, the phenotype differs from these previously described phenotypes and the misshapen cells are most likely formed at the expense of the wild-type pericardial cells. It is possible that the loss of normal pericardial cells is an extension of the observed cell-shape phenotype which if extreme might lead to apoptosis. Further investigation is needed to establish if this loss of normal pericardial cells is due to apoptosis or if these cells contribute to other tissues and/or organs.

Interestingly, over-expression of *Mef2* (at high or low levels) during the embryonic development only is sufficient to derail the normal development of the pericardial cells during larval life. And the observed phenotypes are more severe with the earlier of the two used drivers. (For an overview of all scored pericardial cell numbers refer to Table 3.5.1.) This stresses the importance of preventing any *Mef2* function in the pericardial cells during embryogenesis. *Him* is known to counteract Mef2 function in other mesodermal tissues.

In addition to the changes in cell shape, I have also observed individual contractions of these non-globular but hand-GFP positive cells (refer to the movies on the DVD submitted with this dissertation). These contractions occur intermittently in bundled bursts of activity. The majority of the contractions follow each other in very short and fast succession until all activity again ceases for an indeterminate amount of time. As far as I have been able to tell from visual observations under the microscope, these contractions are independent of body wall and gut contractions (refer to the movies on the DVD with this dissertation). This is surprising as pericardial cells are defined as non-myogenic cells and are not known to express any myogenic markers, nor has any contractile activity been described for them. I have included several movies on a DVD showing this phenotype with this thesis. While not every animal contained these contractile pericardial cells, animals usually had more than one of these very aberrant pericardial cells. These contractions are more likely to happen in animals that carry the

stronger expressing version of the *UAS-Mef2* construct. I have only observed contractions in cells that appear more "spindle-like". As some of the GFP is present in the cytoplasm of the cells, I was also able to observe a regular striped arrangement of stronger and less strong GFP accumulation within the cytoplasm in these contracting cells (Figure 3.5.10). The cytoplasm of the pericardial cells of the control larvae and that of normally shaped cells in the *Mef2* over-expressing larvae have an even and relatively low degree of GFP fluorescence. This observed structure might reflect some form of organisation within the cells, possibly the cytoskeleton and/or the cell membrane are affected by the actin arrangements that a contractile cell needs.

Currently, little is known about the expression of marker genes in the larval or adult pericardial cells, but it would be very interesting to see which genes become de-regulated as a result of over-expressing *Mef2* in the (embryonic) pericardial cells. A first step towards proving that these cells have indeed been altered towards a myogenic fate would be a Phalloidin stain to visualize any filamentous actin present in these cells.

These data show how important the exclusion of Mef2 function from the embryonic pericardial cells is. While I have shown that by stage 13 *Mef2* and *Him* expression are mutually exclusive, there is evidence that the *Mef2* transcript is expressed in the pericardial cells earlier (Lilly et al., 1994; Nguyen et al., 1994). It is highly likely that *Him* is part of a fail-safe mechanism that ensures that any Mef2 still present in the pericardial cells after these early stages of development cannot derail the differentiation of the cells.

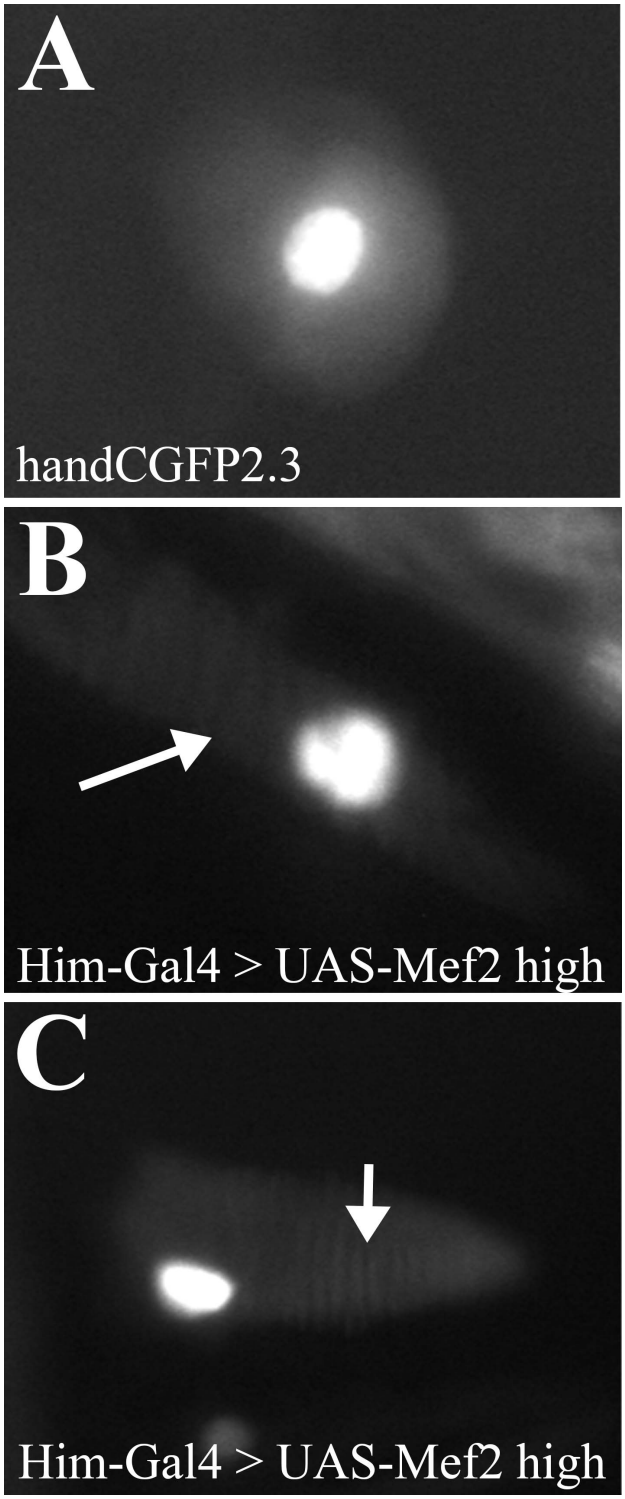


Figure 3.5.10: Images from the movies show the internal organisation visible in some pericardial cells.
The arrows point towards the striped internal organisation that can be seen in the pericardial cells of some third instar larvae that over-express Mef2.
A pericardial cell of a *handCGFP2.3* third instar control animal; **B, C** pericardial cell of third instar larvae over-expressing the stronger expressing version of the *UAS-Mef2* construct under the control of the *Him-Gal4* driver.

3.6 The role of *zfh-1* in *Drosophila* heart development

The gene *zinc-finger-homeodomain-1* (*zfh-1*) has a very similar expression pattern to the *Him* gene that is the focus of this work. Like *Him*, *zhf-1* is initially expressed pan-mesodermally after gastrulation to then be limited to the heart precursors and the adult muscle precursors (AMPs) within the mesoderm during germ-band retraction (see section 3.2.1 and Lai et al., 1991). During embryonic stage 15 both, *Him* and *zfh-1* expression changes. The expression of the *Him* transcript in the heart is reduced and eventually not visible anymore, while *zfh-1* expression is lost from the Eve-positive pericardial cells. Within the heart, no other genes than *Him* and *zfh-1* are expressed exclusively within the embryonic pericardial cells and change their expression pattern at this point of embryogenesis are known. Zhf-1 is a DNA-binding transcription factor that has been shown to be necessary for the differentiation of the Eve-positive pericardial cells (Su et al., 1999). Both genes, *Him* and *zfh-1*, are associated with the down-regulation of Mef2 function in mesodermal tissues (Postigo et al., 1999; Liotta et al., 2007). These similarities in expression pattern and function raise the question if both of these genes, *Him* and *zhf-1*, have redundant functions or if they possibly might work together during the development of the pericardial cells and the AMPs.

3.6.1 *zfh-1* loss-of-function

zhf-1^2 null mutant animals die during the last stages of embryonic development (Lai et al., 1993). Johnson et al. (2007) describe the embryonic heart phenotype of the *zfh-1*^{2} null mutant as falling in either of two categories, the heart of these embryos often show a “kink” in the heart (in or around segment A4) and an “uneven, spindly structure elsewhere in the heart” while the second phenotype category is described as having “breaks in the anterior parts of the heart and possible duplication of cardioblasts”. They also report a loss of Odd- and Tin-positive pericardial cells and a “greatly reduced expression domain of Pericardin” (Johnson et al., 2007). This loss of pericardial cells is reminiscent of the phenotype I have observed in the *Him* null mutant animals, thus furthering the possibility that both genes might be involved in the same processes.

Johnson et al. also raise the possibility that the *zfh-1*^{2} null mutant might not be a true null allele as they can detect some *zfh-1* expression if they use antibodies generated against different parts of the protein (the original Zfh-1 antibody produced by Lai et al., 1993 did not detect any *zfh-1* expression in these null embryos, this is also the antibody I have used). It is thus possible that this line is only a very strong hypomorph and still expresses a truncated version of the Zfh-1 protein that might be partially functional. However, as they also state, the *zfh-1*^{2} line is the strongest loss-of-function allele available for *zfh-1* and is the line most other studies have used to study the role of *zfh-1* in the heart.

In order to assess the *zfh-1*^{2} null allele in the context of my work, I undertook the entire embryonic cell counts with the same markers I used previously. Figure 3.6.1 shows the results of this analysis for the ß3-tubulin-positive cardioblasts. As previously described in the literature, I have observed breaks and kinks in the anterior part of the dorsal vessel in about 60 % of the mutant embryos and also a disorganisation of the cardioblasts, their arrangement is not as orderly as in the wild-type embryo (compare Figures 3.6.1 A and B). However, the numbers of ß3-tubulin-positive cardioblasts and the numbers of total cardioblasts are not significantly affected in these embryos. This result is also very similar to those described in previously published work. However, these studies did not report cell counts for all heart cell markers in *zfh-1*^{2} null mutants. Two different labs have reported "nearly normal" expression patterns for Myosin and Mef2 in the cardioblasts of the *zfh-1*^{2} null mutant (Lai et al., 1993; Su et al., 1999). My results add a third marker gene to these two marker genes. As neither of these two studies gives exact cell numbers, my results go further than their general statements and I have specific numbers to support these observations. I have, however, observed a high degree of disorganisation of the ß3-tubulin-positive cardioblasts, which is at similar to the general heart phenotype described for the *zfh-1*^{2} null mutant. In a wild-type embryo one would expect to see the repetitive two plus four pattern in the abdominal segments A1 to A6. This pattern is disturbed in approximately 80 % of all *zfh-1*^{2} null mutants. All embryos with a disrupted *ß3-tubulin* expression pattern contain hemisegments with fewer than four ß3-tubulin-positive cardioblasts in a row. In addition I have observed more than four ß3-tubulin-positive cardioblasts in a row in 55 % of all scored embryos (Figure 3.6.1. B). As the total number of cardioblasts and that of the ß3-tubulin-positive cardioblasts are unaffected in *zfh-1*^{2} mutants, this indicates an irregular arrangement of

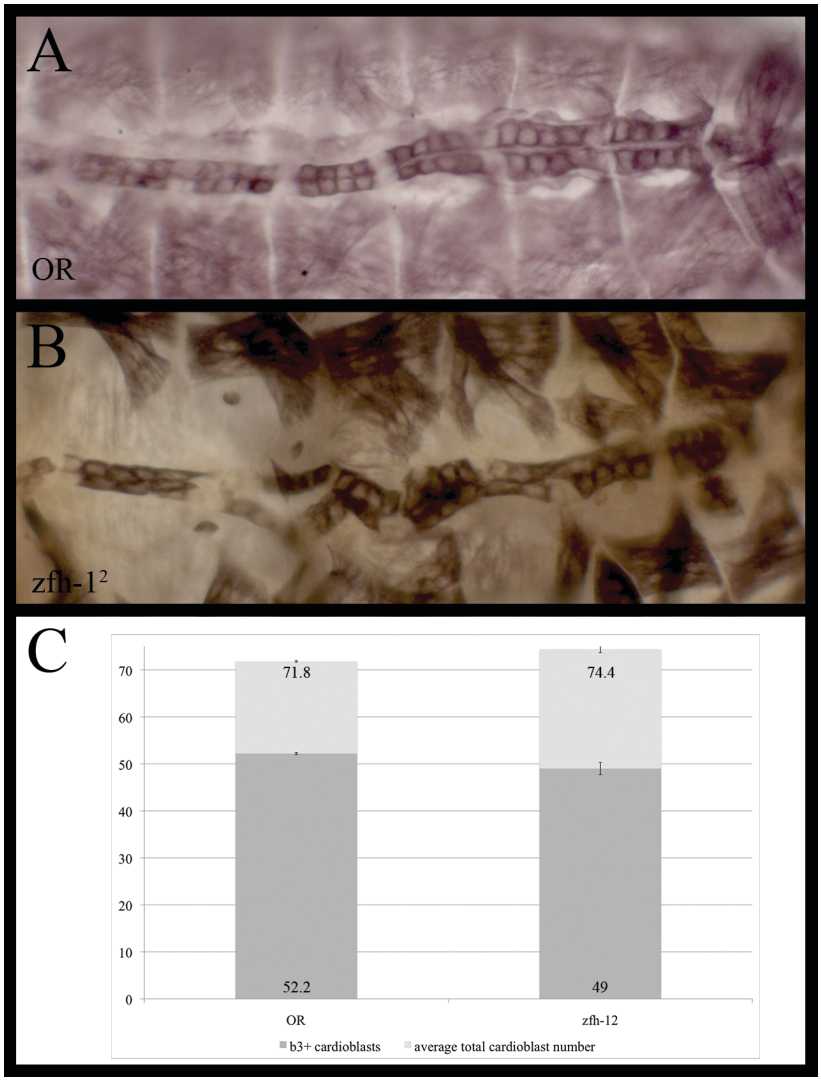


Figure 3.6.1: Expression pattern of ß3-tubulin in embryos that are homozygous deficient for *zfh-1*.
Stage 16/17 embryos were labelled for the ß3-tubulin protein and the number of ß3-tubulin-positive cardioblasts was counted.
A Wild-type embryo; **B** ß3-tubulin expression in an embryo homozygous null mutant for *zfh-1*; **C** Histogram showing the scored numbers for the ß3-tubulin-positive cardioblasts and the total number of cardioblasts in all the controls (OR) and the *zfh-1²* deficiency. The error bars indicate the standard error; there is no statistically significant difference in regard to the wild-type.

the cells. This observation points into the direction that *zfh-1*^{2} might be involved in the proper arrangement and/or organisation of the cardioblasts.

Figure 3.6.2 shows the expression pattern of Eve in wild-type embryos and those homozygous for the *zfh-1*^{2} null allele. As can be seen in the images, the Eve-positive pericardial cells are severely affected in the *zfh-1* null mutant embryos; the loss of Eve-positive pericardial cells is more pronounced in these mutants than in the *Him* null mutants. The loss of Eve-positive pericardial cells is substantiated by the scoring results, which shows a significant loss of Eve-positive pericardial cells (see chart in Figure 3.6.2 C). This result is very close to the previously published results by Su et al. (1999), who describe a complete loss of Eve-positive pericardial cells in all mutant embryos. I have, however, seen a variation with this described loss of Eve-positive pericardial cells in *zfh-1*^{2} mutants. I have observed mutant embryos with no Eve-positive pericardial cells at all to *zfh-1*^{2} mutant embryos that show a nearly normal distribution of Eve-positive pericardial cells. These two different results can only be explained by biological variance between the individual embryos.

The expression pattern of Odd in the *zfh-1*^{2} null mutant embryos is shown in Figure 3.6.3. With the exception of the typical *zfh-1*^{2} phenotype, i.e. kinks and breaks in the anterior part of the dorsal vessel, the Odd-positive pericardial cells seem generally unaffected by the loss of *zfh-1*. Counting the Odd-positive pericardial cells also shows no significant difference when compared to wild-type. Similarly, the Su et al. (1999) paper states that other than the Eve-positive pericardial cells no other group of heart cells appears to be affected by the loss of *zfh-1*. However, they only analysed the expression pattern of Eve, Mef2 and Pericardin (Prc). As the *pericardin* gene encodes a protein secreted into the extra-cellular space from the pericardial cells, its expression pattern is only of limited usefulness to determine the gain or loss of pericardial cells. The 2007 study of Johnson et al. investigated the pericardial cell population in more detail and they report a loss of Odd- and Tin-positive pericardial cells in embryos homozygous for the *zfh-1*^{2} null allele. This is in direct contradiction with my findings that the Odd-positive pericardial cells are not affected. I did however notice when I compared my stainings to the staining in Johnson et al. (2007), their staining for the Odd protein appears to be generally weaker. It is possible that the difference in the results between Johnson et al and my analysis is due to differences in the staining levels. If the level of Odd expression is affected in *zfh-1*^{2} null embryos, it is possible that I let my reaction develop for longer and was thus able to see more Odd-positive

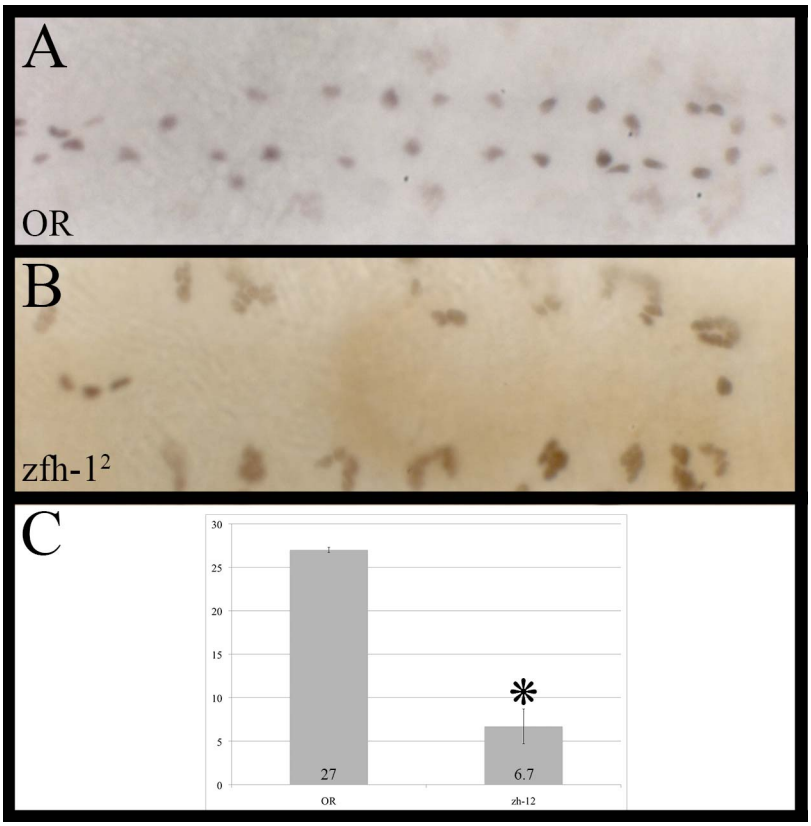


Figure 3.6.2: Expression pattern of Eve in embryos that are homozygous deficient for *zfh-1*.
Stage 16/17 embryos were labelled for the Eve protein and the number of Eve-positive pericardial cells was counted.
A Wild-type embryo; **B** Eve expression in an embryo homozygous null mutant for *zfh-1*; **C** Histogram showing the obtained numbers for the Eve-positive pericardial cells in the two controls (OR, handCGFP2.3) and *zfh-1* deficiency. The error bars indicate the standard error and the star indicates a statistically significant difference in regard to the wild-type.

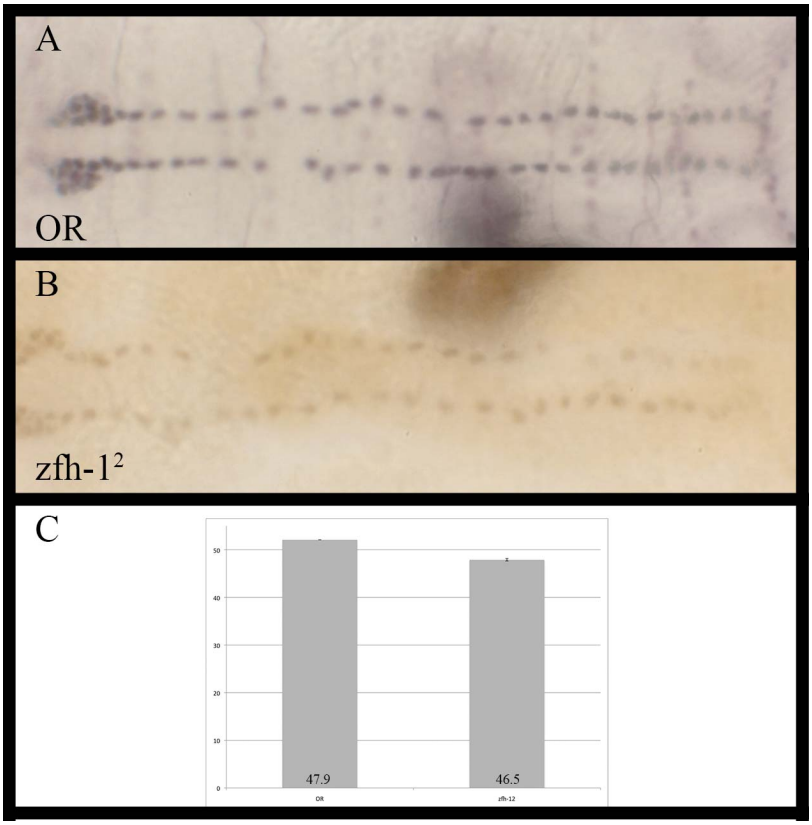


Figure 3.6.3: Expression pattern of Odd in embryos null mutant for *zfh-1*.
Stage 16/17 embryos were labelled for the Odd protein and the number of Odd-positive pericardial cells was counted.
A Wild-type embryo; **B** Odd expression in an embryo homozygous deficient for *zfh-1*; **C** Histogram showing the numbers of Odd-positive pericardial cells in the two controls (*OR, handCGFP2.3*) and *zfh-1* null embryos. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

pericardial cells. Johnson et al. also report that the number of Odd-positive pericardial cells is normal in stage 13 embryos and only reduced in stage 17 embryos. But as I also only selected embryos at this late stage for my analysis, I do not think that this is a possible explanation for the discrepancy between the two results.

Thus, I believe that while the Eve-positive pericardial cells are affected in the *zfh-1*^{2} null mutants, the Odd-positive pericardial cells and the number of cardioblasts are not (for an overview of all pericardial cell number scored in this section see Table 3.7.1). A phenotype of a not correctly organised heart has been described before for the whole heart in *zfh-1*^{2} null mutants (i.e. kinks and breaks; Lai et al., 1993). At least at cardioblast level this organisational phenotype is also visible at the level of individual cell arrangement, while the organisation of the pericardial cells appears unaffected. The regular pattern of ß3-tubulin expression is disrupted in *zfh-1*^{2} null mutants. Similarly to the results I have described for the loss of *Him*, the loss of *zfh-1* also affects some of the dorsal somatic muscles. This similarity points to a role for both genes in the specification of the shared precursor for the Eve-positive pericardial cells and the dorsal muscle precursors.

3.6.2 *zfh-1* over-expression

Johnson et al. (2007) also analysed the effects of ectopically over-expressing *zfh-1*. They describe an increased expression of Pericardin and from this conclude an increase in pericardial cells. Johnson et al. analysed the over-expression in the *zfh-1*^{2} null mutant background and used the *24B-Gal4* driver, which is expressed earlier in the mesoderm than the driver I used; they also do not show any *zfh-1* expression pattern for these experiments. I have extended their experiments by working in a wild-type background and using the *hand-Gal4* driver that is active from late embryonic stage 12, assaying all available markers I used before and following the development of the pericardial cells throughout the larval life in order to see if the increase in pericardial cells I expected to see based on the previous reports is maintained throughout larval life.

Figure 3.6.4 shows the expression pattern of the Zfh-1 protein in the embryos over-expressing *zfh-1* under the control of the *hand-Gal4* driver. The number of Zfh-1-positive cells appears similar to that seen in the wild-type and there are no

obvious organisational problems. This is contrary to what I was expecting. I would have expected to see *zfh-1* expression in the cardioblasts, which are visible and can be identified by their unique shape and position, under DIC optics, or the visceral muscles as the *hand-Gal4* driver is also active in these tissues. As I have verified both the driver line and the line carrying the *UAS-zfh-1* construct separately, this leads me to the conclusion that *zfh-1* is strongly controlled even after transcription and that either the Zfh-1 protein or its RNA is not stable in these cells and has no immediate effect. It is also possible that this particular *UAS-zfh-1* construct only results in low levels of ectopic expression, which cannot be detected by the Zfh-1 antibody.

The images in Figure 3.6.5 show the results for the ß3-tubulin-positive cardioblasts. From the expression pattern of ß3-tubulin alone, I am not able to differentiate between the wild-type embryos and those over-expressing *zhf-1*. Organisation of the cardioblasts and cell shape and size are identical to that observed in the wild-type embryo. Counting the cells also did not show a significant difference between the two types of embryos as can be seen in the graph in Figure 3.6.5 C.

Figure 3.6.6 shows the expression pattern of Eve in wild-type embryos and those over-expressing *zfh-1* under the control of the *hand-Gal4* driver. Again, from the expression pattern of Eve alone, I am not able to differentiate between the wild-type embryos and those over-expressing *zfh-1*. The chart in Figure 3.6.6 C shows the results of counting the Eve-positive pericardial cells, which shows no significant difference between the wild-type and the over-expressing embryos in regard to the number of cells.

Contrary to my previous two results, Odd is expressed in more cells in the embryo over-expressing *zhf-1* (Figure 3.6.7). While the cell shape and size remains roughly the same, there are more cells spread out along the length of the heart. However, this does not alter the "pearls-on-a-string" organisation that is typical for the Odd-positive pericardial cells. Counting of these Odd-positive pericardial cells revealed a statistically significant increase in cell number. This result is similar to the result published in Johnson et al. 2007, who describe "ectopic Tin- and Odd-positive pericardial cells" in their over-expression experiments. Ectopically over-expressing *zfh-1* in the mesoderm induces an increase in Odd-positive pericardial cells but does not affect the Eve-positive pericardial cells or the ß3-tubulin-positive cardioblasts. Only the Odd-positive (and according to Johnson et al. the Tin-positive) pericardial cells are thus susceptible to increased levels of Zfh-1 in the later stages of embryonic development.

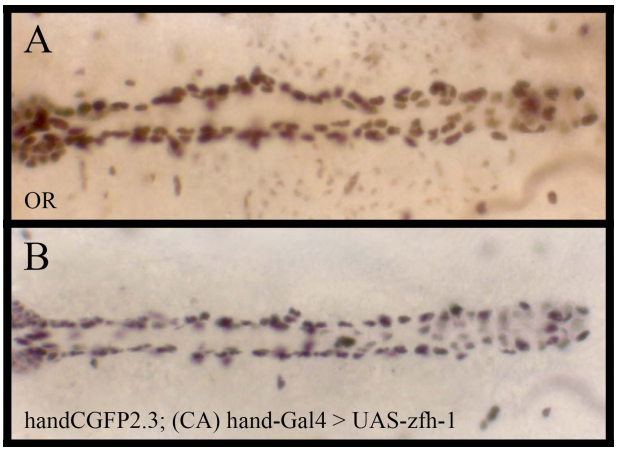


Figure 3.6.4: Zfh-1 expression in embryos over-expressing *zfh-1* in the heart.
Stage 16/17 embryos were labelled for the distribution of the Zfh-1 protein.
A OR wild-type embryo; **B** Zfh-1 expression in an embryo over-expressing *zfh-1* under the control of the *hand-Gal4* driver.

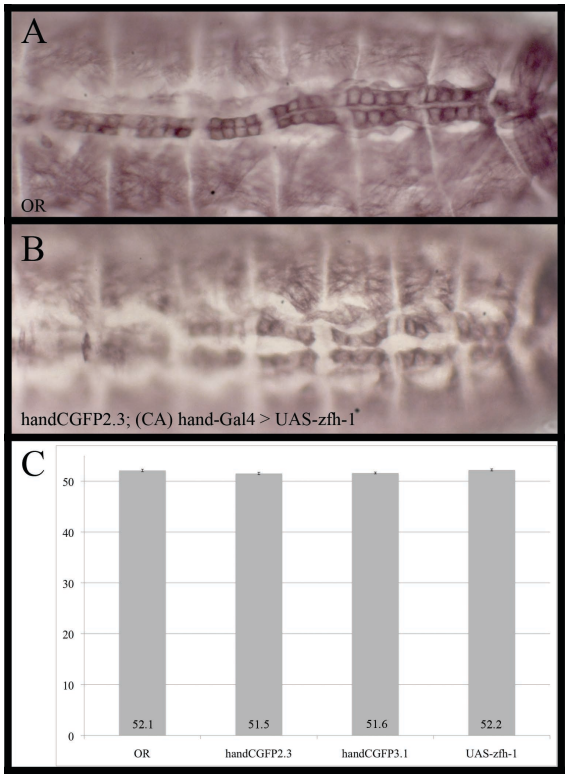


Figure 3.6.5: Expression pattern of ß3-tubulin in embryos that over-express *zfh-1* in the heart.
Stage 16/17 embryos were labelled for the ß3-tubulin protein and the number of ß3-tubulin-positive cardioblasts was counted.
A Wild-type embryo; **B** ß3-tubulin expression in an embryo ectopically over-expressing *zfh-1* under the control of the *hand-Gal4* driver; **C** Histogram showing the numbers of ß3-tubulin-positive pericardial cells in all three controls (*OR, handCGFP2.3, handCGFP3.1*) and the *zfh-1* over-expression study. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

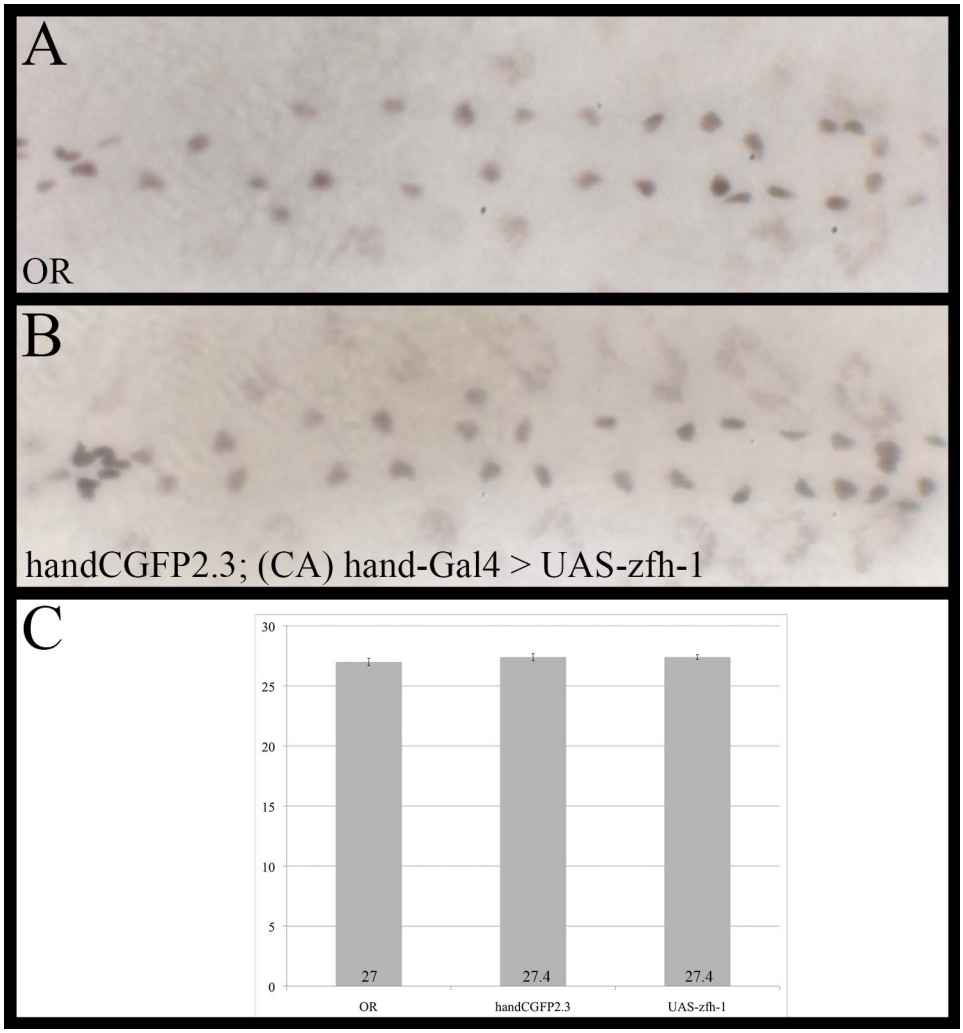


Figure 3.6.6: Expression pattern of Eve in embryos that over-express *zfh-1* in the heart.
Stage 16/17 embryos were labelled for the Eve protein and the number of Eve-positive pericardial cells was counted.
A Wild-type embryo; **B** Eve expression in an embryo over-expressing *zfh-1* under the control of the *hand-Gal4* driver; **C** Histogram showing the numbers of Eve-positive pericardial cells in the two controls (*OR, handCGFP2.3*) and the *zfh-1* over-expression study. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

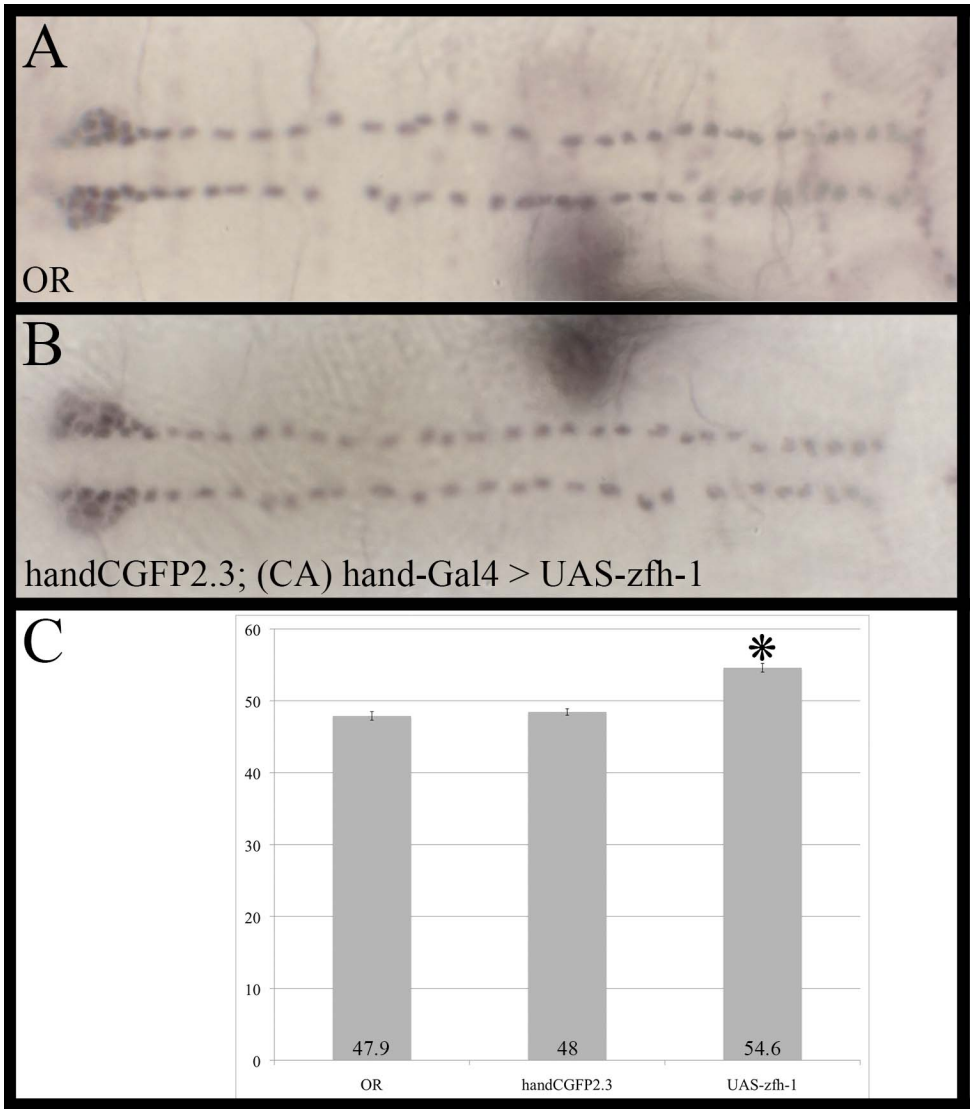


Figure 3.6.7: Expression pattern of Odd in embryos expressing ectopic *zfh-1* in the heart.
Stage 16/17 embryos were labelled for the Odd protein and the number of *Odd-positive* pericardial cells was counted.
A Wild-type embryo; **B** Odd expression in an embryo ectopically over-expressing *zfh-1* under the control of the *hand-Gal4* driver; **C** Histogram showing the numbers of Odd-positive pericardial cells in the two controls (*OR, handCGFP2.3*) and the *zfh-1* over-expression study. The error bars indicate the standard error and stars indicate a statistically significant increase compared to wild-type.

The likely explanation for this are the different precursors for the different subsets of pericardial cells. The precursors of the Eve-positive pericardial cells are the earliest precursors to be specified; it is thus likely that a different genetic program targets these precursors.

Figure 3.6.8 shows the hand-GFP-positive pericardial cells during the first larval instar when *zfh-1* is over-expressed under the control of the *hand-Gal4* driver. There is very little difference in the expression of the hand-GFP construct between the wild-type animals and those that over-express *zfh-1*. Organisation, cell shape and size appear normal even when *zfh-1* is over-expressed in the heart. This is surprising as the embryonic experiments showed an increase in Odd-positive pericardial cells. A possible explanation for these results is that more than the four pericardial cells per hemisegment in a wild-type embryo gain Odd-expression during embryogenesis. This would allow for an unchanged total number of pericardial cells.

The second instar larvae that express ectopic *zfh-1* also do not show an increase in pericardial cells when compared to wild-type larvae (Figure 3.6.9). Counting of the hand-GFP-positive pericardial cells in the second instar larvae confirms there is no significant difference between wild-type or *zfh-1* over-expressing larvae (see chart in Figure 3.6.9 D).

By the third instar, the number of pericardial cells has increased markedly (see Figure 3.6.10). Again it is not possible to say if these cells derive from the pericardial cells that are "lost" in wild-type larvae or if they are generated anew from another cell-type. There are supernumerary cells especially in the aorta and the anterior part of the heart proper (see arrows in Figure 3.6.10 C) while in the wild-type third instar larvae the anterior pericardial cells appear to be spaced further apart. Counting of the number of hand-GFP-positive pericardial cells confirms this impression. There is a statistically significant increase in pericardial cell number of about ten cells per animal if *zfh-1* is over-expressed (see Table 3.7.1 for an overview off all cell numbers in this section). This increase in pericardial cells in the third instar larva is reminiscent of the increase of pericardial cells I observed when over-expressing *Him* under the control of the same driver. However, the phenotypes of the $zfh\text{-}1^2$ and *Him* mutants that I have described in the embryos and in earlier larval stages differ, this shows that both genes most likely affect different aspects of pericardial cell development which result in very similar

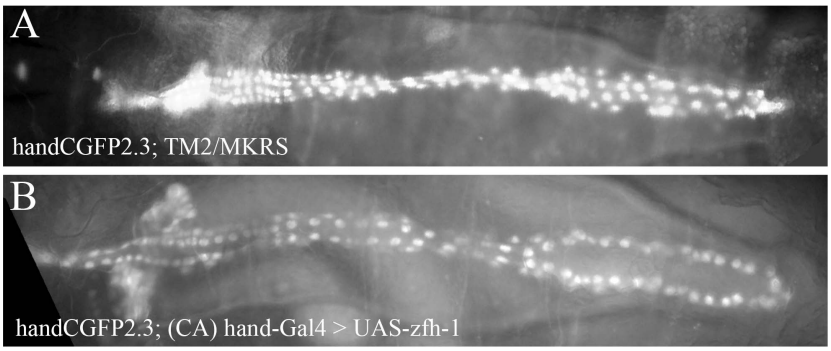


Figure 3.6.8: Expression pattern of the *hand-GFP* reporter construct in first instar larvae over-expressing *zfh-1* in the heart.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the *GFP* expression pattern.
A *handCGFP2.3* reporter construct, wild-type control; **B** GFP expression pattern in a first instar larvae over- expressing *zfh-1* under the control of the *hand-Gal4* driver.

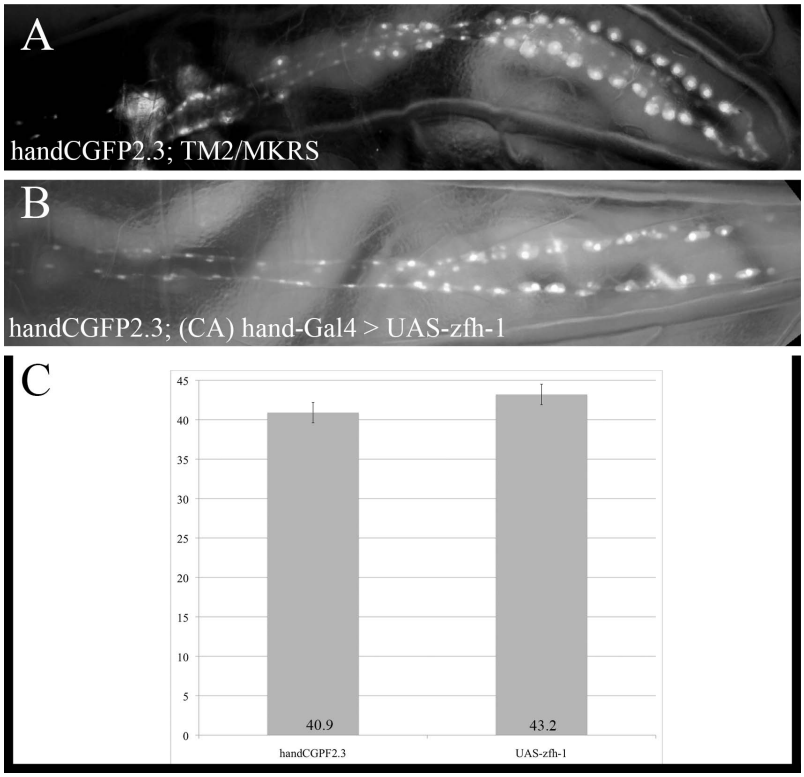


Figure 3.6.9: Expression pattern of the *hand-GFP* reporter construct in second instar larvae expressing ectopic *zfh-1* in the heart.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the hand-GFP expression pattern and the number of GFP-positive pericardial cells was counted.
A *handCGFP2.3* reporter construct, wild-type control; **B** GFP expression pattern in a second instar larvae over-expressing the *zfh-1* under the control of the *hand-Gal4* driver; **C** Histogram showing the numbers of hand-GFP-positive pericardial cells in the two controls (*handCGFP2.3, handCGFP3.1*) and over-expression of *zfh-1* in the second instar. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

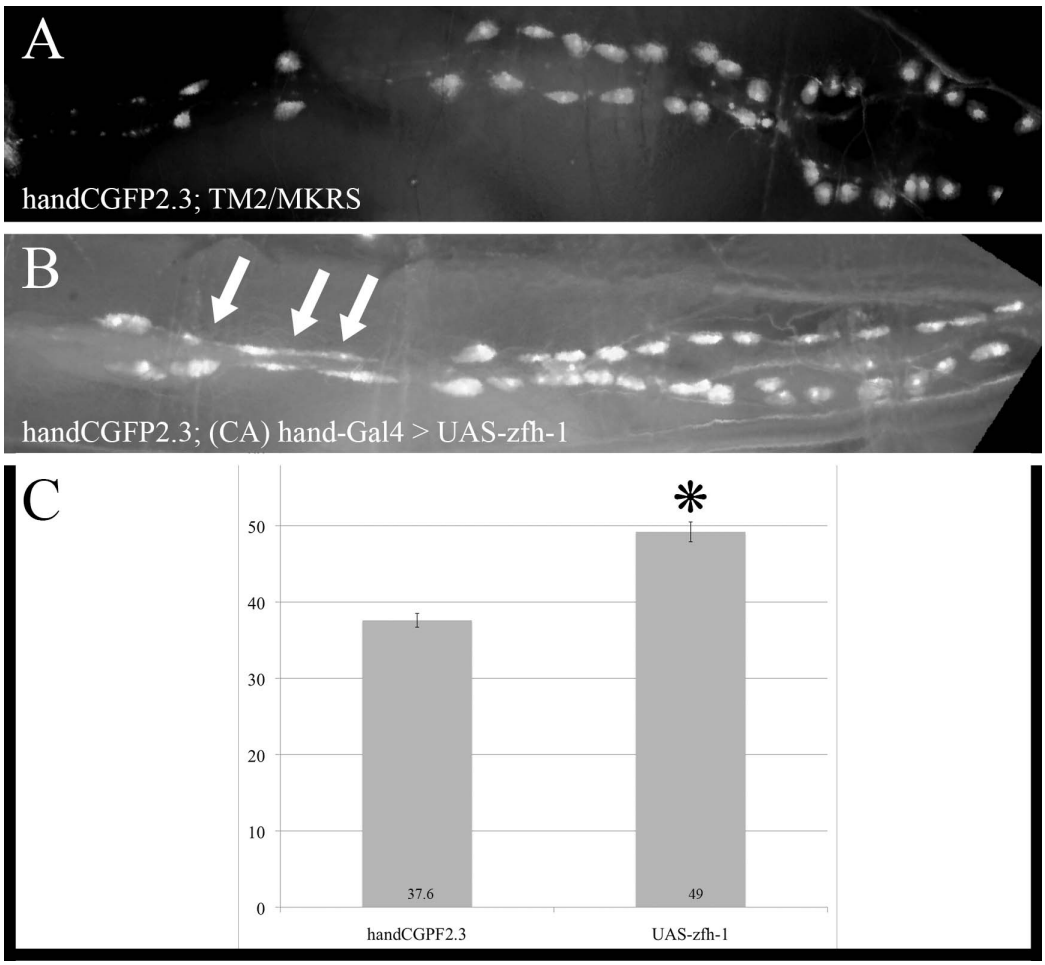


Figure 3.6.10: Expression pattern of the hand-GFP reporter construct in third instar larvae over-expressing *zfh*-1 in the heart.
Live larvae were photographed under the microscope (Zeiss Axiovision) to visualize the GFP expression pattern and the number of GFP-positive pericardial cells was counted.
A *handCGFP2.3* reporter construct, wild-type control; **B** GFP expression pattern in a third instar larvae over-expressing *zfh-1* under the control of the *hand-Gal4* driver; **D** Histogram showing the numbers of handGFP-positive pericardial cells in the two controls (*handCGFP2.3, handCGFP3.1*) and the larvae over-expressing *zfh-1* in the third instar. The error bars indicate the standard error and the star indicates a statistically significant increase compared to wild-type.

phenotypes by the end of larval development. This demonstrates once more how closely the *Him* and *zfh-1* expression patterns and phenotypes resemble each other.

3.7 *Him* and *zfh-1* double null mutant

The *Him* and *zfh*-1 null mutant phenotypes and expression pattern are, as described above, very similar to each other. This leads to the question if these two genes might work as part of the same process during mesoderm development. In order to investigate this question, I analysed the embryonic phenotype of flies doubly null-mutant for *zfh-1*^{2} and the three available *Him* null alleles. I chose to analyse all three *Him* null alleles as there is a slight variation between the three alleles, thus it possible that not all three alleles show the same phenotype when combined with the *zfh-1*^{2} allele. The embryonic lethality of the *zfh-1*^{2} null mutant is maintained in these double mutants, so that I could only analyse the embryonic heart.

Figure 3.7.1 shows the expression pattern of the ß3-tubulin-positive cardioblasts in embryos doubly mutant for *zfh-1*^{2} and *Him*. All three different *Him-zhf-1* double null mutant lines show the basic phenotype of kinks and the occasional break that has also been observed in the single *zfh-1* null mutant. It is however noticeable that the number of ß3-tubulin-positive cardioblasts varies between individual hemisegments from one to six. I have seen a similar effect on the organisation of the ß3-tubulin-positive cardioblast organisation in the single *zfh-1*^{2} mutants, thus this phenotype is most likely due to the loss of *zfh-1*. Depending on which *Him* null mutant line is used for these double mutations, there are slight but no statistically significant differences in the number of scored ß3-tubulin-positive cardioblasts. The total number of cardioblasts as counted on embryos stained for ß3-tubulin using DIC optics is also not statistically different to the number observed in wild-type embryos. This is a different phenotype to what I observed in the single *Him* null mutants which all lost ß3-tubulin-positive cardioblasts and also showed a reduction in the number of total cardioblasts. Clearly, the *zfh-1* phenotype is dominant over that of the *Him* null mutants with regard to the ß3-tubulin-positive cardioblasts. This would imply that, during development of at least the cardioblasts, the Zfh-1 protein is active in a process that is upstream of any need of *Him*. However in both *zfh*-1 single and *zfh-1;; Him* double mutants, the ß3-tubulin-

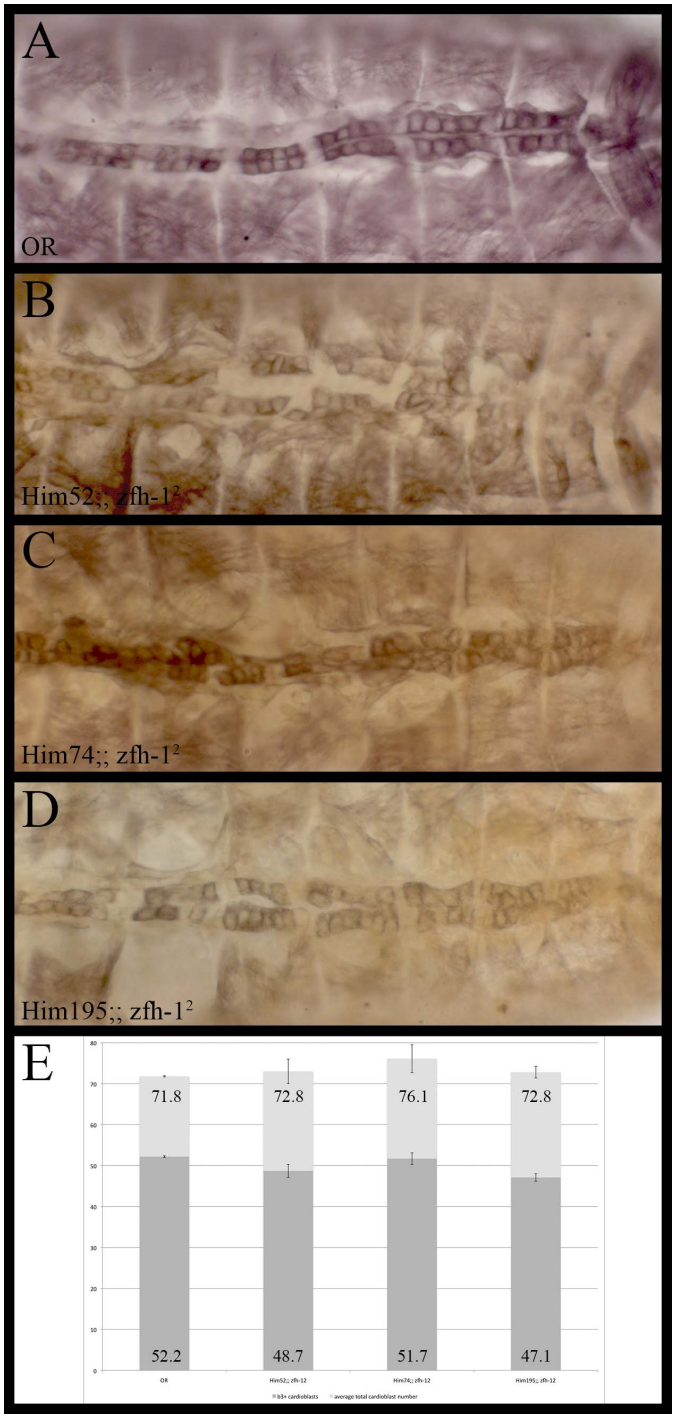


Figure 3.7.1: Expression pattern of ß3-tubulin in *Him;; zfh-1²* double null mutant embryos.
Stage 16/17 embryos were labelled for the ß3-tubulin protein and the number of ß3-tubulin-positive cardioblasts was counted.
A Wild-type embryo; **B** ß3-tubulin expression in a *Him 52;; zfh-1²* null mutant embryo; **C** the ß3-tubulin expression pattern in a *Him 74;; zfh-1²* null mutant embryo; **D** ß3-tubulin expression pattern in a *Him 195;; zfh-1²* null mutant embryo; **E** Histogram showing the numbers of ß3-tubulin-positive cardioblasts and the total number of cardioblasts in the controls and the three different *Him;; zfh-1²* double null lines. The error bars indicate the standard error; there is no statistically significant difference compared to wild-type.

positive cardioblasts show a disordered arrangement when compared to wild-type embryos.

The expression pattern of *eve* in embryos double mutant for *Him* and *zfh-1* is shown in Figure 3.7.2. As can be seen these embryos have fewer Eve-positive pericardial cells in all available lines tested. Approximately one third of the embryos in each line display a severe loss of Eve-positive pericardial cells (defined as less than 10 Eve-positive pericardial cells per whole embryo) while others appear to have a nearly normal distribution of Eve-positive pericardial cells, similar to what I observed in the single *zfh-1* null mutants. Counting of the Eve-positive pericardial cells shows that each doubly homozygous line loses a significant amount of Eve-positive pericardial cells (see chart in Figure 3.7.2 E). This is also the case for all of the single null mutant lines. While the average scored Eve-positive pericardial cell number is generally lower in the *Him-zfh-1* double mutants than in the single *Him* null mutants, the loss of Eve-pericardial cells observed in the double mutants is not as severe as the loss of these cells I observed in the single $zfh\text{-}1^2$ mutants. This places the phenotype of the double mutants between that of the *zfh-1* and *Him* single null mutants with regard to the Eve-positive pericardial cells.

In the $zfh\text{-}1^2$ single mutants the Odd-positive pericardial cells were not affected in my analysis. In *Him* null mutant embryos I observe a loss of these cells. All double mutants also show a significant loss to the same extent as seen in the single *Him* null mutants (see chart in Figure 3.7.3 E). In all three analysed *Him-zfh-1* double mutants there are gaps within the arrangement of the Odd-positive pericardial cells along the side of the cardiac tube as can be seen in Figure 3.7.3 (see arrows). This is most severe in the $Him195;;zfh\text{-}1^2$ double mutant (see Figure 3.7.3 D). The Odd-positive pericardial cells appear to be more sensitive to the loss of *Him* than to the loss of *zfh-1* as they are not affected in the $zfh\text{-}1^2$ single null embryos I analysed.

The experiments and results described above demonstrate the differences between the individual sub-sets of heart cells in regard to their sensitivity to the loss of either or both *Him* and *zfh-1*. For an overview of all cell numbers discussed in this section see Table 3.7.1. While the cardioblasts seem to be more sensitive to the loss of *zfh-1* than *Him*, the Odd-positive pericardial cells are more sensitive to the loss of *Him* and the effect of the loss of both genes on the Eve-positive pericardial cells is not pushed into the direction of either gene.

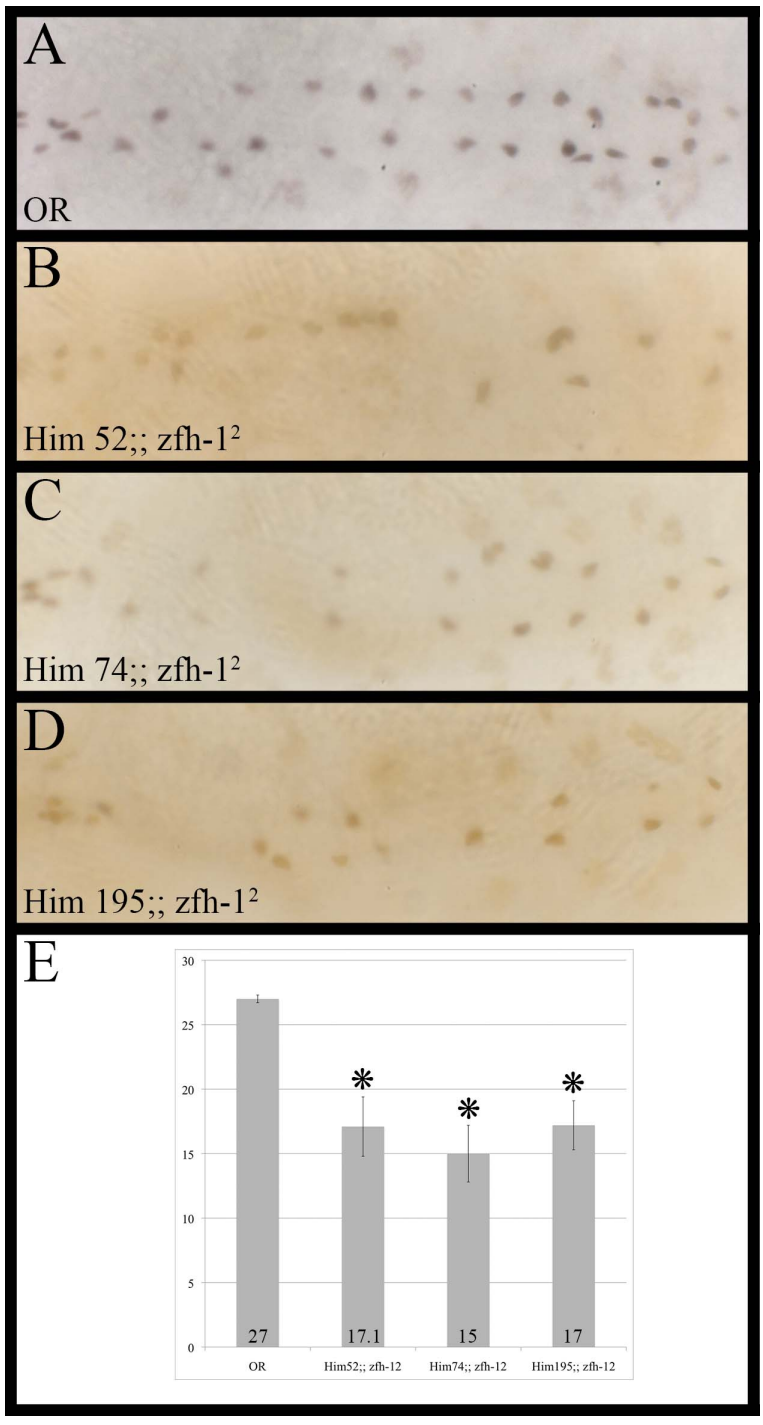


Figure 3.7.2: Expression pattern of Eve in *Him;; zfh-1²* double *null* mutant embryos.
Stage 16/17 embryos were labelled for the Eve protein and the number of *Eve-positive* pericardial cells was counted.
A Wild-type embryo; **B** Eve expression in a *Him 52;; zfh-1²* double *null* mutant embryo; **C** the Eve expression pattern in a *Him 74;; zfh-1²* double *null* mutant embryo; **D** Eve expression pattern in a *Him 195;; zfh-1²* double *null* embryo; **E** Histogram showing the numbers of *Eve-positive* pericardial cells in the control (*OR*) and the *Him;; zfh-1²* double null lines. The error bars indicate the standard error and stars indicate a statistically significant decrease compared to wild-type.

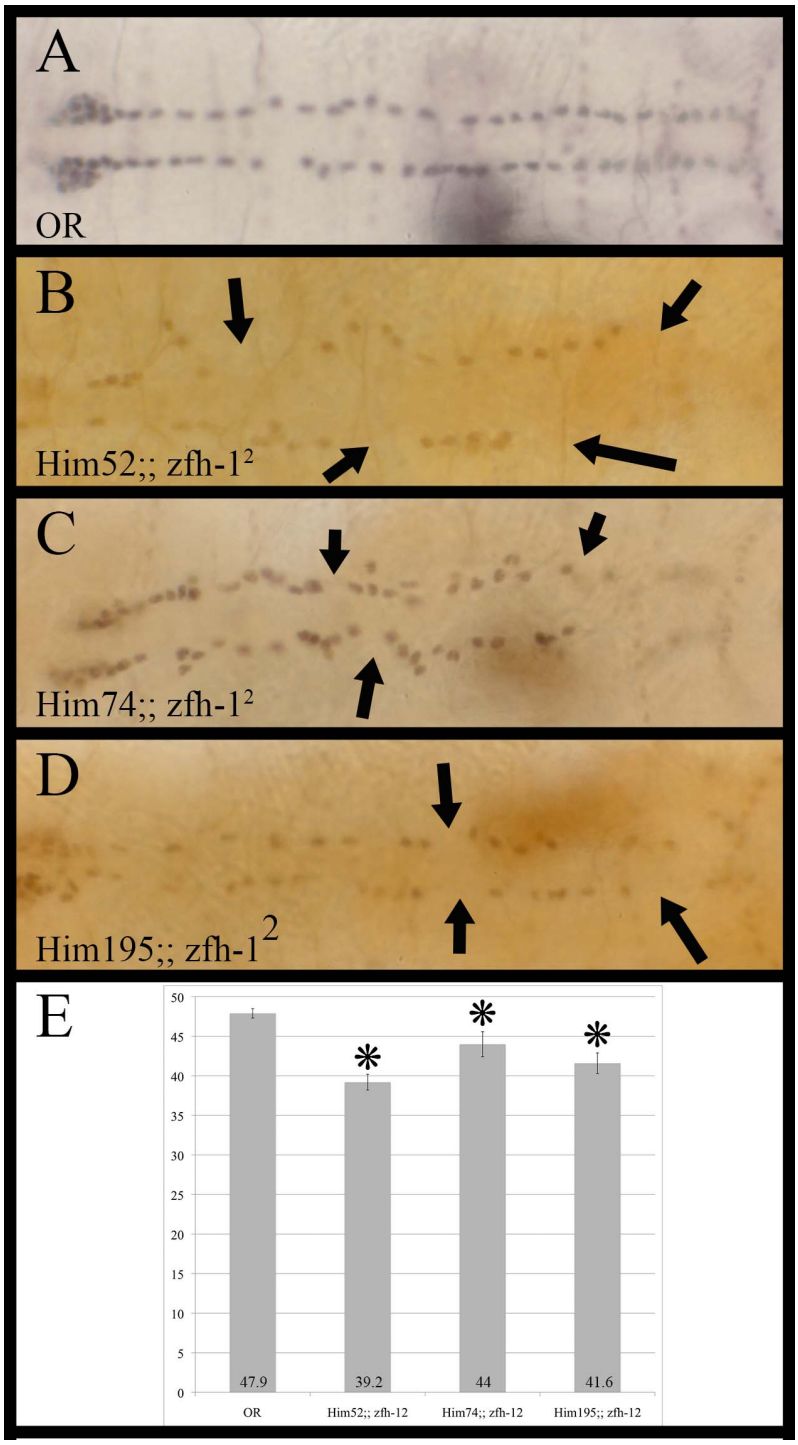


Figure 3.7.3: Expression pattern of Odd in *Him;; zfh-1²* double null mutant embryos.
Stage 16/17 embryos were labelled for the Odd protein and the number of Odd-positive pericardial cells was counted. The arrows point towards gaps in the Odd expression pattern.
A Wild-type embryo; **B** Odd expression in a *Him 52;; zfh-1²* double null mutant embryo; **C** the Odd expression pattern in a *Him 74;; zfh-1²* double null mutant embryo; **D** Odd expression pattern in a *Him 195;; zfh-1²* embryo; **E** Histogram showing the numbers of Odd-positive pericardial cells in the control (*OR*) and the *Him;; zfh-1²* null lines. The error bars indicate the standard error and stars indicate a statistically significant decrease compared to wild-type.

Table 3.7.1: Heart cell numbers obtained from all *zfh-1* loss-of-function and gain-of-function experiments described in this section. OPCs: Odd-positive pericardial cells; EPCs: Eve-positive pericardial cells, CB: total cardioblast number, β3: β3tubulin-positive cardioblast, n/a: data not available, a grey underlay indicates a significant change in cell number when compared to wild type.

	pericardial cells		cardioblasts		larval stages	
	OPCs	EPCs	β3	CB	2nd instar	3rd instar
OR	47.9	27	52.2	71.8	n/a	n/a
handCGFP 2.3	48.5	27.4	n/a	n/a	45.9	38.3
handCGFP 3.1	48.1	26.5	n/a	n/a	43.4	37.6
zfh-1^{2}	46.5	6.7	49	74.4	n/a	n/a
hand-Gal4 x UAS-zfh-1	54.6	27.4	52.2	n/a	43.2	49
Him 52;; zfh-1^{2}	39.2	17.1	48.7	72.8	n/a	n/a
Him 74;; zfh-1^{2}	44	15	51.7	76.1	n/a	n/a
Him 195;; zfh-1^{2}	41.6	17	47.1	72.8	n/a	n/a

If both genes work together in the same pathway, I would expect to see a more dominant phenotype of one above the other gene. If, however, both genes work through different pathways that target the same goal, I would expect a more severe phenotype in the animals doubly homozygous for both genes, ie. a more dramatic loss of Eve expressing heart cells as both individual mutations cause a loss of Eve-positive pericardial cells. The double mutants exhibit the phenotype of the *zfh-1*^{2} null mutants in the cardioblasts. This implies that in these cells *zfh-1* is necessary for *Him* function. As Zfh-1 has been described as a pericardial cell specific factor (Johnson et al., 2007) and both, *zfh-1* and *Him*, are excluded from the cardioblast very soon after their specification, this interaction has to take place during the very early stages of mesoderm development.

For the pericardial cells the situation is slightly different. In the Eve-positive pericardial cells I observe a loss. The degree of this loss is intermediate between that of the individual null mutants. This does not allow me to say which of the two genes is dominant. My results for the Odd-positive pericardial cells point towards a dominance of the *Him* phenotype over the *zfh-1*^{2} phenotype as the double mutants exhibit the phenotype I observed in the *Him* null mutants, while I did not detect any changes in the number of Odd-positive pericardial cells in the *zfh-1*^{2} null mutants. It is possible that depending on the identity of the individual pericardial cell, *Him* and *zfh-1* are of different importance and that one phenotype might supersede the other in these particular pericardial cells. This would be especially interesting if it can be shown that the different groups of pericardial cells are fulfilling different functions during larval and/or adult life. However, one needs to take into consideration that the differentiation of any cell type is under the control of many different factors and is likely to include networks that contain redundancies and feed-back loops.

3.8 Discussion

In this chapter I have analysed and proven the importance of *Him* expression and function during heart development. My results show that *Him* plays an important role during the early specification of all heart cells as the loss of *Him* leads to reduction in

the number of pericardial cells and cardioblasts (Table 3.4.1). The exclusion of *Him* expression from the cardioblasts is of equal importance for the development of these cells as prolonged *Him* expression in these cells leads to a decrease in *myosin*. In addition to this, the levels of *Him* expression within the pericardial cells are also crucial for the correct development of the pericardial cells in the embryo and larva. An increase in *Him* expression in the pericardial cells leads to pericardial cell hyperplasia that cannot be corrected during larval life (Table 3.4.1). The results of both the *Him* loss-of-function and gain-of-function experiments also show that the loss of *Him* expression towards the end of embryogenesis is important to allow for the development of the correct number of larval pericardial cells. A possible explanation for all these results is that *Him* is part of a genetic program that prevents the premature and thus wrong differentiation of heart cells, first separating the cardioblasts from the pericardial cells and then, by not being expressed anymore, "allowing" the pericardial cells to undergo their correct development during larval life. As Johnson et al. (2007) have shown that an increase in pericardial cell numbers interferes with the ability of the heart to pump haemolymph during larval life, there is a "physiological need" for these two possible "check-points" during pericardial cell development to ensure the correct number of pericardial cells is specified.

I have also demonstrated that Him likely interacts with Mef2 in similar ways to those described in the embryonic somatic mesoderm and the development of the flight musculature (Liotta et al., 2007; Soler and Taylor, 2009). The expression of *Mef2* during the embryonic development of the pericardial cells leads to a severe phenotype in the later larval stages (Tables 3.5.1, 3.5.2 and 3.5.3). Not only is there an increase in the number of pericardial cells as marked by the hand-GFP reporter gene and changes in the shape in a proportion of these cells but the physiology of these cells is also severely altered from a non-contractile cell to a cell that is capable of independent and individual contractions. This not only shows how important Him function is to prevent any activity of the possible remnants of the early *Mef2* expression in the pericardial precursor cells but it also demonstrates that Mef2 on its own is capable of inducing a myogenic fate in mesodermal but non-myogenic cells.

In this chapter I have also described the cardiac expression pattern of *Him* in the developing embryo and larvae. The *Him* transcript becomes less and less visible from embryonic stage 15 onwards, until it is not detectable anymore. During larval life, I have been able to detect GFP expression (under the control of various *Him* enhancer

constructs) in an area that is closely associated with the heart of second and third instar larvae. These cells have a different shape and appearance to the previously described hand-GFP positive pericardial cells and I am currently not clear of the identity of these cells. As far as I am aware these cells have not been described previously and might represent a currently unkown subset of cardiac cells. This larval expression pattern is only reproduced with the *Him-Gal4 L4-1* driver line and not with the *Him-Gal4 L3-5* driver line that was used throughout my work. This implicates the 1 kb region directly upstream of the transcriptional start site in the regulation of the *GFP* expression in these unidentified cells.

Analysis of the single *zfh-1*^{2} null mutant and *Him-zfh-1* double mutants has emphasised the different identities of the sub-sets of heart cells as different groups, identified by different marker expression (Odd, Eve, Tin, Lbe); they show varying responses to the individual or combined loss of *Him* and/or *zfh-1* (Table 3.7.1). Both *zfh-1* and *Him* are, despite only being expressed in the very early precursors of the cardioblasts, of importance for the development of the cardioblasts as the loss of *Him* leads to a loss in cardioblasts. However, in *zfh-1*^{2}*-Him* double null mutants this loss is "rescued", demonstrating that without *zfh-1*, *Him* has little to no influence on the cardioblasts. Together with the results I have described for the *Him* null mutant larvae this demonstrates that not only the presence but also the absence and more so the exact points when expression of a gene starts and stops are of importance and immediately connected with its function.

Chapter 4
The Pericardial Cell enhancer

4.1 Introduction

Knowledge of the specific expression pattern of a gene is a very useful tool when analysing its regulation. As shown in Chapter 3, *Him* is necessary for the correct development of all heart precursors and especially for the pericardial cells and its exclusion from the cardioblasts is equally important to allow for their specification. Both pericardial cells and cardioblasts are closely related and in the case of the Seven-up-positive cardioblasts and two of the Odd-positive pericardial cells they share an immediate progenitor (see Chapter 1). Thus the regulation of *Him* is interesting and will further the understanding of which regulatory mechanisms differentiate the cardioblasts and pericardial cells.

The *Drosophila* Genome Project (BDGP) and its associated bioinformatic tools facilitate the search for potential regulators further. The comparison of equivalent sequences from different *Drosophilae*, called phylogenetic footprinting, enables identification of regions that are conserved across those species. This conservation in turn indicates that the sequence potentially has some form of functional constraint applied towards it. One of the likeliest functional constrains put on non-coding sequences is transcriptional regulation. Regions identified in such an analysis are thus likely to contain transcription factor binding sites.

The availability of online transcription factor databases and the research done in this area by others allows for a more targeted approach in finding which putative transcription factor binding sites are contained within the conserved regions. In this chapter, I describe the identification of an enhancer fragment capable of recapitulating the *Him* expression pattern *in vivo* in the heart. I will then describe the bioinformatic analysis performed on this sequence and the conserved regions and putative transcription factor binding sites identified.

4.2 *Him* enhancer regions

The gene *Him* is located on the X-Chromosome between the genes *Her* and CG33639. The upstream region between *Her* and *Him* is 4311bp long and the downstream region between *Him* and CG33639 is 2333bp long (Tweedie et al., 2009; www.flybase.org). Taking advantage of two restriction sites (EcoRI and XhoI) in the upstream sequence of this gene, a fragment spanning from -3734 to -930 was further investigated (S. McConnell & M. Taylor, unpublished results). This so called *p43 (Eco/Xho)* reporter contains this part of the upstream region and drives expression in the *Him* pattern. A similar section of the upstream region between *Her* and *Him* has also been shown to contain regulatory elements that recapitulate *Him* expression in the third instar wing disc (Rebeiz, M. et al., 2002).

The ~2.8 kb of the *p43 (Eco/Xho)* fragment were broken down into sub-constructs, called Set 1 through Set 4 (see Figure 4.2.1; S. McConnell & M. Taylor, unpublished results). All constructs overlap each other at the end by approximately 70 to 80 bp to ensure that binding sites at the start and end of each construct are not destroyed by dividing up the ~2.8 kb fragment. The construct Set 4 is further sub-divided into the Set 2 and Set 3 constructs.

All sub-constructs were cloned in front of a minimal promotor and a ß-Galactosidase reporter. The expression pattern of the ß-Galactosidase reporter was analysed for all four constructs (Figure 4.2.2). In embryos carrying the Set 1 enhancer construct no reporter expression was observed. The other three enhancer constructs show ß-Galactosidase reporter expression.

The ß-Galactosidase expression pattern of Set 4 enhancer is expressed in all the pericardial cells (S. McConnell & M. Taylor, unpublished data) and thus recapitulates the expression of the *Him* transcript (Figure 4.2.2 D). Analysis of the Set 2 enhancer expression in the heart showed that it is not expressed in the Eve-positive pericardial cells (Figure 4.2.2 B; S. McConnell & M. Taylor, unpublished data). As the Set 4 enhancer, the Set 3 enhancer is expressed in all pericardial cells (Figure 4.2.2 C; S. McConnell & M. Taylor, unpublished data).

The Set3 enhancer is the smallest of these fragments (574 bp) that is capable of driving reporter expression in the same pattern in the developing *Drosophila* dorsal vessel as the *Him* transcript. This fragment is also the closest to the transcriptional start

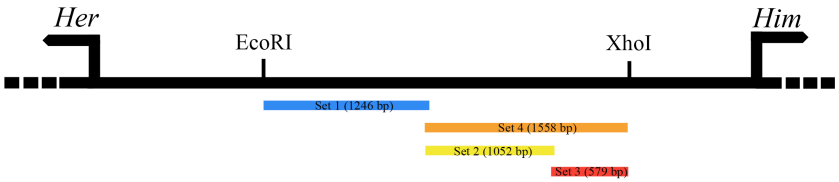


Figure 4.2.1: Overview of reporter constructs available at the beginning of my study.

Schematic representation of the *Him* upstream region. The sequence between the EcoRI and XhoI restriction enzyme sites was subdivided into four reporter constructs: Set 1 (blue), Set 2 (yellow), Set 3 (red, =PCE) and Set 4 (orange). The schematic is drawn to the scale of approximately 750 bp per 1 cm.

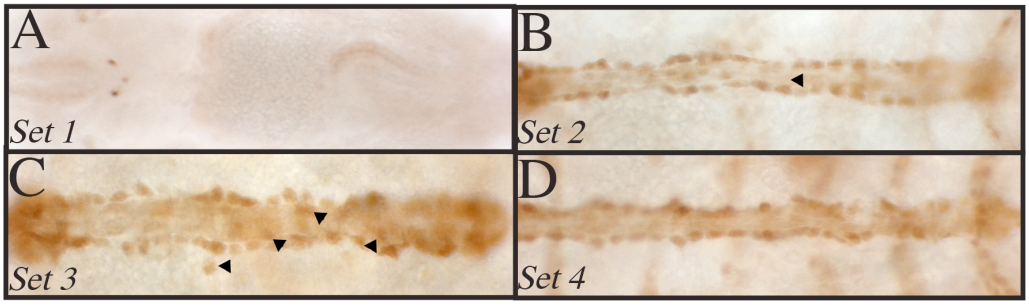


Figure 4.2.2: Overview of the expression pattern of the reporter constructs available at the beginning of my study.

Visualization of the ß-Galactosidase reporter.
A The Set 1 enhancer does not drive reporter gene expression in the heart. **B** The Set 2 enhancer drives reporter gene expression in the majority of the pericardial cells of the heart but not in the Eve-positive pericardial cells (see arrow heads and compare with C). **C** The PCE (also called Set3) enhancer drives reporter gene expression in all pericardial cells. The Eve-positive pericardial cells are indicated by arrow heads. **D** The Set 4 enhancer drives reporter gene expression in all pericardial cells.

site of the *Him* gene with a distance of approximately 930 bp. Because of this, I focussed my analysis of regulatory regions for *Him* on the Set 3 enhancer fragment and renamed it Pericardial Cell Enhancer, PCE.

4.3 The Pericardial Cell Enhancer (PCE) is expressed in all pericardial cells.

In order to establish that the Pericardial Cell Enhancer is expressed in the same heart cells as the *Him* transcript, I co-labelled embryos transgenic for PCE (line 6) with an RNA probe for *Him* and a ß-Galactosidase antibody (Figure 4.3.1). For the same experiment with a different transgenic PCE line (34) refer to the Appendix. The initial expression pattern of the PCE construct was tested in three different lines (6, 34 and 41; see Appendix) which all have the same PCE expression pattern in the pericardial cells.

The PCE construct recapitulates the expression pattern of the *Him* transcript. In late stage 12 embryos all heart cells expressing *Him* also express the PCE (Figure 4.3.1 A-F). The same is true for stage 13 *Drosophila* embryos (Figure 4.3.1 G-L). From stage 15 onwards, there is a discrepancy between the PCE and wild type *Him* expression (Figure 4.3.1 M-O). While *Him* expression decreases and eventually ceases completely by stage 17, the PCE reporter is still strongly visible at this stage (not shown).

This discrepancy is presumed to be due to the higher stability of the ß-Galactosidase protein than that of the *Him* RNA. The expression pattern of the *ß-Galactosidase* RNA resembles that of the *Him* RNA expression more closely than the ß-Galactosidase protein in these late embryonic stages. As the comparison of the *ß-Galactosidase* RNA versus its protein in Figure 4.3.2 shows, the protein outlasts the RNA considerably (compare 4.3.2 J and K).

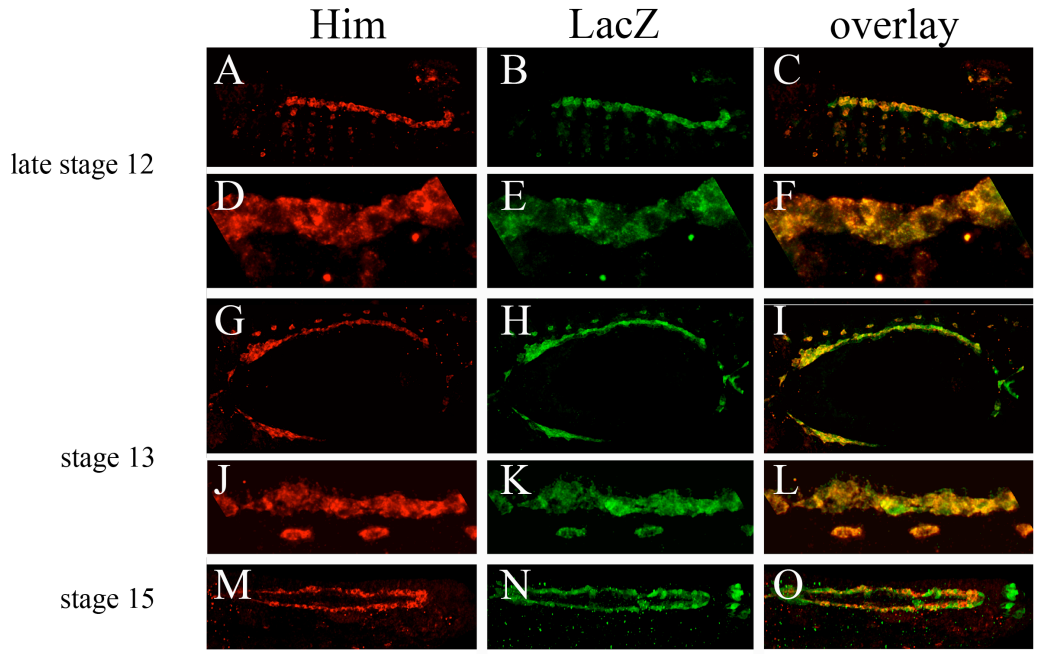


Figure 4.3.1: *Him* and the PCE reporter are expressed in the same heart cells.
Him fluorescent RNA *in situ* and anti-βGalactosidase immuno-stain on embryos transgenic for the PCE construct (line 6; for images of line 34 see Appendix). Images A - F show a lateral view of a late stage 12 embryo; Images G - L show a dorsal view of a stage 13 embryo and Images M - O show a dorsal view of a stage 15 embryo.
A, D, G, J, M fluorescent *Him* RNA *in situ* (red); **B, E, H, K, N** anti-βGalactosidase antibody (green); **C, F, I, L, O** overlay.

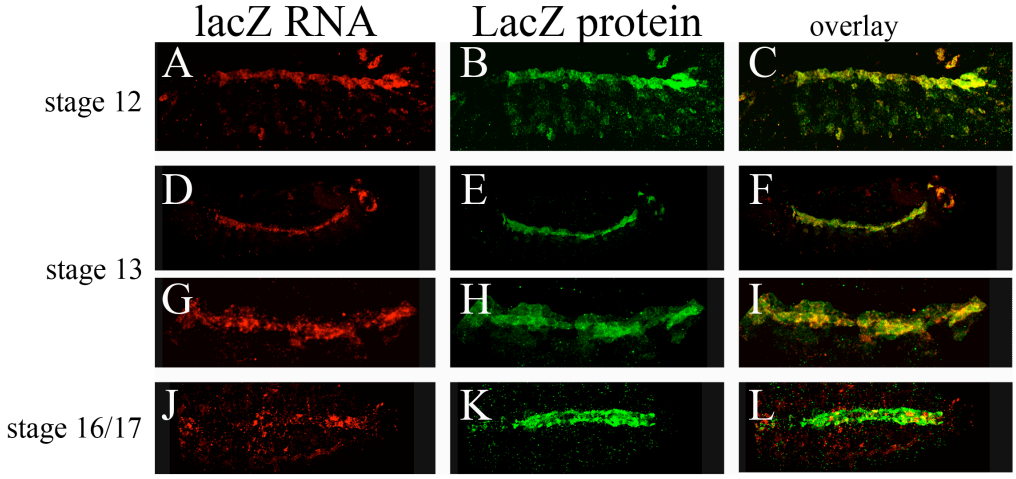


Figure 4.3.2: The *βGalactosidase RNA* is less stable than its protein.
βGalactosidase fluorescent RNA *in situ* (red) and immuno-stain for its protein (green). Images A, B, C show a lateral view of a late stage 12 embryo; Images D, E and F show a lateral view of a stage 13 embryo; G, H and I show the same embryo as in D, E and respectively at a higher magnification; J, K, L show a dorsal view of a stage 16/17 embryo.
A, D, G, J fluorescent *βGalactosidase* RNA *in situ* (red); **B, E, H, K** anti-βGalactosidase antibody (green); **C, F, I, L** overlay.

4.4 The phylogenetic footprint of the PCE sequence reveals three conserved areas.

The community of *Drosophila* researchers has published and annotated the sequence of *D. melanogaster* and *D. pseudoobscura* (Adams et al., 2000, Celniker et al., 2002, Richards et al., 2005) and 10 further species (*D. sechellia, D. simulans, D. yakuba, D. erecta, D. ananassae, D. persimilis, D. willistoni, D. mojavensis, D. virilis* and *D. grimshawi*; *Drosophila* 12 Genomes Consortium, 2007). These data have made detailed comparisons of the different sequences for corresponding regions possible. The interpretation of these so-called "Phylogenetic Footprints" (reviewed by Dickmeis and Müller, 2004) is based on the theory that a functional sequence has more evolutionary pressure to remain the same as a non-functional sequence, thus functional sequences should remain more similar across species than non-functional sequences.

At the time of my analysis the assembly of the ten additional *Drosophila* species was not yet complete and published. Because of this I used the *D. melanogaster* PCE sequence and the BLAST (discontiguous Basic Local Alignment Search Tool) to search the NCBI trace archive (http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?) for the raw sequence project sequences. I edited the retrieved sequences and analysed them using various alignment tools. I retrieved sequence data for ten of the *Drosophila* species with the sequence of *D. melanogaster* (*D. sechellia* sequences were unavailable at the time, for the different species see figure 1.3.1).

Even though the *Drosophilae* genomes were not yet assembled at the time of this analysis, a later check of the assembled data supports these results (Figure 4.4.1). Using the Vista Browser 2.0 (Dubchak et al., 2009), I retrieved the intergenic region between the *Him* and *Her* genes. However, even this much later data only contains alignment data for 5 *Drosophila* species (*D. simulans, D. yakuba, D. erecta, D. ananassae* and *D. pseudoobscura*). As can be seen in Figure 4.4.1 A and B the PCE region is one of two instances where conservation is evident in all six *Drosophila* species used in this alignment. If the conservation diagram generated by the Vista Browser is compared to the alignment I generated (Figure 4.4.2), it is obvious that using more species and retrieving the corresponding sequences manually produces an alignment with higher resolution. While the area I identified is conserved between D. melanogaster, D. yakuba and D. erecta, its conservation in the other three species (D.

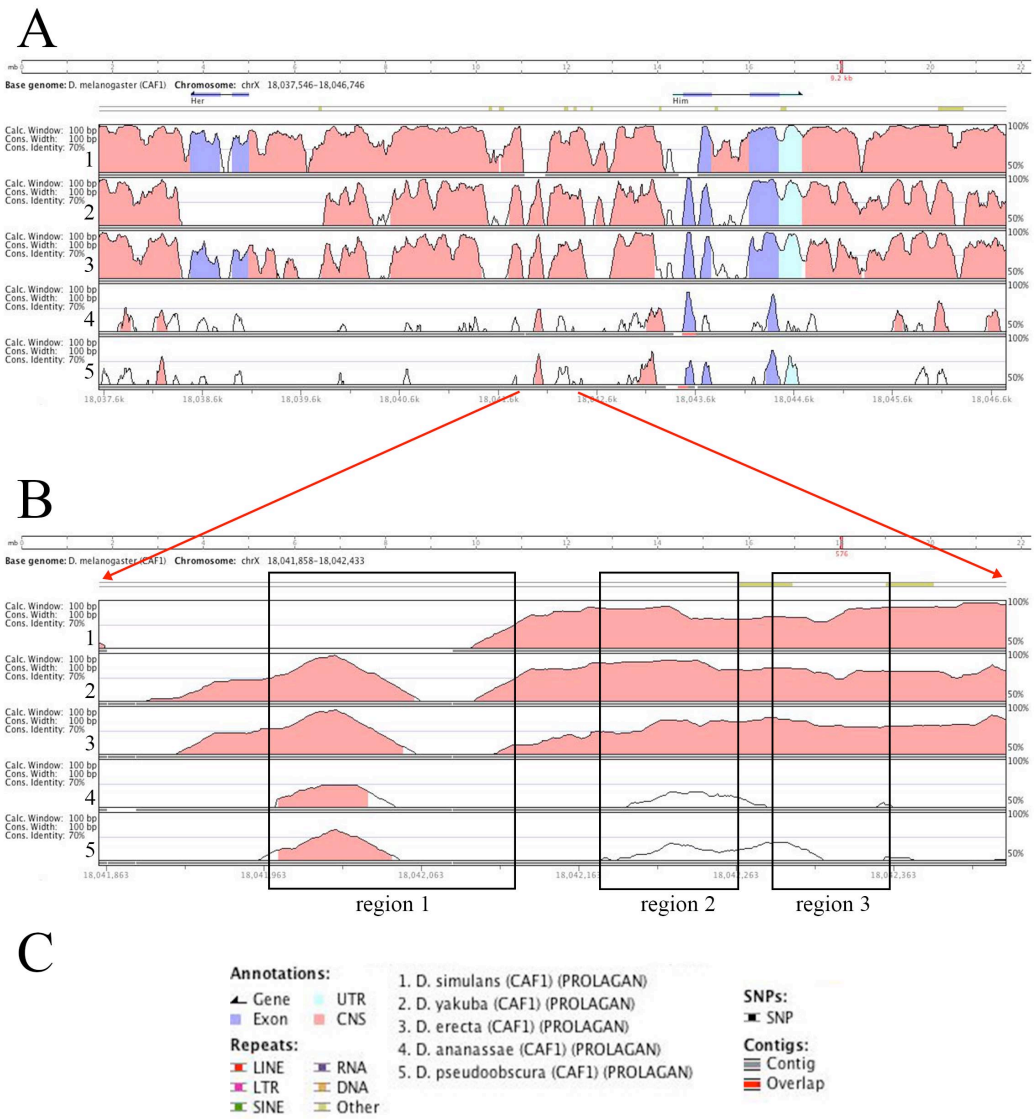
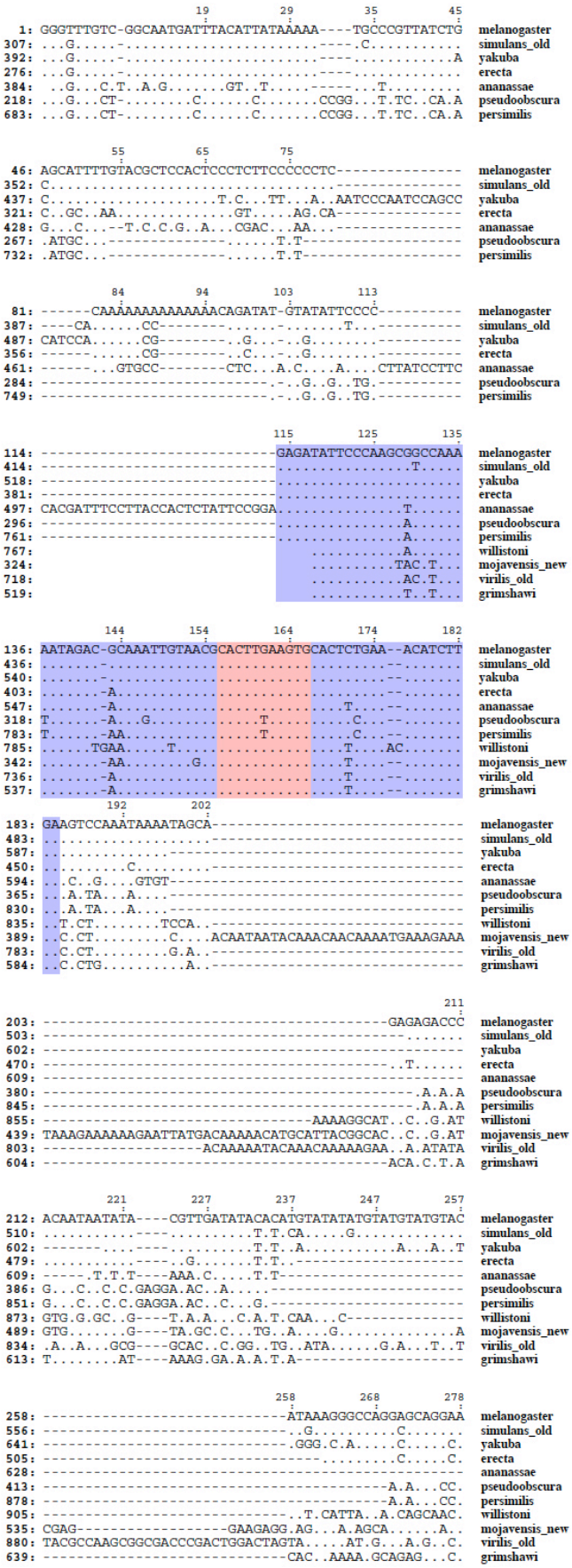


Figure 4.4.1: Conservation of the upstream region of *Him* as generated by the Vista Browser 2.0.

Multiple species alignment, based on the assembled sequences of the 12 Drosophila Genome Consortium (2007), generated by the whole-genome alignment browser Vista 2.0 (Dubchak et al., 2009; http://pipeline.lbl.gov/cgi-gin/gateway2?bg=droMel_caf1&selector=visatpoint)

A Vista plot of the complete intergenic region between the *Him* and *Her* genes. **B** Vista plot for the Set 3 sequence. Arrows indicate the approximate position of this fragment within the upstream sequence of *Him* and the boxes show the approximate positions of the three conserved regions as identified in Figure 4.4.1 A. **C** Legend for the annotation of the Vista Browser 2.0.


melanogaster
simulans_old
yakuba
erecta
ananassae
pseudoobscura
persimilis
willistoni
mojavensis_new
virilis_old
grimshawi

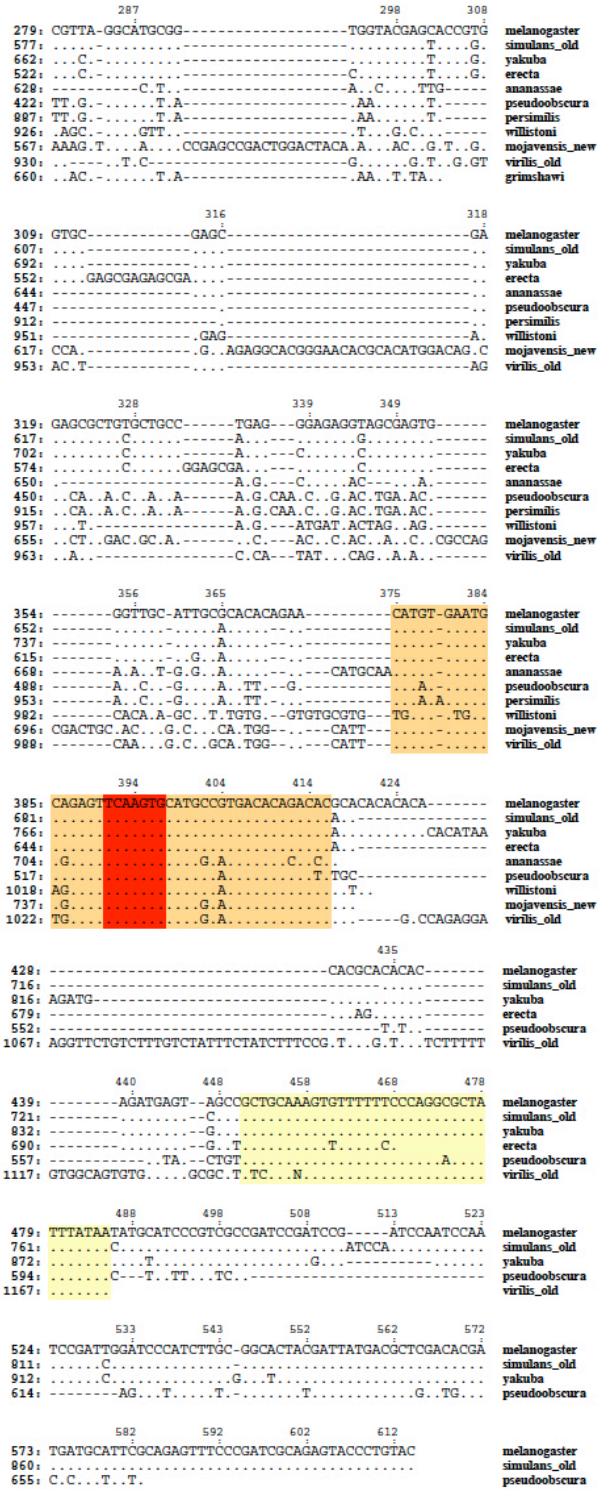


Figure 4.4.2: Phylogenetic Footprint of the PCE sequence across 11 *Drosophila* species.

MultiPipmaker alignment of the *D. melanogaster* PCE sequences with the putative PCE sequences I identified in the ten additional *Drosophila* species. The conserved regions are shaded: region 1 in purple, region 2 in orange and region 3 in yellow. The pink and red shaded sequences are putative Tinman binding sites.

simulans, D. ananassae, D. pseudoobscura) is less high (see boxes in Figure 4.4.1). Based on the results of the Vista Browser alone, I would not have identified the second and third conserved region. This illustrates that, despite the availability of such alignment tools as the Vista Browser, it is still helpful to manually generate multi-species alignments for specific region to achieve more detailed results.

In addition to the multiple alignments, I also generated pairwise alignments between *D. melanogaster* and each of the other ten *Drosophila* species to further rule out any algorithm specific artifacts. For these pairwise alignments, I again used both local and global alignment algorithms (ClustalW, VISTA (Dubchak et al., 2000; Mayor et al., 2000; Bray et al., 2003; http://pipeline.lbl.gov/cgi-bin/gateway2) and lalign (Huang, X. and Miller, M., 1991; http://www.ch.embnet.org/software/LALIGN_form.html). Each of these alignments also highlighted the same conserved regions.

The alignments identified three conserved regions within the PCE sequence, region 1 (purple underlay in Figure 4.4.2), region 2 (orange underlay in Figure 4.4.2) and region 3 (yellow underlay in Figure 4.4.2). Region 1 has a total length of 71 bp, region 2 of 40 bp and region 3 of 35 bp. Because of the high conservation level of these regions, I concentrated on these three conserved regions in my next step.

4.5 *In silico* search for transcriptions factor binding sites

To facilitate the search for transcription factor binding sites, I again made use of the information gathered in online databases. In order to find possible regulating factors for the transcription of *Him*, I analysed the PCE sequence of *Him* to find possible binding sites of known transcription factors. The PCE sequence was screened with the vertebrate and insect transcription factor binding sites derived from the TRANSFAC database (Wingende et al., 2001) and for binding sites of proteins found by a literature search for *Drosophila* mesodermal transcription factors (S. Elgar & M. Taylor, unpublished data, see Appendix).

With the help of the MatInspector program (Quandt et al., 1995; Cartharius et al., 2005), the *D. melanogaster* PCE sequence (which contains the conserved regions 1 to 3) was screened for possible binding sites. MatInspector uses the binding sites contained in the TRANSFAC database in its search process. The TRANSFAC database is mainly comprised of vertebrate transcription factor sites; however, due to the known

conservation of the properties of transcription factors between species, it is possible to identify potential transcription factor binding sites by this method. Since this database contains mainly vertebrate transcription factors and only very few *Drosophila* transcription factor binding sites, candidate *Drosophila* transcription factor binding sites were compiled into a list through a manual literature search (S. Elgar & M. Taylor, unpublished data). Candidate *Drosophila* transcription factors were chosen according to their spatial and temporal expression pattern during the development of the embryo. Binding sites were only included within this list if there was experimental evidence for their activity in the developing mesoderm. A sequence file was created for these sites and the PCE sequence was screened using the DS Gene Software suite (see appendix for the list of used binding consenses).

Figure 4.5.1 shows the three conserved regions of the PCE sequence and the binding sites for potential transcription factor binding sites identified in this analysis. For the conserved region 1 the analysis yielded the following candidate transcription factor binding sites: Tinman (pink underlay), Medea (grey box), vertebrate Oct-1 (blue box) and Bagpipe (green box).

The *Drosophila* homolog of the vertebrate Oct-1 is *nubbin*. Both (Oct-1 and *Drosophila nubbin*) are involved in cell fate determination (Neuman and Cohen, 1998; Irvine, 1999). *Drosophila nubbin* is mainly expressed in the ectoderm and plays an important role during neurogenesis and neurospecification. So far very little is known about its binding preferences. According to the only published DNA footprinting analysis it appears to have a fairly degenerate consensus (Neumann and Cohen, 1998). In their assay, they found that Nubbin protects four different sequences. The consensus sequence used by the TRANSFAC position weight matrix for the vertebrate Oct-1 (NNNNTGCAAATNAN; matrix identifier V$OCT1_Q6) approximately matches one of the four sequences protected by Nubbin (TTATGtAAgTAACC) identified by Neumann and Cohen (1998). However, two bases at position 8 and 11 (lower cases in the protected sequence) of the vertebrate Oct-1 consensus sequence are not identical to the sequence protected by Nubbin. It needs to be noted that the two mismatched bases are within the core binding area for Oct-1 as stated by Neuman and Cohen (1998). Thus, it is possible that these mismatches interfere with Oct-1 and/or Nubbin binding and that this site is not a functional binding site.

The conserved region 2 contains many of the same putative binding sites found in region 1 (see Figure 4.5.1). As in region 1, there is a Tinman binding site in

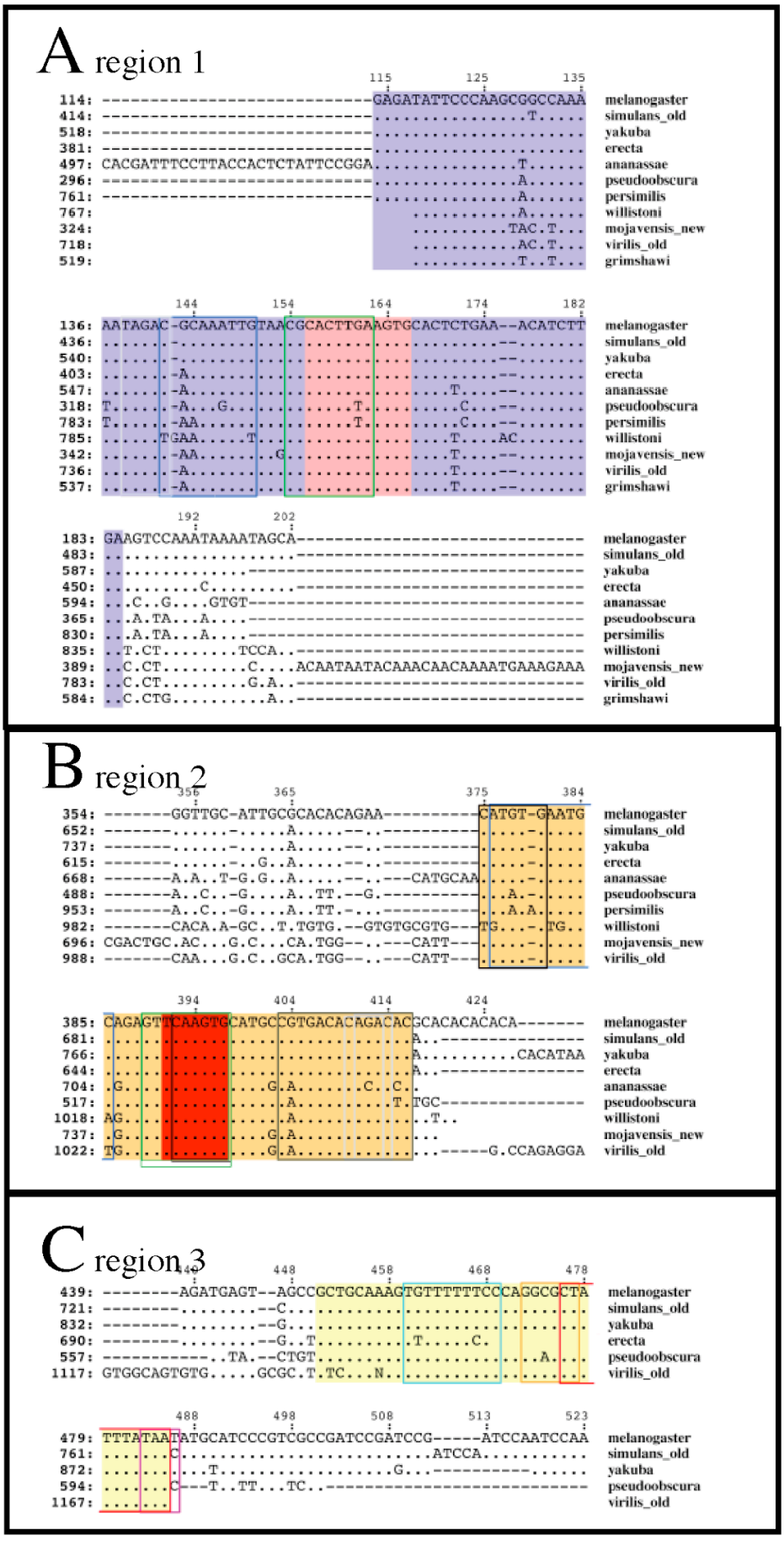

A region 1
B region 2
C region 3
melanogaster
simulans_old
yakuba
erecta
ananassae
pseudoobscura
persimilis
willistoni
mojavensis_new
virilis_old
grimshawi

Figure 4.5.1: Identified transcription factor binding sites in the conserved regions of the PCE.
The sequence of the conserved regions was screened for potential binding sites using MatInspector/ TRANSFAC database and a list manually compiled from literature. **A** Results for region 1. The potential Tinman binding site has a pink underlay, other identified putative binding sites are boxed: grey for Medea, blue for Oct-1 and green for Bagpipe. **B** shows the results for region 2. The potential Tinman binding site has a red underlay and all the other found putative binding sites are boxed: black for E-boxes, green for Bagpipe, dark grey for AP-1, light grey for Medea and blue for Oct-1. **C** shows the results for region 3. All potential binding sites are boxed: light blue for Dorsal, orange for Brinker, green for Mef2 and pink for Zfh-1.

combination with a potential Bagpipe binding site and a vertebrate Oct-1 site in close proximity upstream. Towards the 3' end of region 2 there are two Medea sites. Medea is a Smad4 protein that directly activates Tinman in response to Dpp signalling and is thus a potential regulator for *Him* (Chen and Fishmann, 2000; Cripps and Olson, 2002).

Overlapping this putative site is a vertebrate AP-1 site. AP-1 is a family of transcription factors that contain a basic region leucine zipper (bZip) and include the Jun, Fos and ATF subfamilies. A *Drosophila* equivalent of AP-1 family is *kayak*, which is involved in dorsal closure (Noselli and Agnes, 1999; Harden, 2002). However, Kayak is a single gene product and contrary to the gene products of the vertebrate AP-1 family, it does not need to dimerize to function.

The conserved region 3 contains potential Dorsal, Brinker, Mef-2 and Zfh-1 binding sites in the *D. melanogaster* sequence (see Figure 4.5.1 C). Yet all these candidate sites for transcription factor binding are either a weak match for the consensus sequences or are poorly conserved between the different *Drosophila* species. The potential Zfh-1 binding site is not conserved in *D. simulans*; the *Drosophila* species evolutionary closest to *D. melanogaster*. Furthermore, the last position of the possible Zfh-1 binding site is also not conserved in *D. pseudoobscura*. The Dorsal site is not conserved in *D. erecta* as two of the basepairs are different in this species. At the time of analysis there were no *D. mojavensis* sequence data available for this region.

4.6 Discussion

As shown in Figure 4.2.1, the PCE is the smallest enhancer fragment I have identified that is capable of driving expression of the ß-Galactosidase reporter in the endogenous *Him* expression pattern. This adds a good tool for the study of pericardial cells, as the PCE construct unambiguously identifies all pericardial cells of the *Drosophila* embryo. There is no similar line available for *zfh-1*, the only other known gene expressed in all pericardial cells during embryogenesis.

However, the PCE expression in late embryos (stage 15 onwards) does not resemble the expression pattern of the *Him* transcript. This is due to the higher stability of the ß-Galactosidase protein than the *Him* RNA. The comparison of the ß-Galactosidase protein and its RNA transcript illustrates this. If one compares the

expression of both RNAs, *Him* and *ß-Galactosidase*, it is obvious that the ß-Galactosidase RNA expression starts to decrease at a similar stage of the embryo as the *Him* RNA. Hendren et al. (2007) describe a similar effect: "We reasoned that the stable *ß*-Galactosidase protein perdured in the Svp cells at later stages of development. This was confirmed by analysis of *lacZ* expression by in situ hybridisation in stage 15 transgenic embryos."

Of the other possible enhancer regions tested for heart expression, Set 2 and Set 4 also showed ß-Galactosidase expression (Figure 4.2.2). As Set 2 and PCE (= Set 3) together make up Set 4 and each is capable of driving *Him*-like reporter expression in the *Drosophila* heart, it is likely that Set 2 and Set 3 resemble two different enhancers. A search for potential binding sites in Set 2 identified one putative Tinman binding site. The enhancer contained within Set 2 might also be a so-called "shadow enhancer" that provides a "safety-net" to ensure the appropriate expression pattern of developmentally important genes in the case of disturbance or disruption of the "main" enhancer (Perry et al., 2010). An alignment of the *D. melanogaster* Set 2 and PCE (= Set 3) sequence might possibly delineate any elements contained in both elements and that are essential for *Him* expression in the pericardial cells and should be considered for the future analysis of the complete *Him* regulation. It is possible that similar structures will facilitate the identification of regulatory elements needed for expression in all pericardial cells.

The phylogenetic footprinting performed on the PCE sequence revealed three regions of interest. The high levels of conservation across the eleven *Drosophila* species for these areas indicate a functional constraint that allowed these sequences to remain relatively unchanged during evolution. This functional constraint imposed upon this region is likely to be, at least in part, the transcriptional regulation of *Him* (The Flybase Consortium, 2003).

One needs to keep in mind that the phylogenetic footprints used for my analysis are, except for *D. melanogaster* and *D. pseudoobscura*, based on single "traces", i.e. single sequence reads of the sequencing projects and not the assembled genome. Thus, based on the trace data alone, it is not possible to say with absolute certainty if the region analysed in this study is placed directly upstream of the gene homologous to *Him* in the different *Drosophila* species. However, because of the similar results I obtained using the Vista Browser (see Figure 4.4.1) and the high level of conservation in the analysed PCE region of *Him* between all eleven *Drosophila* species, it is very unlikely

that the sequences I used to generate these alignments are not located in the upstream region of the *Him* homologs.

In most cases, the BLAST search of the NCBI trace archive yielded several traces for each species. As inaccurate sequence data is often generated in the beginning or towards the end of a sequence read, all traces retrieved for one species were aligned with each other. The resulting consensus sequence for each species used for the phylogenetic footprints is the result of these alignments (see chapter 2). Since there were very few discrepancies between the aligned single traces, I am confident that the used sequence data are correct.

The phylogenetic footprint identified three regions with a high degree of conservation. Studies have shown that approximately 22 to 26 % of the non-coding sequence between *D. melanogaster* and *D. virilis* are conserved and the average length of a stretch of conserved sequence is about 19 bp (Bergman and Kreitman, 2001). The conserved region 1 of the PCE sequence has a length of 72 bp, region 2 has a length of 42 bp and region 3 is 32 bp long. In their paper Bergman and Kreitman state that the majority of conserved non-coding sequences (intragenic and intronic) in *Drosophila* have a length ranging from 8 to 75 bp, with the distribution heavily tilted towards the shorter lengths. Thus, all three conserved areas of the PCE are within this range; the region 1 is however at the "longer" end of the range. The occurrence of conserved non-coding sequences of this length is fairly improbable, which makes this region and also the conserved region 2 highly interesting to me as their extra-ordinary length of conserved sequences might be due to the importance of these regions for the regulation of *Him*.

I based the computational search for possible transcription factors in the *D. melanogaster* sequence on the phylogenetic alignment of the PCE sequence and focussed especially on the three conserved regions. The sequence of the PCE element was searched for potential transcription factor binding sites with a manually generated set of binding sites (S. Elgar & M. Taylor, unpublished data, see appendix) and with the help of the Match program available from the TRANSFAC database (http://www.gene-regulation.com/pub/programs.html#match).

The search using the TRANSFAC database is only of use for *Drosophila* transcription factors that have similar specificities as their vertebrate orthologs, since the database contains mainly transcription factors for vertebrates and only very little data on *Drosophila* or other insect transcription factors. However, it has been shown for

Mef-2 for example, that the *Drosophila* and the mouse protein have identical specificities (Molkentin et al., 1996a; Brand, 1997). As there are many similarities in heart development between vertebrates and *Drosophila* on both molecular, morphological and physiological levels (Zaffran and Frasch, 2002 Medioni et al., 2009; Cammarato et al., 2011; Qian and Bodmer, 2012), it is possible that factors, especially if they are already known to be involved in heart or mesoderm development, are available in the vertebrate data but have yet to be characterised during *Drosophila* heart development. This makes the TRANSFAC database a useful resource for my study.

The transcription factors included in the manually generated list were chosen according to their expression pattern and, if known, their implication in heart development in *Drosophila*. In each of the two approaches, transcription factors were only included if there was experimental evidence for their binding characteristics.

To reduce the possibility of identifying too many false positive putative binding sites, stringency criteria were assigned to the consensus sequences used for the search. The threshold was defined at a detection rate of more than one site per 50 bp for all binding sites; sites that occurred more often were considered to be of too low stringency and were not used in the analysis. Exceptions to this are the sites taken from the papers of Knirr et al. (2001), Halfon et al. (2000) and Halfon et al. (2002). The binding sites described within these reports were included in the search file for putative binding sites as described by the authors, since their functionality and importance in *Drosophila* heart development has been proven and it is unlikely that many, if any, of the detected sites are false positive binding sites.

Potential Tinman binding sites were identified in the conserved region 1 and 2 by both search methods. Tinman is a homeobox protein that is necessary for heart development (Azpiazu and Frasch, 1993; Bodmer, 1993). Tinman is expressed in a sub-set of cardioblasts as well as in a sub-set of pericardial cells (Bodmer, 1993). The Tinman protein is a good candidate regulator for *Him* transcription. In their paper Hendren et al. (2007) find that “most of the cardiac enhancers for these Tinman target genes contain at least two Tinman binding sites within 300 bp of each other, which are critical to normal cardiac function”. The putative Tinman binding sites found in the conserved regions of the *PCE* are 224 bp apart and thus well within the criteria mentioned in this paper. This supports the possibility of Tinman as a regulator of *Him*.

Although Tinman is a very likely candidate regulator for the transcriptional regulation of *Him*, it cannot be the only factor involved in this process, since Tinman is

expressed in both pericardial cells and cardioblasts. *Him*, however, is only expressed in the pericardial cells of the heart. Furthermore, Tinman is only expressed in a sub-set of pericardial cells, while *Him* is expressed in all pericardial cells. Thus, it is necessary to search for further potential regulators of the PCE element.

Other putative transcription factor binding sites discovered in the conserved regions include two Oct-1, two Bagpipe, one AP-1 site, several E-Boxes and several Medea sites. The conserved regions 1 and 2 both contain an Oct-1 site. Even though both identified sites do not match the published consensus sequence perfectly, the sites are still possible within the position weight matrix used to search the PCE sequence and the un-conserved nucleotides in each possible site (in *D. pseudoobscura*, *D. virilis* and *D. ananassae*) would only weaken the binding according to the Selex data contained in the TRANSFAC matrix (V$OCT1-Q6) used to identify this sequence.

The sequence of the Bagpipe binding site (Zaffran et al., 2002) is in a large part identical to the sequence of the Tinman binding site (cgCACTTAG vs. CACTTGA for Tinman), which increases the likelihood of finding a Bagpipe site superimposed on a Tinman site considerably. Bagpipe has been shown to promote the differentiation of the visceral mesoderm (Azpiazu and Frasch, 1993; Furlong et al., 2001; Jagla et al., 2001; Cripps and Olsen, 2002), which, like the heart, also derives from the dorsal mesoderm. Because of its distinctive role in the development of the visceral mesoderm and the absence of *Him* expression in the visceral mesoderm at any developmental stage of the *Drosophila* embryo, it is unlikely that Bagpipe is a transcriptional activator of *Him* expression. It is however possible that Bagpipe is involved in *Him* repression in the developing visceral mesoderm and that if Bagpipe is present, it binds to the combined Tinman/Bagpipe site and thus precludes Tinman binding and activation of *Him*.

E-boxes are the binding sites for transcription factors with a bHLH (basic helix-loop-helix) motif. The bHLH transcription factors are found throughout the eukaryotic kingdom and are involved in the development of many tissues (Atchley and Fitch, 1997; Ledent and Vervoot, 2001; Zinzen et al., 2009; Ozdemir et al., 2011). So far, most of the described bHLH transcription factors have been shown to play important roles in the differentiation and specification of cells and tissues (Molkentin and Olson, 1996a; Molkentin and Olson, 1996b). The recognised core sequence of bHLH factors is CANNTG (Ferré-D'Amaré et al., 1993; Ellenberger et al., 1994; Ma et al., 1994). The two nucleotides in the middle of this sequence are necessary for the specificity of the binding and vary according to the bHLH factor.

The bHLH transcription factors known to be involved in *Drosophila* mesoderm development include Twist (Bate et al., 1999; Furlong et al., 2001; Cripps and Olson, 2002; Ozdemir et al., 2011), Nautilus (Keller et al., 1997) and Hand (Kölsch and Paululat, 2002). Thus, the detection of E-boxes within the enhancer fragment of *Him* might be expected.

The identified AP-1 site is based on the data available for the vertebrate transcription factor AP-1 (activator protein 1). The vertebrate AP-1 protein consists of dimers of the Jun/Fos protein (Jonat et al., 1990; Ameyar et al., 2003). A *Drosophila* equivalent of AP-1 is the gene *kayak*. The *kayak* gene is expressed in the ectoderm and is involved in the dorsal closure of the embryo (Noselli and Agnes, 1999). The process of dorsal closure and the formation of the heart tube along the midline of the embryo are closely connected. Thus, there is a possibility for Kayak to be involved in the formation of the heart through signalling cues from the ectoderm to the mesoderm.

The transcription factor Medea is part of the transforming growth factor beta (TGFß) receptor-signalling pathway in the nucleus (Wisotzkey et al., 1998; Lall and Patel, 2001). Among others, it has also been shown to interact genetically with Dpp and Mad (Mothers against dpp). Mad and Medea are both part of the Smad family of transcription factors. Some functions of the Medea protein are in the dorsal-ventral axis specification and in the development of the *Drosophila* heart (Chen and Fishmann, 2000; Cripps and Olson, 2002). Knirr and Frasch, 2001 show that Dpp is capable of activating the even-skipped enhancer through activating the phosphorylated form of Mad (pMad), this is also necessary to maintain *eve* transcription *in vivo*. However, because of the short published Medea binding sequence of 4 nucleotides (GTCT, described in Knirr and Frasch, 2001), this site is found approximately every 256 bp. This high probability of finding this sequence within any other sequence makes it very hard to distinguish between potential genuine Medea sites and the average occurrence of this sequence. The degree of conservation of a potential site between the species might be able to help discriminate genuine from “false-positive” Medea sites. Because of its implication in *Drosophila* heart development and its location within the highly conserved regions of the PCE element, Medea remains an interesting potential regulator of *Him* transcription.

Within the conserved region 3 of the PCE, the discovered binding sites were less well conserved. This is most likely due to the fact that the degree of conservation of this region is generally lower than that of the other two conserved regions. Furthermore,

there is less sequence data available for this region. The Dorsal binding site is fairly well conserved in all *Drosophila* species, there are only two nucleotide differences in *D. erecta*. The Dorsal protein is important for the establishment of the dorsoventral axis of the developing embryo (Simpson, 1983; Lall and Patel, 2001). Around stage 7 and later Dorsal is expressed in the mesoderm and it is thus possible that it plays a role in the early development of the presumptive heart mesoderm (Rushlow et al., 1989; Roth et al., 1989).

Brinker is a transcriptional repressor (Lall and Patel, 2001), involved in the Dpp pathway (Affolter et al., 2001) and expressed in the mesoderm. This potential binding site is not conserved in *D. pseudoobscura*. It is not possible to say anything about the conservation of this site in *D. ananassae*, *D. persimilis*, *D. willistoni*, *D. mojavensis* or *D. grimshawi*, since at the time of analysis it was not yet possible to retrieve sequence data for this part of the PCE fragment.

The potential Mef2 binding site found in the third conserved region of the PCE matches the used consensus site only poorly. Again it is not possible to say anything about its conservation in *D. ananassae*, *D. persimilis*, *D. willistoni*, *D. mojavensis* or *D. grimshawi* as the sequence data was lacking at the time of the analysis. Mef2 is an important transcription factor involved in mesoderm development and is expressed in all cardioblasts of the heart and might be expressed early in the precursors of the pericardial cell lineage (Bour et al., 1995; Taylor et al., 1995; Bate et al., 1999; Furlong et al., 2001). As *Him* is only expressed in the pericardial cells of the heart, Mef2 is only a candidate for very early transcriptional regulation, if at all.

The fourth potential binding site found in region 3 of the PCE element is a Zfh-1 binding site. Zfh-1 is a transcription factor that is expressed in the majority of the pericardial cells of the *Drosophila* heart; its early expression pattern includes all pericardial cells (Lai et al., 1991; Su et al., 1999); and it has an expression pattern very similar to that of *Him*. Thus it would be a promising candidate to regulate *Him* transcription. But the consensus for this site has a length of only 4 nucleotides and because of this it is hard to distinguish between genuine sites and the average occurrence of this site within any sequence as already described for the Medea sites. Additionally, this short sequence is not conserved in *D. simulans* and *D. pseudoobscura* (no sequence data is available for *D. ananassae*, *D. persimilis*, *D. willistoni*, *D. mojavensis* or *D. grimshawi*). As *D. simulans* is closely related to *D. melanogaster*, this deviation from the *D. melanogaster* sequence decreases the likelihood that this Zfh-1

site is of functional importance for the regulation of *Him*. Postigo et al. (1999) show that Zfh-1 also binds to E-boxes of the CACCTG sequence; there is not a potential binding site with this sequence in the whole of the PCE sequence either.

The bioinformatic analysis of the PCE sequence described above has generated several interesting starting points for the functional analysis of it. The high degree of conservation of several of the identified possible transcription factor binding sites highlights these sites and factors as potential regulators. Furthermore, the discovery of the conserved regions within the PCE might lead to the discovery of so far unidentified transcription factor binding sites. In the following two chapters I will describe the deletion and mutational analysis I conducted to relate the results described in this chapter to their possible function *in vivo*. In their recent papers, both Halfon et al. (2011) and Biggin (2011) stress the importance to "rigorously assess" any data generated by computational analyses.

Chapter 5

Tinman regulates *Him* expression

5.1 Introduction

As described in the previous chapter the transcription factor Tinman is implicated as a possible regulatory factor for *Him*. *tinman* null mutants do not develop a heart (Azpiazu and Frasch, 1993; Bodmer, 1993). Its vertebrate ortholog Nkx2.5 has been implicated in vertebrate heart development and, furthermore, in congenital heart disease (Schott et al., 1998; Prall et al., 2002). Tinman is the earliest known marker for heart development in *Drosophila* (Bodmer, 1993). Known Tinman targets include *ß3-tubulin, hand, seven-up, pannier, Mef-2, midline, Dsur* and *Toll* (Akasaka et al., 2006; Gajewski et al., 1997; Gajewski et al., 2001; Han and Olson, 2005; Kremser et al., 1999; Ryan et al., 2007; Ryu et al., 2011; Wang et al., 2005). So far, no Tinman target that is exclusively expressed in pericardial cells of the heart has been identified.

The bioinformatic analysis described in the previous chapter points out several potential Tinman binding sites in two different conserved regions of the PCE. In addition to these results, several separate ChIP studies have found that Tinman binds to a region upstream of the *Him* gene (Liu et al., 2009; Junion et al., 2012; Jin et al., 2013). These results strongly indicate Tinman as a regulator of *Him* transcription; however, as pointed out in Biggin (2011) and Halfon et al. (2011), it is still necessary to prove the functional importance of Tinman for *Him* regulation and identify the actual relevant binding sites.

In the conserved region 1 there are two overlapping putative Tinman binding sites, one on the plus and one on the minus strand. There is one substitution (G to T) in the second last position in both sites. The potential Tinman binding on the minus strand is not the ideal consensus (CACTT***G***A), the G in the second last position is exchanged for a C. However, this sequence (CACTT***C***A) has been shown to be functional and is considered as a potential Tinman binding site (Gajewski et al., 2001). Another putative Tinman binding site is located within the conserved region 2 of the PCE. This is a single site on the plus strand, which is completely conserved between all eleven *Drosophila* species used for the phylogenetic footprint in chapter four. The occurrence

of two or more functional Tinman binding sites in close proximity (ca. 300 bp) is a prominent feature in most known Tinman targets (Hendren et al., 2007).

5.2 *Him* expression in *tinman* null mutants

In wild-type stage 10 and 11 embryos *Him* is expressed in the dorsal mesoderm, including the very top of the mesodermal crest in which the heart precursors are specified. *tinman* null mutants (tin^{EC40}) lack *Him* expression in the dorsal mesoderm (see Figure 5.2.1; compare arrows in B and D). There are two possible explanations for the lack of *Him* expression: either Tinman is necessary for early *Him* expression in the dorsal mesoderm or the cells that express *Him* are not present in the *tinman* null mutant embryos. The Tinman ChIP-on-chip analysis of the Furlong lab (Liu et al., 2009) showed that Tinman binds to a *Him* upstream sequence *in vivo*. According to the data given in this publication, the regions bound by Tinman do not include the PCE sequence. I have, however, identified a few discrepancies between the annotation of their data and the data of the GBrowse tool of Flybase that I based my results on, thus it is possible that the identified region actually corresponds to the PCE. The region identified as binding Tinman in the 2012 paper by Junion et al. contains the PCE sequence identified in chapter four, as does the region indicated in the Jin et al., 2013 paper. Liu et al. (2009) have classified *Him* as a somatic muscle gene and do not refer to *Him* expression in the heart. None of the cited studies have confirmed the ChIP results with any expression studies to link Tinman function with *Him* expression.

tinman null mutants do not form a heart and at stage 12/13 lack precursors for it (Bodmer, 1993). However, in *tinman* null mutants the spreading and migration of the dorsal mesoderm is unaffected and the lack of visceral and cardiac structures is due to the failure to specify these tissues (Bodmer; 1990) and thus the precursor cells should still be present and be exposed to other developmental "cues" during stage 10 of embryonic development (e.g. Decapentaplegic, Wingless).

Heartless (Htl) is a FGF receptor necessary for the migration of the mesoderm and is expressed in the early mesoderm (stage 9) and its expression is maintained in the developing dorsal mesoderm until stage 12 (Shishido et al., 1993). Within the dorsal mesoderm, *Heartless* is only expressed in a distinct cluster of cells that includes the

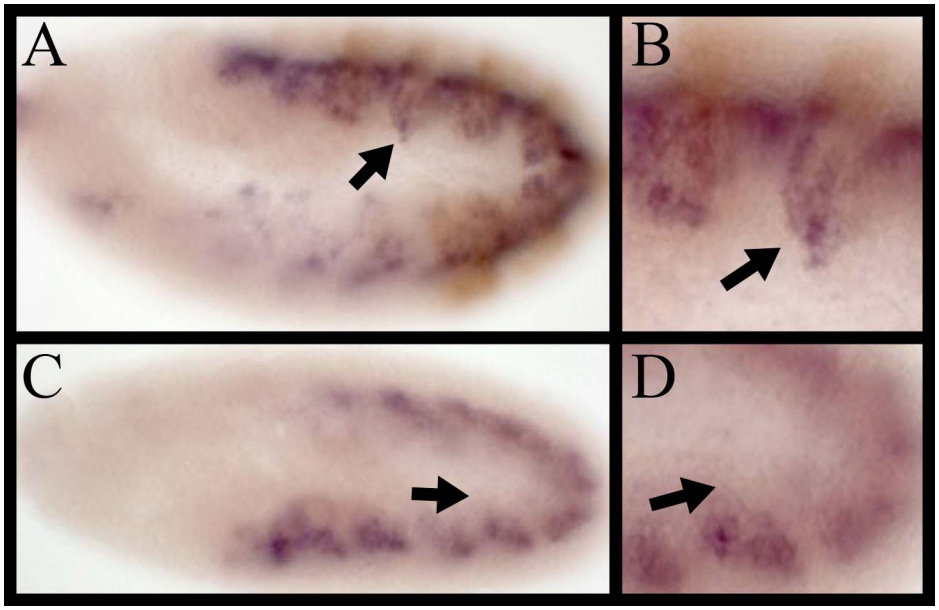


Figure 5.2.1: *tinman* null mutants lack Him expression in the dorsal mesoderm.
Him RNA *in situ*; homozygous *tinman* null mutants were identified by the absence of LacZ reporter of the balancer. B and D are close-ups of A and C, respectively.
A, B Control embryo (not a homozygous mutant for *tinman* as can be seen by the LacZ staining (brown)). The arrow indicates the heart precursors labelled by *Him* (blue). **C, D** Homozyogous *tinman* null mutant. The arrow indicates the location of the heart precursors at the very top of the mesodermal crest.

precursors of the Eve-positive pericardial cells and the precursors of the somatic muscles DA1 and DO2 from stage 10 onwards (Shishido et al., 1993; Michelson et al., 1998a; Halfon et al., 2000). Michelson et al., 1998b also state that "their experiments indicate that Heartless acts independently (or possibly upstream) of *tinman"*. Based on these data, *Heartless* expression should be unaffected in *tinman* null mutants, but so far this has not been tested.

Stumps (in the literature also referred to as Heartbroken and Dumbfounded) is expressed in a very similar pattern in the dorsal mesoderm to *Heartless* and in the same cell cluster as *Heartless* (Casal and Leptin, 1996; Halfon et al., 2000; Halfon et al., 2002). *Stumps* is also part of the FGF receptor signalling pathway and is a downstream target of Heartless and is involved in activation of the MAPK kinase cascade in response to FGF signalling (Michelson et al., 1998b; Vincent et al., 1998). Michelson et al., 1998(b) state that Heartless acts independently of Tinman regulation with a high probability. As *Stumps* is downstream of *Heartless* in the signalling cascade, it is likely that *Stumps* expression will also be unaffected in *tinman* null mutants.

Upon verifying that *Stumps* expression is present in *tinman* null embryos (see arrows in Figure 5.2.2 I), I used *Stumps* to show that the early heart precursors are present in the dorsal mesoderm of *tinman* null mutants but that these cells do not express *Him* (Figure 5.2.2 D). While *Stumps* is expressed in the early heart precursors of *tinman* null mutant embryos, it is noteworthy to mention, that its expression appears less strong than in the wild-type embryos. As these stains were done in parallel and are thus comparable this implicates that Tinman function is integrated into the *Heartless* signalling cascade after activation of the Heartless receptor.

As can be seen in Figure 5.2.2, Stumps and the *Him* transcript co-localize in the early heart progenitors of wild-type embryos (see arrows in Figure 5.2.2 C). *Him* expression can be seen in the progenitors of the later Eve-positive cells (marked by Stumps) and a few cells in the surrounding dorsal mesoderm are likely to be the precursors of the remaining heart cells. As expected, the expression of *Stumps* is present in *tinman* null embryos (see arrow in Figure 5.2.2 F). In these embryos, however, the *Him* RNA transcript is not expressed in the precursors of the heart cells. Consequently, the *Him* transcript and Stumps do not co-localize in the heart precursors of the dorsal mesoderm. This indicates that the early *Him* expression in the developing heart cells is *tinman* dependent.

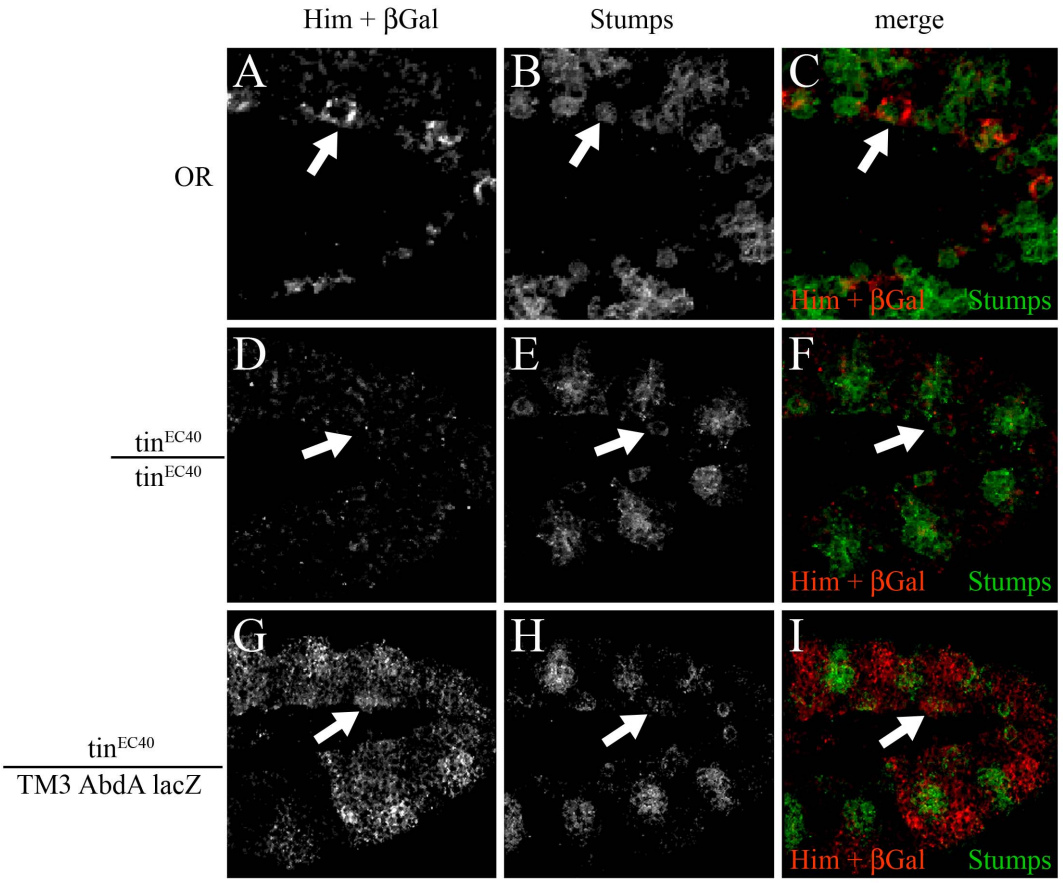


Figure 5.2.2: The heart precursors for the Eve-positive pericardial cells are present, but do not express *Him*, in stage 10 *tinman* null mutant embryos.

Confocal microscopy images of wild-type (*OR*) and *tinman* homo- and heterozygous mutant embryos stained for the *Him* and *βGal* RNA transcripts (red, Fluorescent *in situ* hybridisation) and Stumps antibody (green). C, F and I are merges of the single channel images.

A, B, C Wild-type embryo showing the *Him, βGal* and Stumps expression. The arrow indicates overlapping expression. **D, E, F** Homozygous *tinman* null mutant shows the lack of *Him* transcript while the Stumps protein is expressed in a wild-type like manner (indicated by the arrow). **G, H, I** Heterozygous *tinman* mutant embryo demonstrates the ßGal pattern and the overlap of the *Him* and Stumps expression patterns as in the wild-type embryo (arrow).

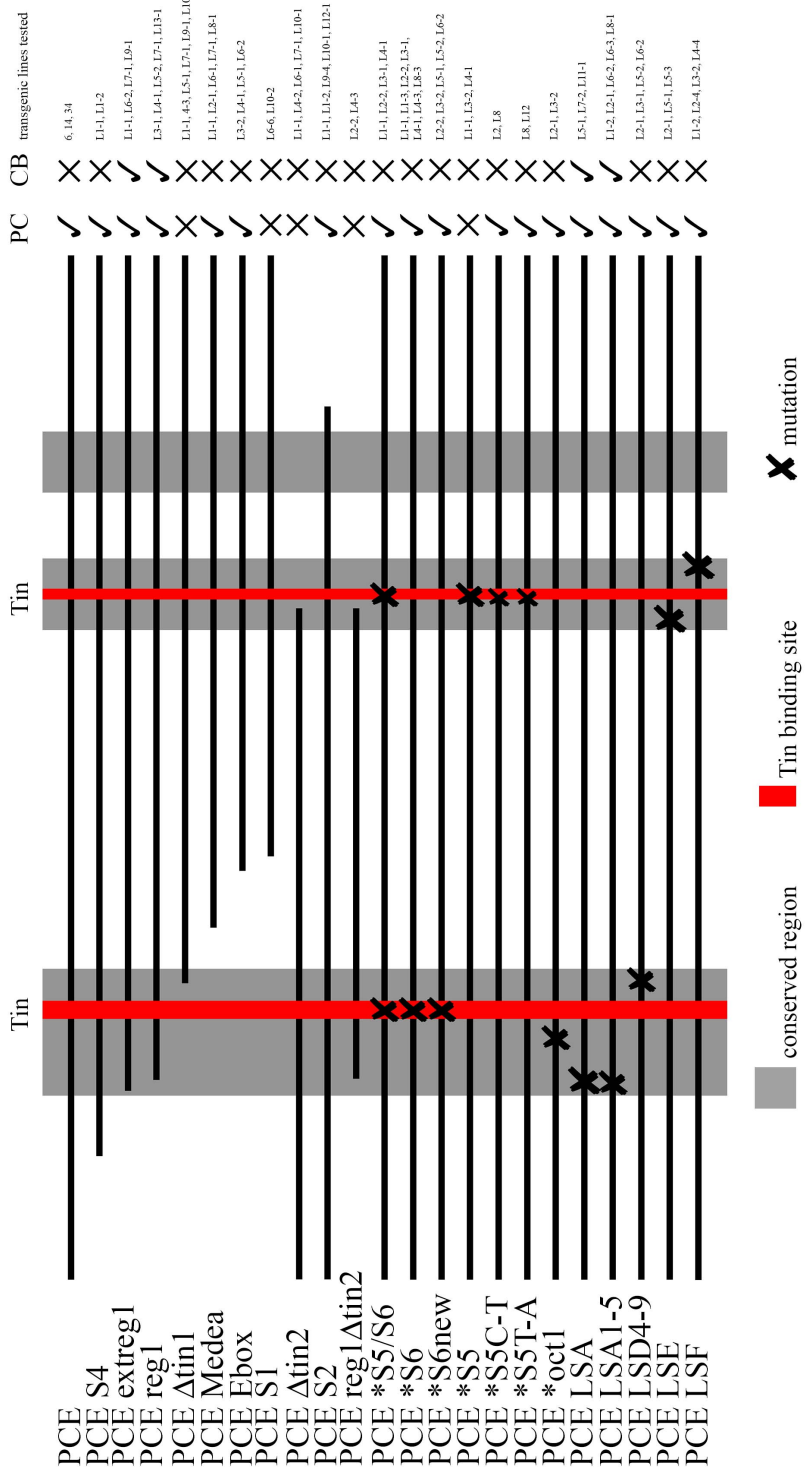


Figure 5.3.1: Schematic representation of all PCE enhancer constructs and their expression in the heart.

The conserved regions are shaded in grey and the Tinman binding sites are indicated by the red stripe. Crosses represent the position of the introduced mutation. PC: pericardial cells, CB: cardioblast, tick/cross: the enhancer construct is (not) expressed in these cells. The schematic is drawn in scale; the length of the PCE is 574 basepairs.

5.3 Deletion of the conserved putative Tinman binding sites

Deleting a sequence of interest in the context of a reporter construct is a common approach in trying to elucidate the function of a sequence of interest *in vivo*. The question asked is whether the expression pattern of the reporter construct with the deletion is altered when compared to the original reporter construct. Several papers describing Tinman target genes have utilized this approach in their analyses (for example: Gajewski et al., 1997; Wang et al., 2005; Ryu et al., 2011). Figure 5.3.1 is a schematic representation of all enhancer constructs generated in this study.

My initial test of whether Tinman is involved in *Him* regulation was to delete the sequences in the PCE up to and including the potential Tinman binding sites. If Tinman does activate *Him* transcription through any of the putative sites, deletion of these sites should abolish or reduce reporter activity if the sites do not function redundantly with each other. I made two different deletion constructs, one with a 5' deletion of 165 bp (called PCE Δtin1) and the second with a 198 bp 3' deletion (called PCE Δtin2; see Figure 5.3.1 for an overview of all constructs generated in this study; all enhancer constructs are linked to a βGal reporter).

Both the PCE Δtin1 and PCE Δtin2 deletion constructs completely abolish reporter expression in the pericardial cells. As Figure 5.3.2 shows, there is no visible staining of the heart in the transgenic embryos of either construct (see stars in Figure 5.3.2). All embryos shown in Figure 5.3.2 were stained at and for the same time and with the same reagents. This ensures that the lack of staining in the embryos carrying the transgenic construct is not due to a mistake in the procedure and that the staining intensity (if present) is comparable. A further internal control for the correctly functioning procedure is the ectopic expression of the LacZ reporter gene in PCE Δtin2 (see arrowheads in Figure 5.3.2 D). In the fly line carrying the PCE Δtin2 deletion construct, the LacZ protein is expressed in very large, round cells that are located dorsally to the heart. Judging by the size, shape and position of the cells ectopically expressing LacZ, they are part of the amnioserosa of the developing embryos.

These two deletions show that both Tinman binding sites are possibly functionally relevant in the transcriptional regulation of *Him*. Contrary to many published examples (Gajewski et al., 1999; Akasaka et al., 2006; Ryan et al., 2007), the loss of one binding site within the PCE sequence is sufficient to abolish reporter gene

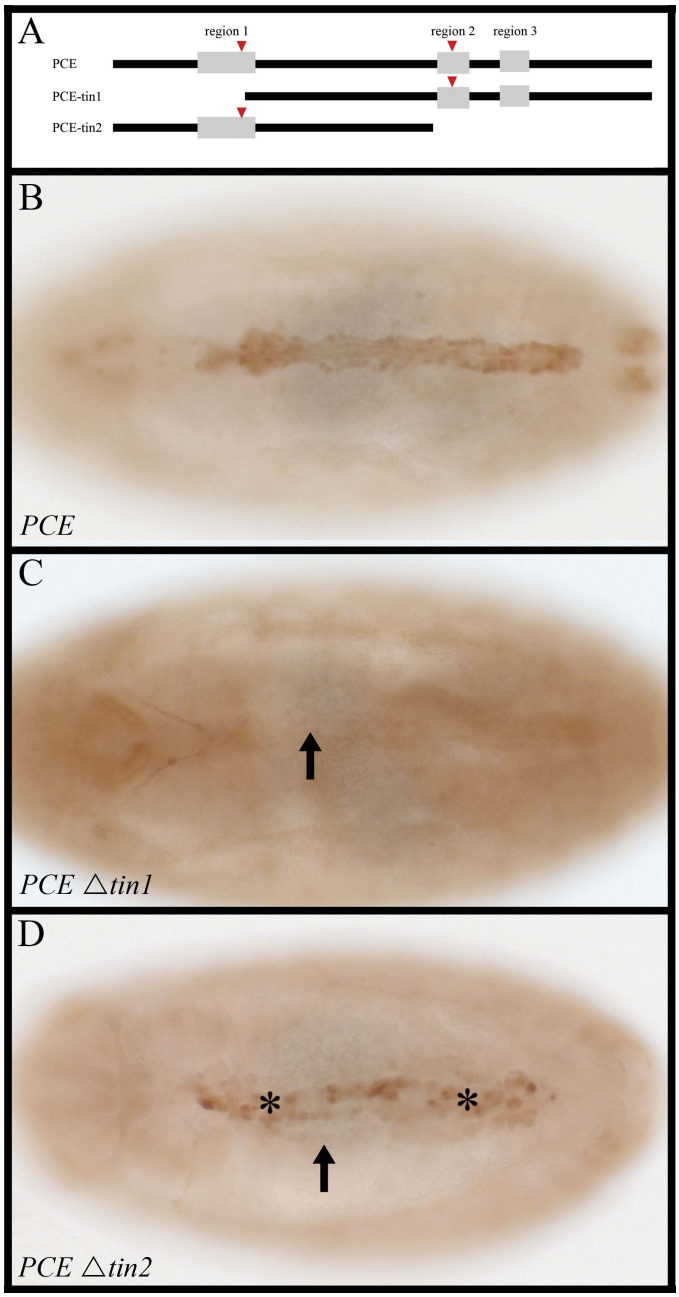


Figure 5.3.2: Deleting the potential Tinman binding sites in the PCE (Pericardial Cell Enhancer) abolishes reporter gene expression in the heart.

A Schematic representation of the constructs used in this Figure; the grey rectangles represent the conserved regions (see chapter 4) and the red triangles indicate the locations of the putative Tinman binding sites. **B** LacZ reporter gene expression for the full-length wild-type PCE construct. **C** LacZ reporter gene expression is lost when the 5' sequence up to and including the first Tinman binding site is deleted. The arrow points to the location of the heart. **D** Deleting the second Tinman binding site and the following 3' sequence abolishes reporter gene expression in the heart (position indicated by the arrow) but induces ectopic expression in amnioserosa cells (indicated by stars).

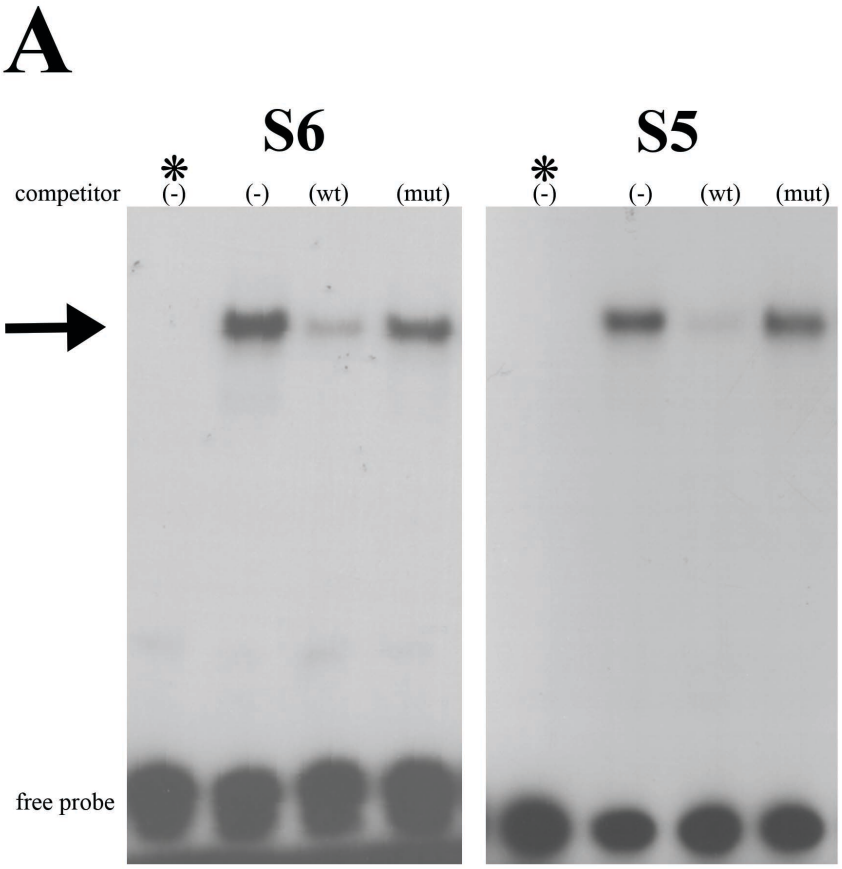
expression. However, further investigation is needed as both deletions remove large stretches of sequence, which could possibly contain other transcription factor binding sites or other regulatory elements that cause the loss of reporter expression.

5.4 Electrophoretic Mobility Shift Assays (EMSA)

Electophoretic Mobility Shift Assays (EMSAs) are designed to test if a protein specifically and selectively binds to a labelled DNA sequence *in vitro* (e.g. Mohun et al., 1989; for a review see Jiang et al., 2009). Adding un-labelled DNA of the same or mutated sequence tests the specificity of the binding. This method relies on the ability of a native polyacrylamide gel to separate complexes by size; a DNA-protein complex will be larger than just the oligonucleotide on its own, and thus will run slower through the gel matrix when a voltage is applied. By labelling the DNA sequence the difference of the migratory distance between the bound and unbound labelled DNA oligonucleotide becomes visible.

I used this system to test if *in vitro* translated Tinman protein is capable of binding to the S5 (the 5' site, deleted in PCE Δtin1) and S6 (the 3' site, deleted in PCE Δtin2) putative binding sites contained within the PCE sequence. The oligonucleotide sequence and the introduced mutations were designed as 24-mers as previously described (Chen and Schwartz, 1995; Gajewski et al., 2001) by including a minimum of 8 bp of the surrounding sequence flanking the putative Tinman binding sites (Figure 5.4.1 B). The mutations I used for this assay have previously been described and shown to inhibit Tinman binding by Chen and Schwartz, 1995 and Gajewski et al., 2001 (see figure 5.4.1 B for the introduced mutations).

Figure 5.4.1 A shows the results of the EMSA for each potential binding site. The arrow points to the area containing any DNA probe bound by the Tinman protein. Figure 5.4.1 shows that the oligonucleotide for the S6 binding site (towards the 5’ end of the PCE, for the location of the S6 site within the PCE refer to Figure 5.5.3) specifically binds to the *in vitro* translated Tinman protein. The unprogrammed lysate control (indicated by a star above the lane in Figure 5.4.1 A) shows no binding to the oligonucleotide through proteins naturally contained within the lysate used to translate the Tinman protein. Upon adding an excess (100-fold) of the unlabelled wild-type



B

S6
tin1 Wt (+) 5'-TTGTAACG***CACTTGAAGTG***CACTCTGA-3'
tin1 Wt (-) 5'-TCAGAGTG***CACTTCAAGTG***CGTTACAA-3'

tin1 Mut (+) 5'-TTGTAACG***GT***CT***AC***AAGTGCACTCTGA-3'
tin1 Mut (-) 5'-TCAGAGTGCACTT***GT***AG***AC***CGTTACAA-3'

S5
tin2 Wt (+) 5'-TGCAGAGT***TCAAGTGC***ATGCCGTG-3'
tin2 Wt (-) 5'-CACGGCAT***GCACTTGA***ACTCTGCA-3'

tin2 Mut (+) 5'-GATCTGA***GT***AC***AG***GTAAGGGCTTC-3'
tin2 Mut (-) 5'-GAAGCCCTTAC***CT***GT***AC***TCAGATC-3'

Figure 5.4.1: Tinman binds to the S5 and S6 sites *in vitro*.
Electrophoretic Mobility Shift Assays (EMSAs) using *in vitro* translated Tinman protein and radioactively labelled oligonucleotides. (-) no competitor; (wt) 100x wild-type competitor; (mut) 100x specifically mutated competitor; * unprogrammed lysate control.
A Tinman protein binds *in vitro* to the Tinman sites in region 1 (S6) and region 2 (S5). The binding can be competed by adding 100-fold excess of the wild-type oligo (labelled (wt)), but not by adding 100-fold excess of the muatated oligo (labelled (mut)) to the reaction. **B** Sequences of the oligos used in the assay. The Tinman binding sites are indicated in bold and italic; the introduced mutations are shown in bold red.

oligonucleotide into the mix (labelled (wt) in Figure 5.4.1 A), the band in the lane containing the labelled oligonucleotide, which has been bound by Tinman is considerably less strong than the band in the lane without added competitor (labelled (-) in Figure 5.4.1 A). This is due to the fact that both labelled and unlabelled probe compete for the available Tinman protein and the excess of unlabelled probe compared to labelled probe results in less available labelled probe that is able to bind the Tinman protein. However, adding a 100-fold excess of unlabelled, mutated oligonucleotide (marked (mut) in Figure 5.4.1 A) to the mix does not compete with the wild-type oliognucleotide for binding; the band has the same intensity as the un-competed band.

Figure 5.4.1 A also shows the behaviour of the S5 oligonucleotide in the same assay (for the location of the S5 site within the PCE refer to Figure 5.5.2). Similarly to the S6 oligonucleotide, the wild-type S5 oligonucleotide binds to the Tinman protein specifically as the binding can be competed with a 100-fold excess of unlabelled wild-type oligonucleotide (labelled (wt)), but not when a 100-fold excess of unlabelled mutated oligonucleotide (labelled (mut)) is added.

In vitro, Tinman binds to both potential binding sites. When the EMSAs for the S6 and S5 site are compared, the competition with the wild-type oligonucleotide is more successful for the S5 site, however, both EMSAs were done separately so they cannot be directly compared to each other and this observation could be just an artefact, this needs to be verified with further experiments such as competing the binding of one binding site with the unlabelled wild-type oligonucleotide of the other Tinman binding site. The ability of Tinman to bind both binding site sequences *in vitro* strengthens the idea that Tinman may indeed play a role in the *in vivo* regulation of *Him* transcription.

Each experiment only generates one shifted complex per reaction, there are no larger (i.e. slower) or smaller (i.e. faster) complexes visible on the X-ray film. This indicates that Tinman binds to the sequences in only one form; either as a monomer, a homodimer or as a heterodimer (with a protein present in the lysate). Further studies are needed to resolve the possibilities of binding partners for Tinman *in vitro* and also *in vivo*. Zaffran and Frasch (2005) have shown that Tinman can form homodimers and that it can also form heterodimers with Binou, so that the appearance of one single band does not allow for the exclusion of either possibility. Tinman has also been shown to act as either a transcriptional activator or repressor through physical interaction with either the p300 coactivator or the Groucho corepressor (Choi et al., 1999). Furthermore, several studies have shown that various combinations of transcription factors with

Tinman (such as Pannier, Serpent, Dorsocross, dTCF and pMad) are necessary for the complete activation of target genes but so far no further evidence for direct physical interactions between Tinman and these transcription factors has been presented (Han & Olson, 2005; Zaffran & Frasch, 2005; Junion et al., 2012).

5.5 Point-mutation of the potential Tinman binding site

Another very informative approach to analysing the importance of a specific sequence is mutating a specific region in the context of the complete enhancer and then placing this before a reporter gene. This approach is widely used in the study of transcription factors (for example: Akasaka et al., 2006; Gajewski et al., 2001; Ryu et al., 2011, Wang et al., 2005). The hypothesis behind this approach is that a transcription factor requires a specific DNA sequence to bind to the enhancer fragment and that by mutating the putative binding site the ability of a regulating transcription factor to bind to the DNA is prevented and thus the expression pattern of the reporter gene is altered.

In order to test if the S5 and S6 Tinman binding sites are of importance *in vivo*, I mutated both sites. The S6 site was mutated from CACTTGAAGTG to ***GT***CT***AC***AAGTG (termed *S6) or ***GT***CT***A***G***T***AGTG (termed *S6new) and the S5 site from TCAAGTG to T***GT***AG***AC*** (termed *S5). The *S6 and *S5 mutations are identical to the mutations used to test the binding specificity in the EMSA, and have been previously used and described as capable of disturbing Tinman binding by Gajewski et al., 2001.

The sequences created by the introduction of the mutations were screened with the same tests used in the bioinformatic analysis of the PCE sequence (described in Chapters two and four) to ensure that no known "new" binding sites for transcription factors were introduced by creating the mutations. The *S6 mutation introduces an E-box (*GT*CT*AC***AAGTG**) into the original sequence (CACTT**GAAGTG**). For this reason, I created the additional *S6new mutation. Instead of the G to C mutation, the *S6new mutation changes the following A to a T (*GT*CT*A***G*T**AGTG**). According to the Tinman position weight matrices (see Figure 5.5.1) available (Zinzen et al., 2009) this mutation should also interrupt binding of the Tinman protein. For all following constructs, I tested several independent insertion lines for their expression pattern and

show the most representative here (see Figure 5.3.1 for details on constructs and tested lines).

Figure 5.5.2 shows the results of the *S5 point mutation. All transgenic fly lines were stained at the same time and the same reagents were used to allow for comparability. As can be seen in Figure 5.5.2 (arrows), all expression of the reporter construct is lost in the heart. This shows that the S5 site is indeed crucial for regulation *in vivo,* and it also implies that the introduced mutation interferes with its ability to bind its regulating factor.

Figure 5.5.3 shows the results of the *S6 and the *S6new point mutations. In each case the PCE control and the mutation *(*S6 or *S6new) were stained at the same time and with the same reagents. The *S6 mutation does not completely abolish reporter activity. Figure 5.5.3 (C and E) shows that reporter activity is reduced. There are several possible explanations for this reduction. It is possible that the S6 site is only responsible for modulating the strength of reporter gene expression and thus mutation of it does not completely abolish the reporter activity. It is also possible that the observed reporter activity is due to the E-box introduced into the sequence through the mutation. This E-box has the same sequence (CAAGTG) as the E-box contained within un-mutated S5 binding site. The importance of this E-box sequence for reporter activation cannot be ruled out, as the mutation introduced into the S5 site also affects this E-box and the S5 mutation abolishes expression. It is also possible that the introduced mutation created a currently unidentified binding site that in combination with its transcription factor is capable of mimicking wild-type reporter expression at a reduced intensity.

To my knowledge, the *S6new mutation does not introduce any new binding sites into the PCE sequence while still affecting the S6 binding site; the *in vitro* analysis of the *S6new mutation has yet to be done. Reporter activity for the *S6new mutation is slightly reduced from the PCE control but not as weak as the *S6 mutation (see Figure 5.5.3 G and I). Again, this could be due to the fact that the function of the S6 site is only in modulating the reporter activity and not an "on/off" regulatory switch. It is also possible that neither of the S6 mutations, *S6 and *S6new, affect a binding site involved in the activation of the reporter gene but introduce a new currently unknown binding site into the PCE sequence. In all tested lines, expression in the *S6 mutation appears to be weaker than for the *S6new mutation (compare 5.5.3 C with G and E with the

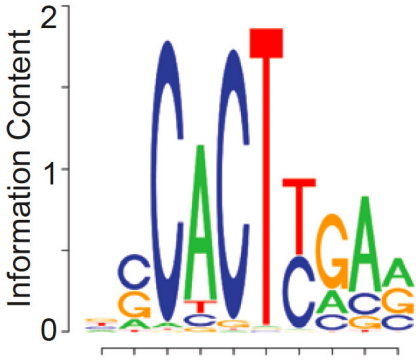


Figure 5.5.1 Tinman Position Weight Matrix (taken from Zinzen et al., 2009)
The Position Weight Matrix (PWM) for the Tinman binding displays the most likely base found at each position of the consensus binding sequence.

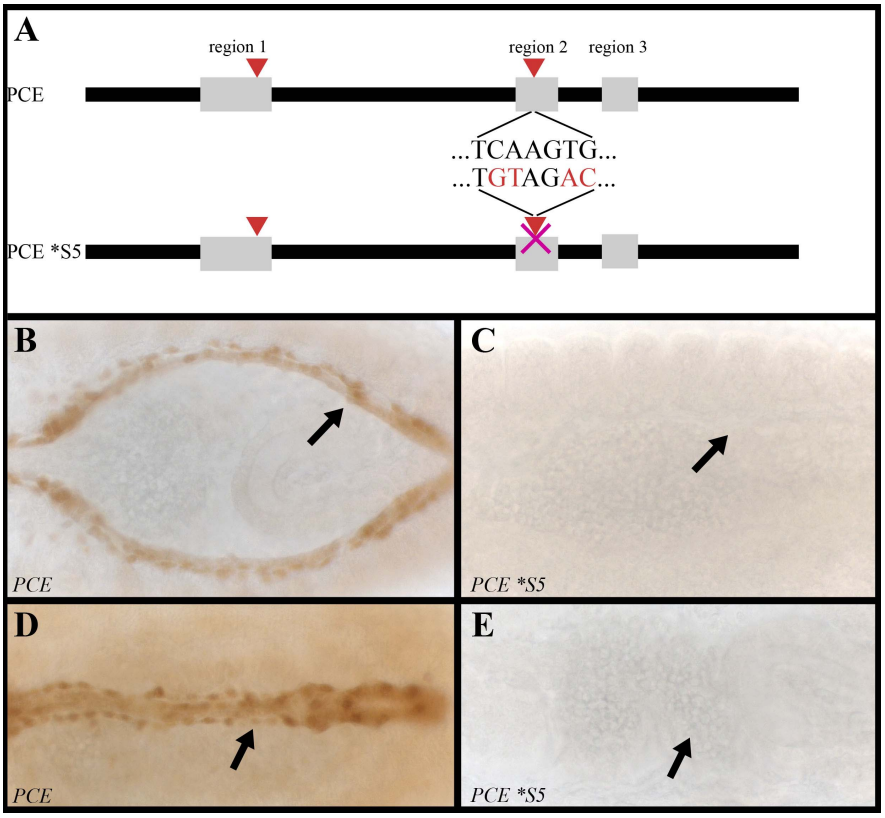


Figure 5.5.2: Mutation of the putative Tinman binding site S5 abolishes reporter gene expression.
A Schematic representation of the constructs used in this Figure; the grey rectangles represent the conserved regions and the red triangles indicate the locations of the putative Tinman binding sites (refer to chapter 4). The mutated bases are given in red. **B, D** LacZ reporter gene expression in the heart (arrows) for the full-length wild-type PCE construct at different stages of embryonic development B: stage 14, D: stage 17. **C, E** LacZ reporter gene expression is lost when the potential S5 Tinman binding site is mutated; the arrows indicate the position of the heart. C: stage 14, E: stage 17.

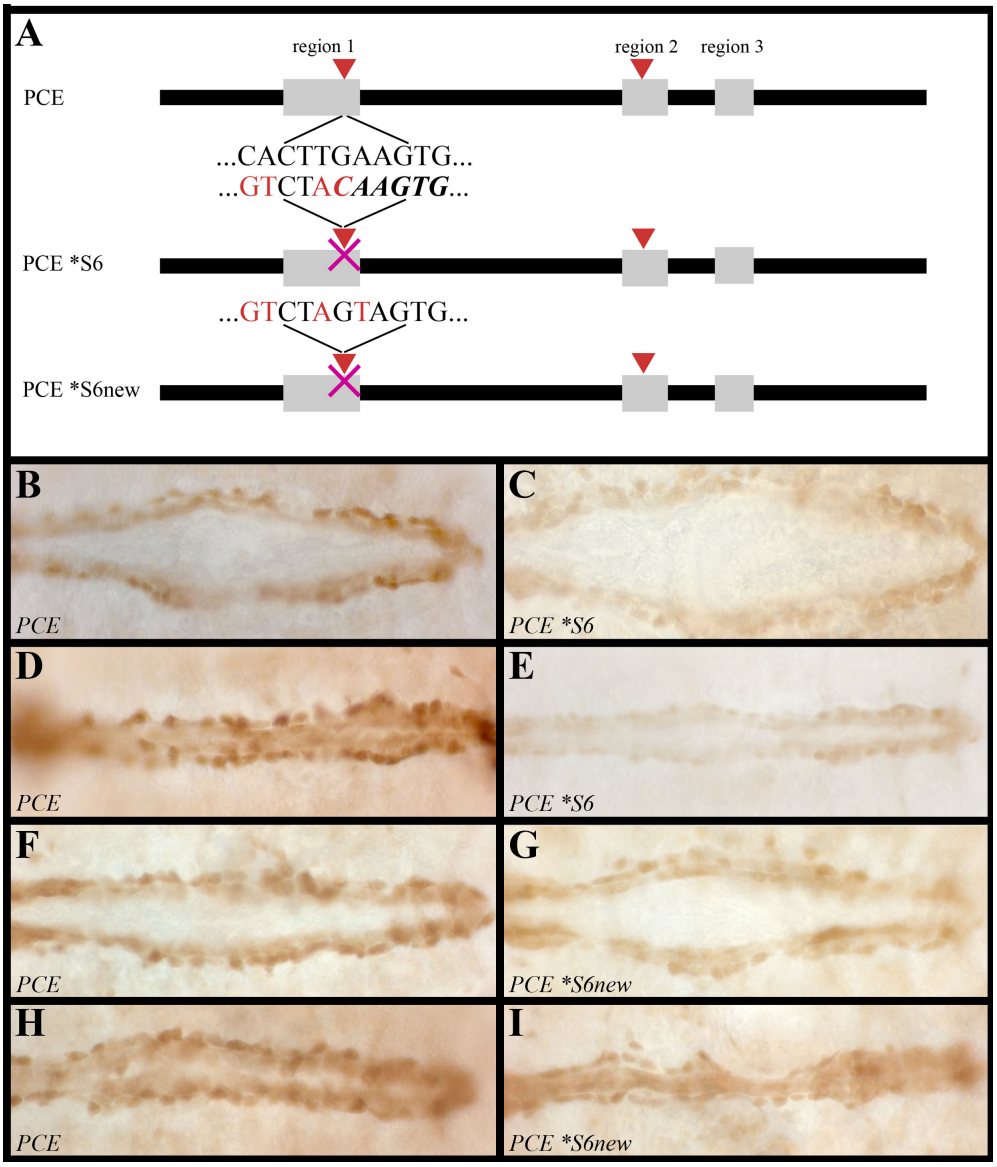


Figure 5.5.3 Mutation of the putative Tinman binding site S6 does not abolish reporter gene expression.

A Schematic representation of the constructs used in this Figure; the grey rectangles represent the conserved regions and the red triangles indicate the locations of the putative Tinman binding sites (please also refer to chapter 4). The mutated bases are given in red. **B, D, F, H** LacZ reporter gene expression in the heart for the full-length wild-type PCE construct at different stages of embryonic development B, F stage 15, D, H stage 17. **C, E** Reporter gene expression is greatly reduced but not lost when the *S6 mutation is introduced into the PCE sequence. C stage 15, E stage 17. **G, I** Introduction of the *S6new mutation into the PCE sequence reduces reporter gene expression to a lesser degree than the *S6 mutation. G stage 15, I stage 17.

I). This implies that either the *S6new mutation is not sufficient to prevent the binding of transcription factor or that indeed the E-box that was introduced with the *S6 mutation has an affect on this site.

Figure 5.5.4 shows the results for the combined *S5/*S6 double mutation. As can be seen in 5.5.4 (C) and (E) the double mutation does not eliminate reporter gene expression as the single **S5* mutation does. In fact, the expression pattern for the *S5/*S6 double mutant closely resembles the expression pattern I observed for the single *S6 mutation (compare Figures 5.5.3 C and E with 5.5.4 C and E). As each generated construct was sequenced and the introduced mutations verified, I can rule out that the *S5 mutation is not present in this enhancer construct. However, as the *S6 mutation introduced in this double mutant is identical to the old *S6 mutation (and thus introduces the additional Ebox), it is highly likely that the observed expression pattern for the *S5/*S6 reporter construct is due to this *S6 mutation.

The *S6new mutation only alters the reporter gene expression a little bit in intensity while the *S6 mutation weakens reporter expression considerably. There are several possible explanations for this: the *S6new mutation only changes three of the four bases of the E-box contained within the Tinman binding site consensus sequence. As the **S6new* mutation is a previously "untested" and unpublished mutation, it is possible that it retains or introduces more functionality than the original *S6 mutation and thus allows for stronger reporter gene expression. I believe that neither S6 mutation affects which pericardial cells express the reporter gene, but to validate this results double-labelling experiments with pericardial cell markers such as Tinman, Even-skipped and Odd-skipped are still necessary. From these results, I conclude that the function of the S6 site is either in modulating the level of expression only or that neither of the two mutations introduced are involved in the regulation of the PCE. For a complete understanding of the importance of the S6 Tinman site, the analysis of the *S5/*S6new double mutant will be necessary.

The S6 site could also be a redundant duplication of the S5 site. In their 2003 study, Dermitzaki et al. demonstrate "that there are sequences within the regulatory regions that are close enough to becoming a functional site that a few substitutions can make them potentially functional". This might also be the case for the S6 site, especially if the fact of two overlapping binding sites on the two different DNA strands is taken into consideration;

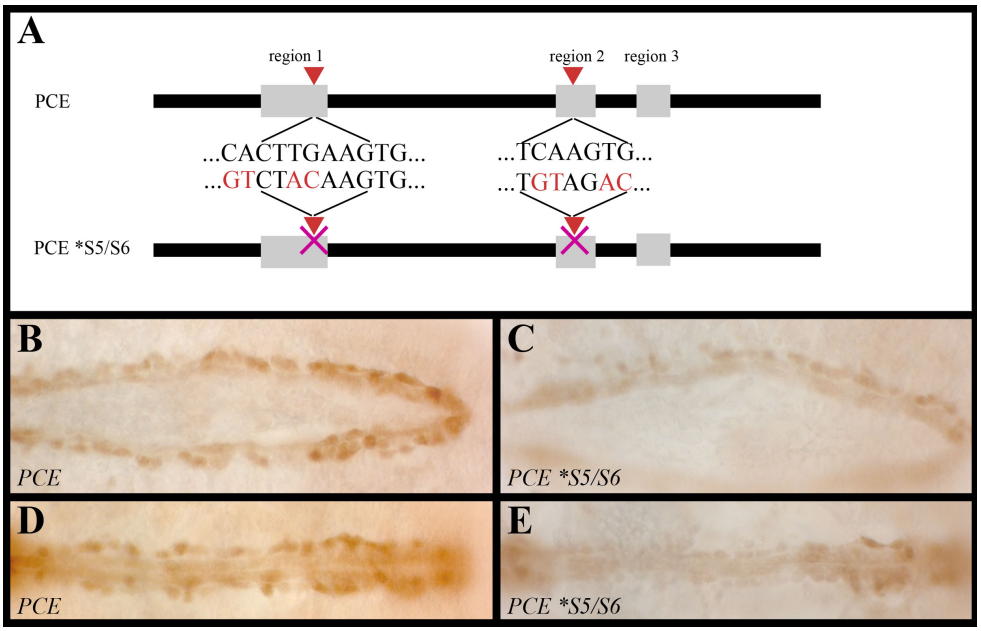


Figure 5.5.4: Effect of the *S5/*S6 double mutation on reporter gene expression.

A Schematic representation of the constructs used in this Figure; the grey rectangles represent the conserved regions and the red triangles indicate the locations of the putative Tinman binding sites (please also refer to chapter 4). The mutated bases are given in red. **B, D** LacZ reporter gene expression in the heart for the full-length wild-type PCE construct at different stages of embryonic development B stage 15, D stage 17. **C, E** Reporter gene expression is greatly reduced but not lost when the *S5/*S6 double mutation is introduced into the PCE sequence. C stage 15, E stage 17.

so far nobody has described a functional Tinman binding site arranged in this way. The S5 site is, however, essential for reporter gene expression as mutation of this site abolishes all reporter activity in the heart and acts as an "on/off" switch. The binding of Tinman to the S5 site activates gene expression in the pericardial cells. This gene expression is lacking if the S5 site is mutated in a way known to interfere with Tinman binding. These data again point towards Tinman as a regulator of *Him* expression.

5.6 Tinman over-expression induces ectopic *Him* expression

The Gal4/UAS-system allows one to express genes at times and in places where they are normally not expressed. If this "mis-expressed" protein is capable of inducing the expression of other proteins normally not expressed at this new location this is a good indication that these newly expressed proteins are direct or indirect targets of the "mis-expressed" protein. In addition, the expression domain of any protein induced by this "mis-expressed" factor is a good indication for further regulatory factors that might be involved in its regulation.

The above-described method is a commonly used approach to demonstrate that a specific protein is upstream of the gene of interest in the regulatory cascade. It follows that if Tinman is capable of activating *Him* transcription, then "mis-expression" of *tinman* should also induce *Him* expression in the tissues that *tinman* expression is driven in and *Him* is not normally expressed in.

I have used the *Da-Gal4*, *Mef2-Gal4* and *En-Gal4* drivers for this experiment. *Da-Gal4* induces uniform expression throughout the embryo (Wodarz et al., 1995), the *Mef2-Gal4* driver is specific for the somatic mesoderm from stage 7 until the end of embryogenesis (Ranganayakulu et al., 1996) and the *En-Gal4* driver drives expression in a specific ectodermal stripe (Gaumer et al., 2000; Sandmann et al., 2006).

In all three cases, the over-expression of *tinman* in these tissues causes ectopic expression of the *Him* transcript. The *Da-Gal4* driver induces ectopic *Him* expression in the somatic mesoderm (arrows in Figure 5.6.1), especially in the ventral areas, similar to the described expression pattern of the wild-type *Him* transcript in stage 12 (see chapter 3) but this expression is maintained in later stages. In addition to the mesodermal

expression, the *Da-Gal4* driver also induces ectopic *Him* expression in a few ectodermal areas overlying the somatic mesoderm and the amnioserosa (see arrowheads in Figure 5.6.1 A). The *Mef2-Gal4* driver has a more specific expression pattern that is limited to the somatic mesoderm than the ubiquitous *Da-Gal4* driver. Thus, ectopic *Him* expression in these embryos is limited to the muscles and within this expression domain, ectopic *Him* expression can mainly be found in the middle and ventral somatic muscles towards the posterior of the embryo. The *Da-Gal4* driver is not only expressed in the mesoderm. When *tinman* expression is driven in the pattern of the *Da-Gal4* driver, ectopic *Him* expression can additionally be observed in the ectoderm (arrow in Figure 5.6.1 C). Again more cells ectopically expressing *Him* can be found towards the ventral part of the somatic musculature. Ectopic *Him* expression is also often stronger in the posterior three para-segments of the embryo, indicating that this region contains regulatory factors that facilitate *Him* expression. The Hox genes *Ubx* and *abdA* involved in the establishment of the anterior-posterior axis are good candidates for this function along the anterior-posterior axis (Lo et al., 2002; Lovato et al., 2002; Ponzielli et al., 2002; Perrin et al., 2004, Ryan et al., 2005). The *En-Gal4* driver drives expression in very distinct ectodermal stripes (see Figure 5.6.1 E). However, *Him* expression is only induced in a sub-set of cells within this stripe (compare Figure 5.6.1 E and D). Ectopic *Him* is expressed in a one cell wide stripe immediately on the boundary of each para-segment. This corresponds to a one-cell-wide line of the two-cell-wide *Engrailed* expression domain. Determination of the exact anterior-posterior position of the ectopic *Him* expression within the *Engrailed* expression domain will require a double-labelling experiment. It has been shown that both one-cell-wide stripes within the *Engrailed* domain express *Mothers against Dpp* (*Mad*, Einers et al., 2009) but that the anterior of the two regions within the *Engrailed* domain is subject to the Wingless signalling pathway, while the posterior region is subject to signals from the EGF receptor (DER; O'Keefe et al., 1997, Dilks and DiNardo, 2010). *Him* expression is also mostly induced in the middle of the dorsal-ventral axis of each segment with less pronounced ectopic *Him* expression in the dorsal and ventral areas of the embryo. This area corresponds to the areas of *u-shaped*, *dorsocross* and high (but not highest) *decapentaplegic* expression (Fossett et al., 2000; Fossett et al., 2001; Hamaguchi et al., 2004; Reeves and Stathopoulos, 2009).

Notably, with all three used drivers the ectopic expression domain of *Him* is smaller than the expression domain of *tinman*. This implies that *Him* expression needs

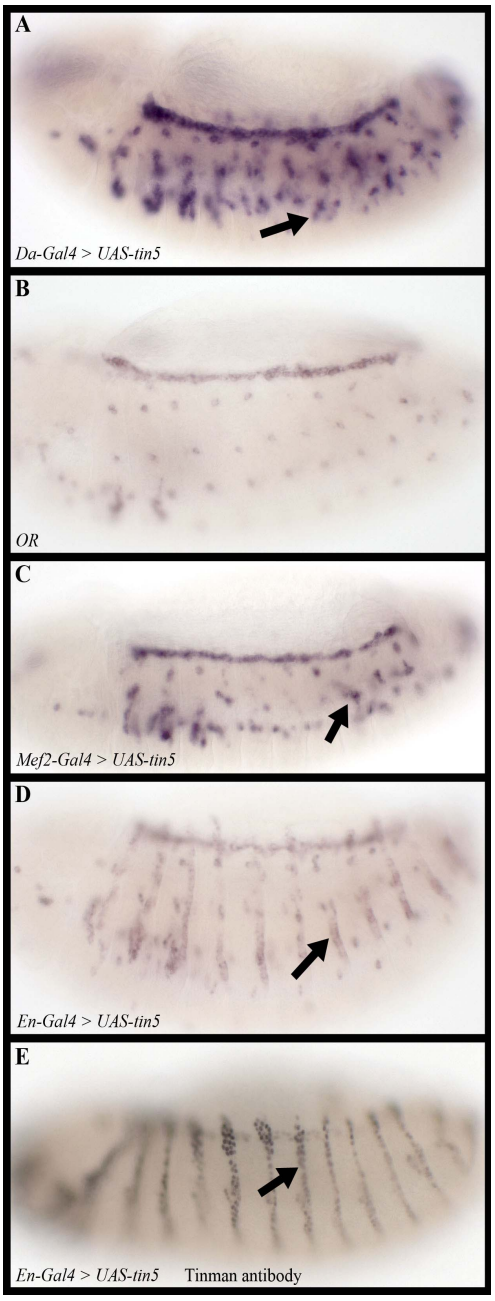


Figure 5.6.1: Ectopic Tinman expression induces ectopic *Him* expression.
Him RNA *in situ* (except where noted) on embryos ectopically expressing Tinman under the control of the Gal4/UAS system.
A Tinman expression construct under the control of the *Da-Gal4* driver induces ectopic *Him* transcript expression in the somatic mesoderm (arrow) and some ectodermal tissues. **B** *Him* RNA *in situ* on wild-type embryos (OR). **C** Driving Tinman expression with the *Mef2-Gal4* construct induces ectopic *Him* expression in the somatic mesoderm (arrow). **D** Tinman expression under the control of the *En-Gal4* driver induces ectopic *Him* expression in distinct ectodermal stripes (arrow). **E** Visualisation of the ectopic ectodermal expression (arrow) of the Tinman protein when driven by the *En-Gal4* driver.

further input by other transcription factors or that *Him* expression is repressed in the other cells over-expressing *tinman*. Based on these over-expression experiments the most likely candidate for *Him* regulators is Decapentaplegic, as it is expressed in the dorsal and lateral regions of the embryo and known to be involved in heart development. It is further possible that the *dorsocross* genes, GATA-factors (in combination with *u-shaped*) and Hox genes such as *abdA* are also involved in *Him* regulation.

In order to show that Tinman is activating reporter transcription through the PCE and not some other regulatory region, I have repeated the tinman over-expression experiments in the presence of the wild-type PCE reporter using the *Da-Gal4* and the *Mef2-Gal4* drivers. The expression pattern of the PCE reporter protein shows that *tinman* over-expression indeed induces ectopic reporter expression. Over-expression of *tinman* by the *Mef2-Gal4* driver causes ectopic PCE expression in the developing somatic muscles, mainly in the dorsal and lateral muscles. However, the over-expression of *tinman* disturbs the wild-type muscle pattern so that I am unable to assign in which specific muscles the PCE reporter is induced in (Figure 5.6.2 F). The ectopic expression of the PCE is more pronounced in the dorsal region of each segment, once again implicating Decapentaplegic as a potential further regulator, but it can also be seen in single cells of the mesoderm (see Figure 5.6.2 B). The ectopic PCE expression caused by using the *Da-Gal4* driver shown in Figure 5.6.2 can again be found in single cells of the somatic mesoderm, some ectodermal cells, in the cells of the gut constrictions and in cells along the ventral midline. I have also observed ectopic PCE expression in the cardioblasts with this driver. I did not observe this if the *Mef2-Gal4* driver was used. This difference in *Him* expression within the same tissue can be caused by the temporal differences in the driver activity and by the different "strength" of the two drivers.

The PCE expression induced by over-expression of *tinman* is more widespread than the observed ectopic expression of the *Him* transcript while using the same drivers (compare Figure 5.6.2 B and F with Figure 5.6.2 A and E). This is likely due to the fact that the PCE sequence has been removed from its context, which might contain repressor binding sites, chromosomal structures and other sequences of regulatory importance that limit the expression of the intrinsic *Him* gene. The context of the PCE enhancer appears to be of little or no importance in a wild-type environment as the PCE

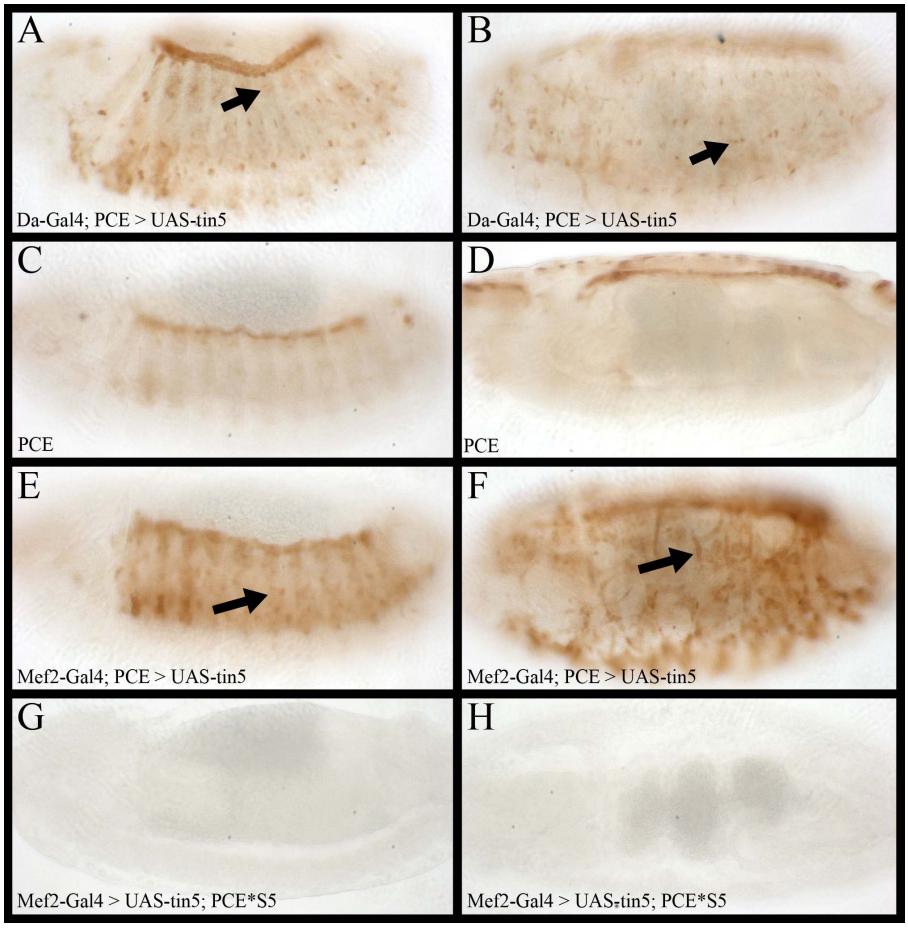


Figure 5.6.2 Tinman over-expression triggers ectopic *Him* reporter gene expression.

Visualisation of the LacZ reporter gene expression in embryos expressing ectopic Tinman under the control of the UAS/Gal4 system. A, C, E, G stage 13 and B, D, F, H: stage 15/16

A, B Tinman expression driven by the *Da-Gal4* driver induces ectopic PCE reporter gene expression in the somatic mesoderm (arrow) and the cardioblasts. **C, D** PCE reporter gene expression in a wild-type background. **E, F** When controlled by *Mef2-Gal4*, ectopic Tinman is able to induce ectopic PCE reporter gene expression in the developing somatic mesoderm (arrows). **G, H** *Mef2-Gal4* driven ectopic expression of Tinman is not able to induce expression of the PCE carrying the *S5 mutation.

region on its own is able of reproducing the expression pattern of the *Him* transcript without additional expression in other tissues. These over-expression results point to the presence of such sites that are capable of limiting *Him* expression in undesired mesodermal and ectodermal tissues if at least one activating transcription factor is present in a wider than usual expression domain.

To further establish that it is indeed the S5 site within the PCE through which Tinman acts, I have over-expressed *tinman* in the presence of the PCE*S5 reporter construct. The S5 site in this construct is mutated and I have previously shown that this mutation abolishes all reporter activity in a wild-type background (Figure 5.6.2 G and H). The staining for both the PCE and PCE*S5 experiments were done in parallel to have a positive staining control and allow for comparability between the two. Figure 5.6.2 G and H show that over-expression of *tinman* cannot compensate for the lack of the S5 binding site in any tissues in these embryos. This implies that the S6 site cannot compensate for a non-functional S5 binding site. It is however difficult to rule out any role of the S6 site in the presence of a functional S5 binding site.

Both of these experiments taken together show that the PCE is responsive to Tinman and, furthermore, that the S5 site is necessary for the transcriptional activation of the PCE through Tinman. Together with the previous results showing that the endogenous *Him* gene is also induced by over-expressing *tinman*, it is highly likely that this response of *Him* transcription to *tinman* expression occurs through the S5 site in a wild-type *in vivo* context as well. This proves that the S5 Tinman binding site is indeed of functional relevance while the S6 site does not appear to be necessary for *Him* regulation *in vivo*.

5.7 The effect of expressing a dominant-negative form of Tinman on *Him* expression

As previously mentioned, *tinman* null mutants do not form a heart (Bodmer, 1993). This complicates the analysis of any potential Tinman target gene in embryos mutant for *tinman*. The targeted expression of a dominant-negative form of Tinman

using the UAS/Gal4-system allows for a conditional loss-of-function in specific cells without detrimental effects on any other developing tissues or cells of the animal.

By expressing dominant-negative Tinman in the heart cells after the early specification for the cardiac field has taken place, it is possible to impair Tinman function while retaining early heart development until the dominant-negative form of Tinman becomes active (Han et al., 2002; Han et al., 2005). This will allow me to see from which stage onwards Tinman is necessary for *Him* expression. If Tinman is necessary only at the very early stages of heart development, *Him* expression will be unaffected in the embryos expressing dominant-negative Tinman. If, however, Tinman is also necessary to maintain *Him* expression during the middle and late stages of heart development, I should then be able to observe a decrease and/or lack of *Him* transcription.

The transgenic fly line carrying the UAS-dominant-negative-Tinman (UAS-DN-Tin) construct was generated in the Bodmer lab (Han et al., 2002) by fusing the Engrailed repressor domain to the Tinman homeodomain (amino acid 291-370) according to the strategy described by Fu et al. (1998). When driven with the eme900-Gal4 line, which drives expression in the mesodermal Eve pattern, Han et al. (2002) describe a "significant but variable reduction of Eve expression" but do not comment on whether this loss of Eve expression is due to cell loss or not.

I have used the *hand-Gal4, TinCΔ4-Gal4* and *Him-Gal4 L3-5* drivers to express the dominant-negative form of Tinman. The *hand-Gal4* line drives expression in all heart cells from late stage 12 onwards (Sellin et al., 2009; D. Popichenko & A. Paululat, unpublished data, see Appendix). The *TinCΔ4-Gal4* driver is specific for the four ß3-tubulin-positive cardioblasts of the heart from stage 13 onwards (Bodmer lab, K. Jagla pers. communication, see Appendix). The *Him-Gal4* line drives expression in the heart precursors from stage 11/12 onwards and then in all pericardial cells of the developing heart (see Appendix).

When the dominant-negative form of Tinman is driven with the *Him-* or *hand-Gal4* lines the *Him* transcript is still present at reduced levels as it still can be visualized in the *Drosophila* embryo (see Figure 5.7.1). If the dominant-negative form of Tinman is driven with the *TinCΔ4-Gal4* line, the *Him* transcript levels are not reduced. While some of the variation between embryos in each experiment has to be attributed to the

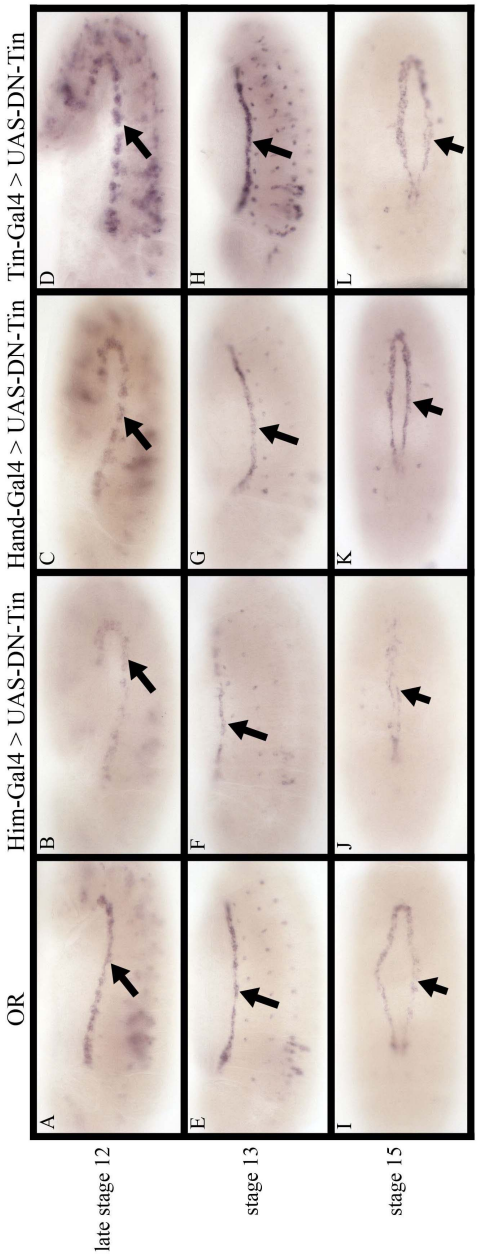


Figure 5.7.1: Expression of a dominant-negative form of Tinman in the mesoderm reduces, but does not abolish, *Him* expression.
Him RNA *in situ* on developing embryos expressing dominant-negative Tinman under the control of three different mesodermal Gal4 drivers.
A, E, I *Him* RNA *in situ* on wild-type (OR*)* embryos, **B, F, J** *Him* expression is reduced when dominant-negative Tinman is under the control of *Him-Gal4*. **C, G, K** When *hand-Gal4* drives the expression of the dominant-negative form of Tinman, *Him* expression is slightly reduced. **D, H, L** Driving the dominant-negative form of Tinman with *TinCΔ4-Gal4* does not affect *Him* expression levels.
A, B, C, D late stage 12, E, F, G, H stage 13, I, J, K, L stage 15.

mechanisms of the UAS/Gal4 system and the most likely incomplete inhibition of Tinman activity, the result that *Him* transcription is not abolished remains constant with all Gal4 driver lines used in this study.

The *TinCΔ4-Gal4* driver, which has no effect on *Him* transcription, is specific to a subset of cardioblasts and starts expressing the Gal4 protein latest of all three drivers used, thus the dominant-negative form of Tinman and *Him* should not be co-expressed in this experiment which can thus serve as an additional control. The *Him-Gal4* driver, which has the strongest effect on the levels of *Him* transcript, is the earliest active driver in the group of driver lines I used. This indicates that while Tinman is necessary at early developmental stages (from this experiment up to stage 12 and see section 5.2) to activate *Him* transcription, at later stages Tinman is, at most, only one contributing factor towards maintaining *Him* transcription. The reason for this early requirement of Tinman might tie in with the fact that the early dorsal mesoderm contains the shared precursors of cardioblast and pericardial cells.

Many embryos showed a loss of heart cells reminiscent of the complete heart loss in the *tinman* null mutant when the dominant-negative form of Tinman is expressed, irrespective of which driver I used. This loss is least severe when the *TinCΔ4-Gal4* driver is used. In their 2002 study, Han et al. described reduced Even-skipped expression if an enhancer fragment specific for the Eve-positive pericardial cells and the dorsal muscle DA1 drives the dominant-negative form of Tinman. They have not described this phenotype any further but the likely conclusion from their experiment is that these cells are lost in a similar fashion as the heart cells in a *tinman* null mutant. These similar phenotypes that both Han et al. (2002) and I have observed, indicate that the dominant-negative Tinman is functioning. Because of the loss of heart cells, it is impossible to determine if the loss of Tinman function after the initial steps of heart development is specific to certain sub-sets of pericardial cells by the position of these cells alone.

In order to establish if the PCE behaves similarly to the endogenous *Him* transcript, I have repeated the previous experiment with PCE in the background using the *hand-Gal4* driver only. All staining shown in Figure 5.7.2 were done in parallel, allowing for comparability of the intensity of the reporter gene expression. As can be seen in Figure 5.7.2, if the levels of reporter gene expression are compared between the

control embryos (carrying the *hand-Gal4* driver with the PCE in the background) and those embryos expressing the dominant-negative form of Tinman, it is obvious that the latter embryos express less of the reporter gene. In these embryos reporter gene expression is reduced considerably from late stage 12 onwards which coincides with the beginning activity of the *hand-Gal4* driver. This shows that under these circumstances the PCE recapitulates the behaviour of the endogenous *Him* transcript. The reduction of the PCE expression appears to be stronger than the degree I observed for the endogenous *Him* transcript when the dominant-negative form of Tinman was expressed under the control of the *hand-Gal4* driver. However, to be able to truly compare the degree at which the endogenous *Him* transcript and the PCE LacZ reporter gene lose expression strength it will be necessary to compare the level of the *Him* transcript to that of the LacZ transcript and not the translated protein, while keeping in mind that both transcripts might also have differing stabilities.

To further illustrate this point, I have double-labelled the *PCE; hand-Gal4 > UAS-DN-Tin* embryos for both the endogenous *Him* transcript and the LacZ reporter gene. Figure 5.7.3 shows again that the expression for both, endogenous *Him* and the reporter gene, is reduced when compared to the control and that gaps of staining within the heart are present. As it is possible that these gaps in staining are caused by the complete loss of heart cells in this region and not by absence of the *Him* transcript or the expressed reporter gene, I have also counterstained these embryos for Zn finger homeodomain 1 (Zfh-1) to identify pericardial cells. Zfh-1 has, so far, not been shown to be a Tinman target and is thus suitable as a marker for pericardial cells in these conditions. At stage 13, Zfh-1 is expressed in all pericardial cells (Lai et al., 1991; Su et al., 1999; Johnson et al., 2003) and, as shown in the previous chapter, Zfh-1 and the *Him* transcript co-localize in the pericardial cells. Figure 5.7.3 shows that there are indeed Zfh-1 positive cells present that do not express either the *Him* transcript or the LacZ reporter (see arrows in Figure 5.7.3). This demonstrates that this system, if the dominant-negative form of Tinman is induced early enough, does allow for the formation of heart cells that lack *tinman* but express other established cardiac markers and that these non-Tinman expressing cells do not express *Him* or the PCE reporter gene. The loss of Tinman function at later stages of development (stage 12 onwards) thus leads to a general reduction of *Him* expression in the remaining heart cells and in few cases can also completely abolish *Him* expression. As this later phenotype was not

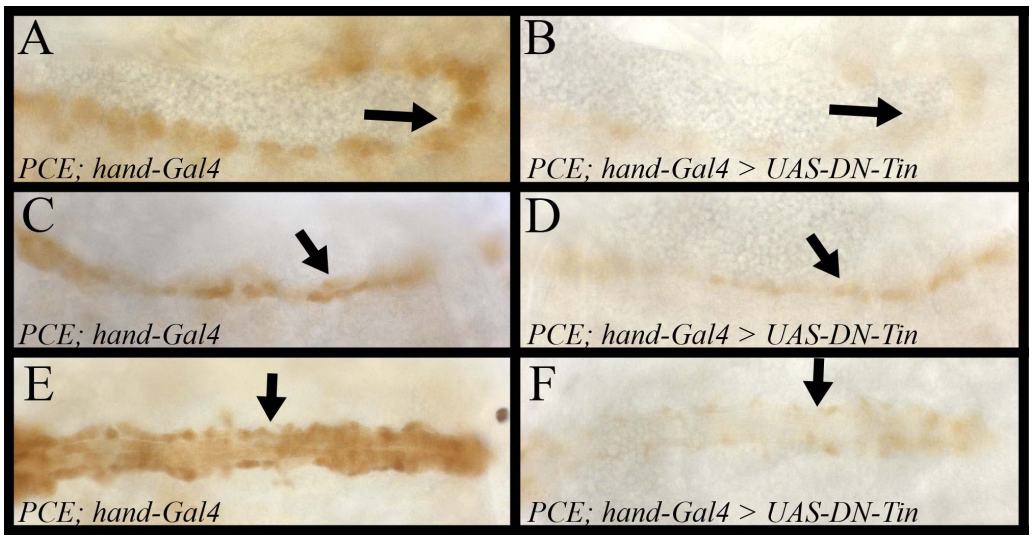


Figure 5.7.2: Expression of a dominant-negative form of Tinman reduces reporter gene expression.
Visualisation of the PCE LacZ reporter gene expression during embryonic development when dominant-negative Tinman is driven by the *hand-Gal4* construct. The arrows point to developing heart.
A, C, E "Normal" expression of the PCE reporter gene with the *hand-Gal4* driver in the background. **B, D, F** The intensity of the LacZ reporter gene is greatly reduced if the dominant-negative form of Tinman is expressed under the control of the *hand-Gal4* driver.
A, B stage 12, C, D stage 14, E, F stage 17.

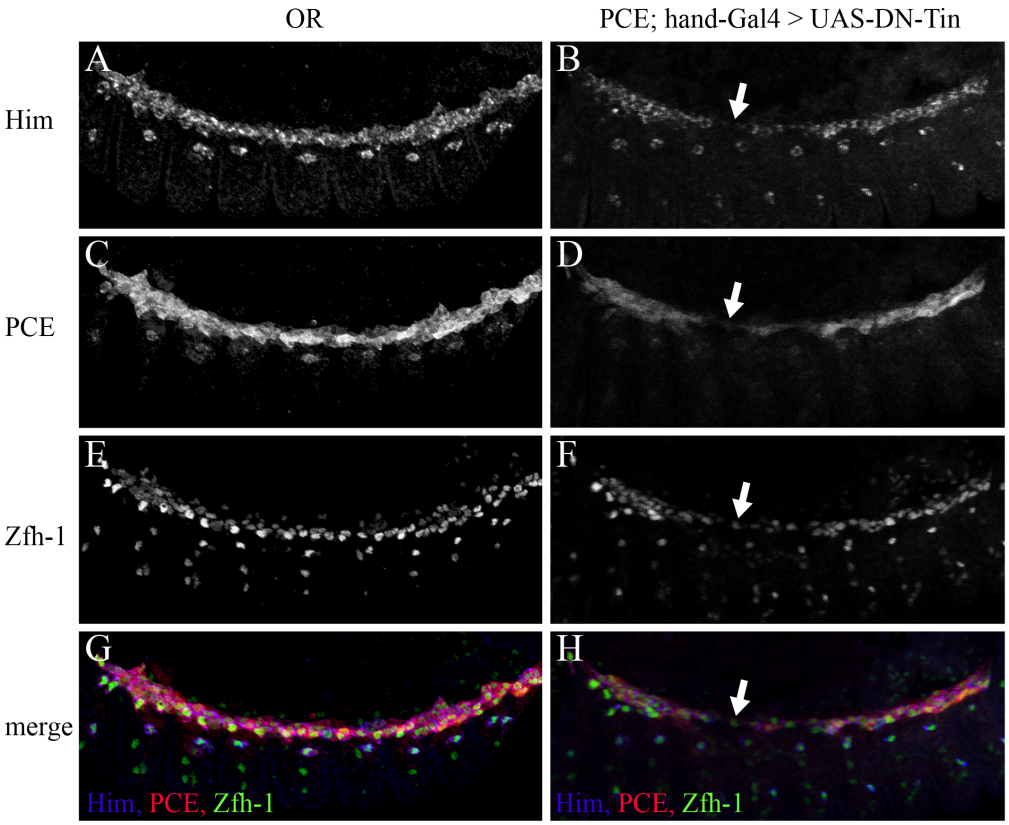


Figure 5.7.3: *Him* transcript, PCE reporter gene and Zfh-1 protein expression co-localise in embryos expressing the dominant-negative form of Tinman under the control of *hand-Gal4*.
Confocal microscopy images of embryos triple labelled for the *Him* RNA transcript, LacZ reporter gene and Zfh-1 expression. A, B: *Him* FISH, C,D: LacZ expression, E, F: Zfh-1 expression, G, H: merge of all three images, *Him* staining in blue, LacZ in red and Zfh-1 in green.
A, C, E, G Wild-type expression pattern of the *Him* transcript, PCE and Zfh-1 on a stage 13 OR embryo. **B, D, F, H** The expression of the *Him* transcript and PCE is reduced in embryos expressing the dominant-negative form of Tinman under the control of *hand-Gal4*. The arrow indicates a single Zfh-1 positive nucleus that does not express *Him* or the PCE reporter gene.

observed very often, it is possible that the time point at which Tinman function is essential for *Him* expression is not much earlier than the onset of driver activity which if a cell were to lag behind could generate a pericardial cell lacking *Him* expression. This result also shows that the expression of *zfh-1* is most likely independent of late Him presence.

5.8 *Him* expression in embryos that completely lose late Tinman function

The removal of Tinman function is most likely incomplete when the dominant-negative form of Tinman is expressed and this needs to be taken into consideration in the interpretation of the above results. Another way to analyse the importance of Tinman function for *Him* expression is the use of the *tinABD-1B2; tin^{346}/TM3 evelacZ* fly line generated in R. Bodmer and M. Frasch's labs (Zaffran et al., 2006). These flies carry a time-defined Tinman rescue element (*tinABD-1B2*) on the second chromosome within a *tinman* null mutant background over a "blue balancer" on the third chromosome (*tin^{346}/TM3 evelacZ*). These flies express *tinman* during the early stages of embryonic development but lose *tinman* expression during stage 11, allowing for a normal early specification and development of the mesoderm and the dorsal mesoderm that contains the heart progenitors until the loss of Tinman is triggered (Zaffran et al., 2006). The loss of Tinman at later developmental stages in this fly line is "complete" and does not depend on protein interactions like the dominant-negative form of Tinman. The loss of Tinman expression in these embryos also occurs, according to the literature, slightly earlier as in the embryos expressing the dominant-negative Tinman with the *hand-* (late stage 12 onwards) or *Him-Gal4* (stage 11/12 onwards) drivers.

As can be seen in Figure 5.8.1, while *Him* expression levels in the *tinABD-1B2; tin346/TM3 eveLacZ* embryos are slightly reduced, the *Him* transcript is present in these embryos. Similar to a *tinman* null mutant and to the dominant-negative experiment, heart cells are also disorganised or missing from stage 11/12 embryos onwards as described by Zaffran et al. (2006). While *Him* expression is affected to a low degree as can be seen by the reduced intensity of the staining in these embryos, neither this nor the dominant-negative experiment completely abolish *Him* expression in the developing

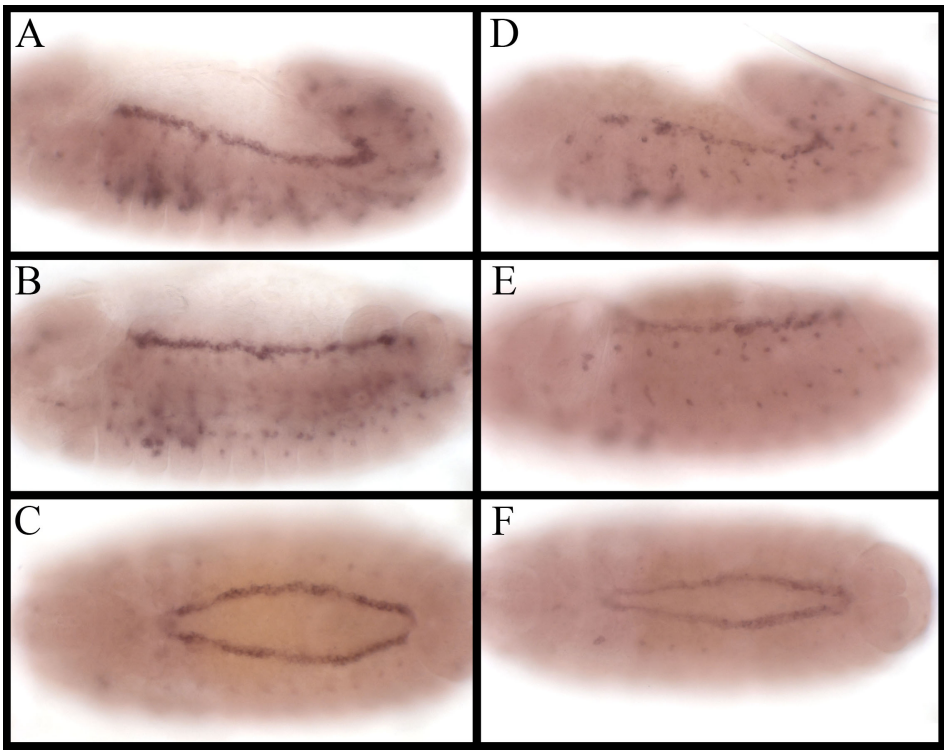


Figure 5.8.1: *Him* transcript expression levels are weakened if Tinman function is lost after stage 11.
Expression pattern of *Him* in wild-type and *tinABD-1B2; tin^{346}/TM3 eve LacZ* transgenic embryos.
A, B, C Wild-type expression pattern of the *Him* transcript. **D, E, F** Loss of *tinman* transcription from stage 11 (*tinABD-1B2; tin^{346}/TM3 eve LacZ*) onwards reduces *Him* transcript levels only slightly.
A, D stage 12, B, E: stage 13, **C, F** stage 15

embryo. This indicates that, while Tinman is absolutely necessary to activate *Him* expression during the early stages of heart development (stages 10, 11; see section 5.2), Tinman is only a contributing factor in the maintenance of *Him* expression during later stages of heart development. These results are also supported by the experiments detailed in section 5.7. There is however the possibility that some Tinman protein might still be present during the later stages of development even though its transcription has stopped as a protein is often more stable than its RNA transcript and there will also be a delay in the loss of Tinman function following the loss of *tinman* expression (Zaffran et al., 2006).

5.9 Discussion

Tinman is one of the earliest markers for the developing dorsal mesoderm and the *Drosophila* heart, which develops from this dorsal mesoderm. Here, I have shown that Tinman is a necessary activator of *Him* transcription in the early stages of heart development and that Tinman contributes towards the full level of *Him* expression during the later stages of heart development. In cases where the function of Tinman is reduced or abolished in these later stages (stage 11 or slightly later), the *Him* transcript is still expressed in the majority of the remaining pericardial cells but at reduced levels. Tinman function is, however, still necessary for the expression of *Him* in some pericardial cells as it is possible to detect the occasional single pericardial cells that does not express the *Him* transcript but does express the pericardial cell marker Zfh-1. In the future it might be interesting to screen for any differences in the susceptibility of the subsets of the pericardial cells to the loss of *Him* expression if Tinman function is interfered with. I would expect that the Tin- and Eve-positive pericardial cells are more likely to be affected in these experiments than the Odd-positive pericardial cells as *tinman* is not expressed in the Odd-pericardial cells after the initial specification of the dorsal mesoderm and heart precursors. However, any such experiments are complicated by the loss of heart cells (and any positional information derived from the full set of heart cells) that accompanies the loss of tinman and thus these studies will need careful co-labelling with markers for the individual sub-sets of pericardial cells. This is further complicated by the fact that some of the commonly used markers for these cells are dependent on Tinman function for their expression, e.g. Even-skipped and Seven-up

(Knirr and Frasch, 2001; Han, Z. et al., 2002; Jagla, 2002) and that Tinman itself is a marker for a sub-set of pericardial cells.

In this chapter I have also shown that the PCE is able to reproduce the expression patterns described for the endogenous *Him* transcript (see chapter 3) on its own and also in response to alterations in Tinman function. The PCE is thus not only able to faithfully reproduce the cardiac *Him* expression pattern in a wild-type background but also reacts faithfully to changes of at least one known *Him* regulator, Tinman. In addition, the PCE fragment is the only known way to label all embryonic pericardial cells until late embryonic stages, when the expression of *zfh-1* is lost from the Eve-positive pericardial cells. Thus, the PCE is a useful tool for the study of the complete pericardial cell population during their complete embryonic life.

Chapter 6

Further regulatory input in *Him* expression

6.1 Introduction

Tinman cannot be the only regulator of *Him* expression as shown by the expression pattern of the PCE construct. If this was the case *Him* expression should mirror *tinman* expression exactly. However, *tinman* is expressed in both, cardioblasts and pericardial cells and additionally, it is not expressed in all pericardial cells (Bodmer et al., 1990; Bodmer, 1993). Only a sub-set of six of the ten pericardial cells express *tinman*. The remaining four pericardial cells, the Odd-positive cells, lose *tinman* expression during stage 12 of the embryonic development (Ward and Coulter, 2000; Ward and Skeath, 2000). However, *Him* and PCE expression persists in these cells after the loss of *tinman* and *Him* and the PCE are not expressed in the cardioblasts (see introduction and chapter 3).

Usually, several different transcription factors confer the temporal and spatial specificity to one downstream target, so it is highly unlikely that Tinman is the only regulating input involved in *Him* expression. All this points to the necessity of at least one further activator for the remaining four pericardial cells that also express Odd-skipped and at least one repressor for the cardioblasts in which *tinman* is also active, but *Him* is not. It is also possible that *Him* expression in all pericardial cells is initiated by Tinman during the early stages of heart development, but then needs further cues to maintain its expression in at least the Odd-positive pericardial cells. The results from chapter 5 (sections 5.2, 5.7 and 5.8) show indeed that while the presence of Tinman is essential during the early stages of heart development, it is of lesser importance at later stages of heart development.

As none of the putative regulators found in the bio-informatic analysis were very strong candidates for this additional activator (see chapter 4) based on the conservation of the potential transcription factor binding sites, I decided to do a combination of a deletion analysis and a “Linker Scan” analysis of the PCE to identify further regulators. McKnight and Kingsberry (1982) and Mohun et al. (1989) describe the technique of “Linker Scanning mutation”. I based both of these approaches on the regions of

conservation I identified during my computational analysis of the PCE sequence (see chapter 4).

This approach is based on the assumption that conservation in transcribed and non-transcribed DNA regions occurs because of a functional constraint placed upon these sequences. This functional constraint could be the binding of a transcriptional regulator to the DNA in non-coding areas. By replacing stretches of conserved sequence with different, unrelated sequences, I aimed to uncover regions of regulatory importance, as the expression pattern of the reporter gene in the Linker Scan constructs should be altered if this is the case. To test this hypothesis I mutated sequences of 5, 10 or 11 bp lengths within the context of the complete PCE sequence and then analysed the resulting construct and its expression pattern *in vivo*.

The results I present here are a combination of the results of the Linker Scan analysis and a deletion analysis based on the conserved regions identified during the bio-informatic analysis.

6.2 *Him* expression is repressed in the cardioblasts.

As explained above, the fact that Tinman is an activator for *Him* expression generates a requirement for at least one repressing factor to exclude *Him* expression from the cardioblasts.

The first conserved region with the PCE sequence is relatively close to the 5' end of the wild-type PCE construct (see chapter 4 and also Figure 6.2.1). The PCE S4 is a deletion construct that removes the most 5' 67 bp of PCE sequence but retains the first conserved region and some of the non-conserved 5' sequence. Both reporter constructs, PCE and PCE S4, have the same expression pattern (see Figure 6.2.1). However, if the upstream 5' deletion is extended up to the first conserved region (the PCE reg1 reporter construct), the construct shows ectopic reporter gene expression in all cardioblasts (see Figure 6.2.1 C). The observed ectopic expression in the PCE reg1 construct is at a slightly reduced level when compared to the wild-type PCE construct but above any background levels. This PCE reg1 construct was made based on the computational analysis of the PCE enhancer (chapter 4) and starts exactly at the beginning of the first

conserved region. Based upon the detection of the ectopic expression in the cardioblasts I investigated the area around this deletion and upstream of it further.

The logical conclusion from these results is that there is a binding site that is located 5' to the region where the first conserved region begins. Assuming that the degree of conservation I observed for this area has a functional importance and in order to rule out that I inadvertently disturbed a binding site at the very beginning of the PCE reg1 deletion construct, I created a further deletion construct. This extends the PCE reg1 construct by 6 bp in its 5' direction and is called PCE extreg1. If this construct is not expressed in the cardioblasts, the area responsible for preventing PCE expression in the cardioblasts is located at the very beginning of the first conserved region. Interestingly, the PCE extreg1 construct also induces ectopic reporter gene expression in the cardioblasts at similar levels observed for the PCE reg1 construct (see Figure 6.2.1). This indicates that there is a repressor binding site upstream of the conserved region1 in the non-conserved sequence up to the start of the PCE S4 construct.

Based on these results, I re-evaluated the results of my bioinformatic analysis of the 39 bp between the PCE extreg1 and the PCE S4 constructs again for the presence of known sites for transcription factors that are expressed in the cardioblast. Even with less stringent criteria (accepting more mismatches to the consensus sequences and ignoring the degree of conservation between the *Drosophila* species) the region deleted between the PCE reg1 and PCE S4 constructs contained no promising candidates for regulatory factors. Interestingly, this stretch of 39 bp does contain a stretch of 14 consecutive Adenines. Consecutive A-stretches in DNA are known to alter the 3-dimensional structure of the DNA and introduce bends (Koo et al., 1986). These consecutive Adenine stretches have been shown to have an influence on the regulation of transcription (Lavigne et al., 1997; Olivarez-Zavaleta et al., 2006). Adenine-Thymine base pairing also only requires two hydrogen bonds, making it easier to separate the two complimentary DNA strands from each other. This is likely to have an influence on the surrounding sequences and possibly might have a regulatory significance. For a comprehensive review of the role of Adenine tracts see Haran and Mohanty, 2009.

Within this region I also created two Linker Scan constructs (called PCE LSA and PCE LSA1-5) that target the start of the first conserved region of the PCE sequence. Both constructs cover the same region and alter the sequence at the very beginning of the first conserved region. The PCE LSA construct mutates 9 bp from tattcccaagc to GCtCTAGaGCA and the PCE LSA1-5 construct mutates the first 5 bp from tattc to

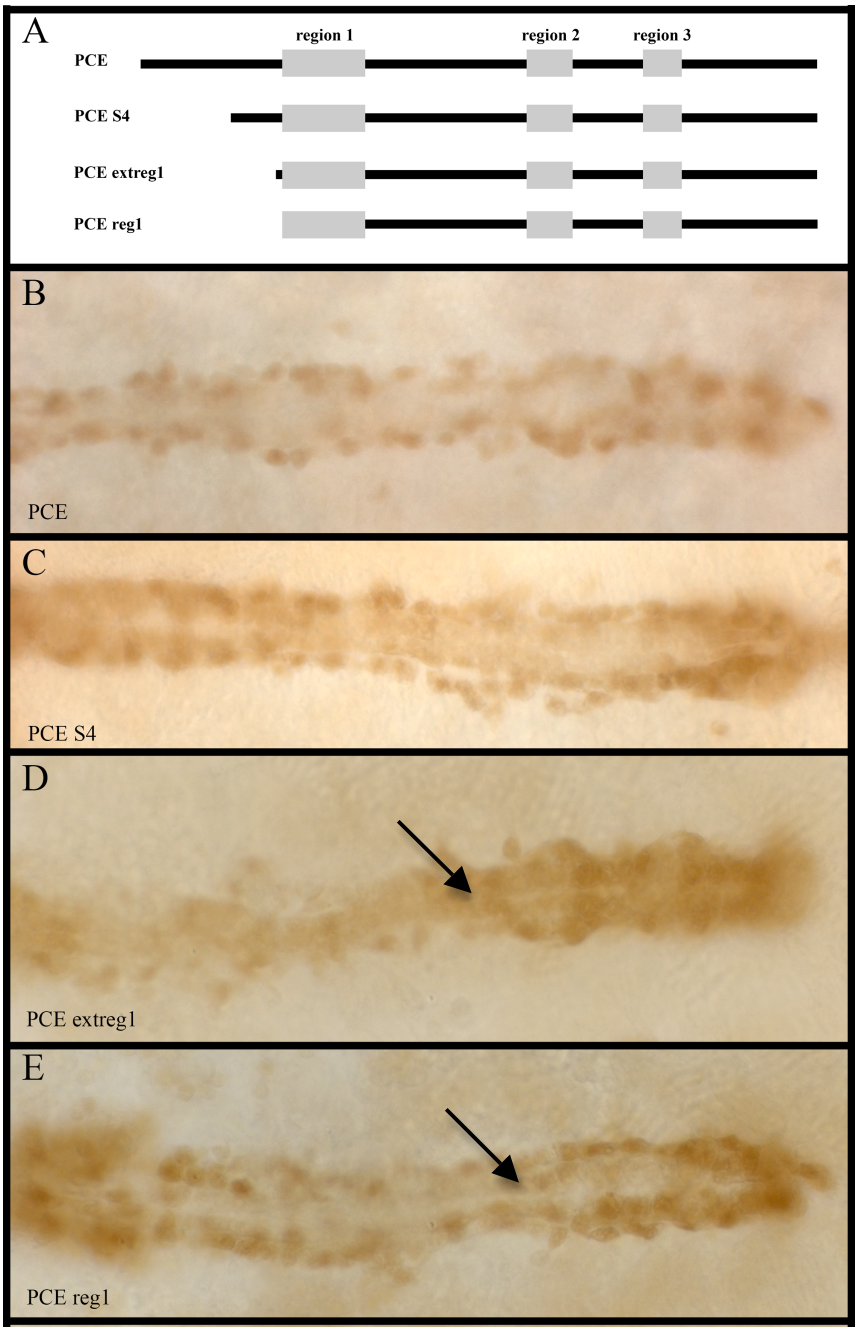


Figure 6.2.1: The deletion constructs PCE reg1 and PCE extreg1 show ectopic expression in the cardioblasts.
A Schematic representation of the reporter constructs used in this Figure. **B** Wild-type expression pattern of the full length PCE reporter construct in the heart. **C** The expression pattern of the PCE S4 reporter construct is not distinguishable from that of the full length construct. **D** The PCE extreg1 deletion causes ectopic expression in all the cardioblasts (arrow). **E** The PCE reg1 construct also shows ectopic expression in the cardioblasts (arrow), this deletion is 5 bp shorter than the PCE extreg1 deletion.

GCGGA. Each mutated sequence was screened with the same tests that were used for the computational analyses in Chapter 4 to rule out the introduction of any already known binding sites by the introduced mutation. As can be seen in Figure 6.2.2 both constructs introduce strong ectopic expression in the cardioblasts, similar to what I observed for the PCE reg1 and PCE extreg1 constructs. However, different to what I observed in those constructs the ectopic expression in the cardioblasts in the PCE LSA and PCE LSA1-5 constructs appears very strong and robust. There are several possible explanations for the difference in reporter gene expression strength between the PCE LSA, PCE LSA1-5, PCE reg1 and PCE extreg1 constructs. All fly lines were generated through random integration of the construct into the fly genome, thus it is not predictable if the integration happens in a transcriptionally accessible part of the genome or not. However, for each construct multiple transgenic lines were analysed (see Figure 5.3.1). Furthermore, it needs to be taken into consideration that the Linker Scan constructs carry the mutation in the context of the complete PCE sequence while the PCE reg1 and PCE extreg1 deletion constructs are missing part of the 5' sequence of the PCE. This deleted region could potentially contain binding sites for factors that are necessary for the strong expression of the reporter construct. It is also possible that, as all four constructs target a small area, only the Linker Scan constructs (PCE LSA and PCE LSA1-5) affect the core binding area of the cardioblast repressor while the deletion constructs (PCE reg1 and PCE extreg1) have a lesser effect on the ability of the repressor to bind to this region, thus quite likely causing a weaker reporter expression in the cardioblasts.

A further explanation for the difference in expression could be the presence of two separate binding sites for the repressor. The results of the PCE S4 and PCE reg1 constructs implies that the negative regulator lies downstream of the PCE S4 cut-off and upstream of the first base of the PCE extreg1 construct. Re-evaluation of the sequence region targeted by the PCE LSA and PCE LSA1-5 constructs showed a putative binding sites for Scalloped and a GATA site which I initially discarded under the applied stringency criteria (see material and methods, chapter 2) as both were neither good matches for the described consensus nor were the sites completely conserved between the different *Drosophila* species used to generate the alignment for the foot-print (see chapter 4). However, the above described results point strongly to the importance of this region. This has led me to investigate these two putative *Him* regulators despite the above-mentioned initial reservations.

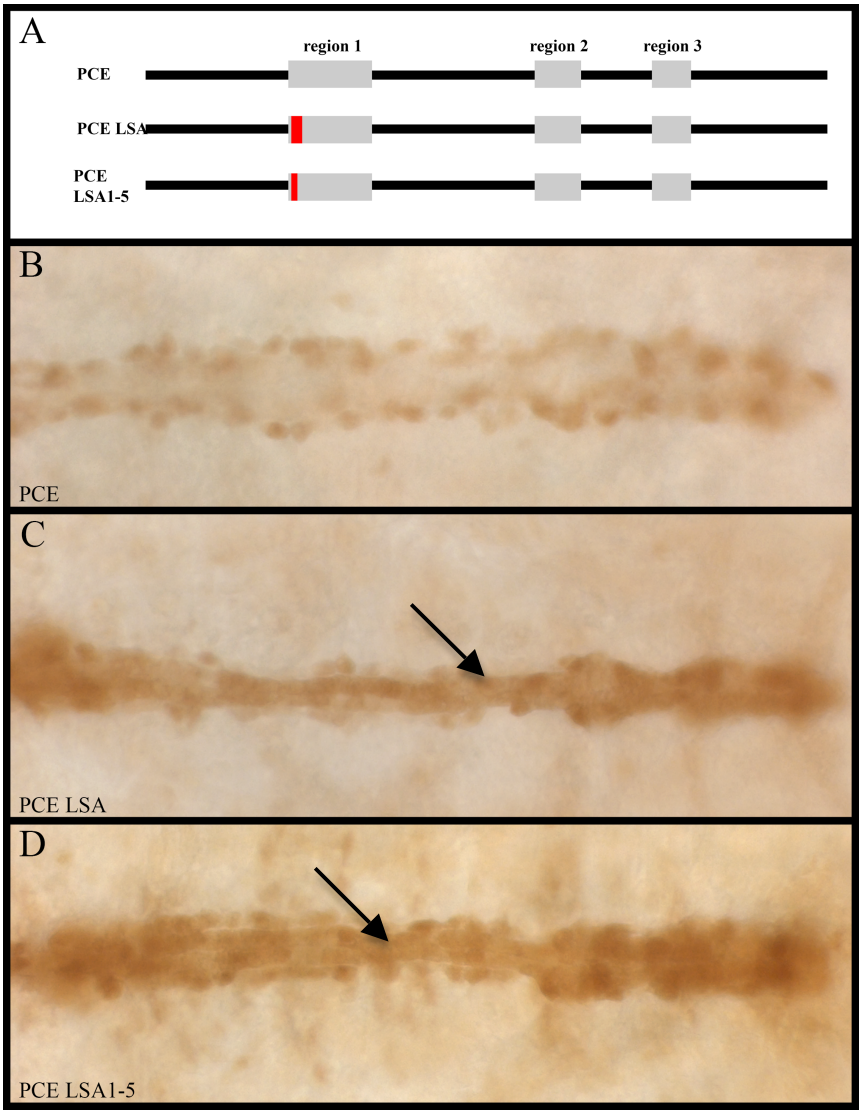


Figure 6.2.2: Mutation of the beginning of the conserved region 1 also induces ectopic expression in the cardioblasts (arrows).
A Schematic representation of the reporter constructs used in this Figure. **B** Wild-type expression pattern of the full length PCE reporter construct in the heart. **C** Mutating 9 bp at the beginning of the conserved region 1 as done in the PCE LSA construct induces ectopic expression in all cardioblasts. **D** Shortening this mutation to 5 bp still induces ectopic expression in all cardioblasts.

The first of these two sites is a Scalloped binding consensus. The Scalloped binding consensus as it occurs in the PCE sequence has a number of substituted bases (AGATATTC in the PCE sequence vs. WVDWATKY as published in Halder et al., 1998), which are likely to at least render any binding of it to this sequence weaker. While the majority of these substitutions have been previously described, the site has not been described in the form it has in the PCE sequence (Halder et al., 1998; Halder and Carroll, 2001). The core ATG binding motif is also not conserved in the PCE sequence. In addition to these factors, the putative binding site is not conserved in two of the screened *Drosophila* species, *D. mojavensis* and *D. virilis*. The Scalloped protein is the only member of the Transcriptional Enhancer Factor-1 (TEF-1) family of proteins in *Drosophila* (Campbell et al., 1992). Scalloped has mainly been described in *Drosophila* wing and bristle development (Campbell et al., 1991; Campbell et al., 1992; Bray, 1999). In order to test if Scalloped is expressed in the heart, I generated a RNA probe for the *scalloped* gene (from DGRC cDNA clone RE23308). Figure 6.2.3 shows the results of this RNA *in situ*. My *scalloped* probe generates a lot of background and only a small amount of specific staining. I observed some expression in the somatic mesoderm but not in the developing embryonic heart. Based on this experiment, I cannot rule out any *scalloped* expression in the heart as the quality of the probe is too poor. Indeed, Deng et al. (2009) report Scalloped expression in the cardioblasts and the developing somatic mesoderm at stage 13 and Guss et al. (2013) report Scalloped expression in the mesoderm. As the majority of the heart cells are generated in stages 11 and 12, this reported expression is slightly later than necessary for *Him* repression in the cardioblasts, but, as it is specific for the cardioblasts and nothing is stated about earlier expression of *scalloped* in the developing heart or dorsal mesoderm, the results of Deng et al. make Scalloped a strong candidate. Further experiments to analyse the expression of *Him* and PCE in scalloped null or hypomorph mutants are necessary to verify this.

Members of the GATA transcription factor family are expressed in the developing heart and regulate the expression of cardiac specific genes in vertebrates (Molkentin, 2000; Latinkic et al., 2003). The only GATA factor that is known to be involved in the development of the embryonic *Drosophila* heart is Pannier (Gajewski et al., 1999; Gajewski et al., 2001; Alvarez et al., 2003; Klinedinst and Bodmer, 2003). *pannier* is expressed in both, the dorsal ectoderm and the dorsal mesoderm. Additionally, *pannier* is expressed uniformly in all heart cells of *Drosophila*, both

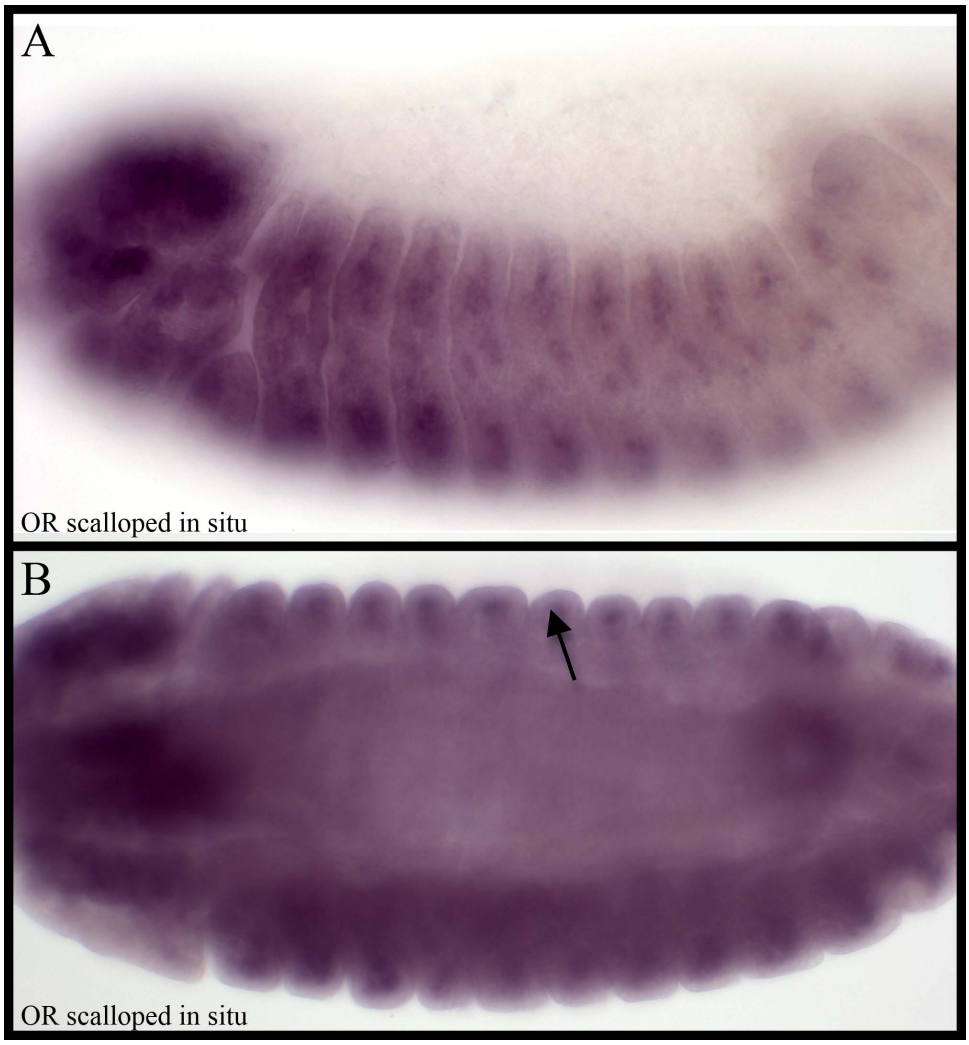


Figure 6.2.3: The *scalloped* transcript is expressed in the somatic mesoderm but not in the cardiac mesoderm.
A Lateral view of a wild-type embryo (OR) labelled for the *scalloped RNA* transcript. Expression can be seen in a segmentally repeated pattern. **B** Dorsal view of the same embryo as in (A). The *scalloped* transcript is expressed in the mesoderm (arrow) and not in the ectoderm of the embryo.

cardioblasts and pericardial cells. So far, Pannier has only been described as an activating factor for transcription. The uniform expression of pannier throughout the dorsal ecto- and mesoderm does not necessarily imply that the function of the Pannier protein is the same in all mesodermal cells as GATA-factors have been shown to function together with co-factors. U-shaped is the main co-factor described as working together with Pannier and is, like *pannier*, expressed in both the dorsal ecto- and mesoderm (for reviews see: Cantor and Orkin, 2004; Sorrentino et al., 2005). U-shaped is mainly required for the maintenance of gene expression and the correct differentiation of both cardiac cell types (Klinedinst and Bodmer, 2003). It is possible that different heart cells express different co-factors for Pannier and thus could generate different specificities for different cell types. This would enable Pannier to function as a transcriptional repressor of *Him* in the cardioblasts without repressing *Him* expression in the pericardial cells.

In order to test my hypothesis that Pannier (together with a co-factor) is responsible for repressing *Him* expression in the cardioblasts, I initially tried to analyse the expression of the *Him* transcript in *pannier* null mutant embryos (pnr^{VX6}). However, as described previously, the loss of heart cells in these animals is nearly complete from an early stage (embryonic stage 12) onwards, making the analysis of the *Him* expression pattern in these embryos very difficult (see Figure 6.2.4 B). Similar observations have previously been published (Heitzler et al., 1996, Klinedinst and Bodmer, 2003). The identity of the remaining heart cells has yet to be studied using the expression of marker genes and it is currently not possible to identify if specific sub-sets of the cardioblasts and/or pericardial cells are more affected by the loss of *pannier* than others. As U-shaped is the most common co-factor for Pannier, I also analysed the *Him* expression pattern in embryos null mutant for *u-shaped* ($Df(2)ush^{rev18}$; see Figure 6.2.4.C; Cubadda et al., 1997). Klinedinst and Bodmer, 2003 describe this line as losing its heart precursors from embryonic stage 11. In my experiment, I obtained similar results. I observed large gaps in the heart making an analysis of the *Him* expression pattern impossible. Next, I tried to analyse the *Him* expression pattern in embryos carrying a hypomorphic allele of *u-shaped* (ush^2) as its phenotype is less severe (Fossett et al., 2000). As can be seen in Figure 6.2.4 D, while there are fewer gaps in the developing heart, the general anatomy of the embryo is still quite disturbed, making an in depth analysis of the *Him* expression pattern impossible. However, the remainder of the *Him*

expression pattern is close to its wild-type pattern, indicating that *u-shaped* most likely is not involved in the transcriptional regulation of *Him*.

Another approach to reduce Pannier and U-shaped activity is the *UAS-pnr^{D4}* expression construct generated in the Bodmer lab (Klinedinst and Bodmer, 2003). In this construct the ability of Pannier to interact with at least one of its co-factors, U-shaped, is inhibited. When under the control of the pan-mesodermal driver *twi-Gal4; 24B-Gal4* this leads to a stronger and broader activation of *tinman* expression, indicating that U-shaped functions in restricting *tinman* expression in this context (Klinedinst and Bodmer, 2003). As can be seen in Figure 6.2.4 E *Him* expression is stronger in these embryos. The band of *Him* expression in the pericardial cells is wider than in wild-type animals, but I have not been able to observe any ectopic expression of the *Him* transcript in the cardioblasts or other tissues that would normally not express *Him*. The observed wider band of *Him* expression is most likely a secondary effect on the *Him* expression due to the reported effect this experiment has on the expression of *tinman* and the increase in all heart cells associated with it (Klinedinst and Bodmer, 2003). This observation is based on the position and shape of the cells stained positive for the *Him* transcript, in order to unambiguously identify these cells as pericardial cells a double-labelling experiment with a pericardial cell marker such as Zhf-1 and a cardioblast marker such as Mef2, will be necessary.

My next approach for analysing *Him* expression in a *pannier* loss-of-function background was to utilize the Gal4/UAS-system to ectopically express a dominant negative form of Pannier in the heart cells. The *UAS-DN-Pnr* expression construct was generated in the Bodmer lab (Klinedinst and Bodmer, 2003). The drivers I used were *Mef2-Gal4* and *hand-Gal4* as both are expressed in all cardioblasts and *hand-Gal4* is additionally expressed in all pericardial cells. This should allow me to see if a (partial) removal of Pannier from the cardioblasts is capable of inducing ectopic *Him* expression indicating a repressor function for Pannier and also possibly give me an indication if removal of Pannier in the pericardial cells leads to a reduction of *Him* expression in the pericardial cells as Pannier is also a possible activator for *Him* expression in the pericardial cells. As can be seen in Figure 6.2.5 there is no loss of *Him* expression in the pericardial cells if driven with *hand-Gal4* ruling out Pannier as an activating influence on *Him* expression in the pericardial cells. The driver I used for this experiment needs to be taken into consideration. *hand-Gal4* starts Gal4 expression relatively late, a repeat of

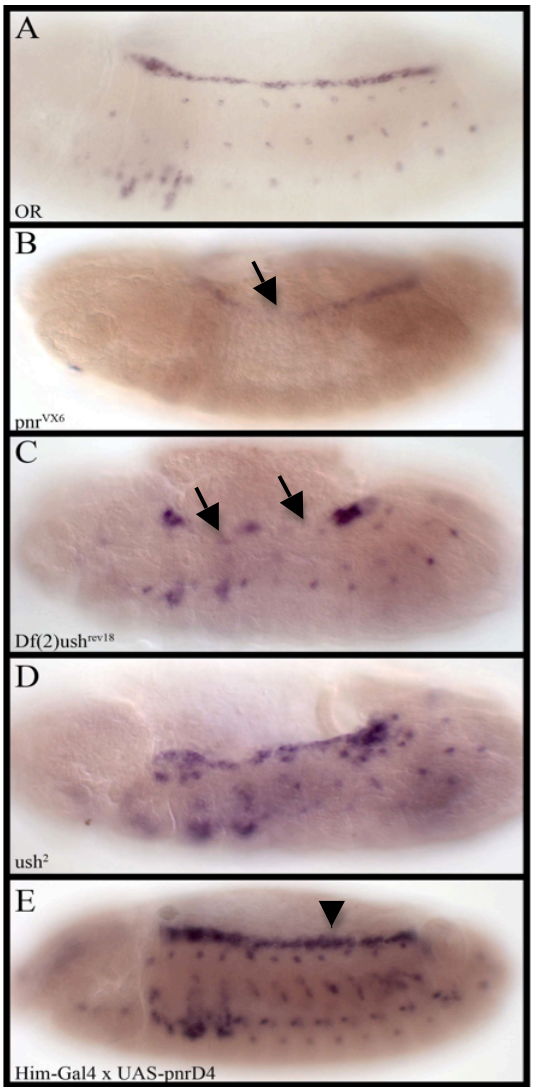


Figure 6.2.4: *Him* expression in *pannier* and *u-shaped* mutants.
A Lateral view of a stage 13 wild-type embryo (*OR*) labelled for the *Him* RNA transcript. **B** *Him* expression pattern in an embryo null mutant for *pannier* (*pnr*^{VX6}). The gaps (arrow) in the expression pattern are caused by pronounced cell loss in the dorsal regions of the embryo. **C** Labelling the *Him* transcript in an embryo null mutant for *u-shaped* (*Df(2)ush*^{ref18}) shows a very severe loss of the cells of the dorsal mesoderm, there are gaps (arrow) within the *Him* expression pattern. **D** *Him* expression pattern in a *u-shaped* hypomorphic embryo (*ush*^{2}). While less severe than the previous two shown fly lines, the heart development is still too disturbed to allow for the analysis of the expression pattern of the *Him* transcript. The remaining *Him* expression pattern is very similar to wild-type. **E** Expanded expression pattern of the *Him* transcript in the heart region (arrowhead) in a stage 13 embryo expressing the *UAS-pnrD4* construct under the control of the *Him-Gal4* driver. Ectopic *Him* expression in the cardioblasts was not observed.

When the dominant-negative form of Pannier is expressed in the cardioblasts using *Mef2-* and *hand-Gal4* drivers, there appears to be some additional expression in these cells as can be seen in Figure 6.2.5 (arrows). However, although the expression of the dominant-negative form of Pannier leads to a less severe phenotype than in the previously described *pannier* mutants, there is still a certain degree of cell loss in the heart which in turn results in gaps and organisationally differences in the heart structure. As cardioblasts lose their direct neighbours and cellular context within the heart, they also seem to lose their distinctive cell shape. This makes it difficult to unambiguously identify any cell as a cardioblast or pericardial cell without an additional cell type marker. To aid in the identification of the different heart cells, I double-labelled embryos driving the dominant-negative form of Pannier under the control of *Mef2-Gal4* and *hand-Gal4* with a *Him* FISH and ß3-tubulin fluorescent immuno-stain which is expressed in four of the six cardioblasts and the developing somatic mesoderm. If the loss of some Pannier function leads to ectopic *Him* expression in the cardioblasts and does not affect the expression of *ß3-tubulin* in these cells, I should be able to observe co-expression of the *Him* transcript and the ß3-tubulin protein in the cardioblasts. As can be seen in Figure 6.2.6 this is not the case, the *Him* transcript and the ß3-tubulin protein do not co-localize in the same cells in the heart. Thus, Pannier is most likely not involved in the regulation of *Him* expression in the ß3-tubulin-positive cardioblasts. As *ß3-tubulin* is not expressed in the remaining two cardioblasts per hemisegment, I cannot comment upon the role of Pannier in these two cells. I also cannot rule out that my observed results are due to non-complete removal of functional Pannier from the cells the dominant-negative form was expressed in, as this indirect approach will not have removed all Pannier activity.

In summary, these results indicate the possibility of two separate, as of now still unidentified, repressors acting on the 5' end of the PCE sequence to limit *Him* expression in the cardioblasts. The distance between the beginning of the PCE extreg1 construct and the first mutated basepair of the PCE LSA1-5 is 11 bp. This length is exactly one turn of the DNA double helix, thus it is possible that complexes that rely on binding sites further away are still capable of binding. While I cannot rule out a large repressor complex that requires both regions to be present and intact to function, the distance of 11 bp between the beginning of the PCE extreg1 construct and the first mutated basepair of the PCE LSA1-5 construct is larger than the most common length

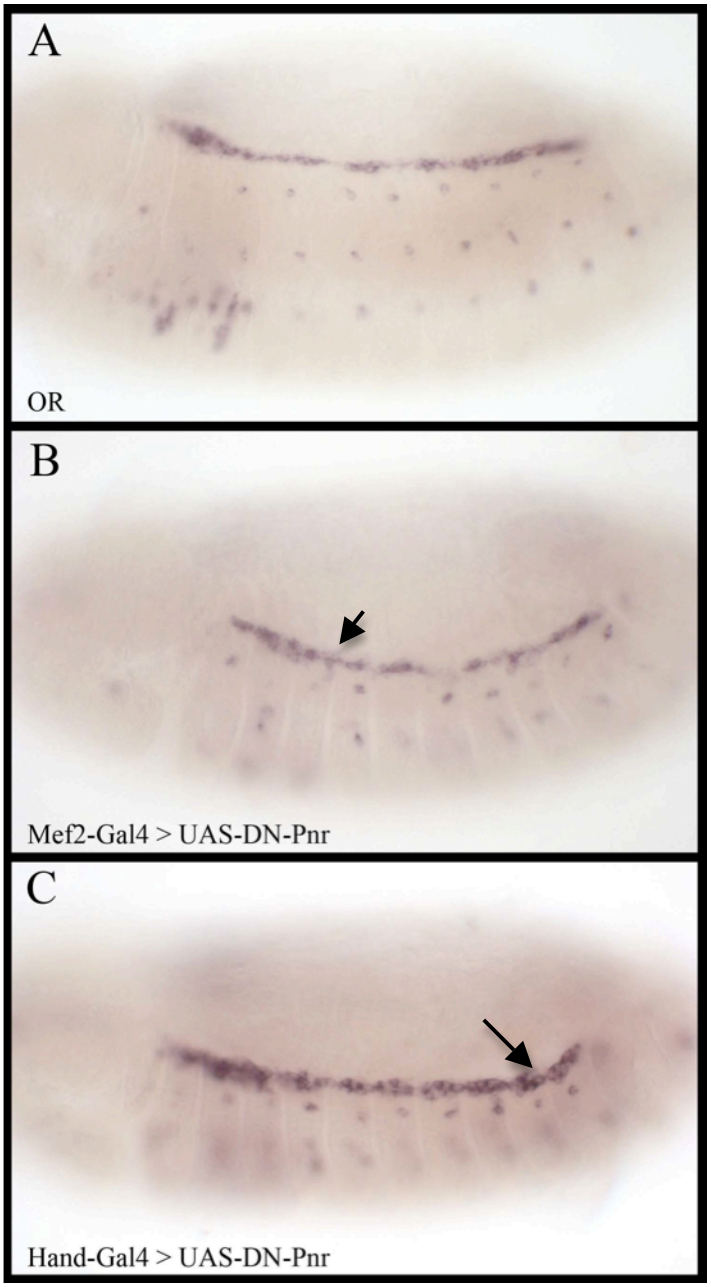


Figure 6.2.5: *Him* expression in embryos expressing a dominant-negative form of Pannier.
A Lateral view of a stage 13 wild-type embryo (OR) labelled for the *Him* RNA transcript. **B** *Him* expression pattern in an embryo expressing the dominant-negative form of Pannier under the control of the *Mef2-Gal4* driver. There is an increase in *Him* expression (arrow). **C** *Him* expression pattern in an embryo expressing the dominant-negative form of Pannier under the control of the *hand-Gal4* driver. Similar to (B), there is an increase in *Him* expression, which is most likely due to general cardiac hyperplasia (arrow).

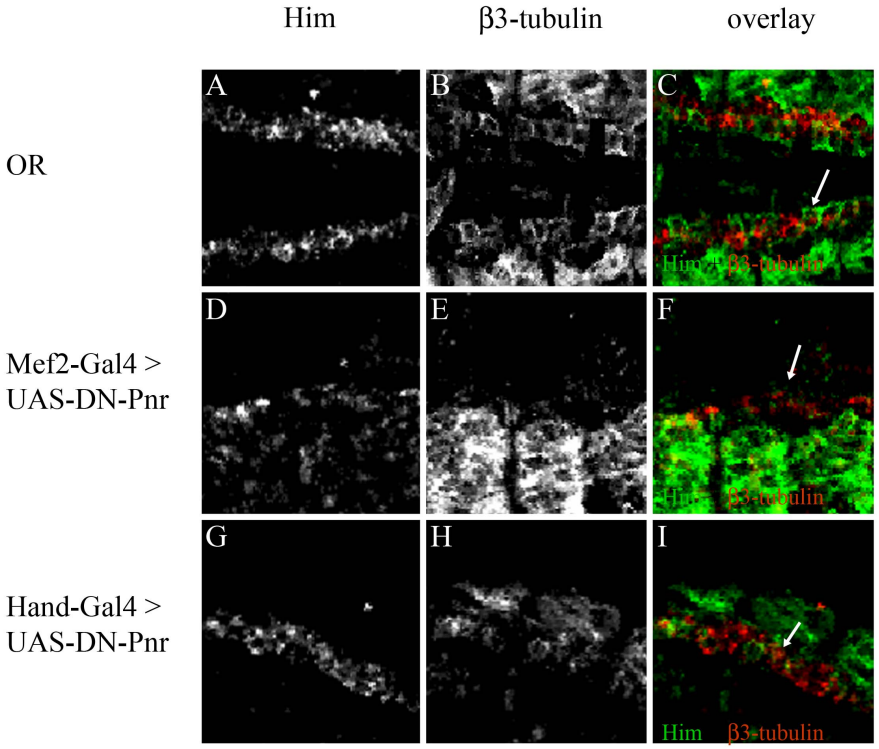


Figure 6.2.6: The dominant-negative form of Pannier does not induce ectopic *Him* expression in the cardioblasts.
Fluorescent *In Situ* Hybridisation (FISH*)* for *Him* (green in overlay) and fluorescent antibody labelled for ß3-tubulin (red in overlay).
A-C Stage 14/15 wild-type embryo (*OR*). *Him* and ß3-tubulin expression do not overlap (arrow). A *Him* FISH, B ß3-tubulin antibody, C overlay **D-F** Stage 13 embryo expressing the dominant-negative form of Pannier under the control of the *Mef2-Gal4* driver. *Him* and ß3-tubulin expression do not overlap (arrow). D *Him* FISH, E ß3-tubulin antibody, F overlay. **G-I** Stage 13/14 embryo expressing the dominant-negative form of Pannier under the control of the *hand-Gal4* driver. The *Him* transcript and the ß3-tubulin antibody do not colocalize (arrow). G *Him* FISH, H ß3-tubulin antibody, I overlay.

of a typical binding site, the core motif of which is usually between five and ten basepairs of length. This strengthens the argument for two separate binding sites. One also needs to consider that these regions can have a function in shaping the "landscape" of the DNA, and not necessarily in transcription factor binding. For example 13 bp into the PCE S4 sequence there is a stretch of fourteen continuous adenine bases. As previously mentioned, this long repeat of one base will change the landscape of the DNA and it is possible that this three dimensional structure due to the long Adenine repeat is necessary to prevent PCE and *Him* expression in the cardioblasts. In this context it is also notable to mention again that the ectopic expression observed in the cardioblasts is slightly weaker in the PCE reg1 and PCE extreg1 constructs than in the PCE LSA and PCE LSA1-5 constructs, further substantiating the possibility of two different regulatory mechanisms at work or that one mechanism necessary for *Him* expression in the cardioblasts is affected in two different ways in these four constructs.

6.3 *Him* expression is repressed in the amnioserosa

Additional analysis identified a further repressor active in the dorsal part of the embryo. Two separate 3' deletion constructs reveal at least one repressor that is necessary to exclude reporter gene expression from the amnioserosa cells. As mentioned in chapter 5, the PCE Δtin2 construct induces ectopic expression in the amnioserosa cells while losing expression of its reporter in the pericardial cells (see Figures 5.3.1 and 6.3.1). A smaller deletion construct, PCE S2, made previously in the lab (S. McConnell and M. Taylor, unpublished results) also shows ectopic, but slightly weaker, expression in the amnioserosa (see Figure 6.3.1). PCE S2 only deletes the 3' 40 bp of the PCE sequence while the PCE Δtin2 construct deletes the 3' 198 bp of the PCE sequence. As both deletion constructs show a similar ectopic expression pattern in the amnioserosa cells, the most likely explanation for this is that the most 3' 40 bp of the PCE sequence contain a repressor binding site. Again, it is possible that there are two different repressors or a large complex at work as the shorter deletion construct gives weaker ectopic expression than the larger deletion. However, the computational analysis (see chapter 4) has not highlighted any potential known binding sites necessary for the maintenance of the amnioserosa. Possibilities are the members of the *u-shaped* family

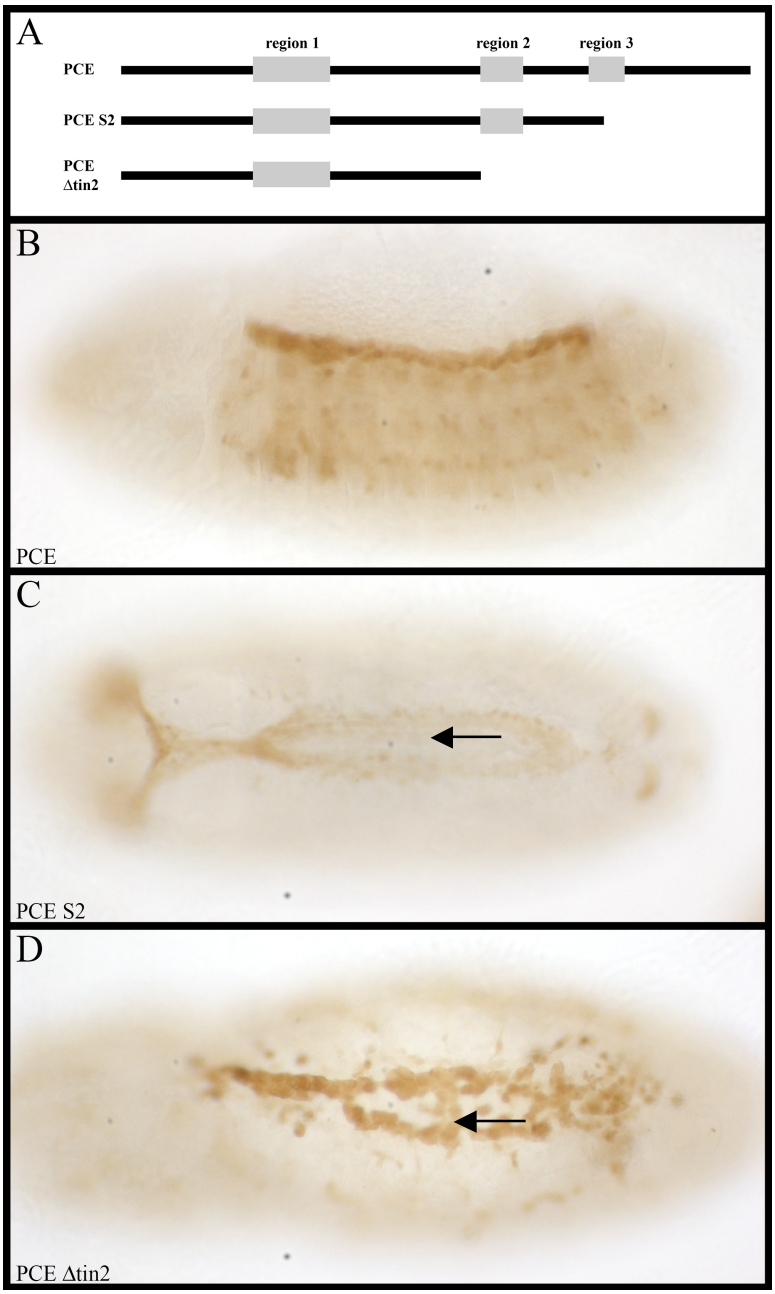


Figure 6.3.1: The 3' end of the PCE construct contains an amnioserosa repressor.
A Schematic representation of the reporter constructs used in this Figure. **B** Wild-type expression pattern of the full length PCE reporter construct in the heart. Lateral view. **C** The expression pattern of the PCE S2 reporter construct shows ectopic expression in the amnioserosa cells (arrow) while the reporter expression in the heart is lost. Dorsal view. **D** Animals carrying the PCE Δtin2 construct show similar ectopic expression in the amnioserosa (arrow). Dorsal view.

of genes (*u-shaped, hindsight, serpent, tail-up, dorsocross1, dorsocross2 and dorsocross3*) as all of them have been shown to be necessary for the maintenance of the amnioserosa (Reim et al., 2003; Hamaguchi et a., 2004); in mutants for these genes the amnioserosa cells enter apoptosis and the amnioserosa degenerates (Frank and Rushlow, 1996). Further investigation into the role of any of these genes in *Him* regulation might uncover a second repressor.

6.4 Hand does not activate *Him* expression

The bioinformatic analysis described in chapter 4 also highlighted several E-boxes within the PCE sequence. The only bHLH transcription factor known to be expressed in the heart is Hand. Hand (*h*eart, *a*utonomic *n*ervous system, neural crest-*d*erived cell types) is a bHLH transcription factor expressed in all heart cells as well as in the lymph and garland gland, the visceral mesoderm and a few cells of the nervous system (Moore et al., 2000; Kölsch and Paululat, 2002). The consensus binding sequence for bHLH transcription factors is called an E-box and has the general sequence of CANNTG (Ferré-D'Amaré et al., 1993; Ellenberger et al., 1994; Ma et al., 1994). The two core basepairs and the surrounding sequences give an E-box its specificity for different bHLH transcription factors (Dang et al., 1992; Blackwell et al., 1993). It is currently unknown which sequence the two core base pairs need to have in a Hand E-box.

Within the PCE sequence there are four E-box sequences. Two of these four E-box sequences are contained within the consensus for the Tinman (*CA*CT*TG*A). The third of the E-boxes is located in the non-conserved sequence between the conserved regions 1 and 2 and the fourth E-box is located just at the beginning of the conserved region 2 of the PCE sequence.

In order to see if Hand is at all involved in the regulation of *Him*, I analysed the expression of the *Him* and *hand* transcripts in the deficiency line w^{1118}*; Df(2L)Exel7046* (BDSC, Bloomington stock 7819). This line is a large deletion from 31B1 to 31D9 of the left arm of the second chromosome that contains the *hand* gene and 14 other genes. As can be seen in Figure 6.4.1 embryos homozygous for this *BL7819* deficiency do not

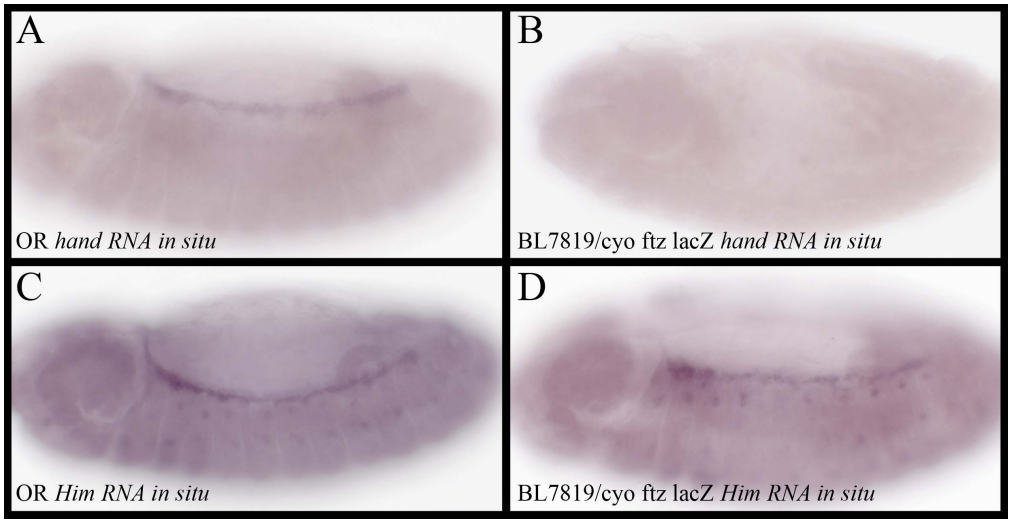


Figure 6.4.1: Hand does not regulate *Him* expression.
hand (A,B) and *Him* (C,D) RNA *In Situ* Hybridisation on wild-type and hand-deficient embryos.
A Expression pattern of the *hand* transcript in wild-type embryos (*OR*). **B** Embryos null mutant for the *hand* gene do not show any *hand* transcript expression. **C** *Him* transcript expression pattern on wild-type embryos (*OR*). **D** The expression pattern of the *Him* transcript is unaltered in embryos null mutant for the *hand* gene.

express any *hand* transcript (compare 6.4.1 B to the wild-type in 6.4.1 A) confirming that this line is deficient for *hand*. However, *Him* expression is completely normal in embryos homozygous for this deficiency (see 6.4.1 D). This result indicates that *Hand* (or any of the other genes affected by the deficiency) is not involved in the activation of *Him* expression.

6.5 Relevance of the E-box sequence contained within the Tinman binding consensus

While examining the presence of the E-boxes within the PCE sequence, I noted that the Tinman consensus also contains an E-box. The presence of E-boxes in the Tinman consensus binding site has so far not been analysed. The only reference regarding the E-box contained in the most common Tinman consensus that I have been able to find is within the discussion of the Durocher et al. (1996) paper. All previously published mutations that have been shown to disturb Tinman binding, target the four main basepairs of the E-box sequence within the Tinman binding consensus.

In order to investigate if the Tinman binding consensus is functional *in vivo* without the E-box, I targeted the S5 site, which I have shown to be necessary for PCE activation and to be responsive to Tinman, with two further point-mutations in separate constructs. In each construct I mutated only one basepair within the PCE sequence. The PCE*S5C-T (T***T**AAGTG) changes the first C of the E-box sequence within the S5 Tinman binding site to a T. This should generate a still functional Tinman binding site while destroying the E-box. The thus created Tinman binding site has been identified and described as functional *in vivo* by Kremser et al., 1999; Gajewski et al, 1997; Han and Olson, 2005. Ozdemir et al. (2011) describe how the introduction of a single basepair mutation into the E-box sequence (from *C*ANNTG to ***G***ANNTG) within the *rho* enhancer limits reporter expression to a narrower domain and was thus successful in interfering between the binding site and its activating protein.

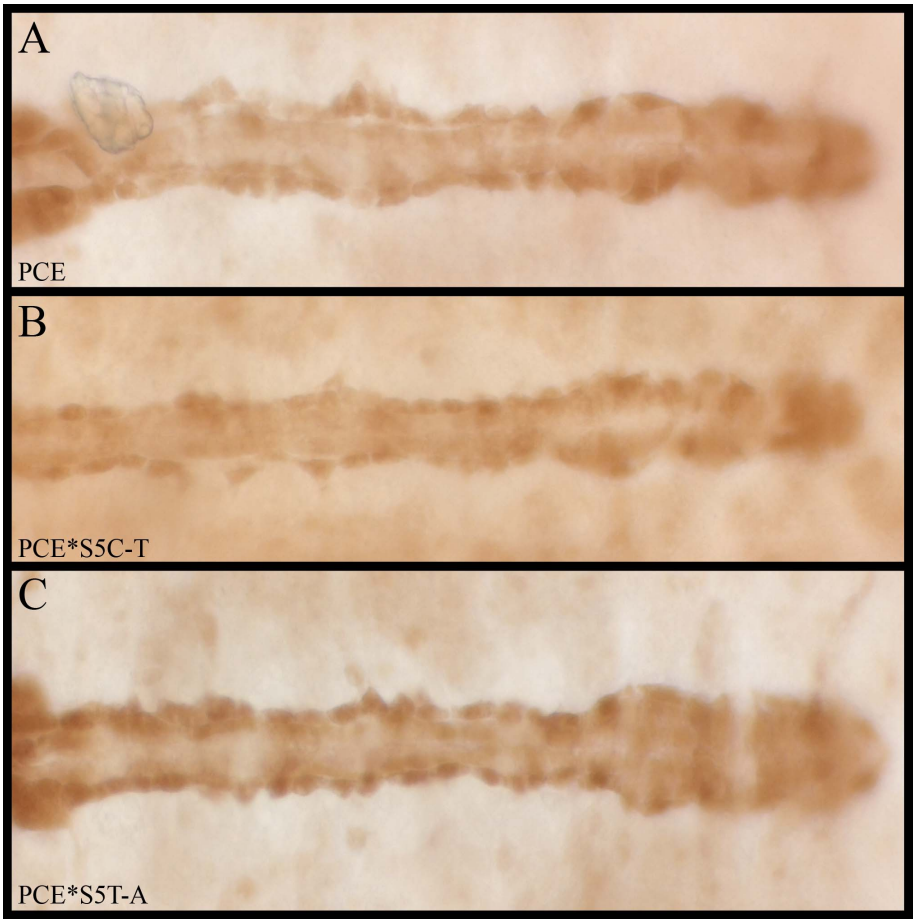


Figure 6.5.1: Single basepair mutations introduced into the PCE sequence to test the importance of the E-box contained within the Tinman binding consensus.

A Wild-type expression pattern of the full length PCE reporter construct in the heart. **B** Changing the first C of the Tinman binding site to a T, maintaining a functional Tinman site but mutating the E-box does not alter the reporter gene expression pattern. **C** Mutating the 5' T to an A, which leaves the E-box intact, also does not alter the expression pattern of the reporter genes.

The PCE*S5T-A mutation (***A***CAAGTG) mutates the first base of the S5 Tinman binding site from a T to an A. This should not affect the core of 6 bp of the E-box binding while destroying the ability of Tinman to bind to this site. According to the position weight matrix shown in Figure 5.5.1 (see chapter 5; Zinzen et al., 2009) it is highly unlikely for an A to occur in this position in a functional Tinman binding site. As can be seen in Figure 6.5.1 neither of the so mutated reporter construct results in a change of the expression pattern of the reporter gene in the heart of *Drosophila* embryos. As neither mutation appears to have an effect *in vivo*, there are two conclusions that can be drawn. A possible explanation is that the presence of the E-box within the Tinman binding site sequence is of no direct importance for *Him* expression. If the site is still capable of binding Tinman, the destruction of the E-box has no effect on the expression of *Him* in the pericardial cells. It is also possible that mutating just one basepair of the whole Tinman/E-box consensus site is not sufficient to interfere with any transcription factor binding. This is supported by the fact that the four basepair mutation of the Tinman consensus binding sequence described in chapter 5 renders the reporter construct inactive and non-responsive to Tinman. Consequently, I expected a similar result for the PCE *ST-A construct, which was created with the aim of destroying any ability of Tinman to bind to the S5 site. An EMSA will be necessary to verify if these mutations affect Tinman binding or not *in vitro*.

A further explanation of these results is that both Tinman and a bHLH factor are capable of activating *Him* transcription and *Him* transcription is only lost if both factors cannot bind to the S5 site. The identity of the bHLH factor is currently unknown, a likely remaining candidate is Twist. Initially the ideal Twist E-box was described as CATATG (Halfon et al., 2000), but recent studies show that *in vivo* Twist favours a CA or TG core (Zinzen et al., 2009, Ozdemir et al., 2011). According to these studies, Twist is unlikely to bind to the E-box contained within the Tinman binding consensus. As there are no other ways of delineating the E-box from the Tinman consensus binding site, of which 6 of the 7 bp are part of both the E-box and the Tinman consensus binding site, I could not pursue this any further.

6.6 Further activator sites are contained within the PCE sequence

As explained previously, Tinman alone is not sufficient to explain the *Him* expression pattern in the pericardial cells; at least one further activator is needed. Combining the results from all available PCE enhancer constructs highlights the region between the PCE Δreg1 and PCE Δtin1 deletion constructs (see Figure 6.6.1).

The PCE Δreg1 construct exhibits wild-type like reporter gene expression in the pericardial cells (and ectopic expression in the cardioblasts) while reporter gene expression is completely lost in animals carrying the PCE Δtin1 deletion construct (see Figure 6.6.2). My previous experiments (chapter 5) have eliminated the S6 site as a major input into the transcriptional regulation of the PCE, the activator must be located upstream of this site. The difference in length between these two constructs is only 53 bp. This area is well conserved as can be seen in the computational analysis of the PCE sequence in chapter 4. These 53 bp correspond to the beginning of the conserved region 1. As explained in part 6.2 the first 5 bp of the conserved region 1 are important to repress activity in the cardioblasts. From the analysis of this region I can exclude the first 11 bp as important for the activation (see Figures 6.2.2 and 6.6.1, PCE LSA construct) in the pericardial cells. This leaves me with 5 bp upstream of the mutated PCE LSA construct and 38 bp between the end of this mutated construct (PCE LSA) and the beginning of the PCE Δtin1 construct that should contain the potential activator binding site (see Figure 6.6.1). This sequence contains a consensus binding site for the vertebrate Oct-1 protein. The *Drosophila* homolog of Oct-1 is *nubbin*. Both, the vertebrate Oct-1 and its *Drosophila* homolog Nubbin, have been shown to be involved in cell fate determination (Neuman and Cohen, 1998; Irvine, 1999). *Drosophila nubbin* is mainly expressed in the ectoderm and plays an important role during neurogenesis, neurospecification and wing development (Ng et al., 1995; Yeo et al., 1995). So far, very little is known about its binding preferences. According to the only published DNA footprinting analyis, Nubbin has a fairly degenerate binding consensus (Neuman and Cohen, 1998). In their paper, Neuman and Cohen describe four different sequences that Nubbin protects. The putative binding site in the PCE sequence was identified with the vertebrate Oct-1 consensus sequence in the TRANSFAC public database. The

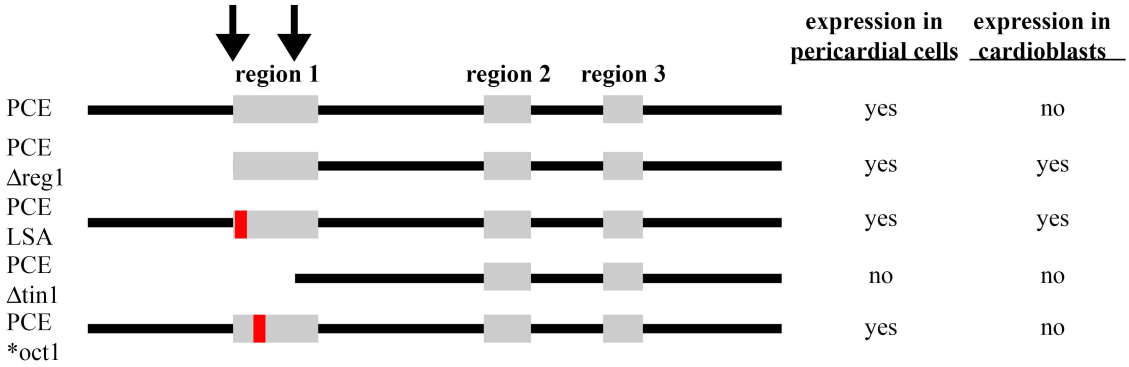


Figure 6.6.1: Expression construct overview highlighting the area (between the arrows) where a further activator for *Him* should be located.

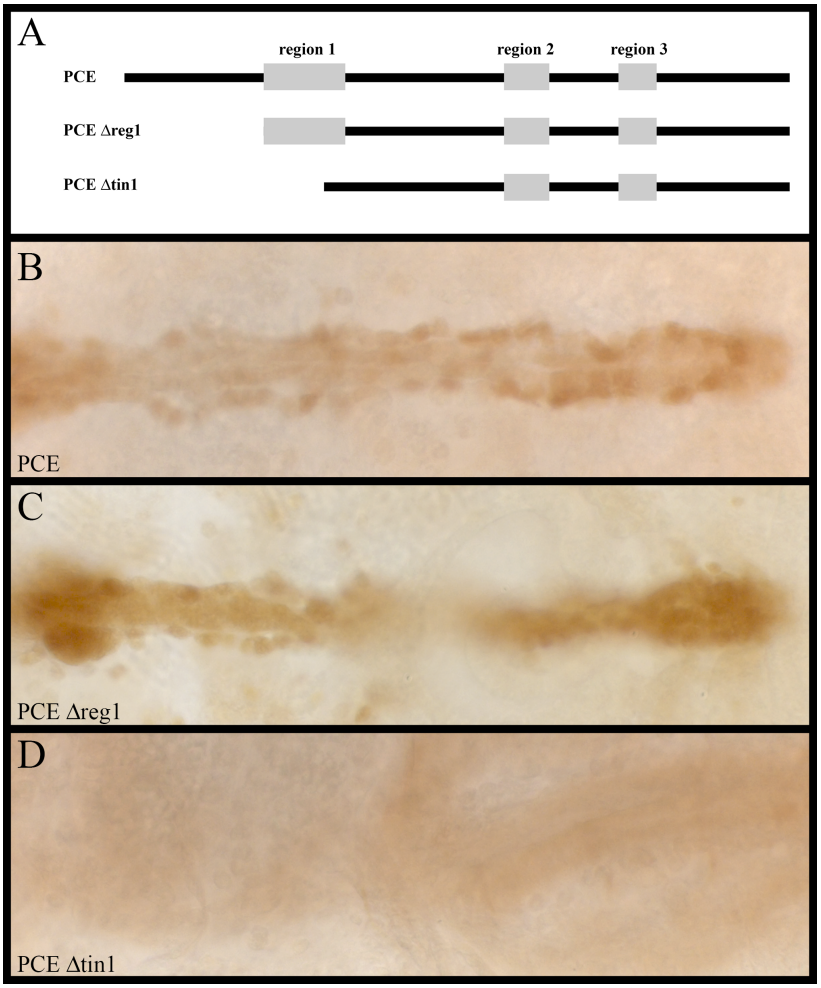


Figure 6.6.2: Between the cut-off points for the PCE Δreg1 and PCE Δtin1 construct lies another activator site.
A Schematic representation of the reporter constructs used in this Figure. **B** Wild-type expression pattern of the full length PCE reporter construct in the heart. **C** The expression pattern of the PCE Δreg1 reporter construct is not distinguishable from that of the full length construct in the pericardial cells. **D** The deletion construct PCE Δtin1 lacks any reporter gene expression in the heart.

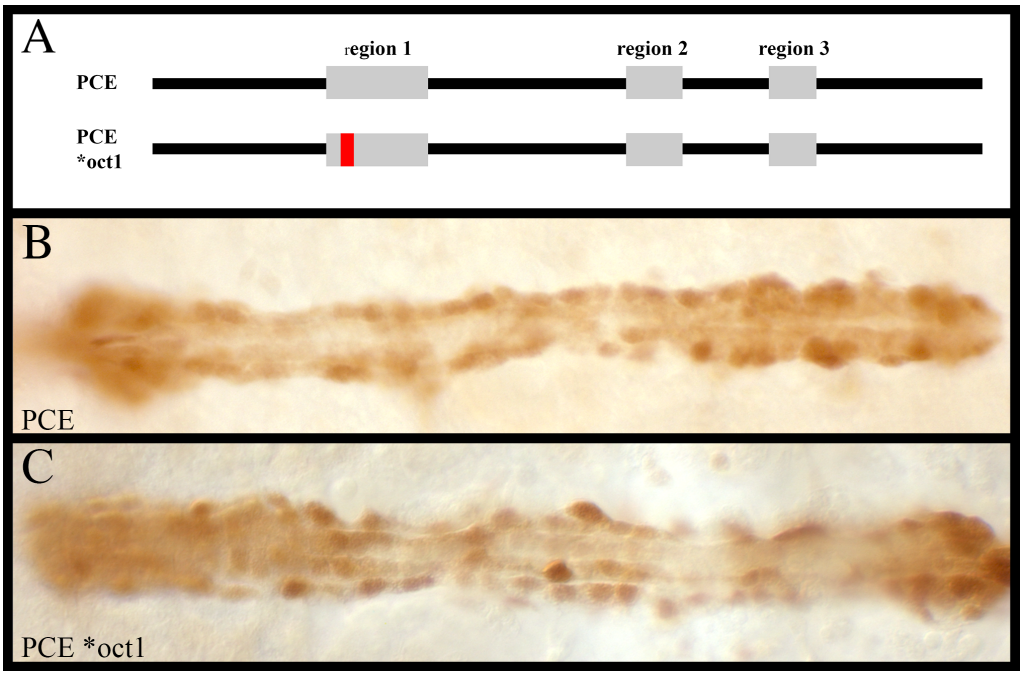


Figure 6.6.3: Mutation of the vertebrate Oct-1 consensus in the PCE sequence does not alter reporter gene expression of the PCE.
A Schematic representation of the reporter constructs used in this Figure. **B** Wild-type expression pattern of the full length PCE reporter construct in the heart. **C** Stage 16/17 embryonic heart of an animal transgenic for the PCE *oct1 construct. Mutating the putative Oct-1/Nubbin binding site does not alter the construct's expression pattern.

consensus sequence used by the TRANSFAC position weight matrix (NNNNATGCAAATNAN; matrix identifier V$OCT1_Q6) matches only one of the four protected sequences (TTATGtAAgTAACC) found by Neumann and Cohen except for two bases at position eight and eleven (lower cases in the protected sequence) of the vertebrate Oct-1 consensus sequence. So far, there are no reports of mesodermal expression for *nubbin*. However there are cases where the expression of one factor in a different germlayer is crucial for mesoderm development in *Drosophila* through an indirect effect of gene transcription (for example see Carmena et al., 2002). Despite this, I mutated five base pairs at the beginning of the vertebrate Oct-1 binding consensus found in the PCE sequence from *CGCAA*ATTG to ***ATCCG***ATTG. Again, this mutation was checked with the bioinformatic tools available to me to rule out the introduction of any additional known binding sites. If this sequence of the PCE is important for its regulation, this mutation should alter the expression pattern of the reporter gene. As shown in Figure 6.6.3 this is not the case. Reporter gene expression is unaffected in embryos transgenic for this construct; its expression in the heart of the developing embryos is identical to that of the wild-type PCE construct.

The previous experiment rules out this area as essential for the activation of gene expression. This points towards the 5' 11 basepairs upstream and the 19 basepairs downstream of the binding consensus for the vertebrate Oct-1 protein and the initial 5 basepairs of the PCE Δreg1 reporter construct. Further experiments aiming to find a second activator of the PCE and *Him* should concentrate on analysing these areas. The 11 basepairs 5' of the vertebrate Oct-1 consensus contain a potential Medea binding site. Medea is a downstream component of the Dpp signalling cascade (Hudson et al., 1998; Wisotzkey et al., 1998). Dpp itself is necessary for mesoderm development (see chapter 1) and has been shown to be important for the development of the embryonic *Drosophila* heart (Frasch 1995). This makes the 11 basepairs upstream of the Oct-1 binding consensus very interesting. As it is possible to mutate all 11 basepairs in one enhancer construct, this should be one of the next constructs to be made and tested.

6.7 Other sequence areas of PCE that can be excluded

Initially, I chose to create further enhancer constructs based on the degree of conservation uncovered in my bioinformatic foot-printing analysis and search for potential binding sites (see chapter 4). Using this approach, I made numerous different PCE constructs and several deletion constructs not yet mentioned in this dissertation. For all these constructs I analysed either the distribution of the LacZ reporter protein or the *lacZ* transcript in the developing embryo as the protein is usually more stable and might obscure a more transient regulatory event that might be visible when analysing the expression pattern of the transcript. Constructs based upon this approach that have already been mentioned are the PCE Δreg1, PCE Δtin1, PCE Δtin2, PCE LSA, PCE LSA1-5 and PCE*oct1 (for an overview Figure including all created reporter constructs please refer to the Figure 5.3.1). These constructs will not be discussed again. With the exception of the PCE*oct1 and the reporter constructs known to show no expression or ectopic expression (PCE Δtin2, PCE LSA1-5, PCE LSA1-5) all these constructs were also analysed for the expression pattern of the lacZ RNA and no differences were found in regard to the protein expression pattern.

The PCE LSD4-9 enhancer construct mutates five consecutive base pairs towards the 3' end of the first conserved region; these 5 bp are not deleted within the PCE Δtin1 construct and are completely conserved in the bioinformatic footprint (see chapter 4). I chose to analyse the expression pattern of the transcript rather than that of the reporter protein to ensure that I would not miss any later events during heart development. As mentioned above, the stability of the LacZ protein might obscure more dynamic changes in the expression pattern. These changes are more readily visible if one screens for the expression pattern of the less stable *lacZ* transcript instead. As can be seen in Figure 6.6.4 B mutating the 5 bp of the PCE LSD4-9 enhancer construct does alter the expression pattern of the reporter gene when assayed for the distribution of the *lacZ* reporter transcript. In general the transcript is distributed similarly as in animals carrying the wild-type PCE construct. However, the expression pattern of the *lacZ* transcript appears much less uniform and shows repeated areas of lower and higher expression levels. This indicates that the sequence mutated in the PCE LSD4-9 construct might be involved in a part of the regulation of the PCE construct and *Him* expression.

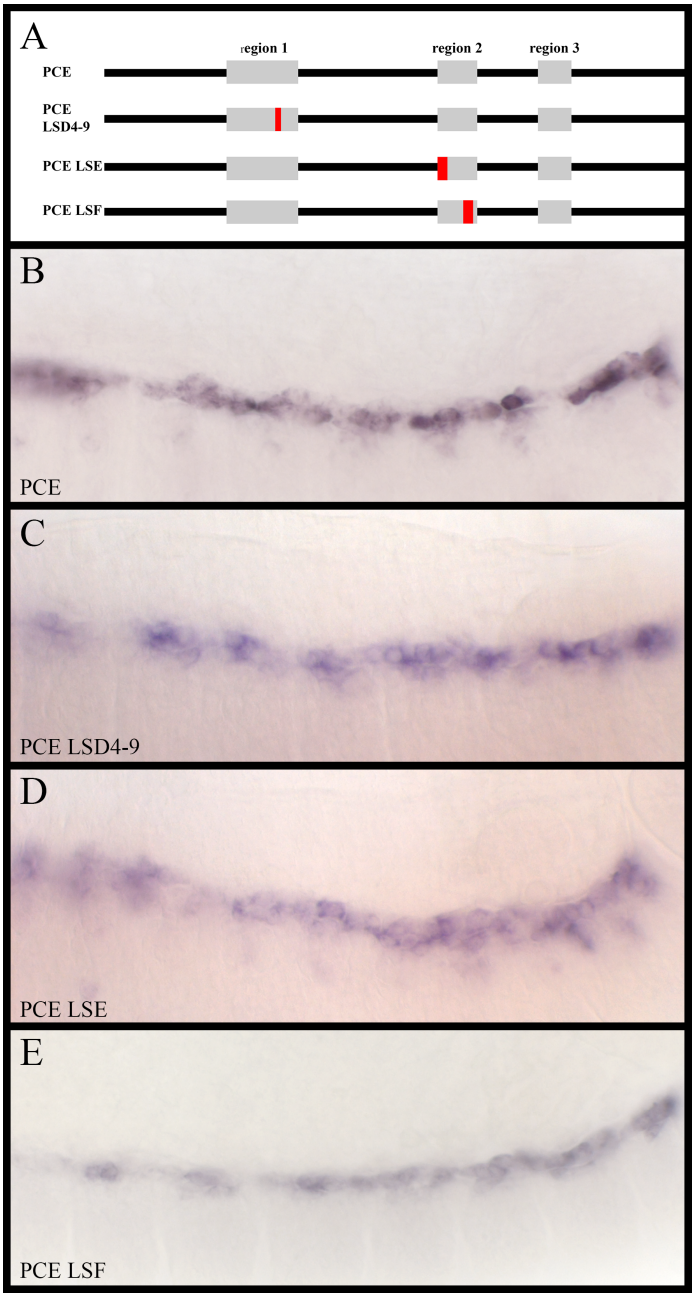


Figure 6.6.4: Further areas of the PCE sequence, which are not involved in its expression.
lacZ In Situ Hybridisation on stage 13 embryos.
A Schematic representation of the reporter constructs used in this Figure. **B** Wild-type expression pattern of the full length PCE reporter construct in the heart. **C** Mutation of 5 well-conserverd basepairs 3' from the putative Tin1 site alters the levels of the expression of the PCE LSD4-9 construct. **D** Mutation of the first 10 bp of the second conserved region also does not alter the expression pattern of the PCE LSE construct. **E** Changing the sequence of 10 bp towards the 3' end of the second conserved region does not alter the expression pattern of the PCE LSF construct.

Two further mutated constructs, termed PCE LSE and PCE LSF, mutate 11 and 10 bp 5' and 3' of the S5 Tinman binding site in the second conserved region of PCE. Figure 6.6.4 shows the transcript expression pattern of the reporter gene for both constructs. Similarly to the PCE LSD4-9 construct these two constructs do not show any alterations to their expression pattern when compared to the wild-type PCE construct. The area covered in the PCE LSE construct covers an E-box and a second binding consensus for the vertebrate Oct-1 protein and the PCE LSF mutation affects the first four basepairs of a putative AP-1 binding site (see chapter 4). A *Drosophila* equivalent to the vertebrate AP-1 family is the gene *kayak*, which is involved in the process of dorsal closure (Noselli and Agnes, 1999; Harden, 2002). As mutation of any of these does not affect the expression pattern of the reporter gene, all of these areas can be excluded from the analysis of regions involved in the regulation of the *PCE* fragment and of *Him*.

The stretch of "unconserved" sequence between the first and second conserved region contains several potentially interesting putative transcription factor binding sites. It contains two Medea sites in short succession followed by a Zfh-1 site and an E-box. In order to analyse this area, I created two further deletion constructs. The first one of these two constructs is called PCE E-box. The construct design is based upon two separate observations: The deletion construct PCE S1 previously created in the lab by Stewart McConnell deletes the 5' 237 bp of the PCE sequence (McConnell and Taylor, unpublished results). As can be seen in Figure 6.6.5 this construct shows no reporter gene activity. While scanning the whole PCE sequence for E-boxes, I also found that this deletion exactly cuts off an E-box located in the unconserved sequence stretch between the first and second conserved regions of the PCE. The sequence of this particular E-box (aCACATGt) has previously been shown to be capable of binding Twist and to be of functional importance for the regulation of *even-skipped* (Halfon et al., 2000). Recent studies by Zinzen et al. (2009) and Ozdemir et al. (2011) have shown that in vivo Twist prefers a CA E-box core sequence and that Twist also has a preference for the surrounding nucleotides (a leading-A and lagging-T residue). These surrounding nucleotides for this E-box are within these preferences; thus it is a good candidate site for possible activation through Twist.

In order to determine if this E-box is of any importance to the regulation of the PCE fragment, I created the PCE E-box construct. This construct extends the PCE S1 construct by 8 bp in the 5' direction, restoring the E-box and the 3 bp upstream of it. If

this E-box is of relevance for the regulation of the PCE, this should restore at least some aspects of the PCE reporter expression. As can be seen in Figure 6.6.5, *Drosophila* embryos transgenic for this construct show a very faint reporter expression in the developing heart. This faint reporter gene expression has been verified in several independent lines (see Figure 5.3.1). This shows that this E-box is of importance for the expression of the reporter gene. It also shows that just the presence of the S5 Tinman binding site and this E-box is not enough to achieve wild-type expression levels. Interestingly, the longer PCE Δtin1 construct does not result in reporter expression (see figure 5.3.1 and figure 5.3.2). This indicates that there are likely further repressor binding sites located in the area between these two deletions. If these binding sites are included within the enhancer construct, the E-box is of less importance than the currently unidentified activator that binds further 5'. It is thus possible that this E-box is of importance for modulating the expression levels of the wild-type. Twist is a good candidate transcription factor for binding this site.

A little further upstream from this E-box my bio-informatic analysis has revealed two potential Medea and one potential Zfh-1 binding site. These sites are also in a non-conserved region. As discussed earlier in this chapter Medea is a possible regulator and Zfh-1 is the only other known factor that is expressed in all pericardial cells (up to stage 14/15, Johnson et al., 2003; Su et al., 1999; Lai et al., 1991). I created a second deletion construct (PCE Medea) that extends the PCE E-box construct by adding another 35 bp in the 5' direction to include these potential binding sites. Again, I assumed that if any of these sites were of importance *in vivo*, I would be able to observe a change in expression pattern. As can be seen in Figure 6.6.5 including these three potential binding sites results in a very faint expression of the reporter gene, similar to that observed for the PCE E-box construct. This low expression could be explained if several Medea binding sites (some of which are further towards the 5' end of the PCE construct and not included in the PCE Medea construct) are needed to achieve full expression levels. It is also possible that, as the expression levels of both the PCE Medea and PCE E-box construct are both low, the weak levels of expression are completely due to the E-Box included in both constructs and the missing activator binding site I mentioned before.

From these results for the PCE E-box and PCE Medea constructs, I conclude that the sequence of the 48 bp upstream of the starting point of the PCE S1 deletion construct is possibly involved in modulating PCE reporter gene expression, potentially

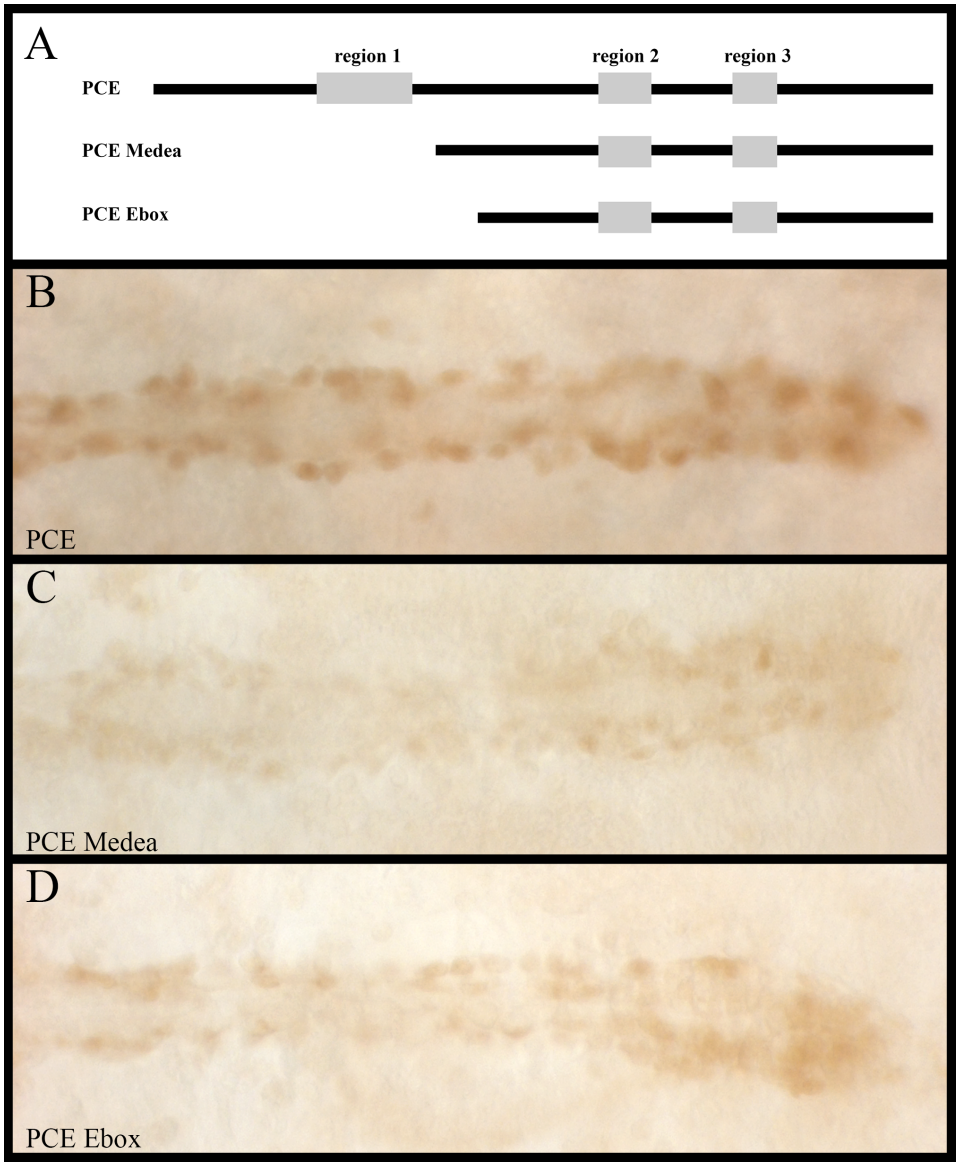


Figure 6.6.5: The PCE Medea and PCE E-box deletion constructs show a very faint reporter expression.
A Schematic representation of the reporter constructs used in this Figure. **B** Wild-type expression pattern of the full length PCE reporter construct in the heart. **C** *Drosophila* embryo transgenic for the PCE Medea deletion construct which deletes the 5' PCE sequence up to the 2 putative Medea binding sites in the unconserved sequence between the first and second conserved region. There is a faint but clear reporter expression visible. **D** The PCE E-box reporter construct deletes even more of the 5' PCE sequence up to the E-box in the unconserved sequence between the conserved regions 1 and 2. It also shows faint but distinct reporter activity.

making sure that the correct expression levels are reached. It is notable that even though the S5 Tinman binding site (see chapter 5) is included in all these constructs, this site alone is not capable of inducing reporter gene expression. This points, as previously discussed, to the sequence between the PCE Δreg1 and the PCE Δtin1 deletion constructs. With a high probability this region contains a second major activator involved in PCE and *Him* regulation. This activator will have a similar importance for PCE and *Him* expression as the S5 Tinman binding site.

6.8 Discussion

I have shown that, as expected, in addition to Tinman, several other signals are integrated through the PCE to regulate *Him* expression. These results are in agreement with the results from the Junion et al. (2012) paper who have found that a region upstream of the transcriptional start site of *Him* that contains the PCE fragment described in my study binds to all five transcription factors (pMad, dTCF, Dorsocross, Pannier, Tinman) they tested (supplementary information, Junion et al., 2012). Using a combined deletion and Linker Scan analysis I have been able to identify an area for the binding of a second activator. This second activator is of similar importance as Tinman. The area, in which this activator binds, is within the first conserved region and through my analysis I can already exclude several short stretches of sequence within this 53 bp region. Decapentaplegic signalling and especially Medea are very strong candidates for this second activator of *Him* expression. *decapentaplegic* expression from the ectoderm is known to have a regulatory influence on the activation of several mesodermal genes during embryonic development (Taylor et al., 1995; Riechmann et al., 1997; Yin et al., 1997) and the Eve-positive pericardial cells are formed underneath the ectodermal stripe with the highest Decapentaplegic expression (Dunin Borkowski et al., 1995).

As stated at the beginning of this chapter, *Him* expression requires a repressor in the cardioblasts as *tinman* is expressed in both, the pericardial cells and the cardioblasts, while *Him* is only expressed in the pericardial cells. My analysis of the PCE sequence has proven this hypothesis to be true and has, additionally, also revealed the presence of a second site involved in repressing *Him* expression in the cardioblasts and a further repressor responsible for preventing *Him* expression in the amnioserosa. Two candidate

regulators for the cardioblast repressor, Pannier and Scalloped, are, according to my results, not involved in excluding *Him* expression from the cardioblasts. However, the results from Deng et al. (2009) and Guss et al. (2013) indicate a role for *scalloped* in *Drosophila* mesoderm and heart development and thus I cannot completely rule out that Scalloped might be involved in the repression of *Him* expression in the cardioblasts. Investigations into whether the removal of Scalloped by RNAi has an effect on the *Him* expression pattern are ongoing.

Additionally I cannot rule out that a multi-protein-complex is necessary to limit *Him* expression in the cardioblasts as there are two repressor binding sites in close proximity to each other. It is possible to separate these two sites from one another by different enhancer constructs that delete or mutate different areas, thus implicating that both sites are equally and independently of each other necessary to exclude *Him* expression in the cardioblasts. As I have demonstrated in chapter 3, it is necessary to exclude Him from the cardioblasts to allow for the correct expression of Mef2-target genes like *myosin*.

The fact that *Him* expression is also actively repressed in the amnioserosa sheds light on which other potential positive regulators are necessary for *Him* expression as *tinman* is not expressed in the amnioserosa cells. It is thus necessary that any activating factor of *Him* expression is also expressed in the amnioserosa. This once again, furthers the possibility of Decapentaplegic signalling having an influence on *Him* expression. This factor is equally important for *Him* regulation as Tinman, as the removal of the 5' region of the PCE sequence does not show high levels of reporter expression (see PCE Ebox and Medea constructs), despite the intact S5 Tinman binding site being included in these experiments.

I have also found evidence for factors augmenting *Him* activation. Deletion of the 5' sequences of the PCE between the conserved regions 1 and 2 reduces PCE reporter expression considerably but does not abolish reporter activity completely. The most likely site for this factor is the Twist E-box located in the middle of the unconserved sequence between the first and second conserved regions. As this E-box sequence has been described to bind Twist (Halfon et al., 2000) and Twist is known to be an early activator of mesodermal genes and specifically a *tinman* activator (Riechmann et al., 1997), it is possible that these two genes act in concert to achieve the correct level of *Him* expression. I have been able to show that Hand, a bHLH protein with prominent expression in the developing heart from stage 12 onwards (Kölsch and

Paululat, 2002), is most likely not involved in the regulation of *Him* expression and the only other known bHLH protein expressed only in the *Drosophila* mesoderm is Twist.

Furthermore I have tried to investigate the importance of the E-box found within the Tinman binding consensus within the context of the PCE. My results allow for two possible interpretations; either the E-box within the Tinman consensus is of no importance for *Him* regulation or mutating only one base within the six basepair consensus is not sufficient to destroy its binding ability. In their 2010 paper, the modENCODE Consortium et al. report that a difference of one basepair within a binding site might be enough to differentiate between a Sna-motif in a HOT-spot and a Sna-bound region and in their 2011 paper, Ozdemir et al. show that a single basepair mutation in a Twist E-box is sufficient to reduce expression of the *rho* enhancer. If this were indeed the case for transcription factor binding at enhancers this result would indicate that the E-box contained within the Tinman binding consensus is of little importance for Tinman binding.

In all, the results presented in this chapter should allow for the identification of further regulators of *Him* expression, which will allow us to fine-tune the position of *Him* within the regulatory mechanisms of *Drosophila* heart development beyond it being a Tinman target gene.

Vector Map

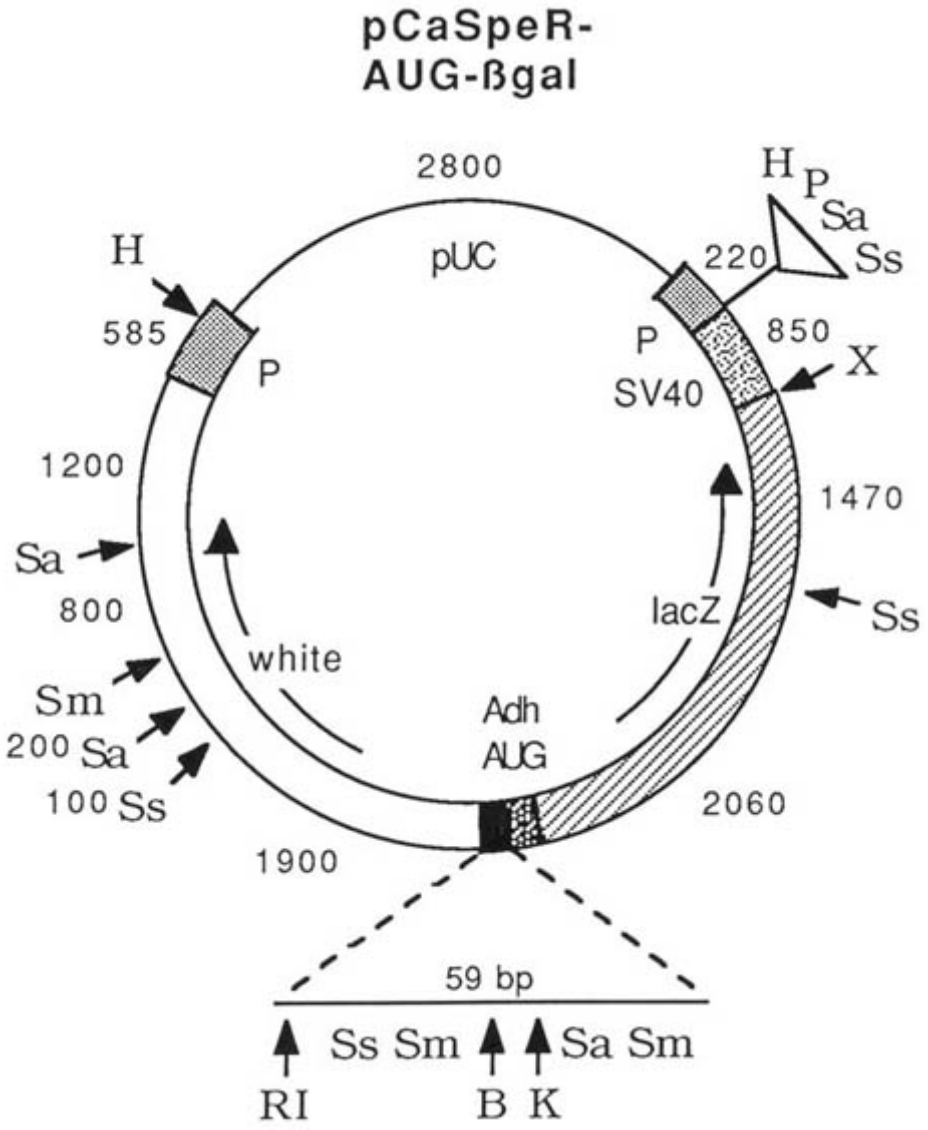


Figure A1: Vector map for the pCaSpeR-AUG-ßGal vector used to create transgenic fly lines (Thummel et al., 1988).

Embryonic expression pattern of Gal4 drivers

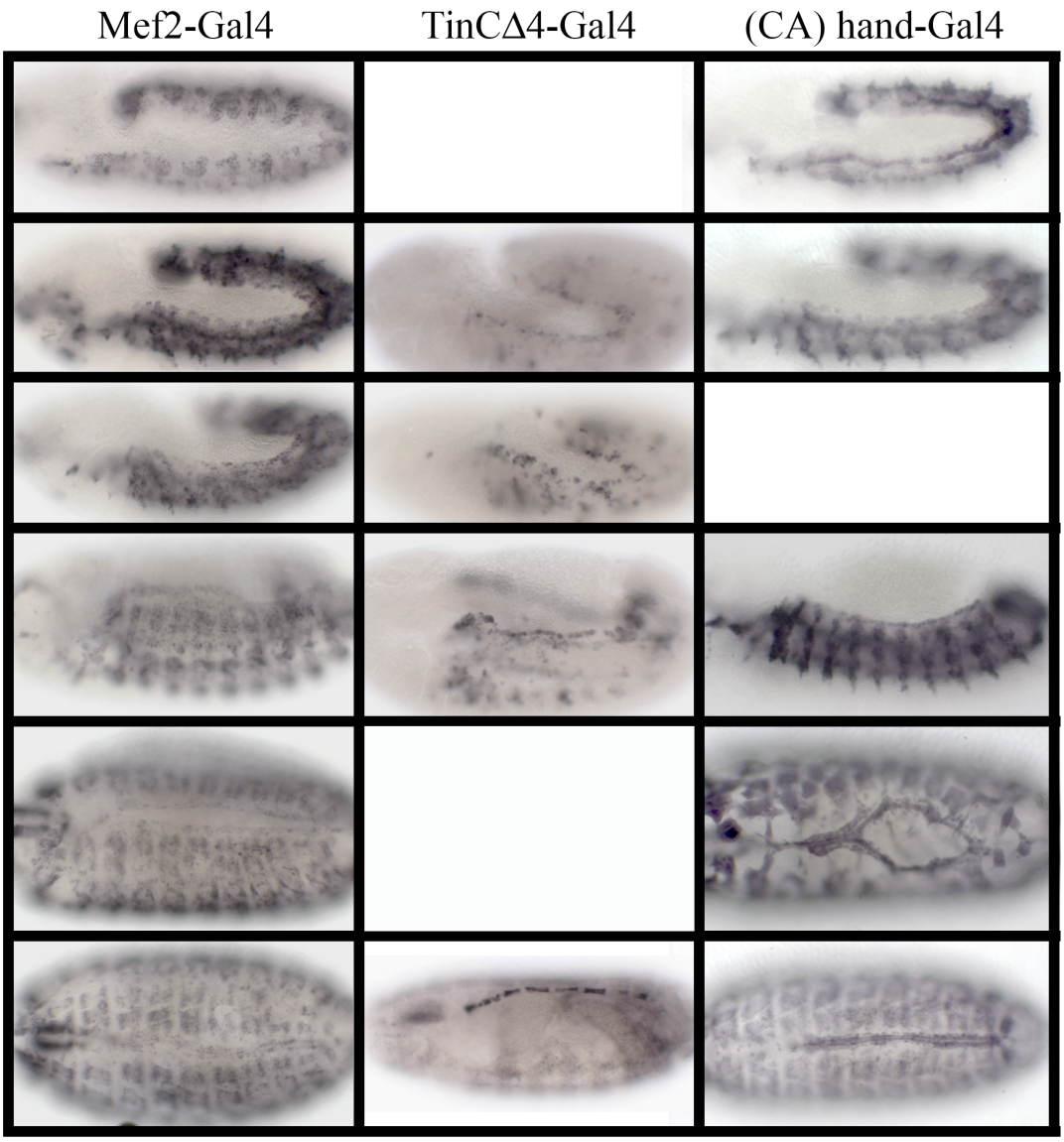


Figure A2: Embryonic expression pattern of the Mef2, hand, and TinCΔ4-Gal4 drivers used in this study.
These embryos are labelled for GFP expression as this driver test was done with an UAS-GFP construct.

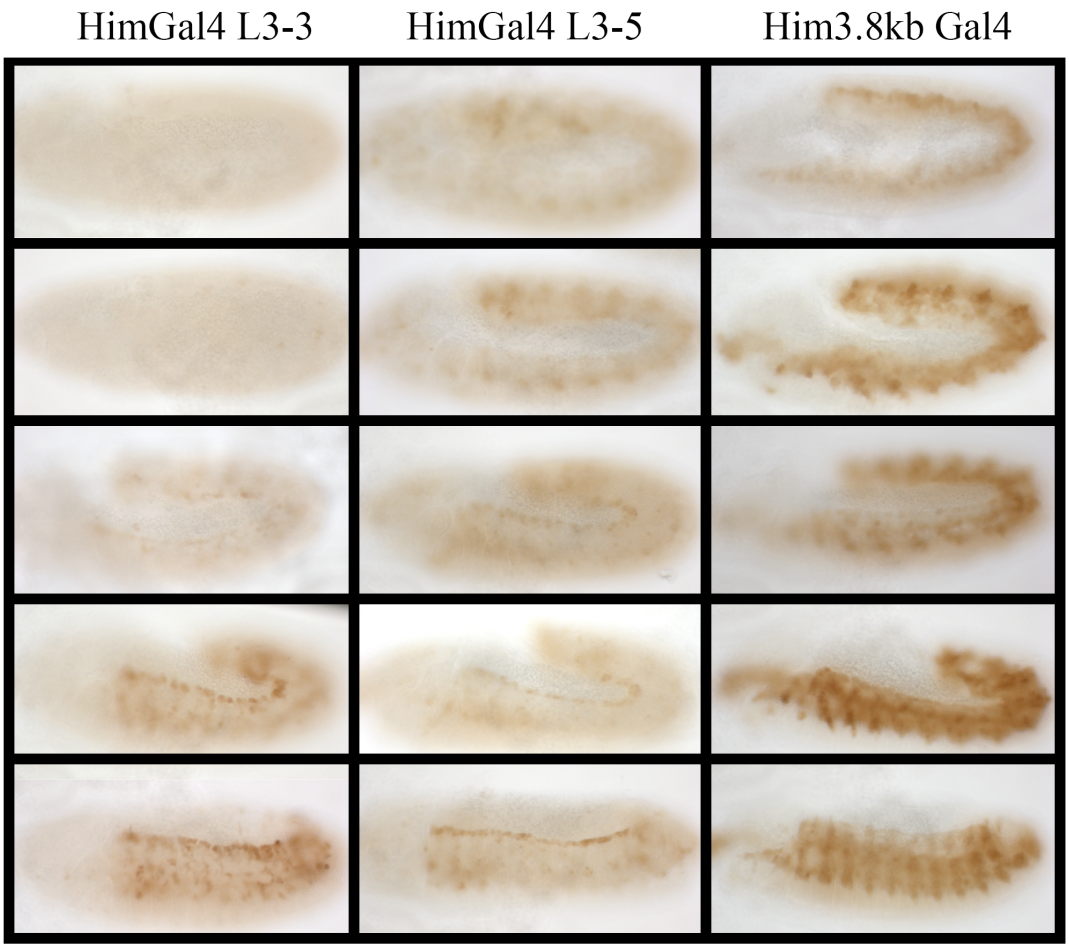


Figure A3: Embryonic expression pattern of the different Him-Gal4 driver lines available.
These embryos are labelled for GFP expression as this driver test was done with an UAS-GFP construct. I used the Him-Gal4 L3-5 throughout this study as it reproduces the expression pattern of the *Him* transcript best.

Expression pattern of the *lacZ* transcript of three PCE

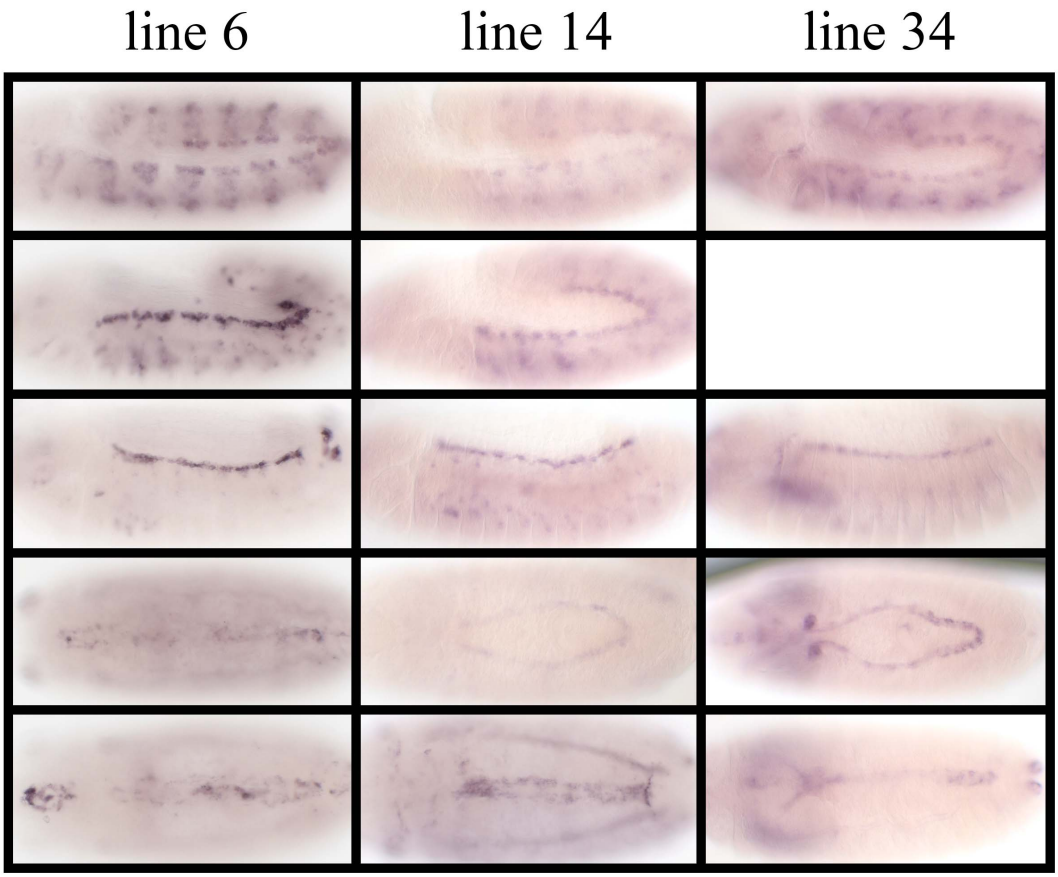


Figure A4: Three independent transgenic PCE fly lines show the same expression pattern for the *lacZ* transcript.
In-situ-hybridisation of the PCE lines 6, 14 and 34. All three constructs show thesame expression pattern of the *lacZ* transcript throughout development.

LacZ reporter expression in the PCE line 34

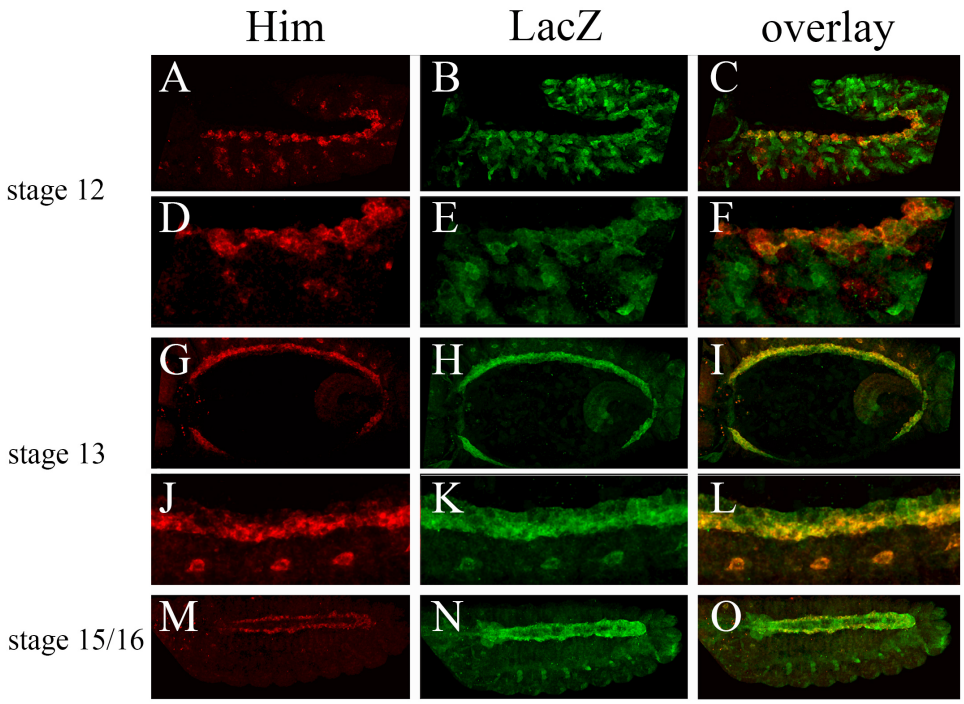


Figure A5: *Him* and the PCE reporter (line 34) are expressed in the same heart cells.
Him fluorescent RNA *in situ* and anti-βGalactosidase counterstain on embryos transgenic for the PCE construct (line 34). **A, D, G, J, M** fluorescent *Him* RNA *in situ* (red); **B, E, H, K, N** anti-βGalactosidase antibody (green); **C, F, I, L, O** overlay.

Manually compiled list of *Drosophila* mesoderm transcription factors (S. Elgar & M. Taylor, unpub. data)

Factor	**Sequence**
twist	CAKVTG
zhf-1a	CACCTG
zhf-1b	CTAATYRRNTT
sup.airless	RTGRGAR
daughterless	CACCTG
ETS	CGCGGAGC
Tinman	TCAAGTG
Mef2	YTAWWWWTAR
dTCF/pangolin	CTTTGA
Ladybird	YTAAYTAG
Eve1	TCAATTCAAT
Eve2	TCAGCACCG
Mad/Medea	GCCGNSB
Dorsal	GRGAAANCC
Glivertebrate	TGGGTGGTC
Lameduck	CGCAGNGTTT
Paired	GWCAGGS
Homeodomain/minimal	TTAAT
Homeodomain/full	TAATNNATTA
Forkhead/biniou	RTAAYA
Eve_high	ACATTAAAGT
Eve_med	WCAKYWAMDY
Ebox	CANNTG
Krupple_meme_high	AACGGGTTAA
Krupple_meme_med	AMSGGRTTAW
Krupple_matrix_high	AACGGGTTAA
Krupple_matrix_med	AMYGGGTTAW
Su(H)_matrix_med	YGTGRGAAM
Su(H)_matrix_low	GTGRGAA
Dorsal_matrix1_ungapped	GGHHWTTMC
Dorsal_matrix1_gapped	GGHHTTMC
Dorsal_meme_high	GGGAAAACCC
Dorsal_meme_med	GGGAAAAYCC
Pannier	GATA
Bagpipe	TCAAGTKC
Paired_High	CGATTAAG
Paired_Med	CRRTKMHR
Tinman	TCAAGTG
Biniou_high	RTAAAYAAA
Biniou_low	TAAAYA
d/TCF_high	CCTTTGATCTT
d/TCF_med1	TTTGAT
d/TCF_med2	CTTTGA

Mad_high	GCCGNCG
Scalloped_high	TGGAATGT
Scalloped_med	WVDWATKY
Brinker	GGCGYY
Tinman	TCAAGTG
Mef2	YTAWWWWTAR
Mad1	GGCGGC
Mad2	GTCGTTGG
Mad3	GCGGGGGC
PanA	CATCAA
Tin1	CAATTAA
FD	TAAACA
MedA	GTCT
Mad4	GGTTCGGCCGCAGAT
MedB	GCCGC
Tin2	CCAAGTG
MedC	GGCGGGC
PanB	ATCAAAG
Mad6	GCGACG
TinA	ACTTCAC
PanC	TTCACAG
TinB	CACTTAA
PanD	CACTTAA
MedD	TCTG
MedE	GCCGG
TinC	CCCTTGA
MedF	GGCGGCC
Zfh1	TAAT
PanE	CCTTTGGAT
TinD	CACTTGA
PanF	TCCTTTAAT
PanG	GATCTAAT

Buffer and Solutions

AP buffer

100 mM	Tris (ph 9.5)
100 mM	NaCl
50 mM	$MgCl_2$
0.1 %	Tween-20

Apple juice agar plates for embryo collection (2 l)

24 g	agar (Brian Drewitt)
	in 1 l dH_2O and boil in microwave until agar is dissolved
24 g	sucrose
	in 500 ml dH_2O and heat until dissolved, then add
500 ml	apple juice
	dissolve 2 g Nipagen (p-Hydroxybenzoic acid methyl ester, Sigma)
	in 10 ml ethanol
	mix everything together and heat to just under boiling point
	pour into 55 mm petridishes (Fisherbrand)
	leave to set and then store at 4 °C

Blocking solution for fluorescent antibody stain

1 x	PBT
2 %	goat serum (Sigma)

Blocking solution for RNA probe quantification

10 %	blocking reagent (w/v; Roche Applied Science) in maleic acid buffer
	autoclave, do not filter

EMSA binding buffer (5 x; Kremser et al., 1999)

20 mM	Tris/Cl (pH 7.5)
200 mM	NaCl
0.1 %	Trition X-100
5 mM	DTE
0.02 %	BSA (w/v)
5 %	glycerol (w/v)

Fluorescent Mounting Medium

90 %	glycerol in PBS
25 mg/ml	DABCO (1,4-Diazabicyclo[2.2.2]octane)

Fly Food Recipe (60 l)

42 l	dH_2O
504 g	agar (Brian Drewitt)
	start heating and stir frequently
135 g	Nipagen (Hydroxybenzoic acid methyl ester; Sigma)
1.529 l	analytical grade ethanol
	dissolve and then add to the heating agar mix

4.0232 kg Maize (organic, Dove's Farm)
1048 g yeast
4.32 kg dextrose
12 l dH_20
dissolve and ensure everything is mixed well
when the agar mixture reaches ca. 70 °C add the maize-yeast-dextrose mix.
Heat until the temperature reaches 85 °C while stirring frequently, then add
4.272 l dH_2O (at room temperature)
wait until the temperature cools down to 68 °C or less, then add
210 ml proprionic acid
mix well.
Dispense food.

Hybridisation buffer for whole mount in-situ hybridisation

50 % Formamide
4 x SSC
1 x Denhardt's
250 µg/ml tRNA
250 µg/ml salmon testis DNA
50 µg/ml heparin
0.1 % Tween-20

Hybridization buffer for whole mount FISH

50 ml Formamide
25 ml 20x SSC
0.1 ml heparin (50mg/ml)
1 ml 10 % TritonX-100
bring to 100 ml with DEPC treated H_2O

Maleic acid buffer

100 mM maleic acid (1 M stock)
150 mM NaCl (1 M stock)
adjust to pH 7.5 with solid NaOH
filter before adding blocking reagent

LB agar

Prepare LB medium and add
15 g Bacto Agar per liter
sterilize by autoclaving

LB medium (Luria-Bertani Medium; 1 l)

950 ml dH2O
10 g tryptone
5 g yeast extract
10 g NaCl
dissolve
adjust to pH 8 with 5 N NaOH
adjust volume to 1l sterilize by autoclaving

Loading buffer for agarose gels

5 % (w/v)	Xylencyanol
5 % (w/v)	Bromphenolblue
40 % (v/v)	glycerol
	in 1 x TAE buffer

Loading buffer for SDS gels

100 mM	Tris-Cl (pH 6.8)
4 % (w/v)	SDS
0.2 % (w/v)	bromophenol blue
20 % (v/v)	glycerol
200 mM	ß-mercaptoethanol

PBS (phosphate buffered saline; 1 x)

140 mM	NaCl
6.5 mM	KCl
2.5 mM	Na_2HPO_4
1.5 mM	KH_2PO_4
	adjust to pH 7.5 and autoclave

PBT (1 x)

1 x	PBS
0.1 %	Tween-20

PBS-TX (1 x)

1 x	PBS
0.1 %	Triton X-100

Paraformaldehyde (4 %; in PBS)

2.142 g	Pearl NaOH
85 ml	dH_2O
	dissolve
20 g	Paraformaldehyde
	add dissolved NaOH
	heat to a max of 60°C
	allow to cool
10.58 g	$NaH_2PO_4x2H_2O$
400 ml	dH2O
	dissolve
	add to cooled Paraformaldehyde solution
	adjust to pH 7 with NaOH if necessary
	store in aliquots at -20°C

Ribofix solution for embryo fixing (2.5ml)

250 µl	10 x PBS
250 µl	0.5 M EGTA (pH 8.0)
600 µl	dH_2O
1.4 ml	16 % formaldehyde, methanol-free (Polysciences)

Spradling buffer (1x)

1mM	Sodium Phosphate buffer pH 6.8
0.5mM	KCl

SOB medium (1 l)

950 ml	dH_2O
20 g	tryptone
5 g	yeast extract
0.5 g	NaCl
	dissolve
10 ml	250 mM KCl
	adjust to pH 7.0 with 5 N NaOH
	adjust total volume to 1 l
	sterilize by autoclaving
	then add
5 ml	2 M $MgCl_2$ (sterile)

SOC medium (1 l)

950 ml	dH_2O
20 g	tryptone
5 g	yeast extract
0.5 g	NaCl
	dissolve
10 ml	250 mM KCl
	adjust to pH 7.0 with 5 N NaOH
	adjust total volume to 1 l
	sterilize by autoclaving
	allow to cool below 60 °C, then add
20 ml	1 M glucose (sterile)
5 ml	2 M $MgCl_2$ (sterile)

STET-buffer

10 mM	Tris (pH 8.0)
25 mM	EDTA
100 mM	NaCl
0.1 %	Triton X-100
8 %	sucrose
	sterilize by autoclaving or filtering

TBE (500 ml, 10 x)

54 g	Tris Base
27.5 g	Boric Acid
20 ml	0.5 M EDTA (pH 8.0)
	bring to 500 ml with dH_2O

TE buffer

10 parts	Tris/HCl (pH 8.0)
0.1 parts	EDTA (pH 8.0)

TFB1 buffer

100 mM	$RbCl_2$
50 mM	$MnCl_2 \cdot 4H_20$
30 mM	KAc
10 mM	$CaCl_2 \cdot 2H_20$
5 %	glycerol

filter sterilize, store at 4 °C in a sterile bottle

TFB2 buffer

10 mM	MOPS
10 mM	$RbCl_2$
75 mM	$CaCl_2 \cdot 2H_20$
15 %	glycerol

adjust to pH 6.8 – 7.0
filter sterilize, store at 4 °C in a sterile bottle

TNE buffer (oligo annealing)

10 mM	Tris (pH 8.0)
100 mM	NaCl
1 mM	EDTA

TXN

0.7 %	NaCal
0.04 %	TritonX-100

in dH_2O

Wash buffer for RNA probe quantification

0.1 M	NaCl
0.05 M	$MgCl_2$
0.1 M	Tris pH 9.5
0.1 %	Tween-20

Wash buffer for whole mount in-situ hybridisation

50 %	Formamide
2 x	SSC
0.1 %	Tween-20

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RD, Scheeler F, Shen H, Shue BC, Sidén-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, WoodageT, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM, Venter JC. (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185-95.

Abmeyr, M., Wisniewska, M. and Weitzmann, J.B. (2003). A role for AP-1 in apoptosis: the case for and against. *Biochimie* **85**, 747-52.

Affolter, M., Mary, T., Vigano, M.A. and Jazwinska, A. (2001). Nuclear interpretation of Dpp signaling in Drosophila. *EMBO J* **20**, 3298-305.

Akasaka, T., Klinedinst, S., Ocorr, K., Bustamante, E.L., Kim, S.K. and Bodmer, R. (2006). The ATP-sensitive potassium (KATP) channel-encoded dSUR gene is required for *Drosophila* heart function and is regulated by *tinman*. *Proc Natl Acad Sci USA* **103**, 11999-2004.

Alvarez, A.D., Shi, W., Wilson, B.A. and Skeath, J.B. (2003). Pannier and pointedP2 act sequentially to regulate Drosophila heart development. *Dev* **120**, 3015-3026.

Arora, K. and Nüsslein-Volhard, C. (1992). Altered mitotic domains reveal fate map changes in *Drosophila* embryos mutant for zygotic dorsoventral patterning genes. *Development* **114**, 1003-24.

Atchley, W.R. and Fitch, W.M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proc Natl Acad Sci USA* **94**, 5172-76.

Azpiazu, N. and Frasch, M. (1993). Tinman and bagpipe: two homeobox genes that determine cell fates in the dorsal mesoderm of *Drosophila*. *Genes Dev* **7**, 5172-76.

Bartlett, H., Veenstra, G.J. and Weeks, D.L. (2010). Examining the cardiac NK-2 genes in early heart development. *Pediatr Cardiol*. **31**, 335-41.

Bate, M., Landgraf, M. and Ruiz Gomez Bate, M. (1999). Tinman and bagpipe, two homeobox genes that determine cell fates in the dorsal mesoderm of Drosophila. *Genes Dev* **10**, 3183-94.

Baylies, M.K., Bate, M. and Gomez, M.R. (1998). Myogenesis: A view from *Drosophila*. *Cell* **93**, 921-27.

Behrens, A.N., Iacovino, M., Lohr, J.L., Ren, Y., Zierold, C., Harvey, R.P., Kyba, M., Garry, D.J. and Martin, C.M. (2013). Nkx2-5 mediates differential cardiac differentiation through interaction with Hoxa10. *Stem Cells Dev* [Epub ahead of print].

Bergman, C.M. and Kreitman, M. (2001). Analysis of conserved noncoding DNA in Drosophila reveals similar constraints in intergenic and intronic sequences. *Genome Research* **11**, 1335-45.

Biggin, M.D. (2011). Animal transcription networks as highly connected, quantitative continua. *Dev Cell* **21**, 611-26.

Bird, C.P., Stranger, B.E. and Dermitzakis, E.T. (2006). Functional variation and evolution of non-coding DNA. *Curr. Opin. Genet. Dev.* **16**, 559-64.

Blackwell, T.K., Huang, J., Ma, L., Kretzner, F.W., Alt, R.N., Eisenmann, R.N. and Weintraub, H. (1993). Binding of Myc proteins to canonical and noncanonical DNA sequences. *Mol. Cell Biol.* **13**, 5216-24.

Bodmer, R. (1993). The gene Tinman is required for specification of the heart and visceral muscles in *Drosophila*. *Dev* **118**, 719-29.

Bodmer, R. (1995). Heart development in *Drosophila* and its relationship to vertebrates. *TMC* **5**, 21-28.

Bodmer, R. and Frasch, M. (1999). Genetic determination of Drosophila heart development. In: *Heart Development* (R.P. Harvey and Rosenthal, Eds.), Academic Press, San Diego, 65-90.

Bodmer, R. and Venkatesh, T.V. (1998). Heart development in *Drosophila* and Vertebrates: conservatsion of molecular mechanisms. *Dev Gen* **22**, 181-86.

Bodmer, R., Jan, L.Y. and Jan, Y.N. (1990). A new homeobox-containing gene, msh-2, is transiently expressed early during mesoderm formation of Drosophila. *Development* **110**, 661-69.

Boulay, J.L., Dennefeld, C. and Alberga, A. (1987). The *Drosophila* developmental gene snail encodes a protein with nucleic acid binding fingers. *Nature* **330**, 395-398.

Bour, B.A., O'Brien, M.A., Lockwood, W.L., Goldstein, E.S., Bodmer, R., Taghert, P.H., Abmayr, S.M. and Nguyen, H.T. (1995). Drosophila MEF2, a transcription factor that is essential for myogenesis. *Genes Dev* **9**, 730-41.

Brand, A.H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Dev.* **118**, 401-15.

Brand, N.J. (1997). Myocyte enhancer factor 2 (MEF2). *Int J Biochem Cell Biol* **29**, 1467-70.

Brantjes, H., Roose, J., van de Wetering, M. and Clevers, H. (2001). All Tcf HMG box transcription factors interact with Groucho-related co-repressors. *Nuc. Acids Res.* **29**, 1420-19.

Brantjes, H., Roose, J., van de Wetering, M. and Clevers, H. (2001). All Tcf HMG box transcription factors interact with Groucho-related co-repressors. *Nuc Acids Res* **29**, 1410-19.

Bray, N., Dubchak, I. and Pachter, L. (2003). AVID: A global alignment program. *Genome Res* **13**, 97-102.

Bray, S. (1999). Drosophila development: Scalloped and Vestigal take wing. *Curr. Biol.* **9**, R245-R247.

Brodu, V., Mugat, B., Roignant, J-Y., Lepesant, J-A. and Antoniewski, C. (1999). Dual requirement for the EcR/USP nuclear receptor and the dGATAb factor in an ecdysone response in *Drosophila melanogaster*. *Mol. Cell. Biol.* **19**, 5732-42.

Bruneau, B.G. (2008). The developmental genetics of congenital heart disease. *Nature* **451**, 943-48.

Bruneau, B.G., Nemer, G., Schitt, J.P., Charron, F., Robitaille, L., Caron, S., Conner, D.A., Gessler, M., Nemer, M., Seidman, C.E. and Seidman, J.G. (2001). A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell* **106**, 709-21.

Cammarato, A., Ahrens, C.H., Alayari, N.N., Qeli, E., Rucker, J., Reedy, M.C., Zmasek, C.M., Gucek, M., Cole, R.N., Van Eyk, J.E., Bodmer, R., O'Rourke, B., Bernstein, S.I., Foster, D.B. (2011). A mighty small heart: the cardiac proteome of adult *Drosophila melanogaster*. *PLoS One* **6**, e18497.

Cammarato, A., Ahrens, C.H., Alayari, N.N., Qeli, E., Rucker, J., Reedy, M.C., Zmasek, C.M., Gucek, M., Cole, R.N., Van Eyk, J.E., Bodmer, R., O'Rourke, B., Bernstein, S.I. and Foster, D.B. (2011) A mighty small heart: the cardiac proteome of adult *Drosophila melanogaster*. *PLoS one* **6**, e18497.

Campbell, S., Inamdar, M., Rodrigues, V., Vijay Raghavan, K., Palazzolo, M.J. and Chovnick, A. (1992). The scalloped gene encodes a novel, evolutionary conserved transcription factor required for sensory organ differentiation in Drosophila. *Genes Dev*. **6**, 367-379.

Campbell, S.D., Duttaroy, A., Katzen, A.L. and Chovnick, A. (1991). Cloning and characterization of the scalloped region of Drosophila melanogaster. *Genetics* **127**, 367-380.

Campos-Ortega, J.A. and Hartenstein, V. (1985). The Embryonic Development of Drosophila melanogaster. *Berlin: Springer Verlag*.

Cantor, A.B. and Orkin, S.H. (2004). Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. *Semin Cell Dev Biol.* **16**, 117-128.

Carmena, A., Buff, E., Halfon, M.S., Gisselbrecht, S., Jiménez, F., Baylies, M.K. and Michelson, A.M. (2002) Reciprocal regulatory interactions between the Notch and Ras signaling pathways in the Drosophila embryonic mesoderm. *Dev Biol* **244**, 226-242.

Carmena, A., Murugasu-Oei, B., Menon, D., Jiménez, F. and Chia, W. (1998). Inscutable and numb mediate asymmetric muscle progenitor cell divisions during *Drosophila* myogenesis. *Genes & Dev* **12**, 304-315.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. and Werner, T. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**, 2933-42.

Casal, J. and Leptin, M. (1996). Identification of novel genes in Drosophila reveals the complex regulation of early gene activity in the mesoderm. *Proc. Natl. Acad. Sci. USA* **93**, 10327-32.

Cavallo, R.A., Cox, R.T., Moline, M.M.; Roose, J., Polevoy, G.A., Clevers, H., Peifer, M. and Brjsovec, A. (1998). Drosophila Tcf and Groucho interact to repress Wingless signaling activity. *Nature* **395**, 604-8.

Celniker SE, Wheeler DA, Kronmiller B, Carlson JW, Halpern A, Patel S, Adams M, Champe M, Dugan SP, Frise E, Hodgson A, George RA, Hoskins RA, Laverty T, Muzny DM, Nelson CR, Pacleb JM, Park S, Pfeiffer BD, Richards S, Sodergren EJ, Svirskas R, Tabor PE, Wan K, Stapleton M, Sutton GG, Venter C,

Celniker SE, Wheeler DA, Kronmiller B, Carlson JW, Halpern A, Patel S, Adams M, Champe M, Dugan SP, Frise E, Hodgson A, George RA, Hoskins RA, Laverty T, Muzny DM, Nelson CR, Pacleb JM, Park S, Pfeiffer BD, Richards S, Sodergren EJ, Svirskas R, Tabor PE, Wan K, Stapleton M, Sutton GG, Venter C, Weinstock G, Scherer SE, Myers EW, Gibbs RA, Rubin GM. (2002). Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence. *Genome Biol* **3**, RESEARCH0079

Celniker, S.E., Dillon, L.A.L., Gerstein, M.B., Gunsalus, K.C., Henikoff, S., Karpen, G.H., Kellis, M., Lai, E.C., Lieb, J.D., MacAlpine, D.M., Mickelm, G., Paino, F., Snyder, M., Stein, L., White, K.P. and Waterston, R.H., for the modENCODE Consortium (2009). Unlocking the secrets of the genome. *Nature* **459**, 927-30.

Chartier, A., Zaffran, S., Astier, M., Sémériva, M. and Gratecos, D. (2002). Pericardin, a Drosophila type IV collagen-like protein is involved in the morphogenesis and maintenance of the heart epithelium during dorsal ectoderm closure. *Development* **129**, 3241-53.

Chasen, R., Anderson, K.V. (1993). Maternal control of dorsal-ventral polarity and pattern in the embryo. In: *The development of Drosophila melanogaster*, Vol 1, M. Bate and A. Martinez-Arias, eds. Cold Spring Harbor, NY: CSH Laboratory Press, pp. 387-424.

Chen, C.Y. and Schwartz, R.J. (1995). Identification of Novel DNA Binding Targets and Regulatory Domains of a Murine Tinman Homeodomain Factor, *nkx-2.5**. *JBC* **270**, 15628-33.

Chen, J.N. and Fishman, M.C. (2000). Genetics of heart development. *Trends Genet* **16**, 383-8

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res*. **31**, 3497-500.

Choi, C.Y., Lee, Y.M., Young, H.K., Park, T., Jeon, B.H., Schulz, R.A. and Kim, Y. (1999). The homeodomain transcription factor NK-4 acts as either a transcriptional activator or repressor and interacts with the p300 coactivator and the Groucho corepressor. *JBC* **274**, 31543-52.

Christoffels, V.M., Habets, P.E., Franco, D., Campione, M., de Jong, F., Lamers, W.H., Bao, Z.Z., Palmer, S., Biben, C., Harvey, R.P. and Moorman, A.F. (2000). Chamber formation and morphogenesis in the developing mammalian heart. *Dev Biol* **223**, 266-78.

Cripps, R.M. and Olson, E.N. (2002) Control of cardiac development by an evolutionary conserved transcriptional network. *Dev Biol*. **246**, 14-28.

Cripps, R.M. and Olson, E.N. (2002). Control of cardiac development by an evolutionary conserved transcriptional network. *Dev Biol*. **246**, 14-28.

Crossley, A.C. (1972). The ultrastructure and function of pericardial cells and other nephrocytes in an insect: *Calliphora erythrocephala*. *Tissue Cell* **4**, 529-60.

Cubadda, Y., Heitzler, P., Ray, R.P., Bourouis, M., Ramain, P., Gelbart, W., Simpson, P. and Haenlin, M. (1997). U-shaped encodes a zinc finger protein that regulates the proneural genes achaete and scute during the formation of bristles in Drosophila. *Genes Dev*. **11**, 3083-95.

Curtis, N.J., Ringo, J.M. and Dowse, H.B. (1999). Morphology of the pupal heart, adult heart, and associated tissues in the fruit fly, *Drosophila melanogaster*. *J Morphol* **240,** 225-35.

Dang, C.V., Dolde, C., Gillison, M.L. and Kato, G.J. (1992). Discrimination

Das, D., Ashoka, D., Rajaguru, A. and Inamdar, M. (2008). Gene expression analysis in post-embryonic pericardial cells of Drosophila. *Gene Expr Patt*. **8**, 199-205.
Deng, H., Hughes, S.C., Bell, J.B and Simmonds, A.J. (2009). Alternative requirements for Vestigal, Scalloped, and Dmef2 during muscle differentiation in *Drosophila melanogaster*. *MBC* **20**, 256-69.
Denholm, B. and Skaer, H. (2009). Bringing together components of the fly renal system. *Curr Opin Genet Dev* **19**, 526-32.
Dermitzakis, E.T., Bergman, C.M and Clark, A.G. (2003). Tracing the evolutionary history of *Drosophila* regulatory regions with models that identify transcription factor binding sites. *Mol Biol Evol* **20**, 703-714.
DeVelasco, B., Shen, J., Go, S. and Hartenstein, V. (2004). Embryonic development of the *Drosophila* corpus cardiacum, a neuroendocrine gland with similarity to the vertebrate pituitary, is controlled by sine oculis and glass. *Dev. Biol*. **274**, 280-94.
Dickmeis, T. and Müller, F. (2004) The identification and functional characterization of conserved regulatory elements in developmental genes. *Briefings in Functional Genomics and Proteomics* **3 (4)**, 332-350.
Dilks, S.A. and DiNardo, S. (2010). Non-cell-autonomous control of denticle diversity in the *Drosophila* embryo. *Dev* **137**, 1395-1404.
Drosophila 12 Genomes Consortium (2007). Evolution of genes and genomes on the Drosophila phylogeny. *Nature*. **450(7167)**, 203-18.
Dubchak, I., Brundo, M., Loots, G.G., Pachter, L., Mayor, C., Rubin, E.M. and Fraser, K.A. (2000). Active conservation of non-coding sequences revealed by three-way species comparisions. *Genome Res* **10**, 1304-06.
Dubchak, I., Poliakov, A., Kislyuk, A. and Brudno, M. (2009). Multiple whole-genome alignments without a reference organism. *Genome Res* **19**, 682-29.
Dunin Borokowski, O.M., Brown, N.H. and Bate, M. (1995). Anterior-posterior subdivision and the diversification of the mesoderm in *Drosophila*. *Dev* **121**, 4183-93.
Durocher, D., Charron, F., Warren, R., Schwartz, R.J. and Nemer, M. (1997). The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *EMBO J*. **16**, 5687-96.
Durocher, D., Chen, C.Y., Ardati, A., Schwartz, R.J. and Nemer, M. (1996). The atrial natriuretic factor promoter is a downstream target for Nkx-2.5 in the myocardium. *MCB* **16**, 4648-4655.
Dutta, D., Bloor, J.W., Ruiz-Gómez, M., VijayRaghavan, K. and Kiehart, D.P. (2002). Real-Time imaging of morphogenetic movements in *Drosophila* using Gal4-UAS-driven expression of GFP fused to the Actin-binding domain of Moesin. *Genesis* **34**, 146-51.
Eivers, E., Fuentealba, L.C., Sander, V., Clemens, J.C., Hartnett, L. et al. (2009) Mad Is Required for Wingless signaling in wing development and segment patterning in *Drosophila*. *PLoS ONE* **4**, e6543.
Elgar, S.J., Han, J. and Taylor, M.V. (2008). Mef2 activity levels differentially affect gene expression during Drosophila muscle development. *PNAS* **105**, 918-23.
Ellenberger, T., Fass, D., Arnaud, M. and Harrsion, S.C. (1994). Crystal structure of the transcription factor E47: E-box recognition by a basic helix-loop-helix dimer. *Genes Dev* **8**, 970-80.
Evans, C.J., Hartenstain, V. and Banerjee, U. (2003). Thicker than blood: conserved mechanisms in *Drosophila* and vertebrate hematopoiesis. *Dev. Cell* **5**, 673-690.
Ferré-D'Amaré, A.R., Prendergast, G.C., Ziff, E.B. and Burley, S.K. (1993). Recognition by Max of its cognate DNA through a dimeric b-HLH-Z domain. *Nature* **363**, 38-45.

Fossett, N., Tevosian, S.G., Gajeswski, K., Zhang, Q., Orkin, S.H. and Schulz, R.A. (2001). The Friend of GATA proteins U-shaped, FOG-1, and FOG-2 function as negative regulators of blood, heart and eye development in *Drosophila*. *PNAS* **98**, 7342-47

Fossett, N., Zhang, Q., Gajewski, K., Choi, C., Kim, Y. and Schulz, R.A. (2000). The multitype zinc-finger protein U-shaped functions in heart cell specification in the Drosophila embryo. *Proc. Natl. Acad. Sci. USA* **97**, 7348-53.

Frank, L.H. and Rushlow, C. (1996). A group of genes required for maintenance of the amnioserosa tissue in Drosophila. *Development* **122**, 1343-52.

Frasch, M. (1995). Induction of visceral and cardiac mesoderm by ectodermal Dpp in the early Drosophila embryo. *Nature* **374**, 464-67.

Frasch, M. (1999). Intersecting signaling and transcriptional pathways in *Drosophila* heart specification. *Cell & Developmental Biology* **10**, 61-71.

Frasch, M. and Levine, M. (1987). Complementary patterns of even-skipped and fushi-tarazu expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. *Genes Dev*. **1**, 981-995.

Fu, Y., Yan, W., Mohun, T.J. and Evans, S.M. (1998). Vertebrate *tinman* homologues XNkx2-3 and XNkx2-5 are required for heart formation in a functionally redundant manner. *Dev* **125**, 4439-49.

Furlong, E.E.M., Andersen, E.C., Null, B., White, K.P. and Scott, M.P. (2001). Patterns of gene expression during Drosophila mesoderm development. *Science* **293**, 1629-33.

Gajewski, K., Fossett, N., Molkentin, J.D. and Schulz, R.A. (1999). The zinc finger proteins Pannier and GATA4 fucntion as cardiogenic factors in Drosophila. *Dev* **126**, 4579-5688.

Gajewski, K., Kim, Y., Choi, C.Y. and Schulz, R.A. (1998). Combinatorial control of *Drosophila mef2* gene expression in cardiac and somatic muscle cell lineages. *Dev Genes Evol*. **208**, 382-92.

Gajewski, K., Kim, Y., Lee, Y.M., Olson, E.N. and Schulz, R.A. (1997). D-mef2 is a target for Tinman activation during *Drosophila* heart development. *EMBO J* **16 (3)**, 515-22.

Gajewski, K., Zhang, Q., Choi, C.Y., Fossett, N., Dang, A., Kim, Y.H., Kim, Y. and Schulz, R.A. (2001). Pannier is a transcriptional target and partner of Tinman during *Drosophila* cardiogenesis. *Dev Biol*. **233**, 425-36.

Gasch, A., Hinz, U., Leiss, D. and Renkawitz-Pohl, R. (1988). The expression of -1 and -3 tubulin genes of *Drosophila melanogaster* is spatially regulated during embryongenesis. *Mol. Gen. Genet*. **211**, 8-16.

Gaumer, S., Guénal, I., Brun, S., Théodore, L. and Mignotte, B. (2000). Bcl-2 and Bax mammalian regulators of apoptosis are functional in *Drosophila*. *Cell Death Diff* **7**, 804-14.

Guner-Ataman, B., Paffett-Lugassy, N., Adams, M.S., Nevis, K.R., Jahangiri, L., Obregon, P., Kikuchi, K., Poss, K.D., Burns, C.E. and Burns, C.G. (2013). Zebrafish second heart field development relies on progenitor specification in anterior lateral plate mesoderm and *nkx2.5* function. *Dev* **140**, 1353-63.

Gunthrope, D., Beatty, K.E. and Taylor, M.V. (1999). Different levels, but not different isoforms, of the *Drosophila* transcription factor DMEF2 affect distinct aspects of muscle differentiation. *Dev Biol*. **215**, 130-145.

Guss, K.A., Benson, M., Gubitosi, N., Brondell, K., Broadie, K. and Skeath, J.B. (2013). Expression and function of Scalloped during *Drosophila* development. *Dev Dyn*, DOI 10.1002/dvdy.23942

Haenlin, M., Cubadda, Y., Blondeau, F., Heitzler, P., Lutz, Y., Simpson, P. and Ramain, P. (1997). Transcriptional activity of Pannier is regulated negatively by heterodimerization of the GATA DNA-binding domain with a cofactor encoded by the *u-shaped* gene of *Drosophila*. Genes Dev **11**, 3096-108.

Halder, G. and Carroll, S.B. (2001). Binding of the Vestigial co-factor switches the DNA-target selectivity of the Scalloped selector protein. *Dev* **128**, 3295-305.

Halder, G., Polaczyk, P., Kraus, M.E., Hudson, A., Kim, J., Laughon, A. and Carroll, S. (1998). The Vestigial and Scalloped proteins act together to directly regulate wing-specific gene expression in *Drosophila*. *Genes Dev* **12**, 3900-09.

Halfon, M.S., Carmena, A., Gisselbrecht, S., Sackerson, C.M., Jimenez, F., Baylies, M.K. and Michelson, A.M. (2000). Ras pathway specificity is determined by the integration of multiple signal-activated and tissue-restricted transcription factors. *Cell* **103**, 63-74.

Halfon, M.S., Grad, Y., Church, G.M. and Michelson, A.M. (2002). Computation-based discovery of related transcriptional regulatory modules and motifs using an experimentally validated combinatorial model. *Genome Res* **12**, 1019-28.

Halfon, M.S., Zhu, Q., Brennan, E.R. and Zhou, Y. (2011). Erroneous attribution of relevant transcription factor binding sites despite successful prediction of cis-regulatory modules. '**12**, 578-87.

Hamaguchi, T., Yabe, S., Uchiyama, H. and Murakami, R. (2004). *Drosophila* Tbx6-related gene, *Dorsocross*, mediates high levels of Dpp and Scw signal required for the development of amnioserosa and wing disc primordium. *Dev Biol* **265**, 355-68.

Han, Z. and Bodmer, R. (2003). Myogenic cells fates are antagonized by Notch only in asymmetric lineages of the *Drosophila* heart, with or without cell division. *Dev* **120**, 3039-51.

Han, Z. and Olson, E.N. (2005). Hand is a direct target of Tinman and GATA factors during Drosophila cardiogenesis and hematopoiesis. *Dev* **132**, 3525-36.

Han, Z., Fujioka, M., Su, M., Liu, M., Jaynes, J.B. and Bodmer, R. (2002). Transcriptional Integration of Competence Modulated by Mutual Repression Genereates Cell-Type Specificity within the Cardiogenic Mesoderm. *Dev Biol*. **252**, 225-40.

Haran, T.E. and Mohanty, U. (2009). The unique structur of A-tracts and intrinsic DNA bending. *Quaterly Reviews of Biophysics* **42**, 41-81.

Harden, N. (2002). Signalling pathways directing the movement and fusuion of epithelial sheets: lessons from dorsal closure in Drosophila. *Differentiation* **70**, 181-203.

Hartenstein, V. and Mandal, L. (2006). The blood/vascular system in a phylogenetic perspective. *BioEssays* 28, 1203-10.

Hartl, D.L. and Lozoskaya, E.R. (1994). Genome evolution: between the nucleosome and chromosome. *EXS* **69**, 579-92.

Harvey, R.P. (1996). Nk-2 homeobox genes and heart development. *Dev Biol* **178**, 203-16.

Harvey, R.P. (1998). Links in the left/right axial pathway. *Cell* **94**, 273-76.

Harvie, P.D., Filippova, M. and Bryant, P.J. (1998). Genes expressed in the ring gland, the major endocrine organ of *Drosophila melanogaster*. *Genetics* **149**, 217-31.

Hasson, P., Müller, B., Basler, K. and Paroush, Z. (2001). Brinker requires two corepressors for maximal and versatile repression in Dpp signaling. *EMBO* **20**, 57-25-36.

Heitzler, P., Haenlin, M., Ramain, P., Calleja, M. and Simpson, P. (1996). A genetic analysis of pannier, a gene necessary for viability of dorsal tissues and bristle

positioning in Drosophila. *Genetics* **143**, 1271-86.
Hendren, J.D., Shah, A.P., Arguelles, A.M. and Cripps, R.M. (2007). Cardiac expression of the Drosophila Sulphonylurea receptor gene is regulated by an intron enhancer dependent upon the NK homeodomain factor Tinman. *Mech Dev* **124 (6)**, 416-26.
Holz, A., Bossinger, B., Strasser, T., Janning, W. and Klapper, R. (2003). The two origins of hemocytes in *Drosophila*. *Dev* **130**, 4955-4962.
Hoogaars, W.M.H., Barnett, P., Moorman, A.F.M. and Christoffels, V.M. (2007). T-box factors determine cardiac design. *Cell. Mol. Life Sci* **64**, 646-60.
Huang, X., M. Miller. (1991). A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math*. **12**, 337-357.
Hudson, J.B., Podos, S.D., Keith, K., Simpson, S.L. and Ferguson, E.L. (1998). The Drosophila Medea gene is required downstream of dpp and encodes a functional homolog of human Smad4. *Development* **125**, 1407-1420.
Ip, Y.T., Park, R.E., Kosman, D., Yazdanbakhsh, K. and Levine, M. (1992). Dorsal-twist interactions establish snail expression in the presumptive mesoderm of the *Drosophila* embryo. *Genes Dev* **6**, 1518-30.
Irvine, K.D. (1999). Fringe, Notch, and making developmental boundaries. *Curr Opin Genet Dev*. **9**, 434-41.
Jagla, K., Bellard, M. and Frasch, M. (2001). A cluster of Drosophila homeobox genes involved in mesoderm differentiation programs. *Bioessays* **23**, 125-33.
Jagla, K., Jagla, T., Heitzler, P., Dretzen, G., Bellard, F. and Bellard, M. (1997). Ladybird, a tandem of homeobox genes that maintain late wingless expression in terminal and dorsal epidermis of the *Drosophila* embryo. *Dev* **124**, 91-100.
Jagla, T., Bidet, Y., Da Ponte, J.P., Dastugue, B. and Jagla, K. (2002). Cross-repressive interactions of identity genes are essential for proper specification of cardiac and muscular fates in *Drosophila*. *Dev* **129**, 1037-47.
Jensen, D.E. and Buckner, B.A. (1973). Soluble and particle-bound trehalase in extracts from larvae, pupae, and adults of *Musca domestica*. *J Insect Physiol* **19**, 2255-59.
Jiang, D., Jarrett, H.W. and Haskins, W.E. (2009). Methods for proteomic analysis of transcription factors. *Journal of Chromatography A* **1216**, 6881-89.
Jiang, J., Kosman, D., Ip, Y.T. and Levine, M. (1991). The dorsal morphogen gradient regulates the mesoderm determinant twist in early *Drosophila* embryos. *Genes Dev* **5**, 1881-91.
Jin, H., Stojnic, R., Adryan, B., Ozdemir, A., Stathopoulos, A. and Frasch, M. (2013). Genome-wide screens for *in vivo* Tinman binding sites identify cardiac enhancers with diverse functional architectures. *PLOS Genetics* **9**, e1003195.
Johnson, A.N., Bergman, C.M., Kreitman, M. and Newfeld, S.J. (2003). Embryonic enhancers in the dpp disk region regulate a second round of Dpp signaling from the dorsal ectoderm to the mesoderm that represses Zfh-1 expression in a subset of pericardial cells. *Dev Biol*. **262**, 137-151.
Johnson, A.N., Burnett, L.A., Sellin, J., Paululat, A. and Newfeld, S.J. (2007). Defective decapentaplegic signaling results in heart overgrowth and reduced cardiac output in Drosophila. *Genetics* **176**, 1609-24.
Jonat, C., Rahmsdorf, H.J., Park, K.K., Cato, A.C., Gebel, S., Ponta, H. and Herrlich, P. (1990). Antitumor promotion and anti-inflammation: down-modulation of AP-1 (Fos/Jun) activity by glucocorticoid hormone. *Cell* **62**, 1189-204.

Junion, G., Spivakov, M., Giradot, C., Braun, M., Gustafson, E.H., Birney, E. and Furlong, E.E.M. (2012). A transcription factor collective defines cardiac cell fate and reflects lineage history. *Cell* **148**, 473-486.

Kambysellis, M.P. and Wheeler, M.R. (1972). Banded polytene chromosomes in pericardial cells of *Drosophila*. *J Hered* **63**, 214-15.

Keller, C.A., Erickson, M.S. and Abmayr, S.M. (1997). Misexpression of nautilus induces myogenesis in cardioblasts and alters the pattern of somatic muscle fibres. *Dev Biol* **181**, 197-212.

Klinedinst, S.L. and Bodmer, R. (2003). Gata factor Pannier is required to establish competence for heart progenitor formation. *Dev* **130**, 3027-38.

Knirr, S. and Frasch, M. (2001). Molecular integration of inductive and mesoderm-intrinsic inputs governs even-skipped enhancer activity in a subset of pericardial and dorsal muscle progenitors. *Dev Biol* **238**, 13-26.

Kölsch, V. and Paululat, A. (2002). The highly conserved cardiogenic bHLH factor Hand is specifically expressed in circular visceral muscle progenitor cells and in all cell types of the dorsal vessel during Drosophila embryogenesis. *Dev Genes Evol* **212**, 473-85.

Koo, H.S., Wu, H.M. and Crothers, D.M. (1986). DNA bending at adenine.thymine tracts. Nature **320**, 501-06.

Krejci, A. and Bray, S. (2007). Notch activation stimulates transient and selective binding of Su(H)/CSL to target enhancers. Genes Dev **12**, 1322-7**.**

Kremser, T., Gajewski, K., Schulz, R.A., Renkawitz-Pohl, R. (1999). Tinman regulates the transcription of the beta3 tubulin gene (betaTub60D) in the dorsal vessel of *Drosophila*. *Dev Biol*. **216**, 327-39.

Kuo, C.T., Morrisey, E.E. Anandappa, R., Sigrist, K., Lu, M.M., Parmacek, M.S., Soudais, C. and Leiden, J.M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes Dev*. **11**, 1048-60**.**

Lai, Z., Fortini, M.E. and Rubin, G.M. (1991). The embryonic expression pattern of zfh-1 and zfh-2, two Drosophila genes encoding novel zinc finger homeodomain proteins. *Mech Dev*. **34**, 123-34.

Lai, Z., Rushton, E., Bate, M. and Rubin, G.M. (1993). Loss of function of the *Drosophila* zfh-1 gene results in abnormal development of mesodermally derived tissues. *PNAS* **90**, 4122-4126.

Lall, S. and Patel, N.H. (2001). Conservation and divergence in molecular mechanisms of axis formation. *Annu Rev Genet* **35**, 407-37.

Latinkic, B.V., Kotecha, S. and Mohun, T.J. (2003). Induction of cardiomyocytes by GATA4 in Xenopus ectodermal explants. *Development* **130**, 3865-76.

Lavigne, M., Roux, P., Buc, H. and Schaeffer, F. (1997). DNA curvature controls termination of plus strand DNA synthesis at the centre of HIV-1 genome. *J Mol Biol* **266**, 507-24.

Lawrence, N., Dearden, P., Hartley, D., Roose, J., Clevers, H. and Martinez Arias, A. (2000). dTcf antagonizes Wingless signaling during the development and patterning of the wing in Drosophila. *Int. J. Dev. Biol*. **44**, 749-56.

Ledent, V. and Vervoort, M. (2001). The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Res* **11**, 757-70.

Lehmacher, C., Abeln, B. and Paululat, A. (2012). The ultrastructure of *Drosophila* heart cells. *Arthropod Structure & Development* 41, 459-474.

Lehmacher, C., Tögel, M., Pass, G. and Paululat, A. (2009). The Drosophila wing

Leptin, M. and Grunewald, B. (1990). Cell shape changes during gastrulation in *Drosophila*. *Dev* **110**, 73-84.

Lilly, B., Galewski, S., Firulli, A.B., Schulz, R.A. and Olson, E.N. (1994). D-Mef2: A MADS box transcription factor expressed in differentiating mesoderm and muscle cell lineages during *Drosophila* embryogenesis. *PNAS* **91**, 5662-66.

Lilly, B., Zhao, B., Ranganayakulu, G., Paterson, B.M., Schulz, R.A. and Olson, E.N. (1995). Requirement of MADS domain transcription factor D-MEF2 for muscle formation in Drosophila. *Science* **267**, 688-93.

Lin, M-H., Bour, B.A., Abmayr, S.M. and Storti, R.V. (1997). Ectopic expression of MEF2 in the epidermis induces expression of muscle genes and abnormal muscle development in *Drosophila*. *Dev Biol*. **182**, 240-255.

Lints, T.J., Parsons, L.M., Hartley, L., Lyons, I. and Harvey, R.P. (1993). Nkx-2.5: a novel murine homeobox gene expressed inearly heart progenitor cells and their myogenic descendants. *Dev* **119**, 419-13.

Liotta, D., Han, J., Elgar, S., Garvey, C., Han Z. and Taylor, M.V. (2007). The *Him* Gene Reveals a Balance of Inputs Controlling Muscle Differentiation in *Drosophila*. *Cur*. *Biol*. **17**, 1409-1413.

Liu, J., Qian, L., Wessells, R.J., Bidet, Y., Jagla, K. and Bodmer, R. (2006). Hedgehog and RAS pathways cooperate in the anterior-posterior specification and positioning of cardiac progenitor cells. *Dev Biol* **290**, 373-385.

Liu, Y.H., Jakobsen, J.S., Valentin, G., Amarantos, I., Glimour, D.T. and Furlong, E.E. (2009). A systematic analysis of Tinman function reveals Eya and JAK-STAT signaling as essential regulators of muscle development. *Dev Cell* **16**, 280-91.

Lo, P.C. and Frasch, M. (2001). A role for the COUP-TF-related gene seven-up in the diversification of cardioblast identities in the dorsal vessel of *Drosophila*. **Mech Dev 104**, 49-60.

Lo, P.C., Zaffran, S., Sénatore, S. and Frasch, M. (2007). The *Drosophila Hand* gene is required for remodeling of the developing adult heart and midgut during metamorphosis. *Dev Biol* **311**, 287-96.

Lo, P.C.H., Skeath, J.B., Gajewski, K., Schulz, R. and Frasch, M. (2002). Homeotic genes autonomously specifiy the anteroposterior subdivision of the *Drosophila* dorsal vessel into aorta and heart. *Dev Biol* **251**, 307-19.

Lovato, T.T., Nguyen, T.P., Molina, M.R. and Cripps, R.M. (2002). The Hox gene *abdominal-A* specifies heart cell fate in the *Drosophila* dorsal vessel. *Dev* **129*, 5019-27*.**

Lyons, I., Parsons, L.M., Hartley, L., Li, R., Andrews, J.E., Robb, L. and Harvey, R.P. (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeobox gene Nkx2-5. *Genes Dev* **9**, 1654-66.

Ma, P.C., Rould, M.A., Weintraub, H. and Pabo, C.O. (1994). Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. *Cell* **77**, 451-59.

Mandal, L., Banerjee, U. and Hartenstein, V. (2004). Evidence for a fruit fly hemangioblast and similarities between lymph-gland hematopoiesis in fruit fly and mammal aorta-gonadal-mesonephros mesoderm. *Nature Genetics* **36**, 1019-23.

Maruyama, R. and Andrew, D.J. (2012). *Drosophila* as a model for epithelial tube formation. *Dev Dyn* **241**, 119-35.

Mayor, C., Brundo, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S. and Dubchak, I. (2000). VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**, 1046-7.

McKnight, S.L. and Kingsbury, R. (1982). Transcriptional control signals of a eukaryotic protein coding gene. *Science* **217**, 316-24.

Medioni, C., Sénatore, S., Salmand, P-A., Lalevée, N., Perrin, L. and Sémériva, M. (2009). The fabulous destiny of the *Drosophila* heart. *Curr Opin Gene Dev* **19**, 518-25.

Medioni, C., Senatore, S., Salmand, P., Lalevée, N., Perrin, L. and Sémériva, M. (2009). The fabulous destiny of the *Drosophila* heart. *Curr Opin Gen & Dev* **19**, 518-525.

Michelson, A.M., Gisselbrecht, S., Buff, E. and Skeath, J.B. (1998a). Heartbroken is a specific downstream mediator of FGF receptor signaling in *Drosophila*. *Dev* **125**, 4379-89.

Michelson, A.M., Gisselbrecht, S., Zhou, Y., Baek, K.-H. and Buff, E.M. (1998b). Dual Functions of the Heartless Fibroblast Growth Factor Receptor in Development of the *Drosophila* Embryonic Mesoderm. *Developmental Genetics* **22**, 212-229.

Miskolczi-McCallum, C.M. Scavetta, R.J., Svendsen, P.C., Soanes, K.H. and Brook, W.J. (2005). The *Drosophila* melanogaster T-box genes *midline* and *H15* are conserved regulators of heart development. Dev Biol **278**, 459-72.

Mohun, T.J., Garrett, N. and Taylor, M.V. (1989). Temporal and tissue-specific expression of the proto-oncogene *c-fos* during development in *Xenopus laevis*. Development **107**, 835-846.

Mohun, T.J., Taylor, M.V., Garrett, N. and Gurdon, J.B. (1989). The CArG promoter sequence is necessary for muscle-specific transcription of the cardiac *actin* gene in *Xenopus* embryos. EMBO J **8**, 1153-61.

Molina, M.R. and Cripps, R.M. (2001). Ostia, the inflow tracts of the Drosophila heart, develop from a genetically distinct subset of pericardial cells. *Mech Dev* **109**, 51-59.

Molkentin, J.D. (2000) The zinc finger-containing transcription factors GATA-4, -5 and -6. *JBC* **275**, 38949-52.

Molkentin, J.D. and Olson, E.N. (1996a). Combinatorial control of muscle development by basic helix-loop-helix and MADS-box transcription factors. *Proc Natl Acad Sci USA* **93**, 366-73.

Molkentin, J.D. and Olson, E.N. (1996b). Defining the regulatory networks for muscle development. *Curr Opin Genet Dev* **6**, 445-53.

Molkentin, J.D., Firulli, A.B., Black, B.L., Martin, J.F., Hustad, C.M., Copeland, N., Jenkins, N. Lyons, G. and Olsen, E.N. (1996). MEF2B is a potent transactivator expressed in early myogenic lineages. *Mol Cell Biol* **16**, 3814-24.

Molkentin, J.D., Lin, Q., Duncan, S.A. and Olsen, E.N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes Dev*. **11*, 1061-72*.**

Monier, B., Astier, M., Sémériva, M. and Perrin, L. (2005). Steroid-dependent modification of Hox function drives myocyte reprogramming in the *Drosophila* heart. *Dev*. **132**, 5283-93.

Moore, A.W., Barbel, S., Jan, L.Y. and Jan, Y.N. (2000). A genomewide survey of basic helix-loop-helix factors in Drosophila. *Proc. Natl. Acad. Sci. USA* **97 (19)**, 10436-441.

Narlikar, L., Sakabe, N.J., Blanski, A.A. et al. (2010). Genome-wide discovery of human heart enhancers. *Genome Res* **20**, 381-92.

Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., Li, Z., Ishii, H., Spokony, R.F., Chen, J., Hwang, L., Cheng, C., Auburn, R.P., Davis, M.B., Domanus, M.,

Nègre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., Li, Z., Ishii, H., Spokony, R.F., Chen, J., Hwang, L., Cheng, C., Auburn, R.P., Davis, M.B., Domanus, M., Shah, P.K., Morrison, C.A., Zieba, J., Suchy, S., Senderowicz, L., Victorsen, A., Bild, N.A., Grundstad, A.J., Hanley, D., MacAlpine, D.M., Mannervik, M., Venken, K., Bellen, H., White, R., Gerstein, M., Russell, S., Grossman, R.L., Ren, B., Posakony, J.W., Kellis, M. and White, K.P. (2011). A cis-regulatory map of the *Drosophila* genome. *Nature* **471**, 527-31.

Neumann, C.J. and Cohen, S.M. (1998). Boundary formation in Drosophila wing: Notch activity attenuated by the POU protein Nubbin. *Science* **281**, 409-13.

Ng, M., Diaz-Benjumea, F.J. and Cohen, S.M. (1995). Nubbin encodes a POU-domain protein required for proximal-distal patterning in the Drosophila wing. *Development* **121**, 589-99.

Nguyen, H.T., Bodmer, R., Abmayr, S.M., McDermott, J.C. and Spoerel, N.A. (1994). D-mef2: A *Drosophila* mesoderm-specific MADS box-containing gene with a biphasic expression profile during embryogenesis. *PNAS* **91**, 7520-24.

Nishimura, M., Ocorr, K., Bodmer, R. and Cartry, J. (2011). *Drosophila* as a model to study cardiac aging. *Exp Gerontol* **46**, 326-30.

Noselli, S. and Agnes, F. (1999). Roles of the JNK signaling pathway in Drosophila morphogenesis. *Curr Opin Genet Dev* **9**, 466-72.

O'Keefe, L., Dougan, S.T., Gabay, L., Raz, E., Shilo, B.Z. and DiNardo, S. (1997). Spitz and Wingless, emanating from distinct borders, cooperate to establish cell fate across the Engrailed domain in the *Drosophila* epidermis. *Dev* **124**, 4837-45.

Obbard, D.J., MacIennan, J., Kim, K-W., Rambaut, A., O'Grady, P. and Jiggins, F.M. (2012). Estimating Divergence Dates and Substitution Rates in the *Drosophila* phylogeny. *Mol. Biol. Evol.* **29**, 3459-73.

Olivarez-Zavaleta, N., Jáuregui, R. and Merino, E. (2006). Genome analysis of *Escherichia coli* promoter sequences evidences that DNA static curvature plays a more important role in gene transcription than has previously been anticipated. *Genomics* **87**, 329-337.

Olson, E.N. and Srivastava, D. (1996). Molecular pathways controlling heart development. *Science* **272**, 671-6.

Olson, E.N., Perry, M. and Schulz, R.A. (1995). Regulation of muscle differentiation by the MEF familiy of MADS box transcription factors. *Dev Biol* **172**, 2-14.

Ozdemir, A., Fisher-Aylor, K.I., Pepke, S. et al. (2011). High resolution mapping of Twist to DNA in *Drosophila* embryos: Efficient functional analysis and evolutionary conservation. *Genome Res.* **21**, 566-77.

Ozdemir, A., Fisher-Aylor, K.I., Pepke, S. et al. (2011). High resolution mapping of Twist to DNA in *Drosophila* embryos: efficient functional analysis and evolutionary conservation. *Genome Res* **21**, 566-77.

Pan, D.J., Huang, J.D. and Courey, A.J. (1991). Functional analysis of the *Drosophila* twist promoter reveals a dorsal-binding ventral activator region. *Genes Dev* **5**, 1892-901.

Park, M., Yaich, L.E. and Bodmer, R. (1998). Mesodermal cell fate decisions in *Drosophila* are under the control of the lineage genes numb, Notch, and sanpodo. *MoD* **75**, 117-126.

Perrin, L., Monier, B., Ponzielli, R., Astier, M. and Sémériva, M. (2004). *Drosophila* cardiac tube organogenesis requires multiple phases of Hox activity. *Dev Biol* **272**, 419-31.

Perry, M.W., Boettiger, A.N., Bothma, J.P and Levine, M. (2010). Shadow

Ponzielli, R., Astier, M., Chartier, A., Gallert, A., Thérond, P. and Sémériva, M. (2002). Heart tube patterning in *Drosophila* requires integration of axial and segmental information provided by the Bithorax Complex genes and hedgehog signaling. *Dev* **129**, 4509-21.

Popichenko, D., Sellin, J., Bartkuhn, M. and Paululat A. (2007). Hand is a direct target of the forkhead transcription factor Biniou during Drosophila visceral mesoderm differentiation. *BMC Developmental Biology* **7:49**.

Postigo, A.A., Ward, W., Skeath, J.B. and Dean, D.C. (1999). Zfh-1, the Drosophila homologue of ZEB, is a transcriptional repressor that regulates somatic myogenesis. *Molec Cell Biol* **19**, 7255-63.

Potthoff, M.J. and Olson, E.N. (2007). MEF2: a central regulator of diverse developmental programs. *Dev* **134**, 4131-4140.

Powell, J.R. (1997). Progres and prospects in evolutionary biology: the *Drosophila* Model. Oxford University Press, New York.

Prall, O.W., Elliott, D.A. and Harvey, R.P. (2002). Developmental paradigms in heart disease: insights from *tinman*. *Ann Med*. **34**, 148-56.

Qian, L. and Bodmer, R. (2012). Probing the polygenic basis of cardiomyopathies in *Drosophila*. *J Cell Mol Med* **16**, 972-77.

Qian, L., Liu, J. and Bodmer, R. (2005). *Neuromancer* Tbx20-related genes (*H15/midline*) promote cell fate specification and morphogenesis of the *Drosophila* heart. Dev Biol. **279**, 509-24.

Qian, L., Mohapatra, B., Akasaka, T., Liu, J., Ocorr, K., Towbin, J.A. and Bodmer, R. (2008). Transcription factor neuromancer/TBX20 is required for cardiac function in *Drosophila* with implications for human heart disease. *PNAS* **105**, 19833-38.

Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res*. **23**, 4878-84.

Ramain, P., Heitzler, P., Haenlin, M. and Simpson, P. (1993). *pannier*, a negative regulator of *achaete* and *scute* in *Drosophila*, encodes a zinc finger protein with homology to the vertebrate transcription factor GATA-1. *Dev* **119**, 1277-1291.

Ranganayakulu, G., Schulz, R.A. and Olson, E.N. (1996). Wingless signaling induces *nautilus* expression in the ventral mesoderm of the *Drosophila* embryo. *Dev Biol*. **176**, 143-48.

Ranganayakulu, G., Zhao, B., Dokidis, A., Molkentin, J. D., Olson, E. N. and Schulz, R. A. (1995). A series of mutations in the D-MEF2 transcription factor reveal multiple functions in larval and adult myogenesis in *Drosophila*. *Dev Biol* **171**, 169-81.

Reamon-Buettner, S.M. and Borlak, J. (2010). Nkx2-5: An Update on this hypermutable homeodomain protein and its role in human Congenital Heart Disease (CHD). *Hum Mutat* **31**, 1185-94.

Rebeiz, M., Reeves, N.L. and Posakony, J.W. (2002) SCORE: a computational approach to the identification of cis-regulatory modules and target genes in whole-genome sequence data. Site clustering over random expectation. *PNAS* **99**, 9888-93.

Reeves, G.T. and Stathopoulos, A. (2009). Graded dorsal and differential gene regulation in the *Drosophila* embryo. *CSH Perspectives* **1**:a000836

Reim, I. and Frasch, M. (2005). The Dorsocross T-box genesare key components of the regulatory network controlling early cardiogenesis in Drosophila. Dev **132**, 4911-4925.

Reim, I., Lee, H.-H. and Frasch, M. (2003). The T-box-encoding Dorsocross genes function in amnioserosa development and the patterning of the dorsolateral germ band downstream of Dpp. *Development* **130**, 3187-204.

Reim, I., Mohler, J.P. and Frasch, M. (2005). *Tbx20*-related genes, *mid* and *H15*, are required for *tinman* expression, proper patterning, and normal differentiation of cardioblasts in *Drososphila*. *Mech Dev* **122**, 1056-59.

Richards S, Liu Y, Bettencourt BR, Hradecky P, Letovsky S, Nielsen R, Thornton K, Hubisz MJ, Chen R, Meisel RP, Couronne O, Hua S, Smith MA, Zhang P, Liu J, Bussemaker HJ, van Batenburg MF, Howells SL, Scherer SE, Sodergren E, Matthews BB, Crosby MA, Schroeder AJ, Ortiz-Barrientos D, Rives CM, Metzker ML, Muzny DM, Scott G, Steffen D, Wheeler DA, Worley KC, Havlak P, Durbin KJ, Egan A, Gill R, Hume J, Morgan MB, Miner G, Hamilton C, Huang Y, Waldron L, Verduzco D, Clerc-Blankenburg KP, Dubchak I, Noor MA, Anderson W, White KP, Clark AG, Schaeffer SW, Gelbart W, Weinstock GM, Gibbs RA. (2005) Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and cis-element evolution. *Genome Research* **15**, 1-18

Riechmann, V., Irion, U., Wilson, R., Grosskortenhaus, R. and Leptin, M. (1997). Control of cell fates and segmentation in the *Drosophila* mesoderm. *Development* **124**, 2915-22.

Rizki, T.M. (1978). The circulatory system and associated cells and tissues. *In: Genetics and biology of Drosophila*, (eds. M. Ashburner and T.R.F. Wright), Academic Press, London, 398-451.

Roth, S., Stein, D. and Nüsslein-Volhard, C. (1989). A gradient of nuclear localization of the dorsal protein determines dorsoventral pattern in the Drosophila embryo. *Cell* **59**, 1189-1202.

Rubin, G.M. and Spradling, A.C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* 22, 348-53.

Rugendorff, A., Younossi-Hartenstein, A. and Hartenstein, V. (1994). Embryonic origin and differentiation of the *Drosophila* heart. **Roux's Arch Dev Biol 203**, 266-280.

Rusch, J. and Levine, M. (1997). Regulation of a dpp target gene in the Drosophila embryo. *Development* **124**, 303-11.

Rushlow, C.A., Han, K., Manley, J.L. and Levine, M. (1989). The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila. *Cell* **59**, 1165-77.

Rushton, E., Drysdale, R., Abmayr, S.M., Michelson, A.M. and Bate, M. (1995). Mutations in a novel gene, myoblast city, provide evidence in support of the founder cell hypothesis for *Drosophila* muscle development. *Dev* **121**, 1979-88.

Russo, C.A., Takezaki, N. and Nei, M. (1995). Molecular phylogeny and divergence times of drosophilid species. *Mol Biol Evol* **12**, 391-404.

Ryan, K. and Chin, A.J. (2003). T-box genes and cardiac development. *Birth Defects Res., Part C Embryo Today* **69**, 25-37.

Ryan, K.M., Hendren, J.D., Helander, L.A. and Cripps, R.M. (2007). The NK homeodomain transcription factor Tinman is a direct activator of *seven-up* in the *Drosophila* dorsal vessel. *Dev Biol* **302**, 694-702.

Ryan, K.M., Hoshizaki, D.K. and Cripps, R.M. (2005). Homeotic selector genes control the patterning of seven-up expressing cells in the *Drosophila* dorsal vessel. *Mech Dev* **122**, 1023-33.

Sandmann, T., Jensen, L.J., Jakobsen, J.S., Karzynski, M.M., Eichenlaub, M.P., Bork, P. and Furlong, E.E.M. (2006). A temporal map of transcription factor activitiy: Mef2 directly regulates target genes at all stages of muscle development. *Dev Cell* **10**, 797-807.

Schott, J.J.; Benson, D.W., Basson, C.T., Pease, W., Silberbach, G.M., Moak, J.P., Maron, B.J., Seidman, C.E. and Seidman, J.G. (1998). Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science* **281**, 108-11.

Schwartz, S., Zhang, Z. and Frazer, K.A. (2000). PipMaker—A Web Server for Aligning Two Genomic DNA Sequences. *Genome Res*. **10**, 577-586.

Sellin, J., Albrecht, S., Kölsch, V. and Paululat, A. (2006). Dynamics of heart differentiation, visualized utilizing heart enhancer elements of the Drosophila *melanogaster* bHLH transcription factor Hand. *Gene Exp*. *Patt*. **6**, 360-375.

Sellin, J., Drechsler, M., Nguyen, H.T. and Paululat, A. (2009). Antagonistic function of Lmd and Zfh1 fine tunes cell fate decisions in the Twi and Tin positive mesoderm of *Drosophila* melanogaster. *Dev Biol* **326**, 444-55.

Shah, A.P., Nongthomba, U., Tanaka, K. K. K., Denton, M.L.B., Meadows, M.M., Bancroft, N., Molina, M.R. and Cripps, R.M. (2011) Cardiac remodeling in *Drosophila* arises from changes in actin gene expression and from a contribution of lymph gland-like cells to the heart musculature. *MoD* **128**, 222-233.

Shishido, E., Higashijima, S-I., Emori, Y. and Saigo, K. (1993). Two FGF-receptor homologues of *Drosophila*: one is expressed in meodermal primordium in early embryos. *Dev* **117**, 751-761.

Simpson, P. (1983). Maternal-zygotic gene interactions during formation of the dorsoventral pattern in *Drosophila* embryos. *Genetics* **105**, 615-32.

Siva, K. and Inamdar, M.S. (2006). Rudhira is a cytoplasmic WD40 protein expressed in mouse embryonic stem cells and during embryonic erythropoiesis. *Gene Expr*. *Patterns* **6**, 225-34.

Soler, C. and Taylor, M.V. (2009). The *Him* gene inhibits the development of *Drosophila* flight muscles during metamorphosis. **Mech Dev** 126, 595-603.

Soler, C., Han, J. and Taylor M.V. (2012). The conserved transcription factor Mef2 has multiple roles in adult *Drosophila* musculature formation. *Dev* **139,** 1270-75.

Sorrentino, R.P., Gajewski, K.M. and Schulz, R.A. (2005). GATA factors in Drosophila heart and blood cell development. *Semin Cell Dev Biol*. **16**, 107-116.

Staehling-Hampton, K., Hoffmann, F.M., Baylies, M.K., Rushton, E. and Bate, M. (1994). Dpp induces mesodermal gene expression in *Drosophila*. *Nature* **372**, 783-6.

Su, M.-T., Fujioka, M., Goto, T. and Bodmer, R. (1999). The *Drosophila* homeobox genes zfh-1 and even-skipped are required for cardiac- specific differentiation of a numb-dependent lineage decision. *Development* **126**, 3241-51.

Tamura, K., Subramanian, S. and Kumar, S. (2004). Temporal Patterns of Fruit Fly (*Drosophila*) Evolution Revealed by Mutation Clocks. *Mol*. *Biol*. *Evol*. **21**, 36-44.

Tanaka, M., Kasahara, H., Bartunkova, S., Schinke, M., Komuro, I., Knagaki, H., Lee, Y., Lyons, G.E. and Izumo, S. (1998). Vertebrate homologs of Tinman and bagpipe: roles of the homeobox genes in cardiobascular development. *Dev Genet* **22**, 239-49.

Taylor, M.V. (2000). A novel *Drosophila*, mef2-regulated muscle gene isolated in a substractive hybridization-based molecular screen using small amounts of zygotic mutant RNA. *Dev Biol*. **220**, 37-52.

Taylor, M.V., Beatty, K.E., Hunter, H.K. and Baylies, M.K. (1995). Drosophila MEF2 is regulated by twist and is expressed in both the primodial and differentiated cells of the embryonic somatic, visceral and heart musculature. *Mech Dev* **50**, 29-41.

The FlyBase Consortium (2003). The FlyBase database of the Drosophila genome projects and community literature. *Nucleic Acids Research* **31**, 172-75. http://flybase.org/

The modENCODE Consortium et al. (2010). Identification of Functional Elements and Regulatory Circuits by *Drosophila* modENCODE. *Science* **330**, 1787-97.

Thisse, B., Stoetzel, C., Gorostiza-Thisse, C. and Perrin-Schmitt, F. (1988). Sequence of the twist gene and nuclear localization of its protein in endomesodermal cells of early *Drosophila* embryos. *EMBO J*, 2175-83.

Thummel, C.S., Boulet, A.M. and Lipshitz, H.D. (1988). Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. *Gene*. **74**, 445-56.

Tian, F., Shah, P.K., Liu, X., Negre, N., Chen, J., Karpenko, O., White, K.P. and Grossman, R.L. (2009). Flynet: a genomic resource for *Drosophila melanogaster* transcriptional regulatory networks. *Bioinformatics* **35**, 3001-4.

Tögel, M., Pass, G. and Paululat, A. (2008). The Drosophila wing hearts originate from pericardial cells and are essential for wing maturation. *Dev Biol*. **318**, 29-37.

Tweedie, S., Ashburner, M., Falls, K., Leyland, P., McQuilton, P., Marygold, S., Millburn, G., Osumi-Sutherland, D.,Schroeder, A., Seal, A., Zhang, H. and The FlyBase Consortium.(2009) FlyBase: enhancing Drosophila Gene Ontology annotations. *Nucleic Acids Research* **37**, 555-559

Vincent, S., Wilson, R., Coelho, C., Affolter, M. and Leptin, M. (1998). The *Drosophila* protein Dof is specifically required for FGF Signaling. *Molecular Cell* **2**, 515-525.

Wang, J., Tao, Y., Reim, I., Gajewski, K., Frasch, M. and Schulz, R.A. (2005). Expression, regulation, and requirement of the toll transmembrane protein during dorsal vessel formation in *Drosophila melanogaster*. *Mol Cell Biol*. **25**, 4200-10.

Ward, E.J. and Coulter, D.E. (2000). Odd-skipped is expressed in multiple tissues during Drosophila embryogenesis. Mech Dev **96**, 233-236.

Ward, E.J. and Skeath, J.B. (2000). Characterization of a novel subset of cardiac cells and their progenitors in the Drosophila embryos. *Development* **127**, 4959-69.

Wasserthal, L.T. (2007). *Drosophila* flies combine periodic heartbeat reversal with a circulation in the anterior body mediated by a newly discovered anterior pair of ostial valves and "venous" channels. *J Exp Biol* **210**, 3707-19.

Weavers, H., Prieto-Sánchez, S., Grawe, F., Garcia-López, A., Wilsch-Bräuninger, M., Ruiz-Gómez, M., Skaer, H. and Denholn, B. (2009). The insect nephrocyte is a podocyte-like cell with a filtration slit diaphragm. *Nature* **457**, 322-36.

Wharton, K.A., Ray, R.P. and Gelbart, W.M. (1993). An activity gradient of *decapentaplegic* is necessary for the specification of dorsal pattern elements in the *Drosophila* embryo. *Development* **117**, 807-22.

Wingender, E., Chen, X., Fricke, E., Geffers, R., Hehl, R., Liebich, I., Krull, M., Matys, V., Michael, H., Ohnhauser, R., Pruss, M., Schacherer, F., Thiele, S. and Urbach, S. (2001). The TRANSFAC system on gene expression regulation. *Nucleic Acids Res* **29**, 281-3.

Wisotzkey, R.G., Mehra, A., Sutherland, D.J., Dobens, L.L., Liu, X., Dohrmann, C., Attisano, L. and Raftery, L.A. (1998). Medea is a Drosophila Smad4 homolog that is differentially required to potentiate DPP responses. *Development* **125**, 1433-45.

Wittkopp, P.J. (2006). Evolution of *cis*-regulatory sequence and function in Diptera. *Heredity* **97**, 139-47.

Wodarz, A., Hinz, U., Engelbert, M. and Knust, E. (1995). Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of *Drosophila*. *Cell* **82**, 67-76.

Wu, X., Golden, K. and Bodmer, R. (1990). Heart development in *Drosophila* requires the segment polarity gene *wingless*. *Dev*. *Biol*. **169**, 619-628.

Yeo, S.L., Lloyd, A. and Kozak, K. (1995). On the functional overlap between two Drosophila POU homeo domain genes and the cell fate specification of a CNS neural precursor. *Genes Dev*. **9**, 1223-36.

Yin, Z. and Frasch, M. (1989). Regulation and function of *tinman* during dorsal mesoderm induction and heart specification in *Drosophila*. *Dev Gen* **22**, 187-200.

Yin, Z., Xu, X.L. and Frasch, M. (1997). Regulation of the twist target gene Tinman by modular cis-regulatory elements during early mesoderm development. *Development* **124**, 4971-82.

Zaffran, S. and Frasch, M. (2002). Early signals in cardiac development. *Circ Res* **91**, 457-69.

Zaffran, S. and Frasch, M. (2005). The homeodomain of Tinman mediates homo- and heterodimerization of NK proteins. *Biochem Biophys Res Commun*. **334**, 361-9.

Zaffran, S., Astier, M., Gratecos, D., Guillen, A. and Sémériva M. (1995). Cellular interactions during heart morphogenesis in the *Drosophila* embryo. *Biol Cell* 84, 13-24.

Zaffran, S., Reim, I., Qian, L., Lo, P.C., Bodmer, R. and Frasch, M. (2006). Cardioblast-intrinsic Tinman activity controls proper diversification and differentiation of myocardial cells in *Drosophila*. *Development* **133**, 4073-83.

Zeitouni, B., Sénatore, S., Séverac, D., Aknin, C., Sémériva, M. and Perrin, L. (2007). Signalling pathways involved in adult heart formation revealed by gene expression profiling in *Drosophila*. *PLoS Genetics* **3**, 1907-21.

Zhuang, S., Shao, H., Guo, F., Trimble, R., Pearce, E. and Abmayr, S.M. (2009). Sns and Kirre, the *Drosophila* orthologs of Nephrin and Neph1, direct adhesion, fusion and formation of a slit diaphragm-like structure in insect nephrocytes. *Dev* **136**, 2335-44.

Zinzen, R.P., Girardot, C., Gagneur, J., Braun, M. and Furlong, E.E.M. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* **462**, 65-72.

Venkatesh, T.V., Park, M., Ocorr, K., Nemaceck, J., Golden, K., Wemple, M. and Bodmer, R. (2000). Cardiac enhancer activity of the homeobox gene *tinman* depends on the CREB consensus binding sites in *Drosophila*. *Genesis* **26**, 55-66.

Leiss, D., Hinz, U., Gasch, A., Mertz, R. and Renkawitz-Pohl, R. (1988). *ß3 tubulin* expression characterizes the differentiating mesodermal germ layer during *Drosophila* embryogenesis. *Dev* **104**, 525-31.