

Systems medicine Lecture notes

Uri Alon (Spring 2019)

<https://youtu.be/INzjiUtrcmo>

Lecture 4

Two-gland feedback in the stress-hormone axis generates seasonal clocks and explains clinical phenomena with a timescale of months

[Avichai Tendler, Yael Korem, Alon Bar, Omer Karin, Lior Maimon, Tomer Milo, Michal Shilo, Avi Mayo, Uri Alon, 2019]

Introduction

So far we discussed a single tissue controlling its own size, with beta-cells as our main example. We now explore what happens when *two* tissues control each other's sizes. We will see that this creates a feedback loop with a timescale of months, which can produce damped oscillations. We will explore such feedback in the HPA axis- the bodies' main stress-response pathway whose output is the hormone **cortisol**. The HPA pathway is clinically important: it is used to suppress the immune system using cortisol-analogous called **steroids**, and is also involved in problems of chronic stress and in mood disorders such as major depression and bipolar disorder.

We will see how the two-gland feedback loop creates seasonal hormone oscillations with specific phases. These seasonal oscillations provide a proper hormonal set-point for different seasons. The two-gland feedback also explains clinical phenomena with month's timescale, such as the side-effects of steroid treatment and certain forms of post-partum depression. Finally, we will explore the possible origin of mood fluctuations that last months in bipolar disorder.

The HPA axis controls the response to physical and psychological stress

When we wake up we get a surge of energy (... although for some of us this takes a bit of time), thanks to a morning surge in the blood concentration of the hormone **cortisol**. We get a similar surge of cortisol if we suddenly see a threat, such as a large bear coming at us. The surge focuses the mind and gets us ready to run or fight. We get a cortisol surge even if we *predict* that a threat is coming, such as a serious psychological stress. The hormone cortisol takes minutes to act, and a surge lasts for about 90 min. Additional stress responses within seconds occur through the nervous system and secretion of adrenalin (which we won't discuss). Without cortisol and adrenaline, we wouldn't have the energy to stand up.

The hormone system that controls cortisol is composed of a series of three glands, which we will denote H, P and A, and is thus called **the HPA axis** (Fig 4.1). H stands for the hypothalamus, a brain region between our ears that has neural inputs from many other brain regions, and converts this information to hormones that act on the rest of the body. H is activated by various emergencies and stresses, including low glucose, low blood pressure, inflammation, alcohol and psychological stresses. It is also regulated by the brain's circadian clock which keeps track of the time of day.

All of these neural inputs are processed in H, and combined¹ into the output of H: an increase the rate of secretion of the hormone CRH, which we will call x_1 . This hormone is secreted into a private blood vessel, called the portal system, which flows into the next gland P, a pea-sized gland called the pituitary, which lies at the bottom of the skull.

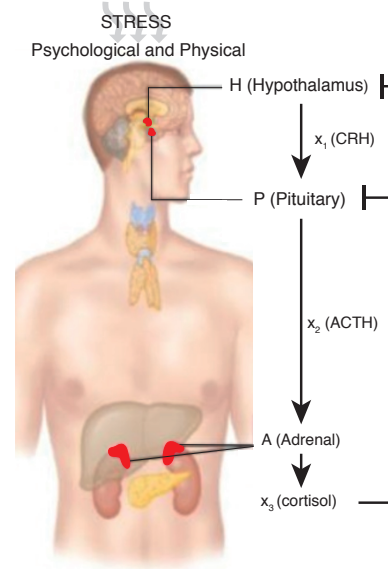


Figure 4.1

In P, the hormone x_1 stimulates cells to secrete the hormone x_2 (ACTH) into the blood stream. Together with x_2 , P also secretes beta-endorphin, which is a pain-killer and causes euphoria (similar to morphine).

The messenger hormone x_2 flows with the blood and reached the third gland of the axis, A, the adrenal cortex. The word 'cortex' refers to the outer layer of the adrenal gland that lies on top of the kidneys. The cells in A secrete the final hormone, cortisol, denoted x_3 . Cortisol closes a negative feedback loop by inhibiting the rate of production and secretion of its two upstream hormones, x_1 and x_2 (Fig 4.1).

Cortisol is a small fat-like hormone (steroid hormone). Because it is fatty, it can penetrate the membranes of all of the cells in the body. It is sensed by special receptors inside the cells. When these receptors bind cortisol, they go into the cell nucleus, bind the DNA at specific sites, and cause the expression of genes that respond to the stress. Cortisol has many effects, including increasing the production of glucose, increasing blood pressure and suppressing inflammation. It thus gets the body ready to respond to the stress.

Sudden stresses cause an **acute** response, in which cortisol levels rise from their normal range of 100-300 nM to close to 1000nM. These pulses last about 90 min, the half-life of cortisol. If the stress lasts for weeks or more, as in psychological worries, cortisol causes the symptoms of **chronic stress**: weight gain, high blood pressure, risk of heart problems, bone loss and depression. Chronic high cortisol also causes cognitive changes, including increased sensitivity to negative stimuli and heightened anxiety, as well as decreased learning (technically, decreased inhibition of the amygdala by the prefrontal cortex, and damage to the hippocampus). That is one reason we take nice deep sighs of relief in our class (deep breaths work on the fast adrenaline system, and

¹ How the inputs are combined is an open question: is it a product or a sum of the inputs?

also cause a nice perception of the class as a pleasant place, which reduces stress). The HPA axis is in fact implicated in mood disorders, as we will discuss later in this chapter.

Cortisol for optimal functioning thus needs to be in a middle range. Insufficient levels result in low energy, and excessive levels over time cause the symptoms of chronic stress.

The classic model for the HPA axis explains responses on the scale of hours, but not on the scale of months

The classic description of the HPA axis described so far is what you will find in textbooks. Let's write it down mathematically so we can explore its properties. Our main conclusion is that the model shows the expected response to acute stress, but cannot explain phenomena on the scale of weeks-months. We will later expand the model by adding control of gland sizes, and open up a world of phenomena on the month's timescale.

The first hormone x_1 is secreted by H upon input u , at a rate q_1 per cell, and is degraded with a timescale of minutes at rate α_1 . This can be described as follows, using the same logic as we did for the glucose minimal model of lecture 1:

$$\frac{dx_1}{dt} = q_1 \frac{H u}{x_3} - \alpha_1 x_1.$$

the inhibitory effect of x_3 on the secretion of x_1 is described² by the term $1/x_3$, so that the more x_3 , the less production of x_1 .

The hormone x_1 acts on P cells to induce secretion of the second hormone x_2 . Secretion is inhibited by x_3 . The hormone x_2 has a half-life of $1/\alpha_2 = 10$ minutes, and so

$$\frac{dx_2}{dt} = q_2 \frac{P x_1}{x_3} - \alpha_2 x_2$$

The hormone x_2 in turn induces the cells of the adrenal cortex A to secrete x_3 , whose half-life is $1/\alpha_3=90\text{min}$

$$\frac{dx_3}{dt} = q_3 A x_2 - \alpha_3 x_3$$

Solving the steady-state, we set $d/dt=0$. The cortisol steady-state depends on all parameters, and rises with the input u and the sizes of the two glands A and P

$$(1) x_3^{st} = \left(\frac{q_1 q_2 q_3}{a_1 a_2 a_3} u HPA \right)^{\frac{1}{3}} \sim u^{\frac{1}{3}} P^{\frac{1}{3}} A^{\frac{1}{3}}$$

Similarly, the steady states of the other two hormones also depend on all parameters. x_2 rises with the size of P but drops with the size of A, because A inhibits its upstream hormone x_1

$$(2) x_2^{st} \sim u^{\frac{1}{3}} A^{-\frac{2}{3}} P^{\frac{1}{3}}$$

² Binding of cortisol affects production as $1/(K + x_3)$ where K is the binding affinity to the receptor- as explained in Appendix A. The $1/x_3$ term is a good approximation when x_3 concentration exceeds K, as it does for cortisol whose blood levels are on the order of 100nM and K is on the order of 10nM.

$$(2) \quad x_1^{st} \sim u^{\frac{2}{3}} P^{-\frac{1}{3}} A^{-\frac{1}{3}}$$

These equations correctly show that the hormone levels rise when stress input u occurs, with a timescale of minutes-hours.

The equations do not explain phenomena on the scale of weeks-months:

- (i) effect of taking steroids to suppress inflammation – if taken for weeks or more, these drugs shouldn't be stopped at once. Steroids need to be gradually tapered down, to avoid risk of dangerously low levels of cortisol.
- (ii) HPA hormone imbalances that occur for months after pregnancy.
- (iii) the months' timescale of depression and bipolar disorder.

Moreover, the glands A and P are made of cells that divide and are removed (the H gland is in the brain and, like most brain neuronal tissues does not show sizable turnover of cells- most neurons don't divide). To maintain proper size, the A and P glands must have feedback control as discussed in lecture 2. Since the glands act in a series of hormone secretions, it turns out that they control each other's growth. This situation is described next.

The HPA axis shows a feedback loop in which two glands control each other's size

The cells of the A and P glands proliferate and are removed with a typical turnover time of about 2 months. They have feedback loops that balance cell proliferation and removal. The input to these loops are the HPA hormones. The hormone x_1 acts not only to induce secretion of x_2 by P cells, it also increases the cells proliferation rate. Similarly, x_2 acts to induce x_3 secretion and also to induce adrenal cortex cell proliferation (Fig 4.2).

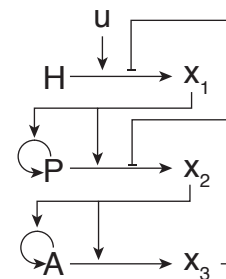


Figure 4.2

Thus, the HPA axis shows two occurrences of the motif we saw for insulin-glucose: the input signal controls both secretion and proliferation. In the HPA axis, the two motifs are stacked on top of each other.

The equations for P and A mass are the difference between proliferation and removal

$$(4) \quad \frac{dP}{dt} = P(b_P x_1 - a_P)$$

$$(5) \quad \frac{dA}{dt} = A(b_A x_2 - a_A)$$

Here a_P, a_A are the cell removal rates ($a_A, a_P \sim 1/2$ months), and b_P, b_A are the hormone-dependent proliferation rates.

These two equations have important consequences for the system on the scale of months. Here we come to the important concept of **separation of timescales**. The hormones work on the timescale of hours, which is much faster than the timescale of months for changes in gland size A and P. Thus, the steady states from Eq (1-3) correctly describe what happens if input u is presented

for a few hours. But if u is present for weeks, A and P sizes change gradually, and the system finds its full steady state (Fig 3). Thus, Eq 1-3 are more accurately described as 'quasi-steady-states'.

When we solve the steady-states of the 'slow equations' Eq 4-5, we find that the steady-state hormone levels become much less sensitive to the physiological parameters of the HPA axis. In particular, the only way to get steady state is $x_{1,st} = a_P/b_P$, $x_{2,st} = a_A/b_A$.

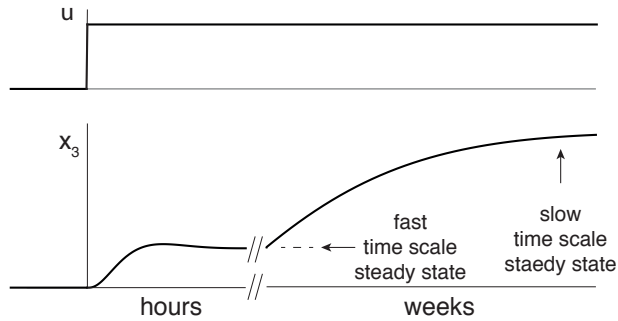


Figure 4.3

These hormone levels thus do not depend at all on production per cell q_1, q_2, q_3 , removal rates etc. Cortisol similarly turns out to have a simple steady state level robust to almost all model parameters (by plugging in x_1^{st} to Eq 1):

$$(6) x_3^{st} = \frac{q_1 b_P}{\alpha_1 a_P} u$$

Nicely, cortisol depends on the input u , as it should, to provide response to stress inputs. It does not, however, depend on q_2, q_3 , which vary, for example, with blood volume. As for the case of glucose, the circuit protects cortisol levels from variations in many parameters, by means of slow growth and shrinkage of the glands.

Equations for the timescale of months can be derived by a quasi-steady state approximation for the hormones. This means using the fact that on the timescale of months, hormones can be assumed to be at their fast steady state levels, Eq 1-3. Plugging in x_1 and x_2 in to Eq 4-5, shows a negative feedback loop between P and A

$$(7) \frac{dP}{dt} = P(c_P u^{\frac{2}{3}} P^{-\frac{1}{3}} A^{-\frac{1}{3}} - a_P)$$

$$(8) \frac{dA}{dt} = A(c_A u^{\frac{1}{3}} P^{\frac{1}{3}} A^{-\frac{2}{3}} - a_A)$$

P acts to increase the size of A , by the action of x_2 (Fig 4). In contrast A acts to shrink P , because the hormone it secretes, cortisol, reduces x_1 which is the growth factor for P .

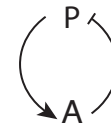


Figure 4.4

Mis-sensing mutants exemplify difference between A and P

We can briefly ask about resistance to mis-sensing mutants, which according to lectures 2 and 3 pose a threat to take over the glands and hyper-secrete the hormones. There appears to be autoimmune removal of mutants from A , as evidenced by the autoimmune disease that attacks A cells called Addison's disease (prevalence of about 10^{-4} in the population). The adrenal cortex is

gradually destroyed, and when 90% of it is gone, clinical symptoms set in due to lack of cortisol. The treatment is to take cortisol pills.

In contrast, the pituitary, P, is a gland that very rarely get autoimmune diseases (estimated at less than 500 cases in 50 years). It is predicted therefore, according to lecture 3, that hyper-secreting mutant expansions will be relatively common. Indeed, the pituitary P shows tumors at rate that is about 100 times larger per cell than the adrenal A. Some of these tumors hyper-secrete the hormone x_2 , and cause high cortisol (Cushing's syndrome, which causes weight gain and depression).

The P and A glands form a negative feedback loop that can oscillate with a seasonal timescale

For the remainder of this lecture, we focus on the timescale of months. The negative feedback loop of two glands can show damped oscillations. If A is large, it reduces P. Smaller P reduces A, making P grow back and so on (Fig 4.5). Each of these steps has a delay due to the tissue turnover time, about 2 months. The overall time for a full cycle is on the order of a year .

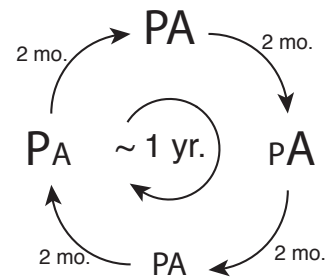


Figure 4.5

When Avichai Tendler in his PhD with me got this year timescale, we got goosebumps: a year is important for organisms because of the changes in seasons. It is useful to have an internal clock that can keep track of seasons, as we will discuss below.

To understand this clock, it helps me to make an analogy of the A-P feedback loop to a mass on a spring. Suppose the mass is at rest at its equilibrium point (Fig 4.6). If you pull the mass down, the spring will stretch and cause a force pointing back to the equilibrium point. If you let go of the mass it will return, but at equilibrium will have high velocity v , so it will overshoot and compress the spring. The compressed spring pushes the mass back down, and so on, to obtain sustained oscillations (Fig 6).

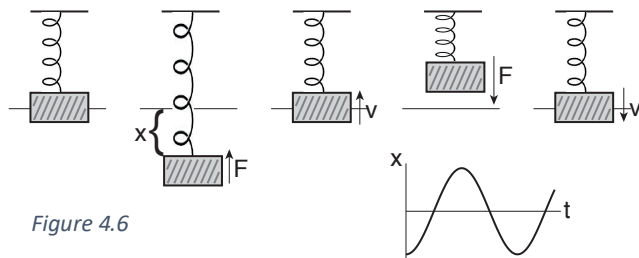


Figure 4.6

To describe this motion, we can use Hooke's law of the spring: the force grows with the extension of the spring, $F = -kx$. This force, according to Newton's law, $F=ma$, changes the acceleration a , which is the rate of change of velocity $a = dv/dt$. Thus $dv/dt = -k/m x$. Velocity is the rate of change of position x , so that $dx/dt = v$. This is like a negative feedback loop in which v increases x but x decreases v (Fig 4.7). Thus, x and v are like A and P, and the negative feedback can act as an oscillator.

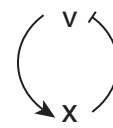


Figure 4.7

When you simulate the slow equations for A and P you don't get sustained oscillations, but instead damped oscillations that decay away. To get damping in the spring analogy, we can put the spring in a container with a viscous fluid like



Figure 4.8

honey. The honey's drag adds a force proportional to velocity: $dv/dt = -kx - b v$ (Fig 4.8). The spring settles down to its equilibrium point.

With the spring analogy we can understand what happens to the HPA axis when it gets inputs that change over a year, like the length of the day. The seasons entrain the oscillator to a period of one year. It's like putting the honey container on a platform that oscillates up and down: the spring picks up the platform's frequency (one year) and begins to oscillate with the same frequency.

Due to this year-scale frequency, the two-gland feedback loop can be **entrained** (driven) effectively by signals that vary over the year. This includes daylight, day-length and other environmental factors.

Indeed, animals have seasonal behaviors and physiological changes across the year. Birds and mammals migrate, change their feathers/fur, and mate in certain seasons. The signals for this seasonal cycle have to do with day length variation- days are shortest in December 21 and longest in June 21 (Fig 4.9). Day length affects a hormone in the brain, called melatonin, which feeds into the P gland.

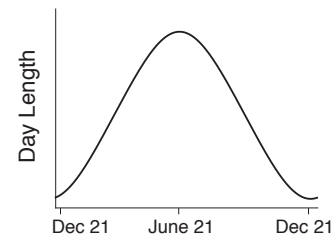


Figure 4.9

But animals do not depend only on external signals such as day-length to figure out which season it is. Animals have an internal clock that keeps track of seasons. Such an internal clock is crucial, for example, to avoid mistaking a few cloudy days in summer as winter. The proof that such a clock exists was found when animals were kept in constant day-length conditions for years. They still showed the seasonal changes, although not with a 12-month period, but with a 10-month period. This inner clock is thus called a circa—annual clock (circa=about, annual=year). The mechanism for this seasonal clock is not well characterized, but the pituitary is thought to be involved, due to its response to a day-length dependent melatonin signal (Jackson and Jansen, 1991; Lincoln *et al.*, 2006; Migaud *et al.*, 2011; Cordero, Brorsen and McFarlane, 2012; Dardente, Hazlerigg and Ebling, 2014; Hut, Dardente and Riede, 2014; Dawson, 2015; Wood *et al.*, 2015; Lewis and Ebling, 2017; Wood and Loudon, 2018).

The long timescale of the $A - P$ oscillator may provide some of the inertia of this internal clock. Since the A-P circuit is a damped oscillator, it cannot, by itself, explain how the seasonal clock runs also in animals kept under constant conditions for years (Strogatz, 2000). But a yearly oscillating input can drive the oscillator to show sizable seasonal oscillations.

To study the effects of entrainment by yearly cycles of day-length, we can study the HPA axis with inputs that vary with a period of one year, $u(t) = 1 + u_0 \cos(\omega t)$. The peak input occurs at the shortest day of the year, December 21. It is worth noting that the seasonal input depends on geography: The amplitude u_0 which increases in north and south latitudes, and is about 2 times larger in London than in Tel Aviv (Teets, 2003).

Using the slow equations, the HPA hormones are predicted to oscillate with a yearly period (Fig 3a), with larger amplitudes at higher latitudes (Fig 3b), reaching about 20% and 30% peak-to-trough in London and Stockholm respectively.

the model makes two specific predictions (Fig 4.10):

- (i) peak cortisol should not occur at the time of peak input (December 21) *but rather be delayed by about two months*. This delay is due to the delay in the feedback loop, and is determined by only two parameters: the P and A tissue turnover times a_A and a_P . The model therefore predicts a cortisol peak at late winter/spring. Such a shift cannot occur in a model without gland dynamics (such as Eq 1-3): when entrained by a yearly input, such models predict a peak at the time of peak input, December 21. The shift and amplitude can be derived analytically from linearized equations.
- (ii) Cortisol and x_2 (ACTH) are in anti-phase, with x_2 peaking in late summer/fall. This antiphase cannot occur in a model without gland size changes, because in such models the two hormones vary together.

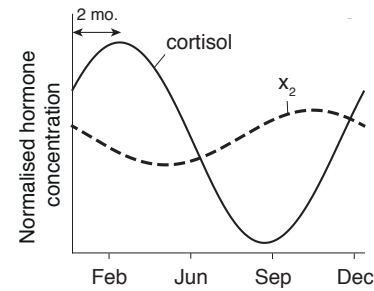


Figure 4.10

Large scale electronic medical record (EMR) datasets show the predicted oscillation phases

To test such predictions requires data on cortisol and x_2 collected in different months. Here we can see two types of data sources in systems medicine- focused 'small data' and exploratory 'big-data'.

A small-data study is usually done with a focused research scope. An example is the UK Whitehall study on thousands of British civil servants. One striking finding is that the lower the civil service rank, the higher the cortisol level. Low rank means higher stress, and indeed socioeconomic status is a strong predictor of health and longevity across cultures.

When tested for seasonality, saliva cortisol levels (meta-analysis with $n=18,698$ subjects) was maximal in winter-spring and minimal in fall (Miller *et al.*, 2016). There was a 17% peak to trough amplitude between spring and fall saliva cortisol. A Swedish study of a dozen people found a saliva cortisol maximum in spring with a larger amplitude of 30% (Persson *et al.*, 2008), as predicted by the northern latitude of Sweden.

In addition to these specific studies on cortisol, one can utilize 'big-data'. A major revolution in systems medicine is the increased availability of large, searchable medical datasets. These are called **electronic medical records (EMR)**. For example, Israel's largest health insurer (Clalit) includes data about half of the Israeli population over 15 years (2002-2017) totaling 46 million life-years, with broad socioeconomic and ethnic representation. The data includes disease codes, drugs purchased, and blood tests. Thus, one can test many hypotheses or look for patterns. For seasonality, you can compare blood tests according to the month of the year they were taken.

A major issue with EMR datasets is **ascertainment bias**. The medical tests were done for a clinical reason, as opposed to a uniform sample of the population as attempted in many 'small-data' experiments. To address this bias requires ways to filter-out people with medical conditions that can confound the results. For the purpose of detecting seasonality, for example, one can remove

all people with chronic disease according to international disease codes. One can also filter out, for each blood test, individuals that took a drug that affects that test. Finding such drugs can be done from the ERM data itself, by finding drugs that significantly affect test results on average. Finally, one can remove tests which include indications for infection to avoid effects of yearly infection cycles (SI).

When these filters are used, a clear picture emerges (Fig 4.11). In both males and females, blood cortisol (n=104,255 blood tests) shows a peak in spring and a minimum in fall, as predicted. The timing of the peak and trough correspond to tissue turnover parameters of about two months. The oscillation amplitude of $5 \pm 0.7\%$ matches the predicted amplitude at the 31° latitude of Israel.

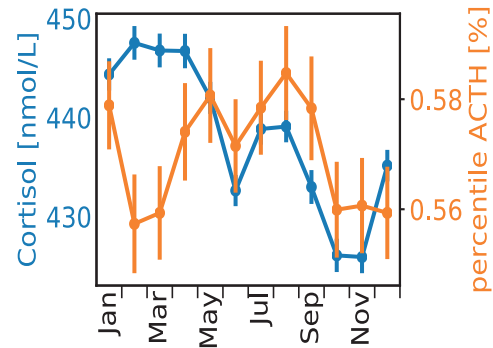


Figure 4.11

There are fewer x_2 (ACTH) tests in the dataset (n=14,073), and hence more noise, yet it appears that x_2 peaks in summer as predicted (Fig 4.11, orange line).

Since the HPA axis is implicated in mood disorders (Watson and Mackin, 2006; Herbert, 2013), the seasonal variation in the HPA axis may contribute to **seasonal affective disorder (SAD)** (Rosen *et al.*, 1990; Mersch *et al.*, 1999; Avery *et al.*, 2001), also known as ‘winter blues’. This is a common disorder especially at high latitudes (1% prevalence in Florida and 10% in Alaska). The spring cortisol peak might explain why the SAD symptoms peak in spring. For example, suicide rates peak in April (Bridges, Yip and Yang, 2005). Likewise, seasonal HPA oscillations can explain the seasonal component of major depression and other mood disorders.

Other HP axes control major tradeoffs- reproduction, growth, metabolism and stress- and show similar seasonal oscillations

The HPA axis is one of four axes that together set the balance between the four major fitness components of living organisms: reproduction, metabolism, growth and stress. The hypothalamus H and pituitary P participate in all four. Each axis has its own x_1, x_2 and x_3 hormones. The hypothalamus H secretes x_1 hormones to induce the pituitary cells to secrete x_2 hormones that activate an effector organ for each axis to secrete x_3 . That is why P is sometimes called “the master gland”.

In the reproduction axis, for example, the x_2 -hormones are FSH and LH, and the x_3 hormones are estrogen (from the ovaries in females) and testosterone (from the testes in males). For metabolism, the x_2 hormone is TSH, and the x_3 hormone is thyroxin secreted from the thyroid gland.

As in the HPA axis, all of these axes show the feedback loop in which x_3 inhibits secretion of x_1 and x_2 . The cells in these glands also proliferate and thus, as in the HPA axis, growth control feedback are at play: for example, thyroid gland growth is regulated by its upstream hormone $x_2 = \text{TSH}$.

This leads to the prediction that the x_2 -hormones should peak in summer/fall and the x_3 hormones in winter-spring. Indeed, Clalit EMR data shows that in both the thyroid axis and the reproduction axis, x_2 -hormone blood test levels (TSH, LH, FSH) show a peak in late summer, about 1-2 months after June 21 (Fig 12). The x_3 hormones (free thyroid hormone T3, estradiol, testosterone) show a late-winter/spring peak (Fig 4. 13), as predicted. Oscillation amplitudes for most hormones are in the range of 4-10%. Some of these tests exceed 6 million data points(!), such as TSH tests, with error bars smaller than the dots in Fig 4.13.

In contrast to the sizable hormone oscillations, blood tests for key metabolites such as glucose, calcium and other important compounds and ions, show smaller seasonal oscillations with amplitudes on the order of 0.03-1%. Thus hormones work hard to keep essential metabolites relatively constant.

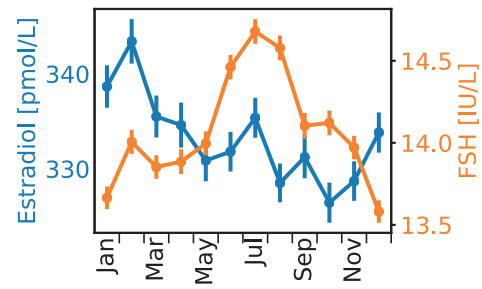


Figure 4.12

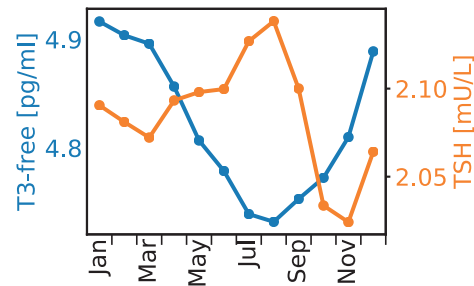


Figure 4.13

Model for steroid withdrawal explains overshoot and recovery over months

Another way to study the HPA oscillator is to pull it away from equilibrium and watch it recover, just like pulling the mass on a spring away from its equilibrium point. Let us therefore consider situations that perturb the HPA axis, reasoning that the two-gland feedback loop should show months-scale recovery dynamics from perturbations.

One common situation occurs when cortisol levels are high for weeks or more, and then are suddenly lowered. This occurs in the medical situation in which people take glucocorticoids (long-lasting cortisol analogues such as dexamethasone or prednisone) for extended periods and then go off the drugs (Koch-Weser and Byyny, 1976; Dixon and Christy, 1980). Glucocorticoid steroids are given to reduce inflammation or suppress immune responses, as in asthma and autoimmune diseases.

It is important not to stop steroid treatment all at once. If you do, the patient will show dangerously low cortisol. This is called “steroid withdrawal”. Thus, one must gradually reduce doses over months.

Steroid treatment can be modelled by external (exogenous) cortisol D added to the equations (by replacing x_3 by $x_3 + D$ in Eq 1 and 2). In this case, both pituitary and adrenal gland sizes shrink, due to inhibition of x_1 and x_2 by D . It’s as if the glands ‘think’ there is too much cortisol and shrink to return it to baseline. This matches the observation of atrophied and involuted adrenal in extended glucocorticoid treatment (Nicolaidis *et al.*, 2000).

Thus, simulations of steroid withdrawal begin with small gland sizes A and P , and exogenous cortisol D set to zero. This is like pulling the spring hard, and then letting it go. The dynamics show that P recovers first and A follows it. As a result, x_2 overshoots after 3 months and goes back to normal after 9 months, whereas cortisol is abnormally low and recovers more slowly after 9 months (Fig 4.14). The overshoot of x_2 is due to the release of x_1 inhibition, which causes P to grow. Only when P returns to normal size is x_2 sufficient to allow A to grow and recover.

Such overshoot dynamics were found by Graber et al, in another example of a “small-data” study which followed $n=14$ patients who went off prolonged steroid treatment (GRABER *et al.*, 1965). Patients showed an overshoot in x_2 and a slower recovery of cortisol over 9 months (Fig 4.14), in agreement with the model.

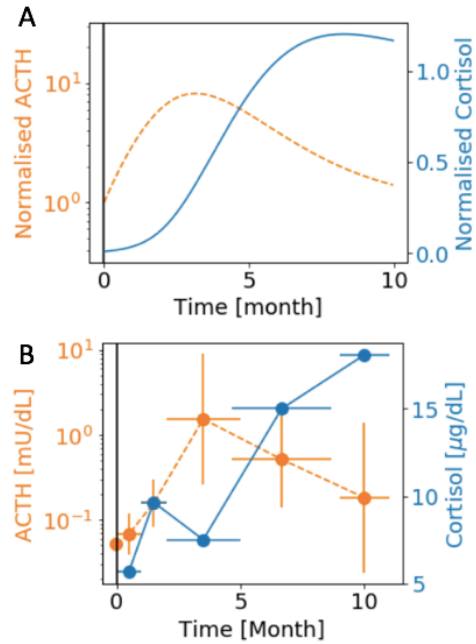


Figure 4.14

Model for postpartum dynamics shows hormonal imbalance over several weeks

Another major perturbation of the HPA axis is pregnancy and the dynamics after delivery, called **post-partum dynamics**. During pregnancy, a special organ in the mother, called the placenta, provides nutrients to the fetus and also secretes hormones to the mother’s bloodstream. One of the main effects of pregnancy on the HPA axis is secretion of x_1 by the placenta at a rate that rises with time, especially in the third trimester (CAMPBELL *et al.*, 1987; Smith and Thomson, 1991; Thomson, 2013). We ignore other effects, for example, increases in cortisol binding proteins which can affect x_3 activity (Hammond, 2016).

The rise in x_1 can be modelled by adding an exogenous x_1 source to the x_1 equation (Eq 1). The rise in x_1 causes an enlargement of both A and P in the model.

At delivery, the placenta exits the mother’s body, and the x_1 source stops all at once. The levels of x_2 fall to normal levels within a day. In contrast, cortisol levels stay high for weeks (Fig 15a). The reason for this hormone imbalance is that cortisol steady-state depends positively on both gland sizes $x_3 \sim P^{\frac{1}{3}}A^{\frac{1}{3}}$ (P supplies x_2 and A supplies cortisol) whereas x_2 depends positively on P but inversely on A : $x_2 \sim P^{\frac{1}{3}}A^{-\frac{2}{3}}$ (the inverse dependence is due to inhibition of x_2 secretion by cortisol). Thus, because both glands are large at delivery, they synergize to increase cortisol but cancel out in x_2 (ACTH). Only when the glands return to normal size does cortisol return to baseline.

These predictions agree with small-data measurements of x_1 (CRH) and x_2 (ACTH, Fig 4.15b), that indicate that these hormones return to baseline within days after delivery (Laatikainen *et al.*, 1987), whereas plasma and urinary cortisol remains high 2-3 months after delivery (Jung *et al.*, 2011).

Big-data EMR records from Clalit cortisol blood tests in women as a function of gestation week ($n = 10,876$) also show that cortisol rises during pregnancy as expected, and shows a slow decline over months post-partum (Fig 4.15c). Although this data is prone to ascertainment bias (cortisol blood tests are not standard and are done in specific conditions), the timescale of the decline should be relatively unbiased, and matches predictions with a half-life of two months (Fig 4.15).

This hormonal mismatch in the weeks after delivery contributes to some forms of postpartum depression (Andrews-Fike, 1999; Seth, Lewis and Galbally, 2016). High cortisol and normal x_2 in the weeks after birth is similar to the hormone imbalance in patients with major depression.

Bipolar disorder has a timescale of weeks-months, and this timescale can be generated by tissue size fluctuations

To end this lecture, let's use these concepts to explore a mystery in psychiatry. We consider a major mood disorder called **bipolar disorder (BD)**. BD, formerly called manic-depression, is characterized by periods of months of depression with negative thoughts and low energy, and periods of mania with excessive high energy, irritability and poorly-thought-out decisions. BD is a leading cause of suicide and self-damage, and causes suffering to about 2% of the world's population. As in most psychiatric conditions, the biological understanding of BP is still lacking. It has a genetic component, but no genes of large effect have been identified. Medication such as lithium can stabilize moods for some people, but the mechanism is not clear.

We focus on the timescale of BD: what sets a weeks-months timescale for mood swings? Neurons work in seconds, hormones in hours, and gene expression changes take a day. Few processes can supply such a month's timescale.

We can explore the hypothesis that the month's timescale is due to tissue size variations. The glands of the HPA axis are natural candidates. The HPA axis is known to be dysregulated in major depression (MD) and BD: people with MD and BD have high cortisol and enlarged adrenal cortex. Conversely, high cortisol and enlarged adrenal causes depression (as in Cushing's syndrome mentioned above). Historically, the one of the first clues for the existence of a stress pathway came from autopsies of suicide victims which showed enlarged adrenals- this was before anyone knew of hormones.

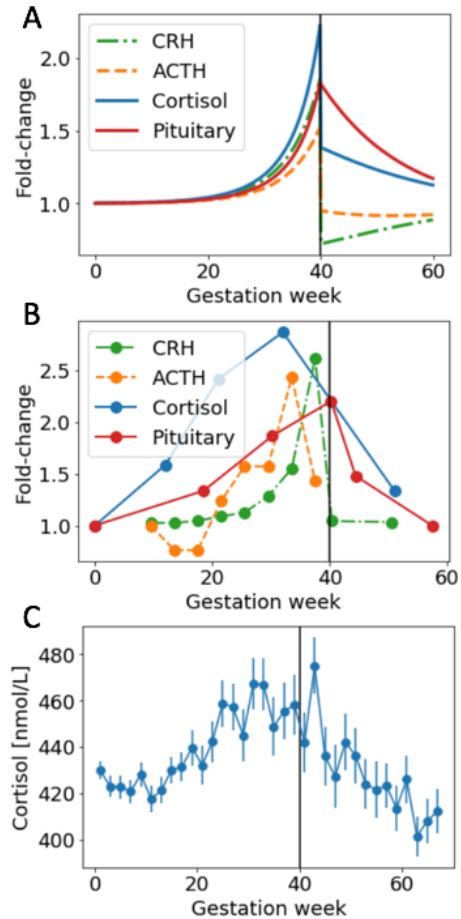


Figure 4.15

Often, BP episodes are preceded by stressful or joyous life events. To make a conceptual model for BP, we therefore need to schematically conceptualize psychological stresses. To do so, we simply note that life events lead to perceptions in the brain (Fig 4.16). Perceptions are person-specific and depend on experience and narratives. For example, a life event such as the victory of a political candidate in an election causes different perceptions in different people: some may grieve, others may rejoice, while others don't care. Perceptions affect our behavior which in turn affects life events. The HPA axis output, cortisol, also affects perceptions as discussed above.

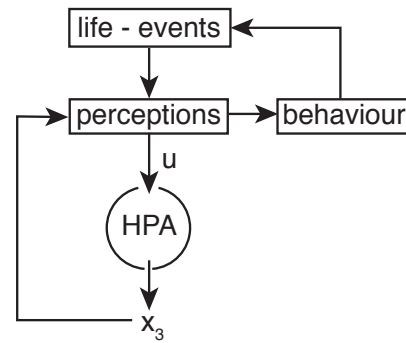


Figure 4.16

Let's suppose that people with BD tend to perceive life events in ways that create larger input signals to H, a kind of psychological excitability. We can model these inputs, based on fluctuating perceptions, as a noise signal u with a standard deviation σ . In people with BD, we can assume that σ is larger than in people without BD.

Such a noisy input to the HPA two-gland feedback loop causes a fascinating phenomenon: the loop shows noisy oscillations in gland sizes and in hormones, with a typical timescale of months. The larger the noisy input amplitude σ , the larger the amplitude of these oscillations. To see this, we can use the spring analogy: put the spring in its honey container on the back of a flatbed truck driving on a rough gravel road (Fig 4.17). The spring picks up vibrational frequencies close to its natural (resonance) frequency and starts making noisy, erratic oscillations (Fig 4.18).

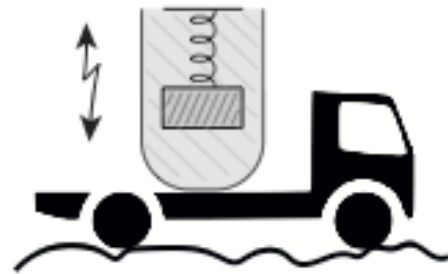


Figure 4.17

Note that these oscillations are due to stresses, and are different from the seasonal oscillations discussed above. Seasonal oscillations are small, and ride on top of the day-day stress responses. To see seasonal oscillations you need to average over many people, in order to average out the stress-signals of different people.

Thus, the noisy oscillations of the HPA axis might account for some of the inertia and timescale of mood swings in BD (Fig 4.18). Similar oscillations in the HP-thyroid axis may cause energy ups and downs typical of BD. A question for current research is how (and whether) the hormones and gland sizes exactly map to mania and depression.

If such HPA oscillations indeed turn out to be important for BD, it points at a way to consider treatment. One can build a controller- an algorithm that administers cortisol at the right phase of the oscillations, in a way that tends to dampen the oscillations out (Fig 4.19). This is similar to an effect

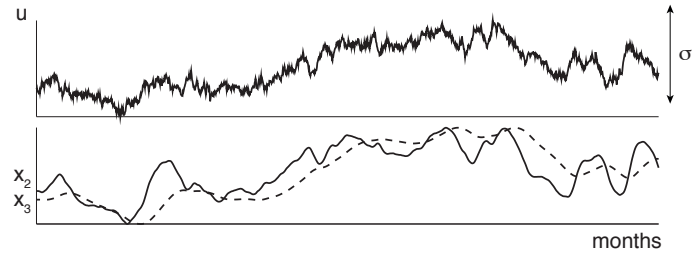


Figure 4.18

known to anyone who ever rode on a swing: if you kick at the right time you can gradually stop a swing from swinging. Indeed, simulations suggest that such a controller can reduce oscillation amplitudes by a large factor (Fig 19). These directions are a current research project in our group.

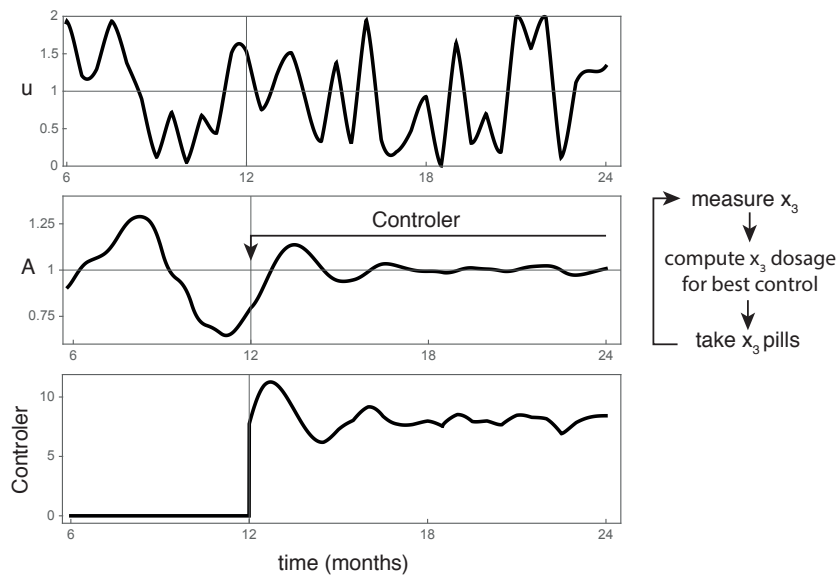


Figure 4.19

To sum up, when two tissues control each other's sizes, oscillatory phenomenon on the timescale of months can occur. The two-gland oscillator allow the HPA stress pathway to synchronize with the seasons, making different hormonal set-points for different times of the year. The cost of this ability is hormone imbalance after delivery and possibly noisy mood swing oscillations in BP. More generally, the body can be considered as an ensemble of interacting organs that constantly adjust their size and activity to changing conditions and to the states of other organs.

Appendix A: Michaelis-Menten equation

A.1 Binding of a repressor to a promoter

This appendix provides a simplified introduction to basic models in biochemistry. We will begin with understanding the interaction of a molecule like cortisol, X, with its receptor, the glucocorticoid receptor R. The molecule X binds to R to form a complex, [XR]. The receptor R can thus be either free, R, or bound, [XR], resulting in a **conservation equation**:

$$R + [XR] = R_T \quad (\text{A.1.1})$$

where R_T is the total concentration of receptor in the cell (about 5000 R per human cell). A single molecule in a cell means a concentration of ~ 1 pM.

X and its target R diffuse in the cell and occasionally collide to form the complex [XR]. This process can be described by mass-action kinetics: X and R collide and bind each other at a rate k_{on} . The rate of complex formation is thus proportional to the collision rate, given by the product of the concentrations of X and R:

$$\text{rate of complex formation} = k_{on} X R$$

The complex [XR] falls apart (dissociates) at a rate k_{off} . The rate of change of [XR] based on these collision and dissociation processes is described by

$$d[XR]/dt = k_{on} X R - k_{off} [XR] \quad (\text{A.1.2})$$

The rate parameter for the collisions, k_{on} , describes how many collision events occur per second per protein at a given concentration of R, and thus has units of 1/time/concentration. It is useful to remember that k_{on} in biochemical reactions is often limited by the rate of collisions of a diffusing molecule hitting a protein-size target, and has a diffusion-limited value of about $k_{on} \sim 10^8 - 10^9 \text{ M}^{-1} \text{ sec}^{-1}$, independent of the details of the reaction.

The off-rate k_{off} , on the other hand, has units of 1/time and can vary over many orders of magnitude for different reactions, because k_{off} is determined by the strength of the chemical bonds that bind X and R.

The kinetics of Equation A.1.2 approach a steady-state in which concentrations do not change with time, $d[XR]/dt = 0$. Solving Equation A.1.2 at steady-state, we find that the balance between the collision of X and R and the dissociation of [XR] leads to the chemical equilibrium equation:

$$K_d [XR] = X R \quad (\text{A.1.3})$$

where K_d is the **dissociation constant**.

$$K_d = k_{off}/k_{on}$$

The dissociation constant K_d has units of concentration. The larger the dissociation constant, the higher the rate of dissociation of the complex, that is, the weaker the binding of X and R

Solving for the concentration of free DNA sites, R, using Equations. A.1.1 and A.1.3, we find $K_d (R_T - R) = X R$, which yields

$$\frac{R}{R_T} = \frac{1}{1+X/K_d} \quad (\text{A.1.4})$$

For many reactions, [XR] complexes dissociate within less than 1 min (that is, $k_{off} > 1 \text{ min}^{-1}$). Therefore, we can average over times much longer than 1 min and consider R/R_T as the probability that site D is free, averaged over many binding and unbinding events.

The probability that the receptor is free, R/R_T , is thus a decreasing function of the concentration of X. When $X = 0$, the receptor is always free, $R/R_T = 1$. The receptor has a 50% chance of being free, $R/R_T = 1/2$, when $X = K_d$.

For cortisol receptor, $K_d \sim 20 \text{ nM}$, and cortisol blood concentration is the hundreds of nM in the morning and about a hundred at night. Hence $X \gg K_d$, and we can approximate $1/(1+X/K_d)$ by K_d/X . This is the rationale for using $1/x$ terms for inhibition in the HPA model equations.