Computerised Modelling for Developmental Biology

Bertens, L.M.F.

Cover Page

Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/19772 holds various files of this Leiden University dissertation.

Author: Bertens, Laura M.F.
Title: Computerised modelling for developmental biology : an exploration with case studies
Date: 2012-09-12

Computerised modelling for developmental biology
-
An exploration with case studies

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 12 september 2012
klokke 11.15

door

Laura Bertens
geboren te Zeist, Nederland, in 1981

To my dad

TABLE OF CONTENTS

INTRODUCTION

Many studies in developmental biology rely on the construction, simulation and analysis of models. These models vary from schematic drawings, representing *e.g.* regulatory pathways, to complex systems of equations, enabling accurate quantitative analyses of processes. New modelling techniques are constantly being developed, as well as new applications for already existing techniques. In particular, a growing interest can be observed in employing computational modelling approaches to biological phenomena, exemplified by the recently emerged field of computational biology.

This dissertation presents a broad view of modelling approaches for developmental biology. In the first chapter an overview is given of modelling properties, relevant when selecting a particular modelling approach, followed by a review of frequently used modelling techniques in the field of developmental biology.

The first chapter establishes a context for the subsequent chapters, 2 to 6, in which a series of case studies is presented for various modelling approaches. The approaches have been selected on the basis of their particular merits for developmental biology and each represents a different category from the review in chapter 1. Since computation is becoming increasingly important for developmental biology, the case studies presented in this dissertation make use of computational and/or computer-assisted techniques. In chapters 2 and 3, 3D reconstructions are used to study heart development in the turtle. Chapter 4 describes a general ontology system for vertebrate heart development, incorporating the use of 3D reconstructions. Finally, chapters 5 and 6 present a modelling solution for the formation of biological gradients, using the algorithmic approach of Petri nets, combined with the equation based approach of differential equations.

In each of the case studies, optimization of captured knowledge and functionality was strived after, by taking advantage of the specific characteristics of the modelling approach. As a result, new ways of utilizing the techniques were developed, in particular for ontologies and Petri nets (*cf.* chapters 4, 5 and 6), and new insights were gained into biological phenomena, in particular for the cardiac development of the turtle (*cf.* chapter 2 and 3). Furthermore, attention was paid to integration of modelling approaches, in order to extend the functionality of the resulting models. Chapter 7 addresses this aspect of integration, for each of the separate instances. In this way the dissertation presents a broad investigation of new approaches to modelling structures and processes in developmental biology.

CHAPTER 1. MODELLING TECHNIQUES FOR DEVELOPMENTAL BIOLOGY – AN OVERVIEW

τό τε γὰρ μιμεῖσθαι σύμφυτον τοῖς ἀνθρώποις ἐκ παίδων ἐστὶ καὶ τούτῳ διαφέρουσι τῶν ἄλλων ζῴων ὅτι μιμητικώτατόν ἐστι καὶ τὰς μαθήσεις ποιεῖται διὰ μιμήσεως τὰς πρώτας

For it is an instinct of humans, from childhood, to engage in mimesis and this distinguishes them from other animals: man is the most mimetic and it is through mimesis that he gains his earliest insights.

- Aristotle, *Poetica* IV, v.4-8

1.1 MODELLING IN BIOLOGY

A first step in trying to understand the world around us is simplification. We tend to create simplified mental images of both concrete and abstract entities. Looking up at the night sky, for instance, we reduce the infinite number of stars to a set of easily recognisable constellations. We simplify things to a basic set of features, which convey all the information necessary to us. However, decisions about which information is necessary depend entirely on the situation at hand.

Furthermore we try and distinguish patterns in the information we perceive. If a pattern is underlying a series of events or objects, there is no need to remember each of the separate instances, for remembering the pattern itself suffices. Recognizing these patterns enables us to predict future events. Our knowledge of the movement of the planets, for instance, illustrates the use of understanding patterns in observed phenomena.

So, in order to comprehend concrete and abstract concepts, we build (mental) models; instead of using all information at hand we select the most relevant components and construct a simplified version of the entity we want to understand. The word 'model' is defined in the dictionary as "a small object, usually built to scale, that represents in detail another, often larger object" or "a schematic description of a system, theory or phenomenon that accounts for its properties and may be used for further study of its characteristics"[1]. Although clearly distinct, these two definitions both

---

[1] *The American Heritage College Dictionary,* Boston and New York, 1993.

focus on the representation of complex systems in a simplified manner. Here, as in science in general, we are particularly interested in the second definition[2].

Biological sciences are often concerned with complex systems and studying these would be unthinkable without the use of models. For centuries scientists have built models to aid the understanding of life in all its appearances. This modelling has taken on many forms in diverse fields in the biological sciences, including the 19th-century papier mâché models of Louis Thomas Jerôme Auzoux, the double helix DNA model by Watson and Crick, but also for instance statistical models in palaeontology and the modelling by taxonomic trees in evolutionary biology. Here we will focus on one particular research field, developmental biology, and the use of (computer assisted) modelling in this field. Firstly, model properties will be discussed, important in choosing a method and making modelling decisions. Subsequently a basic outline will be given of the different types of modelling that play a role in studying development.

1.2 MODEL PROPERTIES

The way we model an entity depends on the requirements of the model and the aspects of the structure or the process on which the model focuses. Which features of the studied entity are important to us and which can be ignored given the specific study? When we are interested in gene expression of immune cells in zebrafish we might omit the cells' metabolism from the model, but when studying solely the KREBS cycle the gene expression is less relevant.

Along with the topic of study, the choice of what to include and how to model the information is also dictated by the purpose of the model. Models are built for various reasons: to educate, to test hypotheses, to examine which assumptions about a system best fit the observed data, to predict future events, or to circumvent experiments that are too costly, time consuming or ethically undesirable. All these motives call for different properties of the modelling method. Here the most important sets of properties to be considered when building a model will be introduced and discussed. An overview of these properties is provided in Table 1.1.

---

[2] It is important to establish an unambiguous understanding of the concept 'model' in this text. The word has many uses, not only in daily life but also in the life sciences. Apart from artificial man-made models there is another very important type of model in biology: animal models. When studying human diseases and physiological or developmental processes it is rarely possible to obtain experimental data from humans; therefore animals are used that resemble the human physiological system under study. This type of model will, however, be excluded from the overview presented in this chapter and the term 'model' refers exclusively to artificial, man-made models, used to represent biological systems.

First of all, models can be either **descriptive** or **mechanistic**. Descriptive models represent observed data, but without modelling the underlying processes that produce these data. For instance, a graph of the numbers of cells in a frog embryo, plotted against time, describes the cell division but does not give us any information about the process. Therefore we cannot determine which variables play a role in the process and the model cannot be used to extrapolate and predict new data. A mechanistic model, on the other hand, describes the underlying process and includes the relevant variables and can thus be used to predict future events. Were we to model the genetic pathways underlying the cell divisions, the resulting model would be mechanistic. While mechanistic models teach us more about the workings of developmental processes, frequently too little is known about a process to construct such a model and constructing a descriptive model is the best option.

Secondly, models can be either **static** or **dynamic**. As the name implies static models describe static non-changing states, without a time component. It is evident that these models are by their very nature descriptive, as opposed to mechanistic. Dynamic models, on the other hand, describe the changes in a system or entity over time (either the simulated time of the modeled process or the running time of the model). They can be descriptive or mechanistic. The Petri net model describing the formation of biological gradients over time, presented in chapter 5 and 6, is an example of a mechanistic dynamic model.

Another defining feature of a model is whether it is **qualitative** or **quantitative**. Many processes can be understood and modelled in a qualitative way, without using exact experimental data. In the case of gradient formation for instance (*cf.* chapter 5), building a model to describe the process does not require exact quantitative data. As such, a qualitative model is a good starting point for modelling a new study, for which experimental data are still incomplete. When, on the other hand, experimental data are available, a quantitative model will allow statistical analysis and make quantitative predictions about future states. Qualitative models can be transformed into quantitative models, by including exact variable values; this is explicitly described for gradient formation in chapter 6.

This distinction between qualitative and quantitative models is directly related to the distinction between **theoretical** and **practical** (or **conceptual** and **concrete**, respectively). Theoretical models are aimed at furthering the theoretical understanding of a system or concept. They do not call for numerical accuracy and are often generic to a group of systems of similar mode of operation. Practical models are used for specific practical situations, in which decisions are made using model predictions and analyses. They are specific to a particular system and include detailed quantitative data. The distinction theoretical versus practical is linked to the trade-off between model generality and model (prediction) strength; a basic model with little

detail will be applicable to a wide range of similar processes, but cannot accurately predict future events for any of those. A highly detailed model, however, has greater predictive power, but only for a specific instance of the process under study. The choice for one or the other depends entirely on the purpose of the model, theoretical or practical.

Models representing system behaviour can do so in a **discrete** or a **continuous** manner. Discrete models represent the behaviour of systems in distinct spatial and/or temporal steps. Continuous models, on the other hand, represent space and/or time in the system's behaviour in a continuous manner, in which no separate steps are discerned. Particular attention is paid to these characteristics and the translation from one to the other in chapter 6.

Furthermore, dynamic models can be either **deterministic** or **nondeterministic**. Whereas series of events in deterministic models are determined at the onset and are predictable, nondeterministic and stochastic models allow multiple outcomes of a process. Stochastic models include the element of probability: events are not set, but depend on levels of probability. In developmental biology nondeterminism can arise through the concept of concurrency (*cf.* chapter 5 and 6) and interaction between process components, while stochastic variables can reflect environmental influences external to the system under study; for instance, the way a fertilized frog egg cell comes to lie on the substrate can be considered a stochastic factor and determines where the animal and vegetal poles of the cell will arise.

MODEL PROPERTIES		
descriptive	versus	mechanistic
static	versus	dynamic
qualitative	versus	quantitative
theoretical	versus	practical
discrete	versus	continuous
deterministic	versus	nondeterministic

Table 1.1. An overview of model properties relevant to modelling in developmental biology.

Apart from these fundamental distinctions there are several other model features that have bearing on the uses of a model. Models can be **temporal** (either discrete or continuous), like the cell division graph mentioned above. **Spatial** models include elements of space, *e.g.* 3-dimensional models of biological structures. These features are not mutually exclusive, in fact many models in developmental biology are both; for instance a model visualising the cell movement of neural crest cells during early development is both spatial and temporal.

1.3 VERBAL AND VISUAL MODELS – NON-COMPUTATIONAL MODELLING

Now that the most significant model properties have been discussed, we will take a look at the different types of models used in developmental biology. There are many ways to divide models into groups, for instance based on the features described above (static models, dynamic models, etcetera). In this overview four types of models are discussed: verbal, visual, algorithmic and equation based models. The last two types are both inherently computational and most often computer-assisted. The first two are not in essence computational and are therefore distinguished in this overview from the computational modelling techniques. However, computation is often possible and useful for these types as well and many verbal and visual models used in studying developmental biology are in fact computer-assisted and/or computational. Since computation is becoming increasingly important for developmental biology, the case studies presented in chapters 2 to 6 make use of computational and/or computer-assisted techniques.

This section will deal with verbal and visual models, *i.e.* model types that are not by definition computational. In 1.4 computational models will be discussed, divided in algorithm and equation based model types. An overview of all four categories and their model types is provided in Table 1.2. For all categories presented here emphasis will be placed on computational and computer-assisted types of modelling since these are rapidly gaining importance in developmental biology. The categorization of modelling types is of course not strict and many models can be considered hybrid and can therefore be assigned to more than one category.

1.3.1 Verbal models

On the most basic level we represent objects and concepts with words. We use language to refer to entities and to enable communication. As such, it is an elementary modelling system. Apart from our everyday, natural vocabulary, we have constructed **terminologies** for many specific research fields, including developmental biology.

VERBAL	- terminologies - ontologies

VISUAL	**SCHEMATIC REPRESENTATIONS**
	   - 2D – schematic drawings
	   - 3D – tangible/virtual 3D models
	   - time – time series of schematic
	      representations
	DIAGRAMS
	   - informal diagrams
	   - formal diagrams / graphical visualisations of
	      computational models
	GRAPHICAL REPRESENTATIONS OF TABULAR DATA
	   - charts

ALGORITHM BASED	- automata based models:
	    • cellular automata
	    • Lindenmayer systems
	- boolean networks
	- process algebra/calculi
	- Petri nets

EQUATION BASED	- differential equations:
	    • gene regulatory networks
	    • reaction diffusion systems
	    • diffusion models
	- enzyme kinetics
	- organismal growth models

Table 1.2. An overview of the categories and types of models presented in this chapter.

Classification of biological entities is by no means a new phenomenon. Throughout the ages people have tried to define formal classifications and terminologies; as early as the fourth century BC, Aristotle wrote extensively about nature, classifying animals and plants (*cf. Historia Animalium*) in a system which has influenced classifications all throughout the Medieval period and the Renaissance. During the age of enlightenment the very nature of science and experimental studies changed and scholars focused their energy on systematically collecting information and, more importantly, sharing the new-found knowledge. Diderot's *Encyclopédie* (1751-1772) was as such an invaluable document, as was Linnaeus' *Systema Naturae* (1735). With modern science becoming more international, conventional terminologies and definitions are crucial in communication between scientists of a particular research field, and formal systems of knowledge are indispensible.

The field of developmental biology relies heavily on terminologies and therefore stands to gain from optimizing the efficient use of this knowledge. A useful modern method of formally representing a terminology, containing the relevant objects, concepts and entities and the relationships between these terms, is collecting this knowledge in an **ontology** (Gruber, 1993). Put another way, an ontology is a verbal model of a domain, comprising the knowledge of all concepts within that particular domain and the way these are related. The knowledge is structured in the ontology in 'classes' (analogous to encyclopaedic lemmas), which are linked in a hierarchical system by 'IS_A' relationships (*e.g.* 'cow' IS_A class in the superclass 'vertebrates'). In addition to the IS_A relationship further 'properties' (relationships between the classes) can be defined within the ontology. In this way, ontologies allow researchers to describe an area of research, using a shared standardized vocabulary, and to classify the concepts and relationships in this area into groups and hierarchies.

Ontologies are becoming increasingly important in biology in general (Dmitrieva, 2011), as exemplified by the successful Gene Ontology, GO (The Gene Ontology Consortium, 2000), and the Sequence Ontology (Eilbeck *et al.*, 2005). In an attempt to oversee and coordinate the ever growing number of biological ontologies, the Open Biomedical Ontologies consortium (OBO) was founded (Smith *et al.*, 2007; Bodenreider and Stevens, 2006). By providing shared principles, such as the use of a common representation in the OBO format or the Web Ontology Language (OWL, https://www.w3.org/TR/owl-features), the consortium helps make ontologies interoperable and of greater use to biologists. Amongst other ontologies, OBO contains a number of specific developmental ontologies, dealing with the development of model species such as *Arabidopsis thaliana*, *Caenorhabditis elegans* and *Drosophila melanogaster*. These ontologies include information about the different anatomical structures present at particular developmental stages.

The most basic goal of ontologies is to formally capture the knowledge of a specific research area and to provide researchers with a standardized vocabulary to communicate. In addition to this, they can also be used to link terms in the vocabulary to other resources in the research field, such as literature and images, as is possible with for instance the GO. Over the last few years progress has been made on implementation of a Semantic Web for the life sciences (Bodenreider and Stevens, 2006). Nowadays ontologies are an integral part of bioinformatics and developmental biology, so much so that the leading journal *Bioinformatics* now dedicates an entire section to this field.

In essence, ontologies are descriptive rather than mechanistic and static rather than dynamic. However, since ontologies by their very nature provide great freedom in relating terms using relevant properties (for instance concerning processes taking place between structures under differing circumstances), they can be used to include mechanistic and dynamic information. An example of this can be found in

chapter 4, where dynamic information on heart physiology, under differing circumstances, has been added to an ontology of heart anatomy and development. Furthermore, while ontologies can be used entirely qualitatively by including solely structure names and qualitative properties, it is also possible to add quantitative information about the structures in data properties. This also allows one to transform a theoretical model into a practical one. Since any semantic content can be modelled in an ontology, the information can be both temporal and spatial.

In chapter 4 we introduce our use of ontologies to capture knowledge about the anatomy, development and physiology of the vertebrate heart. This serves as a case study for extending the functionality of ontologies for developmental biology; as will be explained in more depth in chapter 4, it incorporates the whole range of information types, from static, qualitative knowledge to dynamic, quantitative information in both temporal and spatial domains. By connecting the ontology to 3D models (*cf.* 4.4.4) we integrate verbal modelling and visual modelling. This integration of different types of models will be further discussed in chapter 7.

1.3.2 Visual models

In the biological sciences visual information is very important, more so than in the mathematical, physical or chemical sciences, where abstract formulas can more easily be used to describe and study the systems at hand. Biological textbooks illustrate our knowledge of biological structures with drawings and diagrams and more and more journal publications refer to websites with supplementary figures and visual aids, to be downloaded or viewed online (*e.g.* Noble, 2002). Not only can visual models help clarify information, they can also lead to new insights that would be hard to achieve through other means (Iwasa, 2010); in science, as in life in general, a picture is worth more than a thousand words.

Before looking at the different types of visual models, I would like to stress that we are dealing with modelling techniques, not visualisation techniques in a broader sense. In biology the concept of 'visualisation' encompasses (in addition to visual modelling) both techniques with which biology can be made visible, *i.e.* a wide range of microscopy modalities, and techniques with which this visual information can be captured, *i.e.* imaging techniques. Whereas these techniques try and capture reality without simplification or selection, models specifically reduce information from reality to a set of essential features. These simplifications can have various purposes; they can be used in education, to describe our knowledge of biology, or to present hypotheses.

There are various types of visual models in developmental biology. Here I present a division into three basic categories: schematic representations, diagrams and graphical representations of tabular data, *cf.* Table 1.2.

Schematic representations

The first category comprises schematic representations; these are simplified yet faithful reproductions of biological reality. Examples in 2D are the **schematic drawings** of the turtle heart shown in Figs. 2.1, 2.2 and 2.3 (in the next chapter), drawings of animals or plants found in identification guides and cutaway drawings of anatomical structures. They are true to life (albeit in a simplified form) and are widely used to clarify descriptions or present hypotheses about the nature of structures in (developmental) biology. In three dimensions, schematic representations can be **tangible 3D models**, like the papier-mâché models made by Louis Auzoux in the 19th century or the plastic models used by medical specialists to help patients understand; however, they can also be **virtual 3D models**, like the 3D computer reconstructions of hearts presented in chapters 2 and 3. Virtual 3D modelling has become increasingly important over the last decennia with new advances in computer science leading to more sophisticated software, *e.g.* Mimics (www.materialise.com/mimics), Amira (www.amira.com), BioVis3D (www.biovis3d.com) and the software used to render the models in chapters 2 and 3, TDR3Dbase (Verbeek and Huijsmans, 1998). Computer graphics enable us to display 3D structures using sophisticated rendering algorithms. Two of those commonly used algorithms are volume rendering and surface rendering. Whereas volume rendering does not explicitly define the surface geometry, surface rendering represents the surface by polygons, most often using triangulation (Verbeek and Huijsmans, 1998; Eils and Athale, 2003), *cf.* chapter 2. In addition to making 3D structures more insightful, the conversion of 2D image sections to 3D surface models allows the user to perform quantitative analyses on 3D features such as volume and shape (Eils and Athale, 2003).

These 2D and 3D schematic representations are descriptive and static. They are qualitative and generally theoretical, although, as mentioned, quantitative analyses can be feasible in 3D models. Furthermore they are essentially spatial models, without a temporal aspect. However, by putting several of these representations (either 2D or 3D) of different (developmental) moments together a temporal element can be introduced. Chapter 2 presents a developmental series of turtle heart development, using models of different developmental stages. In this way the static representations are combined in a temporal model, albeit still consisting of static, discrete elements. Other examples of these types of **time series of schematic representations** would be cell or life cycle sequences. By animating the development in a continuous sequence (a movie), the model becomes dynamic. Nowadays **animations** of biological events can be found on numerous scientific websites, for instance the website of *Molecular Movies* (www.molecularmovies.com). Even so, all models in this category, including the

temporal ones, are still in essence descriptive rather than mechanistic; they merely illustrate processes and structures as opposed to explaining them.

<u>Diagrams</u>

The second category of visual models contains diagrams; here I will define a diagram as an abstract graphical representation of a process[3]. The keyword in this definition is 'abstract'; whereas schematic representations are true to life and directly reflect reality, diagrams reflect information about reality in an abstract, indirect way, using conventional symbols.

Two main uses of diagrams can be distinguished in developmental biology. Firstly they can be used informally to illustrate our knowledge of different steps of a process. As such, **informal diagrams** are particularly helpful in the relatively new field of systems biology. Signal transduction pathways, gene regulatory networks and metabolic pathways are frequently illustrated in diagrams. Symbolic representations of the elements in the process clarify the reactions taking place. These models are informative and help us visualise processes but cannot be used for computational analysis.

Secondly, diagrams can be used as **graphical visualisations of formal computational models**. In their most elementary form these diagrams are graphs consisting of nodes, connected by arcs. Depending on the focus of the model the nodes can represent (amongst others) classes, objects, processes or states. These diagrams can be classified as **class diagrams**, **object diagrams**, **flow charts** or **state diagrams,** respectively. As described in 1.4, there exists a wide range of computational modelling techniques which allow the user to model and analyse processes underlying biological events. Many of these techniques have a graphical component, illustrating the modelled process (Biermann *et al.*, 2004). The modelling technique of Petri nets, which is discussed in depth in chapters 5 and 6, is for instance characterized by the combination of graphical notation and exact mathematical definitions, to which we will come back in 1.4.1. Here the diagram, showing the entities and events of the process, contributes to understanding the model. In addition to this graphical component, which focuses on production and consumption of local resources and their transitions, the information contained in a Petri net can also be visualised in state diagrams, showing the different states and state changes of the system as a whole. Both visualisations belong to the category of formal diagrams.

---

[3] Note that in colloquial use the difference between a schematic representation and a diagram is rather hard to define. Here, I will not dispute the exact meaning of these words; the narrow definitions given are solely meant to clarify the classifications of model types presented here.

Another example of a diagram connected to an underlying formal model would be the visualisation of an ontology; here the focus lies on the classes and their mutual relationships and this type of diagram could therefore be categorized as a class diagram (we will look at the visualisation of ontologies in more depth in chapter 7).

An important modelling system consisting of formal diagrams is the object-oriented Unified Modelling Language (UML, www.uml.org). UML comprises fourteen types of diagrams, amongst which class diagrams, object diagrams and activity diagrams (which are related to flow charts). By assigning formal meanings to visual symbols, this language and formal diagrams in general allow the user to formally communicate knowledge. As is often the case with modelling techniques, UML was developed for different purposes (the modelling of computing systems), but has recently been shown to be useful in the field of biology as well (Roux-Rouquié *et al.*, 2004).

Since diagrams are most frequently used to present the development of a process, they often have a time element. This is not necessarily the case however, as exemplified by ontology visualisations and certain types of UML models. Also informal models depicting pathways and networks in systems biology cannot be said to have a true time component; rather, different processes, that are in some way linked, are shown in one diagram. On the other hand informal diagrams that represent a clear sequence of events with time indications, like life cycles, are in fact temporal models. It should be underlined that the difference with the temporal sequences mentioned under 'Schematic representations' is that diagrams consist of abstract symbols, whereas the temporal sequences discussed earlier are composed of truthful representations.

Graphical representations of tabular data

The third and last category of visual models consists of graphical representations of tabular data; tabular data are alphanumerical data (quantitative of qualitative) that can be stored in rows and columns in a database, *e.g.* concentration levels of proteins over a distance in a cell culture, but also the names of research group members in charge of cleaning the lab over a series of months. These types of data can be modelled in a variety of **charts**, such as **histograms**, **bar charts**, **graphs**, **scatter plots** or **cartograms**. Since these charts are simplified visual representations of selective information from reality, I include them in our overview of visual models. These models can be multi-dimensional; simple charts plotting two variables are two-dimensional, but more dimensions can be added for more variables. Models of tabular data are descriptive and static; they are practical (pertaining to specific situations) and can be either quantitative or qualitative, modelling spatial or temporal information, or both.

1.4 ALGORITHMIC AND EQUATION BASED MODELS – COMPUTATIONAL MODELLING

As we have seen, verbal and visual models have been used for centuries in the field of developmental biology. More recently new ways of modelling biology have been developed; integration of biology with the field of mathematics and the relatively new field of computing science has led to models that allow us to study biology through calculation and simulation. For this new area of biology the term 'computational biology' has been coined, which is defined by the NIH as "the development and application of data-analytical and theoretical methods, mathematical modelling and computational simulation techniques to the study of biological, behavioural and social systems" (NIH, 2000)[4]. This definition is followed throughout this chapter. The essential and binding element of the models discussed in the remainder of our overview is therefore computation.

For computational modelling the same holds true as for the entire field of models in developmental biology: models can be categorised in more than one way. Here a distinction is made between **algorithmic process models** and **equation based models**. While both types are concerned with biological processes, the difference between models based on algorithms and those based on equations is a difference of focus on **operation** versus **denotation** (Fischer, 2007; Priami, 2009); *i.e.* equations abstract away from the actual process and solely describe the changes in terms of variable values when a system moves from one state to another, making them denotational. They model systems by functions which transform input into output. Algorithms, on the other hand, focus on the process and describe how and why the system changes between states, which makes these models operational. The distinction between algorithmic process models and equation based models corresponds to the distinction Fischer (2007) makes between computational and mathematical models, respectively; since the terms ´mathematical´ and ´computational´ can both be used in various and sometimes interchangeable ways, these classifications are avoided here and the word ´computational´ is used only in the broad sense, as specified by the definition given in this chapter.

The distinction between algorithm and equation based models is directly related to three important pairs of polar opposites in process models :

1) **discrete versus continuous**; algorithmic models describe a system in discrete state spaces (*i.e.* the set of all states a system can be in), whereas state spaces in equation based models are generally continuous (although they can be discretised, *cf.*6.1). However, there are hybrid techniques, for instance Hybrid Petri Nets (David and Alla,

---

[4] This definition is slightly circular, for it does not explain the term ´computational´. Here we understand this term to refer to using either calculation or a computer.

2001; Matsuno *et al.*, 2003), in which both discrete state spaces and continuous variables are used.

2) **concurrent versus sequential**; most processes in developmental biology are characterised by concurrency (*cf.* chapters 5 and 6), making it an important factor in the choice of a modelling technique. Algorithmic models are typically applied when modelling cause and effect which can lead to concurrency or nondeterminism. Equations on the other hand lend themselves more readily for modelling sequential processes, in which input information determines global behaviour. These models are strictly functional (*i.e.* based solely on mathematical functions; specific input always yields the same output) and are therefore deterministic.

3) **averaged versus individual**; equation based models generalise the process behaviour by describing the average global behaviour of the system. By contrast, for algorithms each simulation describes the exact behaviour of one execution of the process.

With these distinctions in mind, we will now look at some algorithmic and equation based models, commonly used in developmental biology.

1.4.1 Algorithmic process models

Algorithmic process models focus on describing the dynamics of systems. As such they are particularly useful in developmental biology, where researchers are most often interested in the biological process itself. Over the past decades several algorithmic modelling techniques have been developed, often originating in the computing sciences, and have come to be used in the field of bioinformatics.

In a highly simplified form the behaviour of a system can be described as follows: at the start of the process an input state is provided and this is followed by an output state; in this context a state is a particular configuration of the information in the system. This simplification forms the basis for the development of **automata theory** (Hopcroft *et al.*, 2000). In this field processes are seen as automata, which are abstract **state machines** for which a number of states is provided along with a set of rules, which govern transitions from one state to another, taking the current state as input. In addition to the set of states and transition rules also one or more initial states and final states are specified. The process starts off in an initial state and for each next step the provided rules determine the next state of the system, until a final state is reached.

John von Neumann extended this system and developed **cellular automata** (Von Neumann, 1966), which have several biological applications. Cellular automata consist of a grid of identical, interdependent automata, referred to as cells. Again each

of the cells is in a certain state and its next state is determined by a set of rules. Here the rules take as their input not only the current state of that particular cell but also the states of a set of neighbourhood cells (and possibly external factors); this allows local interaction between cells to be modelled, making the technique suitable for modelling more complex systems, characterized by concurrent and locally determined behaviour. In developmental biology cellular automata have commonly been used to study the formation of patterns, such as pigmentation patterns in seashells (Meinhardt, 1982; Plath *et al.,* 1997).

A related modelling system was developed by Aristid Lindenmayer in 1968, the **Lindenmayer system** or **L-system** (Lindenmayer, 1968). This system resembles the system of cellular automata but while cellular automata only change the states of cells but never remove or add cells, Lindenmayer systems allow the appearance of new cells and the disappearance of existing cells. Put differently, cellular automata use *state transition rules*, while Lindenmayer systems use *rewriting rules*. This makes Lindenmayer systems interesting for biologists looking to model growth and development; due to this feature L-systems have been used successfully to model plant development, *e.g.* through arborisation (Prusinkiewicz, 1995).

Another related approach is the **boolean network**, which in a sense is a generalization of the cellular automata. The network consists of nodes, analogous to the cells in cellular automata, but while the next states of cells in cellular automata depends on the states of the direct neighbours, nodes in a boolean network can be connected to and therefore influenced by any other node in the network. The nodes can be in one of two states, one or zero, and as in cellular automata the next state of a particular node is determined by the states of the connected nodes (and the current state of that node), depending on certain logical functions. This system was originally developed by Stuart Kauffman to model genetic regulatory networks (Kauffman, 1969), in which genes are either active (in state 1) or inactive (in state 0) and interact through the network. These modelling systems can be used more generally to study the interaction of biological networks, consisting of *e.g.* proteins and genes. Although this technique strongly simplifies the biological process, it can still yield useful results regarding pathways and gene regulatory networks (Fischer and Henzinger, 2007). A downside to this technique is that the networks cannot be integrated into larger models, *i.e.* they cannot be constructed from smaller sub-parts which are combined in a hierarchical structure.

In contrast to this, **process algebras**, or **process calculi**, allow a hierarchical organization of sub-parts, *i.e.* compositional modelling. This field of modelling is concerned with studying concurrent or distributed systems by algebraic means, *i.e.* the methods of algebra are applied to the behaviour of a system (Baeten, 2005). The compositionality of the system implies that the behaviour of the system as a whole is expressed through the behaviour of its components and their interaction. This

interaction can be synchronous or asynchronous (with components changing states independently), the latter resulting in nondeterminism since different sequences of events might lead to different outcomes (Fischer and Henzinger, 2007). These aspects in particular make process algebras suitable for modelling biological behaviour; in biology, as in distributed systems, many processes are active simultaneously and compete over the use of resources while also cooperating to accomplish a common goal (Priami and Quaglia, 2005). Process algebras have been used successfully in modelling biological pathways and networks within the fields of systems biology and developmental biology (Priami and Quaglia, 2005; Fischer and Henzinger, 2007).

A strong emphasis on concurrency and local independence can also be found in **Petri nets**, which have been mentioned before in 1.3.2 and will be discussed more thoroughly in chapters 5 and 6. A Petri net is an abstract model of information flow. It has both a graphical representation and a formal mathematical definition. The model consists of places (representing passive entities like resources), transitions (representing actions or events) and the flow relation between them. Graphically the net is presented as a bipartite directed graph with places depicted as circles and transitions depicted as rectangles; these two types of nodes are connected by edges, depicting the flow relations. The state of the net is represented by the distribution of tokens (black dots) over the places and the dynamic behaviour of the system is represented by the production and removal of these tokens, governed by occurrences of the transitions. The graphical component makes the net intuitively understandable, while the formal definition enables precise analysis. Petri nets are well-established as models of concurrent and distributed systems and have recently gained importance in the field of biology, in particular biochemistry and developmental biology (Steggles *et al.*, 2007; Heiner *et al.*, 2008; Krepska *et al.*, 2008).

This list of algorithmic process models for developmental biology presented here is far from complete. In addition to lesser known models already in use, new modelling techniques are constantly being developed. Furthermore, already existing modelling techniques, used for instance in computing sciences, are being 're-used' for biological purposes on a regular basis. As is obvious from their focus on processes, algorithmic models are dynamic. There are both quantitative or qualitative algorithmic models (*cf.* chapter 6), both descriptive or mechanistic (depending on whether their representation of the underlying biological process is true to life or solely meant to produce correct outcomes), theoretical or practical and deterministic or nondeterministic. Finally, they may or may not include temporal and spatial elements.

1.4.2 Equation based models

The final category of models, discussed in this overview, comprises the equation based models. These do not model the stepwise executing of natural processes the way algorithms do and allow exact quantitative analysis. As such they complement the qualitative process models. By combining the analysing strength of equation based models with the operational focus of algorithmic models, developmental processes can be studied in detail from different perspectives.

Equations can help study a great number of processes in biology in general and developmental biology in particular. Of special importance are **differential equations**, which constitute by far the most used formalism to model dynamical systems (De Jong, 2002; Priami and Quaglia, 2004; De Jong and Geiselmann, 2005) and have been applied to a variety of biological phenomena, of which gene regulatory networks, reaction-diffusion systems and diffusion models will be discussed here.

The first category of cellular processes, relevant to developmental biology and modeled using differential equations, comprises **gene regulatory networks**. The process of gene expression takes place in two steps: transcription (in which 'reading out' a string of DNA produces a complementary mRNA copy in the cell nucleus) and translation (in which proteins are produced by 'reading out' the mRNA at the ribosomes in the cell). Both processes are influenced by additional reactions and proteins; in order for transcription to take place a polymerase has to attach to the promotor site of the string of DNA and, upon transcription, to detach at the terminator site. This binding of the polymerase to promotor is regulated by other proteins, both repressor proteins (inhibiting transcription) and activator proteins (actively promoting transcription). These repressor and activator proteins are themselves products of gene transcription and translation and the combined interactions between these proteins and gene expression patterns are captured in gene regulatory networks. As is obvious from this description, these networks can become very complex, all the more so since they often include positive or negative (self) feedback loops. A simple example of such a loop is a gene which codes for a protein that actively inhibits the expression of that same gene.

These complex processes play an important role in the field of systems and developmental biology and a vast amount of models exist, mainly focused on differential equations (DEs) (De Jong, 2002; De Jong and Geiselmann, 2005). Nonlinear ordinary DEs model gene regulation by rate equations, in which the rates of change in space and/or time of one component in the system is expressed as a function of the concentrations of other components. The regulatory interactions of the network under study are captured by functional and differential relations between these concentration variables. Depending on the number of independent variables of the unknown function,

either ordinary DEs (in case of a single variable) or partial DEs (in case of multiple variables) are used.

Nonlinear ordinary DEs give an adequate description of the dynamics of gene regulation, but can be mathematically difficult to analyse. Linear ordinary DEs simplify the system and are easier to analyse, but this is countered by a reduced ability to include essential properties of the system (De Jong and Geiselmann, 2005). An intermediate solution is provided by piecewise-linear DEs, which are globally nonlinear and locally linear. They simplify the system, by abstracting biochemical details and approximating the continuous sigmoid curves of the gene behaviour in discontinuous step functions, but still retain the ability to adequately represent the system (De Jong and Geiselmann, 2002).

Apart from gene regulation many other biological processes can also be modelled using differential equations. Another type of model concerned with biochemical interactions is the **reaction-diffusion system**. This model was proposed by Alan Turing in his influential paper of 1952 (Turing, 1952), in which he showed how the interaction of two substances with different diffusion rates can lead to pattern formation. In developmental biology this corresponds to a system of chemicals, called morphogens, diffusing through a tissue and interacting, which results in stable patterns underlying morphogenesis; patterns arise, of alternating high and low concentration areas of a morphogen, accounting for instance for pigment stripes (Gilbert, 2000). The wavelength of these patterns is determined by the diffusion constants and reaction rates of the different morphogens involved. A reaction-diffusion system can be modelled by partial differential equations, consisting of a diffusion component and a reaction component.

The reaction-diffusion system can be simplified by leaving out the reaction component, which results in a **diffusion model**. This describes the spread of particles from areas of higher concentration to areas of lower concentration, like the spreading of a drop of ink in water. In 1970, Francis Crick proposed to apply this model to the biological process of gradient formation (Crick, 1970). This is a pivotal developmental process, which plays an important role in the establishment of body axes and morphogenesis. He modelled this process as a source-sink system, in which a morphogen gets produced at the source, spreads through Brownian motion and gets removed again at the sink. This model has formed the basis for many DE models describing gradient formation (Gregor *et al.*, 2005; Kicheva *et al.*, 2007; Yu *et al.*, 2009). We will come across this use of DEs or gradient formation in chapter 6, in which the modelling of this process by DEs will be complemented by a Petri net model.

In addition to differential equation models, several other equaion based models are used for biology. Essential to many processes in the cell are enzyme-

mediated biochemical reactions, in which a substrate undergoes a reaction, instigated by an enzyme. The dynamics of these are described by **models of enzyme kinetics**. As early as 1913 Leonor Michaelis and Maude Leonora Menten developed an equation to describe the rate of single-substrate enzyme kinetics, based on the reaction rate, the concentration of the substrate and the Michaelis constant of that substrate (Menten and Michaelis, 1913). The importance of this model stems from its general applicability; it describes the reaction behaviour of thousands of enzymes and therefore allows biologists to study and predict enzyme reactions in a wide range of biological processes (Jungck, 1997).

Finally, the important developmental process of growth is described by **models of organismal growth**. Here a distinction can be made between isometric and allometric growth. Isometric growth refers to an increase in volume and size, while retaining the original proportions, *i.e.* the shape is preserved because all parts grow at the same speed. This growth is for instance seen in the spiral growth of shells and can be expressed in equation models (Gilbert, 2000). Allometric growth or allometry refers to the process in which some parts grow at a different rate than others, thereby changing the overall proportions. Human allometry can easily be seen when comparing the head-limb ratio of babies to that of adults. Another clear example is the growth of the male fiddler crab; its claws are initially the same size, but as the crab grows the crushing claw becomes proportionally larger (Gilbert, 2000). The underlying laws of these types of growth can again be captured in mathematical equations.

These equation based models are first and foremost quantitative and deterministic. Their abstraction from the actual processes makes them descriptive and often practical rather than theoretical. Both spatial and temporal information can be included in these models; each of these properties will be discussed in chapter 6, when we will compare an algorithmic modelling technique with an equation based technique in order to develop new methods to combine the strengths of these approaches.

———

This concludes the overview of models used for developmental biology. It goes without saying that this modest chapter can by no means offer an exhaustive overview; many other models are used for a wide range of developmental phenomena. The outline of the dissertation follows the theoretical overview, provided in this introductory chapter. Each of the four model categories is addressed in a case study. Chapters 2 and 3 present several 3D models, visualising aspects of heart development in the turtle *Emys orbicularis*. These visual models illustrate the importance of 3D modelling in understanding complex anatomical structures. As such they belong to the

class of schematic representations, within the category of visual models. Chapter 4 concerns the construction of a multi-species ontology of the vertebrate heart, exemplifying the category of verbal models. In chapters 5 and 6 two modelling approaches of the process of gradient formation are presented, both in the modelling framework of Petri nets; while chapter 5 focuses on a qualitative approach, chapter 6 integrates parameters of differential equations into the Petri net model. These chapters are therefore connected to the categories of algorithmic and equation based models, respectively. Apart from exemplifying the different modelling categories, each of the studies presented in the following chapters also features integration of modelling methods. In the final chapter we look into these different types of integration, encountered in the presented studies.

CHAPTER 2.   A VISUAL MODEL: 3D RECONSTRUCTIONS OF CARDIAC DEVELOPMENT IN THE TURTLE *EMYS ORBICULARIS*

> The scientist gave a superior smile before replying, "What is the tortoise standing on?" "You're very clever, young man, very clever," said the old lady. "But it's turtles all the way down!"
>
> - Stephen Hawking, *A brief history of time*

We begin our series of case studies by looking at a visual model of developmental anatomy: a series of 3D reconstructions of the turtle heart. There are significant benefits to using a computer-assisted technique in studying development and the development of the turtle heart is particularly suited to illustrate these benefits; the embryonic turtle heart is very small, necessitating a high level of magnification, and is spatially complex. While schematic drawings, based on histological sections, can be misleading, 3D computer reconstructions closely resemble reality and allow us to get a better understanding of the spatial properties of the structure. Furthermore, by providing the models online with our TDR-viewer applet, the reader is able to interactively explore the reconstructions as well as the corresponding histological sections, which optimizes the communication of information by these models.

Based on: Bertens L.M.F., Richardson M.K., Verbeek F.J., 'Analysis of cardiac development in the turtle *Emys orbicularis* (Testudines: emidydae) using 3-D computer modelling from histological sections', *Anatomical record* 293:7 (2010), 1104-1114.

LIST OF ABBREVIATIONS

A, atrium; AA, aortic arch; ACV, anterior cardinal vein; AVC, atrioventricular canal; AVV, atrioventricular valve; CA, *cavum arteriosum*; CAVV, cushion tissue forming the atrioventricular valve; CP, *cavum pulmonale*; CCV, common cardinal vein; CV, *cavum* ventrale; DC, distal cushions; F, foramen; HS, horizontal septum; HSt, heart stalk; IAS, interatrial septum; IHC, inner heart curve; IVC, intraventricular canal; LAA, left aortic arch; LA, left atrium; LAVC, left atrioventricular canal; MC, mesenchymal cap; OFT, outflow tract; PA, pulmonary artery; PC, proximal cushions; PCV, posterior cardinal vein; PM, pectinate muscle; PV, pulmonary vein; RAA, right aortic arch; RA, right atrium; SV, sinus venosus; TC, trabeculae carneae; V, ventricle; VS, vertical septum; VV, possible venous valve of the superior cardinal vein.

2.1 INTRODUCTION

Many features of cardiac development in turtles correspond to those in birds and mammals (Goodrich, 1958; Holmes, 1976; Quiring, 1933; Shaner, 1962; Greil, 1903). These shared features include cardiac looping, ventricular trabeculation, formation of an inner heart curvature and presence of a major pair of atrioventricular endocardial cushions that give rise to the atrioventricular valves. Other aspects of the turtle heart, however, differ greatly from the avian and mammalian heart. The septation of the ventricle for instance is remarkably different in the turtle and the developmental origins of this septation are still not fully understood. As opposed to the fully separated hearts of mammals and birds, the turtle heart is only partially separated by two incomplete septa. An attempt at explaining the development of these horizontal and vertical septa (see below) has been made by Holmes (1976), but no conclusive study of this process exists.

The aim of this chapter is twofold; the main biological purpose is to elucidate the embryonic development of the horizontal septum in the turtle, which is shown to be closely linked to the looping of the heart. To this end we present a detailed developmental series of cardiogenesis in *Emys orbicularis*, the European pond terrapin, clarifying the development of the horizontal septum and the looping of the heart tube. Because of the intricate 3-dimensional nature of cardiogenesis and especially the looping, studying histological sections in merely two dimensions does not suffice. For that reason we have complemented the histological studies with computerised high-resolution 3D reconstructions of three different developmental stages, produced with our in-house software environment, TDR-3Dbase (Verbeek *et al.*, 1995; Verbeek and Huijsmans, 1998; Verbeek, 1999; Verbeek and Boon, 2002).

The second aim of this chapter is therefore to underscore the indispensable nature of high-resolution 3D reconstructions in studying embryonic morphology. Our software is particularly well suited for solving complex biological problems in three dimensions, as it enables the construction of high-resolution 3D models, directly from histological sections, and allows us to distinguish between different structures in the heart and study them separately and from every viewpoint. To offer the reader full access to the material, we provide the 3D reconstructions along with additional images on a public website: http://bio-imaging.liacs.nl/galleries/. Here the models can be viewed interactively using our 3D model browser, TDR-viewer, a Java applet which complements our TDR-3Dbase software package and enables online viewing of 3D reconstructions. This ensures optimal insight into the obtained results and serves as an easy reference for anyone studying cardiogenesis. To the best of our knowledge, this is the first application of these modern techniques to the study of the chelonian heart, enabling the study of all structures separately.

Besides the findings on the ventricular septation we have, in the course of this study, also come across an interesting finding regarding the developmental origin of the pulmonary vein, which adds to the current debate on the origin of this vein in mice, humans and chicks and is described in the conclusions.

Below, a review of the available literature on the adult chelonian heart is provided, followed by a detailed description of the developmental series, using both histological studies and 3D reconstructions. The construction methods of the models are explained and finally the results of the study are discussed, considering both new insights into the ventricular septation of the turtle heart and the relevance of using 3D reconstruction techniques in supporting biology.

2.2 CURRENT KNOWLEDGE OF THE ADULT TURTLE HEART

The reptilian order Testudines includes the aquatic and semi-aquatic 'turtles', as well as the terrestrial 'tortoises' (Family Testudinidae). We are concerned here with a species of the former group, the European pond turtle, *Emys orbicularis* L. For a long time, researchers considered the turtle heart a transitional stage between the single circulatory system, found in most fish, and the double circulatory system of lungfish and tetrapods, including mammals and birds. As such it was considered inefficient and 'unfinished', only a stepping stone in the evolution from using gills (*i.e.* branchiate respiration) to using lungs (*i.e.* pulmonate respiration) for gaseous exchange. This idea originated from the observation that most reptiles have only partly separated ventricular *cava*, whereas mammals and birds have completely separated ventricles. At present it is commonly agreed that the turtle heart is in fact a highly specialized organ, adapted to its typical function in turtles — animals that possess lungs, but spend much of their time under water.

Like other tetrapods, the adult turtle possesses two thin-walled atria lined with pectinate muscles and separated by an interatrial septum (*cf.* Fig. 2.1). The sinus venosus is partly absorbed into the right atrium and receives the systemic veins, (Fig. 2.2; Gasch, 1888; O'Donoghue, 1918; Rau, 1924). The pulmonary vein opens into the left atrium. The atria open through the atrioventricular canal. This canal is divided into right and left atrioventricular junctions by the free margin of the atrial septum (Johansen and Burggren, 1980; Burggren and Warburton, 1994), from which hang two leaflets with rudimentary chordae tendinae. Nayak *et al.* (1995) suggest that the two leaflets correspond functionally to the septal leaflet of the human right atrioventricular valve

Cranial

RA —
RAA —
LAA —

CV —

HS —
CP —

PA
LA

LAVC

CA
AVV
VS

Caudal
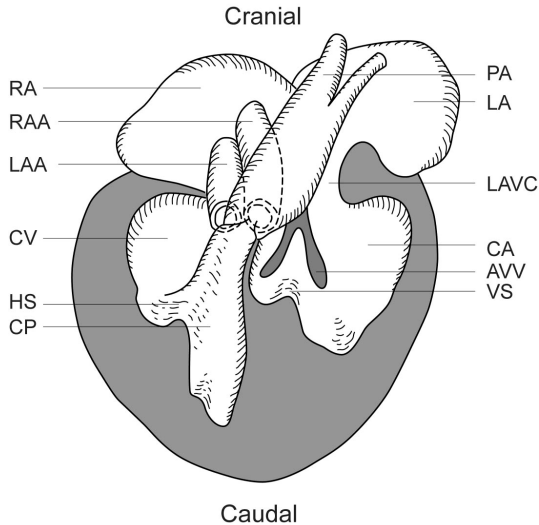
Figure 2.1. Schematic drawing of the adult testudine heart, shown in ventral view (drawing based on Hicks and Wang, 1996). For the meaning of abbreviations, see List of abbreviations.

and the aortic leaflet of the human left atrioventricular valve. The valve leaflets are single-flapped and membranous and develop from endocardial cushions which meet and fuse along the free end of the developing interatrial septum (Goodrich, 1958). The valves project into the ventricular opening (Goodrich, 1958).

The interior surface of the apical region of the ventricle is heavily trabeculated and the ventricular cavity is partially divided by two incomplete septa into three cavities (Fig. 2.1, 2.2 and 2.3). A large muscular septum, the 'horizontal septum' demarcates the cavity known as the *cavum venosum* from the *cavum pulmonale*, while a smaller 'vertical septum' separates the *cavum venosum* from the *cavum arteriosum;* in the figures the cavities are referred to as CV, CA and CP.

The horizontal septum (in the horizontal plane, from left to right sides of the ventricle) is complete on the caudal end of the ventricle, but incomplete on the cranial end. It is sometimes called 'Muskelleiste' in German, or 'muscular ridge' in English (Webb *et al.*, 1971; Burggren, 1988; Hicks and Wang, 1996; Van Mierop and Kutsche, 1984; Van Mierop and Kutsche, 1985). However, like Holmes (Holmes, 1976) we consider these terms rather confusing, since 'Muskelleiste' can be translated as 'muscular ridge', which in turn is a term used for both vertical septum and in general for all trabeculae carneae. The septum is also referred to as interventricular septum, but this is equally misleading, since this term is also commonly used for the vertical septum. Therefore we will call this septum, as do many other authors, the horizontal septum.
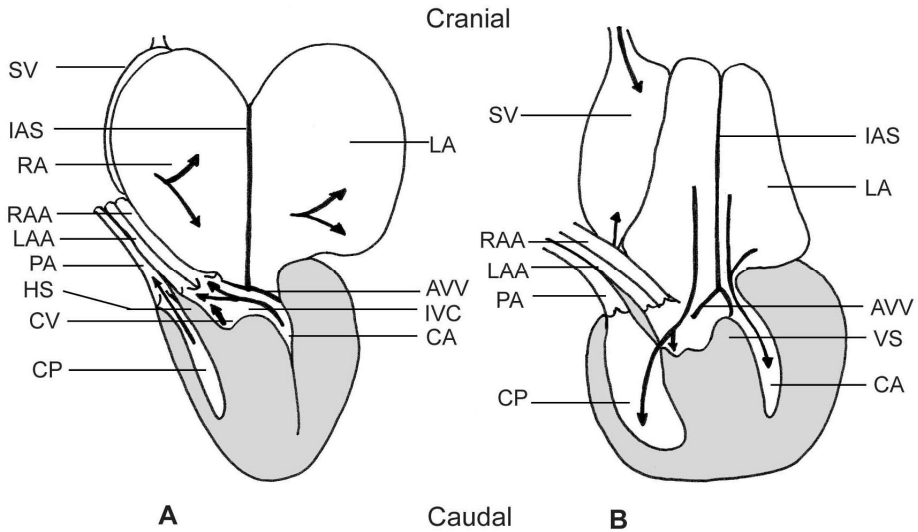
Figure 2.2. Schematic drawing of the blood flow in the chelonian heart shown in ventral view, during atrial systole (A) and ventricular systole (B); note that in order to illustrate the blood flow, the *cavum pulmonale* has been schematically drawn next to the *cavum venosum*, as opposed to ventro-laterally (drawing based on Webb, 1971).
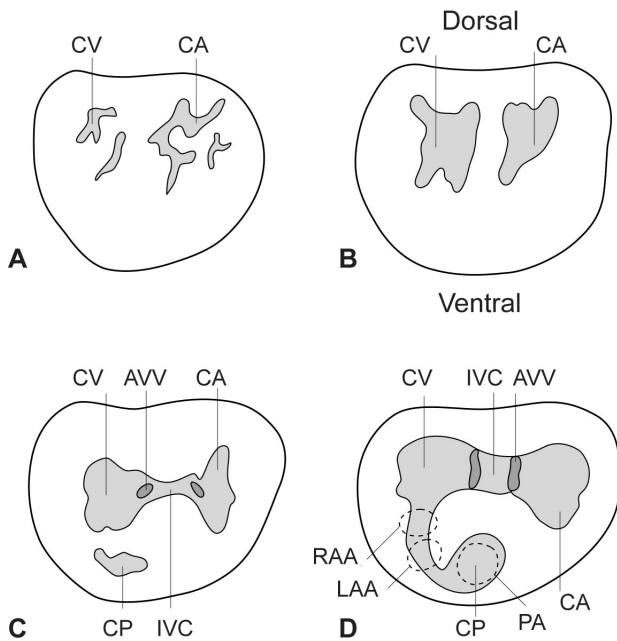
Figure 2.3. Schematic drawings of transverse sections through the adult testudine heart (shown in cranial view), at different planes of section along the craniocaudal axis; starting near the ventricular apex, at the caudal end of the heart, in A, and progressing in the cranial direction up to the intraventricular canal, in D (drawing based on White, 1968).

The vertical septum runs dorso-ventrally, in the same sagittal plane as the interatrial septum (which means it runs vertically when the animal is on its feet). It is complete at the caudal apex of the ventricle, but incomplete at the cranial end of the ventricle, leaving a passage open between the *cavum venosum* and the *cavum arteriosum.* The septum is thought to develop from one of the trabeculae carneae and has a similar structure as the trabeculae (Holmes, 1976). Its size differs between species (Webb *et al.*, 1971; Holmes, 1976). Its relation to the interventricular septum in chick and mouse has recently been studied on the gene level by Koshiba-Takeuchi *et al.* (2009); they found a significant difference in the gradient formation of *Tbx5* in the ventricle between reptiles on the one hand and chick and mouse on the other, suggesting that a steep left-right gradient early in development is essential for ventricular septation. In reptiles this gradient is delayed and less pronounced, resulting in varying degrees of partial septation by the vertical septum.

The *cavum arteriosum* lies on the left side of the vertical septum. This cavity receives blood from the left atrium through the left atrio-ventricular canal. Over the free margin of the vertical septum, we find a passage from the *cavum arteriosum* on the left to the adjacent cavity, the *cavum venosum*, on the right. This passage is situated under the atrioventricular valves and is called the intraventricular canal (Johansen and Burggren, 1980). The canal was first described by White (White, 1968), but there, in our opinion erroneously, called the interventricular canal. Using the term 'interventricular' implies the existence of two completely separated ventricles, which is the case in mammals, birds and crocodilians, but not in turtles. This intraventricular canal is the only exit from the *cavum arteriosum*, since this cavity does not connect directly to any arteries leaving the heart. During atrial systole the intraventricular canal is closed by the atrio-ventricular valves, which are forced caudally, into the ventricle, by the pressure of the blood flowing from atria into the ventricle. In this way all blood from the left atrium is pumped into the *cavum arteriosum* (White, 1968; Johansen and Burggren, 1980; Hicks and Wang, 1996). This process is illustrated in Figure 2.2A.

The *cavum venosum* lies to the right of the atrio-ventricular canal and according to Webb (Webb *et al.*, 1971) it is actually more of a canal than a cavity for retention of blood, but others claim that this is in fact the largest cardiac cavity (Burggren, 1988). Together the *cavum arteriosum* and the *cavum venosum* are sometimes called the *cavum dorsale* (White, 1968), since they occupy the most dorso(-lateral) part of the ventricle (*cf.* Fig. 2.3). As can be seen in Figure 2.2A blood from the left atrium is pumped into the *cavum venosum* during atrial systole and continues into the *cavum pulmonale* (described below). The *cavum venosum* is separated from the *cavum arteriosum* by the atrio-ventricular valves, which are forced caudally (White, 1968; Johansen and Burggren, 1980; Hicks and Wang, 1996).

The third ventricular cavity is the *cavum pulmonale*. This cavity is also referred to as the *cavum ventrale* (White, 1968). It lies deeper (*i.e.* protrudes more apically) than the other cavities, lies ventro-laterally (on the right) of the other *cava* and is connected to the *cavum venosum*. The *cavum pulmonale*, separated from the rest of the heart by the horizontal septum, has a peculiar location, since neither of the atrio-ventricular canals is directly connected to it. Therefore, the only way blood can reach the *cavum pulmonale* is through the *cavum venosum*, over the free edge of the horizontal septum (Holmes, 1976), as seen in Figure 2.2B.

In turtles and squamates, three arterial trunks arise from the ventricle, a right and a left aortic arch and a pulmonary trunk (Van Mierop and Kutsche, 1984; a thorough description is given in 3.3.3, along with several visualisations of the arches in Fig. 3.3). Each of these trunks possesses a pair of bicuspid semilunar valves (Fig. 2.2; Goodrich, 1958; Burggren, 1988), originating from endocardial cushion tissue in the outflow tract. The carotico-systemic aortic arch, *i.e.* the right aorta, begins at the *cavum venosum*, dorsal to the horizontal septum (Fig. 2.3). The systemic arch, *i.e.* the left aorta, begins somewhat to the right and slightly more ventrally, near the free margin of the horizontal septum, but still originating from the *cavum venosum*. The pulmonary trunk arises ventrally to the horizontal septum from the *cavum pulmonale* (Holmes, 1976).

During ventricular systole deoxygenated blood from the *cavum pulmonale* enters the pulmonary trunk, whereas oxygenated blood from the *cavum arteriosum* enters the aortic arches, as shown in Figure 2.2B (White, 1968; Johansen and Burggren, 1980; Hicks and Wang, 1996). However, during periods of apnoea an increase in the pulmonary vascular resistance results in a shunt, which causes a decrease of pulmonary blood flow. This means that when the turtle is under water (and will therefore not benefit from pulmonary circulation) most of the blood in the *cavum venosum* will end up in the aortic arches, as opposed to the pulmonary trunk (White, 1968; Hicks and Wang, 1996).

In crocodilians, possessing complete interventricular septa, a similar shunting of blood can be found; like other reptiles, crocodilians possess two aortic arches, of which the left emerges from the right ventricle (along with the pulmonary arch). Communication between these aortic arches is possible through a foramen between the two arches, the foramen of Panizza, and an aortic anastomosis (Van Mierop and Kutsche, 1984; Van Mierop and Kutsche, 1985; Axelsson *et al.*, 1996; Axelsson, 2001; Eme *et al.*, 2009). Through this communication shunting can arise, resulting in a pulmonary bypass; this resembles the shunting found in turtles, due to changes in pulmonary vascular resistance.

During embryonic development, the turtle outflow tract initially possesses two proximal and two distal cushions; two additional distal cushions arise later (Hart, 1968;

Langer, 1894; Van Mierop and Kutsche, 1984). Similarly, in the mammalian heart two proximal cushions and four distal ones are observed (Webb *et al.*, 2003); however, the development and fusion of these cushions differs from that seen in reptiles (Van Mierop and Kutsche, 1984). A full account of these differences is given in chapter 3, in which we will look at the development of the cushions and the septation of the distal component of the outflow tract.

2.3 MATERIALS AND METHODS

The process of constructing the 3D models can be described in three main steps. First, suitable specimens were prepared and sectioned. Subsequently, the sections were digitized in order to build an acquisition database. Finally, this acquisition database was used to construct a database containing delineations of the anatomical domains relevant to the study. This database contains the 3D model of the heart and was used to generate the visualisations. For each model a database was created.

Preparation of the embryos

A total of 12 embryos was studied. Three embryos were used to construct the 3D models, of stages 8, 10 and 15 (staging according to Yntema, 1968). These stages were chosen because they span the period in which the basic plan of the heart is appearing, but has not fully differentiated yet. In addition to these three stages, serial histological sections of 9 embryos were also studied of the following Yntema stages: 6, 8, 9, 10, 11, 12, 14, 15 and 16. Stages 6 to 11 correspond to the following stages in chick (HH stages; Hamburger and Hamilton, 1951), mouse (Theiler stages; Theiler, 1989) and human (Carnegie stages; O'Rahilly and Müller, 1987): Yntema 6 corresponds to HH 9, Theiler 13, Carnegie 10. Yntema 8 corresponds to HH 16-17, Theiler 14, Carnegie 11. Yntema 9 corresponds to HH 18, Theiler 14, Carnegie 11. Yntema 10 corresponds to HH 19-20, Theiler 15, Carnegie 12. And Yntema 11 corresponds to HH 20, Theiler 13, Carnegie 10. Gravid females were collected under license from the French government. Standard injection with oxytocin was used to induce laying. Eggs were placed on a layer of sand in an incubator at 25-30°C. Embryos were fixed in Bouin's fluid for 2 days and embedded in Fibrowax according to standard protocols, serially sectioned at 7µm and stained with Haematoxylin, Eosin and Alcian Blue.

Acquisition of section images

Images of the sections containing the heart (of the three embryos used for the 3D modelling) were made using a Zeiss Axioskop microscope (Zeiss, Jena Germany), equipped with a JAI M10-RS CCD camera (JAI, Denmark) and a Marzhauser stage controller (Marzhauser, Germany) using the MAC 4000 (Marzhauser, Germany). The

MAC 4000 connected to the computer through the RS-232 interface. The CCD camera, in combination with a PCVision frame grabber (Imaging Technology, MA, USA), was used for image digitization. Our acquisition software, 3Dacq, version 2.0 (Verbeek *et al.*, 1998; Verbeek and Huijsmans, 1998; Verbeek, 1999; Verbeek and Boon, 2002), controlled both the MAC 4000 and the PCVision frame grabber. Using the video overlay option of the PCVision frame grabber, images were aligned prior to acquisition. The metadata was stored in an XML acquisition database referring to the section images which were stored in PNG format.

<u>Construction of 3D models</u>

The annotation and subsequent visualisation of the resulting 3D models was realized with a software suite developed in our group for this purpose, the TDR-3Dbase program (Verbeek *et al.*, 1995; Verbeek and Huijsmans, 1998; Verbeek, 1999; Verbeek and Boon, 2002). The acquisition database, with the section images and other relevant information, is seamlessly integrated in TDR-3Dbase; all relevant information is directly transferred. In each of the section images the relevant anatomical structures, labelled with separate colours and names, were traced using a WACOM LCD tablet (PL series, WACOM, Europe). Each contour was stored as an annotation to the section image in a TDR-3Dbase database. This database was then used in TDR-3Dbase to visualise the anatomical structures of the heart, using surface models (Verbeek *et al.*, 1995). These surface models were derived through an automated triangulation procedure using the contour models in the database.

The TDR-3Dbase software allows one to work in different geometrical representations; the storage is optimized for each representation and allows easy publication on internet. The data do not have to be transformed as is the case with other software environments (Ruthensteiner and Hess; 2008). The 3D models are made available via internet using our 3D model browser application, TDR-viewer (Potikanond and Verbeek, 2012). This viewer is a java applet and requires that java 3D is installed on the local host, *i.e.* the user's computer. Instructions and links are provided from the gallery pages. The viewer is able to use all the available geometrical model information so that the user has maximal freedom in inspecting the 3D models. The viewer opens with a 2D view (Fig. 2.4C) while on request the 3D view is loaded. Selections of the separate anatomical domains can be made from the interface and will be instantaneously altered; the 3D views are interactive (Fig 2.4A) and allow for active exploration of the 3D model information (Fig 2.4B).

2.4 RESULTS AND DISCUSSION

The 3D models, together with the histological studies, provide a developmental overview of the heart of *Emys orbicularis*. In this section a full description is given of this developmental series, with a focus on the looping of the heart and the development of the horizontal septum. Several figures illustrate the major findings; all material has been made available on http://bio-imaging.liacs.nl/galleries/. Here one can find additional figures and animation sequences and view the 3D reconstructions interactively, using the TDR-viewer.

Below a description of the developmental features is given for each of the studied stages, some based on histological sections, others on 3D models. This is followed by some technical comments on the different models and the morphological structures presented in them.

Developmental anatomy

In this section we discuss the morphological findings from our histological sections and, for stages 8, 10 and 15, the 3D models. From laying till hatching 26 stages have been described by Yntema (1968), which can be divided in three periods: presomite period (stages 0-3), somite period (stages 4-10) and limb period (stages 11-26). In order to allow comparison of the described data to development in other animals, some defining characteristics have been given for each described stage; complete descriptions can be found in Yntema (1968).

Stage 6

At this stage eight pairs of somites can be seen and a small cranial neuropore persists. The cardiac endocardial tubes have just fused. The dorsal mesocardium of the atrium is present, but mesenchyme cannot yet be found between its layers. The pericardial coelom can be seen, as well as small amounts of cardiac jelly. No other developmental characteristics were discernible at this stage.

Stage 8 (Fig. 2.5A and 2.6A)

Fourteen pairs of somites are present at this stage and the neural folds are completely closed cranially. Cardiac looping is well underway and clearly visible in the 3D model (Fig. 2.5A). The S-looping, described by Männer *et al.* (2000, 2009), is in progress, but is still in the early phase, corresponding to the third phase in the chick heart; the proximal part of the outflow tract is still on the right side of the heart and the common atrium on the left. The caudal end of the original heart tube, now on the dorsal side of the heart, has ended up in a fold, which is forming the atria. In the cranial direction along the tube lie the ventricle and outflow tract, which are still continuous regions. In this part of the

tube a distinct bending process is taking place. On leaving the primitive ventricle the tube first passes cranially, then bends caudally and to the left, which forms a first, proximal bend. Subsequently it bends again, cranially and to the left, thereby forming a

A

B

C
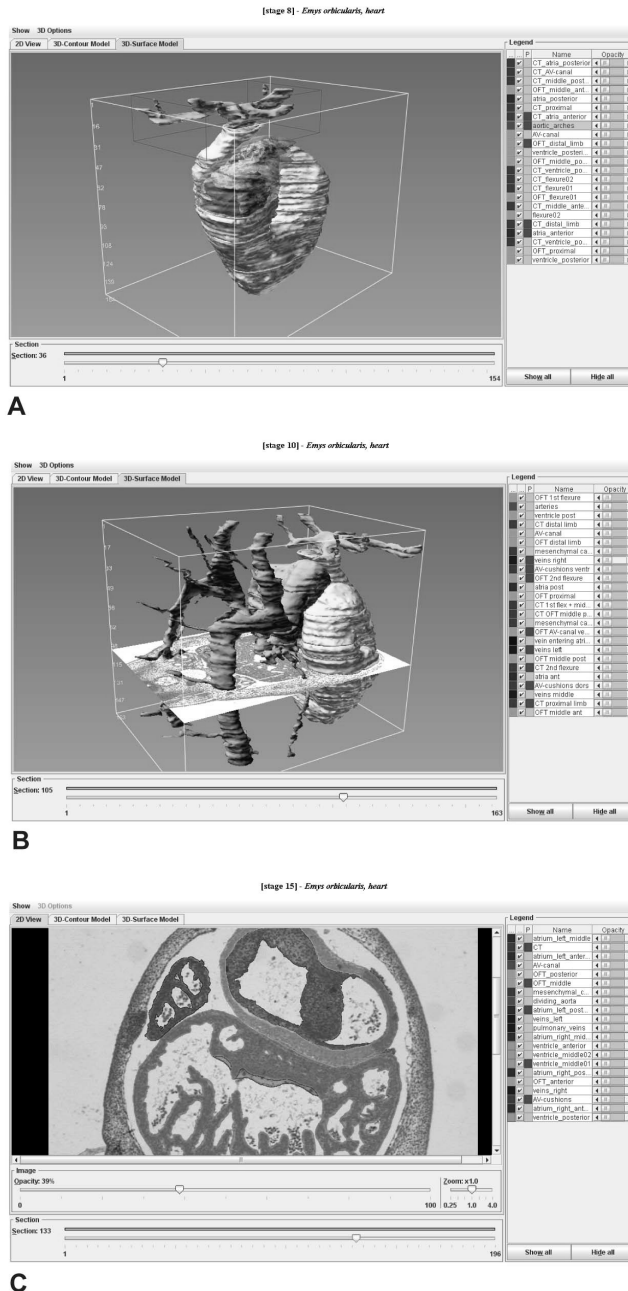
Figure 2.4. Screenshots of the TDR-viewer showing the three 3 reconstructions in different modes. In (A) a surface representation of stage 8 is shown; in addition to the standard interactive visualisation the opacity of each of the anatomical domains can be changed interactively. Anatomical domains can also be deselected so that internal structures come into view. In (B) a surface representation of stage 10 is shown in combination with a 2D cursor that corresponds to the original section. And in (C) a 2D screenshot of stage 15 is shown; here the legend indicates all anatomical domains in the entire model, with the ones seen in this section image highlighted using green boxes.
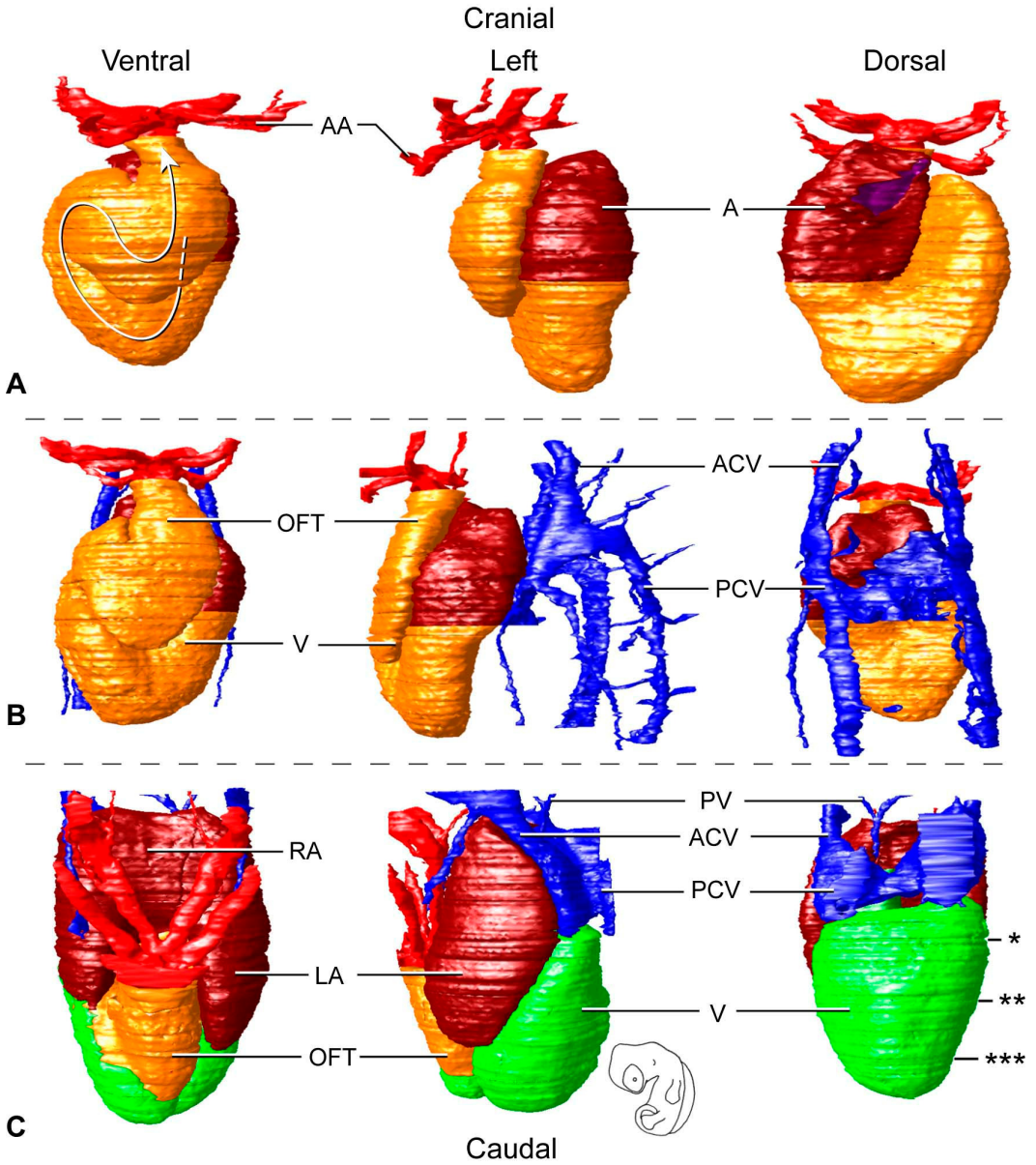
Figure 2.5. Overview of the 3D reconstructions of three *Emys orbicularis* embryos, Yntema stages 8 (A), 10 (B) and 15 (C), shown in ventral, left and dorsal view. The ventral view of A shows the direction of the looping of the heart tube (indicated by the arrow). The asterisks in the dorsal view of C correspond to the cutaway views in figure 10. The schematic drawing of the embryo, in a lateral left view, corresponds to the orientation of the left view of the models. The color labeling of the models is explained in the List of abbreviations.
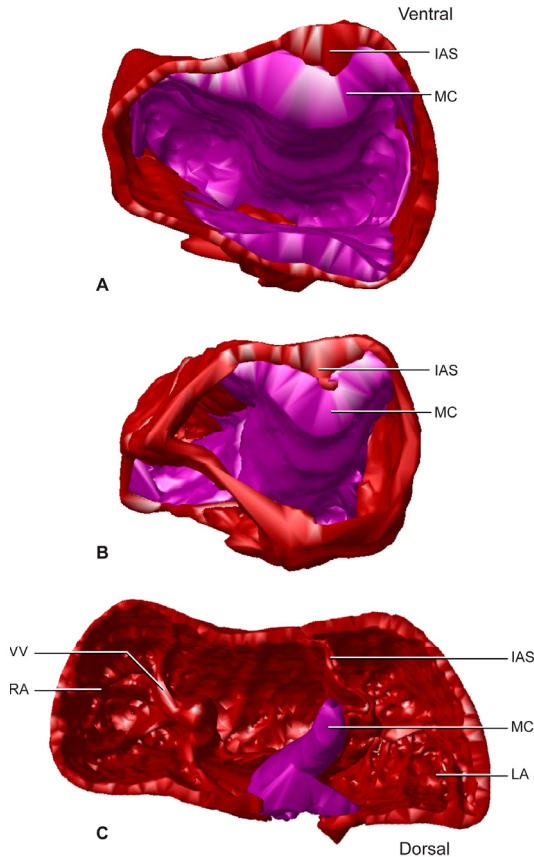
Figure 2.6. Cutaway views of the 3D reconstructions of three *Emys orbicularis* embryos, Yntema stages 8 (A), 10 (B) and 15 (C), showing (part of) the atria and mesenchymal cap in caudal view.

second, distal bend. The bending is starting to separate the outflow tract and *cavum pulmonale* from the rest of the ventricle and is responsible for the future development of the horizontal septum (explained below). In the outflow tract two distal and two proximal cushions are visible. Cardiac jelly is still present in the ventricle. Small trabeculae carneae are just becoming visible. None of these trabeculae, however, can yet be identified as the vertical septum. Furthermore, the *cavum arteriosum* and *cavum venosum* are still indistinguishable.

The interatrial septum is barely indicated as a slightly thickened ridge in the ventral part of the atrial wall. This ridge, the presumptive septum primum, is covered with cushion tissue, forming the 'mesenchymal cap' (Fig. 2.6A). Although the endocardium overlying the pulmonary pit is not yet perforated, the pulmonary ridges are thickened. The atria are still in open connection with the ventricle. In the passage between common atrium and ventricle, two thick endocardial cushions can be seen, continuous with the mesenchymal cap, which will go on to form the atrioventricular valve leaflets at a later stage.

Cranial

HSt

A
IHC
CAVV

1 mm
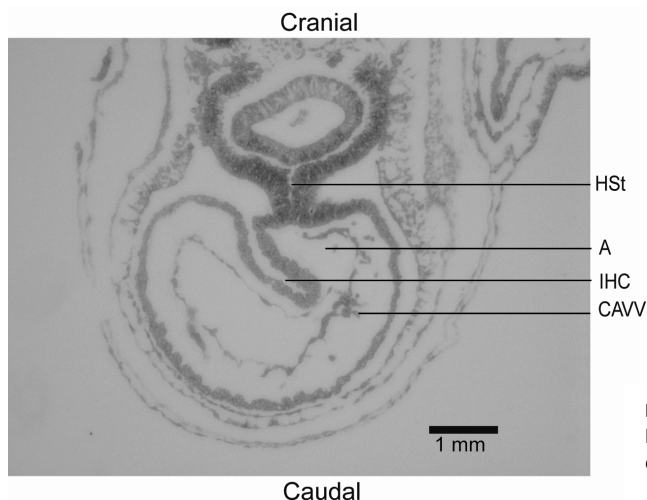
Caudal

Figure 2.7. Section image of the heart of a stage 9 *Emys orbicularis* embryo in caudal view.

Stage 9 (Fig. 2.7)

Nineteen pairs of somites are present, the lens pit is visible and the first pharyngeal slit is open. Cardiac jelly is still present in both the atria and the ventricular part of the heart, with early signs of cell delamination in the future atrioventricular and ventricular region. Trabeculae carneae are clearly visible in some embryos at this stage, but still developing in others.

Stage 10 (Fig. 2.5B and 2.6B)

At this stage twenty-four pairs of somites have developed, the first two pharyngeal slits are open and all four limb buds are present. At this stage cardiac looping is still in progress and the distal bend of the heart tube is moving towards the apex of the ventricle (Fig.2.5B). The heart can still be seen as one folded tube and is in the phase of late S-looping, the fourth phase as described by Männer (2000, 2009); the outflow tract is moving to the left and the common atrium has expanded to the right. In the ventricle trabeculae are still developing and the vertical septum is not yet evident. The *cavum arteriosum* and *cavum venosum* cannot be distinguished. In some embryos trabeculae were only just arising, while in others some trabeculae were found to be two or three times as high as they were wide. In the outflow tract the two proximal and two distal cushions were very clear.

In the common atrium the interatrial septum is forming (Fig. 2.6B). The dorsal mesocardium has not yet lumenised to form the pulmonary vein. Cardiac jelly is still present in both the atrium and the ventricle, but is diminishing. In the atrium pectinate muscles have started to develop. Atrioventricular cushions are now distinct and will go
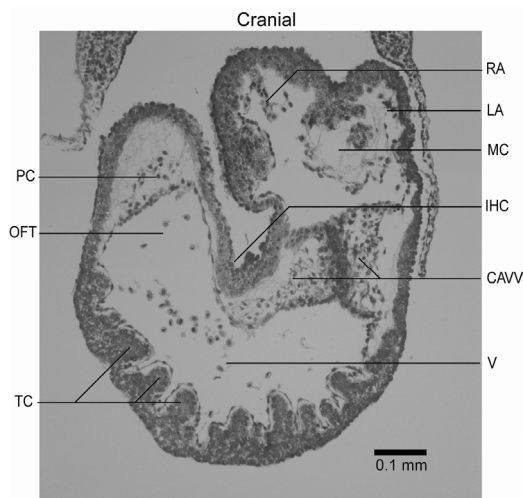
Figure 2.8. Section image of the heart of a stage 11 *Emys orbicularis* embryo in caudal view.
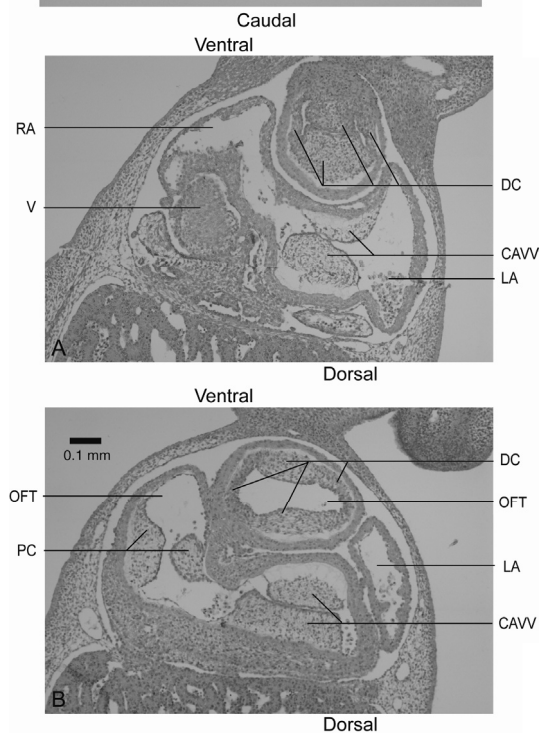
Figure 2.9. Section images of the heart of a stage 14 *Emys orbicularis* embryo in caudal view; A lies cranial to B.

on to form the atrioventricular valve leaflets and the atrioventricular septum. Cell delamination is still in its early stages, with only a few cells delaminating in the outflow tract and the atrioventricular cushion tissue. In one of the four embryos cushions were noted in the outflow tract. The outflow tract is still an undivided channel.

Stage 11 (Fig. 2.8)

Thirty-one pairs of somites are present, the first pharyngeal slit is still open, but the second is covered by the hyoid arch. The cervical flexure has increased prior to turning of the embryo onto its left side.

The atrium is now partially divided by the atrial septum, which extends approximately half way through the common atrium. The mesenchymal cap is still present and is almost completely acellular at this stage, as can be seen in Fig. 2.8. The dorsal mesocardium has advanced since the previous stage and now contains vacuoles, joining to form the pulmonary vein. The pulmonary vein opens through the pulmonary pit into the left atrium, immediately left of the septum primum. In the atria, cardiac jelly is still very prominent and in the atrioventricular canal the cushions are even more apparent. Endocardial cell delamination is very clear in the inner parts of the cushions (Fig. 2.8).

In the ventricle hardly any cardiac jelly remains in the trabeculated area. The trabeculae are now up to three to four times as high as they are wide. In the distal part of the outflow tract two endocardial cushions, a ventral and a dorsal one, are clearly visible and, as in the atrioventricular cushions, cell delamination can be seen in the inner parts. The inner heart curvature is lined with endocardial tissue, connecting the ventral cushion in the outflow tract to the superior atrioventricular cushion. The aorto-pulmonary septum is not present at this stage.

Stage 12

The embryo now lies on its left side, the pharyngeal slits have disappeared, the retina is pigmented and the apical ridge is beginning to form in the forelimb bud.

The septum primum in the common atrium has not developed much since the previous stage and still spans only half of the length of the atrium. The pulmonary vein can be traced from both lung buds to the left atrium. In the atria small pectinate muscles can be seen. Cardiac jelly is also still present in the atria. In the atrioventricular cushions cell delamination has advanced little if any.

In the ventricle trabeculae are very prominent. The endocardial cushions in the outflow tract are more heavily populated with cells than before. The distal cushion spirals to the left as it extends proximally. The outflow tract is still undivided at this stage.

Stage 14 (Fig. 2.9A and B)

The maxillary and lateral nasal processes have fused and the forelimb is in early peddle stage, with the digital plate vaguely indicated.

The right and left sinus valve ridges are both present at this stage. The septum primum now spans approximately two thirds to three quarters of the common atrium. The mesenchymal cap is completely acellular. In one of the three embryos a small additional septum was noted in the atrium, possibly the venous valve of the superior cardinal vein. The pectinate muscles in the atria are now more distinct, generally as high as they are wide and the cardiac jelly has vanished. The atrioventricular cushions are prominent and completely infiltrated with cells. The cushions have not yet fused (Fig. 2.9A and 2.B), but seem to be touching, although this could be due to the phase in which the cardiac activity has been arrested.

In the outflow tract six endocardial cushions can be seen, two proximal and four distal cushions (Fig. 2.9A and B). The distal cushions account for the H-shaped form of the lumen, when seen in transverse sectioning. Like the atrioventricular cushions the proximal outflow tract cushions are completely infiltrated with cells. The cushions approach each other, but are not yet touching.

Stage 15 (Fig. 2.10A, B and C; 2.11A, B and C)

The cervical sinus is closed and the digital plate of the forelimb is well formed with no digital grooves present.

By now the distal bend of the original heart tube has reached the apex of the ventricle and the walls of both have merged. The heart is in the fifth phase described by Männer (2000, 2009), the phase of cardiac septation. The caudal apex of the bend has formed the *cavum pulmonale*, continuous on the cranial end with the outflow tract (Fig. 2.5C). The merged wall between the *cavum pulmonale* and the rest of the ventricle has formed an incomplete septum, the horizontal septum. This is clearly visible in the cutaway views of the 3D model of this stage, shown in (Fig. 2.12). The *cavum pulmonale* can be seen on the ventral side of the incomplete septum, at the base of the outflow tract, as a distinct cup-shaped chamber. It is almost completely separated from the rest of the ventricle, save for a small passage on the cranial side, connecting it to the *cavum venosum*. From this developmental series it can be concluded that the *cavum pulmonale* develops at the proximal end of the outflow tract and the horizontal septum is formed through bending of the tube (*cf.* 2.5).
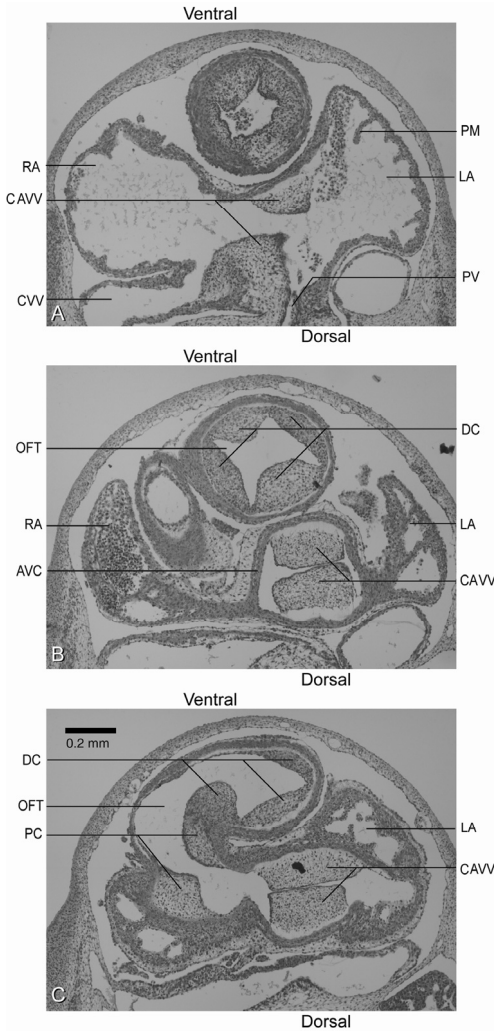
Figure 2.10. Section images of the heart of a stage 14-15 *Emys orbicularis* embryo in caudal view, progressing from cranial (A) to caudal (C).
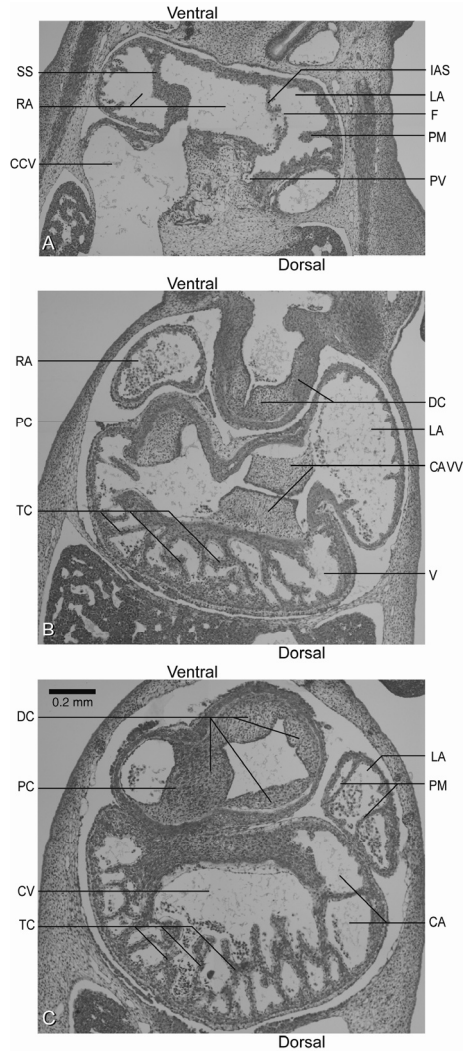
Figure 2.11. Section images of the heart of a stage 15 *Emys orbicularis* embryo in caudal view, progressing from cranial (A) to caudal (C).

It is interesting to note that the *cavum pulmonale* has almost no trabeculae at this stage, while the rest of the ventricle is heavily trabeculated, with trabeculae five to seven times as high as they are wide (Fig. 2.11B and C). One of these trabeculae is slightly bigger and its dorso-ventral base forms a septum at the apex of the ventricle.
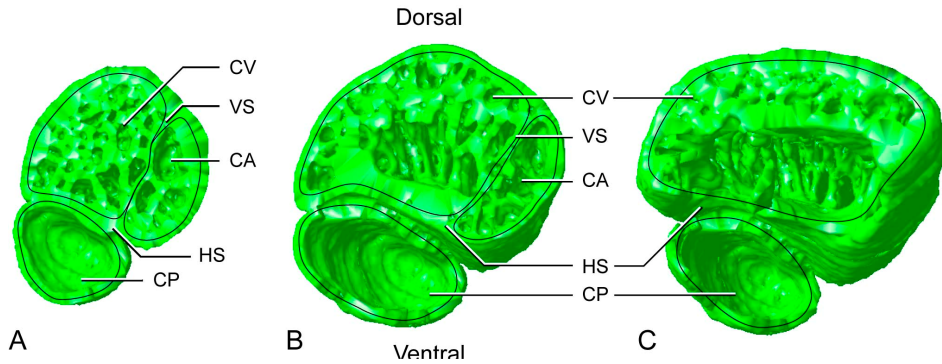
Figure 2.12. Cutaway views of the 3D reconstruction of a stage 15 *Emys orbicularis* embryo, showing only the ventricle in cranial view, at three craniocaudal levels. The asterisks in Fig. 2.3 correspond to the planes of section (* corresponds to A, ** to B and *** to C).
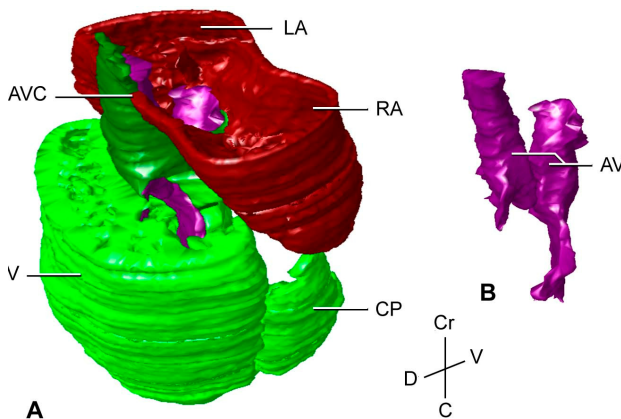
Figure 2.13. Cutaway view of the 3D reconstruction of the stage 15 *Emys orbicularis* embryo, showing parts of atria, ventricle and atrioventricular canal in a ventrolateral view in A. B shows the atrioventricular valves separately in the same plane of view.
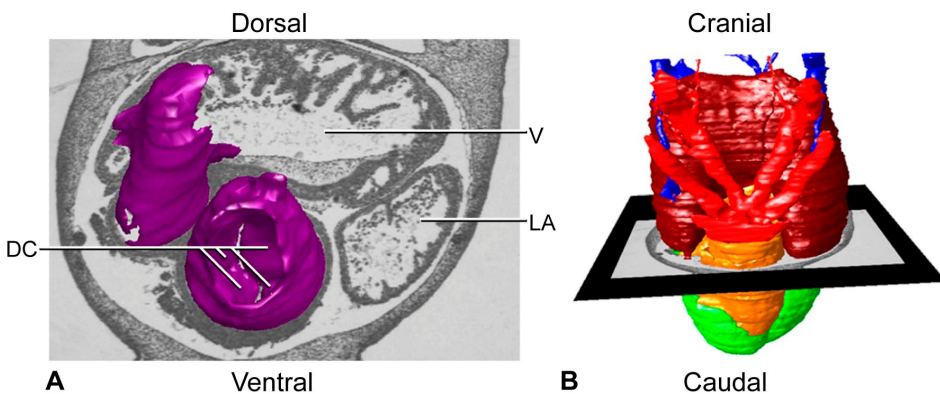
Figure 2.14. Cutaway view of the 3D reconstruction of the stage 15 *Emys orbicularis* embryo, showing the distal cushions in the outflow tract, together with a section image in A. The position of the section image in relation to the entire heart is shown in B.

This is the vertical septum, dividing the *cavum venosum* from the much smaller *cavum arteriosum* (Fig. 2.11C and 2.12). This septum is about half the size of the horizontal septum. It is not covered by a mesenchymal cap and is therefore more of a muscular ridge than a true interventricular septum.

The atrioventricular cushions are still developing and by now approximately two thirds or three quarters of their bulk is infiltrated by cells. The cushions are touching but have not yet fused (Fig. 2.10B, 2.10C and 2.11B). The atria and ventricle are no longer in open connection, but are separated by an atrioventricular canal, in which the atrioventricular cushions are suspended (Fig. 2.10B, 2.13 and 2.15). The ventral atrioventricular cushion is still connected at one end to the proximal cushion tissue in the outflow tract and at the other end to the mesenchymal cap (as can be seen in the animation sequence of the cushion tissue, wich can be found at htpp://bio-imaging.liacs.nl/galleries/). In the outflow tract the four distal and the two proximal cushions are now clearly visible (Fig. 2.10B and 2.14). Two of the distal cushions are nearly fusing and the most distal of the proximal cushions is continuous with one of the distal cushions (Fig.2.10C and 2.11C); a more detailed account of this is given in section 3.3.3.

The atrial septum now extends three quarters of the total distance to the atrioventricular canal and its two limbs are easily distinguishable (Fig. 2.6C). The mesenchymal cap is still present, but very small and is completely infiltrated with cells. In one of the four embryos, perforations in the atrial septum were seen. Immediately left of the septum is the pulmonary pit, in which the lumen of the pulmonary vein is now unequivocally present (Fig. 2.11A).

Early stage 16

This is the latest stage in developmental series available to us in this analysis and it resembles the preceding stage in most aspects. The septum primum remains approximately the same. The pectinate muscles have grown a little and are now two or three times as high as they are wide. The trabeculae are still about five to seven times as high as they are wide. Both the atrioventricular cushions and the outflow tract cushions are completely infiltrated with cells. The cushions are clearly touching but still have not fused.

Source data and annotation of the 3D models

In each of the 3D models the following structures have been labelled: atria, ventricle, outflow tract, aortic arches and endocardial cushion tissue (Fig. 2.5). In two models (Fig. 2.5B en C), the veins entering the heart have also been traced. All of the structures have been divided into smaller sub-structures enabling the user to study smaller parts separately and from the inside.

In this study the models were based on sectioned embryos and the recognition of different tissue types was facilitated by staining methods. Alternatively, a non-invasive technique could have been used; given the size of the *Emys* heart, micro Magnetic Resonance Imaging (µMRI) could have been a good candidate. However, a µMRI study of an adult heart (not published) has taught us that the level of detail is significantly lower than could be achieved using invasive techniques, combined with histological staining. For this reason, physical sectioning was preferred.

The manual delineation of the anatomical domains was, for this specific section material, favoured over automated methods. Given the standard staining of the sections, functional differences between parts of the same tissue type could not have been distinguished automatically; *e.g.* by using automated tracing the atria and ventricle
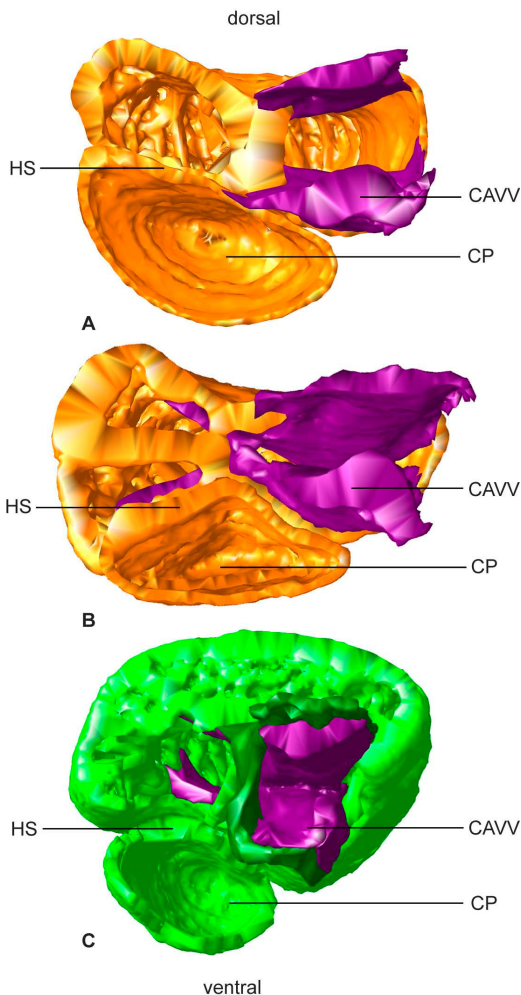
Figure 2.15. Cutaway views of the 3D reconstructions of stages 8 (A), 10 (B) and 15 (C) *Emys orbicularis* embryos, showing the caudal parts of the ventricle in a cranial view, including the ventricle and cushion tissue which will form the AV-valves. The division between the *cavum pulmonale* and the rest of the ventricle, by the developing horizontal septum, is visible.

would all have become part of one structure, based on their tissue type. Manual delineation allowed us to distinguish between structures based on their functionality. Atria and ventricle are therefore annotated as different structures, which makes the models more useful. Furthermore the structures are more detailed and more reliable than would have been possible using computerised tracing; irrelevant material (*e.g.* blood remaining in the lumina) or artefacts from staining and sectioning would have been falsely included using automatic tracing.

For the stage 8 model (Fig. 2.5A), 154 section images were used. The ventricle and outflow tract could not be clearly demarcated from one another at this stage, and so have been assigned the same colour label (orange). The common atrium has not yet divided into right and left parts and is therefore shown as one structure. The veins entering the heart were similar to those seen at stage 10 and were not included in this model.

The stage 10 model (Fig. 2.5B) was constructed from 163 section images. The outflow tract and ventricle are still not completely developed and are therefore presented in orange as one structure. Although the interatrial septum is beginning to form, the atria are not yet easily distinguishable and are therefore also shown as one structure. The arterial arches leaving the heart and the veins entering it are shown in detail, as well as the cushion tissue.

In the model of stage 15 (Fig. 2.5C), built from 196 section images, the ventricle and outflow tract are clearly separated and have each been given a different colour (green and orange, respectively). The interatrial septum is almost completely closed (save for several foramina) and it is now possible to distinguish between the left and right atrium. We have therefore assigned different structure labels to the atria (as opposed to one label for the common atrium in the other models), thereby enabling the user to study the atria simultaneously and separately. Since the dividing line between the atria is very clear from the outside of the model, no difference in colour between the two structure labels was deemed necessary. As in the previous model, veins, arteries and cushion tissue are visible. The atrioventricular canal, which at this stage can be distinguished from the ventricle and atria, is labelled as a separate structure.

All section images, annotated domains and 3D representations are available on the website. Our interactive TDR-viewer application (see Fig. 2.4 and website) provides the user with a wealth of functionalities, which help understand the 3-dimensional structure of the hearts. The 2D view allows the user to browse through the annotated sections (Fig. 2.4C). In a section image, annotated domains can be selected or hidden and the opacity of selected structures can be adapted. Selecting a particular domain from the legend results in the selection of the first section image in which this domain

occurs; in this manner quick inspection and exploration of specific structures is made possible. A slider below enables browsing through all consecutive section images.

The model can be inspected in 3D using the 3D contour or surface view. It can be moved around and inspected from all viewpoints, both as a whole and only showing selected domains. In the surface representation the opacity of each of the anatomical domains can be changed interactively (Fig. 2.4A). A 2D cursor can be added, showing the original section image (Fig. 2.4B). Switching from the 3D to the 2D view allows one to inspect of the annotated areas in that particular section image.

2.5 CONCLUSION

This study yields important new insights into the developmental origins of the horizontal septum in the turtle. From the 3D models it becomes clear that the septation is closely linked to the looping of the heart. This looping is responsible for the formation of the horizontal septum and the demarcation of the *cavum pulmonale* and outflow tract. The looping of the turtle heart is quite distinct from that found in human development. In humans, the part of the heart tube leaving the ventricle has one clear bend, originally termed the 'bayonet bend' (Orts LLorca *et al.*, 1982) and later described as the 'dog-leg' bend (Webb *et al.*, 2003); this separates the distal part of the outflow tract from the proximal part. The proximal part of the outflow tract, leaving the ventricle, is relatively straight. The two developing ventricles lie next to each other (seen in a more or less coronal plane) and become separated by an interventricular septum, through expansion of the ventricles, while the medial walls of the ventricles grow together gradually and fuse.

The situation in the turtle is markedly different. Here the part of the heart tube leaving the primitive ventricle does not have one but two very clear bends. The tube first passes cranially, bends caudally and to the left, forming a first, proximal bend; it then bends a second time, cranially and to the left, forming a second, distal bend. These bends are very clear in the cranial views of the 3D reconstructions of stages 8 and 10 (Fig. 2.5A and B).

Because of this difference in heart looping the establishment of septa in the primitive ventricle in humans and turtles occurs in a strikingly different way. In the turtle the second, distal bend of the tube comes to lie ventrally to the ventricle and where the walls of bend and ventricle touch they merge to form the horizontal septum, which can be seen in Figures 2.5, 2.12 and 2.15. We can see the distal bend moving gradually in the caudal direction during development, to reach the level of the apex of the ventricle. The lumen of this bend becomes the *cavum pulmonale* and is connected to the rest of the ventricle over the free margin of the horizontal septum.

So whereas the ventricles in the developing human heart lie next to each other in the coronal plane, the *cavum pulmonale* lies in front of the rest of the ventricle in the turtle heart. And whereas the human interventricular septum forms through expansion of the ventricles, the horizontal septum in the turtle forms through merging of the ventral part of the ventricular wall with the dorsal wall of the distal (*i.e.* second) bend in the outflow tract.

Our findings contradict the theory proposed by Holmes (1976) on the development of the horizontal septum. He suggests that the cranial part of this septum derives from the spiral fold of the conus arteriosus, while the caudal part derives from hypertrophied trabeculae that join the cranial part. Van Mierop and Kutsche (1984) also suggest that a cranial part of the septum may be formed by condensation of trabeculae. It is clear from Figures 2.5, 2.12 and 2.15 however, that the septum consists of merged walls and that the walls of the *cavum pulmonale* are smooth and do not contain any trabeculae.

Another interesting biological finding concerns the developmental origin of the pulmonary vein. The development of this vein in human, mouse and chick has been the matter of some debate (Blom *et al.*, 2001; Webb *et al.*, 2001). This debate centres around the relationship between the vein and the systemic venous sinus (sinus venosus). Although immunohistological studies have suggested that the human pulmonary vein originates from the sinus venosus (Blom *et al.*, 2001), recent studies in chick and human indicate that the vein canalizes as a newly formed channel within the developing posterior mediastinum and uses the remnants of the dorsal mesocardium to establish a direct connection with the atrium (Webb *et al.*, 2001; Webb *et al.*, 2000).

Our findings in the turtle support these recent findings in chick and human. The sections show that the pulmonary vein in turtles also derives from the mediastinal myocardium and not from the sinus venosus. It forms in the pulmonary pit of the left atrium as was described for the chick (Webb *et al.*, 2003). The sinus in reptiles is partly absorbed into the right atrium and receives the systemic veins (left and right precaval veins and postcaval veins) and coronary sinus (Gasch, 1888; O'Donoghue, 1918; Rau, 1924). The left atrium receives only the pulmonary vein or veins (Gasch, 1888; O'Donoghue, 1918; Rau, 1924).

With respect to 3D reconstruction techniques this study illustrates the indispensable nature of these techniques when studying morphogenesis in the turtle heart. Many biological studies would benefit from high-resolution 3D reconstructions. Our software, which allows high-resolution modelling and manual discernment between functionally separate structures, is therefore of great use. By enabling the readers to study the models in an interactive way, using our TDR-viewer application, we hope to go beyond a descriptive reporting of our findings, using merely static 2D images, and to

invite the readers to join the ongoing search for a more thorough understanding of cardiogenesis in the turtle.

CHAPTER 3. A VISUAL MODEL AT HIGH RESOLUTION: OUTFLOW TRACT DEVELOPMENT IN THE TURTLE SPECIES *EMYS ORBICULARIS*

'What *is* the use of repeating all that stuff,' the Mock Turtle interrupted, 'if you don't explain it as you go on? It's by far the most confusing thing *I* ever heard!'

- Lewis Carroll, *Alice's adventures in Wonderland*

The next case study is an elaboration on the study of the turtle heart development presented in the previous chapter. The transition can be considered as zooming in to a higher level of magnification. Again a series of 3D models has been constructed, in order to study the spatial aspects of developmental anatomy, but while the previous chapter was based on models of the entire heart, the current study focuses on reconstructions of one particular element of the heart: the outflow tract. To this end, high resolution section images were acquired and used as the basis for the reconstructions, enabling a detailed study of the development of this structure. In this way new insights were gained concerning the development of the *cavum pulmonale*, the outflow tract and the outflow tract cushions. The case study exemplifies the use of high resolution modelling for developmental biology.

Manuscript, to be submitted as: Bertens, L.M.F., Richardson, M.K., Poelman, R.E., Verbeek, F.J., 'Outflow tract development in the turtle *Emys orbicularis*'.

3.1 INTRODUCTION

Numerous studies have examined development of the mammalian and chicken hearts. However, far less is known about the development of the reptilian heart (*i.e.* diapsids, excluding the birds). While several studies have addressed the anatomy and physiology of the adult reptilian heart, they often arrive at conflicting conclusions, mostly regarding ventricular septation (Webb *et al.*, 1971; Holmes, 1975; Johansen and Burggren, 1980; White, 1986; Burggren and Warburton, 1994; Bertens *et al.*, 2010). Studying the development of the anatomy and physiology is important in this respect because it can shed new light on these issues. Furthermore, with the current interest in phylogenetic relationships between reptiles, birds and mammals, and particularly the placement of the turtles (Chelonia) within reptiles, (Webb, 1979; Farmer, 1999; Lyson, 2011) studies in evolutionary development are highly relevant.

In the previous chapter a series of 3D reconstructions was presented of different developmental stages in the cardiogenesis of the turtle *Emys orbicularis*. The early looping of the turtle heart was described, along with the finding that the double loop gave rise to the *cavum pulmonale*. This hypothesis has implications for the positioning of the valves derived from the distal cushions in the outflow tract, the septation of this tract and the roles of the endocardial cushions and aortopulmonary septum in this process. Furthermore, in the adult heart the ventricle is functionally three-chambered and this has led to the conventional assumption that the presumptive ventricle becomes partially divided into three cavities. Interestingly, this arrangement of three ventricular *cava* strongly resembles the ventricular morphology of Squamata (snakes and lizards) and differs substantially from the partitioned ventricle of Archosauria (crocodilians and birds). However, our findings raise the possibility that the turtle ventricle may not be three-chambered in origin, but in fact two-chambered, with the *cavum pulmonale* arising from the outflow tract. This would represent a type of cardiac septation unknown in any other vertebrate, meriting further investigation of the outflow tract development.

In the current chapter, the development of the outflow tract of the turtle heart is examined using 3D reconstructions, assembled from histological sections of the turtle *Emys orbicularis* (the European pond turtle), at different stages of embryonic development. Subsequently, the findings are discussed in relation to current theories on chelonian and reptilian outflow tract development.

3.1.1 Current knowledge of outflow tract development in reptiles

Chapter 2 provides an overview of the developmental anatomy of the turtle heart (section 2.2). Here, further aspects of ouflow tract (OFT) development in reptiles are

discussed. Note that 'reptiles' is used here with the knowledge that it is a paraphyletic group, and that the placement of the turtle is controversial. However, it is clear that the crocodile and chick (Archosauria) among reptiles have completely septated ventricles, while the Chelonia (turtles and tortoises) and the Squamata (lizards and snakes) have a partially septated ventricle with three cavities (*cavum arteriosum*, *venosum* and *pulmonale*). We will now further discuss the similar pattern of cardiac development between Chelonia and Squamata.

In the adult heart of these species, three arterial trunks leave the ventricle: right and left aortae and the pulmonary trunk. These arise from the septation of a common OFT. At the root of each of the trunks in the formed heart, semilunar valves, with two leaflets, are present. The brachiocephalic artery branches off from the right aortic arch and gives off the subclavian and carotid arteries.

The connection of these arterial trunks to the ventricular outlet arises from the septation of the OFT and the development of valves, guarding the roots of the trunks. These developmental processes are not yet fully understood in the reptilian heart. It is, however, clear that these processes differ significantly from the mammalian situation. In both the mammalian and the reptilian heart, a curvature arises in the developing OFT, described for the human heart as a 'bayonet bend' by Orts Llorca *et al.* (1982) or a 'dog-leg bend' by Webb *et al.* (2003). This bend demarcates the transition between proximal and distal components of the OFT. Previous researchers have referred to these structures as: (i) conus or bulbus and (ii) truncus, respectively. In the turtle, the proximal component of the OFT comes to lie ventrally to the ventricle and forms the *cavum pulmonale* (Bertens *et al.*, 2010). This has important implications for the development of the valves of the arterial trunks and the septation of the OFT.

Both the proximal and distal components of the OFT contain endocardial cushions. The distal endocardial cushions give rise to the leaflets and sinuses of the semilunar valves, which are located during development at the dog-leg bend. The number of cushions in the distal and proximal components of the OFT has been shown to differ between taxonomic groups and no consensus exists as yet concerning the relations between the cushions in different species (Qayyum *et al.*, 2001; Webb *et al.*, 2003; Okamoto *et al.*, 2010). Goodrich (1958), Shaner (1962) and Hart (1968) all give accounts of OFT development in the reptilian heart. The authors agree that four cushions are found in the distal component of the OFT and two in the proximal component. Both in descriptions and 3D reconstructions we follow the numbering of the cushions used by Shaner (1962) and Hart (1968); one of the distal cushions is markedly bigger and this cushion is designated number DC1. Numbering continues in clockwise fashion when looking at a rostral view. The proximal cushions are designated by Hart (1968) as PCA and PCB (corresponding to proximal ridges 1 and 4 in Shaner, 1962).

Hart (1968) states that DC1 and DC3 arise at a later stage than DC2 and DC4. This temporal differentiation between the cushions is relevant in relation to the different distal and intercalated cushions observed in mammals and birds. DC1 and DC4 are said to merge very early in development at their proximal end with PCA and PCB, respectively (Shaner, 1962; Hart, 1968). The location of the proximal cushions and their fusion with the distal cushions in the reptilian heart, as described by Shaner (1962) and Hart (1968), differs from the situation observed by Qayyum *et al.* (2001) in the chick heart; in early development three distal cushions arise in the avian heart, later followed by two intercalated cushions. Of the three initial cushions the first two to develop correspond to DC1 and DC3 in reptiles. Furthermore, DC1 is seen to merge with PCA, but unlike in the reptilian OFT, DC3 does not merge with PCB.

In addition to the endocardial cushions, the aortopulmonary septum (AP septum) is said to play an important role in septating the distal end of the OFT and the allocation of the cushions between the resulting arterial trunks (Goodrich, 1958; Shaner, 1962; Hart, 1968). The AP septum is formed by migratory neural crest cells, and is visible as dense ectomesenchymal tissue on histological sections (Hart, 1975; Bartelings and Gittenberger, 1989; Qayyum *et al.*, 2001; Webb *et al.*, 2003). The main body of the septum runs through the distal end of DC1 in the direction of the heart. It gives off two lateral prongs and is described by Shaner (1962) as being Y-shaped. One of the prongs extends from DC1 in between DC2 and DC3, the other one in between DC3 and DC4. In this way each of the developing arterial trunks receives one complete distal cushion and a fragment of DC1. These cushions form the valve leaflets providing each of the three arterial trunks with a bicuspid semilunar valve (Shaner, 1962). So far this pattern of reptilian septation has been illustrated only in schematic drawings of histological sections (Goodrich, 1958; Shaner, 1962; Hart, 1968; Holmes, 1975). No clear 3D visualisation has been presented in literature. Our 3D reconstructions of the turtle OFT provide a more informative view of the septum and the cushions in the dividing distal OFT.

3.2 MATERIALS AND METHODS

Preparation of the embryos

Five developmental stages were selected, Yntema stages 8, 15, 16+/17, 19 (Yntema, 1968) and an adult stage (18 months). Of these, the histological sections of the stage 8 and stage 15 embryos have also been used in reconstructing models presented in chapter 2; here the same histological sections have been used, but new reconstructions have been made, of different structures.

The embryos were selected for the presence of significant developmental aspects of the OFT. Gravid females were collected under license from the French government. Standard injection with oxytocin was used to induce laying. Eggs were places on a layer of sand in an incubator at 25-30°C. Embryos were fixed in Bouin's fluid for 2 days and embedded in Fibrowax according to standard protocols, serially sectioned at 7 μm and stained with Haematoxylin, Eosin and Alcian Blue. The material examined is summarized in table 3.1.

Yntema (1968) stage	Embryo code	Plane of section	Shown in figure
8	CP245	transverse	3.2A; 3.3E
15	CP46	transverse	3.1A-C; 3.2B,C
16+/17	CP450	transverse	3.4F,G
19	CP436	transverse	3.1D,F; 3.4A-E
adult, 18 months	CP434	coronal	3.3A-D; 3.1G

Table 3.1. Overview of embryos used in the study.

Acquisition of section images

Sections containing the outflow tract were digitized using the microscope and camera setup described in 2.3. Section images were captured with our inhouse software 3D acq, version 2.0 (Verbeek and Boon, 2002); using this software an image stack was created for each of the specimens. Using the video overlay function of the PCVision frame grabber, images were aligned prior to acquisition. The metadata was stored in an XML acquisition database referring to the section images which were stored in PNG format.

Using the metadata database, a second high resolution image stack was obtained for two of the hearts. 3D acq allows the user to construct section images of higher resolution and at a higher microscopic magnification from separate image tiles which are digitally stitched together. Using the frame grabber coordinates stored in the metadata for the original images, the software can automatically relocate the right position of a previously captured image on a slide. Subsequently, a set of image tiles (number, magnification and overlap between tiles are set by the user) is acquired which, when stitched together, form a high resolution image of the original image. In this way high resolution images were obtained at 16 x magnification, constructed of 4 x 4 image tiles, with 32 pixels overlap. Metadata for these high resolution images was added to the original XML acquisition database and the images were stored in PNG format.

Construction of 3D models

From the obtained section images 3D reconstructons were produced with the inhouse software TDR-3Dbase (Verbeek and Boon, 2002; Bertens *et al.*, 2010). All relevant anatomical structures in the images were traced and labeled with different colors and names. Using these annotations the structures were visualised in surface models, derived through automated triangulation (details on the method can be found in 2.3). The models are provided via internet using our 3D model browser application TDR-viewer (Potikanond and Verbeek, 2012) and can be found here: http://bio-imaging.liacs.nl/galleries/. The viewer allows the user to inspect the annotations and visualisations both in 2D (*i.e.* the histological section images) and 3D. The 3D views are interactive and allow for active exploration of the model information; the user can select anatomical structures to be displayed and a 3D visualisation of this selection is instantaneously rendered.

Anatomical conventions

For the left and right aortic arches leaving the ventrical, two different nomenclatural conventions have arisen, one based on the position at the base of the arterial arches, *i.e.* the point where the arches connect to the ventricle (Shaner, 1962), and the other based on their position further away from the ventricle, after the point of spiralling (Langer, 1894; Hochtstetter, 1901; Goodrich, 1958). Inconsistent identification of structures, as a result of these different terminologies, is exemplified by the description of Holmes (1975), who follows the naming convention according to Goodrich (1958) in his written account and his illustrations of the entire heart, but includes transverse section drawings taken from Shaner (1962), in which left and right aortae are reversed, resulting in inconsistencies within the publication itself. Here (as in the previous chapter), we follow the convention of naming the structures based on their position further away from the heart, since this coincides with the naming of *e.g.* the venae cavae and the left and right pulmonary vein. Therefore, the carotico-systemic arch is identified here as the right aorta and the systemic arch is identified as the left aorta.

3.3 RESULTS

Based on histological sections, we have reconstructed the outflow tract structures in five hearts of the turtle *Emys orbicularis*, at different developmental stages. Using these reconstructions we have examined 1) the way in which septation of the OFT is related to the looping of the heart, 2) the development and fusion of the cushions and 3) the growth and shape of the three-pronged AP septum. We provide all 3D reconstructions discussed in this chapter on the following website: http://bio-imaging.liacs.nl/galleries/. There, the reader can use our 3D-browser application, TDR-viewer (Potikanond and

Verbeek, 2012) to study the modelled anatomy, both as a whole and in separate structures.

First, we address the developmental relation between outflow tract and *cavum pulmonale*, and the implications this has for the positioning of the cushions and the septation of the proximal OFT. This is followed by a detailed description of the development and fusion of the distal, proximal and atrioventricular cushions. Finally, a description is given of the arterial arches and the separation of their roots by the AP septum.

3.3.1 Developmental origin of the cavum pulmonale and its relation to the OFT cushions

In the previous chapter, we described the early looping of the turtle heart and suggested that this double looping gave rise to the *cavum pulmonale*. This finding has implications for the positioning of the valves derived from the distal cushions. It has been proposed that in the mammalian heart the proximal cushions take part in partitioning the ventricular outflow tract, while the distal cushions develop the aortic and pulmonary roots and give rise to the valves guarding these vessels (Van Mierop, 1979). However, theories stating that the valves arise from distal cushions have to account for the shifting position of these valves from their location in the OFT to the arterial roots. Several solutions have been proposed for the mammalian heart, including migration of the aortic root towards the left ventricular outlet (Goor *et al.*, 1972) and absorption of the proximal component of the OFT into the ventricle, resulting in relocation of the aortic root (Anderson *et al.*, 1974; Webb *et al.*, 2003).

The situation in the turtle heart is significantly different. Due to looping of the heart tube at an early developmental stage, the proximal OFT comes to lie ventral to the ventricle (see 2.4). The dorsal wall of the proximal OFT fuses with the ventral wall of the ventricle, forming the horizontal septum (Fig. 3.1C). Through this looping, the proximal component of the OFT develops into the *cavum pulmonale* (CP) and comes to lie in front of the right cavity in the dorsal part of the ventricle, the *cavum venosum* (CV). The observation that the CP derives from looping of a component of the OFT is supported in the current study by the fact that both proximal cushions were found in the CP (Figs. 3.1A and B). The 3D reconstructions provide us with a clear insight into the spatial positioning of these cushions. At Yntema (1968) stages 8 and 15 one of the proximal cushions was seen straddling the horizontal septum, while the other was located in the cranial roof of the CP (Figs. 3.1B and 2). Therefore, we confirm the proximal outflow tract origin of the *cavum pulmonale* and its positional identity as being derived from the proximal component of the OFT.
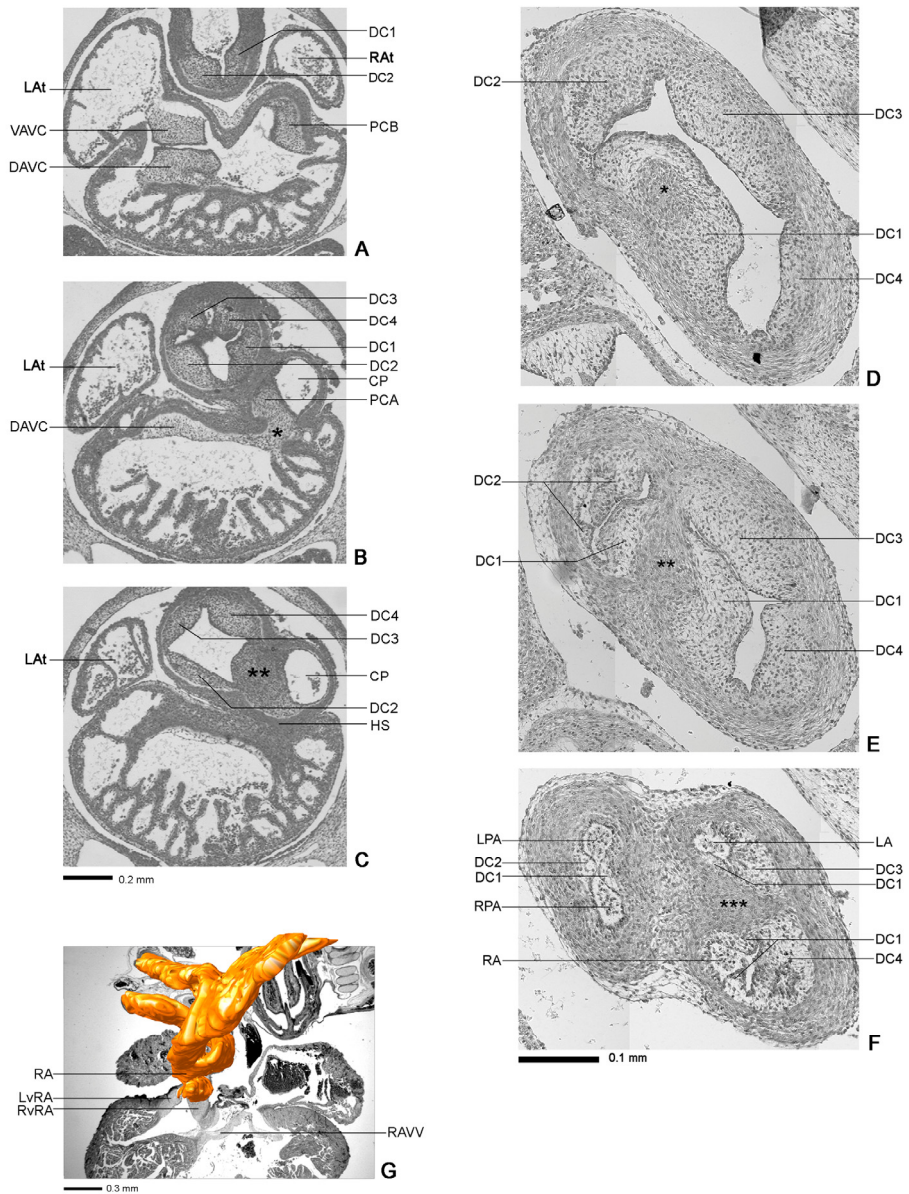
Figure 3.1. Section images of Yntema (1968) stages (A-C) 15, (D-F) 19 and (G) the adult stage (including the 3D reconstruction of the right aortic arch). Abbreviations: CP, *cavum pulmonale*; DAVC, dorsal atrioventricular cushion; DC, distal cushion; HS, horizontal septum; LA, left aorta; LAt, left atrium; LPA, left pulmonary artery; LvRA, left valve of right aorta; PC, proximal cushion; RA, right aorta; RAt, right atrium; RPA, right pulmonary artery; RvRA, right valve of right aorta; VAVC, ventral atrioventricular cushion. The fusion of PCA and the dorsal AV cushion is indicated in B by *; the fusion of DC1 and PCA is indicated by ** in C. The prongs of the AP septum are indicated in D (*, main root), E (**, prong between DC2 and DC3) and F (***, prong between DC3 and DC4).
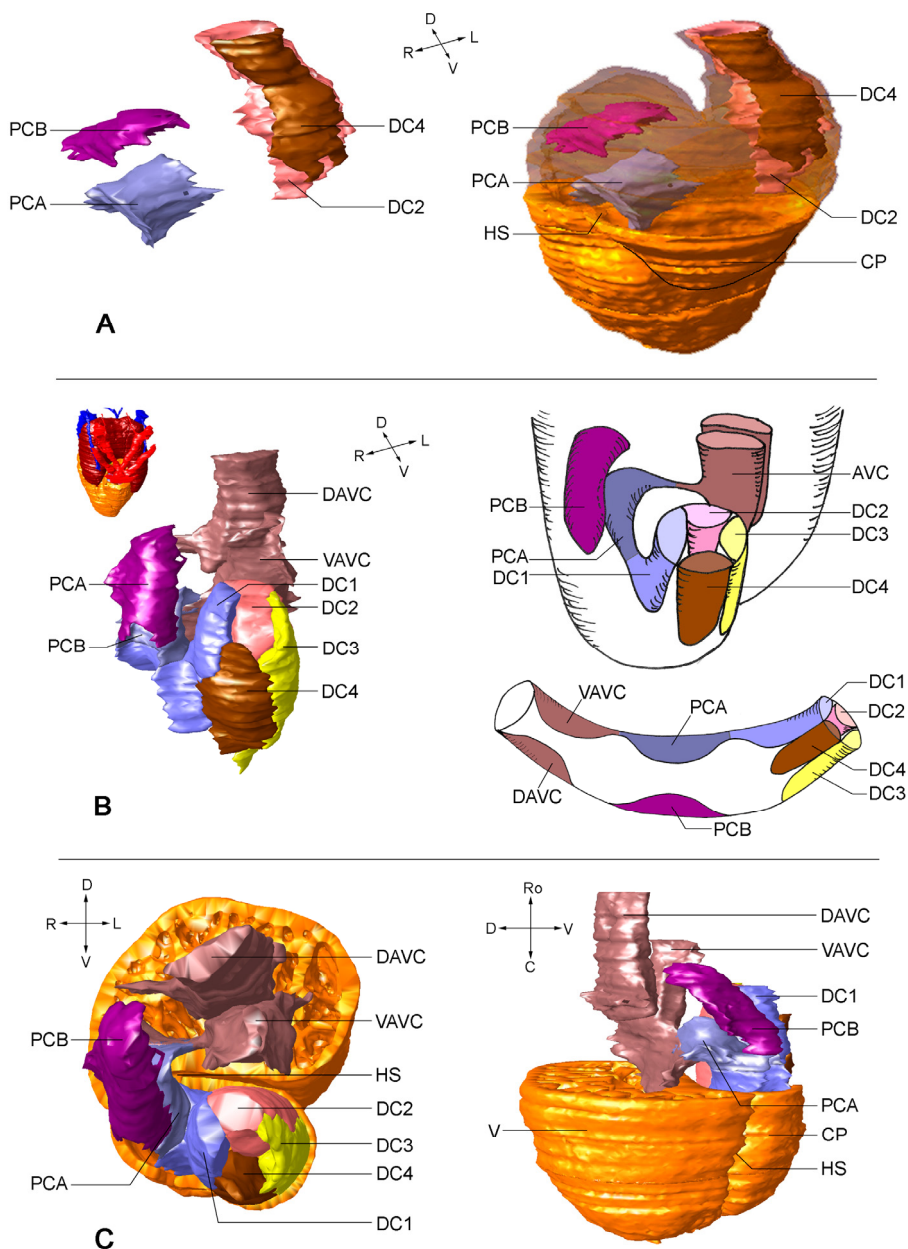
Figure 3.2. Positioning of the distal and proximal OFT cushions. (A) stage 8 in ventrolateral view, isolated (left) and *in situ* with OFT and ventricle, partially transparent (right); (B) stage 15 in ventrolateral view, isolated (left) and schematically represented to clarify the fusion of DC, PC and AVC; (C) stage 15, *in situ* in the caudal half of the ventricle, in rostral view (left) and right lateral view (right). Abbreviations: CP, *cavum pulmonale*; DAVC, dorsal atrioventricular cushion; DC, distal cushion; HS, horizontal septum; PC, proximal cushion; V, ventricle; VAVC, ventral atrioventricular cushion.

The developmental origin of the CP from the proximal OFT has two major implications for the OFT and its cushions. The first concerns the positioning of the distal cushions in the OFT. Unlike the mammalian situation, the distal cushions do not need to be actively relocated within the OFT in order to become aligned with the ventricular outlet. Due to the inclusion of the proximal OFT in the functional ventricle, the distal cushions, and therefore the arterial valves arising from them, come to lie at the outlet of the functional adult ventricle, *i.e.* the distal border of the CP. Meanwhile, the proximal cushions come to lie on the junction between *cavum pulmonale* and the rest of the ventricle, *i.e.* within the functional ventricle. The relative shift in position of the distal cushions, due to the looping, is very clear when comparing the reconstructions of the OFT at stages 8 and 15 (Fig. 3.2). This constitutes a major difference between outflow tract development in mammals and birds on the one hand, and reptiles on the other.

The second implication concerns the septation of the proximal OFT. The separate ventricles of the mammalian heart each become connected to a separate OFT, requiring complete septation of the initial proximal OFT. We found that the proximal component of the turtle OFT does not become divided, since it develops into the undivided CP. The developmental origin of the CP from the OFT, and the two discussed implications, constitute important new insights into the general theory of reptilian heart development.

3.3.2 Development and fusion of distal and proximal cushions

While Hart (1968) states that DC2 and DC4 arise earlier in reptilian development than the other cushions, Qayyum *et al.* (2001) show that DC1 and DC3 develop first in the avian heart. Our findings correspond with those of Hart (1968). At Yntema stage 8, two proximal cushions were found; only two distal cushions were present, and neither of these could be identified as DC1, due to their position in relation to PCA (Fig. 3.2A). We identify these latter two distal cushions as DC2 and DC4, confirming the reported difference between avian and reptilian OFT development. At Yntema stage 15, DC1 and DC3 had developed; at stage 19, DC1 was more prominent than the other three cushions. Several authors describe the fusion of two of the distal cushions with the two proximal cushions in the mammalian heart (Shaner, 1962; Icardo, 1990; Okamoto *et al.*, 2010), as well as in the avian and reptilian heart (Shaner, 1962). At Yntema stage 8, we found the distal cushions and proximal cushions to be clearly distinct. These cushions stayed distinct from the proximal cushions at all stages examined. This is in accord with previous descriptions of DC2 and DC4 (see 3.1), neither of which is described to fuse with a proximal cushion. At Yntema stage 15, we observed a fusion between DC1 and PCA, but no fusion of PCB with any of the distal cushions was found (Figs. 3.2B and C). At stage 16+/17 the fusion between PCA and DC1 was such that the individual cushions

could no longer be distinguished; PCB and DC4 lay much closer than at earlier stages and were nearly touching, but not fusing. At stage 19, PCB and DC4 seemed to fuse with each other, but this could not be confirmed with certainty and will need to be examined further.

At Yntema stages 8 and 15, PCA was seen cranially to PCB and both were in the same horizontal and sagittal planes. At stage 8, neither PCA nor PCB were continuous with the AV cushions, whereas at stage 15 PCA had become continuous with the ventral AV cushion (Figs. 3.2B and C). At the adult stage, tissue at the base of the right AV valve could be seen touching the base of the right aortic leaflet deriving from DC1 (Fig. 3.1G). At all stages examined, PCB was clearly distinct from the AV cushions, which is in conflict with the description by Shaner (1962).

Thus, according to our findings, in the turtle *Emys orbicularis* DC2 and DC4 develop before DC1 and DC3. DC1 fuses with PCA while PCB remains separate from the distal cushions at least until Yntema (1968) stage 16+/17. PCA fuses with the ventral AV cushion between stage 8 and 15, while PCB remains distinct from the AV cushions. This fusion of one distal cusion, one proximal cushion and one AV cushion into a long ridge is also seen in the chick heart (Qayyum *et al.*, 2001). However, the distal cusion that contributes to this fused mass is early-developing in the chick (Qayyum *et al.*, 2001), but develops later in the turtle.

3.3.3 Development of the arterial arches and semilunar valves by means of the aortopulmonary septum

The following description is based on observations in the reconstruction of the adult heart (CP434). The pulmonary artery is positioned ventrally, overlying the *cavum pulmonale*. It bifurcates close to the ventricular outlet into left and right pulmonary arteries. Dorsal to this, the left aortic trunk overlies the *cavum venosum*, and dorsal to this, the right aorta is also continuous with the *cavum venosum*. Close to its origin, the right aorta bifurcates into the aorta proper and the brachiocephalic artery. The latter splits into four smaller arteries: the left and right subclavian arteries ventrally, and dorsal to these, the left and right carotid arteries (Figs. 3.3A-D). The early development of these aortic arches can be seen in the reconstruction of stage 8 (Fig. 3.3E).

Semilunar valves develop at the roots of the arterial trunks, and are derived from the distal endocardial cushions. The allocation of these distal cushions over the arterial roots is brought about by the aortopulmonary septum growing proximally along the OFT and separating the cushions. Up to now, no clear illustrations have been presented in literature of the three-dimensional structure of the reptilian AP septum. Accounts of the AP septum in reptiles have been limited to descriptions, accompanied

by images of histological sections and schematic 2-dimensional drawings (Goodrich, 1958; Shaner, 1962; Hart, 1968).

We have constructed 3D models of this septum at two relevant developmental stages, 16+/17 and 19. These clarify the spatial distribution of what we assume to be neural-crest mesenchyme, of which the prongs of the septum are constituted. For birds and humans the division between the aorta and pulmonary artery has been illustrated as a trouser shape (Bartelings and Gittenberger, 1986; Qayyum *et al.*, 2001), *i.e.* a ridge of tissue from which two prongs extend proximally. However, in agreement with Webb *et al.* (2003) we do not find a clear bar-shaped septal ridge distal to the prongs. The neural-crest derived tissue can be seen to form two wedges between the vessels at the points of their bifurcation. Distal to this point, the arteries are separated by myocardial tissue which is seen to move in between the arches from the outer myocardial sleeve (Figs. 3.1E and F).

As can be seen at stage 19 (Figs. 3.1D-F and 3.4A-E) the mesenchymal tissue extends in three directions; the main root invades DC1 from its distal end and gives off two further prongs. Shaner (1962) describes a Y-shaped septum, but does not specify the direction of the three legs of this Y-shape. From our reconstructions, both secondary prongs can be seen attaching to the main root in a perpendicular direction, *i.e.* in the transverse plane (Figs. 3.1D-F). As such, the tissue resembles a three-pointed caltrop, with each of the prongs perpendicular to the other two (Figs. 3.4D and E). One of these prongs splits off more proximally and this prong divides the aortic arches from the pulmonary arch; together with the main root this constitutes the so-called aortopulmonary septum. It extends from the root between DC2 and DC3, and thereby splits off the pulmonary artery, which therefore contains a segment from DC1 and the complete DC2. It corresponds to the prong extending between DC2 and DC3 in the avian OFT (Qayyum *et al.*, 2001). The position of this split leaves the pulmonary artery furthest removed, both in the ventrodorsal and in the rostrocaudal planes, from the transition between OFT and ventricle, over the free margin of the horizontal septum. This represents the definitive conformation of the adult turtle arterial trunks (Figs. 3.3A-D).

The second prong splits off more distally. It extends between DC3 and DC4 and splits the common aortic channel into left and right aortic arches. This constitutes the so-called aortic septum, which is not found in the avian OFT (Shaner, 1968). The position of the split locates the root of the left aorta just dorsal and to the right of the pulmonary artery; the left aorta contains a segment from DC1 and the complete DC3. The right aorta in turn is located dorsally and to the left of the left aorta and contains a segment from DC1 and the complete DC4.
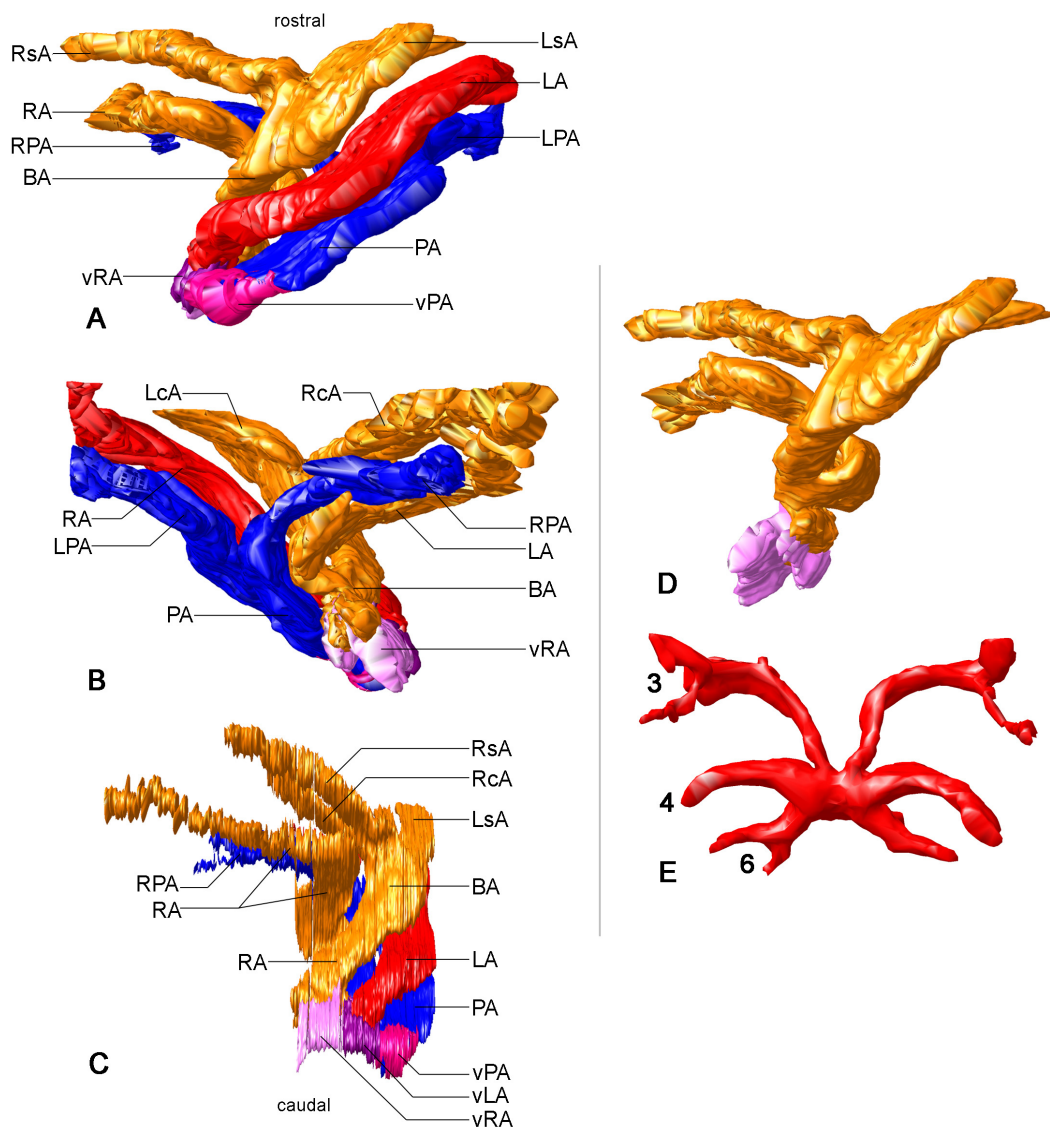
Figure 3.3. Overview of the arterial arches at their origins from the ventricle. (A-D) arterial trunks and their semilunar valves in a reconstruction of the adult heart in ventral (A and D), dorsal (B) and right lateral (C) views; (E) aortic arches 3, 4 and 6 in a reconstruction of stage 8, in ventral view. Abbreviations: BA, brachiocephalic artery; CP, *cavum pulmonale*; CV, *cavum venosum*; LA, left aorta; LcA, left carotid artery; LPA, left pulmonary artery; LsA, left subclavian artery; PA; pulmonary artery; RA, right aorta; RcA, right carotid artery; RPA, right pulmonary artery; RsA, right subclavian artery; v, valves.
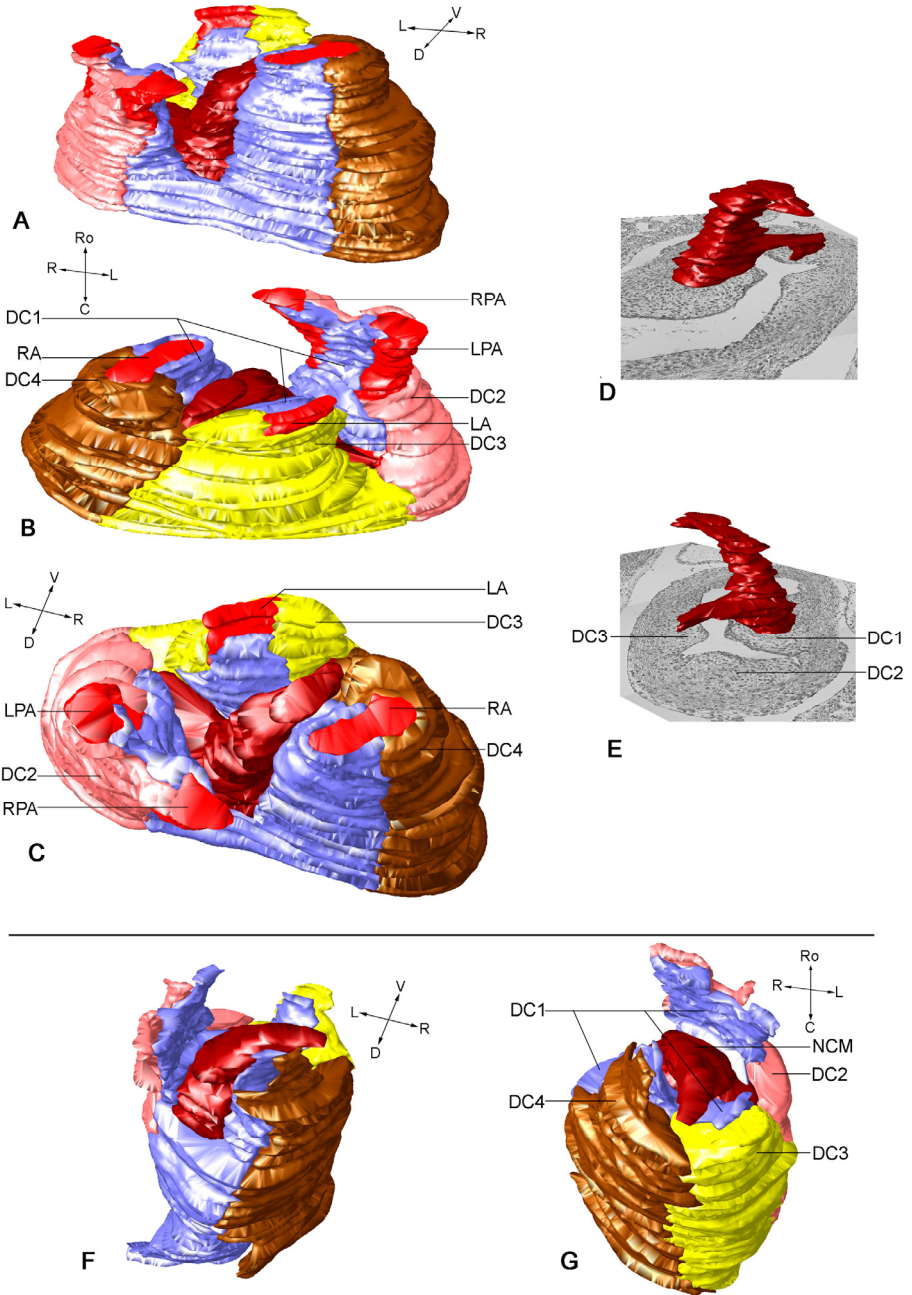
Figure 3.4. Division of the distal cushions by neural crest derived mesenchyme. (A-E) reconstruction of a stage 19 heart showing the aortopulmonary septum and the separation of the distal cushions; (F-G) reconstruction of stage 16+/17 heart showing separation of the distal cushions. Abbreviations: APS, aortopulmonary septum; DC, distal cushion; LA, left aorta; LPA, left pulmonary artery; NCM, neural crest derived mesenchyme; PC, proximal cushion; RA, right aorta; RPA, right pulmonary artery.

3.4 CONCLUSION AND DISCUSSION

We have constructed five 3D models of the outflow tract at different developmental stages in the turtle species *Emys orbicularis*. Examination of these models has yielded significant new insights into the development of the turtle outflow tract. This extends the account, presented in the previous chapter, of the origin of the ventricular cavities or *cava* in this species, by adding additional stages and using high resolution modelling.

We find that the *cavum pulmonale* arises from the proximal OFT segment in the early embryo, although in the adult it is functionally one of the three ventricular cavities. Surprisingly, therefore, this ventricular cavity is developmentally derived from part of the outflow tract that has become displaced into the ventricular region by the looping pattern of the primary heart tube. Therefore the *cavum pulmonale* is functionally, but not developmentally a ventricular structure.

This has important consequences for other aspects of cardiac development. One of these is that the proximal OFT does not become divided, in contrast to the mammalian and avian proximal OFT. Second, the distal OFT cushions do not undergo an active positional shift, but come to lie at the ventricular outlet as a result of the inclusion of the proximal OFT in the functional adult ventricle. Together, these features represent a type of cardiac development not previously described in any vertebrate.

The study presented here is based on 3D reconstructions of histological sections, using high resolution section images. These images were acquired and annotated with our in-house software, and the models are provided to the reader at http://bio-imaging.liacs.nl/galleries/. Schematic drawings can confuse due to their intrinsic level of abstraction. In contrast, 3D reconstructions are more true to nature and allow one to examine a 3-dimensional structure from different viewpoints, thereby clarifying the particular spatial characteristics of the structure. Secondly, by digitally stitching image tiles together high resolution images were obtained, which present a higher microscopic magnification with a higher resolution. Changing solely the magnification or the resolution cannot yield a higher level of detail while retaining the same field of view.

Finally, by allowing the reader full access to both the histological sections and the 3D reconstructions, we hope to optimise the knowledge gained from these models. The viewing applet, available via the website, lets the user browse through the 2-dimensional section images in the image stack, with the option to show our annotations of the relevant structures. In the 3D view, the annotated structures can be examined separately or together and from all viewpoints. In this way the user can verify the findings and further investigate aspects of the outflow tract and relate these to our 3D models of the entire heart, discussed in chapter 2.

CHAPTER 4.   A VERBAL MODEL: AN ONTOLOGY SYSTEM FOR THE VERTEBRATE HEART

διὸ πρῶτον ἡ καρδία φαίνεται διωρισμένη πᾶσι τοῖς ἐναίμοις· ἀρχὴ γὰρ αὕτη καὶ τῶν ὁμοιομερῶν καὶ τῶν ἀνομοιομερῶν

Therefore it is that the heart appears first, distinctly marked off in all the sanguinea, for this is the origin of both homogeneous and heterogeneous parts.

- Aristotle, *De generatione animalium* 740a, v.18

In the previous two chapters we have looked at visual modelling. In this chapter we will shift our attention to verbal modelling. An ontology system is presented, which provides a framework for modelling anatomical, developmental and physiological information. We have chosen to use the vertebrate heart as a proof of concept for the proposed approach, since this facilitates the integration of this study with models from the studies presented in chapters 2 and 3. However, the framework is designed in a generic manner, making it applicable to a wide range of biological structures. In constructing the framework we have taken a novel approach, which has yielded a dynamic ontology system, thereby extending the functionality of ontologies for the field of developmental biology.

4.1 INTRODUCTION

Over the last few years, ontologies have become increasingly important in biology as a formal basis for collecting and sharing knowledge, *cf.* section 1.3.1. Ontologies are used in several fields, such as genetics, biochemistry and anatomy. A vast amount of literature can be found on the use of ontologies in biological sciences, ranging from overview articles (Bodenreider and Stevens, 2006; McEntire, 2002) to textbooks (Baclawski and Niu, 2006; Burger *et al.*, 2008; Staab and Studer, 2004). Journals such as *Computers in biology and medicine* have devoted special issues to bio-ontologies (issues 7-8, vol. 36, 2006) and the well-established journal *Bioinformatics* dedicates a permanent section to databases and ontologies. A collaboration, the OBO foundry, has been set up to bring ontologies in biological and biomedical fields together (Bodenreider and Stevens, 2006; Smith *et al.*, 2007). This foundry contains a number of anatomy ontologies for different species, *e.g.* the *Foundational Model of Anatomy* (FMA; Rosse and Mejino, 2003; Rosse and Mejino, 2008) for humans, the *Mouse Gross Anatomy and Development Ontology* (EMAP; Baldock *et al.*, 2003), the *Amphibian Gross Anatomy* (AAO; Maglia *et al.*, 2007) and the *Zebrafish Anatomy and Development Ontology* (ZFA; Bradford *et al.*, 2011). Each of these ontologies describes the anatomy of the particular species in great detail and is a valuable source for the field of developmental anatomy. However, since they all use different class hierarchies to describe the anatomy, comparative studies between species are often hampered. Therefore we propose an additional approach, which combines the anatomical knowledge of multiple species in one generic class hierarchy. As a case study we use the vertebrate heart. Put another way, the single species ontologies can be described as following a vertical approach (see Fig. 4.1), modelling the entire anatomy of one species. In contrast we use a horizontal approach, modelling the anatomy of one organ system, the heart, for a wide range of species, the subphylum of vertebrates.

The collected knowledge is structured in a system of three ontologies: one comprising all anatomical structures (referred to as the anatomy ontology), another providing developmental staging information (the development ontology) and thirdly the taxonomical species database of the National Center for Biotechnology Information (the NCBI; Sayers *et al.*, 2008). In addition to the ontologies a set of *instances* is used to capture the information for specific species. By founding the class hierarchy of our anatomy ontology on the basic biological distinctions (such as atrial versus ventricular) as opposed to species-specific distinctions (such as right versus left heart chambers) we ensure that any vertebrate species can be added to the system without losing the consistency of the anatomical class hierarchy.

Foundational Model of Anatomy FMA	Mouse Gross Anatomy and Development EMAP	Amphibian Gross Anatomy AAO	Zebrafish Anatomy Ontology ZFA	. . .
Vertebrate Heart Anatomy Ontology				
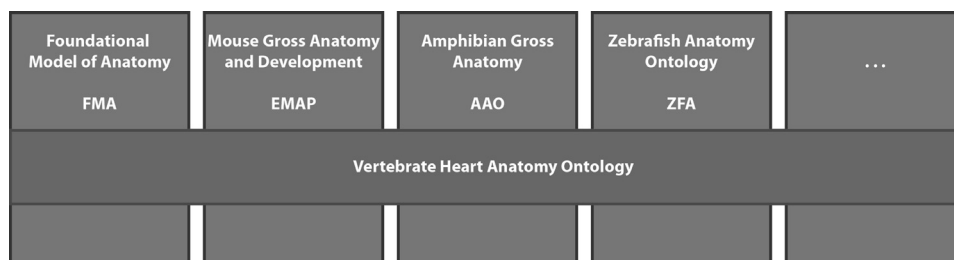
Figure 4.1. Illustration of the vertical approach, adopted by most anatomy ontologies in the OBO Foundry, and the horizontal approach presented here. The one-species ontologies contain the entire anatomy, but of species only, while the proposed ontology contains only data concerning the heart, but includes all vertebrates.

By using these three ontologies, a generic framework is created, in which anatomical, physiological and developmental information for different species can be brought together, by using instances for structures in different species. This makes the system suitable for comparative studies in the field of evolutionary development.

In the remainder of the chapter, we will provide relevant biological background information regarding anatomy, physiology and development, used to construct the knowledge base. Subsequently we will lay out the architecture of the system, discussing the different ontologies, properties and the use of instances. The anatomy and development ontologies have been constructed in the ontology editor *Protégé* and are provided on our website; the reader can download and use the ontologies to include his or her own data or to study and query the content, using an editor such as *Protégé* and/or SPARQL queries. Furthermore an interface is being developed (see section 4.4.4) which allows the user to query the system from our website, without the need for an ontology editor. The ontology system allows one to formulate different types of queries and we describe four of these as case studies in using the system. They concern cross-species comparisons, developmental questions, context-dependent queries about the physiology (using SWRL rules; website SWRL) and 3D visualisation of query results from interface queries. Future additions to the functionality of the system and extensions of the information contained in it will be discussed in section 4.5. Throughout this chapter, whenever we speak of structures, we do so in a biological sense; the terms *classes* and *instances* refer to the elements in the ontology.

4.2 BIOLOGICAL BACKGROUND INFORMATION

The vertebrate heart is a highly diverse structure and although the basic developmental plan is similar for the entire subphylum, the anatomy can vary quite extensively between species. This and the vast body of literature on the topic make it a suitable

case study for our horizontal approach. Through our ontology we aim to provide a framework that can be used to bring together anatomical information for vertebrate species. For this reason three anatomically diverse species have currently been annotated in the system: the zebrafish (*Danio rerio*), the European pond terrapin (*Emys orbicularis*) and the saltwater crocodile (*Crocodylus porosus*). In future, mammalian and avian species will be added to the ontology. Here we will briefly describe the cardiac anatomy and relevant physiology of these species, followed by a short description of developmental staging as used in biology.

4.2.1 Anatomy

In all vertebrates the heart arises as a tube. During development two distinct structures arise in this tube, the atrial and the ventricular chambers. The basic heart can therefore be said to contain four elements: the inflow tract (where blood enters the atria), the atrial compartment, the ventricular compartment and the outflow tract (where blood leaves the ventricle). The classification in the ontology is built on this basic lay-out.

In fish this anatomical plan stays essentially the same; the fish heart has one atrium and one ventricle, connected through the atrio-ventricular canal (the AV canal) and guarded by atrio-ventricular valves (see Fig. 4.2A). Blood enters the atrium through the sinus venosus and leaves the ventricle through the bulbus. In higher vertebrates the heart tube develops into a multi-chambered heart through looping and further dividing. In mammals and birds looping and septation lead to a four-chambered heart, with two
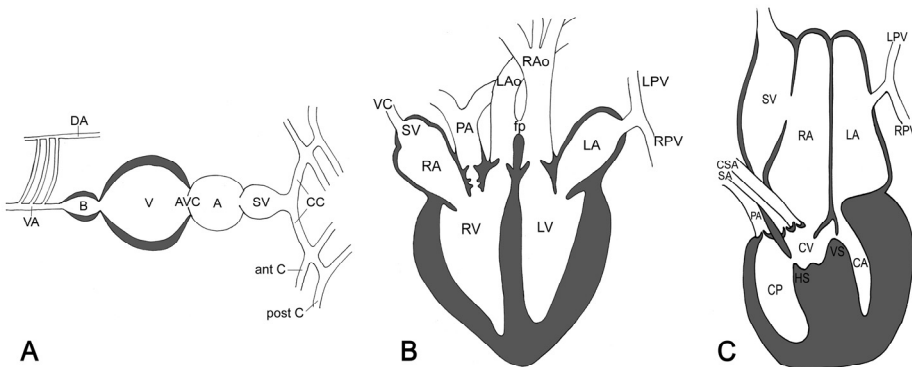
Figure 4.2. Schematic drawings of the hearts of (A) zebrafish, (B) crocodiles and (C) non-crocodilian reptiles. A, atrium; ant C, anterior cardinal; AVC, atrioventricular canal; B, bulbus; CA, *cavum arteriosum*; CC, common cardinal; CP, *cavum pulmonale*; CSA, carotico-systemic aorta; CV, *cavum venosum*; DA, dorsal aorta; fp, foramen of Panizza; LA, left atrium; LAo, left aorta; LPV, left pulmonary vein; LV, left ventricle; PA, pulmonary artery; post C, posterior cardinal; RA, right atrium; RAo, right aorta; RPV, right pulmonary vein; RV, right ventricle; SA, systemic aorta; SV, sinus venosus; V, ventricle; VA, ventral aorta; VC, vena cava.

atria and two ventricles. In reptiles the heart contains two atria and one or two ventricles; the crocodilian order has two completely separated ventricles, divided by an interventricular septum (resembling the mammalian heart but differing significantly in its physiology), whereas the other three reptilian orders have one ventricle, partially divided by two septa into three cavities (see Fig. 4.2B and C). The looping of the non-crocodilian heart is discussed more thoroughly in chapters 2 and 3. As case studies for the generic use of our ontology we have chosen, in addition to the zebrafish, two reptile species representing these two different heart types: the *Crocodylus porosus* of the Crocodilian order and the turtle *Emys orbicularis* of the Testudine order. These two reptile species have very distinct heart structures, not found in other types of vertebrates. As shown in the next paragraph, the classification of the ontology allows users to add these structures without losing consistency in the classification.

4.2.2 Physiology: shunting

In addition to the distinctive anatomy of these hearts the two reptilian species have also been chosen as case studies here for their specific physiology. In both the crocodile and the turtle heart blood flow can be regulated for different physiological situations, according to pressure changes in the different heart compartments (Axelsson *et al.*, 1996; Axelsson *et al.*, 1997; Axelsson, 2001; Bertens *et al.*, 2010; Grigg and Johansen, 1987; Grigg, 1991; Johansen and Burggren, 1980; Hicks and Krosniunas, 1997). The physiology is therefore dependent on environmental factors. We will focus first on the crocodile as a case study for context-dependent querying (as explained in section 4.4.3); this animal is unique in having a completely separated ventricle while at the same time retaining the ability to adapt its blood flow to differing circumstances through shunting (Axelsson, 2001; Burggren, 1987; Axelsson *et al.*, 1996). In addition to the pulmonary artery, the right ventricle also opens into a right aorta. This aorta communicates with the left aorta through the foramen of Panizza (Bertens *et al.*, 2010; Johansen and Burggren, 1980), as can be seen in Fig. 4.2B (structure *fp*). When pressures in the ventricles and aortas change, blood flow can be redirected from the pulmonary to the systemic circuit, causing a pulmonary-to-systemic shunt (a right-to-left or R-L shunt). This allows blood from the right ventricle to bypass the lungs and be recirculated into the systemic circulation (Axelsson, 2001). It is because of this ability to both separate blood flows completely and establish shunts, that Axelsson has described the crocodile heart as probably the most functionally sophisticated found among vertebrates (Axelsson *et al.*, 1996); this makes it an excellent case study for the functionality of our ontology.

Although we focus on the crocodile in presenting the dynamic blood pressure querying of our ontology, this functionality is not limited to this species. Again the

structuring of information in the ontology is kept generic and is therefore applicable to all vertebrates in which shunting is present. The turtle has a very intricate shunting system (*cf.* section 2.2 and Fig. 2.2), including both R-L and L-R shunts, for which information will be added to the ontology. Furthermore, shunting plays an important role during development in for instance the human embryo (between left and right atria through the foramen ovale and between pulmonary artery and aorta through the ductus arteriosus). Since the modelling of shunting in the ontology is generic, shunting via these embryological structures can also be modelled within this framework.

4.2.3 Development

In addition to anatomical structures present in the adult heart, we have chosen to include developmental structures in the ontology system. These are often transient, arising during a particular developmental stage and disappearing (or merging with other structures) at a later stage. For many species, in particular model species such as zebrafish and mouse, conventional staging systems exist; *e.g.* for zebrafish Kimmel *et al.* (1995) have provided names and descriptions of all developmental stages, which are used by scientists all over the world to refer to zebrafish development. Similar stage descriptions have been provided by Yntema (1968) for reptiles, Theiler (1989) for mice, Carnegie (O'Rahilly and Müller, 1987) for humans and Hamburger and Hamilton (1951) for chickens. These staging systems are specific for particular species (or related groups of species) and in our development ontology (described in section 4.3.3) we include the staging systems for the main vertebrate model animals. In addition to these staging systems other features are often used when discussing development; development can be measured in absolute time, *i.e.* hours post fertilization. Alternatively, developmental landmarks can be used, structures that arise at a certain point during development and are used as indicators of developmental progress. For vertebrates the number of somites is often used to this end. As described in section 4.3.3, we use several of these features in our development ontology.

4.3 ARCHITECTURE OF THE SYSTEM

The system is constructed out of four components: three ontologies and a set of instances. In this section we describe these components, starting with the ontology of anatomical structures, followed by a description of its properties and subsequently the ontology of developmental stages. For both ontologies we will explain the biological reasoning behind choosing the classes and properties. Lastly the use of instances is described, including the way these are linked to the first two ontologies as well as to the NCBI database of taxonomy.

4.3.1 The anatomy ontology

We have strived to make the anatomy ontology generically applicable to all vertebrate species. On the highest level in the classification tree a division is made between the superclasses cavitated structures, embryological structures, heart valves, heart septa and tissue types (see Fig. 4.3, presenting a screen dump of the ontology in *Protégé*).

The cavitated structures comprise all anatomical structures that contain a lumen, through which blood can flow, which means that blood flow is located entirely in this superclass. The class of embryological structures comprises all structures that are present during development, but are no longer present in the adult animal. Since some of the embryological structures can also be categorised under one of the other superclasses (*e.g.* the common atrium is an ´embryological structure´ as well as an ´atrial structure´), these classes have been assigned two superclasses and can be found under both. For classes which can have both embryological and adult instances for different species (*e.g.* the sinus venosus, which is a transient structure in humans, but an adult structure in zebrafish), the class itself is not linked to the embryological superclass, but the specific instances are. Having a separate superclass for embryological structures enables the querying of only adult structures for a particular species, by specifically leaving out all structures categorised as embryological. Similarly, the adult structures can be disregarded, in case only structures from development are relevant to a query. Information about developmental staging is also captured in the development ontology (as described below), but by adding the superclass 'embryological structure' to the anatomy ontology we ensure that even when quantitative staging information is not known or included it is still possible to separate the structures qualitatively; furthermore, it means that this information is also available when only the anatomy ontology is used, without linking this to the development ontology.

Heart septa and valves both comprise anatomical structures in the heart which (partially) separate cavitated structures (heart compartments and blood vessels), thereby directing blood flow. The final superclass in the first tree division is of tissue types; this stands apart from the other structures in the sense that its classes are not specific anatomical structures in their own right, but indicate the cell types of which anatomical structures consist. As such they have been assigned their own superclass, next to the anatomical superclasses, allowing the user to link classes and instances to both the relevant anatomical structure and the tissue type.

On the next level in the hierarchy the majority of cavitated structures get divided into arteries, veins and heart compartments, thereby separating the heart proper from the connected blood vessels (which could be extended to include the entire circulation if so required). Subsequently the heart compartments are divided into

Figure 4.3. Screen dump of the partially unfolded anatomy ontology in *Protégé*, showing the main superclasses and several further subdivisions of these; arrows pointing downward indicate that the subclasses of a class are shown, arrows pointing to the right indicate that the subdivisions are not shown.

'inflow', 'atrial', 'ventricular', 'atrioventricular connecting' and 'outflow' structures. This division directly reflects the basic anatomical plan of the vertebrate heart and allows us to add structures of even highly diverse vertebrate species in a consistent system. For instance, even though not all vertebrates have the same number of atrial and ventricular chambers, all have at least one atrial and one ventricular structure. The three species currently annotated in the system illustrate this issue: while the zebrafish has only one atrium and one ventricle, the crocodile has two of both and the turtle has two atria and one ventricle, partially divided into three cavities. In this sense the

ontology differs from the main ontologies in OBO, which focus on the entire anatomy of one species.

Note that general information, applicable to all vertebrates, is added in classes; species-specific information however, is added in instances, as described below. A clear illustration of the importance of this difference is given by the structures in the outflow tract: the outflow tract comprises structures as the bulbus, conus and truncus. These terms are used in diverse ways in different species and by using instances to add the species-specific information in our system we avoid confusion about the definitions; the classes in the anatomy ontology are general, which means we do not have to include different classes of for instance the bulbus for the different uses of the term in different animals. Each instance on the other hand is specific for one species and therefore only has one exact definition, preventing ambiguities. The same applies to the inflow structures, of which for instance the sinus venosus can be either an embryological structure (as in humans), or an adult structure (as in zebrafish).

Further subdivision of the hierarchy follows the same method: at each step down a level in a tree, superclasses are chosen in such a way that each vertebrate anatomical class can be categorised in a consistent way. The ontologies (provided as OWL files), as well as additional information about the entire ontology system, are available at http://bio-imaging.liacs.nl/ontologies/ontologies.html.

Besides being directly based on biological distinctions, the hierarchy in our anatomy ontology is also compatible with the class structuring of CARO (Haendel *et al.*, 2008). CARO is an upper-ontology for anatomy terms. Top-level classes of our anatomy ontology can be placed under CARO classes; the heart is a subclass of the CARO class 'cavitated compound organ' and nearly all of our structures can be placed as sub-classes under the CARO class 'compound organ component'. This is a relevant feature for future integration and interoperability purposes.

4.3.2 Properties used in the anatomy ontology

Apart from the *is_a* relationship with which the class hierarchy is built, information about the relationships between the different classes and instances is provided through a set of properties. These can be either object properties, linking classes or instances together, or data properties, assigning pieces of data to classes or instances, *e.g.* integers, strings, booleans or floats. The properties specify spatial, developmental, taxonomical and physiological information.

To include information about the spatial relationship between structures we use the object properties *part_of* and *bounds*; in *part_of* a distinction is made between

constitutional_part_of and *regional_part_of*, as in the FMA (Rosse and Mejino, 2003; Rosse and Mejino, 2008; Mejino *et al.*, 2003). The property *constitutional_part_of* indicates that a structure is a fundamental element of a bigger structure; the parts are not just positioned in the bigger structure, but the bigger structure consists of the parts (*e.g.* the atrial and ventricular parts of the heart). The property *regional_part_of*, on the other hand, signifies that a structure is positioned in another structure, *e.g.* the atrioventricular valves are positioned in the atrioventricular canal; without the valves the canal is still a valid structure. Furthermore, the property *bounds* denotes that a structure is a boundary of another structure; this is particularly relevant for valves and septa.

Physiological information concerning blood flow is provided through a set of properties. Some of these are object properties, specifying which structures receive blood from which other structures. This can change for different physiological situations by means of blood shunts, as described above, and we use a hierarchy of several properties to enable modelling each of these situations. In addition to the object properties we use a set of data properties, with which the blood pressure values can be stored, for each of the cavitated structures and again for different physiological states. The implementation of these properties and their functionality, based mainly on inheritance rules, will be discussed in section 4.4.3.

Taxonomical information is provided by the *structure_from_species* property; this property links instances of anatomical structures to species codes in the taxonomical database of the NCBI (Sayers *et al.*, 2008). We chose to make use of this comprehensive, well established database, as opposed to including classes for the species in our system in the anatomy ontology itself; in this way we incorporate the knowledge gathered in the NCBI database to optimize the functionality of our system for comparative studies in the field of evolutionary development. Since the NCBI codes are widely accepted, using these makes our system easily accessible and amenable to interoperability approaches; *i.e.* no knowledge is needed of the naming convention in the anatomy ontology itself, instances of particular species can be found using the property *structure_from_species* with the corresponding NCBI code. Furthermore, by using this external database we maintain a division between the generic knowledge in the anatomy ontology and the species specific information, preserving the generic nature of the anatomy framework.

Two types of developmental information are included through object properties; the property *develops_from* links an anatomical structure to the embryological structure that it developed from. Secondly, information about the developmental stage of structures is included using the properties *end_stage* and *start_stage*. The uses of these are explained in the next paragraph, which describes the structure of the development ontology.

4.3.3 The development ontology and its properties

In several anatomy ontologies, such as the ontology of *Human Developmental Anatomy*, the EHDA (Hunter *et al.*,2003), the hierarchy starts with a division in developmental stages. The anatomical structures that are present at that time point are added to the stage structure, using a *part_of* relation. Although this gives a good overview of the entire anatomy for each stage, it means that the same biological structure exists in the ontology as many times as there are stages at which it is present. Not only is this potentially confusing, it also increases the size of the ontology dramatically, since it forces the developer to add all relevant properties for each of the versions of the biological structure.

An alternative solution, adopted here, is to use developmental stages and singular versions of the anatomical structures and to relate these via properties, which denote the start stage and end stage for each of the anatomical structures. This way, querying for a particular developmental stage remains possible, without having to include a separate version of a biological structure for each of the stages it is present in. Some ontologies, such as the ZFA, include the developmental stages in the anatomy ontology itself. We have opted to build a separate ontology containing the developmental stages. Whereas the anatomical classes are generically applicable, the instances and developmental stage classes are specific to individual species. By setting these apart from the anatomy ontology we separate general information from species-specific information, leaving the framework generically applicable. Furthermore the development ontology serves a purpose in its own right by bringing together and relating staging information for diverse species (as described below) and can be used independently from the anatomy ontology. Lastly, the division enables the use of a different external ontology for the developmental stages, if so required.

Instances of classes in the anatomy ontology can be linked to developmental stages in the external ontology by means of the object properties *start_stage* and *end_stage* in the anatomy ontology. Since staging information is not always exactly known for all structures the start and end stage properties both have two additional properties (*_at_earliest* and *_at_latest*). These serve as a *terminus post quem* and *terminus ante quem* and thereby enable the inclusion of inexact staging information. Limiting contributors to the ontology system to providing only exact staging properties results in loss of information, since some structures are only known to appear or disappear after or before a certain stage. Using these additional properties developmental ranges for the presence of structures can be established.

As discussed above, each species or group of related species has its own staging system. We include in the development ontology staging systems for the major vertebrate model species. In some cases information is available to map these systems

onto each other (Schneider and Norton, 1979) and in those cases we have done so with the property *corresponds_to*; note that this property can also be used when only partial mapping information (only for certain stages) is available. However, since developmental processes differ between species (*i.e.* in some species some structures may arise earlier while others arise later) it is often difficult to exactly relate the stages from different staging systems to each other. We therefore also include other features to enable comparison of development between species. The absolute developmental time for the start of a developmental stage, in hours post fertilization, is included using *starts_at_hour*, as well as a relative measure of development in percentages (normalizing time from fertilization to adulthood), using *percentage_of_development* (Bathoorn *et al.*, 2010). Furthermore the number of somites is provided, which is a well established means of assessing vertebrate development, using *somite_count*. In other words we include, in addition to the detailed species specific staging, an absolute time measure, a relative time measure and a time-independent measure using landmarks.

4.3.4 Instances for specific species

We have designed the anatomy ontology as a generic framework, meant to be applicable to all vertebrates as an aid for annotation. The classes therefore do not contain species-specific information. However, we also set out to collect and include information for a number of vertebrate species, either because of their use as model organisms or because of their particular heart anatomy. In doing so we aim to present an ontology system with a predefined set of data, which can be used for comparative studies with data of other species or developmental stages. The anatomical data for the specific species are added to the ontology using instances, known in *Protégé* as individuals. In other anatomy ontologies, dealing with one particular species, the classes are by default species-specific and instances are not used to differentiate between species.

Whenever an instance of a particular class is created, the instance is given the same name as the class, preceded by a two letter code; this code indicates the species to which the instance belongs, *e.g.* DR for a structure of the species *Danio rerio* and EO for *Emys orbicularis*. In this way the class 'Sinus_venosus' in the anatomy ontology can possess several instances for different species, such as 'EO_Sinus_venosus' and 'DR_Sinus_venosus'. These codes are only used to provide unique names for our individuals and are not used semantically. As mentioned before the instances are linked to their corresponding species numbers in the NCBI taxonomical database using properties in the anatomy ontology, allowing a semantic distinction between species.

The hierarchy in the anatomy ontology, together with the instances, represents a gradual transition from general to specific knowledge. We add information on the most general level possible, *i.e.* when for a certain anatomical structure information is known to be valid for all vertebrates possessing that structure, this information is added at the class level; this comprises object properties, synonyms and comments, which we use to provide anatomical definitions. Through inheritance rules this information is also automatically attached to the instances. Information that does not hold for all possible instances of a class is added on instance level only. This notion of hierarchy corresponds to the semantics of the OWL ´SubClassOf relation´ (website OWL-class hierarchies). The inheritance rules also similarly apply to properties; *e.g.* every individual that is linked to another individual using the c*onstitutional_part_of* relation, is also automatically connected with the *part_of* relation. The notion of property hierarchy corresponds with the semantics of the OWL ´SubObjectPropertyOf relation´ (website OWL-property hierarchies). To model the context-dependent behaviour of blood flow we also use multiple properties in a hierarchy reflecting the complex physiological behaviour, as discussed below.

4.4 CASE STUDIES

The broad setup of the system allows for various types of queries. We describe four querying uses here. For each of the uses an example question is formulated and the corresponding SPARQL query is given, along with the result set as produced by the Pellet Reasoner (Sirin *et al.*, 2007). The ontologies are provided as OWL files on our website, http://bio-imaging.liacs.nl/ontologies/ontologies.html, and can be downloaded there. In addition to this we also aim at facilitating the use of the system by presenting an interface on our website.With this interface the user can formulate queries in a web form, which subsequently get translated into SPARQL queries. The design and construction of the interface are discussed in full in section 4.4.4. The most recent stable version of this interface, along with additional information on its use, can also be found on our website. Although currently not all queries described here are feasible using the interface, the full functionality of the system can be employed by performing SPARQL queries directly on the OWL files.

4.4.1 Cross-species comparisons

The main purpose of the horizontal approach is to allow for comparative studies between species. Using the NCBI taxonomy database we can ask questions including taxonomical information. An example of such a query and the way in which it is handled by the system is shown below:

"Find all heart septa present in all Sauropsida species included in the ontology". This informal query gets translated into a SPARQL query, shown below with the result set:

```
SELECT ?x
WHERE {
    ?x rdf:type anatomy:Heart_septum.
    ?x anatomy:structure_from_species ?y.
    ?y rdfs:subClassOf ncbi:NCBITaxon_8457.
}
```

Query Results (10 answers):
```
x
==========================
EO_Aorticopulmonary_septum
EO_Interatrial_septum
EO_Aortic_septum
EO_Horizontal_septum
EO_Septum_primum
EO_Vertical_septum
CP_Aorticopulmonary_septum
CP_Interatrial_septum
CP_Aortic_septum
CP_Interventricular_septum
```

This result set allows one to study differences in particular structures between species of the same phylogenetic group. Using the information contained in the NCBI database, more complex queries are also feasible; questions can for instance be postulated for all species excluding particular taxonomical groups. Also questions regarding evolutionary development could be formulated, such as "Do the similarities in cardiac anatomy of the mouse, the chick and the zebrafish reflect their taxonomic relationships?". To this end taxonomic distances could be compared to graph distances of the particular RDF subgraphs of different species.

4.4.2 Developmental studies

As a proof of concept all relevant developmental structures of the zebrafish have been included in the instance set (based on Kimmel *et al.*, 1995; Hu *et al.*, 2000) and linked to stages in the development ontology. Here we describe an example query, concerning the development of the zebrafish, followed by the result set:

"Find all heart compartments in the zebrafish that develop before Kimmel stage prim-15"

```
SELECT ?x
WHERE {
    ?x rdf:type anatomy:Heart_compartment.
    ?x anatomy:structure_from_species ncbi:NCBITaxon_7955.
    ?x anatomy:start_stage ?y.
    ?y development:order_number ?a.
    development:Kimmel_prim_15 development:order_number ?b.
    FILTER ( ?a < ?b )
}
Query Results (6 answers):
x
==========================================
DR_Presumptive_atrium_primitive_heart_tube
DR_Presumptive_bulbus_arteriosus
DR_Presumptive_atrium_heart_tube
DR_Presumptive_ventricle_heart_tube
DR_Presumptive_sinus_venosus
DR_Presumptive_ventricle_primitive_heart_tube
```

Whereas this query results in a list of all structures that arise before prim-15 (regardless of whether they disappear again and if so, when), similar queries can be formulated for ranges of developmental stages as well, specifying start and end stages within which structures arise and/or disappear, or studying the structures that are present at one particular developmental stage.

Furthermore information about taxonomy and development can be included in queries, providing answers to questions such as: "Does the anterior heart field arise first, in a relative sense (by normalising development time), in mouse or in chick?". For this the percentage of development and the somite number can be used.

4.4.3 Context-dependent queries for physiology

Our ontology has been designed to deal with the physiology of blood flow in different circumstances, which might lead to shunting in particular organisms or during particular stages of development. To model this changing blood flow, we make use of property semantics in OWL and SWRL logic. In our design qualitative information can be included on blood flow between cavitated structures using two properties: *can_receive_blood_from* and *always_receives_blood_from*. Quantitative information

about the blood pressure values of the cavitated structures for different shunting situations can be added using the properties *systolic_pressure_no_shunt*, *systolic_pressure_LR_shunt* and *systolic_pressure_RL_shunt*. To distinguish between different shunting situation the boolean properties *RL_shunt* and *LR_shunt* can be set to true or false. Below we describe how the system uses these properties to model the blood flow for a specified shunting situation, followed by an example query.

In the most stringent case, we know that blood flow between certain structures is the same for different physiological states and therefore these structures are connected using *always_receives_blood_from*; in the most relaxed case on the other hand, when blood flow between two structures only occurs during a specific physiological state and no generally true information can be given for all physiological states at once, the structures are connected using *can_receive_blood_from*. Naturally, due to inheritance rules, structures connected by *always_receives_blood_from* are also automatically connected by *can_receive_blood_from* for all states. A third property, *currently_receives_blood_from* is used dynamically, *i.e.* it is not used to add static, generally valid information about structures, but can be changed by the system, depending on the shunting situation. Whenever a physiological situation is specified, this property is set accordingly, as described below. The hierarchy between the preceding object properties is dictated by the implications each property has. Since we want to model the following logic:

x always_receives_blood_from y => x currently_receives_blood_from y
x currently_receives_blood from y => x can_receive_blood_from y

we place *can_receive_blood_from* on the most general level in the hierarchy and *always_receives_blood_from* on the most specific level. SWRL rules test the occurrence of particular physiological states and make sure that the correct structures are connected for that particular situation by the property *currently_receives_blood_from*. This testing of the physiological situation is done by checking the boolean properties *RL_shunt* and *LR_shunt*; the common physiological state (the default situation) is considered to have no shunting, meaning both boolean properties are set to false. When we want to study a situation in which shunting occurs we can set the corresponding property (for right-to-left or left-to-right shunting) to true in the query. Note that both properties can be set to true, which represents a biologically valid bidirectional intracardiac shunt (Johansen and Burggren, 1980). We are currently working to implement modelling the blood pressure for this situation; the additional properties and rules are not presented here. A high level overview of the implementation of the system is given below:

1. Both shunting properties are set to the situation under study, either true in case of shunting or false (the default state); this is done for the class Cardiovascular_system.

2. Every class that is part of the class Cardiovascular_system 'inherits' the shunting properties and its values.

3. For each class the blood pressure value for that particular state is set in the property *current_systolic_pressure*. This value is assigned by the system on the basis of the shunting and blood pressure properties.

4. When the blood pressure values for all relevant classes have been set for that particular situation, the following rule is used to determine blood flow: if two structures (x, y) are related with a *can_receive_blood_from* property and x has a higher *current_systolic_pressure* than y, then y *currently_receives_blood_from* x.

All blood flow, shunting and blood pressure properties are set manually, except for the properties *currently_receives_blood_from* and *current_systolic_pressure*, which are never used during the addition of new species, but are dynamically assigned by the reasoner using the following SWRL rules:

LR_shunt(?x,false) $\land$ RL_shunt(?x,false) $\land$ systolic_pressure_no_shunt(?x,?p) $\Rightarrow$ current_systolic_pressure(?x,?p).

LR_shunt(?x,true) $\land$ RL_shunt(?x,false) $\land$ systolic_pressure_LR_shunt(?x,?p) $\Rightarrow$ current_systolic_pressure(?x,?p).

LR_shunt(?x,false) $\land$ RL_shunt(?x,true) $\land$ systolic_pressure_RL_shunt(?x,?p) $\Rightarrow$ current_systolic_pressure(?x,?p).

can_receive_blood_from(?x,?y) $\land$ current_systolic_pressure(?x,?p) $\land$ current_systolic_pressure(?y,?q) $\land$ (?p<?q) $\Rightarrow$ currently_receives_blood_from(?x,?y)

Using this system, queries can be formulated regarding the blood flow during the normal physiological state as well as during periods of shunting. An example of a possible biological question is given below, with the SPARQL query and the resulting blood pressure values (with 'a' denoting the right ventricle and 'b' the left ventricle):

"What are the systolic pressures for the right and the left ventricle in a situation without shunting and in a situation with right-to-left shunting?"

```
SELECT ?a ?b
WHERE {
    anatomy:CP_Right_ventricle anatomy:current_systolic_pressure ?a.
    anatomy:CP_Left_ventricle anatomy:current_systolic_pressure ?b.
}
```

No shunt:
Query Results (1 answers):
a | b
=======
22 | 67

Right-to-left shunt:
Query Results (1 answers):
a | b
=======
47 | 50

4.4.4 Querying using the interface and 3D visualisation of query results

Querying the system will result in textual output, listing the classes and/or instances corresponding to the query. As a particular feature of the interface we have extended the querying results with the possibility of visualising the textual results, using previously built 3D reconstructions. Currently, we have 3D reconstructions of the adult zebrafish heart and of several developmental stages of the turtle heart, presented in chapters 2 and 3. All models can be viewed online using out TDR-viewer (*cf.* chapters 2 and 3).

The interface for querying our ontology system is currently under development. However, the intended design and construction have been outlined and are described here. The interface consists of a web form, divided over four tabs, enabling the user to formulate a query. No hierarchy of priority exists between the tabs and the user is free to chose which filters he or she wants to use. While the form is being filled in by the user, the remaining choices the user can make are automatically adjusted by looking up the remaining possibilities, given the user's previous entries. The first tab lets the user select structures of interest. In addition to particular structures we allow groups of structures to be chosen as well, *e.g.* selecting all cavitated structures. A second tab lets the user specify the species he is interested in. Only species that have instances in our ontology can be selected. In another tab the staging system, corresponding to the specified species, is selected (*e.g.* Kimmel for zebrafish or Theiler for mouse). Here the user will be presented with several options for setting the staging restrictions; different staging methods can be used (for instance the stage names for a particular species, the absolute time of development in hours or merely the distinction between adult and embryological structures) and the user can indicate staging with an exact stage or a range of stages. Lastly, there is a fourth tab which lets the user filter the information using the properties linked to structures and instances. At every step in the

form, a SPARQL query is produced, creating a subgraph of all the solutions that still match the restrictions given by the user in the form. When the user submits the query form, a list of all the instances in our ontology is returned. If a 3D reconstruction is available for this species at this stage, a link is provided to the TDR-viewer, with the correct model and structures already selected in the parameters.

The implementation of the form is done using AJAX calls to Java servlets that look up the relevant information from the ontology using the JENA library and SPARQL template queries. While this form is currently focused on 3D visualisations, we aim to extend its functionality to allow more complicated queries, incorporating more detailed information provided in the anatomy and development ontologies, combined with the NCBI database.

The following example illustrates the operation of an interface query. A biologist has a request: "Show all endocardial cushions from the species *Emys orbicularis* that are positioned in a compartment of the heart." This request is filled out in a web form and is internally translated into a query, shown below with result set:

```
SELECT ?x WHERE {
    ?x anatomy:structure_from_species ncbi:NCBITaxon_82168,
    ?x rdf:type anatomy:Endocardial_cushion,
    ?x anatomy:regional_part_of anatomy:Heart_compartment.
}
```

Query Results (4 answers):
x
==
EO_Atrioventricular_cushion
EO_Distal_cushion_outflow_tract
EO_Proximal_cushion_outflow_tract
EO_Messenchymal_cap

This result set is subsequently visualised in the TDR-viewer, a screen dump of which is shown in Fig. 4.4 (see also Fig. 2.4). The viewer renders scenes on the client-side using contour information from TDR-3Dbase. In earlier work by Wong *et al.* (1999) scenes were generated at the server side. Our design focuses instead on the ability to serve multiple clients at the same time with minimal bandwidth.

This architecture has two separate components for the heart terms and their 3D models. In a paper by Köhn *et al.* (2004) a system is described that combines the two types of information into a single ontology. Because we incorporate multiple species, this approach is not feasible for our system.
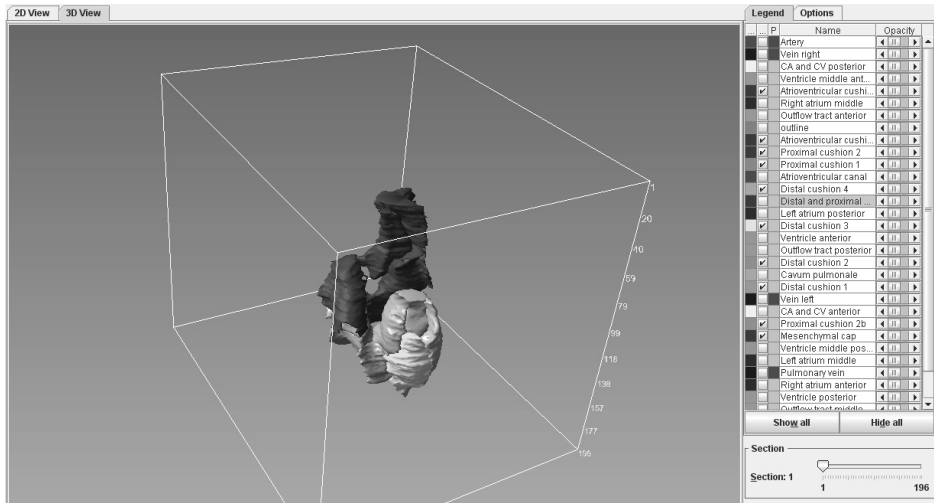
Figure 4.4. Screen dump of the TDR-3D viewer, in which a set of structures is visualised for the turtle species *Emys orbicularis*, corresponding to the query described in the text.

4.5 CONCLUSIONS AND FUTURE WORK

Over the last few decades several ontologies have been created comprising anatomical information. We add to this rapidly growing field by presenting a novel approach centred around organ systems; we have elaborated this approach for the vertebrate heart. As opposed to 'vertical' modelling, focusing on the entire anatomy of one species, we propose a 'horizontal' way of modelling, looking at one organ for a wide range of species simultaneously. This is the key feature of the presented study. The system we have constructed serves a double purpose: the anatomy ontology can be used to annotate information for additional vertebrate species, allowing researchers to supplement the system with information about their own species. Secondly, the instance information that is currently presented for different model animals forms a knowledge base which can be queried to further our knowledge in the developmental, anatomical and physiological aspects of the vertebrate heart. The system is provided on our website and can be used either by directly using the OWL files (to complement or study the information) or by formulating queries using the interface. Note, however, that the visualisation of structures using the TDR-viewer is currently not available, although a static example of the envisioned functionality is provided on the website.

Here we have presented case studies of the vertebrate heart, but the architecture of the system makes it applicable to a wide range of organs and anatomical systems. Additionally, using NCBI accession numbers instead of locally defined name

spaces to distinguish species, opens the system up to semantic interoperability. Using this generic approach we have build an open and easily extendible system.

We are currently incorporating instance information for avian (chick) and mammalian (mouse) hearts, as well as relevant anatomical classes, to the anatomy ontology. For all species in the ontology, developmental information is collected and linked to the development ontology. Furthermore future work will be directed to extending the functionality by providing more querying options; we strive for interoperability with ontologies relating to tissue types phenotypes and diseases. Lastly, the user interface will be amended to allow the user more freedom in formulating complex queries in the field of evolutionary development, anatomy and physiology. New instances of the interface will be made suitable for online use. To this end usability studies will be performed.

CHAPTER 5.   AN ALGORITHMIC PROCESS MODEL: MODELLING GRADIENTS USING PETRI NETS

Omnia mutantur, nihil interit.

Everything changes, nothing perishes.

- Ovid, *Metamorphoses* XV, v.165

The third category of modelling approaches discussed in chapter 1 comprises algorithmic process models, among which Petri nets, and the current chapter presents a case study of the application of Petri nets to the field of developmental biology. We have chosen to model the process of gradient formation, since this process is modular and concurrent in nature and can be placed in a hierarchical structure with other simultaneous and interlinked processes. Petri nets are distinguished by their ability to model these features; the combination of modelling approach and the process to be modelled therefore provides an optimal context to fully explore the possibilities of this approach for developmental biology. Here a qualitative method is applied, not taking into account exact numerical data. In the next chapter a more quantitative approach is applied to the same developmental process.

Based on: Bertens L.M.F, Kleijn J., Koutny M., Verbeek F.J., 'Modelling gradients using Petri nets', in: *Proceedings International Workshop on Biological Processes & Petri Nets (BioPPN) Braga, Portugal, June 21 2010*, 55-69.

5.1 INTRODUCTION

Petri nets (Reisig and Rozenberg, 1998) have been shown to be very promising for molecular and cellular biology, in particular for metabolic, signalling and gene-regulatory networks (see *e.g.* Banks, 2009; Banks *et al.*, 2009; Chaouiya, 2007; Gilbert and Heiner, 2006; Gilbert *et al.*, 2007; Heiner *et al.*, 2008; Koch *et al.*, 2004; Steggles *et al.*, 2006; Talcott and Dill, 2006). We believe Petri nets to be useful for higher level developmental processes as well, *e.g.* on tissue and organ level. Therefore this chapter concerns the use of Petri nets as an abstract modelling tool for higher level processes in the organism, taking cells as central elements. To this end we have selected one developmental process, the formation of a morphogen gradient, as a case study for this approach. This process helps instigate the differentiation of cells along the developing axis in the organism. In early development, gradients are crucial (Wolpert, 2002) and finding a modelling solution for the generic process of gradient formation will not only serve the theoretical goal of investigating the use of Petri nets for developmental biology. It will also, in a more practical sense, be useful for the modelling of other developmental processes in which gradients play a role. By staying very close to the biological sequence of events in gradient formation, rather than focusing on a concrete outcome, the model should be generally applicable and robust. The implications of this approach are further addressed in chapter 6.

Throughout this chapter the emphasis will be on abstraction and modelling decisions, as opposed to implementation of specific biological data (which will be addressed in chapter 6). The main question dealt with in this chapter is: how can Petri nets be used to model higher level developmental processes, which focus on cells as the central units? We present a basic Petri net, modelling gradient formation, which serves as a proof of concept for our approach. In the remainder of this chapter we outline the biological background of gradient formation, we describe our modelling decisions and we present the model. In the last section the possibilities of the model and future work are discussed.

5.2 PT-NETS WITH ACTIVATOR ARCS

For a general introduction to Petri nets we refer to (Reisig and Rozenberg, 1998). Here, we use PT-nets with activator arcs (Kleijn and Koutny, 2007) and a maximally concurrent execution rule (Burkhard, 1983).

Petri nets are defined by an underlying structure consisting of *places* and *transitions*. These basic elements are connected by directed, *weighted arcs*. In the Petri net model considered in this chapter, there are moreover *activator arcs* connecting places to transitions. In modelling, places are usually the passive elements, representing

local states, and transitions the active elements. Here, global states, referred to as *markings*, are defined as mappings assigning to each place a natural number (of *tokens* corresponding to available resources).

A *PTA-net*, is a tuple $N = (P, T, W, Act, m_0)$ such that:
– *P* and *T* are finite disjoint sets, of the *places* and *transitions* of *N*, respectively.

– $W : (T \times P) \cup (P \times T) \to \mathbb{N}$ is the *weight function* of *N*.
– $Act \subseteq P \times T$ is the set of *activator arcs* of *N*.

– $m_0 : P \to \mathbb{N}$ is the *initial marking* of *N*.

In diagrams, places are drawn as circles, and transitions as boxes. Activator arcs are indicated by black-dot arrowheads. If $W(x, y) \geq 1$, then (x, y) is an *arc* leading from *x* to *y*; it is annotated with its weight if this is greater than one. A marking *m* is represented by drawing in each place *p* exactly *m(p)* tokens as small black dots. We assume that each transition *t* has at least one input place (there is at least one place *p* such that $W(p, t) \geq 1$).

When a single transition *t* occurs ('fires') at a marking, it takes tokens from its input places and adds tokens to its output places (with the number of tokens consumed/produced given by the weights of the relevant arcs). Moreover, if there is an activator arc $(p, t) \in Act$, then transition *t* can only be executed at the given marking if *p* contains at least one token, without the implication of tokens in *p* being consumed or produced when t occurs. Thus, the difference with a *self-loop*, *i.e.,* an arc from *p* to *t* and vice versa, is that the activator arc only tests for the presence of tokens in *p*.

We define the executions of *N* in the more general terms of simultaneously occurring transitions. A step is a multiset of transitions $U : T \to \mathbb{N}$. Thus *U(t)* specifies how many times transition *t* occurs in *U*. (Note that if we exclude the empty multiset, single transitions can be considered as minimal steps.) Step *U* is *enabled* (to occur) at a marking *m* if *m* assigns enough tokens to each place for all occurrences of transitions in *U* and, moreover, all places tested through an activator arc by a transition in *U*, contain at least one token. Formally, step *U* is enabled at marking *m* of *N* if, for all $p \in P$:

– $m(p) \geq \Sigma_{t \in T} U(t) \cdot W(p, t)$
– $m(p) \geq 1$ whenever there is a transition *t* such that $U(t) \geq 1$ and $(p, t) \in Act$.

If *U* is enabled at *m*, it can be *executed*, leading to the marking *m'* obtained from *m* through the accumulated effect of all transition occurrences in *U*:

– $m'(p) = m(p) + \Sigma_{t \in T} U(t) \cdot (W(t, p) - W(p, t))$ for all $p \in P$.

Finally, a step U is said to be *max-enabled* at m if it is enabled at m and there is no step U' that strictly contains U (meaning that $U' \neq U$ and $U(t) \leq U'(t)$ for all transitions t) and which is also enabled at m. And we write $m[U\rangle m'$ if U is max-enabled at m and execturion of U at m leads to m'. A (max-enabled) *step sequence* is then a sequence $\sigma = U_1 \ldots U_n$ of non-empty steps U_i such that $m_0 [U_1\rangle m_1 \ldots m_{n-1} [U_n\rangle m_n$, for some markings $m_1, \ldots, m_n$ of N. Then m_n is said to be a reachable marking of N (under the maximally concurrent step semantics).

To conclude this preliminary section, we elaborate on the choice of this particular net model. First, it should be observed that it follows from the above definitions that the semantics allow *auto-concurrency*, the phenomenon that a transition may be executed concurrently with itself. This approach makes it possible to use transitions for a faithful modelling of natural events like the independent (non-sequential) occurrence in vast numbers of a biochemical reaction in a living cell. Note that the degree of auto-concurrency of a transition can easily be controlled by a dedicated place with a fixed, say k, number of tokens connected by a self-loop with that transition implying that never more than k copies of that transition can fire simultaneously.

Activator arcs were introduced in (Janicki and Koutny, 1995) as a means of testing for the presence of at least one token in a place, and so they are similar to other kinds of net features designed for the same reason. We mentioned already self-loops by which the presence of a token in a place can be tested only by a single transition (which 'takes and returns' the token) and not simultaneously by an arbitrary number of transition occurrences in a step. Two other mechanisms related to activator arcs, which do allow such multiple testing are *context arcs* (Montanari and Rossi, 1995) and *read* (or *test*) *arcs* (Vogler, 2002). Both, however, display important differences when compared with activator arcs. A context arc testing for the presence of a token in place p by transition t indicates that after a step in which t participates has been executed, p must still contain a token which precludes the occurrence in the same step of transitions that have p as an input place. A read arc is also different, but less demanding in that there must exist a way to execute sequentially (*i.e.*, one-by-one) all transition occurrences in the step, without violating the read arc specification. In both cases, one can easily see that activator arcs are most permissive since they only check for the presence of a token before the step is executed (this is often referred to as *a priori* testing). We feel that *a priori* testing is more appropriate for biological applications as the 'look ahead' implied by the other two kinds of test arcs is hard to imagine in reality.

Finally, we rely in this chapter on *maximal concurrency* in the steps that are executed which reflects the idea that execution of transitions is never delayed. This may also be viewed as a version of time-dependent Petri nets where all transitions have a

firing duration of 1. However, the maximal concurrency we apply here does not derive from Petri nets with time, but rather from Petri nets with *localities* (Kleijn *et al.*, 2006) leading to *locally maximal* semantics. This semantics is what we plan to use to model other aspects of the development as well. Here one may think of *e.g.* the locally synchronous occurrence (in pulses) of reactions in individual compartments of a cell.

5.3 BIOLOGICAL BACKGROUND AND MODELLING DECISIONS

5.3.1 Mechanisms of biological gradient formation

In biology, the term gradient is used to describe a gradual and directed change in concentration of a morphogen through a group of cells, *e.g.*, a tissue. Morphogens are signalling molecules that cause cells in different places in the body to adopt different fates and thereby help establish embryonic axes. Morphogens are produced in a localized source of a tissue, the source cell(s), and emanate from this region, forming a concentration gradient (Gurdon and Bourillot, 2001; Teleman *et al.*, 2001). A morphogen gradient has an immediate effect on the differentiation of the cells along it; cells are able to 'read' their position along the gradient and determine their developmental fate accordingly. They have a range of possible responses and the morphogen concentration dictates which response will be exhibited (Gurdon and Bourillot, 2001; Teleman *et al.*, 2001).

   The mechanisms by which the morphogen travels through a cell layer have been the topic of some debate and are not yet fully understood. Three mechanisms have been described, shown schematically in Figure 5.1: (A) diffusion through the extracellular matrix (Fischer *et al., 2006; Gurdon et al.*, 1994; Lander *et al.*, 2002), either passively, like a drop of ink in water (Gurdon *et al.*, 1994), or facilitated by receptors on the cell surface which guide the morphogens along (Fischer *et al.*, 2006), as shown in the figure; (B) sequential internalization of the morphogen molecules in vesicles in the cells, a process called endocytosis, and subsequent re-emission (Entchev and Gonzalez-Gaitan, 2002; Fischer *et al.*, 2006; Teleman *et al.*, 2001); (C) direct contact between the cells by means of tentacle-like threads of cytoplasm, called cytonemes, connecting the cells (Gurdon and Bourillot, 2001). These mechanisms are not necessarily mutually exclusive and some studies conclude that a combination of mechanisms underlies the formation of a gradient, (*cf.* Kicheva *et al.* 2007, on which the case study in 6.4 is based). It is important to note that both diffusion and endocytosis take place between neighbouring cells, while cytonemes connect all cells directly to the source. This makes it very different from a modelling perspective, as will be discussed below. In this and the next chapter we will solely be concerned with communication between neighbouring cells, *i.e.* through diffusion and endocytosis.
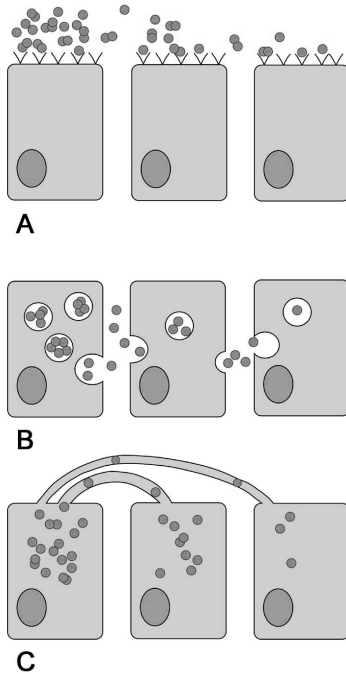
Figure 5.1. Three possible mechanisms for gradient formation: diffusion (A), endocytosis subsequent re-emission (B) and transport through cytonemes (C).

Unfortunately, knowledge of the exact concentrations and shapes of most gradients is often limited. This is mainly due to the transient nature of morphogen gradients and the low concentrations at which they are effective, both of which make it difficult to visualise the morphogens (Gurdon *et al.*, 1994). Many morphogens are rapidly degraded or prevented from binding to receptors by antagonistic proteins (Gurdon *et al.*, 1994). Much of the information on gradients is therefore obtained indirectly, by observing their effect, *i.e.*, the responses of the cells involved (Gurdon *et al.*, 1994). A qualitative approach, such as the one presented in this chapter, circumvents this issue by not relying on exact quantitative data. However, in cases for which quantitative data is available, incorporating this in the model will yield a more detailed and practical model of process under study. Such a quantitative model is presented in chapter 6, along with a case study based on experimental observations.

5.3.2 Modelling decisions

We have chosen **cells as the elementary units** in our model, to be represented by places in the Petri net. Earlier studies (Bonzanni *et al.*, 2009; Krepska *et al.*, 2008; Matsuno *et al.*, 2003) have successfully modelled cell-to-cell signalling, starting from a lower biological level, using places to represent genes and proteins. Although this allows a

high level of detail, it also complicates the net and makes it difficult to identify single cells. In our approach the cellular level represents the intermediate level between the subcellular levels, on which the morphogen signalling between cells takes place, and the tissue/organ level, where whole cell layers may move.

Furthermore, the process lends itself to be modelled using a **modular approach**; for each of the neighbouring biological cell pairs identical modules of places and transitions are used. This makes it easy to change or extend the model or to adjust parameter values according to different experimental data.

We let **tokens represent morphogen levels**, conducted from cells to neighbouring cells by the transitions. By changing the interpretation of the tokens a range of levels is possible, from exact quantitative modelling, in which each token corresponds to a precise morphogen number, to strictly qualitative modelling (Kleijn *et al.*, 2006), in which markings become binary, indicating merely the presence or absence of morphogens in a cell. In an intermediate, semi-qualitative approach, increasing token numbers equal increasing morphogen levels, without exact molecular numbers. Petri nets allow modelling at all these different levels. Biological gradients often work in a rather discrete, semi-qualitative manner; a number of cell responses (such as activation of a particular gene) exists for a given gradient and threshold values in morphogen concentration demarcate the boundaries between these responses, resulting in a stepwise change in cellular behaviour throughout the tissue. Due to this, both semi-qualitative and quantitative ways of modelling can represent biological situations realistically; our Petri net model is applicable to both. Moreover, it is possible to model the formation of a gradient in a quantitative manner, but let other processes which depend on the morphogen levels do so in a semi-qualitative way, by using threshold values. Since we do not use experimental quantitative data in the current chapter, the model can be interpreted as semi-qualitative; in the next chapter an example will be presented of an entirely quantitative approach, in which tokens numbers directly corresponds to numbers of morphogen molecules.

Instead of merely calculating the final distribution of the tokens, we want our net to model the gradual process of morphogen movement through the tissue, *i.e.* to **represent all intermediate steps**. This will allow the user to simulate experiments in which the process is altered while running; *e.g.* grafting experiments, in which parts of the tissue get removed or replaced, can be simulated by taken cells out of the net at a certain moment during the process or depleting them of tokens.

Our model focuses on **local signalling between neighbouring cells**. Therefore we take into account cell-to-cell communication mechanisms, *e.g.* endocytosis and diffusion, but not long distance transport mechanisms, *e.g.* through cytonemes. In the situation of local signalling, the number of morphogens to be transported from one cell

to the next depends solely on the difference in morphogen level between these two neighbouring cells; cells have no 'knowledge' of morphogen transport in other parts of the tissue. In order to accurately reflect this situation we base the computation of transported tokens solely on the difference in token numbers between the neighbouring cells. This makes the model easily scalable, *i.e.* the number of cells in the tissue is irrelevant to the computation and can be adjusted without altering the workings of the model.

5.3.3 Implementation

Often exact quantitative data for the processes of morphogen transport and degradation between neighbouring cells are not known, and these may vary depending on the gradient considered. Therefore we do not discern the molecular mechanisms of diffusion, endocytosis and degradation of morphogens in this model but we introduce a parameter ρ in our model to represent the effective **ratio of concentration levels** between neighbouring cells and to determine the amount of tokens to be transported between places during the simulation of gradient formation. In other words, ρ represents the final ratio of morphogens between neighbouring cells and morphogen degradation is implicit. In the next chapter we present a model in which production, transport and degradation are modelled explicitly.

In the organism, gradient ratios arise passively as a consequence of physical laws. However, to accurately reflect the biological process of gradient formation underlying the spread of morphogens from cell to cell, our formal model has to compute the number of tokens passed on based on the ratio ρ. Hence, the model includes an **explicit separate computational unit** for each pair of neighbouring cells, to perform the necessary calculations. In particular, these parts of the net control the transport of tokens between places. In this way a close relation to the biological process can be maintained in one part of the net, with the underlying computations being performed in the background by another part of the net. At all times, the marking of the places representing biological cells will be consistent with biological observations of (the effect of) the gradient, *i.e.*, the ratio is maintained and places corresponding to cells further away from the source will never have more tokens than places (cells) closer to it.

Another important feature of the model is the **use of concurrent steps** rather than individually occurring transitions. Morphogen transport between cells is not directly influenced by events taking place in non-adjacent cells, which means these processes should be able to take place concurrently and non-adjacent cells can be simultaneously involved in the transport of morphogens. This leads to an execution mode consisting of *concurrent steps*. Moreover, since in the biological situation

morphogens move to the next cell as soon as this is possible, we have chosen to use a **maximally concurrent steps**.

5.4 GRADIENTS AND PETRI NETS

5.4.1 Modelling solution

Following the ideas outlined in the previous section, we will propose a formal model for the formation of a gradient. Our assumptions regarding the biological process of gradient formation are as follows. Given is a segment of k adjacent cells with the i-th cell immediate neighbour of the $(i+1)$-th cell. Morphogens can be transported only between immediate neighbours. Morphogens move from cells with higher concentration to neighbours with lower concentration, as long as the concentration ratio between these cells does not exceed a given gradient ratio $0 < \rho < 1$. We assume that ρ is a rational number, *i.e.*, $\rho = N/M$, where $M > N \geq 1$. Initially, the first cell x_1 (the source) contains a quantity (*i.e.* has a concentration level of) K of a morphogen. These assumptions lead to the following modelling problem.

Given are $k \geq 1$ places $x_1, \dots , x_k$, representing a segment of k cells with place x_i corresponding to the i-th cell. In the initial marking m_0, the first place x_1 contains K tokens and there are no tokens in the other places. In the net modelling the mechanism of gradient formation, we need to shift tokens from x_1 in the direction of the last place x_k. Places and/or transitions may be added, but in such a way that for any reachable marking m the following hold.

1. The number of tokens in the x_i's remains constant, *i.e.*,

$$m(x_1) + \dots + m(x_k) = K \qquad \textbf{token preservation}$$

2. The tokens are distributed monotonically along the sequence of k places, *i.e.*,

$$m(x_1) \geq \dots \geq m(x_k) \qquad \textbf{monotonicity}$$

3. The ratio of the numbers of tokens in two neighbouring places does not exceed ρ, *i.e.*, for every $1 \leq i < k$ with $m(x_i) \geq 1$:

$$(m(x_{i+1})/m(x_i)) \leq \rho \qquad \textbf{ratio}$$

4. Shifting continues until moving even one token would violate the above, *i.e.*, if no tokens are shifted after marking m was reached, then for every $1 \leq i < k$ with $m(x_i) > 1$:

$$(m(x_{i+1})+1)/(m(x_i)-1) > \rho \qquad \textbf{termination}$$

Moreover, the relative position of a place within the sequence plays no role. In particular, the mechanism should be easily scalable and insensitive to the specific values of k and K.

If we look at the above formulation of properties (2) and (3) — monotonicity and preservation of the gradient ratio — and recall that $\rho = N/M$ and $M > N$, it is easy to observe that these two properties are together equivalent to stating that, for every $1 \le i < k$, $N \cdot m(x_i) - M \cdot m(x_{i+1}) \ge 0$. We will call a marking m satisfying this inequality consistent and denote $\alpha_i = N \cdot m(x_i) - M \cdot m(x_{i+1})$, for every $1 \le i < k$. Note that the initial marking is consistent.

Similarly, if we look at the above formulation of properties (2) and (4) — monotonicity and termination — it is easy to observe that together they are equivalent to the statement that, for every $1 \le i < k$, $N \cdot m(x_i) - M \cdot m(x_{i+1}) < M + N$. We will call a consistent marking m satisfying this inequality stable. Note that for a given ρ, k and K, there may be more than one stable marking. For example, if $\rho = 1/2$, $k = 5$ and $K = 111$, then the following are two different stable markings:

x_1	x_2	x_3	x_4	x_5		x_1	x_2	x_3	x_4	x_5
59	29	14	6	3		58	29	14	7	3

We are now ready to propose a generic solution for the above problem. For a given consistent marking m and each $1 \le i < k$, move β_i tokens from x_i to x_{i+1} where $\beta_i \le \left\lfloor \dfrac{\alpha_i}{M + N} \right\rfloor$, and at least one β_i must be non-zero if at least one of the values $\left\lfloor \dfrac{\alpha_i}{M + N} \right\rfloor$ is non-zero. We denote the resulting marking by $m_{\beta_1 \ldots \beta_{k-1}}$.

An intuitive reason for proposing such a mechanism for shifting tokens is that the number of tokens in x_i that are 'balanced' by tokens in x_{i+1} is $(M/N) \cdot m(x_{i+1})$, because each token in x_{i+1} is equivalent to M/N tokens in x_i. Hence there are $m(x_i) - (M/N) \cdot m(x_{i+1})$ unbalanced tokens in x_i. The 'portion' of each unbalanced token that could be safely transferred to x_{i+1} is $N/(M+N)$. Hence in total we may safely transfer $\left\lfloor \dfrac{N}{M + N} \cdot (m(x_i) - \dfrac{M}{N} \cdot m(x_{i+1})) \right\rfloor$ tokens, which is precisely $\left\lfloor \dfrac{\alpha_i}{M + N} \right\rfloor$ tokens.

Clearly, some of the numbers $\beta_1, \ldots, \beta_{k-1}$ can be zero, and by the condition above, all β_i's are zeros if and only if the marking is stable:

Proposition 1. $\beta_1 = \cdots = \beta_{k-1} = 0$ *if and only if m is stable*.

Crucially, by the mechanism proposed consistent markings are always transformed into consistent markings.

Proposition 2. *If m is a consistent marking then $m_{\beta_1 \ldots \beta_{k-1}}$ is also consistent*.

According to the above, any number of tokens not exceeding $\left\lfloor \dfrac{\alpha_i}{M + N} \right\rfloor$ can be moved *simultaneously* from x_i to x_{i+1} (for every $i < k$), and consistency will be

preserved. Clearly, the new consistent marking is different from the previous one if and only if, for at least one i, we have $\beta_i \geq 1$. The idea now is to keep changing the marking on $x_1, \ldots, x_k$ until a marking m has been reached such that $\left\lfloor \dfrac{\alpha_i}{M+N} \right\rfloor = 0$, for all $1 \leq i <$ k, which is equivalent to $\alpha_i < M + N$, for all $1 \leq i < k$. In other words, this m is a stable marking. Since tokens cannot be shifted forever, this procedure will always terminate in a stable marking (formally, we can show this by considering a weighted distance to the end of the chain of the K tokens; it never increases and always decreases in a non-stable state).

Looking now from the point of view of a Petri net implementation of the proposed mechanism, what we are after is a net N_{shift} comprising the places $x_1, \ldots, x_k$ and such that if m is a marking of N_{shift} whose projection on these k places is consistent, then a step U can occur at m if

– it moves at most $\left\lfloor \dfrac{\alpha_i}{M+N} \right\rfloor$ tokens from x_i to x_{i+1}, for all $1 \leq i < k$;

– at least one token is moved from x_i to x_{i+1} for at least one $1 \leq i < k$, unless the projection of m onto $x_1, \ldots, x_k$ is stable.

In fact, in the proposed implementation, we will be preceding the 'token-shifting' with a 'pre-processing' stage which seems to be unavoidable unless one uses some kind of arcs with complex weights depending on the current net marking.

5.4.2 Implementation

In the implementation of the proposed shifting mechanism, as many tokens as possible should be shifted from one neighbour to the next. That means that, at each stage we have $\beta_i = \left\lfloor \dfrac{\alpha_i}{M+N} \right\rfloor$ for every $1 \leq i < k$. Moreover, tokens are shifted from a place without any assumptions whether new tokens will come to that place from its other neighbour. Thus we need to provide a Petri net structure capable of 'calculating' the value of expressions like $\left\lfloor \dfrac{N \cdot m(x_i) - M \cdot m(x_{i+1})}{M+N} \right\rfloor$.

Our proposed gradient forming mechanism distinguishes three phases: I, II and III. An auxiliary net N_{3phase}, shown in Figure 5.2B, is used to schedule the transitions implementing the calculations. It controls these transitions via the places w^{I} and w^{II} and activator arcs. For the full picture of the system one should combine the figures for all

pairs (x_i, x_{i+1}) with a single copy of the net in Figure 5.2B. Note that all places with identical label (in particular w^I, w^{II}, and w^{III}) should be identified. That other parts of the encompassing net model do not interfere with the calculations carried out during phases I and II can be ensured by connecting the relevant transitions with the place w^{III} using activator arcs.

For every $1 \leq i < k$, transition t_i is intended to shift tokens from x_i to x_{i+1} (phase III). To achieve this, we use two disjoint sets of new, auxiliary places, $x'_1, \ldots, x'_k$ and $x''_1, \ldots, x''_k$. These places are initially empty. The idea is to fill x'_i with $N \cdot m(x_i)$ tokens and x''_{i+1} with $M \cdot m(x_{i+1})$ tokens (phase I). The latter are used for the removal of $M \cdot m(x_{i+1})$ tokens from x'_i (phase II). After this, there are α_i tokens remaining in x'_i. Finally, for each group of $N + M$ tokens in x'_i, one token is shifted from x_i to x_{i+1}. The construction (for x_i and x_{i+1}) is shown in Figure 5.2A.

The overall mechanism operates in cycles of three consecutive, maximally concurrent steps such that for every $1 \leq i < k$:

I. Transition c_i, inserts (in $m(x_i)$ auto-concurrent occurrences) $N \cdot m(x_i)$ tokens into x'_i. In the same step, transition c_{i+1}, inserts (in $m(x_i)$ auto-concurrent occurrences) $M \cdot m(x_{i+1})$ tokens into x''_{i+1}. Simultaneously, transitions e'_i and e''_{i+1} empty x'_i and x''_{i+1} of any residual tokens left from the previous cycle.

II. Next, transition d_i (in $M \cdot m(x_{i+1})$ auto-concurrent occurrences) empties x''_{i+1} and leaves in x'_i the difference $\alpha_i = N \cdot m(x_i) - M \cdot m(x_{i+1})$.
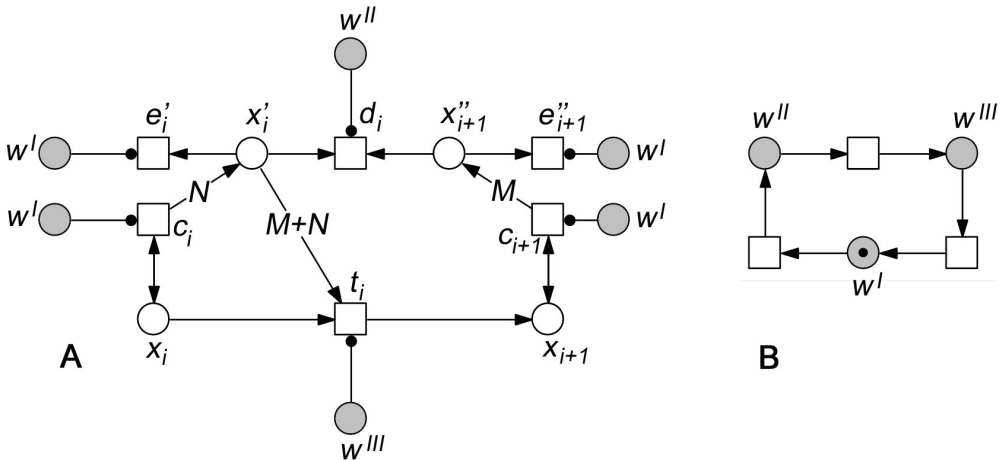

Figure 5.2. (A) The main part of the construction for the solution (note that e''_{i+1} is introduced for later use when one might want to remove or add tokens to the x_i's from 'outside'; in the standard (consistent) situation it is never activated as after phase II, place x''_{i+1} is empty); and (B) the subnet N_{3phase} enforcing the three phases.

III.   In the third step, the occurrences of transition t_i transfer $\beta_i = \left\lfloor \dfrac{\alpha_i}{M+N} \right\rfloor$ tokens

from x_i to x_{i+1}.

Proposition 3. Each cycle results in transferring β_i tokens from x_i to x_{i+1}.

Note that in this implementation with the control net N_{3phase}, neighbouring pairs are either all involved in calculations (step I and II of the cycle) or tokens are transferred between neighbours (step III). During the whole operation of the adjustment process (except for the transfer phase), the token numbers in the places x_i, representing the cells, are unchanged and they can be accessed for reading by other transitions (and thus influence neighbouring cells). In other words, calculations are *orthogonal* to the basic operation of the net (the gradient formation). As an example, let us consider the case when $\rho = 1/2$, $k = 4$ and $K = 100$. Then executing the constructed net in a maximally concurrent manner leads to the following sequence of markings on the x_i after each cycle and eventually to a stable marking:

x_1	100	67	67	60	60	57	57	56	56	55	55	54
x_2	0	33	22	29	25	28	26	27	26	27	26	27
x_3	0	0	11	8	12	10	12	12	13	12	13	13
x_4	0	0	0	3	3	5	5	5	5	6	6	6

The next example shows what happens if we start from a (non-initial) consistent marking (again $\rho = 1/2$):

x_1	200	167	156	...
x_2	50	67	67	...
x_3	0	16	22	...
x_4	0	0	5	...

The construction works without any problems, if we start with a consistent marking. In case $0 > \alpha_i$ for some i, then transition t_i is not executed, but the transitions t_{i-1} and t_{i+1} may still be executed and lead to an adjustment of the marking causing t_i to become active in the next cycle. A further observation is that adding (or removing) tokens at some point, will trigger a re-adjustment process which tries to re-establish the correct ratios between the markings of adjacent places x_i. This process is unpredictable, but to deal with that case we have included transition e''_{i+1} which in the standard (consistent) situation is never activated since then, after phase 2, place x''_{i+1} is empty.

An important characteristics of the proposed solution is that it is purely local and does not assume anything about the number of tokens which may appear in the x_i's nor the length of the chain. In other words, it is truly generic. What's more it also works

if *M* and *N* are different for different pairs of neighbouring places, *i.e.*, if rather than a uniform gradient ratio ρ there is a ratio ρ_i for each pair of neighbours x_i and x_{i+1}.

Another feature of our solution is the maximal concurrency semantics intended to reflect the idea of morphogens (simultaneously) moving from cell to neighbouring cell whenever that is possible. The preliminary sequential semantics model we developed (but not reproduced here) is more complicated as it also needs *inhibitor arcs* which test for absence of tokens (to decide whether or not tokens should still be shifted). Moreover, one needs to decide that x_i either receives or sends tokens at each stage. In a step model it can both receive and send. Also, with the maximal concurrency semantics, the number of states of the model is dramatically reduced. The auxiliary net N_{3phase} is used to partly sequentialise the behaviour in order to separate the pre-processing phases from the actual shifting phase. This net could also have been made local to the main construction of the net in Figure 5.2A, with different copies of it assigned to different localities. This would have given the additional possibility of controlling the degree of synchronisation between different parts of the gradient model by using a locally maximal step semantics.

Finally, we would like to point out that the activator arcs in our implementation are used only to control the calculation and can actually be avoided in case there would be a limit on the number of tokens in each place x_i at any time. (Then the activator arcs can be eliminated basically by having separate copies of N_{3phase} for each $1 \le i < k$, transfer around sufficiently many tokens in a bundle, and replace activator arcs by self-loops). This assumption corresponds to having (or knowing) some capacity bound on the concentration levels of morphogens in a cell and so may be biologically sound.

5.5 CONCLUSION

Starting from gradient formation in the AP axis development in the model organism *Xenopus laevis*, we have presented a novel approach to using Petri nets in developmental biology by focusing on the cellular rather than subcellular levels and abstracting from concrete proteins and genes. This has led to a parameterised Petri net model for the general process of gradient formation through diffusion and endocytosis.

Assumptions regarding gradient formation have been formulated based on essential features of this process as reported in the literature. These assumptions underlie the precise requirements given that should be satisfied by an abstract Petri net model of gradient formation. A crucial point here is the consistency that is maintained during the execution of the model. Hence the realization of the gradient is faithfully reflected. Moreover the close relationship between biological process and evolution of the formal model makes it possible to apply existing Petri net techniques to analyse

what happens during gradient formation. In particular cause-effect relations should be properly reflected in the process semantics of the modelling Petri net (Kleijn and Koutny, 2007; Kleijn and Koutny, 2008).

Another main contribution of this approach is its generic nature, leading to a model that is scalable and applicable to a plethora of specific gradients. Also scalability is a consequence of the faithful reflection of the biological process. Since the final token (morphogen) distribution is not directly computed from the initial amount of morphogen and the length of the chain of cells, but rather simulates the communication between neighbouring cells, the length plays no role in the occurrence of the steps. The model as presented here represents a one-morphogen system without relying on quantitative data, but exact values could be assigned to ratio and individual tokens. Moreover, it provides a basis for simulation of simultaneous gradient formation (different morphogens with different experimental initial markings) and for inhibiting/activating interactions between them. Simulation with actual biological data to validate the model should be a next step; the elaborated model presented in chapter 6 illustrates this. In addition, we will focus our attention on the extension of this still rather basic model to more dimensions, *e.g.*, rather than having just a single line of cells, we consider the spread of morphogens from a source throughout a tissue plane or volume.

In (Bonzanni *et al.*, 2009; Krepska *et al.*, 2008), Petri nets are used to model developmental processes in a way similar to our approach when it comes to the semi-qualitative use of tokens and the use of maximal concurrency. In these papers however, the focus is on subcellular levels. Petri net places are used to represent genes and gene products, where in our approach cells, as basic units in a tissue, are modelled by places. Having cells as basic units should prove to be a useful intermediate position convenient for 'zooming in and out' between subcellular and tissue level. It is our aim to model more subprocesses of the AP axis formation. For instance the different molecular processes underlying diffusion and endocytosis could be modelled in subnets, allowing the user to compare the different effects of these mechanisms. Also the degradation of morphogens could be modelled by a subnet, making the entire process more explicit. In the next chapter a step in that direction is taken, by modelling production, transport and degradation explicitly in the net, by means of parameters derived from an equation based model; a further future development would be to faithfully model the events underlying these subprocesses and thereby have the parameter values arise out of the workings of the subnets, as opposed to providing them to the main as as external information.

CHAPTER 6.   COMBINING ASPECTS FROM ALGORITHMIC AND EQUATION BASED MODELLING: COMPLEMENTING DIFFERENTIAL EQUATION MODELS OF BIOLOGICAL GRADIENTS WITH PETRI NETS

πάντα ῥεῖ

Everything flows.

- Heraclitus

The final chapter again concerns the formation of biological gradients, modelled using Petri nets. However, a new approach was developed, aimed at quantitative modelling of the process, rather than qualitative. In the current study a merger between aspects of Petri nets and differential equations was strived after, in order to extend the possibilities for simulation and analysis of the net. This has particular merits for computational biology, since quantitative analysis can be performed while still retaining an intuitive understanding of the underlying mechanics of the process.

Manuscript, to be submitted as: Bertens L.M.F., Kleijn J., Hille S., Heiner M., Koutny M., Verbeek F.J., 'Combining differential equation models and Petri net modelling for biological gradient formation'.

6.1 INTRODUCTION

Many biological processes, in particular biochemical processes such as metabolic and signal transduction pathways, have been modelled using differential equations (DEs) (Kicheva *et al.*, 2007; Gregor *et al.*, 2005; Yu *et al.*, 2009; Ellner and Guckenheimer, 2006). These mathematical models accurately describe changes in process variables and thereby enable precise quantitative studies, parameter sensitivity analysis of dynamics and bifurcation analysis. They assume the unfolding of processes in continuous time and even continuous space. This allows the deduction of properties of the system mathematically, *e.g.* the existence and stability of steady states, by analyzing the system of DEs. Analysis is complemented by numerical simulations, to investigate the transient behaviour, when the system is moving towards its long-term behaviour. The simulation techniques involve discretisation of the DEs, in time and space. The resulting computational scheme describes the change of state variables in discrete time steps.

In contrast, algorithmic process models aim to describe the mechanisms underlying the changes in the system (Priami, 2009; Ellner and Guckenheimer, 2006). Petri nets, introduced in the previous chapter, exemplify such a modelling framework (Reisig and Rozenberg, 1998; Petri, 1962; Reisig, 2010). This is of particular use to biological studies, because of its origin in the modelling of chemical reactions and molecular interactions, and its ability to model concurrent behaviour, *i.e.* the potentially simultaneous occurrence of events (*cf.* section 5.2), which is a common feature of biological systems (Fischer *et al.*, 2011; Koch *et al.*, 2011). Moreover, Petri nets combine graphical and mathematical elements, making them intuitive to communicate, execute and understand visually, as we saw for the model of gradient formation in the chapter 5. Implementation of a Petri net yields an operational process model. Using analysis tools such as state space exploration and analysis of *e.g.* deadlocks and boundedness properties, the behaviour of processes can be studied. In this way, they provide a view point complementary to DE models.

As for DE models, an interest in modelling biological processes with Petri nets has emerged, especially in the field of systems biology, and new ways to apply this modelling technique to the life sciences are constantly being developed (Banks, 2009; Gilbert and Heiner, 2006; Gilbert *et al.*, 2007; Chaouiya, 2007; Heiner *et al.*, 2008; Krepska *et al.*, 2008; Matsuno *et al.*, 2003; Steggles *et al.*, 2006; Koch *et al.*, 2011). Petri nets have various advantages for modelling processes in the life sciences. They can be used to model various levels of abstraction in a biological process, ranging from subcellular reactions to processes taking place on tissue or organ level. Key elements, *e.g.* passive resources, active events and concurrency, are found on all these levels, making Petri nets a suitable modelling system. Changing focus from lower to higher levels of abstraction in modelling a biological phenomenon is therefore akin to zooming in and out between levels of magnification. Although, similarly, multi-scale modelling is

feasible with DE-models, as *e.g.* employed in advanced models for weather forecasting, it is more involved to implement, with various numerical mathematical problems to overcome. Additionally, Petri nets can be constructed in a modular fashion, using identical building blocks, for instance to construct groups of similar tissue elements or cells. This makes them easily scalable. Specific elements in Petri net theory, such as Coloured nets, allow one to model several processes taking place simultaneously at the same location.

<u>Aim of the study</u>

In this chapter we present our work on combining Petri nets with differential equation models, in order to construct a generic Petri net model for the formation of molecular gradients. In chapter 5 a model has been presented, where the process of gradient formation is modelled as a global decrease in concentration levels of molecules throughout the cells in the tissue, and in which the spread of molecules is governed by a fixed ratio ρ of molecular concentration between neighbouring cells (*cf.* section 5.3.3). This ratio represents the combined effect of molecule transport and degradation in the cells.

In the current chapter we present an elaboration of this earlier model to include parameters of DE modelling. In this way, we continue the investigation of the use of Petri nets for the field of higher level developmental biology. The model presented in chapter 5 can be seen as a proof of concept, in which the modelling method was tested out in an abstract manner, allowing us to thoroughly check the correctness of the underlying assumptions (in particular token preservation, monotonicity, ratio and termination, *cf.* 5.4.1). The model presented in the current chapter builds on this proof of concept and moves from an abstract approach, towards a more detailed and applied approach. Here, the events of molecule production, diffusion and degradation have been modelled explicitly and are governed by individual parameters.

For the process of gradient formation, DE models exist which provide accurate quantitative data about the gradient formation (Kicheva *et al.* 2007; Gregor *et al.*, 2005; Yu *et al.*, 2009). By linking the parameters of the Petri net model to the parameters in the discretised form of such a DE model, the net can be used to produce quantitative data about discrete space and time points in the process, similar to the DE model, while at the same time retaining the advantages of the Petri net framework. In order to validate the Petri net model, we present a case study; from literature we have selected a study in which experimental observations of gradient formation have been modeled using differential equations. We use the parameters from this DE model and show how the execution data from the resulting Petri net correspond to the data obtained from the differential equations.

Since this study forms a continuation of the study presented in chapter 5 the reader is referred to sections 5.2. and 5.3 for Petri net terminology and definitions, the biological background of gradient formation and the modelling decisions. In the remainder of this chapter the discretisation of the DE model is set out, along with the connection of DE parameters to parameters in the Petri net modelling solution, in section 6.2. In 6.3 we present the resulting Petri net model. A case study of gradient formation of the protein Dpp in the fruit fly is used for the validation and presented in section 6.4; finally, conclusions and remarks on future work can be found in section 6.5.

6.2 DERIVATION OF PETRI NET MODEL PARAMETERS FROM A DISCRETISED DE MODEL

In this section the temporally and spatially continuous situation, modelled by a DE model, is translated to a discrete situation, which is subsequently linked to the Petri net solution, described in section 6.3. We consider the following reaction-diffusion equation

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 c}{\partial r^2} - kC , \tag{1}$$

on the one-dimensional interval (0, L). This is used in (Kicheva $et\ al.$, 2007) as the effective equation to describe their data. It consists of the measurement of fluorescence of GFP-labelled morphogens when these form a gradient in a rectangular sample of cell tissue. The morphogens are homogeneously emanating from a source, which is a strip of cells at the left border ($r = 0$) of this rectangle. The morphogens move from the source to the right, $i.e.$ towards $r = L$. The fluorescence measurements are made in multiple vertical layers in the tissue and summed, reducing the situation to two dimensions. The morphogen concentration can be assumed constant in the direction transversal to r, further reducing the situation to one dimension. Thus, $C(r,t)$ represents the areal density of observed morphogen at location r at time t (in molecules/μm^2). D is the effective diffusion coefficient (μm^2/s), combining passive diffusion and possible other transport processes such as endocytosis and active diffusion ($cf.$ section 5.3.1), and k is the degradation rate (s^{-1}). Equation (1) is complemented with an initial condition $C(r,0) = f(r)$ and zero-flux boundary conditions at L and constant influx areal density J_0 through the left side of the sample at $r = 0$, $i.e.$ $D \frac{\partial}{\partial r} C(0) + J_0 = 0$.

The standard procedure of spatial discretisation at equidistant points $0 = r_0 < r_1 < \cdots < r_n = L$, with $l = r_{i+1} - r_i$ and $C_i(t) := C(r_i,t)$, followed by temporal discretisation at time points t_j, in which j represents the number of steps and the steps are equally separated at time intervals Δt (corresponding to a fixed number of n' steps); this yields

$$\frac{\Delta C_i(t_i)}{\Delta t} \approx \frac{D}{l^2}(C_{i-1}(t_j) - 2C_i(t_j) + C_{i+1}(t_j)) - kC_i(t_j) \tag{2}$$

for $i = 1, \dots, n - 1$, where $\Delta C_i(t_j) := C_i(t_{j+1}) - C_i(t_j)$. We take l equal to the cell length and h to the cell height. Multiplying both sides of (2) with the cell area $A = l\,h$ in the plane of observation yields a similar, slightly rewritten expression for the change in the number $m_i = m_i(t_j)$ of molecules in cell i at

time t_j (omitting time dependence):

$$\Delta m_i \approx \frac{D\Delta t}{l^2}(m_{i-1} - m_i) - \frac{D\Delta t}{l^2}(m_i - m_{i+1}) - km_i\Delta t \tag{3}$$

for $i = 1, \dots, n - 1$, with $\Delta m_i = m_i(t_{j+1}) - m_i(t_j)$. Approximation (3) is appropriate when Δt and l are such that $D\Delta t/l^2 < 1$ and $k\Delta t < 1$ are sufficiently small. Equation (3) is complemented by similar equations at $i = 0$ and $i = n$ that incorporate the boundary conditions:

$$\Delta m_0 = J_0 h\Delta t - \frac{D\Delta t}{l^2}(m_0 - m_1) - k\Delta t m_0 \tag{4}$$

$$\Delta m_n = \frac{D\Delta t}{l^2}(m_{n-1} - m_n) - k\Delta t m_n \tag{5}$$

In order to arrive at a Petri net model which incorporates these derivations, we use the abstract Petri net presented in chapter 5 and specify this to include the derived parameters. We start with a Petri net with a *max-enabled* semantics and places $x_1, \dots, x_n$ which represent biological cells (as described in chapter 5). We then extend this with specifications of the three main events in the process of gradient formation, using the given derivations:

I) the first term on the right hand side of (4) represents morphogen **production** in the source and transport to the first cell, x_1 (note that the token preservation from the earlier model no longer applies, *cf.* 5.4.1).

II) the **transport** between neighbouring cells x_i and x_{i+1} is given by the first two terms on the right hand side of (3).

III) lastly, the **degradation** in every x_i is given by the third term on the right hand side of (3).

In other words, the marking of the places x_i (for all places except x_1) after a multiple j of n' steps (in which n' represents one cycle of the modeled process, *cf.* 5.4.2) can be approximated well by the solution of the diffusion equation (1) at times $t_j = j\Delta t$:

$$m_i(jn') \approx h\int_{(i-1)l}^{il} C(r,t_j)dr \approx lh \cdot \frac{1}{2}\Big[C((i-1)l,t_j) + C(il,t_j)\Big] \qquad (6)$$

where we have used the trapezium rule to approximate the integral. In this way we aim to relate the number of molecules m_i to the marking of place x_i, *i.e.* $m(x_i)$. This brings us to the Petri net solution and its exact workings.

6.3 MODELLING SOLUTION

The previous section illustrated the derivation of parameters from the DE model, for the elements production, transport and degradation. In this section we present in detail our Petri net model and describe its workings. We propose a formal, general Petri net model for gradient formation. Given is a segment of n adjacent biological cells with the i-th cell the immediate neighbour of the $(i+1)$-th cell. This is represented in the Petri net by x_1, ..., x_n places. Morphogens are represented by tokens and can be transported only between immediate neighbours. Transitions t'_1, ..., t'_n represent the transport of tokens, in the direction x_1 to x_n. We will focus on one-directional gradient formation, strictly from x_1 to x_n. Figure 6.1A shows the basic structure of the net; here the first neighbouring cells on the left side of the modelled biological tissue are shown, x_1, x_2 and x_3.

In section 6.2 we discussed the derivation of parameters from differential equations for three basic elements in the process. Here we explain the way in which these parameters have been incorporated into the Petri net model.

I) Morphogen **production** and transport from the source to the adjacent cell x_1 are modelled by the transition s and comply to the first term of (4), which is directly translated to the weight arc of s to x_1. The weight is therefore $J_0 h\Delta t$.

II) With the exception of x_1 and x_n the morphogen **transport** from x_i to x_{i+1} follows the first term on the right hand side of (3), which corresponds to the effective diffusion from left to right. In order to incorporate this term into the weight arcs of the Petri net and therefore determine the number of tokens to be transported from x_1 to its neighbour x_{i+1}, additional auxiliary places are used for all $1 \le i \le n$: x'_1, ... , x'_n and x''_1, ..., x''_n. These places are initially empty. The places x'_i and x''_i are filled with $pD\Delta t \cdot m_{x_i}$ tokens and place x'_{i+1} and x''_{i+1} are filled with $pD\Delta t \cdot m_{x_{i+1}}$ tokens. After this first step in which
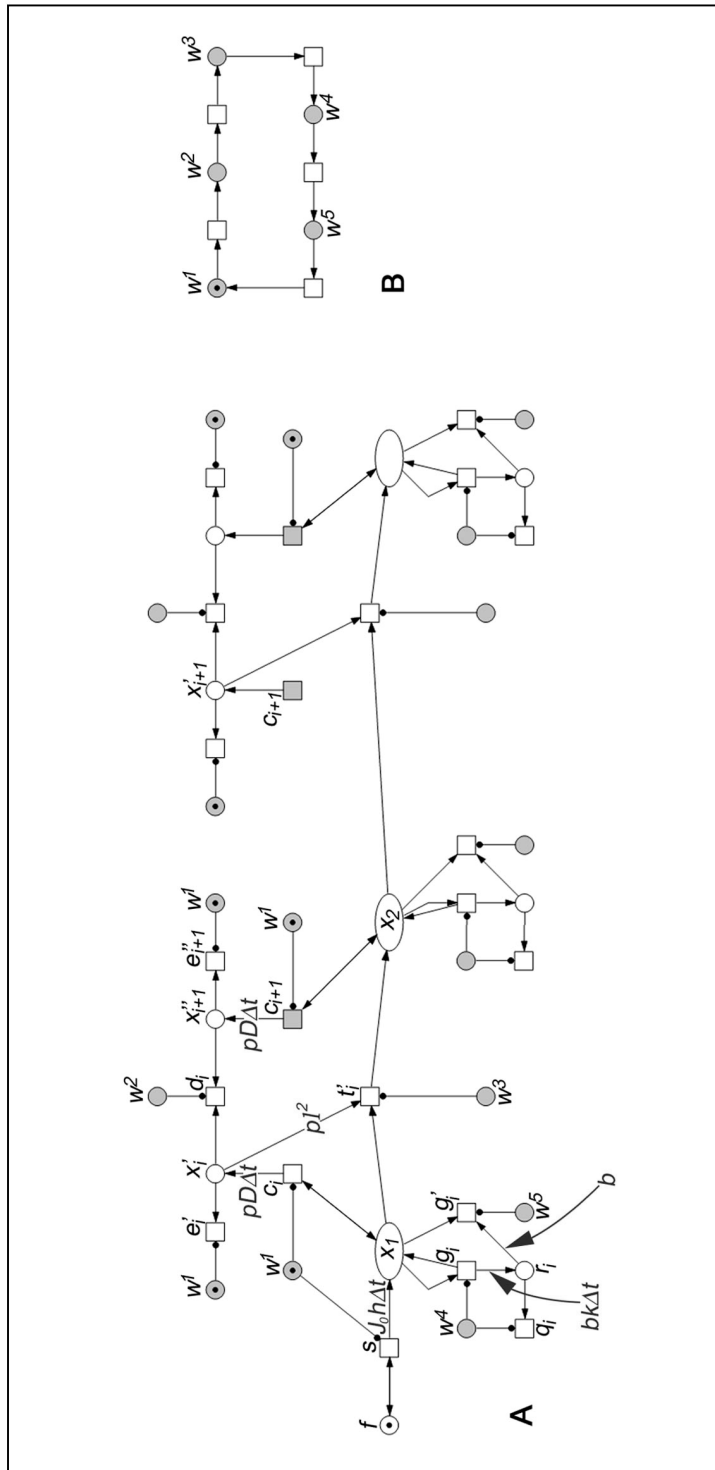
Figure 6.1. The main construction of the net, shown for three neighbouring cells (A) including production of morphogens (transition *s*) and influx of these in the first cell (x_1); the auxiliary net in (B) determines the order of events in the main net. Note that grey places are fusion places.

CHAPTER 6 – COMPLEMENTING DIFFERENTIAL EQUATION MODELS WITH PETRI NETS

places x'_i and x''_{i+1} are filled, these places get depleted simultaneously up to the point where x''_{i+1} is empty, leaving x'_i with a token difference of $pD\Delta t \cdot m_{x_i} - pD\Delta t \cdot m_{x_{i+1}}$.

The number of tokens to be transported from x_i to x_{i+1} is $\beta_i = \dfrac{pD\Delta t \cdot m_{x_i} - pD\Delta t \cdot m_{x_{i+1}}}{pl^2}$.

In other words, for every pl^2 tokens in place x'_i a token is moved by transition t'_i from x_i to x_{i+1}, respectively. Note that the multiplication by p in the previous step is countered here by a division in which the token number is rounded off; this multiplication constant is added to limit deviations caused by rounding off (*cf.* 6.4). Apart from multiplication with and division by p, the steps described here correspond directly to the derivation in (3) without the element of degradation (discussed below), as can be seen from the following:

$$m'_{x_i} = m_{x_i} - \beta_i + \beta_{i-1} = m_{x_i} - \frac{pD\Delta t \cdot m_{x_i} - pD\Delta t \cdot m_{x_{i+1}}}{pl^2} + \frac{pD\Delta t \cdot m_{x_{i-1}} - pD\Delta t \cdot m_{x_i}}{pl^2} \quad (7)$$

$$m'_{x_i} = m_{x_i} - \frac{D\Delta t}{l^2}(m_{x_i} - m_{x_{i+1}}) + \frac{D\Delta t}{l^2}(m_{x_{i-1}} - m_{x_i}) \quad (8)$$

Note that the monotonicity, described for the net in chapter 5, is maintained in the transport of this net. When no degradation takes place, monotonicity holds in the same way as described in 5.4.2.

III) Simultaneously with morphogen transport, morphogen **degradation** also takes place in the cells, which corresponds to the third term on the right hand side in (3). For every x_i, this process is modelled by the transitions g_i and g'_i and the place r_i, which is again an auxiliary place used to determine the number of tokens to be removed from x_i. The place r_i is filled with $bk\Delta t$ tokens (multiplication with b is used to prevent having to round off $k\Delta t$, since due to the small value of k for most biological gradients, this will often lead to 0) and subsequently for each amount of b tokens in r_i a token from x_i is removed, *i.e.* degraded. This results in a degradation element $\dfrac{bk\Delta t}{b} = k\Delta t$, which corresponds with the third element on the right hand side in (3). When k is identical for all x_i, monotonicity is still ensured (as in the case study discussed in 6.4).

These processes of production, transport and degradation take place in a cycle of $n' = 5$ steps. Note that n' was used in 6.2 to obtain the number of steps by multiplying this with the number of cycles, j; thus, in the translation from DE to Petri net parameters, $\Delta t = n'j$. An auxiliary net, shown in Figure 6.1B, is used to regulate these phases and the corresponding transitions. This auxilliary net is similar to the one presented in chapter 5; there, degradation was not modelled in explicit steps (diffusion

and degradation were combined in one parameter) and the cycle was limited to $n' = 3$ steps.

As described in 5.4.2, the auxiliary net controls the transitions via n' places, in this case w^1 to w^5, and activator arcs. For the full picture of the system one should combine a single copy of the auxiliary net to the figures for all of the pairs x_i, x_{i+1}. For clarity's sake the auxiliary net is only drawn once and places with identical labels should be identified as one place (in Fig. 6.1 and 6.3 fusion places are shown in grey). In the auxiliary net a token moves from one place to the next for each step in the cycle and the events in the main net are scheduled by these token movements in the following order, in which the number of the steps corresponds to the number of the place w which contains a token at that point:

1. For $1 \leq i \leq n$, transition c_i fills (in m_{x_i} auto-concurrent occurrences) place x'_i with $pD\Delta t \cdot m_{x_i}$ tokens and transition c_{i+1} fills (in $m_{x_{i+1}}$ auto-concurrent occurrences) place x''_{i+1} (and simultaneously x'_{i+1}) with $pD\Delta t \cdot m_{x_{i+1}}$ tokens. In the same step transitions e'_i and e''_{i+1} empty x'_i and x''_{i+1} of any residual tokens left from the previous cycle. In addition, if place f contains a token, transition s outputs $J_0 \cdot \Delta t \cdot h$ morphogens to x_1.

2. Transition d_i removes tokens from places x'_i and x''_{i+1} in $m_{x''_{i+1}}$ auto-concurrent occurrences, thereby emptying x''_{i+1} and leaving the difference in x'_i ; in other words, $m_{x'_i} = pD\Delta t \cdot m_{x_i} - pD\Delta t \cdot m_{x_{i+1}}$ in.

3. Transition t'_i fires and transports $\alpha/100 l^2$ tokens from x_i to x_{i+1}.

4. In step 4 and 5, the degradation of morphogens in the individual cells is addressed. In this step transition g_i inserts $bk\Delta t \cdot m(x_i)$ tokens into place r_i. Simultaneously, transition q_i empties r_i of any residual tokens left from the previous cycle.

5. Subsequently, transition g'_i removes one token from x_i for each b tokens in r_i.

Due to the auxiliary net, neighbouring pairs of cells are either all involved in calculation steps (steps 1, 2 and 4) or tokens are transferred between or removed from cells. During the calculation steps the token numbers in all places x_i, representing the biological cells, except place x_1 are therefore unchanged and can be accessed for reading by other transitions. In other words calculations are *orthogonal* to the basic operations of gradient formation. Another important feature of this approach is that it is purely local; interactions between neighbouring cells are independent of the token numbers in other cells or the length of the chain of cells.

6.4 A CASE STUDY OF DPP GRADIENT FORMATION TO VALIDATE THE PETRI NET MODEL

For a validation of the Petri net model we use data from (Kicheva *et al.*, 2007). In this study, gradient formation was examined for the protein Dpp (Decapentaplegic), in the wing of the fruit fly, *Drosophila melanogaster*. These proteins were studied as they emanated from a source through the wing epithelium. The gradient could be treated as a series of places. The 3D situation was captured in a stack of images. Firstly, a maximum projection of this stack reduced the tissue to a two-dimensional plane; this could further be reduced to a line of places, since the rectangular region of interest lay parallel to the rectangular source tissue and movement at the lateral sides was negligible *cf.* section 6.2.

Kicheva *et al.* (2007) studied the behaviour of gradient formation and the role played in this by the process of endocytosis, *i.e.* the uptake of particles through membrane vesicles into the cell (*cf.* section 5.3.1), which is known to contribute to the formation of many gradients, in addition to diffusion (Fischer *et al.*, 2011; Gurdon and Bourrillot, 2001; Lander *et al.*, 2002; Scholpp and Brand, 2004; Teleman *et al.*, 2001). To this end the authors created a partial endocytotic block in animals which were mutant for the *shibire* allele and in which the source was rescued by a *shibire*[+] transgene. Using an experimental setup, monitoring fluorescent recovery after photobleaching (known as a FRAP assay), the values for D and k were determined under different experimental conditions of the gradient formation. Here we simulate the process for the *Dpp shibire* mutant at 32°C (Dpp-rescue) and the Dpp control group at 32°C (Dpp).

For these conditions the following values for D and k were found by Kicheva *et al.* (2007) and used in the Petri net model presented here (omitting the standard deviation): for Dpp $D = 0.10$ and $k = 2.52$ and for Dpp-rescue $D = 0.06$ and $k = 1.53$. Based on these values, values for p and b were set at $p = 10^2$ and $b = 10^5$, in order to minimize errors in rounding off. The simulation results from the Petri net model are compared to those predicted by the DE model (1), using the experimentally determined parameter values for D, k, l, j_0 and h as found by Kicheva *et al.* (2007). For this validation the number of cells to be modelled has been set at 30, which is a large enough number to accurately model L, given the current case study. The Petri net therefore describes the situation of a linear array of 30 cells, with a constant influx of morphogen at the left ($r = 0$). At the far right side we assumed that morphogens cannot flow out of the last cell. In our DE model this is represented by zero flux boundary conditions at $r = L$ (see section 6.2; $L = 30l$). Note that this differs from the DE model employed in Kicheva *et al.* (2007), where the array of cells is assumed to extend infinitely far. The finite number of cells in the Petri net has been taken into account in the derivation of parameters from the differential equations (*cf.* 6.2).

Gradient formation is considered finished once a steady state has been reached, *i.e.* a state in which morphogen concentrations stay the same in all cells, due to a balance between production, diffusion and degradation. For the diffusion equation model, the exact steady-state solution $C*$ to the diffusion equation (1) is given by

$$C*(r) = \frac{j_0}{\sqrt{kD}} e^{-\mu r}, \qquad \mu := \sqrt{k/D} \tag{7}$$

In our case, with a finite array of cells and Neumann conditions at $r = L$, (7) requires an additional correction factor: the exact steady state solution becomes

$$C*(r) = \frac{j_0}{\sqrt{kD}} e^{-\mu r} \cdot \frac{1 + e^{2\mu(x-L)}}{1 - e^{-2\mu L}} \tag{8}$$

For comparison of the density description by means of C with the number of tokens in a cell as computed by the Petri net the first must be converted to the number N_k of morphogen molecules in cell k, by means of

$$N_k(t) := h \int_{(k-1)l}^{kl} C(r,t) dr , \tag{9}$$

where l denotes the cell length, h the cell height and $k = 1, 2, \dots , 30$. The integral in (9) is approximated by means of the basic trapezoidal rule, yielding

$$N_k(t) \approx hl \cdot \frac{1}{2} \big[C((k-1)l, t) + C(kl, t) \big]. \tag{10}$$

Here l = 2,6 μm and h = 2,6 μm.

In the Petri net a steady state is reached once the marking of the entire net stays identical for two consecutive step cycles, since the behaviour of the net is deterministic and the parameter values remain unchanged. For each of the experiments the Petri net solution with corresponding parameter values was implemented in the software tool Snoopy (Rohr *et al.*, 2010) and markings of the places $x_1, \dots , x_n$ for every 5 steps were obtained using our in-house analysis tool PetriCalc. Snoopy was used as an interface for the creation of the net, but due to the size of the net and the high numbers of tokens to be processed, analysis was done with PetriCalc. This has been developed based on the analytical methods of Snoopy. The steady state, as reached by the Petri net for Dpp and Dpp-rescue, was found to closely correspond to the steady state given by (8) and (10), with only minor deviations: at most 1.4 % for Dpp-rescue and 0.2 % for Dpp.
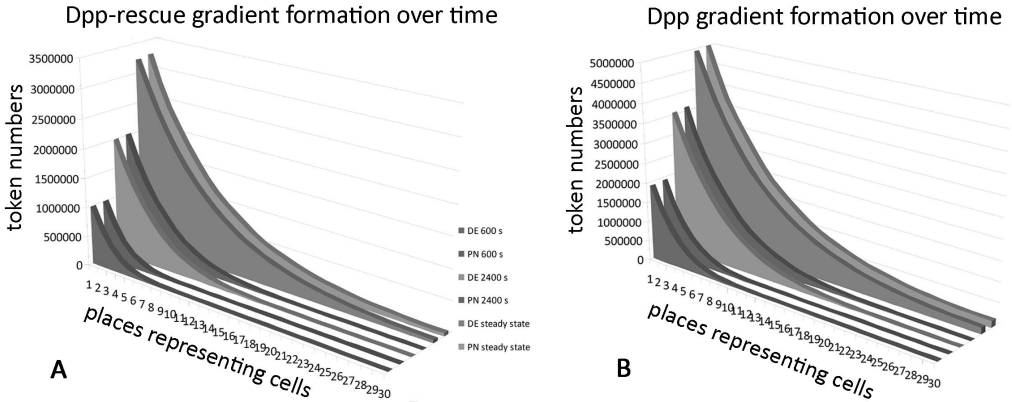
Figure 6.2. Visualisation of different states in the process of gradient formation for Dpp-rescue (A) and Dpp (B), at time points t = 600s, t =2400s and t = 12000s, of which the latter represents the steady state of the process. For each of the time points the token numbers in the places representing cells, *i.e.* x_1 to x_{30}, as predicted by the Petri net model (PN) is compared to the token numbers derived from the concentration levels, predicted by the DE model.

In addition to the steady state, we also compared the gradient formation at moments during the process, t = 600s and t = 2400s. For the DE model, these time-dependent solutions were computed using COMSOL Multiphysics package (version 4.2.0.150) for the finite element. Again the Petri net and the DE model yielded corresponding results, with minor deviations: on average 0.01% for Dpp with a maximum deviation of 0.46% and on average 0.02% for Dpp-rescue with a maximum of 0.2%. In Fig. 6.2 the Petri net marking corresponding to times t = 600s, t = 2400s and the t = 12000 are compared to those predicted by the DE model; the situation at t = 12000s represents the steady state.

6.5 CONCLUSION AND DISCUSSION

We have presented a generic Petri net model for biological gradient formation, based on the model presented in chapter 5. The current model incorporates elements from a discretised DE model for gradient formation. The resulting parametrized Petri net is more versatile and can be fitted to a wide range of instances of the process. This has been shown for the case study of Dpp and Dpp-rescue gradient formation in the fruit fly.

As with DE models, the Petri net allows quantitative analysis of gradient formation. Due to its modular nature the Petri net can easily be adjusted to other observed instances of gradient formation, both with regard to changes in parameter values and to the length of the tissue under study. While the current model represents
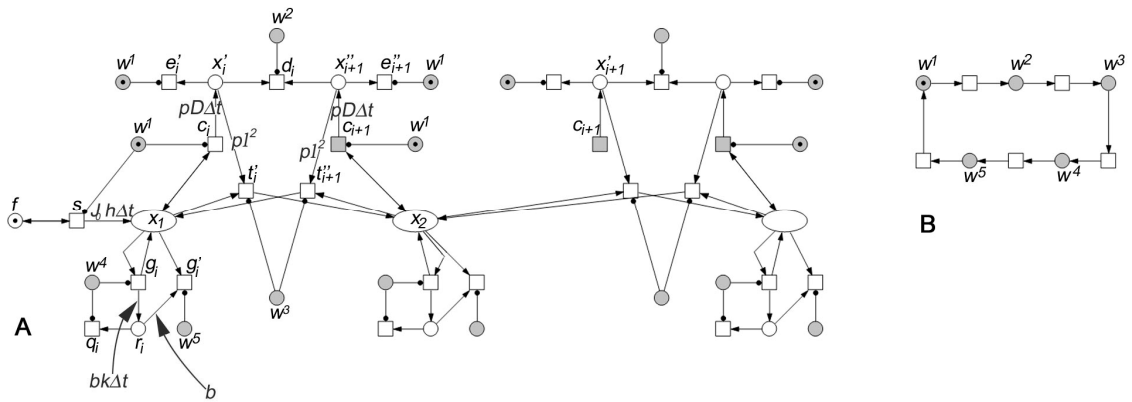
Figure 6.3. The main construction (A) of the net and auxiliary net (B) for the situation of two-directional gradient formation. Note that the main construction shown here differs from the main construction in Fig. 6.1 solely in the addition of transition t''_{i+1}.

the tissue as a one-dimensional structure, *i.e.* a line of cells, the approach is amenable to extension in two and three dimensions, by connecting cells not only in the x-dimension, but also in y- and z-dimensions. In this way the approach increases possibilities for quantitative analysis of a wide range of experimental conditions. Furthermore, by adding one transition, the Petri net can be made to model transport of tokens between neighbouring cells in two directions, as illustrated in Fig. 6.3.

The graphical element of the Petri net framework makes the process insightful to biologists and amenable to experimental setups which interfere with the unfolding of the process. In this way it becomes possible to simulate *e.g.* grafting experiments, in which part of the tissue is removed or replaced and the effects are studied. For gradient formation in particular, experiments have been performed with FRAP assays (Kicheva *et al.*, 2007; Carrero *et al.*, 2003). In these experiments part of a tissue containing fluorescently labelled proteins is locally photobleached, after which recovery of the gradient of fluorescent molecules is studied. Since the structure of the presented Petri net closely resembles the observed biological situation, it can be used to simulate similar experiments using quantitative data, by removing places which correspond to particular biological cells or depleting these of tokens.

Finally, we hope to explore possibilities of building hierarchical nets, using *e.g.* nets-within-nets (Valk, 2004) and/or refinement, to model particular subcellular processes, such as passive and active diffusion through the extracellular space and degradation by means of endocytosis, *cf.* 5.3.1.

The Petri net of gradient formation illustrates an approach that brings together aspects from process and equation based modelling techniques, *i.e.* from Petri nets and

differential equation models; this merger of techniques provides new prospects for simulation and analysis of the process of gradient formation. On a broader scale it underlines the importance of integrative methods for the field of developmental biology, which will be addressed in more detail in the next chapter.

CHAPTER 7.   INTEGRATION OF MODELLING TECHNIQUES

In the previous chapters we have addressed several case studies of modelling techniques, applied to processes in developmental biology. The development of the heart in the turtle *Emys orbicularis* was studied using 3D reconstructions in chapter 2, followed by a detailed study of the outflow tract development, using 3D reconstructions based on high resolution images in chapter 3. Subsequently, the use of ontologies in describing vertebrate anatomy, development and physiology was examined in chapter 4. Finally, gradient formation was modelled using the modelling framework of Petri nets. Both a conceptual, qualitative approach was presented, in chapter 5, and a more concrete, quantitative approach, incorporating aspects of differential equation models, in chapter 6.

The used techniques each have their own particular merits, but also their own particular limitations. The visual reconstructions gave us insight into the three-dimensional structure of the developing turtle heart, but did not tell us much about its specific physiology. The Petri net model of gradient provided us with a way to simulate gradient formation, but did not inform us about the visual appearance of the resulting developmental patterns. In order to extend the functionality of these modelling techniques and try and overcome some of these limitations the goal of the research has been twofold. In the case studies, in particular in chapters 2 and 3, developmental processes or structures have been studied within a biological context. In this way they contributed to our understanding of developmental biology. But at the same time, for each of the cases an integration of modelling techniques was strived after, in order to extend the methodology used in models of developmental biology. With the increase in data production in the life sciences, due to *e.g.* high throughput techniques, novel methods are needed for analysis. Here integration plays an important role, both of data and of methods. The rising interest in these types of integration is illustrated by the publication of a journal specifically dedicated to this topic, the *Journal of Integrative Bioinformatics*. This journal covers data integration, method integration and combinations of both, in particular for the fields of molecular and systems biology. In general, data and method integration are currently focused on the field of systems biology (Hoehndorf *et al.*, 2011; Takai-Igarashi, 2005; Zhang and Verbeek, 2010), dealing mainly with pathway modelling. However, the field of higher level developmental biology (*i.e.* on tissue and organ level) also stands to gain from integration techniques. Method integration is in this case more relevant than data integration, since data increase is less prominent in higher level developmental biology. It is for this reason that we have investigated method integration for this field by means of the presented case studies.

For each of the separate studies the biological context and relevance has been addressed in the corresponding chapters. This concluding chapter will focus on the ways in which integration of modelling techniques has been used or proposed to increase the (accessibility of) knowledge contained within the model system. In chapter 1 four categories of models were presented (verbal, visual, algorithmic and equation based) and the current chapter is based on this same distinction. For each category, integration methods have been investigated, albeit not yet fully implemented or presented for all cases in the previous chapters. Here I present a short overview, looking at these combinations one by one. Additionally, in the case studies we have looked at combinations of models of the same type, for different purposes. These are discussed in section 7.5. At the end of the chapter a table is provided, in which an overview is given of all combinations described here.

7.1 VERBAL AND VISUAL: COMBINING ONTOLOGIES WITH DIAGRAMS AND 3D MODELS

In chapter 4 the functionality of ontologies for developmental biology has been discussed. We have looked at new ways to model knowledge of biological structures using ontologies. One way of doing this was to change the focus of the organization of the captured knowledge, from one particular species to the group of vertebrates in general. Another proposed way of extending the functionality was by integrating other ways of presenting the data. This has been done by including both visual and algorithmic functions, of which the latter will be discussed in the next paragraph.

Two ways of visualising the knowledge can be used for the vertebrate heart ontology. First of all the information captured in the object and data properties (linking classes and instances to one another or to pieces of data respectively) can be visualised in diagrams, as mentioned in 1.3.2. In these **visualisations of ontology information in diagrams**, classes and instances are represented as nodes and the connections between these as edges; these diagrams can be categorized as class diagrams. As can be seen from the blood flow diagram presented as an example in Fig. 7.1, this way of visualising the data helps the user get an overview of information contained in the ontology. When only taking into account the verbal information in the ontology, it is hard to get the full picture of for instance blood flow. Diagrams such as these can be produced for all properties modelled in the ontology. They are particularly useful for our dynamic ontology, in which different physiological situations are modelled; a diagram can provide insight into a particular situation of interest.

A second way in which visual models and ontologies can be combined is by **visualising results of ontology queries in 3D reconstructions**, as discussed in 4.4.4, *cf.*
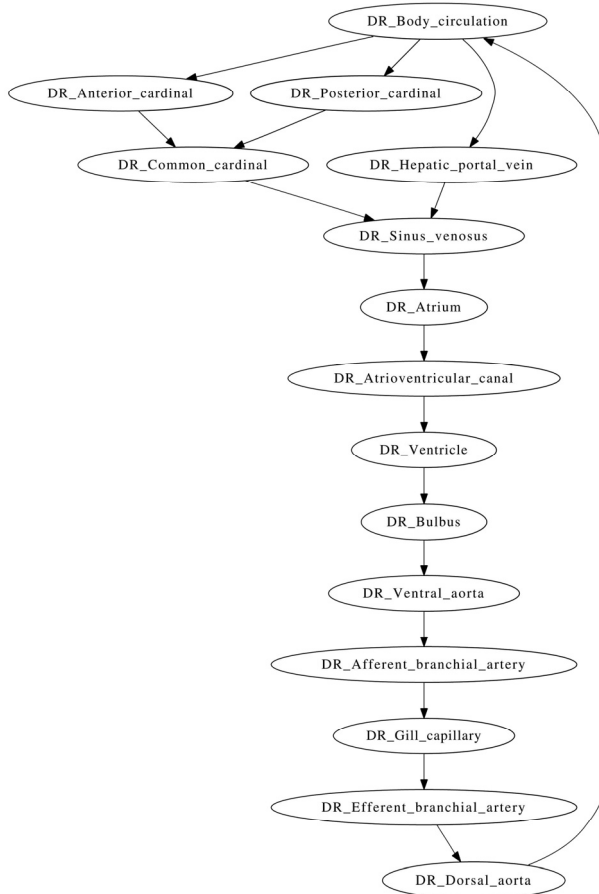
Figure 7.1. Diagrammatic representation of the blood flow information for *Danio rerio*, contained in the ontology system, *cf.* chapter 4.

Fig. 4.4. Combining these methods has benefits both for the 3D reconstructions and for the ontology. By linking the 3D reconstructions to an ontology one is forced to use a standard vocabulary when annotating the structures, in order to identify these with classes and instances in the ontology. This standard vocabulary makes it easier to compare and study different reconstructions and helps prevent miscommunication. The overview of turtle heart development (provided on the website http://bio-imaging.liacs.nl/galleries/), as presented in chapter 2, illustrates this. At the same time, combining these methods also enriches the ontology. A query of the ontology results in a list of names of relevant structures and/or instances and visualising these in a 3D reconstruction provides the user with more information. Other examples of combining ontologies with 3D reconstructions exist (*e.g.* Köhn *et al.*, 2004) and in chapter 4 our particular method and its benefits have been explained.

7.2 VERBAL AND ALGORITHMIC: COMBINING ONTOLOGIES WITH PETRI NETS

The vertebrate heart ontology has been developed with future extensions in mind, in particular with a complementary Petri net model. The information on context-dependent heart physiology, modelled using object and data properties in the ontology, lends itself for modelling in a process oriented method like Petri nets. Our aim is therefore to **connect the ontology to a Petri net simulating the blood flow**, in which the Petri net takes information on blood pressures and shunting from the ontology.

Currently we are working on new tools for the simulation and analysis of Petri nets, for instance using mathematical software. Several tools exist for the implementation and simulation of Petri nets, *e.g.* Snoopy, CPNTools and Pipe2, each restricted by certain limitations. Developing a more generic tool, which allows a range of input file types for the Petri nets and gives the user control over arc functions and firing rules, we hope to overcome these limitations; one of the strengths of Petri nets is the ease with which its theory can be extended to include *e.g.* new arc types or firing rules. This makes the modelling framework easily amenable to particular needs of a research field like developmental biology. In response to the expansion of Petri net theory (*e.g.* with Hybrid Petri nets), existing tools are constantly being adapted and new tools are developed. A tool which allows the user direct access to the theoretical basis will create greater freedom in modelling complex biological processes.

In this way we hope to be able to extract variable values from the ontology and use them in the weight functions of the corresponding Petri net. Consequently, we would combine the semantic properties of ontologies with the ability to model (concurrent) processes of Petri nets.

Combinations of Petri nets and ontologies have been used and reported sporadically for other fields (*e.g.* Takai-Igarashi, 2005; Recker and Indulska, 2007), but there the focus has been different; ontologies were used in these cases to establish a standard 'language' for Petri net components within a particular research area (*e.g.* signalling pathways; Takai-Igarashi, 2005). This systematization of Petri net semantics enables researchers to communicate and share Petri net models more easily. To date no combinations of Petri nets and ontologies have been presented in literature for the field of developmental biology and in particular for physiological processes. Furthermore no methods have been described in which simulation of a Petri net model directly relies on information provided by the ontology in a dynamic way, such as proposed for the vertebrate heart ontology. Developing such a method will help extend the functionality of both ontologies and Petri nets for the field of developmental biology.

7.3 ALGORITHMIC AND VISUAL: COMBINING PETRI NETS WITH 3D MODELS

In chapters 5 and 6 the problem of modelling gradient formation is addressed and a solution is presented with Petri nets. This model serves as a proof of concept for the use of Petri nets for higher level developmental biology. Originally, the biological process selected as a case study was the formation of the anterior-posterior axis (AP-axis) in the frog *Xenopus laevis*. This process comprises several smaller processes, one of which is gradient formation. The model of gradient formation has been constructed with the underlying future aim of combining it with other modelled elements of the AP-axis formation, in order to obtain an all-encompassing Petri net model of the process (this is addressed in more detail in 7.5).

An important aspect of the process of AP-axis formation is the expression of a series of *Hox* genes, which play a pivotal role in the development of all vertebrates. Changes in the order of the *Hox* gene expression in this process can lead to severe developmental aberrations (Wolpert, 2002; McNulty *et al.*, 2005; Stern *et al.*, 2006; Iimura and Pourquié, 2007). The expression patterns of the *Hox* genes as seen in embryonic development can be visualised by staining the gene products using the technique of (fluorescent) *in situ hybridization*. By performing experiments which interfere with normal AP-axis formation (for instance by knocking out *Hox* genes, *cf.* McNulty, 2005; or by removing the organiser mesoderm, *cf.* Jansen *et al.*, 2007), anomalous patterns can be obtained that provide us with new insights in the process. Ideally, a Petri net model of the entire process would be able to reproduce and possibly predict experimental results. When executing the Petri net, it would be helpful if users were provided with visual information about the process in addition to the abstract information on token movement. Therefore we would like to **connect states of the Petri net model to 3D reconstructions with corresponding *Hox* gene expression patterns**.

ectoderm   ■ Hox D1 expression
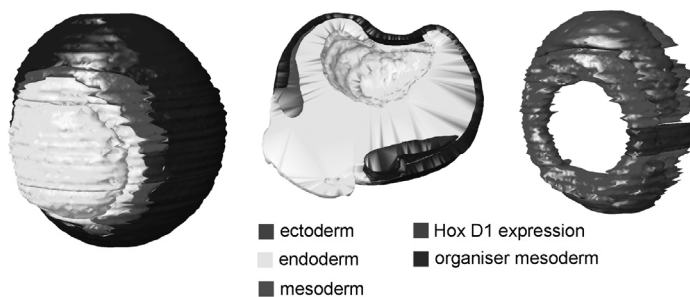endoderm   ■ organiser mesoderm
mesoderm

Figure 7.2. 3D reconstruction of a stage 11 *Xenopus laevis* embryo, showing the three germ layers and the Spemann organiser, as well as the expression pattern of *Hox* gene D1.

To this end we have, in collaboration with the Institute of Biology Leiden (IBL), obtained several sectioned *Xenopus laevis* embryos of different developmental stages in which different *Hox* genes have been visualised by *in situ hybridization*. From section images of these embryos, 3D reconstructions have been created (using the methods described in chapter 2), of which an example is shown in Fig. 7.2. Since in situ hybridization is only possible for two (or sometimes three) genes in one experiment, we cannot visualise expression patterns for more than two genes at a time in one embryo. Therefore work is being done on developing a template model, onto which multiple 3D reconstructions of *Hox* gene expression patterns can be projected simultaneously. The end result of this will allow us to produce visual information on many different situations of gene expression, corresponding to particular states of the Petri net (and thus to particular states in the biological process). In future we hope to connect these visualisations to the Petri net model itself, *i.e.* we would like for the model to present visual output whenever a state is reached for which this information is available.

7.4 EQUATION BASED AND ALGORITHMIC: COMBINING DIFFERENTIAL EQUATIONS WITH PETRI NETS

In chapter 6 we have combined the Petri net model of gradient formation with differential equation models of the same process. As described in 1.4, algorithmic and equation based models differ significantly and both have particular merits. Equations allow precise quantitative analysis of the continuous behaviour (both spatially and temporally) of a process, while an algorithmic approach is better suited to study particular and possibly concurrent events in the process. By **incorporating parameters from differential equations in a Petri net model**, we hope to attain exact analysis and insight into complex processes consisting of different types of events.

While differential equation models describe a process in a continuous manner but can be discretised in time and space, Petri nets start off by describing changes in the system states in discrete space and time steps. Petri nets allow one to model concurrent events and hierarchical coherence of processes, while a hierarchical approach for differential equations, using multiscale modelling, is feasible but quite involved. Thus the two modelling approaches each hold particular advantages for modelling in developmental biology and combining aspects opens up new possibilities in the simulation and analysis of complex developmental processes. The current Petri net model of gradient formation has been designed to accommodate the parameters of existing differential equation models and has been validated using data obtained from these equations (described in detail in chapter 6).

In future we hope to be able to directly implement differential equations in the Petri net, by enabling the weight functions of arcs to include more complex formulas. The current work on simulating Petri nets using mathematical tools (*cf.* section 7.2) is specifically focused on this. This will result in a hybrid model, which will allow more complex computation, but will still operate in discrete temporal and spatial steps.

7.5 INTEGRATING MODELS OF THE SAME MODELLING APPROACH

In addition to integrating different types of models, each of the case studies has also featured combinations of models of the same modelling approach. The vertebrate heart ontology system, described in chapter 4, is deliberately referred to as an ontology *system*, instead of merely an ontology, since this model is a **system of three combined ontologies**. We constructed an ontology for anatomical information and one for developmental information and used the NCBI database ontology for taxonomic information. By separating these different types of information the anatomy and development ontologies could be used on their own, increasing their functionality (as explained in 4.3). Furthermore, using the NCBI database provided us with a standardized way of addressing and including species information which increased the inter-operability of the system. While all the information currently present in the system could have been modelled in one all encompassing ontology, dividing information over separate ontologies provides the system with more freedom for future expansion and integration.

Secondly, 3D reconstructions of cardiogenesis in the turtle *Emys orbicularis*, presented in chapters 2 and 3, were combined to form a developmental series, illustrating essential stages in this process in a chronological order. In addition to these models of cardiogenesis, several other series of models have been produced during the research in its entirety, for instance the *Xenopus laevis* models, mentioned in section 7.3, and series of zebrafish gene expression models of different developmental stages. These models are all presented in an **online atlas of 3D reconstructions** (Potikanond, 2012) which can be browsed and queried, allowing users access to all information in the models as a whole, as opposed to letting them study the models one at a time. Incidentally, this approach is also related to modelling with ontologies, since the database organizes the models and their annotations in a structural way that can be examined in context and queried.

Another way in which 3D reconstructions could be combined is by **linking low resolution 3D reconstructions with high resolution 3D reconstructions** (*i.e.* models constructed from low and high resolution images, respectively). In chapters 2 and 3 we have presented models of the entire turtle heart in low resolution and detailed models

of the outflow tract in high resolution. Apart from the biological insights gained from these models, a second purpose of this approach was to integrate these models in a hierarchical manner. The high resolution models provide a detailed view of one particular structure, the outflow tract, but at the cost of losing the overview of the entire heart and the position of the outflow tract in it; on the other hand, models of the entire heart cannot provide the level of detail necessary to answer particular questions about the development of the outflow tract. The in-house software, 3D acq, used to capture the section images, stores the stage coordinates of the microscope for each image in a database (*cf.* chapter 3), making it possible to relate high and low resolution images of the same section on a pixel level. This in turn can be used to connect structures in an all encompassing low resolution model to high resolution models of these structures. Ideally the visualisation tool would include a function allowing the user to select certain biological structures and zoom in on these.

A final modelling approach for which integration has been investigated concerns Petri net models. As mentioned in 7.3, gradient formation is an important part of AP-axis formation in the species *Xenopus laevis*. In future we aim to construct additional models for the other events in the AP-axis formation and combine these models and to **combine Petri net models of separate processes in an all encompassing model of the entire process**. The other subprocesses of AP-axis formation include *Hox* gene expression, cell layer movement and vertical signalling. We would like to design separate Petri net models for each of these subprocesses, based on the same assumptions and modelling decisions as the current model.

The models used to construct an all encompassing model would be of the same biological level (the same level of magnification as it were). However, as with the 3D reconstructions, a second way of integrating Petri nets is possible, by using models on different levels, creating as it were a possibility to zoom in (*cf.* combining low and high resolution 3D reconstructions). In the case of Petri nets this would amount to **a hierarchical Petri net**, which allows one to refine particular events in the net in more detailed subnets. In analogy with (microscopic) imaging techniques, it would correspond to changing resolutions through different imaging modalities. This can be achieved by identifying a transition or place with a complete net on a lower level of refinement; when for instance such a transition fires, the net inside is activated and generates output to the transition.

The model of gradient formation, presented in chapter 6, is amenable to this approach. Here the process of morphogen dissipation through the tissue has been treated as 'effective diffusion', corresponding to the differential equation model presented by Kicheva *et al.* (2007); no distinction is made between different spreading mechanisms, such as active and passive diffusion or endocytosis. For the current model this generalization suffices, but in case where more detailed information is available this

could be added to the net by means of a subnet specifying all transitions t'_i and t''_i. Similarly, the process of endocytosis, which accounts for the degradation of morphogens, could be specified by connecting subnets to all transitions p'_i. Due to the modular nature of the system only one subnet design is required for each of these detailed mechanisms; general nets for endocytosis and diffusion will serve as building blocks and can be added to each of the cells (or pairs of neighbouring cells).

	VERBAL ONTOLOGY	VISUAL 3D MODEL	ALGORITHMIC PETRI NET	EQUATION BASED DIFFERENTIAL EQUATION
VERBAL ONTOLOGY	system of combined ontologies of the vertebrate heart	- diagrammatic visualisation of information - visualisation of ontology query results	- process model of blood flow	
VISUAL 3D MODEL		- online atlas of 3D reconstructions - linking low and high resolution models	- *Hox* gene expression reconstructions linked to PN	
ALGORITHMIC PETRI NET			- all encompassing model of AP-formation - hierarchical net with specification of diffusion and degradation mechanisms	- Petri net of gradient formation incorporating parameters from differential equations

Table 7.1. Overview of ways in which the modelling techniques, presented in this dissertation, were integrated. Elements in black have been realised, elements in gray are planned and have been taken into account in the modelling decisions of the current models.

———

   This concludes the discussion of integrative modelling methods, used in this research, and the presentation of the research as a whole. In order to investigate and further the use of modelling techniques for developmental biology, four modelling approaches were selected, each on the basis of particular merits for modelling phenomena in developmental biology and each representing a different category from chapter 1. This has led to a broad scope of the study. The 3D reconstructions underlined the significance of spatial information and high resolution imaging in understanding complex developmental anatomy. Secondly, developmental, anatomical and physiological information was structured formally in the vertebrate heart ontology, in order to investigate the uses of formal organization of knowledge for this particular field. Thirdly, the Petri net framework was chosen for its ability to model the dynamics of complex processes, including modular and hierarchical elements, which are particularly relevant to developmental biology. And finally, by combining this with aspects from differential equations commonly used to model developmental processes such as gradient formation, the prospects for simulation and analysis were widened. In each of these projects, the integration of modelling methods has been taken into account as well, as described in this chapter. The combinations of the techniques, used in this research, are summarized in Table 7.1, which includes both projects that have been realised and projects that are planned for future work, as discussed above.

The modelling approaches, discussed in this dissertation, have thus provided us with novel ways to modelling biological phenomena, as well as new insights into developmental biology and have laid the groundwork for integrating diverse modelling methods for developmental biology.

BIBLIOGRAPHY

Anderson R.H., Wilkinson J.L., Arnold R. *et al.,* 'Morphogenesis of bulboventricular malformations I: Consideration of embryogenesis in the normal heart', *British Heart Journal* 36 (1974), 242-255.

Axelsson M., Franklin C.E., Löfman C.O., Nilsson S., Grigg G.C*.,* 'Dynamic anatomical study of cardiac shunting in crocodiles using high-resolution angioscopy', *Journal of Experimental Biology* 199 (1996), 359-365.

Axelsson M., Franklin C.E., Fritsche R., Grigg G.C, Nilsson S., ´The sub-pulmonary conus and the arterial anastomosis as important sites of cardiovascular regulation in the crocodile *Crocodylus porosus´*, *Journal of Experimental Biology* 200 (1997), 807-814.

Axelsson M., 'The crocodilian heart: more controlled than we thought?' *Experimental Physiology* 86:6 (2001), 785-789.

Baclawski K., Niu T., *Ontologies for bioinformatics,* Cambridge 2006.

Baeten J.C.M., 'A brief history of process algebra', *Journal of Theoretical Computer Science* 335 (2005).

Baldock R., Bard J., Burger A., Burton N., Christiansen J., Feng G., Hill B., Houghton D., Kaufman M., Rao J., Sharpe J., Ross A., Stevenson P., Venkataraman S., Waterhouse A., Yang Y., Davidson D., 'EMAP and EMAGE: a framework for understanding spatially organised data', *Neuroinformatics* 1 (2003), 309-325.

Banks, R., *Qualitatively modelling genetic regulatory networks: Petri Net techniques and tools*, Ph.D. Thesis, Newcastle 2009.

Banks R., Khomenko V., Steggles L.J., 'A case for using signal transition graphs for analysing and refining genetic networks', *ENTCS* 227 (2009), 3–19.

Bartelings M.M., Gittenberger A.C., 'Contribution of the aortopulmonary septum to the muscular outlet septum in the human heart', *Acta Morphol. Neerl-Stand* 24 (1986), 181-192.

Bartelings M.M., Gittenberger-de Groot A.C., 'The outflow tract of the heart – embryologic and morphologic correlations', *International Journal of Cardiology* 22 (1989), 289-300.

Bathoorn R., Welten M., Richardson M., Siebes A., Verbeek F.J., 'Frequent episode mining to support pattern analysis in developmental biology', *LNCS* 6282 (2010), 253-263.

Bertens L.M.F., Richardson M.K., Verbeek F.J., 'Analysis of cardiac development in the turtle *Emys orbicularis* (Testudines: Emidydae) using 3D computer modeling from histological sections', *Anatomical Record* 293 (2010), 1101-1114.

Biermann S., Uhrmacher A.M., 'Supporting multi-level models in systems biology by visual methods', *Proceedings of the 18th European Simulation Multiconference, Magdeburg, SCS Europe* (2004) chap.18, 1-8.

Blom N.A., Gittenberger de Groot A.C., Jongeneel T.H., DeRuiter M.C., Poelman R.E., Ottenkamp J., 'Normal development of the pulmonary vein in human embryos and formulation of a morphogenetic concept for sinus venosus defects', *American Journal of Cardiology* 87 (2001), 305-309.

Bodenreider O., Stevens R., 'Bio-ontologies: current trends and future directions', *Briefings in Bioinformatics* 7:3 (2006), 256-274.

Bonzanni N. *et al.*, 'Executing multicellular differentiation: quantitative predictive modelling of *C. elegans* vulval development', *Bioinformatics* 25 (2009).

Bradford Y., Conlin T., Dunn N., Fashena D., Frazer K., Howe D.G., Knight J., Mani P., Martin R., Moxon S.A., Paddock H., Pich C., Ramachandran S., Ruef B.J., Ruzicka L., Bauer Schaper H., Schaper K., Shao X., Singer A., Sprague J., Sprunger B., Slyke C. van, Westerfield M.. 'ZFIN: enhancements and updates to the zebrafish model organism database', *Nucleic Acids Research* 39: suppl 1 (2011), 822-829.

Burger A., Davidson D., Baldock R. (eds.), *Anatomy ontologies for bioinformatics*, London 2008.

Burggren W.W., 'Form and function in Reptilian circulations', *American Zoology* 27 (1987), 5-19.

Burggren W.W., 'Cardiac design in lower vertebrates: what can phylogeny reveal about ontogeny?' *Experientia* 44 (1988), 919-930.

Burggren W.W., Warburton S.J. 'Patterns of form and function in developing hearts: contributions from non-mammalian vertebrates', *Cardioscience* 5 (1994), 183-192.

Burkhard H.-D., 'On Priorities of parallelism: Petri Nets under the maximum firing strategy', *LNCS* 146 (1983), 86–97.

Carrero G., McDonald D., Crawford E., Vries G. de, Hendzel M.J., 'Using FRAP and mathematical modelling to determine the in vivo kinetics of nuclear proteins', *Methods* 29 (2003), 14-28.

Chaouiya C., 'Petri net modelling of biological networks', *Briefings in Bioinformatics* 8 (2007), 210-219.

Crick F., ´Diffusion in embryogenesis´, *Nature* 225 (1970), 420-422.

David R., Alla H., 'On hybrid Petri nets', *Discrete Event Dynamic Systems: Theory and Applications* 11 (2001), 9-40.

Dmitrieva, J.B., *Aspects of ontology visualisation and integration*, Ph.D. Thesis, Leiden 2011.

Eilbeck K., Lewis S.E., Mungall C.J. *et al.*, 'The Sequence Ontology: a tool for the unification of genome annotations', *Genome Biology* 6:5 (2005), article R44.

Eils R., Athale C., 'Computational imaging in cell biology', *The Journal of Cell Biology* 161:3 (2003), 477-481.

Ellner S.P., Guckenheimer J., *Dynamic models in biology*, Princeton 2006.

Eme J., Gwalthney J., Blank J.M., Owerkowicz T., Barron G., Hicks J.W., 'Surgical removal of right-to-left cardiac shunt in the American alligator (*Alligator mississippiensis*) causes ventricular enlargement but does not alter apnoea or metabolism during diving', *Journal of Experimental Biology* 212 (2009), 3553-3563.

Entchev E., Gonzalez-Gaitan M., 'Morphogen gradient formation and vesicular trafficking', *Traffic* 3 (2002), 98–109.

Fischer J., Henzinger T.A., 'Executable cell biology', *Nature Biotechnology* 25:11 (2007), 1239-1249.

Fischer J., Harel D., Henzinger T.A., 'Biology as reactivity', *Communications of the ACM* 54:10 (2011).

Fischer J.A., Eun S.H., Doolan B.T., 'Endoytosis, endosome trafficking, and the regulation of *Drosophila* development', *The Annual Review of Cell and Developmental Biology* 22 (2006), 181–206.

Gasch F.R., Beiträge zur vergleichenden Anatomie des Herzens der Vögel und Reptilien', *Archiv für Naturgeschichte* 54 (1888), 119-152.

The Gene Ontology Consortium, ´Gene Ontology: tool for the unification of biology´, *Nature Genetics* 25 (2000), 25-29.

Gilbert D., Heiner M., 'From Petri nets to differential equations - an integrative approach for biochemical network analysis', *LNCS* 4024 (2006), 181–200.

Gilbert D., Heiner M., Lehrack S., ´A unifying framework for modelling and analysing biochemical pathways using Petri nets´, *LNBI* 4695 (2007), 200–216.

Gilbert S.F., *Developmental biology*, Sunderland 2000.

Goodrich E.S., 'Heart', in: Goodrich E.S. (ed.), *Studies on the structure and development of vertebrates,* London 1958.

Goor D.A., Dische R., Lillehei C.W., 'The conotruncus I: Its normal inversion and conus absorption', *Circulation* 46 (1972), 375-364.

Greil A., 'Beiträge zur vergleichenden Anatomie und Entwicklungsgeschichte des Herzens und des Truncus arteriosus der Wirbeltiere', *Morphologisches Jahrbuch* 275 (1903), 204-216.

Gregor T., Bialek W., De Ruyter R.R. *et al.*, 'Diffusion and scaling during early embryonic pattern formation', *PNAS* 102:51 (2005), 18403-18407.

Grigg G.C, Johansen K., 'Cardiovascular dynamics in *Crocodylus porosus* breathing air and during voluntary aerobic dives', *Journal of Comparative Physiology* 157 (1987), 382-392.

Grigg G.C., 'Central cardiovascular anatomy and function in Crocodilia', in: Wood S.C., Weber R.E., Hargens A.R., Millard R.W. (eds.), *Lung biology in health and disease*, *Volume 56,* New York 1991.

Gruber, T.R., 'A translation approach to portable ontology specifications', *Knowledge Acquisition* 5:2 (1993), 199-220.

Gurdon J.B. *et al.*, 'Activin signalling and response to a morphogen gradient', *Nature* 371 (1994), 487–492.

Gurdon J.B., Bourillot P., ´Morphogen gradient interpretation´, *Nature* 413 (2001), 797–803.

Haendel M.A., Neuhaus F., Osumi-Sutherland D., Mabee P.M., Mejino Jr. J.L.V., Mungall C.J., Smith B., ´CARO – the Common Anatomy Reference Ontology´, in: Burger A., Davidson D., Baldock R. (eds.), *Anatomy ontologies for bioinformatics: principles and practice*, New York 2008.

Hamburger V., Hamilton H.L., 'A series of normal stages in the development of the chick embryo', *Developmental Dynamics* 195 (1951), 231-272.

Hart N.H., 'Formation of septa in the bulbus cordis of a turtle and a lizard', *Journal of Morphology* 125 (1968), 1-21.

Heiner M., Gilbert D., Donaldson R., 'Petri nets for systems and synthetic biology', *SFM 2008*, LNCS 5016 (2008), 215-264.

Hicks J.W., Wang T., 'Functional role of cardiac shunts in reptiles', *Journal of Experimental Zoology* 275 (1996), 204-216.

Hicks J.W., Krosniunas E., 'Physiological states and intracardiac shunting in non-crocodilian reptiles', *Experimental Biology Online* 1:(3 (1997).

Hoehndorf R., Dumontier M., Gennari J.H. *et al.*, ´Integrating systems biology models and biomedical ontologies´, *BMC Systems Biology* 5:124 (2011).

Holmes E.B., 'A reconsideration of the phylogeny of the tetrapod heart', *Journal of Morphology* 147 (1976), 209-228.

Hopcroft J.E., Motwani R., Ullman J.D., *Introduction to automata theory, languages, and computation*, Boston 2000.

Hu N., Sedmera D., Yost H.J., Clark E.B., 'Structure and function of the developing zebrafish heart', *Anatomical Record* 260 (2000), 148-157.

Hunter A., Kaufman M.H., McKay A., Baldock R., Simmen M.W., Bard J.B.L., 'An ontology of human developmental anatomy', *Journal of Anatomy* 203:4 (2003), 347-355.

Icardo J.M., 'Development of the outflow tract', *Annals of the New York Academy of Sciences* 588 (1990), 26-40.

Iimura T., Pourquié O., ´Hox genes in time and space during vertebrate body formation', *Development, Growth and Differentiation* 49 (2007), 265-275.

Iwasa J.H., 'Animating the model figure', *Trends in Cell Biology* 20:12 (2010), 699-704.

Janicki R., Koutny M., 'Semantics of inhibitor nets', *Information and Computation* 12 (1995), 1–16.

Jansen H.J., Wacker S.A., Bardine N. *et al.*, 'The role of the Spemann organizer in anterior-posterior patterning of the trunk', *Mechanisms of Development* 124 (2007), 668-681.

Johansen K., Burggren W.W., 'Cardiovascular function in lower vertebrates', in: Bourne G. (ed.), *Hearts and heart-like organs*, New York 1980.

Jong H. de, 'Modelling and simulation of genetic regulatory systems: a literature review', *Journal of Computational Biology* 9:1 (2002), 67-103.

Jong H. de, Geiselmann J., 'Modelling and simulation of genetic regulatory networks by ordinary differential equations', in: Chen J., Dougherty E.R., Shmulevich I. *et al.* (eds.), *Genomic Signal Processing and Statistics*, New York 2005.

Junck J.R., 'Ten equations that changed biology: mathematics in problem-solving biology curricula', *Bioscene* 23:1 (1997), 11-36.

Kauffman S., 'Metabolic stability and epigenesis in randomly constructed genetic nets', *Journal of Theoretical Biology* 22 (1969), 437-467.

Kicheva A., Pantazis P., Bollenbach T., Kalaidzidis Y., Bittig T., Juelicher F., Gonzalez-Gaitan M., 'Kinetics of morphogen gradient formation', *Science* 315 (2007), 521-525.

Kimmel C.B., Ballard W.W., Kimmel S.R., Ullmann B., Schilling T.F., 'Stages of embryonic development of the zebrafish', *Developmental Dynamics* 203 (1995), 253-310.

Kleijn J., Koutny M., Rozenberg G., 'Towards a Petri net semantics for membrane systems', *LNCS* 3850 (2006), 292–309.

Kleijn J., Koutny M., 'Processes of Petri nets with range testing', *Fundamental Informaticae* 80 (2007), 199–219.

Kleijn J., Koutny M., 'Processes of membrane systems with promoters and inhibitors', *Theoretical Computer Science* 404 (2008), 112–126.

Knublauch H., Fergerson R.W., Noy N.F., Musen M., 'The *Protégé* OWL Plugin: an open development environment for semantic web applications', *LNCS* 3298 (2004), 229–243.

Koch I., Junker B.H., Heiner M., ´Application of Petri net theory for modelling and validation of the sucrose breakdown pathway in the potato tuber', *Bioinformatics* 21 (2004), 1219–1226.

Koch I., Reisig W., Schreiber F. (eds.), *Modelling in systems biology - the Petri net approach*, London 2011.

Köhn S., Lengen R.H. van, Reis G., Bertram M., Hagen H., ´VES: Virtual Echocardiography System´, *Proceedings IASTED Visualization, imaging and image processing VIIP* (2004), 465-471.

Koshiba-Takeuchi K., Mori A.D., Kaynak B.L., Cebra-Thomas J., Susonnik T., Georges R.O., Latham S., Beck L., Henkelman R.M., Black B.L., Olson E.N., Wade J., Takeuchi J.K., Nemer M., Gilbert S.F., Bruneau B.G., 'Reptilian heart development and the molecular basis of cardiac chamber evolution', *Nature* 461 (2009), 95-98.

Krepska E., Bonzanni N., Feenstra A *et al.*, 'Design issues for qualitative modelling of biological cells with Petri nets', *LNBI* 5054 (2008), 48-62.

Lander A.D., Nie Q., Wan F.Y.M., 'Do Morphogen Gradients Arise by Diffusion?', *Developmental Cell* 2 (2002), 785–796.

Langer A., 'Über die Entwicklungsgeschichte des Bulbus cordis bei Amphibien und Reptilien', *Morphologisches Jahrbuch* 21 (1894), 40-67.

Lindenmayer A., 'Mathematical models for cellular interaction in development, Parts I and II', *Journal of Theoretical Biology* 18 (1968), 280-315.

Lopez D., Duran A.C., de Andrés A.V., Guerrero A., Blasco M., Sans-Coma V., 'Formation of cartilage in the heart of the Spanish terrapin, *Mauremys leprosa* (Reptilia, Chelonia)', *Journal of Morphology* 258 (2003), 97-105.

Lyson, T.R., Sperling, E.A., Heimberg, A.M. *et al.*, 'MicroRNAs support a turtle + lizard clade', *Biology Letters* 8 (2011), 104-107.

Maglia A.M., Leopold J.L., Pugener L.A., Gauch S., 'An anatomical ontology for amphibians', *Proceedings Pacific Symposium on Biocomputing* 12 (2007), 367-378.

Männer J., 'Cardiac looping in the chick embryo: a morphological review with special reference to terminological and biomechanical aspects of the looping process', *Anatomical Record* 259 (2000), 248-262.

Männer J. 'The anatomy of cardiac looping: a step towards the understanding of the morphogenesis of several forms of congenital cardiac malformations', *Clinical Anatomy* 22 (2009), 21-35.

Matsuno H., Murakami R., Yamane R., 'Boundary formation by Notch signaling in Drosophila multicellular systems: experimental observations and gene network modelling by genomic object net', *Pacific Symposium on Biocomputing* 8 (2003), 152-163.

McEntire R., 'Ontologies in the life sciences', *The Knowledge Engineering Review* 17:1 (2002), 77-80.

McNulty C.L., Peres J.N., Bardine N. *et al.*, 'Knockdown of the complete Hox paralogous group 1 leads to dramatic hindbrain and neural crest defects', *Development* 136 (2005), 2861-2867.

Meinhardt H., *Models of biological pattern formation*, London 1982.

Mejino Jr. J.L.V., Agoncillo A.V., Rickard K.L., Rosse C., 'Representing complexity in part-whole relationships within the Foundational Model of Anatomy', *Proceedings AMIA 2003 symposium* (2003), 450-454.

Menten, L., Michaelis, M.I., ´Die Kinetik der Invertinwirkung´, *Biochem Z* 49 (1913), 333–369.

Mierop L.H.S. van, Kutsche L.M., 'Comparative anatomy and embryology of the ventricles and arterial pole of the vertebrate heart', in: Nora J.J., Takao A., (eds.), *Congenital heart disease: causes and processes,* New York 1984.

Mierop L.H.S. van, Kutsche L.M., 'Some aspects of comparative anatomy of the heart', in: Johansen K., Burggren W.W. (eds.), *Cardiovascular shunts: phylogenetic, ontogenetic and clinical aspects, Alfred Benzon symposium*, Copenhagen 1985.

Mierop L.H.S. van, 'Morphological development of the heart' (1979), in: Berne R.M., Sperelakis N., Okamoto N., Akimoto N., Hidaka N. *et al.*, 'Formal genesis of the outflow tracts of the heart revisited: previous works in the light of recent observations', *Congenital Anomelies* 50 (2010), 141-158.

Montanari U., Rossi F., 'Contextual Nets', *Acta Informatica* 32 (1995), 545–596.

Nayak V.S., Raveen R., Gladstone M., 'Bicuspid evolution of the arterial and venous atrioventricular valves', *Journal of Heart Valve Disease* 4 (1995), 78-87.

Neumann, von, J., 'Theory of self-reproducing automata', *Urbana*, 1966 (edited and completed by Burks A.W.).

NIH (National Institutes of Health), 'Working definition bioinformatics and computational biology', 17 July 2000, BISTIC Definition Committee.

Noble D., 'Modelling the Heart - from Genes to Cells to the Whole Organ', *Science* 295 (2002), 1678-1682.

O'Donoghue C.H., 'The heart of the leathery turtle, *Dermochelys (Sphargis) coriacea*. With a note on the septum ventriculorum in the reptilian', *Journal of Anatomy* 52 (1918), 467-480.

Oostra R.J., Steding G., Lamers W.H., *Steding's and Viragh's scanning electron microscopy atlas of the developing human heart*, New York 2006.

O'Rahilly R., Müller F., *Developmental stages in human embryos*. Washington 1987.

Orts Llorca F., Puerta Fonella J., Sobrado J., 'The formation, septation and fate of the truncus arteriosus in man', *Journal of Anatomy* 134 (1982), 41-56.

Petri C.A., *Kommunikation mit Automaten,* Ph.D. Thesis, Bonn 1962.

Plath P.J., Plath J.K., Schwietering J., 'Collision patterns on mollusc shells', *Discrete Dynamics in Nature and Society* 1 (1997), 57-76.

Potikanond D., Belmamoune M., Welten M.C.M., Verbeek F.J., ´Visual integration of 3D digital atlas and 3D patterns of gene expression in zebrafish´, *International Symposium on Integrative Bioinformatics, Wageningen* (2011).

Potikanond D., Verbeek F.J., 'Visualization and analysis of 3D gene expression patterns in zebrafish using web services', *Proceedings SPIE* 8294:829412 (2012).

Priami C., Quaglia P., 'Modelling the dynamics of biosystems', *Briefings in Bioinformatics* 5:3 (2004), 259-269.

Priami C., 'Algorithmic systems biology', *Communication of the ACM* 52:5 (2009), 80-88.

Prusinkiewicz P., 'Visual models of morphogenesis', in: Langton C.G. (ed.), *Artificial life*, Cambridge MA 1995.

Qayym S.R., Webb S., Anderson R.H., Verbeek F.J., Brown N.A., Richardson M.K., 'Septation and valvar formation in the outflow tract of the embryonic chick heart', *Anatomical Record* 264 (2001), 273-283.

Quiring D.P., 'The development of the sino-atrial refion of the chick heart', *Journal of Morphology* 55 (1933), 81-118.

Rau A.S., 'Observations on the anatomy of the heart of *Tiliqua scincoides* and *Eunectes murinus*', *Journal of Anatomy* 59 (1924), 60-71.

Recker J., Indulska M., 'An ontology-based evaluation of process modelling with Petri nets', *Interoperability in Business Information Systems* 2:1 (2007), 45-64.

Reisig W., Rozenberg G. (Eds.), 'Lectures on Petri Nets I and II', *LNCS* 1491 and 1492 (1998).

Reisig W., 'Carl Adam Petri 1926-2010 - Visionaer und bedeutender Wissenschaftler´, *Informatik-Spektrum* 33:5 (2010), 514-521.

Rohr C., Marwan W., Heiner M., 'Snoopy - a unifying Petri net framework to investigate biomolecular networks', *Bioinformatics* 26:7 (2010), 974-975.

Rosse C., Mejino Jr. J.L.V., ´A reference guide for biomedical informatics: the Foundational Model of Anatomy´, *Journal of Biomedical Informatics* 36 (2003), 478-500.

Rosse C., Mejino Jr. J.L.V., 'The Foundational Model of Anatomy ontology', in: Burger A., Davidson D., Baldock R. (eds.), *Anatomy ontologies for bioinformatics: principles and practice,* New York 2008.

Roux-Rouquié M., Caritey N., Gaubert L. *et al.*, 'Using the Unified Modelling Language (UML) to guide the systemic description of biological processes and systems', *BioSystems* 75 (2004), 3-14.

Ruthensteiner B., Heß M., 'Embedding 3D models of biological specimens in PDF publications', *Microscopy Research and Technique* 71 (2008), 778-786.

Sayers E.W., Barrett T., Benson D.A., Bryant S.H., Canese K., Chetvernin V., Church D.M., DiCuccio M., Edgar R., Federhen S., Feolo M., Geer L.Y., Helmberg W., Kapustin Y., Landsman D., Lipman D.J., Madden T.L., Maglott D.R., Miller V., Mizrachi I., Ostell J., Pruitt K.D., Schuler G.D., Sequeira E., Sherry S.T., Shumway M., Sirotkin K., Souvorov A., Starchenko G., Tatusova T.A., Wagner L., Yaschenko E., Ye J., 'Database resources of the National Center for Biotechnology Information', *Nucleic Acids Research* 37 (2008), 5-15.

Schneider B.F., Norton S., 'Equivalent ages in rat, mouse and chick embryos', *Teratology* 19 (1979), 273-278.

Scholpp S., Brand M., 'Endocytosis controls spreading and effective signaling range of Fgf8 potein', *Current Biology* 14 (2004), 1834–1841.

Shaner R.F., 'Comparative development of the bulbus and ventricles of the vertebrate heart with special reference to Spitzer's theory of heart malformations', *Anatomical Record* 142 (1962), 519-529.

Sirin E., Parsia B., Grau B.C., Kalyanpur A., Katz Y., 'Pellet: a practical OWL-DL reasoner', *Journal of Web Semantics* 5:2 (2007).

Smith, B., Ashburner, M., Rosse C., Bard C., Bug W., Ceusters W., Goldberg L.J., Eilbeck K., Ireland A., Mungall C.J., The OBI Consortium, Leontis N., Rocca-Serra P., Ruttenberg A., Sansone S., Scheuermann R.H., Shah N., Whetzel P.L., Lewi S., ´The OBO foundry: coordinated evolution of ontologies to support biomedical data integration´, *Nature Biotechnology* 25:11 (2007), 2251-1255.

Staab S., Studer R. (eds.), *Handbook on ontologies*, Berlin 2004.

Steggles L.J., Banks R., Shaw O. *et al*., 'Qualitatively modelling and analysing genetic regulatory networks: a Petri net approach', *Bioinformatics* 23:3 (2007), 336-343.

Stern C.D., Charité J., Deschamps J. *et al.,* 'Head-tail patterning of the vertebrate embryo: one, two or many unresolved problems?', *International Journal of Developmental Biology* 50 (2006), 3-15.

Takai-Igarashi T., 'Ontology based standardization of Petri net modelling for signaling pathways', *In Silico Biology* 5 (2007).

Talcott C., Dill D.L., 'Multiple representations of biological processes', *Transactions on Computational Systems Biology* (2006).

Teleman A.A., Strigini M., Cohen S.M., 'Shaping morphogen gradients', *Cell* 105 (2001), 559–562.

Theiler K., *The house mouse: atlas of embryonic development*, New York 1989.

Tomlin C.J., Axelrod J.D., 'Biology by numbers: mathematical modelling in developmental biology', *Nature Reviews Genetics* 8 (2007), 331–340.

Turing A., 'The chemical basis of morphogenesis', *Philisophical Transactions of the Royal Society London* B237 (1952), 37-72.

Valk R., 'Object Petri nets using the nets-within-nets paradigm', *LNCS* 3098 (2004), 819-848.

Verbeek F.J., Huijsmans D.P., Baeten R.W.A.M., Schoutsen C.M., Lamers W.H., ´Design and implementation of a database and program for 3D reconstruction from serial sections: a data-driven approach´, *Microscopy Research and Technique* 30 (1995), 496-512.

Verbeek FJ, Huijsmans DP., ´A graphical database for 3-D reconstruction supporting (4) different geometrical representations´, in: Wong S.C.T. (ed.), *Medical Image Databases,* Boston 1998.

Verbeek F.J., 'Theory and practice of 3D-reconstructions from serial secions', in: Baldock R.A., Graham J., (eds.), *Image processing, a practical approach*, Oxford 1999.

Verbeek F.J., Boon P.J., 'Three-dimensional and multidimensional microscopy: image acquisition and processing IX', *Proceedings of SPIE* 4621 (2002), 64-76.

Verbeek F.J., Boon P.J., Sloetjes H., Velde R. van der, de Vos N., 'Visualization of complex data sets over Internet: 2D and 3D visualization of the 3D digital atlas of zebrafish development', *Proceedings of SPIE* 4672: Internet Imaging III (2002), 20-29.

Vogler W., 'Partial order semantics and read arcs', *Theoretical Computer Science* 286 (2002), 33–63.

Webb G., Heatwole H., De Bavay J., 'Comparative cardiac anatomy of the Reptilia, I. The chambers and septa of the varanid ventricle', *Journal of Morphology* 134 (1971), 335-350.

Webb S., Brown N.A., Anderson R.H., Richardson M.K., 'Relationship in the chick of the developing pulmonary vein to the embryonic systemic venous sinus', *Anatomical Record* 259 (2000), 67-75.

Webb S., Kanani M., Anderson R.H., Richardson M.K., Brown N.A., 'Development of the human pulmonary vein and its incorporation in the morphologically left atrium', *Cardiology in the Young* 11 (2001), 632-642.

Webb S., Qayyum S.R., Anderson R.H., Lamers W.H., Richardson M.K., 'Septation and separation within the outflow tract of the developing heart', *Journal of Anatomy* 202 (2003), 327-342.

White F.N., 'Functional anatomy of the heart of reptiles', *American Zoologist* 8 (1968), 211-219.

Wolpert L., *Principles of development*, Oxford 2002.

Wong B.A., Rosse C., Brinkley J.F., 'Semi-automatic scene generation using the Digital Anatomist Foundational Model', *Proceedings American medical informatics association fall symposium* (1999), 637-41.

Yntema C.L., 'A series of stages in the embryonic development of Chelydra serpentina', *Journal of Morphology* 125 (1968), 219-252.

Yu S.R., Burkhardt M., Nowak M. *et al.*, 'Fgf8 morphogen gradient forms by a source-sink mechanism with freely diffusing molecules', *Nature* 461 (2009), 533-537.

Zhang, Y., Verbeek, F.J., 'Comparison and integration of target prediction algorithms for microRNA studies', *Journal of Integrative Bioinformatics* 7:3 (2010), 127.

WEBSITES

Amira:
www.amira.com

BioVis3D:
www.biovis3d.com

Matlab:
www.mathworks.nl/products/matlab/?s_cid=global_nav

Mimics:
www.materialise.com/mimics

Molecular Movies:
www.molecularmovies.com

OWL:
https://www.w3.org/TR/owl-features

OWL 2 Web Ontology Language Primer, class hierarchies (accessed on 29-04-2011):
http://www.w3.org/TR/2009/REC-owl2-primer-20091027/#Class_Hierarchies

OWL 2 Web Ontology Language Primer, property hierarchies (accessed on 29-04-2011):
http://www.w3.org/TR/2009/REC-owl2-primer-20091027/#Property_Hierarchies

SWRL: a Semantic Web Rule Language combining OWL and RuleML:
http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/

UML:
www.uml.org

SAMENVATTING

Modelleren is van oudsher een essentieel onderdeel van biologisch onderzoek. Versimpelingen maken de geobserveerde realiteit inzichtelijk en ondersteunen theorievorming. Simulatie en analyse van modellen leiden tot nieuwe inzichten in biologische processen en vormen zo een belangrijke aanvulling op laboratorium-experimenten. Diverse modelleermethoden hebben toepassingsmogelijkheden in de levenswetenschappen en het groeiende belang van computationeel onderzoek, gedurende de afgelopen decennia, heeft ook op dit vakgebied grote invloed gehad.

In dit onderzoek worden verscheidene modelleertechnieken voor het onderzoeksveld van de ontwikkelingsbiologie behandeld, met speciale aandacht voor de integratie van deze methoden. Steeds ligt hierbij de nadruk op het computationele aspect van het modelleren. Hiertoe wordt in het **eerste hoofdstuk** een overzicht geboden van relevante eigenschappen van modelleertechnieken, *e.g.*statisch vs. dynamisch en kwalitatief vs. kwantitatief. Dit wordt gevolgd door een overzicht van modelleermethoden, verdeeld in vier categorieën: verbale en visuele methoden (beide van nature niet-computationeel) en algoritmische en wiskundige methoden (van nature computationeel). Dit hoofdstuk vormt een theoretische basis voor de hierop volgende hoofdstukken, waarin een serie *case studies* aan bod komt.

In het **tweede hoofdstuk** wordt een studie gepresenteerd naar de hartontwikkeling van de schildpad *Emys orbicularis*. Op grond van histologische secties van embryo's in verschillende ontwikkelingsstadia is een serie 3D-reconstructies gemaakt, die de embryonale ontwikkeling van het hart illustreert. Deze modellen passen binnen de categorie van visuele modellen, besproken in het eerste hoofdstuk. Het ventrikel van het schildpadhart bestaat uit drie *cava* die gedeeltelijk van elkaar worden gescheiden door twee septa, het horizontale en het verticale septum. De 3D modellen maken duidelijk hoe het horizontale septum wordt gevormd door een uniek proces van *heart looping*; de proximale component van de *outflow tract* komt ventraal voor het ventrikel te liggen en de wanden van de twee structuren fuseren en vormen zo het septum. De modellen zijn toegankelijk gemaakt op http://bio-imaging.liacs.nl/galleries/; hier is een TDR-viewer beschikbaar om de modellen visueel te bestuderen.

Het **derde hoofdstuk** behandelt een vervolgstudie naar het hart van de schildpad, waarin de aanpak uit het tweede hoofdstuk verder wordt uitgewerkt. Hierin wordt de ontwikkeling van de *outflow tract* nader onderzocht, aan de hand van vijf 3D-reconstructies. De reconstructies hebben een hogere resolutie dan de eerder besproken modellen en bieden daarom een gedetailleerder beeld van de ontwikkeling van de *outflow tract*, de endocardiale kussens en het aortopulmonaire septum in de *tract.* De

modellen onderschrijven de eerdere bevinding dat het *cavum pulmonale* in het volwassen hart een functioneel onderdeel vormt van het ventrikel, maar in de embryonale ontwikkeling deel uitmaakt van de *outflow tract*. De posities van de endocardiale kussens bevestigen dit. Dit ontwikkelingsproces heeft twee belangrijke consequenties: 1) het proximale deel van de *outflow tract* raakt niet opgedeeld, zoals in zoogdieren en vogels; 2) de endocardiale kussens in het distale deel van het ventrikel ondergaan geen actieve verschuiving binnen de *outflow tract*, maar komen door de *heart looping* op de overgang tussen *cavum pulmonale* en *outflow tract* te liggen. Dit representeert een nog niet eerder beschreven ontwikkeling van het vertebrate hart. Het gebruik van 3D modellen is cruciaal voor het begrip van dit proces, aangezien foto´s en schematische beelden in 2D het complexe karakter hiervan niet goed kunnen weergeven. De studie onderscheidt zich bovendien van de studie in het vorige hoofdstuk door de kleinere schaal van de bestudeerde structuren; het modelleren op hoge resolutie is daarom bij uitstek geschikt voor dit onderzoek.

Na de visuele modellen in de twee voorafgaande hoofdstukken, wordt in het **vierde hoofdstuk** een semantisch model gepresenteerd. Het betreft een nieuwe benadering van het gebruik van ontologieën voor ontwikkelingsbiologie: een generiek ontologiesysteem van kennis over de anatomie, ontwikkeling en taxonomie van het vertebrate hart. Drie losse ontologieën zijn gebruikt voor de verschillende typen informatie, deze zijn onderling verbonden en informatie over specifieke vertebrate soorten is aan het systeem toegevoegd; hierdoor is een generiek en makkelijk uitbreidbaar systeem gecreëerd. Dit kan gebruikt worden voor vergelijkende studies naar vertebrate hartanatomie en -ontwikkeling. De methode is hier toegepast op het hart, maar is op gelijke wijze bruikbaar voor andere organen en anatomische structuren. De functionaliteit van het systeem wordt geïllustreerd aan de hand van vier *case studies*, aangaande soortenvergelijking, ontwikkeling, physiologie en 3D visualisatie.

Het **vijfde hoofdstuk** is gewijd aan procesmodellering door middel van de algoritmische modelleertechniek van Petri netten. Hiermee wordt een proces beschreven in een diagram, bestaande uit passieve entiteiten (plaatsen) en actieve gebeurtenissen (transities). Afgezien van de visuele inzichtelijkheid van de modellen, is het hiermee ook mogelijk kwantitatieve analyses uit te voeren; dit maakt de methode bijzonder geschikt voor het modelleren van biologische ontwikkelingsprocessen. In dit hoofdstuk wordt een generiek model gepresenteerd van biologische gradiëntvorming, waarin plaatsen de biologische cellen representeren en transities het transport van morphogene moleculen tussen deze cellen. Door het model nauw te baseren op het biologische proces is het model makkelijk schaalbaar, algemeen toepasbaar en zowel in de eindstaat als in de tussenliggende staten in overeenstemming met de biologische situatie.

In een vervolgstudie op het hierboven beschreven onderzoek is het model van gradiëntvorming uitgebreid. Het **zesde hoofdstuk** bespreekt een nieuwe methode voor het modelleren van dit proces, waarbij de categorieën van algoritmische en wiskundige modelsoorten zijn gecombineerd. Gradiënten worden dikwijls met differentiaal-vergelijkingen beschreven en door aspecten van deze vergelijkingen in het Petri net op te nemen wordt de kwantitatieve kracht van het model groter, terwijl het tegelijkertijd de intuïtieve overzichtelijkheid van het visuele Petri net-diagram behoudt. Voor de validatie van het model zijn concentratiedata uit literatuur, verkregen aan de hand van differentiaalvergelijkingen, vergeleken met resultaten uit het Petri net-model; het model komt overeen met de bestaande wiskundige modellen en daarmee met de gemodelleerde experimentele observaties.

Elk van de *cases studies* uit de vorige hoofdstukken was gericht op het optimaliseren en innoveren van modelleermethoden voor ontwikkelingsbiologische vraagstukken. Hiernaast is ook aandacht besteed aan het integreren van modelleermethoden, om zo de functionaliteit van de modellen te vergroten. Het **zevende hoofdstuk** geeft een overzicht van elk van de mogelijke combinaties van modelleertypen die in dit onderzoek aan de orde zijn gekomen.

PUBLICATIONS

Bertens, L.M.F., Slob, J., Verbeek, F.J., ´A generic organ based ontology system, applied to vertebrate heart anatomy, development and physiology´, *Journal of Integrative Bioinformatics* 8:2 (2011), 167.

Bertens L.M.F, Kleijn J., Koutny M., Verbeek F.J., 'Modelling gradients using Petri nets', in: *Proceedings International Workshop on Biological Processes & Petri Nets (BioPPN) Braga, Portugal, June 21 2010*, 55-69.

Bertens L.M.F., Richardson M.K., Verbeek F.J., 'Analysis of cardiac development in the turtle *Emys orbicularis* (Testudines: emidydae) using 3-D computer modelling from histological sections', *Anatomical Record* 293:7 (2010), 1104-1114.

Yan, K., Bertens, L.M.F., Verbeek, F.J., ´Image registration and realignment using evolutionary algorithms with high resolution 3D model from human liver´, *Proceedings 11th IASTED International Conference on Computer Graphics and Imaging* (2010), track 679-014.

Bertens, L.M.F., Jansen, H.J., Kleijn, J., Koutny, M., Verbeek, F.J., 'Using Petri nets in higher level developmental biology: a case study on the AP axis development in Xenopus laevis', *LIACS Technical Report, 2009-03* (2009).

Brittijn, S.A., Duivesteijn, S.J., Belmamoune, M., Bertens, L.M.F., Bitter, W. *et al.*, ´Zebrafish development and regeneration: new tools for biomedical research´, *International Journal of Developmental Biology* 53:5-6 (2009), 835-850.

Putten, P.W.H. van der, Bertens, L.M.F., Liu, J., Hagen, F, Boekhout, T., Verbeek, F.J., 'Classification of Yeast Cells from Image Features to Evaluate Pathogen Conditions', *Proceedings of SPIE* 6506 (2007), 650601-14.

CURRICULUM VITAE

Laura Bertens was born on May 28 1981, in Zeist, the Netherlands. She studied Classics at Leiden University and completed her Master's degree in May 2007, with a thesis on aspects of the archaeology of Classical Greek religion. Starting in 2001, she studied Biology at Leiden University and in March 2007 she finished her Master's degree with a specialisation in Bio-informatics and Developmental Biology. Her Master's projects focused on the 3D computer visualisation of heart development of the turtle *Emys orbicularis* and the visualisation of processes in the immune system of the zebrafish embryo, using *in situ hybridisation*.

In September 2007 she started her PhD in the research group Imaging and Bioinformatics, part of the Leiden Institute of Advanced Computer Science (LIACS) at Leiden University, under the supervision of Dr. Ir. Fons Verbeek. During this period she also studied Art History at Leiden University. She is currently finishing her Bachelor's degree in Art History and is employed in a research project, concerning the theoretical and practical application of computer ontologies to canons of art history and the humanities.

ACKNOWLEDGEMENTS

I have been supported and helped by more people than I can mention here. I would like to explicitly thank a few.

My parents and brothers, who support and share my interest in both sciences and humanities; with your combined knowledge and love of all things academic, you have helped me solve many of my scientific problems (often by not saying anything).

My research group mates, specifically my long-time colleagues Joris and Amalia, thanks for all the computer support, Monty Python quotes and film nights, in addition to the tremendous help with my research.

Thanks to the Computer Science students who contributed to the Petri net research, in particular Erik Zandvliet and Erik, Jaron and Wouter. I hope your future research will be live, unbounded and conflict-free.

My closest friends, in particular my paranymphs Esther and Floor, who have witnessed both the beginning and the completion of my PhD; thanks for helping me in all areas of life, not in the least my research. Thanks for vrijmibos, endless amounts of tea and pasta salad and Mies and Teun.

My fellow classicists, thanks for reminding me, over coffee and dinner in Leiden and Berlin, that Classical antiquity can be as relevant as state-of-the-art computer analysis.

My fellow art historians, thanks for the continuous interest in what must have been for the most part an unintelligible topic. Thanks for library sessions, Berlin and the opportunity to think outside the borders of predefined scientific fields and integrate my research interests.

Merci les mecs, pour le LFL et les petits déjeuners réconfortants sur les quais de Paris et de Genève.

And thanks to all people not mentioned here, who nonetheless played an important role in completing this work, by either helping me with the content or with the peripheral support. It is true what they say: dimmi con chi vai e ti dirò chi sei.