Biosimulation of water and salt balance

-from minutes to days

Larsen, Peter; Mosekilde, Erik

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Biosimulation of water and salt balance

- from minutes to days

Peter Larsen

PhD thesis

Danmarks Tekniske Universitet

December 1, 2009

Supervisor
Erik Mosekilde
DTU Fysik
Danmarks Tekniske Universitet

Preface

This thesis documents the research performed by Peter Larsen at the Department of Physics at The Technical University of Denmark from September 2006 to September 2009. The thesis is submitted as partial fulfilment of the requirements for the degree of philosophiae doctor. The work was conducted under the supervision of Professor Erik Mosekilde.

The thesis consists of several studies all of which are associated with the general topic *Biosimulation of water and salt balance*. None of the studies are confined to a classical scientific discipline, and they are carried out in the spirit of interdisciplinarity with focus on mathematical analyses of biological systems. The work could not have been done without the help and assistance from a number of people.

My home base during the 3 years was my office in building 309 at DTU. To my supervisor Erik Mosekilde, I am grateful for many interesting discussions on modelling topics and the generality of dynamical phenomena. These discussions have given me the courage to overcome borders between disciplines. A special thanks goes to my PhD-colleagues Christian Bentzen, Wiktor Mazin, and Jakob Laugesen for keeping us afloat through the troubled waters of PhD-research. And I would also like to thank all of my colleagues at the Department of Physics for being the kind and helpful people you are.

During the spring of 2008, I spent 6 months with the group of Professor Hanspeter Herzel at the Institute of Theoretical Biology, Humboldt University in Berlin, Germany. This was a truly inspiring experience. I would like to thank everyone in Hanspeter's group ITB, and especially Hanspeter himself for introducing me to the exciting topic of circadian rhythms. Like Erik, he also made me confident that it is an enriching experience to travel on the many bridges between mathematics, physics, and biology.

In the autumn of 2009, when I was on the look out for experimental

data, I was lucky to establish a contact with the group of Professor Jens Christian Djurhuus at Skejby Hospital in Aarhus. They kindly provided me with an interesting dataset from a clinical study. I am very grateful for that, and I will like to extend my special thanks to Jens Christian Djurhuus, Charlotte Graugaard-Jensen, and Frank Høj. Had it not been for the temporal limits of the project, I would have loved to dive further into their stock of interesting clinical data.

I would also like to thank Lars Erichsen at Ferring Pharmaceuticals A/S for discussions on PK/PD models of ddAVP function. Unfortunately, our collaboration did not make it all the way to the finish line, but the discussion served as inspiration for other parts of the work.

Finally, I would like to thank my family for their support. And last, but still first and foremost, I admire and I am deeply indebted to my wife, Pia, for being a loving, caring and stable support for me throughout the whole period. And thanks to the extension of my deadline, I am proud to say thank you to my little daughter, Freja. She truly makes you aware that life is indeed beautiful.

Peter Larsen

Abstract

Biosimulation of water and salt balance is the topic of this report. Five studies approach the central theme from different angles. The first study focuses on the dynamic regulation of membrane proteins with specific attention to the insertion of aquaporins in the membrane of kidney collecting duct cells. Supported by a dynamical model, we explore the parameter space of exocytosis, endocytosis and transcriptional regulation by the hormone vasopressin. The analysis supports the hypothesis of regulation of both exo- and endocytotic rates during hormonal stimulation. In the second study, we focus on the regulation of water and solute permeability in the kidney. A model of the collecting duct clearly identifies the different roles of water and solute reabsorption, and that the two can be separated by the biophysical reabsorption processes. The third study extends the modelling framework to the whole organism, and we formulate a model of body fluid and solutes, and its regulation by vasopressin. The model successfully replicates results from a simple experiment of water drinking. In the fourth study, we turn the attention to the circadian rhythms of physiological variables in two groups of subjects exposed to different fluid intakes. In spite of different fluid intake, vasopressin levels are not affected, whereas atrial natriuretic peptide levels show larger dispersion during high fluid intake. Using a Fourier score technique, we see indications of disruption of circadian rhythms in plasma potassium levels during high fluid intake. The fifth study also focuses on circadian rhythms. In this study, we look for circadian transcription factors associated with a list of blood pressure regulating genes. The analysis identifies 11 specific transcription factors, out of which there are 5 candidates for circadian transcription factors. The modelling work demonstrates the validity of a mechanism-based approach, and the data analyses with focus on circadian rhythms reveal interesting results and motivate further investigations in the same direction.

Resumé

Biosimulering af væske- og saltbalancen er det gennemgående emne i denne rapport. I fem studier angribes emnet fra forskellige vinkler. Det første studie fokuserer på dynamisk regulering af aquaporiner i membranen på samlerørsceller i nyrerne. Med udgangspunkt i en dynamisk model udforskes parameterrummet for exocytose, endocytose og transkriptionel regulering af hormonet vasopressin. Analysen understøtter hypotesen om samtidig regulering af exo- og endocytose ved hormonel stimulering med vasopressin. I det andet studie fokuseres på regulering af vand- og saltreabsorption i nyrerne. En model af samlerøret identificerer, hvordan vand- og saltreabsorption spiller sammen. Det tredje studie introducerer en model af organismens regulering af væske og salt og respons på væskeindtag via vasopressin. Modellen reproducerer eksperimentelle resultater af væskeindtag på baggrund af uafhængigt estimerede parametre. I det fjerde studie undersøges circadiane rytmer i fysiologiske variable i to grupper af individer udsat for forskellige væskeindtag. På trods af forskelligt væskeindtag er vasopressinniveauet det samme i de to grupper, men atrial natriuretisk peptid niveauer har større dispersion ved højt væskeindtag. Ved hjælp af en Fourier score metode ses indikationer på at den circadiane rytme i kalium plasma koncentrationen forstyrres ved højt væskeindtag. Det femte studie fokuserer ligeledes på circadiane rytmer. Det undersøges hvorvidt circadiane transkriptionsfaktorer er overrepræsenterede i en liste af blodtryksregulerende gener. Analysen identificerer 11 transkriptionsfaktorer, hvoraf 5 er kandidater som circadiane transkriptionsfaktorer. Modelleringsarbejdet eksemplificerer værdien af en mekanisme-baseret tilgang til modellering, og dataanalyserne med fokus på circadiane rytmer afslører interessante resultater og motiverer videre undersøgelser af samspillet med døgnrytmer på tværs af biologiske niveauer.

Contents

Chapter 1

Introduction

Biosimulation is a true interdisciplinary challenge. The barycentre is of course a biological problem, and around it evolves a myriad of analytic, statistical and experimental techniques that all contribute to the description, the quantification and, ultimately, the understanding of the underlying system and its mechanisms. In the current work, the biological problem is water and salt balance of the human organism. The computational approaches are based on a wide variety of experimental data.

One focal point has been the construction of mechanism-based models. The point of departure of this modelling approach is the causal description of the functional mechanism of the system. The models in this work represent an effort to use models as a quantitative tool that integrates background information at various levels. This is not confined to the specific modelling technique as such, but rather to the scope and boundaries of the problem itself. The models provide a methodology to quantitatively combine many sources of information, and they may be used to generate new hypotheses.

But models are not reality. Models are just models. Models are simplified, schematic representations of a system that address problems in time and space. It is a mathematical description that relies on established relations between variables and parameters. Behind each variable and parameter lies a rich volume of scientific knowledge. A model is not a fixed representation, but rather a flexible entity that can be transformed to answer new questions that may arise.

The construction of physiological and biological models is cumbersome and extremely time consuming. In contrast to *e.g.* physics, there is a lack of well-defined and unquestioned, quantitative laws of nature within biology. Every formulation is and should be doubted. Empirical evidence is also

not as strong as in other disciplines. Biological systems are also inherently multi-scale, since living organisms are dependent on the tiniest details at atomic, molecular, cellular levels, and at the same time we know that they do work at larger scales of whole organism and even at ecological and social scales. So the question of which level of organisation that we want to model is absolutely crucial.

Likewise, temporal organisation of biological systems is extremely complex. Consider how blood pressure is regulated. A sudden change from lying down to standing up will require a fast stabilisation of blood pressure that reacts within minutes, and one might ask how this is linked to the daily variation of blood pressure that assures a lower blood pressure during nighttime or even the decade-long development of the serious condition of hypertension. This example illustrates the need for identification of the specific temporal scale in consideration and the need for extracting temporal information. Computation techniques that serve this purpose are also crucial to biosimulation. Often scientists express the opinion that new knowledge is the same as the new data. In my opinion, new data analysis techniques are just as important as new data.

1.1 The origin of thesis

From minutes to days - that is the temporal scale of the current work. From genes to the whole organism - that is the spatial scale. This wide span of scales in time and space is a consequence of many choices made during the last three years. Initially motivated by the modelling work of Erik Mosekilde and others in modelling of nephron autoregulation of blood flow, the kidneys became a central part of the work [11, 53, 87, 103, 104, 105]. Through dialogue with the pharmaceutical company Ferring a need for an improved understanding of the dynamics of hormonal control mechanisms of the kidney became apparent. Ferring produces a variant of *antidiuretic hormone*, the most important hormone in urine flow regulation. This is an important drug for many patients of all age groups that suffer from involuntary urination, especially during the night. This night time lack of control of urination may be caused by a disruption of a natural daily rhythm of antidiuretic hormone secretion. This daily rhythm is called the *circadian rhythm*. It is a fascinating, biological phenomenon that spans all levels of biological organisation. In order to gain a deeper insight, I decided to consider not just urine output regulation and kidney function, but the whole physiological system of water and salt balance on time scales from the acute regulation of urine flow to the circadian oscillations.

1.2 The scope and the goals

The current work focuses on applying the ideas of biosimulation to gain a deeper quantitative understanding of water and salt balance. The scope of the work is the use of computational methods from physiologically-based models to bioinformatics. Relevant data sources are basically anything from DNA sequences to urine samples. The overall goals are:

- Formulation of a dynamical, mechanism-based model of water and salt balance.
- Investigation of various data sources with focus on dynamical properties.

These goals are not particularly concrete, and the more specific challenges that I chose to face were:

- Modelling of hormonal control of kidney membrane properties.
- Modelling of urine flow regulation by the kidney.
- Modelling of body fluid and solute contents at the organism level.
- Analysis of circadian rhythms in clinical data.
- Analysis of circadian rhythms in genomic data.

Through a series of studies, all presented in this dissertation, I will cover all of the above challenges.

1.3 Overview of the thesis

The thesis consists of 10 chapters in total. A short physiological background chapter follows the present introductory chapter. Chapter 3 is a presentation of a model of the regulation of membrane proteins with focus on aquaporin regulation by vasopressin. In Chapter 4 a model of kidney output regulation is formulated and analysed. The models from Chapters 3 and 4 are integrated into a systemic model of urine flow regulation in Chapter 5. The topic of circadian rhythms is introduced in Chapter 6 in combination with a literature study on the evidence for circadian rhythms in water and salt balance. Chapter 7 is an analysis of a clinical dataset from two different groups of subjects with different fluid intake. The data was kindly provided by the group of Jens Christian Djurhuus, at Skejby

Hospital, Denmark. Chapter 8 is a study of the possible role of specific circadian transcription factors in the regulation of blood pressure. This study was performed in collaboration with Professor Hanspeter Herzel, Humboldt University, Berlin. The last two chapters are devoted to a discussion and a final conclusion.

Chapter 2

A brief introduction to water and salt balance

This chapter provides a brief introduction to the physiology of water and salt balance. It serves as a general introduction to the different studies in this work. The physiological topic of water and salt balance is immensely complex. The amount of substances, tissues and genes that are involved in this system is huge. The composition of body fluids is essential to the functioning of physiological mechanisms, and therefore, the organism must be able to maintain fluid composition and volume. Every cell of the organism relies on this regulatory system that faces the challenge of optimising performance under various external conditions. Receptors that monitor fluid volume, blood pressure, plasma osmolarity, and concentrations of specific plasma electrolytes are located throughout the body. The signals from the receptors are usually sent to secretory organs such as the pituitary and adrenal glands, where hormones are released in to the blood plasma. The hormones carry out the desired action at target sites of which some of the most important are located in the kidney.

The starting point for considering water and salt balance will be the causal loop diagram in Figure 2.1. Full arrows indicate stimulation or positive interaction, and dashed arrows indicate inhibition or negative interaction. The input variables to the system, in the form of water and solute intake, are shown at the top in the grey boxes. They feed forward into the main variables of the system, the blood volume, blood osmolarity and blood pressure. Water intake increases blood volume, which in turn dilutes the blood and decreases osmolarity, while large amounts of water causes pressure to increase. Changes in the blood variables stimulate the release

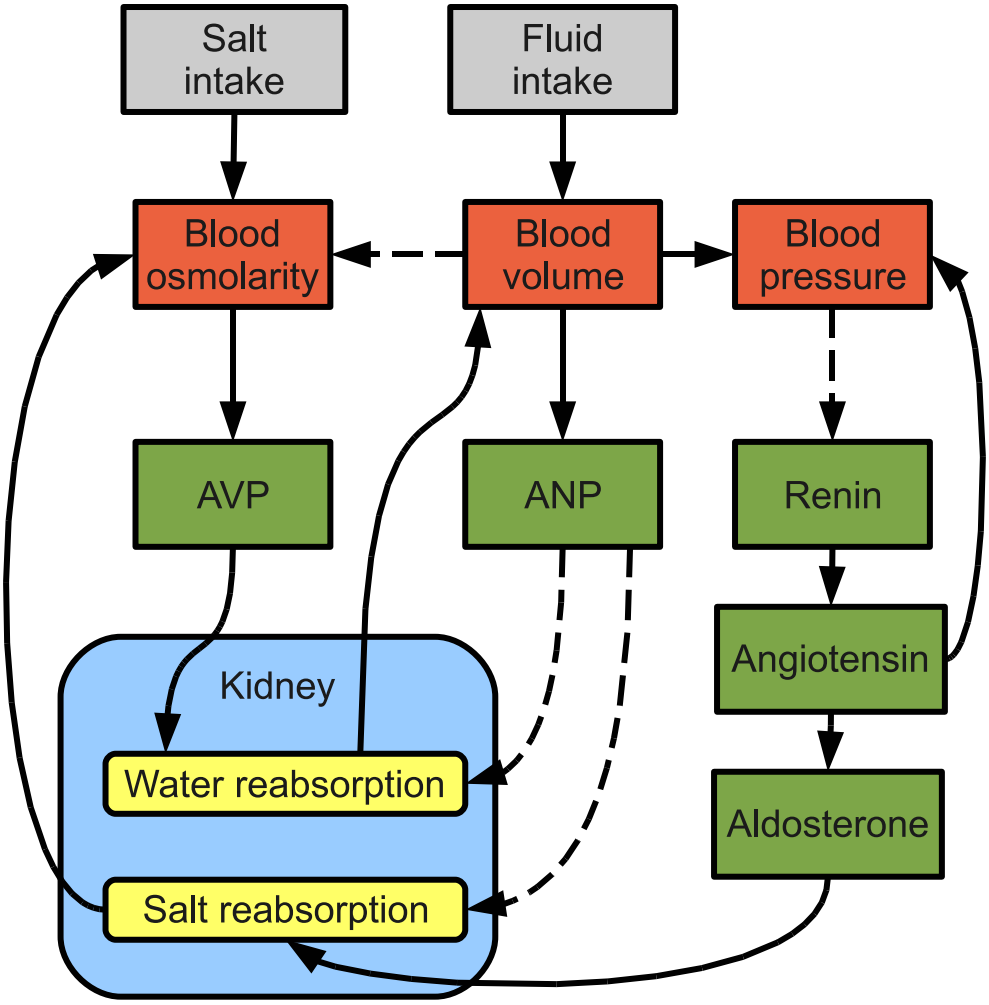


Figure 2.1: Causal loop diagram of water and salt balance. In response to fluid intake the organism adapts the excretatory behaviour of the kidney. The full arrows indicate a positive or excitatory interaction, and the dashed arrows describe a negative or inhibitory one. Hormones serve as regulatory substances that conduct signals in the feedback loops in the control system. Vasopressin (AVP) is mainly responsive to plasma osmolarity changes, atrial natriuretic peptide to volume changes, and the renin-angiotensin-aldosterone system responds to changes in blood pressure. The control mechanisms represented in this causal loop diagram represent the mutual interaction of at least 4 feedback regulations.

of hormones that will counteract the stimulatory change, and thereby constitute a negative feedback loop. The next section introduces the hormonal feedback systems.

2.1 Hormones

Three hormonal systems are shown in the green boxes. In general such hormonal controls are mediators in a negative feedback loop. A negative feedback loop is a mechanism that is activated by a stimulus or perturbation and through a series of reactions counteracts the stimulus.

From the left, it is arginine vasopressin, AVP, also known as antidiuretic hormone, ADH. These two names will be used interchangeably throughout this report. Increases in osmolarity is sensed by osmoreceptors, and causes a stimulation of AVP secretion from the posterior pituitary gland. AVP is a small peptide molecule with a short life time, and it quickly stimulates water reabsorption in the kidney, and hence counteracts the increase in osmolarity. Another effect of AVP in the kidney is the reabsorption of urea.

The second hormone is atrial natriuretic peptide ANP. This hormone is released in response to increases in blood volume, sensed by volume receptors in the atria. It inhibits both water and salt reabsorption in the kidney, and hence increases urine flow of water and solutes. Again a negative feedback mechanism that counteracts the stimulatory signal.

The renin-angiotensin-aldosterone system is shown on the right hand side of the figure. Renin is a catalytic molecule that is released in response to decrease in blood pressure. Renin initiates a cascade of enzymatic reactions that cleave angiotensinogen to angiotensin II. Angiotensin II is a strong vasoconstrictor and it causes an increase in blood pressure - another feedback loop. Furthermore, angiotensin II is a stimulator of release of the hormone aldosterone. Aldosterone is also known as a mineralocorticoid hormone, since it affects the regulation of minerals, in particular sodium and potassium. Aldosterone increases sodium reabsorption in the kidney, and therefore it contributes to the conservation of extracellular fluid volume and pressure.

The kidney is shown as a blue box in Figure 2.1. It is the most important organ in the control of water and salt balance, and it is a target organ of the hormones discussed above. The next section is a short introduction to kidney function.

2.2 The kidney

The kidney plays a very important role in maintaining water and salt balance. The kidney filters blood plasma and excretes waste products through the urine. Control mechanisms allow the kidney to control the composition and volume of the urine. According to the physiological conditions the urine can be very dilute (low osmolarity) or highly concentrated (high osmolarity). Antidiuretic hormone (ADH) enhances the reabsorption of filtered fluid, thereby controlling the amount of fluid excreted. As the name predicts, it has an *antidiuretic* effect. Another hormone, aldosterone, increases the reabsorption of sodium. The presence of aldosterone makes the urine more dilute - less concentrated.

The nephron is the functional unit of the kidney. A human kidney consists of about 1 million nephrons. The nephron contains several functionally distinctive segments. A nephron is presented in Figure 2.2. Blood enters the glomerulus from the afferent arteriole. In the glomerulus the blood is filtered and it flows through the tubular structure of the nephron. Some of the blood continues to flow through the complex vascular structure of the kidney. The first part of the nephron is the proximal tubule, where about 70 % of the tubular fluid is reabsorbed. The reabsorption is isosmotic, meaning that the proximal tubular fluid has the same osmolarity as the surrounding environment. It then enters the hairpin structure called the loop of Henle. The combination of counter-current flow and active transport in the ascending limb causes an osmotic gradient to be maintained along the length of the loop. From the loop of Henle, the fluid flows past the specialised macula densa cells (green). These cells control the flow in the afferent arteriole, thus regulating the blood supply to the glomerulus.

The distal tubule and the collecting duct are the distal part of the nephron, as opposed to the proximal part. In the distal part, many target sites for hormones are located. For example, the membrane bound V_2-receptor for vasopressin in the distal tubule and the principal cells of the collecting duct. Binding of vasopressin to this receptor, initiates the insertion of specific water channels in the cell membrane and allows for the reabsorption of water. This mechanism is the objective of the model presented in Chapter 3. The mineralocorticoid receptor (MR) for aldosterone is also found in the principal cells of the collecting duct. This receptor is a nuclear receptor located intracellularly, and the effect of binding is an increase in membrane abundance of sodium channels.

The filtration rate is very high - approximately 180 l/day. If we excreted the whole total volume filtered in the glomerulus, we would very quickly die of dehydration. Under normal physiological conditions 98-99 %

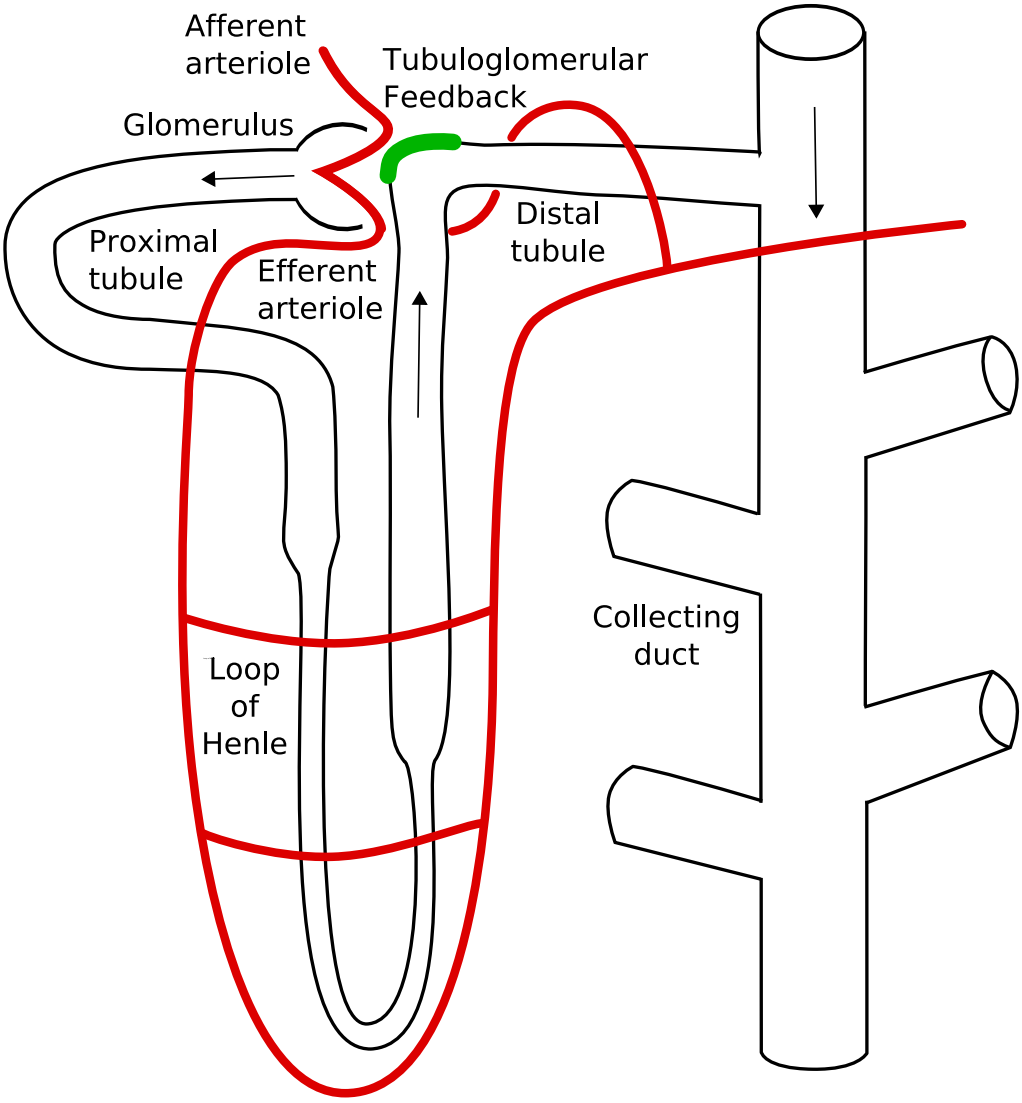


Figure 2.2: Schematic drawing of a nephron - the functional unit of the kidney. Filtration occurs in the glomerulus, and the fluid flows through the proximal tubule and the loop of Henle. Just before the distal tubule, tubuloglomerular feedback mechanism controls the blood supply to the glomerulus. The distal part of the nephron is the distal tubule and the collecting duct. In the distal part the fine tuning of the urine composition and flow occurs through hormonal control mechanisms.

of the fluid is reabsorbed. The kidney must be able to control reabsorption of both water and solutes, and in doing so the urine osmolarity may vary from 50 mOsm/l to 1200 mOsm/l in humans [50]. In other species, especially rodents, the maximum concentrating capacity may be even up to 7600 mOsm/l as the case of the chinchilla [2].

2.3 Summary

The physiology of water and salt balance is very complex. Several hormones act as the active substances of the regulatory feedback loops that control the composition of urine and body fluid. The kidney is the most important organ in water and salt balance regulation, and many interacting biophysical and -chemical phenomena interact to make this organ very efficient at fine-tuning the composition of urine. Each of the following chapters will refer to the diagram in Figure 2.1.

Chapter 3

Regulation of membrane proteins

An important aspect of diuretic regulation is the hormonal control of membrane transport properties of renal tubules. One very important example of this is the regulation of water permeability in collecting duct principal cells by antidiuretic hormone. This process is the main regulator of diuresis, and malfunctioning of this mechanism may cause serious pathologic conditions. The process is highly complex and involves many intermediate steps, and it is still disputed how aquaporin translocation to the apical membrane functions and how it is regulated. The arrow in Figure 3.1 indicates where in the diagram of water and salt balance aquaporin regulation fits in. It is a small part of the big picture, but very important to the functioning of the overall system.

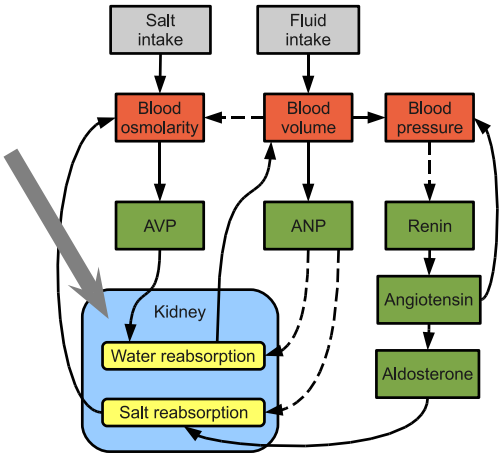


Figure 3.1: The causal loop diagram of water and salt balance presented in Figure 2.1. The problem treated in this chapter deals with the regulation of water reabsorption indicated by the gray arrow. The water transport molecules aquaporins are activated by the hormone vasopressin to decrease urine production by increasing the reabsorption of water.

In this chapter, I will first introduce the cellular mechanisms of aquaporin dynamics. In section 3.1, I will present and discuss an experimental and modelling study by Knepper and Nielsen [63]. I will develop a model that incorporates more aspects of aquaporin regulation, and I will examine the effects of model parameters on the dynamics of the model. Even with insufficient information about system parameters, I will demonstrate how we can investigate the system and gain insight into the dynamic regulation of the system. The model itself is suitable to be incorporated in larger scale models where the dynamic regulation of permeability is of interest.

3.1 Mechanisms of aquaporin regulation

Aquaporins are specialised transporter proteins that, when inserted in the cell membrane, let only pure water molecules pass. Sometimes, they are simply referred to as water channels. Aquaporins are very convenient molecules to make use of in the kidney, since it makes it possible to separate water and its solutes. The road from the aquaporin gene to a functional molecule in the apical membrane of kidney collecting duct cells is long and whindy, yet it is tightly regulated by circulating levels of vasopressin.

Figure 3.2 represents a collecting duct principal cell. At the basolateral membrane circulating AVP (green diamonds) will bind to its receptor (V2R). This is a G-protein coupled receptor, which stimulates adenylate cyclase to produce cyclic adenosine monophosphate (cAMP), which is a small intracellular messenger molecule. Cyclic AMP has many functions, and is involved in a huge variety of intracellular reactions. The purpose of cAMP generation in aquaporin signalling network is, first of all, an activation of protein kinase A (PKA). PKA is involved in numerous signalling networks, and in aquaporin regulation it has two main functions. The first is phosphorylation of aquaporins stored in intracellular vesicles. The phosphorylation is necessary for the exocytosis of aquaporins to the luminal cell membrane. This process is highly complex, and PKA phosphorylation is an initiating step [21, 24]. The process of membrane translocation or exocytosis was demonstrated in a study by Nielsen et al. [81]. The process of membrane insertion of aquaporin, usually takes place with a time scale of tens of minutes [67]. Indirect evidence also points towards a regulation of endocytosis by vasopressin, but the particular pathways involved have not been identified.

The second function of PKA in aquaporin regulation is phosphorylation of several transcriptional regulators of the aquaporin-2 gene. Yasui et al. [132] showed that vasopressin binding to V2R stimulates PKA-

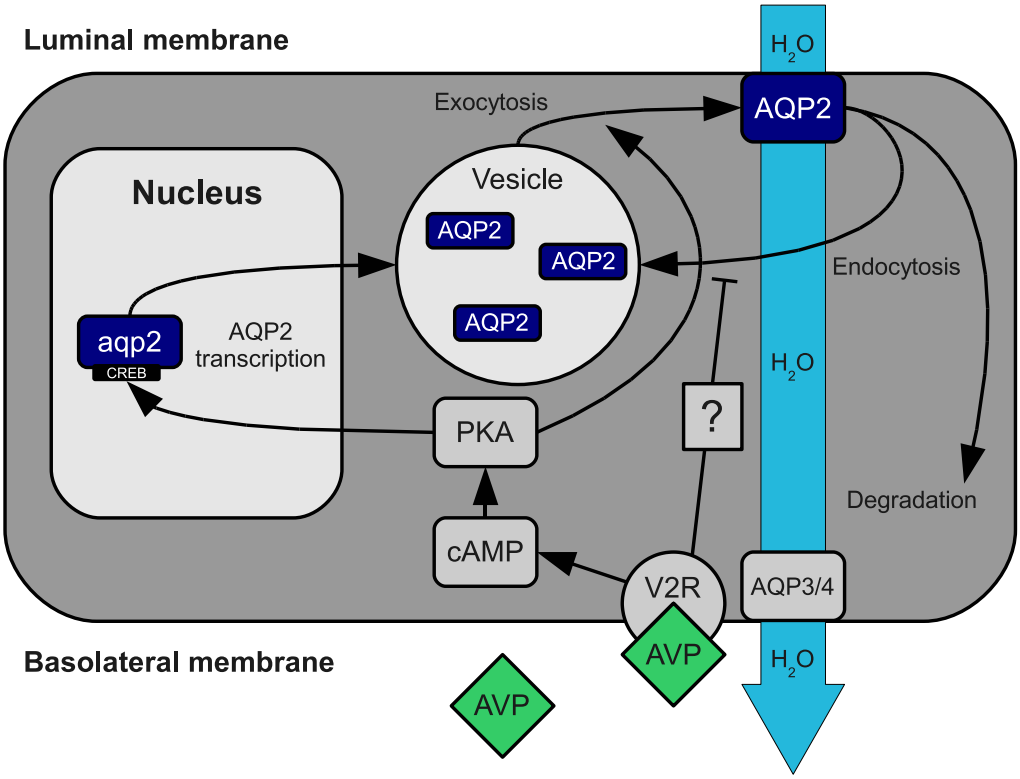


Figure 3.2: Diagram of regulation of aquaporins initiated by the binding of vasopressin (AVP green diamonds) to its receptor (V2R) (grey circle at the basolateral membrane). After binding to V2R several reactions are initiated resulting in (1) a an increase in exocytosis of AQP to the luminal membrane (2) a possible inhibition of endocytosis, and (3) a transcriptional upregulation of aqp2 gene through cAMP-dependent pathway. The effects occur on different time scales.

dependent phosphorylation of CREB (cAMP responsive element binding protein) and c-Fos protein. They also showed that the two phosphorylated proteins exert their effect on transcription through AP1 and CRE binding sites in the proximal aquaporin-2 promoter region. The study of Yasui et al. does not directly address the timescale of the upregulation of transcriptional activity. In another study by Matsumura et al. [75], the effect on aquaporin-2 mRNA by cAMP was studied. They found a steadily increasing mRNA abundance up to 4-fold after 24 hours.

Aquaporins in the plasma membrane share, like all other protein, the inevitable fate of degradation. However, some are conveniently recycled before being chopped up by the lysosome. Through an endocytotic process a proportion of membrane aquaporins reenter the cell cytoplasm, and are

stored in vesicles, ready for reinsertion in the membrane. Several cycles of exo- and endocytosis have been demonstrated with blocking of aquaporin synthesis [59]. A proportion of aquaporins are also directly degraded.

Many investigators have characterised the various parts of aquaporin regulation. However, there exist no quantitative model that integrates all of the above elements.

The Nielsen-Knepper model

The Nielsen-Knepper model was proposed by Knepper and Nielsen in 1993 [63]. In its simplest form the model includes an active and an inactive state for aquaporins. The basic assumption of the model is the total number of proteins is constant. Mathematically this leads to the following system of equations:

$$\frac{dC_A}{dt} = k_A C_I - k_I C_A$$
$$\frac{dC_I}{dt} = k_I C_A - k_A C_I \,,$$

where C_A represents the amount or concentration of active state proteins, and C_I is the inactive proteins. They also assume that

$$C_A + C_I = C_T \,,$$

where C_T is the total number of proteins, which is considered constant. One can solve this equation for C_A and obtain the following equation

$$C_A(t) = \frac{k_A C_T}{k_A + k_I} + \left(C_A(0) - \frac{k_A C_T}{k_A + k_I} \right) \exp\big(-(k_A + k_I)t \big) \,.$$

This equation represent an exponential decay to a steady state value of

$$C_A^* = \frac{k_A C_T}{k_A + k_I} \,, \quad \text{and} \quad C_I^* = \frac{k_I C_T}{k_A + k_I} \,. \tag{3.1}$$

The temporal behaviour is governed by the constant

$$\alpha = k_A + k_I \,. \tag{3.2}$$

Knepper and Nielsen obtain estimates of these two constants that are dependent on the total number of proteins C_T. The dependence can easily be derived by rearranging Equations (3.1) and (3.2) to

$$k_A = \alpha \frac{C_A^*}{C_T} \tag{3.3}$$

$$k_I = \alpha \left(1 - \frac{C_A^*}{C_T} \right) \,. \tag{3.4}$$

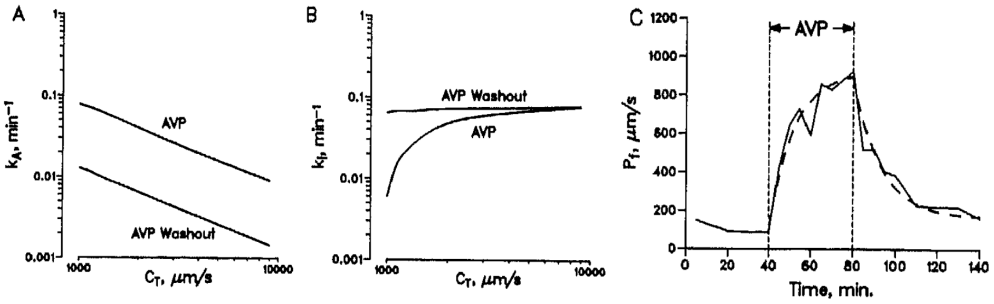


Figure 3.3: Parameter fit of the Nielsen-Knepper model to the experimental time series of water permeability of kidney collecting duct cells during vasopressin stimulation and subsequent washout phase. Fits for k_A and k_I are shown in (a) and (b) for different values of parameter C_T. In (c) the experimental (mean) data are shown (full line) and the fit (dashed line). Figure 6 from Knepper and Nielsen [63].

In Figure 3.3 (Figure 6 from [63]) the estimates of k_A and k_I as a function of C_T are presented in subfigure (a) and (b), respectively. If one compares the expressions in Equations (3.3) and (3.4) to the Figures 3.3(a) and (b), the dependence on C_T is clear. The data from Knepper and Nielsen is presented in Figure 3.3(c) (full line) along with the best fit (dashed line). At 40 min a stimulus of vasopressin is applied to a perfused kidney tubule, and at 80 min the vasopressin is washed out. They obtained a time constant α of approximately $5\,\mathrm{h}^{-1}$ during the stimulation period, and a value of $4.5\,\mathrm{h}^{-1}$ elsewhere. The increase in permeability is about 5-fold. From their parameter estimates Knepper and Nielsen claim that both k_A and k_I are altered by vasopressin stimulation. But from their data, we cannot immediately say what the value of k_I and k_A should be. We lack one crucial piece information - what is the total number of proteins? This question cannot readily be answered. If we had information about both intracellular and extracellular protein, we could fit these data to determine if both k_A and k_I are altered by vasopressin stimulation.

Membrane and intracellular proteins

The question of how many membrane and intracellular proteins that are in the cell is very difficult to answer. Nishimoto et al. [83] measured phosphorylation of aquaporin-2 at the site Ser-256, which has been proven to be necessary for exocytosis. They do not exactly quantify the amount of protein, but only present electrophoresis gels, where it is clear that AVP

Stimulus	Time scale	p/m ratio	Response
On	$5\,h^{-1}$	0.7	5-10 fold
Off	$4.5\,h^{-1}$	3	

Table 3.1: Parameters to fit from Knepper and Nielsen's data. The stimulus is either on or off. The time scale is changed slightly from on to off. The p/m ratio is the pool to membrane ratio. During the on state the p/m ratio is lower, due to the translocation of proteins to the membrane from the pool. The magnitude of the response is 5-10 fold.

enhances phosphorylation of protein without altering the overall protein abundance.

Nielsen et al. [81] used immunogold labeling of aquaporin proteins to quantify the density of aquaporins in intracellular vesicles and the apical plasma membrane in perfused rat collecting ducts in vitro. Their results indicate a ratio of membrane to intracellular protein of 0.3 with no vasopressin stimulus, and 1.4 after 40 min of stimulation. The permeability increase is about 7-fold.

It is not straight forward to incorporate all of this information in a dynamical model. The Nielsen-Knepper model does not account for all of this data, and especially not the long term behaviour of the biological system. To summarise the data by Knepper and Nielsen, I have presented some numbers in Table 3.1 that will serve as guidelines for the dynamic behaviour of the model.

In the following section, I will present a more detailed model and incorporate the above information in relation to the regulation of the exocytosis-endocytosis process.

3.2 A dynamic model of aquaporins

The model presented in this section will incorporate the information presented in the previous section. The model is schematically presented in Figure 3.4. It is composed of two parts. The first part, shown in bright gray, is the exocytosis-endocytosis dynamics. The dynamical behaviour of this part is on the order of minutes up to an hour. The second part, shown in dark gray, adds dynamics of transcription and protein synthesis. The dynamics is slower and on the time scale of hours. In green, the hormone vasopressin is shown. It affects exo- and endocytosis as well as transcription.

The state variable Y_r represents RNA of the membrane protein. The

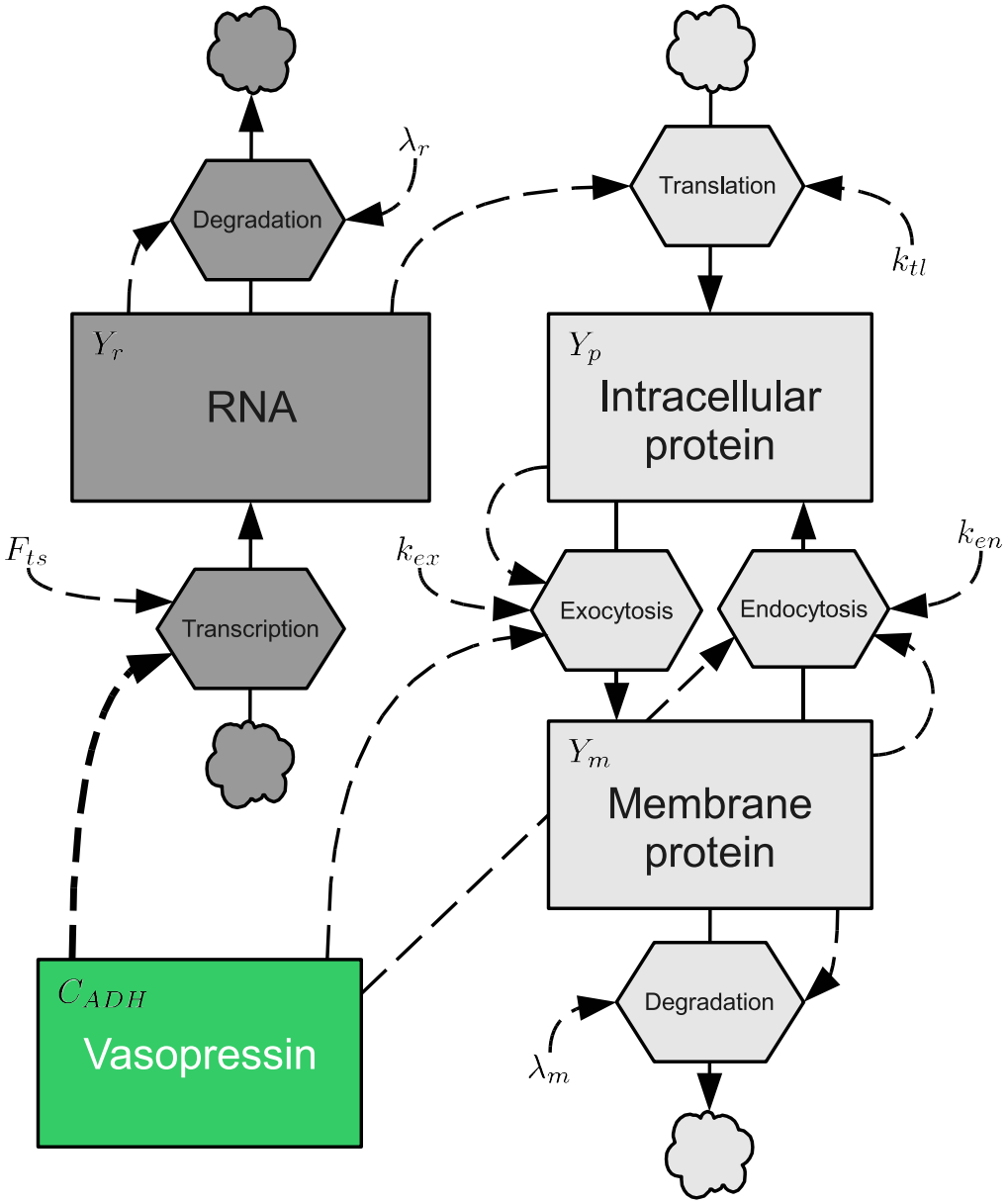


Figure 3.4: Flow diagram of the membrane protein model. Three state variables Y_r, Y_p, and Y_m represent RNA, intracellular protein pool, and membrane protein. RNA is created from the transcription process, and it affects the translation and creation of intracellular protein. The exchange of protein between the intracellular pool and the membrane is caused by exocytotic and endocytotic processes.

amount of Y_r controlled by the transcription rate F_{ts} and the degradation rate λ_r. The amount of intracellular protein is denoted Y_p. It is created from translation of RNA. The amount of protein translated per time unit depends on the parameter k_{tl}, and the amount of RNA Y_r. Membrane protein is denoted Y_m. The exchange between intracellular and membrane protein is controlled by the exocytosis and endocytosis rates k_{ex} and k_{en}. The degradation rate of Y_m is λ_m. In general k's describe rates of transition processes, and λ's describe degradation rates. The model equations are

$$\frac{dY_r}{dt} = F_{ts} - \lambda_r Y_r \tag{3.5}$$

$$\frac{dY_p}{dt} = k_{tl} Y_r - k_{ex} Y_p + k_{en} Y_m \tag{3.6}$$

$$\frac{dY_m}{dt} = k_{ex} Y_p - k_{en} Y_m - \lambda_m Y_m \,. \tag{3.7}$$

The steady states are

$$Y_r^* = \frac{1}{\lambda_r} F_{ts} \tag{3.8}$$

$$Y_p^* = \left(1 + \frac{k_{en}}{\lambda_m}\right) \frac{k_{tl}}{k_{ex}\lambda_r} F_{ts} \tag{3.9}$$

$$Y_m^* = \frac{k_{tl}}{\lambda_r} \frac{1}{\lambda_m} F_{ts} \tag{3.10}$$

All of these are proportional to the steady state of Y_r^* given in Equation (3.8).

The steady state ratio of membrane protein to intracellular protein is given by

$$\frac{Y_p^*}{Y_m^*} = \frac{k_{en} + \lambda_m}{k_{ex}} \,. \tag{3.11}$$

There are a couple of ways to *tune* the system; 1) increasing the total number of proteins by increasing transcription rate, 2) altering the relation between exocytosis and endocytosis of membrane protein. The first effect is a genetic effect that usually has a long time scale of several hours. The second one is usually rapid and occurs within minutes. We note that Y_m^* is independent of both endocytosis rate k_{en} and exocytosis rate k_{ex}. All Y-variables are proportional to the transcription rate. This means that changes in endocytosis or exocytosis can only cause transient changes in membrane protein abundance, while transcriptional up or down regulation cause an overall change. Hence the model design allows for a temporal separation of slow and fast responses to external stimuli.

Exo- and endocytosis

If we ignore for a start the processes of transcriptional and translational regulation and assume that the subsystem describing the production of protein can be neglected. In other words, the supply of protein is considered constant. Hence Equation (3.5) is in steady state ($Y_r^* = F_{ts}/\lambda_r$), and we can consider just Equations (3.6) and (3.7). This is simply a two-dimensional linear system. The system matrix is

$$A = \begin{bmatrix} -k_{ex} & k_{en} \\ k_{ex} & -k_{en} - \lambda_r \end{bmatrix}$$

The eigenvalues $\eta_\pm$ of this matrix are

$$\eta_\pm = \frac{-k_{ex} - k_{en} - \lambda_m \pm \sqrt{(k_{ex} + k_{en} + \lambda_m)^2 - 4k_{ex}\lambda_m}}{2} \tag{3.12}$$

Since all parameter values are positive, both eigenvalues values will always be negative, and for a given initial condition will tend to a steady state, given by Equations (3.9) and (3.10). The interesting part is therefore not the steady state but rather the transient behaviour. Since the system matrix has two negative eigenvalues we know that the solution will be a linear combination of two exponential functions. This also means that two time scales are involved in the system response, namely the time scales defined by the two eigenvalues.

Since the system is linear, all variables a proportional to the system input F_{tl}. We cannot make any *a priori* assumptions about the absolute value of the variables. Hence, it is convenient to have this proportionality. What we are basically interested in is getting the relative changes in Y_m and the temporal response right. The correspondence with water permeability is then taken care of by a proportionality constant. The best state of the system for comparison is the steady state of membrane proteins, as given by Equation (3.10). In the following we consider $F_{tl} = 1$, and $y_{p,m} = Y_{p,m}\lambda_m$.

If we assume λ_m to be small compared to k_{ex} and k_{en}, then we can make a first order Taylor expansion of the square root in Equation (3.12)

$$\begin{aligned} \eta_\pm &= \frac{-k_{ex} - k_{en} - \lambda_m \pm (k_{ex} + k_{en} + \lambda_m)\sqrt{1 - 4\frac{k_{ex}\lambda_m}{(k_{ex}+k_{en}+\lambda_m)^2}}}{2} \\ &= \frac{-k_{ex} - k_{en} - \lambda_m \pm (k_{ex} + k_{en} + \lambda_m) - 2\frac{k_{ex}\lambda_m}{k_{ex}+k_{en}+\lambda_m}}{2}. \end{aligned}$$

Assuming $k_{ex} + k_{en} \gg \lambda_m$, we obtain

$$\begin{aligned} \eta_+ &= -\frac{k_{ex}}{k_{ex} + k_{en}} \lambda_m \\ \eta_- &= -k_{ex} - k_{en} \end{aligned} \tag{3.13}$$

As λ_m tends to zero, η_+ also tends to zero.

In the following sections, I will investigate the dynamical response of the model to stimuli of exocytotic and endocytotic rates. It is commonly accepted that exocytosis is regulated by vasopressin through a cAMP-PKA dependent pathway though many details are not known. Evidence exists that endocytotic regulation occurs, though it remains unidentified.

Response to exocytotic stimulus

I will first look at the response of the system to an increase in exocytosis. The exocytosis rate will be modelled as

$$k_{ex}(t) = \begin{cases} k_{ex,0} * f_{ex} & : t \geq t' \\ k_{ex,0} & : t < t' \end{cases}$$

The parameter f_{ex} is a scalar amplification factor, and in the case of excitation $f_{ex} > 1$. The time t' is the time of onset of stimulation.

Figure 3.5 illustrates the pool to membrane ratio (a) and eigenvalue η_- (b)-(e). The figure illustrates the consequences of choosing different values for parameters k_{ex}, k_{en}, and λ_m. The dashed horizontal lines in (a) correspond to the respective label and Colo of the other subfigures. The grey, dashed line shows $\lambda_m + k_{en} = 4.5\,\mathrm{h}^{-1}$ in all subplots. The eigenvalues in Figure 3.5(c) ($k_{ex} = 5\,\mathrm{h}^{-1}$) are too high to explain the data of Knepper and Nielsen. A "fast" eigenvalue of around 4.5 to $5\,\mathrm{h}^{-1}$ is what we are looking for. So we can discard $k_{ex} = 0.5\,\mathrm{h}^{-1}$. The dashed line in (b), (d) and (e) ($\lambda_m + k_{en} = 4.5\,\mathrm{h}^{-1}$) falls with in that interval of η_-, so these values of k_{ex} seem more reasonable. In the following, we will consider a stimulus where $k_{ex,0} = 0.1\,\mathrm{h}^{-1}$ and $f_{ex} = 10$. This corresponds to moving from Figure 3.5(b) to (d), i.e. from blue to red.

Let's have a look at the magnitude of the response and the time to response for $k_{ex} = 0.1\,\mathrm{h}^{-1}$, $f_{ex} = 10$. The labelled curves in Figure 3.6(a) show the fold increase at maximum response, and Figure 3.6(b) shows the time to maximum response. At the values of k_{en} and λ_m that match our

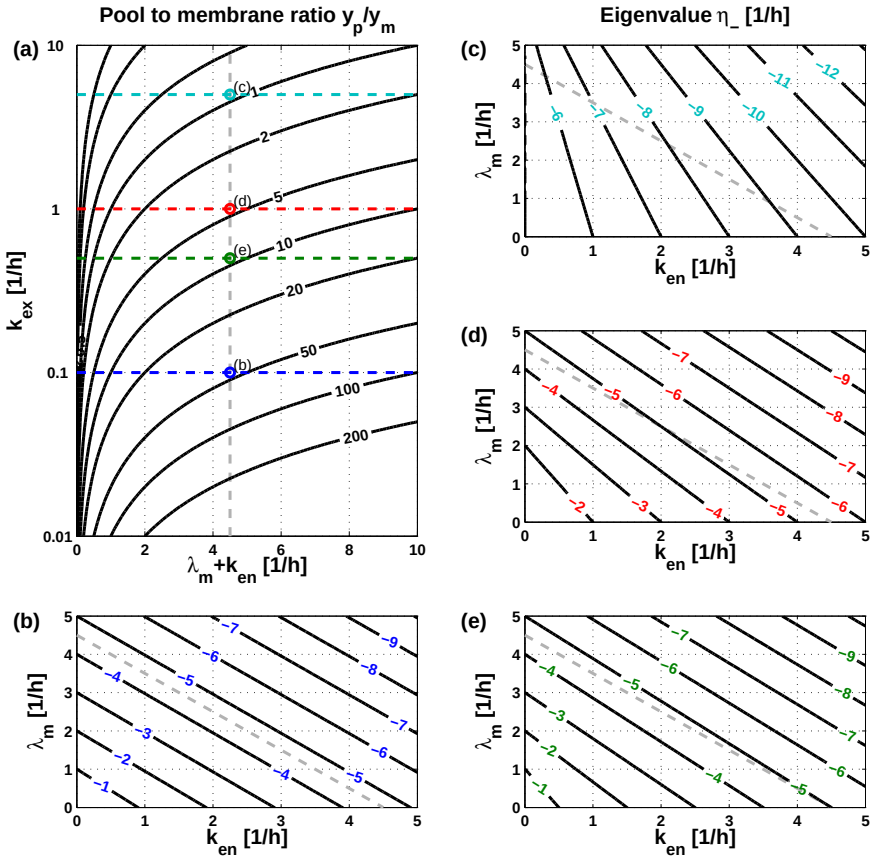


Figure 3.5: Pool to membrane ratio as function of $\lambda_m + k_{en}$ and k_{ex} (a). The crossings of the coloured, horizontal dashed lines at specific k_{ex}-values and the vertical line at $\lambda_m + k_{en} = 4.5\,\mathrm{h}^{-1}$ correspond to the parameter values used for calculating the fast eigenvalue in plots (b)-(e). Applying an exocytotic stimulus corresponds to moving vertically in the pool to membrane ratio plot (a).

time scales (grey dashed line) the response magnitude is relatively insensitive to variations in the range of a 7 to 8-fold increase. The response time however, is very sensitive to changes k_{en} and not so much to λ_m. An increase in k_{en} shortens response time.

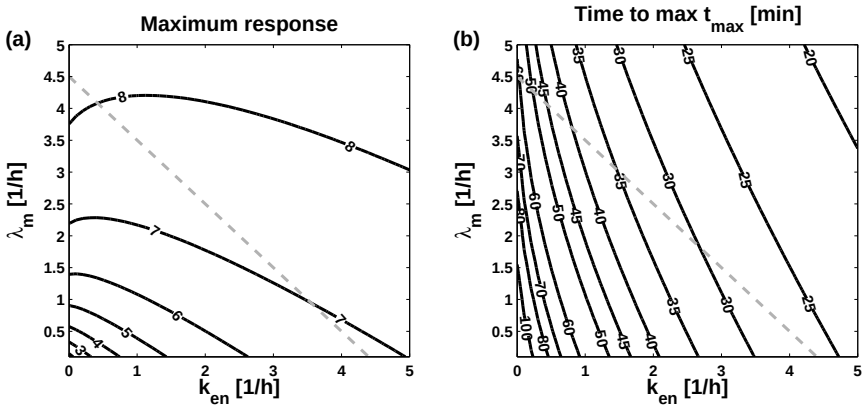


Figure 3.6: Maximum response (a) and time to maximum response (b) for a 10-fold increase for reference values of $k_{ex} = 0.1\,\mathrm{h}^{-1}$ and $f_{ex} = 10$.

The second eigenvalue η_+ also depends on the system parameters. This dependence is shown in Figure 3.7(a) and (b) which correspond to $k_{ex} = 0.1\,\mathrm{h}^{-1}$ and $k_{ex} = 1\,\mathrm{h}^{-1}$. The two figures practically overlap except for a factor of 10 on the labels. The 10-fold increase in k_{ex} corresponds approximately to a 10-fold increase in η_+. The eigenvalue η_+ is not tuneable if we choose parameters to match the fast time scale of the response and the response magnitude, so it imposes a dynamical constraint on the model, and, in principle, a testable hypothesis for experiments. The decay during stimulation should be possible to test experimentally. However, in the pre- and post-stimulation periods the slow time scale would be hard to measure.

Figure 3.8 shows simulations of the step response for various parameter values. The onset of stimulation begins at $t = 1\,\mathrm{h}$. In all curves in Figure 3.8 the green curve corresponds to the reference case with parameter values $k_{en} = 3.5\,\mathrm{h}^{-1}$, $k_{ex,0} = 0.1\,\mathrm{h}^{-1}$, and $\lambda_m = 1\,\mathrm{h}^{-1}$. . The first column (subfigures (a) to (d)) is the normalized membrane population y_m, and the second column (subfigures (e) to (h)) is the intracellular pool population y_p. In subfigures (a) and (e) three different values for λ_m were used $0.2\,\mathrm{h}^{-1}$, $1\,\mathrm{h}^{-1}$, and $5\,\mathrm{h}^{-1}$. The maximal response shows little dependence on λ_m, and it is clear that the smaller λ_m is, the longer it takes before the system

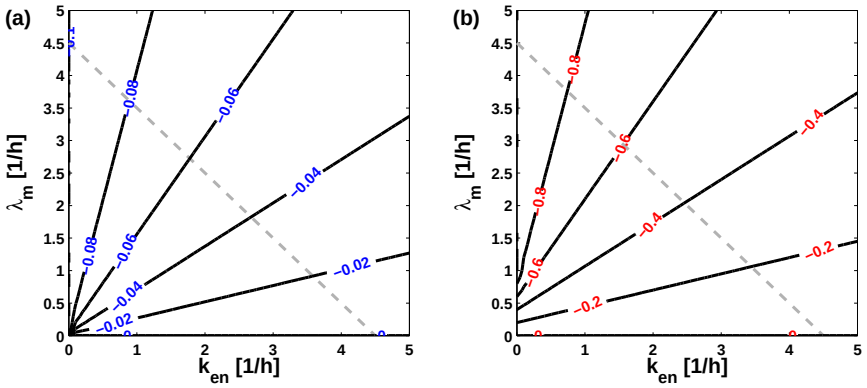


Figure 3.7: The slow eigenvalue η_+ for $k_{ex} = 0.1\,\mathrm{h}^{-1}$ (a) and $k_{ex} = 1\,\mathrm{h}^{-1}$ (b).

reaches steady state. The intracellular population is not particularly sensible to variation in λ_m. Remember that we are considering the population relative to the steady membrane population. Hence a larger λ_m will give rise to a larger y_p. The pre-stimulation value of Y_p^* is in fact equal to the membrane to intracellular protein ratio, since $\frac{Y_p^*}{Y_m^*} = \frac{y_p^*}{y_m^*} = y_p^*$. This gives a measure of how large the pool of available proteins is compared to the membrane population, and is given by Equation (3.11). So for a small λ_m, we have that $y_p \simeq \frac{k_{en}}{k_{ex}}$, which with the reference values is about 10. For $\lambda_m = 0.05\,\mathrm{h}^{-1}$ and $0.5\,\mathrm{h}^{-1}$, we see that the prestimulus y_p is around 10, then λ_m influences the rate of decrease to a steady state. With the large $\lambda_m = 5\,\mathrm{h}^{-1}$ the initial value is larger, and the decrease is far more rapid. The decay at $\lambda_m = 0.05\,\mathrm{h}^{-1}$ is very slow. The "drying out" of the intracellular pool at $\lambda_m = 5\,\mathrm{h}^{-1}$ occurs very rapidly, not leaving enough protein for repeated stimuli. When considering only exocytotic stimulus, the pool to membrane ratio is one feature which is not well reproduced.

In subfigures (b) and (f) the endocytosis rate k_{en} was set to $0.1\,\mathrm{h}^{-1}$, $3.5\,\mathrm{h}^{-1}$, and $7\,\mathrm{h}^{-1}$. In contrast to λ_m, we clearly see that the endocytosis rate has an effect on the magnitude of the response, ranging from a doubling to a 10-fold increase in membrane population. The response also becomes more rapid, and the system stiffer as k_{en} increases. The size of the intracellular pool depends heavily on endocytosis. This is obvious from the affine relation in Equation (3.11). For $k_{en} \gg \lambda_m$ this relation is approximately linear and, hence, for the value $k_{en} = 7\,\mathrm{h}^{-1}$ we have a very large pool size. The decay rate (η_+) is not affected by k_{en}. The highest k_{en}-value renders the system very stiff, and the intracellular pool extremely large. The lowest value leaves the system unresponsive to the stimulus, due to a very small intracellular

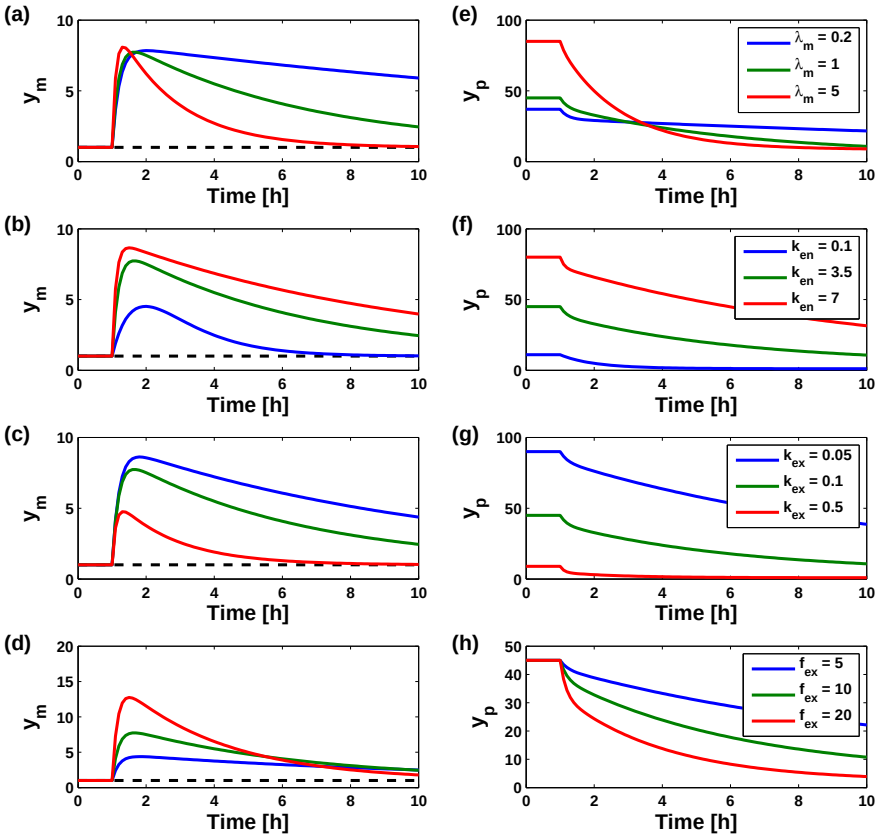


Figure 3.8: Simulation of the exocytosis-endocytosis model for different parameter variations. The left column shows normalized membrane population, and the right column shows the intracellular pool population. Each row shows simulation for three different parameter values of λ_m, k_{en}, k_{ex}, and f_{ex}. The green curve plots the reference case with parameter values $k_{ex} = 0.1\,\mathrm{h}^{-1}$, $k_{en} = 3.5\,\mathrm{h}^{-1}$, $\lambda_m = 1\,\mathrm{h}^{-1}$, and $f_{ex} = 10$. The different parameters affect response magnitude, response time, system stiffness, and the relative size of the intracellular pool.

pool.

Figures 3.8(c) and (g) show the different responses for the varying magnitudes of the stimulus k_{ex}. The response magnitude clearly increases with decreasing k_{ex}. So does the intracellular pool size. This is similar to the effects of increasing k_{en}, however, the rapidness of the response is higher for the lowest magnitude for varying k_{ex}, where as the opposite is the case for varying k_{en}.

The effect of varying the size of the stimulus f_{ex} is shown in subfigures (d) and (h). Increasing f_{ex} leads to a higher maximum membrane population, and the time to maximum response decreases slightly. The intracellular pool changes are similar, and higher stimulus drains the pool more.

Figure 3.9 shows the maximal response $y_{m,max}$ in the left column and the time of maximal response t_{max} in the right one for different sets of parameters. Figures (a) and (d) results for different values of λ_m, figures (b) and (e) for k_{en}, and figures (c) and (f) for k_{ex}. An increase in λ_m and k_{en} cause in increase in response magnitude and a decrease in response time. Increasing k_{ex} cause a decrease in both response magnitude $y_{m,max}$ and response time. This is because with higher k_{ex} more proteins are already located in the membrane compartment, and hence the relative response will not be as pronounced. Increasing f_{ex} causes $y_{m,max}$ to increase in all cases, however, for high values of k_{ex} and low values of k_{en} the increase is not very pronounced. This is again due to the fact that more proteins are located in the membrane before the onset of stimulus.

The exocytotic rate stimulus gives a good understanding of the system behaviour and response in terms of parameters. The main problem in relation to aquaporins is matching time scales, pool to membrane ratio *and* response magnitude at the same time.

Simultaneous regulation of exo- and endocytosis

We now assume that the system response to vasopressin is a simultaneous regulation of exo- and endocytosis, since the pure exocytotic regulation did not prove sufficient.

Figure 3.10 shows the variation in the pool to membrane ratio with k_{ex} and k_{en}. An increase in λ_m generally gives a higher pool to membrane ratio. The lines also become more horizontal, meaning that the ratio becomes less sensitive to changes in k_{en} as λ_m increases.

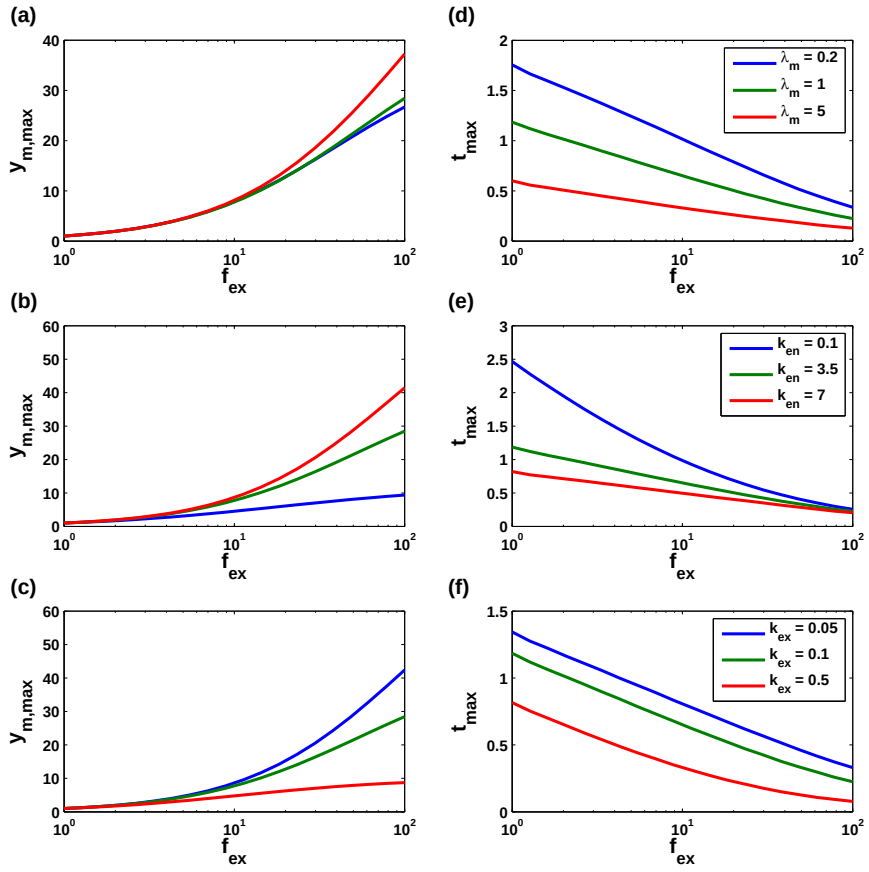


Figure 3.9: Dependence of maximum response and time of maximum response on f_{ex}, λ_m, and k_{en}.

Figure 3.11 shows the largest eigenvalue η_- for the same four values of λ_m. The visual pattern is straight lines from top left corner towards the bottom right corner. As λ_m increases the lines are squeezed towards lower k_{ex} and k_{en}. In subfigure (d) the values are too high for the time constant that we seek. The other three values seem to be applicable for matching the time scale of the problem.

Figure 3.12 shows the smallest of the two eigenvalues η_+. The visual pattern of the lines is similar to the pool to membrane ratio plots with rays emitted from the bottom left hand corner. The absolute value of η_+ increases with increases λ_m.

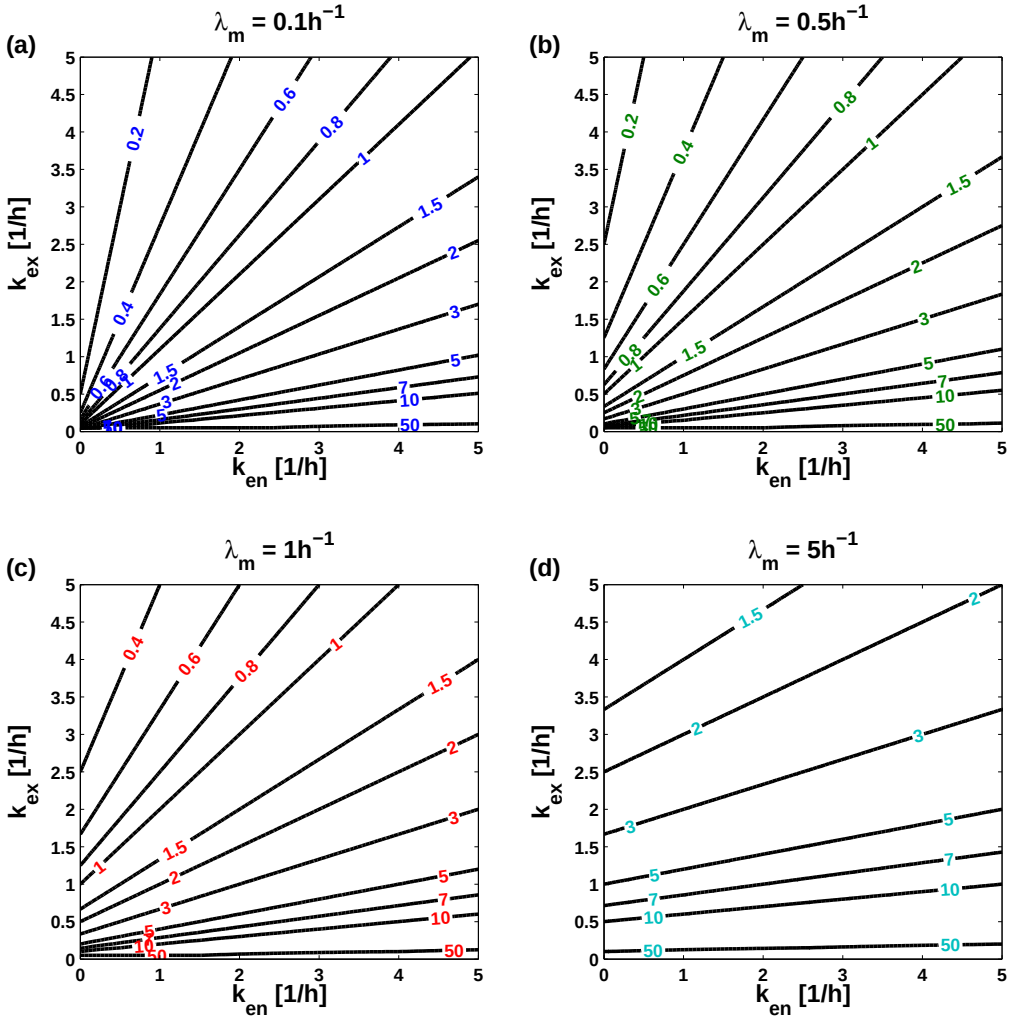


Figure 3.10: Pool to membrane ratio for $\lambda_m = 0.1\,\mathrm{h}^{-1}$ (a), $0.5\,\mathrm{h}^{-1}$ (c), $1\,\mathrm{h}^{-1}$ (b) , and $5\,\mathrm{h}^{-1}$ (d).

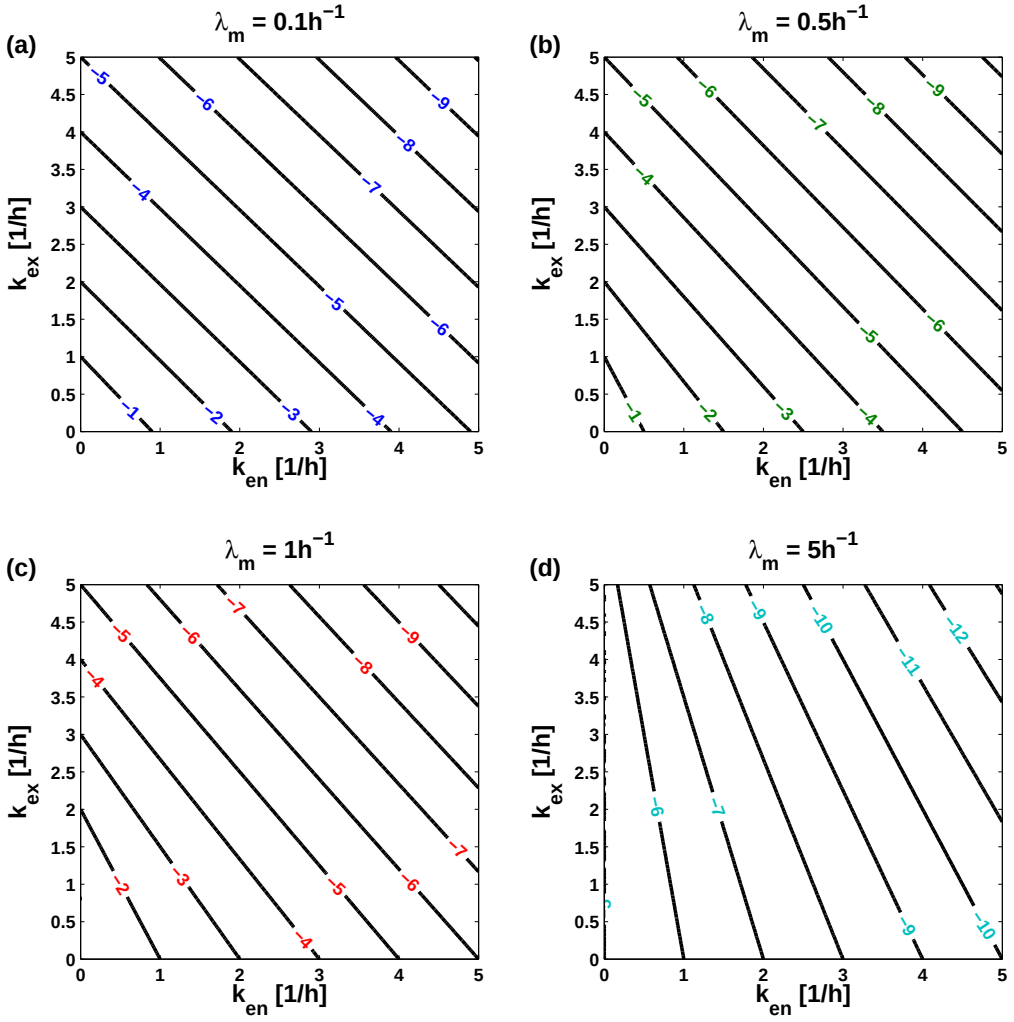


Figure 3.11: Eigenvalue η_- for $\lambda_m = 0.1\,\mathrm{h}^{-1}$ (a), $0.5\,\mathrm{h}^{-1}$ (b), $1\,\mathrm{h}^{-1}$ (c) , and $5\,\mathrm{h}^{-1}$ (d).

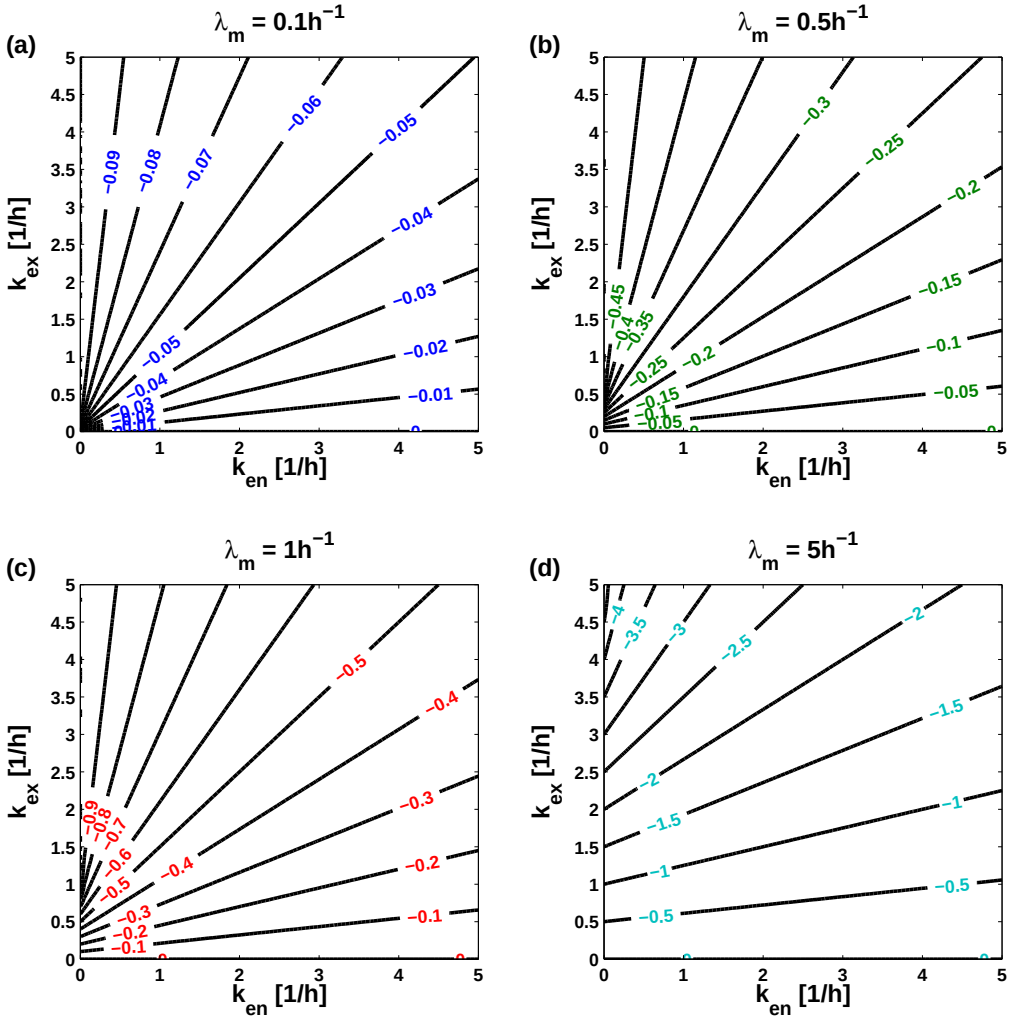


Figure 3.12: Eigenvalue η_+ for $\lambda_m = 0.1\,\mathrm{h}^{-1}$ (a), $0.5\,\mathrm{h}^{-1}$ (b), $1\,\mathrm{h}^{-1}$ (c) , and $5\,\mathrm{h}^{-1}$ (d).

Stimulus	λ_m [h^{-1}]	k_{en} [h^{-1}]	k_{ex} [h^{-1}]
On	0.1	1.98	2.98
Off		3.29	1.33
On	0.5	1.69	3.13
Off		2.97	1.16
On	1	1.33	3.13
Off		2.57	1.19

Table 3.2: Parameter values of k_{en} and k_{ex} that match the criteria of the Knepper and Nielsen data.

Based on the general features of the visual presentation in Figures 3.10-3.12 we can estimate k_{ex} and k_{en} to match the criteria in Table 3.1. This is presented in Figure 3.13 for the same λ_m-values as in the previous figures. The optimal values are found at the intersection of the curve corresponding to a pool to membrane ratio of 3 and $\eta_- = -4.5$ for the pre- or post-stimulation period, and at a ratio of 0.7 and $\eta_- = -5$ for the stimulation period. The arrows indicate the path from pre-stimulation to stimulation. These values can be found by a simple least squares optimisation, and the results are shown in Table 3.2. The arrows in Figure 3.13 indicate the transition from pre-stimulation to stimulation. In all cases it is clear that the transition involves an upregulation of exocytosis and a down regulation of endocytosis as expected. The values in Table 3.2 reveal that for increasing λ_m the corresponding k_{en}-values decrease both in off and on states. On the other hand k_{ex} increases during stimulation and decreases when there is no stimulation. The variation is relatively modest, and we may conclude that in this range of λ_m-values, it is not possible to discriminate between them. Therefore we will have a look at the temporal response.

The temporal response to the obtained parameter values can be seen in Figure 3.14. The transition to the stimulus period occurs at $t = 0$. The main difference of the three simulations is the slow time scale of adaptation. This can be understood by considering the eigenvalue η_+ which approaches zero when λ_m decreases. This is also seen in Figure 3.12. The highest λ_m-value of $1\,h^{-1}$ induces a very rapid decay.

Another observation is the magnitude of the response. There is an approximate 2 to 2.5-fold increase in membrane protein abundance. This is somewhat smaller than the increase in permeability seen in the Knepper and Nielsen data (Table 3.1), where the increase was 5-10 fold. There may be several explanations for this. First of all, it is not certain that the membrane

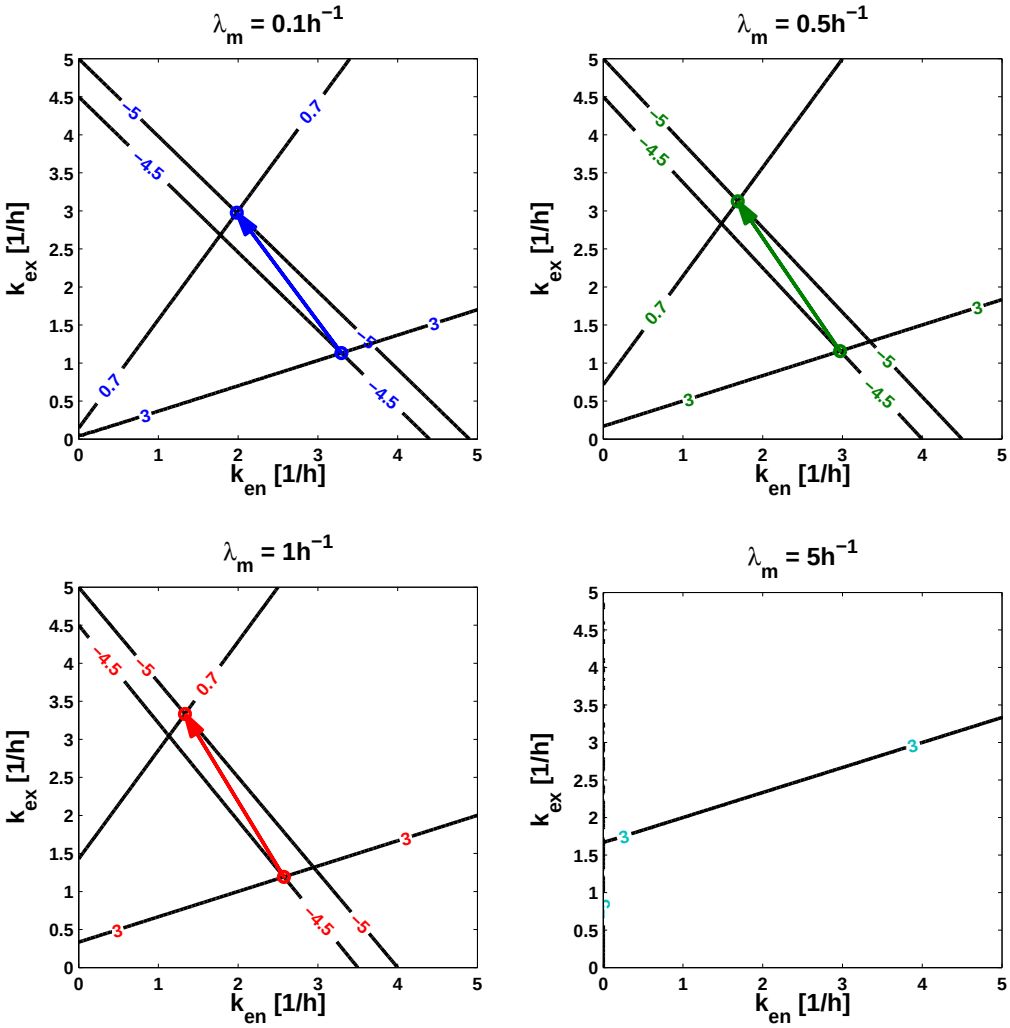


Figure 3.13: Crossing of eigenvalue η_- and pool to membrane ratio for $\lambda_m = 0.1\,\mathrm{h}^{-1}$ (a), $0.5\,\mathrm{h}^{-1}$ (b), $1\,\mathrm{h}^{-1}$ (c) , and $5\,\mathrm{h}^{-1}$ (d). Arrows indicate the transition from pre-stimulation to stimulation.

protein abundance is proportional to the permeability. This is a simple model assumption, and it is possible that permeability increases more than linearly with the amount of membrane protein. Secondly, the data used to estimate the time scales and the data used to estimate protein abundance were taken from different experiments. Though the experimental protocols were very similar, there may be some discrepancy between them. If we assume proportionality between proteins and permeability then we must increase the dynamic range of the model by increasing pool to membrane ratio, when stimulation is not present, and/or decreasing the ratio, when stimulus is present.

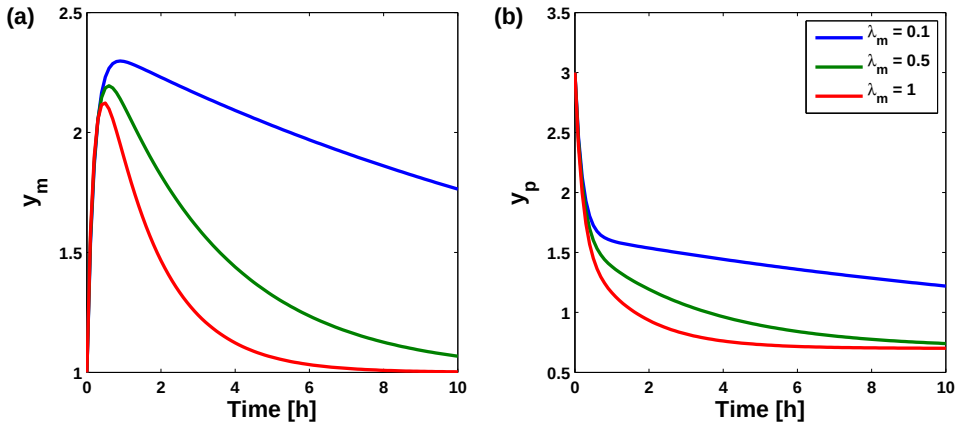


Figure 3.14: Time plots of solutions to the values found in Figure 3.13.

Transcriptional regulation

The analysis of the exo- and endocytosis model showed that during continued stimulation the level of membrane proteins will decrease over time. Therefore it is necessary to have a mechanism that can maintain the level of membrane proteins in spite of the decay. A biological candidate for such a mechanism as an upregulation of transcriptional activity, which ultimately leads to the production of protein. There is good evidence that vasopressin activates transcription of the aquaporin-2 gene as already discussed in Section 3.1. A typical time scale for the transcriptional upregulation is of several hours and much slower than the immediate response. In the following, I will introduce the transcriptional upregulation in the model. This corresponds to the dark grey part in the model diagram in Figure 3.4. The rate that is affected by vasopressin is the transcription rate F_{ts}. If we look

at the model equations in particular Equation (3.5), the time constant here is λ_r. The degradation constant controls the mRNA abundance. Likewise the time constant for the translation process is k_{tl}. The upregulation of transcription itself is due to the complicated signalling pathways that activate or deactivate transcription. Parts of this pathway is stimulated very quickly. The phosphorylation of the transcriptional regulator CREB takes place within few minutes of vasopressin stimulation [132]. Hence we assume that the delay is caused by the transcription and translation processes.

Consider the system matrix of the full model

$$A = \begin{bmatrix} -\lambda_r & 0 & 0 \\ k_{tl} & -k_{ex} & k_{en} \\ 0 & k_{ex} & -k_{en} - \lambda_r \end{bmatrix}$$

The eigenvalues of A are $\eta_r = -\lambda_r$, and $\eta_\pm$ as given by Equation (3.12). The eigenvalues give the time scales of response to perturbations, and in the next section it will become apparent, how they relate to overall system behaviour.

Regulating transcription

In the following, we will assume that vasopressin stimulation does not affect post-transcriptional processes, which in the model are represented by RNA lifetime and translation. These two processes constitute a delay in the upregulation of protein. As previously used for the exocytotic stimulus, a similar expression can be used for a transcriptional stimulus.

$$F_{ts}(t) = \begin{cases} F_{ts,0} * f_{ts} & : t \geq t' \\ F_{ts,0} & : t < t' \end{cases}$$

The value of $F_{ts,0}$ is set to 1 au/h, where au stands for arbitrary unit. We will now investigate the effect of different values of f_{ts}. For the same situations as presented in Figure 3.14 and Table 3.2, I have plotted time courses of combined exo-/endocytotic stimulation and transcriptional regulation in Figure 3.15. The first column (a)-(c) corresponds to a doubling of the transcription rate $f_{ts} = 2$, (d)-(f) correspond to $f_{ts} = 4$, and (g)-(i) to $f_{ts} = 8$. The rows from top to bottom correspond to $\lambda_m = 0.1\,\mathrm{h}^{-1}$, $0.5\,\mathrm{h}^{-1}$, and $1\,\mathrm{h}^{-1}$ respectively. Blue curves are RNA population, green is intracellular pool of protein, and red are the membrane population of proteins.

A first observation is that the blue curves all show an exponential decay to their increased steady state level. The time constant of this process is the first eigenvalue η_r, and is unchanged by variation in λ_m. As was determined

in Figure 3.12, the slow time scale associated with η_+ becomes slower with decreasing λ_m. This is clearly seen by comparing the figures row by row. In the bottom row the equilibrium level is quickly approached, whereas the top and middle rows still have not reached steady state after 10 h. We also see that we can tune the transcriptional stimulus to obtain different types of response. Take Figure (b), where the red curve reaches a maximum and then drops off to a level that is higher than the pre-stimulation equilibrium level. If f_{ts} is larger than the maximum response, as is the case in the second and third column, there will be a slow increase in membrane protein to a level higher than the maximal immediate response.

It is also clear from Figure 3.15 what role each of the three eigenvalues play. The first η_r only relates to the regulation of the RNA abundance. The fast eigenvalue η_- takes care of the dynamics between the pool and the membrane. The slow η_+ is the long term adaptation of the protein populations both intracellular and membrane. Even if the RNA population has reached an equilibrium, the protein population still take some time to adapt to changes.

3.3 Discussion

In this chapter, I have discussed the dynamics of aquaporin regulation. I have developed a dynamic model of exo- and endocytosis and transcriptional regulation. The model analysis has contributed several insightful conclusions on the dynamic response and the dynamic constraints of the system structure. An analysis of the dependence of the eigenvalues on parameters is essential to understanding the system.

The simplest model of an exo- and endocytotic regulation revealed that the system is governed by two time scales - a fast one related to exchange of proteins between an intracellular pool and the membrane, and a slow one that governs the system's approach to equilibrium. Data from Knepper and Nielsen was used to find suitable model parameters that could reproduce both the time scales of the system, and very importantly the relative populations of intracellular and membrane proteins. A very important consequence of this analysis is that the immediate system response to vasopressin must involve both an exocytotic upregulation and an endocytotic downregulation. The visual techniques of plotting eigenvalues and relative population sizes as functions of exo- and endocytosis rates proved very efficient to obtain this insight.

The data by Knepper and Nielsen does not provide us with enough information to determine all system parameters. The slow time scale related to

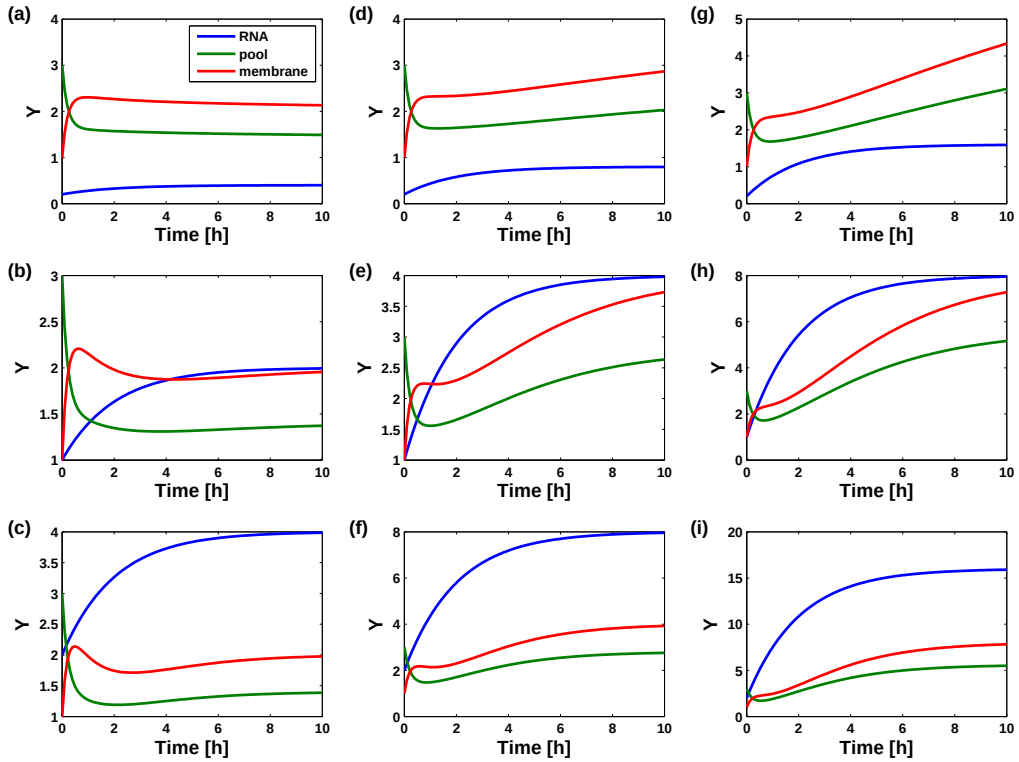


Figure 3.15: Time plots of both exo-/endocytotic stimulation and transcriptional regulation. The first column (a)-(c) corresponds to a doubling of the transcription rate $f_{ts} = 2$, (d)-(f) correspond to $f_{ts} = 4$, and (g)-(i) to $f_{ts} = 8$. The rows from top to bottom correspond to $\lambda_m = 0.1\,\mathrm{h}^{-1}$, $0.5\,\mathrm{h}^{-1}$, and $1\,\mathrm{h}^{-1}$ respectively. Blue curves are RNA population, green is intracellular pool of protein, and red are the membrane population of proteins.

the eigenvalue η_+ would need experimental verification. The model predicts that it heavily depends on the membrane degradation rate.

The simple model was also extended to include the dynamics of transcriptional regulation of the aquaporin gene by vasopressin. This increases the dimensionality of the model, and introduces a new time scale related to a third system eigenvalue. This eigenvalue controls the delay of the response to changes in transcription. The addition of the transcriptional component of the model allows for model-based data analysis of future experiments, where time scales of several hours are considered.

The model is not only applicable to aquaporin regulation. It could easily be adapted to other system that share the characteristics of an exo-/endocytosis type of regulation, where a membrane property can be mea-

sured. This could be other types of transport proteins, as for example epithelial sodium channels ENaC that are both rapidly and transcriptionally controlled by the hormone aldosterone [134].

The identification of probable values for parameters in the dynamic model also allow for the inclusion of this dynamic behaviour in a larger scale model. This will be done in Chapter 5, where the dynamics of membrane properties of the renal collecting duct will be used to study the control of water excretion.

3.4 Summary

A model for regulation of membrane proteins was formulated and its dynamical properties investigated. The parameters of the model were evaluated with data from experiments on water permeability of kidney collecting duct cells by aquaporins. The data stemmed from two related, but independent, experiments providing different types of information - time scales, and relative magnitudes. The model is formulated so that it is possible to describe both fast membrane activation regulation, and slower transcriptional regulation.

Chapter 4

Renal regulation of urine osmolarity

In this chapter, I will present a model of the kidney. The model focuses on the control of urine osmolarity. As indicated in Figure 4.1, we turn our attention to the main controller of the water and salt balance system. The kidney filters the blood and tightly regulates the final composition of the urine. The fine-tuning of the urine occurs in the distal part of the kidney, and this is the part of the kidney we are going to model. More precisely, we will formulate a model of the collecting duct and investigate how urine composition depends on water and solute permeability and renal interstitial osmolarity.

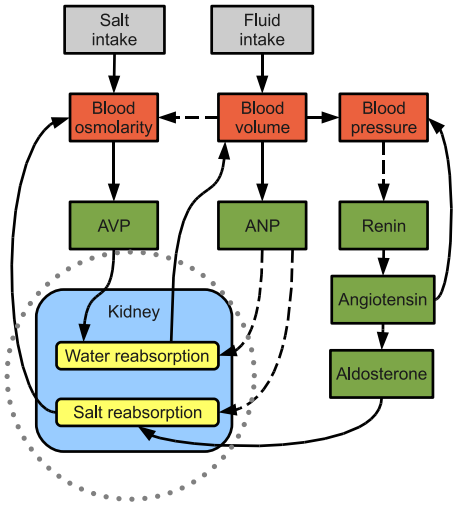


Figure 4.1: The causal loop diagram of water and salt balance presented in Figure 2.1. In this chapter, we focus on the kidney. The kidney controls how much salt and water is reabsorbed and, hence, how much leaves the organism. Kidney function is tremendously complex, and we limit ourselves to studying the regulation in the distal part of the kidney.

4.1 Kidney model

The kidney model focuses on the distal part of nephron function. This part of the nephron is under hormonal control of vasopressin, aldosterone, ANP, and many other substances. It is assumed that the proximal nephron mainly generates an outer medullary concentration gradient, and that it assures a stable supply of tubular fluid of near constant composition. The precise regulation of the final urine is thus left up to hormonal controls in the collecting duct.

The model is presented in Figure 4.2. It consists of two compartments; a collecting duct (yellow part) and the surrounding interstitium. The spatial coordinate x runs along the tubule. The tubule has a radius of r and length L. At $x = 0$ fluid from the distal tubule of the nephron flows into the collecting duct. We consider a volume or water flow Q_W and a solute flow Q_S, and they are functions of x. They are assumed to be in steady state which is a reasonable assumption at time scales of more than a minute. The main solutes of the urine are sodium, potassium, and urea, but here we consider all solutes in one pool, and simply refer to them as solutes with subscript S. The concentration of solutes in the tubular fluid at any position is the ratio of solute and water flow $c_S(x) = \frac{Q_S}{Q_W}$. The walls of the tubule are permeable to water and solutes, and we introduce the corresponding permeabilities G_W and G_S. A reabsorptive flow occurs from the tubule to the interstitium. In the interstitium a solute gradient is present. This is presented by the spatially dependent, interstitial osmolarity $c_i(x)$. The final urine is the flow at $x = L$, *i.e.* $Q_W(L)$ and $Q_S(L)$. The model only represents one collecting duct, and we will extrapolate to a whole kidney by multiplying the flows by the total number collecting ducts in the kidney n. The mathematical formulation of the model builds on the large amount of kidney modelling studies in the literature dating back to Stephenson's work from 1972 onwards [108, 109, 110, 111]. Much of the work found in the literature deals with the creation of a osmolarity gradient in the medulla of the kidney [69, 124, 125].

Equations

The flow in tubules is modelled as a steady state flow. The change in flow with respect to x equals the transmural line flux $J(x)$ counted positive in

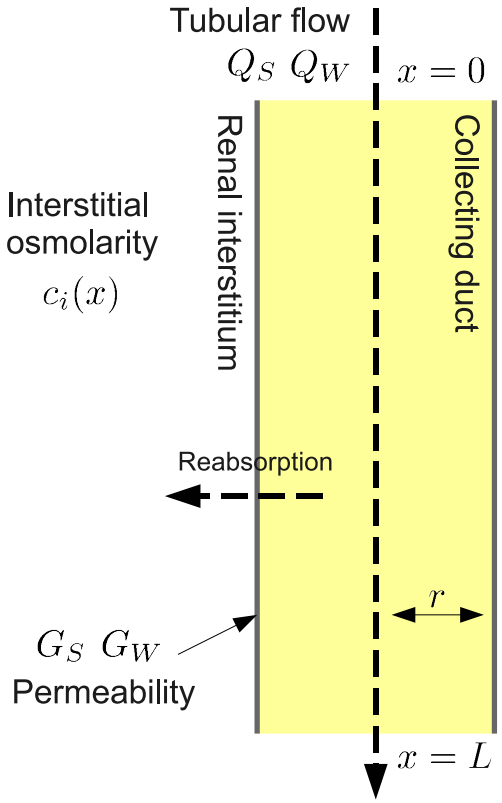


Figure 4.2: Diagram of the single tubule kidney model. We consider the collecting duct as a single tubule with radius r. The tubule is surrounded by the renal interstitium that has an osmolarity $c_i(x)$. The osmolarity increase with the spatial coordinate x. The tubular fluid flows into the collecting duct from the top at $x = 0$, and the final urine comes out at $x = L$. The solute flow is denoted Q_S, and the water or volume flow is Q_W. The tubule walls are permeable to water and solutes and the permeabilities are denoted G_W and G_S, respectively. Along the tubule, reabsorption of water and salt occurs as a function of interstitial osmolarity c_i and membrane permeabilities G_W and G_S. The fluid at $x = L$ is the urine that flows through the ureter to the bladder.

the outward direction from tubule to interstitium (see Figure 4.2)

$$\frac{dQ_W}{dx} = -J_W(x)$$
$$\frac{dQ_S}{dx} = -J_S(x)\,.$$

The transmural water flux J_W is driven by the osmolal gradient across the tubular wall. It is modelled as a simple transmural diffusion process

$$J_W(x) = 2\pi r n M_W G_W \left(c_i(x) - c_S\right)\,,$$

M_W is the molar volume of water, r is the radius of the collecting duct, n is the number of collecting ducts, and G_W is permeability which depends on the presence of aquaporins in the luminal membrane. The solute concentration is $c_S = \frac{Q_S}{Q_W}$. The interstitial concentration c_i is modelled as a linear function of x

$$c_i(x) = c_{i,0} - \frac{x}{L}(c_{i,L} - c_{i,0})\,,$$

where $c_{i,0}$ is the medullary osmolarity at the cortex ($x = 0$), and $c_{i,L}$ is at the papillary tip ($x = L$).

The solute flux is represented as a sodium flux through ENaC channels at the luminal membrane coupled sodium potassium pumps at the basolateral membrane of collecting duct cells. This mechanisms is representative for aldosterone mediated sodium reabsorption [82, 102], and we make the crude assumption that all solute flux occurs in this way. It may lead to an overestimation of diffusive solute transport. The transmural solute flux is

$$J_S(x) = -n2\pi r G_S\left(c_S\right) .$$

This basically corresponds to a diffusion between the tubule and compartment with no solute, which is an approximation of diffusion potential that sodium faces when diffusing into a cell where the sodium concentration around 10 mM.

The final equations are

$$\begin{aligned}\frac{dQ_W}{dx} &= 2\pi r n M_W G_W \left(Q_S/Q_W - c_i(x)\right)\\ \frac{dQ_S}{dx} &= 2\pi r n G_S \left(Q_S/Q_W\right)\end{aligned}$$

The number of parameters in Equations (4.1) and (4.1) can be reduced. We can first normalise the length coordinate x by L, so that $z = x/L$. Then we denote the inflow to the collecting by $Q_{W,0} = Q_{W,0}$, and introduce the normalised flow $q_W = Q_W/Q_{W,0}$. The interstitial concentration at $x = 0$, $c_{i,0}$, is used to normalise the tubular concentration $c'_S = c_S/c_{i,0} = q_S/q_W$, and the same is the case for the interstitial concentration $c'_i = c_i/ci, 0$. The normalised solute flow is $q_S = c'_S q_W$. Combining these normalisations yields

$$\begin{aligned}\frac{dq_W}{dz} &= \frac{2\pi r n L c_{S,0} M_W}{Q_{W,0}} G_W \left(q_S/q_w - c'_i(x)\right)\\ \frac{dq_S}{dz} &= \frac{2\pi r n L}{Q_{W,0}} G_S \left(q_S/q_W\right) .\end{aligned}$$

We now introduce the normalised permeabilities g_W and g_S given by

$$\begin{aligned}g_W &= \frac{2\pi r n L c_{i,0} M_W}{Q_{W,0}} G_W\\ g_S &= \frac{2\pi r n L}{Q_{W,0}} G_S .\end{aligned}$$

to obtain the much simpler equations

$$\frac{dq_W}{dz} = g_W\big(c'_S - c'_i(x)\big) \tag{4.1}$$

$$\frac{dq_S}{dz} = g_S c'_S\,. \tag{4.2}$$

Equations (4.1) and (4.2) will be the subject of investigation in the following.

Parameters

In the previous section a lot of parameters were introduced. The purpose of this section is to get a value for each of them.

The interstitial osmolarity gradient is assumed to have a linear profile. This is a very crude approximation, but still we can put some numbers on the boundary values at $x = 0$ and $x = L$. At the cortical boundary at $x = 0$, the osmolarity is close to that of plasma, and we set it to $c_{i,0} = 300\,\text{mOsm/l}$ [50]. At the papillary tip at $x = L$, the concentration can have a range of values depending on whether the subject is in a diuretic (large urine production) or antidiuretic state (small urine production). The lowest interstitial osmolarity is around 600 mOsm/l and the highest is around 1200 mOsm/l.

The geometrical parameters are a little more cumbersome. According to measurements in rats by Knepper et al. [62] the radius of collecting ducts does not vary with x. We assume that the tubules form basic functional units of the kidney across species, and that we can apply the tubular diameter from rats in humans. The diameter of a collecting duct is thus taken to be $20\,\mu$m [62], corresponding to the radius $r = 10\,\mu$m.

The number of ducts per cross-sectional area varies with x. In the outer medulla this is more or less constant, and it decreases in the inner medulla. For simplicity, we have assumed a constant number of collecting ducts in our model. Knepper et al. [62] provide cross-sectional density number; $500\,\text{mm}^{-2}$ at the inner-outer medullary boundary and $200\,\text{mm}^{-2}$ at the papillary tip in rats. Edwards et al. [38] provide numbers for cross-sectional diameter of the medulla; $17.5\,\text{mm}^2$ at the inner-outer medullary boundary, and $0.54\,\text{mm}^2$ at the papillary tip. This gives about 10,000 ducts at the inner-outer medullary boundary and about 100 at the tip. To get a representative number for our purpose for a rat kidney we take 5000. But still we want a model for a human kidney. A human kidney has about 10 lobes similar to a single rat kidney [52]. So we multiply again by 10, to get to 50,000 for one kidney. Since we have two of them, we end at $n \simeq 100,000$.

The starting point for the length of the tubule L is measurements of human outer kidney dimensions [39]. The average kidney width is around 60 mm. We assume that the length of the ducts are about a quarter of this dimension, so that $L = 15\,\text{mm}$.

The inflow to the collecting is about 10 % of the glomerular filtrate. The glomerular filtrate is approximately 180 l/day corresponding to 2 ml/s [50]. This gives $Q_{W,0} = 0.2\,\text{ml/s}$. The concentration of the fluid entering the distal nephron is taken to be $c_{S,0} = 100\,\text{mOsm/l}$.

Water permeability G_W during strong antidiuresis is on the order of $750\,\mu\text{m/s}$ [68] in rat. The permeability in humans is likely to be lower in light of an overall less efficient urine concentrating mechanism that in rat. We assume that a normal value for a reference state of mild antidiuresis is around $200\,\mu\text{m/s}$ in human collecting duct. Sodium permeability values are normally around $G_S = 1.0 \times 10^{-5}\,\text{cm/s}$ [95]. Finally, the molar volume of water is $M_W = 18\,\text{cm}^3/\text{mol}$. Inserting these numbers we get reference values for the normalised permeabilities

$$g_W \simeq 0.5 \tag{4.3}$$

$$g_S \simeq 0.03\,. \tag{4.4}$$

In the following sections we will use an interval around the values in Equations (4.3) and (4.4) to investigate the effects on urine composition.

4.2 Results

The model defined in Equations (4.1) and (4.2) will be analysed in this section. First, I will look at the effects on the tubule flow of varying the water permeability. This analysis will be extended varying the solute permeability as well. Finally, I will have a look at the effects of varying the osmolar gradient.

Tubular flow and water permeability

Simulations of the model were run for different values of water permeability $g_W \in \{0.1, 0.2, 0.5, 1, 2\}$, and for an interstitial gradient from $c_{i,0} = 300\,\text{mOsm/l}$ to $c_{i,L} = 1200\,\text{mOsm/l}$. The solute permeability is $g_S = 0.03$. The results are presented in Figure 4.3. The volume flow is shown in (a), the solute flow in (b) and the tubular solute concentration in (c). The first observation is that volume flow decreases for all values of g_W. So does the solute flow, while the solute concentration increases. The interstitial gradient is the main reason for this. The gradient is the black dashed line in (c).

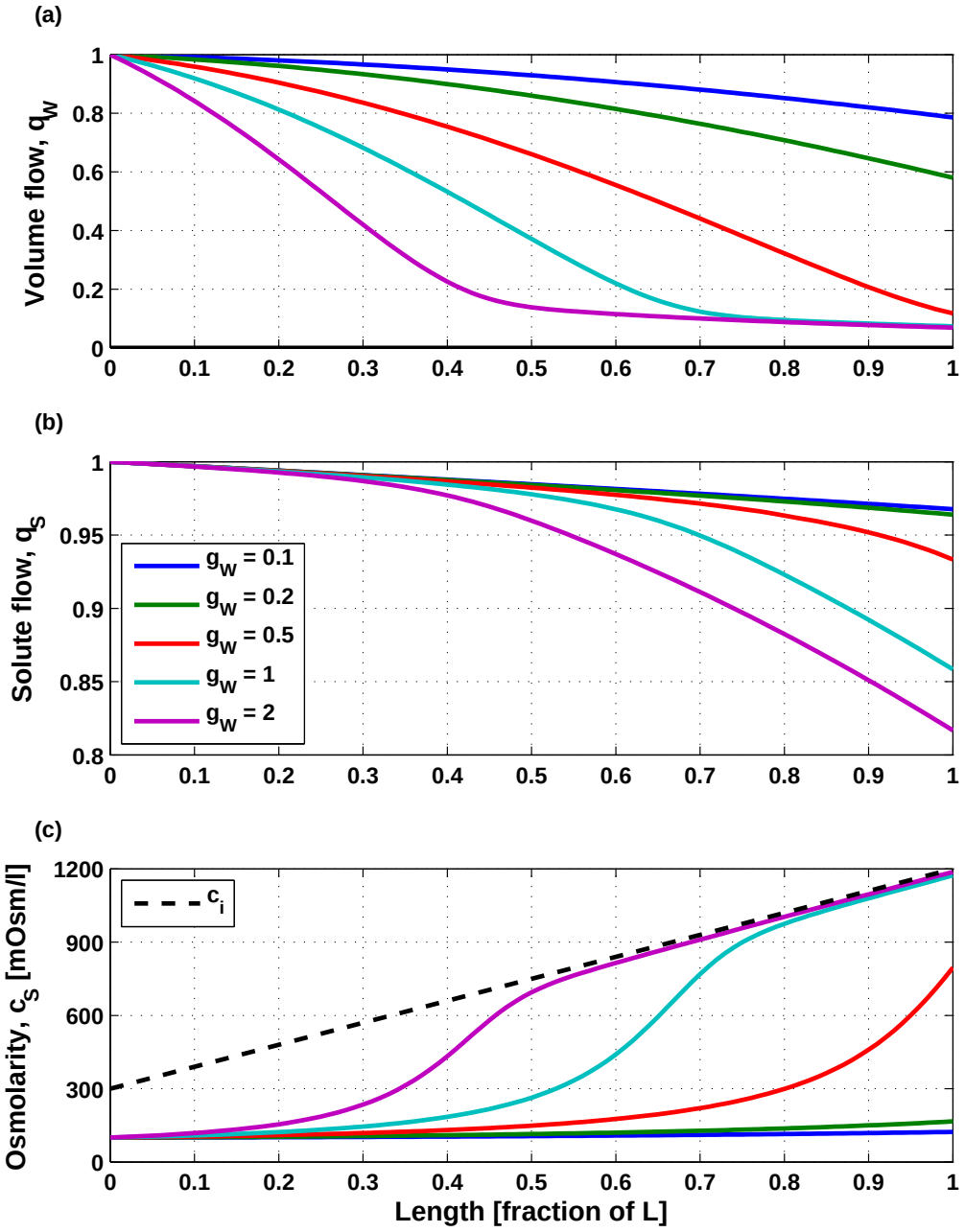


Figure 4.3: Normalised volume flow (a) and solute flow (b) and osmolarity (c) versus tubule length for different values of the normalised water permeability constant g_W. Notice in (c) how the concentration curves cling to the interstitial osmolarity curve (dashed black) as g_W increases. The solute permeability is set to $g_S = 0.03$.

As c_i increases more water is drawn from the tubule to the interstitium, and this increases the concentration in the tubule, which in turn increases the solute reabsorption, that then lowers tubule concentration to increase the osmotic potential for water reabsorption. Hence the two types of reabsorption go hand in hand. If the gradient were constant, the system would reach an equilibrium, but here it keeps increasing and the reabsorption continues.

If one compares to the flow at $x = 0$, then relatively more water is reabsorbed than solute. This just means that the fluid becomes more concentrated, which is basically what the model is supposed to simulate. The higher the water permeability the more water and solute is reabsorbed, but at high $g_W > 1$, the water flow saturates, and only additional solute reabsorption makes further water reabsorption possible. This is quite obvious, since the water flux becomes zero when the c_i and c_S are equal. If then q_S remained constant, which would happen only if $g_S = 0$ or $q_S = 0$, no further water reabsorption could occur.

One notices that three of the curves end up with approximately the same water flow (red, cyan, and purple). In terms of water flux the system is thus relatively insensitive to variations in g_W above 0.5. The final composition of the fluid flow at $x = L$ is what is interesting in terms of water and salt balance, since that is what leaves the organism. In the next section, I will look at the terminal values of the water and solute flow and concentration as a function of the permeabilities.

Urine composition

Figures 4.4 to 4.6 shows the final urine composition as a function of both g_W and g_S. The values for both permeabilities were chosen from 0 to 3-4 times there reference values in Equations (4.3) and (4.4).

In Figure 4.4, it is clear that the water flow is much more dependent on g_W than g_S. That probably does not come as a surprise, but the effect is very pronounced. For example, at $g_W = 2$ the flow is at the minimum, and increasing g_S only slightly alters the picture.

On the other hand, the solute flow shows comparable dependence on both permeabilities. The dependence on g_S is approximately linear, if one looks at a fixed g_W. This basically means that there are two ways of increasing solute reabsorption. The first is by increasing solute permeability. The second works indirectly by increasing water permeability which increase water reabsorption that again increases tubule concentration. And because the higher the tubule concentration is, the higher the transmural solute flux will be. The increase in g_W ultimately leads to an increase in solute reabsorption.

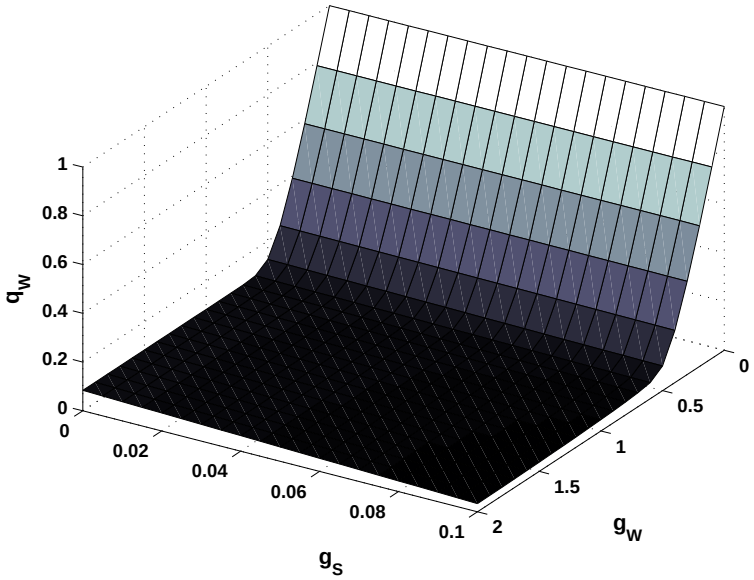


Figure 4.4: Normalised urine volume flow q_W at $x = L$ as a function of water permeability g_W and solute permeability g_S.

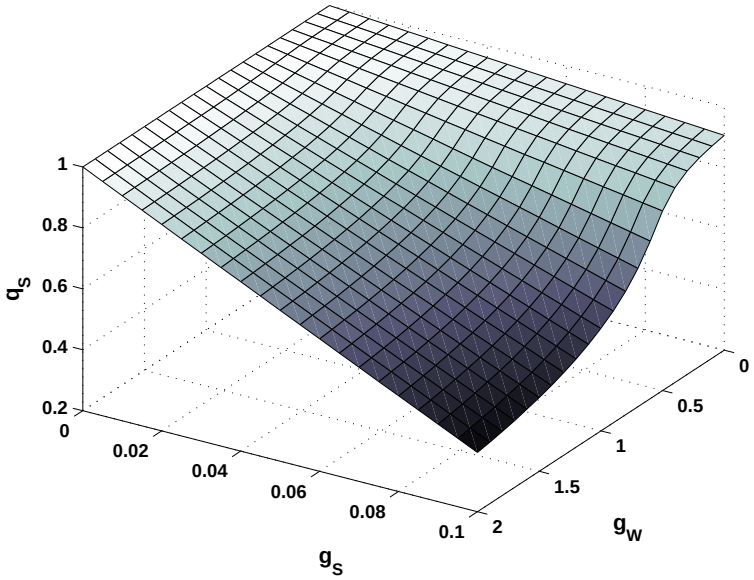


Figure 4.5: Normalised urine solute flow q_S at $x = L$ as a function of water permeability g_W and solute permeability g_S.

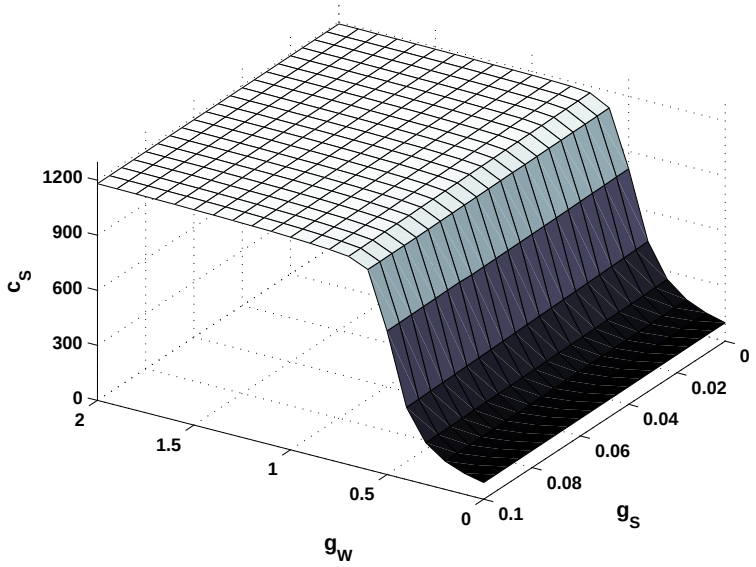


Figure 4.6: Normalised urine osmolarity c_S at $x = L$ as a function of water permeability g_W and solute permeability g_S.

Figure 4.6 is the urine osmolarity at $x = L$. Since it is the ratio of the two surfaces in Figures 4.5 and 4.4, it has similar shape to the q_W-surface. For low g_W-values it is more curved. This is because the relation water reabsorption increases faster with g_W than does the solute reabsorption. At high g_W-values the concentration saturates. This is because the interstitial concentration is a fixed parameter, and it is not possible to obtain a higher concentration than $c_{i,L}$. In the next section, I will look at what happens when we vary the interstitial osmolarity gradient.

Interstitial osmolarity

One parameter that is very important for the concentrating ability of the kidney is $c_{i,L}$. In Figure 4.7 three different values were applied to the system, namely $c_{i,L} = 600\,\text{mOsm/l}$, $900\,\text{mOsm/l}$, and $1200\,\text{mOsm/l}$. The lowest value corresponds to a diuretic state, and the highest one to an antidiuretic state [50]. The solute permeability was fixed to $g_S = 0.03$ as in Figure 4.3. The results are shown in Figure 4.7. Volume flow (a), solute flow (b), and osmolarity (c) are plotted against a range of g_W-values. The general picture is that a larger $c_{i,L}$ leads to decreased flow and increased concentration. If we take the top figure (a) and think of g_W as a control variable, one could say

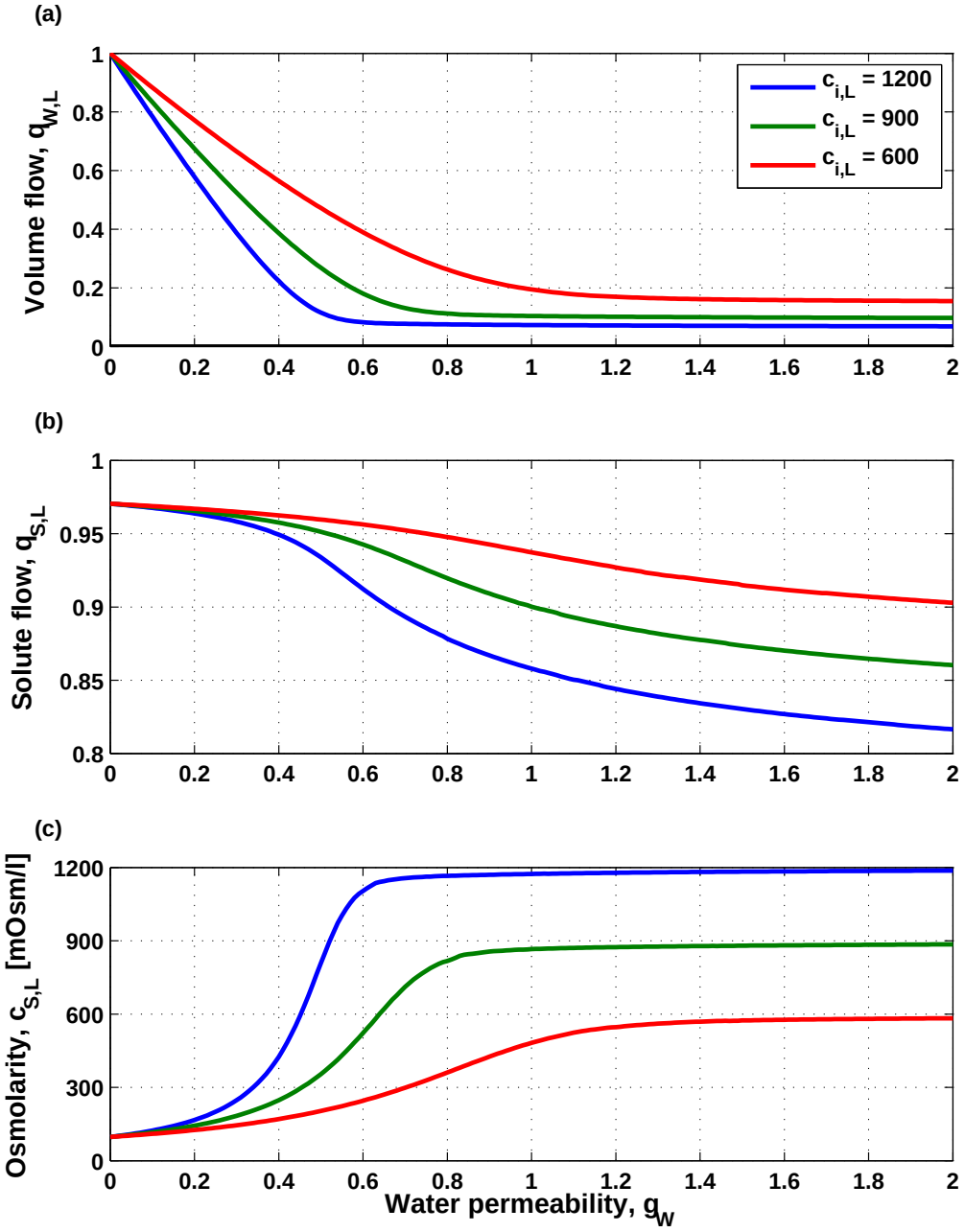


Figure 4.7: Normalised volume flow (a) and solute flow (b) and osmolarity (c) at $x = L$ versus tubule length for different values of the normalised water permeability constant g_W. The solute permeability is set to $g_S = 0.03$.

that the response of the volume flow to changes in g_W becomes stronger as $c_{i,L}$ increases due to the greater steepness of the blue curve. This means that when the interstitial concentration is high, and the body is in an antidiuretic state, then small changes is in permeabilities lead to big changes in urine flow. In a diuretic state, at low $c_{i,L}$, the change in g_W needed to make same reduction in urine flow is much greater. The two characteristic changes of the volume flow curve with $c_{i,L}$, the steepness and the lower minimum flow make physiological sense. When the organism needs to conserve water it is necessary to have a tight control system that can shut off water excretion. When water is in excess, the organism rather shuts off uptake of fluid, while slowly discharging the surplus to approach the right osmolarity and volume. The volume flow (a) and the osmolarity (c) saturate at high g_W-values. This is, however, not the case for the solute flow (b). One could interpret this as a possibility of still regulating solute balance, even though the volume flow already fixed. In this way the organism can to some extent, decouple water and salt excretion.

Another interesting point is that the reference value for water permeability $g_W = 0.5$ lies right where the action is - at the bend of the volume flow curve and the concentration curve in (a) and (c). This seems reasonable since the up- and downregulation of urine flow is must be within reach of "normal" values of the permeability. If the system was saturated at reasonable g_W-values, there would be no regulatory potential in that parameter. This also speaks in favour of the validity of the model.

4.3 Discussion

So can we trust the results from this simple model? The short answer is *yes*. It is not an anatomically exact representation of a kidney, and it is not intended to be so. The magnitudes of the different parameters are carefully chosen to represent a real human, in spite of a less complex geometry. In my opinion, there has to be some general aspects of kidney function that can be treated with models like the one at hand. Otherwise, the kidneys would not work. A paper plane is not an Airbus 380, but still it describes the point of having wings on the big one. The finer geometry of the kidney is a result of many ingenious adaptations to optimise the fluid filtration and excretion system. The fact that the estimated water permeability has a value in the regime where the system can be tightly regulated is interesting. It also suggests that there is scalability between the parameters that go into the normalisation of the permeabilities.

The conclusions that we can draw from the present model are interesting.

At high interstitial concentrations there is a threshold like behaviour of urine flow as a function of permeability. This steep response curve smoothes out for lower interstitial concentrations. Furthermore, there seems to be a possibility of decoupling solute and water flows, when water permeability is high, because of the saturation of the fluid flow. This architecture must be an advantage for the kidney as a whole. One might even postulate that the organism prioritises to fix volume first, and afterwards adjust the composition by different solute flows.

4.4 Summary

In this chapter we investigated a model of the regulation of urine flow by water and solute permeability. The model suggests that volume flow is tightly regulated mainly by water permeability, whereas solute flow is more sensitive to both solute and water permeability. The simple model gives a surprisingly realistic estimate of the regulatory range of water permeability. Furthermore it describes the characteristics of antidiuretic and diuretic states with high and low interstitial concentrations, respectively.

Chapter 5

Systemic model of osmoregulation

A systemic model of water and salt balance considers the regulatory system as a whole and incorporates the components necessary to make a sufficient mathematical description of the dynamical behaviour. In this chapter, I will introduce a model that includes the most important elements of osmoregulation by vasopressin. The parts of the overall control system that I include in the systemic model are enclosed by the gray polygon in the causal loop diagram in Figure 5.1. The main goal of the systemic model is to simulate the response to fluid intake mediated by changes in vasopressin levels and kidney water reabsorption. The membrane protein model from Chapter 3 and the kidney model from Chapter 4 will be integrated in the systemic model.

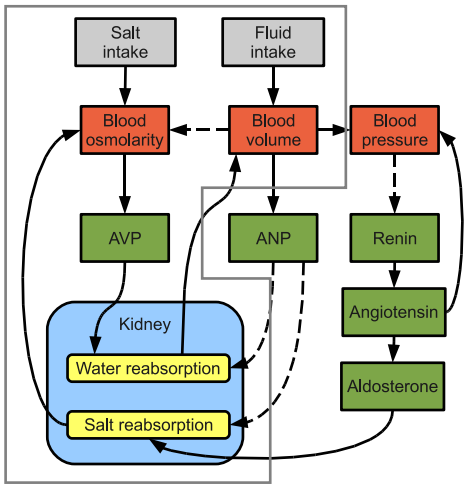


Figure 5.1: The causal loop diagram of water and salt balance presented in Figure 2.1. The model introduced in this chapter takes into account the regulation of osmolarity by vasopressin (AVP). It explicitly includes the kidney model of Chapter 4 and the membrane model of Chapter 3.

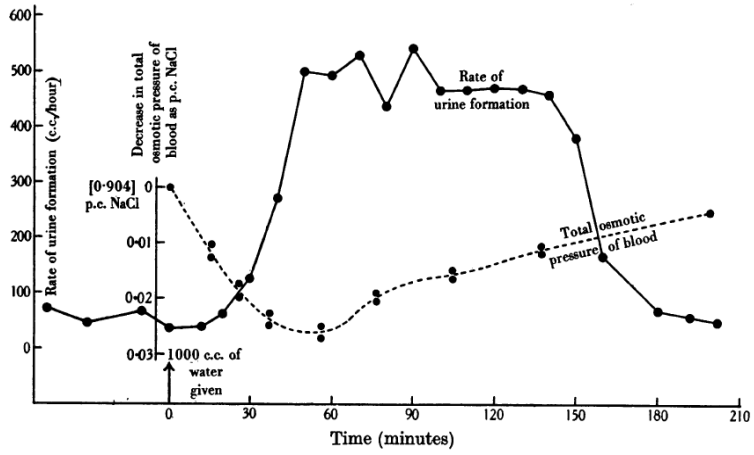


Figure 5.2: Temporal response of urine flow and plasma osmotic pressure to 1 l of water in a human subject. The figure is from Baldes and Smirk [10, Fig. 1].

The response to fluid intake was measured by Baldes and Smirk [10]. From their experiments, the response to drinking 1 l of water is shown in Figure 5.2. Immediately following the fluid intake, the body fluid becomes more dilute, and the plasma osmolarity (dashed curve) starts to decrease. It reaches a minimum at around 60 min after ingestion, and it then slowly increases until it recovers to the reference level at around 210 min. After about 30 min the urine flow starts to increase and reaches a maximum of 500 ml/h at 60 min. After 150 min urine flow decreases again to the level before fluid intake. In the following sections, I will go through the mathematical formulation of the model and argue for the choice of parameter values.

5.1 The model

The model is organised in several parts. The first part is the representation of water and solute in the fluid compartments of the organism. The second part is the modelling of absorption of water from the gut. It is also necessary to define an equation for the active substance, antidiuretic hormone. Finally, the relations that make sure that we can integrate the membrane protein model of Chapter 3 and the kidney model from Chapter 4 are presented.

Figure 5.3 is a flow diagram of the model. In the top right hand corner water and salt intake feed into the system, and a water variable V and a

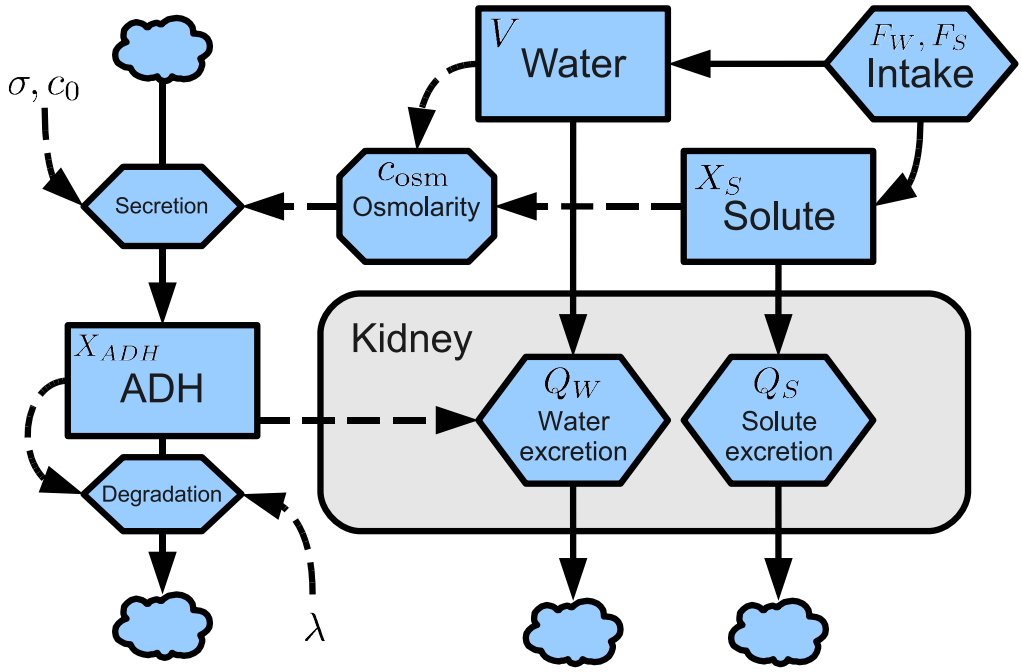


Figure 5.3: Flow diagram of the systemic osmoregulation model.

solute variable X_S represent the state of the body fluid. The osmolarity c_{osm} depends on both solute and water, and an increase in osmolarity stimulates the secretion of antidiuretic hormone ADH. In turn, ADH enhances water reabsorption in the kidney, and consequently decreases urine flow.

Water and solutes

When describing fluid in the body, it is common to distinguish between intra- and extracellular fluid. Blood plasma is part of the extracellular fluid. We will consider two fluid compartments in the model; one for extracellular fluid and one for intracellular fluid. Plasma will be a sub-compartment of extracellular fluid. Indices E, I and P denote the three compartments, respectively. We assume osmotic equilibrium between the compartments, and the equilibrium is attained rather quickly, within a few minutes, and that solely transmembrane movement of water between compartments is responsible for this [50, p. 299]. The total volume is denoted V, and the volumes of each of the compartments are denoted V_I, V_E, and V_P. A similar description of fluid compartments was used by Uttamsingh et al. [119].

The amount of solute in each compartment is denoted $X_{S,j}$, where $j \in \{I, E, P\}$. The compositions of solutes in intracellular and extracellular

spaces are of course different. We will consider the sum of osmotically active substances, such as sodium, potassium, chloride, urea, glucose, etc. as one solute variable. In the extracellular compartment the main constituent of the total osmolarity is sodium. The osmolarity c_{osm} is the concentration of solutes

$$c_{\text{osm}} = c_S = \frac{X_{S,j}}{V_j}\,, \quad j \in \{I, E, P\}\,.$$

The assumption of osmotic equilibrium leads to

$$\frac{X_{S,E}}{V_E} = \frac{X_{S,I}}{V_I} = \frac{X_{S,P}}{V_P}\,.$$

This implies that the intracellular volume V_I is given in terms of the extracellular volume V_E and solute content of both intracellular and extracellular compartments, $X_{S,I}$ and $X_{S,E}$

$$V_I = \frac{X_{S,I}}{X_{S,E}} V_E\,.$$

The total volume is $V = V_E + V_I$, and we get

$$V_I = \frac{1}{1 + X_{S,E}/X_{S,I}} V = \alpha_I V\,,$$

and

$$V_E = \frac{1}{1 + X_{S,I}/X_{S,E}} V = \alpha_E V\,,$$

where $\alpha_I = \frac{1}{1+X_{S,E}/X_{S,I}}$, and $\alpha_E = \frac{1}{1+X_{S,I}/X_{S,E}}$. The plasma volume is assumed to constitute a fixed fraction α_P of extracellular fluid volume

$$V_P = \alpha_P V_E\,.$$

The urine solute flow is created from filtrate of blood plasma, which means that it is taken from the extracellular fluid. Under the assumption of osmotic equilibrium by water movement, this leads to the intracellular solute being constant

$$\frac{dX_{S,I}}{dt} = 0\,.$$

In Chapter 4, we saw that the relative reabsorption and sensitivity to variations of permeabilities was much lower for solutes than for water. In the simulations in Section 5.2, we assume that solute content of the organism is

regulated at a much slower time scale and smaller relative magnitude than volume [16], and we set

$$\frac{dX_{S,E}}{dt} \simeq 0\,.$$

With these simplifying assumptions, we can choose any of the fluid compartments to represent the state of the volume of water in the organism, since they can all be derived from each other. The equation for the total volume V is

$$\frac{dV}{dt} = F_W - Q_W\,,$$

where F_W is the rate at which water enters the body fluid. If it is ingested directly by intravenous injection, the fluid must be added directly to V, but if given orally the fluid passes through an absorption compartment representing the gut. The gut absorption is treated in the next section.

In a normal person of 70 kg about 60 % of the body mass is water. This means that the total volume is around 42 l. Two thirds of this volume is located intracellularly, i.e. inside the cells, corresponding to 28 l. The rest is the extracellular volume of 14 l of which 3 l is blood plasma. In terms of the α's defined above, we get

$$\alpha_E = \frac{2}{3}\,,\ \alpha_I = \frac{1}{3}\,,\ \text{and}\ \alpha_P = \frac{3}{14}\,.$$

Let's assume that this 70 kg person has a plasma osmolarity of 285 mOsm/l. The solute contents are then

$$X_{S,E} = 3990\,\text{mOsm}\,,\ \text{and}\ X_{S,I} = 7980\,\text{mOsm}\,.$$

Absorption from gut

The intake F_W of water occurs through the gastro-intestinal system. This imposes a delay in the system, corresponding to the time it takes for the ingested water to make its way from the mouth, through the stomach and intestines to the body fluids. The absorption of water is modelled by a single "gut" compartment with subscript g. The equation is

$$\frac{dV_g}{dt} = I_W(t) - aV_g\,,$$

where V_g is the water volume in the gut. The rate $I_W(t)$ is the intake rate of water, and a is the absorption rate constant. The time scale of

water absorption from oral ingestion to plasma is on the order of 20 min as determined by D_2O (heavy water) tracer measurements [100]. Similar figures were suggested by Smirk [101] who used another method of weighing the legs of the subject. We set the rate constant to the inverse of the time scale

$$a = \frac{1}{20\,\mathrm{min}} = 3\,\mathrm{h}^{-1}\,.$$

The intake rate F_W in Figure 5.3 is then

$$F_W = aV_g\,.$$

Antidiuretic hormone

Antidiuretic hormone is released in response to changes in plasma osmolarity c_{osm}. The relationship between osmolarity and ADH concentration has been studied by Thompson et al. [114]. They proposed a simple model for the relationship between plasma osmolarity and ADH concentration with two parameters; a threshold osmolarity c_0 and a slope m. The plasma concentration of ADH is given by

$$c_{\mathrm{ADH}} = m(c_{\mathrm{osm}} - c_0)\,. \tag{5.1}$$

Since we implicitly model the plasma volume as well, it is more convenient to consider the total amount of ADH, X_{ADH}, rather than the concentration, and we propose the following dynamic equation for X_{ADH}

$$\frac{dX_{\mathrm{ADH}}}{dt} = \sigma(c_{\mathrm{osm}} - c_0) - \lambda_{\mathrm{ADH}} X_{\mathrm{ADH}}\,, \tag{5.2}$$

where λ_{ADH} is the first order degradation rate of vasopressin expressing the life time of the hormone. The parameter σ is the sensitivity of secretion, and c_0 is the threshold osmolarity as defined above. At equilibrium this Equation (5.2) reduces to

$$X_{\mathrm{ADH}} = \frac{\sigma}{\lambda_{\mathrm{ADH}}}(c_{\mathrm{osm}} - c_0)\,,$$

and division by the plasma volume V_P leads to

$$\frac{X_{\mathrm{ADH}}}{V_P} = \frac{\sigma}{\lambda_{\mathrm{ADH}} V_P}(c_{\mathrm{osm}} - c_0)\,. \tag{5.3}$$

Equation (5.3) is equivalent to Equation (5.1) with

$$c_{\mathrm{ADH}} = \frac{X_{\mathrm{ADH}}}{V_P}\,,$$

and

$$m = \frac{\sigma}{\lambda_{\mathrm{ADH}} V_P}\,,$$

or

$$\sigma = m\lambda_{\mathrm{ADH}} V_P\,. \tag{5.4}$$

Equation (5.4) allows us calculate the sensitivity σ from the parameters m, λ_{ADH} and V_P.

Thompson et al. [114] estimated m and c_0 to

$$c_0 = 285\,\mathrm{mOsm/kg}\,,\ \text{and}\ m = 0.41\,(\mathrm{pg/ml})/(\mathrm{mOsm/kg})\,.$$

They measured osmola*l*ity which is the amount of osmotic particles per kg of solution. Osmolarity is measured per volume, and it is approximately 2 % lower than osmolality, , measured in mOsm/l and mOsm/kg, respectively. We use a value of $c_0 = 278\,\mathrm{mOsm/l}$. They measured ADH concentration in pg/ml, and to get to pmol/l we need to multiply by the molecular weight of ADH, which is 1084 g/mol. Hence $1\,\mathrm{pg/ml} = 1.084\,\mathrm{pmol/l}$ which for our purpose represents the same numeric value.

According to Barlow [12], the plasma half life is approximately 10 min which leads to $\lambda_{\mathrm{ADH}} = 6\,\mathrm{h}^{-1}$. Assuming a plasma volume of $V_P = 3\mathrm{l}$, we arrive at a value for σ

$$\sigma = 7.5\,\frac{\mathrm{pmol}}{\mathrm{mOsm/l{\cdot}h}}\,.$$

Integration of kidney and membrane model

Another task of this modelling work is to integrate the membrane protein model and the kidney model in the systemic model. The first thing to do in this respect, is to couple membrane protein dynamics to permeability, and the second task is to couple hormone concentration to stimulus of membrane proteins.

The membrane protein model is defined in Section 3.2, Equations (3.5)-(3.7). In order to be able to compare responses for different parameter variations of the membrane model, we will normalise the variables Y_r, Y_p, and Y_m by the equilibrium value of Y_m^* given in Equation (3.10). In the subsequent investigation, this is done simply by setting the degradation

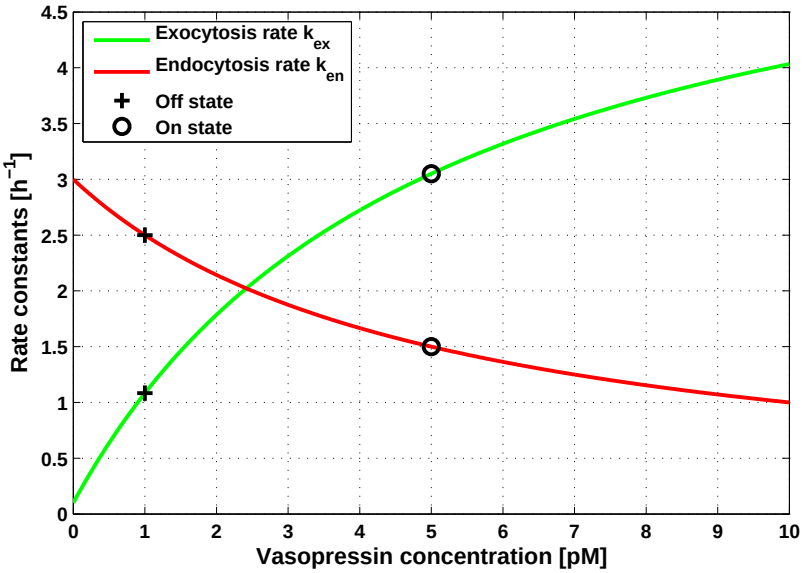


Figure 5.4: Exocytosis rate (green curve) and endocytosis rate (red curve) as functions of vasopressin concentration. The black crosses indicate the "off" state, set at $c_{\mathrm{ADH}} = 1\,\mathrm{pmol/l}$. The black circles indicate the "on" state at $c_{\mathrm{ADH}} = 5\,\mathrm{pmol/l}$, see Section 3.2.

rate of membrane protein λ_m equal to the transcription rate F_{ts}, and the translation rate k_{tl} equal to the RNA degradation rate λ_r. We investigated the effects of various λ_m-values in Section 3.2, $\lambda_m = 0.1, 0.5$, and $1\,\mathrm{h}^{-1}$.

Now in order to couple the membrane protein dynamics to the permeability in the kidney model, we assume proportionality between Y_m and the water permeability G_W. At $G_W = 200\,\mu\mathrm{m/s}$ the kidney is in a mild antidiuretic state as discussed in Section 4.1. This is the state that we will use as starting point for the simulations. We define the proportionality constant p_W, so that

$$G_W = p_W Y_m\,, \tag{5.5}$$

and $p_W = 200\,\mu\mathrm{m/s}$. When simulating the time course of the system, the urine flow will be calculated at each time point as a function of Y_m. The solute permeability is set to $G_S = 1 \times 10^5\,\mathrm{cm/s}$, the reference value used in Section 4.1.

In Section 3.2, we found that it is necessary to regulate both exocytosis rate k_{ex} and endocytosis rate k_{en}. We obtained values for "on" and "off" states in Table 3.2. According to Barlow [12] the vasopressin range of osmoregulation is 1-7 pg/ml, and we pick the "off" state to correspond to 1 pg/ml, and the "on" state to be at 5 pg/ml. The exocytosis rate increases

as a function of c_{ADH}, and the opposite is the case for the endocytosis rate. I choose a simple Michaelis-Menten function for both relationships

$$k_{ex}(c_{\mathrm{ADH}}) = \big(k_{ex,max} - k_{ex,0}\big)\frac{c_{\mathrm{ADH}}}{K_{\mathrm{ADH}} + c_{\mathrm{ADH}}} + k_{ex,0} \tag{5.6}$$

$$k_{en}(c_{\mathrm{ADH}}) = \big(k_{en,max} - k_{en,0}\big)\frac{K_{\mathrm{ADH}}}{K_{\mathrm{ADH}} + c_{\mathrm{ADH}}} + k_{en,0}\,. \tag{5.7}$$

The parameters of Equations (5.6) and (5.7) are not fitted exactly to match the results from Table 3.2. The curves in Figure 5.4 are the plots for the chosen parameter values. Black crosses indicate the "off" states at $c_{\mathrm{ADH}} = 1\,\mathrm{pg/ml}$, and circles indicate "on" states at $c_{\mathrm{ADH}} = 5\,\mathrm{pg/ml}$. For k_{ex} the transition is from "off" at $k_{ex} \simeq 1\,\mathrm{h}^{-1}$ to "on" at $k_{ex} \simeq 3\,\mathrm{h}^{-1}$. For the endocytosis rate the on-off transition is from $k_{ex} =\simeq 2.5\,\mathrm{h}^{-1}$ to $k_{en} \simeq 1.5\,\mathrm{h}^{-1}$.

5.2 Results

Simulations of the experiment by Baldes and Smirk [10] were run. To simulate the drinking of 1 l of water, the initial value for gut volume was set to $V_g = 1\,\mathrm{l}$.

Results are shown in Figure 5.5. The top figure (a) shows urine flow as a function of time, and the bottom figure (b) shows the plasma osmolarity. The blue curve corresponds to $\lambda_m = 0.1\,\mathrm{h}^{-1}$, the green curve to $\lambda_m = 0.5\,\mathrm{h}^{-1}$, the red curve to $\lambda_m = 1\,\mathrm{h}^{-1}$. Recall that λ_m is the lifetime of membrane proteins. In Chapter 3, we were not able to determine the exact value of this parameter, but now we can see what effect it has at a higher organisational level. In this plot, however, the three different λ_m-values give very similar results, and I will get back to their interpretation.

The overall characteristics of the urine flow (a) and plasma osmolarity (b) are similar to the experimental data in Figure 5.2. Especially the time scales match well, whereas the exact shape of the urine flow curve is less well replicated. Let's start with the osmolarity curve. The simulated plasma osmolarity immediately starts decreasing until about 1 h after onset of the experiment. It then slowly increases to the value before the onset of the experiment 285 mOsm/l. In Figure 5.2 the osmolarity axis is given in percent change, and at the minimum the starting value has changed about 2.5 %. In Figure 5.5(c) the minimum value attained is 280 mOsm/l, almost a 2 % change from 285 mOsm/l. The recovery slope starts steeper, and then flattens out, which is also the case in the original experiment. The recovery after 3.5 h in both experiment and simulation is up to about 0.5 % or 1.5 mOsm/l below the initial level.

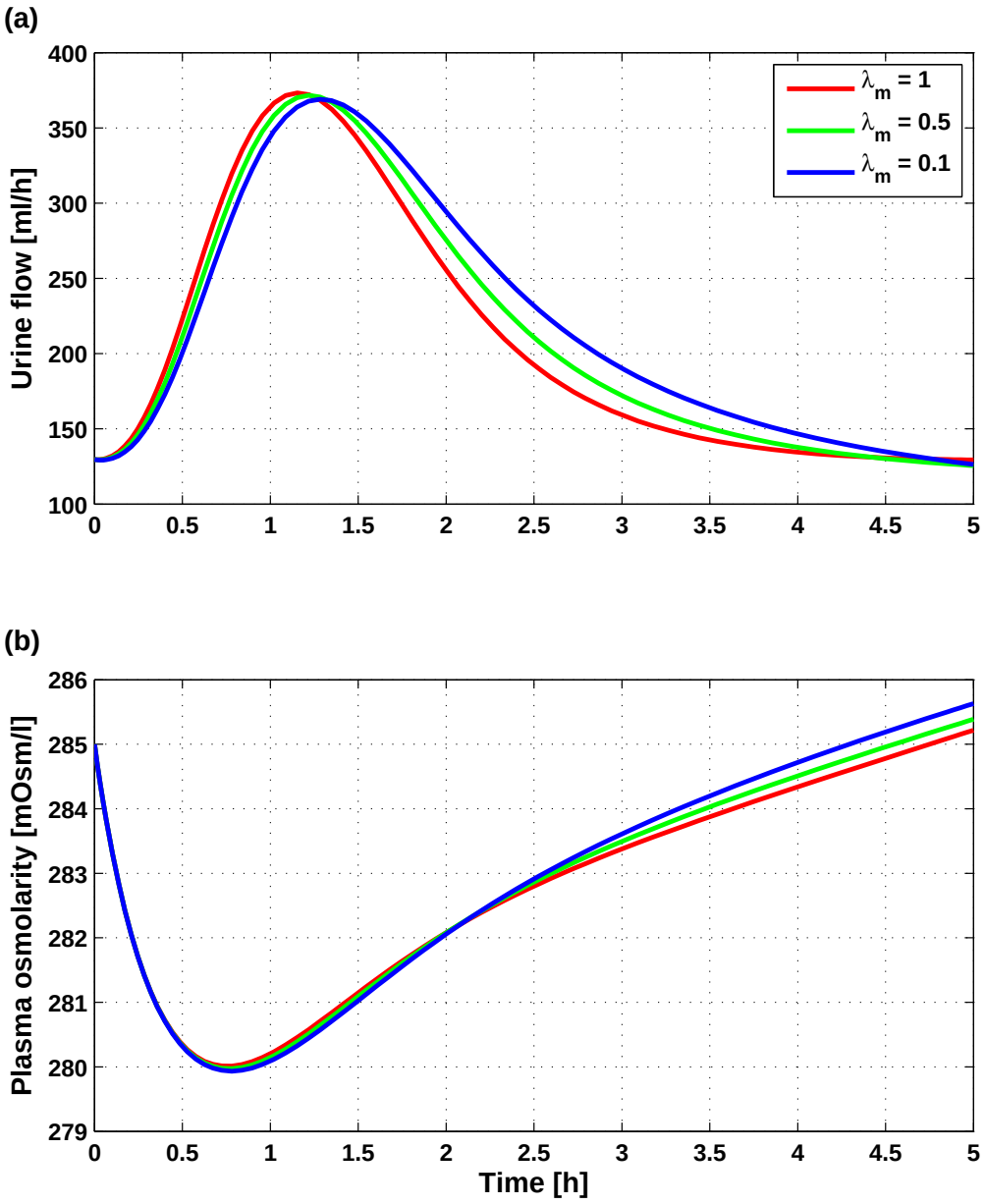


Figure 5.5: Simulation of intake of 1 l of water. Subfigure (a) is the urine production, and (b) is the plasma osmolarity. The different λ_m-values are the same as investigated in Section 3.2, and λ_m is the lifetime of aquaporins in the luminal membrane of kidney collecting duct cells. The simulation captures the dynamics of the experiment by Baldes and Smirk [10] shown in Figure 5.2. The shape of the urine flow curve is slightly different. The simulation has a sharper peak and a slower decay. This is likely to caused by a smoother inactivation of aquaporins in the model than is the case in reality.

The urine flow curve in Figure 5.5(a) initiates a little bit faster than in the experiments. The maximum as attained at about 1 h in both cases. The decrease in urine flow is quite abrupt in the experiment and much less so in the simulation. In both cases it drops to the prestimulus level after 2.5-3 h, so with respect to time scale it seems to match, but the model does not capture the steepness of the experimental response.

The minimum flow rate is lower in the experiment. The state of the subject at onset of the experiment is probably slightly more antidiuretic than what has been used in the simulation. The magnitude of the response is higher in the experiment, and this has to do with the response magnitude of the membrane protein model.

The difference between the three λ_m-curves is interesting. In (a) the red curve reaches maximum the fastest. It also decays the fastest, but faster than the increase in comparison with the two other curves. The red curve is also the slowest at recovering in terms of osmolarity (b). Basically, this is because the response in terms of total volume excreted is smaller.

In Figure 5.6, I have plotted the normalised permeability g_W (a), and the vasopressin concentration c_{ADH}(b) for the same λ_m-values as above. Recall from Equation (5.5) that permeability and Y_m are proportional, and hence the dynamics are exactly the same. The response of the membrane model is a decrease in membrane proteins. The response is nearly a mirror image of the urine flow. The red curve reacts the fastest and also decays the fastest. This is due to the internal mechanism of the membrane model, since in plot (b) one can see that the hormone curves are very similar during the period of the initial response up to 2.5 h. The bottom figure (c) is a phaseplot of the curves in (a) and (b) with c_{ADH} on the abscissa and g_W on the ordinate axis. A phaseplot is convenient for illustrating phase between two variables. It is clear that the higher λ_m, the narrower the curve in (c), and hence the shorter the phase shift. The change is not dramatic in spite of a ten-fold difference in degradation rate λ_m. This is due to the fast eigenvalue of the membrane protein model which is rather insensitive changes in λ_m, as illustrated by the approximation in Equation (3.13).

5.3 Discussion

The present systemic model successfully integrates two models from a different biological level of organisation. The simulations show the ability of the model to reproduce a physiological experiment. All parameters of the systemic model have been estimated from various sources, and it is therefore satisfactory to see that the simulation of the experiment in Figure 5.2 fits so

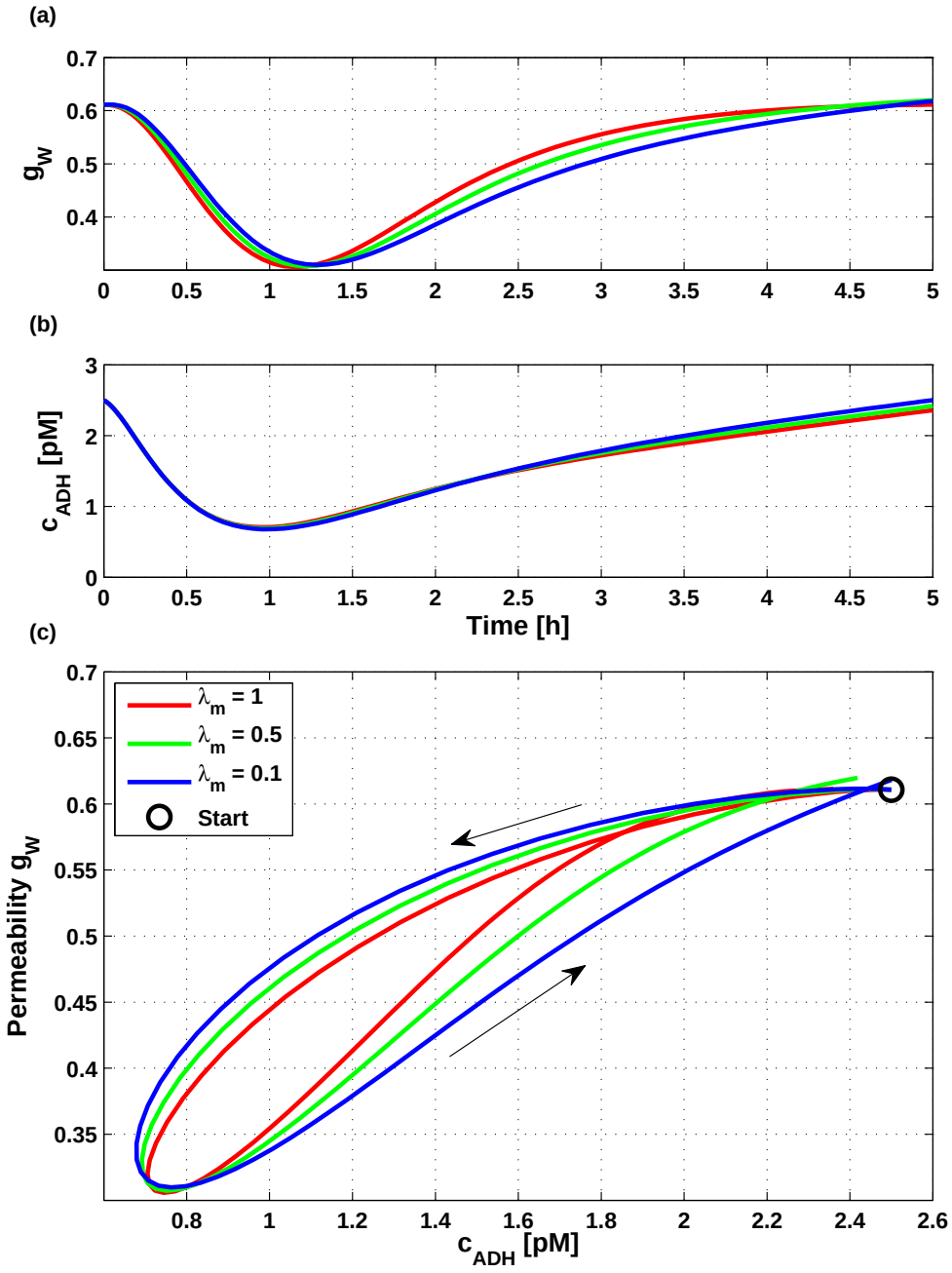


Figure 5.6: Normalised permeability g_W (a) and vasopressin concentration c_{ADH} (b) as a function of time. The phase plot of the same two variables is shown in (c). Three different membrane protein decay rates λ_m were used. The higher decay rate, the quicker the response, and the smaller the phase shift between hormone concentration and permeability.

well. The time scales of the simulation and the experiment are consistent. The magnitudes of urine flow response are a less accentuated than in the experiment.

The initial increase in urine flow in the model is, however, faster than in the experiment. The reason for this is a lack of temporal delay in the causal mechanism:

> *water absorption from gut* $\rightarrow$ *drop in osmolarity* $\rightarrow$
> *drop in vasopressin level* $\rightarrow$ *decrease in membrane permeability*
> $\rightarrow$ *increase in urine flow.*

The model does not have a delay in the mechanism that activates exocytosis and inhibits endocytosis of membrane protein. There is no delay in the adaptation of the urine flow to changes in membrane permeability either. The vasopressin stimulation is a difficult task from a modelling point of view, if the goal is to extrapolate cellular level results to tissue or organ level. The particularities of the biochemical reactions could well introduce an additional delay.

The urine flow model is assumed in steady state which is of course a simplification. However, since it is a steady state model of flows, there is an inherent time component in the model. Let's make an assessment of the time scale of the flow. Take a urine flow of $Q_W = 0.03\,\mathrm{ml/s}$ in n collecting ducts of radius $r = 10\,\mu\mathrm{m}$. The cross-sectional area of 100,000 ducts is about $A \simeq 0.3\,\mathrm{cm}^2$, and the velocity v is then $v = Q_W/A = 1\,\mathrm{mm/s}$. The time it takes to flow through a duct of length $L = 15\,\mathrm{mm}$ is about 15 s. When a change in membrane permeability one would except the flow to have equilibrated after a time comparable to the 15 s, and it is therefore reasonable to assume steady state in this part of the model.

The delay in the absorption part seems reasonable too, since the decrease in osmolarity matches the experiment very well. The delay in vasopressin release, is governed by the lifetime of vasopressin. It is known that is degraded rapidly in plasma, but still it is possible to get a larger delay by decreasing the degradation rate λ_{ADH}.

5.4 Summary

In the present chapter, I introduced a model of the regulation of osmolarity by vasopressin. The model incorporated the two models presented in Chapters 3 and 4 in a framework representing body fluid and solutes. The model successfully simulates a simple physiological experiment, where a subject drinks 1 l of water, and the urine flow and plasma osmolarity are monitored. This result confirms the validity of the modelling approach and the extensive search for parameter values from independent sources.

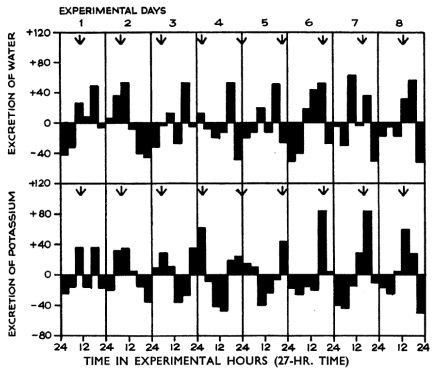
Chapter 6

The circadian rhythm

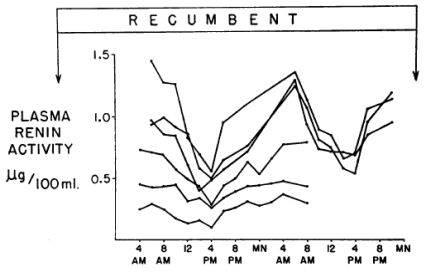
Approximately daily - that is the meaning of circadian. The circadian rhythm is an intrinsic feature of all biological systems, and it manifests itself at all biological levels of an organism. From behaviour to gene expression rhythmic oscillations of about (circa) a day (dia) are present. In this chapter I will present the topic of circadian rhythms and lay the foundation for the two studies in Chapters 7 and 8.

6.1 Rhythms at all levels

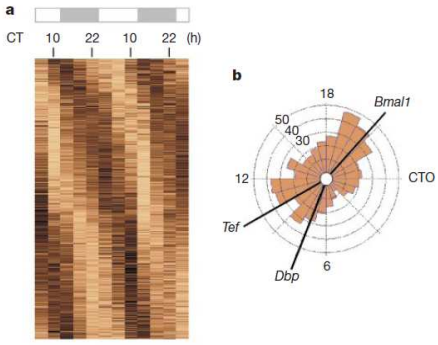
Circadian rhythms are present at all biological levels. Our planet rotates at a 24 h-period and it seems obvious that evolution would have optimised our organism to the shift between day and night. The peculiar thing is that within our organism we have an internal clock that keeps ticking even though we are not exposed to light-dark cycles. In 1957 Lewis and Lobban [72] conducted experiments with two separate groups of individuals on the arctic, Norwegian island of Spitzbergen during the summer period when the sun is still shining at nighttime. Living on strict daily feeding regime, one group was exposed to a 27 h-rhythm, while the other group experienced a 21 h-rhythm. None of the groups were aware of this difference during the study. During the study body temperature was measured, and urinary samples were taken, and the rhythms of excretion were recorded. One example of such a recording is shown in Figure 6.1(a). The top time series is water excretion and the bottom one potassium excretion. The temporal axis shows the time in terms of experimental 27 h-days. Eight experimental days correspond to exactly 9 normal days, and if one looks closely 8 peaks show up in water excretion, while 9 show up in potassium excretion. So does this mean that water excretion is controlled by external



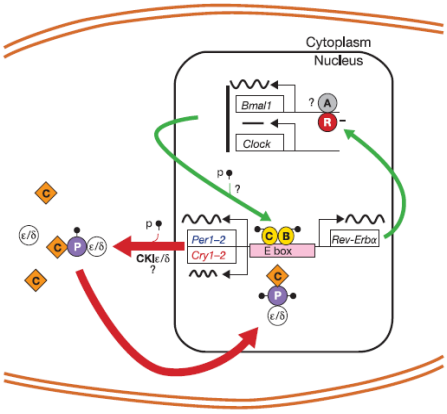
(a) Water and potassium excretion. From [72, Fig. 7].



(b) Renin activity. From [47, Fig. 1].



(c) The core clock. From [112, Fig. 1].



(d) The core clock. From [89, Fig. 1].

Figure 6.1: Circadian rhythms at different biological levels from physiology to genetics.

stimuli, and potassium excretion by some internal pacemaker that sticks to a 24 h? That is probably too much to conclude from just those experiments, but it shows that overall governing of the circadian rhythm is an intriguing interplay between external forcing and internal rhythms.

Another example of circadian rhythmicity is shown in Figure 6.1(b). It is taken from a study by Gordon et al. [47]. They measured the plasma renin activity (PRA) from healthy male subjects for up to two days. During the study the subjects were recumbent in an attempt to exclude influences due to activity rhythms. Renin activity is very sensible to changes in blood pressure and posture. The figure clearly shows dips in the afternoon and peaks in the morning.

The variables studied in these two examples clearly exhibit circadian rhythms. At the same time, they also display one of the most important difficulty of studying circadian rhythms - how do you discriminate between external applied rhythmicity and internal pace-makers. Each experiment claims to exclude one external factor - light-dark cycle and activity rhythm, respectively, but does that mean that the circadian rhythm in excretion variables or in plasma hormone concentration is solely driven by an internal mechanism? Most likely, circadian rhythmicity arises due to a combination of internal and external factors.

Clock controlled genes

An important breakthrough in the understanding of circadian rhythms has occurred within the field of genetics. Several genes in various species have been identified as components of a transcriptional-translational feedback circuit that oscillates at a frequency of about 24 h. These clock genes form the core circadian clock. The core clock is found in the suprachiasmatic nucleus (SCN) just above the optic nerves (chiasmatic) in the brain. The clock runs autonomously, and under normal conditions it is entrained by external influences of which sun light is the most important one. Research performed during the last decade has brought detailed insight about the functioning of the basal clock component which is a genetic regulatory feedback loop in the neurons of the SCN. Some of the genes that are involved are called period, cryptochrome, clock, bmal1. The protein products of period (Per) and cryptochrome (Cry) genes inhibit their own transcription, hence constituting a negative feedback loop. The Clock and Bmal1 genes function as regulators in a positive feedback loop. This is illustrated in Figure 6.1(d). The generation of a rhythm relies on the dynamic characteristics of the system. First of all there is a negative feedback loop. The processes in the loop, such as transcription, translation, protein complex formation, and nuclear-cytoplasmic transport delay the signalling by 6-12 h. This may cause the system to oscillate. Several mathematical models have proven this to be the case [15, 46].

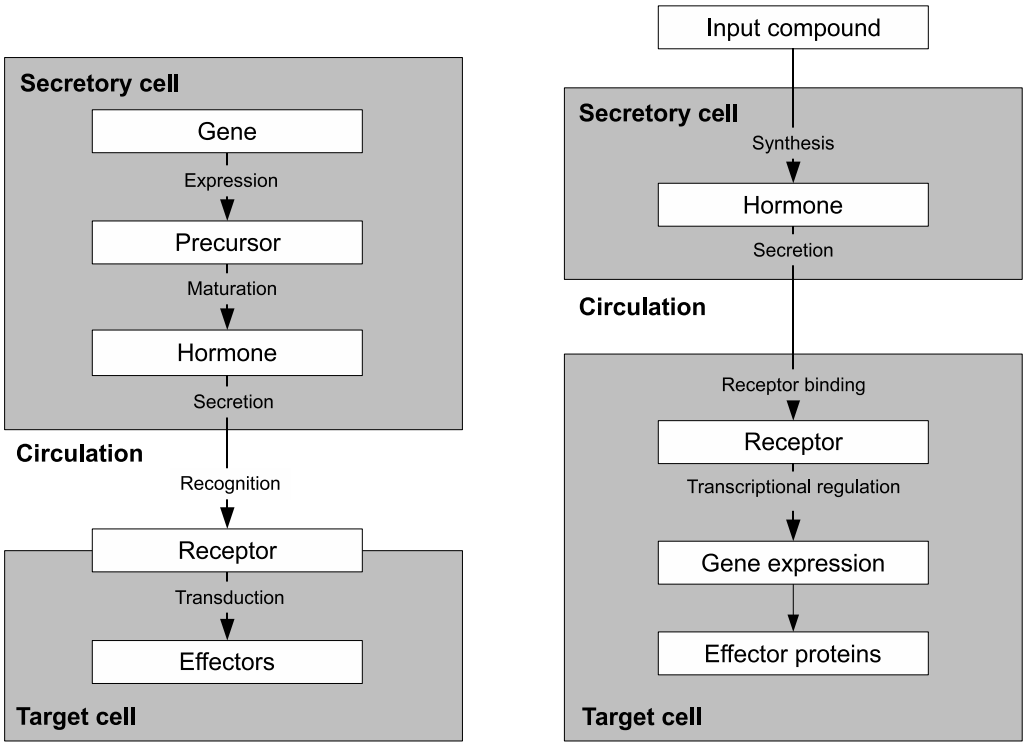
A large proportion of mammalian genes show circadian oscillations in their expression levels. Microarray data from a study by Storch et al. [112] is presented in Figure 6.1(c). The figure displays the mRNA levels (colour coded) of liver specific genes (vertical axis) as a function of time (horizontal axis). There is a clear 24 h-rhythmicity, and interestingly the phase is quite variable among the different genes. These data identify liver specific circadian genes - or clock controlled genes (CCGs) as I will refer to them.

In other organs, other genes show oscillatory behaviour, so there are clock controlled genes specific to liver, heart, adipose tissue, etc. Different authors have claimed that proportions of up to 50 % of all mammalian genes show circadian rhythms in at least one tissue [129]. Questions that arise from considering all of these rhythmic genes are; How are the expression levels controlled? Do specific transcription factors act as *circadian regulators*? In an impressive meta-study of the transcriptional regulation of clock controlled genes Bozek et al. identified several transcription factors as typical for CCGs [23]. They used bioinformatic promoter analysis to identify overrepresented transcription factor binding sites. Typical motifs found were E-box , D-box, and ROR-elements, and they identified transcription factors involved in many processes from the immune system, metabolism, and endocrine regulation underlining that complexity and integrative nature of circadian regulatory mechanisms. Core clock transcription factors, CLOCK:BMAL1, DBP, HLF, E4BP4, CREB and RORα came out as positives. Other putative circadian regulators worth a mention are SP1, WT1, AP-1, AP-2, STAT-proteins and several nuclear receptors. Nuclear receptors have been devoted much attention in the circadian field in part due to an intriguing paper by Yang et al. [131] that documents circadian variations in expression levels in most known nuclear receptors. In the study in Chapter 8, I will compare my results to the study of Bozek et al. [23].

6.2 Circadian water and salt balance

Circadian regulation is clearly present in the control of water and salt balance. Renin, the rate limiting enzyme in the renin-angiotensin system, exhibits oscillatory behaviour. Also vasopressin or antidiuretic hormone is intimately involved in generation of the circadian rhythm as it may function as a neurotransmitter in the suprachiasmatic nucleus. In general, hormonal control system are very likely candidates for the transduction of circadian rhythms from the brain to the peripheral organs [51]. When we take a closer look at the signalling cascade of a typical hormone, we will see that several intermediate steps in hormonal regulation may contribute to generating, maintaining or just transfer a rhythmic input. Figure 6.2 shows to diagrams for two types of hormones - peptide (a) and steroid hormones (b). Peptide hormones are expressed and synthesised in specific cells. They usually go through a precursor state before they are cleaved to the pure hormone. Upon stimulation they are secreted to the circulation. They arrive at a target cell, where they bind to a hormone specific receptor. Inside the cell a signal is transduced to effector molecules. Steroid

hormones are synthethised from cholesterol. The hormone is also secreted into the blood stream when the right stimulus is present. At the target the lipophilic steroids traverse the lipid bilayer of the cell membrane and bind to intracellular nuclear receptors. Nuclear receptors in complex with steroid hormones act as transcriptional regulators and affect gene transcription of target genes. The final effect of steroid hormone stimulation is an increase in effector proteins through increased transcription. In the following, I will review evidence of the circadian nature of three hormonal control systems that control water and salt balance, vasopressin, the renin-angiotensin-aldosterone system, and natriuretic peptides.



(a) Peptide hormone cascade.

(b) Steroid hormone cascade.

Figure 6.2: Schematic representation of the regulatory cascade of peptide hormone function (a) and steroid hormone function (b).

Vasopressin

Vasopressin is a small peptide molecule - a nonapeptide with nine amino acids. It is secreted from the posterior pituitary gland. The main physiological function of vasopressin is to increase water reabsorption in the distal

tubule and collecting ducts in the kidney. It also has vasoconstrictive effects, hence its other name *vaso-pressin*. It is well known that the plasma level of vasopressin has a circadian rhythm with higher levels at night. This conveniently decreases urine output [93].

The regulatory cascade of vasopressin is represented in Figure 6.2(a). The cascade initiates with synthesis by gene expression in magnocellular neurons of the supraoptic and paraventricular nuclei of the hypothalamus. The gene that expresses vasopressin is called *Avp* in mice. The product of the gene expression is a precursor or prohormone called neurophysin II which contains the vasopressin peptide. Other parts of the neurophysin II molecule have no apparent function. The axons of the magnocellular neurons (MCN) project to the posterior pituitary gland. By axonal transport the prohormone is transported in vesicles to this site. Enzymes then cleave the neurophysin molecule, and upon stimulation, vasopressin is released in to the blood stream by exocytosis.

At target sites throughout the organism, the hormone binds to receptors. Three receptors for vasopressin are known; V1R, V2R, and V3R. All receptors are membrane-bound G-coupled receptors. This means that their intracellular effects are mediated by second messengers within the cell. V1R is mainly located in smooth muscle cells, and receptor binding results in vasoconstriction. V2R is located in the basolateral membrane of kidney distal tubule and collecting duct cells. Binding results in insertion of aquaporins in the apical membrane, and hence an increase in water reabsorption. Insertion of urea transporters are also affected by vasopressin binding. V3R receptors are located in variety of cells, for example in ACTH releasing neurons, where binding stimulates ACTH secretion.

In the regulatory cascade of vasopressin, circadian regulation is likely at several levels. The gene expression of vasopressin itself may be regulated. Also the secretion process or the secreting stimuli may be circadianly regulated. The receptor abundance could be rhythmic. The receptor level would be dependent on genetic expression of receptor protein. Finally, the effectors of signalling cascades triggered by vasopressin could be circadianly regulated.

The promoter of the vasopressin gene has several putative circadian regulatory sites [26]. Possibly through the protein kinase A (PKA) pathway and the CREB/cAMP-responsive promoter. Jin et al. [56] report effects of CLOCK and BMAL1 on the vasopressin gene. They also found that the gene is expressed differentially in the SCN where it is rhythmic, and in the SON, where it is not rhythmic.

Renin and angiotensin

The renin-angiotensin system plays a crucial in control of blood pressure and kidney function. Renin is released upon changes in arterial blood pressure, and it functions as a cleaving enzyme that converts the large protein angiotensinogen, abundantly present in the blood, to angiotensin I. A second catalytic reaction cleaves angiotensin I to angiotensin II which is a potent vasoconstrictor. This process is catalysed by the angiotensin converting enzyme simply known as ACE. We already discussed the diurnal renin activity found in 1966 by Gordon et al. [47]. But circadian rhythmicity is not limited to renin. According to Cugini and Lucia [31], all the components of the renin-angiotensin-aldosterone system (RAAS) have marked circadian rhythms with differing phases.

The gene of the renin-angiotensin system that has been mostly investigated is the renin gene, due to the its high expression levels in human embryonic kidney cell lines (HEK cells). TNF-alpha inhibits renin expression, and the responsible mechanism depends on NF-kappaB and binding to a cAMP responsive element (CRE), possibly by interaction with CREB1 [116, 117]. Renin gene expression is also stimulated by PPAR-gamma [115]. Also WT1 is involved in regulation renin expression [107].

Renin is released into the blood stream in secretory vesicles [97]. Release is stimulated by a decrease in renal perfusion pressure. The enzymatically inactive precursor prorenin is also released into the blood plasma, and its concentration is about 10 fold higher than active renin. Prorenin has also been reported as a circadian variable [32].

In recent years, a (pro)renin receptor has been identified. Evidence exists that by binding to the receptor prorenin becomes enzymatically active. Furthermore receptor binding to both prorenin and renin elicits intracellular, angiotensin-like effects [33]. These findings are still very new.

Angiotensin II activates STAT pathways in renal mesangial cells [98]. STAT proteins were also among the identified circadian regulators. Through angiotensin type 1 receptor STAT5 DNA binding was induced, and through receptor type 2 STAT1 DNA binding. PPAR activators (nuclear receptors) are known to be negative regulators of the AT1 receptor gene [27].

Reports of the influence of the renin-angiotensin-system on erythropoeisis suggest that an activated renin-angiotensin system has a stimulatory effect on EPO levels [58]. EPO has also been identified as a clock controlled gene [23].

Aldosterone

Aldosterone is secreted from the cortex of the adrenal glands. It is secreted in response to angiotensin II, ACTH, and to changes in potassium concentration. It is a mineralocorticoid steroid hormone. The term *mineralocorticoid* stems from its involvement in regulating mineral homeostasis, mainly sodium and potassium balance. Intracellularly, aldosterone binds to the mineralocorticoid receptor (MR), a nuclear receptor (NR3C2). The steroid-receptor is then transported in to the cell nucleus where at affects gene-expression and protein synthesis of proteins, e.g. epithelial sodium transporters and sodium-potassium exchangers.

Circadian variations of aldosterone and angiotensin II were observed at the plasma level along with influences on blood pressure and urinary excretion [94]. The same reference also provides clear evidence of circadian variation in aldosterone and angiotensin II levels. Charloux et al. also report circadian rhythms of aldosterone levels, and a clear correlation with plasma renin activity [29]. Their study indicates a strong dependence on the pattern of the sleep-wake cycle.

Aldosterone is a steroid hormone and is synthesised in the adrenal gland as an end product of a series of catalysed transformations of cholesterol. The specific enzyme in aldosterone synthesis is called aldosterone synthase or CYP11B2 [134]. CYP11B2 knockout mice have been developed and their renal function was studied by Makhanova et al. [70, 74]. The mice show impairment of Na^+-reabsorption in the distal nephron and hence salt wasting, decreased blood pressure, and consequent strong renin-angiontensin system activation. Bassett et al. [13] have revealed the involvement of several nuclear receptors in CYP11B2 expression. The steroidogenic factor-1, SF-1, which is known to stimulate steroid synthases, inhibits or has little effect on CYP11B2 expression while it stimulates CYP11B1 expression, which again catalyses the final step of cortisol synthesis [14]. Two other factors, nerve growth factor-induced clone B (NGFIB), Nur-related factor 1 (NURR1), were also reported to stimulate expression of human CYP11B2 [13]. They may be involved in Ang II triggered pathways. The above mentioned factors all show circadian expression patterns according to the nuclear receptor study by Yang et al. [131]. There is evidence that the aldosterone receptor MR is expressed in a circadian fashion in muscle tissue [131]. The gene expression of MR is upregulated by Sp-1 [65]. The primary genetic targets of the MR-aldosterone complex are the Sgnn1 genes that code for the ENaC sodium channel, the serum- and glucocorticoid regulated kinase 1 (sgk1), the corticosteroid hormone-induced factor (CHIF), and also the k-ras gene [42]. MR-mediated transcription is inhibited by NF-kappaB [66].

Natriuretic Peptides

Natriuretic peptides influence water and salt homeostasis by inducing mainly excretion of sodium (natriuresis), vasodilation and lowering of blood pressure. Three types of natriuretic peptides have been discovered - ANP, BNP and CNP. There are also three types of receptors which are NPRA, NPRB, NPRC. NPRA and NPRB are guanylyl cyclase receptors. Binding to them increases intracellular cGMP levels. NPRC binding inhibits adenylyl cyclase activity. ANP and BNP bind primarily to NPRA [80].

Circadian variation in ANP plasma concentration has been documented by Rittig et al. [92] with peak times around midnight. Expression of natriuretic peptides mainly takes place in the atria, but also in the ventricles, pituitary, lungs, hypothalamus, and kidneys, but levels are much lower ($< 1\%$ of total) than in the atria [45, 77]. The ANP promoter contains binding motifs for AP-1, AP-2, CRE, SRE/CarG, Egr-1 [54].

6.3 Summary

Circadian rhythms are definitely a major component of water and salt balance. From excretion variables to gene expression profiles, 24 h-rhythms show their presence. All of the reviewed hormonal components show circadian plasma variations. Several of the previously identified circadian transcriptional regulators are already known to have effects on genes coding for regulatory proteins of water and salt balance.

In the next two chapters, I will approach circadian rhythm in two different ways. In Chapter 7, I will analyse clinical data from patients with varying fluid intakes. In Chapter 8, I will undertake a bioinformatic promoter analysis and look for transcription factors, possibly circadian, common to genes involved in water and salt balance.

Chapter 7

Analysis of circadian data

Physiological time series are a great source information for the study of circadian rhythms. With sufficient temporal resolution and duration of experiments it is possible to identify circadian components of measurements from blood plasma and urine samples. Since the circadian rhythm is a daily oscillation, it is necessary to have a time series with at least 24 h duration, and preferably more. It is also important to have samples collected several times a day to capture the circadian oscillation.

We investigate the effects of increased fluid intake on urine output and blood plasma variables based on clinical data obtained from young, male volunteers. One group was instructed to have a normal fluid intake (30 ml/kgBW/day) and another to have a high fluid intake (60 ml/kgBW/day). The phenomena of interest are indicated in the causal loop diagram in Figure 7.1 by the gray closed curve. Fluid intake is altered and the effects on AVP, ANP, blood plasma, and urine production and their circadian rhythms will be investigated.

The data presented in this chapter was kindly provided by Professor Jens Christian Djurhuus, Skejby Hospital.

Motivation

It is well known that urine production varies between day and night. This is often explained by the control of circadianly oscillating hormone vasopressin (AVP). Subjects suffering from nocturnal enuresis (bedtime wetting) have been shown to have malfunctioning circadian rhythmicity of vasopressin [91, 93], indicating the importance of this type of hormonal control mechanism. As discussed in Chapter 6, another hormone involved in fluid and salt regulation, atrial natriuretic peptide, has also been shown

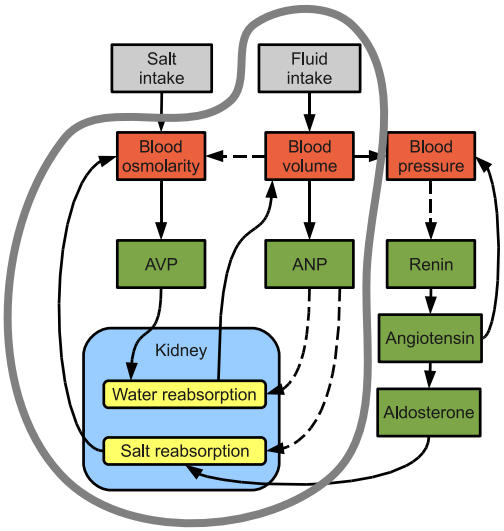


Figure 7.1: The causal loop diagram of water and salt balance presented in Figure 2.1. The parts dealt with in this chapter are enclosed by the gray curve. The hormones AVP and ANP are measured, along with blood volume, osmolarity, and a number of urine variables. The available data will allow us to investigate the response to different fluid intakes.

to oscillate in a circadian manner. In this study, these two hormones will be monitored during two different fluid intake regimes.

A wide variety of analysis techniques will be applied. A first assessment of diurnal variability will be done by statistical comparisons of day-night averages. Diurnal means day to night variations, so it is in principal a special case of circadian rhythmicity with a specific phase. To take into account the temporal nature of the data a technique called repeated measures analysis of variance will be applied to obtain information about differences of temporal behaviour between the two groups. Finally, a so-called Fourier score method will be used to look at the circadian characteristics of the data from each individual.

7.1 Data

Data was recorded from 29 healthy, young male volunteers. They were divided in two groups of 17 and 12 subjects. The first group was instructed to have a fluid intake of 30 ml/kgBW/day. This group will be referred to as the normal fluid intake (NFI) group. The other group was instructed to drink 60 ml/kgBW/day. This group will be called the high fluid intake (HFI) group. Measurements were recorded in a hospital during two days - 48 hours. Subjects were instructed to drink 2/5 of the prescribed fluid intake between 8.00 and 14.00 and again from 14.00 to 20.00. From 8.00 to 23.00 the final 1/5 of was consumed. Standard hospital meals were served at 8.00, 12.00 and 18.00. Prior to the recording period, subjects were instructed to obey their respective fluid intake regimens for three days, in order to assure that they would be in a "circadian steady state", before the initiation of the

experiments. The subjects also had a catheter installed on one of the two nights for the purpose of another study of effects of catheterisation. This was reported not to have any effect on the measured variables. During the night with catheter no nocturia incidents (involuntary nighttime urination) were recorded, since the bladder was automatically emptied.

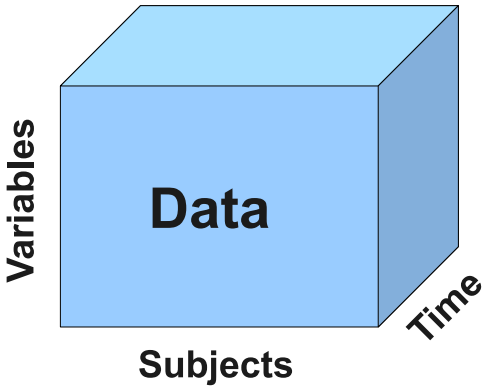


Figure 7.2: Schematic presentation of the type of data.

A simple schematic representation of the type of data in this study is shown in Figure 7.2. Basically the data is 3-dimensional. The first dimension is the number of variables measured. The second dimension is the number of individual subjects from which measurements were taken. The third dimension is time. The temporal dimension is essential for understanding the dynamics of the system. The analysis techniques applied in this study attack the cube from different sides to extract information about the underlying system. Basically, each technique modifies the data axis in different ways in order to infer system behaviour. One example is the grouping of individual subjects on the subject axis. Another one is the averaging of day and night values on the time axis. This way of considering the data allows a us to compare the scope and the results of the different techniques.

Measured variables

A number of quantities were measured in urine and in blood plasma. For urine the volume (Uvol), osmolarity (Uosm), sodium concentration (UNa), potassium concentration (UK), and creatinine concentration (Ucrea) were measured. In blood plasma, osmolarity (Posm) as well concentrations of sodium (PNa), potassium (PK), creatinine (Pcrea), albumine (Palbu),

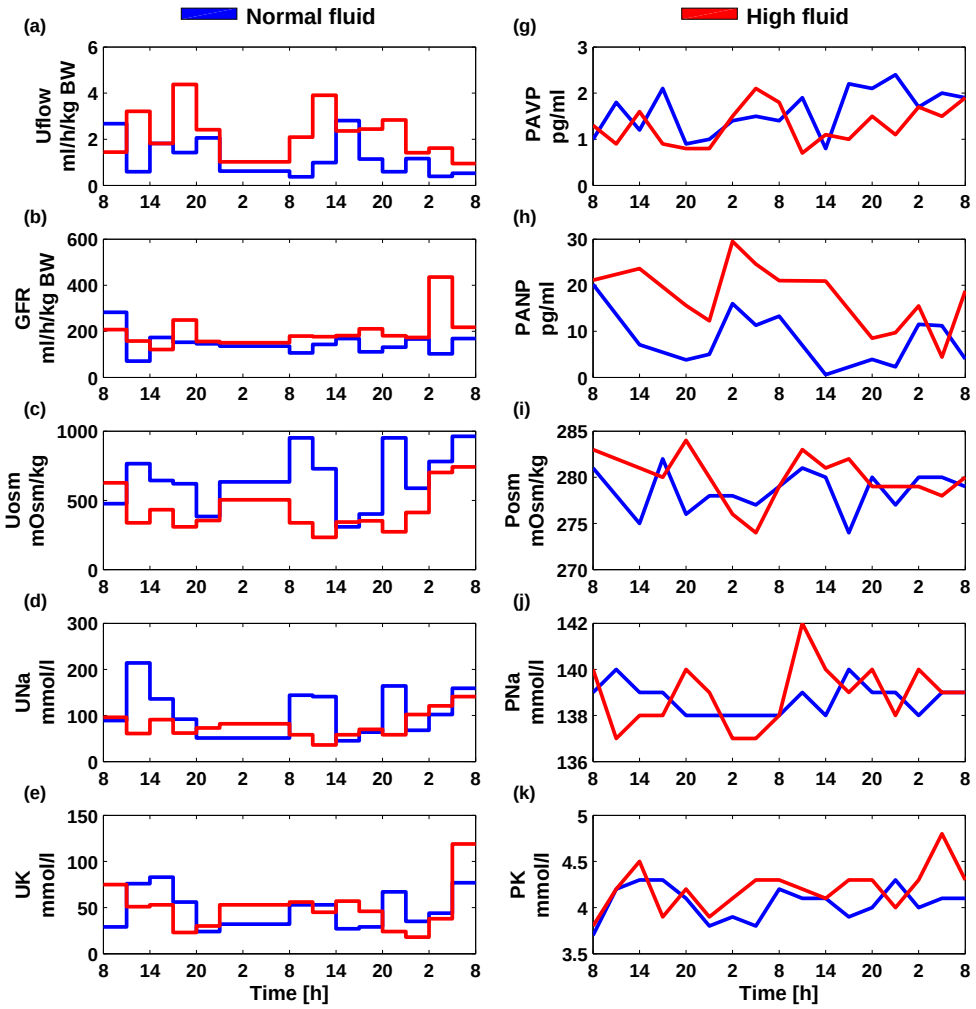


Figure 7.3: Example data from one subject in the normal fluid intake group, and one subject in the high fluid intake group.

and the hormones atrial natriuretic peptide (PANP) and arginine vasopressin were measured (PAVP). All plasma variables were measured during a 48 hour period at 3 hour intervals. Urine samples were also collected every 3 hours, except for one of the two nights, where the urine samples were collected during the whole nighttime period from 23.00 to 8.00. This was the night where the catheter was not installed.

Figure 7.3 shows a selection of the measured variables for one subject in the normal fluid intake group (blue), and on subject in the high fluid intake group (red). In (a) urine flow is shown, and it is clear that the high fluid intake group has higher average flow. The glomerular filtration rate (GFR) is shown in (b), urine osmolarity (Uosm) in (c), urine sodium concentration in (d), and urine potassium in (e). The second column depicts plasma variables; AVP concentration in (g), ANP concentration in (h), plasma osmolarity in (i), and plasma sodium and potassium in (j) and (k). It is not obvious from this figure what to conclude or hypothesise neither about differences between the two groups nor about the temporal characteristics of the individual time series.

Derived variables

From the measured variables, one can calculate additional variables. The urine flow is simply the volume divided by the duration of the collection interval, ΔT,

$$U_{\text{flow}} = U_{\text{vol}}/\Delta T \, .$$

Excretion rates of osmoles, sodium, potassium, and creatinine, can be calculated as urine concentration times urine flow, i.e.

$$E_x = U_x U_{\text{flow}} \, ,$$

where x represents any of the above mentioned variables. Clearance C_x of a substance x is calculated as

$$C_x = \frac{U_x}{\langle P_x \rangle} U_{\text{flow}} \, ,$$

where $\langle P_x \rangle$ is the averaged plasma concentration during the urine collection interval. The glomerular filtration rate is taken to be equivalent to the clearance of creatinine which is standard practice [121].

Fractional excretion gives a measure of how much of the filtered load is excreted. It is the excretion rate divided by the plasma concentration times

the glomerular filtration rate. This reduces to (urine flow cancels out)

$$FE_x = \frac{U_x \langle P_{\text{crea}} \rangle}{\langle P_x \rangle U_{\text{crea}}} .$$

Solute free water clearance $C_{\text{H}_2\text{O}}$ is the difference between urine flow and osmolar clearance

$$C_{\text{H}_2\text{O}} = U_{\text{flow}} - C_{\text{osm}} .$$

The derived and the directly measured variables will be analysed in the following sections. In total there are 18 urine variables (collected in time intervals) and 7 plasma variables (measured instantaneously.

7.2 Analysis techniques

Three approaches will be used to analyse the data. First a simple statistical approach to asses diurnal (day-night differences) characteristics of the data. In reference to Figure 7.2 this technique groups data on both the subject axis and the time axis. The next technique *repeated measures analysis of variance* considers the temporal organisation of the repeated measures between the two groups of patients. Again the subjects are grouped, but the temporal information is treated with more care. The Fourier score technique considers subjects individually. This approach is, in principle, better suited to identify causal relations and to identify outliers.

Diurnal statistics

Day and night averages were calculated for the two groups. For the urine variables, day-time was considered to be the measurements from 8.00-11.00, 11.00-14.00, 14.00-17.00, 17.00-20.00, and 20.00-23.00. For the plasma variables day-time average values were calculated from the 11.00, 14.00, 17.00, 20.00 measurements. Student's t-test was used to compare values between the two groups, and a significance level of $p = 0.05$ was used. Whole day average were also calculated.

Repeated measures analysis of variance

Repeated measures analysis of variance (RMANOVA) was applied to investigate, in more detail, the temporal variation in the data. There are many variants of the RMANOVA, and here we will use a multiple sample RMANOVA with time effects, group effects and interaction between

time and group as presented by Davis [34, Chapter 5]. The well-established ANOVA framework allows for statistical testing of each of three effects and, hence, there are three p-values; for temporal variation p_T away from a mean level, differences in mean between groups p_G, and for the difference of temporal profile of the mean curves of the two groups, the so-called group-time interaction p_{GT}. In the latter test, the less parallel the mean curves of the groups are, the more significant the test. Since the ANOVA test demands a complete dataset with no missing values, such values were replaced by interpolated values. This practice is common for this type of analysis with relatively few missing values, but it does, of course, introduce spurious elements in the data.

Fourier scores

The repeated measures analysis of variance analyses temporal characteristics of groups. By assuming a grouping of subjects, we cannot extract information about the individual subject, and how measured variables relate for each of them. I will introduce a technique called *Fourier score* to perform an individually based analysis of circadian components of the measured plasma and urine variables. A similar approach has been used to analyse cell cycle data [43] though with a different background model.

The name Fourier score stems from the fact that we consider the 24-hour component of the Fourier spectrum of the time series. We need to consider two different types of data. The first one is instantaneously measured signals that represent a momentary state of the system. This is, for example, the blood plasma variables. The urine variables, on the other hand, are collected during a period of time, and hence the measured quantity is equivalent to an integral of an underlying signal. In the following, I will derive expressions for the Fourier score for both types of signals.

The score from each signal does not give us information about whether it is high or low. To evaluate this, each score is compared to an empirical background distribution, and a P-value is calculated as the corresponding percentile. See Figure 7.8 for an illustration. The way of generating background sequences carefully takes into account the fact that we wish to test for temporal coordination in the data by random permutation of the order of the points.

Instantaneously measured signal

The instantaneously measured signal is the more straight forward to treat. For a signal $p(t)$ the Fourier score is defined as

$$FS\big[p(t)\big] = \sum_i \exp\left(i2\pi\frac{t_i}{T}\right) p(t_i)\Delta t_i\,, \tag{7.1}$$

where T is the considered period, here $T = 24\,\mathrm{h}$. FS is a complex number, and it contains information both about the magnitude of the oscillatory component and its phase. In order to compare the magnitude of the score across different measured variables, the signal is first normalised by subtracting its mean and dividing by the standard deviation

$$\hat{p}(t) = \frac{p(t) - \bar{p}(t)}{\sigma_p}\,,$$

where $\hat{p}$ is the normalised signal, $\bar{p}$ is the mean, and σ_p is the standard deviation.

Integrated signal

In the case of the urine signals, consider an instantaneous urine flow rate $u(t)$, the total collected urine volume is the integral of the flow rate

$$U(t) = \int_{t_0}^{t} u(t)dt. \tag{7.2}$$

Consider the discrete collected samples U_i collected during intervals $\Delta t_i = t_i - t_{i-1}$ for $i = 1 \ldots N$. Then $U(t)$ can be written as

$$U(t_i) = \sum_{j=1}^{i} U_i\,. \tag{7.3}$$

We want to know if the instantaneous signal $u(t)$ has a 24 h-component like in equation (7.1). In order to deal with the integrated variable, we start by considering the continuous Fourier component as calculated by the following integral over the time interval of the data series t_0 to t_N

$$FS\big[u(t)\big] = \int_{t_0}^{t_N} \exp\left(i2\pi\frac{t}{T}\right) u(t)dt\,.$$

Integration by parts gives

$$\begin{aligned} FS &= \left[U(t)\exp\left(i2\pi\frac{t}{T}\right)\right]_{t_0}^{t_N} - i2\pi\int_{t_0}^{t_N}\exp\left(i2\pi\frac{t}{T}\right)U(t)dt \\ &= \underbrace{U(t_N)\exp\left(i2\pi\frac{t_N}{T}\right)}_{I} - i2\pi\underbrace{\int_{t_0}^{t_N}\exp\left(i2\pi\frac{t}{T}\right)U(t)dt}_{II}\,. \end{aligned} \tag{7.4}$$

Part I in equation (7.4) was derived using equation (7.2). Part II in equation (7.4) can be approximated by considering the discrete representation of $U(t)$ in equation (7.3)

$$II = \sum_{i=1}^{N}\exp\left(i2\pi\frac{t_i - \Delta t_i/2}{T}\right)U(t_i)\Delta t_i\,.$$

The reason for shifting the complex exponential by $\Delta t_i/2$ is to minimise the phase shift introduced by the discretisation. With this derivation it is also possible to calculate Fourier scores for integrated signals, and it will be applied to the urine variables.

Background distribution

The background sequences are simply generated by permutation of the temporal order of the points. Let I denote the ordered set $0, 1, 2, ...N$, and consider the ordered sequence (time series)

$$p_i = p(t_i) \quad \text{for } i \in I.$$

Now let I_{bg} be a random permutation of I, i.e. a set with the same elements but different order. A background sequence $p_{bg,i}$ is simply

$$p_{bg,i} = p(t_i) \quad \text{for } i \in I_{bg}.$$

In this study 10,000 background series are generated, and for each of these the Fourier score is calculated. The 95^{th} percentile of the empirical distribution of the background Fourier scores is the 5 % significance level.

7.3 Results

The results from the large data set are presented in this section. They will be presented in the same order as the methods with some extra analyses and tests that give additional information.

Nocturia nights

As would be expected the number of nocturia nights is nearly double in the high fluid intake group, 11 nights, compared to the normal fluid intake group, 5 nights. In the high fluid group this means that during 92 % (11/12)) of all measured nights a nocturia incidence occurred, whereas the figure for the normal fluid intake group was 29 %.

Average comparisons

Mean levels were calculated for all variables in both groups. Results are shown in third and sixth columns (titled '24-h') in table 7.3. Values are mean±SE, and significantly different results are shown in bold type ($p < 0.05$). Not surprisingly, urine flow was significantly higher in the high fluid intake group, 2.32 ml/h/kgBW versus 1.2 ml/h/kgBW in the normal fluid intake group. All measured urine concentrations (osmolarity, sodium, potassium, creatinine) were significantly lower in the HFI group. There were slightly higher osmolar and potassium excretions in the high fluid intake group. This was also the case for clearance and fractional excretion of osmoles and potassium. No differences in these excretion variables were observed neither for sodium or nor for creatinine. Interestingly, glomerular filtration rate was similar in the two groups, in spite of the different fluid intakes and urine outputs. Solute free water clearance was significantly higher in the HFI, though still negative. The average levels of none of the measured plasma concentrations showed significant differences.

Day-night comparisons

Day versus night averages were tested against each other in the two groups. Also day values from the two groups and night values from the two groups were tested with Student's t-test. All values are listed as mean±SE in table 7.3. The boxed values indicate significant results in the day vs night test. Bold face indicates significance in the NFI vs HFI tests, as for the 24-h levels.

	Normal Fluid Intake			High Fluid Intake		
	Day	Night	Overall	Day	Night	Overall
Uflow (ml/h/kg BW)	**1.49±0.1**	**0.75±0.059**	**1.2±0.065**	**2.91±0.092**	**1.34±0.14**	**2.32±0.056**
Uosm (mOsm/kg)	**548±24**	**674±37**	**595±24**	**332±19**	**458±30**	**379±20**
UNa (mmol/l)	**105±5.8**	**106±6.8**	**105±5.5**	**63.5±5.5**	**66.8±4.6**	**64.7±4.7**
UK (mmol/l)	**56±3.8**	44.4±5.3	**51±4**	**37±3.2**	34.7±4.2	**36.1±3.1**
Ucrea (mmol/l)	**10.5±0.61**	**16.5±1**	**12.8±0.56**	**5.12±0.37**	**9.71±1**	**6.86±0.54**
Eosm (mOsm/h/kg BW)	0.694±0.042	0.467±0.026	**0.602±0.03**	0.811±0.039	0.507±0.034	**0.696±0.031**
ENa (mmol/h/kg BW)	0.14±0.011	0.0774±0.0058	0.115±0.0079	0.165±0.013	0.0846±0.011	0.134±0.0096
EK (mmol/h/kg BW)	**0.0703±0.0058**	0.0301±0.0032	**0.0539±0.0039**	**0.0911±0.0073**	0.0385±0.0042	**0.0709±0.0057**
Ecrea (mmol/h/kg BW)	0.0121±0.00044	0.0109±0.00043	0.0115±0.00023	0.0122±0.00049	0.0102±0.00036	0.0115±0.00036
Cosm (ml/h/kg BW)	2.47±0.15	1.67±0.094	**2.15±0.11**	2.88±0.14	1.81±0.12	**2.48±0.11**
CNa (ml/h/kg BW)	1±0.082	0.556±0.041	0.825±0.057	1.19±0.093	0.611±0.08	0.966±0.069
CK (ml/h/kg BW)	**17.4±1.4**	7.59±0.77	**13.5±0.97**	**22.3±1.7**	9.46±0.98	**17.4±1.3**
Ccrea (ml/h/kg BW)	143±4.5	139±8.9	140±4	150±6.7	136±8	145±6.2
FEosm (%)	1.76±0.095	1.24±0.078	1.56±0.08	2.04±0.09	1.37±0.09	1.79±0.085
FENa (%)	0.705±0.056	0.413±0.034	0.592±0.042	0.809±0.048	0.459±0.054	0.675±0.044
FEK (%)	12.3±0.95	5.45±0.5	**9.61±0.7**	15.2±1.1	7.19±0.81	**12.1±0.82**
CH2O (ml/h/kg BW)	**-0.992±0.094**	**-0.908±0.071**	**-0.95±0.075**	**0.0283±0.12**	**-0.474±0.1**	**-0.158±0.097**
GFR (ml/h/kg BW)	143±4.5	139±8.9	140±4	150±6.7	136±8	145±6.2
Posm (mOsm/kg)	280±0.46	281±0.69	280±0.47	281±0.49	281±1	281±0.55
PNa (mmol/l)	139±0.34	139±0.28	139±0.27	139±0.36	139±0.37	139±0.33
PK (mmol/l)	4.05±0.022	3.92±0.044	4±0.026	4.1±0.036	3.96±0.045	4.05±0.037
PAVP (pg/ml)	1.16±0.099	1.13±0.12	1.16±0.099	1.17±0.057	1.11±0.086	1.19±0.049
PANP (pg/ml)	12.2±0.88	12.7±1.2	13.2±0.99	18.1±3.4	17±2.5	17.8±2.9
Pcrea (μmol/l)	84.7±2.8	81.7±2.4	82.9±2.5	81±1.7	80.4±2.5	79.7±1.6
Palbu (μmol/l)	584±35	559±34	567±34	628±9.2	601±7.6	613±8

Table 7.1: Overall, day and night averages in the two groups. Values are mean ± SE.

Urine flow and osmolarity both showed significant day-night variation, as well as differences in both mean day and night levels between the two groups. Neither urinary sodium nor potassium concentration showed day-night differences in neither of the two groups. The solute excretion variables for osmoles, sodium and potassium all varied in day-night levels. Day-day and night-night comparisons between the groups only gave a significant difference for the daytime level of potassium excretion, being higher in the HFI group. The picture repeats itself for the clearance variables and fractional excretion. Glomerular filtration rate showed no significant difference in any of the tests. Solute free water clearance had no day-night variation in the NFI group, but only the HFI group, where the daytime value even became positive. The only plasma variable that showed a day-night variation was potassium.

Time profiles and RMANOVA

The results of the repeated measures analysis of variance (RMANOVA) are plotted in Figure 7.4 for urine variables and in Figure 7.5 for plasma variables. The time profiles are shown as mean±SD, the mean represented as a solid line and standard deviations as error bars. The p-values corresponding to each variable are indicated above the figure. The p-values p_G, p_T, and p_{GT} refer to tests for differences between groups, time points, and group-time interaction, respectively.

Again, urine flow and concentrations (Uflow, UNa, Uosm, and UK) are significantly different between the two groups ($p_G < 0.05$). Urine flow and concentrations of osmoles and potassium show significant temporal variation ($p_T < 0.05$), whereas sodium concentration (UNa) does not.

When it comes to the excretion variables, there is only significant difference between the groups for EK. However, for Eosm $p_G = 0.057$, so it is very close to the significance level. All excretion variables have significant temporal variation.

The picture is qualitatively the same for clearance variables, where p_G for osmolar clearance Cosm is 0.061, so close to the significance level. Potassium clearance CK shows both time and group variation. Solute-free water clearance shows temporal and group variation, as well as difference in group-time interaction, meaning that the profiles are not parallel. And that is indeed the case, as can been seen in Figure 7.4. In the HFI group there is a large increase in C_{H_2O} from about 17.00 to 23. In this period C_{H_2O} even becomes positive, indicating the production of large amounts of very dilute urine.

Fractional excretion of osmoles, sodium, and potassium (FEosm, FENa,

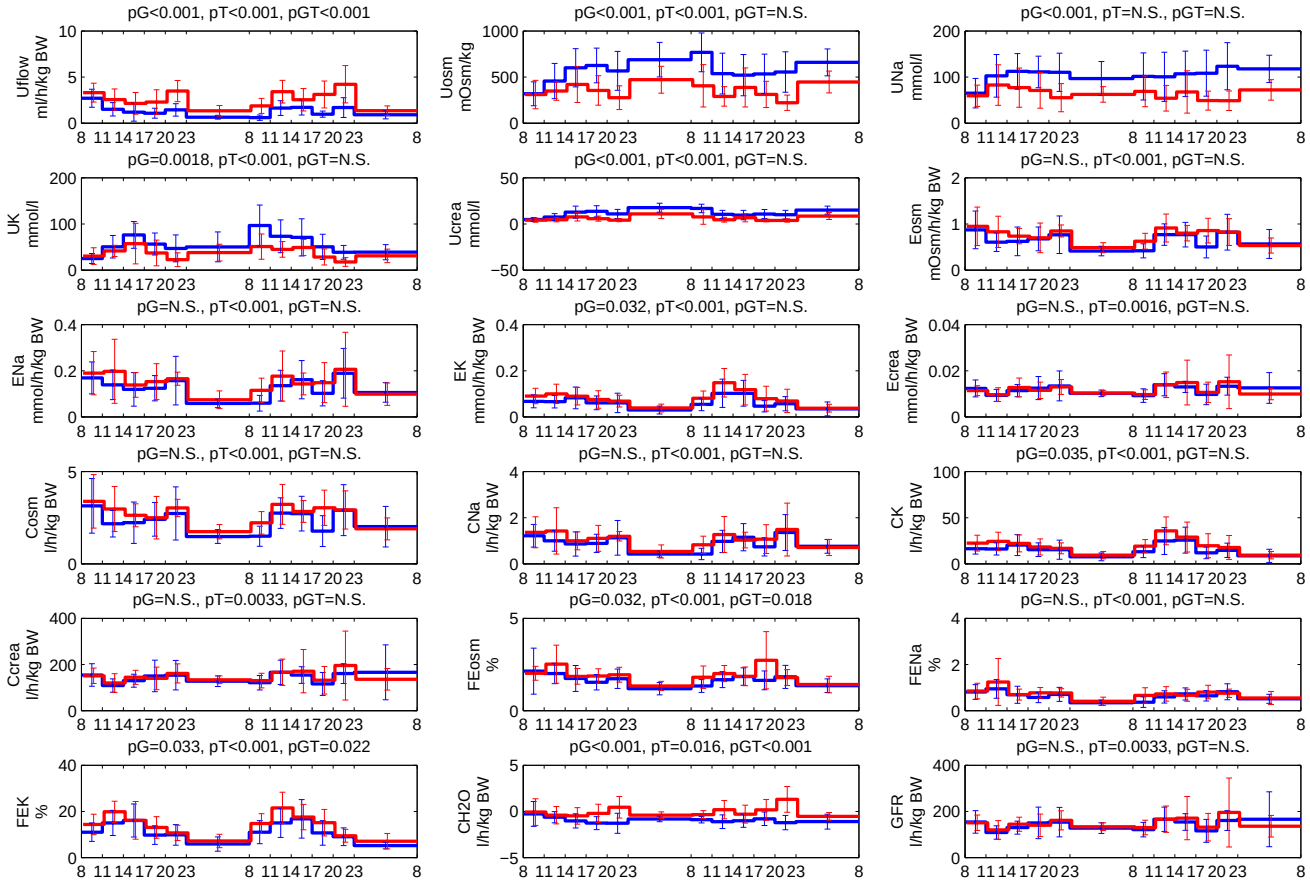


Figure 7.4: 48-h profiles of urine variables. RMANOVA p-values are indicated above each figure.

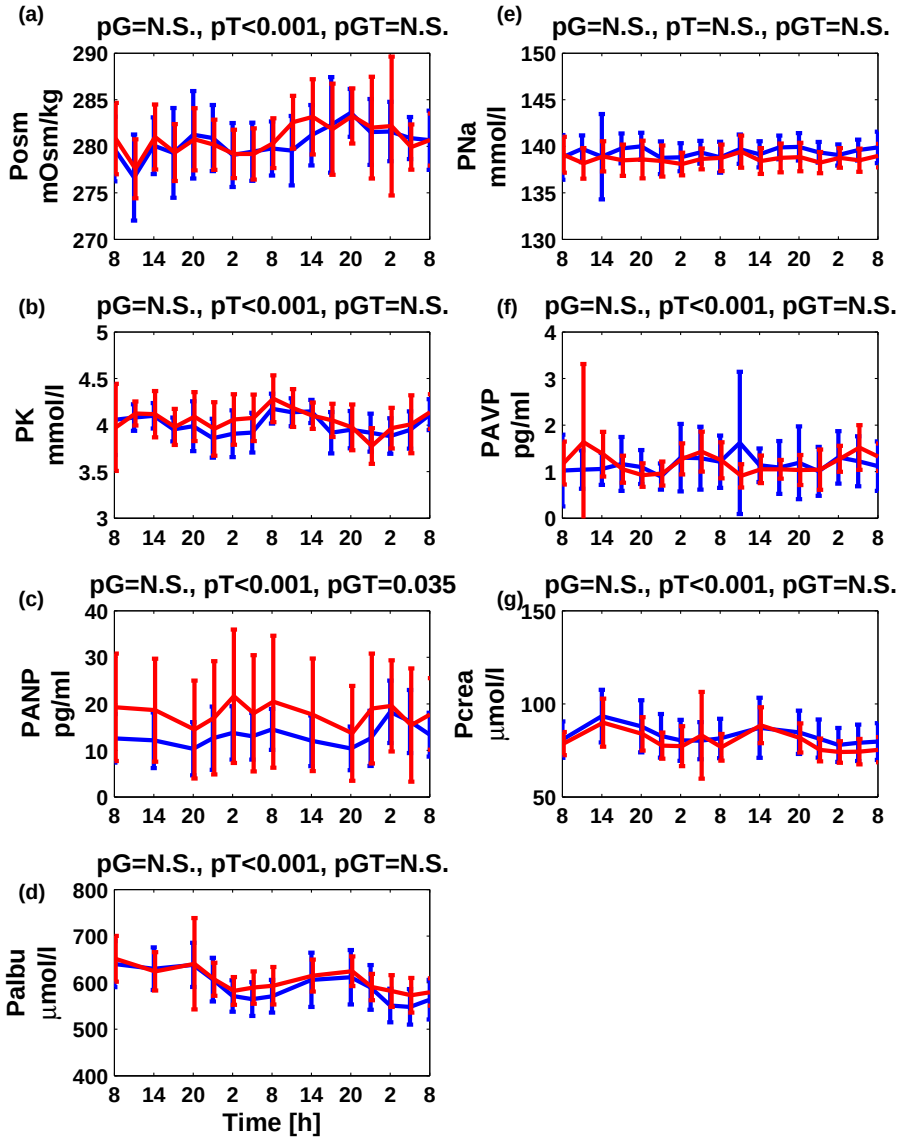


Figure 7.5: 48-h profiles of plasma variables. RMANOVA p-values are indicated above each figure.

FEK) all have significant group and time p-values. The case of FEosm and FEK group-time interaction is also significant. This seems to be due mainly to a larger fractional excretion during day-time.

All plasma variables have significant time p-values, except for PNa. There are no significant group p-values, but PANP has $p_{GT} = 0.035$. This may be due to the higher variance of PANP in the HFI group, as can be seen in Figure 7.5. The red error bars are clearly larger than the blue ones. PAVP shows no significant difference neither between the groups nor for group-time interaction.

Dispersion of hormone levels

The RMAOV did not reveal any differences between average levels of neither AVP nor ANP. By looking at the PAVP and PANP plots in Figure 7.5, one might suspect that the scale of variations or the dispersion of the two groups differs. In order to investigate this, Figure 7.6 shows box plots of PAVP (a) and PANP (b). Blue boxes are NFI, and red ones are HFI. The spread of the data in (a) seems similar, whereas in (b) the red distribution seems wider. The Ansari-Bradley test was used to determine whether the samples from the two groups stem from distributions with different dispersion. Since the Ansari-Bradley test is a non-parametric distribution, it requires no assumptions of normality [3]. The test result for PANP is significant with a p-value way below the 5 % significance level, at $p_{\mathrm{AB}} = 6 \times 10^{-6}$.

Fourier scores

All variables were analysed with the Fourier score technique - the simple one for plasma variables, and the integrated one for urine variables. For each variable 10000 background sequences were generated and evaluated to obtain a P-score. An overall view of significant Fourier scores is shown in Figure 7.9, but I will start with the example of plasma potassium levels. Figure 7.7 shows the time series of subjects with significant Fourier scores for plasma potassium (a) and insignificant scores (b). Again blue curves are normal fluid intake, and red curves are high fluid intake. It is quite clear from (a) that the selected series have a 24 h-oscillation with nadir in the late evening and a peak in the early morning. Also a larger proportion are from the normal fluid group than the high fluid group - more precisely 9 out of 17 (53 %) and 3 out 12 (25 %).

One way of checking the validity of the method is by plotting the background distribution generated from the background sequences. This is done

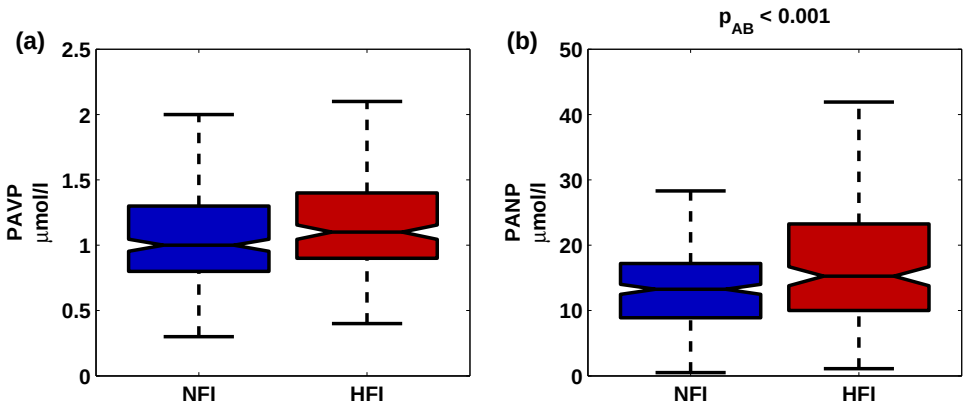


Figure 7.6: Boxplots of all AVP (a) and ANP (b) measurements in the two groups - blue is normal fluid intake, and red is high fluid intake. The distribution of AVP points are very similar and the dispersion is not different by the Ansari-Bradley test. On the other hand, the ANP points have different dispersion as revealed by the AB-test ($p_{\mathrm{AB}} < 0.001$). The interpretation of this is that in the high fluid intake group, there is a tendency to have much higher amounts of circulating ANP in response to the increased amount of fluid.

in Figure 7.8 for subject 9 with significant score (a), and subject 8 with insignificant score (b). The two distributions are very similar, and neither the Wilcoxon ranksum test for medians nor the Ansari-Bradley test for dispersion show significant difference between the two distribution. In (a) the Fourier score is 9.49 (green line), and it is well above the 5 % significance level. In (b) the Fourier score is 3.4 and below the significance level.

The fact that the two series are evaluated against very similar distributions, underlines the fact that the method of background sequence generation efficiently alters mainly temporal coordination, and it is not biased towards high scoring nor low scoring subjects.

The overall picture of Fourier scores of all individuals and all variables. Blue background shading is the normal fluid intake, and red is the high fluid intake group. Crosses denote significance at the 5 % level and circles denote 10 % significance. The first result is that more plasma variable series have significant Fourier score. The poor performance of the urine variables is quite striking. The over all distribution between the two groups (blue and red areas) seem not to differ significantly. As already mentioned above, there is a larger proportion of normal fluid intake subjects with circadian plasma potassium than there are high fluid intake subjects. Raising the

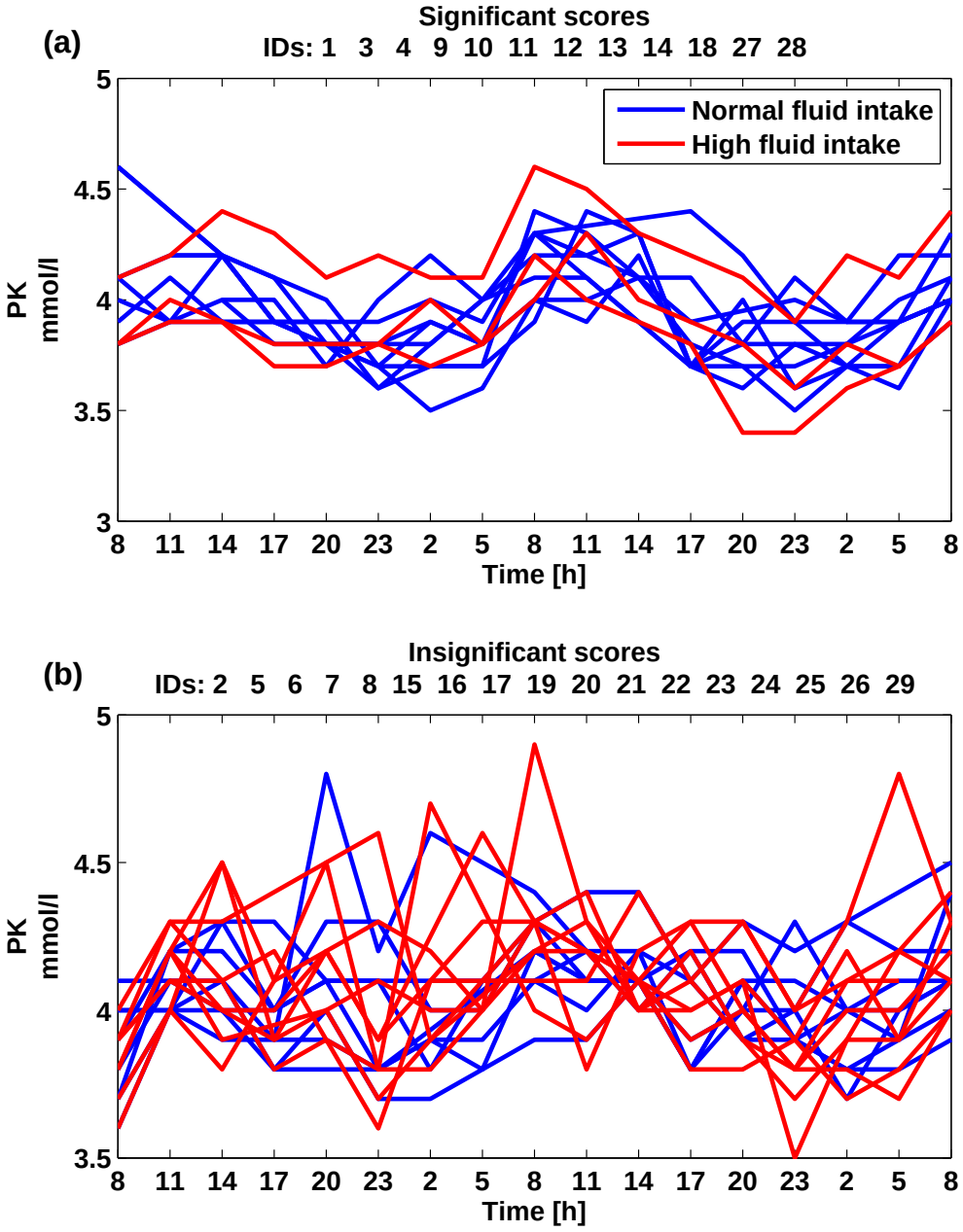


Figure 7.7: Time series of subjects with significant Fourier scores (a) and insignificant scores in (b) for plasma potassium levels. Normal fluid intake group are shown in blue and the high fluid intake group in red. The data in (a) clearly show a more rhythmic pattern.

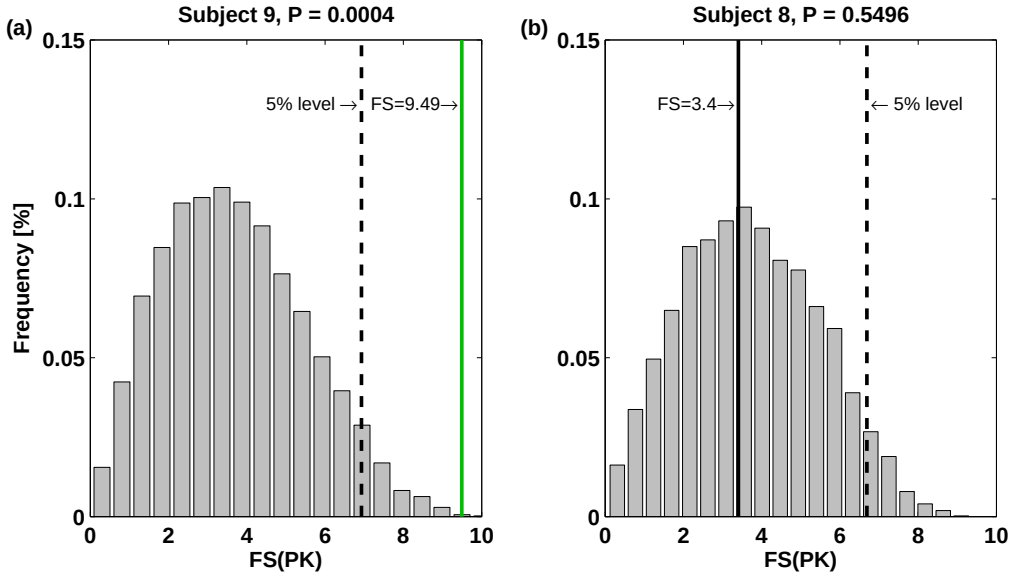


Figure 7.8: Two examples of determination of significance of Fourier score for plasma potassium. The histogram shows the empirical distribution of Fourier scores from the background sequences. The vertical dashed line shows the 5 % significance level. In (a) the data for subject 9 is shown, and the Fourier score is at a value of 9.49 (green value) well above the significance level and, hence, this score as considered significant. For subject 8 the Fourier score is below the significance level (black line in (b)).

significance threshold to 10 % does alter the picture very much.

7.4 Discussion

It is clear that the organism adapts the urine flow to the fluid intake. But the plasma data show that concentrations of sodium, potassium and osmolarity remain unchanged when fluid intake is increased. The question that one might ask, is how this adaptation occurs. The above data analysis adresses this question.

First of all, one might think that an increased fluid intake would lead to a lower plasma osmolarity and in turn an inhibition of vasopressin secretion, ultimately resulting in less water reabsorption in the distal part of the nephrons of the kidney. This is not the case, since plasma osmolarity is unchanged between the groups, and also plasma vasopressin is unchanged. This mechanism may, however, play a role on shorter time scales not de-

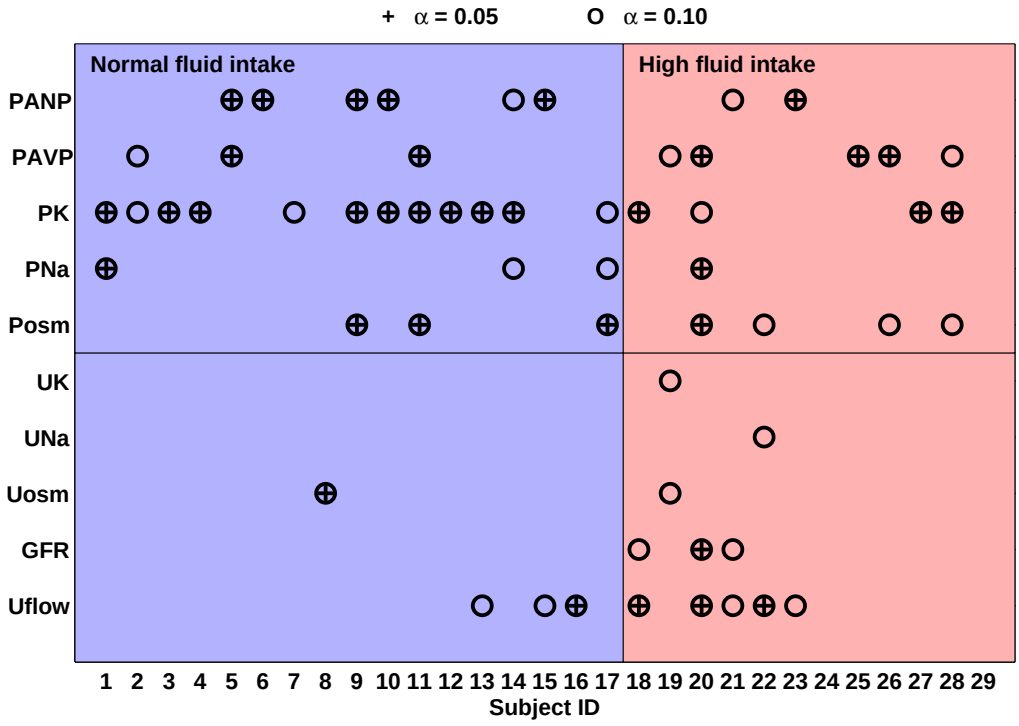


Figure 7.9: Significant Fourier score marked with + or ∘ for 5 % and 10 % significance, respectively. Subject IDs are on the horizontal axis, and a representative selection of the measured variables are on the vertical axis. The top half is plasma variables, and the bottom half is urine variables. It is clear that in spite of day-night averages showing diurnal differences in the urine variables, these are not detected by the Fourier score method. Relatively more plasma variables are identified as circadian, but no clear pattern can be identified. Plasma potassium (PK) is the variable with most significant cases.

tectable with the present sampling scheme. By this, I mean time scales shorter than the sampling interval of 3 h.

One might also speculate that glomerular filtration increases with fluid intake. This parameter is not different in the two groups, which is also coherent with the fact that glomerular filtration is very tightly regulated under various conditions [50].

If glomerular filtration is similar, then the regulation of urine output must be due to reabsorption processes in the tubular segments of the nephron. One obvious candidate that could be responsible for this regulation is vasopressin. But as has already been mentioned, there is no significant difference in the concentration of antidiuretic hormone between the two groups nor

in the scale of vasopressin concentrations. On the other hand, if one takes into account that vasopressin release is a function of plasma osmolarity, and that this is also not significantly different in the two groups, then one would expect similar vasopressin levels which is indeed the case.

Atrial natriuretic peptide could also play a role in the urine output regulation. The variance of ANP levels is much larger in the HFI group. One might interpret this as the system being more sensitive to volume expansion, which is one well-known stimulus of ANP release due to increased refilling of the atria.

What other mechanisms could be responsible for the adaption to higher fluid intake? Some kind of adaptation must occur.

- Relative alteration in the salt regulatory hormone levels like ANP and aldosterone. We have already seen that the spread of ANP levels is higher, thus the system is more responsive in the high fluid intake case. Plasma potassium circadian rhythm is also altered indicating a change in its major regulator aldosterone.

- A reduction of the osmolar gradient of the kidney medulla. This would result in less water reabsorption, and hence larger urine volume. A downregulation of urea transporters could be responsible for this.

- A downregulation of aquaporins in the collecting ducts. This would reduce water permeability of the principal collecting duct cells and thus decrease water reabsorption.

- Inhibition of vasopressin function by other substances. Substances, *e.g.* prostaglandin E2, are known to interfere with vasopressin. PGE2 inhibits production of intercellular cAMP, ultimately resulting in a decrease in aquaporin exocytosis.

- Pressure diuresis due to an increase in blood pressure in the HFI caused by larger blood volume. This has not been measured, so we cannot tell whether or not a pressure increase is present.

For far the most of the variables the absence or presence of circadian rhythms of is not affected by increased fluid intake. For solute-free water clearance the rhythm is inversed, i.e. in the HFI group it has its maximum during the day-time, whereas in the night-time in the NFI. The only plasma variable that showed a significant diurnal was potassium. This could be due to diurnal rhythm of aldosterone.

The Fourier score analysis revealed significant circadian rhythms, especially in plasma potassium for a large proportion of the subjects. By

applying an individual based analysis technique, it is revealed how large the intersubject variability is. This is of course always an issue in the analysis of clinical data. As was seen in the histograms Figure 7.8, the method precisely adresses the identification of temporal coordination in a signal.

The method did not prove effective for the analysis of urine variables. The accuracy of the method may be improved by shortening collection intervals or by extending the duration of the experiment to several days.

7.5 Summary

At this stage we can conclude that urine flow adapts to a higher fluid intake. Plasma osmolarity and solute levels are also tightly controlled under the same conditions. The most well-known antidiuretic substance, vasopressin, is *not* downregulated upon higher fluid intake, but its plasma values remain unchanged. Since glomerular filtration is also unchanged, this means that the magnitude of the renal response to vasopressin is diminished, while the analysis shows that the range of vasopressin levels remains unchanged. Plasma potassium was identified to be circadian in more than half of the normal fluid intake subjects and only a quarter of the high fluid intake subjects. This indicates a possible disruption of the circadian rhythm of potassium as a consequence of increased fluid intake. Most likely the adaptation to high fluid intake that takes place is in interaction between many factors, and the most promising candidates are alterations of ANP and aldosterone levels, and an adaptation of the osmotic environment in the kidney.

Chapter 8

Circadian gene regulation

This chapter is an investigation of transcriptional regulation of genes associated with water and salt balance. Two studies motivate the present work. In Chapter 6, I presented evidence for circadian rhythms in water and salt balance. At the same time circadian rhythm are known to be present at the level of transcription of genes. The relevance of circadian transcriptional control for different physiological systems and tissues were investigated in a study by Bozek et al. [23].

Combining the clinical and the genetic way of considering circadian rhythms is the purpose of the present study. A slight shift from water and salt balance to blood pressure regulation was necessary in order to select relevant genes in a convenient and reproducible way. The questions that we seek to answer are:

- Is there evidence for circadian, transcriptional regulation in blood pressure control?

- Are there specific transcription factor related to blood pressure regulating genes?

The computational framework employed in this study is called transcription factor binding site analysis. As the name states, it is a search for binding sites in the promoter regions of genes.

This chapter starts with a motivation for the study. Then the computational methods are presented, immediately followed by the results and their biological significance. A further investigation of the distribution of scores of transcription factors and promoter sequences follows before a short discussion.

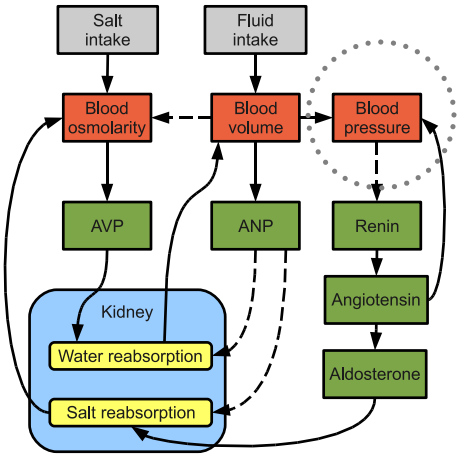


Figure 8.1: The causal loop diagram of water and salt balance presented in Figure 2.1. The genes considered in this chapter are associated with the regulation of blood pressure which is indicated with the dotted circle. The purpose of the study is to investigate circadian, transcriptional regulation and its role in this complex regulatory system.

8.1 Motivation

First of all, it is well known that circadian rhythms play are intimately involved in the regulation water and salt balance. This is also the case for the closely related regulation of blood pressure. In Figure 8.1, the dotted circle indicates blood pressure as the turning in the causal loop diagram in this study. And blood pressure is indeed circadian. One of the most common life style diseases of the Western world, hypertension or simply high blood pressure, is in many cases associated with a disruption of the circadian rhythm. In normotensive subjects blood pressure *dips* at night, and some hypertensive subjects are called *non-dippers* due to the lack of the nightly drop [126]. There is an extensive literature on the effects of hypertensive drugs such as β- and ACE-blockers on circadian rhythmicity - see for example the review by Lemmer [71].

In more general terms, the analysis addresses the problem of identifying transcriptional regulators that are common for genes related to a specific biological function. Some authors have addressed the inverse problem - do genes with common transcription factors share biological function [19, 20, 55]? The conclusion if these studies is that there is a higher degree of association to specific biological functions among genes with common transcriptional regulators. These findings have inspired other authors to find regulators common to genes that show circadian oscillations in their expression levels. One study of *clock controlled genes* by Yan et al. [129] estimated that about 50 % of all mammalian genes show circadian oscillations in at least one type of tissue.

Bozek et al. [23] used transcription factor binding site analysis of clock controlled genes to identify transcription factors common to this more or

TRANSFAC ID	Name
V$AHRARNT_02	AhR:Arnt
V$AHRHIF_Q6	AHRHIF
V$AP2_Q6	AP-2
V$AP2_Q6_01	AP-2
V$AP2_Q3	AP-2
V$CEBP_Q2_01	C_EBP
V$CEBP_Q3	C/EBP
V$CEBPA_01	C/EBPalpha
V$CEBPB_02	C/EBPbeta
V$MYCMAX_B	c-Myc:Max
V$CREBP1_01	CRE-BP1
V$E2F_Q3	E2F
V$E2F1_Q4	E2F-1
V$E4BP4_01	E4BP4
V$EGR_Q6	EGR
V$EVI1_04	Evi-1
V$HLF_01	HLF
V$HMGIY_Q3	HMGIY
V$HNF1_01	HNF-1
V$HNF1_C	HNF-1
V$IRF_Q6	IRF
V$IRF2_01	IRF-2
V$IRF7_01	IRF-7
V$KROX_Q6	KROX
V$MEF2_01	MEF-2
V$MEF2_04	MEF-2
V$MEF2_Q6_01	MEF-2
V$MEIS1AHOXA9_01	MEIS1A:HOXA9
V$NRF1_Q6	Nrf-1
V$OCT1_07	Oct-1
V$PAX4_04	Pax-4
V$SP1_Q6_01	Sp1
V$SP1_Q6	Sp1
V$SP1_Q4_01	Sp1
V$SP1_Q2_01	Sp1
V$SP1_01	Sp1
V$STAT5A_04	STAT5A
V$TATA_C	TATA
V$VBP_01	VBP
V$WT1_Q6	WT1
V$ZF5_01	ZF5

Table 8.1: Overrepresented transcription factors in a set of 2065 clock controlled genes as determined by Bozek et al. [23]. The factors are involved in numerous physiological processes, *e.g.* immune response, endocrine regulation and metabolism. TRANSFAC ID is quoted in the first column and the common name of the transcription factor.

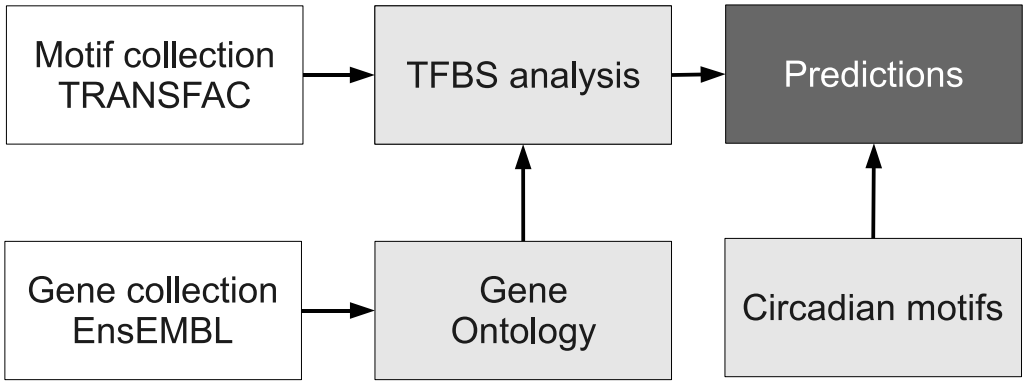


Figure 8.2: Flow diagram for the bioinformatic analysis.

vaguely defined group of genes. From a collection of six temporal microarray studies [36, 48, 78, 86, 112, 118] they compiled a list of 2065 clock controlled genes with clear temporal component. These genes were then subdivided into several sublists according to the tissue of the experimental sample and the circadian phase of the measurement. The results indicate that transcriptional regulators of clock controlled genes are involved in a many functional systems including metabolism, immune response and many parts of the endocrine system. Table 8.1 contains the factors that were overrepresented in the list 2065 clock controlled genes, and this will be the primary check list for circadian transcription factors.

8.2 Methodology

The methodological approach is divided into several parts. A flow diagram of the methodology is presented in Figure 8.2. Starting from the left, data acquisition is the first step. We need first a set of transcription factor motifs and a set of genes. I have used the databases *TRANSFAC* [127] and *EnsEMBL* [30] for this task. The data serves as input to a promoter analysis of selected genes and transcription factors. This is also called a transcription factor binding site analysis (TFBS). The promoter analysis relies on determination of statistically overrepresented binding sites, and the resulting factors will finally be discussed and compared with other studies on circadian regulation [23] (Table 8.1) and the literature. Below I present the different steps in more detail.

Data selection

The point of departure of the analysis is the selection of a set genes related to blood pressure regulation. This is not a straight forward task, and an objective method for selecting such a list is not readily available. In the spirit of reproductivity, I use the *Gene Ontology* [4, 73] for this purpose. The Gene Ontology is a dictionary with associations of genes and terms related to biological processes, molecular function, and cellular localisation of target proteins. The term *water and salt balance* does not exist in the Gene Ontology, and therefore the focus will be on genes related to *blood pressure regulation*.

The main reasons for using the Gene Ontology was to enable a reproducible approach, that could also be applied to other biological problems. The Gene Ontology contains annotations of properties of genes and gene products. It is organised as a directed acyclic graphic, where all relations of a given term are defined as parent relations. This creates a direction in the graph. E.g. *biological process* is a parent term of *biological regulation*, and vice versa for the child relation. The direction of the graph allows for straight forward determination of all parent terms of a given term. However, finding all child terms is more cumbersome. The Perl module `go-perl` provides functionality for parsing ontology type formats. Furthermore, it provides recursive search algorithm that allows for identification of all children or parent terms of a given term. Using the function `getRecursiveChildren()` with a parent term as input, one obtains all of its nested children. To further classify selected genes, the *molecular function* branch of the GO was used. For any given *biological process* in the GO this approach could be used.

Transcription factor frequency matrices were extracted from the database TRANSFAC. The raw list of matrices was filtered by species, and all matrices associated with the species *homo sapiens* were selected. The matrix representation will be dealt with in more detail below (see Figure 8.3).

Promoter analysis

Promoter analysis is based on pattern recognition in the promoter region of a gene. The promoter region is the region in the DNA sequence around the transcription start site. Basically a promoter analysis takes two inputs; (1) a promoter sequence and (2) a position frequency matrix. In the following, I will introduce these two terms.

Promoter sequences

A promoter of a gene is generally thought of as the region proximal to the transcription start site [1]. In the promoter region transcriptional regulatory proteins may bind to specific sites in the DNA string and stimulate or inhibit gene transcription. The extent of the promoter region is not a well defined quantity. Binding of regulatory proteins may occur right next to the transcription start or several thousand base pairs away. The most proximal region to the transcription has binding sites for the so-called transcription initiation complex at -10 bp and -35 bp, and regions at both sides of the transcription start site contain binding sites for other regulatory proteins [40, 123]. In the following we will investigate promoter regions 300 bp upstream and 200 bp downstream.

Promoter sequences were download from the *EnsEMBL* database through the Perl API (EnsEMBL 50). We chose to work with with a large promoter region (3000 bp upstream and 2000 bp downstream of the TSS) and a short one (300 bp upstream and 200 bp downstream of the TSS). The reason for choosing up- and downstream regions is that recent reports indicate that cis-regulatory elements are found in both up- and downstream regions respective to the TSS [40]. Similar choices of regions have been used by other authors [22, 23].

Mathematical representation of a binding site

A transcription factor binds to a DNA string at a transcription factor binding site which is just a substring of DNA with width W. The binding site can be sequenced experimentally. As is often the case in biological experiments their will be some extent of variability in the sequenced string. From one or several empirical profiles of the site, one can derive matrices expressing the frequencies of the four nucleotides at each position in the binding site. This is presented in Figure 8.3, which is taken fromWasserman and Sandelin [122]. The example in the figure is the MEF-2 transcription factor. Eight aligned sequences a the starting point (a). At each position in the binding site the nucleotides are counted and give rise to a position frequency matrix PFM (c), where one dimension is the four nucleotides, and the other dimension is the binding site positions. The position frequency matrix can also be presented by a position weight matrix PWM by taking the logarithm of the normalised frequency distribution of the PFM. This gives a measure of the information content. If only one nucleotide is present at a specific site in the experimental sequences, the information content is optimal. Sequence logos (f) are a convenient graphical representation of

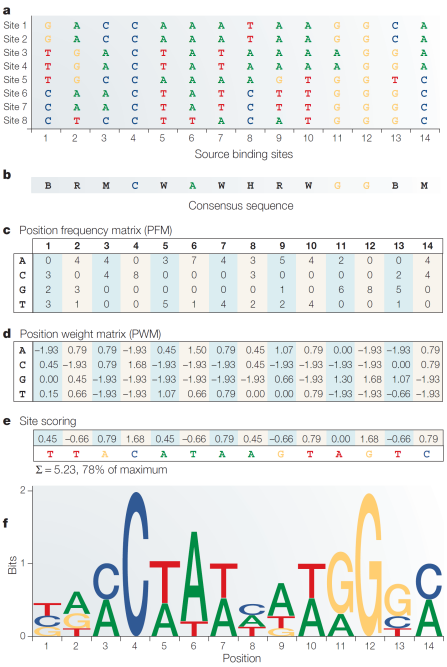


Figure 8.3: Mathematical representation of a transcription factor binding site. Aligned experimental sequences are shown in (a). Based on majority of counts of single nucleotides, one obtains consensus sequence (b). The position frequency matrix PFM (c) gives the frequency of nucleotides at each binding site position. The position weight matrix PWM (d) is a conversion of the PFM to log scale giving a measure of the information content of each nucleotide at each site. The site scoring (e) is simply the maximum weight at each site, and the sequence logo (f) graphically displays the information of PWM. From [122, Box 1].

the PWM. The PFM represents a probability of finding a nucleotide at a specific position in the binding site.

The search for binding sites

The search for binding sites using PFMs and promoter sequences is basically a problem of pattern recognition. The approach here is adopted from [41]. Let's consider a motif M of width W at location L in a promoter sequence S. Let $q(k, L_k)$ denote the frequency of nucleotide L_k in the PFM at position k. The likelihood ratio of the whole binding site is then

$$X_1(M, L) = \prod_{k=1}^{W} \frac{q(k, L_k)}{p(L_k)},$$

where $p(L_k)$ is a background probability of the nucleotide. The background probability may be estimated in various ways, but we will consider it to be the frequency in the sequence S. If there is a high similarity between PFM and sequence, the score will be high.

The single site score X_1 is haunted by a large number of false positive predictions, *i.e.* a high score at a site where no physical binding occurs *in vivo*. However, there is positive correlation between the level of scores in a sequence and the actual transcriptional regulation [123]. It is just too noisy

at the single site level. Hence we seek to get a score for a whole sequence. We calculate the average score X_2 of all sites L in the whole sequence S

$$X_2(M,S) = \frac{1}{N_S}\sum_{L\in S} X_1(L)\,. \tag{8.1}$$

This score is a measure of the association between a transcription factor and a promoter sequence. Different approaches to evaluating such a score for a transcription factor in a given sequence have been compared by Bodén and Bailey [20], and they found that the the average score X_2 of all sites in the whole sequence gave the best results. Still a single promoter may not be enough to obtain convincing results, and we expand the score to a set of promoters from different genes. Frith et al. [41] propose to make all possible collections of i promoters from the total of N promoters for $i = 1 \ldots N$. There are ${}^{N}C_i$ ways to do this, and the collection of all sets of i out N sequences is denoted A. Then for define the following score dependent on matrix M and i

$$X_3(M,i) = \frac{1}{{}^{N}C_i}\sum_{A}\prod_{S\in A} X_2(S)\,.$$

Finally, we sum over all $i = 1 \ldots N$ to obtain

$$X_4(M) = \frac{1}{N}\sum_{i=1}^{N} X_3(i)\,. \tag{8.2}$$

The logarithm of X_4 is the final score evaluated by *Clover* introduced in the paper by Frith et al. [41]. The evaluation of overrepresentation is done by supplying a set of background sequences. From this set, N sequences of similar length to the original set are drawn at random, and X_4 is calculated for this random set. This procedure is repeated typically 1000 times, and a P-value at 5 % is determined as the 95^{th} percentile of the resulting distribution of scores.

The approach explained above is implemented in the program *Clover* [41]. The software needs two inputs: 1) a list of promoters sequences and 2) a list of positional weight matrices (PWM) for transcription factor binding sites. Additionally one can provide a set of background sequences. The PWMs are used to scan through the promoter sequences shifting one nucleotide. The output is a score and a P-value for each matrix, the P-value representing the proportion of samples with higher score than our selected list. The background sequences were the corresponding promoters of all 21528 protein-coding human genes from EnsEMBL.

8.3 Results

The results obtained from the Gene Ontology and transcription factor binding site analysis with *Clover* are presented in the following. The biological significance of the findings are also discussed.

Gene selection

The first step is finding the genes of interest. By using the Gene Ontology, the gene selection is reproducible, and as annotations are added to the Gene Ontology in the future, further genes can easily be included in the study. Searching the Gene Ontology O for genes annotated with the term for blood pressure regulation, *GO:0008217*, results in 63 human genes. These genes can be subgrouped according to their molecular functions: receptor activity (29 genes), transporter activity (8), catalytic activity (25), hormone activity (9), peptidase activity (7), and enzyme regulation activity (4). These categories classify 59 of the 63 genes. Some genes also appear with more than one molecular function. The diagram in Figure 8.4 lists the genes by their EnsEMBL short names and the partition by molecular function. The list of genes is shown in Table A.1 in Appendix A.

The Gene Ontology-based gene selection performs quite well, and the genes selected represent well-known physiological systems. This would of course also be suspected, since the entries in the Gene Ontology are based on published, experimental evidence from the literature. In the endocrine part A, B, and C-type natriuretic peptides (NPPA, NPPB, NPPC) were selected along with their receptors (NPR1, NPR2, NPR3). Components of the renin-angiotensin system were included; renin (REN - peptidase activity), angiotensinogen (AGT - hormone/enzyme regulation activity), and aldosterone synthase (CYP11B2 - catalytic activity). The cortisol pacifier HSD11B2 that oxidises cortisol to inactive cortisone to give aldosterone competitive advantage in binding to its receptor is found in the catalytic activity box [44]. Angiotensin receptors (AGTR1, AGTR2 - receptor activity) are included. The renin receptor (ATP6AP2 - receptor activity) has also found its way into the blood pressure regulation section.

In the transporter activity box, we stumble upon, for example, the sodium-potassium transporting ATPase subunit alpha-1 ATP1A1. This molecule is essential to the regulation of intra- and extracellular solutes. It is known to be upregulated during, for example, aldosterone stimulation [42]. Also the potassium-chloride transporter SLC12A6 is found here.

In the hormone section, apart from those already mentioned, we also find genes such as endothelin (END) and urotensin II (UTS2). Both are

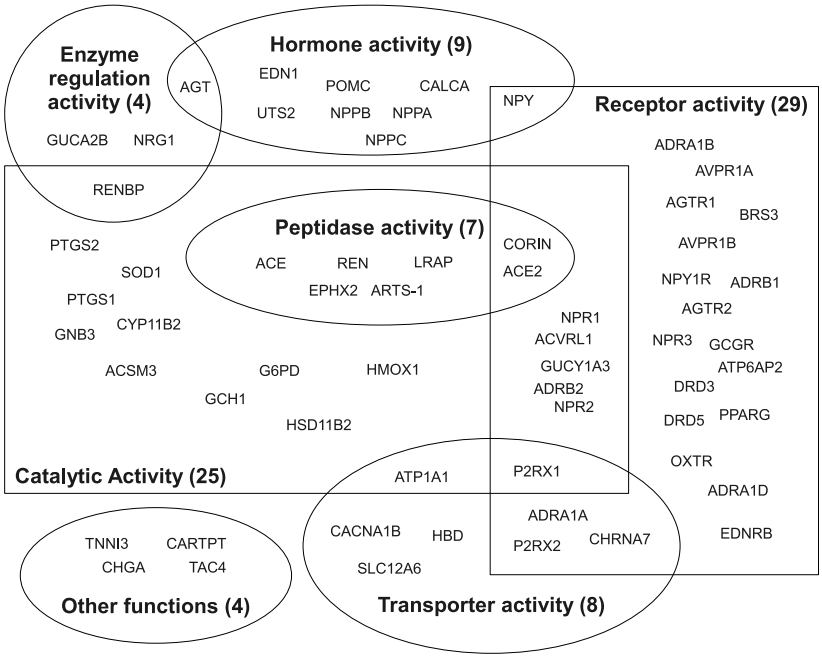


Figure 8.4: Venn diagram of 63 blood pressure regulating genes selected from the Gene Ontology. The genes have been grouped by their molecular function. Some genes have several functions.

known to exert effects on reabsorptive processes in the kidney and the vasculature [76, 106]. The calcium regulatory hormone calcitonin (CALCA) is also associated with blood pressure regulation.

Still the hormone vasopressin has not yet been annotated as a blood pressure regulating gene, and it is therefore not found in the list. However, both of its vascular receptors AVPR1A and AVPR1B are included in receptor activity. Vasopressin is closely related to the hormone oxytocin, and again we find the receptor (OXTR) but not the hormone itself. Nuclear receptor PPAR-γ (PPARG) is present. Nuclear receptors have been shown to exhibit circadian oscillations in many tissues [131]. We also find three types of adrenergic receptors (ADRA1A, ADRA1B, ADRB2) as evidence for adrenergic control of blood pressure. Prostaglandins also come into the picture by the presence in the catalytic section of two prostaglandin synthases PTGS1 and PTGS2.

Overrepresented factors

Clover was run on the list of blood pressure regulating genes with the 440 PWMs. A scatter plot of scores versus P-values is shown in Figure 8.5. The red dots indicate the points below the significance level. This procedure resulted in a total of 17 matrices with P-values of less than 0.05, which are shown in Table A.2 in appendix A.

However, one has to take into account that PWMs may, to some extent, be redundant. By visual inspection of similarities in the matrices and sequence logos, we narrowed the field down to 11 factors, namely TATA, CREB, WT1, GATA1, ZNF219, MAZ, CRX, AP2, IPF1, NRSF/E, and HAND1:E47. For example, for V$NRSE_B and V$NRSF_01 the positional weight matrices are very similar. They are shown as sequence logos in Figure 8.6. Both are 21 nucleotides long and the information content is basically indistinguishable. The matrix for V$NRSF_Q4 is also grouped with NRSF/E. The other factors with similar matrices are the four TATA-box related VTBP_Q6, VTATA_C, V$TATA_01, and V$TFIIA_Q6, and the the GATA ones V$GATA1_04 and V$GATA_Q6.

The final 11 factors are listed in Table 8.2 along with their scores and P-values. In the following, I will discuss the putative roles of these factors.

The TATA related are most often associated with transcription initiation and the core transcriptional machinery [123]. TATA binding protein (TBP) was also found in the circadian transcription factor (see Table 8.1). Since TATA boxes are common to many promoters, the overrepresentation may indicate a role in circadian regulation.

CREB (cAMP responsive element binding protein) plays a role in transcriptional regulation of many of the genes in water and salt balance as discussed in Chapter 6. In the circadian literature cAMP and CREB often pops up, and some authors claim it is a core signalling component of the circadian clock [85, 96]. As an intracellular second messenger, cAMP is involved in many different processes, and it seems natural that it would be the case for blood pressure regulation also.

Wilms' tumor protein, WT1, is one of the high scoring factors. This factor is particularly interesting, because it is known to be involved in a number of kidney specific processes - underlining the importance of the kidneys in blood pressure regulation. Its name stems from its relation to Wilms' tumor which is a pediatric, kidney specific tumor [28, 130]. The protein is known for its involvement in urogenital development [35] in a number of mouse models. It is also known to be involved in transcriptional regulation of renin [107] though at sites very distant (8-11kbp) from the transcription start site. Here we only consider 300bp upstream. The fact

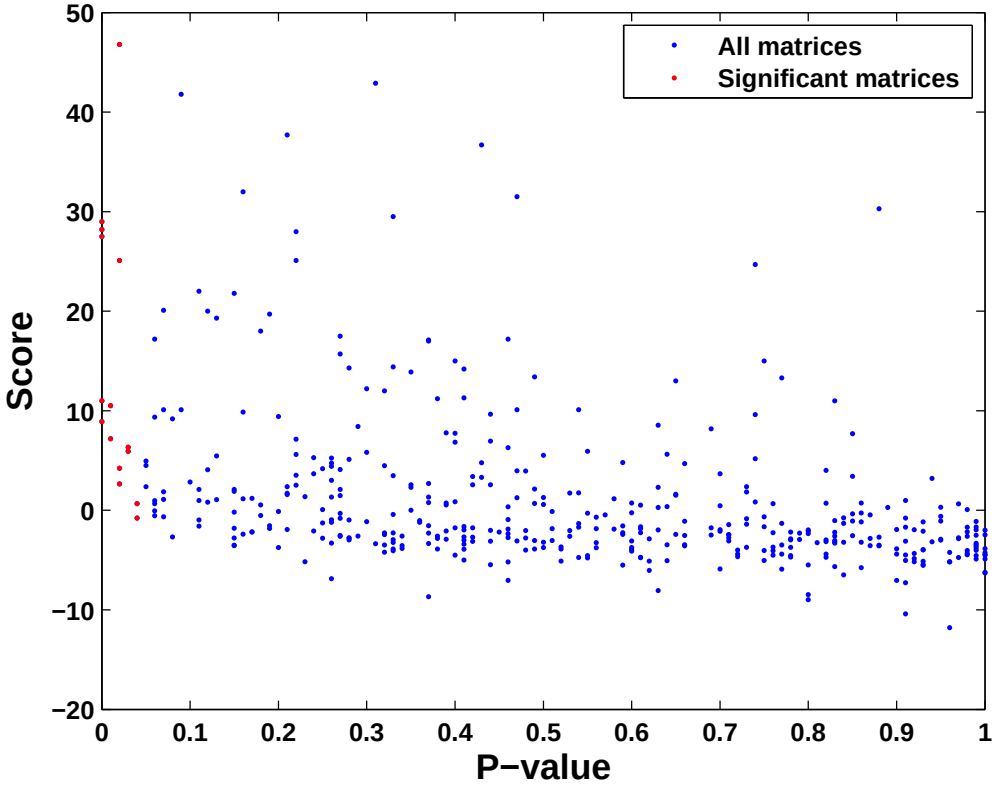


Figure 8.5: Scores vs P-value for each of the 440 matrices. Each dot corresponds to a matrix. The red dots indicate matrices with scores below the significance level of $p = 0.05$.

that it shows up in this list of many blood pressure regulating genes may indicate that it plays a more versatile role in this regulatory system with many functions yet to be discovered.

Neuron restrictive silencing factor and element (NRSF/NRSE) negatively regulate many neuronal genes in stem and progenitor cells [57]. Malfunctioning gene expression in Huntington's disease may caused by overexpression of NRSF, and it is therefore a likely pharmaceutical target [90]. The NRSF/NRSE system has not previously been associated with blood pressure regulation.

The transcription insulin promoter factor 1 (IPF1) is of course a regulator of the insulin gene [84]. It is not previously known to regulate the blood pressure regulating genes, and it is therefore interesting to find it here. It may be a link between two very complex regulatory systems.

The cone-rod homeobox protein CRX is also in our list. Differentiation

Factor	Raw score	P-value
TBP	29	< 0.01
TATA	28.2	< 0.01
NRSF/NRSE	8.9	< 0.01
ZNF219	10.5	0.01
MAZ	46.8	0.02
CREB	2.65	0.02
WT1	25.1	0.02
CRX	6.33	0.03
HAND1:E47	5.92	0.03
GATA	-0.77	0.04
AP2	4.94	0.05
IPF1	2.37	0.05

Table 8.2: Overrepresented factors in 63 blood pressure regulating genes.

of photoreceptor is one known function, and no previous record of relation to blood pressure regulation have been found.

ZNF219 is a so-called zinc finger protein, and the targets of this factor are not well known. It has recently been associated with WT1 [61] as another of our factors MAZ. MAZ is known to activate transcription of PNMT [128], an enzyme necessary for the production of adrenalin. The relation to blood pressure regulation is not obvious. AP2 stands for activator protein 2, and its targets are many and diverse. GATA also has many targets, one of which is endothelial development [120]. That may possibly be a connection to blood pressure regulating genes.

HAND1:E47 is a protein complex of two transcription factors. HAND1 alone and also the complex interestingly bind strongly to E-boxes [64], which is a very common motif among circadian transcription factors [23].

As previously mentioned, promoter analysis of genes with circadian variation in the mRNA levels have been reported in the literature [23, 129]. 5 of the 11 factors have also been identified as overrepresented in the study by Bozek et al. [23]. Not all of them are in the list in Table 8.1, but can be found in other tables from the same reference. These are CREB, AP2, IPF1 , WT1 and TBP. All of these factors may be play a role in circadian regulation of blood pressure regulating genes.

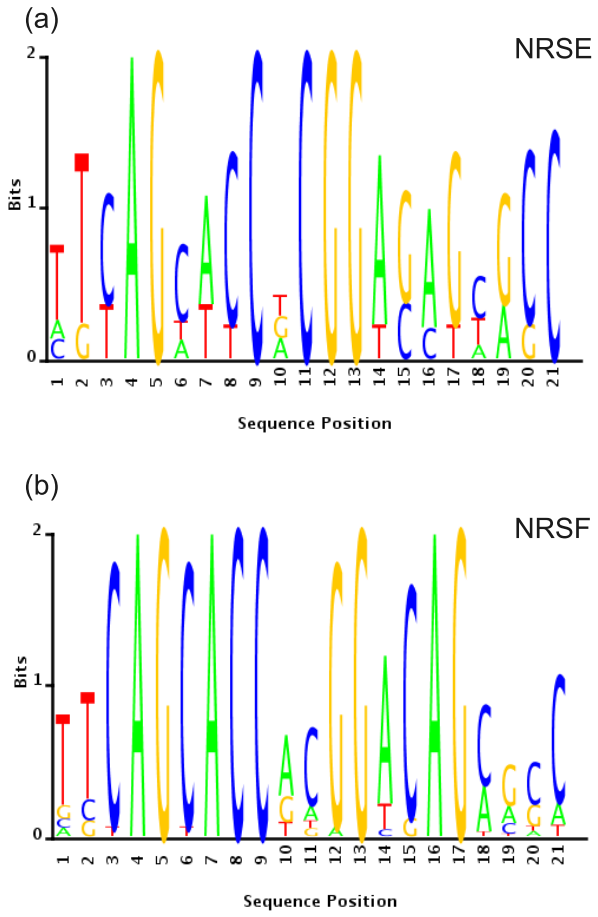


Figure 8.6: Sequence logos for the positional weight matrices NRSE (a) and NRSE (b). By visual inspection they are very similar.

8.4 Distribution of scores

The promoter analysis proved to be a useful tool for evaluating transcription factors associated with blood pressure regulation. The results generated by the analysis tool *Clover* are based on the score X_4 in Equation (8.2). It is not strictly necessary to calculate X_4 to obtain results for the overrepresentation of a given transcription factor in a given set of genes. The score X_2 in Equation (8.1) gives the association of a transcription factor to a single promoter. If we take all F factors available and all P promoters of the genome, we can make a map of associations of dimension $F \times P$. Such a map is shown in Figure 8.7 for all 440 human matrices from TRANSFAC on the

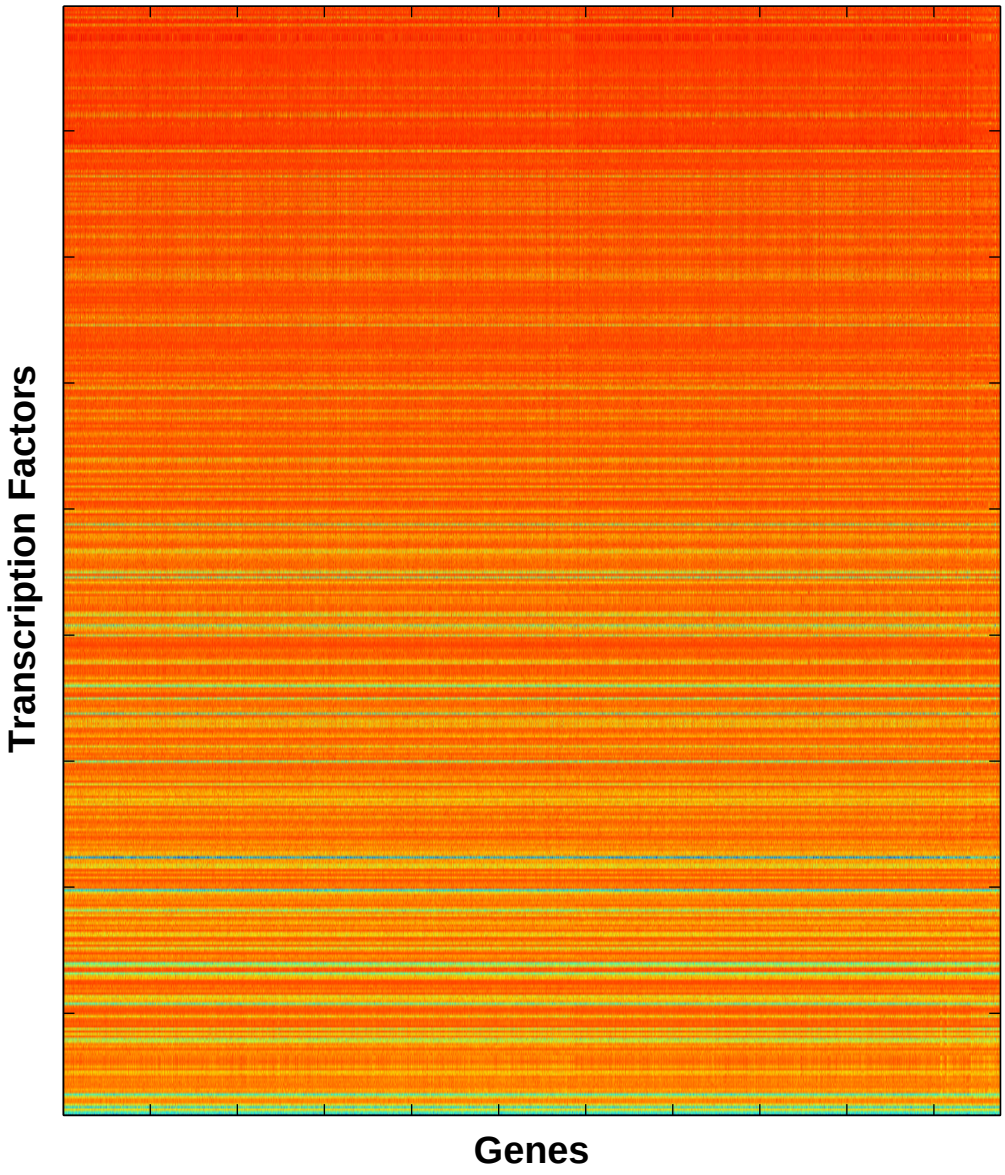


Figure 8.7: Heat map of transcription factor/gene associations. Each dot corresponds to the score of a given transcription factor positional weight matrix in a given gene promoter. Red correspond to a high score, and blue to a low scores. The vertical axis has 440 points for the transcription factors, and the horizontal axis has 21528 points for all protein-coding genes in the human genome. By selecting a subset of genes, one can statistically evaluate overrepresentation of a specific transcription factor in the subset against the rest of the genes.

vertical axis, and all 21528 protein-coding human genes from EnsEMBL. Each point in the map corresponds to a given factor and a given gene. I call it a heat map, since the hot red colour indicates a high score, and the cold blue colour indicates a low score. One may immediately notice the horizontal stripes in the map. This indicates that in general, across the whole genome, a specific factor has a low score. So does this mean that the single factor is less useful than others? No, it does not. Factors have different binding sites and these are of differing lengths, and they also differ in how well they have been determined experimentally. So one cannot really compare across the vertical axis of the figure, since the matrix representation of the transcription factor binding sites is not very homogeneous.

On the other hand, the promoters sequences have a more comparable representation, and this we have already exploited to get the results from *Clover*. By selecting a subset of genes, as for example the 63 blood pressure regulating genes, we can evaluate the distribution of scores in the subset against the distribution of scores in the rest of the protein-coding genes. A suitable test for this is the Wilcoxon ranksum test, which is a test for the null hypothesis that the medians of the two distribution are equal. This procedure is illustrated in Figure 8.8. Subfigure (a) and (b) show the collection of distributions of scores for each transcription factor (TF). The z-axis is the empirical density. In (a) it is the distributions from scores from all genes (except the selected genes). The distributions of scores from the selected genes are shown in (b). If we take out one distribution from (a) and one from (b) as for example the one indicated with red, we can compare the two and statistically test if they are different. Such an example is shown in Figure 8.8(c) for the factor Wilm's tumor protein WT1. This factor was discussed in Section 8.3. The blue distribution is the background distribution, and the green one is the distribution in the selected genes. The median of the scores in the green distribution is statistically different from the blue one by the Wilcoxon ranksum test with a p-value of 0.048 which tells us that WT1 is overrepresented in the selected genes. This is repeated for all 440 factors, and this approach leads to 10 overrepresented factors, which are indicated in the Figures 8.8(a) and (b) by black curves. These factors are shown in Table 8.3.

The results of the analysis by using the Wilcoxon ranksum test differ from the results already obtained with Clover. Among the ten factors are, however, some of the same regulators which we may consider as more robust predictions. These factors are WT1, TBP, TATA, MAZ, and CRX. Among the new ones is COUP, chicken ovalbumin upstream protein, which in complex with TFII is regulated by high glucose levels in endothelial cells [25] - a possible link to diabetes mellitus. COUP:TFII is also involved in regulat-

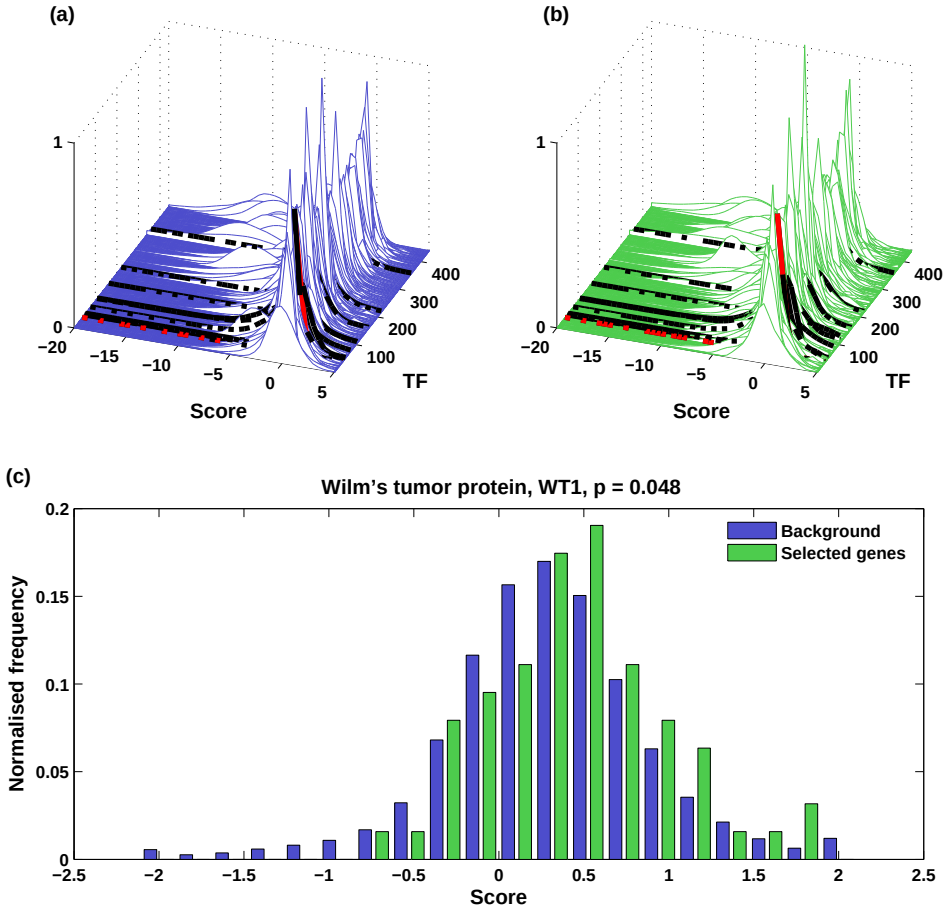


Figure 8.8: Distributions of scores for each transcription factor (TF) presented as estimated probability densities. In (a) the distributions are taken from the whole genome, and in (b) it is taken from the selected set of 63 blood pressure regulating genes. The black lines indicate overrepresented factor in the selected genes. One of these is Wilm's tumor protein WT1, and it is represented in by a red curve in (a) and (b), and the corresponding distributions in blue and green are compared in (c). With a p-value of 0.048 we can conclude by Wilcoxon ranksum test that the two distribution are different in median, and that WT1 is overrepresented in the set of promoters of blood pressure regulating genes.

TRANSFAC ID	Name	p-value
V$WT1_Q6	WT1	0.048
V$TBP_Q6	TBP	0.001
V$MZF1_02	MZF1	0.035
V$MAZ_Q6	MAZ	0.020
V$TATA_C	TATA	0.007
V$TATA_01	TATA	0.002
V$TBP_01	TBP	0.012
V$CRX_Q4	CRX	0.041
V$COUP_01	COUP	0.023
V$MEIS1AHOXA9_01	MEIS1A:HOXA9	0.013

Table 8.3: Overrepresented factors found by Wilcoxon ranksum test of score distributions in whole genome and 63 blood pressure regulating genes. For illustration see Figure 8.8.

ing the specificity of arterial and venous endothelial cells during development [113] which is directly important for the regulation of blood pressure. MEIS1A:HOXA9 can be found among the circadian factors in Table 8.1, and MZF1 is present in the list for heart specific circadian transcription factors in the supplementary material of the paper by Bozek et al. [23].

8.5 Discussion

The overall goal of this study was rather ambitious - building a bridge from genetics to physiology across the troubled waters of circadian rhythms in water and salt balance. I cannot say that the construction work is completed, but still we have a preliminary assessment of whether or not it is possible to one day complete the project and gain an in-depth understanding of these interconnected, intriguing regulatory mechanisms.

The methods applied in this section are based on statistical overrepresentation in a set of selected genes. Statistical conclusions are not factual, but are subject to inherent uncertainty. The reason for the use of statistical measures is the high degree of inaccuracy in determination of binding sites. In fact, a linear DNA sequence does not contain a lot of information. Compared to a data series of a continuous variable with the same sample rate, the DNA string only has 4 possible states at each data point as opposed to a continuous spectrum. In this study, I have relied on other sources of information to give additional confirmation of the findings.

About half of the factors identified were also identified in the study by Bozek et al. [23]. This is very interesting, since it supports the hypothesis that the circadian rhythms in blood pressure regulation are found throughout the biological levels of organisation from genes to organism. This is of course not at strict proof (as if there were any in biology), and the specific mechanisms and the complex genetic and signalling networks are not identified by the applied technique.

The physiological relevance of the analysis is underlined in particular by the interesting finding of Wilm's tumor protein. It is a regulator that is found to be involved in an increasing number of processes related to the kidney and blood pressure regulating substances such as renin. Two factors, IPF1 and COUP-TFII, indicate a possible connection to diabetes mellitus and regulation of glucose and insulin.

So what can be done to enhance the accuracy of promoter analysis techniques? This is a question of increasing the information available to distinguish between false positives and true positives. One crucial piece of information is the wrapping of DNA - the chromatin structure. This part of the chromosomes has proven to be highly modifiable, and it may even store information in addition to the "pure" genetic information of the DNA string. The chromatin structure is made up of histone molecules, and these proteins may be modified. When DNA is transcribed the chromatin structure has to open up to allow access to the DNA string for transcription factors, and the modifications are thought to be able to make it harder to open up the DNA and hence certain genes can be silenced by preventing access. Furthermore, the nucleotides of the DNA may be modified by addition of a methyl group - so-called DNA methylation. The two types of changes are called epigenetic modifications, and they can be inherited. Integrating epigenetic information would greatly reduce the number of false positive predictions by only considering accessible regions of the DNA.

8.6 Summary

In this study the transcriptional regulation of blood pressure was investigated with focus on circadian rhythms. 11 transcription factors were identified as overrepresented in a set of 63 genes associated with blood pressure regulation. Of these 11 factors, 5 were also found in a study of circadian, transcriptional regulation by Bozek et al. [23]. The factor Wilm's Tumor Protein 1 (WT1) was among the overrepresented factor, and this underlines its importance in blood pressure regulation, where it is already known to be involved in renin transcription and in kidney related diseases. Furthermore,

a more general methodology of evaluating transcription factor-gene association was introduced, and the results were, to some extent, overlapping with the first analysis. The analysis supports the hypothesis of circadian, transcriptional regulation in blood pressure regulating genes.

Chapter 9

Discussion

As the title of this thesis states, the common thread of the preceding chapters is *Biosimulation of water and salt balance*. The approaches have been very different, and one might even claim that I have used a sort of scattergun technique across computational disciplines from ordinary differential equations to bioinformatics. This is not untrue, and it is also very fruitful.

Models

The models of Chapters 3-5 are all mechanism-based models. In principle, the models equations were defined from known causal relations, and all of the model parameters were determined from data sources different from the simulated experiment in Figure 5.2. This approach is time-consuming, but it is also gives a lot more information than a data-driven technique. Let's assume that we had started by fitting the data in Figure 5.2 to some model involving a time scale of onset of urine production. We would be able only to determine the time scale of onset of urine production. With the mechanism-based approach, we have based the model on several intermediate steps from ingestion of water to urine production onset. We can, in principle, evaluate the influence of every mechanism included in the model and overcome statistical standard issues of observability and identifiability.

In Chapter 5, I successfully managed to integrate models from several levels of biological organisation. Starting with a cellular model of aquaporins to a kidney model, and finally the systemic model. The challenge for such an approach is not a question of what to include in the model, but rather what to exclude from the model. I am not the first researcher that attempts to construct a model of water and salt balance and related regulatory systems. The most famous example in the physiology literature

is the model of blood pressure regulation by Guyton et al. [49]. The main objective of that model was of course to model blood pressure, a variable that I chose not to model. The Guyton model is extremely complex, and in my opinion it is too complex to analyse in detail and understand. Already with the simple model, relative to the Guyton model, of Chapter 5 the task of analysing the influence and importance of all parameters is huge. The choice of simplicity versus complexity is a major issue in any type of modelling work.

The blood pressure-oriented approach used by of Guyton has been questioned by several authors [16, 17, 18, 88, 99]. Guyton proposes blood pressure as a main regulator of both volume and sodium excretion, so-called pressure diuresis and pressure natriuresis. However, under normal conditions there is no clear evidence that changes in sodium intake affects blood pressure. So in terms of regulatory feedback blood pressure would be excluded as the mediator of the urine flow response. Basically, it is a question of which physiological variables are controlled and which stimulate the regulatory feedback systems. Some variables are very tightly regulated such as plasma sodium and osmolarity (see Section 7.3). The basic concept of negative feedback regulation relies on perturbing a controlled variable that through a series of reactions eliminates its own perturbation. But if the variable is so tightly regulated, one might ask how the feedback mechanism is activated. One answer is that we do not know the variable observed by the physiological system. Another answer of more hypothetical origin, is that there is a sensing of changes in intake that precedes the changes in osmolarity, for example. Such a mechanism could be a humoral signal from the gut whenever a certain solute is present in large quantities. Such a gastro-intestinal signalling system has been suggested for the case of potassium homeostasis by Youn and McDonough [133]. The precise regulation of water and salt balance under conditions of highly variable intake makes this type of regulatory system very plausible.

Time scales

Understanding time scales is essential to improve the understanding of water and salt balance. In Chapter 5, we saw that the time scale of aquaporin regulation, the time scale of water absorption, and that of vasopressin degradation were sufficient to describe the response to fluid up to 3-4 h. But from the current work we cannot immediately say anything about response to salt intake. Are the "salt dynamics" at the same time scale? It is generally accepted that regulation by ADH is quite fast, and that salt reabsorption

processes are subject to a slower regulation [16]. This we did not see explicitly in the kidney model, but the model results still indicate that this is very probable. Recall that the urine volume flow dependence on water permeability was very strong, where as the solute flow response was sensitive to both solute and water permeability. Now, if the water permeability is a fast process that quickly adjust the water flow, then there is time for the solute permeability to be adjusted to the appropriate level.

Breaching the gap from the relatively fast response to fluid intake and all the way to circadian rhythms is challenging. Especially the transition from diuretic to antidiuretic states is interesting in this respect. In Chapter 7, we saw that the level of ADH was unchanged in spite of an elevated fluid intake. This also makes sense from the point of view of plasma osmolarity, since it also was not different between the two groups. So how does ADH regulate differently in the two groups? Now, recall Figure 4.7(a). The urine flow at different interstitial osmolarities reveals that the relation between flow and ADH dependent permeability can be shifted by changing the interstitial osmolarity. This is very likely the case, and this is a shift that must occur within at most a couple of days (judging from the data in Chapter 7), and very likely shorter. A tour of kidney and vasopressin literature leads one to a collection of papers by Atherton and coworkers [5, 6, 7, 8, 9]. They dissected rat kidneys at different time points after and during vasopressin stimulation. During vasopressin stimulation they found an interesting slow change at the order of 4-5 h, where the renal interstitial gradient was slowly increased. This build up of solute in the interstitium is, in my opinion crucial for the maintenance of an antidiuretic state, and it cannot be explained by the current distal kidney model of Chapter 4. To capture this type of dynamics, it is necessary to include a time dependent interstitium in the model.

Data

The experimental achievements within molecular and cellular biology that form the basis for bioinformatic data generation are impressive. But that does not mean that the data is straight-forward to analyse or comprehend. I would like to quote Professor Hanspeter Herzel from Humboldt University. He expressed how ungrateful an endeavour in bioinformatics can be at times, and he concluded “Bioinformatics is a search for very weak signals in huge amounts of noisy information”. When we later discussed inconsistency in cell biological research he added “Well, cell biology is just a collection of case studies”, referring to the low degree of reproducibility in cell biology. I think these two statements coin the inherent uncertainty of dealing with biological

data. Promoter analysis, as in Chapter 8, is far from an exact science. Therefore it is necessary to support such results with either literature search or further supportive analyses. I believe that the way to deal with such issues is by combining various sources of information, from independent experiments if possible. This was done indirectly in Chapter 8 by using the Gene Ontology, by using TRANSFAC and EnsEMBL databases, and by studying the literature. Though ideally (some) bioinformaticians would like to explain biology by p-values, I believe that in order to deal with the inherent complexity of biological systems, we have to rely on a wider range of information sources than just pure numeric data.

The data quality of the time series used in Chapter 7 is also very good. However, from a dynamical systems point of view, it would be great to measure more often, but in practice this may not be possible. If blood samples are taken too often, it may start to influence the blood volume directly, and urine samples are not easily collected at precise intervals, if the subject is not using a catheter. In order to study circadian rhythms more profoundly, it would also be nice to have longer data series, since we only looked at two consecutive days in the study. This does of course have limitations as well, due to the amount of time one can keep volunteers hospitalised for measurements.

Models and data

The "real world" relevance of the current research is interesting. The simulation of drinking 1l of water in itself, will very likely not be directly applicable in a drug development process. But the prospect of using such models is great. With a mechanism-based model several hypotheses can be tested *in silico*, and hopefully speed up research work. A model is, in principle, a way to integrate information across different time scales and biological levels. A mathematical model is often thought of as a fixed entity, but I see it as a very flexible. When new information becomes available the model should be updated, extended or altered in another way to account for the new insights. Take the systemic model of Chapter 5. The model does not include the regulation of solute permeability by aldosterone. The model design allows us to include a small dynamic model of sodium channels comparable to the aquaporin model. One could also extend the model by including the renin-angiotensin system or another hormonal control systems. Building and testing a model is, however, an investment of a lot of time (and time = money) and resources, and without the necessary patience it will not be possible to benefit from using models as a development tool.

Chapter 10

Conclusion

In response to the overall goals presented in the introduction, we can conclude that a mechanism-based model of water and salt has been formulated, and that it fits well to experimental data. We have also been successful in applying analysis techniques to study the circadian rhythms of water and solute balance.

A model for regulation of membrane proteins was formulated and its dynamical properties investigated in Chapter 3. The parameters of the model were evaluated with data from experiments on water permeability of kidney collecting duct cells by aquaporins. The data were taken from two related, but independent, experiments providing information on time scales and relative magnitudes. The model was formulated so that it was possible to describe both fast membrane activation regulation, and slower transcriptional regulation. The parameter fit to experimental data indicate that both exocytosis and endocytosis are regulated during vasopressin activated aquaporin translocation.

In Chapter 4, we investigated a model of the regulation of urine flow by water and solute permeability. The model suggested that volume flow is tightly regulated mainly by water permeability, whereas solute flow is more sensitive to both solute and water permeability. The simple model gives a surprisingly realistic estimate of the regulatory range of water permeability. Furthermore, it describes the characteristics of antidiuretic and diuretic states with high and low interstitial concentrations, respectively. This model is a good example of how general characteristics can be captured by simple models.

In Chapter 5, I introduced a model of the regulation of osmolarity by vasopressin. The models presented in Chapters 3 and 4 were incorporated in the systemic model in a framework representing body fluid and solutes. Sim-

ulations successfully reproduced a simple physiological experiment, where a subject drinks 1 l of water, and the urine flow and plasma osmolarity are monitored. This results confirms the validity of the modelling approach and that it is possible to determine parameter values from independent sources.

From the study in Chapter 7, we can conclude that urine flow adapts to a higher fluid intake. Plasma osmolarity and solute levels also tightly controlled under the same conditions. The most well-known antidiuretic substance, vasopressin, is *not* downregulated upon higher fluid intake, but its plasma values remain unchanged. Since glomerular filtration is also unchanged, this means that the magnitude of the renal response to vasopressin is diminished, while the analysis shows that the range of vasopressin levels remains unchanged. Plasma potassium was identified to be circadian in more than half of the normal fluid intake subjects and only a quarter of the high fluid intake subjects. This indicates a possible disruption of the circadian rhythm of potassium as a consequence of increased fluid intake. Most likely the adaptation to high fluid intake that takes place is in interaction between many factors, and the most promising candidates are alterations of ANP and aldosterone levels, and an adaptation of the osmotic environment in the kidney.

In the bioinformatic study in Chapter 8 the transcriptional regulation of blood pressure was investigated with focus on circadian rhythms. 11 transcription factors were identified as overrepresented in a set of 63 genes associated with blood pressure regulation. Of these 11 factors, 5 were also found in a study of circadian, transcriptional regulation by Bozek et al. [23]. The factor Wilm's Tumor Protein 1 (WT1) was among the overrepresented factors, and this underlines its importance in blood pressure regulation, where it is already known to be involved in renin transcription and in kidney related diseases. Furthermore, a more general methodology of evaluating transcription factor-gene association was introduced, and the results were, to some extent, overlapping with the first analysis. The analysis supports the hypothesis of circadian, transcriptional regulation in blood pressure regulating genes.

This present work is composed of many bits and pieces. The work sets the scene for further investigations in mechanism-based modelling of water and salt balance, and it motivates the search for the importance of circadian rhythms across levels of organisation in biology.

Bibliography

[1] Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter (2008). *Molecular Biology of the Cell* (Fifth ed.). Garland Science.

[2] Alpern, R. J. and S. C. Hebert (2007). *Seldin and Giebisch's The Kidney: Physiology & Pathophysiology: Physiology and Pathophysiology: Vols. 1 & 2*. Academic Press.

[3] Ansari, A. R. and R. A. Bradley (1960). Rank-sum tests for dispersion. *Annals of Mathematical Statistics 31*(4), 1174–1189.

[4] Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, and G. Sherlock (2000, May). Gene ontology: tool for the unification of biology. the gene ontology consortium. *Nat Genet 25*(1), 25–29.

[5] Atherton, J. C., R. Green, and S. Thomas (1971, Mar). Influence of lysine-vasopressin dosage on the time course of changes in renal tissue and urinary composition in the conscious rat. *J Physiol 213*(2), 291–309.

[6] Atherton, J. C., R. Green, S. Thomas, and J. A. Wood (1972, May). Time course of changes in renal tissue and urinary composition after cessation of constant infusion of lysine vasopressin in the conscious, hydrated rat. *J Physiol 222*(3), 583–595.

[7] Atherton, J. C., M. A. Hai, and S. Thomas (1968a, Jul). The time course of changes in renal tissue composition during mannitol diuresis in the rat. *J Physiol 197*(2), 411–428.

[8] Atherton, J. C., M. A. Hai, and S. Thomas (1968b, Jul). The time course of changes in renal tissue composition during water diuresis in the rat. *J Physiol 197*(2), 429–443.

[9] Atherton, J. C., M. A. Hai, and S. Thomas (1969). Acute effects of lysine vasopressin injection (single and continuous) on urinary composition in the conscious water diuretic rat. *Pflugers Arch 310*(4), 281–296.

[10] Baldes, E. J. and F. H. Smirk (1934, Aug). The effect of water drinking, mineral starvation and salt administration on the total osmotic pressure of the blood in man, chiefly in relation to the problems of water absorption and water diuresis. *J Physiol 82*(1), 62–74.

[11] Barfred, M., E. Mosekilde, and N.-H. Holstein-Rathlou (1996). Bifurcation analysis of nephron pressure and flow regulation. *Chaos 6*(3), 280–7.

[12] Barlow, M. (2002). Vasopressin. *Emergency Medicine 14*(3), 304–314.

[13] Bassett, M. H., T. Suzuki, H. Sasano, P. C. White, and W. E. Rainey (2004, Feb). The orphan nuclear receptors nurr1 and ngfib regulate adrenal aldosterone production. *Mol Endocrinol 18*(2), 279–290.

[14] Bassett, M. H., Y. Zhang, C. Clyne, P. C. White, and W. E. Rainey (2002, Apr). Differential regulation of aldosterone synthase and 11beta-hydroxylase transcription by steroidogenic factor-1. *J Mol Endocrinol 28*(2), 125–135.

[15] Becker-Weimann, S., J. Wolf, H. Herzel, and A. Kramer (2004, Nov). Modeling feedback loops of the mammalian circadian oscillator. *Biophys J 87*(5), 3023–3034.

[16] Bie, P. (2009, Jan). Blood volume, blood pressure and total body sodium: internal signalling and output control. *Acta Physiol (Oxf) 195*(1), 187–196.

[17] Bie, P. and M. Damkjaer (2009, Oct). Renin secretion and total body sodium: Pathways of integrative control. *Clin Exp Pharmacol Physiol epub*, 1.

[18] Bie, P., S. Wamberg, and M. Kjolby (2004, Aug). Volume natriuresis vs. pressure natriuresis. *Acta Physiol Scand 181*(4), 495–503.

[19] Blüthgen, N., S. M. K. basa, and H. Herzel (2005). Inferring combinatorial regulation of transcription in silico. *Nucleic Acids Res 33*(1), 272–279.

[20] Bodén, M. and T. L. Bailey (2008, Jul). Associating transcription factor-binding site motifs with target go terms and target genes. *Nucleic Acids Res 36*(12), 4108–4117.

[21] Boone, M. and P. M. T. Deen (2008, Sep). Physiology and pathophysiology of the vasopressin-regulated renal water reabsorption. *Pflugers Arch 456*(6), 1005–1024.

[22] Bozek, K., S. M. Kielbasa, A. Kramer, and H. Herzel (2007). Promoter analysis of mammalian clock controlled genes. *Genome Informatics 18*, 65–74.

[23] Bozek, K., A. Relógio, S. M. Kielbasa, M. Heine, C. Dame, A. Kramer, and H. Herzel (2009). Regulation of clock-controlled genes in mammals. *PLoS ONE 4*(3), e4882.

[24] Brown, D. (2003). The ins and outs of aquaporin-2 trafficking. *Am J Physiol Renal Physiol 284*(5), F893–901.

[25] Brunssen, C., S. Korten, M. Brux, S. Seifert, J. Roesler, S. R. Bornstein, H. Morawietz, and W. Goettsch (2009, Oct). Coup-tfii is regulated by high glucose in endothelial cells. *Horm Metab Res* -, –.

[26] Burbach, J. P., S. M. Luckman, D. Murphy, and H. Gainer (2001, Jul). Gene regulation in the magnocellular hypothalamo-neurohypophysial system. *Physiol Rev 81*(3), 1197–1267.

[27] Calkin, A. C. and M. C. Thomas (2008). Ppar agonists and cardiovascular disease in diabetes. *PPAR Res 2008*, 245410.

[28] Call, K. M., T. Glaser, C. Y. Ito, A. J. Buckler, J. Pelletier, D. A. Haber, E. A. Rose, A. Kral, H. Yeger, and W. H. Lewis (1990, Feb). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 wilms' tumor locus. *Cell 60*(3), 509–520.

[29] Charloux, A., C. Gronfier, E. Lonsdorfer-Wolf, F. Piquard, and G. Brandenberger (1999, Jan). Aldosterone release during the sleep-wake cycle in humans. *Am J Physiol 276*(1 Pt 1), E43–E49.

[30] Clamp, M., D. Andrews, D. Barker, P. Bevan, G. Cameron, Y. Chen, L. Clark, T. Cox, J. Cuff, V. Curwen, T. Down, R. Durbin, E. Eyras, J. Gilbert, M. Hammond, T. Hubbard, A. Kasprzyk, D. Keefe, H. Lehvaslaiho, V. Iyer, C. Melsopp, E. Mongin, R. Pettett, S. Potter, A. Rust, E. Schmidt, S. Searle, G. Slater, J. Smith, W. Spooner, A. Stabenau, J. Stalker, E. Stupka, A. Ureta-Vidal, I. Vastrik, and E. Birney (2003, Jan). Ensembl 2002: accommodating comparative genomics. *Nucleic Acids Res 31*(1), 38–42.

[31] Cugini, P. and P. Lucia (2004). Funzione ritmica circadiana del sitema renina-angiotensina-aldosterone: consuntivo delle nostre ricerche. *La Clinica Terapeutica 155*(7-8), 287–291.

[32] Cugini, P., E. Salandi, G. Murano, M. Centanni, and D. Scavo (1983, Aug). Inactive renin of human plasma is a circadian variable. *Biochem Med 30*(1), 119–126.

[33] Danser, A. H. J., W. W. Batenburg, and J. H. M. van Esch (2007, May). Prorenin and the (pro)renin receptor–an update. *Nephrol Dial Transplant 22*(5), 1288–1292.

[34] Davis, C. S. (2003). *Statistical Methods for the Analysis of Repeated Measurements*. Springer.

[35] Discenza, M. T. and J. Pelletier (2004, Feb). Insights into the physiological role of wt1 from studies of genetically modified mice. *Physiol Genomics 16*(3), 287–300.

[36] Duffield, G. E., J. D. Best, B. H. Meurers, A. Bittner, J. J. Loros, and J. C. Dunlap (2002, Apr). Circadian programs of transcriptional activation, signaling, and protein turnover revealed by microarray analysis of mammalian cells. *Curr Biol 12*(7), 551–557.

[37] Eaton, D. C. and J. P. Pooler (2004). *Vander's Renal Physiology* (6th ed.). Lange Physiology Series. McGraw-Hill.

[38] Edwards, A., M. J. Delong, and T. L. Pallone (2000, Feb). Interstitial water and solute recovery by inner medullary vasa recta. *Am J Physiol Renal Physiol 278*(2), F257–F269.

[39] Emamian, S. A., M. B. Nielsen, J. F. Pedersen, and L. Ytte (1993, Jan). Kidney dimensions at sonography: correlation with age, sex, and habitus in 665 adult volunteers. *AJR Am J Roentgenol 160*(1), 83–86.

[40] ENCODE Project Consortium (2007, Jun). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature 447*(7146), 799–816.

[41] Frith, M. C., Y. Fu, L. Yu, J.-F. Chen, U. Hansen, and Z. Weng (2004). Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Res 32*(4), 1372–1381.

[42] Fuller, P. J. and M. J. Young (2005, Dec). Mechanisms of mineralocorticoid action. *Hypertension 46*(6), 1227–1235.

[43] Futschik, M. E. and H. Herzel (2008, Apr). Are we overestimating the number of cell-cycling genes? the impact of background models on time-series analysis. *Bioinformatics 24*(8), 1063–1069.

[44] Gaeggeler, H.-P., E. Gonzalez-Rodriguez, N. F. Jaeger, D. Loffing-Cueni, R. Norregaard, J. Loffing, J.-D. Horisberger, and B. C. Rossier (2005, Apr). Mineralocorticoid versus glucocorticoid receptor occupancy mediating aldosterone-stimulated sodium transport in a novel renal cell line. *J Am Soc Nephrol 16*(4), 878–891.

[45] Gardner, D. G., S. Chen, D. J. Glenn, and C. L. Grigsby (2007, Mar). Molecular biology of the natriuretic peptide system: implications for physiology and hypertension. *Hypertension 49*(3), 419–426.

[46] Geier, F., S. Becker-Weimann, A. Kramer, and H. Herzel (2005). Entrainment in a model of the mammalian circadian oscillator. *Journal of Biological Rhythms 20*(1), 83–93.

[47] Gordon, R. D., L. K. Wolfe, D. P. Island, and G. W. Liddle (1966, Oct). A diurnal rhythm in plasma renin activity in man. *J Clin Invest 45*(10), 1587–1592.

[48] Grundschober, C., F. Delaunay, A. Pühlhofer, G. Triqueneaux, V. Laudet, T. Bartfai, and P. Nef (2001, Dec). Circadian regulation of diverse gene products revealed by mrna expression profiling of synchronized fibroblasts. *J Biol Chem 276*(50), 46751–46758.

[49] Guyton, A. C., T. G. Coleman, A. W. Cowley, J. F. Liard, R. A. Norman, and R. D. Manning (1972, Dec). Systems analysis of arterial pressure regulation and hypertension. *Ann Biomed Eng 1*(2), 254–281.

[50] Guyton, A. C. and J. E. Hall (2006). *Textbook of Medical Physiology* (11th ed.). Elsevier Health Sciences.

[51] Hastings, M., J. S. O'Neill, and E. S. Maywood (2007). Circadian clocks: regulators of endocrine and metabolic rhythms. *J Endocrinol 195*(2), 187–98.

[52] Hodson, J. (1972, Apr). The lobar structure of the kidney. *Br J Urol 44*(2), 246–261.

[53] Holstein-Rathlou, N.-H. and D. J. Marsh (1994). Renal blood flow regulation and arterial pressure fluctuations: A case study in nonlinear dynamics. *Physiological Reviews 74*(3), 637–681.

[54] Houweling, A. C., M. M. van Borren, A. F. M. Moorman, and V. M. Christoffels (2005, Sep). Expression and regulation of the atrial natriuretic factor encoding gene nppa during development and disease. *Cardiovasc Res 67*(4), 583–593.

[55] Hvidsten, T. R., B. Wilczyński, A. Kryshtafovych, J. Tiuryn, J. Komorowski, and K. Fidelis (2005, Jun). Discovering regulatory binding-site modules using rule-based learning. *Genome Res 15*(6), 856–866.

[56] Jin, X., L. P. Shearman, D. R. Weaver, M. J. Zylka, G. J. de Vries, and S. M. Reppert (1999, Jan). A molecular mechanism regulating rhythmic output from the suprachiasmatic circadian clock. *Cell 96*(1), 57–68.

[57] Johnson, D. S., A. Mortazavi, R. M. Myers, and B. Wold (2007, Jun). Genome-wide mapping of in vivo protein-dna interactions. *Science 316*(5830), 1497–1502.

[58] Kato, H., J. Ishida, S. Imagawa, T. Saito, N. Suzuki, T. Matsuoka, T. Sugaya, K. Tanimoto, T. Yokoo, O. Ohneda, F. Sugiyama, K. ichi Yagami, T. Fujita, M. Yamamoto, M. Nangaku, and A. Fukamizu (2005, Dec). Enhanced erythropoiesis mediated by activation of the renin-angiotensin system via angiotensin ii type 1a receptor. *FASEB J 19*(14), 2023–2025.

[59] Katsura, T., D. A. Ausiello, and D. Brown (1996, Mar). Direct demonstration of aquaporin-2 water channel recycling in stably transfected llc-pk1 epithelial cells. *Am J Physiol 270*(3 Pt 2), F548–F553.

[60] Keener, J. and J. Sneyd (1998). *Mathematical Physiology*. New York, NY, USA: Springer-Verlag New York, Inc.

[61] Kim, M. K.-H., T. J. McGarry, P. O. Broin, J. M. Flatow, A. A.-J. Golden, and J. D. Licht (2009, Jul). An integrated genome screen

identifies the wnt signaling pathway as a major target of wt1. *Proc Natl Acad Sci U S A 106*(27), 11154–11159.

[62] Knepper, M. A., R. A. Danielson, G. M. Saidel, and R. S. Post (1977, Nov). Quantitative analysis of renal medullary anatomy in rats and rabbits. *Kidney Int 12*(5), 313–323.

[63] Knepper, M. A. and S. Nielsen (1993). Kinetic model of water and urea permeability regulation by vasopressin in collecting duct. *Am J Physiol Renal Physiol 265*(2), F214–224.

[64] Knöfler, M., G. Meinhardt, S. Bauer, T. Loregger, R. Vasicek, D. J. Bloor, S. J. Kimber, and P. Husslein (2002, Feb). Human hand1 basic helix-loop-helix (bhlh) protein: extra-embryonic expression pattern, interaction partners and identification of its transcriptional repressor domains. *Biochem J 361*(Pt 3), 641–651.

[65] Kolla, V. and G. Litwack (1999, Apr). Upregulation of mineralocorticoid- and glucocorticoid-receptor gene expression by sp-i. *Mol Cell Biol Res Commun 1*(1), 44–47.

[66] Kolla, V. and G. Litwack (2000, Nov). Inhibition of mineralocorticoid-mediated transcription by nf-kappab. *Arch Biochem Biophys 383*(1), 38–45.

[67] Kuwahara, M. and A. S. Verkman (1989, Aug). Pre-steady-state analysis of the turn-on and turn-off of water permeability in the kidney collecting tubule. *J Membr Biol 110*(1), 57–65.

[68] Layton, A. T., T. L. Pannabecker, W. H. Dantzler, and H. E. Layton (2004, Oct). Two modes for concentrating urine in rat inner medulla. *Am J Physiol Renal Physiol 287*(4), F816–F839.

[69] Layton, H. E., E. B. Pitman, and M. A. Knepper (1995). A dynamic numerical method for models of the urine concentrating mechanism. *SIAM Journal on Applied Mathematics 55*(5), 1390–1418.

[70] Lee, G., N. Makhanova, K. Caron, M. L. S. Lopez, R. A. Gomez, O. Smithies, and H.-S. Kim (2005, Jun). Homeostatic responses in the adrenal cortex to the absence of aldosterone in mice. *Endocrinology 146*(6), 2650–2656.

[71] Lemmer, B. (2007). Circadian rhythm regulations of the cardiovascular system. studies in rats and mice. *IEEE Eng Med Biol Mag 26*(6), 30–32.

[72] Lewis, P. R. and M. C. Lobban (1957). Dissociation of Diurnal Rhythms in Human Subjects Living on Abnormal Time Routines. *Q J Exp Physiol Cogn Med Sci 42*(4), 371–386.

[73] Lomax, J. (2005, Sep). Get ready to go! a biologist's guide to the gene ontology. *Brief Bioinform 6*(3), 298–304.

[74] Makhanova, N., G. Lee, N. Takahashi, M. L. S. Lopez, R. A. Gomez, H.-S. Kim, and O. Smithies (2006, Jan). Kidney function in mice lacking aldosterone. *Am J Physiol Renal Physiol 290*(1), F61–F69.

[75] Matsumura, Y., S. Uchida, T. Rai, S. Sasaki, and F. Marumo (1997, Jun). Transcriptional regulation of aquaporin-2 water channel gene by camp. *J Am Soc Nephrol 8*(6), 861–867.

[76] McDonald, J., M. Batuwangala, and D. G. Lambert (2007). Role of urotensin ii and its receptor in health and disease. *J Anesth 21*(3), 378–389.

[77] McGrath, M. F. and A. J. de Bold (2005, Jun). Determinants of natriuretic peptide gene expression. *Peptides 26*(6), 933–943.

[78] Miller, B. H., E. L. McDearmon, S. Panda, K. R. Hayes, J. Zhang, J. L. Andrews, M. P. Antoch, J. R. Walker, K. A. Esser, J. B. Hogenesch, and J. S. Takahashi (2007, Feb). Circadian and clock-controlled regulation of the mouse transcriptome and cell proliferation. *Proc Natl Acad Sci U S A 104*(9), 3342–3347.

[79] Mohammadi, B. and J.-H. Saïac (2003). *Pratique de la Simulation Numérique*. Dunod, Paris.

[80] Nakayama, T. (2005, Feb). The genetic contribution of the natriuretic peptide system to cardiovascular diseases. *Endocr J 52*(1), 11–21.

[81] Nielsen, S., C. L. Chou, D. Marples, E. I. Christensen, B. K. Kishore, and M. A. Knepper (1995, Feb). Vasopressin increases water permeability of kidney collecting duct by inducing translocation of aquaporin-cd water channels to plasma membrane. *Proc Natl Acad Sci U S A 92*(4), 1013–1017.

[82] Niisato, N., A. Taruno, and Y. Marunaka (2007, Sep). Aldosterone-induced modification of osmoregulated enac trafficking. *Biochem Biophys Res Commun 361*(1), 162–168.

[83] Nishimoto, G., M. Zelenina, D. Li, M. Yasui, A. Aperia, S. Nielsen, and A. C. Nairn (1999, Feb). Arginine vasopressin stimulates phosphorylation of aquaporin-2 in rat renal tissue. *Am J Physiol 276*(2 Pt 2), F254–F259.

[84] Ohlsson, H., K. Karlsson, and T. Edlund (1993, Nov). Ipf1, a homeodomain-containing transactivator of the insulin gene. *EMBO J 12*(11), 4251–4259.

[85] O'Neill, J. S., E. S. Maywood, J. E. Chesham, J. S. Takahashi, and M. H. Hastings (2008, May). camp-dependent signaling as a core component of the mammalian circadian pacemaker. *Science 320*(5878), 949–953.

[86] Panda, S., M. P. Antoch, B. H. Miller, A. I. Su, A. B. Schook, M. Straume, P. G. Schultz, S. A. Kay, J. S. Takahashi, and J. B. Hogenesch (2002). Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell 109*, 307–320.

[87] Pavlov, A., O. Sosnovtseva, A. Ziganshin, N.-H.-H. Holstein-Rathlou, and E. Mosekilde (2002). Multiscality in the dynamics of coupled chaotic systems. *Physica A 316*(1-4), 233–249.

[88] Reinhardt, H. W. and E. Seeliger (2000, Dec). Toward an integrative concept of control of total body sodium. *News Physiol Sci 15*, 319–325.

[89] Reppert, S. M. and D. R. Weaver (2002). Coordination of circadian timing in mammals. *Nature 418*(6901), 935.

[90] Rigamonti, D., C. Mutti, C. Zuccato, E. Cattaneo, and A. Contini (2009, Sep). Turning rest/nrsf dysfunction in huntington's disease into a pharmaceutical target. *Curr Pharm Des 1*, 1.

[91] Rittig, S., U. B. Knudsen, M. Jønler, J. P. Nørgaard, E. B. Pedersen, and J. C. Djurhuus (1989). Adult enuresis. the role of vasopressin and atrial natriuretic peptide. *Scand J Urol Nephrol Suppl 125*, 79–86.

[92] Rittig, S., U. B. Knudsen, J. P. Nørgaard, H. Gregersen, E. B. Pedersen, and J. C. Djurhuus (1991, Apr). Diurnal variation of plasma atrial natriuretic peptide in normals and patients with enuresis nocturna. *Scand J Clin Lab Invest 51*(2), 209–217.

[93] Rittig, S., U. B. Knudsen, J. P. Nørgaard, E. B. Pedersen, and J. C. Djurhuus (1989, Apr). Abnormal diurnal rhythm of plasma vasopressin and urinary output in patients with enuresis. *Am J Physiol 256*(4 Pt 2), F664–F671.

[94] Rittig, S., T. B. Matthiesen, E. B. Pedersen, and J. C. Djurhuus (2006). Circadian variation of angiotensin ii and aldosterone in nocturnal enuresis: Relationship to arterial blood pressure and urine output. *The Journal of Urology 176*(2), 774–780.

[95] Sands, J. M., H. Nonoguchi, and M. A. Knepper (1988, Sep). Hormone effects on nacl permeability of rat inner medullary collecting duct. *Am J Physiol 255*(3 Pt 2), F421–F428.

[96] Sassone-Corsi, P. (1998, Jan). Coupling gene expression to camp signalling: role of creb and crem. *Int J Biochem Cell Biol 30*(1), 27–38.

[97] Schweda, F., U. Friis, C. Wagner, O. Skott, and A. Kurtz (2007, Oct). Renin release. *Physiology (Bethesda) 22*, 310–319.

[98] Seebach, F. A., T. Welte, X. Y. Fu, L. H. Block, and M. Kashgarian (2001, Jun). Differential activation of the stat pathway by angiotensin ii via angiotensin type 1 and type 2 receptors in cultured human fetal mesangial cells. *Exp Mol Pathol 70*(3), 265–273.

[99] Seeliger, E., E. Safak, P. B. Persson, and H. W. Reinhardt (2001, Dec). Contribution of pressure natriuresis to control of total body sodium: balance studies in freely moving dogs. *J Physiol 537*(Pt 3), 941–947.

[100] Sharpe, K., L. Ward, J. Cichero, P. Sopade, and P. Halley (2007, Jul). Thickened fluids and water absorption in rats and humans. *Dysphagia 22*(3), 193–203.

[101] Smirk, F. H. (1933, May). The rate of water absorption in man and the relationship of the water load in tissues to diuresis. *J Physiol 78*(2), 113–126.

[102] Snyder, P. M. (2005, Dec). Minireview: regulation of epithelial na+ channel trafficking. *Endocrinology 146*(12), 5079–5085.

[103] Sosnovtseva, O., A. Pavlov, E. Mosekilde, and N.-H. Holstein-Rathlou (2002). Bimodal oscillations in nephron autoregulation. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics 66*(6), 061909/1–061909/7.

[104] Sosnovtseva, O., A. Pavlov, E. Mosekilde, and N.-H. Holstein-Rathlou (2003). Bimodal dynamics in nephron autoregulation. *Physics and Control, 2003. Proceedings. 2003 International Conference 1*, 282–287.

[105] Sosnovtseva, O., A. Pavlov, E. Mosekilde, N.-H. Holstein-Rathlou, and D. Marsh (2005). Double-wavelet approach to studying the modulation properties of nonstationary multimode dynamics. *Physiological Measurement 26*(4), 351–362.

[106] Stauffer, B. L., C. M. Westby, and C. A. DeSouza (2008, Jul). Endothelin-1, aging and hypertension. *Curr Opin Cardiol 23*(4), 350–355.

[107] Steege, A., M. Fähling, A. Paliege, A. Bondke, K. M. Kirschner, P. Martinka, C. Kaps, A. Patzak, P. B. Persson, B. J. Thiele, H. Scholz, and R. Mrowka (2008, Aug). Wilms' tumor protein (-kts) modulates renin gene transcription. *Kidney Int 74*(4), 458–466.

[108] Stephenson, J. L. (1972). Concentration of urine in a central core model of the renal counterflow system. *Kidney International 2*(2), 85–94.

[109] Stephenson, J. L. (1973a). Concentrating Engines and the Kidney: I. Central Core Model of the Renal Medulla. *Biophys. J. 13*(6), 512–545.

[110] Stephenson, J. L. (1973b). Concentrating Engines and the Kidney: II. Multisolute Central Core Systems. *Biophys. J. 13*(6), 546–567.

[111] Stephenson, J. L. (1976). Concentrating engines and the kidney. III. Canonical mass balance equation for multinephron models of the renal medulla. *Biophys. J. 16*(11), 1273–1286.

[112] Storch, K.-F., O. Lipan, I. Leykin, N. Viswanathan, F. C. Davis, W. H. Wong, and C. J. Weitz (2002). Extensive and divergent circadian gene expression in liver and heart. *Nature 417*(6884), 78–83.

[113] Swift, M. R. and B. M. Weinstein (2009, Mar). Arterial-venous specification during development. *Circ Res 104*(5), 576–588.

[114] Thompson, C. J., J. Bland, J. Burd, and P. H. Baylis (1986, Dec). The osmotic thresholds for thirst and vasopressin release are similar in healthy man. *Clin Sci (Lond) 71*(6), 651–656.

[115] Todorov, V. T., M. Desch, N. Schmitt-Nilson, A. Todorova, and A. Kurtz (2007, Nov). Peroxisome proliferator-activated receptor-gamma is involved in the control of renin gene expression. *Hypertension 50*(5), 939–944.

[116] Todorov, V. T., S. Völkl, J. Friedrich, L. A. Kunz-Schughart, T. Hehlgans, L. Vermeulen, G. Haegeman, M. L. Schmitz, and A. Kurtz (2005, Jul). Role of creb1 and nfkappab-p65 in the down-regulation of renin gene expression by tumor necrosis factor alpha. *J Biol Chem 280*(26), 24356–24362.

[117] Todorov, V. T., S. Völkl, M. Müller, A. Bohla, J. Klar, L. A. Kunz-Schughart, T. Hehlgans, and A. Kurtz (2004, Jan). Tumor necrosis factor-alpha activates nfkappab to inhibit renin transcription by targeting camp-responsive element. *J Biol Chem 279*(2), 1458–1467.

[118] Ueda, H. R., W. Chen, A. Adachi, H. Wakamatsu, S. Hayashi, T. Takasugi, M. Nagano, K. ichi Nakahama, Y. Suzuki, S. Sugano, M. Iino, Y. Shigeyoshi, and S. Hashimoto (2002, Aug). A transcription factor response element for gene expression during circadian night. *Nature 418*(6897), 534–539.

[119] Uttamsingh, R., M. Leaning, J. Bushman, E. Carson, and L. Finkelstein (1985). Mathematical model of the human renal system. *Medical & Biological Engineering & Computing 23*(6), 525–535.

[120] Val, S. D. and B. L. Black (2009, Feb). Transcriptional control of endothelial cell development. *Dev Cell 16*(2), 180–195.

[121] Warberg, J. (2005). *Human Fysiologi - en grundbog* (5th ed.). Polyteknisk Forlag.

[122] Wasserman, W. W. and A. Sandelin (2004, Apr). Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet 5*(4), 276–287.

[123] Watson, J. D., T. A. Baker, S. P. Bell, A. Gann, M. Levine, and R. Losick (2008). *Molecular Biology of the Gene* (6th ed.). Cold Spring Harbor Laboratory Press.

[124] Wexler, A. S., R. E. Kalaba, and D. J. Marsh (1991a). Three-dimensional anatomy and renal concentrating mechanism. I. Modeling results. *Am J Physiol Renal Physiol 260*(3), F368–383.

[125] Wexler, A. S., R. E. Kalaba, and D. J. Marsh (1991b). Three-dimensional anatomy and renal concentrating mechanism. II. Sensitivity results. *Am J Physiol Renal Physiol 260*(3), F384–394.

[126] White, W. B. (2007, Oct). Importance of blood pressure control over a 24-hour period. *J Manag Care Pharm 13*(8 Suppl B), 34–39.

[127] Wingender, E., P. Dietze, H. Karas, and R. Knüppel (1996, Jan). Transfac: a database on transcription factors and their dna binding sites. *Nucleic Acids Res 24*(1), 238–241.

[128] Wong, D. L. (2003). Why is the adrenal adrenergic? *Endocr Pathol 14*(1), 25–36.

[129] Yan, J., H. Wang, Y. Liu, and C. Shao (2008, Oct). Analysis of gene regulatory networks in the mammalian circadian rhythm. *PLoS Comput Biol 4*(10), e1000193.

[130] Yang, L., Y. Han, F. S. Saiz, F. S. Saiz, and M. D. Minden (2007, May). A tumor suppressor and oncogene: the wt1 story. *Leukemia 21*(5), 868–876.

[131] Yang, X., M. Downes, R. T. Yu, A. L. Bookout, W. He, M. Straume, D. J. Mangelsdorf, and R. M. Evans (2006). Nuclear receptor expression links the circadian clock to metabolism. *Cell 126*(4), 801–810.

[132] Yasui, M., S. M. Zelenin, G. Celsi, and A. Aperia (1997, Apr). Adenylate cyclase-coupled vasopressin receptor activates aqp2 promoter via a dual effect on cre and ap1 elements. *Am J Physiol 272*(4 Pt 2), F443–F450.

[133] Youn, J. H. and A. A. McDonough (2009, Aug). Recent advances in understanding integrative control of potassium homeostasis. *Annu Rev Physiol 71*, 381–401.

[134] Zöllner, S., K. H. Hwang, B. Wilzewski, C. Carapito, E. Leize-Wagner, A. V. Dorsselaer, and R. Bernhardt (2008, Oct). Aldosterone: From biosynthesis to non-genomic action onto the proteome. *Steroids 73*(9-10), 966–972.

Appendix A

Promoter analysis data

This appendix contains data and results related to the promoter analysis in chapter 8. Table A.1 is a list of the 63 genes selected from the gene ontology. Tables A.2-A.8 are the results of the clover analysis. Table A.2 is for all 63 genes, and the subsequent tables are the results for the various subgroups defined by figure 8.4.

Ensembl Gene ID	Associated Gene Name
ENSG00000005187	ACSM3
ENSG00000043591	ADRB1
ENSG00000044012	GUCA2B
ENSG00000049247	UTS2
ENSG00000073756	PTGS2
ENSG00000078401	EDN1
ENSG00000095303	PTGS1
ENSG00000100292	HMOX1
ENSG00000100604	CHGA
ENSG00000102032	RENBP
ENSG00000102239	BRS3
ENSG00000108405	P2RX1
ENSG00000110680	CALCA
ENSG00000111664	GNB3
ENSG00000113389	NPR3
ENSG00000115138	POMC
ENSG00000120907	ADRA1A
ENSG00000120915	EPHX2
ENSG00000120937	NPPB
ENSG00000122585	NPY
ENSG00000129991	TNNI3
ENSG00000130234	ACE2
ENSG00000131979	GCH1
ENSG00000132170	PPARG
ENSG00000135744	AGT
ENSG00000136160	EDNRB
ENSG00000139567	ACVRL1
ENSG00000140199	SLC12A6
ENSG00000142168	SOD1
ENSG00000143839	REN
ENSG00000144891	AGTR1
ENSG00000145244	CORIN
ENSG00000148408	CACNA1B
ENSG00000151577	DRD3
ENSG00000157168	NRG1
ENSG00000159640	ACE
ENSG00000159899	NPR2
ENSG00000160211	G6PD
ENSG00000163273	NPPC
ENSG00000163399	ATP1A1
ENSG00000164116	GUCY1A3
ENSG00000164128	NPY1R
ENSG00000164307	ARTS-1
ENSG00000164308	LRAP
ENSG00000164326	CARTPT
ENSG00000166148	AVPR1A
ENSG00000169252	ADRB2
ENSG00000169418	NPR1
ENSG00000169676	DRD5
ENSG00000170214	ADRA1B
ENSG00000171873	ADRA1D
ENSG00000175206	NPPA
ENSG00000175344	CHRNA7
ENSG00000176358	TAC4
ENSG00000176387	HSD11B2
ENSG00000179142	CYP11B2
ENSG00000180772	AGTR2
ENSG00000180914	OXTR
ENSG00000182220	ATP6AP2
ENSG00000187848	P2RX2
ENSG00000188170	HBD
ENSG00000198049	AVPR1B
ENSG00000215644	GCGR

Table A.1: List of genes of children terms of GO:0008217 regulation of blood pressure

Transfac ID	Raw score	P-value
Overrepresented		
V$TBP_Q6	29	0
V$TATA_01	28.2	0
V$TATA_C	27.5	0
V$NRSF_01	8.9	0
V$TFIIA_Q6	11	0
V$NRSF_Q4	7.19	0.01
V$ZNF219_01	10.5	0.01
V$NRSE_B	4.23	0.02
V$MAZ_Q6	46.8	0.02
V$CREB_Q2_01	2.65	0.02
V$WT1_Q6	25.1	0.02
V$CRX_Q4	6.33	0.03
V$HAND1E47_01	5.92	0.03
V$GATA_Q6	-0.77	0.04
V$GATA1_04	0.67	0.04
V$AP2_Q3	4.94	0.05
V$IPF1_Q4_01	2.37	0.05
Underrepresented		
V$MYB_Q5_01	-2.9	0.95
V$ELK1_01	0.31	0.95
V$FXR_Q3	-4.22	0.96
V$LXR_Q3	-5.18	0.96
V$AIRE_02	-11.8	0.96
V$HIF1_Q3	-5.16	0.96
V$MIF1_01	-4.74	0.97
V$RBPJK_Q4	-2.76	0.97
V$HNF3_Q6_01	0.64	0.97
V$NFY_Q6	-2.84	0.97
V$BACH2_01	-4.46	0.98
V$GABP_B	-1.69	0.98
V$FOXO3_01	-4.33	0.98
V$SRY_02	-2.61	0.98
V$NFY_01	-2.17	0.98
V$SOX9_B1	-4.04	0.98
V$PEA3_Q6	0.08	0.98
V$BRCA_01	-1.95	0.99
V$NKX25_Q5	-2.5	0.99
V$HIF1_Q5	-4.9	0.99
V$SOX10_Q6	-1.67	0.99
V$DR4_Q2	-4.02	0.99
V$ALPHACP1_01	-3.61	0.99
V$AP2REP_01	-1.12	0.99
V$SMAD4_Q6	-3.47	0.99
V$HNF6_Q6	-4.56	0.99
V$AP1_Q2_01	-3.32	0.99
V$ERR1_Q2	-3.8	0.99
V$XBP1_01	-6.23	1
V$YY1_Q6_02	-1.99	1
V$PEBP_Q6	-4.45	1
V$FREAC4_01	-6.29	1
V$NFY_C	-4.51	1
V$NRF1_Q6	-4.9	1
V$NFMUE1_Q6	-4.26	1
V$YY1_Q6	-2.47	1
V$HES1_Q2	-3.85	1

Table A.2: Over- and underrepresented motifs in set of 63 blood pressure regulation annotated genes.

Transfac ID	Raw score	P-value
Overrepresented		
V$TBP_01	9.33	0
V$TATA_01	14.9	0
V$TATA_C	11.4	0
V$TBP_Q6	9.42	0
V$MYC_Q2	1.39	0.01
V$CREB_Q2_01	2.67	0.02
V$HAND1E47_01	3.87	0.02
V$ATF3_Q6	3.41	0.02
V$AP1FJ_Q2	2.67	0.02
V$CRX_Q4	3.74	0.03
V$RSRFC4_01	1.94	0.03
V$SMAD3_Q6	3.78	0.03
V$DEC_Q1	2.79	0.03
V$CLOCKBMAL_Q6	0.92	0.03
V$MEF2_02	2.48	0.03
V$USF2_Q6	0.7	0.04
V$MEF2_Q6_01	5.12	0.04
V$PPAR_DR1_Q2	3.94	0.04
V$AP2ALPHA_02	0.38	0.05
V$SRF_Q6	0.3	0.05
V$OTX_Q1	2.24	0.05
Underrepresented		
V$PR_Q2	-1.73	0.96
V$GCNF_01	-7.75	0.97
V$MYB_Q6	-1.95	0.97
V$AREB6_04	-1.79	0.98

Table A.3: Over- and underrepresented motifs in set of 9 hormone activity annotated genes.

Transfac ID	Raw score	P-value
Overrepresented		
V$TBP_Q6	6.36	0.01
V$GATA1_05	2.45	0.01
V$WHN_B	0.27	0.01
V$AP2_Q3	3.71	0.01
V$FXR_IR1_Q6	1.55	0.02
V$PR_Q2	3.96	0.02
V$NFKB_Q6	3.42	0.02
V$TATA_01	4.63	0.03
V$TBX5_Q5	3.33	0.03
V$E2F1_Q3_01	1.06	0.03
V$MEIS1_01	2.4	0.03
V$HNF4_Q6	3.01	0.03
V$CREB_Q3	0.04	0.04
V$BLIMP1_Q6	6.64	0.05
V$P53_01	-4.56	0.05
Underrepresented		
V$MYCMAX_B	-2.85	0.96
V$PAX5_02	-4.11	0.96
V$ARNT_02	-6.44	0.96
V$HSF1_01	-2.53	0.97
V$NRF1_Q6	-3.82	0.97
V$OCT1_02	-4.4	0.97
V$SREBP1_01	-5	0.98
V$USF_01	-5.41	0.98
V$HIF1_Q3	-4.15	0.98
V$CREB_01	-4.82	0.98
V$AML1_01	-2.76	0.98
V$YY1_Q6	-2.05	0.99
V$MYCMAX_03	-5.77	0.99
V$MYCMAX_01	-8.24	0.99
V$ATF1_Q6	-3.3	0.99
V$MYCMAX_02	-4.22	0.99
V$TFE_Q6	-3.69	0.99
V$SMAD4_Q6	-2.76	0.99
V$AP2ALPHA_02	-4.47	0.99
V$ATF_B	-4.38	0.99
V$TAL1BETAE47_01	-5.92	0.99
V$ARNT_01	-6.55	0.99
V$AML1_Q6	-2.41	1
V$AHRARNT_02	-10.3	1
V$MAX_01	-8.06	1

Table A.4: Over- and underrepresented motifs in set of 7 peptidase activity annotated genes.

Transfac ID	Raw score	P-value
Overrepresented		
V$ZNF219_01	10.8	0
V$NRSF_01	7.81	0.01
V$E2F1_Q3_01	0.19	0.03
V$NRSE_B	2.11	0.03
V$PAX2_02	12.4	0.03
V$GATA1_04	0.88	0.04
V$MAZ_Q6	21.3	0.04
V$WT1_Q6	11.7	0.05
V$AIRE_01	3.14	0.05
Underrepresented		
V$DEC_Q1	-2.42	0.95
V$HMEF2_Q6	-5.49	0.95
V$TEL2_Q6	-1.25	0.96
V$NFY_C	-3.7	0.96
V$MTF1_Q4	-3.84	0.96
V$CEBP_Q3	-0.29	0.96
V$SOX9_B1	-3.8	0.96
V$MAF_Q6_01	-1.74	0.96
V$FOXO1_02	-4.2	0.96
V$MEF2_Q6_01	-3.09	0.97
V$ELK1_02	-2.02	0.97
V$AR_01	-5.12	0.97
V$MYC_Q2	-3.7	0.97
V$YY1_Q6	-1.79	0.97
V$FREAC4_01	-5.62	0.97
V$BRCA_01	-1.78	0.97
V$AP2REP_01	-0.88	0.97
V$ATF4_Q2	-3.33	0.97
V$SZF11_01	-3.9	0.97
V$PEA3_Q6	-1.09	0.98
V$HNF4_Q6_03	-0.3	0.98
V$RBPJK_Q4	-2.96	0.98
V$ELK1_01	-1.79	0.98
V$GCM_Q2	-3.31	0.98
V$TCF11MAFG_01	-5.13	0.98
V$NRF1_Q6	-4.35	0.98
V$CEBP_Q2_01	-2.3	0.99
V$DR4_Q2	-3.76	0.99
V$HLF_01	-4.92	0.99
V$TAXCREB_01	-5.5	0.99
V$AP1_Q2_01	-3.03	0.99
V$HES1_Q2	-3.3	0.99
V$ERR1_Q2	-3.71	0.99
V$ELF1_Q6	-0.47	0.99
V$FREAC3_01	-5.13	0.99
V$FOXO4_02	-4.89	0.99
V$YY1_Q6_02	-2.21	0.99
V$FOXO3_01	-5.12	1
V$CREL_01	-3.06	1
V$SOX10_Q6	-2.04	1
V$AREB6_01	-3.62	1
V$NKX25_Q5	-3.07	1
V$MYCMAX_B	-3.92	1
V$NFKAPPAB65_01	-4.5	1
V$CEBPDELTA_Q6	-3.74	1
V$NFMUE1_Q6	-4.19	1

Table A.5: Over- and underrepresented motifs in set of 29 receptor activity annotated genes.

Transfac ID	Raw score	P-value
Overrepresented		
V$AP2_Q3	8.71	0
V$TBP_Q6	15	0
V$PU1_Q6	24.1	0.01
V$SP1_Q6_01	25.2	0.01
V$GATA_C	4.17	0.01
V$SP1_01	11	0.02
V$GATA1_05	2.65	0.02
V$NFKB_Q6	4.06	0.02
V$TATA_01	6.95	0.03
V$TATA_C	9.24	0.03
V$GATA1_04	1.6	0.03
V$LUN1_01	-0.98	0.03
V$PAX2_02	11.6	0.03
V$GATA4_Q3	15.4	0.04
V$TGIF_01	2.6	0.04
V$TBX5_Q5	3.24	0.04
V$AREB6_04	5.6	0.04
V$CRX_Q4	3.85	0.04
V$CEBP_Q3	10.2	0.04
Underrepresented		
V$PITX2_Q2	-3.04	0.96
V$PAX6_01	-5.39	0.97
V$AP1_Q2	-2.79	0.97
V$AP1_Q4	-3.1	0.97
V$STAT3_01	-10.9	0.97
V$XBP1_01	-5.5	0.97
V$MRF2_01	-3.66	0.97
V$AP1FJ_Q2	-2.81	0.98
V$ALPHACP1_01	-3.63	0.98
V$ERR1_Q2	-3.52	0.98
V$MIF1_01	-4.47	0.98
V$HMEF2_Q6	-5.91	0.99
V$GATA2_01	-2.98	0.99
V$SMAD4_Q6	-3.33	0.99
V$RORA1_01	-4.92	1
V$SOX9_B1	-4.38	1
V$BACH2_01	-4.5	1

Table A.6: Over- and underrepresented motifs in set of 25 catalytic activity annotated genes.

Transfac ID	Raw score	P-value
Overrepresented		
V$TFIIA_Q6	6.17	0
V$CLOCKBMAL_Q6	2.56	0
V$OCT1_02	3.27	0.01
V$USF_C	1.7	0.02
V$USF_Q6	2.29	0.02
V$PAX6_01	1.16	0.03
V$PPARA_02	0.92	0.03
V$RFX_Q6	0.99	0.04
V$TFE_Q6	1.38	0.04
V$EBOX_Q6_01	3.2	0.04
V$PAX1_B	-1.72	0.05
Underrepresented		
V$E2A_Q2	-1.7	0.96
V$YY1_Q6_02	-1.67	0.96
V$TGIF_01	-3.11	0.96
V$HES1_Q2	-2.36	0.96
V$AREB6_02	-2.37	0.96
V$HSF_Q6	-2.91	0.96
V$LEF1TCF1_Q4	-2.97	0.97
V$SMAD3_Q6	-2.14	0.97
V$AFP1_Q6	-3.01	0.97
V$STAT3_02	0.41	0.97
V$LUN1_01	-10.3	0.97
V$MYB_Q5_01	-1.88	0.97
V$PXR_Q2	-2.47	0.97
V$NKX61_01	-4.42	0.98
V$XBP1_01	-5.13	0.98
V$ATF4_Q2	-2.86	0.98
V$WHN_B	-4.52	0.98
V$GCNF_01	-8.31	0.99
V$SMAD4_Q6	-2.86	0.99
V$NFAT_Q4_01	-1.76	0.99
V$AP1_Q2	-2.71	0.99
V$AP1_Q2_01	-2.7	0.99
V$NFY_C	-4.13	0.99
V$ATF3_Q6	-3.88	0.99
V$NFAT_Q6	-1.87	0.99
V$AP1FJ_Q2	-2.67	1
V$AP1_Q4	-3.09	1
V$SMAD_Q6_01	-2.68	1
V$AP2_Q3	-2.49	1
V$STAT6_02	0.24	1
V$MYB_Q3	-2.51	1
V$TAXCREB_01	-5.29	1

Table A.7: Over- and underrepresented motifs in set of 8 transporter activity annotated genes.

Transfac ID	Raw score	P-value
Overrepresented		
V$MZF1_02	6.61	0
V$HNF1_01	9.06	0
V$FREAC7_01	9.15	0
V$RSRFC4_01	4.89	0
V$FOX_Q2	7.69	0
V$HNF3_Q6	8.75	0
V$MEF2_02	5.82	0
V$HFH3_01	7.3	0.01
V$TBP_01	5.58	0.01
V$ER_Q6	4.13	0.01
V$HNF1_C	6.19	0.01
V$AML1_Q6	2.71	0.01
V$HFH4_01	5.55	0.01
V$HNF6_Q6	4.24	0.01
V$FOXJ2_01	8.41	0.01
V$HNF1_Q6	6.52	0.01
V$HNF3_Q6_01	8.4	0.01
V$TATA_01	5.08	0.02
V$GATA4_Q3	5.3	0.02
V$FOXO4_01	4.37	0.02
V$MEIS1_01	3.21	0.02
V$FOXD3_01	6.19	0.02
V$HNF1_Q6_01	6.79	0.02
V$AFP1_Q6	2.94	0.02
V$CDP_01	0.6	0.02
V$AML1_01	1.7	0.02
V$FOXO3_01	3.51	0.02
V$HNF3ALPHA_Q6	4.99	0.03
V$FREAC3_01	3.19	0.03
V$AP2ALPHA_02	1.39	0.03
V$TFIIA_Q6	2.75	0.03
V$POU3F2_01	2.51	0.03
V$CDP_02	-0.15	0.03
V$NKX3A_01	2.17	0.03
V$FOXO1_01	3.63	0.03
V$FAC1_01	5.59	0.03
V$FREAC2_01	2.53	0.03
V$WT1_Q6	2.59	0.04
V$E4BP4_01	0.19	0.04
V$NKX62_Q2	4.06	0.04
V$RSRFC4_Q2	1.83	0.04
V$NKX61_01	2.38	0.04
V$GR_Q6	3.48	0.04
V$OSF2_Q6	1.67	0.04
V$ETS2_B	4.04	0.05
V$LHX3_01	1.43	0.05
V$CDX_Q5	3.87	0.05
V$STAT_01	2.44	0.05
Underrepresented		
V$GATA6_01	-2.27	0.95
V$E2F_Q4	-2.7	0.96
V$CREBP1CJUN_01	-5.37	0.96
V$HSF1_01	-2.27	0.96
V$HIF1_Q3	-3.74	0.96
V$CREB_02	-3.15	0.97
V$HOXA7_01	-3.23	0.97
V$NFKB_Q6_01	-3.4	0.98
V$NRF1_Q6	-3.89	0.99
V$XBP1_01	-5.28	0.99
V$CREL_01	-2.6	0.99
V$NFKAPPAB65_01	-4.28	1

Table A.8: Over- and underrepresented motifs in set of 4 enzyme regulator activity annotated genes.